

Método GMRES

M. Sepúlveda

10 de Junio, 2021



Método de Residuos Mínimos Generalizado

- En contraste con CG, Residuos Mínimos Generalizado (GMRES) es usado para sistemas no simétricos.



Método de Residuos Mínimos Generalizado

- En contraste con CG, Residuos Mínimos Generalizado (GMRES) es usado para sistemas no simétricos.
- En contraste con CG, Residuos Mínimos Generalizado (GMRES) es usado para sistemas no simétricos. El propósito de GMRES es el de minimizar la norma $\|\cdot\|_2$ del residuo sobre K_k . Es decir la GMRES iteración está caracterizada por

$$\mathbf{x}_k \in \mathbf{x}_0 + K_k(\mathbf{r}_0, \mathbf{A})$$

$$\|b - \mathbf{A}\mathbf{x}_k\| = \min \{\|b - \mathbf{A}\mathbf{u} \mid \mathbf{u} \in \mathbf{x}_0 + K_k(\mathbf{r}_0, \mathbf{A})\|\}$$

$$\text{con } K_k(\mathbf{r}_0, \mathbf{A}) = \text{span} [\mathbf{r}_0, \mathbf{A}\mathbf{r}_0, \dots, \mathbf{A}^{k-1}\mathbf{r}_0]$$



Método de Residuos Mínimos Generalizado

- En contraste con CG, Residuos Mínimos Generalizado (GMRES) es usado para sistemas no simétricos.
- En contraste con CG, Residuos Mínimos Generalizado (GMRES) es usado para sistemas no simétricos. El propósito de GMRES es el de minimizar la norma $\|\cdot\|_2$ del residuo sobre K_k . Es decir la GMRES iteración está caracterizada por

$$\mathbf{x}_k \in \mathbf{x}_0 + K_k(\mathbf{r}_0, \mathbf{A})$$

$$\|b - \mathbf{A}\mathbf{x}_k\| = \min \{\|b - \mathbf{A}\mathbf{u} \mid \mathbf{u} \in \mathbf{x}_0 + K_k(\mathbf{r}_0, \mathbf{A})\|\}$$

$$\text{con } K_k(\mathbf{r}_0, \mathbf{A}) = \text{span} [\mathbf{r}_0, \mathbf{A}\mathbf{r}_0, \dots, \mathbf{A}^{k-1}\mathbf{r}_0]$$

- $\mathbf{r}_0 \neq 0$ implica $1 \leq \dim K_k(\mathbf{r}_0, \mathbf{A}) \leq k$.



Método de Residuos Mínimos Generalizado

- En contraste con CG, Residuos Mínimos Generalizado (GMRES) es usado para sistemas no simétricos.
- En contraste con CG, Residuos Mínimos Generalizado (GMRES) es usado para sistemas no simétricos. El propósito de GMRES es el de minimizar la norma $\|\cdot\|_2$ del residuo sobre K_k . Es decir la GMRES iteración está caracterizada por

$$\mathbf{x}_k \in \mathbf{x}_0 + K_k(\mathbf{r}_0, \mathbf{A})$$

$$\|b - \mathbf{A}\mathbf{x}_k\| = \min \{\|b - \mathbf{A}\mathbf{u} \mid \mathbf{u} \in \mathbf{x}_0 + K_k(\mathbf{r}_0, \mathbf{A})\|\}$$

$$\text{con } K_k(\mathbf{r}_0, \mathbf{A}) = \text{span} [\mathbf{r}_0, \mathbf{A}\mathbf{r}_0, \dots, \mathbf{A}^{k-1}\mathbf{r}_0]$$

- $\mathbf{r}_0 \neq 0$ implica $1 \leq \dim K_k(\mathbf{r}_0, \mathbf{A}) \leq k$.
- Existe un m mínimo tal que $K_k(\mathbf{r}_0, \mathbf{A})$ es \mathbf{A} -invariante, esto es $\mathbf{A}K_m(\mathbf{r}_0, \mathbf{A}) \subset K_m(\mathbf{r}_0, \mathbf{A})$ o equivalentemente $\mathbf{A}^m\mathbf{r}_0 \in K_m(\mathbf{r}_0, \mathbf{A})$ de modo que

$$\dim K_{m+1}(\mathbf{r}_0, \mathbf{A}) = \dim K_m(\mathbf{r}_0, \mathbf{A}) < m + 1.$$



