

Chapitre 4 : MÉTHODES ITÉRATIVES POUR SYSTÈMES LINÉAIRES

1 Méthodes stationnaires

- Définition et contexte
- Méthodes stationnaires : principe
- Méthodes stationnaires : convergence
- Partitionnement de A
- Méthode de Jacobi
- Méthode de Jacobi : exemple
- Méthode de Gauss-Seidel
- Méthode de Gauss-Seidel : exemple

2 Méthodes non stationnaires

- Énergie du système
- Minimisation d'énergie
- Méthode du gradient
- Méthode du gradient conjugué

3 Contrôle de convergence et critères d'arrêt

- Contrôle de convergence
- Critères d'arrêt

4 Annexe

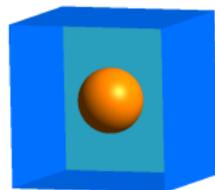
DÉFINITION ET CONTEXTE

PROBLÈME : On considère la résolution des systèmes réguliers (voir Chapitre 2).

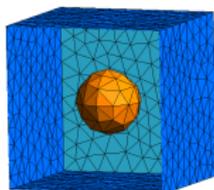
DÉFINITION : Une méthode est **itérative** si elle génère une séquence de solutions approchées du problème considéré (en l'occurrence un système régulier d'équations linéaires).

SCHÉMA : $\mathbf{x}^{(0)} \xrightarrow{\text{itération 1}} \mathbf{x}^{(1)} \xrightarrow{\text{itération 2}} \dots \xrightarrow{\text{itération } k} \mathbf{x}^{(k)} \xrightarrow{\text{itération } k+1}$

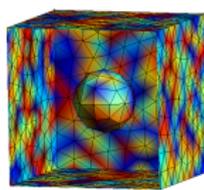
CONTEXTE : Souvent des systèmes linéaires sont des variantes discrètes de problèmes continus. Lorsqu'on approche un problème continu par un problème discret, on commet une **erreur de discrétisation** (souvent quantifiable) ; par conséquent on peut souvent se contenter d'une **solution approchée** du problème discret, pour autant que l'erreur ainsi commise reste inférieure à l'erreur de discrétisation.



discrétisation



résolution
& interpol.



problème continu
(ex. $-\Delta u = f$)

système d'équations
(ex. $\mathbf{Ax} = \mathbf{b}$)

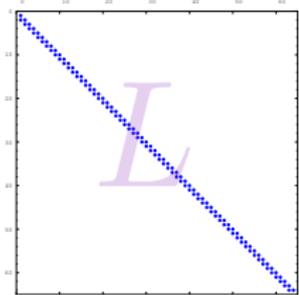
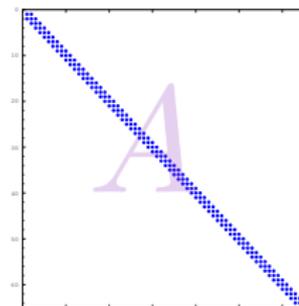
solution
(ex. \mathbf{x} & u)

CONTEXTE (SUITE)

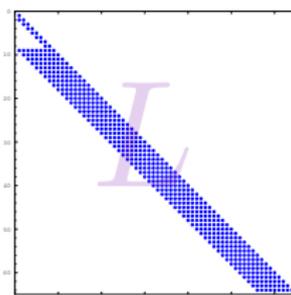
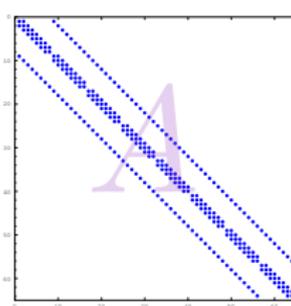
La matrice du système issue d'une application est souvent **creuse**, alors que sa factorisation LU peut être **sensiblement plus dense**, engendrant un surcoût important en mémoire et en temps de calcul. Ce surcoût est souvent évitable dans le cas des méthodes itératives.

EXEMPLE

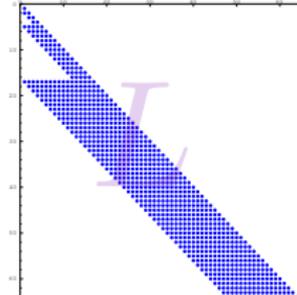
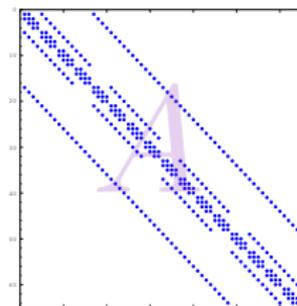
POISSON 1D



POISSON 2D



POISSON 3D



CONTEXTE (SUITE)

La matrice du système issue d'une application est souvent **creuse**, alors que sa factorisation LU peut être **sensiblement plus dense**, engendrant un surcoût important en mémoire et en temps de calcul. Ce surcoût est souvent évitable dans le cas des méthodes itératives.

