Introduction à la modélisation multivariée des actifs et la théorie des matrices aléatoires

10 janvier 2012

Table des matières

1 Introduction
- Théorie du portefeuille : Markowitz
- Quelques lois multivariées

2 Asymptotique des grandes matrices aléatoires
- Loi du semi-cercle
- Loi de Marchenko-Pastur

3 Preuves
- Lemmes préliminaires
- Loi du semi-cercle
- Loi de Marchenko-Pastur

4 Références

Les stocks sont-ils corrélés ?

On note $(r_1, \cdots, r_N)$ la loi des rendements des stocks du SP500.
On note I la loi du rendement de l'indice SP500.
Ordres de grandeur : $\sigma_I^2 = \sqrt{E[I^2]} \sim 10^{-2}$ et pour tout i $\sigma^2 = \sqrt{E[r_i^2]} \sim 2.10^{-2}$.
On introduit la corrélation moyenne inter-stock :

$$\overline{\rho} = \frac{1}{N(N-1)} \sum_{i \neq j} E[r_i r_j] / \sigma^2$$

On en déduit l'ordre de grandeur de $\overline{\rho}$: $\overline{\rho} \approx \sigma_I^2 / \sigma^2 \approx 0.25$ d'où la nécessité de faire des modèles multivariés (Théorie du portefeuille, estimation du risque, etc...)

Théorie du portefeuille : Markowitz

On suppose que l'on veut investir au temps t dans N stocks de rendements $(r_1(t), \cdots, r_N(t))$ de moyenne $(\mu_1, \cdots, \mu_N)$. On considère les quantités suivantes (pour une allocation $(w_1, \cdots, w_N)$) :

- Gain moyen : $\mathcal{G} = \sum_{i=1}^{N} w_i \mu_i$.
- Risque (mesuré par la variance) : $\mathcal{R}^2 = \sum_{i,j=1}^{N} w_i w_j C_{i,j}$,

où $(C_{i,j})_{1 \leq i,j \leq N}$ est la matrice des covariances.

Théorie du portefeuille : Markowitz

Programme de Markowitz : minimiser $\mathcal{R}$ avec une contrainte sur le gain $\mathcal{G}$ ou bien maximiser $\mathcal{G}$ avec une contrainte de risque sur $\mathcal{R}$. En particulier, cela nécessite une bonne connaissance de la matrice $(C_{i,j})_{1\leq i,j\leq N}$. En diagonalisant la matrice C, on obtient les modes marché $(e^\alpha)_{1\leq\alpha\leq N}$ (associés à $\lambda^1 \geq \cdots \geq \lambda^N$) :

$$e^\alpha = \frac{1}{\sqrt{\lambda^\alpha}} \sum_{j=1}^N e_j^\alpha r_j, \quad r_j = \sum_{\alpha=1}^N \sqrt{\lambda^\alpha} e_j^\alpha e^\alpha$$

Dans la décomposition ci-dessus, les variables e^α sont décorrélées : $E[e^\alpha e^\beta] = \delta_{\alpha=\beta}$. En pratique, on trouve :

$$e^1 = \frac{1}{\sqrt{\lambda^1}} \sum_{j=1}^N r_j, \quad \lambda^1 = \frac{1}{N} \sum_{i,j=1}^N C_{i,j}$$

Théorie du portefeuille : Markowitz

On introduit donc la matrice de covariance empirique $(\overline{C}_{i,j})_{1 \leq i,j \leq N}$:

$$\overline{C}_{i,j} = \frac{1}{T} \sum_{i=1}^{T} r_i(t) r_j(t)$$

Si $N << T$, le régime asymptotique est atteint ; malheureusement, en finance, N est comparable à T, ce qui amène à considérer l'asymptotique $N \to \infty$ et $T/N \underset{N \to \infty}{\to} q$ (théorie des grandes matrices aléatoires).

Quelques exemples de loi multivariée : loi Gaussienne

Densité de la loi Gaussienne :

$$f(r_1, \cdots, r_N) = \frac{1}{\sqrt{(2\pi)^N \det C}} e^{-\frac{1}{2}\sum_{i,j=1}^{N} r_i (C^{-1})_{i,j} r_j},$$

où $C = (C_{i,j})_{1 \leq i,j \leq N}$ est la matrice des covariances.

Quelques exemples de loi multivariée : loi de Student multivariée

Densité de la loi de Student multivariée :

$$f(r_1, \cdots, r_N) = \frac{\Gamma(\frac{N+\mu}{2})}{\Gamma(\frac{\mu}{2})\sqrt{(\mu\pi)^N \det(\tilde{C})}} \frac{1}{(1 + \frac{1}{\mu}\sum_{i,j=1}^{N} r_i(\tilde{C}^{-1})_{i,j} r_j)^{\frac{N+\mu}{2}}},$$

où $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$.

