Cociente de Rayleigh

Sea A una matriz cuadrada simétrica $d \times d$. El **cociente de Rayleigh** de un vector x se define como el número:

$$R(x) = \frac{x^T A x}{x^T x}, \qquad x \neq 0. \tag{1}$$

Supongamos que los autovalores de A satisfacen que

$$\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_d \tag{2}$$

y $v_1, ..., v_d$ sus respectivos autovectores ortogonales dos a dos y todos normalizados tal que $\|v_i\|_2 = 1$, norma euclídea. Entonces, para $x \neq 0$, existen $a_i \in \mathbb{R}$ con $i = 1, \ldots, d$ tal que:

$$x = a_1 v_1 + \cdots + a_d v_d \quad \text{y} \quad \|x\|_2^2 = a_1^2 + \cdots + a_d^2. \tag{3}$$

Así,

$$Ax = a_1 \lambda_1 v_1 + \cdots + a_d \lambda_d v_d$$

y

$$x^t A x = a_1^2 \lambda_1 + \cdots + a_d^2 \lambda_d. \tag{4}$$

De esta ecuación tengo:

$$x^t A x = \lambda_1 (a_1^2 + \cdots + a_d^2) + (\lambda_2 - \lambda_1) a_2^2 + \cdots + (\lambda_d - \lambda_1) a_d^2$$

Por (3) tengo que $\lambda_i - \lambda_1 \leq 0$ para $i = 2, \ldots, d$, entonces

$$x^t A x \leq \lambda_1 (a_1^2 + \cdots + a_d^2) \Rightarrow R(x) \leq \lambda_1 \tag{5}$$

Por otro lado de (4)

$$x^t A x = \lambda_d (a_1^2 + \cdots + a_d^2) + (\lambda_1 - \lambda_d) a_1^2 + \cdots + (\lambda_{d-1} - \lambda_d) a_d^2$$

Por (3) tengo que $\lambda_i - \lambda_d \geq 0$ para $i = 1, \ldots, d-1$, entonces

$$x^t A x \geq \lambda_d (a_1^2 + \cdots + a_d^2) \Rightarrow R(x) \geq \lambda_d.$$

Recuperando el resultado (5), se concluye que si se tiene (2) entonces

$$\lambda_1 \geq R(x) \geq \lambda_d \tag{6}$$

Observación: $R(x)$ puede ser negativo si alguno de los autovalores lo es.

Supongamos ahora que autovalores $\lambda_1, \cdots, \lambda_d$ satisfacen

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_d|, \tag{7}$$

donde $v_1, ..., v_d$ son sus autovectores normalizados en la norma euclidea. Supongamos que parte de un iterante inicial $x_0 = x$ que tiene componente no nula en la dirección del autovector dominante

v_1, es decir, $a_1 \neq 0$ en (3). Por recurrencia defino $x_n = Ax_{n-1}$, $n \geq 1$, es decir, $x_n = A^n x_0$. Por tanto, su cociente de Raileigh vale

$$R(x_n) = \frac{(x_n)^T A x_n}{(x_n)^T x_n} = \frac{(x_0)^T A^{2n+1} x_0}{(x_0)^T A^{2n} x_0} = \frac{a_1^2 \lambda_1^{2n+1} + \cdots + a_d^2 \lambda_d^{2n+1}}{a_1^2 \lambda_1^{2n} + \cdots + a_d^2 \lambda_d^{2n}}. \tag{8}$$

Así,

$$R(x_n) - \lambda_1 = \frac{a_2^2 \lambda_2^{2n}(\lambda_2 - \lambda_1) + \cdots + a_d^2 \lambda_d^{2n}(\lambda_d - \lambda_1)}{a_1^2 \lambda_1^{2n} + \cdots + a_d^2 \lambda_d^{2n}}. \tag{9}$$

