Nonparametric Econometrics

Methods I

Aman Ullah

University of California, Riverside

Books

- **A. Pagan and A. Ullah**, Nonparametric Econometrics, Cambridge University Press, 1999.
- **Q. Li and J. Racine**, Nonparametric Econometrics: Theory and Practice, Princeton University Press, 2007.
- **D. Henderson and C.F. Parmeter**, Applied Nonparametric Econometrics, Cambridge University Press, 2015.
- **J. Fan and I. Gilbels**, Local Polynomial Modelling and Its Applications, Chapman and Hall, 1996.
- **B.W. Silverman,** Density Estimation for Statistics and Data Analysis, Chapman and Hall, 1986.
- **B.L.S. Prakasa-Rao,** Nonparametric Functional Estimation, Academic Press, 1983.

Survey Papers

- Zongwu Cai and Yongmiao Hong, Some Recent Developments in Nonparametric Finance, *Advances in Econometrics* (Racine and Li), 2009.
- Zongwu Cai and Jingping Gu and Qi Li, Some Recent Developments on Nonparametric Econometrics, *Advances in Econometrics* (Racine and Li), 2009.
- J. Racine and A. Ullah, Nonparametric Econometrics, *Palgrave Handbook of Econometrics* (Mills and Patterson), Palgrave Macmillan, 2006.
- J.L. Powell, Estimation of Semiparametric Models, *Handbook of Econometrics*, Vol. IV (Engle and MaFadden), 1994.

Topics

1. Kernel Estimation of Density
2. Conditional Mean (Regression)
3. Time Series Conditional Variance (Volatility) and Conditional Correlations
4. Nonparametric Hypothesis Testing
5. Semiparametric Models
6. Empirical Examples: Financial Time Series Models of Volatility and Correlations, Earning Functions, Income Distributions, Panel Data

Softwares

- R-program

  http://www.r-project.org/

- Racine 'np', 'npRmpi', 'crs' packages available from above link.
- Eviews,Stata

Functions (Models) in Econometrics

$\{Y, X\}$,

$$
\begin{aligned}
Y &= E(Y|X) + U \\
&= m(X) + U
\end{aligned}
$$

- $f(X)$, $f(Y)$, $f(X, Y)$, $f(Y|X)$
- $m(X) = E(Y|X = x)$: REGRESSION FUNCTION
- $\beta(x) = \frac{\partial m(x)}{\partial x}$: REGRESSION COEFFICIENT FUNCTION
- $C(x) = \frac{\partial^2 m(x)}{\partial x^2}$: CURVATURE FUNCTION
- $\sigma^2(x) = V(Y|X = x)$: VARIANCE (VOLATILITY)
- $\sigma_{Y,Z}(x) = cov(Y, Z|X = x)$: COVARIANCE FUNCTION
- LIKELIHOOD FUNCTION, SCORE FUNCTION

$$m(x) = E(y|x) = \int_y yf(y|x)dy = \int_y y\frac{f(y,x)}{f(x)}dy$$

$$V(y|x) = E[(y - m(x)^2)|x] = \int_y (y - m(x))^2 \frac{f(y,x)}{f(x)}dy$$

$$\begin{aligned}
cov(y_1, y_2|x) &= E[(y_1 - m_1(x))(y_2 - m_2(x))|x] \\
&= \int_{y_1}\int_{y_2} (y_1 - m_1(x))(y_2 - m_2(x))\frac{f(y_1, y_2, x)}{f(x)}dy_1 dy_2 \\
m_1(x) &= E(y_1|x), \quad m_2(x) = E(y_2|x)
\end{aligned}$$

Distribution Function-Parametric

X : logwage

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$\{X_i\}$, $i = 1, 2, ..., n$

$$\hat{f}(x) = \frac{1}{\hat{\sigma}\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\hat{\mu}}{\hat{\sigma}}\right)^2}$$

$\hat{\mu} = \bar{x} = \frac{1}{n}\sum x_i$; $\hat{\sigma}^2 = \frac{1}{n}\sum(x_i - \bar{x})^2$

Calculate for each $x_1, x_2, ..., x_n$.

Distribution Function-Nonparametric [Data Based]

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K_h(x_i - x)$$

$$\begin{aligned} K_h(x_i - x) &= K\left(\frac{x_i - x}{h}\right) = I\left(\frac{x_i - x}{h}\right) = 1 \text{ if } -\frac{1}{2} \le \frac{x_i - x}{h} \le \frac{1}{2} \\ &= 0 \text{ if otherwise} \end{aligned}$$

$$\begin{aligned} \hat{f}(x) &= \frac{\text{number of data } x_i \text{ in } [x - \frac{h}{2}, x + \frac{h}{2}]}{nh} = \frac{n^*}{nh} \\ &= \text{per unit relative frequency (proportion)} \end{aligned}$$

Distribution Function-Nonparametric [Data Based]

Empirical Density (Local Histogram): Jumps at the end of interval and 0 derivatives elsewhere.

