

Modèles graphiques gaussiens et sélection de modèles

Nicolas Verzelen

Directeur de thèse : Pascal Massart

Université Paris-Sud 11
INRIA Projet Select

17 décembre 2008

Modèle graphique non orienté

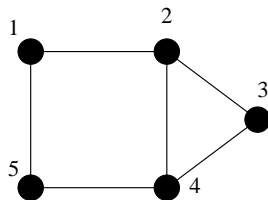
On considère $\mathbf{X} = (X_1, \dots, X_{p+1}) \sim \mathcal{N}_{p+1}(0, \Sigma)$

Σ inversible

$\Gamma := \{1, \dots, p+1\}$

$G = (\Gamma, E)$ graphe non orienté fini

$ne_G(a)$: voisins de a dans G .



Modèle graphique non orienté

On considère $X = (X_1, \dots, X_{p+1}) \sim \mathcal{N}_{p+1}(0, \Sigma)$

Σ inversible

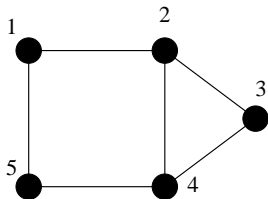
$\Gamma := \{1, \dots, p+1\}$

$G = (\Gamma, E)$ graphe non orienté fini

$ne_G(a)$: voisins de a dans G .

X vérifie la propriété de **Markov local** en a par rapport à G si

$$(X_a \perp\!\!\!\perp X_{-\{a, ne_G(a)\}}) | X_{ne_G(a)}$$



Modèle graphique non orienté

On considère $X = (X_1, \dots, X_{p+1}) \sim \mathcal{N}_{p+1}(0, \Sigma)$

Σ inversible

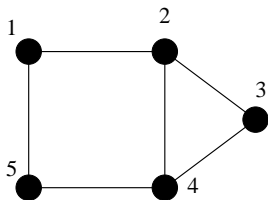
$\Gamma := \{1, \dots, p+1\}$

$G = (\Gamma, E)$ graphe non orienté fini

$ne_G(a)$: voisins de a dans G .

X vérifie la propriété de **Markov local** en a par rapport à G si

$$(X_a \perp\!\!\!\perp X_{-\{a, ne_G(a)\}}) | X_{ne_G(a)}$$

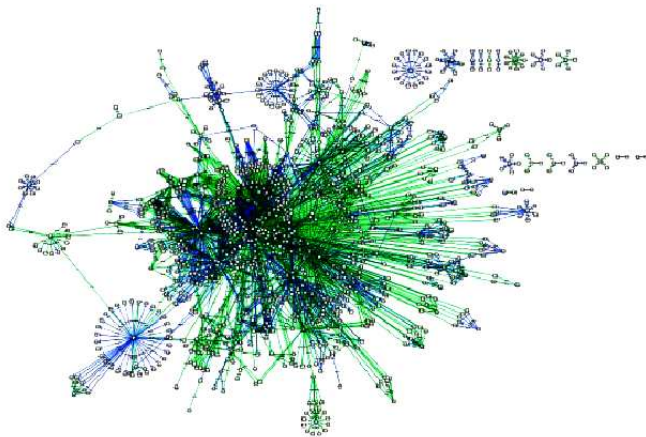


X est un **modèle graphique gaussien** par rapport à G



X vérifie la propriété de Markov local pour tout sommet $a \in \Gamma$.

Ex. 1 : Réseau de régulation génique d'E. Coli

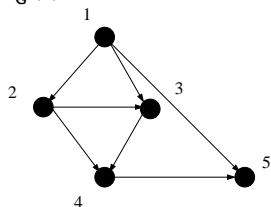


Modèle graphique orienté

$\vec{G} = (\Gamma, \vec{E})$ graphe orienté fini acyclique

$\text{pa}_{\vec{G}}(a)$: parents de a dans \vec{G}

$\text{nd}_{\vec{G}}(a)$: non-descendants de a dans \vec{G} .

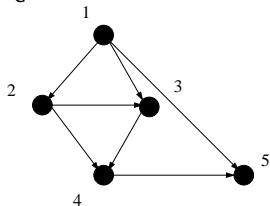


Modèle graphique orienté

$\vec{G} = (\Gamma, \vec{E})$ graphe orienté fini acyclique

$\text{pa}_{\vec{G}}(a)$: parents de a dans \vec{G}

$\text{nd}_{\vec{G}}(a)$: non-descendants de a dans \vec{G} .



X vérifie la propriété de **Markov local orienté** en a par rapport à \vec{G} si

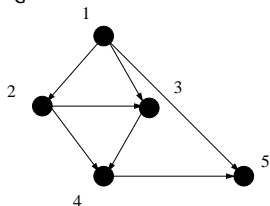
$$(X_a \perp\!\!\!\perp X_{\text{nd}_{\vec{G}}(a)}) | X_{\text{pa}_{\vec{G}}(a)}$$

Modèle graphique orienté

$\vec{G} = (\Gamma, \vec{E})$ graphe orienté fini acyclique

$\text{pa}_{\vec{G}}(a)$: parents de a dans \vec{G}

$\text{nd}_{\vec{G}}(a)$: non-descendants de a dans \vec{G} .



X vérifie la propriété de **Markov local orienté** en a par rapport à \vec{G} si

$$(X_a \perp\!\!\!\perp X_{\text{nd}_{\vec{G}}(a)} \mid X_{\text{pa}_{\vec{G}}(a)})$$

X est un **modèle graphique gaussien orienté** par rapport à \vec{G}



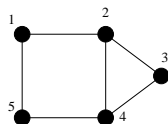
X vérifie la propriété de Markov local orienté pour tout sommet $a \in \Gamma$.

Remarque : Les noeud sont numérotés de façon compatible avec \vec{G} .

Propriété

Modèle graphique non orienté

$$\Sigma_{a,b}^{-1} = 0 \iff (X_a \perp\!\!\!\perp X_b) | X_{-\{a,b\}}.$$

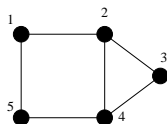


	1	2	3	4	5
1	*	*	0	0	*
2	*	*	*	*	0
3	0	*	*	*	0
4	0	*	*	*	*
5	*	0	0	*	*

Propriété

Modèle graphique non orienté

$$\Sigma_{a,b}^{-1} = 0 \iff (X_a \perp\!\!\!\perp X_b) | X_{-\{a,b\}}.$$



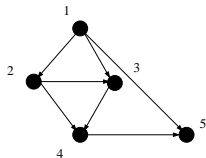
	1	2	3	4	5
1	*	*	0	0	*
2	*	*	*	*	0
3	0	*	*	*	0
4	0	*	*	*	*
5	*	0	0	*	*

Modèle graphique orienté

Décomposition de Cholesky :

$$\Sigma^{-1} = T^* D^{-1} T$$

$$T_{a,b} = 0 \iff (X_a \perp\!\!\!\perp X_b) | X_{<b}$$

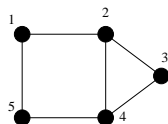


	1	2	3	4	5
1	1	0	0	0	0
2		1	0	0	0
3			1	0	0
4	0			1	0
5		0	0		1

Propriété

Modèle graphique non orienté

$$\Sigma_{a,b}^{-1} = 0 \iff (X_a \perp\!\!\!\perp X_b) | X_{-\{a,b\}}$$



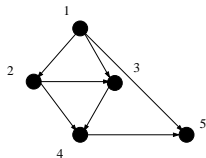
	1	2	3	4	5
1	*	*	0	0	*
2	*	*	*	*	0
3	0	*	*	*	0
4	0	*	*	*	*
5	*	0	0	*	*

