The Eigenvalue Problem

The Eigenvalue Problem

> *Def.* **The Eigenvalue Problem**
>
> For an $(n \times n)$ matrix A, find all scalars λ s.t. the equation:
>
> $$A\mathbf{x} = \lambda\mathbf{x} \tag{1}$$
>
> has a **nonzero** solution, $\mathbf{x}$. Such a scalar λ is called an **eigenvalue** of A, and any **nonzero** $(n \times 1)$ vector $\mathbf{x}$ satisfying the equation above is called an **eigenvector** corresponding to λ.
>
> *Note.* The equation above is equivalent to
>
> $$(A - \lambda I)\mathbf{x} = \theta, \ \mathbf{x} \neq \theta. \tag{2}$$
>
> Therefore, the eigenvalue problem consists of 2 parts:
>
> 1. Find all scalars λ such that $A - \lambda I$ is singular.
> 2. Given a scalar λ such that $A - \lambda I$ is singular, find all nonzero vectors $\mathbf{x}$ such that $(A - \lambda I)\mathbf{x} = \theta$.

> *Th.* Let A be an $(n \times n)$ matrix. Then $\det(A - tI)$ is a polynomial of degree n in t.

> *Def.* **Characteristic Polynomial**
>
> Let A be an $(n \times n)$ matrix. The nth-degree polynomial, $p(t)$, given by
>
> $$p(t) = \det(A - tI) \tag{3}$$
>
> is called the **characteristic polynomial** for A.

> *Th.* Let A be an $(n \times n)$ matrix, and let p be the characteristic polynomial for A. Then the eigenvalues of A are precisely the roots of $p(t) = 0$. This equation is called the **characteristic equation**.

> *Note.*
>
> 1. An $(n \times n)$ matrix can have no more than n distinct eigenvalues.
> 2. An $(n \times n)$ matrix always has at least one eigenvalue (possible complex).
> 3. The characteristic polynomial can be rewritten as $p(t) = a\prod_{i=1}^{n}(t - r_i)$. The number of times the factor $(t - r)$ appears in the factorization of $p(t)$ given above is called the **algebraic multiplicity** of r.

Special Results

> *Th.* Let A be an $(n \times n)$ matrix, and let λ be an eigenvalue of A. Then
>
> 1. λ^k is an eigenvalue of A^k, $k = 2, 3, \cdots$
> 2. If A is nonsingular, then $1/\lambda$ is an eigenvalue of A^{-1}
> 3. If α is any scalar, then $\lambda + \alpha$ is an eigenvalue of $A + \alpha I$.

> *Th.* Let A be an $(n \times n)$ matrix. Then A and A^T have the same eigenvalues.

> *Th.* Let A be an $(n \times n)$ matrix. Then A is singular if and only if $\lambda = 0$ is an eigenvalue of A.
>
> *Proof.* From the definition of eigenvalue problem.

> *Th.* Let $T = (t_{i,j})$ be an $(n \times n)$ triangular matrix. Then the eigenvalues of T are the diagonal entries, $t_{1,1}, t_{2,2}, \ldots, t_{n,n}$.

Eigenvectors and Eigenspaces

If λ is an eigenvalue of A, then the eigenvectors corresponding to λ are precisely the **nonzero** vectors in the null space of $A - \lambda I$.

> *Def.* Let A be an $(n \times n)$ matrix. If λ is an eigenvalue of A, then:
>
> - The null space of $A - \lambda I$ is denoted by E_λ and is called the **eigenspace** of λ.
> - The dimension of E_λ is called the **geometric multiplicity** of λ.
>
> *Note*: The geometric multiplicity of λ is never larger than the algebraic multiplicity of λ.

Defective Matrix

If A is an $(n \times n)$ matrix and if some eigenvalue of A has a geometric multiplicity that is less than its algebraic multiplicity, then A will not have a set of n linearly independent eigenvectors. Such a matrix is called defective.

> *Def.* Let A be an $(n \times n)$ matrix. If there is an eigenvalue λ of A such that the geometric multiplicity of λ is less than the algebraic multiplicity of λ, then A is called a **defective** matrix.

> *Th.* Let $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$ be eigenvectors of an $(n \times n)$ matrix A corresponding to distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$. That is,
>
> $$\begin{gathered} A\mathbf{u}_j = \lambda_j \mathbf{u}_j \ \ \forall\, j = 1, 2, \ldots, k. \ \ k \le n \\ \lambda_i \ne \lambda_j \ \ \forall\, i \ne j. \end{gathered} \tag{4}$$
>
> Then $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ is a linearly independent set
>
> *Corollary.* Let A be an $(n \times n)$ matrix. If A has n distinct eigenvalues, then A has a set of n linearly independent eigenvectors.

