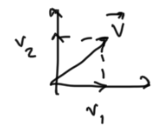
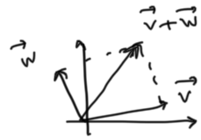
Primer on Linear Algebra

Vectors : tuple of elements

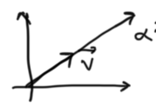
$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = [v_1, v_2]^T \quad , \quad \vec{v} \in \mathbb{R}^2$$



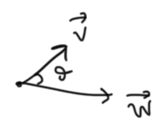
- vector addition : $\vec{v} + \vec{w} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} + \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} v_1 + w_1 \\ v_2 + w_2 \end{bmatrix}$



- multiplication by a scalar : $\alpha \in \mathbb{R}$, $\alpha \vec{v} = \begin{bmatrix} \alpha v_1 \\ \alpha v_2 \end{bmatrix}$



- elementwise multiplication : $\vec{v} \odot \vec{w} = \begin{bmatrix} v_1 w_1 \\ v_2 w_2 \end{bmatrix}$
- linear combination : $c_1 \vec{v} + c_2 \vec{w} = c_1 \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} + c_2 \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} c_1 v_1 + c_2 w_1 \\ c_1 v_2 + c_2 w_2 \end{bmatrix}$
- dot/inner product : $\vec{v}^T \vec{w} = \underset{1\times 2}{[v_1 \; v_2]} \underset{2 \times 1}{\begin{bmatrix} w_1 \\ w_2 \end{bmatrix}} = v_1 w_1 + v_2 w_2 = \sum_{i=1}^{2} v_i w_i$
- Norms :
  - Euclidean / L_2 / Length norm : $\|\vec{v}\|_2 = \sqrt{\vec{v}^T v} = \sqrt{v_1^2 + v_2^2}$
  - L_∞ / maximum norm : $\|\vec{v}\|_\infty = \max_i |v_i|$
  - L_1 norm : $\|\vec{v}\|_1 = \sum_i |v_i|$
- Unit vector : $\|\vec{u}\|_2 = 1$, $\vec{u} = \dfrac{\vec{v}}{\|\vec{v}\|_2}$
- Cosine similarity : $\cos\theta = \dfrac{\vec{v}^T \vec{w}}{\|\vec{v}\|_2 \|\vec{w}\|_2}$



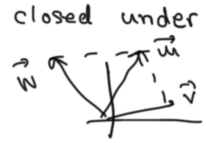
$\Longrightarrow$ $\vec{v}$ and $\vec{w}$ are orthogonal ($\vec{v} \perp \vec{w}$) iff $\begin{cases} \cos\theta = 0 \\ \vec{v}^T \vec{w} = 0 \end{cases}$

- Schwarz inequality : $|\vec{v}^T \vec{w}| \leq \|\vec{v}\|_2 \|\vec{w}\|_2$

⊛ Vector space / Linear Space : closed under <u>addition</u> + <u>scalar mult</u>

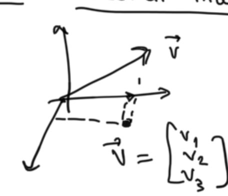
$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$, $\vec{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$, $\vec{v}, \vec{w} \in \mathbb{R}^2$

$\vec{v} \in \mathcal{V}$



closed under addition

$\mathbb{R}^2$



- Linear (in)dependence : Assume a finite collection of vectors :

$$\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k \in V$$

Then define the linear combination:

$$\vec{v} = c_1\vec{x}_1 + c_2\vec{x}_2 + \ldots + c_k\vec{x}_k = \sum_{i=1}^{k} c_i\vec{x}_i$$

<u>Definition</u>:

$\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k$ are linearly dependent if there exist a <u>non-trivial</u> linear combination for which $\sum_{i=1}^{k} c_i\vec{x}_i = 0$ with at least one $c_i \neq 0$.

If only the trivial solution exist, i.e. $c_1 = c_2 = \ldots = c_k = 0$ then $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k$ are linearly independent.

