1

Let H be nxn simple regression

By what we learnt in simple regression $(X'X)^{-1} = \begin{pmatrix} \sum \frac{x_i^2}{n} & -\bar{x} \\ -\bar{x} & 1 \end{pmatrix} \frac{1}{SXX}$

Then by definition of h_{ij}, $h_{ij} = x_i'(X'X)^{-1}x_j$

$$= (1 \ \ x_i) \frac{1}{SXX} \begin{pmatrix} \sum \frac{x_i^2}{n} & -\bar{x} \\ -\bar{x} & 1 \end{pmatrix} \begin{pmatrix} 1 \\ x_j \end{pmatrix}$$

$$= \left(\frac{\sum x_i^2}{nSXX} - \frac{\bar{x} x_i}{SXX} \quad \frac{-\bar{x}}{SXX} + \frac{x_i}{SXX} \right) \begin{pmatrix} 1 \\ x_j \end{pmatrix}$$

$$= \frac{\sum x_i^2}{nSXX} - \frac{\bar{x} x_i}{SXX} + \frac{-\bar{x} x_j}{SXX} + \frac{x_i x_j}{SXX}$$

Let $j = i$ then, $h_{ii} = \frac{\sum x_i^2}{nSXY} - \frac{\bar{x} x_i + \bar{x} x_i - x_i^2}{SXX}$

$$= \frac{\sum x_i^2}{nSXX} - \frac{\bar{x}^2}{SXX} - \frac{\bar{x} x_i + \bar{x} x_i - x_i^2 - \bar{x}^2}{SXX}$$

$$= \frac{\sum x_i^2 - n\bar{x}^2}{nSXX} + \frac{(x_i - \bar{x})^2}{SXX}$$

$$= \frac{1}{n} + \frac{(x_i - \bar{x})^2}{SXX}$$

2 High $h_{ii} \longrightarrow$ high $(x_i - \bar{x})^2 \longrightarrow$ some points in the scatterplot are very far from other points (outliers).

3 Suppose for X, $x_1 = 1$. $x_2 = x_3 = \cdots = x_n = 0$.

so $\bar{x} = \frac{1}{n}$. $SXX = \sum (x_i - \bar{x})^2 = (\frac{1}{n})^2 (n-1) + (1 - \frac{1}{n})^2$

$$= \frac{n-1}{n^2} + \frac{(n-1)^2}{n^2}$$

$$= \frac{n-1+n^2-2n+1}{n^2}$$

$$= \frac{n^2-n}{n^2}$$

$$= \frac{n-1}{n}$$

Hence $h = \frac{1}{n} + \frac{(1-\frac{1}{n})^2}{1-\frac{1}{n}} = 1$

Done