

I Généralités sur les matrices symétriques réelles

Q1. Si A est orthodiagonalisable, alors il existe D diagonale et P orthogonale telle que $A = PDP^\top$. Alors

$$A^\top = (PDP^\top)^\top = PD^\top P^\top = PDP^\top = A$$

puisque $D^\top = D$ et $(P^\top)^\top = P$. Donc $A \in \mathcal{S}_n(\mathbb{R})$.

Réciproquement, toute matrice symétrique réelle est orthodiagonalisable, c'est la version matricielle du théorème spectral.

I.A – Un exemple dans $\mathcal{M}_3(\mathbb{R})$ $A_1 = \begin{pmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{pmatrix}$.

Q2. On remarque que $C_1(A_1) + C_3(A_1) = A_1 \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 7 \\ 0 \\ 7 \end{pmatrix} = 7 \times \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, donc $X_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ est un vecteur propre de A_1 associé à la valeur propre $\lambda_1 = 7$.

Q3. • La matrice $A_1 - 7I_3 = \begin{pmatrix} -4 & -2 & 4 \\ -2 & -1 & 2 \\ 4 & 2 & -4 \end{pmatrix}$ est de rang 1, donc par le théorème du rang, $\dim(E_7(A_1)) = 2$, et $E_7(A_1) = \text{Vect}(X_1, X_2)$ avec $X_2 = \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}$ (relation pour $X_2 : C_1 = 2C_2$).

• Comme A_1 est symétrique réelle, elle est (ortho)diagonalisable, et la somme de ses valeurs propres (réelles) est égale à sa trace. On a le nombre 7 comme valeur propre double, et il reste une valeur propre simple λ_2 telle que $2 \times 7 + \lambda_2 = \text{tr}(A_1) = 12$, donc $\lambda_2 = -2$. Ainsi, $\text{sp}(A_1) = \{-2, 7\}$, avec -2 simple et 7 double.

Q4. • Les sous-espaces propres d'une matrice symétrique réelle sont deux à deux orthogonaux.

On en déduit ici que $E_{-2}(A_1) = (E_7(A_1))^\perp = \{X \in \mathcal{M}_{3,1}(\mathbb{R}), x + z = 0 \text{ et } x - 2y = 0\}$ (on traduit $X \perp X_1$ et $X \perp X_2$ avec le produit scalaire canonique).

Donc $X \in E_{-2}(A_1) \Leftrightarrow \begin{cases} x + z = 0 \\ x - 2y = 0 \end{cases} \Leftrightarrow \begin{cases} y = \frac{1}{2}x \\ z = -x \end{cases}$ Donc $E_{-2}(A_1) = \text{Vect}(X_3)$ avec $X_3 = \begin{pmatrix} 1 \\ 1/2 \\ -1 \end{pmatrix}$.

Rq : On peut aussi faire comme d'habitude et chercher une relation entre les colonnes de $A_1 + 2I_3$.

• On n'a pas encore une base orthogonale de $\mathcal{M}_{3,1}(\mathbb{R})$ (X_1 et X_2 appartiennent au même sous-espace propre, ils n'ont pas été choisis orthogonaux). On peut appliquer le procédé de Gram-Schmidt à (X_1, X_2) , inutile de mettre X_3 dans la liste pour orthogonaliser, il est déjà orthogonal à X_1 et X_2 .

Ou bien, en calculant le produit vectoriel de X_3 et X_1 , on obtient $X'_2 = \begin{pmatrix} 1/2 \\ -2 \\ -1/2 \end{pmatrix}$ et on a bien $A_1 X'_2 = 7X'_2$.

• Ainsi, (X_1, X'_2, X_3) constitue une base orthogonale de vecteurs propres de A_1 , il ne reste plus qu'à normer ces vecteurs pour orthodiagonaliser A_1 . On a donc

$$A_1 = PDP^\top = PDP^{-1}, \text{ avec } D = \begin{pmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & -2 \end{pmatrix} \text{ et } P = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{18}} & \frac{2}{3} \\ 0 & \frac{-2\sqrt{2}}{\sqrt{9}} & \frac{1}{3} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{18}} & \frac{-2}{3} \end{pmatrix}.$$

I.B – Un exemple dans $\mathcal{M}_n(\mathbb{R})$

Q5. $\phi : (P, Q) \mapsto \phi(P, Q) = \int_0^1 P(t)Q(t)dt$ est bien définie (intégrale d'une fonction continue sur un segment).

* *Symétrie* claire.

* *Bilinéarité* : linéarité à gauche découlant de la linéarité de l'intégrale ; bilinéarité par symétrie.

* *Positivité* : $\forall P \in \mathbb{R}[X], \phi(P, P) = \int_0^1 P^2(t)dt \geq 0$ car intégrale d'une fonction positive.

- * *Stricte positivité* : Si $\phi(P, P) = 0$, alors la fonction $t \mapsto P(t)^2$ étant **continue**, positive et d'intégrale nulle sur $[0, 1]$, elle est identiquement nulle sur $[0, 1]$. Donc le polynôme P a une **infinité de racines**, c'est le polynôme nul.

