Eigenvalues and Eigenvectors

P. Sam Johnson

May 26, 2017

We study eigenvalues and eigenvectors associated with a **complex squrare 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, Engineering, Economics and Statistics.

For any $n \times n$ matrix A, consider the polynomial

$$\chi_A(\lambda) := |\lambda I - A| = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix}. \qquad (1)$$

Clearly this is a monic polynomial of degree n.

By the fundamental theorem of algebra, $\chi(A)$ has exactly n (not necessarily distinct) roots.

$\chi_A(\lambda)$	the **characteristic polynomial** of A
$\chi_A(\lambda) = 0$	the **characteristic equation** of A
the roots of $\chi_A(\lambda)$	the **characteristic roots** of A
distinct roots of $\chi_A(\lambda)$	the **spectrum** of A

- The constant terms and the coefficient of λ^{n-1} in $\chi_A(\lambda)$ are $(-1)^n|A|$ and $tr(A)$.
- The sum of the characteristic roots of A is $tr(A)$ and the product of the characteristic roots of A is $|A|$.
- Since $\lambda I - A^T = (\lambda I - A)^T$, characteristic polynomials of A and A^T are the same.
- Since $\lambda I - P^{-1}AP = P^{-1}(\lambda I - A)P$, similar matrices have the same characteristic polynomials.

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

Just like determinant, characteristic polynomial canbe defined for a linear operator ϕ on a vector space V as the characteristic polynomial of the matrix of ϕ with respect to any basis of V.

Suppose A and B are matrices of a linear operator ϕ with respect to bases B_1 and B_2 of V respectively. Then $\chi_A(\lambda) = \chi_B(\lambda)$.

Theorem

Let A and B be matrices of orders $m \times n$ and $n \times m$ respectively, where $m \leq n$. Then $\chi_{BA}(\lambda) = \lambda^{n-m}\chi_{AB}(\lambda)$.

Proof. Let $r = rank(A)$. There exist non-singular matrices P and Q such that

$$PAQ = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} \text{ and } Q^{-1}BP^{-1} = \begin{bmatrix} C & D \\ E & G \end{bmatrix},$$

where C is of order $r \times r$. Then

$$PABP^{-1} = \begin{bmatrix} C & D \\ 0 & 0 \end{bmatrix} \text{ and } Q^{-1}BAQ^{-1} = \begin{bmatrix} C & 0 \\ E & 0 \end{bmatrix}.$$

Hence

$$\chi_{AB}(\lambda) = \chi_{PABP^{-1}}(\lambda) = \begin{vmatrix} \lambda I_r - C & -D \\ 0 & \lambda I_{m-r} \end{vmatrix} = |\lambda I_r - C|\lambda^{m-r}$$

and

$$\chi_{BA}(\lambda) = \chi_{QBAQ^{-1}}(\lambda) = \begin{vmatrix} \lambda I_r - C & 0 \\ -E & \lambda I_{n-r} \end{vmatrix} = |\lambda I_r - C|\lambda^{n-r}.$$

Thus $\chi_{BA}(\lambda) = \lambda^{n-m}\chi_{AB}(\lambda)$.

- For any two $n \times n$ matrices A and B, the characteristic polynomials of AB and BA are the same.
- If AB is not square, the non-zero characteristic roots of AB are the same as those of BA.

Definition

A complex number α is an **eigenvalue** *of A if there exists $x \neq 0$ in $\mathbb{C}^n$ such that $Ax = \alpha x$. Any such (non-null) x is an* **eigenvector** *of A corresponding to the eigenvalue α.*

When we say that x is an eigenvector of A we mean that x is an eigenvector of A corresponding to some eigenvalue of A.

Two observations:

- α is an eigenvalue of A iff the system $(\alpha I - A)x = 0$ has a non-trivial solution.
- α is a characteristic root of A iff $\alpha I - A$ is singular.

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.

Equivalent names:

Eigenvalues	proper values, latent roots, etc.
Eigenvectors	characteristic vectors, latent vectors, etc.

Theorem

Let β an eigenvalue of A and $f(\lambda)$ be a polynomial. Then $f(\beta)$ is an eigenvalue of $f(A)$.

Proof. Let x be an eigenvector of A corresponding to β. Then $Ax = \beta x$. Premultiplying by A, we get $A^2x = \beta^2 x$. Proceeding like this we get $A^k x = \beta^k x$ for all $k \geq 0$, so $f(A)x = f(\beta)x$. Since $x \neq 0$, $f(\beta)$ is an eigenvalue of $f(A)$.

