Gradient Conjugué

$\vec{x}^i$ vecteur
x_i scalaire

Considérons A, une matrice symétrique définie positive.

- (Ax, y) définit un produit scalaire.
- $\bar{x}$ est la solution du système $Ax = b$
- La fonctionnelle $J(x) = \frac{1}{2}(Ax, x) - (b, x)$ réalise son minimum en $\bar{x}$.

Il est équivalent de minimiser $J(x)$ et la fonctionnelle $E(x)$ définit par

$$E(x) = \|x - \bar{x}\|_A = (A(x - \bar{x}), x - \bar{x})$$

car $E(x) = (Ax, x) - 2\underbrace{(x, A\bar{x})}_{\substack{=(A\bar{x}, x) \\ =(b, x)}} + (A\bar{x}, \bar{x})$

$$= 2\,J(x) + \underbrace{(A\bar{x}, \bar{x})}_{\text{Constante}}$$

Pour la méthode du gradient conjugué, on choisira comme critère d'optimalité :

$$\min_{x^p \in x^0 + K_p(A, r^0)} E(x^p)$$

$$\Leftrightarrow \min_{x^p \in x^0 + K_p(A, r^0)} \|x^p - \bar{x}\|_A$$

Pour définir la méthode, on doit construire une base de $K_p(A, r^0)$.

On note

$$\underbrace{W(p)}_{\in \mathcal{M}_{np}(\mathbb{R})} = \left(w^1 \quad w^2 \quad \dots \quad w^p \right) \begin{cases} \text{avec } (w^i)_{i \in 1..n} \text{ forment} \\ \text{une base de } K_p(A, r^0) \\ \text{avec } \underline{w^i \in \mathbb{R}^n} \ \forall i \in 1..p \end{cases}$$

Dans cette base, le critère d'optimalité se réécrit :

$$\min_{x^p \in x^0 + K_p(A, r^0)} \| x^p - \bar{x} \|_A = \min_{x^p \in x^0 + \underbrace{K_p(A, r^0)}} \left(A(x^p - \bar{x}), x^p - \bar{x} \right)$$

avec $x^p = x^0 + \sum_{i=1}^{p} z_i w^i \qquad \underline{z_i \in \mathbb{R}}$

$$= x^0 + W(p) Z^p \quad \text{avec } Z^p = \begin{pmatrix} z_1 \\ \vdots \\ z_p \end{pmatrix}$$

D'où

$$\min_{x^p \in x^0 + K_p(A, r^0)} \| x^p - \bar{x} \|_A \iff \min_{z^p \in \mathbb{R}^p} \left(A(x^0 + W(p) z^p - \bar{x}), x^0 + W(p) z^p - \bar{x} \right)$$

$$\iff \min_{z^p \in \mathbb{R}^p} \left(\underbrace{A(x^0 - \bar{x})}_{-r^0} + A W(p) z^p, \underbrace{x^0 - \bar{x}}_{-A r^0} + W(p) z^p \right)$$

$$\iff \min_{z^p \in \mathbb{R}^p} \left(A W(p) z^p, W(p) z^p \right) - \left(r^0, W(p) z^p \right) + \left(r^0, A r^0 \right) - \left(r^0, W(p) z^p \right)$$

$$\iff \min_{z^p \in \mathbb{R}^p} \left(A W(p) z^p - 2 r^0, W(p) z^p \right) + \underbrace{(r^0, A r^0)}_{\text{constante}}$$

$$\iff \min_{z^p \in \mathbb{R}^p} \left(W^T_{(p)} A W(p) z^p - 2 W^T_{(p)} r^0, z^p \right)$$

