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

2×2

$\lambda_1 > \lambda_2$

$4 \qquad 1$

$\vec{u}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad u_2 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$

$A\vec{u} = \lambda \vec{u}$

$\underbrace{(A - \lambda I)}\vec{u} = \vec{0}$

$\det(A - \lambda I) = 0$

Characteristic equation

Power Iteration

dominant eigenvalue λ_1, $\vec{u}_1$ eigenvector

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

$x_0 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

$\hat{\lambda}_1 = 6/2 = 3 \qquad x_1 = \underline{Ax_0} = \begin{pmatrix} 5 \\ 6 \end{pmatrix} \longrightarrow \begin{pmatrix} 0.83 \\ 1 \end{pmatrix}$

$\hat{\lambda}_1 = \frac{22}{6} = 3.66 \qquad x_2 = Ax_1 = \begin{pmatrix} 21 \\ 22 \end{pmatrix} \rightarrow \begin{pmatrix} 0.95 \\ 1 \end{pmatrix}$

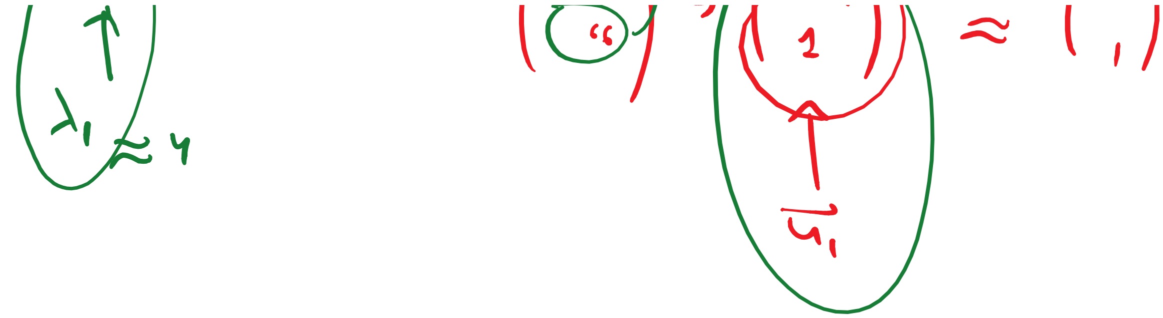
$\frac{86}{22} = 3.909 \qquad x_3 = Ax_2 = \begin{pmatrix} 85 \\ 86 \end{pmatrix} \rightarrow \begin{pmatrix} 0.988 \\ 1 \end{pmatrix}$

$\frac{342}{86} = 3.977 \qquad x_4 = Ax_3 = \begin{pmatrix} 341 \\ 342 \end{pmatrix} \rightarrow \begin{pmatrix} 0.997 \\ 1 \end{pmatrix}$

$\frac{1366}{342} = \underline{\underline{3.994}} \qquad x_5 = Ax_4 = \begin{pmatrix} 1365 \\ 1366 \end{pmatrix} \rightarrow \begin{pmatrix} 0.999 \\ 1 \end{pmatrix} \approx \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

$A\vec{u} = \lambda \vec{u}$

$A(5 \cdot \vec{u}) = \lambda(5 \cdot \vec{u})$



$A_1 A_2 \ldots A_d$ (d)

n D

$$A' = c_1 A_1 + c_2 A_2 + \cdots + c_d A_d$$

linear combination

$\underbrace{c_1, c_2, \ldots, c_d}$ you have to choose these

unsupervised → PCA ← max. variance

supervised LDA

direction that separates the class (POS vs. NEG)

non-linear features → kernel

$$c_1 A_1^2 + c_2 A_1 \cdot A_2 + c_3 \cdot \exp(A^3) + c_4 \cdot \log(A_4) + \ldots$$