Donc ϕ est bien un produit scalaire sur $\mathbb{R}_{n-1}[X]$.

- Q6.** Pour $(i, j) \in \llbracket 0, n-1 \rrbracket^2$, on a $h_{i,j} = \phi(X^i, X^j) = \int_0^1 t^{i+j} dt = \frac{1}{i+j+1}$, donc (en faisant attention que la première ligne correspond à $i = 0$ dans l'expression précédente, etc) :

$$H = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \cdots & \frac{1}{n+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{1}{n} & \frac{1}{n+1} & \frac{1}{n+1} & \cdots & \frac{1}{2n-1} \end{pmatrix} \in \mathcal{S}_n(\mathbb{R})$$

- Q7.** En posant $U = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$, on obtient $U^\top H U = \sum_{i=1}^n \sum_{j=1}^n h_{i-1,j-1} u_i u_j = \sum_{1 \leq i,j \leq n} \phi(X^{i-1}, X^{j-1}) u_i u_j$.

Rq : Pour éviter le décalage d'indice, on peut poser $U = (u_0 \ \cdots \ u_{n-1})^\top$, la somme va alors de 0 à $n-1$.

Avec la bilinéarité on peut encore écrire : $U^\top H U = \phi \left(\sum_{i=1}^n u_i X^{i-1}, \sum_{j=1}^n u_j X^{j-1} \right)$.

- Q8.** La symétrie du produit scalaire et la définition des coefficients de H donnent la symétrie de H .

$H \in \mathcal{S}_n^{++}(\mathbb{R})$ ssi $\text{Sp}(H) \subset \mathbb{R}_+^*$ ou encore ssi $U^\top H U > 0$ pour tout $U \neq 0$. On va évidemment utiliser la deuxième caractérisation ici, avec Q7.

$U^\top H U = \phi \left(\sum_{i=1}^n u_i X^{i-1}, \sum_{j=1}^n u_j X^{j-1} \right)$. En introduisant le polynôme $P_U = \sum_{k=1}^n u_k X^{k-1} \in \mathbb{R}_{n-1}[X]$, on peut donc écrire :

$$U^\top H U = \phi(P_U, P_U).$$

Si $U \neq 0$, alors le polynôme P_U a au moins un coefficient non nul, donc $P_U \neq 0$ et $\phi(P_U, P_U) > 0$ car le produit scalaire est défini positif. On a bien : $\forall U \in \mathcal{M}_{n,1}(\mathbb{R}), U \neq 0, U^\top H U > 0$.

On en déduit que $H \in \mathcal{S}_n^{++}(\mathbb{R})$ (on a vérifié la définition). Par propriété du cours, on en déduit que $\text{Sp}(H) \subset \mathbb{R}_+^*$.

I.C – Rayon spectral

Si $A \in \mathcal{M}_n(\mathbb{R})$ de spectre non vide, $\rho(A) = \max_{\lambda \in \text{sp}(A)} |\lambda|$ (rayon spectral).

- Q9.** • *Classique* : Si A est nilpotente, alors $\text{sp}(A) = \{0\}$.

En effet, la matrice A ne peut être inversible puisqu'il existe $p \in \mathbb{N}^*$ tel que $A^p = 0_n$. Donc $0 \in \text{sp}(A)$, et $\text{sp}(A) \neq \emptyset$.

D'autre part, X^p est un polynôme annulateur de A , donc le spectre de A est inclus dans l'ensemble des racines de X^p : $\text{sp}(A) \subset \{0\}$.

- Comme $\text{sp}(A) = \{0\}$, $\rho(A) = 0$.

- Q10.** L'ensemble C est la sphère de centre 0 et de rayon 1 pour la norme euclidienne canonique sur $\mathcal{M}_{n,1}(\mathbb{R})$, et toute sphère est une partie fermée de l'espace vectoriel normé dans lequel elle se trouve (cours).

On peut également utiliser la caractérisation séquentielle : soit $(U_n)_{n \in \mathbb{N}}$ est une suite d'éléments de C , convergente dans $\mathcal{M}_{n,1}(\mathbb{R})$ vers $V \in \mathcal{M}_{n,1}(\mathbb{R})$, alors $U_n^\top \xrightarrow{n \rightarrow +\infty} V^\top$ et $U_n^\top U_n \xrightarrow{n \rightarrow +\infty} V^\top V$.

Mais on a aussi : $\forall n \in \mathbb{N}, U_n^\top U_n = 1$. Donc par unicité de la limite : $V^\top V = 1$, i.e. $V \in C$.

- Q11.** L'application $f : \begin{cases} \mathcal{M}_{n,1}(\mathbb{R}) & \rightarrow \mathbb{R} \\ U & \mapsto |U^\top A U| \end{cases}$ est continue.

En effet, l'application $\alpha : U \mapsto AU$ est continue car c'est un endomorphisme de l'espace vectoriel de dimension finie $\mathcal{M}_{n,1}(\mathbb{R})$, l'application $\beta : (U, V) \mapsto U^\top V = (U|V)$ est continue car c'est une forme bilinéaire sur l'espace de dimension finie $\mathcal{M}_{n,1}(\mathbb{R})$, enfin la valeur absolue $v : x \mapsto |x|$ est continue sur $\mathbb{R}$, donc $f : U \mapsto v(\beta(U, \alpha(U)))$ est continue par composition.

