

Vecteurs gaussiens

Xavier Bay

Septembre 2025

Table des matières

1	Motivation et contexte	2
2	Variable gaussienne et loi normale univariée	2
2.1	Définition	2
2.2	Propriétés	2
3	Vecteur gaussien et loi normale multivariée	3
3.1	Définition	3
3.2	Décomposition d'un vecteur gaussien	3
4	Simulation d'un vecteur gaussien	6
5	Propriétés fondamentales	7
6	Bruit blanc gaussien	7
7	Un jeu de données réelles pour illustrer	8

1 Motivation et contexte

La loi gaussienne (ou normale) univariée joue un rôle fondamental en Statistique et, plus largement, en Science des Données. Elle apparaît naturellement avec la Loi des Grands Nombres (LGN) et le Théorème Central Limite (TCL) qui en fournit une quantification précise faisant apparaître la célèbre densité de **de Moivre, Laplace et Gauss**. La **loi normale multivariée ou multigaussienne** est une généralisation multidimensionnelle de la loi normale. Quant à un **vecteur gaussien**, c'est une généralisation multidimensionnelle d'une variable gaussienne ou normale (univariée).

2 Variable gaussienne et loi normale univariée

2.1 Définition

Une variable aléatoire réelle X est dite gaussienne (ou normale) si sa loi (ou distribution) est gaussienne, laquelle est décrite par une densité de probabilité de la forme suivante :

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

Notation : $X \sim \mathcal{N}(\mu, \sigma^2)$

Paramètres

- $\mu \in \mathbb{R}$ est l'espérance de X : $\mathbb{E}[X] = \mu$
- σ^2 est la variance de X : $\text{Var}(X) = \sigma^2$, $\sigma = \sqrt{\text{Var}(X)}$ supposé > 0 étant l'écart-type de X

Convention : $X = \mu$ constante sera encore dite normale $\mathcal{N}(\mu, 0)$ (toute la masse de probabilité est concentrée au point μ , la variance étant nulle)

Loi normale centrée-réduite ou standard : c'est la loi $\mathcal{N}(0, 1)$ de moyenne nulle et de variance 1 (écart-type unité)

2.2 Propriétés

- Symétrie : X et $2\mu - X$ ont même loi
- (Centrage et réduction) $Z = \frac{X-\mu}{\sigma}$ est de loi $\mathcal{N}(0, 1)$ et donc $X = \mu + \sigma Z$ avec $Z \sim \mathcal{N}(0, 1)$
- (Stabilité pour la somme) Toute combinaison linéaire de variables aléatoires normales **indépendantes** est encore de loi normale : si X_1, \dots, X_n sont indépendantes et les $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$ alors $X = X_1 + \dots + X_n$ est de loi $\mathcal{N}(\mu, \sigma^2)$ avec $\mu = \sum_{i=1}^n \mu_i$ et $\sigma^2 = \sum_{i=1}^n \sigma_i^2$

3 Vecteur gaussien et loi normale multivariée

3.1 Définition

Un vecteur aléatoire $X \in \mathbb{R}^d$ est dit **gaussien** de moyenne $\mu \in \mathbb{R}^d$ et de matrice de covariance $\Gamma \in \mathbb{R}^{d \times d}$ si sa loi admet la densité suivante

$$f_X(x) = \frac{1}{\sqrt{\det(\Gamma)} \sqrt{2\pi}^d} \exp\left(-\frac{1}{2}(x - \mu)^T \Gamma^{-1}(x - \mu)\right),$$

lorsque la matrice Γ est **inversible**. C'est la loi normale multivariée notée $\mathcal{N}(\mu, \Gamma)$. Cette loi est également définie lorsque Γ n'est pas inversible (0 valeur propre de Γ) (cf. infra).

