Linear Systems

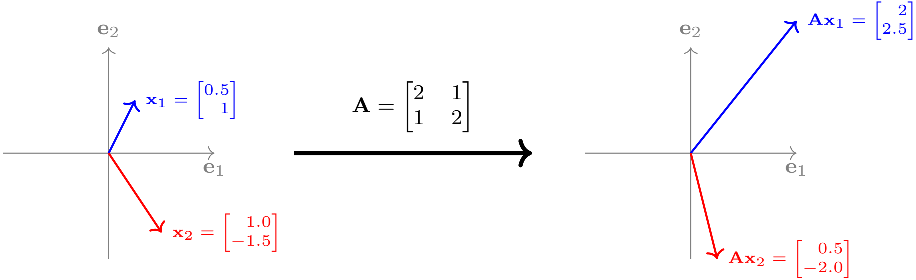
Eigenvalues and Eigenvectors

Sivakumar Balasubramanian

Department of Bioengineering
Christian Medical College, Bagayam
Vellore 632002

Linear transformation

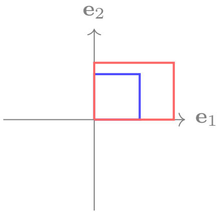
- Matrices represent linear transformations, $\mathbf{A} \in \mathbb{R}^{m \times n}$ represents a transformation $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$.
$$\mathbf{y} = T(\mathbf{x}) = \mathbf{A}\mathbf{x}, \quad \mathbf{x} \in \mathbb{R}^n \text{ and } \mathbf{y} \in \mathbb{R}^m$$
- Consider a linear transformation $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$.
In general, T scales and rotates the vector $\mathbf{x}$ to produce $\mathbf{y}$.



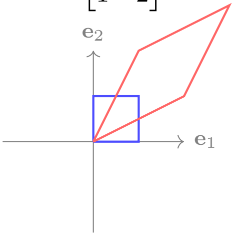
Linear transformation

An easier way is to look at what happens to the standard basis $\{\mathbf{e}_i\}_{i=1}^n$.

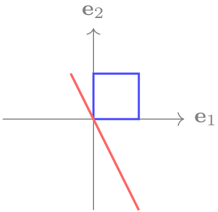
$$\mathbf{A} = \begin{bmatrix} 1.75 & 0 \\ 0 & 1.25 \end{bmatrix}$$



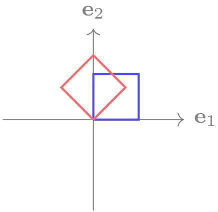
$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$



$$\mathbf{A} = \begin{bmatrix} -0.5 & 1 \\ 1 & -2 \end{bmatrix}$$



$$\mathbf{A} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$



Linear transformation in different basis

- Consider a basis $V = \{\mathbf{v}_i\}_{i=1}^{n}$ for $\mathbb{R}^n$. Let $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \ldots & x_n \end{bmatrix}^\top \in \mathbb{R}^n$ be the representation of $\mathbf{x}$ in the standard basis.
  Representation of $\mathbf{x}$ in V is,

$$\mathbf{x}_V = \sum_{i=1}^{n} x_{vi}\mathbf{v}_i, \;\; \mathbf{x}_V = \begin{bmatrix} x_{v1} & x_{v2} & \ldots & x_{vn} \end{bmatrix}^\top$$

- We can go back and forth between these two representations in the following way,

$$\mathbf{x} = \mathbf{V}\mathbf{x}_V \quad \text{and} \quad \mathbf{x}_V = \mathbf{V}^{-1}\mathbf{x}; \;\; \text{where, } \mathbf{V} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \ldots & \mathbf{v}_n \end{bmatrix}$$

- When V is an orthonormal basis, then the algebra gets simpler,

$$\mathbf{x} = \mathbf{V}\mathbf{x}_V \quad \text{and} \quad \mathbf{x}_V = \mathbf{V}^\top\mathbf{x}$$

Linear transformation in different basis

Consider a linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ represented by the matrix $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$.

