

INSTITUTE OF ECONOMIC STUDIES, FACULTY OF SOCIAL SCIENCES CHARLES UNIVERSITY IN PRAGUE (established 1348)



ROBUST STATISTICS ^{AND} ECONOMETRICS

INSTITUTE OF ECONOMIC STUDIES FACULTY OF SOCIAL SCIENCES CHARLES UNIVERSITY IN PRAGUE

JAN ÁMOS VÍŠEK

Week 12

Content of lecture



Disqualifying classical regression analysis by ignoring ...

- ... broken orthogonality condition
- ... collinearity

Content of lecture



- ... broken orthogonality condition
- ... collinearity
- 2 Significance of individual explanatory variable
 - In for LWS-estimation
 - ... simulations of *p*-values

... broken orthogonality condition ... collinearity

Content



- ... broken orthogonality condition
- ... collinearity

- ... for LWS-estimation
 - ... simulations of p-values

... broken orthogonality condition ... collinearity

Content



... broken orthogonality condition ... collinearity

Ignoring the orthogonality condition - breaking the consistency of $\hat{eta}^{(OLS,n)}$

We have discussed it in details in the tenth lecture.



... broken orthogonality condition ... collinearity

Ignoring the orthogonality condition - breaking the consistency of $\hat{eta}^{(OLS,n)}$

We have discussed it in details in the tenth lecture.

We have robustified the *Instrumental Variables* by means of *implicit weighting*, i. e. we have defined

Instrumental Weighted Variables (IWV)

as a solutions of the normal equations:



... broken orthogonality condition ... collinearity

Ignoring the orthogonality condition - breaking the consistency of $\hat{m{eta}}^{(OLS,n)}$

We have discussed it in details in the tenth lecture.

We have robustified the Instrumental Variables by means of implicit weighting, i.e. we have defined

Instrumental Weighted Variables (IWV)

as a solutions of the normal equations:

$$\sum_{i=1}^n w\left(F_n(|r_\ell(\beta)|)\right) Z_i(Y_i - X_i'\beta) = 0.$$

But there are still at least two questions:

Firstly, can we robustify *Instrumental Variables* by some other way, employing some other idea of robustification

of the classical econometric methods?

... broken orthogonality condition ... collinearity

Ignoring the orthogonality condition - breaking the consistency of $\hat{m{eta}}^{(OLS,n)}$

We have discussed it in details in the tenth lecture.

We have robustified the Instrumental Variables by means of implicit weighting, i.e. we have defined

Instrumental Weighted Variables (IWV)

as a solutions of the normal equations:

$$\sum_{i=1}^n w\left(F_n(|r_\ell(\beta)|)\right) Z_i(Y_i - X_i'\beta) = 0.$$

But there are still at least two questions:

Firstly, can we robustify *Instrumental Variables* by some other way, employing some other idea of robustification

of the classical econometric methods?

Secondly (and more importantly), how we learn that we should use IWV?

... broken orthogonality condition ... collinearity

Answering the first question

We can define *Instrumental M-estimators*, say $\hat{\beta}^{(IM,n)}$, as a solution of:

$$\sum_{i=1}^{n} Z_{i}\psi\left(\frac{\mathbf{Y}_{i}-\mathbf{X}_{i}'\beta}{\hat{\sigma}_{n}}\right)=\mathbf{0}$$

where Z_i 's have to be selected so that $EZ_1\psi(e_1\sigma_{e_1}^{-1}) = 0$ and $|EZ_{1j}(X_{1j} - EX_{1j})|$ as large as possible for all j = 1, 2, ..., p.

... broken orthogonality condition ... collinearity

Answering the second question

Let's make some preparation steps for the Hausman test Under hypothesis of orthogonality:

$$\mathcal{L}_{\mathrm{as}}\left(\sqrt{n}(\hat{eta}^{(IM,n)}-\hat{eta}^{(M,n)})
ight)=\mathcal{N}(0,C^{H})$$

with

$$C^{H} = \mathbf{E}\psi^{2}(e_{1}\sigma_{e_{1}}^{-1}) \left[Q_{Z}^{-1}\mathbf{E} \left\{ Z_{1}Z_{i}^{\prime} \right\} [Q_{Z}^{-1}]^{\prime} - Q^{-1}\mathbf{E}^{-1}\psi^{\prime}(e_{1}\sigma_{e_{1}}^{-1}) \right].$$

... broken orthogonality condition ... collinearity

Answering the second question

Let's make some preparation steps for the Hausman test Under hypothesis of orthogonality:

with

$$\mathcal{L}_{as}\left(\sqrt{n}(\hat{\beta}^{(IM,n)}-\hat{\beta}^{(M,n)})\right)=\mathcal{N}(0,C^{H})$$

$$C^{H} = \mathbf{E}\psi^{2}(e_{1}\sigma_{e_{1}}^{-1}) \left[Q_{Z}^{-1}\mathbf{E} \left\{ Z_{1}Z_{i}^{\prime} \right\} \left[Q_{Z}^{-1} \right]^{\prime} - Q^{-1}\mathbf{E}^{-1}\psi^{\prime}(e_{1}\sigma_{e_{1}}^{-1}) \right]$$

(Notice that the second term in parentheses is $\hat{\beta}^{(M,n)}$)

- this is a preparatory considerations for Hausman test.)

... broken orthogonality condition ... collinearity

Answering the second question

Let's continue in the preparation steps for the Hausman test: <u>Under alternative</u>:

 $\mathcal{L}_{\mathrm{as}}(\sqrt{n}\left(\hat{\beta}^{(\textit{IM},n)} - \hat{\beta}^{(\textit{M},n)} - Q^{-1}\boldsymbol{E}\left[X_{1}\psi(\boldsymbol{e}_{1}\sigma_{\boldsymbol{e}_{1}}^{-1})\right]\right)) = \mathcal{N}(0,C^{A})$

with

$$\begin{split} \boldsymbol{C}^{A} &= \boldsymbol{Q}^{-1} \left[\boldsymbol{E} \left\{ \psi^{2}(\boldsymbol{e}_{1} \boldsymbol{\sigma}_{\boldsymbol{e}_{1}}^{-1}) \boldsymbol{X}_{1} \boldsymbol{X}_{1}^{\prime} \right\} \right. \\ & - \boldsymbol{E} \left\{ \psi(\boldsymbol{e}_{1} \boldsymbol{\sigma}_{\boldsymbol{e}_{1}}^{-1}) \boldsymbol{X}_{1} \right\} \boldsymbol{E} \left\{ \psi(\boldsymbol{e}_{1} \boldsymbol{\sigma}_{\boldsymbol{e}_{1}}^{-1}) \boldsymbol{X}_{1}^{\prime} \right\} \right] \boldsymbol{Q}^{-1} \\ & - \boldsymbol{Q}^{-1} \boldsymbol{E} \left\{ \boldsymbol{X}_{1} \boldsymbol{Z}_{1}^{\prime} \psi^{2}(\boldsymbol{e}_{1} \boldsymbol{\sigma}_{\boldsymbol{e}_{1}}^{-1}) \right\} \left[\boldsymbol{Q}_{\boldsymbol{Z}}^{-1} \right]^{\prime} \\ & - \boldsymbol{Q}_{\boldsymbol{Z}}^{-1} \boldsymbol{E} \left\{ \boldsymbol{Z}_{1} \boldsymbol{X}_{1}^{\prime} \psi^{2}(\boldsymbol{e}_{1} \boldsymbol{\sigma}_{\boldsymbol{e}_{1}}^{-1}) \right\} \left[\boldsymbol{Q}^{-1} \right]^{\prime} \\ & + \boldsymbol{E} \left\{ \psi^{2}(\boldsymbol{e}_{1} \boldsymbol{\sigma}_{\boldsymbol{e}_{1}}^{-1}) \right\} \boldsymbol{Q}_{\boldsymbol{Z}}^{-1} \boldsymbol{E} \left\{ \boldsymbol{Z}_{1} \boldsymbol{Z}_{1}^{\prime} \right\} \left[\boldsymbol{Q}_{\boldsymbol{Z}}^{-1} \right]^{\prime}. \end{split}$$

... broken orthogonality condition ... collinearity

Answering the second question

Let's continue in the preparation steps for the Hausman test: <u>Under alternative</u>:

$$\mathcal{L}_{\mathrm{as}}(\sqrt{n}\left(\hat{\beta}^{(\textit{IM},n)} - \hat{\beta}^{(\textit{M},n)} - Q^{-1}\textit{I\!\!E}\left[X_{1}\psi(e_{1}\sigma_{e_{1}}^{-1})\right]\right)) = \mathcal{N}(0,C^{A})$$

with

$$C^{A} = Q^{-1} \left[E \left\{ \psi^{2}(e_{1}\sigma_{e_{1}}^{-1})X_{1}X_{1}' \right\} - E \left\{ \psi(e_{1}\sigma_{e_{1}}^{-1})X_{1}' \right\} E \left\{ \psi(e_{1}\sigma_{e_{1}}^{-1})X_{1}' \right\} \right] Q^{-1} - Q^{-1} E \left\{ X_{1}Z_{1}'\psi^{2}(e_{1}\sigma_{e_{1}}^{-1}) \right\} [Q_{Z}^{-1}]' - Q_{Z}^{-1} E \left\{ Z_{1}X_{1}'\psi^{2}(e_{1}\sigma_{e_{1}}^{-1}) \right\} [Q^{-1}]' + E \left\{ \psi^{2}(e_{1}\sigma_{e_{1}}^{-1}) \right\} Q_{Z}^{-1} E \left\{ Z_{1}Z_{1}' \right\} [Q_{Z}^{-1}]'.$$

Under hypothesis we can show that $C^A \rightarrow C^H$.

... broken orthogonality condition ... collinearity

How we learn that the orthogonality condition is broken?



... broken orthogonality condition ... collinearity

How we learn that the orthogonality condition is broken?

The classical econometrics offered the Hausman test.

... broken orthogonality condition ... collinearity

How we learn that the orthogonality condition is broken?

The classical econometrics offered the Hausman test. $q = \sqrt{n} \left(\hat{\beta}^{(IV,n)} - \hat{\beta}^{(LS,n)} \right), \qquad \hat{X} = Z \left(Z'Z \right)^{-1} Z'X$ Put $V = \left(\frac{1}{n}\hat{X}'\hat{X}\right)^{-1} - \left(\frac{1}{n}X'X\right)^{-1} \text{ and } \lambda = \mathbf{E}\left\{X_1'e_1\right\} V^{-1}\mathbf{E}\left\{X_1e_1\right\}.$

... broken orthogonality condition ... collinearity

How we learn that the orthogonality condition is broken?



... broken orthogonality condition ... collinearity

How we learn that the orthogonality condition is broken?



(Notice that under the hypothesis χ^2 is central,

while under the alternative we have some parameter λ of noncentrality.)

