Corrigés des épreuves du concours HEC 2008

Jean Mallet
Professeur en classes préparatoires, lycée Montaigne (Paris)

Michel Miternique
Professeur en classes préparatoires, lycée Jean-Baptiste Corot (Savigny sur Orge) et IPESUP (Paris)

Voie scientifique

BCE

BANQUE COMMUNE D'EPREUVES

Concepteur : H.E.C.

OPTION SCIENTIFIQUE

MATHEMATIQUES I

La présentation, la lisibilité, l'orthographe, la qualité de la rédaction, la clarté et la précision des raisonnements entreront pour une part importante dans l'appréciation des copies.
Les candidats sont invités à encadrer dans la mesure du possible les résultats de leurs calculs.
Ils ne doivent faire usage d'aucun document : **l'utilisation de toute calculatrice et de tout matériel électronique est interdite.**
Seule l'utilisation d'une règle graduée est autorisée.

Dans tout le problème, n et p désignent deux entiers vérifiant $1 \leqslant p \leqslant n$. On note $\mathcal{M}_{n,p}(\mathbb{R})$ l'espace vectoriel des matrices à n lignes et p colonnes, à coefficients réels. La transposée d'une matrice A de $\mathcal{M}_{n,p}(\mathbb{R})$ est notée ${}^tA$. Lorsqu'une matrice A est inversible, on note A^{-1} son inverse.
Dans tout le problème, on identifie les deux espaces vectoriels $\mathcal{M}_{n,1}(\mathbb{R})$ (respectivement $\mathcal{M}_{p,1}(\mathbb{R})$) et $\mathbb{R}^n$ (resp. $\mathbb{R}^p$), c'est-à-dire qu'on identifie un vecteur (point) de $\mathbb{R}^n$ (resp. $\mathbb{R}^p$) avec le vecteur-colonne de ses coordonnées dans la base canonique de $\mathbb{R}^n$ (resp. $\mathbb{R}^p$).
On munit $\mathbb{R}^n$ (resp. $\mathbb{R}^p$) de sa structure euclidienne canonique, et pour tous vecteurs u et v de $\mathbb{R}^n$ (resp. $\mathbb{R}^p$), on note $\langle u, v\rangle = {}^tuv$ leur produit scalaire, et $||u||$ la norme de u associée.

Pour tout i de $[\![1, n]\!]$, on note f_i une fonction définie sur $\mathbb{R}^p$ à valeurs réelles, et de classe C^2 sur $\mathbb{R}^p$. Soit F la fonction définie sur $\mathbb{R}^p$, à valeurs réelles, par : $F(x_1, x_2, \ldots, x_p) = \frac{1}{2}\sum_{i=1}^{n}[f_i(x_1, x_2, \ldots, x_p)]^2$.

Autrement dit, si $X = (x_1, \ldots, x_p)$ est un point de $\mathbb{R}^p$, on a : $F(X) = \frac{1}{2}\sum_{i=1}^{n} f_i^2(X) = \frac{1}{2}||f(X)||^2$, en notant $f(X)$ le vecteur $(f_1(X), \ldots, f_n(X))$.

Le problème a pour objet l'étude de quelques aspects mathématiques liés à la recherche du minimum de la fonction F.

Partie I. Gradient et hessienne

Pour tout point $X = (x_1, x_2, \ldots, x_p)$ de $\mathbb{R}^p$, on rappelle que :

- le *gradient* de F au point X, noté $\nabla F(X)$, est le vecteur de $\mathbb{R}^p$ suivant :

$$\nabla F(X) = \left(\frac{\partial F}{\partial x_1}(X), \ldots, \frac{\partial F}{\partial x_p}(X)\right)$$

- la *matrice hessienne* de F au point X, notée $\nabla^2 F(X)$, est la matrice symétrique de $\mathcal{M}_p(\mathbb{R})$ suivante :

$$\nabla^2 F(X) = \left(\frac{\partial^2 F}{\partial x_k \partial x_j}(X)\right)_{1\leqslant k,j\leqslant p}$$

Pour tout point $X = (x_1, \dots, x_p)$ de $\mathbb{R}^p$, on note $J(X)$ la matrice de $\mathcal{M}_{n,p}(\mathbb{R})$ définie par :

$$J(X) = \left(\frac{\partial f_i}{\partial x_j}(X)\right)_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant p}}$$

dans laquelle i désigne l'indice de ligne et j l'indice de colonne. On pose : $G(X) = {}^tJ(X)J(X)$.

Si X est un point de $\mathbb{R}^p$ vérifiant $\nabla F(X) \neq 0$, on dit qu'un vecteur h de $\mathbb{R}^p$ est *une direction de décroissance* de F en X, si on a : $\langle \nabla F(X), h\rangle < 0$.

Dans les trois exemples suivants, on suppose que p est égal à 2.

1. *Un premier exemple.*
On considère les deux fonctions f_1 et f_2 définies sur $\mathbb{R}^2$ par : $f_1(x_1,x_2) = x_1^2 + x_2 + 1$, et $f_2(x_1,x_2) = x_1 + x_2^2 + 1$.

a) Justifier que F est de classe C^2 sur $\mathbb{R}^2$. Calculer, en tout point (x_1,x_2) de $\mathbb{R}^2$, le gradient $\nabla F(x_1,x_2)$.

b) Montrer que le système d'équations qui permet de déterminer les éventuels points critiques de F, peut se mettre sous la forme suivante :

$$\begin{cases} 2x_1^3 + 2x_1x_2 + 3x_1 + x_2^2 + 1 = 0 \\ (x_1 - x_2)(2x_1^2 + 2x_1x_2 + 2x_2^2 - x_1 - x_2 + 3) = 0 \end{cases}$$

c) Établir, pour tout (x_1,x_2) de $\mathbb{R}^2$, l'inégalité : $2x_1^2 + 2x_1x_2 + 2x_2^2 - x_1 - x_2 + 3 > 0$. En déduire que l'unique point critique de F est $(-1/2, -1/2)$.

d) Déterminer, en tout point (x_1,x_2) de $\mathbb{R}^2$, la matrice hessienne $\nabla^2 F(x_1,x_2)$. En déduire que F admet un minimum local en $(-1/2,-1/2)$.

e) On note pour tout point X de $\mathbb{R}^2$, $\nabla^2 f_1(X)$ et $\nabla^2 f_2(X)$ respectivement, les matrices hessiennes de f_1 et f_2 au point X. Préciser la matrice $J(X)$. Exprimer ${}^tJ(X)f(X)$ et $G(X) + \sum_{i=1}^{2} f_i(X)\nabla^2 f_i(X)$ en fonction de $\nabla F(X)$ et $\nabla^2 F(X)$ respectivement.

2. *Un deuxième exemple.*
Soit $a = (a_1, \dots, a_n)$, $b = (b_1, \dots, b_n)$ et $c = (c_1, \dots, c_n)$ trois vecteurs non nuls donnés de $\mathbb{R}^n$, tels que la famille (a,b) soit libre.
Pour tout i de $[\![1,n]\!]$, la fonction f_i est définie sur $\mathbb{R}^2$ par : $f_i(x_1,x_2) = a_ix_1 + b_ix_2 - c_i$.

a) Exprimer, pour tout point (x_1,x_2) de $\mathbb{R}^2$, le gradient $\nabla F(x_1,x_2)$ à l'aide de $x_1, x_2, ||a||, ||b||, \langle a,b\rangle, \langle a,c\rangle$ et $\langle b,c\rangle$.

b) Justifier l'inégalité : $||a||^2 \times ||b||^2 - \langle a,b\rangle^2 > 0$. En déduire que la fonction F possède un unique point critique $(\widehat{x_1}, \widehat{x_2})$.
Exprimer $\widehat{x_1}$ et $\widehat{x_2}$ en fonction de $||a||, ||b||, \langle a,b\rangle, \langle a,c\rangle$ et $\langle b,c\rangle$.

c) Calculer, en tout point (x_1,x_2) de $\mathbb{R}^2$, la matrice hessienne $\nabla^2 F(x_1,x_2)$; en déduire que F admet un minimum local en $(\widehat{x_1}, \widehat{x_2})$.

d) En utilisant la structure euclidienne de $\mathbb{R}^n$, montrer que F admet un minimum global en $(\widehat{x_1}, \widehat{x_2})$.

3. *Un troisième exemple.*
On suppose que $c_1, c_2, \dots, c_n$ sont n réels donnés non tous égaux. On note $\bar{c}$ et s^2 respectivement, la moyenne arithmétique et la variance de la série statistique $(c_i)_{1\leqslant i\leqslant n}$.
Pour tout i de $[\![1,n]\!]$, la fonction f_i est définie sur $\mathbb{R}^2$ par : $f_i(x_1,x_2) = x_1 + x_2 - c_i$.

a) Déterminer les points critiques de F.

b) Soit $(\widehat{x_1}, \widehat{x_2})$ un point critique de F. Exprimer $F(\widehat{x_1}, \widehat{x_2})$ en fonction de s^2. Montrer, pour tout (x_1,x_2) de $\mathbb{R}^2$, l'égalité : $F(x_1,x_2) - F(\widehat{x_1}, \widehat{x_2}) = \frac{n}{2}(x_1 + x_2 - \bar{c})^2$.

c) En déduire la nature des points critiques de F. Ce résultat était-il prévisible ?

4. *Retour au cas général.*
Soit $X = (x_1, \dots, x_p)$ un point de $\mathbb{R}^p$.

a) Exprimer $\nabla F(X)$ en fonction de ${}^tJ(X)$ et de $f(X)$.

b) Pour tout i de $[\![1,n]\!]$, on note $\nabla^2 f_i(X)$ la matrice hessienne de f_i au point X.
Établir la formule : $\nabla^2 F(X) = G(X) + \sum_{i=1}^{n} f_i(X)\nabla^2 f_i(X)$.

Partie II. Une approximation de F

Dans cette partie, on conserve les définitions et les notations de la partie I, et on suppose que X est un vecteur **fixé** de $\mathbb{R}^p$ vérifiant : $\nabla F(X) \neq 0$.
Pour tout vecteur $h = (h_1, h_2, \dots, h_p)$ de $\mathbb{R}^p$, on pose : $\ell(h) = f(X) + J(X)h$ et $L(h) = \frac{1}{2}||\ell(h)||^2$.

1. Établir, pour tout h de $\mathbb{R}^p$, l'égalité : $L(h) = F(X) + {}^th\nabla F(X) + \frac{1}{2}{}^th\, G(X)h$.

2. Soit P une matrice symétrique de $\mathcal{M}_p(\mathbb{R})$.
a) Justifier que P est diagonalisable.
b) On note $\theta_1, \dots, \theta_p$ les valeurs propres de P, et on pose : $\theta = \max\limits_{1\leqslant j\leqslant p} |\theta_j|$. Montrer, pour tout vecteur h de $\mathbb{R}^p$, l'inégalité suivante : $|{}^thPh| \leqslant \theta||h||^2$.

3. a) Écrire un développement limité à l'ordre 2 de la fonction F au point X.
b) En déduire, à l'aide de la question 2.b, que l'on a : $\lim\limits_{||h||\to 0} \frac{F(X+h) - L(h)}{||h||} = 0$.

Pour X fixé de $\mathbb{R}^p$, on dit que $L(h)$ est une approximation à l'ordre 2 de $F(X+h)$ lorsque $||h||$ tend vers 0.

4. On note : $G(X) = (g_{i,j}(X))_{1\leqslant i,j\leqslant p}$. Soit φ_1 et φ_2 deux fonctions définies sur $\mathbb{R}^p$ par : $\varphi_1(h) = {}^th\nabla F(X)$ et $\varphi_2(h) = {}^thG(X)h$.
a) Montrer que pour tout j de $[\![1,p]\!]$, on a : $\frac{\partial \varphi_1}{\partial h_j}(h) = \frac{\partial F}{\partial x_j}(X)$ et $\frac{\partial \varphi_2}{\partial h_j}(h) = 2\sum_{i=1}^{p} g_{i,j}(X)h_i$.
b) En déduire que le gradient $\nabla L(h)$ de L en h, est donné par : $\nabla L(h) = \nabla F(X) + G(X)h$.
c) Soit $\nabla^2 L(h)$ la matrice hessienne de L en h. Établir la formule : $\nabla^2 L(h) = G(X)$.

5. Soit J une matrice de $\mathcal{M}_{n,p}(\mathbb{R})$.
a) Montrer que la matrice ${}^tJJ$ est diagonalisable et que ses valeurs propres sont positives ou nulles.
b) Montrer que lorsque la matrice ${}^tJJ$ est inversible, le rang de la matrice J est égal à p.

6. Montrer que si la fonction L admet des points critiques $\widehat{h}$, alors ceux-ci vérifient l'inéquation : $\langle \widehat{h}, \nabla F(X)\rangle \leqslant 0$.

7. On suppose que la matrice $G(X)$ est inversible.
a) Montrer que L admet un unique point critique $\widehat{h}$ donné par : $\widehat{h} = -(G(X))^{-1} \times {}^tJ(X)f(X)$.
b) Établir que $\widehat{h}$ est une direction de décroissance de F en X. En déduire que L admet un minimum local en $\widehat{h}$.

Partie III. Une décomposition d'une matrice rectangulaire

Afin de réduire les inconvénients liés à l'inversion de la matrice $G(X)$, on remplace celle-ci par la matrice $G(X) + \mu I$, où μ désigne un paramètre réel strictement positif, et I la matrice identité d'ordre p. Certains résultats d'algèbre linéaire permettent alors de substituer à l'inversion d'une matrice, le calcul plus simple d'une somme de matrices.

Soit J une matrice non nulle de $\mathcal{M}_{n,p}(\mathbb{R})$.

1. Montrer qu'il existe une matrice V orthogonale de $\mathcal{M}_p(\mathbb{R})$, un entier q tel que $1 \leqslant q \leqslant p$, et des réels $\lambda_1, \lambda_2, \dots, \lambda_q$ tels que $\lambda_1 \geqslant \lambda_2 \geqslant \dots \geqslant \lambda_q > 0$, qui vérifient l'égalité : ${}^tV{}^tJJV = D$, où $D = (d_{i,j})_{1\leqslant i,j\leqslant p}$ est définie par : $d_{i,i} = \lambda_i$ si $1 \leqslant i \leqslant q$, et $d_{i,j} = 0$ sinon. Si $q < p$, on pose : $\lambda_{q+1} = \dots = \lambda_p = 0$.
Pour tout i de $[\![1,p]\!]$, on note V_i la i-ième colonne de V.

2. a) Montrer que le rang de ${}^tJJ$ est égal à q.

b) Montrer que, pour tout i de $[\![1,q]\!]$, JV_i est un vecteur propre de la matrice $J\,{}^tJ$ associé à la valeur propre λ_i. En déduire que les matrices ${}^tJJ$ et $J\,{}^tJ$ ont les mêmes valeurs propres non nulles.

c) Soit $(Y_1, \dots, Y_r)$ une base du sous-espace propre de ${}^tJJ$ associée à une valeur propre λ non nulle. Montrer que la famille $(JY_1, \dots, JY_r)$ est une famille libre de $\mathcal{M}_{n,1}(\mathbb{R})$.

d) En déduire que les sous-espaces propres de ${}^tJJ$ et de $J\,{}^tJ$ associés à la même valeur propre non nulle sont de même dimension, et que le rang de $J\,{}^tJ$ est égal à q.

3. On pose, pour tout i de $[\![1,q]\!]$: $U_i = \dfrac{1}{\sqrt{\lambda_i}}JV_i$.

a) Montrer que la famille $(U_1, \dots, U_q)$ est une famille orthonormée de vecteurs propres de $J\,{}^tJ$.

b) En déduire qu'il existe une base orthonormée $(U_1, \dots, U_q, U_{q+1}, \dots, U_n)$ de $\mathcal{M}_{n,1}(\mathbb{R})$, formée de vecteurs propres de $J\,{}^tJ$.

4. On note U la matrice de $\mathcal{M}_n(\mathbb{R})$ telle que, pour tout i de $[\![1,n]\!]$, la i-ième colonne de U est la matrice-colonne U_i de $\mathcal{M}_{n,1}(\mathbb{R})$.
Soit $S = (s_{i,j})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant p}}$ la matrice de $\mathcal{M}_{n,p}(\mathbb{R})$ définie par : $s_{i,i} = \sqrt{\lambda_i}$ si $1 \leqslant i \leqslant p$ et $s_{i,j} = 0$ sinon.
Établir l'égalité matricielle suivante : $S = {}^tUJV$. En déduire l'égalité : $J = US{}^tV$.

5. a) Montrer que la matrice $({}^tJJ + \mu I)$ est inversible.

b) On note $R = (r_{i,j})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant n}}$ la matrice de $\mathcal{M}_{p,n}(\mathbb{R})$ définie par : $r_{i,i} = \dfrac{\sqrt{\lambda_i}}{\lambda_i + \mu}$ si $1 \leqslant i \leqslant p$ et $r_{i,j} = 0$ sinon.
Établir la formule suivante : $({}^tJJ + \mu I)^{-1} \times {}^tJ = VR{}^tU$.

c) En déduire l'égalité : $({}^tJJ + \mu I)^{-1} \times {}^tJ = \displaystyle\sum_{i=1}^{q} \frac{\sqrt{\lambda_i}}{\lambda_i + \mu} V_i{}^tU_i$

6. Soit X un vecteur **fixé** de $\mathbb{R}^p$ vérifiant : $\nabla F(X) \neq 0$.
Pour tout vecteur h de $\mathbb{R}^p$, on pose : $M(h) = L(h) + \dfrac{\mu}{2}||h||^2$.

a) Montrer que : $\displaystyle\lim_{||h||\to 0} \frac{F(X+h) - M(h)}{||h||} = 0$.

b) Calculer, pour tout h de $\mathbb{R}^p$, le gradient $\nabla M(h)$ et la matrice hessienne $\nabla^2 M(h)$ de M en h.

c) En appliquant les résultats des questions précédentes à la matrice $J(X)$, montrer que M admet un unique point critique $h^\star$. Donner une expression de $h^\star$ qui utilise les résultats de la question 5.c.

d) Montrer que M admet un minimum local en $h^\star$.

À partir de ce minimum local $h^\star$ de M (ou du minimum local $\widehat{h}$ de L), on pourrait utiliser une méthode algorithmique permettant, sous certaines conditions, d'approcher avec une précision donnée un minimum local de la fonction F

Exemple 1

a)

Les applications f_1 et f_2 sont des applications polynomiales définies sur $\mathbb{R}^2$, elles sont donc de classe C^2 sur $\mathbb{R}^2$.

$F(x_1, x_2) = \frac{1}{2}(f_1^2(x_1, x_2) + f_2^2(x_1, x_2))$

$$\begin{aligned}\frac{\partial F}{\partial x_1}(x_1, x_2) &= \frac{\partial f_1}{\partial x_1} f_1(x_1, x_2) + \frac{\partial f_2}{\partial x_1} f_2(x_1, x_2)\\ &= 2x_1(x_1^2 + x_2 + 1) + (x_1 + x_2^2 + 1)\\ &= 2x_1^3 + 2x_1x_2 + 3x_1 + x_2^2 + 1\end{aligned}$$

$$\begin{aligned}\frac{\partial F}{\partial x_2}(x_1, x_2) &= \frac{\partial f_1}{\partial x_2} f_1(x_1, x_2) + \frac{\partial f_2}{\partial x_2} f_2(x_1, x_2)\\ &= (x_1^2 + x_2 + 1) + 2x_2(x_1 + x_2^2 + 1)\\ &= x_1^2 + 2x_1x_2 + 2x_2^3 + 3x_2 + 1\end{aligned}$$

$$\boxed{\nabla F(x_1, x_2) = (2x_1^3 + 2x_1x_2 + 3x_1 + x_2^2 + 1, x_1^2 + 2x_1x_2 + 2x_2^3 + 3x_2 + 1)}$$

b)

(x_1, x_2) est un point critique de F si et seulement si $\nabla F(x_1, x_2) = 0$

$$\nabla F(x_1, x_2) = 0 \iff \begin{cases} 2x_1^3 + 2x_1x_2 + 3x_1 + x_2^2 + 1 &= 0\\ x_1^2 + 2x_1x_2 + 2x_2^3 + 3x_2 + 1 &= 0 \quad L_2 \longleftarrow L_2 - L_1\end{cases}$$

$$\iff \begin{cases} 2x_1^3 + 2x_1x_2 + 3x_1 + x_2^2 + 1 &= 0\\ 2(x_2^3 - x_1^3) + x_1^2 - x_2^2 + 3(x_2 - x_1) &= 0\end{cases}$$

$$\iff \begin{cases} 2x_1^3 + 2x_1x_2 + 3x_1 + x_2^2 + 1 &= 0\\ (x_2 - x_1)(2(x_1^2 + x_1x_2 + x_2^2) - (x_1 + x_2) + 3 &= 0\end{cases}$$

$$\iff \begin{cases} 2x_1^3 + 2x_1x_2 + 3x_1 + x_2^2 + 1 &= 0 \quad \mathbf{(1)}\\ (x_2 - x_1)(2x_1^2 + 2x_1x_2 + 2x_2^2 - x_1 - x_2 + 3) &= 0 \quad \mathbf{(2)}\end{cases}$$

c)

Considérons $2x_1^2 + 2x_1x_2 + 2x_2^2 - x_1 - x_2 + 3 = 2x_1^2 + x_1(2x_2 - 1) + 2x_2^2 - x_2 + 3$ comme un trinôme en x_1. Son discriminant Δ vaut $(2x_2 - 1)^2 - 8(2x_2^2 - x_2 + 3) = -12x_2^2 + 4x_2 - 23$. Il est alors immédiat que le discriminant de ce nouveau trinôme est négatif ; ce trinôme reste strictement négatif , $\Delta < 0$, d'où $(2x_2 - 1)^2 - 8(2x_2^2 - x_2 + 3) = -12x_2^2 + 4x_2 - 23 = 2x_1^2 + 2x_1x_2 + 2x_2^2 - x_1 - x_2 + 3 > 0$

Le trinôme $2x_1^2 + x_1(2x_2 - 1) + 2x_2^2 - x_2 + 3$ n'a pas de racines réelles ; il est donc constamment du signe du coefficient de x_1^2 :

$$\forall (x_1, x_2] \in \mathbb{R}^2,\ 2x_1^2 + 2x_1x_2 + 2x_2^2 - x_1 - x_2 + 3 > 0.$$

L'équation **(2)** donne $x_1 = x_2$ et l'équation **(1)** s'écrit alors $2x_1^3 + 3x_1^2 + 3x_1 + 1 = 0$.

