Structures Markoviennes cachées et modèles à corrélations conditionnelles dynamiques: extensions et applications aux corrélations d'actifs financiers.

Philippe Charlot[1]

[1]GREQAM, 2 rue de la Charité, 13236 Marseille cedex 02, France

25 novembre 2010

1. Introduction
2. Model with Hierarchical HMM
3. Model with Factorial HMM
4. Model with HMDT
5. Conclusion

Motivations and contributions

Q: *Why dynamic correlations?*
A: several empirical studies about stock market behaviors (Longin and Solnik (1995, 1996 and 2001)) show that the hypothesis of constant correlations is not realistic.

Q: *The literature on multivariate GARCH Model is huge, so why another model?*
A: because only a little of them deals with regime switching.

Q: *What is the goal?*
A: a better understanding of the global dynamic of the correlation by introducing new tools in the Markov-switching framework.

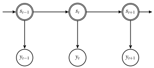
Q: *Contributions*
A: Using specification from Graphical models theory, we present new regime-switching specification for dynamic correlation models.

Dynamic Conditional Correlation models

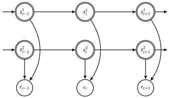
- Brief History:
  - conditional correlations model is originally presented by Bollerslev (1990)
  - extensions to dynamic case suggested by Engle & Sheppard (2001) ans Tse & Tsui (2002)
- Basic Framework
  - assume that the stochastic process $\mathbf{r}_t$ with K elements is generated by:
  $$\mathbf{r}_t|\mathcal{F}_{t-1} \sim \mathcal{L}(0, H_t)$$
  with $\mathcal{F}_{t-1}$ denotes the information set generated by past observations
  - $\mathcal{L}$ is a p.d.f. with mean equal to zero and conditional variance H_t follows :
  $$H_t = D_t R_t D_t$$
  with $D_t = diag(h_{1,t}^{1/2}, ..., h_{K,t}^{1/2})$, a diagonal matrix of the standard deviation
  - standardized residuals are expressed as $\epsilon_t = D_t^{-1}\mathbf{r}_t$ which lead to:
  $$\mathbb{E}_{t-1}[\epsilon_t \epsilon_t'] = D_t^{-1} H_t D_t^{-1} = R_t$$
- Huge literature: this specification has given rise to many extensions.

What is graphical models theory?

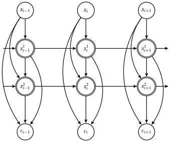
- **Definition:**
  - "Graphical models are a marriage between probability theory and graph theory." *Michael I. Jordan*
  - oriented acyclic graph where valuations are probabilities
- **Applications**
  - hanwritting recognition, source separation, genetics, robot localization, etc.
  - recently introduced in finance with applications in algorithmic trading
- **Framework:**
  - one of the simplest Graphical Model: the Hidden Markov Model (HMM)
    - In a HMM, each observation y_t is linked to state of a hidden Markov chain s_t via a probability
  - Many extensions of HMM have emerged for modeling complex structures:



(a) HMM.

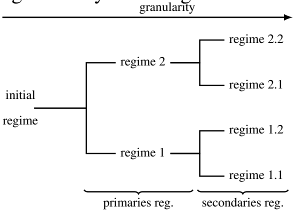


(b) Factorial HMM.



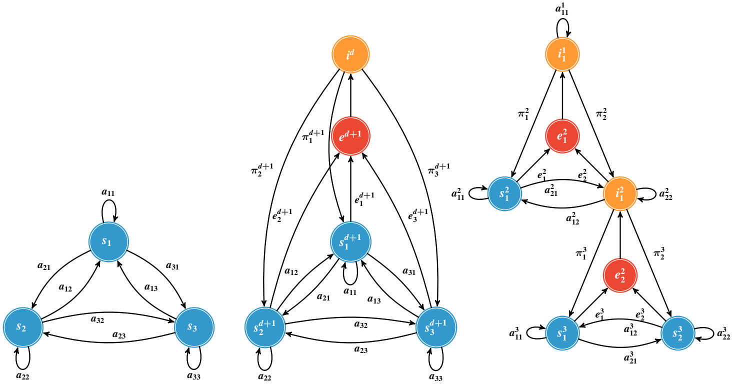
(c) Hidden Markov Decision Tree.

- Regime switching DCC model:
  - Deterministic transition (STAR): STCC of Silvennoinen and Teräsvirta (2005)
  - Stochastic transition (Markov-switching): RSDC of Pelletier (2006)
- How can we modeling multiple switches?
  - the idea is to intruduce more flexibility in the swithing mechanic
  - STCC becomes DSTCC by introducing another transition around the first one
- Q: Is it possible to do this in a Markov-Switching setup?
  A: Yes! with the *Hierarchical Hidden Markov Model...*
  - build a tree-like structure to introduce sub-regimes...
  - ... in order to increase the granularity of the regimes:



  - ... and having a state hierarchy.

