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

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

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

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

- Pour demontrer la conv. de cette methode 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 $(f_1, .., f_n)$ 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 :

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

• Pour calculer facilement les α_j^* les vecteurs $(f_1, \ldots, f_n)$ sont choisis (19)
A - conjugués i.e

$$f_i \cdot A f_l = 0 \qquad i \neq l$$

Pour obtenir que $\boxed{\alpha_j^* = \dfrac{f_j \cdot b}{f_j \cdot A f_j}}$

En effet : $f_j \cdot b = f_j A x^*$

$$= f_j \cdot A \left(\sum_{i=1}^{n} \alpha_i^* f_i \right)$$

$$= \sum_{i=1}^{n} \alpha_i^* f_j \cdot A f_i = \alpha_j^* f_j \cdot A f_j$$

$$\Rightarrow \boxed{\alpha_j^* = \frac{f_j \cdot b}{f_j \cdot A f_j}}.$$

L'itération est definie par :

$$\boxed{\begin{aligned} & X^{(0)}_{(0)} = 0 \\ & X_k = \sum_{j=1}^{k} \alpha_j^* f_j \qquad k \geq 1 \end{aligned}}$$

i.e le passage de X_{k-1} à X_k et à partir du vecteur de direction f_k. Donc :

$$\boxed{X_k = X_{k-1} + \alpha_k^* f_k}$$

Géométriquement cela signifie : 1/ Choisir une direction de descente f_k.
2/ Minimiser J dans cette direction.

$$J(x_{k-1}+\alpha_k f_k) = \min_{\alpha\in\mathbb{R}} J(x_{k-1}+\alpha f_k)$$

or d'après (2*) :

$$J(x_{k-1}+\alpha f_k) = J(x_{k-1}) + \alpha \underbrace{(Ax_{k-1}-b)^t}_{-r_{k-1}} \cdot f_k + \frac{1}{2}\alpha^2 f_k^t A f_k$$

La direction la plus descendue est "Steepest descent" :

$$-\nabla J(x_{k-1}) = b - Ax_{k-1} = r_{k-1}$$

- On peut essayer de choisir $f_k = r_{k-1}$, mais les f_j ne sont pas A-conjugués.

On écrit donc :

$$f_k = r_{k-1} - \sum_{j=1}^{k-1} \beta_j \cdot f_j, \quad \text{avec } \boxed{\beta_j = \frac{f_j \cdot A r_{k-1}}{f_j \cdot A f_j}}$$

Le vecteur le plus proche de r_{k-1} ds $\|\cdot\|_2$.

- Est ce qu'on a CV si $f_k = 0$? oui parce que

$$\text{si } f_k = 0 \Rightarrow r_{k-1} = \sum_{j=1}^{k-1} \beta_j f_j$$

$$r_{k-1}\cdot r_{k-1} = \left(\sum_{j=1}^{k-1} \beta_j f_j\right)\cdot r_{k-1} = 0 \Rightarrow r_{k-1} = 0$$

$$\Rightarrow Ax_{k-1} = b \Rightarrow x_{k-1} = x^* \text{ d'où la conv.}$$

- Les résidus r_j sont orthogonaux i.e : $\boxed{r_j \cdot r_k = 0 \quad j<k}$

Donc l'algorithme est :

$x_0 = 0$

Pour $k = 1, \dots, n$ faire

$r_{k-1} = b - Ax_{k-1}$;

Si $r_{k-1} = 0$ Alors écrire $x_{k-1} = x^*$ et sortir.

Si $k = 1$ Alors $f_1 = r_0$

Sinon $f_k = r_{k-1} - \sum_{j=1}^{k-1} \frac{f_j \cdot A r_{k-1}}{f_j \cdot A f_j} f_j$

$$\boxed{\begin{aligned} \alpha_k^* &= \frac{b \cdot f_k}{f_k \cdot A f_k} \\ x_k &= x_{k-1} + \alpha_k^* f_k \end{aligned}}$$

Remarque : Dans chaque itération on a un certain nombre de multiplication Matrice-Vecteur et les $(f_i)_{i=\overline{1,n}}$ doivent être tous sauvegarder

$\Rightarrow$ L'algorithme ~~et moin efficace~~. (manque d'efficacité)

• Un peut de changement ~~et~~ sont necessaire pour rendre cet algorithme plus efficace. et simplification

