Processus de diffusion : aspects probabilistes et statistiques

Marek Musiela

THESE

présentée à

l'Université Scientifique et Médicale de Grenoble

pour obtenir le grade de
DOCTEUR ES SCIENCES
«Mathématiques»

par

MUSIELA Marek

PROCESSUS DE DIFFUSION

ASPECTS PROBABILISTES ET STATISTIQUES

Thèse soutenue le 10 juillet 1984 devant la commission d'examen.

A. BERNARD	Président
R. CARMONA	Examinateurs
A. LE BRETON	Examinateurs
B. MAISONNEUVE	Examinateurs
E. PARDOUX	Examinateurs
T. PHAM DINH	Examinateurs

UNIVERSITE SCIENTIFIQUE ET MEDICALE DE GRENOBLE

Année universitaire 1982-1983

Président de l'Université : M. TANCHE

MEMBRES DU CORPS ENSEIGNANT DE L'U.S.M.G.

(RANG A)

SAUF ENSEIGNANTS EN MEDECINE ET PHARMACIE

PROFESSEURS DE 1ère CLASSE

ARNAUD Paul	Chimie organique
ARVIEU Robert	Physique nucléaire I.S.N.
AUBERT Guy	Physique C.N.R.S.
AYANT Yves	Physique approfondie
BARBIER Marie-Jeanne	Electrochimie
BARBIER Jean-Claude	Physique expérimentale C.N.R.S. (labo de magnétisme)
BARJON Robert	Physique nucléaire I.S.N.
BARNOUD Fernand	Biosynthèse de la cellulose-Biologie
BARRA Jean-René	Statistiques - Mathématiques appliquées
BELORISKY Elie	Physique
BENZAKEN Claude (M.)	Mathématiques pures
BERNARD Alain	Mathématiques pures
BERTRANDIAS Françoise	Mathématiques pures
BERTRANDIAS Jean-Paul	Mathématiques pures
BILLET Jean	Géographie
BONNIER Jean-Marie	Chimie générale
BOUCHEZ Robert	Physique nucléaire I.S.N.
BRAVARD Yves	Géographie
CARLIER Georges	Biologie végétale
CAUQUIS Georges	Chimie organique
CHIBON Pierre	Biologie animale
COLIN DE VERDIERE Yves	Mathématiques pures
CRABBE Pierre (détaché)	C.E.R.M.O.
CYROT Michel	Physique du solide
DAUMAS Max	Géographie
DEBELMAS Jacques	Géologie générale
DEGRANGE Charles	Zoologie
DELOBEL Claude (M.)	M.I.A.G. Mathématiques appliquées
DEPORTES Charles	Chimie minérale
DESRE Pierre	Electrochimie
DOLIQUE Jean-Michel	Physique des plasmas
DUCROS Pierre	Cristallographie
FONTAINE Jean-Marc	Mathématiques pures
GAGNAIRE Didier	Chimie physique

.../...

GASTINEL Noël	Analyse numérique - Mathématiques appliquées
GERBER Robert	Mathématiques pures
GERMAIN Jean-Pierre	Mécanique
GIRAUD Pierre	Géologie
IDELMAN Simon	Physiologie animale
JANIN Bernard	Géographie
JOLY Jean-René	Mathématiques pures
JULLIEN Pierre	Mathématiques appliquées
KAHANE André (détaché DAFCO)	Physique
KAHANE Josette	Physique
KOSZUL Jean-Louis	Mathématiques pures
KRAKOWIAK Sacha	Mathématiques appliquées
KUPTA Yvon	Mathématiques pures
LACAZE Albert	Thermodynamique
LAJZEROWICZ Jeannine	Physique
LAJZEROWICZ Joseph	Physique
LAURENT Pierre	Mathématiques appliquées
DE LEIRIS Joël	Biologie
LLIBOUTRY Louis	Géophysique
LOISEAUX Jean-Marie	Sciences nucléaires I.S.N.
LOUP Jean	Géographie
MACHE Régis	Physiologie végétale
MAYNARD Roger	Physique du solide
MICHEL Robert	Minéralogie et pétrographie (géologie)
MOZIERES Philippe	Spectrométrie - Physique
OMONT Alain	Astrophysique
OZENDA Paul	Botanique (biologie végétale)
PAYAN Jean-Jacques (détaché)	Mathématiques pures
PEBAY PEYROULA Jean-Claude	Physique
PERRIAUX Jacques	Géologie
PERRIER Guy	Géophysique
PIERRARD Jean-Marie	Mécanique
RASSAT André	Chimie systématique
RENARD Michel	Thermodynamique
RICHARD Lucien	Biologie végétale
RINAUDO Marguerite	Chimie CERMAV
SENGEL Philippe	Biologie animale
SERGERAERT Francis	Mathématiques pures
SOUTIF Michel	Physique
VAILLANT François	Zoologie
VALENTIN Jacques	Physique nucléaire I.S.N.
VAN CUTSEN Bernard	Mathématiques appliquées
VAUQUOIS Bernard	Mathématiques appliquées
VIALON Pierre	Géologie

PROFESSEURS DE 2ème CLASSE

ADIBA Michel	Mathématiques pures
ARMAND Gilbert	Géographie

.../...

AURIAULT Jean-Louis	Mécanique
BEGUIN Claude (M.)	Chimie organique
BOEHLER Jean-Paul	Mécanique
BOITET Christian	Mathématiques appliquées
BORNAREL Jean	Physique
BRUN Gilbert	Biologie
CASTAING Bernard	Physique
CHARDON Michel	Géographie
COHENADDAD Jean-Pierre	Physique
DENEUVILLE Alain	Physique
DEPASSEL Roger	Mécanique des fluides
DOUCE Roland	Physiologie végétale
DUFRESNOY Alain	Mathématiques pures
GASPARD François	Physique
GAUTRON René	Chimie
GIDON Maurice	Géologie
GIGNOUX Claude (M.)	Sciences nucléaires I.S.N.
GUITTON Jacques	Chimie
HACQUES Gérard	Mathématiques appliquées
HERBIN Jacky	Géographie
HICTER Pierre	Chimie
JOSELEAU Jean-Paul	Biochimie
KERCKOVE Claude (M.)	Géologie
LE BRETON Alain	Mathématiques appliquées
LONGEQUEUE Nicole	Sciences nucléaires I.S.N.
LUCAS Robert	Physiques
LUNA Domingo	Mathématiques pures
MASCLE Georges	Géologie
NEMOZ Alain	Thermodynamique (CNRS - CRTBT)
OUDET Bruno	Mathématiques appliquées
PELMONT Jean	Biochimie
PERRIN Claude (M.)	Sciences nucléaires I.S.N.
PFISTER Jean-Claude (détaché)	Physique du solide
PIBOULE Michel	Géologie
PIERRE Jean-Louis	Chimie organique
RAYNAUD Hervé	Mathématiques appliquées
ROBERT Gilles	Mathématiques pures
ROBERT Jean-Bernard	Chimie physique
ROSSI André	Physiologie végétale
SAKAROVITCH Michel	Mathématiques appliquées
SARROT REYNAUD Jean	Géologie
SAXOD Raymond	Biologie animale
SOUTIF Jeanne	Physique
SCHOOL Pierre-Claude	Mathématiques appliquées
STUTZ Pierre	Mécanique
SUBRA Robert	Chimie
VIDAL Michel	Chimie organique
VIVIAN Robert	Géographie

Je suis reconnaissant à Monsieur A. BERNARD de l'intérêt qu'il a témoigné pour ce travail, et pour l'honneur qu'il me fait en acceptant de présider le Jury.

Monsieur A. LE BRETON a assuré la direction de ce travail sans ménager ni son temps ni son énergie. Je lui exprime ici ma profonde gratitude tant pour les nombreuses heures qu'il m'a consacrées, ses encouragements et conseils toujours déterminants que pour l'aide constante qu'il m'a apportée sur de tout autres plans.

Monsieur B. MAISONNEUVE s'est intéressé de près à ce travail. Je le remercie très sincèrement pour ses critiques et conseils amicaux ainsi que sa participation au Jury.

Je désire également remercier Messieurs R. CARMONA, E. PARDOUX, PHAM DINH TUAN pour l'attention qu'ils ont portée à mes travaux, pour leurs suggestions et pour l'honneur qu'ils me font en participant au Jury.

Mes remerciements vont aussi à tous mes amis qui, à des titres divers, m'ont apporté leur soutien.

Madame HOTTELLIER a assuré la dactylographie du manuscrit avec beaucoup de soin. Je l'en remercie vivement.

Mes remerciements vont enfin à Messieurs ANGUILLE et IGLESIAS et au service de reprographie qui ont effectué le tirage de cette thèse.

TABLE DES MATIERES

ANNEXE - REGROUPEMENT DES PUBLICATIONS

INTRODUCTION

Dans ce travail nous menons une étude probabiliste et statistique de processus de diffusions multidimensionnels. L'idée directrice était d'élaborer des outils probabilistes utiles en statistique et de les utiliser pour résoudre certains problèmes d'estimation de paramètres.

Ainsi, si on envisage d'appliquer la méthode du maximum de vraisemblance, il est important de savoir si les lois associées aux observations sont dominées par une certaine mesure. On sait qu'une condition nécessaire et suffisante pour l'absolue continuité de mesures de diffusion est donnée en termes de convergence presque sûre (p.s.)d'une certaine fonctionnelle intégrale . On est alors conduit à analyser le comportement asymptotique p.s. de telles fonctionnelles.

Une fois l'absolue continuité assurée, on peut envisager de construire l'estimateur de maximum de vraisemblance du paramètre. L'étude de la convergence de cet estimateur passe par celle du comportement asymptotique de quantités du type $\langle M\rangle_t^{-1} M_t$, où (M_t) est une martingale locale continue et $(\langle M\rangle_t)$ est son processus de variation quadratique. On peut alors utiliser une version de la loi forte des grands nombres pour les martingales. Si (M_t) est réelle on sait que $\langle M\rangle_t^{-1} M_t$ converge p.s. vers zéro sur l'ensemble $\{\langle M\rangle_\infty = \infty\}$. Ainsi on est encore conduit à étudier le comportement asymptotique de certaines fonctionnelles intégrales.

Nous tentons donc de donner un nouvel éclairage sur des problèmes de nature probabiliste et d'autres de nature statistique concernant les processus de diffusion multidimensionnels. Dans une première partie, nous nous intéressons à des conditions nécessaires et suffisantes pour la convergence p.s. ou la divergence p.s. de certaines fonctionnelles intégrales de tels processus. Dans une deuxième partie nous étudions certains problèmes d'estimation de paramètres pour les diffusions. Nous renvoyons à l'introduction de chacune de ces parties pour une présentation plus détaillée de leurs contenus respectifs.

P A R T I E   I

ETUDE DE CERTAINES FONCTIONNELLES DE PROCESSUS DE DIFFUSION MULTIDIMENSIONNELS

Dans cette partie, nous nous intéressons aux deux problèmes suivants pour des fonctionnelles intégrales du type :

$$I_T = \int_0^T f(t, X_t)\, dt$$

où (X_t) est un processus de diffusion de durée de vie S, f est une application à valeurs réelles positives et T est un temps d'arrêt $T \leq S$:

- Sous quelles conditions peut-on affirmer que $I_T = \infty$ presque sûrement ou au contraire que $I_T < \infty$ presque sûrement ?

- Sous quelles conditions peut-on assurer que la variable aléatoire I_T est intégrable ou même admet des moments exponentiels ?

D'abord nous fixons le cadre de l'étude et nous présentons les bases utiles concernant la technique de substitution pour les mesures de diffusion (cf. KABANOV, LIPTSER, SHIRYAEV [27]). Ensuite nous énonçons les résultats obtenus pour des diffusions multidimentionnelles non homogènes, en les situant par rapport aux résultats classiques pour l'explosion ou la non explosion (cf. KHAS'MINSKII [30, 31], NARITA [50-52]) la récurrence ou la transience (cf.KHAS'MINSKII [30, 31], BHATTACHARYA [7, 8]).

Puis nous envisageons des situations particulières dont celle des diffusions unidimensionnelles homogènes pour lesquelles les résultats sont très explicites et couvrent tous les cas possibles ; ils complètent et généralisent ceux connus concernant l'explosion et la non explosion (cf. FELLER [19], Mc KEAN Jr [43]) et donnent un nouvel éclairage sur des problèmes de nature purement analytique (cf. COPPEL [14]). Enfin nous terminons par quelques remarques concernant le problème de Dirichlet.

CHAPITRE I

DEFINITIONS,
NOTIONS ET OUTILS DE BASE

§ 1. Définitions et notations

Soit D un sous-ensemble ouvert et connexe de $\mathbb{R}^d$ et $D_\delta = D \cup \{\delta\}$, où $\delta = \infty(D)$, le compactifié de D. Soit Ω l'ensemble de toutes les fonctions continues ω de $\mathbb{R}_+$ dans D_δ arrêtées au premier instant d'atteinte de δ (si $\omega(t) = \delta$ alors $\omega(t') = \delta$ pour $t' \geq t$). Désignons par $(X_t)_{t\geq 0}$ le processus des coordonnées sur Ω, par F_t la tribu $(X_s ; s \leq t)$ engendrée sur Ω par les coordonnées X_s pour $s \leq t$ et par F la tribu $F_\infty = \sigma(X_s ; s \in \mathbb{R}_+)$. Soit enfin $W = \mathbb{R}_+ \times \Omega$, $(Y_t)_{t\geq 0}$ le processus défini sur W par $Y_t(c, \omega) = (c + t, \omega(t))$ et G_t et G respectivement les tribus $G_t = \sigma(Y_s ; s \leq t)$ et $G = G_\infty = \sigma(Y_s ; s \in \mathbb{R}_+)$ sur W.

Si U est un sous-ensemble ouvert (pour la topologie trace) de $[c, \infty[\times D$ pour un certain $c \in \mathbb{R}_+$ nous posons

$$S(U) = \inf \{s \geq 0 ; Y_s \notin U\} \quad (\inf \emptyset = \infty).$$

Il est clair que S(U) est un temps d'arrêt de $(G_t)_{t\geq 0}$. La durée de vie de $(Y_t)_{t\geq 0}$ est $S = S(\mathbb{R}_+ \times D) = \inf \{s \geq 0, X_s = \delta\}$.

Nous supposons que les applications $b : \mathbb{R}_+ \times D \to \mathbb{R}^d$ et $\sigma : \mathbb{R}_+ \times D \to \mathbb{R}^d \otimes \mathbb{R}^r$ sont continues et que pour tout $R > 0$ il existe une constante $c_R > 0$ telle que

$$|b(t,x) - b(t,y)| + |\sigma(t,x) - \sigma(t,y)| \leq c_R |x-y|$$

pour tout $t \in \mathbb{R}_+$ et $x, y \in D$ vérifiant $|x| + |y| \leq R$. Pour des vecteurs ou des matrices x, y les symboles $|x|$ et $x \cdot y$ désignent respectivement la norme et le produit scalaire euclidiens ; de même x^* désigne le ou la transposé(e) de x.

Soit $a = \sigma\sigma^*$ et soit Λ l'opérateur différentiel du second ordre sur $\mathbb{R}_+ \times D$ défini par

$$\Lambda = \mathbb{D}_t + L \quad \text{où} \quad L = \frac{1}{2} a \cdot \mathbb{D}_x^2 + b \cdot \mathbb{D}_x .$$

Pour tout $y \in \mathbb{R}_+ \times D_\delta$, soit Q_y la loi de probabilité sur (W, G) telle que $Q_y\{Y_0 = y\} = 1$ et

$$(g(Y_t) - \int_0^t \Lambda\ g(Y_s)\ ds)$$

est une $((G_t), Q_y)$ - martingale locale sur $[0, S[$ pour toute fonction $g \in C^{1,2}(\mathbb{R}_+ \times D)$ (l'application $(t, x) \to g(t,x)$ est une fois continûment dérivable en t, deux fois en x). $(W, G, (G_t), (X_t), S, (Q_y))$ est la réalisation canonique d'un processus de diffusion de Markov sur D engendré par l'opérateur différentiel Λ. Si les fonctions $a(t,\cdot)$ et $b(t,\cdot)$ ne dépendent pas de t, alors $(W, G, (G_t), (X_t), S, (P_x))$ où $P_x = Q_{0,x}$, $x \in D_\delta$, est une diffusion homogène de générateur infinitésimal L (considéré sur D). L'existence et l'unicité de Q_y pour tout $y \in \mathbb{R}_+ \times D_\delta$ découlent de résultats classiques concernant les équations différentielles stochastiques (cf. IKEDA, WATANABE [26] et NARITA [50] par exemple) et la représentation de DOOB de diffusions arrêtées (cf. DYNKIN [16])

Nous utilisons les notations suivantes : pour une fonction ξ définie sur W et Q_y - intégrable le symbole $Q_y\xi$ désigne l'espérance $\int_W \xi dQ_y$; pour une fonction f à valeurs réelles et un sous-ensemble B non vide de son domaine de définition $\sup_B f$ et $\inf_B f$ représentent respectivement le supremum et l'infimum de f sur B.

§ 2. Substitution de mesure

Supposons que c soit une fonction continue de $\mathbb{R}_+ \times D$ dans $\mathbb{R}^d$ uniformément localement lipschitzienne par rapport à x. Définissons les processus stochastiques

$$(1) \quad M_t = X_t - \int_0^t b(Y_s)\ ds$$

et

$$(2) \quad N_t = \exp(\int_0^t c(Y_s)\cdot dM_s - \frac{1}{2}\int_0^t c\cdot ac(Y_s)ds),\ t < S.$$

Le processus (M_t) est une $((G_t), Q_y)$ martingale locale sur $[0,S[$ (cf. MAISONNEUVE [42]) dont la variation quadratique tensorielle et le processus $(\int_0^t a(Y_s)\, ds)$. De plus (N_t) est une $((G_t), Q_y)$ - martingale locale exponentielle sur $[0, S[$ qui a $(\int_0^t N_s^2\, c\cdot ac(Y_s)\, ds)$ pour processus de variation quadratique. Si $\Delta_n = [0, n[\times D_n$ où (D_n) est une suite croissante de sous-ensembles ouverts bornés de D telle que $\cup \bar{D}_n = D$, le théorème de Girsanov assure que, pour tout n, sous $N_{S(\Delta_n)} Q_y$, $(X_t - \int_0^t (b+ac)(Y_s)ds,\ t < S(\Delta_n))$ est une martingale locale bornée. Cela montre que la mesure de diffusion Q_y^c associée à l'opérateur $\Lambda + ac\cdot \mathbb{D}_x$ satisfait à

$$Q_y^c = N_{S(\Delta_n)}\, Q_y \quad \text{sur} \quad G_{S(\Delta_n)} .$$

En fait pour tout temps d'arrêt $\tau \leq S$ on a la décomposition suivante :

$$(3) \quad Q_y^c = N_\tau\, Q_y + Q_y^c\{\{N_\tau = \infty\} \cap \cdot\ \} \text{ sur } G_\tau .$$

De plus, dans cette décomposition, on peut remplacer $\{N_\tau = \infty\}$ par $\{\int_0^\tau c\cdot ac\ (Y_s)ds = \infty\}$ (cf. KABANOV, LIPTSER, SHIRYAEV [27] pour des détails). Par conséquent nous pouvons énoncer :

Proposition 2.1.

Soit $(W, G, (G_t), (X_t), S, (Q_y^c))$ la réalisation canonique d'un processus de diffusion sur D engendré par l'opérateur différentiel $\Lambda + ac \cdot \mathbb{D}_x$. Soit $\tau \leq S$ un temps d'arret de (G_t). Alors

(i) la mesure Q_y^c est localement absolument continue par rapport à la mesure Q_y $(= Q_y^0)$ et N, donnée par (1) et (2), est la densité locale,

(ii) on a la décomposition (3) ;

(iii) la mesure Q_y^c est absolument continue par rapport à Q_y sur G_τ si et seulement si

$$Q_y^c\{\int_0^\tau c\cdot ac(Y_s)\ ds < \infty\} = 1,$$

(iv) la mesure Q_y^c est étrangère à la mesure Q_y sur G_τ si et seulement si

$$Q_y^c\{\int_0^\tau c\cdot ac(Y_s)ds = \infty\} = 1.$$

Remarque 2.1.

Soit $u \in C^{1,2}(\mathbb{R}_+ \times D)$ à valeurs non nulles et $c = \frac{\mathbb{D}x\, u}{u}$. La densité locale N de Q_y^c par rapport à Q_y peut alors s'écrire sous la forme

$$(4) \quad N_t^u = \frac{u(Y_t)}{u(Y_0)} \exp\left(-\int_0^t \frac{\Lambda u}{u}(Y_s)ds\right).$$

Cette représentation de N sera utilisée dans la suite.

CHAPITRE II

INEGALITES DIFFERENTIELLES ET CRITERES RADIAUX POUR L'ETUDE DES FONCTIONNELLES INTEGRALES

Etant donnée une application borélienne f de $\mathbb{R}_+ \times D$ dans $\mathbb{R}_+$, nous posons

$$I_t = \int_0^{t\wedge S} f(Y_s)ds \quad .$$

Nous obtenons des critères de divergence ou de convergence (= finitude) presque sûre (p.s.) de I_S (de $I_{S(\Delta)}$: voir § 1) et nous donnons une condition suffisante pour l'existence de certains de ses moments exponentiels. Les critères de divergence ou de convergence p.s., qui sont formulés en termes d'inégalités différentielles, généralisent ceux de STROOCK et VARADHAN [56] et de NARITA [50-52] ; ils relient le problème de nature probabiliste à un problème de nature purement analytique. Les critères radiaux obtenus généralisent les tests classiques de KHAS'MINSKII [30] de non-explosion et d'explosion (cf. aussi AZENCOTT [3], BHATTACHARYA [7,8], FRIEDMAN [21], Mc KÉAN Jr [43]).

§ 1. Inégalités différentielles

Soit Δ un sous-ensemble ouvert connexe non vide de $[c, \infty[\times D$ pour un certain $c \in \mathbb{R}_+$. Soit de plus U un sous-ensemble compact de Δ tel que $\inf_U f > 0$ ($\inf \emptyset = \infty$). Nous pouvons alors énoncer

Proposition 1.1. (cf. MUSIELA [47])

Supposons qu'il existe un réel $\lambda > 0$ et une fonction $u \geq 0$, $u \in C^{1,2}(\Delta)$ tels que

$$\Lambda u \leq \lambda f \; \phi \circ u \quad \text{sur } \Delta - U,$$

où ϕ est une application croissante différentiable de $\mathbb{R}_+$ dans $\mathbb{R}_+$ telle que

$$\int_0^\infty \frac{dt}{1+\phi(t)} = \infty .$$

Alors pour $y \in \Delta$ on a Q_y p.s.

$$\{\overline{\lim_{t\uparrow S(\Delta)}} u(Y_t) = \infty\} \subset \{I_{S(\Delta)} = \infty\}.$$

En particulier $I_{S(\Delta)} = \infty$ Q_y p.s. si $\lim_{\Delta \ni z \to \infty(\Delta)} u(z) = \infty.$

Proposition 1.2. (cf. MUSIELA [47])

Supposons qu'il existe un réel $\lambda > 0$ et une fonction $u > 0$, $u \in C^{1,2}(\Delta)$ tels que

$$\Lambda u \geq \lambda f\, u \quad \text{sur} \quad \Delta.$$

Alors pour $y \in \Delta$ on a

$$Q_y\{I_{S(\Delta)} < \infty\} \geq \frac{u(y)}{\sup_\Delta u}.$$

Nous formulons maintenant un résultat concernant l'existence de moments de l'intégrale étudiée (cf. KHAS'MINSKII [29, 31] pour le cas non dégénéré).

Proposition 1.3. (cf. MUSIELA [47])

(i) S'il existe un réel $\lambda > 0$ et une fonction $u \in C^{1,2}(\Delta)$ tels que

$$\Lambda u \geq \lambda f \quad \text{sur } \Delta$$

alors on a

$$Q_y\, I_{S(\Delta)} \leq \frac{1}{\lambda} (\sup_\Delta u - u(y)),\ y \in \Delta.$$

(ii) Si on a

$$\alpha = \sup_\Delta Q_y I_{S(\Delta)} < 1$$

alors on a aussi

$$Q_y \exp(I_{S(\Delta)}) \leq \frac{1}{1-\alpha}\ (< \infty),\ y \in \Delta.$$

Désignant par π l'application projection de $\mathbb{R}_+ \times D$ sur D on en déduit :

<u>Corollaire 1.1.</u> (cf. MUSIELA [47])

Si $\pi(\Delta)$ est borné et s'il existe $z \in \mathbb{R}^d$ tel que pour un certain réel $\lambda > 0$

$$\frac{1}{2}\, z\cdot az + z\cdot b \geq \lambda \quad \text{sur } \Delta$$

alors on a

$$Q_y S(\Delta) < \infty \ , \ y \in \Delta .$$

Utilisant la propriété de Markov de (Y_t) sous Q_y nous pouvons démontrer :

<u>Proposition 1.4.</u> (cf. MUSIELA [47])

Posons $q(y) = Q_y\{I_{S(\Delta)} < \infty\}$, $y \in \Delta$. Supposons que

(a) $Q_y\{S(U) < \infty\} = 1$ pour tout $y \in U$ et pour tout sous-ensemble U de Δ tel que $\pi(U)$ soit borné et $\overline{U} \subset \Delta$,

(b) $\lim\limits_{\pi\Delta \ni \pi y \to \infty(\pi\Delta)} q(y) = 1.$

Alors on a $q \equiv 1$ sur Δ.

Considérons une fonction $u \in C^{1,2}(\mathbb{R}_+ \times D)$ à valeurs non nulles et posons $c = \frac{\mathbb{D}_x u}{u}$. Désignons, comme dans le § 2 du chapitre I, par Q_y^c la mesure de diffusion associée à l'opérateur différentiel $\Lambda + ac\cdot\mathbb{D}_x$. Comme la densité locale de Q_y^c par rapport à Q_y est donnée par (4), on peut facilement démontrer :

<u>Proposition 1.5.</u> (cf. MUSIELA [47])

Supposons que pour un certain réel $\lambda > 0$ on ait

$$\frac{\Lambda u}{u} \leq \lambda f \text{ sur } \Delta$$

et que $\underline{\lim}_{n\to\infty} Q_y^c \frac{1}{u} (Y_{S_n}) = 0$

pour une certaine suite (S_n) de temps d'arrêt telle que $S_n < S(\Delta)$, $S_n \nearrow S(\Delta)$. Alors on a

$$Q_y \{I_{S(\Delta)} = \infty\} = 1, \ y \in \Delta.$$

Proposition 1.6. (cf. MUSIELA [47])

Supposons que pour un certain réel $\lambda > 0$ on ait

$$\frac{\Lambda u}{u} \geq \lambda f \quad \text{sur} \quad \Delta.$$

Alors, pour tout $y \in \Delta$, pour toute suite (S_n) de temps d'arrêt telle que $S_n < S(\Delta)$, $S_n \nearrow S(\Delta)$, on a

$$Q_y \{I_{S(\Delta)} < \infty\} \geq \overline{\lim}_{n\to\infty} Q_y^c \frac{u(y)}{u(Y_{S_n})} .$$

§ 2. Critères radiaux

Dans la suite de ce chapitre nous supposons que $D = \mathbb{R}^d$ et que pour tout $x \in \mathbb{R}^d$ les fonctions $a(\cdot, x)$ et $b(\cdot, x)$ sont bornées. Nous sommes alors en mesure de fournir des critères plus explicites de divergence ou de convergence de l'intégrale I_S, où $S = \inf \{t \geq 0 \ ; \ X_t = \delta\}$. Nous étudions aussi l'intégrale $I_{T_a \wedge T_b}$, où

$$T_a = \inf\{t \geq 0 \ ; \ |X_t - z| = (2a)^{1/2}\}, \ z \in \mathbb{R}^d, \ 0 \leq a < b \leq \infty.$$

Soit $z \in \mathbb{R}^d$ et r un réel $r > 0$. Supposons qu'il existe des fonctions continues

$$\overline{\alpha} : [r, \infty[\to]0, \infty[$$
$$\overline{\beta} : [r, \infty[\to \mathbb{R}$$
$$\underline{\gamma} : [r, \infty[\to \mathbb{R}_+$$

telles que pour tout $t \in \mathbb{R}_+$ et tout $x \in \mathbb{R}^d$ avec $|x| \geq (2r)^{1/2}$, on ait

$$x \cdot a(t,\, x+z)x \leq \overline{\alpha}\left(\frac{|x|^2}{2}\right),$$

$$\text{(5)} \quad \operatorname{tr} a(t,\, x+z) + 2\, x \cdot b(t,\, x+z) \leq x \cdot a(t,\, x+z)x\ \overline{\beta}\left(\frac{|x|^2}{2}\right),$$

$$\underline{\gamma}\left(\frac{|x|^2}{2}\right) \leq f(t,\, x+z).$$

Définissons pour $t \geq r$

$$\overline{e}(t) = \exp\left(-\int_r^t \overline{\beta}(\eta)d\eta\right) \; ; \; \overline{m}(t) = 2\int_r^t \frac{\underline{\gamma}}{\overline{e}\,\overline{\alpha}}(\eta)\, d\eta \; ;$$

$$\overline{s}(t) = \int_r^t \overline{e}(\eta)\, d\eta \; ; \; \overline{k}(t) = \int_r^t \overline{e}\, \overline{m}(\eta)\, d\eta \, .$$

Soit $B(z,\, r) = \{x \in \mathbb{R}^d : |x-z| < (2r)^{1/2}\}$. Nous démontrons alors un résultat qui généralise le test de non explosion de KHAS'MINSKII.

<u>Proposition 2.1.</u> (cf. MUSIELA [47])

Si $\overline{k}(\infty) = \infty$ et si $\inf\limits_{\mathbb{R}_+ \times B(z,r)} f > 0$, alors pour tout $y \in \mathbb{R}_+ \times \mathbb{R}^d$ on a :

$$Q_y\{I_S = \infty\} = 1.$$

Soit encore $z \in \mathbb{R}^d$, $r > 0$ et supposons maintenant qu'il existe des fonctions continues

$\underline{\alpha} : [r,\, \infty[\to]0,\, \infty[$

$\underline{\beta} : [r,\, \infty[\to \mathbb{R}$

$\overline{\gamma} : [r,\, \infty[\to \mathbb{R}_+$

telles que pour tout $t \in \mathbb{R}_+$ et tout $x \in \mathbb{R}^d$ avec $|x| \geq (2r)^{1/2}$ on ait

$$\underline{\alpha}\left(\frac{|x|^2}{2}\right) \leq x \cdot a(t,\, x+z)x \, ,$$

(6) $x \cdot a(t, x+z)x \, \underline{\beta}(\frac{|x|^2}{2}) \leq \text{tr } a(t, x+z) + 2x \cdot b(t, x+z)$,

$f(t, x+z) \leq \overline{\gamma}(\frac{|x|^2}{2})$

Définissons pour $t \geq r$

$$\underline{e}(t) = \exp(-\int_r^t \underline{\beta}(\eta)d\eta) \; ; \; \underline{m}(t) = 2\int_r^t \frac{\overline{\gamma}}{\underline{e}\,\underline{\alpha}}(\eta)d\eta \; ;$$

$$\underline{s}(t) = \int_r^t \underline{e}(\eta)d\eta \; ; \; \underline{k}(t) = \int_r^t \underline{e}\,\underline{m}(\eta)d\eta \; ;$$

et

$$\overline{\delta}(t) = \frac{\overline{e}\,\overline{m}}{\underline{e}}(t) \; ; \; \underline{\delta}(t) = \frac{\underline{e}\,\underline{m}}{\overline{e}}(t).$$

L'étude qui suit est conduite sous l'hypothèse complémentaire :

(A) Pour tout sous-ensemble ouvert U de $\mathbb{R}_+ \times \mathbb{R}^d$ tel que $\pi(U)$ est borné et tout $y \in U$ on a $Q_y S(U) < \infty$.

Une condition simple sous laquelle (A) est automatiquement satisfaite est donnée dans le corollaire 1.1 ci-dessus. Ainsi si les fonctions a et b ne dépendent pas de t et si a(x) est définie positive pour tout $x \in \mathbb{R}^d$ alors (A) est vérifiée. Si de plus $f \equiv 1$ alors le résultat suivant se réduit au test d'explosion de KHAS'MINSKII (cf. Mc KEAN [43] par exemple).

Proposition 2.2. (cf. MUSIELA [47])

Supposons l'hypothèse (A) satisfaite. Si $\underline{k}(\infty) < \infty$, alors pour tout $y \in \mathbb{R}_+ \times \mathbb{R}^d$ on a

$$Q_y\{I_S < \infty\} = 1.$$

Dans la suite de cette section nous supposons que les fonctions continues $\underline{\alpha}$, $\underline{\beta}$, $\overline{\gamma}$ et $\overline{\alpha}$, $\overline{\beta}$, $\underline{\gamma}$ satisfont à (5) et (6) respectivement sur l'intervalle $]0, \infty[$. Nous étudions l'intégrale $I_{T_a \wedge T_b}$ où $T_a = \inf\{t \geq 0 : |X_t - z| = (2a)^{1/2}\}$, $z \in \mathbb{R}^d$, $0 \leq a < b \leq \infty$. Remarquons que $T_\infty = \lim_{b \to \infty} T_b = S$ et $T_0 = \inf\{t \geq 0 : X_t = z\}$.

Proposition 2.3. (cf. MUSIELA [47])

Supposons l'hypothèse (A) satisfaite. Si x est tel que $0 < (2a)^{1/2} < |x-z| < (2b)^{1/2} < \infty$ et $t \in \mathbb{R}_+$ alors pour $y = (t,x)$ on a

$$\frac{\overline{s}(b)-\overline{s}(\frac{|x-z|^2}{2})}{\overline{s}(b) - \overline{s}(a)} \leq Q_y\{T_a < T_b\} \leq \frac{\underline{s}(b)-\underline{s}(\frac{|x-z|^2}{2})}{\underline{s}(b) - \underline{s}(a)}$$

et

$$\overline{k}(b)-\overline{k}(\frac{|x-z|^2}{2}) - (\overline{k}(b)-\overline{k}(a))\frac{\underline{s}(b)-\underline{s}(\frac{|x-z|^2}{2})}{\underline{s}(b) - \underline{s}(a)} \leq Q_y I_{T_a \wedge T_b} \leq$$

$$\underline{k}(b)-\underline{k}(\frac{|x-z|^2}{2}) - (\underline{k}(b)-\underline{k}(a))\frac{\overline{s}(b)-\overline{s}(\frac{|x-z|^2}{2})}{\overline{s}(b) - \overline{s}(a)}$$

Nous envisageons maintenant l'étude de l'intégrale $I_{T_a \wedge T_b}$ où $0 \leq a < b \leq \infty$ (0 et ∞ compris).

Proposition 2.4. (cf. MUSIELA [47])

Supposons l'hypothèse (A) satisfaite. Soit $y = (t, x)$, où $t \in \mathbb{R}_+$ et $x \in B(z, b) - \overline{B(z, a)}$.

1) Si $\overline{k}(a) = \infty$, $\overline{k}(b) = \infty$, alors $Q_y\{I_{T_a \wedge T_b} = \infty\} = 1$.

2) Si $\overline{k}(a) < \infty$, $\overline{k}(b) = \infty$, alors $\{I_{T_a \wedge T_b} < \infty\} = \{T_a < T_b\}$ Q_y p.s.

Si de plus :

a) $\underline{s}(a) > -\infty$, $\underline{s}(b) < \infty$, alors

$$\frac{\overline{s}(b)-\overline{s}\left(\frac{|x-z|^2}{2}\right)}{\overline{s}(b) - \overline{s}(a)} \leq Q_y\{I_{T_a \wedge T_b} < \infty\} \leq \frac{\underline{s}(b)-\underline{s}\left(\frac{|x-z|^2}{2}\right)}{\underline{s}(b) - \underline{s}(a)},$$

b) $\overline{s}(b) = \infty$, alors $Q_y\{I_{T_a \wedge T_b} < \infty\} = 1$,

c) $\overline{s}(b) = \infty$, $\overline{\delta}(b) = \infty$, alors $Q_y\, I_{T_a \wedge T_b} = \infty$

3) Si $\overline{k}(a) = \infty$, $\overline{k}(b) < \infty$, alors $\{I_{T_a \wedge T_b} < \infty\} = \{T_b < T_a\}$ Q_y p.s.

Si de plus :

a) $\underline{s}(a) > -\infty$, $\underline{s}(b) < \infty$, alors

$$\frac{\underline{s}\left(\frac{|x-z|^2}{2}\right)-\underline{s}(a)}{\underline{s}(b) - \underline{s}(a)} \leq Q_y\{I_{T_a \wedge T_b} < \infty\} \leq \frac{\overline{s}\left(\frac{|x-z|^2}{2}\right)-\overline{s}(a)}{\overline{s}(b) - \overline{s}(a)}$$

b) $\overline{s}(a) = -\infty$, alors $Q_y\{I_{T_a \wedge T_b} < \infty\} = 1$,

c) $\overline{s}(a) = -\infty$, $\overline{\delta}(a) = -\infty$, alors $Q_y\, I_{T_a \wedge T_b} = \infty$.

Nous pouvons aussi démontrer :

<u>Proposition 2.5.</u> (cf. MUSIELA [47])

Supposons l'hypothèse (A) satisfaite et $y = (t, x) \in \mathbb{R}_+ \times (B(z,b)-\overline{B(z,a)})$
Supposons que l'une des conditions suivantes soit vérifiée :

(a) $\underline{k}(a) < \infty$, $\underline{k}(b) < \infty$
(b) $\underline{k}(a) < \infty$, $\underline{k}(b) = \infty$, $\overline{s}(b) = \infty$,
(c) $\underline{k}(a) = \infty$, $\overline{s}(a) = -\infty$, $\underline{k}(b) < \infty$.

Alors $Q_y\{I_{T_a \wedge T_b} < \infty\} = 1$. Si de plus $\underline{\delta}(b) < \infty$ dans (b) ou $\underline{\delta}(a) > -\infty$ dans (c), alors $Q_y I_{T_a \wedge T_b} < \infty$.

Dans la suite nous désignons par $\underline{k}_1$, $\underline{\delta}_1$, $\overline{\delta}_1$ les fonctions $\underline{k}$, $\underline{\delta}$, $\overline{\delta}$ respectivement lorsque $f = \underline{\gamma} = \overline{\gamma} \equiv 1$.

Des résultats précédents, faisant $f \equiv 1$, on peut déduire des critères pour la transience, la récurrence, la récurrence positive et la récurrence nulle de diffusions dégénérées satisfaisant à (A) (cf. BHATTACHARYA [7,8] pour le cas non dégénéré). Si $\underline{s}(\infty) < \infty$ alors la diffusion est transiente. Si $\overline{s}(\infty) = \infty$ la diffusion est récurrente. Si $\overline{s}(\infty) = \infty$, $\underline{\delta}_1(\infty) < \infty$ ou $\overline{s}(\infty) = \infty$, $\overline{\delta}_1(\infty) = \infty$ alors elle est respectivement récurrente positive ou récurrente nulle.

Maintenant, utilisant la proposition 1.3 nous pouvons démontrer :

<u>Proposition 2.6.</u> (cf. MUSIELA [47])

Supposons l'hypothèse (A) satisfaite et l'une des conditions suivantes vérifiée :

(a) $\underline{k}(a) < \infty$, $\underline{k}(b) < \infty$ et $\int_{t_0}^{b} \int_{t_0}^{t} \underline{\delta}(du)\, \overline{s}(dt) < 1$,

où t_0 est donné par $\underline{\delta}(t_0) \int_a^b \overline{s}(du) = \int_a^b \underline{\delta}(u)\, \overline{s}(du)$,

(b) $\underline{k}(a) < \infty$, $\underline{k}(b) = \infty$, $\overline{s}(b) = \infty$, $\underline{\delta}(b) < \infty$ et $\int_a^b \int_t^b \underline{\delta}(du)\overline{s}(dt) < 1$,

(c) $\underline{k}(a) = \infty$, $\overline{s}(a) = -\infty$, $\underline{\delta}(a) > -\infty$, $\underline{k}(b) < \infty$ et $\int_a^b \int_a^t \underline{\delta}(du)\overline{s}(dt) < 1$.

Alors il existe une constante c telle que pour $y \in \mathbb{R}_+ \times (B(z,b) - \overline{B(z,a)})$ on a

$$Q_y \exp(I_{T_a \wedge T_b}) \le c < \infty.$$

Utilisant le résultat précédent on peut déduire :

Corollaire 2.1. (cf.MUSIELA [47])

Supposons l'hypothèse (A) satisfaite et l'une des conditions suivantes vérifiée

(a) $\underline{k}(a) < \infty$, $\underline{k}(b) < \infty$,

(b) $\underline{k}(a) < \infty$, $\underline{k}(b) = \infty$, $\overline{s}(b) = \infty$, $\underline{\delta}(b) < \infty$ et $\int_a^b \int_t^b \underline{\delta}(du)\overline{s}(dt) < \infty$,

(c) $\underline{k}(a) = \infty$, $\overline{s}(a) = -\infty$, $\underline{\delta}(a) > -\infty$, $\underline{k}(b) < \infty$ et $\int_a^b \int_a^t \underline{\delta}(du)\overline{s}(dt) < \infty$.

Alors il existe un réel $0 < \lambda < \infty$ tel que pour $y \in \mathbb{R}_+ \times (B(z,b)-\overline{B(z,a)})$ on a

$$Q_y \exp(\lambda I_{T_a \wedge T_b}) \leq \lambda^{-1} < \infty.$$

Corollaire 2.2. (cf.MUSIELA [47])

Supposons l'hypothèse (A) satisfaite et l'une des situations suivantes réalisée :

(a) $\underline{k}_1(a) < \infty$, $\underline{k}_1(b) < \infty$, $\lambda < (\int_{t_0}^b \int_{t_0}^t \underline{\delta}_1(du)\ \overline{s}(dt))^{-1}$,

où t_0 est donné par $\underline{\delta}_1(t_0) \int_a^b \overline{s}(du) = \int_a^b \underline{\delta}_1(u)\ \overline{s}(du)$,

(b) $\underline{k}_1(a) < \infty$, $\underline{k}_1(b) = \infty$, $\overline{s}(b) = \infty$, $\underline{\delta}_1(b) < \infty$ et

$\lambda < (\int_a^b \int_t^b \underline{\delta}_1(du)\ \overline{s}(dt))^{-1}$, $\quad (\infty^{-1} = 0)$

(c) $\underline{k}_1(a) = \infty$, $\overline{s}(a) = -\infty$, $\underline{\delta}_1(a) > -\infty$, $\underline{k}_1(b) < \infty$ et

$\lambda < (\int_a^b \int_a^t \underline{\delta}_1(du)\ \overline{s}(dt))^{-1}$.

Alors il existe une constante c telle que pour $y \in \mathbb{R}_+ \times (B(z,b) - \overline{B(z,a)})$ on a

$$Q_y \exp(\lambda\, T_a \wedge T_b) \leq c < \infty.$$

Exemple

Supposons que $b(t,x) \equiv 0$ et que $a(t,x) \equiv I$. Alors sous $P_x = Q_{0,x}$, $x \in \mathbb{R}^d$, $(X_t)_{t\geq 0}$ est un mouvement brownien standard. Supposons que $d \geq 2$, le cas $d = 1$ étant étudié dans le chapitre suivant. Considérons l'intégrale

$$I_t = \int_0^t \gamma\left(\frac{|X_s|^2}{2}\right) ds,$$

où la fonction $\gamma :]0, \infty[\to \mathbb{R}_+$ est continue et non identiquement nulle. On peut voir facilement que avec $z = 0$ et $r > 0$

$$\overline{e}(t) = \underline{e}(t) = \left(\frac{r}{t}\right)^{d/2}$$

$$\delta(t) \overset{\text{def}}{=} \overline{\delta}(t) = \overline{m}(t) = \underline{\delta}(t) = \underline{m}(t) = r^{-\frac{d}{2}} \int_r^t \eta^{\frac{d}{2}-1} \gamma(\eta)\, d\eta$$

$$k(t) \overset{\text{def}}{=} \overline{k}(t) = \underline{k}(t) = \int_r^t \eta^{-\frac{d}{2}} \int_r^\eta \xi^{\frac{d}{2}-1} \gamma(\xi)\, d\xi\, d\eta.$$

Supposons d'abord que $d = 2$, $0 < a < \infty$ et $x \in \mathbb{R}^2 - \overline{B(0, a)}$.

Nous pouvons énoncer les assertions suivantes :

(a) $P_x\{I_{T_a} < \infty\} = 1$,

(b) si $\int_a^\infty \gamma(t)\, dt = \infty$, alors $P_x I_{T_a} = \infty$,

(c) si $\int_a^\infty \gamma(t)\, dt < \infty$, alors $P_x I_{T_a} < \infty$,

(d) si $\int_a^\infty \frac{1}{t} \int_t^\infty \gamma(s) ds\, dt < 1$, alors $P_x \exp(I_{T_a}) \leq c < \infty$.

Supposons maintenant que $d \geq 3$, $0 < a < \infty$, $x \in \mathbb{R}^d - \overline{B(0,\, a)}$.

Nous pouvons alors énoncer :

(a) si $k(\infty) = \infty$, alors $P_x\{I_{T_a} < \infty\} = \left(\frac{\sqrt{2a}}{|x|}\right)^{d-2}$

(b) si $k(\infty) < \infty$, alors il existe un réel $\lambda > 0$ tel que $P_x \exp(\lambda\, I_{T_a}) \leq c < \infty$,

(c) si $k(\infty) < \infty$ et $\int_{t_0}^{\infty} t^{-\frac{d}{2}} \int_{t_0}^{t} s^{\frac{d}{2}-1} \gamma(s)\, ds\, dt < 1$, où t_0 est donné par $\delta(t_0) = (\frac{d}{2} - 1)\, a^{\frac{d}{2}-1} \int_a^{\infty} \delta(t)\, t^{-\frac{d}{2}}\, dt$, alors $P_x \exp(I_{T_a}) \leq c < \infty$.

Supposons enfin que $d \geq 3$, $0 < a < b < \infty$, $x \in B(0,\, b) - \overline{B(0,\, a)}$.

Alors, si $\lambda \leq \dfrac{\frac{d}{2} - 1}{b - a}$, on a $P_x \exp(\lambda\, T_a \wedge T_b) \leq c < \infty$.

C H A P I T R E III

FONCTIONNELLES INTEGRALES DE DIFFUSIONS LINEAIRES ET DE DIFFUSIONS UNIDIMENSIONNELLES HOMOGENES

Dans ce chapitre, nous considérons des cas particuliers de diffusions et de fonctionnelles pour lesquels nous donnons une description plus précise des propriétés de divergence et de convergence. Il s'agit d'abord de fonctionnelles quadratiques de diffusion linéaires puis de fonctionnelles de diffusions unidimensionnelles homogènes pour lesquelles les résultats couvrent et généralisent ceux connus pour l'explosion et la non-explosion (cf. FELLER [19], Mc KEAN Jr. [43]).

§ 1. Fonctionnelles quadratiques de diffusions linéaires

Nous supposons ici que $a(t, x) = A(t)$, $b(t, x) = B(t)x$, $t \in \mathbb{R}_+$, $x \in \mathbb{R}^d$ où A et B sont des applications continues de $\mathbb{R}_+$ dans $\mathbb{R}^d \otimes \mathbb{R}^d$. Alors on sait que $S = \infty$ Q_y p.s. pour tout $y \in \mathbb{R}_+ \times \mathbb{R}^d$ et que le processus $(X_t)_{t\geq 0}$ est gaussien sous $P_x = Q_{0,x}$ pour $x \in \mathbb{R}^d$, avec comme fonction moyenne $(\phi(t)x)_{t\geq 0}$ et fonction de covariance $(K(s,t))_{s,t\geq 0}$, où

$$\phi(t) = I + \int_0^t B(s)\, \phi(s)\, ds,$$

$$K(s,t) = \phi(s) \int_0^{s\wedge t} (\phi^{-1} A\, \phi^{-1*})\, (\eta) d\eta\, \phi^*(t).$$

Nous cherchons à étudier la finitude de l'intégrale

$$I_\infty = \int_0^\infty X_t \cdot Q(t)\, X_t\, dt$$

où $(Q(t))_{t\geq 0}$ est une fonction continue de $\mathbb{R}_+$ dans l'ensemble des matrices $d \times d$ symétriques positives.

Désignons par $\Gamma_0 \in \mathbb{R}^d \otimes \mathbb{R}^d$ une matrice symétrique définie positive et soit Γ l'application de $\mathbb{R}_+$ dans $\mathbb{R}^d \otimes \mathbb{R}^d$, solution de l'équation de Riccati matricielle

$$\Gamma' = A + B\Gamma + \Gamma B^* - \Gamma Q \Gamma \ ; \ \Gamma(0) = \Gamma_0.$$

Proposition 1.1. (cf. MUSIELA [47])

Si K(t,t) est définie positive pour tout t > 0, on a, pour tout $x \in \mathbb{R}^d$

(i) $P_x\{I_\infty = \infty\} = 1$ si et seulement si $\int_0^\infty \operatorname{tr} Q\Gamma(t)dt = \infty$

(ii) $P_x\{I_\infty < \infty\} = 1$ si et seulement si $\int_0 \operatorname{tr} Q\Gamma(t)dt < \infty$.

Supposons maintenant que A(t) = A et B(t) = B pour tout $t \in \mathbb{R}_+$. Alors K(t,t) est définie positive pour tout t > 0 si et seulement si rang $[A^{1/2}, BA^{1/2}, \dots, B^{d-1} A^{1/2}] = d$ i.e. de façon équivalente la paire $[B, A^{1/2}]$ est contrôlable. Utilisant le résultat précédent on peut déduire :

Proposition 1.2. (cf. LE BRETON et MUSIELA [39])

Supposons que la paire $[B, A^{1/2}]$ soit contrôlable. Alors pour tout $x \in \mathbb{R}^d$

$$P_x\{\int_0^\infty X_t \cdot QX_t \, dt = \infty\} = 1$$

pour toute matrice symétrique postive Q (Q ≠ 0).

§ 2. Diffusions unidimensionnelles homogènes

Nous supposons ici que $D =]\ell, r[$ $(-\infty \leq \ell < r < \infty)$ est un intervalle ouvert de $\mathbb{R}$ et que les fonctions à valeurs réelles a et b ne dépendent pas de t i.e. a(t,x) = a(x) et b(t,x) = b(x), $t \in \mathbb{R}_+$, $x \in D$. Nous supposons de plus a > 0 sur D.

Alors $(W, G, (G_t), (X_t), S, (P_x))$ avec $P_x = Q_{0,x}$, $x \in D$, est la réalisation canonique d'une diffusion homogène de générateur infinitésimal

$$L = \frac{1}{2} a \, D_x^2 + b \, D_x.$$

Nous envisageons l'étude de l'intégrale

$$I_S = \int_0^S f(X_t)dt$$

pour une fonction f continue de $\mathbb{R}$ dans $\mathbb{R}_+$ non identiquement nulle.

Choisissons un réel $c \in D$ tel que $f(c) > 0$ et définissons les fonctions

$$e(x) = \exp(-\int_c^x \frac{2b}{a}(y)dy) \; ; \; m(x) = 2\int_c^x \frac{f}{ea}(y)dy \; ;$$

$$s(x) = \int_c^x e(y)dy \; ; \; k(x) = \int_c^x em(y)dy.$$

Posons $p(x) = P_x\{I_S < \infty\}$, $x \in D$. Lorsque $f \equiv 1$, les parties "si" de (i) et (ii-a) de l'énoncé suivant se réduisent aux tests classiques de Feller de non explosion et d'explosion.

Proposition 2.1. (cf. MUSIELA [46])

On a les assertions suivantes :

(i) $p \equiv 0$ si et seulement si $k(\ell) = \infty$ et $k(r) = \infty$.

(ii) $p \equiv 1$ si et seulement si l'une des situations suivantes est réalisée

(a) $k(\ell) < \infty$, $k(r) < \infty$,

(b) $k(\ell) < \infty$, $k(r) = \infty$, $s(r) = \infty$,

(c) $k(\ell) = \infty$, $s(\ell) = -\infty$, $k(r) < \infty$.

(iii) Si $k(\ell) < \infty$, $k(r) = \infty$, $s(r) < \infty$, alors $p(x) = \dfrac{s(r) - s(x)}{s(r) - s(\ell)}$.

(iv) Si $k(\ell) = \infty$, $s(\ell) > -\infty$, $k(r) < \infty$, alors $p(x) = \dfrac{s(x) - s(\ell)}{s(r) - s(\ell)}$.

En ce qui concerne l'espérance $P_x I_S$ nous avons

Proposition 2.2. (cf. MUSIELA [47])

(i) Soit $x \in D$. On a $P_x I_S < \infty$ si et seulement si l'une des situations suivantes est réalisée :

(a) $k(\ell) < \infty$, $k(r) < \infty$,
(b) $k(\ell) < \infty$, $k(r) = \infty$, $s(r) = \infty$, $m(r) < \infty$,
(c) $k(\ell) = \infty$, $s(\ell) = -\infty$, $m(\ell) > -\infty$, $k(r) < \infty$.

(ii) Soit $x \in D$. $P_x I_S$ est donnée dans chacun des cas (a), (b) et (c) de (i) respectivement par

(a') $k(\ell) \dfrac{s(r) - s(x)}{s(r) - s(\ell)} + k(r) \dfrac{s(x) - s(\ell)}{s(r) - s(\ell)} - k(x)$,

(b') $k(\ell) + m(r)\ (s(x) - s(\ell)) - k(x)$,

(c') $k(r) - m(\ell)\ (s(r) - s(x)) - k(x)$.

Le résultat précédent et la proposition 1.3 du chapitre II permettent d'obtenir des conditions suffisantes pour l'existence de moments exponentiels de I_S.

Proposition 2.3. (cf. MUSIELA [47])

Supposons que l'une des conditions suivantes soit vérifiée :

(a) $k(\ell) < \infty$, $k(r) < \infty$ et $\int_{x_0}^{r} \int_{x_0}^{y} m(dt)\ s(dy) < 1$,

où x_0 est donné par $m(x_0) \int_{\ell}^{r} s(dy) = \int_{\ell}^{r} m(y)\ s(dy)$,

(b) $k(\ell) < \infty$, $k(r) = \infty$, $s(r) = \infty$, $m(r) < \infty$ et $\int_{\ell}^{r} \int_{y}^{r} m(dt)\ s(dy) < 1$,

(c) $k(\ell) = \infty$, $s(\ell) = -\infty$, $m(\ell) > -\infty$, $k(r) < \infty$ et $\int_{\ell}^{r} \int_{\ell}^{y} m(dt)\ s(dy) < 1$

Alors il existe une constante positive c telle que pour tout $x \in D$ on a :

$$P_x \exp(I_S) \leq c < \infty.$$

Désignons par k_1 et m_1 les fonctions k et m respectivement lorsque $f \equiv 1$. Le résultat suivant est une conséquence immédiate de la proposition précédente.

Corollaire 2.1. (cf. MUSIELA [47])

Supposons que l'une des conditions suivantes soit vérifiée :

(a) $k_1(\ell) < \infty$, $k_1(r) < \infty$ et $\lambda < \left(\int_{x_0}^{r} \int_{x_0}^{y} m_1(dt)\, s(dy)\right)^{-1}$,

où x_0 est donné par $m_1(x_0) \int_{\ell}^{r} s(dy) = \int_{\ell}^{r} m(y)\, s(dy)$,

(b) $k_1(\ell) < \infty$, $k_1(r) = \infty$, $s(r) = \infty$, $m_1(r) < \infty$ et $\lambda < \left(\int_{\ell}^{r} \int_{y}^{r} m_1(dt)\, s(dy)\right)^{-1}$,

(c) $k_1(\ell) = \infty$, $s(\ell) = -\infty$, $m_1(\ell) > -\infty$, $k_1(r) < \infty$ et $\lambda < \left(\int_{\ell}^{r} \int_{\ell}^{y} m_1(dt)\, s(dy)\right)^{-1}$.

Alors il existe une constante positive c telle que pour tout $x \in D$ on a :

$$P_x \exp(\lambda S) \leq c < \infty.$$

Exemple

Supposons que $b(x) = 0$ et $a(x) = 1$. Alors $e(x) = 1$, $s(x) = x-c$, $m(x) = 2\int_c^x f(y)dy$ et $k(x) = 2\int_c^x \int_c^y f(z)dz\, dy$. Si $D = \mathbb{R}$, alors $P_x\{I_S = \infty\} = 1$. Supposons maintenant que $D =]\ell, \infty[$, avec $\ell > -\infty$. Nous pouvons démontrer qu'alors :

(a) $P_x\{I_S < \infty\} = 1$,

(b) $P_x\, I_S < \infty$ si et seulement si $\int_{\ell}^{\infty} f(y)dy < \infty$,

(c) si $\int_{\ell}^{\infty} \int_{y}^{\infty} f(t)dt\, dy < \frac{1}{2}$, on a $P_x \exp(I_S) < \infty$.

Supposons enfin que $D =]\ell, r[$ avec $-\infty < \ell < r < \infty$. Alors si $\int_{x_0}^{r} \int_{x_0}^{y} f(t)dt\, dy < \frac{1}{2}$, où x_0 est donné par $k'(x_0)(r-\ell) = k(r)-k(\ell)$, on a $P_x \exp(I_S) < \infty$.

Remarque 2.1.

Notons que sous les hypothèses de la proposition 2.3 la fonction $x \to u(x) = P_x \exp(I_S)$ appartient à $C^2(D)$ et vérifie l'équation $(L + f)u = 0$ sur D. De plus si (a) est satisfaite alors $u(\ell) = u(r) = 1$. Par suite, l'équation a une solution ne s'annulant pas sur $D \cup \{\ell, r\}$. Cela améliore les résultats classiques (cf. COPPEL [14] p. 20 et 60 par exemple). De même si (b) (resp (c)) est satisfaite, alors $u(\ell) = 1$ (resp. $u(r) = 1$) et l'équation a une solution ne s'annulant pas sur $D \cup \{\ell\}$ (resp. $D \cup \{r\}$) (cf. COPPEL [14] p. 28).

Supposons maintenant que l'équation $(L + f)u = 0$ ait une solution ne s'annulant pas sur $D \cup \{\ell\}$. Alors il existe (cf. COPPEL [14] p. 7) une solution u ne s'annulant pas et telle que $\int_c^r \frac{e}{u^2}(y)dy = \infty$. De plus, pour toute solution ne s'annulant pas v, linéairement indépendante de u on a $\int_c^r \frac{e}{v^2}(y)dy < \infty$. Bien sûr la solution u est unique à un facteur constant près. Elle est appelée solution principale. Désignons alors par u la solution principale normalisée par $u(\ell) = 1$. On a la caractérisation suivante de u.

Proposition 2.4. (cf. MUSIELA [46])

Supposons que l'équation $(L + f)\, u = 0$ ait une solution ne s'annulant pas sur $D \cup \{\ell\}$. La solution principale normalisée est donnée par

$$u(x) = P_x(\exp(I_{S^\ell}),\ S^\ell < S^r),$$

où $S^a = \inf\{t \geq 0\ ;\ X_t = a\}$. Pour toute solution v ne s'annulant pas sur $D \cup \{\ell\}$, normalisée par $v(\ell) = 1$, linéairement indépendante de u, on a $u < v$ sur D.

Remarque 2.2.

Pour $L = \frac{1}{2}\,\mathbb{D}^2$ une assertion analogue a été obtenue par CHUNG et VARADHAN [13]. Pour un L général le résultat précédent a été annoncé sans démonstration par KHAS'MINSKII [28]. A notre connaissance aucune preuve n'en a été fournie jusqu'à présent dans la littérature.

§ 3. Une remarque sur le problème de Dirichlet

Dans cette section $P_x = Q_{0,x}$, $x \in \mathbb{R}^d$ désigne la mesure de diffusion homogène associée à l'opérateur

$$L = \frac{1}{2}\, a\cdot\mathbb{D}_x^2 + b\cdot\mathbb{D}_x$$

les fonctions a de $\mathbb{R}^d$ dans $\mathbb{R}^d \otimes \mathbb{R}^d$ et b de $\mathbb{R}^d$ dans $\mathbb{R}^d$ ne dépendant pas de t.

Soit D un sous-ensemble ouvert connexe de $\mathbb{R}^d$. Considérons le problème consistant à trouver une fonction u telle que

$$(7)\qquad \begin{aligned} (L+g)u &= 0 && \text{sur } D\\ u &> 0 && \text{sur } D\\ u &\equiv 1 && \text{sur } \partial D \end{aligned}$$

où la fonction g de $\mathbb{R}^d$ dans $\mathbb{R}$ est borélienne et bornée sur les compacts.

D'un point de vue probabiliste, si $P_x\{T(\partial D) < S\} > 0$ où $T(\partial D) = \inf\{t \geq 0 \;;\; X_t \in \partial D\}$, pour $x \in D$, on peut espérer qu'une solution soit donnée par l'expression

$$u(x) = P_x(e_{T(\partial D)} \;;\; T(\partial D) < S),$$

où

$$e_{T(\partial D)} = \exp\left(\int_0^{T(\partial D)} g(X_t)\,dt\right).$$

Soit $v \in C^2(D)$ une solution de (7) et P_x^v la mesure de diffusion associée à l'opérateur $L + a \frac{\mathbb{D}_x v}{v} \cdot \mathbb{D}_x$. Comme la densité locale de P_x^v par rapport à P_x est donnée par (4) on peut montrer que pour $x \in D$

$$u(x) \leq v(x)\ P_x^v\{T(\partial D) < S\}$$

De plus si $P^v\{T(\partial D) < S\} \in C^2(D)$ alors $v(x)\ P_x^v\{T(\partial D) < S\}$ est solution de (7).

Soit maintenant $v \in C^2(D)$ une solution de (7) telle que $P_x^v\{T(\partial D) < S\} = 1$, $x \in D$. Alors la mesure P_x^v est absolument continue par apport à P_x sur $G_{T(\partial D)}$ et

$$u(x) = P_x(e_{T(\partial D)},\ T(\partial D) < S) =$$

$$= P_x\left(\frac{v(x)}{v(X_{T(\partial D)})}\ N^v_{T(\partial D)}\ ,\ T(\partial D) < S\right) = v(x).$$

Enfin si u est solution de (7) alors $P_x^u\{T(\partial D) < S\} = 1$. Cela est un analogue multidimensionnel de la proposition 2.4.