- first proposed by Fine and al. (1998) to generalize the HMM model
- idea: build a stochastic process with several levels by adopting a tree structure to obtain an interlacing of regimes
- How? by using different types of states:
  - emitting states: produce observations
  - internal states: abstract states for building hierarchy
  - exiting states: allow to quit a level of the tree
- Graphically: HMM vs Basic HHMM vs Complete HHMM



- Formally, a HHMM can be represented as the process $\{\mathcal{Y}_t, \mathcal{Q}_t\}_{t\in\mathbb{N}}$ with :
  - $\{\mathcal{Y}_t\}_{t\in\mathbb{N}}$ is the process followed by the observations
  - $\{\mathcal{Q}_t\}_{t\in\mathbb{N}}$ is a homogeneous first order Markov chain.

- Each state of a HHMM q_i^d at level d belongs to the set $\mathcal{Q} = \{\mathcal{S}, \mathcal{I}, \mathcal{E}\}$ where:
  - $\mathcal{S}$ is the set of emitting states
  - $\mathcal{I}$ the set of internal states
  - $\mathcal{E}$ the set of exiting states

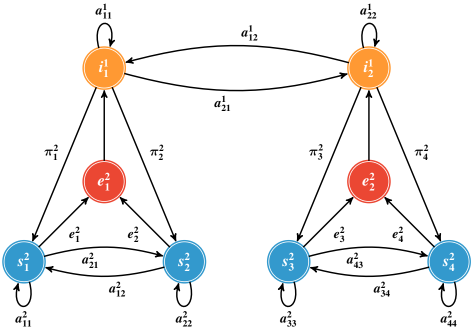
- The dynamic of the hidden structure is driven by three probabilities:
  - $e_i^d = \mathbb{P}[q_{t+1}^d | q_t^{d+1}]$ vertical probability to leads from a child $(ch(k)$ to its parent $pa(k))$
  - $\pi_i^d = \mathbb{P}[q_{t+1}^{d+1} | q_t^d]$ vertical probability to lead from a parent state to its child
  - $\mathcal{A}_k^d = (a_k^d(i,j))$: is the matrix of horizontal transition matrix

- The transition matrix is stochastic if for each sub-model depending of parent k:

$$\sum_{j\in ch(k)} a_k^d(i,j) + e_i^d = 1 \text{ and } \sum_{i\in ch(k)} \pi_i^d = 1$$

where $i, j \in ch(k)$ are two states with parents k.

- Recall our objective: capture thinner nuances in the dynamics of the regime
- we use a HHMM to increase the granularity of the regimes with a 2-levels tree:



- We have two type of regimes:
  - primary regime, defined by emitting states i^1_1 and i^2_2
  - secondary regime, defined by $s^2_i, i = 1, ..., 4$
  - the primaries regimes are then the result of a combination of the secondarie's

- mechanic of the model:
  - The pair of emitting states (s_1^2, s_2^2) and (s_3^2, s_4^2) define sub-HMM
  - The link between these sub-models is provided by the abstract states i_1^1 and i_2^1
  - Like Silvennoinen & Teräsvirta (2007), the correlation process is bounded by four states of constant correlations over time.

- Transition matrix
  - for the two sub-models:

$$A_1^2 = \begin{bmatrix} a_{11}^2 & a_{12}^2 \\ a_{21}^2 & a_{22}^2 \end{bmatrix} \text{ and } A_2^2 = \begin{bmatrix} a_{33}^2 & a_{34}^2 \\ a_{43}^2 & a_{44}^2 \end{bmatrix}$$

    and verified constraints:

$$\begin{cases} a_{11}^2 + a_{21}^2 + e_1^2 = 1 \\ a_{12}^2 + a_{22}^2 + e_2^2 = 1 \end{cases} \text{ and } \begin{cases} a_{33}^2 + a_{43}^2 + e_3^2 = 1 \\ a_{34}^2 + a_{44}^2 + e_4^2 = 1 \end{cases}$$