① On peut augmenter l'efficacité on observons que :

$$\boxed{f_j \cdot A r_{k-1} = 0 \qquad j < k-1} \quad (*)$$

et l'expression des f_k ds l'algorithme devient uniquement :

$$\boxed{f_k = r_{k-1} - \frac{f_{k-1} \cdot A r_{k-1}}{f_{k-1} \cdot A f_{k-1}} f_{k-1}.}$$

(*) vient de :

$$\begin{aligned} f_j \cdot A r_{k-1} &= \alpha_j^* A f_j \cdot r_{k-1} / \alpha_j^* \\ &= (A x_j - A x_{j-1}) \cdot r_{k-1} / \alpha_j^* \\ &= (A x_j - b + b - A x_{j-1}) \cdot r_{k-1} / \alpha_j^* \\ &= (-r_j + r_{j-1}) \cdot r_{k-1} / \alpha_j^* = 0 \end{aligned}$$

$A x_j = A x_{j-1} + \alpha_j A f_j$

argument (2*)

② On peut aussi augmenter l'efficacité de l'algo. si on ecrit :

ⓐ $$\alpha_k^* = \frac{b \cdot f_k}{f_k \cdot A f_k} = \frac{(A x_{k-1} + r_{k-1}) \cdot f_k}{f_k \cdot A f_k} = \boxed{\frac{r_{k-1} \cdot f_k}{f_k \cdot A f_k}}$$

car $$\frac{A x_{k-1} \cdot f_k}{f_k \cdot A f_k} = \frac{x_{k-1} \cdot A f_k}{f_k \cdot A f_k} = \frac{\left(\sum_{j=1}^{k-1} \alpha_j^* f_j\right) \cdot A f_k}{f_k \cdot A f_k} = 0$$

Donc $\alpha_k^* = \dfrac{r_{k-1} \cdot f_k}{f_k \cdot A f_k} = \dfrac{r_{k-1} \cdot \left(r_{k-1} - \sum_1^{k-2} \beta_j f_j\right)}{f_k \cdot A f_k}$

$$= \frac{r_{k-1} \cdot r_{k-1}}{f_k \cdot A f_k}$$

L'expression de f_k :

(b) $\dfrac{f_{k-1} \cdot A r_{k-1}}{f_{k-1} \cdot A f_{k-1}} = \dfrac{f_{k-1} \cdot A r_{k-1}}{f_{k-1} \cdot A r_{k-2}}$ puisque $f_{k-1} = r_{k-2} - \sum_{j=1}^{k-2} f_j \cdot \beta_j$.

$$= \frac{(r_{k-2} - r_{k-1}) \cdot r_{k-1}}{(r_{k-2} - r_{k-1}) \cdot r_{k-2}} \quad \text{(comme ds (2*))}$$

$$= - \frac{r_{k-1} \cdot r_{k-1}}{r_{k-2} \cdot r_{k-2}} \quad \text{(par l'orthogonalité des résidus).}$$

(c) $r_k = r_{k-1} - \alpha_k^* \cdot A f_k$ [ça vient de $x_k = x_{k-1} + \alpha_k^* f_k$

$\underbrace{b - A x_k}_{r_k} = \underbrace{b - A x_{k-1}}_{r_{k-1}} - \alpha_k^* A f_k$]

La forme finale de l'algorithme du G.C :

Maintenant notre algorithme peut être écrit sous une forme **plus efficace** comme suit : (forme finale) :

$x_0 = 0$

$r_0 = b$

Pour $k = 1, \dots, n$ faire

Si $r_{k-1} = 0$ Alors Imprimer $x_{k-1} = x^*$ et sortir

Si $k = 1$ Alors $f_1 = r_0$

Sinon $f_k = r_{k-1} + \dfrac{r_{k-1} \cdot r_{k-1}}{r_{k-2} \cdot r_{k-2}} f_{k-1}$

$\alpha_k^* = \dfrac{r_{k-1} \cdot r_{k-1}}{f_k \cdot A f_k}$

$$x_k = x_{k-1} + \alpha_k^* f_k$$

$$r_k = r_{k-1} - \alpha_k^* A f_k$$

Fin k

Sortir x_k.

. Remarque : La seule matrice vecteur multiplication à faire c'est $A f_k$ (dans chaque itération) et on sauvegarde les vecteurs x_k, r_k, f_k, $A f_k$ et les scalaires $r_k \cdot r_k$, $r_{k-1} \cdot r_{k-1}$.

