EVERYTHING YOU NEED TO KNOW ABOUT LINEAR REGRESSION BUT WERE AFRAID TO ASK

FRANK WOOD

Linear regression is Gaussian conditional density estimation the mean of the output distribution conditionally dependent on the input value. The basic mathematical setup looks like

$$\mathbf{y} = \mathbf{\Phi}\boldsymbol{\beta} + \boldsymbol{\epsilon} \tag{1}$$

where $\mathbf{y}' = [y_1, \ldots, y_N]$, $\boldsymbol{\beta}' = [\beta_1, \ldots, \beta_P]$, $\boldsymbol{\epsilon}' = [\epsilon_1, \ldots, \epsilon_N]$,

$$\mathbf{\Phi} = \begin{bmatrix} \phi_1(x_1) & \cdots & \phi_P(x_1) \\ \vdots & \vdots & \vdots \\ \phi_1(x_N) & \cdots & \phi_P(x_N) \end{bmatrix}$$

and $\phi_p(x_n)$ is the p^{th} "feature" of input x_n. The matrix form of simple linear regression is recovered by setting $\phi_1(\cdot) = \cdot$ and $\phi_2(\cdot) = 1$.