  - for the first level:

$$A^1 = \begin{bmatrix} a_{11}^1 & a_{21}^1 \\ a_{12}^1 & a_{22}^1 \end{bmatrix}$$

    which verifies:

$$a_{11}^1 + a_{12}^1 = 1 \text{ and } a_{21}^1 + a_{22}^1 = 1$$

  - π_i^2 probabilities, $i = 1, ..., 4$ must verify:

$$\pi_1^2 + \pi_2^2 = 1 \text{ and } \pi_3^2 + \pi_4^2 = 1$$

- Estimation: Multi-step estimation (Engle & Sheppard (2001))
- idea: convert HHMM into a HMM representation
  - Xie's method: build the transition matrix by successive layers
    - vertical transitions are given by:

$$\pi_q = \prod_{d=1}^{D} \pi^d_{q^d}, \; q = 0, \ldots, Q^D - 1$$

    - each layer represents a level of the hierarchy:

$$\tilde{a}^d(q', q) = \prod_{i=d}^{D} e^i_{q'^i} \pi^i_{q^i} \cdot a(q'^d, q^d)$$

    - aggregation of these probabilities leads to the so-called *hypertransition matrix*:

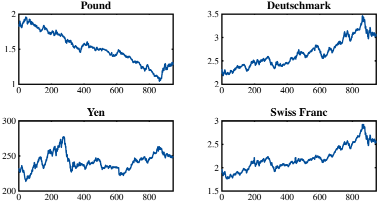
$$\tilde{A}(q', q) = \sum_{i=1}^{D} \tilde{a}^i(q', q)$$

- allow the use of classical filtration tools (Hamilton, Kim) for ML estimation

- Estimate correlations between exchange rate data

Plotted series:

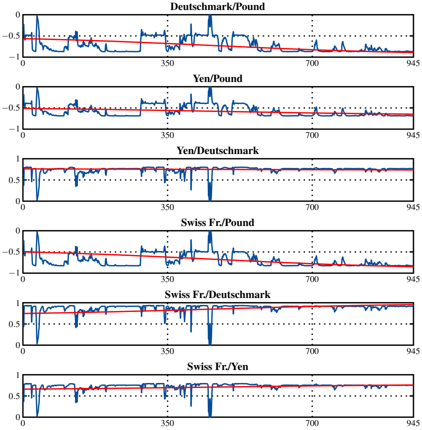
- 4 × 946 observations over the period 1/10/81 to 28/6/85
- week-days close exchange rates against US dollar for Pound, Deutschmark, Yen and Swiss-Franc



Legend for estimated correlations:

- in blue: HRSDC
- in red: TVDSTCC

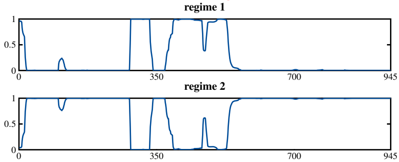
Estimated correlations:



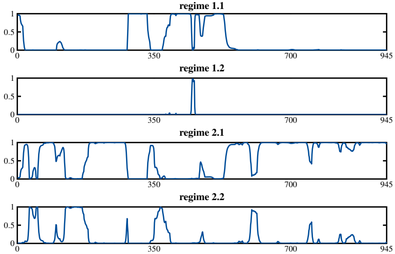
Smoothed probabilities:

- The model clearly identify two primaries regimes...
- ..which are clearly a combination of two sub-regimes
- decomposition of regime 1 exhibits a very punctual sub-regime 1.2
  - correlation are equals to zero
  - extreme but briefly
- capture small elements of the dynamics
- provide a better understanding of the dynamics

Primaries regimes:



Secondaries regimes:



- Existing models
  - Pelletier '06, Billio et al. '05, Hass-Mittnik-Paolella '08 (HMP08)
  - Drawback: in these models the covariance/correlation matrix are assumed to have the same switching date.

- Objective: introducing individual switching dynamic conditions for each element of the correlation matrix.

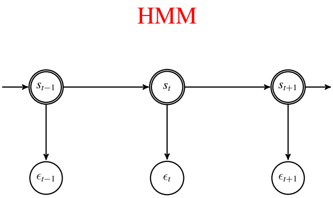
- Starting from HMP '08 models
  - Recall the specification:

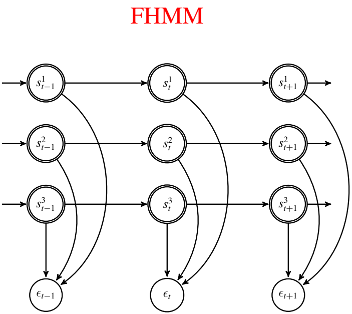
$$\begin{pmatrix} Q_{1,t} \\ \vdots \\ Q_{N,t} \end{pmatrix} = \begin{pmatrix} \Omega_1 \\ \vdots \\ \Omega_N \end{pmatrix} + \begin{pmatrix} a_1 \\ \vdots \\ a_N \end{pmatrix} \epsilon_{t-1}\epsilon'_{t-1} + \begin{pmatrix} b_1 \\ \vdots \\ b_N \end{pmatrix} \odot \begin{pmatrix} Q_{1,t-1} \\ \vdots \\ Q_{N,t-1} \end{pmatrix}$$

  - does not require approximations
  - *most natural* approach for multi-regime GARCH
  - K covariance processes evolve independantly which make easier the identification of low- and high-volatility periods.