P A R T I E II

PROBLEMES D'ESTIMATION DE PARAMETRES POUR LES PROCESSUS DE DIFFUSION MULTIDIMENSIONNELS

Dans cette partie de notre travail nous nous intéressons au problème statistique d'estimation du paramètre $\theta \in \Theta \subset \mathbb{R}^p$ de la diffusion $P_x^\theta = Q_{0,x}^\theta$ associée à l'opérateur différentiel

$$\Lambda^\theta = \mathbb{D}_t + L^\theta, \; L^\theta = \frac{1}{2}\, a\cdot\mathbb{D}_x^2 + (b + ac_\theta)\cdot\mathbb{D}_x .$$

Les fonctions continues $a : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$, $b, c_\theta : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$, $\theta \in \Theta$, sont supposées localement lipschitziennes en x uniformément par rapport à t. Nous nous proposons d'estimer θ au vu de l'observation du processus (X_t) jusqu'à un temps d'arrêt $T < S$. La structure statistique à étudier est alors

$$(W, G_T, \{P_x^\theta \; ; \; \theta \in \Theta \subset \mathbb{R}^p\}).$$

Soit P_x la diffusion associée à l'opérateur différentiel

$$\Lambda = \mathbb{D}_t + L \; , \; L = \frac{1}{2}\, a\cdot\mathbb{D}_x^2 + b\cdot\mathbb{D}_x .$$

Comme $P_x^\theta\{\int_0^T c_\theta\cdot ac_\theta(Y_t)dt < \infty\} = 1$, d'après la proposition 2.1 du chapitre I, partie I, la mesure P_x^θ est absolument continue par rapport à la mesure P_x sur G_T et la densité est donnée par

$$(8) \qquad \exp(\int_0^T c_\theta(Y_t)\cdot dM_t - \frac{1}{2}\int_0^T c_\theta\cdot a\, c_\theta(Y_t)dt)$$

où

$$M_t = X_t - \int_0^t b(Y_s)ds.$$

Cela permet d'aborder le problème de construction d'estimateurs de θ par des méthodes de statistique classique. Si, par exemple, $c_\theta(t,x) = c(t,x)\theta$, où la fonction continue $c : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ est localement lipschitzienne en x uniformément par rapport à t, alors la méthode du maximum de vraisemblance conduit à envisager un estimateur du paramètre θ qui soit solution de l'équation

$$\int_0^T c^*(Y_t)(dX_t - b(Y_t)dt) = \int_0^T c^* ac(Y_t)dt\; \theta .$$

Evidemment, si la matrice $\int_0^T c^* ac(Y_t)dt$ est définie positive P_x^θ p.s., alors l'estimateur de maximum de vraisemblance est donné par

$$(9) \qquad \hat{\theta}_T = \left(\int_0^T c^* ac(Y_t)dt\right)^{-1} \int_0^T c^*(Y_t)(d\, X_t - b(Y_t)dt).$$

Nous montrons que cette matrice est effectivement définie positive en particulier pour les diffusions linéaires hypoelliptiques (cf. chapitre I § 4) et pour les diffusions bilinéaires elliptiques (cf. chapitre II § 3).

Si la fonction $c_\theta(t,x)$ ne dépend pas linéairement de θ le problème est plus compliqué. Dans le chapitre I § 3 nous étudions un cas particulier de dépendance non linéaire en θ. Nous construisons un estimateur qui est fonction d'une statistique exhaustive minimale et complète pour θ.

Une fois construit un estimateur $\hat{\theta}_T$ il convient de cerner ses propriétés. Nous abordons d'une part l'étude de propriétés asymptotiques lorsque $T \nearrow S$ et d'autre part celle de propriétés d'optimalité pour T fixé. Nous analysons également certains plans séquentiels d'estimation du paramètre. Voici une bibliographie permettant de situer nos résultats sans prétendre à l'exhaustivité.

Les propriétés asymptotique d'estimateurs des paramètres de processus de diffusion ont été abondamment étudiées. Les cas de diffusions linéaires elliptiques et hypoelliptiques ont été considérés par ARATO [1, 2], LE BRETON [33], LE BRETON et MUSIELA [39]. Des modèles bilinéaires ont été traités par TARASKIN [57], LE BRETON et MUSIELA [37] dans le cas unidimensionnel et par MUSIELA [45] dans le cas multidimensionnel. La situation générale d'une dépendance linéaire du coefficient de translation par rapport au paramètre a été abordé par TARASKIN [58], BROWN et HEWITT [11], LEE et KOZIN [40]. Dans tous ces articles, l'hypothèse d'ergodicité des diffusions correspondantes était essentielle. D'autres travaux ont été conduits sans cette hypothèse. Ainsi NOVIKOV [53] et FEIGIN [17,18] ont utilisé des méthodes de martingales pour étudier l'estimateur de maximum de vraisemblance du paramètre réel arbitraire θ pour un processus d'Ornstein-Uhlenbeck unidimensionnel gouverné par l'opérateur différentiel $L^{\theta} = \frac{1}{2} \mathbb{D}_x^2 + \theta x \mathbb{D}_x$. Le résultat de convergence de Novikov a été généralisé par LE BRETON et MUSIELA [39] pour les diffusions linéaires hypoelliptiques d-dimensionnelles et par MUSIELA [45] pour les diffusions bilinéaires elliptiques. Récemment KUTOYANTS [32] a utilisé le concept de LE CAM de familles de distributions localement asymptotiquement gaussiennes pour étudier les propriétés asymptotiques des estimateurs de maximum de vraisemblance et de Bayes dans le cas d'une dépendance non linéaire de la fonction $c_{\theta}(t,x)$ par rapport au paramètre θ. Pour d'autres informations sur des études asymptotiques d'estimateurs on peut consulter le livre de BASAWA et PRAKASA RAO [4]. Les propriétés d'optimalité des estimateurs pour T fixé ont été moins étudiées. Un estimateur sans biais de la translation d'un processus de WIENER observé jusqu'à un temps d'arrêt T a été construit par FEREBEE [20]. Dans le cas où T est déterministe ARATO [2], GRENANDER [22], HAJEK [23, 24] et HOLEVO [25] ont considéré le problème d'estimation sans biais de variance minimum. Les meilleurs estimateurs sans biais des coefficients de régression et des composantes de la variance ont été construits par MUSIELA et ZMYSLONY [48,49].

Construire un plan séquentiel d'estimation du paramètre θ consiste à choisir à la fois un temps d'observation T (aléatoire) et un estimateur $\hat{\theta}_T$. On cherche naturellement un couple $(T, \hat{\theta}_T)$ qui soit optimal en un certain sens dans une certaine classe de tels plans. NOVIKOV [53] a comparé des méthodes séquentielles et non séquentielles d'estimation du paramètre θ d'un processus d'Ornstein-Uhlenbeck unidimensionnel ($L = \frac{1}{2} \mathbb{D}_x^2 + \theta x \mathbb{D}_x$). DVORETZKY, KIEFER et WOLFOWITZ [15] ont montré que pour la translation d'un processus de WIENER les plans séquentiels à temps déterministes sont minimax pour la fonction de perte quadratique. ROZANSKI [54] a étudié des estimations séquentielles minimax pour la moyenne d'un processus d'Ornstein-Uhlenbeck. MUSIELA [43] et LE BRETON et MUSIELA [39] ont construit, pour des diffusions linéaires des plans séquentiels sans biais de variance minimum minimax et admissibles. Le cas général de la dépendance linéaire du coefficient de translation par rapport au paramètre a été abordé par LIPTSER et SHIRYAEV [41] pour d = 1 et par BOROBEICIKOV et KONEV [9] pour d arbitraire.

Cette partie est organisée de la manière suivante :

Le chapitre I est consacré aux problèmes d'estimation pour les processus gaussiens markoviens. Nous montrons d'abord que, sous certaines conditions de régularité, tout tel processus peut être considéré comme la solution d'un problème "linéaire" de martingales. Puis, utilisant les résultats de la partie I, nous étudions les propriétés d'optimalité d'estimateurs.

Le chapitre II traite de l'étude des diffusions bilinéaires multidimensionnelles. Nous donnons des conditions nécessaires et suffisantes pour la stationnarité au second-ordre. Nous construisons les meilleurs estimateurs linéaires des coefficients de régression et le filtre linéaire optimal. Pour les diffusions bilinéaires homogènes nous étudions les propriétés asymptotiques de l'estimateur de maximum de vraisemblance.

Dans le chapitre III nous examinons certains problèmes d'estimation pour une diffusion homogène unidimensionnelle P_x^θ associée à l'opérateur différentiel $L^\theta = \frac{1}{2} a \mathbb{D}_x^2 + (b + ac\theta)\mathbb{D}_x$. Nous étudions, dans tous les cas possibles, le comportement asymptotique de l'estimateur de maximum de vraisemblance $\hat{\theta}_T$ défini par (9) et nous déterminons des plans séquentiels optimaux pour l'estimation de θ.

C H A P I T R E I

DIFFUSIONS LINEAIRES

Soit P une mesure gaussienne sur (W, G) de moyenne m et de covariance K. Nous supposons que $K(t) = K(t, t)$ est définie positive pour tout $t \in \mathbb{R}_+$ et que les fonctions m', K' et $K_1(\cdot) = \lim_{h\to 0^+} \frac{1}{h} (K(\cdot + h, \cdot) - K(\cdot))$ sont définies et continues. Définissons la famille de lois de probabilité $\{P_x : x \in \mathbb{R}^d\}$ par $P_x(\cdot) = P(\cdot | X_0 = x)$.

§ 1. Processus de Gauss-Markov comme solutions d'un problème de martingales

Supposons que la famille $\{P_x : x \in \mathbb{R}^d\}$ soit markovienne (cf. IKEDA et WATANABE [26], par exemple). Par un théorème de corrélation normale (cf. LIPTSER et SHIRYAEV [41], par exemple) on a, pour $0 \le s \le t$

$$P_x(X_t | X_s) = m(t) + K(t,s)\ K^{-1}(s)(X_s - m(s))$$

et

$$P_x(X_s - P_x X_s)(X_t - P_x X_t)^* = (K(s) - K(s,0)K^{-1}(0)K(0,s))K^{-1}(s)K(s,t).$$

De plus, pour $0 \le s \le u \le t$

$$P_x(X_s - P_x X_s)(X_t - P_x X_t)^* = P_x(X_s - P_x X_s)P_x(X_t - P_x X_t | X_u)^* =$$

$$(K(s) - K(s,0)K^{-1}(0)K(0,s))K^{-1}(s)K(s,u)K^{-1}(u)K(u,t).$$

Cela implique que pour $0 \le s \le u \le t$

$$K(s,t) = K(s,u)\ K^{-1}(u)\ K(u,t).$$

Définissons les fonctions $A = K' - K_1 - K_1^*$, $B = K_1\ K^{-1}$ et $b = m' - Bm$. Il est clair que $m' = b + Bm$, $K' = A + BK + KB^*$ et $\frac{\partial}{\partial t} K(t,s) = B(t)K(t,s)$, $t > s$.

Nous pouvons alors démontrer

Proposition 1.1 (cf. MUSIELA et ZMYSLONY [48])

Le processus stochastique $(X_t - \int_0^t (b(s) + B(s)\, X_s)ds)$ est une $((G_t), P_x)$ martingale locale sur $\mathbb{R}_+$ ayant $(\int_0^t A(s)ds)$ pour processus de variation quadratique tensorielle. De plus, pour tout $x \in \mathbb{R}^d$, $P_x\{X_0 = x\} = 1$ et le processus $(g(Y_t) - \int_0^t \Lambda\, g(Y_s)ds)$, où

$$\Lambda = \mathbb{D}_t + L,\ L = \frac{1}{2}\, a\cdot\mathbb{D}_x^2 + b\cdot\mathbb{D}_x$$

$$a(t, x) = A(t),\ b(t, x) = b(t) + B(t)x$$

est une $((G_t), P_x)$ martingale locale sur $\mathbb{R}_+$ pour tout $g \in C^{1,2}(\mathbb{R}_+ \times \mathbb{R}^d)$.

§ 2. Estimation de coefficients de régression

Supposons que la fonction moyenne m de la mesure gaussienne P soit de la forme

$$m(t) = M(t)\theta\ ,\ M(0) = 0,$$

où la fonction $M : \mathbb{R}_+ \to \mathbb{R}^d \otimes \mathbb{R}^p$ appartient à $C^1(\mathbb{R}_+)$ $(M \not\equiv 0)$ et $\theta \in \Theta \subset \mathbb{R}^p$. Supposons que $A(t) = K'(t) - K_1(t) - K_1^*(t)$ soit définie positive pour tout $t > 0$. Désignons par $P_x(\cdot) = P(\cdot | X_0 = x)$ et supposons que pour tout $\theta \in \Theta$ la famille de lois de probabilité $\{P_x : x \in \mathbb{R}^d\}$ soit markovienne.

Soit $T > 0$ un nombre réel. Nous envisageons d'estimer le paramètre θ au vu de l'observation du processus (X_t) sur l'intervalle $[0,T]$ sous P_0^θ. Alors la structure statistique à étudier est

$$(W, \mathcal{G}_T, \{P_0^\theta : \theta \in \Theta\}),$$

où P_0^θ est la mesure de diffusion associée à l'opérateur

$$\Lambda^\theta = \mathbb{D}_t + L^\theta \ , \ L^\theta = \frac{1}{2} A \cdot \mathbb{D}_x^2 + (Bx + (M'-BM)\theta) \cdot \mathbb{D}_x,$$

avec $A = K' - K_1^* - K_1$, $B = K_1 K^{-1}$.

Alors, utilisant (9) nous pouvons démontrer :

Proposition 2.1. (cf. MUSIELA et ZMYSLONY [48])

Supposons que l'intérieur de Θ soit non vide et que la matrice $\Sigma_T = \int_0^T (M'-BM)^* A^{-1} (M'-BM)(t)dt$ soit définie positive. Alors

(10) $$\hat{\theta}_T = \Sigma_T^{-1} \int_0^T (M'-BM)^* A^{-1}(t)(dX_t - B(t)X_t dt)$$

est l'estimateur sans biais de variance minimum de θ.

Exemple 2.1.

Soit $A(t) = I$ et $B(t) = 0$. Alors P_0^θ est la mesure de Wiener de translation $M'(t)\theta$. La proposition précédente assure que si l'intérieur de Θ est non vide et si la matrice $\int_0^T M'^* M'(t)dt$ est définie positive, alors $(\int_0^T M'^* M'(t)dt)^{-1} \int_0^T M'^*(t)dX_t$ est l'estimateur sans biais de variance minimum de θ.

§ 3. Estimation de coefficients de régression et de composantes de la variance

Soit $T > 0$ un nombre réel et considérons maintenant la structure statistique

$$(W, G_T, \{P_0^\theta : \theta \in \Theta\})$$

où P_0^θ est la mesure gaussienne de moyenne

$$m(t) = M(t)\beta \ , \ M(0) = M(T) = 0,$$

et de covariance

$$K(s,t) = st\,K + s \wedge t\,I.$$

Nous supposons que la fonction $M : \mathbb{R}_+ \to \mathbb{R}^d \otimes \mathbb{R}^p$ appartient à $C^1(\mathbb{R}_+)$ et que la matrice K est définie positive et s'écrit $K = \Sigma_{i=1}^k \sigma_i V_i$, où $V_1, \dots, V_k$ sont des matrices symétriques linéairement indépendantes.

Supposons de plus que $\beta \in \Xi \subset \mathbb{R}^p$, $\sigma = (\sigma_1, \dots, \sigma_k)^* \in \Sigma \subset \mathbb{R}^k$, ($\theta = \binom{\beta}{\sigma} \in \Xi \times \Sigma = \Theta$) les intérieurs de Ξ et Σ étant non vides. Nous cherchons à estimer le paramètre θ. Remarquons que sous P_0^θ le processus (X_t) peut se représenter sous la forme :

$$X_t = M(t)\beta + tN + w_t,$$

où

- . w est un processus de Wiener standard,
- . N est un vecteur gaussien centré de covariance K, F_0-mesurable,
- . w et N sont indépendants.

Sous P_0^θ le processus (X_t) est donc gaussien markovien. De plus, comme $K'(t) = 2tK+I$ et $K_1(t) = tK$ on a $A(t) = I$ et $B(t) = K(I+tK)^{-1}$. D'après la proposition 1.1 , P_0^θ est la mesure de diffusion associée à l'opérateur

$$\Lambda^\theta = \mathbb{D}_t + L^\theta, \quad L^\theta = \frac{1}{2}\,\Delta + (B(t)x + (M'-BM)(t)\theta)\cdot\mathbb{D}_x\,.$$

D'après (8), pour tout $\theta \in \Theta$, la mesure P_0^θ est absolument continue par rapport à la mesure P_0^0 sur G_T et la densité est donnée par

$$(11)\quad c(\theta)\, h(T, X_T) \exp(-\frac{1}{2T}\, X_T^*(I+TK)^{-1}\, X_T + \beta^* \int_0^T M'^*(t)\, dX_t),$$

où

$$\log c(\theta) = -\frac{1}{2}\int_0^T \operatorname{tr} K(I+tK)^{-1}dt - \frac{1}{2}\beta^* \int_0^T (M'-BM)^*(M'-BM)(t)dt\ \beta$$

et

$$\log h(T, X_T) = -\frac{1}{2T} X_T^* X_T.$$

Nous allons pouvoir exhiber une statistique exhaustive minimale pour $\{P_0^\theta : \theta \in \Theta\}$. Remarquons que l'ensemble $\{I + TK : K = \Sigma_{i=1}^n \sigma_i V_i, \sigma \in \Sigma\}$ contient un sous-ensemble ouvert de la variété linéaire $V_I = I + V$, où V est le sous-espace de $\mathbb{R}^d \otimes \mathbb{R}^d$ engendré par $V_1, \ldots, V_k$. Soit W_I la plus petite variété telle que

$$\{(I + TK)^{-1} : \sigma \in \Sigma\} \subset W_I.$$

Alors il existe une matrice W_0 telle que $W_I = W_0 + W$, où W est un sous-espace parallèle à W_I. Soit $W_1, \ldots, W_r$ une base de W. Alors $(I + TK)^{-1}$ peut être représentée de manière unique sous la forme

$$(I + TK)^{-1} = W_0 + \Sigma_{i=1}^r c_i(\sigma) W_i.$$

D'après (11) et les considérations précédentes on obtient :

<u>Proposition 3.1.</u> (cf. MUSIELA et ZMYSLONY [49])

Supposons que la matrice $\int_0^T M'^* M'(t)dt$ soit définie positive. Alors le vecteur

$$Z_T = \left(\left(\int_0^T M'^*(t)dX_t\right)^*, X_T^* W_i X_T, i = 1, \ldots, r\right)$$

est une statistique exhaustive minimale pour $\{P_0^\theta : \theta \in \Theta\}$.

Nous donnons maintenant des conditions nécessaires et suffisantes pour que la statistique exhaustive minimale Z_T soit complète.

Un sous-espace D de l'espace des matrices symétriques est dit quadratique si $V \in D$ implique que $V^2 \in D$ (cf. SEELY [55]).

Proposition 3.2. (cf. MUSIELA et ZMYSLONY [49])

Supposons que la matrice $\int_0^T M'^* M'(t)dt$ soit définie positive. Alors les conditions suivantes sont équivalentes :

(a) la statistique Z_T est complète,
(b) V est un sous-espace quadratique,
(c) V = W.

Nous sommes alors en mesure de décrire le meilleur estimateur sans biais de θ. Soit $N = (\mathrm{tr}\, V_i\, V_j)$, $i,j = 1,\ldots, k$ une matrice symétrique. Remarquons que N^{-1} existe compte tenu de ce que $V_1,\ldots, V_k$ sont linéairement indépendantes. En outre, posons $U_T = (U_T^1,\ldots, U_T^k)^*$ où $U_T^i = \frac{1}{T^2} X_T^* V_i X_T - \frac{1}{T} \mathrm{tr}\, V_i$, $i = 1,\ldots, k$. Nous pouvons démontrer :

Proposition 3.3. (cf. MUSIELA et ZMYSLONY [49])

Supposons que la matrice $\int_0^T M'^* M'(t)dt$ soit définie positive.

Alors $\hat{\theta}_T = \binom{\hat{\beta}_T}{\hat{\sigma}_T}$, où

$$\hat{\beta}_T = \int_0^T M'^* M'(t)dt^{-1} \int_0^T M'^*(t)dX_t \quad \text{et} \quad \hat{\sigma}_T = N^{-1} U_T$$

est l'estimateur sans biais de variance minimum de θ si et seulement si V est un sous-espace quadratique.

Remarque 3.1.

Tout sous-espace quadratique D est une algèbre de Jordan pour $A \circ B = \frac{1}{2}(AB + BA)$.

§ 4. Propriétés asymptotiques d'estimateurs pour les paramètres de diffusions gaussiennes homogènes hypoelliptiques

Soit $T > 0$ un nombre réel. Nous étudions ici la structure statistique

$$(W, G_T, \{P_0^\theta : \theta \in \mathbb{R}^d \otimes \mathbb{R}^d\})$$

où P_0^θ est la mesure de diffusion homogène associée à l'opérateur différentiel

$$L^\theta = \frac{1}{2} A \cdot \mathbb{D}_x^2 + \theta x \cdot \mathbb{D}_x$$

Rappelons que, sous P_0^θ, le processus (X_t) est gaussien centré et de covariance

$$K(s,t) = e^{\theta s} \int_0^{s \wedge t} e^{-\theta u} A e^{-\theta^* u} du \, e^{\theta^* t}.$$

Nous supposons que $K(t,t)$ est définie positive pour tout $t > 0$ c'est-à-dire que la diffusion est hypoelliptique (cf. CHALEYAT-MAUREL et ELIE [12]). Nous cherchons à estimer le paramètre $\theta \in \mathbb{R}^d \otimes \mathbb{R}^d$. D'après (8) la densité de P_0^θ par rapport à P_0^0 sur G_T est donnée par

$$\exp(\operatorname{tr} A^+ (\int_0^T dX_t X_t^* \theta^* - \frac{1}{2} \theta \int_0^T X_t X_t^* dt \, \theta^*)),$$

où A^+ désigne la pseudo-inverse de Moore-Penrose de la matrice A. Nous pouvons alors démontrer :

<u>Proposition 4.1.</u> (cf. LE BRETON et MUSIELA [36, 39])

La matrice $\int_0^T X_t X_t^* dt$ est P_0^θ p.s. définie positive. L'estimateur de maximum de vraisemblance de θ est donné par

$$\hat{\theta}_T = \int_0^T dX_t X_t^* (\int_0^T X_t X_t^* dt)^{-1} .$$

Nous étudions maintenant les propriétés asymptotiques de l'estimateur $\hat{\theta}_T$.

Proposition 4.2. (cf. LE BRETON et MUSIELA [36, 39]).

Supposons que θ soit une matrice stable. Alors l'estimateur $\hat{\theta}_T$ est fortement convergent i.e.

$$\lim_{T\to\infty} \hat{\theta}_T = \theta \quad P_0^\theta \quad \text{p.s.}$$

De plus, le vecteur aléatoire $T^{\frac{1}{2}} \operatorname{vec}(\hat{\theta}_T - \theta)$ est asymptotiquement gaussien de moyenne nulle et de matrice de covariance $\left(\int_0^\infty e^{\theta s} A e^{\theta^* s} ds\right)^{-1} \otimes A$.

Lorsque la matrice θ n'est pas supposée stable nous ne savons pas, en général, répondre positivement à la question de la convergence de l'estimateur $\hat{\theta}_T$. Cependant nous avons résolu le problème dans la situation particulière suivante : supposons que $\theta = \Sigma_{i=1}^p \theta_i B_i$ où $\theta_i \in \mathbb{R}$ et B_i sont des matrices connues telles que pour $i \neq j$ on ait

$$B_i^* A^+ B_j + B_j^* A^+ B_i = 0 \quad \text{et} \quad AA^+ B_i \neq 0, \; i = 1, \ldots, p.$$

Proposition 4.3. (cf. LE BRETON et MUSIELA [39]).

L'estimateur de maximum de vraisemblance $\hat{\theta}_{i,T}$ de θ_i donné par

$$\hat{\theta}_{i,T} = \frac{\int_0^T X_t^* B_i^* A^+ dX_t}{\int_0^T X_t^* B_i^* A^+ B_i X_t dt}$$

est fortement convergent i.e.

$$\lim_{T\to\infty} \hat{\theta}_{i,T} = \theta_i \quad P_0^\theta \quad \text{p.s.}$$

Exemple 4.1.

On sait (cf. par exemple BASAWA et PRAKASA RAO [4]) que l'axe instantané de rotation de la terre se déplace par rapport à l'axe mineur de l'ellipsoïde terrestre. Ce déplacement comprend une part périodique et une part de fluctuations. Cette dernière peut raisonnablement être modélisée par une mesure de diffusion homogène P_0^θ associée à l'opérateur infinitésimal

$$L^\theta = \frac{1}{2}\Delta + \theta x \cdot \mathbb{D}_x,$$

où $\theta = \theta_1 B_1 + \theta_2 B_2$ et

$$B_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad , \quad B_2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} .$$

Comme $B_1^* B_2 + B_2^* B_1 = 0$, les estimateurs de maximum de vraisemblance $\hat\theta_{1,T}$ et $\hat\theta_{2,T}$ de θ_1 et θ_2 sont donnés par :

$$\hat\theta_{1,T} = \frac{\int_0^T X_t^1\, dX_t^1 + \int_0^T X_t^2\, dX_t^2}{\int_0^T ((X_t^1)^2 + (X_t^2)^2)dt},$$

$$\hat\theta_{2,T} = \frac{\int_0^T X_t^1\, dX_t^2 - \int_0^T X_t^2\, dX_t^1}{\int_0^T ((X_t^1)^2 + (X_t^2)^2)dt}.$$

Il était acquis (cf. TARASKIN [58]) que ces estimateurs sont convergents lorsque la matrice θ est stable. En vertu de la proposition 4.3 nous pouvons affirmer que ces estimateurs sont encore fortement convergents lorsque cette hypothèse n'est plus satisfaite.

Pour d'autres résultats dans cette direction on peut consulter BELLACH [5, 6] et LE BRETON [34].

§ 5. Estimation séquentielle

Construire un plan séquentiel pour l'estimation d'un paramètre consiste (cf. LIPTSER et SHIRYAEV [41]) à choisir à la fois un temps d'observation (aléatoire) $\tilde{T}$ et un estimateur $\tilde{\theta}_T$. Nous cherchons un plan séquentiel $\delta = (T, \tilde{\theta}_T)$ qui soit optimal en un certain sens dans une classe de tels plans. Nous nous plaçons dans le cas d'un paramètre θ réel et utilisons des critères de comparaison des plans séquentiels définis à partir de la fonction de perte quadratique usuelle et d'une fonction de coût.

Supposons que le processus (X_t) soit unidimensionnel et que la mesure P_0^θ soit du type défini dans le § 2. avec $d = p = 1$.

Nous nous intéressons d'abord à la fonction de perte

$$L(\theta, \delta) = (\tilde{\theta}_T - \theta)^2 + C(T)$$

où $\delta = (T, \tilde{\theta}_T)$ est un plan séquentiel et C est une fonction de coût. Nous supposons que C est positive continue et telle que $\lim_{t\to\infty} C(t) = \infty$.

Désignons par D l'ensemble de tous les plans séquentiels δ qui ont une fonction de risque $R(\theta, \delta) = P_0^\theta\, L(\theta, \delta)$ finie pour tout $\theta \in \Theta$. Un plan séquentiel $\delta_0 \in D$ est dit minimax dans la classe D si

$$\sup_\Theta R(\theta, \delta_0) = \inf_D \sup_\Theta R(\theta, \delta).$$

Considérons $\delta_0 = (T_0, \hat{\theta}_{T_0})$ le plan séquentiel à temps déterministe T_0 tel que

$$\frac{1}{\int_0^{T_0} (M'-BM)^2 A^{-1}(t)\,dt} + C(T_0) = \inf_T \left(\frac{1}{\int_0^{T} (M'-BM)^2 A^{-1}(t)\,dt} + C(T)\right)$$

et $\hat{\theta}_{T_0}$ est donné par (10). Nous pouvons alors énoncer

Proposition 5.1. (cf. MUSIELA [44])

Le plan séquentiel à temps déterministe δ_0 est minimax dans la classe D.

Plaçons nous maintenant dans la situation où le coût d'observation C n'est pas pris en compte i.e. où le risque de δ est donné par $R(\theta, \delta) = P_0^\theta(\tilde{\theta}_T - \theta)^2$. Il est clair qu'il est ici nécessaire d'imposer des restrictions aux temps d'arrêt pouvant intervenir dans un plan séquentiel si on veut éviter que le temps d'arrêt optimal soit infini p.s.

Soit alors $T_0 > 0$ un nombre réel. Désignons par $D(T_0)$ l'ensemble de tous les plans séquentiels $\delta = (T, \tilde{\theta}_T)$ pour lesquels le risque $R(\theta, \delta) = P_0^\theta(\tilde{\theta}_T - \theta)^2$ est fini pour tout $\theta \in \Theta$ et tels que

$$P_0^\theta \int_0^T (M'-BM)^2 A^{-1}(t)dt \leq \int_0^{T_0} (M'-BM)^2 A^{-1}(t)dt \ , \ \theta \in \Theta.$$

Un plan séquentiel $\delta_1 \in D(T_0)$ est dit meilleur que $\delta_2 \in D(T_0)$ si $R(\theta, \delta_1) \leq R(\theta, \delta_2)$ pour tout $\theta \in \Theta$ et si cette inégalité est stricte pour au moins un $\theta \in \Theta$. Un plan séquentiel $\delta \in D(T_0)$ est dit admissible dans $D(T_0)$ s'il n'existe aucun plan dans $D(T_0)$ qui soit meilleur que δ. Enfin un plan séquentiel $\delta_0 \in D(T_0)$ est dit minimax dans $D(T_0)$ si

$$\sup_\Theta R(\theta, \delta_0) = \inf_{D(T_0)} \sup_\Theta R(\theta, \delta).$$

Soit alors $\delta_0 = (T_0, \hat{\theta}_{T_0})$ le plan séquentiel à temps déterministe où $\hat{\theta}_{T_0}$ est donné par (10). Nous pouvons démontrer

Proposition 5.2. (cf. MUSIELA [44])

(i) Si $\Theta = \mathbb{R}$, alors le plan séquentiel δ_0 est admissible et minimax dans $D(T_0)$;

(ii) Si $\Theta =]\theta_0, \infty[$, $\theta_0 > -\infty$, alors δ_0 is minimax dans $D(T_0)$;

(iii) Le plan δ_0 est sans biais de variance minimum parmi tous les plans sans biais dans $D(T_0)$.

Supposons désormais que le processus (X_t) soit d-dimensionnel et que, pour $\theta \in \Theta \subset \mathbb{R}$, P_0^θ est la mesure de diffusion homogène associée à l'opérateur différentiel

$$L^\theta = \frac{1}{2} A \cdot \mathbb{D}_x^2 + \theta B x \cdot \mathbb{D}_x$$

où $A A^+ B \neq 0$. Nous nous plaçons donc sous des hypothèses analogues à celles de la proposition 4.3 mais avec $p = 1$. Supposons que la variance de X_t sous P_0, $\theta \in \Theta$, soit définie positive pour tout $t > 0$.

Soit $H > 0$ un nombre réel. Désignons par $D(H)$ l'ensemble de tous les plans séquentiels $\delta = (T, \tilde{\theta}_T)$ pour lesquels le risque $R(\theta, \delta) = P_0^\theta(\tilde{\theta}_T - \theta)^2$ est fini pour tout $\theta \in \Theta$ et tels que

$$P_0^\theta \int_0^T X_t^* B^* A^+ B X_t \, dt \leq H, \ \theta \in \Theta.$$

Soit alors $\delta_H = (T_H, \hat{\theta}_{T_H})$ le plan séquentiel défini par le temps d'arrêt

$$T_H = \inf\{t \geq 0 : \int_0^t X_s^* B^* A^+ B X_s \, ds = H\}$$

et l'estimateur

$$\hat{\theta}_{T_H} = \frac{1}{H} \int_0^{T_H} X_t^* B^* A^+ dX_t .$$

Remarquons que $\hat{\theta}_{T_H}$ est en fait l'estimateur de maximum de vraisemblance de θ construit au vu de l'observation du processus (X_t) sur l'intervalle aléatoire $[0, T_H]$. Nous pouvons démontrer

Proposition 5.3. (cf. LE BRETON et MUSIELA [39])

Le plan séquentiel δ_H a les propriétés suivantes :

(i) Sous P_0^θ, $\hat{\theta}_{T_H}$ est gaussien de moyenne θ et de variance $\frac{1}{H}$.

(ii) Si $\Theta = \mathbb{R}$ alors δ_H est admissible et minimax dans D(H) ;

(iii) Si $\Theta =]\theta_0, \infty[$, $\theta_0 > -\infty$, alors δ_H est minimax dans D(H) ;

(iv) δ_H est sans biais de variance minimum parmi tous les plans sans biais dans D(H) ;

(v) Il existe $\lambda > 0$ tel que $P_0^\theta \exp(\lambda\, T_H) < \infty$.

C H A P I T R E   II

DIFFUSIONS BILINEAIRES

Dans ce chapitre, nous considérons une diffusion multidimensionnelle bilinéaire non homogène $P_x = Q_{0,x}$, $x \in \mathbb{R}^d$, associée à l'opérateur infinitésimal

$$\Lambda = \mathbb{D}_t + L \, , \; L = \frac{1}{2} \, a \cdot \mathbb{D}_x^2 + b \cdot \mathbb{D}_x$$

où

$$a(t,x) = \Sigma_{i=1}^{p} (A_i(t)x + a_i(t))(A_i(t)x + a_i(t))^*$$

(12)

$$b(t,x) = B(t)x + b(t).$$

Les fonctions b, $a_i : \mathbb{R}_+ \to \mathbb{R}^d$ et B, $B_i : \mathbb{R}_+ \to \mathbb{R}^d \otimes \mathbb{R}^d$, $i = 1,\ldots, p$ sont supposées continues.

D'abord nous étudions les propriétés du second-ordre du processus (X_t) sous P_x et donnons des conditions nécessaires et suffisantes pour sa stationnarité au second-ordre. Puis nous obtenons une représentation linéaire de (X_t) et construisons les meilleurs estimateurs linéaires des coefficients de régression et le filtre linéaire optimal. Enfin, dans le cas homogène, nous étudions les propriétés asymptotiques d'estimateurs de maximum de vraisemblance.

§ 1. Propriétés du second-ordre

Soit μ une loi de probabilité sur $(\mathbb{R}^d, B(\mathbb{R}^d))$ de moyenne m_0 et de variance K_0 et $P_\mu = \int_{\mathbb{R}^d} P_x \, \mu(dx)$.

Nous étudions la stationnarité au second-ordre du processus (X_t) sous P_μ. Remarquons que sous P_μ (X_t) est bien du second ordre et notons par m et K respectivement sa moyenne et sa covariance.

Nous pouvons démontrer :

Proposition 1.1. (cf. LE BRETON et MUSIELA [35, 38])

Les fonctions moyenne et covariance de (X_t) sous P_μ sont données par

$$m' = Bm + b, \quad m(0) = m_0$$

$$K(s,t) = \begin{cases} \phi_s \, \phi_t^{-1} \, K(t,t) & \text{si} \quad t \leq s \\ K(s,s) \, (\phi_t \, \phi_s^{-1})^* & \text{si} \quad s \leq t \end{cases}$$

où $(K(t,t))_{t \geq 0}$ vérifie

$$K' = BK + KB^* + \Sigma_{i=1}^{p} A_i \, K \, A_i^* + \Sigma_{i=1}^{p} (A_i m + a_i)(A_i m + a_i)^*, \quad K(0,0) =$$

La fonction de covariance satisfait pour tout $0 \leq s \leq u \leq t$ la propriété

$$K(s,t) = K(s,u) \, K^+(u,u) \, K(u,t)$$

où K^+ désigne la pseudo-inverse de Moore-Penrose de la matrice K. Cette analogie avec le cas gaussien-markovien (cf. partie II, chapitre I, § 1.) permet de démontrer

Proposition 1.2. (cf. LE BRETON et MUSIELA [35, 38]

Le processus (X_t) est stationnaire au second-ordre sous P_μ si et seulement si

$$B(t) \, m_0 + b(t) = 0 \qquad \text{pour tout } t \in \mathbb{R}_+$$

et s'il existe une matrice constante B telle que les conditions suivantes soient satisfaites pour tout $t \in \mathbb{R}_+$

$$B(t) \, K_0 K_0^+ = BK_0 K_0^+ = K_0 K_0^+ \, B = B,$$

$$BK_0 + K_0B^* + \Sigma_{i=1}^p A_i(t)K_0A_i^*(t) + \Sigma_{i=1}^p (A_i(t)m_0 + a_i(t))(A_i(t)m_0 + a_i(t))^* = 0.$$

Si les conditions ci-dessus sont vérifiées alors la fonction de covariance K du processus (X_t) est donnée par

$$K(s,t) = \begin{cases} e^{B(s-t)}K_0 & \text{si} \quad s \geq t \\ K_0\, e^{B^*(t-s)} & \text{si} \quad s \leq t . \end{cases}$$

§ 2. Représentation linéaire, filtrage linéaire optimal et meilleurs estimateurs linéaires des coefficients de régression.

Suivant les idées développées par Liptser et Shiryaev dans le chapitre 15 de leur livre [41] nous définissons la notion de représentation linéaire d'une mesure de diffusion :

Définition 2.1.

La mesure de diffusion $\bar{P}_x$ associée à l'opérateur différentiel

$$\bar{\Lambda} = \mathbb{D}_t + \bar{L} \quad , \quad \bar{L} = \frac{1}{2}\,\bar{A} \cdot \mathbb{D}_x^2 + (\bar{B}x + \bar{b}) \cdot \mathbb{D}_x$$

où les fonctions $\bar{A}, \bar{B} : \mathbb{R}_+ \to \mathbb{R}^d \otimes \mathbb{R}^d$ et $\bar{b} : \mathbb{R}_+ \to \mathbb{R}^d$ sont continues est appelée représentation linéaire de la mesure de diffusion P_x si les fonctions moyenne et covariance de (X_t) sous P_x et $\bar{P}_x$ coïncident.

Soit alors P_x la mesure de diffusion non-homogène définie par (12) et $\bar{P}_x$ la mesure de diffusion associée à l'opérateur différentiel

$$\bar{\Lambda} = \mathbb{D}_t + \bar{L} \quad , \quad \bar{L} = \frac{1}{2}\,\bar{A} \cdot \mathbb{D}_x^2 + (Bx + b) \cdot \mathbb{D}_x \ ,$$

où

(13) $\bar{A} = K'' - BK - KB^*$

et

(14) $K' = BK + KB^* + \Sigma_{i=1}^{p} A_i K A_i^* + \Sigma_{i=1}^{p} (A_i m + a_i)(A_i m + a_i)^*$, $K(0) = 0$,

$m' = Bm + b$, $m(0) = x$.

Nous pouvons démontrer

Proposition 2.1. (cf. LE BRETON et MUSIELA [35, 38])

La mesure $\overline{P}_x$ est une représentation linéaire de P_x.

Nous allons maintenant utiliser cette représentation linéaire pour résoudre un problème de filtrage.
Supposons qu'on observe un processus (U_t) tel que

$$(U_t - U_0 - \int_0^t (B_0(s) X_s + b_0(s)) ds)$$

soit une $((G_t), \overline{P}_x)$ et $((G_t), P_x)$ martingale locale sur $\mathbb{R}_+$ de variation quadratique $(\int_0^t B_1(s) ds)$.

Nous supposons que les fonctions $b_0 : \mathbb{R}_+ \to \mathbb{R}^q$, $B_0 : \mathbb{R}_+ \to \mathbb{R}^q \otimes \mathbb{R}^d$ et $B_1 : \mathbb{R}_+ \to \mathbb{R}^q \otimes \mathbb{R}^q$ $(B_1 B_1 > 0)$ sont continues.

Nous disons que $(\hat{X}_t)$ est, sous P_x, le filtre linéaire optimal de (X_t) au vu de l'observation (U_t) si $(\hat{X}_t)$ est, sous $\overline{P}_x$, le filtre optimal de (X_t) au vu de l'observation (U_t).

Proposition 2.2. (cf. LE BRETON et MUSIELA [35, 38]

Sous P_x, le filtre linéaire optimal $(\hat{X}_t)$ de (X_t) au vu de l'observation (U_t) est donné par

$$d\hat{X}_t = (B(t)\hat{X}_t + b(t))dt + \gamma B_0^*(B_1 B_1^*)^{-1}(t)(dU_t - (B_0(t)\hat{X}_t + b_0(t))dt),$$

$\hat{X}_0 = x$, où la variance $\gamma(t) = \overline{P}_x(X_t - \hat{X}_t)(X_t - \hat{X}_t)^*$ vérifie

$$\gamma' = \overline{A} + B\gamma + \gamma B^* - \gamma B_0^*(B_1 B_1^*)^{-1} B_0 \gamma, \ \gamma(0) = 0.$$

Nous envisageons l'étude de la stabilité du filtre linéaire optimal. Nous supposons que B_0, B_1 et B ne dépendent pas de t et que le processus (X_t) est stationnaire au second-ordre sous P_μ pour une certaine loi de probabilité μ de variance K .

Soit $\bar{A} = -BK_0 - K_0B^*$ et soit $(\gamma(t))$ la solution de $\gamma' = \bar{A} + B\gamma + \gamma B^* - \gamma B_0(B_1B_1^*)^{-1}B_0\gamma$.

Proposition 2.3. (cf. LE BRETON et MUSIELA [35, 38])

Supposons que la paire (B, B_0) soit observable et que la paire $(B, \bar{A}^{1/2})$ soit contrôlable. Alors la limite $\lim_{t\to\infty} \gamma(t)$ existe, ne dépend pas de la valeur initiale γ_0 et est l'unique (dans la classe de matrices symétriques positives) solution de l'équation de Riccati matricielle

$$\bar{A} + B\gamma + \gamma B^* - \gamma B_0^*(B_1B_1^*)^{-1} B_0\gamma = 0 .$$

Envisageons maintenant l'utilisation de la représentation linéaire dans le cadre du problème d'estimation. Soit $T > 0$ un nombre réel. Considérons la structure statistique

$$(W, G_T, \{P_0^\theta : \theta \in \Theta\})$$

où P_0^θ est la mesure de diffusion associée à l'opérateur différentiel

$$\Lambda^\theta = \mathbb{D}_t + L^\theta \ , \quad L^\theta = \frac{1}{2}\, a\cdot\mathbb{D}_x^2 + b_\theta\cdot\mathbb{D}_x$$

a étant comme dans (12), $b_\theta(t,x) = B(t)x + (M'-BM)(t)\theta$, la fonction $M : \mathbb{R}_+ \to \mathbb{R}^d \otimes \mathbb{R}^p$ appartenant à $C^1(\mathbb{R}_+)$ et $\theta \in \Theta \subset \mathbb{R}^p$; Nous supposons que $\bar{A}(t)$ et $K(t)$, données par (13) et (14) respectivement, sont définies positives pour tout $t > 0$.

Nous disons que l'estimateur $\hat{\theta}_T$ est, sous P_0^θ, le meilleur estimateur linéaire sans biais du paramètre θ si $\hat{\theta}_T$ est, sous $\bar{P}_0^\theta$, l'estimateur sans biais de variance minimum.

D'après la proposition 2.1 , chapitre I, partie II, nous pouvons énoncer :

Proposition 2.4.

Supposons que l'intérieur de Θ soit non vide et que la matrice $\Sigma_T = \int_0^T (M' - BM)^* \bar{A}^{-1} (M' - BM)(t)dt$ soit définie positive. Alors

$$\hat{\theta}_T = \Sigma_T^{-1} \int_0^T (M' - BM)^* \bar{A}^{-1}(t)(dX_t - B(t) X_t dt)$$

est le meilleur estimateur linéaire sans biais de θ.

§ 3. Propriétés asymptotiques d'estimateurs de maximum de vraisemblance dans le cas d-dimensionnel elliptique homogène

Nous considérons ici la structure statistique

$$(W, G_T, \{P_0^\theta : \theta \in \mathbb{R}^q\})$$

où $T > 0$ est un nombre réel et P_0^θ est une mesure de diffusion homogène associée à l'opérateur

$$L^\theta = \frac{1}{2} a \cdot \mathbb{D}_x^2 + B_\theta x \cdot \mathbb{D}_x ,$$

où $a(x) = \Sigma_{k=1}^r A_k xx^* A_k^* + AA^*$, AA^* est définie positive, $B_\theta = \Sigma_{k=1}^q \theta_k B_k$ et les matrices $d \times d$ $B_1, \ldots, B_q$ sont linéairement indépendantes.
En ce qui concerne l'estimation du paramètre $\theta = (\theta_1, \ldots, \theta_q)^* \in \mathbb{R}^q$ nous pouvons démontrer

Proposition 3.1. (cf. MUSIELA [45])

La matrice $((\int_0^T X_t^* B_k^* a^{-1}(X_t) B_\ell X_t dt))$ est P_0^θ p.s. définie positive.
L'estimateur de maximum de vraisemblance du paramètre θ est donné par

$$(15) \quad \hat{\theta}_T = ((\int_0^T X_t^* B_k^* a^{-1}(X_t) B_\ell X_t dt))^{-1} (\int_0^T X_t^* B_k^* a^{-1}(X_t) dX_t) .$$

Rappelons que (cf. BROCKETT [10]) si la matrice $B_\theta \otimes I + I \otimes B_\theta + \Sigma_{k=1}^{r} A_k \otimes A_k$ est stable, alors il existe une unique loi de probabilité μ_θ invariante pour (P_x^θ) qui est centrée du second ordre et qui possède une densité p_θ qui est C^∞. En outre, μ_θ est gaussienne de variance K si et seulement si $A_k K + K A_k^* = 0$, $k = 1,\dots, r$ et $B_\theta K + K B_\theta + \Sigma_{k=1}^{r} A_k K A_k^* + AA^* = 0$. Nous pouvons aussi montrer que si $B_\theta = \theta I$, $r = 1$, $A_1 = aI$, $a \neq 0$ où θ et a sont des réels tels que $2\theta + a^2 < 0$ alors

$$p_\theta(x) = C(\theta)(1 + a^2 x^*(AA^*)^{-1} x)^{-\frac{d+1}{2} + \frac{\theta}{a^2}}$$

Les propriétés asymptotiques de l'estimateur $\hat{\theta}_T$ donné par (15) sont les suivantes :

Proposition 3.2. (cf. MUSIELA [45])

Si la matrice $B_\theta \otimes I + I \otimes B_\theta + \Sigma_{k=1}^{r} A_k \otimes A_k$ est stable, alors l'estimateur $\hat{\theta}_T$ défini par (15) est fortement convergent i.e.

$$\lim_{T\to\infty} \hat{\theta}_T = \theta \qquad P_0^\theta \quad \text{p.s.}$$

En outre, le vecteur aléatoire $T^{1/2}(\hat{\theta}_T - \theta)$ est asymptotiquement gaussien de moyenne nulle et de matrice de covariance

$$((\int_{\mathbb{R}^d} x^* B_k^* a^{-1}(x) B_\ell x \mu_\theta(dx)))^{-1}.$$

Utilisant la proposition 2.1 du chapitre II, partie I, nous pouvons aussi démontrer

Proposition 3.3. (cf. MUSIELA [45])

Si $\det B_k \neq 0$, $k = 1,\dots, q$ et si pour tout $x \in \mathbb{R}^d$ et $k \neq \ell$, $k,\ell = 1,\dots, q$ on a $x^* B_k^* a^{-1}(x) B_\ell x = 0$, alors l'estimateur de maximum de vraisemblance $\hat{\theta}_T$ défini par (15) est fortement convergent.

Sans hypothèse d'ergodicité (même sous les conditions de la proposition précédente) nous ne savons pas, en général, déterminer la distribution asymptotique de $\hat{\theta}_T$. En vue de donner un aperçu du problème nous donnons maintenant les résultats d'une étude concernant un cas particulier en dimension d = 2.

Soit d = q = 2, r = 1 et

$$B_1 = A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad , \qquad B_2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad ,$$

$A_1 = aB_2$, $a \in \mathbb{R}$. Alors $\hat{\theta}_T = (\hat{\theta}_{1,T}, \hat{\theta}_{2,T})^*$ est donné par

$$\hat{\theta}_{1,T} = V_{0,T}^{-1}\left(\int_0^T X_t^1 \, dX_t^1 + \int_0^T X_t^2 \, dX_t^2\right) ,$$

$$\hat{\theta}_{2,T} = V_{a,T}^{-1} \int_0^T \frac{X_t^1 \, dX_t^2 - X_t^2 \, dX_t^1}{1 + a^2((X_t^1)^2 + (X_t^2)^2)}$$

où

$$V_{a,T} = \int_0^T \frac{(X_t^1)^2 + (X_t^2)^2}{1 + a^2((X_t^1)^2 + (X_t^2)^2)} \, dt$$

Nous pouvons démontrer

Proposition 3.4. (cf. MUSIELA [45])

(i) L'estimateur $\hat{\theta}_T$ est fortement convergent.

(ii) Si $2\theta_1 + a^2 \neq 0$ alors la variable aléatoire $V_{0,T}^{1/2}(\hat{\theta}_{1,T} - \theta_1)$ est asymptotiquement gaussienne centrée de variance égale à un.

(iii) Si $2\theta_1 + a^2 = 0$, alors la distribution de $V_{0,T}^{1/2}(\hat{\theta}_{1,T} - \theta_1)$ ne dépend pas de T et est la distribution (qui n'est pas gaussienne) de la variable aléatoire

$$\frac{(W_1^1)^2 + (W_1^2)^2 - 2}{2\int_0^1((W_t^1)^2 + (W_t^2)^2)\,dt}$$

où W^1 et W^2 sont des mouvements browniens standards indépendants.

(iv) Les variables aléatoires $V_{0,T}^{1/2}(\hat{\theta}_{2,T} - \theta_1)$ et $V_{a,T}^{1/2}(\hat{\theta}_{2,T} - \theta_2)$ sont indépendantes.

(v) La distribution de $V_{a,T}^{1/2}(\hat{\theta}_{2,T} - \theta_2)$,est gaussienne centrée de variance un.

§ 4. Propriétés asymptotiques d'estimateurs du maximum de vraisemblance dans le cas unidimensionnel dégénéré

La structure statistique étudiée est

$$(W,\ G_T,\ \{P_0^\theta\ :\ \theta = \binom{m}{\beta} \in \mathbb{R}^2\})$$

où P_0^θ est la mesure de diffusion homogène associée à l'opérateur infinitésimal

$$L^\theta = \frac{1}{2}\,(Ax + a)^2\,\mathbb{D}_x^2 + \beta(m - x)\,\mathbb{D}_x\,.$$

Nous estimons $\theta = \binom{m}{\beta}$ au vu de l'observation de (X_t) sur l'intervalle de temps fixé $[0,T]$. L'estimateur de maximum de vraisemblance $\hat{\theta}_T = \binom{\hat{m}_T}{\hat{\beta}_T}$ est donné par :

$$\hat{m}_T = \frac{S_0(T)I_2(T) - S_1(T)I_1(T)}{S_0(T)I_1(T) - S_1(T)I_2(T)}\ ,$$

(16)

$$\hat{\beta}_T = \frac{S_0(T)I_1(T) - S_1(T)I_0(T)}{I_0(T)I_2(T) - I_1^2(T)}\ ,$$

où

$$I_j(T) = \int_0^T X_t^j(AX_t + a)^{-2}\,dt\ ,\ j = 0,\ 1,\ 2\ ,$$

$$S_j(T) = \int_0^T X_t^j(AX_t + a)^{-2}\,dX_t\ ,\ j = 0,\ 1\ .$$

Nous étudions les propriétés asymptotiques de $\hat{\theta}_T$ lorsqu'il existe une unique mesure de probabilité invariante μ_θ.

Proposition 4.1. (cf. LE BRETON et MUSIELA [37])

Supposons que l'une des conditions suivantes soit satisfaite :

(a) $A = 0,\ a \neq 0,\ \beta > 0$
(b) $A \neq 0,\ Am + a \neq 0,\ \beta > \frac{1}{2} A^2$.

Alors il existe une unique loi de probabilité μ_θ invariante pour (P_x^θ). Dans le cas (a) μ_θ est gaussienne de moyenne m et de variance $a^2/2\beta$. Dans le cas (b) μ_θ est la distribution d'une variable aléatoire U^{-1} sign $(\frac{a}{A} + m) - \frac{a}{A}$ où U est de loi $\Gamma(1 + 2\beta A^{-2}, (2\beta|Am + A|)^{-1}|A|^3)$.

En outre, l'estimateur de maximum de vraisemblance $\hat{\theta}_T$ défini par (16) est fortement convergent et le vecteur $T^{1/2}(\hat{\theta}_T - \theta)$ est asymptotiquement gaussien centré et de matrice de covariance

$$\begin{bmatrix} -1 & 0 \\ m & \beta \end{bmatrix}^{-1} Q_\theta^{-1} \begin{bmatrix} -1 & m \\ 0 & \beta \end{bmatrix}^{-1}$$

où

$$Q_\theta = \int_{\mathbb{R}} f(x)\ f^*(x)\ \mu_\theta(dx)$$

avec

$$f(x) = \begin{bmatrix} x(Ax + a)^{-1} \\ (Ax + a)^{-1} \end{bmatrix} .$$

Pour une étude détaillée de ce problème on pourra consulter LE BRETON et MUSIELA [37].

CHAPITRE III

DIFFUSIONS HOMOGENES UNIDIMENSIONNELLES

Dans ce chapitre, nous considérons le problème d'estimation du paramètre $\theta \in \Theta \subset \mathbb{R}$ de la diffusion homogène P_x^θ associée à l'opérateur

$$L^\theta = \frac{1}{2}\, a\, \mathbb{D}_x^2 + (b + ac\,\theta)\,\mathbb{D}_x .$$

Les fonctions a, b, c : $\mathbb{R} \to \mathbb{R}$ sont supposées localement lipschitziennes et $a > 0$.

Nous cherchons à estimer θ au vu de l'observation du processus (X_t) jusqu'à un temps d'arrêt $T < S$. D'après (9) l'estimateur de maximum de vraisemblance $\hat{\theta}_T$ est donné par :

$$\hat{\theta}_T = \frac{\int_0^T c(X_t)\,(dX_t - b(X_t)\,dt)}{\int_0^T c^2\, a(X_t)\,dt}$$

D'abord nous étudions dans tous les cas possibles le comportement asymptotique de $\hat{\theta}_T$ quand $T \to S$. Puis, dans le cas particulier où $b = 0$, nous montrons que $\hat{\theta}_T$ est P_x^θ p.s. convergent et nous construisons des plans séquentiels optimaux pour l'estimation de θ.

§ 1. Comportement asymptotique de $\hat{\theta}_T$

Soit $M_t = \int_0^t c(X_s)\,(dX_s - (b + ac\,\theta)(X_s)\,ds)$. Le processus (M_t) est une $((G_t), P_x^\theta)$ martingale locale sur $[0, S[$ de processus de variation quadratique $(\int_0^t c^2\, a(X_s)\,ds)$. De plus $\hat{\theta}_T = \theta + \langle M\rangle_T^{-1} M_T$. Alors, en vertu de la loi forte des grands nombres pour les martingales, on a :

$$\lim_{T\to S} \hat{\theta}_T = \theta + \langle M\rangle_S^{-1} M_S\, \mathbb{1}_{\{\langle M\rangle_S < \infty\}} \qquad P_x^\theta \quad \text{p.s.}$$

Soit $r \in \mathbb{R}$ tel que $c(r) \neq 0$. Définissons les fonctions

$$e_\theta(x) = \exp(-\int_r^x 2(\frac{b}{a} + \theta c)(y)dy), \quad m_\theta(x) = 2\int_r^x \frac{c^2}{e_\theta}(y)dy,$$

(18)

$$s_\theta(x) = \int_r^x e_\theta(y)dy, \quad k_\theta(x) = \int_r^x e_\theta\, m_\theta(y)dy.$$

D'après les considérations ci-dessus, utilisant la proposition 2.1 , chapitre III, partie I, il vient :

Proposition 1.1.

(i) Si $k_\theta(-\infty) = \infty$ et $k_\theta(\infty) = \infty$, alors pour tout $x \in \mathbb{R}$

$$P_x^\theta\{\lim_{T\to S} \hat{\theta}_T = \theta\} = 1.$$

(ii) Si $k_\theta(-\infty) < \infty$, $k_\theta(\infty) = \infty$, $s_\theta(\infty) < \infty$, alors pour tout $x \in \mathbb{R}$

$$P_x^\theta\{\lim_{T\to S} \hat{\theta}_T = \theta\} = \frac{s_\theta(x) - s_\theta(-\infty)}{s_\theta(\infty) - s_\theta(-\infty)}$$

et

$$P_x^\theta\{\lim_{T\to S} \hat{\theta}_T = \theta + \langle M\rangle_S^{-1} M_S\, \mathrm{I\!I}_{\{\langle M\rangle_S<\infty\}}\} = \frac{s_\theta(\infty) - s_\theta(x)}{s_\theta(\infty) - s_\theta(-\infty)}.$$

(iii) Si $k_\theta(-\infty) = \infty$, $s_\theta(-\infty) > -\infty$, $k_\theta(\infty) < \infty$ alors pour tout $x \in \mathbb{R}$

$$P_x^\theta\{\lim_{T\to S} \hat{\theta}_T = \theta\} = \frac{s_\theta(\infty) - s_\theta(x)}{s_\theta(\infty) - s_\theta(-\infty)}$$

et

$$P_x^\theta\{\lim_{T\to S} \hat{\theta}_T = \theta + \langle M\rangle_S^{-1} M_S\, \mathrm{I\!I}_{\{\langle M\rangle_S<\infty\}}\} = \frac{s_\theta(x) - s_\theta(-\infty)}{s_\theta(\infty) - s_\theta(-\infty)}.$$

(iv) Si l'une des situations suivantes est réalisée

(a) $k_\theta(-\infty) < \infty$, $k_\theta(\infty) < \infty$,

(b) $k_\theta(-\infty) < \infty$, $k_\theta(\infty) = \infty$, $s_\theta(\infty) = \infty$

(c) $k_\theta(-\infty) = \infty$, $s_\theta(-\infty) = -\infty$, $k_\theta(\infty) < \infty$,

alors pour tout $x \in \mathbb{R}$

$$P_x^\theta\{\lim_{T\to S} \hat{\theta}_T = \theta + \langle M\rangle_S^{-1} M_{S-} \mathrm{I\!I}_{\{\langle M\rangle_S < \infty\}}\} = 1.$$

Dans la suite nous supposons que

$$a = \sigma^2,\ b = \frac{1}{2}\sigma\sigma',\ c = \sigma^{-1},$$

où $\sigma \in C^1(\mathbb{R})$ et $\sigma > 0$. On peut vérifier que les fonctions e_θ, m_θ, s_θ, k_θ définies en (18) sont données par :

$$e_\theta(x) = \frac{\sigma(r)}{\sigma(x)} \exp(-2\theta \int_r^x \frac{1}{\sigma}(y)dy),$$

$$m_\theta(x) = \frac{1}{\theta\sigma(r)}(\exp(2\theta \int_r^x \frac{1}{\sigma}(y)dy) - 1),$$

$$s_\theta(x) = \frac{\sigma(r)}{2\theta}(1 - \exp(-2\theta \int_r^x \frac{1}{\sigma}(y)dy)),$$

$$k_\theta(x) = \frac{1}{\theta}\int_r^x \frac{1}{\sigma}(y)dy + \frac{1}{2\theta^2}(\exp(-2\theta \int_r^x \frac{1}{\sigma}(y)dy) - 1),$$

où $m_0(x) = \lim_{\theta\to 0} m_\theta(x)$ (resp. $s_0(x) = \lim_{\theta\to 0} s_\theta(x)$ et $k_0(x) = \lim_{\theta\to 0} k_\theta(x)$). Notons que l'estimateur de maximum de vraisemblance $\hat{\theta}_T$ défini par (17) s'écrit :

$$\hat{\theta}_T = \frac{\int_{X_0}^{X_T} \frac{1}{\sigma}(y)dy}{T}.$$

Alors utilisant la proposition précédente nous pouvons démontrer

Proposition 1.2.

(i) Si $\int_{-\infty}^{r} \frac{1}{\sigma}(y)\,dy = \infty$ et $\int_{r}^{\infty} \frac{1}{\sigma}(y)\,dy = \infty$, alors pour tout θ, $x \in \mathbb{R}$

$$P_x\{\lim_{T\to S} \hat{\theta}_T = \theta,\ S = \infty\} = 1.$$

(ii) Supposons que $\int_{-\infty}^{r} \frac{1}{\sigma}(y)\,dy = \infty$ et $\int_{r}^{\infty} \frac{1}{\sigma}(y)\,dy < \infty$.

Si $\theta < 0$, alors pour tout $x \in \mathbb{R}$

$$P_x^{\theta}\{\lim_{T\to S} \hat{\theta}_T = \theta,\ S = \infty\} = 1 - \exp(2\theta \int_{x}^{\infty} \frac{1}{\sigma}(y)\,dy)$$

et

$$P_x^{\theta}\{\lim_{T\to S} \hat{\theta}_T = -\infty,\ S < \infty\} = \exp(2\theta \int_{x}^{\infty} \frac{1}{\sigma}(y)\,dy).$$

Si $\theta \geq 0$, alors pour tout $x \in \mathbb{R}$

$$P_x^{\theta}\{\lim_{T\to S} \hat{\theta}_T = \frac{1}{S}\int_{x}^{\infty} \frac{1}{\sigma}(y)\,dy,\ S < \infty\} = 1.$$

(iii) Supposons que $\int_{-\infty}^{r} \frac{1}{\sigma}(y)\,dy < \infty$ et $\int_{r}^{\infty} \frac{1}{\sigma}(y)\,dy = \infty$.

Si $\theta > 0$, alors pour tout $x \in \mathbb{R}$

$$P_x^{\theta}\{\lim_{T\to S} \hat{\theta}_T = \theta,\ S = \infty\} = 1 - \exp(-2\theta \int_{-\infty}^{x} \frac{1}{\sigma}(y)\,dy)$$

et

$$P_x^{\theta}\{\lim_{T\to S} \hat{\theta}_T = \infty,\ S < \infty\} = \exp(-2\theta \int_{-\infty}^{x} \frac{1}{\sigma}(y)\,dy).$$

Si $\theta \leq 0$, alors pour tout $x \in \mathbb{R}$

$$P_x^{\theta}\{\lim_{T\to S} \hat{\theta}_T = -\frac{1}{S}\int_{-\infty}^{x} \frac{1}{\sigma}(y)\,dy,\ S < \infty\} = 1.$$

(iv) Si $\int_{-\infty}^{r} \frac{1}{\sigma}(y)dy < \infty$ et $\int_{r}^{\infty} \frac{1}{\sigma}(y)dy < \infty$, alors pour tout θ, $x \in \mathbb{R}$

$$P_x^\theta\{\lim_{T\to S} \hat{\theta}_T = \frac{\int_x^\infty \frac{1}{\sigma}(y)dy}{S} \mathbb{I}_{\{X_S=\infty\}} - \frac{\int_{-\infty}^x \frac{1}{\sigma}(y)dy}{S} \mathbb{I}_{\{X_S=-\infty\}}, S < \infty\} = 1.$$

§ 2. Etude du cas b = 0

Dans ce paragraphe nous supposons que $b = 0$. Montrons d'abord que $k_\theta(-\infty) = k_\theta(\infty) = \infty$ pour tout $\theta \in \mathbb{R}$. Comme $k_0(\infty) = \infty$ et pour $\theta \neq 0$

$$k_\theta(x) = \frac{1}{2\theta^2} \int_r^x e_\theta(y) \int_r^y \frac{1}{e_\theta} \left(\frac{e_\theta'}{e_\theta}\right)^2 (z)dz\, dy$$

on a $k_\infty(\infty) = \infty$ pourvu que $s_\theta(\infty) = \infty$. Supposons que $s_\theta(\infty) < \infty$. Comme pour $x > r$

$$\int_r^x \frac{1}{e_\theta} \left(\frac{e_\theta'}{e_\theta}\right)^2 (y)dy \geq \frac{\left(\frac{1}{e_\theta}(x) - 1\right)^2}{\int_r^x \frac{1}{e_\theta}(y)dy}$$

et $\lim_{x\to\infty} e_\theta(x) = 0$, on a aussi pour $x \geq x_0$

$$k_\theta(x) \geq c_1 + c_2 \log \int_r^x \frac{1}{e_\theta}(y)dy$$

pour certaines constantes $c_1, c_2 > 0$. Par suite $k_\theta(\infty) = \infty$ pour tout $\theta \in \mathbb{R}$. De la même manière, il est possible de montrer que $k_\theta(-\infty) = \infty$ pour tout $\theta \in \mathbb{R}$. Utilisant cela nous pouvons démontrer :

Proposition 2.1.

