

## Table of Symbols

Symbol	Typical meaning
$a, b, c, \alpha, \beta, \gamma$	Scalars are lowercase
$\mathbf{x}, \mathbf{y}, \mathbf{z}$	Vectors are bold lowercase
$\mathbf{A}, \mathbf{B}, \mathbf{C}$	Matrices are bold uppercase
$\mathbf{x}^\top, \mathbf{A}^\top$	Transpose of a vector or matrix
$\mathbf{A}^{-1}$	Inverse of a matrix
$\langle \mathbf{x}, \mathbf{y} \rangle$	Inner product of $\mathbf{x}$ and $\mathbf{y}$
$\mathbf{x}^\top \mathbf{y}$	Dot product of $\mathbf{x}$ and $\mathbf{y}$
$B = (\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$	(Ordered) tuple
$\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3]$	Matrix of column vectors stacked horizontally
$\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$	Set of vectors (unordered)
$\mathbb{Z}, \mathbb{N}$	Integers and natural numbers, respectively
$\mathbb{R}, \mathbb{C}$	Real and complex numbers, respectively
$\mathbb{R}^n$	$n$ -dimensional vector space of real numbers
$\forall x$	Universal quantifier: for all $x$
$\exists x$	Existential quantifier: there exists $x$
$a := b$	$a$ is defined as $b$
$a =: b$	$b$ is defined as $a$
$a \propto b$	$a$ is proportional to $b$ , i.e., $a = \text{constant} \cdot b$
$g \circ f$	Function composition: “ $g$ after $f$ ”
$\Longleftrightarrow$	If and only if
$\implies$	Implies
$\mathcal{A}, \mathcal{C}$	Sets
$a \in \mathcal{A}$	$a$ is an element of set $\mathcal{A}$
$\emptyset$	Empty set
$\mathcal{A} \setminus \mathcal{B}$	$\mathcal{A}$ without $\mathcal{B}$ : the set of elements in $\mathcal{A}$ but not in $\mathcal{B}$
$D$	Number of dimensions; indexed by $d = 1, \dots, D$
$N$	Number of data points; indexed by $n = 1, \dots, N$
$\mathbf{I}_m$	Identity matrix of size $m \times m$
$\mathbf{0}_{m,n}$	Matrix of zeros of size $m \times n$
$\mathbf{1}_{m,n}$	Matrix of ones of size $m \times n$
$\mathbf{e}_i$	Standard/canonical vector (where $i$ is the component that is 1)
$\dim$	Dimensionality of vector space
$\text{rk}(\mathbf{A})$	Rank of matrix $\mathbf{A}$
$\text{Im}(\Phi)$	Image of linear mapping $\Phi$
$\ker(\Phi)$	Kernel (null space) of a linear mapping $\Phi$
$\text{span}[\mathbf{b}_1]$	Span (generating set) of $\mathbf{b}_1$
$\text{tr}(\mathbf{A})$	Trace of $\mathbf{A}$
$\det(\mathbf{A})$	Determinant of $\mathbf{A}$
$ \cdot $	Absolute value or determinant (depending on context)
$\ \cdot\ $	Norm; Euclidean, unless specified
$\lambda$	Eigenvalue or Lagrange multiplier
$E_\lambda$	Eigenspace corresponding to eigenvalue $\lambda$

Symbol	Typical meaning
$\mathbf{x} \perp \mathbf{y}$	Vectors $\mathbf{x}$ and $\mathbf{y}$ are orthogonal
$V$	Vector space
$V^\perp$	Orthogonal complement of vector space $V$
$\sum_{n=1}^N x_n$	Sum of the $x_n$ : $x_1 + \dots + x_N$
$\prod_{n=1}^N x_n$	Product of the $x_n$ : $x_1 \cdot \dots \cdot x_N$
$\boldsymbol{\theta}$	Parameter vector
$\frac{\partial f}{\partial x}$	Partial derivative of $f$ with respect to $x$
$\frac{df}{dx}$	Total derivative of $f$ with respect to $x$
$\nabla$	Gradient
$f_* = \min_x f(x)$	The smallest function value of $f$
$x_* \in \arg \min_x f(x)$	The value $x_*$ that minimizes $f$ (note: $\arg \min$ returns a set of values)
$\mathcal{L}$	Lagrangian
$\mathcal{L}$	Negative log-likelihood
$\binom{n}{k}$	Binomial coefficient, $n$ choose $k$
$\mathbb{V}_X[\mathbf{x}]$	Variance of $\mathbf{x}$ with respect to the random variable $X$
$\mathbb{E}_X[\mathbf{x}]$	Expectation of $\mathbf{x}$ with respect to the random variable $X$
$\text{Cov}_{X,Y}[\mathbf{x}, \mathbf{y}]$	Covariance between $\mathbf{x}$ and $\mathbf{y}$ .
$X \perp\!\!\!\perp Y \mid Z$	$X$ is conditionally independent of $Y$ given $Z$
$X \sim p$	Random variable $X$ is distributed according to $p$
$\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$	Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$
$\text{Ber}(\mu)$	Bernoulli distribution with parameter $\mu$
$\text{Bin}(N, \mu)$	Binomial distribution with parameters $N, \mu$
$\text{Beta}(\alpha, \beta)$	Beta distribution with parameters $\alpha, \beta$

### Table of Abbreviations and Acronyms

Acronym	Meaning
e.g.	Exempli gratia (Latin: for example)
GMM	Gaussian mixture model
i.e.	Id est (Latin: this means)
i.i.d.	Independent, identically distributed
MAP	Maximum a posteriori
MLE	Maximum likelihood estimation/estimator
ONB	Orthonormal basis
PCA	Principal component analysis
PPCA	Probabilistic principal component analysis
REF	Row-echelon form
SPD	Symmetric, positive definite
SVM	Support vector machine