2. Méthodes de Type Minimisation

Lorsque la matrice A est SDP. La solution du système linéaire $Ax=b$ est le point où la fonctionnelle $J : \mathbb{R}^n \to \mathbb{R}$ (forme Quadratique) :

$$\boxed{J(X) = \frac{1}{2} X \cdot AX - b \cdot X} \quad (*)$$

$$= \frac{1}{2} \sum_{i,j=1}^{n} a_{ij} \cdot x_i x_j - \sum_{i=1}^{n} b_i x_i$$

atteint son minimum. Les méthodes du gradient à pas fixe, du gradient à pas optimal et du gradient conjugué permettent d'approcher le minimum de la fonctionnelle J de $(*)$.

- Elles nécéssitent uniquement le calcul de produits d'une matrice par un vecteur et de produits scalaires.
- Elle sont particulièrement adaptés au cas des matrices creuses.

Lemme : $\boxed{A x^* = b \iff J(x^*) = \min_{x \in \mathbb{R}^n} J(x)}$

$(\Rightarrow)$ x^* est solution de $Ax=b$. Alors pour X qtq on a :

$$J(X) = \frac{1}{2} X \cdot AX - b \cdot X = J(X^* + h), \quad h = X - X^*$$

$$= J(X^*) + \frac{1}{2}(X - X^*) \nabla J(X^*) + \underbrace{\frac{1}{2}(X - X^*) \cdot A(X - X^*)}_{>0}$$

> 0 si $X \neq X^*$ $\forall X \in \mathbb{R}^n$

Puisque A est symétrique.

$$J(X) = \frac{1}{2}\left(\sum_{i,j=1}^{n} a_{ij} \cdot x_i x_j - \cancel{2} \sum_{i=1}^{n} b_i x_i\right)$$

$$\frac{\partial J}{\partial x_k}(X) = \frac{2}{2} \sum_{i=1}^{n} a_{ki} x_i - \frac{2 b_k}{2}$$

$$\nabla J(x^*) = \frac{\cancel{2}}{\cancel{2}}(AX^* - b) = 0$$

$$\Rightarrow J(X) \geq J(X^*) \quad \forall Y \in \mathbb{R}^n$$

$J(x^* + h) = \frac{1}{2}(x^* + h) \cdot A(x^* + h)$
$- b \cdot (x^* + h)$
$= \frac{1}{2} x^* \cdot Ax^* + \frac{1}{2} x^* \cdot Ah$
$+ \frac{1}{2} h \cdot Ax^* + \frac{1}{2} h \cdot Ah$
$- b \cdot x^* - b \cdot h$
$= \boxed{J(x^*) + (Ax^* - b) h + \frac{1}{2} h \cdot Ah} \quad (2*)$

$h \neq 0$

$\Leftarrow$ X^* réalise le minimum de J ds $\mathbb{R}^n$, on a :

$$J(X^* + \varepsilon Y) \geq J(X^*) \qquad \forall \varepsilon \in \mathbb{R} \quad \forall Y \in \mathbb{R}^n$$

En utilisant (2*) on trouve que

$$J(X^*) + \varepsilon Y \cdot (AX^* - b) + \frac{\varepsilon^2}{2} Y \cdot AY \geq J(X^*)$$

i.e $$\varepsilon Y \cdot (AX^* - b) + \frac{\varepsilon^2}{2} Y \cdot AY \geq 0$$

Pour $\varepsilon > 0$ ça devient : $$Y \cdot (AX^* - b) + \frac{\varepsilon}{2} Y \cdot AY \geq 0$$

En faisant tendre $\varepsilon \to 0$: $$\boxed{Y \cdot (AX^* - b) \geq 0 \quad \forall Y \in \mathbb{R}^n} \quad ①$$

Même Argument pour $J(X^* - \varepsilon Y) \geq J(X^*)$ on trouve

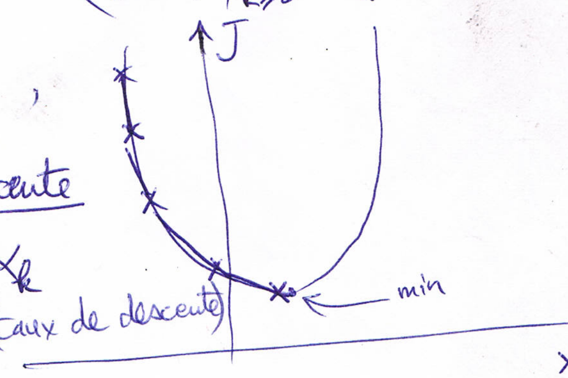
$$-Y \cdot (AX^* - b) \geq 0 \Rightarrow \boxed{Y \cdot (AX^* - b) \leq 0 \quad \forall Y \in \mathbb{R}^n} \quad ②$$

$$① + ② \Rightarrow \boxed{AX^* - b = 0}.$$

• Donc lorsque A est symétrique, définie positive (SDP), on peut chercher la solⁿ de $Ax = b$ en minimisant la fonctionnelle $J(\cdot)$ plutôt qu'en utilisant les méthodes du chapitre précédent.