(i) L'estimateur de maximum de vraisemblence $\hat{\theta}_T$ donné en (17) (avec $b = 0$) est P_x^θ fortement convergent pour tout θ, $x \in \mathbb{R}$

(ii) Le plan séquentiel $\delta_H = (T_H, \hat{\theta}_{T_H})$, où

$$T_H = \inf\{t \geq 0 \int_0^t c^2 a(X_s)ds = H\}, \quad \hat{\theta}_{T_H} = \frac{1}{H}\int_0^{T_H} c(X_t)dX_t$$

a les propriétés suivantes :

(a) $\hat{\theta}_{T_H}$ est gaussien de moyenne θ et de variance $\frac{1}{H}$ sous P_x^θ.

(b) Si $\Theta = \mathbb{R}$, alors δ_H est admissible et minimax dans l'ensemble $D(H)$ de tous les plans séquentiels $\delta = (T, \hat{\theta}_T)$ pour lesquels le risque $P_x^\theta(\tilde{\theta}_T - \theta)^2$ est fini et $P_x^\theta \int_0^T c^2\, a(X_t)dt \leq H$ pour tout $\theta \in \mathbb{R}$.

(c) Si $\Theta =]\theta_0, \infty[$, $\theta_0 > -\infty$, alors δ_H est minimax dans $D(H)$.

(d) δ_H est sans biais de variance minimum parmi tous les plans sans biais dans $D(H)$.

Exemple

Supposons que $a(x) = (1+x^2)^2$ et $c(x) = \dfrac{x}{1+x^2}$.

L'estimateur de maximum de vraisemblance $\hat{\theta}_T$ est donné par :

$$\hat{\theta}_T = \frac{1}{2}\left(\frac{\log(1+X_T^2) - \log(1+X_0^2) + T}{\int_0^T X_t^2\, dt} - 1\right)$$

Il est intéressant de remarquer que $P_x^0\{S = \infty\} = 1$ et $P_x^1\{S < \infty\} = 1$. On peut même calculer la distribution de S sous P_x^1 : en effet sa transformée de Laplace est donnée par

$$P_x^1 \exp(-\lambda S) = \frac{\mathrm{ch}((2\lambda)^{1/2} \operatorname{Arc\,tg} x)}{\mathrm{ch}((2\lambda)^{1/2} \frac{\pi}{2})}, \quad \lambda > 0.$$

De plus, pour $0 \leq \lambda < \frac{1}{2}$, on a

$$P_x^1 \exp(\lambda S) = \frac{\cos((2\lambda)^{1/2} \operatorname{Arc\,tg} x)}{\cos((2\lambda)^{1/2} \frac{\pi}{2})}.$$

B I B L I O G R A P H I E

[1] M. ARATO

On the parameter estimation of processes satisfying a stochastic differential equation (in Russian) Studia Sci. Math. Hungar. 5, 11-15 (1970).

[2] M. ARATO

On the statistical examination of continuous state Markov processes, I, II, III, IV, Selected Transl. in Math. Statist. and Probability 4, 203-288 (1978).

[3] R. AZENCOTT

Behavior of diffusion semi-groups at infinity, Bull. Soc. Math. France, 102, 193-240 (1974).

[4] I.V. BASAWA and B.L.S. PRAKASA RAO

Statistical Inference for Stochastic Processes, Academic Press, London 1980.

[5] B. BELLACH

Consistency, Asymptotic Normality and Asymptotic Efficiency of the Maximum-Likelihood-Estimator in Linear Stochastic Differential Equations. Math. Operationsforch. Statist., Ser. Statistics, vol. 11, n° 2, 227-266, (1980).

[6] B. BELLACH

Parameter Estimators in Linear Stochastic Differential Equations and their Asympotic Properties, Math.Operationsforch. Statist., Ser. Statistics, vol. 14, n° 1, 141-191 (1983).

[7] R. N. BHATTACHARYA

Criteria for recurrence and existence of invariant measures for multidimensional diffusions. Ann. Probab., 6, 541-553 (1980).

[8] R.N. BHATTACHARYA and S. RAMASUBRAMANIAN

Recurrence and ergodicity of diffusions. J. of Multivariate Analysis 12, 95-122 (1982).

[9] S.E. BOROBEICIKOV and W. W. KONEV

Construction of sequential plans for parameters of recurrent type processes (in Russian) Math. Statist. and Appl., Publ. of Tomsk Univ., vol. 6, 72-81 (1980).

[10] R.W. BROCKETT

Parametrically stochastic linear differential equations, in : R. J. B. Wets, Ed., Stochastic systems : Modeling, Identification and Optimization, I, Math. Prog. Study, 5, 8-21, Armsterdam, North-Holland (1976).

[11] B.M. BROWN and J.I. HEWITT

Asymptotic likelihood theory for diffusion processes, J. Appl. Prob. 12, 228-238 (1975).

[12] M. CHALEYAT-MAUREL and L. ELIE

Diffusions gaussiennes, S.M.F., Astérisque, 84-85, 225-279 (1981).

[13] K.L. CHUNG and S.R.S. VARADHAN

Kac functional and Schrödinger equation, Studia Math. 68, 249-260 (1980).

[14] W.A. COPPEL

Disconjugacy, Lecture Notes in Math. 220, Springer-Verlag, Berlin, Heidelber, New-York, 1971.

[15] A. DVORETZKY, J. KIEFER and J. WOLFOWITZ

Sequential decision problems for processes with continuous time parameter. Problems of estimation. Ann. Math. Statist. 24, 403-415 (1953).

[16] E.B. DYNKIN

Markov Processes, Springer-Verlag, Berlin, Göttingen, Heidelberg, 1965.

[17] P.D. FEIGIN

Maximum likelihood estimation for continuous time stochastic processes, Adv. Appl. Prob. 8, 712-736 (1976).

[18] P.D. FEIGIN

Some comments concerning a curious singularity, J. Appl. Prob. 16, 440-444 (1979).

[19] W. FELLER

The parabolic differential equations and the associated semigroups of transformations, Annals of Mathematics 55, 468-519 (1952).

[20] B. FEREBEE

An unbiased estimator for the drift of a stopped Wiener process, J. Appl. Prob. 20, 94-102 (1983).

[21] A. FRIEDMAN

Stochastic Differential Equations and Applications. Vol. 1, 2, Academic Press, New-York, 1976.

[22] U. GRENANDER

Abstract Inference. John Wiley & Sons New-York, 1981.

[23] J. HAJEK

On a simple linear model in Gaussian processes, Proc. of the Second Prague Conf. on Prob. Th., 185-197, (1962).

[24] J. HAJEK

On linear statistical problems in stochastic processes. Czechoslovak Math. J., 12, 404-444 (1962).

[25] A.S. HOLEVO

On estimates of regression coefficients (in Russian) Th. of Probab. Appl. 14, 79-104 (1969).

[26] N. IKEDA and S. WATANABE

Stochastic differential Equations and Diffusion Processes, North-Holland Publishing Company, Armsterdam, Oxford, New-York ; Kodansha LTD. Tokyo, 1981.

[27] J.M. KABANOV, R.S. LIPTSER, A.N. SHIRYAEV

Absolute continuity and singularity of locally absolutely continuous probability distributions. (in Russian) I ; II, Math. Sbornik V. 107 (149) 364-415 (1978) ; 108 (150) n° 1, 32-61 (1979).

[28] R.Z. KHAS'MINSKII

Probability distribution for functionals of trajectories of a diffusion type stochastic processes. (in Russian), Dokl Akad. Naak SSSR, 104, 1, 22-25 (1955).

[29] R.Z. KHAS'MINSKII

On positive solutions of the equation $(\Lambda + g)u=0$.Theor.Probab. Appl.Vol.4,n°3,309-318 (1959).

[30] R.Z. KHAS'MINSKII

Ergodic properties of recurrent diffusion processes and stabilization of the Cauchy problem for parabolic equations. Theor.Probab.Appl.Vol.5,n°2 ,179-196 (1960).

[31] R.Z. KHAS'MINSKII

Stochastic stability of differential equations. Sijthoff & Noordhoff, Rockville, 1980.

[32] Y.A. KUTOYANTS

Parameter Estimation for Random Processes, Felderman Verlag, Berlin, to appear.

[33] A. LE BRETON

Parameter estimation in a vector linear stochastic differential equation, in : Transactions of the Seventh Prague Conference on Information Theory. Statistical Decision Functions and Random Processes, vol. A. 353-366 (1977).

[34] A. LE BRETON

Propriétés asymptotiques et estimation des paramètres pour les diffusions gaussiennes homogènes hypoelliptiques dans le cas purement explosif. C.R. Acad. Sci. Paris, to appear.

[35] A. LE BRETON and M. MUSIELA

Moments et filtrage linéaire d'un système stochastique bilinéaire à temps continue, C.R. Acad. Sc. Paris t. 292, I, 83-86 (1981).

[36] A. LE BRETON and M. MUSIELA

Estimation des paramètres pour les diffusions gaussiennes homogènes hypoelliptiques, C.R. Acad. Sc. Paris t. 294, I, 341-344 (1982).

[37] A. LE BRETON and M. MUSIELA

A study of a one dimensional bilinear differential model for stochastic processes. Probab. and Math. Statist. Vol. 4, Fasc. 1 (1983).

[38] A. LE BRETON and M. MUSIELA

A look at a bilinear model for multidimensional stochastic systems in continuous time, Statistics & Decisions 1, 285-303 (1983).

[39] A. LE BRETON and M. MUSIELA

Some parameter estimation problems for hypoelliptic homogeneous Gaussian diffusions. Proc. Banach Math. Center, Warsaw, to appear.

[40] T.S. LEE and F. KOZIN

Almost sure asymptotic likelihood theory for diffusion processes. J. Appl. Prob. 14, 527-537 (1977).

[41] R.S. LIPTSER and A.N. SHIRYAEV

Statistics of Random Processes I, II, Springer-Verlag, New-York, 1978.

[42] B. MAISONNEUVE

Une mise au point sur les martingales locales continues définies sur un intervalle stochastique. Séminaire de Probabilité XI, Lecture Notes in Mathematics 581, Springer-Verlag, Berlin (1977)

[43] H.P.. Mc KEAN JR.

Stochastic Integrals. Academic Press. New-York, London, 1969.

[44] M. MUSIELA

On sequential estimation of parameters of continuous Gaussian Markov processes. Probab. and Math. Statist. Vol. 2, Fasc. 1, 37-53 (1981).

[45] M. MUSIELA

Parameter estimation for bilinear elliptic diffusions. Statistics & Decisions, to appear.

[46] M. MUSIELA

On Kac functionals of one dimensional diffusions. Stochastic Process. Appl., to appear.

[47] M. MUSIELA

Divergence, convergence and moments of some integral functionals of diffusions, submited for publication.

[48] M. MUSIELA and R. ZMYSLONY

Estimation of regression parameters of Gaussian Markov processes. Lecture Notes in Statistics 2. Mathematical Statistics and Probability Theory. Springer-Verlag, New-York, Heidelberg, Berlin, 330-341, (1980).

[49] M. MUSIELA and R. ZMYSLONY

Estimation for some classes of Gaussian Markov processes. Lecture Notes in Statistics 2. Mathematical Statistics and Probability Theory. Springer-Verlag, New-York, Heidelberg, Berlin, 318-329 (1980).

[50] K. NARITA

On explosion and growth order of inhomogeneous diffusion processes. Yokohama Math. J. Vol. 28, 45-57 (1980).

[51] K. NARITA

Remarks on non explosion theorem for stochastic differential equations. Kodai Math. J. 5, 395-401 (1982).

[52] K. NARITA

No explosion cirteria for stochastic differential equations. J. Math. Soc. Japan. Vol. 34, n° 2, 191-203 (1982).

[53] A.A. NOVIKOV

Sequential estimation of the parameters of diffusion processes. Theor. Prob. Appl., Vol. 16, n° 2, 394-396 (1971).

[54] R. ROZANSKI

Minimax sequential estimation of the mean of an Ornstein-Uhlenbeck process. Zastos. Mat.,to appear.

[55] J. SEELY

Quadratic subspaces and completeness.
Ann. Math. Statist., 42, 710-721 (1971).

[56] D.W. STROOCK and S.R.S. VARADHAN

Multidimensional Diffusion Processes. Springer-Verlag, Berlin, Heidelberg, New-York, 1979.

[57] A.F. TARASKIN

The asymptotic normality of stochastic integrals and estimates of the coefficient of diffusion process transfer, in : Mathematical Physics n° 8 Naukova Dumka, Kiev, 149-163 (1970).

[58] A.F. TARASKIN

On the asymptotic normality of vector valued stochastic integrals and estimates of drift parameters of a multidimensional diffusion process. Theory Probab. and Math. Statist. 2, 209-224 (1974).

A N N E X E

Moments et filtrage linéaire d'un système stochastique bilinéaire à temps continu (*) - Note de A. Le Breton et M. Musiela

On considère un système dynamique régi par une équation différentielle stochastique bilinéaire vectorielle. D'abord on calcule les deux premiers moments du processus d'état et on donne des conditions nécessaires et suffisantes pour sa stationnarité au second ordre. Ensuite on détermine les équations du filtre linéaire optimal et on étudie leur comportement asymptotique.

We consider a dynamical system governed by a multidimensional bilinear stochastic differential equation. First we compute the mean and covariance functions of the state process and we give necessary and sufficient conditions for its second-order stationarity. Then we provide the equations of the optimal linear filter and we study their asymptotic behaviour.

1. PROCESSUS D'ETAT D'UN SYSTEME STOCHASTIQUE BILINEAIRE.

Considérons un système stochastique dont le processus d'état dans $\mathbb{R}^n$, $(X_t \,;\, t \geq 0)$, est solution de l'équation différentielle (au sens de Itô) :

(1) $$dX_t = [A_o(t)X_t + a_o(t)]\,dt + \sum_{j=1}^{d} [A_j(t)X_t + a_j(t)]\,dW_t^j \;;\; t \geq 0 \;;\; X_o = X(0) ,$$

où $W = (W^1, \dots, W^d)'$ est un mouvement brownien standard dans $\mathbb{R}^d$, les applications déterministes $A_j = ((A_j^{k,\ell} \,;\, k, \ell = 1, \dots, n))$ et $a_j = (a_j^1, \dots, a_j^n)'$, $j = 0, 1, \dots, d$ sont mesurables et telles que, pour tout $T > 0$, $k, \ell = 1, \dots, n$, on a :

$$\int_0^T |A_o^{k,\ell}(t)|\,dt < \infty\ ,\ \int_0^T |a_o^k(t)|\,dt < \infty\ ,\ \int_0^T |A_j^{k\ell}(t)|^2 dt < \infty\ ,\ \int_0^T |a_j^k(t)|^2 dt < \infty;\ j=1,\ldots,c$$

et l'état initial $X(0)$ est un vecteur aléatoire indépendant de W, admettant une espérance mathématique $m(0)$ et une matrice de covariances $K(0)$ non-singulière.

Désignons par $(\Phi_t\ ;\ t \geq 0)$ le processus solution de l'équation différentielle stochastique matricielle

$$(2) \qquad d\Phi_t = A_o(t)\,\Phi_t\,dt + \sum_{i=1}^{d} A_i(t)\,\Phi_t\,dW_t^i\ ;\ t \geq 0\ ;\ \Phi_o = I_n$$

où I_n désigne la matrice unité $n \times n$. Il est alors facile de montrer (cf. par exemple ([1])) que posant $D_t = \det \Phi_t$, $t \geq 0$, on a

$$D_t = \exp\{\int_0^t \operatorname{Tr} A_o(s)\,ds - \frac{1}{2}\sum_{j=1}^{d}\int_0^t \operatorname{Tr} A_j^2(s)\,ds + \sum_{j=1}^{d}\int_0^t \operatorname{Tr} A_j(s)\,dW_s^j\}\ ;\ t \geq 0\ .$$

Ainsi pour tout $t \geq 0$, Φ_t est inversible presque sûrement et on obtient aisément :

1.1. LEMME - Le processus d'état $(X_t\ ;\ t \geq 0)$ du modèle (1) est donné par

$$X_t = \Phi_t\{X(0) + \int_0^t \Phi_s^{-1}[a_o(s) - \sum_{j=1}^{d} A_j(s)\,a_j(s)]\,ds + \sum_{j=1}^{d}\int_0^t \Phi_s^{-1}\,a_j(s)\,dW_s^j\}\ ;\ t \geq 0$$

où $(\Phi_t\ ;\ t \geq 0)$ est définie par (2). De plus si $m_t = E(X_t)$, $t \geq 0$ on a

$$(3) \qquad \dot{m}_t = A_o(t)\,m_t + a_o(t)\ ;\ t \geq 0\ ;\ m_o = m(0)\ .$$

On peut obtenir (cf.([2])) une autre représentation du processus d'état, plus commode pour le calcul de sa fonction de covariance :

1.2. LEMME - Le processus d'état $(X_t\ ;\ t \geq 0)$ du modèle (1) est donné par

(4) $$X_t = \pi_t \circ V_t + m_t \quad ; \; t \geq 0 \; ,$$

où $(m_t \, ; \, t \geq 0)$ est donné par (3) , $(\pi_t \, ; \, t \geq 0)$ est la matrice fondamentale de (3) et $(V_t \, ; \, t \geq 0)$ est solution de l'équation différentielle stochastique vectorielle :

$$dV_t = \pi_t^{-1} \circ \sum_{j=1}^{d} \{ A_j(t) [\pi_t V_t + m_t] + a_j(t) \} \, dW_t^j \; ; \; t \geq 0 \; ; \; V_o = X(0) - m(0) \; .$$

2. FONCTION DE COVARIANCE ET STATIONNARITE AU SECOND ORDRE.

En utilisant la représentation (4) du processus d'état on démontre [2] le résultat suivant complétant l'énoncé 8.5.5. de [3] :

2.1. PROPOSITION - Si $K(t,s) = E[(X_t - m_t)(X_s - m_s)']$ et $K_t = K(t,t)$, $t,s \geq 0$, $(X_t \, ; \, t \geq 0)$ étant le processus d'état du modèle (1), on a :

$$K(t,s) = \begin{cases} \pi_t \cdot \pi_s^{-1} \, K_s & \text{si} \quad t \geq s \\ K_t \cdot (\pi_s \cdot \pi_t^{-1})' & \text{si} \quad t \leq s \end{cases}$$

où $(K_t \, ; \, t \geq 0)$ est donné par :

(5) $$\dot{K}_t = A_o(t) K_t + K_t A_o'(t) + \sum_{j=1}^{d} A_j(t) K_t A_j'(t) + \sum_{j=1}^{d} [A_j(t) m_t + a_j(t)] [A_j(t) m_t + a_j(t)]' \; ;$$
$$t \geq 0 \; ; \; K_o = K(0) \; .$$

On est alors en mesure d'obtenir [2] une caractérisation de la stationnarité au second ordre du processus d'état :

2.2. PROPOSITION - Le processus d'état $(X_t \, ; \, t \geq 0)$ du modèle (1) est stationnaire au second ordre si et seulement si il existe une matrice constante $\bar{A}_o$ et un vecteur constant $\bar{a}_o$ tels que

$$\bar{A}_o\, m(0) + \bar{a}_o = 0$$

et que les conditions suivantes sont vérifiées pour presque tout $t \geq 0$:

(i) $A_o(t) \equiv \bar{A}_o \quad , \quad a_o(t) \equiv \bar{a}_o$

(ii) $\bar{A}_o . K(0) + K(0) . \bar{A}'_o + \sum_{j=1}^{d} A_j(t) K(0) A'_j(t) + \sum_{j=1}^{d} [A_j(t)m(0)+a_j(t)]\,[A_j(t)m(0)+a_j(t)]' = 0$

Lorsque les conditions sont vérifiées la fonction de covariance de $(X_t \,;\, t \geq 0)$ est donnée par

$$(6) \quad K(t,s) = \begin{cases} e^{\bar{A}_o . (t-s)}\, K(0) & \text{si } t \geq s \\ K(0) . e^{\bar{A}'_o . (s-t)} & \text{si } t \leq s \,. \end{cases}$$

Un lemme (2) permet de compléter ce résultat :

2.3. LEMME - Soit $\bar{A}_o, \ldots, \bar{A}_d$ $d+1$ matrices $n \times n$ telles que la matrice $\bar{A}_o \otimes I_n + I_n \otimes \bar{A}_o + \sum_{j=1}^{d} \bar{A}_j \otimes \bar{A}_j$ soit une matrice stable. Alors pour toute matrice $n \times n$ symétrique définie non-négative M l'équation matricielle

$$(7) \qquad \bar{A}_o . K + K . \bar{A}'_o + \sum_{j=1}^{d} \bar{A}_j K \bar{A}'_j + M = 0$$

admet pour solution une matrice K^* symétrique définie non-négative. De plus si M est définie positive, K^* est définie positive.

L'énoncé suivant est alors un corollaire immédiat de la Proposition 2.2. :

2.4. COROLLAIRE - Soit $\bar{A}_o, \ldots, \bar{A}_d$ $d+1$ matrices $n \times n$ vérifiant l'hypothèse du lemme 2.3. et $a_1, \ldots, a_d$ d vecteurs de $\mathbb{R}^n$. Alors si K(0) est la solution K^* de l'équation matricielle (7) avec

$$M = \sum_{j=1}^{d} [\bar{A}_j m(0) + \bar{a}_j] [\bar{A}_j m(0) + \bar{a}_j]' ,$$

le processus d'état $(X_t ; t \geq 0)$ du modèle

$$dX_t = \bar{A}_o [X_t - m(0)] dt + \sum_{j=1}^{d} [\bar{A}_j X_t + \bar{a}_j] dW_t^j ; t \geq 0 ; X_o = X(0) ,$$

est stationnaire au second ordre, de fonction de covariance donnée par (6).

3. REPRESENTATION LINEAIRE ET FILTRAGE LINEAIRE.

Le processus d'état $(X_t ; t \geq 0)$ du modèle (1) admet ([2]) une représentation linéaire au sens suivant :

3.1. PROPOSITION - Il existe un processus de Wiener au sens large (cf([4]) Ch. 15) W^* dans $\mathbb{R}^n$ tel que le processus d'état $(X_t ; t \geq 0)$ du modèle (1) admette la représentation

$$dX_t = [A_o(t) X_t + a_o(t)] dt + \Sigma^{1/2}(t) dW_t^* ; t \geq 0 ; X_o = X(0)$$

où

$$(8) \quad \Sigma(t) = \sum_{j=1}^{d} [A_j(t) K_t A_j'(t) + (A_j(t) m_t + a_j(t))(A_j(t) m_t + a_j(t))'] ; t \geq 0 ,$$

$(m_t ; t \geq 0)$ et $(K_t ; t \geq 0)$ étant donnés par (3) et (5).

3.2. REMARQUES - Dans le cas particulier où le processus est stationnaire au second ordre on a

$$\Sigma(t) \equiv -\bar{A}_o K(0) - K(0) . \bar{A}_o' ; t \geq 0 .$$

On peut s'assurer de la cohérence des énoncés 2.1. et 3.1. avec les théorèmes 15.1 et 15.2. de ([4]).

Envisageons maintenant le problème suivant de filtrage linéaire d'un système bilinéaire : on cherche à filtrer l'état X_t du système régi par (1), linéairement (au sens de ([4]) 15.2.1) à partir de l'observation sur $[0,t]$ du processus $(Y_t \; ; \; t \geq 0)$ gouverné par l'équation

$$(9) \qquad dY_t = [B_o(t) X_t + b_o(t)]\, dt + B_1(t)\, dZ_t \quad ; \; t \geq 0 \; ; \; Y_o = Y(0) \; ,$$

où Z est un mouvement brownien standard dans $\mathbb{R}^p$ indépendant de W, les applications $B_o = ((B_o^{k,\ell} \; ; k=1,\ldots,p \; ; \ell=1,\ldots,n))$, $b_o = (b_o^1,\ldots,b_o^p)'$ et $B_1 = ((B_1^{k,\ell} \; ; k,\ell=1,\ldots,p))$ vérifient les mêmes hypothèses que A_o, a_o et A_1 respectivement et l'état initial $Y(0)$ est un vecteur aléatoire du second ordre tel que $(X(0), Y(0))$ est indépendant de (W,Z). On suppose de plus que la matrice $B_1(t) \cdot B_1^*(t)$ est non-singulière pour tout $t \geq 0$. Le théorème 15.3 de ([4]) et la proposition 3.1. permettent alors de démontrer ([2]) :

3.3. PROPOSITION - Le filtre linéaire optimal $\hat{X}_t$ de l'état X_t du système régi par (1) à partir de l'observation sur $[0,t]$ du processus gouverné par (9) est défini par les équations :

$$d\hat{X}_t = [A_o(t)\hat{X}_t + a_o(t)]\, dt + \gamma_t \cdot B_o'(t)[B_1(t) B_1'(t)]^{-1} [dY_t - \{B_o(t)\hat{X}_t + b_o(t)\} dt]$$

$$(10) \qquad \dot{\gamma}_t = A_o(t)\gamma_t + \gamma_t A_o'(t) + \Sigma(t) - \gamma_t B_o'(t)[B_1(t) \cdot B_1'(t)]^{-1} B_o(t)\gamma_t \; ; \; t \geq 0 \; ,$$

avec

$$\hat{X}_o = m(0) + \mathrm{Cov}(X(0), Y(0))\, \mathrm{Cov}^+(Y(0),Y(0)) \; \{Y(0) - E(Y(0))\}$$

$$\gamma_o = K(0) - \mathrm{Cov}(X(0), Y(0)) \cdot \mathrm{Cov}^+(Y(0),Y(0))\, \mathrm{Cov}(Y(0),X(0)) \; .$$

$(\Sigma(t) \; ; \; t \geq 0)$ étant donnée par (8).

De plus on a

$$\gamma_t = E\{[X_t - \hat{X}_t]\,[X_t - \hat{X}_t]'\} \; .$$

Dans le cas autonome $(A_j(t) \equiv \bar{A}_j, a_j(t) \equiv \bar{a}_j, B_j(t) \equiv \bar{B}_j, b_j(t) \equiv \bar{b}_j$, $j = 0,\dots,d)$ on peut, en utilisant le théorème 16.2 de (4), démontrer le résultat suivant concernant le comportement asymptotique de la matrice γ_t :

3.4. PROPOSITION - Supposons le système (1) - (9) autonome, le couple $(\bar{A}_o, \bar{B}_o)$ observable et le processus d'état du modèle (1) stationnaire au second ordre. Alors si le couple $(\bar{A}_o, [-\bar{A}_o . K(0) - K(0) . \bar{A}'_o]^{1/2})$ est contrôlable, $(\gamma_t ; t \geq 0)$ étant donné par (10), la limite $\lim_{t \to +\infty} \gamma_t$ existe. Cette limite ne dépend pas de γ_o et est l'unique (dans la classe des matrices symétriques définies positives) solution γ de l'équation matricielle :

$$\bar{A}_o \gamma + \gamma \bar{A}_o - \bar{A}_o . K(0) - K(0) . \bar{A}'_o - \gamma \bar{B}'_o [\bar{B}_1 . \bar{B}'_1]^{-1} \bar{B}_o \gamma = 0 .$$

(*) Travail effectué pendant un séjour de M. Musiela au Laboratoire I.M.A.G. Grenoble.

(1) I. VRKOČ, Commentationes Mathematicae Universitatis Carolinae, 19, 1, 1978, p.141-146.

(2) A. LE BRETON et M. MUSIELA, à paraître.

(3) L. ARNOLD, Stochastic Differential equations (Theory and Applications), J. Wiley, 1974.

(4) R.S. LIPTSER et A.N. SHIRYAYEV, Statistics of Random Processes II (Applications), Springer-Verlag 1978.

A. LE BRETON
Laboratoire IMAG
B.P. 53 X GRENOBLE-CEDEX
FRANCE

M. MUSIELA
IMPAN
Kopernika 18
51-617 WROCLAW
POLOGNE

Estimation des paramètres pour les diffusions gaussiennes homogènes hypoelliptiques [1]. Note de Alain Le Breton et Marek Musiela.

Le problème d'estimation des paramètres d'une diffusion gaussienne homogène a été étudié dans le cas où le générateur différentiel associé est elliptique (cf. par exemple [2], [5]). Nous montrons ici qu'il est possible d'étendre la démarche proposée dans [5] au cas de l'hypothèse d'hypoellipticité.

Parameter estimation for hypoelliptic homogeneous gaussian diffusions [1]. Note by Alain Le Breton and Marek Musiela.

The problem of parameter estimation for an homogeneous gaussian diffusion has been studied in the case when the corresponding differential generator is elliptic (cf. for example [2], [5]). Here we show that the method used in [5] may be extended to the case when hypoellipticity only is assumed.

1. INTRODUCTION. - Considérons une diffusion gaussienne homogène $X = (X_t ; t \geq 0)$ associée au générateur différentiel

$$L = \frac{1}{2} \sum_{i,j=1}^{n} a_{ij} \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i,j=1}^{n} b_{ij} x_j \frac{\partial}{\partial x_i}$$

où $A = ((a_{ij}))$ et $B = ((b_{ij}))$ sont deux matrices $n \times n$ constantes. Une telle diffusion X vérifie l'équation différentielle stochastique

(1) $$dX_t = BX_t \, dt + A^{1/2} \, dW_t \quad ; \quad t \geq 0 ,$$

où $A = A^{1/2} . (A^{1/2})'$ (C' désignant la transposée de la matrice C) et $(W_t ; t \geq 0)$ est un mouvement brownien standard r-dimensionnel si $A^{1/2}$ est une matrice rectangulaire $n \times r$.

Nous nous intéressons au problème d'estimation des paramètres inconnus A et B du générateur différentiel au vu de l'observation sur un intervalle de temps [0,T] d'une trajectoire de la diffusion issue de zéro au temps zéro i.e. de la solution de (1) avec condition initiale $X_o \equiv 0$. Ce problème a été abordé dans [2] et [5] sous l'hypothèse d'ellipticité i.e. det $A \neq 0$. Nous nous proposons ici de montrer que la démarche utilisée dans [5] peut être étendue au cas de l'hypothèse d'hypoellipticité. Nous supposons désormais que cette hypothèse est satisfaite ou, de façon équivalente (cf. par exemple [3], [4]) que la paire $(B, A^{1/2})$ est contrôlable i.e. rang $[A^{1/2}, BA^{1/2}, \dots, B^{n-1} A^{1/2}] = n$.

Une estimation de la matrice A peut être obtenue à l'aide du processus de la variation quadratique [X] du processus X. De façon précise A peut être calculée avec probabilité un sur tout intervalle de temps fini [0,T] par

$$A = \frac{1}{T} [X]_T .$$

Par suite nous pouvons envisager le problème de l'estimation de la matrice B lorsque A est supposée connue sans que cela soit restrictif, A ayant éventuellement été préalablement calculée.

Nous montrons que la méthode du maximum de vraisemblance peut être utilisée, qu'elle fournit un estimateur $\hat{B}_T$ tel que la paire $(\hat{B}_T, A^{1/2})$ est contrôlable et que de plus, si B est stable, cet estimateur est fortement convergent et asymptotiquement gaussien.

2. UN RESULTAT PRELIMINAIRE. - Enonçons d'abord un lemme qui est à la base de notre approche :

2.1. LEMME. - La matrice $\int_0^T X_t X'_t \, dt$ est presque sûrement strictement positive définie.

PREUVE. - Il s'agit de montrer que l'ensemble

$$\Omega = \bigcup_{h \in \mathbb{R}^n \setminus \{0\}} \left[h' \left(\int_0^T X_t X_t' \, dt \right) h = 0 \right]$$

est négligeable. Compte tenu de ce que X est presque sûrement à trajectoires continues, si Ω_1 est défini par

$$\Omega_1 = \bigcup_{h \in \mathbb{R}^n \setminus \{0\}} \ \bigcap_{s \in [0,T]} \left[\langle X_s , h \rangle = 0 \right] .$$

l'ensemble $\Omega \setminus \Omega_1$ est négligeable ; il suffit donc de montrer que Ω_1 est lui-même négligeable. Soit alors une suite $0 < t_1 < \ldots < t_n \leq T$; on a

$$\Omega_1 \subset \Omega_1^* = [\det [X_{t_1}, \ldots , X_{t_n}] = 0] .$$

Ainsi, comme le sous-ensemble de $(\mathbb{R}^n)^n$ $\{(V_1, \ldots , V_n) : \det[V_1, \ldots , V_n] = 0\}$ est de mesure de Lebesgue nulle, le résultat du lemme sera acquis si nous montrons que le vecteur aléatoire gaussien $(X'_{t_1}, \ldots , X'_{t_n})'$ est de loi non dégénérée. Or nous pouvons écrire, E désignant la matrice identité $n \times n$,

$$\begin{pmatrix} X_{t_1} \\ \vdots \\ X_{t_n} \end{pmatrix} = \begin{pmatrix} e^{Bt_1} & & 0 \\ & \ddots & \\ 0 & & e^{Bt_n} \end{pmatrix} \begin{pmatrix} E & & 0 \\ \vdots & \ddots & \\ E & \cdots & E \end{pmatrix} \begin{pmatrix} \int_0^{t_1} e^{-Bs} A^{1/2} \, dW_s \\ \vdots \\ \int_{t_{n-1}}^{t_n} e^{-Bs} A^{1/2} \, dW_s \end{pmatrix}$$

où les vecteurs aléatoires $\int_{t_i}^{t_{i+1}} e^{-Bs} A^{1/2} \, dW_s$, $i = 0, \ldots , n-1$, $(t_o = 0)$, sont indépendants de matrices de covariances respectives $\int_{t_i}^{t_{i+1}} e^{-Bs} A e^{-B's} ds$, $i = 0, \ldots , n-1$, lesquelles sont strictement définies positives car la paire $(B, A^{1/2})$ est contrôlable (cf. par exemple [3], [6]).

2.2. REMARQUE. - Comme d'après (1) on a :

$$\int_0^T dX_t\, X'_t = B \int_0^T X_t\, X'_t\, dt + A^{1/2} \int_0^T dW_t\, X'_t$$

on a aussi, désignant par C^+ la pseudo-inverse de la matrice C,

$$\int_0^T \{d[E-A^{1/2}.(A^{1/2})^+]X_t\}X'_t = [E-A^{1/2}.(A^{1/2})^+]\, B \int_0^T X_t\, X'_t\, dt$$

car CC^+ n'est autre que la matrice de projection orthogonale sur le sous-espace engendré par les colonnes de C. Utilisant l'égalité (cf. [6]) $A^{1/2}.(A^{1/2})^+ = A.A^+$ et le lemme 2.1. précédant on obtient :

$$(E-AA^+)B = \{\int_0^T [d(E-AA^+)X_t]X'_t\}.\{\int_0^T X_t\, X'_t\, dt\}^{-1} . \tag{2}$$

Notons que l'intégrale stochastique du second membre de l'égalité (2) est en fait une intégrale ordinaire puisque le processus $((E-AA^+)X_t\ ;\ t \geq 0)$ est à variation localement finie. En définitive, de même que A, la matrice $(E-AA^+)B$ peut être calculée avec probabilité un sur un intervalle de temps fini ; cela nous permet donc de supposer que cette matrice est elle aussi connue.

3. ESTIMATION PAR LA METHODE DU MAXIMUM DE VRAISEMBLANCE.

Considérons un processus $Y = (Y_t\ ;\ t \geq 0)$ solution de l'équation différentielle stochastique

$$dY_t = (E-AA^+)BY_t\, dt + A^{1/2}\, dW^*_t\ ;\ t \geq 0\ ;\ Y_o \equiv 0\ ,$$

où $(W^*_t\ ;\ t \geq 0)$ est un mouvement brownien standard r-dimensionnel. L'équation $A^{1/2}\alpha(x) = A^{1/2}.(A^{1/2})^+x$ admettant pour tout vecteur x de $\mathbb{R}^n$ la solution $\alpha(x) = (A^{1/2})^+x$, la loi du processus $(X_t\ ;\ t \in [0,T]$ est absolument continue par rapport à celle du processus $(Y_t\ ;\ t \in [0,T])$ avec pour densité (cf. [6])

$$\exp \{ \int_0^T (AA^+ BX_t)' A^+ dX_t - \frac{1}{2} \int_0^T (AA^+ BX_t)' A^+(2B-AA^+)X_t \, dt \} .$$

La loi de $(Y_t ; t \in [0,T])$ ne dépendant que des matrices connues A et $(E-AA^+)B$ est une mesure dominante pour la structure statistique associée au problème d'estimation posé ; la fonction de Log-vraisemblance peut s'écrire sous la forme

$$\mathrm{Tr} \{ A^+ [\int_0^T dX_t X'_t B' - \frac{1}{2} \int_0^T BX_t X'_t B' dt] \} .$$

Le résultat suivant est alors une conséquence immédiate du Lemme 2.1.

3.1. PROPOSITION. - Un estimateur de maximum de vraisemblance de la matrice B est donné par

$$\hat{B}_T = [\int_0^T dX_t X'_t] . [\int_0^T X_t X'_t \, dt]^{-1} . \tag{3}$$

3.2. REMARQUE. - D'après (1) on peut écrire

$$\hat{B}_T = B + A^{1/2} [\int_0^T dW_t X'_t] [\int_0^T X_t X'_t \, dt]^{-1} . \tag{4}$$

Cela assure, en vertu de la contrôlabilité de la paire $(B, A^{1/2})$, que la paire $(\hat{B}_T, A^{1/2})$ est elle-même contrôlable ; ainsi l'estimateur $(\hat{B}_T, (\frac{1}{T} [X]_T)^{1/2})$ est à valeurs dans l'espace des paramètres. Notons de plus que le calcul $(E-AA^+) \hat{B}_T$ redonne la matrice $(E-AA^+)B$ définie par (2).

4. PROPRIETES ASYMPTOTIQUES DE L'ESTIMATEUR. - Nous faisons maintenant l'hypothèse supplémentaire que la matrice B est stable. Nous sommes alors assurés (cf. [9]) de l'existence et de l'unicité d'une mesure de probabilité invariante pour la diffusion gaussienne. Cette mesure

est gaussienne centrée de matrice de covariances non dégénérée

(5) $$Q_{A,B} = \int_0^{+\infty} e^{Bs} \; A \, e^{B's} \; ds$$

i.e. solution de l'équation de Lyapunov

$$BQ + QB' = -A \; .$$

Utilisant l'égalité (4), un résultat (cf. [7]) concernant l'ergodicité des diffusions et un résultat (cf. [8]) sur la normalité asymptotique d'intégrales stochastiques, il est facile de montrer :

4.1. PROPOSITION. - L'estimateur $\hat{B}_T$ défini par (3) est convergent presque sûrement. De plus le vecteur aléatoire $T^{1/2}$. vec$(\hat{B}_T - B)$ est asymptotiquement gaussien d'espérance nulle et de matrice de covariance $Q_{A,B}^{-1} \otimes A$ où $Q_{A,B}$ est donnée par (5).

4.2. REMARQUE. - Notons que si on suppose que l'espace des paramètres est l'ensemble des couples (B,A) tels que d'une part la paire $(B, A^{1/2})$ est contrôlable et que d'autre part B est une matrice stable, l'estimateur $\hat{B}_T$ défini par (3) n'est pas un estimateur de maximum de vraisemblance ; en fait dans ces conditions un tel estimateur n'existe pas. La matrice $\hat{B}_T$ n'est pas en général stable mais la probabilité qu'elle le soit tend vers un lorsque T augmente indéfiniment. Dans le cas où A est non dégénérée, il est possible de modifier $\hat{B}_T$ de manière à obtenir un estimateur de B qui soit stable de la manière suivante : on a

$$\int_0^T dX_t \, X'_t + \int_0^T X_t \, dX'_t = X_T \, . \, X'_T - [X]_T \; ,$$

de sorte qu'on a aussi

$$\left(\int_0^T dX_t \, X'_t - \frac{1}{2} X_T \, . \, X'_T\right) \left(\int_0^T X_t \, X'_t \, dt\right)^{-1} \frac{1}{T} \int_0^T X_t \, X'_t \, dt \; +$$

$$+ \frac{1}{T} \int_0^T X_t \; X'_t \, dt \left(\int_0^T X_t X'_t \, dt\right)^{-1} \left(\int_0^T X_t \, dX'_t - \frac{1}{2} X_T \, X'_T\right) = -A \; .$$

Cela montre que l'équation de Lyapunov

$$\tilde{B}_T Q + Q \tilde{B}_T' = -A ,$$

où

$$\tilde{B}_T = \left(\int_0^T dX_t X_t' - \frac{1}{2} X_T . X_T' \right) \left(\int_0^T X_t X_t' \right)^{-1}$$

admet pour solution la matrice strictement définie positive $\frac{1}{T} \int_0^T X_t X_t' \, dt$.

Alors le théorème de Lyapunov assure que $\tilde{B}_T$ est une matrice stable. Il est de plus clair que l'estimateur $\tilde{B}_T$ a les mêmes propriétés asymptotiques que $\hat{B}_T$.

[1] Travail effectué pendant un séjour de M. Musiela au Laboratoire I.M.A.G.

[2] M. ARATO, Studia Scientiarum Mathematicarum Hungaria, 5, 1970, p.11-15.

[3] M. CHALEYAT-MAUREL et L. ELIE, S.M.F., Astéristique, 84-85, 1981, p.255-279.

[4] K. ICHIHARA and H. KUNITA, Zeitschrift f.W., 30, 1974, p.235-254.

[5] A. LE BRETON, in Trans. Seventh Prague Conf., Vol.A, 1977, p.353-366.

[6] R.S. LIPTSER and A.N. SHIRYAYEV, Statistics of Random Processes, Springer-Verlag, 1978.

[7] G. MARUYAMA and H. TANAKA, Memoirs Fac. Sc., Kyusyu Univ., Ser.A, Vol. 19, n°2, 1957, p.118-142.

[8] A.F. TARASKIN, in Theory of Stochastic Processes, Vol.1, J. Wiley, 1974, p.136-151.

[9] M. ZAKAI and J. SNYDERS, J. Differential Equations, 8, 1970, p.27-33.

A.L.B. : Laboratoire I.M.A.G., B.P. 53 X, 38041 Grenoble-Cédex ;

M.M. : I.M.P.A.N., Kopernika 18, 51-617 Wroclaw, Pologne.

A STUDY OF AN ONE-DIMENSIONAL BILINEAR * DIFFERENTIAL MODEL FOR STOCHASTIC PROCESSES

BY

A. LE BRETON (GRENOBLE) AND M. MUSIELA (WROCLAW)

ABSTRACT. - This paper is concerned with a study of an one-dimensional bilinear differential model for stochastic processes in continuous time. We provide conditions for second-order and strict-sense stationarities of the state process. We obtain a linear representation of the state process, we derive the optimal linear filter and we investigate its asymptotic behaviour. We consider the problem of parameter estimation for the autonomous version of the model. By use of the quadratic variation of the process we compute the diffusion coefficient parameters. In the reduced model, under the additional assumption that the parameters of the diffusion coefficient are known, we use the maximum likelihood method and the method of moments in order to estimate the drift coefficient parameters. We prove consistency and asymptotic normality of the estimates.

1. INTRODUCTION. Bilinear deterministic models for dynamical systems in discrete time (see e.g. A. ISIDORI [6]) and in continuous time (see e.g. C. BRUNI and al [4]) have been intensively studied. Analogous stochastic models have also been considered in time series analysis (see e.g. C.W. GRANGER and A. ANDERSEN [5]) and in the theory of stochastic differential equations (see e.g. L. ARNOLD [2]).

---

* Work done during a stay of M. Musiela in Laboratoire I.M.A.G. Grenoble.

In this paper we try to develop a probabilistic and statistical study of an one-dimensional stochastic model which is given by a bilinear differential equation (in the Itô sense) of the following form

$$(1.1) \qquad dX_t = [A(t)\,X_t + a(t)]\,dt + [B(t)\,X_t + b(t)]\,dW_t \;;\; t \geq 0 \;;\; X_o = X(0) \;,$$

where

- $W = (W_t \,;\, t \geq 0)$ is a standard brownian motion in $\mathbb{R}$ defined on some basic probability space $(\Omega, \mathcal{A}, (\mathcal{A}_t)_{t\geq 0}, P)$

- the deterministic functions A, a, B and b are measurable and satisfy for every $T > 0$ the following conditions

$$\int_0^T |A(t)|\,dt < \infty \;;\; \int_0^T |a(t)|\,dt < \infty \;;\; \int_0^T |B(t)|^2 dt < \infty \;;\; \int_0^T |b(t)|^2 dt < \infty$$

- the initial state $X(0)$ is a random variable, defined on $(\Omega, \mathcal{A}, P)$, which is independent of W and admits an expectation $m(0)$ and a variance $K(0)$.

Section 2 is devoted to the analysis of the state process $(X_t; t\geq 0)$ of model (1.1). We compute its mean and covariance functions ; we obtain a linear representation of $(X_t \,;\, t\geq 0)$ with respect to a wide-sense Wiener process (cf. R.S. LIPTSER and A.N. SHIRYAYEV [11]) ; we give also conditions for second-order and strict-sense stationarities in model (1.1) (cf. K. ITO and M. NISIO [7]).

In section 3, by use of the linear representation and results in R.S. LIPTSER and A.N. SHIRYAYEV [11] , we provide the equations for the optimal linear filter of $(X_t \,;\, t\geq 0)$; moreover, under the assumption that the state process is second-order stationary, we study the behaviour and the stability of the filter error.

In section 4 we deal with the problem of parameter estimation in the autonomous version of the model when a strict sense stationary solution exists, the state process being observed in continuous time. We use a two steps procedure (cf. M. ARATO [1], B.M. BROWN and J.I. HEWITT [3], A. LE BRETON [8]) : first the diffusion coefficient parameters are estimated by use of the quadratic variation of the observed process and then the maximum likelihood method and the method of moments provide consistent and asymptotically normal estimates of the drift coefficient parameters (cf. T.S. LEE and F. KOZIN [10]).

2. PROPERTIES OF THE STATE PROCESS. In this section we summarize the properties of the state process generated by model (1.1) which will be used in the next parts. For more details one can see [9] § 2, 3A and 4A.

A. Moments and linear representation.

The existence and uniqueness of the solution process $(X_t ; t \geq 0)$ of equation (1.1) is ensured by general results on stochastic differential equations (see e.g. [11] Ch.4) and the conditions listed in section 1. The Itô formula provides (see e.g. [2] Th.8.4.2. or [9] Th. 2.1.) that $(X_t ; t \geq 0)$ is given by

$$X_t = \Phi_t \{ X(0) + \int_0^t \Phi_s^{-1} [a(s) - B(s) b(s)] ds + \int_0^t \Phi_s^{-1} b(s) dW_s \} ; t \geq 0$$

where

$$\Phi_t = \exp\{ \int_0^t [A(s) - \frac{1}{2} B^2(s)] ds + \int_0^t B(s) dW_s \} \ ; \ t \geq 0 \ .$$

Denote by $(\psi_t ; t \geq 0)$ the mean function of $(\Phi_t ; t \geq 0)$ i.e.

$$\psi_t = \exp \{ \int_0^t A(s) \ ds \} \ ; \ t \geq 0 \ .$$

The second-order structure of $(X_t ; t \geq 0)$ is described in the following

Lemma 2.1. - The mean function $(m_t = EX_t \; ; \; t \geq 0)$ is given by

$$(2.1) \qquad m_t = \psi_t \, [m(0) + \int_0^t \psi_s^{-1} \, a(s)\, ds] \quad ; \quad t \geq 0 \; .$$

The covariance function $(K(t,s) = E(X_t - m_t)(X_s - m_s) \; ; \; t \geq 0 \; , \; s \geq 0)$ is equal to

$$(2.2) \qquad K(t,s) = \psi_t \, \psi_s^{-1} \, K_s \quad ; \quad t \geq s$$

where the variance function $(K_t = K(t,t) \; ; \; t \geq 0)$ is expressed by :

$$(2.3) \qquad \begin{aligned} K_t &= \exp\{ \int_0^t [\, 2A(u) + B^2(u)]\, du \} \, \{K(0) + \\ &+ \int_0^t \exp\{ - \int_0^s [\, 2A(u) + B^2(u)]\, du\} \, [B(s) m_s + b(s)]^2 ds\} ; t \geq 0 \; . \end{aligned}$$

Proof - Formula (2.1) follows immediately from (1.1). The derivation of equations (2.2) and (2.3) can be based on the representation

$$X_t - m_t = \psi_t \, V_t \quad ; \quad t \geq 0$$

where

$$dV_t = \{B(t) \, V_t + \psi_t^{-1} [B(t) \, m_t + b(t)] \} \, dW_t \quad ; \quad t \geq 0 \; ; \; V_o = X(0) - m(0) \; .$$

Simple computations provide

$$(2.4) \quad E(V_t \, V_s) = K(0) + \int_0^s \{ B^2(u) \, E(V_u^2) + \psi_u^{-2} \, [B(u) \, m_u + b(u)]^2 \} du \; ; t \geq s \geq 0 \; .$$

Setting $t = s$, multiplying by ψ_s^2 and noting that $\psi_u^2 \; E(V_u^2) = K_u$, we obtain

$$K_s = \psi_s^2 \left\{ K(0) + \int_0^s \psi_u^{-2} \, [B^2(u) \, K_u + \{ B(u) \, m_u + b(u) \}^2] \, du \right\} \; ; \; s \geq 0 \; .$$

This leads to (2.3) and moreover, together with (2.4), shows that $\psi_s^{-2}\; E(V_t V_s) = K_s$; $t \geq s$. Finally (2.2) holds since $K(t,s) = \psi_t \psi_s E(V_t V_s)$. ■

Following up the ideas presented in [11] Ch.15 we provide now a linear representation of the state process (for a detailed proof see [9] § 3A.).

Lemma 2.2. - There exists a wide-sense Wiener process $W^* = (W_t^* \; ; \; t \geq 0)$ in $\mathbb{R}$ which is uncorrelated with $X(0)$ and such that the state process $(X_t \, ; t \geq 0)$ admits the representation

(2.5) $$dX_t = [\, A(t) X_t + a(t)]\, dt + [B^2(t) K_t + \{ B(t)\, m_t + b(t) \}^2]^{1/2}\, dW_t^* \; ; \; t \geq 0 \; ; \; X_o = X(0) \, .$$

Proof - It is easy to verify that Lemma 2.1. ensures that the assumptions of Th.15.2 in [11] are satisfied for the process $(X_t \; ; t \geq 0)$; then representation (2.5) holds. Moreover, from the proof of Th.15.2 [11], one can see that W^* is uncorrelated with $X(0)$. ■

B. Second-order and strict-sense stationarities.

A characterization of second-order stationarity of the state process is given in

Lemma 2.3. - The process $(X_t \; ; t \geq 0)$ generated by model (1.1) is second-order stationary if and only if either it is deterministic with $X(0) = m(0)$ a.s. and $A(t)\; m(0) + a(t) = B(t)\; m(0) + b(t) = 0$ almost everywhere or there exist two constants A and a such that

(2.6) $$A m(0) + a = 0$$

and the following conditions hold for almost every $t \geq 0$

(2.7) $$A(t) = A \; ; \quad a(t) = a$$

(2.8) $$[2A+B^2(t)]\,K(0) + [B(t)\,m(0)+b(t)]^2 = 0 .$$

Moreover, if conditions (2.6) - (2.8) are satisfied, then the covariance function of $(X_t\ ;\ t\geq 0)$ is given by

(2.9) $$K(t,s) = e^{A|t-s|}\,K(0)\ ;\quad t\geq 0\ ,\ s\geq 0 .$$

Proof - If $K(0) = 0$ (i.e. $X(0) = m(0)$ a.s.) then $(X_t\ ;t\geq 0)$ is second-order stationary if and only if $m_t = m(0)$ and $K_t = 0\ ;\ t\geq 0$. In view of (2.1) and (2.3) this is true if and only if

$$\int_0^t [A(s)\,m(0)+a(s)]\,ds = \int_0^t \psi_s^{-2}\,[B(s)\,m(0)+b(s)]^2 ds = 0\ ;\ t\geq 0$$

or equivalently

$$A(t)\ m(0) + a(t) = B(t)\ m(0) + b(t) = 0$$

almost everywhere.

Let us now consider the case when $K(0) > 0$. It is clear from Lemma 2.1 that conditions (2.6) - (2.8) ensure second-order stationarity and that (2.9) holds. Conversely, if $(X_t\ ;\ t\geq 0)$ is second-order stationary then $K_t = K(0)$ and $K(t,s) = R(t-s)\ ;\ t\geq 0\ ,\ s\geq 0$. Then (2.2) implies that for $t\geq s$

$$\int_s^t A(u)\,du = \operatorname{Log}\frac{R(t-s)}{K(0)}$$

from which we obtain that there exists a constant A such that $A(t) = A$ almost everywhere. Consequently $\psi_t = e^{At}\ ;\ t\geq 0$ and, since $m_t = m(0)\ ;$ $t\geq 0$, formula (2.1) leads to $a(t) = -A\,m(0)$ almost everywhere. Moreover, since $K_t = K(0)\ ;\ t\geq 0$, formula (2.3) implies that (2.8) holds almost everywhere. ∎

Examples 2.4. - (a) If $A(t) \equiv A < 0$, $a(t) \equiv a$, $B(t) \equiv (-2A)^{1/2}$,$b(t) \equiv \frac{B}{A} a$; $t \geq 0$, then, for $m(0) = -\frac{a}{A}$ and every $K(0)$, the process is second-order stationary.

(b) If $A(t) \equiv A$, $a(t) \equiv a$, $B(t) \equiv B$, $b(t) \equiv b$; $t \geq 0$ with $2A + B^2 < 0$, then, for $m(0) = -\frac{a}{A}$ and $K(0) = \frac{(Ab - Ba)^2}{A^2 |2A + B^2|}$, the process is second-order stationary.

(c) If $A(t) \equiv A < 0$, $a(t) \equiv 0$, $B(t) \equiv (-2A)^{1/2} \operatorname{Sin} t$, $b(t) = (-2A)^{1/2} \operatorname{Cos} t$; $t \geq 0$, then, for $m(0) = 0$ and $K(0) = 1$, the process is second-order stationary.

Remark 2.5. - If the state process generated by (1.1) is second-order stationary with $K(0) > 0$, then, from Lemma 2.2 and 2.3 we obtain that there exists a wide-sense Wiener process W^* and a constant $A \leq 0$ such that

$$dX_t = A[X_t - m(0)]\,dt + (-2AK(0))^{1/2}\, dW_t^* \;;\; t \geq 0 \;,\; X_o = X(0) \;.$$

Now we consider the autonomous version of model (1.1)

$$(2.10) \qquad dX_t = (AX_t + a)\,dt + (BX_t + b)\, dW_t \;;\; t \geq 0 \;;\; X_o = X(0) \;.$$

We study the problem of existence of a strict-sense stationary solution with first two moments. First, let us eliminate the deterministic case. In view of Lemma 2.3 we know that the solution process of (2.10) is strictly stationary with $K(t) \equiv 0$ if and only if $A\,m(0) + a = B\,m(0) + b = 0$ and $X(0) = m(0)$ a.s. . Now let us work with $K(0) > 0$ and, taking into account Lemma 2.3, assume that $2A + B^2 \leq 0$. In fact, when $B = 0$, in order to avoid the trivial case

(C0) $A = a = B = b = 0$ i.e. $X_t \equiv X(0)$; $t \geq 0$,

we shall assume

(C1) $B = 0$, $A < 0$, $b \neq 0$.

Moreover, when $B \neq 0$, since the deterministic case is eliminated, we have

(C2) $B \neq 0$, $2A + B^2 < 0$, $Ab \neq aB$.

We are able to prove the following

Lemma 2.6. - In each case (C1) or (C2) there exists a unique invariant probability distribution $\mu_{a,A}^{b,B}$ for the Markov process generated by (2.10). In case (C1) $\mu_{a,A}^{b,B}$ is the gaussian distribution $N\left(-\frac{a}{A}, \frac{b^2}{2|A|}\right)$ and in case (C2) it is the distribution of a random variable of the form $\left\{\text{sign}\left(\frac{b}{B} - \frac{a}{A}\right) . U^{-1} - \frac{b}{B}\right\}$ where U has the gamma distribution $\Gamma\left(1 - \frac{2A}{B^2}, \frac{B^2}{2|A| \, |\frac{b}{B} - \frac{a}{A}|}\right)$. Moreover, if $X(0)$ is distributed along $\mu_{a,A}^{b,B}$ then the state process $(X_t ; t \geq 0)$ is a second-order (stationary) process with

$$(2.11) \qquad m_t \equiv -\frac{a}{A} \quad ; \; K(t,s) = \frac{(Ab - aB)^2}{A^2 |2A + B^2|} . e^{A|t-s|} \; ; \; s \geq 0 \; , \; t \geq 0 \; .$$

Proof - In case (C1) one can write

$$d(X_t + \frac{a}{A}) = A(X_t + \frac{a}{A})\, dt + b \; dW_t \; ; \; t \geq 0 \; ; \; X_o = X(0) \; .$$

Since, as it is well known, an Ornstein-Uhlenbeck process admits a unique invariant probability measure which moreover is gaussian, the result follows.

In case (C2) one can write

$$X_t = Z_t - \frac{b}{B} \quad ; \quad t \geq 0$$

where

$$dZ_t = (A Z_t + \tilde{a})\,dt + B Z_t\, dW_t \;;\; t \geq 0 \;;\; Z_o = X(0) + \frac{b}{B}$$

and

$$\tilde{a} = a - \frac{bA}{B} \neq 0 \, .$$

Let us assume that $\tilde{a}$ is positive. From

(2.12) $$Z_t = \Phi_t \{ Z_o + \tilde{a} \int_0^t \Phi_s^{-1} ds \} \;;\quad t \geq 0$$

it is clear that if a stationary probability distribution for $(Z_t \,;\, t \geq 0)$ exists, it is necessarily concentrated on $]0,+\infty[$. It is easy to show (by use for instance of results of [7] § 13) that $]0,+\infty[$ is a non-singular interval of positive recurrent type for the Markov process generated by (2.12) and therefore that a unique stationary distribution confined on $]0,+\infty[$ for this process exists. Moreover (see e.g. [13] p.274 or [16] example F) this distribution admits a density f which satisfies the Pearson equation

$$\frac{\dot{f}(z)}{f(z)} = \frac{2}{B^2} \cdot \frac{(A - B^2)\, z + \tilde{a}}{z^2}$$

and then is of the form

$$K \cdot z^{2(AB^{-2}-1)} \cdot \exp\{ - \frac{2\tilde{a}}{B^2} \cdot \frac{1}{z} \} \;;\quad z > 0 \, .$$

It follows that the stationary distribution for $(Z_t \,;\, t \geq 0)$ is that of a random variable U^{-1} where U has the $\Gamma\left(1 - \frac{2A}{B^2} \, , \, \frac{B^2}{2\tilde{a}}\right)$ distribution. Moreover, simple computations show that, since $2A + B^2 < 0$,

$$E(U^{-1}) = \frac{b}{B} - \frac{a}{A} \quad ; \quad \mathrm{var}(U^{-1}) = \frac{(Ab - aB)^2}{A^2 \, |2A + B^2|} \, .$$

Finally, coming back to $(X_t \, ; t \geq 0)$, one easily obtains the announced result. Similar arguments lead to the result when $\tilde{a} < 0$. ■

3. OPTIMAL LINEAR FILTERING OF THE STATE PROCESS. In this section we consider the linear filtering problem (in the sense of [11] Ch.15) for the state X_t of the process generated by (1.1) by use of observations on $[0,t]$ of the process $(Y_t \, ; t \geq 0)$ which is given by

$$dY_t = [A_1(t) \, X_t + a_1(t)] \, dt + B_1(t) \, d\tilde{W}_t \; ; \; t \geq 0 \; ; \; Y_o = Y(0) \; .$$

Here $\tilde{W} = (\tilde{W}_t \, ; t \geq 0)$ is a standard brownian motion in $\mathbb{R}$ independent of W , the functions A_1 , a_1 , B_1 satisfy analogous conditions as A, a, B respectively and the initial state $Y(0)$ is a square integrable random variable such that $(X(0), Y(0))$ is independent of $(W, \tilde{W})$. We assume also that $B_1(t) \neq 0$ for all $t > 0$.

A. Equations for the optimal linear filter.

We have the following

<u>Theorem</u> 3.1. - The optimal linear filter $\hat{X}_t$ of the state X_t from the observation $(Y_s \, ; 0 \leq s \leq t)$ is given by

$$d\hat{X}_t = [A(t) \, \hat{X}_t + a(t)] dt + \gamma_t \, \frac{A_1(t)}{B_1^2(t)} \, [\, dY_t - \{A_1(t) \, \hat{X}_t + a_1(t)\} dt \,] \; ; \; t > 0 \; ,$$

$$\text{(3.1)} \qquad \dot{\gamma}_t = 2A(t) \, \gamma_t + \Sigma_t^2 - \{\frac{A_1(t)}{B_1(t)}\}^2 \, \gamma_t^2 \; ; \; t > 0$$

with

$$\hat{X}_o = m(0) + Cov(X(0) \, ; Y(0)) \, . \, Cov^+(Y(0) \, ; Y(0)) \, \{Y(0) - E(Y(0))\} \; ,$$

$$\gamma_o = K(0) - Cov^2(X(0) \, ; Y(0)) \, Cov^+(Y(0) \, ; Y(0))$$

and

$$\Sigma_t^2 = B^2(t) \, K_t + [\, B(t) \, m_t + b(t) \,]^2 \; ; \; t \geq 0$$

while

$$\gamma_t = E\{(X_t - \hat{X}_t)^2\} \;;\quad t \geq 0 .$$

Proof - Taking into account Lemma 2.2, the result follows directly by use of Theorem 15.3 in [11]. ∎

B. Behaviour of the filter error.

Let $(\gamma_t^x \,;\, t \geq 0)$ denote the solution of (3.1) when $\gamma_0 = x$ i.e. $\gamma_t^x = E[(X_t - \hat{X}_t)^2]$ when

$$K(0) - \mathrm{Cov}^2(X(0)\,;\,Y(0))\,\mathrm{Cov}^+(Y(0)\,;\,Y(0)) = x \geq 0 .$$

Then the following holds.

Theorem 3.2. - Assume that the state process $(X_t \,;\, t \geq 0)$ is non deterministic second-order stationary and that the equation for the observation is autonomous with $A_1 \neq 0$. Then we have

(i) under (C0)

$$\gamma_t^x = \begin{cases} \left[\dfrac{1}{x} + \left(\dfrac{A_1}{B_1}\right)^2 t\right]^{-1} & ;\quad t \geq 0 \qquad \text{if} \quad x > 0 \\ 0 & ;\quad t \geq 0 \qquad \text{if} \quad x = 0 \end{cases}$$

and, for every $x \geq 0$,

$$\lim_{t \nearrow \infty} \searrow \gamma_t^x = 0 .$$

(ii) under (C1) or (C2)

$$\gamma_t^x = \frac{\gamma^*(x-\gamma_*) - \gamma_*(x-\gamma^*)\, e^{-2\lambda_* t}}{(x-\gamma_*) - (x-\gamma^*)\, e^{-2\lambda_* t}} \;;\quad t \geq 0 \;,\; x \geq 0$$

where

$$\gamma^{*} = \frac{B_1^2}{A_1^2}\,(A+\lambda_*) \quad ; \quad \gamma_{*} = \frac{B_1^2}{A_1^2}\,(A-\lambda_*)$$

and

$$\lambda_* = \left[A^2 + 2\,|A|\,K(0)\,\frac{A_1^2}{B_1^2}\right]^{1/2} .$$

Moreover, for every $x \le \gamma^*$,

$$\lim_{t \nearrow \infty} \nearrow \gamma_t^x = \gamma^*$$

and, for every $x \ge \gamma^*$,

$$\lim_{t \nearrow \infty} \searrow \gamma_t^x = \gamma^* .$$

Proof - Under (C0) the results are obvious. Under (C1) or (C2), since $\Sigma_t^2 \equiv -2AK(0)$ and $2A\gamma^* + 2\,|A|\,K(0) - \frac{A_1^2}{B_1^2}\,(\gamma^*)^2 = 0$, either $x = \gamma^*$ and then $\gamma_t^x = \gamma^*$; $t \ge 0$, or $x \ne \gamma^*$ and then the function $(\gamma_t^x\, ; t \ge 0)$ satisfies

$$\frac{\dot{\gamma}_t^x}{-\frac{A_1^2}{B_1^2}\,(\gamma_t^x)^2 + 2A\gamma_t^x - 2AK(0)} = 1 \ ; \ t \ge 0 \ ; \ \gamma_o^x = x .$$

In that case, if $F(y)$ is an integral of $\left\{-\frac{A_1^2}{B_1^2}\,y^2 + 2Ay - 2AK(0)\right\}^{-1}$, then one has $F(\gamma_t) = t + \text{Constant}$. One can easily compute $F(y)$ and then invert the last equation in order to obtain $(\gamma_t^x\, ; t \ge 0)$ in the form stated in the theorem. The last assertion is a simple consequence of that form. ∎

Remarks 3.3. - (a) Theorem 3.2 precises Theorem 16.2 [11] in the one-dimensional case.

