Utrecht, 8 november 2017

Optimal Iterative Methods

Gerard Sleijpen

Department of Mathematics

http://www.staff.science.uu.nl/~sleij101/

$\mathbf{Ax} = \mathbf{b}$ $\qquad\qquad$ $\{\lambda \in \mathbb{C} \mid \mathbf{Av} = \lambda \mathbf{v}\} \subset \mathcal{E} \subset \mathbb{C}$

Construct iteratively $\mathbf{r}_k = \mathbf{b} - \mathbf{A}\mathbf{x}_k$ with $\|\mathbf{r}_k\|_2$ small

$\mathbf{x}_k$ gets a free ride: $\mathbf{r}_k = \mathbf{r}_{k-1} - \alpha_k \mathbf{A}\mathbf{u}_{k-1}$, $\mathbf{x}_k = \mathbf{x}_{k-1} + \alpha_k \mathbf{u}_{k-1}$

Richardson: $\alpha \in \mathbb{C}$, $\mathbf{r}_k = \mathbf{r}_{k-1} - \alpha \mathbf{A}\mathbf{r}_{k-1} = (\mathbf{I} - \alpha \mathbf{A})^k \mathbf{r}_0$

Krylov subspace: $\mathcal{K}_k(\mathbf{A}, \mathbf{r}_0) = \text{span}(\mathbf{r}_0, \mathbf{A}\mathbf{r}_0, \ldots, \mathbf{A}^{k-1}\mathbf{r}_0)$
$= \{p(\mathbf{A})\mathbf{r}_0 \mid p \text{ pol. degree } < k\}$

Richardson: $\alpha_k \in \mathbb{C}$, $\mathbf{r}_k = \mathbf{r}_{k-1} - \alpha_k \mathbf{A}\mathbf{r}_{k-1} = p_k(\mathbf{A})\mathbf{r}_0$
with $p_k(\lambda) = (1 - \alpha_1 \lambda) \cdot \ldots \cdot (1 - \alpha_k \lambda)$

Selection α_k: static, **dynamic**

- $\alpha_k = \alpha_{k \bmod \ell}$: with $p(\lambda) = (1 - \alpha_1 \lambda) \cdots (1 - \alpha_\ell \lambda)$, $p_{j\ell} = p^j$
and $\max_{\zeta \in \mathcal{E}} |p(\zeta)|$ as small as possible.

- **Chebyshev iteration**: $p_k = \widetilde{T}_k$, $\mathbf{r}_{k+1} = \widetilde{\alpha}_k \mathbf{r}_k - \widetilde{\beta}_k \mathbf{A}\mathbf{r}_k - \widetilde{\gamma}_{k-1}\mathbf{r}_{k-1}$

- **Local Minimal Residual**: $\alpha_k = \text{argmin}_\alpha \|\mathbf{r}_{k-1} - \alpha \mathbf{A}\mathbf{r}_{k-1}\|_2$

- **Generalized Conjugate Residuals**: $\mathbf{R}_k \equiv [\mathbf{r}_0, \ldots, \mathbf{r}_{k-1}]$
$\mathbf{r}_k = \mathbf{r}_{k-1} - \mathbf{A}\mathbf{R}_k \vec{\alpha}_{k-1}$ with $\vec{\alpha}_{k-1} \equiv \text{minarg}_{\vec{\alpha} \in \mathbb{C}^k} \|\mathbf{r}_{k-1} - \mathbf{A}\mathbf{R}_k \vec{\alpha}\|_2$

Program Lecture 6

- Krylov basis & Hessenberg matrices
- Arnoldi's decomposition
- Linear systems and Arnoldi's decomposition
- GMRES and FOM
- Convergence
- Eigenvalues and Arnoldi's decomposition
- Arnoldi's method
- Convergence
- Stability issues in Arnoldi's decomposition
- Summary

Program Lecture 6

- Krylov basis & Hessenberg matrices
- Arnoldi's decomposition
- Linear systems and Arnoldi's decomposition
- GMRES and FOM
- Convergence
- Eigenvalues and Arnoldi's decomposition
- Arnoldi's method
- Convergence
- Stability issues in Arnoldi's decomposition
- Summary

Krylov subspace methods

$\mathbf{Ax} = \mathbf{b}$. $\mathbf{x}_0 = 0$.

Find $\mathbf{x}_k \in \mathcal{K}_k(\mathbf{A}, \mathbf{r}_0)$ such that

$$\mathbf{r}_k = \mathbf{b} - \mathbf{A}\mathbf{x}_k = \mathbf{b} - \mathbf{A}q_{k-1}(\mathbf{A})\mathbf{b} = p_k(\mathbf{A})\mathbf{b}$$

is small in some sense.

$\mathcal{K}_k(\mathbf{A}, \mathbf{r}_0)$ is a **search subspace**.

Idea.

- Compute orthonormal basis $\mathbf{v}_1, \ldots, \mathbf{v}_k$ of $\mathcal{K}_k(\mathbf{A}, \mathbf{r}_0)$.
- Assemble the $\mathbf{v}_j$ in the matrix $\mathbf{V}_k \equiv [\mathbf{v}_1, \ldots, \mathbf{v}_k]$.
- Compute $\mathbf{x}_k$ as $\mathbf{V}_k \vec{y}_k$ with $\vec{y}_k$ a k-vector.
- Find $\vec{y}_k$ such that $\mathbf{A}\mathbf{V}_k \vec{y}_k \approx \mathbf{r}_0 = \|\mathbf{r}_0\|_2 \mathbf{V}_k e_1$

Orthonormal basis $\mathcal{K}_k(\mathbf{A}, \mathbf{r}_0)$

Construction: recursively

How to **expand** $\mathbf{V}_k$, that is,
how to expand $\mathbf{v}_1, \ldots, \mathbf{v}_k$ to a basis for $\mathcal{K}_{k+1}(\mathbf{A}, \mathbf{r}_0)$?

Exercise. Suppose $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is a **Krylov basis**, i.e.,

$$\mathcal{K}_j(\mathbf{A}, \mathbf{r}_0) = \text{span}(\mathbf{v}_1, \ldots, \mathbf{v}_j) \text{ for all } j = 1, \ldots, k$$
$$\Rightarrow \quad \mathcal{K}_{k+1}(\mathbf{A}, \mathbf{r}_0) = \text{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{A}\mathbf{v}_k).$$

$\mathbf{w} \equiv \mathbf{A}\mathbf{v}_k$ can be used for expanding the search subspace. For stability reasons, $\mathbf{w}$ is modified first before expanding the basis.

Orthonormal basis $\mathcal{K}_k(\mathbf{A}, \mathbf{r}_0)$

Recall $\mathbf{V}_k \equiv [\mathbf{v}_1, \ldots, \mathbf{v}_k]$.

Suppose $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is an orthonormal Krylov basis $\mathcal{K}_k(\mathbf{A}, \mathbf{r}_0)$.

Compute $\mathbf{v}_{k+1}$ by **orthogonalising $\mathbf{A}\mathbf{v}_k$ against $\mathbf{V}_k$**:

- **Expand**: $\mathbf{w} = \mathbf{A}\mathbf{v}_k$,
- **Orthogonalize**: $\widetilde{\mathbf{v}} = \mathbf{w} - \mathbf{V}_k \vec{h}'_k$ with $\vec{h}'_k = \mathbf{V}_k^* \mathbf{w}$,
- **Normalize**: $\mathbf{v}_{k+1} = \widetilde{\mathbf{v}}/\nu_k$ with $\nu_k = \|\widetilde{\mathbf{v}}\|_2$.

Orthonormal basis $\mathcal{K}_k(\mathbf{A}, \mathbf{r}_0)$

Recall $\mathbf{V}_k \equiv [\mathbf{v}_1, \ldots, \mathbf{v}_k]$.

Suppose $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is an orthonormal Krylov basis $\mathcal{K}_k(\mathbf{A}, \mathbf{r}_0)$.

Compute $\mathbf{v}_{k+1}$ by **orthogonalising $\mathbf{A}\mathbf{v}_k$ against $\mathbf{V}_k$**:

- **Expand**: $\mathbf{w} = \mathbf{A}\mathbf{v}_k$,
- **Orthogonalize**: $\widetilde{\mathbf{v}} = \mathbf{w} - \mathbf{V}_k \vec{h}'_k$ with $\vec{h}'_k = \mathbf{V}_k^* \mathbf{w}$,
- **Normalize**: $\mathbf{v}_{k+1} = \widetilde{\mathbf{v}}/\nu_k$ with $\nu_k = \|\widetilde{\mathbf{v}}\|_2$.

