

APPROXIMATION NUMERIQUE ET OPTIMISATION

G. ALLAIRE

18 Septembre 2018

CHAPITRE III (début)

- ☞ Formulations variationnelles.
- ☞ Approximation variationnelle.
- ☞ Éléments finis.
- ☞ Résolution de systèmes linéaires.

modèle plus compliqué

Problème aux limites

$$\left\{ \begin{array}{l} -\operatorname{div}(K(x)\nabla u) + \alpha(x)u = f \quad \Omega \\ u = 0 \quad \partial\Omega \end{array} \right.$$

Comment peut-on résoudre:

$$\left\{ \begin{array}{ll} -\Delta u = f & \text{dans } \Omega \\ u = 0 & \text{sur } \partial\Omega \end{array} \right.$$

En dimension 1, si $\Omega = (0, L)$, alors

$$u(x) = Cx - \int_0^x \int_0^y f(z) dz dy \quad \text{avec} \quad C = L^{-1} \int_0^L \int_0^y f(z) dz dy$$

Mais en dimension plus grande ? Pas de formule explicite...

Idée principale: formulation variationnelle.

Formulation variationnelle en 1-d

Problème aux limites de Dirichlet:

$$(EDP) \quad \begin{cases} -u'' = f & \text{dans } (0, L) \\ u(0) = u(L) = 0 & \text{sur le bord .} \end{cases}$$

Formulation variationnelle: trouver

$u \in V_0 = \{\phi \in C^1[0, L] \text{ tel que } \phi(0) = \phi(L) = 0\}$ qui vérifie

$$(FV) \quad \int_0^L u'(x) \phi'(x) dx = \int_0^L f(x) \phi(x) dx \quad \forall \phi \in V_0.$$

Proposition. Soit $u \in C^2[0, L]$. Alors u est solution de (EDP) **si et seulement si** $u \in V_0$ est solution de (FV).

Démonstration

(EDP) \Rightarrow (FV)

Soit $\varphi \in V_0$. On multiplie l'EDP par φ :

$$\begin{aligned}
 - \int_0^L u'' \varphi \, dx &= \int_0^L f \varphi \, dx && \text{Par intégration par parties} \\
 &= \int_0^L u' \varphi' \, dx - \underbrace{u'(L)}_0 \underbrace{\varphi(L)}_0 + \underbrace{u'(0)}_0 \underbrace{\varphi(0)}_0 && \text{car } \varphi \in V_0
 \end{aligned}$$

On conclut que $\forall \varphi \in V_0$

$$\int_0^L u' \varphi' \, dx = \int_0^L f \varphi \, dx$$

(FV) \Rightarrow (EDP)

Même calcul à l'envers!

$$\int_0^L u' \varphi' \, dx = - \int_0^L u'' \varphi \, dx + \underbrace{u'(L)}_0 \underbrace{\varphi(L)}_0 - \underbrace{u'(0)}_0 \underbrace{\varphi(0)}_0 \quad \text{car } \varphi \in V_0$$

$$\Rightarrow \int_0^L (u'' + f) \varphi \, dx = 0 \quad \forall \varphi \in V_0 \quad \Rightarrow \quad \left. \begin{aligned} & -u'' = f \text{ sur } [0, L] \end{aligned} \right\} \text{(EDP)}$$

$$\text{or } u \in V_0 \Rightarrow u(0) = u(L) = 0$$

Lemme. Soit $g(x)$ une fonction continue sur $[0, L]$. Si elle vérifie

$$\int_0^L g(x) \phi(x) dx = 0 \quad \text{pour tout } \phi \in V_0,$$

alors $g \equiv 0$ sur $[0, L]$.

Démonstration. Si $g \neq 0$, alors $\exists x_0 \in [0, L]$ tq $g(x_0) \neq 0$. Par exemple $g(x_0) > 0$.

