Advanced Numerical Methods	20111219

Lecture Notes 0: Krylov subspace methods

Lecturer: Che-Rung Lee *Scribe: Tzu-Wei Huang*

1 Polynomial of matrix A

1.1 Characteristic polynomial of matrix A

$$P_A(x) = |A - xI|.$$

example: given

$$A = \begin{bmatrix} 1 & 4 \\ 3 & 2 \end{bmatrix},$$

$$P_A(x) = |A - xI| = \begin{vmatrix} 1-x & 4 \\ 3 & 2-x \end{vmatrix} = x^2 - 3x - 10.$$

Let $P_A(x) = 0$, then $x_1 = -5, x_2 = 2$, which are the eigenvalues of matrix A. *This polynomial encodes several important properties of the matrix, most notably its eigenvalues, its determinant and its trace. [wikipedia]*

1.2 Cayley–Hamilton theorem

Let $P_A(x)$ be the Characteristic polynomial of matrix A, then

$$P_A(A) = 0.$$

try it:

$$\begin{aligned} P_A(A) &= A^2 - 3A - 10I \\ &= \begin{bmatrix} 13 & 12 \\ 9 & 16 \end{bmatrix} - 3 \begin{bmatrix} 1 & 4 \\ 3 & 2 \end{bmatrix} - 10 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

A *bogus* proof: $p(A) = det(AI - A) = det(A - A) = 0$.
formal proof:

$$\begin{aligned}
P_A(x) &= \sum_{i=1}^{n} a_i x^i \\
\Rightarrow P_A(A) &= \sum_{i=1}^{n} a_i A^i \\
&= \sum_{i=1}^{n} a_i (X\Lambda X^{-1})^i \\
&= \sum_{i=1}^{n} a_i X \Lambda^i X^{-1} \\
&= X(\sum_{i=1}^{n} a_i \Lambda^i) X^{-1} \\
&= X \begin{bmatrix} \sum_{i=1}^{n} a_i \lambda_1^i & & & & \\ & \sum_{i=1}^{n} a_i \lambda_2^i & & \text{\huge 0} & \\ & \text{\huge 0} & \sum_{i=1}^{n} a_i \lambda_3^i & & \\ & & & \ddots & \\ & & & & \sum_{i=1}^{n} a_i \lambda_n^i \end{bmatrix} X^{-1} \\
&= X \begin{bmatrix} P_A(\lambda_1) & & & & \\ & P_A(\lambda_2) & & \text{\huge 0} & \\ & \text{\huge 0} & P_A(\lambda_3) & & \\ & & & \ddots & \\ & & & & P_A(\lambda_n) \end{bmatrix} X^{-1} \\
&= 0,
\end{aligned}$$

Since $\lambda_1, \lambda_2, ..., \lambda_n$, are root of $P_A(x) = 0$.

1.3 express a matrix by its characteristic polynomial

We can use polynomials of A to express any functions of A if A is diagonalizable.

$$f(x) = x^{-1} = a_1 x^1 + ... + a_i x^i.$$
$$A^{-1} = f(A).$$

example: given

$$A = \begin{bmatrix} 1 & 4 \\ 3 & 2 \end{bmatrix}$$

the eigenvalues of A are $\lambda_1 = 5, \lambda_2 = -2$.

$$P(A) = A^{-1} = x^2 + bx + c$$
$$P(5) = f(5) = 1/5 = (1/5)^2 + (1/5)b + c$$

$$P(-2) = f(-2) = -1/2 = (-1/2)^2 + (-1/2)b + c$$

$$\Rightarrow b = -2.9, c = -10.3$$

$$\Rightarrow P(x) = x^2 - 2.9x - 10.3$$

$$P(A) = A^2 - 2.9A - 10.3I = A^{-1}$$

.(check by hand)

We can replace the operation on matrix A by with a corresponding polynomial, the coeffients of the polynomial are computed similarly.

