La méthode GMRES (Generalized Minimal RESidual method)

- Si la matrice A ds le système $Ax=b$ est inversible non définie positive et non symétrique on utilise la méthode GMRES.
- L'idée et de construire une suite X_k qui minimise la norme Euclidienne du résidu $b-AX$ ds le ss espace affine X_0+K_k ($K_k = \text{Vect}(r_0, Ar_0, \ldots, A^{k-1}r_0)$ le $k^{\text{ième}}$ ss espace de Krylov engendré par r_0) i.e :

$$\boxed{\|b-AX_k\|_2 = \inf_{X \in X_0+K_k} \{\|b-AX\|_2\}}$$

Le vecteur X_k est appelé le $k^{\text{ième}}$ vecteur itéré de la méthode GMRES.(*) Tout d'abord nous construisons l'espace de Krylov indépendamment de la recherche de la solution approchée X_k ds cet espace, grâce à l'algorithme d'Arnoldi. (i.e construire la base $(v_1, \ldots, v_k)$)

- Pour $0 \leq k \leq n$ on note $r_k = b - AX_k$, on pose

$$w_1 = r_0 \,, \qquad h_{1,0} = \beta = \|r_0\|_2 \text{ et puis pour } k = 1, 2, \ldots$$

on construit par récurrence des vecteurs v_k tels que $\|v_k\|_2 = 1$ selon l'algorithme d'Arnoldi :

(*) Les vecteurs $r_0, Ar_0, \ldots, A^{k-1}r_0$ sont presque linéairement dependents. Arnoldi a construit une autre base (orthonormée) de K_k formée par les vecteurs $\longrightarrow (v_1, \ldots, v_k)$

Algorithme d'Arnoldi

- Si $\omega_k = 0 \Rightarrow$ $\cancel{x_k = x^*}$ ~~et~~ on arrête l'itération.
- Si $\omega_k \neq 0$ on pose $v_k = \dfrac{\omega_k}{\|\omega_k\|_2}$

$$h_{i,k} = (v_i, Av_k) \qquad 1 \leq i \leq k$$

$$\omega_{k+1} = Av_k - \sum_{i=1}^{k} h_{i,k} \cdot v_i$$

$$h_{k+1,k} = \|\omega_{k+1}\|_2$$

GMRES Version d'Arnoldi

- On definit $\tilde{H} = (h_{ij})_{\substack{1 \leq i \leq k+1 \\ 1 \leq j \leq k}}$ matrice rectangulaire
- Calculer $z_k : \min_y \{\|\beta e_1 - \tilde{H}_k y\|_2\}$ (Π^de des moindre carrées QR)
- $X_k = X_0 + V_k z_k$ $\qquad V_k = [v_1, \ldots, v_k]$ matrice rectangulaire

<u>Propriétés</u> : 1) Si les vecteurs ω_l ; $l = 2, k+1$ sont non nuls alors la famille de vecteurs $(v_1, \ldots, v_{k+1})$ est une <u>base orthonormée</u> de K_{k+1} qui vérifie :

(i) $(v_i, Av_j) = h_{ij}$ si $1 \leq i \leq j \leq k+1$

(ii) $(v_{j+1}, Av_j) = h_{j+1,j}$ si $1 \leq j \leq k$

(iii) $(v_i, Av_j) = 0$ si $j+2 \leq i \leq k+1$

<u>Preuve</u> : Puisque $X_0 \neq X^*$, on a $r_0 \neq 0$ et donc on definit $v_1 = \dfrac{r_0}{\|r_0\|_2}$. Ainsi $v_1 \in K_1$ et $\|v_1\|_2 = 1$.

Soit $1 \leq l \leq k$. Supposons que $(v_1, \ldots, v_l)$ soit une base orthonormée de K_l. Par construction, le vecteur $w_{l+1} \in AK_l \subset K_{l+1}$ et vérifie, pour $1 \leq j \leq l$:

$$(v_j, w_{l+1}) = (v_j, Av_l) - \sum_{i=1}^{k} h_{i,l}(v_j, v_i)$$
$$= (v_j, Av_l) - h_{j,l} = 0. \quad (\text{par déf de } h_{j,l})$$

Par conséquent, le vecteur $v_{l+1} = \dfrac{w_{l+1}}{\|w_{l+1}\|_2} \in K_{l+1}$ et vérifie

$(v_j, v_{l+1}) = 0$ pour $1 \leq j \leq l$.