(b) The stability property $\lim_{t \nearrow \infty} \gamma_t^x = \gamma^*$ still holds (cf. [9] Th.3.7) if model (1.1) is autonomous and such that a second-order stationary solution exists even if the state process $(X_t\ ; t \geq 0)$ itself is not assumed to be second-order stationary.

4. PARAMETER ESTIMATION IN THE AUTONOMOUS MODEL. In this part of the paper we want to demonstrate how the two steps procedure (cf. section 1) works in model (2.10) and also to compare the method of moments with that of maximum likelihood for estimating the drift coefficient parameters.
We assume that one observes the state process of (2.10) in continuous time and we are interested in estimation of all the parameters a, A, b and B or some functions of these parameters.

Let $C(\mathbb{R}_+ ; \mathbb{R})$ be the canonical space endowed with the σ-algebra $\mathcal{B} = \sigma\{\pi_s\ ; s \geq 0\}$. generated by the coordinates functions $\pi_s, s \geq 0$. Let $P^{b,B}_{\nu,a,A}$ be the distribution of the state process of (2.10) when the distribution of $X(0)$ is ν and let ${}_TP^{b,B}_{\nu,a,A}$ be the restriction of $P^{b,B}_{\nu,a,A}$ to the σ-algebra $\mathcal{B}_T = \sigma\{\pi_s\ ; 0 \leq s \leq T\}$. Then the statistical space under study when one observes the process from 0 to T is :

$$(C(\mathbb{R}_+ ; \mathbb{R}) , \mathcal{B}_T , \{{}_TP^{b,B}_{\nu,a,A}\ ; (\nu , a , A, b, B) \in \Lambda\}) ,$$

where Λ is some subset of the set $\mathcal{P} \times \mathbb{R}^4$ while $\mathcal{P}$ is the set of probability measures on $\mathbb{R}$.

A. First step : parameter estimation in the diffusion coefficient.

By use of the quadratic variation of the observed process one can identify the diffusion coefficient of the model previously to the estimation of the parameters in the drift coefficient. We know that, if $(\langle \pi \rangle (t) ; t \geq 0)$

is the quadratic variation of the canonical process, then the following equalities hold for every $t \in]0,T]$, $(\nu, a, A, b, B) \in \Lambda$:

$$\langle \pi \rangle (t) = \lim_{N \to +\infty} \sum_{i=1}^{2^N} \left(\pi_{\frac{it}{2^N}} - \pi_{(i-1)\frac{t}{2^N}} \right)^2$$

$$= \int_0^t (b + B\pi_s)^2 \, ds \qquad {}_T P^{b,B}_{\mu,a,A} - \text{a.s.}$$

So one is able to compute, by use of the observation of the state process on a finite time, the functions of parameters b and B on which depend the distribution of the process. First, since $B = 0$ if and only if $(t^{-1} \cdot \langle \pi \rangle (t) \, ; t>0)$ is a constant function, one can identify wether $B = 0$ or not. Now, when $B = 0$, one can compute b^2 by $b^2 = t^{-1} \cdot \langle \pi \rangle (t)$. Similarly, when $B \neq 0$, since the system

$$\int_0^t (b + B\pi_s)^2 \, ds = \langle \pi \rangle (t) \; ; \; t \in]0,T] \; ,$$

holds ${}_T P^{b,B}_{\nu,a,A}$ a.s. for every ν , a and A , one can for instance first compute $\frac{b}{B}$ by solving the equation

$$[t_2 \cdot \langle \pi \rangle (t_1) - t_1 \cdot \langle \pi \rangle (t_2)] \frac{b^2}{B^2} + 2 [\langle \pi \rangle (t_1) \int_0^{t_2} \pi_s \, ds - \langle \pi \rangle (t_2) \int_0^{t_1} \pi_s \, ds] \frac{b}{B} +$$

$$+ [\langle \pi \rangle (t_1) \int_0^{t_2} \pi_s^2 \, ds - \langle \pi \rangle (t_2) \int_0^{t_1} \pi_s^2 \, ds] = 0$$

and then compute B^2 by

$$B^2 = \frac{\langle \pi \rangle (t_1)}{\frac{b^2}{B^2} t_1 + \int_0^{t_1} \pi_s^2 \, ds + \frac{2b}{B} \int_0^{t_1} \pi_s \, ds} .$$

Consequently we are allowed to assume that the diffusion coefficient is known, because it would have been eventually previously computed (with probability one) on some finite time interval.

B. Second step : parameter estimation in the drift coefficient.

In the following we are concerned with the reduced statistical model :

$$(C(\mathbb{R}_+ , \mathbb{R}) , \mathcal{B}_T , \{ {}_T P_{\nu,a,A} ; (\nu,a,A) \in \Theta \}) ,$$

where Θ is some subset of $\mathcal{P} \times \mathbb{R}^2$ and parameters b and B have been cancelled in ${}_T P^{b,B}_{\nu,a,A}$ (and will be in $P^{b,B}_{\nu,a,A}$) since these parameters are now fixed known. Since we deal with the case when the model admits a stationary second-order distribution, we consider a parametrization of (2.10) given in terms of the natural parameters of the corresponding process. Taking into account the discussion in section 2B, when $2A + B^2 < 0$, it is better to set :

$$\beta = -A \; ; \; m = -\frac{a}{A} \quad ; \; \sigma^2 = \frac{(b+Bm)^2}{2\beta - B^2}$$

and then to write model (2.10) in the following form :

(4.1) $$dX_t = \beta(m-X_t)dt + (b+BX_t)\,dW_t \; ; \; t \geq 0 \; ; \; X_o = X(0) ,$$

with $\beta > \frac{B^2}{2}$. A second-order stationary solution of (4.1) has mean and covariance functions given by :

$$m_t \equiv m \; ; \quad K(t,s) = \sigma^2 . e^{-\beta|t-s|} \; ; \; t \geq 0 \; , \; s \geq 0 .$$

A strict sense stationary solution of (4.1) corresponds to $\nu = \nu_{\beta,m}$, where $\nu_{\beta,m} = \mu^{b,B}_{\beta m,-\beta}$ is given in Lemma 2.6. When we shall consider the case $B \neq 0$, we shall assume that $m + \frac{b}{B} > 0$ (since the case $m + \frac{b}{B} < 0$ is quite similar) and moreover that $\mathcal{N}$ is a subset of the set $\mathcal{P}(\frac{b}{B})$ of probability measures on $\mathbb{R}$ which are concentrated on $]-\frac{b}{B} , +\infty[$.
Finally, rather to look for estimates of a and A we shall try to estimate

β, m and σ^2 . Then we are concerned with the statistical model :

(4.2) $$(C(\mathbb{R}_+ ; \mathbb{R}) \ , \ \mathcal{B}_T \ , \ \{ {}_T P_{\nu,\beta,m} \ ; \ (\nu,\beta,m) \in \mathcal{C} \}) \ ,$$

where $\mathcal{C}$ is some subset of $\mathcal{P} \times \mathbb{R}_+^* \times \mathbb{R}$ (in fact of $\mathcal{P}(\frac{b}{B}) \times]\frac{B^2}{2}, +\infty[\times]-\frac{b}{B}, +\infty[$ when $B \neq 0$) , ${}_T P_{\nu,\beta,m}$ (resp. $P_{\nu,\beta,m}$) standing for ${}_T P_{\nu,\beta m,-\beta}$ (resp. $P_{\nu,\beta m,-\beta}$) .

An immediate consequence of GIRSANOV's Theorem (see e.g. [11] is that if ν is absolutely continuous with respect to some fixed probability measure ν_o , then the statistical space (4.2) is dominated by ${}_T P_{\nu_o,0,0}$ with the likelihood function

$$\frac{d_T P_{\nu,\beta,m}}{d_T P_{\nu_o,0,0}} = H_{\nu_o,\nu}(\pi_o) \ . \exp\{ L_T(\beta,m) \} \ ,$$

where

$$H_{\nu_o,\nu}(x) = \frac{d\nu}{d\nu_o}(x) \ ; \quad x \in \mathbb{R}$$

and

$$L_T(\beta,m) = \int_0^T \beta(m-\pi_t)(b+B\pi_t)^{-2} \, d\pi_t - \frac{1}{2} \int_0^T \beta^2 (m-\pi_t)^2 (b+B\pi_t)^{-2} dt \ ,$$

the stochastic integral being defined with respect to ${}_T P_{\nu_o,0,0}$.

One can write

$$L_T(\beta,m) = \beta m \, S_o(T) - \beta S_1(T) - \frac{\beta^2}{2} m^2 I_o(T) + \beta^2 m I_1(T) - \frac{\beta^2}{2} I_2(T) \ ,$$

where

$$I_j(T) = \int_0^T \pi_t^j (b+B\pi_t)^{-2} \, dt \ ; \ j=0,1,2$$

and the stochastic integrals

$$S_j(T) = \int_0^T \pi_t^j (b+B\pi_t)^{-2} \, d\pi_t \ ; \quad j=0,1$$

are given explicitely by

$$S_o(T) = \begin{cases} b^{-2}[\pi_T - \pi_o] & \text{if} \quad B = 0 \\ B^{-1}[(b+B\pi_o)^{-1} - (b+B\pi_T)^{-1}] + B[b I_o(T) + B I_1(T)] & \text{if} \quad B \neq 0 , \end{cases}$$

and

$$S_1(T) = \begin{cases} \frac{b^{-2}}{2}[\pi_T^2 - \pi_0^2 - b^2 T] & \text{if} \quad B = 0 \\ \frac{T}{2} + B^{-2}[\mathrm{Log}(\frac{b}{B} + \pi_T) - \mathrm{Log}(\frac{b}{B} + \pi_o)] & \\ \quad - B^{-2} b[(b+B\pi_o)^{-1} - (b+B\pi_T)^{-1}] - b[b I_o(T) + B I_1(T)] & \text{if } B \neq 0. \end{cases}$$

Then, if ν is known, it is easy to see that the maximum likelihood estimate of (β, m) is given by

$$(4.3) \quad \begin{cases} \hat{m}_T = \dfrac{S_o(T) \,.\, I_2(T) - S_1(T) \,.\, I_1(T)}{S_o(T) \,.\, I_1(T) - S_1(T) \,.\, I_2(T)} \\ \\ \hat{\beta}_T = \dfrac{S_o(T) \,.\, I_1(T) - S_1(T) \,.\, I_o(T)}{I_o(T) \,.\, I_2(T) - I_1^2(T)} \end{cases}$$

The asymptotic properties of these estimates are described in the following

<u>Theorem</u> 4.1. - The estimate $(\hat{\beta}_T, \hat{m}_T)$ given by (4.3) is strongly consistent and is asymptotically normal i.e. with respect to $P_{\nu,\beta,m}$ the following limits hold

$$\lim_{T \to +\infty} \text{a.s.} \; (\hat{\beta}_T, \hat{m}_T) = (\beta, m)$$

$$\lim_{T \to +\infty} \mathcal{L}\; T^{1/2} \left[\begin{pmatrix} \hat{\beta}_T \\ \hat{m}_T \end{pmatrix} - \begin{pmatrix} \beta \\ m \end{pmatrix} \right] = N\left(0, \begin{bmatrix} -1 & 0 \\ m & \beta \end{bmatrix}^{-1} \Lambda^{-1}_{\beta,m} \begin{bmatrix} -1 & m \\ 0 & \beta \end{bmatrix}\right)^{-1}$$

where

$$\Lambda_{\beta,m} = \int_{\mathbb{R}} \begin{bmatrix} x(b+Bx)^{-1} \\ (b+Bx)^{-1} \end{bmatrix} \begin{bmatrix} x(b+Bx)^{-1} \\ (b+Bx)^{-1} \end{bmatrix} d\nu_{\beta,m}(x) .$$

Proof - If $L_T^{(1)}(\beta,m)$ stands for the gradient of $L_T(\beta,m)$, we can write

$$L_T^{(1)}(\beta,m) = \begin{bmatrix} -1 & m \\ 0 & \beta \end{bmatrix} \left\{ \begin{bmatrix} S_1(T) \\ S_o(T) \end{bmatrix} - \beta m \begin{bmatrix} I_1(T) \\ I_o(T) \end{bmatrix} + \beta \begin{bmatrix} I_2(T) \\ I_1(T) \end{bmatrix} \right\}.$$

Similarly, if $L_T^{(2)}(\beta,m)$ denotes the matrix of second-order partial derivatives of $L_T(\beta,m)$, we have

$$L_T^{(2)}(\beta,m) = - \begin{bmatrix} -1 & m \\ 0 & \beta \end{bmatrix} \begin{bmatrix} I_2(T) & I_1(T) \\ I_1(T) & I_o(T) \end{bmatrix} \begin{bmatrix} -1 & 0 \\ m & \beta \end{bmatrix} +$$

$$+ \begin{bmatrix} 0 & S_o(T) - \beta m I_o(T) + \beta I_1(T) \\ S_o(T) - \beta m I_o(T) + \beta I_1(T) & 0 \end{bmatrix}$$

Usual arguments in the maximum likelihood method will provide the announced results if we prove that, with respect to $P_{\nu,\beta,m}$, the following limits hold :

(4.4) $$\lim_{T\to+\infty} \text{a.s.} \ \frac{1}{T} L_T^{(1)}(\beta,m) = 0$$

(4.5) $$\lim_{T\to+\infty} \text{a.s.} \ \frac{1}{T} L_T^{(2)}(\beta,m) = - \begin{bmatrix} -1 & m \\ 0 & \beta \end{bmatrix} \Lambda_{\beta,m} \begin{bmatrix} -1 & 0 \\ m & \beta \end{bmatrix} = -\Gamma_{\beta,m}$$

(4.6) $$\lim_{T\to+\infty} \mathcal{D} \ \frac{1}{T^{1/2}} L_T^{(1)}(\beta,m) = N(0, \Gamma_{\beta,m}) .$$

Let us first note that the limits

(4.7) $$\lim_{T\to+\infty} \frac{1}{T} \cdot I_j(T) = \int x^j (b+Bx)^{-2} \, d\nu_{\beta,m}(x) < +\infty \ ; \ j=0,1,2$$

hold because of Lemma 2.6 and of the ergodic properties of the process under study that it ensures (cf. Th.4.1 [12]). Now let us notice that one can write

$$(4.8) \qquad d\pi_t = \beta(m-\pi_t)\, dt + (b+B\pi_t)\, d\tilde{W}_t \ ; \ t \geq 0$$

where $(\tilde{W}_t \ ; t \geq 0)$ is some brownian motion with respect to $P_{\nu,\beta,m}$. It follows that

$$(4.9) \qquad S_j(T) = \beta m\, I_j(T) - \beta I_{j+1}(T) + \int_0^T \pi_t^j \,(b+B\pi_t)^{-1}\, d\tilde{W}_t \ ; \ j = 0,1 \ .$$

Then, using Lemma 17.4 of [11] and taking into account (4.7) for $j = 0,2$, we obtain

$$\lim_{T \to +\infty} \text{a.s.} \left\{ \frac{1}{T} \left[S_j(T) - \beta m\, I_j(T) + \beta I_{j+1}(T) \right] \right\} = 0 ; j = 0,1 \ ,$$

what implies (4.4) and (4.5).

In order to prove (4.6), because of (4.9), we have only to show that

$$\lim_{T \to +\infty} \mathcal{L}\ T^{-1/2} \int_0^T \begin{bmatrix} \pi_t \\ 1 \end{bmatrix} (b+B\pi_t)^{-1}\, dW_t = N(0, \Lambda_{\beta,m}) \ .$$

This fact follows from (4.7) and the results of [14] or [15]. ∎

Remarks 4.2. - (a) These results can be applied, for instance, in the case when $\nu = \delta_{x(0)}$ for $x(0)$ given (with $x(0) > -\frac{b}{B}$ in the case $B \neq 0$). When ν is not known it is clear that one can still use the estimate $(\hat{\beta}_T, \hat{m}_T)$ defined by (4.3) ; in that case it can be considered as approximate maximum likelihood estimate (see [1] and [9] for the classical stationary Gauss-Markov case).

(b) One can use the consistent estimate

$$\hat{\sigma}_T^2 = \frac{(b + \hat{m}_T B)^2}{2\hat{\beta}_T - B^2}$$

for the parameter σ^2 . It is possible to obtain the asymptotic distribution for $(\hat{\beta}_T, \hat{m}_T, \hat{\sigma}_T^2)$ (see [9] for explicite formulas of the asymptotic covariances in cases $B = 0$ and $B \neq 0$).

Now let us use the method of moments for parameter estimation. Let $\tilde{m}_T$ and $\tilde{\sigma}_T^2$ be defined on (4.2) by

$$(4.10) \qquad \begin{cases} \tilde{m}_T = \frac{1}{T} J_1(T) = \frac{1}{T} \int_0^T \pi_t \, dt \\ \tilde{\sigma}_T^2 = \frac{1}{T} J_2(T) - \tilde{m}_T^2 = \frac{1}{T} \int_0^T (\pi_t - \tilde{m}_T)^2 \, dt . \end{cases}$$

We have the following

Theorem 4.3. - The estimate $(\tilde{m}_T, \tilde{\sigma}_T^2)$ given by (4.10) is strongly consistent and the estimate $\tilde{m}_T$ is asymptotically normal with

$$\lim_{T \to +\infty} \mathcal{L}\, T^{1/2} (\tilde{m}_T - m) = N(0, \frac{2\sigma^2}{\beta}) .$$

Moreover if $B = 0$ or if $B \neq 0$ and $\beta > \frac{3B^2}{2}$ one also has

$$\lim_{T \to +\infty} \mathcal{L}\, T^{1/2} \left[\begin{pmatrix} \tilde{m}_T \\ \tilde{\sigma}_T^2 \end{pmatrix} - \begin{pmatrix} m \\ \sigma^2 \end{pmatrix} \right] = N\left(0, \psi_{\beta,m} \Delta_{\beta,m} \psi'_{\beta,m}\right)$$

where

$$\psi_{\beta,m} = \begin{bmatrix} 1/\beta & 0 \\ \frac{2[B(mB+b) - \beta m]}{\beta(2\beta - B^2)} & \frac{2}{2\beta - B^2} \end{bmatrix}$$

and

$$\Delta_{\beta,m} = \int_{\mathbf{R}} \begin{bmatrix} b + Bx \\ x(b+Bx) \end{bmatrix} \begin{bmatrix} b + Bx \\ x(b+Bx) \end{bmatrix}' d\nu_{\beta,m}(x) .$$

Proof - The ergodic properties of the process under study provide the consistency of the estimate since

(4.11) $$\lim_{T\to+\infty} \text{ a.s.} \frac{1}{T} J_i(T) = \int x^i \, d\nu_{\beta,m}(x) < +\infty \; ; \; i = 1,2 \, .$$

Now let us look at $(\tilde{m}_T - m)$ and $(\tilde{\sigma}_T^2 - \sigma^2)$: by use of the representation (4.8) and the fact that

$$\sigma^2 + m^2 = \frac{b^2 + 2m(\beta m + bB)}{2\beta - B^2}$$

it is easy to prove that $T^{1/2}\begin{pmatrix}\tilde{m}_T - m \\ \tilde{\sigma}_T^2 - \sigma^2\end{pmatrix}$ is given by :

$$\begin{bmatrix} \frac{1}{\beta} \cdot \frac{1}{T^{1/2}} (\pi_0 - \pi_T) \\ \frac{1}{2\beta - B^2} \cdot \frac{1}{T^{1/2}} [\pi_0^2 - \pi_T^2 + \frac{2(\beta m + bB)}{\beta} (\pi_0 - \pi_T)] + \frac{\tilde{m}_T + m}{\beta} \cdot \frac{1}{T^{1/2}} (\pi_T - \pi_0) \end{bmatrix} +$$

(4.12)

$$+ \begin{bmatrix} 1/\beta & 0 \\ \frac{2(\beta m + bB)}{\beta(2\beta - B^2)} - \frac{\tilde{m}_T + m}{\beta} & \frac{2}{2\beta - B^2} \end{bmatrix} \frac{1}{T^{1/2}} \int_0^T \begin{bmatrix} b + B\pi_t \\ \pi_t(b + B\pi_t) \end{bmatrix} d\tilde{W}_t \, .$$

Moreover, $E_{\nu,\beta,m}$ standing for expectation with respect to $P_{\nu,\beta,m}$, setting $m_{i,t} = E_{\nu,\beta,m}(\pi_t^i)$; $i = 1,2$, we have :

$$m_{1,t} - m = e^{-\beta t}(m_{1,0} - m)$$

and

$$m_{2,t} - (m^2 + \sigma^2) = e^{(-2\beta + B^2)t} \, [K(0) - (m^2 + \sigma^2) + 2(\beta m + bB)(m_{1,0} - m) \int_0^t e^{(\beta - B^2)s} ds] .$$

Then the limit $\lim_{T\to+\infty} m_{2,T} = m^2 + \sigma^2$ holds, what implies that

(4.13) $$\lim_{T\to+\infty} E_{\nu,\beta,m} \left|\frac{1}{T^{1/2}} (\pi_0^i - \pi_T^i)\right| = 0 \quad ; \quad i=1,2 .$$

Formula (4.12) together with (4.11), (4.13) and the results of [14] ensure that

$$\lim_{T\to+\infty} \mathcal{L}\, T^{1/2} (\tilde{m}_T - m) = N\left(0, \frac{1}{\beta^2} \int (b+Bx)^2 \, d\nu_{\beta,m}(x)\right)$$

where

$$\int_{\mathbb{R}} (b+Bx)^2 \, d\nu_{\beta,m}(x) = b^2 + B^2(m^2+\sigma^2) + 2bBm = 2\sigma^2\beta .$$

Moreover, if $B=0$ or $B \neq 0$ and $\beta > \frac{3B^2}{2}$, since we have $\int x^4 \, d\nu_{\beta,m}(x) < +\infty$, the ergodic property

$$\lim_{T\to+\infty} \text{a.s. } \frac{1}{T} \int_0^T \pi_t^j \, dt = \int x^j \, d\nu_{\beta,m}(x) < +\infty \quad ; \quad j=1,2,3,4$$

hold. Then, by the same arguments, the last assertion in the lemma holds because

$$\lim_{T\to+\infty} \text{a.s. } \left\{ \frac{2(\beta m + bB)}{\beta(2\beta - B^2)} - \frac{\tilde{m}_T + m}{\beta} \right\} = \frac{2[B(mB+b) - \beta m]}{\beta(2\beta - B^2)} . \blacksquare$$

<u>Remarks</u> 4.4. - (a) It is easy to see that in the case $B=0$ the asymptotic variance of $\tilde{m}_T$ is the same as that of $\hat{m}_T$ and oppositely in the case $B \neq 0$ the first one is greater than the second one.

(b) One can use the consistent estimate

$$\tilde{\beta}_T = \frac{(b + B\tilde{m}_T)^2}{2\tilde{\sigma}^2} + \frac{B^2}{2}$$

for the parameter β. It is still possible to obtain the asymptotic distribution for $(\tilde{\beta}_T, \tilde{m}_T, \tilde{\sigma}_T^2)$. (see [9] for explicite formulas of the asymptotic covariances).

REFERENCES

[1] M. ARATO - On the statistical examination of continuous state Markov processes, I. Selected Transl. in Math. Statist. and Probability, 14, (1978), p.203-225.

[2] L. ARNOLD - Stochastic differential equations (Theory and applications), J. Wiley, 1974.

[3] B.M. BROWN and J.I. HEWITT - Asymptotic likelihood theory for diffusion processes. J.Appl. Prob., 12, (1975), p.228-238.

[4] C. BRUNI, G. Di PILLO and G. KOCH - Bilinear systems : an appealing class of "nearly linear" systems in theory and applications. I.E.E.E. Trans. Automat. Contr. Vol. A.C-19, 4, (1974), p.334-348.

[5] C.W. GRANGER and A. ANDERSEN - Non-linear time series modelling- In : Applied Time Series Analysis - D.F. FINDLEY ed., Academic Press, New-York, 1978.

[6] A. ISIDORI - Direct construction of minimal bilinear realizations from non-linear input output maps. I.E.E.E. Trans. Automat. Contr., Vol AC-18, 6, (1973), p.626-631.

[7] K. ITO and M. NISIO - On stationary solutions of a stochastic differential equation. J. Math. Kyoto Univ., 4-1, (1964), p.1-75.

[8] A. LE BRETON - Parameter estimation in a linear stochastic differential equation. In : Trans. 7 th Prague Conf. and 1974 E.M.S. - Acad. Publ. House of the Czechoslovak Acad. of Sc., Vol.A, (1977), p.353-366

[9] A. LE BRETON and M. MUSIELA - A study of an one-dimensional bilinear differential model for stochastic processes, R.R. n°221 Labo. I.M.A.G., Univ. Grenoble 1, France, 1980.

[10] T.S. LEE and F. KOZIN - Almost sure asymptotic lilelihood theory for diffusion processes. J. Appl. Prob., 14, (1977), p.527-537.

[11] R.S. LIPSER and A.N. SHIRYAYEV - Statistics of random processes I, II, Springer-Verlag, Berlin, 1978.

[12] G. MARUYAMA and H. TANAKA - Some properties of one-dimensional diffusion processes. Memoirs of the Faculty of Sciences Kyusyu Univ. Ser.A., 11, 2, (1957), p.117-141.

[13] Yu. V. PROHOROV and Yu. A. ROZANOV - Probability theory, Springer-Verlag, Berlin, 1969.

[14] A.F. TARASKIN - On the asymptotic normality of vector stochastic integrals and estimates of the drift parameters of a many-dimensional diffusion process. Teoriya veroyatnostei i matematicheskaya statistika, 2, Kiev, Izdatel' stro Kiev. Gos. Univ., (1970), p.205-220.

[15] A.F. TARASKIN - Some limit theorems for stochastic integrals. In : Theory of Stochastic Processes, vol.1, J. Wiley, 1974.

[16] E. WONG - The construction of a class of stationary Markoff processes. In : Proc. Symp. in Appl. Math., 16, Ann. Math. Soc., 1964.

A. LE BRETON

Laboratoire I.M.A.G.
B.P. 53 X
38041 Grenoble-Cédex, France

M. MUSIELA

I.M.P.A.N.
ul Kopernika 18
51-617 Wroclaw, Poland

A LOOK AT A BILINEAR MODEL FOR MULTIDIMENSIONAL STOCHASTIC SYSTEMS IN CONTINUOUS TIME (*)

Alain Le Breton , Marek Musiela

Abstract. - This paper is concerned with a study of a multimensional bilinear differential model for stochastic processes in continuous time. Explicit formulas for the mean and covariance functions of the state process are given. Necessary and sufficient conditions for second-order stationarity are provided. A linear representation for the state process is obtained. The optimal linear filter is derived and its asymptotic behaviour is investigated.

1 . Introduction

Deterministic bilinear models for dynamical systems in discrete time (see e.g. T. GOKA and al [8], A. ISIDORI [10]) and in continuous time (see e.g. P. d'ALESSANDRO and al [1], C. BRUNI and al [7]) have been intensively studied. Analogous stochastic models have also been considered in time series analysis (see e.g. C.W. GRANGER and A. ANDERSEN [9], T.D. PHAM and L.T. TRAN [14]) and in continuous time stochastic systems theory (see e.g. L. ARNOLD [2], [3], R.W. BROCKETT [5], [6]).

In [11] we have led a probabisistic and statistical study of a one-dimensional bilinear differential model for stochastic processes.

Key words and phrases : stochastic differential equation, bilinear, second-order stationarity, optimal linear filtering. A.M.S. 1979 Subject classification, Primary 60 H 10, Secondary 93 E 03, (*) Work done during a stay of M. Musiela in Laboratoire I.M.A.G. Grenoble.

In the present paper we consider a multidimentional model which is given by a bilinear stochastic differential equation (in the Itô sense) of the following form

$$dX_t = [A_o(t)X_t + a_o(t)]dt + \sum_{j=1}^{d} [A_j(t)X_t + a_j(t)]dW_t^j \;;\; t\geq 0 \;;\; X_o = X(0) \quad (1.1)$$

where

. $W = (W^1, \dots, W^d)'$ is a standard brownian motion in R^d defined on some basic probability space $(\Omega, \mathcal{A}, (\mathcal{A}_t)_{t\geq 0}, P)$.

. The deterministic functions $A_j = (A_j^{\kappa,\ell} \;;\; \kappa,\ell = 1,\dots,n)$ and $a_j = (a_j^1, \dots, a_j^n)'$, $j=0,\dots,d$ are measurable and such that for every $T>0$, $\kappa,\ell=1,\dots,n$, the following conditions are satisfied :

$$\int_0^T |A_0^{\kappa,\ell}(t)|\,dt < \infty, \quad \int_0^T |a_o^{\kappa}(t)|\,dt < \infty,$$

$$\int_0^T |A_j^{\kappa,\ell}(t)|^2 dt < \infty, \quad \int_0^T |a_j^{\kappa}(t)|^2 dt < \infty \;;\; j=1,\dots,d.$$

. The initial state $X(0)$ is a randomvariable, defined on $(\Omega, \mathcal{A}, P)$, which is independent of W and admits an expectation $m(0)$ and a covariance matrix $K(0)$.

Section 2 is devoted to the analysis of the state process of model (1.1). Using a method similar to that of R.S. LIPTSER and A.M. SHIRYAYEV [13] for linear systems, we compute the mean and covariance functions of the process.

In section 3 we give necessary and sufficient conditions for second-order stationarity of the process generated by model (1.1).

In section 4 we combine the results of the previous sections with those of R.S. LIPTSER and A.N. SHIRYAYEV [13] concerning linear stochastic differential equations. First we obtain a linear representation of the state process of model

(1.1) with respect to a wide-sense Wiener process. We derive the equations for the optimal linear filter. In the case when the equation for observatons is autonomous and the state process is second-order stationary, we show, under usual conditions, that this filter is stable.

The last part of the paper consists of Appendix A and Appendix B containing the proofs of two auxilary results in section 3.

2. The mean and covariance functions.

The existence and uniqueness of the solution process of equation (1.1) is ensured by general results on stochastic differential equations (see e.g. [13] Ch. 4) and the conditions listed in section 1. Let us denote by $(\Phi_t \ ; \ t\geq 0)$ the solution process of the matrix homogeneous equation

$$d\Phi_t = A_o(t)\Phi_t dt + \sum_{j=1}^{d} A_j(t)\Phi_t dW_t^j \ ; \ t\geq 0 \ ; \ \Phi_0 = I_n \qquad (2.1)$$

where I_n is the n x n unit matrix. Setting $D_t = \det \Phi_t$, $t\geq 0$ one can prove (cf. [15]) that

$$D_t = \exp\{\int_0^t tr(A_0(s))ds - \frac{1}{2}\sum_{j=1}^{d} \int_0^t tr(A_j^2(s))ds + \sum_{j=1}^{d} \int_0^t tr(A_j(s))dW_s^j\}; \ t\geq 0,$$

where tr(A) stands for the trace of the square matrix A. This implies that Φ_t is regular for all $t\geq 0$ with probability one. Moreover, the Itô formula provides

THEOREM 2.1

The state process $(X_t \ ; \ t\geq 0)$ generated by model (1.1) is given by

$$X_t = \Phi_t\{X(0) + \int_0^t \Phi_s^{-1}[a_0(s) - \sum_{j=1}^{d} A_j(s)a_j(s)]ds + \\ + \int_0^t \Phi_s^{-1} \sum_{j=1}^{d} a_j(s)\, dW_s^j\} \; ; \; t\geq 0, \tag{2.2}$$

where $(\Phi_t, t\geq 0)$ is defined by (2.1).
(for the proof see [2] Th. 8.4.2).

Denote by $(m_t \; ; \; t\geq 0)$ and $(\pi_t \; ; \; t\geq 0)$ the mean functions of $(X_t \; ; \; t\geq 0)$ and $(\Phi_t \; ; \; t\geq 0)$ respectively. In view of (1.1) and (2.1) it is clear that

$$\dot{m}_t = A_0(t)m_t + a_0(t) \; ; \; t\geq 0 \; ; \; m_0 = m(0) \tag{2.3}$$

and

$$\dot{\Pi}_t = A_0(t)\Pi_t \; ; \; t\geq 0 \; ; \; \Pi_0 = I_n \; . \tag{2.4}$$

<u>REMARK 2.2</u>

The above equations and all others which are written in differential form are considered as integral equations.
In order to obtain analogous formulas for the covariance function

$$K(t,s) = E(X_t - m_t)(X_s - m_s)' \; ; \; t\geq 0 \; , \; s\geq 0 \; ,$$

and the variance function

$$K(t,t) = K_t \; ; \; t\geq 0 \; ,$$

of the state process $(X_t \; ; \; t\geq 0)$, we shall use the following

LEMMA 2.3

The state process (X_t ; $t \geq 0$) of model (1.1) can be written in the form

$$X_t = \Pi_t V_t + m_t \;;\; t \geq 0 \qquad (2.5)$$

where (m_t ; $t \geq 0$) and (Π_t ; $t \geq 0$) are given by (2.3) and (2.4) respectively, and (V_t ; $t \geq 0$)is the solution process of the stochastic differential equation

$$dV_t = \Pi_t^{-1} \sum_{j=1}^{d} \{A_j(t)[\Pi_t V_t + m_t] + a_j(t)\} dW_t^j \;;\; t \geq 0; V_0 = X(0) - m(0). \quad (2.6)$$

Proof : Consider the stochastic process

$$Y_t = \Pi_t V_t + m_t \;;\; t \geq 0.$$

Using the Itô formula and formulas (2.3), (2.4) and (2.6) we obtain

$$dY_t = (\dot{\Pi}_t V_t + \dot{m}_t)\, dt + \Pi_t dV_t =$$

$$= \{A_0(t)\; [\Pi_t V_t + m_t] + a_0\; (t)\} dt + \sum_{j=1}^{d} \{A_j(t)[\Pi_t V_t + m_t] + a_j(t)\} dW_t^j =$$

$$= [A_0(t) Y_t + a_0(t)] dt + \sum_{j=1}^{d} [A_j(t) Y_t + a_j(t)]\, dW_t^j \;;\; t \geq 0 .$$

Moreover, we have

$$Y_0 = \Pi_0 V_0 + m_0 = X(0) - m(0) + m(0) = X(0) = X_0 .$$

Thus , by the uniqueness of the solution of (1.1), the assertion of the lemma follows.∎

In view of (2.5) it is clear that

$$K(t,s) = \Pi_t\; E\; V_t\; V_s'\; \Pi_s' \;;\; t \geq 0 \;;\; s \geq 0 \,.$$

A formula for the term $E\, V_t\, V_s'$ is given in

LEMMA 2.4.

For $t \geq 0$, $s \geq 0$

$$E\, V_t\, V_s' = K(0) + \int_0^{t\wedge s} \sum_{j=1}^{d} \Pi_u^{-1}\, A_j(u)\, \Pi_u G_u \Pi_u'\, A_j'(u)\, (\Pi_u^{-1})'\, du + $$
$$+ \int_0^{t\wedge s} \sum_{j=1}^{d} \Pi_u^{-1}\, [A_j(u) m_u + a_j(u)]\, [A_j(u) m_u + a_j(u)]'\, (\Pi_u^{-1})'\, du,$$

where

$$\dot{G}_t = \sum_{j=1}^{d} \{\Pi_t^{-1}\, A_j(t)\, \Pi_t\, G_t\, \Pi_t'\, A_t'(t)\, (\Pi_t^{-1})' +$$
$$+ \Pi_t^{-1} [A_j(t) m_t + a_j(t)]\, [A_j(t) m_t + a_j(t)]'\, (\Pi_t^{-1})'\} \;;\; t \geq 0$$

while $G_t = EV_t\, V_t'$; $t \geq 0$, $(m_t\ ;\ t \geq 0)$ and $(\Pi_t\ ;\ t \geq 0)$ are given by (2.3) and (2.4) respectively.

Proof : Simple computations based on representation (2.6) of $(V_t\ ;\ t \geq 0)$ provide the first assertion. The second statement immediately follows by setting $t = s$ in the first. ∎

Now we can prove

THEOREM 2.5

The covariance function of the state process of model (1.1) is given by

$$K(t,s) = \begin{cases} \Pi_t\, \Pi_s^{-1}\, K_s & \text{if } t \geq s \\ K_t (\Pi_s\, \Pi_t^{-1})' & \text{if } t \leq s \end{cases} \qquad (2.7)$$

where $(K_t\ ,\ t \geq 0)$ satisfies

$$\dot{K}_t = A_0(t)\, K_t + K_t\, A_0'(t) + \sum_{j=1}^{d} A_j(t)\, K_t\, A_j'(t) +$$
$$+ \sum_{j=1}^{d} [A_j(t) m_t + a_j(t)]\, [A_j(t) m_t + a_j(t)]' \;;\; t \geq 0,\; K_0 = K(0) \qquad (2.8)$$

while $(m_t\ ;\ t \geq 0)$ and $(\Pi_t\ ;\ t \geq 0)$ are defined by (2.3) and (2.4) respectively.

Proof : From Lemma 2.4 and the fact that

$K_t = K(t,t) = \Pi_t G_t \Pi_t'$ it follows

$$\dot{K}_t = \dot{\Pi}_t\ G_t\ \Pi_t' + \Pi_t\ G_t\ \dot{\Pi}_t' + \Pi_t\ \dot{G}_t\ \Pi_t' =$$

$$= A_0(t)\ \Pi_t\ G_t\ \Pi_t' + \Pi_t\ G_t\ \Pi_t'\ A_0'(t) +$$

$$+ \sum_{j=1}^{d} A_j(t)\ \Pi_t G_t \Pi_t' A_j'(t) + \sum_{j=1}^{d} [A_j(t) m_t + a_j(t)]\ [A_j(t) m_t + a_j(t)]'$$

This leads to (2.8). Moreover, since equation (2.8) is linear, its unique solution satisfies

$$K_s = \Pi_s K(0) \Pi_s' + \Pi_s \{ \int_0^s \sum_{j=1}^{d} \Pi_u^{-1}\ A_j(u) K_u\ A_j'(u)\ (\Pi_u^{-1})'\ du +$$

$$+ \int_0^s \sum_{j=1}^{d} \Pi_u^{-1} [A_j(u) m_u + a_j(u)]\ [A_j(u) m_u + a_j(u)]'\ (\Pi_u^{-1})'\ du \} \Pi_s' .$$

So, in view of Lemma 2.4, for $t \geq s$, one has

$$\Pi_s\ E\ V_t\ V_s'\ \Pi_s' = K_s .$$

Finally

$$K(t,s) = \Pi_t\ E\ V_t\ V_s'\ \Pi_s' = \Pi_t\ \Pi_s^{-1}\ \Pi_s\ E V_t\ V_s'\ \Pi_s' = \Pi_t\ \Pi_s^{-1}\ K_s$$

what concludes the proof of (2.7) when $t \geq s$. The case $s \leq t$ can be studied in the same way. ◻

REMARKS 2.6

(a) Theorem 2.5 completes Theorem 8.5.5 in [2]. It can be compared with the results concerning example 1 § 5 in [4].

(b) If A_o does not depend on t we have

$$K(t,s) = e^{A_o(t-s)}\ K_s \ ;\ t \geq s.$$

If moreover $m(0) = 0$, the system is autonomous (A_j ; $j=1,\dots,d$ do not depend on t) and homogeneous ($a_j = 0$; $j=1,\dots,d$) and the matrices A_j, A_j', $j=0,\dots,d$ and $K(0)$ commute with each other, then it is easy to see that

$$K_t = \exp \{[A_0 + A_0' + \sum_{j=1}^{d} A_j A_j']t\} K(0) .$$

So, if $A_0 + A_0' + \sum_{j=1}^{d} A_j A_j' = 0$, then the state process of model (1.1) is a zero-mean second-order stationary process with covariance function

$$K(t,s) = e^{A_0(t-s)} K(0) \; ; \; t \geq s .$$

3. Second-order stationarity

In order to obtain a characterization of second-order stationarity in the general case, we need some preliminary results. Let $(\Sigma(t) \; ; \; t\geq 0)$ be defined by

$$\Sigma(t) = \sum_{j=1}^{d} \{A_j(t) \, K_t A_j'(t) + [A_j(t)m_t + a_j(t)][A_j(t)m_t + a_j(t)]'\} ; \quad t\geq 0 \quad (3.1)$$

where $(m_t \; ; \; t\geq 0)$ and $(K_t \; ; \; t\geq 0)$ are given by (2.3) and (2.8) respectively.

Let $(\Sigma^{1/2}(t) \; ; \; t\geq 0)$ denote a n x n matrix valued measurable function such that

$$\Sigma^{1/2}(t) \, (\Sigma^{1/2}(t))' = \Sigma(t) \; ; \; t\geq 0. \quad (3.2)$$

Moreover let $R(t,s)$ be given by

$$R(t,s) = K(t,s) \, K_s^+ \; , \; s\leq t \quad (3.3)$$

where $(K(t,s) \; ; \; s\leq t)$ and $(K_s \; ; \; s\geq 0)$ are given by (2.7) and (2.8), while K_s^+ stands for the pseudo-inverse matrix of K_s.

Then we have

LEMMA 3.1

If $(R(t,s) \; ; \; s\leq t)$ is defined by (3.3), then the following equality holds for $s\leq u\leq t$:

$$R(t,s) = R(t,u) \, R(u,s). \quad (3.4)$$

Proof : Equations (2.3), (2.7) and (2.8) show that the state process $(X_t \; ; \; t\geq 0)$ has the same second-order structure as a Gauss-Markov process $(\tilde{X}_t \; ; \; t\geq 0)$ satisfying

$$d\tilde{X}_t = [A_0(t) \, \tilde{X}_t + a_0(t)]dt + \Sigma^{1/2}(t) \, d\tilde{W}_t \; ; \; t\geq 0 ,$$

where $(\Sigma^{1/2}(t)\ ;\ t\geq 0)$ is defined by (3.2) and $(\widetilde{W}_t\ ;\ t\geq 0)$ is some standard brownian motion in R^n (see Theorem 15.1 in [13]). Then equation (3.4) follows from considerations just before Theorem 15.2 § 15.1.7 of [13]. ∎

Let P denote the matrix of the orthogonal projection on some linear subspace of R^n and $(\Pi_t\ ;\ t\geq 0)$ be defined by (2.4). Then the following holds :

LEMMA 3.2

The matrix valued function $(\Pi_t P\ ;\ t\geq 0)$ satisfies the semi-group property

$$\Pi_t\ P\ \Pi_s\ P = \Pi_{t+s}\ P\ ;\ s\geq 0\ ,\ t\geq 0 \qquad (3.5)$$

if and only if there exists a constant n x n matrix A_0 such that $A_0\ P = A_0$ and one of the two following equivalent properties

$$\Pi_t\ P = P\ \Pi_t\ P = e^{A_0 t}P = Pe^{A_0 t}\ ;\ t\geq 0 \qquad (3.6)$$

$$A_0(t)P = A_0P = PA_0 \text{ for almost every } t\geq 0 \qquad (3.7)$$

is satisfied.

(for the proof see Appendix A).

Now we are able to prove

THEOREM 3.3

The state process $(X_t\ ;\ t\geq 0)$ generated by model (1.1) is second-order stationary if and only if

$$A_0(t)m(0) + a_0(t) = 0 \text{ for almost every } t\geq 0 \qquad (3.8)$$

and there exists a constant matrix A_0 such that the following conditions hold for almost every $t\geq 0$

$$A_0(t)\ K(0)\ K^+(0) = A_0K(0)\ K^+(0) = K(0)\ K^+(0)\ A_0 = A_0 \qquad (3.9)$$

$$A_0\ K(0) + K(0)\ A_0' + \sum_{j=1}^{d} A_j(t)\ K(0)\ A_j'(t) + \sum_{j=1}^{d} [A_j(t)m(0) + a_j(t)]\ [A_j(t)m(0) + a_j(t)]' = 0 \qquad (3.10)$$

If conditions (3.8) - (3.10) are satisfied, then the covariance function of $(X_t\ ;\ t\geq 0)$ is given by

$$K(t,s) = \begin{cases} e^{A_0(t-s)} K(0) & \text{if } t\geq s \\ K(0)e^{A_0'(s-t)} & \text{if } t\leq s . \end{cases} \quad (3.11)$$

<u>Proof</u> : First we prove that conditions (3.8) - (3.10) are sufficient. Note that since the solution of (2.3) is unique, condition (3.8) leads to $m_t \equiv m(0)$. Moreover, since condition (3.8) implies that $A_0(t)\ K(0) = A_0K(0)$ and the solution of (2.8) is unique, condition (3.10) leads to $K_t \equiv K(0)$.
It is also clear that Lemma 3.2 and condition (3.9) imply

$\Pi_t P = e^{A_0 t} P = Pe^{A_0 t}$, $t\geq 0$, where $P = K(0)\ K^+(0)$.
Consequently $P = \Pi_t^{-1} e^{A_0 t} P = \Pi_t^{-1} P e^{A_0 t}$; $t\geq 0$ and then
$\Pi_t^{-1} P = Pe^{-A_0 t} = e^{-A_0 t} P$, $t\geq 0$, what, together with representation (2.7), provides (3.11).

Conversely, if $(X_t\ ;\ t\geq 0)$ generated by (1.1) is second-order stationary then $m_t = m(0)$, $K_t = K(0)$; $t\geq 0$, and taking into account Lemma 3.1, for $s\leq u\leq t$,

$R(t-s,0) = R(t,s) = R(t,u)\ R(u,s) = R(t-u,0)\ R(u-s,0)$.

Using (2.7) and (3.3) we obtain that

$(\Pi_t P\ ;\ t\geq 0)$
where $P = K(0)\ K^+(0)$, is a semi-group. Now from Lemma 3.2 it follows that there exists a constant matrix A_0 such that $A_0(t)P = A_0 P = PA_0 = A_0$ for almost every $t\geq 0$. This, together with (2.3) and (2.8), provides the result. ∎

<u>REMARK 3.4</u>

In the case when K(0) is regular the set of conditions (3.8)-(3.10) is equivalent to the existence of a constant matrix A_0 and a constant vector a_0 such that $A_0 m(0) + a_0 = 0$ and, for almost every $t\geq 0$, $A_0(t) = A_0$, $a_0(t) = a_0$, and condition

(3.10) holds (cf. [12]).

The above theorem can be completed by use of the following

LEMMA 3.5

Let $A_0,\dots,A_d$ be n x n matrices such that the matrix $A_0 \otimes I_n + I_n \otimes A_0 + \sum_{j=1}^{d} A_j \otimes A_j$ is stable. For every n x n symmetric non-negative definite matrix M, the matrix equation

$$A_0 K + K A_0' + \sum_{j=1}^{d} A_j K A_j' + M = 0 \qquad (3.12)$$

admits a unique n x n symmetric non-negative definite solution matrix K(M). Moreover, if M is positive definite, then K(M) is also positive definite.

(for the proof, see Appendix B ; see also Theorem 2 in [6]).

The following statement is then an immediate consequence of Theorem 3.3.

COROLLARY 3.6

Let $A_0,\dots,A_d$ be n x n matrices such that the matrix

$$A_0 \otimes I_n + I_n \otimes A_0 + \sum_{j=1}^{d} A_j \otimes A_j$$

is stable and let $a_1,\dots,a_d$ be vectors in R^n. If K(0) is the solution of (3.12) with

$$M = \sum_{j=1}^{d} [A_j m(0) + a_j]\,[A_j m(0) + a_j]',$$

then the state process $(X_t \,;\, t \geq 0)$ generated by the model

$$dX_t = A_0[X_t - m(0)]dt + \sum_{j=1}^{d} [A_j X_t + a_j] dW_t^j \;;\; t \geq 0,\; X_0 = X(0)$$

is second-order stationary with the covariance function given by (3.11).

4. Linear representation and optimal linear filtering

Following up the ideas presented in [13] Ch. 15 we shall now provide a linear representation of the state process $(X_t \,;\, t \geq 0)$ generated by (1.1) and then investigate the optimal linear filtering problem of that process.

THEOREM 4.1

There exists a wide-sense Wiener process $W^* = (W_t^* \,;\, t \geq 0)$

in R^n such that the state process $(X_t \; ; \; t \geq o)$ generated by model (1.1) admits the representation

$$dX_t = [A_0(t)X_t + a_0(t)]dt + \Sigma^{1/2}(t)\, dW_t^* \; ; \; t \geq o ; X_o = X(0) \quad (4.1)$$

where $(\Sigma^{1/2}(t) \; ; \; t \geq o)$ is defined by (3.2). Moreover, X(0) and W^* are uncorrelated.

Proof :We verify that all assumptions in Theorem 15.2 of [13] are satisfied with Γ, a_1, B_1, replaced respectively by K, A_o, Σ. Lemma 3.1 shows that condition (15.40) in [13] holds. Since (K_t) and (m_t) are continuous, condition (1) in [13] is verified. Formulas (2.7), (2.8) and (2.3) clearly imply respectively conditions (2), (3), (4) in [13]. Then representation (4.1) follows. Moreover from the proof of Theorem 15.2 in [13], it is easy to see that the last assertion holds. ◻

Combining Theorems 3.3 and 4.1 we obtain

COROLLARY 4.2

If the state process $(X_t \; ; \; t \geq o)$ generated by model (1.1) is second-order stationary, then there exists a wide-sense Wiener process W^* in R^n and a constant n x n matrix A_0 satisfying (3.9) almost everywhere such that the representation

$$dX_t = A_0[X_t - m(0)]dt + \Sigma^{1/2} dW_t^* \; ; \; t \geq o \; ; \; X_0 = X(0),$$

holds with

$$\Sigma^{1/2}(\Sigma^{1/2})' = \Sigma = -A_0K(0) - K(0)A_0'. \quad (4.2)$$

Proof : If $(X_t \; ; \; t \geq o)$ is second-order stationary then, from Theorem 4.1 and equalities (3.1), (3.10), representation(4.1) holds with $\Sigma(t) = -A_0K(0) - K(0)\, A_0'$. From (3.8)

$$A_0(t)X_t + a_0(t) = A_0(t)\, [X_t - m(0)]$$

holds almost everywhere. Moreover, since clearly, almost surely

$$X_t - m(0) = K(0)\, K^+(0)\, [X_t - m(0)]$$

for every t, condition (3.9) implies that $A_0(t)X_t + a_0(t)$ can be replaced by $A_0[X_t - m(0)]$ in (4.1).

Now we consider the linear filtering problem(in the sense of [13] Ch. 15) for the state process $(X_t\ ;\ t\geq 0)$ of model (1.1) by use of observation on $[0,t]$ of the process $(Y_t\ ;\ t\geq 0)$ which is given by

$$dY_t = [B_0(t)\ X_t + b_0(t)]dt + B_1(t)\ dZ_t\ ;\ t\geq 0\ ;\ Y_0 = Y(0)$$

where

. $Z = (Z^1,\dots,Z^q)$ is a standard brownian motion in R^q, defined on the basic space $(\Omega,\mathcal{Q},(\mathcal{Q}_t)_{t\geq 0},P)$, independent of W.

. the deterministic functions $B_0=((B^{k,\ell};k=1,\dots,q\ ;\ \ell=1\dots n))$, $b_0=(b_0^1,\dots,b_0^q)'$ and $((B_1^{k,\ell};\ k,\ell=1\dots q))$ satisfy analagous conditions as A_0, a_0 and A_1 respectively. Moreover, the matrix $B_1(t)\ B_1'(t)$ is regular for every $t\geq 0$.

. the initial state $Y(0)$ is a second-order random vector defined on $(\Omega,\mathcal{Q},P)$ such that $(X(0),\ Y(0))$ is independent of (W,Z).

It is easy to see that Theorem 4.1 above and Theorem 15.3 in [13] provide

<u>THEOREM 4.3</u>

The optimal linear filter $\hat{X}_t$ of the state process X_t from the observation of the process Y_t is given by

$$d\hat{X}_t=[A_0(t)\hat{X}_t+a_0(t)]dt+\gamma_t B_0'(t)[B_1(t)B_1'(t)]^{-1}\{dY_t-[B_0(t)\hat{X}_t+b_0(t)]dt\}$$

$$\dot{\gamma}_t=A_0(t)\gamma(t)+\gamma_t A_0'(t)+\Sigma(t)-\gamma_t B_0'(t)[B_1(t)B_1'(t)]^{-1}B_0(t)\gamma_t \qquad (4.3)$$

with

$$\hat{X}_0=m(0) + Cov(X(0),\ Y(0))\ Cov^+(Y(0),\ Y(0))\ [Y(0)-EY(0)]$$

$$\gamma_0=K(0) - Cov(X(0),\ Y(0))\ Cov^+(Y(0),\ Y(0))\ Cov\ (Y(0),\ X(0))$$

where $(\Sigma(t)\ ;\ t\geq 0)$ is given by (3.1) while $\gamma_t=E(X_t-\hat{X}_t)\ (X_t-\hat{X}_t)'$.

Let us now look at the stability of the optimal linear filter. We assume that the equation for the observation is autonomous (i.e. B_0, B_1 and b_1 do not depend on t) and that the state process $(X_t;\ t\geq 0)$ is second-order stationary. Let A_0 denote the constant matrix given by Corollary 4.2. The following result is a simple consequence of Theorem 16.2 in [13] and Corollary 4.2.

THEOREM 4.4

Assume that the pair (A_0, B_0) is observable and that the pair $(A_0, \Sigma^{1/2})$, where $\Sigma^{1/2}$ is defined by (4.2), is controllable. Then, if $(\gamma_t\ ;\ t\geq 0)$ is defined by (4.3) with $\Sigma_t = -A_0K(0) - K(0)A_0'$, $t\geq 0$, the limit $\lim_{t\to\infty} \gamma_t$ exists. This limit does not depend on the initial value and is the unique (in the class of symmetric positive definite matrices) solution γ of the matrix equation

$$A_0\gamma + \gamma A_0' - A_0K(0) - K(0)A_0' - \gamma B_0'[B_1\ B_1']^{-1}B_0\ \gamma = 0.$$

APPENDIX A. - Proof of Lemma 3.2

Let us first show that equation (3.5) implies that there exists A_0 such that $A_0P = A_0$ and (3.6) is satisfied.

Note that $P^2 = P$ and

(*) $P\Pi_t P P\Pi_s P = P\Pi_t P\Pi_s P = P\Pi_{t+s}P\ ;\ t\geq 0\ ,\ s\geq 0.$

There exists also an orthogonal matrix B such that

$$B'PB = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix},$$

where r is the rank of P. Moreover

$$B'\Pi_s PB = \begin{bmatrix} U_{11}(s) & U_{12}(s) \\ U_{21}(s) & U_{22}(s) \end{bmatrix}$$

where $U_{11}(s)$, $U_{22}(s)$, $U_{12}(s)$, $U_{21}(s)$ are $r \times r$, $(n-r)\times(n-r)$, $r\times(n-r)$ and $(n-r)\times r$ matrices respectively. This implies that

$$B'P\Pi_s PB = \begin{bmatrix} U_{11}(s) & U_{12}(s) \\ 0 & 0 \end{bmatrix}$$

what, together with (*), leads to

$U_{11}(s)\ U_{11}(t) = U_{11}(t+s)\ ;\ s\geq 0\ ,\ t\geq 0\ ;\ U_{11}(0) = I_r,$

$U_{11}(t)\ U_{12}(s) = U_{12}(t+s)\ ;\ s\geq 0\ ,\ t\geq 0\ ;\ U_{12}(0) = 0\ .$

Consequently, because $(U_{11}(t)\ ;\ t\geq 0)$ is a semi-group with the starting point $U_{11}(0) = I_r$, there exists a matrix A such that $U_{11}(t) = e^{At}\ ;\ t\geq 0$.

Moreover, taking into account

$U_{12}(t) = U_{11}(t) \quad U_{12}(0) = 0 \; ; \; t \geq 0,$

it is clear that

$$B'\Pi_t PB = \begin{bmatrix} e^{At} & 0 \\ U_{21}(t) & U_{22}(t) \end{bmatrix} ; t \geq 0$$

But $(B'\Pi_t PB \; ; \; t \geq 0)$ is also a semi-group what gives

$U_{22}(t) \; U_{22}(s) = U_{22}(t+s) \; ; \; t \geq 0 \; ; \; s \geq 0 \; ; \; U_{22}(0) = 0$

$U_{21}(t) \; e^{As} + U_{22}(t) \; U_{21}(s) = U_{21}(t+s) \; ; \; t > 0 \; ; \; s > 0 \; ; \; U_{21}(0) = 0$

Now it is easy to deduce that

$$B'\Pi_t PB = B'P\Pi_t PB = \begin{bmatrix} e^{At} & 0 \\ 0 & 0 \end{bmatrix} ; \; t \geq 0,$$

and, setting

$$\tilde{A} = \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix}$$

to obtain

$$B'\Pi_t PB = B'P\Pi_t PB = e^{\tilde{A}t} - \begin{bmatrix} 0 & 0 \\ 0 & I_{n-r} \end{bmatrix}$$

This shows that, setting $A_0 = B \tilde{A} B'$

$$\Pi_t P = P\Pi_t P = B\left(e^{\tilde{A}t} - \begin{bmatrix} 0 & 0 \\ 0 & I_{n-r} \end{bmatrix}\right) B' =$$

$$= Be^{\tilde{A}t}B' - (I_n - P) = e^{A_0 t} - (I_n - P).$$

Consequently $\Pi_t P = P\Pi_t P = PP\Pi_t P = Pe^{A_0 t} = P\Pi_t PP = e^{A_0 t} P.$

Moreover, since $\tilde{A} \; B' \; P \; B = \tilde{A}$, one has $A_0 P = A_0$.

- Let us prove now that conditions (3.6) and (3.7) are equivalent. If (3.6) holds, then $A_0 \; e^{A_0 t} \; P = P \; A_0 \; e^{A_0 t}$; $t > 0$ and $A_0 P = PA_0$.

Moreover it is clear that, for almost every $t \geq 0$,

$A_0(t) \; \Pi_t P = A_0(t) \; Pe^{A_0 t} = PA_0 \; e^{A_0 t}$

what provides (3.7).

Conversely, note that from the proof of Theorem 4.10 in [13] $\Pi_t = \lim_{k\to\infty} \Pi_t^k$, where $\Pi_t^0 = I_n$ and

$\Pi_t^{k+1} = I_n + \int_0^t A_0(s)\, \Pi_s^k\, ds$, $t\geq 0$; $k = 0,1,\ldots$.

This implies that $\Pi_t P = \lim_{k\to\infty} \Pi_t^k P$, $t\geq 0$. Moreover, because of (3.7), we have

$\Pi_t^k P = (I_n + A_0 t + \ldots + A_0^k \frac{t^k}{k!})\, P$; $t\geq 0$.

Then we obtain $\Pi_t P = P\, \Pi_t P = e^{A_0 t} P = P\, e^{A_0 t}$.

Finally the fact that (3.6) leads to (3.5) is quite obvious.∎

APPENDIX B. - Proof of Lemma 3.5.

Since the matrix $L = A_0 \otimes I_n + I_n \otimes A_0 + \sum_{j=1}^{d} A_j \otimes A_j$ is stable and equation (3.12) can be written

$L \text{ vec } K = - \text{vec } M$

the solution K(M) exists and is unique. Now let us prove that it is non-negative definite.

Let $(U_t,\ t\geq 0)$ be the solution of the equation

$\dot{U}_t = A_0 U_t + U_t A_0' + \sum_{j=1}^{d} A_j U_t A_j'$; $t\geq 0$, $U_0 = M$.

Since L is stable and

$\text{vec } U_t = \exp Lt.\ \text{vec } M$

it is clear that $\int_0^{+\infty} U_t dt$ exists , $\lim_{t\to+\infty} U_t = 0$ and

$$- U_0 = A_0 \int_0^{+\infty} U_t dt + \int_0^{+\infty} U_t dt\, A_0' + \sum_{j=1}^{d} A_j \int_0^{+\infty} U_t dt\, A_j'.$$

Thus by uniqueness of the solution of (3.12) we obtain that $K(M) = \int_0^{+\infty} U_t dt$.

Taking derivative of $e^{A_0 t} H_t\, e^{A_0' t}$ where

$\dot{H}_t = \sum_{j=1}^{d} e^{-A_0 t} A_j\, e^{A_0 t} H_t\, e^{A_0' t} A_j' e^{-A_0' t}$; $t\geq 0$; $H_0 = M$,

by the uniqueness argument one gets

$U_t = e^{A_0 t} H_t\, e^{A_0' t}$.

Let us define now $H_t^0 = M$ and

$$H_t^{n+1} = \int_0^t \sum_{j=1}^{d} e^{-A_0 u} A_j e^{A_0 u} H_u^n e^{A_0' u} A_j' e^{-A_0' u} du \; ; \; t \geq 0.$$

Proceeding as in the proof of Theorem 4.10 [13], one can show that the series $\sum_{n=0}^{\infty} H_t^n$ converges uniformly on every interval [0,T] and that the solution of the above equation for H_t is uniquely given by

$$H_t = \sum_{n=0}^{\infty} H_t^n \; , \; t \geq 0.$$

Consequently

$$K(M) = \int_0^{\infty} \sum_{n=0}^{\infty} e^{A_0 t} H_t^n e^{A_0' t} dt$$

and, in order to prove that K(M) is non-negative definite, it is sufficient to prove that $e^{A_0 t} H_t^n e^{A_0' t}$ is non-negative definite for every $n = 0,1,2,\ldots$ and $t \geq 0$.

It is clear that $e^{A_0 t} H_t^0 e^{A_0' t} = e^{A_0 t} M e^{A_0' t}$ is non-negative definite (if M is positive definite, then it is positive definite) and because

$$e^{A_0 t} H_t^{n+1} e^{A_0' t} = \int_0^t \sum_{j=1}^{d} e^{A_0 (t-u)} A_j e^{A_0 u} H_u^n e^{A_0' u} A_j' e^{A_0' (t-u)} du$$

the assertion of the lemma follows. ◻

REFERENCES

[1] P. d' ALESSANDRO, A. ISIDORI and A. RUBERTI - Realization and structure theory of bilinear dynamical systems, SIAM J. Control 12, 3, 517-535 (1974).

[2] L. ARNOLD - Stochastic differential equations, Theory and applications, New-York, John Wiley (1974).

[3] L. ARNOLD - Qualitative theory of stochastic non linear systems, Tech. Rept. N° 69, Palaiseau, France, Centre de Math. Appl., Ecole Polytechnique, (1980).

[4] R. BOUC and E. PARDOUX - Moments of semi-linear random evolutions, Tech. Rept. N° 80-5, Marseille, France, Publications de Math. Appl., Univ. Provence, (1980).