Smooth Density

(i) $K(\frac{x_i - x}{h}) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x_i - x}{h})2}$, $-\infty < \frac{x_i - x}{h} < \infty$
(ii) $K(\frac{x_i - x}{h}) = \frac{3}{4}[1 - (\frac{x_i - x}{h})^2]$, $\left|\frac{x_i - x}{h}\right| \leq 1$
$K(\frac{x_i - x}{h}) \uparrow$ if $\frac{x_i - x}{h} \uparrow$; $K(\frac{x_i - x}{h}) \downarrow$ if $\frac{x_i - x}{h} \downarrow$
$h = 1.06\sigma_x n^{-1/5}$, $1.06\sigma_x n^{-1/(q+4)}$.

PROPERTIES

Assumptions (p.21)

1. $i.i.d.$
2. $f(x)$ is continuous and bounded
3. Second order Kernel:
$\int K(\psi)d\psi = 1, \int \psi K(\psi)d\psi = 0, \int \psi^2 K(\psi)d\psi = \mu_2 < 0$
4. $h \longrightarrow 0,\ nh \longrightarrow \infty$ as $n \longrightarrow \infty$

$$\begin{aligned} BIAS(\hat{f}(x)) &= \frac{h^2}{2} f^{(2)}(x)\mu_2 = O(h^2) \\ V(\hat{f}(x)) &= \frac{1}{nh} f(x) \int K^2(\psi)d\psi = O(\frac{1}{nh}) \end{aligned}$$

Theorem 2.2, p.23.

PROPERTIES

$$MSE(\hat{f}(x)) = (BIAS)^2 + V(\hat{f}(x))$$

$$\begin{aligned} IMSE &= MISE = \int_x MSE(\hat{f}(x))dx \quad (2.46) \\ &= \int_x (BIAS)^2 dx + \int_x V(\hat{f}(x))dx \end{aligned}$$

$$h_{opt} = cn^{-\frac{1}{5}} \qquad (2.49)$$

$$c = [\frac{\int K^2(\psi)d\psi}{\mu_2^2 \int (f^{(2)}(x))^2 dx}]^{1/5}$$

$f(x) \longrightarrow N(\mu, \sigma_x^2)$, $K(\psi) \longrightarrow N(0,1)$
$c = 1.06\sigma_x$

2-step (plug-in): $c = \hat{c}$, substitute $\hat{f}^{(2)}(x)$ for $f^{(2)}(x)$
$K_{opt} = K(\psi) = \frac{3}{4}(1-\psi^2)$; $-1 \leq \psi \leq 1 \qquad (2.61)$.

CROSS-VALIDATION

p.51

$$
\begin{aligned}
& \min_c \int_x (\hat{f}(x) - f(x))^2 dx \\
= \quad & \min_c ISE \\
= \quad & \min_c \left[\int_x \hat{f}^2(x) dx - \frac{2}{nh(n-1)} \sum_i \sum_{j \neq i} K\left(\frac{x_j - x_i}{h}\right) \right]
\end{aligned}
$$

where $h = cn^{-1/5}$.

MULTIVARIATE

$X : \hat{f}(x)dx = \hat{f}(x)h = \frac{1}{n}\sum_{i=1}^{n} K(\frac{x_i - x}{h})$

$X_1 : \hat{f}(x_1)dx_1 = \hat{f}(x_1)h = \frac{1}{n}\sum_{i=1}^{n} K(\frac{x_{1i} - x_1}{h})$

$X_1, X_2 : \hat{f}(x_1, x_2)dx_1 dx_2 = \hat{f}(x_1, x_2)h^2 = \frac{1}{n}\sum_{i=1}^{n} K(\frac{x_{1i} - x_1}{h}, \frac{x_{2i} - x_2}{h})$

$X = [X_1, X_2, ..., X_q]; \; x = (x_1, x_2, ..., x_q)$

$\hat{f}(x_1, x_2, ..., x_q)h^q = \hat{f}(x)h^q = \frac{1}{n}\sum_{i=1}^{n} K(\frac{x_{1i} - x_1}{h}, \frac{x_{2i} - x_2}{h}, ..., \frac{x_{qi} - x_q}{h})$ or

$$\hat{f}(x) = \frac{1}{nh^q}\sum_{i=1}^{n} K(\frac{x_i - x}{h}).$$

ASYMPTOTIC PROPERTIES

ASSUMPTIONS:

A.1. Let $\Re$ be the class of all Borel-Measurable bounded real functions $K(x), x = (x_1, ..., x_q)'$.
(i) $\int K(x)dx = 1$, (ii) $\int |K(x)|\, dx < \infty$, (iii) $\|x\|^q |K(x)| \to 0$ as $\|x\| \to \infty$, (iv) $\sup |K(x)| < \infty$,where $\|x\|$ is E.Norm.

Examples: $K(x) = (2\pi)^{-q/2} \exp\{-\frac{1}{2}(x'x)\}$. $K(x) = 2^{-q} \prod^{q} I(x_j)$, $I(x_j) = 1$, if $|x_j| < 1$; $I(x_j) = 0$ otherwise

A.2. $h_n = h \to 0$

A.3. $nh^q \to \infty$ as $n \to \infty$

- Weak (pointwise) Consistency: $P \lim \hat{f}(x) = f(x)$ at every c.p. of $f(x)$
- $E\hat{f}(x) \to f(x)$
- $V(\hat{f}(x)) \to \frac{1}{nh^q} f(x) \int K^2(w)dw \to 0$ as $n \to \infty$.