Les vecteurs $v_1, \ldots, v_{l+1}$ sont de norme 1 et 2 à 2 orthogonaux. Ils forment donc un système orthonormé de $l+1$ vecteurs de K_{l+1}, qui par construction, est un ss espace vect. de $\mathbb{R}^n$ de dim. au plus égale à $l+1$. Le système $(v_1, \ldots, v_{l+1})$ est donc une base orthonormée de K_{l+1}.

(i) Par déf. on a $(v_i, Av_j) = h_{ij}$ pour $1 \leq i \leq j \leq k+1$

De plus $(v_{j+1}, Av_j) = (v_{j+1}, w_{j+1}) + \sum_{i=1}^{j} h_{i,j}(v_{j+1}, v_i) = \|w_{j+1}\|_2 = h_{j+1,j}$

car $(v_1, \ldots, v_{j+1})$ est une base orthonormée et $v_{j+1} = \dfrac{w_{j+1}}{\|w_{j+1}\|_2}$ donc (ii)

Enfin, si $1 \leq i, j \leq k+1$ et si $j+2 \leq i$ alors $Av_j \in AK_j \subset K_{j+1} \subset K_{i-1}$ donc il est $\perp$ à v_i et (iii) en résulte. ∎

Nous définissons la matrice rectangulaire

$$V_k = [v_1, \dots, v_k] \in \mathbb{R}^{n \times k}$$

et la matrice $\tilde{H}_k$ $\underbrace{(\text{sous})}$ la forme :

H_k (matrice de Hessenberg) Supérieure.

$$\tilde{H}_k = \begin{pmatrix} h_{1,1} & \dots & \dots & h_{1k} \\ h_{2,1} & h_{2,2} & \dots & h_{2,k} \\ 0 & h_{32} & h_{33} \dots & h_{3,k} \\ \vdots & & \ddots & h_{kk} \\ 0 & \dots & 0 & h_{k+1,k} \end{pmatrix} \in \mathbb{R}^{k+1,k}$$

où la definition des h_{ij} est donneé par l'algorithme d'Arnoldi. Comme on a

$$A v_j = \sum_{1 \leq i \leq j+1} h_{ij} v_i$$

on déduit la propriété $\boxed{A V_k = V_{k+1} \tilde{H}_k}$. (*) voir notes ③

- On cherche $X_k \in X_0 + \text{Vect}(v_1, \dots, v_k)$ tel que r_k soit le plus petit possible en norme Euclidienne, i.e X_k peut s'écrire sous la forme :

$$\boxed{X_k = X_0 + V_k z_k}, \quad z_k \in \mathbb{R}^k \text{ sont les coef}^{\text{ts}} \text{ à déterminer.}$$

(voir note ②)

D'où

$$r_k = b - AX_k = r_0 - AV_k z_k = r_0 - V_{k+1} \tilde{H}_k z_k = V_{k+1}(d_{k+1} - \tilde{H}_k z_k)$$

où on a noté $d_{k+1} = (\underbrace{\|r_0\|_2}_{=\beta}, 0, \dots, 0)^T \in \mathbb{R}^{k+1} = \beta e_1$

<u>Lemme</u> : Le minimum de r_k est caractérisé par :

$$\|d_{k+1} - \tilde{H}_k z^k\|_2 \leq \|d_{k+1} - \tilde{H}_k w_k\|_2, \quad z_k \in \mathbb{R}^k, \ \forall w_k \in \mathbb{R}^k.$$

En effet : $\|r_k\|_2 = \|V_{k+1}(d_{k+1} - \tilde{H}_k w_k)\|_2 = \|d_{k+1} - \tilde{H}_k w_k\|_2$ car les (v_i) sont orthonormés

• La solution de ce problème de minimisation est bien connue et s'obtient grace à la méthode de factorisation QR :

$$\tilde{H}_k = Q_k \cdot R_k \qquad Q_k \in \mathbb{R}^{k+1 \times k+1},\ R_k \in \mathbb{R}^{k+1 \times k}$$

