## Linear Algebra and Calculus!

Jeremy Irvin and Daniel Spokoyny

#### **Eigenvalues and Eigenvectors**

• Given a square matrix  $A \in \mathbb{R}^{n \times n}$ , we say  $\lambda \in \mathbb{C}$  is an <u>eigenvalue</u> of A and  $x \in \mathbb{C}^n$  is the corresponding <u>eigenvector</u> if

$$Ax = \lambda x, \quad x \neq 0.$$

Intuitively, this means multiplying A by x results in a new vector in the same direction as x but scaled by λ.

#### **Eigenvalues and Eigenvectors**

- Note that if *x* is an eigenvector, then *cx* is an eigenvector for any complex *c*.
- We can rewrite the equation above as

$$(\lambda I - A)x = 0, \quad x \neq 0.$$

•  $(\lambda I - A)x = 0$  has a non-zero solution iff  $(\lambda I - A)$  has a non-empty nullspace, which only happens if  $(\lambda I - A)$  is singular, ie,  $|(\lambda I - A)| = 0.$ 

#### **E-values and E-vectors Properties**

• The trace of *A* is equal to the sum of its eigenvalues,

$$\mathrm{tr}A = \sum_{i=1}^{N} \lambda_i.$$

• The determinant of *A* is equal to the product of its eigenvalues

$$|A| = \prod_{i=1}^{n} \lambda_i.$$

- The rank of *A* is equal to the number of nonzero eigenvalues of *A*.
- The eigenvalues of a diagonal matrix are just the diagonal entries.

#### Diagonalization

• We can write all the eigenvector equations simultaneously as

$$AX = X\Lambda.$$

with  $X \in \mathbb{R}^{n \times n}$  the eigenvectors of A and a diagonal matrix  $\Lambda$  whose entries are the eigenvalues A, ie,

$$X \in \mathbb{R}^{n \times n} = \begin{bmatrix} | & | & | & | \\ x_1 & x_2 & \cdots & x_n \\ | & | & | & | \end{bmatrix}, \quad \Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_n).$$

#### Diagonalization

 If the eigenvectors of A are linearly independent, then X will be invertible, so

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

We say that *A* is <u>diagonalizable</u>.



#### **Quadratic Forms**

- Given any symmetric matrix  $A \in \mathbb{R}^{n \times n}$ and vector  $x \in \mathbb{R}^n$ , the scalar value is called a  $x^T A x$  quadratic form.
- Explicitly, we have

$$x^{T}Ax = \sum_{i=1}^{n} x_{i}(Ax)_{i} = \sum_{i=1}^{n} x_{i}\left(\sum_{j=1}^{n} A_{ij}x_{j}\right) = \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij}x_{i}x_{j} \quad .$$



#### **Definite Matrices**

• A is positive definite if for all non-zero vectors  $x \in \mathbb{R}^n$ 

 $x^T A x > 0.$ 

• A is <u>negative definite</u> if for all non-zero vectors  $x \in \mathbb{R}^n$ 

 $x^T A x < \mathbf{0}.$ 

• Positive and negative definite matrices are full rank and thus invertible.

• For any matrix  $A \in \mathbb{R}^{m \times n}$ ,  $A^T A$  is positive semidefinite.

**E-values and E-vectors of Symmetric Matrices** 

- Let  $A \in \mathbb{R}^{n \times n}$  be any symmetric matrix:
  - All eigenvalues of *A* are real.
  - The non-collinear eigenvectors of *A* are orthonormal.
  - Thus we can decompose A:
      $A = U\Lambda U^T$  where U is an orthogonal matrix.

**E-values and E-vectors of Symmetric Matrices** 

 We can use this to show that definiteness only depends on sign of eigenvalues:

$$x^{T}Ax = x^{T}U\Lambda U^{T}x = y^{T}\Lambda y = \sum_{i=1}^{n} \lambda_{i}y_{i}^{2}$$

where  $y = U^T x$ .

• For any quadratic form  $x^T A x$  subject to  $||x||_2^2 = 1$ , its maximum value is the maximum eigenvalue of A, and its minimum value is the minimum eigenvalue of A.

• Goal: Given any matrix  $A \in \mathbb{R}^{m \times n}$ , find orthogonal matrices U and V such that

$$A = U\Sigma V^T$$

• If A diagonalizable,

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

- If *A* positive semidefinite,
  - $A = U\Lambda U^T$

- See blackboard pictures.
- We know we can find an orthonormal basis for the rowspace of *A*.
- Can we find one that is mapped into an orthonormal basis for the column space of *A*?



• Goal: Find an orthonormal basis  $v_1, ..., v_r$  for the row space of A such that

$$Av_1 = \sigma_1 u_1, \dots, Av_r = \sigma_r u_r$$

where  $u_1, ..., u_r$  is an orthonormal basis for the column space of A.



• In matrix notation,

 $A\begin{bmatrix} | & | & | & | & | \\ v_1 & \cdots & v_r & v_{r+1} & \cdots & v_m \\ | & | & | & | & | \end{bmatrix} = \begin{bmatrix} | & | & | & | & | \\ u_1 & \cdots & u_r & u_{r+1} & \cdots & u_n \\ | & | & | & | & | \end{bmatrix} \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_r & \\ \hline & & 0 & 0 \end{bmatrix}$ 

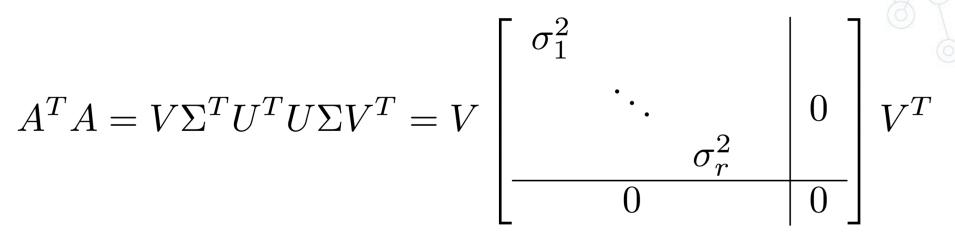
where  $v_{r+1}, ..., v_m$  orthonormal basis for the null space of A, and  $u_{r+1}, ..., u_n$  an orthonormal basis for the null space of  $A^T$ .