Or $f : z^p \longmapsto \left(W^T(p) A W(p) z^p - 2 W^T(p) r^0 , z^p \right)$

est minimum lorsque

$$W^T(p) A W(p) z^p - 2 W^T(p) r^0 = 0$$

$$\Leftrightarrow \quad W^T(p) A W(p) z^p = 2 W^T(p) r^0 \qquad *$$

Afin d'isoler facilement z^p, on va chercher à determiner $W(p)$ de sorte que

$$W^T(p) A W(p) = D(p)$$

où $D(p)$ est une matrice diagonale définie positive

Remarque

Si on arrive à trouver une telle base $W(p)$ tq $W^T(p) A W(p) = D(p)$ alors

à l'étape p : $z^p = 2 D^{-1}(p) W^T(p) r^0$ par $*$

$$z^p = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_p \end{pmatrix} = 2 \begin{pmatrix} \frac{1}{d_{11}} & 0 & \cdots & \cdots & 0 \\ 0 & \frac{1}{d_{22}} & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & \frac{1}{d_{pp}} \end{pmatrix} \underbrace{\begin{pmatrix} (w^1)^T \\ \vdots \\ (w^p)^T \end{pmatrix}}_{w_i \in \mathbb{R}^n} \begin{pmatrix} r_0 \\ r_1 \\ \vdots \\ r_{p-1} \end{pmatrix}$$

à l'étape p+1 :

$$z^{p+1} = \begin{pmatrix} z_1 \\ \vdots \\ z_p \\ z_{p+1} \end{pmatrix} = 2 \begin{pmatrix} \frac{1}{d_{11}} & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & & & \vdots \\ \vdots & & \frac{1}{d_{pp}} & & 0 \\ 0 & \cdots & \cdots & 0 & \frac{1}{d_{p+1p+1}} \end{pmatrix} \begin{pmatrix} (w^1)^T \\ \vdots \\ (w^p)^T \\ (w^{p+1})^T \end{pmatrix} \begin{pmatrix} r_0 \\ \vdots \\ r_{p-1} \\ r_p \end{pmatrix}$$

À l'étape $p+1$, les composantes 1 à p de z^p reste inchangé,

$$z^{p+1} = \begin{pmatrix} z^p \\ z_{p+1} \end{pmatrix}$$

$$z^{p+1} = 2 \begin{pmatrix} D(p) \\ \frac{1}{d_{p+1\,p+1}} \end{pmatrix} \begin{pmatrix} W(p)^T \\ (w^{p+1})^T \end{pmatrix} \begin{pmatrix} r^{p-1} \\ r_p \end{pmatrix}$$

De plus

$$\begin{aligned} x^{p+1} &= x^0 + W(p+1)\, z^{p+1} \\ &= x^0 + \begin{pmatrix} W(p) & w^{p+1} \end{pmatrix} \begin{pmatrix} z^p \\ z_{p+1} \end{pmatrix} \\ &= x^0 + W(p)\, z^p + w^{p+2} z_{p+2} \\ &= x^p + w^{p+2} z_{p+1} \end{aligned}$$

$\boxed{\text{Construction de la base } W(p)}$

proposition

Le vecteur $r^p = b - Ax^p$ est orthogonale à $K_p(A, r^0)$ et appartient à $K_{p+1}(A, r^0)$. Il est donc colinéaire à v^{p+1} (base orthonormale d'Arnoldi).

