Université Claude Bernard - Lyon 1 Semestre de printemps 2019-2020

Analyse matricielle-L3

Feuille d'exercices n° 3 - corrigé des exercices 2 à 4

MOINDRES CARRÉS, RECHERCHE DE VALEURS PROPRES

Corrigé de l'exercice 2

Dans cet exercice, on généralise ce qu'on a vu à l'exercice précédent. À l'exercice 1, on a $n = 3$ et $d = 1$.

1. On suppose ici qu'on a échantillonné, c'est-à-dire sélectionné, n points sur le graphe d'une fonction $f : \mathbb{R} \longrightarrow \mathbb{R}$. On connaît donc n points $(t_i, y_i) = (t_i, f(t_i))$. On va chercher une approximation du graphe de f par le graphe d'un polynôme P de degré au plus d qui passe "le plus près possible" des n points (t_i, y_i). Par l'expression "le plus près possible" on entend ici qu'on veut minimiser la somme des erreurs quadratiques $\sum_{i=1}^{n} \|P(t_i) - y_i\|_2^2$. Cette somme n'est autre que la norme au carré du vecteur $\begin{pmatrix} P(t_1) - y_1 \\ \vdots \\ P(t_n) - y_n \end{pmatrix}$ et le problème qu'on veut résoudre est donc bien de trouver le polynôme P de degré au plus d tel que

$$\left\| \begin{pmatrix} P(t_1) - y_1 \\ \vdots \\ P(t_n) - y_n \end{pmatrix} \right\| = \min_{Q \in \mathbb{R}_d[X]} \left\| \begin{pmatrix} Q(t_1) - y_1 \\ \vdots \\ Q(t_n) - y_n \end{pmatrix} \right\|$$

La forme générale d'un polynôme P de degré au plus d est $P(X) = \sum_{k=0}^{d} \alpha_k X^k$ et, étant donné $i \in [\![1, n]\!]$, on peut voir la valeur du polynôme au point t_i comme un produit matriciel :

$$P(t_i) = \sum_{k=0}^{d} \alpha_k t_i^k = \begin{pmatrix} 1 & t_i & t_i^2 & \cdots & t_i^d \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_d \end{pmatrix}$$

On définit alors la matrice

$$A = \begin{pmatrix} 1 & t_1 & t_1^2 & \cdots & t_1^d \\ 1 & t_2 & t_2^2 & \cdots & t_2^d \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & t_n & t_n^2 & \cdots & t_n^d \end{pmatrix}$$

et on observe que trouver un polynôme $P(X) = \sum_{k=0}^{d} \alpha_k X^k$ qui minimise la quantité $\sum_{i=1}^{n} \|P(t_i) - y_i\|_2^2$ équivaut à trouver le vecteur $(\alpha_0, \ldots, \alpha_d)$ qui minimise la norme du vecteur

$$A \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_d \end{pmatrix} - \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

c'est-à-dire qu'on cherche la solution $\alpha = (\alpha_0, \dots, \alpha_d)$ du problème

$$\left\| \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} - A \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_d \end{pmatrix} \right\| = \min_{z \in \mathbb{R}^{d+1}} \left\| \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} - A \begin{pmatrix} z_0 \\ z_1 \\ \vdots \\ z_d \end{pmatrix} \right\|$$

D'après les résultats du cours, la solution satisfait l'équation normale :

$$A^T A \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_d \end{pmatrix} = A^T \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

2. Pour $d = 0$ on a $A = (1 \cdots 1)^T$, $A^T A = n$ et $A^T \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \sum_{i=1}^n y_i$ donc l'unique solution du problème est $\alpha_0 = \frac{1}{n} \sum_{i=1}^n y_i$, c'est-à-dire la moyenne des y_i.

