

# Statistical Machine Learning

## Lecture 02: Linear Algebra Refresher

**Kristian Kersting**

**TU Darmstadt**

Summer Term 2020

# Today's Objectives

- Make you remember Linear Algebra!
- I know this is mostly easy but some of you may have forgotten all of it...
- Covered Topics:
  - Vectors, Matrices
  - Linear Transformations

# Outline

## 1. Vectors

## 2. Matrices

## 3. Operations and Linear Transformations

## 4. Wrap-Up

# Outline

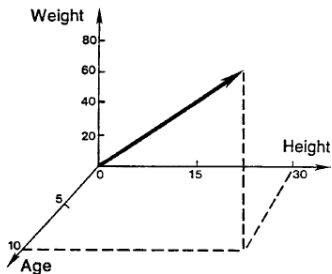
## 1. Vectors

## 2. Matrices

## 3. Operations and Linear Transformations

## 4. Wrap-Up

# Vectors



$$\text{Joe} = \begin{bmatrix} 37 \\ 72 \\ 175 \end{bmatrix}, \text{ Mary} = \begin{bmatrix} 10 \\ 30 \\ 61 \end{bmatrix}, \text{ Carol} = \begin{bmatrix} 25 \\ 65 \\ 121 \end{bmatrix}, \text{ Brad} = \begin{bmatrix} 66 \\ 67 \\ 155 \end{bmatrix}, \text{ Joe} = \begin{bmatrix} 37 \\ 72 \\ 175 \\ 8 \\ 1946 \end{bmatrix}$$

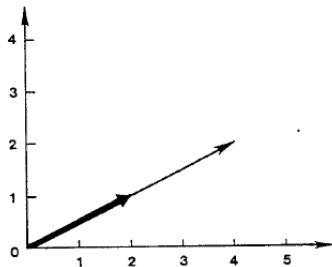
# What can you do with vectors?

## ■ Multiplication by a scalar $c\mathbf{v}$

$$2 \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$$

$$5 \begin{bmatrix} -3 \\ 4 \\ 1 \end{bmatrix} = \begin{bmatrix} -15 \\ 20 \\ 5 \end{bmatrix}$$

$$c\mathbf{v} = c \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} cv_1 \\ \vdots \\ cv_n \end{bmatrix}$$



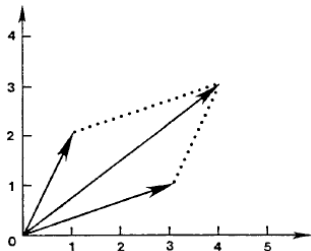
# What can you do with vectors?

## ■ Addition of vectors $\mathbf{v}_1 + \mathbf{v}_2$

$$\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 4 \end{bmatrix}$$

$$\begin{bmatrix} 2 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 3 \\ -3 \end{bmatrix} = \begin{bmatrix} 5 \\ -1 \end{bmatrix}$$

$$\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{bmatrix}$$



# Linear Combination of Vectors

- By positive recombination we can obtain:

$$\blacksquare \mathbf{u} = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_n \mathbf{v}_n$$

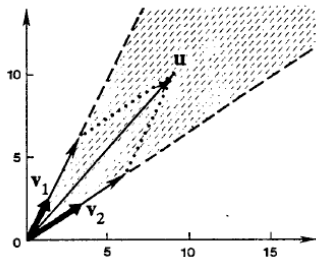
- Examples:

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ and } \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ and } \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix} \text{ and } \begin{bmatrix} -1 \\ 3 \end{bmatrix}$$

$$\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} 9 \\ 10 \\ 0 \end{bmatrix}$$





# Inner Product and Length of a Vector

## ■ Inner Product

$$\blacksquare \mathbf{v} = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}, \mathbf{w} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

$$\blacksquare \mathbf{v} \cdot \mathbf{w} = \mathbf{v}^T \mathbf{w} = (3 \cdot 1) + (-1 \cdot 2) + (2 \cdot 1) = 3$$

## ■ Length of a vector (Frobenius norm)

$$\blacksquare \|\mathbf{v}\| = (\mathbf{v} \cdot \mathbf{v})^{1/2}$$

$$\blacksquare \|c\mathbf{v}\| = |c| \|\mathbf{v}\|$$

$$\blacksquare \|\mathbf{v}_1 + \mathbf{v}_2\| \leq \|\mathbf{v}_1\| + \|\mathbf{v}_2\| \text{ (triangle inequality)}$$

