Professional program in AI

Introduction to vectors and matrices

May 30, 2019

Vectors

- An ordered list of numbers

$$\underline{a} = \begin{bmatrix} -2 \\ -5 \\ 1.3 \\ \sqrt{10} \end{bmatrix}$$

entries
components
coefficients
elements

$a_1 = -2$, $a_2 = -5$

$a_3 = 1.3$, $a_4 = \sqrt{10}$

- scalars

- the size of the list / number of entries is the size/length of the vector.

- $\mathbb{R}^n$: set of all n-vectors with entries from $\mathbb{R}$

  $\underline{a} \in \mathbb{R}^n$ means a is an n-vector

Vectors

- Equality of vectors $\underline{a}, \underline{b} \in \mathbb{R}^n$ $\quad \underline{a} = \underline{b}$ if $a_i = b_i$ for $i = 1, 2, \dots, n$

- zero vector, ones vector, canonical unit vectors

$$\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \qquad \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{bmatrix}, \dots$$

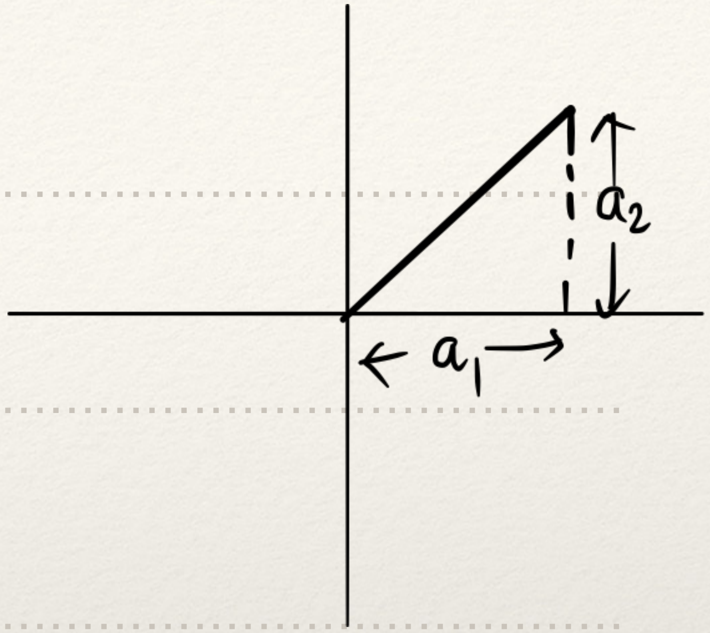
- Adding vectors, scalar multiple, linear combination

$$\underline{a} + \underline{b} \qquad \alpha \underline{a} \qquad \alpha \underline{a} + \beta \underline{b}$$

Vectors

Examples: location / displacement

"2-d"