[5] R. W. BROCKETT - Lie algebras and Lie groups in control theory, in : D. Q. Mayne and R. W. Brockett, Eds., Geometric methods in system theory, Dordrecht, Reidel, (1973).

[6] R.W. BROCKETT - Parametrically, stochastic linear differential equations, in : R. J. B. Wets, Ed., Stochastic systems : Modeling, Identification and Optimization, I, Math. Prog. Study,5 Amsterdam, North-Holland (1976).

[7] C. BRUNI, G. di PILLO and G. KOCH - Bilinear systems : an appealing class of "nearly linear" systems in theory and applications, I.E.E.E. Trans. Automat. Contr. A.C. 19,4, 334-348, (1974).

[8] T. GOKA, T.J. TARN and J. ZABORSKY - On the controllability of a class of discrete bilinear systems, Automatica 9, 615-622, (1973).

[9] C.W. GRANGER and A. ANDERSEN - Non linear time series modelling, in : D.F. Findley, Ed., Applied time series analysis, New-York, Academic Press, (1978).

[10] A. ISIDORI - Direct construction fo minimal linear realizations from non-linear input-output maps, I.E.E.E. Trans. Automat. Contr. A.C. 18,6, 626-631, (1973).

[11] A. LE BRETON and M. MUSIELA - A study of an one-dimensional bilinear differential model for stochastic processes, R.R. n° 221, Grenoble, France, Laboratoire IMAG, Univ. Grenoble 1, (1980).

[12] A. LE BRETON and M. MUSIELA - Moments et filtrage linéaire d'un système bilinéaire à temps continu, C.R. Acad. Sc. Paris t. 292, I, 83-86, (1981).

[13] R.S. LIPTSER and A.N. SHIRYAYEV - Statistics of Random Processes, I, II, New-York, Springer-Verlag, (1978).

[14] T.D. PHAM and L.T. TRAN - Quelques résultats sur les modèles bilinéaires de séries chronologiques, C.R. Acad. Sc. Paris t. 290, A, 335-338, (1980).

[15] I. VRKOC - Liouville formula for systems of linear homogeneous Itô stochastic differential equations, Commentationes Mathematicae Univ. Carolinae, 19, 1, 141-146, (1978).

Le Breton Alain
Laboratoire I.M.A.G.
B.P. 53 X
38041 Grenoble-Cédex
France

Musiela Marek
I.M.P.A.N.
Kopernika 18
51-617 Wroclaw
Poland

SOME PARAMETER ESTIMATION PROBLEMS FOR HYPOELLIPTIC HOMOGENEOUS GAUSSIAN DIFFUSIONS (*)

A. LE BRETON
Laboratoire I.M.A.G
B.P. 53 X
38041 GRENOBLE Cedex
FRANCE

and

M. MUSIELA
I.M.P.A.N.
Kopernika 18
51-617 Wroclaw
POLAND

ABSTRACT- This paper is concerned with the statistical problem of parameter estimation for hypoelliptic homogeneous gaussian diffusions. Since quadratic forms of the processes under study play a central role, some of their properties are proved first. Then the maximum likelihood method is used to derive ordinary and sequential plans for parameter estimation and characteristics of these plans are studied.

I - INTRODUCTION

Stochastic linear models for dynamical systems in continuous time have been intensively studied as well in the litterature concerning systems theory ([3], [18], ...) as in that about statistics of random processes ([1], [10], ...)

In the present paper we are concerned with a multidimensional model which is defined by an autonomous linear stochastic differential equation of the following form :

$$(I.1) \quad dX_t^x = AX_t^x\,dt + G\,dW_t \quad ; \quad t \geq 0 \quad ; \quad X_0^x = x$$

where

. $W = (W^1, \ldots, W^r)'$ is some standard brownian motion in $\mathbb{R}^r$

(*) Work done during a stay of M.Musiela in Laboratoire I.M.A.G., Grenoble.

defined on some basic probability space $(\Omega, \mathcal{A}, (\mathcal{A}_t)_{t\geq 0}, P)$,

. A and G are $n \times n$ and $n \times r$ constant matrices,

. x stands for any initial state of the system in $\mathbb{R}^n$.

Let us recall that (cf.[10]) for every $x \in \mathbb{R}^n$, the proces $X^x = (X_t^x, t \geq 0)$ is a gaussian Markov process with mean function

$$m_t^x = EX_t^x = e^{At} x \quad ; \quad t \geq 0 \quad ,$$

and covariance function (not depending on x)

$$K_{t,s} = E\left\{\left(X_t^x - EX_t^x\right)\left(X_t^x - EX_t^x\right)'\right\}$$

$$= \begin{cases} e^{A(t-s)} K_s & \text{if} \quad t \geq s \\ K_t \; e^{A'(s-t)} & \text{if} \quad t \leq s \end{cases}$$

where the variance function $(K_t \; ; \; t \geq 0)$ is given by

$$\dot{K}_t = AK_t + K_t A' + GG' \quad ; \quad t \geq 0 \quad ; \quad K_0 = 0$$

or equivalently

$$K_t = e^{At} \int_0^t e^{-As} \, GG' \, e^{-A's} \, ds \quad e^{A't} \quad ; \quad t \geq 0 \quad .$$

Moreover $(X^x \; ; \; x \in \mathbb{R}^n)$ is an homogeneous gaussian diffusion corresponding to the differential generator

$$L = \frac{1}{2} \sum_{i,j=1}^{n} b_{ij} \frac{\partial^2}{\partial x_i \, \partial x_j} + \sum_{i,j=1}^{n} a_{ij} \, x_j \frac{\partial}{\partial x_i}$$

where $A = ((a_{ij}))$ and $B = ((b_{ij})) = G.G'$. It is known (cf. [6] ,[7]) that if the differential generator is hypoelliptic that is equivalently that the pair $[A,G]$

is controllable or

(H1) rank $[G, AG, \dots, A^{n-1}G] = n$

then, for every $0 \leq s < t$, the integral $\int_s^t e^{-Au} GG' e^{-A'u} du$ is a positive definite matrix and in particular, for every $t > 0$, the covariance matrix K_t is regular. Obviously this occurs for example when G is a $n \times n$ regular matrix.

It is also known (cf. [7], [19]) that if (H1) is satisfied together with the assumption

(H2) A is a stable matrix

then there exists a unique invariant probability measure for the diffusion $(X^x \; ; \; x \in \mathbf{R}^n)$; this measure is gaussian with mean zero and nonsingular covariance matrix

$$(I.2) \quad K_{A,B} = \lim_{t \to +\infty} K_t = \int_0^{+\infty} e^{As} GG' e^{A's} ds$$

which is the unique nonnegative definite symmetric matrix satisfying the Lyapunov equation

$$AK + KA' + GG' = 0 \ .$$

Moreover from [11] it follows that under (H1) - (H2) the diffusion $(X^x \; ; x \in \mathbf{R}^n)$ is ergodic.

Here we are interested in the statistical problem of estimating the parameters A and $B = GG'$ of the diffusion when they are unknown, in view of the observation on some time interval $[0,T]$ of one trajectory. This problem has been studied in [1] and [8] under the assumption that the differential generator is elliptic i.e. $\det B \neq 0$ and sequential procedures have been investigated for the one(resp. two) dimensional case cf.[10] (resp. cf.[13]).

In the present work we show that the methods used in these papers may be extended in some ways under the weaker assumption (H1). Since preliminary results which are of interest on their own merits are needed in statistical considerations, the paper is organized as follows :

. Section 2 is devoted to the computation of some Radon-Nikodym derivatives and to the sudy of quadratic forms of the observed process .

. Section 3 is concerned with parameter estimation by the maximum likelihood method.

. In section 4 sequential schemes for parameter estimation are studied.

II - Preliminary results

Let $C = C(R_+ ; R^n)$ be the space of all continuous functions from R_+ into R^n and, for every $T > 0$, C_T the σ-algebra generated on C by the coordinates π_t for $0 \le t \le T$. If $Z = (Z_t ; t \ge 0)$ is some random process with continuous sample paths, μ_Z^T will stand for the restriction to C_T of the probability distribution induced by Z on C.

II.1 Radon-Nikodym derivatives.

The following lemma will be important in future considerations :

II.1.1 Lemma - Let $(\Gamma(t) ; t \ge 0)$ be some $n \times n$ matrix valued function which is absolutely continuous with derivative $\dot{\Gamma}(t)$ and such that for every $T > 0$

$$\int_0^T |\Gamma_{k\ell}(t)| \, dt < +\infty \qquad ; k, \ell = 1, \dots, n ,$$

where $\Gamma(t) = ((\Gamma_{k\ell}(t)))$. Let $Y^x = (Y_t^x ; t \ge 0)$ be the solution process of the stochastic differential equation

$$\text{(II.1)} \quad dY_t^x = [A + GG'\Gamma(t)] \, Y_t^x \, dt + G \, d\eta_t \quad ; \quad t \ge 0 \quad ; \quad Y_0^x = x ,$$

where $(\eta_t ; t \ge 0)$ is some r-dimensional brownian motion. Then the measure $\mu_{X^x}^T$ is absolutely continuous with respect to the measure $\mu_{Y^x}^T$ with following Radon-Nikodym derivative :

$$\frac{d\mu_{X^x}^T}{d\mu_{Y^x}^T} = \exp\left\{ -\int_0^T \pi'_t \, \Gamma'(t) \, GG^+ d\pi_t + \frac{1}{2} \int_0^T \pi'_t \, \Gamma'(t) GG^+ [2A + GG'\Gamma(t)] \, \pi_t \, dt \right\} .$$

If moreover $\Gamma(t)$ is a symmetric matrix for $t \ge 0$, then the Radon-Nikodym derivative can be written

$$\frac{d\mu^T_{X}x}{d\mu^T_{Y}x} = \exp\left\{\frac{1}{2}\int_0^T \mathrm{Tr}\,(G'\Gamma(t)G)\,dt + \frac{1}{2}x'\,\Gamma(0)x - \frac{1}{2}\,\pi'_T\,\Gamma(T)\,\pi_T\right\} \times$$

$$\times \exp\left\{\frac{1}{2}\int_0^T \pi'_T\,[\dot{\Gamma}(t) + A'\Gamma(t) + \Gamma(t)A + \Gamma(t)GG'\Gamma(t)]\,\pi_t\,dt\right\}.$$

Proof - Note that the equation

$$Ax - (A + GG'\Gamma(t))\,x = G\alpha_t(x)$$

admits the solution

$$\alpha_t(x) = -G'\Gamma(t)x \quad .$$

Then (cf. [10]) one has $\mu^T_{X}x \ll \mu^T_{Y}x$ and moreover

$$\frac{d\mu^T_{X}x}{d\mu^T_{Y}x} = \exp\left\{\int_0^T [A\pi_t - \{A + GG'\Gamma(t)\}\,\pi_t]'[GG']^+\,d\pi_t - \right.$$

$$\left. - \frac{1}{2}\int_0^T [A\pi_t - \{A + GG'\Gamma(t)\}\,\pi_t]'[GG']^+[A\pi_t + \{A + GG'\Gamma(t)\}\pi_t]\,dt\right\}.$$

Using properties of pseudoinverses, namely $[GG']\,[GG']^+ = GG^+$ (cf. [10]), the first assertion in the lemma immediately follows. Now, if $\Gamma(t)$ is symmetric, the stochastic integral can be computed as follows :

$$\int_0^T \pi'_t\,\Gamma'(t)GG^+\,d\pi_t = \int_0^T \pi'_t\,\Gamma(t)\,GG^+\,d\pi_t$$

$$= \int_0^T \pi'_t\,\Gamma(t)\,d\pi_t - \int_0^T \pi'_t\,\Gamma(t)\,(E-GG^+)\,d\pi_t$$

$$= \int_0^T \pi'_t\,\Gamma(t)\,d\pi_t - \int_0^T \pi'_t\,\Gamma(t)\,(E-GG^+)\,A\,\pi_t\,dt$$

where E stands for the $n \times n$ identity matrix. By the Itô formula one gets :

$$\int_0^T \pi'_t\, \Gamma(t)\, d\pi_t = \frac{1}{2} \left\{ \int_0^T \pi'_t\, \Gamma(t)\, d\pi_t + \int_0^T d\pi'_t\, \Gamma(t)\, \pi_t \right\}$$

$$= \frac{1}{2} \left\{ \pi'_T\, \Gamma(T)\, \pi_T - x'\, \Gamma(0)\, x - \int_0^T \mathrm{Tr}\,(G'\Gamma(t)G)\, dt - \int_0^T \pi'_t\, \dot{\Gamma}(t)\, \pi_t\, dt \right\} .$$

Then one has

$$\frac{d\mu^T_{Xx}}{d\mu^T_{Yx}} = \exp \left\{ \frac{1}{2} \int_0^T \mathrm{Tr}\,(G'\Gamma(t)G)\, dt + \frac{1}{2}\, x'\,\Gamma(0)\, x - \frac{1}{2}\, \pi'_T\, \Gamma(T)\, \pi_T \right\} \times$$

$$\times \exp \left\{ \frac{1}{2} \int_0^T \pi'_t\, \dot{\Gamma}(t)\, \pi_t\, dt \right\} \times$$

$$\times \exp \left\{ \int_0^T \pi'_t\, \Gamma(t)\, (E-GG^+)\, A\pi_t\, dt + \frac{1}{2} \int_0^T \pi'_t\, \Gamma(t)\, GG^+ \,[2A + GG'\,\Gamma(t)]\, \pi_t\, dt \right\} .$$

The last term within brackets can be written :

$$\frac{1}{2} \int_0^T \pi'_t\, \Gamma(t)(E-GG^+)\, A\pi_t\, dt + \frac{1}{2} \int_0^T \pi'_t\, A'\, (E-GG^+)\, \Gamma(t)\, \pi_t\, dt$$

$$+ \frac{1}{2} \int_0^T \pi'_t\, \Gamma(t)\, GG^+ A\pi_t\, dt + \frac{1}{2} \int_0^T \pi'_t\, A'\, GG^+\, \Gamma(t)\, \pi_t\, dt$$

$$+ \frac{1}{2} \int_0^T \pi'_t\, \Gamma(t)\, GG'\, \Gamma(t)\, \pi_t\, dt$$

that is

$$\frac{1}{2} \int_0^T \pi'_t\, [A'\,\Gamma(t) + \Gamma(t)\, A + \Gamma(t)\, GG'\, \Gamma(t)]\, \pi_t\, dt$$

what completes the proof of the last assertion. ■

Now let us state the following corollary which will be useful in the maximum likelihood approach for parameter estimation :

II.1.2 Corollary - Let $\tilde{X}^x = (\tilde{X}^x_t \ ; \ t \geq 0)$ the solution process of the stochastic diffrential equation

$$\text{(II.2)} \quad d\tilde{X}^x_t = (E-GG^+)\, A\, \tilde{X}^x_t\, dt + G\, d\eta_t \ ; \ t \geq 0 \ ; \ \tilde{X}^x_0 = x$$

where $(\eta_t ; t \geq 0)$ is as in lemma I.1.1.Then $\mu^T_{X^x} << \mu^T_{\tilde{X}^x}$ with the following Radon-Nikodym derivative

$$\frac{d\mu^T_{X^x}}{d\mu^T_{\tilde{X}^x}} = \exp\left\{ Tr\ [(GG')^+ \left\{ \int_0^T d\pi_t\, \pi'_t\, A' - \frac{1}{2} A \int_0^T \pi_t\, \pi'_t\, dt\ A' \right\}] \right\}.$$

Proof - Applying the first part of lemma I.1.1 with $\Gamma(t) = -(GG')^+ A \ , \ t \geq 0$, one get

$$\frac{d\mu^T_{X^x}}{d\mu^T_{\tilde{X}^x}} = \exp\left\{ \int_0^T \pi'_t\, A'(GG')^+ (GG')(GG')^+\, d\pi_t - \right.$$

$$- \frac{1}{2}\int_0^T \pi'_t\, A'\, (GG')^+ (GG')\, (GG')^+\, A\pi_t\, dt -$$

$$\left. - \frac{1}{2} \int_0^T \pi'_t\, A(GG')^+ [E-(GG')(GG')^+]\, A\pi_t\, dt \right\}$$

$$= \exp\left\{ \int_0^T \pi'_t\, A'(GG')^+\, d\pi_t - \frac{1}{2}\int_0^T \pi'_t\, A'(GG')^+\, A\pi_t\, dt \right\}$$

what can be written as in the corollary. ∎

II.2. Quadratic forms of the process.

Since the quadratic integral $\int_0^T X^x_t \quad X^{x'}_t\, dt$ will play a central role in the following, now we investigate some its properties.

II.2.1 Lemma - Under assumption (H1) the matrix

$\int_0^T X_t^x \; X_t^{x'} dt$ is almost surely positive definite. If moreover (H2) is

satisfied then, $K_{A,B}$ being defined by (I.2), one has :

$$\lim_{T \to +\infty} \frac{1}{T} \int_0^T X_t^x \; X_t^{x'} dt = K_{A,B} \text{ almost surely .}$$

Proof - In order to prove the first assertion one has to show that the set

$$\Omega_0 = \bigcup_{h \in \mathbb{R}^n \setminus \{0\}} [h' (\int_0^T X_t^x \; X_t^{x'} dt) h = 0]$$

is negligible. It is clear that since X^x has continuous sample paths it is sufficient to prove that

$$\Omega_1 = \bigcup_{h \in \mathbb{R}^n \setminus \{0\}} \; \bigcap_{s \in [0,T]} [<X_s^x, h> = 0]$$

is negligible. Let $0 = t_0 < t_1 < \dots < t_n \leq T$; since

$$\Omega_1 \subset \Omega_1^* = \left\{ \det [X_{t_1}^x, \dots, X_{t_n}^x] = 0 \right\}$$

and the subset of $(\mathbb{R}^n)^n$

$$\left\{ (v_1, \dots, v_n) \in (\mathbb{R}^n)^n : \det [v_1, \dots, v_n] = 0 \right\}$$

has Lebesgue measure zero, the assertion in the lemma will be proved if one asserts that the gaussian random vector $(X_{t_1}^{x\prime}, \dots, X_{t_n}^{x\prime})'$ is non singular. But one can write

$$\begin{pmatrix} X^x_{t_1} \\ \vdots \\ X^x_{t_n} \end{pmatrix} = \begin{pmatrix} e^{At_1} & & 0 \\ & \ddots & \\ 0 & & e^{At_n} \end{pmatrix} \begin{pmatrix} E & & 0 \\ \vdots & \ddots & \\ E & \cdots & E \end{pmatrix} \begin{pmatrix} x + \int_0^{t_1} e^{-As} G dW_s \\ \int_{t_1}^{t_2} e^{-As} G dW_s \\ \int_{t_{n-1}}^{t_n} e^{-As} G dW_s \end{pmatrix},$$

where the random vectors $\int_{t_i}^{t_{i+1}} e^{-As} G dW_s$, $i = 0, \dots, n$ are independent with respective covariance matrices $\int_{t_i}^{t_{i+1}} e^{-As} GG' e^{-A's} ds$.

Since it has been noticed (cf. §I) that these matrices are positive definite when (H1) is satisfied, the proof of the first statement is completed. The second assertion is a simple consequence of the ergodicity of the diffusion $(X^x ; x \in \mathbb{R}^n)$ which has been stated (cf. §I) under (H1) - (H2) , since, for the invariant measure μ , one has

$$\int_{\mathbb{R}^n} yy' \, d\mu(y) = K_{A,B} = \int_0^{+\infty} e^{As} GG' e^{A's} ds \quad . \blacksquare$$

From the second part of lemma II.2.1. it follows that under (H1) - (H2) for every symmetric non negative matrix $S \neq 0$

$$\lim_{T \to +\infty} \int_0^T X_t^{x'} S X_t^x \, dt = +\infty \quad \text{a. s..}$$

Now we shall show that in order that this limit holds, assumption (H2) is not essential. We look at the Laplace transform of $\int_0^T X_t^x X_t^{x'} dt$ i.e the functional φ_x^T defined on the set $\mathcal{S}$ of symmetric non negative definite matrices S by

$$\text{(II.3)} \quad \varphi_x^T(S) = E\left\{ \exp\left[-\operatorname{Tr} S \int_0^T X_t^x X_t^{x'} dt\right] \right\}$$

$$= E\left\{ \exp\left[-\int_0^T X_t^{x'} S X_t^x \, dt\right] \right\} .$$

II.2.2 Lemma - Under assumption (H1) the functional φ_x^T is given by :

$$\varphi_x^T(S) = \exp\left\{ \frac{1}{2} \int_0^T \mathrm{Tr}\,(G'\gamma_t^{-1}(S)G)\,dt + \frac{1}{2}x'x \right\} \times$$

$$\times \exp\left\{ \frac{1}{2} x'\phi_T'(S)\,(E+\gamma_T^{-1}(S)\Delta_T(S))^{-1}\,\gamma_T^{-1}(S)\,\phi_T(S)x \right\} \times$$

$$\times \left[\det\left\{ E + \gamma_T^{-1}(S)\,\Delta_T(S) \right\} \right]^{-\frac{1}{2}} \qquad ; S \in \mathcal{S} ,$$

where $(\gamma_t(S)\ ;\ t \geq 0)$ is a positive definite matrix valued function which is defined by the equation

$$\text{(II.4)} \qquad \dot{\gamma}_t = A\gamma_t + \gamma_t A' + GG' - 2\gamma_t S \gamma_t \quad ; t \geq 0 \ ; \gamma_0 = E \quad ,$$

and $\Delta_t(S) = \phi_T(S) \int_0^T \phi_t^{-1}(S)\, GG'\, \phi_t'^{-1}(S)\,dt \quad \phi_T'(S)$

with $\dot{\phi}_t(S) = (A + GG'\gamma_t^{-1}(S))\,\phi_t(S) \ ; t \geq 0 \ ; \phi_0(S) = E$.

Proof - The existence of a unique solution to (II.4) is established in [10] (see also [14]) ; the fact that γ_t is invertible comes from (H1)(cf. [10]). It is easy to see that setting $\Gamma(t) = \gamma_t^{-1}(S)$ one has

$$\dot{\Gamma}(t) + A'\Gamma(t) + \Gamma(t)A + \Gamma(t)GG'\Gamma(t) - 2S = 0 \quad ; t \geq 0 \ ; \ \Gamma(0) = E \ .$$

Then, taking into account the second part of lemma II 1.1, one obtains

$$\varphi_x^T(S) = E\left\{ \exp\left[-\int_0^T Y_t^x S Y_t^{x'}\,dt\right] \quad \frac{d\mu_{X^x}^T}{d\mu_{Y^x}^T}(Y^x) \right\}$$

$$= E\left\{ \exp\left\{ -\int_0^T Y_t^x S Y_t^{x'}\,dt + \frac{1}{2}\int_0^T \mathrm{Tr}\,(G'\gamma_t^{-1}(S)G)\,dt + \frac{1}{2}x'x - \right.\right.$$

$$\left.\left. -\frac{1}{2} Y_T^{x'}\gamma_T^{-1}(S)\,Y_T^x + \int_0^T Y_t^x S Y_t^{x'}\,dt \right\} \right.$$

$$= \exp\left\{ \frac{1}{2} \int_0^T \mathrm{Tr}\,(G'\, \gamma_t^{-1}(S)\, G)\, dt + \frac{1}{2} x' x \right\} \times E \exp\left\{ - \frac{1}{2} Y_T^{x'}\, \gamma_T^{-1}(S)\, Y_T^{x} \right\} .$$

Moreover from (II.1) we know that Y_T^x is a gaussian random vector with mean

$$m_x^T(S) = \phi_T(S)\, x$$

and covariance matrix

$$\Delta_T(S) = \phi_T(S) \int_0^T \phi_t^{-1}(S)\, GG'\, \phi_t'^{-1}(S)\, dt\, \phi_T'(S)$$

with $(\phi_t(S) \; ; t \geq 0)$ as in the statement. Then (cf. Lemma11.6 in russian version of [10]) one gets

$$E(\exp\{-\frac{1}{2} Y_T^{x'}\, \gamma_T^{-1}(S)\, Y_T^x\})$$

$$= \exp\left\{ -\frac{1}{2} m_x^{T'}(S)\, (E + \gamma_T^{-1}(S)\, \Delta_T(S))^{-1}\, \gamma_T^{-1}(S)\, m_x^T(S) \right\} \times$$

$$\times \left\{ \det\,[E + \gamma_T^{-1}(S)\, \Delta_T(S)] \right\}^{-\frac{1}{2}}$$

and the proof is completed . ∎

Now we are able to prove the following statement :

II.2.3 Corollary - If (H1) is satisfied, then, for every $x \in \mathbb{R}^n$ and $S \in \mathcal{S}$, $S \neq 0$, there exist strictly positive constants $\alpha_x(S)$ and $\beta(S)$ such that

$$\varphi_x^T(S) \leq \alpha_x(S) \cdot e^{-\beta(S)T} , \quad T \geq 0 .$$

Proof. First, since $\gamma_T^{-1}(S) \in \mathcal{S}$ and $\Delta_T(S) \in \mathcal{S}$, it is clear that

$$\exp\left\{\frac{1}{2} x' x\right\} \times \exp\left\{-\frac{1}{2} x' \phi'_T(S) (E + \gamma_T^{-1}(S) \Delta_t(S))^{-1} \gamma_T^{-1}(S) \phi_T(S) x\right\} \leq$$

$$\leq \exp\left\{\frac{1}{2} x' x\right\} .$$

So we only have to prove that

$$\text{(II.5)} \quad \exp\left\{\frac{1}{2} \int_0^T \operatorname{Tr} (G' \gamma_t^{-1}(S) G)\, dt\right\} \times \left(\det [E + \gamma_T^{-1}(S) \Delta_T(S)]\right)^{-\frac{1}{2}} \leq$$

$$\leq K(S)\, e^{-\beta(S)T}$$

for some positive constants $K(S)$ and $\beta(S)$.

But, setting

$$\Lambda_T(S) = \int_0^T \phi_t^{-1}(S)\, GG'\, \phi_t'^{-1}(S)\, dt,$$

we have

$$\det \, [E + \gamma_T^{-1}(S) \Delta_T(S)] =$$

$$= \det [E + \gamma_T^{-1}(S)\, \phi_T(S)\, \Lambda_T(S) \phi'_T(S)] \geq$$

$$\geq 1 + \det [\gamma_T^{-1}(S)\, \phi_T(S) \Lambda_T(S) \phi'_T(S)] =$$

$$= [\det \phi_T(S)]^2 \, \frac{\det \Lambda_T(S) + \dfrac{\det \gamma_T(S)}{(\det \phi_T(S))^2}}{\det \gamma_T(S)} .$$

Then the first member of (II.5) is bounded by

$$\exp\left\{\frac{1}{2}\int_0^T \operatorname{Tr}(G'\gamma_t^{-1}(S)G)\,dt\right\} \times \left(\det \phi_T(S)\right)^{-1} \times$$

$$(\text{II}.6) \qquad \times\left[\frac{\det \gamma_T(S)}{\det \Lambda_T(S) + \dfrac{\det \gamma_T(S)}{(\det \phi_T(S))^2}}\right]^{\frac{1}{2}} .$$

But, since

$$\det \phi_T(S) = \exp\left\{\int_0^T \operatorname{Tr}(A + GG'\gamma_t^{-1}(S)\,dt\right\} ,$$

one gets for the product of the first two terms in (II.6);

$$\exp\left\{-\int_0^T \operatorname{Tr}(A + GG'\gamma_t^{-1}(S))\,dt\right\}$$

or, by use of (II.4) ,

$$\exp\left\{-\int_0^T \operatorname{Tr}\gamma_t(S)\,S\,dt - \frac{1}{2}\int_0^T \operatorname{Tr}\dot{\gamma}_t(S)\gamma_t^{-1}(S)\,dt\right\}$$

that is, taking into account the fact that

$$\frac{d}{dt}(\gamma_t^{-1}(S)) = -\gamma_t^{-1}(S)\ \dot{\gamma}_t(S)\ \gamma_t^{-1}(S) \ ; \ \gamma_0^{-1}(S) = E ,$$

the expression

$$\exp\left\{-\int_0^T \operatorname{Tr}\gamma_t(S)\,S\,dt\right\}[\det \gamma_t(S)]^{-\frac{1}{2}} .$$

Now (II.6) can be written

$$\left[\det \Lambda_T(S) + \frac{\det \gamma_T(S)}{(\det \phi_T(S))^2} \right]^{-\frac{1}{2}} \times \exp \left\{ - \int_0^T \mathrm{Tr}\, \gamma_t(S)\; S\, dt \right\}$$

and in order to get (II.5) it remains to prove that

$$\operatorname*{Inf}_{t \geq 0} \left\{ \det \Lambda_t(S) + \frac{\det \gamma_t(S)}{(\det \phi_t(S))^2} \right\} > 0$$

and

$$\operatorname*{Inf}_{t \geq 0} \mathrm{Tr} \left| \gamma_t(S)\; S \right| > 0 \;.$$

The first fact follows from that $\det \Lambda_t(S)$ increases with t, is zero if and only if $t = 0$ and that $\dfrac{\det \gamma_t(S)}{(\det \phi_t(S))^2}$ is positive with value 1 for $t = 0$.

In order to prove the second point let us consider $(\sigma_t(S)\;,\; t \geq 0)$ the solution of

$$\dot{\sigma}_t = A\sigma_t + \sigma_t A' + GG' - 2\sigma_t S \sigma_t \;\; ; \;\; t \geq 0 \;\; ; \;\; \sigma_0 = 0 \;.$$

It is known (cf. [17]) that, since (H1) is satisfied, $\sigma_t(S)$ is monotone non decreasing and moreover, for every $t > 0$, $\sigma_t(S)$ is positive definite. It is easy to see that, setting $\delta_t(S) = \gamma_{T-t}(S) - \sigma_{T-t}(S)$, one has

$\delta_T(S) = E$ and for $t \geq 0$:

$$\dot{\delta}_t(S) + (A - 2\sigma_{T-t}(S)\, S)\, \delta_t(S) + \delta_t(S)\,(A - 2\sigma_{T-t}(S) S)' - 2\delta_t(S)\, S \delta_t(S) = 0$$

So (cf. [17]), for every $t \in [0,T]$, $\delta_t(S)$ is non negative definite and therefore for every $t \geq 0$ the matrix $\gamma_t(S) - \sigma_t(S)$ is also non negative definite. All these properties lead to what is needed. ∎

III-Maximum likelihood estimation of the drift parameter

Here we investigate the problem of estimating the unknown parameters A and $B = GG'$ of the diffusion process under consideration in view of the observation on $[0,T]$ of one trajectory starting from zero at time zero since in view of the preliminary results this is not really a restriction. So we, start with a process $X = (X_t \;; t \geq 0)$ satisfying (I.1) with $x=0$.

An estimation of matrix B can be obtained by use of the quadratic variation $[X]$ of process X.Precisely B can be computed with probability one on every finite time interval $[0,T]$ by $B=\frac{1}{T}[X]_T$.So we can consider the problem of estimating the matrix A when B is assume known without any restriction,B having eventually been previously computed.First we investigate the case when the matrix A is completely unknown(for which the results have been announced in[9])and ther the case when it is known up to a multiplicative constant.

III.1-The case when the drift parameter is completely unknown.

Since, from (I.1), one has

$$\int_0^T dX_t \; X_t' = A \int_0^T X_t X_t' \, dt + G \int_0^T dW_t \; X_t'$$

one also has

$$\int_0^T \left\{ d(E - BB^+) X_t \right\} X_t' = (E - BB^+) A \int_0^T X_t X_t' \, dt$$

because $BB^+ = GG^+$ is nothing but the matrix of the orthogonal projection on the subspace of $\mathbb{R}^n$ generated by the columns of G . Then, by use of Lemma II.2.1, under (H1) one gets

(III.1) $(E - BB^{+})A = \left\{ \int_0^T [d(E - BB^{+})X_t X_t'] \right\} \left\{ \int_0^T X_t X_t' \, dt \right\}^{-1}$.

Let us notice that the stochastic integral in the second member of (III.1) is in fact an ordinary integral since the process $((E - BB^{+}) X_t \; ; \; t \geq 0)$ has locally bounded variation. Finally, as B , the matrix $(E - BB^{+}) A$ can be computed with probability one on some finite time interval ; this allows us to assume that this matrix is known too.

Now, from Corollary II.1.2, the measure $\mu^T_{\tilde{X}}$, which only depends on known matrices B and $(E - BB^{+}) A$ can be considered as a dominating measure for the statistical space associated with the concerned estimation problem ; the log-likelihood function can be written in the following form

$$Tr \left\{ B^{+} [\int_0^T dX_t X_t' \;\; A' - \frac{1}{2} A \int_0^T X_t X_t' \, dt \;\; A'] \right\} .$$

Then the next result is an immediate consequence of Lemma II.2.1 :

III.1.1 Proposition - Under (H1) a maximum likelihood estimator of the matrix A is given by

(III.2) $\hat{A}_T = [\int_0^T dX_t X_t'] [\int_0^T X_t X_t' dt]^{-1}$.

III.1.2 Remark - From (I.1) one can write

(III.3) $\hat{A}_T = A + G [\int_0^T dW_t \; X_t'] [\int_0^T X_t X_t' \, dt]^{-1}$.

This ensures, because of the controllability of the pair [A,G] that the pair $[\hat{A}_T , G]$ is itself controllable ; so the estimator $[\hat{A}_T \, , \, (\frac{1}{T} [X]_T)^{\frac{1}{2}}]$ takes its values in the parameter space. Moreover let us notice that the matrix $(E - BB^{+})\hat{A}_T$ provides again the matrix $(E - BB^{+})A$ defined by (III.1).

Now we state the asymptotic properties of the estimator:

III.1.3. Proposition. Under (H1) and (H2) the estimator $\hat{A}_T$ defined by (III.2) is strongly consistent i.e.

$$\lim_{T\to+\infty} \hat{A}_T = A \qquad \text{a.s..}$$

Moreover the random vector $T^{\frac{1}{2}}$ vec $(\hat{A}_T - A)$ is asymptotically normally distributed with mean zero and covariance matrix $K_{A,B}^{-1} \otimes B$ where $K_{A,B}$ is given by (I.2).

Proof. We start from the decomposition of the estimator given in equation (III.3). By lemma II.2.1 one gets that

$$\lim_{T\to+\infty} \left[\frac{1}{T}\int_0^T X_t X_t' \, dt\right]^{-1} = K_{A,B}^{-1} \qquad \text{a.s..}$$

Moreover, since for every coordinate $(X_t^i \,;\, t \geq 0)$ of the observed process X one ha

$$\lim_{T\to+\infty} \frac{1}{T}\int_0^T (X_t^i)^2 \, dt = K_{A,B}^{i,i} > 0 \quad \text{a.s.},$$

by Theorem 4.1 [4], one obtains

$$\lim_{T\to+\infty} \frac{1}{T}\int_0^T dW_t \, X_t' = 0 \text{ a.s.}$$

Then the first assertion is proved.

Now, still from (III.3), one can write

$$\sqrt{T}\,\text{vec}\,(\hat{A}_T - A) = \left\{\left[\frac{1}{T}\int_0^T X_t X_t' \, dt\right]^{-1} \otimes G\right\} \text{vec}\left\{\frac{1}{\sqrt{T}}\int_0^T dW_t \, X_t'\right\}$$

where

$$\operatorname{vec}\left\{\frac{1}{\sqrt{T}}\int_0^T dW_t\, X_t'\right\} = \frac{1}{\sqrt{T}}\int_0^T (X_t \otimes E)\, dW_t \ .$$

From Lemma II.2.1 and

$$\int_0^T (X_t \otimes E)\,(X_t \otimes E)'\, dt$$

$$= \int_0^T (X_t \otimes E)\,(X_t' \otimes E)\, dt = \int_0^T (X_t X_t') \otimes E\, dt$$

one gets

$$\lim_{T\to+\infty} \frac{1}{T}\int_0^T (X_t \otimes E)\,(X_t \otimes E)'\, dt = K_{A,B} \otimes E \ .$$

Then, by use of the results of [16] concerning the asymptotic normality of stochastic integrals, one deduces that

$$\lim_{T\to+\infty} \frac{1}{\sqrt{T}} \operatorname{vec}\left(\frac{1}{\sqrt{T}}\int_0^T dW_t\, X_t'\right) = N(0, K_{A,B} \otimes E)$$

where the limit stands in the sense of convergence in probability distribution.

It follows that in the same sense $\sqrt{T}\operatorname{vec}(\hat{A}_T - A)$ converges to the gaussian distribution with mean zero and covariance matrix

$$(K_{A,B}^{-1} \otimes G)\,(K_{A,B} \otimes E)\,(K_{A,B}^{-1} \otimes E)\,(K_{A,B}^{-1} \otimes G)'$$

$$= \left((K_{A,B}^{-1} K_{A,B}) \otimes (G\ E)\right)(K_{A,B}^{-1} \otimes G')$$

$$= (E \otimes G)\,(K_{A,B}^{-1} \otimes G') = K_{A,B}^{-1} \otimes (GG')$$

$$= K_{A,B}^{-1} \otimes B \ . \blacksquare$$

III.1.4 Remark. Note that if one assumes that the parameter space is that of pairs [A,B] such that on one hand [A,G] is controllable and on the other hand A is a stable matrix, then the estimator $\hat{A}_T$ defined by (III.2) is not a maximum likelihood estimator ; in fact under these conditions such an estimator does not exist. The matrix $\hat{A}_T$ is not stable in general but the probability that it is goes to one when T increases to infinity.
In the case when B is non singular it is possible to modify $\hat{A}_T$ in order to obtain an estimator of A which is stable: one has

$$\int_0^T dX_t X'_t + \int_0^T X_t \, dX'_t = X_T X'_T - [X]_T$$

so

$$\left(\int_0^T dX_t X'_t - \frac{1}{2} X_T X'_T\right)\left(\int_0^T X_t X'_t \, dt\right)^{-1} \frac{1}{T}\int_0^T X_t X'_t \, dt +$$
$$+ \frac{1}{T}\int_0^T X_t X'_t \, dt \left(\int_0^T X_t X'_t \, dt\right)^{-1}\left(\int_0^T X_t \, dX'_t - \frac{1}{2} X_T X'_T\right) = - B \, .$$

This shows that the Lyapunov equation

$$\tilde{A}_T Q + Q\tilde{A}'_T = - B$$

where

$$\tilde{A}_T = \left(\int_0^T dX_t X'_t - \frac{1}{2} X_T X'_T\right)\left(\int_0^T X_t X'_t \, dt\right)^{-1} ,$$

admits the positive definite matrix $\frac{1}{T}\int_0^T X_t X'_t \, dt$ as a solution, what implies that $\tilde{A}_T$ is stable (see [15]) . Moreover it is clear that the estimator $\tilde{A}_T$ has the same asymptotic properties as $\hat{A}_T$.

So in some sense, under asumptions (H1) - (H2) the problem of parameter estimation is completely solved. The question is : does the maximum likelihood estimator given by (III.2) still converge if one drops assumption (H2) ?We have no answer for this in general but the problem has been positively solved in the one dimensional case (cf. [10]) .

We are now going to look at the problem of parameter estimation in the case when the drift coeeficient is known up to an unknown multiplicative constant ; the results will include those cited before concerning the one-dimensional case.

III.2 - The case when the drift matrix is known up to multiplicative constant.

Here we assume that A belongs to the set $\{\theta A_o ; \theta \in \mathbb{R}\}$ where A_o is some known $n\times n$ matrix and θ is an unknown parameter that one has to estimate. Since if $BB^+A_o = 0$ then θ can be computed from (III.1), we shall also assume that $BB^+A_o \neq 0$.

Because the Log-likelihood function is equal to

$$\theta \int_o^T X_t' A_o' B^+ d X_t - \frac{1}{2} \theta^2 \int_o^T X_t' A_o' B^+ A_o X_t dt,$$

the maximum likelihood estimator of θ is given by

$$\text{(III.4)} \qquad \hat{\theta}_T = \frac{\int_o^T X_t' A_o' B^+ dX_t}{\int_o^T X_t' A_o' B^+ A_o X_t dt} .$$

III.2.1 - Proposition : Under (H 1) the estimator $\hat{\theta}_T$ defined by (III.4) is strongly consistent i.e.

$$\lim_{T\to\infty} \hat{\theta}_T = \theta \quad \text{a.s.},$$

Proof : As before, from (I.1) one can write

$$\hat{\theta}_T = \theta + \frac{\int_o^T X_t' A_o' B^+ G dW_t}{\int_o^T X_t' A_o' B^+ A_o X_t dt} .$$

But $(\int_o^t X_s' A_o' B^+ A_o X_s ds,\ t \geq 0)$ is the quadratic variation process of the martingale $(\int_o^t X_s' A_o' B^+ G dW_s,\ t \geq 0)$ and

$$\lim_{T\to\infty} \int_0^T X_t' A_0' B^+ A_0 X_t \, dt = \infty \quad \text{a.s.}$$

(see Corollary II.2.3). So, by the analogue for continuous time martingales of the strong law of large numbers (see [4], Theorem 4.1, p. 394) the assertion follows. ∎

III.2.2 - Remark - Note that

$$\int_0^T X_t' A_0' B^+ dX_t = \frac{1}{2} (X_T' B^+ A_0 X_T - T \operatorname{tr} BB^+ A_0) +$$

$$+ \frac{1}{4} \operatorname{tr} (A_0' B^+ - B^+ A_0) \left(\int_0^T dX_t X_t' - \int_0^T X_t dX_t'\right)$$

So, if B^+A_0 is symmetric, then

$$\hat{\theta}_T = \frac{X_T' B^+ A_0 X_T - T \operatorname{tr} BB^+ A_0}{2\int_0^T X_t' A_0' B^+ A_0 X_t dt} .$$

Moreover, if B^+A_0 is skew symmetric, then

$$\hat{\theta}_T = \frac{\operatorname{tr} A' B^+ \left(\int_0^T dX_t X_t' - \int_0^T X_t dX_t'\right)}{2\int_0^T X_t' A_0' B^+ A_0 X_t dt} .$$

III.2.3 - Remark - The results of this section (under H 1) may be easily generalized to the following case. Let $A = \sum_{i=1}^{p} \theta_i A_i$, where A_i, $i = 1,\ldots,p$ are known matrices such that $BB^+A_i \neq 0$, $i = 1,\ldots,p$ and

$$A_i' B^+ A_j + A_j' B^+ A_i = 0 \; ; \; i \neq j .$$

In this case, the maximum likelihood estimator of θ_i, $i = 1,\ldots,p$ is given by

$$\hat{\theta}_{i,T} = \frac{\int_0^T X_t' A_i' B^+ dX_t}{\int_0^T X_t' A_i' B^+ A_i X_t dt}$$

and it is clear that for all $i = 1,\ldots,p$

$$\lim_{T\to\infty} \hat{\theta}_{i,T} = \theta_i \qquad \text{a.s.}.$$

Now we give an example of 2-dimensional process which is a model for so called geophysical problem.

III.2.4 - Example - It is known (see [2], [13]) that the instantaneous axis of rotation of the earth is displaced with respect to the minor axis of the terrestrial ellipsoid. This displacement consists of a periodic part and a fluctuating part. The latter can be assumed to be a solution of the system of stochastic differential equations of the form

$$dX_1(t) = \theta_1 X_1(t)dt - \theta_2 X_2(t)dt + g dw_1(t),$$

$$dX_2(t) = \theta_2 X_1(t)dt + \theta_1 X_2(t)dt + g dw_2(t),$$

where $(w_1(t)\ ;\ t \geq 0)$ and $(w_2(t)\ ;\ t \geq 0)$ are two independent Wiener processes, θ_1 and θ_2 are unknown and g^2 is known. It is clear that the above system of equations can be written in the form (I.1) with

$$A = \begin{pmatrix} \theta_1 & -\theta_2 \\ \theta_2 & \theta_1 \end{pmatrix}, \quad G = g\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Moreover, because $A = \theta_1 A_1 + \theta_2 A_2$, where

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

and $A_1^{*}A_2 + A_2^{*}A_1 = 0$ the maximum likelihood estimators of θ_1 and θ_2 are given by

$$\hat{\theta}_{1,T} = \frac{\int_0^T X_1(t)\,dX_1(t) + \int_0^T X_2(t)\,dX_2(t)}{\int_0^T [X_1^2(t)+X_2^2(t)]dt},$$

$$\hat{\theta}_{2,T} = \frac{\int_0^T X_1(t)\,dX_2(t) - \int_0^T X_2(t)\,dX_1(t)}{\int_0^T [X_1^2(t)+X_2^2(t)]dt}.$$

It has been shown (see [4]) that these estimators are consistent when the solution process

$$X_t = \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix}$$

is assumed to be stationary. Taking into account Remark III.2.3 we can assert that these estimators are still strongly consistent when one drops the assumption of stationarity.

IV - SEQUENTIAL ESTIMATION

As in section III.2 we assume that A belongs to the set $\{\theta A_o ; \theta \in \mathbb{R}\}$, where A_o is some known $n\times n$ matrix such that $BB^+A_o \neq 0$. Here we deal with the problem of sequential estimation of unknown parameter θ. Mathematical statement of the problem and some details concerning the case of continuous time observations may be found for example in [10], [12].

Let H be a non-negative number. Define the stopping time

$$\text{(IV.1)} \qquad \tau(H) = \inf\{t : \int_0^t X_s' A_o' B^+ A_o X_s \, ds = H\}.$$

In the case of a one dimensional process X it has been shown (cf. [10]) that $E\tau^n(H) < \infty$ for all n. We shall prove the following stronger result.

IV.1 - Lemma - Under (H1) there exists $\delta > 0$ such that

$$E \exp\{\delta \tau(H)\} < \infty .$$

Proof : From Corollary II.2.3 it follows that there exist positive constants α and β such that

$$E \exp \{- \int_0^T X_t' A_o' B^+ A_o X_t \, dt\} \le \alpha e^{-\beta T} .$$

So

$$P(\tau(H) \ge T) = P\left(\int_0^T X_t' A_o' B^+ A_o X_t \, dt \le H\right) \le$$

$$\le e^H E \exp\{- \int_0^T X_t' A_o' B^+ A_o X_t \, dt\} \le \alpha e^H e^{-\beta T} .$$

Now let $0 < \delta < \beta$. Because

$$e^{\delta T} P(\tau(H) \ge T) \le e^H \alpha e^{-(\beta-\delta)T}$$

we have also

$$E \exp\{\delta\tau(H) = \delta \int_0^\infty e^{\delta t} P(\tau(H) \ge t) \, dt$$

$$\le \delta\alpha e^H \int_0^\infty e^{-(\beta-\delta) t} dt = \frac{\delta\alpha e^H}{\beta-\delta} < \infty$$

what completes the proof. ∎

Define the sequential plan

$$(\tau(H), \hat{\theta}_{\tau(H)}),$$

where $\tau(H)$ is given by (IV.1) and

$$\hat{\theta}_{\tau(H)} = \frac{1}{H} \int_0^{\tau(H)} X_t' A_o' B^+ \, d X_t$$

Note that $\hat{\theta}_{\tau(H)}$ is in fact the maximum likelihood estimator of θ based on the observation of the process X on the random time interval $[0,\tau(H)]$. It is also easy to see that the case studied here covers those considered in [10]. We shall prove now :

IV.2 - Proposition - Under (H 1) the sequential plan $(\tau(H), \hat{\theta}_{\tau(H)})$ has the following properties for all $\theta \in \mathbb{R}$:

(i) $\hat{\theta}_{\tau(H)}$ is normally distributed with $E\hat{\theta}_{\tau(H)} = \theta$ and var $\hat{\theta}_{\tau(H)} = \frac{1}{H}$;

(ii) in the class of sequential plans $(\tau, \tilde{\theta})$ such that :

$$E\,\tilde{\theta}^2 < \infty \quad ; \quad E \int_0^\tau X_t' A' B^+ A_o X_t \, dt \le H$$

the sequential plan $(\tau(H), \hat{\theta}_{\tau(H)})$ is admissible and minimax, with respect to the quadratic loss function.

Proof : Part (i) follows from equality

$$\hat{\theta}_{\tau(H)} = \theta + \frac{1}{H} \int_0^{\tau(H)} X_t' A_o' B^+ G \, dW_t$$

and the fact that $(\int_0^{\tau(H)} X_t' A_o' B^+ G \, d\, W_t,\ H \ge 0)$ is a Wiener process.

To prove admissibility note that for all sequential plans $(\tau, \tilde{\theta})$ such that $E\,\tilde{\theta}^2 < \infty$ and $E \int_0^\tau X_t' A_o' B^+ A_o X_t \, dt \le H$:

$$E(\tilde{\theta}-\theta)^2 \ge \frac{(1+b'(\theta))^2}{H} + b^2(\theta),$$

where $b(\theta) = E(\tilde{\theta}-\theta)$. This is a simple generalization of the Cramer-Rao inequality (see [10]). Next, suppose $(\tau(H), \hat{\theta}_{\tau(H)})$ inadmissible and show that $b(\theta) \equiv 0$ is the only function satisfying

$$(1+b'(\theta))^2 + H\, b^2(\theta) \le 1.$$

This leads to admissibility. Moreover, $(\tau(H), \hat{\theta}_{\tau(H)})$ has constant risk so it is also minimax. ∎

IV.3 - Remark - Assume (H 1) and, as in Remark III.2.3, that $A = \sum_{i=1}^{p} \theta_i A_i$, where A_i are known matrices such that

$BB^+A_i \neq 0$ and $A_i'B^+A_j + A_j'B^+A_i = 0 \quad i \neq j \ ; \ i,j = 1,\dots,p.$
It is clear that one may use, in order to estimate $\theta_i \ ; \ i = 1,\dots,p,$ the sequential plans

$$(\tau_i(H), \hat{\theta}_{i,\tau_i(H)}),$$

where

$$\tau_i(H) = \inf\{t : \int_0^t X_s' A_i' B^+ A_i X_s \, ds = H\}$$

and

$$\hat{\theta}_{i,\tau_i(H)} = \frac{1}{H} \int_0^{\tau_i(H)} X_t' A_i' B^+ \, d X_t .$$

The estimator $\hat{\theta}_{i,\tau_i(H)}$ is normally distributed with $E\hat{\theta}_{i,\tau_i(H)} = \theta_i$ and $\mathrm{Var}\ \hat{\theta}_{i,\tau_i(H)} = \frac{1}{H}$(*). Moreover, by the Cramer-Rao inequality, we have that the sequential plan $(\tau_i(H), \hat{\theta}_{i,\tau_i(H)})$ is minimum variance unbiased in the class of unbiased sequential plans $(\tau_i, \tilde{\theta}_i)$ such that

$$E\ \tilde{\theta}_i^2 < \infty, \quad E \int_0^{\tau_i} X_t' A_i' B^+ A_i X_t \, dt \leq H.$$

By the same argument as in Lemma IV.1, one may also show that for all $i = 1,\dots,p$ there exists a constant $\delta_i > 0$ such that

$$E \exp\{\delta_i \tau_i(H)\} < \infty.$$

Finally note that in the case of Example III.2.4, the above sequential plans are nothing but

$$(\tau(H), \hat{\theta}_{i,\tau(H)})$$

where

$$\tau(H) = \inf\{t : \int_0^t [X_1^2(s) + X_2^2(s)]ds = H\}$$

(*) In fact $\hat{\theta}_H = (\hat{\theta}_{1,\tau_1(H)},\dots,\hat{\theta}_{p,\tau_p(H)})'$ is normally distributed with $E\hat{\theta}_H = \theta$ and $\mathrm{Var}\ \hat{\theta}_H = \frac{1}{H} E$ and then $\hat{\theta}_{1,\tau_1(H)},\dots,\hat{\theta}_{p,\tau_p(H)}$ are independent.

and

$$\hat{\theta}_{1,\tau(H)} = \frac{1}{H}\left[\int_0^{\tau(H)} X_1(s)\,dX_1(s) + \int_0^{\tau(H)} X_2(s)\,dX_2(s)\right]$$

$$\hat{\theta}_{2,\tau(H)} = \frac{1}{H}\left[\int_0^{\tau(H)} X_1(s)\,dX_2(s) - \int_0^{\tau(H)} X_2(s)\,dX_2(s)\right]$$

which have been derived in [13].

For some related questions one can also consult [5].

REFERENCES

[1] M. ARATO ; On the parameter estimation of processes satisfying a stochastic differential equation , (in Russian),Studia Sci. Math. Hungar. ,5 (1970) ,11-15 .

[2] M. ARATO ; On the statistical examination of continuous state Markov processes, I, II, III, IV, Selected Transl. in Math. Statist. and Probability, 4(1978).

[3] L. ARNOLD and W. KLIEMANN ; Qualitative theory of stochastic systems, Preprint form : A.T. Bharudra-Reid (ed.) : Probabilistic Analysis and Related Topics, Vol. 3, Academic Press, New-York (1981).

[4] I.V. BASAWA and B.L.S. PRAKASA RAO ; Statistical Inference for Stochastic Processes, Academic Press, London (1980).

[5] S.E. BOROBEICIKOV and W.W. KONEV ; Construction of sequential plans for parameters of recurrent type processes ; Mathematical Statistics and Applications, Publ. of Tomsk Univ., Vol. 6, (1980), 72-81.

[6] M.CHALEYAT-MAUREL et L. ELIE ; Diffusions gaussiennes ; S.M.F., Astérisque, 84-85, (1981), 255-279.

[7] K. ICHIHARA and H. KUNITA ; A Classification of the Second Order Degenerate Elliptic Operators and its Probabilistic Characterization, Z. Wahrch. verw. Gebiete 30(1974), 235-254.

[8] A. LE BRETON, Parameter estimation in a vector linear stochastic differential equation. In "Transactions of the Seventh Prague Conference on Information Theory. Statistical Decision functions and Random Processes", Vol. A (1977), 353-366.

[9] A. LE BRETON et M. MUSIELA ; Estimation des paramètres pour les diffusions homogènes hypoelliptiques, C.R.Acad.Sci.Paris Ser. I, t. 294 (1982),341-344.

[10] R.S. LIPTSER and A.N. SHIRYAYEV ; Statistics of Random Processes, I, II, Springer-Verlag, New-York (1978).

[11] G. MARUYAMA and H. TANAKA ; Ergodic properties of n-dimensional recurrent Markov processes, Memoirs Fac. Sc., Kyusyu Univ. Ser. A, Vol. 13, n° 2, (1959), 157-172.

[12] M. MUSIELA ; On sequential estimation of parameters of continuous Gaussian Markov processes. Probability and Mathematical Statistics (1982).

[13] A.A. NOVIKOV ; Sequenital estimation of the parameters of diffusion processes. Theor. Prob. Appl., Vol. 16, n°2, (1971), 394-396.

[14] J. RODRIGEZ-CANABAL; The Geometry of the Riccati Equation, Stochastics, Vol. 1, (1973), 129-149.

[15] J. SNYDERS and M.ZAKAI; On nonnegative solutions of the equation AD+DA' = -C, SIAM J. Appl. Math. Vol. 18, N°3, (1970), 704-714.

[16] A.F. TARASKIN ; Some limit theorems for stochastic integrals in theory of Stochastic Processes, Vol. 1, John Wiley, (1974), 136-151.

[17] W.M. WONHAM ; On a matrix Riccati equation of stochastic control, SIAM J. Control, Vol. 6, n°4, (1968), 681-697.

[18] J. ZABCZYK ; Controllability of stochastic linear systems ; Systems and Control Letters, Vol. 1, n°1, (1981), 25-31.

[19] M. ZAKAI and J. SNYDERS ; Stationary probabilistic measures for linear differential equations driven by white noise, J. Differential Equations, 8(1970), 27-33.

On sequential estimation of parameters of continuous Gaussian Markov processes

by

Marek Musiela (Wrocław)

1. Introduction

Recently a number of authors have studied various estimators of parameters of stochastic processes and nonasymptotic optimal properties of such estimators. In particular Arato [1] and Hajek [7] have investigated nonsequential minimum variance unbiased estimators for parameters of Gaussian processes. Novikov [8] has compared sequential and nonsequential methods of estimation for a shift parameter of a diffusion Gaussian process. Dvoretzky, Kiefer and Wolfowitz [5] have shown that for the Poisson process, the negative-binomial process, the gamma process and the Wiener process fixed time sequential plans are minimax under the weighted quadratic loss function. Magiera [14] has extended these results of Dvoretzky, Kiefer and Wolfowitz to a class of processes which contains all the processes considered by these authors.

In this paper we consider continuous Gaussian Markov process $y = (y(t),\ t \geq 0)$ with mean $m(t)$ and covariance $K(t, s)$ and we assume that

$$m(t) = \Theta \varphi(t) + \psi(t),$$

where $\varphi(t)$ and $\psi(t)$ are known, while Θ is unknown. We deal with the problem of sequential estimation of Θ when $K(t, s)$ is known. Also, postulating

$$K(t, s) = \exp\left\{\int_s^t (\alpha p(u) + q(u))\, du\right\} K(s, s),$$

where $p(t)$ and $q(t)$ are known we estimate α. Comparing the sequential plans the usual quadratic loss function and the quadratic loss function plus the cost function will be used. Admissible, minimax and minimum variance unbiased sequential plans for estimation of Θ and α will be given.

2. Absolute continuity of measures

Throughout the paper we assume that the derivatives $m'(t)$ and $K'(t, t)$ exist for all $0 \leq t < \infty$. Moreover, we assume that

$$K_1(t) = \lim_{h \downarrow 0} \frac{K(t+h, t) - K(t, t)}{h}$$

exists for all $0 \leq t < \infty$.

Let

$$A(t) = K_1(t) K^+(t, t),$$

$$B(t) = K'(t, t) - 2K_1(t),$$

$$a(t) = m'(t) - A(t) m(t),$$

where $K^+ = K^{-1}$ for $K \neq 0$ and $K^+ = 0$ for $K = 0$. Assume that

$$\int_0^t (|a(u)| + |A(u)| + B(u))\, du < \infty$$

for all $0 \leq t < \infty$. Let $\{F_t\}$ be the family of the σ-fields generated by random variables $\{y(s): s \leq t\}$. Under the above assumptions [17] there exists a Wiener process $w = (w(t), F_t)$ such that

$$y(t) = y(0) + \int_0^t (a(u) + A(u)y(u))du + \int_0^t B^{1/2}(u)dw(u) . \qquad (2.1)$$

This implies that process y is a semimartingale with a Gaussian martingale component. This fact is useful in the absolute continuity considerations below.

Let $C[0, t]$ denote the space of all continuous functions $c: [0, t] \longrightarrow (-\infty, \infty)$ and let $\mathcal{B}_t$ denote the σ-field of Borel subsets of $C[0, t]$ endowed with the norm topology. Let C be the space of all continuous functions $C: [0, \infty) \longrightarrow (-\infty, \infty)$ and let $\mathcal{B}$ denote the σ-field of Borel subsets of C relative to the topology of uniform convergence on compact subsets.

A function $\tau : C \longrightarrow [0, \infty]$ is said to be a stopping time if

$$\left\{ c: \tau(c) \leq t \right\} \in \mathcal{B}_t$$

for every $t \geqslant 0$.

Now we define a new Gaussian process

$$V^x(t) = x + \int_0^t B^{1/2}(u)\, dw(u), \quad x \in (-\infty, \infty).$$

Let μ_y and μ_{V^x} be the measures induced by y and V^x, respectively, i.e. $\mu_y(B) = P(y \in B)$ and $\mu_{V^x}(B) = P(V^x \in B)$, for $B \in \mathcal{B}$. Moreover, let $\mu_{\tau, y}$ and μ_{τ, V^x} denote the restrictions of μ_y and μ_{V^x} to the σ-field

$$\mathcal{B}_\tau = \left\{ B \in \mathcal{B} : \{\tau \leq t\} \cap B \in \mathcal{B}_t \right\},$$

respectively. With this notation $\mu_{\infty, y} = \mu_y$. Now let $\nu(\cdot)$ be a measure defined on the σ-field $\mathcal{B}$ by

$$\nu(B) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mu_{V^x}(B) \exp\left\{-\frac{1}{2}x^2\right\} dx.$$

The restriction ν_τ of ν to σ-field $\mathcal{B}_\tau$ is given by

$$\nu_\tau(B) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mu_{\tau, V^x}(B) \exp\left\{-\frac{1}{2}x^2\right\} dx.$$

Using some results of Kabanov, Lipcer, Shiryayev (cf [9] Section 10, [12] Chapter 7) concerning conditions for absolute continuity and singularity of measures induced by semimartingales with Gaussian martingale component it is possible to prove the following theorem.

Theorem 2.1.

(i) If

$$\int_0^\infty \left[(m'(u))^2 + A^2(u)\, K(u, u)\right] B^+(u)\, du < \infty ,$$

then $\mu_{\tau,y} \ll \nu_\tau$ for every stopping time τ.

(ii) If for all $0 \leq t < \infty$

$$\int_0^t \left[(m'(u))^2 + A^2(u)\, K(u, u)\right] B^+(u)\, du < \infty$$

and

$$\int_0^\infty \left[(m'(u))^2 + A^2(u)\, K(u, u)\right] B^+(u)\, du = \infty ,$$

then

$$\mu_y(\tau < \infty) = 1 \quad \text{iff} \quad \mu_{\tau,y} \ll \nu_\tau ,$$

and

$$\mu_y(\tau = \infty) = 1 \quad \text{iff} \quad \mu_{\tau,y} \perp \nu_\tau .$$

(iii) If $\mu_{\tau,y} \ll \nu_\tau$, then the density function is given by

$$\frac{d\,\mu_{\tau,y}}{d\,\nu_\tau}(y) = \frac{1}{\sqrt{K(0,0)}} \exp\left\{\frac{1}{2}\left[(y(0))^2 - \frac{(y(0) - m(0))^2}{K(0,0)}\right]\right\} \times$$

$$\times \exp\left\{ \int_0^\tau (a(u) + A(u)\, y(u))\, B^+(u)\, dy(u) - \right.$$

$$\left. - \frac{1}{2} \int_0^\tau (a(u) + A(u)\, y(u))^2\, B^+(u)\, du \right\} .$$

Now we give an application of Theorem 2.1.

Example 1

Assume in addition that y is stationary. Then $m(t) = m$ and $K(t, s) = \sigma^2 \exp\left\{-\beta|t-s|\right\}$, where $\sigma^2 > 0$

and $\beta > 0$. In this particular case $a(t) = \beta m$, $A(t) = -\beta$, $B(t) = 2\sigma^2\beta$ so that

$$y(t) = y(0) + \int_0^t (\beta m - \beta y(u))\, du + \sqrt{2\sigma^2\beta}\; w(t),$$

where $w(t)$ is a Wiener process. Theorem 1 implies that

$$\mu_y(\tau < \infty) = 1 \quad \text{iff} \quad \mu_{\tau,y} << \nu_\tau$$

and that

$$\frac{d\mu_{\tau,y}}{d\nu_\tau}(y) = \frac{1}{\sigma} \exp\left\{\frac{1}{2}\left[(y(0))^2 - \frac{(y(0)-m)^2}{\sigma^2}\right]\right\} \times$$
$$\times \exp\left\{\frac{1}{2\sigma^2\beta}\left[\int_0^\tau (\beta m - \beta y(u))\, dy(u) - \frac{1}{2}\int_0^\tau (\beta m - \beta y(u))^2 du\right]\right\}$$

is the density function.

3. Estimation of θ

Recall that $m(t)$ is of the form

$$m(t) = \theta\, \varphi(t) + \psi(t).$$

We assume that the derivatives φ' and ψ' exist and that for $0 \leq t < \infty$

$$\int_0^t \left[(\varphi'(u))^2 + (\psi'(u))^2 + A^2(u)\, K(u, u)\right] B^+(u)\, du < \infty .$$

We consider the family

$$\left\{\mu_y^\theta : \theta \in \Theta \subset (-\infty, \infty)\right\}$$

of Gaussian Markov measures with the mean function $m(t)$ and the covariance operator $K(t, s)$. For each $\theta \in \Theta$ let $\mu_{\tau,y}^\theta$ be the restriction of the measure μ_y^θ to the σ-field $\mathcal{B}_\tau$. Index θ indicates that the distribution μ_y^θ of y depends upon $\theta \in \Theta$, where Θ is an open interval on the real line.

