

Companion to Machine Learning

Rohan Kumar

Contents

0	Notation	3
0.1	Data	3

0 Notation

0.1 Data

$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_M \end{pmatrix}$: data point corresponding to a column vector of M features

$\bar{\mathbf{x}} = \begin{pmatrix} 1 \\ x_1 \\ x_2 \\ \dots \\ x_M \end{pmatrix}$: concatenation of 1 with the vector \mathbf{x}

$\mathbf{X} = \begin{pmatrix} x_{1,1} & \dots & x_{1,N} \\ \dots & \dots & \dots \\ x_{M,1} & \dots & x_{M,N} \end{pmatrix}$: dataset consisting of N data points and M features

$\bar{\mathbf{X}} = \begin{pmatrix} 1 & \dots & 1 \\ x_{1,1} & \dots & x_{1,N} \\ \dots & \dots & \dots \\ x_{M,1} & \dots & x_{M,N} \end{pmatrix}$: concatenation of a vector of 1's with the matrix \mathbf{X}

y : output target (regression) or label (classification)

$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_N \end{pmatrix}$: vector of outputs for a dataset of N points

\mathbf{x}_* : test input / unknown input

\mathbf{y}_* : predicted output

N : Number of data points in the dataset

M : Number of a features in a data point

$\mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \\ \dots \\ w_M \end{pmatrix}$

$\mathbf{w}^T = (w_1, w_2, \dots, w_M)$ or $(w_0, w_1, w_2, \dots, w_M)$ w_0 multiplies the first entry of $\bar{\mathbf{x}}$ (bias)

Note: bold symbols represents a vector