EXEMPLE

POISSON 1D	POISSON 2D	POISSON 3D
	dimension de A	
4096	4096	4096
	MÉMOIRE	
	élém. non nuls A	
$1.2 \cdot 10^4$	$2.0 \cdot 10^4$	$2.7 \cdot 10^4$
	élém. non nuls L	
$8.1 \cdot 10^3$ ($\times 0.66$)	$2.6 \cdot 10^5$ ($\times 13$)	$1.0 \cdot 10^6$ ($\times 37$)
	TEMPS DE CALCUL (s)	
	lu(A)	
0.0015	0.062	0.26
— ¹	pcg(A, b, $1e-4$, 1000)	
	0.10 (131 itér.)	0.035 (36 itér.)

1. La méthode n'a pas convergé en 1000 itérations.

CONTEXTE (SUITE)

La matrice du système issue d'une application est souvent **creuse**, alors que sa factorisation LU peut être **sensiblement plus dense**, engendrant un surcoût important en mémoire et en temps de calcul. Ce surcoût est souvent évitable dans le cas des méthodes itératives.

EXEMPLE

POISSON 1D	POISSON 2D	POISSON 3D
	dimension de A	
15625	15625	15625
	MÉMOIRE	
	élém. non nuls A	
$4.7 \cdot 10^4$	$7.7 \cdot 10^4$	$1.1 \cdot 10^5$
	élém. non nuls L	
$3.1 \cdot 10^4$ ($\times 0.66$)	$1.9 \cdot 10^6$ ($\times 25$)	$9.4 \cdot 10^6$ ($\times 89$)
	TEMPS DE CALCUL (s)	
	lu(A)	
0.04	0.49	3.7
— ¹	pcg(A, b, 1e-4, 1000)	
	0.43 (234 itér.)	0.12 (56 itér.)

1. La méthode n'a pas convergé en 1000 itérations.

MÉTHODES STATIONNAIRES : PRINCIPE

Nombreuses méthodes itératives pour la résolution des systèmes linéaires

$$A\mathbf{x} = \mathbf{b} \quad (1)$$

exploitent le fait que la différence

$$\mathbf{e} = \mathbf{x} - \mathbf{x}^{(k)}$$

entre la solution exacte \mathbf{x} (qu'on ne connaît pas !) et la solution approchée $\mathbf{x}^{(k)}$ (à l'itération k) satisfait

$$A\mathbf{e} = \underbrace{A\mathbf{x}}_{\mathbf{b}} - A\mathbf{x}^{(k)}; \quad (2)$$

le membre de droite $\mathbf{r}^{(k)} = \mathbf{b} - A\mathbf{x}^{(k)}$ est appelé **résidu**.

Plusieurs méthodes itératives **stationnaires** utilisent une approximation

$$B \approx A$$

et résolvent

$$B\mathbf{e}^{(k)} = \mathbf{b} - A\mathbf{x}^{(k)}; \quad (3)$$

au lieu de (2) ; la correction $\mathbf{e}^{(k)}$ obtenue n'étant en général pas exacte, elle définit la nouvelle approximation

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{e}^{(k)}. \quad (4)$$

NOTE : L'approche peut être compétitive si (3) est plus facile à résoudre que (2) ; autrement dit, si «inverser» B est moins coûteux que «inverser» A .

MÉTHODES STATIONNAIRES : CONVERGENCE

Plusieurs méthodes itératives **stationnaires** pour la résolution de

$$A\mathbf{x} = \mathbf{b}$$

construisent donc une suite de solutions approchées $\mathbf{x}^{(k)}$ en résolvant

$$B\mathbf{e}^{(k)} = \mathbf{b} - A\mathbf{x}^{(k)}, \quad B \approx A, \quad (3)$$

suivi d'une correction

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{e}^{(k)}. \quad (4)$$

Les propriétés de convergence découlent de l'observation suivante

$$\begin{aligned} \mathbf{x} - \mathbf{x}^{(k+1)} &= (\mathbf{x} - \mathbf{x}^{(k)}) - \underbrace{(\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)})}_{\mathbf{e}^{(k)}} \\ &= (\mathbf{x} - \mathbf{x}^{(k)}) - B^{-1}(\mathbf{b} - A\mathbf{x}^{(k)}) \quad (\text{grâce à (3)}) \\ &= (\mathbf{x} - \mathbf{x}^{(k)}) - B^{-1}A(\mathbf{x} - \mathbf{x}^{(k)}) \quad (A\mathbf{x} = \mathbf{b}) \\ &= (I - B^{-1}A)(\mathbf{x} - \mathbf{x}^{(k)}); \end{aligned}$$

la matrice

$$E = I - B^{-1}A$$

est appelée la **matrice d'itération**.