Theorem

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

Proof. Let β an eigenvalue of A and let $f(\lambda) = \lambda^2 - \lambda$. Then $f(A) = A^2 - A = 0$. By previous theorem, $f(\beta) = 0$. Hence β is 0 or 1.

More generally, if β is an eigenvalue of a matrix A and $f(\lambda)$ is any polynomial such that $f(A) = 0$, then $f(\lambda) = 0$.

If α is an eigenvalue of A, the set of all eigenvectors of A corresponding to α, together with 0, forms $N(\alpha I - A)$, called the **eigen space** of A corresponding to α and is denoted by $ES(A, \alpha)$.

$dim[ES(A, \alpha)]$ is called the **geometric multiplicity** of α with respect to A. Note that $ES(A, 0) = N(A)$ and $ES(A, \alpha) \subseteq C(A)$ if $\alpha \neq 0$.

Another type of multiplicity of an eigenvalue α of A:

The number of times α appears as a root of the characteristic equation of A. This is called the **algebraic multiplicity** of α with respect to A.

Relation between the two multiplicities:

Let V be a vector space having dimension n.

- The sum of albraic multiplicities is equal to the dimension of V, n.
- If $\alpha_1, \alpha_2, \ldots, \alpha_k$ are the distinct eigenvalues of an $n \times n$ matrix A with geometric multiplicities $n_1, n_2, \ldots, n_k$ respectively, then $n_1 + \cdots + n_k \leq n$.

Theorem

For any eigenvalue α of A, the algebraic multiplicity of α with respect to A is not less than the geometric multiplicity of α with respect to A. That is, $sim[ES(A, \alpha)]$ is at most the algebraic multiplicity of α with respect to A. (or) The algebraic multiplicity of α with respect to A is at least $sim[ES(A, \alpha)]$.

Proof of the theorem

Let $\{x_1, x_2, \ldots, x_k\}$ be a basis of $ES(A, \alpha)$ and $\{x_1, x_2, \ldots, x_n\}$ an extension to a basis of $\mathbb{C}^n$. Then $P := [x_1 : x_2 : \cdots : x_n]$ is non-singular and

$$\begin{aligned} P^{-1}AP &= P^{-1}[Ax_1 : Ax_2 : \cdots : Ax_n] \\ &= P^{-1}[\alpha x_1 : \alpha x_2 : \cdots : \alpha x_n : Ax_{k+1} : \cdots : Ax_n]. \end{aligned}$$

Since for each $j = 1, 2, \ldots, k$, $P^{-1}(\alpha x_j) = \alpha P^{-1} P_{*j} = \alpha e_j$.

$P^{-1}AP = \begin{bmatrix} \alpha I_k & B \\ 0 & C \end{bmatrix}$ for some matrices B and C.

Hence $\chi_A(\lambda) = \chi_{P^{-1}AP}(\lambda) = (\lambda - \alpha)^k \chi_C(\lambda)$.

Thus the number of times α appears as a root of the characteristic equation of A is at least $k = dim[ES(A, \alpha)]$.

Let α be an eigenvalue of A.

α is **regular**	the algebraic and the geometric multiplicities of α with respect to A are equal
α is **simple**	the algebraic multiplicity of α with respect to A is 1

Note that every simple eigenvalue is regular.

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.

Corollary

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

We have seen that if AB is a square matrix then every nonzero eigenvalue of AB is also an eigenvalue of BA with the same algebraic multiplicity.

We now show that the geometric multiplicity also remains the same.

Theorem

Let α be a nonzero 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.

Proof of the theorem

Note that $x \in ES(A, \alpha)$, then $ABx = \alpha x$. Hence $BABx = \alpha Bx$, so $Bx \in ES(A, \alpha)$. Similarly, if $x \in ES(A, \alpha)$, then $BAx = \alpha x$. Hence $ABAx = \alpha Ax$, so $Ax \in ES(A, \alpha)$.

Let $\{x_1, x_2, \ldots, x_r\}$ be a basis of $ES(A, \alpha)$. Then $\{Bx_1, Bx_2, \ldots, Bx_r\}$ be a basis of $ES(BA, \alpha)$.