• Principe de la méthode du Gradient : • Construire une suite X_k, $k \geq 0$ telle que la fonctionnelle $(J(X_k))_{k \geq 0}$ ↘ strictement décroissante

i.e $J(X_{k+1}) < J(X_k)$ ......,



à chaque itération d'une méthode de descente on détermine un vecteur p_k (direction de descente) et un scalaire α_k (taux de descente) telle que

$$X_{k+1} = X_k + \alpha_k p_k$$

$$J(X_{k+1}) = J(X_k + \alpha_k P_k) = \frac{1}{2}(X_k + \alpha_k P_k) \cdot A(X_k + \alpha_k P_k) - b \cdot (X_k + \alpha_k P_k)$$

$$= \frac{1}{2} X_k \cdot AX_k + \frac{\alpha_k}{2}(X_k \cdot AP_k) + \frac{\alpha_k}{2} P_k \cdot AX_k + \frac{1}{2}\alpha_k^2 P_k \cdot AP_k - b \cdot X_k - \alpha_k b \cdot P_k$$

$$= J(X_k) + \alpha_k (AX_k \cdot P_k) - \alpha_k b \cdot P_k + \frac{1}{2}\alpha_k^2 P_k \cdot AP_k$$

(2*)
$$= J(X_k) + \alpha_k \underbrace{(AX_k - b)}_{\nabla J(X_k)} \cdot P_k + \frac{1}{2}\alpha_k^2 P_k \cdot AP_k$$

- La partie principale de l'accroissement de J est majorée par le module du gradient fois le module de P_k. ($P_k = -\nabla J(X_k)$):

Donc la direction de la plus grande discente est donnée par

L'itération prend la forme : *Le residu* : $r_k = b - AX_k$

$$X_{k+1} = X_k - \alpha_k \nabla J(X_k), \quad \nabla J(X_k) = AX_k - b = -r_k$$ (résidu à l'étape k)

Donc $$\boxed{X_{k+1} = X_k + \alpha_k \cdot r_k}$$

- La méthode du Gradient à pas fixe : Le paramètre α_k est indépendant de k. L'algorithme s'écrit :

$$\boxed{X_{k+1} = X_k + \alpha\, r_k}$$

Notons M et ρ deux réels > 0 définis par :

$$M = \sup_{X \in \mathbb{R}^n, X \neq 0} \frac{X \cdot AX}{X \cdot X}, \quad \rho = \inf_{X \in \mathbb{R}^n, X \neq 0} \frac{X \cdot AX}{X \cdot X}$$

Proposition : (Convergence de la Méthode) : Si $0 < \alpha < \frac{2\rho}{M^2}$ la méthode du Gradient à pas fixe converge.

- $X_{k+1} - X = X_k - X - \alpha A(X_k - X)$

$$\|X_{k+1} - X\|_2^2 = \left((X_{k+1}-X), (X_{k+1}-X)\right) = \left(X_k - X - \alpha A(X_k - X), X_k - X - \alpha A(X_k - X)\right)$$
$$= (X_k - X, X_k - X) - 2\alpha (X_k - X)\cdot A(X_k - X) + \alpha^2 \left(A(X_k - X), A(X_k - X)\right)$$
$$= \|X_k - X\|_2^2 - 2\alpha\left(X_k - X \cdot A(X_k - X)\right) + \alpha^2 \|A(X_k - X)\|_2^2$$
$$\leq \|X_k - X\|_2^2 + (-2\alpha\rho + \alpha^2 M^2)\|X_k - X\|_2^2$$
$$\leq (1 - 2\alpha\rho + \alpha^2 M^2)\|X_k - X\|_2^2$$

si $0 < \alpha < \frac{2\rho}{M^2} \Leftarrow 1 - 2\alpha\rho + \alpha^2 M^2 < 1$ | d'où la cv.

