

Eigenvalues, Eigenvectors, and Matrix Decompositions

$\mathbf{A} \Rightarrow q(\lambda) = |\mathbf{A} - \lambda \mathbf{I}|$ is a polynomial of degree n .

For example,

$$\begin{aligned} \mathbf{A} \Rightarrow |\mathbf{A} - \lambda \mathbf{I}| &= \begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} \\ &= (a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21} \\ &= \lambda^2 + (-a_{11} - a_{22})\lambda + a_{11}a_{22} - a_{12}a_{21}. \end{aligned}$$

The definition of $|A - \lambda I|$ implies that the coefficient on λ^n is $(-1)^n$.

$$\begin{vmatrix} a_{11} - \lambda & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} - \lambda & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} - \lambda \end{vmatrix}$$

The n (not necessarily distinct) roots of $q(\lambda)$ (i.e., solutions to $q(\lambda) = 0$) are called the eigenvalues of A .

The Fundamental Theorem of Algebra guarantees n roots, and the Polynomial Factorization Theorem implies that

$$q(\lambda) = |\mathbf{A} - \lambda \mathbf{I}| = \prod_{i=1}^n (\lambda_i - \lambda)$$

for the eigenvalues $\lambda_1, \dots, \lambda_n$.

Example:

Suppose $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, then

$$\begin{aligned} |A - \lambda I| &= \begin{vmatrix} 2 - \lambda & 1 \\ 1 & 2 - \lambda \end{vmatrix} \\ &= (2 - \lambda)(2 - \lambda) - 1 \\ &= \lambda^2 - 4\lambda + 3 \\ &= (3 - \lambda)(1 - \lambda). \end{aligned}$$

Eigenvalues are

$$\lambda_1 = 3, \lambda_2 = 1 \because (3 - \lambda)(1 - \lambda) = 0 \iff \lambda = 3 \text{ or } \lambda = 1.$$

Example:

Suppose $A = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$, then

$$\begin{aligned} |A - \lambda I| &= \begin{vmatrix} 3 - \lambda & 0 \\ 0 & 3 - \lambda \end{vmatrix} \\ &= (3 - \lambda)(3 - \lambda) \\ &= \lambda^2 - 6\lambda + 9. \end{aligned}$$

Eigenvalues are $\lambda_1 = 3, \lambda_2 = 3 \because (3 - \lambda)(3 - \lambda) = 0 \iff \lambda = 3$.

In the previous example, the eigenvalue 3 is said to have algebraic multiplicity 2.

The algebraic multiplicity of eigenvalue λ^* is

$$\sum_{i=1}^n \mathbb{I}[\lambda_i = \lambda^*].$$

Note that the roots of $q(\lambda)$ are not necessarily real numbers.

For example,

$$\begin{aligned}\mathbf{A} &= \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \Rightarrow \mathbf{A} - \lambda \mathbf{I} = \begin{bmatrix} -\lambda & 1 \\ -1 & -\lambda \end{bmatrix} \\ &\Rightarrow |\mathbf{A} - \lambda \mathbf{I}| = \lambda^2 + 1 \\ &= (\sqrt{-1} - \lambda)(-\sqrt{-1} - \lambda) \\ \therefore \lambda_1 &= \sqrt{-1}, \lambda_2 = -\sqrt{-1}.\end{aligned}$$

If \mathbf{A} is symmetric, then all roots of $q(\lambda) = |\mathbf{A} - \lambda \mathbf{I}|$ are real numbers.

λ is an eigenvalue of A

$$\iff |\mathbf{A} - \lambda \mathbf{I}| = 0$$

$\iff \mathbf{A} - \lambda \mathbf{I}$ is singular

$$\iff \text{rank}(\mathbf{A} - \lambda \mathbf{I}) < n$$

$$\iff \exists \mathbf{x} \neq \mathbf{0} \ni (\mathbf{A} - \lambda \mathbf{I})\mathbf{x} = \mathbf{0}$$

$$\iff \exists \mathbf{x} \neq \mathbf{0} \ni \mathbf{A}\mathbf{x} - \lambda \mathbf{x} = \mathbf{0}$$

$$\iff \exists \mathbf{x} \neq \mathbf{0} \ni \mathbf{A}\mathbf{x} = \lambda \mathbf{x}.$$

Such a vector $x \neq \mathbf{0}$ satisfying $Ax = \lambda x$ is called an eigenvector of A _{$n \times n$} corresponding to eigenvalue λ .

Note that if \mathbf{x} is an eigenvector corresponding to eigenvalue λ , then $c\mathbf{x}$ is also an eigenvector $\forall c \in \mathbb{R}$.

$$\because \mathbf{A}\mathbf{x} = \lambda\mathbf{x} \Rightarrow c\mathbf{A}\mathbf{x} = c\lambda\mathbf{x} \Rightarrow \mathbf{A}(c\mathbf{x}) = \lambda(c\mathbf{x}).$$

It is customary to take eigenvalues to be unit norm, i.e., we take

$$c = \frac{1}{\|\mathbf{x}\|}$$

so that the eigenvector $c\mathbf{x}$ satisfies

$$\|c\mathbf{x}\| = 1.$$

If λ is an eigenvalue of \mathbf{A} , the set of vectors satisfying $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ is called the eigenspace of the eigenvalue λ .

