Project II

Linear Regression

Given N values $\{x_1, x_2, \dots, x_N\}$, it is desired to find a linear function $f(x_i) = w.x_i + b$ to map these source values to associated target values $\{y_1, y_2, \dots, y_N\}$. Considering the difference of y_i and $f(x_i)$ as the error of i^{th} instance (i.e., e_i) and $\vec{e}^T = [e_1, e_2, \dots, e_N]$ as the error vector, the mapping function $f(.)$ is modeled in order to minimize $\|\vec{e}\|_d$ whereas, d is the degree of the norm. For a simple regression dataset (can be find in UCI repository), try to solve this problem for each feature separately (i.e., the selected feature forms x_i and the target value forms y_i). For each $d = 0, 1, 2, \infty$, analyze and report the norm of error vector respect to each feature. Dear students should consider different datasets.

Regards,

M. Taheri