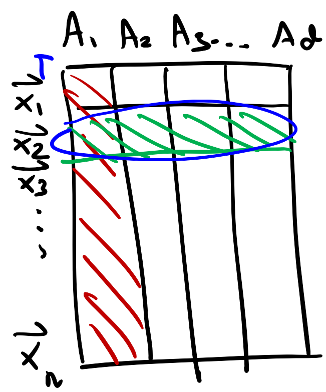
Linear Algebra → Gilbert Strang

$D =$



A_j : attributes / features / Variables

$d \equiv$ dimensionality

$\vec{x}_i$: data points, instances

$n \equiv$ # of points in the data D
(sample size)

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

Data / Data Matrix

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

$\vec{x}_i \in \mathbb{R}^d$

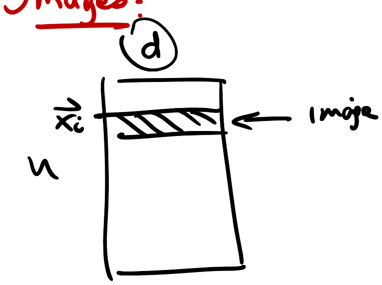
Attributes → Continuous / Discrete $\equiv$ Numeric

→ Symbolic (sex, Edu . . .)
Categorical

$\underbrace{\qquad\qquad\qquad}_{\text{Convert into numeric attributes}}$

$d \equiv k \times k$
square image

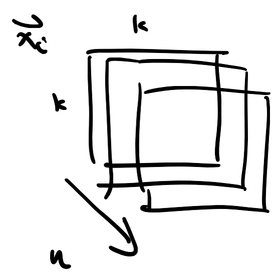
Images:



matrix
(2-way / 2D)

flattened data

$\Longrightarrow$



data tensor (3D)

Education :

→ Domain = { HS, BS, MS, PhD }

HS → 0, BS → 1, MS → 2, PhD → 3

avg = 2 ≡ MS ?

Domain → a set of binary variables

Edu-HS,	Edu-BS,	Edu-MS,	Edu-PhD
0/1	0/1	0/1	0/1

One-hot encoding

HS : (1, 0, 0, 0)

BS : (0, 1, 0, 0)

MS : (0, 0, 1, 0)

PhD : (0, 0, 0, 1)

	Edu
x_1	MS
x_2	HS
x_3	PhD

	Temp	HS	BS	MS	PhD	Press
x_1		0	0	1	0	
x_2		1	0	0	0	
x_3		0	0	0	1	

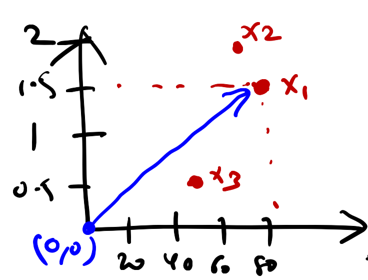
D

	A_1 Temp	A_2 Pressure
x_1	80.0	1.5
x_2	75.0	2
x_3	50.0	0.5
⋮		
x_n		

(y)

A_2 Pressure

scatter plot



Temp (x) A_1

$\mathbb{R}^2$

x_1 = point = (80.0, 1.5)

$\vec{x}_1$ ≡ vector (column) = $\begin{pmatrix} 80.0 \\ 1.5 \end{pmatrix}$

$\vec{x}_1^{T} = (80.0,\ 1.5)$

transpose operator

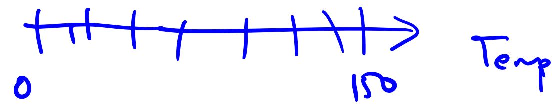
$$D = \begin{pmatrix} \vec{x}_1^{T} \\ \vec{x}_2^{T} \\ \vdots \\ \vec{x}_n^{T} \end{pmatrix} = \begin{pmatrix} | & | & & | \\ A_1 & A_2 & \cdots & A_d \\ | & | & & | \end{pmatrix}$$

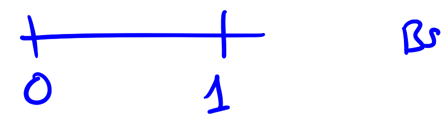
Numpy (deals with rows!)