D'après le **théorème des bornes atteintes**, f admet donc un maximum sur la partie fermée bornée C de $\mathcal{M}_{n,1}(\mathbb{R})$.

- Q12.** En supposant toujours $\text{sp}(A) \neq \emptyset$, il existe une valeur propre réelle de A telle que $|\lambda| = \rho(A)$.
 Soit U_0 un vecteur propre unitaire associé à cette valeur propre. Alors $U_0 \in C$ et $AU_0 = \lambda U_0$.
 Donc $U_0^\top AU_0 = \lambda U_0^\top U_0 = \lambda$, puis $|U_0^\top AU_0| = \rho(A)$. Finalement,

$$\rho(A) = f(U_0) \leq \max_{U \in C} f(U) = \max_{U \in C} |U^\top AU|.$$

I.D – Rayon spectral d’une matrice symétrique

- Q13.** Si A est symétrique réelle, elle est orthodiagonalisable : il existe $P \in \mathcal{O}_n(\mathbb{R})$ et $D = \text{diag}(\lambda_1, \dots, \lambda_n)$ diagonale réelle telles que $A = PDP^\top$.

Alors $\rho(A) = \max_{1 \leq i \leq n} |\lambda_i|$ donc il existe $i_0 \in \llbracket 1, n \rrbracket$ tel que $\rho(A) = |\lambda_{i_0}|$, i.e. $|\lambda_{i_0}| \geq |\lambda_i|$ pour tout $i \in \llbracket 1, n \rrbracket$.

Pour tout $U \in C$, on a $U^\top AU = U^\top PDP^\top U = V^\top DV$ en posant $V = P^\top U$.

Comme $PP^\top = I_n$, on déduit $V^\top V = U^\top PP^\top U = U^\top U = 1$ et, si on pose $V = (v_1 \ \dots \ v_n)^\top$, alors

$$|U^\top AU| = |V^\top DV| = \left| \sum_{i=1}^n \lambda_i v_i^2 \right| \leq \sum_{i=1}^n |\lambda_i| v_i^2 \leq |\lambda_{i_0}| \cdot \sum_{i=1}^n v_i^2 = |\lambda_{i_0}| V^\top V = |\lambda_{i_0}| = \rho(A)$$

L’inégalité précédente est valable pour tout U de norme 1, donc $\max_{U \in C} |U^\top AU| \leq \rho(A)$. Et la valeur $|\lambda_{i_0}| = \rho(A)$ est atteinte si on prend un vecteur propre unitaire U_0 associé à la valeur propre λ_{i_0} d’après **Q12**.

Rq : On peut aussi passer par l’endomorphisme de $\mathcal{M}_{n,1}(\mathbb{R})$ canoniquement associé à A , il est autoadjoint. On considère alors une base orthonormée de vecteurs propres et on écrit plus facilement le produit scalaire de deux vecteurs quelconques à l’aide de cette base orthonormée.

- Q14.** Si $\text{sp}(A) \subset \mathbb{R}^+$, en reprenant les notations introduites dans la question précédente, on a

$$U^\top AU = V^\top DV = \sum_{i=1}^n \lambda_i v_i^2 \geq 0$$

puisque les λ_i sont positifs. Donc $|U^\top AU| = U^\top AU$ et $\max_{U \in C} (U^\top AU) = \rho(A)$.

- Q15.** L’application ρ est bien définie sur $\mathcal{S}_n(\mathbb{R})$ (toute matrice symétrique réelle a un spectre non vide, inclus dans $\mathbb{R}$).

* On a bien $\rho(A) \in \mathbb{R}^+$ pour tout $A \in \mathcal{S}_n(\mathbb{R})$, donc ρ est une application positive.

* $\rho(A) = 0 \iff \max_{\lambda \in \text{sp}(A)} |\lambda| = 0 \iff \text{sp}(A) = \{0\}$.

Cela équivaut à $A = 0_n$ puisque A est diagonalisable avec 0 comme seule valeur propre. Donc ρ vérifie l’axiome de séparation.

* Si $A \in \mathcal{S}_n(\mathbb{R})$ et $\alpha \in \mathbb{R}$, alors $\text{sp}(\alpha A) = \alpha \text{sp}(A) = \{\alpha \lambda ; \lambda \in \text{sp}(A)\}$, donc

$$\rho(\alpha A) = \max_{\lambda \in \text{sp}(A)} (|\alpha| |\lambda|) = |\alpha| \cdot \max_{\lambda \in \text{sp}(A)} |\lambda| = |\alpha| \rho(A),$$

car $|\alpha|$ ne dépend pas de $\lambda \in \text{Sp}(A)$. Donc ρ vérifie l’homogénéité.