Convergence de la méthode G.C : On remarque au départ que la recherche du minimum de $J(\cdot)$ $\Leftrightarrow$ à la recherche du minimum de $E(\cdot)$ avec :

$$E(x) = \frac{1}{2}(x - x^*) \cdot A(x - x^*)$$

$$= \frac{1}{2} x \cdot Ax + \frac{1}{2} x^* \cdot Ax^* - x \cdot Ax^*$$

$$= J(x) + \underbrace{\frac{1}{2} x^* \cdot Ax^*}_{>0}$$

$E(\cdot)$ et $J(\cdot)$ atteignent leur minimum au même point. Puisque $Ax - Ax^* = Ax - b = -r(x)$

$x - x^* = -A^{-1} r(x)$

on peut écrire que :

$$E(x) = r(x) \cdot A^{-1} r(x).$$

Donc

$X_k = \sum_{j=1}^{k} \alpha_j^* f_j$

$$(x^* - x_k) \cdot A(x^* - x_k) = \min_{x \in P_k} \{(b - Ax) \cdot A^{-1}(b - Ax)\}$$

$$P_k = \left\{ \sum_{j=1}^{k} \gamma_j A^{j-1} b \,,\ (\gamma_j)_1^k \text{ qlq} \right\}$$

espace Krylov engendré par b

Puisque

$$f_1 = b = (r_0)$$

$$f_2 = r_1 + (\text{terms en } f_1) = -Ab + [b \text{ term}]$$

$$f_3 = r_2 + (\text{termes en } f_1, f_2) = A^2 b + (b, Ab, \ldots)$$

$\Rightarrow (f_1, \dots, f_k)$ et $(b, Ab, A^2b, \dots, A^{k-1}b)$ engendre le même sous espace P_k. (espace de Krylov engendré par b) Vect $= K_k$ sous espace vect. de $\mathbb{R}^n$. $(X_k \in X_0 + K_k)$.

$$(X^* - X_k) \cdot A(X^* - X_k) = \min_{X \in P_k} \{(b - AX) \cdot A^{-1}(b - AX)\}$$

$$= \min_{\gamma_1, \dots, \gamma_k} \left\{\left(b - \sum_{j=1}^{k} \gamma_j A^j b\right) \cdot A^{-1}\left(b - \sum_{j=1}^{k} \gamma_j A^j b\right)\right\}$$

$$= \min_{\gamma_1, \dots, \gamma_k} [P_k(A) b] \cdot A^{-1} [P_k(A) b]$$

avec $P_k(t) = 1 + \sum_{j=1}^{k} \gamma_j t^j$, $P_k(0) = 1$.

Si λ_j sont les valeurs propres de A, u_j les vecteurs propres correspondants (orthogonaux) et $b = \sum_{j=1}^{n} \eta_j u_j$

Alors

$$P_k(A) b = \sum_{j=1}^{n} P_k(\lambda_j) \eta_j u_j$$

$$A^{-1}[P_k(A) b] = \sum_{j=1}^{n} \frac{P_k(\lambda_j) \eta_j u_j}{\lambda_j}$$

et

$$\{P_k(A) b\} \cdot A^{-1} \{P_k(A) b\} = \sum_{j=1}^{n} \frac{P_k(\lambda_j)^2}{\lambda_j} \eta_j^2 .$$

Donc on a :

$$(X^* - X_k) \cdot A(X^* - X_k) = \min_{\gamma_i} \left\{\sum_{j=1}^{n} \frac{P_k(\lambda_j)^2}{\lambda_j} \eta_j^2\right\}$$

Le minimum est trouvé sur tous les polynômes $P_k(t) \in \mathcal{P}_k$ qui satisfait $P_k(0) = 1$. Un certain nbre de remarque peuvent être faite à cette étape, concernant la conv. de la Mde du G.C

a) Si A a seulement s valeurs propres distinctes $\lambda_1, \dots, \lambda_s$ $(s \leq n)$

On peut trouver un polynôme de degré s qui a les $(\lambda_i)_{i=1}^{s}$ comme racines et $=1$ en zéro i.e

$$P_s(t) = \frac{(\lambda_1 - t) \cdots (\lambda_s - t)}{\lambda_1 \lambda_2 \cdots \lambda_s}$$

tel que $(X^* - X_s).A(X^* - X_s) = 0$ et la solⁿ est trouvé après s itérations.

