Paper Study: Eigenvectors from Eigenvalues

Yu-Hsuan Huang
asd00012334@hotmail.com

December 2019

Abstract

This is a report on studying, "Eigevectors from Eigenvalues," presented by Terrence Tao et al. Roughly speaking, the literature reproves a matrix identity with various different technique. This reports would include some methodologies used to prove the identity, and what we have learnt from it.

1 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 1. *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,

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

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

Theorem 1 (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,

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

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

The identity gives a direct way to determine componentwise amplitudes, $|v_{ij}|$ of an eigenvector directly from the eigenvalues of A and its principal minors.

2 Methodology

There are various technique capable of deriving the identity. From the algebraic perspective, we could start from adjugate proof, Crammer's rule to exterior algebraic proof, which gradually relaxed explicit constructions into existential proof. From the analytic perspective, proof with Crammer's rule and perturbative analysis treat eigen pairs as meromorphic objects and study their local behaviors.

2.1 Spectral Theorem

It is well knwon that any Hermitian matrix could be unitarily diagonalized. Indeed, the assumption could be relaxed to normal matrices, which is defined as any square matrix A that commutes with its adjoint $A^*A = AA^*$. Hermitian matrix is just a special case of normal matrix. This is due to Schur decomposition, in which any complex valued square matrix is untarily similar to a upper triangular matrix,

$$A = QUQ^*,$$

where Q is unitary and U is upper trinagular. Then

$$QU^*UQ^* = AA^* = A^*A = QU^*UQ^*,$$

would yield $UU^* = U^*U$, thus U must be diagonal and the Schur decomposition itself admits a unitary diagonalization.

2.2 Adjugate Matrix Proof

The adjugate matrix mimics matrix inverse in an explicit manner,

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

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

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

Since matrix determinant is just product of eigenvalues, the adjugate matrix could be written in terms of eigenvalues. Also, since A^{adj} and A commutes, they also share the same set of eigenvectors. These interesting fact makes adjugate matrix well suited for proving the eigenvector eigenvalue identity. By double sided evaluation we have,

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

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

2.3 Cramer's Rule Proof

There is a taste of both analysis and algebra within the proof with Cramer's rule. Since Hermitian matrices with simple spectrum are dense among all Hermitian matrices, 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,

$$\begin{aligned} 2\pi\sqrt{-1}\frac{\prod_{1\le k\le n-1}(\lambda - \lambda_k(M_j))}{\prod_{1\le k\le n, k\ne i}(\lambda - \lambda_k(A))} &= \oint_{C_{\lambda_i(A)}} \frac{\prod_{1\le k\le n-1}(\lambda - \lambda_k(M_j))}{\prod_{1\le k\le n}(\lambda - \lambda_k(A))} \\ &= \oint_{C_{\lambda_i(A)}} \sum_{1\le i\le n} \frac{|v_{ij}|^2}{\lambda - \lambda_i(A)} = 2\pi\sqrt{-1}\,|v_{ij}|^2. \end{aligned}$$

2.4 Perturbative Analysis Proof

This proof consider the rank one perturbation $A+\epsilon e_j e_j^*$ where e_j are standard basis vector for $1 \le j \le n$. 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.

2.5 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,

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

As a $\mathbb{C}$ 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,

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

For convenience, we denote $V^{\wedge k} := \mathsf{Span}\{v_1 \wedge \cdots \wedge v_k : v_i \in V\}$ 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,

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

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

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

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.

2.6 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,

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

Theorem 2. *Suppose self adjoint $A : \mathbb{C}^n \to \mathbb{C}^n$ annihilate v, for each unit $f \in \mathbb{C}^n$, we have,*

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

Proof. 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,

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

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. □

2.7 Cauchy Binet Type Formula Proof

Lemma 1. *Suppose A Hermitian with a zero eigenvalue, say $\lambda_n(A)$ and $B \in \mathbb{C}^{n\times n}$, then,*

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

Firstly, with proper selection of coordinate, we might assume $i = n, j = 1$ in Theorem 1. 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,

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

which conclude the proof.

3 Disscussion

We give some comment on these various proof techniques as well as the applicability of this identity in practical use. We also discuss some good points mentioned in the paper.