Modèle graphique orienté

Décomposition de Cholesky :

$$\Sigma^{-1} = T^* D^{-1} T$$

$$T_{a,b} = 0 \iff (X_a \perp\!\!\!\perp X_b) | X_{<b}$$

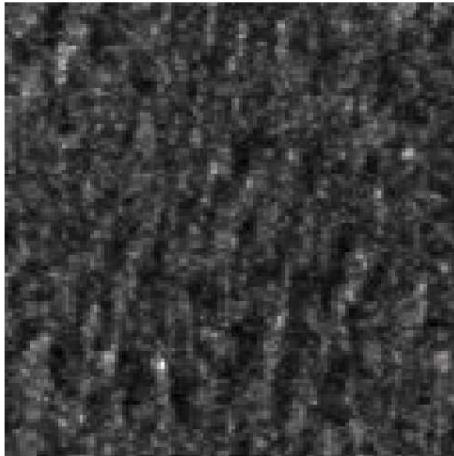


	1	2	3	4	5
1	1	0	0	0	0
2		1	0	0	0
3			1	0	0
4	0			1	0
5		0	0		1

Double intérêt :

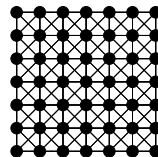
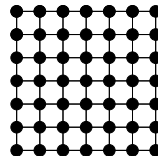
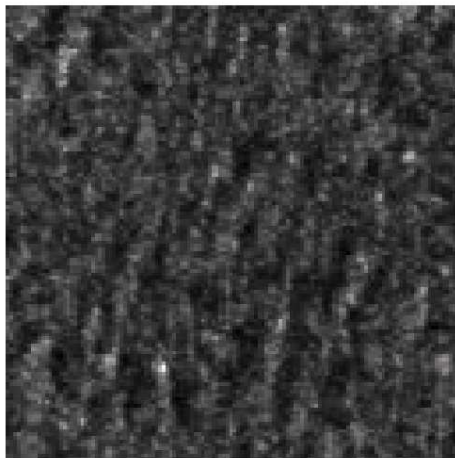
- **Interprétabilité** scientifique du graphe
- **Réduction** de la dimension

Ex. 2 : Analyse d'image



Rue et Tjelmeland (2002)

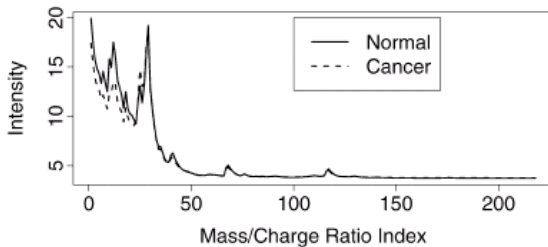
Ex. 2 : Analyse d'image



Rue et Tjelmeland (2002)

Ex. 3 : Spectrométrie de masse

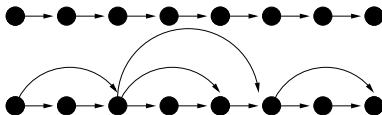
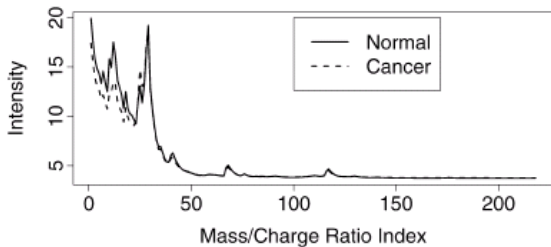
Profils protéiques



Adam et al. (2002), Levina et al. (2008).

Ex. 3 : Spectrométrie de masse

Profils protéiques



Adam et al. (2002), Levina et al. (2008).

Objectifs

DONNÉES : $X \sim \mathcal{N}_{p+1}(0, \Sigma)$.

n observations de $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_{p+1})$

Objectifs

DONNÉES : $X \sim \mathcal{N}_{p+1}(0, \Sigma)$.
 n observations de $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_{p+1})$

QUESTIONS :

- 1 Estimation ou test de la **structure**
- 2 Estimation de la **distribution**.

Objectifs

DONNÉES : $X \sim \mathcal{N}_{p+1}(0, \Sigma)$.
 n observations de $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_{p+1})$

QUESTIONS :

- 1 Estimation ou test de la **structure**
- 2 Estimation de la **distribution**.

CADRE :

- Grande dimension $p \gg n$ (ex : microarray)
- Optimalité et adaptation
- Universalité
- Flexibilité
- Questions numériques

Plan de l'exposé

Analyse de structure

- Test de graphe

- Estimation du graphe

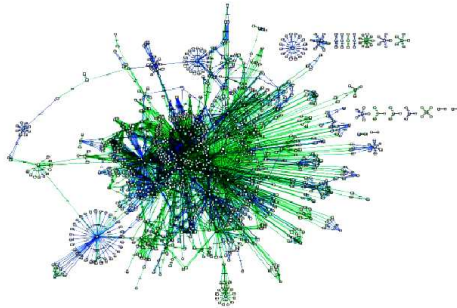
Analyse de covariance

- Vecteur non-stationnaire

- Champs spatiaux

Conclusion

Test de validation de graphe



Données d'expression : $p + 1$ gènes
 n observations de $(\mathbf{X}_1, \dots, \mathbf{X}_{p+1})$

BUT : **tester** que l'on n'a pas oublié de régulation entre 2 gènes.

Test de validation de graphe

Graphe : $G = (\Gamma, E)$ avec $\Gamma = \{1, \dots, p + 1\}$

DONNÉES : n observations de $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_{p+1}) \sim \mathcal{N}(0, \Sigma)$

OBJECTIF : tester

H_0 : « X est un modèle graphique gaussien par rapport à G . »

Moyen : Tester pour chaque sommet a :

$H_{0,a}$: « X suit la propriété de Markov local en a par rapport à G »

$$X_a \perp\!\!\!\perp X_{-\{a, \text{ne}_G(a)\}} \mid X_{\text{ne}_G(a)}$$

Approche

Régression de X_a par rapport X_{-a} : $X_a = \sum_{b \in \Gamma \setminus \{a\}} \theta_b^a X_b + \epsilon_a$, $\epsilon_a \perp\!\!\!\perp X_{-a}$

Test de voisinage pour le sommet a : tester

$$H_{0,a} : \ll X_a \perp\!\!\!\perp X_{-\{a, \text{ne}_{\mathbf{G}}(a)\}} \mid X_{\text{ne}_{\mathbf{G}}(a)} \gg$$

$$\iff H_{0,a} : \ll \theta_{-\text{ne}_{\mathbf{G}}(a)}^a = 0 \gg$$

Approche

Régression de X_a par rapport X_{-a} : $X_a = \sum_{b \in \Gamma \setminus \{a\}} \theta_b^a X_b + \epsilon_a$, $\epsilon_a \perp\!\!\!\perp X_{-a}$