Etudions le polynôme P défini par $\forall x \in \mathbb{R},\ P(x) = 2x^3 + 3x^2 + 3x + 1$

$P'(x) = 3(2x^2 + 2x + 1)$; $\forall x \in \mathbb{R},\ P'(x) > 0$ car son discriminant $\delta = -4$.

On en déduit que P est une bijection de $\mathbb{R}$ dans $\mathbb{R}$ (car P est continue, strictement croissante).

Conclusion : P admet une unique racine réelle ; on constate facilement que $P(-\frac{1}{2}) = 0$

Finalement l'équation $\nabla F(x_1,x_2)=0$ admet une unique solution $(x_1,x_2)=(-\frac{1}{2},-\frac{1}{2})$

La fonction F admet un point critique et un seul $(-\frac{1}{2},-\frac{1}{2})$.

d)

La fonction F est de classe C^2 sur $\mathbb{R}^2$, on peut donc appliquer le théorème Schwarz.

$$\begin{aligned}\frac{\partial^2 F}{\partial x_1^2}(x_1,x_2) &= 6x_1^2+2x_2+3\\ \frac{\partial^2 F}{\partial x_1\partial x_2}=\frac{\partial^2 F}{\partial x_2\partial x_1}(x_1,x_2) &= 2x_1+2x_2\\ \frac{\partial^2 F}{\partial x_2^2}(x_1,x_2) &= 2x_1+6x_2^2+3\end{aligned}$$

On a alors $\nabla^2 F(x_1,x_2)=\begin{pmatrix}6x_1^2+2x_2+3 & 2x_1+2x_2\\ 2x_1+2x_2 & 2x_1+6x_2^2+3\end{pmatrix}$.

Au point critique, $\nabla^2 F(-\frac{1}{2},-\frac{1}{2})=\begin{pmatrix}\frac{7}{2} & -2\\ -2 & \frac{7}{2}\end{pmatrix}$

$(s^2-rt)(-\frac{1}{2},-\frac{1}{2})=4-\frac{49}{4}=-\frac{33}{4}<0$ et $r>0$.

La fonction F admet au point $(-\frac{1}{2},-\frac{1}{2})$ un minimum local

e)

- $J(X)=\begin{pmatrix}\frac{\partial f_1}{\partial x_1}(X) & \frac{\partial f_1}{\partial x_2}(X)\\ \frac{\partial f_2}{\partial x_1}(X) & \frac{\partial f_2}{\partial x_2}(X)\end{pmatrix}=\begin{pmatrix}2x_1 & 1\\ 1 & 2x_2\end{pmatrix}$

On remarque que ${}^tJ(X)=J(X)$.

$$\begin{aligned}{}^tJ(X)f(X) &= \begin{pmatrix}2x_1 & 1\\ 1 & 2x_2\end{pmatrix}\begin{pmatrix}f_1(X)\\ f_2(X)\end{pmatrix}\\ &= \begin{pmatrix}2x_1 & 1\\ 1 & 2x_2\end{pmatrix}\begin{pmatrix}x_1^2+x_2+1\\ x_1+x_2^2+1\end{pmatrix}\\ &= \begin{pmatrix}2x_1^3+2x_1x_2+3x_1+x_2^2+1\\ 2x_2^3+2x_1x_2+3x_2+x_1^2+1\end{pmatrix}\end{aligned}$$

On remarque que ${}^tJ(X)f(X)=\nabla F(X)$ puiqu'on a identifié $\mathbb{R}^2$ et $\mathcal{M}_{2,1}(\mathbb{R})$

- $G(X)={}^tJ(X)J(X)=\begin{pmatrix}4x_1^2+1 & 2x_1+2x_2\\ 2x_1+2x_2 & 4x_2^2+1\end{pmatrix}$

$\frac{\partial^2 f_1}{\partial x_1^2}(X)=2$; $\frac{\partial^2 f_1}{\partial x_1\partial x_2}=\frac{\partial^2 f_1}{\partial x_2\partial x_1}$ (d'après Schwarz) $=0$; $\frac{\partial^2 f_1}{\partial x_2^2}(X)=0$; donc

$$\nabla^2 f_1(X)=\begin{pmatrix}2 & 0\\ 0 & 0\end{pmatrix}$$

$\frac{\partial^2 f_2}{\partial x_1^2}(X)=0$; $\frac{\partial^2 f_2}{\partial x_1\partial x_2}=\frac{\partial^2 f_2}{\partial x_2\partial x_1}$ (d'après Schwarz) $=0$; $\frac{\partial^2 f_2}{\partial x_2^2}(X)=2$; donc

$$\nabla^2 f_1(X)=\begin{pmatrix}0 & 0\\ 0 & 2\end{pmatrix}$$

-

$$\begin{aligned}G(X)+\sum_{i=1}^{2}f_i(X)\nabla^2 f_i(X) &= \begin{pmatrix}4x_1^2+1 & 2x_1+2x_2\\ 2x_1+2x_2 & 4x_2^2+1\end{pmatrix}+(x_1^2+x_2+1)\begin{pmatrix}2 & 0\\ 0 & 0\end{pmatrix}\\ &\quad+(x_1+x_2^2+1)\begin{pmatrix}0 & 0\\ 0 & 2\end{pmatrix}\\ &= \begin{pmatrix}4x_1^2+1+2x_1^2+2x_2+2 & 2x_1+2x_2\\ 2x_1+2x_2 & 4x_2^2+1+2x_1+2x_2^2+2\end{pmatrix}\\ &= \begin{pmatrix}6x_1^2+2x_2+3 & 2x_1+2x_2\\ 2x_1+2x_2 & 6x_2^2+2x_1+3\end{pmatrix}\end{aligned}$$

$$\boxed{G(X)+\sum_{i=1}^{2}f_i(X)\nabla^2 f_i(X)=\nabla^2 F(X)}$$

Exemple 2

a)

$$\begin{aligned}F(x_1,x_2) &= \frac{1}{2}\sum_{i=1}^{n}(a_ix_1+b_ix_2-c_i)^2\\ \frac{\partial F}{\partial x_1}(x_1,x_2) &= \sum_{i=1}^{n}a_i(a_ix_1+b_ix_2-c_i)=x_1\sum_{i=1}^{n}a_i^2+x_2\sum_{i=1}^{n}a_ib_i-\sum_{i=1}^{n}a_ic_i\\ &= ||a||^2x_1+\langle a,b\rangle x_2-\langle a,c\rangle\\ \frac{\partial F}{\partial x_2}(x_1,x_2) &= \sum_{i=1}^{n}b_i(a_ix_1+b_ix_2-c_i)=x_1\sum_{i=1}^{n}a_ib_i+x_2\sum_{i=1}^{n}b_i^2-\sum_{i=1}^{n}b_ic_i\\ &= \langle a,b\rangle x_1+||b||^2x_2-\langle b,c\rangle\end{aligned}$$

$$\boxed{\nabla F(x_1,x_2)=(||a||^2x_1+\langle a,b\rangle x_2-\langle a,c\rangle,\langle a,b\rangle x_1+||b||^2x_2-\langle b,c\rangle)}$$

b)

On sait, d'après l'inégalité de Cauchy-Schwarz, que $|\langle a,b\rangle|\leq ||a||||b||$. On sait aussi qu'il y a égalité si et seulement si les vecteurs a et b sont liés. D'après les hypothèses, les vecteurs a et b sont libres, donc l'inégalité est stricte :

$$|\langle a,b\rangle|<||a||||b||,\text{ ce qui équivaut à }\langle a,b\rangle^2<||a||^2||b||^2$$

- Le couple (x_1,x_2) est un point critique de F si et seulement si $\nabla F(x_1,x_2)=0$

$$\nabla F(x_1,x_2)=0 \iff \begin{cases}||a||^2x_1+\langle a,b\rangle x_2 &= \langle a,c\rangle\\ \langle a,b\rangle x_1+||b||^2x_2 &= \langle b,c\rangle \quad L_2\longleftarrow ||a||^2L_2-\langle a,b\rangle L_1\end{cases}$$

opération permise car $||a||^2\neq 0$

$$\iff \begin{cases}||a||^2x_1+\langle a,b\rangle x_2 &= \langle a,c\rangle\\ (||a||^2||b||^2-\langle a,b\rangle^2)x_2 &= ||a||^2\langle b,c\rangle-\langle a,b\rangle\langle a,c\rangle\end{cases}$$

D'après le point précédent, $(||a||^2||b||^2-\langle a,b\rangle^2)>0$.

Le système est de Cramer : il admet donc une unique solution notée $(\widehat{x_1},\widehat{x_2})$

- Calcul de la solution

On a tout de suite $\widehat{x_2}=\frac{||a||^2\langle b,c\rangle-\langle a,b\rangle\langle a,c\rangle}{||a||^2||b||^2-\langle a,b\rangle^2}$.

La première ligne du système va donner $\widehat{x_1}$

$$\widehat{x_1} = \frac{1}{||a||^2}\left(\langle a,c\rangle - \langle a,b\rangle \frac{||a||^2\langle b,c\rangle - \langle a,b\rangle\langle a,c\rangle}{||a||^2||b||^2 - \langle a,b\rangle^2}\right)$$

opération permise car $||a||^2 \neq 0$

$$= \frac{1}{||a||^2}\frac{1}{||a||^2||b||^2 - \langle a,b\rangle^2}\left(\langle a,c\rangle||a||^2||b||^2 - \langle a,c\rangle\langle a,b\rangle^2 - ||a||^2\langle a,b\rangle\langle b,c\rangle + \langle a,b\rangle^2\langle a,c\rangle\right)$$

$$= \frac{1}{||a||^2}\frac{1}{||a||^2||b||^2 - \langle a,b\rangle^2}\left(||a||^2\left(||b||^2\langle a,c\rangle - \langle a,b\rangle\langle b,c\rangle\right)\right)$$

D'où les résultats :

$$\boxed{\widehat{x_1} = \frac{||b||^2\langle a,c\rangle - \langle a,b\rangle\langle b,c\rangle}{||a||^2||b||^2 - \langle a,b\rangle^2} \text{ et } \widehat{x_2} = \frac{||a||^2\langle b,c\rangle - \langle a,b\rangle\langle a,c\rangle}{||a||^2||b||^2 - \langle a,b\rangle^2}}$$

c)

$$\nabla^2 F(x_1,x_2) = \begin{pmatrix} \frac{\partial^2 F}{\partial x_1^2}(x_1,x_2) & \frac{\partial^2 F}{\partial x_1 \partial x_2}(x_1,x_2) \\ \frac{\partial^2 F}{\partial x_2 \partial x_1}(x_1,x_2) & \frac{\partial^2 F}{\partial x_2^2}(x_1,x_2) \end{pmatrix}.$$

$$= \begin{pmatrix} ||a||^2 & \langle a,b\rangle \\ \langle a,b\rangle & ||b||^2 \end{pmatrix}$$

$$(s^2 - rt)(\widehat{x_1},\widehat{x_2}) = \langle a,b\rangle^2 - ||a||^2||b||^2 < 0$$

$$\boxed{(s^2-rt)(\widehat{x_1},\widehat{x_2}) < 0 \text{ et } r(\widehat{x_1},\widehat{x_2}) > 0 : F \text{ possède en } (\widehat{x_1},\widehat{x_2}) \text{ un minimum local}}$$

d)

Considérons le vecteur $Y = (Y_1, \ldots, Y_i, \ldots, Y_n)$ de $\mathbb{R}^n$ où $\forall i \in [\![1,n]\!]$, $Y_i = (a_i x_1 + b_i x_2 - c_i)$. On constate qu'alors $F(X) = \frac{1}{2}||Y||^2$.

Si l'on considère la matrice $A \in \mathcal{M}_{n,2}(\mathbb{R})$ suivante : $A = \begin{pmatrix} a_1 & b_1 \\ \vdots & \vdots \\ a_i & b_i \\ \vdots & \vdots \\ a_n & b_n \end{pmatrix}$ et $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathcal{M}_{2,1}(\mathbb{R})$,

on a immédiatement $AX = \begin{pmatrix} a_1x_1 + b_1x_2 \\ \vdots \\ a_ix_1 + b_ix_2 \\ \vdots \\ a_nx_1 + b_nx_2 \end{pmatrix}$

Introduisons maintenant le vecteur $C = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbb{R})$, on obtient

$$\begin{pmatrix} a_1x_1 + b_1x_2 - c_1 \\ \vdots \\ a_ix_1 + b_ix_2 - c_i \\ \vdots \\ a_nx_1 + b_nx_2 - c_n \end{pmatrix} = AX - C.$$

$$\boxed{F(X) = \frac{1}{2}||AX - C||^2}$$

- La matrice A est de rang 2 puisque les vecteurs a et b sont libres. D'après le cours sur les moindres carrés, $||AX - C||$ sera minimale pour l'unique vecteur $\widehat{X}$ de $\mathbb{R}^2$ donné par : $\widehat{X} = ({}^tAA)^{-1} \times {}^tAC$.

Vérifions que $\widehat{X} = (\widehat{x_1}, \widehat{x_2})$: cela prouvera que le minimum est global.

$${}^tAA = \begin{pmatrix} a_1 & \ldots & a_n \\ b_1 & \ldots & b_n \end{pmatrix}\begin{pmatrix} a_1 & b_1 \\ \vdots & \vdots \\ a_n & b_n \end{pmatrix} = \begin{pmatrix} ||a||^2 & \langle a,b\rangle \\ \langle a,b\rangle & ||b||^2 \end{pmatrix}$$

- Calcul de $({}^tAA)^{-1}$ (on sait que cette matrice existe car A, donc ${}^tA$ sont inversibles) . Pour cela résolvons le système : ${}^tAA\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ d'inconnues x et y.

$${}^tAA\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \iff \begin{cases} ||a||^2 x + \langle a,b\rangle y &= \alpha \\ \langle a,b\rangle x + ||b||^2 y &= \beta \end{cases}$$

On a déjà résolu ce système à la question **b)** : il suffit de changer $\langle a,c\rangle$ en α, $\langle b,c\rangle$ en β et x_1 en x, x_2 en y : les formules

$x_1 = \dfrac{||b||^2\langle a,c\rangle - \langle a,b\rangle\langle b,c\rangle}{||a||^2||b||^2 - \langle a,b\rangle^2}$ et $x_2 = \dfrac{||a||^2\langle b,c\rangle - \langle a,b\rangle\langle a,c\rangle}{||a||^2||b||^2 - \langle a,b\rangle^2}$ donnent

$x = \dfrac{||b||^2\alpha - \langle a,b\rangle\beta}{||a||^2||b||^2 - \langle a,b\rangle^2}$ et $y = \dfrac{||a||^2\beta - \langle a,b\rangle\alpha}{||a||^2||b||^2 - \langle a,b\rangle^2}$

Matriciellement

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{||a||^2||b||^2 - \langle a,b\rangle^2}\begin{pmatrix} ||b||^2 & -\langle a,b\rangle \\ -\langle a,b\rangle & ||a||^2 \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

$$({}^tAA)^{-1} = \frac{1}{||a||^2||b||^2 - \langle a,b\rangle^2}\begin{pmatrix} ||b||^2 & -\langle a,b\rangle \\ -\langle a,b\rangle & ||a||^2 \end{pmatrix}$$

$${}^tAC = \begin{pmatrix} a_1 & \ldots & a_n \\ b_1 & \ldots & b_n \end{pmatrix}\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} \langle a,c\rangle \\ \langle b,c\rangle \end{pmatrix}$$

$$\widehat{X} = \frac{1}{||a||^2||b||^2 - \langle a,b\rangle^2}\begin{pmatrix} ||b||^2 & -\langle a,b\rangle \\ -\langle a,b\rangle & ||a||^2 \end{pmatrix}\begin{pmatrix} \langle a,c\rangle \\ \langle b,c\rangle \end{pmatrix}$$

$$= \frac{1}{||a||^2||b||^2 - \langle a,b\rangle^2}\begin{pmatrix} ||b||^2\langle a,c\rangle - \langle a,b\rangle\langle b,c\rangle \\ ||a||^2\langle b,c\rangle - \langle a,b\rangle\langle a,c\rangle \end{pmatrix}$$

$$\boxed{\widehat{X} = \begin{pmatrix} \widehat{x_1} \\ \widehat{x_2} \end{pmatrix} : F \text{ admet en } (\widehat{x_1},\widehat{x_2}) \text{ un minimum global}}$$

Exemple 3

a)

$$F(X) = \frac{1}{2}\sum_{i=1}^{n}(x_1 + x_2 - c_i)^2$$

$$\frac{\partial F}{\partial x_1}(X) = \sum_{i=1}^{n}(x_1 + x_2 - c_i) = n(x_1 + x_2 - \overline{c}) \quad \text{car } \overline{c} = \frac{1}{n}\sum_{i=1}^{n} c_i$$

$$\frac{\partial F}{\partial x_2}(X) = \sum_{i=1}^{n}(x_1 + x_2 - c_i) = n(x_1 + x_2 - \overline{c})$$

$$\boxed{\begin{array}{l} \nabla F(X) = (n(x_1 + x_2 - \overline{c}), n(x_1 + x_2 - \overline{c})) \\ \text{Les points critiques de } F \text{ vérifient : } x_1 + x_2 = \overline{c} \end{array}}$$

b)

Soit $(\widehat{x_1},\widehat{x_2})$ un point critique de F.

$F(\widehat{x_1},\widehat{x_2}) = \frac{1}{2}\sum_{i=1}^{n}(\widehat{x_1}+\widehat{x_2}-c_i)^2 = \frac{1}{2}\sum_{i=1}^{n}(\overline{c}-c_i)^2$. Or par définition, $s^2 = \frac{1}{n}\sum_{i=1}^{n}(\overline{c}-c_i)^2$. Donc

$$\boxed{F(\widehat{x_1},\widehat{x_2}) = \frac{ns^2}{2}}$$

- $$\begin{aligned} F(x_1,x_2)-F(\widehat{x_1},\widehat{x_2}) &= \frac{1}{2}\sum_{i=1}^{n}(x_1+x_2-c_i)^2 - \frac{ns^2}{2} \\ &= \frac{1}{2}\sum_{i=1}^{n}(x_1+x_2-\overline{c}+\overline{c}-c_i)^2 - \frac{ns^2}{2} \\ &= \frac{1}{2}\sum_{i=1}^{n}(x_1+x_2-\overline{c})^2 + \frac{1}{2}\sum_{i=1}^{n}(\overline{c}-c_i)^2 + \frac{2}{2}\sum_{i=1}^{n}(x_1+x_2-\overline{c})(\overline{c}-c_i) - \frac{ns^2}{2} \\ &= \frac{n}{2}(x_1+x_2-\overline{c})^2 + \frac{ns^2}{2} + (x_1+x_2-\overline{c})\sum_{i=1}^{n}(\overline{c}-c_i) - \frac{ns^2}{2} \\ &= \frac{n}{2}(x_1+x_2-\overline{c})^2 + (x_1+x_2-\overline{c})\underbrace{\left(n\overline{c}-\sum_{i=1}^{n}c_i\right)}_{=0} \end{aligned}$$

$$\boxed{F(x_1,x_2)-F(\widehat{x_1},\widehat{x_2}) = \frac{n}{2}(x_1+x_2-\overline{c})^2 \geq 0}$$

c)

On vient de voir que $\forall(x_1,x_2)\in\mathbb{R}^2,\ F(x_1,x_2)-F(\widehat{x_1},\widehat{x_2})\geq 0$

$$\boxed{\text{Aux points critiques , } F \text{ orésente des minima globaux}}$$

Il était prévisible que F ne pouvait pas présenter de maximum globaux car $\lim_{x_1\to+\infty} F(x_1,x_2) = +\infty$.