Quelques exemples de loi multivariée : loi de Student multivariée

On peut écrire la loi de Student multivariée comme $\sigma(r_1, \cdots, r_N)$ où :

- σ est une volatilité commune de loi inverse gamma : $\sqrt{\frac{\mu}{2s}}$ où s a pour densité $f(s) = \frac{1}{\Gamma(\mu/2)} s^{\mu/2-1} e^{-s}$.
- $(r_1, \cdots, r_N)$ est un vecteur Gaussien de covariance $\tilde{C}$

Loi du semi-cercle

Soit $(X_{i,j})_{1\leq i\leq j\leq N}$ une suite i.i.d. de variables aléatoires centrées de variance σ^2. Si $j < i$, on note $X_{i,j} = X_{j,i}$. On considère la matrice symétrique $X^{(N)} = (X_{i,j}/\sqrt{N})_{1\leq i,j\leq N}$ et on note $(\lambda_i^{(N)})_{1\leq i\leq N}$ le spectre de $X^{(N)}$.

Theorem (Wigner, loi du semi-cercle)

On alors la convergence presque sûre suivante de la mesure empirique :

$$\frac{1}{N}\sum_{i=1}^{N}\delta_{\lambda_i^{(N)}} \underset{N\to\infty}{\Rightarrow} \frac{1}{2\pi\sigma^2}\sqrt{4\sigma^2-\lambda^2}1_{|\lambda|\leq 2\sigma}d\lambda$$

Loi de Marchenko-Pastur

Soit $(X_{i,j})_{1\leq i\leq N, 1\leq j\leq T}$ une suite i.i.d. de variables aléatoires centrées de variance σ^2. On considère la matrice symétrique $X^{(N)}$ avec $X^{(N)} = (X_{i,j}/\sqrt{N})_{1\leq i\leq N, 1\leq j\leq T}$ et on note $(\lambda_i^{(N)})_{1\leq i\leq N}$ le spectre de $X^{(N)}$.

Theorem (Marchenko-Pastur)

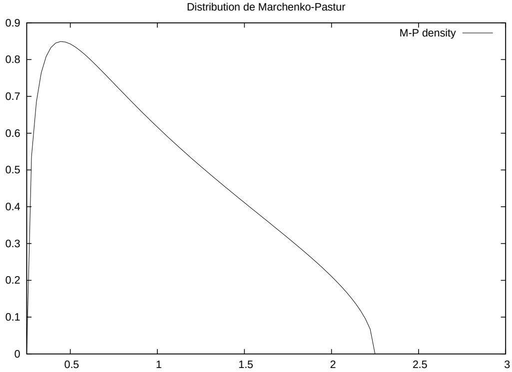
On alors la convergence presque sûre suivante de la mesure empirique lorsque $N\to\infty$ et $T/N \underset{N\to\infty}{\to} q \geq 1$:

$$\frac{1}{N}\sum_{i=1}^{N}\delta_{\lambda_i^{(N)}} \underset{N\to\infty}{\Rightarrow} \frac{q}{2\pi\sigma^2\lambda}\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}1_{\lambda\in[\lambda_-,\lambda_+]}d\lambda$$

où $\lambda_- = \sigma^2(1-\sqrt{1/q})^2$ et $\lambda_+ = \sigma^2(1+\sqrt{1/q})^2$.

Loi de Marchenko-Pastur : graphique

On choisit $\sigma^2 = 1$, $N = 500$ (Stocks du SP500) et $T = 2000$ (8 ans de données).



Preuve : Schur complement formula

Soit $A^{(N+1)} = (A_{i,j})_{0\leq i,j\leq N}$ une matrice de taille $N+1$. On note $(A^{(N+1)})^{-1}_{0,0}$ l'élément $(0,0)$ de $(A^{(N+1)})^{-1}$ et $A^{(N)}$ la matrice de taille N : $(A_{i,j})_{1\leq i,j\leq N}$.