Note. With $\vec{h}_k \equiv (\vec{h}'^{\mathsf{T}}_k, \nu_k)^{\mathsf{T}}$, we have

$$\mathbf{A}\mathbf{v}_k = \mathbf{w} = \mathbf{V}_k \vec{h}'_k + \mathbf{v}_{k+1}\nu_k = [\mathbf{V}_k, \mathbf{v}_{k+1}] \begin{bmatrix} \vec{h}'_k \\ \nu_k \end{bmatrix} = \mathbf{V}_{k+1} \vec{h}_k$$

Orthonormal basis $\mathcal{K}_k(\mathbf{A}, \mathbf{r}_0)$

Recall $\mathbf{V}_k \equiv [\mathbf{v}_1, \ldots, \mathbf{v}_k]$.

Suppose $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is an orthonormal Krylov basis $\mathcal{K}_k(\mathbf{A}, \mathbf{r}_0)$.

Compute $\mathbf{v}_{k+1}$ by **orthogonalising $\mathbf{A}\mathbf{v}_k$ against $\mathbf{V}_k$**:

- **Expand**: $\mathbf{w} = \mathbf{A}\mathbf{v}_k$,
- **Orthogonalize**: $\widetilde{\mathbf{v}} = \mathbf{w} - \mathbf{V}_k \vec{h}'_k$ with $\vec{h}'_k = \mathbf{V}_k^* \mathbf{w}$,
- **Normalize**: $\mathbf{v}_{k+1} = \widetilde{\mathbf{v}}/\nu_k$ with $\nu_k = \|\widetilde{\mathbf{v}}\|_2$.

Note. With $\vec{h}_k \equiv (\vec{h}'^{\mathsf{T}}_k, \nu_k)^{\mathsf{T}}$, we have

$$\mathbf{A}\mathbf{v}_k = \mathbf{w} = \mathbf{V}_k \vec{h}'_k + \mathbf{v}_{k+1}\nu_k = [\mathbf{V}_k, \mathbf{v}_{k+1}] \begin{bmatrix} \vec{h}'_k \\ \nu_k \end{bmatrix} = \mathbf{V}_{k+1} \vec{h}_k$$

Assemble $\mathbf{A}[\mathbf{V}_{k-1}, \mathbf{v}_k] = [\mathbf{V}_k, \mathbf{v}_{k+1}] \begin{bmatrix} \underline{H}_{k-1} & \vec{h}'_k \\ 0 \ldots 0 & \nu_k \end{bmatrix}$

Orthonormal basis $\mathcal{K}_k(\mathbf{A}, \mathbf{r}_0)$

Recall $\mathbf{V}_k \equiv [\mathbf{v}_1, \ldots, \mathbf{v}_k]$.

Suppose $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is an orthonormal Krylov basis $\mathcal{K}_k(\mathbf{A}, \mathbf{r}_0)$.

Compute $\mathbf{v}_{k+1}$ by **orthogonalising $\mathbf{A}\mathbf{v}_k$ against $\mathbf{V}_k$**:

- **Expand**: $\mathbf{w} = \mathbf{A}\mathbf{v}_k$,
- **Orthogonalize**: $\widetilde{\mathbf{v}} = \mathbf{w} - \mathbf{V}_k \vec{h}'_k$ with $\vec{h}'_k = \mathbf{V}_k^* \mathbf{w}$,
- **Normalize**: $\mathbf{v}_{k+1} = \widetilde{\mathbf{v}}/\nu_k$ with $\nu_k = \|\widetilde{\mathbf{v}}\|_2$.

Note. With $\vec{h}_k \equiv (\vec{h}'^{\mathsf{T}}_k, \nu_k)^{\mathsf{T}}$, we have

$$\mathbf{A}\mathbf{v}_k = \mathbf{w} = \mathbf{V}_k \vec{h}'_k + \mathbf{v}_{k+1}\nu_k = [\mathbf{V}_k, \mathbf{v}_{k+1}] \begin{bmatrix} \vec{h}'_k \\ \nu_k \end{bmatrix} = \mathbf{V}_{k+1}\vec{h}_k$$

Assemble $\mathbf{A}[\mathbf{V}_{k-1}, \mathbf{v}_k] = [\mathbf{V}_k, \mathbf{v}_{k+1}] \begin{bmatrix} \underline{H}_{k-1} & \vec{h}'_k \\ 0 \ldots 0 & \nu_k \end{bmatrix}$

Orthonormal basis $\mathcal{K}_k(\mathbf{A}, \mathbf{r}_0)$

Recall $\mathbf{V}_k \equiv [\mathbf{v}_1, \ldots, \mathbf{v}_k]$.

Suppose $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is an orthonormal Krylov basis $\mathcal{K}_k(\mathbf{A}, \mathbf{r}_0)$.

Compute $\mathbf{v}_{k+1}$ by **orthogonalising $\mathbf{A}\mathbf{v}_k$ against $\mathbf{V}_k$**:

- **Expand**: $\mathbf{w} = \mathbf{A}\mathbf{v}_k$,
- **Orthogonalize**: $\widetilde{\mathbf{v}} = \mathbf{w} - \mathbf{V}_k \vec{h}'_k$ with $\vec{h}'_k = \mathbf{V}_k^* \mathbf{w}$,
- **Normalize**: $\mathbf{v}_{k+1} = \widetilde{\mathbf{v}}/\nu_k$ with $\nu_k = \|\widetilde{\mathbf{v}}\|_2$.

Note. With $\vec{h}_k \equiv (\vec{h}'^{\mathsf{T}}_k, \nu_k)^{\mathsf{T}}$, we have

$$\mathbf{A}\mathbf{v}_k = \mathbf{w} = \mathbf{V}_k \vec{h}'_k + \mathbf{v}_{k+1}\nu_k = [\mathbf{V}_k, \mathbf{v}_{k+1}] \begin{bmatrix} \vec{h}'_k \\ \nu_k \end{bmatrix} = \mathbf{V}_{k+1} \vec{h}_k$$

Assemble $$\mathbf{A}[\mathbf{V}_{k-1}, \mathbf{v}_k] = [\mathbf{V}_k, \mathbf{v}_{k+1}] \begin{bmatrix} \underline{H}_{k-1} & \vec{h}'_k \\ 0 \ldots 0 & \nu_k \end{bmatrix}$$

Theorem.

Orthogonalising $\mathbf{A}\mathbf{v}_j$ against $\mathbf{V}_j$ for $j = 1, \ldots, k$ leads to

$$\mathbf{A}\mathbf{V}_k = \mathbf{V}_{k+1}\,\underline{H}_k,$$

with $\mathbf{V}_k$ orthonormal, spanning $\mathcal{K}_k(\mathbf{A}, \mathbf{v}_1)$,

$\underline{H}_k$ is $(k+1) \times k$ upper Hessenberg.

Note. The matrix $\underline{H}_k$ comes for **free** in the orthogonalisation process.

Application for solving $\mathbf{A}\mathbf{x} = \mathbf{r}_0$. Try $\mathbf{x}_k = \mathbf{V}_k\,\vec{y}_k$.

Find $\vec{y}_k$ such that $\mathbf{A}\,\mathbf{V}_k\,\vec{y}_k \approx \mathbf{r}_0 = \rho_0\,\mathbf{V}_k\,e_1$ with $\rho_0 \equiv \|\mathbf{r}_0\|_2$

$\rightsquigarrow$ Find $\vec{y}_k$ such that $\mathbf{V}_{k+1}\,\underline{H}_k\,\vec{y}_k \approx \rho_0\,\mathbf{V}_{k+1}\,e_1$ if $\rho_0\,\mathbf{v}_1 = \mathbf{r}_0$

$\rightsquigarrow$ Find $\vec{y}_k$ such that $\underline{H}_k\,\vec{y}_k \approx \rho_0\,e_1$ if $\rho_0\,\mathbf{v}_1 = \mathbf{r}_0$

Details later.

Hessenberg and Krylov

Hessenberg matrices and Krylov subspaces are intimately related.

Theorem. Consider the relation $\mathbf{A}\mathbf{V}_k = \mathbf{V}_{k+1}\,\underline{H}_k$,

where $\mathbf{V}_{k+1} = [\mathbf{V}_k, \mathbf{v}_{k+1}]$ is $n \times (k+1)$,

and $\underline{H}_k$ is $(k+1) \times k$.

Then, $\mathbf{v}_1, \ldots, \mathbf{v}_k$ form a **Krylov basis** for $\mathcal{K}_k(\mathbf{A}, \mathbf{v}_1)$

i.e., $\mathbf{V}_j$ spans $\mathcal{K}_j(\mathbf{A}, \mathbf{r}_0)$ for all $j = 1, \ldots, k$,

$\Leftrightarrow$ $\underline{H}_k$ is Hessenberg.

In **Arnoldi's decomposition**, $\mathbf{V}_k$ is selected to be orthonormal (to ease computations and to enhance stability).

Arnoldi's method:

orthonormalise $\mathbf{A}\mathbf{v}_k$ against $\mathbf{V}_k$ to obtain $\mathbf{v}_{k+1}$ all k.