Note that the eigenspace

$$\{\mathbf{x} : \mathbf{A}\mathbf{x} = \lambda\mathbf{x}\} = \{\mathbf{x} : (\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}\} = \mathcal{N}(\mathbf{A} - \lambda\mathbf{I}).$$

The dimension of the eigenspace of λ is known as the geometric multiplicity of λ .

For symmetric matrices, it turns out that algebraic multiplicity of an eigenvalue is the same as the geometric multiplicity.

Show that eigenvectors corresponding to distinct eigenvalues of a symmetric matrix A are orthogonal.

$n \times n$

Fact V7:

If $\mathcal{S} \subseteq \mathbb{R}^n$ is a vector space of dimension $k > 0$, then there exist a set of mutually orthonormal vectors that forms a basis for \mathcal{S} .

Proof: Homework problem.

It follows that for each distinct eigenvalue λ of a matrix \mathbf{A} satisfying $\mathbf{A} = \mathbf{A}'$, there exists a set of orthonormal eigenvectors with $\dim(\mathcal{N}(\mathbf{A} - \lambda\mathbf{I}))$ elements.

These orthonormal eigenvectors form a basis for the eigenspace of λ .

Furthermore, the collection of all such eigenvectors corresponding to the eigenvalues of A is a mutually orthonormal set of n vectors.

Let $\mathbf{q}_1, \dots, \mathbf{q}_n$ denote these n orthonormal eigenvectors corresponding to eigenvalues $\lambda_1, \dots, \lambda_n$ respectively.

Let $\mathbf{Q} = [\mathbf{q}_1, \dots, \mathbf{q}_n]$.

Let $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_n)$.

$\because \mathbf{q}_1, \dots, \mathbf{q}_n$ are mutually orthonormal, $\mathbf{Q}'\mathbf{Q} = \mathbf{I}$.

Moreover, because \mathbf{Q} is $n \times n$, it follows that \mathbf{Q}' is \mathbf{Q}^{-1} .

Thus,

$$\mathbf{Q}\mathbf{Q}' = \mathbf{Q}'\mathbf{Q} = \mathbf{I}.$$

Note that

$$\mathbf{A}\mathbf{q}_i = \lambda_i\mathbf{q}_i \quad \forall i = 1, \dots, n$$

$$\iff \mathbf{A}\mathbf{Q} = \mathbf{Q}\mathbf{\Lambda}$$

$$\iff \mathbf{A}\mathbf{Q}\mathbf{Q}' = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}'$$

$$\iff \mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}' = \sum_{i=1}^n \lambda_i\mathbf{q}_i\mathbf{q}_i'$$

This result is known as the Spectral Decomposition Theorem:

If \mathbf{A} is a symmetric matrix, then $\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}'$, where \mathbf{Q} is an orthogonal matrix and $\mathbf{\Lambda}$ is a diagonal matrix.

Result A.19:

Let A be a symmetric matrix. Then $rank(A)$ is equal to the number of nonzero eigenvalues of A .

Proof: Homework problem.

Suppose \mathbf{A} has eigenvalues $\lambda_1, \dots, \lambda_n$. Then

(a) $\text{trace}(\mathbf{A}) = \sum_{i=1}^n \lambda_i$, and

(b) $|\mathbf{A}| = \prod_{i=1}^n \lambda_i$.

Proof: Homework problem.

Quadratic Forms

Suppose $\mathbf{A} = [a_{ij}]$ and $\mathbf{x} = [x_1, \dots, x_n]'$.

$\mathbf{x}'\mathbf{A}\mathbf{x} = \sum_{i=1}^n \sum_{j=1}^n x_i x_j a_{ij}$ is known as a quadratic form.

Nonnegative Definite and Positive Definite Matrices

A symmetric matrix A is nonnegative definite (NND) if and only if

$$\mathbf{x}'\mathbf{A}\mathbf{x} \geq 0 \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

A symmetric matrix A is positive definite (PD) if and only if

$$\mathbf{x}'\mathbf{A}\mathbf{x} > 0 \quad \forall \mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}.$$

Suppose A is a symmetric matrix.

$n \times n$

Prove that A is NND if and only if all eigenvalues of A are nonnegative.

Prove that A is PD if and only if all eigenvalues of A are positive.

Suppose \mathbf{A} is a symmetric NND matrix.

$n \times n$

Show there exists a symmetric matrix \mathbf{B} such that $\mathbf{B}\mathbf{B} = \mathbf{A}$.

