Volume 2, Issue 12, December 2012 ISSN: 2277 128X

International Journal of Advanced Research in Computer Science and Software Engineering

Research Paper

Available online at: www.ijarcsse.com

Alternate Proof of Cayley-Hamilton Theorem

Paramjeet Sangwan [1] , Dr. G.N. Verma [2]

Ph.D Scholar CMJ University, Shillong (Meghalaya) [1]

Director- Principal, Sri Sukhmani Institute of Engineering and Technology, Dera Bassi (Punjab) [2]

Abstract: **The main focus of the paper revolves around bringing new insight to the theorem through alternate derivations. It provides and analyzes proofs via Schur's Triangularization, a variation of topological proofs. Properties of upper- triangular matrices are used in proofs via Schur's Triangularization. The Topological proof uses continuity properties and matrix norms.**

***Keywords*: Cayley- Hamilton Theorem, topology, Matrices, Eigen values.**

I. INTRODUCTION

Defintion 1.1. If A is an n×n matrix ,then the characteristic polynomial of A is defined to be $P_A(x)= det(xI-A)$. This is a polynomial in x of degree n with leading term x^n . the constant term c_0 of a polynomial q(x) is interpreted as c_0 I in q(A).

Theorem 1.2 (Cayley – Hamilton Theorem). If A is an n×n matrix ,then $p_A(A)=0$, the zero matrix.

Theorem 1.3 If q≠0 is a quaternion of the form q-= a+bi+cj+dk with a,b,c,d, being real , then $q^2 - 2aq + (a^2 +b^2+c^2+d^2)=0$

$$q^{-1} = \frac{\bar{q}}{|q|}$$

$$= \frac{a-bi-cj-dk}{a^2+b^2+c^2+d^2}$$

$$= \frac{2a}{a^2+b^2+c^2+d^2} - \frac{a+bi+cj+dk}{a^2+b^2+c^2+d^2}$$

$$= \frac{1}{a^2+b^2+c^2+d^2}(2a-q)$$

$$\Rightarrow a^2+b^2+c^2+d^2 = 2aq - q2$$

$$\Rightarrow q2 - 2aq + (a^2+b^2+c^2+d^2) = 0$$

If one represents a quaternion q= a+bi+cj+dk as a matrix,

$$A = \begin{bmatrix} a+bi & c+di \\ -c+di & a-bi \end{bmatrix},$$

$P_A(A) = A^2 - 2aA + (a^2+b^2+c^2+d^2)I = 0$, and the polynomial given in Theorem 1.3 is characteristic polynomial of A

II. GENERALIZATION OF CAYLEY HAMILTON THEOREM

Theorem 2.1 (Cayley-Hamilton Theorem). For any n × n Matrix A, $P_A(A)=0$.

Proof. Let D(x) be the matrix with polynomial entries $D(x)= adj(xI_n-A)$, So $D(x)(xI-A)= det(xI_n-A)I_n$. Since each entry in D(x) is the determinant of an (n-1) X((n-1) submatrix of (xI_n-A), each entry of D(x) is a polynomial of degree less than or equal to n-1. It folowws that there exist matrices $D_0, D_1,\ldots D_{n-1}$ with entries from C such that $D(x) = D_{n-1} x^{n-1}+\ldots+ D_1 x +D_o$. Then the matrix equation follows

$det(xI_n-A) I_n = (x I_n-A) adj(xI_n-A) = (xI_n-A)D(x)$

Substituting $p_A(x)= det(xI_n-A)$, (and using the fact that scalars commute with matrix)

$X^nI_n+ b_{n-1}x^{n-1}I_n+\ldots+ b_1xI_n+b_0I_n$

$= p_A(x) I_n= det(xI_n-A)I_n$

$=(xI_n-A)adj(xI_n-A)$

$=(xI_n-A)(x^{n-1}D_{n-1}+\ldots+xD_1+D_o)$

$=x^{nD}{}_{n-1}- x^{n-1}AD_{n-1}+ x^{n-1}D_{n-2}-x^{n-2}AD_{n-2}+\ldots+ xD_0-AD_o$

$=x^{nD}{}_{n-1}+ x^{n-1}(-AD_{n-1}+ D_{n-2}) +\dots\dots+(-AD_1+D_o) -AD_o$

Since two polynomials are equal if and only if their coefficients are equal , the coefficient matrices are equal ; that is , $I_n=D_{n-1}$, $b_{n-1}I_n= (-AD_{n-1}+D_{n-2})$,……,$b_1I_n=-(AD_1+D_0)$, and $b_0I_n=-AD_0$. This means that A may be substituted for the variable x in the equation (2.1) to conclude

