Kernel estimator:
$$\omega(u,v) \triangleq \frac{1}{h} k(\frac{u-v}{h})$$
with $h > 0$.

Problem 1.

The local polynomial estimator can be derived as

$$(\tilde{\beta}_0, \ldots, \tilde{\beta}_k) = \underset{\beta_0,\ldots,\beta_k}{\arg\min} \frac{1}{N} \sum_{i=1}^{n} \sum_{j=1}^{m_i} (Y_{ij} - \sum_{j=0}^{k_i} \frac{g^{(j)}(t)}{j!} (t_{ij} - t)^j)^2 k_h(t - t_{ij})$$

where

$$k_h(v) = \frac{1}{h} k(\frac{v}{h}).$$

It implies that $\tilde{g}_h(t) = \tilde{\beta}_0$.

Show that, when $k = 1$: (local linear fit)

$$\tilde{g}_h(t) = \frac{\sum_{i=1}^{n} \sum_{j=1}^{m_i} \omega(t, t_{ij}) Y_{ij}}{\sum_{i=1}^{n} \sum_{j=1}^{m_i} \omega(t, t_{ij})}$$

where

$$\omega(t, t_{ij}) = k_h(t - t_{ij})\{S_{n,2} + (t - t_{ij}) S_{n,2}\},$$

with

$$S_{n,j} = \sum_{i=1}^{n} \sum_{j=1}^{m_i} k_h(t - t_{ij})(t - t_{ij})^j.$$

Solution.

$$Q = \sum_{i=1}^{n} \sum_{j=1}^{m_i} (Y_{ij} - \sum_{j=0}^{k_i} \frac{g^{(j)}(t)}{j!} (t_{ij} - t)^j)^2 k_h(t - t_{ij})$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{m_i} (Y_{ij}^2 - \beta_0^2 - 2\beta_0 Y_{ij} + \beta_1^2 (t_{ij} - t)^2 - 2\beta_1 Y_{ij}(t_{ij} - t) + 2\beta_0 \beta_1 (t_{ij} - t)) k_h(t - t_{ij})$$

$$\frac{\partial Q}{\partial \beta_0} = \sum_{i=1}^{n} \sum_{j=1}^{m_i} (-2Y_{ij} - 2\beta_0 + 2\beta_1 (t_{ij} - t)) k_h(t - t_{ij}) \overset{\text{set}}{=} 0$$

$$\frac{\partial Q}{\partial \beta_1} = \sum_{i=1}^{n} \sum_{j=1}^{m_i} (-2Y_{ij}(t_{ij} - t) + 2\beta_1 (t_{ij} - t)^2 + 2\beta_0 (t_{ij} - t)) \overset{\text{set}}{=} 0$$

$$\begin{pmatrix} S_{n,0} & S_{n,1} \\ S_{n,1} & S_{n,2} \end{pmatrix} \begin{pmatrix} \tilde{\beta}_0 \\ \tilde{\beta}_1 \end{pmatrix} =$$

$$\begin{pmatrix} \sum\sum k_h(t - t_{ij}) & \sum\sum k_h(t - t_{ij}) \\ \sum\sum k_h(t - t_{ij})(t_{ij} - t) & \sum\sum k_h(t - t_{ij})(t_{ij} - t)^2 \end{pmatrix} \begin{pmatrix} \tilde{\beta}_0 \\ \tilde{\beta}_1 \end{pmatrix} = \begin{pmatrix} \sum\sum Y_{ij} k_h(t - t_{ij}) \\ \sum\sum Y_{ij} k_h(t - t_{ij})(t_{ij} - t)) \end{pmatrix}$$

$$\begin{pmatrix} \tilde{\beta}_0 \\ \tilde{\beta}_1 \end{pmatrix} = \begin{pmatrix} S_{n,0} & S_{n,1} \\ S_{n,1} & S_{n,2} \end{pmatrix}^{-1} \begin{pmatrix} \sum\sum k_h(t-t_{ij})Y_{ij} \\ \sum\sum k_h(t-t_{ij})(t_{ij}-t)Y_{ij} \end{pmatrix}$$

$$= \frac{1}{S_{n,0}S_{n,2}-S_{n,1}^2} \begin{pmatrix} S_{n,2} & -S_{n,1} \\ -S_{n,1} & S_{n,0} \end{pmatrix} \begin{pmatrix} \sum\sum k_h(t-t_{ij})Y_{ij} \\ \sum\sum k_h(t-t_{ij})(t_{ij}-t))Y_{ij} \end{pmatrix}$$