... broken orthogonality condition ... collinearity

Hausman specification test for *M*-estimators

Theorem: Under hypothesis & alternative:

Put $\tilde{q} = \sqrt{n} \left(\hat{\beta}^{(IM,n)} - \hat{\beta}^{(M,n)} \right)$ and $\tilde{W} = \tilde{q}' \left[C^A \right]^{-1} \tilde{q}$. Then $\mathcal{L}_{as}(\tilde{W}) = \chi^2(p, \tilde{\lambda})$

with

 $\tilde{\lambda} = \boldsymbol{E} \left[\boldsymbol{X}_{1}^{\prime} \boldsymbol{\psi}(\boldsymbol{e}_{1} \sigma_{\boldsymbol{e}_{1}}^{-1}) \right] \left[\boldsymbol{Q}^{-1} \right]^{\prime} \left[\boldsymbol{C}^{\boldsymbol{A}} \right]^{-1} \boldsymbol{Q}^{-1} \boldsymbol{E} \left[\boldsymbol{X}_{1} \boldsymbol{\psi}(\boldsymbol{e}_{1} \sigma_{\boldsymbol{e}_{1}}^{-1}) \right].$

For details see:

... broken orthogonality condition ... collinearity

Hausman specification test for *M*-estimators

Theorem: Under hypothesis & alternative:

Víšek, J. Á. (1998): Robust specification test. Proceedings of Prague Stochastics'98 (eds. Hušková, M. & Lachout, P.), Union of Czechoslovak Mathematicians and Physicists, 581 - 586.

Víšek, J. Á. (1998): Robust instrumental variables and specification test. *PRASTAN 2000, Proceedings of conference "Mathematical statistics, numerical mathematics and their application" (eds. Kalina, M., Minárová, M., Nánásiová, O.), 133 - 164.*

... broken orthogonality condition ... collinearity

Hausman specification test for *M*-estimators

Theorem: Under hypothesis & alternative:

Víšek, J. Á. (1998): Robust specification test. Proceedings of Prague Stochastics'98 (eds. Hušková, M. & Lachout, P.), Union of Czechoslovak Mathematicians and Physicists, 581 - 586.

Víšek, J. Á. (1998): Robust instrumental variables and specification test. *PRASTAN 2000, Proceedings of conference "Mathematical statistics, numerical mathematics and their application" (eds. Kalina, M., Minárová, M., Nánásiová, O.), 133 - 164.*

(I believe that for LWS the same result can be derived

- but an exact proof is still to be written.)

... broken orthogonality condition ... collinearity

Content



... broken orthogonality condition ... collinearity

Recognizing the collinearity and estimating the model



... broken orthogonality condition ... collinearity

Recognizing the collinearity and estimating the model

Condition number $\eta = \frac{\max_{1 \le i \le p} \sqrt{\lambda_i}}{\min_{1 \le i \le p} \sqrt{\lambda_i}}$

where λ_i 's are eigenvalues of X'X. If η is larger than a treshold,

... broken orthogonality condition ... collinearity

Recognizing the collinearity and estimating the model



 $\mathbf{Y} = (\mathbf{X} \mathbf{X} + \mathbf{0} \cdot \mathbf{I}) \mathbf{X} \mathbf{Y},$

Biased - but $\delta \neq 0$ is not the sufficient explanation

... broken orthogonality condition ... collinearity

Recognizing the collinearity and estimating the model



because when X'X is regular matrix, also $X'X + \delta \cdot I$ is regular.

... broken orthogonality condition ... collinearity

Estimating the model under collinearity

An alternative estimator:

The (ordinary) least squares with constraints



... broken orthogonality condition ... collinearity

Estimating the model under collinearity

An alternative estimator:

The (ordinary) least squares with constraints

$$\hat{\beta}^{(OLS,n,C)} = \arg\min_{\beta \in R^{p}} \left\{ \sum_{i=1}^{n} \left(Y_{i} - X_{i}^{\prime} \beta \right)^{2}; C \cdot \beta + \kappa = 0 \right\}$$

They can be employed at least in two situations:



... broken orthogonality condition ... collinearity

Estimating the model under collinearity

An alternative estimator:

The (ordinary) least squares with constraints

$$\hat{\beta}^{(OLS,n,C)} = \arg\min_{\beta \in R^{p}} \left\{ \sum_{i=1}^{n} \left(Y_{i} - X_{i}^{\prime} \beta \right)^{2}; C \cdot \beta + \kappa = 0 \right\}$$

They can be employed at least in two situations:



... broken orthogonality condition ... collinearity

Estimating the model under collinearity

An alternative estimator:

The (ordinary) least squares with constraints

$$\hat{\beta}^{(OLS,n,C)} = \underset{\beta \in R^{p}}{\operatorname{arg\,min}} \left\{ \sum_{i=1}^{n} \left(Y_{i} - X_{i}^{\prime} \beta \right)^{2}; C \cdot \beta + \kappa = 0 \right\}$$

They can be employed at least in two situations:

Depressing the influence of the collinearity

and

improving the combinations of forecasts. A few words about the second possibility

... broken orthogonality condition ... collinearity

Combining forecasts to reach an improvement

Bates, J. M., C. W. J. Granger (1969): The combination of forecasts. Operational Research Quarterly, 20, 451-468.

Clemen, R. T. (1986):

Linear constraints and efficiency of combined forecasts. Journal of Forecasting, 6, 31 - 38.



... broken orthogonality condition ... collinearity

Combining forecasts to reach an improvement

Bates, J. M., C. W. J. Granger (1969): The combination of forecasts. *Operational Research Quarterly, 20, 451-468.*

Clemen, R. T. (1986):

Linear constraints and efficiency of combined forecasts. Journal of Forecasting, 6, 31 - 38.

$$C = \left[\begin{array}{ccc} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 1 \end{array} \right], \quad \kappa = \left[\begin{array}{ccc} 0 \\ 1 \end{array} \right]$$

... broken orthogonality condition ... collinearity

Combining forecasts to reach an improvement

Bates, J. M., C. W. J. Granger (1969): The combination of forecasts. *Operational Research Quarterly, 20, 451-468.*

Clemen, R. T. (1986):

Linear constraints and efficiency of combined forecasts. Journal of Forecasting, 6, 31 - 38.

$$C = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 1 \end{bmatrix}, \quad \kappa = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$
$$\hat{\beta}_1^{(LS,C,n)} = 0 \qquad \sum_{j=2}^p \hat{\beta}_j^{(LS,C,n)} = 1.$$

i.e.

... broken orthogonality condition ... collinearity

Combining forecasts to reach an improvement

Hendry, D. F., M. P. Clements (2004): Pooling of forecasts. Econometrics Journal 7, is 1, 1 - 31.

Araújo, M. B., M. New (2007): Ensemble forecasting of species distributions.

Trends in Ecology & Evolution, 22, 42-47.

(both having more than 150 references)

Jore, A. S., J. Mitchell, S. P. Vahey (2010): Combining forecast densities from VARs with uncertain instabilities.

Special Issue of J. of Applied Econometrics 25, 621-634.

Wonga, K. K. F., H. Song, S. F. Witta and D. C. Wua (2007): Tourism forecasting: To combine or not to combine?

Tourism Management 28, 1068-1078.
... broken orthogonality condition ... collinearity



... broken orthogonality condition ... collinearity

Estimating the model under collinearity

We are looking for:

$$\hat{\beta}^{(OLS,n,C)} = \operatorname{arg\,min}_{\beta \in R^p} \left\{ \sum_{i=1}^n \left(Y_i - X'_i \beta \right)^2; \quad C \cdot \beta + \kappa = 0 \right\}.$$

If κ (can be assumed) random, we can write (for some $\delta \neq 0$)

$$\begin{bmatrix} \mathbf{Y} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{X} \\ \delta \cdot \mathbf{C} \end{bmatrix} \cdot \beta + \begin{bmatrix} \varepsilon \\ \delta \cdot \kappa \end{bmatrix}$$

... broken orthogonality condition ... collinearity

Estimating the model under collinearity

We are looking for: $\hat{\beta}^{(OLS,n,C)} = \arg\min_{\beta \in \mathbb{R}^{p}} \left\{ \sum_{i=1}^{n} \left(Y_{i} - X_{i}^{\prime} \beta \right)^{2}; C \cdot \beta + \kappa = 0 \right\}.$ If κ (can be assumed) random, we can write (for some $\delta \neq 0$) $\begin{bmatrix} \mathbf{Y} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{X} \\ \delta \cdot \mathbf{C} \end{bmatrix} \cdot \beta + \begin{bmatrix} \varepsilon \\ \delta \cdot \kappa \end{bmatrix}$ $\hat{\beta}^{(OLS,n,C)} = \left\{ \begin{bmatrix} X \\ \delta \cdot C \end{bmatrix}' \cdot \begin{bmatrix} X \\ \delta \cdot C \end{bmatrix} \right\}^{-1} \cdot \begin{bmatrix} X \\ \delta \cdot C \end{bmatrix}' \cdot \begin{bmatrix} Y \\ 0 \end{bmatrix}$

... broken orthogonality condition ... collinearity

Estimating the model under collinearity

We are looking for:

$$\hat{\beta}^{(OLS,n,C)} = \arg\min_{\beta \in R^{p}} \left\{ \sum_{i=1}^{n} \left(Y_{i} - X_{i}^{\prime} \beta \right)^{2}; \quad C \cdot \beta + \kappa = 0 \right\}.$$

If κ (can be assumed) random, we can write (for some $\delta \neq 0$)

$$\begin{bmatrix} \mathbf{Y} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{X} \\ \delta \cdot \mathbf{C} \end{bmatrix} \cdot \beta + \begin{bmatrix} \varepsilon \\ \delta \cdot \kappa \end{bmatrix}$$
$$\hat{\beta}^{(OLS,n,C)} = \left\{ \begin{bmatrix} \mathbf{X} \\ \delta \cdot \mathbf{C} \end{bmatrix}' \cdot \begin{bmatrix} \mathbf{X} \\ \delta \cdot \mathbf{C} \end{bmatrix} \right\}^{-1} \cdot \begin{bmatrix} \mathbf{X} \\ \delta \cdot \mathbf{C} \end{bmatrix}' \cdot \begin{bmatrix} \mathbf{Y} \\ \mathbf{0} \end{bmatrix}$$
$$\hat{\beta}^{(OLS,n,C)} = \begin{bmatrix} \mathbf{X}'\mathbf{X} + \delta^2 \cdot \mathbf{C}'\mathbf{C} \end{bmatrix}^{-1} \cdot \mathbf{X}'\mathbf{Y}.$$

(To keep a possibility to follow the explanation, a part of slide is rewritten on the next slide.)

... broken orthogonality condition ... collinearity

We are looking for:

$$\hat{\beta}^{(OLS,n,C)} = \arg \min_{\beta \in \mathbb{R}^{p}} \left\{ \sum_{i=1}^{n} (Y_{i} - X'_{i}\beta)^{2}; C \cdot \beta + \kappa = 0 \right\}$$
(1)
and it gives

$$\hat{\beta}^{(OLS,n,C)} = [X'X + \delta^{2} \cdot C'C]^{-1} \cdot X'Y.$$
(2)

... broken orthogonality condition ... collinearity

We are looking for:

$$\hat{\beta}^{(OLS,n,C)} = \underset{\beta \in \mathbb{R}^{p}}{\arg \min} \left\{ \sum_{i=1}^{n} (Y_{i} - X_{i}'\beta)^{2}; C \cdot \beta + \kappa = 0 \right\}$$
(1)
and it gives

$$\hat{\beta}^{(OLS,n,C)} = \left[X'X + \delta^{2} \cdot C'C \right]^{-1} \cdot X'Y.$$
(2)
So it seems that $\hat{\beta}^{(OLS,n,C)}$ is a special case of $\hat{\beta}^{(R,n)} = (X'X + \delta \cdot I)^{-1} X'Y$, or vice versa.