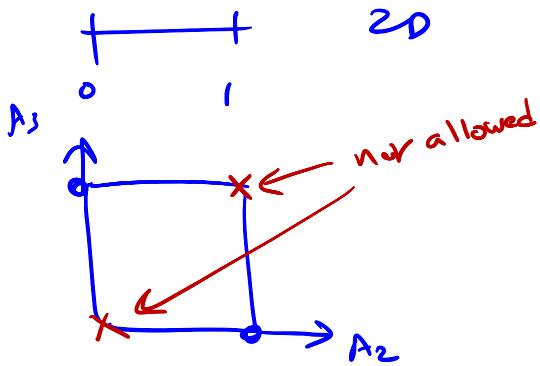
$$\vec{x}_1 = \begin{array}{c} \\ \mathbb{R} \\ A_1 \end{array}\begin{array}{c} \\ \left(97.8, \right. \\ \\ \end{array} \begin{array}{c} A_2\ A_3\ A_4\ A_5 \\ \left[\underbrace{0,1,0,0}_{\text{binary}}\right] \end{array} \begin{array}{c} \mathbb{R} \\ \left. 120 \right) \\ A_6 \end{array}$$

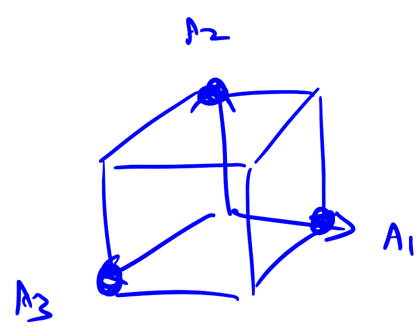
$A_1, A_6 \in \mathbb{R}$

$A_{2,3,4,5} \in \{0,1\}$

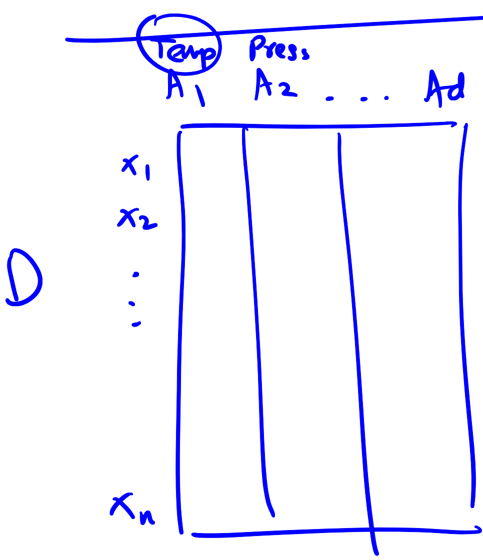








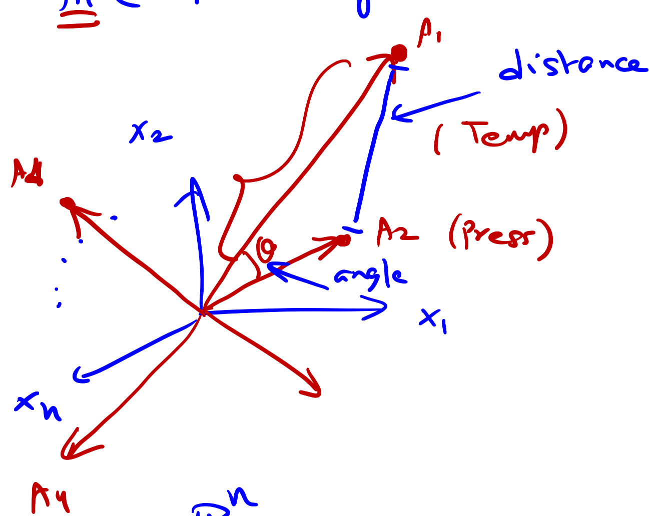
gap: binary vs continuous (relaxed)



Temp

$$\begin{pmatrix} 80.0 \\ 75.0 \\ 50.0 \\ \vdots \end{pmatrix}$$

$\underline{A_1} \in \mathbb{R}^n$ feature vector



$\mathbb{R}^n$