Quantile Regression

Ruibin Xi

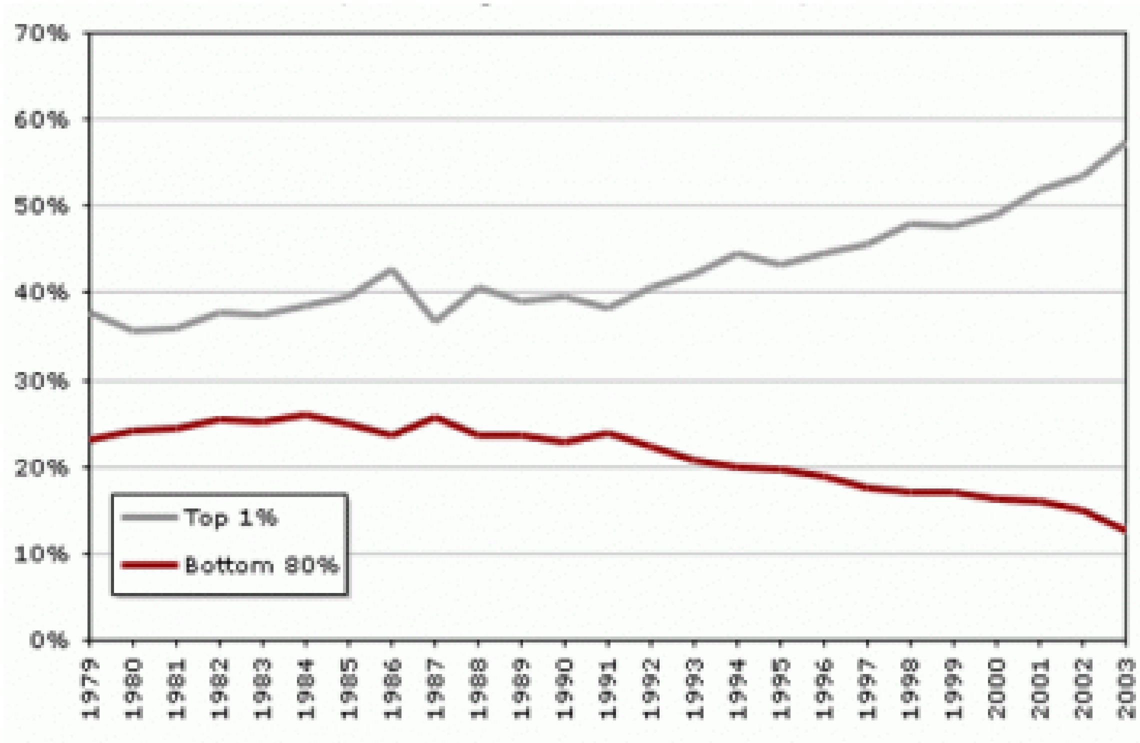
Average derivative estimation (1)

- In mean regression, we focus on estimating the conditional mean

$$\mu(\mathbf{x}) = E(Y|\mathbf{x})$$

- In linear regression, the partial derivatives $\partial\mu(\mathbf{x})/\partial x_i$ are assumed to be constant.
- They are of primary interest because they measure how much the mean response change as the *i*th covariate perturb one unit.

Average derivative estimation (2)
Rich got richer, poor got poorer?



Share of capital income earned by top 1% and bottom 80%, 1979-2003 (Shapiro and Friedman 2006)

Average derivative estimation (3)

- For quantile regression, we may consider the average gradient

$$\beta_\alpha = (\beta_{\alpha 1}, \ldots, \beta_{\alpha d}) = E(\nabla \theta_\alpha(\mathbf{X}))$$

where $\theta_\alpha(\mathbf{X})$ is the conditional αth of Y given **X**

Average derivative estimation (4)