* Soient A et B dans $\mathcal{S}_n(\mathbb{R})$. Alors, pour tout $U \in C$, on a

$$|U^\top (A+B)U| = |U^\top AU + U^\top BU| \leq |U^\top AU| + |U^\top BU| \leq \rho(A) + \rho(B)$$

on a mené la suite d’inégalités jusqu’à obtenir un majorant indépendant de U .

Cette majoration étant valable pour tout $U \in C$, on en déduit que $\rho(A+B) \leq \rho(A) + \rho(B)$ grâce à **Q13**.

Donc ρ vérifie l’inégalité triangulaire.

Donc ρ est une norme sur l’espace vectoriel $\mathcal{S}_n(\mathbb{R})$.

II Matrice de covariance

II.A – L'énoncé suppose que $\mathbb{E}(Y)$ et Σ_Y sont bien définies, donc les Y_i^2 sont toutes d'espérances finies.

Q16. • La matrice Σ_Y est symétrique puisque $\sigma_{j,i} = \text{cov}(Y_j, Y_i) = \text{cov}(Y_i, Y_j) = \sigma_{i,j}$ pour tout (i, j) .

• De plus, pour $(i, j) \in \llbracket 1, n \rrbracket^2$, $\sigma_{i,j} = \text{cov}(Y_i, Y_j) = \mathbb{E}\left((Y_i - \mathbb{E}(Y_i))(Y_j - \mathbb{E}(Y_j))\right)$.

En notant $M = (Y - \mathbb{E}(Y))(Y - \mathbb{E}(Y))^\top$, matrice carrée de taille n , on a $M_{i,j} = (Y_i - \mathbb{E}(Y_i))(Y_j - \mathbb{E}(Y_j))$.

Par définition (énoncé), la matrice $\mathbb{E}(M)$ est calculée terme à terme, donc pour tout $(i, j) \in \llbracket 1, n \rrbracket$:

$$(\mathbb{E}(M))_{i,j} = \mathbb{E}(M_{i,j}) = \mathbb{E}\left((Y_i - \mathbb{E}(Y_i))(Y_j - \mathbb{E}(Y_j))\right) \stackrel{\text{déf de cov}}{=} \text{cov}(Y_i, Y_j) = \sigma_{i,j}$$

ce qui prouve l'égalité entre les matrices : $\mathbb{E}(M) = \mathbb{E}\left((Y - \mathbb{E}(Y))(Y - \mathbb{E}(Y))^\top\right) = (\sigma_{i,j})_{1 \leq i, j \leq n} = \Sigma_Y$.

• Soit $U = (u_1 \ \cdots \ u_n)^\top \in \mathcal{M}_{n,1}(\mathbb{R})$ un vecteur constant.

Pour $(i, j) \in \llbracket 1, n \rrbracket^2$, le coefficient d'indice (i, j) de Σ_{Y+U} est, par bilinéarité de la covariance,

$$\text{cov}(Y_i + u_i, Y_j + u_j) = \text{cov}(Y_i, Y_j) + \text{cov}(Y_i, u_j) + \text{cov}(u_i, Y_j) + \text{cov}(u_i, u_j) = \text{cov}(Y_i, Y_j) = \sigma_{i,j}$$

car lorsque l'une des deux variables aléatoires est constante, la covariance est nulle ($\text{cov}(Y_i, u_j) = \mathbb{E}(Y_i u_j) - \mathbb{E}(Y_i)\mathbb{E}(u_j) = u_j \mathbb{E}(Y_i) - \mathbb{E}(Y_i)u_j = 0$). Ainsi, $\Sigma_{Y+U} = \Sigma_Y$.

Q17. Posons $Z(\omega) = \begin{pmatrix} Z_1(\omega) \\ \vdots \\ Z_p(\omega) \end{pmatrix}$ pour tout $\omega \in \Omega$, et $M = (m_{i,j})_{1 \leq i \leq p, 1 \leq j \leq n}$.

• Par définition, pour tout $i \in \llbracket 1, p \rrbracket$: $Z_i = \sum_{j=1}^n m_{i,j} Y_j$, combinaison linéaire de variables d'espérances finies.

Par linéarité de l'espérance, chaque Z_i est d'espérance finie et $\mathbb{E}(Z_i) = \sum_{j=1}^n m_{i,j} \mathbb{E}(Y_j)$ pour tout i . Donc matriciellement : $\mathbb{E}(Z) = M \times \mathbb{E}(Y)$.

• La matrice Σ_Y est bien définie par hypothèse, donc les $\text{cov}(Y_i, Y_j)$ existent toutes.

Pour tout $(i, j) \in \llbracket 1, p \rrbracket^2$, Z_i^2 et Z_j^2 sont des combinaisons linéaires de termes de la forme $Y_k Y_l$, tous d'espérances finies. Donc Z_i^2 et Z_j^2 sont d'espérances finies, ce qui justifie l'existence de $\text{cov}(Z_i, Z_j)$.

Par bilinéarité de la covariance : $\text{cov}(Z_i, Z_j) = \text{cov}\left(\sum_{k=1}^n m_{i,k} Y_k, \sum_{l=1}^n m_{j,l} Y_l\right) = \sum_{k=1}^n \sum_{l=1}^n m_{i,k} \sigma_{k,l} m_{j,l}$.