Cas général

a)

Par définition, $\nabla F(X) = \left(\frac{\partial F}{\partial x_1}(X),\ldots,\frac{\partial F}{\partial x_p}(X))\right)$.

$$\begin{aligned} F(X) &= \frac{1}{2}\sum_{i=1}^{n} f_i^2(X) \\ \forall j\in[\![1,p]\!],\ \frac{\partial F}{\partial x_i}(X) &= \sum_{j=1}^{n}\frac{\partial f_j}{\partial x_i}(X)f_j(X) \end{aligned}$$

Or $J(X) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(X) & \cdots & \frac{\partial f_1}{\partial x_j}(X) & \cdots & \frac{\partial f_1}{\partial x_p}(X) \\ \vdots & & \vdots & & \vdots \\ \frac{\partial f_i}{\partial x_1}(X) & \cdots & \frac{\partial f_i}{\partial x_j}(X) & \cdots & \frac{\partial f_i}{\partial x_p}(X) \\ \vdots & & \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(X) & \cdots & \frac{\partial f_n}{\partial x_j}(X) & \cdots & \frac{\partial f_n}{\partial x_p}(X) \end{pmatrix}$, donc

$${}^tJ(X) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(X) & \cdots & \frac{\partial f_j}{\partial x_1}(X) & \cdots & \frac{\partial f_n}{\partial x_1}(X) \\ \vdots & & \vdots & & \vdots \\ \frac{\partial f_1}{\partial x_i}(X) & \cdots & \frac{\partial f_j}{\partial x_i}(X) & \cdots & \frac{\partial f_n}{\partial x_i}(X) \\ \vdots & & \vdots & & \vdots \\ \frac{\partial f_1}{\partial x_p}(X) & \cdots & \frac{\partial f_j}{\partial x_p}(X) & \cdots & \frac{\partial f_n}{\partial x_p}(X) \end{pmatrix}$$

$${}^tJ(X)f(X) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(X) & \cdots & \frac{\partial f_j}{\partial x_1}(X) & \cdots & \frac{\partial f_n}{\partial x_1}(X) \\ \vdots & & \vdots & & \vdots \\ \frac{\partial f_1}{\partial x_i}(X) & \cdots & \frac{\partial f_j}{\partial x_i}(X) & \cdots & \frac{\partial f_n}{\partial x_i}(X) \\ \vdots & & \vdots & & \vdots \\ \frac{\partial f_1}{\partial x_p}(X) & \cdots & \frac{\partial f_j}{\partial x_p}(X) & \cdots & \frac{\partial f_n}{\partial x_p}(X) \end{pmatrix} \begin{pmatrix} f_1(X) \\ \vdots \\ f_j(X) \\ \vdots \\ f_n(X) \end{pmatrix}$$

${}^tJ(X)f(X)$ est une colonne dont la i ème ligne est $\sum_{j=1}^{n}\frac{\partial f_j}{\partial x_i}(X)f_j(X)$. Comme on assimile les colonnes aux listes, on peut dire que ${}^tJ(X)f(X)$ est une $n-$liste dont le i ème terme est $\sum_{j=1}^{n}\frac{\partial f_j}{\partial x_i}(X)f_j(X)$. On reconnaît le i ème terme de $\nabla F(X)$

$$\boxed{\nabla F(X) = {}^tJ(X)f(X)}$$

b)

Par définition, $\forall X\in\mathbb{R}^p,\ \nabla^2F(X) = \left(\frac{\partial^2F(X)}{\partial x_k\partial x_j}\right)\in\mathcal{M}_p(\mathbb{R})$. Cette matrice est symétrique d'après le théorème de Schwarz puisque F est de classe C^2 sur $\mathbb{R}^p$. Explicitons un peu :

La k ème ligne de $\nabla^2F(X)$ est $\left(\frac{\partial^2F(X)}{\partial x_k\partial x_1},\ldots,\frac{\partial^2F(X)}{\partial x_k\partial x_j},\ldots,\frac{\partial^2F(X)}{\partial x_k\partial x_p}\right)$

D'autre part, pour tout $j\in[\![1;p]\!]$

$\frac{\partial F(X)}{\partial x_j} = \sum_{i=1}^{n}\frac{\partial f_i(X)}{\partial x_j}f_i(X)$, donc $\forall(k,j)\in(\ [\![1;p]\!]\ )^2$,

$$\frac{\partial^2F(X)}{\partial x_k\partial x_j} = \sum_{i=1}^{n}\frac{\partial f_i(X)}{\partial x_j}\frac{\partial f_i(X)}{\partial x_k} + \sum_{i=1}^{n}\frac{\partial^2 f_i(X)}{\partial x_k\partial x_j}f_i(X) \qquad \mathbf{(1)}$$

- $\forall i\in[\![1;n]\!],\ \nabla^2 f_i(X) = \begin{pmatrix} \frac{\partial^2 f_i(X)}{\partial x_1^2} & \cdots & \cdots & \frac{\partial^2 f_i(X)}{\partial x_1\partial x_p} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ \frac{\partial^2 f_i(X)}{\partial x_p\partial x_1} & \cdots & \cdots & \frac{\partial^2 f_i(X)}{\partial x_p^2} \end{pmatrix} \in\mathcal{M}_p(\mathbb{R})$

$\forall(k,j)\in(\ [\![1;p]\!]\ )^2$, $\frac{\partial^2F(X)}{\partial x_k\partial x_j}$ est le terme général de $\nabla^2F(X)$

$\sum_{i=1}^{n}\frac{\partial f_i(X)}{\partial x_j}\frac{\partial f_i(X)}{\partial x_k}$ est le terme général de ${}^tJ(X)J(X)$

$\frac{\partial^2 f_i(X)}{\partial x_k\partial x_j}f_i(X)$ est le terme général de $f_i(X)\nabla^2 f_i(X)$, donc, d'après l'égalité **(1)**, le terme général de $\nabla^2(F(X)$ est égal au terme général de ${}^tJ(X)J(X)$ + la somme, pour $i \in [\![1;n]\!]$, du terme général de $f_i(X)\nabla^2 f_i(X)$

$$\boxed{\text{Conclusion : } \nabla^2(F(X) = {}^tJ(X)J(X) + \sum_{i=1}^{n} f_i(X)\nabla^2 f_i(X) = G(X) + \sum_{i=1}^{n} f_i(X)\nabla^2 f_i(X)}$$

Partie II

1)

$$\begin{aligned} L(h) &= \frac{1}{2}||l(h)||^2 = \frac{1}{2}{}^tl(h)l(h) \\ &= \frac{1}{2}{}^t(f(X)+J(X)h)(f(X)+J(X)h) \\ &= \frac{1}{2}({}^tf(X)+{}^th{}^tJ(X))(f(X)+J(X)h) \\ &= \frac{1}{2}({}^tf(X)f(X)+{}^tf(X)J(X)h+{}^th{}^tJ(X)f(X)+{}^th{}^tJ(X)J(X)h) \\ &= \frac{1}{2}(||f(X)||^2+{}^tf(X)J(X)h+{}^th{}^tJ(X)f(X)+{}^thG(X)h) \end{aligned}$$

$J(X) \in \mathcal{M}_{n,p}(\mathbb{R})$, $h \in \mathcal{M}_{p,1}(\mathbb{R})$, donc $J(X)h \in \mathcal{M}_{n,1}(\mathbb{R})$
$f(X) = (f_i(X), \dots, f_n(X)) \in \mathbb{R}^n$ que l'on assimile à $\mathcal{M}_{n,1}(\mathbb{R})$; il en résulte que ${}^tf(X)J(X)h \in \mathcal{M}_1(\mathbb{R})$, donc ${}^tf(X)J(X)h \in \mathbb{R}$.

Même raisonnement pour ${}^th{}^tJ(X)f(X)$. Donc ${}^th{}^tJ(X)f(X) = {}^t\Big({}^th{}^tJ(X)f(X)\Big) = {}^tf(X)J(X)h$

Il en résulte que $L(h) = \frac{1}{2}(||f(X)||^2 + 2{}^th{}^tJ(X)f(X) + {}^thG(X)h)$

D'après I–4–b), ${}^tJ(X)f(X) = \nabla F(X)$ et on sait que $\frac{1}{2}||f(X)||^2 = F(X)$, donc

$$\boxed{\begin{aligned} L(h) &= F(X) + {}^th{}^tJ(X)f(X) + \frac{1}{2}{}^thG(X)h \\ &= F(X) + {}^th\nabla F(X) + \frac{1}{2}{}^thG(X)h \qquad \text{d'après I–4–a)} \end{aligned}}$$

2–a)

La matrice P est symétrique réelle de $\mathcal{M}_p(\mathbb{R})$, elle est donc diagonalisable en base orthonormée.

2–b)

Il existe donc une matrice orthogonale Q appartenant à $\mathcal{M}_p(\mathbb{R})$, il existe une matrice diagonale D appartenant à $\mathcal{M}_p(\mathbb{R})$ telles que $P = QDQ^{-1} = QD{}^tQ$ (puisque Q est orthogonale).
${}^thPh = {}^thQD{}^tQh = {}^t({}^tQh)D{}^tQh$.
Posons $Y = {}^tQh$; Y appartient à $\mathcal{M}_{p,1}(\mathbb{R})$. On peut écrire $Y = (y_1, \dots, y_p)$ en identifiant, comme le dit l'énoncé, les éléments de $\mathbb{R}^p$, les matrices colonnes de $\mathcal{M}_{p,1}(\mathbb{R})$ et les matrices lignes de $\mathcal{M}_{1,p}(\mathbb{R})$.

L'égalité précedente devient ${}^thPh = {}^tYDY = \displaystyle\sum_{k=1}^{p} \theta_k y_k^2$.

L'inégalité triangulaire donne $|{}^thPh| \leq \displaystyle\sum_{k=1}^{p} |\theta_k| y_k^2$

Or par hypothèse $\forall k \in [\![1;p]\!]$, $|\theta_k| \leq \theta$. On multiplie ces inégalités par $y_k^2 \geq 0$ et on les ajoute ; il vient $\displaystyle\sum_{k=1}^{p} |\theta_k| y_k^2 \leq \sum_{k=1}^{p} \theta y_k^2 \leq \theta \sum_{k=1}^{p} y_k^2$, soit finalement

$$\boxed{|{}^thPh| \leq \theta||Y||^2}$$

3–a)

F est de classe C^2 sur $\mathbb{R}^p$, son développement à l'ordre 2 existe et on a $\forall h \in \mathbb{R}^p$,
$F(X+h) = F(X) + \langle \nabla F(X), h\rangle + \frac{1}{2}q_X(h) + ||h||^2\varepsilon(h)$ avec $\lim\limits_{h\to 0} \varepsilon(h) = 0$ et où q_X est la forme quadratique associée à la hessienne de F au point X.

3–b)

$\langle \nabla F(X), h\rangle = {}^th\nabla F(X)$ et $q_X(h) = {}^th\nabla^2 F(X)h$.

D'après I–4–b), $q_X(h) = {}^th\Big(G(X) + \displaystyle\sum_{i=1}^{n} f_i(X)\nabla^2 f_i(X)\Big)h$, donc

$$\begin{aligned} F(X+h) &= F(X) + {}^th\nabla F(X) + \frac{1}{2}{}^th\Big(G(X) + \sum_{i=1}^{n} f_i(X)\nabla^2 f_i(X)\Big)h + ||h||^2\varepsilon(h) \\ &= \underbrace{F(X) + {}^th\nabla F(X) + \frac{1}{2}{}^thG(X)h}_{=L(h)} + \frac{1}{2}{}^th\Big(\sum_{i=1}^{n} f_i(X)\nabla^2 f_i(X)\Big)h + ||h||^2\varepsilon(h) \end{aligned}$$

$$F(X+h) = L(h) + \frac{1}{2}{}^th\Big(\sum_{i=1}^{n} f_i(X)\nabla^2 f_i(X)\Big)h + ||h||^2\varepsilon(h) \qquad \textbf{(2)}$$

Pour tout indice $i \in [\![1;n]\!]$, la matrice $f_i(X)\nabla^2 f_i(X)$ est symétrique réelle à cause du théorème de Schwarz, donc la matrice $\frac{1}{2}\Big(\displaystyle\sum_{i=1}^{n} f_i(X)\nabla^2 f_i(X)\Big)$ est aussi symétrique réelle.

Posons $P = \frac{1}{2}\Big(\displaystyle\sum_{i=1}^{n} f_i(X)\nabla^2 f_i(X)\Big)$; l'égalité **(2)** s'écrit

$$F(X+h) = L(h) + {}^thPh + ||h||^2\varepsilon(h)$$

$$\begin{aligned} \forall h \neq 0,\ \left|\frac{F(X+h)-L(h)}{||h||}\right| &= \left|\frac{{}^thPh + ||h||^2\varepsilon(h)}{||h||}\right| \\ &\leq \frac{1}{||h||}\Big(|{}^thPh| + ||h||^2|\varepsilon(h)|\Big) \\ &\leq \frac{1}{||h||}|{}^thPh| + ||h|||\varepsilon(h)| \\ &\leq \frac{\theta||h||^2}{||h||} + ||h|||\varepsilon(h)| \qquad \text{d'après 2–b)} \\ &\leq \theta||h|| + ||h|||\varepsilon(h)| \end{aligned}$$

$\lim\limits_{h\to 0}(\theta||h|| + ||h|||\varepsilon(h)|) = 0$ donc par encadrement

$$\boxed{\lim_{h\to 0}\left|\frac{F(X+h)-L(h)}{||h||}\right| = 0}$$

4–a)

$\varphi_1(h) = {}^th\nabla F(X) = \displaystyle\sum_{i=1}^{p} h_i \frac{\partial F}{\partial x_i}(X)$; donc $\dfrac{\partial \varphi_1}{\partial h_j}(h) = \dfrac{\partial F}{\partial x_j}(X)$

$\varphi_2(h) = \displaystyle\sum_{i=1}^{p} g_{i,i}(X)h_i^2 + 2\sum_{1\leq i<k\leq p} g_{i,k}(X)h_ih_k$ d'après la formule classique du développement d'une forme quadratique lorsque l'on connait sa matrice.

Soit $j \in [\![1;p]\!]$; réécrivons la deuxième somme de ce développement en faisant apparaître les termes en h_j

$$\sum_{1\le i<k\le p} g_{i,k}(X)h_ih_k = \sum_{1\le i<j\le p} g_{i,j}(X)h_ih_j + \sum_{1\le j<k\le p} g_{j,k}(X)h_jh_k + \sum_{\substack{1\le i<k\le p\\ i\ne j,k\ne j}} g_{i,k}(X)h_ih_k$$

$$\frac{\partial\varphi_2(h)}{\partial h_j} = 2g_{j,j}(X)h_j + 2\sum_{1\le i<j\le p} g_{i,j}(X)h_i + 2\sum_{1\le j<i\le p} g_{j,i}(X)h_i\ ;$$

la dernière somme vaut $2\sum\limits_{1\le j<i\le p} g_{i,j}(X)h_i$ car G est symétrique, donc

$$\frac{\partial\varphi_2(h)}{\partial h_j} = 2\sum_{i=1}^{p} g_{i,j}(X)h_i$$

4–b)

On a $L(h) = F(X) + \varphi_1(h) + \frac{1}{2}\varphi_2(h)$; par linéarité du gradient, $\nabla L(h) = \nabla\varphi_1(h) + \frac{1}{2}\nabla\varphi_2(h)$.

Or $\nabla\varphi_1(h) = \nabla F(X)$ d'après a).

$$\begin{aligned}\frac{1}{2}\nabla\varphi_2(h) &= \left(\frac{\partial\varphi_2}{\partial h_1}(X),\dots,\frac{\partial\varphi_2}{\partial h_j}(X),\dots,\frac{\partial\varphi_2}{\partial h_p}(X)\right)\\ &= \left(\sum_{i=1}^{p} g_{i,1}(X)h_i,\dots,\sum_{i=1}^{p} g_{i,j}(X)h_i,\dots,\sum_{i=1}^{p} g_{i,p}(X)h_i\right)\\ &= \begin{pmatrix}\sum_{i=1}^{p} g_{i,1}(X)h_i\\ \vdots\\ \sum_{i=1}^{p} g_{i,j}(X)h_i\\ \vdots\\ \sum_{i=1}^{p} g_{i,p}(X)h_i\end{pmatrix}\end{aligned}$$

en identifiant les lignes aux colonnes

$$\frac{1}{2}\nabla\varphi_2(h) = G(X)h$$

$$\boxed{\text{Finalement, } \nabla L(h) = \nabla F(X) + G(X)h}$$

5–a)

La matrice ${}^tJJ$ est une matrice symétrique réelle appartenant à $\mathcal{M}_p(\mathbb{R})$. Elle est donc diagonalisable en base orthonormée. Soit λ une valeur propre de ${}^tJJ$ et $Y \in \mathcal{M}_{p,1}(\mathbb{R})$ un vecteur propre associé.

${}^tJJY = \lambda Y \Longrightarrow {}^tY{}^tJJY = \lambda{}^tYY$, ce qui s'exprime aussi ${}^t(JY)JY = \lambda{}^tYY$, ou encore $||JY||^2 = \lambda||Y||^2$

$Y \ne 0 \Longrightarrow ||Y||^2 > 0$, on a alors $\lambda = \dfrac{||JY||^2}{||Y||^2}$, donc $\boxed{\lambda \ge 0}$

5–b)

Il s'agit de montrer que $\operatorname{Ker}{}^tJJ = \{0\} \Longrightarrow \operatorname{Ker} J = \{0\}$.

Montrons d'une manière générale que $\operatorname{Ker}{}^tJJ = \operatorname{Ker} J$

• Soit $Y \in \operatorname{Ker}{}^tJJ$: $Y \in \mathcal{M}_{p,1}(\mathbb{R})$ et ${}^tJJY = 0$. Cela implique ${}^tY{}^tJJY = 0$, ou encore $||JY||^2 = 0$. On a $JY = 0$, donc $Y \in \operatorname{Ker} J$.

• Soit $Y \in \operatorname{Ker} J$: $Y \in \mathcal{M}_{p,1}(\mathbb{R})$ et $JY = 0$. On multiplie par ${}^tJ$ à gauche et on a ${}^tJJY = 0$: $Y \in \operatorname{Ker}{}^tJJ$.

On a bien l'égalité $\operatorname{Ker}{}^tJJ = \operatorname{Ker} J$.

$$\boxed{\text{Donc si } {}^tJJ \text{ est inversible, son noyau est nul, donc celui de } J \text{ aussi et le rang de } J \text{ égal } p}$$

6)

Si $\widehat{h}$ est un point critique de L, alors $\nabla L(\widehat{h}) = 0$.

D'après 4–b), on a alors : $\nabla F(X) + G(X)\widehat{h} = 0$, soit $G(X)\widehat{h} = -\nabla F(X)$; donc ${}^t\widehat{h}G(X)\widehat{h} = -{}^t\widehat{h}\nabla F(X) = -\langle\widehat{h}, \nabla F(X)\rangle$.

Ceci s'écrit aussi ${}^t\widehat{h}{}^tJJ\widehat{h} = -\langle\widehat{h}, \nabla F(X)\rangle$, et finalement $||J(X)\widehat{h}||^2 = -\langle\widehat{h}, \nabla F(X)\rangle$.

$$\boxed{\text{On a immédiatement } \langle\widehat{h}, \nabla F(X)\rangle \le 0}$$

7–a)

Si la matrice $G(X)$ est inversible, alors l'équation $G(X)\widehat{h} = -\nabla F(X)$ admet une unique solution : $\widehat{h} = -(G(X))^{-1}\nabla F(X)$. Or d'après 4–b), $\nabla F(X) = {}^tJ(X)f(X)$, donc

$$\boxed{\widehat{h} = -(G(X))^{-1} \times {}^tJ(X)f(X)}$$

7–b)

Montrons que $J(X)\widehat{h} \ne 0$ pour avoir $\langle\widehat{h}, \nabla F(X)\rangle < 0$

Si $J(X)\widehat{h} = 0$, alors $\widehat{h} \in \operatorname{Ker} J(X)$. Donc d'après 5–b), $\widehat{h} \in \operatorname{Ker}{}^tJ(X)J(X)$, c'est-à-dire $\widehat{h} \in \operatorname{Ker} G(X)$. Par hypothèse, $G(X)$ est inversible, donc $\operatorname{Ker} G(X) = \{0\}$. On aurait donc $\widehat{h} = 0$.

