Notation

Throughout this book, we adhere to the following notational conventions. Note that some of these symbols are placeholders, while others refer to specific objects. As a general rule of thumb, the indefinite article "a" indicates that the symbol is a placeholder and that similarly formatted symbols can denote other objects of the same type. For example, "x: a scalar" means that lowercased letters generally represent scalar values.

Numerical Objects

- x: a scalar
- $\mathbf{x}$: a vector
- $\mathbf{X}$: a matrix
- X: a general tensor
- $\mathbf{I}$: an identity matrix—square, with 1 on all diagonal entries and 0 on all off-diagonals
- x_i, $[\mathbf{x}]_i$: the i^th element of vector $\mathbf{x}$
- x_{ij}, $x_{i,j}$,$[\mathbf{X}]_{ij}$, $[\mathbf{X}]_{i,j}$: the element of matrix $\mathbf{X}$ at row i and column j.

Set Theory

- $\mathcal{X}$: a set
- $\mathbb{Z}$: the set of integers
- $\mathbb{Z}^+$: the set of positive integers
- $\mathbb{R}$: the set of real numbers
- $\mathbb{R}^n$: the set of n-dimensional vectors of real numbers
- $\mathbb{R}^{a\times b}$: The set of matrices of real numbers with a rows and b columns
- $|\mathcal{X}|$: cardinality (number of elements) of set $\mathcal{X}$
- $\mathcal{A}\cup\mathcal{B}$: union of sets $\mathcal{A}$ and $\mathcal{B}$
- $\mathcal{A}\cap\mathcal{B}$: intersection of sets $\mathcal{A}$ and $\mathcal{B}$
- $\mathcal{A}\setminus\mathcal{B}$: set subtraction of $\mathcal{B}$ from $\mathcal{A}$ (contains only those elements of $\mathcal{A}$ that do not belong to $\mathcal{B}$)

Functions and Operators

- $f(\cdot)$: a function
- $\log(\cdot)$: the natural logarithm (base e)
- $\log_2(\cdot)$: logarithm with base 2
- $\exp(\cdot)$: the exponential function
- $1(\cdot)$: the indicator function, evaluates to 1 if the boolean argument is true and 0 otherwise
- $1_X(z)$: the set-membership indicator function, evaluates to 1 if the element z belongs to the set X and 0 otherwise
- $(\cdot)^\top$: transpose of a vector or a matrix
- X^{-1}: inverse of matrix X
- $\odot$: Hadamard (elementwise) product
- $[\cdot, \cdot]$: concatenation
- $\|\cdot\|_p$: L_p norm
- $\|\cdot\|$: L_2 norm
- $\langle x, y \rangle$: dot product of vectors x and y
- $\sum$: summation over a collection of elements
- $\prod$: product over a collection of elements
- $\stackrel{\mathrm{def}}{=}$: an equality asserted as a definition of the symbol on the left-hand side

Calculus

- $\frac{dy}{dx}$: derivative of y with respect to x
- $\frac{\partial y}{\partial x}$: partial derivative of y with respect to x
- $\nabla_x y$: gradient of y with respect to x
- $\int_a^b f(x)\,dx$: definite integral of f from a to b with respect to x
- $\int f(x)\,dx$: indefinite integral of f with respect to x

Probability and Information Theory

- X: a random variable
- P: a probability distribution
- $X \sim P$: the random variable X has distribution P
- $P(X = x)$: the probability assigned to the event where random variable X takes value x

- $P(X \mid Y)$ P(X|Y): the conditional probability distribution of X X given Y Y
- $p(\cdot)$ p(·): a probability density function (PDF) associated with distribution P
- $E[X]$ E[X]: expectation of a random variable X X
- $X \perp Y$ X⊥Y: random variables X X and Y Y are independent
- $X \perp Y \mid Z$ X⊥Y|Z: random variables X X and Y Y are conditionally independent given Z Z
- σ_X σX: standard deviation of random variable X X
- $\mathrm{Var}(X)$ Var(X): variance of random variable X X, equal to σ^2_X σX2
- $\mathrm{Cov}(X, Y)$ Cov(X,Y): covariance of random variables X X and Y Y
- $\rho(X, Y)$ ρ(X,Y): the Pearson correlation coefficient between X X and Y Y, equals $\frac{\mathrm{Cov}(X, Y)}{\sigma_X \sigma_Y}$ Cov(X,Y)σXσY
- $H(X)$ H(X): entropy of random variable X X
- $D_{\mathrm{KL}}(P\|Q)$ DKL(P‖Q): the KL-divergence (or relative entropy) from distribution Q Q to distribution P