On reconnaît bien l'écriture du coefficient d'indices (i, j) de la matrice produit $M \Sigma_Y M^\top$, cqfd.

II.B – Propriété des valeurs propres

P , matrice de passage de la base canonique à une base orthonormée de vecteurs propres de Σ_Y , existe bien car Σ_Y est symétrique réelle donc orthodagonalisable. $X = P^\top Y = (X_1 \ \cdots \ X_n)^\top$.

Q18. L'orthodagonalisation de la matrice symétrique Σ_Y s'écrit : $P^{-1} \Sigma_Y P = P^\top \Sigma_Y P = D$, où D est diagonale.

On a d'après **Q17** : $\Sigma_X = P^\top \Sigma_Y (P^\top)^\top = P^\top \Sigma_Y P$, donc $\Sigma_X = D$ est diagonale.

Q19. D'après **Q18**, $\text{sp}(\Sigma_Y) = \text{sp}(\Sigma_X)$ car les deux matrices sont semblables.

Or, les vp de la matrice diagonale Σ_X sont ses coefficients diagonaux, qui sont les $\text{cov}(X_i, X_i) = \mathbb{V}(X_i) \geq 0$.

Ainsi, $\text{sp}(\Sigma_Y) \subset \mathbb{R}^+$, autrement dit (puisque Σ_Y est symétrique) $\Sigma_Y \in \mathcal{S}_n^+(\mathbb{R})$.

Q20. Les matrices Σ_X et Σ_Y sont semblables, donc ont la même trace. Donc les vecteurs aléatoires X et Y ont la même variance totale : $\mathbb{V}_T(X) = \text{tr}(\Sigma_X) = \text{tr}(\Sigma_Y) = \mathbb{V}_T(Y)$.

II. C – Étude de la réciproque

Q21. $D = \text{diag}(\lambda_1, \dots, \lambda_n)$, coefficients diagonaux positifs. On considère n variables aléatoires réelles $Z_1, \dots, Z_n$, deux à deux indépendantes (pour que les $\text{cov}(Z_i, Z_j)$ soient nulles lorsque $i \neq j$, coefficients non diagonaux).

Si de plus les variables $Z_1, \dots, Z_n$ ont pour variances respectives $\lambda_1, \dots, \lambda_n$, alors $\Sigma_Z = \text{diag}(\lambda_1, \dots, \lambda_n) = D$.

On prend par exemple $Z_i \sim \mathcal{P}(\lambda_i)$ lorsque $\lambda_i > 0$ (de variance λ_i), et $Z_i = 1$ (constante, valeur arbitraire) lorsque $\lambda_i = 0$ car la variance d'une variable constante est nulle.

Q22. On orthodagonalise A , soit $A = P D P^\top$ avec $P \in \mathcal{O}_n(\mathbb{R})$ et $D = \text{diag}(\lambda_1, \dots, \lambda_n)$, les λ_i étant positifs.

D'après **Q21.**, il existe un vecteur aléatoire Z à valeurs dans $\mathcal{M}_{n,1}(\mathbb{R})$ tel que $\Sigma_Z = D$. D'après **Q17** ou **Q18**, $Y = P Z$ vérifie alors $\Sigma_Y = P \Sigma_Z P^\top = P D P^\top = A$.

II.D – $U = (u_1 \ \cdots \ u_n)^\top \in \mathcal{M}_{n,1}(\mathbb{R})$; $X = U^\top Y \in \mathcal{M}_1(\mathbb{R}) = \mathbb{R}$.

Q23. X est une variable aléatoire réelle, $X = \sum_{i=1}^n u_i Y_i$ en posant $Y = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}$: X est une combinaison linéaire de variables aléatoires réelles admettant une variance, elle admet donc aussi une variance et

$$\mathbb{V}(X) = \mathbb{V}\left(\sum_{i=1}^n u_i Y_i\right) = \sum_{i=1}^n \mathbb{V}(u_i Y_i) + 2 \sum_{1 \leq i < j \leq n} \text{cov}(u_i Y_i, u_j Y_j) = \sum_{i=1}^n u_i^2 \mathbb{V}(Y_i) + \sum_{i \neq j} u_i u_j \text{cov}(Y_i, Y_j).$$

On peut également écrire $\mathbb{V}(X) = \sum_{1 \leq i, j \leq n} u_i u_j \text{cov}(Y_i, Y_j)$ car $\mathbb{V}(Y_i) = \text{cov}(Y_i, Y_i)$ pour tout i .

Par ailleurs : $U^\top \Sigma_Y U = \sum_{1 \leq i, j \leq n} u_i u_j (\Sigma_Y)_{i,j} = \sum_{1 \leq i, j \leq n} u_i u_j \text{cov}(Y_i, Y_j)$, donc $U^\top \Sigma_Y U = \mathbb{V}(X) = \mathbb{V}(U^\top Y)$.

Remarque. Avec les mêmes notations, $\mathbb{E}(X) = U^\top \mathbb{E}(Y)$ par linéarité de l'espérance. On s'en servira en **Q27**.

