33

11.4 Influential diagnostics

34

some notes

35

Full model $\quad M: \; Y|X \sim (X\beta, \sigma^2 I_n)$

$$\text{rank}(X_{n\times k}) = k$$

Leave-one-out model

$$M_{(-t)}: \; Y_{(-t)}|X_{(-t)} \sim (X_{(-t)}\beta, \sigma^2 I_n)$$

$$\text{rank}(X_{(-t)}) = k$$

(ASSUMPTION: $m_{tt} > 0$)

R: influence.measures 36

11.4.1 DFBETAS

37

- comparison of LSE of β under models M and $M_{(-t)}$

M: $\hat{\beta} = (\hat{\beta}_0, \dots, \hat{\beta}_{k-1})^T = (X^TX)^{-1}X^TY$

$M_{(-t)}$: $\hat{\beta}_{(-t)} = (\hat{\beta}_{(-t),0}, \dots, \hat{\beta}_{(-t),k-1})^T = (X_{(-t)}^T X_{(-t)})^{-1} X_{(-t)}^T Y_{(-t)}$

Lemma 11.3:

$$\hat{\beta} - \hat{\beta}_{(-t)} = \frac{u_t}{m_{tt}} (X^TX)^{-1} x_t$$

$$\rightarrow \text{DFBETA}_{t,j} := \hat{\beta}_j - \hat{\beta}_{(-t),j} = \frac{u_t}{m_{tt}} \cdot v_t^T x_t$$

$$\boxed{V := (X^TX)^{-1} = \begin{pmatrix} v_0^T \\ \vdots \\ v_{k-1}^T \end{pmatrix} = \begin{pmatrix} v_{00} & \cdots & v_{0\,k-1} \\ \vdots & & \\ v_{k-1,0} & \cdots & v_{k-1,k-1} \end{pmatrix}}$$

remember: $S.E.(\hat{\beta}_j) = \sqrt{MSe\, v_{jj}}$, $S.E.(\hat{\beta}_{(-t),j}) = \sqrt{MSe_{,(-t)}\, \underline{v_{(-t),jj}}}$

from matrix $\rightarrow (X_{(-t)}^T X_{(-t)})^{-1}$

$\rightarrow \text{DFBETAS}_{t,j}$ = standardized $\text{DFBETA}_{t,j}$

38

$$\text{DFBETAS}_{t,j} := \frac{\hat{\beta}_j - \hat{\beta}_{(-t),j}}{\sqrt{MSe_{,(-t)}\, v_{jj}}} = \frac{u_t}{m_{tt}\sqrt{MSe_{,(-t)}\, v_{jj}}} v_t^T x_t$$

R software: t-th observation influential w.r.t. estimation of $\beta_j \iff |\text{DFBETAS}_{t,j}| > 1$.

R example: Cars2004

39

11.4.2 DFFITS

- comparison of LSE of $\mu_t := \mathbb{E}(Y_t \mid X_t = x_t) = x_t^T \beta$ under models M and $M_{(-t)}$

M: $\hat{y}_t = x_t^T \hat{\beta}$, $\hat{\beta} = (X^T X)^{-1} X^T Y$

$M_{(-t)}$: $\hat{y}_{[t]} = x_t^T \hat{\beta}_{(-t)}$, $\hat{\beta}_{(-t)} = (X_{(-t)}^T X_{(-t)})^{-1} X_{(-t)}^T Y_{(-t)}$

While using $\hat{\beta}_{(-t)} = \hat{\beta} - \frac{u_t}{m_{tt}} (X^T X)^{-1} x_t$ (Lemma 11.3):

$$\hat{y}_{[t]} = x_t^T \left\{ \hat{\beta} - \frac{u_t}{m_{tt}} (X^T X)^{-1} x_t \right\} = \underbrace{x_t^T \hat{\beta}}_{\hat{y}_t} - \frac{u_t}{m_{tt}} \underbrace{x_t^T (X^T X)^{-1} x_t}_{h_{tt}}$$

$$\overset{\#}{=} \hat{y}_t - u_t \frac{h_{tt}}{m_{tt}}$$

$\rightarrow$ $DFFIT_t = \hat{y}_t - \hat{y}_{[t]} = u_t \frac{h_{tt}}{m_{tt}}$

(mind difference with deleted residual $\hat{\mu}_t^{out} = Y_t - \hat{y}_{[t]}$)

remember: $\widehat{var}(\hat{y}_t \mid X) = MS_e \cdot h_{tt}$

$\rightarrow$ $DFFITS_t$ = standardized $DFFIT_t$

$$DFFITS_t = \frac{\hat{y}_t - \hat{y}_{[t]}}{\sqrt{MS_{e,(-t)}\, h_{tt}}} = \frac{h_{tt}}{m_{tt}} \cdot \frac{u_t}{\sqrt{MS_{e,(-t)}\, h_{tt}}}$$

$$= \sqrt{\frac{h_{tt}}{m_{tt}}} \cdot \frac{u_t}{\sqrt{MS_{e,(-t)}\, m_{tt}}} = \sqrt{\frac{h_{tt}}{m_{tt}}} \cdot T_t$$

studentized residual $\downarrow$

Rsoftware: t^{th} observation influential w.r.t. estimation of $\mu_t = \mathbb{E}(Y_t \mid X_t = x_t)$ $\Leftrightarrow$ $|DFFITS_t| > 3 \cdot \sqrt{\frac{k}{n-k}}$

