

# Modélisation bayésienne hiérarchique pour l'estimation de matrice de covariance - Application à la gestion actif-passif de portefeuilles financiers

Mathilde Bouriga, Olivier Féron, Jean-Michel Marin, Christian Robert

## ► To cite this version:

Mathilde Bouriga, Olivier Féron, Jean-Michel Marin, Christian Robert. Modélisation bayésienne hiérarchique pour l'estimation de matrice de covariance - Application à la gestion actif-passif de portefeuilles financiers. 42èmes Journées de Statistique, 2010, Marseille, France, France. 2010. <inria-00494745>

HAL Id: inria-00494745

<https://hal.inria.fr/inria-00494745>

Submitted on 24 Jun 2010

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# MODÉLISATION BAYÉSIENNE HIÉRARCHIQUE POUR L'ESTIMATION DE MATRICE DE COVARIANCE

Application à la gestion actif-passif de portefeuilles financiers

Mathilde Bouriga<sup>1,2</sup>, Olivier Féron<sup>2</sup>, Jean-Michel Marin<sup>3</sup> et Christian P.Robert<sup>1</sup>

<sup>1</sup> Université Paris Dauphine, Paris, France

<sup>2</sup> EDF R&D, Clamart, France

<sup>3</sup> Université, Montpellier, France

*EDF R&D - 1, avenue du Général de Gaulle - 92141 Clamart Cedex - France*

**Résumé :** Ce papier concerne l'estimation de matrices de covariance dans le cas où le nombre de données utilisées pour l'estimation est faible par rapport à la dimension du problème et où les méthodes d'estimation classiques fondées sur le Maximum de Vraisemblance sont peu robustes.

Nous proposons une méthode d'estimation non supervisée fondée sur une modélisation bayésienne hiérarchique du problème d'estimation de matrice de covariance : on pose une loi Inverse Wishart *a priori* pour la matrice, conditionnellement aux hyperparamètres sur lesquels on pose des *a priori* de référence. On considère une matrice cible de structure diagonale. L'estimateur bayésien associé au coût entropique sera approché par des méthodes de type Monte Carlo par chaînes de Markov (MCMC). On comparera empiriquement notre estimateur avec celui obtenu par Maximum de Vraisemblance. L'aspect régularisant de la méthode sera étudié en l'appliquant sur données financières dans un cadre de gestion actif-passif, où le nombre de données est faible par rapport à la taille de la matrice.

**Abstract :** We are in the context where a limited number of observations are available for estimating a large-dimensional normal covariance matrix. In this framework, the traditional estimator - the sample covariance matrix - isn't robust. The resort to bayesian methods affords to regularize the estimator.

The proposed estimator is built from a hierarchical Bayesian model with a shrinkage type prior for covariance matrices : we consider an Inverse Wishart prior on the matrix and we embed a structure in its scale matrix, we put non-informative prior measures on the hyperparameters. This prior shrinks toward a diagonale form. The problem of estimating is under the Stein's loss function and a Markov Chain Monte Carlo sampling scheme is used to implement posterior inference in the proposed model. We compare empirically the resulting Bayes estimator with the traditional estimator, in terms of frequentist risk. The regularizing effect is illustrated by applying the method to a dataset that is high-dimensional relative to the sample size, in a framework of asset and liability management.

MOTS-CLÉS : Méthodes bayésiennes - Statistique mathématique

# 1 Contexte, motivations et objectif

L'étude de l'évolution de portefeuilles financiers est fondée sur une modélisation probabiliste de produits financiers (actions, obligations,...) dans le futur. Dans le but de déterminer les portefeuilles optimaux, la théorie classique de Markowitz [4] repose sur la construction d'une courbe d'efficience dans un univers rendement/risque. La matrice de covariance entre les produits financiers joue un rôle essentiel dans la décision des portefeuilles optimaux. Cette matrice est inconnue et doit être estimée. Bien souvent seul un faible nombre de données historiques peut être utilisé pour estimer une matrice de grande dimension. Dans ce contexte, les méthodes d'estimation classiques sont peu robustes, les matrices estimées mal conditionnées, ce qui peut être dimensionnant dans le calcul des risques portés par les portefeuilles.

De nombreux travaux liés à la théorie de la décision ont été développés pour l'estimation de matrice de covariance. Ils sont principalement fondés sur des méthodes de *shrinkage* qui consistent à réduire la variance des valeurs propres de la matrice de covariance classique [3, 5, 7, 8]. Cela revient généralement à imposer une structure à la matrice. L'utilisation de méthodes bayésiennes, de type shrinkage [1, 2, 6, 12], va permettre de renforcer la robustesse de l'estimateur.

