

Advanced Machine Learning

Cristian Perez Jensen

October 26, 2024

Note that these are not the official lecture notes of the course, but only notes written by a student of the course. As such, there might be mistakes. The source code can be found at github.com/cristianpjensen/eth-cs-notes. If you find a mistake, please create an issue or open a pull request.

Contents

List of symbols

\doteq	Equality by definition
\approx	Approximate equality
\propto	Proportional to
\mathbb{N}	Set of natural numbers
\mathbb{R}	Set of real numbers
$i : j$	Set of natural numbers between i and j . <i>I.e.</i> , $\{i, i+1, \dots, j\}$
$f : A \rightarrow B$	Function f that maps elements of set A to elements of set B
$\mathbb{1}\{\text{predicate}\}$	Indicator function (1 if predicate is true, otherwise 0)
$\boldsymbol{v} \in \mathbb{R}^n$	n -dimensional vector
$\boldsymbol{M} \in \mathbb{R}^{m \times n}$	$m \times n$ matrix
\boldsymbol{M}^\top	Transpose of matrix \boldsymbol{M}
\boldsymbol{M}^{-1}	Inverse of matrix \boldsymbol{M}
$\det(\boldsymbol{M})$	Determinant of \boldsymbol{M}
$\frac{\mathrm{d}}{\mathrm{d}x}f(x)$	Ordinary derivative of $f(x)$ w.r.t. x at point $x \in \mathbb{R}$
$\frac{\partial}{\partial \boldsymbol{x}}f(\boldsymbol{x})$	Partial derivative of $f(\boldsymbol{x})$ w.r.t. \boldsymbol{x} at point $\boldsymbol{x} \in \mathbb{R}^n$
$\nabla_{\boldsymbol{x}}f(\boldsymbol{x}) \in \mathbb{R}^n$	Gradient of $f : \mathbb{R}^n \rightarrow \mathbb{R}$ at point $\boldsymbol{x} \in \mathbb{R}^n$
$\nabla_{\boldsymbol{x}}^2f(\boldsymbol{x}) \in \mathbb{R}^{n \times n}$	Hessian of $f : \mathbb{R}^n \rightarrow \mathbb{R}$ at point $\boldsymbol{x} \in \mathbb{R}^n$

References