Paper Study: Eigenvectors from Eigenvalues

Yu-Hsuan Huang

National Chiao Tung University

December 2019

Overview

Today you will be expected to learn the eigenvector eigenvalue identity re-discovered by Terrence Tao et al. with examples and 5 different proof techniques used in the paper and some backgrounds required to understand them.

- Introduction
- Adjugate Matrix Proof
- Cramer's Rule Proof
- Perturbative Analysis Proof
- Exterior Algebra and Coordinate Free Proof
- Cauchy Binet Type Formula Proof
- Further Discussion

Introduction

Eigenvalues and eigenvectors perhaps are one of the most fundamental objects in linear algebra. In regular textbooks, it is usually defined as following.

Definition

Given a linear operator $L : V \to V$, an eigenvalue $\lambda \in \mathbb{C}$ and its corresponding eigenvector $v \in V$ satisfies,

$$Lv = \lambda v.$$

One text-book approach for deriving an eigen pair is to first solve for eigenvalue λ that satisfies the characteristic equation,

$$p(\lambda) := \det(\lambda I - A) = 0.$$

Then solve for the eigen space V_λ by Gaussian elimination,

$$V_\lambda := \ker(\lambda I - A).$$

Introduction

Tao et al. explicitly wrote the following identity, which may look quite surprising at the first glance.

Theorem (Eigenvector Eigenvalue Identity)

Suppose $A \in \mathbb{C}^{n\times n}$ is a Hermitian matrix and denote the i-th eigenpair as,

$$Av_i = \lambda_i(A)v_i.$$

Then the following identity holds,

$$|v_{ij}|^2 \prod_{1\leq k\leq n: k\neq i} (\lambda_i(A) - \lambda_k(A)) = \prod_{1\leq k\leq n-1} (\lambda_i(A) - \lambda_k(M_j)),$$

where M_j is the j, j-minor of A.

Introduction

Consider the following matrix,

$$A = \begin{pmatrix} 5 & -1 \\ -1 & 5 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 4 & 0 \\ 0 & 6 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{pmatrix},$$

with minor $M_1 = \begin{pmatrix} 5 \end{pmatrix}$ and $M_2 = \begin{pmatrix} 5 \end{pmatrix}$,

$$\begin{aligned} \frac{4-5}{4-6} &= \frac{1}{2} = |v_{1,1}|^2, \\ \frac{4-5}{4-6} &= \frac{1}{2} = |v_{1,2}|^2 = |v_{2,1}|^2, \\ \frac{6-5}{6-4} &= \frac{1}{2} = |v_{2,2}|. \end{aligned}$$

Adjugate Matrix Proof

The adjugate matrix mimics matrix inverse in an explicit manner,

$$A^{\text{adj}}_{ij} := (-1)^{i+j} \det M_{ji}.$$

One general fact often used to annihilate matrix A is that,

$$A^{\text{adj}} A = A A^{\text{adj}} = \det(A) I.$$

Since A^{adj} and A commutes, they share the same set of eigenvectors. By double sided evaluation we have,

$$\det(\lambda I - M_j) = (\lambda I - A)^{\text{adj}}_{jj} = \sum_{1 \le i \le n} \left(\prod_{1 \le k \le n, k \ne i} (\lambda - \lambda_k(A)) \right) v_{ij} v^*_{ij}.$$

Substitute $\lambda = \lambda_i(A)$ and rewrite determinant as eigenvalues, the proof is done.

Adjugate Matrix Proof

Consider $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, then $A^{\text{adj}} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$. Indeed, $A^{\text{adj}}A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = AA^{\text{adj}}$ commutes and admits a simultaneous diagonalization,

$$A = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{pmatrix},$$

$$A^{\text{adj}} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{pmatrix}.$$

Also, $\sigma(A) = \{2, 0\}$ and we could check,

$$\det(\lambda_1(A)I - M_1) = 1$$

$$= \frac{1}{2}(2-0) + \frac{1}{2}(2-2) = \sum_{1 \le i \le n} \left(\prod_{1 \le k \le n, k \ne i} (\lambda - \lambda_k(A)) \right) v_{ij} v_{ij}^*.$$

Cramer's Rule Proof

It suffices to consider the case of A with simple spectrum. Two way evaulating the j, j component of the following inverse matrix would yield,

$$\frac{\det(\lambda I - M_j)}{\det(\lambda I - A)} = (\lambda I - A)^{-1}_{j,j} = \sum_{1 \le i \le n} \frac{v_{ij} v_{ij}^*}{\lambda - \lambda_i(A)}, \text{ for } \lambda \notin \sigma(A).$$

