



Numerical Methods of Linear Algebra for Sparse Matrices

Lecture 14

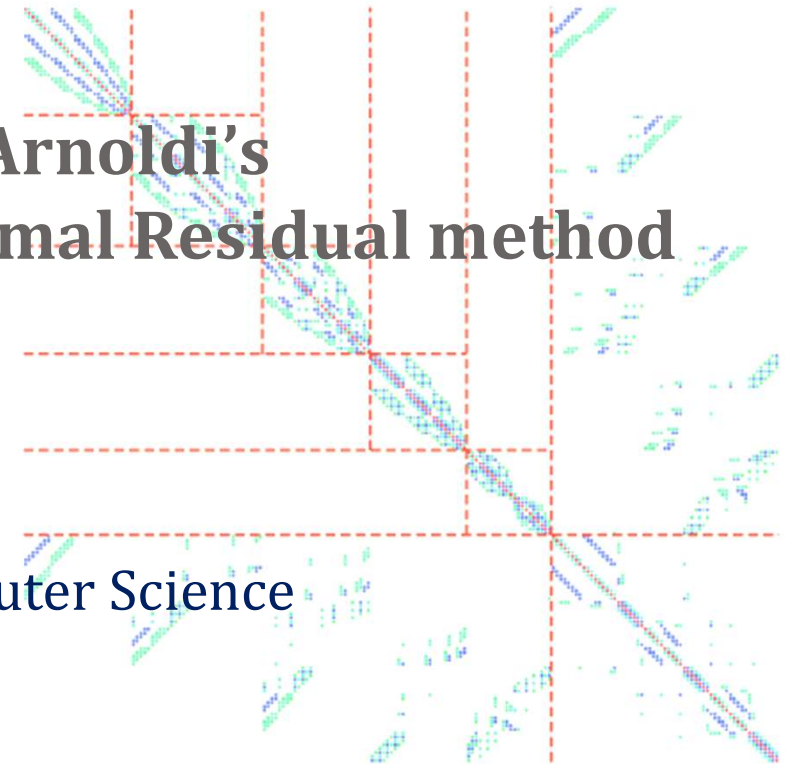
Krylov subspace methods based on Arnoldi's orthogonalization. Generalized Minimal Residual method (GMRES)

Anna Nasedkina

Department of Mathematical Modeling

Institute of Mathematics, Mechanics and Computer Science

Southern Federal University



Krylov Subspace methods based on Arnoldi's orthogonalization

Generalized Minimal Residual method (GMRES)

Methods for solving the linear system $Ax=b$ based on Arnoldi orthogonalization

- Classical examples of methods based on Arnold orthogonalization process are FOM (Full Orthogonalization Method) and **GMRES (Generalized Minimal Residual)**
- In these methods, the solution is searched in the form $x_m = x_0 + \delta_m$, $\delta_m \in K_m(A, r_0)$ is correction vector,
 $K_m(A, r_0) = \text{span}\{r_0, Ar_0, A^2r_0, \dots, A^{m-1}r_0\}$ is Krylov subspace
- Orthogonal basis of $K_m(A, r_0)$ is produced by Arnoldi process, where the basis vectors are the columns of V_m :
 $\delta_m \in K_m(A, r_0) \Leftrightarrow \exists y \in \mathbb{C}^m: \delta_m = V_m y$
- FOM uses the first part of Arnoldi relation: $AV_m = V_m H_m + w_m e_m^T$ and solves the linear system of the size m with matrix H_m
- **GMRES uses the second part of Arnoldi relation:** $AV_m = V_{m+1} \overline{H}_m$ and solves the linear system of size the $m + 1$ by m with matrix \overline{H}_m

Derivation of Generalized Minimal Residual method (GMRES)

Consider projection method with Krylov subspace as the search subspace.