Thus,

$$\tilde{\beta}_0 = \frac{S_{n,2}\sum\sum k_h(t-t_{ij})Y_{ij} - S_{n,1}\sum\sum k_h(t-t_{ij})(t_{ij}-t)Y_{ij}}{S_{n,0}S_{n,2}-S_{n,1}^2}$$

$$= \frac{\sum\sum k_h(t-t_{ij})(S_{n,2}-S_{n,1}(t_{ij}-t))Y_{ij}}{\sum\sum k_h(t-t_{ij})\sum\sum k_h(t-t_{ij})(t_{ij}-t)^2 - \sum\sum k_h(t-t_{ij})(t_{ij}-t)\sum\sum k_h(t-t_{ij})(t_{ij}-t)}$$

$$= \frac{\sum\sum k_h(t-t_{ij})\{S_{n,2}+(t-t_{ij})S_{n,1}\}Y_{ij}}{\sum\sum k_h(t-t_{ij})\sum\sum k_h(t-t_{ij})(t_{ij}-t)^2 + \sum\sum k_h(t-t_{ij})(t-t_{ij})\sum\sum k_h(t-t_{ij})(t_{ij}-t)}$$

$$= \frac{\sum\sum k_h(t-t_{ij})\{S_{n,2}+(t-t_{ij})S_{n,1}\}Y_{ij}}{\sum\sum k_h(t-t_{ij})\{S_{n,2}+(t-t_{ij})S_{n,1}\}}$$

$$= \frac{\sum\sum \omega(t,t_{ij})Y_{ij}}{\sum\sum \omega(t,t_{ij})}$$

∎

Problem 2.

Explain

1. Total regression coefficient estimator

2. Partial regression coefficient estimator

3. Sequential regression coefficient estimator

and when can they be the same.

Solution. The total regression coefficient estimator refers to the coefficients estimated in a regression model when you consider all the predictor variables simultaneously to predict the outcome variable.

The partial regression coefficient estimator measures the expected change in the dependent variable associated with a one-unit change in an independent variable holding the other independent variables constant.

Sequential regression coefficients refer to coefficients estimated in a stepwise regression procedure, which selects variables in a sequence based on certain criteria.

The three estimators are the same in the simple linear regression model because there is only one dependent variable to be considered, thus the three estimators measure the relationship between the single independent variable and the dependent variable in the same way.

∎

Problem 3. Time Series Correlation Structure

The first-order autoregressive model with equally spaced time points:

$$\varepsilon_{i,j} = \rho\varepsilon_{i,j-1} + Z_{i,j} \text{ with } Z_{i,j} \overset{\text{iid}}{\sim} (0, (1-\rho^2)\sigma^2), |\rho| < 1.$$

Show that

$$\mathbf{Cov}[\varepsilon_{i,j}, \varepsilon_{i,j-k}] = \rho^k\sigma^2$$

Solution.

$$
\begin{aligned}
\mathbf{Var}[\varepsilon_{i,j}] &= \mathbb{E}[\varepsilon_{i,j}\varepsilon_{i,j}] \\
&= \mathbb{E}[(\rho\varepsilon_{i,j-1} + Z_{i,j})(\rho\varepsilon_{i,j-1} + Z_{i,j})] \\
&= \mathbb{E}[\rho^2\varepsilon_{i,j-1}^2 + 2\rho\varepsilon_{i,j-1}Z_{i,j} + Z_{i,j}^2] \\
&= (1-\rho^2)\mathbf{Var}[\varepsilon_{i,j}] + 0 + \mathbf{Var}[Z_{i,j}]
\end{aligned}
$$

