0 - Course Notations

November 10, 2020

0.1 Course Notations

For simplicity and avoiding confusion, we shall stick to the following notations throughout our course. Note that these notations may vary across disciplines and even person to person. I will try to use most common notations when possible.

Symbol / Notations	Typical meaning
$a, b, c, \alpha, \beta, \gamma$	Scalars are lowercase
$\mathbf{x}, \mathbf{y}, \mathbf{z}$	Vectors are bold lowercase
$\mathbf{X}, \mathbf{Y}, \mathbf{Z}$	Matrices are bold uppercase
$\mathbf{x}^\top, \mathbf{X}^\top$	Transpose of a vector or matrix
$\mathbf{X^{-1}}$	Inverse of a matrix
$\langle \mathbf{x}, \mathbf{y} \rangle$	Inner product of **x** and **y**
$\mathbf{x}^\top \mathbf{y}$	Dot product of **x** and **y**
$\mathbb{Z}$	set of integers
$\mathbb{R}$	set of real numbers
$\mathbb{R}^n$	n-dimensional vector space of real numbers
$\mathbf{x} \in \mathbb{R}^n$	x is member of n-dimensional vector space of real numbers, i.e., x has n features
$\forall x$	for all x
$\exists x$	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., $a = \text{constant} * b$
$\iff$	if and only if
$\implies$	implies
I_m	Identity matrix of size $m \times m$
$0_{m,n}$	Matrix of zeros of size $m \times n$
$I(a = b)$	Indicator function; True will evaluate to 1, and False will evaluate to 0
$rk(\mathbf{A})$	Rank of matrix $\mathbf{A}$
$tr(\mathbf{A})$	Trace of matrix $\mathbf{A}$
$det(\mathbf{A})$	Determinant of matrix $\mathbf{A}$
$\lVert a \rVert$	Norm of a; Euclidean unless specified
λ	Eigenvalue or Lagrange multiplier or learning rate

Symbol / Notations	Typical meaning
α	Equality lagrange multiplier or learning rate
β	Inequality lagrange multiplier
θ	Model weights
w	Model weights
π	Model weights
$f(x)$	Function of x
∂	Partial derivatives
d	Derivatives
$f'(x)$	Derivatives of $f(x)$
Δ	Delta, i.e., differences
∇	Gradient
$\mathscr{L}$	Lagrangian
$\mathcal{L}$	Negative log-likelihood
$\mathbb{V}_X[x]$	Variance of x with respect to the random variable X
$\mathbb{E}_X[x]$	Expectation of x with respect to the random variable X
$\mathbb{E}_X[x]$	Expectation of x with respect to the random variable X
μ	Mean
$\bar{x}$	Mean of x
Σ	Covariance
$Cov_{X,Y}[x, y]$	Covariance between x and y
σ	Standard deviation
$p(x)$	Probability of x
$p(x\|y)$	Probability of x given y
$p(x\|y; \theta)$	Probability of x given y parametrize by θ
$X \sim p$	Random variable X is distributed according to p
$\mathcal{N}(\mu, \pm)$	Gaussian distribution with mean μ and covariance Σ
$\sum$	Summation
$\prod$	Products

Course-Specific Notations	Meaning
M	Number of samples; indexed by $m = 1, \cdots, M$
N	Number of features; indexed by $n = 1, \cdots, N$
K	Number of classes / clusters; indexed by $k = 1, \cdots, K$
$a \ \times \ b$	Matrix shape of (a, b), i.e., a rows, b columns
$\mathbf{x}$	Vector of a sample with shape of n

Course-Specific Notations	Meaning
$\mathbf{x^{(1)}}, \mathbf{x^{(i)}}$	First sample; i-th sample
$\mathbf{x_1}, \mathbf{x_i}$	First feature; i-th feature
$\mathbf{x_1^{(1)}}, \mathbf{x_i^{(1)}}$	First feature of first sample; i-th feature of first sample
$\mathbf{X}$	Matrices are all samples, with shape $M \times N$, i.e., $\mathbf{X}$ shall have m rows of samples, and n columns of features
$\mathbf{y}$	Vector of targets with shape of m

Acronym	Meaning
e.g.,	For example
i.e.,	That is
i.i.d.	Independent, identically distributed

[ ]: