Overview

- Polynomial approximation
- Eigenvalue approximation
- Arnoldi lemniscates
- GMRES
- Applications:
  - the Harwell-Boeing Matrix Collection
  - JupyterLab, Fortran, f2py, FortranMagic

- $\boldsymbol{x} \in \mathcal{K}_n \Rightarrow \boldsymbol{x} = c_0 \boldsymbol{b} + c_1 \boldsymbol{A}\boldsymbol{b} + \cdots + c_{n-1} \boldsymbol{A}^{n-1} \boldsymbol{b}$, Introduce $q(z) = c_0 + c_1 z + \cdots + c_n z^{n-1}$

$$\boldsymbol{x} = q(\boldsymbol{A})\,\boldsymbol{b} = \boldsymbol{K}_n\,\boldsymbol{c} = \begin{pmatrix} \boldsymbol{b} & \boldsymbol{A}\boldsymbol{b} & \ldots & \boldsymbol{A}^{n-1}\boldsymbol{b} \end{pmatrix} \boldsymbol{c}$$

- $\boldsymbol{Q}_n$ is an orthonormal basis for $\mathcal{K}_n$, $\boldsymbol{Q}_n \boldsymbol{R}_n = \boldsymbol{K}_n$ is constructed by Arnoldi iteration, and $\boldsymbol{H}_n = \boldsymbol{Q}_n^* \boldsymbol{A} \boldsymbol{Q}_n$ is the restriction of operator $\mathcal{A}$ encoded by matrix $\boldsymbol{A}$ to $\mathcal{K}_n$. Let $-\boldsymbol{y}$ denote coordinates of $\boldsymbol{x}$ in $\boldsymbol{Q}_n$, $\boldsymbol{K}_n \boldsymbol{c} = -\boldsymbol{Q}_n \boldsymbol{y} = \boldsymbol{I}\boldsymbol{x}$. Note that $-\boldsymbol{y}$ is a representation of $q(z)$.

Theorem. *Let $\boldsymbol{A} \in \mathbb{C}^{m \times m}$, $\boldsymbol{b} \in \mathbb{C}^m$, and P^n denote the space of monic polynomials of degree n. If $\boldsymbol{K}_n = \begin{pmatrix} \boldsymbol{b} & \boldsymbol{A}\boldsymbol{b} & \ldots & \boldsymbol{A}^{n-1}\boldsymbol{b} \end{pmatrix}$ is of full rank, the solution of*

$$\min_{p^n \in P^n} \|p^n(\boldsymbol{A})\,\boldsymbol{b}\|$$

is the characteristic polynomial

$$p_{\boldsymbol{H}_n}(z) = \det(z\boldsymbol{I} - \boldsymbol{H}_n)$$

of $\boldsymbol{H}_n = \boldsymbol{Q}_n^* \boldsymbol{A} \boldsymbol{Q}_n$, with $\boldsymbol{Q}_n = \begin{pmatrix} \boldsymbol{q}_1 & \ldots & \boldsymbol{q}_n \end{pmatrix}$ constructed by Arnoldi iteration, i.e.,

$$p_{\boldsymbol{H}_n} = \arg \min_{p^n \in P^n} \|p^n(\boldsymbol{A})\,\boldsymbol{b}\|.$$

Polynomial approximation theorem proof

Proof. Consider some $p \in P^n$, i.e., some monic polynomial of degree n,

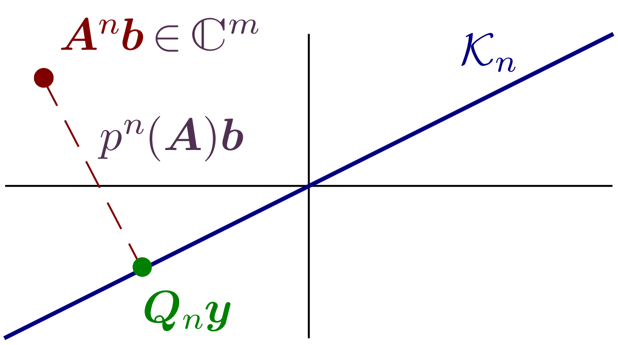
$$p^n(z) = z^n + c_{n-1} z^{n-1} + \cdots + c_1 z + c_0 = z^n + q(z) \Rightarrow p^n(\boldsymbol{A})\boldsymbol{b} = \boldsymbol{A}^n \boldsymbol{b} - \boldsymbol{Q}_n \boldsymbol{y},$$

and the polynomial approximation problem $\min_{p^n \in P^n} \| p^n(\boldsymbol{A})\, \boldsymbol{b} \|$ can be restated as

$$\min_{\boldsymbol{y}} \| \boldsymbol{A}^n \boldsymbol{b} - \boldsymbol{Q}_n \boldsymbol{y} \|,$$

a least-squares problem, with $\boldsymbol{K}_n \boldsymbol{c} = -\boldsymbol{Q}_n \boldsymbol{y} = \boldsymbol{I}\boldsymbol{x} = q(\boldsymbol{A})\, \boldsymbol{b}$. Recall that $(\boldsymbol{Q}\boldsymbol{H}\boldsymbol{Q}^*)^n = \boldsymbol{Q}\boldsymbol{H}^n \boldsymbol{Q}^*$.

