

# ANALYSE NUMÉRIQUE SV

## SYSTÈMES LINÉAIRES

Simone Deparis

EPFL Lausanne – MATH

Printemps 2022



# FORMULATION DU PROBLÈME

On appelle **système linéaire d'ordre  $n$**  ( $n$  entier positif), une expression de la forme

$$Ax = b,$$

où  $A = (a_{ij})$  est **une matrice** de taille  $n \times n$  **donnée**,  $b = (b_j)$  est un **vecteur** colonne également **donné** et  $x = (x_j)$  est le **vecteur des inconnues** du système. La relation précédente équivaut aux  $n$  équations

$$\sum_{j=1}^n a_{ij}x_j = b_i, \quad i = 1, \dots, n.$$

La matrice  $A$  est dite **régulière (non singulière)** si  $\det(A) \neq 0$ . On a **l'existence et l'unicité** de la solution  $x$  (pour n'importe quel vecteur  $b$  donné) si et seulement si la matrice associée au système linéaire est régulière.

- Méthodes itératives : définitions
- Méthode de Richardson
- Méthodes de Jacobi et de Gauss-Seidel  
Critères de convergence
- Méthodes du Gradient et du Gradient Conjugués  
Critères de convergence

# MÉTHODES ITÉRATIVES

Résoudre un système linéaire  $Ax = b$  par une méthode itérative consiste à construire une suite de vecteurs  $x^{(k)}$ ,  $k \geq 0$ , de  $\mathbb{R}^n$  qui converge vers la solution exacte  $x$ , c'est-à-dire :

$$\lim_{k \rightarrow \infty} x^{(k)} = x$$

pour n'importe quelle donnée initiale  $x^{(0)} \in \mathbb{R}^n$ .

On peut considérer la relation de récurrence suivante :

$$x^{(k+1)} = Bx^{(k)} + g, \quad k \geq 0 \tag{1}$$

où  $B$  est une matrice bien choisie (dépendante de  $A$ ) et  $g$  est un vecteur (dépendant de  $A$  et de  $b$ ), qui vérifient la relation (de consistance)

$$x = Bx + g. \tag{2}$$

Étant donné que  $x = A^{-1}b$ , on obtient  $g = (I - B)A^{-1}b$ ; la méthode itérative est donc complètement définie par la matrice  $B$  qui est appelée *matrice d'itération*. En définissant l'erreur à l'itération  $k$  comme

$$e^{(k)} = x - x^{(k)},$$

on obtient la relation de récurrence :

$$e^{(k+1)} = Be^{(k)} \quad \text{et donc} \quad e^{(k+1)} = B^{k+1}e^{(0)}, \quad k = 0, 1, \dots$$

On peut montrer que  $\lim_{k \rightarrow \infty} e^{(k)} = 0$  pour tout  $e^{(0)}$  (et donc pour tout  $x^{(0)}$ ) si et seulement si

$$\rho(B) < 1,$$

où  $\rho(B)$  est le *rayon spectral* de la matrice  $B$ , défini par

$$\rho(B) = \max |\lambda_i(B)|$$

et  $\lambda_i(B)$  sont les valeurs propres de la matrice  $B$ .

Plus la valeur de  $\rho(B)$  est petite, moins il est nécessaire d'effectuer d'itérations pour réduire l'erreur initiale d'un facteur donné. Plus précisément :

$$\|e^{(k+1)}\| \leq \rho(B) \|e^{(k)}\| \quad \text{et donc} \quad \|e^{(k+1)}\| \leq \rho(B)^{k+1} \|e^{(0)}\|, \quad k = 0, 1, \dots$$