Lemme (Schur complement formula)

On a alors :

$$(A^{(N+1)})^{-1}_{0,0} = \frac{1}{A_{0,0} - \sum_{i,j=1}^{N} A_{i,0}A_{0,j}(A^{(N)})^{-1}_{j,i}}$$

Preuve : Résolvante d'une mesure

Si $\mu(d\lambda)$ est une mesure de probabilités, on introduit sa résolvante $L(z)$ pour $z \in \mathbb{C} \setminus \mathbb{R}$:

Définition

$$L(z) = \int_{\mathbb{R}} \frac{\mu(d\lambda)}{z - \lambda}$$

Preuve : Résolvante d'une mesure

Si $\mu(d\lambda) = f(\lambda)d\lambda$ a une densité, on a :

Lemme

$$f(a) = -\lim_{\epsilon \to 0^+} \frac{1}{\pi} Im(L(a + i\epsilon))$$

Preuve : Formule du semi-cercle

Soit $z \in \mathbb{C} \setminus \mathbb{R}$. On procède par récurrence sur N. On considère donc la matrice symétrique $X^{(N+1)} = (X_{i,j}/\sqrt{N+1})_{0 \leq i \leq j \leq N}$. On applique la Schur complement formula avec :

$$A^{(N+1)}(z) = zI_{N+1} - X^{(N+1)}.$$

On obtient donc :

$$(A^{(N+1)}(z))^{-1}_{0,0} = \frac{1}{z - \frac{X_{0,0}}{\sqrt{N+1}} - \frac{1}{N+1}\sum_{i,j=1}^{N} X_{i,0}X_{0,j}(A^{(N)}(z))^{-1}_{i,j}}$$

Preuve : Formule du semi-cercle

On admet que :

$$\frac{1}{N}\sum_{i=1}^{N}\delta_{\lambda_i^{(N)}} \underset{N\to\infty}{\sim} \frac{1}{N}E[\sum_{i=1}^{N}\delta_{\lambda_i^{(N)}}]$$
$$\underset{N\to\infty}{\Rightarrow} f(\lambda)d\lambda$$

On en déduit que :

$$\frac{1}{N+1}\sum_{i=0}^{N}\frac{1}{z-\lambda_i^{(N+1)}} = \frac{1}{N+1}\sum_{i=0}^{N}(A^{(N+1)}(z))^{-1}_{i,i}$$
$$\underset{N\to\infty}{\sim} \frac{1}{N+1}E[\sum_{i=0}^{N}(A^{(N+1)}(z))^{-1}_{i,i}]$$
$$= E[(A^{(N+1)}(z))^{-1}_{0,0}]$$

Preuve : Formule du semi-cercle

On a (en notant $(\lambda_i^{(N)})_{1\leq i\leq N}$ les valeurs propres réelles de $X^{(N)}$) :

$$
\begin{aligned}
&E[|\frac{1}{N+1}\sum_{i\neq j=1}^{N} X_{i,0}X_{0,j}(A^{(N)}(z))^{-1}_{i,j}|^2] \\
&= \frac{2}{(N+1)^2}\sum_{i\neq j=1}^{N} E[X_{0,0}^2]^2 E[|(A^{(N)}(z))^{-1}_{i,j}|^2] \\
&\leq 2\frac{E[X_{0,0}^2]^2}{(N+1)^2}\mathrm{tr}((zI_N - X^{(N)})^{-1}(\bar{z}I_N - X^{(N)})^{-1}) \\
&= 2\frac{E[X_{0,0}^2]^2}{(N+1)^2}\sum_{i=1}^{N}\frac{1}{|z-\lambda_i^{(N)}|^2} \\
&\leq 2\frac{E[X_{0,0}^2]^2}{(N+1)|Im(z)|^2}
\end{aligned}
$$

Preuve : Formule du semi-cercle

Si l'on note $L(z) = \lim_{N\to\infty} \frac{1}{N+1} \sum_{i=0}^{N} \frac{1}{z-\lambda_i^{(N+1)}}$, on obtient :

$$
\begin{aligned}
L(z) &= \lim_{N\to\infty} \frac{1}{N+1} \sum_{i=0}^{N} \frac{1}{z - \lambda_i^{(N+1)}} \\
&\underset{N\to\infty}{\sim} E[\frac{1}{z - \frac{X_{0,0}}{\sqrt{N+1}} - \frac{1}{N+1}\sum_{i,j=1}^{N} X_{i,0}X_{0,j}(A^{(N)}(z))^{-1}_{i,j}}] \\
&\underset{N\to\infty}{\sim} E[\frac{1}{z - \frac{1}{N+1}\sum_{i=1}^{N} X_{i,0}^2(A^{(N)}(z))^{-1}_{i,i}}] \\
&\underset{N\to\infty}{\sim} E[\frac{1}{z - \frac{\sigma^2}{N+1}\sum_{i=1}^{N}(A^{(N)}(z))^{-1}_{i,i}}] \\
&= \frac{1}{z - \sigma^2 L(z)}.
\end{aligned}
$$