$$\mathbf{Var}[\varepsilon_{i,j}] = \frac{\mathbf{Var}[Z_{i,j}]}{1-\rho^2} = \sigma^2$$

$$
\begin{aligned}
\mathbf{Cov}[\varepsilon_{i,j}, \varepsilon_{i,j-k}] &= \mathbf{Cov}[\rho\varepsilon_{i,j-1} + Z_{i,j}, \varepsilon_{i,j-k}] \\
&= \mathbf{Cov}[\rho\varepsilon_{i,j-1}, \varepsilon_{i,j-k}] + \mathbf{Cov}[Z_{i,j}, \varepsilon_{i,j-k}] \\
&= \mathbf{Cov}[\rho\varepsilon_{i,j-1}, \varepsilon_{i,j-k}] + 0 \\
&= \mathbf{Cov}[\rho^2\varepsilon_{i,j-2} + \rho Z_{i,j-1}, \varepsilon_{i,j-k}] \\
&= \mathbf{Cov}[\rho^2\varepsilon_{i,j-2}, \varepsilon_{i,j-k}] \\
&\ \ \vdots \\
&= \mathbf{Cov}[\rho^k\varepsilon_{i,j-k}, \varepsilon_{i,j-k}] \\
&= \rho^k\mathbf{Var}[\varepsilon_{i,j-k}] \\
&= \rho^k\sigma^2
\end{aligned}
$$

∎

<div style="border: 1px solid #f8c; border-radius: 8px;">

Problem 4. Cross-Sectional Data

For Cross-Sectional Data

$$Y = \mathbf{X}' \beta + e, e \overset{\text{iid}}{\sim} (0, \sigma^2 \mathbf{I}_n),$$

prove that the least square estimator $\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'Y$ is the best linear unbiased estimator (BLUE) of β.

</div>

Solution. To prove that the least square estimator $\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'Y$ is the best linear unbiased estimator (BLUE) of β, is to show that

(a) $\mathbb{E}[\hat{\beta}|\mathbf{X}] = \beta$,

(b) $\mathbf{Var}[\tilde{\beta}|\mathbf{X}] - \mathbf{Var}[\hat{\beta}|\mathbf{X}] \succeq 0$,

where $\tilde{\beta} = \mathbf{A}'\mathbf{X}$ is an arbitrary linearly unbiased estimator for β.

....

Define

$$\mathbf{A}^{*\prime} := (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'.$$

The Least Square estimator

$$\hat{\beta} := (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'Y)$$
$$= \mathbf{A}^{*\prime}Y.$$

Let $\tilde{\beta}$ be an arbitrary linearly unbiased estimator for β

$$\tilde{\beta} = \mathbf{A}'Y,$$

with

$$\mathbb{E}[\tilde{\beta}|\mathbf{X}] = \mathbb{E}[\mathbf{A}'Y|\mathbf{X}]$$
$$= \mathbb{E}[\mathbf{A}'(\mathbf{X}\beta + e)|\mathbf{X}]$$
$$= \mathbf{A}'\mathbf{X}\beta$$
$$= \beta,$$

which requires that

$$\mathbf{A}'\mathbf{X} = \mathbf{I}_k.$$

If $\mathbf{A} = \mathbf{A}^*$, we have $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X} = \mathbf{I}_k$. Thus the LSE $\hat{\beta}$ is unbiased.

$$\mathbf{Var}[\tilde{\beta}|\mathbf{X}] = \mathbb{E}[(\tilde{\beta} - \mathbb{E}[\tilde{\beta}|\mathbf{X}])(\tilde{\beta} - \mathbb{E}[\tilde{\beta}|\mathbf{X}])'|\mathbf{X}]$$
$$= \mathbb{E}[(\tilde{\beta} - \beta)(\tilde{\beta} - \beta)'|\mathbf{X}]$$

$$\mathbf{Var}[\tilde{\beta}|\mathbf{X}] = \mathbf{Var}[\mathbf{A}'Y|\mathbf{X}]$$
$$= \mathbf{Var}[\mathbf{A}'e|\mathbf{X}]$$
$$= \mathbf{A}'\mathbf{Var}[e|\mathbf{X}]\mathbf{A}$$
$$= \mathbf{A}'\sigma^2\mathbf{I}_n\mathbf{A}$$
$$= \sigma^2\mathbf{A}'\mathbf{A}$$