Paramètres

- Espérance : $\mathbb{E}[X] = \mu$ est le vecteur des moyennes des composantes X_i
- Covariance : $\text{Cov}(X) = \mathbb{E}((X - \mu)(X - \mu)^T) = \Gamma$ est la matrice symétrique positive formée des variances sur la diagonale $\Gamma_{ii} = \text{Cov}(X_i, X_i) = \text{Var}(X_i) = \sigma_i^2$ et des covariances des composantes prises 2 à 2 en dehors de la diagonale :

$$\Gamma_{ij} = \text{Cov}(X_i, X_j) \quad (i \neq j)$$

Cas de la dimension $d = 1$: $\Gamma = \text{Var}(X) = \sigma^2$, $\Gamma^{-1} = \frac{1}{\sigma^2}$, $(x - \mu)^T \Gamma^{-1}(x - \mu) = \frac{x^2}{\sigma^2}$ et enfin $\sqrt{\det(\Gamma)} = \sigma > 0$, on retrouve bien l'expression de la densité normale univariée.

Cas général : c'est la matrice Γ qui permet de décrire toute la variabilité du vecteur aléatoire X dans l'espace en intégrant les dépendances éventuelles puisque $\Gamma_{ij} = \sigma_i \sigma_j \rho(X_i, X_j)$ où

$$\rho(X_i, X_j) = \frac{\text{Cov}(X_i, X_j)}{\sigma_i \times \sigma_j} \in [-1, +1]$$

est le **coefficient de corrélation linéaire** entre X_i et X_j .

En pratique, on utilise souvent la paramétrisation $\Gamma_{ij} = \sigma_i \sigma_j \rho_{ij}$. Par exemple, en dimension $d = 2$, il faut se donner 3 paramètres $\sigma_1 > 0$, $\sigma_2 > 0$ et $\rho \in]-1, +1[$ ($\rho = \rho_{12} = \rho_{21}$) pour définir une matrice de covariance Γ inversible (soit une matrice symétrique définie positive) :

$$\Gamma = \begin{pmatrix} \sigma_1^2 & \rho \sigma_1 \sigma_2 \\ \rho \sigma_1 \sigma_2 & \sigma_2^2 \end{pmatrix}.$$

Voir figure 1 pour une illustration.

3.2 Décomposition d'un vecteur gaussien

La matrice Γ étant symétrique définie positive, on peut la diagonaliser en base ortho-normée :

$$\Gamma = \mathbf{U} \Lambda \mathbf{U}^T$$

avec $\mathbf{U}^T \mathbf{U} = \mathbf{I}_d$ et $\Lambda = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_d \end{bmatrix}$ où $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d > 0$ sont les valeurs propres (> 0) ordonnées de la plus grande à la plus petite.

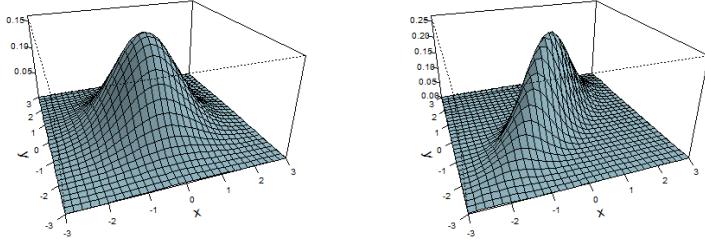


FIGURE 1 – Densité gaussienne bivariée avec $\mu = 0$, $\sigma_1 = \sigma_2 = 1$
 $\rho = 0$ (à gauche) et $\rho = 0.8$ (à droite)