Consider a vector $\mathbf{x} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$. What is $\mathbf{y} = \mathbf{A}\mathbf{x}$?

$$\mathbf{y} = \mathbf{A}\mathbf{x} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$$

Now, consider a basis $V = \left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\}$ for $\mathbb{R}^2$. The representation of $\mathbf{x}, \mathbf{y}$ in V is,

$$\mathbf{x}_V = \mathbf{V}^{-1}\mathbf{x} = \frac{1}{3} \begin{bmatrix} 1 & -2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 4 \\ 1 \end{bmatrix}, \quad \mathbf{y}_V = \mathbf{V}^{-1}\mathbf{y} = \frac{1}{3} \begin{bmatrix} 8 \\ -1 \end{bmatrix}$$

Now, if we apply the linear transformation T on $\mathbf{x}_V$ will we get $\mathbf{y}_V$?

$$\mathbf{A}\mathbf{x}_V = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \frac{1}{3} \begin{bmatrix} -4 \\ 1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} -4 \\ 5 \end{bmatrix} \neq \mathbf{y}_V$$

Representation of a linear transformation T is basis dependent!

Similarity transformation

- Linear transformations represented in one basis represent a different transformation in another basis. This issue can be addressed by keeping track of the basis one is working in.

- Let $\mathbf{x}, \mathbf{y}$ be representations in the standard basis. Changing basis to V, gives us $\mathbf{x}_V, \mathbf{y}_V$.

$$\mathbf{y}_V = \mathbf{V}^{-1}\mathbf{y} = \mathbf{V}^{-1}\mathbf{A}\mathbf{x} = \mathbf{V}^{-1}\mathbf{A}\mathbf{V}\mathbf{x}_V = \mathbf{A}_V\mathbf{x}_V$$

Similarity transformation

- Check if this works with the example in the previous slide. $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$; $\mathbf{V} = \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}$. Determine $\mathbf{A}_V$ and check that $\mathbf{y}_V = \mathbf{A}_V \mathbf{x}_V$.

Similarity transformation

- A linear transformation $\hat{T}$ is represented as $\mathbf{A}_V$ in V. What is its representation in the standard basis? E.g.: $\mathbf{A}_V = \begin{bmatrix} -4 & 1 \\ 1 & 2 \end{bmatrix}$; $\mathbf{V} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$. Determine $\mathbf{A}$. If $\mathbf{x}_V = \begin{bmatrix} \frac{1}{2} & 2 \end{bmatrix}^\top$. What is $\mathbf{y}_V$ and $\mathbf{y}$?

Similarity transformation

- Two matrices $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$ are called *similar* matrices, if there exists a non-singular matrix $\mathbf{Q}$, such that,

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

- The transformation represented by $\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q}$ is called the *similarity transformation*.

- *Similar* matrices represent the same linear transformation in different basis.

- When $\mathbf{Q}$ is an orthogonal matrix, we have $\mathbf{B} = \mathbf{Q}^{\top}\mathbf{A}\mathbf{Q}$.

Complex Vectors and Matrices

- Similar to $\mathbb{R}^n$, we can have $\mathbb{C}^n$. $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_{r1} + jx_{i1} \\ x_{r2} + jx_{i2} \\ \vdots \\ x_{rn} + jx_{in} \end{bmatrix}$
- Vector addition and scalar mulitplication are the same. The scalar is a complex number.
- Additive identity, and scalar multiplication identity are the same. So is the **standard basis** $\{\mathbf{e}_i\}_{i=1}^n$
- **Linear independence**: The set $\{\mathbf{v}_i\}_{i=1}^n$ with $\mathbf{v}_i \in \mathbb{C}^n$ is linearly independent, if $\sum_{i=1}^n c_i \mathbf{v}_i = 0, \implies c_i = 0, \ \forall 1 \leq i \leq n, \ \ c_i \in \mathbb{C}$
- **Inner product**: $\mathbf{x}^H \mathbf{y} = \begin{bmatrix} \overline{x}_1 & \overline{x}_2 & \dots & \overline{x}_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \sum_{i=1}^n \overline{x}_i y_i$