- Tool: we use Factorial HMM model to introduce individual switch for the elements of the correlation matrix.

- Origine: introduced by Gharamani & Jordan (*Machine Learning*, 1997)
- generalized the HMM approach by representing the hidden state in a factored form





- Each hidden state is factored into multiple hidden state variables evolving in parallel.
- Each chain has similar dynamics to a basic hidden Markov model.

Specification: we start from HMP '08 with diagonal specification

- let $s_t = (s_t^1, ..., s_t^M)$ with s_t^i a first order Markov chain
- the conditional covariance process of the standardized residuals can be written as:

$$Q_{t,s_t} = C_{s_t} + A_{s_t}\epsilon_{t-1}\epsilon_{t-1}' A_{s_t}' + B_{s_t} Q_{t-1,s_t} B_{s_t}'$$

  with:

$$A_{s_t} = \text{diag}(a_{s_t^1}, a_{s_t^2}, ..., a_{s_t^M})$$

  and similarly:

$$B_{s_t} = \text{diag}(b_{s_t^1}, b_{s_t^2}, ..., b_{s_t^M})$$

Stationnarity conditions:

- $a_{i,s_t^m} a_{i,s_t^m} + b_{i,s_t^m} b_{i,s_t^m} < 1$ with $m = 1, ..., M$ and $i = 1, ...K$
- C_{s_t} symetric and positive definite and the initial covariance matrix to be positive definite.

Equivalent HMM representation:

- any FHMM can be converted in a HMM
- equivalent relationship:

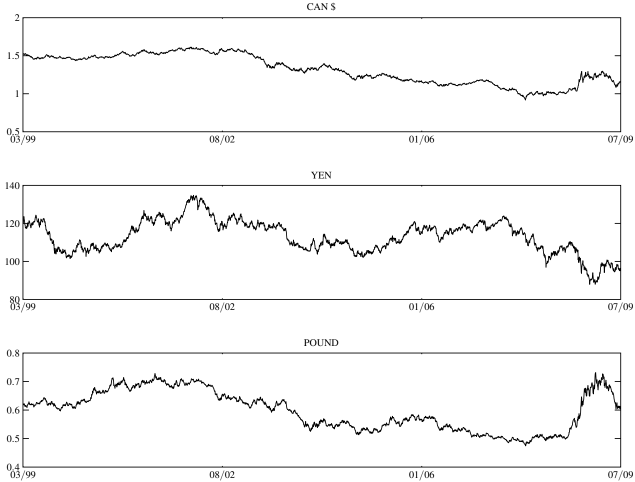
$$\mathbb{P}[s^1_{t+1} = i, ..., s^1_{t+1} = i | s^1_t = j, ..., s^1_t = j] = \prod_{m=1}^{M} \mathbb{P}[s^m_{t+1} = i | s^m_t = j]$$

- Assuming the M Markov chains has N states, our model has M transition matrix of size $N \times N$.
- HMM representation given by: $\Upsilon = \bigotimes_{i=1}^{M} \mathrm{P}^{\mathrm{i}}$
- our M chain with N transition can be represented by a HMM with a $N^M \times N^M$ transition matrix

Estimation with ML:

- Multi-step estimation of Engle-Shephard
- allow the use of classical filtration tools (Hamilton, Kim)

We apply our model on daily dataset with Canadian dollar, Yen and Pound against the US dollar from march 1999 to july 2009 (2697 observations).



Comparaison of two models

- diagonal DCC *à-la* HMP '08 with two regimes:

$$Q_{t,s_t} = \begin{bmatrix} q1_{s_t} & q2_{s_t} & q3_{s_t} \\ & q5_{s_t} & q6_{s_t} \\ & & q7_{s_t} \end{bmatrix} + \begin{bmatrix} a1_{s_t} & 0 & 0 \\ 0 & a2_{s_t} & 0 \\ 0 & 0 & a3_{s_t} \end{bmatrix} \epsilon_{t-1}\epsilon_{t-1}' \begin{bmatrix} a1_{s_t} & 0 & 0 \\ 0 & a2_{s_t} & 0 \\ 0 & 0 & a3_{s_t} \end{bmatrix}' + \begin{bmatrix} b1_{s_t} & 0 & 0 \\ 0 & b2_{s_t} & 0 \\ 0 & 0 & b3_{s_t} \end{bmatrix} Q_{t-1,s_t} \begin{bmatrix} b1_{s_t} & 0 & 0 \\ 0 & b2_{s_t} & 0 \\ 0 & 0 & b3_{s_t} \end{bmatrix}'$$