Ortonormalización por el Método de Arnoldi

- Una característica clave del método GMRES es que usa vectores ortonormales para generar los espacios $K_k(\mathbf{r}_0, \mathbf{A})$, para $k = 1, \dots, m$. Para ello se usa el algoritmo de Alnordi (1951):



Ortonormalización por el Método de Arnoldi

- Una característica clave del método GMRES es que usa vectores ortonormales para generar los espacios $K_k(\mathbf{r}_0, \mathbf{A})$, para $k = 1, \dots, m$. Para ello se usa el algoritmo de Alnordi (1951):

Dado $\mathbf{r}_0 \neq 0$, sea $\beta = \|\mathbf{r}_0\|$, $\mathbf{v}_1 = \mathbf{r}_0/\beta$.
para $k = 1, 2, \dots$

$$\mathbf{u} = \mathbf{A}\mathbf{v}_k.$$

para $i = 1, 2, \dots, k$

$$| \quad h_{ik} = \mathbf{v}_i^T \mathbf{u}.$$

$$\mathbf{w}_k = \mathbf{u} - \sum_{i=1}^k h_{ik} \mathbf{v}_i, \text{ y } h_{k+1,k} = \|\mathbf{w}_k\|$$

Si $h_{k+1,k} = 0$, entonces $m = k$, y stop.

Si no

$$| \quad \mathbf{v}_{k+1} = \mathbf{w}_k / h_{k+1,k}.$$

• Algoritmo:



Ortonormalización por el Método de Arnoldi

- Una característica clave del método GMRES es que usa vectores ortonormales para generar los espacios $K_k(\mathbf{r}_0, \mathbf{A})$, para $k = 1, \dots, m$. Para ello se usa el algoritmo de Alnordi (1951):



Ortonormalización por el Método de Arnoldi

- Una característica clave del método GMRES es que usa vectores ortonormales para generar los espacios $K_k(\mathbf{r}_0, \mathbf{A})$, para $k = 1, \dots, m$. Para ello se usa el algoritmo de Alnordi (1951):

Dado $\mathbf{r}_0 \neq 0$, sea $\beta = \|\mathbf{r}_0\|$, $\mathbf{v}_1 = \mathbf{r}_0/\beta$.
para $k = 1, 2, \dots$

$$\mathbf{u} = \mathbf{A}\mathbf{v}_k.$$

para $i = 1, 2, \dots, k$

$$| \quad h_{ik} = \mathbf{v}_i^T \mathbf{u}.$$

$$\mathbf{w}_k = \mathbf{u} - \sum_{i=1}^k h_{ik} \mathbf{v}_i, \text{ y } h_{k+1,k} = \|\mathbf{w}_k\|$$

Si $h_{k+1,k} = 0$, entonces $m = k$, y stop.

Si no

$$| \quad \mathbf{v}_{k+1} = \mathbf{w}_k / h_{k+1,k}.$$

• Algoritmo:



La base ortonormal del Espacio de Krylov

- El método de Arnoldi termina con el mismo índice m de la \mathbf{A} -invarianza de $K_m(\mathbf{r}_0, \mathbf{A})$ (dem. en detalle en Stoer). De ello se deduce por inducción que el método de Arnoldi genera vectores $\mathbf{v}_1, \dots, \mathbf{v}_m$ tal que

$$\text{span}[\mathbf{v}_1, \dots, \mathbf{v}_k] = K_k(\mathbf{r}_0, \mathbf{A}) = \text{span}[\mathbf{r}_0, \mathbf{A}\mathbf{r}_0, \dots, \mathbf{A}^{k-1}\mathbf{r}_0]$$

para $k \leq m$.



