E2 212: Cayley-Hamilton Theorem

Bharath B. N.

2 Oct. 2012

1 Statement of the Cayley-Hamilton Theorem

Let $f(x) = \sum_{i=0}^{m} a_i x^i$ be an m^{th} degree polynomial in $\mathcal{C}$ with $a_i \in \mathcal{C}$. Using this one can define a polynomial in $A \in \mathcal{C}^{m \times m}$ as

$$f(A) := \sum_{i=0}^{m} a_i A^i, \tag{1}$$

where A^i is the i^{th} power of the matrix A with $A^0 := I_m$. For $\lambda \in \mathcal{C}$, the characteristic polynomial of the matrix A is defined as

$$f_A(\lambda) := \det(A - \lambda I). \tag{2}$$

Note that

$$f_A(\lambda) = \det(A - \lambda I) = \sum_{i=1}^{m} b_i \lambda^i, \tag{3}$$

where the coefficients b_i, $i = 1, 2, \ldots, m$ depend on the entries of the matrix A. We say that the matrix A is annihilated by the polynomial $f_A(\lambda)$ if $f_A(A) = 0$.

Exercise 1: Prove that $\det(A - \lambda I)$ is a continuous function in the entries of the matrix A.

Exercise 2: Let the matrix $A \in \mathcal{C}^{m \times m}$ be equivalent to a diagonal matrix. Then, prove that $f_A(A) = 0$.

Now, let us state the main theorem.

Theorem 1 *Any $m \times m$ matrix A is annihilated by its characteristic polynomial, i.e.,*

$$f_A(A) = 0. \tag{4}$$

2 Proof of the Cayley-Hamilton Theorem

First, let us look at the idea behind the proof.

Supposing that there exist matrices $A_1, A_2, \ldots$ such that

1. $f_{A_i}(A_i) = 0$ for all $i = 1, 2, 3$, etc.

2. $A = \lim_{n\to\infty} A_n$, i.e., $\|A - A_n\| \to 0$ as $n \to \infty$ for any norm $\| * \|$.

Now, let us look at the consequence of the above assumptions. Denote the set of all $m \times m$ matrices by $\mathcal{M}_m$. Since the characteristic polynomial $f : \mathcal{M}_n \to \mathcal{C}$ is a continuous function, we have

$$f_A(A) = \lim_{n\to\infty} f_{A_n}(A_n) = 0, \tag{5}$$

which gives the desired result! Note that the equality above follows from the fact that all norms are equivalent in a finite dimensional vector space. Thus, we need to construct a sequence of matrices that satisfy (1) and (2) above.

It is easy to see that property (1) above is satisfied if the matrices A_i, $i = 1, 2, 3, ...$ have distinct eigenvalues (why?). Property (2) is satisfied if A_i, $i = 1, 2, 3, ...$ become "sufficiently" close to A as i becomes large. The task of finding such a sequence of matrices become easy due to the Schur's triangularization theorem.

2.1 Complete Proof

Let $A \in \mathcal{C}^{m\times m}$. By Schur's triangularization theorem, we have,

$$A = UTU^H, \tag{6}$$

where U is a unitary matrix, T is an upper triangle matrix with the diagonal entries being the eigenvalues of A. Denote the (ij)-th entry of T by t_{ij}. Now, construct the sequence A_n, $n = 1, 2, 3, ...$ as follows:

$$A_n = UT_nU^H, \tag{7}$$

where t_{ij}^n (the $ij - th$ entry of T_n) is given by

$$t_{ij}^n = \begin{cases} t_{ij} & \text{if } i \neq j \\ t_{ii} + \frac{\epsilon_i}{n} & \text{if } i = j, \end{cases} \tag{8}$$

where $0 < \epsilon_i < \infty$, $i = 1, 2, \ldots, m$ are chosen such that the diagonal entries of T_n are distinct. Note that the matrices T_n and T differ only in the diagonal entries. Since the diagonal entries of T_n are the *distinct* eigenvalues of the matrix A_n, A_n is diagonalizable for all $n \geq 1$. Thus, each matrix A_n satisfies $f_{A_n}(A_n) = 0$. It is easy to see that for any norm $\| * \|$, we have

$$\begin{aligned} \|A - A_n\|^2 &= \|T - T_n\|^2 \\ &\leq K\|T - T_n\|_F^2 \text{ for some } 0 < K < \infty \text{ (by equivalence of norms)} \\ &\leq K\frac{1}{n^2}\sum_{i=1}^{m} \epsilon_i^2 \\ &\leq K\frac{m}{n^2}\max_{1\leq i\leq m} \epsilon_i^2 \\ &\to 0 \text{ as } n \to \infty. \end{aligned}$$

Using the above and the continuity of the characteristic polynomial, it implies that

$$f_A(A) = \lim f_{A_n}(A_n) = 0.$$

This completes the proof. $\square$