TOOL

Notations for Machine Learning

Specific Notation

Name	Notation	Remarks
The set of real numbers	$\mathbb{R}$	
D-dimensional vector space	$\mathbb{R}^d$	data points
Training point	$\mathbf{x}_i$	$\mathbf{x}_i \in \mathbb{R}^d$ (i-th data point)
Data entry	$\mathbf{x}_{ij}$ or $[\mathbf{x}_i]_j$	the j-th entry of the i-th data point
Number of data points	n	
Dimension of data point	d	
Function	f	xs
Label	y_i	the label of the i-th data point
Label vector	$\mathbf{y}$	an n-dimensional vector, where the i-th entry is y_i
Coefficient vector	$\mathbf{w}$	weight vector of a linear model
Bias or Offset	b	bias term of a linear model
Data matrix	$\mathbf{X}$	an $(n \times d)$ real matrix. its i-th row is $\mathbf{x}_i^T$
First order partial deriative	$\frac{\partial f}{\partial \mathbf{x}_i}$	first order partial deriative of f with respect to $\mathbf{x}_i$
Second order partial deriative	$\frac{\partial^2 f}{\partial \mathbf{x}_i \partial \mathbf{x}_j}$	second order partial deriative of f with respect to $\mathbf{x}_i, \mathbf{x}_j$
Gradient with respect to $\mathbf{x}$	$\nabla_{\mathbf{x}} f$	$\nabla_{\mathbf{x}} f$ is a vector where the i-th entry is $\frac{\partial f}{\partial \mathbf{x}_i}$
Hessian with respect to $\mathbf{x}$	$\nabla^2_{\mathbf{x}} f$	a matrix where the (i,j)-th entry is $\frac{\partial^2 f}{\partial \mathbf{x}_i \partial \mathbf{x}_j}$