KOMINFO digitalent

PRO Professional Academy

Digital Talent Scholarship 2022

Math For Machine Learning - PCA

Lead a sprint through the Machine Learning Track

Agenda

- Statistics of Datasets
- Inner Products

Are your students ML-ready?

Recap

Pelajaran hari ini

- Kita akan memahami PCA dari sudut pandang geometri.
- Kita akan belajar cara summarize datasets dengan statistik dasar.
- Perubahan mean dan variance saat kita shift atau scale original dataset.
- Akan ada jupyter notebook yang membantu pemahaman matematika dengan coding.

Introduction

Principal Component Analysis (PCA) adalah salah satu algoritma dimensionality reduction terpenting dalam Machine Learning.

PART 1

Statistics

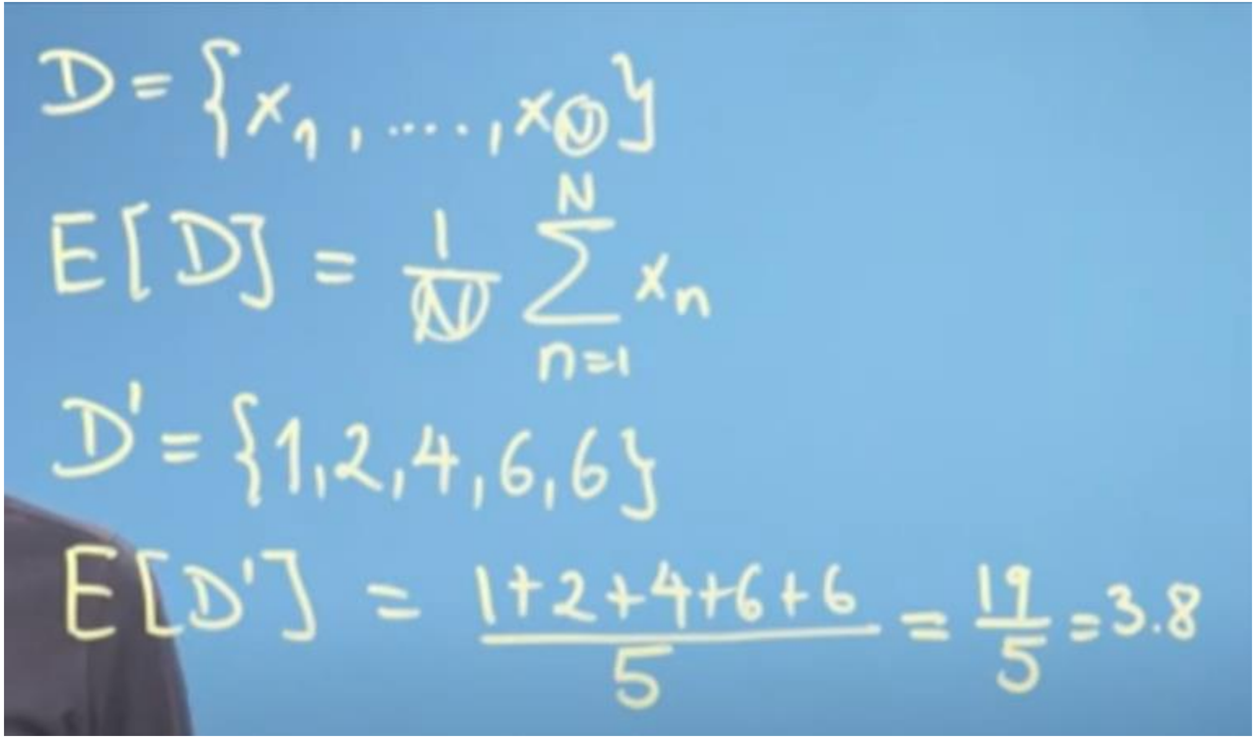
[Calculating the Mean, Variance and Standard Deviation, Clearly Explained!!! - YouTube]

[Covariance, Clearly Explained!!! - YouTube]

Objektif Pembelajaran

- Compute statistik dasar dari sebuah data sets
- Memahami efek dalam linear transformations terhadap means dan (co)variances
- Menghitung means/variances dari linearly transformed data sets
- Membuat kode yang dapat merepresentasikan gambar sebagai vector
- Membuat kode yang dapat menghitung statistik dasar sebuah datasets