Complex Vectors and Matrices

- **Length**: $\|\mathbf{x}\|_2^2 = \mathbf{x}^H\mathbf{x} = \sum_{i=1}^{n} \overline{x}_i x_i = \sum_{i=1}^{n} |x_i|^2$
- **Orthogonality**: $\mathbf{x}^H\mathbf{y} = 0$
- Complex matrices have complex entries. $\mathbf{A} \in \mathbb{C}^{m\times n}$ such that $a_{ij} \in \mathbb{C},\ \forall 1 \leq i \leq m,\ 1 \leq j \leq n$
- The transpose operation is generalized to conjugate transpose known as the *Hermitian*. $\mathbf{A}^H = \overline{\mathbf{A}}^\top$.
- The idea of symmetric matrices $\mathbb{R}^{n\times n}$ are now generalized to $\mathbb{C}^{n\times n}$ as $\mathbf{A} = \mathbf{A}^H$. Such matrices are called **Hermitian** matrices.
- Orthogonal matrices in the complex case are called **Unitary** matrices, $\mathbf{U}^H\mathbf{U} = \mathbf{I} \implies \mathbf{U}^{-1} = \mathbf{U}^H$.

Eigenvectors and Eigenvalues

- Any linear transformation represented by $\mathbf{A} \in \mathbb{C}^{n \times n}$ has vectors that satisfy the following property,

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}, \quad \mathbf{x} \in \mathbb{C}^n, \lambda \in \mathbb{C}, \ \mathbf{x} \neq \mathbf{0}$$

  where, λ and $\mathbf{x}$ are called the eigenvalue and the associated eigenvector of $\mathbf{A}$.

- Any such pair $(\lambda, \mathbf{x})$ is called the eigenpair of $\mathbf{A}$.

- These are important for understanding and solving linear differential and difference equations:

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{A}\mathbf{x}(t) \quad \text{and} \quad \mathbf{x}[n+1] = \mathbf{A}\mathbf{x}[n]$$

Eigenvectors and Eigenvalues

Consider the differential equation, $\frac{d\mathbf{x}(t)}{dt} = \mathbf{A}\mathbf{x}\left(t\right) = \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} \mathbf{x}\left(t\right)$. Let us assume that the solution is of the form, $\mathbf{x} = e^{\lambda t}\hat{\mathbf{x}}$. Then we have,

$$\frac{d\mathbf{x}\left(t\right)}{dt} = e^{\lambda t}\mathbf{A}\hat{\mathbf{x}} = e^{\lambda t}\lambda\hat{\mathbf{x}} \implies \mathbf{A}\hat{\mathbf{x}} = \lambda\hat{\mathbf{x}}$$

$$\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} = \begin{bmatrix} \lambda\hat{x}_1 \\ \lambda\hat{x}_2 \end{bmatrix} \implies \begin{bmatrix} 2-\lambda & 1 \\ 1 & 4-\lambda \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

where, $\hat{\mathbf{x}} \in N\left(\mathbf{A} - \lambda\mathbf{I}\right)$.

Eigenvectors and Eigenvalues

This problem can be solved by, $\det(\mathbf{A} - \lambda\mathbf{I}) = |\mathbf{A} - \lambda\mathbf{I}| = 0$

$$(2-\lambda)(4-\lambda) - 1 = \lambda^2 - 6\lambda + 7 = 0 \implies \lambda = 3 \pm \sqrt{2}$$

$\mathbf{A}\hat{\mathbf{x}} = \left(3+\sqrt{2}\right)\hat{\mathbf{x}} \implies \hat{\mathbf{x}} = \begin{bmatrix} -1+\sqrt{2} \\ 1 \end{bmatrix}$ and $\mathbf{A}\hat{\mathbf{x}} = \left(3-\sqrt{2}\right)\hat{\mathbf{x}} \implies \hat{\mathbf{x}} = \begin{bmatrix} -1-\sqrt{2} \\ 1 \end{bmatrix}$