   Pour $d = 1$ on a

$$A = \begin{pmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_n \end{pmatrix}, \quad A^T A = \begin{pmatrix} n & \sum_{i=1}^n t_i \\ \sum_{i=1}^n t_i & \sum_{i=1}^n t_i^2 \end{pmatrix}, \quad A^T \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n t_i y_i \end{pmatrix}$$

   On est ramené au système à deux équations et deux inconnues

$$\begin{pmatrix} n & \sum_{i=1}^n t_i \\ \sum_{i=1}^n t_i & \sum_{i=1}^n t_i^2 \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n t_i y_i \end{pmatrix}$$

   Ce sytème admet une unique solution si, et seulement si, $n \sum_{i=1}^n t_i^2 \neq (\sum_{i=1}^n t_i)^2$. Cette solution s'écrit :

$$\alpha_0 = \frac{\sum_{i=1}^n t_i \sum_{i=1}^n t_i y_i - \sum_{i=1}^n t_i^2 \sum_{i=1}^n y_i}{(\sum_{i=1}^n t_i)^2 - n \sum_{i=1}^n t_i^2}, \qquad \alpha_1 = \frac{\sum_{i=1}^n t_i \sum_{i=1}^n y_i - n \sum_{i=1}^n t_i y_i}{(\sum_{i=1}^n t_i)^2 - n \sum_{i=1}^n t_i^2}$$

   (Remarque : on pourra faire le lien avec la remarque de la fin de l'exercice 1).

3. **<u>Méthode 1</u>** : On a

$$A^T A = \begin{pmatrix} n & \sum_{i=1}^{n} t_i & \sum_{i=1}^{n} t_i^2 & \cdots & \sum_{i=1}^{n} t_i^d \\ \sum_{i=1}^{n} t_i & \sum_{i=1}^{n} t_i^2 & \sum_{i=1}^{n} t_i^3 & \cdots & \sum_{i=1}^{n} t_i^{d+1} \\ \sum_{i=1}^{n} t_i^2 & \sum_{i=1}^{n} t_i^3 & \sum_{i=1}^{n} t_i^4 & \cdots & \sum_{i=1}^{n} t_i^{d+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} t_i^{d-2} & \sum_{i=1}^{n} t_i^{d-1} & \sum_{i=1}^{n} t_i^{d} & \cdots & \sum_{i=1}^{n} t_i^{2d-2} \\ \sum_{i=1}^{n} t_i^{d-1} & \sum_{i=1}^{n} t_i^{d} & \sum_{i=1}^{n} t_i^{d+1} & \cdots & \sum_{i=1}^{n} t_i^{2d-1} \\ \sum_{i=1}^{n} t_i^{d} & \sum_{i=1}^{n} t_i^{d+1} & \sum_{i=1}^{n} t_i^{d+2} & \cdots & \sum_{i=1}^{n} t_i^{2d} \end{pmatrix}$$

Cherchons les conditions sous lesquelles $A^T A$ est inversible. Notons $C_0, \dots, C_d$ les colonnes de cette matrice et supposons qu'il existe $\alpha_0, \dots, \alpha_d \in \mathbb{R}$ tels que

$$\sum_{k=0}^{d} \alpha_k C_k = 0$$

On définit le polynôme $P(x) = \sum_{k=0}^{d} \alpha_k x^k$ et on observe que la première ligne de l'égalité matricielle $\sum_{k=0}^{d} \alpha_k C_k = 0$ équivaut à $\sum_{i=0}^{n} P(t_i) = 0$, la seconde ligne équivaut à $\sum_{i=0}^{n} t_i P(t_i) = 0$, la troisième à $\sum_{k=i}^{n} t_i^2 P(t_i) = 0$, etc., jusqu'à la dernière ligne qui équivaut à $\sum_{i=0}^{n} t_i^d P(t_i) = 0$. On prend alors la combinaison linéaire de ces $(d+1)$-égalités pondérées par les coefficients $\alpha_0, \cdots, \alpha_d$ et on obtient $\sum_{k=0}^{d} \alpha_k \sum_{i=0}^{n} t_i^k P(t_i) = 0$ qui se récrit comme $\sum_{i=0}^{n} P(t_i) \sum_{k=0}^{d} \alpha_k t_i^k = \sum_{i=0}^{n} [P(t_i)]^2 = 0$. On en déduit que

$$P(t_i) = 0, \qquad \forall i \in [\![0, n]\!].$$

P est un polynôme de degré au plus d s'annulant en $t_1, \dots, t_n$, qui sont distincts par hypothèse. Une condition suffisante pour que P soit le polynôme nul est que $n > d$, et on peut conclure sous cette hypothèse que tous les α_k sont nuls donc $A^T A$ est inversible.

