

Courant-Fischer

Ce développement est assez coton à faire rentrer en 15 minutes, voyez bien ce que vous escamotez ou pas. Le but est d'au moins d'obtenir que les valeurs propres des matrices symétriques sont des fonctions 1-lipschitziennes.

I. Contexte : une autre preuve du Théorème spectral (pas dans le dev)

On pose \mathbb{S} la sphère unité de \mathbb{R}^n pour la norme euclidienne.

Lemme I.0.1.

Toute matrice symétrique admet au moins une valeur propre réelle qui s'écrit :

$$\max_{X \in \mathbb{S}} \langle X, MX \rangle .$$

Preuve. Soit M symétrique. Alors l'application $\varphi : \begin{cases} \mathbb{S} & \rightarrow \mathbb{R} \\ X & \mapsto \langle X, MX \rangle \end{cases}$ est continue car bilinéaire en dimension finie, et sa source est compacte car \mathbb{S} est fermé et borné en dimension finie. Elle est à valeurs dans \mathbb{R} donc admet un maximum, atteint en $V \in \mathbb{S}$.

\mathbb{S} est une sous-variété différentiable de \mathbb{R}^n , et l'espace tangent de V dans \mathbb{S} est $(\mathbb{R}V)^\perp$.¹ φ est en fait différentiable en V . En effet, si $H \in (\mathbb{R}V)^\perp$, alors $\varphi(V+H) - \varphi(V) = \langle H, MV \rangle + \langle V, MH \rangle + O(\|H\|)$ donc $d\varphi_V(H) = \langle H, MV \rangle + \langle V, MH \rangle = 2 \langle H, MV \rangle$ car M est symétrique. Comme φ est maximale en V , sa différentielle est nulle en ce point. Donc $\forall H \in (\mathbb{R}V)^\perp, \langle H, MV \rangle = 0$. Mais alors $MV \in ((\mathbb{R}V)^\perp)^\perp = \mathbb{R}V$. Donc V est un vecteur propre de M pour une certaine valeur propre réelle. \square

Remarque I.0.2. On va voir que c'est la plus grande valeur propre de M . Par le même raisonnement, $\min_{X \in \mathbb{S}} \langle X, MX \rangle$ est aussi une valeur propre de M , qui est la plus petite.

Théorème I.0.3. Théorème Spectral

Toute matrice symétrique réelle est orthodiagonalisable dans \mathbb{R} .

Preuve. Si le théorème est vrai en dimension $n - 1$, alors l'on prend une matrice $M \in \mathcal{S}_n(\mathbb{R})$, qui admet un vecteur propre $v \in \mathbb{R}^n$, et l'on pose $V = (\mathbb{R}v)^\perp$. Comme M est symétrique et stabilise $\mathbb{R}v$, elle stabilise V . Il y a donc une matrice $M' \in \mathcal{S}_{n-1}(\mathbb{R})$ telle que M est orthogonalement semblable à

$$\begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & M' & \\ 0 & & & \end{pmatrix}$$

Il suffit donc d'orthodiagonaliser M' pour orthodiagonaliser M , et cela est possible par hypothèse de récurrence. \square

II. Courant-Fischer, inégalités de Weyl et conséquences

On munit $E := \mathbb{R}^n$ du produit scalaire $\langle X, Y \rangle = \sum_{k=1}^n x_k y_k$. On note $\mathbb{S} = \{x \in \mathbb{R}^n \mid \langle x, x \rangle = 1\}$. On note $q_M(X) = \langle X, MX \rangle$. Le max de q_M sur $\mathbb{S} \cap E$ s'écrit de façon succincte $\max_{\mathbb{S} \cap E} q_M$. On note \mathcal{E}_k l'ensemble des sev de dimension k .

1. Sans la théorie des sous-variétés, on peut remplacer la notion d'espace tangent par le fait de simplement considérer des arcs tracés sur la sphère. Ils sont faciles à construire : simplement prendre une fonction affine, et normaliser pour que ça soit tracé sur \mathbb{S} .

