— Toutes les variables aléatoires qui interviennent dans ce problème sont réelles et définies sur un même espace probabilisé $(\Omega, \mathcal{A}, P)$ où P peut dépendre de paramètres réels inconnus a, b, σ etc ; elles admettent toutes une espérance et une variance : si J désigne l'une de ces variables aléatoires, on note $E(J)$ son espérance et $V(J)$ sa variance.
Si J_1, J_2 et J_1+J_2 sont des variables aléatoires à densité, on admet alors l'existence de la covariance de J_1 et J_2, notée $\mathrm{Cov}\,(J_1,\, J_2)$, qui est définie par la formule : $\mathrm{Cov}\,(J_1,\, J_2) = \dfrac{1}{2}\left(V(J_1+J_2) - V(J_1) - V(J_2)\right)$.
On admet que les covariances de variables aléatoires à densité vérifient les mêmes règles de calcul que celles des variables aléatoires discrètes.
— Pour tout (k,ℓ) de $(\mathbb{N}^*)^2$, on note $\mathcal{M}_{k,\ell}(\mathbb{R})$ l'ensemble des matrices à k lignes et ℓ colonnes à coefficients réels ; on note $\mathcal{M}_k(\mathbb{R})$ l'ensemble des matrices carrées d'ordre k.
— On note ${}^tQ$ la transposée d'une matrice Q.
— Dans tout le problème, n désigne un entier supérieur ou égal à 3.

L'objectif du problème est l'étude de quelques propriétés du modèle de régression linéaire élémentaire.

Partie A - Quelques résultats statistiques et algébriques

On considère une population d'individus statistiques dans laquelle on étudie deux caractères quantitatifs $\mathcal{X}$ et $\mathcal{Y}$. On extrait de cette population, un échantillon de n individus sélectionnés selon des valeurs choisies du caractère $\mathcal{X}$ et numérotées de 1 à n.
Pour tout i de $[\![1;n]\!]$, les réels x_i et y_i sont les observations respectives de $\mathcal{X}$ et de $\mathcal{Y}$ pour l'individu i de l'échantillon. On suppose que les réels $x_1, x_2, \cdots, x_n$ ne sont pas tous égaux.
Soit a et b deux paramètres réels. On pose pour tout i de $[\![1;n]\!]$, $u_i = y_i - (ax_i + b)$. $\quad (*)$

(1). On note $\overline{x}$ (resp. $\overline{y}$) et s_x^2 (resp. s_y^2), la moyenne empirique et la variance empirique de la série statistique $(x_i)_{1\leqslant i\leqslant n}$ (resp.$(y_i)_{1\leqslant i\leqslant n}$) ; on rappelle que $\overline{x} = \dfrac{1}{n}\sum_{i=1}^{n} x_i$ et $s_x^2 = \dfrac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2$.

(a). Montrer que $s_x^2 > 0$.

(b). Etablir les formules : $\sum_{i=1}^{n}(x_i - \overline{x})y_i = \sum_{i=1}^{n}(x_i y_i) - n\overline{x}\,\overline{y}$ et $\sum_{i=1}^{n}(x_i - \overline{x})^2 = \sum_{i=1}^{n}(x_i^2) - n\overline{x}^2$.

(c). On pose pour tout i de $[\![1;n]\!]$: $\alpha_i = \dfrac{(x_i - \overline{x})}{n s_x^2}$. Montrer que : $\sum_{i=1}^{n}\alpha_i = 0$, $\sum_{i=1}^{n}\alpha_i x_i = 1$ et $\sum_{i=1}^{n}\alpha_i^2 = \dfrac{1}{n s_x^2}$

(2). On pose : $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbb{R})$, $u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbb{R})$, $\theta = \begin{pmatrix} a \\ b \end{pmatrix} \in \mathcal{M}_{2,1}(\mathbb{R})$ et $M = \begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix} \in \mathcal{M}_{n,2}(\mathbb{R})$.
Les n relations $(*)$ s'écrivent sous la forme matricielle suivante $y = M\theta + u$.

(a). Quel est le rang de la matrice M ?

(b). Calculer la matrice ${}^tMM$ et justifier son inversibilité.
On exprimera les coefficients de ${}^tMM$ en fonction de s_x^2, $\overline{x}$ et n.

(3). L'espace vectoriel $\mathbb{R}^n$ est muni de sa structure euclidienne canonique. Soit $\mathcal{F}$ le sous-espace vectoriel engendré par les vecteurs $(x_1, \cdots, x_n)$ et $(1 \cdots, 1)$ de $\mathbb{R}^n$. On note K la matrice du projecteur orthogonal de $\mathbb{R}^n$ sur $\mathcal{F}$ dans la base canonique de $\mathbb{R}^n$ et $G = I - K$, où I désigne la matrice identité de $\mathcal{M}_n(\mathbb{R})$.

(a). On cherche les matrices $\theta = \begin{pmatrix} a \\ b \end{pmatrix}$ de $\mathcal{M}_{2,1}(\mathbb{R})$ qui minimisent $\sum_{i=1}^{n} u_i^2 = \sum_{i=1}^{n}(y_i - (ax_i + b))^2$.
Montrer que ce problème admet une unique solution $\hat{\theta} = \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix}$ et qu'elle vérifie la relation ${}^tMM\hat{\theta} = {}^tMy$.

(b). Montrer que $\hat{a} = \sum_{i=1}^{n}\alpha_i y_i$ et $\hat{b} = \overline{y} - \hat{a}\overline{x}$.

(c). Exprimer K en fonction de M et de ${}^tM$.

(d). Soit $\hat{u}$ la matrice colonne de $\mathcal{M}_{n,1}(\mathbb{R})$ de composantes $\hat{u}_1$, $\hat{u}_2$, $\cdots$, $\hat{u}_n$ définie par $\hat{u} = y - M\hat{\theta}$.
Montrer que : $\hat{u} = Gy = Gu$.

(e). En déduire les égalités : ${}^t\hat{u}\hat{u} = \sum_{i=1}^{n}\hat{u}_i^2 = {}^t yGy = {}^t uGu$.

Partie B - Le modèle de régression linéaire

Le contexte et les notations sont ceux de la partie I. Dans cette partie, on cherche à modéliser les fluctuations aléatoires du caractère $\mathcal{Y}$ sur l'échantillon.
Les hypothèses du modèle de régression linéaire élémentaire sont les suivants :
— les réels a et b sont des paramètres inconnus.
— pour tout i de $[\![1;n]\!]$, la valeur x_i du caractère $\mathcal{X}$ est connue et la valeur y_i du caractère $\mathcal{Y}$ est la réalisation d'une variable aléatoire Y_i.
— pour tout i de $[\![1;n]\!]$, Y_i est la somme d'une composante déterministe $ax_i + b$, fonction affine de la valeur x_i, et d'une composante aléatoire U_i.
— les variables aléatoires U_1, U_2, $\cdots$, U_n sont mutuellement indépendantes, de même loi, possèdent une densité, et pour tout i de $[\![1;n]\!]$: $E(U_i) = 0$ et $V(U_i) = \sigma^2$, où le paramètre inconnu σ est strictement positif.
Le modèle de régression linéaire s'écrit alors : pour tout i de $[\![1;n]\!]$, $Y_i = ax_i + b + U_i$ (1).
L'objectif consiste à estimer les paramètres inconnus a, b et σ^2 du modèle (1).
On pose pour tout $n \geqslant 3$: $\overline{Y}_n = \dfrac{1}{n}\sum_{i=1}^{n} Y_i$ et $\overline{U}_n = \dfrac{1}{n}\sum_{i=1}^{n} U_i$.

(4). On note A_n et B_n les deux variables aléatoires définies par $A_n = \sum_{i=1}^{n}\alpha_i Y_i$ et $B_n = \overline{Y}_n - A_n\overline{x}$, où le réel α_i a été défini dans la question **??**.

(a). Montrer que A_n et B_n sont des estimateurs sans biais de a et b respectivement.

(b). Etablir les formules suivantes : $V(A_n) = \dfrac{\sigma^2}{n s_x^2}$ et $V(B_n) = \left(1 + \dfrac{\overline{x}^2}{s_x^2}\right)\dfrac{\sigma^2}{n}$.

(c). Calculer $\mathrm{Cov}\,(A_n,\, B_n)$.

(5). Dans cette question uniquement, l'entier n n'est plus fixé.
On suppose l'existence de $\lambda = \lim\limits_{n\to+\infty}\dfrac{1}{n}\sum_{i=1}^{n} x_i$ et $\mu^2 = \lim\limits_{n\to+\infty}\dfrac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2$, avec $(\lambda, \mu) \in \mathbb{R}\times\mathbb{R}_+^*$.
Montrer que les deux suites $(A_n)_{n\geqslant 3}$ et $(B_n)_{n\geqslant 3}$ convergent en probabilité vers a et b respectivement.