$$\mathbf{Var}[\hat{\beta}|\mathbf{X}] = \mathbf{Var}[\mathbf{A}^{*\prime} Y|\mathbf{X}]$$
$$= \mathbf{Var}[\mathbf{A}^{*\prime} e|\mathbf{X}]$$
$$= \mathbf{A}^{*\prime}\,\mathbf{Var}[e|\mathbf{X}]\,\mathbf{A}^{*}$$
$$= \mathbf{A}^{*\prime}\,\mathbb{E}[ee^{\prime}|\mathbf{X}]\,\mathbf{A}^{*}$$
$$= \mathbf{A}^{*\prime}\,\sigma^2\mathbf{I}_n\,\mathbf{A}^{*}$$
$$= \sigma^2\,\mathbf{A}^{*\prime}\,\mathbf{A}^{*}$$

$$\mathbf{Var}[\tilde{\beta}|\mathbf{X}] - \mathbf{Var}[\hat{\beta}|\mathbf{X}] = \sigma^2\,\mathbf{A}^{\prime}\,\mathbf{A} - \sigma^2\,\mathbf{A}^{*\prime}\,\mathbf{A}^{*}$$
$$= \sigma^2(\mathbf{A}^{\prime}\,\mathbf{A} - \mathbf{A}^{*\prime}\,\mathbf{A}^{*})$$
$$= \sigma^2(\mathbf{A}^{\prime}\,\mathbf{A} - (\mathbf{X}^{\prime}\,\mathbf{X})^{-1}\,\mathbf{X}^{\prime}((\mathbf{X}^{\prime}\,\mathbf{X})^{-1}\,\mathbf{X}^{\prime})^{\prime})$$
$$= \sigma^2(\mathbf{A}^{\prime}\,\mathbf{A} - (\mathbf{X}^{\prime}\,\mathbf{X})^{-1}\,\mathbf{X}^{\prime}\,\mathbf{X}(\mathbf{X}^{\prime}\,\mathbf{X})^{-1})$$
$$= \sigma^2(\mathbf{A}^{\prime}\,\mathbf{A} - (\mathbf{X}^{\prime}\,\mathbf{X})^{-1})$$
$$= \sigma^2(\mathbf{A}^{\prime}\,\mathbf{A} - \mathbf{I}_k(\mathbf{X}^{\prime}\,\mathbf{X})^{-1}\mathbf{I}_k)$$
$$= \sigma^2(\mathbf{A}^{\prime}\,\mathbf{A} - \mathbf{A}^{\prime}\,\mathbf{X}(\mathbf{X}^{\prime}\,\mathbf{X})^{-1}\,\mathbf{X}^{\prime}\,\mathbf{A})$$
$$= \sigma^2(\mathbf{A}^{\prime}[\mathbf{I}_n - \mathbf{X}(\mathbf{X}^{\prime}\,\mathbf{X})^{-1}\,\mathbf{X}^{\prime}]\,\mathbf{A})$$
$$= \sigma^2\,\mathbf{A}^{\prime}\,\mathbf{M}\,\mathbf{A}\,.$$

For any nonzero vector λ, we have

$$\lambda^{\prime}(\mathbf{Var}[\tilde{\beta}|\mathbf{X}] - \mathbf{Var}[\hat{\beta}|\mathbf{X}])\lambda = \sigma^2\lambda^{\prime}\,\mathbf{A}^{\prime}\,\mathbf{M}\,\mathbf{A}\,\lambda$$
$$= \sigma^2\lambda^{\prime}\,\mathbf{A}^{\prime}\,\mathbf{M}\,\mathbf{M}\,\mathbf{A}\,\lambda$$
$$= \sigma^2(\mathbf{M}\,\mathbf{A}\,\lambda)^{\prime}(\mathbf{M}\,\mathbf{A}\,\lambda) \geq 0,$$

implying that $\mathbf{Var}[\tilde{\beta}|\mathbf{X}] - \mathbf{Var}[\hat{\beta}|\mathbf{X}]$ is positive semi-definite.