1) linear

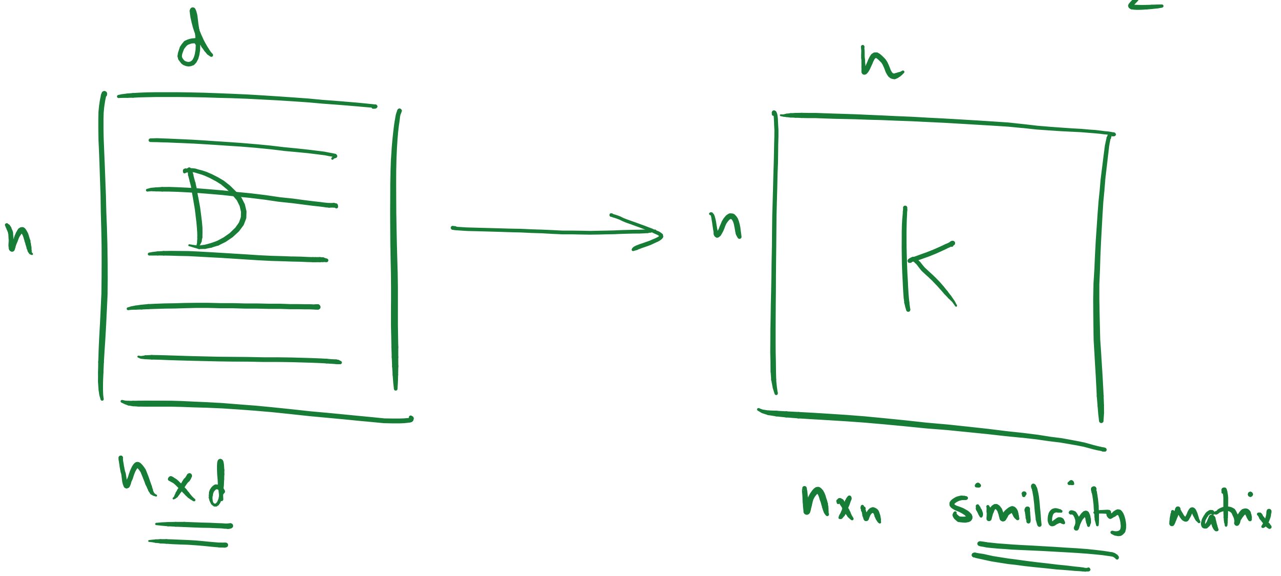
2) quadratic → exact powers A 2 $A_1^2,\ A_1 \cdot A_2$