- Consider the model

$$Y = \mu(\mathbf{X}) + \tau[\,\mu(\mathbf{X})]^{\lambda}\varepsilon$$

ε and $\mathbf{X}$ are independent

ε has continuous distribution function F_ε

the mean of ε is zero

τ and λ are real parameters

Average derivative estimation (5)

- Let e_α be the αth quantile of F_ε

$$\theta_\alpha(\mathbf{x}) = \mu(\mathbf{x}) + \tau[\mu(\mathbf{x})]^\lambda e_\alpha$$

$$\nabla\theta_\alpha(\mathbf{x}) = \nabla\mu(\mathbf{x}) + \tau\lambda[\mu(\mathbf{x})]^{\lambda-1}\nabla_\mu(\mathbf{x})e_\alpha$$

$$\beta_\alpha = E(\nabla\mu(\mathbf{X})) + \tau\lambda E\{[\mu(\mathbf{X})]^{\lambda-1}\nabla\mu(\mathbf{X})\}e_\alpha$$

- If assuming $F_\varepsilon = \Phi$

$$d = \tau = \lambda = 1$$

$$\mu(x) = \gamma_1 + \gamma_2 x$$

then

$$\beta_\alpha = [1 + \Phi^{-1}(\alpha)]\gamma_2$$

$$\beta_{0.1} = -0.282\gamma_2, \qquad \beta_{0.5} = \gamma_2, \qquad \beta_{0.9} = 2.282\gamma_2$$

Average derivative estimation (6)

- Taking derivative on both side of the above equation and taking expectation, we get

$$
\begin{aligned}
E(\omega(\boldsymbol{X})\nabla\theta_\alpha(\boldsymbol{X})) &= \left[\int g'(\gamma^t \boldsymbol{x})\omega(\boldsymbol{x})f(\boldsymbol{x})d\boldsymbol{x}\right]\gamma \\
&= \beta\gamma
\end{aligned}
$$

Average derivative estimation (7)

- We may estimate

$$E(\omega(\boldsymbol{X})\nabla\theta_\alpha(\boldsymbol{X})) = \int \nabla\theta_\alpha(\boldsymbol{X})\omega(\boldsymbol{x})f(\boldsymbol{x})d\boldsymbol{x}$$

by $\hat{\beta}_1 = n^{-1}\sum\{\nabla\hat{\theta}(\mathbf{X}_i)\}w(\mathbf{X}_i)$,

where $\nabla\hat{\theta}(\mathbf{X}_i)$ is a nonparametric estimator of the gradient of the conditional quantile $\theta(\mathbf{x})$

Average derivative estimation (8)

- By integration by parts,

$$\begin{aligned} E(\omega(\boldsymbol{X})\nabla\theta_\alpha(\boldsymbol{X})) &= \int \nabla\theta_\alpha(\boldsymbol{X})\omega(\boldsymbol{x})f(\boldsymbol{x})d\boldsymbol{x} \\ &= -\int \theta_\alpha(\boldsymbol{X})\nabla\{\omega(\boldsymbol{x})f(\boldsymbol{x})\}d\boldsymbol{x} \end{aligned}$$

- An alternative estimator is

$$\begin{aligned} \hat{\beta}_2 &= -\frac{1}{n}\sum_{i=1}^{n}\hat{\theta}(\mathbf{X}_i)\frac{\nabla w(\mathbf{X}_i)\hat{f}(\mathbf{X}_i) + w(\mathbf{X}_i)\nabla\hat{f}(\mathbf{X}_i)}{\hat{f}(\mathbf{X}_i)} \\ &= -\frac{1}{n}\sum_{i=1}^{n}\hat{\theta}(\mathbf{X}_i)\Big\{\nabla w(\mathbf{X}_i) + w(\mathbf{X}_i)\hat{\ell}(\mathbf{X}_i)\Big\}, \end{aligned}$$

where $\hat{\ell}(\mathbf{X}_i) = \nabla\hat{f}(\mathbf{X}_i)/\hat{f}(\mathbf{X}_i)$, $\hat{f}$ and $\nabla\hat{f}$ are nonparametric estimator of the density and its derivative