La base ortonormal del Espacio de Krylov

- El método de Arnoldi termina con el mismo índice m de la \mathbf{A} -invarianza de $K_m(\mathbf{r}_0, \mathbf{A})$ (dem. en detalle en Stoer). De ello se deduce por inducción que el método de Arnoldi genera vectores $\mathbf{v}_1, \dots, \mathbf{v}_m$ tal que

$$\text{span}[\mathbf{v}_1, \dots, \mathbf{v}_k] = K_k(\mathbf{r}_0, \mathbf{A}) = \text{span}[\mathbf{r}_0, \mathbf{A}\mathbf{r}_0, \dots, \mathbf{A}^{k-1}\mathbf{r}_0]$$

para $k \leq m$.

- De modo que $m = \max\{k \geq 1 \mid \dim K_k(\mathbf{r}_0, \mathbf{A}) = k\}$.



La base ortonormal del Espacio de Krylov

- El método de Arnoldi termina con el mismo índice m de la \mathbf{A} -invarianza de $K_m(\mathbf{r}_0, \mathbf{A})$ (dem. en detalle en Stoer). De ello se deduce por inducción que el método de Arnoldi genera vectores $\mathbf{v}_1, \dots, \mathbf{v}_m$ tal que

$$\text{span}[\mathbf{v}_1, \dots, \mathbf{v}_k] = K_k(\mathbf{r}_0, \mathbf{A}) = \text{span}[\mathbf{r}_0, \mathbf{A}\mathbf{r}_0, \dots, \mathbf{A}^{k-1}\mathbf{r}_0]$$

para $k \leq m$.

- De modo que $m = \max\{k \geq 1 \mid \dim K_k(\mathbf{r}_0, \mathbf{A}) = k\}$.
- Las $n \times k$ matrices $\mathbf{V}_k = [\mathbf{v}_1, \dots, \mathbf{v}_k]$, con $k \leq m$, tienen columnas ortonormales de modo que $\mathbf{V}_k^T \mathbf{V}_k = \mathbf{I}_k$, formando una base de $K_k(\mathbf{r}_0, \mathbf{A})$,

$$K_k(\mathbf{r}_0, \mathbf{A}) = \{\mathbf{V}_k \mathbf{y} \mid \mathbf{y} \in \mathbb{R}^k\}$$



La base ortonormal del Espacio de Krylov

- El método de Arnoldi termina con el mismo índice m de la \mathbf{A} -invarianza de $K_m(\mathbf{r}_0, \mathbf{A})$ (dem. en detalle en Stoer). De ello se deduce por inducción que el método de Arnoldi genera vectores $\mathbf{v}_1, \dots, \mathbf{v}_m$ tal que

$$\text{span}[\mathbf{v}_1, \dots, \mathbf{v}_k] = K_k(\mathbf{r}_0, \mathbf{A}) = \text{span}[\mathbf{r}_0, \mathbf{A}\mathbf{r}_0, \dots, \mathbf{A}^{k-1}\mathbf{r}_0]$$

para $k \leq m$.

- De modo que $m = \max\{k \geq 1 \mid \dim K_k(\mathbf{r}_0, \mathbf{A}) = k\}$.
- Las $n \times k$ matrices $\mathbf{V}_k = [\mathbf{v}_1, \dots, \mathbf{v}_k]$, con $k \leq m$, tienen columnas ortonormales de modo que $\mathbf{V}_k^T \mathbf{V}_k = \mathbf{I}_k$, formando una base de $K_k(\mathbf{r}_0, \mathbf{A})$,

$$K_k(\mathbf{r}_0, \mathbf{A}) = \{\mathbf{V}_k \mathbf{y} \mid \mathbf{y} \in \mathbb{R}^k\}$$

- Cada $\mathbf{x} \in \mathbf{x}_0 + K_k(\mathbf{r}_0, \mathbf{A})$, por lo tanto, tiene una representación de la forma $\mathbf{x} = \mathbf{x}_0 + \mathbf{V}_k \mathbf{y}$ con un vector único $\mathbf{y} \in \mathbb{R}^k$.



