Inégalité de Kantorovich

Salim Rostam

13 juillet 2014

Ce document a pour but de faire un peu plus connaître une preuve de l'inégalité de Kantorovich que je n'ai pas vu durant la prépa agreg.

Soit A une matrice symétrique définie positive de taille $n \in \mathbb{N}^*$. On note $0 < \lambda_1 \leq \cdots \leq \lambda_n$ les valeurs propres de A.

Théorème (inégalité de Kantorovich)**.** *Pour tout vecteur $x \in \mathbb{R}^n$ on a l'inégalité suivante :*

$$\langle Ax, x\rangle\langle A^{-1}x, x\rangle \leq \frac{1}{4}\left(\sqrt{\frac{\lambda_1}{\lambda_n}} + \sqrt{\frac{\lambda_n}{\lambda_1}}\right)^2 \|x\|^4$$

Remarque. Par l'inégalité de Cauchy–Schwarz appliquée au produit scalaire associé à A (*i.e.* $\langle x, y\rangle_A := \langle Ax, y\rangle$) on obtient que $\langle Ax, x\rangle\langle A^{-1}x, x\rangle = \|x\|_A^2\|A^{-1}x\|_A^2 \geq \langle A^{-1}x, x\rangle_A^2 = \langle x, x\rangle^2 = \|x\|^4$; le théorème fournit une inégalité dans l'autre sens.

Notations. Pour démontrer le théorème, remarquons tout d'abord qu'il suffit de le démontrer pour x de norme 1 (on utilise ensuite la bilinéarité du produit scalaire). En décomposant x dans une base orthonormale (e_i) de vecteur propres de A (avec e_i le vecteur propre associé à λ_i), on obtient :

$$\langle Ax, x\rangle = \sum_{i=1}^{n} \lambda_i x_i^2 \tag{1}$$

Or, comme $Ae_i = \lambda_i e_i$ on a $\lambda_i^{-1}e_i = A^{-1}e_i$ donc on a également :

$$\langle A^{-1}x, x\rangle = \sum_{i=1}^{n} \lambda_i^{-1} x_i^2 \tag{2}$$

Ainsi, si l'on note X la variable aléatoire qui prend la valeur λ_i avec une probabilité de x_i^2 (note [1]), les équations (1) et (2) deviennent :

$$\langle Ax, x\rangle = \mathbb{E}[X]$$
$$\langle A^{-1}x, x\rangle = \mathbb{E}\left[\frac{1}{X}\right]$$

1. C'est une variable aléatoire bien définie car $0 \leq x_i^2$ et comme x est de norme 1 on a $\sum x_i^2 = 1$.

et l'on cherche donc à majorer $\mathbb{E}[X]\mathbb{E}\left[\frac{1}{X}\right]$; on va en fait simplement chercher une majoration de $\mathbb{E}\left[\frac{1}{X}\right]$.

Parachutes. Par définition de la variable X et comme les λ_i sont triées par ordre croissant, on a $0 < \lambda_1 \leq X \leq \lambda_n$ presque sûrement. Ainsi, on a p.s. :

$$\begin{aligned} 0 &\leq (X - \lambda_1)(\lambda_n - X) \\ &= -\lambda_1\lambda_n + (\lambda_1 + \lambda_n)X - X^2 \\ 0 &\leq -\lambda_1\lambda_n + (\lambda_1 + \lambda_n - X)X \end{aligned}$$

donc on en déduit que p.s. :

$$\frac{1}{X} \leq \frac{\lambda_1 + \lambda_n - X}{\lambda_1\lambda_n}$$

En passant à l'espérance, on obtient donc :

$$\begin{aligned} \mathbb{E}[X]\mathbb{E}\left[\frac{1}{X}\right] &\leq \mathbb{E}[X]\frac{\lambda_1 + \lambda_n - \mathbb{E}[X]}{\lambda_1\lambda_n} \\ &= -\frac{1}{\lambda_1\lambda_n}\left[\mathbb{E}[X]^2 - (\lambda_1 + \lambda_n)\mathbb{E}[X]\right] \\ &= -\frac{1}{\lambda_1\lambda_n}\left[\left((\mathbb{E}[X] - \frac{\lambda_1 + \lambda_n}{2}\right)^2 - \frac{(\lambda_1 + \lambda_n)^2}{4}\right] \\ &= -\frac{1}{\lambda_1\lambda_n}\left(\mathbb{E}[X] - \frac{\lambda_1 + \lambda_n}{2}\right)^2 + \frac{(\lambda_1 + \lambda_n)^2}{4\lambda_1\lambda_n} \\ \mathbb{E}[X]\mathbb{E}\left[\frac{1}{X}\right] &\leq \frac{(\lambda_1 + \lambda_n)^2}{4\lambda_1\lambda_n} = \frac{1}{4}\left(\sqrt{\frac{\lambda_1}{\lambda_n}} + \sqrt{\frac{\lambda_n}{\lambda_1}}\right)^2 \end{aligned}$$

et voilà !

Références

[1] ALPARGU Gülhan, *The Kantorovich inequality, with some extensions and with some statistical applications.* McGill University, Montréal. Juillet 1996.