Search subspace is Krylov subspace $K_m(A, r_0) = \text{span}\{r_0, Ar_0, A^2r_0, \dots, A^{m-1}r_0\}$

Subspace of constraints in GMRES is $L_m = AK_m = K_m(A, r_0)$

$V_m = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \dots & v_m \\ | & | & & | \end{pmatrix}$ is the basis in K_m

Approximate solution is given by $x_m = x_0 + \delta_m = x_0 + V_m y_m$; $\delta_m \in K_m(A, r_0)$

δ_m is chosen so that $\|r_m\|_2 \rightarrow \min$

We take the second part of equality in Arnoldi relation: $AV_m = V_{m+1} \overline{H}_m$

and the first basis vector as normalized initial residual: $v_1 = \frac{r_0}{\|r_0\|_2}$, $\beta = \|r_0\|_2$,

then $r_0 = \beta v_1$. Recall the formula for r_m and substitute the previous relations in it:

$$r_m = b - Ax_m = b - A(x_0 + V_m y_m) = b - Ax_0 - AV_m y_m = r_0 - AV_m y_m = \beta v_1 - V_{m+1} \overline{H}_m y_m$$

Derivation of GMRES

We got $r_m = \beta v_1 - V_{m+1} \overline{H}_m y_m$

We can represent v_1 as $v_1 = \begin{pmatrix} | & | & & | & | \\ v_1 & v_2 & \dots & v_m & v_{m+1} \\ | & | & & | & | \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \dots \\ 0 \end{pmatrix} = V_{m+1} e_1$ (e_1 is the unit ort)

Hence $r_m = \beta v_1 - V_{m+1} \overline{H}_m y_m = \beta V_{m+1} e_1 - V_{m+1} \overline{H}_m y_m = V_{m+1} (\beta e_1 - \overline{H}_m y_m)$

Here $\overline{H}_m \in \mathbb{C}^{(m+1) \times m}$, $y_m \in \mathbb{C}^m$, $e_1 \in \mathbb{R}^{m+1}$ is the unit ort

Consider the formula for residual norm.

$\|r_m\|_2 = \|V_{m+1} (\beta e_1 - \overline{H}_m y_m)\|_2 = \|\beta e_1 - \overline{H}_m y_m\|_2$, because $\|V_{m+1}\|_2 = 1$ as V_{m+1} has orthonormal columns.

$\|r_m\|_2 = \|\beta e_1 - \overline{H}_m y_m\|_2 \rightarrow \min$

Consider the problem of minimizing the functional $J(y) = \|\beta e_1 - \overline{H}_m y\|_2 \rightarrow \min$.

$y_m = \arg \min_y \|\beta e_1 - \overline{H}_m y\|_2$. It will be when $\|\beta e_1 - \overline{H}_m y_m\|_2 = 0$, i.e. $\overline{H}_m y_m = \beta e_1$

Small system in GMRES

We find y from solving the system $\overline{H}_m y = \beta e_1$.

Here $\overline{H}_m \in \mathbb{C}^{(m+1) \times m}$, $y \in \mathbb{C}^m$, $e_1 \in \mathbb{R}^{m+1}$ is the unit ort

This system is *overdefined*: it has $m+1$ equations and m unknowns

For overdefined system $\overline{H}_m y = \beta e_1$ we have least-squares problem: $\|\beta e_1 - \overline{H}_m y\|_2 \rightarrow \min$.

$y^T = y_m^T = (y_m^1, y_m^2, \dots, y_m^m)$ are unknowns in $m+1$ equations:

$$\left\{ \begin{array}{l} h_{11}y_m^1 + h_{12}y_m^2 + \dots + h_{1m}y_m^m = \beta \\ h_{21}y_m^1 + h_{22}y_m^2 + \dots + h_{2m}y_m^m = 0 \\ \dots \\ h_{m,m-1}y_m^{m-1} + h_{mm}y_m^m = 0 \\ h_{m+1,m}y_m^m = 0 \end{array} \right.$$

