Overview

- Nonsymmetric systems revisited
- Biconjugate gradient
- Preconditioning

- Arnoldi iteration: $AQ_n = Q_{n+1}\tilde{H}_n \Rightarrow Aq_n = h_{1n}q_1 + \cdots + h_{n+1,n}q_{n+1}$, long recurrence
- Lanczos, $A = A^T$: $AQ_n = Q_{n+1}\tilde{T}_n \Rightarrow Aq_n = \beta_{n-1}q_{n-1} + \alpha_n q_n + \beta_n q_{n+1}$, 3-term
- Approaches to obtaining a short-term recurrence algorithm for non-symmetric matrices:
  - Conjugate gradient applied to normal equations (CGN)

$$Ax = b \Rightarrow A^*Ax = A^*b \Leftrightarrow Mx = y, M = A^*A = M^*, y = A^*b$$

Recall convergence estimate

$$\frac{\|r_n\|_2}{\|r_0\|_0} \leqslant 2\left(\frac{\sqrt{\kappa_M}-1}{\sqrt{\kappa_M}+1}\right)^2 = 2\left(\frac{\kappa_A-1}{\kappa_A+1}\right)^2, \mathcal{O}(\kappa_A) \text{ iterations}$$

  - Biorthogonalization: use $A = VTV^{-1}$ (compare to Lanczos for $A = A^T$, $A = QTQ^T$)

$$A = VTV^{-1} \qquad A^* = V^{-*}T^*(V^{-*})^{-1}$$
$$(V^{-*})^*V = W^*V = I \text{ replaces } QQ^* = I$$
$$W = (\begin{array}{ccc} w_1 & \ldots & w_m \end{array}), V = (\begin{array}{ccc} v_1 & \ldots & v_m \end{array}), w_i^* v_j = \delta_{ij}$$

- Full biorthogonalization: $\boldsymbol{W}^* \boldsymbol{V} = \boldsymbol{V}^* \boldsymbol{W} = \boldsymbol{I}$
- Partial biorthgonalization: $\boldsymbol{W}_n^* \boldsymbol{V}_n = \boldsymbol{V}_n^* \boldsymbol{W}_n = \boldsymbol{I}_n$, $n \leqslant m$

Arnoldi	$\boldsymbol{A} \boldsymbol{Q}_n = \boldsymbol{Q}_{n+1} \tilde{\boldsymbol{H}}_n$	
Lanczos	$\boldsymbol{A} \boldsymbol{Q}_n = \boldsymbol{Q}_{n+1} \tilde{\boldsymbol{T}}_n$	
Biorthogonalization	$\boldsymbol{A} \boldsymbol{V}_n = \boldsymbol{V}_{n+1} \tilde{\boldsymbol{T}}_n$	$\boldsymbol{V}_n, \boldsymbol{W}_n \in \mathbb{C}^{m \times n}$
	$\boldsymbol{A}^* \boldsymbol{W}_n = \boldsymbol{W}_{n+1} \tilde{\boldsymbol{S}}_n$	$\tilde{\boldsymbol{S}}_n, \tilde{\boldsymbol{T}}_n \in \mathbb{C}^{(n+1) \times n}$
	$\boldsymbol{T}_n = \boldsymbol{S}_n^* = \boldsymbol{W}_n^* \boldsymbol{A} \boldsymbol{V}_n$	

- Theoretical framework: two Krylov spaces from arbitrary starting vectors $\boldsymbol{v}_1, \boldsymbol{w}_1$

$$\boldsymbol{v}_n \in \langle \boldsymbol{v}_1, \boldsymbol{A} \boldsymbol{v}_1, \ldots, \boldsymbol{A}^{n-1} \boldsymbol{v}_1 \rangle, \boldsymbol{w}_n \in \langle \boldsymbol{w}_1, \boldsymbol{A}^* \boldsymbol{w}_1, \ldots, (\boldsymbol{A}^*)^{n-1} \boldsymbol{w}_1 \rangle$$

- Applications:
  - Eigenvalues: Ritz values, i.e., eigenvalues of $\boldsymbol{T}_n$ can rapidly converge to eigenvalues of $\boldsymbol{A}$
  - Biconjugate gradient to solve $\boldsymbol{A} \boldsymbol{x} = \boldsymbol{b}$