Notons (u_1, \dots, u_d) la nouvelle base orthonormée formée par les vecteurs colonnes de la matrice U (base de vecteurs propres, cf. $U^T U = I_d$ et $\Gamma U = \Lambda U$). Ecrivons que

$$X = \mu + Y_1 u_1 + \dots + Y_d u_d,$$

ou, sous forme matricielle, $X = \mu + UY$ en posant $Y = U^T(X - \mu)$. Les nouvelles variables

$$Y_j = \langle X - \mu | u_j \rangle = u_j^T (X - \mu) = U_{1j}(X_1 - \mu_1) + \dots + U_{dj}(X_d - \mu_d)$$

sont les coordonnées euclidiennes du vecteur centré $X - \mu$ dans la base des vecteurs propres : elles sont **indépendantes** et de **lois respectives** $\mathcal{N}(0, \lambda_i)$. Ainsi, on peut écrire que

$$X = \mu + Y_1 u_1 + \dots + Y_d u_d = \sqrt{\lambda_1} Z_1 u_1 + \dots + \sqrt{\lambda_d} Z_d u_d,$$

avec Z_1, \dots, Z_d variables i.i.d. $\mathcal{N}(0, 1)$. Sous forme matricielle,

$$X = \mu + \Sigma Z,$$

en posant $\Sigma = U\sqrt{\Lambda}$. Le vecteur Z est de loi $\mathcal{N}(0, I_d)$, l'analogue en dimension d de la loi normale centrée-réduite en dimension 1. La décomposition ci-dessus généralise l'écriture $X = \mu + \sigma Z$ de la dimension 1 qui consiste à centrer et réduire X pour obtenir Z . Remarquer que $\Gamma = \Sigma \Sigma^T$ et donc $\mathcal{N}(\mu, \Gamma) = \mathcal{N}(\mu, \Sigma \Sigma^T)$ version multidimensionnelle de la loi $\mathcal{N}(\mu, \sigma^2)$.

Définissons maintenant un vecteur gaussien en toute généralité (la définition donnée en terme de densité de probabilité gaussienne suppose que la matrice de covariance Γ qui est toujours une matrice symétrique positive soit de plus inversible, c'est-à-dire définie positive). Reprenons la décomposition du vecteur X dans le repère associé aux vecteurs propres de Γ :

$$X = \mu + Y_1 u_1 + \dots + Y_d u_d,$$

où les variables $Y_j = \langle X - \mu | u_j \rangle$ sont les coordonnées du vecteur centré $X - \mu$ dans la base des vecteurs propres. Dans le cas général (X vecteur aléatoire pas nécessairement gaussien), les nouvelles variables Y_1, \dots, Y_d sont (centrées) et non linéairement corrélées : $\text{Cov}(Y_k, Y_l) = 0$ si $k \neq l$. De plus, on a $\text{Var}(Y_j) = u_j^T \Gamma u_j = \lambda_j$. En particulier $Y_j = 0$ si $\lambda_j = 0$. Le cas Γ dégénérée ou non inversible correspond exactement à la situation où $\lambda_1 \geq \dots \geq \lambda_r > \lambda_{r+1} = \dots = \lambda_d = 0$ où $r < d$ est la dimension de l'image $\text{Im}\Gamma$ de la matrice Γ et $d - r$ la dimension de son noyau $\text{Ker}\Gamma$ (sous-espace propre associé à la valeur propre 0). On a donc la décomposition

$$X = \mu + Y_1 u_1 + \dots + Y_r u_r,$$

qui montre en particulier que le vecteur aléatoire X a seulement r degrés de liberté et appartient au sous-espace affine $\mu + \text{Im}\Gamma$. Le vecteur est dit **gaussien** si les variables Y_1, \dots, Y_d sont indépendantes (pas seulement non linéairement corrélées) et de lois normales respectives $\mathcal{N}(0, \lambda_j)$. Cette définition est cohérente avec celle donnée en terme de densité qui correspond au cas où $r = d$ et des valeurs propres toutes > 0 .

