STA 221: LECTURE 2

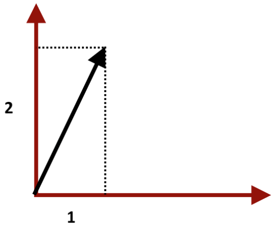
KRISHNA BALASUBRAMANIAN

(University of California, Davis)

- A vector has a direction and a "magnitude" (norm)
- Example (2-norm):

$$\boldsymbol{x} = [1, 2]^T, \quad \|\boldsymbol{x}\|_2 = \sqrt{1^2 + 2^2} = \sqrt{5}$$

- Properties satisfied by a vector norm ($\|\cdot\|$)
  - $\|\boldsymbol{x}\| \geq 0$ and $\|\boldsymbol{x}\| = 0$ if and only if $\boldsymbol{x} = 0$
  - $\|\alpha \boldsymbol{x}\| = |\alpha| \|\boldsymbol{x}\|$ (homogeneity)
  - $\|\boldsymbol{x} + \mathbf{y}\| \leq \|\boldsymbol{x}\| + \|\mathbf{y}\|$ (triangle inequality)



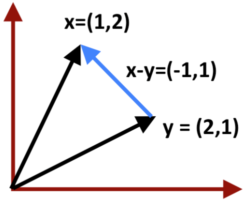
EXAMPLES OF VECTOR NORMS

- $\boldsymbol{x} = [x_1, x_2, \cdots, x_n]^T$
- $\|\boldsymbol{x}\|_2 = \sqrt{|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2}$ (2-norm)
- $\|\boldsymbol{x}\|_1 = |x_1| + |x_2| + \cdots + |x_n|$ (1-norm)
- $\|\boldsymbol{x}\|_p = \sqrt[p]{|x_1|^p + |x_2|^p + \cdots + |x_n|^p}$ (p-norm)
- $\|\boldsymbol{x}\|_\infty = \max_{1 \leq i \leq n} |x_i|$ (∞-norm)

- $x = [1, 2], y = [2, 1]$

  $\boldsymbol{x} - \mathbf{y} = [1 - 2, 2 - 1] = [-1, 1]$
- Distance:

$$\|\boldsymbol{x} - \mathbf{y}\|_2 = \|[-1, 1]\|_2 = \sqrt{2}$$
$$\|\boldsymbol{x} - \mathbf{y}\|_1 = 2$$
$$\|\boldsymbol{x} - \mathbf{y}\|_\infty = 1$$



- $\boldsymbol{x} = [x_1, \cdots, x_n]^T$, $\mathbf{y} = [y_1, , \cdots, y_n]^T$
- Inner product:

$$\boldsymbol{x}^T \mathbf{y} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n = \sum_{i=1}^{n} x_i y_i$$

- $\boldsymbol{x}^T \boldsymbol{x} = \|\boldsymbol{x}\|_2^2, \quad \|\boldsymbol{x} - \mathbf{y}\|_2^2 = (\boldsymbol{x} - \mathbf{y})^T (\boldsymbol{x} - \mathbf{y})$
- Orthogonal:

  $x \perp y \Leftrightarrow \boldsymbol{x}^T \mathbf{y} = 0$

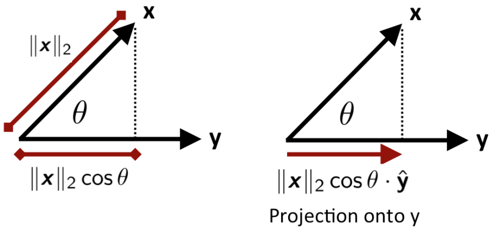
  ($\boldsymbol{x}$ and $\mathbf{y}$ are orthogonal to each other)

- $\boldsymbol{x}^T \mathbf{y} = \|\boldsymbol{x}\|_2 \|\mathbf{y}\|_2 \cos(\theta)$

$$\cos\theta = \frac{\boldsymbol{x}^T \mathbf{y}}{\|\boldsymbol{x}\|_2 \|\mathbf{y}\|_2} \Rightarrow \quad \|\boldsymbol{x}\|_2 \cos\theta = \boldsymbol{x}^T \left(\frac{\mathbf{y}}{\|\mathbf{y}\|_2}\right) = \boldsymbol{x}^T \widehat{\mathbf{y}}$$

- Projection of $\boldsymbol{x}$ onto $\mathbf{y}$:

$$\|\boldsymbol{x}\|_2 \cos\theta \cdot \widehat{\mathbf{y}} = \underbrace{\widehat{\mathbf{y}}\widehat{\mathbf{y}}^T}_{\text{projection matrix}} \boldsymbol{x}$$



- Suppose we have 3 vectors $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3$