Theorem 2.1. assertes that if $\mu_y^\Theta(\tau < \infty) = 1$ for all $\Theta \in \Theta$, then $\mu_{\tau,y}^\Theta \ll \nu_\tau$ for all $\Theta \in \Theta$ and the density function is given by

$$(3.1) \quad \frac{d\mu_{\tau,y}^\Theta}{d\nu_\tau}(y) = S(\tau, y)\exp\left\{-\frac{1}{2}\Theta^2 u(\tau) + \Theta\lambda(\tau, y)\right\},$$

where

$$(3.2) \quad u(\tau) = \varphi^2(0)K^+(0,0) + \int_0^\tau (\varphi'(u) - A(u)\varphi(u))^2 B^+(u)\,du,$$

$$(3.3) \quad \lambda(\tau, y) = (y(0) - \psi(0))\varphi(0)K^+(0,0) + \int_0^\tau (\varphi'(u) - A(u)\varphi(u))B^+(u)\big[dy(u) - (\psi'(u) + A(u)(y(u) - \psi(u)))\,du\big],$$

while

$$S(\tau, y) = \frac{1}{\sqrt{K(0,0)}}\exp\left\{\frac{1}{2}\left[(y(0))^2 - \frac{(y(0) - \psi(0))^2}{K(0,0)}\right]\right\}\times$$

$$(3.4) \quad \times\exp\left\{\int_0^\tau (\psi'(u) + A(u)(y(u) - \psi(u)))B^+(u)\,dy(u) - \frac{1}{2}\int_0^\tau (\psi'(u) + A(u)(y(u) - \psi(u)))^2 B^+(u)\,du\right\}.$$

Having an explicite formula for the density function we may use the maximum likelihood method to study sequential plans for estimation of Θ.

Let τ be a stopping time with respect to $\{\mathcal{B}_t\}$. A function $f\colon [0, \infty] \times C \longrightarrow (-\infty, \infty)$ is called an estimator of Θ if $f(\tau(\cdot), \cdot)$ is $\mathcal{B}_\tau$ measurable for every τ. A pair $\delta = (\tau, f)$, where τ is a stopping time and f is an estimator, is called a sequential plan.

We limit our considerations to a quadratic loss function $L(\Theta, \delta) = (f - \Theta)^2 + H(\tau)$, where H is a cost function. We assume that $H(t)$ is nonnegative, lower semicontinuous and such that

$$\lim_{t \to \infty} H(t) = \infty .$$

Let $\mathcal{D}$ denote the set of all those sequential plans $\delta = (\tau, f)$ that have a finite risk function

$$R(\Theta, \delta) = E_{\Theta}\left[(f - \Theta)^2 + H(\tau)\right]$$

for all $\Theta \in \Theta$. The expectation is here taken with respect to μ_y^{Θ}.

A sequential plan $\delta = (\tau, f)$ is said to be minimax if

$$\sup_{\Theta \in \Theta} R(\Theta, \delta) = \inf_{\delta \in \mathcal{D}} \sup_{\Theta \in \Theta} R(\Theta, \delta).$$

Suppose that a prior probability distribution $\pi(\Theta)d\Theta$ of Θ is given. The integral

$$r(\pi, \delta) = \int_{\Theta} R(\Theta, \delta)\pi(\Theta)\,d\Theta$$

is called the Bayesian risk of δ, provided it exists.

A sequential plan $\delta = (\tau, f)$ is said to be Bayes with respect to π if

$$r(\pi, \delta) = \inf_{\delta \in \mathcal{D}} r(\pi, \delta).$$

First we consider the case $\Theta = (-\infty, \infty)$. In view of (3.1) the maximum likelihood estimator of Θ is given by

$$\hat{\Theta}_{\tau} = \frac{\lambda(\tau, y)}{u(\tau)} .$$

In latter considerations we use the fact that $\hat{\Theta}_{\tau}$ is a limit of Bayes estimators. To prove this we introduce a sequence of normal prior distributions with densities

$$\pi_n(\Theta) = \sqrt{\frac{u_n}{2\pi}} \exp\left\{- u_n \frac{\Theta^2}{2}\right\}.$$

According to (3.1) the density function of the posterior probability distribution is given by

$$\pi_n(\theta | y) = \frac{\exp\left\{-\frac{\theta^2}{2}(u(\tau) + u_n) + \theta\lambda(\tau, y)\right\}}{\int_{-\infty}^{\infty} \exp\left\{-\frac{\theta^2}{2}(u(\tau) + u_n) + \theta\lambda(\tau, y)\right\} d\theta} .$$

Using simple algebraic manipulations we obtain

$$\pi_n(\theta | y) = \sqrt{\frac{u(\tau) + u_n}{2\pi}} \exp\left\{-\frac{u(\tau) + u_n}{2}\left(\theta - \frac{\lambda(\tau, y)}{u(\tau) + u_n}\right)^2\right\} .$$

Since

$$r(\pi_n, f | y) = \int_{-\infty}^{\infty} (f - \theta)^2 \pi_n(\theta | y)\, d\theta$$

attains its minimum value at

$$\hat{\theta}_\tau^n = \int_{-\infty}^{\infty} \theta\, \pi_n(\theta | y)\, d\theta ,$$

the Bayes estimator with respect to π_n is given by

$$\hat{\theta}_\tau^n = \frac{\lambda(\tau, y)}{u(\tau) + u_n} .$$

Clearly, if $\lim_{n \to \infty} u_n = 0$, then $\lim_{n \to \infty} \hat{\theta}_\tau^n = \hat{\theta}_\tau$.

A simple calculation shows that the posterior risk of the estimator $\hat{\theta}_\tau^n$ is equal to

$$r(\pi_n, \hat{\theta}_\tau^n | y) = \frac{1}{u(\tau) + u_n} .$$

Now we proceed to sequential estimation of θ .

Because

$$\inf_{\delta \in \mathfrak{D}} r(\pi_n, \delta) = \inf_{\tau} E_\theta\left(\frac{1}{u(\tau) + u_n} + H(\tau)\right),$$

the problem of finding Bayes sequential plan reduces to the problem of minimizing

$$E_\theta\left(\frac{1}{u(\tau) + u_n} + H(\tau)\right)$$

with respect to τ . It is clear that a fixed time sequential plan $\delta_n = (T_n, \hat{\theta}^n_{T_n})$, where T_n is determined by

$$\frac{1}{u(T_n)+u_n} + H(T_n) = \inf_T \left(\frac{1}{u(T)+u_n} + H(T)\right),$$

is Bayes with respect to Π_n.

Now let $\delta_o = (T_o, \hat{\theta}_{T_o})$ be a fixed time sequential plan with

$$\hat{\theta}_{T_o} = \frac{\lambda(T_o, y)}{u(T_o)}$$

and with T_o determined by

$$\frac{1}{u(T_o)} + H(T_o) = \inf_T \left(\frac{1}{u(T)} + H(T)\right).$$

Theorem 3.1.

Plan $\delta_o = (T_o, \hat{\theta}_{T_o})$ is minimax. Moreover, $\hat{\theta}_{T_o}$ is normally distributed with mean value θ and variance $1/u(T_o)$.

Proof.

Using (2.1), (3.2) and (3.3) gives

$$\lambda(T_o, y) - \theta\, u(T_o) = (y(0)-\theta\varphi(0)-\psi(0))\varphi(0)\, K^+(0,0) + \\ + \int_0^{T_o} (\varphi'(u) - A(u)\varphi(u))(B^{1/2}(u))^+\, dw(u).$$

(3.5) $$E_\theta (y(0) - \theta\varphi(0) - \psi(0))\varphi(0)\, K^+(0,0) \times \\ \times \int_0^{T_o} (\varphi'(u) - A(u)\varphi(u))(B^{1/2}(u))^+\, dw(u) = 0.$$

Thus the assertion concerning the distribution of $\hat{\theta}_{T_o}$ follows.

A simple calculation shows that

$$\sup_{\theta \in \Theta} R(\theta, \delta_0) = \inf_{\delta \in \mathcal{D}} r(F_n, \delta) + \frac{1}{u(T_0)} + H(T_0) -$$

$$- \frac{1}{u(T_n)+u_n} - H(T_n) \leq \inf_{\delta \in \mathcal{D}} \sup_{\theta \in \Theta} R(\theta, \delta) +$$

$$+ \frac{1}{u(T_0)} + H(T_0) - \frac{1}{u(T_n)+u_n} - H(T_n).$$

Moreover, if $\lim_{n \to \infty} u_n = 0$, then

$$\lim_{n \to \infty} \left(\frac{1}{u(T_n)+u_n} + H(T_n) \right) = \frac{1}{u(T_0)} + H(T_0).$$

Hence

$$\sup_{\theta \in \Theta} R(\theta, \delta_0) \leq \inf_{\delta \in \mathcal{D}} \sup_{\theta \in \Theta} R(\theta, \delta),$$

and δ_0 is minimax.

Example 2

Let y be defined as in Example 1. From the theorems established above the following results may be easily deduced.

For all $\theta \in \Theta$

$$\mu_y^{\theta}(\tau < \infty) = 1 \quad \text{iff} \quad \mu_{\tau, y} \ll \nu_{\tau}.$$

The density function is given by

$$\frac{d\mu^{\theta}_{\tau, y}}{d\nu_{\tau}}(y) = g(\tau, y) \exp\left\{ -\frac{1}{2}\theta^2 u(\tau) + \theta \lambda(\tau, y) \right\},$$

where

$$u(\tau) = \frac{2 + \beta\tau}{2\sigma^2} \quad \text{while} \quad \lambda(\tau, y) = \frac{y(0) + y(\tau) + \beta \int_0^{\tau} y(u)\,du}{2\sigma^2}$$

The maximum likelihood estimator of θ is given by

$$\hat{\theta}_{\tau} = \frac{y(0) + y(\tau) + \beta \int_0^{\tau} y(u)\,du}{2 + \beta\tau}$$

The fixed-time sequential plan $\delta_0 = (T_0, \hat{\theta}_{T_0})$, where T_0 is determined by

$$\frac{2\sigma^2}{2+\varrho T_0} + H(T_0) = \inf_T \left[\frac{2\sigma^2}{2+\varrho T} + H(T) \right].$$

is minimax.

The problem of estimation of θ has been also in a recent paper considered by Różański [20] for a stationary Gaussian Markov process.

All the above results have been derived under the assumption that the risk of δ is given by the formula

$$R(\theta, \delta) = E_\theta \left[(f - \theta)^2 + H(\tau) \right],$$

where H represents the cost function. Now we proceed also to consider sequential estimation of θ assuming that the cost of observations is not taken into account, i.e. that the risk of δ is given by

$$\tilde{R}(\theta, \delta) = E_\theta (f - \theta)^2.$$

Clearly, in this case it is necessary to impose additional restrictions on the stopping times considered. Otherwise the optimal stopping time τ would be equal to $+\infty$ with probability one.

Let $\mathcal{D}(T)$ denote the set of all sequential plans $\delta = (\tau, f)$ for which $\tilde{R}(\theta, \delta)$ is finite and $E_\theta u(\tau) \leq u(T)$ holds for all $\theta \in \Theta$.

If the function $(\varphi'(u) - A(u)\varphi(u))^2 B^+(u)$ is non-increasing and if

$$E_\theta \tau \leq T,$$

then

$$E_\theta u(\tau) \leq u(T).$$

A sequential plan $\delta_1 = (\tau_1, f_1)$ is said to be better than $\delta_2 = (\tau_2, f_2)$ if

$$\tilde{R}(\theta, \delta_1) \leq \tilde{R}(\theta, \delta_2)$$

for all θ and a strict inequality for at least one $\theta \in \Theta$.

A sequential plan $\delta \in \mathcal{D}(T)$ is said to be admissible among $\mathcal{D}(T)$ if there is no other plan in $\mathcal{D}(T)$ which is better than δ.

A sequential plan δ is said to be minimax if

$$\sup_{\theta \in \Theta} \tilde{R}(\theta, \delta) = \inf_{\delta \in \mathcal{D}(T)} \sup_{\theta \in \Theta} \tilde{R}(\theta, \delta).$$

Function $b(\theta) = E_\theta(f - \theta)$ is called the bias function of $\delta = (\tau, f)$. If $b(\theta) = 0$, then $\delta = (\tau, f)$ is said to be unbiased.

A sequential plan $\delta = (\tau, f)$ is said to be best unbiased if it is unbiased and if

$$\tilde{R}(\theta, \delta') \geq \tilde{R}(\theta, \delta)$$

for all $\theta \in \Theta$ and all unbiased sequential plans δ' in $\mathcal{D}(T)$.

We prove that $\delta_T = (T, \hat{\theta}_T) \in \mathcal{D}(T)$ is admissible and minimax. To establish this we need a lemma which may be considered as an analogue to the classical Cramér-Rao inequality. It is as follows.

<u>Lemma 3.1</u>

If $\delta = (\tau, f)$ is a sequential plan and if

$$\int_{\theta_1}^{\theta_2} E_\theta(f^2 + u(\tau))\, d\theta < \infty$$

for $\theta_1 < \theta_2$, $\theta_1, \theta_2 \in \Theta$,

then

$$E_{\theta_2} f - E_{\theta_1} f =$$

$$(3.6) \quad = \int_{\theta_1}^{\theta_2} E_\theta f \Big\{ (y(0) - \theta\varphi(0) - \psi(0))\varphi(0) K^+(0, 0) +$$

$$+ \int_0^{\tau} (\varphi'(t) - A(t)\varphi(t))(B^{1/2}(t))^+ \, dw(t) .$$

Moreover, if

$$\int_{\theta_1}^{\theta_2} E_\theta u(\tau) \, d\theta > 0,$$

then

$$(3.7) \quad \int_{\theta_1}^{\theta_2} \tilde{R}(\theta, \delta) \, d\theta \geqslant \int_{\theta_1}^{\theta_2} b^2(\theta) \, d\theta +$$

$$+ \frac{(\theta_2 - \theta_1 + b(\theta_2) - b(\theta_1))^2}{\int_{\theta_1}^{\theta_2} E_\theta u(\tau) \, d\theta} .$$

Proof.

Note that

$$E_{\theta_2} f - E_{\theta_1} f =$$

$$= \int_C f(c^x) \int_{\theta_1}^{\theta_2} \frac{d}{d\theta} \frac{1}{\sqrt{K(0,0)}} \exp\left\{- \frac{(x - \theta\varphi(0) - \psi(0))^2}{2K(0,0)}\right\} \times$$

$$\times \frac{d\mu^\theta_{\tau, y^x}}{d\mu_{\tau, v^x}} (c^x) \, d\theta \, d\nu_\tau(x, c^x),$$

where $d\nu_\tau(x, c^x) = \frac{1}{\sqrt{2\pi}} \, d\mu_{\tau, v^x}(c^x) \exp\left\{- \frac{x^2}{2}\right\} dx.$

By (3.1)

$$\int_{\theta_1}^{\theta_2} \int_C \Big| f(c^x) \frac{d}{d\theta} \frac{1}{\sqrt{K(0,0)}} \exp\left\{ -\frac{(x-\theta\varphi(0)-\psi(0))^2}{2K(0,0)} \right\} \times$$

$$\times \frac{d\mu^{\theta}_{\tau, y^x}}{d\mu_{\tau, y^x}}(c^x) \Big| d\nu_{\tau}(x, c^x)\, d\theta =$$

$$= \int_{\theta_1}^{\theta_2} E_\theta \Big| f \Big\{ (y(0) - \theta\varphi(0) - \psi(0))\varphi(0) K^+(0,0) +$$

$$+ \int_0^{\tau} (\varphi'(t) - A(t)\varphi(t))(B^{1/2}(t))^+ dw(t) \Big| \, d\theta \leq$$

$$\leq \left(\int_{\theta_1}^{\theta_2} E_\theta f^2 \, d\theta \right)^{1/2} \left(\int_{\theta_1}^{\theta_2} E_\theta u(\tau)\, d\theta \right)^{1/2} < \infty .$$

Now Fubini's theorem yields (3.6). To complete the proof note that

$$(E_{\theta_2} f - E_{\theta_1} f)^2 \leq \int_{\theta_1}^{\theta_2} E_\theta (f - E_\theta f)^2 d\theta \int_{\theta_1}^{\theta_2} E_\theta u(\tau)\, d\theta =$$

$$= \left\{ \int_{\theta_1}^{\theta_2} E_\theta (f-\theta)^2 \, d\theta - \int_{\theta_1}^{\theta_2} (E_\theta (f-\theta))^2 \, d\theta \right\} \times$$

$$\times \int_{\theta_1}^{\theta_2} E_\theta u(\tau)\, d\theta .$$

Theorem 3.2

If $\Theta = (-\infty, \infty)$, then $\delta_T = (T, \hat{\theta}_T)$ is admissible and minimax.

Proof.

Suppose that δ_T is not admissible. Then, there exists a sequential plan $\delta = (\tau, f)$ such that

$$\tilde{R}(\theta, \delta) \leq \tilde{R}(\theta, \delta_T) = \frac{1}{u(T)}$$

with a strict inequality for at least one θ . Since $\sup_{\theta_1 \leq \theta \leq \theta_2} E_\theta u(\tau) \leq u(T) < \infty$ the assumption of Lemma 3.1 is fulfilled. Hence, according to (3.7)

$$(3.8)\quad \frac{1}{u(T)} \geq \sup_{\theta_1 \leq \theta \leq \theta_2} \tilde{R}(\theta, \hat{\delta}) \geq \frac{1}{\theta_2 - \theta_1} \int_{\theta_1}^{\theta_2} \tilde{R}(\theta, \hat{\delta})\, d\theta \geq$$

$$\geq \frac{1}{\theta_2 - \theta_1} \int_{\theta_1}^{\theta_2} b^2(\theta)\, d\theta + \frac{\left(1 + \frac{b(\theta_2) - b(\theta_1)}{\theta_2 - \theta_1}\right)^2}{u(T)} .$$

Now we show that $b(\theta) \equiv 0$ is the only function satisfying this inequality.

Function $b(\theta)$ is non-increasing because

$$\left(1 + \frac{b(\theta_2) - b(\theta_1)}{\theta_2 - \theta_1}\right)^2 \leq 1.$$

Moreover, $b(\theta)$ is bounded. To prove this we consider first the case $b(\theta) \geq 0$. Then

$$b^2(\theta_2) \leq \frac{1}{\theta_2 - \theta_1} \int_{\theta_1}^{\theta_2} b^2(\theta)\, d\theta \leq \frac{1}{u(T)}$$

for every $\theta_1 \leq \theta_2$. Similarly, $b(\theta)$ is bounded when $b(\theta) \leq 0$. Because b is non-increasing there exists at most one value θ_0 such that $b(\theta) \geq 0$ for $\theta \leq \theta_0$ and $b(\theta) \leq 0$ for $\theta \geq \theta_0$. Considering these two intervals separately we easily establish that $b(\theta)$ is bounded.

Note that there exists a sequence $\{\theta_n\}$ such that $\lim_{n \to \infty} \theta_n = \infty$ and

$$\lim_{n\to\infty} \frac{b(\theta_n) - b(\theta_{n-1})}{\theta_n - \theta_{n-1}} = 0.$$

Suppose, on the contrary, that there exist $\varepsilon > 0$ and θ^* such that for every $\theta_2 \geq \theta_1 \geq \theta^*$

$$\frac{b(\theta_2) - b(\theta_1)}{\theta_2 - \theta_1} < -\varepsilon .$$

Then for every $\theta > \theta_1$

$$b(\theta) < -\varepsilon(\theta - \theta_1) + b(\theta_1).$$

This shows that b can not be bounded.

Substituting $\{\theta_n\}$ into (3.8) we have

$$\frac{1}{\theta_n - \theta_{n-1}} \int_{\theta_{n-1}}^{\theta_n} b^2(\theta)\, d\theta \leq$$

$$\leq \frac{1}{u(T)} \frac{b(\theta_{n-1}) - b(\theta_n)}{\theta_n - \theta_{n-1}} \left(2 + \frac{b(\theta_n) - b(\theta_{n-1})}{\theta_n - \theta_{n-1}}\right).$$

Moreover, for sufficiently large n (such that $\theta_{n-1} > \theta^*$)

$$\min(b^2(\theta_{n-1}), b^2(\theta_n)) \leq \frac{1}{\theta_n - \theta_{n-1}} \int_{\theta_{n-1}}^{\theta_n} b^2(\theta)\, d\theta ,$$

so that

$$\lim_{n\to\infty} b(\theta_n) = 0.$$

Similar we may prove that there exists a sequence $\{\tilde{\theta}_n\}$ such that $\lim_{n\to\infty} \tilde{\theta}_n = -\infty$ and $\lim_{n\to\infty} b(\tilde{\theta}_n) = 0.$

Using the fact that $b(\theta)$ is non-increasing and tends to zero when θ tends to $\pm\infty$ we see that $b(\theta) \equiv 0.$

In view of (3.8) it is clear that $\sup_{\Theta_1 \le \Theta \le \Theta_2} \tilde{R}(\Theta, \delta) = \frac{1}{u(T)}$ for every $\Theta_1 < \Theta_2$. This implies that $\tilde{R}(\Theta, \delta) = \frac{1}{u(T)}$ for all Θ. Thus the admissibility of δ_T is proved.

Minimaxity of δ_T follows easily from the fact that δ_T has a constant risk. Indeed suppose that δ_T is not minimax. Then there exists a sequential plan $\delta = (\tau, f)$ such that

$$\sup_{\Theta} \tilde{R}(\Theta, \delta) < \sup_{\Theta} \tilde{R}(\Theta, \delta_T) = \frac{1}{u(T)} .$$

This implies that $\tilde{R}(\Theta, \delta) < \tilde{R}(\Theta, \delta_T)$ for all Θ. This shows that δ_T is not admissible.

It is interesting to note that in case the parameter space is truncated δ_T is minimax, but not admissible. For example δ_T is worse than $\delta_T^* = (T, \max(\Theta_0, \hat{\Theta}_T))$ when $\Theta = (\Theta_0, \infty)$.

<u>Theorem 3.3</u>

If $\Theta = (\Theta_0, \infty)$, then $\delta_T = (T, \hat{\Theta}_T)$ is minimax.

Proof.

Suppose that δ_T is not minimax. Then there exists a plan $\delta = (\tau, f)$ such that

$$\sup_{\Theta \geq \Theta_0} \tilde{R}(\Theta, \delta) < \frac{1}{u(T)} .$$

Hence, $\tilde{R}(\Theta, \delta) \le \frac{1}{u(T)} - \varepsilon$ for all $\Theta \geq \Theta_0$ and some $\varepsilon > 0$. It is easy to see that $b(\Theta)$ is bounded.

Since

$$(f + \frac{b(\Theta_2) - b(\Theta_1)}{\Theta_2 - \Theta_1})^2 \le 1 - \varepsilon u(T)$$

we obtain after a simple calculation that

$$\frac{b(\theta_2) - b(\theta_1)}{\theta_2 - \theta_1} < -\varepsilon \frac{u(T)}{1+\sqrt{1-\varepsilon u(T)}}$$

holds for every $\theta_2 > \theta_1 > \theta_0$. This implies that $b(\theta)$ is unbounded. Thus δ_T is minimax.

Now we assume that the parameter space Θ is an open interval on the real line and consider best unbiased sequential plans for θ. As mentioned earlier δ_T is unbiased.

Theorem 3.4

Plan $\delta_T = (T, \hat{\theta}_T)$ is best among all unbiased plans in $\mathcal{D}(T)$.

Proof.

This assertion follows in a straightforward way from (3.8).

Example 3

From Theorem 3.1 it follows that $\delta_T = (T, \hat{\theta}_T)$, where

$$\hat{\theta}_T = \frac{y(0) + y(T) + \beta \int_0^T y(u)\, du}{2 + \beta T}$$

is admissible and minimax among the class of plans $\delta = (\tau, f)$ which satisfy the following two conditions

$$E_\theta f^2 < \infty, \quad E_\theta \tau \leq T ; \qquad \theta \in (-\infty, \infty).$$

If the parameter space is truncated δ_T is minimax but not admissible. Finally, Theorem 3.4 shows that δ_T is a best unbiased plan.

4. Estimation of α

Note that the covariance operator $K(t, s)$ of the stochastic process y defined in Section 1 is equal to

$$K(t, s) = \exp\left\{\int_s^t A(u)\, du\right\} K(s, s),$$

where

$$K(s, s) = \exp\left\{2 \int_0^s A(u)\, du\right\}\left\{K(0, 0) + \right.$$

$$\left. + \int_0^s \exp\left\{-2 \int_0^u A(\vartheta)\right\} d\vartheta\, B(u)\, du\right\}.$$

In this section we postulate that

$$A(t) = \alpha p(t) + q(t)$$

and consider the problem of estimation of the parameter α. We assume that α ranges over an open interval Ω on the real line. Functions p and q are known and such that

$$\int_0^t \left[(m'(u))^2 + (p^2(u) + q^2(u))\, K(u, u)\right] B^+(u)\, du < \infty$$

for $t < \infty$ and that

$$\int_0^\infty p^2(u)\, K(u, u)\, B^+(u)\, du = \infty.$$

Since we use here the same methods as in the case of estimation of θ we omit the proofs.

Consider the stopping time

$$\tau_T(y) = \inf\left\{t: Z(t, y) > T\right\},$$

where

$$Z(t, y) = \int_0^t p^2(u)(y(u) - m(u))^2\, B^+(u)\, du.$$

It is easy to see that $\tau_T(y)$ is nondecreasing, with respect to T,

$$\mu_y^{\alpha}(\tau_T < \infty) = 1, \qquad \alpha \in \Omega,$$

for all $T < \infty$ and

$$\mu_y^{\alpha}(\lim_{T \to \infty} \tau_T = \infty) = 1, \qquad \alpha \in \Omega.$$

Consider the sequential plan

$$\rho_T = (\tau_T(y), \frac{1}{T} \eta(\tau_T, y)),$$

where

$$\eta(\tau_T, y) = \int_0^{\tau_T} p(u)(y(u) - m(u)) B^+(u) \Big[dy(u) -$$

$$- (m'(u) + q(u)(y(u) - m(u)))du\Big].$$

The estimator $\frac{1}{T} \eta(\tau_T, y)$ has a normal distribution with expectation α and variance $\frac{1}{T}$. The risk function including the cost term is now of the form

$$R(\alpha, \delta) = E_{\alpha}\left[(f - \alpha)^2 + H(Z(\tau, y))\right],$$

where H is defined as in Section 3. Assuming that $\Omega = (-\infty, \infty)$ the following theorem may be established.

<u>Theorem 4.1</u>

Plan $\rho_T = (\tau_T, \frac{1}{T}\eta(\tau_T, y))$, where T is determined by

$$\frac{1}{T} + H(T) = \inf_{t \geq 0} (\frac{1}{t} + H(t))$$

is minimax.

If the risk function

$$R(\alpha, \delta) = E_{\alpha}(f - \alpha)^2$$

which does not take into account the cost of observations is postulated, one may establish, by using arguments similar as those in Section 3, the following results.

Let $\mathcal{D}(T)$ denote the set of all sequential plans $\delta = (\tau, f)$ such that $\tilde{R}(\alpha, \delta)$ is finite and $E_\alpha Z(\tau, y) \leq T$ holds for all $\alpha \in \Omega$.

Theorem 4.2

(i) If $\Omega = (-\infty, \infty)$, then ρ_T is admissible, minimax and best unbiased among $\mathcal{D}(T)$.
(ii) If the parameter space is truncated, say $\Omega = (\alpha_0, \infty)$, then ρ_T is minimax and best unbiased among $\mathcal{D}(T)$.
(iii) If Ω is an open interval, then ρ_T is best among all unbiased plans in $\mathcal{D}(T)$.

As already mentioned the optimal stopping time τ_T is finite. Moreover, Theorem 4.3 below assertes that all moments of τ_T are finite when

$$0 < \inf_t \frac{|P(t)|}{B(t)} = a, \qquad b = \sup_t \frac{|P(t)|}{B(t)} < \infty,$$

$$0 < \inf_t B(t) = c, \qquad d = \sup_t \frac{|h(t)|}{B(t)} < \infty.$$

Theorem 4.3

Under these qualifications there exist for every $n = 1, 2, \ldots$ constants a_n, b_n and c_n depending upon a, b, c and d only such that

$$E_\alpha \tau_T^n \quad (a_n |\alpha|^n + b_n)\, T^n + c_n\, T^{n/2}.$$

This theorem may be establish by using some ideas of Wognik (Theorem 17.7. in [12]) and Musiela ([15]).

Acknowledgement

I wish to thank A. Kozek and the referees for useful comments and discussions.

References

[1] M. Arató, On the statistical examination of continuous state Markov processes, I, II, III, IV, Selected Transl. in Math. Statist. and Probability, 4(1978).

[2] C.R. Blyth, On minimax statistical decision procedures and their admissibility, Ann. Math. Statist., 22(1951), 22-42.

[3] C.R. Blyth and D.N. Roberts, On inequalities of Cramér--Rao type and admissibility proofs, Proc. Sixth Berkeley Symp. Math. Statist. Prob., 1(1970), 17-30.

[4] D.G. Chapman and H. Robbins, Minimum variance estimation without regularity assumptions, Ann. Math. Statist., 22(1951), 581-586.

[5] A. Dvoretzky, J. Kiefer and J. Wolfowitz, Sequential decision problems for processes with continuous time parameter, Problems of estimation, Ann. Math. Statist., 24 24(1953), 403-415.

[6] I. Frantz and R. Magiera, On sequential plans for the exponential class of processes, Zastosowania Matematyki, 16, 2(1978), 153-165.

[7] J. Hajek, On a simple linear model in Gaussian processes, Proc. Second Prague Conf. Prob. Th., (1962).

[8] J.L. Hodges and E.L. Lehmann, Some applications of the Cramér-Rao inequality, Proc. Sixth Berkeley Symp. Math. Statist. Prob., 2(1951), 13-22

[9] Ju. Kabanov, R. Lipcer and A. Shiryayev, Absolute continuity and singularity of locally absolutely continuous probability distributions I/II, (in Russian) Matem. Sb., T 107(149), N 3(11) (1978), 346-415/ T 108(150), N 1 (1979), 32-61.

[10] A.M. Kagan, Ju.V. Linnik and S.R. Rao, Characterization problems in mathematical statistics, (in Russian) Moscow (1972).

[11] J. Kiefer, On minimum variance estimators, Ann. Math. Statist., 23(1952), 627-629.

[12] R. Lipcer and A. Shiryayev, Statistics of random processes I/II. Springer-Verlag, Berlin (1977/1978).

[13] R. Magiera, On the inequality of Cramér-Rao type in sequential estimation theory, Zastosow. Matem., 14, 2(1974), 228-235.

[14] R. Magiera, On minimax estimation for the exponential class of processes, Zastosow. Matem., 15, 4(1977), 445--454.

[15] M. Musiela, Sequential estimation of parameters of a stochastic differential equation, Math. Operationsforsch. Statist., Ser. Statistics, 8(1978), N 4, 483--498.

[16] M. Musiela, Two examples of admissible and minimax sequential plans for parameters of stochastic processes, Preprint 156. Institute of Mathematics Polish Academy of Sciences (1978).

[17] M. Musiela and R. Zmyślony, Estimation of regression parameters of Gaussian Markov processes, Preprint 155. Institute of Mathematics Polish Academy od Sciences (1978).

[18] A. Novikov, Sequential estimation of parameters of diffussion processes (in Russian) Teor. Verojatnost. i Primenen., 16, 2(1971), 394-396.

[19] R. Różański, A characterization of efficient sequential plans for Markov stationary Gaussian processe, Zastosow. Matem., to appear.

[20] R. Różański, Minimax sequential estimation of the mean of an Ornstein-Uhlenbeck process, Zastosow. Matem., to appear.

[21] S. Trybuła, Sequential estimation in processes with independent increaments, Dissertationes Mathematicae 60(1968).

PARAMETER ESTIMATION FOR BILINEAR ELLIPTIC DIFFUSIONS

M. MUSIELA

Institute of Mathematics, Polish Academy of Sciences
Kopernika 18, 51-617 Wroclaw, Poland (*)

ABSTRACT. - This paper is concerned with the statistical problem of parameter estimation for bilinear elliptic diffusions. Asymptotic properties of a maximum likelihood estimator are investigated in the ergodic and non-ergodic cases.

KEY WORDS AND PHRASES. - Stochastic differential equation, bilinear, diffusion processes, maximum likelihood estimation.

A.M.S. 1980 Subject classification, Primary 60 H 10, Secondary 62 A 10.

(*) Presently : Laboratoire I.M.A.G., BP 53 X, 38041 Grenoble-Cédex, France.

1. INTRODUCTION.

Asymptotic properties of maximum likelihood estimators for parameters of diffusion processes have been intensively studied. Linear elliptic and hypoelliptic cases have been considered by Arato [1,2], Le Breton [10], Le Breton and Musiela [13]. A one-dimensional bilinear model has been investigated by Taraskin [19], Le Breton and Musiela [11]. The general case of linear dependence of the drift coefficient on a parameter has been considered by Taraskin [20], Brown and Hewitt [5], Lee and Kozin [14]. In all these papers ergodicity of corresponding diffusions was an important assumption. Novikov in [17] and Feigin in [6,7] have used martingale methods to study the maximum likelihood estimator of the arbitrary real parameter θ in the one-dimensional Ornstein-Uhlenbeck motion governed by the differential generator

$$\frac{1}{2} D^2 + \theta x D , \quad \left(D = \frac{d}{dx} \right) . \tag{1}$$

Novikov's result on convergence of the estimator has been generalized by Le Breton and Musiela [13] to the case of linear d-dimensional hypoelliptic diffusions.

In the present paper we consider bilinear elliptic diffusions of the following type.

Let Ω be the set of all continuous functions from $\mathbb{R}_+$ to $\mathbb{R}^d$. We denote $(X(t))_{t\geq 0}$ the coordinate process, $(\mathcal{F}_t)_{t\geq 0}$ its natural right continuous filtration and $\mathcal{F} = \bigvee_{t\geq 0} \mathcal{F}_t$. Moreover, let

$$a(x) = ((a_{ij}(x))) = \Sigma_{k=1}^{r} A_k \, xx' \, A_k' + AA' ,$$

$$b_\theta(x) = (b_i^\theta(x)) = B_\theta x ,$$

$$B_\theta = \Sigma_{k=1}^{q} \theta_k B_k , \quad \theta = (\theta_k) \in \mathbb{R}^q ,$$

where $d \times d$ matrices $B_1, \ldots, B_q$ are known and linearly independent and $\det AA' > 0$.

We define P^x_θ, $x \in \mathbb{R}^d$, $\theta \in \mathbb{R}^q$, to be the probability law on $(\Omega, \mathcal{F})$ such that

$$P^x_\theta \{X(0) = x\} = 1$$

and the process

$$M(t) = X(t) - \int_0^t b_\theta(X(s))\, ds \tag{2}$$

is an $(\mathcal{F}_t)$-local martingale on $\mathbb{R}_+$ with $(\int_0^t a(X(s))ds)$ as tensor quadratic variation process. $(\Omega, \mathcal{F}, \mathcal{F}_t, X(t), P^x_\theta)$ is the canonical realization of a diffusion process with differential operator

$$L = \frac{1}{2} \Sigma^d_{i,j=1} \; a_{ij} D_i D_j + \Sigma^d_{j=1} b^\theta_j D_j \;. \tag{3}$$

Parameters AA' and $\Sigma^r_{k=1} A_k \otimes A_k$ of the matrix valued function $(a(x) \; ; \; x \in \mathbb{R}^d)$ can be computed, with probability one on every finite time interval $[0,T]$, by use of the tensor quadratic variation process $(\int_0^t a(X(s))\,ds \;,\; 0 \le t \le T)$. Therefore we shall investigate the problem of estimating θ, in view of the observation of one trajectory on $[0,T]$, when $a(\cdot)$ is assumed known. We study convergence and asymptotic distribution for the maximum likelihood estimator (MLE) of the drift parameter θ.

2. MAXIMUM LIKELIHOOD ESTIMATION OF θ.

It is well known (cf. Liptser and Shiryaev [15]) that the restriction $P^x_{\theta,T}$ of P^x_θ to $\mathcal{F}_T$ is absolutely continuous with respect to $P^x_{0,T}$ for all $\theta \in \mathbb{R}^q$. Moreover, the Log-likelihood equation can be written in the form

$$\left(\left(\int_0^T X'(t) B_k' a^{-1}(X(t)) B_\ell X(t)\, dt\right)\right) \theta = \left(\int_0^T X'(t) B_k' a^{-1}(X(t))\, dX(t)\right). \quad (4)$$

The following lemma will be useful in order to compute a MLE of θ.

Lemma 1.

The matrix

$$\int_0^T X(t)\, X'(t) \otimes a^{-1}(X(t))\, dt$$

is P_θ^X almost surely positive definite.

Proof. In order to prove the assertion one has to show that the set

$$\Omega_o = \bigcup_{h \in \mathbb{R}^{d^2} - \{0\}} \{ h' \int_0^T X(t)\, X'(t) \otimes a^{-1}(X(t))\, dt\ h = 0 \}$$

is negligible. Since X has continuous sample paths it is sufficient to prove that

$$\Omega_1 = \bigcup_{h \in \mathbb{R}^{d^2} - \{0\}} \ \bigcap_{t \in [0,T]} \{h'\, X(t)\, X'(t) \otimes a^{-1}(X(t))h = 0\}$$

is negligible. But

$$h'X(t)\, X'(t) \otimes a^{-1}(X(t))\, h = (\text{vec } H)'\, X(t) X'(t) \otimes a^{-1}(X(t))\ \text{vec } H =$$

$$\text{tr } H' a^{-1}(X(t))\ H X(t)\ X'(t) = X'(t)\ H'\ a^{-1}(X(t))\ H X(t)\ ,$$

where H is square $d \times d$ matrix such that $\text{vec } H = h$ (see Henderson and Searle [8] for the definition and properties of vec operator). Therefore

$$\Omega_1 = \bigcup_{H \neq 0}\ \bigcap_{t \in [0,T]} \{H X(t) = 0\}\ .$$

It is also clear that

$$\Omega_1 \subset \Omega_2 = \bigcup_{h \in \mathbb{R}^d - \{0\}} \bigcap_{t \in [0,T]} \{ h' X(t) = 0 \} .$$

Now let $0 = t_o < t_1 < \cdots < t_d \leq T$. Since

$$\Omega_2 \subset \Omega_3 = \{ \det [X(t_1), \ldots, X(t_d)] = 0 \}$$

and the subset of $\mathbb{R}^{d^2}$.

$$\{ (V_1', \ldots, V_d')' \in \mathbb{R}^{d^2} : \det [V_1, \ldots, V_d] = 0 \}$$

has Lebesgue measure zero, the assertion will be proved if one shows that the random vector $(X'(t_1), \ldots, X'(t_d))'$ has a density with respect to Lebesgue measure in $\mathbb{R}^{d^2}$. This follows from ellipticity of the differential operator L defined in (3) and the fact that P_θ^X is Markov (cf. Ikeda and Watanabe [9]). ∎

The next result is then an immediate consequence of (4) and the above Lemma 1.

Proposition 1.

A maximum likelihood estimator $\hat{\theta}_T$ of the parameter θ is given by

$$\hat{\theta}_T = \left(\left(\int_0^T X'(t) B_k' a^{-1}(X(t)) B_\ell X(t)\, dt \right) \right)^{-1} \left(\int_0^T X'(t) B_k' a^{-1}(X(t))\, dX(t) \right) . \quad (5)$$

Proof. Since matrices $B_1, \ldots, B_q$ are linearly independent and

$$\int_0^T X'(t) B_k' a^{-1}(X(t)) B_\ell X(t)\, dt = (\mathrm{vec}\, B_k)' \int_0^T X(t) X'(t) \otimes a^{-1}(X(t))\, dt\ \mathrm{vec}\, B_\ell ,$$

then $((\int_0^T X'(t)\, B_k'\, a^{-1}(X(t))\, B_\ell\, X(t)\, dt))$ is a Gram matrix for vectors vec $B_1, \dots,$ vec B_q . Therefore it is almost surely positive definite. ∎

3. CONVERGENCE OF THE M L E OF θ.

We shall consider separately two cases : when the observed diffusion is ergodic and when the matrix $((\int_0^T X'(t)\, B_k'\, a^{-1}(X(t))\, B_\ell X(t)\, dt))$ is diagonal.

3.1. Ergodic case.

In the following Lemma we list some known results (cf. Brockett [4]).

Lemma 2.

If the matrix $B_\theta \otimes I + I \otimes B_\theta + \Sigma_{k=1}^{r} A_k \otimes A_k$ is stable, then

(i) there exists a unique invariant zero-mean probability measure μ for P_θ^x,

(ii) $\int_{\mathbb{R}^d} x'x\, d\mu(x) < \infty$,

(iii) μ possesses a density p which is C^∞ and satisfies the steady state Fokker-Planck equation,

(iv) P_θ^x is ergodic.

Now we can prove

Proposition 2.

If the matrix $B_\theta \otimes I + I \otimes B_\theta + \Sigma_{k=1}^{r} A_k \otimes A_k$ is stable, then the M L E $\hat{\theta}_T$ defined in (5) is strongly consistent i.e.

$$\lim_{T \to \infty} \hat{\theta}_T = \theta \quad , \qquad P_\theta^x \text{ a.s.}.$$

Proof. First we note that for $k, \ell = 1, \ldots, q$

$$\int_{\mathbb{R}^d} x' B_k' a^{-1}(x) B_\ell x \, d\mu(x) \le \frac{[\operatorname{tr} B_k B_k' \operatorname{tr} B_\ell B_\ell']^{\frac{1}{2}}}{\operatorname{tr} AA'} \int_{\mathbb{R}^d} x'x \, d\mu(x) < \infty .$$

Next we shall prove that the matrix $((\int_{\mathbb{R}^d} x' B_k' a^{-1}(x) B_\ell x \, d\mu(x)))$ is positive definite. It is sufficient to show that the matrix $\int_{\mathbb{R}^d} xx' \otimes a^{-1}(x) \, d\mu(x)$ is positive definite. We assume now that for some $h \in \mathbb{R}^{d^2} - \{0\}$ $h' \int_{\mathbb{R}^d} xx' \otimes a^{-1}(x) \, d\mu(x) h = 0$. Then, for $\operatorname{vec} H = h$ we have also $\int_{\mathbb{R}^d} x'H'a^{-1}(x) \, Hx \, d\mu(x) = 0$ what implies that $x'H' a^{-1}(x) Hx = 0$ μ a.e. . Therefore $Hx = 0$ μ a.e. and there exists $h_1 \in \mathbb{R}^d - \{0\}$ such that $h_1' x = 0$ μ a.e. . This would imply that $Q = \int_{\mathbb{R}^d} xx' \, d\mu(x)$ is not positive definite. But on the other hand we know (cf. Brockett [4] or Le Breton and Musiela [12]) that Q is the unique positive definite solution of the equation

$$B_\theta Q + Q B_\theta' + \Sigma_{k=1}^{r} A_k Q A_k' + AA' = 0 .$$

Then the matrix $((\int_{\mathbb{R}^d} x' B_k' a^{-1}(x) B_\ell x \, d\mu(x))$ is positive definite.

To prove consistency note that (2) and (5) yield

$$\hat{\theta}_T - \theta = ((\int_0^T X'(t) B_k' a^{-1}(X(t)) B_\ell X(t) \, dt))^{-1} (\int_0^T X'(t) B_k' a^{-1}(X(t)) \, dM(t)) . \quad (6)$$

Moreover, since P_θ^x is ergodic, we obtain that

$$\lim_{T\to\infty} \frac{1}{T} \left(\left(\int_0^T X'(t)\, B_k'\, a^{-1}(X(t))\, B_\ell\, X(t)\, dt \right)\right) = \left(\left(\int_{\mathbb{R}^d} x'\, B_k'\, a^{-1}(x)\, B_\ell\, x\, d\mu(x) \right)\right)$$

P_θ^x a.s. and using the analogue for continuous time martingales of the strong law (cf. Basawa and Prakasa Rao [3] Theorem 4.1. p.394 for example) we have also

$$\lim_{T\to\infty} \frac{1}{T} \left(\int_0^T X'(t)\, B_k'\, a^{-1}(X(t))\, dM(t) \right) = 0 \ , \quad P_\theta^x \text{ a.s.}$$

Therefore the conclusion of the proposition follows. ◻

3.2. Diagonal case.

In this paragraph we shall study convergence of the M L E without ergodicity assumption. We begin by the following

Lemma 3.

For any $d \times d$ matrix H such that $\det H \neq 0$

$$P_\theta^x \left\{ \int_0^\infty X'(t)\, H'\, a^{-1}(X(t))\, H X(t)\, dt = \infty \right\} = 1 \ . \tag{7}$$

Proof. First note that (cf. Musiela [16]) if there exists a number $\lambda > 0$ and a strictly positive function $\varphi \in C^2(\mathbb{R}^d)$ such that

$$\lim_{|x|\to\infty} \varphi(x) = \infty \ ,$$

$$L\varphi(x) \le \lambda\, x'\, H'\, a^{-1}(x)\, H x\, \varphi(x) \quad , \ x \in \mathbb{R}^d \ ,$$

then (7) holds. Now such a φ can be constructed by choosing $\varphi \in C^2(\mathbb{R}^d)$ so that $\varphi \ge 1$ and $\varphi(x) = u\left(\frac{|x|^2}{2}\right)$ for $|x| \ge (2r)^{1/2}$, where u is a solution of the second order differential equation

$$\frac{1}{2}(C_1 x + C_2)\,[\,D^2 u + C_3\, Du] = C_4 u\ , \quad x \geq r\ . \quad ■$$

The main result of this paragraph is the following

Proposition 3.

If $\det B_k \neq 0$, $k = 1,\ldots,q$ and for all $x \in \mathbb{R}^d$ $x' B_k' a^{-1}(x) B_\ell x = 0$, $k \neq \ell$, $k, \ell = 1,\ldots,q$, then the M L E $\hat{\theta}_T$ given in (5) is strongly consistent.

Proof. It is clear that

$$\hat{\theta}_T = \left(\frac{\int_0^T X'(t)\, B_k'\, a^{-1}(X(t))\, dX(t)}{\int_0^T X'(t)\, B_k'\, a^{-1}(X(t))\, B_k X(t)\, dt} \right) \tag{8}$$

and, as before, taking into account (6) we can write

$$\hat{\theta}_{T,k} - \theta_k = \frac{\int_0^T X'(t)\, B_k'\, a^{-1}(X(t))\, dM(t)}{\int_0^T X'(t)\, B_k'\, a^{-1}(X(t))\, B_k\, X(t)\, dt}\ .$$

Since $(\int_0^t X'(s)\, B_k'\, a^{-1}(X(s))\, B_k\, X(s)\, ds\ ,\ t \geq 0)$ is the quadratic variation process of martingale $(\int_0^t X'(s)\, B_k'\, a^{-1}(X(s))\, dM(s)\ ,\ t \geq 0)$ and (7) holds with $H = B_k$, $k = 1,\ldots,q$ then using the martingale version of the strong law of large numbers we get the assertion. ■

Example 1.

Let $d = q = 2$, $r = 1$ and $B_1 = A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $B_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, $A_1 = a\, B_2$, $a \in \mathbb{R}$. It is easy to see that the assumptions of Proposition 3 are fulfiled. Therefore $\hat{\theta}_T = (\hat{\theta}_{1,T}, \hat{\theta}_{2,T})'$, given in (8) , is strongly consistent. Note that

$$\hat{\theta}_{1,T} = \frac{\int_0^T [X_1(t)\, dX_1(t) + X_2(t)\, dX_2(t)]}{\int_0^T [X_1^2(t) + X_2^2(t)]\, dt} , \tag{9}$$

$$\hat{\theta}_{2,T} = \frac{\int_0^T \frac{X_1(t)\, dX_2(t) - X_2(t)\, dX_1(t)}{1 + a^2 [X_1^2(t) + X_2^2(t)]}}{\int_0^T \frac{X_1^2(t) + X_2^2(t)}{1 + a^2 [X_1^2(t) + X_2^2(t)]}\, dt} , \tag{10}$$

where $X(t) = (X_1(t) , X_2(t))'$.

4. ASYMPTOTIC DISTRIBUTION OF THE M L E OF θ .

As in Section 3 we shall first study the case when the observed diffusion is ergodic.

4.1. Ergodic case.

Proposition 4.

If the matrix $B_\theta \otimes I + I \otimes B_\theta + \Sigma_{k=1}^r A_k \otimes A_k$ is stable, then the random vector $T^{1/2}(\hat{\theta}_T - \theta)$ is asymptotically normally distributed with mean zero and covariance matrix $((\int_{\mathbf{R}^d} x' B_k' a^{-1}(x) B_\ell x \, d\mu(x)))^{-1}$, where μ is the unique invariant probability measure for P_θ^x.

Proof. On can deduce, using results of Taraskin [20], that the random vector $T^{-\frac{1}{2}}(\int_0^T X'(t) B_k' a^{-1}(X(t)) \, dM(t))$ is asymptotically normally distributed with mean zero and covariance matrix $((\int_{\mathbf{R}^d} x' B_k' a^{-1}(x) B_\ell x \, d\mu(x)))$. Therefore, the conclusion follows from (6). ∎

We shall complete the above proposition by two examples in which invariant measures are known explicitely.

Example 2 (Brockett [4]).

μ is Gaussian with variance Q if and only if $A_k Q + Q A_k' = 0$, $k = 1, \ldots, r$, and $B_\theta Q + Q B_\theta' + \Sigma_{k=1}^r A_k Q A_k' + AA' = 0$.

Example 3.

If $B_\theta = \theta I$, $r = 1$, $A_1 = aI$, $a \neq 0$, where θ and a are real such that $2\theta + a^2 < 0$ and I is the identity matrix, then $d\mu(x) = p(x)\,dx$, where

$$p(x) = C(1 + a^2 x'(AA')^{-1} x)^{-\frac{d+1}{2} + \frac{\theta}{a^2}}.$$

This result may be proved directly by verification of the steady state Fokker-Planck equation.

4.2. A two-dimensional diagonal example.

Without ergodicity assumption, even in the diagonal case, we were not able to find the asymptotic distribution of the M L E of θ. In order to give an insight into the problem we shall investigate now the two-dimensional diffusion process $X(t) = (X_1(t), X_2(t))'$ considered in Example 1. Let us recall that the M L E $\hat{\theta}_T = (\hat{\theta}_{1,T}, \hat{\theta}_{2,T})'$, where $\hat{\theta}_{1,T}$ and $\hat{\theta}_{2,T}$ are given in (9) and (10), respectively, is strongly consistent.

We introduce the following notation

$$V_{1,T} = \int_0^T [X_1^2(t) + X_2^2(t)]\, dt ,$$

$$V_{2,T} = \int_0^T \frac{X_1^2(t) + X_2^2(t)}{1+a^2[X_1^2(t)+X_2^2(t)]}\, dt .$$

Proposition 5.

(i) If $2\theta_1 + a^2 \neq 0$, then the random variable $V_{1,T}^{\frac{1}{2}}(\hat{\theta}_{1,T} - \theta_1)$ is asymptotically normally distributed with mean zero and variance equal to one.

(ii) If $2\theta_1 + a^2 = 0$, then $V_{1,T}^{\frac{1}{2}}(\hat{\theta}_{1,T} - \theta_1)$ converges in distribution to the distribution (which is not Gaussian) of the random variable

$$\frac{W_1^2(1) + W_2^2(1) - 2}{2\int_0^1 [W_1^2(t) + W_2^2(t)]\, dt} ,$$

where W_1 and W_2 are independent Brownian motions.

(iii) Random variables $V_{1,T}^{\frac{1}{2}}(\hat{\theta}_{1,T}-\theta_1)$ and $V_{2,T}^{\frac{1}{2}}(\hat{\theta}_{2,T}-\theta_2)$ are independent.

(iv) The distribution of $V_{2,T}^{\frac{1}{2}}(\hat{\theta}_{2,T}-\theta_2)$ is normal with mean zero and variance equal to one.

Proof. First we define the processes

$$\rho(t) = X'(t)\, X(t) ,$$

$$\tilde{W}_1(t) = \int_0^t \rho^{-\frac{1}{2}}(s)\, X'(s)\, dM(s)$$

$$\tilde{W}_2(t) = \int_0^t \left[\frac{\rho(s)}{1+a^2\rho(s)}\right]^{-\frac{1}{2}} X'(s)\, B_2'\, a^{-1}(X(s))\, dM(s)$$

and we denote by $(\mathcal{F}.^{\rho})$ and $(\mathcal{F}.^{\tilde{W}_2})$ the completed filtrations determined, respectively, by ρ and $\tilde{W}_2$.
It is easy to see that $\tilde{W}_1$ and $\tilde{W}_2$ are independent Brownian motions and that

$$d\rho(t) = [(2\theta_1+a^2)\rho(t)+2]\, dt + 2\, |\rho(t)|^{\frac{1}{2}}\, d\tilde{W}_1(t) \quad , \ t \geq 0 .$$

$$\rho(0) = x_1^2 + x_2^2 .$$

Since the above equation has a unique strong solution (cf. Ikeda and Watanabe [9], Theorem 3.2. p.168), then $(\mathcal{F}.^{\rho}) = (\mathcal{F}.^{W_1})$ and the stochastic processes ρ and $\tilde{W}_2$ are independent. Moreover, we know also (cf. Pitman and Yor [18]) that diffusion obtained by taking the sum of squares of two independent Ornstein-Uhlenbeck processes governed by the differential generator (1) with $\theta = 2\theta_1 + a^2$ satisfies the same equation.

Since $a(x)x = x$, then $V_{1,T}^{\frac{1}{2}}(\hat{\theta}_{1,T}-\theta_1) = V_{1,T}^{-\frac{1}{2}} \int_0^T \rho^{\frac{1}{2}}(t)\, d\tilde{W}_1(t)$

and (i) follows (cf. Basawa and Prakasa Rao [3], Theorem 4.2. p.395).

If $2\theta_1 + a^2 = 0$, then $\int_0^T \rho^{\frac{1}{2}}(t)\, d\tilde{W}_1(t) = \frac{1}{2}[\rho(T) - x_1^2 - x_2^2 - 2T]$.

Therefore $V_{1,T}^{\frac{1}{2}}(\theta_{1,T} - \theta_1) = \frac{1}{2} V_{1,T}^{-\frac{1}{2}} [\rho(T) - x_1^2 - x_2^2 - 2T]$ what yields (ii).

To get (iii) and (iv) it is sufficient to note that

$V_{2,T}^{\frac{1}{2}}(\hat{\theta}_{2,T} - \theta_2) = V_{2,T}^{-\frac{1}{2}} \int_0^T [\frac{\rho(t)}{1+a^2\rho(t)}]^{\frac{1}{2}} d\tilde{W}_2(t)$ and then to use independence of ρ and $\tilde{W}_2$. ▪

5. ACKNOWLEDGEMENT.

We would like to express our gratitude to Professor A. Le Breton for numerous enlightening discussions on the subject of this paper.

REFERENCES

[1] M. Arato - On the parameter estimation of processes satisfying a stochastic differential equation, Studia Sci. Math. Hungar.,5, 11-15 (1970).

[2] M. Arato - On the statistical examination of continuous state Markov processes, I, II, III, IV, Selected Transl. in Math. Statist. and Probability, 4 (1978).

[3] I.V. Basawa and B.L.S. Prakasa Rao - Statistical Inference for Stochastic Processes, Academic Press, London (1980).

[4] R.W. Brockett - Parametrically stochastic linear differential equations, in : R.J.B. Wets, Ed., Stochastic systems : Modeling, Identification and Optimization, I, Math. Prog. Study, 5,8-21, Amsterdam, North-Holland (1976).

[5] B.M. Brown and J.I. Hewitt - Asymptotic likelihood theory for diffusion processes, J.Appl. Prob. 12, 228-238 (1975).

[6] P.D. Feigin - Maximum likelihood estimation for continuous-time stochastic processes, Adv. Appl. Prob. 8, 712-736 (1976).

[7] P.D. Feigin - Some comments concerning a curious singularity, J. Appl. Prob. 16, 440-444 (1979).

[8] H.V. Henderson and S.R. Searle - Vec and vech operators for matrices, with some uses in Jacobians and multivariate statistics, The Canadian Journal of Statistics Vol.7. N°1., 65-81 (1979).

[9] N. Ikeda and S. Watanabe - Stochastic Differential Equations and Diffusion Processes, North-Holland, Amsterdam/Kadansha LTD, Tokyo (1981).

[10] A. Le Breton - Parameter estimation in a vector linear stochastic differential equation, in : Transactions of the Seventh Prague Conference on Information Theory. Statistical Decision Functions and Random Processes, Vol.A. 353-366 (1977).

[11] A. Le Breton and M. Musiela - A study of an one-dimensional differential model for stochastic processes, Prob. and Math. Statist. to appear.

[12] A. Le Breton and M. Musiela - A look at a bilinear model for multidimensional stochastic systems in continuous time, Stochastics and Decisions, to appear.

[13] A. Le Breton and M. Musiela - Some parameter estimation problems for hypoelliptic homogeneous Gaussian diffusions, Proc. Banach Math. Center, Warsaw (1981) to appear.

[14] T.S. Lee and F. Kozin - Almost sure asymptotic likelihood theory for diffusion processes, J. Appl. Prob. 14, 527-537 (1977).

[15] R.S. Liptser and A.N. Shiryaev - Statistics of Random Processes, I, II, Springer-Verlag, New-York (1978).

[16] M. Musiela, forthcoming paper.

[17] A.A. Novikov - Sequential estimation of the parameters of diffusion processes, Theor. Prob. Appl., Vol.16, N°2, 394-396 (1971).

[18] J. Pitman and M. Yor - A decomposition of Bessel bridges, Z. Wahrscheinlichkeitstheorie verw. Gebiete, 59, 425-457 (1982).

[19] A.F. Taraskin - The asymptotic normality of stochastic integrals and estimates of the coefficient of diffusion process transfer, in : Mathematical Physics N°8 Nankova Dumka, Kiev, 149-163, (1970).

[20] A.F. Taraskin - On the asymptotic normality of vector-valued stochastic integrals and estimates of drift parameters of a multidimensional diffusion process, Theory of Prob. and Math. Statist. 2, 209-224 (1974).

ON KAC FUNCTIONALS OF ONE DIMENSIONAL DIFFUSIONS

M. MUSIELA

Institute of Mathematics, Polish Academy of Sciences,
Kopernika 18, 51-617 Wroclaw, Poland (*)

ABSTRACT. - This paper uses martingale calculus in order to study multiplicative Kac functionals. Probabilistic representation of a solution of the Schrödinger equation with non necessarily negative potential is obtained. Necessary and sufficient conditions for the a.s. convergence and the a.s. divergence of some integrals are given.

KEY WORDS AND PHRASES. - Stochastic integrals, diffusion processes, multiplicative functionals, Schrödinger operator. A.M.S. 1980 Subject classification. Primary 60 J 60, Secondary 60 J 57.

(*) Presently : Laboratoire I.M.A.G., BP 53 X, 38041 Grenoble-Cédex, France.

1. INTRODUCTION.

Let $I =]\ell, r[$ be an open interval in $\mathbb{R}$ $(-\infty \leq \ell < r \leq \infty)$. Let Ω be the set of all continuous functions from $\mathbb{R}_+$ to $I \cup \{\ell, r\}$ which are stopped at the first exit time from I. We denote $(X_t)_{t \geq 0}$ the coordinate process, $(\mathcal{F}_t)_{t \geq 0}$ its natural right continuous filtration and $\mathcal{F} = \bigvee_{t \geq 0} \mathcal{F}_t$. For an open interval $J \subset I$ $S^J = \inf\{t : X_t \notin J\}$ is the first exit time of X from J. The lifetime of X is $T = S^I$.

Let a and b be sufficiently smooth real valued functions on I such that $a > 0$. Let P^x, $x \in I$, be the probability law on $(\Omega, \mathcal{F})$ such that $P^x\{X_o = x\} = 1$ and the process

$$M_t = X_t - \int_0^t b(X_s)\,ds$$

is an $(\mathcal{F}_t)$ - local martingale on $[0,T[$ with $(\int_0^t a(X_s)\,ds)$ as increasing process. $(\Omega, \mathcal{F}, \mathcal{F}_t, X_t, T, P^x)$ is the canonical realization of a diffusion process with differential operator

$$L = \frac{1}{2} a D^2 + b D .$$

Now let g be a real continuous function on I. In this paper we study the multiplicative functional

$$e_t = \exp\left(\int_0^t g(X_s)\,ds\right) \tag{1}$$

introduced by M. Kac in [11]. For a large class of non necessarily negative potentials g we prove that $P^x(e_{S^J})$ $(= \int e_{S^J}\,dP^x)$ is a solution of the Schrödinger equation

$$(L + g)\,u = 0 \tag{2}$$

on J. Moreover, we give for a continuous $f \geq 0$ and all intervals $J \subset I$, necessary and sufficient conditions for

$$P^x \{ \int_0^{S^J} f(X_t)\, dt = \infty \} = 1$$

and

$$P^x \{ \int_0^{S^J} f(X_t)\, dt < \infty \} = 1 .$$

2. THE DIFFUSION ASSOCIATED WITH THE SCHRODINGER EQUATION.

Now we establish a connection between the mean function $P^x(e_{S^J})$ of the multiplicative functional (1) and the Schrödinger equation (2) on an interval $J =]\alpha, \beta[$, $\ell < \alpha < \beta < r$.

Let $v \in C^2(I)$ be a nonvanishing real function. It is well known that the process

$$N_t = \frac{v(X_t)}{v(x)} \exp(-\int_0^t \frac{Lv}{v}(X_s)\, ds)$$

is a positive $(\mathcal{F}_t)$-local martingale on $[0, T[$. Moreover for all $t > 0$, $x \in J$ and $J =]\alpha, \beta[$, $\ell < \alpha < \beta < r$, $P^x(N_{t \wedge S^J}) = 1$.

Now let $P^x_{t \wedge S^J}$ be the restriction of P^x to $\mathcal{F}_{t \wedge S^J}$. Define $\tilde{Q}^x_{t \wedge S^J} = N_{t \wedge S^J} P^x_{t \wedge S^J}$, $x \in J$, a probability law on $(\Omega, \mathcal{F}_{t \wedge S^J})$.

By the Girsanov theorem, under $\tilde{Q}^x_{t \wedge S^J}$, the process

$M_s - \int_0^s \frac{d \langle M, N \rangle_u}{N_u} = X_s - \int_0^s (b + a\frac{v'}{v})(X_u)\, du$ is an $(\mathcal{F}_s)$-martingale on $[0, t \wedge S^J[$ with $(\int_0^s a(X_u)\,du)$ as increasing process. Consequently, for the canonical realization $(\Omega, \mathcal{F}, \mathcal{F}_t, X_t, T, Q^x)$ of a diffusion process with differential operator

$$L + a \frac{v'}{v} D ,$$

$\tilde{Q}^x_{t \wedge S^J} = Q^x_{t \wedge S^J}$. Therefore diffusion Q^x is locally absolutely continuous with respect to P^x with N as the local density.

Now we fix $J =]\alpha, \beta[$, $\ell < \alpha < \beta \leq r$, we note $S = S^J$ and we assume that equation (2) is disconjugate on $J \cup \{\alpha, \beta\}$ or equivalently (cf. Coppel [5]) it has a solution without zeros on $J \cup \{\alpha, \beta\}$.

Proposition 1.

