ENGMAE 200A: Engineering Analysis I

Matrix Eigenvalue Problems

Instructor: Dr. Ramin Bostanabad
DEFINITIONS

• What is the difference between the results of these multiplications:

\[
\begin{bmatrix}
6 & 3 \\
4 & 7
\end{bmatrix}
\begin{bmatrix}
5 \\
1
\end{bmatrix}
\text{ vs. }
\begin{bmatrix}
6 & 3 \\
4 & 7
\end{bmatrix}
\begin{bmatrix}
3 \\
4
\end{bmatrix}
\]

• A matrix eigenvalue problem considers the vector equation:

\[\mathbf{Ax} = \lambda \mathbf{x}\] (1)

• Here \(\mathbf{A}\) is a given square matrix, \(\lambda\) an unknown scalar, and \(\mathbf{x}\) an unknown vector. In a matrix eigenvalue problem, the task is to determine \(\lambda\)’s and \(\mathbf{x}\)’s
• \( \lambda \)'s for which (1) has a solution \( x \neq 0 \) is called: an eigenvalue, characteristic value, or latent root.

• Solutions \( x \neq 0 \) of (1) are called the eigenvectors, characteristic vectors of \( A \) corresponding to that eigenvalue \( \lambda \).

• The set of all the eigenvalues of \( A \) is called the spectrum of \( A \). We shall see that the spectrum consists of at least one eigenvalue and at most of \( n \) numerically different eigenvalues.

• The largest of the absolute values of the eigenvalues of \( A \) is called the spectral radius of \( A \).
We use an example:

\[ A = \begin{bmatrix} 6 & 3 \\ 4 & 7 \end{bmatrix} \]

\[ Ax = \begin{bmatrix} -5 & 2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \]

1. \[ (A - \lambda I)x = 0 \]

2. \[ -5x_1 + 2x_2 = \lambda x_1 \]
   \[ 2x_1 - 2x_2 = \lambda x_2 \]

3. \[ (-5 - \lambda)x_1 + 2x_2 = 0 \]
   \[ 2x_1 + (-2 - \lambda)x_2 = 0 \]

4. \[ (A - \lambda I)x = 0 \]

5. \[ D(\lambda) = \det(A - \lambda I) = \begin{vmatrix} -5 - \lambda & 2 \\ 2 & -2 - \lambda \end{vmatrix} \]
   \[ = (-5 - \lambda)(-2 - \lambda) - 4 = \lambda^2 + 7\lambda + 6 = 0 \]
• $D(\lambda)$ is the **characteristic determinant** or, if expanded, the characteristic polynomial.
• $D(\lambda) = 0$ the **characteristic equation** of A.

**Example Continued:**
• Solutions of $D(\lambda) = 0$ are $\lambda_1 = -1$ and $\lambda_2 = -6$. These are the eigenvalues of A.
• Eigenvectors for $\lambda_1 = -1$:

• Eigenvectors for $\lambda_1 = -6$: 
What is the difference between the results of these multiplications:

\[
\begin{bmatrix}
6 & 3 \\
4 & 7
\end{bmatrix}
\begin{bmatrix}
5 \\
1
\end{bmatrix}
\text{ vs. }
\begin{bmatrix}
6 & 3 \\
4 & 7
\end{bmatrix}
\begin{bmatrix}
3 \\
4
\end{bmatrix}
\]

Obtain the eigenvalues and eigenvectors.
GENERAL CASE

• How to find the eigenvalues and eigenvectors of:

\[ \mathbf{A} \mathbf{x} = \lambda \mathbf{x} \quad (1) \]

\[
\begin{align*}
a_{11}x_1 + \cdots + a_{1n}x_n &= \lambda x_1 \\
a_{21}x_1 + \cdots + a_{2n}x_n &= \lambda x_2 \\
\vdots & \quad \vdots \\
a_{n1}x_1 + \cdots + a_{nn}x_n &= \lambda x_n.
\end{align*}
\]

\[
(A - \lambda I) \mathbf{x} = 0.
\]

\[
D(\lambda) = \det(A - \lambda I) = \begin{vmatrix}
  a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\
  a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\
  \vdots & \vdots & \ddots & \vdots \\
  a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda
\end{vmatrix} = 0.
\]
EXAMPLE

Find the eigenvalues and eigenvectors:

\[
\mathbf{A} = \begin{bmatrix}
-2 & 2 & -3 \\
2 & 1 & -6 \\
-1 & -2 & 0
\end{bmatrix}
\]