Test de voisinage pour le sommet a : tester

$$H_{0,a} : \langle X_a \perp\!\!\!\perp X_{-\{a, \text{ne}_{\mathbf{G}}(a)\}} | X_{\text{ne}_{\mathbf{G}}(a)} \rangle$$

$$\iff H_{0,a} : \langle \theta_{-\text{ne}_{\mathbf{G}}(a)}^a = 0 \rangle$$

$$Y = \sum_{i=1}^p \theta_i Z_i + \epsilon$$

avec :

- $\theta \in \mathbb{R}^p$ inconnu
- $\epsilon \sim \mathcal{N}(0, \sigma^2)$ et $\sigma^2 = \text{var}(Y|Z)$ inconnue
- $(Z_i)_{1 \leq i \leq p} \sim \mathcal{N}(0, \Psi)$ avec Ψ inconnue mais inversible.
- ϵ indépendant de Z

Approche

Régression de X_a par rapport X_{-a} : $X_a = \sum_{b \in \Gamma \setminus \{a\}} \theta_b^a X_b + \epsilon_a$, $\epsilon_a \perp\!\!\!\perp X_{-a}$

Test de voisinage pour le sommet a : tester

$$H_{0,a} : \langle X_a \perp\!\!\!\perp X_{-\{a, ne_G(a)\}} | X_{ne_G(a)} \rangle$$

$$\iff H_{0,a} : \langle \theta_{-ne_G(a)}^a = 0 \rangle$$

$$Y = \sum_{i=1}^p \theta_i Z_i + \epsilon$$

avec :

- $\theta \in \mathbb{R}^p$ inconnu
- $\epsilon \sim \mathcal{N}(0, \sigma^2)$ et $\sigma^2 = \text{var}(Y|Z)$ inconnue
- $(Z_i)_{1 \leq i \leq p} \sim \mathcal{N}(0, \Psi)$ avec Ψ inconnue mais inversible.
- ϵ indépendant de Z

$$V \subset \mathcal{I} = \{1, \dots, p\}$$

Test du support de θ :

donnés : n observations de (\mathbf{Y}, \mathbf{Z}) .

$$H_0 : \langle \theta_{-V} = 0 \rangle$$

Description de la procédure

↔ Modèle de régression à design fixe. *Baraud et al.* (2003)

Soit $m \subset \mathcal{I} \setminus V$

tester $H_0 : \langle \theta_{-V} = 0 \rangle$. contre l'alternative $H_{1,m} : \langle \theta_{-\{V,m\}} = 0 \rangle$

Description de la procédure

↪ Modèle de régression à design fixe. *Baraud et al.* (2003)

Soit $m \subset \mathcal{I} \setminus V$

tester $H_0 : \langle \theta_{-V} = 0 \rangle$. contre l'alternative $H_{1,m} : \langle \theta_{-\{V,m\}} = 0 \rangle$

On note

- $d = |V|$, $N_m = n - d - |m|$
- Π_V la projection sur l'espace engendré par les \mathbf{Z}_i , $i \in V$
- $\|\cdot\|_n$: norme euclidienne $/\sqrt{n}$

Test de Fisher :

$$\phi_m(\mathbf{Y}, \mathbf{Z}) = \frac{N_m \|\Pi_{V \cup m} \mathbf{Y} - \Pi_V \mathbf{Y}\|_n^2}{|m| \|\mathbf{Y} - \Pi_{V \cup m} \mathbf{Y}\|_n^2}$$

On rejette H_0 si $\phi_m(\mathbf{Y}, \mathbf{Z}) >$ seuil

Sous H_0 , $\phi_m(\mathbf{Y}, \mathbf{Z}) \sim \text{Fisher}(|m|, N_m)$

Description (suite)

\mathcal{M} une classe de modèles tel que $\forall m \in \mathcal{M}, n \geq 1 + |V| + |m|$.

Définition

Soit

$$T_\alpha = \sup_{m \in \mathcal{M}} \left\{ \phi_m(\mathbf{Y}, \mathbf{Z}) - \bar{F}_{|m|, N_m}^{-1}(\alpha_m(\mathbf{Z})) \right\}$$

- $\bar{F}_{|m|, N_m}(u)$ est la probabilité qu'une distribution de Fisher de ddl $|m|$ et N_m soit plus grande que u .
- $\alpha_m(\mathbf{Z})$ collection de nombres dans $]0, 1[$

On rejette $H_0 : \ll \theta_{-V} = 0 \gg$ si $T_\alpha > 0$

\rightsquigarrow choisir collection $\{\alpha_m(\mathbf{Z}), m \in \mathcal{M}\}$

Définition des poids

- (P_1) (Bonferroni) : (α_m) ne dépend pas de \mathbf{Z} et :

$$\sum_{m \in \mathcal{M}} \alpha_m = \alpha$$

Sous (P_1) : $\mathbb{P}_{H_0}(T_\alpha > 0) \leq \alpha$

Définition des poids

- (P_1) (Bonferroni) : (α_m) ne dépend pas de \mathbf{Z} et :

$$\sum_{m \in \mathcal{M}} \alpha_m = \alpha$$

Sous (P_1) : $\mathbb{P}_{H_0}(T_\alpha > 0) \leq \alpha$

- (P_2) : Pour tout $m \in \mathcal{M}$, $\alpha_m(\mathbf{Z}) = q_{\mathbf{Z}, \alpha}$, le α -quantile de la distribution de

$$\inf_{m \in \mathcal{M}} \bar{F}_{|m|, N_m} \left(\frac{\|\Pi_{V \cup m}(\epsilon) - \Pi_V(\epsilon)\|_n^2 / |m|}{\|\epsilon - \Pi_{V \cup m}(\epsilon)\|_n^2 / N_m} \right),$$

conditionnellement au design \mathbf{Z} . \rightsquigarrow Calcul par Monte-Carlo

Sous (P_2) : $\mathbb{P}_{H_0}(T_\alpha > 0) = \alpha$

Définition des poids

- (P_1) (Bonferroni) : (α_m) ne dépend pas de \mathbf{Z} et :

$$\sum_{m \in \mathcal{M}} \alpha_m = \alpha$$

Sous (P_1) : $\mathbb{P}_{H_0}(T_\alpha > 0) \leq \alpha$

- (P_2) : Pour tout $m \in \mathcal{M}$, $\alpha_m(\mathbf{Z}) = q_{\mathbf{Z}, \alpha}$, le α -quantile de la distribution de

$$\inf_{m \in \mathcal{M}} \bar{F}_{|m|, N_m} \left(\frac{\|\Pi_{V \cup m}(\epsilon) - \Pi_V(\epsilon)\|_n^2 / |m|}{\|\epsilon - \Pi_{V \cup m}(\epsilon)\|_n^2 / N_m} \right),$$

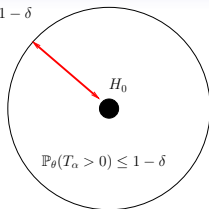
conditionnellement au design \mathbf{Z} . \rightsquigarrow Calcul par Monte-Carlo

Sous (P_2) : $\mathbb{P}_{H_0}(T_\alpha > 0) = \alpha$

Proposition

Le test T_α sous la procédure P_2 est uniformément plus puissant que le test T_α sous la procédure P_1 avec $\alpha_m = \alpha/|\mathcal{M}|$ pour tout $m \in \mathcal{M}$.

$$\mathbb{P}_\theta(T_\alpha > 0) \geq 1 - \delta$$



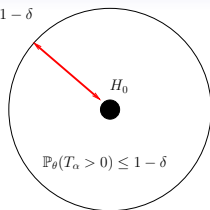
$\Theta[k, p]$: Vecteurs $\theta \in \mathbb{R}^{\mathcal{I}}$ avec au plus k coordonnées non nulles.