Llamo

$$K = \max_{j=2,\ldots,d} \{|\lambda_j - \lambda_1 1\},$$

entonces de (9) se tiene

$$\begin{aligned}
|R(x_n) - \lambda_1| &\leq K \frac{a_2^2 \lambda_2^{2n} + \cdots + a_d^2 \lambda_d^{2n}}{a_1^2 \lambda_1^{2n} + \cdots + a_d^2 \lambda_d^{2n}} \\
&\leq K \frac{\lambda_2^{2n}}{\lambda_1^{2n}} \frac{a_2^2 + \cdots + a_d^2 \lambda_d^{2n} \lambda_2^{-2n}}{a_1^2 + \cdots + a_d^2 \lambda_d^{2n} \lambda_1^{-2n}} \\
&\leq K \frac{\lambda_2^{2n}}{\lambda_1^{2n}} \frac{a_2^2 + \cdots + a_d^2}{a_1^2} \\
&\leq K(x_0, A) \left(\frac{\lambda_2}{\lambda_1} \right)^{2n},
\end{aligned}$$

donde $K(x_0, A) > 0$ y solo depende de la matriz A y del iterante inicial x_0, nunca de n. Observar que utilizando el cociente de Rayleigh es un buen método para aproximarse al autovalor dominante: **Suponiendo (7) y dado un iterante inicial x_0 que tiene componente no nula en la dirección del autovector dominante v_1 entonces para**

$$x_n = Ax_{n-1} \ \textbf{se tiene que} \ R(x_n) \to \lambda_1 \tag{10}$$

Además, por (9) se tiene

$$\begin{aligned}
\frac{R(x_{n+1}) - \lambda_1}{R(x_n) - \lambda_1} &= \frac{a_2^2 \lambda_2^{2n+2}(\lambda_2 - \lambda_1) + \cdots + a_d^2 \lambda_d^{2n+2}(\lambda_d - \lambda_1)}{a_2^2 \lambda_2^{2n}(\lambda_2 - \lambda_1) + \cdots + a_d^2 \lambda_d^{2n}(\lambda_d - \lambda_1)} \frac{a_1^2 \lambda_1^{2n} + \cdots + a_d^2 \lambda_d^{2n}}{a_1^2 \lambda_1^{2n+2} + \cdots + a_d^2 \lambda_d^{2n+2}} \\
&= \frac{\lambda_2^2}{\lambda_1^2} \frac{a_2^2 \lambda_2^{2n}(\lambda_1 - \lambda_2) + \cdots + a_d^2 \lambda_d^{2n+2} \lambda_2^{-2}(\lambda_1 - \lambda_d)}{a_2^2 \lambda_2^{2n}(\lambda_1 - \lambda_2) + \cdots + a_d^2 \lambda_d^{2n}(\lambda_1 - \lambda_d)} \frac{a_1^2 \lambda_1^{2n} + \cdots + a_d^2 \lambda_d^{2n}}{a_1^2 \lambda_1^{2n} + \cdots + a_d^2 \lambda_d^{2n+2} \lambda_1^{-2}}
\end{aligned}$$

Utilizando (7), resulta

$$\begin{aligned}
\left| \frac{R(x_{n+1}) - \lambda_1}{R(x_n) - \lambda_1} \right| &\leq \frac{\lambda_2^2}{\lambda_1^2} \frac{a_2^2 \lambda_2^{2n}(\lambda_1 - \lambda_2) + \cdots + a_d^2 \lambda_d^{2n}(\lambda_1 - \lambda_d)}{a_2^2 \lambda_2^{2n}(\lambda_1 - \lambda_2) + \cdots + a_d^2 \lambda_d^{2n}(\lambda_1 - \lambda_d)} \frac{a_1^2 \lambda_1^{2n} + \cdots + a_d^2 \lambda_d^{2n}}{a_1^2 \lambda_1^{2n}} \\
&\leq \frac{\lambda_2^2}{\lambda_1^2} \frac{a_1^2 + a_2^2 (\lambda - 2/\lambda_1)^{2n} + \cdots + a_d^2 (\lambda_d/\lambda_1)^{2n}}{a_1^2} \\
&\leq \left(\frac{\lambda_2}{\lambda_1} \right)^2 \frac{a_1^2 + \cdots + a_d^2}{a_1^2}.
\end{aligned}$$