Complex Eigenvalues and Eigenvectors

When the characteristic equation has complex roots.

> *Def.* **The Norm of Complex Vector**
>
> The norm of a complex vector $\mathbf{x}$ (denoted by $\|\mathbf{x}\|$) is defined in terms of $\mathbf{x}$ and $\bar{\mathbf{x}}$:
>
> $$\|\mathbf{x}\| = \sqrt{\bar{\mathbf{x}}^T \mathbf{x}}. \tag{5}$$

> *Th.* Let A be a real $(n \times n)$ matrix with an eigenvalue λ and corresponding eigenvector $\mathbf{x}$. Then $\bar{\lambda}$ is also an eigenvalue of A, and $\bar{\mathbf{x}}$ is an eigenvector corresponding to it.

> *Th.* If A is an $(n \times n)$ real symmetric matrix, then all the eigenvalues of A are real.

Similarity Transformations and Diagonalization

Similarity

We are interested in identifying classes of matrices that have the same eigenvalues.

In particular, let A be an $(n \times n)$ matrix, and let S be a **nonsingular** $(n \times n)$ matrix and let $B = S^{-1}AS$. Let $p(t)$ be the characteristic polynomial of B. Then we have:

$$\begin{aligned} p(t) &= \det(S^{-1}AS - tI) = \det(S^{-1}AS - tS^{-1}S) \\ &= \det[S^{-1}(A - tI)S] = \det(S^{-1})\det(A - tI)\det(S) \\ &= \det(S^{-1})\det(S)\det(A - tI) \\ &= \det(A - tI) \end{aligned}$$

> *Def.* **Similarity**
>
> The $(n \times n)$ matrices A and B are said to be **similar** if there is a nonsingular $(n \times n)$ matrix S such that $B = S^{-1}AS$.

> *Th.* If A and B are similar $(n \times n)$ matrices, then A and B have the same eigenvalues. Moreover, these eigenvalues have the same algebraic multiplicity.
>
> *Note.* Although similar matrices always have the same characteristic polynomial, it is not true that two matrices with the same characteristic polynomial are necessarily similar.
>
> *Remark.* The only matrix similar to the identity matrix is I itself.
>
> *Note.* Similar matrices do not generally have the same eigenvector. If $B = S^{-1}AS$ and $B\mathbf{x} = \lambda\mathbf{x}$, then $S^{-1}AS\mathbf{x} = \lambda\mathbf{x}$ or $A(S\mathbf{x}) = \lambda(S\mathbf{x})$.

Diagonalization

Computations involving an $(n \times n)$ matrix A can often be simplified if we know that A is similar to a diagonal matrix. Suppose $S^{-1}AS = D$, where D is a diagonal matrix. we have:

$$\begin{aligned} D^k &= (S^{-1}AS)^k \\ &= S^{-1}ASS^{-1}AS \cdots AS \\ &= S^{-1}A^kS \end{aligned}$$

Then we have:

$$A^k = SD^kS^{-1} \tag{6}$$

> *Def.* **Diagonalizable Matrix**
>
> Let A be an $(n \times n)$ matrix and A is similar to a diagonal matrix, we say that A is **diagonalizable**.

How to determine whether a matrix is diagonalizable?

> *Th.* An $(n \times n)$ matrix A is diagonalizable if and only if A possesses(has) a set of n linearly independent eigenvectors.

If A is an $(n \times n)$ diagonalizable matrix and let matrix $S = [\mathbf{S}_1, \mathbf{S}_2, \cdots, \mathbf{S}_n]$ composed by the eigenvectors of A. We have:

$$S^{-1}AS = D \tag{7}$$

where

$$D = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \tag{8}$$

is a diagonal matrix whose elements are eigenvalues of A corresponding to the eigenvectors in S .

Proof.

$$\begin{aligned} AS &= [A\mathbf{S}_1, A\mathbf{S}_2, \cdots, A\mathbf{S}_n] \\ &= [\lambda_1\mathbf{S}_1, \lambda_2\mathbf{S}_2, \cdots, \lambda_n\mathbf{S}_n] \end{aligned}$$

and

$$\begin{aligned} SD &= [S\mathbf{D}_1, S\mathbf{D}_2, \cdots, S\mathbf{D}_n] \\ &= [\lambda_1 \mathbf{S}_1 + 0\mathbf{S}_2 + \cdots + 0\mathbf{S}_n, \cdots] \\ &= [\lambda_1 \mathbf{S}_1, \lambda_2 \mathbf{S}_2, \cdots, \lambda_n \mathbf{S}_n] \end{aligned}$$

thus

$$AS = SD \text{ or } S^{-1}AS = D \tag{9}$$

> *Corollary.* Let A be an $(n \times n)$ matrix with n distinct eigenvalues. Then A is diagonalizable.