... broken orthogonality condition ... collinearity

We are looking for:

$$\hat{\beta}^{(OLS,n,C)} = \underset{\beta \in R^{p}}{\arg \min} \left\{ \sum_{i=1}^{n} (Y_{i} - X'_{i}\beta)^{2}; C \cdot \beta + \kappa = 0 \right\}$$
(1)
and it gives

$$\hat{\beta}^{(OLS,n,C)} = [X'X + \delta^{2} \cdot C'C]^{-1} \cdot X'Y.$$
(2)
So it seems that $\hat{\beta}^{(OLS,n,C)}$ is a special case of $\hat{\beta}^{(R,n)} = (X'X + \delta \cdot I)^{-1} X'Y$, or vice versa.
Why is this conclusion wrong ?
Also the derivation of $\hat{\beta}^{(OLS,n,C)}$ needs some attention !!

... broken orthogonality condition ... collinearity

We are looking for:

$$\hat{\beta}^{(OLS,n,C)} = \underset{\beta \in R^{\rho}}{\arg \min} \left\{ \sum_{i=1}^{n} (Y_{i} - X'_{i}\beta)^{2}; C \cdot \beta + \kappa = 0 \right\}$$
(1)
and it gives

$$\hat{\beta}^{(OLS,n,C)} = \left[X'X + \delta^{2} \cdot C'C \right]^{-1} \cdot X'Y.$$
(2)
So it seems that $\hat{\beta}^{(OLS,n,C)}$ is a special case of $\hat{\beta}^{(R,n)} = (X'X + \delta \cdot I)^{-1} X'Y$, or vice versa.
Why is this conclusion wrong ?
Also the derivation of $\hat{\beta}^{(OLS,n,C)}$ needs some attention !!
The answer on the second sentence is straightforward:
Notice that (1) doesn't depend on δ while (2) does.

... broken orthogonality condition ... collinearity

We are looking for:

$$\hat{\beta}^{(OLS,n,C)} = \underset{\beta \in \mathbb{R}^{p}}{\arg \min} \left\{ \sum_{i=1}^{n} (Y_{i} - X'_{i}\beta)^{2}; C \cdot \beta + \kappa = 0 \right\}$$
(1)
and it gives

$$\hat{\beta}^{(OLS,n,C)} = \left[X'X + \delta^{2} \cdot C'C \right]^{-1} \cdot X'Y.$$
(2)
So it seems that $\hat{\beta}^{(OLS,n,C)}$ is a special case of $\hat{\beta}^{(R,n)} = (X'X + \delta \cdot I)^{-1} X'Y$, or vice versa.
Why is this conclusion wrong ?
Also the derivation of $\hat{\beta}^{(OLS,n,C)}$ needs some attention !!
The answer on the second sentence is straightforward:
Notice that (1) doesn't depend on δ while (2) does.
(As κ is not too much specified, the solution of (1) is also solution for

$$\hat{\beta}^{(OLS,n,C)} = \underset{\beta \in \mathbb{R}^{p}}{\arg \min} \left\{ \sum_{i=1}^{n} (Y_{i} - X'_{i}\beta)^{2}; \delta \cdot C \cdot \beta + \tilde{\kappa} = 0 \right\}$$
for some $\tilde{\kappa}$ and it affects of course $\hat{\beta}^{(OLS,n,C,\delta)}$.

... broken orthogonality condition ... collinearity

We are looking for:

$$\hat{\beta}^{(OLS,n,C)} = \underset{\beta \in R^{p}}{\arg \min} \left\{ \sum_{i=1}^{n} (Y_{i} - X'_{i}\beta)^{2}; C \cdot \beta + \kappa = 0 \right\}$$
(1)
and it gives

$$\hat{\beta}^{(OLS,n,C)} = \left[X'X + \delta^{2} \cdot C'C \right]^{-1} \cdot X'Y.$$
(2)
So it seems that $\hat{\beta}^{(OLS,n,C)}$ is a special case of $\hat{\beta}^{(R,n)} = (X'X + \delta \cdot I)^{-1} X'Y$, or vice versa.
Why is this conclusion wrong ?
Also the derivation of $\hat{\beta}^{(OLS,n,C)}$ needs some attention !!
The answer on the second sentence is straightforward:
Notice that (1) doesn't depend on δ while (2) does.
(As κ is not too much specified, the solution of (1) is also solution for

$$\hat{\beta}^{(OLS,n,C)} = \underset{\beta \in R^{p}}{\arg \min} \left\{ \sum_{i=1}^{n} (Y_{i} - X'_{i}\beta)^{2}; \delta \cdot C \cdot \beta + \tilde{\kappa} = 0 \right\}$$
for some $\tilde{\kappa}$ and it affects of course $\hat{\beta}^{(OLS,n,C,\delta)}$.
The answer on the question is a bit more complicated:

... broken orthogonality condition ... collinearity

Estimating the model under collinearity

Consider the constraints $C\beta + \kappa = 0$ (it implies that C has p columns).



... broken orthogonality condition ... collinearity

Estimating the model under collinearity

Consider the constraints $C\beta + \kappa = 0$ (it implies that C has p columns).

What is (or what can be) the number of rows of C (or, equivalently, the dimension of κ)?



... broken orthogonality condition ... collinearity

Estimating the model under collinearity

Consider the constraints $C\beta + \kappa = 0$ (it implies that C has p columns).

- What is (or what can be) the number of rows of C (or, equivalently, the dimension of κ)?
 - The rows of C are (linearly) independent,

otherwise we delete some of them.

... broken orthogonality condition ... collinearity

Estimating the model under collinearity

Consider the constraints $C\beta + \kappa = 0$ (it implies that *C* has *p* columns). What is (or what can be) the <u>number of rows of *C*</u> (or, equivalently, the <u>dimension of κ </u>) ? The rows of *C* are (linearly) independent, otherwise we delete some of them. Denote the number of rows of *C*, i. e. the rank(*C*) by ℓ .

... broken orthogonality condition ... collinearity

Estimating the model under collinearity

Consider the constraints $C\beta + \kappa = 0$ (it implies that C has p columns). What is (or what can be) the number of rows of C (or, equivalently, the dimension of κ)? The rows of C are (linearly) independent, otherwise we delete some of them. Denote the number of rows of C, i.e. the rank(C) by ℓ . If $\ell = p$, then the *p*-tuple of equations $C \cdot \beta + \kappa = 0$ identifies uniquely β and hence the whole extremal problem has solution given by constraints. It is a nonsence.

... broken orthogonality condition ... collinearity

Estimating the model under collinearity

Consider the constraints $C\beta + \kappa = 0$ (it implies that C has p columns). What is (or what can be) the number of rows of C (or, equivalently, the dimension of κ)? The rows of C are (linearly) independent, otherwise we delete some of them. Denote the number of rows of C, i.e. the rank(C) by ℓ . If $\ell = p$, then the *p*-tuple of equations $C \cdot \beta + \kappa = 0$ identifies uniquely β and hence the whole extremal problem has solution given by constraints. It is a nonsence, hence $\ell < p$.

... broken orthogonality condition ... collinearity

Estimating the model under collinearity

Consider the constraints $C\beta + \kappa = 0$ (it implies that C has p columns). What is (or what can be) the number of rows of C (or, equivalently, the dimension of κ)? The rows of C are (linearly) independent, otherwise we delete some of them. Denote the number of rows of C, i.e. the rank(C) by ℓ . If $\ell = p$, then the *p*-tuple of equations $C \cdot \beta + \kappa = 0$ identifies uniquely β and hence the whole extremal problem has solution given by constraints. It is a nonsence, hence $\ell < p$.

We have of course also $\dim(\kappa) = \ell$.

... broken orthogonality condition ... collinearity

Solving the "extremal problem with constraints"

So we have C of type $\ell \times p$. Then there is a matrix \tilde{C} (of type $p - \ell \times p$) so that $\begin{bmatrix} C \\ \tilde{C} \end{bmatrix}$ is regular and $C' \cdot \tilde{C} = 0$. Let's give an example.

... broken orthogonality condition ... collinearity

Solving the "extremal problem with constraints"

So we have *C* of type $\ell \times p$. Then there is a matrix \tilde{C} (of type $p - \ell \times p$) so that $\begin{bmatrix} C \\ \tilde{C} \end{bmatrix}$ is regular and $C' \cdot \tilde{C} = 0$. Let's give an example.

C = [1, 1, 1, 1, 1].

Let



... broken orthogonality condition ... collinearity

Solving the "extremal problem with constraints"

So we have C of type $\ell \times p$. Then there is a matrix C (of type $p - \ell \times p$) so that $\begin{bmatrix} C \\ \tilde{C} \end{bmatrix}$ is regular and $C' \cdot \tilde{C} = 0$. Let's give an example. Let $C = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}$ The matrix \tilde{C} can be selected as $\tilde{C} = \begin{vmatrix} 1 & -1 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & 0 & -1 \end{vmatrix}.$

... broken orthogonality condition ... collinearity

Solving the "extremal problem with constraints"

Now, let β^* fulfills constraints, i. e. $C\beta^* + \kappa = 0$ and define the mapping $\beta(\lambda) : R^{p-\ell} \rightarrow \{\beta \in R^p : C\beta + \kappa = 0\}$ by $\beta(\lambda) = \tilde{C}' \cdot \lambda + \beta^*.$



... broken orthogonality condition ... collinearity

Solving the "extremal problem with constraints"

Now, let β^* fulfills constraints, i. e. $C\beta^* + \kappa = 0$ and define the mapping $\beta(\lambda) : R^{p-\ell} \rightarrow \{\beta \in R^p : C\beta + \kappa = 0\}$ by

$$\beta(\lambda) = \tilde{C}' \cdot \lambda + \beta^*.$$

• Let's verify that it is mapping into $\{\beta \in \mathbb{R}^p : C\beta + \kappa = 0\}$.

... broken orthogonality condition ... collinearity

Solving the "extremal problem with constraints"

Now, let β^* fulfills constraints, i. e. $C\beta^* + \kappa = 0$ and define the mapping $\beta(\lambda) : R^{p-\ell} \rightarrow \{\beta \in R^p : C\beta + \kappa = 0\}$ by $\beta(\lambda) = \tilde{C}' \cdot \lambda + \beta^*.$ • Let's verify that it is mapping into $\{\beta \in R^p : C\beta + \kappa = 0\}$. We have for any $\lambda \in R^{p-\ell}$ $C \cdot \beta(\lambda) = \underbrace{C\tilde{C}'}_{0} \cdot \lambda + C\beta^* = \kappa.$

... broken orthogonality condition ... collinearity

Solving the "extremal problem with constraints"

Now, let β^* fulfills constraints, i. e. $C\beta^* + \kappa = 0$ and define the mapping $\beta(\lambda) : R^{p-\ell} \rightarrow \{\beta \in R^p : C\beta + \kappa = 0\}$ by $\beta(\lambda) = \tilde{C}' \cdot \lambda + \beta^*.$ • Let's verify that it is mapping into $\{\beta \in R^p : C\beta + \kappa = 0\}$. We have for any $\lambda \in R^{p-\ell}$ $C \cdot \beta(\lambda) = C\tilde{C}' \cdot \lambda + C\beta^* = \kappa.$

The mapping is one-to-one.