Such a matrix is called the symmetric square root of \mathbf{A} and is denoted by $\mathbf{A}^{1/2}$.

Result A.20 (Cholesky Decomposition):

Suppose \mathbf{A} is a symmetric matrix. \mathbf{A} is PD iff there exists a nonsingular lower triangular matrix \mathbf{L} such that

$$\mathbf{A} = \mathbf{L}\mathbf{L}'.$$

(\implies) If $n = 1$, $\mathbf{A} = [a]_{1 \times 1}$. Moreover,

$$\begin{aligned} \mathbf{A} \text{ is PD } &\Rightarrow x'ax > 0 \quad \forall x \neq 0 \\ &\Rightarrow ax^2 > 0 \quad \forall x \neq 0 \\ &\Rightarrow a > 0. \end{aligned}$$

If we take $\mathbf{L} = [\sqrt{a}]$, then $\mathbf{LL}' = \mathbf{A}$, \mathbf{L} is a nonsingular lower triangular matrix, and the result holds.

Now suppose the result holds for $n = 1, \dots, k$ for some integer $k \geq 1$.

Suppose \mathbf{A} is any $(k + 1) \times (k + 1)$ matrix. Partition \mathbf{A} as

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_k & \mathbf{a}_k \\ \mathbf{a}'_k & a \end{bmatrix}$$

A is PD $\Rightarrow A_k$ is PD \because

$$\mathbf{x}'\mathbf{A}\mathbf{x} > 0 \quad \forall \mathbf{x} \neq \mathbf{0}$$

$\Rightarrow \mathbf{x}'\mathbf{A}\mathbf{x} > 0 \quad \forall \mathbf{x} \neq \mathbf{0}$ with the last component 0

$$\Rightarrow [\mathbf{y}' \ 0]\mathbf{A} \begin{bmatrix} \mathbf{y} \\ 0 \end{bmatrix} > 0 \quad \forall \mathbf{y} \neq \mathbf{0}$$

$$\Rightarrow \mathbf{y}'\mathbf{A}_k\mathbf{y} > 0 \quad \forall \mathbf{y} \neq \mathbf{0}.$$

$\therefore \exists \mathbf{L}_k$ nonsingular and lower triangular $\ni \mathbf{A}_k = \mathbf{L}_k \mathbf{L}'_k$.

Now let $\mathbf{l}_k = \mathbf{L}_k^{-1} \mathbf{a}_k$.

$\therefore \mathbf{A}$ is PD,

$$\begin{bmatrix} \mathbf{l}'_k \mathbf{L}_k^{-1} & -1 \end{bmatrix} \begin{bmatrix} \mathbf{A}_k & \mathbf{a}_k \\ \mathbf{a}'_k & a \end{bmatrix} \begin{bmatrix} \mathbf{L}_k'^{-1} \mathbf{l}_k \\ -1 \end{bmatrix} > 0 \Rightarrow$$

$$\Rightarrow \mathbf{l}'_k \mathbf{L}_k^{-1} \mathbf{A}_k \mathbf{L}'_k{}^{-1} \mathbf{l}_k - 2\mathbf{l}'_k \mathbf{L}_k^{-1} \mathbf{a}_k + a > 0$$

$$\Rightarrow \mathbf{l}'_k \mathbf{L}_k^{-1} \mathbf{L}_k \mathbf{L}'_k{}^{-1} \mathbf{l}_k - 2\mathbf{l}'_k \mathbf{l}_k + a > 0$$

$$\Rightarrow -\mathbf{l}'_k \mathbf{l}_k + a > 0.$$

Now let $l = \sqrt{-\mathbf{l}'_k \mathbf{l}_k + a}$.

Then

$$\mathbf{L} = \begin{bmatrix} \mathbf{L}_k & \mathbf{0} \\ \mathbf{l}'_k & l \end{bmatrix}$$

is lower triangular and nonsingular ($\because l > 0 \Rightarrow L$ is full rank) and

$$\begin{aligned}
LL' &= \begin{bmatrix} \mathbf{L}_k & 0 \\ \mathbf{l}'_k & l \end{bmatrix} \begin{bmatrix} \mathbf{L}'_k & \mathbf{l}_k \\ 0 & l \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{L}_k \mathbf{L}'_k & \mathbf{L}_k \mathbf{l}_k \\ \mathbf{l}'_k \mathbf{L}'_k & \mathbf{l}'_k \mathbf{l}_k + l^2 \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{A}_k & \mathbf{L}_k \mathbf{L}_k^{-1} \mathbf{a}_k \\ \mathbf{a}'_k \mathbf{L}'_k^{-1} \mathbf{L}'_k & \mathbf{l}'_k \mathbf{l}_k + a - \mathbf{l}'_k \mathbf{l}_k \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{A}_k & \mathbf{a}_k \\ \mathbf{a}'_k & a \end{bmatrix} \\
&= \mathbf{A}.
\end{aligned}$$

