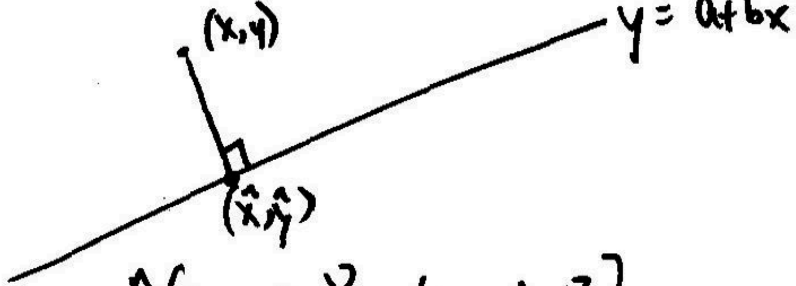
Stat 523
5-14-26

Orthogonal least squares

①



Minimize $\sum_{i=1}^{n}\left[(x_i-\hat{x}_i)^2 + (y_i-\hat{y}_i)^2\right]$

②

Given a point (x,y) not on the line $y=a+bx$,

Find the point $(\hat{x},\hat{y})$ on the line, closest to (x,y)

Let D = squared distance from (x,y) to $(\hat{x},\hat{y})$

$$= (x-\hat{x})^2 + (y-\hat{y})^2$$

$$= (x-\hat{x})^2 + (y-(a+b\hat{x}))^2$$

$$\frac{dD}{d\hat{x}} = \cancel{2}(x-\hat{x})(-1) + \cancel{2}(y-a-b\hat{x})(-b) \overset{\text{set}}{=} 0$$

$$x-\hat{x}+by-ab-b^2\hat{x} = 0$$

$$\hat{x} = \frac{x+by-ab}{(1+b^2)}$$

So $\hat{x} = \frac{x+by-ab}{(1+b^2)}$ and

$$\hat{y} = a + b\left(\frac{x+by-ab}{1+b^2}\right)$$

3

Our target function that we need to minimize is

$$\sum_{i=1}^{n}\left[\left(x_i - \frac{x_i+by_i-ab}{1+b^2}\right)^2 + \left(y_i - a - b\left(\frac{x_i+by_i-ab}{1+b^2}\right)\right)^2\right]$$

$$= \frac{1}{1+b^2}\sum_{i=1}^{n}(y_i - a - bx_i)^2$$

4

Fix b. Then $\hat{a} = \bar{y} - b\bar{x}$

Substitute:

$$\frac{1}{1+b^2}\sum_{i=1}^{n}\left(y_i - (\bar{y} - b\bar{x}) - bx_i\right)^2$$

$$= \frac{1}{1+b^2}\sum_{i=1}^{n}\left[(y_i-\bar{y}) - b(x_i-\bar{x})\right]^2$$

$$= \frac{1}{1+b^2}\left[\sum(y_i-\bar{y})^2 + b^2\sum(x_i-\bar{x})^2 - 2b\sum(x_i-\bar{x})(y_i-\bar{y})\right]$$

$$= \frac{1}{1+b^2}\left[S_{yy} + b^2 S_{xx} - 2b S_{xy}\right]$$

(5)

Take the derivative w.r.t. b:

$$\frac{(1+b^2)[2bS_{xx}-2S_{xy}]-[S_{yy}+b^2S_{xx}-2bS_{xy}](2b)}{(1+b^2)^2} \overset{set}{=} 0$$

$$2bS_{xx}-2S_{xy}+\cancel{2b^3S_{xx}}-\underline{2b^2S_{xy}}$$
$$-2bS_{yy}-\cancel{2b^3S_{xx}}+\underline{4b^2S_{xy}} = 0$$

$$b^2S_{xy}+b(S_{xx}-S_{yy})-S_{xy}=0$$

(6)

$$\hat{b} = \frac{S_{yy}-S_{xx} \pm \sqrt{(S_{xx}-S_{yy})^2+4S_{xy}^2}}{2S_{xy}}$$

and $\hat{a} = \bar{y}-\hat{b}\,\bar{x}$

Next up: measurement error model

(7)

The EIV regression model

"errors in variables" also called

"measurement errors" model

$i=1,\dots,n \qquad X_i \sim N(\mu_i, \sigma_x^2)$

$Y_i \sim N(\beta_0+\beta_1\mu_i, \sigma_y^2)$

Assume $X_1,\dots,X_n, Y_1,\dots,Y_n$ are independent

$$L = \left(\frac{1}{\sigma_x\sqrt{2\pi}}\right)^n e^{-\frac{1}{2\sigma_x^2}\sum(x_i-\mu_i)^2}\left(\frac{1}{\sigma_y\sqrt{2\pi}}\right)^n e^{-\frac{1}{2\sigma_y^2}\sum(y_i-\beta_0-\beta_1\mu_i)^2}$$

Let $\quad \lambda = \frac{\sigma_x^2}{\sigma_y^2} \qquad \sigma_y = \frac{\sigma_x}{\sqrt{\lambda}}$