Por lo tanto, tengo que la velocidad de convergencia al autovalor dominante del método (14) es de orden dos, $\mathcal{O}(\lambda_2/\lambda_1)^2$, es decir, si llamo $e_n = |R(x_n) - \lambda_1|$ entonces

$$e_{n+1} \leq C \left(\frac{\lambda_2}{\lambda_1} \right)^2 e_n. \tag{11}$$

Otro método para calcular el autovalor dominante es el **método de la potencia**. Se toma un vector de arranque x_0 no cero de arranque que tiene componente no nula en la dirección del autovector dominante v_1, llamo $M(x_0)$ la primera componente de x_0 que tenga mayor módulo y defino $y_0 = x_0/M(x_0)$. Entonces, construímos una sucesión de vectores $\{y_k\}_k$ y de escalares $\{M(x_k)\}_k$ tal que para $k \geq 1$:

$$x_k = Ay_{k-1}, \qquad y_k = \frac{x_k}{M(x_k)},$$

tal que cuando $k \to \infty$

$$y_k \to v_1, \qquad M(x_k) \to \lambda_1.$$

Observar que

$$x_1 = Ay_0 = \frac{Ax_0}{M(x_0)} \Rightarrow M(x_1) = \frac{M(Ax_0)}{M(x_0)}.$$

Por lo tanto,

$$y_1 = \frac{x_1}{M(x_1)} = \frac{M(x_0)}{M(Ax_0)} \frac{Ax_0}{M(x_0)} = \frac{Ax_0}{M(Ax_0)}$$

Ahora demostremos por inducción que

$$M(x_n) = \frac{M(A^n x_0)}{M(A^{n-1}x_0)}, \qquad y_n = \frac{A^n x_0}{M(A^n x_0)}. \tag{12}$$

Suponiendolo cierto hasta $n-1$, entonces

$$x_n = Ay_{n-1} = A\frac{A^{n-1}x_0}{M(A^{n-1}x_0)} = \frac{A^n x_0}{M(A^{n-1}x_0)}.$$

Por lo tanto, efectivamente

$$M(x_n) = \frac{M(A^n x_0)}{M(A^{n-1}x_0)},$$

y entonces

$$y_n = \frac{x_n}{M(x_n)}) = \frac{M(A^{n-1}x_0)}{M(A^n x_0)} \frac{A^n x_0}{M(A^{n-1}x_0)} = \frac{A^n x_0}{M(A^n x_0)},$$

como queríamos demostrar.

Notemos que si calculamos el coeficiente de Rayleigh de y_n, resulta por (12)

$$R(y_n) = \frac{y_n^T A y_n}{y_n^T y_n} = \frac{(A^n x_0)^T A (A^n x_0)}{(A^n x_0)^T (A^n x_0)} = \frac{x_0^T A^{2n+1} x_0}{x_0^T A^{2n} x_0}.$$

Entonces por (8), (14) y (11), $R(y_n) \to R(v_1) = \lambda_1$ con convergencia de orden 2, mejor convergencia que $M(x_n) \to \lambda_1$, que solamente es de orden 1.

¿Como podemos converger con un coeficiente de Rayleigh a un autovalor simple pero que no es en general dominante? Supongamos que los autovalores de A satisfacen que

$$\lambda_1 > \lambda_2 \geq cdots \geq \lambda_d. \tag{13}$$

donde λ_1 no tiene porque satisfacer (7). Si consideramos $\widetilde{A} = A + kI$, donde I es la matriz identidad, sus autovalores son $\lambda_i + K$. Tomando k suficiente grande tal que $\lambda_d + k \geq 0$ y entonces estamos cumpliendo la hipótesis (7) para los autovalores de $\widetilde{A}$. Por tanto, tomando x_0 que tenga componente no nula en la dirección del autovector v_1,

$$x_n = \widetilde{A}x_{n-1} \text{ se tiene que } R(x_n) = \frac{(x_0)^T \widetilde{A}^{2n+1} x_0}{(x_0)^T \widetilde{A}^{2n} x_0} \to \lambda_1 + k. \tag{14}$$

Un K que nos asegura este resultado es:

$$k = \max_{i=1,\dots,d} \{|a_{i1}| + \cdots + |a_{id}|\},$$

donde a_{ij} son los coeficientes de la matriz A. Con este k la matriz $\widetilde{A}$ es semidefinida positiva.