Dans ce papier, nous proposons un estimateur de la matrice de covariance bien conditionné et stable dans le temps. L'approche proposée s'inspire des méthodes classiques de shrinkage approchées sous un angle bayésien. L'objectif consiste alors à construire des lois *a priori* qui ne reflètent pas, dans un premier temps, une information *a priori* particulière, mais qui visent à régulariser comme les méthodes classiques de shrinkage. De plus, la méthode proposée est non supervisée : les hyperparamètres sont modélisés *a priori* par des lois non-informatives et sont conjointement estimées avec la matrice de covariance.

## 2 Modélisation bayésienne proposée pour le problème d'estimation

Dans cette section, nous nous plaçons dans le cadre général où les données  $\mathbf{y}$  sont modélisées par une loi gaussienne centrée multivariée :

$$\mathbf{y}_i \sim \mathcal{N}(\mathbf{0}, \Sigma) , \mathbf{y}_i \in \mathbb{R}^p , \forall i = 1, \dots, n, \quad (1)$$

dont le paramètre inconnu est la matrice de covariance  $\Sigma$ . Nous nous plaçons, de plus, dans un cadre d'estimation bayésienne, qui s'inspire de la modélisation hiérarchique adoptée dans [2]. Cependant, dans notre approche, nous proposons un modèle hiérarchique qui évite toute influence des hyperparamètres dans l'inférence menée sur  $\Sigma$ .

Nous considérons une loi *a priori* classique pour les matrices aléatoires : une loi inverse Wishart, que l'on centre sur une matrice diagonale, proportion de l'identité :

$$\Sigma | \alpha, \beta \sim \mathcal{IW}(\alpha(\beta - p - 1)I, \beta) , \alpha > 0 , \beta > p + 1 \text{ pour que l'espérance existe.} \quad (2)$$

Cette loi dépend d'hyperparamètres,  $\alpha$  et  $\beta$ , que l'on souhaite estimer conjointement avec  $\Sigma$ . Dans le but de proposer une méthode d'estimation non supervisée, nous considérons une loi *a priori* de référence pour  $\alpha$  et nous étudierons différentes lois *a priori* de référence pour  $\beta$  de la forme suivante :

$$\pi(\alpha) = \frac{1}{\alpha} \mathbb{I}_{]0;+\infty[}(\alpha) \text{ et } \pi(\beta) = \frac{1}{\beta^\delta} \mathbb{I}_{]p+1;+\infty[}(\beta). \quad (3)$$

Ces *a priori* sont en fait des mesures *a priori*, la loi *a posteriori* peut alors être impropre.

L'estimateur proposé  $\hat{\Sigma}$  minimisera le risque bayésien, risque fréquentiste moyenné sur les valeurs de  $\Sigma$  selon  $\pi(\Sigma)$ , basé sur la fonction de coût de Stein :

$$L(\Sigma, \hat{\Sigma}) = \text{tr}(\hat{\Sigma}\Sigma^{-1}) - \log \det(\hat{\Sigma}\Sigma^{-1}) - p.$$

### 3 Étude théorique du modèle et mise en pratique de la procédure d'estimation

La loi jointe *a posteriori* résultante de  $(\Sigma, \alpha, \beta)$  est définie en (4) :

$$\pi(\Sigma, \alpha, \beta/\mathbf{y}) \propto \frac{|\Sigma^{-1}|^{\frac{\beta+n+p+1}{2}} \exp(-\frac{1}{2}\text{tr}(\Sigma^{-1}(S + \alpha(\beta - p - 1)I))) \alpha^{\frac{p\beta}{2}-1} (\beta - p - 1)^{\frac{p\beta}{2}}}{2^{\frac{p(n+\beta)}{2}} \beta^\delta \Gamma_p(\frac{\beta}{2})}. \quad (4)$$

On peut montrer que cette loi est propre pour  $\delta$  supérieure ou égale à 2.

La loi jointe *a posteriori*  $\pi(\Sigma|\mathbf{y})$  n'étant pas standard, l'estimateur proposé ne s'obtient pas analytiquement. Il est approché par une méthode de Monte Carlo qui nécessite de générer des échantillons selon la loi  $\pi(\Sigma|\mathbf{y})$ .

