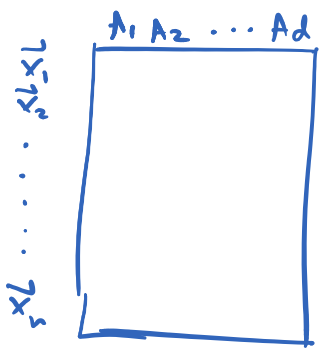
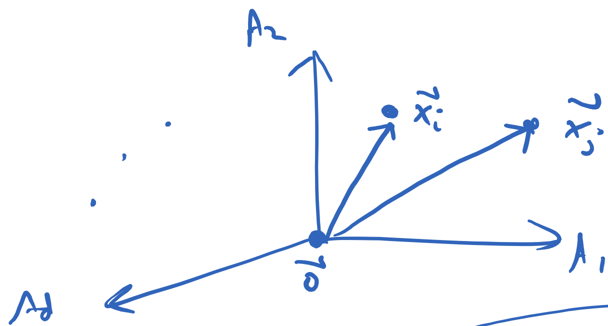


$D \in \mathbb{R}^{n \times d}$

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

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



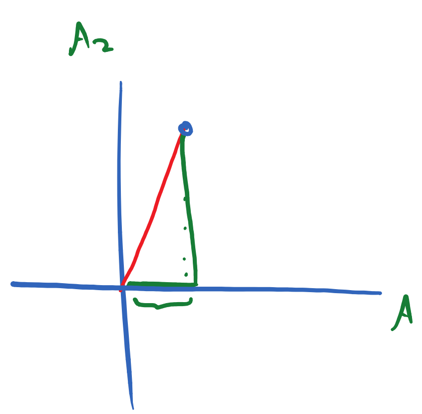
$\underbrace{L_2}_{\text{Euclidean}}$ Norm : $\|\vec{x}_i\|_2 = \sqrt{x_{i1}^2 + x_{i2}^2 + \dots + x_{id}^2} = \left(\sum_{j=1}^{d} |x_{ij}|^2\right)^{1/2}$

$\vec{x}_i = (x_{i1}, x_{i2}, \dots, x_{id})^T$

L_p-norm : $\|\vec{x}_i\|_p = \left(\sum_{j=1}^{d} |x_{ij}|^p\right)^{1/p}$

$\underbrace{L_1}_{\text{Manhattan}} = |x_{i1}| + |x_{i2}| + \dots + |x_{id}|$

$p \geq 0$



$x_1 \in \mathbb{R}^d = (x_{11}, x_{12}, \dots x_{1d})^T$

$x_2 \in \mathbb{R}^d = (x_{21}, x_{22}, \dots x_{2d})^T$

dot-product : $\vec{x}_i \cdot \vec{x}_j = \boxed{\sum_{j=1}^{d} (x_{ij} \cdot x_{2j})}$

	A_1	A_2
x_1	1	2

	A_1	A_2
x_1	1	2
x_2	2	3

$$\vec{x}_1 \cdot \vec{x}_2 = 1.2 + 2.3 = 2+6 = 8$$

$$\vec{x}_i \in \mathbb{R}^d \qquad \vec{x}_i \in \mathbb{R}^{d \times 1}$$

$$\vec{x}_i \cdot \vec{x}_j = \underline{\underline{x_i^T x_j}}$$

$$\vec{x}_i = \begin{pmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{id} \end{pmatrix} \qquad \vec{x}_j = \begin{pmatrix} x_{j1} \\ x_{j2} \\ \vdots \\ x_{jd} \end{pmatrix}$$

$d \times 1$ $\qquad$ $d \times 1$

$$\cancel{\vec{x}_i \times \vec{x}_j}$$

$$\vec{x}_i^T \times \vec{x}_j$$

$$\begin{pmatrix} x_{i1} & x_{i2} & \cdots & x_{id} \end{pmatrix} \qquad \begin{pmatrix} x_{j1} \\ x_{j2} \\ \vdots \\ x_{jd} \end{pmatrix}$$

