Last time: $\qquad A \in \mathbb{C}^{n\times n}$, eigenvalues $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|$
$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad v_1, \quad v_2, \quad \cdots, \quad v_n$

- Krylov subspaces

Given $x_0 \in \mathbb{C}^n$, $\|x_0\|_2 = 1$

$$
\begin{aligned}
&\mathcal{K}_k(A, x_0) \\
&= \operatorname{span}\{x_0, Ax_0, \cdots, A^{k-1}x_0\} \\
&= \{p(A)x_0 \in \mathbb{C}^k : p(A) = \sum_{l=0}^{k-1} t_l A^l \text{ is a polynomial of degree} \leq k-1\}
\end{aligned}
$$

- Krylov matrix

$$K_k(A, x_0) = \begin{bmatrix} | & | & & | \\ x_0 & Ax_0 & \cdots & A^{k-1}x_0 \\ | & | & & | \end{bmatrix} = Q_k R_k, \quad Q_k \in \mathbb{C}^{n\times k}.$$

- Rayleigh–Ritz projection

$$B_k z = \tilde{\lambda} z$$

$B_k = Q_k^* A Q_k \in \mathbb{C}^{k\times k}$

$\tilde{\lambda}$ Ritz value, $\quad \tilde{v} = Q_k z$ Ritz vector

- Arnoldi's iteration

$$x_0 \longrightarrow Ax_0 \longrightarrow A^2x_0 \longrightarrow \cdots \longrightarrow A^{k-1}x_0$$

$$
\begin{tikzcd}
x_0 \arrow[d, Rightarrow] & Aq_1 \arrow[d, Rightarrow] & Aq_2 \arrow[d, Rightarrow] & \cdots & Aq_{k-1} \arrow[d, Rightarrow] \\
q_1 \arrow[ur] & q_2 \arrow[ur] & q_3 \arrow[ur] & \cdots \arrow[ur] & q_k
\end{tikzcd}
$$

<u>Implementation</u>:

Given $x_0 \in \mathbb{C}^n$, $k \geq 1$,

$q_1 \leftarrow x_0 / \|x_0\|_2$

For $i = 1, \cdots, k$

$\quad w_i \leftarrow Aq_i$

For $j = 1, \dots, i$

$h_{ji} \leftarrow q_j^* w_i$

$w_i \leftarrow w_i - h_{ji} q_j$

end

$h_{i,i+1} = \|w_i\|_2$

if $h_{i,i+1} = 0$, break.

$q_{i+1} = w_i / \|w_i\|_2$

end

Output: $q_1, \dots, q_k, q_{k+1}, h_{ji}$

Remark: If $h_{k+1,k} = 0$ for some k, Arnoldi iteration breakdowns

But this is a lucky breakdown because the Ritz values and vectors are eig. values and vectors of A.

- Arnoldi decomposition: $Q_k \in \mathbb{C}^{n \times k}$ orthonormal columns, H_k upper Hessenberg

$$A Q_k = Q_k H_k + h_{k+1,k} q_{k+1} e_k^*$$

$q_{k+1}^* H_k = 0$,

$e_k = (0, \dots, 0, 1)^T$

$$\Rightarrow Q_k^* A Q_k = H_k$$

- Special case: when A is Hermitian, $Q_k^* A Q_k$ is Hermitian,

$\Rightarrow H_k$ is Hermitian $\Rightarrow H_k$ is tridiagonal

$\alpha_j \equiv h_{jj}, \quad \beta_j \equiv h_{j+1,j}$

$$H_k := T_k = \begin{bmatrix} \alpha_1 & \beta_1 & & & \\ \beta_1 & \alpha_2 & \beta_2 & & \\ & \ddots & \ddots & \ddots & \\ & & \beta_{k-2} & \alpha_{k-1} & \beta_{k-1} \\ & & & \beta_{k-1} & \alpha_k \end{bmatrix} \in \mathbb{C}^{k \times k}$$

Implementation: (Lanczos)

Given $x_0 \in \mathbb{C}^n$, $k \geq 1$,

$q_1 \leftarrow x_0 / \|x_0\|_2$, $\beta_0 = 0$, $q_0 = 0$

For $i = 1, \dots, k$

$w_i \leftarrow A q_i$

$\alpha_i = q_i^* w_i$

$w_i \leftarrow w_i - \beta_{i-1} q_{i-1} - \alpha_i q_i$

$\beta_i = \|w_i\|_2$ if $\beta_i = 0$, break

$q_{i+1} = w_i / \beta_i$

end

- Cost $\approx$ ($9n$ Flops + cost for Aq) $\times k$
- Storage $\approx kn$ for Q_n

- What exactly are $\tilde{\lambda}$ and $\tilde{v} = Q_k z$?

