ESL exercise 7.6: Show that for an additive-error model, the effective degrees of freedom for the k-nearest-neighbors regression fit is N/k

* Additive error model: $Y = f(X) + \varepsilon$, $E(\varepsilon) = 0$, $Var(\varepsilon) = \sigma^2_\varepsilon$

* We know that for a linear smoother $\hat{Y} = SY$

then $\sum_{i=1}^{N} Cov(\hat{y}_i, y_i) = \text{trace}(S) \cdot \sigma^2_\varepsilon$

* Can we write k-NN as a linear smoother?

We know that $\hat{f}(X_0) = \frac{1}{k} \sum_{i:\, x_i \in N_k(x_0)} Y_i$ — so it is a linear function of Y's and thus a linear smoother.

Here S has elements $S_{ij} = \frac{1}{k}$ if $x_j \in N_k(x_i)$

$N \times N$

we now look at all N obs

to predict

0 if not in $N_k(x_i)$

$1/k$ if in

$$\begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_N \end{bmatrix} = \begin{bmatrix} \frac{1}{k} & & & \\ & \frac{1}{k} & & \\ & & \ddots & \\ & & & \frac{1}{k} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}$$

$$df(S) = \text{trace}(S) = \sum_{i=1}^{N} \frac{1}{k} = \underline{\underline{\frac{N}{k}}}$$