(8)

$$L = (2\pi)^{-n}(\sigma_x^2)^{-\frac{n}{2}}\left(\frac{\sigma_x^2}{\lambda}\right)^{-\frac{n}{2}} \cdot$$

$$e^{-\frac{1}{2\sigma_x^2}\Big[\underbrace{\sum_{i=1}^n (x_i-\mu_i)^2 + \lambda\sum_{i=1}^n (y_i-\beta_0-\beta_1\mu_i)^2}_{\star}\Big]}$$

Maximizing L with respect to μ_i is equivalent

to minimizing $\star$ w.r.t. μ_i

$$\frac{\partial\star}{\partial\mu_i} = \cancel{2}(x_i-\mu_i)(\cancel{-1}) + \cancel{2}\lambda(y_i-\beta_0-\beta_1\mu_i)(\cancel{-}\beta_1) \overset{set}{=} 0$$

$$x_i + \lambda\beta_1 y_i - \lambda\beta_0\beta_1 = \mu_i + \lambda\beta_1^2\mu_i$$

$$\hat{\mu}_i = \frac{x_i + \lambda\beta_1(y_i - \beta_0)}{1 + \lambda\beta_1^2}$$

Substitute this into $\star$:

$$\star = \frac{\lambda}{1+\lambda\beta_1^2} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2$$

Let $\beta_0^* = \sqrt{\lambda}\,\beta_0$, $\beta_1^* = \sqrt{\lambda}\,\beta_1$, $y_i^* = \sqrt{\lambda}\,y_i$

then $$\star = \frac{\sum (y_i^* - \beta_0^* - \beta_1^* x_i)^2}{1 + {\beta_1^*}^2}$$

So β_0^* and β_1^* are the solutions to the orthogonal least squares problem (solved last time)

Then $\hat{\beta}_0 = \frac{1}{\sqrt{\lambda}}\beta_0^*$, $\hat{\beta}_1 = \frac{1}{\sqrt{\lambda}}\beta_1^*$

$$\ln L = -n\ln(2\pi) - \frac{n}{2}\ln\sigma_x^2 - \frac{n}{2}\ln\sigma_y^2 - \frac{1}{2\sigma_x^2}\sum(x_i - \hat{\mu}_i)^2 - \frac{1}{2\sigma_y^2}\sum(y_i - \hat{\beta}_0 - \hat{\beta}_1\hat{\mu}_i)^2$$

$$\frac{\partial \ln L}{\partial \sigma_x^2} = -\frac{n}{2}\frac{1}{\sigma_x^2} - \frac{1}{2}\frac{-1}{(\sigma_x^2)^2}\sum(x_i - \hat{\mu}_i)^2 \overset{\text{set}}{=} 0$$

$$\hat{\sigma}_x^2 = \frac{\sum (x_i - \hat{\mu}_i)^2}{n}$$

Similarly, $$\hat{\sigma}_y^2 = \frac{\sum (y_i - \hat{\beta}_0 - \hat{\beta}_1 \hat{\mu}_i)^2}{n}$$

and $$\hat{\lambda} = \frac{\hat{\sigma}_x^2}{\hat{\sigma}_y^2}$$

12.2 Show that the extrema of

$$f(b) = \frac{1}{1+b^2}\left[S_{yy} - 2bS_{xy} + b^2 S_{xx}\right]$$

are given by

$$b = \frac{-(S_{xx} - S_{yy}) \pm \sqrt{(S_{xx} - S_{yy})^2 + 4S_{xy}^2}}{2S_{xy}}.$$

Show that the "+" solution gives the minimum of $f(b)$.

12.4 Consider the MLE of the slope in the EIV model

$$\hat{\beta}(\lambda) = \frac{-(S_{xx} - \lambda S_{yy}) + \sqrt{(S_{xx} - \lambda S_{yy})^2 + 4\lambda S_{xy}^2}}{2\lambda S_{xy}},$$

where $\lambda = \sigma_\delta^2/\sigma_\epsilon^2$ is assumed known.

(a) Show that $\lim_{\lambda\to 0}\hat{\beta}(\lambda) = S_{xy}/S_{xx}$, the slope of the ordinary regression of y on x.

(b) Show that $\lim_{\lambda\to\infty}\hat{\beta}(\lambda) = S_{yy}/S_{xy}$, the reciprocal of the slope of the ordinary regression of x on y.

(c) Show that $\hat{\beta}(\lambda)$ is, in fact, monotone in λ and is increasing if $S_{xy} > 0$ and decreasing if $S_{xy} < 0$.

(d) Show that the orthogonal least squares line ($\lambda = 1$) is always between the lines given by the ordinary regressions of y on x and of x on y.

(e) The following data were collected in a study to examine the relationship between brain weight and body weight in a number of animal species.

Species	Body weight (kg) (x)	Brain weight (g) (y)
Arctic fox	3.385	44.50
Owl monkey	.480	15.50
Mountain beaver	1.350	8.10
Guinea pig	1.040	5.50
Chinchilla	.425	6.40
Ground squirrel	.101	4.00
Tree hyrax	2.000	12.30
Big brown bat	.023	.30

Calculate the MLE of the slope assuming the EIV model. Also, calculate the least squares slopes of the regressions of y on x and of x on y, and show how these quantities bound the MLE.