Santiago Novoa Pérez *000125089*

Desigualdad de Kantorovich: Sea H una matriz simétrica definida positiva de dimensión nxn, para cualquier vector x se satisface:

$$\frac{(x^T x)^2}{(x^T Hx)(x^T H^{-1}x)} \geq \frac{4\lambda_{min}\lambda_{max}}{(\lambda_{min}+\lambda_{max})^2} = \text{L.D.}$$

donde λ_{min} y λ_{max} son, respectivamente, el valor propio menor y mayor de H.

Prueba

Se utilizaran los siguientes resultados:

1. $\forall H$ matriz simétrica $\exists P$ ortogonal tal que $H = P^T DP$. Sea $x = Py$, entonces $y^T Hy = y^T P^T DPy = x^T Dx = \sum_{i=1}^n \lambda_i x_i^2$ con λ_i los valores propios de H(y de D).
2. Si λ_i's son los valores propios de H no-singular (por ejemplo, S.P.D.), entonces $\frac{1}{\lambda_i}$'s son los valores propios de H^{-1}. Además, $y^T H^{-1}y = x^T D^{-1}x = \sum_{i=1}^n \frac{x_i^2}{\lambda_i}$
3. Sea f una función convexa ($\frac{d^2f}{dx^x} > 0$ si existe o $f(tx+(1-t)y) \leq tf(x)+(1-t)f(y) \quad \forall t \in [0,1]$), entonces se cumple que $f(z) \leq \sum_{i=1}^n \alpha_i f(x_i)$ donde z es la combinación convexa de x_i ($\sum_{i=1}^n \alpha_i x_i$)
4. $x^* = \sum_{i=1}^n \alpha_i x_i$ es una combinación convexa de x_i's si $\alpha_i \in [0,1]$ y $\sum_{i=1}^n \alpha_i = 1$
5. Si H es S.P.D, todos sus valores propios son mayores que cero, además, $\lambda_i = t\lambda_1 + (1-t)\lambda_n$ con algún $t \in [0,1]$, y λ_1, λ_n los valores propios más chicos y más grandes respectivamente.

Sean $y \in \mathbb{R}^n$ y $H \in \mathbb{R}^{nxn}$ tales que H es S.P.D. Por (1) y (2) se tiene que:

$$L.I. = \frac{(y^T y)^2}{(y^T Hy)(y^T H^{-1}y)} = \frac{(x^T x)^2}{(x^T Dx)(x^T D^{-1}x)} = \frac{(\sum_{i=1}^n x_i^2)^2}{(\sum_{i=1}^n x_i^2\lambda_i)(\sum_{i=1}^n \frac{x_i^2}{\lambda_i})}$$

Por comodidad y para facilitar la manipulación algebrŕica, se tomará la siguiente expresión:

$$L.I. = \frac{\sum x_i^2}{\sum \lambda_i x_i^2} / \frac{\sum \frac{x_i^2}{\lambda_i}}{\sum x_i^2}$$

Por el mismo motivo, se define $\xi_i = \frac{x_i^2}{\sum_{i=1}^n xi^2}$. Notar que por (4) ξ_i forma una combinación convexa de λ_i's para el numerador y otra de $\frac{1}{\lambda_i}$'s para el denominador.

$$L.I. = \left[\sum \frac{1}{\xi_i\lambda_i}\right] / \left[\sum \frac{\xi_i}{\lambda_i}\right]$$

Nota: Lo que se busca con esto es encontrar una expresión de L.I. tal que L.I $\leq$ L.D. Para hacer esto, se pasará por ecuaciones más simples intermedias, empezando por funciones convexas.