 $AV = U\Sigma$ 

- But because  $\mathcal{N}(A)$  and  $\mathcal{R}(A)$  are orthogonal complements, V is orthogonal.
- Similarly, *U* is orthogonal.
- Therefore

# $AV = U\Sigma \Rightarrow A = U\Sigma V^T$

How do we find U and V?
Trick:



Since A<sup>T</sup> A is positive semidefinite, V is the orthogonal matrix of eigenvectors of A<sup>T</sup> A, and its eigenvalues are the squares of the diagonal entries of Σ.

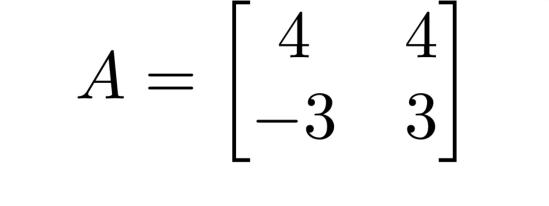
 $\int \sigma_1^2$ 

• Similarly,

$$AA^T = U\Sigma^T V^T V\Sigma U^T = U$$

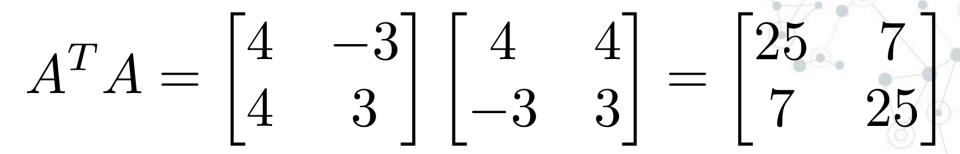
$$\begin{bmatrix}
1 & & & & & \\
& \ddots & & & 0 \\
& & \sigma_r^2 & & \\
& & 0 & & 0
\end{bmatrix}
U^T$$

• So to find *U*, simply find the eigenvectors of *AA*<sup>*T*</sup>.



Find orthonormal  $v_1, v_2$  in the row space of  $A(\mathbb{R}^2)$  and orthonormal  $u_1, u_2$  in the column space of  $A(\mathbb{R}^2)$ , and  $\sigma_1, \sigma_2 > 0$  such that

### $Av_1 = \sigma_1 u_1, Av_2 = \sigma_2 u_2$



#### Therefore the eigenvectors of $A^T A$ are

$$\begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}, \begin{bmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}$$

with eigenvalues 32 and 18 respectively.

# $AA^{T} = \begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix} \begin{bmatrix} 4 & -3 \\ 4 & 3 \end{bmatrix} = \begin{bmatrix} 32 & 0 \\ 0 & 18 \end{bmatrix}$

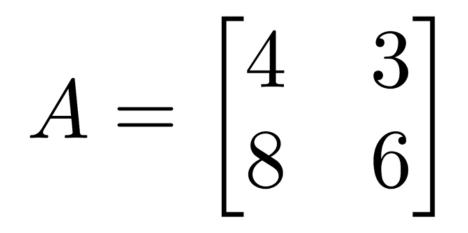
# Therefore the eigenvectors of $A^T A$ are $\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

again with eigenvalues 32 and 18 respectively (is this surprising?)

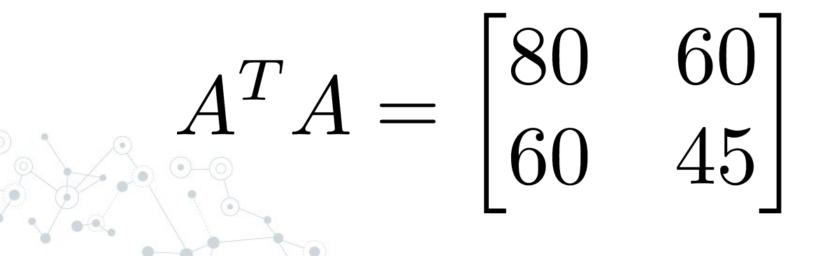
#### Therefore:

## $A = \begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{32} & 0 \\ 0 & \sqrt{18} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = U\Sigma V^T$





#### See blackboard for geometric intuition.



#### SVD Example 2.

• Therefore the eigenvalues of  $A^T A$  are 0 and 125. Hence:

$$A = \begin{bmatrix} 4 & 3 \\ 8 & 6 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{5}}{5} & \frac{2\sqrt{5}}{5} \\ \frac{2\sqrt{5}}{5} & -\frac{\sqrt{5}}{5} \end{bmatrix} \begin{bmatrix} \sqrt{125} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{4}{5} & \frac{3}{5} \\ \frac{3}{5} & -\frac{4}{5} \end{bmatrix} = U\Sigma V^T$$



#### So who cares?

- Why is SVD even useful?
- SVD can be used for **dimensionality** reduction - given high dimensional data, one can use SVD to represent the data using less dimensions, while still capturing the most significant (largest eigenvalues) features.

• We will see important applications later.