MÉTHODES STATIONNAIRES : CONVERGENCE (SUITE)

L'écart entre la solution exacte \mathbf{x} et la solution approchée $\mathbf{x}^{(k)}$ à l'itération k satisfait donc

$$\mathbf{x} - \mathbf{x}^{(k+1)} = E(\mathbf{x} - \mathbf{x}^{(k)}), \quad (5)$$

avec $E = I - B^{-1}A$.

THÉORÈME 1 : Soit E une matrice $n \times n$. Si toutes les valeurs propres λ_i de E satisfont $|\lambda_i| < 1$, $i = 1, \dots, n$, alors la suite définie par

$$\mathbf{w}^{(k+1)} = E \mathbf{w}^{(k)}$$

converge vers $\mathbf{0}$ pour tout choix de $\mathbf{w}^{(0)}$.

DÉMONSTRATION. Nous nous limiterons au cas où la matrice E est diagonalisable. Dans ce cas, il existe une matrice P telle que

$$P^{-1}EP = \text{diag}(\lambda_i) =: \Lambda.$$

On a alors (quel que soit le choix pour $\|\cdot\|$)

$$\|\mathbf{w}^{(k)}\| = \|E^k \mathbf{w}^{(0)}\| = \|P \Lambda^k P^{-1} \mathbf{w}^{(0)}\| \leq \underbrace{\|P\| \|P^{-1}\|}_{\text{const}} \|\mathbf{w}^{(0)}\| \underbrace{\left(\max_i |\lambda_i| \right)^k}_{\rightarrow 0}.$$

PARTITIONNEMENT DE A

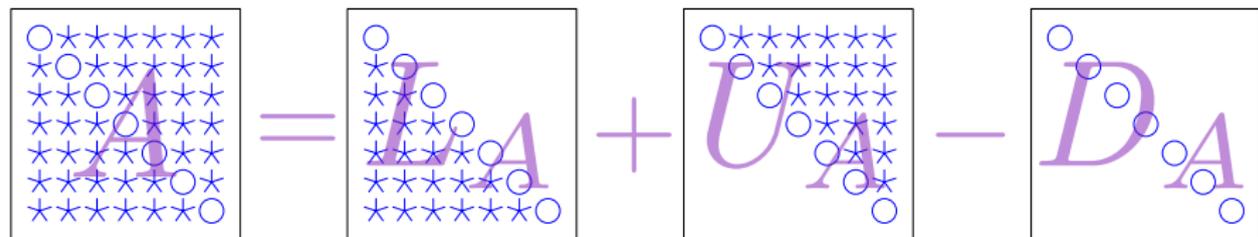
Pour une matrice $A = (a_{ij})$ donnée on dénote par :

- D_A la matrice diagonale avec $(D_A)_{ii} = a_{ii}$
(commande `diag(diag(A))`);
- L_A la matrice triangulaire inférieure avec $(L_A)_{ij} = a_{ij}$ pour tout $i \geq j$
(commande `tril(A)`);
- U_A la matrice triangulaire supérieure avec $(U_A)_{ij} = a_{ij}$ pour tout $i \leq j$
(commande `triu(A)`);

En particulier

$$A = L_A + U_A - D_A.$$

SCHEMA :



MÉTHODE DE JACOBI

Plusieurs méthodes itératives **stationnaires** pour la résolution de

$$A\mathbf{x} = \mathbf{b}$$

construisent une séquence d'approximations $\mathbf{x}^{(k)}$ en résolvant

$$B\mathbf{e}^{(k)} = (\mathbf{b} - A\mathbf{x}^{(k)}), \quad B \approx A, \quad (3)$$

suivi d'une correction

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{e}^{(k)}. \quad (4)$$

La méthode de Jacobi requiert $a_{ii} \neq 0$, $i = 1, \dots, n$, et utilise

$$B_J = D_A,$$

ce qui mène à la suite des solutions approchées

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + D_A^{-1}(\mathbf{b} - A\mathbf{x}^{(k)}).$$

ou encore (comme les éléments diagonaux de $D_A^{-1}A$ valent 1) :

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1, j \neq i}^n a_{ij} x_j^{(k)} \right), \quad i = 1, \dots, n.$$