preuve

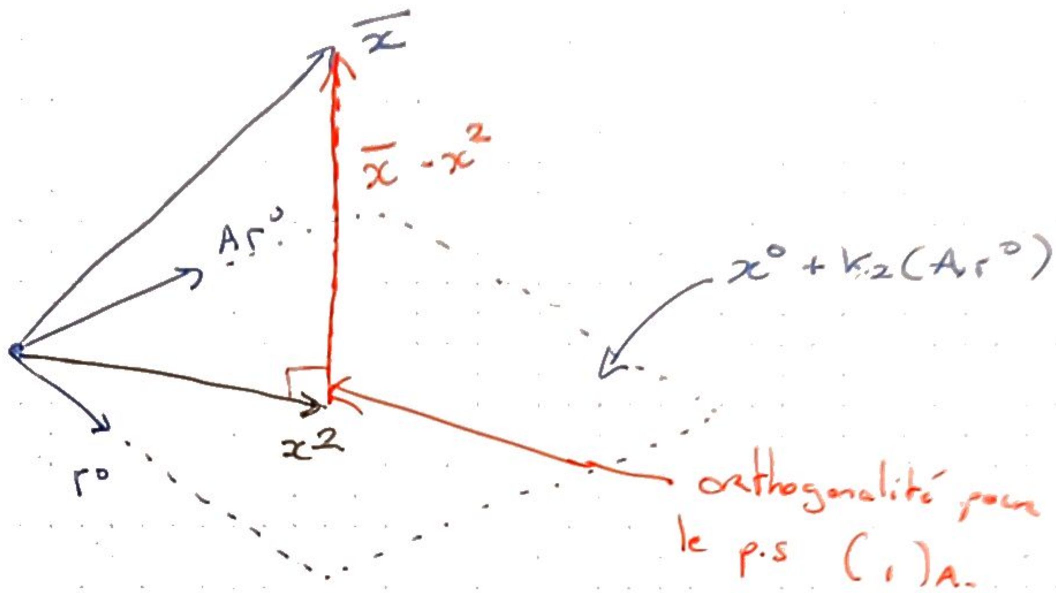
- Soit x^p qui réalise $\min\limits_{x \in x^0 + K_p(A, r^0)} \| x - \bar{x} \|_A$

particulier

quelconque

x^p est la projection orthogonale de $\bar{x}$ sur $x^0 + K_p(A, r^0)$ pour le produit scalaire $(\,,\,)_A$

Par exemple pour $K_2(A, r^0) = (r^0, Ar^0)$



donc $x^p \in x^0 + K_p(A, r^0)$

et $(\overline{x} - x^p, y)_A = 0 \quad \forall y \in K_p(A, r^0)$

$\Leftrightarrow \quad (\underbrace{A\overline{x}}_{b} - Ax^p, y) = 0 \quad \forall y \in K_p(A, r^0)$

$\Leftrightarrow \quad (r^p, y) = 0 \quad \forall y \in K_p(A, r^0)$

donc $r^p \perp K_p(A, r^0)$

- $x^p \in K_p(A, r^0) \Leftrightarrow \exists (\alpha_i)_{i \in 1..p}$ tq

$$x^p = \sum_{i=1}^{p} \alpha_i A^{i-1} r^0$$

$$r^p = b - Ax^p = b - \sum_{i=1}^{p} \alpha_i A^i r^0 \in K_{p+1}(A, r^0)$$

- Rappelons que $(v^1\ v^2 \dots v^{p+1})$ est la base orthonormée de $K_{p+1}(A, r^0)$ formée par l'algorithme d'Arnoldi

donc $\exists (\alpha_i)_{i \in 1..p+1}$ tq $r^p = \sum_{i=1}^{p+1} \alpha_i v^i$.

Puisque $r^p \perp K_p(A, r^0)$ alors

$$(r^p, v^i) = 0 \quad \forall i \in 1..p$$

$$\Leftrightarrow \quad \alpha_i = 0 \quad \forall i \in 1..p$$

donc $r^p = \alpha_{p+1} v^{p+1}$

$\Leftrightarrow$ r^p colinéaire à v^{p+1}. $\boxtimes$

On a donc

$\in M_{np}(\mathbb{R}) \rightarrow G(p) = (r^0\ r^1 \dots r^{p-1})$ $\quad r^i \in \mathbb{R}^n \ \forall i \in 0..p-1$

qui est une base <u>orthogonale</u> de $K_p(A, r^0)$

$(r^i, r^j) = 0 \quad \forall i \neq j$

Notons $H(p) = G^T(p) A\, G(p)$. $\in M_p(\mathbb{R})$

On a $H(p)$ qui est
- symétrique
- définit positive
- tridiagonale.