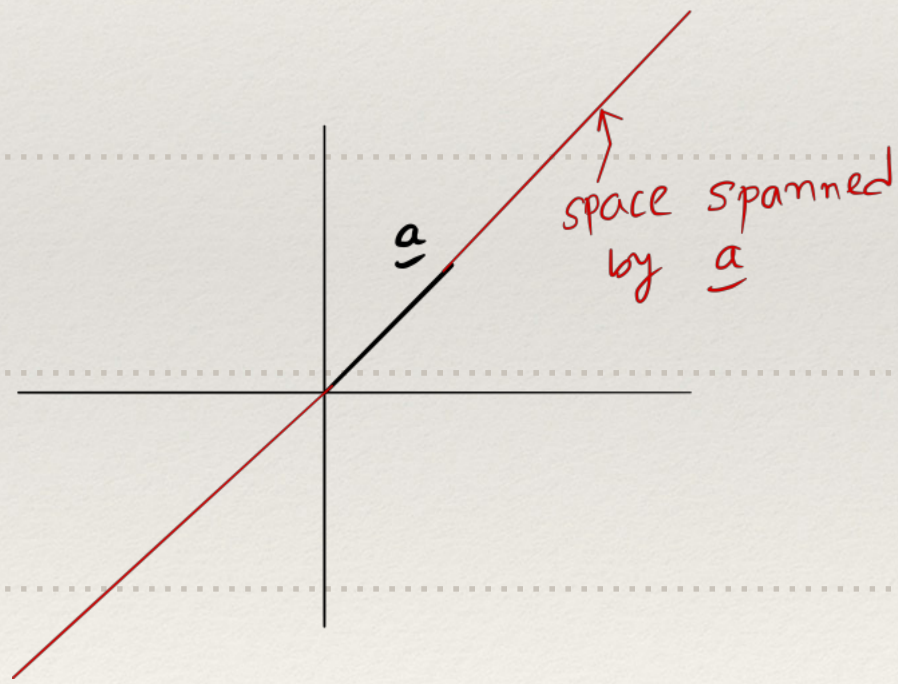


Other examples: time series, images, video, word count / histogram, portfolio, Features

Span

The space spanned by $v_1, v_2, \ldots, v_m \in \mathbb{R}^n$ is

$$\left\{ \sum_{i=1}^{m} \alpha_i v_i \;\middle|\; \alpha_i \in \mathbb{R} \right\}$$



Inner product

Inner product of vectors $\underline{v}_1 = \begin{bmatrix} v_{11} \\ v_{12} \\ \vdots \\ v_{1n} \end{bmatrix}$ and

$\underline{v}_2 = \begin{bmatrix} v_{21} \\ v_{22} \\ \vdots \\ v_{2n} \end{bmatrix}$ is given by

$$\underline{v}_1^T \underline{v}_2 = \langle \underline{v}_1, \underline{v}_2 \rangle = \sum_{i=1}^{n} v_{1i} \cdot v_{2i}$$

Examples: Sum, average, weighted average, polynomials, portfolio, sentiment analysis

Inner product

Linear functions can be written as inner products

- Taylor series

$$f(x_1, x_2, \dots, x_m) \approx f(0) + \sum_{i=1}^{m} \frac{\partial f}{\partial x_i} x_i$$

$$\approx f(0) + \nabla f^T \underline{x}$$

- Regression

$$\hat{y} = \underbrace{\underline{x}^T}_{\text{feature vector}} \underbrace{\underline{\beta}}_{\text{weights}} + \underbrace{v}_{\text{bias}}$$

Norm

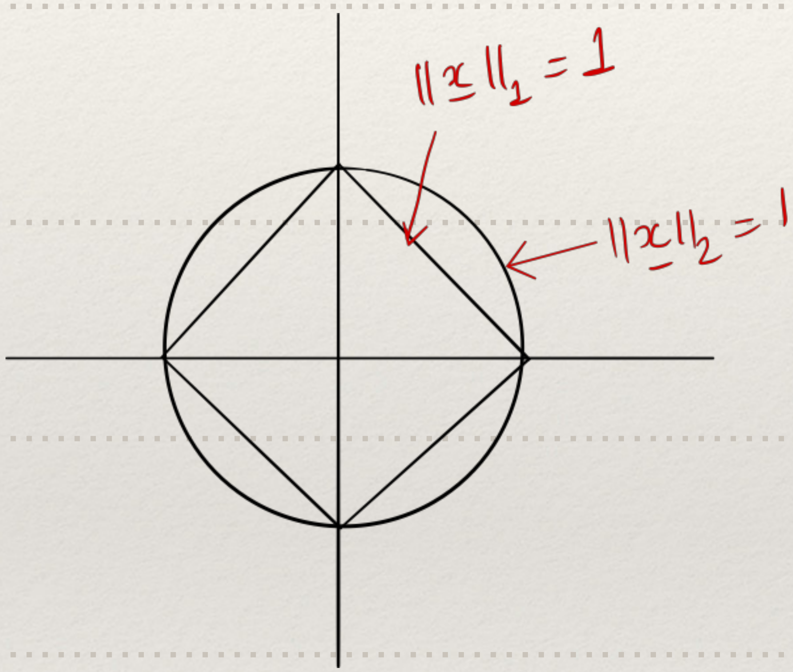
The (Euclidean) norm of $\underline{x} \in \mathbb{R}^n$ is

