Summary of Notation

Symbol	**Meaning**
$\exists$	there exists
$\forall$	for all
$A \cup B$	the union of sets A and B
$A \cap B$	the intersection of sets A and B
$\bigcup_i A_i$	the union of the sets A_i
$A \subseteq B$	A is a subset of B
$a \in A$	element a is a member of set A
$\neg$	logical not
$\vee$	logical or
$\wedge$	logical and
$\sum_{i=1}^{n} x^{(i)}$	$x^{(1)} + x^{(2)} + ... + x^{(n)}$
$\prod_{i=1}^{n} x^{(i)}$	$x^{(1)} \times x^{(2)} \times ... \times x^{(n)}$
$\mathbb{N}$	the set of natural numbers
$\mathbb{R}$	the set of real numbers
$\mathbb{R}^n$	the set of n-dimensional vectors over $\mathbb{R}$
$\mathbb{R}^{n \times m}$	the set of $(n \times m)$-dimensional matrices over $\mathbb{R}$

Symbol	Meaning
$\left\{x^{(i)} \middle\vert \texttt{Condition}\right\}_{i=1}^{n}$	a set of objects $x^{(1)}, x^{(2)}, ..., x^{(n)}$, defined such that `Condition` is fulfilled
$f : \mathbb{X} \to \mathbb{V}$	a function f which maps from the space $\mathbb{X}$ to the space $\mathbb{V}$
$f : x \mapsto f(x)$	a function f which maps an input x to an output $f(x)$
$\mathbb{I}[\texttt{Condition}]$	indicator function (equals 1 if `Condition` is true, and 0 otherwise)
δ_{ij}	Kronecker delta (equals 1 if $i = j$, and 0 otherwise)
$\mathbf{x} = \underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$	column vector comprising elements x_1, x_2,...,x_n
$\mathbf{x}^T = [x_1, x_2, \ldots, x_n]$	row vector comprising elements x_1, x_2,...,x_n
x_i or $[\mathbf{x}]_i$	the ith element of a vector $\mathbf{x}$
$\langle \mathbf{x}, \mathbf{v} \rangle$ or $\mathbf{x} \cdot \mathbf{v}$ or $\mathbf{x}^T \mathbf{v}$	inner or scalar or dot product: $\sum_{i=1}^{n} x_i v_i$
$\Vert\mathbf{x}\Vert_2$ or $\Vert\mathbf{x}\Vert$	ℓ_2-norm or two-norm of $\mathbf{x}$: $\sqrt{\sum_{i=1}^{n} x_i^2}$
$\Vert\mathbf{x}\Vert_1$ or $\vert\mathbf{x}\vert$	ℓ_1-norm or one-norm of $\mathbf{x}$: $\sum_{i=1}^{n} \vert x_i \vert$
$\mathbf{A} = \underline{A} = \begin{bmatrix} A_{11} & A_{12} & \ldots & A_{1m} \\ A_{21} & A_{22} & \ldots & A_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \ldots & A_{nm} \end{bmatrix}$	matrix comprising elements $\{A_{ij}\}_{i=1,j=1}^{n,m}$
A_{ij} or $[\mathbf{A}]_{ij}$	the i,jth element of a matrix $\mathbf{A}$
$\mathbf{I}_n$	the n by n identity matrix

Symbol	**Meaning**
$\mathbf{A}^{-1}$	the inverse matrix of matrix $\mathbf{A}$
$\mathbf{A}^T$	the transpose of matrix $\mathbf{A}$
$\det(\mathbf{A})$ or $\lvert\mathbf{A}\rvert$	the determinant of matrix $\mathbf{A}$
$\mathrm{tr}(\mathbf{A})$	the trace of matrix $\mathbf{A}$
$\mathrm{diag}(\mathbf{x})$	the diagonal matrix with diagonal elements corresponding to the elements of the vector $\mathbf{x}$
$\mathbf{A} \succ 0$	the matrix $\mathbf{A}$ is positive definite
$\mathbf{A} \succeq 0$	the matrix $\mathbf{A}$ is positive semi-definite
$\mathbf{A} \prec 0$	the matrix $\mathbf{A}$ is negative definite
$\mathbf{A} \preceq 0$	the matrix $\mathbf{A}$ is negative semi-definite
$\frac{df(x)}{dx}$ or $f'(x)$	the derivative of a function $f : \mathbb{R} \to \mathbb{R}$ at x
$\frac{d^2f(x)}{dx^2}$ or $f''(x)$	the second derivative of a function $f : \mathbb{R} \to \mathbb{R}$ at x
$\frac{\partial f(\mathbf{x})}{\partial x_i}$	the partial derivative of a function $f : \mathbb{R}^n \to \mathbb{R}$ at $\mathbf{x}$ with respect to x_i
$\nabla_{\mathbf{x}} f(\mathbf{x}) = \left[\frac{\partial f(\mathbf{x})}{\partial x_1}, \frac{\partial f(\mathbf{x})}{\partial x_2}, \ldots, \frac{\partial f(\mathbf{x})}{\partial x_n}\right]^T$	the gradient of a function $f : \mathbb{R}^n \to \mathbb{R}$ at $\mathbf{x}$
$\nabla^2_{\mathbf{x}} f(\mathbf{x})$ or $\mathcal{H}(\mathbf{x})$	the Hessian of a function $f : \mathbb{R}^n \to \mathbb{R}$ at $\mathbf{x}$
$\int f(x)dx$	the integral of a function $f : \mathbb{R} \to \mathbb{R}$
$\min_x f(x)$	the minimal value of $f(x)$
$\max_x f(x)$	the maximal value of $f(x)$

Symbol	Meaning
$\operatorname{argmin}_x f(x)$	the set $\{x \vert f(x) = \min_z f(z)\}$
$\operatorname{argmax}_x f(x)$	the set $\{x \vert f(x) = \max_z f(z)\}$
$x \sim \mathcal{D}$	x is the outcome of some random variable (say, $\mathcal{X}$) drawn from a probability distribution $\mathcal{D}$
$\mathcal{S} \sim \mathcal{D}^n$	$\mathcal{S} = \left\{x^{(i)}\right\}_{i=1}^n$ is a data set, the elements of which are the outcome of some random variable (say, $\mathcal{X}$) drawn i.i.d. from a probability distribution $\mathcal{D}$
$\mathbb{P}(\mathcal{X} = x)$ or $\mathbb{P}_{\mathcal{D}}(\mathcal{X} = x)$ or $\mathbb{P}_{x\sim\mathcal{D}}(x)$	the probability that an outcome associated with a random variable, $\mathcal{X}$, drawn from a probability distribution, $\mathcal{D}$, assumes a value x
$\mathbb{E}[f(\mathcal{X})]$ or $\mathbb{E}_{\mathcal{D}}[f(\mathcal{X})]$	the expected value or mean of a random variable, $f(\mathcal{X})$, generated from the action of a function, f, on some random variable, $\mathcal{X}$, with outcomes $x \sim \mathcal{D}$
$\mathbb{E}_{\mathcal{S}}[f(\mathcal{X})] = \frac{1}{n}\sum_{i=1}^n f(x^{(i)})$	the sample mean of the output of some function, f, evaluated on a data set, $\mathcal{S} = \left\{x^{(i)}\right\}_{i=1}^n$, the elements of which are the outcomes associated with a random variable, $\mathcal{X}$