# Angles between Vectors

- The angle between vectors is defined by

$$\cos \theta = \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{v}\| \|\mathbf{w}\|} = \frac{\sum_{i=1}^n v_i w_i}{\left(\sum_{i=1}^n v_i^2\right)^{1/2} \left(\sum_{i=1}^n w_i^2\right)^{1/2}}$$

- Example:

- Find the angle between vectors  $\mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$  and  $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

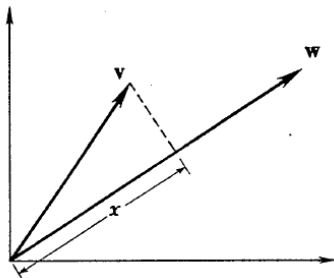
- $\mathbf{v}_1 \cdot \mathbf{v}_2 = 1$ ,  $\|\mathbf{v}_1\| = 1$ ,  $\|\mathbf{v}_2\| = \sqrt{2}$

- $\cos \theta = \frac{1}{1\sqrt{2}} = 0.707$ ,  $\theta = \pi/4$

# Projections of Vectors: Basic Idea

- What is a projection of  $\mathbf{v}$  onto  $\mathbf{w}$ ?
- Formally

$$\begin{aligned}
 x &= \|\mathbf{v}\| \cos \theta \\
 &= \|\mathbf{v}\| \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{v}\| \|\mathbf{w}\|} \\
 &= \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{w}\|}
 \end{aligned}$$



- Note that  $x$  is a **not** a vector!

# Vector Transpose, Inner and Outer Products

## ■ Vector Transpose

$$\blacksquare \mathbf{v} = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}, \mathbf{v}^T = [ 3 \quad 1 \quad 2 ]$$

## ■ Inner Product

$$\blacksquare \mathbf{v}^T \mathbf{u} = [ 3 \quad 1 \quad 2 ] \begin{bmatrix} 0 \\ 4 \\ 1 \end{bmatrix} = [ 6 ]$$

## ■ Outer Product

$$\blacksquare \mathbf{w} \mathbf{v}^T = \begin{bmatrix} 1 \\ 4 \\ 0 \end{bmatrix} [ 3 \quad 1 \quad 2 ] = \begin{bmatrix} 3 & 1 & 2 \\ 12 & 4 & 8 \\ 0 & 0 & 0 \end{bmatrix}$$

# Outline

1. Vectors

**2. Matrices**

3. Operations and Linear Transformations

4. Wrap-Up

# Matrices

## ■ Examples

■  $\mathbf{M} = \begin{bmatrix} 3 & 4 & 5 \\ 1 & 0 & 1 \end{bmatrix}$ , 2x3 matrix

■  $\mathbf{N} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ , 3x3 matrix

■  $\mathbf{P} = \begin{bmatrix} 10 & -1 \\ -1 & 27 \end{bmatrix}$ , 2x2 matrix

# What can you do with Matrices?

## ■ Multiplication by Scalars

$$3 \cdot \mathbf{M} = 3 \begin{bmatrix} 3 & 4 & 5 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 9 & 12 & 15 \\ 3 & 0 & 3 \end{bmatrix}$$

## ■ Addition of Matrices

$$\mathbf{M} + \mathbf{N} = \begin{bmatrix} 3 & 4 & 5 \\ 1 & 0 & 1 \end{bmatrix} + \begin{bmatrix} -1 & 0 & 2 \\ 4 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 7 \\ 5 & 1 & 0 \end{bmatrix}$$

- Addition is only defined for matrices with the same dimensions.

## ■ Transpose of a Matrix

$$\mathbf{M}^T = \begin{bmatrix} 3 & 4 & 5 \\ 1 & 0 & 1 \end{bmatrix}^T = \begin{bmatrix} 3 & 1 \\ 4 & 0 \\ 5 & 2 \end{bmatrix}$$

# Matrix-Vector multiplication

- Multiplication of a Vector by a Matrix

$$\mathbf{u} = \mathbf{W}\mathbf{v} = \begin{bmatrix} 3 & 4 & 5 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \cdot 1 + 4 \cdot 0 + 5 \cdot 2 \\ 1 \cdot 1 + 0 \cdot 0 + 1 \cdot 2 \end{bmatrix} = \begin{bmatrix} 13 \\ 3 \end{bmatrix}$$

- Think of it as

$$\begin{bmatrix} | & & | \\ \mathbf{w}_1 & \dots & \mathbf{w}_n \\ | & & | \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} v_1 \mathbf{w}_1 + & | & \\ \dots & & \\ v_n \mathbf{w}_n \end{bmatrix}$$