CONVERGENCE : D'après Théorème 1, la suite $\mathbf{x}^{(k)}$ converge vers la solution si

$$\max_i |\lambda_i(I - D_A^{-1}A)| < 1.$$

MÉTHODE DE JACOBI : EXEMPLE

EXEMPLE : (système de l'Exemple 1 dans la Motivation)

$$\begin{cases} 2x_1 & -x_2 & & = & 1 \\ -x_1 & +2x_2 & -x_3 & = & 1 \\ & -x_2 & +x_3 & = & 1 \end{cases} \quad \text{avec } \mathbf{x}^{(0)} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Une itération de la méthode de Jacobi :

$$\begin{cases} x_1^{(k+1)} & = & \frac{1}{2}(1 + x_2^{(k)}) \\ x_2^{(k+1)} & = & \frac{1}{2}(1 + x_1^{(k)} + x_3^{(k)}) \\ x_3^{(k+1)} & = & 1 + x_2^{(k)} \end{cases}$$

et donc

$$\underbrace{\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}}_{\mathbf{x}^{(0)}} \Rightarrow \underbrace{\begin{pmatrix} 0.5 \\ 0.5 \\ 1 \end{pmatrix}}_{\mathbf{x}^{(1)}} \Rightarrow \underbrace{\begin{pmatrix} 0.75 \\ 1.25 \\ 1.5 \end{pmatrix}}_{\mathbf{x}^{(2)}} \Rightarrow \dots \Rightarrow \underbrace{\begin{pmatrix} 2.95\dots \\ 4.93\dots \\ 5.91\dots \end{pmatrix}}_{\mathbf{x}^{(30)}}$$

la vraie solution est $\mathbf{x} = \begin{pmatrix} 3 \\ 5 \\ 6 \end{pmatrix}$.

MÉTHODE DE GAUSS-SEIDEL

Plusieurs méthodes itératives **stationnaires** pour la résolution de

$$A\mathbf{x} = \mathbf{b}$$

construisent une séquence d'approximations $\mathbf{x}^{(k)}$ en résolvant

$$B\mathbf{e}^{(k)} = (\mathbf{b} - A\mathbf{x}^{(k)}), \quad B \approx A, \quad (5)$$

suivi d'une correction

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{e}^{(k)}. \quad (6)$$

La méthode de Gauss-Seidel requiert $a_{ii} \neq 0$, $i = 1, \dots, n$, et utilise comme approximation

$$B_{GS} = L_A,$$

ce qui mène à la suite des solutions approchées

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + L_A^{-1}(\mathbf{b} - A\mathbf{x}^{(k)}).$$

ou encore :

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j < i} a_{ij} x_j^{(k+1)} - \sum_{j > i} a_{ij} x_j^{(k)} \right), \quad i = 1, \dots, n.$$

CONVERGENCE (sans démonstration) : la suite $\mathbf{x}^{(k)}$ converge vers la solution \mathbf{x} du système si A est symétrique définie positive.

MÉTHODE DE GAUSS-SEIDEL : EXEMPLE

EXEMPLE : (système de l'Exemple 1 dans la Motivation)

$$\begin{cases} 2x_1 & -x_2 & & = 1 \\ -x_1 & +2x_2 & -x_3 & = 1 \\ & -x_2 & +x_3 & = 1 \end{cases} \text{ avec } \mathbf{x}^{(0)} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

une itération de la méthode de Gauss-Seidel :

$$\begin{cases} x_1^{(k+1)} & = \frac{1}{2}(1 + x_2^{(k)}) \\ x_2^{(k+1)} & = \frac{1}{2}(1 + x_1^{(k+1)} + x_3^{(k)}) \\ x_3^{(k+1)} & = 1 + x_2^{(k+1)} \end{cases}$$

et donc

$$\underbrace{\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}}_{\mathbf{x}^{(0)}} \Rightarrow \underbrace{\begin{pmatrix} 0.5 \\ 0.75 \\ 1.75 \end{pmatrix}}_{\mathbf{x}^{(1)}} \Rightarrow \underbrace{\begin{pmatrix} 0.875 \\ 1.81\dots \\ 2.81\dots \end{pmatrix}}_{\mathbf{x}^{(2)}} \Rightarrow \dots \Rightarrow \underbrace{\begin{pmatrix} 2.94 \\ 4.92 \\ 5.92 \end{pmatrix}}_{\mathbf{x}^{(15)}}$$

la vraie solution est $\mathbf{x} = \begin{pmatrix} 3 \\ 5 \\ 6 \end{pmatrix}$.