$P_A(A) = A^n+ b_{n-1} A^{n-1}+\dots\dots+ b_1A+b_0I_n$

$=A^nD_{n-1}+A^{n-1}(-AD_{n-1} +D_{n-2})+\dots\dots+A(-AD_1+D_0)-AD_0$

$=A^nD_{n-1}-A^nD_{n-1}+A^{n-1}D_{n-2}- A^{n-1}D_{n-2}+\dots\dots\dots AD_0-AD_0$

$=0$

This proves the theorem

III. PROOF THROUGH SCHUR'S TRIANGULARIZATION

This proof synthesizes work of Issai Schur, According to this, If $S_1S_2,\dots\dots,S_n$ are $n \times n$ upper –triangular matrices such that *(i,i)* element of S_i is zero for all I, then $S_1S_2\dots\dots S_n=0$

*Proof. (*Induction of n)

For n=1 , there is nothing to be proved since $S_1=0$

For n=2 , $S_1=\begin{bmatrix} 0 & x \\ 0 & y \end{bmatrix}$, $S_2=\begin{bmatrix} u & v \\ 0 & 0 \end{bmatrix}$, where x,y,u,and v are scalars . It is clear that $S_1S_2=0$

Assume now that the theorem is true for some integer m, then $S1S2\dots\dots\dots S_mS_{m+1}= T\, S_{m+1}$,

Where, $S_1=\begin{bmatrix} T_i & a_i \\ \bar{0} & x_1 \end{bmatrix}$, ………., $S_m=\begin{bmatrix} T_m & a_m \\ \bar{0} & x_m \end{bmatrix}$

And $T_1,\dots\dots\dots T_m \in M_m$ are upper triangular matrices such that the (i,i) element of T_i is zero.

Then,

$T= S_1S_2\dots\dots S_m= \begin{bmatrix} T_1T_2\dots\dots\dots T_M & u_m \\ \bar{0} & x \end{bmatrix}=\begin{bmatrix} 0_m & u_m \\ \bar{0} & x \end{bmatrix}$ and $S_{m+1} =\begin{bmatrix} A_m & t_m \\ \bar{0} & 0 \end{bmatrix}$

Where $0_m \in M_m$, $A_m \in M_m$ is upper triangular , u_m and t_m are vector columns of order m X 1, $\bar{0}$ is zero row of order 1 X m , and x is a scalar. The Proof is complete.

Main Theorem.(Cayley- Hamilton Theorem).

Let $p_A(t)$ be the characteristic polynomial of $A \in M_m$. Then $P_A(A)=0$

Proof. Since $p_A(t)$ is of degree n with leading coefficient 1 and the roots of $p_A(t)$ are precisely the eigen values $\lambda_1\dots\dots\dots\dots, \lambda_n$ of A, counting multiplicities , factor $p_A(t)$ as

$P_A(t) =(t- \lambda_1)(t- \lambda_2)\dots\dots\dots(t- \lambda_m)$

Using Schur's Theorem , write A as

$A= UTU^*$

Where T is upper triangular with λ_i in the ith diagonal position , i=1,……n. The theorem follows.

$P_A(A)$

$= P_A(UTU^*)=(UTU^*- \lambda_1I)(UTU^*- \lambda_2I)\dots\dots\dots(UTU^*- \lambda_nI)$

$=[U(T- \lambda_1I)U^*][U(T- \lambda_2I)U^*]\dots\dots\dots\dots [U(T- \lambda_nI)U^*]$

$=U[(T- \lambda_1I)(T- \lambda_2I)\dots\dots\dots\dots\dots\dots\dots\dots (T- \lambda_nI)]U^*$

$=0$

The last equality follows from theorem 2.1.

IV. TOPOLOGICAL PROOF

This is most concise alternate proof to the Frobenius's proof.

Theorem 3.1. The set $D_n= \{A \epsilon M_n | A \text{ is diagonalizable}\}$ of diagonalizable matrices dense in M_n

Proof. Fix $\varepsilon > 0$. It is sufficient to show that , given any matrix $A \epsilon M_n$, there exists diagonalizable matrix B such that $|||A-B|||_2 < \varepsilon$

Where $|||A-B|||_2$ is the Frobenius norm given by

$$|||A-B|||_2= \left(\sum_{i,j=1}^{n}\left|a_{ij}\right|^2\right)^{1/2}$$

Given $A\epsilon M_n$,let A=UTU* where U is unitary and T is upper –Triangular , possible Schur's Triangularization theorem. Define B=A+ UCU* , where

$$C_{fg}=\begin{Bmatrix} 0 & if\ f\neq g \\ \frac{\varepsilon\delta}{2f\sqrt{n}} & if\ f=g \end{Bmatrix}$$

Choose δ>o , so that B will have distinct eigen values, hence B will be diagonalizable .Let $\sigma(A)=\{\lambda_f, f=1,2,\ldots\ldots\ldots n\}$ and $\lambda_{i+1}\geq\lambda_f$. Define δ as

$$\delta<\min_{\lambda_f\neq\lambda_g}\left\{\frac{(\lambda_g-\lambda_f)2\sqrt{n}}{\varepsilon}\left(\frac{1}{f}-\frac{1}{g}\right)\right\}\ and\ \delta<1$$