$$\begin{aligned}
& p^n(\boldsymbol{A})\boldsymbol{b} = \boldsymbol{A}^n \boldsymbol{b} - \boldsymbol{Q}_n \boldsymbol{y} \perp \mathcal{K}_n \Rightarrow && \boldsymbol{Q}_n^* p^n(\boldsymbol{A})\boldsymbol{b} = \boldsymbol{0}_n. \\
& \boldsymbol{A} = \boldsymbol{Q}\boldsymbol{H}\boldsymbol{Q}^*,\ \boldsymbol{Q} = \begin{pmatrix} \boldsymbol{Q}_n & \boldsymbol{U} \end{pmatrix} && \boldsymbol{H} = \begin{pmatrix} \boldsymbol{H}_n & \times \\ \times & \times \end{pmatrix} \\
& \boldsymbol{Q}_n^* p^n(\boldsymbol{A})\boldsymbol{b} = \boldsymbol{Q}_n^* p^n(\boldsymbol{Q}\boldsymbol{H}\boldsymbol{Q}^*)\boldsymbol{b} = && \boldsymbol{Q}_n^* \boldsymbol{Q} p^n(\boldsymbol{H})\boldsymbol{Q}^* \boldsymbol{b} \\
& \boldsymbol{Q}^* \boldsymbol{b} = \begin{pmatrix} \boldsymbol{q}_1 & \ldots & \boldsymbol{q}_n & \ldots & \boldsymbol{q}_m \end{pmatrix} \boldsymbol{b} = && \|\boldsymbol{b}\|\, \boldsymbol{e}_1 \\
& \boldsymbol{Q}_n^* \boldsymbol{Q} = \begin{pmatrix} \boldsymbol{I}_n & \boldsymbol{0}_{n,m-n} \end{pmatrix} && \\
& \begin{pmatrix} \boldsymbol{I}_n & \boldsymbol{0}_{n,m-n} \end{pmatrix} p^n(\boldsymbol{H})\boldsymbol{e}_1 = \boldsymbol{0}_n \Rightarrow && (p^n(\boldsymbol{H}))_{1:n,1} = 0 \\
& (p^n(\boldsymbol{H}))_{1:n,1} = (p^n(\boldsymbol{H}_n))_{:,1} = 0 && \\
& \text{Cayley-Hamilton} \Rightarrow && p^n = p_{\boldsymbol{H}_n}
\end{aligned}$$



□

- Solution of $\min_{p^n \in P^n} \|p^n(\boldsymbol{A})\boldsymbol{b}\|$ is characteristic polynomial of $\boldsymbol{H}_n = \boldsymbol{Q}_n^* \boldsymbol{A} \boldsymbol{Q}_n \Rightarrow$ the eigenvalues of $\boldsymbol{H}_n$ are useful approximations of eigenvalues of $\boldsymbol{A}$
- Eigenvalues of $\boldsymbol{H}_n$ are denoted $\{\theta_j\}$ $j = 1, \ldots, n$, and called the *Ritz* (Arnoldi) values (The Rayleigh quotient is also known as a Rayleigh-Ritz approximation)
- Invariance properties:
  - $\boldsymbol{A} \rightarrow \boldsymbol{A} + \sigma \boldsymbol{I} \Rightarrow \theta_j \rightarrow \theta_j + \sigma$
  - $\boldsymbol{A} \rightarrow \sigma \boldsymbol{A} \Rightarrow \theta_j \rightarrow \sigma \theta_j$
  - $\boldsymbol{A} \rightarrow \boldsymbol{U} \boldsymbol{A} \boldsymbol{U}^*$ and $\boldsymbol{b} \rightarrow \boldsymbol{U}\boldsymbol{b} \Rightarrow \theta_j \rightarrow \theta_j$
- *Lemniscate* is curve in $\mathbb{C}$ for which $|p(z)| = C$
- Arnoldi iteration lemniscate

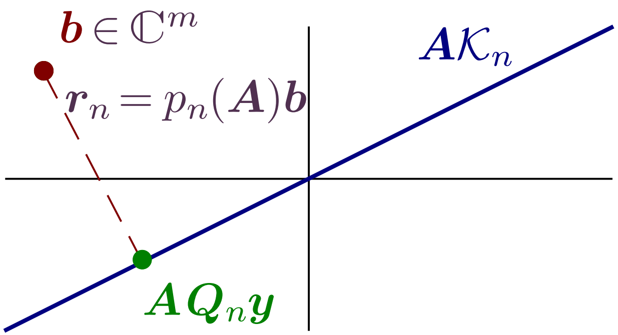
$$C = \frac{\|p^n(\boldsymbol{A})\,\boldsymbol{b}\|}{\|\boldsymbol{b}\|} = \frac{\|p_{\boldsymbol{H}_n}(\boldsymbol{A})\,\boldsymbol{b}\|}{\|\boldsymbol{b}\|}$$