- Characteristic equation: \(-\lambda^3 - \lambda^2 + 21\lambda + 45 = 0\)
- Roots: \(\lambda_1 = 5, \lambda_2 = \lambda_3 = -3\).
- Form \(\mathbf{A} - \lambda \mathbf{I}\) and then use Gauss elimination:

\[
\mathbf{A} - \lambda \mathbf{I} = \mathbf{A} - 5\mathbf{I} = \begin{bmatrix}
-7 & 2 & -3 \\
2 & -4 & -6 \\
-1 & -2 & -5
\end{bmatrix} \rightarrow \begin{bmatrix}
0 & 2 & -3 \\
0 & -24/7 & -48/7 \\
0 & 0 & 0
\end{bmatrix}
\]

- Hence, the rank is 2.
EXAMPLE CONTINUED

• Choosing $x_3 = -1$ we have $x_2 = 2$ from $-\frac{24}{7} x_2 - \frac{48}{7} x_3 = 0$ and then $x_1 = 1$ from $-7x_1 + 2x_2 - 3x_3 = 0$: $x_1 = [1 \ 2 \ -1]^T$.

• For $\lambda = -3$ the characteristic matrix:

\[ A - \lambda I = A + 3I = \begin{bmatrix} 1 & 2 & -3 \\ 2 & 4 & -6 \\ -1 & -2 & 3 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \]
\[
A - \lambda I = A + 3I = \begin{bmatrix}
1 & 2 & -3 \\
2 & 4 & -6 \\
-1 & -2 & 3
\end{bmatrix} \rightarrow \begin{bmatrix}
1 & 2 & -3 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{bmatrix}
\]

- From \(x_1 + 2x_2 - 3x_3 = 0\) we have \(x_1 = -2x_2 + 3x_3\). Choosing \(x_2 = 1, x_3 = 0\) and \(x_2 = 0, x_3 = 1\), we obtain two linearly independent eigenvectors of \(A\) corresponding to \(\lambda = -3\):
Find the eigenvalues and eigenvectors:

\[
A = \begin{bmatrix}
0 & 1 \\
-1 & 0
\end{bmatrix}
\]

- What are the eigenvalues and eigenvectors of \(A^T\)?
SOME THEOREMS

• **Theorem 1:** The eigenvalues of a square matrix $A$ are the roots of $D(\lambda) = 0$.
  - An $n \times n$ matrix has at least one eigenvalue and at most $n$ numerically different eigenvalues.

• **Theorem 2:** If $w$ and $x$ are eigenvectors of a matrix $A$ corresponding to the same eigenvalue $\lambda$, so are $w + x$ (provided $x \neq -w$) and $kx$ for any $k \neq 0$.
  - Eigenvectors corresponding to one and the same eigenvalue $\lambda$ of $A$, together with $0$, form a vector space, called the *eigenspace* of $A$ corresponding to that $\lambda$.
  - *An eigenvector $x$ is determined only up to a constant factor.* So, we can *normalize* $x$. 
EXAMPLE ON LAND USE REVISITED

The 2004 state of land use in a city of 60 $mi^2$ of built-up area is:

C: Commercially Used 25%

I: Industrially Used 20%

R: Residentially Used 55%

- The transition probabilities for 5-year intervals are given by $A$ and remain practically the same over the time considered.

$$A = \begin{bmatrix} 0.7 & 0.1 & 0 \\ 0.2 & 0.9 & 0.2 \\ 0.1 & 0 & 0.8 \end{bmatrix}$$

To C

To I

To R
What is the **limit state**?

From C  From I  From R

\[
A = \begin{bmatrix}
0.7 & 0.1 & 0 \\
0.2 & 0.9 & 0.2 \\
0.1 & 0 & 0.8 \\
\end{bmatrix}
\]

- Definition of limit state
- How to find it systematically
EIGENVALUES OF SPECIAL MATRICES

Symmetric, Skew-Symmetric, and Orthogonal:

A real square matrix $A = [a_{jk}]$ is called symmetric if transposition leaves it unchanged:

(1) $A^T = A$, thus $a_{kj} = a_{jk},$

skew-symmetric if transposition gives the negative of $A$:

(2) $A^T = -A$, thus $a_{kj} = -a_{jk},$

orthogonal if transposition gives the inverse of $A$:

(3) $A^T = A^{-1}.$
Theorem: Eigenvalues of Symmetric and Skew-Symmetric Matrices

(a) *The eigenvalues of a symmetric matrix are real.*

(b) *The eigenvalues of a skew-symmetric matrix are pure imaginary or zero.*

Examples:

\[
A = \begin{bmatrix}
-3 & 1 & 5 \\
1 & 0 & -2 \\
5 & -2 & 4
\end{bmatrix} \quad B = \begin{bmatrix}
0 & 9 & -12 \\
-9 & 0 & 20 \\
12 & -20 & 0
\end{bmatrix}
\]
Orthogonal transformations: Transformations like $y = Ax$ where $A$ is an orthogonal matrix.

