INTRODUCTION

COLOUR KEY
Helpful
Important
Definition
Theorem
Example

Engagement:
- Questions encouraged - use chat
- Break in middle
- Can you read my writing okay?

Recordings + notes to be released tonight.

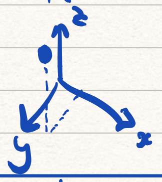
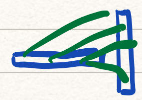
Session Breakdown

1: Theory	2: Practice	3: Application
- Preliminaries - Decompositions - The SVD - Low-rank approximations - Geometric interpretation	- Eigenvector recap - Existence of SVD - Comparison of implementations - The economy SVD - Optimal truncation	- Image compression - Linear regression - Dimensionality reduction - Wider uses - Limitations

Missing content: randomised SVD, alignment, RPCA, efficient algorithms

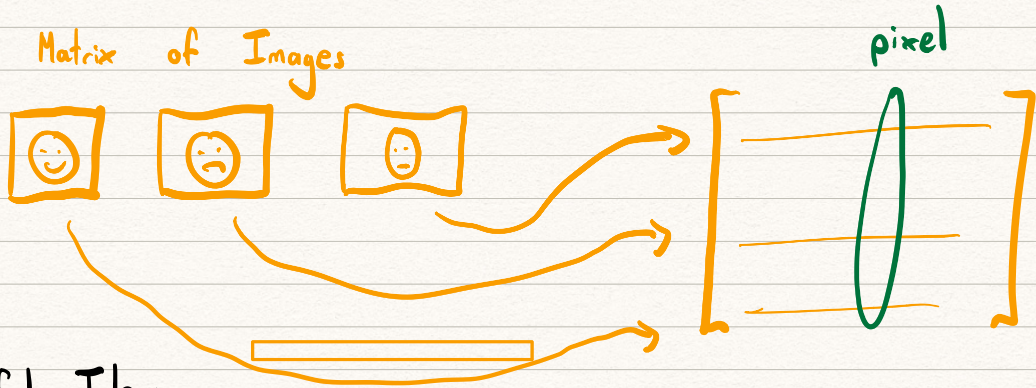
PRELIMINARIES

Definitions

Term	Notation	Geometric	Algebraic	Data-centric
Vector	$v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$ $v^T = (v_1\ v_2 \cdots v_n)$	Point $\mathbb{R}^n$ $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ 	Element of vector space. 8 axioms ("rules")	Data! Observation, predictor, response
Dot Product	$u \cdot v$ $u^T v$ $\langle u, v \rangle$	"Angle" between two vectors	$u \cdot v = \sum_{i=1}^{n} u_i v_i$ $u^T v =$ 	/
Matrix	$A_{(m \times n)} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$	Linear Transformation	Linear Map	$A = \begin{pmatrix} - & u_1^T & - \\ & \vdots & \\ - & u_m^T & - \end{pmatrix} = \begin{pmatrix} \vert & & \vert \\ v_1 & \cdots & v_n \\ \vert & & \vert \end{pmatrix}$
Rank	$\text{rank}(A)$	Dimension of a transformed sphere.	# of linearly independent columns / rows in matrix	Non-redundant observational predictors $\begin{array}{ccc} m_1 & m_2 & m_3 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{array}$

Orthogonal	/	Rotation/reflection	$A^TA = AA^T = I$ [A must square]	"Uncorrelated after centre" $\Sigma(x_i - \cancel{\bar{x}}^{\,0})(y_i - \cancel{\bar{y}}^{\,0})$

Ex: Matrix of Images



Helpful Ideas

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 7 & 10 \\ 15 & 22 \\ 23 & 34 \end{bmatrix}$$

(3×2) (2×2) ✓

(3×2)

$$\underset{(m\times n)}{A} \quad \underset{(n\times p)}{B} \quad \rightarrow \quad \underset{(m\times p)}{AB}$$

$$(AB)_{ij} = A_{i1}B_{1j} + A_{i2}B_{2j} + \cdots + A_{in}B_{nj} = \sum_{k=1}^{n} A_{ik}B_{kj} = a_i \cdot b_j$$

$$A = \begin{bmatrix} - & a_1^T & - \\ - & a_2^T & - \\ & \vdots & \\ - & a_m^T & - \end{bmatrix} \qquad B = \begin{bmatrix} | & | & & | \\ b_1 & b_2 & \cdots & b_p \\ | & | & & | \end{bmatrix}$$

$$AB = \begin{bmatrix} a_1\cdot b_1 & \cdots & a_1\cdot b_p \\ \vdots & \ddots & \vdots \\ a_m\cdot b_1 & \cdots & a_m\cdot b_p \end{bmatrix}$$

