École des Ponts ParisTech

Estimation de variance-covariance par méthode sparse

Par H. Andres, M. Bouazza, M. Karpe et C. Zeng

La détermination des structures de dépendance entre différents actifs ou facteurs de risques est au cœur des problèmes de modélisation financière multidimensionnels. Par exemple, lorsque l'on considère un portefeuille d'actifs, la théorie de Markowitz énonce qu'il est important que le portefeuille soit diversifié et donc que les actifs soient le moins corrélé possible.

Mathématiquement, la variable naturelle pour modéliser cette structure de dépendance est la matrice de covariance. En effet, dans le cadre de variables gaussiennes, la covariance est suffisante à décrire les structures de corrélation (ce qui n'est pas le cas pour d'autres distributions). De plus, bien que l'on s'intéresse à la covariance et que le critère étudié se base sur une modélisation gaussienne, l'information obtenue est en fait plus riche et peut s'appliquer à d'autres distributions.

Cependant, les données empiriques que nous exploitons sont affectées par du bruit qui biaise l'estimation et les méthodes classiques fournissent alors de piètres résultats.

Objectif et méthode de résolution

Le but est donc d'obtenir une matrice de covariance débarrassée du bruit, dite *sparse* en raison d'un grand nombre de coefficients nuls, à partir d'une matrice de covariance bruitée. Ce problème peut s'écrire comme un problème d'optimisation dans lequel on cherche à maximiser la log-vraisemblance de la solution en pénalisant le nombre de zéros dans la matrice de covariance inverse. On impose, de plus, des contraintes sur ses valeurs propres permettant d'une part de s'assurer que la matrice est positive, et d'autre part de borner davantage la solution, compte tenu d'informations qu'on aurait a priori sur le problème. Dans le cadre gaussien, ce dernier se formule ainsi :

$$\begin{aligned} &\text{maximize } f(X) := \log\det X - <\Sigma, X> - \rho\mathbf{Card}(X) \\ &\text{subject to } \alpha\mathbf{I}_n \leq X \leq \beta\mathbf{I}_n \\ &X \in \mathbf{S}_n; \Sigma \in \mathbf{S}_n^+; \rho, \alpha, \beta > 0 \end{aligned}$$

Ce problème est néanmoins qualifié de NP-difficile, ce qui signifie qu'on ne peut pas le résoudre informatiquement en temps raisonnable – cela est notamment dû à la non-convexité de la fonction f(X). Il est donc nécessaire de le transformer. Pour cela, on applique au problème initial des méthodes de relaxation convexe. Ces méthodes permettent de mettre le problème sous une forme pour laquelle il existe des algorithmes de résolution numérique. L'algorithme utilisé ici est l'algorithme de Nesterov (2005).

