COM521500
Math. Methods for SP I
Lecture 2: Eigenvalues and Eigenvectors

The Basics

Let $\mathbf{A}$ be an $n \times n$ matrix. The eigenvalue problem is to find a vector $\mathbf{v}$ such that

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}, \quad \mathbf{v} \neq \mathbf{0}$$

The scalar λ is called an **eigenvalue** of $\mathbf{A}$, and the vector $\mathbf{v}$ is called an **eigenvector** of $\mathbf{A}$.

The eigen-equation can be rewritten as

$$(\mathbf{A} - \lambda \mathbf{I})\mathbf{v} = \mathbf{0},$$

which is satisfied if and only if $\mathbf{A} - \lambda \mathbf{I}$ is singular. Thus,

$$\det(\mathbf{A} - \lambda \mathbf{I}) = 0 \qquad (*)$$

Eq. (*) is called the **characteristic equation** of $\mathbf{A}$, and $\det(\mathbf{A} - \lambda \mathbf{I})$ is called the **characteristic polynomial** of $\mathbf{A}$.

The function $\det(\mathbf{A} - \lambda \mathbf{I})$ is a polynomial of degree n, which means that it can be factored as

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

where $\lambda_1, \ldots, \lambda_n$ are the roots of the polynomial.

Thus,

$$\mathbf{A}\mathbf{v}_i = \lambda_i \mathbf{v}_i, \quad i = 1, \ldots, n$$

where $\mathbf{v}_i$ is the eigenvector associated with the eigenvalue λ_i, for $i = 1, \ldots, n$.

Some remarks:

1. For any eigenvector $\mathbf{v}_i$, any vector $c\mathbf{v}$, $c \in \mathbb{R}$ is also an eigenvector. Often, the eigenvectors are assumed to be normalized; i.e., $\|\mathbf{v}_i\|_2 = 1$.
2. For an $\mathbf{A} \in \mathbb{R}^{n\times n}$, it is possible to have complex-valued λ_i's (recall a root of a real-valued polynomial can be complex-valued). Under such circumstances the eigenvectors may be complex-valued too. Hence, it is convenient for us to study the case of $\mathbf{A} \in \mathbb{C}^{n\times n}$, which subsumes the case of $\mathbf{A} \in \mathbb{R}^{n\times n}$.

Similarity, and Diagonalizability

A matrix $\mathbf{B} \in \mathbb{C}^{n\times n}$ is **similar** to a matrix $\mathbf{A} \in \mathbb{C}^{n\times n}$ if there exists a nonsingular matrix $\mathbf{S} \in \mathbb{C}^{n\times n}$ such that

$$\mathbf{B} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}$$

Property 2.1 If $\mathbf{A}$ and $\mathbf{B}$ are similar, then the characteristic polynomial of $\mathbf{A}$ is same as that of $\mathbf{B}$.

Property 2.2 If $\mathbf{A}$ and $\mathbf{B}$ are similar, then they have the same eigenvalues, and the same determinant.

A matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ is said to be **diagonalizable** if it is similar to a diagonal matrix.

In other words, for a diagonalizable matrix $\mathbf{A}$ we can find a nonsingular matrix $\mathbf{S}$ such that

$$\mathbf{A} = \mathbf{S}\mathbf{D}\mathbf{S}^{-1}$$

where $\mathbf{D}$ is a diagonal matrix.

Theorem 2.1 A matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ is diagonalizable if and only if there is a set of n linearly independent vectors, each of which is an eigenvector of $\mathbf{A}$.

Distinct eigenvalues

Property 2.3 Suppose that $\{\lambda_1, \ldots, \lambda_k\}$, $k \leq n$, is a set of distinct eigenvalues; i.e, $\lambda_i \neq \lambda_j$ for $i \neq j$, $i, j \in \{1, 2, \ldots, k\}$. Then, the corresponding set of eigenvectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is linearly independent.

It follows that if all the eigenvalues of $\mathbf{A}$ are distinct, then $\mathbf{A}$ is diagonalizable.