ASYMPTOTIC NORMALITY

$$\sqrt{nh^q}(\hat{f}(x) - f(x) - B) \sim N(0, f(x) \int K^2(\psi) d\psi)$$

where $B = \frac{h^2}{2}\mu_2 f^{(2)}(x)$, $\sqrt{nh}h^2 \longrightarrow 0$ as $n \longrightarrow \infty$; 95% C.I. for $f(x)$: $\hat{f}(x) \pm 1.96\sqrt{V(\hat{f}(x))}$.

"*CURSE OF DIMENSIONALITY*"

Reduction in Bias (Higher Order Kernels)

If $K \in \Re_r$ are kernels which could take negative/positive values such that first $(r-1)$ moments are zero.
Earlier case was $r = 2$.

$$Bias(\hat{f}(x)) = O(h^r),\ Bias(\hat{M}(x)) = O(h^r).$$

$MSE(\hat{f}(x)) = O(h^{2r}) + O(\frac{1}{nh^q})$ where $h \propto n^{-1/(2r+q)}$. Same for $\hat{m}(x)$.

$MSE = O(n^{-2r/(2r+q)}) \longrightarrow O(n^{-1})$, if r is large.

$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K(\frac{x_i - x}{h}) = \frac{1}{nh}\sum_{i=1}^{n} K(\psi_i)$ where $\psi_i = \frac{x_i - x}{h}$, $\left|\frac{dx}{d\psi_i}\right| = h$.

1.

$$\begin{aligned}\int_x \hat{f}(x)dx &= \frac{1}{nh}\sum_{i=1}^{n}\int_x K(\frac{x_i - x}{h})dx \\ &= \frac{1}{nh}\sum_{i=1}^{n}\int_{\psi_i} K(\psi_i)hd\psi_i \\ &= \frac{1}{n}\sum_{i=1}^{n} 1 \\ &= 1.\end{aligned}$$

2. $\hat{f}(x) = \frac{1}{n}\sum_{i=1}^{n} z_i$, $z_i = \frac{1}{h}K(\frac{x_i - x}{h})$.

$$
\begin{aligned}
E\hat{f}(x) &= Ez_1 \\
&= \frac{1}{h}\int_{x_1} K(\frac{x_1 - x}{h})f(x_1)dx_1 \\
&= \int_{\psi} K(\psi)f(x + h\psi)d\psi \\
&= \int_{\psi} K(\psi)[f(x) + h\psi f^{(1)}(x) + \frac{h^2\psi^2}{2}f^{(2)}(x) + \cdots)d\psi \\
&\simeq f(x) + \frac{h^2}{2}\mu_2 f^{(2)}(x).
\end{aligned}
$$

$$
\begin{aligned}
V(\hat{f}(x)) &= \frac{1}{n^2}\sum_{i=1}^{n} V(z_i) \\
&= \frac{V(z_1)}{n} \\
&= \frac{1}{n}[Ez_1^2 - (Ez_1)^2] \\
&= \frac{1}{n}\int \frac{1}{h^2}K^2(\frac{x_1 - x}{h}f(x_1)dx_1) - \frac{1}{n}(Ez_1)^2 \\
&= \frac{1}{nh}\int_{\psi} K^2(\psi)f(x + h\psi)d\psi - \frac{1}{n}(Ez_1)^2 \\
&\simeq \frac{1}{nh}f(x)\int_{\psi} K(\psi)d\psi.
\end{aligned}
$$

Cumulative Distribution Function

$$
\begin{aligned}
\hat{F}(x) &= \int_{-\infty}^{x} \hat{f}(t)dt \\
&= \int_{-\infty}^{x} \frac{1}{nh}\sum_{i=1}^{n} K(\frac{t-x_i}{h})dt \\
&= \frac{1}{n}\sum_{i=1}^{n}\int_{-\infty}^{x} \frac{1}{h}K(\frac{t-x_i}{h})dt.
\end{aligned}
$$

Let $\frac{t-x_i}{h} = \psi_i$.

$$
\begin{aligned}
\hat{F}(x) &= \frac{1}{n}\sum_{i=1}^{n}\int_{-\infty}^{\frac{x-x_i}{h}} K(\psi_i)d\psi_i \\
&= \frac{1}{n}\sum_{i=1}^{n} G(\frac{x-x_i}{h})
\end{aligned}
$$

and $G(z) = \int_{-\infty}^{z} K(\psi)d\psi$.

Cumulative Distribution Function

$$
\begin{aligned}
E\hat{F}(x) &= EG(\frac{x-x_i}{h}) = \int_{-\infty}^{\infty} G(\frac{x-x_i}{h}) f(x_i) dx_i \\
&= h\int_{-\infty}^{\infty} G(z) f(x-hz) dz = -\int_{-\infty}^{\infty} G(z) dF(x-hz) \\
&= -[G(z)F(x-hz)]_{-\infty}^{\infty} + \int_{-\infty}^{\infty} K(z) F(x-hz) dz \\
&= \int_{-\infty}^{\infty} K(z) F(x-hz) dz \\
&= \int_{-\infty}^{\infty} K(z)[F(x) - F^{(1)} hz + \frac{1}{2} h^2 z^2 F^{(2)}(x) + \cdots] dz \\
&= F(x) + \frac{1}{2}\mu_2 h^2 F^{(2)}(x) + o(h^2) \\
&\simeq F(x) + \frac{1}{2}\mu_2 h^2 F^{(2)}(x)
\end{aligned}
$$

where $\frac{x-x_i}{h} = z$. Thus, $BIAS(\hat{F}(x)) = \frac{1}{2}\mu_2 h^2 F^{(2)}(x)$.