- our FHMM specification with three chains of two regimes:

$$Q_{t,s_t^{1,2,3}} = \overline{Q}_{s_t^{1,2,3}} + \begin{bmatrix} a1_{s_t^1} & 0 & 0 \\ 0 & a2_{s_t^2} & 0 \\ 0 & 0 & a3_{s_t^3} \end{bmatrix} \epsilon_{t-1}\epsilon_{t-1}' \begin{bmatrix} a1_{s_t^1} & 0 & 0 \\ 0 & a2_{s_t^2} & 0 \\ 0 & 0 & a3_{s_t^3} \end{bmatrix}' + \begin{bmatrix} b1_{s_t^1} & 0 & 0 \\ 0 & b2_{s_t^2} & 0 \\ 0 & 0 & b3_{s_t^3} \end{bmatrix} Q_{t-1,s_t^{1,2,3}} \begin{bmatrix} b1_{s_t^1} & 0 & 0 \\ 0 & b2_{s_t^2} & 0 \\ 0 & 0 & b3_{s_t^3} \end{bmatrix}'$$

HMP: variance targeting is not feasable and each Q_{t,s_t} needs 7 parameters

FHMM-DCC: 8 intercepts matrix to estimate with HMM representation.

- hint.: estimate the constants of the two extrem cases (all covariances are in regime 1, and all are in regime 2), and then construct the six other intercepts.
- Let

$$\overline{Q}_{s_t^1=1,s_t^2=1,s_t^3=1} = \begin{bmatrix} q1_{s_t^1=1} & q2_{s_t^1=1,s_t^2=1} & q3_{s_t^1=1,s_t^3=1} \\ & q5_{s_t^2=1} & q6_{s_t^2=1,s_t^3=1} \\ & & q7_{s_t^3=1} \end{bmatrix}$$

- Exemple:

$$\overline{Q}_{s_t^1=1,s_t^2=2,s_t^3=1} = \begin{bmatrix} q1_{s_t^1=1} & q2_{s_t^1=1,s_t^2=2} & q3_{s_t^1=1,s_t^3=1} \\ & q5_{s_t^2=2} & q6_{s_t^2=2,s_t^3=1} \\ & & q7_{s_t^3=1} \end{bmatrix}$$

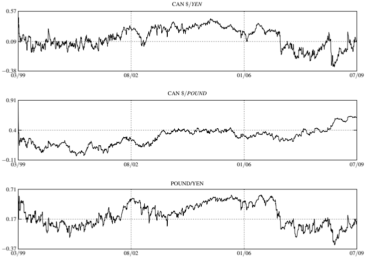
  black elements are coming from the constants of the two extrem cases. red elements corresponds to the additional elements needed to define the intercept in that case.

- and so on for the others...

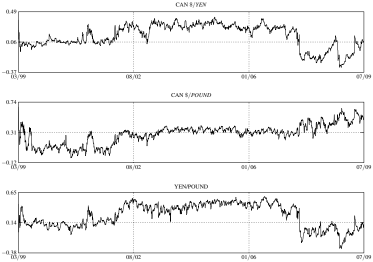
- Number of parameters for the correlations:
  - HMP: 28
  - FHMM-DCC: 42

- Estimated correlations:

HMP

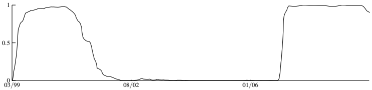


FHMM-DCC

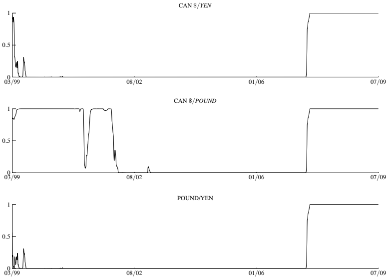


Comparing Smoothed probabilities

HMP:



FHMM-DCC:

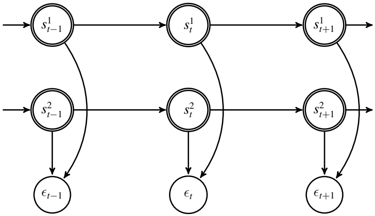


- Contribution: this paper presents a tree-structured dynamic correlation model
- Existing litterature
  - Audrino-Barine-Audresi '06: rolling window averaged conditional correlation estimator based on tree-structured GARCH
  - Audrino-Trojani '05: tree-based model in two step:
    - conditional variance are extracted with tree-structured GARCH
    - conditional correlations computed from standardized residuals and based on a second tree-structured dynamic
  - Dellaportas-Vrontos '07: study volatility and co-volatility asymmetries using a sequence of binary decisions rules to exhibit multivariate thresholds.
- commun feature of theses approaches:
  - all based on the idea of binary tree
  - Time series are recurcively partionned using binary decisions
  - deterministic approach
- Objective: we propose en extension of the DCC based on a stochastic decision tree linking univariate volatility and correlations.

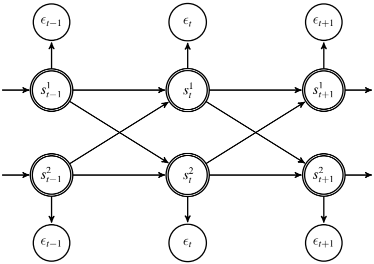
- Origin: Hidden Markov Decision Tree (HMDT) has been proposed by Jordan, Ghahramani and Saul (1997)

- HMDT = Factorial HMM + Coupled HMM

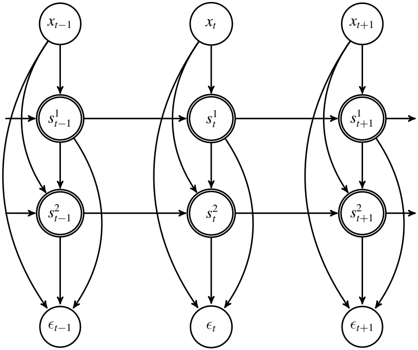
Factorial HMM



Coupled HMM



- Finally, a HMDT looks like:



- factorial decomposition provides a factorised state space.
- hierarchy is done via a coupling transition matrix
- input x_{t-1} is optional

- Objective: study the relationship between univariate volatility and correlation with a tree containing two levels.
  - the first level discriminates between low/high volatility
  - the second level discriminates between low/high correlations
- Hidden structure of the first level
  - we partition the space of the univariate conditional variance in two subspaces, low and high variance,
  - the k^{th} time serie has a 2×2 transition matrix:

$$\mathrm{P}_{\mathrm{vol}}^{\mathrm{k}} = \begin{bmatrix} p_{11}^{k} & 1-p_{22}^{k} \\ 1-p_{11}^{k} & p_{22}^{k} \end{bmatrix}$$

  - individual transition matrix can be aggregate to constitute the first level of the decision tree:

$$\mathrm{P}_{\mathrm{vol}} = \bigotimes_{i=1}^{K} \mathrm{P}_{\mathrm{vol}}^{\mathrm{i}}$$

  - all the dynamics of the K univariate volatilities containing 2 states with a single transition matrix of size $2^K \times 2^K$.
- Hidden structure of the second level
  - we partition the space of the univariate conditional variance in two subspaces, low and high correlations,
  - the decision step is represented by a 2-by-2 transition matrix:

$$\mathrm{P}_{\mathrm{cor}} = \begin{bmatrix} p_{11}^{c} & 1-p_{22}^{c} \\ 1-p_{11}^{c} & p_{22}^{c} \end{bmatrix}$$

- Global partition of the space: the partition of the space is represented by a $2^{K+1} \times 2^{K+1}$ transition matrix P:

$$\mathrm{P} = \mathrm{P}_{\text{vol}} \otimes \mathrm{P}_{\text{cor}}$$

- Coupling matrix: using Abstract Markov chain
  - attribute a weight to the decision related to the correlation given the decision of the univariate volatility
  - transition probability given by:

$$\mathrm{P}_{\text{coupl}} = \begin{bmatrix} c_{11} & 1-c_{22} \\ 1-c_{11} & c_{22} \end{bmatrix}$$