For example for $m = 4$ the matrix form of the system is:

$$\begin{pmatrix} h_{11} & h_{12} & h_{13} & h_{14} \\ h_{21} & h_{22} & h_{23} & h_{24} \\ 0 & h_{32} & h_{33} & h_{34} \\ 0 & 0 & h_{43} & h_{44} \\ 0 & 0 & 0 & h_{54} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} \beta \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Generalized Minimal Residual method (GMRES)

1. Take initial guess x_0 , calculate initial residual $r_0 = b - Ax_0$, calculate residual norm $\beta = \|r_0\|_2$, and the first vector $v_1 = \frac{r_0}{\beta}$
2. Define $\overline{H}_m = \{h_{ij}\} \in \mathbb{C}^{(m+1) \times m}$, $V_m \in \mathbb{C}^{n \times m}$, $V_{m+1} \in \mathbb{C}^{n \times (m+1)}$. Set $\overline{H}_m = 0$
3. Build the basis for Krylov subspace using Arnoldi process:
Loop for j from 1 to m .
 - 3.1) $w_j = Av_j$
 - 3.2) Loop for i from 1 to j $h_{ij} := (w_j, v_i)$, $w_j := w_j - h_{ij}v_i$
 - 3.3) $h_{j+1,j} := \|w_j\|_2$. If $h_{j+1,j} = 0 \Rightarrow m := j$, go to step 4, else
 - 3.4) $v_{j+1} := \frac{w_j}{h_{j+1,j}}$
4. Find y_m as a minimizer of the functional $J(y) = \|\beta e_1 - \overline{H}_m y\|_2$
(solve overdetermined system $\overline{H}_m y = \beta e_1$)
5. **Calculate residual r_m** , estimate residual norm: $\|r_m\|_2 < \varepsilon$.
6. Calculate new approximate solution $x_m = x_0 + V_m y_m$

The residual in GMRES can be calculated **without** general formula $r_m = b - Ax_m$

Solving overdefined system in GMRES

For overdefined system $\overline{H}_m y = \beta e_1$ we have least-squares problem:

$$\|\beta e_1 - \overline{H}_m y\|_2 \rightarrow \min$$

This problem can be solved by transforming the Hessenberg matrix \overline{H}_m into upper triangular form.

This can be done via:

- 1) Givens rotations
- 2) Housholder reflections.

The idea for both is to perform QR-factorization $\overline{H}_m = QR$, where $\overline{H}_m \in \mathbb{C}^{(m+1) \times m}$, $Q \in \mathbb{C}^{(m+1) \times (m+1)}$ is unitary and $R \in \mathbb{C}^{(m+1) \times m}$ is upper triangular.

In this lecture we consider *Givens rotations* to obtain QR-factorization for \overline{H}_m . Givens rotations will also give us the formula to calculate the residual norm $\|r_m\|_2$ inexpensively.

Solving overdefined system in GMRES

Consider the system $\overline{H}_m y = \beta e_1$ and QR-factorization $\overline{H}_m = QR$.

We have least-squares problem: $\|\beta e_1 - \overline{H}_m y\|_2 \rightarrow \min$

$$\begin{aligned} \|r_m\|_2 &= \|\beta e_1 - \overline{H}_m y\|_2 = \|\beta e_1 - QRy\|_2 = \|\beta Q Q^H e_1 - QRy\|_2 = \\ &= \|Q(\beta Q^H e_1 - Ry)\|_2 = \|\beta Q^H e_1 - Ry\|_2, \text{ because } Q \text{ is unitary and } \|Q\|_2 = 1 \end{aligned}$$

$$\|\beta Q^H e_1 - Ry\|_2 \rightarrow \min$$

Here $y \in \mathbb{C}^m$, $R \in \mathbb{C}^{(m+1) \times m}$ is upper triangular matrix with zero last row, $Q^H \in \mathbb{C}^{(m+1) \times (m+1)}$, $e_1 \in \mathbb{C}^{m+1}$, $\beta = \|r_0\|_2$ is a scalar $\Rightarrow \beta Q^H e_1 \in \mathbb{C}^{m+1}$