SIMILARITY

Let $z = \frac{x-x_i}{h}$,

$$\begin{aligned}
EG^2(\frac{x-x_i}{h}) &= \int_{-\infty}^{\infty} G^2(\frac{x-x_i}{h}) f(x_i) dx_i \\
&= h \int_{-\infty}^{\infty} G^2(z) f(x-hz) dz \\
&= -\int_{-\infty}^{\infty} G^2(z) dF(x-hz) \\
&= 2\int_{-\infty}^{\infty} G(z)K(z)F(x-hz)dz \\
&= 2\int_{-\infty}^{\infty} G(z)K(z)[F(x) - hzF^{(1)}(x)]dz + O(h^2) \\
&= F(x) - \lambda hf(x) + O(h^2),
\end{aligned}$$

$\lambda = 2\int zG(z)K(z)dz$ and
$2\int_{-\infty}^{\infty} G(z)K(z)dz = \int_{-\infty}^{\infty} dG^2(z) = G^2(\infty) - G^2(-\infty) = 1$

$$
\begin{aligned}
V(\hat{F}(x)) &= \frac{1}{n} V[G(\frac{x-x_i}{h})] \\
&= \frac{1}{n}[EG^2(\frac{x-x_i}{h}) - (EG(\frac{x-x_i}{h}))^2] \\
&= \frac{1}{n} F(x)(1-F(x)) - \frac{\lambda f(x) h}{n} + o(\frac{h}{n}).
\end{aligned}
$$

Hence

$$
MSE(\hat{F}(x)) \simeq \frac{1}{n} F(x)(1-F(x)) + h^4 (\frac{\mu_2}{2})^2 (F^{(2)}(x))^2 - \lambda f(x) \frac{h}{n}
$$

$$
IMSE(\hat{F}(x)) = \int_x MSE(\hat{F}(x))^2 dx \simeq \frac{\lambda_1}{n} - \frac{\lambda_2 h}{n} + \lambda_3 h^4
$$

where $h_0 = \lambda_0 n^{-1/3}$, $\lambda_0 = (\lambda_2/\lambda_3)^{1/3}$ and
$\sqrt{n}(\hat{F}(x) - F(x)) \sim N(0, F(x)(1-F(x)))$.
Stochastic Dominance: Linton, Whang, Maasoumi (2005, RES).

TESTING

$H_0 : f(x) = g(x)$, $H_1 : f(x) \neq g(x)$

TEST STATISTIC

$$I = \int_x (\hat{f}(x) - \hat{g}(x))^2 dx \sim N(0, V) \text{ under } H_0$$

- $H_0 : f_1(x) = f_2(x)$; Panel Data
- $H_0 : f(x) = f(-x)$ Symmetry
- $H_0 : f(y, x) = f(x)f(y)$, $f(z) = g(z)$
- $H_0 : f(x) = f(x, \theta)$.

Fan and Ullah (1998, JNS), Su and White (2008, ET; 2007, JE).

WHAT IS THE "TRUE" MODEL FOR THIS DATA?

THIS IS THE SUBJECT OF NONPARAMETRIC ECONOMETRICS (DATA BASED MODELING).

$\alpha + X\beta$: PARAMETRIC MODEL.

(1) $\{Y, X\}, X \in R'$ Data $\{y_i, x_i\}, i = 1, 2, ..., n$

$$m(x) = E(Y|X = x) = \int_y yf(y|x)dy = \int_y y\frac{f(y,x)}{f(x)}dy$$

$$\begin{aligned} \hat{m}(x) &= \int_y y\frac{\hat{f}(y,x)}{\hat{f}(x)}dy \\ &= \int_y y\frac{\frac{1}{nh^2}\sum_{i=1}^n K\left(\frac{y_i - y}{h}\right)K\left(\frac{x_i - x}{h}\right)}{\frac{1}{nh}\sum_{i=1}^n K\left(\frac{x_i - x}{h}\right)}dy \\ &= \frac{\sum_{i=1}^n y_i K\left(\frac{x_i - x}{h}\right)}{\sum_{i=1}^n K\left(\frac{x_i - x}{h}\right)} \end{aligned}$$

NADARAYA/WATSON (1964, SANKHYA)

NP ESTIMATOR (LOCAL FIT)

$m(x) = E(Y|X = x)$; $Y = m(x) + u$

$$\begin{aligned} y_i &= m(x_i) + u \\ &= m(x) + (x_i - x)m^{(1)}(x) + \frac{(x_i - x)^2}{2} m^{(2)}(x) + \cdots + u_i \\ &\simeq m(x) + u_i^*; \end{aligned}$$

and $\frac{x_i - x}{h} = \psi_i$; $x_i - x = h\psi_i = O(h)$.