Average derivative estimation (9)

- Leave one out estimator

$$\hat{f}(\mathbf{X}_i) = \frac{1}{(n-1)h_n^d} \sum_{j \neq i} W\left(\frac{\mathbf{X}_j - \mathbf{X}_i}{h_n}\right),$$

$$\nabla\hat{f}(\mathbf{X}_i) = \frac{1}{(n-1)h_n^{d+1}} \sum_{j \neq i} W^{(1)}\left(\frac{\mathbf{X}_j - \mathbf{X}_i}{h_n}\right),$$

Penalized univariate smoothing method (1)

- Assume that

$$y = f(x) + \epsilon$$

- In case of the mean regression, we may consider estimate the function f bv minimizing

$$\mathrm{RSS}(f, \lambda) = \sum_{i=1}^{N} \{y_i - f(x_i)\}^2 + \lambda \int \{f''(t)\}^2 dt,$$

$\lambda = 0:$ f can be any function that interpolates the data.

$\lambda = \infty:$ the simple least squares line fit, since no second derivative can be tolerated.

- It can be shown that for $\lambda \in (0, \infty)$, there is a unique minimizer, which is a natural cubic spline with knots at x_i

Splines (1)

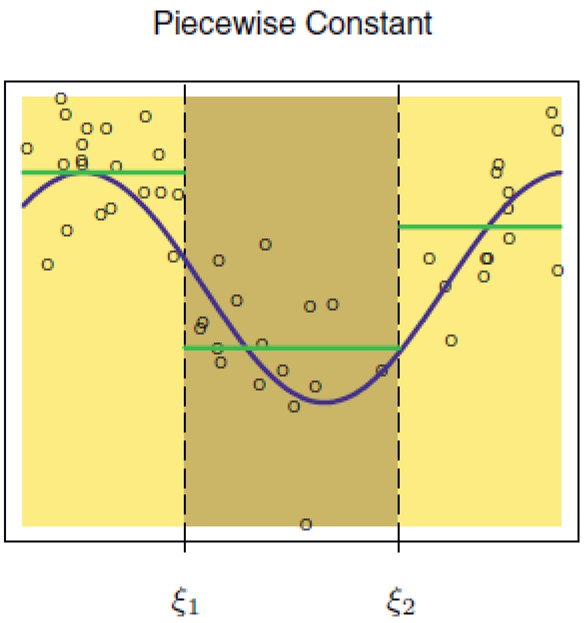
- Suppose that we can to estimate a function f(.) with a piecewise polynomial function
- In the simplest case, we can just estimate f by a piecewise constant function
- We may write