ENERGIE DU SYSTÈME

Pour les (trois) algorithmes suivants, on suppose que A est **symétrique définie positive**; on peut alors définir la **norme énergie** comme

$$\|\mathbf{v}\|_A = \sqrt{\mathbf{v}^T A \mathbf{v}}.$$

Il s'agit bien d'une norme (elle respecte en particulier l'inégalité triangulaire) même si on n'utilisera que son caractère positif :

$$\|\mathbf{v}\|_A > 0 \quad \forall \mathbf{v} \neq \mathbf{0}.$$

Pour une approximation donnée $\mathbf{x}^{(k)}$ de la solution, l'«énergie» de la différence avec la solution exacte \mathbf{x} est

$$f(\mathbf{x}^{(k)}) := \|\mathbf{x} - \mathbf{x}^{(k)}\|_A^2 = \mathbf{x}^{(k)T} A \mathbf{x}^{(k)} - 2 \mathbf{x}^{(k)T} \underbrace{A \mathbf{x}}_{\mathbf{b}} + \underbrace{\mathbf{x}^T A \mathbf{x}}_{\text{const}}.$$

En particulier

$$f(\mathbf{x}^{(k)}) \geq 0 \quad \text{et} \quad f(\mathbf{x}^{(k)}) = 0 \quad \Leftrightarrow \quad \mathbf{x}^{(k)} = \mathbf{x}.$$

Une manière de s'assurer que la suite $\mathbf{x}^{(k)}$, $k = 1, \dots$, se rapproche de la solution exacte est donc de choisir $\mathbf{x}^{(k+1)}$ tel que $f(\mathbf{x}^{(k)}) > f(\mathbf{x}^{(k+1)})$.

MINIMISATION D'ÉNERGIE

Les méthodes basées sur la minimisation d'énergie construisent la suite des solutions approchées $\mathbf{x}^{(k)}$, $k = 1, \dots$, tout en minimisant leur «énergie»

$$f(\mathbf{x}^{(k)}) = \mathbf{x}^{(k)T} A \mathbf{x}^{(k)} - 2 \mathbf{x}^{(k)T} \mathbf{b} + \text{const.}$$

PRINCIPE : Supposons qu'on ait choisi une direction $\mathbf{p} \neq \mathbf{0}$ et que la forme de la solution approchée à l'itération $k + 1$ est

$$\mathbf{x}^{(k+1)}(\alpha) = \mathbf{x}^{(k)} + \alpha \mathbf{p}, \quad \alpha \in \mathbb{R}.$$

Quelle valeur de α minimise $f(\mathbf{x}^{(k+1)})$?

Comme

$$f(\mathbf{x}^{(k+1)}(\alpha)) = (\mathbf{x}^{(k)} + \alpha \mathbf{p})^T A (\mathbf{x}^{(k)} + \alpha \mathbf{p}) - 2(\mathbf{x}^{(k)} + \alpha \mathbf{p})^T \mathbf{b} + \text{const};$$

la valeur de α au point critique satisfait

$$\frac{1}{2} \frac{df(\mathbf{x}^{(k+1)}(\alpha))}{d\alpha} = \mathbf{p}^T A (\mathbf{x}^{(k)} + \alpha \mathbf{p}) - \mathbf{p}^T \mathbf{b} = \alpha \mathbf{p}^T A \mathbf{p} - \mathbf{p}^T \underbrace{(\mathbf{b} - A \mathbf{x}^{(k)})}_{\mathbf{r}^{(k)}} = 0,$$

et donc

$$\alpha = \frac{\mathbf{p}^T \mathbf{r}^{(k)}}{\mathbf{p}^T A \mathbf{p}}.$$

MINIMISATION D'ÉNERGIE (SUITE)

Pour une solution approchée $\mathbf{x}^{(k)}$ et une direction \mathbf{p} données, la valeur de α qui minimise l'énergie $f(\mathbf{x}^{(k+1)})$ de

$$\mathbf{x}^{(k+1)}(\alpha) = \mathbf{x}^{(k)} + \alpha\mathbf{p}$$

est donc

$$\alpha = \frac{\mathbf{p}^T \mathbf{r}^{(k)}}{\mathbf{p}^T A \mathbf{p}}.$$

On notera en particulier que, comme

$$\mathbf{x}^{(k+1)}(0) = \mathbf{x}^{(k)},$$

on a toujours, avec $\mathbf{x}^{(k+1)}$ correspondant à α optimal,

$$f(\mathbf{x}^{(k+1)}) \leq f(\mathbf{x}^{(k)});$$

l'erreur $\mathbf{x} - \mathbf{x}^{(k)}$ sur la solution est donc toujours décroissante avec k (en norme énergie) quel que soit le choix de \mathbf{p} (cela ne garantit pas pour autant que $\mathbf{x}^{(k)} \rightarrow \mathbf{x}$ pour $k \rightarrow \infty$, pourquoi?).