b) Si les valeurs propres de A sont $0 < \lambda_1 < \lambda_2 < \cdots < \lambda_n$. On sait que le polynôme $\{P_k(t)$ de degree k tq $P_k(0) = 1\}$ qui a le plus petit maximum sur $[\lambda_1, \lambda_n] \ni t$ est :

$\lambda_1 < t < \lambda_n$; $\mu = \dfrac{-(\lambda_1 + \lambda_n) + 2t}{\lambda_n - \lambda_1} \in [-1, 1]$

$$P_k^*(t) = \hat{T}_k(t) = \frac{T_k\left(\overbrace{\frac{+2t + (\lambda_n + \lambda_1)}{\lambda_n - \lambda_1}}^{\mu}\right)}{T_k\left(\frac{+\lambda_n + \lambda_1}{\lambda_n - \lambda_1}\right)}, \quad \hat{T}_k(0) = 1$$

et le max $P_k^*(t)$ sur $[\lambda_1, \lambda_n]$ et , $|T_k(\mu)| = |\cos(k \arccos(\mu))| \leq 1$

$$\leq \frac{1}{T_k\left(\frac{+\lambda_n + \lambda_1}{\lambda_n - \lambda_1}\right)}.$$

Donc
$$(X^* - X_k) \cdot A(X^* - X_k) \leq \frac{1}{\left\{T_k\left(\frac{\lambda_n + \lambda_1}{\lambda_n - \lambda_1}\right)\right\}^2} \sum_{j=1}^{n} \frac{\eta_j^2}{\lambda_j}$$

$$= \frac{1}{\left\{T_k\left(\frac{\lambda_n + \lambda_1}{\lambda_n - \lambda_1}\right)\right\}^2} b.A^{-1}b$$

$$\overset{(*)}{=} \frac{1}{\left\{T_k\left(\frac{\lambda_n + \lambda_1}{\lambda_n - \lambda_1}\right)\right\}^2} X^*.AX^*$$

Mais on a que $\lambda_1 X.X \leq X.AX \leq \lambda_n X.X \quad \forall X \in \mathbb{R}^N$

L'erreur relative s'écrit :

$$\frac{\|x^* - x_k\|_2}{\|x^*\|_2} \leq \frac{\sqrt{\lambda_n}}{\sqrt{\lambda_1}} \cdot \frac{1}{T_k\left(\frac{\lambda_n+\lambda_1}{\lambda_n-\lambda_1}\right)}$$

$$\left(\nu = \frac{\lambda_n+\lambda_1}{\lambda_n-\lambda_1} \right)$$

- On peut démontrer que $T_k(\nu) \geq \frac{\left(\nu + \sqrt{\nu^2-1}\right)^k}{2}$ $(|\nu| \geq 1)$ exercice

Et parce que $\underbrace{\frac{\lambda_n+\lambda_1}{\lambda_n-\lambda_1} + \sqrt{\left(\frac{\lambda_n+\lambda_1}{\lambda_n-\lambda_1}\right)^2 - 1}}_{\nu + \sqrt{\nu^2-1}} = \frac{\sqrt{\lambda_n}+\sqrt{\lambda_1}}{\sqrt{\lambda_n}-\sqrt{\lambda_1}} \geq 1$ (exercice), $T_k(\nu) \geq \frac{1}{2}\left(\nu+\sqrt{\nu^2-1}\right)^k$

$$\Rightarrow \frac{\|x^* - x_k\|_2}{\|x^*\|_2} \leq 2\frac{\sqrt{\lambda_n}}{\sqrt{\lambda_1}}\left(\frac{\sqrt{\lambda_n}-\sqrt{\lambda_1}}{\sqrt{\lambda_n}+\sqrt{\lambda_1}}\right)^k$$

$$= 2\sqrt{K(A)}\left(\frac{\sqrt{K(A)}-1}{\sqrt{K(A)}+1}\right)^k \quad \left(K(A) \overset{\text{Cond.}}{=} \frac{\lambda_n(A)}{\lambda_1(A)} \right)$$

erreur relative

<u>Remarques</u> ① Il n'est pas nécessaire de connaître les valeurs propres de A pour appliquer la méthode du Gradient (pas de paramètre à estimer).