$$f(X) = \sum_{m=1}^{3} \beta_m h_m(X)$$

$$h_1(X) = I(X < \xi_1),$$

$$h_2(X) = I(\xi_1 \leq X < \xi_2),$$

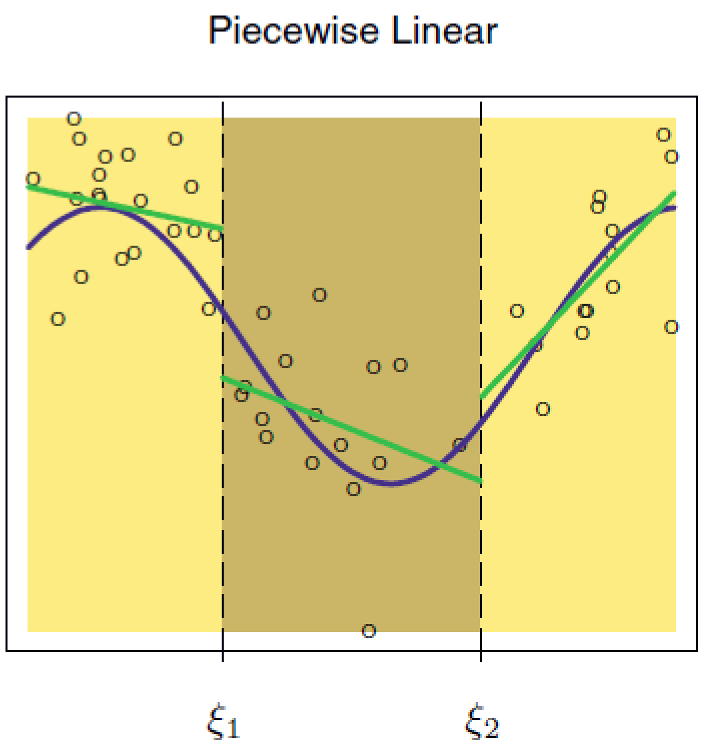
$$h_3(X) = I(\xi_2 \leq X).$$



Splines (2)

- We may also fit f by a piecewise linear function, which requires 3 additional base function

$$h_{m+3} = h_m(X)X, \; m = 1, \ldots, 3$$

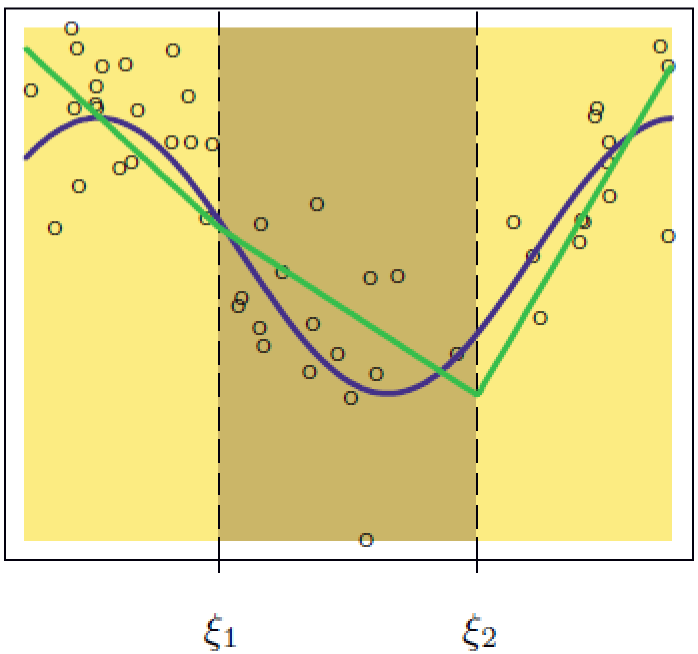


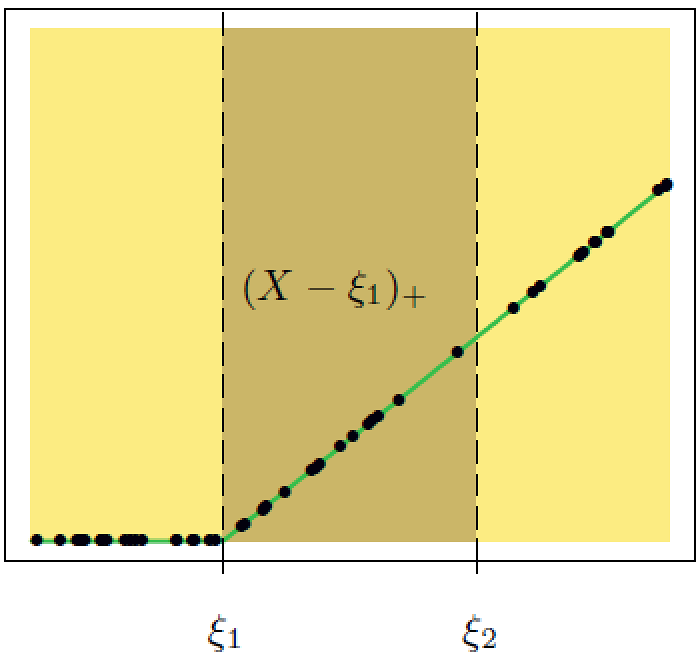
Splines (3)