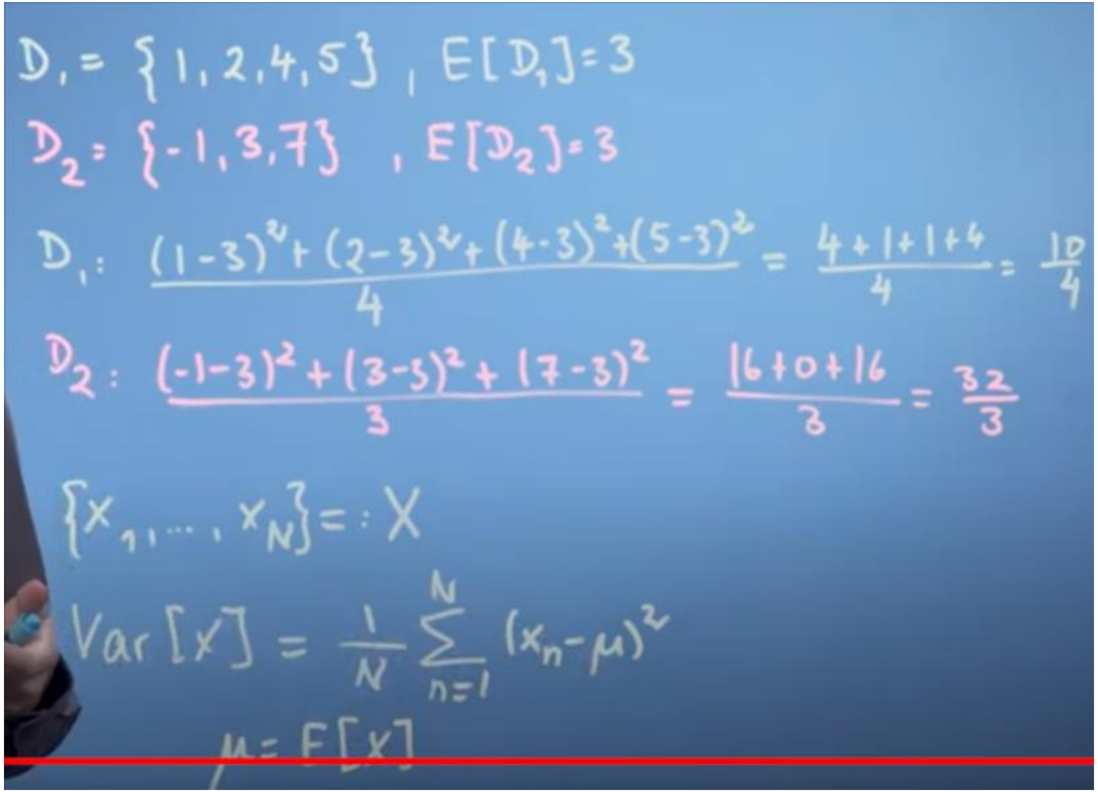
Mean



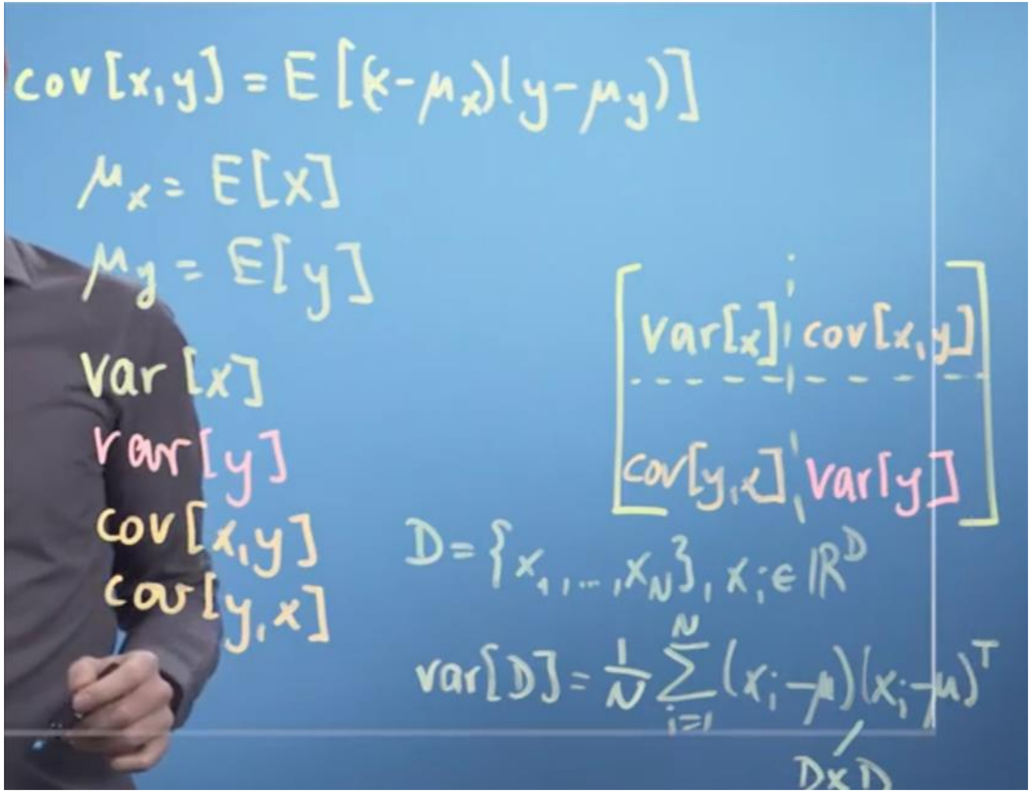
D = Data

E = Expected Value

Variance


$D_1 = \{1, 2, 4, 5\}, \quad E[D_1] = 3$
$D_2 = \{-1, 3, 7\}, \quad E[D_2] = 3$
$D_1: \frac{(1-3)^2 + (2-3)^2 + (4-3)^2 + (5-3)^2}{4} = \frac{4+1+1+4}{4} = \frac{10}{4}$
$D_2: \frac{(-1-3)^2 + (3-3)^2 + (7-3)^2}{3} = \frac{16+0+16}{3} = \frac{32}{3}$
$\{x_1, \dots, x_N\} =: X$
$Var[x] = \frac{1}{N} \sum_{n=1}^{N} (x_n - \mu)^2$
$\mu = E[x]$

Covariance


cov[x,y] = E[(x-μx)(y-μy)]
μx = E[x]
μy = E[y]
var[x]
var[y]
cov[x,y]
cov[y,x]
var[x] cov[x,y]
cov[y,x] var[y]
D={x1,...,xN}, xi ∈ IR^D
var[D] = 1/N Σ_{i=1}^{N} (xi-μ)(xi-μ)^T