N-W (LCLS):

$$\begin{aligned} \hat{m}(x) &= \min_{m(x)} \sum_{i=1}^{n} (y_i - m(x_i))^2 K(\frac{x_i - x}{h}) \\ &= \frac{\sum y_i K(\frac{x_i - x}{h})}{\sum K(\frac{x_i - x}{h})} \end{aligned}$$

LLLS:

$$\begin{aligned} y_i &\simeq m(x) + (x_i - x)m^{(1)}(x) + u_i^* \\ &\simeq z_i(x)\delta(x) + u_i^*; \end{aligned}$$

$z_i(x) = [1 \;\; x_i - x]$, $\delta(x) = [m(x) \;\; m^{(1)}(x)]'$

$$\hat{\delta}(x) = \min_{\delta(x)} \sum_{i=1}^{n} [y_i - z_i(x)\delta(x)]^2 K(\frac{x_i - x}{h})$$

$$\hat{\delta}(x) = \begin{bmatrix} \hat{m}(x) \\ \hat{\beta}(x) \end{bmatrix} = (Z'K(x)Z)^{-1}Z'K(x)y$$

where $Z_i = Z_i(x)$, $K(x) = Diag(K(\frac{x_1 - x}{h}), ..., K(\frac{x_n - x}{h}))$.

$\hat{m}(x) = [1 \;\; 0]\hat{\delta}(x)$; $\hat{\beta}(x) = [0 \;\; 1]\hat{\delta}(x)$ and

$\hat{\beta}(x)$: VARYING COEFFICIENT ESTIMATOR.

VARIANCE: Nonparametric Estimation

$$\begin{aligned} V(y|x) &= E[(y-m(x))^2|x] \\ &= \frac{\sum(y_i-m(x_i))^2 K(\frac{x_i-x}{h})}{\sum K(\frac{x_i-x}{h})} \end{aligned}$$

with $m(x) = E(y|x)$.

$y = m(x) + u$

$$\begin{aligned} V(y|x) &= V(u|x) \\ &= \frac{\sum \hat{u}_i^2 K(\frac{x_i-x}{h})}{\sum K(\frac{x_i-x}{h})} \end{aligned}$$

is the Conditional Variance [Heteroskedasticity]

MISSPECIFICATION TEST

$H_0 : Y = \alpha + X\beta + u$ or $y = m(x, \theta) + u$
$H_1 : y = m(x) + u$

$\Rightarrow H_0 : E(u|x) = 0;\ H_1 : E(u|x) \neq 0$

$$
\begin{aligned}
H_0 : E(u|x) = 0 &\Rightarrow E[uE(u|x)] = 0 \\
&\Rightarrow E[um^*(x)] = 0 \\
&\Rightarrow E[um^*(x)f(x)] = 0
\end{aligned}
$$

TEST STATISTIC

$$
\begin{aligned}
I &= \frac{1}{n}\sum_{i=1}^{n} u_i m^*(x_i) f(x_i) \\
&= \frac{1}{n}\sum_{i=1}^{n} \hat{u}_i \hat{m}^*(x_i) \hat{f}(x_i) \\
&= \frac{1}{n(n-1)h}\sum_{i=1}^{n}\sum_{\substack{j\neq i\\ j=1}}^{n} \hat{u}_i \hat{u}_j K(\frac{x_j - x_i}{h}) \sim N(0, V^*)
\end{aligned}
$$

Li-Wang (J. Econometrics, 1998).

LLLS ESTIMATION OF VARYING (FUNCTIONAL) COEFFICIENTS

LLLS: $\sum(y_i - x_i\beta(x))^2 K(\frac{x_i - x}{h})$
$\hat{\beta}(x) = (X'K(x)X)^{-1}X'K(x)y$

$y_i = x_i\beta(z_i) + u_i = m(x_{i,}z_i) + u_i$

$\sum(y_i - x_i\beta(z))^2 K(\frac{z_i - z}{h})$

$\hat{\beta}(z) = (X'K(z)X)^{-1}X'K(z)y$

Examples: Parametric Models

THRESHOLD AR (TAR): TONG (90)

$$y_i = y_{i-1}\beta(y_{i-d}) + u_i$$

$$\begin{aligned}\beta(y_{i-d}) &= \beta_1 \text{ if } |y_{i-d}| \geq c \\ &= \beta_2 \text{ if } |y_{i-d}| < c\end{aligned}$$

$$\beta(y_{i-d}) = \beta_1 I(|y_{i-d}| \geq c) + \beta_2 I(|y_{i-d}| < c)$$

CROSS-VALIDATION

$$y_i = \hat{m}(x_i) + \hat{u}_i$$

$$Min. \sum \hat{u}_i^2 \text{ with respect to } h$$

Also CV for both h and p (local polynomial degree)
Hall and Racine (2015, JE).

$$\hat{m}(x_i) = \hat{m}(x) + (x_i - x)\hat{m}^{(1)}(x) + (x_i - x)^2\hat{m}^{(2)}(x) + \cdots + (x_i - x)^p\hat{m}^{(p)}(x$$

LLLS PROPERTIES

$$\sqrt{nh^q}(\hat{m}(x) - m(x) - B(x)) \sim N(0, \frac{\sigma^2(x)}{f(x)} \int_\psi K^2(\psi) d\psi)$$

where $B(x) = \frac{1}{2}\mu_2 h^2 m^{(2)}(x)$.

$$\sqrt{nh^{q+2}}(\hat{\beta}(x) - \beta(x) - B_1(x)) \sim N(0, \frac{\sigma^2(x)}{f(x)} \int_\psi (K^{(1)}(\psi))^2 d\psi)$$

where $B_1(x) = \frac{1}{2}\mu_2 h^2 m^{(3)}(x)$.