Par continuité, on peut choisir $x_0 \in]0, L[\Rightarrow \exists \varepsilon > 0$ tq $]x_0 - \varepsilon, x_0 + \varepsilon[\subset]0, L[$

Soit φ à support compact dans $]x_0 - \varepsilon, x_0 + \varepsilon[$, $\varphi \neq 0$ et $\varphi \geq 0$

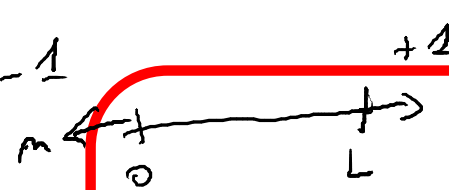
$$\int_0^L g(x) \varphi(x) dx > 0 \quad \text{contradiction} \Rightarrow g \equiv 0 \text{ sur } [0, L].$$

Remarque. Si on veut résoudre un problème avec des conditions aux limites de Dirichlet **non homogènes**

$$\begin{cases} -u'' = f & \text{dans } (0, L) \\ u(0) = \alpha, u(L) = \beta & \text{sur le bord} \end{cases}$$

alors on pose $u(x) = \tilde{u}(x) + (\alpha(L - x) + \beta x)/L$ et on cherche \tilde{u} solution de

$$\begin{cases} -\tilde{u}'' = f & \text{dans } (0, L) \\ \tilde{u}(0) = \tilde{u}(L) = 0 & \text{sur le bord} \end{cases}$$



$$\frac{\partial u}{\partial n} = g \text{ sur } \partial \Omega$$

↓
V. n

Condition aux limites de Neumann et Fourier

Condition aux limites de Fourier: étant donné un paramètre $\kappa \geq 0$

$$(EDP) \quad \begin{cases} -u'' = f & \text{dans } (0, L) \\ -u'(0) + \kappa u(0) = \alpha, \quad u'(L) + \kappa u(L) = \beta & \text{sur le bord.} \end{cases}$$

Les conditions aux limites de Neumann correspondent à $\kappa = 0$.

Formulation variationnelle: trouver $u \in V = C^1[0, L]$ qui vérifie

$$(FV) \quad \int_0^L u'(x) \phi'(x) dx + \kappa (u(L)\phi(L) + u(0)\phi(0)) = \int_0^L f(x) \phi(x) dx + \beta\phi(L) + \alpha\phi(0) \quad \forall \phi \in V.$$

Proposition. Soit $u \in C^2[0, L]$. Alors u est solution de (EDP) **si et seulement si** $u \in V$ est solution de (FV).

Remarques

- ☞ Dans une formulation variationnelle on a besoin de dériver une fois de moins (**pratique !**).
- ☞ La condition aux limites de Dirichlet apparait dans la définition de l'espace V_0 . On parle de condition aux limites **essentielle** (ou explicite).
- ☞ La condition aux limites de Neumann ou de Fourier n'apparait pas dans la définition de l'espace V . On parle de condition aux limites **naturelle** (ou implicite).
- ☞ Dans le cas $\kappa = 0$ (condition aux limites de Neumann), il ne peut exister de solution que s'il y a **équilibre des sources**

$$\int_0^L f(x) dx + \beta + \alpha = 0$$

Démonstration

(EDF) \Rightarrow (FV) Soit $\varphi \in C^1[0, L]$ on multiplie l'équation par φ

$$-\int_0^L u'' \varphi \, dx = \int_0^L f \varphi \, dx$$

$$\int_0^L u' \varphi' \, dx = \underbrace{-u'(L) \varphi(L)}_{\kappa u(L) \varphi(L) - \beta \varphi(L)} + \underbrace{u'(0) \varphi(0)}_{\kappa u(0) \varphi(0) - \alpha \varphi(0)}$$

ou les CL $-u'(0) + \kappa u(0) = \alpha$ et $u'(L) + \kappa u(L) = \beta$

$$\Rightarrow \int_0^L u' \varphi' \, dx + \kappa (u(L) \varphi(L) + u(0) \varphi(0)) = \int_0^L f \varphi \, dx + \alpha \varphi(0) + \beta \varphi(L) \quad \forall \varphi \in C^1$$