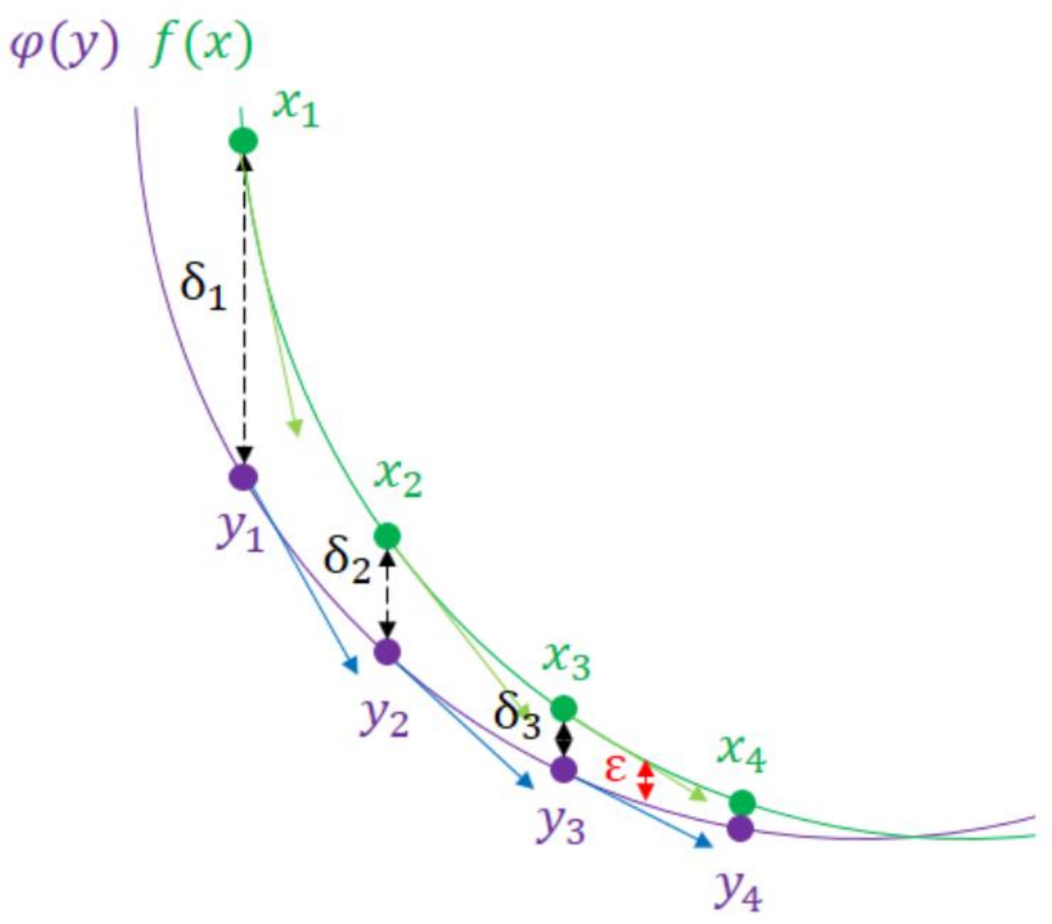


Fig. 1 : Illustration de l'algorithme de Nesterov avec f(x) fonction exacte, ϕ(y) fonction approchée.

L'algorithme de Nesterov s'inspire de méthodes d'optimisation récentes et efficaces pour déterminer l'extremum d'une fonction ϕ(y) qui est proche de la fonction f(X). En réduisant l'écart entre la solution exacte et la solution approchée par itérations successives (cf. fig. 1), on obtient un résultat très proche de la solution exacte.

Vérification de l'algorithme

Afin de s'assurer de la bonne implémentation de l'algorithme, nous l'avons testé sur une matrice bruitée très simple. Celle-ci a été construite de la façon suivante :

Etape 1 : On construit une matrice A dont les coefficients diagonaux sont égaux à 1 et dont quelques coefficients non diagonaux (c'est-à-dire un nombre négligeable par rapport à la taille de la matrice choisie), tirés aléatoirement selon une probabilité uniforme, sont égaux à 1 ou -1 avec équiprobabilité. Les coefficients non diagonaux tirés aléatoirement sont recopiés symétriquement par rapport à la diagonale de façon à obtenir une matrice diagonale.

Si la matrice alors obtenue n'est pas inversible, on répète la procédure. La matrice finalement créée correspond à l'inverse d'une matrice de covariance sans bruit mais avec quelques covariances non nulles, c'est-à-dire que seuls quelques variables sont corrélées.

Etape 2 : On construit une seconde matrice V dont les coefficients suivent une loi uniforme sur [-1;1] puis on la symétrise. Cette matrice correspond au bruit.

Etape 3 : On constitue une matrice de covariance B bruitée en sommant l'inverse de A avec σV où σ permet d'intensifier ou d'atténuer le bruit.

Nous avons appliqué l'algorithme de Nesterov sur cette matrice test avec $\varepsilon = 10^{-5}, \rho = 0.5, \alpha = 10^{-1}, \beta = 10$ et $\sigma = 0.15$.

Afin de présenter les résultats obtenus par l'algorithme de Nesterov pour la matrice test, nous représentons les matrices en entrée et en sortie de l'algorithme par des tableaux de pixels où chaque pixel a une couleur qui dépend de la valeur du coefficient auquel il est associé.

Les résultats obtenus sont très satisfaisants. En effet, la matrice obtenue par l'algorithme ressemble très fortement à la matrice sans bruit initiale.