$1 \times d$ $\qquad$ $d \times 1$

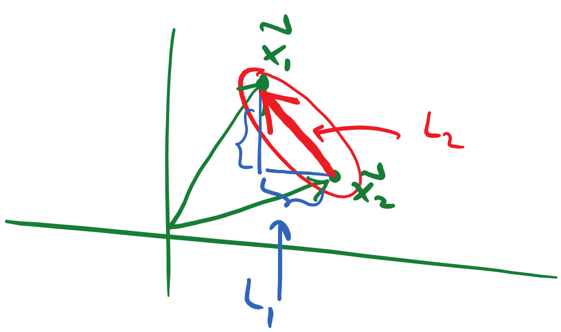
$= 1 \times 1$ or Scalar Value matrix

$$\|\vec{x}\|_2 = (\vec{x} \cdot \vec{x})^{1/2}$$

$$= (\vec{x}^T \vec{x})^{1/2}$$

$$\|\vec{x}\|^2 = \vec{x}^T \vec{x}$$

dot product give the squared magnitude



pair-wise

similarities

reverse

distance

$$d(\vec{x}_1, \vec{x}_2) = \sqrt{\underline{(x_{11} - x_{21})}^2 + \underline{(x_{12} - x_{22})}^2 + \cdots + \underline{(x_{1d} - x_{2d})}^2}$$

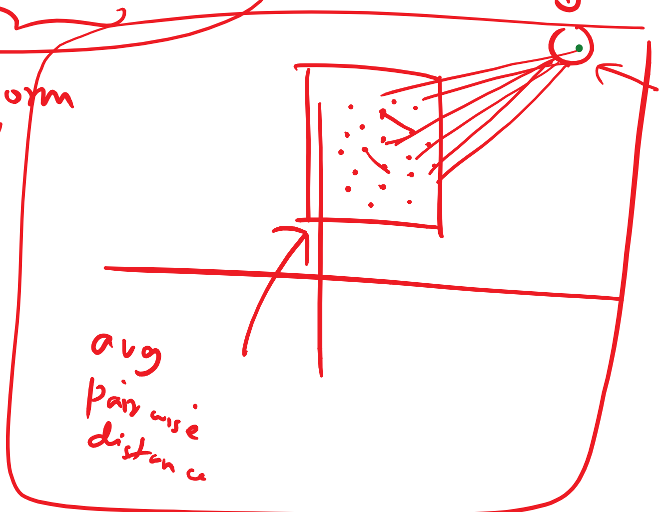
	A_1	A_2
$\vec{x}_1$	1	2

$\vec{x}_1$	1	2
$\vec{x}_2$	2	3

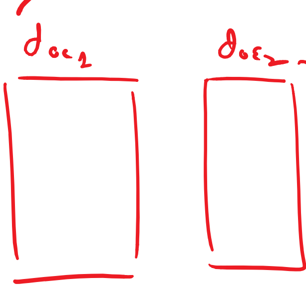
$$\sqrt{(1-2)^2+(2-3)^2} = \sqrt{1+1} = \sqrt{2}$$

$$d(\vec{x}_1, \vec{x}_2) = \underbrace{\|\vec{x}_1 - \vec{x}_2\|}_{\text{norm}}$$

Induces a distance



$L_2 \rightarrow$ Non-robustness



English

$$d_1 = (\overset{w_1}{1}\,\overset{w_2}{0}\,1\,0\,0\,0\,0\,0\,1\ldots)^T$$

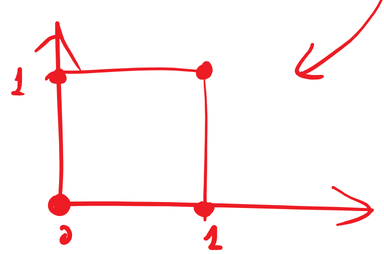
words in English $\leftarrow$ 1 million long feature set or attributes

