NOTES DE COURS

METHODES ITERATIVES DE KRYLOV

1 Méthode de projection

1.1 Généralités

Considérons un système linéaire inversible admettant $\bar{u}$ comme unique solution du système

$$Au = b, \tag{1}$$

où A est une matrice carrée inversible d'ordre n.
On se donne une solution initiale u_0 ou encore une approximation de la solution $\bar{u}$. Le résidu associé au vecteur u_0 est donné par

$$r_0 = b - Au_0.$$

Le système linéaire (1) peut aussi s'écrire sous forme résiduelle:
trouver $\bar{u} \in R^n$ vérifiant la relation

$$(b - A\bar{u}, v) = 0 \quad \forall v \in R^n. \tag{2}$$

En d'autres termes, cela signifie que

$$b - A\bar{u} \perp R^n \quad \text{ou encore} \quad b - A\bar{u} \in (R^n)^\perp \equiv \{0\}.$$

Le principe de base d'une méthode de projection consiste à se donner un sous-espace K et de chercher une approximation u de la solution dans l'espace $u_0 + K$ vérifiant une condition d'orthogonalité plus faible que (2)

$$b - Au \perp K.$$

On peut réécrire cette condition d'orthogonalité sous la forme:
déterminer $u \in u_0 + K$ vérifiant:

$$(Au, v) = (b, v) \quad \forall v \in K$$

Le lecteur notera une grande similitude avec la méthode de Galerkin utilisée dans la méthode des éléments finis. La condition ci-dessus porte le nom de condition de Galerkin.
Plus généralement, on se donnera deux sous-espaces K et L de même dimension. On recherchera une solution approchée $u \in u_0 + K$ vérifiant la condition d'orthogonalité

$$b - Au \perp L.$$

De même cette relation peut s'écrire sous la forme:
déterminer $u \in u_0 + K$ vérifiant:

$$(Au, v) = (b, v) \quad \forall v \in L. \tag{3}$$

Cette dernière relation porte le nom de condition de Petrov-Galerkin.

Voici l'interprétation géométrique de la condition:

d'où l'appellation de méthode de projection.

En résumé, une méthode de projection est une méthode itérative de résolution des systèmes linéaires qui, à chaque itération,

- construit deux sous-espaces K et L,
- recherche le nouvel itéré u dans $u_0 + K$ satisfaisant la condition de Petrov-Galerkin

$$(Au, v) = (b, v) \quad \forall v \in L.$$

1.2 Forme matricielle

Dans cette sous-section. nous allons poser le problème sous forme matricielle.
Soit $V = [v_1, v_2, \ldots, v_m]$ une matrice d'ordre $n \times m$ construite à partir de m vecteurs $v_1, v_2, \ldots, v_m$ formant une base du sous-espace K. De même, on notera par $W = [w_1, w_2, \ldots, w_m]$ la matrice déduite d'une base du sous-espace L.
Un vecteur v de K, $v = \sum_{i=1}^{m} y_i v_i \in K$, s'écrira

$$v = Vy$$

où $y = (y_1, y_2, \ldots, y_m) \in R^m$. De même, un vecteur w quelconque de L peut s'écrire

$$w = Wz \in L.$$

Le problème consiste à trouver la solution approchée $u = u_0 + Vy$ où $y \in R^m$.
La condition de Petrov-Galerkin s'écrit, sous forme corrective,

$$\begin{array}{rcl} (Av, w) & = & (r_0, w) \quad \forall w \in L \\ (AVy, Wz) & = & (r_0, Wz) \quad \forall z \in R^m \\ (W^t AVy, z) & = & (W^t r_0, z) \quad \forall z \in R^m \\ W^t AVy & = & W^t r_0 \end{array}$$

Par conséquent, la solution correspondant à la condition de Petrov-Galerkin est donnée par

$$u = u_0 + V(W^t AV)^{-1} W^t r_0.$$

Voici le prototype d'un algorithme basé sur la méthode de projection.
En itérant jusqu'à convergence, faire:

1. choisir deux sous-espaces K_m et L_m de R^n de même dimension m,
2. construire les bases V_m et W_m associées aux sous-espaces K_m et L_m,
3. évaluer le résidu courant: $r = b - Au$,
4. résoudre le système linéaire en dimension m: $y = (W_m^t AV_m)^{-1} W_m^t r$,
5. mettre à jour la solution: $u = u + V_m y$

1.3 Deux familles de méthodes de projection

Dans cette sous-section, nous allons étudier, de mainière générale, deux familles de méthode de projection.
Avant d'expliciter les choix des sous-espaces K et L, nous allons décrire, de manière générale, deux grandes familles de méthodes:

- $L = K$: ceci conduira aux méthodes Full Orthogonalization Method (FOM) et du gradient conjugué (CG).
- $L = AK$: ceci conduira aux méthodes GMRES et MINRES.