# CONSTRUCTION D'UNE MÉTHODE ITÉRATIVE

Une approche générale pour construire une méthode itérative est basée sur la décomposition de la matrice  $A$  :

$$A = P - (P - A)$$

où  $P$  est une matrice inversible appelée *préconditionneur* de  $A$ .  
 Alors,

$$Ax = b \quad \Leftrightarrow \quad Px = (P - A)x + b$$

qui est de la forme (2) en posant

$$B = P^{-1}(P - A) = I - P^{-1}A \quad \text{et} \quad g = P^{-1}b.$$

On peut définir la méthode itérative correspondante

$$P(x^{(k+1)} - x^{(k)}) = r^{(k)} \quad k \geq 0$$

où  $r^{(k)}$  désigne le *résidu* à l'itération  $k$  :  $r^{(k)} = b - Ax^{(k)}$

On peut généraliser cette méthode de la manière suivante :

$$P(x^{(k+1)} - x^{(k)}) = \alpha_k r^{(k)} \quad k \geq 0 \quad (3)$$

où  $\alpha_k \neq 0$  est un paramètre pour améliorer la convergence de la suite  $x^{(k)}$ .

La méthode (3) est appelée *méthode de Richardson*.

La matrice  $P$  doit être choisie de telle manière que le coût de la résolution de (3) soit assez faible. Par exemple, une matrice  $P$  diagonale ou triangulaire vérifierait ce critère.



# LA MÉTHODE DE JACOBI

Si les éléments diagonaux de  $A$  sont non nuls, on peut poser

$$P = D = \text{diag}(a_{11}, a_{22}, \dots, a_{nn})$$

$D$  étant la partie diagonale de  $A$  :

$$D_{ij} = \begin{cases} 0 & \text{si } i \neq j \\ a_{ij} & \text{si } i = j. \end{cases}$$

La méthode de Jacobi correspond à ce choix avec  $\alpha_k = 1$  pour tout  $k$ .  
 On déduit alors :

$$Dx^{(k+1)} = b - (A - D)x^{(k)} \quad k \geq 0.$$

Par composantes :

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left( b_i - \sum_{j=1, j \neq i}^n a_{ij} x_j^{(k)} \right), \quad i = 1, \dots, n. \quad (4)$$

La méthode de Jacobi peut s'écrire sous la forme générale

$$x^{(k+1)} = Bx^{(k)} + g,$$

avec

$$B = B_J = D^{-1}(D - A) = I - D^{-1}A, \quad g = g_J = D^{-1}b.$$

# LA MÉTHODE DE GAUSS-SEIDEL

Cette méthode est définie par la formule suivante :

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

Cette méthode correspond à (1) avec  $\alpha_k = 1$  ( $\forall k \geq 0$ ) et  $P_{GS}$  la matrice triangulaire inférieure

$$\begin{cases} P_{ij} = a_{ij} & \text{si } i \geq j \\ P_{ij} = 0 & \text{si } i < j \end{cases}$$

(partie triangulaire inférieure de  $A$  avec la diagonale).

On peut écrire cette méthode sous la forme (3), avec la matrice d'itération  $B = B_{GS}$  donnée par

$$B_{GS} = P_{GS}^{-1}(P_{GS} - A)$$

et

$$g_{GS} = P_{GS}^{-1}b.$$

## CRITÈRES DE CONVERGENCE

On a la relation suivante :

Si  $A$  est une matrice *symétrique définie positive*, alors

$$\frac{\|x^{(k)} - x\|}{\|x\|} \leq K(A) \frac{\|r^{(k)}\|}{\|b\|}. \quad (5)$$

L'erreur relative à la  $k$ -ième itération peut être majorée par le résidu relatif multiplié par le conditionnement de  $A$ .

En particulier, si  $K(A) \approx 1$ , une petite valeur de la norme du résidu correspond à une petite valeur de la norme de l'erreur ; si  $K(A) \gg 1$ , cette relation peut être fausse.

