MT09 A25 - Feuille de TD n° 8 bis Approximation aux moindres carrés (suite)

Exercice 1 : régression polynomiale et régularisation de Tykhonov

Soit $(x_i, y_i)_{i=1,\dots,m} \in \mathbb{R}^2$ un nuage de points. Soit $n \leq m-1$ et $\mu \geq 0$. Soit

$$f(x) = u_0 + u_1 x + \dots + u_n x^n$$

un polynôme de régression. On note $\boldsymbol{u}=(u_0,u_1,...,u_n)^T\in\mathbb{R}^{n+1}$. On cherche à minimiser la fonction de perte

$$\mathscr{L}_{\mu}(\boldsymbol{u}) = \frac{1}{2} \sum_{i=1}^{m} \left(f(x_i) - y_i \right)^2 + \frac{1}{2} \mu \left((u_0 - \bar{y})^2 + \sum_{i=1}^{n} u_i^2 \right)$$

où
$$\bar{y} = \frac{1}{m} \sum_{i=1}^{m} y_i$$
.

1. Réécrire $\mathcal{L}_{\mu}(\boldsymbol{u})$ sous la forme

$$\mathscr{L}_{\mu}(u) = \frac{1}{2} ||Au - y||^2 + \frac{1}{2} \mu ||u - u^0||^2$$

où l'on précisera la matrice A et sa taille, ainsi que les vecteurs \boldsymbol{y} et \boldsymbol{u}^0 (et leur taille).

2. Montrer que les solutions u_{μ} réalisant le minimum de

$$\min_{oldsymbol{u} \in \mathbb{R}^{n+1}} \, \mathscr{L}_{\mu}(oldsymbol{u})$$

sont solutions du système linéaire

$$(A^T A + \mu I_{n+1}) \boldsymbol{u}_{\mu} = A^T \boldsymbol{y} + \mu \boldsymbol{u}^0.$$

- 3. Pour $\mu > 0$, que peut-on dire de la matrice $A^T A + \mu I_{n+1}$?
- 4. Que vaut $\lim_{\mu \to \infty} u_{\mu}$? Quelle est la fonction de régression correspondante?
- 5. Que vaut

$$\operatorname{Cond}_{\|.\|_2}(A^T A + \mu I_{n+1}) ?$$

(exprimer le résultat en fonction des valeurs propres de ${\cal A}^T{\cal A}$). Que vaut

$$\lim_{\mu \to \infty} \operatorname{Cond}_{\|.\|_2} (A^T A + \mu I_{n+1}) ?$$

6. Que préconisez-vous comme choix de μ ?

Exercice 2: factorisation QR "économique" et moindres carrés

Soit $A \in \mathcal{M}_{mn}(\mathbb{R})$, $m \geq n$ et $\mathbf{y} \in \mathbb{R}^n$. On suppose que $\operatorname{rg}(A) = n$. On cherche à résoudre les équations normales

$$A^T A \boldsymbol{u} = A^T \boldsymbol{y}$$

de manière rapide et le mieux conditionné possible. On note A_j les vecteurs colonnes de A. On rappelle le procédé d'orthogonalisation de Gram-Schmidt ici appliqué aux n vecteurs A_j :

$$egin{aligned} oldsymbol{q}_1 &= rac{oldsymbol{A}_1}{\|oldsymbol{A}_1\|_2}, \ & \hat{oldsymbol{q}}_2 &= oldsymbol{A}_2 - raket{oldsymbol{A}_2, oldsymbol{q}_1} oldsymbol{q}_1, \qquad oldsymbol{q}_2 &= rac{\hat{oldsymbol{q}}_2}{\|\hat{oldsymbol{q}}_2\|_2}, \end{aligned}$$

:

$$\hat{oldsymbol{q}}_n = oldsymbol{A}_n - \sum_{i=1}^{n-1} \langle oldsymbol{A}_n, oldsymbol{q}_j
angle \, oldsymbol{q}_j, \qquad oldsymbol{q}_n = rac{\hat{oldsymbol{q}}_n}{\|\hat{oldsymbol{q}}_n\|_2}.$$

1. Montrer que A s'écrit

$$A = Q\tilde{R}$$

avec $\tilde{R} \in \mathcal{M}_n(\mathbb{R})$ matrice triangulaire supérieure et $Q = [q_1, ..., q_n] \in \mathcal{M}_{mn}(\mathbb{R})$ telle que $Q^TQ = I_n$.

- 2. A-t-on $QQ^T = I_m$?
- 3. La matrice \tilde{R} est-elle inversible?
- 4. Montrer que les équations normales peuvent s'écrire de manière équivalente

$$\tilde{R}\,\boldsymbol{u} = Q^T\boldsymbol{y}.$$

Comment résoudre ce système?

Commentaire. On peut montrer que

$$\operatorname{Cond}_{\|.\|_2}(\tilde{R}) = \sqrt{\operatorname{Cond}_{\|.\|_2}(A^T A)},$$

ce qui est un véritable apport du point de vue de la robustesse de la solution aux erreurs d'arrondis ou aux incertitudes sur les données.

Exercice 3: projection orthogonale

Soit A_1 et y deux vecteurs de \mathbb{R}^n .

- 1. Comment caractérise-t-on la projection orthogonale z^* de y sur vect (A_1) ?
- 2. Résoudre

$$\min_{\alpha \in \mathbb{R}} \quad \frac{1}{2} \| \boldsymbol{A}_1 \alpha - \boldsymbol{y} \|_2^2$$

par les équations normales. Soit α^* le scalaire qui réalise ce minimum.

- 3. Exprimez z^* en fonction de A_1 et α^* .
- 4. Soit

$$\hat{\boldsymbol{A}}_1 = \frac{\boldsymbol{A}_1}{\|\boldsymbol{A}_1\|_2}.$$

Exprimez z^* en fonction de \hat{A}_1 et y.