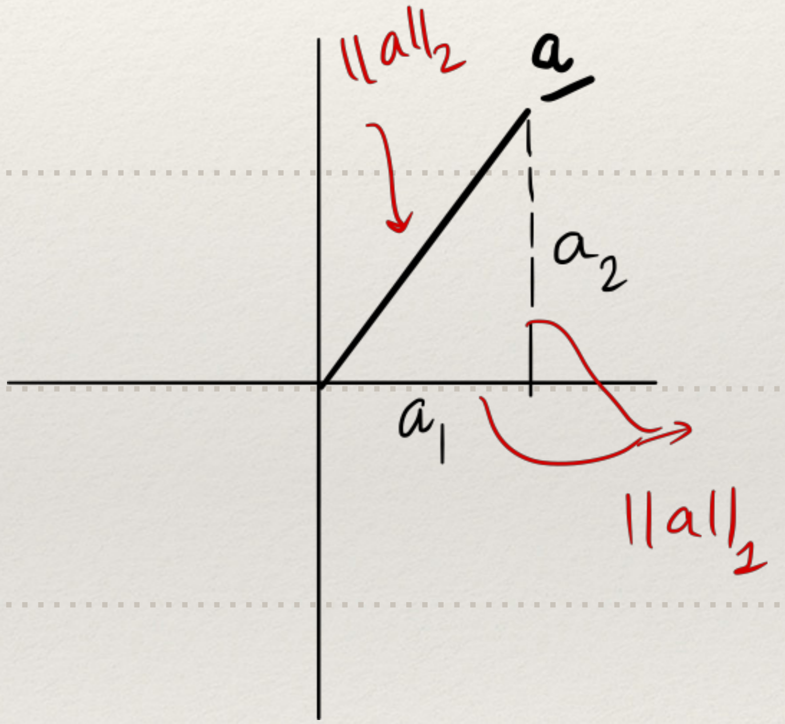
$$\|\underline{x}\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2} = \sqrt{\underline{x}^T \underline{x}}$$

Distance, homogenity, norm of sum, rms and variance, Triangle inequality

Other norms

$$\|\underline{x}\|_1 = \sum_{i=1}^{n} |x_i| , \qquad \|\underline{x}\|_\infty = \max_{i=1}^{n} \{|x_i|\}$$

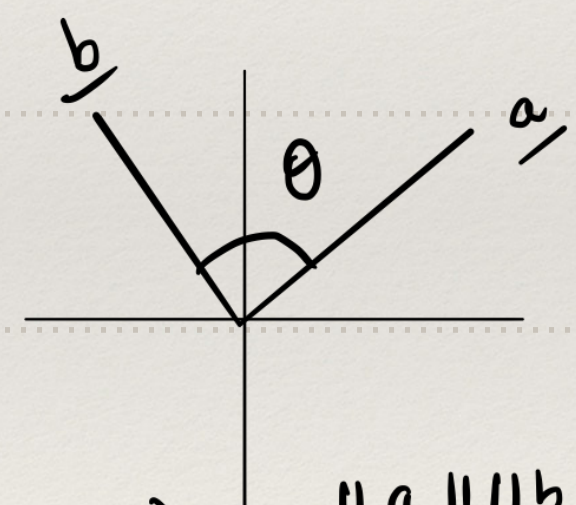




Cauchy Schwartz inequality

For any two vectors $\underline{v}_1, \underline{v}_2 \in \mathbb{R}^n$,

$$|\underline{v}_1^T \underline{v}_2| \leq \|\underline{v}_1\|_2 \|\underline{v}_2\|_2$$



$$(\underline{a}, \underline{b}) = \|\underline{a}\| \|\underline{b}\| \cos\theta$$

Orthogonal vectors

$\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_m \in \mathbb{R}^n$ is an orthogonal set of vectors if $v_i^T v_j = 0 \quad \forall\ i \neq j$.