QUANTILE ESTIMATION

$y_i = x_i\beta + u_i$

$$\begin{aligned}
\sum |u_i| &= \sum |y_i - x_i\beta| = \sum_{u_i \geq 0} u_i - \sum_{u_i \leq 0} u_i \\
&= \sum_{u_i \geq 0} |u_i| + \sum_{u_i \leq 0} |u_i| \\
&= \sum u_i I(u_i \geq 0) - \sum u_i I(u_i < 0) \\
&= \sum u_i (1 - I(u_i < 0)) - \sum u_i I(u_i < 0) \\
&= \sum u_i (1 - 2I(u_i < 0)) = \sum u_i (0.5 - I(u_i < 0)) \\
&= \sum_{u_i \geq 0} |u_i| \frac{1}{2} + \sum_{u_i \leq 0} |u_i| \frac{1}{2} \\
&= \theta \sum_{u_i \geq 0} |u_i| + (1 - \theta) \sum_{u_i \leq 0} |u_i|
\end{aligned}$$

$\rightarrow \theta$·quantile

QUANTILE ESTIMATION

- $\theta-$th quantile of $F_Y = P(Y \leq y)$, $q_Y(\theta)$ is the solution of

$$P(Y \leq y) = F_Y(q) = \theta = \int_{-\infty}^{q} f(y)dy$$

$$q_Y(\theta) = F_Y^{-1}(\theta)$$

- $\theta E_{y>q}\left|y-q\right| + (1-\theta)E_{y<q}\left|y-q\right| =$
$\theta \int_{y>q} \left|y-q\right| dF_Y(y) + (1-\theta)\int_{y<q}\left|y-q\right| dF_Y(y)$

$$\begin{aligned}
\text{DWR to } q &= \theta \int_{y>q} (y-q)dF_Y(y) + (1-\theta)\int_{y<q}(y-q)dF_Y(y) \\
0 &= -\theta \int_{y>q} dF_Y(y) + (1-\theta)\int_{y<q} dF_Y(y) \\
&= -\theta[1-F_Y(q)] + (1-\theta)F_Y(q) \\
&= -\theta + F_Y(q)
\end{aligned}$$

QUANTILE REGRESSION

(1) $y = E(y|x) + u$; $y = \alpha + X\beta + u$, $\min_{\alpha,\beta} \sum \hat{u}_i^2$: LS

(2) $y = q_\theta(y|x) + u = \alpha_\theta + x\beta_\theta u + u$
$\min_{\alpha,\beta}[\theta \sum_{u_i \geq 0} |u_i| + (1-\theta) \sum_{u_i<0} |u_i|]$
$\theta = 0.5$ is Median Regression Estimator

$\min \frac{1}{2}[\sum_{u_i \geq 0} |u_i| + \sum_{u_i<0} |u_i|] = \min_{\alpha,\beta} \sum |u_i|$

(3) NP: $\min_{\alpha,\beta}[\theta \sum_{u_i \geq 0} |u_i| \, K(\frac{x_i - x}{h}) + (1-\theta) \sum_{u_i<0} |u_i| \, K(\frac{x_i - x}{h})]$
Su and Ullah (2008, SS)

QUANTILE REGRESSION

$$
\begin{aligned}
S(\beta,\theta) &= \frac{1}{T}[\theta \sum_{u_t \geq 0} |u_t| + (1-\theta) \sum_{u_t<0} |u_t|] \\
&= \frac{1}{T}[\theta \sum_{u_t \geq 0} u_t - (1-\theta) \sum_{u_t<0} u_t] \\
&= \frac{1}{T}\sum_t [\theta - I(u_t < 0)] u_t \\
& \frac{1}{T}\sum_t \rho_\theta(u_t).
\end{aligned}
$$

ADDITIVE REGRESSIONS

$$\begin{aligned} y_i &= m(x_i) + u_i \\ &= m(x_{i1,\dots,}x_{iq}) + u_i \\ &= m_1(x_{i1}) + m_2(x_{i2}) + \cdots + m_q(x_{iq}) + u_i \end{aligned}$$

where $Em_s(x_{is}) = 0$ for identification.
Then

$$m_1(x_{i1}) = \int m(x_{i1}, x_{i\underline{2}})\hat{f}(x_{i\underline{2}})dx_{i\underline{2}}$$

$$\hat{m}_1(x_{i1}) = \int m(x_{i1}, x_{i\underline{2}})d\hat{F}(x_{i\underline{2}})$$

$$\hat{m}_1(x_{i1}) = \frac{1}{n}\sum_{j=1}^{n} \hat{m}(x_{i1}, x_{j\underline{2}}).$$