- If we require continuity on the knots, we put some constraints (2 constraints in the above exmple) on the coefficient of

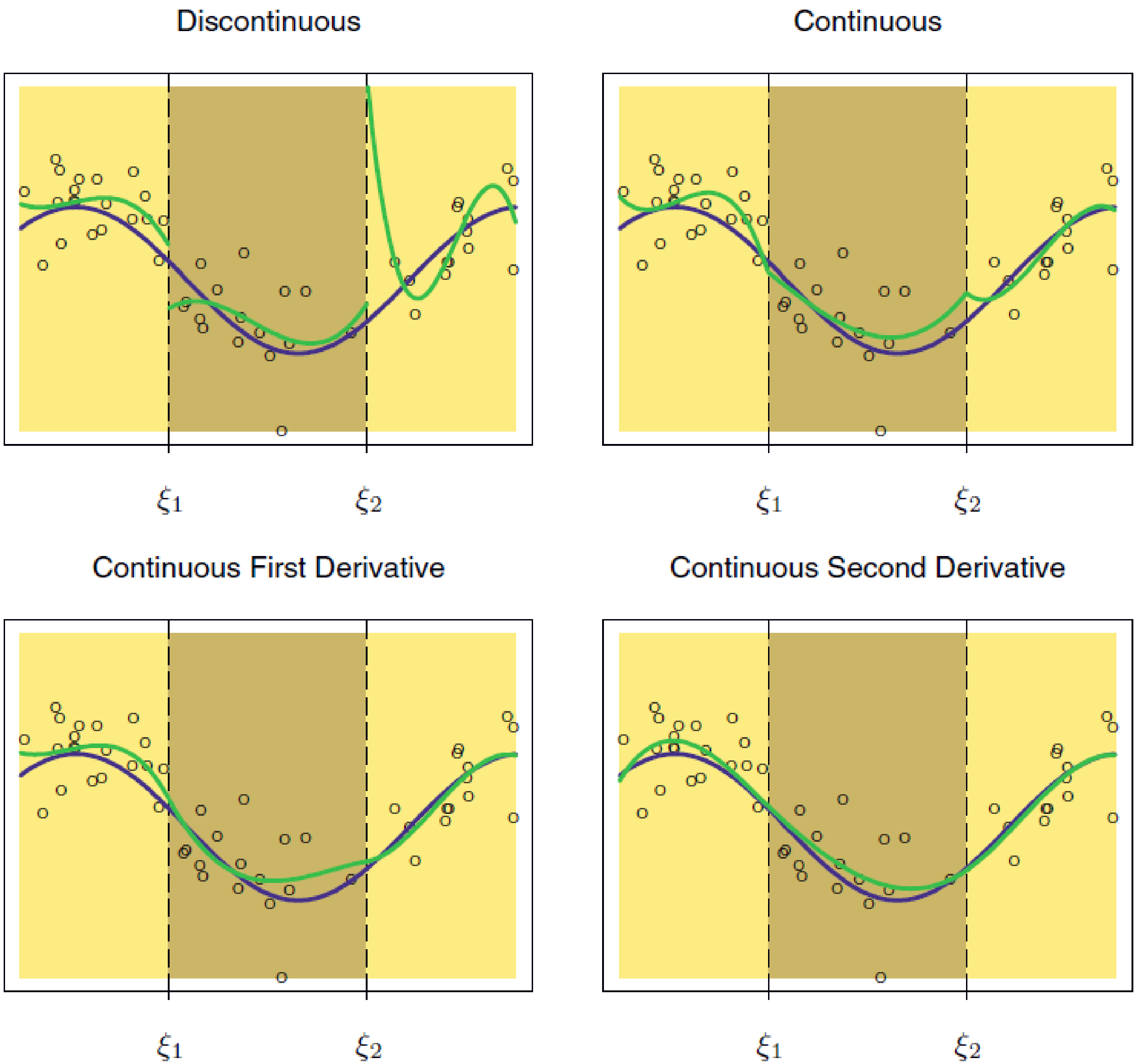
$$f(X) = \sum_{m=1}^{M} \beta_m h_m(X),$$