  $\boldsymbol{x}_1 = \alpha_2 \boldsymbol{x}_2 + \alpha_3 \boldsymbol{x}_3 \quad \Rightarrow \quad \boldsymbol{x}_1$ is linearly dependent on $\boldsymbol{x}_2$ and $\boldsymbol{x}_3$
- When are $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n$ linearly independent?
  $\alpha_1 \boldsymbol{x}_1 + \alpha_2 \boldsymbol{x}_2 + \cdots + \alpha_n \boldsymbol{x}_n = 0$ if and only if
  $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$
- A vector space is a set of vectors that is closed under vector addition & scalar multiplications.
  If $\boldsymbol{x}_1, \boldsymbol{x}_2 \in V$, then $\alpha_1 \boldsymbol{x}_1 + \alpha_2 \boldsymbol{x}_2 \in V$
- A basis of the vector space is the maximal set of vectors in the subspace that are linearly independent of each other.
- An orthogonal basis is a basis where all basis vectors are orthogonal to each other.
- Dimension of the vector space: number of vectors in the basis.

- ▷ This is an m by n matrix: $A \in \mathbb{R}^{m\times n}$:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

- ▷ Its job is to do linear transformations:

$$\begin{aligned} A : \mathbb{R}^m &\to \mathbb{R}^n \\ \boldsymbol{x} &\to A\boldsymbol{x} \\ \alpha\boldsymbol{x} + \beta\mathbf{y} &\to A(\alpha\boldsymbol{x} + \beta\mathbf{y}) = \alpha A\boldsymbol{x} + \beta A\mathbf{y} \text{ (Linear Transform)} \end{aligned}$$

- Popular matrix norm: Frobenius norm

$$\|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2}$$

- Matrix norms satisfy following properties:

$$\|A\| \geq 0 \text{ and } \|A\| = 0 \text{ if and only if } A = 0 \text{ (positivity)}$$
$$\|\alpha A\| = |\alpha|\|A\| \text{ (homogeneity)}$$
$$\|A + B\| \leq \|A\| + \|B\|$$

- Column rank of A: the dimension of column space (vector space formed by column vectors)
- Row rank of A: the dimension of row space
- Column rank = row rank := rank (always true)
- Examples:

Rank 2 matrix:

$$\begin{bmatrix} 1 & 0 & 1 \\ -2 & -3 & 1 \\ 3 & 3 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -2 & -3 \\ 3 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix}$$

Rank 1 matrix:

$$\begin{bmatrix} 1 & 1 & 0 & 2 \\ -1 & -1 & 0 & -2 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 2 \end{bmatrix}$$

- For a m by m matrix A, if

$$A\mathbf{y} = \lambda \mathbf{y},$$

  then we say

  λ is an eigenvalue of A

  $\mathbf{y}$ is the corresponding eigenvector
- Eigen decomposition defined only for square matrices.
- Eigenvalues in general can be real or complex numbers.

- Consider $A \in \mathbb{R}^{m \times m}$ to be a square, symmetric matrix. The eigenvalue decomposition of A is:

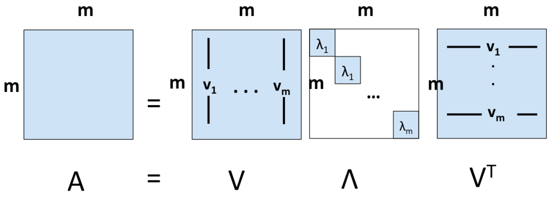
$$A = V \Lambda V^T, \quad V^T V = I (V \text{ is unitary}), \quad \Lambda \text{ is diagonal}$$

- $A = V \Lambda V^T \Rightarrow AV = V \Lambda$
  $\Rightarrow A\boldsymbol{v}_i = \lambda_i \boldsymbol{v}_i, \quad \forall i = 1, \cdots, m$
- Each $\boldsymbol{v}_i$ is an eigenvector, and each λ_i is an eigenvalue
- Usually, we assume the diagonal numbers are organized in descending order:

$$\Lambda = \text{diag}(\lambda_1, \lambda_2, \cdots, \lambda_m),$$
$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m$$

- Eigenvalue decomposition is unique when there are m unique eigenvalues.

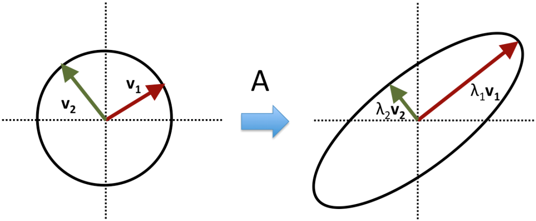
EIGEN DECOMPOSITION



- ▷ Each eigenvector $\boldsymbol{v}_i$ will be mapped to $A\boldsymbol{v}_i = \lambda \boldsymbol{v}_i$ after the linear transform:

  Scaling without changing the direction of eigenvectors
- ▷ $A\boldsymbol{x} = \sum_{i=1}^{m} \lambda_i \boldsymbol{v}_i (\boldsymbol{v}_i^T \boldsymbol{x})$

  Project $\boldsymbol{x}$ to eigenvectors, and then scaling each vector



Visualization of matrix as transformation.