H_0 : $\theta = 0$ contre H_1 : $\theta \in \Theta[k, p] \setminus \{0\}$

$\mathcal{M}(k, p)$: Classe des sous-ensembles à k éléments parmi p .

$$\|\theta\|^2 := \text{var}[\sum_{i \in \mathcal{I}} \theta_i Z_i]$$

$$\mathbb{P}_\theta(T_\alpha > 0) \geq 1 - \delta$$



$\Theta[k, p]$: Vecteurs $\theta \in \mathbb{R}^{\mathcal{I}}$ avec au plus k coordonnées non nulles.

H_0 : $\theta = 0$ contre H_1 : $\theta \in \Theta[k, p] \setminus \{0\}$

$\mathcal{M}(k, p)$: Classe des sous-ensembles à k éléments parmi p .

$$\|\theta\|^2 := \text{var}[\sum_{i \in \mathcal{I}} \theta_i Z_i]$$

Proposition

T_α est le test avec la procédure P_1 , la classe $\mathcal{M}(k, p)$, $\alpha_m = \alpha/|\mathcal{M}|$. Si

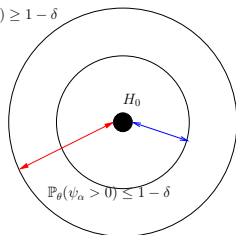
$$n \geq \square[\alpha, \delta] k \log \left(\frac{ep}{k} \right),$$

alors $\mathbb{P}_\theta(T_\alpha > 0) \geq 1 - \delta$ pour tout $\theta \in \Theta[k, p]$ tel que

$$\frac{\|\theta\|^2}{\text{var}(Y|Z)} \geq \frac{\square[\alpha, \delta]}{n} \left[k \log \left(\frac{ep}{k} \right) \vee \sqrt{k} \right]$$

Vitesse minimax de séparation

$$\mathbb{P}_\theta(\psi_\alpha > 0) \geq 1 - \delta$$

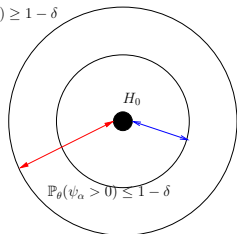


ψ_α : test $H_0 : \theta = 0$. $H_1 : \theta \in \Theta[k, p] \setminus \{0\}$.

Vitesse de séparation du test ψ_α :
 $\rho(\psi_\alpha, \Theta[k, p], \delta)$

Vitesse minimax de séparation

$$\mathbb{P}_\theta(\psi_\alpha > 0) \geq 1 - \delta$$



ψ_α : test $H_0 : \theta = 0$. $H_1 : \theta \in \Theta[k, \rho] \setminus \{0\}$.

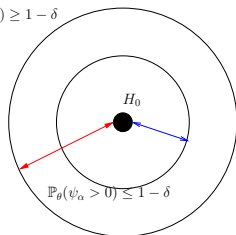
Vitesse de séparation du test ψ_α :
 $\rho(\psi_\alpha, \Theta[k, \rho], \delta)$

Vitesse minimax de séparation :

$$\rho(\Theta, \delta) := \inf_{\psi_\alpha} \rho(\psi_\alpha, \Theta[k, \rho], \delta) .$$

Vitesse minimax de séparation

$$\mathbb{P}_\theta(\psi_\alpha > 0) \geq 1 - \delta$$



ψ_α : test $H_0 : \theta = 0$. $H_1 : \theta \in \Theta[k, p] \setminus \{0\}$.

Vitesse de séparation du test ψ_α :
 $\rho(\psi_\alpha, \Theta[k, p], \delta)$

Vitesse minimax de séparation :

$$\rho(\Theta, \delta) := \inf_{\psi_\alpha} \rho(\psi_\alpha, \Theta[k, p], \delta) .$$

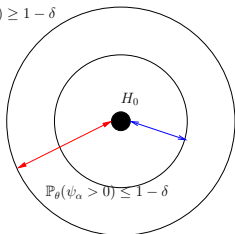
Théorème

Supposons que les covariables Z_i sont indépendantes et que $k \ll \sqrt{p}$, alors

$$\rho(\Theta[k, p], \delta) \geq \square(\alpha, \delta) \frac{k \log(p)}{n}$$

Vitesse minimax de séparation

$$\mathbb{P}_\theta(\psi_\alpha > 0) \geq 1 - \delta$$



ψ_α : test $H_0 : \theta = 0$. $H_1 : \theta \in \Theta[k, p] \setminus \{0\}$.

Vitesse de séparation du test ψ_α :
 $\rho(\psi_\alpha, \Theta[k, p], \delta)$

Vitesse minimax de séparation :

$$\rho(\Theta, \delta) := \inf_{\psi_\alpha} \rho(\psi_\alpha, \Theta[k, p], \delta) .$$

Théorème

Supposons que les covariables Z_i sont indépendantes et que $k \ll \sqrt{p}$, alors

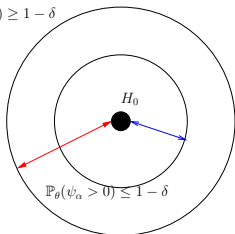
$$\rho(\Theta[k, p], \delta) \geq \square(\alpha, \delta) \frac{k \log(p)}{n}$$

Commentaires :

- Si $k \ll \sqrt{p}$, le test T_α défini précédemment est approximativement minimax.

Vitesses minimax de séparation

$$\mathbb{P}_\theta(\psi_\alpha > 0) \geq 1 - \delta$$



$$\mathbb{P}_\theta(\psi_\alpha > 0) \leq 1 - \delta$$

ψ_α : test $H_0 : \theta = 0$. $H_1 : \theta \in \Theta[k, p] \setminus \{0\}$.

Vitesse de séparation du test ψ_α :
 $\rho(\psi_\alpha, \Theta[k, p], \delta)$

Vitesse minimax de séparation :

$$\rho(\Theta, \delta) := \inf_{\psi_\alpha} \rho(\psi_\alpha, \Theta[k, p], \delta) .$$

Théorème

Supposons que les covariables Z_i sont indépendantes et que $k \ll \sqrt{p}$, alors

$$\rho(\Theta[k, p], \delta) \geq \square(\alpha, \delta) \frac{k \log(p)}{n}$$

Commentaires :

- Si $k \ll \sqrt{p}$, le test T_α défini précédemment est approximativement minimax.
- Si $\sqrt{p} \leq k \leq p$ et $n > (1 + \gamma)p$, la vitesse minimax est de l'ordre de $\frac{\sqrt{p}}{n}$.
- Si $k < \sqrt{p}$ mais proche de \sqrt{p} , les deux bornes ne coïncident pas (à un facteur $\log p$ près).

Estimation du graphe

DONNÉES : n observations $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_{p+1})$.

OBJECTIF : **Estimer** le graphe G .

Estimation du graphe

DONNÉES : n observations $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_{p+1})$.

OBJECTIF : **Estimer** le graphe G .

Meinshausen et Bühlmann (2006), Wille et Bühlmann (2006), Yuan and Lin (2007), Barnerjee et al. (2008),...

Estimation du graphe

DONNÉES : n observations $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_{p+1})$.

OBJECTIF : **Estimer** le graphe G .