1. $y = x\beta + u$, $V(u|x) = \begin{bmatrix} \hat{\sigma}^2(x_1) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \hat{\sigma}^2(x_n) \end{bmatrix} = \hat{\Sigma}$ and
$\hat{\beta}(X'\hat{\Sigma}(x)X)^{-1}X'\hat{\Sigma}^{-1}(x)y$. Also, $\mathrm{H}_0 : E(u^2|z) = m(z) = \sigma^2$,
$\hat{u}^2 = m(z) + V$.Su and Ullah (2013, ET)
2. $y_i = x_i\beta + u_i$, $f(u_i)$ unknown
$L(\beta) = \prod_{i=1}^{n} f(u_i)$
$\log L(\beta) = \sum \log f(u_i) = \sum \log \hat{f}(u_i)$, $\hat{f}(u_i) = \frac{1}{nh}\sum_{j=1}^{n} K(\frac{u_j - u_i}{h})$,
$u_j = y_j - x_j\beta$
$\max_\beta \log L(\beta)$. Engle and Gonzalez-Rivera (1991, JBES).
3. $y_t = x_t\beta + u_t$, $u_t = m(u_{t-1}) + \epsilon_t$
$y_t = x_t\beta + m(u_{t-1}) + \epsilon_t$
$y_t - \hat{m}(\hat{u}_{t-1}) = x_t\beta + \epsilon_t$
Su and Ullah (2006, ET): $y_t = m(x_t) + u_t$, $u_t = m(u_{t-1}) + \epsilon_t$
Test: see Hong (1996, Econometrica), Lee and Hong (2001, ET).

Semiparametric Models

$$\begin{aligned}
y_i &= x_{i1}\beta + m(x_{i2}) + u_i \\
E(y_i|x_{i2}) &= E(x_{i1}|x_{i2})\beta + m(x_{i2}) \\
y_i - E(y_i|x_{i2}) &= (x_{i1} - E(x_{i1}|x_{i2}))\beta + u_i \\
y_i^* &= x_{i1}^*\beta + u_i \\
y_i^{**} &= y_i - x_{i1}\hat{\beta} = m(x_{i2}) + u_i
\end{aligned}$$

Robinson (1988, Econometrica)
$\sqrt{n}$ convergence of $\hat{\beta}$.

(NP) Nonparametric Dummy Dependent Regression

$y^* = x\beta - u, u \sim (0, 1)$

$$\begin{aligned} y &= 1 \text{ if } y^* \geq 0 \text{ or } u \leq x\beta \\ y &= 0 \text{ if } y^* < 0 \text{ or } u > x\beta \end{aligned}$$

$$\begin{aligned} y &= E(y|x\beta) + v \\ &= F(x\beta) + v \\ &= \frac{e^{x\beta}}{1 + e^{x\beta}} + v \rightarrow \text{logit} \\ &= \int_{-\infty}^{x\beta} f(t)dt + v \rightarrow \text{probit} \\ &= \frac{\sum y_i K(\frac{x_i\beta - x\beta}{h})}{\sum K(\frac{x_i\beta - x\beta}{h})} + v \rightarrow \text{NP} \end{aligned}$$

Estimate β by NLS(Ichimura 1993, JE) or ML(Klein-Spady)

Nonparametric Dummy Dependent Regression

Klein-Spady (1993, Econometrica):
$\log L = \sum[F(x_i\beta)y_i + (1 - y_i)\log(1 - F(x_i\beta))]$

Manski's Score: $\sum[I(x_i\beta > \epsilon_i)y_i + (1 - y_i)(1 - I(x_i\beta \leq \epsilon_i))]$

Horowitz: Replace $I(x_i\beta > \epsilon_i)$ by $K(\frac{x_i\beta}{h})$

Parametric Regression $\mathbf{R}^2$

$$
\begin{aligned}
Y_t &= X_t\theta + U_t \quad t = 1, \ldots, n \\
\min_{\theta} \sum \left(Y_t - X_t\theta\right),\ \hat{\theta} &= (X'X)^{-1}X'Y \\
Y_t &= X_t\hat{\theta} + \hat{U}_t \\
&= \hat{Y}_t + \hat{U}_t \\
Y_t - \bar{Y} &= \hat{Y}_t - \bar{Y} + \hat{U}_t \\
\sum \left(Y_t - \bar{Y}\right)^2 &= \sum \left(\hat{Y}_t - \bar{Y}\right)^2 + \sum \hat{U}_t^2
\end{aligned}
$$

ANOVA Decomposition

$$
\begin{aligned}
TSS &= ESS + RSS \\
R^2 &= \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}
\end{aligned}
$$

Normal Equations

$$
\sum \hat{U}_t = 0,\ \sum \hat{U}_t X_t = 0
$$

Nonparametric Regression Estimation

$$\begin{aligned} Y_t &= m(X_t) + U_t \\ &\simeq m(x) + (X_t - x)' \beta(x) + U_t \\ &= X_{tx}' \delta(x) + U_t \end{aligned}$$

where $X_{tx} = \begin{bmatrix} 1 & (X_t - x)' \end{bmatrix}'$, $\delta(x) = \begin{bmatrix} m(x) & \beta'(x) \end{bmatrix}'$, X_t is $p \times 1$.

