Question 1) State the Markowitz optimization problem and solve it.

- 2 dates : $t=0$ et $t=1$
- 1 ASR : $1 \to 1+r$
- N actifs risqués : $P_{i,0} \to P_{i,1}$ $i=1,..,N$
- Rendement de l'actif i entre $t=0$ et $t=1$: $r_i = \frac{P_{i,1}}{P_{i,0}} - 1 := y_i - 1$ en posant $y_i = \frac{P_{i,1}}{P_{i,0}}$ et $Y = (y_1,..,y_N)'$

$$\boxed{\begin{array}{l} E[Y] := \mu \\ Var(Y) := \Omega \end{array}}$$

Portefeuille contenant une quantité a_0 en ASR et $a = (a_1,..,a_N)'$ en actifs risqués

$$\Rightarrow \begin{cases} V_0 = a_0 + \sum_{i=1}^{N} a_i P_{i,0} \\ V_1 = a_0(1+r) + \sum_{i=1}^{N} a_i P_{i,1} = a_0(1+r) + \sum_{i=1}^{N} a_i y_i P_{i,0} = a_0(1+r) + a' \operatorname{diag}[P_0] Y \end{cases}$$

$$\boxed{\begin{array}{l} E[V_1(a_0,a)] = a_0(1+r) + a' \operatorname{diag}[P_0]\mu \\ Var(V_1(a_0,a)) = a' \operatorname{diag}[P_0]\, \Omega \operatorname{diag}[P_0] a \end{array}}$$

On pose $\boxed{w_a = \operatorname{diag}[P_0] a = \text{Composition en € du ptf}}$

$$\Rightarrow \boxed{Var(V_1(a_0,a)) = w_a' \Omega w_a}$$

on peut montrer que la contrainte doit être saturée $\Rightarrow$ "="

Markowitz Problem : $\max_{(a_0,a)} a_0(1+r) + a' \operatorname{diag}[P_0]\mu$ sc $\begin{cases} a' \operatorname{diag}[P_0]\, \Omega \operatorname{diag}[P_0] a \leq \sigma^2 \\ a_0 + a' \operatorname{diag}[P_0] e = v \end{cases}$

$$\Leftrightarrow \max_{w_a} (v - w_a' e)(1+r) + w_a' \mu \quad sc \quad w_a' \Omega w_a = \sigma^2$$

$$\Leftrightarrow \max_{w_a} w_a' (\mu - (1+r)e) \quad sc \quad w_a' \Omega w_a = \sigma^2$$

$$\Leftrightarrow \boxed{\max_{w_a} w_a' \tilde{\mu} \quad sc \quad w_a' \Omega w_a = \sigma^2} \quad \text{où} \quad \boxed{\tilde{\mu} = \mu - (1+r)e}$$

Resolution

$$\mathcal{L}(w_a, \lambda) = w_a' \tilde{\mu} - \frac{\lambda}{2}(w_a' \Omega w_a - \sigma^2)$$

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial w_a} = \tilde{\mu} - \lambda \Omega w_a = 0 \\ \frac{\partial \mathcal{L}}{\partial \lambda} = -\frac{1}{2}(w_a' \Omega w_a - \sigma^2) = 0 \end{cases} \Leftrightarrow \begin{cases} w_a = \frac{1}{\lambda} \Omega^{-1} \tilde{\mu} & (1) \\ \sigma^2 = w_a' \Omega w_a & (2) \end{cases}$$

$$(1) \to (2) \Rightarrow \sigma^2 = \frac{1}{\lambda^2} \tilde{\mu}' \Omega^{-1} \Omega \Omega^{-1} \tilde{\mu} \Leftrightarrow \sigma^2 = \frac{1}{\lambda^2} \tilde{\mu}' \Omega^{-1} \tilde{\mu} \Leftrightarrow \boxed{\lambda = \frac{1}{\sigma}(\tilde{\mu}' \Omega^{-1} \tilde{\mu})^{1/2}}$$ ~ Aversion par le risque

$$(1) \Leftrightarrow \operatorname{diag}[P_0] a = \frac{1}{\lambda} \Omega^{-1} \tilde{\mu} \Leftrightarrow \boxed{a = \frac{1}{\lambda} \operatorname{diag}[P_0]^{-1} \Omega^{-1} \tilde{\mu}}$$

$$a_0 = v - a' \operatorname{diag}[P_0] e \Leftrightarrow \boxed{a_0 = v - \frac{1}{\lambda} \tilde{\mu}' \Omega^{-1} e}$$