<u>Preuve</u>

- $H^T(p) = (G^T_{(p)} A\, G(p))^T = G^T(p) A^T G(p) = G^T(p) A\, G(p)$

$\Rightarrow H(p)$ symétrique.

- Soit $x \in \mathbb{R}^p \neq 0$

$$\Rightarrow (G^T(p) A G(p) x, x) = (A \underbrace{G(p)x}_{Y}, \underbrace{G(p)x}_{Y}) \geq 0$$ car A définie positive

$\Leftarrow$ Supposons que $(G^T_{(p)} A G(p) x, x) = 0$

$$\Leftrightarrow (A G(p) x, G(p) x) = 0$$

Comme $A \in S^{++}(\mathbb{R})$ alors

$$G(p) x = 0 \qquad G(p) = (r^0 \; \dots \; r^{p-1})$$

$$\Leftrightarrow \sum_{i=1}^{p} x_i r^{i-1} = 0$$

où x_i sont les coefficients de x dans $\mathbb{R}^p$

Puisque $G(p)$ forme une une base de $K_p(A, r^0)$ alors $r^{i-1} \neq 0 \; \forall i \in 1..p$ donc $x_i = 0 \; \forall i \in 1..p$

- $(H(p))_{ij} = (H(p) e_i, e_j)$

$$= (A G(p) e_i, G(p) e_j)$$

$$(r^0 \; r^1 \dots r^i \dots r^{p-1}) \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \leftarrow i = r^i$$

$$= (A r^i, r^j)$$

OR

$$r^j \perp K_j(A, r^0)$$

et $A r^i = A(b - A \underbrace{x^i}_{\in x^0 + K_i(A, r^0)})$

$$= \underbrace{Ab - \underbrace{A^2 x^i}_{\in x^0 + K_{i+2}(A, r^0)}}_{\in x^0 + K_{i+2}(A, r^0)}$$

Donc si $j \geq i+2$

$$(Ar^i, r^j) = 0$$

d'où $(H(p))_{ij} = 0 \quad \forall j \geq i+2$

et comme $H(p)$ est symétrique

$$(H(p))_{ij} = 0 \quad \forall |i-j| \geq 2$$

Donc $H(p)$ est tridiagonale ▨

On a donc $H(p) = G^T(p) A G(p)$ définit positive.

$\Rightarrow$ on peut décomposer $H(p)$ avec Cholesky.

$$H(p) = G^T(p) A G(p) = \tilde{L} \tilde{L}^T$$
$$= L D L^T$$

Comme $H(p)$ est tridiagonale on a L triangulaire bidiagonale à diagonale unité.

$$L = \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ l_{21} & 1 & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & l_{p-1\,p} & 1 \end{pmatrix} \in \mathcal{M}_p(\mathbb{R})$$

et $D = \text{diag}(\tilde{L}_{11}, \ldots, \tilde{L}_{pp})$

Finalement, on construit la base $W(p)$ à l'aide de la relation :

$$W(p) = G(p)(L^T)^{-1}$$

$$\Leftrightarrow \quad W(p) L^T = G(p)$$

$$\underbrace{\begin{pmatrix} w_{11} & w_{12} & \cdots & w_{1p} \\ w_{21} & w_{22} & \ddots & \vdots \\ \vdots & \vdots & & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{np} \end{pmatrix}}_{n \times p} \underbrace{\begin{pmatrix} 1 & l_{12} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ & & & 1 & l_{p-1,p} \\ 0 & \cdots & & 0 & 1 \end{pmatrix}}_{p \times p} = \underbrace{\begin{pmatrix} r^0 & r^1 & r^2 \cdots & r^{p-1} \end{pmatrix}}_{n \times p}$$

