AM160 Homework 1

1 Problem 1

1.1 Subproblem A

Consider n observations, from a data distribution, D, given as: (x_1, y_1), (x_2, y_2), $(x_2, y_2) \cdots$ (x_n, y_n). We formulate a linear regression problem with feature matrix, A, containing p features using the basis function map $\Phi : R^1 \to R^p$. So A is given by:

$$\mathrm{A}_{(n\times p)} = \begin{bmatrix} \Phi_1(x_1) & \Phi_2(x_1) & \Phi_3(x_1) & \cdots & \Phi_p(x_1) \\ \Phi_1(x_2) & \Phi_2(x_2) & \Phi_3(x_2) & \cdots & \Phi_p(x_2) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \Phi_1(x_n) & \Phi_2(x_n) & \Phi_3(x_n) & \cdots & \Phi_p(x_n) \end{bmatrix}$$

The unknown parameters, we want to estimate is given by $\Theta_{p\times 1}$. Prove that, if $p \gg n$, then the solution to: $A\Theta = y$, which is given by $\Theta^* = A^\dagger y$ has the minimum 2-norm amongst all the other infinite solutions.