Rexample: Cars 2004

40

41

42-43

11.4.3 Cook distance

- comparison of LSE of $\mu = E(Y|X)$ under models M and $M_{(-t)}$

M: $\hat{Y} = X\hat{\beta}$, $\hat{\beta} = (X^TX)^{-1}X^TY$

$M_{(-t)}$: $\hat{Y}_{(-t\cdot)} = X\hat{\beta}_{(-t)}$, $\hat{\beta}_{(-t)} = (X_{(-t)}^T X_{(-t)})^{-1} X_{(-t)}^T Y_{(-t)}$

Mind difference

$$\hat{Y}_{(-t\cdot)} = \begin{pmatrix} x_1^T\hat{\beta}_{(-t)} \\ \vdots \\ x_n^T\hat{\beta}_{(-t)} \end{pmatrix}, \quad \hat{Y}_{[\cdot]} = \begin{pmatrix} x_1^T\hat{\beta}_{(-1)} \\ \vdots \\ x_n^T\hat{\beta}_{(-n)} \end{pmatrix}$$

$\hat{Y}_{(-t)} = X_{(-t)} \cdot \hat{\beta}_{(-t)}$ = subvector of $\hat{Y}_{(-t\cdot)}$

Lemma 11.3: $\hat{\beta} - \hat{\beta}_{(-t)} = \frac{U_t}{m_{tt}}(X^TX)^{-1}x_t$

$$\Rightarrow \hat{Y} - \hat{Y}_{(-t\cdot)} = X(\hat{\beta} - \hat{\beta}_{(-t)}) = \frac{U_t}{m_{tt}} X(X^TX)^{-1}x_t$$

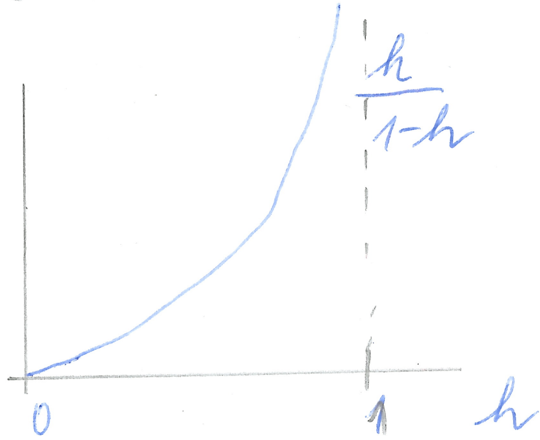
$$\Rightarrow \|\hat{Y} - \hat{Y}_{(-t\cdot)}\|^2 = \frac{U_t^2}{m_{tt}^2} \cdot \underbrace{x_t^T \underbrace{(X^TX)^{-1}X^TX(X^TX)^{-1}}_{(X^TX)^{-1}} x_t}_{h_{tt}} = \frac{U_t^2}{m_{tt}^2} h_{tt}$$

Cook distance = standardized (by $k \cdot MSe$) $\|\hat{Y} - \hat{Y}_{(-t\cdot)}\|^2$

Cook distance$_t$ $= D_t := \frac{1}{k \cdot MS_e} \cdot \| \hat{Y} - \hat{Y}_{(-t\cdot)} \|^2$ 45

$$= \frac{1}{k} \cdot \frac{h_{tt}}{m_{tt}} \underbrace{\frac{U_t^2}{MS_e \cdot m_{tt}}}_{(U_t^{std})^2} = \frac{1}{k} \cdot \frac{h_{tt}}{m_{tt}} (U_t^{std})^2$$

$0 < h_{tt} = 1 - m_{tt} < 1$



$$D_t = \frac{1}{k} \cdot \underbrace{\frac{h_{tt}}{m_{tt}}}_{\text{high for leverage points}} \cdot (U_t^{std})^2 \rightarrow \text{high for outliers}$$

Cook distance $\Longleftrightarrow$ conf. ellipsoid for β 46

$$D_t = \frac{1}{kMS_e} \| X\hat{\beta} - X\hat{\beta}_{(-t)} \|^2 = \frac{1}{kMS_e} (\hat{\beta}_{(-t)} - \hat{\beta})^T X^T X (\hat{\beta}_{(-t)} - \hat{\beta})$$