DxD

[Range, variance and standard deviation as measures of dispersion | Khan Academy]

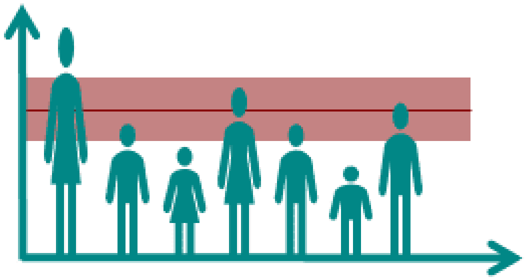
Spread

The 5 main values of measuring spread:

- Range, bottom to top
- Quartiles, bagi empat
- Interquartile range (IQR), Q1 - Q3
- Variance (σ^2),
- Standard deviation (σ), aka. Simpangan baku

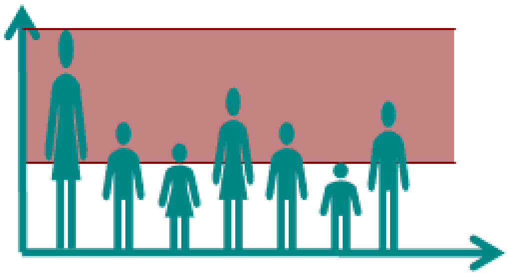
Spreads

Standard deviation / variance



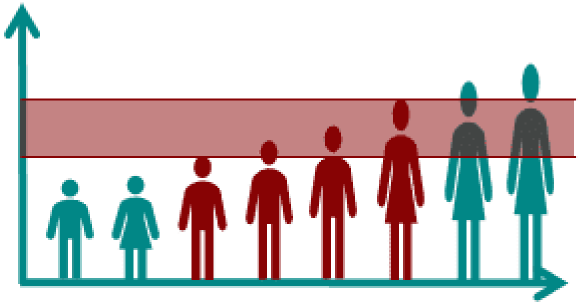
Average distance of all measured values from the mean value

Range



Distance between lowest and highest value of a distribution

Quantile distance



Spectrum in which the middle 50% of the values lie.