(FV) \Rightarrow (EDF)

$$-\int_0^L u'' \varphi \, dx + (u'(L) + \kappa u(L)) \varphi(L) + (-u'(0) + \kappa u(0)) \varphi(0) = \int_0^L f \varphi \, dx + \alpha \varphi(0) + \beta \varphi(L)$$

Soit $\varphi \in C_c^1[0, L]$ (à support compact) en particulier $\varphi(0) = \varphi(L) = 0$

$$\Rightarrow \int_0^L (u'' + f) \varphi \, dx = 0 \quad \Rightarrow \text{(lemme précédent)} \quad u'' + f = c \text{ sur } [0, L]$$

$$\forall \varphi \in C^1 \Rightarrow u'(L) + \kappa u(L) = \beta \quad \text{et} \quad -u'(0) + \kappa u(0) = \alpha$$

Formulation variationnelle abstraite

Dans tous les cas une formulation variationnelle (FV) se met sous la forme générale suivante: trouver $u \in V$ tel que

$$a(u, \phi) = L(\phi) \quad \forall \phi \in V$$

avec

- ☞ V un espace vectoriel réel,
- ☞ $L(\cdot)$ une forme linéaire sur V (c'est-à-dire une application linéaire de V dans \mathbb{R}).
- ☞ $a(\cdot, \cdot)$ une forme bilinéaire sur V (c'est-à-dire une application de $V \times V$ dans \mathbb{R} , linéaire par rapport à chacun de ses deux arguments).
- ☞ Dans ce cours on ne démontre pas l'existence d'une solution de la FV.

Dans le cas Dirichlet on a $V = V_0 = \{\phi \in C^1[0, L] \text{ tel que } \phi(0) = \phi(L) = 0\}$,

$$a(u, \phi) = \int_0^L u'(x) \phi'(x) dx \quad \text{et} \quad L(\phi) = \int_0^L f(x) \phi(x) dx,$$

tandis que pour Fourier/Neumann on a $V = C^1[0, L]$,

$$a(u, \phi) = \int_0^L u'(x) \phi'(x) dx + \kappa(u(L)\phi(L) + u(0)\phi(0))$$

et

$$L(\phi) = \int_0^L f(x) \phi(x) dx + \beta\phi(L) + \alpha\phi(0)$$

Remarque (importante pour les éléments finis). Les formulations variationnelles sont toujours bien définies si on remplace $C^1[0, L]$ par $\{\phi \in C^0[0, L] \text{ dérivable par morceaux}\}$.

Approximation variationnelle

Soit une formulation variationnelle: trouver $u \in V$ tel que $a(u, \phi) = L(\phi)$ pour tout $\phi \in V$.

Soit V_h un sous-espace vectoriel de V , de dimension finie.

On appelle **approximation variationnelle interne** le problème: trouver $u_h \in V_h$ tel que

$$a(u_h, \phi_h) = L(\phi_h) \quad \forall \phi_h \in V_h$$

Lemme. On suppose que la forme bilinéaire est **coercive**

$$\exists \nu > 0 \text{ tel que } a(v, v) \geq \nu \|v\|^2 \quad \forall v \in V.$$

Alors il existe une unique solution de l'approximation variationnelle interne qui, de plus, s'obtient en résolvant un système linéaire.