- GMRES:

$$\boldsymbol{x} \in \mathcal{K}_n = \mathcal{R}(\boldsymbol{K}_n), \qquad \boldsymbol{K}_n = \begin{pmatrix} \boldsymbol{b} & \boldsymbol{A}\boldsymbol{b} & \ldots & \boldsymbol{A}^{n-1}\boldsymbol{b} \end{pmatrix}$$
$$\boldsymbol{r}_n \perp \mathcal{A}\mathcal{K}_n = \mathcal{R}(\boldsymbol{A}\,\boldsymbol{K}_n)$$

- BCG:

$$\boldsymbol{x} \in \mathcal{K}_n = \mathcal{R}(\boldsymbol{K}_n), \quad \boldsymbol{K}_n = \begin{pmatrix} \boldsymbol{b} & \boldsymbol{A}\boldsymbol{b} & \ldots & \boldsymbol{A}^{n-1}\boldsymbol{b} \end{pmatrix}$$
$$\boldsymbol{r}_n \perp \mathcal{L}_n = \mathcal{R}(\boldsymbol{L}_n), \quad \boldsymbol{L}_n = \begin{pmatrix} \boldsymbol{w}_1 & \boldsymbol{A}^*\boldsymbol{w}_1 & \ldots & (\boldsymbol{A}^*)^{n-1}\boldsymbol{w}_1 \end{pmatrix}$$

Algorithm

$\boldsymbol{x}_0 = \boldsymbol{0}$, $\boldsymbol{p}_0 = \boldsymbol{r}_0 = \boldsymbol{b}$, $\boldsymbol{q}_0 = \boldsymbol{s}_0$ arbitrary

for $n = 1, 2, 3, \ldots$

$$\alpha_n = (\boldsymbol{s}_{n-1}^*\,\boldsymbol{r}_{n-1}) / (\boldsymbol{q}_{n-1}^*\boldsymbol{A}\,\boldsymbol{p}_{n-1})$$
$$\boldsymbol{x}_n = \boldsymbol{x}_{n-1} + \alpha_n\,\boldsymbol{p}_{n-1}$$
$$\boldsymbol{r}_n = \boldsymbol{r}_{n-1} - \alpha_n\,\boldsymbol{A}\,\boldsymbol{p}_{n-1}$$
$$\boldsymbol{s}_n = \boldsymbol{s}_{n-1} - \bar{\alpha}_n\,\boldsymbol{A}^*\,\boldsymbol{q}_{n-1}$$
$$\beta_n = (\boldsymbol{s}_n^*\,\boldsymbol{r}_n) / (\boldsymbol{s}_{n-1}^*\,\boldsymbol{r}_{n-1})$$
$$\boldsymbol{p}_n = \boldsymbol{r}_n + \beta_n\,\boldsymbol{p}_{n-1}$$
$$\boldsymbol{q}_n = \boldsymbol{s}_n + \bar{\beta}_n\,\boldsymbol{q}_{n-1}$$

- Convergence of Krylov methods on $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ depends on spectrum $\Lambda(\boldsymbol{A})$
- Preconditioning idea: modify the spectrum

$$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}, \boldsymbol{M}\ \mathrm{non-singular} \Rightarrow \boldsymbol{M}^{-1}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{M}^{-1}\boldsymbol{b}$$

  with operator $\boldsymbol{M}^{-1}$ denoting solution of $\boldsymbol{M}\boldsymbol{y}=\boldsymbol{c}$

- Choice of $\boldsymbol{M}$
  - Ideally, $\boldsymbol{M}^{-1}=\boldsymbol{A}^{-1} \Rightarrow \boldsymbol{I}\boldsymbol{x}=\boldsymbol{M}^{-1}\boldsymbol{b}$
  - In practice choose $\boldsymbol{M}^{-1} \cong \boldsymbol{A}^{-1}$:
    - $\rightarrow$ partial $\boldsymbol{L}\boldsymbol{U}$ factorization
    - $\rightarrow$ coarse grid approximation (multigrid)
    - $\rightarrow$ low-order discretization
    - $\rightarrow$ operator splitting
    - $\rightarrow$ dimensional splitting