1. [20 pts] (2.2 in the textbook) The reflection boundary condition is a typical boundary condition for smoothing by a moving average or a binomial filter. Other boundary conditions, however, are also available. The periodic boundary condition is also a typical one. This boundary condition is following:

$$Y_0 = Y_n, Y_{-1} = Y_{n-1}, \cdots, Y_{-m+1} = Y_{n-m+1}$$
$$Y_{n+1} = Y_1, Y_{n+2} = Y_2, \cdots, Y_{n+m} = Y_m$$

(a) Construct a R function for smoothing by a moving average based on this boundary condition.
(b) Create a function for smoothing by a binomial filter based on this boundary condition.
(c) Using the functions in (a) and (b), smooth the 20 data below. m can make various values.
$2.46, -0.59, 1.14, -0.94, 0.62, -0.63, -0.43, 2.30, 1.29, 0.25, 1.92, -0.17, 0.22, -2.13, -3.03,$
$-1.29, -3.24, 1.04, -0.64, 1.85$
(d) Using the functions in (a) and (b), smooth the 20 data below.
$-0.63, -0.43, 2.30, 1.29, 0.25, 1.92, -0.17, 0.22, -2.13, -3.03, -1.29, -3.24, 1.04, -0.64,$
$1.85, 2.46, -0.59, 1.14, -0.94, 0.62$
The data presented in (d) is produced by the periodic shift of the data in (c). Confirm that this periodic shift does not affect the essential behaviour of the estimates.

2. [10 pts] (2.6 in the textbook) Alter the code (F) in section 2.7 to acquire intuitive understanding of fitting a spline by the least squares.
(a) Adopt $\{1, 10.3, 16.5, 22.8, 31\}$ as the positions of knots of a linear spline. Note that 1 and 31 are end knots which is not included in `knots=` option in R
(b) Using the code from (a), observe the responses of the estimates given by the movements of data and describe the findings obtained. Construct a graph of the values of elements of a hat matrix in a manner similar to those of figure 2.17, 2.19.

3. [10 pts] (2.7 in the textbook) Consider the code (G) in section 2.7 to obtain the basic concepts of local linear regression.
(a) Create an R code for producing a graph to show how the estimates given by local linear regression are derived; the graph should be similar to that in figure 2.26.
(b) Using the following data, apply the code in (a) with various bandwidth.
$9.6, 12.8, 14.6, 15.6, 15.5, 15.1, 15.6, 13.8, 13.9, 16.1,$
$17.3, 18, 19, 9, 20, 19.9, 18.2, 15.8, 11.2, 9.6, 15.8$
$16.7, 17.5, 13.7, 15.7, 20.6, 21.2, 16.7, 16, 20.7, 17.6$

4. [15 pts] (2.9 in the textbook) Utilize the code (H) in section 2.7 to obtain the basic concepts of the smoothing spline.
(a) Smooth the data in Q. 3 in this homework using the various smoothing parameters.
(b) Alter some values of the data and observe the effects of such alteration. On the basis of these results, calculate and illustrate the values of elements of a hat matrix of the smoothing spline. Confirm that the resultant hat matrix is symmetric.
(c) Compute the values of elements of a hat matrix directly using $(\boldsymbol{I} + \lambda \boldsymbol{S})^{-1}$ and compare these values with those obtained in (b).

5. [25 pts] (2.15 in the textbook) Consider the diagonal elements of a hat matrix of a simple regression.

(a) Assume that the predictor values are $\{X_1, X_2, \cdots, X_n\}$ (they are not necessarily equispaced) and that the values of a target variable are $\{Y_1, \cdots, Y_n\}$. The regression equation ($y = a_0 + a_1 x$) are derived by minimizing $E_{simple} = \sum_{i=1}^{n}(a_0 + a_1 X_i - Y_i)^2$.

Prove that the resultant regression coefficients are written as

$$\hat{a}_0 = \bar{Y} - \hat{a}_1 \bar{X}, \quad \hat{a}_1 = S_{xy}/S_{xx},$$

where $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$, $S_{xx} = \sum_{i=1}^{n}(X_i - \bar{X})^2$ and $S_{xy} = \sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})$.

(b) Assume that among $\{Y_1, \cdots, Y_n\}$, only Y_k is replaced by $Y_k + \Delta Y_k$. $\{X_1, \cdots, X_n\}$ remains the same. The estimate corresponding to $Y_k + \Delta Y_k$ is defined as $\hat{Y}_k + \Delta\hat{Y}_k$.

Derive the two equations below.

$$\hat{Y}_k = \bar{Y} + \frac{(X_k - \bar{X})S_{xy}}{S_{xx}}$$

$$\hat{Y}_k + \Delta\hat{Y}_k = \bar{Y} + \frac{\Delta Y_k}{n} + \frac{(X_k - \bar{X})S_{xy} + (X_k - \bar{X})^2 \Delta Y_k}{S_{xx}}.$$

(c) Using the result of (b), obtain

$$\Delta\hat{Y}_k = \frac{\Delta Y_k}{n} + \frac{(X_k - \bar{X})^2 \Delta Y_k}{S_{xx}}.$$

(d) On the basis of (c), explain/derive

$$[\boldsymbol{H}]_{kk} = \frac{1}{n} + \frac{(X_k - \bar{X})^2}{S_{xx}}.$$

(e) Using the result in (d), confirm that $\sum_{i=1}^{n}[\boldsymbol{H}]_{ii} = 2$.