Orthogonalisation

Terminology.
If $\mathbf{V}$ is an $n \times k$ orthonormal matrix and $\mathbf{w}$ is an n vector, then, with **orthonormalise w against V**, we mean: construct an n-vector $\mathbf{v}$ and a $(k+1)$-vector $\vec{h}$ such that

$$\mathbf{v} \perp \mathbf{V}, \quad \|\mathbf{v}\|_2 = 1, \quad \mathbf{w} = [\mathbf{V}, \mathbf{v}]\, \vec{h}$$

Notation. $[\mathbf{v}, \vec{h}\,] = \texttt{Orth}(\mathbf{V}, \mathbf{w})$

Use a stable variant of Gram–Schmidt.

Note that the last coordinate of $\vec{h}$ is 0 if $\mathbf{w}$ is in the span of $\mathbf{V}$: in such a case (and if $k < n$), we select $\mathbf{v}$ to be a (random) normalized vector orthogonal to $\mathbf{V}$ (we insist on expanding to avoid stagnation in subsequential steps).

Program Lecture 6

- Krylov basis & Hessenberg matrices
- Arnoldi's decomposition
- Linear systems and Arnoldi's decomposition
- GMRES and FOM
- Convergence
- Eigenvalues and Arnoldi's decomposition
- Arnoldi's method
- Convergence
- Stability issues in Arnoldi's decomposition
- Summary

Arnoldi's decomposition

$$\mathbf{A}\mathbf{V}_{k-1} = \mathbf{V}_k \underline{H}_{k-1},$$

with $\mathbf{V}_k$ $n \times k$ orthonormal, $\underline{H}_{k-1}$ $k \times (k-1)$ Hessenberg.

Expand the decomposition to $\mathbf{A}\mathbf{V}_k = \mathbf{V}_{k+1} \underline{H}_k$.

Notation. $[\mathbf{V}_{k+1}, \underline{H}_k] = \texttt{ArnStep}(\mathbf{A}, \mathbf{V}_k, \underline{H}_{k-1})$

$$\mathbf{w} = \mathbf{A}\mathbf{v}_k$$

$$[\mathbf{v}_{k+1}, \vec{h}_k] = \texttt{Orth}(\mathbf{V}_k, \mathbf{w})$$

$$\mathbf{V}_{k+1} = [\mathbf{V}_k, \mathbf{v}_{k+1}]$$

$$\underline{H}_k = \begin{bmatrix} \underline{H}_{k-1} \\ \vec{0}^*_{k-1} \end{bmatrix}, \quad \underline{H}_k \leftarrow [\underline{H}_k, \vec{h}_k]$$

Program Lecture 6

- Krylov basis & Hessenberg matrices
- Arnoldi's decomposition
- Linear systems and Arnoldi's decomposition
- GMRES and FOM
- Convergence
- Eigenvalues and Arnoldi's decomposition
- Arnoldi's method
- Convergence
- Stability issues in Arnoldi's decomposition
- Summary

Linear systems and Arnoldi's decomposition

Arnoldi's decomposition: $\mathbf{A}\mathbf{V}_k = \mathbf{V}_{k+1}\,\underline{H}_{k+1}$

The columns of $\mathbf{V}_k$ form an orthonormal basis for $\mathcal{K}_k(\mathbf{A}, \mathbf{v}_1)$.

Ax = b. Put $\rho_0 \equiv \|\mathbf{b}\|_2$.

Take $\mathbf{x}_0 = \mathbf{0}$ and $\mathbf{v}_1 = \mathbf{b}$, and form Arnoldi's decomp.

Hence, $\mathbf{A}\,\mathbf{V}_k\,\vec{y}_k = \mathbf{V}_{k+1}\,\underline{H}_k\,\vec{y}_k$ and

$$\mathbf{r}_k = \mathbf{b} - \mathbf{A}\mathbf{x}_k = \mathbf{V}_{k+1}(\rho_0\,e_1 - \underline{H}_k\,\vec{y}_k) \text{ and}$$
$$\|\mathbf{r}_k\|_2 = \|\rho_0\,e_1 - \underline{H}_k\,\vec{y}_k\|_2$$

Observation. The norm of the residual $\mathbf{r}_k$ can be computed in k-dimensional space without computing $\mathbf{r}_k$.

How to solve $\underline{H}_k\,\vec{y}_k = \rho_0\,e_1$?

Program Lecture 6

- Krylov basis & Hessenberg matrices
- Arnoldi's decomposition
- Linear systems and Arnoldi's decomposition
- GMRES and FOM
- Convergence
- Eigenvalues and Arnoldi's decomposition
- Arnoldi's method
- Convergence
- Stability issues in Arnoldi's decomposition
- Summary

Linear systems and Arnoldi's decomposition

Arnoldi's decomposition: $\mathbf{A}\mathbf{V}_k = \mathbf{V}_{k+1}\,\underline{H}_{k+1}$

The columns of $\mathbf{V}_k$ form an orthonormal basis for $\mathcal{K}_k(\mathbf{A}, \mathbf{v}_1)$.

Ax = b. Put $\rho_0 \equiv \|\mathbf{b}\|_2$.

Take $\mathbf{x}_0 = \mathbf{0}$ and $\mathbf{v}_1 = \mathbf{b}$, and form Arnoldi's decomp.

Hence, $\mathbf{A}\,\mathbf{V}_k\,\vec{y}_k = \mathbf{V}_{k+1}\,\underline{H}_k\,\vec{y}_k$ and

$$\mathbf{r}_k = \mathbf{b} - \mathbf{A}\mathbf{x}_k = \mathbf{V}_{k+1}(\rho_0\,e_1 - \underline{H}_k\,\vec{y}_k) \text{ and}$$
$$\|\mathbf{r}_k\|_2 = \|\rho_0\,e_1 - \underline{H}_k\,\vec{y}_k\|_2$$

GMRES: solve $\underline{H}_k\,\vec{y}_k = \rho_0\,e_1$ in the least square sense.

Linear systems and Arnoldi's decomposition

Arnoldi's decomposition: $\mathbf{A}\mathbf{V}_k = \mathbf{V}_{k+1}\,\underline{H}_{k+1}$

The columns of $\mathbf{V}_k$ form an orthonormal basis for $\mathcal{K}_k(\mathbf{A}, \mathbf{v}_1)$.

Ax = b. Put $\rho_0 \equiv \|\mathbf{b}\|_2$.

Take $\mathbf{x}_0 = \mathbf{0}$ and $\mathbf{v}_1 = \mathbf{b}$, and form Arnoldi's decomp.

Hence, $\mathbf{A}\,\mathbf{V}_k\,\vec{y}_k = \mathbf{V}_{k+1}\,\underline{H}_k\,\vec{y}_k$ and

$$\mathbf{r}_k = \mathbf{b} - \mathbf{A}\mathbf{x}_k = \mathbf{V}_{k+1}(\rho_0\,e_1 - \underline{H}_k\,\vec{y}_k) \text{ and}$$
$$\|\mathbf{r}_k\|_2 = \|\rho_0\,e_1 - \underline{H}_k\,\vec{y}_k\|_2$$

FOM: solve $H_k\,\vec{y}_k = \rho_0\,e_1$;
H_k is the $k \times k$ upper block of $\underline{H}_k$.

[Saad and Schultz '86]

Generalized Minimal Residuals

Proposition. With $\mathbf{x}_k \equiv \mathbf{V}_k \vec{y}_k$ and $\mathbf{r}_k \equiv \mathbf{b} - \mathbf{A}\mathbf{x}_k$,

$\vec{y}_k$ solves $\underline{H}_k \vec{y}_k = \rho_0 e_1$ in least square sense

$\Leftrightarrow$ $\|\mathbf{r}_k\|$ minimizes $\|\mathbf{b} - \mathbf{A}\tilde{\mathbf{x}}\|_2$ over all $\tilde{\mathbf{x}} \in \mathcal{K}_k(\mathbf{A}, \mathbf{b})$.

Select $k_{\max}$ and tol
Set $\rho_0 = \|\mathbf{b}\|_2$, $\mathbf{V}_1 = [\mathbf{b}/\rho_0]$, $\underline{H}_0 = [\,]$
for $k = 1, \ldots, k_{\max}$ do
  Break if $\rho_k <$ tol
  $[\mathbf{V}_{k+1}, \underline{H}_k] = \mathtt{ArnStep}(\mathbf{A}, \mathbf{V}_k, \underline{H}_{k-1})$
  Solve $\underline{H}_k \vec{y}_k = \rho_0 e_1$ in least square sense
  $\rho_k = \|\rho_0 e_1 - \underline{H}_k \vec{y}_k\|_2$
end for
$\mathbf{x} = \mathbf{V}_k \vec{y}_k$.