Therefore, we can conclude that the least square estimator $\hat{\beta} = (\mathbf{X}^{\prime}\,\mathbf{X})^{-1}\,\mathbf{X}^{\prime}\,Y$ is the best linear unbiased estimator (BLUE) of β.

Projection Matrix:
$$\mathbf{P} := \mathbf{X}\,(\mathbf{X}^{\prime}\,\mathbf{X})^{-1}\,\mathbf{X}^{\prime}$$

Orthogonal Projection Matrix:

$$\mathbf{M} := \mathbf{I}_n - \mathbf{P} = \mathbf{I}_n - \mathbf{X}\,(\mathbf{X}^{\prime}\,\mathbf{X})^{-1}\,\mathbf{X}^{\prime}$$

We can show that $\mathbf{P}$ and $\mathbf{M}$ are both symmetric and idempotent. That is

$$\mathbf{P}^{\prime} = \mathbf{P}$$
$$\mathbf{P}\mathbf{P} = \mathbf{P}$$
$$\mathbf{M}^{\prime} = \mathbf{M}$$
$$\mathbf{M}\,\mathbf{M} = \mathbf{M}$$

$$
\begin{aligned}
\mathbf{P}' &= [\mathbf{X}\,(\mathbf{X}'\,\mathbf{X})^{-1}\,\mathbf{X}']' \\
&= [\mathbf{X}']'[\mathbf{X}\,(\mathbf{X}'\,\mathbf{X})^{-1}]' \\
&= \mathbf{X}[(\mathbf{X}'\,\mathbf{X})^{-1}]'[\mathbf{X}]' \\
&= \mathbf{X}\,[(\mathbf{X}'\,\mathbf{X})']^{-1}[\mathbf{X}]' \\
&= \mathbf{X}\,[\mathbf{X}'\,\mathbf{X}]^{-1}[\mathbf{X}]' \\
&= \mathbf{P} \quad \text{symmetric}
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{P}\mathbf{P} &= \underbrace{\mathbf{P}\,\mathbf{X}}_{\mathbf{X}}(\mathbf{X}'\,\mathbf{X})^{-1}\,\mathbf{X}' \\
&= \mathbf{X}\,(\mathbf{X}'\,\mathbf{X})^{-1}\,\mathbf{X}' \\
&= \mathbf{P} \quad \text{idempotent}
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{M}' &= [\mathbf{I}_n - \mathbf{P}]' \\
&= [\mathbf{I}_n]' - [\mathbf{P}]' \\
&= \mathbf{I}_n - \mathbf{P} \\
&= \mathbf{M} \quad \text{symmetric}
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{M}\,\mathbf{M} &= [\mathbf{I}_n - \mathbf{P}][\mathbf{I}_n - \mathbf{P}] \\
&= [\mathbf{I}_n - \mathbf{P}]\mathbf{I}_n - [\mathbf{I}_n - \mathbf{P}]\mathbf{P} \\
&= \mathbf{I}_n - \mathbf{P} - \mathbf{P} + \mathbf{P} \\
&= \mathbf{I}_n - \mathbf{P} \\
&= \mathbf{M} \quad \text{idempotent}
\end{aligned}
$$

∎

Problem 5.

GLM: $Y = X\beta + \varepsilon$, where $\varepsilon \sim (0, \sigma^2 V)$. Let $A = I - X(X^TX)^{-1}X^T$ and B be the $N \times (N-(p+1))$ matrix such that $BB^T = A$ and $B^TB = I_{N-(p+1)}$.
For any fixed V, the MLE of β is $\hat{\beta}(V) = GY$, with $G = (X^TV^{-1}X)^{-1}X^TV^{-1}$ Show that

$$
|GG^T - GBB^TG^T|^{\frac{1}{2}} = |X^TX|^{-\frac{1}{2}}
$$

Solution.