Difference between the first and the third quartile

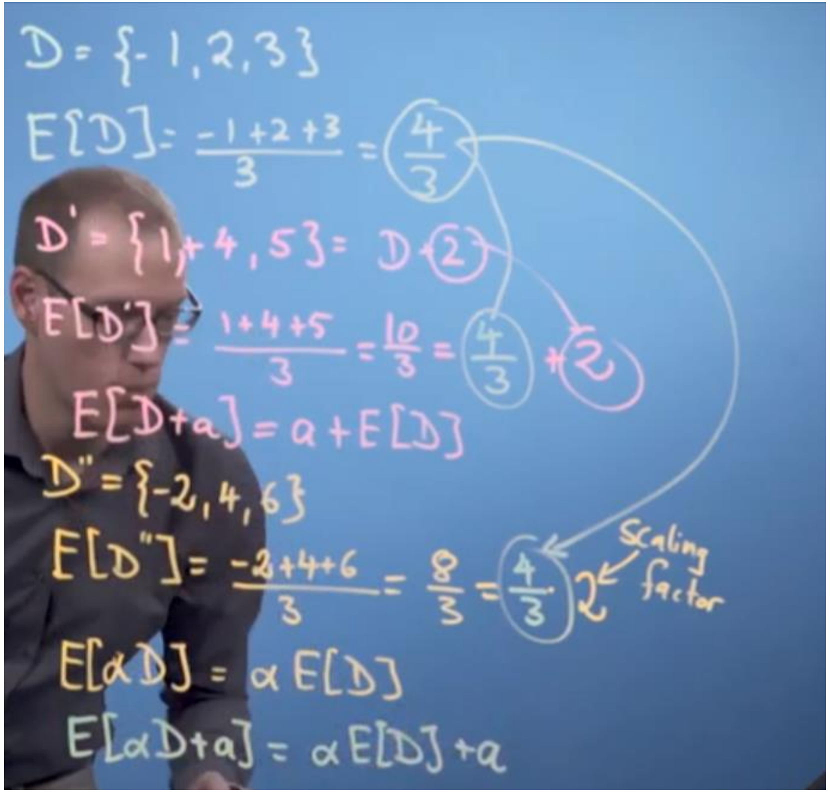
Variance and Standard Deviation

Variance and Standard Deviation Formula

	Population	Sample
Variance	$\sigma^2 = \frac{\sum_{i=1}^{N} (x_i - \mu)^2}{N}$	$S^2 = \frac{\sum_{i=1}^{n} (x_i - \overline{x})^2}{n - 1}$
Standard Deviation	$\sigma = \sqrt{\frac{\sum_{i=1}^{N} (x_i - \mu)^2}{N}}$	$S = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \overline{x})^2}{n - 1}}$

[Linear transformations | Matrix transformations | Linear Algebra | Khan Academy - YouTube]

Linear Transformation of Datasets



Linear Transformation of Datasets

$$\mathrm{var}[D] = \mathrm{var}[D + a]$$

$$\mathrm{var}[\alpha D] = \alpha^2 \mathrm{var}[D]$$

$$D = \{x_1, \ldots, x_N\}, \quad x_i \in \mathbb{R}^p$$

$$A x_i + b$$

$$\mathrm{var}[AD + b] = A \, \mathrm{var}[D] A^T$$

DEMO Mean Covariance Effect on Linear Transformation

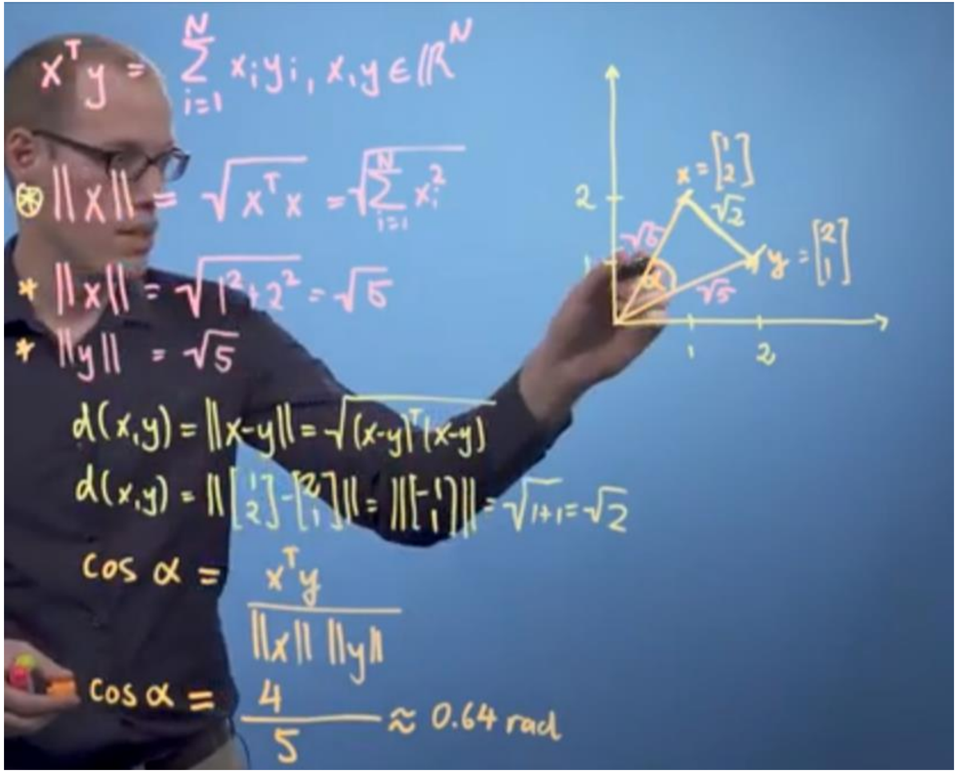
PART 2

Inner Product

Objektif Pembelajaran

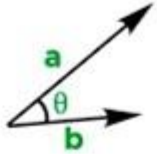
- Memahami apa itu inner products
- Menghitung sudut dan jarak dengan menggunakan inner products
- Menulis kode untuk menghitung jarak dan sudut antar data
- Memahami properti inner products
- Mengetahui bahwa orthogonality tergantung pada inner product