Solving overdefined system in GMRES

After Givens rotations the system $\overline{H}_m y = \beta e_1$ will be transformed into the system $Ry = \beta Q^H e_1$ with $m + 1$ equations and m unknowns in the following form:

$$\begin{pmatrix} \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & 0 & \times & \times \\ 0 & 0 & 0 & \times \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} y^1 \\ y^2 \\ \dots \\ y^m \end{pmatrix} = \begin{pmatrix} \times \\ \times \\ \dots \\ \times \\ \times \end{pmatrix}$$

The last element will be used to estimate residual norm

y^1, y^2, \dots, y^m are the unknown components of vector y , crosses \times represent nonzero elements

Right-hand side vector $\beta Q^H e_1$ is completely full, but the last $(m + 1)$ -th equation is zero (zero row of the matrix) corresponds to nonzero right-hand side. So this system cannot be solved exactly, but we can minimize

the norm $\|\beta Q^H e_1 - Ry\|_2$.

Givens rotations

Consider the overdetermined system that we solve in GMRES

$$\overline{H}_m y = \beta e_1 = \overline{g}_0$$

We'll apply Givens rotations for the case when $m = 4$ (in real space):

$$\overline{H}_4 y = \overline{g}_0$$

$$\begin{pmatrix} h_{11} & h_{12} & h_{13} & h_{14} \\ h_{21} & h_{22} & h_{23} & h_{24} \\ 0 & h_{32} & h_{33} & h_{34} \\ 0 & 0 & h_{43} & h_{44} \\ 0 & 0 & 0 & h_{54} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} \beta \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

We want to make under-diagonal $h_{21}, h_{32}, h_{43}, h_{54}$ elements zero.

To make four elements zero, we need to apply four rotations.

In general for m under-diagonal elements $h_{21}, \dots, h_{m,m-1}, h_{m+1,m}$

we need to apply m rotations.

Givens rotations: step 1

1. Apply the first rotation Ω_1 : for $\overline{H}_4 y = \overline{g}_0$ we multiply both sides by Ω_1

$$\Omega_1 \overline{H}_4 y = \Omega_1 \overline{g}_0$$

$$\overline{H}_4^{(1)} = \Omega_1 \overline{H}_4 = \begin{pmatrix} c_1 & s_1 & 0 & 0 & 0 \\ -s_1 & c_1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} h_{11} & h_{12} & h_{13} & h_{14} \\ h_{21} & h_{22} & h_{23} & h_{24} \\ 0 & h_{32} & h_{33} & h_{34} \\ 0 & 0 & h_{43} & h_{44} \\ 0 & 0 & 0 & h_{54} \end{pmatrix} = \begin{pmatrix} h_{11}^{(1)} & h_{12}^{(1)} & h_{13}^{(1)} & h_{14}^{(1)} \\ 0 & h_{22}^{(1)} & h_{23}^{(1)} & h_{24}^{(1)} \\ 0 & h_{32} & h_{33} & h_{34} \\ 0 & 0 & h_{43} & h_{44} \\ 0 & 0 & 0 & h_{54} \end{pmatrix}.$$

$h_{21}^{(1)} = -s_1 h_{11} + c_1 h_{21} = 0$ and $c_1^2 + s_1^2 = 1$. So we can take $d_1 = \sqrt{h_{11}^2 + h_{21}^2}$, $c_1 = \frac{h_{11}}{d_1}$, $s_1 = \frac{h_{21}}{d_1}$

to make $h_{21}^{(1)} = 0$

$$\overline{g}_1 = \Omega_1 \overline{g}_0 = \begin{pmatrix} c_1 & s_1 & 0 & 0 & 0 \\ -s_1 & c_1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} c_1 \beta \\ -s_1 \beta \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Givens rotations: step 2