II.E – Image de Σ_Y . $r = \text{rg}(\Sigma_Y)$.

Q24. Si $r = n$, Σ_Y est inversible (l'endomorphisme canoniquement associé est bijectif) donc $\text{Im}(\Sigma_Y) = \mathcal{M}_{n,1}(\mathbb{R})$. Alors $\{Y - \mathbb{E}(Y) \in \text{Im } \Sigma_Y\}$ est l'événement certain, probabilité 1 effectivement.

Q25. *C'est une propriété générale des matrices symétriques réelles : si $S \in \mathcal{S}_n(\mathbb{R})$, alors les sous-espaces $\ker(S)$ et $\text{Im}(S)$ sont supplémentaires orthogonaux.*

Justification : Si $V = SU \in \text{Im}(S)$ et $W \in \ker(S)$, alors $(V | W) = (SU | W) = (U | S^\top W) = (U | SW) = 0$.

Cela entraîne que ces deux sous-espaces sont en somme directe (orthogonale). Enfin, le théorème du rang permet de conclure que la somme de leurs dimensions est la dimension de $\mathcal{M}_{n,1}(\mathbb{R})$. Donc $\mathcal{M}_{n,1}(\mathbb{R}) = \ker(S) \oplus \text{Im}(S)$.

Q26. $d = \dim(\ker \Sigma_Y)$, $(V_1, \dots, V_d)$ base orthonormée de $\ker \Sigma_Y$.

Introduisons le vecteur constant $U = \mathbb{E}(Y)$. Alors d'après **Q23.**, pour tout $j \in \llbracket 1, d \rrbracket$:

$$\mathbb{V}\left(V_j^\top (Y - \mathbb{E}(Y))\right) = \mathbb{V}\left(V_j^\top (Y - U)\right) = V_j^\top \Sigma_{Y-U} V_j = V_j^\top \Sigma_Y V_j$$

en utilisant la fin de **Q16**. Enfin, $V_j \in \ker(\Sigma_Y)$ donc $\Sigma_Y V_j = 0$ et $\mathbb{V}\left(V_j^\top (Y - \mathbb{E}(Y))\right) = 0$.

Q27. • La variable aléatoire réelle $V_j^\top (Y - \mathbb{E}(Y))$ a une variance nulle, elle est donc presque sûrement constante : $\mathbb{V}(Z) = \mathbb{E}((Z - \mathbb{E}(Z))^2) = 0$ entraîne que pour toute valeur non nulle de $(Z - \mathbb{E}(Z))^2$, la probabilité associée est nulle (écrire la somme de termes positifs définissant l'espérance de cette variable positive), donc Z est presque sûrement égale à $\mathbb{E}(Z)$.

• Or d'après la fin de **Q23**, la linéarité de l'espérance donne $\mathbb{E}\left(V_j^\top (Y - \mathbb{E}(Y))\right) = V_j^\top \cdot \mathbb{E}(Y - \mathbb{E}(Y)) = V_j^\top \cdot 0 = 0$, donc la valeur de cette constante est 0. Ainsi la variable aléatoire est presque sûrement nulle :

$$\mathbb{P}\left(V_j^\top (Y - \mathbb{E}(Y)) = 0\right) = 1, \text{ pour tout } j \in \llbracket 1, d \rrbracket.$$

Q28. D'après **Q25.**, on a : $\{Y - \mathbb{E}(Y) \in \text{Im } \Sigma_Y\} = \{Y - \mathbb{E}(Y) \in (\ker \Sigma_Y)^\perp\} = \bigcap_{j=1}^d \left\{V_j^\top (Y - \mathbb{E}(Y)) = 0\right\}$.

$$\text{Donc } \mathbb{P}\left(Y - \mathbb{E}(Y) \in \text{Im } \Sigma_Y\right) = \mathbb{P}\left(\bigcap_{j=1}^d \left\{V_j^\top (Y - \mathbb{E}(Y)) = 0\right\}\right) = 1 - \mathbb{P}\left(\bigcup_{j=1}^d \overline{\left\{V_j^\top (Y - \mathbb{E}(Y)) = 0\right\}}\right).$$

Par sous-additivité de la probabilité :

$$\mathbb{P}\left(\bigcup_{j=1}^d \overline{\left\{V_j^\top (Y - \mathbb{E}(Y)) = 0\right\}}\right) \leq \sum_{j=1}^d \left(1 - \mathbb{P}\left(\left\{V_j^\top (Y - \mathbb{E}(Y)) = 0\right\}\right)\right) = d - d = 0.$$

Donc $\mathbb{P}\left(\bigcup_{j=1}^d \overline{\left\{V_j^\top (Y - \mathbb{E}(Y)) = 0\right\}}\right) = 0$, nécessairement, et ainsi $\mathbb{P}(Y - \mathbb{E}(Y) \in \text{Im } \Sigma_Y) = 1$ cqfd.

III Maximisation de la variance

$q_Y(U) = \mathbb{V}(U^\top Y)$ pour $U \in \mathcal{M}_{n,1}(\mathbb{R})$ tel que $U^\top U = 1$.

