Lecture 22

The Krylov subspace and a nice proprety in terms of the solution of Ax = b.
—> Cayley Hamilton Theorem.

The problem with a simple basis for the Krylov subspace.
—> The simple basis for the subspace is highly ill-conditioned.

Building an orthogonal basis for the subspace with the Arnoldi and Lanczos processes.
—> This will look a lot like Gram-Schmidt.

HW 9 - Q4

Recall the Krylov subspace

$\mathbb{K}_k(A,\underline{b}) = \text{span}(\underline{b}, A\underline{b}, \ldots, A^k\underline{b})$

Why are we looking at this?

- x_k in Neumann/Richardson lies in $\mathbb{K}_k$
- it shows up a lot!
- there are awesome relationships to polynomials that will show up soon.

For a "quick" property of $\mathbb{K}_k$

Theorem:

Suppose that

$A^k\underline{b} \in \mathbb{K}_{k-1}(A,\underline{b})$

Then $\underline{x} = A^{-1}\underline{b}$ must be in $\mathbb{K}_{k-1}(A,\underline{b})$ too.

<=>

If $X = [\underline{b}, A\underline{b}, \ldots A^k\underline{b}]$ is singular. Then

$\underline{x} = X\underline{y}$ for some $\underline{y}$.

Let

$X = [\underline{b}, A\underline{b}, \ldots, A^{k-1}\underline{b}]$

be a basis for $\mathbb{K}_{k-1}(A,\underline{b})$

We know

$\underline{z} = A^k\underline{b} \in \mathbb{K}_{k-1}(A,\underline{b})$

<=>

$\underline{z} = X\underline{y}$ for some $\underline{y}$.

This means

$A^{k+1}\underline{b} \in \mathbb{K}_{k-1}(A,\underline{b})$ too

and also $A^{k+j}\underline{b}$ for $j = 1, 2, \ldots$

Why is

$A^{k+1}\underline{b} \in \mathbb{K}_{k-1}(A,\underline{b})$?

$A \cdot A^k\underline{b} = A \cdot \sum_{j=0}^{k-1} y_j A^j\underline{b}$

$= \sum_{j=0}^{k-1} y_j A^{j+1}\underline{b}$

$= \underbrace{y_0 A\underline{b} + y_1 A^2\underline{b} + \ldots + y_{k-2}A^{k-1}\underline{b}}_{\in \mathbb{K}_{k-1}(A,\underline{b})} + \underbrace{y_{k-1}A^k\underline{b}}$

$y_{k-1}A^k\underline{b} = y_{k-1} \cdot \sum_{j=0}^{k-1} y_j A^j\underline{b}$

$y_0 \to y_0 y_{k-1}$

$y_1 \to y_0 + y_1 y_{k-1}$

$\vdots$

$y_{k-1} \to y_{k-2} + y_{k-1}^2$.

Good practice - figure out coeffs for $A^{k+2}\underline{b}$

The Cayley-Hamilton Theorem.

Let A be a square matrix. There is a degree n poly $q(x)$ s.t.

$q(A) = 0$.

$A = 5$.

$q(x) = (x-5)$ <=>

$q(A) = (A-5) = 0$.

$A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$

$q(x) = (x-1)(x-2)$

$q(A) = (A-I)(A-2I)$

$= \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}$

$= 0$.

In general. If A has n distinct λ vals. Then

$q(x) = \prod_{i=1}^{n} (x - \lambda_i)$.

$q(A) = \prod_{i=1}^{n} (A - \lambda_i I)$

We will use $q(x)$ to find $p(x)$ st. $p(A) = A^{-1}$.

Note, if we show $p(x)$ exists, then we have proven our theorem.

$x = A^{-1}\underline{b} = p(A)\underline{b}$

$= \sum_{j=0}^{n-1} c_j A^j\underline{b}$

$= c_0\underline{b} + c_1 A\underline{b} + \ldots + c_{n-1}A^{n-1}\underline{b} =$

$\underbrace{c_0\underline{b} + \ldots c_{k-1}A^{k-1}\underline{b}}_{\substack{\in \mathbb{K}_{k-1}(A,\underline{b}) \\ \text{(by defn)}}} + \underbrace{c_k A^k\underline{b} + \ldots c_{n-1}A^{n-1}\underline{b}}_{\substack{\in \mathbb{K}_k(A,\underline{b}) \\ \text{by our result}}}$

<=>

$x \in \mathbb{K}_{k-1}(A,\underline{b})$.

We have

$q(A) = 0$.

$A^{-1}q(A) = 0$.

$= A^{-1}[c_0 I + c_1 A + c_2 A^2 + \ldots c_n A^n] = 0$

$= c_0 A^{-1} + c_1 I + c_2 A + \ldots c_n A^{n-1} = 0$

$A^{-1} = \frac{-1}{c_0}[c_1 I + c_2 A + \ldots c_n A^{n-1}]$

$= \frac{-c_1}{c_0} I - \frac{c_2}{c_0} A - \ldots - \frac{c_n}{c_0} A^{n-1}$

$= p(A)$

$\boxtimes$

Is $c_0 = 0$?

