B24 Aug 11 Lec 1 Notes

Proof: Cayley-Hamilton theorem

Case 1: Assume A is a diagonal matrix

$$A = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

then $P_A(\lambda) = (\lambda_1 - \lambda) \cdots (\lambda_n - \lambda)$, so

$$P_A(A) = \underbrace{(\lambda_1 I - A)}_{\substack{\text{diagonal} \\ \text{with } 0 \text{ in} \\ 1,1 \text{ entry}}} \cdots \underbrace{(\lambda_n I - A)}_{\substack{\text{diagonal} \\ \text{with } 0 \text{ in} \\ n,n \text{ entry}}} = 0$$

Thus the composition of the diagonal matrices would be 0.

Case 2: Assume A is diagonalizable, i.e.

$$A = QDQ^{-1}$$

for D diagonal.

Then $P_A(\lambda) := \det(A - \lambda I) = \det(D - \lambda I) = P_D(\lambda)$

And so

$$\begin{aligned} P_A(A) = P_D(A) &= (\lambda_1 I - A) \cdots (\lambda_n I - A) \\ &= (\lambda_1 I - QDQ^{-1}) \cdots (\lambda_n I - QDQ^{-1}) \\ &= Q(\lambda_1 I - D)\cancel{Q^{-1}} \cdots \cancel{Q}(\lambda_n I - D)Q^{-1} \\ &= Q(\lambda_1 I - D) \cdots (\lambda_n I - D)Q^{-1} \\ &= Q0Q^{-1} \qquad \text{from case 1} \\ &= 0 \end{aligned}$$

Case 3: A is upper triangular

If the diagonal entries of A are distinct, then A is diagonalizable, and so case 2 applies. If the diagonal entries of A are not distinct, we argue as follows.

For each $k \in \mathbb{N}$, let A_k be a matrix such that each entry of A_k converges to the corresponding entry of A as $k \to \infty$, and A_k has distinct diagonal entries.

e.g. $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, $A_1 = \begin{bmatrix} .9 & 1 \\ 0 & 1.1 \end{bmatrix}$, $A_2 = \begin{bmatrix} .99 & 1 \\ 0 & 1.01 \end{bmatrix}$

So for each k, A_k is diagonalizable, so $P_{A_k}(A_k) = 0$. Moreover,

$$P_{A_k}(\lambda) \xrightarrow{k \to \infty} P_A(\lambda), \text{ and so}$$

$$\underbrace{P_{A_k}(A_k)}_{\text{constant sequence } 0} \xrightarrow{k \to \infty} P_A(A), \text{ and so } P_A(A) = 0$$

Proof (Continued...):

Case 4:

For any matrix A, there exists upper triangular D and invertible Q so that:

$$A = QDQ^{-1}$$

(Q, D are not necessarily real even if A is)

hence the reasoning of case 2 applies:

$$P_A(\lambda) = P_D(\lambda) = c(\lambda - \lambda_1)\cdots(\lambda - \lambda_n)$$

So

$$\begin{aligned} P_A(A) = P_D(A) &= c(A - \lambda_1 I)\cdots(A - \lambda_n I) \\ &= c(QDQ^{-1} - \lambda_1 I)\cdots(QDQ^{-1} - \lambda_n I) \\ &= c\, Q(D - \lambda_1 I)\cancel{Q^{-1}}\cdots\cancel{Q}(D - \lambda_n I)Q^{-1} \\ &= c\, Q\underbrace{P_D(D)}_{=0 \text{ since } D \text{ is upper triangular}}Q^{-1} \\ &= c\, Q\, 0\, Q^{-1} = 0 \end{aligned}$$

$\boxtimes$

Definition:

$\sigma(A)$ is the set of eigenvalues of A.

Theorem: Spectral mapping theorem

Let A be a square matrix, and p a polynomial. Then

$$p(\sigma(A)) = \sigma(p(A))$$

e.g. $A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$, $p(z) = z^2 + 1$

$\sigma(A) = \{1, 3\}$ $\qquad$ $p(\sigma(A)) = \{1^2+1, 3^2+1\} = \{2, 10\}$

$$\begin{aligned} P(A) &= A^2 + I \\ &= \begin{bmatrix} 1 & 8 \\ 0 & 9 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 8 \\ 0 & 10 \end{bmatrix} \end{aligned}$$

$\therefore \sigma(p(A)) = p(\sigma(A))$

Proof:

Let $p(\lambda) = c_n \ell^n + \dots + c_1 \ell + c_0$

Case 1: $p(\sigma(A)) \subseteq \sigma(p(A))$

Let $\lambda \in \sigma(A)$. WTS $p(\lambda) \in \sigma(p(A))$

Let $x \neq 0$ s.t. $Ax = \lambda x$. Then:

$$\begin{aligned} p(A)x &= (c_n A^n + \dots + c_1 A + c_0 I)x \\ &= c_n \lambda^n x + \dots + c_1 \lambda x + c_0 x \\ &= (c_n \lambda^n + \dots + c_1 \lambda + c_0) x \\ &= p(\lambda) x \end{aligned}$$

So $p(\lambda) \in \sigma(p(A))$

Case 2. $p(\sigma(A)) \supseteq \sigma(p(A))$

Let $u \in \sigma(p(A))$.

Define $q(z) := p(z) - u$

Then $q(A) = p(A) - uI$ is not invertible, since an eigenvector for $p(A)$ is sent to 0 by $q(A)$.

Let $q(z) = c(z - z_1) \cdots (z - z_n)$

So $q(A) = c(A - z_1 I) \cdots (A - z_n I)$

one of the factors $A - z_j I$ must be non-invertible, i.e. z_j is an eigenvalue of A.

and $p(z_j) = q(z_j) + u = u$

i.e. $u = p(z_j) \in p(\sigma(A))$

$\boxtimes$