... broken orthogonality condition ... collinearity

Solving the "extremal problem with constraints"

Now, let β^* fulfills constraints, i. e. $C\beta^* + \kappa = 0$ and define the mapping $\beta(\lambda) : R^{p-\ell} \rightarrow \{\beta \in R^p : C\beta + \kappa = 0\}$ by $\beta(\lambda) = \tilde{C}' \cdot \lambda + \beta^*.$ • Let's verify that it is mapping into $\{\beta \in R^p : C\beta + \kappa = 0\}$. We have for any $\lambda \in R^{p-\ell}$ $C \cdot \beta(\lambda) = C\tilde{C}' \cdot \lambda + C\beta^* = \kappa.$

• The mapping is one-to-one. Assume that for a pair $\lambda_1, \lambda_2 \in \mathbb{R}^{p-\ell}$ $\beta(\lambda_1) = \beta(\lambda_2)$

which means that $\tilde{C}' \cdot (\lambda_1 - \lambda_2) = 0$.

... broken orthogonality condition ... collinearity

Solving the "extremal problem with constraints"

Now, let β^* fulfills constraints, i. e. $C\beta^* + \kappa = 0$ and define the mapping $\beta(\lambda) : R^{p-\ell} \rightarrow \{\beta \in R^p : C\beta + \kappa = 0\}$ by $\beta(\lambda) = \tilde{C}' \cdot \lambda + \beta^*$. • Let's verify that it is mapping into $\{\beta \in R^p : C\beta + \kappa = 0\}$. We have for any $\lambda \in R^{p-\ell}$ $C \cdot \beta(\lambda) = C\tilde{C}' \cdot \lambda + C\beta^* = \kappa$.

 The mapping is one-to-one. Assume that for a pair λ₁, λ₂ ∈ R^{p-ℓ} β(λ₁) = β(λ₂) which means that C̃' ⋅ (λ₁ - λ₂) = 0.
 As the matrix C̃ is of full rank, we have λ₁ - λ₂ = 0.

... broken orthogonality condition ... collinearity

Solving the "extremal problem with constraints"

Now, let β^* fulfills constraints, i. e. $C\beta^* + \kappa = 0$ and define the mapping $\beta(\lambda) : R^{p-\ell} \rightarrow \{\beta \in R^p : C\beta + \kappa = 0\}$ by $\beta(\lambda) = \tilde{C}' \cdot \lambda + \beta^*$. • Let's verify that it is mapping into $\{\beta \in R^p : C\beta + \kappa = 0\}$. We have for any $\lambda \in R^{p-\ell}$ $C \cdot \beta(\lambda) = C\tilde{C}' \cdot \lambda + C\beta^* = \kappa$.

• The mapping is one-to-one. Assume that for a pair $\lambda_1, \lambda_2 \in \mathbb{R}^{p-\ell}$ $\beta(\lambda_1) = \beta(\lambda_2)$ which means that $\tilde{C}' \cdot (\lambda_1 - \lambda_2) = 0$. As the matrix \tilde{C} is of full rank, we have $\lambda_1 - \lambda_2 = 0$.

• The mapping is on $\{\beta \in R^p : C\beta + \kappa = 0\}$.

... broken orthogonality condition ... collinearity

Solving the "extremal problem with constraints"

Now, let β^* fulfills constraints, i. e. $C\beta^* + \kappa = 0$ and define the mapping $\beta(\lambda) : R^{p-\ell} \rightarrow \{\beta \in R^p : C\beta + \kappa = 0\}$ by $\beta(\lambda) = \tilde{C}' \cdot \lambda + \beta^*.$ • Let's verify that it is mapping into $\{\beta \in R^p : C\beta + \kappa = 0\}.$

We have for any $\lambda \in \mathbb{R}^{p-\ell}$ $C \cdot \beta(\lambda) = \underbrace{CC'}{} \cdot \lambda + C\beta^* = \kappa$.

• The mapping is one-to-one. Assume that for a pair $\lambda_1, \lambda_2 \in \mathbb{R}^{p-\ell}$ $\beta(\lambda_1) = \beta(\lambda_2)$ which means that $\tilde{C}' \cdot (\lambda_1 - \lambda_2) = 0$.

As the matrix \hat{C} is of full rank, we have $\lambda_1 - \lambda_2 = 0$.

• The mapping is on $\{\beta \in \mathbb{R}^p : C\beta + \kappa = 0\}$. We have for any $\overline{\beta} \in \{\beta \in \mathbb{R}^p : C\beta + \kappa = 0\}$ $C(\overline{\beta} - \beta^*) = 0$.

... broken orthogonality condition ... collinearity

Solving the "extremal problem with constraints"

Now, let β^* fulfills constraints, i. e. $C\beta^* + \kappa = 0$ and define the mapping $\beta(\lambda) : R^{p-\ell} \rightarrow \{\beta \in R^p : C\beta + \kappa = 0\}$ by $\beta(\lambda) = \tilde{C}' \cdot \lambda + \beta^*.$ • Let's verify that it is mapping into $\{\beta \in R^p : C\beta + \kappa = 0\}$. We have for any $\lambda \in R^{p-\ell}$ $C \cdot \beta(\lambda) = C\tilde{C}' \cdot \lambda + C\beta^* = \kappa.$

The mapping is one-to-one. Assume that for a pair λ₁, λ₂ ∈ R^{p-ℓ} β(λ₁) = β(λ₂) which means that Č' · (λ₁ - λ₂) = 0.
As the matrix Č is of full rank, we have λ₁ - λ₂ = 0.
The mapping is on {β ∈ R^p : Cβ + κ = 0}. We have for any β ∈ {β ∈ R^p : Cβ + κ = 0}. Hence β - β* ⊥ C.

... broken orthogonality condition ... collinearity

Solving the "extremal problem with constraints"

Now, let β^* fulfills constraints, i. e. $C\beta^* + \kappa = 0$ and define the mapping $\beta(\lambda): \mathbb{R}^{p-\ell} \to \{\beta \in \mathbb{R}^p : \mathbb{C}\beta + \kappa = 0\}$ by $\beta(\lambda) = \mathbf{C}' \cdot \lambda + \beta^*.$ Let's verify that it is mapping into $\{\beta \in \mathbb{R}^p : \mathbb{C}\beta + \kappa = 0\}$. We have for any $\lambda \in \mathbb{R}^{p-\ell}$ $C \cdot \beta(\lambda) = C\tilde{C}' \cdot \lambda + C\beta^* = \kappa$. The mapping is one-to-one. Assume that for a pair $\lambda_1, \lambda_2 \in \mathbb{R}^{p-\ell}$ $\beta(\lambda_1) = \beta(\lambda_2)$ which means that $\tilde{C}' \cdot (\lambda_1 - \lambda_2) = 0$. As the matrix \tilde{C} is of full rank, we have $\lambda_1 - \lambda_2 = 0$. • The mapping is on $\{\beta \in \mathbb{R}^p : C\beta + \kappa = 0\}$. We have for any $\bar{\beta} \in \{\beta \in \mathbb{R}^p : C\beta + \kappa = 0\}$ $C(\bar{\beta} - \beta^*) = 0$. Hence $\overline{\beta} - \beta^* \perp C$. So, $\exists \lambda \in \mathbb{R}^{p-\ell}$ such that $\overline{\beta} - \beta^* = \widetilde{C}' \cdot \lambda$ (combination of columns of \widetilde{C}'), i.e. $\bar{\beta} = \tilde{C}' \cdot \lambda + \beta^*$.

... broken orthogonality condition ... collinearity

An example

Let $\beta^0 = (1, -2, 3, -4, 5)'$ and consider the constraint: $\hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3 + \hat{\beta}_4 + \hat{\beta}_5 = 3.$ Then $C = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}$ and $\kappa = \begin{bmatrix} 3 \end{bmatrix}$.

... broken orthogonality condition ... collinearity

An example

Let $\beta^0 = (1, -2, 3, -4, 5)'$ and consider the constraint: $\hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3 + \hat{\beta}_4 + \hat{\beta}_5 = 3.$ Then $C = [1 \ 1 \ 1 \ 1 \ 1 \]$ and $\kappa = [3].$ The matrix \tilde{C} and β^* can be selected as $\tilde{C} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & 0 & -1 \end{bmatrix} \text{ and } \beta^* = \begin{bmatrix} 1 & -2 & 3 & -4 & 5 \end{bmatrix}'.$

... broken orthogonality condition ... collinearity

Solving the "extremal problem with constraints" (continued)

Let again β^* fulfills constraints, i. e. $C\beta^* + \kappa = 0$ and put

$$ilde{Y} = Y - X eta^*, \quad ilde{X} = X \cdot ilde{C}'.$$



... broken orthogonality condition ... collinearity

Solving the "extremal problem with constraints" (continued)

Let again β^* fulfills constraints, i. e. $C\beta^* + \kappa = 0$ and put

$$\begin{split} \widetilde{Y} &= Y - X\beta^*, \quad \widetilde{X} = X \cdot \widetilde{C}'. \\ \text{Then for any } \lambda \in R^{n-\ell} \quad \text{and} \quad \beta(\lambda) = \widetilde{C}' \cdot \lambda + \beta^* \text{ we have:} \\ \forall i \quad \widetilde{Y}_i - \widetilde{X}'_i \lambda = Y_i - X'_i \beta^* - X'_i \cdot \widetilde{C}' \lambda = Y_i - X'_i \left(\widetilde{C}' \lambda + \beta^*\right) = Y_i - X'_i \beta(\lambda). \end{split}$$

... broken orthogonality condition ... collinearity

Solving the "extremal problem with constraints" (continued)

Let again β^* fulfills constraints, i. e. $C\beta^* + \kappa = 0$ and put $\tilde{Y} = Y - X\beta^*$. $\tilde{X} = X \cdot \tilde{C}'$. Then for any $\lambda \in \mathbb{R}^{n-\ell}$ and $\beta(\lambda) = \tilde{C}' \cdot \lambda + \beta^*$ we have: $\forall i \quad \tilde{Y}_i - \tilde{X}'_i \lambda = Y_i - X'_i \beta^* - X'_i \cdot \tilde{C}' \lambda = Y_i - X'_i \left(\tilde{C}' \lambda + \beta^* \right) = Y_i - X'_i \beta(\lambda).$ It means that $\hat{\lambda}^{(OLS,n)} = \underset{\lambda \in B^{p-\ell}}{\operatorname{arg\,min}} \left\{ \sum_{i=1}^{n} \left(\tilde{Y}_{i} - \tilde{X}_{i}^{\prime} \lambda \right)^{2} \right\} = \underset{\lambda \in B^{p-\ell}}{\operatorname{arg\,min}} \left\{ \sum_{i=1}^{n} \left(Y_{i} - X_{i}^{\prime} \beta(\lambda) \right)^{2} \right\}$ $= \inf_{\overline{\beta} \in \{\beta \in R^{p}: C\beta + \kappa = 0\}} \left\{ \sum_{i=1}^{n} \left(Y_{i} - X_{i}^{\prime} \overline{\beta} \right)^{2} \right\}.$