1.4 interpolation and approximation of polynomial

- Interpolation

$$P(x) = x^n + a_{n-1}x^{n-1} + ... + a_1 x + a_0$$

find $a_{n-1}, ..., a_0$, such that for all λ_i, $P(\lambda_i) = f(\lambda_i)$.

$$\begin{bmatrix} \lambda_1^{n-1} & \lambda_1^{n-2} & \dots & \lambda_1 & 1 \\ \lambda_2^{n-1} & \lambda_2^{n-2} & \dots & \lambda_2 & 1 \\ \vdots & & & & \vdots \\ \lambda_n^{n-1} & \lambda_n^{n-2} & \dots & \lambda_n & 1 \end{bmatrix} * \begin{bmatrix} a_{n-1} \\ a_{n-2} \\ \vdots \\ a_1 \\ a_0 \end{bmatrix} = \begin{bmatrix} f(\lambda_1) - \lambda_1^n \\ f(\lambda_2) - \lambda_2^n \\ \vdots \\ f(\lambda_n) - \lambda_n^n \end{bmatrix}$$

That is the form of $Ax = b$, where $A \in \mathbb{R}^{n \times n}; x, b \in \mathbb{R}^{n \times 1}$.

- Approximation use low-dimensional polynomial to approximate high-dimensional poltnomial.

$$P(x) = a_k x^k + a_{k-1}x^{k-1} + ... + a_1 x + a_0$$

find $a_{n-1}, ..., a_0$, such that for all λ_i, $P(\lambda_i) \approx f(\lambda_i)$.

$$\begin{bmatrix} \lambda_1^{k} & \lambda_1^{k-2} & \dots & \lambda_1 & 1 \\ \lambda_2^{k} & \lambda_2^{k-2} & \dots & \lambda_2 & 1 \\ \vdots & & & & \vdots \\ \lambda_n^{k} & \lambda_n^{k-2} & \dots & \lambda_n & 1 \end{bmatrix} * \begin{bmatrix} a_{k} \\ a_{k-1} \\ \vdots \\ a_1 \\ a_0 \end{bmatrix} \approx \begin{bmatrix} f(\lambda_1) \\ f(\lambda_2) \\ \vdots \\ f(\lambda_n) \end{bmatrix}$$

That is the form of $Ax \approx b$, where $A \in \mathbb{R}^{n \times k}; x \in \mathbb{R}^{k \times 1}; b \in \mathbb{R}^{n \times 1}$, and $k < n$. We can treat it as a least squares problem and solve x for the polynomial coefficients.

1.5 Krylov subspace

An m-dimentional Krylov subspace of A is defined as follows:

$$\mathcal{K}_m(A, \vec{q}_1) = span\{\vec{q}_1, A\vec{q}_1, A^2\vec{q}_1, ..., A^{m-1}\vec{q}_1\}.$$

Different $\vec{q}_1$ results in different subspace $\mathcal{K}_m(A\vec{q}_1)$.

Since the above operation is numerically unstable, we use Arnoldi mehtod to generate an equvalent subspace. The basic idea of Arnoldi method is similar to Gram-Schmidt process.

for $i = 1...m$ **do**
- $\vec{y_i} = A\vec{q_i}$
- orthogonalize $\vec{y_i}$ against current subspace Q_i such that
- $\vec{y_i} = Q_i\vec{h_i} + \vec{z}$
- **if** $||\vec{z}||$ *euqals* 0 **then**
  - reach invariant subspace
  - *breaktheloop*
- **end**
- $\vec{q}_{i+1} = \vec{z}/||\vec{z}||$
- $Q_{i+1} = [Q_i\ q_{i+1}]$

end

Then $span\{\vec{q}_1, \vec{q}_2, \vec{q}_3, ..., \vec{q}_m\} = span\{\vec{q}_1, A\vec{q}_1, A^2\vec{q}_1, ..., A^{m-1}\vec{q}_1\}$.