<u>In practice</u>:

- $\vec{x}_1, \ldots, \vec{x}_k$ are linearly dependent
  - if at least one of them is zero
  - if two vectors identical
  - if at least one of the $\vec{x}_i$ can be written as a linear combination of all the other vectors, i.e. $\vec{x}_i = \sum_{i \neq j} c_j x_j$
- if $\vec{x}_1, \ldots, \vec{x}_k$ vectors are linearly independent,

then they <u>form a basis</u> in V

$\Longrightarrow$ that any $\vec{v} \in V$ can be written as a linear combination of $\vec{x}_1, \ldots, \vec{x}_k$, i.e. $\vec{v} = \sum_{i=1}^{k} c_i\vec{x}_i$

<u>Matrices</u>: collections of vectors, linear systems of equations, linear map

$$\underset{m\times n}{A} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ \vdots & & \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}, \quad a_{ij} \in \mathbb{R}, \; i = 1, \ldots, m, \; j = 1, \ldots, n$$

columns

rows

$$A \in \mathbb{R}^{m\times n} \longrightarrow \vec{a} \in \mathbb{R}^{m \cdot n}$$

neighboring dimensions should match!

- <u>Multiplication</u>: $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{n\times k}$, $\underset{m\times k}{C} = \underset{m\times \underline{\underline{n}}}{A}\,\underset{\underline{\underline{n}}\times k}{B} \in \mathbb{R}^{m\times k}$

$$\underline{C_{ij}} = \sum_{l=1}^{n} a_{il} b_{lj}, \quad i = 1, \ldots, m, \; j = 1, \ldots, k$$

i.e. taking the dot-product of the i-th row of A and j-th column of B.

⊛ Not a commutative operation, i.e. $AB \neq BA$
Not an element-wise operation, i.e. $c_{ij} \neq a_{ij} b_{ij}$

- Addition: $A, B \in \mathbb{R}^{m \times n}$, $C = A + B \in \mathbb{R}^{m \times n}$, $c_{ij} = a_{ij} + b_{ij}$

- Identity matrix: $I_n = \begin{bmatrix} 1 & & & 0 \\ & 1 & & \\ & & 1 & \\ 0 & & & 1 \end{bmatrix} \in \mathbb{R}^{n \times n}$, $AI = IA = A$

- Associativity: $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{n \times p}$, $C \in \mathbb{R}^{p \times q}$, then

$$\underbrace{\overset{m \times p}{(AB)}\, \overset{p \times q}{C}}_{\substack{m\times n \ n \times p \\ m \times q}} = \underbrace{A\, \overset{n \times q}{(BC)}}_{\substack{m \times n \ n \times p \ p \times q \\ m \times q}}$$

- Distributivity: $A, B \in \mathbb{R}^{m \times n}$, $C, D \in \mathbb{R}^{n \times p}$

$$\begin{cases} (A+B)C = AC + BC \\ A(C+D) = AC + AD \end{cases}$$

- Inverse: $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{n \times m}$, if $\overset{m \times m}{AB} = I_m$
then $B := A^{-1}$ is the inverse of A.

Remark: Not all matrices are invertible!

- Transpose: $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{n \times m}$ if $b_{ij} = a_{ji}$ then B is the traspose of A, i.e. $B = A^T$

⊛ if A symmetric then $A = A^T$.

- Matrix rank: dimension of the vector space generated by its columns (or equivalently by its rows).
i.e. rank(A) := # linearly independent column vectors

- Nullspace: is formed by the remaining linearly dependent column vect

Linear Systems

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \end{aligned}\right\} \quad \underset{m\times n}{\overset{\text{known}}{A}}\ \underset{n\times 1}{\overset{\text{unknown}}{\vec{x}}} = \underset{n\times 1}{\vec{b}} \to \text{known}$$

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \\ a_{m1} & \dots & a_{mn} \end{bmatrix}, \quad \vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

In general, a linear system could have:

- infinitely many solution
- single unique solution
- no solution.

Elementary transformation: (they keep the solution the same, but they can transform the system into a simpler form

- Exchange two equation
- Multiplication with a non-zero constant.
- Addition of an equation (row) to another equation (row)

⊛ Are the basis for many analytical/numerical methods (e.g. Gauss elimination, LU decomposition)

- Determining the inverse A^{-1} is only possible when A is a square and invertible matrix (which is not often the case!)

Otherwise, under the assumption that A has linearly independent columns (i.e. is full-rank), then we can still solve the system:

$$\underset{m\times n}{A}\ \underset{n\times 1}{\vec{x}} = \underset{m\times 1}{\vec{b}} \Rightarrow \underbrace{\underbrace{A^T A}_{n\times m\ \ m\times n}}_{n\times n}\ \underset{n\times 1}{\vec{x}} = \underbrace{\underbrace{A^T \vec{b}}_{n\times m\ \ m\times 1}}_{n\times 1} \Rightarrow \vec{x} = \underbrace{(A^T A)^{-1} A^T}_{\text{Moore-Penrose pseudo-inverse}} \vec{b}$$