Dot Product


$x^T y = \sum_{i=1}^{N} x_i y_i, \ x, y \in \mathbb{R}^N$
$\|x\| = \sqrt{x^T x} = \sqrt{\sum_{i=1}^{N} x_i^2}$
$\|x\| = \sqrt{1^2 + 2^2} = \sqrt{5}$
$\|y\| = \sqrt{5}$
$d(x,y) = \|x-y\| = \sqrt{(x-y)^T (x-y)}$
$d(x,y) = \left\| \begin{bmatrix} 1 \\ 2 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\| = \left\| \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\| = \sqrt{1+1} = \sqrt{2}$
$\cos \alpha = \frac{x^T y}{\|x\| \|y\|}$
$\cos \alpha = \frac{4}{5} \approx 0.64$ rad
$x = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$
$y = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$
$\sqrt{2}$
$\sqrt{5}$
$\sqrt{5}$
α
1
2
1
2

Dot Product

The Vector Dot Product



$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos\theta$$

$$\cos\theta = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}||\mathbf{b}|}$$

[PCA - Unit 9 - Section 9.5 - Dot Product, Angle Bewteen Vectors and Work - YouTube]

Inner Product Definition and Properties

Def.

$x, y \in V$

$\langle \cdot , \cdot \rangle : V \times V \rightarrow \mathbb{R}$

- symmetric
- positive definite
- bilinear

Bilinear

$x, y, z \in V, \ \lambda \in \mathbb{R}$

$\langle \lambda x + z, y \rangle = \lambda \langle x, y \rangle + \langle z, y \rangle$

$\langle x, \lambda y + z \rangle = \lambda \langle x, y \rangle + \langle x, z \rangle$

Pos. def.

$\langle x, x \rangle \geq 0 , \ \langle x, x \rangle = 0 \Leftrightarrow x = 0$

Symmetry

$\langle x, y \rangle = \langle y, x \rangle$

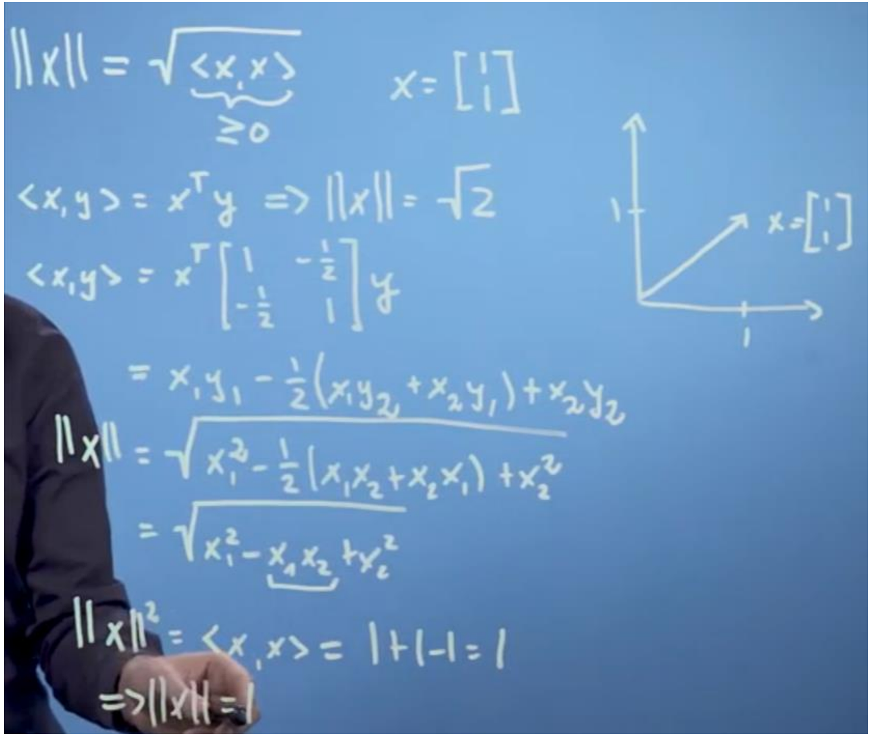
$\langle x, y \rangle = x^T I y$ ← dot product