Lien avec l'Analyse en Composantes Principales (ACP) La première direction propre u_1 est celle de plus grande variance ou variabilité du vecteur X (égale à λ_1), la seconde u_2 celle de plus grande variance (égale à λ_2) dans le sous-espace orthogonal à la première direction, etc. On a ainsi réalisé une décomposition de la variabilité globale du vecteur gaussien X (analyse de la variance ou ANOVA - ANalysis Of VAriance) puisque

$$\mathbb{E}[\|X - \mu\|^2] = \sigma_1^2 + \dots + \sigma_d^2 = \text{Tr}(\Gamma) = \lambda_1 + \dots + \lambda_d.$$

Dans le repère affine orthonormé $(\mu ; (u_1, \dots, u_d))$ d'origine le point μ , notons $y = U^T(x - \mu)$ le vecteur des nouvelles coordonnées. On a donc $x - \mu = Uy$ et la densité du vecteur gaussien au point x s'écrit simplement :

$$f_X(x) = \frac{1}{\sqrt{\lambda_1 \times \dots \lambda_d} \sqrt{2\pi}^d} \exp\left(-\frac{1}{2} \left(\frac{y_1^2}{\lambda_1^2} + \dots + \frac{y_d^2}{\lambda_d^2} \right)\right) = f_{Y_1}(y_1) \times \dots \times f_{Y_d}(y_d).$$

Ainsi, dans le cas bivarié $X = (X_1, X_2) \sim \mathcal{N}(\mu, \Gamma)$, les courbes d'iso-densité sont des ellipses d'équation $\frac{y_1^2}{\lambda_1^2} + \frac{y_2^2}{\lambda_2^2} = c$ (voir figure 2 et figure 3). En dimension $d = 3$, ce sont des ellipsoïdes (voir figure 4), et en dimension quelconque, des hyper-ellipsoïdes.

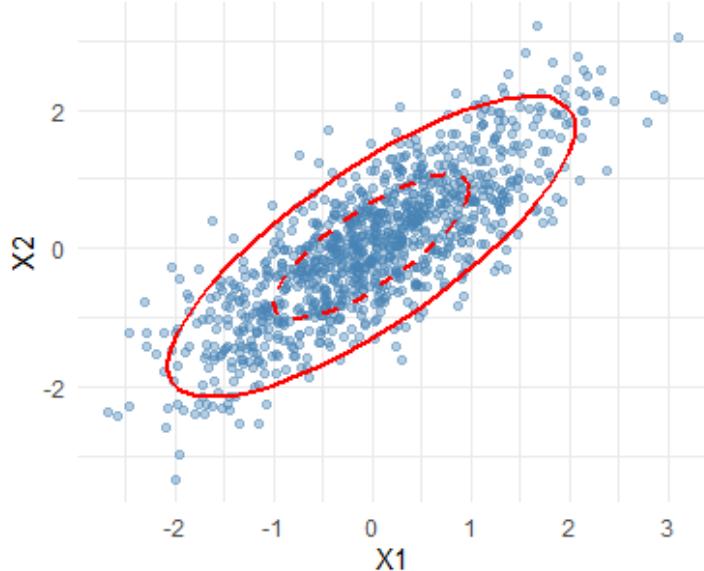


FIGURE 2 – Nuage d'individus de distribution gaussienne bivariée et ellipses de concentration à 50 % et 95 %

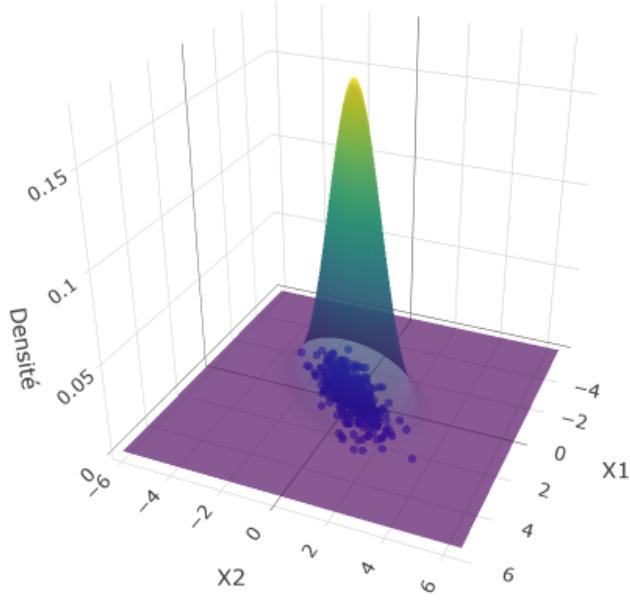


FIGURE 3 – Nuage d’individus et densité pour une loi normale bivariée.