(6). **(a).** On pose pour tout i de $[\![1;n]\!]$: $\hat{U}_i = Y_i - A_n x_i - B_n$. Calculer $E(\hat{U}_i)$.

(b). Etablir l'égalité : $\sum_{i=1}^{n} \hat{U}_i^2 = \sum_{i=1}^{n} (U_i - \overline{U}_n)^2 - n s_x^2 (A_n - a)^2$.

(c). Calculer $E\left(\sum_{i=1}^{n} \hat{U}_i^2\right)$. En déduire un estimateur sans biais de σ^2.

Partie C - Hypothèse de normalité et prévision

Le contexte et les notations de cette partie sont ceux des parties I et II. De plus, on suppose dans cette partie que pour tout i de $[\![1;n]\!]$, la variable aléatoire U_i suit une loi normale $\mathcal{N}(0,\sigma^2)$.

On pose $Y = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}$ et $U = \begin{pmatrix} U_1 \\ \vdots \\ U_n \end{pmatrix}$. Le modèle (1) de la partie II s'écrit alors matriciellement : $Y = M\theta + U$.

Soit W_1, W_2, $\cdots$, W_q ($q \in \mathbb{N}^*$), q variables aléatoires réelles définies sur $(\Omega, \mathcal{A}, P)$. On définit le vecteur aléatoire $(W_1, W_2, \cdots, W_q)$ à valeurs dans $\mathbb{R}^q$; en associant à tout ω de Ω le vecteur $(W_1(\omega), W_2(\omega), \cdots, W_q(\omega))$ de $\mathbb{R}^q$. On dit que le vecteur aléatoire $(W_1, W_2, \cdots, W_q)$ est *normal* si pour tout q-uplet $(\rho_1, \rho_2, \cdots, \rho_q)$ de nombres réels, différent de $(0, 0, \cdots, 0)$, la variable aléatoire $\sum_{i=1}^{n} \rho_i W_i$ suit une loi normale de variance non nulle.

Dans le cas où le vecteur $(W_1, W_2, \cdots, W_q)$ est normal, on admet que les variables aléatoires W_1, W_2, $\cdots$, W_q sont mutuellement indépendantes si et seulement si pour tout (i,j) de $[\![1;q]\!]^2$ avec $i \neq j$, $\mathrm{Cov}(W_i, W_j) = 0$.

(7). **(a).** Montrer que le vecteur aléatoire $(Y_1, Y_2, \cdots, Y_n)$ est normal mais que le vecteur $(Y_1 - \overline{Y}_n, Y_2 - \overline{Y}_n, \cdots, Y_n - \overline{Y}_n)$ ne l'est pas.

(b). Déterminer la loi de chacune des variables aléatoires A_n et B_n. Le vecteur aléatoire (A_n, B_n) est-il normal ?

(8). Soit S une matrice inversible de $\mathcal{M}_n(\mathbb{R})$. On note T la matrice colonne des composantes du vecteur aléatoire $(T_1, T_2, \cdots, T_n)$ telle que $T = SU$.

(a). Montrer que le vecteur $(T_1, T_2, \cdots, T_n)$ est normal.

(b). On suppose que la matrice S est orthogonale i.e. ${}^tS = S^{-1}$. Montrer que T_1, T_2, $\cdots$, T_n sont mutuellement indépendantes.

(9). Soit $\hat{U}_1, \hat{U}_2, \cdots, \hat{U}_n$ les variables aléatoires qui ont été définies dans la question 6.
On note $\hat{U}$ la matrice colonne de composantes $\hat{U}_1, \hat{U}_2, \cdots, \hat{U}_n$ définie par $\hat{U} = Y - M\begin{pmatrix} A_n \\ B_n \end{pmatrix}$

(a). Montrer que $\hat{U} = GU$, où la matrice G a été définie dans la question 3.

(b). Justifier l'existence d'une matrice orthogonale R de $\mathcal{M}_n(\mathbb{R})$et d'une matrice diagonale D de $\mathcal{M}_n(\mathbb{R})$, telles que $G = RD^tR$. Quels sont les éléments diagonaux de D ?

(c). Soit Z la matrice colonne de composantes Z_1, Z_2, $\cdots$, Z_n définie par $Z = {}^t RU$. Quelle est la loi de $\sum_{i=1}^{n-2} Z_i^2$?

(d). En déduire que la variable aléatoire $\sum_{i=1}^{n} \hat{U}_i^2$ suit une loi $\Gamma\left(2\sigma^2, \frac{n-2}{2}\right)$.

(e). Soit p un réel donné vérifiant $0 < p < 1$. Etablir l'existence d'un réel c_n ne dépendant pas des paramètres inconnus a, b et σ^2, tel que $P\left(\left[\sum_{i=1}^{n} \hat{U}_i^2 \geqslant c_n \sigma^2\right]\right) = p$.

Dans les questions 10 et 11, on suppose qu'une $(n+1)^{\text{ème}}$ valeur de $\mathcal{X}$, notée x_{n+1}, est choisie mais que la valeur correspondante y_{n+1} de $\mathcal{Y}$ est inconnue. On suppose que y_{n+1} est la réalisation d'une variable aléatoire Y_{n+1} qui vérifie $Y_{n+1} = ax_{n+1} + b + U_{n+1}$, où les variables aléatoires $U_1, U_2, \cdots, U_{n+1}$ sont mutuellement indépendantes et de même loi $\mathcal{N}(0,\sigma^2)$.

(10). On pose pour tout n-uplet $r = (r_1, r_2, \cdots, r_n)$ de $\mathbb{R}^n$: $\hat{Y}_{n+1}^{(r)} = \sum_{i=1}^{n} r_i Y_i$.

L'ensemble $\left\{\hat{Y}_{n+1}^{(r)};\ r \in \mathbb{R}^n\right\}$ est l'ensemble des "prédicteurs linéaires" de Y_{n+1}.

(a). Soit g la fonction définie sur $\mathbb{R}^n$ à valeurs réelles, telle que pour tout $r = (r_1, r_2, \cdots, r_n)$ de $\mathbb{R}^n$,
$g(r_1, r_2, \cdots, r_n) = \sum_{i=1}^{n} r_i^2$. On rappelle que pour tout i de $[\![1;n]\!]$: $\alpha_i = \frac{(x_i - \overline{x})}{n s_x^2}$.
Montrer que la fonction g admet un minimum absolu sous les contraintes $\sum_{i=1}^{n} r_i = 1$ et $\sum_{i=1}^{n} x_i r_i = x_{n+1}$, atteint en l'unique point $r^* = (r_1^*, r_2^*, \cdots, r_n^*)$, où pour tout i de $[\![1;n]\!]$, $r_i^* = \dfrac{1}{n} + (x_{n+1} - \overline{x})\alpha_i$

(b). Montrer que parmi les prédicteurs linéaires $\hat{Y}_{n+1}^{(r)}$ de Y_{n+1} , qui vérifient $E(\hat{Y}_{n+1}^{(r)}) = E(Y_{n+1})$ pour tout (a,b) de $\mathbb{R}^2$, $\hat{Y}_{n+1}^{(r^*)}$ est celui qui a la plus petite variance.
Vérifier que $\hat{Y}_{n+1}^{(r^*)} = A_n x_{n+1} + B_n$.

(11). **(a).** Déterminer la loi de la variable aléatoire $Y_{n+1} - (A_n x_{n+1} + B_n)$.

(b). On note Φ la fonction de répartition de la loi $\mathcal{N}(0,1)$. Soit p un réel donné vérifiant $\dfrac{1}{2} < p < 1$.
Justifier l'existence d'un réel d_n, que l'on exprimera à l'aide de Φ^{-1}, ne dépendant pas de a, b et σ^2, tel que $P\Big(\big[\,|Y_{n+1} - (A_n x_{n+1} + B_n)| \leqslant d_n \sigma\big]\Big) = p$.

(c). En déduire, à l'aide de la question 9.e), un intervalle dont les bornes ne dépendent que des $(Y_i)_{1 \leqslant i \leqslant n}$, des $(x_i)_{1 \leqslant i \leqslant n+1}$, de c_n et d_n, qui contiennent Y_{n+1} avec une probabilité supérieure ou égale à $2p-1$.
S'agit-il d'un intervalle de confiance au sens usuel du terme ?

Éléments de correction

[pb]

(1). **(a).** — s_x^2 est la somme de carrés, donc $s_x^2 \geqslant 0$.
— Si $s_x^2 = 0$ alors $\forall i \in [\![1;n]\!]$, $x_i = \overline{x}$. Donc les réels x_1, x_2, $\cdots$, x_n sont égaux. Par hypothèse, les réels x_1, x_2, $\cdots$, x_n ne sont pas tous égaux, donc par contraposée $s_x^2 \neq 0$.
Ainsi $s_x^2 > 0$.

