Matrices d'intervalle et normes

Jérémie Dumas
École Normale Supérieure de Lyon

Janvier 2012

1 Généralités

1.1 Introduction

Dans un rapport technique assez récent [FRL11], Farhadsefat et al. discutent de la complexité de calcul d'une norme pour les matrices d'intervalle. Nous avons vu en cours que le calcul de la norme $\|\cdot\|_{\infty,1}$ est NP-dur pour des matrices de $\mathbb{R}^{m\times n}$, il l'est donc aussi pour la norme induite sur les matrices d'intervalle de $\mathbb{IR}^{m\times n}$. Nous verrons alors comment calculer la norme induite $|||\cdot|||$ par évaluation sur 2^{mn} points d'un rectangle de dimension $m \times n$, ou 2^{m+n} dans le cas de la norme quadratique. Enfin nous terminerons par présenter le résultat principal de [FRL11], qui permet d'évaluer une norme induite en un seul point d'une matrice d'intervale, dans le cas des normes absolues.

1.2 Définitions

Les notations reprennent celles vues en cours, sur la représentation inf-sup et centre-rayon des matrices d'intervalle. On notera $\mathbf{A} = [\underline{A}, \overline{A}] = [A_c - \Delta, A_c + \Delta]$ la matrice d'intervalle de $\mathbb{IR}^{m\times n}$ qui représente l'ensemble $\{A, \underline{A} \leqslant A \leqslant \overline{A}\}$, où les inégalités sont entendues composante par composante. Les opérations d'addition et de multiplication par un scalaire sont les mêmes que vues en cours, à savoir $\mathbf{A} \diamond \mathbf{B} = Hull\{A \diamond B, A \in \mathbf{A}, B \in \mathbf{B}\}$. On ne s'intéresse pas à la multiplication de matrices ici

L'espace $\mathbb{IR}^{m\times n}$ des matrices d'intervalle à m lignes et n colonne n'est pas à proprement parler un espace vectoriel, car la multiplication par un scalaire n'est pas distributive ($(\alpha+\beta)\mathbf{A} \neq \alpha\mathbf{A} + \beta\mathbf{A}$ en général), et les matrices ne possèdent toujours d'inverses pour l'addition ($\mathbf{A} + \mathbf{B} = [0, 0]$ n'est possible que si $\mathbf{A} = \mathbf{B} = [0, 0]$).

1.3 Norme induite

On introduit néanmoins une définition de norme dans l'espace $\mathbb{IR}^{m\times n}$ de la manière suivante : une fonction $|||\cdot||| : \mathbb{IR}^{m\times n} \to \mathbb{R}$ est appelée une *norme de matrice d'intervalle* si elle vérifie les points suivants :

1. $\forall \mathbf{A}, |||\mathbf{A}||| \geqslant 0$ et $|||\mathbf{A}||| = 0 \Leftrightarrow \mathbf{A} = [0,0]$
2. $\forall \mathbf{A}, \mathbf{B}, |||\mathbf{A}+\mathbf{B}||| \leqslant |||\mathbf{A}||| + |||\mathbf{B}|||$
3. $\forall \mathbf{A}, \alpha, |||\alpha\mathbf{A}||| = |\alpha|\,|||\mathbf{A}|||$

Farhadsefat et al. montrent alors pour toute norme de matrice $\|\cdot\|$ dans $\mathbb{R}^{m\times n}$, la norme $|||\cdot|||$ induite sur $\mathbb{IR}^{m\times n}$ par $|||\mathbf{A}||| = \sup\{\|A\|, A \in \mathbf{A}\}$ est bien une norme au sens défini précédemment.

Une norme induite $|||\cdot|||$ se caractérise alors de manière unique par la propriété :

$$|||\mathbf{A}||| = \max\{|||[A,A]|||, A \in \mathbf{A}\}$$

Auquel cas la norme inductante de $\mathbb{R}^{m\times n}$ peut se retrouver en posant $\|A\| = |||[A,A]|||$.

2 Calcul des normes

2.1 Cas général

Étant donnée une matrice Z qui vérifie $|Z| = E = [1]_{i,j}$ et $\mathbf{A} = [A_c - \Delta, A_c + \Delta]$, on introduit la notation $A_Z = A_c + Z \circ \Delta$, où $\circ$ dénote le produit de Hadamard (composante par composante). Montrons alors le résultat suivant :

Théorème 2.1. *Si $|||\cdot|||$ est une norme de matrice d'intervalle induite par $\|\cdot\|$, alors $|||\boldsymbol{A}||| = \max_{|Z|=E}\|A_Z\|$ pour chaque $\boldsymbol{A} \in \mathbb{IR}^{m\times n}$.*

Démonstration. En voyant $\mathbf{A}$ comme un rectangle de $\mathbb{IR}^{m\times n}$, on peut exprimer un point intérieur $A \in \mathbf{A}$ comme une somme positive des coins du pavé : $A = \sum_{|Z|} \lambda_Z A_Z$, où les λ_Z sont $\geqslant 0$ de somme égale à 1. De fait, on déduit que $\|A\| \leqslant \max\|A_Z\| \sum \lambda_Z = \max\|A_Z\|$. En passant au sup, on trouve que $|||\mathbf{A}||| \leqslant \max\|A_Z\|$.