Meinshausen et Bühlmann (2006), Wille et Bühlmann (2006), Yuan and Lin (2007), Barnerjee et al. (2008),...

Approche de *Meinshausen et Bühlmann* :

- Estimation de $ne_G(a)$ pour tout $a \in G$:
- $X_a = \sum_{b \in \Gamma \setminus \{a\}} \theta_b^a X_b + \epsilon_a$, \rightsquigarrow Estimation support de θ^a
- $\hat{\theta}^{a,\lambda} := \arg \min \|\mathbf{X}_a - \sum_{b \neq a} \theta_b \mathbf{X}_b\|_n^2 + \lambda \|\theta\|_1$

Estimation du graphe

DONNÉES : n observations $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_{p+1})$.

OBJECTIF : **Estimer** le graphe G .

Meinshausen et Bühlmann (2006), Wille et Bühlmann (2006), Yuan and Lin (2007), Barnerjee et al. (2008),...

Approche de *Meinshausen et Bühlmann* :

- Estimation de $ne_G(a)$ pour tout $a \in G$:
- $\mathbf{X}_a = \sum_{b \in \Gamma \setminus \{a\}} \theta_b^a \mathbf{X}_b + \epsilon_a$, \rightsquigarrow Estimation support de θ^a
- $\hat{\theta}^{a,\lambda} := \arg \min \|\mathbf{X}_a - \sum_{b \neq a} \theta_b \mathbf{X}_b\|_n^2 + \lambda \|\theta\|_1$

$$Y = \sum_{i=1}^p \theta_i Z_i + \epsilon$$

OBJECTIF : Estimation du support de θ

Estimation par moindres carrés

Critère des moindres carrés : $\gamma_n(\theta') := \|\mathbf{Y} - \sum_{i=1}^p \theta'_i \mathbf{Z}_i\|_n^2$.

Soit $m \subset \{1, \dots, p\}$ et $|m| < n$,

$$\hat{\theta}_m := \arg \min_{\theta_{-m}=0} \gamma_n(\theta') = \mathbf{Z}_m^* (\mathbf{Z}_m^* \mathbf{Z}_m)^{-1} \mathbf{Z}_m \mathbf{Y}$$

Moindres carrés intégrés : $l(\theta', \theta) := \mathbb{E} [\sum_{i=1}^p (\theta'_i - \theta_i) \mathbf{Z}_i]^2$.

Estimation par moindres carrés

Critère des moindres carrés : $\gamma_n(\theta') := \|\mathbf{Y} - \sum_{i=1}^p \theta'_i \mathbf{Z}_i\|_n^2$.

Soit $m \subset \{1, \dots, p\}$ et $|m| < n$,

$$\hat{\theta}_m := \arg \min_{\theta_{-m}=0} \gamma_n(\theta') = \mathbf{Z}_m^* (\mathbf{Z}_m^* \mathbf{Z}_m)^{-1} \mathbf{Z}_m \mathbf{Y}$$

Moindres carrés intégrés : $l(\theta', \theta) := \mathbb{E} [\sum_{i=1}^p (\theta'_i - \theta_i) \mathbf{Z}_i]^2$.

$$\begin{aligned} \text{Risque : } \mathbb{E} [l(\hat{\theta}_m, \theta)] &= l(\theta_m, \theta) + [l(\theta_m, \theta) + \sigma^2] \frac{|m|}{n-|m|-1} \\ &= \ll \text{biais} \gg + \ll \text{variance} \gg. \end{aligned}$$

Objectif : **minimiser** $\mathbb{E} [l(\hat{\theta}_m, \theta)]$ sur une collection \mathcal{M} de modèles.

Estimation par moindres carrés

Critère des moindres carrés : $\gamma_n(\theta') := \|\mathbf{Y} - \sum_{i=1}^p \theta'_i \mathbf{Z}_i\|_n^2$.

Soit $m \subset \{1, \dots, p\}$ et $|m| < n$,

$$\hat{\theta}_m := \arg \min_{\theta_{-m}=0} \gamma_n(\theta') = \mathbf{Z}_m^* (\mathbf{Z}_m^* \mathbf{Z}_m)^{-1} \mathbf{Z}_m \mathbf{Y}$$

Moindres carrés intégrés : $l(\theta', \theta) := \mathbb{E} [\sum_{i=1}^p (\theta'_i - \theta_i) \mathbf{Z}_i]^2$.

$$\begin{aligned} \text{Risque : } \mathbb{E} [l(\hat{\theta}_m, \theta)] &= l(\theta_m, \theta) + [l(\theta_m, \theta) + \sigma^2] \frac{|m|}{n - |m| - 1} \\ &= \ll \text{biais} \gg + \ll \text{variance} \gg. \end{aligned}$$

Objectif : **minimiser** $\mathbb{E} [l(\hat{\theta}_m, \theta)]$ sur une collection \mathcal{M} de modèles.

Problèmes :

- design aléatoire
- Variance σ^2 inconnue.

Évaluation sans biais du risque (*Mallows(73)*)

$$\gamma(\theta') := \text{var} \left(Y - \sum_{i=1}^p \theta'_i X_i \right) \Rightarrow l(\theta', \theta) = \gamma(\theta') - \gamma(\theta)$$

$$\begin{aligned} \text{« Risque » :} \quad \mathbb{E} \left[\gamma(\hat{\theta}_m) \right] &= [l(\theta_m, \theta) + \sigma^2] \left(1 + \frac{|m|}{n - |m| - 1} \right) \\ \text{Perte empirique :} \quad \mathbb{E} \left[\gamma_n(\hat{\theta}_m) \right] &= [l(\theta_m, \theta) + \sigma^2] \left(1 - \frac{|m|}{n} \right) . \end{aligned}$$

Évaluation sans biais du risque (Mallows(73))

$$\gamma(\theta') := \text{var} \left(Y - \sum_{i=1}^p \theta'_i X_i \right) \Rightarrow l(\theta', \theta) = \gamma(\theta') - \gamma(\theta)$$

$$\begin{aligned} \text{« Risque » :} \quad \mathbb{E} \left[\gamma(\hat{\theta}_m) \right] &= [l(\theta_m, \theta) + \sigma^2] \left(1 + \frac{|m|}{n - |m| - 1} \right) \\ \text{Perte empirique :} \quad \mathbb{E} \left[\gamma_n(\hat{\theta}_m) \right] &= [l(\theta_m, \theta) + \sigma^2] \left(1 - \frac{|m|}{n} \right). \end{aligned}$$

Heuristique :

$$\gamma(\hat{\theta}_m) \approx \gamma_n(\hat{\theta}_m) \left[1 + 2 \frac{|m|}{n - |m|} + \frac{|m|(|m| - 1)}{(n - |m| - 1)(n - |m|)} \right].$$

Évaluation sans biais du risque (Mallows(73))

$$\gamma(\theta') := \text{var} \left(Y - \sum_{i=1}^p \theta'_i X_i \right) \Rightarrow l(\theta', \theta) = \gamma(\theta') - \gamma(\theta)$$

$$\begin{aligned} \text{« Risque » :} \quad \mathbb{E} \left[\gamma(\hat{\theta}_m) \right] &= [l(\theta_m, \theta) + \sigma^2] \left(1 + \frac{|m|}{n - |m| - 1} \right) \\ \text{Perte empirique :} \quad \mathbb{E} \left[\gamma_n(\hat{\theta}_m) \right] &= [l(\theta_m, \theta) + \sigma^2] \left(1 - \frac{|m|}{n} \right). \end{aligned}$$

Heuristique :

$$\gamma(\hat{\theta}_m) \approx \gamma_n(\hat{\theta}_m) \left[1 + 2 \frac{|m|}{n - |m|} + \frac{|m|(|m| - 1)}{(n - |m| - 1)(n - |m|)} \right].$$

$$\begin{aligned} \text{Critère pénalisé} \quad \hat{m} &:= \arg \min_{m \in \mathcal{M}} \gamma_n(\hat{\theta}_m) [1 + \text{pen}(m)] \\ \hat{m} &:= \arg \min_{m \in \mathcal{M}} \text{Crit}(m). \end{aligned}$$

\mathcal{M}_p^d : collection de tous les sous-ensembles de cardinal inférieur ou égal à d .