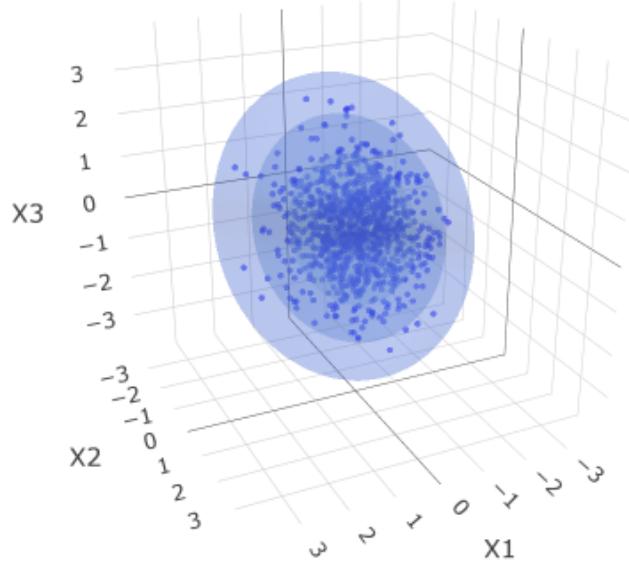


FIGURE 4 – Nuage d’individus et ellipsoïdes de concentration pour une loi normale en dimension 3.

4 Simulation d’un vecteur gaussien

Soit X un vecteur gaussien de moyenne μ et de covariance Γ . Pour simuler ce vecteur, on peut utiliser la décomposition

$$X = \mu + \Sigma Z,$$

avec $\Sigma = U\sqrt{\Lambda}$ et Z de loi $\mathcal{N}(0, I_d)$. Mais cela demande de diagonaliser la matrice de covariance. D’autres décompositions sont bien plus efficaces comme la décomposition de

Cholesky

$$X = \mu + SZ,$$

où S une matrice triangulaire inférieure telle que $\Gamma = SS^T$ et $Z \sim \mathcal{N}(0, I_d)$. Voir la figure 3 montrant des simulations (nuage d'individus) et la densité gaussienne correspondante. Sous R, on simule Z de composantes i.i.d. $\mathcal{N}(0, 1)$ avec la fonction rnorm().

5 Propriétés fondamentales

1. **Stabilité par transformation linéaire ou affine** : si $X \sim \mathcal{N}(\mu, \Gamma)$ et $A \in \mathbb{R}^{p \times d}, b \in \mathbb{R}^p$, alors

$$Y = AX + b \sim \mathcal{N}(A\mu + b, A\Gamma A^T)$$

est un vecteur gaussien de dimension p . C'est une propriété très précieuse en pratique pour l'analyse de tout modèle statistique linéaire avec hypothèse gaussienne sur le bruit. Ce sera le cas avec la Régression Linéaire Multiple mais aussi avec les modèles ARMA en Séries Temporelles. Dans le cas particulier où $A = u^T$ est une forme linéaire ($A \in \mathbb{R}^{1 \times d}$) et b un scalaire, la combinaison linéaire (affine) $Y = u^T X + b$ formée avec les composantes de X (nouvelle variable) est de loi normale univariée $\mathcal{N}(u^T \mu + b, u^T \Gamma u)$.