D'après 4–b), $\nabla L(0) = \nabla F(X) \ne 0$. Cela est contradictoire.

Conclusion : $J(X)\widehat{h} \ne 0$, donc $||J(X)\widehat{h}|| > 0$ et par conséquent $\langle\widehat{h}, \nabla F(X) < 0\rangle = -||J(X)\widehat{h}||^2 < 0$ d'après 6).

$$\boxed{\widehat{h} \text{ est une direction de décroissance de } F(X)}$$

Soit Y un vecteur non nul de $\mathbb{R}^p$.

${}^tY\nabla^2L(\widehat{h})Y = {}^tYG(X)Y$ (d'après 4–c)), donc ${}^tY\nabla^2L(\widehat{h})Y = ||J(X)Y||^2$ (car $G(X) = {}^tJ(X)J(X)$.

Or $G(X)$ inversible implique $J(X)$ inversible et $J(X)$ inversible et $Y \ne 0$ impliquent $J(X)h \ne 0$

$$\boxed{\begin{array}{l}\text{Donc } \forall Y \in \mathbb{R}^p\ (Y \ne 0) \Longrightarrow ({}^tY\nabla^2L(\widehat{h})Y = ||J(X)Y||^2 > 0)\text{. C'est une condition}\\ \text{suffisante pour que } \widehat{h} \text{ soit un minimum local d'après le développement limité à l'ordre 2.}\end{array}}$$

Partie III

1)

La matrice ${}^tJJ$ est une matrice appartenant à $\mathcal{M}_p(\mathbb{R})$, symétrique, donc diagonalisable en base orthonormée. Il existe une matrice diagonale $D \in \mathcal{M}_p(\mathbb{R})$, il existe une matrice $V \in \mathcal{M}_p(\mathbb{R})$, orthogonale (${}^tV = V^{-1}$) telles que $D = {}^tV{}^tJJV$.

On sait que les valeurs propres de ${}^tJJ$ sont positives ou nulles (d'après II–5–a)). Rangeons ces valeurs propres dans l'ordre décroissant large, notons q la somme des dimensions des sous-espaces propres associés aux valeurs propres non nulles (il y en a car ${}^tJJ$ est **diagonalisable et non nulle**). Si $q < p$, alors $\dim\operatorname{Ker}{}^tJJ = p - q > 0$ et il y a $p - q$ zéros sur la diagonale de D. Si l'on note $\lambda_1, \dots, \lambda_q, \lambda_{q+1}, \dots, \lambda_p$ les termes diagonaux de D, on peut ecrire $D = \begin{pmatrix}\lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_q & & & \\ & & & \lambda_{q+1} & & \\ & & & & \ddots & \\ & & & & & \lambda_p\end{pmatrix}$

avec $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_q > 0$ et $\lambda_{q+1} = \ldots = \lambda_p = 0$

2–a)

Le rang de la matrice ${}^tJJ$ est égal au rang de D puisque les deux matrices sont semblables.

$$\boxed{\text{rang}({}^tJJ) = q}$$

2–b)

On sait que $\forall i \in [\![1;q]\!], V_i$ est un vecteur propre de ${}^tJJ$ associé à la valeur propre $\lambda_i > 0$: ${}^tJJV_i = \lambda_i V_i$.

Donc $(J^tJ)JV_i = \lambda_i JV_i$.

Si $JV_i = 0$, alors $V_i \in \text{Ker}\, J$, donc $V_i \in \text{Ker}\,{}^tJJ$ d'après II–5–b) ; le vecteur V_i serait un vecteur propre de ${}^tJJ$ associé à la valeur propre 0, **ce qui est faux.**

Donc JV_i est un vecteur propre de J^tJ associé à la valeur propre non nulle λ_i.

Toute valeur propre non nulle de ${}^tJJ$ est une valeur propre non nulle de J^tJ. Par symétrie des rôles joués par J et ${}^tJ$, on en déduit que toute valeur propre non nulle de J^tJ est une valeur propre non nulle de ${}^tJJ$.

$$\boxed{\text{Les matrices } J^tJ \in \mathcal{M}_n(\mathbb{R}) \text{ et } {}^tJJ \in \mathcal{M}_p(\mathbb{R}) \text{ ont les mêmes valeurs propres non nulles}}$$

2–c)

Nous supposerons que la famille $((Y_1, \ldots, Y_r)$ **est une base orthogonale du sous-espace propre de** ${}^tJJ$ **associé à une valeur propre** $\lambda > 0$**, ce qui semble avoir été omis dans l'énoncé.**

Soit $(Y_1, \ldots, Y_r)$ une base orthogonale de vecteurs propres du sous-espace propre de ${}^tJJ$ associé à une valeur propre $\lambda > 0$.

Soit $(\alpha_1, \ldots, \alpha_r) \in \mathbb{R}^r \ / \ \sum_{k=1}^{r} \alpha_k JY_k = 0$. Prenons le produit scalaire de ces deux membres avec JY_i pour $i \in [\![1;r]\!]$. On obtient, après avoir utilisé la linéarité par rapport à la première variable du produit scalaire,

$$\begin{aligned}
\sum_{k=1}^{r} \alpha_k \langle JY_k, JY_i\rangle = 0 \quad &\iff \sum_{k=1}^{r} \alpha_k {}^t(JY_k)JY_i = 0 \\
&\iff \sum_{k=1}^{r} \alpha_k {}^tY_k {}^tJJY_i = 0 \\
&\iff \sum_{k=1}^{r} \alpha_k \lambda {}^tY_k Y_i = 0 \quad \text{puisque } Y_i \text{ est un vecteur propre de } {}^tJJ \text{ associé à } \lambda \\
&\iff \alpha_i \lambda ||Y_i||^2 = 0
\end{aligned}$$

puisque les vecteurs Y_k sont deux à deux orthogonaux.

$\lambda > 0$ et $||Y_i||^2 > 0$; l'égalité précédente implique $\alpha_i = 0$

$$\boxed{\forall (\alpha_1, \ldots, \alpha_r) \in \mathbb{R}^r \ / \ \sum_{k=1}^{r} \alpha_k JY_k = 0 \Longrightarrow \alpha_k = 0 \text{ pour tout } k \in [\![1;r]\!] : \text{la famille } (JY_k) \text{ est libre}}$$

2–d)

Soit E_λ le sous-espace propre de ${}^tJJ$ associé à la valeur propre $\lambda > 0$ et notons r sa dimension. Soit $(V_1, \ldots, V_r) \in \left(\mathcal{M}_{p,1}(\mathbb{R})\right)^r$ une base orthogonale de E_λ.

On vient de voir que la famille $(JV_1, \ldots, JV_r) \in \left(\mathcal{M}_{n;1}(\mathbb{R})\right)^r$ est une famille libre de E'_λ, sous-espace propre de J^tJ associé à $\lambda > 0$. Si l'on note r' sa dimension, on en déduit $r \leq r'$. Par la symétrie de rôles joués par ${}^tJ$ et J, on en déduit que $r' \leq r$.

Conclusion $r = r'$.

Puisque ${}^tJJ$ est diagonalisable, son image est égale à la somme des sous-espaces propres associés aux valeurs propres non nulles. De même pour J^tJ. D'après le résultat précédent, et puisque ces deux matrices ont les mêmes valeurs propres non nulles, on conclut

$$\boxed{\text{Les matrices } {}^tJJ \in \mathcal{M}_p(\mathbb{R}) \text{ et } J^tJ \in \mathcal{M}_n(\mathbb{R}) \text{ ont le même rang } q}$$

3–a)

$\forall i \in [\![1;q]\!],\ \lambda_i > 0$, donc $U_i = \dfrac{1}{\sqrt{\lambda_i}} JV_i$ existe et appartient à $\mathbb{R}^n$ ou $\mathcal{M}_{n,1}(\mathbb{R})$. Soit $(i,j) \in (\ [\![1;q]\!]\ )^2$.

$$\begin{aligned}
\langle U_i, U_j\rangle &= \frac{1}{\sqrt{\lambda_i}\sqrt{\lambda_j}} \langle JV_i, JV_j\rangle \\
&= \frac{1}{\sqrt{\lambda_i}\sqrt{\lambda_j}} {}^t(JV_i)JV_j = \frac{1}{\sqrt{\lambda_i}\sqrt{\lambda_j}} {}^tV_i {}^tJJV_j \\
&= \frac{1}{\sqrt{\lambda_i}\sqrt{\lambda_j}} {}^tV_i({}^tJJ)V_j \\
&= \frac{1}{\sqrt{\lambda_i}\sqrt{\lambda_j}} \lambda_j {}^tV_i V_j \quad \text{car } V_j \text{ est un vecteur propre de } {}^tJJ \text{ associé à } \lambda_j
\end{aligned}$$

Les vecteurs V_i et V_j sont des vecteurs colonnes de la matrice V, qui est une matrice orthogonale, donc ces deux vecteurs sont normés et orthogonaux dès que $i \neq j$. D'où le résultat :

$$\boxed{\langle U_i, U_j\rangle = \begin{cases} 0 & \text{si } i \neq j \\ \dfrac{\lambda_j}{\sqrt{\lambda_j}\sqrt{\lambda_j}} = 1 & \text{si } j = i \end{cases}}$$

La famille (U_k) $1 \leq k \leq q$ est une famille orthonormale de $\mathbb{R}^n$, formée de vecteurs propres de J^tJ car V_k est un vecteur propre de ${}^tJJ$, donc JV_k est un vecteur propre de J^tJ, donc U_k aussi.

3–b)

$(U_1, \ldots, U_q)$ est une famille libre (orthonormale) de q vecteurs propres de J^tJ, associés aux valeurs propres non nulles de J^tJ. Donc $\forall j \in [\![1;q]\!],\ U_j \in \text{Im}(J^tJ)$.

La dimension de cette image est q d'après 2–c), donc $(U_1, \ldots, U_q)$ **est une base orthonormée de** $\text{Im}\, J^tJ$ **; c'est aussi une base orthonormée de la somme des sous-espaces propres associés aux valeurs propres non nulles de** J^tJ.

Puisque la matrice J^tJ est diagonalisable (en base orthonormée d'ailleurs) les sous-espaces propres sont supplémentaires. Le sous-espace propre associé à la valeur 0 est $\text{Ker}(J^tJ)$.

$\mathbb{R}^n = \text{Im}(J^tJ) \bigoplus \text{Ker}(J^tJ)$; d'ailleurs ces deux sous-espaces sont orthogonaux. Donc si l'on prend une base orthonormée $(U_{q+1}, \ldots, U_n)$ de $\text{Ker}(J^tJ)$, la famille $(U_1, \ldots, U_q, U_{q+1}, \ldots, U_n)$ est une base orthonormée de $\mathbb{R}^n$

4)

Soit $S' = {}^tUJV$, on remarque que $S' \in \mathcal{M}_{n,p}(\mathbb{R})$ car $V \in \mathcal{M}_p(\mathbb{R}),\ J \in \mathcal{M}_{n,p}(\mathbb{R}), U \in \mathcal{M}_n(\mathbb{R})$. Ecrivons $s'_{i,j}$ son terme général (on a $1 \leq i \leq n,\ 1 \leq j \leq p$).

$s'_{i,j}$ est le résultat du produit de la i ème ligne de ${}^tU$, qui est ${}^tU_i$, et de la j ème colonne de JV, qui est JV_j. Donc $\boxed{s'_{i,j} = {}^tU_i JV_j = \langle U_i, JV_j\rangle}$

- Pour $1 \leq j \leq q, s'_{i,j} = \langle U_i, \sqrt{\lambda_j} U_j\rangle$ d'après la définition des U_j , donc $s'_{i,j} = \sqrt{\lambda_j}\langle U_i, U_j\rangle$, ce qui donne

$\forall j \in [\![1;q]\!],\ s'_{i,j} = 0$ pour $i \neq j$ et $s'_{j,j} = \sqrt{\lambda_j}$ puisque U_j est normé.

- Pour $j \geq q+1$, le vecteur JV_j est nul ; en effet, V_j est un vecteur propre de ${}^tJJ$ associé à la valeur propre 0, donc c'est un vecteur de $\text{Ker}\,{}^tJJ$ d'après la question 1) ; d'après la question II–5–a) on sait que V_j est aussi un vecteur de $\text{Ker}\, J$, donc $JV_j = 0$ et par conséquent $\langle U, JV_j\rangle = 0$

Donc $s'_{i,j} = \begin{cases} \sqrt{\lambda_j} & \text{si } i = j \text{ et } 1 \leq j \leq q \\ 0 & \text{sinon} \end{cases}$. On retrouve le terme général de S : $S = {}^tUJV$

Les matrices U et V sont orthogonales (ce sont les matrices dont les colonnes sont constituées des vecteurs de 2 bases orthonormées de $\mathbb{R}^n$ et de $\mathbb{R}^p$).

Dans le résultat précédent $S = {}^tUJV$, multiplions les deux termes par U à gauche et ${}^tV$ à droite ; il vient $U^tUJV^tV = US^tV$, soit $\boxed{J = US^tV}$ puisque $U^tU = I_n$ et $V^tV = I_p$

5–a)

Soit $\mu > 0$

${}^tJJ = {}^t(US^tV)US^tV = V^tS^tUUS^tV = V^tSS^tV$ car $U^tU = I_n$. Or la matrice V est orthogonale, donc l'égalité précédente traduit le fait que ${}^tJJ$ et ${}^tSS$ sont semblables dans un changement de bases orthonormées.

$${}^tSS = \begin{pmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_q & & & \\ & & & \lambda_{q+1} & & \\ & & & & \ddots & \\ & & & & & \lambda_p \end{pmatrix} \text{ avec } \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_q > 0 \text{ et } \lambda_k = 0 \text{ pour } k > q$$

$${}^tSS + \mu I_p = \begin{pmatrix} \lambda_1+\mu & & & & & \\ & \ddots & & & & \\ & & \lambda_q+\mu & & & \\ & & & \lambda_{q+1}+\mu & & \\ & & & & \ddots & \\ & & & & & \lambda_p+\mu \end{pmatrix}$$

$\forall i \in [\![1;p]\!]$, $\lambda_i \geq 0$ et $\mu > 0$ impliquent que tous les termes diagonaux sont strictement positifs.

$\boxed{\text{La matrice } {}^tSS + \mu I_p \text{ est inversible, donc } {}^tJJ + \mu I_p \text{ aussi puisque les deux sont semblables}}$

5–b)

Notons $\Delta = {}^tJJ + \mu I_n$.

$$\begin{aligned} \Delta &= V({}^tSS + \mu I_p) \times {}^tV \quad \text{donc} \\ \Delta^{-1} &= V({}^tSS + \mu I_p)^{-1} \times {}^tV \quad \text{car } {}^tV = V^{-1}, \text{ donc} \\ \Delta^{-1} \times {}^tJ &= V({}^tSS + \mu I_p)^{-1} \times {}^tV^tJ \\ &= V({}^tSS + \mu I_p)^{-1} \times {}^tV({}^tUS^tV) \quad \text{d'après 4)} \\ &= V({}^tSS + \mu I_p)^{-1} \times {}^tVV^tS^tU \\ &= V({}^tSS + \mu I_p)^{-1} \times {}^tS^tU \quad \text{puisque } V \text{ est orthogonale} \end{aligned}$$

Pour avoir la relation demandée : $({}^tJJ + \mu I_p)^{-1} \times {}^tJ = VR^tU$ il suffit de vérifier que

$$({}^tSS + \mu I_p)^{-1} \times {}^tS = R.$$

Remarquons que cette égalité équivaut à $({}^tSS + \mu I_p)R = {}^tS$. C'est celle là que nous allons vérifier.

Notons $r_i = \dfrac{\sqrt{\lambda_i}}{\lambda_i + \mu}$ pour $1 \leq i \leq p$

$$R = \begin{pmatrix} r_1 & 0 & \ldots & & & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & & & \vdots & & \vdots \\ \vdots & & & & & & & \\ & & & \ddots & \ddots & \vdots & & \vdots \\ 0 & \ldots & & \ldots & r_p & 0 & \ldots & 0 \end{pmatrix} \in \mathcal{M}_{p,n}(\mathbb{R})$$

$$S = \begin{pmatrix} \sqrt{\lambda_1} & 0 & & 0 \\ 0 & \ddots & \ddots & \\ & \ddots & \ddots & 0 \\ 0 & & 0 & \sqrt{\lambda_p} \\ 0 & \ldots & \ldots & 0 \\ \vdots & & & \vdots \\ & & & \\ 0 & \ldots & \ldots & 0 \end{pmatrix}$$

$${}^tS = \begin{pmatrix} \sqrt{\lambda_1} & 0 & \ldots & & & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & & & \vdots & & \vdots \\ \vdots & & & & & & & \\ & & & \ddots & \ddots & \vdots & & \vdots \\ 0 & \ldots & & \ldots & \sqrt{\lambda_p} & 0 & \ldots & 0 \end{pmatrix} \in \mathcal{M}_{p,n}(\mathbb{R})$$

$${}^tSS + \mu I_p = \begin{pmatrix} \lambda_1+\mu & & & & \\ & \ddots & & & \\ & & \lambda_q+\mu & & \\ & & & \ddots & \\ & & & & \lambda_p+\mu \end{pmatrix}$$

avec $\lambda_k = 0$ pour $k \geq q+1$

$$({}^tSS + \mu I_p)R = \begin{pmatrix} \lambda_1+\mu & & & & \\ & \ddots & & & \\ & & \lambda_q+\mu & & \\ & & & \ddots & \\ & & & & \lambda_p+\mu \end{pmatrix} \begin{pmatrix} r_1 & 0 & \ldots & & & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & & & \vdots & & \vdots \\ \vdots & & & & & & & \\ & & & \ddots & \ddots & \vdots & & \vdots \\ 0 & \ldots & & \ldots & r_p & 0 & \ldots & 0 \end{pmatrix}$$

$$= \begin{pmatrix} r_1(\lambda_1+\mu) & 0 & \ldots & & & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & & & \vdots & & \vdots \\ \vdots & & & & & & & \\ & & & \ddots & \ddots & \vdots & & \vdots \\ 0 & \ldots & & \ldots & r_p(\lambda_p+\mu) & 0 & \ldots & 0 \end{pmatrix} = \begin{pmatrix} \sqrt{\lambda_1} & 0 & \ldots & & & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & & & \vdots & & \vdots \\ \vdots & & & & & & & \\ & & & \ddots & \ddots & \vdots & & \vdots \\ 0 & \ldots & & \ldots & \sqrt{\lambda_p} & 0 & \ldots & 0 \end{pmatrix}$$

En effet, les $n-p$ dernières colonnes de R sont nulles, donc celles de $({}^tSS + \mu I_p)R$ aussi et pour les p premières cela revient à faire un produit de matrices diagonales.

