Eigenvalues and Eigenvectors
An Introduction

P. Sam Johnson

November 3, 2014

Overview

We discuss eigenvalues and eigenvectors associated with a complex square matrix. These are useful in the study of canonical forms of a matrix under similarity and in the study of quadratic forms.

They have applications in many subjects like Geometry, Mechanics, Astronomy, Enigneering, Economics and Statistics.

Throughout in the lecture we take the base field to be $\mathbb{C}$ (the set of all complex numbers) except in a few places where we take it to be $\mathbb{R}$ (the set of all real numbers).

'A' will denote an $n \times n$ matrix unless specified otherwise.

For any $n \times n$ matirix, consider the polynomial

$$\chi_A(\lambda) := |\lambda I - A|.$$

Clearly this is a monic polynomial of degree n.

So, by the fundamental theorem of algebra, $\chi_A(\lambda)$ has exactly n (not necessarily distinct) roots, usually denoted by $\lambda_1, \lambda_2, \ldots, \lambda_n$. Clearly

$$\chi_A(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n).$$

$\chi_A(\lambda)$ is called the **characteristic polynomial** of A and $\chi_A(\lambda) = 0$ is called the **characteristic equation** of A.

The n roots of $\chi_A(\lambda)$ are called the **characteristic roots** of A. The **spectrum** of A is the set of distinct characteristic roots of A.

Finding the characteristic roots of a matrix is not easy in general, since there is **no easy way** of finding the roots of a polynomial of degree greater than 3.

Clearly, a complex number α is a characteristic root of A iff $\alpha I - A$ is singular.

In particular, 0 is a characteristic root of A iff A is singular. The constant term in $\chi_A(\lambda)$ is $(-1)^n|A|$.

The coefficient of λ^{n-1} in $\chi_A(\lambda)$ is $-tr(A)$, where $tr(A)$ is the trace of A, the sum of the diagonals of A.

The coefficient of λ^k in $\chi_A(\lambda)$ is $(-1)^{(n-k)}$ times the sum of all the $(n-k)$-rowed principal minors of A.

The sum of the characteristic roots of A is $tr(A)$.

The product of the characteristic roots of A is $|A|$.

If A is a (upper or lower) triangular matrix, then $\chi_A(\lambda) = \prod_{i=1}^{n}(\lambda - a_{ii})$ and the characteristic roots of A are the diagonal entries of A.

Since $\lambda I - A^T = (\lambda I - A)^T$, characteristic polynomials of A and A^T are the same.

Theorem

Similar matrices have the same characteristic polynomial.

Theorem

Let A and B be matrices of orders $m \times n$ and $n \times m$ respectively, where $m \leq n$. Then $\chi_{BA}(\lambda) = \lambda^{n-m}\chi_{AB}(\lambda)$.
For any two $n \times n$ matrices A and B, the characteristic polynomials of AB and BA are the same.
If AB is square, the non-zero characteristic roots of AB are the same as those of BA.

Definition

A complex number α is said to be an **eigenvalue** *of A if there exists a non-null vector $x \in \mathbb{C}^n$ such that $Ax = \alpha x$.*
For such an x, called as an **eigenvector corresponding to the eigenvalue** α, *the line $\{\beta x : \beta \in \mathbb{C}\}$ is invariant under the map $x \mapsto Ax$.*
By an eigenvector of A we mean an eigenvector of A corresponding to some eigenvalue of A.

Theorem

A number α is an eigenvalue of A iff α is a characteristic root of A.

The preceding theorem shows that eigenvalues are the same as characteristic roots. However, **by 'the characteristic roots of A' we mean the n roots of the characteristic polynomial of A whereas 'the eigenvalues of A' would mean the distinct characteristic roots of A.**

Eigenvalues are also known as proper values, latent roots etc. and eigenvectors are also called the characteristic vectors, latent vectors etc.

Theorem

Let $f(\lambda)$ be a polynomial and β an eigenvalue of A. Then $f(\beta)$ is an eigenvalue of $f(A)$.
More generally, it follows from the preceding theorem that if β is an eigenvalue of a matrix A and $f(\lambda)$ is any polynomial such that $f(A) = 0$, then $f(\beta) = 0$.

Each eigenvalue of an idempotent matrix A is 0 or 1.

Definition

Let α be an eigenvalue of A. Then the subspace of $\mathbb{C}^n$ consisting of all eigenvectors of A corresponding to α together with 0 is called the eigenspace of A corresponding to α and is denoted by $ES(A, \alpha)$.
The dimension of $ES(A, \alpha)$ is called the **geometric multiplicity** *of α with respect to A.*

Note that $ES(A, 0) = N(A)$, and $ES(A, \alpha) \subseteq \mathcal{C}(A)$ if $\alpha \neq 0$. Clearly the geometric multiplicity of an eigenvalue α of A is at least 1.

