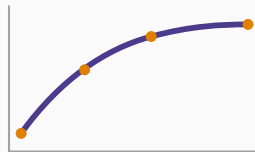


Économétrie financière

Chapitre 2 : Analyse en composantes principales



Jaouad Madkour

Faculté d'Économie et de Gestion de Tanger · 27 août 2026

Avant de commencer

Construire les composantes principales

Synthèse

Avant de commencer

Pourquoi ce chapitre ?

Quand des dizaines, voire des centaines, d'actifs sont fortement corrélés (une courbe des taux, un univers d'actions d'un même secteur), estimer une matrice de covariance complète devient instable et peu maniable. L'analyse en composantes principales (ACP) construit, à partir des données elles-mêmes, un petit nombre de **facteurs statistiques** orthogonaux qui capturent l'essentiel de la variabilité commune — sans aucune hypothèse économique préalable.

Construire les composantes
principales

Composantes principales

Soit Σ la matrice de covariance ($n \times n$) des rendements de n actifs, **symétrique** donc diagonalisable (chapitre 6, algèbre linéaire) : $\Sigma = W\Lambda W^\top$, où W est orthogonale (ses colonnes sont les vecteurs propres, normés) et $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ contient les valeurs propres, ordonnées $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$ (Σ est semi-définie positive). Les **composantes principales** sont $Z = W^\top X$: chaque Z_k est une combinaison linéaire des rendements originaux, et

$$\text{Var}(Z_k) = \lambda_k, \quad \text{Cov}(Z_k, Z_l) = 0 \ (k \neq l).$$

Une rotation des axes, rien de plus

L'ACP ne fait que **tourner** le repère : au lieu de décrire les rendements par les n actifs originaux (corrélés), on les décrit par n facteurs orthogonaux (non corrélés), ordonnés par variance décroissante. Aucune information n'est perdue tant qu'on garde les n composantes — le gain apparaît en **tronquant**.

Part de variance expliquée

La part de la variance totale $\sum_k \lambda_k = \text{tr } \Sigma$ (chapitre 3, algèbre linéaire) expliquée par les m premières composantes est

$$\frac{\lambda_1 + \dots + \lambda_m}{\lambda_1 + \dots + \lambda_n}.$$

En pratique, on retient $m \ll n$ composantes expliquant, par exemple, 95% de la variance, et l'on modélise chaque rendement comme $X_i \approx \sum_{k=1}^m w_{ik} Z_k$ – un modèle factoriel **statistique**, structurellement identique à celui du chapitre 1, mais dont les facteurs Z_k n'ont pas nécessairement d'interprétation économique directe.

Exemple : deux actifs corrélés

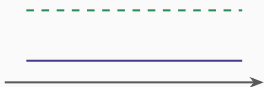
Deux actifs ont pour matrice de covariance $\Sigma = \begin{pmatrix} 4 & 3 \\ 3 & 4 \end{pmatrix}$ (en $\%^2$). Valeurs propres : $\text{tr } \Sigma = 8$,

$\det \Sigma = 7$, donc $\lambda^2 - 8\lambda + 7 = 0 \Rightarrow \lambda \in \{1, 7\}$. Vecteur propre pour $\lambda_1 = 7$: $(1, 1)/\sqrt{2}$ – la **composante commune** (les deux actifs bougent ensemble), qui capture $7/8 = 87.5\%$ de la variance totale. Vecteur propre pour $\lambda_2 = 1$: $(1, -1)/\sqrt{2}$ – la composante de **spread** (écart relatif), bien plus marginale.

Application emblématique : la courbe des taux

Niveau, pente, courbure

Appliquée aux variations quotidiennes des taux le long d'une courbe (2, 5, 10, 30 ans...), l'ACP révèle presque toujours la même structure : la **première** composante (très dominante, souvent $> 80\%$ de la variance) correspond à un déplacement parallèle de la courbe (**niveau**); la **deuxième**, à un changement de **pente** (les taux courts et longs bougent en sens opposés); la **troisième**, à un changement de **courbure**. Trois facteurs suffisent en général à expliquer plus de 95% de la variance d'une courbe des taux.



(a) niveau (CP1)



(b) pente (CP2)



(c) courbure (CP3)

Application en gestion : couverture et allocation par facteurs

Un gérant obligataire peut couvrir un portefeuille contre le risque de **niveau** (duration) sans se soucier des mouvements de pente ou de courbure, en ne neutralisant que l'exposition à la première

Synthèse

L'essentiel du chapitre 2

- L'ACP diagonalise la matrice de covariance $\Sigma = W\Lambda W^\top$; les composantes principales $Z = W^\top X$ sont non corrélées, de variance λ_k .
- Ordonnées par λ_k décroissant, les premières composantes concentrent l'essentiel de la variance : c'est la base de la réduction de dimension.
- Application phare : courbes de taux — niveau, pente, courbure expliquent typiquement $> 95\%$ de la variance.

Vers le chapitre 3

Les modèles factoriels (chapitres 1 et 2) supposent une matrice de covariance **connue** ou **estimée sur une fenêtre fixe**. Comment cette matrice évolue-t-elle dans le temps ? Les modèles classiques de volatilité et de corrélation apportent une première réponse.