3.1 Algebraic Proof Technique

One advantage using adjugate matrix for proving is that it avoids matrix inversion issue, which makes it the simplest proof in the whole paper. Furthermore, by construction, adjugate matrix preserve normal property of A, which would generalize the proof such that it also holds if we only require A to be normal instead of Hermitian. This technique could be used to prove another off diagonal variant,

$$(-1)^{j+j'} \det(\lambda_i(A)(I_n)_{jj'} - M_{jj'}) = \left(\prod_{1 \le k \le n, k \ne i} (\lambda_i(A) - \lambda_k(A)) \right) v_{ij} \bar{v_{ij'}},$$

where $(I_n)_{jj'}$ is the j, j' minor of I_n. Then we could proof the identity without explicitly mention the adjugate matrix, the Crammer's rule kicks in and essentially is doing the same thing as the adjugate matrix, except now we require invertibility. Finally, my favorite proof using exterior algebra, could mitigate the matrix representation. It is because values such as determinant and spectrum of minors could be thought as hyper volume and formalized as norm within exterior algebra. This method gives good geometric interpretation. In fact, exterior algebra, equipped on smooth manifold, would yield exterior derivative, which eventually leads to a special kind of cochain complex that for de Rham cohomology.

3.2 Analytic Proof Technique

The adjugate matrix proof mitigate any problem about invertibility, while we turn to us Crammer's rule to represent inverse matrix $(\lambda I - A)^{-1}$, we must constraint $\lambda \notin \sigma(A)$. On the other hand, due to differentiability of such matrix w.r.t. λ, those undefined points actually captures properties regarding the spectrum of A. Indeed, we use Contour integral to calculating residue around each poles, which eliminate singularities in the equation. This is generally called the resolvent formalism, and is widely used to analyze spectral properties with complex analysis tools.

On the other hand, the perturbative analysis proof, is relatively direct. It studys the errors introduced by shifting the original matrix with a (usually rank one) ϵ error. The first order terms are usually easy to derive and in this case also mitigates invertibility issues.

3.3 Reduction

In the analysis proofs, we assume that A is equipped with simple spectrum, which may mitigate several issues,

- When the algebraic multiplicity of A is greater than one for some eigenvalue, the interlacing behavior under perturbation seems unclear.
- When considering singularity of $(\lambda I - A)^{-1}$, there will be non-simple poles locating at where the repeated eigenvalues are.

On the other hand, in the Cauchy Binet type formula proof, we do an assumption that $i = n, j = 1$ and $\lambda_n(A)$. The former part could be done by permuting eigenvalues and choice of good coordinate, in another word, for general case we might simply substitute A with QAQ^* where Q is unitary. The latter part is due to translation symmetry. It is because when we substitute A with $A + \lambda I$ then both the spectrum of A and minors M_j will be shifted the same amount. Thus the equation is still equivalent with the old one,

$$
\begin{aligned}
|v_{ij}|^2 \prod_{1\le k\le n: k\ne i} (\lambda_i(A+\lambda I) - \lambda_k(A+\lambda I)) &= |v_{ij}|^2 \prod_{1\le k\le n: k\ne i} (\lambda_i(A) - \lambda_k(A)) \\
&= \prod_{1\le k\le n-1} (\lambda_i(A) - \lambda_k(M_j)) \\
&= \prod_{1\le k\le n-1} (\lambda_i(A+\lambda I) - \lambda_k(M_j+\lambda I)).
\end{aligned}
$$

3.4 Applicability

There is absolute square enclosed outside $|v_{ij}|$, therefore the formula helps to calculate only the magnitude of v_{ij}. Obviously, it is not possible to completely restore the phase since we could replace v_{ij} with $-v_{ij}$ and the resulting matrix is still the same. This is when we would loss information on the global phase. On the other hand, the paper mentioned that it is possible to restore the local phase. For instance, suppose we are now considering real symmetric matrices, applying the identity 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.

The identity itself is not practical in use if we want a fast algorithmic for spectral decomposition of general matrices. This is because, in order to apply the formula to obtain magnitudes, one needs to compute the entire spectrum together with those of the minors M_j. On the other hand, the paper also mentioned that, if we are given structured matrices, for example, tridiagonal ones, then practical algorithm could be expected. We might say that applications of this identity in practical computation is rather indirect.