Outer Product

$$u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \qquad v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

$$u \otimes v = uv^T = \begin{bmatrix} u_1 v_1 & \cdots & u_1 v_n \\ u_2 v_1 & \cdots & u_2 v_n \\ \vdots & \ddots & \vdots \\ u_n v_1 & \cdots & u_n v_n \end{bmatrix}$$

$$C = \begin{bmatrix} | & | & & | \\ c_1 & c_2 & \cdots & c_n \\ | & | & & | \end{bmatrix} \qquad D = \begin{bmatrix} — & d_1 & — \\ & \vdots & \\ — & d_n & — \end{bmatrix}$$

$$\begin{aligned} CD &= c_1 \otimes d_1 + \cdots + c_n \otimes d_n \\ &= c_1 d_1^T + \cdots + c_n d_n^T \\ &= e_1 \cdot [\;] + e_2 \cdot [\;] + e_3 \underbrace{[\;]}_{\text{rank}=1} + \cdots + e_n [\;] \end{aligned}$$

$$CED$$

$$\| $$

$$\begin{bmatrix} e_1 & & & 0 \\ & e_2 & & \\ & & \ddots & \\ 0 & & & e_n \end{bmatrix}$$

DECOMPOSITIONS

$$x^2 - x - 6 = (x-3)(x+2)$$

A decomposition is a factorisation of a matrix into product

Ex. Eigendecomposition

$$A = PDP^{-1}$$

diagonal (D) invertible (P)

$$A^{100} = (PDP^{-1})^{100} = (PD\cancel{P^{-1}})(\cancel{P}DP^{-1}) \cdots (PDP^{-1})$$

$$= PD^{100}P^{-1}$$

Problems:

- Doesn't always exist
- Only makes sense square matrices

Singular Value Decomposition (SVD)

$$\underset{(p \times n)}{X} = \begin{bmatrix} | & | & & | \\ x_1 & x_2 & \cdots & x_n \\ | & | & & | \end{bmatrix}$$

$x_i \in \mathbb{R}^p$

$p >> n$

Importances

$$= U \Sigma V^T = \underset{(p \times p)}{\begin{bmatrix} | & & | \\ u_1 & \cdots & u_p \\ | & & | \end{bmatrix}} \underset{(p \times n)}{\begin{bmatrix} \sigma_1 & & 0 \\ & \sigma_2 & \\ 0 & & \ddots \; \sigma_n \\ \hline & 0 & \end{bmatrix}} \underset{(n \times n)}{\begin{bmatrix} - & v_1^T & - \\ & \vdots & \\ - & v_n^T & - \end{bmatrix}}$$

"eigen" face

Mixtures of $u_1 \cdots u_p$ needed to reconstruct x_1

"Time series" for how the contribution of an eigenface changes we observe

U, V orthogonal

$U^T U = U U^T = I$

$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 0$

Properties of SVD

- Exists
- Unique (up to sign)
- If X is real then U, V real
- # σ's not zero $\equiv$ rank(X)

$$X = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots \sigma_n u_n v_n^T$$

$$= \sigma_1 \left[\; \right] + \sigma_2 \left[\; \right] + \cdots + \sigma_n u_n v_n^T$$