(b). — On a :

$$\begin{aligned} \sum_{i=1}^{n}(x_i - \overline{x})y_i &= \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} \overline{x} y_i \qquad \text{par linéarité} \\ &= \sum_{i=1}^{n} x_i y_i - n\overline{x}\,\overline{y} \qquad \text{par définition de } \overline{y} \end{aligned}$$

— Donc $\sum_{i=1}^{n}(x_i-\overline{x})y_i = \sum_{i=1}^{n}(x_iy_i) - n\overline{x}\,\overline{y}$

$$\begin{aligned}\sum_{i=1}^{n}(x_i-\overline{x})^2 &= \sum_{i=1}^{n}x_i^2 - 2x_i\overline{x}+\overline{x}^2\\ &= \sum_{i=1}^{n}x_i^2 - 2\overline{x}\sum_{i=1}^{n}x_i + n\overline{x}^2 \qquad \text{par linéarité}\\ &= \sum_{i=1}^{n}x_i^2 - 2n\overline{x}^2 + n\overline{x}^2 \qquad \text{par définition de } \overline{x}\\ &= \sum_{i=1}^{n}x_i^2 - n\overline{x}^2\end{aligned}$$

Donc $\sum_{i=1}^{n}(x_i-\overline{x})^2 = \sum_{i=1}^{n}(x_i^2) - n\overline{x}^2$.

(c). — On a : $\sum_{i=1}^{n}\alpha_i = \sum_{i=1}^{n}\frac{(x_i-\overline{x})}{ns_x^2} = \frac{1}{ns_x^2}\left(\sum_{i=1}^{n}x_i - n\overline{x}\right) = \frac{1}{ns_x^2}(n\overline{x}-n\overline{x})$

Donc $\sum_{i=1}^{n}\alpha_i = 0$.

— On a :

$$\begin{aligned}\sum_{i=1}^{n}\alpha_i x_i &= \sum_{i=1}^{n}\frac{(x_i-\overline{x})}{ns_x^2}x_i\\ &= \sum_{i=1}^{n}\frac{x_i-\overline{x}}{ns_x^2}(x_i-\overline{x}+\overline{x})\\ &= \frac{1}{ns_x^2}\left(\sum_{i=1}^{n}(x_i-\overline{x})^2 + \overline{x}(x_i-\overline{x})\right)\\ &= \frac{1}{ns_x^2}\left(ns_x^2 + \overline{x}\sum_{i=1}^{n}(x_i-\overline{x})\right)\\ &= \frac{1}{ns_x^2}\left(ns_x^2 + \overline{x}\left(\sum_{i=1}^{n}x_i - n\overline{x}\right)\right)\\ &= 1\end{aligned}$$

Donc $\sum_{i=1}^{n}\alpha_i x_i = 1$.

— On a :

$$\begin{aligned}\sum_{i=1}^{n}\alpha_i^2 &= \sum_{i=1}^{n}\frac{(x_i-\overline{x})^2}{(ns_x^2)^2}\\ &= \frac{1}{(ns_x^2)^2}\sum_{i=1}^{n}(x_i-\overline{x})^2\\ &= \frac{1}{(ns_x^2)^2}ns_x^2\\ &= \frac{1}{ns_x^2}\end{aligned}$$

Donc $\sum_{i=1}^{n}\alpha_i^2 = \frac{1}{ns_x^2}$.

(2). (a). Comme les réels $x_1, x_2, \cdots, x_n$ ne sont pas tous égaux, les vecteurs $\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}$ et $\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}$ ne sont pas proportionnels. Donc le rang de la matrice M est égal à 2.

(b). — On a : ${}^tMM = \begin{pmatrix}x_1 & \cdots & x_n\\ 1 & \cdots & 1\end{pmatrix}\begin{pmatrix}x_1 & 1\\ \vdots & \vdots\\ x_n & 1\end{pmatrix} = \begin{pmatrix}\sum_{i=1}^{n}x_i^2 & \sum_{i=1}^{n}x_i\\ \sum_{i=1}^{n}x_i & n\end{pmatrix} = \begin{pmatrix}ns_x^2+n\overline{x}^2 & n\overline{x}\\ n\overline{x} & n\end{pmatrix} = n\begin{pmatrix}s_x^2+\overline{x}^2 & \overline{x}\\ \overline{x} & 1\end{pmatrix}$

— D'après la question 1a), $s_x^2 > 0$. Donc les vecteurs $\begin{pmatrix}s_x^2+\overline{x}^2\\ \overline{x}\end{pmatrix} = \begin{pmatrix}s_x^2\\ 0\end{pmatrix} + \overline{x}\begin{pmatrix}\overline{x}\\ 1\end{pmatrix}$ et $\begin{pmatrix}\overline{x}\\ 1\end{pmatrix}$ ne sont pas proportionnels.

Donc le rang de ${}^tMM$ est égal à 2. Or ${}^tMM \in \mathcal{M}_2(\mathbb{R})$. Ainsi ${}^tMM$ est inversible.

(3). (a). Remarquons que $\sum_{i=1}^{n}u_i^2 = \|y-M\theta\|^2$ et $M \in \mathcal{M}_{n,2}(\mathbb{R})$ et $\operatorname{rg}(M) = 2$.

En utilisant le théorème des moindres carrés .

D'après la problème des moindres carrés, il existe un unique vecteur θ tel que $\|y-M\theta\|^2$ soit minimale. Alors $\hat{\theta}$ vérifie ${}^tMM\hat{\theta} = {}^tMy$.

En adaptant la démonstration au problème .

D'après le théorème de caractérisation du projeté orthogonal par les distances, il existe un unique vecteur t de $\operatorname{Im}(M)$ minimisant $\|y-M\theta\|^2$. t est alors le projeté orthogonal de y sur $\operatorname{Im}(M)$.

— Comme $t \in \operatorname{Im}(M)$, il existe un vecteur $\hat{\theta}$ de $\mathcal{M}_{2,1}(\mathbb{R})$ tel que $t = M\hat{\theta}$.

— Comme $\operatorname{rg}(M) = 2$, alors d'après le théorème du rang $\dim\operatorname{Ker}(M) = 0$ donc M est injective. Ainsi $\hat{\theta}$ est unique.

— Ainsi il existe un unique vecteur $\hat{\theta}$ de $\mathcal{M}_{2,1}(\mathbb{R})$, tel que $\|y-M\theta\|^2$ soit minimale.

Et $M\hat{\theta}$ est le projeté orthogonal de y sur $\operatorname{Im}M$. Alors $y-M\hat{\theta}$ est orthogonal à tout vecteur de $\operatorname{Im}(M)$.

Pour tout vecteur w de $\mathcal{M}_{2,1}(\mathbb{R})$, ${}^t(Mw)(y-M\hat{\theta}) = 0$ ou encore ${}^tw\big({}^tMy - {}^tMM\hat{\theta}\big)$. Donc le vecteur ${}^tMy - {}^tMM\hat{\theta}$ est orthogonal à tout vecteur w de $\mathcal{M}_{2,1}(\mathbb{R})$. Ainsi ${}^tMy = {}^tMM\hat{\theta}$.

Ainsi ce problème admet une unique solution $\hat{\theta} = \begin{pmatrix}\hat{a}\\ \hat{b}\end{pmatrix}$ et qu'elle vérifie la relation ${}^tMM\hat{\theta} = {}^tMy$.

(b). D'après la question précédente, ${}^tMM\hat{\theta} = {}^tMy$ c'est-à-dire $\begin{pmatrix}(s_x^2+\overline{x}^2)n\hat{a}+\overline{x}n\hat{b}\\ n\overline{x}\hat{a}+n\hat{b}\end{pmatrix} = \begin{pmatrix}\sum_{i=1}^{n}x_iy_i\\ \sum_{i=1}^{n}y_i\end{pmatrix}$.

Donc $(s_x^2+\overline{x}^2)\hat{a}+\overline{x}\hat{b} = \frac{1}{n}\sum_{i=1}^{n}x_iy_i$ et $\hat{b} = \overline{y}-\hat{a}\overline{x}$.

Alors $(s_x^2+\overline{x}^2)\hat{a} = \frac{1}{n}\sum_{i=1}^{n}x_iy_i - \overline{x}\hat{b} = \frac{1}{n}\sum_{i=1}^{n}x_iy_i - \overline{x}\,\overline{y} + \overline{x}^2\hat{a}$

ou encore $s_x^2\hat{a} = \frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})y_i$ d'après la question 1b) .

Ainsi $\hat{a} = \frac{1}{ns_x^2}\sum_{i=1}^{n}(x_i-\overline{x})y_i = \sum_{i=1}^{n}\alpha_iy_i$ par définition de α_i. Finalement $\hat{a} = \sum_{i=1}^{n}\alpha_iy_i$ et $\hat{b} = \overline{y}-\hat{a}\overline{x}$.