- Dimensions:  $\mathbf{W} \in \mathbb{R}^{M \times N}$ ,  $\mathbf{v} \in \mathbb{R}^{N \times 1}$ ,  $\mathbf{u} \in \mathbb{R}^{M \times 1}$

- Hence

$$\mathbf{u} = v_1 \mathbf{w}_1 + v_2 \mathbf{w}_2 + v_3 \mathbf{w}_3 = 1 \begin{bmatrix} 3 \\ 1 \end{bmatrix} + 0 \begin{bmatrix} 4 \\ 0 \end{bmatrix} + 2 \begin{bmatrix} 5 \\ 1 \end{bmatrix} = \begin{bmatrix} 13 \\ 3 \end{bmatrix}$$



# Matrix-Matrix multiplication

- Multiplication of a Matrix by a Matrix

$$\mathbf{C} = \mathbf{AB} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} =$$
$$\begin{bmatrix} 1 \cdot 1 + 2 \cdot 3 + 3 \cdot 5 & 1 \cdot 2 + 2 \cdot 4 + 3 \cdot 6 \\ 4 \cdot 1 + 5 \cdot 3 + 6 \cdot 5 & 4 \cdot 2 + 5 \cdot 4 + 6 \cdot 6 \end{bmatrix} = \begin{bmatrix} 22 & 28 \\ 49 & 64 \end{bmatrix}$$

- Dimensions:  $\mathbf{A} \in \mathbb{R}^{M \times N}$ ,  $\mathbf{B} \in \mathbb{R}^{N \times K}$ ,  $\mathbf{C} \in \mathbb{R}^{M \times K}$
- Verifying the right dimensions is an important sanity checker when working with matrices

# Matrix Inverse

- Definition for square matrices  $\mathbf{W} \in \mathbb{R}^{n \times n}$

$$\mathbf{W}^{-1}\mathbf{W} = \mathbf{W}\mathbf{W}^{-1} = \mathbf{I}$$

$$\mathbf{W}^{-1} = \frac{1}{\det \mathbf{W}} \mathbf{C}^T$$

where  $\mathbf{C}$  is the **cofactor matrix** of  $\mathbf{W}$ .

- If  $\mathbf{W}^{-1}$  exists, we say  $\mathbf{W}$  is **nonsingular**.

# Matrix Inverse

- A condition for invertibility is that **the determinant has to be different than zero**.
- For an intuition consider the following linear transformation matrix

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

- Applying this transformation to a vector gives

$$\mathbf{v}' = \mathbf{A}\mathbf{v} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = v_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + v_2 \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} v_1 \\ 0 \end{bmatrix} = \begin{bmatrix} v'_1 \\ v'_2 \end{bmatrix}$$

- This transformation removes one dimension from  $\mathbf{v}$  and projects it as a point along the first dimension.

# Matrix Inverse

- Can we from  $\mathbf{A}$  and  $\mathbf{v}' = [v'_1 \ v'_2]^T$  recover the initial vector  $\mathbf{v}$ ?
- We have the following linear system of equations

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} v'_1 \\ v'_2 \end{bmatrix} = \begin{bmatrix} v_1 \\ 0 \end{bmatrix}$$

- While there is only one solution for  $v_1$ , there are *infinitely many solutions* for  $v_2$ . This means we cannot recover the initial value of  $v_2$ .
- On the contrary, a nonsingular matrix, such as the identity matrix, admits one solution.

# Matrix Inverse

## ■ Example

$$\mathbf{W} = \begin{bmatrix} 1 & 1/2 \\ -1 & 1 \end{bmatrix}, \mathbf{W}^{-1} = \begin{bmatrix} 2/3 & -1/3 \\ 2/3 & 2/3 \end{bmatrix}$$

## ■ Verify it!

$$\mathbf{W}\mathbf{W}^{-1} = \begin{bmatrix} 1 & 1/2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 2/3 & -1/3 \\ 2/3 & 2/3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\mathbf{W}^{-1}\mathbf{W} = \begin{bmatrix} 2/3 & -1/3 \\ 2/3 & 2/3 \end{bmatrix} \begin{bmatrix} 1 & 1/2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

# Matrix Pseudoinverse

How can we invert a matrix  $\mathbf{J} \in \mathbb{R}^{n \times m}$  that is not squared?

