Mathematics for Machine Learning 2.4-2.5

def group: Consider a set G and an operation $\otimes : G \times G \to G$ defined on G. Then $G := (G, \otimes)$ is called a group if the following hold: $\otimes$: Kronecker Product

$A \otimes B = \begin{pmatrix} aB & bB \\ cB & dB \end{pmatrix}$

1) closure of G under $\otimes$: $\forall x, y \in G : x \otimes y \in G$
2) Associativity: $\forall x, y, z \in G : (x \otimes y) \otimes z = x \otimes (y \otimes z)$
3) neutral element: $\exists e \in G \, \forall x \in G : x \otimes e = x$ and $e \otimes x = x$
4) Inverse element: $\forall x \in G \, \exists y \in G : x \otimes y = e$ and $y \otimes x = e$, where e is the neutral element. we often write x^{-1} to denote the inverse element of x.
   The inverse element is defined with respect to the operation $\otimes$ and does not necessarily mean $\frac{1}{x}$
5) Abelian Group: $\forall x, y \in G : x \otimes y = y \otimes x$, then $G = (G, \otimes)$ is an Abelian group (commutative)

$\mathbb{N}$ - natural numbers $\mathbb{Z}$ - integers $\mathbb{R}$ - real numbers $\mathbb{I}$ - imaginary numbers

$(\mathbb{Z}, +)$ is an abelian group (positive integers)

$(\mathbb{N}_0, +)$ not a group, inverse elements are missing

$(\mathbb{Z}, \cdot)$ not a group. inverse elements for any $z \in \mathbb{Z}, z \neq \pm 1$ are missing

$(\mathbb{R}, \cdot)$ not a group, 0 does not posess an inverse element.

$(\mathbb{R}^{m \times n}, +)$, the set of $m \times n$ matrices is Abelian

$(\mathbb{R}^{n \times n})$, the set of $n \times n$ matrices
- closure and associativity follow from the definition of matrix multiplication.
- Neutral element: The identity matrix I_n is the neutral element with respect to matrix multiplication "$\cdot$" in $(\mathbb{R}^{n \times n}, \cdot)$
- Inverse element: If the inverse exists, (A is regular), then A^{-1} is the inverse element of $A \in \mathbb{R}^{n \times n}$, and in exactly this case $(\mathbb{R}^{n \times n}, \cdot)$ is a group. called the general linear group

Def (General Linear Group): The set of regular (invertible) matrices $A \in \mathbb{R}^{n\times n}$ is a group with respect to matrix multiplication is called the general linear group $GL(n,\mathbb{R})$. Since matrix multiplication is not commutative, the group is not Abelian.

Def (Vector Space): A real-valued vector space $V = (V, +, \cdot)$ is a set V with two operations

$+ : V \times V \to V$ vector addition

$\cdot : \mathbb{R} \times V \to V$ multiplication by scalars

where:

1.) $(V, +)$ is an Abelian group

2) Distributivity:

1) $\forall \lambda \in \mathbb{R}, x, y \in V : \lambda \cdot (x+y) = \lambda x + \lambda y$

"for all scalars λ that are Real numbers, vectors x, y that exist in the Vector Space V, $(x+y)$ scaled by λ = $\lambda x + \lambda y$"

2) $\forall \lambda, \psi \in \mathbb{R}, x \in V : (\lambda + \psi) \cdot x = \lambda \cdot x + \psi x$.

3) Associativity (outer): $\forall \lambda, \psi \in \mathbb{R}, x \in V : \lambda(\psi x) = (\lambda\psi)x$

4) Neutral element with respect to outer operation: $\forall x \in V : 1 \cdot x = x$

$\Rightarrow$ zero vector $[0, \dots 0]$

$\mathbb{R}^{n,\, n\times 1,\, 1\times n}$ allows us to write n-tuples as column vectors. x denotes a column vector

x^T, the transpose of x denotes a row vector

Def (Vector Subspace). Let $V = (V, +, \cdot)$ be a vector space and $\mathcal{U} \subseteq V$, $\mathcal{U} \neq 0$. Then $U = (\mathcal{U}, +, \cdot)$ is called vector subspace of V (or linear subspace) if U is a vector space with vector space operations $+, \cdot$ restricted to $\mathcal{U} \times \mathcal{U}$ and $\mathbb{R} \times \mathcal{U}$. We write $U \subseteq V$ to denote a subspace U of V.

To determine whether $(\mathcal{U}, +, \cdot)$ is a subspace of V we need to show:

1) $\mathcal{U} \neq \emptyset$ (empty set). in particular: $0 \in \mathcal{U}$

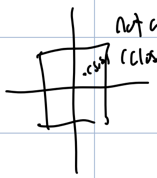
2) closure of U:

- with respect to outer operation: $\forall \lambda \in \mathbb{R}\ \forall x \in \mathcal{U} : \lambda x \in \mathcal{U}$
- with respect to inner operation: $\forall x, y \in \mathcal{U} : x + y \in \mathcal{U}$

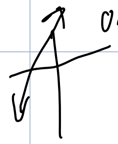
- For every vector space V, the trivial subspaces are V itself and $\{0\}$.

- A subspace of $\mathbb{R}^n$ <u>must</u> include 0: The homogenous system of linear equations $Ax=0$ with n unknowns is a subspace of $\mathbb{R}^n$. $Ax=b, b\neq 0$ is <u>not</u> a subspace of $\mathbb{R}^n$.

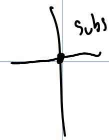
- The intersection of arbitrarily many subspaces is a subspace itself.



not a subspace on $\mathbb{R}^2$ (closure) $(2,2)$ scaled by 100 no longer in subset. closure broken.



0 not in subset. not a subspace



subspace. closure holds, 0 is member

Every subspace $U \subseteq (\mathbb{R}^n, +, \cdot)$ is the solution space of a homogenous system of linear equations $Ax=0$ for $x \in \mathbb{R}^n$

Def: Linear Combination: Consider a vector space V and a finite number of vectors $x_1, \dots x_k \in V$. Then every $v \in V$ of the form

$$v = \lambda_1 x_1 + \dots + \lambda_k x_k = \sum_{i=1}^{k} \lambda_i x_i \in V$$

with $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ is a linear combination of the vectors $x_1 \dots x_k$.

Def: Linear (in) dependence: Let us consider a vector space V with $k \in \mathbb{N}$ and $x_1, \dots x_k \in V$. If there is a non-trivial linear combination, such that $0 = \sum_{i=1}^{k} \lambda_i x_i$ with at least one $\lambda_i \neq 0$, the vectors $x_1 \dots x_k$ are linearly dependent. If only the trivial solution $\lambda_1 \dots \lambda_k = 0$ exists, the vectors $x_1 \dots x_k$ are linearly independent.

Linear independence is important because a set of linearly independent vectors have no redundancy. In simpler terms, if one vector cannot be described in terms of another, the set of those two vectors is linearly independent.

Properties to determine linear independency
- k vectors are either linearly dependent or linearly independent. There is no 3rd option.
- If at least one of the vectors $x_1, \dots x_k$ is 0, then they are linearly dependent. The same holds if two vectors are identical.
- The vectors $\{x_1, \dots, x_k : x_i \neq 0, i = 1, \dots, k\}, k \geq 2$, are linearly dependent iff one of them is a linear combination of the others.

$$\begin{bmatrix} 1 & 1 & -1 \\ 2 & 1 & -2 \\ -3 & 0 & 1 \\ 4 & 2 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$ linearly independent.

In a vector space V, m linear combinations of k vectors are linearly dependent if $m > k$ (if columns > rows, linearly dependent)