Solution du devoir maison : Statistique pour les big data (année 2021-2022)

Patrick Tardivel,
Université de Bourgogne, Dijon

Exercice 1. *Soient $\tau \geq 0$, $v \in \mathbb{R}^p$, on définit v^τ et $v^{\overline{\tau}}$ par*

$$\forall i \in \{1,\dots,p\}, v_i^\tau = v_i \mathbf{1}(|v_i| > \tau) \text{ et } v_i^{\overline{\tau}} = v_i \mathbf{1}(|v_i| \leq \tau).$$

1) Montrer que $v = v^\tau + v^{\overline{\tau}}$ et que $\|v\|_1 = \|v^\tau\|_1 + \|v^{\overline{\tau}}\|_1$.

Soit $A \in \mathbb{R}^{n\times p}$ et $x \in \mathbb{R}^p$ un vecteur identifiable relativement à A et à la norme ℓ_1.

1. *Montrer que x^τ est un vecteur identifiable relativement à A et à la norme ℓ_1. Pour résoudre cette question, on pourra remarquer que $Az = Ax^\tau$ si et seulement si $A(z + x^{\overline{\tau}}) = Ax$, appliquer la question 1) et utiliser simplement la définition "d'identifiabilité".*

Solution : 1) Soit $i \in \{1,\dots,p\}$ alors,

$$v_i^\tau + v_i^{\overline{\tau}} = v_i \mathbf{1}(|v_p| > \tau) + v_i \mathbf{1}(|v_p| \leq \tau) = v_i(\underbrace{\mathbf{1}(|v_p| > \tau) + \mathbf{1}(|v_p| \leq \tau)}_{=1}) = v_i.$$

Ainsi, $v = v^\tau + v^{\overline{\tau}}$. Par ailleurs, comme $\mathrm{supp}(v^\tau) \cap \mathrm{supp}(v^{\overline{\tau}}) = \emptyset$ et que $\mathrm{supp}(v^\tau) \cup \mathrm{supp}(v^{\overline{\tau}}) = \{1,\dots,p\}$ on a

$$\|v\|_1 = \sum_{1\leq i\leq p} |v_i| = \sum_{i\in supp(v^\tau)} |v_i| + \sum_{i\in supp(v^{\overline{\tau}})} |v_i| = \|v^\tau\|_1 + \|v^{\overline{\tau}}\|_1.$$

2) Soit $z \in \mathbb{R}^p$ tel que $Az = Ax^\tau$, remarquons que

$$Az = Ax^\tau \Leftrightarrow Az + Ax^{\overline{\tau}} = Ax^\tau + Ax^{\overline{\tau}} \Leftrightarrow A(z + x^{\overline{\tau}}) = Ax.$$

Comme x is identifiable relativement à A et à la norme ℓ_1 alors $\|z + x^{\overline{\tau}}\|_1 \geq \|x\|_1$. Ainsi,

$$\begin{aligned} & \|z + x^{\overline{\tau}}\|_1 \geq \|x\|_1 \\ \Rightarrow\quad & \|z\|_1 + \|x^{\overline{\tau}}\|_1 \geq \|x\|_1 = \|x^\tau\|_1 + \|x^{\overline{\tau}}\|_1 \\ \Rightarrow\quad & \|z\|_1 \geq \|x^\tau\|_1. \end{aligned}$$

Ce qui montre que x^τ est identifiable relativement à A et à la norme ℓ_1.

Exercice 2. *Soient $A \in \mathbb{R}^{n\times p}$ et $s \in \{-1,0,1\}^p$. Pour vérifier que s est identifiable relativement à A et à la norme ℓ_1, on résout numériquement le problème poursuite de base du système linéaire d'équations $Ax = b$ avec $b = As$. On suppose que la solution poursuite de base est unique et on note x^* la solution théorique de ce problème. Une méthode numérique (par exemple celle que vous avez implémentée avec Monsieur Dupuis) donne une solution approchée x^{num} du problème poursuite de base et on suppose que $\|x^* - x^{num}\|_\infty < 1/2$. On pose $arrond(x^{num})$ le vecteur obtenu en arrondissant chaque composante de x^{num} à l'entier le plus proche.*

1) Justifier que lorsque s est identifiable relativement à A et à la norme ℓ_1 alors $arrond(x^{num}) = s$

2) Montrer que lorsque s n'est identifiable relativement à A et à la norme ℓ_1 alors $arrond(x^{num}) \neq s$. Pour cette question on pourra montrer que $arrond(x^{num})$ est (toujours) identifiable identifiable relativement à A et à la norme ℓ_1.