If the equation (2) is disconjugate on $J \cup \{\alpha, \beta\}$, then for all $x \in J$

$$P^x(e_S) = \frac{v(x)}{v(\alpha)} \frac{q(\beta) - q(x)}{q(\beta) - q(\alpha)} + \frac{v(x)}{v(\beta)} \frac{q(x) - q(\alpha)}{q(\beta) - q(\alpha)} , \quad (3)$$

where v is a nonvanishing on $J \cup \{\alpha, \beta\}$ solution of (2) and

$$q(x) = \int_c^x \exp\left(-\int_c^y \frac{2b}{a}(z)\,dz\right) v^{-2}(y)\,dy , \quad c \in J , \quad (4)$$

is the natural scale for Q^x.

Proof.

Since $\frac{Lv}{v} + g = 0$ on $J \cup \{\alpha, \beta\}$, then $N_{t \wedge S} = \frac{v(X_{t \wedge S})}{v(x)} e_{t \wedge S}$.

Moreover, since

$$Q^x\{\int_0^S N_s^{-2}\, d\langle N\rangle_s < \infty\} = Q^x\{\int_0^S a(\frac{v'}{v})^2 (X_s)\, ds < \infty\} = 1$$

we have also (cf. Kabanov, Liptser, Shiryaev [10]) $Q_S^x \ll P_S^x$. Consequently

$$P^x(e_S) = P^x(\frac{v(x)}{v(X_S)} N_S) = v(x)\, Q^x(v^{-1}(X_S)) .$$

Then, the assertion is proved because (cf. Ikeda and Watanabe [8] for example)

$$Q^x\{X_S = \alpha\} = \frac{q(\beta) - q(x)}{q(\beta) - q(\alpha)} , \quad Q^x\{X_S = \beta\} = \frac{q(x) - q(\alpha)}{q(\beta) - q(\alpha)} .$$

Proposition 2.

If the equation (2) is disconjugate on $J \cup \{\alpha, \beta\}$, then the function $P^x(e_S)$ is the solution of

$$(L+g)\, u = 0 \quad , \quad x \in J \;, \quad u(\alpha) = u(\beta) = 1 .$$

Proof. It is sufficient to show that if v is a nonvanishing solution of (2), then vq also solves (2).

Remark 3.

The method presented in Proposition 1 (change of measure) seems to be very fruitful not only in the case of stopped functionals. We shall give several examples in which we calculate $P^x(e_t)$.

a) Let $L = \frac{1}{2} D^2$ and $g(x) = -\frac{1}{2} x^2 - \frac{1}{2}$. It is easy to see that $v(x) = \exp(\frac{1}{2} x^2)$ satisfies (2) on $\mathbb{R}$ and

$$P^x(e_t) = v(x)\ Q^x(v^{-1}(X_t)) .$$

Since under Q^x X_t is Gaussian $N(e^t x, \frac{1}{2}(e^{2t}-1))$ after simple calculations we obtain the Cameron-Martin formula

$$P^x(\exp(-\frac{1}{2}\int_0^t X_s^2\, ds)) = (\operatorname{ch} t)^{-\frac{1}{2}} \exp(-\frac{1}{2} x^2 \operatorname{th} t) .$$

b) Let $L = \frac{1}{2} D^2$ and $g(x) = -3 \operatorname{th}^2 x + 1$. Because $v(x) = (\operatorname{ch} x)^{-2}$ satisfies (2) on $\mathbb{R}$ we have $P^x(e_t) = v(x)\ Q^x(v^{-1}(X_t))$, where (X_t) under Q^x is a diffusion with the differential operator $\frac{1}{2} D^2 - 2 \operatorname{th} x\, D$. Now it is not difficult to see that

$$Q^x(\operatorname{ch}^2 X_t) = 1 + e^{-2t}(\operatorname{sh}^2 x + \frac{1}{2}(e^{2t}-1))$$

what leads to

$$P^x(\exp(-3\int_0^t \operatorname{th}^2 X_s\, ds)) = (e^t + e^{-t}\operatorname{sh}^2 x + \operatorname{sh} t)\, e^{-2t} \operatorname{ch}^{-2} x .$$

c) If $L = \frac{1}{2} D^2 + \operatorname{th} x\, D$ and $g(x) = -\frac{1}{2} x^2$, then with $v(x) = \operatorname{ch}^{-1} x \exp(\frac{1}{2} x^2)$ we have that under Q^x X_t is Gaussian

$N(e^t x \, , \, \frac{1}{2}(e^{2t}-1))$ and

$$P^x(\exp(-\frac{1}{2}\int_0^t X_s^2\, ds)) = ch^{-1}x \; ch^{-\frac{1}{2}}t \; ch(x \, ch^{-1}t) \; \exp(-\frac{1}{2}\, t - \frac{1}{2}(x^2-1)\, th\, t) \; .$$

d) If $L = \frac{1}{2}(1+x^2)\, D^2 + x D$, then

$$P^x(\exp(-2\int_0^t X_s^2(1+X_s^2)^{-1}\, ds)) = (e^{-t}x^2 - e^{-t} + 2)\; e^{-t}(1+x^2)^{-1} \; .$$

3. NEGATIVE POTENTIALS.

In this section we assume $g = -f$, $f \geq 0$. We start by the following useful

Lemma 4.

Let v be the unique solution of

$$(L-f)\, u = 0 \; , \quad u(c) = 1 \; , \quad u'(c) = 0 \tag{5}$$

on I , where $c \in I$, is such that $f(c) > 0$. Then

(i) $1 + k \leq v \leq \exp(k)$, where

$$k(x) = 2\int_c^x \exp(-\int_c^y \frac{2b}{a}(z)\, dz) \int_c^y \exp(\int_c^z \frac{2b}{a}(t)\, dt) \; \frac{f}{a}(z)\, dz\, dy \; ,$$

(ii) $v' > 0$ on $]c, r[$ and $v' < 0$ on $]\ell, c[$,

(iii) $q(r) < \infty$ and $q(\ell) > -\infty$, where q is defined in (4).

To prove (i) see Ikeda and Watanabe [8] Lemma 3.1. p.363 or Marini and Zezza [15]. For (ii) and (iii) consult Coppel [5] or Marini and Zezza [15].

As a consequence of (3) and the above Lemma we have

Proposition 5.

(i) If $J =]\alpha, r[$, $\ell<\alpha$, and $k(r) = \infty$, then $P^x(e_{S^J})$ is a solution, on J, of $(L-f)u = 0$, $u(\alpha) = 1$.

(ii) If $J =]\alpha, r[$, $\ell<\alpha$, and $k(r) < \infty$, then $P^x(e_{S^J})$ is a solution, on J, of $(L-f)u = 0$, $u(\alpha) = u(r) = 1$.

Similar assertions hold for $J =]\ell, \beta[$, $\beta < r$, and $J =]\ell, r[$.

The main result of this section is the following

Proposition 6.

Let $A_t = \int_0^t f(X_s)\,ds$, $t \in [0,T[$ and let $S = S^J$, where $J =]\alpha, \beta[$, $\ell \le \alpha < \beta \le r$. Then

(i) $P^x\{A_S = \infty\} = 1$, $x \in J$, if and only if $k(\alpha) = \infty$ and $k(\beta) = \infty$.

(ii) $P^x\{A_S < \infty\} = 1$, $x \in J$, if and only if one of the following cases occurs

(a) $k(\alpha) < \infty$, $k(\beta) < \infty$,

(b) $k(\alpha) < \infty$, $k(\beta) = \infty$, $p(\beta) = \infty$,

(c) $k(\alpha) = \infty$, $p(\alpha) = -\infty$, $k(\beta) < \infty$,

where

$$p(x) = \int_c^x \exp\left(-\int_c^y \frac{2b}{a}(z)\,dz\right) dy \quad , \; c \in J \; , \tag{6}$$

is the natural scale for P^x .

(iii) If $k(\alpha) < \infty$, $k(\beta) = \infty$, $p(\beta) < \infty$, then

$$P^x\{A_S < \infty\} = \frac{p(\beta) - p(x)}{p(\beta) - p(\alpha)} \quad , \; x \in J \; .$$

(iv) If $k(\alpha) = \infty$, $p(\alpha) > -\infty$, $k(\beta) < \infty$, then

$$P^x\{A_S < \infty\} = \frac{p(x) - p(\alpha)}{p(\beta) - p(\alpha)} \quad , \; x \in J \; .$$

Proof. We shall consider separately the following cases

1) $k(\alpha) = \infty$, $k(\beta) = \infty$

2) $k(\alpha) < \infty$, $k(\beta) < \infty$

3) $k(\alpha) < \infty$, $k(\beta) = \infty$

4) $k(\alpha) = \infty$, $k(\beta) < \infty$.

Case 1. If $k(\alpha) = k(\beta) = \infty$, then $\alpha = \ell$, $\beta = r$ and $S = T$. Moreover, from Lemma 4 (i) it follows that $v(\ell) = v(r) = \infty$. Consequently we have $P^x\{v(X_T) = \infty\} = 1$. But $v(X_{t \wedge T})\, e_{t \wedge T}$ admits a finite limit what implies $P^x(e_T) = 0$ and the assertion (i) part "if".

Case 2. Now we show that if $k(\alpha) < \infty$ and $k(\beta) < \infty$, then

$P^x\{A_S < \infty\} = 1$. Let $u_\lambda(x) = P^x(\exp(-\lambda A_S))$.

It is sufficient to prove that $u_{0+}(x) = 1$, $x \in J$. Let v_λ , $\lambda > 0$, be a solution of (5) with λf instead of f . It is clear that $1 + \lambda k \leq v_\lambda \leq \exp(\lambda k)$ (see Lemma 4(i)). Therefore the conclusion follows from (3).

Note that since u_λ , $\lambda > 0$, is a solution of $(L - \lambda f)u = 0$, $u(\alpha) = u(\beta) = 1$, and the distribution of A_S lies in the Bondensson class (cf. Bondensson [1]) an analytic continuation argument (cf. Kent [12]) yields an alternative direct proof.

Case 3. If $k(\beta) = \infty$, then $\beta = r$ and $J =]\alpha, r[$. Let $S^a = \inf\{t : X_t = a\}$, $\ell \leq a \leq r$. For $\alpha < a < r$ we have

$$\{A_S < \infty\} \cap \{S^\alpha < S^r\} = \{A_S < \infty\} \cap \{S^a < S^\alpha < S^r\} \cup$$

$$\cup \{A_{S^\alpha \wedge S^a} < \infty\} \cap \{S^\alpha < S^a \wedge S^r\} .$$

Moreover, because $k(\alpha) + k(a) < \infty$ we know also that

$$P^x\{A_{S^\alpha \wedge S^a} < \infty\} = 1$$

what leads to

$$P^x\{A_S < \infty , S^\alpha < S^r\} = P^x\{S^\alpha < S^r\} , x \in J . \tag{7}$$

On the other hand $P^x\{v(X_S) = \infty , S^\alpha > S^r\} = P^x\{S^\alpha > S^r\}$, $x \in J$, and analogously to "case 1" we obtain $P^x(e_S , S^\alpha > S^r) = 0$, $x \in J$. Consequently, for all $x \in J$, $P^x\{A_S = \infty , S^\alpha > S^r\} = P^x\{S^\alpha > S^r\}$ what together with (7) yields

$$P^x\{A_S < \infty\} = P^x\{S^\alpha < S^r\} .$$

To conclude note that if $p(r) = \infty$, then $P^x\{S^\alpha < S^r\} = 1$ and if $p(r) < \infty$, then

$$P^x\{S^\alpha < S^r\} = \frac{p(r) - p(x)}{p(r) - p(\alpha)} .$$

Case 4. It is sufficient to interchange the roles of α and β to prove that $P^x\{A_S < \infty\} = P^x\{S^\beta < S^\ell\}$.

Remark 7.

If $f \equiv 1$ and $J = I$, then parts "if" of (i) and (ii-a) are reduced to the classical Feller's test for explosions (cf. Feller [6], Mc Kean Jr. [16]). For the transient case see also Khas'minskii [13] .

4. MORE ON THE SCHRODINGER EQUATION.

In this section we study the Schrödinger equation (2) on an interval $J =]\alpha, r[$, $\ell < \alpha$. The case $J =]\ell, \beta[$, $\beta < r$, may be treated analogously. First we state a result which is standard in the theory of linear differential equations (cf. Coppel [5]).

Lemma 8.

Suppose the equation (2) has a solution nonvanishing on $J \cup \{\alpha\}$. Then there exists a nonvanishing solution v such that

$$\int_c^r \exp\left(-\int_c^y \frac{2b}{a}(z)\, dz\right) v^{-2}(y)\, dy = \infty , \quad c \in J . \tag{8}$$

Moreover, for any linearly independent nonvanishing solution u

$$\int_c^r \exp\left(-\int_c^y \frac{2b}{a}(z)\,dz\right) u^{-2}(y)\,dy < \infty \ , \quad c \in J \ . \tag{9}$$

Obviously the solution v is uniquely determined up to a constant factor. It is called the principal solution of (2). Now we denote by v the principal solution of (2) normalized by $v(\alpha) = 1$. We have the following probabilistic characterization of v .

Proposition 9.

Suppose the equation (2) has a solution nonvanishing on $J \cup \{\alpha\}$.

(i) The normalized principal solution is given by

$$v(x) = P^x(e_{S^\alpha}, \ S^\alpha < S^r) \ ,$$

where $S^a = \inf\{t : X_t = a\}$, $\ell \le a \le r$. If $p(r)$ given by (6) is infinite, then $v(x) = p^x(e_{S^\alpha})$.

(ii) For any normalized nonvanishing on $J \cup \{\alpha\}$ solution u of (2) which is linearly independent of v , one has $v < u$ on J .

Proof.

(i) If v is the principal solution of (2) then (8) holds and for all $x \in J$ $Q^x\{S^\alpha < S^r\} = 1$. Moreover, by Proposition 6 (ii-b) $Q^x_{S^J} \ll P^x_{S^J}$ and

$$P^x(e_{S^\alpha}, S^\alpha < S^r) = P^x(e_{S^J}, S^\alpha < S^r) = P^x\left(\frac{v(x)}{v(X_{S^J})} N_{S^J}, S^\alpha < S^r\right) =$$

$= v(x)\, Q^x\{S^\alpha < S^r\} = v(x)$. If $p(r) = \infty$, then $P^x\{S^\alpha < S^r\} = 1$, $x \in J$.

(ii) Let u be a normalized, linearly independent of v, positive solution of (2). Let q be given by (4) with u instead of v. Since (9) holds, then $q(r) < \infty$ and

$$Q^x\{S^\alpha < S^r\} = \frac{q(r) - q(x)}{q(r) - q(\alpha)} < 1$$

for all $x \in J$. Moreover, let $J_n =]\alpha, r_n[$, $r_n < r$, $\lim_{n\to\infty} r_n = r$. By the Fatou lemma we have

$$v(x) = P^x(e_{S^\alpha}, S^\alpha < S^r) \le \lim_{n\to\infty} P^x(e_{S^{J_n}}, S^\alpha < S^{r_n}) = u(x)\, Q^x\{S^\alpha < S^r\} < u(x)$$

for all $x \in J$ what yields (ii).

Remark 10.

For $L = \frac{1}{2} D^2$ analogous assertion has been obtained by Chang and Varadhan [4]. For general L the above Proposition was announced without proof by Khas'minskii [13]. As far as we know there has not appeared in the literature any proof of this result.

5. NONVANISHING SOLUTIONS AND EXPONENTIAL MOMENTS OF EXIT TIMES.

Conditions for existence of a nonvanishing solution of the Schrödinger equation has been studied by many authors.

Probabilistic approach (cf. Chung [2], Chung and Rao [3], Khas'minskii [13,14]) roughly speaking gives finiteness of $P^x(e_{S^J})$ as a necessary and sufficient condition.

The classical approach (cf. Coppel [5]) yieldsmore explicit sufficient conditions. For the reader's convenience we present here the following

Lemma 11.

Let $J =]\alpha,\beta[$, $\ell < \alpha < \beta < r$. If $g \geq 0$ and

$$\int_\alpha^\beta \frac{g}{p'a}(x)\,dx \leq \frac{2}{p(\beta) - p(\alpha)} ,$$

then the equation (2) has a nonvanishing solution (is disconjugate) on $J \cup \{\alpha,\beta\}$.

For the proof see Coppel [5] p.64.

Remark 12.

Consider the eigenvalue problem

$$(L+\lambda)u = 0 \quad \text{in} \quad J =]\alpha,\beta[\quad , \ell < \alpha < \beta < r$$
$$u(\alpha) = u(\beta) = 1 \quad . \tag{10}$$

It is well known that the principal eigenvalue λ_o (i.e. such that $\mathrm{Re}\,\lambda \geq \lambda_o$ for any eigenvalue λ of (10)) is given by

$$\lambda_o = \sup\{\lambda \geq 0 \; ; \; \sup_{x \in J} P^x(\exp(\lambda S^J)) < \infty \}$$

(cf. Friedman [7]). Using the above Lemma we obtain the following minoration

$$\lambda_o \geq \frac{2}{(p(\beta)-p(\alpha)) \int_\alpha^\beta (p'a)^{-1}(x)\,dx} \quad . \tag{11}$$

Note that from the above inequality one can obtain R. Durrett's result extensively used by Chung and Varadhan in [4]. It is also clear that if $p(r) < \infty$ and $\int_{\alpha}^{r} (p'a)^{-1}(x)\,dx < \infty$, then for $J =]\alpha, r[$, $\ell < \alpha$, $P^x(\exp(\lambda S^J)) < \infty$ for all $x \in J$ and

$$\lambda \leq \frac{2}{(p(r)-p(\alpha)) \int_{\alpha}^{r} (p'a)^{-1}(x)\,dx} .$$

Now let $F(x,y) = \int_x^y \sigma^{-1}(z)\,dz$, where $a = \sigma^2$, $\sigma > 0$, and let $L = \frac{1}{2}(\sigma D)^2$. Since $F(x, X_t)$ is a Brownian motion starting at zero it is clear that (cf. Itô and Mc Kean Jr. [9])

$$P^x(\exp(-\lambda S^J)) = \frac{\operatorname{ch}\left(\sqrt{2\lambda}\ \frac{F(x,\beta)-F(\alpha,x)}{2}\right)}{\operatorname{ch}\left(\sqrt{2\lambda}\ \frac{F(\alpha,\beta)}{2}\right)} , \qquad (12)$$

where $J =]\alpha, \beta[$, $\ell < \alpha < \beta < r$. Moreover, for $0 \leq \lambda < \frac{\pi^2}{2F^2(\alpha,\beta)}$

$$P^x(\exp(\lambda S^J)) = \frac{\cos\left(\sqrt{2\lambda}\ \frac{F(x,\beta)-F(\alpha,x)}{2}\right)}{\cos\left(\sqrt{2\lambda}\ \frac{F(\alpha,\beta)}{2}\right)}$$

what shows that minoration (11) can be improved. Finally, note that if $F(\ell, r) < \infty$, then $P^x(S^I = T < \infty) = 1$ and $P^x(\exp(-\lambda T))$ is given by (12) with ℓ and r instead of α and β, respectively.

Acknowledgement. I wish to thank Professor A. Le Breton not only for a number of helpful conversations but also for his hospitality and assistance during my stay in Grenoble. I express my hearty thanks to Professor B. Maisonneuve for suggesting many improvements of a previous version of the paper. Thanks are also due to the referee for his useful comments.

REFERENCES

[1] L. Bondensson, Classes of infinitely divisible distributions and densities. Z. Wahrsch. verw. Gebiete 57 (1981), 39-71.

[2] K.L. Chung, Lectures from Markov Processes to Brownian Motion, Springer-Verlag, New-York, Heidelberg, Berlin, 1982.

[3] K.L. Chung and K.M. Rao, Sur la théorie du potentiel avec la fonctionnelle de Feynman-Kac, C.R. Acad. Sc. Paris, t.290, Série A, (1980), 629-631.

[4] K.L. Chung and S.R.S. Varadhan, Kac functional and Schrödinger equation, Studia Math. 68 (1980), 249-260.

[5] W.A. Coppel, Disconjugacy, Lecture Notes in Math. 220, Springer-Verlag, Berlin, Heidelberg, New-York, 1971.

[6] W. Feller, The parabolic differential equations and the associated semigroups of transformations, Annals of Mathematics 55 (1952), 468-519.

[7] A. Friedman, Stochastic Differential Equations and Applications, Volume 2, Academic Press, New-York, San Francisco, London 1976.

[8] N. Ikeda and S. Watanabe, Stochastic Differential Equations and Diffusion Processes, North-Holland Publishing Company, Amsterdam, Oxford, New-York; Kodansha LTD. Tokyo, 1981.

[9] K. Itô and H.P. Mc Kean Jr., Diffusion Processes and their Sample Paths, Springer-Verlag, Berlin, Heidelberg, New-York, 1965.

[10] Ju.M.Kabanov, R.S. Liptser, A.N. Shiryaev, Absolute continuity and singularity of locally absolutely continuous probability distributions. I ; II , Mathem. Sbornik V.107 (149), N°3(11), (1978), 364-415 ; 108(150), N° 1 (1979), 32-61.

[11] M. Kac, On some connections between probability theory and differential and integral equations, Proc. Second Berkeley Symposium (1951) 189-215.

[12] J.T. Kent, The spectral decomposition of a diffusion hitting time, The Annals of Probability, Vol.10, N° 1, (1982), 207-219.

[13] R.Z. Khas'minskii, Probability distribution for functionals of trajectories of a diffusion type stochastic process, Dokl. Akad. Nauk SSSR, 104, 1, (1955), 22-25.

[14] R.Z. Khas'minskii, On positive solutions of the equation Uu + Vu = 0 , Theory of Probability and its Applications, Vol. IV, N° 3, (1959), 309-318.

[15] M. Marini and P. Zezza, On the asymptotic behaviour of the solutions of a class of second-order linear differential equations, J. of Differential Equations 28, (1978) 1-17.

[16] H.P. Mc Kean Jr., Stochastic Integrals, Academic Press. New-York and London, 1969.

DIVERGENCE, CONVERGENCE AND MOMENTS OF SOME INTEGRAL FUNCTIONALS OF DIFFUSIONS

M. MUSIELA
Laboratoire IMAG-TIM 3
BP : 68
38402 SAINT-MARTIN D'HERES CEDEX
FRANCE

ABSTRACT

In this paper we study some integral functionals of diffusions. We obtain criteria for their divergence and their convergence and we investigate the existence of their moments.

AMS 1980 subject classification. Primary 60 J 60, Secondary 60 J 55.

Key words and phrases. diffusion processes, additive functionals, hitting times.

1. INTRODUCTION

Let D be an open, connected subset of $\mathbb{R}^d$ and let $D_\delta = D \cup \{\delta\}$ with $\delta = \infty(D)$ be the one point compactification of D.

Let Ω be the set of all continuous functions ω of $\mathbb{R}_+$ into D_δ which are stopped at the first time of hitting δ : if $\omega(t) = \delta$, then $\omega(t') = \delta$ for $t' \geq t$. Denote $(X_t)_{t\geq 0}$ the coordinate process, $F_t = \sigma(X_s \; ; \; s \leq t)$, $F = F_\infty$. Moreover, let $W = \mathbb{R}_+ \times \Omega$, $Y_t(c, \omega) = (c+t, \omega(t))$, $G_t = \sigma(Y_s \; ; \; s \leq t)$ $(= \sigma(Y_0, X_s, s \leq t))$, $G = G_\infty$.

If U is an open (for the relative topology) subset of $[c,\infty[\times D$ for some $c \in \mathbb{R}_+$ we set $S(U) = \inf \{s \geq 0 \; ; \; Y_s \notin U\}$ $(\inf \emptyset = \infty)$. $S(U)$ is a stopping time of (G_t). The lifetime of (Y_t) is $S = S(\mathbb{R}_+ \times D) = \inf\{s \geq 0 \; ; \; X_s = \delta\}$.

Assume that the functions $b : \mathbb{R}_+ \times D \to \mathbb{R}^d$ and $\sigma : \mathbb{R}_+ \times D \to \mathbb{R}^d \otimes \mathbb{R}^r$ are continuous and that for any $R > 0$ there exists a constant $c_R > 0$ such that :

$$|b(t,x) - b(t,y)| + |\sigma(t,x)-\sigma(t,y)| \leq c_R \; |x-y|$$

for all $t \in \mathbb{R}_+$ $x,y \in D$ provided $|x| + |y| \leq R$.

For the vectors (matrices) x,y the symbols $|x|$ and $x\cdot y$ stand for the Euclidean norm and the Euclidean scalar product, respectively. The symbol x^* stands for the transpose of x.

Let $a = \sigma\sigma^*$ and let Λ be the second order differential operator on $\mathbb{R}_+ \times D$:

$$\Lambda = \mathbb{D}_t + L, \quad L = \frac{1}{2}\, a\cdot\mathbb{D}_x^2 + b\cdot\mathbb{D}_x$$

For $y \in \mathbb{R}_+ \times D_\delta$, let Q_y be the probability on (W,G) such that $Q_y\{Y_0 = y\} = 1$ and $(g(Y_t) - \int_0^t \Lambda\, g(Y_s)ds)$ is a (G_t, Q_y)-local martingale on $[0, S[$ for each $g \in C^{1,2}(\mathbb{R}_+ \times D)$. $(W, G, G_t, X_t, S, Q_y$

is the canonical realization of an inhomogeneous strongly Markov diffusion process on D generated by the differential operator Λ. If the functions $a(t,\cdot)$ and $b(t,\cdot)$ do not depend on t, then (W, G, G_t, X_t, S, P_x) with $P_x = Q_{0,x}$, $x \in D_\delta$, is a homogeneous diffusion on D with infinitesimal generator L.

Existence and uniqueness of Q_y, for all $y \in \mathbb{R}_+ \times D_\delta$, follow from the classical results on stochastic differential equations (cf. Ikeda and Watanabe [6] and Narita [15] for example) and Doob's representation of stopped diffusions (cf. Dynkin [5]).

Let the function $f : \mathbb{R}_+ \times D \to \mathbb{R}_+$ be Borel measurable. In this paper we investigate the integral :

$$(1.1) \qquad I_t = \int_0^{t \wedge S} f(Y_s)\, ds \quad , \quad 0 \le t \le \infty.$$

We obtain criteria ensuring its a.s. divergence or its a.s. convergence. We give also sufficient conditions for the existence of exponential moments of the integral. Criteria which are formulated in terms of differential inequalities generalize those of Stroock and Varadhan [19] and Narita [15–17] and relate the problem to a purely analytical one. The radial criteria generalize the classical Khas'minskii's [9,10] tests of non explosion or explosion (see also Azencott [1], Bhattacharya [2,3] Mc Kean Jr [13]).

We shall use the following notation. For a Q_y-integrable function ξ on W the symbol $Q_y\xi$ stands for the expectation $\int_W \xi \, d\, Q_y$. For a real function f defined on a non empty set B $\sup_B f$ and $\inf_B f$ stand respectively for the supremum and the infinmum of f over B.

2. GENERAL CRITERIA

In this section we discuss a few simple criteria which relate the problem to a purely analytical one.

Let Δ be an open connected non empty subset of $[c, \infty[\times D$ for some $c \in \mathbb{R}_+$. Moreover, let U be a compact subset of Δ such that $\inf_U f > 0$ ($\inf \emptyset = \infty$). We have the following

Theorem 2.1.

Assume that there exist $\lambda > 0$ and a positive function $u \in C^{1,2}(\Delta)$ such that $\Lambda u \leq \lambda f \phi \circ u$ on $\Delta - U$, where $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ is increasing (= nondecreasing), differentiable, such that $\int_0^\infty \frac{dt}{1+\phi(t)} = \infty$.

Then for $y \in \Delta$ one has Q_y a.s.

$$\{\overline{\lim_{t\uparrow S(\Delta)}}\, u(Y_t) = \infty\} \subset \{I_{S(\Delta)} = \infty\}.$$

In particular if $\lim_{\Delta \ni z \to \infty(\Delta)} u(z) = \infty$, then

$$Q_y\{I_{S(\Delta)} = \infty\} = 1 .$$

Proof

a) $\phi \equiv 1$, $U = \emptyset$. We may assume $u_0 = \inf_\Delta u > 0$. Define :

$$(2.1) \qquad N_t^u = u(Y_t) \exp\left(-\int_0^t \frac{\Lambda u}{u}(Y_s)\, ds\right) .$$

Under Q_y, N^u is a positive local martingale on $[0, S(\Delta)[$ (cf. Maisonneuve [12]). Thus N_t^u tends to a finite limit a.s. as $t \to S(\Delta)$. Since

$$N_t^u \geq u(Y_t) \exp\left(-\frac{\lambda}{u_0} I_t\right)$$

we have $\lim_{t\uparrow S(\Delta)} u(Y_t) < \infty$ a.s. on $\{I_{S(\Delta)} < \infty\}$.

b) If $\phi \equiv 1$ and $U \neq \emptyset$, one has

$$\Lambda u \leq \frac{\max\{1, \sup_U \Lambda u\}}{\inf_U f} f \quad \text{on } U$$

and the conclusion follows form case a).

c) In the general case, consider $\phi = \int_0^{\cdot} \frac{dt}{1+\phi(t)}$.

A simple computation shows that :

$$\Lambda\ \phi \circ u \le \frac{\Lambda u}{1 + \phi \circ u} \le \lambda f \frac{\phi \circ u}{1 + \phi \circ u} \le \lambda f \quad \text{on} \quad \Delta - U .$$

Since $\phi(\infty) = \infty$, the proof follows from case b).

Now we discuss the convergence of the integral $I_{S(\Delta)}$.

Theorem 2.2.

Assume that there exist $\lambda > 0$ and $u \in C^{1,2}(\Delta)$ such that $u > 0$ and $\Lambda u \ge \lambda f u$ on Δ. Then for $y \in \Delta$

$$Q_y\{I_{S(\Delta)} < \infty\} \ge \frac{u(y)}{\sup_\Delta u} .$$

Proof

Let (S_n) be an increasing sequence of stopping times reducing N^u, given by (2.1), under Q_y. Then

$$u(y) = Q_y N^u_{S_n} \le (\sup_\Delta u)\, Q_y \exp(-\lambda I_{S_n})$$

and

$$\frac{u(y)}{\sup_\Delta u} \le Q_y \exp(-\lambda I_{S_n}) .$$

The assertion follows by taking lim as $n \to \infty$.

Concerning the moments we have

Theorem 2.3.

(i) If there exist $\lambda > 0$ and $u \in C^{1,2}(\Delta)$ such that $\Lambda u \ge \lambda f$, then

$$Q_y I_{S(\Delta)} \le \lambda^{-1} (\sup_\Delta u - u(y)) , \; y \in \Delta .$$

(ii) If $\alpha = \sup_{y\in\Delta} Q_y\, I_{S(\Delta)} < 1$, then

$$Q_y \exp(I_{S(\Delta)}) \le \frac{1}{1-\alpha} \quad (< \infty) , \; y\in\Delta .$$

Proof

Let (S_n) reduce the local martingale $(u(Y_t) - \int_0^t \Lambda\, u(Y_s)ds$, $t < S(\Delta))$ under Q_y. We have $Q_y\, u(Y_{S_n}) - u(y) \ge \lambda\, Q_y\, I_{S_n}$, $y \in \Delta$ and hence $\lambda\, Q_y\, I_{S_n} \le \sup_\Delta u - u(y)$. The second statement follows from Khas'minskii's criterion of solvability [8].

Let π be the projection of $\mathbb{R}_+ \times D$ on D.

Corollary 2.1.

If $\pi(\Delta)$ is bounded and if there exists $z \in \mathbb{R}^d$ such that $\frac{1}{2}\, z\cdot az+z\cdot b \ge \lambda$ on Δ for some $\lambda > 0$, then $Q_y\, S(\Delta) < \infty$, $y \in \Delta$.

Proof

The function u defined on Δ by $u(s,x) = \exp(z\cdot x)$ satisfies

$$\Lambda\, u = (\frac{1}{2}\, z\cdot az + z\cdot b)u \ge \lambda \inf_\Delta u > 0 \text{ on } \Delta.$$

The conclusion follows from Theorem 2.3(i).

We use the strong Markov property of (Y_t) under Q_y to prove the following result.

Theorem 2.4.

Let $q(y) = Q_y\{I_{S(\Delta)} < \infty\}$, $y \in \Delta$. Assume that

a) $Q_y\{S(U) < \infty\} = 1$ for each $y \in U$ and each open subset U of Δ such that $\pi(U)$ is bounded and $\overline{U} \subset \Delta$,

b) $\lim_{\pi(\Delta)\ni\pi(y)\to\infty(\pi(\Delta))} q(y) = 1.$

Then $q \equiv 1$ on Δ.

Proof

Let (U_n) be a increasing sequence of open subsets of Δ such that $\bigcup_n \bar{U}_n = \Delta$ and for each $n \in \mathbb{N}$ $\pi(U_n)$ is bounded. Let $y \in \Delta$. Since $Q_y\{S(U_n) < \infty\} = 1$, we have $q(y) = Q_y\, q(Y_{S(U_n)})$ by the strong Markov property of (Y_t) under Q_y. But $q(Y_{S(U_n)}) \xrightarrow{n} 1$ Q_y a.s. by assumption b) and the statement follows.

We finish this section by criteria of divergence or convergence obtained via change of measure. Let $u \in C^{1,2}(\mathbb{R}_+ \times D)$ be real non vanishing. Let $c = \frac{\mathbb{D}_x u}{u}$ and $Z_t = \frac{1}{u(Y_0)} N_t^u$, where N is given by (2.1). Under Q_y one has a.s.

$$Z_t = \exp\left(\int_0^t c(Y_s)\cdot dM_s - \frac{1}{2}\int_0^t c\cdot ac(Y_s)ds\right), \quad t < S,$$

where $(M_t \equiv X_t - \int_0^t b(Y_s)ds)$ is a local martingale on $[0,S[$ with quadratic variation $(\int_0^t a(Y_s)ds)$. Let $\Delta_n = [0,n[\times D_n$, where D_n is an increasing sequence of bounded open subsets of D such that $\bigcup_n \bar{D}_n = D$. It follows from Girsanov's theorem that the process $(X_t - \int_0^t (b+ac)(Y_s)ds)$ stopped at $S(\Delta_n)$ is a bounded martingale under $Z_{S(\Delta_n)} Q_y$.

This proves that the diffusion measures Q_y^u associated with the operator $\Lambda + ac\cdot\mathbb{D}_x$ satisfy

$$Q_y^u = Z_{S(\Delta_n)}\, Q_y \text{ on } G_{S(\Delta_n)}.$$

Note that if we set $Z_S = \overline{\lim_{t\uparrow S}}\, Z_t$ we have the following decomposition :

$$Q_y^u = Z_\tau\, Q_y + Q_y^u\{\{Z_\tau = \infty\} \cap \cdot\} \quad \text{on } G_\tau$$

for each stopping time $\tau \leq S$. Moreover, in this decomposition we can replace $\{Z_\tau = \infty\}$ by $\{\int_0^\tau c\cdot ac(Y_t)dt = \infty\}$ (see Kabanov, Liptser, Shiryaev [7] for details).

Theorem 2.5.

Suppose that $\frac{\Lambda u}{u} \leq \lambda f$ on Δ for some $\lambda > 0$ and that $\lim_{n\to\infty} Q_y^u \frac{1}{u}(Y_{S_n}) = 0$ for some sequence (S_n) of stopping times such that $S_n < S(\Delta)$, $S_n \uparrow S(\Delta)$. Then $Q_y\{I_{S(\Delta)} = \infty\} = 1$, $y \in \Delta$.

Proof

Taking $\lim_{n\to\infty}$ in the inequality

$Q_y \exp(-\lambda I_{S(\Delta)}) \leq Q_y \frac{N_{S_n}}{u(Y_{S_n})} = u(y)\, Q_y^u \frac{1}{u}\,(Y_{S_n})$ yields the conclusion.

Theorem 2.6.

Suppose that $\frac{\Lambda u}{u} \geq \lambda f$ on Δ for some $\lambda > 0$. Then for any $y \in \Delta$ and any sequence (S_n) of stopping times such that $S_n < S(\Delta)$, $S_n \uparrow S(\Delta)$, we have :

$$Q_y\{I_{S(\Delta)} < \infty\} \geq \overline{\lim_{n\to\infty}}\, Q_y^u \frac{u(y)}{u(Y_{S_n})} .$$

Proof

Note that

$$Q_y\{I_{S(\Delta)} < \infty\} \geq Q_y \exp(-\lambda I_{S(\Delta)}) =$$

$$\lim_{n\to\infty} Q_y \exp(-\lambda I_{S_n}) \geq \overline{\lim_{n\to\infty}}\, Q_y \frac{N_{S_n}^u}{u(Y_{S_n})} = \overline{\lim_{n\to\infty}}\, Q_y^u \frac{u(y)}{u(Y_{S_n})}$$

what proves the assertion.

3. RADIAL CRITERIA

Now we shall use results of Section 2 in order to obtain more explicit criteria of divergence or convergence of the integral I_S, where $S = \inf\{t \geq 0 ; X_t = \delta\}$. We shall study also the integral $I_{T_a \wedge T_b}$, where $T_a = \inf\{t \geq 0 ; |X_t - z| = (2a)^{1/2}\}$, $z \in \mathbb{R}^d$, $0 \leq a < b \leq \infty$.

We consider only the case $D = \mathbb{R}^d$. We assume that the continuous functions $b : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^r$ are uniformly in t locally Lipschitz continuous in x and that for all $x \in \mathbb{R}^d$ the functions $a(\cdot,x)$ and $b(\cdot,x)$ are bounded ($a=\sigma\sigma^*$).

Let $z \in \mathbb{R}^d$, $r > 0$. We assume also that there are continuous fuctions :

$$\bar{\alpha} : [r, \infty[\to]0, \infty[$$
$$\bar{\beta} : [r, \infty[\to \mathbb{R}$$
$$\underline{\gamma} : [r, \infty[\to \mathbb{R}_+$$

such that for all $|x| \geq (2r)^{1/2}$, $t \in \mathbb{R}_+$

$$x \cdot a(t,x+z)x \leq \bar{\alpha}\left(\frac{|x|^2}{2}\right)$$

$$\text{(3.1)} \quad \operatorname{tr} a(t,x+z) + 2x\cdot b(t,x+z) \leq x\cdot a(t,x+z)x\, \bar{\beta}\left(\frac{|x|^2}{2}\right),$$

$$\underline{\gamma}\left(\frac{|x|^2}{2}\right) \leq f(t,x+z).$$

Define for $t \geq r$

$$\bar{e}(t) = \exp\left(-\int_r^t \bar{\beta}(u)du\right), \quad \bar{m}(t) = 2\int_r^t \frac{\underline{\gamma}}{\bar{e}\,\bar{\alpha}}(u)\, du,$$

$$\bar{s}(t) = \int_r^t \bar{e}(u)\, du, \quad \bar{k}(t) = \int_r^t \bar{e}\,\bar{m}(u)\, du.$$

Let $B(z,r) = \{x : |x-z| < (2r)^{1/2}\}$. Our following result reduces to Khas'minskii's test of non explosion (cf. Stroock, Varadhan [19] for example) if $f \equiv 1$.

<u>Theorem 3.1.</u>

If $\bar{k}(\infty) = \infty$ and $\inf\limits_{\mathbb{R}_+ \times B(z,r)} f > 0$, then for $y \in \mathbb{R}_+ \times \mathbb{R}^d$

$$Q_y\{I_S = \infty\} = 1.$$

Proof

Let $u \in C^2(\mathbb{R}^d)$ be such that $u(x) = \bar{k}(\frac{|x-z|^2}{2})$ on $B^c(z,r)$. It is easy to see that on $\mathbb{R}_+ \times B^c(z,r)$

$$\Lambda u(t,x) = \frac{1}{2}(x-z)\cdot a(t,x)(x-z)\ \bar{k}''\ (\frac{|x-z|^2}{2}) +$$

$$(\frac{1}{2}\operatorname{tr} a(t,x) + (x-z)\cdot b(t,x))\ \bar{k}'\ (\frac{|x-z|^2}{2}) \le$$

$$\frac{(x-z)\cdot a(t,x)(x-z)}{\bar{\alpha}\ (\frac{|x-z|^2}{2})}\ \gamma(\frac{|x-z|^2}{2}) \le f(t,x).$$

Moreover, since $\sup\limits_{\mathbb{R}_+ \times B(z,r)} \Lambda u = c < \infty$ then $\Lambda u \le \lambda f$

on $\mathbb{R}_+ \times B(z,r)$ with $\lambda = \dfrac{\max\{1,c\}}{\inf\limits_{\mathbb{R}_+ \times B(z,r)} f}$

Therefore $\Lambda u \le \max\{1,\lambda\} f$ on $\mathbb{R}_+ \times \mathbb{R}^d$ and $\lim\limits_{|x|\to\infty} u(x) = \infty$ what, using Theorem 2.1 , yields the conclusion.

As before we fix $z \in \mathbb{R}^d$, $r > 0$ and we assume that there are continuous functions :

$\underline{\alpha} : [r, \infty[\to]0, \infty[$,
$\underline{\beta} : [r, \infty[\to \mathbb{R}$,
$\bar{\gamma} : [r, \infty[\to \mathbb{R}_+$

such that for all $|x| \ge (2r)^{1/2}$, $t \in \mathbb{R}_+$,

$$\underline{\alpha}(\frac{|x|^2}{2}) \le x\cdot a(t,x+z)x ,$$

(3.2) $$x\cdot a(t,x+z)x\ \underline{\beta}(\frac{|x|^2}{2}) \le \operatorname{tr} a(t,x+z) + 2x\cdot b(t,x+z),$$

$$f(t,x+z) \le \bar{\gamma}\ (\frac{|x|^2}{2}) .$$

We define for $t \geq r$ the functions

$$\underline{e}(t) = \exp(-\int_r^t \underline{\beta}(u)du), \quad \underline{m}(t) = 2\int_r^t \frac{\overline{\gamma}}{\underline{e}\,\underline{\alpha}}(u)\,du$$

$$\underline{s}(t) = \int_r^t \underline{e}(u)\,du, \quad \underline{k}(t) = \int_r^t \underline{e}\,\underline{m}(u)\,du,$$

and

$$\overline{\delta}(t) = \frac{\overline{e}\,\overline{m}}{\underline{e}}(t), \quad \underline{\delta}(t) = \frac{\underline{e}\,\underline{m}}{\overline{e}}(t).$$

Now we introduce an additional assumption which will be used in the following study.

(A) For each open subset U of $\mathbb{R}_+ \times \mathbb{R}^d$ such that $\pi(U)$ is bounded and for each $y \in U$ $Q_y\, S(U) < \infty$.

A simple sufficient condition for (A) to hold is given in Corollary 2.1. Therefore if the functions a and b do not depend on t and a(x) is positive definite for all $x \in \mathbb{R}^d$, then (A) holds. If moreover $f \equiv 1$, then the next result reduces to Khas'miskii's test of explosion (cf. Mc Kean [13] for example).

<u>Theorem 3.2.</u>

Assume (A) holds. If $\underline{k}(\infty) < \infty$, then for $y \in \mathbb{R}_+ \times \mathbb{R}^d$

$$Q_y\{I_S < \infty\} = 1.$$

<u>Proof</u>

Let $u \in C^2(\mathbb{R}^d)$ be such that $u(x) = \underline{k}\left(\frac{|x-z|^2}{2}\right)$ on $B^c(z,r)$. Therefore on $\mathbb{R}_+ \times B^c(z,r)$ we have

$$\Lambda\, u(t,x) = \frac{1}{2}(x-z)\cdot a(t,x)(x-z)\,\underline{k}''\left(\frac{|x-z|^2}{2}\right) +$$

$$\left(\frac{1}{2}\operatorname{tr} a(t,x) + (x-z)\cdot b(t,x)\right)\underline{k}'\left(\frac{|x-z|^2}{2}\right) \geq$$

$$\frac{1}{2}(x-z)\cdot a(t,x)(x-z)\left(\underline{k}''\left(\frac{|x-z|^2}{2}\right) + \underline{\beta}\left(\frac{|x-z|^2}{2}\right) k'\left(\frac{|x-z|^2}{2}\right)\right) \geq$$

$$f(t,x) \geq \frac{\underline{k}\left(\frac{|x-z|^2}{2}\right)}{\underline{k}(\infty)} f(t,x) = \frac{1}{\underline{k}(\infty)} f(t,x)\ u(x).$$

Let $T_a = \inf\{t \geq 0\ ;\ |X_t - z| = (2a)^{1/2}\}$, $z \in \mathbb{R}^d$, $a \in \mathbb{R}_+$. Since (A) holds then for $y = (t,x)$ with $t \in \mathbb{R}_+$ and $(2r)^{1/2} < |x-z| < (2R)^{1/2} < \infty$ we have

$$Q_y\ u(X_{T_r \wedge T_R}) \exp\left(-\frac{I_{T_r \wedge T_R}}{\underline{k}(\infty)}\right) \geq u(x).$$

But

$$Q_y\ u(X_{T_r \wedge T_R}) \exp\left(-\frac{I_{T_r \wedge T_R}}{\underline{k}(\infty)}\right) =$$

$$Q_y\left(u(X_{T_r}) \exp\left(-\frac{I_{T_r}}{\underline{k}(\infty)}\right)\ ;\ T_r < T_R\right) +$$

$$Q_y\left(u(X_{T_R}) \exp\left(-\frac{I_{T_R}}{\underline{k}(\infty)}\right)\ ;\ T_R < T_r\right) =$$

$$\underline{k}(R)\ Q_y\left(\exp\left(-\frac{I_{T_R}}{\underline{k}(\infty)}\right)\ ;\ T_R < T_r\right)$$

Consequently on $\mathbb{R}_+ \times B^c(z,r)$

$$\underline{k}(\infty)\ Q_y\left(\exp\left(-\frac{I_S}{\underline{k}(\infty)}\right)\ ;\ S < T_r\right) \geq \underline{k}\left(\frac{|x-z|^2}{2}\right)$$

This proves that

$$\lim_{|\pi y| \to \infty} Q_y\{I_S < \infty\} \geq \lim_{|\pi y| \to \infty} Q_y\left(\exp\left(-\frac{I_S}{\underline{k}(\infty)}\right)\ ;\ S < T_r\right) \geq$$

$$\lim_{|x| \to \infty} \frac{\underline{k}\left(\frac{|x-z|^2}{2}\right)}{\underline{k}(\infty)} = 1$$

and the assertion follows from Theorem2.4.

From now until the end of this section we suppose that the continuous functions $\underline{\alpha}$, $\underline{\beta}$, $\overline{\gamma}$ and $\overline{\alpha}$, $\overline{\beta}$, $\underline{\gamma}$ satisfy (3.1) and (3.2), respectively, on the interval $]0, \infty[$. We shall study the integral $I_{T_a \wedge T_b}$, where $T_a = \inf\{t \geq 0 ; |X_t - z| = (2a)^{1/2}\}$, $z \in \mathbb{R}^d$, $0 \leq a < b \leq \infty$. Note that $T_\infty = \lim_{b \to \infty} T_b = S$ and $T_0 = \inf\{t \geq 0 ; X_t = z\}$.

Theorem 3.3.

Assume (A) holds. If $0 < (2a)^{1/2} < |x-z| < (2b)^{1/2} < \infty$ and $t \in \mathbb{R}_+$, then for $y = (t,x)$ we have

$$(3.3)\quad \frac{\overline{s}(b) - \overline{s}(\frac{|x-z|^2}{2})}{\overline{s}(b) - \overline{s}(a)} \leq Q_y\{T_a < T_b\} \leq \frac{\underline{s}(b) - \underline{s}(\frac{|x-z|^2}{2})}{\underline{s}(b) - \underline{s}(a)}$$

and

$$\overline{k}(b) - \overline{k}(\frac{|x-z|^2}{2}) - (\overline{k}(b) - \overline{k}(a)) \frac{\underline{s}(b) - \underline{s}(\frac{|x-z|^2}{2})}{\underline{s}(b) - \underline{s}(a)} \leq$$

$$(3.4)\quad Q_y\, I_{T_a \wedge T_b} \leq$$

$$\underline{k}(b) - \underline{k}(\frac{|x-z|^2}{2}) - (\underline{k}(b) - \underline{k}(a)) \frac{\overline{s}(b) - \overline{s}(\frac{|x-z|^2}{2})}{\overline{s}(b) - \overline{s}(a)}$$

Proof

To prove the first statement note that the functions $\overline{u}(x) = \overline{s}(\frac{|x-z|^2}{2})$ and $\underline{u}(x) = \underline{s}(\frac{|x-z|^2}{2})$ satisfy $\Lambda \overline{u} \leq 0$ and $\Lambda \underline{u} \geq 0$, respectively, on $\mathbb{R}_+ \times (\mathbb{R}^d - \{z\})$. This implies that

$$\overline{s}(a)\, Q_y\{T_a < T_b\} + \overline{s}(b)\, Q_y\{T_b < T_a\} \leq \overline{s}(\frac{|x-z|^2}{2}) \leq$$

$$\underline{s}(\frac{|x-z|^2}{2}) \leq \underline{s}(a)\, Q_y\{T_a < T_b\} + \underline{s}(b)\, Q_y\{T_b < T_a\}$$

To get the conclusion note that the functions $\overline{v}(x) = \overline{k}(\frac{|x-z|^2}{2})$ and $\underline{v}(x) = \underline{k}(\frac{|x-z|^2}{2})$ satisfy $\Lambda \overline{v} \leq f$ and $\Lambda \underline{v} \geq f$, respectively, on $\mathbb{R}_+ \times (\mathbb{R}^d - \{z\})$.

In the following we shall study the integral $I_{T_a \wedge T_b}$ for $0 \leq a < b \leq \infty$ (0 and ∞ included).

Theorem 3.4.

Assume (A) holds. Denote $y = (t,x)$, where $t \in \mathbb{R}_+$ and $x \in B(z,b) - \overline{B(z,a)}$.

1) If $\overline{k}(a) = \infty$, $\overline{k}(b) = \infty$, then $Q_y\{I_{T_a \wedge T_b} = \infty\} = 1$.
2) If $\overline{k}(a) < \infty$, $\overline{k}(b) = \infty$, then $\{I_{T_a \wedge T_b} < \infty\} = \{T_a < T_b\}$ Q_y a.s.

If additionally :

a) $\underline{s}(a) > -\infty$, $\underline{s}(b) < \infty$, then

$$\frac{\overline{s}(b)-\overline{s}(\frac{|x-z|^2}{2})}{\overline{s}(b) - \overline{s}(a)} \leq Q_y\{I_{T_a \wedge T_b} < \infty\} \leq \frac{\underline{s}(b)-\underline{s}(\frac{|x-z|^2}{2})}{\underline{s}(b) - \underline{s}(a)}$$

b) $\overline{s}(b) = \infty$, then $Q_y\{I_{T_a \wedge T_b} < \infty\} = 1$
c) $\overline{s}(b) = \infty$, $\overline{\delta}(b) = \infty$, then $Q_y\ I_{T_a \wedge T_b} = \infty$.

3) If $\overline{k}(a) = \infty$, $\overline{k}(b) < \infty$, then $\{I_{T_a \wedge T_b} < \infty\} = \{T_b < T_a\}$ Q_y a.s.

If additionally :

a) $\underline{s}(a) > -\infty$, $\underline{s}(b) < \infty$, then

$$\frac{\underline{s}(\frac{|x-z|^2}{2})-\underline{s}(a)}{\underline{s}(b) - \underline{s}(a)} \leq Q_y\{I_{T_a \wedge T_b} < \infty\} \leq \frac{\overline{s}(\frac{|x-z|^2}{2})-\overline{s}(a)}{\overline{s}(b) - \overline{s}(a)}$$

b) $\overline{s}(a) = -\infty$, then $Q_y\{I_{T_a \wedge T_b} < \infty\} = 1$

c) $\overline{s}(a) = -\infty$, $\overline{\delta}(a) = -\infty$, then $Q_y\ I_{T_a \wedge T_b} = \infty$.

Proof

The first assertion follows from Theorem 2.1. To prove the second one note that

$$Q_y\{I_{T_a \wedge T_b} < \infty ; T_a < T_b\} = Q_y\{T_a < T_b\}$$

Moreover, again by Theorem 2.1 , we have

$$Q_y\{T_a > T_b\} = Q_y\{\overline{k}(\frac{|X_{T_a \wedge T_b} - z|^2}{2}) = \infty, T_a > T_b\}$$

$$= Q_y\{I_{T_a \wedge T_b} = \infty, T_a > T_b\}.$$

Therefore Q_y a.s.

$$\{I_{T_a \wedge T_b} < \infty\} = \{T_a < T_b\}.$$

Statements a) and b) follow from (3.3). To prove c) note that by (3.4) we have

$$Q_y \, I_{T_a \wedge T_b} \geq \overline{k}(a) - \overline{k}(\frac{|x-z|^2}{2}) +$$

$$(\underline{s}(\frac{|x-z|^2}{2}) - \underline{s}(a)) \lim_{t \to b} \overline{\delta}(t) = \infty.$$

In order to get the final conclusion it is sufficient to interchange the roles of a and b.

Using the function $\underline{k}$ we prove

Theorem 3.5.

Suppose (A) holds and $y = (t,x) \in \mathbb{R}_+ \times (B(z,b) - \overline{B(z,a)})$. Assume that one of the following conditions is satisfied :

a) $\underline{k}(a) < \infty$, $\underline{k}(b) < \infty$
b) $\underline{k}(a) < \infty$, $\underline{k}(b) = \infty$, $\overline{s}(b) = \infty$
c) $\underline{k}(a) = \infty$, $\overline{s}(a) = -\infty$, $\underline{k}(b) < \infty$.

Then $Q_y\{I_{T_a \wedge T_b} < \infty\} = 1$. If additionally $\underline{\delta}(b) < \infty$ in b) and $\underline{\delta}(a) > -\infty$ in c) then $Q_y\{I_{T_a \wedge T_b}\} < \infty$.

Proof

If a) holds, then $Q_y I_{T_a \wedge T_b} < \infty$ by (3.4). Assume b) holds. From (3.3) it follows that $Q_y\{T_a < T_b\} = 1$. Note that, since $Q_y\{T_b = S = \infty\} = 1$ then $Q_y\{T_a < \infty\} = 1$. Consequently $Q_y\{I_{T_a \wedge T_b} < \infty\} = Q_y\{I_{T_a} < \infty\} = 1$. If moreover $\underline{\delta}(b) < \infty$, then by (3.4) we have

$$(3.5\ ) \qquad Q_y\, I_{T_a \wedge T_b} \le \underline{k}(a) - \underline{k}\left(\frac{|x-z|^2}{2}\right) + \underline{\delta}(b)\left(\overline{s}\left(\frac{|x-z|^2}{2}\right) - \overline{s}(a)\right) < \infty$$

Analogously one can show that if c) holds and $\underline{\delta}(a) > -\infty$, then

$$(3.6\ ) \qquad Q_y\, I_{T_a \wedge T_b} \le \underline{k}(b) - \underline{k}\left(\frac{|x-z|^2}{2}\right) - \underline{\delta}(a)\left(\overline{s}(b) - \overline{s}\left(\frac{|x-z|^2}{2}\right)\right) < \infty$$

In the following we denote by $\underline{k}_1$, $\underline{\delta}_1$, $\overline{\delta}_1$ the functions $\underline{k}$, $\underline{\delta}$, $\overline{\delta}$, respectively, with $f = \underline{\gamma} = \overline{\gamma} \equiv 1$.

From the above results, putting $f \equiv 1$, one can obtain a criteria for transience, recurrence, positive recurrence and null recurrence of degenerated diffusions satisfying (A) (cf. Bhattacharya [2,3] for the nondegenerated case). If $\underline{s}(\infty) < \infty$, then the diffusion is transient. If $\overline{s}(\infty) = \infty$, then it is recurrent. If $\overline{s}(\infty) = \infty$, $\underline{\delta}_1(\infty) < \infty$ and $\overline{s}(\infty) = \infty$, $\overline{\delta}_1(\infty) = \infty$, then it is, respectively, positive recurrent and null recurrent.

Now using Theorem 2.3 and inequalities (3.4) - (3.6) we prove

Theorem 3.6.

Assume (A) holds. If one of the following conditions holds :

a) $\underline{k}(a) < \infty$, $\underline{k}(b) < \infty$ and $\int_{t_0}^{b} \int_{t_0}^{t} \underline{\delta}(du)\overline{s}(dt) < 1$, where t_0 is given by $\underline{\delta}(t_0) \int_a^b \bar{s}(du) = \int_a^b \underline{\delta}(u)\ \bar{s}(du)$,

b) $\underline{k}(a) < \infty$, $\underline{k}(b) = \infty$, $\bar{s}(b) = \infty$, $\underline{\delta}(b) < \infty$ and $\int_a^b \int_t^b \underline{\delta}(du)\bar{s}(dt) < 1$,

c) $\underline{k}(a) = \infty$, $\bar{s}(a) = -\infty$, $\underline{\delta}(a) > -\infty$, $\underline{k}(b) < \infty$ and $\int_a^b \int_a^t \underline{\delta}(du)\bar{s}(dt) < 1$,

then there exists a constant c such that for
$y = (t,x) \in \mathbb{R}_+ \times (B(z,b) - \overline{B(z,a)}$ we have

$$Q_y \exp(I_{T_a \wedge T_b}) \leq c < \infty.$$

Proof

Assume a) holds and define the function $\phi(t) = k(b) - k(t) - (\underline{k}(b) - \underline{k}(a)) \dfrac{\bar{s}(b)-\bar{s}(t)}{\bar{s}(b)-\bar{s}(a)}$. Since $\phi(t) \leq \int_{t_0}^b \int_{t_0}^t \underline{\delta}(du)\ \bar{s}(dt)$ for all $t \in [a,b]$ the assertion follows from (3.4) and Theorem 2.3 (ii).

Now assume b) holds and note that

$$\underline{k}(a) - \underline{k}(t) + \underline{\delta}(b)(\bar{s}(t) - \bar{s}(a)) = \int_a^t \int_u^b \underline{\delta}(dv)\bar{s}(du).$$

Consequently, if $\int_a^b \int_t^b \underline{\delta}(du)\ \bar{s}(dt) < 1$, then using (3.5) and Theorem 2.3 (ii) we get the conclusion.

The above result yields

Corollary 3.1.

Assume (A) holds. If one of the following conditions hold :

a) $\underline{k}(a) < \infty$, $\underline{k}(b) < \infty$

b) $\underline{k}(a) < \infty$, $\underline{k}(b) = \infty$, $\bar{s}(b) = \infty$, $\underline{\delta}(b) < \infty$
and $\int_a^b \int_t^b \underline{\delta}(du)\ \bar{s}(dt) < \infty$,

c) $\underline{k}(a) = \infty$, $\bar{s}(a) = -\infty$, $\underline{\delta}(a) > -\infty$, $\underline{k}(b) < \infty$
and $\int_a^b \int_a^t \underline{\delta}(du)\ \bar{s}(dt) < \infty$,

then there exist $\lambda > 0$ and $c < \infty$ such that for $y \in \mathbb{R}_+ \times (B(z,b) - \overline{B(z,a)})$ we have

$$Q_y \exp(\lambda I_{T_a \wedge T_b}) \leq c < \infty.$$

Corollary 3.2.

Suppose (A) holds. Assume that one of the following cases occurs :

a) $\underline{k}_1(a) < \infty$, $\underline{k}_1(b) < \infty$, $\lambda < \left(\int_{t_0}^b \int_{t_0}^t \underline{\delta}_1(du)\ \bar{s}(dt)\right)^{-1}$,
where t_0 is given by $\underline{\delta}_1(t_0) \int_a^b \bar{s}(du) = \int_a^b \underline{\delta}_1(u)\ \bar{s}(du)$

b) $\underline{k}_1(a) < \infty$, $\underline{k}_1(b) = \infty$, $\bar{s}(b) = \infty$, $\underline{\delta}_1(b) < \infty$ and
$\lambda < \left(\int_a^b \int_t^b \underline{\delta}_1(du)\ \bar{s}(dt)\right)^{-1}$, $(\infty^{-1} = 0)$

c) $\underline{k}_1(a) = \infty$, $\bar{s}(a) = -\infty$, $\underline{\delta}_1(a) > -\infty$, $\underline{k}_1(b) < \infty$ and
$\lambda < \left(\int_a^b \int_a^t \underline{\delta}_1(du)\ \bar{s}(dt)\right)^{-1}$.

Then there exists $c < \infty$ such that for $y \in \mathbb{R}_+ \times (B(z,b) - \overline{B(z,a)})$

$$Q_y \exp(\lambda\ T_a \wedge T_b) \leq c < \infty.$$

Example

We suppose now that $b(t,x) \equiv 0$ and $a(t,x) \equiv I$. In this case under $P_x = Q_{0,x}$, $x \in \mathbb{R}^d$, $(X_t)_{t \geq 0}$ is a d-dimensional standard Brownian motion. We assume that $d \geq 2$. The case of $d = 1$ will be investigated separetly in the next section. Consider the integral

$$I_t = \int_0^t \gamma\left(\frac{|X_s|^2}{2}\right) ds,\ 0 < t < \infty$$

where the function $\gamma :]0,\infty[\to \mathbb{R}_+$ is continuous. Note that with $z = 0$ and $r > 0$ we have

$$\overline{e}(t) = \underline{e}(t) = (\frac{r}{t})^{d/2}$$

$$\delta(t) = \overline{\delta}(t) = \overline{m}(t) = \underline{\delta}(t) = \underline{m}(t) = r^{-\frac{d}{2}} \int_r^t \eta^{\frac{d}{2}-1} \gamma(\eta)\, d\eta$$

$$k(t) = \overline{k}(t) = \underline{k}(t) = \int_r^t \frac{1}{\eta^{d/2}} \int_r^\eta \xi^{\frac{d}{2}-1} \gamma(\xi)\, d\xi\, d\eta$$

Denote, as before, $T_a = \inf\{t \geq 0 \; ; \; |X_t| = (2a)^{1/2}\}$ and $B(a) = \{x \in \mathbb{R}^d \; ; \; |x| < (2a)^{1/2}\}$. The above results yield the following statements.

Let $d = 2$, $0 < a < \infty$ and $x \in \mathbb{R}^2 - \overline{B(a)}$. Then

a) $P_x\{I_{T_a} < \infty\} = 1$,

b) if $\int_a^\infty \gamma(t)dt = \infty$, then $P_x\, I_{T_a} = \infty$,

c) if $\int_a^\infty \gamma(t)dt < \infty$, then $P_x\, I_{T_a} < \infty$,

d) if $\int_a^\infty \frac{1}{t} \int_t^\infty \gamma(s)ds\, dt < 1$, then $P_x \exp(I_{T_a}) \leq c < \infty$

Let $d \geq 3$, $0 < a < \infty$ and $x \in \mathbb{R}^d - \overline{B(a)}$. Then

a) if $k(\infty) = \infty$, then $P_x\{I_{T_a} < \infty\} = (\frac{\sqrt{2a}}{|x|})^{d-2}$

b) if $k(\infty) < \infty$, then there exists $\lambda > 0$ such that

$$P_x \exp(\lambda\, I_{T_a}) \leq c < \infty$$

c) if $k(\infty) < \infty$ and $\int_{t_0}^\infty t^{-d/2} \int_{t_0}^t s^{\frac{d}{2}-1} \gamma(s)\, ds\, dt < 1$

where t_0 is given by $\delta(t_0) = (\frac{d}{2} - 1)\, a^{\frac{d}{2}-1} \int_a^\infty \delta(t)\, t^{-\frac{d}{2}}\, dt$,

then $P_x \exp(I_{T_a}) \leq c < \infty$.

Let $d \geq 3$, $0 < a < b < \infty$ and $x \in B(b) - \overline{B(a)}$. If $\lambda \leq \dfrac{\frac{d}{2} - 1}{b - a}$; then $P_x \exp(\lambda\, T_a \wedge T_b) \leq c < \infty$

4. PARTICULAR CASES

In this section we consider particular cases of diffusions and functionals in which we give more complete description of divergence or convergence.

