MATH-517: Assignment 2

Valerio Viscovo

04/10/2025

Theoretical exercise

1) Proof of the Weighted Least Squares Solution

Let's prove that the solution to the minimization problem

$$\begin{aligned}\hat{\beta}(x) = (\hat{\beta}_0(x), \hat{\beta}_1(x)) &= \underset{\beta\in\mathbb{R}^2}{\operatorname{argmin}} \sum_{i=1}^{n} \{Y_i - \beta_0 - \beta_1(X_i - x)\}^2 K\left(\frac{X_i - x}{h}\right) \\ &= \underset{\beta\in\mathbb{R}^2}{\operatorname{argmin}} \left((\mathbf{Y} - \mathbf{X}\beta)^T \mathbf{W}(\mathbf{Y} - \mathbf{X}\beta)\right)\end{aligned}$$

is given by the weighted least squares estimator:

$$\hat{\beta} = (\mathbf{X}^t\mathbf{W}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{W}\mathbf{Y},$$

where the matrices and vectors are defined as:

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix} \in \mathbb{R}^{n\times 1}, \quad \mathbf{X} = \begin{pmatrix} 1 & X_1 - x \\ \vdots & \vdots \\ 1 & X_n - x \end{pmatrix} \in \mathbb{R}^{n\times 2},$$

$$\mathbf{W} = \operatorname{diag}\left(K\left(\frac{X_1 - x}{h}\right), \dots, K\left(\frac{X_n - x}{h}\right)\right) \in \mathbb{R}^{n\times n}.$$

The objective function in matrix form is:

$$L(\beta) = (\mathbf{Y} - \mathbf{X}\beta)^t\mathbf{W}(\mathbf{Y} - \mathbf{X}\beta)$$

Expanding the quadratic form yields:

$$L(\beta) = \mathbf{Y}^t\mathbf{W}\mathbf{Y} - \beta^t\mathbf{X}^t\mathbf{W}\mathbf{Y} - \mathbf{Y}^t\mathbf{W}\mathbf{X}\beta + \beta^t\mathbf{X}^t\mathbf{W}\mathbf{X}\beta$$

Since $\mathbf{W}^t = \mathbf{W}$ (as it is a diagonal matrix) and the middle two terms are scalars which are transposes of each other ($\mathbf{Y}^t\mathbf{W}\mathbf{X}\beta = (\mathbf{Y}^t\mathbf{W}\mathbf{X}\beta)^t = \beta^t\mathbf{X}^t\mathbf{W}\mathbf{Y}$), the function simplifies to:

$$L(\beta) = \mathbf{Y}^t\mathbf{W}\mathbf{Y} - 2\mathbf{X}^t\mathbf{W}\mathbf{Y}\beta + \beta^t\mathbf{X}^t\mathbf{W}\mathbf{X}\beta$$

Taking the derivative with respect to β and setting it to zero (the first-order condition):

$$\frac{\partial L}{\partial \beta} = -2\mathbf{X}^t\mathbf{W}\mathbf{Y} + 2\mathbf{X}^t\mathbf{W}\mathbf{X}\beta = \mathbf{0}$$

Solving for $\hat{\beta}$:

$$\hat{\beta} = (\mathbf{X}^t\mathbf{W}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{W}\mathbf{Y}$$

The second derivative is $\frac{\partial^2 L}{\partial \beta^2} = 2\mathbf{X}^t\mathbf{W}\mathbf{X}$, which is positive definite (since $\mathbf{W}$ is positive definite due to the kernel function), confirming that $\hat{\beta}$ is a minimum.

The estimator $\hat{m}(x)$ is the first component of $\hat{\beta}$, so $\hat{m}(x) = \hat{\beta}_0(x)$. Denoting the first row of $(\mathbf{X}^t\mathbf{W}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{W}$ as $(w_{n,1}, \dots, w_{n,n})$, we have:

$$\hat{\beta}_0 = \sum_{i=1}^{n} w_{n,i} Y_i$$

Thus, the Local Linear Regression is a Linear Smoother.

2) Derivation of the Explicit Expression for the Weights

As demonstrated:

$$\begin{pmatrix} \hat{\beta}_0(x) \\ \hat{\beta}_1(x) \end{pmatrix} = (\mathbf{X}^t\mathbf{W}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{W}\mathbf{Y}$$

We compute $\mathbf{X}^t\mathbf{W}\mathbf{X}$:

$$\mathbf{X}^t\mathbf{W}\mathbf{X} = \begin{pmatrix} 1 & \dots & 1 \\ X_1 - x & \dots & X_n - x \end{pmatrix} \mathbf{W} \begin{pmatrix} 1 & X_1 - x \\ \vdots & \vdots \\ 1 & X_n - x \end{pmatrix}$$

Multiplying the first two matrices ($\mathbf{X}^t\mathbf{W}$):

$$= \left(\begin{pmatrix} K\left(\frac{X_1 - x}{h}\right) & \dots & K\left(\frac{X_n - x}{h}\right) \\ (X_1 - x)K\left(\frac{X_1 - x}{h}\right) & \dots & (X_n - x)K\left(\frac{X_n - x}{h}\right) \end{pmatrix} \begin{pmatrix} 1 & X_1 - x \\ \vdots & \vdots \\ 1 & X_n - x \end{pmatrix} \right)$$