Théorème II.0.1. Principe variationnel de Courant-Fischer

Soit $M \in \mathcal{S}_n(\mathbb{R})$. Alors si $\lambda_1, \dots, \lambda_n$ sont ses valeurs propres énumérées dans l'ordre croissant et avec multiplicités, alors :

$$\lambda_k = \min_{V \in E_k} \max_{\mathbb{S} \cap V} q_M = \max_{V \in E_{n-k+1}} \min_{\mathbb{S} \cap V} q_M.$$

Preuve. On pose e_1, \dots, e_n une base orthonormée correspondant aux $\lambda_1, \dots, \lambda_n$. Comme e_1, \dots, e_n est une base orthonormée, pour $x = \sum_{i=1}^n x_i e_i \in \mathbb{S}$, on a la très importante formule suivante :

$$q_M(x) = \left\langle \sum_{i=1}^n x_i e_i, \sum_{i=1}^n \lambda_i x_i e_i \right\rangle = \sum_{i=1}^n x_i^2 \lambda_i.$$

Soit $k \in \llbracket 1, n \rrbracket$. On pose $V_k = \text{Vect}(e_1, \dots, e_k) \in E_k$. Si $x \in V_k \cap \mathbb{S}$, on a $\sum_{i=1}^k x_i^2 = 1$ d'où l'on tire :

$$q_M(x) = \sum_{i=1}^k x_i^2 \lambda_i \leq \sum_{i=1}^k x_i^2 \lambda_k = \lambda_k. \text{ En outre, } q_M(e_k) = \lambda_k. \text{ Donc } \max_{V_k \cap \mathbb{S}} q_M = \lambda_k.$$

Le raisonnement précédent a montré que $\lambda_k \geq \inf_{V \in E_k} \max_{\mathbb{S} \cap V} q_M$. Soit $V \in E_k$. Pour montrer que $\max_{\mathbb{S} \cap V} q_M \geq \lambda_k$, on remarque d'abord que $V \cap \text{Vect}(e_k, \dots, e_n)$ est non nul. En effet, la formule de Grassmann donne que $\dim(V \cap \text{Vect}(e_k, \dots, e_n)) = \dim(V) + n - k + 1 - \dim(V + \text{Vect}(e_k, \dots, e_n)) \geq k + n - k + 1 - n = 1$. En prenant alors $x = \sum_{i=k}^n x_i e_i \in V \cap \text{Vect}(e_k, \dots, e_n) \in \mathbb{S}$, on trouve que

$$q_M(x) = \sum_{i=k}^n \lambda_i |x_i|^2 \geq \sum_{i=k}^n |x_i|^2 \lambda_k = \lambda_k. \text{ Ainsi, } \max_{\mathbb{S} \cap V} q_M \geq \lambda_k. \text{ C'est vrai pour tout } V \text{ donc on a prouvé}$$

que $\min_{V \in E_k} \max_{V \cap \mathbb{S}} q_M = \lambda_k$ (et que ce min existe).

La deuxième identité du théorème se prouve en appliquant la première à $-M$, et en voyant que pour $i \in \llbracket 1, n \rrbracket$, $\lambda_i(-M) = -\lambda_{n-i+1}(M)$. **Attention, vous n'avez probablement pas le temps d'écrire cet argument. À mentionner à l'oral, c'est tout.** \square

Le théorème de Courant-Fischer donne au calcul des valeurs propres un aspect linéaire qui donne des informations sur les valeurs propres de $M + N$ en fonction de celles de M et de N . Il s'agit cependant de choisir de bons indices afin de permettre l'utilisation de la formule de Grassmann, qui est l'outil privilégié pour manipuler ensemble des sous-espaces de E lorsqu'on ne connaît que leur dimension.

Corollaire II.0.2. Inégalités de Weyl

Soit M, N deux éléments de $\mathcal{S}_n(\mathbb{R})$. Alors, avec des notations évidentes :

$$\text{Si } i + j = n + k, \text{ alors } \lambda_i(M) + \lambda_j(N) \geq \lambda_k(M + N)$$

$$\text{Si } i + j = 1 + k, \text{ alors } \lambda_i(M) + \lambda_j(N) \leq \lambda_k(M + N).$$