Question 2) Give the CAPM equations for expectation and variance of returns, providing accurate definitions of all quantities. What does the CAPM tell us about the relationship between risk and returns in finance? How can you prove the validity of CAPM in practice? (Give main mathematical elements without entering in details)

⓪ Notations

- $Z = Y - (1+r)e$ = Vecteur des rendements net des actifs risqués
- $\tilde{\mu} = \mu - (1+r)e$
- $Z(a)$ = Vecteur de rentabilité net de la partie risquée d'un ptf de composition a en AR

$$\hookrightarrow Z(a) = \frac{a'P_1}{a'P_0} - (1+r) = \frac{a'\operatorname{diag}[P_0](Z+(1+r)e)}{a'\operatorname{diag}[P_0]e} - (1+r) = \frac{a'\operatorname{diag}[P_0]Z}{a'\operatorname{diag}[P_0]e} = \frac{w_0'Z}{w_0'e}$$

- Portefeuille de marché = Ptf contenant tous les AR du marché pondéré par leur capitalisation

① Equations du CAPM

$$Z = \beta Z(a^*_m) + \varepsilon$$

- a^*_m = Ptf de marché
- β = Vecteur des Bêtas i.e $\dfrac{Cov(Z, Z(a^*_m))}{Var(Z(a^*_m))}$
- $\mathbb{E}[\varepsilon] = Cov(\varepsilon, Z(a^*_m)) = 0$

$$\Rightarrow \boxed{\begin{aligned} &\mathbb{E}[Z_i] = \beta_i \mathbb{E}[Z(a^*_m)] \\ &Var(Z_i) = \beta_i^2 Var(Z(a^*_m)) + \sigma_i^2 \end{aligned}}$$

② Commentaires

L'espérance du rendement d'un titre que de son beta et de l'espérance de rendement du ptf de marché

$\Rightarrow$ l'idée selon laquelle plus il y a de risque, plus il y a d'espérance de rendement est fausse.

$\Rightarrow$ Le rendement d'un actif ne dépends que du risque systémique

$$Var(Z_i) = \beta_i^2 Var(Z(a^*_m)) + \sigma_i^2$$

Risque systémique = dépendant de l'économique Non diversifiable et rémunérateur

Risque idiosyncratique = propre à l'actif i et décorrélé des rendements de l'économie
- Diversifiable et non rémunéré

③ Tester le CAPM

Pour chaque AR on est dans le cadre de la regression linéaire (MCO) : $Z_{i,t} = \alpha_i + \beta_{i,t} Z_{i,t}(a^*_m) + \varepsilon_{i,t}$ donc on peut tester la nullité de α_i.

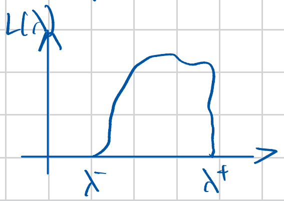
Question 3) State the Marcenko-Pastur theorem (no need to give the density of the Marcenko-Pastur law, just call it L). How can this result be used in finance ?

Théorème de marcenko-Pastur :

Notons :

- Ω_N = Matrice de variance-covariance des rendements
- $E = \frac{1}{N} \sum_{t=1}^{N} y_t y_t'$ la corrélation empirique

Supposons $N, T \to +\infty$ tq $\frac{N}{T} = q < 1$. Supposons $\Omega_N = I_N$. Alors p.s on a la convergence de la mesure empirique associée aux vp de E vers la loi de Marcenko Pastur qui a une densité $L(\lambda)$



Utile en finance pour estimer E : $E = \sum_{k=1}^{N} \tilde{\lambda}_k \tilde{v}^k \tilde{v}^{k'}$ $\quad \tilde{\lambda}_1 \geq \dots \geq \tilde{\lambda}_N$.

On ne fait pas confiances aux valeurs propres trop petites $\Rightarrow$ on va poser $\begin{cases} \tilde{\lambda}^k_{clean} = a & \text{pour } k > k^* \\ \tilde{\lambda}^k_{clean} = \tilde{\lambda}_k & \text{pour } k < k^* \end{cases}$

Le théorème de Marcenko-Pastur permet d'avoir une idée de k^*. En effet k^* sera le nombre de vp $> \lambda^+$.

