Partie 1

- 1) a) **Réponse 1** : Pour $(a, b, c) \in \mathbb{R}^3$, $F(a, b, c)$ est la somme des carrés des distances des points $M_i(t_i, y_i)$ à la parabole d'équation $y = a + bt + ct^2$: (faire une figure)

$$F(a, b, c) = \sum_{i=1}^n (y_i - a - bt_i - ct_i^2)^2$$

Minimiser $F(a, b, c)$ revient à chercher la parabole qui réalise le meilleur ajustement au sens des moindres carrés.

Réponse 2 : Dans l'espace euclidien $\mathbb{R}^n$ muni du produit scalaire usuel.

$$\text{En notant } Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \quad U_0 = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \quad U_1 = \begin{pmatrix} t_1 \\ \vdots \\ t_n \end{pmatrix} \quad U_2 = \begin{pmatrix} t_1^2 \\ \vdots \\ t_n^2 \end{pmatrix}$$

$$F(a, b, c) = \|Y - aU_0 - bU_1 - cU_2\|^2$$

Le minimum de F est la distance de Y au sous-espace $\text{Vect}(U_0, U_1, U_2) = E$ il est réalisé pour $(\hat{a}, \hat{b}, \hat{c})$ tel que.

$$P_E(Y) = \hat{a} U_0 + \hat{b} U_1 + \hat{c} U_2 \quad (\text{Projeté orthogonal de } Y \text{ sur } E)$$

C'est cette deuxième approche qui motive la première partie de cette épreuve de concours. Nous la reverrons en cours dans le cas de la régression linéaire.

Remarque du correcteur :

$P_E(Y)$ est l'unique vecteur de E tel que $Y - P_E(Y)$ est orthogonal à U_1, U_2 et U_3 .

Ce qui se traduit avec les notations de ce devoir par la relation : $T^\top (Y - T\hat{\beta}) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$

Le projeté orthogonal sur E est le vecteur $T\hat{\beta}$.

Il n'était pas nécessaire de savoir cela pour faire cette partie.

- b) F est une fonction polynomiale en a, b et c donc F admet des dérivées partielles par rapport à a, b et c ,

$$F(a, b, c) = \sum_{i=1}^n (y_i - a - bt_i - ct_i^2)^2$$

$$\text{donc } \begin{cases} \frac{\partial F}{\partial a}(a, b, c) = \sum_{i=1}^n -2(y_i - a - bt_i - ct_i^2) \\ \frac{\partial F}{\partial b}(a, b, c) = \sum_{i=1}^n -2t_i(y_i - a - bt_i - ct_i^2) \\ \frac{\partial F}{\partial c}(a, b, c) = \sum_{i=1}^n -2t_i^2(y_i - a - bt_i - ct_i^2) \end{cases}$$

$$2) \text{ a) } Y - T\beta = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} - \begin{pmatrix} 1 & t_1 & t_1^2 \\ \vdots & \vdots & \vdots \\ 1 & t_n & t_n^2 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} \text{ donc } Y - T\beta = \begin{pmatrix} y_1 - a - bt_1 - ct_1^2 \\ \vdots \\ y_n - a - bt_n - ct_n^2 \end{pmatrix}$$

$$\text{or } \frac{\partial F}{\partial a}(a, b, c) = -2 \sum_{i=1}^n (y_i - a - bt_i - ct_i^2) \quad (\text{d'après 1)b}).$$

$$\text{donc } \boxed{\frac{\partial F}{\partial a}(a, b, c) = -2 \left\langle Y - T\beta ; \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \right\rangle}$$

b) De même

$$\begin{aligned} -2 \left\langle Y - T\beta ; \begin{pmatrix} t_1 \\ \vdots \\ t_n \end{pmatrix} \right\rangle &= -2 \begin{pmatrix} y_1 - a - bt_1 - ct_1^2 \\ \vdots \\ y_n - a - bt_n - ct_n^2 \end{pmatrix} \begin{pmatrix} t_1 \\ \vdots \\ t_n \end{pmatrix} \\ &= -2 \sum_{i=1}^n t_i (y_i - a - bt_i - ct_i^2) \end{aligned}$$

$$\text{donc } \frac{\partial F}{\partial b}(a, b, c) = -2 \left\langle Y - T\beta ; \begin{pmatrix} t_1 \\ \vdots \\ t_n \end{pmatrix} \right\rangle$$