Analysons ces deux cas:

1. $L_m = K_m$
   Supposons que la matrice A soit symétrique et définie-positive. Dans ce cas, l'équation (2) s'interprète comme étant la condition d'optimalité du problème de minimisation:
   trouver $u_m \in u_0 + K_m$ tel que

   $$J(u_m) = \min_{v \in K_m} J(u_0 + v)$$

   où la fonctionnelle J est définie par

   $$J(v) = \frac{1}{2}(Av, v) - (b, v).$$

   En effet, la condition d'optimalité du problème de minimisation s'écrit

   $$\begin{aligned} (\nabla J(u_m), v) &= (Au_m - b, v) = 0 \quad \forall v \in K_m \\ b - Au_m &\perp K_m \end{aligned}$$

   On peut aussi donner une autre interprétation de la méthode de projection. En effet, posons

   $$E(u_0 + v) = (A(u_0 + v - \bar{u}), u_0 + v - \bar{u}) = \|u_0 + v - \bar{u}\|_A^2$$

   C'est-à-dire l'erreur mesurée à l'aide de la norme induite par A.

La condition optimalité de ce problème de minimisation est donnée par

$$\begin{aligned}
(\nabla E(u_m), v) &= 2(A(u_0 + v_m - \bar{u}), v) \\
&= 2(A(u_0 + v_m) - b, v) \quad \forall v \in K_m \\
&= 0.
\end{aligned}$$

Par conséquent, la solution u_m donnée par la méthode de projection admet la caractérisation suivante:

$$E(u_m) = \min_{v \in K_m} E(u_0 + v)$$

Remarques

On obtient la méthode du gradient conjugué si on prend $K_m = [\nabla J(u_1), \nabla J(u_2), \ldots, \nabla J(u_{m-1})]$. De même, la méthode de descente optimale s'inscrit dans cette famille. En effet, on prendra comme sous-espace $K_m = [\nabla J(u_{m-1})] = [r_{m-1}]$.

2. $L_m = AK_m$:
Dans ce cas, on peut écrire le problème approché (2) sous la forme

$$(Au_m, Av) = (b, Av) \quad \forall v \in L_m$$

Mais ceci est la condition d'optimalité du problème de minimisation:
trouver $u_m \in u_0 + K_m$ tel que

$$J(u_m) = \min_{v \in K_m} J(u_0 + v) = \min_{v \in K_m} \frac{1}{2} \|A(u_0 + v) - b\|_2^2$$

En effet, on a que

$$\frac{1}{2}\|A(u_0 + \epsilon v) - b\|^2 = \frac{1}{2}(A(u_0 + \epsilon v) - b, A(u_0 + \epsilon v) - b)$$

On dérive cette relation par rapport à ϵ et ensuite on pose $\epsilon = 0$. On obtient:

$$(\nabla J(u_m), v) = (Au_m - b, Av) = 0 \quad \forall v \in K_m$$

2 Méthodes itératives de Krylov

Ces méthodes itératives sont des méthodes de projection basées sur un choix particulier du sous-espace K. Le choix de K est lié au théorème de Cayley-Hamilton.

Théorème de Cayley-Hamilton

Soient A une matrice carrée d'ordre n et $h(\lambda) \equiv \det(\lambda Id - A)$ le polynôme caractéristique de A. On a que

$$h(A) = 0.$$

2.1 Généralités

Si on utilise ce théorème, on obtient l'existence d'un polynôme de degré $n-1$ vérifiant

$$A^{-1} = p(A).$$

Ceci signifie que la solution du problème (1) s'écrit maintenant

$$v = p(A)r_0 \quad \text{ou encore} \quad u = u_0 + p(A)r_0$$

Ceci suggère de définir les sous-espaces suivants de R^n.

Définition: sous-espaces de Krylov

$$\mathcal{K}_m(A, r_0) = \mathcal{K}_m = [r_0, Ar_0, A^2 r_0, \ldots, A^{m-1} r_0] = \{p(A)r_0 | p \in P_{m-1}\}$$

où P_{m-1} est l'espace des polynômes de degré $m-1$. Généralement, on a que $dim\ K_m = m$. D'une manière plus précise, cette dimension est donnée par

$$dim\ K_m = \min\{m, \text{ degré du polynôme minimal de } A\}$$

De plus, ces espaces vérifient les inclusions suivantes

$$\mathcal{K}_1 = [r_0] \subset \mathcal{K}_m \subset \mathcal{K}_{m+1} \subset \mathcal{K}_n = R^n.$$