(c). K est la matrice de la projection sur $\mathcal{F} = \mathrm{Im}\,(M)$.
Alors pour tout vecteur t de $\mathcal{M}_{n,1}(\mathbb{R})$, $Kt \in \mathrm{Im}\,(M)$ et $t - Kt \in (\mathrm{Im}\,(M))^{\perp}$.
Il existe donc un vecteur t' de $\mathcal{M}_{2,1}(\mathbb{R})$ tel que $Kt = Mt'$.
Et pour tout vecteur w de $\mathcal{M}_{2,1}(\mathbb{R})$, ${}^t(Mw)(t-Kt) = 0$. Alors la vecteur ${}^tM(t-Kt)$ est orthogonal à tout vecteur de $\mathcal{M}_{2,1}(\mathbb{R})$. Donc ${}^tM(t-Kt) = 0$ i.e. ${}^tMt = {}^tMKt$.
En remplaçant dans cette dernière égalité Kt par son expression en fonction de M, ${}^tMt = {}^tMMt'$.
Or ${}^tMM$ est inversible. Donc $t' = ({}^tMM)^{-1}{}^tMt$. Or $Kt = Mt'$.
Donc $Kt = M({}^tMM)^{-1}{}^tMt$. Cette dernière égalité est vraie pour tout vecteur t de $\mathcal{M}_{n,1}(\mathbb{R})$. Ainsi $K = M({}^tMM)^{-1}{}^tM$.

(d). D'après la question 3a), ${}^tMM\hat{\theta} = {}^tMy$.
Or d'après la question 2b), ${}^tMM$ est inversible, donc $\hat{\theta} = ({}^tMM)^{-1}{}^tMy$.
Par définition $\hat{u} = y - M\hat{\theta}$. Donc $\hat{u} = y - M({}^tMM)^{-1}{}^tMy$.
D'après la question précédente, $K = M({}^tMM)^{-1}{}^tM$, donc $\hat{u} = y - Ky = (I-K)y$.
Donc $\hat{u} = Gy$. Comme $G = I - K$ et que K est la projection orthogonale sur $\mathrm{Im}\,(M)$, alors G est la projection orthogonale sur $(\mathrm{Im}\,(M))^{\perp}$. Donc le noyau de G est $\mathrm{Im}\,(M)$.
Par définition, $y = M\theta + u$. Alors $Gy = GM\theta + Gu = Gu$ car $M\theta \in \mathrm{Im}\,(M)$. Ainsi $\hat{u} = Gy = Gu$.

(e). Par définition $\hat{u} = y - M\hat{\theta}$. D'après la question précédente $\hat{u} = Gy$. Alors ${}^t\hat{u}\hat{u} = {}^tyGy - {}^t\hat{\theta}{}^tMGy$.
Or Gy est un élément de l'orthogonal de $\mathrm{Im}\,(M)$et $M\hat{\theta}$ est un élément de $\mathrm{Im}\,(M)$ donc ${}^t\hat{\theta}{}^tMGy = 0$. Donc ${}^t\hat{u}\hat{u} = {}^tyGy$. De même par définition $\hat{u} = y - M\hat{\theta}$ et $y = M\theta + u$ donc $\hat{u} = u + M(\theta - \hat{\theta})$. D'après la question précédente $\hat{u} = Gu$. Alors ${}^t\hat{u}\hat{u} = {}^tuGu + {}^t(\theta-\hat{\theta}){}^tMGu = {}^tuGu$ car $M(\theta - \hat{\theta}) \in \mathrm{Im}\,(M)$. Donc ${}^t\hat{u}\hat{u} = {}^tuGu$.

Partie B - Le modèle de régression linéaire

(4).(a). — Les variables aléatoires Y_i admettent une espérance, donc A_n admet une espérance.

$$\begin{aligned}
E(A_n) &= \sum_{i=1}^{n} \alpha_i E(Y_i) \qquad \text{par linéarité de l'espérance}\\
&= \sum_{i=1}^{n} \alpha_i \big(ax_i + b + E(U_i)\big) \qquad \text{car } Y_i = ax_i + b + U_i\\
&= \sum_{i=1}^{n} \alpha_i \big(ax_i + b\big) \qquad \text{car } E(U_i) = 0\\
&= a\sum_{i=1}^{n} \alpha_i x_i + b\sum_{i=1}^{n} \alpha_i\\
&= a \qquad \text{car d'après la question 1c) } \sum_{i=1}^{n} \alpha_i = 0,\ \sum_{i=1}^{n} \alpha_i x_i = 1
\end{aligned}$$

— On a :

$$\begin{aligned}
E(B_n) &= E(\overline{Y}_n) - \overline{x}E(A_n) \qquad \text{par linéarité de l'espérance}\\
&= \frac{1}{n}\sum_{i=1}^{n} E(Y_i) - a\overline{x}\\
&= \frac{1}{n}\sum_{i=1}^{n} ax_i + b + E(U_i) - a\overline{x} \qquad \text{car } Y_i = ax_i + bU_i\\
&= \frac{1}{n}\sum_{i=1}^{n} ax_i + b - a\overline{x} \qquad \text{car } E(U_i) = 0\\
&= a\overline{x} + b - a\overline{x}\\
&= b
\end{aligned}$$

Donc A_n et B_n sont des estimateurs sans biais de a et b respectivement.

(b). — Comme les variables aléatoires U_1,U_2, $\cdots$, U_n sont mutuellement indépendantes, Y_1, Y_2, $\cdots$,Y_n sont mutuellement indépendantes. Donc $V(A_n) = \sum_{i=1}^{n} \alpha_i^2 V(Y_i)$.
Or $Y_i = ax_i + b + U_i$. Donc $V(Y_i) = V(U_i) = \sigma^2$.
Alors $V(A_n) = \sum_{i=1}^{n} \alpha_i^2\sigma^2 = \dfrac{\sigma^2}{ns_x^2}$ d'après la question 1c) de la partie I.

— Par définition de $\overline{Y_n}$ et de A_n, $B_n = \overline{Y}_n - A_n\overline{x} = \sum_{i=1}^{n}\left(\frac{1}{n} - \alpha_i\overline{x}\right)Y_i$.
Par indépendance des variables aléatoires Y_1, Y_2, $\cdots$, Y_n, $V(B_n) = \sum_{i=1}^{n}\left(\frac{1}{n} - \alpha_i\overline{x}\right)^2 V(Y_i)$.
Or $V(Y_i) = \sigma^2$. Donc $V(B_n) = \sigma^2\sum_{i=1}^{n}\left(\frac{1}{n} - \alpha_i\overline{x}\right)^2$.

$$\begin{aligned}
\text{Par ailleurs } \sum_{i=1}^{n}\left(\frac{1}{n} - \alpha_i\overline{x}\right)^2 &= \sum_{i=1}^{n}\left(\frac{1}{2} - 2\frac{1}{n}\alpha_i\overline{x} + \alpha_i^2\overline{x}^2\right)\\
&= \frac{1}{n} - 2\frac{1}{n}\overline{x}\sum_{i=1}^{n}\alpha_i + \overline{x}^2\sum_{i=1}^{n}\alpha_i^2\\
&= \frac{1}{n} + \overline{x}^2\frac{1}{ns_x^2}
\end{aligned}$$

Ainsi $V(B_n) = \sigma^2\dfrac{1}{n}\left(1 + \frac{\overline{x}^2}{s_x^2}\right)$. Finalement $V(A_n) = \dfrac{\sigma^2}{ns_x^2}$ et $V(B_n) = \left(1 + \frac{\overline{x}^2}{s_x^2}\right)\frac{\sigma^2}{n}$.

(c). Déterminons d'abord l'expression de $A_n + B_n$ en fonction de Y_i, pour ensuite déterminer la variance de $A_n + B_n$ et enfin trouver la covariance de (A_n, B_n) à l'aide la formule
$\mathrm{Cov}\,(A_n,\, B_n) = \dfrac{1}{2}\big(V(A_n + B_n) - V(A_n) - V(B_n)\big)$.
Expression de $A_n + B_n$.

$$\begin{aligned}
A_n + B_n &= \sum_{i=1}^{n}\alpha_i Y_i + \frac{1}{n}\sum_{i=1}^{n} Y_i - \sum_{i=1}^{n}\alpha_i\overline{x}Y_i\\
&= \sum_{i=1}^{n}\left(\frac{1}{n} + \alpha_i - \alpha_i\overline{x}\right)Y_i
\end{aligned}$$

Calcul de $V(A_n+B_n)$.

Les variables aléatoires $Y_1, \cdots, Y_n$ sont mutuellement indépendantes.

Donc $V(A_n+B_n)=\sum_{i=1}^{n}\left(\frac{1}{n}+\alpha_i-\alpha_i\overline{x}\right)^2 V(Y_i)$.