... broken orthogonality condition ... collinearity

Robustifying the Ordinary Least squares with constraints

It means that for

$$\hat{\lambda}^{(OLS,n)} = \underset{\lambda \in \mathcal{R}^{n-\ell}}{\operatorname{arg\,min}} \left\{ \sum_{i=1}^{n} \left(\tilde{Y}_{i} - \tilde{X}_{i}^{\prime} \lambda \right)^{2} \right\},$$

 $\hat{eta}(\lambda) = \tilde{C}' \cdot \hat{\lambda}^{(OLS,n)} + \beta^*$ solves
... broken orthogonality condition ... collinearity

Robustifying the Ordinary Least squares with constraints

It means that for $\hat{\lambda}^{(OLS,n)} = \underset{\lambda \in R^{p-\epsilon}}{\operatorname{arg\,min}} \left\{ \sum_{i=1}^{n} \left(\tilde{Y}_{i} - \tilde{X}_{i}' \lambda \right)^{2} \right\},$ $\hat{\beta}(\lambda) = \tilde{C}' \cdot \hat{\lambda}^{(OLS,n)} + \beta^{*} \text{ solves}$ $\hat{\beta}(\lambda) = \underset{\beta \in R^{p}}{\operatorname{arg\,min}} \left\{ \sum_{i=1}^{n} \left(Y_{i} - X_{i}' \beta \right)^{2}; C \cdot \beta + \kappa = 0 \right\}.$

... broken orthogonality condition ... collinearity

Estimating robustly the model under collinearity

Robustifying the least squares with constraints:

The least weighted squares with constraints

$$\hat{\beta}^{(LWS,n,C)} = \arg\min_{\beta \in \mathbb{R}^{p}} \left\{ \sum_{i=1}^{n} w\left(\frac{i-1}{n}\right) r_{(i)}^{2}(\beta); \quad C \cdot \beta + \kappa = 0 \right\}$$

... broken orthogonality condition ... collinearity

Estimating robustly the model under collinearity

Robustifying the least squares with constraints:

The least weighted squares with constraints

$$\hat{\beta}^{(LWS,n,C)} = \underset{\beta \in R^{\rho}}{\operatorname{arg\,min}} \left\{ \sum_{i=1}^{n} w\left(\frac{i-1}{n}\right) r_{(i)}^{2}(\beta); \quad C \cdot \beta + \kappa = 0 \right\}$$

They can be calculated as LWS without constraints for data $\left(\tilde{Y}, \tilde{X}\right)$

- that's all.

... broken orthogonality condition ... collinearity

Numerical study

We have generated 1000 data-sets, as follows

 $\left\{ \left\{ X_{i}^{(k)}, \ e_{i}^{(k)}, \ \varepsilon_{i}^{(k)} \right\}_{i=1}^{100} \right\}_{k=1}^{1000} \text{ and } \left\{ \left\{ \sigma_{i}^{(k)} \right\}_{i=1}^{100} \right\}_{k=1}^{1000} \right\}_{k=1}^{1000}$

with $X_i^{(k)} = \left(X_{i1}^{(k)}, X_{i2}^{(k)}, X_{i3}^{(k)}\right)'$'s, $e_i^{(k)}$'s and $\varepsilon_i^{(k)}$'s normally distributed, and $\sigma_i^{(k)}$'s uniformly distributed over [0.5, 3.5],

 $U_{ij}^{(k)} = X_{ij}^{(k)}, j = 1, 2, 3 \text{ and } U_{i,3+\ell}^{(k)} = 0.5 * X_{(i,\rho-1-\ell)}^{(k)} + 0.5 * X_{(i,\rho-2-\ell)}^{(k)} + \eta \cdot \varepsilon_{i\ell}^{(k)},$ for $\ell = 1, 2, i = 1, 2, ..., 100, k = 1, 2, ..., 1000.$

... broken orthogonality condition ... collinearity

Numerical study

We have generated 1000 data-sets, as follows

$$\left\{ \left\{ X_{i}^{(k)}, \ e_{i}^{(k)}, \ \varepsilon_{i}^{(k)} \right\}_{i=1}^{100} \right\}_{k=1}^{1000} \text{ and } \left\{ \left\{ \sigma_{i}^{(k)} \right\}_{i=1}^{100} \right\}_{k=1}^{1000} \right\}_{k=1}^{1000}$$

with $X_i^{(k)} = \left(X_{i1}^{(k)}, X_{i2}^{(k)}, X_{i3}^{(k)}\right)'$'s, $e_i^{(k)}$'s and $\varepsilon_i^{(k)}$'s normally distributed, and $\sigma_i^{(k)}$'s uniformly distributed over [0.5, 3.5],

 $U_{ij}^{(k)} = X_{ij}^{(k)}, j = 1, 2, 3 \text{ and } U_{i,3+\ell}^{(k)} = 0.5 * X_{(i,\rho-1-\ell)}^{(k)} + 0.5 * X_{(i,\rho-2-\ell)}^{(k)} + \eta \cdot \varepsilon_{i\ell}^{(k)},$ for $\ell = 1, 2, i = 1, 2, ..., 100, k = 1, 2, ..., 1000.$ Then we calculated for $\beta^0 = [1, -2, 3, -4, 5]'$

 $W_{i}^{(k)} = \sum_{j=1}^{p} U_{ij}^{(k)} \cdot \beta_{j}^{0} + e_{i}^{(k)} \cdot \sigma_{i}^{(k)} \text{ and employed data } \left\{ \left\{ W_{i}^{(k)}, U_{i}^{(k)} \right\}_{i=1}^{100} \right\}_{k=1}^{100}.$

... broken orthogonality condition ... collinearity



... broken orthogonality condition ... collinearity

Employing one constraint



... broken orthogonality condition ... collinearity

Employing one constraint



Notice that the rows of matrix \tilde{C} are not orthogonal each to other.

... broken orthogonality condition ... collinearity

Numerical study (continued)

TABLE 1

The disturbances are heteroscedastic ($0.5 \le \sigma_i^2 \le 3.5$) and independent from explanatory variables.

... broken orthogonality condition ... collinearity

Numerical study (continued)

TABLE 1

- The disturbances are heteroscedastic ($0.5 \le \sigma_i^2 \le 3.5$) and independent from explanatory variables.
- 2 Data are not contaminated but there is a collinearity, the mean value of condition numbers was equal to 41.25.

... broken orthogonality condition ... collinearity

Numerical study (continued)

TABLE 1

- The disturbances are heteroscedastic $(0.5 \le \sigma_i^2 \le 3.5)$ and independent from explanatory variables.
- 2 Data are not contaminated but there is a collinearity, the mean value of condition numbers was equal to 41.25.
- The level of robustness was fixed: the number of observations *h* taken into account by LTS was 95% of *n*, the weight function *w* had h = 85% and g = 95% of *n*.

... broken orthogonality condition ... collinearity

Numerical study (continued)

TABLE 1

- The disturbances are heteroscedastic $(0.5 \le \sigma_i^2 \le 3.5)$ and independent from explanatory variables.
- 2 Data are not contaminated but there is a collinearity, the mean value of condition numbers was equal to 41.25.
- The level of robustness was fixed: the number of observations *h* taken into account by LTS was 95% of *n*, the weight function *w* had h = 85% and g = 95% of *n*.
- The collinearity is depressed by one constraint condition (see previous slide).

... broken orthogonality condition ... collinearity

	10 11-5	1.2010	STA DY		
$\hat{\beta}^{OLS}_{(MSE(\hat{\beta}^{OLS}))}$	1.00(0.808)	-2.00 _(1.619)	3.01 _(0.855)	-4.01 _(3.413)	5.00 _(3.228)
$\hat{\beta}^{LWS}_{(MSE(\hat{\beta}^{LWS}))}$	1.02 _(0.851)	-1.94 _(1.745)	3.04 _(0.966)	-4.08 _(3.839)	4.96 _(3.378)
$\hat{\beta}^{LTS}$ (MSE($\hat{\beta}^{LTS}$))	1.03 _(0.868)	-1.93 _(1.760)	3.05 _(0.961)	-4.09 _(3.813)	4.94 _(3.455)
$\hat{\beta}^{OLSC}_{(MSE(\hat{\beta}^{OLSC}))}$	1.00 _(0.797)	-1.99 _(1.601)	3.01 _(0.839)	-4.01 _(3.351)	5.00 _(3.187)
$\hat{\beta}^{LWSC}_{(MSE(\hat{\beta}^{LWSC}))}$	1.03 _(0.829)	-1.94 _(1.697)	3.04 _(0.912)	-4.07 _(3.633)	4.94 _(3.303)
$\hat{\beta}^{LTSC}_{(MSE(\hat{\beta}^{LTSC}))}$	1.03 _(0.851)	-1.94(1.703)	3.03 _(0.946)	-4.06(3.766)	4.94 _(3.387)

The other tables for the mean values of the condition numbers 1.16, 9.95, 19.95, 31.01, 32.79, 48.97, 61.22, 69.78 and 98.47 are available on http://samba.fsv.cuni.cz/~visek/Constraints_Heteroscedasticity_Numerical_Study/visek_NumericalStudy.

... broken orthogonality condition ... collinearity

What about to employing two constraints ?

Now we consider two constraints:



... broken orthogonality condition ... collinearity

What about to employing two constraints ?

Now we consider two constraints:

 $\sum_{j=1}^{n} \hat{\beta}_{j} = 3 \quad \text{and} \quad \hat{\beta}_{1} + \hat{\beta}_{2} = -1.$ Then $C = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \kappa = \begin{bmatrix} 3 \\ -1 \end{bmatrix}.$ The matrix \tilde{C} and β^{*} can be selected as $\tilde{C} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 & -1 \end{bmatrix} \quad \text{and} \quad \beta^{*} = \begin{bmatrix} 1 & -2 & 3 & -4 & 5 \end{bmatrix}'.$

Notice that the rows of matrix C are not orthogonal each to other. The same is true about the rows of \tilde{C} .

... broken orthogonality condition ... collinearity

Numerical study(continued)

TABLE 2

Nearly the same framework as in previous case - except of underlined.

The disturbances are heteroscedastic $(0.5 \le \sigma_i^2 \le 3.5)$ and independent from explanatory variables.

... broken orthogonality condition ... collinearity

Numerical study(continued)

TABLE 2

- The disturbances are heteroscedastic ($0.5 \le \sigma_i^2 \le 3.5$) and independent from explanatory variables.
- 2 Data are not contaminated but there is a collinearity, the mean value of condition numbers was equal to 32.79.

... broken orthogonality condition ... collinearity

Numerical study(continued)

TABLE 2

- The disturbances are heteroscedastic ($0.5 \le \sigma_i^2 \le 3.5$) and independent from explanatory variables.
- 2 Data are not contaminated but there is a collinearity, the mean value of condition numbers was equal to 32.79.
- The level of robustness was fixed: the number of observations *h* taken into account by LTS was 95% of *n*, the weight function *w* had h = 85% and g = 95% of *n*.

... broken orthogonality condition ... collinearity

Numerical study(continued)

TABLE 2

- The disturbances are heteroscedastic ($0.5 \le \sigma_i^2 \le 3.5$) and independent from explanatory variables.
- 2 Data are not contaminated but there is a collinearity, the mean value of condition numbers was equal to 32.79.
- The level of robustness was fixed: the number of observations *h* taken into account by LTS was 95% of *n*, the weight function *w* had h = 85% and g = 95% of *n*.
- The collinearity is depressed by two constraints (see previous slide).