$\left(3+\sqrt{2}, \begin{bmatrix} -1+\sqrt{2} \\ 1 \end{bmatrix}\right)$ and $\left(3-\sqrt{2}, \begin{bmatrix} -1-\sqrt{2} \\ 1 \end{bmatrix}\right)$ are the eigenpairs of $\mathbf{A}$.

Eigenvalues and Eigenvectors

- We can find the eigenpairs using the same approach for $\mathbf{A} \in \mathbb{C}^{n \times n}$, $\det(\mathbf{A} - \lambda \mathbf{I}) = 0 = p(\lambda)$.

- $p(\lambda)$ is the characteristic polynomial of $\mathbf{A}$, and $p(\lambda) = 0$ is the characteristic equation.

- The eigenvalues are the roots of the polynomial $p(\lambda)$, and the $\mathbf{x}$ in $(\mathbf{A} - \lambda \mathbf{I})\mathbf{x} = 0$ for the different λs are the corresponding eigenvectors.

Eigenvalues and Eigenvectors

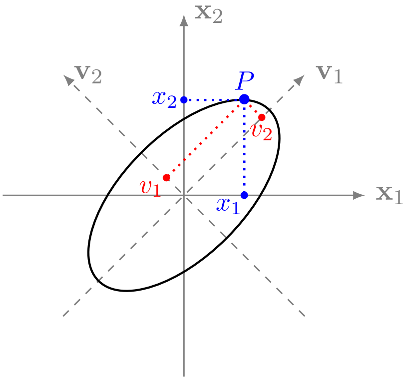
Compute the eigenpairs for the following matrices: $\begin{bmatrix} 2 & 1 \\ -1 & 3 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$.

Eigenvalues and Eigenvectors

Compute the eigenpairs for the following matrices: $\begin{bmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix}$, $\begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix}$.

Diagonalization of a matrix

Often the right choice of basis can simplify an equation or the analysis of a problem. For example,



The equation of the ellipse in standard basis is:

$$3x_1^2 + 3x_2^2 - 2x_1x_2 = 1$$

This has a much simpled representation in the dashed coordinate frame.

$$4v_1^2 + 2v_2^2 = 1$$

The use of similarity transformation to simplify a matrix is at the heart of diagonalization.

Diagonalization of a matrix

- Consider a matrix $\mathbf{A}$ with n eigenpairs $\{(\lambda_i, \mathbf{x}_i)\}_{i=1}^{n}$.

$$\mathbf{A}\begin{bmatrix}\mathbf{x}_1 & \mathbf{x}_2 & \ldots & \mathbf{x}_n\end{bmatrix} = \begin{bmatrix}\lambda_1\mathbf{x}_1 & \lambda_2\mathbf{x}_2 & \ldots & \lambda_n\mathbf{x}_n\end{bmatrix}$$

$$\mathbf{AX} = \mathbf{X}\begin{bmatrix}\lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n\end{bmatrix} = \mathbf{X\Lambda}$$

- If the eignevectors are linearly independent, then we have $\mathbf{X}^{-1}\mathbf{AX} = \mathbf{\Lambda}$

Diagonalization of a matrix

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ represented by $\mathbf{A} = \begin{bmatrix} 8 & 1 \\ 2 & 7 \end{bmatrix}$. Diagonalize this matrix. What does $\mathbf{A}$ do to $\mathbf{x} = \begin{bmatrix} 3 & 4 \end{bmatrix}^\top$?

Diagonalization of a matrix

What about $\mathbf{A} = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}$?