Sea

$$f(\lambda) : [\lambda_1, \lambda_n] \longrightarrow \mathbb{R}$$

la función

$$f(\lambda) = \frac{1}{\lambda}$$

Es fácil notar que f es convexa en $\mathbb{R}^+$. Por lo anterior y por (5), f es convexa en $[\lambda_1, \lambda_n]$. Por (3) se tiene que $f(z) \leq \sum \alpha_i f(x_i)$. En particular, para $\lambda = \sum \xi_i \lambda_i$. Entonces, $f(\lambda) \leq \sum \xi_i f(\lambda_i)$. Por (5), $f(\lambda_i) = f(t\lambda_1 + (1-t)\lambda_n) \implies f(\lambda) \leq \sum \xi_i (\frac{t}{\lambda_i} + \frac{(1-t)}{\lambda_n})$

Lo que se acaba de lograr es encontrar una expresión que va a ayudar a simplificar el L.I. Es fácil notar que, para λ_i , t toma el siguiente valor:

$$\lambda_i = t\lambda_1 + (1-t)\lambda_n$$

$$\iff t = \frac{\lambda_i - \lambda_n}{\lambda_1 - \lambda_n}$$

$$\therefore f(\lambda) \leq \sum \xi_i \left[\frac{\lambda_i - \lambda_n}{\lambda_1 - \lambda_n} (\frac{1}{\lambda_1}) - \frac{\lambda_1 - \lambda_i}{\lambda_1 - \lambda_n} (\frac{1}{\lambda_n}) \right]$$

$$= \sum \xi_i \frac{\lambda_1 + \lambda_n - \lambda_i}{\lambda_1 \lambda_n} \qquad (6)$$

De hecho, se puede ver que el numerador $(f(\lambda))$ siempre será menor o igual al denominador $(\sum \xi_i f(\lambda_i))$ por lo que para toda λ_*, su punto en la curva $(f(\lambda_*))$ será menor o igual a la combinación convenza de los puntos en la curva. $\therefore L.I. \geq \frac{1}{\lambda} * \frac{1}{\sum \xi_i f(\lambda_i)}$.

Volviendo a la expresión del L.I. se tiene la siguiente equivalencia:

$$\frac{\frac{1}{\sum \xi_i \lambda_i}}{\sum \frac{\xi_i}{\lambda_i}} \geq \frac{\frac{1}{\lambda}}{\sum \xi_i f(\lambda_i)} = \frac{1}{\lambda} \frac{1}{\sum \xi_i f(\lambda_i)} \ \ldots \ g(\lambda)$$

Ya llegado a este punto, se intentará minimizar la expresión anterior (derivable) para terminar la demostración.

$$\frac{d}{d\lambda}g = 0 \iff \frac{d}{d\lambda^*}\lambda^*(\lambda_1 + \lambda_n - \lambda^*)^{-1} = 0$$

$$\iff \lambda_1 + \lambda_n - 2\lambda^* = 0$$

$$\iff \lambda^* = \frac{\lambda_1 + \lambda_n}{2}$$

Revisando la segunda derivada(porque $\frac{d}{d\lambda}$ es derivable), se encuentra que:

$$\frac{d^2}{d\lambda}g(\lambda^*) = 2\lambda_1\lambda_n \left[\frac{\lambda_1 + \lambda_n}{2}(\frac{\lambda_1 + \lambda_n}{2})\right]^{-2} > 0$$

$\therefore \lambda^*$ minimiza la función g
Es decir,

$$\frac{(y^Ty)^2}{(y^THy)(y^TH^{-1}y)} = \frac{(\sum_{i=1}^n x_i^2)^2}{(\sum_{i=1}^n x_i^2\lambda_i)(\sum_{i=1}^n \frac{x_i^2}{\lambda_i})}$$

$$\geq \frac{1}{\lambda}\left(\frac{1}{\sum \frac{\xi_i}{\lambda_i}}\right)$$

$$\geq \min \frac{1}{\lambda(\frac{\lambda_1 + \lambda_n - \lambda}{\lambda_1\lambda_n})}$$

$$= \frac{\lambda_1\lambda_n}{(\frac{\lambda_1+\lambda_n}{2})(\lambda_1 + \lambda_n - (\frac{\lambda_1+\lambda_n}{2}))}$$

$$= \frac{4\lambda_1\lambda_n}{\lambda_1^2 + 2\lambda_1\lambda_n + \lambda_2^2}$$

$$= \frac{4\lambda_1\lambda_n}{(\lambda_1 + \lambda_n)^2} \qquad \square$$