Preuve : Formule du semi-cercle

On obtient donc pour $z = a + ib$ avec $a, b > 0$:

$$L(z) = \frac{z - \sqrt{z^2 - 4\sigma^4}}{2\sigma^2}.$$

On obtient facilement que $-\frac{1}{\pi} Im(L(a + i\epsilon)) \underset{\epsilon \to 0^+}{\to} \frac{\sqrt{4\sigma^2 - a^2}}{2\pi\sigma^2} 1_{|a| \leq 2\sigma}$

Preuve : Formule de Marchenko-Pastur

On obtient la loi de Marchenko-Pastur par une méthode similaire. On commence par le lemme suivant pour X une matrice $N \times T$ $(N \leq T)$:

Lemme

$$det \begin{pmatrix} \lambda I_T & -{}^tX \\ -X & \lambda I_N \end{pmatrix} = det(I - X{}^tX)^2 \lambda^{T-N}$$

Pour étudier la mesure spectrale de $X^{(N)}{}^tX^{(N)}$, il suffit donc d'étudier celle de $\begin{pmatrix} \lambda I_T & -{}^tX^{(N)} \\ -X^{(N)} & \lambda I_N \end{pmatrix}$.

Preuve du lemme

Il suffit de constater que :

$$\begin{pmatrix} \lambda I_T & -{}^tX \\ -X & \lambda I_N \end{pmatrix} \begin{pmatrix} \lambda I_T & 0 \\ X & \lambda I_N \end{pmatrix} = \begin{pmatrix} \lambda^2 I_T - \lambda {}^tXX & -\lambda {}^tX \\ 0 & \lambda^2 I_N \end{pmatrix}$$

On conclut en prenant les determinants.

Preuve : Formule de Marchenko-Pastur

Soit $z \in \mathbb{C} \setminus \mathbb{R}$. On procède par récurrence sur N. On considère donc la matrice $X^{(N+1)} = (X_{i,j}/\sqrt{N+1})_{0 \leq i \leq N, 1 \leq j \leq T}$. On applique la Schur complement formula avec :

$$A^{(N+1)}(z) = \begin{pmatrix} zI_T & -{}^tX^{(N+1)} \\ -X^{(N+1)} & zI_{N+1} \end{pmatrix}$$

Preuve : Formule de Marchenko-Pastur

On obtient donc :

$$(A^{(N+1)}(z))^{-1}_{0,0} = \frac{1}{z - \frac{1}{N+1}\sum_{i,j=1}^{N} X_{i,1}X_{1,j}(A^{(N,1)}(z))^{-1}_{T+i,T+j}}$$

où $A^{(N,1)}(z)$ est une matrice de même loi que
$A^{(N)}(z) = \begin{pmatrix} zI_T & -{}^tX^{(N)} \\ -X^{(N)} & zI_N \end{pmatrix}$

Preuve : Formule de Marchenko-Pastur

On obtient aussi la formule suivante :

$$
\begin{aligned}
&(A^{(N+1)}(z))^{-1}_{N+T+1,N+T+1} \\
&= \frac{1}{z - \frac{1}{N+1}\sum_{i,j=1}^{T} X_{N,i}X_{N,j}(A^{(N)}(z))^{-1}_{i,j}}
\end{aligned}
$$

où $A^{(N,2)}(z)$ est une matrice de même loi que $A^{(N)}(z)$.

Preuve : Formule de Marchenko-Pastur

On admet que :

$$\frac{1}{T}\sum_{i=1}^{T}(A^{(N)}(z))^{-1}_{i,i} \underset{N\to\infty}{\sim} \frac{1}{T}E[\sum_{i=1}^{T}(A^{(N)}(z))^{-1}_{i,i}]$$
$$\underset{N\to\infty}{\to} L_1(z)$$

et de même que $\frac{1}{N}\sum_{i=T+1}^{T+N}(A^{(N)}(z))^{-1}_{i,i} \underset{N\to\infty}{\to} L_2(z)$

Preuve : Formule de Marchenko-Pastur

En procédant comme pour la loi du semi-cercle, on en déduit les équations suivantes pour L_1, L_2 :

$$L_1(z) = \frac{1}{z - \sigma^2 L_2(z)}$$

et

$$L_2(z) = \frac{1}{z - q\sigma^2 L_1(z)}$$

On conclut en résolvant les équations, etc...

Références

Livres et articles généraux sur les matrices aléatoires et ses applications en finance :

- Bouchaud, Potters : *Theory of Financial Risk and Derivative Pricing*, Cambridge University Press, Cambridge (2003).
- Bouchaud, Potters : Financial applications of Random Matrix Theory, à paraitre dans *Handbook on Random Matrix Theory*, disponible à http ://arxiv.org/abs/0910.1205.
- Anderson, Guionnet, Zeitouni : *An Introduction to Random Matrices*, Cambridge University Press (2010).