Solution : On pose $\overline{z} = arrond(x^{num})$ et $z = x^{num}$. Si s est identifiable relativement à A et à la norme ℓ_1 alors, $x^* = s$ et par hypothèse sur la distance entre x^{num} et x^* on a $\|z - s\|_\infty < 1/2$. Ainsi

— Lorsque $s_i = 1$ on a $z \in]0,5; 1,5[$ d'où $\overline{z}_i = 1$.

— Lorsque $s_i = 0$ on a $z \in]-0,5; 0,5[$ d'où $\overline{z}_i = 0$.

— Lorsque $s_i = -1$ on a $z \in]-1,5; -0,5[$ d'où $\overline{z}_i = -1$.

Ainsi, lorsque s est identifiable relativement à A et à la norme ℓ_1 on a $\overline{z} = s$.

Montrons que $\overline{z}$ est identifiable relativement à A et à la norme ℓ_1. On peut remarquer le faits suivant

— Si $\overline{z}_i > 0$ alors, $z_i \geq 0,5$ et donc $x_i^* > 0$ (car $|z_i - x_i^*| < 1/2$). De même, si $\overline{z}_i < 0$ alors, $z_i \leq -0,5$ et donc $x_i^* < 0$.

— Clairement, si $x_i^* = 0$ alors $z \in]-0,5; 0,5[$ d'où $\overline{z}_i = 0$. Ainsi, $zeros(\overline{z}) \subset zeros(x^*)$ ou encore $\mathrm{supp}(\overline{z}) \subset \mathrm{supp}(x^*)$.

Soit $h \in \ker(A)$ et comme x^* est identifiable alors

$$\begin{aligned}
0 &\geq |\sum_{i\in supp(x^*)} sign(x_i^*)h_i| - \sum_{i\notin supp(x^*)} |h_i| \\
0 &\geq \underbrace{|\sum_{i\in supp(x^*)} sign(x_i^*)h_i| + \sum_{i supp(x^*)\setminus supp(\overline{z})} |h_i|}_{\geq |\sum_{i\in supp(\overline{z})} sign(x_i^*)h_i|} - \underbrace{(\sum_{i supp(x^*)\setminus supp(\overline{z})} |h_i| + \sum_{i\notin supp(x^*)} |h_i|)}_{=\sum_{i\notin supp(\overline{z})} |h_i|} \\
0 &\geq |\sum_{i\in supp(\overline{z})} sign(x_i^*)h_i| - \sum_{i\notin supp(\overline{z})} |h_i|
\end{aligned}$$

Exercice 3 (Exercice optionnel). *Soient $A := (A_1|\ldots|A_p) \in \mathbb{R}^{n\times p}$ où $\|A_1\|_2 = \cdots = \|A_p\|_2 = 1$, $b \in \mathrm{col}(A)$, x^* une solution du système linéaire $Ax = b$, $I := \mathrm{supp}(x^*)$, A_I la matrice dont les colonnes sont $(A_i)_{i\in I}$ et $\mathrm{signe}(x_I^*) = (\mathrm{signe}(x_i^*))_{i\in I}$. On veut prouver l'implication suivante*

$$\|x^*\|_0 \leq (1 + 1/M(A))/2 \Rightarrow \forall j \notin I, |A_j^T A_I (A_I^T A_I)^{-1} \mathrm{signe}(x_I^*)| \leq 1. \qquad (1)$$

Montrons dans un premier temps le résultat suivant qui localise les valeurs propres.