If $\lambda_h=\lambda_i$

$$\lambda_h+\frac{\varepsilon\delta}{2h\sqrt{n}}=\lambda_i+\frac{\varepsilon\delta}{2h\sqrt{n}}$$

$$\neq\ \lambda_i+\frac{\varepsilon\delta}{2i\sqrt{n}}$$

If $\lambda_h\neq\lambda_i$

$$\frac{(\lambda_h-\lambda_i)2\sqrt{n}}{\varepsilon\left(\frac{1}{h}-\frac{1}{i}\right)}\geq\min_{\lambda_g\geq\lambda_f}\left\{\frac{(\lambda_g-\lambda_f)2\sqrt{n}}{\varepsilon}\left(\frac{1}{f}-\frac{1}{g}\right)\right\}>\ \delta$$

$$\Rightarrow(\lambda_i-\lambda_h)>\frac{\varepsilon\delta}{2i\sqrt{n}}\left(\frac{1}{h}-\frac{1}{i}\right)$$

$$\Rightarrow\lambda_i+\frac{\varepsilon\delta}{2i\sqrt{n}}>\lambda_h+\frac{\varepsilon\delta}{2h\sqrt{n}}$$

The Diagonal entries of T+C are distinct from the choice of , on B has distinct eigen values and is diagonalizable. Then

$|||A-B|||_2=|||A-(UCU^*|||_2$

$=|||C|||_{2,}$ because $|||A-B|||_2$ is unitarily invariant

$=\sqrt{\sum_{f=1}^{n}\left(\frac{\varepsilon\delta}{2f\sqrt{n}}\right)^2}$ because C is Diagonal $\leq\sqrt{\sum_{f=1}^{n}\left(\frac{\varepsilon\delta}{2}\right)^2}$

$=\frac{\varepsilon\delta}{2}$

$<\varepsilon$ since every $\delta<1$

Example 3.1 If A is diagonalizable , then $p_A(A)=0$

Proof. A=PDP^{-1}, where P is invertible.

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

And λi *are the eigen value of A.Then,*

$P_A(A)$ = (A- λ_1I)(A- λ_2I)………..(A- λ_nI)

= P(D- λ_1I)(D- λ_2I)………..(D- λ_nI)P^{-1}

= P.0.P^{-1}

= 0

Main Theorem.(Cayley Hamilton).*If $P_A(t)$ is the characteristics polynomial of A then $P_A(A)$ =0*

Proof. Let P_n be the space of polynomials of degree n or less, with the Zariski topology. From Example 3.1 , the Cayley-Hamilton theorem is proved for all diagonalizable matrices. The mapping Ω: M_n X P_N →M_n given by Ω(A,f(x))=f(A) is continuous , and the mapping ϕ: M_n→ M_n X P_N given by $\phi(A)$ =(A, pA(x)) is continuous . Hence the composition

Ωo ϕ: M_n→$M_{n,}$

Which is given by Ωo ϕ(A)= p_A(A) is continuous. This mapping is identically zero on a dense subject of M_n, so by continuity vanishes everywhere.

V. CHARACTERSTICS

The Cayley –Hamilton theorem is one of the most powerful and classical matrix theory theorem. Many application derive their results from this theorem. To understand the scope of this theorem , alternate proofs were used. Each proof helped to understand how intertwined areas of mathematics are with respect to matrices and the characteristics polynomial.

VI. APPLICATION OF CAYLEY– HAMILTON THEOREM

A very common application of the Cayley- Hamilton Theorem is to use it to find A^n usually for the large powers of n. However many of the techniques involved require the use of the eigen values of A.

REFERENCES

[1] William A. Adkins and Mark G. Davidson, *The Cayley Hamilton and Frobenious theorems via the Laplace Transformation*, Linear Algebra and its Applications 371(2003), 147-152.

[2] Arthur Cayley, *A memoir on the theory of Matrices*, available from http:// www.jstor.org, 1857.

[3] Jeffrey A. Rosoff, *A topological Proof of the Cayley- Hamilton Theorem*, Missouri J. Math .Sc. 7(1995), 63-67.

[4] Wikipedia , Arthur Cayley, Available from http:// en.wikipedia .org, 2004

[5] Wikipedia, William Rowan Hamilton, Available from http:// en. wikipedia .org, 2005

[6] D.R. Wilkins, *Linear Operators and the Cayley –Hamilton Theorem* , available from http://www.maths.tod.i.e/pub/histMath/people/Hamilton, 2005

[7] Raghib Abu –Saris and Wajdi Ahmad, *Avoiding Eigen values in Computing Matrix Powers*. The Mathematical Association of America 112(2005), 450-454.