Démonstration

Trouver $u_h \in V_h$ tel que $a(u_h, \varphi_h) = L(\varphi_h) \quad \forall \varphi_h \in V_h$

$\dim V_h < +\infty \Rightarrow \exists$ une base $(\varphi_i)_{1 \leq i \leq m}$ de V_h

on développe $u_h = \sum_{i=1}^m u_i \varphi_i$. On choisit $\varphi_h = \varphi_j$

$$\Rightarrow \sum_{i=1}^m u_i a(\varphi_i, \varphi_j) = L(\varphi_j) \quad \text{On introduit } b = (L(\varphi_j))_{1 \leq j \leq m} \in \mathbb{R}^m$$

$$U = (u_i)_{1 \leq i \leq m} \in \mathbb{R}^m$$

matrice $K = (a(\varphi_j, \varphi_i))_{1 \leq i, j \leq m}$

$$\Rightarrow \boxed{KV = b} \Leftrightarrow (FV) \text{ sur } V_h$$

Remarque $a(v, v) = K \cdot v \geq c \|v\|^2 \geq c \|v\|^2$

$$\text{avec } v = \sum_{i=1}^m v_i \varphi_i$$

$c > 0$ et équivalents des normes en dim finie

$$\Rightarrow \det K \neq 0 \Leftrightarrow K \text{ inversible}$$

Analogie en dim finie

$x, b \in \mathbb{R}^m$, A une matrice $m \times n$

Résoudre $Ax = b$ est équivalent à

$$Ax \cdot y = b \cdot y \quad \forall y \in \mathbb{R}^m$$



comme une FV

↑
comme une EDP

Erreur d'approximation

Lemme de Céa. On suppose que la forme bilinéaire est **coercive** mais aussi **continue** au sens où

$$\exists M > 0 \text{ tel que } |a(v, w)| \leq M \|v\| \|w\| \quad \forall v, w \in V.$$

Soit u la solution de la formulation variationnelle et soit u_h la solution de l'approximation variationnelle interne. On a

$$\|u - u_h\| \leq \frac{M}{\nu} \inf_{v_h \in V_h} \|u - v_h\|.$$

Démonstration

Eléments finis en 1-d

En 1-d, un maillage du segment $(0, 1)$ est une collection de points (appelés **sommets ou noeuds**) $(x_j)_{0 \leq j \leq n+1}$ tels que

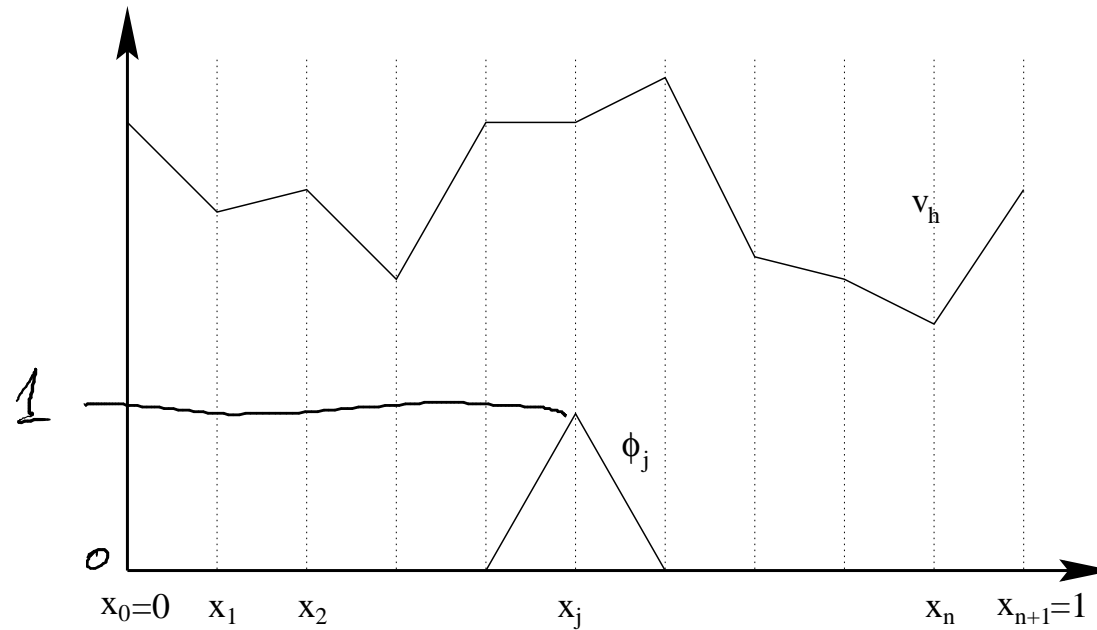
$$x_0 = 0 < x_1 < \dots < x_n < x_{n+1} = 1.$$