- With each vector $x$ in $\mathbb{R}^n$ such a transformation assigns a vector $y$ in $\mathbb{R}^n$. For instance, the plane rotation through an angle $\theta$

$$y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

is an orthogonal transformation.
Theorem: Invariance of Inner Products

An orthogonal transformation preserves the value of the inner product:

\[ \mathbf{a} \cdot \mathbf{b} = \mathbf{a}^T \mathbf{b} = [a_1 \ldots a_n] \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} \]

That is, for any \( \mathbf{a} \) and \( \mathbf{b} \) in \( \mathbb{R}^n \), orthogonal \( n \times n \) matrix \( \mathbf{A} \), and \( \mathbf{u} = \mathbf{Aa} \), \( \mathbf{v} = \mathbf{Ab} \) we have \( \mathbf{u} \cdot \mathbf{v} = \mathbf{a} \cdot \mathbf{b} \).

Hence the transformation also preserves the length or norm of any vector \( \mathbf{a} \) in \( \mathbb{R}^n \)

\[ \|\mathbf{a}\| = \sqrt{\mathbf{a}^T \mathbf{a}} \]
• **Orthonormality of Column and Row Vectors:** A real square matrix is orthogonal if and only if its column vectors \(\mathbf{a}_1, \ldots, \mathbf{a}_n\) (and also its row vectors) form an **orthonormal system**:

\[
\mathbf{a}_j^T \mathbf{a}_k = \begin{cases} 
0 & \text{if } j \neq k \\
1 & \text{if } j = k 
\end{cases}
\]

• **Determinant of an Orthogonal Matrix:** Has the value +1 or −1.

• **Eigenvalues of an orthogonal matrix \(\mathbf{A}\):** Real or complex conjugates in pairs and have absolute value 1.
Orthogonal matrix:

\[
A = \begin{bmatrix}
2 & 1 & 2 \\
-2 & 2 & 1 \\
1 & 2 & -2 \\
\end{bmatrix} \times \frac{1}{3}
\]
**Theorem:** If an \( n \times n \) matrix \( A \) has \( n \) **distinct** eigenvalues, then \( A \) has a basis of eigenvectors \( x_1, \ldots, x_n \) for \( \mathbb{R}^n \).

**Theorem:** A symmetric matrix has an orthonormal basis of eigenvectors for \( \mathbb{R}^n \).

**Example:**

\[
A = \begin{bmatrix}
-2 & 2 & -3 \\
2 & 1 & -6 \\
-1 & -2 & 0
\end{bmatrix}
\]

\[
\lambda_1 = 5, \quad \lambda_2 = -3, \quad \lambda_3 = -3
\]

\[
x_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \quad x_2 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \quad x_3 = \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}
\]
**Definition:** An $n \times n$ matrix $\hat{A}$ is called **similar** to an $n \times n$ matrix $A$ if

$$\hat{A} = P^{-1}AP$$

for some (nonsingular!) $n \times n$ matrix $P$. This transformation, which gives $\hat{A}$ from $A$, is called a **similarity transformation**.

**Theorem:** If $\hat{A}$ is similar to $A$, then $\hat{A}$ has the same eigenvalues as $A$. Furthermore, if $x$ is an eigenvector of $A$, then $y = P^{-1}x$ is an eigenvector of $\hat{A}$ corresponding to the same eigenvalue.
Let’s revisit the matrix in slide 2:

\[ A = \begin{bmatrix} 6 & 3 \\ 4 & 7 \end{bmatrix} \]

Let’s choose \( P = \begin{bmatrix} 1 & 3 \\ 1 & 4 \end{bmatrix} \) which gives \( P^{-1} = \begin{bmatrix} 4 & -3 \\ -1 & 1 \end{bmatrix} \)

Then:
\[
\hat{A} = \begin{bmatrix} 4 & -3 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 6 & -3 \\ 4 & -1 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}
\]
Theorem: If an $n \times n$ matrix $A$ has a basis of eigenvectors, then

$$D = X^{-1}AX$$

is diagonal, with the eigenvalues of $A$ as the entries on the main diagonal. Here $X$ is the matrix with these eigenvectors as column vectors.