Carrying out the final matrix multiplication:

$$= \begin{pmatrix} \sum_{i=1}^n K\left(\frac{X_i - x}{h}\right) & \sum_{i=1}^n (X_i - x)K\left(\frac{X_i - x}{h}\right) \\ \sum_{i=1}^n (X_i - x)K\left(\frac{X_i - x}{h}\right) & \sum_{i=1}^n (X_i - x)^2 K\left(\frac{X_i - x}{h}\right) \end{pmatrix}$$

Using the notation $S_{n,k} = \frac{1}{nh}\sum_{i=1}^{n}(X_i - x)^k K\left(\frac{X_i - x}{h}\right)$ we have:

$$\mathbf{X}^t\mathbf{W}\mathbf{X} = nh\begin{pmatrix} S_{n,0} & S_{n,1} \\ S_{n,1} & S_{n,2} \end{pmatrix}$$

The inverse of $\mathbf{X}^t\mathbf{W}\mathbf{X}$ is:

$$(\mathbf{X}^t\mathbf{W}\mathbf{X})^{-1} = \frac{1}{nh}\begin{pmatrix} S_{n,0} & S_{n,1} \\ S_{n,1} & S_{n,2} \end{pmatrix}^{-1} = \frac{1}{nh(S_{n,0}S_{n,2} - S_{n,1}^2)}\begin{pmatrix} S_{n,2} & -S_{n,1} \\ -S_{n,1} & S_{n,0} \end{pmatrix}$$

Therfore,

$$\begin{aligned}(\mathbf{X}^t\mathbf{W}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{W} = {} & \frac{1}{nh(S_{n,0}S_{n,2} - S_{n,1}^2)}\begin{pmatrix} S_{n,2} & -S_{n,1} \\ -S_{n,1} & S_{n,0} \end{pmatrix} \\ & \cdot \begin{pmatrix} K\left(\frac{X_1-x}{h}\right) & \dots & K\left(\frac{X_n-x}{h}\right) \\ (X_1 - x)K\left(\frac{X_1-x}{h}\right) & \dots & (X_n - x)K\left(\frac{X_n-x}{h}\right) \end{pmatrix}\end{aligned}$$

The estimator $\hat{\beta}_0(x)$ is found by taking the dot product of the first row of the matrix $(\mathbf{X}^t\mathbf{W}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{W}$ with the vector $\mathbf{Y}$. Hence, the weights $w_{n,i}(x)$ are the entries of the first row of $(\mathbf{X}^t\mathbf{W}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{W}$:

$$\begin{aligned} w_{n,i}(x) &= \frac{1}{nh(S_{n,0}S_{n,2} - S_{n,1}^2)}\left[S_{n,2}K\left(\frac{X_i - x}{h}\right) - S_{n,1}(X_i - x)K\left(\frac{X_i - x}{h}\right)\right] \\ &= \frac{S_{n,2} - S_{n,1}(X_i - x)}{nh(S_{n,0}S_{n,2} - S_{n,1}^2)}K\left(\frac{X_i - x}{h}\right)\end{aligned}$$

3) Proof that the Weights Satisfy $\sum_{i=1}^{n} w_{n,i}(x) = 1$

The sum of the weights is:

$$\sum_{i=1}^{n} w_{n,i}(x) = \sum_{i=1}^{n}\left(\frac{S_{n,2} - S_{n,1}(X_i - x)}{nh(S_{n,0}S_{n,2} - S_{n,1}^2)}K\left(\frac{X_i - x}{h}\right)\right)$$

Factoring out the common terms:

$$= \frac{1}{nh(S_{n,0}S_{n,2} - S_{n,1}^2)}\sum_{i=1}^{n}\left[S_{n,2}K\left(\frac{X_i - x}{h}\right) - S_{n,1}(X_i - x)K\left(\frac{X_i - x}{h}\right)\right]$$

Separating the summation:

$$= \frac{1}{nh(S_{n,0}S_{n,2} - S_{n,1}^2)}\left[S_{n,2}\sum_{i=1}^{n}K\left(\frac{X_i - x}{h}\right) - S_{n,1}\sum_{i=1}^{n}(X_i - x)K\left(\frac{X_i - x}{h}\right)\right]$$

Using the definitions of $S_{n,0}$ and $S_{n,1}$:

$$\sum_{i=1}^{n} K\left(\frac{X_i - x}{h}\right) = nhS_{n,0}$$

$$\sum_{i=1}^{n} (X_i - x)K\left(\frac{X_i - x}{h}\right) = nhS_{n,1}$$

Substituting these back into the expression:

$$\sum_{i=1}^{n} w_{n,i}(x) = \frac{1}{nh(S_{n,0}S_{n,2} - S_{n,1}^2)} \left[S_{n,2}(nhS_{n,0}) - S_{n,1}(nhS_{n,1})\right]$$

$$= \frac{nh(S_{n,0}S_{n,2} - S_{n,1}^2)}{nh(S_{n,0}S_{n,2} - S_{n,1}^2)} = 1$$

This proves that the weights satisfy the necessary property: $\sum_{i=1}^{n} w_{n,i}(x) = 1$.

Practical exercise

Description of the Simulation Study

Intervening Quantities