III.A – Un exemple dans $\mathcal{M}_{3,1}(\mathbb{R})$. $A_2 = \text{diag}(9, 5, 4)$.

Q29. Soit $Y = \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix}$, où Y_1, Y_2, Y_3 sont trois variables aléatoires réelles mutuellement indépendantes, de variances respectives 9, 5 et 4. Alors $\Sigma_Y = A_2$. On peut aussi se contenter d'invoquer **Q21**.

Q30. Pour tout $U \in C$, i.e. $U^\top U = 1$, on a $q_Y(U) = \mathbb{V}(U^\top Y) = U^\top \Sigma_Y U = U^\top A_1 U$ d'après **Q23**. Comme A_1 est symétrique à valeurs propres strictement positives, elle est symétrique définie positive donc on a déjà $U^\top A_1 U > 0$ pour tout $U \in C$ ($U \in C \Leftrightarrow \|U\| = 1$ donc $U \neq 0$).

Alors d'après **Q14** : $\max_{U \in C} q_Y(U) = \max_{U \in C} |q_Y(U)| = \rho(A_1) = \max(\text{sp}(A_1)) = 9$.

III.B – Cas général

Q31. Comme Σ_Y est symétrique à valeurs propres positives,

$$q_Y(U) = U^\top \Sigma_Y U = |U^\top \Sigma_Y U| \text{ pour tout } U \in C.$$

Il résulte de **Q14**. l'existence d'un maximum de q_Y sur C : $\max_{U \in C} q_Y(U) = \rho(\Sigma_Y) = \max(\text{sp}(\Sigma_Y))$.

Comme en **Q12.**, si λ est la plus grande valeur propre de Σ_Y (ici, elles sont toutes réelles positives), si $U_0 \in C$ est un vecteur propre associé, alors $q_Y(U_0) = \lambda = \rho(\Sigma_Y) = \max_{U \in C} q_Y(U)$, soit

$$\max_{U \in C} \mathbb{V}(U^\top Y) = \mathbb{V}(U_0^\top Y).$$

III.C – Étude d'un exemple

$\sigma, \gamma \in \mathbb{R}_+^*$; $\sigma_{i,i} = \sigma^2$ pour tout $i \in \llbracket 1, n \rrbracket$ et $\sigma_{i,j} = \sigma^2 \gamma$ pour $(i, j) \in \llbracket 1, n \rrbracket^2, i \neq j$.

Q32. Par l'inégalité de Cauchy-Schwarz sur la covariance, on doit avoir $\text{cov}(Y_1, Y_2)^2 \leq \mathbb{V}(Y_1)\mathbb{V}(Y_2)$, soit $\sigma_{1,2}^2 \leq \sigma_{1,1}\sigma_{2,2}$, soit ici $\sigma^4 \gamma^2 \leq \sigma^4$, ce qui entraîne $\gamma^2 \leq 1$, puis $\gamma \leq 1$.

On a facilement $\Sigma_Y = \sigma^2 (\gamma J + (1 - \gamma)I_n)$.

Q33. La matrice J est symétrique donc diagonalisable, elle est de rang 1 donc elle admet 0 comme valeur propre avec un sous-espace propre $E_0(J) = \text{Ker}(J)$ de dimension $n - 1$ par le théorème du rang.

Sa trace est n , donc n est valeur propre, le sous-espace propre associé étant de dimension 1, et $\text{sp}(J) = \{0, n\}$.

En posant $V = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$, on a $JV = nV$, donc V est un vecteur propre de J associé à sa valeur propre de module maximal, à savoir n .

Q34. Si $A \in \mathcal{M}_n(\mathbb{R})$, et si a et b sont deux réels, on montre facilement que

$$\text{sp}(aA + bI_n) = \{a\lambda + b ; \lambda \in \text{sp}(A)\}.$$

L'expression de Σ_Y obtenue en **Q32.** et l'étude du spectre de J faite en **Q33.** montrent alors que

$$\text{sp}(\Sigma_Y) = \left\{ \sigma^2(1 - \gamma), \sigma^2(1 + (n - 1)\gamma) \right\}.$$

Donc $\rho(\Sigma_Y) = \max(\text{sp}(\Sigma_Y)) = \sigma^2(1 + (n - 1)\gamma)$. D'après **Q31.**, pour maximiser q_Y sur C , on prend un vecteur propre unitaire de Σ_Y associé à la valeur propre maximale $\rho(\Sigma_Y)$, et l'expression de Σ_Y en fonction de J nous montre qu'il suffit pour cela de prendre un vecteur propre unitaire de J associé à sa valeur propre maximale n .

En prenant $U_0 = \frac{1}{\sqrt{n}} V = \frac{1}{\sqrt{n}} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$, on a un vecteur $U_0 \in C$ tel que $q_Y(U_0) = \max_{U \in C} q_Y(U)$, ce que l'on recherchait.