[Saad and Schultz '86]

Full Orthogonalisation Method

Proposition. With $\mathbf{x}_k \equiv \mathbf{V}_k \vec{y}_k$ and $\mathbf{r}_k \equiv \mathbf{b} - \mathbf{A}\mathbf{x}_k$,

$\vec{y}_k$ solves $H_k \vec{y}_k = \rho_0 e_1 \quad \Leftrightarrow \quad \mathbf{r}_k \perp \mathcal{K}_k(\mathbf{A}, \mathbf{b})$.

```
Select k_max and tol
Set ρ_0 = ||b||_2, V_1 = [b/ρ_0], H_0 = [ ]
for k = 1, ..., k_max do
    Break if ρ_k < tol
    [V_{k+1}, H_k] = ArnStep(A, V_k, H_{k-1})
    Solve H_k y_k = ρ_0 e_1 for y_k
    ρ_k = ||ρ_0 e_1 − H_k y_k||_2
end for
x = V_k y_k .
```

[Saad and Schultz '86]

Full Orthogonalisation Method

Proposition. With $\mathbf{x}_k \equiv \mathbf{V}_k \vec{y}_k$ and $\mathbf{r}_k \equiv \mathbf{b} - \mathbf{A}\mathbf{x}_k$,

$\vec{y}_k$ solves $H_k \vec{y}_k = \rho_0 e_1 \quad \Leftrightarrow \quad \mathbf{r}_k \perp \mathcal{K}_k(\mathbf{A}, \mathbf{b}) \quad \Leftrightarrow \quad \mathbf{b} - \mathbf{A}\mathbf{V}_k \vec{y}_k \perp \mathbf{V}_k.$

```
Select k_max and tol
Set ρ_0 = ||b||_2, V_1 = [b/ρ_0], H_0 = []
for k = 1, ..., k_max do
    Break if ρ_k < tol
    [V_{k+1}, H_k] = ArnStep(A, V_k, H_{k-1})
    Solve H_k y_k = ρ_0 e_1 for y_k
    ρ_k = ||ρ_0 e_1 − H_k y_k||_2
end for
x = V_k y_k .
```

[Saad and Schultz '86]

Generalized Minimal Residuals

Proposition. With $\mathbf{x}_k \equiv \mathbf{V}_k \vec{y}_k$ and $\mathbf{r}_k \equiv \mathbf{b} - \mathbf{A}\mathbf{x}_k$,

$\vec{y}_k$ solves $\underline{H}_k \vec{y}_k = \rho_0 e_1$ in least square sense

$\Leftrightarrow$ $\|\mathbf{r}_k\|$ minimizes $\|\mathbf{b} - \mathbf{A}\tilde{\mathbf{x}}\|_2$ over all $\tilde{\mathbf{x}} \in \mathcal{K}_k(\mathbf{A}, \mathbf{b})$.

$\Leftrightarrow$ $\mathbf{b} - \mathbf{A}\mathbf{V}_k \vec{y}_k \perp \mathbf{A}\mathbf{V}_k$.

GMRES & FOM

Notes.

- "Solve $\underline{H}_k \vec{y}_k = \rho_0 e_1$ in the least square sense"
  and "solve $H_k \vec{y}_k = \rho_0 e_1$"

are problems in k-space, where $k \ll n$: **costs are 'negligible'**

- There is no need to compute the residuals $\mathbf{r}_k$:

computing residual-norm does not require computing $\mathbf{r}_k$

$$\rho_k = \|\mathbf{r}_k\|_2 = \|\rho_0 e_1 - \underline{H}_k \vec{y}_k\|_2$$

- Computation of the approximate solution

only if residual accuracy has been reached.

GMRES & FOM

Let $\vec{\gamma}_{k+1} = (1, \gamma_2, \ldots, \gamma_k, \gamma_{k+1})^\mathsf{T} = (\vec{\gamma}_k^\mathsf{T}, \gamma_{k+1})^\mathsf{T}$ such that

$$\vec{\gamma}_{k+1}^* \underline{H}_k = \vec{0}_k^*$$

Note that $\vec{\gamma}_{k+1}$ can be computed by recursive updating.

Proposition.

$$\|\mathbf{r}_k^{\mathrm{GMRES}}\|_2 = \frac{\rho_0}{\|\vec{\gamma}_{k+1}\|_2} \qquad \|\mathbf{r}_k^{\mathrm{FOM}}\|_2 = \frac{\rho_0}{|\gamma_{k+1}|}$$

Note. [Sleijpen vd Eshof '04]

- Even the k-dimensional system has to be solved

only (once) at reaching residual accuracy.

GMRES

Select $k_{\max}$ and tol

Set $\rho_0 = \|\mathbf{b}\|_2$, $\mathbf{V}_1 = [\mathbf{b}/\rho_0]$, $\underline{H}_0 = [\,]$, $\vec{\gamma}_1 = (1)$

for $k = 1, \ldots, k_{\max}$ do

  Break if $\rho_k <$ tol

  $[\mathbf{V}_{k+1}, \underline{H}_k] = \mathtt{ArnStep}(\mathbf{A}, \mathbf{V}_k, \underline{H}_{k-1})$

  Update $\vec{\gamma}_{k+1}$ st $\vec{\gamma}_{k+1}^{*}(\underline{H}_k\, e_k) = 0$

  $\rho_k = \rho_0 / \sqrt{\rho_{k-1}^{-2} + |\gamma_{k+1}|^2}$

end for

Solve $\underline{H}_k\, \vec{y}_k = \rho_0\, e_1$ for $\vec{y}_k$

$\mathbf{x} = \mathbf{V}_k\, \vec{y}_k$.

GMRES versus GCR

Both methods are mathematically equivalent, that is, in exact arithmetic, they have the same residuals (residual norms) at step k and the same approximate solutions.

GMRES is the most efficient method that gives the approximate solution from $\mathcal{K}_k(\mathbf{A}, \mathbf{b})$ with **smallest residual $\|\cdot\|_2$-norm**. GCR needs $\approx$ twice as many AXPYs as GMRES. Assuming the same form of Gram-Schmidt is used.

GCR trivially extends to a flexible variant (injecting an $\mathbf{u}_k \neq \mathbf{r}_k$ does not hamper convergence). A flexible variant of **GMRES lacks the efficiency advantage**.

Both methods suffer from **growing computational costs per step and growing memory requirements** with increasing step numbers.

GMRES versus GCR

GMRES relies on
an orthonormal Krylov basis $\mathbf{v}_1, \ldots, \mathbf{v}_k$ for $\mathcal{K}_k(\mathbf{A}, \mathbf{r}_0)$.

GCR relies on
an $\mathbf{A}^*\mathbf{A}$-orthogonal Krylov basis $\mathbf{u}_1, \ldots, \mathbf{u}_k$ for $\mathcal{K}_k(\mathbf{A}, \mathbf{r}_0)$.

$$\mathbf{u}_j \perp_{A^*A} \mathbf{u}_i \qquad \Leftrightarrow \qquad 0 = \mathbf{u}_i^* \mathbf{A}^* \mathbf{A} \mathbf{u}_j = \mathbf{c}_i^* \mathbf{c}_j.$$

GMRES uses the same basis $\mathbf{v}_1, \ldots, \mathbf{v}_k$ for expansion
as for extraction.

GCR expands with $\mathbf{r}_0, \ldots, \mathbf{r}_k$, and

extracts with $\mathbf{u}_0, \ldots, \mathbf{u}_k$
(after $\mathbf{A}^*\mathbf{A}$-orthogonalisation).

GMRES is a Krylov subspace method,
Flexible GCR is a subspace method.

GMRES versus GCR

In contrast to GCR,
GMRES exploits the Krylov–Hessenberg structure.

Advantage. More efficiency.
If for mathematical reasons an inner product is zero, then there is no need to compute it: thus, saving flops.

Disadvantage. Sensitive to perturbations that affect the Krylov structure.
Due to perturbations, the actual value of such an inner product may not be zero

Example. If at step k in GCR "$\mathbf{u}_k = \mathbf{r}_k$" is replaced by "Select a random $\mathbf{u}_k$", then convergence is delayed by 1 step. If at step k in GMRES $\mathbf{A}\mathbf{v}_k$ is replaced by a random vector then GMRES stagnates forever.

Pertubations may come from inexact MVs, variable preconditioners, etc.

Program Lecture 6

- Krylov basis & Hessenberg matrices
- Arnoldi's decomposition
- Linear systems and Arnoldi's decomposition
- GMRES and FOM
- Convergence
- Eigenvalues and Arnoldi's decomposition
- Arnoldi's method
- Convergence
- Stability issues in Arnoldi's decomposition
- Summary

Convergence

With the same initial residual $\mathbf{r}_0$:

$$\|\mathbf{r}_k^{\text{GMRES}}\|_2 = \|\mathbf{r}_k^{\text{GCR}}\|_2 \leq \|\mathbf{r}_k^{\text{pol}}\|_2$$

Among all Krlylov subspace methods, GMRES (and) GCR find the approximate solution in the Krylov subspaces, with smallest residual 2-norm, i.e., using k MVs, GMRES (and GCR) find the solution with smallest residual 2-norm.