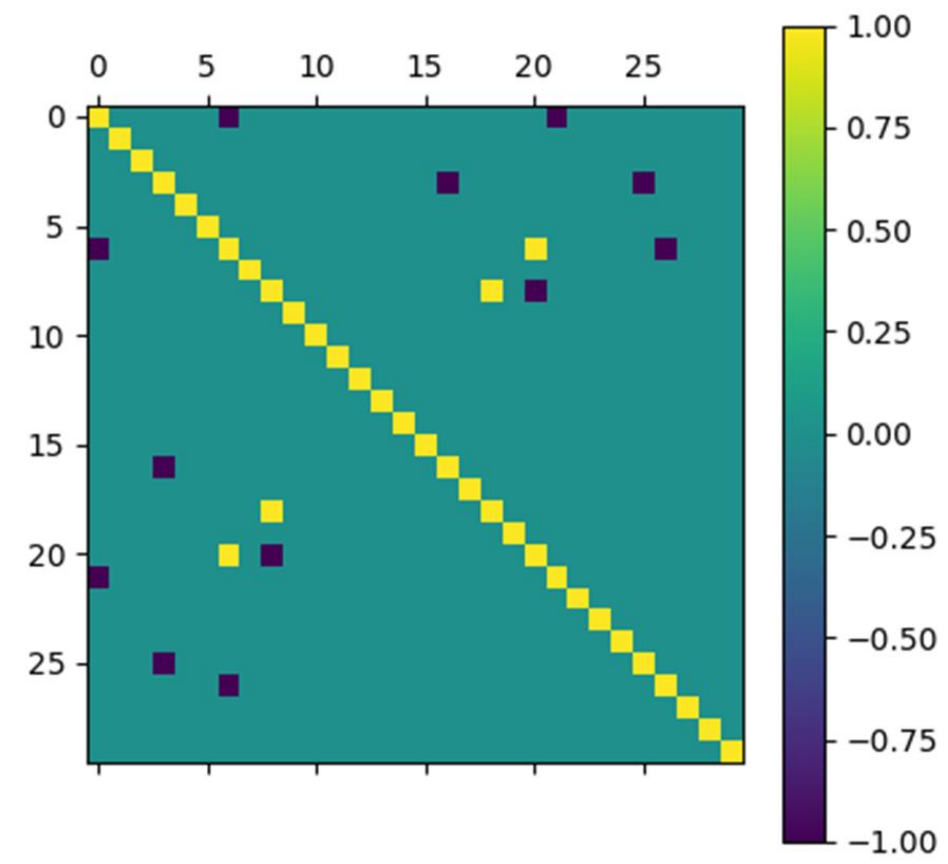


Fig. 2 : Matrice A de l'étape 1.