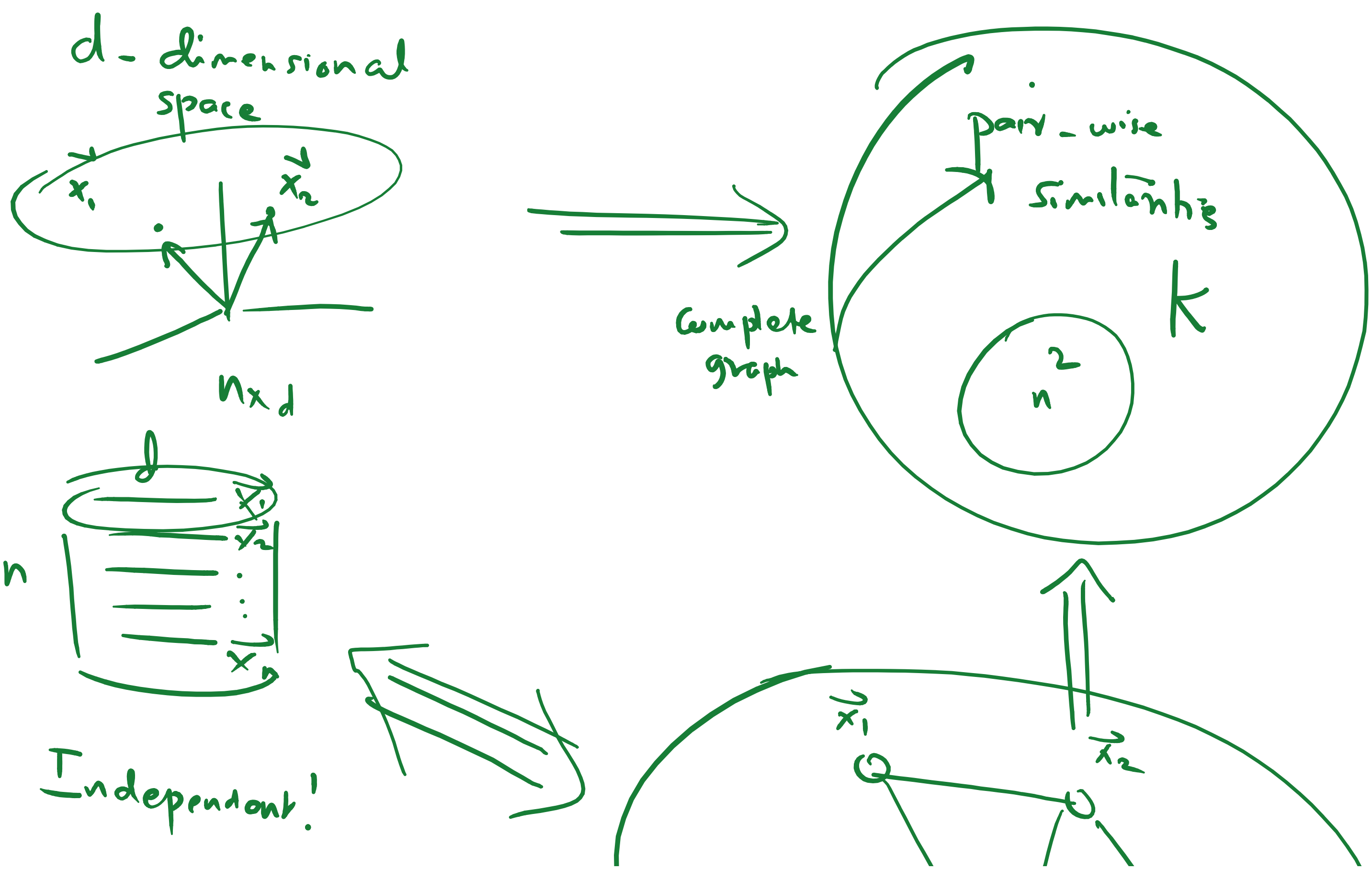
2) quadratic

homogeneous → exact powers of 2 A_1^2, $A_1 \cdot A_2$

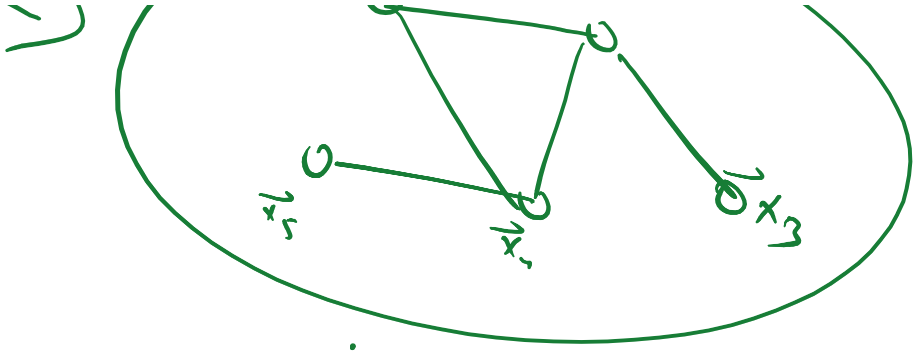
inhomogeneous → powers up to 2

$$\underbrace{1}_{0},\ \underbrace{A_1, A_2, \dots, A_d}_{1},\ \underbrace{A_1^2, A_2^2, \dots A_d^2, A_1 A_2 \dots}_{2}$$



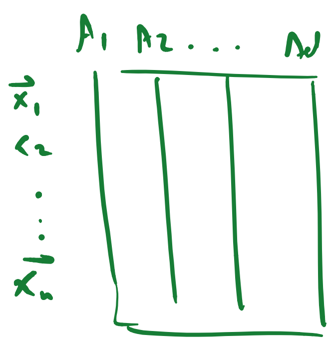


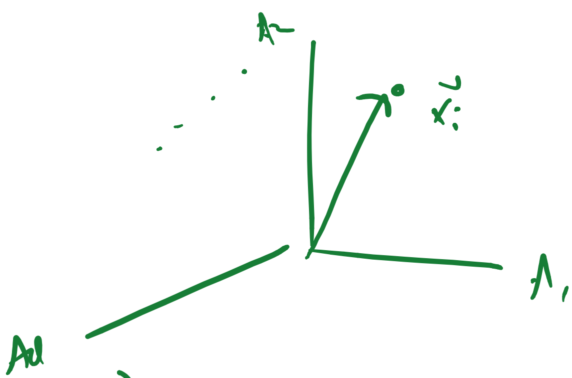
independent!



inter-linked (sparse)

$D \in \mathbb{R}^{n \times d}$





"extrinsic" d-dimensional space

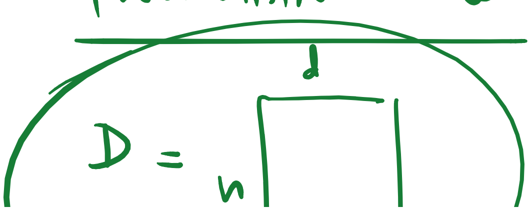
"true"? dimensionality → smaller subspace