$$3) \text{ Soit } \hat{\beta} = \begin{pmatrix} \hat{a} \\ \hat{b} \\ \hat{c} \end{pmatrix} \text{ une solution de } \begin{cases} \frac{\partial F}{\partial a}(a, b, c) = 0 \\ \frac{\partial F}{\partial b}(a, b, c) = 0 \\ \frac{\partial F}{\partial c}(a, b, c) = 0 \end{cases},$$

on en déduit alors avec les résultats de la question 2) que :

$$\begin{cases} \begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix} (Y - T\hat{\beta}) = 0 \\ \begin{pmatrix} t_1 & t_2 & \cdots & t_n \end{pmatrix} (Y - T\hat{\beta}) = 0 \\ \begin{pmatrix} t_1^2 & t_2^2 & \cdots & t_n^2 \end{pmatrix} (Y - T\hat{\beta}) = 0 \end{cases}$$

Ce qui s'écrit aussi :

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ t_1 & t_2 & \cdots & t_n \\ t_1^2 & t_2^2 & \cdots & t_n^2 \end{pmatrix} (Y - T\hat{\beta}) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\text{d'où : } T^\top (Y - T\hat{\beta}) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \text{ ou encore : } T^\top Y = T^\top T\hat{\beta}$$

on a bien :

$$\boxed{\text{Si } \hat{\beta} \text{ est un point critique de } F \text{ alors } T^\top T\hat{\beta} = T^\top Y}$$

4) a) *Raisonnons par double inclusion.*

- Soit $X \in \text{Ker}(A)$, on sait que $AX = 0$ ou en déduit que $A^\top AX = 0$, autrement dit $X \in \text{Ker}(A^\top A)$

$$\boxed{\text{ker}(A) \subset \text{ker}(A^\top A)}$$

- Soit $X \in \text{ker}(A^\top A)$, on sait que $A^\top AX = 0$ ou en déduit que $X^\top A^\top AX = 0$, autrement dit : $\|AX\|^2 = 0$ ce qui entraîne $AX = 0$ et ainsi $X \in \text{Ker}(A)$

$$\boxed{\text{ker}(A^\top A) \subset \text{ker}(A)}$$

En conclusion :

$$\boxed{\text{ker}(A) = \text{ker}(A^\top A)}$$

b) $A \in \mathcal{M}_{n,p}(\mathbb{R})$ donc $A^\top A \in \mathcal{M}_p(\mathbb{R})$ donc en appliquant le théorème du rang à $X \mapsto A^\top AX$ on obtient

$$\boxed{\dim(\ker(A^\top A)) + \text{rg}(A^\top A) = p}$$

c) $A \in \mathcal{M}_{n,p}(\mathbb{R})$ donc en appliquant le théorème du rang à $X \mapsto AX$ on obtient

$$\boxed{\dim(\ker(A)) + \text{rg}(A) = p}$$

Remarque : On rappelle que $X \mapsto AX$ est une application de $\mathcal{M}_{p,1}(\mathbb{R})$ dans $\mathcal{M}_{n,1}(\mathbb{R})$

d)

$$\begin{aligned} \text{rg}(A) &= p - \dim(\ker(A)) && \text{d'après c)} \\ &= p - \dim(\ker(A^\top A)) && \text{d'après a)} \\ &= \text{rg}(A^\top A) && \text{d'après b)} \end{aligned}$$

de plus $A^\top A \in \mathcal{M}_p(\mathbb{R})$ donc $\boxed{\text{Si } \text{rg}(A) = p \text{ alors } A^\top A \text{ est inversible.}}$

On sait que $\text{rg}(A) = \text{rg}(A^\top A)$ et $A^\top A \in \mathcal{M}_p(\mathbb{R})$ donc

si $A^\top A$ est inversible alors $\text{rg}(A) = p$.

En conclusion :

$$\boxed{\text{rg}(A) = p \text{ si, et seulement si, } A^\top A \text{ est inversible.}}$$

5) Si $T^\top T$ est inversible, sachant que $T^\top T \hat{\beta} = T^\top Y$ il vient $\hat{\beta} = (T^\top T)^{-1} T^\top Y$

$$\boxed{\text{Si } T^\top T \text{ est inversible alors } \hat{\beta} = (T^\top T)^{-1} T^\top Y}$$

Attention : T n'est pas une matrice carrée, elle n'a pas de raison d'être inversible.