- <u>La méthode du Gradient à pas optimal</u> : A chaque étape le paramètre α_k change de valeur. Il doit être solution de

$$\boxed{\text{Trouver } \alpha_k \text{ tq } J(X_k + \alpha_k r_k) = \inf_{\alpha \in \mathbb{R}} J(X_k + \alpha r_k)}$$

or d'après (2*)

$$J(X_k + \alpha r_k) = J(X_k) - \alpha\, r_k \cdot r_k + \frac{\alpha^2}{2} r_k \cdot A r_k$$

$$\Rightarrow J'_\alpha(X_k + \alpha r_k) \underset{(*)}{=} 0 \Rightarrow \boxed{\alpha_k = \frac{r_k \cdot r_k}{r_k \cdot A r_k}}$$

- Donc <u>l'itération</u> prend la forme :

$$\boxed{X_{k+1} = X_k + \frac{r_k \cdot r_k}{r_k \cdot A r_k} r_k \quad ; \quad k \geq 0}$$

<u>Remarques</u> 1 Les residus successives sont orthogonaux i.e

$$r_k \cdot r_{k+1} = 0 \quad (*) \qquad \text{(verifiez!)}$$

2 on trouve que :

$$J(X_k) - J(X_{k+1}) \geq \frac{\rho}{2}\|X_k - X_{k+1}\|_2^2$$

- La méthode est aussi convergente si $0 < \alpha < \frac{2}{\lambda_n}$.

<u>Pas fixe</u> : $X_{k+1} = X_k - \alpha (A X_k - b)$

$$\Rightarrow X_{k+1} = \underbrace{(I - \alpha A)} X_k + \alpha b$$

Conv. ssi $\rho(I - \alpha A) < 1$.

Supposons les val. propres de A sont rangeés par ordre croissant

$$0 < \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$$

$\Rightarrow$ Les valeurs propres de $I - \alpha A$ sont

$$1 - \alpha \lambda_n \leq \dots \leq \underline{1 - \alpha \lambda_1}$$

$$\rho(I - \alpha A) = \max_{1 \leq i \leq n} \{ |1 - \alpha \lambda_i| \}$$

or $|1 - \alpha \lambda_i| < 1$ ssi $-1 < 1 - \alpha \lambda_i < 1$

$\Rightarrow$ $1 - \alpha \lambda_i > -1 \Rightarrow \alpha \lambda_i < 2 \Rightarrow \alpha < \frac{2}{\lambda_i}$

$1 - \alpha \lambda_i < 1 \Rightarrow \alpha \lambda_i > 0 \Rightarrow \alpha > 0$

$$0 < \alpha < \frac{2}{\lambda_i} \qquad 0 < \lambda_1 \leq \lambda_n$$

$$\Rightarrow \boxed{0 < \alpha < \frac{2}{\lambda_n}}$$

car $$J(X_k)-J(X_{k+1}) = \frac{1}{2}(X_k-X_{k+1})\cdot A(X_k-X_{k+1}) + r_{k+1}\cdot \underbrace{(X_{k+1}-X_k)}_{\alpha_k r_k}$$

$$\left((r_k, \alpha_k r_k) - \frac{\alpha_k^2}{2}(Ar_k, r_k) = \frac{\alpha_k}{2}\|r_k\|_2^2 \right)$$

$$= \frac{1}{2}(X_k-X_{k+1})\cdot A(X_k-X_{k+1})$$

$$\geq \frac{\rho}{2}\|X_k-X_{k+1}\|_2^2$$

- Pour demontrer la conv. de cette méthode nous écrivons:

$$\alpha\|X_{k+1}-X\|_2^2 \leq (X_k-X)\cdot A(X_k-X)$$

$$= (X_k-X)\cdot r_k$$

$Ax-b$

$$\leq \|X_k-X\|\cdot\|r_k\|_2$$

$$\|X_{k+1}-X\|_2 \leq \frac{\|r_k\|_2}{\alpha} \quad (2*)$$

Mais $\|r_k\|_2^2 = r_k\cdot(r_k-r_{k+1}) \leq \|r_k\|_2\|r_k-r_{k+1}\|_2$.