Quadratic forms of linear diffusions

We suppose now that $D = \mathbb{R}^d$ and that $a(t,x) = A(t)$, $b(t,x) = B(t)x$, $t \in \mathbb{R}_+$, $x \in \mathbb{R}^d$, where $A, B : \mathbb{R}_+ \to \mathbb{R}^d \otimes \mathbb{R}^d$ are continuous. In this case $S = \infty\ Q_y$ a.s. for $y \in \mathbb{R}_+ \times \mathbb{R}^d$ and the process $(X_t)_{t \geq 0}$ is Gaussian under $P_x = Q_{0,x}$, $x \in \mathbb{R}^d$, with mean $(\phi(t)x)$ and covariance K given by :

$$\phi(t) = I + \int_0^t B\phi(r)\, dr,$$

$$K(s,t) = \phi(s) \int_0^{s \wedge t} (\phi^{-1} A \phi^{-1*})(r) dr\ \phi^*(t).$$

We assume that $K(t,t)$ is positive definite for all $t > 0$. We wish to study the finiteness or infiniteness of I_∞, given by (1.1.), in the case where

$$f(t,x) = x \cdot Q(t)x\ ,\ t \in \mathbb{R}_+,\ x \in \mathbb{R}^d.$$

Here $Q(t)$ is a symmetric positive (≥ 0) $d \times d$ matrix which is continuous in t.

Let $\Gamma_0 \in \mathbb{R}^d \otimes \mathbb{R}^d$ be symmetric and positive definite(> 0) and let $\Gamma : \mathbb{R}_+ \to \mathbb{R}^d \otimes \mathbb{R}^d$ be the solution of the matrix Riccati equation

$$\dot{\Gamma} = A + B\Gamma + \Gamma B^* - \Gamma Q \Gamma,\ \Gamma(0) = \Gamma_0.$$

Let $D = \psi^{-1}(\Gamma + R)\psi^{-1*}$, where the functions $\psi, R : \mathbb{R}_+ \to \mathbb{R}^d \otimes \mathbb{R}^d$ are defined like ϕ, K respectively but with $B + A\Gamma^{-1}$ instead of B ($\Gamma(t)$ is positive definite for all $t \in \mathbb{R}_+$).

Moreover, let $g_K(x)$ stands for the density of Gaussian measure with mean 0 and variance K and let

$$u(t,x) = (2\pi\det\Gamma(t))^{-1/2}\exp(-\tfrac{1}{2}\int_0^t \mathrm{tr}A\Gamma^{-1}(s)ds)g^{-1}_{\Gamma(t)}(x).$$

Since $\Lambda u = \frac{1}{2}fu$ on $\mathbb{R}_+ \times \mathbb{R}^d$ one has

$$P_x\exp(-\tfrac{1}{2}I_t) = P_x\frac{N^u_t}{u(t,X_t)} = P^u_x\frac{u(0,x)}{u(t,X_t)},$$

where N^u is defined in (2.1) and P^u_x is a diffusion measure with the operator $\Lambda + A\Gamma^{-1}\cdot\mathbb{D}_x$. But, under P^u_x, (X_t) is Gaussian with mean $(\psi(t)x)$ and variance $R(t)$ (which is >0 since the distributions of X_t under P_x and P^u_x are equivalent and $K(t)>0$). Therefore

$$P^u_x\, g_{\Gamma(t)}(X_t) = g_{\Gamma(t)} * g_{R(t)}(\psi(t)x) = g_{(\Gamma + R)(t)}(\psi(t)x) =$$

$$(\det\psi(t))^{-1}g_{D(t)}(x) = \exp(-\int_0^t \mathrm{tr}(B + A\Gamma^{-1})(s)ds)g_{D(t)}(x).$$

Consequently $P_x\exp(-\frac{1}{2}I_t) =$

$$g^{-1}_{\Gamma(0)}(x)(\det\Gamma(0)^{-1}\det\Gamma(t))^{1/2}\exp(-\int_0^t \mathrm{tr}(B + A\Gamma^{-1})(s)ds)g_{D(t)}(x) =$$

$$g^{-1}_{D(0)}(x)\exp(\int_0^t \mathrm{tr}(\tfrac{1}{2}\dot{\Gamma}\Gamma^{-1} - B - \tfrac{1}{2}A\Gamma^{-1})(s)ds)g_{D(t)}(x) =$$

$$g^{-1}_{D(0)}(x)g_{D(t)}(x)\exp(-\tfrac{1}{2}\int_0^t \mathrm{tr}Q\Gamma(t)dt). \tag{4.1}$$

Moreover, $\dot{D} = -\psi^{-1}\Gamma Q\Gamma\psi^{-1*}$ what implies that $\lim_{t\to\infty} D(t)$ exists $(=D(\infty))$. Finally

$$D(\infty) \geq \int_0^\infty \psi^{-1}A\psi^{-1*}(s)ds > 0.$$

From the above considerations it follows

<u>Theorem 4.1.</u>

If $K(t) \equiv K(t,t)$ is positive definite for all $t>0$, then for all $x\in\mathbb{R}^d$ one has

$$P_x\{I_\infty = \infty\} = 1 \quad \text{iff} \quad \int_0^\infty \mathrm{tr}\,Q\Gamma(t)dt = \infty,$$

$$P_x\{I_\infty < \infty\} = 1 \quad \text{iff} \quad \int_0^\infty \mathrm{tr}\,Q\Gamma(t)dt < \infty.$$

Proof

Note that $\inf_{t\geq 0} \det D(t) > 0$. Therefore, using (4.1), we have that $P_x \exp(-\frac{1}{2} I_\infty) = 0$ and consequently $P_x\{I_\infty = \infty\} = 1$ if $\int_0^\infty \mathrm{tr} Q\Gamma(t)dt = \infty$. Consider the linear space $W_0 = \{0,\omega) : \omega(t) \neq \delta \ \forall\ t \in \mathbb{R}_+$ and $I_\infty(\omega) < \infty\}$. Since (X_t) is Gaussian under P_x, one has $P_x\{W_0\} = P_x\{I_\infty < \infty\} = 0$ or 1. Moreover, if $\int_0^\infty \mathrm{tr} Q\Gamma(t)dt < \infty$, then by (4.1) we obtain that $P_x\{I_\infty < \infty\} \geq$

$$P_x \exp(-\frac{1}{2} I_\infty) = g^{-1}_{D(0)}(x) g_{D(\infty)}(x) \exp(-\frac{1}{2}\int_0^\infty \mathrm{tr} Q\Gamma(t)dt) > 0$$

what proves the assertion.

Remark 4.1.

For Bessel's quadratic functionals J.Pitman and M.Yor [18] obtained a different necessary and sufficient conditions.

Suppose now that the functions A and B are constant (i.e. $A(t)=A$, $B(t)=B$ for all $t \in \mathbb{R}_+$).Then $K(t)>0$ for each $t>0$ if and only if the $d \times d^2$ block matrix $[A,BA,\ldots,B^{d-1}A]$ has rank d or equivalently the pair $[B,A^{1/2}]$ is controllable.From Theorem 4.1 it follows the following result of Le Breton and Musiela [11].

Corollary 4.1.

Assume that the pair $[B,A^{1/2}]$ is controllable.Then for all $x \in \mathbb{R}^d$

$$P_x \{ \int_0^\infty X_t \cdot QX_t dt = \infty \} = 1$$

for each symmetric positive $d \times d$ matrix Q ($Q \neq 0$).

Proof

To get the assertion it is sufficient to choose the function Γ constant i.e. $\Gamma(t) = \Gamma_0$ for all $t \in \mathbb{R}_+$,where Γ_0 is a positive solution of the matrix equation

$$A + B\Gamma + \Gamma B^* - \Gamma Q\Gamma = 0 .$$

One dimensional homogeneous diffusions

We suppose now that $D =]\ell, r[$ $(-\infty \le \ell < r \le \infty)$ is an open interval in $\mathbb{R}$ and that the real functions $a(t,\cdot)$ $(a > 0)$ and $b(t,\cdot)$ do not depend on t. In this case (W, G, G_t, X_t, S, P_x) with $P_x = Q_{0,x}$, $x \in D$, is the canonical realization of a homogeneous diffusion with infinitesimal generator $L = \frac{1}{2}\, a\, \mathbb{D}_x^2 + b\, \mathbb{D}_x$. We wish to study the integral

$$I_S = \int_0^S f(X_t)\, dt$$

in the case where the function $f : D \to \mathbb{R}_+$ is continuous.

Choose $c \in D$ such that $f(c) > 0$ and define the functions

$$e(x) = \exp\left(-\int_c^x \frac{2b}{a}(y)dy\right),\quad m(x) = 2\int_c^x \frac{f}{ea}(y)\, dy,$$

$$s(x) = \int_c^x e(y)dy\ ,\quad k(x) = \int_c^x em(y)\, dy.$$

For $f \equiv 1$ parts "if" of (i) and (ii-a) of the following result (cf. Musiela [14]) reduce to the classical Feller's tests of non explosion or explosion.

Theorem 4.2.

Let $p(x) = P_x\{I_S < \infty\}$. The following assertions hold.

(i) $p \equiv 0$ iff $k(\ell) = \infty$ and $k(r) = \infty$.

(ii) $p \equiv 1$ iff one of the following cases occurs

(a) $k(\ell) < \infty$, $k(r) < \infty$,

(b) $k(\ell) < \infty$, $k(r) = \infty$, $s(r) = \infty$,

(c) $k(\ell) = \infty$, $s(\ell) = -\infty$, $k(r) < \infty$.

(iii) If $k(\ell) < \infty$, $k(r) = \infty$, $s(r) < \infty$, then $p(x) = \dfrac{s(r)-s(x)}{s(r)-s(\ell)}$.

(iv) If $k(\ell) = \infty$, $s(\ell) > -\infty$, $k(r) < \infty$, then $p(x) = \dfrac{s(x)-s(\ell)}{s(r)-s(\ell)}$.

Analogously to Theorem 3.5 one can prove

Theorem 4.3.

(i) $P_x I_S < \infty$, $x \in D$ iff one of the following cases occurs

(a) $k(\ell) < \infty$, $k(r) < \infty$,
(b) $k(\ell) < \infty$, $k(r) = \infty$, $s(r) = \infty$, $m(r) < \infty$,
(c) $k(\ell) = \infty$, $s(\ell) = -\infty$, $m(\ell) > -\infty$, $k(r) < \infty$.

(ii) $P_x I_S$, $x \in D$, is given by

$$k(\ell)\,\frac{s(r)-s(x)}{s(r)-s(\ell)} + k(r)\,\frac{s(x)-s(\ell)}{s(r)-s(\ell)} - k(x),$$

or by

$$k(\ell) + m(r)(s(x) - s(\ell)) - k(x)$$

or by

$$k(r) - m(\ell)(s(r) - s(x)) - k(x)$$

in cases (a), (b) and (c) of (i), respectively.

The above result together with Theorem 2.3 permit to obtain sufficient conditions for the existence of exponential moments of I_S.

Theorem 4.4.

Suppose that one of the following conditions holds

(a) $k(\ell) < \infty$, $k(r) < \infty$ and $\int_{x_0}^{r}\int_{x_0}^{y} m(dt)\ s(dy) < 1$, where x_0 is given by $m(x_0)\int_{\ell}^{r} s(dt) = \int_{\ell}^{r} m(t)\ s(dt)$,

(b) $k(\ell) < \infty$, $k(r) = \infty$, $s(r) = \infty$, $m(r) < \infty$ and $\int_{\ell}^{r}\int_{y}^{r} m(dt)s(dy) < 1$,

(c) $k(\ell) = \infty$, $s(\ell) = -\infty$, $m(\ell) > -\infty$, $k(r) < \infty$ and $\int_{\ell}^{r}\int_{\ell}^{y} m(dt)s(dy) < 1$.

Then there exists a constant C such that for $x \in D$ we have

$$P_x \exp(I_S) \le C < \infty.$$

Example

Assume that $b \equiv 0$ and $a \equiv 1$. Then $e \equiv 1$, $s(x) = x-c$, $m(x) = 2\int_c^x f(y)dy$ and $k(x) = 2\int_c^x \int_c^y f(z)dz\, dy$.

If $D = \mathbb{R}$ then $P_x\{I_S = \infty\} = 1$.
If $D =]\ell, \infty[$ with $\ell > -\infty$, then $P_x\{I_S < \infty\} = 1$.
Moreover, $P_x I_S < \infty$ if and only if $\int_\ell^\infty f(y)dy < \infty$.
Finally, if $\int_\ell^\infty \int_y^\infty f(t)dt\, dy < \frac{1}{2}$, then $P_x \exp(I_S) \le C < \infty$.
For $D =]\ell, r[$ with $-\infty < \ell < r < \infty$ we have :

if $\int_{x_0}^r \int_{x_0}^y f(t)dt\, dy < \frac{1}{2}$, where x_0 is given by

$$k'(x_0)(r-\ell) = k(r) - k(\ell), \text{ then } P_x \exp(I_S) \le C < \infty.$$

Remark 4.2.

Note that under assumptions of Theorem 4.4 the function $u(x) = P_x \exp(I_S)$ belongs to $C^2(D)$ and satisfies the equation $(L+f)u = 0$ on D. Moreover, if (a) holds, then $u(\ell) = u(r) = 1$, and the equation has a solution non vanishing on $D \cup \{\ell, r\}$. This improves classical results (cf. Coppel [4] p. 20 and 64 for example). If (b) (resp. (c)) holds, then $u(\ell) = 1$ (resp. $u(r) = 1$) and the equation has a solution non vanishing on $D \cup \{\ell\}$ (resp. $D \cup \{r\}$) (cf. Coppel [4] p. 28).

Acknowledgement

The author would like to thank Professors A. LE BRETON and B. MAISONNEUVE for comments and suggestions throughout the course of this research.

REFERENCES

[1] R. AZENCOTT
Behavior of diffusion semi-groups at infinity,
Bull. soc. Math. France, 102, 193-240 (1974).

[2] R.N. BHATTACHARYA
Criteria for recurrence and existence of invariant measures for multidimensional diffusions.
Ann. Probab., 6, 541-553 (1980). Correction Note, Ibid. 8, 1194-95 (1980).

[3] R.N. BHATTACHARYA, S. RAMASUBRAMANIAN
Recurrence and ergodicity of diffusions.
J. of Multivariate Analysis 12, 95-122 (1982).

[4] W. A. COPPEL
Disconjugacy, Lecture Notes in Math. 220, Springer-Verlag, Berlin, Heidelberg, New-York, 1971.

[5] E.B. DYNKIN
Markov Processes,
Springer-Verlag, Berlin, Göttingen, Heidelberg, 1965.

[6] N. IKEDA and S. WATANABE
Stochastic Differential Equations and Diffusion Processes,
North-Holland, Amsterdam Oxford, New-York ; Kodansha LTD, Tokyo, 1981.

[7] Ju. M. KABANOV, R.S. LIPTSER, A.N. SHIRYAEV
Absolute continuity and singularity of locally absolutely continuous probability distributions. I ; II, Mathem. Sbornik 107 (149), n° 3 (11) 364-415 (1978) ; 108 (150), n° 1, 32-61 (1979

[8] R.Z. KHAS'MINSKII
On positive solutions of the equation $\Lambda u + Vu = 0$
Theor. Probability Appl., vol. 4, n° 3, 309-318 (1959).

[9] R.Z. KHAS'MINSKII
Ergodic properties of recurrent diffusion processes and stabilizat of the solution of the Cauchy problem for parabolic equations,
Theor. Probability Appl., Vol. 5, n° 2, 179-196 (1960).

[10] R.Z. KHAS'MINSKII
Stochastic stability of differential equations,
Sijthoff and Noordhoff, Rockville, 1980.

[11] A. LE BRETON and M. MUSIELA
Some parameter estimation problems for hypoelliptic homogeneous Gaussian diffusions, Proc. Banach Math. Center, Warsaw, (to appear).

[12] B. MAISONNEUVE
Une mise au point sur les martingales locales continues definies sur un intervalle stochastique. Séminaire de Probabilités XI, Lecture Notes in Mathematics 581, Springer-Verlag, Berlin (1977).

[13] H.P. Mc KEAN Jr.
Stochastic integrals,
Academic Press, New-York, London, 1969.

[14] M. MUSIELA
On Kac functionals of one dimensional diffusions,
Stochastic Process. Appl., (to appear).

[15] K. NARITA
On explosion and growth order of inhomogeneous diffusion processes,
Yokohama Math. J. Vol. 28, 45-57 (1980).

[16] K. NARITA
Remarks on non explosion theorem for stochastic differential equations, Kodaï Math. J. 5, 395-401 (1982).

[17] K. NARITA
No explosion criteria for stochastic differential equations,
J. Math. Soc. Japan, vol. 34, n° 2, 191-203 (1982).

[18] J. PITMAN and M. YOR
A decomposition of Bessel bridges,
Z. Wahrscheinlichkeitstheorie verw. Gebiete, 59, 425-457 (1982).

[19] D.W. STROOCK and S.R.S. VARADHAN
Multidimensional Diffusion Processes, Springer-Verlag, Berlin, Heidelberg, New-York, 1979.

ESTIMATION OF REGRESSION PARAMETERS OF GAUSSIAN MARKOV PROCESSES

by

Marek Musiela and Roman Zmyślony
Polish Academy of Sciences, Wrocław

1. Introduction

Let $\{\Omega, \mathcal{F}, \mathcal{P}\}$ be a complete probability space and let for $t \in [0,\infty)$

$$y(t) = X(t)\beta + z(t), \quad y(0) = 0 \tag{1}$$

be an n-dimensional continuous stochastic process. $X(t)$ stands for a matrix of known functions, β for a p-dimensional vector of unknown parameters and $z(t)$ for an n-dimensional Gaussian Markov process with expectation zero and covariance matrix $R(t,s)$. Let $R_0(t)=R(t,t)$. The following limits

$$X_1(t) = \lim_{h \to 0} \frac{1}{h} (X(t+h) - X(t)), \tag{2}$$

$$R_1(t) = \lim_{h \to 0} \frac{1}{h} (R_0(t+h) - R_0(t)) \tag{3}$$

and

$$R_2(t) = \lim_{h \to 0+} \frac{1}{h} (R(t+h,t) - R_0(t)) \tag{4}$$

are assumed to exist for $t \in [0,\infty)$ and to be continuous. In this paper we study the problem of estimation of β when one sample path $y(t)$ on $[0, T]$ is available. Two cases are considered, one deals with a known covariance matrix $R(t,s)$ and the other one with an unknown covariance matrix $R(t,s)$.

This model is somewhat more general than the model considered by

Hajek [7, 8]. He assumed that y is a one dimensional process and that β is an unknown number. In order to find a sufficient statistic Hajek used methods of Hilbert spaces. In 1969 Holevo considered the estimation of regression parameters of process (1) in the one dimensional case. Moreover, he assumed that the process z(t) is stationary. At the begining of the seventies modern theory of martingales and stochastic integrals has been used in many statistical problems arising in stochastic processes, (see [1, 2, 3, 5, 10, 11, 13]). Using the new techniques we obtain best unbiased estimators of regression coefficients.

2. Two auxiliary lemmas

First we prove two lemmas concerning the covariance operator which will be used repeatedly. Let A^+ stand for the Moore-Penrose general inverse of matrix A.

Lemma 1. Whatever be u, $s \leqslant u \leqslant t$, we have

$$R(t,s) = R(t,u)\, R_0^+(u)\, R(u,s). \tag{5}$$

P r o o f: Since z(t) belongs to the image of $R_0(t)$ and since $R_0(t)\, R_0^+(t)$ is the projection on $R_0(t)$, it is clear that $z(t) = R_0(t)\, R_0^+(t)\, z(t)$. From the above and a theorem on normal correlation [13] it follows that for $u \leqslant t$

$$E(z(t)|z(u)) = R(t,u)\, R_0^+(u)\, z(u). \tag{6}$$

Because z(t) is a Markov process it is clear that $R(t,s) = E(E(z(t)|z(u))z'(s))$ for $s \leqslant u$. Thus the lemma follows from (6).

Now let

$$B(t) = R_1(t) - R_2(t) - R_2'(t), \tag{7}$$

where $R_1(t)$ and $R_2(t)$ are defined by (3) and (4), respectively. Symbol A' stands for the transpose of matrix A.

Lemma 2. Matrix $B(t)$ is non-negative definite.

P r o o f: Simple calculations show that

$$B(t) = \lim_{h \to 0^+} \frac{1}{h} E(z(t+h) - z(t))(z(t+h) - z(t))'.$$

Because $\frac{1}{h} E(z(t+h) - z(t))(z(t+h) - z(t))'$ is non-negative definite for every $h > 0$ the lemma follows.

3. Differential representation of the process

In this section we show that the process defined in (1) is a solution of a stochastic differential equation. Let $\bar{y}(t)$ be a solution of the following equation

$$d\,\bar{y}(t) = (A(t)\beta + C(t)\,\bar{y}(t))\,dt + B^{\frac{1}{2}}(t)\,d\,\bar{W}(t), \tag{8}$$
$$\bar{y}(0) = 0,$$

where

$$A(t) = X_1(t) - R_2(t)\,R_0^+(t)\,X(t), \tag{9}$$

$$C(t) = R_2(t)'\,R_0^+(t), \tag{10}$$

while $\bar{w}(t)$ is an n-dimensional Wiener process. Moreover, let $\bar{m}(t)$, $\bar{R}(t,s)$ and $\bar{R}_0(t)$ be the expectation, the covariance and the variance matrices of $y(t)$, respectively.

Lemma 3. The distributions of the processes $y(t)$ and $\bar{y}(t)$ coincide.

P r o o f: Because $\bar{y}(t)$ is a Gaussian process it is sufficient to prove that the first two moments of $\bar{y}(t)$ and $y(t)$ coincide. One can prove that $\bar{m}(t)$ and $\bar{R}(t,s)$ are solution of the following differential equations (compare [13], Theorem 15.2).

$$\frac{d\,\bar{m}(t)}{dt} = A(t)\beta + C(t)\,\bar{m}(t), \quad \bar{m}(0) = 0, \tag{11}$$

$$\frac{d\,\bar{R}_o(t)}{dt} = C(t)\,\bar{R}_o(t) + \bar{R}_o(t)\,C'(t) + B(t), \quad \bar{R}_o(0) = 0, \tag{12}$$

and

$$\bar{R}(t,s) = \begin{cases} D(t)\,D^{-1}(s)\,\bar{R}_o(s) & \text{if} \quad t \geqslant s, \\ \bar{R}_o(t)\,D^{-1}(t)\,D(s) & \text{if} \quad t < s, \end{cases} \tag{13}$$

where

$$\frac{d\,D(t)}{dt} = C(t)\,D(t), \qquad D(0) = I_n. \tag{14}$$

Here I_n stands for the identity matrix. Letting $f(t) = \bar{m}(t) - X(t)\beta$ we obtain from (9) and (11) that

$$\frac{d\,f(t)}{dt} = C(t)\,f(t), \; f(0) = 0\,.$$

Thus we have $\bar{m}(t) = X(t)\beta$. Moreover, letting $G(t) = \bar{R}_o(t) - R_o(t)$, and using (4), (10) and (12) we obtain

$$\frac{d\,G(t)}{dt} = R_2(t)(R_o^+(t)\,\bar{R}_o(t) - I_n) + (\bar{R}_o(t)\,R_o^+(t) - I_n)R_2'(t),$$
$$G(0) = 0. \tag{15}$$

By Lemma 1 we have $R_2(t) = R_2(t)\,R_o^+(t)\,R_o(t)$. Since $R_o^+(t) = R_o^+(t)\,R_o(t)\,R_o^+(t)$ we can rewrite (15) in the following form

$$\frac{d\,G(t)}{dt} = C(t)\,G(t) + G(t)\,C'(t), \quad G(0) = 0.$$

This implies that $\bar{R}_o(t) = R_o(t)$.

Finally, we prove that $\bar{R}(t,s) = R(t,s)$. Without loss of generality we may assume that $s < t$. From Lemma 1 it follows that $\partial R(t,s)/\partial t = C(t)\,R(t,s)$. On the other hand in view of (13) and (14) we have

$$\frac{\partial \bar{R}(t,s)}{\partial(t)} = \frac{d\,D(t)}{dt}\,D^{-1}(s)\,\bar{R}_o(s) = C(t)\,\bar{R}(t,s).$$

This terminates the proof of the lemma.

Now we shall prove that the process defined in (1) is the solution of a stochastic differential equation.

Theorem 1. There exists a standard Wiener process w(t) such that for $t \in [0,T]$

$$d\,y(t) = (A(t)\beta + C(t)\,y(t))\,dt + B^{\frac{1}{2}}(t)\,d\,w(t), \quad y(0) = 0. \qquad (16)$$

P r o o f: From Lemma 3 it follows that the distributions of the processes

$$u(t) = y(t) - \int_0^t (A(u)\beta + C(u)\,y(u))\,du \qquad (17)$$

and

$$\bar{u}(t) = \bar{y}(t) - \int_0^t (A(u)\beta + C(u)\,y(u))\,du$$

coincide. Using (8) we have

$$\bar{u}(t) = \int_0^t B^{\frac{1}{2}}(u)\,d\,\bar{w}(u).$$

This implies that u(t) has independent increaments. For $t \in [0,T]$ let

$$w(t) = \int_0^t (B^{\frac{1}{2}}(s))^+\,d\,u(s) + \int_0^t (I_n - B^{\frac{1}{2}}(s))^+\,B^{\frac{1}{2}}(s)\,d\,v(s),$$

where v(s) is a standard Wiener process which is independent of y(t). From Lemma 10.4 in [13] it follows that w(t) is a Wiener process and

$$u(t) = \int_0^t B^{\frac{1}{2}}(u)\,d\,w(u).$$

This together with (17) completes the proof of the theorem.

4. Sufficient statistics and estimation

In this section we assume that the vector β runs over a subset $\Theta \subset R^p$ such that the interior of Θ is non-empty. Moreover, we assume that the covariance matrix is known. In this case we shall prove that there exists a complete sufficient statistic for the vector of parameters β.

Let $\mathcal{C}$ stand for the space of continuous functions $f:[0,T] \to R^n$ endowed with the topology of uniform convergence and let $\mathcal{B}$ denote the σ-field of Borel subsets of $\mathcal{C}$. Moreover, let

$$\mu^{\beta}(B) = P_{\beta}(y \in B) \quad \text{and} \quad \mu(B) = P(u \in B),$$

where $B \in \mathcal{B}$, while y and u are defined by (1) and (17), respectively.

Theorem 2. Suppose that

$$\Sigma(T) = \int_0^T A'(t)\, B^+(t)\, A(t)\, dt$$

is a positive definite matrix. If int $\Theta \neq \emptyset$, then

$$S(T) = \int_0^T A'(t)\, B^+(t)\, [d\, y(t) - C(t)\, y(t)\, dt] \tag{18}$$

is a complete and sufficient statistic for $\{\mu^{\beta} \mid \beta \in \Theta\}$.

P r o o f: Since $y(t)$ has the representation given by (16) it follows from Theorem 7.20 in [13] that $\mu^{\beta} \ll \mu$ for every $\beta \in \Theta$ and that

$$\frac{d\,\mu^{\beta}(y)}{d\mu} = \exp\Big\{ \int_0^T (A(t)\beta + C(t)\, y(t))'\, B^+(t)\, d\, y(t) +$$

$$- \frac{1}{2} \int_0^T (A(t)\beta + C(t)\, y(t))' B^+(t) (A(t)\beta + C(t)\, y(t))\, dt \Big\}. \tag{19}$$

After a simple calculation the factorization theorem [12] implies that $S(T)$ defined by (18) is sufficient. Moreover, formula (18) implies that $S(T)$ is normally distributed $N(\Sigma(T)\beta, \Sigma(T))$. So, $S(T)$ is complete because $\Sigma(T)$ is positive definite and int Θ is non-empty.

Note that Theorem 1 is valid provided $\Sigma(T)$ is positive definite. In the following lemma we give a simple characterization of this assumption.

Lemma 4. Let $M(t)$ be an $n \times n$ non-negative definite symmetric matrix for all $t \in [0,T]$. Let the mapping $t \to M(t)$ be continuous. The integral $\int_0^T M(t)\, dt$ is positive definite if and only if there exist $t_1, \ldots, t_k \in (0,T)$, where $k \leq n(n+1)/2 + 1$, such that $\sum_{i=1}^{k} M(t_i)$ is positive definite.

P r o o f: Suppose that $\int_0^T M(t)\, dt$ is positive definite. Let $\mathcal{M}^*$ stand for the convex hull of $\mathcal{M} = \{M(t) \mid t \in (0,T)\}$. Since $\frac{1}{T}\int_0^T M(t)\, dt \in \mathcal{M}^*$

it follows from the Caratheodory theorem that there exist numbers $t_1, \dots, t_k \in [0,T]$, where $k \leqslant n(n+1)/2 + 1$, and $p_i > 0$, $\sum_{i=1}^{k} p_i = 1$, such that

$$\int_0^T M(t)\, dt = T \sum_{i=1}^{k} p_i\, M(t_i).$$

Since $M(t_i)$ is non-negative definite and since $T \sum_{i=1}^{k} p_i\, M(t_i)$ is positive definite $\sum_{i=1}^{k} M(t_i)$ is positive definite, too.

Now, let $\sum_{i=1}^{k} M(t_i)$ be positive definite. Then the mapping $(s_1, \dots, s_k) \longrightarrow \sum_{i=1}^{k} M(s_i)$ is continuous. Hence, there exists $\varepsilon > 0$ such that $\sum_{i=1}^{k} M(s_i)$ is positive definite for every $s_i \in (t_i - \varepsilon,\ t_i + \varepsilon)$, $i = 1, \dots, k$. It is clear that

$$\int_0^T (M(t)x, x)\, dt \geqslant \int_0^{\varepsilon} \left(\sum_{i=1}^{k} M(t_i + t)x, x \right) dt > 0.$$

This terminates the proof.

Now we consider the problem of estimation of β and $Ey(t) = X(t)\beta$.

Theorem 3. Suppose that $\sum(T)$ is positive definite. If $\operatorname{int} \Theta$ is non-empty, then

$$\hat{\beta}_T = \sum\nolimits^{-1}(T) S(T) \tag{20}$$

is the best unbiased estimator of β.

P r o o f: Since $\sum(T)$ is positive definite, $\hat{\beta}_T$ is an unbiased estimator of β [14]. Moreover, in view of Theorem 2, $S(T)$ is a complete and sufficient statistic. Hence the Rao-Blackwell Theorem implies that $\hat{\beta}_T$ is the best unbiased estimator of β.

Remark 1. Note that $X(t) \sum^{-1}(T)\ S(T)$ is the best unbiased estimator of $Ey(t)$.

Remark 2. If $\lim_{T\to\infty} \sum^{-1}(T) = 0$, then $\hat{\beta}_T$ is a consistent estimator of β, i.e $\lim_{T\to\infty} \hat{\beta}_T = \beta$ a.s. for each $\beta \in \Theta$.

When $R(t,s)$ is unknown μ depends on $R(t,s)$, thus μ does not dominate μ^{β}. In order to find out whether or not μ dominates μ^{β} we need to have an explicit form of $B(\cdot)$. We shall prove that there exists $\mathcal{F}_t^y = \sigma(y(s), s \leq t)$ - measurable modification of B, say $\hat{B}$, such that B and $\hat{B}$ are indistinguishable (see [4] for definition)

Let $t_k^{(n)} = \frac{kT}{2^n}$, where $k = 0, 1, \ldots, 2^n$.

Define

$$Q_n(t) = \sum_{k \in I_n(t)} \left(y(t_{k+1}^n) - y(t_k^n)\right)\left(y(t_{k+1}^n) - y(t_k^n)\right)',$$

where $I_n(t) = \left\{k:\ t_k^n \leq t\right\}$. Clearly, $Q_n(t)$ converges to $Q(t)$ with probability one, where $Q(t) = \int_0^t B(u)du$ (see [6], [13]). Since $B(t)$ is continuous it follows that

$$\hat{B}(t) = \lim_{h \to 0} \frac{1}{h}\left(Q(t+h) - Q(t)\right)$$

is a $\mathcal{F}_t^y$ measurable modification of $B(t)$. So we have proved the following lemma.

Lemma 5. Processes $\hat{B}(\cdot)$ and $B(\cdot)$ are indistinguishable.

Remark 3. By Lemma 5 process $B(\cdot)$ can be determined from an observation of the process in the interval $[0,T]$. So, we can assume for futher considerations that $B(t)$ is known. If this is the case we have, in view of (16), that $\mu^{\beta} << \mu$ for every β and $R(t,s)$. Moreover, if $C(t) = R_2(t)\ R_o(t)$ is known, it is easy to see that $S(T)$ is still a sufficient statistic for μ^{β}, and that an estimator of β can be obtained in the same way as in (20) by putting $\hat{B}(t)$ instead of $B(t)$.

5. Examples

Three examples of applications of our results will be given. When $z(t)$ is a Wiener process the best estimators have an analogous form as those obtained in the case of the classical linear models. We deal with this situation in Example 1. Example 2 illustrates how one may apply Lemma 5 for certain Gaussian processes. In these two examples $C(t) = 0$. Example 3 is free of this assumption, however it deals with the case when a best estimator exists and does not depend on $C(\cdot)$.

Example 1. Let $y(t) = X(t)\beta + w(t)$ be a continuous stochastic process, where $w(t)$ is an n-dimensional Wiener process. Then $A(t) = X_1(t)$, $B(t) = I$ and $C(t) = 0$. If $\int_0^T X_1'(t)\, X_1(t)\, dt$ is positive definite, then, in view of (20),

$$\hat{\beta}_T = \left(\int_0^T X_1'(t)\, X_1(t)\, dt \right)^{-1} \int_0^T X_1'(t)\, d\, y(t).$$

In the particuler case when $X(t) = t\,X$ this formula reduces to $\hat{\beta}_T = \frac{1}{T}\,(X'X)^{-1}\, X'y(T)$.

Example 2. A covariance operator $R(\cdot,\cdot)$ of a Gaussian Markov process can be represented as $R(t,s) = U(t)\, V(s)$ for $t \geqslant s$, where $U(t)$ and $V(s)$ are $n \times n$ matrices. This result was obtained by Timoszyk [15] in the one dimensional case. For a generalization to the n-dimensional case see [13].

Let us now consider a process for which $R(t,s) = V(s)$. Assume that $V(\cdot)$ is unknown, and that $V_1(t) = \lim_{h \to 0} \frac{1}{h}\,(V(t+h) - V(t))$ is continuous. From Lemma 5 it follows that there exists $\mathcal{F}_t^y$-measurable process $\hat{V}(t)$ such that V and $\hat{V}$ are indistinguishable. If

$$\int_0^T X_1'(t)\, \hat{V}_1^{+}(t)\, X_1(t)\, dt$$

is positive definite then the best unbiased estimator of β is of the following form

$$\hat{\beta}_T = \left(\int_0^T X_1(t)\ V_1^+(t)\ X_1(t)\,dt \right)^{-1} \int_0^T X_1(t)\ V_1^+(t)\ d\,y(t).$$

E x a m p l e 3. Let us consider the process $y(t) = t\,x + w(t)$ where $w(t)$ is a Wiener process, while x is normally distributed with an unknown expectation β and a known covariance matrix Γ. Finally, we assume that x and $w(\cdot)$ are independent. Then $y(t) = t\beta + z(t)$, where $z(t) = t(x-\beta) + w(t)$ is a Gaussian Markov process. For this process we have

$$R(t,s) = t\,s\,\Gamma + t \wedge s\,I.$$

Easy calculations show that $A(t) = (I + t\,\Gamma)^{-1}$, $B(t) = I$ and that $C(t) = \Gamma(I + t\Gamma)^{-1}$. Moreover, using

$$\frac{d\,A^{-1}(t)}{dt} = -\,A^{-1}(t)\,\frac{d\,A(t)}{dt}\,A^{-1}(t),$$

we obtain

$$\sum(T) = T(I + T\,\Gamma)^{-1}.$$

In view of (21) and (20) we have

$$\hat{\beta}_T = \frac{1}{T}(I + T\Gamma) \int_0^T (I+t\Gamma)^{-1}\left[d\,y(t) - \Gamma(I+t\Gamma)^{-1}y(t)\,dt\right]. \qquad (22)$$

From the Ito formula it follows that

$$d(I + t\Gamma)^{-1}(y(t) - t\beta) = (I + t\Gamma)^{-1}d\,w(t). \qquad (23)$$

Combining (21), (22), (23) and using Theorem 1 we find

$$\hat{\beta}_T = \frac{1}{T}\,y(T).$$

Because $\hat{\beta}$ is independent of Γ, these results hold also when Γ is unknown.

Acknowledgement. We wish to thank T. Bednarski and A. Kozek for useful comments and discussions.

References

[1] A a l e n, O., Statistical inference for a family of counting processes, Ph. D. dissertation, Univ. of California, Berkeley. Reprinted by the Copenhagen University Institute of Mathematical Statistics (1975).

[2] A r a t o, M., Exact formulas for densities of measures of elementary Gaussian processes, Studia Scientiarum Mathematicarum Hungarica, 5, 17-27 (1970).

[3] - On estimation of parameters of stochastic linear differential equations, ibidem, 5, 11-15 (1970).

[4] D e l l a c h e r i e, C., Capacites et processus stochastiques, Springer - Verlag Berlin 1972.

[5] D i o n, J. P. and K e i d i n g, N., Statistical inference in branching processes, Preprint No 10, Institute of Mathematical Statistics University of Copenhagen (1977).

[6] F i s k, D. L., Sample quadratic variation of simple second order martingales. Z. Wahrscheinlichkeitstheorie verw. Geb., 6, 273-278 (1966).

[7] H a j e k, J., On a simple linear model in Gaussian processes, Proc. of the Second Prague Conf. on Prob. Th. (1962).

[8] - On linear statistical problems in stochastic processes, Czechoslovak Math. J., 12, 404-444 (1962).

[9] H o l e v o, A. S., On estimates of regression coefficients, Theory of Probability and its Applications 14, 79-104 (1969).

[10] L e B r e t o n, A. On continuous and discrete sampling for parameter estimation in diffusion type processes, Mathematical Programming Study 5, 124-144 (1976).

[11] L e B r e t o n, A., Parameter estimation in a linear stochastic differential equation, Trans. of the Seventh Prague Conference on Information Theory (1974).

[12] L e h m a n n, E. L., Testing Statistical Hypotheses, New York 1959.

[13] L i p c e r, R. S. and S h i r y a y e v, A. N., Statistics of random processes I/II, Springer - Verlag Berlin 1977, 1978.

[14] R a o, C. R., Linear statistical inference and its applications, New York 1973.

[15] T i m o s z y k, W., A characterization of Gaussian processes that are Markovian, Colloquium Mathematicum, 30, 157-167(1974).

ESTIMATION FOR SOME CLASSES OF GAUSSIAN MARKOV PROCESSES

by

Marek Musiela and Roman Zmyślony

Polish Academy of Sciences, Wrocław

1. Introduction

In the paper we consider a class of Gaussian Markov processes $y(t)$. For a fixed $t \in [0,T]$, the random vector $y(t)$ may be interpreted as a vector of observations in a classical linear model. We give a minimal sufficient statistic and obtain a simple characterization of its completeness. Using the terminology of stochastic integrals we give explicit formulas for estimators of regression coefficients and variance components. Moreover, we prove that they are the best unbiased estimators. Such problems are considered in detection, modulation, communication and control. We use here the same notation and terminology as in [7].

2. A minimal sufficient statistic

Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a complete probability space and let $\{\mathcal{F}_t\}$, $t \geq 0$, be a nondecreasing right continuous family of σ- fields contained in $\mathcal{F}$. Moreover, let $y = (y(t), \mathcal{F}_t)$ be an n-dimensional continuous stochastic process and let $W = (W(t), \mathcal{F}_t)$ be an n-dimensional Wiener process. We assume that the processes y and W are connected by the equation

$$y(t) = X(t)\beta + t\,e + \omega(t). \tag{1}$$

Here $X(t)$ is an $n \times p$ matrix of known twice - continuously differentiable functions, β is a $p \times 1$ vector of parameters running over $\Theta \subset R^p$ and e is an $n \times 1$ normal random vector. We assume that e is $\mathcal{F}_0$ measurable with expectation zero and covariance $\Gamma = \sum_{i=1}^{k} \sigma_i V_i$, where $\sigma = (\sigma_1, \ldots, \sigma_k)'$ is a vector of parameters running over $\Omega \subset R^k$, while $V_1, \ldots, V_k$ are known symmetric matrices. Without loss of generality we may assume that $V_1, \ldots, V_k$ are linearly independent. Moreover, we assume that Θ and Ω have non-empty interiors in R^p and R^k, respectively, and e and W are independent. Throughout the paper we assume also that

$$X(0) = X(T) = 0. \tag{2}$$

First we shall derive a minimal sufficient statistic for parameters β and σ. It is convenient for us to use for this problem the factorization theorem [5]. So we have to find a density function of the measure generated by the process y. Note that $z(t) = t\,e + W(t)$ is a Gaussian Markov process. Moreover, the covariance operator of $z(t)$ has the following form

$$R(t,s) = t \cdot s\,\Gamma + t \wedge s\, I, \tag{4}$$

where I stands for the identity matrix. In view of Theorem 1 in [7] (see also [6]) there exists a Wiener process $v = (v(t), \mathcal{F}_t)$ such that

$$d\,y(t) = (A(t)\beta + \Gamma(I + t\Gamma)^{-1} y(t))dt + dv(t), \tag{5}$$

where

$$A(t) = X_1(t) - \Gamma(I + t\Gamma)^{-1} X(t), \tag{6}$$

while the elements of matrix $X_1(t)$ are the first derivatives of elements of $X(t)$, i.e. $X_1(t)$ is the derivative of $X(t)$.

Now let $\mathcal{C}$ denote the space of all continuous functions from $[0,T]$ to R^n and let $\mathcal{B}$ denote the σ-field of Borel subsets of $\mathcal{C}$. Moreover, let $\mu^{\beta,\Gamma}$ and μ stand for the measures induced by processes $y = (y(t), \mathcal{F}_t)$ and $v = (v(t), \mathcal{F}_t)$, respectively; that is $\mu^{\beta,\Gamma}(B)=P(y\in B)$ and $\mu(B) = P(v\in B)$ for every $B\in\mathcal{B}$. Finally, let $\frac{d\mu^{\beta,\Gamma}}{d\mu}$ denote the Radon-Nikodym derivative.

Theorem 1. The measure $\mu^{\beta,\Gamma}$ is absolutely continuous with respect to the measure μ for every β and Γ. Moreover, the Radon-Nikodym derivative has the form

$$\frac{d\mu^{\beta,\Gamma}}{d\mu}(y) = c(\beta,\Gamma)\ h(T,\ y(T))\times$$

$$\times\exp\left\{-\frac{1}{2T}\ y'(T)(I+T\Gamma)^{-1}y(T) + \beta'\int_0^T X_1'(t)\ dy(t)\right\}, \tag{7}$$

where

$$\log c(\beta,\Gamma) = -\frac{1}{2}\int_0^T \operatorname{tr}\Gamma(I+t\Gamma)^{-1}dt - \frac{1}{2}\beta'\int_0^T A'(t)\ A(t)dt\,\beta$$

and

$$\log h(T,\ y(T)) = -\frac{1}{2T}y'(T)\ y(T).$$

P r o o f: Since y has the differential representation given by (5) $\mu^{\beta,\Gamma}$ is absolutely continuous with respect to μ for every β and Γ and

$$\frac{d\mu^{\beta,\Gamma}}{d\mu}(y) = \exp\left\{\int_0^T \left(A(t)\beta + \Gamma(I+t\Gamma)^{-1}y(t)\right)\ dy(t) + \right.$$

$$\left. -\frac{1}{2}\int_0^T (A(t)\beta + \Gamma(I+t\Gamma)^{-1}y(t))'(A(t)\beta + \Gamma(I+t\Gamma)^{-1}\ y(t))\ dt\right\}, \tag{8}$$

(see [6], [7]).

Now, let $f(t,\ x)$, $t\in[0,T]$, $x\in R^n$ stand for the real function defined by formula

$$f(t, x) = \frac{1}{2} x'\Gamma(I+t\Gamma)^{-1} x + \beta' A'(t) x +$$
$$- \frac{1}{2} \int_0^t tr\Gamma(I+s\Gamma)^{-1} ds - \frac{1}{2}\beta' \int_0^t A'(s) A(s) ds\beta. \tag{9}$$

Let $A_1(t)$ and $X_2(t)$ stand for the derivatives of $A(t)$ and $X_1(t)$, respectively. From (6) we have

$$A_1'(t) = X_2(t) - A'(t)\Gamma(I + t\Gamma)^{-1} \tag{10}$$

The Ito formula and formulas (5) and (8) yield

$$df(t, y(t)) = - \frac{1}{2} y'(t)\Gamma(I+t\Gamma)^{-1}\Gamma(I+t\Gamma)^{-1} y(t) +$$
$$+ \beta' A_1'(t)y(t) - \frac{1}{2} \beta' A'(t) A(t)\beta +$$
$$+ (A(t)\beta + \Gamma(I+t\Gamma)^{-1}y(t))' (A(t)\beta + \Gamma(I+t\Gamma)^{-1}y(t))dt +$$
$$+ (A(t)\beta + \Gamma(I+t\Gamma)^{-1}y(t))' d v(t). \tag{11}$$

In view of (9), (10) and (11) we have

$$f(T, y(T)) - \beta'\int_0^T X_2'(t) \; d v(t) +$$
$$= \int_0^T (A(t)\beta + \Gamma(I+t\Gamma)^{-1} y(t))' d v(t) + \tag{12}$$
$$+ \frac{1}{2}\int_0^T (A(t)\beta + \Gamma(I+t\Gamma)^{-1}y(t))' (A(t)\beta + \Gamma(I+t\Gamma)^{-1}y(t)) dt.$$

From the Ito formula it follows that

$$\beta' X_1'(T)y(T) - \beta' \int_0^T X_2'(t) y(t) dt = \beta' \int_0^T X_1'(t) dy(t). \tag{13}$$

Moreover, it is clear that

$$\Gamma(I+t\Gamma)^{-1} = \frac{1}{t}(I - (I+t\Gamma)^{-1}). \tag{14}$$

Finally, using (2), (5), (8), (12), (13) and (14) we obtain the required result.

Note that in view of (7) and the factorization theorem [5] it follows that $((\int_0^T X_1'(t) dy(t))', y'(T))$ is a sufficient statistic for the family $\{\mu^{\beta,\Gamma}\}$. However, in general, this statistic is not minimal sufficient. Now, we state a lemma which will be used to the minimal sufficiency and completeness of this statistic.

<u>Lemma 1.</u> Vectors $\int_0^T X_1'(t)\, dy(t)$ and $y(T)$ are independent.

P r o o f: Note that $\int_0^T X_1'(t)\, dy(t)$ and $y(T)$ are Gaussian random vectors and that $E\, y(T) = 0$. Then it is sufficient to prove that

$$E \int_0^T X_1'(t)\, dy(t)\, y'(T) = 0. \tag{15}$$

The Ito formula yields

$$E \int_0^T X_1'(t) dy(t) y'(T) = X_1'(T) E y(T) y'(T) - \int_0^T X_2'(t) E y(t) y'(T) dt. \tag{16}$$

But, in view of (2), it follows that for every $t \in [0,T]$ we have

$$E\, y(t)\, y'(T) = tI + tT\Gamma. \tag{17}$$

So, using (16) and (17) we obtain

$$E \int_0^T X_1'(t)\, dy(t)\, y'(T) = X_1'(T)(I + T\Gamma)T - \\ - \int_0^T X_2'(t)(tI + tT\Gamma)\, dt. \tag{18}$$

Since the expression on the right hand side of (18) is equal to $\int_0^T X_1'(t)\, dt\, (I + T\Gamma)$ the lemma follows from (2) and (15).

Now, we shall find a minimal sufficient statistic for $\{\mu^{\beta,\Gamma}\}$. Note that set $\left\{I + T\Gamma \,\middle|\, \Gamma = \sum_{i=1}^{k} \sigma_1 V_i, \sigma \in \Omega\right\}$ contains an open subset of a linear manifold $\vartheta = I + \vartheta_1$, where ϑ_1 is a subspace generated by $V_1, \ldots, V_k$. Let ω be the smallest manifold such that

$$\left\{(I + T\Gamma)^{-1} \,\middle|\, \sigma \in \Omega\right\} \subset \omega .$$

Then there exists a matrix W_0 such that $\omega = W_0 + \omega_1$, where ω_1 is a subspace parallel to ω.

Let $W_1, \ldots, W_r$ be a basis for ω_1. Then $(I + T\Gamma)^{-1}$ can be uniquely represented in the following way

$$(I + T\Gamma)^{-1} = \sum_{i=1}^{r} c_i(\sigma)\, W_i + W_o. \tag{19}$$

Theorem 2. Let $\int_0^T X_1'(t)\, X_1(t)\, dt$ be positive definite. Then the vector

$$Z(T) = \left(\left(\int_0^T X_1'(t)\, dy(t)\right)', y'(T)\, W_i y(T),\ i = 1,\dots, r\right) \tag{20}$$

is a minimal sufficient statistic for $\{\mu^{\beta,\Gamma}\}$.

P r o o f: In view of (19) the density function given by (7) can be represented in the following form

$$\frac{d\mu^{\beta,\Gamma}}{d\mu}(y) = c(\beta,\Gamma) h_o(T,y(T)) \exp\left\{-\frac{1}{2T} \sum_{i=1}^{r} c_i(\sigma) y'(T) W_i\, y(T) + \right.$$

$$\left. + \beta' \int_0^T X_1'(t)\, dy(t)\right\}, \tag{21}$$

where $c(\beta,\Gamma)$ is defined by (7) and

$$\log h_o(T, y(T)) = -\frac{1}{2T} y'(T)\, (I + W_o)\, y(T).$$

Note that $c_1(\sigma),\dots, c_r(\sigma), \beta_1,\dots,\beta_p$ are linearly independent. Moreover, from Lemma 1 it follows that the distribution of $Z(T)$ defined by (20) is not degenerated and the coordinates of this vector are linearly independent. Thus, in view of Theorems 5.2 and 5.3 in [4], we obtain the required result.

3. Completeness and estimation

In this section we shall consider the problem of estimation of parameters β and σ^2. First we shall derive a necessary and sufficient conditions for the completeness of the minimal sufficient statistic $Z(T)$ defined in (20).

Recall that a subspace $\mathcal{D}$ of all symmetric matrices is called a quadratic subspace if $V \in \mathcal{D}$ implies $V^2 \in \mathcal{D}$ (see [8]).

Theorem 3. Let $\int_0^T X_1'(t)\, X_1(t)\, dt$ be positive definite. Then, the following conditions are equivalent:

(a) statistic $Z(T)$ is complete,

(b) ϑ_1 is a quadratic subspace,

(c) $\vartheta_1 = \omega_1$.

P r o o f: The equivalence of (b) and (c) follows from Theorem 5 given in the Appendix provided we put $V_0 = I$. To prove that (b) implies (c) we note that from Corollary 1 in the Appendix it follows that $r = k$ and that the set $\Omega_0 = \{(c_1(\sigma), \dots, c_k(\sigma))' \mid \sigma \in \Omega\}$ contains an open subset of R^k, where $c_i(\sigma)$ are defined by (19). In view of Lemma 1 the density function of $Z(T)$ is given by (21). Then, $Z(T)$ is exponentially distributed and the natural parameter space $\{(\beta', c'(\sigma)) \mid \beta \in \Theta, \sigma \in \Omega\}$ contains an open subset of R^{p+k}. Hence the statistic $Z(T)$ is complete (see [5]).

Now, suppose that ϑ_1 is not equal to ω_1. Since $\vartheta_1 \subset \omega_1$, there exists a vector $W \in \omega_1$, $W \neq 0$, such that $tr(W(I + T\Gamma)) = 0$ for all $\Gamma = \sum_{i=1}^{k} \sigma_i V_i$. Moreover, $E\, y'(T)\, W\, y(T) = tr(W\, E\, y(T)\, y'(T)) =$

$= T\, tr\,(W(I + T\Gamma))$. Hence $E\, y'(T)\, W\, y(T) = 0$. Because $y'(T)\, W\, y(T)$ is a function of the minimal sufficient statistic $Z(T)$ and because $P(y'(T)\, W\, y(T) \neq 0) = 1$, statistic $Z(T)$ is not complete.

Remark 1. If ϑ_1 is a quadratic subspace, then

$$Z_1(T) = ((\int_0^T X_1'(t)\, dy(t))', \; y'(T)\, V_i\, y(T),\; i = 1, \dots, k)$$

is a minimal sufficient and complete statistic for $\{\mu^{\beta,\Gamma}\}$.

Now, we are able to describe the best unbiased estimators (B U E for short) of β and σ. An estimator $\hat{\beta}$ of β is called unbiased if $E\,\hat{\beta} = \beta$ for every β and σ. Moreover, an unbiased estimator $\hat{\beta}$ of β is called B U E if the matrix $E(\bar{\beta} - \beta)(\bar{\beta} - \beta)' - E(\hat{\beta} - \beta)(\hat{\beta} - \beta)'$ is nonnegative definite for every unbiased estimator $\bar{\beta}$ of β and for every β and σ.

Let $A = (a_{ij})$ stand for a symmetric matrix, where $a_{ij} = \operatorname{tr}(V_i V_j)$ for $i,j = 1,\dots, k$. Note that A^{-1} exists because $V_1,\dots, V_k$ are linearly independent. Moreover, let $u(T) = (u_1(T),\dots, u_k(T))'$ stand for a random vector, where $u_i(T) = \frac{1}{T^2}\, y'(T)\, V_i\, y(T) - \frac{1}{T} \operatorname{tr} V_i$ for $i = 1,\dots, k$. We prove the following

Theorem 4. Let $\int_0^T X_1'(t)\, X_1(t)\, dt$ be positive definite. Then

(i) $\hat{\beta}_T = \left(\int_0^T X_1'(u)\, X_1(u)\, du\right)^{-1} \int_0^T X_1'(t)\, dy(t)$

and

(ii) $\hat{\sigma}_T = A^{-1}\, u(T)$

are the B U E for β and σ, respectively, if and only if ϑ_1 is a quadratic subspace.

P r o o f: Note that

$$E \int_0^T X_1'(t)\, dy(t) = \int_0^T X_1'(t)\, X_1(t)\, dt\, \beta.$$

Hence in view of (i) we have $E\hat{\beta}_T = \beta$. Moreover, after simple calculations we obtain

$$E\, u(T) = \frac{1}{T}(TA\sigma + v - v) = A\sigma,$$

where $v = (\operatorname{tr} V_1,\dots, \operatorname{tr} V_k)'$. This implies that $E\,\hat{\sigma}_T = \sigma$. Since ϑ_1 is a quadratic subspace $Z(T)$ is complete. Thus, $\hat{\beta}_T$ and $\hat{\sigma}_T$ are the B U E because they are functions of a complete sufficient statistic (see Remark 1).

Now let $\hat{\beta}_T$ and $\hat{\sigma}_T$ be the B U E for β and σ, respectively. Then, in particular, $\hat{\sigma}_T$ is the best unbiased estimator in the class of quadratic estimators of the form $y'(T)\, A\, y(T)$, where A is a symmetric matrix. But from Theorem 2 in [10] it follows that the subspace generated by

$$\left\{ I + T\Gamma \;\middle|\; \Gamma = \sum_{i=1}^{k} \sigma_i\, V_i,\ \sigma \in \Omega \right\}$$

is a quadratic subspace of all symmetric matrices. This implies that ϑ_1 is also a quadratic subspace.

Remark 2. In view of Theorem 3 in [9] there exists B U E for σ if and only if there exists a best quadratic unbiased estimator for σ.

4. Appendix

Let V_0 be an $n \times n$ regular and symmetric matrix and let ϑ_1 be a linear subspace of all $n \times n$ symmetric matrices. Then

$$\vartheta = V_0 + \vartheta_1 = \{V_0 + V \mid V \in \vartheta_1\}$$

is a linear manifold. If $V_0 \in \vartheta_1$, then $\vartheta = \vartheta_1$. Let $\mathcal{M}$ be the set of all $M \in \vartheta_1$ such that $V_0 + M$ is regular and let ω be the smallest linear manifold containing the set $\{V^{-1} \mid V \in V_0 + \mathcal{M}\}$. Since $V_0^{-1} \in \omega$, the manifold ω can be represented in the form $\omega = V^{-1} + \omega_1$, where ω_1 is a subspace parallel to ω. Let $\mathcal{L}$ be the set of all $L \in \omega_1$ such that the matrix $V_0^{-1} + L$ is regular. Moreover, let $(V_0^{-1} + \mathcal{L})^{-1}$ stand for the set of all V^{-1} such that $V \in V_0^{-1} + \mathcal{L}$. We shall describe the subspaces $\mathcal{M}$ for which the mapping $\pi(V) = V^{-1}$ from $V_0 + \mathcal{M}$ into $V_0^{-1} + \mathcal{L}$ is open and one to one. For any subspace $\mathcal{D}$ of all symmetric matrices and for any symmetric matrix V let $V\mathcal{D}V = \{VBV \mid B \in \mathcal{D}\}$. Let Σ stand for a symmetric matrix.

Definition 1. A subspace $\mathcal{D}$ of all symmetric matrices is called a Σ-quadratic subspace if $A \in \mathcal{D}$ implies $A\Sigma A \subset \mathcal{D}$.
If Σ is the identity matrix then $\mathcal{D}$ is called a quadratic subspace (see [8]). Each Σ - quadratic subspace is a Jordan Algebra, where

$$A \circ B = (A\Sigma B + B\Sigma A)/2.$$

Theorem 5. The following conditions are equivalent

(i) $\vartheta_1 = V_0\, \omega_1\, V_0$,

(ii)

(iii) $V_0 + \mathfrak{M} = (V_0^{-1} + \mathcal{L})^{-1}$,

(iv) ϑ_1 is V_0^{-1} - quadratic subspace.

P r o o f: The equivalence of (i) and (ii) is obvious. We prove that (i) implies (iii). It is sufficient to show that $(V_0^{-1}+\mathcal{L})^{-1} \subset V_0 + \mathfrak{M}$. If $L \subset \mathcal{L}$, then $V_0 + L$ is regular and in view of (i) we have that $(V_0^{-1} + L)^{-1} = V_0(V_0 + V)^{-1} V_0$ for some $V \in \vartheta_1$. So, from (ii) it follows $V_0(V_0 + V)^{-1} V_0 \subset \vartheta$. Since the above matrix is regular we conclude that it belongs to $V_0 + \mathfrak{M}$. Now we prove that (iii) implies (i). If $\Gamma \in \vartheta_1$, then we have that $V_0 - \lambda\Gamma$ is regular for sufficiently small λ. Moreover, we have that

$$(V_0 - \lambda\Gamma)^{-1} = \sum_{n=0}^{\infty} (V_0^{-1} \lambda \Gamma)^n V_0^{-1} \tag{22}$$

belongs to $V_0^{-1} + \mathcal{L}$, where A^0 stands for the identity matrix. Hence we have

$$\sum_{n=1}^{\infty} (V_0^{-1} \lambda\Gamma)^n V_0^{-1} \in \omega_1 .$$

So, it is clear that

$$V_0^{-1} \Gamma V_0^{-1} + \sum_{n=2}^{\infty} \lambda^{n-1} (V_0^{-1}\Gamma)^n V_0^{-1} \in \omega_1 .$$

It means that $\vartheta_1 \subset V_0 \omega_1 V_0$. Similary, one can prove $\vartheta_1 \supset V_0 \omega_1 V_0$ putting in (22) V_0^{-1} instead of V_0 and taking $\Gamma \in \omega_1$.

We prove that (iv) implies (i). Let ϑ_1 be a V_0^{-1} quadratic subspace. It is clear that for every $n \geqslant 1$ we have $(V V_0^{-1} V)^n \in \vartheta_1$ provided $V \in \vartheta_1$. Since $(V_0 + V)^{-1}$ can be represented in the form $V_0^{-1} + \sum_{n=1}^{k} a_n (V_0^{-1} V V_0^{-1})^n$, for some a_n and k, it is clear that $\omega_1 \subset V_0^{-1} \vartheta_1 V_0^{-1}$. Thus, because $\omega_1 \supset V_0^{-1} \vartheta_1 V_0^{-1}$ holds we have $\omega_1 = V_0^{-1} \vartheta_1 V_0^{-1}$, Now, we prove that (i) implies (iv). Let $V \in \vartheta_1$. Then $\Gamma = V_0^{-1} V V_0^{-1}$ is an element of ω_1 and for sufficiently small λ we have that $V_0^{-1} - \lambda V_0^{-1} V V_0^{-1}$ is regular. Since conditions (i) and (iii) are equivalent we have $(V_0^{-1} - \lambda V_0 V V_0)^{-1} \in V_0 + \mathfrak{M}$. This implies that $V V_0^{-1} V + \sum_{i=2}^{\infty} \lambda^{i-1} (V V_0)^i V \in \vartheta_1$. Thus letting λ tend to zero we obtain $V V_0^{-1} V \in \vartheta_1$. This terminates the proof of the theorem.

We have the following consequence of Theorem 5.

Corollary 1. If ϑ_1 is a V_0^{-1} - quadratic subspace, then function $\Pi: V_0 + \mathfrak{M} \longrightarrow V_0^{-1} + \mathcal{L}$ defined by $\Pi(V) = V^{-1}$ is an open mapping.

References

[1] A r a t o , M., On the statistical examination of continuous state Markov processes I, II, III,IV, Selected Transl.in Math. Statist. and Probability, Vol, 14 (1978).

[2] B a r n d o r f f - N i e l s e n, O., Exponential families; exact theory, Aarhus University Mathematics Institute, Various Publication Series, No. 19 (1970).

[3] L e B r e t o n, A.,Parameter estimation in a linear stochastic differential equation, Trans. of the Seventh Prague Conference on Information Theory (1974).

[4] J e n s e n, S. T., Covariance hypotheses which are linear in both the covariance and the inverse covariance, Institute of Mathematical Statistics, Univ. of Copenhagen, Preprint,1(1975).

[5] L e h m a n n, E. L., Testing statistical hypotheses,New York 1959.

[6] L i p c e r, R. S. and S h i r y a y e v, A. N. Statistics of random processes I/II, Springer-Verlag Berlin 1977 (1978).

[7] M u s i e l a, M. and Z m y ś l o n y, R., Estimation of regression parameters of Gaussian Markov processes,in this volume, Preprint No 155, Institute of Polish Academy of Sciences(1978).

[8] S e e l y, J. Quadratic subspaces and completeness. Ann. Math. Statist., 42, 710-721 (1971).

[9] S e l l y, J. Minimal sufficient statistics and completeness for multivariate normal families, Sankhya, Ser. A, 39, 170-185 (1977).

[10] Z m y ś l o n y, R., On estimation of parameters in linear models. Applicationes Mathematicae, 22, 271-276 (1976).

DERNIERE PAGE DE LA THESE

AUTORISATION DE SOUTENANCE

DOCTORAT D'ETAT

Vu les dispositions de l'article 5 de l'arrêté du 16 avril 1974,

Vu les rapports de M. R. Carmona ....

M. A. Le Breton ....

M. B. Maisonneuve

M. E. Pardoux ....

M. ...Marek MUSIELA.............................. est autorisé à présenter une thèse en vue de l'obtention du grade de DOCTEUR D'ETAT ES SCIENCES.

Fait à Grenoble, le 27 JUIN 1984

UNIVERSITE SCIENTIFIQUE ET MEDICALE DE GRENOBLE

Le Président de l'U.S.M.G.

Le Président

M. TANCHE