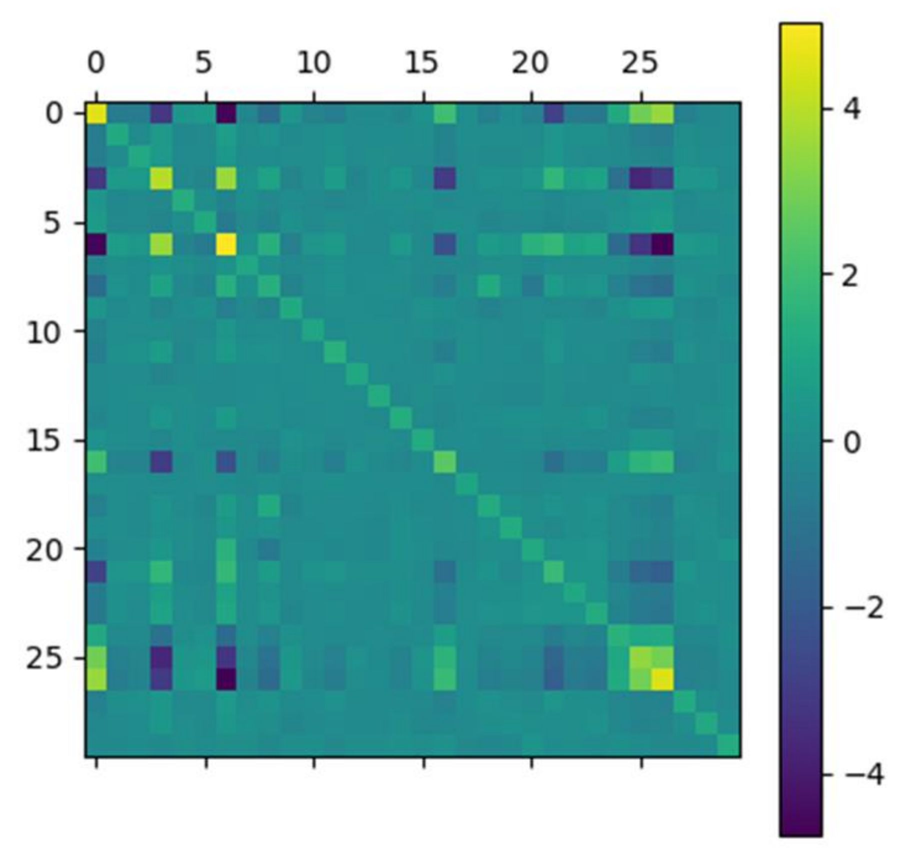


Fig. 3 : Matrice B de l'étape 3.

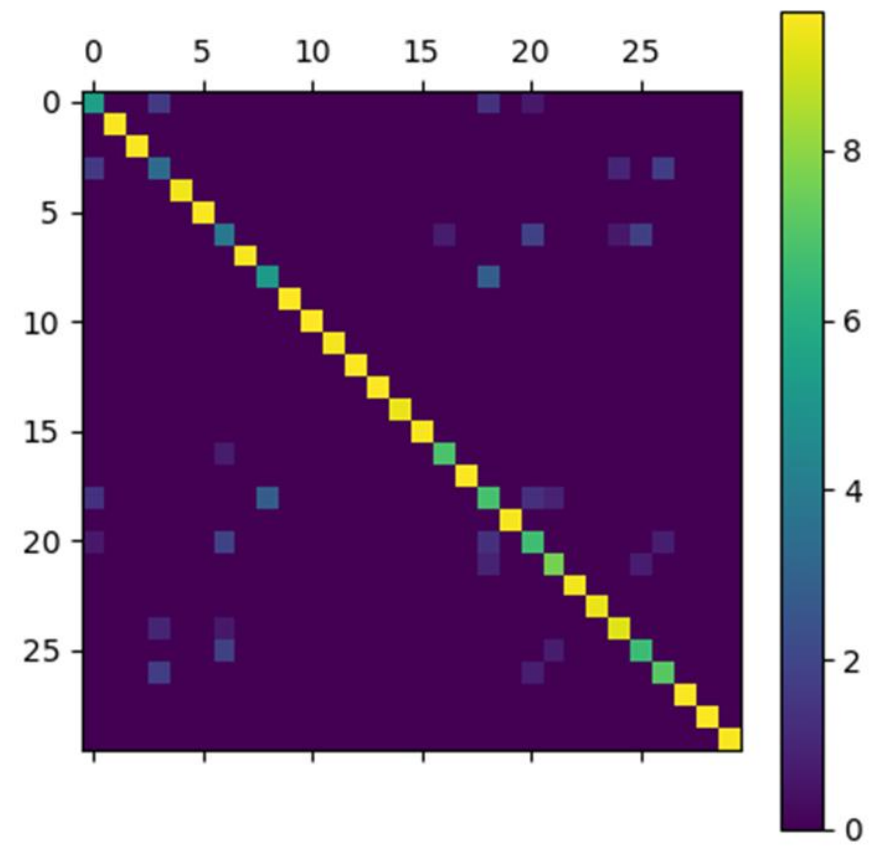


Fig. 4 : Matrice obtenue en appliquant l'algorithme de Nesterov à B.

Application à l'analyse de taux d'intérêts

Nous avons ensuite appliqué notre algorithme à une matrice de covariance obtenue empiriquement à partir de données de variations de taux d'intérêts.

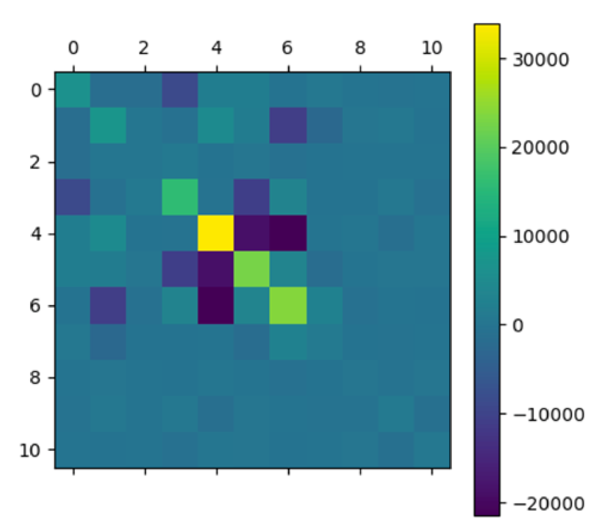


Fig. 5a : Matrice de covariance empirique bruitée.