Or $V(Y_i)=\sigma^2$. Donc $V(A_n+B_n)=\sigma^2\sum_{i=1}^{n}\left(\frac{1}{n}+\alpha_i-\alpha_i\overline{x}\right)^2$.

$$\begin{aligned}\text{Par ailleurs } \sum_{i=1}^{n}\left(\frac{1}{n}+\alpha_i-\alpha_i\overline{x}\right)^2 &= \sum_{i=1}^{n}\left(\frac{1}{2}+(1-\overline{x})^2\alpha_i^2+2\frac{1}{n}(1-\overline{x})\alpha_i\right)\\ &= \frac{1}{n}+(1-\overline{x})^2\sum_{i=1}^{n}\alpha_i^2+2\frac{1}{n}(1-\overline{x})\sum_{i=1}^{n}\alpha_i\\ &= \frac{1}{n}+(1-\overline{x})^2\frac{1}{ns_x^2}\end{aligned}$$

Ainsi $V(A_n+B_n)=\frac{\sigma^2}{n}\left(1+(1-\overline{x})^2\frac{1}{s_x^2}\right)$.

Calcul de $\mathrm{Cov}\,(A_n,\,B_n)$.

$$\begin{aligned}\mathrm{Cov}\,(A_n,\,B_n) &= \frac{1}{2}\left(V(A_n+B_n)-V(A_n)-V(B_n)\right)\\ &= \frac{1}{2}\left[\frac{\sigma^2}{n}\left(1+(1-\overline{x})^2\frac{1}{s_x^2}\right)-\frac{\sigma^2}{ns_x^2}-\left(1+\frac{\overline{x}^2}{s_x^2}\right)\frac{\sigma^2}{n}\right]\\ &= \frac{\sigma^2}{2ns_x^2}\left((1-\overline{x})^2-1-\overline{x}^2\right)\\ &= \frac{-\overline{x}\sigma^2}{ns_x^2}\end{aligned}$$

Ainsi $\mathrm{Cov}\,(A_n,\,B_n)=\frac{-\overline{x}\sigma^2}{ns_x^2}$.

(5). $s_x^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^2$. Par hypothèse dans cette question $s_x^2 \underset{n\to+\infty}{\sim} \mu^2$ Donc $\lim_{n\to+\infty} ns_x^2=+\infty$.

Or $V(A_n)=\frac{\sigma^2}{ns_x^2}$. Donc $\lim_{n\to+\infty} V(A_n)=0$.

D'après l'inégalité de Bienaymé-Tchebychev appliquée à la variable aléatoire A_n qui admet un moment d'ordre 2 :

$$\forall\varepsilon>0,\quad P\Big(|A_n-E(A_n)|\geqslant\varepsilon\Big)\leqslant\frac{V(A_n)}{\varepsilon^2}$$

Or $E(A_n)=a$. Donc par encadrement $\lim_{n\to+\infty} P\Big(|A_n-a|\geqslant\varepsilon\Big)=0$. Ainsi la suite $(A_n)_{n\geqslant 3}$ converge en probabilité vers a. D'après la question $4b$), $V(B_n)=\left(1+\frac{\overline{x}^2}{s_x^2}\right)\frac{\sigma^2}{n}$.

Par hypothèse dans cette question $\lim_{n\to+\infty}\overline{x}=\lambda$ et $\lim_{n\to+\infty} s_x^2=\mu^2$, donc $\lim_{n\to+\infty}\left(1+\frac{\overline{x}^2}{s_x^2}\right)=1+\frac{\lambda^2}{\mu^2}$.

Donc $\lim_{n\to+\infty} V(B_n)=0$.

Or d'après l'inégalité de Bienaymé-Tchebychev appliquée à la variable aléatoire B_n qui admet un moment d'ordre 2 : $\forall\varepsilon>0,\quad P\Big(|B_n-E(B_n)|\geqslant\varepsilon\Big)\leqslant\frac{V(B_n)}{\varepsilon^2}$.

Or $E(B_n)=b$. Donc par encadrement $\lim_{n\to+\infty} P\Big(|B_n-b|\geqslant\varepsilon\Big)=0$. Ainsi la suite $(B_n)_{n\geqslant 3}$ converge en probabilité vers b.

(6). (a). Par linéarité de l'espérance $E(\hat{U}_i)=E(Y_i)-x_iE(A_n)-E(B_n)$.

Or $E(Y_i)=ax_i+b$ car $Y_i=ax_i+b+U_i$ et $E(U_i)=0$. Et $E(A_n)=a$ et $E(B_n)=b$.

Donc $E(\hat{U}_i)=0$.

(b). Remarquons que $\hat{U}_i=Y_i-A_nx_i-B_n=ax_i+b+U_i-A_nx_i-\overline{Y}_n+A_n\overline{x}$.

Or $\overline{Y}_n=\frac{1}{n}\sum_{i=1}^{n}Y_i=\frac{1}{n}\sum_{i=1}^{n}(ax_i+b+U_i)=a\overline{x}+b+\overline{U}_n$.

Donc $\hat{U}_i=U_i-\overline{U}_n-(x_i-\overline{x})(A_n-a)$.

$$\begin{aligned}\sum_{i=1}^{n}\hat{U}_i^2 &= \sum_{i=1}^{n}\big(U_i-\overline{U}_n-(x_i-\overline{x})(A_n-a)\big)^2\\ &= \sum_{i=1}^{n}(U_i-\overline{U}_n)^2+\sum_{i=1}^{n}(x_i-\overline{x})^2(A_n-a)^2-2\sum_{i=1}^{n}(x_i-\overline{x})(U_i-\overline{U}_n)(A_n-a)\\ &= \sum_{i=1}^{n}(U_i-\overline{U}_n)^2+ns_x^2(A_n-a)^2-2\sum_{i=1}^{n}(x_i-\overline{x})(U_i-\overline{U}_n)(A_n-a)\end{aligned}$$

Or $U_i-\overline{U}_n=\frac{1}{n}\sum_{j=1}^{n}(U_i-U_j)=\frac{1}{n}\sum_{j=1}^{n}\big(Y_i-Y_j-a(x_i-x_j)\big)=Y_i-\overline{Y}_n-a(x_i-\overline{x})$

et $(x_i-\overline{x})=ns_x^2\alpha_i$.

$$\begin{aligned}\text{Donc } \sum_{i=1}^{n}(x_i-\overline{x})(U_i-\overline{U}_n) &= \sum_{i=1}^{n}(x_i-\overline{x})(Y_i-\overline{Y}_n)-a\sum_{i=1}^{n}(x_i-\overline{x})^2\\ &= ns_x^2\left(\sum_{i=1}^{n}\alpha_iY_i-\sum_{i=1}^{n}\alpha_i\overline{Y}_n\right)-ans_x^2\\ &= ns_x^2A_n-ns_x^2a\\ &= ns_x^2(A_n-a)\end{aligned}$$

Ainsi $\sum_{i=1}^{n}\hat{U}_i^2=\sum_{i=1}^{n}(U_i-\overline{U}_n)^2-ns_x^2(A_n-a)^2$.

(c). Par linéarité de l'espérance, $E\left(\sum_{i=1}^{n}\hat{U}_i^2\right)=\sum_{i=1}^{n}E\big((U_i-\overline{U}_n)^2\big)-ns_x^2E\big((A_n-a)^2\big)$.

$$\begin{aligned}\text{Or } \sum_{i=1}^{n}E\big((U_i-\overline{U}_n)^2\big) &= \sum_{i=1}^{n}E(U_i^2)+nE(\overline{U}_n^2)-2\sum_{i=1}^{n}E(U_i\overline{U}_n)\\ &= n\sigma^2+nE(\overline{U}_n^2)-2E\left(\sum_{i=1}^{n}U_i\overline{U}_n\right)\quad \text{car } E(U_i)=0 \text{ alors } V(U_i)=\\ &= n\sigma^2+nE(\overline{U}_n^2)-2E\left(n\overline{U}_n^2\right)\\ &= n\sigma^2-nE(\overline{U}_n^2)\end{aligned}$$

$$\begin{aligned}
\text{Mais } E(\overline{U}_n^2) &= \frac{1}{2}E\left(\sum_{i=1}^n U_i \sum_{j=1}^n U_j\right)\\
&= \frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n E(U_iU_j)\\
&= \frac{1}{2}\left(\sum_{i=1}^n E(U_i^2)+\sum_{i=1}^n\sum_{\substack{j=1\\j\neq i}}^n E(U_i)E(U_j)\right) \qquad \text{car } U_i \text{ et } U_j\\
&= \frac{\sigma^2}{n}\\
\text{Donc } \sum_{i=1}^n E\big((U_i-\overline{U}_n)^2\big) &= (n-1)\sigma^2
\end{aligned}$$

Et $E\big((A_n-a)^2\big)=E\big((A_n-E(A_n))^2\big)=V(A_n)=\frac{\sigma^2}{ns_x^2}$. Donc $E\left(\sum_{i=1}^n \hat{U}_i^2\right)=(n-2)\sigma^2$.