- More conveniently, we may write the base functions





Splines (4)



Piecewise
cubic polynomial fit

$$h_1(X) = 1,$$
$$h_2(X) = X,$$

$$h_3(X) = X^2$$
$$h_4(X) = X^3$$

$$h_5(X) = (X - \xi_1)_+^3,$$
$$h_6(X) = (X - \xi_2)_+^3.$$

Splines (5)

- An order-M spline with knots $\xi_j,\ j = 1, \ldots, K$ is a piecewise-polynomial of order M (degree M-1), has continuous derivative up to order M-2
- A cubic spline has order M=4
- The bases are

$$
\begin{aligned}
h_j(X) &= X^{j-1},\ j = 1, \ldots, M, \\
h_{M+\ell}(X) &= (X - \xi_\ell)_+^{M-1},\ \ell = 1, \ldots, K.
\end{aligned}
$$

Natural cubic splines

- In addition to requiring the function to have continuous derivatives on the knots, we require the function is linear beyond the boundary points
- If the function is represented as

$$f(X) = \sum_{j=0}^{3} \beta_j X^j + \sum_{k=1}^{K} \theta_k (X - \xi_k)_+^3.$$

it can be shown that

$$\begin{aligned} \beta_2 &= 0, \qquad & \textstyle\sum_{k=1}^{K} \theta_k &= 0, \\ \beta_3 &= 0, \qquad & \textstyle\sum_{k=1}^{K} \xi_k \theta_k &= 0. \end{aligned}$$

B-splines (1)

Let $\xi_0 < \xi_1$ and $\xi_K < \xi_{K+1}$ be two *boundary* knots.

$\tau_1 \leq \tau_2 \leq \cdots \leq \tau_M \leq \xi_0;$

$\tau_{j+M} = \xi_j, \ j = 1, \cdots, K;$

$\xi_{K+1} \leq \tau_{K+M+1} \leq \tau_{K+M+2} \leq \cdots \leq \tau_{K+2M}.$

$$B_{i,1}(x) = \begin{cases} 1 & \text{if } \tau_i \leq x < \tau_{i+1} \\ 0 & \text{otherwise} \end{cases}$$

for $i = 1, \ldots, K + 2M - 1.$

$$B_{i,m}(x) = \frac{x - \tau_i}{\tau_{i+m-1} - \tau_i} B_{i,m-1}(x) + \frac{\tau_{i+m} - x}{\tau_{i+m} - \tau_{i+1}} B_{i+1,m-1}(x)$$

for $i = 1, \ldots, K + 2M - m.$

B-splines (2)