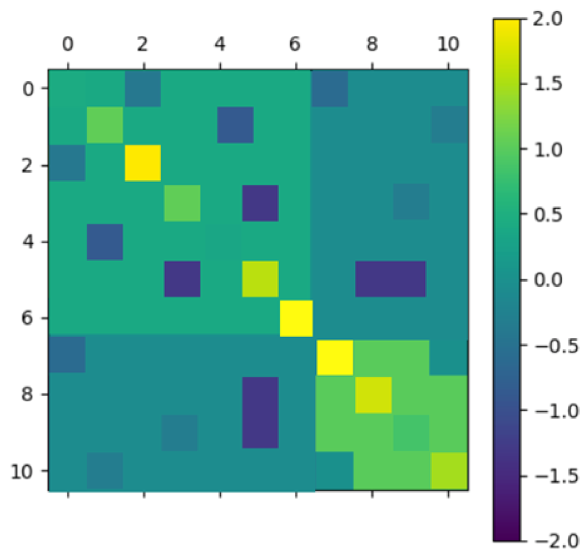
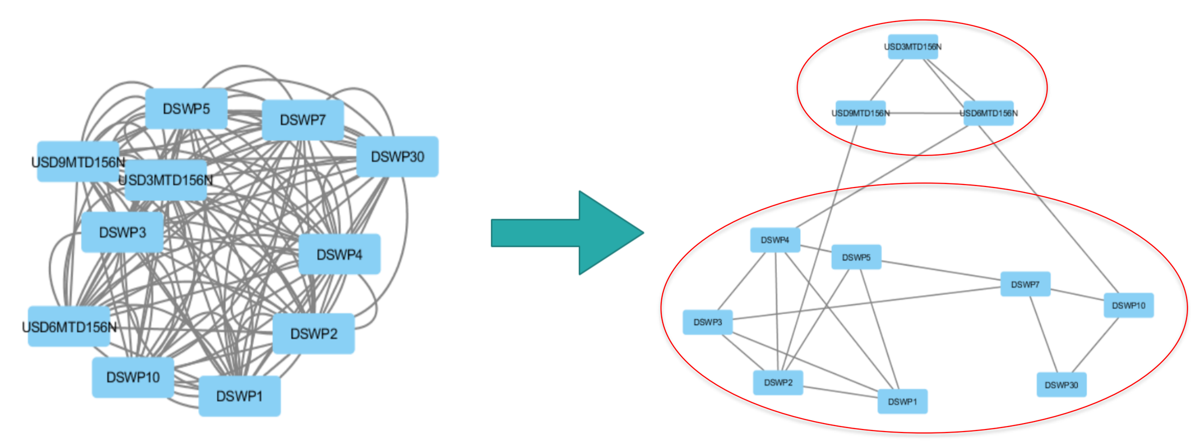


Fig. 5b : Matrice de covariance sélectionnée

Après avoir appliqué l'algorithme, la matrice sélectionnée présente une structure globalement diagonale par blocs. Pour interpréter cette structure, on utilise une représentation graphique où les nœuds sont les différents actifs et on trace une arête entre deux nœuds s'ils sont corrélés (i.e. de covariance non nulle).



Ainsi, nous obtenons deux clusters qui montrent que les swaps sont corrélés en fonction de leur maturité. On trouve dans le cluster du haut, des maturités de 3, 6 et 9 mois et dans celui du bas, des maturités de 1 à 30 ans. Ces différentes maturités correspondent à de différents marchés et de différents besoins de financement.

Analyse de la méthode présentée

L'estimation de variance-covariance par méthode sparse est un outil puissant et générique pour gérer les contraintes et garder une certaine tractabilité au niveau des données. Cette méthode aide à l'extraction d'informations fines à partir de données bruitées, lesquelles données n'auraient pas pu être analysées correctement sans avoir enlevé la contribution du bruit.

Bibliographie :
B. A. d'Aspremont, O. Banerjee, L. El Ghaoui, *First-Order Methods for Sparse Covariance Selection*