# Exercices avec corrigé succinct du chapitre 3

(Remarque : les références ne sont pas gérées dans ce document, par contre les quelques **??** qui apparaissent dans ce texte sont bien définis dans la version écran complète du chapitre 3)

## Exercice III.1

1. Ecrire le problème de la régression linéaire comme un problème de moindres carrés : plus précisément on se donne une famille de points $(t_i,b_i)_{1\leq i\leq m}$ (les $t_i$ étant distincts) et on cherche à faire passer une droite le plus près possible de ces points.
2. La question précédente conduit à une fonction de deux variables à minimiser. On admet que ce minimum est donné en annulant les deux dérivées partielles. Donner le système linéaire de deux équations à deux inconnues ainsi obtenu.

**Solution :**

1. Soit $y = \alpha + \beta t$, l'équation de la droite considérée. Le problème de régression linéaire s'écrit

$$\min_{(\alpha,\beta)\in\mathbb{R}^2} E(\alpha,\beta)$$

   où

$$E(\alpha,\beta) = \sum_{i=1}^{m}(\alpha + \beta t_i - b_i)^2.$$

   La solution $(\alpha^*,\beta^*)$ donne les coefficients de la droite solution du problème de régression linéaire et elle vérifie

$$E(\alpha^*,\beta^*) = \min_{(\alpha,\beta)\in\mathbb{R}^2} E(\alpha,\beta).$$

2. Calculons les deux dérivées partielles

$$\begin{cases} \frac{\partial E}{\partial \alpha}(\alpha,\beta) = \sum_{i=1}^{m} 2(\alpha + \beta t_i - b_i) \\ \\ \frac{\partial E}{\partial \beta}(\alpha,\beta) = \sum_{i=1}^{m} 2(\alpha + \beta t_i - b_i)t_i \end{cases}$$

   La solution $(\alpha^*,\beta^*)$ du problème de régression linéaire est donc donnée par la solution des deux équations linéaires obtenues en regroupant les termes

$$\begin{cases} \alpha^* m + \beta^* \sum_{i=1}^{m} t_i = \sum_{i=1}^{m} b_i \\ \\ \alpha^* \sum_{i=1}^{m} t_i + \beta^* \sum_{i=1}^{m} t_i^2 = \sum_{i=1}^{m} b_i t_i \end{cases}$$

□

## Exercice III.2

On cherche à approcher les données $(t_i,b_i)_{1\leq i\leq m}$ (les $t_i$ étant distincts) à l'aide d'un polynôme de degré inférieur ou égal à $n-1$, on suppose $m \geq n$, on note

$$p(t) = \alpha_1 + \alpha_2 t + \ldots + \alpha_n t^{n-1},$$

et on cherche les coefficients $\alpha_1,\alpha_2,\ldots,\alpha_n$ qui minimisent

$$E(\alpha_1,\alpha_2,\ldots,\alpha_n) = \sum_{i=1}^{m}(p(t_i) - b_i)^2.$$

Ecrire le problème de moindres carrés sous forme matricielle. Montrer alors que la matrice $A$ est bien de rang $n$. (On rappelle que la matrice carrée de Van der Monde V, $v_{ij} = t_i^{j-1}$, est inversible si tous les $t_i$ sont distincts).

**Solution :** Soit

$$p(t) = \alpha_1 + \alpha_2 t + \ldots + \alpha_n t^{n-1},$$

un polynôme de degré $n-1$. Pour que ce polynôme approche les données $(t_i,b_i)_{i=1,\ldots,m}$ le plus près possible, il doit minimiser la quantité suivante

$$E(\alpha_1,\alpha_2,\ldots,\alpha_n) = \sum_{i=1}^{m}(p(t_i) - b_i)^2 = \|Ax - b\|^2.$$