$$d_2 = (1\,0\,0\,0\,0\,1\,0\,1\ldots)^T$$

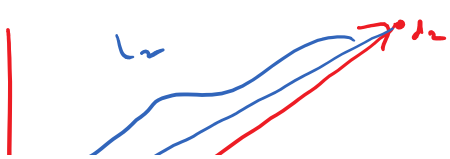
Sparse vectors

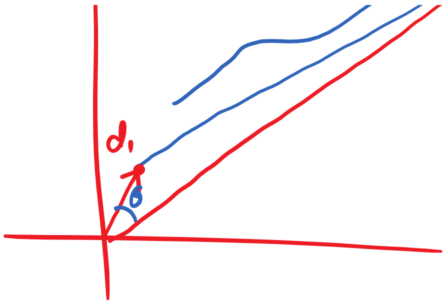
$$\text{dist} = \|d_1 - d_2\|$$

not very effective



nodes in a hypercube.





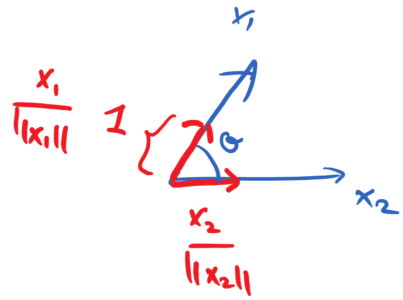
angle between two vectors

$$\cos\theta = \frac{\vec{x}_1 \cdot \vec{x}_2}{\|\vec{x}_1\| \cdot \|\vec{x}_2\|}$$

smallest angle

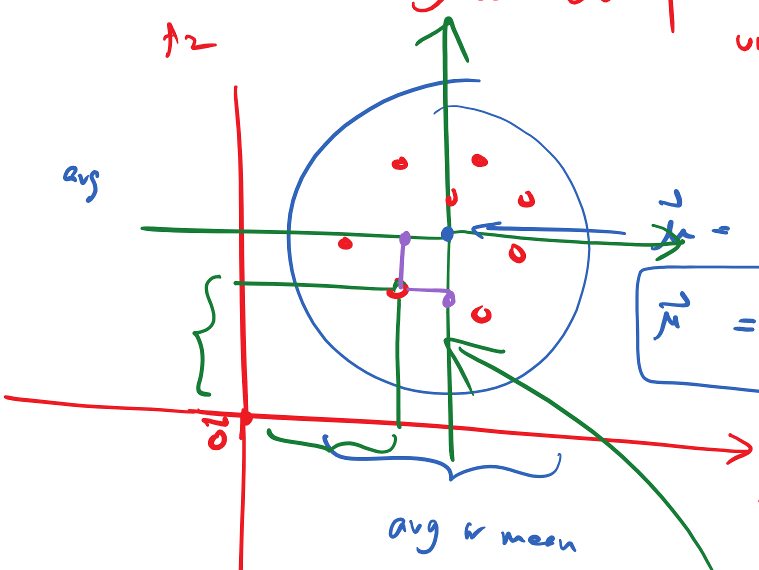
$\theta \in [0^\circ, 180^\circ]$

$$= \left(\frac{\vec{x}_1}{\|\vec{x}_1\|}\right)^T \left(\frac{\vec{x}_2}{\|\vec{x}_2\|}\right)$$



Unit vector in the direction of $\vec{x}_1$

the dot product between the unit vectors!



$$\vec{\mu} = \left(\text{avg } A_1,\ \text{avg } A_2\right)^T$$

$$\boxed{\vec{\mu} = \frac{1}{n}\left(\sum_{i=1}^{n} \vec{x}_i\right)}$$

mean

New co-ordinate frame

$\vec{z}_i = \vec{x}_i - \vec{\mu}$

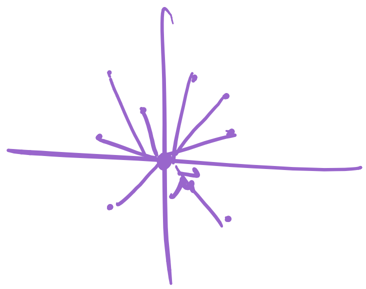
centered vector

$$\text{avg } z_i = \frac{1}{n} \sum_{i=1}^{n} \vec{z}_i = \vec{0}$$

total variance

$$\text{var}(D) = \frac{1}{n} \sum^{n} \left(\|\vec{x}_i - \vec{\mu}\|^2\right)$$