Nonparametric Regression R^2 (Local)

$$
\begin{aligned}
Y_t &= X_{tx}'\hat{\delta}(x) + \hat{U}_t \\
&= \hat{Y}_{tx} + \hat{U}_t \\
Y_t - \bar{Y} &= \hat{Y}_{tx} - \bar{Y} + \hat{U}_t \\
\left(Y_t - \bar{Y}\right)^2 &= \left(\hat{Y}_{tx} - \bar{Y}\right)^2 + \hat{U}_t^2 + 2\left(\hat{Y}_{tx} - \bar{Y}\right)\hat{U}_t \\
\sum \left(Y_t - \bar{Y}\right)^2 K_h\left(X_t - x\right) &= \sum \left(\hat{Y}_{tx} - \bar{Y}\right)^2 K_h\left(X_t - x\right) + \sum \hat{U}_t^2 K_h\left(X_t - x\right)
\end{aligned}
$$

Local ANOVA

$$
TSS(x) = ESS(x) + RSS(x)
$$

Normal Equations

$$
\begin{aligned}
\sum \hat{U}_t K_h\left(X_t - x\right) &= 0, \ \sum \hat{U}_t\left(X_t - x\right) K_h\left(X_t - x\right) = 0 \\
R^2(x) &= \frac{ESS(x)}{TSS(x)}
\end{aligned}
$$

Nonparametric Regression R^2 (Global)

$$
\begin{aligned}
TSS(x) &= ESS(x) + RSS(x) \\
\sum \left(Y_t - \bar{Y}\right)^2 K_h(X_t - x) &= \sum \left(\hat{Y}_{tx} - \bar{Y}\right)^2 K_h(X_t - x) + \sum \hat{U}_t^2 K_h(X_t - x) \\
[\hat{Y}_{tx} &= X_{tx}'\hat{\delta}(x) = X_{tx}'\left(X_x' W_x X_x\right)^{-1} X_x' W_x Y] \\
\int TSS(x)\,dx &= \int ESS(x)\,dx + \int RSS(x)\,dx
\end{aligned}
$$

Global ANOVA

$$
\begin{aligned}
TSS &= ESS + RSS \\
R^2 &= \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}
\end{aligned}
$$

$ESS = Y'MH^*MY$, $H^* = \int H_x dx$, $H_x = W_x X_x \left(X_x' W_x X_x\right)^{-1} X_x' W_x$
$M = I_n - L$, where L is Matrix of $1/n$.

Goodness of Fit / Model Selection Measures

$$
\begin{aligned}
R_q^2 &= \frac{ESS_q}{TSS} = 1 - \frac{RSS_q}{TSS} \\
&= 1 - \frac{Y'\left(I_n - H_q^*\right)Y}{Y'MY} \\
R_{Adj}^2 &= 1 - \frac{RSS_q / \left(n - \mathrm{tr} H_q^*\right)}{TSS / (n-1)} \\
AIC &= \log\left(RSS_q\right) + 2\mathrm{tr}\left(H_q^*\right)/n \\
BIC &= \log\left(RSS_q\right) + (\log n)\,\mathrm{tr}\left(H_q^*\right)/n
\end{aligned}
$$

Global R^2 (other definition)

$$
\begin{aligned}
y &= m(x) + u \\
V(y) &= V(m(x)) + V(u) \\
V(y) &= V[E(y|x)] + E[V(y|x)] \\
R^2 &= V[m(x)] \ / \ V(y) \\
&= 1 - \frac{E(y - m(x))^2}{V(y)} \\
\hat{R}^2 &= 1 - \frac{\frac{1}{n}\sum (y_i - \hat{m}(x_i))^2}{\frac{1}{n}\sum (y_i - \bar{y})^2} \\
\tilde{R} &= R^2 I\left[\frac{1}{n}\sum (y_i - \bar{y})^2 \geq \frac{1}{n}\sum (y_i - \hat{m}(x_i))^2\right]
\end{aligned}
$$

Uses of R^2: Goodness of Fit

: Testing based on R^2 (LM type Tests)

Su, Ullah (2013, ET), Yao and Ullah (2013, JSPI)

Discrete Data / Mixed Data

$X : f(x)$?
X : Continuous

$$\hat{f}(x) = \frac{1}{nh}\sum_{1}^{n} K(\frac{x_i - x}{h}) = \frac{1}{nh}\sum_{1}^{n} K(x_i, x, h)$$

Discrete

$$\begin{aligned} \hat{f}(x) &= \frac{1}{n}\sum_{1}^{n} L(x_i, x, \lambda) \\ L(x_i, x, \lambda) &= 1 - \lambda \qquad x_i = x \\ &= \frac{\lambda}{c-1} \qquad x_i \neq x \end{aligned}$$

for $x_i \in \{0, 1, \ldots, c-1\}$, $0 \le \lambda \le \frac{c-1}{c}$.For $\lambda = \frac{c-1}{c}$, $L(\cdot) = 1/c$.
For Regression: x_i is discrete

$$\begin{aligned} y_i &= m(x_i) + u_i \\ \hat{m}(x) &= \frac{\sum y_i L(x_i, x, \lambda)}{\sum L(x_i, x, \lambda)} \rightarrow \text{N-W} \end{aligned}$$

Discrete Data / Mixed Data

$L(x_i, x, \lambda) = 1$ if $x_i = x$ and λ otherwise.
Mixed Data $x_i = \left[\begin{array}{cc} x_{i1}, & x_{i2} \end{array}\right]$, x_{i1}: C and x_{i2}: D

$$K(\frac{x_i - x}{h}) = K(\frac{x_{i1} - x_1}{h}) L(x_{i2}, x_2, \lambda)$$