Remarque

Faisons le lien entre les sous-espaces de Krylov et l'algorithme du gradient conjugué. On rappelle que l'itéré u_m est caractérisé par

$$J(u_m) = \min_{\alpha_i} J(u_{m-1} + \sum_{i=0}^{m-1} \alpha_i \nabla J(u_i))$$

et que l'on a les relations:

$$\begin{aligned} \nabla J(u_m) &= \nabla J(u_{m-1}) + Aw_m \\ r_m &= r_{m-1} - Aw_m \end{aligned}$$

Or $r_0 = -\nabla J(u_0)$. Ce qui implique que $[\nabla J(u_0)] = \mathcal{K}_1$. De plus, $u_1 \in u_0 + \mathcal{K}_1$.
De même, en utilisant les relations ci-dessus, on a que $w_1 = \alpha_0 r_0 + \alpha_1 r_1 \in [r_0, Ar_0] = \mathcal{K}_2$ et $u_2 = u_1 + w_1 \in u_1 + [r_0, Ar_0] = u_0 + \mathcal{K}_2$ et ainsi de suite. Par conséquent, l'itéré u_m vérifie

$$J(u_m) = \min_{\alpha_i} J(u_{m-1} + \sum_{i=0}^{m-1} \alpha_i \nabla J(u_i)) = \min_{v \in \mathcal{K}_m} J(u_0 + v)$$

2.2 Procédé d'Arnoldi

Le but est de construire une base orthonormale du sous-espace de Krylov $\mathcal{K}_m(A, r_0)$. Cette base sera par la suite utilisée pour résoudre les problèmes de moindres carrés rencontrés ci-dessus. Il est clair que la factorisation complète donnée par Householder est à éviter car $n >> m$. Le calcul de Q est trop coûteux et gourmant en mémoire car il faudrait stocker la matrice pleine d'ordre n.
Le procédé d'Arnoldi est un procédé d'othogonalisation analogue à Gram-Schmidt mais qui est adapté à la structure particulière des sous-espaces de Krylov.

Procédé d'Arnoldi-GS:

$$\begin{cases} v_1 &= r_0/\|r_0\| \\ \\ \tilde{v}_{j+1} &= Av_j - \sum_{i=1}^{j} (Av_j, v_i) v_i \\ \\ v_{j+1} &= \tilde{v}_{j+1}/\|\tilde{v}_{j+1}\| \end{cases}$$

Il est facile de vérifier que ce procédé conduit à une base orthonormale de $\mathcal{K}_m(A, r_0)$. Par contre, il est instable tout comme le procédé original de Gram-Schmidt. On préfère utiliser une variante connue sous le nom de procédé de Gram-Schmidt modifié. Ce procédé est basé sur l'observation que

$$\tilde{v}_{j+1} = P_E A v_j$$

où $E = [v_1, v_2, v_3, \ldots, v_j]^\perp$, i.e. le sous-espace engendré par les j vecteurs v_1 à v_j et P_E dénote le projecteur sur le sous-espace E. Or ce projecteur peut être obtenu d'une manière différente:

$$P_E = P_{[v_j]^\perp} \ldots P_{[v_2]^\perp} P_{[v_1]^\perp}$$

Ceci conduit à l'algorithme de Gram-Schmidt modifié

2.3 Procédé d'Arnoldi-MGS:

1. $v_1 = r_0/\|r_0\|$
2. Pour $j = 1, \ldots, m$; faire:
$$w = Av_j$$
3. Pour $i = 1, \ldots, j$; faire:
$$w = w - (w, v_i)v_i$$
4. $v_{j+1} = w/\|w\|$

Cette algorithme est moins sensible aux erreurs d'arrondie.

Introduisons quelques notations.
On notera par
$$V_m = [v_1, v_2, v_3, \ldots, v_m]$$
la matrice *orthogonale* d'ordre $n \times m$ et
$$\bar{H}_m = (h_{ij}) \text{ avec } h_{ij} = \begin{cases} (Av_j, v_i) & i \leq j \\ \|\tilde{v}_{j+1}\| & i = j+1 \end{cases}$$
La matrice $\bar{H}_m$ est une matrice de Heissenberg de la forme
$$\bar{H}_m = \begin{pmatrix} h_{11} & h_{12} & h_{13} & h_{14} & \ldots & h_{1m} \\ h_{21} & h_{22} & h_{23} & h_{24} & \ldots & h_{2m} \\ 0 & h_{32} & h_{33} & h_{34} & \ldots & h_{3m} \\ \vdots & 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & h_{m,m-1} & h_{m,m} \\ 0 & 0 & 0 & \ldots & 0 & h_{m+1,m} \end{pmatrix}$$
La sous matrice formée des m premières lignes sera notée par H_m.
L'intérêt de ces matrices proviennent du fait qu'elles permettent une écriture matricielle du procédé d'Arnoldi.