Also:

$$D^m = X^{-1}A^mX \quad (m = 2, 3, \ldots).$$
Diagonalize:

\[
A = \begin{bmatrix}
7.3 & 0.2 & -3.7 \\
-11.5 & 1.0 & 5.5 \\
17.7 & 1.8 & -9.3
\end{bmatrix}
\]

Solution:
Definition: A quadratic form \( Q \) in the components \( x_1, \ldots, x_n \) of a vector \( \mathbf{x} \) is a sum \( n^2 \) of terms:

\[
Q = \mathbf{x}^\top \mathbf{A} \mathbf{x} = \sum_{j=1}^{n} \sum_{k=1}^{n} a_{jk} x_j x_k
\]

\[
= a_{11} x_1^2 + a_{12} x_1 x_2 + \cdots + a_{1n} x_1 x_n
\]

\[
+ a_{21} x_2 x_1 + a_{22} x_2^2 + \cdots + a_{2n} x_2 x_n
\]

\[
+ \cdots + a_{nn} x_n^2
\]

\( \mathbf{A} = [a_{jk}] \) is called the **coefficient matrix** of the form. We may assume that \( \mathbf{A} \) is symmetric (?).
EXAMPLE

Let

$$x^T A x = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 3 & 4 \\ 6 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= 3x_1^2 + 4x_1x_2 + 6x_2x_1 + 2x_2^2$$

$$= 3x_1^2 + 10x_1x_2 + 2x_2^2.$$

Here $4 + 6 = 10 = 5 + 5$. So:

$$x^T C x = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 3 & 5 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= 3x_1^2 + 5x_1x_2 + 5x_2x_1 + 2x_2^2$$

$$= 3x_1^2 + 10x_1x_2 + 2x_2^2.$$
MANIPULATING QUADRATIC FORMS

• Say you are given something like this:
  \[ 3x_1^2 + 10x_1 x_2 + 2x_2^2 \]

• How can you convert it to a canonical form:
  \[ \alpha y_1^2 + \beta y_2^2 \]

• Applications are many! We will look into an important one.
• Symmetric coefficient matrix $A$ has an orthonormal basis of eigenvectors (Theorem on slide 21). So, if we take these as column vectors, we obtain a matrix $X$ that is orthogonal, so that $X^{-1} = X^T$: 

$$A = XDX^{-1} = XDX^T.$$ 

• Substitution: $Q = x^T XDX^T x$.

• Set $X^T x = y$. Since $X^{-1} = X^T$, we have $X^{-1} x = y$ and so $x = Xy$.

• We also have $x^T X = (X^T x)^T = y^T$ and $X^T x = y$.

• Now $Q$ becomes

$$Q = y^T Dy = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2$$
**Theorem:** The substitution $x = Xy$ transforms a quadratic form

$$ Q = x^T A x = \sum_{j=1}^{n} \sum_{k=1}^{n} a_{jk} x_j x_k $$

$$ (a_{kj} = a_{jk} ) $$

to the principal axes form or **canonical form**

$$ Q = y^T D y = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2 $$

$\lambda_1, \ldots, \lambda_n$ are the (not necessarily distinct) eigenvalues of the (symmetric!) matrix $A$, and $X$ is an orthogonal matrix with corresponding eigenvectors $x_1, \ldots, x_n$, respectively, as column vectors.
EXAMPLE

What type of conic section the following quadratic form represents

\[ Q = 17x_1^2 - 30x_1x_2 + 17x_2^2 = 128. \]

Solution. We have \( Q = x^T A x \), where

\[ A = \begin{bmatrix} 17 & -15 \\ -15 & 17 \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \]

Characteristic equation \((17 - \lambda)^2 - 15^2 = 0\) with roots \( \lambda_1 = 2, \lambda_2 = 32 \). So:

\[ Q = 2y_1^2 + 32y_2^2. \quad \rightarrow \quad \frac{y_1^2}{8^2} + \frac{y_2^2}{2^2} = 1. \]

What is the direction of the principal axes in the \( x_1x_2 \)-coordinates?
A square matrix $A = [a_{kj}]$ is called

- **Hermitian** if $\bar{A}^T = A$, that is, $\bar{a}_{kj} = a_{jk}$
- **skew-Hermitian** if $\bar{A}^T = -A$, that is, $\bar{a}_{kj} = -a_{jk}$
- **unitary** if $\bar{A}^T = A^{-1}$

These are generalizations of symmetric, skew-symmetric, and orthogonal matrices in complex spaces.

For example, (Theorem on invariance of Inner Product): The unitary transformation $y = Ax$ with a unitary matrix $A$, preserves the value of the inner product and norm.
generalizing the theorems

- The eigenvalues of a Hermitian matrix (and thus of a symmetric matrix) are real.
- The eigenvalues of a skew-Hermitian matrix (and thus of a skew-symmetric matrix) are pure imaginary or zero.
- The eigenvalues of a unitary matrix (and thus of an orthogonal matrix) have absolute value 1.