Krylov subspace methods are attractive if $\mathbf{A}^{-1}$ can be well approximated by matrix polynomials $q(\mathbf{A})$ with q a polynomial of low degree.

General convergence statements for general situations are hard to interpret.

'Basic convergence conditions' are:
1) the eigenvalues of $\mathbf{A}$ cluster away from 0, and
2) the eigenvector basis is not very ill conditioned.

Convergence

With the same $\mathbf{r}_0$: $\|\mathbf{r}_k^{\text{GMRES}}\|_2 = \|\mathbf{r}_k^{\text{GCR}}\|_2 \leq \|\mathbf{r}_k^{\text{pol}}\|_2$

'Basic convergence conditions' are:
1) the eigenvalues of $\mathbf{A}$ cluster away from 0, and
2) the eigenvector basis is not very ill conditioned.

Proposition. If eigs $\mathbf{A}$ in $[\lambda_-, \lambda_+] \subset (0, \infty)$, then

$$\|\mathbf{r}_k^{\text{GMRES}}\|_2 \leq \|\mathbf{r}_k^{\text{Cheb}}\|_2 \leq (2\,\mathcal{C}_E\,\|\mathbf{r}_0\|_2) \exp\left(-\frac{2k}{\sqrt{\mathcal{C}}}\right),$$

where $\mathcal{C} \equiv \frac{\lambda_+}{\lambda_-}$ and $\mathcal{C}_E$ is the condition number of the eigenvectors (i.e, $\mathcal{C}_E \equiv \|\mathbf{V}\|_2\,\|\mathbf{V}^{-1}\|_2$ if $\mathbf{AV} = \mathbf{V}\Lambda$).

Note. It is easy to construct *sophisticated* examples with $\mathcal{C}_E$ extremely large (as 10^{100}). Actually, even for small $\frac{\lambda_+}{\lambda_-}$, convergence can be arbitrarily slow.

Convergence

With the same $\mathbf{r}_0$: $\|\mathbf{r}_k^{\text{GMRES}}\|_2 = \|\mathbf{r}_k^{\text{GCR}}\|_2 \leq \|\mathbf{r}_k^{\text{pol}}\|_2$

'Basic convergence conditions' are:

1) the eigenvalues of $\mathbf{A}$ cluster away from 0, and

2) the eigenvector basis is not very ill conditioned.

In the following two examples

$\mathbf{Ax} = \mathbf{b} \equiv \mathbf{e}_1$, $\mathcal{K}_k(\mathbf{A}, \mathbf{b}) = \text{span}(\mathbf{e}_1, \ldots, \mathbf{e}_k)$, and,

if $k < n$, then $\|\mathbf{x} - \widetilde{\mathbf{x}}_k\|_2 \geq 1$ for any $\widetilde{\mathbf{x}}_k \in \mathcal{K}_k(\mathbf{A}, \mathbf{b})$.

Example 1. Let $\mathbf{A}$ be the 'circular matrix':
$\mathbf{Ae}_k \equiv \mathbf{e}_{k+1}$ for $k < n$ and $\mathbf{Ae}_n \equiv \mathbf{e}_1$. In particular, $\mathbf{x} = \mathbf{e}_n$.

Note that, in this example: $\mathcal{C}_E = 1$ ($\mathbf{A}$ is unitary),
but **the eigenvalues cluster around** 0.

Convergence

With the same $\mathbf{r}_0$: $\|\mathbf{r}_k^{\text{GMRES}}\|_2 = \|\mathbf{r}_k^{\text{GCR}}\|_2 \leq \|\mathbf{r}_k^{\text{pol}}\|_2$

'Basic convergence conditions' are:
1) the eigenvalues of $\mathbf{A}$ cluster away from 0, and
2) the eigenvector basis is not very ill conditioned.

In the following two examples
$\mathbf{Ax} = \mathbf{b} \equiv \mathbf{e}_1$, $\mathcal{K}_k(\mathbf{A}, \mathbf{b}) = \text{span}(\mathbf{e}_1, \ldots, \mathbf{e}_k)$, and,
if $k < n$, then $\|\mathbf{x} - \tilde{\mathbf{x}}_k\|_2 \geq 1$ for any $\tilde{\mathbf{x}}_k \in \mathcal{K}_k(\mathbf{A}, \mathbf{b})$.

Example 2. With $\mathbf{Se}_k \equiv \mathbf{e}_{k+1}$ for $k < n$ and $\mathbf{Se}_n \equiv \mathbf{0}$, let $\mathbf{A} = \mathbf{I} - \mathbf{S}$. Then $\mathbf{x} = \mathbf{1}$.

Note that the eigenvalues cluster away from 0: 1 is the only eigenvalue (and has algebraic multiplicity n, geometric mult. 1), while $\mathcal{C}_E = \infty$ ($\mathbf{A}$ has one huge Jordan block).

Convergence

With the same $\mathbf{r}_0$: $\|\mathbf{r}_k^{\text{GMRES}}\|_2 = \|\mathbf{r}_k^{\text{GCR}}\|_2 \leq \|\mathbf{r}_k^{\text{pol}}\|_2$

'Basic convergence conditions' are:
1) the eigenvalues of $\mathbf{A}$ cluster away from 0, and
2) the eigenvector basis is not very ill conditioned.

A modification of the above examples shows that any monotonic convergence curve is possible with unitary matrices and also with any eigenvalue distribution:

Proposition. For $\rho_0 \geq \rho_1 \geq \ldots \geq \rho_{n-1} \geq \rho_n = 0$,
with $\mathbf{x}_0 = \mathbf{0}$ and $\mathbf{A}$ is $n \times n$, consider the statement

($*$) for a $\mathbf{b}$ and $\mathbf{Ax} = \mathbf{b}$, we have $\|\mathbf{r}_k^{\text{GMRES}}\|_2 = \rho_k$ all k.

Then a) ($*$) holds for some unitary matrix $\mathbf{A}$, and

b) given $\lambda_1, \ldots, \lambda_n$ in $\mathbb{C}$, ($*$) holds for some matrix $\mathbf{A}$ with eigenvalues $\lambda_1, \ldots, \lambda_n$.

Convergence

With the same $\mathbf{r}_0$: $\|\mathbf{r}_k^{\text{GMRES}}\|_2 = \|\mathbf{r}_k^{\text{GCR}}\|_2 \leq \|\mathbf{r}_k^{\text{pol}}\|_2$

'Basic convergence conditions' are:
1) the eigenvalues of **A** cluster away from 0, and
2) the eigenvector basis is not very ill conditioned.

These conditions are not necessary for good convergence.

Example 3. If **A** is an $n \times n$ block diagonal matrix with diagonal blocks of size $k \times k$ all equal to some D, then GMRES finds the exact solution in k steps: GMRES converges quickly if $k \ll n$.

The eigensystem of D determines the eigensystem of **A**. If D is block diagonal, with one diagonal block equal to the matrix from example 1 (with $n = k/2$) and the other equal to the matrix from example 2, then $\mathcal{C}_E = \infty$ and the eigenvalues cluster around 0.

Program Lecture 6

- Krylov basis & Hessenberg matrices
- Arnoldi's decomposition
- Linear systems and Arnoldi's decomposition
- GMRES and FOM
- Convergence
- Eigenvalues and Arnoldi's decomposition
- Arnoldi's method
- Convergence
- Stability issues in Arnoldi's decomposition
- Summary

Eigenvalues and Arnoldi's decomposition

$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$.

Find a normalized $\mathbf{x}_k \in \mathcal{K}_k(\mathbf{A}, \mathbf{r}_0)$ such that

$$\mathbf{r}_k = \lambda^{(k)}\mathbf{x}_k - \mathbf{A}\mathbf{x}_k \quad \text{with} \quad \lambda^{(k)} \equiv \mathbf{x}_k^*\mathbf{A}\mathbf{x}_k$$

is small in some sense and $\lambda^{(k)}$ 'almost' has the desired properties

Arnoldi's decomposition: $\mathbf{A}\mathbf{V}_k = \mathbf{V}_{k+1}\underline{H}_k$.

$$\mathbf{r}_k = \mathbf{V}_{k+1}(\lambda^{(k)}\vec{y}_k - \underline{H}_k\,\vec{y}_k), \qquad \lambda^{(k)} = \vec{y}_k^*\,H_k\,\vec{y}_k.$$

Note

$$\|\mathbf{r}_k\|_2 = \|\lambda^{(k)}\vec{y}_k - \underline{H}_k\,\vec{y}_k\|_2$$

The computation of $\|\mathbf{r}_k\|_2$ and $\lambda^{(k)}$ is in k-space!