1) Soit $Q \in \mathbb{R}^{n\times n}$. Si $\lambda \in \mathbb{R}$ est une valeur propre de Q alors $\lambda \in [Q_{ii} - \sum_{j\neq i}|Q_{ij}|, Q_{ii} + \sum_{j\neq i}|Q_{ij}|]$ pour un certain $i \in \{1,\ldots,n\}$.

En utilisant la question i), nous sommes en mesure d'établir l'inégalité (1).

1) Supposons que $\|x^*\|_0 \leq (1 + 1/M(A))/2$. Démontrer que la plus grande valeur propre de $(A_I^T A_I)^{-1}$ est inférieure à $2/(M(A) + 1)$.

2) En utilisant la question 2), prouver que $\|x^*\|_0 \leq (1 + 1/M(A))/2$ implique l'inégalité suivante

$$\forall j \notin I, |A_j^T A_I (A_I^T A_I)^{-1} \mathrm{signe}(x_I^*)| \leq 1.$$

Solution : 1) Soit $v \in \mathbb{R}^n \setminus \{0\}$ un vecteur propre de Q pour la valeur propre λ et $i_0 \in \{1,\ldots,n\}$ tel que

$|v_{i_0}| = \|v\|_\infty$. Comme la composante i_0 du vecteur Qv vaut λv_{i_0} on en déduit les égalités suivantes :

$$Q_{i_0 i_0} v_{i_0} + \sum_{j \neq i_0} Q_{i_0 j} v_j = \lambda v_{i_0} \Rightarrow \lambda = Q_{i_0 i_0} + \sum_{j \neq i_0} Q_{i_0 j} \frac{v_j}{v_{i_0}}.$$

En majorant et en minorant le terme $\sum_{j \neq i_0} Q_{i_0 j} \frac{v_j}{v_{i_0}}$ par $\pm \sum_{j \neq i_0} |Q_{i_0 j}|$ on obtient l'encadrement de λ souhaité :

$$Q_{i_0 i_0} - \sum_{j \neq i_0} |Q_{i_0 j}| \leq \lambda \leq Q_{i_0 i_0} + \sum_{j \neq i_0} |Q_{i_0 j}|.$$

2) On remarque que $A_I^T A_I$ est une matrice symétrique (donc diagonalisable) de dimension $\|x^*\|_0$ dont les composantes diagonales sont toutes égales à 1 et dont les composantes à l'extérieure de la diagonale sont inférieures en valeur absolue à $M(A)$. Ainsi, d'après la question i), la plus petite valeur propre λ de $A_I^T A_I$ vérifie

$$\lambda \geq 1 - (\|x^*\|_0 - 1) M(A) \geq 1 - \frac{M(A)+1}{2} = \frac{1 - M(A)}{2}.$$

Comme λ^{-1} est la plus grande valeur propre $(A_I^T A_I)^{-1}$; la plus grande valeur propre de cette matrice est donc inférieure à $2/(M(A)+1)$.

3) D'après l'inégalité de Cauchy-Schwartz on a

$$|A_j^T A_I (A_I^T A_I)^{-1} \text{signe}(x_I^*)| \leq \|A_j^T A_I\|_2 \|(A_I^T A_I)^{-1} \text{signe}(x_I^*)\|_2.$$

— Par définition de la cohérence mutuelle, $A_j^T A_I$ est un vecteur dont les composantes sont toutes inférieures en valeur absolue à $M(A)$ ainsi
$$\|A_j^T A_I\|_2 \leq \sqrt{\|x^*\|_0 M(A)^2}$$

— Comme la plus grande valeur propre de $(A_I^T A_I)^{-1}$ est inférieure à $2/(M(A)+1)$ ainsi

$$\|(A_I^T A_I)^{-1} \text{signe}(x_I^*)\|_2 \leq \frac{2}{M(A)+1} \|\text{signe}(x_I^*)\|_2 = \frac{2}{M(A)+1} \sqrt{\|x^*\|_0}.$$

Par conséquent, on obtient l'inégalité souhaitée :

$$|A_j^T A_I (A_I^T A_I)^{-1} \text{signe}(x_I^*)| \leq \|x^*\|_0 \frac{2M(A)}{M(A)+1} = 1.$$