Let $\tilde{\lambda}$, $\tilde{v} = Q_k z$ be Ritz pair, i.e. $H_k z = \tilde{\lambda} z$, $\|z\|_2 = 1$

$$\Rightarrow \quad A\tilde{v} = \tilde{\lambda}\tilde{v} + h_{k+1,k}(e_k^* z)\, q_{k+1}$$

$$\Rightarrow \quad (A+E)\tilde{v} = \tilde{\lambda}\tilde{v}, \quad \|\tilde{v}\|_2 = 1 \qquad (*)$$

$$E = -h_{k+1,k}(e_k^* z)\, q_{k+1}\tilde{v}^*, \quad \|E\|_2 = |h_{k+1,k}|\,|z_k|, \quad z = (z_1, \cdots, z_k)^T$$

The Ritz pairs are eigenpairs of a perturbed eigenvalue problem

Recall the following theorem from Lecture 11:

<u>Thm</u> (Bauer – Fike)

If μ is an eig. val of $A+E$ and

$$V^{-1}AV = D = \mathrm{diag}(\lambda_1, \cdots, \lambda_n) \qquad (K_p(V) = \|V\|_p \|V^{-1}\|_p)$$

then $\min_{1\le i\le n} |\lambda_i - \mu| \le K_p(V)\,\|E\|_p \quad \forall\, p \in [1, +\infty)$ ☒

Apply Bauer – Fike to $(*)$, we have

Arnoldi: $\min_{1\le i\le n} |\lambda_i - \tilde{\lambda}| \le K_p(V)\,|h_{k+1,k}|\,|z_k|$

Lanczos: $\min_{1\le i\le n} |\lambda_i - \tilde{\lambda}| \le |\beta_k|\,|z_k|$

The approximation quality of H_k's eigenvalues as a function of k?

In the Hermitian setting, this can be made a little bit more precise. Let $\lambda_1 \geq \cdots \lambda_n$ be eig. val. of A.
$v_1, \cdots, v_n$ be eig. vec. of A.

<u>Thm</u> (Kaniel-Paige-Saad) Let $\|x_0\|_2 = 1$ in Krylov.
Suppose the Lanczos iteration are performed and T_k is obtained as a tridiagonal matrix. If $\theta_1 = \lambda_1(T_k)$, then

$$\lambda_1 \geq \theta_1 \geq \lambda_1 - (\lambda_1 - \lambda_n)\left(\frac{\tan(\phi_1)}{C_{k-1}(1+2\rho_1)}\right)^2$$

where $\cos(\phi_1) = x_0^* v_1$,

$$\rho_1 = \frac{\lambda_1 - \lambda_2}{\lambda_2 - \lambda_n}$$

properties of chebyshev: $|p(x)| \leq 1 \ \forall x \in [-1,1]$ but $|p(x)|$ grows rapidly outside $[-1,1]$

and $C_{k-1}(x)$ is the Chebyshev polynomial of degree $k-1$.

$$C_k(x) = 2x C_{k-1}(x) - C_{k-2}(x), \quad C_0 = 1, \quad C_1 = x$$

Pf: We use the following characterization of the dominant eigenvalue of an Hermitian matrix

$$\theta_1 = \max_{z \neq 0} \frac{z^* T_k z}{z^* z} \longleftarrow \text{Rayleigh quotient}$$

Use $T_k = Q_k^* A Q_k$, we have

$$\theta_1 = \max_{z \neq 0} \frac{(Q_k z)^* A (Q_k z)}{(Q_k z)^* (Q_k z)}$$

$$= \max_{w \in K_k(A, x_0)} \frac{w^* A w}{w^* w}$$

$$= \max_{p \in \mathcal{P}_{k-1}(\mathbb{C})} \frac{x_0^* p(A) A p(A) x_0}{x_0^* p(A)^2 x_0}$$

maximize over polynomial degree $\leq k-1$ $\longrightarrow$

Now let $x_0 = d_1 v_1 + \cdots + d_n v_n$. where $d_i = x_0^* v_i$

$$\frac{x_0^* p(A) A\, p(A) x_0}{x_0^* p(A)^2 x_0} = \frac{\sum_{i=1}^{n} d_i^2\, p(\lambda_i)^2 \lambda_i}{\sum_{i=1}^{n} d_i^2\, p(\lambda_i)^2} \qquad (**)$$

$$\geq \frac{\lambda_1 d_1^2 p(\lambda_1)^2 + \lambda_n \sum_{i=2}^{n} d_i^2 p(\lambda_i)^2}{d_1^2 p(\lambda_1)^2 + \sum_{i=2}^{n} d_i^2 p(\lambda_i)^2}$$

$$= \lambda_1 - (\lambda_1 - \lambda_n) \frac{\sum_{i=2}^{n} d_i^2 p(\lambda_i)^2}{d_1^2 p(\lambda_1)^2 + \sum_{i=2}^{n} d_i^2 p(\lambda_i)^2}$$

$$\geq \lambda_1 - (\lambda_1 - \lambda_n) \frac{\sum_{i=2}^{n} d_i^2 p(\lambda_i)^2}{d_1^2 p(\lambda_1)^2}$$