Le maillage est dit **uniforme** si les points x_j sont équidistants, c'est-à-dire que

$$x_j = jh \quad \text{avec} \quad h = \frac{1}{n+1}, \quad 0 \leq j \leq n+1.$$

On note \mathbb{P}_k l'ensemble des polynômes réels de degré inférieur ou égal à k .

Eléments finis de Lagrange \mathbb{P}_1



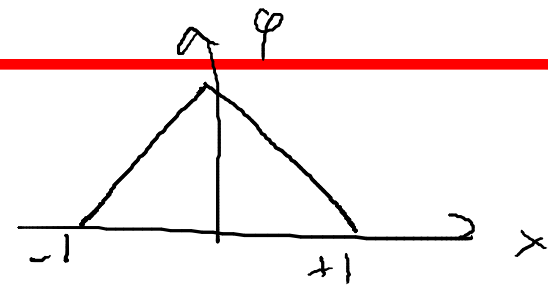
La méthode des éléments finis \mathbb{P}_1 est la méthode d'approximation variationnelle interne avec

$$V_h = \{v \in C[0, 1] \text{ tel que } v|_{[x_j, x_{j+1}]} \in \mathbb{P}_1 \text{ pour tout } 0 \leq j \leq n\},$$

et son sous-espace

$$V_{0h} = \{v \in V_h \text{ tel que } v(0) = v(1) = 0\}.$$

Base de V_h



Soit la “fonction chapeau”

$$\phi(x) = \begin{cases} 1 - |x| & \text{si } |x| \leq 1, \\ 0 & \text{si } |x| > 1. \end{cases}$$

Pour un maillage uniforme, on définit les fonctions de base

$$\phi_j(x) = \phi\left(\frac{x - x_j}{h}\right) \quad 0 \leq j \leq n + 1.$$

Lemme. L'espace V_h est un espace vectoriel de dimension $n + 2$, et toute fonction $v_h \in V_h$ est définie de manière unique par ses valeurs aux sommets

$$v_h(x) = \sum_{j=0}^{n+1} v_h(x_j) \phi_j(x) \quad \forall x \in [0, 1].$$

Lagrange

De même, V_{0h} est un espace vectoriel de dimension n , et toute fonction $v_h \in V_{0h}$ vérifie $v_h(x) = \sum_{j=1}^n v_h(x_j) \phi_j(x) \quad \forall x \in [0, 1].$

Démonstration

$$\begin{cases} -u'' = f \text{ sur } (0,1) \\ u(0) = u(1) = 0 \end{cases}$$

Résolution pratique

Trouver $u_h \in V_h = \{ \varphi \in C[0,1] \mid \varphi|_{[x_j, x_{j+1}]} \in \mathbb{P}_1 \text{ et } \varphi(0) = \varphi(1) = 0 \}$ tel que

$$a(u_h, \varphi) = L(\varphi) \quad \forall \varphi \in V_h$$

$$L(\varphi) = \int_0^1 f \varphi \, dx \quad \text{et} \quad a(u, v) = \int_0^1 u' v' \, dx \stackrel{\text{Sym.}}{=} a(v, u)$$

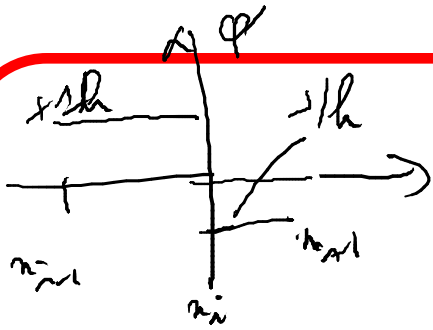
$(FV_h) \Leftrightarrow$

$$K \mathcal{U} = b$$

$$K = (a(\varphi_j, \varphi_i))_{1 \leq i, j \leq n}$$

$$b_i = \int_0^1 f \varphi_i \, dx$$

$$a(\varphi_i, \varphi_j) = \begin{cases} 0 & \text{si } |i-j| > 1 \\ \chi(?) & \text{si } |i-j| \leq 1 \end{cases}$$