La variable aléatoire $\frac{1}{n-2}\sum_{i=1}^n \hat{U}_i^2$ est un estimateur sans biais de σ^2.

Partie C - Hypothèse de normalité et prévision

(7). (a). — Comme $Y=M\theta+U$, alors $\forall i\in[\![1;n]\!]$, $Y_i=ax_i+b+U_i$.
U_i suit une loi normale $\mathcal{N}(0,\sigma^2)$. Donc Y_i suit une loi normale $\mathcal{N}(ax_i+b,\sigma^2)$.
Soit $(\rho_1,\cdots,\rho_q)$ un q-uplet différent de $(0,\cdots,0)$.
Alors si $\rho_i\neq 0$, ρ_iY_i suit une loi normale $\mathcal{N}\big(\rho_i(ax_i+b),\rho_i^2\sigma^2\big)$.

Donc $\sum_{i=1}^q \rho_iY_i$ est la somme de variables aléatoires indépendantes suivant une loi normale. Alors $\sum_{i=1}^q \rho_iY_i$ suit une loi normale. Donc le vecteur aléatoire $(Y_1,\cdots,Y_n)$ est normal.

— Remarquons que $\sum_{i=1}^n (Y_i-\overline{Y}_n)=\sum_{i=1}^n Y_i-n\overline{Y}_n=0$.

La variable aléatoire $\sum_{i=1}^n (Y_i-\overline{Y}_n)$ ne suit pas une loi normale. Donc le vecteur $(Y_1-\overline{Y}_n,Y_2-\overline{Y}_n,\cdots,Y_n-\overline{Y}_n)$ n'est pas normal.

(b). — Les réels x_1, x_2, $\cdots$, x_n ne sont pas tous égaux, donc au moins un des réels α_i est non nul. Or le vecteur aléatoire $(Y_1,\cdots,Y_n)$ est normal. Donc A_n suit une loi normale.
Or $E(A_n)=a$ et $V(A_n)=\frac{\sigma^2}{ns_x^2}$. Ainsi A_n suit une loi normale $\mathcal{N}\left(a,\frac{\sigma^2}{ns_x^2}\right)$.

— $B_n=Y_n-A_n\overline{x}=\sum_{i=1}^n\left(\frac{1}{n}-\overline{x}\alpha_i\right)Y_i$.

Or $\sum_{i=1}^n \frac{1}{n}-\overline{x}\alpha_i=1$ car $\sum_{i=1}^n \alpha_i=0$. Il existe donc un entier i tel que $\frac{1}{n}-\overline{x}\alpha_i\neq 0$.
Donc B_n suit une loi normale.
Or $E(B_n)=b$ et $V(B_n)=\left(1+\frac{\overline{x}^2}{s_x^2}\right)\frac{\sigma^2}{n}$. Donc B_n sui une loi normale $\mathcal{N}\left(b,\left(1+\frac{\overline{x}^2}{s_x^2}\right)\frac{\sigma^2}{n}\right)$.

— Soit (ρ_1,ρ_2) de $\mathbb{R}^2\backslash\{(0,0)\}$.
$\rho_1A_n+\rho_2B_n=\sum_{i=1}^n\big(\rho_1\alpha_i+\rho_2(1-\overline{x}\alpha_i)\big)Y_i$.

Supposons que $\forall i\in[\![1;n]\!]$, $\rho_1\alpha_i+\rho_2(1-\overline{x}\alpha_i)=0$.
Alors $\forall i\in[\![1;n]\!]$, $\alpha_i=\frac{\rho_2}{\overline{x}\rho_2-\rho_1}$. Alors les réels α_1, α_2, $\cdots$, α_n sont égaux.
Or $\sum_{i=1}^n \alpha_i=0$ et il existe un réel α_i non nul. Donc les réels α_1, α_2, $\cdots$, α_n ne sont pas tous égaux. Donc il existe un entier i tel que $\rho_1\alpha_i+\rho_2(1-\overline{x}\alpha_i)\neq 0$.
Ainsi $\rho_1A_n+\rho_2B_n$ suit une loi normale. Donc le vecteur aléatoire (A_n,B_n) est normal.

(8). (a). Notons $s_{i,i}$ le coefficient de la ligne i et de la colonne j de la matrice S.
Alors $\forall i\in[\![1;n]\!]$, $T_i=\sum_{j=1}^n s_{i,j}U_j$.
Soit $(\rho_1,\rho_2,\cdots,\rho_n)$ un élément de $\mathbb{R}^n$ différent de $(0,\cdots,0)$.
$\sum_{i=1}^n \rho_iT_i=\sum_{i=1}^n\sum_{j=1}^n \rho_is_{i,j}U_j=\sum_{j=1}^n\left(\sum_{i=1}^n \rho_is_{i,j}\right)U_j$.
Or $\sum_{i=1}^n \rho_is_{i,j}$ est le coefficient de la $j^{\text{ème}}$ colonne du vecteur de $\mathcal{M}_{1,n}(\mathbb{R})$ $\begin{pmatrix}\rho_1 & \rho_2 & \cdots & \rho_n\end{pmatrix}S$.
Comme S est inversible et $\begin{pmatrix}\rho_1 & \rho_2 & \cdots & \rho_n\end{pmatrix}\neq\begin{pmatrix}0 & 0 & \cdots & 0\end{pmatrix}$,
le vecteur $\left(\sum_{i=1}^n \rho_is_{i,1} \quad \sum_{i=1}^n \rho_is_{i,2} \quad \cdots \quad \sum_{i=1}^n \rho_is_{i,n}\right)\neq\begin{pmatrix}0 & 0 & \cdots & 0\end{pmatrix}$.
Or le vecteur aléatoire $(U_1,U_2,\cdots,U_n)$ est normal.
Donc $\sum_{i=1}^n \rho_iT_i=\sum_{j=1}^n\left(\sum_{i=1}^n \rho_is_{i,j}\right)U_j$ suit une loi normale. Ainsi le vecteur $(T_1,T_2,\cdots,T_n)$ est normal.

(b). Soit $(i,j)\in[\![1;n]\!]^2$ tel que $i\neq j$.
$T_i=\sum_{k=1}^n s_{i,k}U_k$ et $T_j=\sum_{\ell=1}^n s_{j,\ell}U_\ell$.

$$\begin{aligned}
\text{Cov}\,(T_i,\,T_j) &= \text{Cov}\left(\sum_{k=1}^n s_{i,k}U_k,\,\sum_{\ell=1}^n s_{j,\ell}U_\ell\right)\\
&= \sum_{k=1}^n\sum_{\ell=1}^n s_{i,k}s_{j,\ell}\text{Cov}\,((,\,U_k)\,U_\ell)\\
&= \sum_{k=1}^n s_{i,k}s_{j,k}\text{Cov}\,(U_k,\,U_k) \qquad \text{car } U_k \text{ et } U_\ell \text{ sont indépendantes}\\
&= \sigma^2\sum_{k=1}^n s_{i,k}s_{j,k}
\end{aligned}$$

Or $\sum_{k=1}^n s_{i,k}s_{j,k}$ est le coefficient de la $i^{\text{ème}}$ ligne et $j^{\text{ème}}$ colonne de la matrice S^tS.
Or $S^tS=I_n$. Donc si $i\neq j$, $\sum_{k=1}^n s_{i,k}s_{j,k}=0$.
d'après la question précédente, $(T_1,\cdots,T_n)$ est normal. Donc T_1, T_2, $\cdots$, T_n sont mutuellement indépendantes.

(9). (a). Remarquons que $GU=GY$ car $Y-U=M\theta\in\text{Im}\,M$ et G est la matrice de la projection orthogonale sur $(\text{Im}\,M)^\perp$.
Or $GY=Y-KY$ et $\hat{U}=Y-M\begin{pmatrix}A_n\\B_n\end{pmatrix}$.
Montrons alors que $M\begin{pmatrix}A_n\\B_n\end{pmatrix}$ est le projeté orthogonal de Y sur $\text{Im}\,M$. Pour cela

nous montrerons que $M\begin{pmatrix}A_n\\B_n\end{pmatrix}$ est un élément de $\operatorname{Im} M$ et que $Y - M\begin{pmatrix}A_n\\B_n\end{pmatrix}$ est un élément de $(\operatorname{Im} M)^\perp$

— $M\begin{pmatrix}A_n\\B_n\end{pmatrix}$ est l'image par M de $\begin{pmatrix}A_n\\B_n\end{pmatrix}$ donc un élément de $\operatorname{Im} M$.