$\langle x, y \rangle = x^T A y$

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

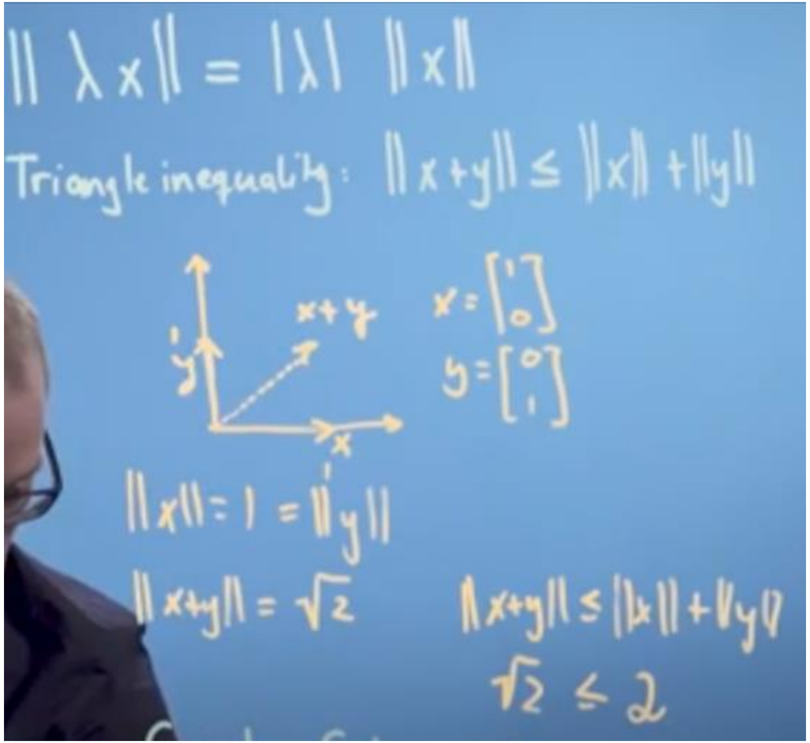
$\rightarrow 2x_1 y_1 + x_2 y_1 + x_1 y_2 + 2 x_2 y_2$