Claim: $\{Bx_1, Bx_2, \ldots, Bx_r\}$ is a linearly independent set. Suppose $\sum_{i=1}^{r} \beta_i Bx_i = 0$ for all $i = 1, 2, \ldots, r$. Then $\{Bx_1, Bx_2, \ldots, Bx_r\}$ is a linearly independent set. Hence $dim[ES(BA, \alpha)] \geq r = dim[ES(A, \alpha)]$.

Thus geometric multiplicity of α with respect to $BA \geq$ geometric multiplicity of α with respect to AB.

By symmetry the reverse inequality holds and equality follows.

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 vectors. Then there exists an eigenvector y of A belonging to the span of $\{x, Ax, A^x, \ldots\}$.

Theorem

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

Proof. We prove by induction on n. If $n = 1$, the result holds trivially. So assume it for matrices for order $n - 1$. Let A be of order n. Let α be an eigenvalue of A; x be an eigenvector of A corresponding to α, and P be a non-singular matrix with x as the first column.

Then $P^{-1}AP = \begin{bmatrix} \alpha & y^T \\ 0 & C \end{bmatrix}$, for some $y \in \mathbb{C}^{n-1}$ and $C \in \mathbb{C}^{n-1} \times C^{n-1}$.

By induction hypothesis, there exists a non-singular matrix W of order $n-1$ such that $T := W^{-1}CW$ is upper triangular.

$Q := \begin{bmatrix} 1 & 0 \\ 0 & W \end{bmatrix}$ is non-singular, so PQ is non-singular, and

$$(PQ)^{-1}A(PQ) = \begin{bmatrix} 1 & 0 \\ 0 & W^{-1} \end{bmatrix} = \begin{bmatrix} \alpha & y^T \\ 0 & C \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & W \end{bmatrix} = \begin{bmatrix} \alpha & y^TW \\ 0 & T \end{bmatrix}$$

is upper triangular.

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

Theorem

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

Proof. As any matrix is similar to a diagonal matrix, there exists a non-singular matrix P such that $T := P^{-1}AP$ is upper triangular. Since A and T have the characteristic roots, we may take $t_{ii} = \lambda_i$, for $i = 1, 2, \ldots, n$.

By induction on k, we have $T^k := P^{-1}A^kP$, for all $k \geq 0$. if $f(\lambda) = a_0 + a_1\lambda + \cdots + a_s\lambda_s$, we have

$$\begin{aligned} f(T) &= a_0 I + a_1 T + \cdots + a_s T^s \\ &= a_0 P^{-1}P + a_1 P^{-1}AP + \cdots + a_s P^{-1}A^sP \\ &= P^{-1}(a_0 I + a_1 T + \cdots + a_s T^s)P \\ &= P^{-1}f(A)P. \end{aligned}$$

Hence $f(T)$ is upper triangular with $f(t_{11}, t_{22}, \ldots, t_{nn}$ as the diagonal entries, hence the characteristic roots of $f(A)$ are $f(\lambda_1), f(\lambda_2), \ldots, f(\lambda_n)$.

Corollary

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 $f(A)$ is said to **annihilate** A if $F(A) = 0$. If f annihilates A, αf also annihilates A.

For any squae matrix A, there exists a non-zero annihilating polynomial. This also follows from the fact that $I, A, \ldots, A^{n^2}$ are linearly dependent in $F^{n \times n}$.

Does there exist a monic polynomial annihilating A? The answer is affirmative by the following theorem.

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

Simple proof? We have $\chi_A(\lambda) = |\lambda I - A|$. Replace λ by A, shall we get the Cayley - Hamilton theorem.

Two main uses of Cayley-Hamilton theorem

1. To evaluate large powers A.
2. To evaluate a polynomial in A with large degree even if A is singular.
3. To express A^{-1} as a polynomial in A whereas A is non-singular.

Definition

A monic polynomial of the least degree which annihilates A is called a **minimal polynomial** *of A, denoted by $m(\lambda)$.*

Minimal polynomial of A is unique. Suppose k is the minimum degree of a nonzero polynomial annihilating A and f & g are two monic polynomials of degree k annihilating A.

Then $h = f - g$ also annihilates A and has degree less than k, so $h = 0$ and $f = g$.

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

Theorem

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

Proof. Let $f(\lambda)$ be the minimal polynomial of A and let $g(A) = 0$. Since $f \neq 0$, there exist polynomials $q(\lambda)$ and $r(\lambda)$ such that $g(\lambda) = f(\lambda)a(\lambda) + r(\lambda)$ where $deg(r(\lambda)) < deg(f(\lambda))$.