■ Left-Pseudo Inverse  $\mathbf{J}^\# \mathbf{J} = \underbrace{(\mathbf{J}^T \mathbf{J})^{-1} \mathbf{J}^T}_{\text{left multiplied}} \mathbf{J} = \mathbf{I}_m$

- Works if  $\mathbf{J}$  has full column rank

■ Right-Pseudo Inverse  $\mathbf{J} \mathbf{J}^\# = \mathbf{J} \underbrace{\mathbf{J}^T (\mathbf{J} \mathbf{J}^T)^{-1}}_{\text{right multiplied}} = \mathbf{I}_n$

- Works if  $\mathbf{J}$  has full row rank

# Outline

1. Vectors

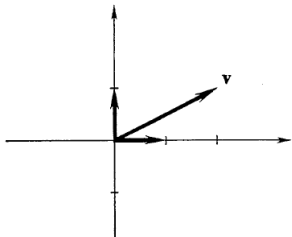
2. Matrices

**3. Operations and Linear Transformations**

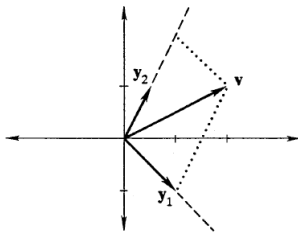
4. Wrap-Up

# Change of Basis

- Basis as **Unit Vectors**



- New Basis (vectors  $\mathbf{y}_1$  and  $\mathbf{y}_2$ )



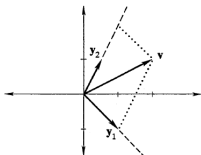
- Coordinates of vector  $\mathbf{v}$  in the original coordinate system (with unit basis vectors)

$$\mathbf{v} = c_1 \mathbf{y}_1 + \dots + c_n \mathbf{y}_n = \mathbf{Y} \mathbf{v}^*$$

- Where  $\mathbf{v}^*$  holds the coordinates in the **new** coordinate system.
  - To get the coordinates of  $\mathbf{v}^*$  (in the new basis) we just apply the inverse transformation



# Change of Basis - Example



- We have

$$\mathbf{y}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \mathbf{y}_2 = \begin{bmatrix} 1/2 \\ 1 \end{bmatrix}$$

- Thus

$$\mathbf{Y} = \begin{bmatrix} 1 & 1/2 \\ -1 & 1 \end{bmatrix}, \mathbf{Y}^{-1} = \begin{bmatrix} 2/3 & -1/3 \\ 2/3 & 2/3 \end{bmatrix}$$

$$\mathbf{v}^* = \mathbf{Y}^{-1}\mathbf{v} = \begin{bmatrix} 2/3 & -1/3 \\ 2/3 & 2/3 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = 2 \begin{bmatrix} 2/3 \\ 2/3 \end{bmatrix} + 1 \begin{bmatrix} -1/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

- $\mathbf{v}^*$  holds the **coordinates** in the **new basis**

# Change of Basis for a Linear Transformation

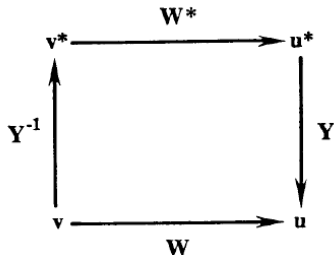
- We know

$$\mathbf{v} = \mathbf{Y} \mathbf{v}^* \quad \mathbf{u} = \mathbf{W} \mathbf{v} \quad \mathbf{u}^* = \mathbf{Y}^{-1} \mathbf{u}$$

- Plugging these together

$$\begin{aligned} \mathbf{u}^* &= \mathbf{Y}^{-1} \mathbf{u} \\ &= \mathbf{Y}^{-1} \mathbf{W} \mathbf{v} \\ &= \mathbf{Y}^{-1} \mathbf{W} \mathbf{Y} \mathbf{v}^* \\ &= \mathbf{W}^* \mathbf{v}^* \end{aligned}$$

$$\mathbf{W}^* = \mathbf{Y}^{-1} \mathbf{W} \mathbf{Y}$$



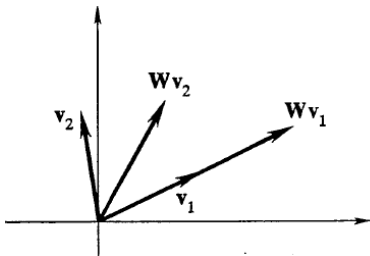
- To apply a transformation  $\mathbf{W}$  to the vector  $\mathbf{v}^*$  in the new basis:

1. Convert it to the unit basis:  $\mathbf{Y} \mathbf{v}^*$
2. Apply the transformation:  $\mathbf{W}(\mathbf{Y} \mathbf{v}^*)$

3. Convert the result back to the new basis space:  $\mathbf{Y}^{-1}(\mathbf{W}(\mathbf{Y} \mathbf{v}^*))$

# Eigenvectors and Eigenvalues

- Some vectors  $\mathbf{v}$  change only their length when multiplied by a matrix  $\mathbf{W}$



$$\begin{bmatrix} 4 & -1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$\begin{bmatrix} 3 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

- These vectors are called **eigenvectors** and the scaling factor is called **eigenvalues**.
- They obey the relation  $\mathbf{W}\mathbf{v} = \lambda\mathbf{v}$
- Eigenvectors are defined for a particular transformation matrix  $\mathbf{W}$ .