2. Apply the second rotation Ω_2 : for $\overline{H}_4^{(1)} y = \overline{g}_1$ we multiply both sides by Ω_2

$$\Omega_2 \overline{H}_4^{(1)} y = \Omega_2 \overline{g}_1$$

$$\overline{H}_4^{(2)} = \Omega_2 \overline{H}_4^{(1)} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & c_2 & s_2 & 0 & 0 \\ 0 & -s_2 & c_2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} h_{11}^{(1)} & h_{12}^{(1)} & h_{13}^{(1)} & h_{14}^{(1)} \\ 0 & h_{22}^{(1)} & h_{23}^{(1)} & h_{24}^{(1)} \\ 0 & h_{32} & h_{33} & h_{34} \\ 0 & 0 & h_{43} & h_{44} \\ 0 & 0 & 0 & h_{54} \end{pmatrix} = \begin{pmatrix} h_{11}^{(2)} & h_{12}^{(2)} & h_{13}^{(2)} & h_{14}^{(2)} \\ 0 & h_{22}^{(2)} & h_{23}^{(2)} & h_{24}^{(2)} \\ 0 & 0 & h_{33}^{(2)} & h_{34}^{(2)} \\ 0 & 0 & h_{43} & h_{44} \\ 0 & 0 & 0 & h_{54} \end{pmatrix}.$$

$h_{32}^{(2)} = -s_2 h_{22}^{(1)} + c_2 h_{32} = 0$ and $c_2^2 + s_2^2 = 1$. So we can take $d_2 = \sqrt{(h_{22}^{(1)})^2 + h_{32}^2}$, $c_2 = \frac{h_{22}^{(1)}}{d_2}$, $s_2 = \frac{h_{32}}{d_2}$

to make $h_{32}^{(2)} = 0$

$$\overline{g}_2 = \Omega_2 \overline{g}_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & c_2 & s_2 & 0 & 0 \\ 0 & -s_2 & c_2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} c_1 \beta \\ -s_1 \beta \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} c_1 \beta \\ -c_2 s_1 \beta \\ s_2 s_1 \beta \\ 0 \\ 0 \end{pmatrix}$$

Givens rotations: step 3

3. Apply the third rotation Ω_3 : for $\overline{H}_4^{(2)} y = \overline{g}_2$ we multiply both sides by Ω_3

$$\Omega_3 \overline{H}_4^{(2)} y = \Omega_3 \overline{g}_2$$

$$\overline{H}_4^{(3)} = \Omega_3 \overline{H}_4^{(2)} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & c_3 & s_3 & 0 \\ 0 & 0 & -s_3 & c_3 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} h_{11}^{(2)} & h_{12}^{(2)} & h_{13}^{(2)} & h_{14}^{(2)} \\ 0 & h_{22}^{(2)} & h_{23}^{(2)} & h_{24}^{(2)} \\ 0 & 0 & h_{33}^{(2)} & h_{34}^{(2)} \\ 0 & 0 & h_{43} & h_{44} \\ 0 & 0 & 0 & h_{54} \end{pmatrix} = \begin{pmatrix} h_{11}^{(3)} & h_{12}^{(3)} & h_{13}^{(3)} & h_{14}^{(3)} \\ 0 & h_{22}^{(3)} & h_{23}^{(3)} & h_{24}^{(3)} \\ 0 & 0 & h_{33}^{(3)} & h_{34}^{(3)} \\ 0 & 0 & 0 & h_{44}^{(3)} \\ 0 & 0 & 0 & h_{54} \end{pmatrix}.$$

$$h_{43}^{(3)} = -s_3 h_{33}^{(2)} + c_3 h_{43} = 0 \text{ and } c_3^2 + s_3^2 = 1. \text{ So we can take } d_3 = \sqrt{(h_{33}^{(2)})^2 + h_{43}^2}, c_3 = \frac{h_{33}^{(2)}}{d_3}, s_3 = \frac{h_{43}}{d_3}$$

to make $h_{43}^{(3)} = 0$

$$\overline{g}_3 = \Omega_3 \overline{g}_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & c_3 & s_3 & 0 \\ 0 & 0 & -s_3 & c_3 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} c_1 \beta \\ -c_2 s_1 \beta \\ s_2 s_1 \beta \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} c_1 \beta \\ -c_2 s_1 \beta \\ c_3 s_2 s_1 \beta \\ -s_3 s_2 s_1 \beta \\ 0 \end{pmatrix}$$