Matrice de rigidité

$$a(\varphi_i, \varphi_i) = \int_{x_{i-1}}^{x_i} |\varphi_i'|^2 dx = \frac{2h}{h} = 2$$

$$a(\varphi_i, \varphi_{i+1}) = \int_{x_i}^{x_{i+1}} \frac{-1}{h} \frac{1}{h} dx = \frac{-1}{h}$$

On définit la matrice de rigidité

$$\mathcal{K}_h = \left(\int_0^1 \phi_j'(x) \phi_i'(x) dx \right)_{1 \leq i, j \leq n},$$

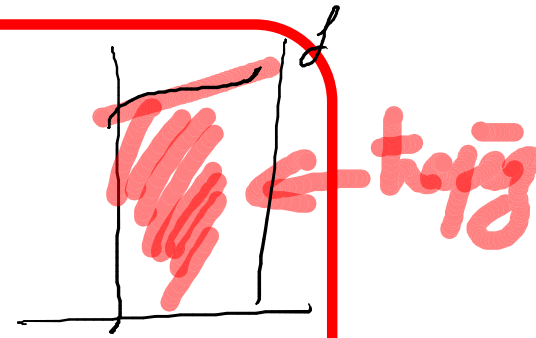
Lemme. La matrice \mathcal{K}_h est tridiagonale

$$\mathcal{K}_h = h^{-1} \begin{pmatrix} 2 & -1 & & & 0 \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ 0 & & & -1 & 2 \end{pmatrix}.$$

$\underbrace{\hspace{15em}}_n$

Démonstration

Aspects pratiques



Formules de quadrature, résolution de système linéaire...

= calcul d'intégrales approché (par exemple pour calculer $\int_0^1 f(x) dx$)

1°) formule des trapèzes $\int_{x_j}^{x_{j+1}} f(x) dx \approx \frac{h}{2} (f(x_j) + f(x_{j+1}))$ ← avantage

2°) formule de Simpson ---

3°) formule du point milieu

$$\int_{x_j}^{x_{j+1}} f(x) dx \approx h f\left(\frac{x_j + x_{j+1}}{2}\right)$$

Résolution des systèmes linéaires

$$Ax = b$$

- ☞ **Problème crucial** en temps CPU et place mémoire dans le calcul scientifique.
- ☞ On ne calcule jamais la matrice **inverse** A^{-1} !
- ☞ Deux classes de méthodes: **directes** et **itératives**.
- ☞ On ne stocke pas toute la matrice A qui est souvent **creuse** (contient beaucoup d'éléments nuls).

Méthodes directes

Méthode d'élimination de Gauss: si on ne pivote pas, on obtient une factorisation $A = LU$ avec

$$L = \begin{pmatrix} 1 & 0 & \dots & 0 \\ l_{2,1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ l_{n,1} & \dots & l_{n,n-1} & 1 \end{pmatrix}, \quad U = \begin{pmatrix} u_{1,1} & \dots & \dots & u_{1,n} \\ 0 & u_{2,2} & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & u_{n,n} \end{pmatrix}$$

Méthode de Cholesky: si la matrice A est symétrique définie positive, alors il existe une matrice triangulaire inférieure B telle que

$$A = BB^*$$

$$Ax = b$$

Résolution par remontée-descente: on résout successivement (et facilement !)

$Ly = b$ (descente) puis $Ux = y$ (remontée). ← $O(\frac{n^2}{2})$

Compte d'opérations: Gauss $N_{op} \approx n^3/3$, Cholesky $N_{op} \approx n^3/6$.

