# HW 2 – Navya Bhat

2.4.1

(a) Flexible model performs better when n is large and no. of predictors is small.

→ For flexible method like knn, we can find sufficient no. of neighbours to accurately estimate f when n is large & p small.

(b) When p is large and no. of observation is small, flexible model performs worse as it suffers from the curse of dimensionality.

(c) When the relationship b/w predictors & response is highly non linear, flexible method performs better as it can adapt to complex patterns of f, reducing bias & MSE.

d) When variance of error terms is high, inflexible method performs better as flexible models may suffer from overfitting given the higher noise.

2.4.7 $X_1 = X_2 = X_3 = 0$.

Eucliden distance:

$$\sqrt{(X_1 - X_1^*)^2 + (X_2 - X_2^*)^2 + (X_3 - X_3^*)^2}$$

(a)

| Obs. | $X_1$ | $X_2$ | $X_3$ | Y | Eucliden Distance |
|---|---|---|---|---|---|
| 1 | 0 | 3 | 0 | R | 3 |
| 2 | 2 | 0 | 0 | R | 2 |
| 3 | 0 | 1 | 3 | R | $\sqrt{10} = 3.16$ |
| 4 | 0 | 1 | 2 | G | $\sqrt{5} = 2.23$ |
| 5 | -1 | 0 | 1 | G | $\sqrt{2} = 1.414$ |
| 6 | 1 | 1 | 1 | R. | $\sqrt{3} = 1.73$ |

(b) $K=1$; only 1 neighbor $\rightarrow$ observation 5 since it has smallest distance.

Prediction: GREEN as $P(Y|X=5) = 1$ is maximised for $Y =$ green

(c) $K = 3 \rightarrow$ Neighbors: Obs 5, 6, 2

$$P(\text{Red} \mid X_1 = X_2 = X_3 = 0) = \frac{2}{3} = 67\%$$

P maximised ~~&~~ when $Y =$ Red.

Prediction $\rightarrow$ RED

(d) If Bayes decision boundary is highly non-linear then we need a more flexible model to minimize MSE.

$K \rightarrow$ Small