$q(x) = \prod (x - \lambda_i)$

$= c_0 + c_1 x + \ldots + c_n x^n$

$q(0) = c_0$?

$q(0) = \prod_{i=1}^{n} (-\lambda_i)$

$q(0) \neq 0$ because

$\lambda_i \neq 0$ for a non-sing A.

Demo!

Recap!

If X_k is the matrix

$[\underline{b}, A\underline{b}, \ldots, A^k\underline{b}]$

then $\kappa(X_k) \to \infty$

(in general)

Why? Columns of X_k are the (scaled) iterates of the power method.

Those converge! So columns of X_k converge to some limit -> X_k will get ill-conditioned

The Lanczos + Arnoldi Process are a way to build a better conditioned basis for $\mathbb{K}_k(A,b)$.

At a useful level, Arnoldi + Lanczos compute a matrix decomp.

Arnoldi decomp

A: n x n, Given $\underline{b}$, find

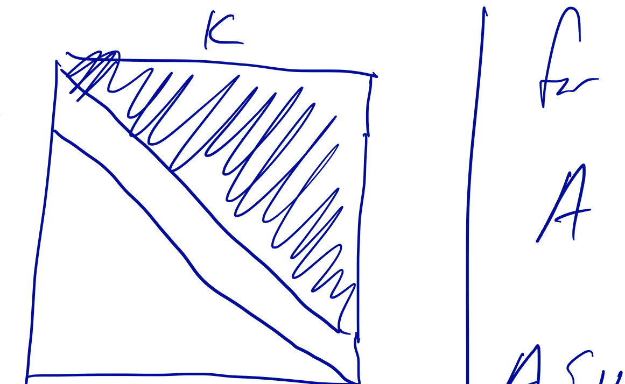
$AV_k = V_{k+1}H_{k+1,k}$

V_k = n by k, orthog

$H_{k+1,k}$ is a k+1 by k upper Hessenberg matrix.

$V_k = [\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_k]$ and $\underline{v}_1 = \rho\underline{b}$

$H_{k+1,k} =$



upper-tri + one more diag.

for $k = 1$

$AV_1 = V_2 H_{2,1}$

$A[\underline{v}_1] = [\underline{v}_1\ \underline{v}_2]\begin{bmatrix} H_{11} \\ H_{21} \end{bmatrix}$

Set $\underline{v}_1 = \underline{b}/\|\underline{b}\|$ so $\underline{v}_1^T\underline{v}_1 = 1$

$A\underline{v}_1 = H_{11}\underline{v}_1 + H_{21}\underline{v}_2$

$\underline{v}_2^T\underline{v}_1 = 0$.

So $H_{11} = \underline{v}_1^T A\underline{v}_1$

So $H_{21}\underline{v}_2 = A\underline{v}_1 - H_{11}\underline{v}_1$.

Then $H_{21} = \|A\underline{v}_1 - H_{11}\underline{v}_1\|_2$

$\underline{v}_2 = \frac{A\underline{v}_1 - H_{11}\underline{v}_1}{H_{21}}$

Check

$H_{21}\underline{v}_1^T\underline{v}_2 = \underline{v}_1^T A\underline{v}_1 - H_{11}\underbrace{\underline{v}_1^T\underline{v}_1}_{=1}$

$\underline{v}_1^T A\underline{v}_1$

$= 0$

$A[\underline{v}_1\ \underline{v}_2] = [\underline{v}_1\ \underline{v}_2\ \underline{v}_3]\begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \\ 0 & H_{32} \end{bmatrix}$

Quiz.

What are $\underline{v}_3$, H_{12}, H_{22}, H_{32}?

Let $\underline{y} = A\underline{v}_2 = H_{12}\underline{v}_1 + H_{22}\underline{v}_2 + H_{32}\underline{v}_3$

$\underline{v}_1^T\underline{y} = \underline{v}_1^T H_{12}\underline{v}_1 + 0 + 0$

$= H_{12}$

$\underline{v}_2^T\underline{y} = \quad = 0 + H_{22}\underline{v}_2^T\underline{v}_2 + 0$

$= H_{22}$

~~$\underline{v}_3^T\underline{y}$~~ $H_{32} = \|\underline{y} - H_{12}\underline{v}_1 - H_{22}\underline{v}_2\|$

$\underline{v}_3 = (\underline{y} - H_{12}\underline{v}_1 - H_{22}\underline{v}_2)/H_{32}$

```
Alg for Arnoldi
Input Square $A$, $\underline{b}$
Let $\underline{v}_1 = \underline{b} / \|\underline{b}\|$

for $i = 1$ to $k$
    $\underline{y} = A \underline{v}_i$
    for $j = 1$ to $k$
        $H_{j,i} = \underline{v}_j^T \underline{y}$
        $\underline{y} = y - H_{j,i} \underline{v}_j$
    End
    $H_{i+1,i} = \|y\|$
    $\underline{v}_{i+1} = \underline{y} / H_{i+1,i}$
End.
```

CS 515 - Purdue University - Computer Science - David F. Gleich - 2019