Question 4) Define the Ridge estimator and provide its closed-form formula with proof. What is the interest of the Ridge estimator compared to ordinary least squares? Define the Lasso estimator. What is the interest of Lasso estimator compared to Ridge? How do we choose the regularization parameter (λ) of the Ridge and Lasso estimators? Give one example of a situation where Ridge or Lasso estimators are useful in finance.

① Ridge estimator

$$\boxed{\hat{\beta} \in \underset{\beta}{\operatorname{argmin}} \|Y - X\beta\|^2 + \lambda \|\beta\|^2}$$

CPO : Notons $F(\beta) := \|Y - X\beta\|^2 + \lambda \|\beta\|^2$

$$\nabla F(\beta) = -2X'(Y - X\beta) + 2\lambda\beta = 0 \iff \boxed{\hat{\beta}^{Ridge} = (X'X + \lambda I_p) X'Y}$$

Intérêt de ridge :

- Si $n \leq p$ (Nombre de variable explicative) $\Rightarrow (X'X)$ non inversible $\Rightarrow$ MCO ne marche plus $\Rightarrow$ Ridge permet de corriger ce pb grâce au terme $+\lambda I_p$
- Permet de contrôler les β_j pour qu'ils ne soient pas trop grand
- L'estimateur ridge (même s'il est biaisé) peut avoir un risque quadratique plus faible que les MCO.

② Lasso estimator :

$$\boxed{\hat{\beta}^{Lasso} = \underset{\beta}{\operatorname{argmin}} \|Y - X\beta\|^2 + \|\beta\|_1}$$

- Permet d'avoir des $\hat{\beta}$ sparses i.e avec beaucoup de 0 pour faire de la selection de variable alors que ridge shrink les coefficients mais sans les mettre à 0.

③ Selection de λ : On fait de la cross-validation pour avoir une bonne valeur de λ.

Question 5) In the PCA of a cloud of data points in $\mathbb{R}^p$ given by the $n \times p$ array X, how does one build the best sub-vector space of $\mathbb{R}^p$ with dimension k to project the array ?

On considère la matrice de variance-covariance entre les variables : $V = \frac{1}{n} X'X$

On appelle inertie portée par un SEV : $I_F = \frac{1}{n} \sum_{i=1}^{n} d^2(\vec{0}, P_F x_i)$

ACP : Pour k fixé, on cherche le SEV F_k d'inertie maximale. Connaissant F_k, on peut montrer que F_{k+1} est de la forme $F_{k+1} = F_k \oplus^{\perp} v$ ⇒ Pour trouver F_k, on les cherche les uns après les autres

On peut montrer que le sev F_k d'inertie max est $Vect(\vec{v}_1, .., \vec{v}_k)$ et l'inertie associée est $\sum_{j=1}^{k} \lambda_j$
où $\lambda_1 \geq \lambda_2 \geq .. \geq \lambda_n$ sont les vp associées aux $\vec{vp}$ $\vec{v}_1, .., \vec{v}_n$ de la matrice V.

Question 6) Consider you have access to interest rates curves between 2000 and 2023, one curve every month, with 10 maturity points for each curve, from 1 year to 10 years. Describe how you would analyze the results of a PCA applied to this dataframe and what you expect to see.

- 23 ans de données ⇒ $N = 23 \times 12 = 276$ mois de données
- Chaque courbe a 10 points ⇒ 1 courbe = un vecteur de $\mathbb{R}^{10}$

On pose $X \in \mathcal{M}_{N,10}$ la matrice de nos données
→ Une ligne = 1 mois
→ une colonne = une des 10 maturités ⇒ 10 vp

On commence par centrer et réduire nos données

Pour analyser la PCA, on trace le graphe des vp pour déterminer combien d'axes on peut garder
Pour des données de taux, on s'attend à conserver les 3 premiers axes (PC1, PC2 et PC3)
Typiquement, on s'attend à ce que :

- PC1 = Parallel shift ie est-ce que toute la courbe se déplace vers le haut ou vers le bas
- PC2 = Pente de la courbe
- PC3 : Courbature de la courbe