Repeated eigenvalues

In the case where there are, say, r repeated eigenvalues, then a linearly independent set of r eigenvectors for those eigenvalues exists, provided that

$$\text{rank}(\mathbf{A} - \lambda\mathbf{I}) = n - r \qquad (*)$$

Example: Show that

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

satisfies Condition (*).

Assume that the eigenvectors of $\mathbf{A}$ are linear independent. From Theorem 2.1, we obtain the following eq. known as the **eigendecomposition**:

$$\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$$

where

$$\mathbf{V} = [\ \mathbf{v}_1, \ldots, \mathbf{v}_n\ ]$$

$$\mathbf{\Lambda} = \text{Diag}(\lambda_1, \ldots, \lambda_n) = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

Orthogonality

Two vectors $\mathbf{x}, \mathbf{y}$ (either real or complex valued) are said to be **orthogonal** if

$$< \mathbf{x}, \mathbf{y} >= 0$$

A set of vectors $\{\mathbf{x}_1, \dots, \mathbf{x}_m\}$ is said to be **orthogonal** if $< \mathbf{x}_i, \mathbf{x}_k >= 0$ for all $i \neq k$.

A set of vectors $\{\mathbf{x}_1, \dots, \mathbf{x}_m\}$ is said to be **orthonormal** if $< \mathbf{x}_i, \mathbf{x}_k >= 0$ for all $i \neq k$, and $\|\mathbf{x}_i\|_2^2 = 1$ for all i.

Property: An orthogonal set of vectors is linearly independent.

A matrix $\mathbf{U} \in \mathbb{C}^{n \times n}$ is said to be **unitary** if

$$\mathbf{U}^H \mathbf{U} = \mathbf{I}$$

A matrix $\mathbf{U} \in \mathbb{R}^{n \times n}$ is said to be **orthogonal** if

$$\mathbf{U}^T \mathbf{U} = \mathbf{I}$$

From the definition, a unitary (orthogonal) matrix is a matrix where its columns form an orthonormal set of vectors.

Some properties for unitary (orthogonal) matrices:

1. $\mathbf{U}^{-1} = \mathbf{U}^H$.
2. $\mathbf{U}\mathbf{U}^H = \mathbf{I}$.
3. The rows of $\mathbf{U}$ form an orthonormal set of vectors.

Symmetric and Hermitian matrices

A matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is **symmetric** when

$$\mathbf{A} = \mathbf{A}^T$$

A matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ is **Hermitian** when

$$\mathbf{A} = \mathbf{A}^H$$

A real sym. matrix is a Hermitian matrix.

Note: Symmetric and Hermitian matrices are very frequently encountered in SP, and hence their eigenvector/ eigenvalue properties deserve particular attention.

Property 2.4 The eigenvalues of a Hermitian matrix are real.

Property 2.5 Let $\mathbf{A}$ be a Hermitian matrix, and suppose that all the eigenvalues of $\mathbf{A}$ are distinct. Then, the eigenvectors of $\mathbf{A}$ are mutually orthogonal.

We conclude that under the distinct eigenvalue assumption, the eigendecomposition of a Hermitian $\mathbf{A}$ is

$$\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^H \qquad (\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T \text{ for real sym. } \mathbf{A})$$

Now the question remained is: can a Hermitian (real sym.) matrix have eigendecomposition?

This question has been answered by **Schur triangularization theorem**.

Theorem 2.2 (Schur triangularization) Given a matrix $\mathbf{A} \in \mathbb{C}^{n\times n}$, there exists a unitary matrix $\mathbf{U} \in \mathbb{C}^{n\times n}$ such that

$$\mathbf{U}^H \mathbf{A} \mathbf{U} = \mathbf{T}$$

where $\mathbf{T}$ is an upper triangular matrix with main diagonal $\mathrm{diag}(\mathbf{T}) = [\ \lambda_1, \ldots, \lambda_n\ ]^T$.