On retrouve bien la matrice ${}^tS$

5–c)

Gardons les notations $r_i = \dfrac{\sqrt{\lambda_i}}{\lambda_i + \mu}$ pour $1 \leq i \leq p$; on remarque que $r_i = 0$ pour $i > q$.

$$V = \begin{pmatrix} v_{1,1} & \ldots & v_{1,p} \\ \vdots & & \vdots \\ v_{p,1} & \ldots & v_{p,p} \end{pmatrix} \in \mathcal{M}_p(\mathbb{R}) \ ; \ R = \begin{pmatrix} r_1 & 0 & \ldots & & & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & & & \vdots & & \vdots \\ \vdots & & & & & & & \\ & & & \ddots & \ddots & \vdots & & \vdots \\ 0 & \ldots & & \ldots & r_p & 0 & \ldots & 0 \end{pmatrix} \in \mathcal{M}_{p,n}(\mathbb{R}) \text{ et}$$

$$U = \begin{pmatrix} u_{1,1} & \ldots & u_{1,n} \\ \vdots & & \vdots \\ u_{n,1} & \ldots & u_{n,n} \end{pmatrix} \in \mathcal{M}_n(\mathbb{R}) \ ;$$

$$VR^tU=\begin{pmatrix} r_1v_{1,1} & r_2v_{1,2} & \dots & & r_pv_{1,p} & 0 & \dots & 0\\ r_1v_{2,1} & r_2v_{2,2} & \dots & & r_pv_{2,p} & \vdots & & \vdots\\ \vdots & & & & & & & \\ & & & \ddots & \ddots & \vdots & & \vdots\\ r_1v_{p,1} & r_2v_{p,2} & & \dots & r_pv_{p,p} & 0 & \dots & 0\end{pmatrix}\begin{pmatrix} u_{1,1} & \dots & u_{j,1} & \dots & u_{n,1}\\ \vdots & & \vdots & & \vdots\\ u_{1,i} & \dots & u_{j,i} & \dots & u_{n,i}\\ \vdots & & \vdots & & \vdots\\ u_{1,n} & \dots & u_{j,n} & \dots & u_{n,n}\end{pmatrix}$$

Soit $\alpha \in [\![1,p]\!]$ et $\beta \in [\![1;n]\!]$. Le terme général d'indice (α,β) de ce produit est $\boxed{\sum_{k=1}^{p} r_k v_{\alpha,k} u_{\beta,k}.}$

$$V_i{}^tU_i=\begin{pmatrix} v_{1,i}\\ \vdots\\ v_{\alpha,i}\\ \vdots\\ v_{p,i}\end{pmatrix}\times\begin{pmatrix} u_{1,i} & \dots & u_{\beta,i} & \dots & u_{n,i}\end{pmatrix}$$

Son terme d'indice (α,β) est $v_{\alpha,i}u_{\beta,i}$ Le terme général d'indice (α,β) de la somme $\sum_{i=1}^{p} r_iV_i{}^tU_i$ est

$\boxed{\sum_{i=1}^{p} r_i v_{\alpha,i} u_{\beta,i}}$

Les deux résultats encadrés sont égaux puisque les indices sont muets. De plus la somme s'arrête à q puisque $r_i = 0$ pour $i > q$.

$$\boxed{VR^tU=\sum_{i=1}^{q} r_iV_i{}^tU_i \text{ ; donc } ({}^tJJ+\mu I_p)^{-1}\times{}^tJ=\sum_{i=1}^{q}\frac{\sqrt{\lambda_i}}{\lambda_i+\mu}V_i{}^tU_i}$$

6–a)

$F(X+h)-M(h)=F(X+h)-L(h)-\dfrac{\mu}{2}||h||^2$; donc

$\forall h\neq 0,\ \dfrac{F(X+h)-M(h)}{||h||}=\dfrac{F(X+h)-L(h)}{||h||}-\dfrac{\mu||h||}{2}$

On sait que $\lim\limits_{h\to 0}\dfrac{F(X+h)-L(h)}{||h||}=0$ d'après II–3–a) donc

$$\boxed{\lim_{h\to 0}\frac{F(X+h)-M(h)}{||h||}=0}$$

6–b)

$$\begin{aligned} M(h) &= L(h)+\frac{\mu}{2}||h||^2\\ &= F(X)+{}^th\nabla F(X)+\frac{1}{2}{}^thG(X)h+\frac{\mu}{2}{}^thh \qquad \text{d'après II–1)}\\ &= F(X)+{}^th\nabla F(X)+\frac{1}{2}{}^th(G(X)+\mu I_p)h\end{aligned}$$

D'après II–4–b) et c) on peut écrire

$\nabla M(h)=\nabla F(X)+(G(X)+\mu I_p)h=\nabla L(h)+\mu h$

$\nabla^2M(h)=G(X)+\mu I_p=\nabla^2L(h)+\mu I_p.$

6–c)

Les points critiques h^* de M vérifient $\nabla M(h^*)=0$, soit $(G(X)+\mu I_p)h^*=-\nabla F(X)$ et, puisque $(G(X)+\mu I_p)$ est inversible,

$$h^*=-(G(X)+\mu I_p)^{-1}\nabla F(X) \text{ avec } \nabla F(X)={}^tJ(X)f(X)$$

Puique $\nabla F(X)\neq 0$, on en conclut $-(G(X)+\mu I_p)^{-1}\nabla F(X)\neq 0$ (car $(G(X)+\mu I_p)$ est inversible), donc $h^*\neq 0$

$h^*=-(G(X)+\mu I_p)^{-1}\nabla F(X)$ est non nul et c'est l'unique point critique de M
D'après 5–c), $h^*=-\Big(\sum_{i=1}^{q}\dfrac{\sqrt{\lambda_i(X)}}{\lambda_i(X)+\mu}V_i(X){}^tU_i(X)\Big)f(X)$

6–d)

$$\begin{aligned}\langle h^*,\nabla F(X)\rangle &= {}^th^*\nabla F(X)\\ &= -{}^th^*(G(X)+\mu I_p)h^*\\ &= -{}^th^*G(X)h^*-\mu{}^th^*h^*\\ &= -||J(X)h^*||^2-\mu||h^*||^2\end{aligned}$$

car $G(X)={}^tJ(X)J(X)$.

h^* est une direction de décroissance de F en X.

Soit $Y\in\mathbb{R}^p$, non nul.

${}^tY\nabla^2M(h^*)Y={}^tY(G(X)+\mu I)Y=||J(X)Y||^2\mu+||Y||^2>0$

C'est une condition suffisante pour que M admette en h^* un minimum local

BCE

BANQUE COMMUNE D'EPREUVES

Concepteur : H.E.C.

OPTION ECONOMIQUE

MATHEMATIQUES III

La présentation, la lisibilité, l'orthographe, la qualité de la rédaction, la clarté et la précision des raisonnements entreront pour une part importante dans l'appréciation des copies.
Les candidats sont invités à encadrer dans la mesure du possible les résultats de leurs calculs.
Ils ne doivent faire usage d'aucun document : **l'utilisation de toute calculatrice et de tout matériel électronique est interdite.**
Seule l'utilisation d'une règle graduée est autorisée.

EXERCICE

Étant donné un entier n supérieur ou égal à 2, on considère un nuage de n points du plan, c'est-à-dire un n-uplet $((x_1,y_1),(x_2,y_2),\ldots,(x_n,y_n))$ d'éléments de $\mathbb{R}^2$. On suppose que les réels $x_1,x_2,\ldots,x_n$ (resp. $y_1,y_2,\ldots,y_n$) ne sont pas tous égaux.

On appelle moyenne arithmétique $\overline{x}$ et écart-type σ_x du n-uplet $x=(x_1,\ldots,x_n)$, les réels suivants :

$$\overline{x}=\frac{1}{n}\sum_{k=1}^{n}x_k \quad \text{et} \quad \sigma_x=\sqrt{\frac{1}{n}\sum_{k=1}^{n}(x_k-\overline{x})^2}$$

On définit de même la moyenne arithmétique $\overline{y}$ et l'écart-type σ_y du n-uplet $y=(y_1,\ldots,y_n)$.

La covariance $\mathrm{cov}(x,y)$ et le coefficient de corrélation linéaire $r(x,y)$ du couple (x,y) sont donnés par :

$$\mathrm{cov}(x,y)=\frac{1}{n}\sum_{k=1}^{n}(x_k-\overline{x})(y_k-\overline{y}) \quad \text{et} \quad r(x,y)=\frac{\mathrm{cov}(x,y)}{\sigma_x\times\sigma_y}$$

Soit f la fonction définie sur $\mathbb{R}^2$ à valeurs réelles qui, à tout couple (a,b) de $\mathbb{R}^2$, associe le réel $f(a,b)$ tel que :

$$f(a,b)=\sum_{k=1}^{n}(ax_k+b-y_k)^2$$

1. Justifier que f est de classe C^2 sur $\mathbb{R}^2$.

2. a) Écrire le système d'équations (S) permettant de déterminer les points critiques de f.

b) Résoudre le système (S). En déduire que f admet un unique point critique $(\hat{a},\hat{b})$ que l'on exprimera en fonction de $\overline{x},\overline{y},\sigma_x^2$ et $\mathrm{cov}(x,y)$.

c) Montrer que ce point critique correspond à un minimum local de f.

d) Établir la formule suivante : $f(\hat{a},\hat{b})=n\sigma_y^2(1-r^2(x,y))$

3. a) Montrer que l'on a : $|r(x,y)|\leqslant 1$.

b) Que peut-on dire du nuage de points lorsque $|r(x,y)|=1$?

PROBLÈME

Dans tout le problème, N désigne un entier naturel fixé supérieur ou égal à 2, et p un réel fixé de l'intervalle $]0,1[$. On pose : $q=1-p$. Soit n un entier naturel quelconque.

Dans une population de N individus, on s'intéresse à la propagation d'un certain virus. Chaque jour, on distingue dans cette population trois catégories d'individus : en premier lieu, les individus sains, c'est-à-dire ceux qui ne sont pas porteurs du virus, ensuite les individus qui viennent d'être contaminés et qui sont inoffensifs pour les autres, et enfin, les individus contaminés par le virus et qui sont contagieux.

Ces trois catégories évoluent jour après jour selon le modèle suivant :

- chaque jour n, chaque individu sain peut être contaminé par n'importe lequel des individus contagieux ce jour avec la même probabilité p, ces contaminations éventuelles étant indépendantes les unes des autres ;
- un individu contaminé le jour n devient contagieux le jour $n+1$;
- chaque individu contagieux le jour n redevient sain le jour $n+1$.

On note alors X_n le nombre aléatoire d'individus contagieux le jour n.
On remarquera que si, pour un certain entier naturel i, on a $X_i=0$, alors on a aussi $X_{i+1}=0$.
Les variables aléatoires $X_0,X_1,\ldots,X_n,\ldots$ sont supposées définies sur un même espace probabilisé $(\Omega,\mathcal{A},P)$, et $E(X_n)$ désigne, pour tout n de $\mathbb{N}$, l'espérance de X_n.

Partie I. Un cas particulier

Dans cette partie uniquement, on suppose que l'on a : $N=3$ et $p=1/3$.
On considère les matrices S et R suivantes :

$$S=\begin{pmatrix}9&4&4&9\\0&4&5&0\\0&1&0&0\\0&0&0&0\end{pmatrix}, \quad R=\begin{pmatrix}0&1&-6&1\\1&0&5&0\\-1&0&1&0\\0&-1&0&0\end{pmatrix}$$

L'ensemble $\mathcal{M}_{4,1}(\mathbb{R})$ des matrices colonnes à quatre lignes est confondu avec l'espace vectoriel $\mathbb{R}^4$.

1. Montrer que la matrice R est inversible et calculer son inverse R^{-1}.

2. a) Montrer que les réels $-1,0,5$ et 9 sont des valeurs propres de S.

b) Calculer le produit matriciel $R^{-1}SR$.

c) En déduire, pour tout n de $\mathbb{N}$, l'expression de la matrice S^n en fonction de n (on pose $S^0=I$, où I désigne la matrice identité de $\mathcal{M}_4(\mathbb{R})$).

3. Soit n un entier fixé de $\mathbb{N}$.

a) Déterminer la loi de probabilité conditionnelle de X_{n+1} sachant l'événement $[X_n=0]$.

b) Déterminer la loi de probabilité conditionnelle de X_{n+1} sachant l'événement $[X_n=3]$.

c) Vérifier que la loi de probabilité conditionnelle de X_{n+1} sachant l'événement $[X_n=1]$ (resp. $[X_n=2]$) est la loi binomiale de paramètres $(2,\frac{1}{3})$ (resp. $(1,\frac{5}{9})$).

d) On note $E(X_{n+1}/X_n=i)$ l'espérance de la loi de probabilité conditionnelle de X_{n+1} sachant l'événement $[X_n=i]$. Déterminer les valeurs respectives de $E(X_{n+1}/X_n=1)$ et $E(X_{n+1}/X_n=2)$.

4. On suppose, *uniquement dans cette question*, que X_0 suit la loi binomiale de paramètres $(3,\frac{1}{3})$.

a) Déterminer la loi de X_1 et calculer $E(X_1)$.

b) Vérifier la formule suivante : $E(X_1)=\sum_{i=0}^{3}E(X_1/X_0=i)\times P([X_0=i])$.

5. Pour tout entier naturel n, on considère le vecteur U_n de $\mathbb{R}^4$ défini par :

$$U_n=\begin{pmatrix}u_n\\v_n\\w_n\\t_n\end{pmatrix}=\begin{pmatrix}P([X_n=0])\\P([X_n=1])\\P([X_n=2])\\P([X_n=3])\end{pmatrix}$$

a) Déterminer une relation entre u_n, v_n, w_n, et t_n.

b) À l'aide de la formule des probabilités totales appliquée au système complet d'événements $([X_n = i])_{0\leqslant i\leqslant 3}$, déterminer une matrice M de $\mathcal{M}_4(\mathbb{R})$ indépendante de n, telle que : $U_{n+1} = MU_n$.

c) Exprimer M en fonction de S. En déduire les valeurs propres de M.

d) Donner l'expression des réels u_n et v_n en fonction de n, v_0 et w_0.

6. On pose : $F = \bigcup_{n=0}^{+\infty} [X_n = 0]$.

a) Que signifie l'événement F ?

b) Montrer que le virus finit par disparaître presque sûrement, quelle que soit la loi de la variable aléatoire initiale X_0.

Partie II. Le cas général

On suppose que pour tout entier naturel n et pour tout entier i de $[\![0, N]\!]$, on a : $P([X_n = i]) > 0$. On suppose également que pour tout couple (i, j) de $[\![0, N]\!]^2$, le réel $q_{i,j}$ défini par : $q_{i,j} = P_{[X_n=i]}([X_{n+1} = j])$, est indépendant de n.
Soit Q la matrice de $\mathcal{M}_{N+1}(\mathbb{R})$ définie par : $Q = (q_{i,j})_{0\leqslant i,j\leqslant N}$.

1. a) Déterminer, pour tout j de $[\![0, N]\!]$, les probabilités $q_{0,j}$ et $q_{N,j}$. De même, déterminer pour tout i de $[\![0, N]\!]$, la probabilité $q_{i,N}$.

b) Justifier que si l'on a $j > N - i$, alors $q_{i,j} = 0$.

c) Montrer que pour tout i de $[\![1, N-1]\!]$, la loi de probabilité conditionnelle de X_{n+1} sachant $[X_n = i]$, est une loi binomiale dont on déterminera les paramètres.

2. a) Montrer que 1 est valeur propre de la matrice Q.

b) Soit λ une valeur propre de Q, et $V = \begin{pmatrix} V(0) \\ V(1) \\ \vdots \\ V(N) \end{pmatrix}$ un vecteur propre associé à λ.

On pose : $|V(i)| = \max_{0\leqslant j\leqslant N} |V(j)|$. Justifier que la composante $V(i)$ n'est pas nulle, puis, en examinant la ligne i du système $QV = \lambda V$, montrer que l'on a : $|\lambda| \leqslant 1$.

3. On pose pour tout entier naturel n : $U_n = \begin{pmatrix} P([X_n = 0]) \\ P([X_n = 1]) \\ \vdots \\ P([X_n = N]) \end{pmatrix}$.

À l'aide de la formule des probabilités totales, déterminer en fonction de Q une matrice M de $\mathcal{M}_{N+1}(\mathbb{R})$ indépendante de n et vérifiant, pour tout n de $\mathbb{N}$, la relation : $U_{n+1} = MU_n$.

On suppose jusqu'à la fin de la partie II que la matrice M est diagonalisable, et que $\mathcal{B} = (V_0, \dots, V_N)$ est une base de vecteurs propres de M telle que, pour tout k de $[\![0, N]\!]$, le vecteur propre V_k est associé à une valeur propre λ_k.

De plus, on suppose que : $\lambda_0 = 1$, $V_0 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, et que pour tout k de $[\![1, N]\!]$, on a : $|\lambda_k| < 1$.

4. On décompose alors le vecteur U_0 sur la base $\mathcal{B}$: $U_0 = \sum_{k=0}^{N} \alpha_k V_k$.

a) Déterminer, pour tout n de $\mathbb{N}$, la décomposition du vecteur U_n sur la base $\mathcal{B}$.

b) On note, pour tout couple (k, i) de $[\![0, N]\!]^2$, $V_k(i)$ la $(i+1)$-ième composante du vecteur V_k. Exprimer, pour tout n de $\mathbb{N}$ et pour tout i de $[\![0, N]\!]$, la probabilité de l'événement $[X_n = i]$ en fonction des réels α_k, λ_k et $V_k(i)$ $(k \in [\![0, N]\!])$.

c) Montrer que, pour tout i de $[\![1, N]\!]$, on a : $\lim_{n\to+\infty} P([X_n = i]) = 0$.

d) En déduire que le virus finit par disparaître presque sûrement, quelle que soit la loi de la variable aléatoire initiale X_0.

Partie III. Estimations ponctuelle et par intervalle de confiance de p

On suppose que le paramètre p, qui exprime la probabilité qu'un individu contagieux transmette le virus à un individu sain, est inconnu, et on cherche à l'estimer. On rappelle que : $q = 1 - p$.
Pour m entier supérieur ou égal à 1, on considère un m-échantillon $(Y_1, Y_2, \dots, Y_m)$ de variables aléatoires indépendantes, de même loi de Bernoulli de paramètre p, définies sur un espace probabilisé $(\Omega, \mathcal{A}, P)$.

On pose : $\overline{Y_m} = \frac{1}{m} \sum_{i=1}^{m} Y_i$.

Dans toute la suite de cette partie, on note ε un réel strictement positif quelconque.

1. a) Montrer que $\overline{Y_m}$ est un estimateur sans biais de p ; déterminer son risque quadratique.

b) À l'aide de l'inégalité de Bienaymé-Tchebycheff, montrer que l'intervalle $\left[\overline{Y_m} - \sqrt{\frac{5}{m}}, \overline{Y_m} + \sqrt{\frac{5}{m}}\right]$ est un intervalle de confiance de p au niveau de confiance 0.95.

2. Soit θ un réel positif.
a) Établir l'égalité suivante : $P([\overline{Y_m} - p \geqslant \varepsilon]) = P([e^{m\theta\overline{Y_m}} \geqslant e^{m\theta(p+\varepsilon)}])$.

b) Montrer que si T est une variable aléatoire discrète finie à valeurs positives d'espérance $E(T)$, et a un réel strictement positif, on a l'inégalité : $P([T \geqslant a]) \leqslant \frac{E(T)}{a}$.

c) Soit g la fonction définie sur $\mathbb{R}^+$ par : $g(x) = \ln(pe^x + q)$. Déduire des questions précédentes, l'inégalité suivante : $P([\overline{Y_m} - p \geqslant \varepsilon]) \leqslant e^{m(g(\theta)-\theta(p+\varepsilon))}$.

d) Montrer que la fonction g est de classe C^2 sur $\mathbb{R}^+$ et vérifie, pour tout x de $\mathbb{R}^+$, l'inégalité : $|g''(x)| \leqslant \frac{1}{4}$.

e) En déduire l'inégalité suivante : $g(\theta) \leqslant \theta p + \frac{\theta^2}{8}$.

f) Étudier les variations de la fonction h définie sur $\mathbb{R}^+$ par : $h(x) = \frac{x^2}{8} - \varepsilon x$.
En déduire l'inégalité : $P([\overline{Y_m} - p \geqslant \varepsilon]) \leqslant e^{-2m\varepsilon^2}$.

3. On pose : $\overline{W_m} = \frac{1}{m} \sum_{i=1}^{m} (1 - Y_i)$. Établir l'inégalité : $P([\overline{W_m} - q \geqslant \varepsilon]) \leqslant e^{-2m\varepsilon^2}$.

4. a) Déduire des questions 2 f) et 3, l'inégalité suivante : $P([|\overline{Y_m} - p| \geqslant \varepsilon]) \leqslant 2e^{-2m\varepsilon^2}$.

b) Sachant que $\ln(0.025) \approx -3.688$, calculer $2e^{-2m\varepsilon^2}$ pour $\varepsilon = \sqrt{\frac{1.844}{m}}$. En déduire un nouvel intervalle de confiance de p au niveau de confiance 0.95. Comparer cet intervalle de confiance à celui obtenu à la question 1.b. Conclure.