... broken orthogonality condition ... collinearity

	100	A VALER OF MAL	N. Contraction		
$\hat{\beta}^{OLS}_{(MSE(\hat{\beta}^{OLS}))}$	1.00 _(0.559)	-2.01 _(1.179)	2.98 _(0.609)	-3.97 _(3.402)	5.00 _(2.224)
$\hat{\beta}^{LWS}_{(MSE(\hat{\beta}^{LWS}))}$	1.02 _(0.543)	-2.00 _(1.086)	2.99 _(0.546)	-3.97 _(2.172)	4.96 _(2.154)
$\hat{\beta}^{LTS}$ (MSE($\hat{\beta}^{LTS}$))	1.01 _(0.557)	-2.00 _(1.165)	2.99 _(0.594)	-3.98 _(2.360)	4.97 _(2.219)
$\hat{\beta}OLSC$ (MSE($\hat{\beta}OLSC$))	1.01(0.115)	-2.01 _(0.115)	2.99 _(0.455)	-3.97 _(1.795)	4.98 _(0.449)
$\hat{\beta}^{LWSC}$ (MSE($\hat{\beta}^{LWSC}$))	1.01 _(0.108)	-2.01 _(0.108)	2.99 _(0.425)	-3.97 _(1.686)	4.98 _(0.425)
$\hat{\beta}^{LTSC}_{(MSE(\hat{\beta}^{LTSC}))}$	1.01 _(0.112)	-2.01 _(0.112)	2.99 _(0.444)	-3.97 _(1.760)	4.98 _(0.443)

The other tables for the mean values of the condition numbers 1.16, 9.95, 19.95, 31.01, 41.25, 48.97, 61.22, 69.78 and 98.47 are available again on http://samba.fsv.cuni.cz/~visek/Constraints_Heteroscedasticity_Numerical_Study/visek_NumericalStudy.

... broken orthogonality condition ... collinearity

	10	A REPORT	N. Con		
$\hat{\beta}^{OLS}_{(MSE(\hat{\beta}^{OLS}))}$	1.00 _(0.559)	-2.01 _(1.179)	2.98 _(0.609)	-3.97 _(3.402)	5.00 _(2.224)
$\hat{\beta}^{LWS}_{(MSE(\hat{\beta}^{LWS}))}$	1.02 _(0.543)	-2.00 _(1.086)	2.99 _(0.546)	-3.97 _(2.172)	4.96 _(2.154)
$\hat{\beta}^{LTS}_{(MSE(\hat{\beta}^{LTS}))}$	1.01 _(0.557)	-2.00 _(1.165)	2.99 _(0.594)	-3.98 _(2.360)	4.97 _(2.219)
$\hat{\beta}^{OLSC}_{(MSE(\hat{\beta}^{OLSC}))}$	1.01 _(0.115)	-2.01 _(0.115)	2.99 _(0.455)	-3.97 _(1.795)	4.98 _(0.449)
$\hat{\beta}^{LWSC}$ (MSE($\hat{\beta}^{LWSC}$))	1.01 _(0.108)	-2.01 _(0.108)	2.99 _(0.425)	-3.97 _(1.686)	4.98 _(0.425)
$\hat{\beta}^{LTSC}_{(MSE(\hat{\beta}^{LTSC}))}$	1.01 _(0.112)	-2.01 _(0.112)	2.99 _(0.444)	-3.97 _(1.760)	4.98 _(0.443)

The other tables for the mean values of the condition numbers 1.16, 9.95, 19.95, 31.01, 41.25, 48.97, 61.22, 69.78 and 98.47 are available again on http://samba.fsv.cuni.cz/~visek/Constraints_Heteroscedasticity_Numerical_Study/visek_NumericalStudy.

The results are much better than in Table 1 - compare MSE's.

... broken orthogonality condition ... collinearity

What about to employ two orthogonal constraints ?

Let $\beta^0 = (1, -2, 3, -4, 5)'$ and consider two orthogonal constraints:



... broken orthogonality condition ... collinearity

What about to employ two orthogonal constraints ?

Let $\beta^0 = (1, -2, 3, -4, 5)'$ and consider two orthogonal constraints:

 $\hat{\beta}_1 + \hat{\beta}_2 = -1 \quad \text{and} \quad \hat{\beta}_3 + \hat{\beta}_4 = -1.$ Then $C = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix} \quad \text{and} \quad \kappa = \begin{bmatrix} -1 \\ -1 \end{bmatrix}.$ The matrix \tilde{C} and β^* can be selected as $\tilde{C} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \beta^* = \begin{bmatrix} 1 & -2 & 3 & -4 & 5 \end{bmatrix}'.$

... broken orthogonality condition ... collinearity

How the transformation of variables looks like ? An example

We have $\begin{aligned}
\tilde{Y}_{i} &= Y_{i} - X_{i}^{\prime}\beta^{*} \quad \text{and} \quad \tilde{X}_{i} &= \tilde{C}X_{i} \\
\text{i. e.} \quad \tilde{X}_{i1} &= X_{i1} - X_{i2} , \quad \tilde{X}_{i2} &= X_{i3} - X_{i4} \quad \text{and} \quad \tilde{X}_{i3} &= X_{i5}. \\
\text{Finally,} \quad \hat{\beta} &= \tilde{C}^{\prime}\hat{\lambda} + \beta^{*} \\
\text{i. e.} \quad \hat{\beta}_{1} &= \hat{\lambda}_{1} + 1, \quad \hat{\beta}_{2} &= -\hat{\lambda}_{1} - 2, \quad \hat{\beta}_{3} &= \hat{\lambda}_{2} + 3 \\
\hat{\beta}_{4} &= -\hat{\lambda}_{2} - 4 \quad \text{and} \quad \hat{\beta}_{5} &= \hat{\lambda}_{3} + 5.
\end{aligned}$

... broken orthogonality condition ... collinearity

Numerical study_(continued)

TABLE 3

Nearly the same framework as in previous case - except of underlined.

The disturbances are heteroscedastic $(0.5 \le \sigma_i^2 \le 3.5)$ and independent from explanatory variables.

... broken orthogonality condition ... collinearity

Numerical study(continued)

TABLE 3

- The disturbances are heteroscedastic $(0.5 \le \sigma_i^2 \le 3.5)$ and independent from explanatory variables.
- 2 Data are not contaminated but there is a collinearity, the mean value of condition numbers was equal to 32.95.

... broken orthogonality condition ... collinearity

Numerical study(continued)

TABLE 3

- The disturbances are heteroscedastic $(0.5 \le \sigma_i^2 \le 3.5)$ and independent from explanatory variables.
- 2 Data are not contaminated but there is a collinearity, the mean value of condition numbers was equal to 32.95.
- The level of robustness was fixed: the number of observations *h* taken into account by LTS was 95% of *n*, the weight function *w* had h = 85% and g = 95% of *n*.

... broken orthogonality condition ... collinearity

Numerical study(continued)

TABLE 3

- The disturbances are heteroscedastic $(0.5 \le \sigma_i^2 \le 3.5)$ and independent from explanatory variables.
- 2 Data are not contaminated but there is a collinearity, the mean value of condition numbers was equal to 32.95.
- The level of robustness was fixed: the number of observations *h* taken into account by LTS was 95% of *n*, the weight function *w* had h = 85% and g = 95% of *n*.
- We have employed two orthogonal constraints (see previous slide).

... broken orthogonality condition ... collinearity

$\hat{\beta}^{OLS}_{(MSE(\hat{\beta}^{OLS}))}$	1.05 _(0.593)	-1.91 _(1.167)	3.04 _(0.579)	-4.08 _(2.297)	4.91 _(2.359)
$\hat{\beta}^{LWS}_{(MSE(\hat{\beta}^{LWS}))}$	1.04 _(0.573)	-1.93 _(1.065)	3.03 _(0.525)	-4.06 _(2.076)	4.92 _(2.279)
$\hat{\beta}^{LTS}$ (MSE($\hat{\beta}^{LTS}$))	1.04 _(0.615)	-1.91 _(1.174)	3.05 _(0.572)	-4.11 _(2.275)	4.93 _(3.435)
$\hat{\beta}^{OLSC}$ (MSE($\hat{\beta}^{OLSC}$))	1.00 _(0.005)	-2.00 _(0.005)	3.00 _(0.021)	-4.00 _(0.021)	5.01 _(0.015)
$\hat{\beta}^{LWSC}_{(MSE(\hat{\beta}^{LWSC}))}$	1.00 _(0.004)	-2.00 _(0.004)	3.00 _(0.019)	-4.00 _(0.019)	5.01 _(0.014)
$\hat{\beta}^{LTSC}_{(MSE(\hat{\beta}^{LTSC}))}$	1.00 _(0.004)	-2.00 _(0.004)	3.00 _(0.021)	-4.00 _(0.021)	5.01 _(0.015)

The other tables for some other mean values of the condition numbers are available again on http://samba.fsv.cuni.cz/~visek/Constraints_Heteroscedasticity_Numerical_Study/visek_ NumericalStudy.

... broken orthogonality condition ... collinearity

$\hat{\beta}^{OLS}_{(MSE(\hat{\beta}^{OLS}))}$	1.05 _(0.593)	-1.91 _(1.167)	3.04 _(0.579)	-4.08 _(2.297)	4.91 _(2.359)
$\hat{\beta}^{LWS}_{(MSE(\hat{\beta}^{LWS}))}$	1.04 _(0.573)	-1.93 _(1.065)	3.03 _(0.525)	-4.06 _(2.076)	4.92 _(2.279)
$\hat{\beta}^{LTS}$ (MSE($\hat{\beta}^{LTS}$))	1.04 _(0.615)	-1.91 _(1.174)	3.05 _(0.572)	-4.11 _(2.275)	4.93 _(3.435)
$\hat{\beta}^{OLSC}$ (MSE($\hat{\beta}^{OLSC}$))	1.00 _(0.005)	-2.00 _(0.005)	3.00 _(0.021)	-4.00 _(0.021)	5.01 _(0.015)
$\hat{\beta}^{LWSC}_{(MSE(\hat{\beta}^{LWSC}))}$	1.00 _(0.004)	-2.00 _(0.004)	3.00 _(0.019)	-4.00 _(0.019)	5.01 _(0.014)
$\hat{\beta}^{LTSC}_{(MSE(\hat{\beta}^{LTSC}))}$	1.00 _(0.004)	-2.00 _(0.004)	3.00 _(0.021)	-4.00 _(0.021)	5.01 _(0.015)

The other tables for some other mean values of the condition numbers are available again on http://samba.fsv.cuni.cz/~visek/Constraints_Heteroscedasticity_Numerical_Study/visek_ NumericalStudy.

The results are again even much better than in Table 2 - compare MSE's.