Program Lecture 6

- Krylov basis & Hessenberg matrices
- Arnoldi's decomposition
- Linear systems and Arnoldi's decomposition
- GMRES and FOM
- Convergence
- Eigenvalues and Arnoldi's decomposition
- Arnoldi's method
- Convergence
- Stability issues in Arnoldi's decomposition
- Summary

Approximate eigenpairs

- **Ritz values.**

$$H_k\,\vec{y}_k = \vartheta_k\,\vec{y}_k \qquad \Leftrightarrow \qquad \mathbf{r}_k = (\vartheta_k\mathbf{I} - \mathbf{A})\mathbf{V}_k\,\vec{y}_k \perp \mathbf{V}_k$$

k solution pairs $(\vartheta_k^{(1)}, \vec{y}_k^{(1)}), (\vartheta_k^{(2)}, \vec{y}_k^{(2)}), \ldots, (\vartheta_k^{(k)}, \vec{y}_k^{(k)})$.

ϑ_k is a **Ritz value**,
$\vec{y}_k$ is pre-Ritz vector,
$\mathbf{x}_k = \mathbf{V}_k\,\vec{y}_k$ is a **Ritz vector**.

$$H_k\,\vec{y}_k = \vartheta_k\,\vec{y}_k, \qquad \mathbf{x}_k \equiv \mathbf{V}_k\,\vec{y}_k.$$

$$\|\mathbf{r}_k\|_2 = \|\vartheta_k\,\mathbf{x}_k - \mathbf{A}\mathbf{x}_k\|_2 = |h_{k+1,k}\,e_k^*\vec{y}_k|$$

The residual norm comes at no additional costs!
No need to compute $\mathbf{x}_k$ first!

[Arnoldi '52]

Arnoldi's method

Proposition. With $\mathbf{x}_k \equiv \mathbf{V}_k \vec{y}_k$ and $\mathbf{r}_k \equiv \vartheta_k \mathbf{x}_k - \mathbf{A}\mathbf{x}_k$,

$\vec{y}_k$ solves $H_k \vec{y}_k = \vartheta_k \vec{y}_k \quad \Leftrightarrow \quad \mathbf{r}_k \perp \mathcal{K}_k(\mathbf{A}, \mathbf{b})$

Select $k_{\max}$ and tol
Set $\rho_0 = 1$, $\mathbf{V}_1 = [\mathbf{b}/\|\mathbf{b}\|_2]$, $\underline{H}_0 = [\,]$
for $k = 1, \ldots, k_{\max}$ do
  Break if $\rho_k <$ tol
  $[\mathbf{V}_{k+1}, \underline{H}_k] =$ ArnStep$(\mathbf{A}, \mathbf{V}_k, \underline{H}_{k-1})$
  Solve $H_k \vec{y}_k = \vartheta_k \vec{y}_k$ for k eigenpairs $(\vartheta_k^{(i)}, \vec{y}_k^{(i)})$.
  Select a pair, say, $(\vartheta_k, \vec{y}_k)$, $\vec{y}_k \leftarrow \vec{y}_k / \|\vec{y}_k\|_2$
  $\rho_k = |h_{k+1,k}|\,|e_k^* \vec{y}_k|$
end for
$\mathbf{x} = \mathbf{V}_k \vec{y}_k$, $\lambda = \vartheta_k$.

Approximate eigenpairs

- **Ritz values.**

$$H_k\,\vec{y}_k = \vartheta_k\,\vec{y}_k \qquad \Leftrightarrow \qquad \mathbf{r}_k = (\mathbf{A} - \vartheta_k\,\mathbf{I})\mathbf{V}_k\,\vec{y}_k \perp \mathbf{V}_k$$

- **Harmonic Ritz values.**

$$\underline{H_k}^*(H_k\,\vec{y}_k - \vartheta_k\,\vec{y}_k) = \vec{0} \qquad \Leftrightarrow \qquad \mathbf{r}_k = (\mathbf{A} - \vartheta_k\,\mathbf{I})\mathbf{V}_k\,\vec{y}_k \perp \mathbf{A}\mathbf{V}_k$$

k solution pairs $(\vartheta_k, \vec{y}_k)$.

ϑ_k is a harmonic **Ritz value**,
$\vec{y}_k$ is pre-harmonic-Ritz vector,
$\mathbf{u}_k = \mathbf{V}_k\,\vec{y}_k$ is a harmonic **Ritz vector**.

- **Refined Ritz vectors.**

When an approximate eigenvalue is selected $\tau \in \mathbb{C}$,

minimise $\|\underline{H_k}\,\vec{y}_k - \tau\vec{y}_k\|_2 \quad \Leftrightarrow \quad$ minimise $\|(\mathbf{A} - \tau\,\mathbf{I})\mathbf{V}_k\,\vec{y}_k\|_2$

(Harmonic) Ritz values and (GMRES) FOM residuals

Consider the linear system $\mathbf{Ax} = \mathbf{r}$. Take $\mathbf{v}_1 = \mathbf{r}_0/\|\mathbf{r}_0\|_2$.

On the next transparencies a one-one relation is given between Ritz values and FOM residuals and between harmonic Ritz values and GMRES residuals. This relation provides **theoretical insight**: it allows to relate convergence of (harmonic) Ritz values towards eigenvalues to convergence of FOM (GMRES). More details on consequences for the convergence will be discussed next lecture.

Ritz values and FOM residuals

Consider the linear system $\mathbf{A}\mathbf{x} = \mathbf{r}$. Take $\mathbf{v}_1 = \mathbf{r}_0/\|\mathbf{r}_0\|_2$.

If $\mathbf{r}_k$ is the FOM residual,
then $\mathbf{r}_k = p_k(\mathbf{A})\mathbf{r}_0$ for some polynomial p_k of degree k:
p_k is the so-called kth residual **FOM polynomial**.

Theorem. For a $\vartheta \in \mathbb{C}$ we have that

$$\vartheta \text{ is a Ritz value} \quad \Leftrightarrow \quad p_k(\vartheta) = 0.$$

Proof. If $p_k(\vartheta) = 0$,
then $p_k(\lambda) = (\lambda - \vartheta)q(\lambda)$ for some polynomial q of degree $< k$.

$$\mathbf{r}_k = (\mathbf{A} - \vartheta\,\mathbf{I})q(\mathbf{A})\mathbf{r}_0 = (\mathbf{A} - \vartheta\,\mathbf{I})\mathbf{u}_k \perp \mathbf{V}_k$$

where $\mathbf{u}_k \equiv q(\mathbf{A})\mathbf{r}_0 \in \text{span}(\mathbf{V}_k)$. Hence, $(\vartheta, \mathbf{u})$ is a Ritz pair.

A counting argument completes the proof (there are k Ritz values and p_k has k zeros).

Harmonic Ritz values and GMRES residuals

Consider the linear system $\mathbf{Ax} = \mathbf{r}$. Take $\mathbf{v}_1 = \mathbf{r}_0/\|\mathbf{r}_0\|_2$.

If $\mathbf{r}_k$ is the GMRES residual,
then $\mathbf{r}_k = p_k(\mathbf{A})\mathbf{r}_0$ for some polynomial p_k of degree k
p_k is the so-called kth residual **GMRES polynomial**.

Theorem. For a $\vartheta \in \mathbb{C}$ we have that

$$\vartheta \text{ is a harmonic Ritz value} \quad \Leftrightarrow \quad p_k(\vartheta) = 0.$$

Proof. If $p_k(\vartheta) = 0$,
then $p_k(\lambda) = (\lambda - \vartheta)q(\lambda)$ for some polynomial q of degree $< k$.

$$\mathbf{r}_k = (\mathbf{A} - \vartheta\,\mathbf{I})q(\mathbf{A})\mathbf{r}_0 = (\mathbf{A} - \vartheta\,\mathbf{I})\mathbf{u}_k \perp \mathbf{AV}_k$$

where $\mathbf{u}_k \equiv q(\mathbf{A})\mathbf{r}_0 \in \text{span}(\mathbf{V}_k)$. Hence, $(\vartheta, \mathbf{u})$ is an harmonic Ritz pair.

A counting argument completes the proof (there are k harmonic Ritz values and p_k has k zeros).

Program Lecture 6

- Krylov basis & Hessenberg matrices
- Arnoldi's decomposition
- Linear systems and Arnoldi's decomposition
- GMRES and FOM
- Convergence
- Eigenvalues and Arnoldi's decomposition
- Arnoldi's method
- Convergence
- Stability issues in Arnoldi's decomposition
- Summary

Convergence

The space $\mathcal{K}_k(\mathbf{A}, \mathbf{b}) = \text{span}(\mathbf{V}_k)$ contains all vectors that can be computed with $k-1$ steps of (shifted) power method, and also the vectors computed with a $k-1$-degree polynomial filter.

$\Rightarrow$ faster convergence than any polynomial filter method.

A shifted power method, with appropriate shift, is effective in computing eigenpairs with 'extremal', 'isolated' eigenvalues. Arnoldi (without shift) is even more successful in detecting such eigenpairs.

Achieving better convergence also depends on how the approximate eigenpairs are **extracted** from the **search subspaces** span($\mathbf{V}_k$).