② Il n'est pas nécessaire de commencer l'itération par $x_0 = (0, -, 0)$. Si une bonne approximation est connue on peut l'utiliser comme valeur initiale.

La méthode du Gradient conjugué préconditionnée :

Arithmetique inexacte $\Rightarrow$ Perte de la propriété d'A-conjugaison des f_j.

- <u>Conséquence</u> : ① Convergence en n itérations ne pourra être réaliser quand n est assez grande.

② Pour n assez grande, on espère que la méthode converge après k-itérations $k < n$. Donc pour remédier l'effet des erreurs

$$\boxed{M^{-1}Ax = M^{-1}b} \quad (*)$$

au lieu de $Ax=b$ avec M est SDP facilement inversible choisie de telle manière que la méthode du Gradient conjugué converge plus rapide avec arithmétique inexacte. M s'appelle "matrice de préconditionnement".

On considère la connaissance du factorisation de Cholesky de M :

$$M = U^T U \qquad (Mz = d \text{ est simple à résoudre})$$

- Pour appliquer GC à $(*)$, $M^{-1}A$ doit être symétrique. Donc au lieu d'appliquer la méthode du GC à $(*)$, on l'applique à :

$$\boxed{U^{-T} A U^{-1} y = c} \; ; \; c = U^{-T} b$$

(a) $U^{-T} A U^{-1}$ est symétrique.

(b) $U^{-T} A U^{-1}$ a les même valeurs propres que $M^{-1}A$.

(parce que $U^{-1}(U^{-T} A U^{-1}) U = M^{-1}A$).

- Si la solⁿ de ce système est y^* $\Rightarrow$ $\boxed{x^* = U^{-1} y^*}$ solⁿ de $Ax=b$.

Donc on applique l'algorithme du G-C au système :

$$\boxed{\hat{A} y = c} \qquad \hat{A} = U^{-T} A U^{-1}, \; c = U^{-T} b$$

- Choix de M : Parmi les choix très populaire de M c'est : SSOR d'Evans

~~Sur~~ relax
Sym. tri
$0 < \omega < 2$

$$M_{SSOR} = \frac{1}{\omega(2-\omega)} (D - \omega L) D^{-1} (D - \omega L^T)$$

Aucun calcul

- On constate que M est obtenue directement en fonction de A : ni aucun stockage n'est nécessaire.

$$M_{SSOR} = U^T U, \qquad \boxed{U^T = \frac{(D - \omega L) D^{-1/2}}{\sqrt{\omega(2-\omega)}}}$$

$L \cdot L^T$

Lemme : Les valeurs propres de $M_{SSOR}^{-1}A$ sont réelles et dans l'intervalle $]0,1]$.

Preuve : Posons $U^T = \left(\frac{D}{\omega} - L\right)\left(\frac{\omega}{2-\omega}\right)^{\frac{1}{2}} D^{-1/2}$. Alors $M = U^T U$

Posons $\tilde{D} = \frac{2-\omega}{\omega} D$ et $\tilde{L} = \frac{1-\omega}{\omega} D + L$.

Alors nous avons la relation : $A = \tilde{D} - \tilde{L} - \tilde{L}^T$

et $$M = \frac{1}{\omega(2-\omega)} (D - \omega L) D^{-1} (D - \omega L^T)$$

$\pm\left(\frac{1-\omega}{\omega}\right)D$
$$= \left(\frac{1}{\omega} D - L\right)\left(\frac{\omega}{2-\omega}\right) D^{-1} \left(\frac{1}{\omega} D - L^T\right)$$
$$= (\tilde{D} - \tilde{L})(\tilde{D}^{-1})(\tilde{D} - \tilde{L}^T)$$
$$= \tilde{D} - \tilde{L} - \tilde{L}^T + \tilde{L}\tilde{D}^{-1}\tilde{L}^T = A + \tilde{L}\tilde{D}^{-1}\tilde{L}^T \geq A > 0 \quad (*)$$

Soit (λ, u) un couple de valeur propre - vecteur propre de $M^{-1}A$, alors $M^{-1}Au = \lambda u$ donc

$$Au = \lambda M u$$
$$\Rightarrow (u, Au) = \lambda (u, Mu) \Rightarrow 0 < \lambda = \frac{(u, Au)}{(u, Mu)} \leq 1$$

d'après $(*)$.