On a également une estimation (utilisée si  $P \neq I$ ) :

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

# CONVERGENCE : CONDITIONS SUFFISANTES

- Condition nécessaire et suffisante :  $\rho(B) < 1$ .
- Si  $A$  est une matrice *à diagonale dominante stricte* par ligne, c'est-à-dire

$$|a_{ii}| > \sum_{j=1, \dots, n; j \neq i} |a_{ij}|, \quad i = 1, \dots, n.$$

*alors les méthodes de Jacobi et de Gauss-Seidel sont convergentes*

- Soit  $A$  *régulière, tridiagonale et dont les coefficients diagonaux sont tous non-nuls*. Alors les méthodes de Jacobi et de Gauss-Seidel sont toutes les deux soit divergentes soit convergentes. Dans le deuxième cas,  $\rho(B_{GS}) = \rho(B_J)^2$
- Si  $A$  est une matrice *symétrique définie positive*, alors la méthode de Gauss-Seidel converge (la méthode de Jacobi pas forcément).

# LA MÉTHODE DE RICHARDSON PRÉCONDITIONNÉE

Considérons la méthode itérative générale :

$$P(x^{(k+1)} - x^{(k)}) = \alpha_k r^{(k)}, \quad k \geq 0. \quad (6)$$

Cette méthode est appelée **méthode de Richardson stationnaire préconditionnée** si  $\alpha_k = \alpha$  (une constante donnée) ; autrement elle est dite **méthode de Richardson dynamique préconditionnée** quand  $\alpha_k$  peut varier au cours des itérations.

La matrice inversible  $P$  est appelée *préconditionneur* de  $A$ .

RICHARDSON : CHOIX DE  $\alpha_k$ 

Si  $A$  et  $P$  sont **symétriques définies positives**, alors on a deux critères optimaux pour le choix de  $\alpha_k$  :

1. *Cas stationnaire* :

$$\alpha_k = \alpha_{opt} = \frac{2}{\lambda_{min}(P^{-1}A) + \lambda_{max}(P^{-1}A)}, \quad k \geq 0,$$

où  $\lambda_{min}$  et  $\lambda_{max}$  désignent respectivement la plus petite et la plus grande valeur propre de la matrice  $P^{-1}A$ .

2. *Cas dynamique* :

$$\alpha_k = \frac{(z^{(k)})^T r^{(k)}}{(z^{(k)})^T A z^{(k)}}, \quad k \geq 0,$$

où  $z^{(k)} = P^{-1}r^{(k)}$  est le résidu préconditionné.

Cette méthode est aussi appelée **méthode du gradient préconditionné**.



# LA MÉTHODE DE RICHARDSON NON-PRÉCONDITIONNÉE

Si  $P = I$  et  $A$  est symétrique définie positive, on trouve les méthodes :

- de **Richardson stationnaire** si on choisit :

$$\alpha_k = \alpha_{opt} = \frac{2}{\lambda_{min}(A) + \lambda_{max}(A)}. \quad (7)$$

- du **gradient** si :

$$\alpha_k = \frac{(r^{(k)})^T r^{(k)}}{(r^{(k)})^T A r^{(k)}}, \quad k \geq 0. \quad (8)$$

# ALGORITHME DU GRADIENT PRÉCONDITIONNÉ

On peut réécrire plus efficacement la méthode du gradient préconditionné de la manière suivante : soit  $x^{(0)}$ , poser  $r^{(0)} = b - Ax^{(0)}$ , puis pour  $k \geq 0$ ,

$$\begin{aligned} Pz^{(k)} &= r^{(k)} \\ \alpha_k &= \frac{(z^{(k)})^T r^{(k)}}{(z^{(k)})^T A z^{(k)}} \\ x^{(k+1)} &= x^{(k)} + \alpha_k z^{(k)} \\ r^{(k+1)} &= r^{(k)} - \alpha_k A z^{(k)}. \end{aligned}$$

On observe qu'on doit résoudre un système linéaire pour la matrice  $P$  à chaque itération ; donc  $P$  doit être telle que la résolution du système associé soit facile (c'est-à-dire avec un coût raisonnable). Par exemple, on pourra choisir  $P$  diagonale (comme dans le cas du gradient ou de Richardson stationnaire) ou triangulaire.