Preuve. Supposons $i + j = n + k$. On veut montrer que $\lambda_i(M) + \lambda_j(N) \geq \max_{V \in E_{n-k+1}} \min_{\mathbb{S} \cap V} q_{M+N}$. Soit $V \in E_{n-k+1}$ quelconque. On veut donc montrer que $\lambda_i(M) + \lambda_j(N) \geq \min_{\mathbb{S} \cap V} q_{M+N}$. On pose également $V_i \in E_i$ tel que $\lambda_i(M) = \max_{\mathbb{S} \cap V_i} q_M$ et $V_j \in E_j$ tel que $\lambda_j(N) = \max_{\mathbb{S} \cap V_j} q_N$. Il nous suffit donc de trouver un élément $v \in V_i \cap V_j \cap V \cap \mathbb{S}$. En effet, on en tirerait :

$$\begin{aligned} \min_{\mathbb{S} \cap V} q_{M+N} &\leq \langle v, (M + N)v \rangle = \langle v, Mv \rangle + \langle v, Nv \rangle \\ &\leq \max_{V_i \cap \mathbb{S}} q_M + \max_{V_j \cap \mathbb{S}} q_N \\ &= \lambda_i(M) + \lambda_j(N). \end{aligned}$$

L'existence de v se fait encore en appliquant successivement la formule de Grassmann :

$$\begin{aligned} \dim(V_i \cap V_j \cap V) &= \dim(V_i \cap V_j) + \dim(V) - \dim(V_i \cap V_j + V) \\ &= \dim(V_i) + \dim(V_j) + \dim(V) - \dim(V_i + V_j) - \dim(V_i \cap V_j + V) \\ &\geq i + j + (n - k + 1) - n - n = 1. \end{aligned}$$

Si maintenant $i + j = k + 1$, alors $\lambda_i(M) + \lambda_j(N) \leq \lambda_k(M + N)$ équivaut à $-\lambda_{n-i+1}(-M) - \lambda_{n-j+1}(-N) \geq -\lambda_{n-k+1}(-M - N)$ soit $\lambda_{n-k+1}(-M - N) \geq \lambda_{n-i+1}(-M) + \lambda_{n-j+1}(-N)$. Mais ce résultat vient tout juste d'être établi car $(n - k + 1) + n = (n - i + 1) + (n - j + 1)$. **Vous n'aurez pas le temps de faire ce paragraphe, dites simplement à l'oral qu'on applique le point précédent à $-N$ et $-M$.**
 \square

Corollaire II.0.3.

Si $k \in \llbracket 1, n \rrbracket$, alors l'application qui à $M \in \mathcal{S}_n(\mathbb{R})$ associe sa k -ième valeur propre (en comptant les multiplicités) est 1-lipschitzienne. λ_1 est concave et λ_n est convexe.

Preuve. Si k est fixé, alors on veut montrer que pour $M, N \in \mathcal{S}_n(\mathbb{R})$, on a $|\lambda_k(M) - \lambda_k(N)| \leq \|M - N\|_{op}$. Mais on sait que la norme d'opérateur d'une matrice hermitienne $M - N$ vaut :

$$\max (|\lambda_i(M - N)|, i \in \{1, \dots, n\}) = \max (\lambda_n(M - N), -\lambda_1(M - N)).$$

L'intervalle $[\lambda_1(M - N), \lambda_n(M - N)]$ est contenu dans l'intervalle $[-\|M\|, \|M\|]$.² Donc il est suffisant de prouver ces deux inégalités :³

$$\lambda_k(M) - \lambda_k(N) \leq \lambda_n(M - N) \tag{**}$$

$$\lambda_1(M - N) \leq \lambda_k(M) - \lambda_k(N). \tag{*}$$

Ces dernières correspondent exactement aux deux inégalités de Weyl. (*) se montre avec la première inégalité, et (**) avec la seconde, en prenant $i = k$ et $j = 1$ ou n .

Maintenant, la convexité de λ_n se prouve avec (**) en voyant qu'elle implique que $\lambda_n(M) \leq \lambda_n(N) + \lambda_n(M - N)$ soit $\lambda_n(M + N) \leq \lambda_n(M) + \lambda_n(N)$. Donc, si $t \in [0, 1]$, alors $\lambda_n((1 - t)M + tN) \leq \lambda_n((1 - t)M) + \lambda_n(tN) = (1 - t)\lambda_n(M) + t\lambda_n(N)$ par homogénéité. Et puisque $\lambda_1(M) = -\lambda_n(-M)$, on a que $\lambda_1(M)$ est concave⁴.
 \square

2. Ça se voit en une ligne : les bornes de $[\lambda_1(M - N), \lambda_n(M - N)]$ sont de valeur absolue plus petite que $\|M - N\|$.
 3. Qui prouvent ainsi que $\lambda_k(M) - \lambda_k(N)$ est de valeur absolue $\leq \|M - N\|$.
 4. On pouvait aussi réutiliser (*).