Convergence

The space $\mathcal{K}_k(\mathbf{A}, \mathbf{b}) = \text{span}(\mathbf{V}_k)$ contains all vectors that can be computed with $k-1$ steps of (shifted) power method, and also the vectors computed with a $k-1$-degree polynomial filter.

$\Rightarrow$ faster convergence than any polynomial filter method.

A shifted power method, with appropriate shift, is effective in computing eigenpairs with 'extremal', 'isolated' eigenvalues. Arnoldi (without shift) is even more successful in detecting such eigenpairs.

Achieving better convergence also depends on how the approximate eigenpairs are **extracted** from the **search subspaces** span($\mathbf{V}_k$).

Using Ritz-Galerkin, for extremal eigenvalues

(selecting extremal Ritz values)

Arnoldi:shifted power $\sim$ GMRES:Richardson

Program Lecture 6

- Krylov basis & Hessenberg matrices
- Arnoldi's decomposition
- Linear systems and Arnoldi's decomposition
- GMRES and FOM
- Convergence
- Eigenvalues and Arnoldi's decomposition
- Arnoldi's method
- Convergence
- Stability issues in Arnoldi's decomposition
- Summary

Gram–Schmidt orthogonalisation

$[\mathbf{v}, \vec{h}] = \mathrm{Orth}(\mathbf{V}, \mathbf{w})$, then $\mathbf{v} \perp \mathbf{V}$, $\|\mathbf{v}\|_2 = 1$, $\mathbf{w} = [\mathbf{V}, \mathbf{v}]\vec{h}$

Classical Gram–Schmidt

$$\vec{h} = \mathbf{V}^*\mathbf{w}, \quad \mathbf{v} = \mathbf{w} - \mathbf{V}\vec{h}$$
$$\nu = \|\mathbf{v}\|_2, \quad \vec{h} \leftarrow (\vec{h}^{\mathsf{T}}, \nu)^{\mathsf{T}}, \quad \mathbf{v} \leftarrow \mathbf{v}/\nu$$

Loss of stability.

- Sensitive to perturbations on $\mathbf{V}$.
- DOTs and AXPYs introduce roundig errors.
- Scaling by ν amplifies rounding errors if
$$\tan(\angle(\mathrm{span}(\mathbf{V}), \mathbf{w}) = \nu/\|\vec{h}\|_2 \ll 1.$$

Note. Costs of computing $\|\vec{h}\|_2$ are negligible (wrt costs computing $\|\mathbf{v}\|_2$).

Gram–Schmidt orthogonalisation

$[\mathbf{v}, \vec{h}] = \mathsf{Orth}(\mathbf{V}, \mathbf{w})$, then $\mathbf{v} \perp \mathbf{V}$, $\|\mathbf{v}\|_2 = 1$, $\mathbf{w} = [\mathbf{V}, \mathbf{v}]\vec{h}$

Classical Gram–Schmidt

$\mathbf{v} = \mathbf{w}$

for $j = 1, \ldots, k$ do

$h_j = \mathbf{v}_j^* \mathbf{w}$, $\mathbf{v} \leftarrow \mathbf{v} - \mathbf{v}_j h_j$

end for

$\nu = \|\mathbf{v}\|_2$, $\vec{h} = (h_1, h_2, \ldots, h_k, \nu)^\mathsf{T}$, $\mathbf{v} \leftarrow \mathbf{v}/\nu$

Gram–Schmidt orthogonalisation

$[\mathbf{v}, \vec{h}] = \mathrm{Orth}(\mathbf{V}, \mathbf{w})$, then $\mathbf{v} \perp \mathbf{V}$, $\|\mathbf{v}\|_2 = 1$, $\mathbf{w} = [\mathbf{V}, \mathbf{v}]\vec{h}$

Modified Gram–Schmidt

$\mathbf{v} = \mathbf{w}$

`for` $j = 1, \ldots, k$ `do`

$h_j = \mathbf{v}_j^* \mathbf{v}, \quad \mathbf{v} \leftarrow \mathbf{v} - \mathbf{v}_j h_j$

`end for`

$\nu = \|\mathbf{v}\|_2, \quad \vec{h} = (h_1, h_2, \ldots, h_k, \nu)^\mathsf{T}, \quad \mathbf{v} \leftarrow \mathbf{v}/\nu$

Loss of stability.

- Sensitive to perturbations on $\mathbf{V}$
- **Smaller** rounding errors from AXPYs.
- Scaling by ν amplifies rounding errors if $\nu / \|\vec{h}\|_2 \ll 1$

\+ More stable. − Harder to parallelise.

Gram–Schmidt orthogonalisation

$[\mathbf{v}, \vec{h}] = \mathrm{Orth}(\mathbf{V}, \mathbf{w})$, then $\mathbf{v} \perp \mathbf{V}$, $\|\mathbf{v}\|_2 = 1$, $\mathbf{w} = [\mathbf{V}, \mathbf{v}]\vec{h}$

Repeated Gram–Schmidt with **DGKS** criterion

$\vec{h} = \mathbf{V}^*\mathbf{w}$, $\mathbf{v} = \mathbf{w} - \mathbf{V}\vec{h}$
$\nu = \|\mathbf{v}\|_2$, $\mu = \|\vec{h}\|_2$
`while` $\nu \leq \tau\mu$
  $\vec{g} = \mathbf{V}^*\mathbf{v}$, $\mathbf{v} \leftarrow \mathbf{v} - \mathbf{V}\vec{g}$
  $\nu = \|\mathbf{v}\|_2$, $\mu = \|\vec{g}\|_2$, $\vec{h} \leftarrow \vec{h} + \vec{g}$
`end while`
$\vec{h} \leftarrow (\vec{h}^\mathsf{T}, \nu)^\mathsf{T}$, $\mathbf{v} \leftarrow \mathbf{v}/\nu$

Repeat if $\frac{\|\mathbf{v}\|_2}{\|\vec{h}\|_2}$ (=$\tan\angle(\mathbf{w}, \mathbf{W})$) is too small ($\leq$ preselected τ.)

Property. Twice is enough ($\exists$ proof if $\mathbf{V} = [\mathbf{v}_1]$).

Daniel **G**rag **K**aufmann **S**tewart

Gram–Schmidt orthogonalisation

$[\mathbf{v}, \vec{h}] = \mathrm{Orth}(\mathbf{V}, \mathbf{w})$, then $\mathbf{v} \perp \mathbf{V}$, $\|\mathbf{v}\|_2 = 1$, $\mathbf{w} = [\mathbf{V}, \mathbf{v}]\vec{h}$

Repeated Gram–Schmidt with **DGKS** criterion

$\vec{h} = \mathbf{V}^*\mathbf{w}, \quad \mathbf{v} = \mathbf{w} - \mathbf{V}\vec{h}$

$\nu = \|\mathbf{v}\|_2, \quad \mu = \|\vec{h}\|_2$

`while` $\nu \leq \tau\mu$

$\quad \vec{g} = \mathbf{V}^*\mathbf{v}, \quad \mathbf{v} \leftarrow \mathbf{v} - \mathbf{V}\vec{g}$

$\quad \nu = \|\mathbf{v}\|_2, \quad \mu = \|\vec{g}\|_2, \quad \vec{h} \leftarrow \vec{h} + \vec{g}$

`end while`

$\vec{h} \leftarrow (\vec{h}^\mathsf{T}, \nu)^\mathsf{T}, \quad \mathbf{v} \leftarrow \mathbf{v}/\nu$

Loss of stability.

- **Not** sensitive to perturbations on $\mathbf{V}$
- **Smaller** rounding errors from AXPYs.
- Scaling by ν amplifies rounding errors if $\nu/\|\vec{h}\|_2 \ll 1$

Gram–Schmidt orthogonalisation

$[\mathbf{v}, \vec{h}] = \mathrm{Orth}(\mathbf{V}, \mathbf{w})$, then $\mathbf{v} \perp \mathbf{V}$, $\|\mathbf{v}\|_2 = 1$, $\mathbf{w} = [\mathbf{V}, \mathbf{v}]\vec{h}$

Repeated Gram–Schmidt with **DGKS** criterion

$\vec{h} = \mathbf{V}^*\mathbf{w}, \quad \mathbf{v} = \mathbf{w} - \mathbf{V}\vec{h}$

$\nu = \|\mathbf{v}\|_2, \quad \mu = \|\vec{h}\|_2$

`while` $\nu \leq \tau\mu$

$\quad \vec{g} = \mathbf{V}^*\mathbf{v}, \quad \mathbf{v} \leftarrow \mathbf{v} - \mathbf{V}\vec{g}$

$\quad \nu = \|\mathbf{v}\|_2, \quad \mu = \|\vec{g}\|_2, \quad \vec{h} \leftarrow \vec{h} + \vec{g}$

`end while`

$\vec{h} \leftarrow (\vec{h}^\mathsf{T}, \nu)^\mathsf{T}, \quad \mathbf{v} \leftarrow \mathbf{v}/\nu$

\+ Stable (depends on κ, typical value $\kappa = 0.5$).

