TELECOM Nancy *(ESIAL)*
Maths Numériques

feuille 6 : moindres carrés

Exercice 1. *Formulations de problèmes de moindres carrés*

1. On cherche à identifier un modèle physique :
$$q = f(t) = \alpha + \beta \exp(-t^2)$$
dont les paramètres sont (α, β). On dispose d'une série de mesures $(t_i, q_i)_{1\le i\le m}$. Écrire le problème correspondant sous forme standard $\min_x \|Ax - b\|^2$.
2. Même problèmatique mais avec un modèle physique à deux variables d'entrées (notées x et y) et une sortie (notée z) donc de la forme "$z = f(x, y)$". On entreprend une série de m mesures $\{((x_i, y_i), z_i),\ 1 \le i \le m\}$ et l'on essaie de déterminer un modèle raisonable $z = f(x, y)$ donnant donc directement z en fonction de x et y.
   (a) On utilise $f_1(x, y) = a + bx + cy$ comme premier modèle. On détermine ses trois paramètres a, b et c par moindres carrés. Écrire ce problème sous la forme standard :
   $$\min_{v\in\mathbb{R}^3} ||Mv - z||^2$$
   où $v = (a\ \ b\ \ c)^\top$ (il s'agit donc de déterminer la matrice M et le vecteur z).
   (b) Même question avec le modèle à 6 paramètres $f_2(x, y) = f_1(x, y) + dx^2 + exy + fy^2$.

Exercice 2. *Moindres carrés avec poids*
Lorsque les mesures n'ont pas la même précision, nous avons vu en cours qu'il est plus intéressant de minimiser la fonction :
$$E_w(x) = \sum_{i=1}^{m} \left(\frac{f(t_i, x) - y_i}{\sigma_i}\right)^2$$
Lorsque le modèle est linéaire en ses paramètres $f(t, x) = \sum_{j=1}^{n} x_j \varphi_j(t)$

1. Montrer que : $E_w(x) = \|A_w x - b_w\|_2^2$ où $(A_w)_{i,j} = \frac{\varphi_j(t_i)}{\sigma_i}$ et $(b_w)_i = \frac{y_i}{\sigma_i}$.
2. Montrer aussi que : $E_w(x) = \|WAx - Wb\|_2^2$ où W est la matrice diagonale $diag(1/\sigma_1, \dots, 1/\sigma_m)$ et A et b la matrice et le vecteur intervenant dans la fonction des moindres carrés sans pondération ($(A)_{i,j} = \varphi_j(t_i)$ et $b_i = y_i$).

Exercice 3. *Un modèle non-linéaire linéarisable.*
Un modèle souvent rencontré est le modèle exponentiel décroissant $y = Ce^{-\frac{t}{\tau}}$ de paramètres C et τ la constante de temps. Pour identifier ces paramètres on dispose de mesures (t_i, y_i) où $y_i > 0, \forall i$.

1. Ce modèle n'est pas linéaire par rapport au paramètre τ : montrer qu'en passant en log on obtient un modèle linéaire de paramètres $x_1 := \log(C)$ et $x_2 := \frac{1}{\tau}$.
2. Ecrire alors le problème sous la forme standard $\min_x E(x) = \|Ax - b\|$ (i.e. donner la matrice A et le vecteur b) puis écrire les équations normales associées.

Exercice 4. *Notions utilisées dans théorème 1*
On considère un e.v. réel E de dimension m et muni d'un produit scalaire noté $(.|.)$ et la norme associée $\|.\|$ ($\|x\| := \sqrt{(x|x)}$). Soit $\mathcal{E} = (e^1, e^2, \dots, e^m)$ une b.o.n. de E, c'est à dire telle que $(e^i|e^j) = \delta_{i,j}$ le symbole de Kronecker. Tout vecteur $x \in E$ admet donc une décomposition unique sur $\mathcal{E}$: $x = \sum_{i=1}^{m} x_i e^i$ où les m réels x_i sont les composantes de x dans la base $\mathcal{E}$. Attention l'espace E n'est pas forcément $\mathbb{R}^m$, on notera $[x]_{/\mathcal{E}}$ le vecteur de $\mathbb{R}^m$ formé par ces composantes x_i.

