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

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

Solution.

■

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. ■

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

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

Problem 4. Cross-Sectional Data

For Cross-Sectional Data

$$Y_i = X_i^T\beta + \varepsilon_i, \varepsilon \overset{\text{iid}}{\sim} (0, \sigma^2),$$

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

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

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

Problem 6. Serial rorrelation plus 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 $r(0) = \tau^2 > 0$.

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 $F = \frac{MBTSS_1}{MRSS_1}$ of the ANOVA table.

Solution.

■