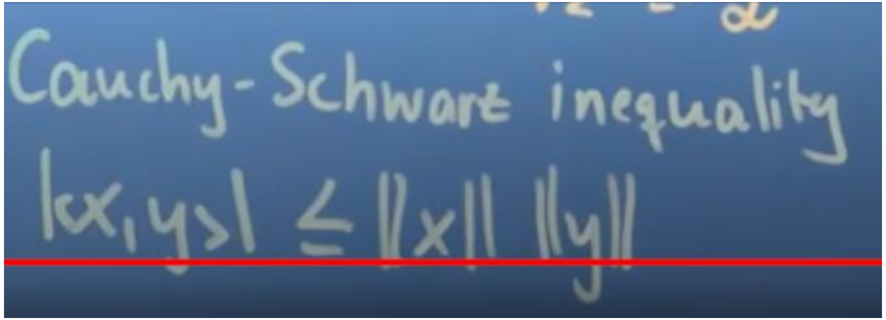
Inner Product: Length of Vectors



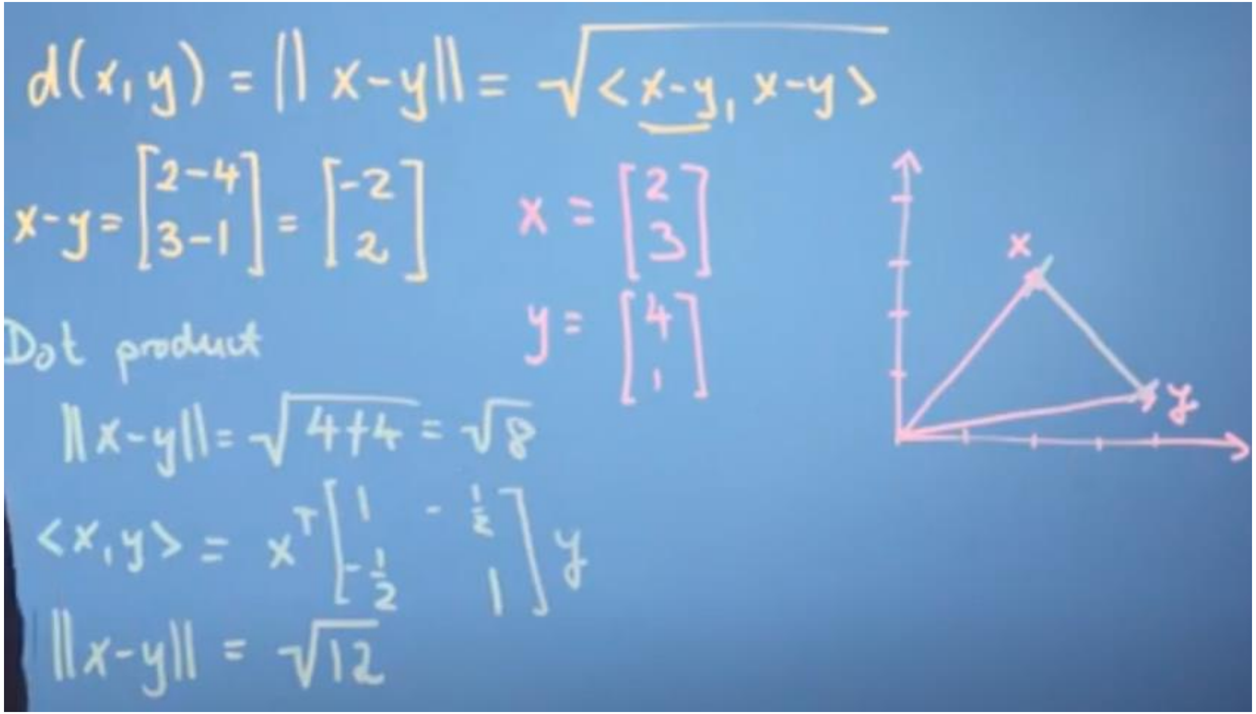
KOMINFO digitalent

Inner Product: Length of Vectors

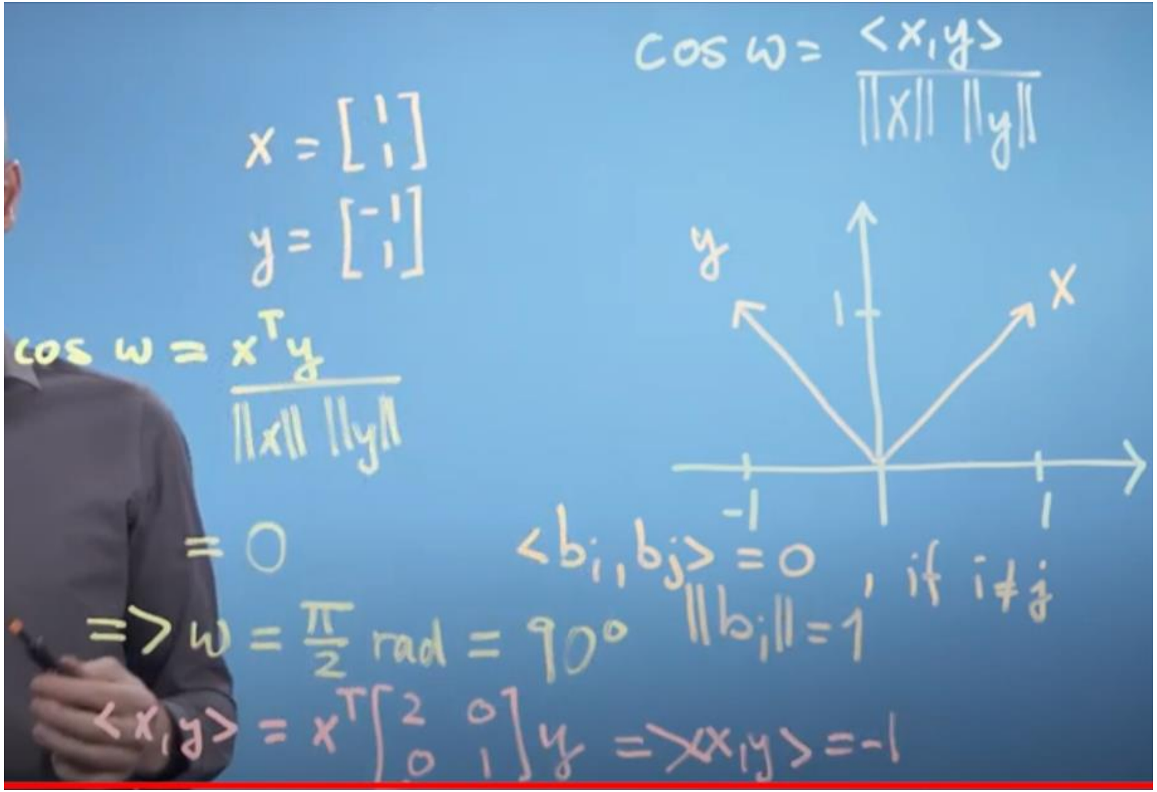




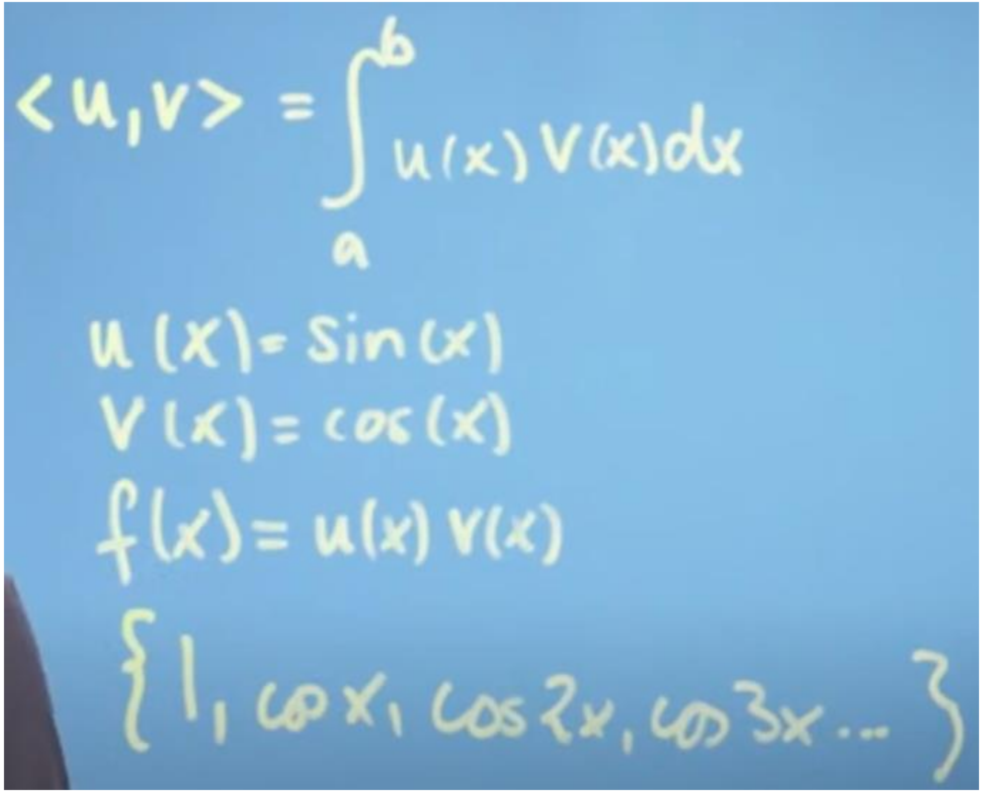
