Lecture 1

$D \equiv$ data matrix

Columns: $X_i \equiv$ attributes, features, variables (r.v.)

d columns

rows: $x_i =$ points, samples, observations, etc.

n points

D : n×d matrix

n×d points $\equiv$ d

Dimensionality reduction

Data lives in $\mathbb{R}^d$ but only needs $\mathbb{R}^2$

Image data matrix: flatten

All vectors are column vectors

$X_i \equiv$ attributes (d)

$x_j =$ point (n)

$x_i = \begin{pmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{in} \end{pmatrix} \in \mathbb{R}^n$

$x_j \in \mathbb{R}^d$ d-dimensional vector

Sample Table

Transpose

row vector

x_j^T = row vector

x_j and x_j^T = a column vector

$x_i^T \equiv$ row vector

d-dimensional space

Vector space

Temp Precip

$x_1 = \begin{pmatrix} 30.5 \\ 11.9 \end{pmatrix}$ $x_2 = \begin{pmatrix} 8.1 \\ -1.1 \end{pmatrix}$ 2-d dataset

x_3^T 21.6 8.3

$\text{Temp}(X_1) = \begin{pmatrix} 30.5 \\ 8.1 \\ \vdots \\ 21.6 \end{pmatrix}$ $\text{Precip}(X_2) = \begin{pmatrix} 11.9 \\ -1.1 \\ \vdots \\ 8.3 \end{pmatrix}$ 1d

$\mathbb{R}^2 \equiv (\mathbb{R}^n)^2$

Scatter plot

Precip (X_2)

Temp (X_1)

Point view (row-wise)

(Temp) X_1

Points in vector

X_1 (Temp)

Temp Precip

Column-wise

Σ vectors

X_1, X_2

in $\mathbb{R}^n$-dim space

dot product (notion of similarity)

$x_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ $x_2 = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$

$\langle x_1, x_2 \rangle \Rightarrow x_1 \cdot x_2 = x_1^T x_2 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}^T \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$

Geometric dot product

$= (x_1 \; y_1) \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$

Real dot product is commutative

$= x_1 x_2 + y_1 y_2$

$= x_2 x_1 + y_2 y_1$

$= x_2^T x_1$

$= x_2 \cdot x_1$

$a = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ $b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$

$a^T b = a_1 b_1 + a_2 b_2$

Scalar

Vector

matrix

tensor

Length (strongly tied to dot product)

$a = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$

Length $= \|a\| = \sqrt{a^T a} = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}$

Euclidean length

L_2-norm

$\|a\|^2 = a^T a$

$a = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ $\|a\| = \sqrt{1^2 + 2^2} = \sqrt{5}$

≈ 2.24

$\sqrt{1^2 + 2^2} = \sqrt{5}$

L_p-norm

$\|a\|_p = \left(|a_1|^p + |a_2|^p + \cdots + |a_n|^p \right)^{1/p}$

$a = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$

$L_1 \equiv$ taxicab $\|a\|_1 \equiv \|a\|_{TC}$ (Euclidean) $\|a\|_2 = \sqrt{25}$

$L_\infty \equiv \|a\|_\infty = \max_i |a_i| = \max_i |a_i|$ $\|a\|_\infty = \max(3, 4) = 4$

distance $a = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$ $b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$

$a = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ $b = \begin{pmatrix} 4 \\ 6 \end{pmatrix}$

$\|a - b\| = \left\| \begin{pmatrix} 1 \\ 2 \end{pmatrix} - \begin{pmatrix} 4 \\ 6 \end{pmatrix} \right\|$

$= \left\| \begin{pmatrix} -3 \\ -4 \end{pmatrix} \right\| = \sqrt{9 + 16}$

$= \sqrt{25}$

$\|a - b\| = \sqrt{(a-b)^T (a-b)}$

Similarity? (kernel methods)

Dimensionality Reduction

Angle: $\cos\theta$?

Unit vector : vector of length 1

$u_a = \frac{a}{\|a\|} = \frac{1}{\|a\|} a$

Unit vector for a

Scalar

$a = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$ $b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$

$\cos\theta = \left(\frac{a}{\|a\|} \right)^T \left(\frac{b}{\|b\|} \right)$

dot product between unit vector