Matriz de Hessenberg

- Las recursiones de Arnoldi pueden formularse de manera compacta en términos de matrices de Hessenberg de $(k + 1) \times k$

$$\tilde{\mathbf{H}}_k := \begin{pmatrix} h_{11} & h_{12} & \cdots & h_{1k} \\ h_{21} & h_{22} & \cdots & h_{2k} \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & h_{kk} \\ 0 & \cdots & 0 & h_{k+1,k} \end{pmatrix}, \quad k \leq m,$$

y su submatriz \mathbf{H}_k eliminando la última columna de $\tilde{\mathbf{H}}_k$.



Matriz de Hessenberg similar a \mathbf{A}

- El algoritmo de Arnoldi implica que

$$\mathbf{A}\mathbf{v}_k = \sum_{i=1}^{k+1} h_{ik} \mathbf{v}_i = \sum_{i=1}^k h_{ik} \mathbf{v}_i + \mathbf{w}_k, \quad (k < m) \quad \text{y} \quad \mathbf{A}\mathbf{v}_m = \sum_{i=1}^m h_{im} \mathbf{v}_i.$$

Equivalente a

$$\mathbf{AV}_m = \mathbf{V}_m \mathbf{H}_m$$



Matriz de Hessenberg similar a \mathbf{A}

- El algoritmo de Arnoldi implica que

$$\mathbf{A}\mathbf{v}_k = \sum_{i=1}^{k+1} h_{ik} \mathbf{v}_i = \sum_{i=1}^k h_{ik} \mathbf{v}_i + \mathbf{w}_k, \quad (k < m) \quad \text{y} \quad \mathbf{A}\mathbf{v}_m = \sum_{i=1}^m h_{im} \mathbf{v}_i.$$

Equivalente a

$$\mathbf{AV}_m = \mathbf{V}_m \mathbf{H}_m$$

- Para $1 \leq k < m$ se tiene

$$\mathbf{AV}_k = \mathbf{V}_k \mathbf{H}_k + \mathbf{w}_k \mathbf{e}_k^T = \mathbf{V}_{k+1} \tilde{\mathbf{H}}_k.$$



Matriz de Hessenberg similar a \mathbf{A}

- El algoritmo de Arnoldi implica que

$$\mathbf{A}\mathbf{v}_k = \sum_{i=1}^{k+1} h_{ik} \mathbf{v}_i = \sum_{i=1}^k h_{ik} \mathbf{v}_i + \mathbf{w}_k, \quad (k < m) \quad \text{y} \quad \mathbf{A}\mathbf{v}_m = \sum_{i=1}^m h_{im} \mathbf{v}_i.$$

Equivalente a

$$\mathbf{AV}_m = \mathbf{V}_m \mathbf{H}_m$$

- Para $1 \leq k < m$ se tiene

$$\mathbf{AV}_k = \mathbf{V}_k \mathbf{H}_k + \mathbf{w}_k \mathbf{e}_k^T = \mathbf{V}_{k+1} \tilde{\mathbf{H}}_k.$$

- Para $1 \leq k \leq m$ se tiene (porque $\mathbf{V}_k^T \mathbf{w}_k = 0$)

$$\mathbf{H}_k = \mathbf{V}^T \mathbf{AV}_k$$



Matriz de Hessenberg similar a \mathbf{A}

- El algoritmo de Arnoldi implica que

$$\mathbf{A}\mathbf{v}_k = \sum_{i=1}^{k+1} h_{ik} \mathbf{v}_i = \sum_{i=1}^k h_{ik} \mathbf{v}_i + \mathbf{w}_k, \quad (k < m) \quad \text{y} \quad \mathbf{A}\mathbf{v}_m = \sum_{i=1}^m h_{im} \mathbf{v}_i.$$

Equivalentе a

$$\mathbf{AV}_m = \mathbf{V}_m \mathbf{H}_m$$

- Para $1 \leq k < m$ se tiene

$$\mathbf{AV}_k = \mathbf{V}_k \mathbf{H}_k + \mathbf{w}_k \mathbf{e}_k^T = \mathbf{V}_{k+1} \tilde{\mathbf{H}}_k.$$

- Para $1 \leq k \leq m$ se tiene (porque $\mathbf{V}_k^T \mathbf{w}_k = 0$)

$$\mathbf{H}_k = \mathbf{V}^T \mathbf{A} \mathbf{V}_k$$

- La matriz \mathbf{H}_m es no singular y rango $\tilde{\mathbf{H}}_k = k$, para $k < m$.