Supposons à présent que $n \leq d$ et considérons le polynôme $P(x) = \prod_{i=0}^{n} (x - t_i)$ qui est donc de degré au plus d. On note $\beta_0, \dots, \beta_d$ les coefficients de sa décomposition dans la base canonique de $\mathbb{R}_d[X]$, i.e., $P(x) = \sum_{k=0}^{d} \beta_k x^k$. Il est clair que ces coefficients ne sont pas tous nuls. En reprenant les notations utilisées ci-dessus on observe que $\sum_{k=0}^{d} \beta_k C_k = 0$ puisque $P(t_i) = 0$ pour tout $i \in [\![0, n]\!]$. Ceci prouve que $A^T A$ n'est pas inversible.

On peut donc conclure que $A^T A$ est inversible si, et seulement si, $n > d$.

Méthode 2 : On peut aussi raisonner en utilisant la propriété vue en cours : $A^T A$ est inversible si, et seulement, si $\ker A = \{0\}$.
Supposons d'abord que $d \geq n$. Par le théorème du rang, on obtient que $\dim(\ker A) = \dim \mathbb{R}^{d+1} - \operatorname{rg} A \geq d + 1 - n \geq 1$ donc $A^T A$ n'est pas inversible.
Supposons ensuite que $d < n$, on distingue alors deux cas :

1. Si $d = n - 1$, A est une matrice carrée, de Vandermonde, on a $\det A = \prod_{i<j}(t_j - t_i) \neq 0$ donc A est inversible et $A^T A$ est inversible.
2. Si $d < n - 1$, A contient une sous-matrice de taille $d + 1$ inversible donc son rang est $d + 1$ et son noyau est réduit à $\{0\}$. Ainsi $A^T A$ est inversible.

On peut conclure que $A^T A$ est inversible si, et seulement si, $n > d$.

Corrigé de l'exercice 3

<u>Méthode 1</u> : factorisation QR de A par la méthode de Householder.

On n'utilise pas l'équation normale pour résoudre le problème aux moindres carrés. On cherche la factorisation QR de A avec $Q \in \mathcal{M}_{4,4}(\mathbb{R})$ orthogonale et $R \in \mathcal{M}_{4,2}(\mathbb{R})$ triangulaire supérieure, en utilisant la méthode de Householder. On verra à la fin comment on utilise cette factorisation pour conclure.

Étape 1. On pose $v_1 = \begin{pmatrix} 2 \\ 0 \\ 2 \\ 1 \end{pmatrix}$, on a $\|v_1\|_2 = 3$. On pose $w_1 = v_1 - 3e_1 = \begin{pmatrix} -1 \\ 0 \\ 2 \\ 1 \end{pmatrix}$. On a $\|w_1\|_2^2 = 6$. On pose $H^{(1)} = I_4 - \dfrac{2}{\|w_1\|_2^2} w_1 w_1^T = \begin{pmatrix} \frac{2}{3} & 0 & \frac{2}{3} & \frac{1}{3} \\ 0 & 1 & 0 & 0 \\ \frac{2}{3} & 0 & -\frac{1}{3} & -\frac{2}{3} \\ \frac{1}{3} & 0 & -\frac{2}{3} & \frac{2}{3} \end{pmatrix}$ et $A^{(2)} = H^{(1)}A = \begin{pmatrix} 3 & \frac{14}{3} \\ 0 & 0 \\ 0 & \frac{2}{3} \\ 0 & -\frac{5}{3} \end{pmatrix}$.

Étape 2. On pose $\widetilde{v}_2 = \begin{pmatrix} 0 \\ \frac{2}{3} \\ -\frac{5}{3} \end{pmatrix}$, on a $\|\widetilde{v}_2\|_2 = \dfrac{\sqrt{29}}{3}$. On pose $\widetilde{w}_2 = \widetilde{v}_2 - \dfrac{\sqrt{29}}{3} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -\frac{\sqrt{29}}{3} \\ \frac{2}{3} \\ -\frac{5}{3} \end{pmatrix}$. On a $\|\widetilde{w}_2\|_2^2 = \dfrac{58}{9}$. On pose $\widetilde{H}^{(2)} = I_3 - \dfrac{2}{\|\widetilde{w}_2\|_2^2} \widetilde{w}_2 \widetilde{w}_2^T = \begin{pmatrix} 0 & \frac{2}{\sqrt{29}} & -\frac{5}{\sqrt{29}} \\ \frac{2}{\sqrt{29}} & \frac{25}{29} & \frac{10}{29} \\ -\frac{5}{\sqrt{29}} & \frac{10}{29} & \frac{4}{29} \end{pmatrix}$ puis