Now, consider a Hermitian $\mathbf{A}$.

$$\begin{aligned}\mathbf{T}^H &= (\mathbf{U}^H \mathbf{A} \mathbf{U})^H \\ &= \mathbf{U}^H \mathbf{A}^H \mathbf{U} = \mathbf{U}^H \mathbf{A} \mathbf{U} = \mathbf{T}\end{aligned}$$

This implies that $\mathbf{T}$ is diagonal, and that $\mathbf{T} = \mathbf{\Lambda}$. Thus,

Theorem 2.3 A Hermitian matrix $\mathbf{A} \in \mathbb{C}^{n\times n}$ can always be diagonalized as

$$\mathbf{V}^H \mathbf{A} \mathbf{V} = \mathbf{\Lambda} \qquad (\mathbf{V}^T \mathbf{A} \mathbf{V} = \mathbf{\Lambda} \text{ for real sym. } \mathbf{A})$$

where $\mathbf{\Lambda} = [\ \lambda_1, \ldots, \lambda_n\ ]$, $\mathbf{V} = [\ \mathbf{v}_1, \ldots, \mathbf{v}_n\ ]$, $\{\ \lambda_i\ \}$ is the set of eigenvalues of $\mathbf{A}$, $\mathbf{v}_i$ is the normalized eigenvector of $\mathbf{A}$ associated with λ_i.

Some properties obtained from Theorem 2.3:

Property 2.6 For Hermitian $\mathbf{A}$, $\mathbf{A}^{-1} = \mathbf{V}\mathbf{\Lambda}^{-1}\mathbf{V}^H$. [note that $\mathbf{\Lambda}^{-1} = \text{Diag}(\frac{1}{\lambda_1}, \ldots, \frac{1}{\lambda_n})$.]

Property 2.7 For Hermitian $\mathbf{A}$, $\text{rank}(\mathbf{A})$ is the number of nonzero eigenvalues.

Matrix Norms

The definition of a matrix norm is equivalent to that of a vector norm.

Specifically, a matrix norm is a function $\|.\| : \mathbb{C}^{m\times n} \to \mathbb{R}$ that satisfies the following properties:

1. $\|\mathbf{X}\| \geq 0$ for all $\mathbf{X} \in \mathbb{C}^{m\times n}$
2. $\|\mathbf{X}\| = 0$ if and only if $\mathbf{X} = \mathbf{0}$
3. $\|\mathbf{X}+\mathbf{Y}\| \leq \|\mathbf{X}\| + \|\mathbf{Y}\|$ for any $\mathbf{X}, \mathbf{Y} \in \mathbb{C}^{m\times n}$
4. $\|c\mathbf{X}\| = |c|\|\mathbf{X}\|$ for $c \in \mathbb{C}$, $\mathbf{X} \in \mathbb{C}^{m\times n}$

Frobenius Norm:

$$\|\mathbf{A}\|_F^2 = \left[\sum_{i=1}^{m}\sum_{k=1}^{n}|a_{ik}|^2\right]^{1/2}$$

Note that

$$\|\mathbf{A}\|_F^2 = \left[\mathrm{tr}(\mathbf{A}^H\mathbf{A})\right]^{1/2}$$

Matrix p-Norms:

$$\|\mathbf{A}\|_p^2 = \sup_{\mathbf{x}\neq\mathbf{0}} \frac{\|\mathbf{Ax}\|_p}{\|\mathbf{x}\|_p} = \max_{\|\mathbf{x}\|_p=1} \|\mathbf{Ax}\|_p$$

For the matrix 2-norm,

$$\|\mathbf{A}\|_2 = \sqrt{\lambda_{max}}$$

where λ_{max} is the largest eigenvalue of $\mathbf{A}^H\mathbf{A}$.

As we will see later in this course, $\|\mathbf{A}\|_2$ is also the largest **singular value** of $\mathbf{A}$.

Some useful inequalities:

1. $\|\mathbf{Ax}\|_p \leq \|\mathbf{A}\|_p \|\mathbf{x}\|_p$

2. Let $\mathbf{Q}$ and $\mathbf{Z}$ be unitary matrices of appropriate sizes.

$$\|\mathbf{QAZ}\|_2 = \|\mathbf{A}\|_2$$
$$\|\mathbf{QAZ}\|_F = \|\mathbf{A}\|_F$$

3. $\|\mathbf{AB}\|_p \leq \|\mathbf{A}\|_p \|\mathbf{B}\|_p$