1 λ dependent on Y

using setup from Hall & Racine 2015

"Data pairs (x_i, y_i) are assumed to be generated by the model, $y_i = g(x_i) + \epsilon_i$ where $x_1, ..., x_n$ are independent and identically distributed as x, with density $f(x)$ supported on a compact interval I, and the experimental errors ϵ_i are independent and identically distributed with zero mean finite variance, independent too of the x_is. " g arbitrary, unspecified

S as estimate of g, $S := \hat{g}(\lambda, X)$

$\hat{g}_\lambda(x_i) = S_{i,\lambda}Y$ where S_i is the i'th row of S_λ, Y the column vector of regressands, and $\hat{g}_\lambda^{-i}(x_i)$ is the estimate for g with the i'th data point removed

minimizing

$$GCV(\lambda) = \frac{1}{n}\sum_{i=1}^{n}\Big(\frac{y_i - \hat{g}_\lambda(x_i)}{1 - tr(S_\lambda)/n}\Big)^2 = \frac{n(Y - S_\lambda Y)^\top (Y - S_\lambda Y)}{(n - tr(S_\lambda))^2}$$

yields

$$\frac{\partial GCV}{\partial \lambda} = 0 \to tr(\frac{\partial S_\lambda}{\partial \lambda})\frac{\hat{\epsilon}^\top \hat{\epsilon}}{n - tr(S_\lambda)} - (\frac{\partial S_\lambda}{\partial \lambda}Y)^\top \hat{\epsilon} = 0$$

where $\hat{\epsilon} = Y - S_\lambda Y$. need $\frac{\partial^2 GCV}{\partial \lambda^2}$ to use newton's method to find λ

2 function spaces

for spline, s(x), is continuous in X and has continuous derivatives in X up to order-1, with $\int (s^{(order-1)})^2 dx < \infty$

for local polynomials with polynomial kernels, the estimating functions are only continuous in X with $\int (p(x)^{(q)})^2 dx < \infty$ for q = 1,...,m , where m is the sum of the degree of local polynomial and twice the degree of the kernel.

for local polynomials with the Gaussian kernel, the estimating functions belong to C^∞ and L^2

so the estimating functions are from different function spaces with different continuity conditions.

3 smoothness

column rank of X_0 increasing in smoothness (ie $[1 \;\; X - X_0 \;\; (X - X_0)^2 \;\; (X - X_0)^3 \;\; ...]$, thus $tr(X_0(X_0^\top X_0)^{-1}X_0^\top)$ is increasing in smoothness

λ can impart smoothness without adding higher order columns to X_0