$H^{(2)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & \frac{2}{\sqrt{29}} & -\frac{5}{\sqrt{29}} \\ 0 & \frac{2}{\sqrt{29}} & \frac{25}{29} & \frac{10}{29} \\ 0 & -\frac{5}{\sqrt{29}} & \frac{10}{29} & \frac{4}{29} \end{pmatrix}$ et $A^{(3)} = H^{(2)}A^{(2)} = \begin{pmatrix} 3 & \frac{14}{3} \\ 0 & 0 \\ 0 & \frac{\sqrt{29}}{3} \\ 0 & 0 \end{pmatrix}$.

Factorisation.

On a donc $A = QR$ avec $Q = H^{(1)}H^{(2)} = \begin{pmatrix} \frac{2}{3} & -\frac{1}{3\sqrt{29}} & \frac{20}{29} & \frac{8}{29} \\ 0 & 0 & \frac{2}{\sqrt{29}} & -\frac{5}{\sqrt{29}} \\ \frac{2}{3} & \frac{8}{3\sqrt{29}} & -\frac{15}{29} & -\frac{6}{29} \\ \frac{1}{3} & -\frac{14}{3\sqrt{29}} & -\frac{10}{29} & -\frac{4}{29} \end{pmatrix}$ et $R = \begin{pmatrix} 3 & \frac{14}{3} \\ 0 & \frac{\sqrt{29}}{3} \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$.

Problème aux moindres carrés. Soit $y \in \mathbb{R}^2$. La matrice Q^T étant orthogonale, on a

$$\|Ay - b\|_2 = \|QRy - b\|_2 = \|Q^TQRy - Q^Tb\|_2 = \|Ry - Q^Tb\|_2.$$

Ainsi, le problème aux moindres carrés se réécrit :

$$\text{« trouver } x \in \mathbb{R}^2 \text{ tel que } \|Rx - Q^Tb\|_2 = \min_{y \in \mathbb{R}^2} \|Ry - Q^Tb\|_2 \text{ ».}$$

On note $R = \begin{pmatrix} \widetilde{R} \\ \hline 0_{\mathcal{M}_2} \end{pmatrix}$ et $Q^Tb = \begin{pmatrix} \widetilde{b_1} \\ \hline \widetilde{b_2} \end{pmatrix}$ avec $\widetilde{R} \in \mathcal{M}_2(\mathbb{R})$, $\widetilde{b_1} \in \mathbb{R}^2$, $\widetilde{b_2} \in \mathbb{R}^2$.

Alors $\|Ry - Q^T b\|^2_{2,\mathbb{R}^4} = \|\widetilde{R}y - \widetilde{b_1}\|^2_{2,\mathbb{R}^2} + \|\widetilde{b_2}\|^2_{2,\mathbb{R}^2}$.
Or $\widetilde{R}$ est inversible, le système (de dimension 2) $\widetilde{R}y = \widetilde{b_1}$ a une unique solution, notée x : x est l'unique solution du problème aux moindres carrés.
En effet, pour tout $y \in \mathbb{R}^2$, $\|Ry - Q^T b\|^2_{2,\mathbb{R}^4} \geq \|\widetilde{b_2}\|^2_{2,\mathbb{R}^2} = \|Rx - Q^T b\|^2_{2,\mathbb{R}^4}$
et $\|Ry - Q^T b\|^2_{2,\mathbb{R}^4} = \|\widetilde{b_2}\|^2_{2,\mathbb{R}^2}$ si et seulement si $\widetilde{R}y = \widetilde{b_1}$ si et seulement si $y = x$.

Le système $\widetilde{R}x = \widetilde{b_1}$ se résout immédiatement une fois calculé $Q^T b = \begin{pmatrix} 3 \\ \dfrac{3}{\sqrt{29}} \\ \dfrac{-35 + 4\sqrt{29}}{29} \\ \dfrac{-14 - 10\sqrt{29}}{29} \end{pmatrix}$. On doit résoudre

$\begin{pmatrix} 3 & \dfrac{14}{3} \\ 0 & \dfrac{\sqrt{29}}{3} \end{pmatrix} x = \begin{pmatrix} 3 \\ \dfrac{3}{\sqrt{29}} \end{pmatrix}$, on a donc $x = \begin{pmatrix} \dfrac{15}{29} \\ \dfrac{9}{29} \end{pmatrix}$.