2. **Lois marginales** : compte conséquence de la propriété précédente, si $X \sim \mathcal{N}(\mu, \Sigma)$ est un vecteur gaussien de taille $d = d_1 + d_2$, alors les sous-vecteurs (X_1, \dots, X_{d_1}) , $(X_{d_1+1}, \dots, X_{d_1+d_2})$ sont encore gaussiens. En particulier, les composantes sont de lois normales (lois marginales simples).
3. **Indépendance et corrélation** : pour un vecteur gaussien X , **si les composantes X_j sont non corrélées deux à deux, alors elles sont (mutuellement) indépendantes**. Cela correspond exactement au cas où la matrice de covariance Γ est diagonale. C'est sans doute la propriété la plus remarquable pour un vecteur gaussien, à savoir que la propriété d'indépendance se ramène à celle de non corrélation linéaire (l'indépendance entraîne toujours la non corrélation linéaire mais la réciproque est grossièrement fausse). Cela n'est pas surprenant puisqu'il suffit seulement des coefficients de corrélations linéaires pour spécifier la distribution conjointe (en dehors des caractéristiques de moyenne et de variance pour les lois marginales)
4. **Lois conditionnelles** : supposons que (X, Y) soit un vecteur aléatoire bidimensionnel. Si l'on cherche à prédire Y à partir de X comme en Régression, il faut calculer les lois conditionnelles de Y sachant $X = x$ pour tout $x \in \mathbb{R}$. Si (X, Y) est un vecteur gaussien, alors ces lois conditionnelles sont des lois normales (univariées) de même variance (variance résiduelle). Ce résultat se généralise au cas de plusieurs variables prédictives et de plusieurs variables à prédire (voir chapitre 2 du cours de Probabilités Avancées)

6 Bruit blanc gaussien

Nous avons vu que tout vecteur gaussien se met sous la forme $X = \mu + \Sigma Z$ avec $Z \sim \mathcal{N}(0, I_d)$. On dit parfois que Z est un bruit blanc gaussien standard en dimension d puisque ses composantes sont **indépendantes** $\mathcal{N}(0, 1)$. En modélisation statistique, c'est en général l'hypothèse par défaut pour rendre compte du bruit présent dans les données

(la partie "purement aléatoire" qui dépend de facteurs aléatoires non pris en compte par le modèle). Il est donc essentiel de bien savoir manipuler de tels vecteurs $Z \sim \mathcal{N}(0, I_d)$.

Propriétés pour un bruit blanc gaussien

1. **(invariance par isométrie)** Soient $Z \sim \mathcal{N}(0, I_d)$ et U une isométrie $U^T U = I_d$. Alors, $UZ \sim \mathcal{N}(0, I_d)$
2. **(invariance par changement de repère orthonormé)** Soient $Z \sim \mathcal{N}(0, I_d)$ et U une matrice correspondant à un changement de base orthonormée (U isométrie). Alors, les nouvelles coordonnées forment encore un bruit blanc gaussien standard, i.e. le vecteur $U^T Z$ est $\mathcal{N}(0, I_d)$
3. **(Théorème de Cochran)** Soient $Z \sim \mathcal{N}(0, I_d)$ et P un projecteur orthogonal, i.e. $P = P^2$ et $P = P^T$. Considérons la décomposition orthogonale $Z = PZ + (I_d - P)Z$ où le vecteur PZ est la projection orthogonale de Z sur $\text{Im}P$ et le vecteur $(I_d - P)Z$ la projection orthogonale sur $(\text{Im}P)^\perp = \text{Im}(I_d - P)$. Alors, les deux vecteurs PZ et $(I_d - P)Z$ sont **indépendants**. De plus, $PZ \sim \mathcal{N}(0, P)$ et $(I_d - P)Z \sim \mathcal{N}(0, I_d - P)$. Enfin, la variable aléatoire $\|PZ\|^2$ est de loi χ_p^2 , loi du khi-deux à p degrés de liberté (ddl) où $p = \dim(\text{Im}P)$. De même, $\|(I_d - P)Z\|^2$ est de loi du khi-deux χ_{d-p}^2 à $(d-p)$ ddl.