total variance



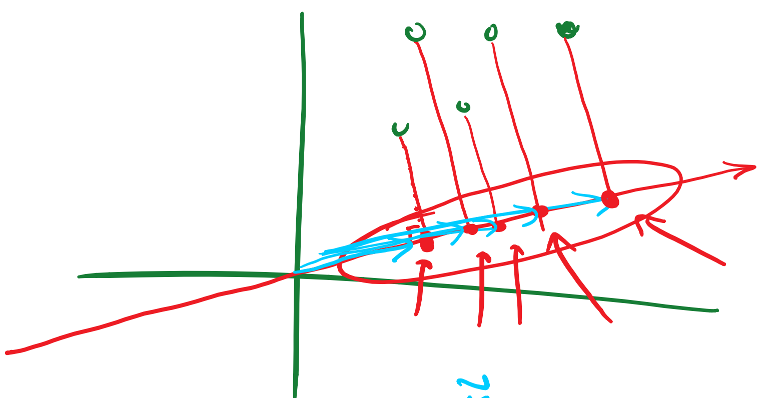
$$\mathrm{var}(D) = \frac{1}{n}\sum_{i=1}^{n}\left(\|\vec{x}_i - \vec{\mu}\|^2\right)$$

$$= \frac{1}{n}\sum_{i=1}^{n}\|z_i\|^2$$

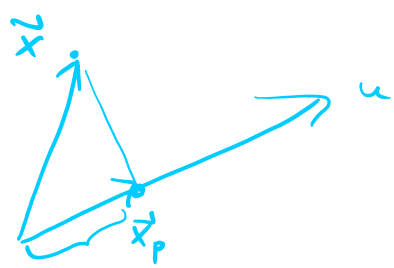
avg squared deviation from mean

$$\underline{\mathrm{MAD}}(D) = \frac{1}{n}\left(\sum_{i=1}^{n}\|\vec{x}_i - \vec{\mu}\|_1\right)$$

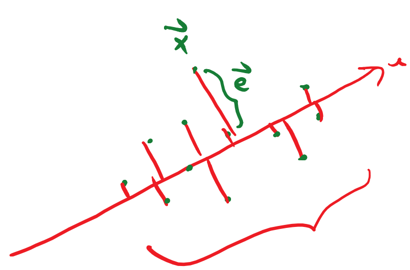
Mean absolute deviation



(new dimension)



$$\vec{x}_p = \left(\frac{\vec{x}\cdot\vec{u}}{\vec{u}\cdot\vec{u}}\right)\vec{u} = \underbrace{\left(\frac{\vec{x}^T\vec{u}}{\|u\|^2}\right)}_{\text{Scalar "length"}}\underbrace{\vec{u}}_{\text{direction}}$$



$\vec{u}$: optimal direction that maximize the variance

: minimize the squared errors

Eigenvectors are special $\underline{\underline{\text{directions}}}$

$A = d \times d$ matrix

on pairwise dimensions

$A = n \times n$ matrix

on pairwise points

$D = {}_n \overset{d}{\square}$

$A \in \mathbb{R}^{d \times d}$

every matrix has an associated linear operator or linear transformation

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

$$\underbrace{\underset{d \times d}{A}\ \underset{d \times 1}{\vec{x}}} = \underset{d \times 1}{\vec{y}}$$

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

$\vec{x} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

$$A\vec{x} = \underset{A}{\begin{pmatrix} 3 & 1 \\ 2 & 2 \end{pmatrix}} \underset{\vec{x}}{\begin{pmatrix} 1 \\ 2 \end{pmatrix}} = \underset{\vec{y}}{\begin{pmatrix} 5 \\ 6 \end{pmatrix}}$$

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

scalar

special "invariant" direction

$\equiv$ eigenvector

eigenvalue

$A \in \mathbb{R}^{d \times d}$

$\longrightarrow$ eigenvalue

$A \in \mathbb{R}^{d \times d}$

$\rightarrow$ d eigenvectors, corresponding eigenvalues

$\vec{u}$: ?