- Specification for volatilities and correlations
  - Specification for univariate volatility: Haas et al. '04
    - $$\begin{pmatrix} h_{1,t} \\ \vdots \\ h_{N,t} \end{pmatrix} = \begin{pmatrix} \omega_1 \\ \vdots \\ \omega_N \end{pmatrix} + \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_N \end{pmatrix} y_{t-1}^2 + \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_N \end{pmatrix} \odot \begin{pmatrix} h_{1,t-1} \\ \vdots \\ h_{N,t-1} \end{pmatrix}$$
    - stationnarity condition implies $\alpha_n + \beta_n < 1$ for each $n = 1, \ldots, N$
  - Specification for the correlations: Haas at al. '08
    - $$\begin{pmatrix} Q_{1,t} \\ \vdots \\ Q_{N,t} \end{pmatrix} = \begin{pmatrix} \Omega_1 \\ \vdots \\ \Omega_N \end{pmatrix} + \begin{pmatrix} a_1 \\ \vdots \\ a_N \end{pmatrix} \epsilon_{t-1}\epsilon_{t-1}' + \begin{pmatrix} b_1 \\ \vdots \\ b_N \end{pmatrix} \odot \begin{pmatrix} Q_{1,t-1} \\ \vdots \\ Q_{N,t-1} \end{pmatrix}$$
    - stationnarities conditions imply the intercept Ω_n to be a positive definite matrix and $a_n + b_n < 1$ for $n = 1, \ldots N$.

- ML estimation in one step:

$$L = -\frac{1}{2}\sum_{t=1}^{T}\left(K\log(2\pi) + \log(|H_t|) + y_t' H_t^{-1} y_t\right)$$

- Equivalent HMM representation:
  - conversion in regular HMM given by:

    $$\mathrm{P}_{\mathrm{reg}} = (\mathrm{P}_{\mathrm{vol}} \otimes \mathrm{P}_{\mathrm{cor}}) \odot (\mathrm{P}_{\mathrm{coupl}} \otimes (\iota\iota'))$$

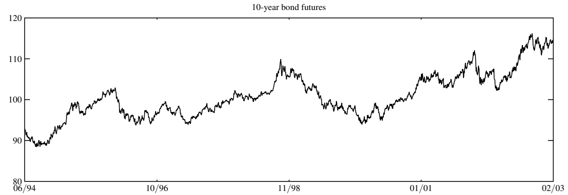
    with ι a vector of ones of length 2^{K+1}.
  - convert a the K-variate problem with $K+2$ Markov chains in a problem with a $2^{K+1} \times 2^{K+1}$ transition matrix.
  - allows the use of Hamilton's filter:

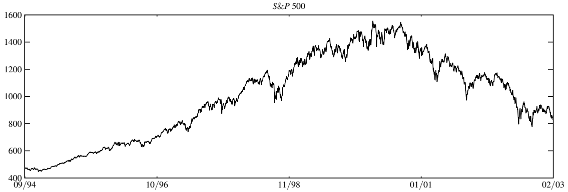
    $$\hat{\xi}_{t|t} = \frac{(\hat{\xi}_{t|t-1} \odot \eta_t)}{\mathbf{1}'(\hat{\xi}_{t|t-1} \odot \eta_t)}$$

    with $\hat{\xi}_{t|t+1} = \mathrm{P}_{\mathrm{reg}} \times \hat{\xi}_{t|t}$
  - and Kim's filter.

Remark: remains a strong optimization problem

We apply our model on bivariate dataset futures prices on 10 year treasury bonds and the S&P 500 from september 1994 to february 2003 (2201 observations).

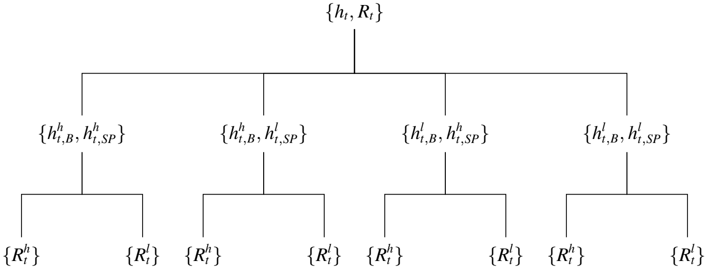




- Number of parameters:
  - first level needs 8 for each series; 12 for the correlations; 2 parameters for the coupling matrix
  - total=30 parameters

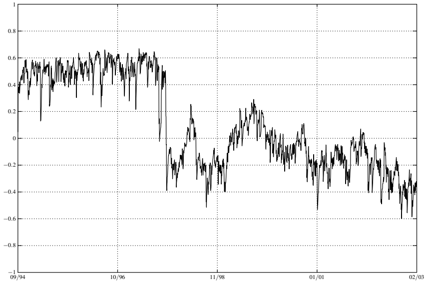
- regular HMM representation: an 8-by-8 transtion matrix

- With two series, the decision process can be summerized as:

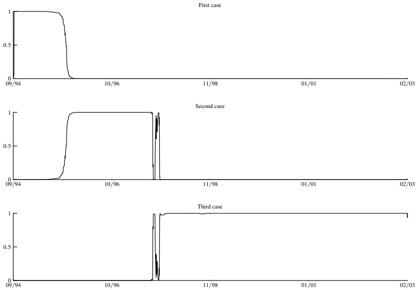


with BD=bond and SP=SP500.