Pour ce faire, nous proposons d'utiliser la loi jointe  $\pi(\Sigma, \alpha, \beta/\mathbf{y})$  et un algorithme de Gibbs pour générer des échantillons suivant cette loi. La loi  $\pi(\Sigma|\alpha, \beta, \mathbf{y})$  dite conditionnelle complète est une loi Inverse Wishart, celle de  $\pi(\alpha|\Sigma, \beta, \mathbf{y})$  est une loi Gamma. La loi conditionnelle complète de  $\beta$  n'est pas standard, un échantillon suivant cette loi est obtenu par un algorithme Métropolis-Hastings à marche aléatoire. Le plan d'échantillonnage général est décrit par l'algorithme 1.

### 4 Résultats attendus et perspectives

Dans un premier temps, l'estimateur bayésien est comparé en termes de risque (fonction de coût de Stein, Mean Squared Error), sur données simulées, avec celui du Maximum de

---

**Algorithme 1** Échantillonneur de Gibbs pour l'estimation de matrice de covariance

---

- 1: Initialisation avec  $k = 0$  et des valeurs arbitraires pour  $\alpha_0$  et  $\Sigma_0$
  - 2:  $k = k + 1$
  - 3: Tirage de  $\Sigma^{(k)} | \alpha^{(k-1)}, \beta^{(k-1)}, \mathbf{y} \sim \mathcal{IW}(\beta^{(k-1)} + n, S + \alpha(\beta^{(k-1)} - p - 1)I)$
  - 4: Tirage de  $\alpha^{(k)} | \Sigma^{(k)}, \beta^{(k-1)}, \mathbf{y} \sim \mathcal{G}(p\beta^{(k-1)}/2, (\beta^{(k-1)} - p - 1)\text{tr}(\Sigma^{-1(k)})/2)$
  - 5: Tirage  $\beta^{(k)} \sim \pi(\beta | \alpha^{(k)}, \Sigma^{(k)}, \mathbf{y})$
  - 6: Retour en 2 sauf si critère d'arrêt
- 

Vraisemblance, en particulier quand la dimension du modèle est importante par rapport au nombre de données.

Dans un deuxième temps, on appliquera cette méthode sur une problématique de gestion actif-passif, les comparaisons se feront sur l'étude de courbes d'efficience associées.

Enfin, on sortira du cadre du *shrinkage* et on rendra la loi *a priori* sur  $\Sigma$  informative, dans le sens où elle portera l'information issue des données historiques et des dires d'experts financiers.

## Références

- [1] M.J. Daniels. A class of shrinkage prior for the dependence structure in longitudinal data. *Journal of Statistical Planning and Inference*, 127(1-2) :119–130, 2005.
- [2] M.J. Daniels and R.E. Kass. Nonconjugate bayesian estimation of covariance matrices and its use in hierarchical models. *Journal of the American Statistical Association*, 94(448) :1254–1263, 1999.
- [3] D.K. Dey and C. Srinivasan. Estimation of a covariance matrix under stein's loss. *The Annals of Statistics*, 13(4) :1581–1591, 1985.
- [4] D. Pachamanova S.M. Focardi F.J. Fabozzi, P.N. Kolm. *Robust Portfolio Optimization and Management*. Wiley, 2007.
- [5] L.R. Haff. Estimation of the inverse covariance matrix : random mixtures of the inverse wishart matrix and the identity. *The Annals of Statistics*, 7(6) :1264–1276, 1979.
- [6] L.R. Haff. Empirical bayes estimation of the multivariate normal covariance matrix. *The Annals of Statistics*, 8(3) :586–597, 1980.

- [7] W. James and C. Stein. Estimation with quadratic loss. *Proceedings of the Fourth Berkeley Symposium on Mathematical and statistical Probabilities*, University of California Press, Berkeley :361–379, 1961.
- [8] O. Ledoit and M. Wolf. Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance*, 10 :603–621, 2002.
- [9] C.P. Robert. *Le choix bayésien. Principes et pratique*. Springer, 2005.
- [10] C.P. Robert and G. Casella. *Monte carlo statistical methods*. Springer, 2009.
- [11] C.P. Robert and J.M. Marin. *Bayesian core*. Springer, 2007.
- [12] M. Smith and R. Kohn. Parsimonious covariance matrix estimation for longitudinal data. *Journal of the American Statistical Association*, 97(460) :1140–1153, 2002.