Real matrix case

If A is a real matrix and α is a real eigenvalue of A then the system $(\alpha I - A)x = 0$ has a non-trivial solution over $\mathbb{R}$ and so there exists a real eigenvector of A corresponding to α. There will, of course, be non-real eigenvectors of A corresponding of α, for example : $\sqrt{-1}\,x$.

Suppose now α is a real eigenvalue of a real matrix A and S is the corresponding eigenspace. Then S has a basis consisting of real vectors which can even be chosen to be orthogonal. This is because the nullity of $\alpha I - A$ over $\mathbb{C}$ is the same as that over $\mathbb{R}$ and the Gram-Schmidt orthogonalization process gives a real orthonormal basis if we start with a real basis.

Definition

Let α be an eigenvalue of A. The number of times α appears as a root of the characteristic equation of A, is called the **algebraic multiplicity** *of α with respect to A.*

Theorem

For any eigenvalue α of A, the algebraic multiplicity of α with respect to A is at least the geometric multiplicity of α with respect to A.

Definition

An eigenvalue α of A is said to be **regular** *if the algebraic and the geometric multiplicities of α with respec to A are equal.*
α is said to be **simple** *eigenvalue of A if the algebraic multiplicity of α with respect to A is 1.*

If α is a simple eigenvalue of A, then α is regular and there exists a unique (upto multiplication by a non-zero scalar) eigenvector of A corresponding to α.

Theorem

Let $\alpha_1, \alpha_2, \ldots, \alpha_k$ be distinct eigenvalues of A and let $x_1, x_2, \ldots, x_k$ be corresponding eigenvectors. Then $x_1, x_2, \ldots, x_k$ are linearly independent.

If $S_1, S_2, \ldots, S_k$ are the eigenspaces corresponding to distinct eigenvalues $\alpha_1, \alpha_2, \ldots, \alpha_k$ of a matrix A, then $S_1 + S_2 + \cdots + S_k$ is direct.

We have seen that if AB is a square matrix then every non-zero eigenvalue of AB is also an eigenvalue of BA with the same algebraic multiplicity. The following result shows that the geometric multiplicity also remains the same.

Theorem

Let α be a non-zero eigenvalue of a square matrix AB, where A and B need not be square. Then α is an eigenvalue of BA with the same geometric multiplicity. If $x_1, x_2, \ldots, x_r$ are linearly independent eigenvectors of AB corresponding to α then $Bx_1, Bx_2, \ldots, Bx_r$ are linearly independent eigenvectors of BA corresponding to α.

Theorem

Let α be a non-zero eigenvalue of a square matrix AB, where A and B need not be square. If $x_1, x_2, \ldots, x_r$ are linearly independent eigenvectors of AB corresponding to α then $Bx_1, Bx_2, \ldots, Bx_r$ are linearly independent eigenvectors of BA corresponding to α.

The above theorem can be used effectively to find eigenvectors of BA when AB is of smaller order than BA, for example if (B, A) is a rank-factorization of a singular matrix.

Theorem

Let x be a non-null vector. Then there exists an eigenvector y of A belonging to the span of $\{x, Ax, A^2x, \ldots\}$.

Note that if A is a real matrix with real eigenvalues and if x is real, then the y obtained is real.

Theorem

Every $n \times n$ matrix A is similar to an upper triangular matrix over $\mathbb{C}$.

The preceding theorem does not hold over $\mathbb{R}$ since a real matrix may not have real eigenvalues

We have proved that if $f(\lambda)$ is a polynomial and β is an eigenvalue of A, then $f(\beta)$ is an eigenvalue of $f(A)$. The preceding theorem is a powerful tool. The following result is a stronger form of the above statement.

Theorem

Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the characteristic roots of A and let $f(\lambda)$ be a polynomial. Then $f(\lambda_1), f(\lambda_2), \ldots, f(\lambda_n)$ are the characteristic roots of $f(A)$.

If A is singular, the algebraic multiplicities of 0 with respect to A^ℓ and with respect to A are equal for any positive integer ℓ.

A polynomial $f(\lambda)$ is said to **annihilate** a matrix A if $f(A) = 0$. The following result is used to evaluate large powers of A or the value of a polynomial with large degree in A.

Theorem (Cayley-Hamilton Theorem)

For every matrix A, the characteristic polynomial of A annihilates A. That is, every matrix satisfies its own characteristic equation.