Now we pick

$$p(x) = C_{k-1}\left(-1 + 2\frac{x - \lambda_n}{\lambda_2 - \lambda_n}\right)$$

want $|p(\lambda_i)|$ small $(i \geq 2)$
but $|p(\lambda_1)|$ large
so rescale $[\lambda_n, \lambda_2]$ to $[-1, 1]$

By the construction of Chebyshev, $|C_k(x)| \leq 1$, $\forall x \in [-1, 1]$

Thus $|p(\lambda_i)| \leq 1$, $\forall\, i = 2, 3, \ldots, n$

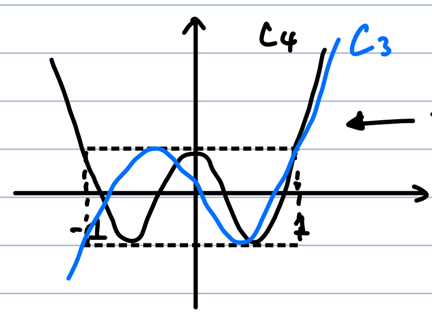
and $p(\lambda_1) = C_{k-1}(1 + 2\rho_1)$

Thus $\sum_{i=2}^{n} d_i^2 p(\lambda_i)^2 \leq \sum_{i=2}^{n} d_i^2 = \sum_{i=2}^{n} d_i^2 - d_1^2 = 1 - d_1^2$

so we have

$$\theta_1 \geq \lambda_1 - (\lambda_1 - \lambda_n) \frac{1 - d_1^2}{d_1^2} \frac{1}{C_{k-1}(1 + 2\rho_1)}$$

Chebyshev polynomial



← rapid growth outside $[-1, 1]$

∎

- Compare with power iteration:

  In power iteration, at the $k-1^{th}$ stage,

$$x_{k-1} = A^{k-1} x_0 = \sum_{i=1}^{n} d_i \lambda_i^{k-1} v_i$$

and eig. val. estimator (Rayleigh quotient of x_{k-1})

$$\gamma_1 = \frac{x_{k-1}^* A x_{k-1}}{x_{k-1}^* x_{k-1}}$$

corresponds to setting $p(x) = x^{k-1}$ in $(**)$

thus

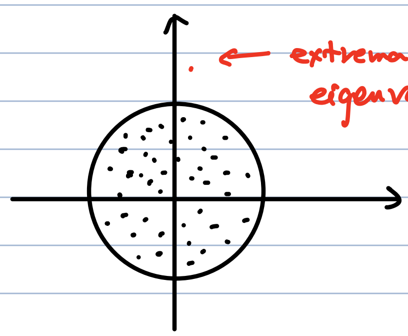
$$\lambda_1 \geq \gamma_1 \geq \lambda_1 - (\lambda_1 - \lambda_n) \tan(\phi_1)^2 \left(\frac{\lambda_2}{\lambda_1}\right)^{2(k-1)}$$

Usually. $C_{k-1}\left(1 + 2\frac{\lambda_1 - \lambda_n}{\lambda_2 - \lambda_n}\right) >> \left(\frac{\lambda_1}{\lambda_1}\right)^{2(k-1)}$

- Remark: The idea can be applied to estimate interior eigenvalues, but the bound is less satisfatory.

It is observed that Lanczos tends to approximate eigenvalues near the edges of spectrum with exp. rate in k.

Remark: For the non-Hermitian case, exponetial convergence of Ritz values to "extremal" eigenvalues can also be proved. But convergence to "interior" eigenvalues can be much slower!



extremal eigenvalues ("separable" by a circle/ellipsoid from all other eig. values)

- Practical Arnoldi / Lanczos

Issues: 1) If k is large, the computation of q_{k+1} involves $O(kn)$ flops in Arnoldi, and memory can quickly run out in both Arnoldi / Lanczos.

hope to : limit k in practice

but still want result be close to eigenvalues!

2) Arnoldi / Lanczos without reorthogonalization suffer from rounding error (orthogonality of $\hat{Q}_k$ is lost) $\Rightarrow$ many issues. such as ghost eigenvalues, i.e. single λ would repeat many times!

combine with other techniques. filtering (shifting).

Solution: Restarting $\Longrightarrow$ implicitly restarting Arnoldi (ARPACK.jl)

- After $k+p$ steps, throw out most of Q_{k+p}, keep only k "best" vectors so far, restart Arnoldi / Lanczos on step k with Q_k (reorthogonalized)

  How to choose "best" vectors?

  — Want to satisfy the form of Arnoldi decomposition

- Naive example: suppose we only want the dominant $|\lambda|$, just keep $k=1$ vector. if $p=0$, $\approx$ power iteration