$\Longleftrightarrow$

$$\begin{pmatrix} w_{11} & w_{11} l_{12} + w_{12} & \cdots & w_{1,p-1} l_{p-1,p} + w_{1p} \\ w_{21} & w_{21} l_{12} + w_{22} & \cdots & w_{2,n-1} l_{p-1,p} + w_{2p} \\ \vdots & \vdots & & \vdots \\ w_{n1} & w_{n1} l_{12} + w_{n2} & & w_{n,p-1} l_{p-1,p} + w_{np} \end{pmatrix} = \begin{pmatrix} r^0 & \cdots & r^{p-1} \end{pmatrix}$$

$$\Longleftrightarrow \left(w^1 \quad w^1 \underbrace{l_{12}}_{=l_{21}} + w^2 \quad \cdots \quad w^{p-1} \underbrace{l_{p-1,p}}_{=l_{p,p-1}} + w^p \right) = \left(r^0 \cdots r^{p-1} \right)$$

On a donc
$$\boxed{\begin{aligned} r^0 &= w^1 \\ r^i &= w^{i+1} + l_{i+1,i}\, w^i \quad \forall i \in 1..p-1 \end{aligned}}$$

Remarque Les $(w^i)_{i \in 1..p}$ forment une base de $K_p(A, r^0)$

et $W_{(p)}^T A W_{(p)} = D(p)$

Les w^i sont A-conjugué :

$$(Aw^i, w^j) = \delta_{ij} d_{ii}$$

- Ainsi
$$\begin{aligned} (Ar^p, w^p) &= (Aw^{p+1} + l_{p+1,p} A w^p, w^p) \\ &= \underbrace{(Aw^{p+1}, w^p)}_{0} + l_{p+1,p} (Aw^p, w^p) \end{aligned}$$

$$\Longleftrightarrow \quad l_{p+1,p} = \frac{(Ar^p, w^p)}{(Aw^p, w^p)}$$

- et $x^{p+1} = x^p + w^{p+1} z_{p+1}$

$$(Ax^{p+1}, w^{p+1}) = (Ax^p + z_{p+1} A w^{p+1}, w^{p+1})$$
$$= (Ax^p, w^{p+1}) + z_{p+1} (A w^{p+1}, w^{p+1})$$

OR $r^p = b - A x^p \Leftrightarrow Ax^p = b - r^p$

donc

$$(b - r^{p+1}, w^{p+1}) = (b - r^p, w^{p+1}) + z_{p+1} (A w^{p+1}, w^{p+1})$$

$$\cancel{(b, w^{p+1})} - \underbrace{(r^{p+1}, \underbrace{w^{p+1}}_{\in K_{p+1}(A, r^0)})}_{0} = \cancel{(b, w^{p+1})} - (r^p, w^{p+1}) + z_{p+1} (A w^{p+1}, w^{p+1})$$

$\in K_{p+1}(A, r^0)$ et $r^{p+1} \perp K_{p+1}(A, r^0)$

donc $(r^p, w^{p+1}) = z_{p+1} (A w^{p+1}, w^{p+1})$

$$\Leftrightarrow z_{p+1} = \frac{(r^p, w^{p+1})}{(A w^{p+1}, w^{p+1})}$$

- De plus $r^{p+1} = b - A x^{p+1}$

$$= b - Ax^p - z_{p+1} A w^{p+1}$$
$$= r^p - z_{p+1} A w^{p+1}$$

Résumé

- $w^1 = r^0$ $\quad r^0 = b - Ax^0$
- $z_{p+1} = \dfrac{(r^p, w^{p+1})}{(A w^{p+1}, w^{p+1})}$
- $x^{p+1} = x^p + z_{p+1} w^{p+1}$
- $r^{p+1} = r^p - z_{p+1} A w^{p+1}$
- $\rho_{p+1,p} = \dfrac{(A r^p, w^p)}{(A w^p, w^p)}$

⚠ pour $p > 1$

- $w^{p+1} = r^p - \rho_{p+1,p} w^p$

Algorithme

$r \leftarrow b - Ax^0$

$w \leftarrow r$

Tant que $\|r\| \leq \varepsilon \|r^0\|$

$$z \leftarrow \frac{(r,w)}{(Aw,w)} \quad \left(\leftarrow \frac{\|r^p\|^2}{(Aw^{p+1}, w^{p+1})} \right)$$

$x \leftarrow x + zw$

$r \leftarrow r - zAw$

$$l \leftarrow \frac{(Ar,w)}{(Aw,w)} \quad \left(\leftarrow \frac{\|r^{p+1}\|^2}{\|r^p\|^2} \right)$$ avec variable en plus pour stocker le résidu d'avant

$w \leftarrow r - lw$

Fin Tant Que

→ Converge en au plus n itérations.