Givens rotations: step 4

4. Apply the fourth and last rotation Ω_4 : for $\overline{H}_4^{(3)} y = \overline{g}_3$ we multiply both sides by Ω_4

$$\Omega_4 \overline{H}_4^{(3)} y = \Omega_4 \overline{g}_3$$

$$\overline{H}_4^{(4)} = \Omega_4 \overline{H}_4^{(3)} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & c_4 & s_4 \\ 0 & 0 & 0 & -s_4 & c_4 \end{pmatrix} \begin{pmatrix} h_{11}^{(3)} & h_{12}^{(3)} & h_{13}^{(3)} & h_{14}^{(3)} \\ 0 & h_{22}^{(3)} & h_{23}^{(3)} & h_{24}^{(3)} \\ 0 & 0 & h_{33}^{(3)} & h_{34}^{(3)} \\ 0 & 0 & 0 & h_{44}^{(3)} \\ 0 & 0 & 0 & h_{54} \end{pmatrix} = \begin{pmatrix} h_{11}^{(4)} & h_{12}^{(4)} & h_{13}^{(4)} & h_{14}^{(4)} \\ 0 & h_{22}^{(4)} & h_{23}^{(4)} & h_{24}^{(4)} \\ 0 & 0 & h_{33}^{(4)} & h_{34}^{(4)} \\ 0 & 0 & 0 & h_{44}^{(4)} \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

$$h_{54}^{(4)} = -s_4 h_{44}^{(3)} + c_4 h_{54} = 0 \text{ and } c_4^2 + s_4^2 = 1. \text{ So we can take } d_4 = \sqrt{(h_{44}^{(3)})^2 + h_{54}^2}, \quad c_4 = \frac{h_{44}^{(3)}}{d_4}, \quad s_4 = \frac{h_{54}}{d_4}$$

to make $h_{54}^{(4)} = 0$

$$\overline{g}_4 = \Omega_4 \overline{g}_3 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & c_4 & s_4 \\ 0 & 0 & 0 & -s_4 & c_4 \end{pmatrix} \begin{pmatrix} c_1 \beta \\ -c_2 s_1 \beta \\ c_3 s_2 s_1 \beta \\ -s_3 s_2 s_1 \beta \\ 0 \end{pmatrix} = \begin{pmatrix} c_1 \beta \\ -c_2 s_1 \beta \\ c_3 s_2 s_1 \beta \\ -c_4 s_3 s_2 s_1 \beta \\ s_4 s_3 s_2 s_1 \beta \end{pmatrix}$$

Givens rotations after m steps

We had: $\overline{H}_m y = \beta e_1 = \overline{g}_0$.

After m rotations we obtained: $\overline{H}_m^{(m)} y = \overline{g}_m$

During rotations we did m multiplications by $\Omega_1, \Omega_2, \dots, \Omega_m$:

$$\Omega_m \Omega_{m-1} \dots \Omega_2 \Omega_1 \overline{H}_m y = \Omega_m \Omega_{m-1} \dots \Omega_2 \Omega_1 \overline{g}_0$$

Define $\Omega_m \Omega_{m-1} \dots \Omega_2 \Omega_1 = Q_m$, Q_m will be unitary square matrix: $Q_m^{-1} = Q_m^H$

$\Omega_m \Omega_{m-1} \dots \Omega_2 \Omega_1 \overline{H}_m = Q_m \overline{H}_m = \overline{H}_m^{(m)} = \overline{R}_m$, \overline{R}_m will be upper triangular with zero last row

$$\Omega_m \Omega_{m-1} \dots \Omega_2 \Omega_1 \overline{g}_0 = Q_m \overline{g}_0 = \overline{g}_m$$