7 Un jeu de données réelles pour illustrer

On considère les prix journaliers à la clôture pour $d = 6$ actions du CAC40 durant l'année 2024, soit au total 256 valeurs du 0/01/2024 au 31/12/2024 pour chacun des titres BNP, CA (Crédit Agricole), Hermès, Kering, LVMH et SG (Société Générale). Sur la figure 6, on a représenté l'évolution sur un même graphe des cours des 3 actions du secteur bancaire.

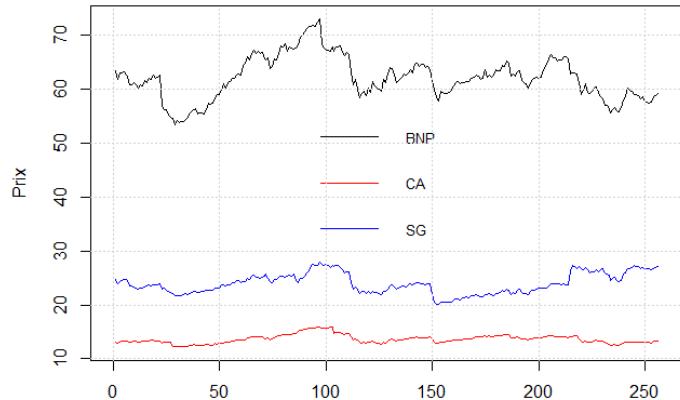


FIGURE 5 – Evolution sur 2024 des cours (en €) de BNP, CA et SG

On s'intéresse au vecteur X de dimension $d = 6$ des taux de rendement logarithmiques des 6 actions (taux de hausse ou de baisse par rapport au cours de la veille) et donc à la variabilité globale de ce vecteur dans \mathbb{R}^6 . Pour cela, on calcule à partir des données

(tableau de taille $n = 255 \times 6$ avec en ligne les jours boursiers et en colonnes les actions) les taux de rendements logarithmiques $\ln\left(\frac{S_i}{S_{i-1}}\right) \approx \frac{S_i - S_{i-1}}{S_{i-1}}$ pour chaque action S (en colonne), ce qui conduit à un nouveau tableau de données de taille $(n - 1) \times d = 255 \times 6$ puisque pour le premier jour boursier (première ligne) on ne peut pas faire le calcul (il faudrait le cours de la veille). La figure 7 permet de visualiser ces taux de rendement logarithmiques pour les 3 actions du secteur bancaire (sous-espace de dimension 3 dans un espace de dimension 6).