1. Montrer que $\forall i, x_i = (x|e^i)$.
2. Montrer que $\forall x, y \in E$, $(x|y) = ([x]_{/\mathcal{E}}, [y]_{/\mathcal{E}})$ où $(., .)$ désigne le produit scalaire canonique de $\mathbb{R}^m$.

Exercice 5. *Complexité de la résolution d'un pmc via les équations normales*

Etant donné un pmc donné sous la forme standard, cad une matrice A de taille $m \times n$ et un vecteur $b \in \mathbb{R}^m$ calculer le coût de la résolution en utilisant les équations normales. Comme la matrice $A^\top A$ est symétrique et définie positive, on peut utiliser une variante de Gauss-LU appelée factorisation de Choleski et qui a une complexité moitiée moindre que celle de Gauss-LU de $(1/6)n^3$ additions et multiplications.

Il s'agit de l'algorithme le plus rapide par contre il est moins stable que l'algorithme QR.

Exercice 6. *Un peu d'algèbre linéaire !*

Soit $A \in \mathbb{R}^{m \times n}$. On note $(.|.)$ le produit scalaire canonique de $\mathbb{R}^n$ et aussi de $\mathbb{R}^m$.

1. Montrer que $\forall x \in \mathbb{R}^n$, $\forall y \in \mathbb{R}^m$ on a $(Ax|y) = (x|A^\top y)$ (rmq : le produit scalaire de gauche est dans $\mathbb{R}^m$, le deuxième dans $\mathbb{R}^n$).
2. Montrer[1] que $KerA = (ImA^\top)^\perp$. D'où : $KerA \oplus ImA^\top = \mathbb{R}^n$.
3. En utilisant le théorème du rang sur la matrice A, déduire du résultat précédent que $rang(A^\top) = rang(A)$.
4. Montrer que $KerA \subset KerA^\top A$
5. Montrer que si $A^\top Ax = 0$ alors $x^\top A^\top Ax = 0$. En écrivant cette dernière égalité sous forme d'une norme au carré, en déduire que $KerA^\top A \subset KerA$ (et donc que $KerA = KerA^\top A$).
6. En déduire (en utilisant le théorème du rang sur la matrice $A^\top A$) que $rang(A^\top A) = rang(A)$.
7. Montrer que $ImA^\top A \subset ImA^\top$. En déduire en utilisant le résultat de la question 3 [2] que $ImA^\top A = ImA^\top$.
8. Avec ce dernier résultat nous pouvons montrer que les équations normales $A^\top Ax = A^\top b$ possèdent toujours (au moins) une solution. Montrer en effet que $\forall b \in \mathbb{R}^m$, $A^\top b \in ImA^\top$. Conclusion ?
9. Lorsque $KerA = \{0\}$ nous avons vu que $A^\top A$ est définie positive et donc inversible : il y a alors solution unique $x^* = (A^\top A)^{-1}A^\top b$. Sinon il existe donc une infinité de solutions. Il faut alors d'autres critères pour sélectionner une solution intéressante parmis toutes les solutions admissibles et souvent on choisit la solution de norme minimale. Cependant ce critère conduit à des calculs assez sophistiqués et une alternative est de résoudre :

   $$(\underbrace{A^\top A + \epsilon I}_{G_\epsilon})x = A^\top b$$

   Montrer que $\forall \epsilon > 0$ la matrice G_ϵ est maintenant bien définie positive (et donc inversible). Rmq : on peut montrer que la solution de ce système linéaire est le minimum de la fonction :

   $$E_\epsilon(x) = \underbrace{\|Ax - b\|_2^2}_{=E(x)} + \epsilon \|x\|_2^2$$

   Pour $\epsilon \to 0$ on privilégie la minimisation de $E(x)$ et pour $\epsilon \to +\infty$ la minimisation de $\|x\|$ (ce qui finit par nous donner 0 !).

---

1. Si V est un s.e.v. d'un e.v. réel E de dimension finie p muni d'un produit scalaire, on appelle $V^\perp$ l'ensemble des vecteurs orthogonaux à tous les vecteurs de V : $V^\perp = \{x \in E : (x|y) = 0,\ \forall y \in V\}$; il est facile de montrer que $V^\perp$ est aussi un s.e.v. de E et que $E = V \oplus V^\perp$ (par exemple en utilisant le théorème de projection) et donc $p = dim(V) + dim(V^\perp)$.
2. Ainsi qu'en utilisant le résultat évident suivant : si un s.e.v. E est inclus dans un autre s.e.v. F et si $dim(E) = dim(F)$ alors $E = F$.