Exercice

1)

f est un polynôme par rapport aux variables a et b , défini sur $\mathbb{R}^2$ et à valeurs dans $\mathbb{R}$

$$\boxed{\text{D'après le cours, } f \text{ est de classe } C^2 \text{ sur } \mathbb{R}^2}$$

2–a)

Un point (a,b) de $\mathbb{R}^2$ est un point critique si et seulement si (a,b) est solution de **(S)** $= \begin{cases} \dfrac{\partial f}{\partial a}(a,b) &= 0 \\ \dfrac{\partial f}{\partial b}(a,b) &= 0 \end{cases}$

$$\begin{cases} \dfrac{\partial f}{\partial a}(a,b) &= 0 \\ \dfrac{\partial f}{\partial b}(a,b) &= 0 \end{cases} \iff \begin{cases} \displaystyle\sum_{k=1}^{n} 2(ax_k+b-y_k)x_k &= 0 \\ \displaystyle\sum_{k=1}^{n} 2(ax_k+b-y_k) &= 0 \end{cases}$$

$$\iff \begin{cases} a\displaystyle\sum_{k=1}^{n} x_k^2+\sum_{k=1}^{n} bx_k-\sum_{k=1}^{n} y_kx_k &= 0 \\ a\displaystyle\sum_{k=1}^{n} x_k+\sum_{k=1}^{n} b-\sum_{k=1}^{n} y_k &= 0 \end{cases}$$

$$\iff \begin{cases} a\displaystyle\sum_{k=1}^{n} x_k^2+b\sum_{k=1}^{n} x_k &= \displaystyle\sum_{k=1}^{n} x_ky_k \\ a\displaystyle\sum_{k=1}^{n} x_k+\sum_{k=1}^{n} b &= \displaystyle\sum_{k=1}^{n} y_k \end{cases}$$

$$\iff \begin{cases} a\displaystyle\sum_{k=1}^{n} x_k^2+bn\overline{x} &= \displaystyle\sum_{k=1}^{n} x_ky_k \\ na\overline{x}+nb &= n\overline{y} \end{cases}$$

$$\iff \begin{cases} na\overline{x}+nb &= n\overline{y} \qquad L_2 \longleftrightarrow L_1 \\ a\displaystyle\sum_{k=1}^{n} x_k^2+bn\overline{x} &= \displaystyle\sum_{k=1}^{n} x_ky_k \end{cases}$$

$$\iff \begin{cases} a\overline{x}+b &= \overline{y} \qquad L_1 \longleftarrow \frac{1}{n}L_1 \\ a\Big(\displaystyle\sum_{k=1}^{n} x_k^2-n(\overline{x})^2\Big) &= \displaystyle\sum_{k=1}^{n} y_kx_k-n\overline{x}\times\overline{y} \qquad L_2 \longleftarrow L_2-\overline{x}L_1 \end{cases}$$

$$\boxed{(\mathbf{S}) \iff \begin{cases} a\overline{x}+b &= \overline{y} \\ a\displaystyle\sum_{k=1}^{n}(x_k^2-(\overline{x})^2) &= \displaystyle\sum_{k=1}^{n}(y_kx_k-\overline{x}\times\overline{y}) \end{cases}}$$

2–b)

Remarques

(1)
$$\begin{aligned} \sigma_x^2 &= \frac{1}{n}\sum_{k=1}^{n}(x_k^2-2\overline{x}x_k+(\overline{x})^2) \\ &= \frac{1}{n}\Big(\sum_{k=1}^{n}x_k^2-2\overline{x}\times\sum_{k=1}^{n}x_k+\sum_{k=1}^{n}(\overline{x})^2\Big) \\ &= \frac{1}{n}\Big(\sum_{k=1}^{n}x_k^2-2n(\overline{x})^2+n(\overline{x})^2\Big) \\ &= \frac{1}{n}\Big(\sum_{k=1}^{n}x_k^2-n(\overline{x})^2\Big) \\ &= \frac{1}{n}\sum_{k=1}^{n}\Big(x_k^2-(\overline{x})^2\Big) \end{aligned}$$

(2)
$$\begin{aligned} \operatorname{cov}(x,y) &= \frac{1}{n}\sum_{k=1}^{n}(x_k-\overline{x})(y_k-\overline{y}) \\ &= \frac{1}{n}\Big(\sum_{k=1}^{n}x_ky_k-\overline{y}\sum_{k=1}^{n}x_k-\overline{x}\sum_{k=1}^{n}y_k+\sum_{k=1}^{n}\overline{x}\times\overline{y}\Big) \\ &= \frac{1}{n}\Big(\sum_{k=1}^{n}x_ky_k-n\overline{y}\times\overline{x}-n\overline{x}\times\overline{y}+n\overline{x}\times\overline{y}\Big) \\ &= \frac{1}{n}\Big(\sum_{k=1}^{n}x_ky_k-n\overline{x}\times\overline{y}\Big) \\ &= \frac{1}{n}\sum_{k=1}^{n}(x_ky_k-\overline{x}\times\overline{y}) \end{aligned}$$

Dans ces conditions, $a\displaystyle\sum_{k=1}^{n}(x_k^2-(\overline{x})^2)=\sum_{k=1}^{n}(y_kx_k-\overline{x}\times\overline{y})$ s'écrit aussi $a\sigma_x^2=\operatorname{cov}(x,y)$. Or $\sigma_x^2\neq 0$ sinon, d'après la définition de σ_x^2 on aurait $\forall k\in[\![1,n]\!],\ x_k=\overline{x}$, tous les x_k seraient égaux, ce qui est contraire à l'hypothèse. On peut diviser par σ_x^2 l'égalité précédente et on a $\boxed{a=\dfrac{\operatorname{cov}(x,y)}{\sigma_x^2}=\widehat{a}}$

L'égalité $a\overline{x}+b=\overline{y}$ s'écrit $\boxed{b=\widehat{b}=\overline{y}-\widehat{a}\,\overline{x}}$

2–c)

La fonction f est de classe C^2 sur l'ouvert $\mathbb{R}^2$, on peut donc lui appliquer le Théorème de Schwarz : $\dfrac{\partial^2 f}{\partial a\partial b}(a,b)=\dfrac{\partial^2 f}{\partial b\partial a}(a,b)$.

Comme $\mathbb{R}^2$ est un ouvert, f ne peut admettre un extremum local qu'en un point critique, donc uniquement au point $(\widehat{a},\widehat{b})$.

Utilisons les notations de Monge. Pour tout couple $(a,b)\in\mathbb{R}^2$,

$$
\begin{aligned}
\frac{\partial^2 f}{\partial a^2}(a,b) &= r(a,b) = 2\sum_{k=1}^{n} x_k^2\\
\frac{\partial^2 f}{\partial a \partial b}(a,b) &= s(a,b) = 2\sum_{k=1}^{n} x_k = 2n\overline{x}\\
\frac{\partial^2 f}{\partial b^2}(a,b) &= t(a,b) = 2\sum_{k=1}^{n} 1 = 2n
\end{aligned}
$$

$$
\begin{aligned}
(s^2 - rt)(a,b) &= 4n^2(\overline{x})^2 - 4n\sum_{k=1}^{n} x_k^2\\
&= 4n\Big(n(\overline{x})^2 - \sum_{k=1}^{n} x_k^2\Big)\\
&= -4n\sum_{k=1}^{n}(x_k^2 - \overline{x})^2\\
&= -4n^2\sigma_x^2 \qquad \text{d'après la remarque (1) faite au 2−b)}
\end{aligned}
$$

$\sigma_x^2 \neq 0 \Longrightarrow \sigma_x^2 > 0$, donc $\boxed{f \text{ présente un extremum local en } (\widehat{a},\widehat{b})}$

Les nombres x_k n'étant pas tous égaux, l'un au moins d'entre eux n'est pas nul, donc $r(a,b) = 2\sum_{k=1}^{n} x_k^2 > 0$

$$\boxed{f \text{ présente un minimum local en } (\widehat{a},\widehat{b})}$$

2– d)

$$
\begin{aligned}
f(\widehat{a},\widehat{b}) &= \sum_{k=1}^{n}(\widehat{a}x_k + \overline{y} - \widehat{a}.\overline{x} - y_k)^2\\
&= \sum_{k=1}^{n}\Big((x_k - \overline{x})\widehat{a} - (y_k - \overline{y})\Big)^2\\
&= \widehat{a}^2\sum_{k=1}^{n}(x_k - \overline{x})^2 - 2\widehat{a}\sum_{k=1}^{n}(x_k - \overline{x})(y_k - \overline{y}) + \sum_{k=1}^{n}(y_k - \overline{y})^2\\
&= \widehat{a}^2 n\sigma_x^2 - 2\widehat{a}n\,\mathrm{cov}(x,y) + n\sigma_y^2\\
&= n\frac{(\mathrm{cov}(x,y))^2}{\sigma_x^2} - n\frac{2(\mathrm{cov}(x,y))^2}{\sigma_x^2} + n\sigma_y^2 \qquad \text{d'après 2−b)}\\
&= n\Big(\sigma_y^2 - \frac{(\mathrm{cov}(x,y))^2}{\sigma_x^2}\Big)\\
&= n(\sigma_y^2 - \sigma_y^2 r^2(x,y))
\end{aligned}
$$

$$\boxed{f(\widehat{a},\widehat{b}) = n\sigma_y^2(1 - r^2(x,y))}$$

3– a)

$f(a,b)$ est une somme de carrés, donc $f(a,b) \geq 0$. En particulier $f(\widehat{a},\widehat{b}) \geq 0$, donc $n\sigma_y^2(1 - r^2(x,y)) \geq 0$.

Les y_k n'étant pas tous égaux, l'un d'entre eux au moins est différent de $\overline{y}$, donc $\sigma_y^2 \neq 0$ donc $\sigma_y^2 > 0$.

L'inégalité précédente équivaut à $1 - r^2(x,y) \geq 0$, donc $r^2(x,y) \leq 1$, donc $\boxed{|r(x,y)| \leq 1}$

3– b)

Si $|r(x,y)| = 1$, alors $f(\widehat{a},\widehat{b}) = 0$, ce qui équivaut successivement à

$$
\begin{aligned}
\sum_{k=1}^{n}(\widehat{a}x_k + \widehat{b} - y_k)^2 &= 0\\
\forall k \in [\![1,n]\!],\ \widehat{a}x_k + \widehat{b} - y_k &= 0\\
\forall k \in [\![1,n]\!],\ y_k &= \widehat{a}x_k + \widehat{b}
\end{aligned}
$$

$$\boxed{\text{Les points } (x_k,y_k) \text{ sont tous sur la droite d'équation } y = \widehat{a}x + \widehat{b}}$$

Probème

Commentaire concernant la remarque $(X_i = 0) \Longrightarrow (X_{i+1} = 0)$ **:**

Cette implication a été contestée en disant : on peut avoir simultanément aucun individu contagieux le jour i et k, $(k > 0)$ individus contaminés, non contagieux ce même jour, ce qui impliquerait que $(X_{i+1} = k)$.

Ceci est faux, cette situation ne peut avoir lieu.

Explication : Si le jour $i+1$, il y a k $(k > 0)$ individus contaminés et contagieux, alors ces k individus ont été contaminés le jour i. Donc ce jour là, il doit y avoir au moins un individu contagieux, pour les contaminer, c'est-à-dire que $X_i \neq 0$.

On vient de montrer : $X_{i+1} \neq 0 \Longrightarrow X_i \neq 0$, ce qui revient à : $X_i = 0 \Longrightarrow X_{i+1} = 0$

Partie I

1)

Résolvons le système **(S)**: $R\begin{pmatrix} x\\ y\\ z\\ t\end{pmatrix} = \begin{pmatrix} a\\ b\\ c\\ d\end{pmatrix}$

$$
\textbf{(S)} \iff \left\{\begin{array}{ll}
y - 6z + t & = a\\
x \quad + 5z & = b\\
-x \quad + z & = c\\
-y & = d
\end{array}\right.
$$

$$
\iff \left\{\begin{array}{ll}
y & = -d\\
x + 5z & = b\\
-x + z & = c\\
-6z + t & = a + d
\end{array}\right.
$$

$$
\iff \left\{\begin{array}{ll}
y & = -d\\
z & = \frac{1}{6}(b + c)\\
x & = z - c\\
t & = a + d + 6z
\end{array}\right. \qquad L_2 \longleftarrow \frac{1}{6}(L_2 + L_3)
$$

$$
\iff \left\{\begin{array}{ll}
x & = \frac{1}{6}(b - 5c)\\
y & = -d\\
z & = \frac{1}{6}(b + c)\\
t & = a + b + c + d
\end{array}\right.
$$

Ceci équivaut à l'égalité matricielle $\begin{pmatrix} x\\ y\\ z\\ t\end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{6} & -\frac{5}{6} & 0\\ 0 & 0 & 0 & -1\\ 0 & \frac{1}{6} & \frac{1}{6} & 0\\ 1 & 1 & 1 & 1\end{pmatrix}\begin{pmatrix} a\\ b\\ c\\ d\end{pmatrix}$

On en déduit $R^{-1} = \begin{pmatrix} 0 & \frac{1}{6} & -\frac{5}{6} & 0 \\ 0 & 0 & 0 & -1 \\ 0 & \frac{1}{6} & \frac{1}{6} & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}$

2–a)

Le réel λ est valeur propre de S si et seulement si le système $(S-\lambda I)\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ n'est pas de Cramer.

Ce système équivaut au système Σ suivant

$$(\Sigma) \quad \begin{cases} (9-\lambda)x + 4y + 4z + 9t &= 0 \\ (4-\lambda)y + 5z &= 0 \\ y - \lambda z &= 0 \\ -\lambda t &= 0 \end{cases}$$

- **Pour $\lambda = -1$:**

$$(\Sigma) \iff \begin{cases} 10x + 4y + 4z + 9t &= 0 \\ 5y + 5z &= 0 \\ y + z &= 0 \\ t &= 0 \end{cases}$$

$$\iff \begin{cases} t = 0 & \\ z = -y & \text{d'après } L_2 \text{ et } L_3 \\ 10x = 0 & \text{en reportant les résultats précédents dans } L_1 \end{cases}$$

$$\iff \begin{cases} x = 0 \\ y \in \mathbb{R} \\ z = -y \\ t = 0 \end{cases}$$

Le système n'est pas de Cramer, il admet d'autres solutions que la solution $(0,0,0,0)$: $\lambda = -1$ **est valeur propre de** S.

Le sous-espace propre associé est :

$E_{-1} = \{(0, y, -y, 0) \in \mathbb{R}^4 \ / \ y \in \mathbb{R}\} = \text{vect}((0,1,-1,0))$

- **Pour $\lambda = 0$:**

$$(\Sigma) \iff \begin{cases} 9x + 4y + 4z + 9t &= 0 \\ 4y + 5z &= 0 \\ y &= 0 \end{cases}$$

$$\iff \begin{cases} 9x + 9t &= 0 \\ y = z &= 0 \\ t &\in \mathbb{R} \end{cases}$$

$$\iff \begin{cases} x &= -t \\ y = z &= 0 \\ t &\in \mathbb{R} \end{cases}$$

Le système n'est pas de Cramer: $\lambda = 0$ **est valeur propre de** S.

Le sous-espace propre associé est :

$E_0 = \{(-t, 0, 0, t) \in \mathbb{R}^4 \ / \ t \in \mathbb{R}\} = \text{vect}((-1,0,0,1)) = \text{vect}((1,0,0,-1))$

- **Pour $\lambda = 5$:**

$$(\Sigma) \iff \begin{cases} 4x + 4y + 4z + 9t &= 0 \\ -y + 5z &= 0 \\ y - 5z &= 0 \\ -5t &= 0 \end{cases}$$

$$\iff \begin{cases} t &= 0 \\ y &= 5z \\ x &= -6z \end{cases}$$

Le système n'est pas de Cramer: $\lambda = 5$ **est valeur propre de** S.

Le sous-espace propre associé est :

$E_5 = \{(-6z, 5z, z, 0) \in \mathbb{R}^4 \ / \ z \in \mathbb{R}\} = \text{vect}((-6,5,1,0))$

- **Pour $\lambda = 9$:**

$$(\Sigma) \iff \begin{cases} 4y + 4z + 9t &= 0 \\ -5y + 5z &= 0 \\ y - 9z &= 0 \\ -9t &= 0 \end{cases}$$

$$\iff \begin{cases} t = 0 & \\ y = -z & \text{d'après } L_1 \\ y = z & \text{d'après } L_2 \\ y = 9z & \text{d'après } L_3 \end{cases}$$

$$\iff \begin{cases} x &\in \mathbb{R} \\ y = z &= 0 \\ t &= 0 \end{cases}$$

Le système n'est pas de Cramer: $\lambda = 9$ **est valeur propre de** S.

Le sous-espace propre associé est :

$E_9 = \{(x, 0, 0, 0) \in \mathbb{R}^4 \ / \ x \in \mathbb{R}\} = \text{vect}((1,0,0,0))$

2–b)

Le calcul n'est pas nécessaire : supposons S associée à un endomorphisme s de $\mathbb{R}^4$ dans la base canonique B de $\mathbb{R}^4$. s admet 4 valeurs propres distinctes , donc s est diagonalisable puisque $\dim \mathbb{R}^4 = 4$ (c'est une condition suffisante). On obtient une base de $\mathbb{R}^4$ formée de vecteurs propres de s en réunissant les bases de E_{-1}, E_0, E_5 et E_9 trouvées dans le **2–a)**. Une telle base est $B' = \Big((0,1,-1,0), (1,0,0,-1), (-6,5,1,0), (1,0,0,0)\Big)$; dans cette base la matrice de s est :

$$D = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 9 \end{pmatrix}$$

La matrice de passage de la base B à la base B' est R. D'après la formule de changement de base pour les endomorphismes, on a $\boxed{D = R^{-1}SR}$

2–c)

L'égalité précédente s'écrit aussi $S = RDR^{-1}$ ce qui implique (c'est maintenant un résultat classique du cours qui peut se montrer facilement par récurrence)

$\forall n \in \mathbb{N}^*, \ S^n = RD^nR^{-1}$ (formule valable pour $n = 0$ avec la convention $S^0 = D^0 = I$)

Calculons explicitement : soit $n \in \mathbb{N}^*$,

$$
\begin{aligned}
S^n &= \begin{pmatrix} 0 & 1 & -6 & 1 \\ 1 & 0 & 5 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} (-1)^n & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 5^n & 0 \\ 0 & 0 & 0 & 9^n \end{pmatrix} \begin{pmatrix} 0 & \frac{1}{6} & -\frac{5}{6} & 0 \\ 0 & 0 & 0 & -1 \\ 0 & \frac{1}{6} & \frac{1}{6} & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 0 & 0 & -6\times 5^n & 9^n \\ (-1)^n & 0 & 5^{n+1} & 0 \\ (-1)^{n+1} & 0 & 5^n & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & \frac{1}{6} & -\frac{5}{6} & 0 \\ 0 & 0 & 0 & -1 \\ 0 & \frac{1}{6} & \frac{1}{6} & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}
\end{aligned}
$$

$$
\boxed{S^n = \begin{pmatrix} 9^n & -5^n + 9^n & -5^n + 9^n & 9^n \\ 0 & \frac{(-1)^n}{6} + \frac{5^{n+1}}{6} & (-1)^{n+1}\frac{5}{6} + \frac{5^{n+1}}{6} & 0 \\ 0 & \frac{(-1)^{n+1}}{6} + \frac{5^n}{6} & (-1)^n\frac{5}{6} + \frac{5^n}{6} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}}
$$

On, peut remarquer que cette formule qui donne explicitement S^n n'est pas valable pour $n = 0$ puisque l'on n'obtient pas I.

3–a)

On a vu que $(X_n = 0) \Longrightarrow (X_{n+1} = 0)$. Donc la variable X_{n+1} conditionnée par l'événement $(X_n = 0)$ ne prend que la valeur 0

$$\boxed{\forall n \in \mathbb{N},\ P_{(X_n=0)}(X_{n+1} = 0) = 1}$$

3–b)

Si $(X_n = 3)$, les 3 individus sont contagieux le jour n, ils sont donc sains le jour suivant d'après l'énoncé. Donc la variable X_{n+1} conditionnée par l'événement $(X_n = 3)$ ne prend que la valeur 0

$$\boxed{\forall n \in \mathbb{N},\ P_{(X_n=3)}(X_{n+1} = 0) = 1}$$

3–c)

Si $X_n = 1$, l'individu contagieux le jour n sera sain le jour suivant : il y aura au plus 2 contagieux le jour $(n + 1)$.

Sous la condition $(X_n = 1)$

- $(X_{n+1} = 0)$ signifie qu'aucun des deux individus sains le jour n n'a été contaminé (ce même jour bien-sûr). Par indépendance des contaminations,

$$P_{(X_n=1)}(X_{n+1} = 0) = \left(\frac{2}{3}\right)^2 = \frac{4}{9}$$

- $(X_{n+1} = 1)$ signifie qu'un seul des deux individus, que nous nommerons A et B, sains le jour n a été contaminé.

$$
\begin{aligned}
P_{(X_n=1)}(X_{n+1} = 1) &= P\Big((A \text{ est contaminé mais pas } B) \cup (B \text{ est contaminé mais pas } A)\Big) \\
&= P(A \text{ est contaminé mais pas } B) + P(B \text{ est contaminé mais pas } A) \\
&\quad \text{car les deux événements sont incompatibles} \\
&= \frac{1}{3} \times \frac{2}{3} + \frac{2}{3} \times \frac{1}{3} \quad \text{par indépendance des contaminations}
\end{aligned}
$$

$$P_{(X_n=1)}(X_{n+1} = 1) = \frac{4}{9}$$

- $(X_{n+1} = 2)$ signifie que les deux individus sains le jour n ont été contaminés ce même jour.

$$P_{(X_n=1)}(X_{n+1} = 2) = \frac{1}{9}$$

On a bien

$$
\begin{aligned}
P_{(X_n=1)}(X_{n+1} = 0) &= \binom{2}{0}\left(\frac{1}{3}\right)^0\left(\frac{2}{3}\right)^2 \\
P_{(X_n=1)}(X_{n+1} = 1) &= \binom{2}{1}\left(\frac{1}{3}\right)^1\left(\frac{2}{3}\right)^1 \\
P_{(X_n=1)}(X_{n+1} = 2) &= \binom{2}{2}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^0
\end{aligned}
$$

$$\boxed{\text{La variable } (X_{n+1}\ /\ X_n = 1) \text{ suit la loi binomiale } B(2; \tfrac{1}{3}) \text{ de paramètres 2 et } \tfrac{1}{3}}$$

Si $(X_n = 2)$, il y aura au plus un contagieux le jour suivant. Nommons C_1 et C_2 les deux individus contagieux le jour n et C l'individu sain le jour n.