$\text{rank}(D) = \#$ of linearly independent rows or columns

Cannot write one column as a linear combination of the others

$\text{rank}(D) \leq \min(n, d)$

Probabilistic View



$D = n\,\lfloor \;\; \rfloor$

data matrix is a sample

Compute statistics $(\hat{\mu})$

sample mean

?? μ

true mean

Unknown Universe or population

Each attribute A_i is a random variable

$A_i : O \to \mathbb{R}$

real value

Set of outcomes/objects

1) discrete (integer) $= \sum$
2) real / Continuous $= \int$

distribution of A

1) discrete : PMF = probability mass function

x of the R.V

$P(A = x)$: probability of x

Unknown !!

$$\boxed{\sum P(A = x) = 1}$$

$$\boxed{\sum_x P(A=x) = 1}$$

2) Continuous random variable A

$$\boxed{P(A=x) = 0 = \frac{1}{\infty}} \; !$$

PDF: probability density function

probability mass over an interval

$$P(a \le x \le b) = \int_a^b f(x)\,dx$$

$f(x)$ is the density function

dataset

D is a random sample

D is a set of IID random variables

population

A_1

A_2

$\vdots$

A_d

$\vec{x}_1$ x_{11}
$\vec{x}_2$ x_{21}
$\vdots$
$\vec{x}_n$ x_{n1}

random variable

IID → Independent
→ identically distributed

random sample of size $\underline{\underline{n}}$
Independent
ID like A_1

<u>statistic</u>

$$\hat{\theta} = f(x_{11}, x_{21}, \ldots x_{n1}) \in \mathbb{R}$$

estimate of the population parameter θ

A = Age

x_1
x_2
$\vdots$
x_n

Sample

Population

$$\mu = E[A]$$

$$\mu = \sum_x x \cdot P(A=x)$$

mean

$$\int_{-\infty}^{\infty} x \cdot f(x)\, dx$$

PDF

$$\boxed{\hat{\mu} = \frac{1}{n} \sum x_i}$$

mean

put a probability mass of $1/n$ on each point of $1/n$ on each point

PDF

"concentration"

median : middle most value

mode : most freq value

---

dispersion

sample variance?

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \hat{\mu})^2$$

$\hat{p}_{MF}$

statistic

asymptotically unbiased

$n \to \infty$

A

$$A = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

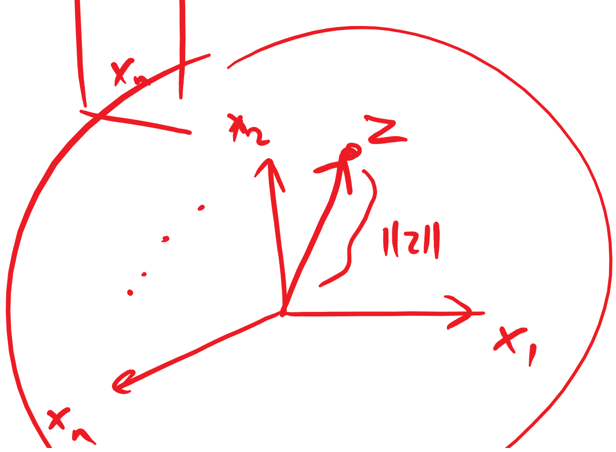
$A \in \mathbb{R}^n$

$$\sigma^2 = E[\underbrace{(A-\mu)^2}_{\text{expected squared deviation from mean}}]$$

$$= \sum_x (x-\mu)^2 \cdot \underbrace{P(A=x)}_{p_{MF}}$$

$Y = (A-\mu)^2$

$E[Y] \equiv$ variance of A



centering

Z centered vector

$$\begin{array}{c} z_1 \\ z_2 \\ \vdots \end{array} \begin{bmatrix} x_1 - \hat{\mu} \\ x_2 - \hat{\mu} \\ \vdots \end{bmatrix}$$

$z_n \; \big[\, x_n - \hat{\mu} \,\big]$

$$\hat{\sigma}^2 = \frac{1}{n}\left(\sum_{i=1}^{n} z_i^2\right) = \frac{1}{n} Z^T Z = \frac{1}{n}\underbrace{\|Z\|^2}_{\text{length!}}$$

avg. sq. length of Z

	A_1	A_2
$\vec{x}_1$	x_{11}	x_{12}
$\vec{x}_2$	x_{21}	x_{22}
$\vdots$		
$\vec{x}_n$		

joint probability distribution

$$P(A_1 = v_1,\ A_2 = v_2)$$