Estimated correlations:



Representative smoothed probabilities:



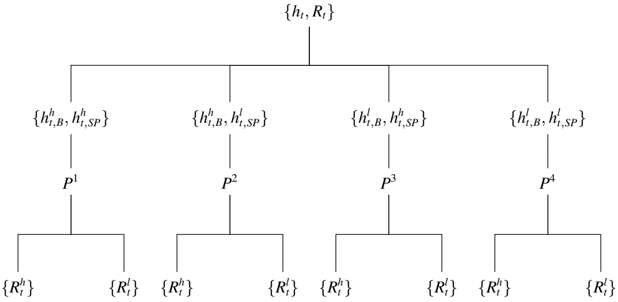
- The model clearly identifies three combinations of regimes:
  1. $\{h^2_{t,B}, h^1_{t,SP}, R^1_t\}$: normal volatility for bond, low volatility for S&P500 and positive correlations regime.
  2. $\{h^2_{t,B}, h^2_{t,SP}, R^1_t\}$: normal volatility for bond, high volatility for S&P500 and positive correlations regime.
  3. $\{h^2_{t,B}, h^1_{t,SP}, R^2_t\}$: normal volatility for bond, low volatility for S&P500 and negative correlations regime.

- Estimated coupling matrix:

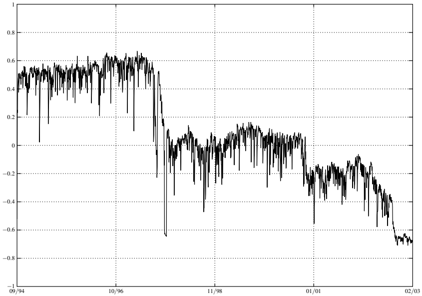
$$\widehat{\mathrm{P}}_{\mathrm{coupl}} = \begin{bmatrix} 0.2522 & 0.1371 \\ 0.7478 & 0.8629 \end{bmatrix}$$

- result means that in general, volatility is associated with negative correlations.

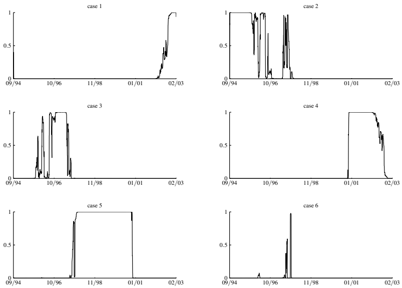
- Deepening of the relationship between first and second level:
- we extend the later specification by introducing a specific relation for all the possible case à the first level.
- graphical representation:



Estimated correlations:



Representative smoothed probabilities:



- Comparing to the previous specification, the introduction of specific coupling matrix increases the explanatory power of the model. Estimation of the model exhibits six cases.

- Strong probability of having positive correlation when:
  - bond and S&P500 are in low volatility regime
  - bond has low volatility and S&P500 has high volatility

- Negative correlations:
  - appears to be associated with a high volatility of the bond
  - more significant when the S&P500 is in low volatility regime

- Our contributions are at the interface between graphical models and dynamic correlations models
  1. special case of RSDC increasing granularity of the regime, based on Hierarchical HMM
  2. Markov-switching DCC where each elements of the correlation matrix have their own switching dynamic, based on the Factorial HMM
  3. Stochastic decision tree to study linkages between unvariate volatility and conditional correlations, based on the Hidden Markov Decision Tree

- Our results show that:
  - classical Markov-switching seems to be sometimes too rigid
  - introducing more flexibility shows new patterns in the dynamic of the correlations

- Directions for further research:
  - asymptotic theory
  - develop more complex specifications
  - improve estimation methods of these models

SELECTED BIBLIOGRAPHY

R.F. Engle, K. Sheppard.
Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH.
UCSD Working Paper, No. 2001-15, 2001.

S. Fine, Y. Singer, N. Tishby.
The Hierarchical Hidden Markov Model : Analysis and Applications.
Machine Learning, 32(1):41-62, 1998.

G. Ghahramani, M.I. Jordan.
Factorial hidden Markov models.
Machine Learning, 29:245-273, 1998.

MI. Jordan, G. Ghahramani, L. Saul
Hidden Markov decision trees.
in *Advances in Neural Information Processing Systems (NIPS) 9*, ed. by M. C. Mozer, M. I. Jordan, and T. Petsche. Cambridge MA: MIT Press.