Sous la condition $(X_n = 2)$

- $(X_{n+1} = 0)$ signifie que (C n' a pas été contaminé par C_1) et que (C n' a pas été contaminé par C_2). Par indépendance des contaminations,

$$P_{(X_n=2)}(X_{n+1} = 0) = \left(\frac{2}{3}\right)^2 = \frac{4}{9}$$

- $(X_{n+1} = 1)$ est l'événement contraire du précédent, donc

$$P_{(X_n=2)}(X_{n+1} = 1) = \frac{5}{9}$$

$$\boxed{\text{La variable } (X_{n+1}\ /\ X_n = 2) \text{ suit la loi de Bernoulli } B(\tfrac{5}{9}) \text{ de paramètre } \tfrac{5}{9}}$$

3–d)

C'est du cours

$$E\Big(X_{n+1}\ /\ X_n = 1) = 2 \times \frac{1}{3} = \boxed{\frac{2}{3}}$$

$$E\Big(X_{n+1}\ /\ X_n = 2) = 1 \times \frac{5}{9} = \boxed{\frac{5}{9}}$$

4–a)

$X_0(\Omega) = [\![0;3]\!]$ et $X_1(\Omega) \subset [\![0;3]\!]$

Utilisons le système complet d'événements $\{\Big(X_0 = i\Big)\ /\ 0 \leq i \leq 3\}$. D'après la formule des probabilités totales, pour tout $j \in [\![0;3]\!]$, on a

$$P(X_1 = j) = \sum_{i=0}^{3} P(X_0 = i) P_{(X_0=i)}(X_1 = j) \text{ avec } P(X_0 = i) = \begin{cases} \left(\frac{2}{3}\right)^3 = \frac{8}{27} & \text{si } i = 0 \\ 3 \times \frac{1}{3} \times \left(\frac{2}{3}\right)^2 = \frac{4}{9} & \text{si } i = 1 \\ 3 \times \left(\frac{1}{3}\right)^2 \times \frac{2}{3} = \frac{2}{9} & \text{si } i = 2 \\ \left(\frac{1}{3}\right)^3 = \frac{1}{27} & \text{si } i = 3 \end{cases},$$

car la variable X_0 suit la loi binomiale $B(3, \frac{1}{3})$.

- $$
\begin{aligned}
P_{(X_0=0)}(X_1 = 0) &= 1 && \text{d'après 3–a)} \\
P_{(X_0=1)}(X_1 = 0) &= \frac{4}{9} && \text{d'après 3–c)} \\
P_{(X_0=2)}(X_1 = 0) &= \frac{4}{9} && \text{d'après 3–c)} \\
P_{(X_0=3)}(X_1 = 0) &= 1 && \text{d'après 3–b)}
\end{aligned}
$$

$$\text{Donc, } P(X_1 = 0) = 1 \times \frac{8}{27} + \frac{4}{9} \times \frac{4}{9} + \frac{4}{9} \times \frac{2}{9} + 1 \times \frac{1}{27} = \frac{51}{81} = \boxed{\frac{17}{27}}$$

- $$\begin{aligned} P_{(X_0=0)}(X_1=1) = P_{(X_0=3)}(X_1=1) &= 0 \quad \text{d'après 3−a) et 3−b)} \\ P_{(X_0=1)}(X_1=1) &= 2\times\frac{1}{3}\times\frac{2}{3}=\frac{4}{9} \quad \text{d'après 3−c)} \\ P_{(X_0=2)}(X_1=1) &= \frac{5}{9} \quad \text{d'après 3−c)} \end{aligned}$$

$$\text{Donc, } P(X_1=1)=\frac{4}{9}\times\frac{4}{9}+\frac{2}{9}\times\frac{5}{9}=\boxed{\frac{26}{81}}$$

- $$\begin{aligned} P_{(X_0=0)}(X_1=2) = P_{(X_0=3)}(X_1=2) &= 0 \quad \text{d'après 3−a) et 3−b)} \\ P_{(X_0=1)}(X_1=2) &= \left(\frac{1}{3}\right)^2=\frac{1}{9} \quad \text{d'après 3−c)} \\ P_{(X_0=2)} = (X_1=2) &= 0 \quad \text{d'après 3−c)} \end{aligned}$$

$$\text{Donc, } P(X_1=2)=\frac{1}{9}\times\frac{4}{9}=\boxed{\frac{4}{81}}$$

- Remarquons que l'on n'a jamais $P(X_{n+1}=3)$ d'après la question 3).

$$\text{Donc, } P(X_1=3)=\boxed{0}$$

$$\boxed{E(X_1)=1\times\frac{26}{81}+2\times\frac{4}{81}=\frac{34}{81}}$$

4−b)

$$\begin{aligned} E(X_1\ /\ X_0=0)&=0 \quad ; \quad P(X_0=0)=\left(\frac{2}{3}\right)^3=\frac{8}{27} \\ E(X_1\ /\ X_0=1)&=\frac{2}{3} \quad ; \quad P(X_0=1)=3\times\left(\frac{2}{3}\right)^2\left(\frac{1}{3}\right)=\frac{4}{9} \\ E(X_1\ /\ X_0=2)&=\frac{5}{9} \quad ; \quad P(X_0=2)=3\times\left(\frac{2}{3}\right)\left(\frac{1}{3}\right)^2=\frac{2}{9} \\ E(X_1\ /\ X_0=3)&=0 \quad ; \quad P(X_0=3)=\left(\frac{1}{3}\right)^3=\frac{1}{27} \end{aligned}$$

(On a utilisé les résultats du 3)).

$$\boxed{\sum_{i=0}^{3}E(X_1\ /\ X_0=i)\times P(X_0=i)=\frac{2}{3}\times\frac{4}{9}+\frac{5}{9}\times\frac{2}{9}=\frac{8}{27}+\frac{10}{81}=\frac{34}{81}=E(X_1)}$$

Remarque : Cette égalité s'appelle la formule de l'espérance totale.

5−a)

La famille $\{(X_n=i)\ /\ 0\leq i\leq 3\}$ est un système complet d'événements, donc $\boxed{u_n+v_n+w_n+t_n=1}$

5−b)

Il en résulte que : $\forall j\in[\![0;3]\!],\ P(X_{n+1}=j)=\displaystyle\sum_{i=0}^{3}P_{(X_n=i)}(X_{n+1}=j)P(X_n=i)$

$$\begin{aligned} P(X_{n+1}=0)&=1\times u_n+\frac{4}{9}\times v_n+\frac{4}{9}\times w_n+1\times t_n \\ P(X_{n+1}=1)&=0\times u_n+\frac{4}{9}\times v_n+\frac{5}{9}\times w_n+0\times t_n \\ P(X_{n+1}=2)&=0\times u_n+\frac{1}{9}\times v_n+0\times w_n+0\times t_n \\ P(X_{n+1}=3)&=0, \end{aligned}$$

toujours à l'aide des résultats du 3). Matriciellement, on obtient $\begin{pmatrix}u_{n+1}\\v_{n+1}\\w_{n+1}\\t_{n+1}\end{pmatrix}=M\begin{pmatrix}u_n\\v_n\\w_n\\t_n\end{pmatrix}$ avec

$$\boxed{M=\begin{pmatrix}1&\frac{4}{9}&\frac{4}{9}&1\\0&\frac{4}{9}&\frac{5}{9}&0\\0&\frac{1}{9}&0&0\\0&0&0&0\end{pmatrix}}$$

5−c)

Il est alors clair que $\boxed{M=\frac{1}{9}S}$

Un réel λ est valeur propre de M si et seulement si il existe une matrice colonne X non nulle appartenant à $\mathcal{M}_{4,1}(\mathbb{R})$ telle que $MX=\lambda X$; ce qui équivaut à $SX=(9\lambda)X$, c'est-à-dire que 9λ est valeur propre de S.

$\lambda\in\text{spect}(M)\Longleftrightarrow 9\lambda\in\{-1,0,5,9\}$, donc $\boxed{\text{spect}(M)=\{-\frac{1}{9},0,\frac{5}{9},1\}}$

5−d)

Par une récurrence classique que nous ne ferons pas, on montre que $\boxed{\forall n\in\mathbb{N},\ U_n=M^nU_0}$

$$\text{Donc } U_n=\frac{1}{9^n}S^nU_0=\begin{pmatrix}1&1-\left(\frac{5}{9}\right)^n&1-\left(\frac{5}{9}\right)^n&1\\0&\frac{1}{6}\left(\frac{-1}{9}\right)^n+\frac{5}{6}\left(\frac{5}{9}\right)^n&-\frac{5}{6}\left(\frac{-1}{9}\right)^n+\frac{5}{6}\left(\frac{5}{9}\right)^n&0\\0&-\frac{1}{6}\left(\frac{-1}{9}\right)^n+\frac{1}{6}\left(\frac{5}{9}\right)^n&\frac{5}{6}\left(\frac{-1}{9}\right)^n+\frac{1}{6}\left(\frac{5}{9}\right)^n&0\\0&0&0&0\end{pmatrix}\begin{pmatrix}u_0\\v_0\\w_0\\t_0\end{pmatrix}$$

D'où

$$\begin{aligned} u_n&=u_0+(1-\left(\frac{5}{9}\right)^n)v_0+(1-\left(\frac{5}{9}\right)^n)w_0+t_0 \\ &=u_0+v_0+w_0+t_0-\left(\frac{5}{9}\right)^nv_0-\left(\frac{5}{9}\right)^nw_0 \end{aligned}$$

$$\boxed{u_n=1-\left(\frac{5}{9}\right)^nv_0-\left(\frac{5}{9}\right)^nw_0}$$

$$v_n=(\frac{1}{6}\left(\frac{-1}{9}\right)^n+\frac{5}{6}\left(\frac{5}{9}\right)^n)v_0+(-\frac{5}{6}\left(\frac{-1}{9}\right)^n+\frac{5}{6}\left(\frac{5}{9}\right)^n)w_0$$

$$\boxed{v_n=\frac{5}{6}\left(\frac{5}{9}\right)^n(v_0+w_0)+\frac{1}{6}\left(\frac{-1}{9}\right)^n(v_0-5w_0)}$$

6−a)

$F=\displaystyle\bigcup_{n=0}^{+\infty}(X_n=0)$. Dire que l'événement F est réalisé c'est dire, par définition d'une réunion, qu'il existe une entier $n\in\mathbb{N}$ tel que l'événement $(X_n=0)$ est réalisé. Cela veut dire qu'il arrive un jour, le jour n, où le virus a disparu.

L'événement F est l'événement " le virus a disparu "

6−b)

On a vu que $(X_n=0)\subset(X_{n+1}=0)$ car toute éventualité qui réalise $(X_n=0)$ réalise aussi $(X_{n+1}=0)$.

F est alors l'union d'une suite d'événements, croissante pour l'inclusion. **D'après le théorème de la limite monotone,**

$P(F)=\displaystyle\lim_{n\to+\infty}P(X_n=0)=\lim_{n\to+\infty}u_n=1$ car

$u_n = 1 - \left(\frac{5}{9}\right)^n v_0 - \left(\frac{5}{9}\right)^n w_0$ et $\left|\frac{5}{9}\right| < 1 \Longrightarrow \lim\limits_{n\to+\infty} \left(\frac{5}{9}\right)^n = 0$

Remarque : Ce résultat ne dépend pas des valeurs $u_0, \ldots, t_0$

$\boxed{P(F) = 1 \text{ veut dire que l'on est quasiment certain que le virus va disparaître}}$

Partie II

1–a)

- $q_{0,j} = P_{(X_n=0)}(X_{n+1} = j) = \begin{cases} 0 & \text{si } j > 0 \\ 1 & \text{si } j = 0 \end{cases}$

car si les individus sont tous sains le jour n, ils le sont encore le jour suivant.

- $q_{N,j} = P_{(X_n=N)}(X_{n+1} = j) = \begin{cases} 0 & \text{si } j > 0 \\ 1 & \text{si } j = 0 \end{cases}$

car si les individus sont tous contagieux le jour n, ils sont sains le jour suivant.

- $q_{i,N} = P_{(X_n=i)}(X_{n+1} = N) = 0$

car si $i > 0$ (il y a i individus contagieux le jour n), ces i individus seront sains le jour suivant : on n'aura pas $X_{n+1} = N$

Si $i = 0$, on a vu l'explication au début du a). On a $(X_n = i) \cap (X_{n+1} = N) = \emptyset$

Remarque : Ceci prouve que l'on n'a pas : $\forall n \in \mathbb{N},\ \forall i \in [\![0; N]\!],\ P(X_n = i) > 0$.

L'énoncé comporte une erreur, car $P(X_{n+1} = N) = \sum\limits_{i=0}^{N} P_{(X_n=i)}(X_{n+1} = N)P(X_n = i) = 0$;

le vrai énoncé doit être $\forall n \in \mathbb{N},\ \forall i \in [\![0; N[\![,\ P(X_n = i) > 0$.

1–b)

Si i individus sont contagieux le jour n, ils sont sains le jour suivant. Donc ce jour-là, numéro $n+1$, il ne peut y avoir, au plus, que $N - i$ individus contagieux.

$\boxed{\text{Donc, si } j > N - i, \text{ la probabilité } q_{i,j} = 0}$

1–c)

D'après ce qui précède, la variable X_{n+1} conditionnée par l'événement $(X_n = i)$ ne peut prendre que les valeurs que $[\![0, N - i]\!]$.

L'observation de $N-i$ individus, autres que les i contagieux, est une suite de $N-i$ épreuves indépendantes, identiques, à deux issues possibles : le succès S " l'individu observé est contagieux le jour $n+1$ " et l'échec $\overline{S}$ "l'individu observé est sain le jour $n+1$ ".

La variable X_{n+1} conditionnée par l'événement $(X_n = i)$ indique le nombre de succès.

Cette variable aléatoire suit donc la loi binomiale de paramètres $N - i$ et $P(S)$.

Soit $C_1, \ldots, C_i$ les individus contagieux le jour n et soit I un individu sain le jour n.

L'événement $\overline{S}$: l'individu I est sain le jour $n+1$ est l'intersection des événements indépendants suivants I_k : I n'est pas contaminé par C_k, pour $k \in [\![1, i]\!]$

$$\overline{S} = \bigcap_{k=1}^{k} I_k.$$

Or la probabilité que I ne soit pas contaminé par C_k le jour n est $q = 1 - p$. Il en résulte que $P(\overline{S}) = q^i$ et donc $P(S) = 1 - q^i$

$\boxed{X_{n+1} \text{ conditionnée par } (X_n = i), \text{ notée } X_{n+1} \ / \ (X_n = i), \text{ suit la loi binomiale } B(N - i, 1 - q^i)}$

2–a)

$$\begin{aligned}
\sum_{j=0}^{N} q_{i,j} &= \sum_{j=0}^{N} P_{(X_n=i)}(X_{n+1} = j) \\
&= \sum_{j=0}^{N} \frac{P(X_n = i \cap X_{n+1} = j)}{P(X_n = i)} \\
&= \frac{1}{P(X_n = i)} \sum_{j=0}^{N} P((X_n = i \cap X_{n+1} = j) \\
&= \frac{P(X_n = i)}{P(X_n = i)} \quad \text{d'après la formule des probabilités totales}
\end{aligned}$$

$$\text{Donc } \sum_{j=0}^{N} q_{i,j} = 1$$

Remarque 1 : Ce résultat est valable si $P(X_n = i) \neq 0$ et cette situation est fausse si $i = N$ et $n > 0$

Si $i = N$ et $n > 0$, alors $(X_n = N) \Longrightarrow (X_{n+1} = 0)$, donc $P_{(X_n=N)}(X_{n+1} = j) = 0$ pour $j \in [\![1, N]\!]$ et $P_{(X_n=N)}(X_{n+1} = 0) = 1$

Dans ce cas là, $\sum\limits_{j=0}^{N} q_{N,j} = \sum\limits_{j=0}^{N} P_{(X_n=N)}(X_{n+1} = j) = P_{(X_n=N)}(X_{n+1} = 0) = 1$

Donc le résultat est encore valable et on peut dire :

$$\boxed{\forall i \in [\![0, N]\!],\ \sum_{j=0}^{N} q_{i,j} = 1}$$

Remarque 2 : Dans les cas $P(X_n = i) \neq 0$, on aurait pu obtenir cette égalité en disant que l'application $P_{(X_n=i)}$ est une probabilité, donc $\sum\limits_{j=0}^{N} P_{(X_n=i)}(X_{n+1} = j) = 1$ puisque la famille $(X_{n+1} = j)\ /\ 0 \leq j \leq N$ est un système complet d'événements.

La somme des termes de chaque ligne de Q vaut donc 1. Donc $Q \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} \sum\limits_{j=0}^{N} q_{1,j} \\ \vdots \\ \sum\limits_{j=0}^{N} q_{N,j} \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$

Ceci prouve que $\boxed{1 \text{ est valeur propre de } Q}$ et que $\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathcal{M}_{n+1,1}(\mathbb{R})$ en est un vecteur propre.

2–b)

Un vecteur propre étant un vecteur non nul, il existe au moins une coordonnée non nulle ; les valeurs absolues des coordonnées sont positives ou nulles et il y en a au moins une qui est **strictement positive** (sinon toutes les coordonnées seraient nulles). Si l'on considère la coordonnée qui a la plus grande valeur absolue, cette coordonnée est non nulle, puisque sa valeur absolue est strictement positive.

Si l'on note $V(i)$ cette coordonnée, on a $|V(i)| = \max\limits_{0 \leq j \leq N} |V(j)|$ et $V(i) \neq 0$.

La ligne numéro i du système $QV = \lambda V$ s'écrit $\sum\limits_{j=0}^{N} q_{i,j}V(j) = \lambda V(i)$, ce qui implique successivement :

$$\begin{aligned}
|\lambda V(i)| &= |\sum_{j=0}^{N} q_{i,j}V(j)| \\
|\lambda||V(i)| &\le \sum_{j=0}^{N} |q_{i,j}V(j)| \quad \text{inégalité triangulaire} \\
|\lambda||V(i)| &\le \sum_{j=0}^{N} q_{i,j}|V(j)| \quad \text{car } q_{i,j} \ge 0 \\
|\lambda| &\le \sum_{j=0}^{N} q_{i,j}\frac{|V(j)|}{|V(i)|} \quad \text{car } |V(i)| > 0 \\
|\lambda| &\le \sum_{j=0}^{N} q_{i,j} \quad \text{car } \frac{|V(j)|}{|V(i)|} \le 1 \text{ et } q_{i,j} \ge 0
\end{aligned}$$

$$\boxed{\text{Si } \lambda \text{ est une valeur propre de } Q, \text{ alors } |\lambda| \le 1 \text{ car } \sum_{j=0}^{N} q_{i,j} = 1 \text{ d'après 2–a)}}$$

3)

La famille $\{(X_n = 0), \ldots, (X_n = N)\}$ est un (pseudo) système complet d'événements (pseudo puisque $P(X_n = N)$ peut être nulle).

$$\begin{aligned}
\forall k \in [\![0, N]\!],\ P(X_{n+1} = k) &= \sum_{j=0}^{N} P_{(X_n=j)}(X_{n+1} = k)P(X_n = j) \\
&= \sum_{j=0}^{N} q_{j,k}P(X_n = j) \\
&= q_{0,k}P(X_n = 0) + q_{1,k}P(X_n = 1) + \ldots, q_{N,k}P(X_n = N) \\
&= \begin{pmatrix} q_{0,k} & \ldots & q_{N,k} \end{pmatrix} \begin{pmatrix} P(X_n = 0) \\ \vdots \\ P(X_n = N) \end{pmatrix}
\end{aligned}$$

produit d'une matrice ligne et d'une matrice colonne

Donc, $\begin{pmatrix} P(X_{n+1} = 0) \\ \vdots \\ \vdots \\ P(X_{n+1} = N) \end{pmatrix} = \begin{pmatrix} q_{0,0} \cdots & \cdots & q_{N,0} \\ \vdots & & \vdots \\ q_{0,k} \cdots & \cdots & q_{N,k} \\ \vdots & & \vdots \\ q_{0,N} \cdots & \cdots & q_{N,N} \end{pmatrix} \begin{pmatrix} P(X_n = 0) \\ \vdots \\ \vdots \\ P(X_n = N) \end{pmatrix}$

$$\boxed{\text{On constate que } U_{n+1} = {}^tQU_n}$$

4–a)

Par des récurrences classiques que nous ne ferons pas, on montre que

$$\forall n \in \mathbb{N},\ U_n = M^nU_0 \text{ et } \forall k \in [\![0; N]\!],\ \forall n \in \mathbb{N},\ M^nV_k = \lambda_k^nV_k$$