$$
\begin{aligned}
GG^T - GBB^TG^T &= GG^T - GAG^T \\
&= GG^T - G[I - X(X^TX)^{-1}X^T]G^T \\
&= GG^T - GG^T + GX(X^TX)^{-1}X^TG^T \\
&= GX(X^TX)^{-1}X^TG^T
\end{aligned}
$$

$$
\begin{aligned}
|GG^T - GBB^TG^T|^{\frac{1}{2}} &= |GX(X^TX)^{-1}X^TG^T|^{\frac{1}{2}} \\
&= |G|^{\frac{1}{2}}|X(X^TX)^{-1}X^T|^{\frac{1}{2}}|G^T|^{\frac{1}{2}} \\
&= |G||X(X^TX)^{-1}X^T|^{\frac{1}{2}} \\
&= |(X^TV^{-1}X)^{-1}X^TV^{-1}||X|^{\frac{1}{2}}|(X^TX)^{-1}|^{\frac{1}{2}}|X^T|^{\frac{1}{2}} \\
&= \frac{1}{|X^TV^{-1}X|}|X^TV^{-1}||X|^{\frac{1}{2}}|(X^TX)^{-1}|^{\frac{1}{2}}|X^T|^{\frac{1}{2}} \\
&= |(X^TX)^{-1}|^{\frac{1}{2}} \\
&= |X^TX|^{-\frac{1}{2}}
\end{aligned}
$$

∎

Problem 6. Serial correlation plus mesurement error

When $\varepsilon(t)$ is stationary, the covariance function or variogram :

$$r(u) = \frac{1}{2} \mathbb{E}[(\varepsilon(t) - \varepsilon(t - u))^2].$$

Sources of random variation in longitudinal data:

Serial correlation + Mesurement error mesurement error: $\varepsilon \sim (0, \sigma^2 H + \tau^2 I_N)$.

Show that the corresponding variogram is $r(u) = \tau^2 + \sigma^2(1 - \rho(u))$ for $u \geq 0$, with $r(0) = \tau^2 > 0$.

Solution.

$$\begin{aligned}
r(u) &= \frac{1}{2} \mathbb{E}[(\varepsilon(t) - \varepsilon(t - u))^2] \\
&= \frac{1}{2} \mathbb{E}[\varepsilon(t)^2] + \frac{1}{2} \mathbb{E}[\varepsilon(t - u)^2] - \mathbb{E}[\varepsilon(t)\varepsilon(t - u)] \\
&= \sigma^2 + \tau^2 - \sigma^2 \rho(u) \\
&= \sigma^2(1 - \rho(u)) + \tau^2,
\end{aligned}$$

where $\rho(0) = 1$.

■

Problem 7. Repeated Measures

$$y_{hij} = \beta_h + r_{hj} + U_{hi} + z_{hij}, h = 1, \ldots, g, i = 1, \ldots, n_h, j = 1, \ldots, m,$$

where β_h is the main effect for the h-th treatment, and r_{hj} is the interaction effect between the h-th treatment and the j-th time with $\sum_{j=1}^{m} r_{hj} = 0, \forall h$.

For the hypothesis $H_0 : \beta_1 = \cdots \beta_g$ find the F statistics of the ANOVA table.

Solution.

Table 1: ANOVA Table

Source of Variation	Sum of Squares (SS)	Degrees of Freedom (df)	Mean Square (MS)
Between Treatments	SSB $= \sum_{h=1}^{g} (n_h m)(\bar{y}_{h\cdot\cdot} - \bar{y}_{\cdots})^2$	df$_{\text{between}} = g - 1$	MSB $=$ SSB/df$_{\text{between}}$
Within Treatments	SSW $= \sum_{h=1}^{g} \sum_{i=1}^{n_h} m(\bar{y}_{hi\cdot} - \bar{y}_{h\cdot\cdot})^2$	df$_{\text{within}} = \sum_{h=1}^{g} n_h - g$	MSW $=$ SSW/df$_{\text{within}}$
Total	SST $= \sum_{h=1}^{g} \sum_{i=1}^{n_h} m(\bar{y}_{hi\cdot} - \bar{y}_{\cdots})^2$	df$_{\text{total}} = \sum_{h=1}^{g} n_h - 1$	

Table 2: F-statistic

F-statistic
F $=$ MSB/MSW

■