Reducción del problema de minimización del Residuo

- Las matrices $\tilde{\mathbf{H}}_k$, \mathbf{H}_k y \mathbf{V}_k permiten una determinación sencilla de la solución x_k del problema de mínimos residuos.



Reducción del problema de minimización del Residuo

- Las matrices $\tilde{\mathbf{H}}_k$, \mathbf{H}_k y \mathbf{V}_k permiten una determinación sencilla de la solución \mathbf{x}_k del problema de mínimos residuos.
- Como se señaló anteriormente, cada $\mathbf{x} \in \mathbf{x}_0 + K_k(\mathbf{r}_0, \mathbf{A})$ se puede escribir en la forma $\mathbf{x} = \mathbf{x}_0 + \mathbf{V}_k \mathbf{y}$ con un vector único $\mathbf{y} \in \mathbb{R}^k$.



Reducción del problema de minimización del Residuo

- Las matrices $\tilde{\mathbf{H}}_k$, \mathbf{H}_k y \mathbf{V}_k permiten una determinación sencilla de la solución \mathbf{x}_k del problema de mínimos residuos.
- Como se señaló anteriormente, cada $\mathbf{x} \in \mathbf{x}_0 + K_k(\mathbf{r}_0, \mathbf{A})$ se puede escribir en la forma $\mathbf{x} = \mathbf{x}_0 + \mathbf{V}_k \mathbf{y}$ con un vector único $\mathbf{y} \in \mathbb{R}^k$.
- Dado que $\mathbf{r}_0 = \beta \mathbf{v}_1 = \beta \mathbf{V}_{k+1} \tilde{\mathbf{e}}_1$ con $\tilde{\mathbf{e}}_1 = [1, 0, \dots, 0]^T \in \mathbb{R}^{k+1}$, se tiene para $k < m$:

$$\begin{aligned}\|\mathbf{b} - \mathbf{Ax}\|_2 &= \|\mathbf{b} - \mathbf{Ax}_0 - \mathbf{Av}_k \mathbf{y}\|_2 \\&= \|\mathbf{r}_0 - \mathbf{V}_{k+1} \tilde{\mathbf{H}}_k \mathbf{y}\|_2 \\&= \|\mathbf{V}_{k+1} (\beta \tilde{\mathbf{e}}_1 - \tilde{\mathbf{H}}_k \mathbf{y})\|_2 \\&= \|\beta \tilde{\mathbf{e}}_1 - \tilde{\mathbf{H}}_k \mathbf{y}\|_2.\end{aligned}$$



Reducción del problema de minimización del Residuo

- Las matrices $\tilde{\mathbf{H}}_k$, \mathbf{H}_k y \mathbf{V}_k permiten una determinación sencilla de la solución \mathbf{x}_k del problema de mínimos residuos.
- Como se señaló anteriormente, cada $\mathbf{x} \in \mathbf{x}_0 + K_k(\mathbf{r}_0, \mathbf{A})$ se puede escribir en la forma $\mathbf{x} = \mathbf{x}_0 + \mathbf{V}_k \mathbf{y}$ con un vector único $\mathbf{y} \in \mathbb{R}^k$.
- Dado que $\mathbf{r}_0 = \beta \mathbf{v}_1 = \beta \mathbf{V}_{k+1} \tilde{\mathbf{e}}_1$ con $\tilde{\mathbf{e}}_1 = [1, 0, \dots, 0]^T \in \mathbb{R}^{k+1}$, se tiene para $k < m$:

$$\begin{aligned}\|\mathbf{b} - \mathbf{Ax}\|_2 &= \|\mathbf{b} - \mathbf{Ax}_0 - \mathbf{Av}_k \mathbf{y}\|_2 \\ &= \|\mathbf{r}_0 - \mathbf{V}_{k+1} \tilde{\mathbf{H}}_k \mathbf{y}\|_2 \\ &= \|\mathbf{V}_{k+1} (\beta \tilde{\mathbf{e}}_1 - \tilde{\mathbf{H}}_k \mathbf{y})\|_2 \\ &= \|\beta \tilde{\mathbf{e}}_1 - \tilde{\mathbf{H}}_k \mathbf{y}\|_2.\end{aligned}$$

