Iterative Methods

Part 2: More about the Arnoldi Iteration

Recap

Power Method

$x_0 =$ arbitrary

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

$$\hat{x}_k = A x_{k-1}$$
$$x_k = \hat{x}_k / \|\hat{x}_k\|$$
$$\rho_k = x_k^* A x_k$$

$$\left.\begin{array}{l} \rho_k \to \lambda_1 \\ x_k \to v_1 \end{array}\right\} \text{largest mod. eigenvalue of } A$$

when $v^T x_0 \neq 0$

and $|\lambda_1| > |\lambda_2|$

Krylov Subspace Method

("Naive" - DO NOT USE!)

$x_0 =$ arbitrary,

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

$$x_k = A x_{k-1}$$
$$QR = \begin{bmatrix} | & & | & | \\ x_1 & \cdots & x_{k-1} & x_k \\ | & & | & | \end{bmatrix}$$
$$A_Q = Q^* A Q$$

solve $A_Q w = \rho w$

$\to$ Q may not be an accurate ONB b/c $K_n = [x_1 \cdots x_n]$ is typically very ill-cond.

How to build Q?

Arnoldi Iteration

$$x \to Ax \to A^2 x \to \cdots \to A^{n-1} x \qquad \text{Naive}$$

$$\begin{array}{ccccccccc} x & & Aq_1 & & Aq_2 & & & & Aq_{n-1} \\ \downarrow & \nearrow & \downarrow & \nearrow & \downarrow & \nearrow & \cdots & \nearrow & \downarrow \\ q_1 & & q_2 & & q_3 & & & & q_n \end{array} \qquad \text{Arnoldi (MGS)}$$

$$\begin{bmatrix} | & | & | & & | \\ x & Ax & A^2x & \cdots & A^{n-1}x \\ | & | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & | & & | \\ q_1 & q_2 & q_3 & \cdots & q_n \\ | & | & | & & | \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & \cdots \\ & r_{22} & \\ & & \ddots \end{bmatrix}$$

$$K_n \qquad = \qquad Q \qquad R$$

Arnoldi produces a QR factorization of K_n ✓

Arnoldi <u>also</u> produces the ONB that makes A upper Hessenberg

$$\hat{H}_n = \begin{bmatrix} H_n \\ h e_n^T \end{bmatrix}$$

$$A \overbrace{\begin{bmatrix} | & | & & | \\ q_1 & q_2 & \cdots & q_n \\ | & | & & | \end{bmatrix}}^{Q_n} = \Bigg[\overbrace{\begin{matrix} | & | & & | \\ q_1 & q_2 & \cdots & q_n \\ | & | & & | \end{matrix}}^{Q_n} \; \underbrace{\overbrace{\begin{matrix} | \\ q_{n+1} \\ | \end{matrix}}^{q_{n+1}}}_{\text{extra column}} \Bigg] \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1n} \\ h_{21} & h_{22} & & \vdots \\ & h_{32} & \ddots & \\ & & \ddots & h_{nn} \\ & & & h_{n+1,n} \end{bmatrix}$$

extra row

$$= Q_n H_n + q_{n+1} (h_{n+1,n} e_n^T)$$

$$\boxed{A_{Q_n} = Q_n^T A Q_n = H_n}$$

RR projection onto $K_n(A, x)$ is simply H_n

Here is the Arnoldi pseudocode that builds Q_n and H_n from A and x.

x = arbitrary, $q_1 = x/\|x\|$
for $k = 1, 2, 3, \dots, n$
    $v = Aq_k$
    for $j = 1, \dots, k$
        $h_{jk} = q_j^* v$
        $v = v - h_{jk} q_j$
    $h_{k+1,k} = \|v\|$
    $q_{k+1} = v / h_{k+1,k}$

Arnoldi Breakdown

← note that $h_{k+1,k} = 0$ means that Q_k is an invariant subspace of A and we can "deflate" a $k \times k$ EVP (homework). Also, it means $\mathrm{rank}(K_n) < n$

Convergence

Rule of Thumb: Arnoldi approximates well-isolated eigenvalues, typically near the edges of the spectrum ("extremal eigenvalues") with exponential accuracy.

The key is a connection to polynomial approx.

If $u \in K_n(A,b)$, then

$$u = c_1 x + c_2 Ax + \cdots + c_{n-1} A^{n-1} x = q(A)x$$

where $q(z) = c_0 + c_1 z + \cdots + c_{n-1} z^{n-1}$
$= c_{n-1}(z - r_1) \cdots (z - r_{n-1})$

The Arnoldi iteration generates a very special polynomial, which solves

P1 Find $p_n \in$ {Monic Polynomials of deg n}

s.t. $\|p_n(A)x\| = \text{minimum}$

<u>Thm</u> IF Arnoldi doesn't break down ($\text{rank}(K_n) = n$), then the unique soln. of P1 is determined by the characteristic polynomial of H_n, i.e.,

$$p(z) = (z - \theta_1)(z - \theta_2) \cdots (z - \theta_n)$$

where $H w_i = \theta_i w_i$ for $i = 1, \ldots, n$.

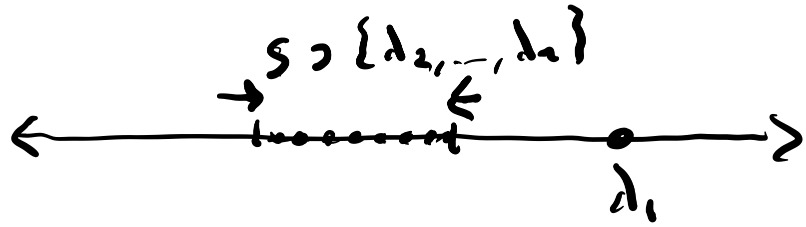
<u>Pf</u>/ L.N.T. Lecture 34

This then provides some intuition about why certain Ritz values tend to approximate isolated eigenvalues of A.

If
$$A = \underbrace{\begin{bmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{bmatrix}}_{V} \underbrace{\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}}_{\Lambda} \underbrace{\begin{bmatrix} - & w_1^* & - \\ & \vdots & \\ - & w_n^* & - \end{bmatrix}}_{V^{-1}}$$

Then
$$p_n(A) = \underbrace{\begin{bmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{bmatrix}}_{V} \underbrace{\begin{bmatrix} p_n(\lambda_1) & & \\ & \ddots & \\ & & p_n(\lambda_n) \end{bmatrix}}_{p_n(\Lambda)} \underbrace{\begin{bmatrix} - & w_1^* & - \\ & \vdots & \\ - & w_n^* & - \end{bmatrix}}_{V^{-1}}$$

Ideally, $p(z)$ would be zero on each $\lambda_1, \ldots, \lambda_m$, but $p_n(z)$ only has n zeros, $\theta_1, \ldots, \theta_n$. How should it place them to solve PI?



Here, $p_n(\lambda)$ should be small on $S \cup \lambda_1$,
so it should put a zero, θ_1, near λ_1
to make $p_n(\lambda_1)$ small.

As n increases, $p_n(\lambda)$ has more
zeros, which it can use to zero out
other (less) isolated eigenvalues.

A Quantitative analysis using potential
theory makes this precise.