For any square matrix A, $I, A, A^2, \ldots, A^{n^2}$ are linearly dependent in $F^{n \times n}$, there exists a non-zero annihilating polynomial.

If f annihilates A, αf also annihilates A, so there exists a monic polynomial annihilating A.

Suppose k is the minimum degree of a non-zero polynomial annihilating A and f and g are to monic polynomials of degree k annihilating A. Then $h = f - g$ also annihilates A and has degree less than k, so $h = 0$ and $f = g$.

The monic polynomial of the least degree which annihilates A is called the **minimum polynomial** of A.

By Cayley-Hamilton theorem the degree of the minimal polynomial of an $n \times n$ matrix A is at most n.

We now show that the minimal polynomial not only has the least degree among the non-zero polynomials annihilating A but also divides each of them.

Theorem

The minimal polynomial of A divides every polynomial which annihilates A. The minimal polynomial of A divides the characteristic polynomial of A.

The preceding theorem shows that once an annihilating polynomial $g(\lambda)$ is known, the search for the minimal polynomial can be restricted to the factors of $g(\lambda)$.

Example

If A is idempotent then $\lambda^2 - \lambda$ annihilates A, so the minimal polynomial of A is $\lambda, \lambda - 1$ or $\lambda^2 - \lambda$. If A is neither 0 nor I it follows that the minimal polynomial of A is $\lambda^2 - \lambda$.

Theorem

A complex number α is a root of the minimal polynomial of A iff α is a characteristic root of A.

The preceding theorem shows that the distinct roots of the minimal polynomial coincide with those of the characteristic polynomial.

If the n characteristic roots of A are distinct, then the minimal polynomial of A coincides with the characteristic polynomial of A. A matrix A with the latter property is said to be **non-derogatory**.

Theorem

The minimal polynomial of a diagonal matrix A is $\prod_{i=1}^{k}(\lambda - d_i)$ where $d_1, d_2, \ldots, d_k$ are the distinct diagonal entries of A.

Caution. The minimal polynomial of a matrix need not be a product of distinct linear factors.

Theorem

Similar matrices have the same minimal polynomial.

We have discussed that every matrix is similar to an upper triangular matrix. But not every matrix is similar to a diagonal matrix.

For example, if $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ is similar to a diagonal matrix D, then both the characteristic roots of D are 0 and so $D = 0$, an impossibility.

We give some necessary and sufficient conditions for a matrix to be similar to a diagonal matrix and study some nice representation of such matrices.

Definition

A matrix is said to be **semi-simple** *or* **diagonable** *if it is similar to a diagonal matrix.*

If A is the matrix of a linear operator ψ on V with respect to some basis, then A is semi-simple iff there is a coordinate system (with the same origin) each of whose coordinate axes is left invariant by ψ.

If A has n linearly independent eigenvectors and P is the matrix formed with these vectors as the columns, then $P^{-1}AP$ is diagonal.

Conversely, if A is similar to a diagonal matrix (A is semi-simple), there there exists an invertible matrix P such that
$P^{-1}AP = D := \text{diag}(d_1, d_2, \ldots, d_n)$.

Then $AP = PD$, so $AP_{*j} = d_j P_{*j}$. Thus the columns of P are linearly independent eigenvectors of A (corresponding to the diagonal entries of D in the same order).

In the following theorem we give some more characteristions of a semi-simple matrix.

Theorem

The following statements about an $n \times n$ matrix A are equivalent :

1. *A is semi-simple,*
2. *the minimal polynomial of A is a product of distinct linear factors or, equivalently, there exists an annihilating polynomial of A which is a product of distinct linear factors,*
3. *all the eigenvalues of A are regular,*
4. *the sum of the eigenspaces of A is $\mathbb{C}^n$,*
5. *A has n linearly independent eigenvectors.*

Theorem (Sufficient conditions)

An $n \times n$ matrix with n distinct eigenvalues is semi-simple.
An idempotent matrix is semi-simple.

In the next theorem we give several useful ways of representing a semi-simple matrix.

Theorem

The following statements about an $n \times n$ matrix A are equivalent :

1. *A is semi-simple and has rank r.*
2. *There exists a non-singular matrix P of order n and a diagonal non-singular matrix D of order r such that $A = P \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} P^{-1}$.*
3. *There exists non-zero scalars $d_1, d_2, \ldots, d_r$ and vectors $u_1, u_2, \ldots, u_r$ and $v_1, v_2, \ldots, v_r \in \mathbb{C}^n$ such that $v_i^T u_j = \delta_{ij}$ for all i, j and $A = \sum_{i=1}^{r} d_i u_i v_i^T$.*
4. *There exist matrices R, S and D of orders $n \times r, r \times n$ and $r \times r$ respectively such that D is diagonal and non-singular, $SR = I$ and $A = RDS$.*

We call each of the representation in the above theorem, a **spectral decomposition** of the semi-simple matrix A.