— Soit Z un vecteur de $\mathcal{M}_{2,1}(\mathbb{R})$. Alors ${}^tZ{}^tM(Y - M\begin{pmatrix}A_n\\B_n\end{pmatrix}) = {}^t\ Z({}^tMY - {}^t MM\begin{pmatrix}A_n\\B_n\end{pmatrix})$.

— ${}^tMY = \begin{pmatrix}\sum_{i=1}^n x_iY_i\\ \sum_{i=1}^n Y_i\end{pmatrix} = \begin{pmatrix}\sum_{i=1}^n x_iY_i\\ n\overline{Y}_n\end{pmatrix}$

— ${}^tMM\begin{pmatrix}A_n\\B_n\end{pmatrix} = \begin{pmatrix}(ns_x^2 + n\overline{x})A_n + n\overline{x}B_n\\ n\overline{x}A_n + nB_n\end{pmatrix} = \begin{pmatrix}(ns_x^2 + n\overline{x})A_n + n\overline{x}B_n\\ n\overline{Y}_n\end{pmatrix}$.

$$\begin{aligned}(ns_x^2 + n\overline{x})A_n + n\overline{x}B_n &= (ns_x^2 + n\overline{x})A_n + n\overline{x}(\overline{Y}_n - \overline{x}A_n)\\ &= ns_x^2A_n + n\overline{x}\overline{Y}_n\\ &= \sum_{i=1}^n (x_i - \overline{x})Y_i + \overline{x}\sum_{i=1}^n Y_i\\ &= \sum_{i=1}^n x_iY_i\end{aligned}$$

Donc ${}^tMM\begin{pmatrix}A_n\\B_n\end{pmatrix} = {}^t\ MY$.

Alors $Y - M\begin{pmatrix}A_n\\B_n\end{pmatrix}$ est orthogonal à tout vecteur de $\operatorname{Im} M$.

Ainsi $KY = M\begin{pmatrix}A_n\\B_n\end{pmatrix}$ et $\hat{U} = Y - KY = GY = GU$.

Finalement $\hat{U} = GU$.

(b). La projection orthogonale sur $(\operatorname{Im} M)^\perp$ est un endomorphisme symétrique.
Donc la matrice de cette projection orthogonale dans la base canonique qui est orthonormée, est symétrique.
Donc G est une matrice symétrique réelle.
Ainsi il existe une matrice orthogonale R et une matrice D diagonale telles que $G = RD{}^tR$.
Comme l'image de G est $(\operatorname{Im} M)^\perp$ de dimension $n-2$ et que les valeurs propres de G sont 0 et 1, alors $n-2$ coefficients diagonaux de D sont égaux à 1 et tous les autres coefficients sont nuls.

(c). Déterminons d'abord la loi de Z_i^2 pour ensuite déterminer la loi de la somme.

— — D'après la question 8a), le vecteur aléatoire Z est normal.
Pour tout entier i de $[\![1;n]\!]$, Z_i est combinaison linéaire des variables aléatoires U_i d'espérance nulle, donc l'espérance de Z_i est nulle.
Et d'après le calcul effectué à la question 8b), en prenant $i = j$, on obtient $V(Z_i) = \sigma^2$.
Donc Z_i suit une loi normale $\mathcal{N}(0,\sigma^2)$.

— Pour tout réel t négatif, $P(Z_i^2 \leqslant t) = 0$.

— Et pour tout réel t strictement positif, $P(Z_i^2 \leqslant t) = P(-\sqrt{t} \leqslant Z_i \leqslant \sqrt{t}) = F_{Z_i}(\sqrt{t}) - F_{Z_i}(-\sqrt{t})$.
Alors la fonction de répartition de Z_i^2 est une fonction continue sur $\mathbb{R}$ et de classe $\mathcal{C}^1$sur $\mathbb{R}^*$, donc Z_i^2 est une variable aléatoire à densité, de densité la fonction

$$t \longmapsto \begin{cases}0 & \text{si } t \leqslant 0\\ \dfrac{1}{\sigma\sqrt{2\pi}}\dfrac{1}{\sqrt{t}}\mathrm{e}^{-\frac{t}{2\sigma^2}} & \text{si } t > 0\end{cases}$$

Donc Z_i^2 suit une loi $\Gamma\left(2\sigma^2, \dfrac{1}{2}\right)$.

— D'après la question 8b), les variables aléatoires $Z_1, Z_2, \cdots, Z_n$ sont mutuellement indépendantes. Alors les variables aléatoires $Z_1^2, Z_2^2, \cdots, Z_{n-2}^2$ sont mutuellement indépendantes et suivent la même loi $\Gamma\left(2\sigma^2, \dfrac{1}{2}\right)$.
Donc $\sum_{i=1}^{n-2} Z_i^2$ suit une loi $\Gamma\left(2\sigma^2, \frac{n-2}{2}\right)$.

(d). Remarquons que $\sum_{i=1}^n \hat{U}_i^2 = {}^t\,\hat{U}\hat{U}$ et que $\hat{U} = GU = RD{}^tRU = RDZ$.
Donc $\sum_{i=1}^n \hat{U}_i^2 = {}^t\ Z{}^tD{}^tRRDZ \underset{\text{car } {}^tRR=I_n}{=} {}^t\ ZDZ = \sum_{i=1}^{n-2} Z_i^2$ car D est une matrice diagonale dont $n-2$ coefficients (on peut supposer que ceux sont les premiers.) sont égaux à 1 et les deux autres à 0. Ainsi $\sum_{i=1}^n \hat{U}_i^2$ suit une loi $\Gamma\left(2\sigma^2, \frac{n-2}{2}\right)$.

(e). Alors la variable aléatoire $\dfrac{1}{\sigma^2}\sum_{i=1}^n \hat{U}_i^2$ suit une loi $\Gamma\left(2, \frac{n-2}{2}\right)$.
La fonction $t \longmapsto P\left(\dfrac{1}{\sigma^2}\sum_{i=1}^n \hat{U}_i^2 \geqslant t\right)$ est une fonction strictement décroissante et continue sur $]0,+\infty]$ donc bijective de $]0,+\infty[$ vers $]0.1[$.
Donc il existe un unique réel c_n tel que $P\left(\dfrac{1}{\sigma^2}\sum_{i=1}^n \hat{U}_i^2 \geqslant c_n\right) = p$ et $c_n \in]0,+\infty[$. Donc il existe un réel c_n ne dépendant pas des paramètres inconnus a, b et σ^2, tel que $P\left(\left[\sum_{i=1}^n \hat{U}_i^2 \geqslant c_n\sigma^2\right]\right) = p$.

(10). (a). — Soit $\mathcal{H} = \left\{(r_1,\cdots,r_n) \in \mathbb{R}^n,\ \sum_{i=1}^n r_i = 0 \text{ et } \sum_{i=1}^n x_ir_i = 0\right\}$.
Alors $\mathcal{H}^\perp = \operatorname{Vect}\big((1,1,\cdots,1),(x_1,x_2,\cdots,x_n)\big)$.
$(r_1,\cdots,r_n)$ est un point critique de g sous les contraintes $\sum_{i=1}^n r_i = 1$ et $\sum_{i=1}^n x_ir_i = x_{n+1}$ si et seulement si $\nabla g(r_1,\cdots,r_n) \in \mathcal{H}^\perp$ et $\sum_{i=1}^n r_i = 1$ et $\sum_{i=1}^n x_ir_i = x_{n+1}$.
Or $\nabla g(r_1,\cdots,r_n) \in \mathcal{H}^\perp$
. si et seulement si $\exists(\alpha,\beta) \in \mathbb{R}^2$ tels que $(2r_1,\cdots,2r_n) = \alpha(1,\cdots,1) + \beta(x_1,\cdots,x_n)$
. si et seulement si $\exists(\alpha,\beta) \in \mathbb{R}^2$ tels que $\forall i \in [\![1;n]\!]$, $r_i = \frac{\alpha+\beta x_i}{2}$.
Alors $(r_1,\cdots,r_n)$ est un point critique de g sous les contraintes $\sum_{i=1}^n r_i = 1$ et

$\sum_{i=1}^{n} x_i r_i = x_{n+1}$ si et seulement si il existe des réels α et β tels que pour tout entier i de $[\![1;n]\!]$, $r_i = \frac{\alpha+\beta x_i}{2}$ et $\sum_{i=1}^{n} \frac{\alpha+\beta x_i}{2} = 1$ et $\sum_{i=1}^{n} x_i \frac{\alpha+\beta x_i}{2} = x_{n+1}$.