Writing everything in terms of eigenvalues would yield,

$$\frac{\prod_{1 \le k \le n-1}(\lambda - \lambda_k(M_j))}{\prod_{1 \le k \le n}(\lambda - \lambda_k(A))} = \sum_{1 \le i \le n} \frac{|v_{ij}|^2}{\lambda - \lambda_i(A)}.$$

Both sides of the equation are meromorphic with poles at $\lambda_i(A)$ for all i. Taking their residues and we should conclude the proof,

$$2\pi\sqrt{-1}\frac{\prod_{1 \le k \le n-1}(\lambda - \lambda_k(M_j))}{\prod_{1 \le k \le n, k \ne i}(\lambda - \lambda_k(A))} = \oint_{C_{\lambda_i(A)}} \frac{\prod_{1 \le k \le n-1}(\lambda - \lambda_k(M_j))}{\prod_{1 \le k \le n}(\lambda - \lambda_k(A))}$$

$$= \oint_{C_{\lambda_i(A)}} \sum_{1 \le i \le n} \frac{|v_{ij}|^2}{\lambda - \lambda_i(A)} = 2\pi\sqrt{-1}\,|v_{ij}|^2$$

Cramer's Rule Proof

Assume $j = 1$, for $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, the left hand side would be,

$$\frac{\prod_{1 \le k \le n-1} (\lambda - \lambda_k(M_j))}{\prod_{1 \le k \le n} (\lambda - \lambda_k(A))} = \frac{\lambda - 1}{(\lambda - 2)(\lambda - 0)},$$

and the right hand side would be,

$$\sum_{1 \le i \le n} \frac{|v_{ij}|^2}{\lambda - \lambda_i(A)} = \frac{\frac{1}{2}}{\lambda - 2} + \frac{\frac{1}{2}}{\lambda - 0}.$$

Indeed, they are equal as rational polynomials. Taking residue on the both side, we obtain,

$$\oint_{C_{\lambda_1(A)}} \frac{\frac{1}{2}}{\lambda - 2} + \frac{\frac{1}{2}}{\lambda - 0} = \pi\sqrt{-1} + 0 = 2\pi\sqrt{-1}\left|v_{1j}\right|^2,$$

$$\oint_{C_{\lambda_2(A)}} \frac{\frac{1}{2}}{\lambda - 2} + \frac{\frac{1}{2}}{\lambda - 0} = 0 + \pi\sqrt{-1} = 2\pi\sqrt{-1}\left|v_{2j}\right|^2.$$

Perturbative Analysis Proof

This proof consider the rank one perturbation $A + \epsilon e_j e_j^*$ where e_j are standard basis vector for $1 \leq j \leq n$. Such perturbation affects both the spectrum and characteristic polynomial of A. For the former, the interlacing behavior occurs as,

$$\lambda_i(A + \epsilon e_j e_j^*) = \lambda_i(A) + \epsilon \left| v_{ij} \right|^2 + O(\epsilon^2).$$

For the latter,

$$p_{A+\epsilon e_j e_j^*}(\lambda) := \det(\lambda I - A - \epsilon e_j e_j^*) = p_A(\lambda) - \epsilon p_{M_j}(\lambda) + O(\lambda^2).$$

By extracting the linear order term for $p_{A+\epsilon e_j e_j^*}(\lambda_i(A + \epsilon e_j e_j^*)) = 0$, we have,

$$\epsilon \left| v_{ij} \right|^2 p_A'(\lambda_i(A)) - \epsilon p_{M_j}(\lambda_i(A)) = 0,$$

which concludes the proof.

Perturbative Analysis Proof

Assume also $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, then

$$\det(\lambda I - A - \epsilon e_1 e_1^*) = \det \begin{pmatrix} \lambda - 1 - \epsilon & 1 \\ -1 & \lambda - 1 \end{pmatrix} = \lambda^2 - (2+\epsilon)\lambda + \epsilon,$$

and we have $\sigma(A + \epsilon e_1 e_1^*) = \left\{ \frac{2+\epsilon \pm \sqrt{\epsilon^2+4}}{2} \right\}$,

$$\lambda_1(A + \epsilon e_1 e_1^*) = \frac{2 + \epsilon + \sqrt{\epsilon^2 + 4}}{2} \in 2 + \frac{\epsilon}{2} + O(\epsilon^2) = 2 + |v_{12}|\,\epsilon + O(\epsilon^2),$$
$$\lambda_2(A + \epsilon e_1 e_1^*) = \frac{2 + \epsilon - \sqrt{\epsilon^2 + 4}}{2} \in \frac{\epsilon}{2} + O(\epsilon^2) = 2 + |v_{22}|\,\epsilon^2 + O(\epsilon^2).$$