Évaluation sans biais du risque (Mallows(73))

$$\gamma(\theta') := \text{var} \left(Y - \sum_{i=1}^p \theta'_i X_i \right) \Rightarrow l(\theta', \theta) = \gamma(\theta') - \gamma(\theta)$$

$$\begin{aligned} \text{« Risque » :} \quad \mathbb{E} \left[\gamma(\hat{\theta}_m) \right] &= [l(\theta_m, \theta) + \sigma^2] \left(1 + \frac{|m|}{n - |m| - 1} \right) \\ \text{Perte empirique :} \quad \mathbb{E} \left[\gamma_n(\hat{\theta}_m) \right] &= [l(\theta_m, \theta) + \sigma^2] \left(1 - \frac{|m|}{n} \right). \end{aligned}$$

Heuristique :

$$\gamma(\hat{\theta}_m) \approx \gamma_n(\hat{\theta}_m) \left[1 + 2 \frac{|m|}{n - |m|} + \frac{|m|(|m| - 1)}{(n - |m| - 1)(n - |m|)} \right].$$

$$\begin{aligned} \text{Critère pénalisé} \quad \hat{m} &:= \arg \min_{m \in \mathcal{M}} \gamma_n(\hat{\theta}_m) [1 + \text{pen}(m)] \\ \hat{m} &:= \arg \min_{m \in \mathcal{M}} \text{Crit}(m). \end{aligned}$$

\mathcal{M}_p^d : collection de tous les sous-ensembles de cardinal inférieur ou égal à d .

\Rightarrow Giraud (2008)

Inégalités oracles

Théorème

Soit $K > 1$. Supposons :

$$\textcircled{1} \quad d \leq \square \frac{n}{\log p},$$

$$\textcircled{2} \quad \text{pen}(m) = K \frac{|m|}{n-|m|} \left(1 + \sqrt{2 \log \left(\frac{ep}{|m|} \right)} \right)^2 \quad \forall m \in \mathcal{M}_p^d,$$

alors l'estimateur $\hat{\theta}_{\hat{m}}$ satisfait

$$\mathbb{E} \left[l(\hat{\theta}_{\hat{m}}, \theta) \right] \leq \square(K) \log p \inf_{m \in \mathcal{M}} \mathbb{E} \left[l(\hat{\theta}_m, \theta) \right] + \tau_n[\theta, K].$$

Inégalités oracles

Théorème

Soit $K > 1$. Supposons :

$$\textcircled{1} \quad d \leq \square \frac{n}{\log p},$$

$$\textcircled{2} \quad \text{pen}(m) = K \frac{|m|}{n-|m|} \left(1 + \sqrt{2 \log \left(\frac{ep}{|m|} \right)} \right)^2 \quad \forall m \in \mathcal{M}_p^d,$$

alors l'estimateur $\hat{\theta}_{\hat{m}}$ satisfait

$$\mathbb{E} \left[l(\hat{\theta}_{\hat{m}}, \theta) \right] \leq \square(K) \log p \inf_{m \in \mathcal{M}} \mathbb{E} \left[l(\hat{\theta}_m, \theta) \right] + \tau_n[\theta, K].$$

Commentaires :

- Borne *non-asymptotique*
- Aucune dépendance en Ψ .
- Adaptation au sens minimax.
- Pas de problèmes de calibration.
- Complexité $|\mathcal{M}|$ simple \Rightarrow procédure praticable.
- Sélection complète \Rightarrow complexité exponentielle.

Vecteur non-stationnaire

DONNÉES : n observations de $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_{p+1})$

OBJECTIF : Estimer Σ grâce aux données

Vecteur non-stationnaire

DONNÉES : n observations de $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_{p+1})$

OBJECTIF : Estimer Σ grâce aux données

Objectifs :

- Analyse en composantes principale :
↪ seuillage de la covariance empirique
Bickel et Levina (2008a,b), *El Karoui* (2008) \Rightarrow convergence en norme d'opérateur.

Vecteur non-stationnaire

DONNÉES : n observations de $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_{p+1})$

OBJECTIF : Estimer Σ grâce aux données

Objectifs :

- Analyse en composantes principale :
 \rightsquigarrow seuillage de la covariance empirique
Bickel et Levina (2008a,b), *El Karoui* (2008) \Rightarrow convergence en norme d'opérateur.
- Analyse linéaire discriminante : $X^+ \sim \mathcal{N}(\mu_+, \Sigma)$, $X^- \sim \mathcal{N}(\mu_-, \Sigma)$
 Décisions $\Rightarrow \frac{(X - \mu_+)^* \Sigma^{-1} (X - \mu_+)}{(X - \mu_-)^* \Sigma^{-1} (X - \mu_-)}$
 \rightsquigarrow estimation de $\Omega = \Sigma^{-1}$.

Vecteur non-stationnaire

DONNÉES : n observations de $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_{p+1})$

OBJECTIF : Estimer Σ grâce aux données

Objectifs :

- Analyse en composantes principale :
 \rightsquigarrow seuillage de la covariance empirique
Bickel et Levina (2008a,b), *El Karoui (2008)* \Rightarrow convergence en norme d'opérateur.
- Analyse linéaire discriminante : $X^+ \sim \mathcal{N}(\mu_+, \Sigma)$, $X^- \sim \mathcal{N}(\mu_-, \Sigma)$
 Décisions $\Rightarrow \frac{(X - \mu_+)^* \Sigma^{-1} (X - \mu_+)}{(X - \mu_-)^* \Sigma^{-1} (X - \mu_-)}$
 \rightsquigarrow estimation de $\Omega = \Sigma^{-1}$.

	Ω creux	T creux
Pas (ou peu) d'ordre	<i>Rothman et al. (2007)</i> <i>Banerjee et al. (2008)</i>	<i>Huang et al. (2006)</i>
Ordre naturel (ex : temporel)		<i>Bickel et Levina (2008a)</i> <i>Levina et al. (2008)</i>

Vecteur non-stationnaire

DONNÉES : n observations de $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_{p+1})$

OBJECTIF : Estimer Σ grâce aux données

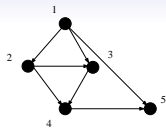
Objectifs :

- Analyse en composantes principale :
 \rightsquigarrow seuillage de la covariance empirique
Bickel et Levina (2008a,b), *El Karoui* (2008) \Rightarrow convergence en norme d'opérateur.
- Analyse linéaire discriminante : $X^+ \sim \mathcal{N}(\mu_+, \Sigma)$, $X^- \sim \mathcal{N}(\mu_-, \Sigma)$
 Décisions $\Rightarrow \frac{(X - \mu_+) * \Sigma^{-1} (X - \mu_+)}{(X - \mu_-) * \Sigma^{-1} (X - \mu_-)}$
 \rightsquigarrow estimation de $\Omega = \Sigma^{-1}$.