$C(\alpha) = (1-\alpha)100\%$ conf. ellipsoid for β based on model M

$$= \{ \beta : \underbrace{\frac{1}{kMS_e} (\beta - \hat{\beta})^T X^T X (\beta - \hat{\beta})}_{= D_t \text{ if } \beta = \hat{\beta}_{(-t)}} < F_{k, n-k}(1-\alpha) \}$$

$D_t \geq F_{k,n-k}(1-\alpha) \iff \hat{\beta}_{(-t)}$ does not lie 47

in a conf. ellipsoid for β

based on model M

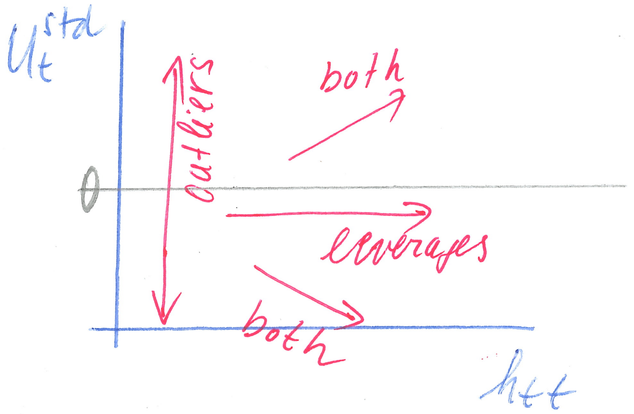
with coverage $(1-\alpha)\,100\%$

or LESS

R software: t^{th} observation influential for LSE of $\mu = \mathbb{E}(Y/X)$ $\Leftrightarrow$ $D_t > F_{k,n-k}(0{,}50)$ 47

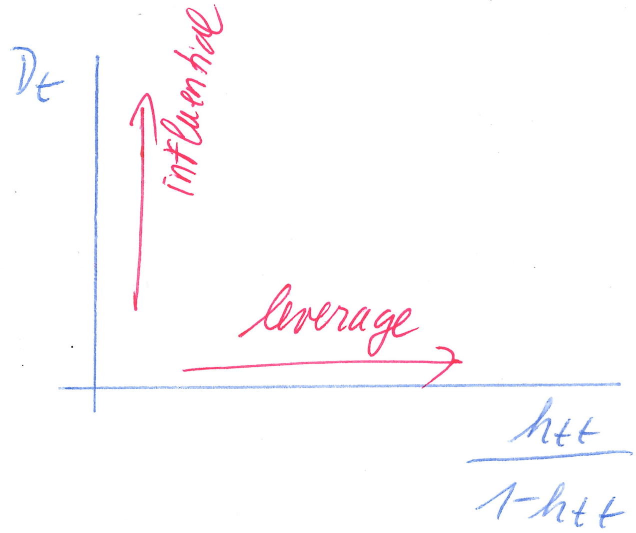
R illustration 48-49

50



red contours
(here not visible)
= area where
$D_t > F_{k,n-k}(1-\alpha)$
for $\alpha = 0{,}50, \ldots$

51



$$D_t = \underbrace{\frac{h_{tt}}{1-h_{tt}}}_{\text{leverage}} \cdot \underbrace{\left(\frac{U_t^{std}}{\sqrt{k}}\right)^2}_{\text{outlier}}$$

contours
$= \frac{U_t^{std}}{\sqrt{k}} = \text{const}$

= outlier part of D_t

11.4.4 COVRATIO

52

- comparison of precision in estimation of β using models M and $M_{(-t)}$

M: $\hat{\beta} = (X^T X)^{-1} X^T Y$

$$\hat{var}(\hat{\beta}|X) = MS_e \cdot (X^T X)^{-1}$$

$M_{(-t)}$: $\hat{\beta}_{(-t)} = (X_{(-t)}^T X_{(-t)})^{-1} X_{(-t)}^T Y_{(-t)}$

$$\hat{var}(\hat{\beta}_{(-t)}|X) = MS_{e,(-t)} (X_{(-t)}^T X_{(-t)})^{-1}$$

$$COVRATIO_t = \frac{\det\{\hat{var}(\hat{\beta}_{(-t)}|X)\}}{\det\{\hat{var}(\hat{\beta}|X)\}}$$ 53

= some calculations (see notes)

$$= \frac{1}{m_{tt}} \left\{ \frac{n-k-(U_t^{std})^2}{n-k-1} \right\}^k$$

R software: t^{th} observation influential

$$\Leftrightarrow |1 - COVRATIO_t| > 3 \cdot \frac{k}{n-k}$$

R illustration 54-55

11.4.5 Final remarks

- see slide