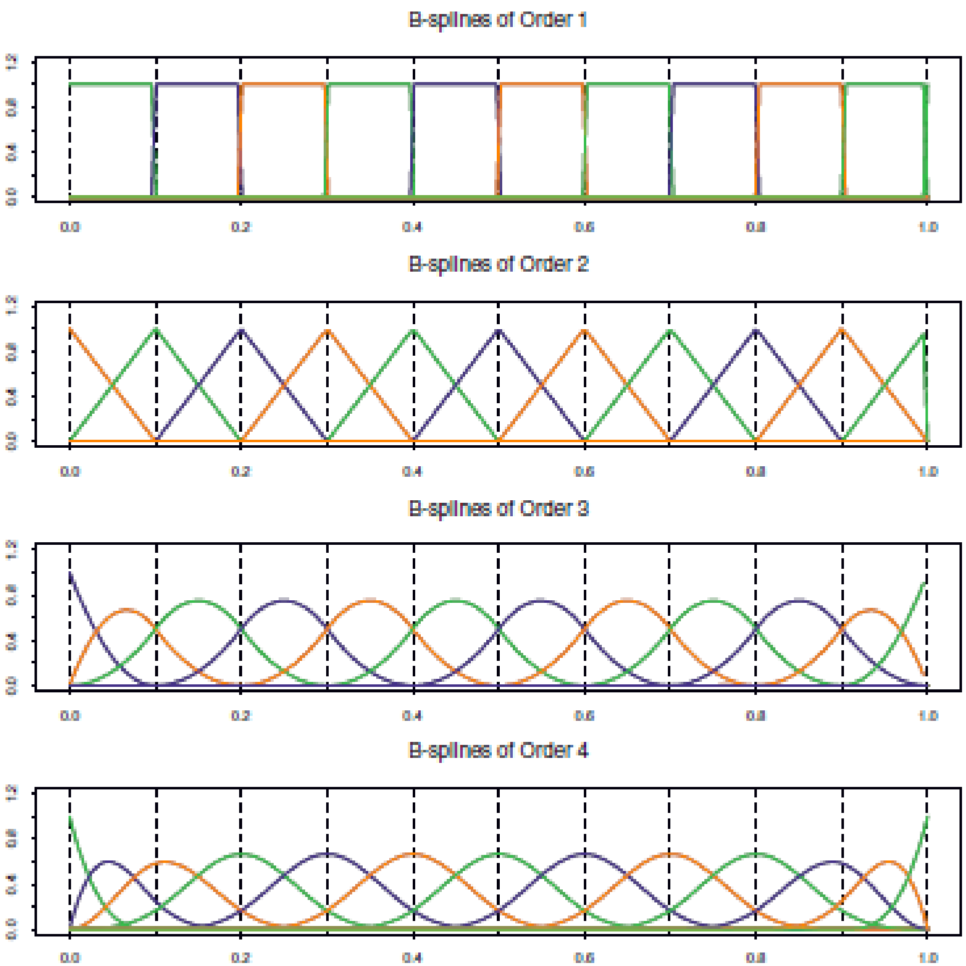


FIGURE 5.20. *The sequence of B-splines up to order four with ten knots evenly spaced from 0 to 1. The B-splines have* local support*; they are nonzero on an interval spanned by $M + 1$ knots.*

Quantile regression-penalized method (1)

- For quantile regression, we may consider

$$\min_{g \in \mathcal{G}} \sum_{i=1}^{n} \rho_\tau(y_i - g(x_i)) + \lambda \int (g''(x))^2 dx.$$

- The solution is also a natural cubic spline

Quantile regression-penalized method (2)

Koenker, Ng, and Portnoy (1994) consider other L_p penalties

$$J(g) = \|g''\|_p = (\int (g''(x))^p)^{1/p}.$$

For $p = 1$

$$\min \sum_{i=1}^{n} \rho_\tau(y_i - g(x_i)) + \lambda \int |g''(x)| dx$$

Quantile regression-penalized method (3)