MINIMISATION D'ÉNERGIE (SUITE)

Pour une solution approchée $\mathbf{x}^{(k)}$ et une direction \mathbf{p} données, la valeur de α qui minimise l'énergie $f(\mathbf{x}^{(k+1)})$ de

$$\mathbf{x}^{(k+1)}(\alpha) = \mathbf{x}^{(k)} + \alpha \mathbf{p}$$

est donc

$$\alpha = \frac{\mathbf{p}^T \mathbf{r}^{(k)}}{\mathbf{p}^T A \mathbf{p}}.$$

ALGORITHME (MINIMISATION D'ÉNERGIE) :

$\mathbf{r}^{(0)} = \mathbf{b} - A\mathbf{x}^{(0)}$ avec $\mathbf{x}^{(0)}$ donné

répéter jusqu'à l'arrêt

choisir $\mathbf{p}^{(k)} \neq \mathbf{0}$

calculer $\alpha_k = \frac{\mathbf{p}^{(k)T} \mathbf{r}^{(k)}}{\mathbf{p}^{(k)T} A \mathbf{p}^{(k)}}$

$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k \mathbf{p}^{(k)}$

$\mathbf{r}^{(k+1)} = \mathbf{r}^{(k)} - \alpha_k A \mathbf{p}^{(k)}$

NOTE : En pratique, la multiplication $A\mathbf{p}^{(k)}$ ne se fait qu'une fois par itération

COÛT : 1 MatVec

(c'est-à-dire une multiplication de A par un vecteur ; ici c'est l'opération la plus coûteuse ($\approx n^2$ flops avec symétrie ; moins pour matrices creuses)).

MÉTHODE DU GRADIENT

Pour la méthode du gradient on choisit \mathbf{p} égale à la direction de la plus grande pente de f au point $\mathbf{x}^{(k)}$. Intuitivement, cela permet d'espérer une réduction significative d'énergie pour $f(\mathbf{x}^{(k+1)})$.

La direction \mathbf{p} est donc donnée par :

$$\mathbf{p} = \nabla f(\mathbf{x}^{(k)}) = 2(A\mathbf{x}^{(k)} - \mathbf{b}) = -2\mathbf{r}.$$

Notons que le facteur devant \mathbf{p} importe peu, il est corrigé par α_k .

ALGORITHME (DU GRADIENT) :

$$\mathbf{r}^{(0)} = \mathbf{b} - A\mathbf{x}^{(0)}$$

répéter jusqu'à l'arrêt

$$\mathbf{p}^{(k)} = \mathbf{r}^{(k)}$$

$$\text{calculer } \alpha_k = \frac{\mathbf{p}^{(k)T} \mathbf{r}^{(k)}}{\mathbf{p}^{(k)T} A \mathbf{p}^{(k)}}$$

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k \mathbf{p}^{(k)}$$

$$\mathbf{r}^{(k+1)} = \mathbf{r}^{(k)} - \alpha_k A \mathbf{p}^{(k)}$$

NOTE : Comme avant, la multiplication $A\mathbf{p}^{(k)}$ ne se fait qu'une fois par itération ;

COÛT : 1 MatVec

DIVISION PAR 0 : le dénominateur de l'expression pour α_k est nul si $\|\mathbf{p}^{(k)}\|_A = 0$ et donc $\mathbf{r}^{(k)} = \mathbf{0}$; mais alors $\mathbf{x}^{(k)}$ est la solution exacte !

MÉTHODE DU GRADIENT CONJUGUÉ

Il s'agit d'une évolution de la méthode du gradient où on enrichit la direction courante $\mathbf{p}^{(k)}$ avec la direction $\mathbf{p}^{(k-1)}$ précédente

$$\mathbf{p}^{(k)} = \mathbf{r}^{(k)} - \beta \mathbf{p}^{(k-1)} .$$

On peut montrer (voir l'annexe) que l'énergie de la nouvelle solution approchée est minimisée par rapport à β si

$$\mathbf{p}^{(k-1)T} A \mathbf{p}^{(k)} = 0 ,$$

et donc si

$$\beta = \frac{\mathbf{p}^{(k-1)T} A \mathbf{r}^{(k)}}{\mathbf{p}^{(k-1)T} A \mathbf{p}^{(k-1)}} .$$

ALGORITHME (DU GRADIENT CONJUGUÉ) :

$$\mathbf{r}^{(0)} = \mathbf{b} - A \mathbf{x}^{(0)} ; \quad \mathbf{p}^{(0)} = \mathbf{r}^{(0)}$$

répéter jusqu'à l'arrêt

$$\text{calculer } \alpha_k = \frac{\mathbf{p}^{(k)T} \mathbf{r}^{(k)}}{\mathbf{p}^{(k)T} A \mathbf{p}^{(k)}}$$