Or $\sum_{i=1}^{n} \frac{\alpha+\beta x_i}{2} = 1$ et $\sum_{i=1}^{n} x_i \frac{\alpha+\beta x_i}{2} = x_{n+1}$

. si et seulement si $\alpha + \overline{x}\beta = \frac{2}{n}$ et $\overline{x}\alpha + (s_x^2 + n\overline{x}^2)\beta = \frac{2x_{n+1}}{n}$

. si et seulement si $\alpha + \overline{x}\beta = \frac{2}{n}$ et $\beta = \frac{2(x_{n+1}-\overline{x})}{ns_x^2}$.

Ainsi $(r_1, \cdots, r_n)$ est un point critique de g sous les contraintes $\sum_{i=1}^{n} r_i = 1$ et $\sum_{i=1}^{n} x_i r_i = x_{n+1}$ si et seulement si pour tout entier i de $[\![1;n]\!]$, $r_i = \frac{1}{2}(\alpha + \overline{x}\beta + (x_i - \overline{x})\beta) = \frac{1}{n} + (x_{n+1} - \overline{x})\alpha_i$.

— Or $\nabla^2 g(r_1, \cdots, r_n) = \begin{pmatrix} 2 & 0 & \cdots & 0 \\ 0 & 2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 2 \end{pmatrix}$. Les valeurs propres de la hessienne de g en tout point de $\mathbb{R}^n$ sont strictement positives, donc g présente un minimum global en son point critique.

Donc la fonction g admet un minimum absolu sous les contraintes $\sum_{i=1}^{n} r_i = 1$ et $\sum_{i=1}^{n} x_i r_i = x_{n+1}$, atteint en l'unique point $r^* = (r_1^*, r_2^*, \cdots, r_n^*)$, où pour tout i de $[\![1;n]\!]$, $r_i^* = \frac{1}{n} + (x_{n+1} - \overline{x})\alpha_i$.

(b). — $\hat{Y}_{n+1}^{(r)} = \sum_{i=1}^{n} r_i Y_i$. Donc $E\big(\hat{Y}_{n+1}^{(r)}\big) = \sum_{i=1}^{n} r_i E(Y_i) = \sum_{i=1}^{n} r_i(ax_i + b)$.
$E\big(\hat{Y}_{n+1}^{(r)}\big) = E(Y_{n+1})$ pour tout (a,b) de $\mathbb{R}^2$ si et seulement si $\forall (a,b) \in \mathbb{R}^2$, $\sum_{i=1}^{n} r_i(ax_i + b) = ax_{n+1} + b$ si et seulement si $\sum_{i=1}^{n} r_i x_i = x_{n+1}$ et $\sum_{i=1}^{n} r_i = 1$.

— $V\big(\hat{Y}_{n+1}^{(r)}\big) = \sum_{i=1}^{n} r_i V(Y_i)$ car les variables aléatoires Y_1, Y_2, $\cdots$, Y_n sont indépendantes.
Donc $V\big(\hat{Y}_{n+1}^{(r)}\big) = \sigma^2 g(r_1, \cdots, r_2)$. D'après la question précédente, parmi les prédicteurs linéaires $\hat{Y}_{n+1}^{(r)}$ de Y_{n+1} , qui vérifient $E(\hat{Y}_{n+1}^{(r)}) = E(Y_{n+1})$ pour tout (a,b) de $\mathbb{R}^2$, $\hat{Y}_{n+1}^{(r^*)}$ est celui qui a la plus petite variance.

—

$$\begin{aligned} \text{Alors } \hat{Y}_{n+1}^{(r^*)} &= \sum_{i=1}^{n} r_i^* Y_i \\ &= \sum_{i=1}^{n} \frac{1}{n} Y_i + \sum_{i=1}^{n} (x_{n+1} - \overline{x})\alpha_i Y_i \qquad \text{car } r_i^* = \frac{1}{n} + (x_{n+1} - \overline{x})\alpha_i \\ &= \overline{Y}_n + x_{n+1} A_n - \overline{x} A_n \\ &= x_{n+1} A_n + B_n \end{aligned}$$

Donc $\hat{Y}_{n+1}^{(r^*)} = A_n x_{n+1} + B_n$.

(11). (a). — $Y_{n+1} - (A_n x_{n+1} + B_n) = Y_{n+1} - \sum_{i=1}^{n} \left(\alpha_i(x_{n+1} - \overline{x}) + \frac{1}{n}\right) Y_i$.
Or le vecteur $(Y_1, \cdots, Y_{n+1})$ est normal. Donc $Y_{n+1} - (A_n x_{n+1} + B_n)$ suit une loi normale.

— $E\big(Y_{n+1} - (A_n x_{n+1} + B_n)\big) = E(Y_{n+1}) - x_{n+1}E(A_n) - E(B_n) = ax_{n+1} + b - ax_{n+1} - b = 0$.
Et $V\big(Y_{n+1} - (A_n x_{n+1} + B_n)\big) = V(Y_{n+1}) + x_{n+1}^2 V(A_n) + V(B_n) + 2x_{n+1}\mathrm{Cov}\,(A_n,\, B_n)$ car Y_{n+1} et $A_n x_{n+1} + B_n$ sont indépendantes.
Donc $V\big(Y_{n+1} - (A_n x_{n+1} + B_n)\big) = \sigma^2 + \frac{x_{n+1}^2 \sigma^2}{ns_x^2} + \left(1 + \frac{\overline{x}^2}{s_x^2}\right)\frac{\sigma^2}{n} + 2x_{n+1}\frac{\overline{x}\sigma^2}{ns_x^2}$.
Posons $\delta_n = 1 + \frac{x_{n+1}^2}{ns_x^2} + \left(1 + \frac{\overline{x}^2}{s_x^2}\right)\frac{1}{n} + 2x_{n+1}\frac{\overline{x}}{ns_x^2}$. δ_n ne dépend pas de a, ni de b ni de σ^2.
Ainsi $Y_{n+1} - (A_n x_{n+1} + B_n)$ suit une loi normale $\mathcal{N}\big(0, \sigma^2\delta_n\big)$.

(b). La variable aléatoire $\frac{1}{\sigma\sqrt{\delta_n}}\Big(Y_{n+1} - (A_n x_{n+1} + B_n)\Big)$ suit une loi normale $\mathcal{N}(0,1)$.
Donc $P\left(\left|\frac{1}{\sigma\sqrt{\delta_n}}\big(Y_{n+1} - (A_n x_{n+1} + B_n)\big)\right| \leqslant \frac{t}{\sqrt{\delta_n}}\right) = \Phi(\frac{t}{\sqrt{\delta_n}}) - \Phi(-\frac{t}{\sqrt{\delta_n}}) = 2\Phi(\frac{t}{\sqrt{\delta_n}}) - 1$.
Or Φ est une fonction strictement croissante et continue sur $\mathbb{R}_+$ donc bijective de $\mathbb{R}_+$ vers $\left[\frac{1}{2}, 1\right[$.
Il existe donc un unique réel positif d_n tel que $\Phi(\frac{d_n}{\sqrt{\delta_n}}) = \frac{p+1}{2}$. Ainsi il existe un réel d_n, $d_n = \sqrt{\delta_n}\Phi^{-1}\left(\frac{p+1}{2}\right)$ tel que $P\Big(\big[\,|Y_{n+1} - (A_n x_{n+1} + B_n)| \leqslant d_n\sigma\big]\Big) = p$.

(c). Remarquons que $[-d_n\sigma + A_n x_{n+1} + B_n \leqslant Y_{n+1} \leqslant d_n\sigma + A_n x_{n+1} + B_n] \cap \left[\sigma^2 c_n \sum_{i=1}^{n} \hat{U}_i^2\right]$ est inclus dans :

$$\left[-d_n\sqrt{\frac{\sum_{i=1}^{n}\hat{U}_i^2}{c_n}} + A_n x_{n+1} + B_n \leqslant Y_{n+1} \leqslant d_n\sqrt{\frac{\sum_{i=1}^{n}\hat{U}_i^2}{c_n}} + A_n x_{n+1} + B_n\right]$$

Or si A et B sont des événements, alors $P(A \cap B) = P(A) + P(B) - P(A \cup B) \geqslant P(A) + P(B) - 1$.
Donc $P\Big((-d_n\sigma + A_n x_{n+1} + B_n \leqslant Y_{n+1} \leqslant d_n\sigma + A_n x_{n+1} + B_n) \cap (\sigma^2 c_n \sum_{i=1}^{n} \hat{U}_i^2)\Big) \geqslant 2p - 1$.
Ainsi $P\left(-d_n\sqrt{\frac{\sum_{i=1}^{n}\hat{U}_i^2}{c_n}} + A_n x_{n+1} + B_n \leqslant Y_{n+1} \leqslant d_n\sqrt{\frac{\sum_{i=1}^{n}\hat{U}_i^2}{c_n}} + A_n x_{n+1} + B_n\right) \geqslant 2p - 1$.
Cet intervalle n'est pas un intervalle de confiance au sens usuel car Y_{n+1} est une variable aléatoire.