## Eigenvectors form a basis

- Let us assume there are  $n$  Eigenvectors and corresponding Eigenvalues

$$\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$$

$$\lambda_1, \lambda_2, \dots, \lambda_n$$

- **Theorem**

- For an  $n \times n$  matrix with eigenvectors  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ , if they correspond to *distinct* eigenvalues  $\lambda_1, \lambda_2, \dots, \lambda_n$ , then the set  $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$  is linearly independent.
- Hence, any vector can be expressed as a linear combination of eigenvectors

$$\mathbf{v} = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_n \mathbf{v}_n$$

# Eigenvectors form a basis

- This means that a transformation  $\mathbf{W}$  applied to a vector  $\mathbf{v}$  can be seen as a linear combination of eigenvectors

$$\begin{aligned}\mathbf{u} &= \mathbf{W} \mathbf{v} \\ &= \mathbf{W}(c_1 \mathbf{v}_1 + \dots + c_n \mathbf{v}_n) \\ &= c_1 \mathbf{W} \mathbf{v}_1 + \dots + c_n \mathbf{W} \mathbf{v}_n \\ &= c_1 \lambda_1 \mathbf{v}_1 + \dots + c_n \lambda_n \mathbf{v}_n\end{aligned}$$

# Linear transformations in Eigen-Basis

- For each eigenvector  $\mathbf{y}_i$ , we have

$$\mathbf{W} \mathbf{y}_i = \lambda_i \mathbf{y}_i$$

- We can summarize them in one equation

$$\mathbf{W} \mathbf{Y} = \mathbf{Y} \mathbf{\Lambda}$$

- In this case, if we apply  $\mathbf{W}$  we just stretch

$$\mathbf{W}^* = \mathbf{Y}^{-1} \mathbf{W} \mathbf{Y} = \mathbf{\Lambda}$$

- It is just a reformulation, but nice!

# Symmetric Matrix

## ■ Definition

- A **squared**  $n \times n$  matrix **A**, is a **symmetric** matrix iff

$$\forall i, j \quad a_{ij} = a_{ji}$$

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

## ■ Some properties

- The inverse  $\mathbf{A}^{-1}$  is also symmetric.
- **A** can be decomposed into  $\mathbf{A} = \mathbf{Q}\mathbf{D}\mathbf{Q}^T$ , where the columns of **Q** are the eigenvectors of **A**, and **D** is a diagonal matrix where the entries are the corresponding eigenvalues.

# Positive (semi-)Definite Matrix

## ■ Definition

- A **squared symmetric**  $n \times n$  matrix **A**, is a **positive definite matrix** if for any vector  $\mathbf{x} \in \mathbb{R}^n$

$$\mathbf{x}^T \mathbf{A} \mathbf{x} > 0$$

- Or **positive semidefinite** if  $\mathbf{x}^T \mathbf{A} \mathbf{x} \geq 0$
- These matrices are important in optimization and machine learning. For instance the covariance matrix is always positive semidefinite.



# Outline

1. Vectors

2. Matrices

3. Operations and Linear Transformations

**4. Wrap-Up**

## 4. Wrap-Up

You know now:

- What vectors and matrices represent
- Which operations you can do with vectors and matrices
- What eigenvectors and eigenvalues are
- How to perform a linear transformation

## Self-Test Questions

- Remember vectors and what you can do with them
- Remember matrices and what you can do with them
- What is a projection? How do you use it?
- How to compute the inverse of a matrix?
- What are Eigenvectors and Eigenvalues?
- What is a change of basis? What is a linear transformation? Are they the same?

# Homework

- Reading Assignment for next lecture
  - Bishop ch. 2
  - Murphy ch. 2
  - MacKay ch. 1, 2

# References

- If you want to grasp better the intuition behind Linear Algebra concepts
  - Essence of Linear Algebra by 3Blue1Brown:  
<https://goo.gl/9wFTgS>
- The Matrix Cookbook
  - <https://www.math.uwaterloo.ca/~hwolkowi/matrixcookbook.pdf>