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

Exercice 1 : régression polynomiale et régularisation de Tykhonov

Soit $\{(x_i, y_i),\ i = 1, ..., m\}$ un nuage de points dans $\mathbb{R}^2$. Soit n un entier, $n \leq m-1$ et un paramètre $\mu \geq 0$. On considère ici un polynôme de régression de degré inférieur ou égal à n :

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

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

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

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

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

   $$\mathscr{L}_\mu(\boldsymbol{u}) = \frac{1}{2} \|A\boldsymbol{u} - \boldsymbol{y}\|_2^2 + \frac{1}{2}\,\mu\,\|\boldsymbol{u} - \boldsymbol{u}^0\|_2^2$$

   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 $\boldsymbol{u}_\mu$ réalisant le minimum de

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

   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 $\lim\limits_{\mu \to \infty} \boldsymbol{u}_\mu$? Quelle est la fonction de régression correspondante ?
5. Que vaut

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

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

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

6. Que préconisez-vous comme choix de μ ?

Exercice 2 : factorisation QR "économique" et moindres carrés

Soit $A \in \mathscr{M}_{mn}(\mathbb{R})$, $m \geq n$ et $\boldsymbol{y} \in \mathbb{R}^m$. On suppose que $\mathrm{rg}(A) = n$. 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 $\boldsymbol{A}_j$ les vecteurs colonnes de A. On rappelle le procédé d'orthogonalisation de Gram-Schmidt ici appliqué aux n vecteurs $\boldsymbol{A}_j$:

$$\begin{aligned}
&\boldsymbol{q}_1 = \frac{\boldsymbol{A}_1}{\|\boldsymbol{A}_1\|_2},\\
&\hat{\boldsymbol{q}}_2 = \boldsymbol{A}_2 - \langle \boldsymbol{A}_2, \boldsymbol{q}_1 \rangle \, \boldsymbol{q}_1, \qquad \boldsymbol{q}_2 = \frac{\hat{\boldsymbol{q}}_2}{\|\hat{\boldsymbol{q}}_2\|_2},\\
&\vdots\\
&\hat{\boldsymbol{q}}_n = \boldsymbol{A}_n - \sum_{j=1}^{n-1} \langle \boldsymbol{A}_n, \boldsymbol{q}_j \rangle \, \boldsymbol{q}_j, \qquad \boldsymbol{q}_n = \frac{\hat{\boldsymbol{q}}_n}{\|\hat{\boldsymbol{q}}_n\|_2}.
\end{aligned}$$

1. Montrer que A s'écrit
$$A = Q\tilde{R}$$
avec $\tilde{R} \in \mathscr{M}_n(\mathbb{R})$ matrice triangulaire supérieure et $Q = \left[\boldsymbol{q}_1, ..., \boldsymbol{q}_n\right] \in \mathscr{M}_{mn}(\mathbb{R})$.
2. Montrez que $Q^T Q = I_n$. 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

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

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 $\boldsymbol{A}_1$ et $\boldsymbol{y}$ deux vecteurs de $\mathbb{R}^n$.

1. Comment caractérise-t-on la projection orthogonale $\boldsymbol{z}^\star$ de $\boldsymbol{y}$ sur $\mathrm{vect}(\boldsymbol{A}_1)$?
2. Résoudre
$$\min_{\alpha \in \mathbb{R}} \quad \frac{1}{2}\|\boldsymbol{A}_1 \alpha - \boldsymbol{y}\|_2^2$$
au moyen des équations normales. Soit $\alpha^\star$ le scalaire qui réalise ce minimum.
3. Exprimez $\boldsymbol{z}^\star$ en fonction de $\boldsymbol{A}_1$ et $\alpha^\star$.
4. Soit
$$\hat{\boldsymbol{A}}_1 = \frac{\boldsymbol{A}_1}{\|\boldsymbol{A}_1\|_2}.$$
Exprimez $\boldsymbol{z}^\star$ en fonction de $\hat{\boldsymbol{A}}_1$ et $\boldsymbol{y}$. Le résultat vous paraît-il cohérent ?