	Ω creux	T creux
Pas (ou peu) d'ordre	<i>Rothman et al.</i> (2007) <i>Banerjee et al.</i> (2008)	<i>Huang et al.</i> (2006)
Ordre naturel (ex : temporel)		<i>Bickel et Levina</i> (2008a) <i>Levina et al.</i> (2008)

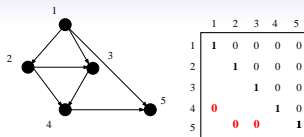
Risque en distance de **Kullback** ?

$S_{\vec{G}}$: Matrices de précision associé à \vec{G} .



	1	2	3	4	5
1	1	0	0	0	0
2		1	0	0	0
3			1	0	0
4				1	0
5					1

$S_{\vec{G}}$: Matrices de précision associées à \vec{G} .



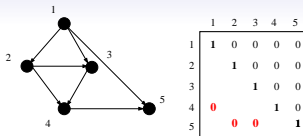
Maximum de vraisemblance : $\hat{\Omega}_{\vec{G}} = \arg \min_{\Omega_1 \in S_{\vec{G}}} \mathcal{L}_n(\Omega_1)$

\vec{G} : collection de graphes orientés.

Critère pénalisé $Crit(\vec{G}) := \mathcal{L}_n(\hat{\Omega}_{\vec{G}}) + \text{pen}(\vec{G})$

$\hat{\vec{G}} := \arg \min_{\vec{G} \in \vec{\mathcal{G}}} Crit(\vec{G}) .$

$S_{\vec{G}}$: Matrices de précision associées à \vec{G} .



Maximum de vraisemblance : $\hat{\Omega}_{\vec{G}} = \arg \min_{\Omega_1 \in S_{\vec{G}}} \mathcal{L}_n(\Omega_1)$

$\vec{\mathcal{G}}$: collection de graphes orientés.

Critère pénalisé $Crit(\vec{G}) := \mathcal{L}_n(\hat{\Omega}_{\vec{G}}) + \text{pen}(\vec{G})$

$\hat{\vec{G}} := \arg \min_{\vec{G} \in \vec{\mathcal{G}}} Crit(\vec{G})$.

$\vec{\mathcal{G}}_{\text{co}}^d$: graphes orientés de degré $\leq d$.

1	0	0	0	0	0	0
	1	0	0	0	0	0
		1	0	0	0	0
			1	0	0	0
				1	0	0
					1	0
						1

$\vec{\mathcal{G}}_{\text{ord}}^d$: graphes orientés ordonnés de degré $\leq d$.

1	0	0	0	0	0	0
	1	0	0	0	0	0
		1	0	0	0	0
			1	0	0	0
				1	0	0
					1	0
						1

Inégalité oracle : $\vec{\mathcal{G}}_{\text{ord}}^d$

Théorème

Soit $K > 1$. Supposons

- $n \geq \square(K)$ et $d \leq \square(K)n$
- $\text{pen}(\vec{G}) = K \sum_{a=1}^{p+1} \frac{|\text{pa}_{\vec{G}}(a)|}{n - |\text{pa}_{\vec{G}}(a)|}, \forall \vec{G} \in \vec{\mathcal{G}}_{\text{ord}}^d$.

alors

$$\mathbb{E} \left[\mathcal{K} \left(\Omega; \hat{\Omega}_{\hat{G}} \right) \right] \leq \square(K) \inf_{\vec{G} \in \vec{\mathcal{G}}_{\text{ord}}^d} \mathbb{E} \left[\mathcal{K} \left(\Omega; \hat{\Omega}_{\vec{G}} \right) \right] + \tau_n .$$

Inégalité oracle : $\vec{\mathcal{G}}_{\text{ord}}^d$

Théorème

Soit $K > 1$. Supposons

- $n \geq \square(K)$ et $d \leq \square(K)n$
- $\text{pen}(\vec{G}) = K \sum_{a=1}^{p+1} \frac{|\text{pa}_{\vec{G}}(a)|}{n - |\text{pa}_{\vec{G}}(a)|}, \forall \vec{G} \in \vec{\mathcal{G}}_{\text{ord}}^d$.

alors

$$\mathbb{E} \left[\mathcal{K} \left(\Omega; \hat{\Omega}_{\hat{G}} \right) \right] \leq \square(K) \inf_{\vec{G} \in \vec{\mathcal{G}}_{\text{ord}}^d} \mathbb{E} \left[\mathcal{K} \left(\Omega; \hat{\Omega}_{\vec{G}} \right) \right] + \tau_n .$$

Commentaires :

- Ne dépend pas des valeurs propres de Σ .
- Calcul rapide
- Calibration
- Adaptation au sens minimax.
- Bornes en risque Frobenius.

Simulation (Levina et al. (2008))

$$p + 1 = 100, n = 100.$$

Pour chaque ligne a , le nombre de termes non nuls de T est tiré uniformément

$$T_{a,b} = 0.5.$$

1	0	0	0	0	0	0	0
	1	0	0	0	0	0	0
0	0	1	0	0	0	0	0
			1	0	0	0	0
0	0			1	0	0	0
0					1	0	0
0	0	0	0			1	0
							1

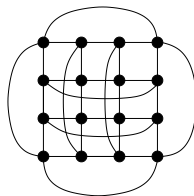
Perte	$K = 3$	N.Lasso	Bickel	Ledoit
Kullback	3.69 ± 0.08	5.74 ± 0.12	8.34 ± 0.65	14.85 ± 0.04
Frob./10 ⁶	1.24 ± 0.15	4.63 ± 0.17	8.17 ± 0.67	12.78 ± 0.01
Op./10 ³	0.73 ± 0.08	1.72 ± 0.04	2.05 ± 0.06	2.36 ± 0.01
Kullback.cov	3.32 ± 0.08	7.05 ± 0.21	14.87 ± 2.47	43.09 ± 0.30
Frob.cov × 10 ³	3.67 ± 0.09	4.81 ± 0.10	7.03 ± 0.42	8.96 ± 0.05
Op.cov × 10 ²	2.25 ± 0.11	2.25 ± 0.05	2.65 ± 0.11	2.95 ± 0.03
Puissance	0.92 ± 0.01	0.78 ± 0.01	0.61 ± 0.05	*
FDR	0.04 ± 0.01	0.36 ± 0.01	0.26 ± 0.04	*

Rapport oracle : **1.25 ± 0.05**.

Champs spatiaux

DONNÉES : 1 observation de $(X_a)_{1 \leq a \leq p+1}$
 X **stationnaire** sur le tore Λ

OBJECTIF : estimer Σ .

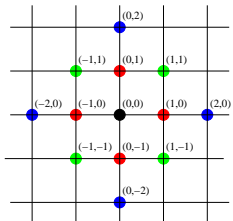
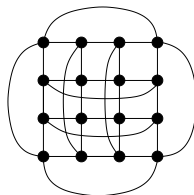


Champs spatiaux

DONNÉES : 1 observation de $(X_a)_{1 \leq a \leq p+1}$
 X **stationnaire** sur le tore Λ

OBJECTIF : estimer Σ .