Inversement, les A_Z sont dans $\mathbf{A}$, donc on a aussi $\max\|A_Z\| \leqslant |||\mathbf{A}|||$. Ce qui prouve bien l'égalité annoncée $|||\mathbf{A}||| = \max\|A_Z\|$. □

2.2 Norme quadratique

Dans le cas de la norme $\|\cdot\|_2$, un résultat analogue au théorème 2.1 est donné, mais ne nécessitant l'évaluation que de 2^{m+n} points de la matrice $\mathbf{A}$, contre 2^{mn} pour le cas général (théorème 2.1). La norme $\|\cdot\|_2$ est définie sur $\mathbb{R}^{m\times n}$ par $\|A\|_2 = \max_{\|x\|_2}\|Ax\|_2$. On a aussi $\|A\|_2 = \sigma_{\max}(A)$, où $\sigma_{\max}(A)$ désigne la plus grande valeur singulière de A.

La suite du papier donne deux théorèmes sur le calcul de la norme $\|\cdot\|_2$ et de sa norme induite, notamment :

Théorème 2.2.
$$\forall A \in \mathbb{R}^{m\times n}, \|A\|_2 = \max_{\|x_1\|_2=\|x_2\|_2=1} x_1^T A x_2.$$

Théorème 2.3.
$$\forall \boldsymbol{A} \in \mathbb{IR}^{m\times n}, \|\!|\boldsymbol{A}|\!\|_2 = \max_{|y|=e_m,|z|=e_n} \|A_c + (yz^T)\circ\Delta\|_2.$$

Le théorème 2.2 se démontre en décomposant A en valeurs singulières, et puis en faisant jouer une inégalité de Cauchy-Schwarz. Pour démontrer le théorème 2.3, on prend une instance $A \in \mathbf{A}$ et on cherche à majorer le scalaire $x_1^T A x_2$ par $x_1^T A_c x_2 + |x_1|^T \Delta |x_2| = x_1^T(A_c + U\Delta V)x_2$, où U et V sont des matrices diagonales avec des ± 1 sur la diagonale. L'autre sens découlant du fait que $A_c + (yz^T)\circ\Delta \in \mathbf{A}$.

2.3 Normes absolues

Pour d'autres normes de matrices usuelles, à savoir les normes $\|\cdot\|_1$, $\|\cdot\|_\infty$, $\|\cdot\|_{1,\infty}$ et $\|\cdot\|_F$ (norme de Frobenius), les auteurs de [FRL11] démontrent qu'il est facile de calculer leurs normes induites sur les matrices d'intervalles. Plus exactement, il suffit d'évaluer la norme $\|\!|\cdot|\!\|$ induite en un seut point de la matrice d'intervalle pour conclure. Cela tient à la propriété qu'on ces normes d'êtres *absolues*, c'est-à-dire que $\||A|\|_\alpha = \|A\|_\alpha$ pour toute $A \in \mathbb{R}^{m\times n}$ et $\alpha \in \{1, \infty, (1,\infty), F\}$, où $|A|$ est entendu composante par composante.

Dans l'essentiel, leur démonstration repose sur la caractérisation que $\|\cdot\|$ est une norme absolue sur $\mathbb{R}^{m\times n}$ ssi $\forall A, B, |A| \leqslant |B| \Rightarrow \|A\| \leqslant \|B\|$. Pour une matrice A_c, notons $\operatorname{sgn}(A_c)$ la matrice définie par $\operatorname{sgn}(A_c)_{i,j} = 1$ si $(A_c)_{i,j} \geqslant 0$ et -1 sinon. On peut alors énoncer le :

Théorème 2.4. *Si $\|\cdot\|$ est une norme absolue, alors la norme induite $\|\!|\cdot|\!\|$ vérifie :*
$$\forall \boldsymbol{A} \in \mathbb{IR}^{m\times n}, \|\!|A|\!\| = \|A_c + \operatorname{sgn}(A_c)\circ\Delta\| = \||A_c| + \Delta\|$$

Démonstration. Sans entrer dans le détail, on note que pour toute matrice $A \in \mathbf{A}$, on a $|A| = |A_c + A - A_c| \leqslant ||A_c| + \Delta|$. Ce qui, d'après la caractérisation précédente des normes absolues, implique que $\|A\| = \||A_c\| + \Delta\|$, et donc $\|\!|\mathbf{A}|\!\| \leqslant \||A_c| + \Delta\|$.

Or, on a aussi $|A_c| + \Delta = \operatorname{sgn}(A_c)\circ A_c + \Delta$. Et donc (en prenant la valeur absolue) $|A_c| + \Delta = |A_c + \operatorname{sgn}(A_c)\circ\Delta|$. D'où également $\||A_c| + \Delta\| = \|A_c + \operatorname{sgn}(A_c)\circ\Delta\|$. De fait, comme $A_c + \operatorname{sgn}(A_c)\circ\Delta \in \mathbf{A}$, cela prouve enfin l'autre sens de l'inégalité (la borne est atteinte). □

Conclusion

Nous pourrions donc résumer l'article [FRL11] de la manière suivante :

- Calcul de $\|\!|\mathbf{A}|\!\|$ possible par un max sur 2^{mn} matrices en général.
- Calcul de $\|\!|\mathbf{A}|\!\|_2$ possible par un max sur 2^{m+n} matrices.
- Calcul de $\|\!|\mathbf{A}|\!\|_\alpha$ pour $\alpha \in \{1, \infty, (1,\infty), F\}$ par une évaluation en un seul point.

Ainsi que les points encore inconnus suivant :

- Le calcul de $\|\cdot\|_2$ est-t-il NP-dur ?
- Construire une norme sur $\mathbb{IR}^{m\times n}$ qui ne soit induite par aucune norme de matrice ?

Références

[FRL11] Raena Farhadsefat, Jiři Rohn, and Taher Lotfi. Norms of Interval Matrices. Technical Report V-1122, Institute of Computer Science, Academy of Sciences of the Czech Republic, 2011.