Q35. D'après **Q31.**, on a $\mathbb{V}(Z) = \mathbb{V}(U_0^\top Y) = \rho(\Sigma_Y) = \sigma^2(1 + (n - 1)\gamma)$, alors que $\mathbb{V}_T(Y) = \text{tr}(\Sigma_Y) = n\sigma^2$. Donc

$$\frac{\mathbb{V}(Z)}{\mathbb{V}_T(Y)} = \frac{1 + (n - 1)\gamma}{n} = 1 - (1 - \gamma)\left(1 - \frac{1}{n}\right).$$

III.D –

Q36. L'application $q_Y : U \mapsto U^\top \Sigma_Y U$ est toujours continue sur $\mathcal{M}_{n,1}(\mathbb{R})$ et à valeurs dans $\mathbb{R}$.

La partie C' est bornée car incluse dans C , fermée comme intersection de deux fermés puisque $C' = C \cap (\text{Vect}(U_0))^\perp$, et elle est non vide car il existe dans l'hyperplan $(\text{Vect}(U_0))^\perp$ des vecteurs unitaires.

Donc q_Y admet un maximum sur C' par le théorème des bornes atteintes.

Q37. Soit $\mathcal{B} = (E_1, \dots, E_n)$ une base orthonormale de vecteurs propres de Σ_Y , associés respectivement aux valeurs propres $\lambda_1, \dots, \lambda_n$. Soit $U \in C$, décomposons-le dans cette base : $U = \sum_{i=1}^n u_i E_i$, on a alors $1 = U^\top U = \|U\|^2 = \sum_{i=1}^n u_i^2$ puisque la base $\mathcal{B}$ est orthonormale.

D'autre part,

$$q_Y(U) = \mathbb{V}(U^\top Y) = U^\top \Sigma_Y U = \sum_{i=1}^n \sum_{j=1}^n u_i u_j E_i^\top \Sigma_Y E_j = \sum_{i=1}^n \lambda_i u_i^2$$

puisque $\Sigma_Y E_j = \lambda_j E_j$ et que $E_i^\top E_j = \delta_{i,j}$ (symbole de Kronecker).

Avec cette écriture, on retrouve que $q_Y(U) \leq \lambda_1 \sum_{i=1}^n u_i^2 = \lambda_1 = \rho(\Sigma_Y)$.

De plus, cette valeur est atteinte si et seulement si $\sum_{i=2}^n (\lambda_1 - \lambda_i) u_i^2 = 0$, donc si et seulement si $u_i = 0$ pour $i \in \llbracket 1, n \rrbracket$ (puisque une somme de réels positifs est nulle si et seulement si chaque terme est nul).

Donc $\max_{U \in C} q_Y(U) = \lambda_1$, et cette valeur maximale est atteinte uniquement si le vecteur unitaire U vaut E_1 ou $-E_1$, le vecteur U_0 de l'énoncé (*qui doit appartenir à C même si ce n'est pas précisé*) est donc nécessairement un de ces deux vecteurs.

Ensuite, $(\text{Vect}(U_0))^\perp$ est l'hyperplan $H = \text{Vect}(E_2, \dots, E_n)$, et un calcul analogue à celui qui précède montre que

$$\max_{U \in C'} q_Y(U) = \max_{U \in C \cap H} q_Y(U) = \lambda_2$$

et que cette valeur est atteinte en prenant $U = U_1 = \pm E_2$ (et seulement pour ces deux vecteurs).

Q38. Si A et B sont deux variables aléatoires réelles admettant un moment d'ordre deux, on a

$$\mathbb{V}(A+B) = \mathbb{V}(A) + 2 \text{cov}(A, B) + \mathbb{V}(B) \quad \text{et} \quad \mathbb{V}(A-B) = \mathbb{V}(A) - 2 \text{cov}(A, B) + \mathbb{V}(B),$$

donc $\text{cov}(A, B) = \frac{1}{4} (\mathbb{V}(A+B) - \mathbb{V}(A-B))$ (identité de polarisation).

Comme U_0 et U_1 sont des vecteurs propres unitaires de Σ_Y associés respectivement aux valeurs propres λ_1 et λ_2 , et que ces deux vecteurs sont orthogonaux, on a

$$U_0^\top \Sigma_Y U_1 = U_1^\top \Sigma_Y U_0 = 0, \text{ puis } U_0^\top \Sigma_Y U_0 = \lambda_1, U_1^\top \Sigma_Y U_1 = \lambda_2,$$

et finalement

$$\begin{aligned} \text{cov}(U_0^\top Y, U_1^\top Y) &= \frac{1}{4} \left(\mathbb{V}((U_0 + U_1)^\top Y) - \mathbb{V}((U_0 - U_1)^\top Y) \right) \\ &= \frac{1}{4} \left((U_0 + U_1)^\top \Sigma_Y (U_0 + U_1) - (U_0 - U_1)^\top \Sigma_Y (U_0 - U_1) \right) \\ &= \frac{1}{4} \left((\lambda_1 + \lambda_2) - (\lambda_1 + \lambda_2) \right) \\ &= 0 \end{aligned}$$

les variables $U_0^\top Y$ et $U_1^\top Y$ sont donc décorréliées.