Orthogonal Matrix

> *Th.* A real symmetric matrix is always diagonalizable.

> *Def.* **Orthogonal Matrix**
>
> A **real** $(n \times n)$ matrix Q is called an **orthogonal** matrix if Q is **invertible** and $Q^{-1} = Q^T$.
>
> Or
>
> A real square matrix Q is **orthogonal** if and only if $Q^TQ = I$.
>
> Or
>
> An $(n \times n)$ matrix $Q = [\mathbf{Q}_1, \mathbf{Q}_2, \cdots, \mathbf{Q}_n]$ is orthogonal if and only if the columns of Q from an **orthonormal (Not Orthogonal)** set of vectors.
>
> *Corollary.* If Q is an $(n \times n)$ orthogonal matrix and if P is formed by rearranging the columns of Q, then P is also an orthogonal matrix.

> *Def.* If P is an $(n \times n)$ matrix formed by rearranging the columns of $I_{n \times n}$, then P is called a **permutation** matrix.
>
> *Note.* Permutation matrix is orthogonal matrix.

> *Th.* Let Q be an $(n \times n)$ orthogonal matrix.
>
> 1. If $\mathbf{x} \in R^n$, then $\|Q\mathbf{x}\| = \|\mathbf{x}\|$. (because $\|Q\mathbf{x}\| = \sqrt{\mathbf{x}^T QTQ\mathbf{x}} = \sqrt{\mathbf{x}^T I\mathbf{x}}$).
> 2. If $\mathbf{x} \in R^n$ and $\mathbf{y} \in R^n$, then $(Q\mathbf{x})^T(Q\mathbf{y}) = \mathbf{x}^T\mathbf{y}$.
> 3. $\det(Q) = \pm 1$.
>
> *Note.* Orthogonal matrix is a rotation linear transformation.

Diagonalization of Symmetric Matrices

Every symmetric matrix can be diagonalized by an orthogonal matrix.

> *Th.* **Schur's Theorem(special case)**

> Let A be an $(n \times n)$ matrix, where A has only real eigenvalues. Then there is an $(n \times n)$ orthogonal matrix Q such that:
>
> $$Q^T AQ = T \tag{10}$$
>
> where T is an $(n \times n)$ **upper-triangular** matrix.
>
> Note that we can rewrite $Q^T AQ = T$ as $A = QTQ^T$, which is called a **Schur decomposition** or **Schur factorization** of A.

> *Th.* Let A be a real $(n \times n)$ matrix.
>
> 1. If A is symmetric, then there is an orthogonal matrix Q such that $Q^T AQ = D$, where D is diagonal.
> 2. If $Q^T AQ = D$, where Q is orthogonal and D is diagonal, then A is a symmetric matrix.

> *Corollary.* Let A be a real $(n \times n)$ symmetric matrix. It is possible to choose eigenvectors $\mathbf{u}_1, \ldots, \mathbf{u}_n$ for A such that $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ is an **orthonormal** basis for R^n.

Spectral Decomposition

Let A be an $(n \times n)$ symmetric matrix with eigenvalues $\lambda_1, \ldots, \lambda n$. Let $\mathbf{u}_1, \ldots, \mathbf{u}_n$ be a corresponding set of orthonormal eigenvectors. where $A\mathbf{u}_i = \lambda_i \mathbf{u}_i$. Matrix A can be expressed in the form:

$$A = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^T + \cdots + \lambda_n \mathbf{u}_n \mathbf{u}_n^T \tag{11}$$

Each $(n \times n)$ matrix $\mathbf{u}_i \mathbf{u}_i^T$ is a rank-one matrix. And the equation above is called **spectral decomposition** of A.

Proof. Let $U = [\mathbf{u}_1, \ldots, \mathbf{u}_n]$, because $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ is an orthonormal vector set, so U is an orthogonal matrix. We have $UU^T = I$. So:

$$A = AUU^T = A \sum_{i=1}^{n} \mathbf{u}_i \mathbf{u}_i^T = \sum_{i=1}^{n} A\mathbf{u}_i \mathbf{u}_i^T \tag{12}$$
$$A\mathbf{u}_i = \lambda_i \mathbf{u}_i$$

Then

$$A = \sum_{i=1}^{n} \lambda_i \mathbf{u}_i \mathbf{u}_i^T \tag{13}$$