- La solución \mathbf{y}_k del problema de mínimos cuadrados

$$\min_{\mathbf{y} \in \mathbb{R}^k} \|\beta \tilde{\mathbf{e}}_1 - \tilde{\mathbf{H}}_k \mathbf{y}\|_2$$

Entrega la solución $\mathbf{x}_k = \mathbf{x}_0 + \mathbf{V}_k \mathbf{y}_k$.



Resultado del método

- Por otro lado, si $k = m$ se tiene que para $\mathbf{x} \in \mathbf{x}_0 + K_m(\mathbf{r}_0, \mathbf{A})$,

$$\begin{aligned}\|\mathbf{b} - \mathbf{Ax}\|_2 &= \|\mathbf{r}_0 - \mathbf{AV}_m\mathbf{y}\|_2 \\ &= \|\mathbf{V}_m(\beta\mathbf{e}_1 - \mathbf{H}_m\mathbf{y})\|_2 \\ &= \|\beta\mathbf{e}_1 - \mathbf{H}_m\mathbf{y}\|_2.\end{aligned}$$

Esto permite demostrar (Stoer)



Resultado del método

- Por otro lado, si $k = m$ se tiene que para $\mathbf{x} \in \mathbf{x}_0 + K_m(\mathbf{r}_0, \mathbf{A})$,

$$\begin{aligned}\|\mathbf{b} - \mathbf{Ax}\|_2 &= \|\mathbf{r}_0 - \mathbf{AV}_m\mathbf{y}\|_2 \\ &= \|\mathbf{V}_m(\beta\mathbf{e}_1 - \mathbf{H}_m\mathbf{y})\|_2 \\ &= \|\beta\mathbf{e}_1 - \mathbf{H}_m\mathbf{y}\|_2.\end{aligned}$$

Esto permite demostrar (Stoer)



Resultado del método

- Por otro lado, si $k = m$ se tiene que para $\mathbf{x} \in \mathbf{x}_0 + K_m(\mathbf{r}_0, \mathbf{A})$,

$$\begin{aligned}\|\mathbf{b} - \mathbf{Ax}\|_2 &= \|\mathbf{r}_0 - \mathbf{AV}_m\mathbf{y}\|_2 \\ &= \|\mathbf{V}_m(\beta\mathbf{e}_1 - \mathbf{H}_m\mathbf{y})\|_2 \\ &= \|\beta\mathbf{e}_1 - \mathbf{H}_m\mathbf{y}\|_2.\end{aligned}$$

Esto permite demostrar (Stoer)

Teorema

El vector \mathbf{x}_m resuelve $\mathbf{Ax} = \mathbf{b}$, $\mathbf{x}_m = \mathbf{A}^{-1}\mathbf{b}$ y $\mathbf{x}_k \neq \mathbf{A}^{-1}\mathbf{b}$ para todo $k < m$: el índice m es el primer índice para el que \mathbf{x}_k resuelve $\mathbf{Ax} = \mathbf{b}$.



Algoritmo

Algoritmo GMRES Básico

$v_1 = \frac{b}{\|b\|}$
for $k = 1 : M$,

$w = Av_k$

for $j = 1 : k$,

$h_{jk} = \langle v_j, w \rangle$

$w = w - h_{jk}v_j$

end

$h_{k+1,k} = \|w\|$

$v_{k+1} = \frac{w}{h_{k+1,k}}$

{Resolver para y_k , por mínimos cuadrados,

$H_k y_k = \beta e_1$

$x_k = V_k y_k$

end



Comentarios adicionales

- La resolución del problema de mínimos cuadrados $\tilde{\mathbf{H}}_k \mathbf{y}_k = \beta \tilde{\mathbf{e}}_1$ de $(k+1) \times k$, involucra una matriz de Hessenberg, y por lo tanto conviene factorizar en QR, usando rotaciones de Givens.