Finally, plug into characteristic polynomial, we do have

$$\epsilon\,|v_{11}|^2\, p_A'(\lambda_1(A)) = \epsilon = \epsilon p_{M_1}(\lambda_1(A)),$$
$$\epsilon\,|v_{21}|^2\, p_A'(\lambda_2(A)) = -\epsilon = \epsilon p_{M_1}(\lambda_2(A)).$$

Exterior Algebra

The notion of directional volume could be introduced by exterior algebra. Formally speaking, one could first consider the tensor algebra over a linear space V is the algebra obtained by adding in tensor product as multiplication,

$$T(V) := \mathbb{C} \oplus \left(\bigoplus_{1 \leq i} V^{\otimes i} \right).$$

As an $\mathbb{C}$ algebra, $T(V)$ contains a double sided ideal,

$$J := \langle v \otimes v : v \in V \rangle \lhd T(V).$$

Its quotient algebra is called an exterior algebra. In convention, we write the multiplication using $\wedge$ instead of $\otimes$, and we may write it as,

$$\bigwedge(V) := T(V)/J.$$

Exterior Algebra

For convenience, we denote $V^{\wedge k} := \mathrm{Span}\{v_1 \wedge \cdots \wedge v_k : v_i \in V\}$ and any vector in it called a k-vector. We say that $\bigwedge(V)$ anti-symmetrize $T(V)$ since for any $u, v \in V$ we have,

$$0 = (u+v) \wedge (u+v) = u \wedge u + v \wedge v + u \wedge v + v \wedge u = u \wedge v + v \wedge u,$$

thus $\wedge$ anti-commutes, i.e. $u \wedge v = -v \wedge u$. Note that if $k > \dim V =: n$, then $V^{\wedge k}$ must be zero. Then we might rewrite the entire algebra as,

$$\bigwedge(V) = \bigoplus_{0 \leq i \leq k} V^{\wedge i},$$

which is finite dimensinal.

Exterior Algebra

For example, assume $V = \mathbb{C}^3 = \text{Span}\{x, y, z\}$, then $V^{\wedge 2} = \text{Span}\{x \wedge y, y \wedge z, z \wedge x\}$ and $V^{\wedge 3} = \{x \wedge y \wedge z\}$. Conisder $u = a_1 x + a_2 y + a_3 z, v = c_1 x + c_2 y + c_3 z \in V$, then

$$\begin{aligned} u \wedge v &= (a_1 c_2 - a_2 c_1) x \wedge y + (a_2 c_3 - a_3 c_2) y \wedge z + (a_3 c_1 - a_1 c_3) z \wedge x \\ &= \begin{vmatrix} a_1 & a_2 \\ c_1 & c_2 \end{vmatrix} x \wedge y + \begin{vmatrix} a_2 & a_3 \\ c_2 & c_3 \end{vmatrix} y \wedge z + \begin{vmatrix} a_3 & a_1 \\ c_3 & c_1 \end{vmatrix} z \wedge x. \end{aligned}$$

Adding in $w = r_1 x + r_2 y + r_3 z \in V$, then,

$$\begin{aligned} u \wedge v \wedge w &= \left(r_3 \begin{vmatrix} a_1 & a_2 \\ c_1 & c_2 \end{vmatrix} + r_2 \begin{vmatrix} a_2 & a_3 \\ c_2 & c_3 \end{vmatrix} - r_1 \begin{vmatrix} a_3 & a_1 \\ c_3 & c_1 \end{vmatrix} \right) x \wedge y \wedge z \\ &= \begin{vmatrix} a_1 & a_2 & a_3 \\ c_1 & c_2 & c_3 \\ r_1 & r_2 & r_3 \end{vmatrix} x \wedge y \wedge z = \pm \text{vol} \left(\frac{V}{\mathbb{Z}u + \mathbb{Z}v + \mathbb{Z}w} \right) x \wedge y \wedge z \end{aligned}$$

Exterior Algebra

Inner product $\langle \cdot, \cdot \rangle$ in V could be taken into $V^{\wedge k}$ in the following manner,

$$\langle u_1 \wedge \cdots \wedge u_k, v_1 \wedge \cdots \wedge v_k \rangle := \det \left(\langle u_i, v_j \rangle \right)_{1 \leq i,j \leq k} .$$

By noticing that, as linear spaces,

$$\dim V^{\wedge k} = \binom{n}{k} = \binom{n}{n-k} = \dim V^{\wedge n-k},$$

we could have a natural isomorphism called the Hodge dual,

$$* : V^{\wedge k} \xrightarrow{\sim} V^{\wedge n-k},$$

by the following equation that provide uniqueness,

$$u \wedge *v = \langle u, v \rangle \omega,$$

where $\omega \in V^{\wedge n}$ is a pre-selected unit n-vector.