We next study another representation of a semi-simple matrix which is unique unlike spectral decomposition.

Theorem

Let A be an $n \times n$ matrix. Then A is semi-simple iff for some k $(1 \leq k \leq n)$ there exist (complex) numbers $\alpha_1, \alpha_2, \ldots, \alpha_k$ and $n \times n$ matrices $E_1, E_2, \ldots, E_k$ such that the following four conditions are satisfied:

1. $A = \alpha_1 E_1 + \alpha_2 E_2 + \cdots + \alpha_k E_k.$
2. $\alpha_1, \alpha_2, \ldots, \alpha_k$ *are distinct and* $E_1, E_2, \ldots, E_k$ *are non-null.*
3. $E_1 + E_2 + \cdots + E_k = I.$
4. $E_i^2 = E_i$ *for* $i = 1, 2, \ldots, k.$

Further, α's and E's are uniquely determined by A thus : $\alpha_1, \alpha_2, \ldots, \alpha_k$ are the distinct eigenvalues of A and if S_i denotes the eigen subspace of A corresponding to α_i, then E_i is the projector into S_i along $\sum_{j \neq i} S_j$.

In the preceding theorem, $E_i^2 = E_i$ for $i = 1, 2, \ldots, k$ can be replaced by any one of the following :

1. $E_i E_j = 0$ whenever $i \neq j$.
2. $\rho(E_1) + \rho(E_2) + \cdots + \rho(E_k) = n$.

The unique representation of a semi-simple matrix A in the form

$$A = \alpha_1 E_1 + \alpha_2 E_2 + \cdots + \alpha_k E_k$$

where $E_1 + E_2 + \cdots + E_k = I$, $E_i^2 = E_i$ for $i = 1, 2, \ldots, k$ and $\alpha_1, \alpha_2, \ldots, \alpha_k$ are distinct and $E_1, E_2, \ldots, E_k$ are non-null, is called the **spectral form** or **spectral representation** of A.

Exercises

- If $\lambda_1, \lambda_2, \dots, \lambda_r$ are the eigenvalues of a matrix A, then prove the following:
  (a) $k\lambda_1, k\lambda_2, \dots, k\lambda_r$ are the eigenvalues of the matrix kA, where k is a non-zero scalar.
  (b) $\lambda_1^p, \lambda_2^p, \dots, \lambda_r^p$ are the eigenvalues of the matrix A^p, where p is any positive integer.
  (c) $1/\lambda_1, 1/\lambda_2, \dots, 1/\lambda_r$ are the eigenvalues of the inverse of A, provided $|A| \neq 0$.
  (d) $|A|/\lambda_1, |A|/\lambda_2, \dots, |A|/\lambda_r$ are the eigenvalues of the adjoint of A.
  (e) $k+\lambda_1, k+\lambda_2, \dots, k+\lambda_r$ are the eigenvalues of the matrix $A+kI$, where k is any scalar.
- Prove that the eigenvalues of a real symmetric matrix are real.
- Prove that the eigenvectors corresponding to a distinct eigenvalues of a real symmetric matrix are orthogonal.

- Prove that the eigenvectors corresponding to distinct eigenvalues of a matrix are linearly independent.
- Prove that an eigenvector cannot correspond to two different eigenvalues.
- Verify Cayley-Hamilton theorem for $\begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}$ and hence evaluate A^3 and A^{-1}.
- Evaluate $A^8 - A^7 + 5A^6 - A^5 + A^4 - A^3 + 6A^2 + A - 2I$ if

$$A = \begin{pmatrix} 1 & 2 & -2 \\ 2 & 5 & -4 \\ 3 & 7 & -5 \end{pmatrix}.$$

- For any positive integer, evaluate A^n, where $\begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}$.

- Diagonalize the matrix $A = \begin{pmatrix} 11 & -4 & -7 \\ 7 & -2 & -5 \\ 10 & -4 & -6 \end{pmatrix}$ and hence find A^3.
- Reduce the matrix $A = \begin{pmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{pmatrix}$ to a diagonal form by orthogonal reduction.

Reference

- A. Ramachandra Rao and P. Bhimasankaram, "*Linear Algebra*", Hindustan Book Agency, 2000.