→ $$z_{p+1} = \frac{(r^p, w^{p+1})}{(Aw^{p+1}, w^{p+1})} = \frac{(r^p, r^p) - l_{p+1,p}(r^p, w^p)}{(Aw^{p+1}, w^{p+1})} = \frac{\|r^p\|^2}{(Aw^{p+1}, w^{p+1})}$$

$\in K_p(A, r^0)$ et $r^p \perp K_p(A, r^0)$

→ $$\begin{aligned}(r^{p+1}, r^p) &= (r^p - z_{p+1} A w^{p+1}, r^p) \\ &= \|r^p\|^2 - z_{p+1}(Aw^{p+1}, r^p) \\ &= \|r^p\|^2 - z_{p+1}(Aw^{p+1}, w^{p+1} + l_{p+1,p} w^p) \\ * &= \|r^p\|^2 - z_{p+1}(Aw^{p+1}, w^{p+1}) + z_{p+1} l_{p+1,p} \underbrace{(Aw^{p+1}, w^p)}_{0} \\ &= \|r^p\|^2 - \frac{\|r^p\|^2}{(Aw^{p+1}, w^{p+1})}(Aw^{p+1}, w^{p+1}) = 0\end{aligned}$$

⇒ les résidus sont orthogonaux pour le p.s usuel

$\longrightarrow \quad l_{p+1,p} = \dfrac{(Ar^p, w^p)}{(Aw^p, w^p)} = \dfrac{(Aw^p, r^p)}{(Aw^p, w^p)}$ pour $p > 1$

Or

- $(Aw^p, r^p) = \dfrac{1}{z_{p}}(r^{p-1} - r^{p}, r^p)$ car $r^{p+1} = r^p - z_{p+1} A w^{p+1}$

$= \dfrac{1}{z_p}\underbrace{(r^{p-1}, r^p)}_{0} - \dfrac{1}{z_p}(r^p, r^p)$

$= \dfrac{-\|r^p\|^2}{z_p}$

- $(Aw^p, w^p) = \dfrac{\|r^{p-1}\|^2}{z_p}$ par $*$

D'où $l_{p+1,p} = \dfrac{-\|r^p\|^2}{\|r^{p-1}\|^2}$ pour $p > 1$

Complexité

Tant que $\|r^p\|^2 \leq \varepsilon \|r^0\|^2$

$z \leftarrow \dfrac{\|r^p\|^2}{(Aw, w)}$

$x \leftarrow x + z w$

$r^{p+1} \leftarrow r^p - z A w$

$L \leftarrow -\dfrac{\|r^{p+1}\|^2}{\|r^p\|^2}$

$w \leftarrow r^{p+1} - L w$

$r^p \leftarrow r^{p+1}$

fin Tant Que

ou $L \leftarrow \dfrac{\|r^{p+1}\|^2}{\|r^p\|^2}$, $w \leftarrow r^{p+1} + L w$

	multiplication / division	addition / soustraction
Aw	$n \times n$	$n \times (n-1)$
(Aw, w)	n	$n-1$
z	1	
x	n	n
r	n	n
$\|r\|^2$	n	$n-1$
L	1	
w	n	n
Total	$n^2 + 5n + 2$	$n^2 + 4n - 2$

Si on itère au maximum n fois, on aboutit à une complexité de l'ordre de $2n^3$

mais avec un bon conditionnement la complexité devient inférieure à n.