En effet on peut écrire

$$\begin{pmatrix} p(t_1) - b_1 \\ p(t_2) - b_2 \\ \ldots \\ p(t_m) - b_m \end{pmatrix} = \begin{pmatrix} \alpha_1 + \alpha_2 t_1 + \ldots + \alpha_n t_1^{n-1} - b_1 \\ \alpha_1 + \alpha_2 t_2 + \ldots + \alpha_n t_2^{n-1} - b_2 \\ \ldots \\ \alpha_1 + \alpha_2 t_m + \ldots + \alpha_n t_m^{n-1} - b_m \end{pmatrix} = A \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \ldots \\ \alpha_n \end{pmatrix} - \begin{pmatrix} b_1 \\ b_2 \\ \ldots \\ b_m \end{pmatrix} = Ax - b,$$

où les matrices $A$ et $x$ sont définies par :

$$A = \begin{pmatrix} 1 & t_1 & \ldots & t_1^{n-1} \\ 1 & t_2 & \ldots & t_2^{n-1} \\ \ldots & \ldots & \ldots & \ldots \\ 1 & t_m & \ldots & t_m^{n-1} \end{pmatrix}, \; x = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}.$$

Dans les données, les points $t_i$ sont tous distincts, ce qui implique que la matrice $V$, constituée des $n$ premières lignes de $A$ est inversible, d'où la matrice $A$ est de rang $n$. □

### Exercice III.3

Donner les équations normales du problème de moindres carrés associé à la régression linéaire. Montrer que l'on retrouve les équations de l'exercice **??**.

**Solution :** La matrice $A$ du problème de régression linéaire s'écrit (voir la correction de l'exercice **??**) :

$$A = \begin{pmatrix} 1 & t_1 \\ 1 & t_2 \\ \ldots & \ldots \\ 1 & t_m \end{pmatrix}.$$

Les équations normales sont

$$A^T A \hat{x} = A^T b$$

ce qui donne en effectuant les produits matriciels

$$\begin{pmatrix} m & \sum_{i=1}^m t_i \\ \sum_{i=1}^m t_i & \sum_{i=1}^m t_i^2 \end{pmatrix} \hat{x} = \begin{pmatrix} \sum_{i=1}^m b_i \\ \sum_{i=1}^m t_i b_i \end{pmatrix}.$$

On retrouve bien ainsi le système de deux équations à deux inconnues de l'exercice **??**. □

### Exercice III.4

Soit $b \in \mathbb{R}^n$, montrer que

$$b^T z = 0, \forall z \in \mathbb{R}^n \;\Leftrightarrow\; b = 0.$$

**Solution :** L'implication $\Leftarrow$ est évidente puisque l'on multiplie 0 par le vecteur $z$.

Supposons maintenant que

$$b^T z = 0, \forall z \in \mathbb{R}^n,$$

alors cette égalité étant vraie pour tout $z$ l'est en particulier pour $z = b$, ce qui donne

$$b^T b = \|b\|^2 = 0.$$

Or la norme d'un vecteur est nulle si et seulement si ce vecteur est nul, ce qui donne

$$b = 0.$$

□

### Exercice III.5

Ecrire l'algorithme de l'orthogononalisation de Schmidt, donnée dans le document **??**.

**Solution :** Cet algorithme est très simple et suppose connues des fonctions telles que norme, produit scalaire ... ce qui est le cas de Scilab.

$E_1 = B_1/\|B_1\|$
**pour** $k = 2, \ldots, n$ **faire**
  $\widetilde{E}_k = B_k - \sum_{j=1}^{k-1} \langle B_k, E_j \rangle E_j$
  $E_k = \widetilde{E}_k / \|\widetilde{E}_k\|$
**fin pour**

□

### Exercice III.6

Soit $E \in \mathcal{M}_{mn}$, une matrice dont les colonnes sont des vecteurs orthonormés de $\mathbb{R}^n$, montrer que $E^T E = I$

**Solution :** On a donc

$$\langle E_i, E_i \rangle = 1, \ \langle E_i, E_j \rangle = 0 \text{ pour } i \neq j,$$

ou ce qui est équivalent

$$E_i^T E_i = 1, \ E_i^T E_j = 0 \text{ pour } i \neq j.$$

Les termes de la matrice (carrée) $C = E^T E$ sont donc

$$c_{ii} = E_i^T E_i = 1, \ \ c_{ij} = E_i^T E_j = 0 \text{ pour } i \neq j,$$

$C$ est donc la matrice identité.