Pour terminer, l'erreur d'approximation vaut

$$\|\widetilde{b_2}\|_{2,\mathbb{R}^2} = \frac{\sqrt{4785}}{29}$$

<u>**Méthode 2**</u> : **utilisation de l'équation normale et de la factorisation QR de $A^T A$.**

On considère la matrice $A = \begin{pmatrix} 2 & 3 \\ 0 & 0 \\ 2 & 4 \\ 1 & 0 \end{pmatrix}$ et on calcule $B = A^T A = \begin{pmatrix} 9 & 14 \\ 14 & 25 \end{pmatrix}$ et $A^T \begin{pmatrix} 1 \\ 2 \\ 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 9 \\ 15 \end{pmatrix}$. La solution du problème de moindres carrés associé à A est la solution de l'équation normale $Bx = \begin{pmatrix} 9 \\ 15 \end{pmatrix}$.
On peut facilement résoudre directement ce système 2x2 mais on va voir, pour illustrer le cours, comment le faire avec la factorisation QR de B obtenue par la méthode de Householder. Une étape suffira puisque la matrice est de taille 2x2.
La première colonne de B est $v_1 = \begin{pmatrix} 9 \\ 14 \end{pmatrix}$, on a $\|v_1\|^2 = 277$.
On pose $w_1 = v_1 - \|v_1\| e_1 = \begin{pmatrix} 9 - \sqrt{277} \\ 14 \end{pmatrix}$ qu'on note $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$. On définit ensuite

$$H = H(w_1) = I_2 - 2\frac{w_1 w_1^T}{\|w_1\|^2} = \frac{1}{\alpha^2 + \beta^2} \begin{pmatrix} \beta^2 - \alpha^2 & -2\alpha\beta \\ -2\alpha\beta & \alpha^2 - \beta^2 \end{pmatrix}$$

et enfin

$$R = HA = \frac{1}{\alpha^2 + \beta^2} \begin{pmatrix} 9(\beta^2 - \alpha^2) - 28\alpha\beta & 14(\beta^2 - \alpha^2) - 50\alpha\beta \\ -18\alpha\beta + 14(\alpha^2 - \beta^2) & -28\alpha\beta + 25(\alpha^2 - \beta^2) \end{pmatrix}$$

Comme on a posé $\alpha = 9 - \sqrt{277}$ et $\beta = 14$, on vérifie que le coefficient en bas à gauche de R est bien nul, c'est-à-dire $-18\alpha\beta + 14(\alpha^2 - \beta^2) = 0$. On peut alors simplifier les autres coefficients en utilisant l'égalité $14(\alpha^2 - \beta^2) = 18\alpha\beta$, ce qui donne

$$R = \frac{1}{\alpha^2 + \beta^2} \begin{pmatrix} \dfrac{277}{9}(\beta^2 - \alpha^2) & -68\alpha\beta \\ 0 & \dfrac{29}{9}(\alpha^2 - \beta^2) \end{pmatrix}$$

Par construction, la matrice H est symétrique et orthogonale donc la décomposition QR de B s'écrit avec $Q = H$, i.e.,

$$B = HR = \frac{1}{\alpha^2+\beta^2}\begin{pmatrix}\beta^2-\alpha^2 & -2\alpha\beta\\ -2\alpha\beta & \alpha^2-\beta^2\end{pmatrix}\frac{1}{\alpha^2+\beta^2}\begin{pmatrix}\frac{277}{9}(\beta^2-\alpha^2) & -68\alpha\beta\\ 0 & \frac{29}{9}(\alpha^2-\beta^2)\end{pmatrix}$$