Calcul de la factorisation LU, colonne par colonne

L_1 et $U_1 = 1$ ère colonnes de L et U

→ L et U

$$L = \begin{pmatrix} \begin{matrix} \times & & & \\ & \times & & \\ & & \times & \\ & & & \times \end{matrix} & \dots & \\ \dots & & & \dots & \\ \dots & & & \dots & \\ \dots & & & \dots & \end{pmatrix} \quad U = \begin{pmatrix} \begin{matrix} \times & & & \\ 0 & \times & & \\ & & \times & \\ & & & \times \end{matrix} & \dots & \\ \dots & & & \dots & \\ \dots & & & \dots & \\ \dots & & & \dots & \end{pmatrix}$$

k ème colonne de $L \times U$

$$\sum_{i=1}^m \left(\sum_{i=1}^k i + \sum_{i=k+1}^m k \approx \frac{k^2}{2} + k(m-k) \right)$$

$$= O\left(\frac{m^3}{3}\right)$$

Méthodes itératives

Soit A une matrice inversible avec $A = M - N$ et M inversible. Pour résoudre $Ax = b$ on définit une méthode itérative basée sur la décomposition (M, N) par

$$\begin{cases} x_0 \text{ donné dans } \mathbb{R}^n, \\ Mx_{k+1} = Nx_k + b \quad \forall k \geq 1. \end{cases}$$

Exemples:

- ☞ **Jacobi:** $M = \text{diag}(A)$ et $N = \text{diag}(A) - A$
- ☞ **Gauss-Seidel:** si $A = D - E - F$ avec D diagonal, E triangulaire inférieur et F triangulaire supérieur, alors $M = D - E$ et $N = F$.
- ☞ **Gradient conjugué:** plus compliqué !

Méthodes itératives

Définition. Une méthode itérative est dite convergente si, pour tout vecteur initial $x_0 \in \mathbb{R}^n$, la suite x_k converge vers la solution exacte x .

Lemme. La méthode itérative converge si et seulement si le rayon spectral de la matrice d'itération $M^{-1}N$ vérifie $\rho(M^{-1}N) < 1$.

Préconditionnement

Idée cruciale !

Soit une matrice inversible C proche de A mais plus facile à inverser. Comme

$$Ax = b \Leftrightarrow C^{-1}Ax = C^{-1}b,$$

on applique une méthode itérative à $C^{-1}A$ en **espérant** que $\rho(M^{-1}N)$ sera plus petit pour $C^{-1}A$ que pour A .

Exemple: $C = \text{diag}(A)$ ou bien $C =$ approximation de LU .

Matrices bandes ou creuses

Une matrice A est dite **bande**, de demie largeur de bande p , si ses éléments vérifient $a_{i,j} = 0$ pour $|i - j| > p$.

Lemme. Les factorisations LU et de Cholesky conservent la structure bande des matrices.

$\nearrow O(mp^2) \searrow$

Une matrice A est stockée sous forme **creuse** si on ne garde en mémoire que ses éléments non nuls.

Gradient conjugué

Soit A une matrice symétrique définie positive, et $x_0 \in \mathbb{R}^n$. Soit (x_k, r_k, p_k) trois suites définies par les relations de récurrence

$$p_0 = r_0 = b - Ax_0, \text{ et pour } 0 \leq k \left\{ \begin{array}{l} x_{k+1} = x_k + \alpha_k p_k \\ r_{k+1} = r_k - \alpha_k A p_k \\ p_{k+1} = r_{k+1} + \beta_k p_k \end{array} \right.$$

avec

$$\alpha_k = \frac{\|r_k\|^2}{Ap_k \cdot p_k} \quad \text{et} \quad \beta_k = \frac{\|r_{k+1}\|^2}{\|r_k\|^2}.$$

Alors, la suite (x_k) de la méthode du gradient conjugué converge en moins de n itérations vers la solution du système linéaire $Ax = b$.