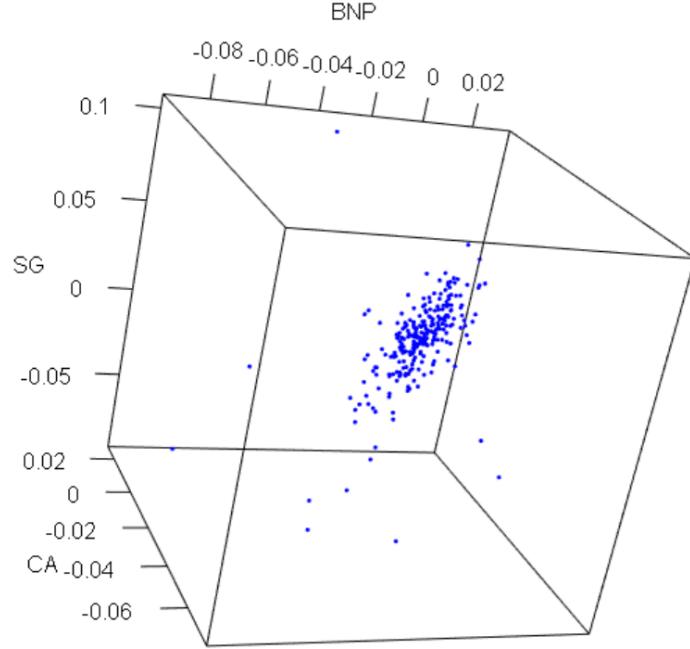


FIGURE 6 – Taux de rendement logarithmiques des actions BNP, CA et SG

Le modèle de base en Finance Quantitative est celui de mouvement brownien géométrique pour les prix des actions, ce qui implique en particulier que le vecteur X soit un vecteur gaussien et donc le sous-vecteur composé des 3 actions du secteur bancaire le soit également. Sur la figure 7, on constate que le nuage des individus semble se concentrer selon des ellipsoïdes mais fait apparaître toutefois quelques individus "extrêmes". On peut faire un boxplot pour s'en convaincre (voir figure 7).

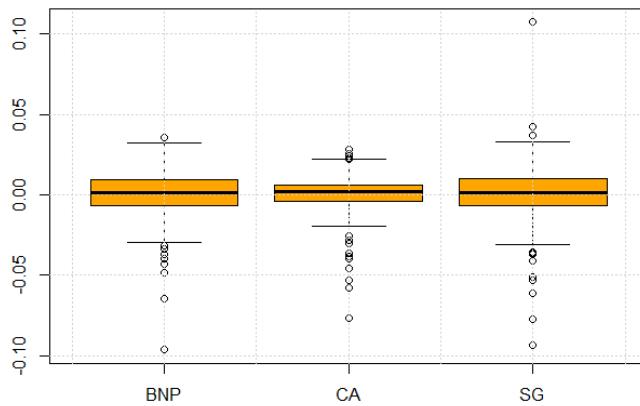


FIGURE 7 – Boîtes à moustaches pour les taux de rendement logarithmiques des actions BNP, CA et SG

On identifie bien des valeurs de taux de baisse ou de hausse très importantes, de l'ordre de -5% , voire -10% sachant que ce sont des variations de cours sur seulement 1 jour ! De nombreuses variantes du modèle gaussien ont été proposées pour tenir compte de ces valeurs extrêmes sachant qu'une distribution normale ne peut pas rendre compte de tels écarts. Néanmoins, le modèle gaussien est très utilisé comme très bonne approximation (linéaire) de la réalité, notamment dans le domaine de la gestion du risque pour l'évaluation et couverture des produits dérivés (cf. modèle de Black et Scholes de 1973).

Terminons en donnant le résultat d'une ACP normée sur les données de taux de rendement.

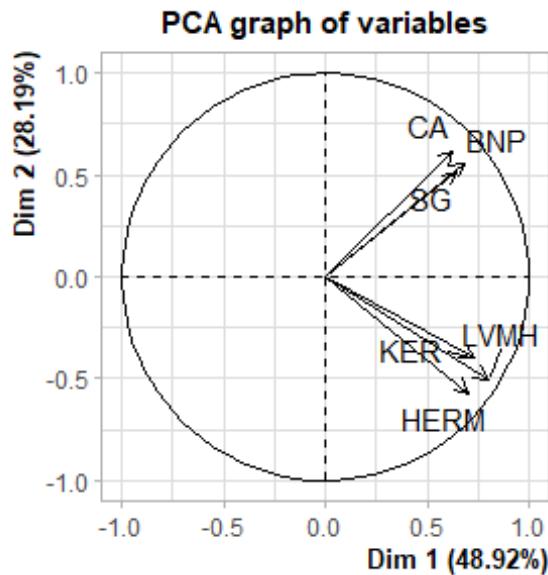


FIGURE 8 – Cercle des corrélations

On constate que près de 80 % de la variabilité est concentrée dans un plan. Comme attendu, les 3 actions du secteur bancaire sont "très" corrélées positivement entre elles et

de même celles du secteur du secteur du luxe. On voit également que ces deux groupes d'actions sont peu corrélés entre eux (angle droit ou presque). La gestion d'actifs met à profit ce type d'analyse pour réduire le risque de perte (principe de diversification). Voir la Théorie moderne du Portefeuille développée par Harry Markowitz (article Wikipedia, etc.).