### THÉORÈME (CAS STATIONNAIRE)

On suppose la matrice  $P$  inversible et les valeurs propres de  $P^{-1}A$  strictement positives et telles que  $\lambda_{\max} = \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n = \lambda_{\min} > 0$ . Alors la méthode de Richardson stationnaire est convergente si et seulement si  $0 < \alpha < 2/\lambda_1$ . De plus, le rayon spectral de la matrice d'itération  $R_\alpha$  est minimal si  $\alpha = \alpha_{\text{opt}}$

$$\alpha_{opt} = \frac{2}{\lambda_{min} + \lambda_{max}},$$

*avec*

$$\rho_{opt} = \frac{\lambda_{max} - \lambda_{min}}{\lambda_{min} + \lambda_{max}}$$

### THÉORÈME (CAS DYNAMIQUE)

$$\alpha_k = \frac{(r^{(k)}, z^{(k)})}{(Az^{(k)}, z^{(k)})}, \quad k \geq 0 \quad (9)$$
$$\mathbf{z}^{(k)} = P^{-1} \mathbf{r}^{(k)}. \quad (10)$$

Si on choisit les coefficients  $\alpha$  de manière optimale, pour les cas stationnaire et dynamique, on peut démontrer que, si  $A$  et  $P$  sont symétriques définies positives, la suite  $\{x^{(k)}\}$  donnée par la méthode de Richardson (stationnaire et dynamique) converge vers  $x$  lorsque  $k \rightarrow \infty$ , et

$$\|\mathbf{x}^{(k)} - \mathbf{x}\|_A \leq \left( \frac{K(P^{-1}A) - 1}{K(P^{-1}A) + 1} \right)^k \|\mathbf{x}^{(0)} - \mathbf{x}\|_A, \quad k \geq 0, \quad (11)$$

où  $\|v\|_A = \sqrt{v^T A v}$  et  $K(P^{-1}A)$  est le conditionnement de la matrice  $P^{-1}A$  :

$$K(C) = \sqrt{\frac{\lambda_{\max}(C^T C)}{\lambda_{\min}(C^T C)}}$$

**Remarque.** Dans le cas de la méthode du gradient ou de Richardson stationnaire (sans préconditionneur) l'estimation de l'erreur devient

$$\|x^{(k)} - x\|_A \leq \left( \frac{K(A) - 1}{K(A) + 1} \right)^k \|x^{(0)} - x\|_A, \quad k \geq 0. \quad (12)$$

**Remarque.** Si  $A$  et  $P$  sont symétriques définies positives, on a

$$K(P^{-1}A) = \frac{\lambda_{\max}(P^{-1}A)}{\lambda_{\min}(P^{-1}A)}.$$

# LA MÉTHODE DU GRADIENT CONJUGUÉ

Une méthode encore plus rapide dans le cas où  $P$  et  $A$  sont **symétriques définies positives** est celle du **gradient conjugué préconditionné** qui s'exprime ainsi :  
soit  $x^{(0)}$  une donnée initiale ; on calcule  $r^{(0)} = b - Ax^{(0)}$ ,  $z^{(0)} = P^{-1}r^{(0)}$ ,  
 $p^{(0)} = z^{(0)}$ , puis pour  $k \geq 0$ ,

$$\begin{aligned}\alpha_k &= \frac{p^{(k)T} r^{(k)}}{p^{(k)T} A p^{(k)}} \\ x^{(k+1)} &= x^{(k)} + \alpha_k p^{(k)} \\ r^{(k+1)} &= r^{(k)} - \alpha_k A p^{(k)} \\ P z^{(k+1)} &= r^{(k+1)} \\ \beta_k &= \frac{(A p^{(k)})^T z^{(k+1)}}{(A p^{(k)})^T p^{(k)}} \\ p^{(k+1)} &= z^{(k+1)} - \beta_k p^{(k)}.\end{aligned}$$

Dans ce cas, l'estimation de l'erreur est donnée par

$$\|x^{(k)} - x\|_A \leq \frac{2c^k}{1 + c^{2k}} \|x^{(0)} - x\|_A, \quad k \geq 0 \quad \text{où} \quad c = \frac{\sqrt{K_2(P^{-1}A)} - 1}{\sqrt{K_2(P^{-1}A)} + 1}. \quad (13)$$