Élément de correction :

Exercice 7. *Un peu d'algèbre linéaire !*
Soit $A \in \mathbb{R}^{m\times n}$. On note $(.|.)$ le produit scalaire canonique de $\mathbb{R}^n$ et aussi de $\mathbb{R}^m$.

1. Montrer que $\forall x \in \mathbb{R}^n$, $\forall y \in \mathbb{R}^m$ on a $(Ax|y) = (x|A^\top y)$ (rmq : le produit scalaire de gauche est dans $\mathbb{R}^m$, le deuxième dans $\mathbb{R}^n$).

   réponse : $(Ax|y) = y^\top Ax = (A^\top y)^\top x = (x|A^\top y)$
2. Montrer que $KerA = (ImA^\top)^\perp$. D'où : $KerA \oplus ImA^\top = \mathbb{R}^n$

   réponse :

$$\begin{aligned}(ImA^\top)^\perp &= \{x \in \mathbb{R}^n : (x|y) = 0, \forall y \in ImA^\top\}\\ &= \{x \in \mathbb{R}^n : (x|A^\top z) = 0, \forall z \in \mathbb{R}^m\}\\ &= \{x \in \mathbb{R}^n : (Ax|z) = 0, \forall z \in \mathbb{R}^m\}\\ &= \{x \in \mathbb{R}^n : Ax = 0\} = KerA\end{aligned}$$

   On en déduit donc que $KerA \oplus ImA^\top = \mathbb{R}^n$ d'où $dim(KerA) + rang(A^\top) = n$.
3. En utilisant le théorème du rang sur la matrice A, déduire du résultat précédent que $rang(A^\top) = rang(A)$.

   réponse : D'après le théorème du rang $dim(KerA) + rang(A) = n$ et donc $rang(A^\top) = rang(A)$.
4. Montrer que $KerA \subset KerA^\top A$.

   réponse : Soit $x \in KerA$, c'est à dire $Ax = 0$. Donc $A^\top Ax = A^\top 0 = 0$ soit $x \in KerA^\top A$
5. Montrer que si $A^\top Ax = 0$ alors $x^\top A^\top Ax = 0$. En écrivant cette dernière égalité sous forme d'une norme au carré, en déduire que $KerA^\top A \subset KerA$ (et donc que $KerA = KerA^\top A$).

   réponse : Si $A^\top Ax = 0$ alors $x^\top A^\top Ax = x^\top 0 = 0$. Or $x^\top A^\top Ax = (Ax)^\top Ax = (Ax|Ax) = \|Ax\|^2$. D'où si $x \in KerA^\top A$ alors $\|Ax\|^2 = 0 \iff Ax = 0$ et donc $x \in KerA$.
6. En déduire (en utilisant le théorème du rang sur la matrice $A^\top A$) que $rang(A^\top A) = rang(A)$.

   réponse : D'après le théorème du rang $dim(KerA^\top A) + rang(A^\top A) = n$. D'où $rang(A^\top A) = n - dim(KerA^\top A) = n - dim(KerA) = rang(A)$.
7. Montrer que $ImA^\top A \subset ImA^\top$. En déduire en utilisant le résultat de la question 3 que $ImA^\top A = ImA^\top$.

   réponse : $z \in ImA^\top A \iff \exists x \in \mathbb{R}^n : z = A^\top Ax$. En posant $y = Ax$, on a trouvé $y \in \mathbb{R}^m$ tel que $z = A^\top y$, c'est à dire que $z \in ImA^\top$, donc $ImA^\top A \subset ImA^\top$. D'après les questions précédentes $rang(A^\top A) = rang(A) = rang(A^\top)$ d'où $Im(A^\top A) = ImA^\top$.
8. Avec ce dernier résultat nous pouvons montrer que les équations normales $A^\top Ax = A^\top b$ possèdent toujours (au moins) une solution. Montrer en effet que $\forall b \in \mathbb{R}^m$, $A^\top b \in ImA^\top A$. Conclusion ?

   réponse : On a par définition que $A^\top b \in ImA^\top$ donc $A^\top b \in ImA^\top A$. Le second membre des e.n. est donc toujours dans l'image de la matrice des e.n. : il existe donc toujours (au moins) une solution !