□

### Exercice III.7

On applique l'orthogonalisation de Schmidt (voir le document **??**), sur les $n$ colonnes $A_1, A_2, ..., A_n$ d'une matrice $A \in \mathcal{M}_{m,n}$ de rang $n$ ($m \geq n$), on obtient les vecteurs $E_1, E_2, ..., E_n$ qui seront les colonnes d'une matrice $E$.

1. Montrer que les colonnes $A_k$ de $A$ peuvent s'écrire

$$A_k = \sum_{j=1}^{k} \alpha_{jk} E_j$$

sans expliciter les scalaires $\alpha_{jk}$.

2. En déduire que $A = ET$, où $T$ est une matrice triangulaire supérieure inversible.

**Solution :**

1. Reprenons l'algorithme d'orthogonalisation de Schmidt, alors

$$E_1 = A_1/\|A_1\| \;\Rightarrow\; A_1 = \alpha_{11}E_1,$$

puis

$$\widetilde{E}_2 = A_2 - \langle A_1,E_1\rangle\, E_1 \text{ et } E_2 = \widetilde{E}_2/\|\widetilde{E}_2\| \;\Rightarrow\; A_2 = \langle A_1,E_1\rangle\, E_1 + \|\widetilde{E}_2\|E_2,$$

ce qui donne

$$A_2 = \alpha_{12}E_1 + \alpha_{22}E_2.$$

De manière générale

$$\widetilde{E}_k = A_k - \sum_{j=1}^{k-1}\langle A_k,E_j\rangle\, E_j \text{ et } E_k = \widetilde{E}_k/\|\widetilde{E}_k\| \;\Rightarrow\; A_k = \sum_{j=1}^{k-1}\langle A_k,E_j\rangle\, E_j + \|\widetilde{E}_k\|E_k,$$

ce qui donne

$$A_k = \sum_{j=1}^{k}\alpha_{jk}E_j.$$

2. Considérons le produit $C = ET$ de deux matrices, $E \in \mathcal{M}_{m,n}$ et $T \in \mathcal{M}_{n,n}$, alors

$$c_{ik} = \sum_{j=1}^{n} e_{ij}t_{jk}.$$

On peut aussi considérer $c_{ik}$ comme le ième élément de la kième colonne de $C$. Alors cette colonne est donnée par

$$C_k = ET_k = \sum_{j=1}^{n} t_{jk}E_j.$$

Si l'on compare avec le résultat de la question précédente :

$$A_k = \sum_{j=1}^{k}\alpha_{jk}E_j,$$

on voit que $t_{jk} = \alpha_{jk}$ pour $j = 1,\ldots,k$ et que $t_{jk} = 0$ pour $j = k+1,\ldots,n$, ce qui correspond à une matrice triangulaires supérieure

$$T = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \ldots & \ldots & \alpha_{1n} \\ 0 & \alpha_{22} & \ldots & \ldots & \alpha_{2n} \\ 0 & \ldots & \vdots & \ldots & \ldots \\ 0 & \ldots & \ldots & \ldots & \alpha_{nn} \end{pmatrix}$$

La matrice $T$ est inversible car $\alpha_{ii} = \|\widetilde{E}_i\|$.

□

## Exercice III.8

1. Montrer que si $Q$ est une matrice orthogonale, alors $Q^T$ est orthogonale.
2. Montrer que si $Q$ est orthogonale, alors $\|Qy\| = \|y\|$ pour tout vecteur $y$ de $\mathbb{R}^n$.