↪ Approximation par des modèles graphiques



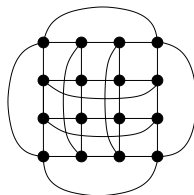
\mathcal{NE} : collection de voisinages ne.

Champs spatiaux

DONNÉES : 1 observation de $(X_a)_{1 \leq a \leq p+1}$
 X **stationnaire** sur le tore Λ

OBJECTIF : estimer Σ .

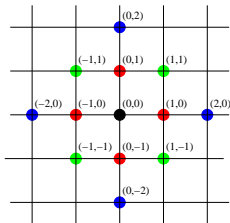
\rightsquigarrow Approximation par des modèles graphiques



Régression conditionnelle :

$$X_{0,0} = \sum_{(i,j) \in \Lambda \setminus \{(0,0)\}} \theta_{i,j} X_{i,j} + \epsilon$$

$$\rightsquigarrow \Sigma^{-1} = \sigma^2(I - C(\theta))$$



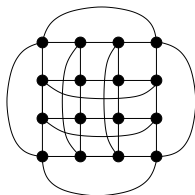
\mathcal{NE} : collection de voisinages ne.

Champs spatiaux

DONNÉES : 1 observation de $(X_a)_{1 \leq a \leq p+1}$
 X **stationnaire** sur le tore Λ

OBJECTIF : estimer Σ .

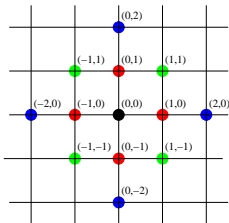
↪ Approximation par des modèles graphiques



Régression conditionnelle :

$$X_{0,0} = \sum_{(i,j) \in \Lambda \setminus \{(0,0)\}} \theta_{i,j} X_{i,j} + \epsilon$$

$$\rightsquigarrow \Sigma^{-1} = \sigma^2(I - C(\theta))$$



Moindres carrés conditionnels (*Besag (1975)*) :

$$\gamma_p(\theta') := \sum_{(i,j) \in \Lambda} \left(X_{i,j} - \sum_{(k,l) \in \Lambda} X_{k,l} \theta_{i-k,j-l} \right)^2$$

$$\hat{\theta}_{ne} := \arg \inf_{\theta' \in \Theta_{ne}} \gamma_p(\theta')$$

\mathcal{NE} : collection de voisinages ne.

Inégalité oracle

$$\hat{n}_e := \arg \inf_{n_e \in \mathcal{N}_E} \left\{ \gamma_p(\hat{\theta}_{n_e}) + \text{pen}(n_e) \right\}$$

Théorème

Soit $K \geq \square$. Si $\text{pen}(m) = K \varphi_{\max}(\Sigma) \frac{|n_e|+1}{p+1}$, alors

$$\mathbb{E} \left[l(\hat{\theta}_{\hat{n}_e}, \theta) \right] \leq \square(K) \inf_{n_e \in \mathcal{N}_E} \{ l(\theta_m, \theta) + \text{pen}(m) \}$$

Inégalité oracle

$$\hat{ne} := \arg \inf_{ne \in \mathcal{NE}} \left\{ \gamma_p(\hat{\theta}_{ne}) + \text{pen}(ne) \right\}$$

Théorème

Soit $K \geq \square$. Si $\text{pen}(m) = K \varphi_{\max}(\Sigma) \frac{|ne|+1}{p+1}$, alors

$$\mathbb{E} \left[l \left(\hat{\theta}_{ne}, \theta \right) \right] \leq \square(K) \inf_{ne \in \mathcal{NE}} \left\{ l(\theta_m, \theta) + \text{pen}(m) \right\}$$

Commentaires :

- Propriétés d'adaptation au sens minimax.
- Calcul rapide

Inégalité oracle

$$\hat{ne} := \arg \inf_{ne \in \mathcal{NE}} \left\{ \gamma_p(\hat{\theta}_{ne}) + \text{pen}(ne) \right\}$$

Théorème

Soit $K \geq \square$. Si $\text{pen}(m) = K \varphi_{\max}(\Sigma) \frac{|ne|+1}{p+1}$, alors

$$\mathbb{E} \left[l \left(\hat{\theta}_{ne}, \theta \right) \right] \leq \square(K) \inf_{ne \in \mathcal{NE}} \left\{ l(\theta_m, \theta) + \text{pen}(m) \right\}$$

Commentaires :

- Propriétés d'adaptation au sens minimax.
- Calcul rapide

Application pratique :

- Extensions à des réseaux non torique
- Calibration : Méthode de la pente *Birgé et Massart* (2007)

Conclusion

Approche :

- Régression linéaire à design gaussien
- Sélection de modèles par pénalisation

Questions :

- Estimation et test de **structure** pour modèles graphiques et **orientés ou non**.
- Estimation de **covariance** pour des modèles graphiques **orientés** (ordre fixé)

Conclusion

Approche :

- Régression linéaire à design gaussien
- Sélection de modèles par pénalisation

Questions :

- Estimation et test de **structure** pour modèles graphiques et **orientés ou non**.
- Estimation de **covariance** pour des modèles graphiques **orientés** (ordre fixé)

Non traité : Estimation de covariance pour des modèles graphiques non orientés

Conclusion

Approche :

- Régression linéaire à design gaussien
- Sélection de modèles par pénalisation

Questions :

- Estimation et test de **structure** pour modèles graphiques et **orientés ou non**.
- Estimation de **covariance** pour des modèles graphiques **orientés** (ordre fixé)

Non traité : Estimation de covariance pour des modèles graphiques non orientés

Propriétés :

- Grande dimension
- **Oracle**, minimax et adaptation
- Pas ou peu d'hypothèse sur Σ
- Procédure rapide si connaissances a priori.
- Problèmes **numériques** en sélection complètes.

Conclusion

Travaux en cours :

- Questions **numériques** :
collaboration avec *C. Giraud et S. Huet*

$$\mathcal{M} \rightsquigarrow \widehat{\mathcal{M}} \rightsquigarrow \widehat{\widehat{m}}$$

Conclusion

Travaux en cours :

- Questions **numériques** :
collaboration avec *C. Giraud et S. Huet*

$$\mathcal{M} \rightsquigarrow \widehat{\mathcal{M}} \rightsquigarrow \widehat{\widehat{m}}$$

- *Design fixe = Design aléatoire ? Variance connue = Variance inconnue ?*

Conclusion

Travaux en cours :

- Questions **numériques** :
collaboration avec *C. Giraud et S. Huet*

$$\mathcal{M} \rightsquigarrow \widehat{\mathcal{M}} \rightsquigarrow \widehat{\widehat{m}}$$

- *Design fixe = Design aléatoire ? Variance connue = Variance inconnue ?*
 \rightsquigarrow Estimation en très grande dimension.

Conclusion

Travaux en cours :

- Questions **numériques** :
collaboration avec *C. Giraud et S. Huet*

$$\mathcal{M} \rightsquigarrow \widehat{\mathcal{M}} \rightsquigarrow \widehat{\widehat{m}}$$

- *Design fixe = Design aléatoire ? Variance connue = Variance inconnue ?*
↪ Estimation en très grande dimension.

Perspectives :

- Adaptation à Ψ en régression ?
- **Estimation** de Ω creux ? Adaptation à l'ordre ?
- Maximum de vraisemblance en stat. spatiales
- Approximations