Donc $\|r_k\|_2 \leq \|r_k-r_{k+1}\|_2$

$$\Rightarrow \|X_{k+1}-X\|_2 \overset{(2*)}{\leq} \frac{\|r_k-r_{k+1}\|_2}{\alpha}$$

$$\leq \frac{c}{\alpha}\|X_k-X_{k+1}\|_2.$$

$$\boxed{\begin{aligned} r_k &= AX_k-b \\ r_{k+1} &= AX_{k+1}-b \\ \Rightarrow r_k-r_{k+1} &= A(X_k-X_{k+1}) \\ \|r_k-r_{k+1}\|_2 &\leq \|A\|\,\|X_k-X_{k+1}\|_2 \end{aligned}}$$

Voir suite (18') & (18'')

- La méthode du Gradient conjugué : L'idée de base et de choisir un ensemble $(f_1,\dots,f_n)$ composé de n vecteurs linéairement indépendants pour former une base de $\mathbb{R}^n$ et puis on écrit la sol^n de $Ax=b$ (notée x^*) en fct de cette base ; i.e :

$$\boxed{x^* = \sum_{j=1}^{n}\alpha_j^* f_j}$$

<u>Proposition</u> : X^* sol^{n} exacte. On note $\|\cdot\|_A$ la norme associée au produit scalaire $(\cdot,\cdot)_A$ tq $(X,Y)_A = (X, AY)$, $X, Y \in \mathbb{R}^n$

~~alors~~ on a

$$\|X_k - X^*\|_A \leq \left(\frac{K_2(A)-1}{K_2(A)+1}\right)^k \|X_0 - X^*\|_A \,, \quad K_2(A) = \frac{\max \lambda}{\min \lambda}$$

<u>Preuve</u> : On remarque que

$$\|X_k - X^*\|_A^2 = (A^{-1} r_k ; r_k)$$

$\|-X_k + X^*\|_A^2 = (-X_k + X^*, (A(-X_k + X^*)$

$r_k = -AX_k + b = -AX_k + AX^*$
$= A(-X_k + X^*)$
$\Rightarrow -X_k + X^* = A^{-1} r_k$

or $$(A^{-1} r_{k+1} ; r_{k+1}) \leq \left(\frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1}\right)^2 (A^{-1} r_k ; r_k) \quad (*)$$

$(0 < \lambda_1 \leq \lambda_1 \leq \dots \leq \lambda_n)$

i.e $$\|X_{k+1} - X^*\|_A^2 \leq \left(\frac{K_2(A)-1}{K_2(A)+1}\right)^2 \|X_k - X^*\|_A^2$$

on obtient Ainsi par recurrence que

$$\boxed{\|X_k - X^*\|_A \leq \left(\frac{K_2(A)-1}{K_2(A)+1}\right)^k \|X_0 - X^*\|_A}$$

; avec $\left(K_2(A) = \frac{\lambda_n}{\lambda_1}\right)$

cond. de A.

<u>Preuve de $(*)$</u>

$(*)$ vient du fait que : $x_{k+1} = x_k + \alpha_k r_k \Rightarrow r_{k+1} = r_k - \alpha_k A r_k$ ($r = b - A(\cdot)$)

par déf. de α_k on a $(r_k, r_{k+1}) = 0$. On en déduit :

$$(A^{-1} r_{k+1}, r_{k+1}) = (A^{-1} r_k - \alpha_k r_k , r_{k+1})$$
$$= (A^{-1} r_k, r_{k+1}) = (A^{-1} r_k, r_k) - \alpha_k (r_k, r_k).$$

Donc $$\frac{(A^{-1} r_{k+1}, r_{k+1})}{(A^{-1} r_k, r_k)} = 1 - \frac{(r_k, r_k)^2}{(A r_k, r_k)(A^{-1} r_k, r_k)}$$

Or d'après l'inégalité de <u>Kantorovich</u> on a :

$$(A r_k, r_k)(A^{-1} r_k, r_k) \leq \frac{(\lambda_1 + \lambda_n)^2}{4 \lambda_1 \lambda_n} (r_k, r_k)^2$$

(18'')

$$1 - \frac{(r_k, r_k)^2}{(Ar_k, r_k)(A^{-1} r_k, r_k)} \leq \left(1 - \frac{4\lambda_1\lambda_n}{(\lambda_1+\lambda_n)^2}\right)$$

$$\leq \left(\frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1}\right)^2$$

i.e $\|X_{k+1} - X^*\|_A^2 \leq \left(\frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1}\right)^2 \| X_k - X^* \|_A^2$. cqfd

$$\left\{ \leq \left(\frac{K_2(A) - 1}{K_2(A) + 1}\right)^2 \|X_k - X^*\|_A^2 . \right\}$$

$\lambda_n^2 + \cancel{2\lambda_1\lambda_n} + \lambda_1^2 - \cancel{4\lambda_1\lambda_n}$

$-2\lambda_1\lambda_n$