λ : ?

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

$$\left(A\vec{u} - \lambda \cdot I\,\vec{u}\right) = \vec{0}$$

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

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

$A = \begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix}$ $\qquad A - \lambda I = \begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$

$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ $\qquad = \begin{bmatrix} 3-\lambda & 1 \\ 2 & 2-\lambda \end{bmatrix}$

$$\det(A - \lambda I) = (3-\lambda)(2-\lambda) - 2 = 0$$

$$\Rightarrow 6 - 5\lambda + \lambda^2 - 2 = 0$$

$$\Rightarrow 1\lambda^2 - 5\lambda + 4 = 0$$

$$\begin{array}{ccc} a & b & c \\ 1 & -5 & 4 \end{array}$$

$$\boxed{\lambda = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}}$$

$$= \frac{5 \pm \sqrt{25 - 16}}{}$$

$$\lambda^2 - 4\lambda - \lambda + 4 = 0$$

$$\lambda(\lambda - 4) - 1(\lambda - 4) = 0$$

$$(\lambda - 1)(\lambda - 4) = 0$$

$$= \frac{5 \pm \sqrt{25 - 16}}{2}$$

$$\lambda = \frac{5 \pm \sqrt{9}}{2} = \frac{5 \pm 3}{2}$$

roots of the polynomial

$$\lambda_1 = 4 \qquad \lambda_2 = 1$$

$\vec{u}_1$? eigenvalue $\vec{u}_2$?

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

$$(A - \lambda I)\vec{u}_1 = 0$$

$$\begin{pmatrix} 3-\lambda & 1 \\ 2 & 2-\lambda \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ 2 & -2 \end{pmatrix} \vec{u} = \vec{0}$$

$\lambda = 4$

$$\vec{u} = \begin{pmatrix} x \\ y \end{pmatrix} \qquad \begin{pmatrix} -1 & 1 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$-x + y = 0$$

$$2x - 2y = 0$$

$$x = y$$

$$\lambda_1 = 4 \quad \vec{u}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\underline{A\vec{u}_1 = \lambda_1 \vec{u}_1}$$

$$\begin{pmatrix} 3 & 1 \\ 2 & 2 \end{pmatrix} \underbrace{\begin{pmatrix} 1 \\ 1 \end{pmatrix}}_{\vec{u}_1} = \begin{pmatrix} 4 \\ 4 \end{pmatrix} = \underbrace{4 \begin{pmatrix} 1 \\ 1 \end{pmatrix}}_{\lambda_1 u_1}$$

$$\lambda_2 = 1$$

$$\begin{pmatrix} 3-\lambda_2 & 1 \\ 2 & 2-\lambda_2 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$2x + y = 0$$

$$y = -2x$$

$$u_2 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$$

$$u_2 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$$

$$\begin{pmatrix} 3 & 1 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ -2 \end{pmatrix} = 1 \begin{pmatrix} 1 \\ -2 \end{pmatrix}$$

$$A \quad \vec{u}_2 \quad \lambda_2 \quad u_2$$

Power iteration

$$A = \begin{pmatrix} 3 & 1 \\ 2 & 2 \end{pmatrix} \qquad x_0 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

$$x_1 = A x_0$$

$$x_2 = A x_1 = A(Ax_0) = A^2 x_0$$

$$x_3 = A x_2 = A(A^2 x_0) = A^3 x_0$$

$$\vdots \qquad \qquad \vdots$$

$$\tilde{x}_n = \vec{u}_1$$

eigenvector $\longrightarrow$ eigenvalue