The system $\overline{R}_m y = \overline{g}_m$ for the case $m=4$:

$$\overline{R}_4 y = \overline{g}_4$$

$$\begin{pmatrix} h_{11}^{(4)} & h_{12}^{(4)} & h_{13}^{(4)} & h_{14}^{(4)} \\ 0 & h_{22}^{(4)} & h_{23}^{(4)} & h_{24}^{(4)} \\ 0 & 0 & h_{33}^{(4)} & h_{34}^{(4)} \\ 0 & 0 & 0 & h_{44}^{(4)} \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} c_1 \beta \\ -c_2 s_1 \beta \\ c_3 s_2 s_1 \beta \\ -c_4 s_3 s_2 s_1 \beta \\ s_4 s_3 s_2 s_1 \beta \end{pmatrix} = \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \\ \gamma_5 \end{pmatrix}.$$

Calculating residual in GMRES

In general: $\bar{R}_m y = \bar{g}_m$ is the system with $m + 1$ equations and m unknowns

$$\begin{pmatrix} h_{11}^{(m)} & h_{12}^{(m)} & \dots & h_{1m}^{(m)} \\ 0 & h_{22}^{(m)} & \dots & h_{2m}^{(m)} \\ & & \dots & \\ 0 & 0 & \dots & h_{mm}^{(m)} \\ 0 & 0 & \dots & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_m \end{pmatrix} = \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \dots \\ \gamma_m \\ \gamma_{m+1} \end{pmatrix}$$

$$\|\bar{R}_m y - \bar{g}_m\|_2 \rightarrow \min$$

Let R_m be the matrix \bar{R}_m with the last zero row deleted and g_m be the vector \bar{g}_m

$$\text{with } \gamma_{m+1} \text{ deleted: } R_m = \begin{pmatrix} h_{11}^{(m)} & h_{12}^{(m)} & \dots & h_{1m}^{(m)} \\ 0 & h_{22}^{(m)} & \dots & h_{2m}^{(m)} \\ & & \dots & \\ 0 & 0 & & h_{mm}^{(m)} \end{pmatrix}, \quad g_m = \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \dots \\ \gamma_m \end{pmatrix}$$

Then we get the system $R_m y = g_m$ with upper triangular square matrix $R_m \in \mathbb{C}^{m \times m}$, $g_m \in \mathbb{C}^m$

This system has m equations and m unknowns. It can be easily solved by backward

substitution. The solution of this system $y_m = R_m^{-1} g_m$ minimizes $\|\bar{R}_m y - \bar{g}_m\|_2$ and

$$\|r_m\|_2 = \|\beta e_1 - \bar{H}_m y_m\|_2 = \|\bar{g}_m - \bar{R}_m y_m\|_2 = |\gamma_{m+1}|$$

Theorems about GMRES

Proposition. Consider $m \ll n$ and $\Omega_i, i = \overline{1, m}$ are the rotation matrices used to transform the upper Hessenberg matrix $\overline{H}_m \in \mathbb{C}^{(m+1) \times m}$ into upper triangular form $\overline{R}_m \in \mathbb{C}^{(m+1) \times m}$,

$$\overline{R}_m = Q_m \overline{H}_m, \text{ where } Q_m = \Omega_m \Omega_{m-1} \dots \Omega_2 \Omega_1.$$

$$g_m = (\gamma_1, \gamma_2, \dots, \gamma_m, \gamma_{m+1})^T \in \mathbb{C}^{m+1}$$

$R_m \in \mathbb{C}^{m \times m}$ is the matrix \overline{R}_m without the last zero row and

$g_m \in \mathbb{C}^m$ is the vector \overline{g}_m without the last component γ_{m+1}

Then the following holds:

1) $rank(AV_m) = rank(R_m)$ and if $r_{mm} = 0$, then A is singular

2) $y_m = R_m^{-1} g_m$ gives minimum to the functional $\|\overline{R}_m y - \overline{g}_m\|_2$

3) The residual vector r_m of the approximate solution x_m computed by GMRES algorithm

is defined as $r_m = V_{m+1} Q_m^H \gamma_{m+1} e_{m+1}$

4) Residual norm $\|r_m\|_2 = |\gamma_{m+1}|$

Theorem 1. If A is nonsingular, then GMRES breaks down

at step j ($h_{j+1,j} = 0$) \Leftrightarrow approximate solution x_j is exact (residual is zero).