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k \mathbf{p}^{(k)}$$

$$\mathbf{r}^{(k+1)} = \mathbf{r}^{(k)} - \alpha_k A \mathbf{p}^{(k)}$$

$$\text{calculer } \beta_k = \frac{(A \mathbf{p}^{(k)})^T \mathbf{r}^{(k+1)}}{\mathbf{p}^{(k)T} A \mathbf{p}^{(k)}}$$

$$\mathbf{p}^{(k+1)} = \mathbf{r}^{(k+1)} - \beta_k \mathbf{p}^{(k)}$$

NOTE : Comme avant, la multiplication $A \mathbf{p}^{(k)}$ ne se fait qu'une fois par itération ;

COÛT : 1 MatVec

DIVISION PAR 0 : (sans démonstration) le dénominateur de l'expression pour α_k (ou β_k) est nul si et seulement si $\mathbf{x}^{(k)}$ est la solution exacte.

GRADIENT CONJUGUÉ : COMMENTAIRES

- La méthode du gradient conjugué est une méthode itérative de référence pour les systèmes dont la matrice A est symétrique définie positive.
- Plusieurs variantes car plusieurs manières de calculer les coefficients α_k , β_k (mathématiquement équivalentes mais numériquement différentes).
- La méthode est surtout efficace si on dispose d'un **préconditionneur** ; une matrice symétrique définie positive $B \approx A$ avec B plus facile à «inverser» que A ;
ex : $B = D_A$ (Jacobi), $B = U_A D_A^{-1} L_A$ (Gauss-Seidel symétrique).
- l'instruction **Octave** correspondante : `pcg`.
- pour l'Exemple 1 de la Motivation la méthode converge avec $\mathbf{x}_0 = \mathbf{0}$ en 3 itérations !

CONTRÔLE DE CONVERGENCE

Comme déjà indiqué, on peut souvent s'arrêter quand une certaine précision (disons absolue) sur la solution (avec la norme $\|\cdot\|$ de votre choix)

$$\|\mathbf{x}^{(k)} - \mathbf{x}\|$$

est atteinte. Comme la solution exacte \mathbf{x} n'est pas connue, on l'évalue indirectement via la norme du résidu, soit

$$\|\mathbf{x} - \mathbf{x}^{(k)}\| = \|A^{-1}\mathbf{r}^{(k)}\|.$$

Or, comme

$$\|A^{-1}\mathbf{r}^{(k)}\| \leq \|A^{-1}\| \|\mathbf{r}^{(k)}\| \quad \text{et} \quad \|\mathbf{r}^{(k)}\| = \|AA^{-1}\mathbf{r}^{(k)}\| \leq \|A\| \|A^{-1}\mathbf{r}^{(k)}\|,$$

on obtient

$$\|A\|^{-1} \|\mathbf{r}^{(k)}\| \leq \|\mathbf{x}^{(k)} - \mathbf{x}\| \leq \|A^{-1}\| \|\mathbf{r}^{(k)}\|.$$

Cela nous permet de relier le résidu à l'erreur absolue sur la solution.

Pour ce qui est de l'erreur relative, une relation similaire existe (généralement plus pessimiste) :

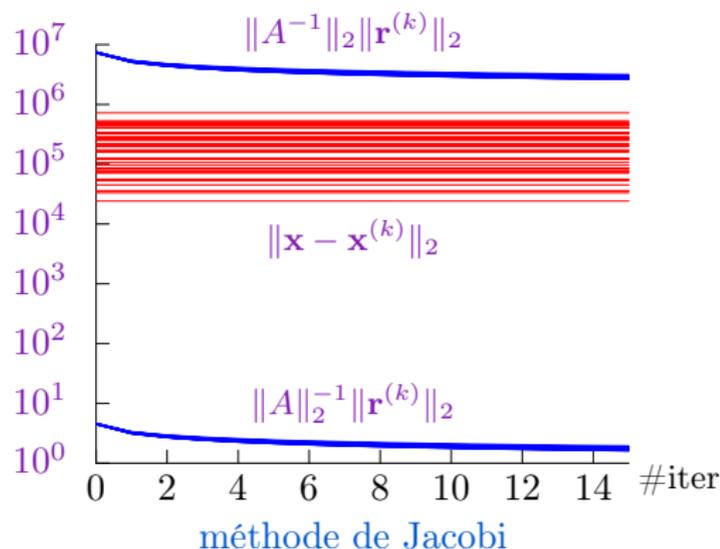
$$\kappa(A)^{-1} \frac{\|\mathbf{r}^{(k)}\|}{\|\mathbf{b}\|} \leq \frac{\|\mathbf{x}^{(k)} - \mathbf{x}\|}{\|\mathbf{x}\|} \leq \kappa(A) \frac{\|\mathbf{r}^{(k)}\|}{\|\mathbf{b}\|}.$$