Comentarios adicionales

- La resolución del problema de mínimos cuadrados $\tilde{\mathbf{H}}_k \mathbf{y}_k = \beta \tilde{\mathbf{e}}_1$ de $(k+1) \times k$, involucra una matriz de Hessenberg, y por lo tanto conviene factorizar en QR, usando rotaciones de Givens.
- El método de Arnoldi al igual que Gram-Schmidt, no es muy estable y para k muy grande tiene a perderse la ortogonalidad de \mathbf{V}_k .



Comentarios adicionales

- La resolución del problema de mínimos cuadrados $\tilde{\mathbf{H}}_k \mathbf{y}_k = \beta \tilde{\mathbf{e}}_1$ de $(k+1) \times k$, involucra una matriz de Hessenberg, y por lo tanto conviene factorizar en QR, usando rotaciones de Givens.
- El método de Arnoldi al igual que Gram-Schmidt, no es muy estable y para k muy grande tiene a perderse la ortogonalidad de \mathbf{V}_k .
- El libro de Stoer y Bulirsch dá varias alternativas y soluciones de mejora: modificar Arnoldi; hacer GMRES incompleto cuasi-mínimo; Precondicionar.



Comentarios adicionales

- La resolución del problema de mínimos cuadrados $\tilde{\mathbf{H}}_k \mathbf{y}_k = \beta \tilde{\mathbf{e}}_1$ de $(k+1) \times k$, involucra una matriz de Hessenberg, y por lo tanto conviene factorizar en QR, usando rotaciones de Givens.
- El método de Arnoldi al igual que Gram-Schmidt, no es muy estable y para k muy grande tiene a perderse la ortogonalidad de \mathbf{V}_k .
- El libro de Stoer y Bulirsch dá varias alternativas y soluciones de mejora: modificar Arnoldi; hacer GMRES incompleto cuasi-mínimo; Precondicionar.
- El error residual relativo está acotado por:

$$\frac{\|\mathbf{r}_n\|}{\|\mathbf{b}\|} \leq \inf_{p \in P_n} \|p(A)\| \leq \kappa(V) \inf_{p \in P_n} \max_{\lambda \in \sigma(A)} |p(\lambda)|$$

que en el caso de una matriz simétrica se reduce a

$$\|\mathbf{r}_n\| \leq \left(\frac{\kappa(A)^2 - 1}{\kappa(A)^2} \right)^{n/2} \|\mathbf{r}_0\|$$



Comentarios adicionales

- La resolución del problema de mínimos cuadrados $\tilde{\mathbf{H}}_k \mathbf{y}_k = \beta \tilde{\mathbf{e}}_1$ de $(k+1) \times k$, involucra una matriz de Hessenberg, y por lo tanto conviene factorizar en QR, usando rotaciones de Givens.
- El método de Arnoldi al igual que Gram-Schmidt, no es muy estable y para k muy grande tiene a perderse la ortogonalidad de \mathbf{V}_k .
- El libro de Stoer y Bulirsch dá varias alternativas y soluciones de mejora: modificar Arnoldi; hacer GMRES incompleto cuasi-mínimo; Precondicionar.
- El error residual relativo está acotado por:

$$\frac{\|\mathbf{r}_n\|}{\|\mathbf{b}\|} \leq \inf_{p \in P_n} \|p(A)\| \leq \kappa(V) \inf_{p \in P_n} \max_{\lambda \in \sigma(A)} |p(\lambda)|$$

que en el caso de una matriz simétrica se reduce a

$$\|\mathbf{r}_n\| \leq \left(\frac{\kappa(A)^2 - 1}{\kappa(A)^2} \right)^{n/2} \|\mathbf{r}_0\|$$

- Código GMRES en Matlab/Octave