Inner Product: Distance between Vectors



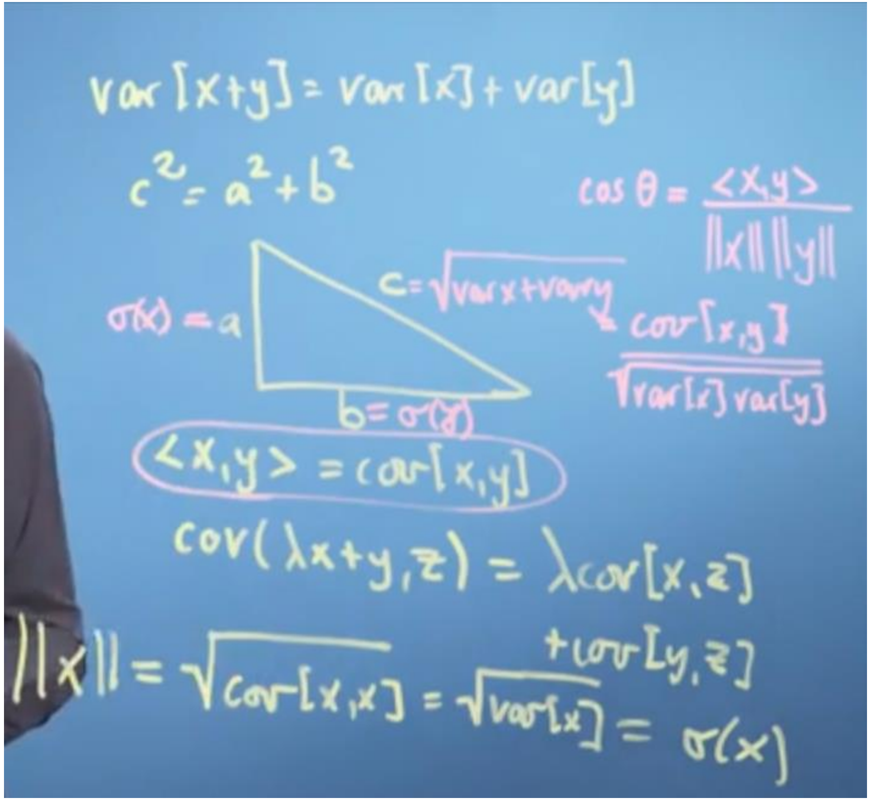
Inner Product: Angles and Orthogonality



Unconventional Inner Product for Function (Advanced)



Unconventional Inner Product (Advanced)



PCA by Imperiall Collage London

Q & A

Thank You