Coordinate Free Proof

This prove improves the adjugate proof in the sense that it does not rely on matrix representation but the intrinsic property of the linear map itself. We first see the target identity, for one side, there is,

$$\prod_{1\le k\le n, k\ne i} \lambda_k(A).$$

By translation symmetry we might assume $\lambda_i(A) = 0$, then the product could be describe as $\Delta_A(v_i)$ being the determinant of the quadratic form,

$$w \mapsto \langle Aw, w\rangle,$$

restricted in $v_i^\perp := \{w \in \mathbb{C}^n : \langle w, v_i\rangle = 0\}$. Similarly we have,

$$\Delta_A(e_j) = \det(M_j) = \prod_{1\le k\le n-1} \lambda_k(M_j).$$

We thus turn to prove the coordinate free version of the identity,

$$|\langle v_i, e_j\rangle|^2 \, \Delta_A(v_i) = \Delta_A(e_j).$$

Coordinate Free Proof

Theorem

Suppose self adjoint $A : \mathbb{C}^n \to \mathbb{C}^n$ *annihilate* v*, for each unit* $f \in \mathbb{C}^n$*, we have,*

$$|\langle v, f\rangle|^2 \, \Delta_A(v) = \Delta_A(f).$$

First of all, for any $x \in \mathbb{C}^n$, $\Delta_A(x) = \langle Aw, w\rangle / \langle w, w\rangle$ for $w \in \left(x^{\perp}\right)^{\wedge n} - \{0\}$ with A extending to $\bigwedge(\mathbb{C}^n)$ naturally. Since f is unit, so is $*f \in \bigwedge(f^{\perp})$, thus,

$$\Delta_A(f) = \langle A(*f), *f\rangle / \langle *f, *f\rangle = \langle A(*f), *f\rangle.$$

The equation turns into,

$$|\langle v, f\rangle|^2 \, \Delta_A(v) = \langle A(*f), *f\rangle,$$

and we turn to argue for a more general case with $g \in \mathbb{C}^n$ also a unit,

$$\langle f, v\rangle \langle v, g\rangle \Delta_A(v) = \langle A(*f), *g\rangle.$$

Coordinate Free Proof

There are three possible cases we need to discuss, due to multi-linearity of left hand side.

- $f \perp v$, in this case since $*f \in \langle v \rangle$, $A(*f) = 0$ by assumption.
- $v \perp g$, in this case since $*g \in \langle v \rangle$, we have $\langle A(*f), *g \rangle = \langle *f, A^*(*g) \rangle = \langle *f, A(*g) \rangle = \langle *f, 0 \rangle = 0$.
- $f = v = g$, in this case $\langle f, v \rangle \langle v, g \rangle \Delta_A(v) = |\langle f, f \rangle|^2 \Delta_A(f) = \Delta_A(f) = \langle A(*f), *f \rangle$.

This conclude the proof. $\square$

Cauchy Binet Type Formula Proof

Lemma

Suppose A Hermitian with a zero eigenvalue, say $\lambda_n(A)$ and $B \in \mathbb{C}^{n\times n}$, then,

$$\prod_{1\leq i\leq n-1} \lambda_i(A) \left|\det\begin{pmatrix} B & v_n \end{pmatrix}\right|^2 = \det\left(B^* AB\right).$$

Firstly, with proper selection of coordinate, we might assume $i = n, j = 1$ in Theorem 2. One could also find that the equation is translation symmetric, i.e. the equation hold if and only if it hold when we substitute A with $A + \lambda I$. Therefore, WLOG we could assume that $\lambda_n(A) = 0$ by translation. Pick $B = \begin{pmatrix} 0 \\ I_{n-1} \end{pmatrix}$, then we immediately get,

$$\prod_{1\leq i\leq n-1} \lambda_i(A) \left|v_{n1}\right|^2 = \det(M_n),$$

which conclude the proof. $\square$

Further Discussion

- There is absolute square enclosed outside $|v_{ij}|$: could restore amplitude, but not global phase (imagine replacing all v_{ij} with $-v_{ij}$).
- What about local phase? Suppose we are now considering real symmetric matrices, applying the identity in a basis that contains $\frac{1}{\sqrt{2}}(e_j + e_{j'})$ would give us magnitude for $\frac{1}{\sqrt{2}}(v_{ij} + v_{ij'})$, together with magnitude of v_{ij} and $v_{ij'}$, we could restore their relative sign.
- Not practical for spectral decomposition of general matrices: require entire spectrum of A and minors M_j.
- Given structured matrices possible for practical use, e.g. tridiagonal matrices.