

# Machine Learning @MIT Lecture Note 6 notes Active learning - Non linear prediction - Kernel 1

Hien Nguyen (hienminhnguyen711@gmail.com)

September 30, 2020

## Question 1.

---

Lecture note 6 Active Learning - Non linear prediction - kernel

ACTIVE LEARNING - Reduce MSE to actively select input point - noise variance  $(A^T A + \lambda I)^{-1} = A^T (A^T A + \lambda I)^{-1} A$   
- noise in response has a large effect on param of linear model - linear model  $\rightarrow$  input should be far from each other as possible - variance in prediction

NON LINEAR PREDICTION, KERNEL - map input to a higher dimensional feature space  $\rightarrow$  non linear function of inputs  $x_1, x_2, \dots, x_n$  - dimension of feature vector

LINEAR REGRESSION AND KERNEL - regularization - regularized least square objective to be minimized  
 $J(\theta) = \frac{1}{2} \sum_{t=1}^n (y_t - \theta^T x_t)^2 + \frac{\lambda}{2} \theta^T \theta$