Automated Method for Nadaraya-Watson Estimates

Original Model:

$$Y_i = m(x_i) + e_i$$

where $\mathrm{Var}(e) = \sigma^2 I$. Because of correlated errors, it is violated.

Formula: Corrected GCV

$$\mathrm{GCV}(\lambda) = \mathrm{RSS}/(1 - \frac{1}{n}\mathrm{tr}(\mathrm{S}_\lambda \mathrm{A}))^2$$

Steps

1. Fit a model with a small bandwidth (but not overfitting)
2. Compute the sample autocorrelations of residuals and estimate A
3. Minimize the corrected GCV

Ref. ALTMAN. *Kernel Smoothing of Data With Correlated Errors*.

Extension: Bimodal Kernel

Main Theorem

$$\mathbb{E}[\mathrm{LCV}(h)] = \frac{1}{n}\mathbb{E}\left[\sum_{i=1}^{n}\left(m(x_i) - \hat{m}_n^{(-i)}(x_i)\right)^2\right] + \sigma^2 - \frac{4K(0)}{nh - K(0)}\sum_{k=1}^{\infty}\gamma_k + o\left(n^{-1}h^{-1}\right)$$

where $\gamma_k := \mathrm{Cov}(e_i, e_{i-k})$.

Steps

1. Estimate $\hat{m}$ using bimodal kernel.
2. Estimate the length l of correlation.
3. Use leave-(2l+1)-out cross-validation with unimodal kernel.

Ref. *Kernel Regression in the Presence of Correlated Errors*

Comparison among Previous Results