Analyse numérique avancée – TD 1

M. Kern

24 février 2022

Thème : Gradient conjugué – préconditionnement – GMRES

Exercice 1 : Convergence du gradient conjugué

Soit $A \in \mathbf{R}^{n\times n}$ une matrice définie positive, et $b \in \mathbf{R}^n$. On utilise la méthode du gradient conjugué pour résoudre le système

$$Ax = b.$$

On rappelle que le problème de minimisation

$$\min_{p\in P_k,\, p(0)=1} \max_{\lambda\in[c,d]} |p(\lambda)| \tag{1}$$

admet une solution $p_k^\star$ qui vérifie l'inégalité suivante (on note $\kappa = d/c$.)

$$|p_k^\star(\lambda)| \leq 2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^k \tag{2}$$

1 Donner une estimation du nombre d'itérations nécessaires pour réduire l'erreur initiale d'un facteur ε donné en utilisant la borne simple fournie par le conditionnement.

2 Dans cette question, A est la matrice obtenue par discrétisation de l'équation de Laplace sur le carré unité par une méthode de différences finies sur une grille régulière de pas h.

Donner un équivalent du conditionnement de A pour h petit. Comment le nombre d'itérations nécessaire pour réduire l'erreur d'un facteur η donné varie-t-il en fonction de h ?

3 On suppose maintenant que les valeurs propres de A, notées $\lambda_1 \leq \cdots \leq \lambda_n$ vérifient

$$c \leq \lambda_1 \leq \cdots \leq \lambda_{n-1} \leq d \ll \lambda_n$$

Donner une borne plus précise sur en utilisant cette information, et en déduire une meilleure estimation du nombre d'itérations.

Indication : utiliser un polynôme de la forme $p(\lambda) = (1 - \alpha\lambda)q(\lambda)$ où α est à choisir, et q résout un problème de minimisation de même type que 1.

Exercice 2 : Coût du gradient conjugué

Soit $A \in \mathbf{R}^{n \times n}$ une matrice définie positive.

1 On suppose d'abord que A est une matrice pleine. Calculer le coût d'une itération du gradient conjugué appliqué à A. Comparer le coût total s'il faut n itérations pour obtenir la solution avec la méthode de Gauss (ou Cholesky).

2 Maintenant on prend pour A la matrice du obtenue par discrétisation de l'équation de Laplace sur le carré unité par une méthode de différences finies sur une grille régulière de pas $h = 1/(N+1)$, avec N entier.

Quel est le coût de la méthode par itération, et quel est le coût total pour réduire l'erreur d'un facteur η (voir exercice précédent) ?

3 Pourquoi est-il, en théorie, suffisant d'étudier la convergence du gradient conjugué pour des matrices diagonales ? Est-ce le cas en pratique ?

Exercice 3 : Convergence de GMRES

On utilise GMRES pour résoudre un système linéaire $Ax = b$, où $A \in \mathbf{R}^{n \times n}$ est une matrice inversible. On notera $(e_1, \dots, e_n)$ les vecteurs de la base canonique de $\mathbf{R}^n$.

1 Dans la suite, on considère la matrice $A \in \mathbf{R}^{4 \times 4}$ dont les colonnes sont e_2, e_3, e_4, e_1, et le second membre $b = e_1$. On démarre les itérations avec $x_0 = 0$.

Représenter la matrice. Quelle est la solution exacte du système ?

2 Déterminer les matrices V_k produites par l'algorithme d'Arnoldi, les matrices réduites H_k, pour $k = 1, \dots, n$.

En déduire les itérés successifs. Que peut-on dire de la convergence de l'algorithme pour cette matrice ?

Pouvez-vous généraliser le résultat ?

Exercice 4 : Résidus de GMRES

On résout le système $Ax = b$ (A est supposée inversible) par la méthode GMRES.

1 Montrer que la condition d'optimalité de GMRES à l'itération k est équivalente à la condition

$$r_k \perp AK_k(A, r_0).$$

Montrer que $V_{k+1}^T A V_k = H_k$.

2 On note r_0 le résidu initial, et on suppose $r_0 \neq 0$. Montrer qu'à la première itération, le résidu r_1 vérifie

$$r_1 = r_0 - \alpha A r_0$$

pour un scalaire α que l'on précisera.

En déduire l'identité

$$\left(\frac{\|r_1\|_2}{\|r_0\|_2} \right)_2 = 1 - \frac{(Ar_0,\, r_0)}{(r_0,\, r_0)} \frac{(Ar_0,\, r_0)}{(Ar_0,\, Ar_0)}$$

Exercice 5 : Factorisation incomplète et matrices de différences finies

On considère une matrice tridiagonale par blocs qui intervient dans la résolution d'une équation elliptique en dimension 2 (sur le carré unité, avec un mailalge régulier de N points sur chaque coté), et on considère le cas général où A n'est pas nécessairement symétrique. On écrit

$$A = L_A + D_A + U_A,$$

où D_A est la diagonale principale de A, et L_A et U_A sont les parties strictement triangulaires inférieures et supérieures respectivement. Toutes les matrices sont $n \times n$, avec $n = N^2$.

On veut calculer une factorisation incomplète de A de la forme

$$M = (L_M + \Delta)\Delta^{-1}(\Delta + U_M),$$

où Δ est une matrice diagonale (inversible) et L_M et U_M sont triangulaires inférieures et supérieures, à diagonale nulle.

La matrice L_M n'a que deux diagonales non-nulles, en position -1 et $-n$, et de même U_M n'a que deux diagonales non-nulles, en position 1 et n.

1 Question préliminaire.

Pour un entier $-n \leq p \leq n$, on note A_p une matrice n'ayant qu'une seule diagonale non-nulle en position p ($p = 0$ correspond à la diagonale principale, $p < 0$ à la partie inférieure, $p > 0$ à la partie supérieure).

Montrer que $A_p A_q = A_{p+q}$. On peut donc trouver l'emplacement du produit de deux matrices diagonales.

2 Quelles sont a priori les diagonales non-nulles de L_M et U_M ? En décomposant A et M en somme selon les diagonales, déterminer sans calcul les matrices L_M et U_M.

3 En ecrivant l'égalité des diagonales de M et de A (pourquoi ?), déterminer les éléments de Δ. On obiendra une relation de récurrence, et il est utile de traduire cette relation en écriture à deux indices, en revenant sur la grille.

Exercice 6 : Opérations avec la matrice transposée en format CSR

Soit $L \in \mathbf{R}^{n\times n}$ une matrice **triangulaire inférieure**. On supposera que les élément diagonaux de L sont tous égaux à 1.

On veut écrire des algorithmes pour résoudre le système triangulaire $L^T x = b$ quand la matrice L est stockée dans le format CSR.

1 Sans se préoccuper du stockage, écrire un algorithme pour calculer la solution du système $L^T x = b$ qui accède aux éléments de L par lignes.

Indication : penser à interpréter le produit $L^T x$ comme une combinaison linéaire des **colonnes** de L^T.

2 Dans cette question L est stockée sous forme CSR, dans 3 tableaux `IL`, `JL`, `VL`. En utilisant la question précédente, écrire un algorithme pour résoudre le système, en exploitant la structure de données.