Cheat sheet on mathematical notation & linear algebra

Greek Alphabet

Letter	Upper case	Lower case
Alpha	A	α
Beta	B	β
Gamma	Γ	γ
Delta	Δ	δ
Epsilon	E	ϵ
Zeta	Z	ζ
Eta	H	η
Theta	Θ	θ
Iota	I	ι
Kappa	K	κ
Lambda	Λ	λ
Mu	M	μ
Nu	N	ν
Xi	Ξ	ξ
O	O	o
Pi	Π	π
Rho	P	ρ
Sigma	Σ	σ
Tau	T	τ
Upsilon	Υ	υ
Phi	Φ	ϕ
Chi	X	χ
Psi	Ψ	ψ
Omega	Ω	ω

Common Symbols and Notation

Symbol	Meaning
∞	Infinity
$\mathbb{R}$	Real numbers
$\mathbb{R}^n$	Set of all real-valued vectors of length n.
$\mathbf{x}$	(bold lower case) $\mathbf{x}$ is a vector, or
$\mathbf{x}$	(alternatively) $\mathbf{x}$ is a vector
$\mathbf{X}$	(bold upper case) $\mathbf{X}$ is a matrix
$\mathbf{x_i}$	The ith element of vector $\mathbf{x}$
$\mathbf{X_{i,j}}$	Element at row i and column j of matrix $\mathbf{X}$
$x \in [1, 10]$	x is a number between 1 and 10
$\approx$	Approximately equal
e	Euler's number (≈ 2.718)
π	Pi (≈ 3.142)
$\forall x \in \mathbb{R}$	Condition holds for all real numbers
$\|x\|$	Absolute value of number x (e.g., $\|x\| = \|-x\|$)
$\|X\|$	Cardinality of set X (e.g., $\|\{a, b, c\}\| = 3$).
$P(X \mid Y)$	Conditional probability of X given Y.
$\Delta(X)$	Set of all probability distributions over set X.

Logical Connectives

Symbol	Meaning
$\neg A$	not A
$A \wedge B$	A AND B
$A \vee B$	A OR B
$A \veebar B$	A XOR B
$A \Rightarrow B$	A logically implies B

Operators

Addition

$$\sum_{n=1}^{4} n = 1 + 2 + 3 + 4 = 10$$

Multiplication

$$\prod_{n=1}^{4} n = 1 \times 2 \times 3 \times 4 = 24$$

Differentiation

$$\frac{\delta}{\delta x} x^3 = 3x^2$$

Argmax The value(s) of x for which $f(x)$ takes its maximum value.

$$\arg\max_{x \in \mathbb{R}} x(10 - x) = 5$$

Logarithms

$$log_2(10) \approx 3.32$$
$$log_{10}(10) = 1$$
$$ln(10) \approx 2.3$$

Vector & Matrix Notation

- vectors are written in bold lower-case $\mathbf{x}$
- matrices are written in bold upper-case $\mathbf{A}$
- use square brackets for vectors and matrices, e.g., $\mathbf{x} = [x_1, \ldots, x_n]$ or

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \qquad \mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$$

  - an $n \times m$ matrix $\mathbf{A}$ can be written as $A = [a_{ij}] \in \mathbb{R}^{n \times m}$, or shorter as: $A \in \mathbb{R}^{n \times m}$
- the **transpose** of a matrix or vector is written as $\mathbf{A}^\intercal$ and $\mathbf{x}^\intercal$
- we use **row-first**
  - first index of a 2-D matrix is the row
  - indices for vectors give the *row*, so that vectors are column vectors
    - by convention, if the context is clear, we can write $\mathbf{x} = [x_1, \ldots, x_n]$ to denote a column vector
    - it would be more precise to write $\mathbf{x} = [x_1, \ldots, x_n]^\intercal$ or

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

Matrix Product

- if $\mathbf{A} \in \mathbb{R}^{n \times m}$ and $\mathbf{B} \in \mathbb{R}^{m \times o}$, the **matrix product** $AB = C \in \mathbb{R}^{n \times o}$ is defined via:

$$c_{ik} = \sum_j a_{ij} b_{jk}$$

- example:

$$\begin{bmatrix} 1 & -2 \\ 0 & 3 \\ -1 & 4 \end{bmatrix} \times \begin{bmatrix} 4 & 1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 4+2 & 1-4 \\ 0-3 & 0+6 \\ -4-4 & -1+8 \end{bmatrix} = \begin{bmatrix} 6 & -3 \\ -3 & 6 \\ -8 & 7 \end{bmatrix}$$

- think of matrix $\mathbf{A}$ with dimensions (n, m) as a **linear mapping** $f_{\mathbf{A}} \colon \mathbb{R}^n \to \mathbb{R}^m$ from vectors of length m to vectors for length n, so that with $\mathbf{x} = [x_1, \ldots, x_m]$:

$$f_{\mathbf{A}}(\mathbf{x}) = \mathbf{A}\mathbf{x}$$

Dot Product & Vector Similarity

- the **dot product** between equal-length vectors $\mathbf{x}$ and $\mathbf{y}$ is:

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^\intercal \mathbf{y} = \sum_i x_i y_i$$

- the **dot-product similarity** of vectors $\mathbf{x}$ and $\mathbf{y}$ is the dot product:

$$\text{DotSim}(\mathbf{x}, \mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$$

- the **cosine similarity** of vectors $\mathbf{x}$ and $\mathbf{y}$ is the dot-product similarity adjusted for vector length:

$$\text{CosineSim}(\mathbf{x}, \mathbf{y}) = \frac{\mathbf{x} \cdot \mathbf{y}}{||\mathbf{x}||\ ||\mathbf{y}||}$$

where $||\mathbf{x}|| = \sqrt{\sum_i x_i^2}$ is a measure of the length of vector $\mathbf{x}$

Further Vector & Matrix Operations

- **vector concatenation** of $\mathbf{x} \in \mathbb{R}^N$ and $\mathbf{y} \in \mathbb{R}^m$ is written as $\mathbf{x} \oplus \mathbf{y} = [x_1, \ldots, x_n, y_1, \ldots, y_m]$
  - We can generalize this notation to many vectors: $\bigoplus_{i=1}^n \mathbf{x}_i = \mathbf{x}_1 \oplus \cdots \oplus \mathbf{x}_n$
- **vector norms**

$$\begin{aligned} \mathbf{v} &= \{1, 2, 3\} \\ |\mathbf{v}| &= 6 \\ |\mathbf{v}|_2 &= \sqrt{14} \end{aligned}$$

- **matrix negation**

$$\mathbf{W} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \quad -\mathbf{W} = \begin{bmatrix} -1 & -2 & -3 \\ -4 & -5 & -6 \\ -7 & -8 & -9 \end{bmatrix}$$

- **matrix addition**

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} -2 & 7 \\ \frac{1}{7} & 1 \end{bmatrix} = \begin{bmatrix} -1 & 9 \\ \frac{22}{7} & 5 \end{bmatrix}$$

- **scalar multiplication**

$$a \times \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 1a & 2a \\ 3a & 4a \end{bmatrix}$$

- **matrix transposition**

$$\mathbf{W} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}, \mathbf{W}^T = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$$

- **matrix determinant**

$$\mathbf{W} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, |\mathbf{W}| = det(\mathbf{W}) = -2$$

- **matrix inversion**

$$\mathbf{W} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \mathbf{W}^{-1} = \begin{bmatrix} -2 & 1 \\ 1.5 & -0.5 \end{bmatrix}$$