- Arnoldi iteration can also be used to solve linear systems, $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, $\mathcal{N}(\boldsymbol{A}) = \{\boldsymbol{0}\}$, $\boldsymbol{x}_* = \boldsymbol{A}^{-1}\boldsymbol{b}$
- GMRES: $\boldsymbol{x}_n = \arg\min_{\boldsymbol{z} \in \mathcal{K}_n} \|\boldsymbol{b} - \boldsymbol{A}\boldsymbol{z}\|$. The following problems are all equivalent:

$\min_{\boldsymbol{z} \in \mathcal{K}_n} \|\boldsymbol{b} - \boldsymbol{A}\boldsymbol{z}\|$ $\quad$ $\min_{\boldsymbol{c} \in \mathbb{C}^n} \|\boldsymbol{A}\boldsymbol{K}_n\boldsymbol{c} - \boldsymbol{b}\|$

$\min_{\boldsymbol{y} \in \mathbb{C}^n} \|\boldsymbol{A}\boldsymbol{Q}_n\boldsymbol{y} - \boldsymbol{b}\|$ $\quad$ $\min_{\boldsymbol{y} \in \mathbb{C}^n} \|\boldsymbol{Q}_{n+1}\tilde{\boldsymbol{H}}_n\boldsymbol{y} - \boldsymbol{b}\|$

$\min_{\boldsymbol{y} \in \mathbb{C}^n} \|\tilde{\boldsymbol{H}}_n\boldsymbol{y} - \boldsymbol{Q}_{n+1}^*\boldsymbol{b}\|$ $\quad$ $\min_{\boldsymbol{y} \in \mathbb{C}^n} \|\tilde{\boldsymbol{H}}_n\boldsymbol{y} - \|\boldsymbol{b}\|\,\boldsymbol{e}_1\|$



- GMRES solves the least squares problem $\min_{\boldsymbol{y} \in \mathbb{C}^n} \|\tilde{\boldsymbol{H}}_n\boldsymbol{y} - \|\boldsymbol{b}\|\,\boldsymbol{e}_1\|$ at each Arnoldi iteration, and furnishes the approximation

$$\boldsymbol{x}_n = \boldsymbol{Q}_n\boldsymbol{y}$$

to $\boldsymbol{x}_* = \boldsymbol{A}^{-1}\boldsymbol{b}$.

GMRES polynomial approximation

- $P_n = \{c_n z^n + c_{n-1} z^{n-1} + \cdots + c_1 z + 1, \boldsymbol{c} \in \mathbb{C}^n\}$
- $\boldsymbol{x}_n = \boldsymbol{Q}_n \boldsymbol{y} = \boldsymbol{K}_n \boldsymbol{c} = \begin{pmatrix} \boldsymbol{b} & \boldsymbol{A}\boldsymbol{b} & \ldots & \boldsymbol{A}^{n-1}\boldsymbol{b} \end{pmatrix} \boldsymbol{c} = q(\boldsymbol{A})\boldsymbol{b}$ (q of degree $n-1$)
- $\boldsymbol{r}_n = \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}_n = (\boldsymbol{I} - \boldsymbol{A} q(\boldsymbol{A}))\boldsymbol{b}$
- Define polynomial $p_n(z) = 1 - z q(z)$ of degree $n \Rightarrow \boldsymbol{r}_n = p_n(\boldsymbol{A})\,\boldsymbol{b}$
- GMRES is therefore equivalent to polynomial approximation problem

$$\min_{p_n \in P_n} \|p_n(\boldsymbol{A})\boldsymbol{b}\|$$

- Invariance properties: $\boldsymbol{A} \to \sigma \boldsymbol{A}, \boldsymbol{b} \to \sigma \boldsymbol{b} \Rightarrow \boldsymbol{r}_n \to \sigma \boldsymbol{r}_n$, $\boldsymbol{A} \to \boldsymbol{U}\boldsymbol{A}\boldsymbol{U}^*, \boldsymbol{b} \to \boldsymbol{U}\boldsymbol{b} \Rightarrow \boldsymbol{r}_n \to \boldsymbol{U}^* \boldsymbol{r}_n$
- Convergence of GMRES determined by $\|p_n(\boldsymbol{A})\|$

$$\frac{\|\boldsymbol{r}_n\|}{\|\boldsymbol{b}\|} \leqslant \inf_{p_n \in P_n} \|p_n(\boldsymbol{A})\|$$

- Convergence rate bounds: $\boldsymbol{A}\boldsymbol{X} = \boldsymbol{X}\boldsymbol{\Lambda}$, $\|p\|_{\Lambda(\boldsymbol{A})} = \sup_{z \in \Lambda(\boldsymbol{A})} |p(z)|$

$$\frac{\|\boldsymbol{r}_n\|}{\|\boldsymbol{b}\|} \leqslant \inf_{p_n \in P_n} \|p_n(\boldsymbol{A})\| \leqslant \kappa(\boldsymbol{X}) \inf_{p_n \in P_n} \|p_n\|_{\Lambda(\boldsymbol{A})}$$