On peut alors résoudre le système $Bx = \begin{pmatrix}9\\15\end{pmatrix}$. On écrit d'abord que

$$Rx = H^T\begin{pmatrix}9\\15\end{pmatrix} = H\begin{pmatrix}9\\15\end{pmatrix} = \frac{1}{\alpha^2+\beta^2}\begin{pmatrix}9(\beta^2-\alpha^2)-30\alpha\beta\\ -18\alpha\beta+15(\alpha^2-\beta^2)\end{pmatrix} = \frac{1}{\alpha^2+\beta^2}\begin{pmatrix}9(\beta^2-\alpha^2)-30\alpha\beta\\ \alpha^2-\beta^2\end{pmatrix}$$

et on est ramené au système triangulaire supérieur

$$\frac{1}{\alpha^2+\beta^2}\begin{pmatrix}\frac{277}{9}(\beta^2-\alpha^2) & -68\alpha\beta\\ 0 & \frac{29}{9}(\alpha^2-\beta^2)\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix} = \frac{1}{\alpha^2+\beta^2}\begin{pmatrix}9(\beta^2-\alpha^2)-30\alpha\beta\\ \alpha^2-\beta^2\end{pmatrix}$$

dont la solution vérifie $y = \frac{9}{29}$, $x = \frac{9}{277}(9 + \frac{258}{29}\frac{\alpha\beta}{\alpha^2-\beta^2}) = \frac{9}{277}(9 + \frac{258}{29}\frac{14}{18}) = \frac{15}{29}$. Ouf !

On calcule $A\begin{pmatrix}\frac{15}{29}\\ \frac{9}{29}\end{pmatrix} = \frac{1}{29}\begin{pmatrix}57\\0\\66\\15\end{pmatrix}$ puis l'erreur d'approximation en norme 2 avec b :

$$\|\frac{1}{29}\begin{pmatrix}57\\0\\66\\15\end{pmatrix} - \begin{pmatrix}1\\2\\3\\1\end{pmatrix}\| = \frac{1}{29}\|(28,-58,-21,-14)\| = \frac{\sqrt{4785}}{29}$$

Corrigé de l'exercice 4

Le théorème de Gerschgorin (également appelé théorème de Gerschgorin-Hadamard) affirme que l'ensemble des valeurs propres d'une matrice A de taille $n \times n$ à coefficients complexes est contenu dans l'union des n disques de Gerschgorin associés à la matrice A. Étant donné $i \in [\![1, n]\!]$ on désigne par i-ème disque de Gerschgorin associé à A l'ensemble

$$D_i = \{z \in \mathbb{C},\ |z - A_{i,i}| \leq \sum_{j \neq i} A_{j,i}\}$$

On démontre le théorème de la façon suivante : soit λ une valeur propre de A et $x = (x_1, \ldots, x_n)$ un vecteur propre associé. On a $Ax = \lambda x$ donc, pour tout $i \in [\![1, n]\!]$,

$$\sum_{j \in [\![1,n]\!]} A_{ij} x_j = \lambda x_i,$$

d'où

$$(\lambda - A_{i,i}) x_i = \sum_{j \in [\![1,n]\!] \setminus \{i\}} A_{ij} x_j.$$

On choisit un indice i_0 tel que $|x_{i_0}| = \max\limits_{i \in [\![1,n]\!]} |x_i|$. Puisque x est un vecteur propre de A, x_{i_0} est non nul et on a :

$$|\lambda - A_{i_0,i_0}| = \frac{1}{|x_{i_0}|} \left| \sum_{j \in [\![1,n]\!] \setminus \{i_0\}} A_{i_0 j} x_j \right| \leq \sum_{j \in [\![1,n]\!] \setminus \{i_0\}} |A_{i_0 j}| \frac{|x_j|}{|x_{i_0}|} \leq \sum_{j \in [\![1,n]\!] \setminus \{i_0\}} |A_{i_0 j}|$$

On en déduit que $\lambda \in D_{i_0} \subset \bigcup\limits_{i \in [\![1,n]\!]} D_i$. La démonstration vaut pour une valeur propre arbitraire de A donc $\sigma(A) \subset \bigcup\limits_{i \in [\![1,n]\!]} D_i$ et le théorème de Gerschgorin est démontré.

On peut énoncer un corollaire facile du théorème de Gerschgorin :

Toute matrice carrée à diagonale strictement dominante est inversible.

En effet, soit $A \in \mathcal{M}_n(\mathbb{C})$ (avec $n \in \mathbb{N}^*$) une telle matrice, on suppose donc que pour tout $i \in [\![1, n]\!]$, $|A_{i,i}| > \sum\limits_{j \neq i} |A_{j,i}|$. Cette hypothèse implique que $0 \notin \bigcup\limits_{i \in [\![1,n]\!]} D_i$ donc, d'après le théorème de Gerschgorin, $0 \notin \sigma(A)$, ce qui prouve que A est inversible.