\+ Easy to parallelize.

\− More costly (# flops = 1 a 2 $\times$ larger)

Stability of the Gram–Schmidt variants

Orthogonalisation recursively applied to the columns of an $n \times k$ matrix $\mathbf{W}$ leads to computed $\widehat{\mathbf{V}}$ and $\widehat{R}$ such that

$$\mathbf{W} + \Delta = \widehat{\mathbf{V}}\,\widehat{R}$$

for some $n \times k$ perturbation matrix Δ with

- $\widehat{R}$ is $k \times k$ upper triangular,
- $\|\Delta\|_F \leq 4\,k^2\,\mathrm{u}\,\|\mathbf{W}\|_F$,
- Loss of orthogonality: $\|\widehat{\mathbf{V}}^*\,\widehat{\mathbf{V}} - I_k\|_2 \leq \kappa\,\mathrm{u}\,(\mathcal{C}_2(\mathbf{W}))^{\ell}$

with

u (=1.1e-16) the **relative machine precision**,

$\mathcal{C}_2(\mathbf{W})$ the **condition number of $\mathbf{W}$**: $(z \in \mathbb{C}^k \backslash \{0\})$

$$\mathcal{C}_2(\mathbf{W}) \equiv \max_z \frac{\|\mathbf{W}z\|_2}{\|z\|_2} / \min_z \frac{\|\mathbf{W}z\|_2}{\|z\|_2}$$

Stability of the Gram–Schmidt variants

Orthogonalisation recursively applied to the columns of an $n \times k$ matrix $\mathbf{W}$ leads to computed $\widehat{\mathbf{V}}$ and $\widehat{R}$ such that

$$\mathbf{W} + \Delta = \widehat{\mathbf{V}}\,\widehat{R}$$

for some $n \times k$ perturbation matrix Δ with

- $\widehat{R}$ is $k \times k$ upper triangular,
- $\|\Delta\|_F \leq 4\,k^2\,\mathrm{u}\,\|\mathbf{W}\|_F$,
- Loss of orthogonality: $\|\widehat{\mathbf{V}}^*\,\widehat{\mathbf{V}} - I_k\|_2 \leq \kappa\,\mathrm{u}\,(\mathcal{C}_2(\mathbf{W}))^\ell$

with κ, ℓ depend on orth. method:

ClassGS: κ of order $\sqrt{kn}$, $\ell = 2$ (conjecture).

ModGS: κ of order $\sqrt{kn}$, $\ell = 1$.

RepGS: κ may depend on $\frac{1}{\tau^k}$ (rarely), $\ell = 0$.

Householder QR: $\kappa = \mathcal{O}(\sqrt{kn})$, $\ell = 0$.

Gram–Schmidt and Arnoldi

Theorem. Modified Gram–Schmidt is suffciently stable for solving linear systems.

Proof. In Arnoldi, the $n \times (k+1)$ matrix $\mathbf{W}$ is

$$\mathbf{W} = [\mathbf{v}_1, \mathbf{A}\mathbf{v}_1, \ldots, \mathbf{A}\mathbf{v}_k] \quad \text{and} \quad R = \begin{bmatrix} 1 & \\ \mathbf{0} & \underline{H}_k \end{bmatrix}.$$

Hence, when solving $\mathbf{A}\mathbf{x} = \mathbf{r}_0$ with $\mathbf{x}_k = \mathbf{V}_k \vec{y}_k$, we have

$$\min \frac{\|\mathbf{W}z\|_2}{\|z\|_2} \leq \frac{\|\mathbf{r}_k\|_2}{\|\mathbf{r}_0\|_2} \qquad (\text{take } z = (-\|\mathbf{r}_0\|_2, \vec{y}_k^{\,\mathsf{T}})^{\mathsf{T}}).$$

Therefore, we have the (sharp) estimate

$$\mathcal{C}_2(\mathbf{W}) \gtrsim \|\mathbf{A}\|_2 \frac{\|\mathbf{r}_0\|_2}{\|\mathbf{r}_k\|_2}.$$

Prop. Eigenvalue computations requires more stability.

Program Lecture 6

- Krylov basis & Hessenberg matrices
- Arnoldi's decomposition
- Linear systems and Arnoldi's decomposition
- GMRES and FOM
- Convergence
- Eigenvalues and Arnoldi's decomposition
- Arnoldi's method
- Convergence
- Stability issues in Arnoldi's decomposition
- Summary

Arnoldi's decomposition based methods

1) **Expansion.** Use recursive expansion for building a Krylov basis $\mathbf{V}_k$ (involves high dimensional operations)

2) **Extraction.** For theoretical analysis, consider a projected problem as:

$$\mathbf{b} - \mathbf{A}\,\mathbf{V}_k\,\vec{y}_k \perp \mathbf{V}_k, \qquad \mathbf{b} - \mathbf{A}\,\mathbf{V}_k\,\vec{y}_k \perp \mathbf{A}\mathbf{V}_k$$

$$\vartheta\,\mathbf{V}_k\,\vec{y}_k - \mathbf{A}\,\mathbf{V}_k\,\vec{y}_k \perp \mathbf{V}_k, \quad \text{or} \quad \vartheta\,\mathbf{V}_k\,\vec{y}_k - \mathbf{A}\,\mathbf{V}_k\,\vec{y}_k \perp \mathbf{A}\mathbf{V}_k$$

For practical computations,

2.a) Form a projected matrix, as $H_k = \mathbf{V}_k^* \mathbf{A} \mathbf{V}_k$. (high dim)

2.b) Use the projected matrix to solve the projected problem for $\vec{y}_k$ in k-space (only k-dimensional operations)

2.c) Assemble $\mathbf{x}_k = \mathbf{V}_k\,\vec{y}_k$. (high dim)

Note. When recursively using Gram-Schmidt to compute the component of $\mathbf{A}\mathbf{v}_k$ that is orthonormal to $\mathbf{V}_k$, the projected matrix H_k comes for free.

Krylov subspace methods

Krylov subspace methods search for approximate solutions in a krylov subspace: the search subspace is a Krylov subspace.

Stages.

- **Expansion.**

  Expand a Krylov basis $\mathbf{v}_1, \ldots, \mathbf{v}_k$ recursively

- **Extraction.**

  Extract an approximate solution from span($\mathbf{V}_k$)

- If space becomes too large

  Shrinking. (Restart)

  For some $\ell < k$, select a Krylov basis $\widetilde{\mathbf{v}}_1, \ldots, \widetilde{\mathbf{v}}_\ell$ in the space span($\mathbf{V}_k$) such that span($\widetilde{\mathbf{V}}_\ell$) contains promising approximations.

Subspace methods or Projection methods

1) **Expansion.** Use recursive expansion for building a Krylov basis $\mathbf{V}_k$ (involves high dimensional operations)

2) **Extraction.** For theoretical analysis, consider a projected problem as:

$$\mathbf{b} - \mathbf{A}\,\mathbf{V}_k\,\vec{y}_k \perp \mathbf{V}_k, \qquad \mathbf{b} - \mathbf{A}\,\mathbf{V}_k\,\vec{y}_k \perp \mathbf{A}\mathbf{V}_k$$

$$\vartheta\,\mathbf{V}_k\,\vec{y}_k - \mathbf{A}\,\mathbf{V}_k\,\vec{y}_k \perp \mathbf{V}_k, \quad \text{or} \quad \vartheta\,\mathbf{V}_k\,\vec{y}_k - \mathbf{A}\,\mathbf{V}_k\,\vec{y}_k \perp \mathbf{A}\mathbf{V}_k$$

For practical computations,

2.a) Form a projected matrix, as $H_k = \mathbf{V}_k^* \mathbf{A} \mathbf{V}_k$. (high dim)

2.b) Use the projected matrix to solve the projected problem for $\vec{y}_k$ in k-space (only k-dimensional operations)

2.c) Assemble $\mathbf{x}_k = \mathbf{V}_k\,\vec{y}_k$. (high dim)

Note. When recursively using Gram-Schmidt to compute the component of $\mathbf{A}\mathbf{v}_k$ that is orthonormal to $\mathbf{V}_k$, the projected matrix H_k comes for free.

Krylov subspace methods

Why searching for approximations in Krylov subspaces?

1) Convergence based on polynomial approximation theory (better than Richardson, Power method, etc.)

2) Krylov structure can be exploited to enhance efficiency.

For instance,

- with Arnoldi's method, the Hessenberg matrix (projected matrix) comes for free.
- if **A** is Hermitian then expansion vectors can efficiently be computed (as in CR, CG, . . .)

Subspace methods

Why searching for approximations in general subspaces?

To allow detection of more effective expansion vectors (as in Flexible GCR, and Flexible GMRES).