Donc, $\forall n \in \mathbb{N},\ U_n = M^n \sum_{k=0}^{N} \alpha_kV_k = \sum_{k=0}^{N} \alpha_kM^nV_k$

$$\boxed{\forall n \in \mathbb{N},\ U_n = \sum_{k=0}^{N} \alpha_k\lambda_k^nV_k}$$

4–b)

La colonne des coordonnées de V_k est $\begin{pmatrix} V_k(0) \\ \vdots \\ V_k(i) \\ \vdots \\ V_k(N) \end{pmatrix}$, donc la colonne des coordonnées de $\alpha_k\lambda_k^nV_k$ est $\begin{pmatrix} \alpha_k\lambda_k^nV_k(0) \\ \vdots \\ \alpha_k\lambda_k^nV_k(i) \\ \vdots \\ \alpha_k\lambda_k^nV_k(N) \end{pmatrix}$, donc celle de $\sum_{k=0}^{N} \alpha_k\lambda_k^nV_k$ est $\begin{pmatrix} \sum_{k=0}^{N} \alpha_k\lambda_k^nV_k(0) \\ \vdots \\ \sum_{k=0}^{N} \alpha_k\lambda_k^nV_k(i) \\ \vdots \\ \sum_{k=0}^{N} \alpha_k\lambda_k^nV_k(N) \end{pmatrix}$

La colonne des coordonnées de U_n est $\begin{pmatrix} P(X_n = 0) \\ \vdots \\ P(X_n = i) \\ \vdots \\ P(X_n = N) \end{pmatrix}$

$$\boxed{\text{L'égalité : } \forall n \in \mathbb{N},\ U_n = \sum_{k=0}^{N} \alpha_k\lambda_k^nV_k \text{ donne } \forall n \in \mathbb{N},\ \forall i \in \mathbb{N},\ P(X_n = i) = \sum_{k=0}^{N} \alpha_k\lambda_k^nV_k(i)}$$

4–c)

$P(X_n = i) = \alpha_0V_0(i) + \sum_{k=1}^{N} \alpha_k\lambda_k^nV_k(i)$ car $\lambda_0 = 1$

$\forall k \in [\![1, N]\!]$, on a $|\lambda_k| < 1$, donc $\lim_{n\to+\infty} \lambda_k^n = 0$

$$\boxed{\text{Il en résulte que : } \forall k \in [\![1, N]\!],\ \forall i \in [\![0, N]\!],\ \lim_{n\to+\infty} P(X_n = i) = \alpha_0V_0(i)}$$

D'après l'énoncé, $V_0(i) = 1$ si $i = 0$ et $V_0(i) = 0$ pour $1 \le i \le N$

$$\boxed{\text{Donc, } \forall i \in [\![1, N]\!],\ \lim_{n\to+\infty} P(X_n = i) = 0}$$

4–d)

La famille $\{(X_n = 0), \ldots, (X_n = N)\}$ est un (pseudo) système complet d'événements, on a :

$$\sum_{i=0}^{N} P(X_n = i) = 1 = P(X_n = 0) + \sum_{i=1}^{N} P(X_n = i)$$

D'après le résultat précédent, $\lim_{n\to+\infty} \sum_{i=1}^{N} P(X_n = i) = 0$, donc $\boxed{\lim_{n\to+\infty} P(X_n = 0) = 1}$

Cela veut dire que l'on est quasiment certain que le virus va disparaître.

Partie III

1–a)

Les variables aléatoires Y_i étant indépendantes et suivant la même loi de Bernoulli de paramètre p, la somme $\displaystyle\sum_{i=1}^{m} Y_i$ suit la loi binomiale $B(m,p)$ de paramètres m et p. Donc $E\Big(\displaystyle\sum_{i=1}^{m} Y_i\Big) = mp$ et $E\Big(\overline{Y_m}\Big) = \dfrac{1}{m}mp = p$

$\overline{Y_m}$ est une fonction des Y_i pour $1 \leq i \leq m$, donc $\overline{Y_m}$ **est un estimateur de** p

$$\boxed{\overline{Y_m} \text{ est un estimateur sans biais de } p}$$

Le risque quadratique est donc égal à $V(\overline{Y_m}) = \dfrac{1}{m^2}V\Big(\displaystyle\sum_{i=1}^{m} Y_i\Big) = \dfrac{1}{m^2}mp(1-p) = \boxed{\dfrac{p(1-p)}{m}}$

1–b)

On veut montrer que $P\Big(\overline{Y_m} - \sqrt{\dfrac{5}{m}} \leq p \leq \overline{Y_m} + \sqrt{\dfrac{5}{m}}\Big) \geq 0.95$

ceci équivaut à $P\Big(|\overline{Y_m} - p| \leq \sqrt{\dfrac{5}{m}}\Big) \geq 0.95$, puis à $-P\Big(|\overline{Y_m} - p| \leq \sqrt{\dfrac{5}{m}}\Big) \leq -0.95$, c'est-à-dire $1 - P\Big(|\overline{Y_m} - p| \leq \sqrt{\dfrac{5}{m}}\Big) \leq 0.05$, et finalement $P\Big(|\overline{Y_m} - p| > \sqrt{\dfrac{5}{m}}\Big) \leq 0.05$.

D'après Bienaymé-Tchebycheff, $P\Big(|\overline{Y_m} - p| > \sqrt{\dfrac{5}{m}}\Big) \leq \dfrac{V(\overline{Y_m})}{\Big(\sqrt{\dfrac{5}{m}}\Big)^2}$ **(1)**

Or $\dfrac{V(\overline{Y_m})}{\Big(\sqrt{\dfrac{5}{m}}\Big)^2} = \dfrac{p(1-p)}{m \times \dfrac{5}{m}} = \dfrac{p(1-p)}{5}$

D'autre part, pour tout $p \in [0;1]$, $p(1-p) \leq \dfrac{1}{4}$ (inégalité classique dite de Bernoulli)

Montrons la :

Posons $f(x) = x(1-x)$ pour $x \in [0;1]$; $f'(x) = 1 - 2x$ d'où le tableau de variations :

x	0	$\frac{1}{2}$	1
$f'(x)$	$+$	0	$-$
f	$\nearrow$	$\frac{1}{4}$	$\searrow$

On a immédiatement $\forall x \in [0;1],\ x(1-x) \leq \dfrac{1}{4}$ et comme $p \in]0;1[$, on a bien $p(1-p) \leq \dfrac{1}{4}$

Par suite, $\dfrac{V(\overline{Y_m})}{\Big(\sqrt{\dfrac{5}{m}}\Big)^2} \leq \dfrac{1}{4 \times 5} = \dfrac{1}{20} = 0.05$

L'inégalité **(1)** donne $P\Big(|\overline{Y_m} - p| > \sqrt{\dfrac{5}{m}}\Big) \leq 0.05$

donc

$$\boxed{P\Big(\overline{Y_m} - \sqrt{\dfrac{5}{m}} \leq p \leq \overline{Y_m} + \sqrt{\dfrac{5}{m}}\Big) \geq 0.95}$$

2–a)

$$\begin{aligned}
\overline{Y_m} - p \geq \varepsilon &\iff m\theta(\overline{Y_m} - p) \geq m\theta\varepsilon \quad \text{car } m\theta > 0\\
&\iff m\theta\overline{Y_m} \geq m\theta(p+\varepsilon)\\
&\iff \exp(m\theta\overline{Y_m}) \geq \exp(m\theta(p+\varepsilon)) \quad \text{car exp est strictement croissante}
\end{aligned}$$

Les événements $(\overline{Y_m} - p \geq \varepsilon)$ et $\Big(\exp(m\theta\overline{Y_m}) \geq \exp(m\theta(p+\varepsilon))\Big)$ sont donc égaux : ils ont même probabilité

$$\boxed{\forall\theta > 0,\ \forall\varepsilon > 0,\ P(\overline{Y_m} - p \geq \varepsilon) = P\Big(\exp(m\theta\overline{Y_m}) \geq \exp(m\theta(p+\varepsilon))\Big)}$$

2–b)

Posons $T(\Omega) = \{t_1, \ldots, t_n\}$, alors $E(T) = \displaystyle\sum_{k=1}^{n} t_k P(T = t_k)$

Posons $I_a = \{k \in [\![1,n]\!] \ /\ t_k \geq a\}$ et $\overline{I_a} = \{k \in [\![1,n]\!] \ /\ t_k < a\}$. Les deux ensembles I_a et $\overline{I_a}$ forment une partition de $[\![1,n]\!]$, l'un de ces deux ensembles pouvant être vide.

$E(T) = \displaystyle\sum_{k\in I_a} t_k P(T = t_k) + \sum_{k\in \overline{I_a}} t_k P(T = t_k)$ avec la convention que si un des deux ensembles I_a ou $\overline{I_a}$ est vide, la somme correspondante est nulle.

La variable T ne prend que des valeurs positives, donc $\displaystyle\sum_{k\in \overline{I_a}} t_k P(T = t_k) \geq 0$ et il s'ensuit

$$E(T) \geq \sum_{k\in I_a} t_k P(T = t_k)$$

$\forall k \in I_a,\ t_k \geq a$, donc $t_k P(T = t_k) \geq aP(T = t_k)$ car $P(T = t_k) \geq 0$.

Sommons ces inégalités pour $k \in I_a$, il vient $\displaystyle\sum_{k\in I_a} t_k P(T = t_k) \geq a\sum_{k\in I_a} P(T = t_k)$, c'est-à-dire $\displaystyle\sum_{k\in I_a} t_k P(T = t_k) \geq aP(T \in I_a)$, puis $E(T) \geq aP(T \geq a)$.

Et comme $a > 0$, on a :

$$\boxed{\forall a > 0,\ P(T \geq a) \leq \dfrac{E(T)}{a}}$$

2–c)

La variable aléatoire $\exp(m\theta\overline{Y_m})$ prend des valeurs positives strictement par propriété de l'exponentielle. les variables Y_i ne prennent que deux valeurs 0 ou 1, donc $\overline{Y_m}$ ne prend qu'un nombre fini de valeurs et par conséquent $\exp(m\theta\overline{Y_m})$ également. On peut donc appliquer à cette variable l'inégalité de 2–b). Cette inégalité s'écrit :

$$P\Big(\exp(m\theta\overline{Y_m}) \geq \exp(m\theta(p+\varepsilon))\Big) \leq \dfrac{E(\exp(m\theta\overline{Y_m})}{\exp(m\theta(p+\varepsilon))}$$

Calculons :

$$\begin{aligned}
E(\exp(m\theta\overline{Y_m}) &= E(\exp(m\theta\frac{1}{m}\sum_{i=1}^{m} Y_i))\\
&= E(\exp(\theta\sum_{i=1}^{m} Y_i))\\
&= E(\prod_{i=1}^{m}\exp(\theta Y_i))
\end{aligned}$$

Les variables Y_i sont indépendantes, donc les variables $\exp(\theta Y_i)$ aussi en tant que fonctions des Y_i ; il en résulte que $E(\exp(m\theta\overline{Y_m})) = E(\displaystyle\prod_{i=1}^{m}\exp(\theta Y_i)) = \prod_{i=1}^{m} E(\exp(\theta Y_i))$.

D'après le théorème du tranfert,

$E(\exp(\theta Y_i)) = \exp(\theta \times 0) \times P(Y_i = 0) + \exp(\theta \times 1) \times P(Y_i = 1) = q + p\exp(\theta)$

Il en résulte que $E(\exp(m\theta\overline{Y_m})) = (q + p\exp(\theta))^m$ et

$$\boxed{\forall(\theta,\varepsilon)\in(\mathbb{R}_+^*)^2,\ P(\overline{Y_m}-p\geq\varepsilon)\leq\frac{(q+p\exp(\theta))^m}{\exp(m\theta(p+\varepsilon))}}$$

Ceci s'écrit aussi

$P(\overline{Y_m}-p\geq\varepsilon)\leq\dfrac{\exp\Big(m\ln(q+p\exp(\theta))\Big)}{\exp(m\theta(p+\varepsilon))}$ ou

$$\boxed{P(\overline{Y_m}-p\geq\varepsilon)\leq\exp\Big(m(g(\theta)-\theta(p+\varepsilon))\Big)}$$

2–d)

L'application $x\longmapsto pe^x+q$ est de classe C^2 sur $\mathbb{R}_+$, comme combinaison linéaire de fonctions de classe C^2 sur $\mathbb{R}_+$ **et à valeurs dans** $\mathbb{R}_+^*$. La fonction ln est de classe C^2 **sur** $\mathbb{R}_+^*$, donc par composition, la fonction $g\ :\ x\longmapsto\ln(pe^x+q)$ est de classe C^2 sur $\mathbb{R}_+$

$\forall x\geq 0,\ g'(x)=\dfrac{pe^x}{q+pe^x}$ et $\boxed{g''(x)=\dfrac{pqe^x}{(pe^x+q)^2}}$

$$\begin{aligned}|g''(x)|\leq\frac{1}{4}&\iff g''(x)\leq\frac{1}{4}\quad\text{car } g''(x)>0\\&\iff\frac{pqe^x}{(pe^x+q)^2}\leq\frac{1}{4}\\&\iff 4pqe^x\leq(pe^x+q)^2\quad\text{on a multiplié par } 4(pe^x+q)^2>0\\&\iff 0\leq(pe^x)^2-2pqe^x+q^2\\&\iff 0\leq(pe^x-q)^2,\quad\text{ce qui est toujours vrai}\end{aligned}$$

$$\boxed{\text{Conclusion : comme il s'agit d'équivalences, }\forall x\geq 0,\ |g''(x)|\leq\frac{1}{4}}$$

2–e)

La fonction g étant de classe C^2 sur $\mathbb{R}_+$, et $|g''|$ étant majorée sur $\mathbb{R}_+$ par $\frac{1}{4}$, on peut appliquer l'inégalité de Taylor-Lagrange à l'ordre 1 sur le segment $[0,\theta]$. On obtient

$$|g(\theta)-g(0)-\theta g'(0)|\leq\frac{\theta^2}{2}\max_{t\in[0;\theta]}|g''(t)|$$

$$\text{ou encore } |g(\theta)-g(0)-\theta g'(0)|\leq\frac{\theta^2}{8}\text{ puisque }\max_{t\in[0;\theta]}|g''(t)|\leq\frac{1}{4}$$

Rappelons qu'un réel est toujours inférieur ou égal à sa valeur absolue, donc

$g(\theta)-g(0)-\theta g'(0)\leq|g(\theta)-g(0)-\theta g'(0)|$, soit encore $g(\theta)-g(0)-\theta g'(0)\leq\dfrac{\theta^2}{8}$

Remarquons que $g(0)=\ln(p+q)=0$ puisque $p+q=1$ et $g'(0)=\dfrac{p}{p+q}=p$. l'inégalité précédente donne alors

$$\boxed{\forall\theta>0,\ g(\theta)\leq p\theta+\frac{\theta^2}{8}}$$

2–f)

$\forall x>0,\ h'(x)=\frac{x}{4}-\varepsilon$. On obtient le tableau de variations suivant :

x	0		4ε		$+\infty$
$h'(x)$		$-$	0	$+$	
h	0	$\searrow$	$-2\varepsilon^2$	$\nearrow$	$+\infty$

$\forall\theta\geq 0,\ g(\theta)-\theta(p+\varepsilon)\leq\theta p+\dfrac{\theta^2}{8}-\theta(p+\varepsilon)$ d'après 2–e), soit encore

$\forall\theta\geq 0,\ g(\theta)-\theta(p+\varepsilon)\leq\dfrac{\theta^2}{8}-\theta\varepsilon$

Comme $m>0,\ \forall\theta\geq 0,\ m(g(\theta)-\theta(p+\varepsilon))\leq m(\dfrac{\theta^2}{8}-\theta\varepsilon)$, et la fonction exp étant croissante,

$\forall\theta\geq 0,\ \exp(m(g(\theta)-\theta(p+\varepsilon)))\leq\exp(m(\dfrac{\theta^2}{8}-\theta\varepsilon))$, soit finalement

$\forall\theta\geq 0,\ \exp(m(g(\theta)-\theta(p+\varepsilon)))\leq\exp(mh(\theta))$

En particulier, pour $\theta=4\varepsilon$, $P(\overline{Y_m}-p\geq\varepsilon)\leq\exp(mh(4\varepsilon))$ d'après 2–c), donc

$$\boxed{P(\overline{Y_m}-p\geq\varepsilon)\leq\exp(-2m\varepsilon^2)}$$

3)

Posons $W_i=1-Y_i$ pour $i\in[\![1,m]\!]$. $Y_i(\Omega)=\{0,1\}\Longrightarrow W_i(\Omega)=\{0,1\}$

$P(W_i=1)=P(Y_i=0)=q$.

Les variables aléatoires W_i suivent la même loi de Bernoulli de paramètre q, elles sont indépendantes car les Y_i le sont. On peut donc leur appliquer l'inégalité du 2–f).

$$\boxed{\forall\varepsilon>0,\ P(\overline{W_m}-q\geq\varepsilon)\leq\exp(-2m\varepsilon^2)}$$

4–a)

$$(|\overline{Y_m}-p|\geq\varepsilon)=\underbrace{(\overline{Y_m}-p\geq\varepsilon)}_{\text{cas où }\overline{Y_m}-p\geq 0}\cup\underbrace{(-(\overline{Y_m}-p)\geq\varepsilon)}_{\text{cas où }\overline{Y_m}-p\leq 0}$$

$$=(\overline{Y_m}-p\geq\varepsilon)\cup(\overline{Y_m}-p)\leq-\varepsilon)$$

Remarquons que :

$$\begin{aligned}\overline{W_m}&=\frac{1}{m}\sum_{i=1}^m(1-Y_i)\\&=\frac{1}{m}\Big(\sum_{i=1}^m 1-\sum_{i=1}^m Y_i)\\&=\frac{1}{m}(m-\sum_{i=1}^m Y_i)\\&=1-\frac{1}{m}\sum_{i=1}^m Y_i\\&=1-\overline{Y_m}\end{aligned}$$

Donc

$$\begin{aligned}(\overline{Y_m}-p\leq-\varepsilon)&=(p-\overline{Y_m}\geq\varepsilon)\\&=(p-(1-\overline{W_m})\geq\varepsilon)\\&=(p-1+\overline{W_m}\geq\varepsilon)\\&=(\overline{W_m}-q\geq\varepsilon)\end{aligned}$$

Donc $(|\overline{Y_m}-p|\geq\varepsilon)=(\overline{Y_m}-p\geq\varepsilon)\cup(\overline{W_m}-q\geq\varepsilon)$.

Ces deux événements sont incompatibles puisque les événements $(\overline{Y_m}-p\geq\varepsilon)$ et $(\overline{Y_m}-p\leq-\varepsilon)$ le sont. Donc

$P(|\overline{Y_m}-p|\geq\varepsilon)=P(\overline{Y_m}-p\geq\varepsilon)+P(\overline{W_m}-q\geq\varepsilon)$, donc

$P(|\overline{Y_m}-p|\geq\varepsilon)\leq\exp(-2m\varepsilon^2)+\exp(-2m\varepsilon^2)$ d'après les questions 2–f) et 3). D'où

$$\boxed{P(|\overline{Y_m}-p|\geq\varepsilon)\leq 2\exp(-2m\varepsilon^2)}$$

4–b)

$$\begin{aligned}2\exp(-2m\varepsilon^2) &= 2\exp(-2m\frac{1.844}{m})\\ &= 2\exp(-3.688)\\ &\simeq 2\exp(\ln(0.025))\end{aligned}$$

$$\boxed{\text{Donc, } 2\exp(-2m\varepsilon^2) \simeq 0.05}$$

Pour $\varepsilon = \sqrt{\frac{1.844}{m}}$,

$$\begin{aligned}P\Big(|\overline{Y_m} - p| \geq \sqrt{\frac{1.844}{m}}\Big) \leq 0.05 &\iff 1 - P\Big(|\overline{Y_m} - p| \geq \sqrt{\frac{1.844}{m}}\Big) \geq 1 - 0.05\\ &\iff P\Big(|\overline{Y_m} - p| < \sqrt{\frac{1.844}{m}}\Big) \geq 0.95\\ &\iff P\Big(\overline{Y_m} - \sqrt{\frac{1.844}{m}} < p < \overline{Y_m} + \sqrt{\frac{1.844}{m}}\Big) \geq 0.95\end{aligned}$$

Un intervalle de confiance au niveau 0.95 **pour** p **est** $I_1 = [\overline{Y_m} - \sqrt{\frac{1.844}{m}}; \overline{Y_m} + \sqrt{\frac{1.844}{m}}]$

Dans B–8), on avait trouvé un intervalle $I_0 = [\overline{Y_m} - \sqrt{\frac{5}{m}}; \overline{Y_m} + \sqrt{\frac{5}{m}}]$

On a $I_1 \subset I_0$ puisque $\sqrt{\frac{1.844}{m}} < \sqrt{\frac{5}{m}}$.

L'intervalle I_1 est plus précis que I_0 puisque pour un même risque la marge d'erreur est plus faible pour I_1 que pour I_0