$(Q_k^T = Q_k^{-1})$: Q_k est une matrice unitaire et R_k est triangulaire supérieur

On a $R_k = \begin{pmatrix} \tilde{R}_k \\ 0 \end{pmatrix}$ avec $\tilde{R}_k \in \mathbb{R}^{k \times k}$ (matrice carrée)

Une méthode pratique pour la résolution du problème de moindre carrés est l'utilisation de la rotation de Givens. On definit les matrices de Givens par

pour éliminer la colonne endessous la diagonale principale de $\tilde{H}_k$.

$$\Omega_i = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & 0 & \\ & & c_i & s_i & & \\ & & -s_i & c_i & & \\ & 0 & & & 1 & \\ & & & & & \ddots \ 1 \end{pmatrix} \begin{array}{l} \\ \\ \leftarrow \text{ligne } i \\ \leftarrow \text{ligne } i+1 \\ \\ \\ \end{array} \quad (k+1)\times(k+1)$$

avec $c_i^2 + s_i^2 = 1$ où $s_i = \dfrac{h_{i+1,i}}{\sqrt{(h_{ii}^{(i-1)})^2 + h_{i+1,i}^2}}$, $c_i = \dfrac{h_{ii}^{(i-1)}}{\sqrt{(h_{ii}^{(i-1)})^2 + h_{i+1,i}^2}}$

$$Q_k = \Omega_k \Omega_{k-1} \cdots \Omega_1$$

donc $\tilde{R}_k = \tilde{H}_k^{(k)} = Q_k \tilde{H}_k$

$$\tilde{g}_k = Q_k(\beta e_1) = (\gamma_1, \dots, \gamma_{k+1})^T$$

$$\min \| \beta e_1 - \tilde{H}_k y \|_2 = \min \| \tilde{g}_k - \tilde{R}_k y \|_2$$

puisque Q_k est unitaire. $y_k = \tilde{R}_k^{-1} \hat{g}_k$. (sys. triangulaire sup)

② • Théorème : Si A est inversible, alors $\exists$ $0 < k < n-1$ tel que $X_k = X^*$.

① • Lemme : Si A est inversible, que $h_{j+1,j} \neq 0$ $1 \leq j \leq k$ et que $h_{k+1,k} = 0$, alors $X_k = X^*$.

Notes

① • Méthode de l'eqt normale : $A^*Ax = A^*b$.

$H = I$, $K = I$, $N = A^*A$.

• Méthode de l'erreur minimale

$H = (AA^*)^{-1}$, $K = A^*A$, $L = A^*$

$(AA^*y = b,\ x = A^*y)$

• Méthode ~~de l'erreur minimale~~ du résidu conjugué généralisé : $AA^*y = b$, $x = A^*y$

$H = I$, $K = (A^{-1})^*$

② • $z \in K_k \Leftrightarrow (y_1, \dots, y_k)$ des réels uniques tels que

$$z = \sum_{i=1}^{k} y_i v_i$$

$\Leftrightarrow$ posons $s = (y_1, \dots, y_n)^T$

$$z = V_k s = \begin{pmatrix} | & | & & \\ v_1 & v_2 & \cdots & v_k \\ | & | & & \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

③ •

Les matrices AV_k et $V_{k+1}\tilde{H}_k$ sont égales.

En effet, pour $1 \le j \le k$: La $j^{\text{ème}}$ colonne de $V_{k+1}\tilde{H}_k = \sum_{i=1}^{k+1} h_{ij} v_i$

$= \sum_{i=1}^{j+1} h_{ij} v_i$ du fait de (iii).

Or la $j^{\text{ème}}$ colonne de $AV_k = Av_j = w_{j+1} + \sum_{i=1}^{j} h_{ij} v_i$

$$= \sum_{i=1}^{j+1} h_{ij} v_i$$

$$\boxed{w_{j+1} = h_{j+1,j} \cdot v_{j+1} = \|w_{j+1}\|_2 \cdot \frac{w_{j+1}}{\|w_{j+1}\|}}$$

Ainsi on a $AV_k = V_{k+1}\tilde{H}_k$.