CONTRÔLE DE CONVERGENCE (SUITE)

$$\|A\|^{-1}\|\mathbf{r}^{(k)}\| \leq \|\mathbf{x}^{(k)} - \mathbf{x}\| \leq \|A^{-1}\|\|\mathbf{r}^{(k)}\|.$$

EXEMPLE : (système de l'Exemple 4 dans la Motivation, 1000 ressorts avec la même constante de rappel)

- 50 membres de droite \mathbf{b} générés de manière aléatoire (via $\mathbf{b} = \text{ones}(1000,1) - 2*\text{rand}(1000,1)$);
- solution exacte \mathbf{x} calculée avec $A \setminus \mathbf{b}$.



CRITÈRES D'ARRÊT

Les quelques critères communément utilisés pour les méthodes itératives.
On arrête la méthode si

- ⊕ la norme du résidu $\mathbf{r}^{(k)} = \mathbf{b} - A\mathbf{x}^{(k)}$ en valeur absolue satisfait

$$\|\mathbf{r}^{(k)}\| \leq \varepsilon_a ;$$

alternativement, on peut utiliser la réduction de cette norme **relativement** à celle du second membre \mathbf{b} :

$$\|\mathbf{r}^{(k)}\| \leq \varepsilon_r \|\mathbf{b}\| ;$$

ε_a , ε_r sont typiquement fournis par l'utilisateur ;

▶ **bonne pratique** : les valeurs de ε_r entre 10^{-5} et 10^{-10} ;

▶ **mauvaise pratique** : attention à l'utilisation du critère $\|\mathbf{r}^{(k)}\| \leq \varepsilon_r \|\mathbf{r}^{(0)}\|$
car $\|\mathbf{r}^{(0)}\|$ dépend du choix de l'approximation initiale $\mathbf{x}^{(0)}$;

⊖ le nombre maximal d'itérations est dépassé ;

⊖ la décroissance du résidu est lente ou inexistante ;

ex : quand la convergence ne peut que se détériorer.

Les deux derniers critères sont utilisés pour arrêter
une méthode qui converge trop lentement !

ANNEXE

On considère la minimisation de

$$f(\mathbf{x}^{(k+1)}) = \mathbf{x}^{(k+1)T} A \mathbf{x}^{(k+1)} - 2 \mathbf{x}^{(k+1)T} \mathbf{b} + \text{const}$$

avec

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k \mathbf{p}^{(k)} \quad (10)$$

et

$$\mathbf{p}^{(k)} = \mathbf{r}^{(k)} - \beta_{k-1} \mathbf{p}^{(k-1)}, \quad (11)$$

sachant que

$$\mathbf{r}^{(k+1)} = \mathbf{r}^{(k)} - \alpha_k A \mathbf{p}^{(k)} \quad (12)$$

et $\mathbf{p}^{(0)} = \mathbf{r}^{(0)}$. Le raisonnement général sur la minimisation d'énergie de la page 15 reste valable et donc

$$\alpha_k = \frac{\mathbf{p}^{(k)T} \mathbf{r}^{(k)}}{\mathbf{p}^{(k)T} A \mathbf{p}^{(k)}}.$$

Par ailleurs, ce choix de α_k implique (via (12))

$$\mathbf{p}^{(k)T} \mathbf{r}^{(k+1)} = 0, \quad (13)$$

les deux vecteurs sont donc orthogonaux.

ANNEXE (SUITE)

Par ailleurs on vérifie que

$$\begin{aligned}\frac{df(\mathbf{x}^{(k+1)})}{d\beta_{k-1}} &= -2\alpha_k \left(\mathbf{p}^{(k-1)T} A\mathbf{x}^{(k+1)} - \mathbf{p}^{(k-1)T} \mathbf{b} \right) \\ &= -2\alpha_k \left(-\mathbf{p}^{(k-1)T} \mathbf{r}^{(k)} + \alpha_k \mathbf{p}^{(k-1)T} A\mathbf{p}^{(k)} \right) \quad (\text{via (11), (10)}).\end{aligned}$$

Le premier terme s'annule en appliquant l'égalité (13) obtenue pour l'itération précédente. L'annulation de la dérivée par rapport à β_{k-1} implique que soit $\alpha_k = 0$ (on peut montrer que cela n'arrive que si $\mathbf{r}^{(k)} = \mathbf{0}$, et donc, si $\mathbf{x}^{(k)}$ est la solution exacte!), soit

$$\mathbf{p}^{(k-1)T} A\mathbf{p}^{(k)} = \mathbf{0},$$

ce qui est bien la condition recherchée.