... broken orthogonality condition ... collinearity

$\hat{\beta}^{OLS}_{(MSE(\hat{\beta}^{OLS}))}$	1.05 _(0.593)	-1.91 _(1.167)	3.04 _(0.579)	-4.08 _(2.297)	4.91 _(2.359)
$\hat{\beta}^{LWS}_{(MSE(\hat{\beta}^{LWS}))}$	1.04 _(0.573)	-1.93 _(1.065)	3.03 _(0.525)	-4.06 _(2.076)	4.92 _(2.279)
$\hat{\beta}^{LTS}$ (MSE($\hat{\beta}^{LTS}$))	1.04 _(0.615)	-1.91 _(1.174)	3.05 _(0.572)	-4.11 _(2.275)	4.93 _(3.435)
$\hat{\beta}^{OLSC}$ (MSE($\hat{\beta}^{OLSC}$))	1.00 _(0.005)	-2.00 _(0.005)	3.00 _(0.021)	-4.00 _(0.021)	5.01 _(0.015)
$\hat{\beta}^{LWSC}_{(MSE(\hat{\beta}^{LWSC}))}$	1.00 _(0.004)	-2.00 _(0.004)	3.00 _(0.019)	-4.00 _(0.019)	5.01 _(0.014)
$\hat{\beta}^{LTSC}_{(MSE(\hat{\beta}^{LTSC}))}$	1.00 _(0.004)	-2.00 _(0.004)	3.00 _(0.021)	-4.00 _(0.021)	5.01 _(0.015)

The other tables for some other mean values of the condition numbers are available again on http://samba.fsv.cuni.cz/~visek/Constraints_Heteroscedasticity_Numerical_Study/visek_ NumericalStudy.

The results are again even much better than in Table 2 - compare MSE's.

Hence we will keep these constraints in the rest of study.

1

... broken orthogonality condition ... collinearity

We are going to make an idea what a contamination can cause.

TABLE 4

Nearly the same framework as in previous case - except of underlined.

The disturbances are heteroscedastic ($0.5 \le \sigma_i^2 \le 3.5$)

and independent from explanatory variables.

... broken orthogonality condition ... collinearity

We are going to make an idea what a contamination can cause.

TABLE 4

Nearly the same framework as in previous case - except of underlined.

The disturbances are heteroscedastic ($0.5 \le \sigma_i^2 \le 3.5$)

and independent from explanatory variables.

2 There is a collinearity,

the mean value of condition numbers was equal to 16.55.

... broken orthogonality condition ... collinearity

We are going to make an idea what a contamination can cause.

TABLE 4

Nearly the same framework as in previous case - except of underlined.

The disturbances are heteroscedastic ($0.5 \le \sigma_i^2 \le 3.5$)

and independent from explanatory variables.

- There is a collinearity, the mean value of condition numbers was equal to <u>16.55</u>.
- The level of robustness was fixed: the number of observations *h* taken into account by LTS was 90% of *n*, the weight function *w* had h = 65% and g = 90% of *n*.

... broken orthogonality condition ... collinearity

We are going to make an idea what a contamination can cause.

TABLE 4

Nearly the same framework as in previous case - except of underlined.

The disturbances are heteroscedastic (0.5 $\leq \sigma_i^2 \leq$ 3.5)

and independent from explanatory variables.

- There is a collinearity, the mean value of condition numbers was equal to <u>16.55</u>.
- The level of robustness was fixed: the number of observations *h* taken into account by LTS was 90% of *n*, the weight function *w* had h = 65% and g = 90% of *n*.
- We have employed two orthogonal constraints.

... broken orthogonality condition ... collinearity

We are going to make an idea what a contamination can cause.

TABLE 4

Nearly the same framework as in previous case - except of underlined.

The disturbances are heteroscedastic ($0.5 \le \sigma_i^2 \le 3.5$)

and independent from explanatory variables.

There is a collinearity, the mean value of cor

the mean value of condition numbers was equal to 16.55.

- The level of robustness was fixed: the number of observations *h* taken into account by LTS was 90% of *n*, the weight function *w* had h = 65% and g = 90% of *n*.
- We have employed two orthogonal constraints.

The 5% contamination by outliers $Y^{(contaminated)} = -2 * Y^{(original)}$
... broken orthogonality condition ... collinearity

$\hat{\beta}^{OLS}_{(MSE(\hat{\beta}^{OLS}))}$	0.02 _(5.943)	-1.44 _(8.821)	3.41 _(4.669)	-3.45 _(16.254)	4.10 _(19.377)	
$\hat{\beta}^{LWS}_{(MSE(\hat{\beta}^{LWS}))}$	1.00(0.160)	-2.02 _(0.314)	2.99 _(0.164)	-3.97 _(0.620)	5.01 _(0.612)	
$\hat{\beta}^{LTS}$ (MSE($\hat{\beta}^{LTS}$))	0.99 _(0.170)	-2.02 _(0.336)	2.99 _(0.182)	-3.97 _(0.692)	5.02 _(0.652)	
$\hat{\beta}^{OLSC}_{(MSE(\hat{\beta}^{OLSC}))}$	0.25 _(0.958)	-1.25 _(0.958)	2.38 _(1.639)	-3.38 _(1.639)	3.66 _(3.231)	
$\hat{\beta}^{LWSC}$ (MSE($\hat{\beta}^{LWSC}$))	1.00(0.005)	-2.00 _(0.005)	3.00 _(0.023)	-4.00 _(0.023)	5.01 _(0.019)	
$\hat{\beta}^{LTSC}_{(MSE(\hat{\beta}^{LTSC}))}$	1.00 _(0.006)	-2.00 _(0.006)	3.00(0.028)	-4.00 _(0.028)	5.01 _(0.023)	

... broken orthogonality condition ... collinearity

$\hat{\beta}^{OLS}_{(MSE(\hat{\beta}^{OLS}))}$	0.02 _(5.943)	-1.44 _(8.821)	3.41 _(4.669)	-3.45 _(16.254)	4.10 _(19.377)		
$\hat{\beta}^{LWS}_{(MSE(\hat{\beta}^{LWS}))}$	1.00 _(0.160)	-2.02 _(0.314)	2.99 _(0.164)	-3.97 _(0.620)	5.01 _(0.612)		
$\hat{\beta}^{LTS}$ (MSE($\hat{\beta}^{LTS}$))	0.99 _(0.170)	-2.02 _(0.336)	2.99 _(0.182)	-3.97 _(0.692)	5.02 _(0.652)		
$\hat{\beta}^{OLSC}_{(MSE(\hat{\beta}^{OLSC}))}$	0.25 _(0.958)	-1.25 _(0.958)	2.38 _(1.639)	-3.38 _(1.639)	3.66 _(3.231)		
$\hat{\beta}^{LWSC}$ (MSE($\hat{\beta}^{LWSC}$))	1.00 _(0.005)	-2.00 _(0.005)	3.00 _(0.023)	-4.00 _(0.023)	5.01 _(0.019)		
$\hat{\beta}^{LTSC}_{(MSE(\hat{\beta}^{LTSC}))}$	1.00 _(0.006)	-2.00 _(0.006)	3.00 _(0.028)	-4.00 _(0.028)	5.01 _(0.023)		

The other tables for some other mean values of the condition numbers are available again on http://samba.fsv.cuni.cz/~visek/Constraints_Heteroscedasticity_Numerical_Study/visek_ NumericalStudy.

The results are again even much better than in Table 2 - compare MSE's.

... broken orthogonality condition ... collinearity

We are going to make an idea what a contamination can cause.

TABLE 4

Nearly the same framework as in previous cases - except of underlined.

There is a collinearity,

the mean value of condition numbers was equal to 32.60.

... broken orthogonality condition ... collinearity

$\hat{\beta}^{OLS}_{(MSE(\hat{\beta}^{OLS}))}$	0.22 _(16.606)	-1.54 _(31.568)	2.12 _(18.144)	-2.90 _(69.279)	3.72 _(64.889)
$\hat{\beta}^{LWS}_{(MSE(\hat{\beta}^{LWS}))}$	1.06 _(0.636)	-1.92 _(1.279)	3.01 _(0.622)	-4.03 _(3.485)	4.88 _(2.541)
$\hat{\beta}LTS (MSE(\hat{\beta}^{LTS}))$	1.04 _(0.686)	-1.95 _(1.286)	3.00 _(0.653)	- 3 .99 _(2.607)	4.91 _(2.700)
$\hat{\beta}^{OLSC}_{(MSE(\hat{\beta}^{OLSC}))}$	0.27 _(0.915)	-1.27 _(0.915)	2.34 _(1.848)	- 3 .34 _(1.848)	3.62 _(3.488)
$\hat{\beta}^{LWSC}_{(MSE(\hat{\beta}^{LWSC}))}$	1.00 _(0.005)	-2.00 _(0.005)	3.00 _(0.026)	-4.00(0.026)	5.00 _(0.021)
$\hat{\beta}^{LTSC}_{(MSE(\hat{\beta}^{LTSC}))}$	1.00 _(0.006)	-2.00 _(0.006)	3.01 _(0.030)	-4.01 _(0.030)	5.00 _(0.024)

... broken orthogonality condition ... collinearity

$\hat{\beta}^{OLS}_{(MSE(\hat{\beta}^{OLS}))}$	0.22 _(16.606)	-1.54 _(31.568)	2.12 _(18.144)	-2.90 _(69.279)	3.72 _(64.889)
$\hat{\beta}^{LWS}_{(MSE(\hat{\beta}^{LWS}))}$	1.06 _(0.636)	-1.92 _(1.279)	3.01 _(0.622)	-4.03 _(3.485)	4.88 _(2.541)
$\hat{\beta}LTS (MSE(\hat{\beta}^{LTS}))$	1.04 _(0.686)	-1.95 _(1.286)	3.00 _(0.653)	- 3 .99 _(2.607)	4.91 _(2.700)
$\hat{\beta}^{OLSC}_{(MSE(\hat{\beta}^{OLSC}))}$	0.27 _(0.915)	-1.27 _(0.915)	2.34 _(1.848)	- 3 .34 _(1.848)	3.62 _(3.488)
$\hat{\beta}^{LWSC}_{(MSE(\hat{\beta}^{LWSC}))}$	1.00 _(0.005)	-2.00 _(0.005)	3.00 _(0.026)	-4.00(0.026)	5.00 _(0.021)
$\hat{\beta}^{LTSC}_{(MSE(\hat{\beta}^{LTSC}))}$	1.00 _(0.006)	-2.00 _(0.006)	3.01 _(0.030)	-4.01 _(0.030)	5.00 _(0.024)

... broken orthogonality condition ... collinearity

What about leverage points.

TABLE 5

Nearly the same framework as in previous cases - except of underlined.

There is a collinearity,

the mean value of condition numbers was equal to 16.55.

... broken orthogonality condition ... collinearity

What about leverage points.

TABLE 5

Nearly the same framework as in previous cases - except of underlined.

There is a collinearity,

the mean value of condition numbers was equal to 16.55.