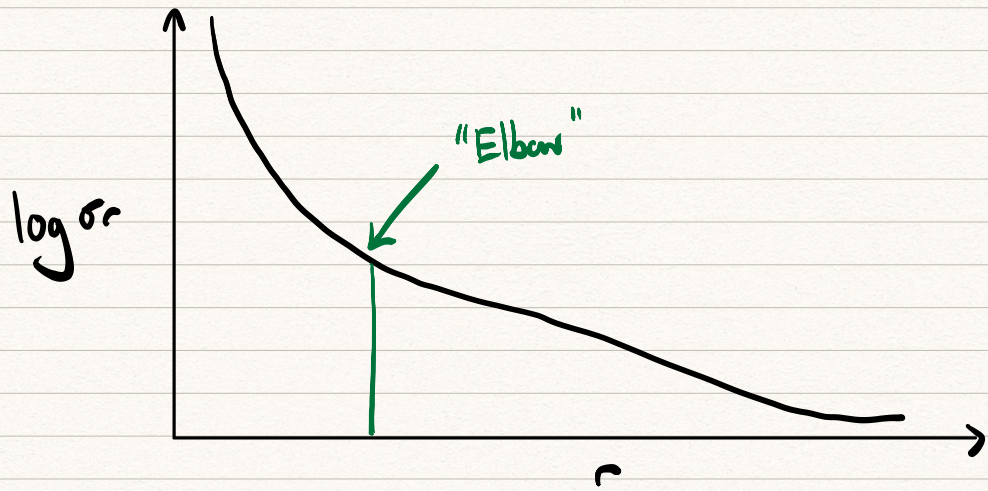
r

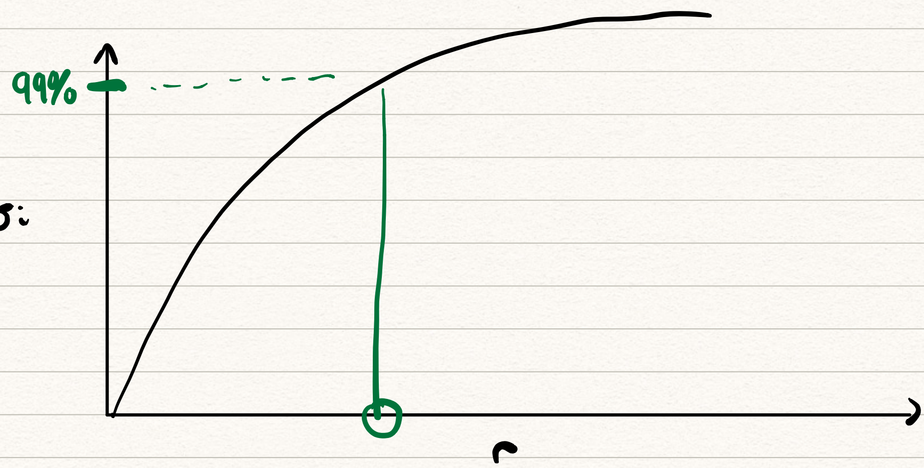
$$\approx \underset{(n \times r)}{\tilde{U}_r} \underset{(r \times r)}{\tilde{\Sigma}_r} \underset{(r \times p)}{\tilde{V}_r^T}$$

Eckart-Young [1936]

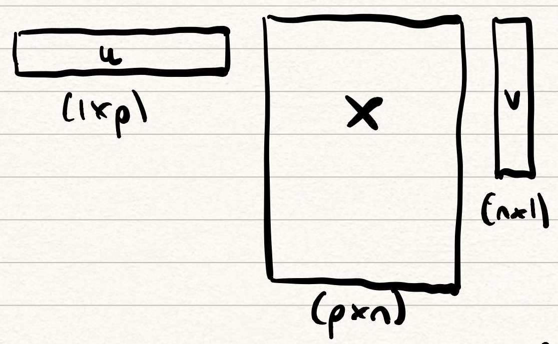
$$\underset{\substack{\tilde{X} \in \mathbb{R}^{p \times n} \\ \text{rank}(\tilde{X}) = r}}{\text{argmin}} \| X - \tilde{X} \|_F = \tilde{U}_r \tilde{\Sigma}_r \tilde{V}_r^T$$

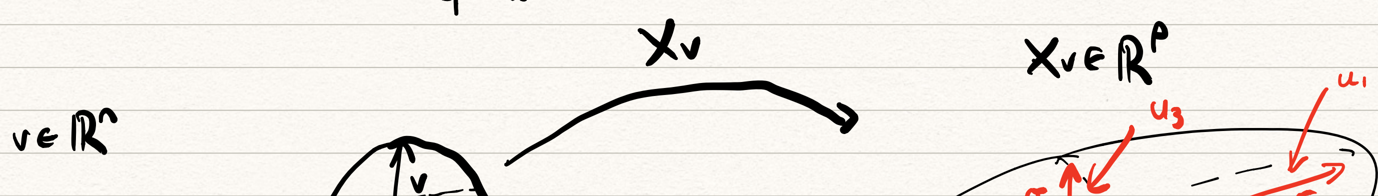
$$\|A\|_F = \sqrt{\sum (A_{ij})^2}$$

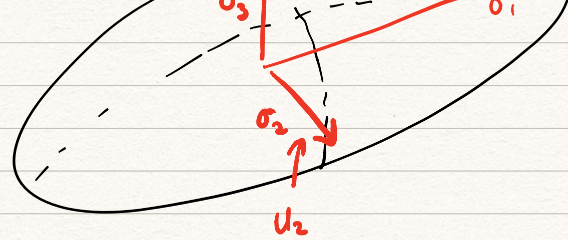


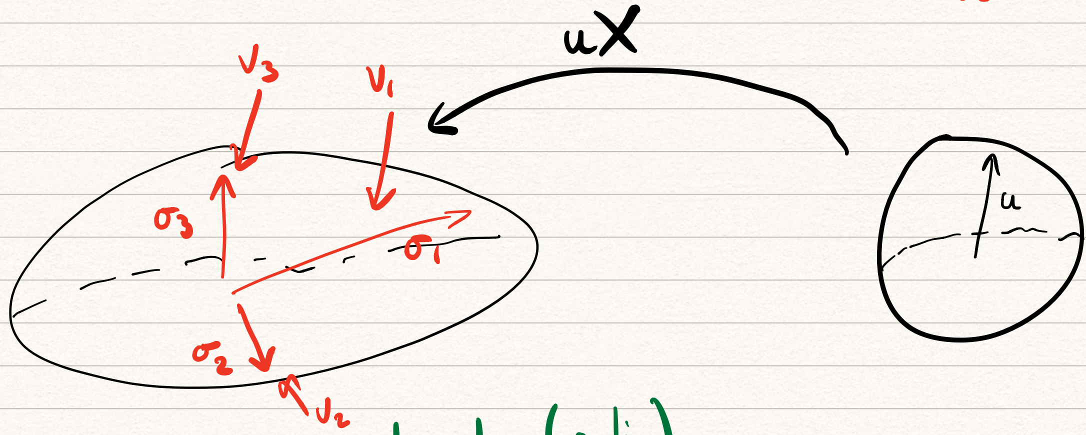


GEOMETRIC INTERPRETATION









$u \in \mathbb{R}^D$

diagonal (scaling)

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

orthogonal
(rotation/reflections)