**Solution :**

1. On a
$$Q \text{ orthogonale} \quad \Leftrightarrow \quad Q^{-1} = Q^T \quad \Leftrightarrow \quad (Q^T)^{-1} = Q \quad \Leftrightarrow \quad Q^T \text{ orthogonale} .$$
2. On va démontrer le résultat pour le carré de l'expression, ce qui est équivalent pour des réels positifs. Dans ces équivalences, on utilise le fait que $Q^T Q = I$.
$$\|Qy\|^2 = (Qy)^T Qy = y^T Q^T Qy = y^T y = \|y\|^2.$$

$\square$

## Exercice III.9

On veut effectuer une régression linéaire sur les points suivants : $(-1{,}0.5)$, $(0.5{,}1)$, $(2{,}2.5)$. Appliquer la méthode "QR" pour résoudre ce problème et utiliser SCILAB, en particulier la procédure "qr", pour faire les calculs.

**Solution :** La matrice $A$ et le vecteur $b$ correspondants à ce problème sont
$$A = \begin{pmatrix} 1 & -1 \\ 1 & 0.5 \\ 1 & 2 \end{pmatrix} , \quad b = \begin{pmatrix} 0.5 \\ 1 \\ 2.5 \end{pmatrix} .$$

Les étapes du calcul sont alors les suivantes :

- calcul de la décomposition QR par "qr" ce qui donne $A = QR$, où $R = \begin{pmatrix} \widetilde{R} \\ 0 \end{pmatrix}$,
- calcul de $Q^T b = \begin{pmatrix} c \\ d \end{pmatrix}$,
- résolution de $\widetilde{R}x = c$, ce qui donne $x = \begin{pmatrix} 1 \\ 0.666667 \end{pmatrix}$
- calcul de l'erreur $\|d\|^2 = 0.1666667$.

$\square$

## Exercice III.10

Soit $A$ une matrice $m \times n$ de rang $n \leq m$. Soient $Q$ une matrice orthogonale et $\widetilde{R}$ une matrice carrée triangulaire supérieure telles que $A = Q \begin{pmatrix} \widetilde{R} \\ 0 \end{pmatrix}$.

Montrer que, si $\chi_2$ désigne le conditionnement calculé à partir de la norme matricielle subordonnée à la norme 2,
$$\chi_2(\widetilde{R}) = \sqrt{\chi_2(A^T A)}.$$

**Solution :** Revoyez le lien entre la norme $\|.\|_2$ et le rayon spectral vu au chapitre 2.
$$(\chi_2(\widetilde{R}))^2 = \|\widetilde{R}\|_2^2 \|\widetilde{R}^{-1}\|_2^2 = \rho(\widetilde{R}^T \widetilde{R}) \rho((\widetilde{R}^{-1})^T \widetilde{R}^{-1}).$$

On remarque que $A^T A = R^T Q^T Q R = R^T R = \widetilde{R}^T \widetilde{R}$ donc
$$\chi_2(A^T A) = \chi_2(\widetilde{R}^T \widetilde{R}) = \|\widetilde{R}^T \widetilde{R}\|_2 \|\|(\widetilde{R}^T \widetilde{R})^{-1}\|_2,$$

or $\widetilde{R}^T\widetilde{R}$ et son inverse sont des matrices symétriques, toujours dans le chapitre 2, on a montré

$$\|R^T R\|_2 = \rho(R^T R),$$

on a donc également

$$\|(\widetilde{R}^T\widetilde{R})^{-1}\|_2 = \rho((\widetilde{R}^T\widetilde{R})^{-1}) = \rho(\widetilde{R}^{-1}(\widetilde{R}^T)^{-1}) = \rho((\widetilde{R}^T)^{-1}\widetilde{R}^{-1})$$

On a utilisé le résultat montré dans le chapitre 2 : $\rho(AB) = \rho(BA)$ on sait d'autre part que $(\widetilde{R}^T)^{-1} = (\widetilde{R}^{-1})^T$, ce qui permet de terminer la démonstration.

Ce résultat est important car dans le cas des équations normales on est conduit à résoudre un système dont la matrice est $A^T A$, dans le cas de la factorisation QR on est amené à résoudre un système dont la matrice est $\widetilde{R}$, comme vous le savez le conditionnement est toujours supérieur à 1 donc la matrice $\widetilde{R}$ a un conditionnement plus faible que la matrice $A^T A$, ce qui est intéressant numériquement.

$\square$