Proof of proposition

△ Prove 1) $\text{rank}(AV_m) = \text{rank}(R_m)$

From Arnoldi relation $AV_m = V_{m+1} \bar{H}_m = V_{m+1} Q_m^H Q_m \bar{H}_m = V_{m+1} Q_m^H \bar{R}_m$

Since $V_{m+1} Q_m^H$ is unitary and \bar{R}_m has the last zero row,

$$\text{rank}(AV_m) = \text{rank}(V_{m+1} Q_m^H \bar{R}_m) = \text{rank}(R_m)$$

If $r_{mm} = 0$, $\text{rank}(R_m) \leq m-1$, so $\text{rank}(AV_m) \leq m-1$

Since V_m has full rank, A must be singular.

2) The solution $y_m = R_m^{-1} g_m$ of the system $R_m y = g_m$

minimizes the functional $\|\bar{R}_m y - \bar{g}_m\|_2$, because

$$\|r_m\|_2 = \|\beta e_1 - \bar{H}_m y\|_2 = \|\bar{g}_m - \bar{R}_m y\|_2 = |\gamma_{m+1}| + \|g_m - R_m y\|_2$$

$$\|g_m - R_m y\|_2 = 0 \text{ when } y = y_m = R_m^{-1} g_m$$

3) During derivation of GRMRES we got $r_m = V_{m+1} (\beta e_1 - \bar{H}_m y_m) \Rightarrow$

$$r_m = V_{m+1} (\beta e_1 - \bar{H}_m y_m) = V_{m+1} Q_m^H Q_m (\beta e_1 - \bar{H}_m y_m) = V_{m+1} Q_m^H (Q_m \beta e_1 - Q_m \bar{H}_m y_m) =$$

$$= V_{m+1} Q_m^H (\bar{g}_m - \bar{R}_m y_m) = V_{m+1} Q_m^H \gamma_{m+1} e_{m+1} \Rightarrow$$

4) $\|r_m\|_2 = \|V_{m+1} Q_m^H \gamma_{m+1} e_{m+1}\|_2 = |\gamma_{m+1}|$, as $V_{m+1} Q_m^H$ has orthonormal columns \square

Modifications of GMRES

As it was with FOM, large m leads to the growth of memory.

The full GMRES algorithm is guaranteed to converge at most n steps, but this would be impractical.

So in order not to increase m to the value of n , there are two variations of GMRES:

- 1) restart GMRES for the fixed value of m
- 2) truncate the orthogonalization in Arnoldi process (see details in Yosef Saad book, p. 180)

Restarted GMRES can stagnate when the matrix is not positive definite.

In practice GMRES is often used with *preconditioning*.

Restarted GMRES(m)

1. Take initial guess x_0 , calculate initial residual $r_0 = b - Ax_0$, calculate residual norm $\beta = \|r_0\|_2$, and the first vector $v_1 = \frac{r_0}{\beta}$
2. Define $\overline{H}_m = \{h_{ij}\} \in \mathbb{C}^{(m+1) \times m}$, $V_m \in \mathbb{C}^{n \times m}$, $V_{m+1} \in \mathbb{C}^{n \times (m+1)}$. Set $\overline{H}_m = 0$
3. Build the basis for Krylov subspace using Arnoldi process.
4. Find y_m as a minimizer of the functional $J(y) = \|\beta e_1 - \overline{H}_m y\|_2$
(solve overdefined system $\overline{H}_m y = \beta e_1$)
5. **Calculate residual r_m** , estimate residual norm: $\|r_m\|_2 < \varepsilon$.
6. Calculate new approximate solution $x_m = x_0 + V_m y_m$
If satisfied, then stop. Else
7. Set $x_0 := x_m$ and go to step 1.