Diagonalization of a matrix: Eigenpairs of special matrices

- A square matrices with a complete set of eigenvectors, i.e. a linearly independent set of n eigenvectors, can be decomposed into the following,

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

Diagonalization of a matrix: Eigenpairs of special matrices

- When $\mathbf{A} \in \mathbb{R}^{n \times n}$ is symmetric, i.e. $\mathbf{A} = \mathbf{A}^\top$,
  - All eigenvalues are real.
  - The matrix poses a complete set of eigenvectors, i.e. they form a linearly independent set.
  - The eigenvectors can be chosen to be orthogonal to each other. When the eigenvalues are distinct, the eigenvectors are orthogonal. But when the eigenvalues are not distinct, we can choose them to be orthogonal.

This gives us, $\mathbf{A} = \mathbf{A}^\top = \mathbf{X}\mathbf{\Lambda}\mathbf{X}^\top$.

Diagonalization of a matrix: Eigenpairs of special matrices

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

Diagonalization of a matrix

- A change of basis to $\mathbf{X}$ simplifies $\mathbf{A}$ to a diagonal matrix, the simplest possible form.
- If a matrix $\mathbf{A}$ has n distinct eigenvalues, then $\mathbf{A}$ can always be diagonalized.
- When there are repeated eigenvalues, we might not always be able to diagonalize a matrix. This happens when there aren't enough eigenvectors. These are called *defective* matrices.

$$\text{Algebraic multiplicity} \neq \text{Geometric multiplicity}$$

where, *algebraic multiplicity* is the number of times the eigenvalue λ is repeated, and *geometric multiplicity* is $\dim N\left(\mathbf{A} - \lambda\mathbf{I}\right)$.

Diagonalization of a matrix

Diagonalize $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$.

Jordan Form

- If $\mathbf{A}$ cannot be diagonalized, the next best thing is the *Jordan form*.
- Let $\mathbf{A}$ have eigenvalues $(\lambda_1, \lambda_2, \ldots \lambda_k)$. We can find a similarity transformation, such that,

$$\mathbf{A} = \mathbf{P}\mathbf{J}\mathbf{P}^{-1}, \;\; \mathbf{J} = \begin{bmatrix} \mathbf{J}(\lambda_1) & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{J}(\lambda_2) & \ldots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \mathbf{J}(\lambda_k) \end{bmatrix}$$

Each $\mathbf{J}(\bullet)$ is associated with an eigenvalue and an eigenvector, and is called a Jordan block, and has the form

$$\mathbf{J}(\lambda_l) = \begin{bmatrix} \lambda_l & 1 & 0 & \ldots & 0 \\ 0 & \lambda_l & 1 & \ldots & 0 \\ 0 & 0 & \lambda_l & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & 1 \\ 0 & 0 & 0 & \ldots & \lambda_l \end{bmatrix}$$

- $\mathbf{J} \in \mathbb{C}^{r \times r}$. $r =$ the algebraic multiplicity of the eigenvalue λ_l.
- $1 =$ the geometric multiplicity of the eigenvalue $\lambda_l = \dim N(\mathbf{A} - \lambda_l \mathbf{I})$.
- A 1-by-1 Jordan block is simply $[\lambda_l]$, corresponding to a eigenvalue with an associated eigenvector.

Jordan Form

- Jordan form of a diagonalizable matrix $\mathbf{A} \rightarrow \mathbf{J} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$

$\lambda = -2$ (AM[1] $= 1$, GM[2] $= 1$), and $\lambda = 11$ (AM $= 2$, GM $= 1$) $\rightarrow \mathbf{J} = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 11 & 1 \\ 0 & 0 & 11 \end{bmatrix}$

[1] AM: Algebraic multiplicity
[2] GM: Geometric multiplicity

Jordan Form

Write down the Jordan form.

$\lambda_1 = 1$ (AM = 2, GM = 1)

$\lambda_2 = 11$ (AM = 3, GM = 2)

$\lambda_3 = 0$ (AM = 3, GM = 1)

$\lambda_4 = -1$ (AM = 2, GM = 2).