Relation importante:

$$AV_m = V_{m+1}\bar{H}_m = V_m H_m + h_{m+1,m} v_{m+1} e_m^t$$

En effet, on a que:

$$\begin{aligned}
\tilde{v}_{j+1} &= Av_j - \sum_{i=1}^{j} h_{i,j} v_i \\
Av_j &= \tilde{v}_{j+1} + \sum_{i=1}^{j} h_{i,j} v_i \\
Av_j &= h_{j+1,j} v_{j+1} + \sum_{i=1}^{j} h_{i,j} v_i.
\end{aligned}$$

C'est-à-dire que les Av_j sont des combinaisons linéaires des $v_1, v_2, \ldots, v_{m+1}$ dont les coefficients sont donnés par la matrice $\bar{H}_m$.

3 Méthode GMRES

Considérons un système linéaire inversible

$$Au = b.$$

Notons par u_0 le vecteur initial (par exemple, $u = 0$) et le résidu initial $r_0 = b - Au_0$.

L'algorithme GMRES est une méthode de projection basée sur les choix de $K_m = \mathcal{K}_m(A, r_0)$ et $L_m = AK_m$.
D'après la section 1, ceci est équivalent à choisir comme solution approchée le vecteur $u_m = u_0 + v_m$ minimisant le résidu

$$\min_{v \in K_m} \|b - A(u_0 + v)\| = \min_{v \in K_m} \|r_0 - Av\| \tag{4}$$

par rapport aux vecteurs dans $u_0 + K_m$, d'où l'appellation de "Generalized Minimal RESidual method".

La solution du problème de moindres carrés (4) se fait par le procédé d'Arnoldi qui construit une base orthonormale du sous-espace de Krylov $\mathcal{K}_m(A, r_0)$.

Conséquences:

- Le résidu de la solution vérifie que
$$\|r_m\| = \|b - A(u_m)\| = \min_{v \in K_m} \|b - A(u_0 + v)\|.$$
- $\|r_{m+1}\| \leq \|r_m\|$
car $\mathcal{K}_m(A, r_0) \subset \mathcal{K}_{m+1}(A, r_0)$.
- L'algorithme converge en au plus n itérations car $\mathcal{K}_m(A, r_0) = R^n$.

Résolution du problème de moindres carrés:

On pose que $v = V_m y$. On a que

$$\begin{aligned} \|r_0 - Av\| &= \|r_0 - AV_m y\| \\ &= \|r_0 - V_{m+1}\bar{H}_m y\| \\ &= \|V_{m+1}(\beta e_1 - \bar{H}_m y)\| \\ &= \|\beta e_1 - \bar{H}_m y\| \end{aligned}$$

Ci-dessus, on a utilisé le fait que les vecteurs de V_m sont orthonormaux, $r_0 = \|r_0\| v_1$ et on a posé $\beta = \|r_0\|$.

Ainsi, le problème de moindres carrés (4) posé sur le sous-espace de Krylov revient à résoudre le problème de moindres carrés sur R^m

$$\min_{y \in R^m} \|\beta e_1 - \bar{H}_m y\|$$

Pour résoudre ce problème, on effectue une factorisation QR à l'aide des matrices de Givens (car la matrice $\bar{H}_m$ est de Heissenberg)

$$\bar{H}_m = Q_m R_m = \bar{Q}_m \bar{R}_m.$$

On a noté par $\bar{Q}_m \bar{R}_m$ la factorisation complète tandis que $Q_m R_m$ est la factorisation réduite i.e R_m est d'ordre $m \times m$. On obtient

$$\begin{aligned} \|\beta e_1 - \bar{H}_m y\| &= \|\beta \bar{Q}_m \bar{Q}_m^t e_1 - \bar{Q}_m \bar{R}_m y\| \\ &= \|\bar{Q}_m(\bar{Q}_m^t \beta e_1 - \bar{R}_m y)\| \\ &= \|\bar{Q}_m^t \beta e_1 - \bar{R}_m y\| \\ &= \|Q_m^t \beta e_1 - R_m y\| \end{aligned}$$

Le minimum est atteint en

$$y_m = R_m^{-1} g_m \qquad \text{où } g_m = Q_m^t \beta e_1.$$

Revenons au calcul de la norme du résidu à l'itération m.
On a que

$$\begin{array}{rcl}
\|r_m\| & = & \|b - Au_m\| = \|r_0 - Av_m\| \\
& = & \|r_0 - AV_m y_m\| \\
& = & \|r_0 - V_m H_m y_m - h_{m+1,m} v_{m+1} e_m^t\| \\
& = & \|V_m(\beta e_1 - H_m y) - h_{m+1,m} v_{m+1} e_m^t\| \\
& = & |h_{m+1,m}| \|v_{m+1}\| |(e_m, y_m)| \\
& = & h_{m+1,m} |(e_m, y_m)| \\
& \leq & h_{m+1,m} \|y_m\| \leq cte \;\; h_{m+1,m}
\end{array}$$