- Another way to penalize the objective function is

$$\mathrm{P}(\mathrm{g}) = \mathrm{V}(\mathrm{g}')$$

- The total variation is defined as

$$V(f) = \sup \sum_{i=1}^{n} |f(x_{i+1}) - f(x_i)|,$$

where the sup is taken over all partitions $a \leq x_1 < \cdots < x_n < b$

- For absolute continuous function

$$V(f) = \int_a^b |f'(x)|dx$$

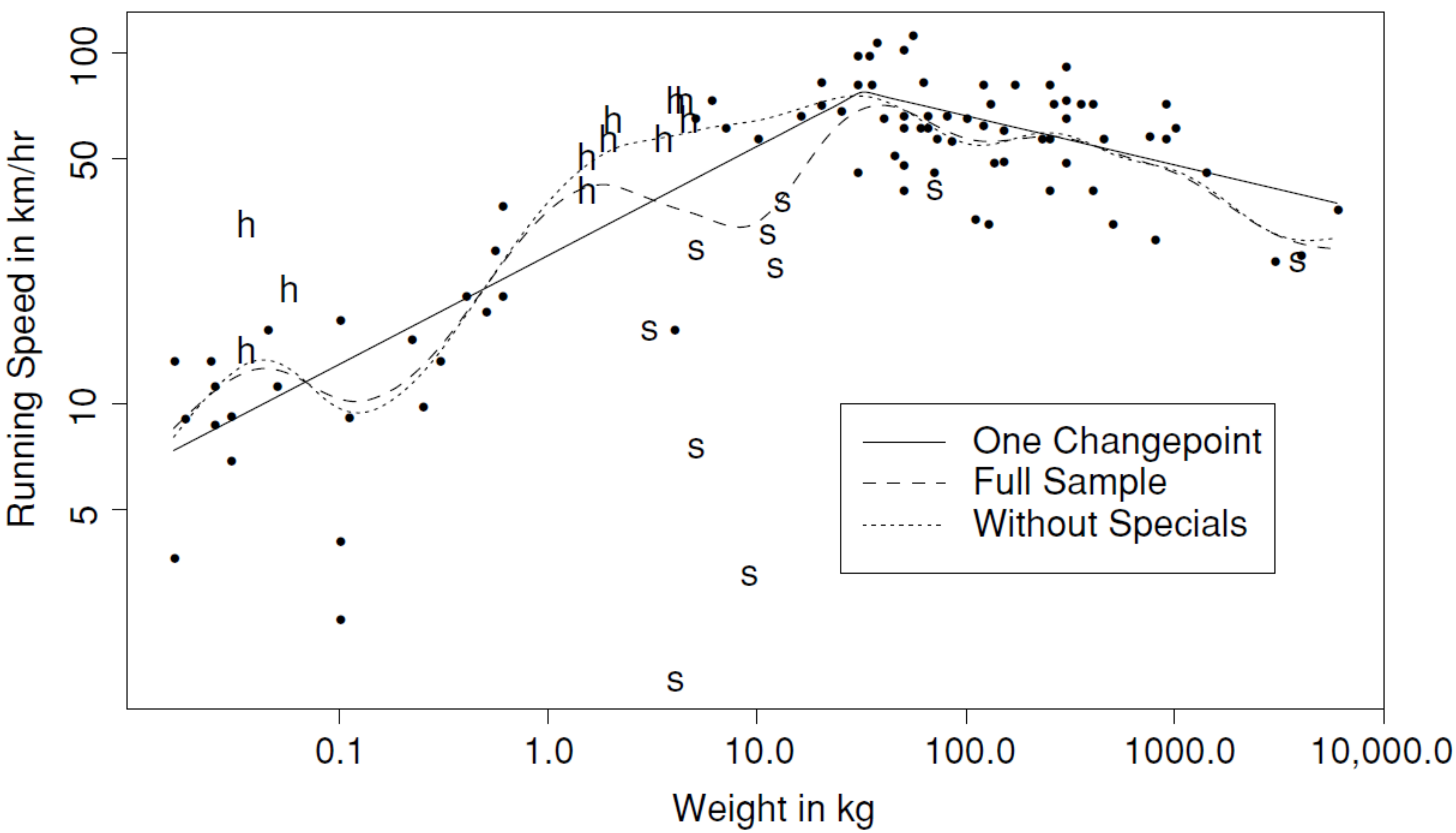
Quantile regression-penalized method (4)

- The new penalized objective function is

$$\min_{g \in \mathcal{G}} \sum_{i=1}^{n} \rho_\tau(y_i - g(x_i)) + \lambda V(g')$$

- The solution is a piecewise linear function

Animal weight VS running speed



h: hoppers

s: specials including sloth, porcupine, hippopotamus