Matrices

- Rectangular array of scalars

- rows / columns

$$\begin{array}{cc} & \text{Column } j \\ \text{row } i & \begin{bmatrix} & & \\ & A_{ij} & \\ & & \end{bmatrix} \end{array}$$

- $m \times n$: m rows, n columns

- Square, wide, tall matrices

- Block matrices

Examples

Matrix operations

- zero matrix, identity matrix, diagonal matrix

- Transpose and addition

- Symmetric matrices

- Matrix norm (Frobenius)

Matrix multiplication

Input - output / system / operator perspective

$$A\underline{x} \longleftarrow \boxed{A} \longleftarrow \underline{x}$$

$$A\underline{x} = \begin{bmatrix} A_1 & \overset{2^{nd}\text{ column of } A}{A_2} & \cdots & A_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \sum_{i=1}^{n} A_i \cdot x_i$$

- outer product, inner product

Matrix multiplication

- $A\underline{x}$ is a linear combination of columns of A
- Each entry of $A\underline{x}$ is an inner product of rows of A with $\underline{x}$
- $A \in \mathbb{R}^{m \times n}$ (m rows n columns),
  $x \in \mathbb{R}^{n}$

Matrix multiplication

$A \in \mathbb{R}^{m \times n}$ $\quad B \in \mathbb{R}^{n \times k}$, $\quad AB \in \mathbb{R}^{m \times k}$

- Each entry of AB is an inner product

- $AB = A\begin{bmatrix} B_1 & B_2 & \cdots & B_k \end{bmatrix} = \begin{bmatrix} AB_1 & AB_2 & \cdots & AB_k \end{bmatrix}$ (lin comb of columns of A)

- $AB = \begin{bmatrix} A_1^T \\ A_2^T \\ \vdots \\ A_m^T \end{bmatrix} B = \begin{bmatrix} A_1^T B \\ A_2^T B \\ \vdots \\ A_m^T B \end{bmatrix}$ (lin comb of rows of B)

- $AB = \begin{bmatrix} A_1 & A_2 & \cdots & A_n \end{bmatrix} \begin{bmatrix} B_1^T \\ B_2^T \\ \vdots \\ B_n^T \end{bmatrix} = \sum_{i=1}^{n} A_i B_i^T$ (sum of outer products)

Matrix spaces

Column space / image space : $\{Ax \mid x \in \mathbb{R}^n\}$

Row space : $\{x^T A \mid x \in \mathbb{R}^m\}$

Null space : $\{x \mid Ax = 0\}$

$im(A)$, $im(A^T)$ and $N(A)$

Every vector in $im(A^T)$ is orthogonal to every vector in $N(A)$.

Orthogonal matrices

$V \in \mathbb{R}^{n\times n}$ is an orthogonal matrix if

$$V^T V = V V^T = I$$

columns of V are orthogonal to each other,

so are the rows.

Matrix inverse

- A matrix B is the inverse of $A \in \mathbb{R}^{n \times n}$ if

$$AB = BA = I,$$ we denote B by A^{-1}.

- Inverse may not exist for all square matrices. If $A\underline{x} = \underline{b}$, then $\underline{x} = A^{-1}\underline{b}$

- $(AB)^T = B^T A^T$, $(AB)^{-1} = B^{-1} A^{-1}$

Examples: product of orthogonal matrices

Complexity

Inner product: $O(n)$

Linear combination: $O(n^2)$

Matrix mult : $O(n^3)$

Norm : $O(n)$

Inverse : $O(n^3)$

For matrices in $\mathbb{R}^{n\times n}$ and vectors in $\mathbb{R}^n$