Then $0 = g(A) = f(A)q(A) + r(A) = r(A)$. Thus $r(\lambda)$ annihilates A. By the minimality of f, $r(\lambda) = 0$, so f divides g.

Thus the minimal polynomial not only has the least degree among the nonzero polynomials annihilating A but also divides each of them.

The minimal polynomial of A divides the characteristic polynomial of A.

How to find the minimal polynomial?

1. Once an annihilating polynomial $g(\lambda)$ is known, the search for the minimal polynomial can be restricted to the factors of $g(\lambda)$.
2. If A is idempotent, then $\lambda^2 - \lambda$ annihilates A, so the minimal polynomial of A is $\lambda, \lambda - 1$, or $\lambda^2 - \lambda$.
3. If A is neither 0 or I, 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.

Proof. α is a root of the minimal polynomial, $m_A(\lambda)$ of A.

Then $m_A(\alpha) = 0$, hence $\chi_A(\alpha) = m_A(\alpha)g(\alpha)$. Thus α is a characteristic root of A.

Converse?

1. The distinct roots of the minimal polynomial coincides with those of the characteristic polynomial.
2. The minimal polynomial of A coincides with the characteristic polynomial of A if A has n distinct characteristic roots. A matrix A with the property is said to be **non-derogatory**.
3. The minimal polynomial of a matrix need not be a product of distinct linear factors.
4. 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 entries of A.

Theorem

Similar matrices have the same minimal polynomial.

Proof. Let $B = P^{-1}AP$. Then $B^k := P^{-1}A^kP$, for all $k \geq 0$ and $f(B) = P^{-1}f(A)P$ for any polynomial f. Thus $f(B) = 0 \iff f(A) = 0$, so A and B have the same minimal polynomial.

$\therefore$ We can define the minimal polynomial of a linear operator ϕ on a vector space V as the minimal polynomial of the matrix of ϕ with respect to any basis of V.

If f is any polynomial and A is the matrix of ϕ with respect to a basis B, then $f(A)$ is the matrix of $f(\phi)$ with respect to B. Thus $f(A) = 0 \iff f(\phi) = 0$, and the minimal polynomial of ϕ is the monic polynomial of the least degree which annihilates ϕ.

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

Example

Suppose $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ *is similar to a diagonal matrix* D*. Since* $\chi_A(\lambda) = \chi_D(\lambda)$*, both the characteristic roots of* D *are* 0*. Thus* $D = 0$*, which is impossible.*

Definition

A matrix is **semi-simple** *or* **diagonalable** *if it is similar to a diagonal matrix.*

Let A be the matrix of a linear operator ϕ on V with respect to some basis.

A is semisimple $\iff$ there is a coordinate system (with the same origin) each of whose coordinate axes is left invariant by ϕ.

Suppose A is semisimple and $P^{-1}AP = D := diag(d_1, d_2, \dots, d_n)$. 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).

Conversely, 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.

Let A be an $n \times n$ matrix. TFAE

1. A is semisimple,
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 eigenvalues of A are regular,
4. the sum of the eigenspaces of A is $\mathbb{C}^n$,
5. A has n linearly independent eigenvectors.

- An $n \times n$ matrix with n distinct eigenvalues is semisimple (because if all the characteristic roots of A are distinct, then each is simple and so regular).
- An idempotent matrix is semisimple because $\lambda(\lambda - 1)$ annihilates an idempotent matrix.

Let A be an $n \times n$ matrix. TFAE.

1. A is semisimple and has rank r.
2. There exists a nonsingular matrix P of order n and a diagonal nonsingular matrix Δ of order r such that $A = P \begin{bmatrix} \Delta & 0 \\ 0 & 0 \end{bmatrix} P^{-1}$.
3. There exist nonzero scalars $\delta_1, \delta_2, \ldots, \delta_n$ and vectors $u_1, u_2, \ldots, u_n$ and $v_1, v_2, \ldots, v_n \in \mathbb{C}^n$ such that $v_i^T u_j = \delta_{ij}$ for all i, j and $A = \sum_{i=1}^n \delta_i u_i v_i^T$.
4. There exist matrices R, S and Δ of orders $n \times r, r \times n$ and $r \times r$ respectively such that D is diagonal and nonsingular, $SR = I$ and $A = R\Delta S$.