2 The 5% contamination by leverage points

 $X^{(contaminated)} = 3 * X^{(original)}, Y^{(contaminated)} = -2 * Y^{(original)}.$

... broken orthogonality condition ... collinearity

$\hat{\beta}^{OLS}_{(MSE(\hat{\beta}^{OLS}))}$	-1.59 _(45.600)	0.68 _(81.603)	0.68 _(45.803)	-1.53 _(155.949)	0.50 _(169.327)
$\hat{\beta}^{LWS}_{(MSE(\hat{\beta}^{LWS}))}$	0.99 _(0.162)	-2.01 _(0.312)	3.01 _(0.163)	-4.01 _(0.634)	5.01 _(0.622)
$\hat{\beta}^{LTS}_{(MSE(\hat{\beta}^{LTS}))}$	1.00 _(0.190)	-2.00 _(0.323)	3.01 _(0.201)	-4.01 _(0.745)	4.99 _(0.717)
$\hat{\beta}^{OLSC}$ (MSE($\hat{\beta}^{OLSC}$))	-1.86 _(10.474)	0.86 _(10.474)	0.70 _(13.424)	-1.70 _(13.424)	0.52 _(27.070)
$\hat{\beta}^{LWSC}$ (MSE($\hat{\beta}^{LWSC}$))	1.00 _(0.005)	-2.00 _(0.005)	3.00 _(0.026)	-4.00 _(0.026)	5.00 _(0.021)
$\hat{\beta}^{LTSC}_{(MSE(\hat{\beta}^{LTSC}))}$		-2.00 _(0.007)	3.01 _(0.037)	-4.01 _(0.037)	4.99 _(0.028)

... broken orthogonality condition ... collinearity

$\hat{\beta}^{OLS}_{(MSE(\hat{\beta}^{OLS}))}$	-1.59 _(45.600)	0.68 _(81.603)	0.68 _(45.803)	-1.53 _(155.949)	0.50 _(169.327)
$\hat{\beta}^{LWS}_{(MSE(\hat{\beta}^{LWS}))}$	0.99 _(0.162)	-2.01 _(0.312)	3.01 _(0.163)	-4.01 _(0.634)	5.01 _(0.622)
$\hat{\beta}^{LTS}_{(MSE(\hat{\beta}^{LTS}))}$	1.00 _(0.190)	-2.00 _(0.323)	3.01 _(0.201)	-4.01 _(0.745)	4.99 _(0.717)
$\hat{\beta}^{OLSC}$ (MSE($\hat{\beta}^{OLSC}$))	-1.86 _(10.474)	0.86 _(10.474)	0.70 _(13.424)	-1.70 _(13.424)	0.52 _(27.070)
$\hat{\beta}^{LWSC}$ (MSE($\hat{\beta}^{LWSC}$))	1.00 _(0.005)	-2.00 _(0.005)	3.00 _(0.026)	-4.00 _(0.026)	5.00 _(0.021)
$\hat{\beta}^{LTSC}_{(MSE(\hat{\beta}^{LTSC}))}$		-2.00 _(0.007)	3.01 _(0.037)	-4.01 _(0.037)	4.99 _(0.028)

... broken orthogonality condition ... collinearity

What about leverage points.

TABLE 6

The same framework as in previous cases - only the mean value of condition numbers was equal to <u>32.60</u>.

$\hat{\beta}^{OLS}_{(MSE(\hat{\beta}^{OLS}))}$	-1.67 _(145.901)	0.47 _(268.296)	0.98(139.653)	-2.16 _(536.245)	1.23 _(570.985)
$\hat{\beta}^{LWS}_{(MSE(\hat{\beta}^{LWS}))}$	0.94 _(0.625)	-2.06 _(1.216)	3.00 _(0.593)	-4.00 _(2.335)	5.12 _(2.471)
$\hat{\beta}^{LTS}$ (MSE($\hat{\beta}^{LTS}$))	0.94 _(0.686)	-2.07 _(1.260)	2.99 _(0.687)	-3.98 _(2.658)	5.11 _(2.672)
$\hat{\beta}^{OLSC}_{(MSE(\hat{\beta}^{OLSC}))}$	-1.61 _(9.285)	0.61 _(9.285)	0.69(13.963)	-1.69 _(13.963)	0.65 _(26.999)
$\hat{\beta}^{LWSC}$ (MSE($\hat{\beta}^{LWSC}$))		-2.00 _(0.005)	3.01 _(0.025)	-4.01 _(0.025)	5.00 _(0.019)
$\hat{\beta}^{LTSC}_{(MSE(\hat{\beta}^{LTSC}))}$		-2.00 _(0.009)	3.00 _(0.045)	-4.00 _(0.045)	5.00 _(0.043)

... broken orthogonality condition ... collinearity

What about leverage points.

TABLE 6

The same framework as in previous cases - only the mean value of condition numbers was equal to <u>32.60</u>.

$\hat{\beta}^{OLS}_{(MSE(\hat{\beta}^{OLS}))}$	-1.67 _(145.901)	0.47 _(268.296)	0.98(139.653)	-2.16 _(536.245)	1.23 _(570.985)
$\hat{\beta}^{LWS}_{(MSE(\hat{\beta}^{LWS}))}$	0.94 _(0.625)	-2.06 _(1.216)	3.00 _(0.593)	-4.00 _(2.335)	5.12 _(2.471)
$\hat{\beta}^{LTS}$ (MSE($\hat{\beta}^{LTS}$))	0.94 _(0.686)	-2.07 _(1.260)	2.99 _(0.687)	-3.98 _(2.658)	5.11 _(2.672)
$\hat{\beta}^{OLSC}_{(MSE(\hat{\beta}^{OLSC}))}$	-1.61 _(9.285)	0.61 _(9.285)	0.69(13.963)	-1.69 _(13.963)	0.65 _(26.999)
$\hat{\beta}^{LWSC}$ (MSE($\hat{\beta}^{LWSC}$))		-2.00 _(0.005)	3.01 _(0.025)	-4.01 _(0.025)	5.00 _(0.019)
$\hat{\beta}^{LTSC}_{(MSE(\hat{\beta}^{LTSC}))}$		-2.00 _(0.009)	3.00 _(0.045)	-4.00 _(0.045)	5.00 _(0.043)

... for LWS-estimation ... simulations of *p*-values

Content

- Disqualitying classical regression analysis by ignoring ...
 - ... broken orthogonality condition
 - ... collinearity
- 2 Significance of individual explanatory variable
 - In for LWS-estimation
 - ... simulations of p-values

... for LWS-estimation ... simulations of *p*-values

Content

- Disqualitying classical regression analysis by ignoring ...
 - ... broken orthogonality condition
 - ... collinearity
- 2 Significance of individual explanatory variable
 - In for LWS-estimation
 - ... simulations of p-values

... for LWS-estimation ... simulations of *p*-values

Significance of explanatory variable - for the Least Weighted Squares

We are going to give an idea of deriving the significance of *individual regressor* - two steps:



... for LWS-estimation ... simulations of *p*-values

Significance of explanatory variable - for the Least Weighted Squares

We are going to give an idea of deriving the significance of *individual regressor* - two steps:

The first one (which we have already seen in the seventh lecture

- hence only a brief repetition):

The Least Weighted Squares $\hat{\beta}^{(LWS,n,W)}$ can be - at any point of a basic probabily space (Ω, \mathcal{A}, P) written as Weighted Least Squares $\hat{\beta}^{(WLS,n,W,\pi)}$.



... for LWS-estimation ... simulations of *p*-values

Significance of explanatory variable - for the Least Weighted Squares

We are going to give an idea of deriving the significance of *individual regressor* - two steps:

The first one (which we have already seen in the seventh lecture

hence only a brief repetition):

The Least Weighted Squares $\hat{\beta}^{(LWS,n,W)}$ can be - at any point of a basic probabily space (Ω, \mathcal{A}, P) written as Weighted Least Squares $\hat{\beta}^{(WLS,n,W,\pi)}$.

The second one:

The classical derivation for significance of *individual regressor for OLS* can be generalised for the *classical WLS* $\hat{\beta}^{(WLS,n,W,\pi)}$.

... for LWS-estimation ... simulations of *p*-values

Showing that $\hat{\beta}^{(LWS,n)}$ is $\hat{\beta}^{(LWS,n,\hat{\pi})}$...

We have seen in the seventh lecture:

 $\frac{\forall \ (\omega \in \Omega) \ \exists \ (\pi = \pi(\omega) = \{\pi_1(\omega), \pi_2(\omega), ..., \pi_n(\omega)\})}{\hat{\beta}^{(LWS, n, W)}(\omega) = \arg \min_{\beta \in R^p} \sum_{i=1}^n w_i \left(Y_{\pi_i} - X'_{\pi_i}\beta\right)^2 = \hat{\beta}^{(WLS, n, W, \pi)}(\omega)$ $= \arg \min_{\beta \in R^p} \sum_{i=1}^n \left(w_i^{\frac{1}{2}}Y_{\pi_i} - w_i^{\frac{1}{2}}X'_{\pi_i}\beta\right)^2.$

... for LWS-estimation ... simulations of *p*-values

Showing that $\hat{\beta}^{(LWS,n)}$ is $\hat{\beta}^{(LWS,n,\hat{\pi})}$...

We have seen in the seventh lecture:

$$\frac{\forall \ (\omega \in \Omega) \ \exists \ (\pi = \pi(\omega) = \{\pi_1(\omega), \pi_2(\omega), ..., \pi_n(\omega)\}) \text{ so that}}{\hat{\beta}^{(LWS, n, W)}(\omega) = \underset{\beta \in R^p}{\operatorname{arg\,min}} \sum_{i=1}^n w_i \left(Y_{\pi_i} - X'_{\pi_i}\beta\right)^2 = \hat{\beta}^{(WLS, n, W, \pi)}(\omega)$$
$$= \underset{\beta \in R^p}{\operatorname{arg\,min}} \sum_{i=1}^n \left(w_i^{\frac{1}{2}}Y_{\pi_i} - w_i^{\frac{1}{2}}X'_{\pi_i}\beta\right)^2.$$

Notice the dependence of π on ω .

... for LWS-estimation ... simulations of *p*-values

Showing that $\hat{\beta}^{(LWS,n)}$ is $\hat{\beta}^{(LWS,n,\hat{\pi})}$...

We have seen in the seventh lecture:

$$\frac{\forall (\omega \in \Omega) \exists (\pi = \pi(\omega) = \{\pi_1(\omega), \pi_2(\omega), ..., \pi_n(\omega)\})}{\hat{\beta}^{(LWS, n, W)}(\omega) = \underset{\beta \in R^p}{\operatorname{arg\,min}} \sum_{i=1}^n w_i (Y_{\pi_i} - X'_{\pi_i}\beta)^2 = \hat{\beta}^{(WLS, n, W, \pi)}(\omega)$$
$$= \underset{\beta \in R^p}{\operatorname{arg\,min}} \sum_{i=1}^n \left(w_i^{\frac{1}{2}}Y_{\pi_i} - w_i^{\frac{1}{2}}X'_{\pi_i}\beta\right)^2.$$

Notice the dependence of π on ω .

Let's recall how we have found it.

... for LWS-estimation ... simulations of *p*-values

Recalling several facts

Let \mathcal{P} be the set of all permutations of integers $\{1, 2, ..., n\}$.



... for LWS-estimation ... simulations of *p*-values

Recalling several facts

Let \mathcal{P} be the set of all permutations of integers $\{1, 2, ..., n\}$.

For any $\pi \in \mathcal{P}, \pi = \{\pi_1, \pi_2, ..., \pi_n\}$ let

$$Y_{\pi} = (Y_{\pi_1}, Y_{\pi_2}, ..., Y_{\pi_n})', X_{\pi} = (X_{\pi_1}, X_{\pi_2}, ..., X_{\pi_n})' \text{ and } \varepsilon_{\pi} = (\varepsilon_{\pi_1}, \varepsilon_{\pi_2}, ..., \varepsilon_{\pi_n})'.$$

... for LWS-estimation

Recalling several facts

Let \mathcal{P} be the set of all permutations of integers $\{1, 2, ..., n\}$.

For any $\pi \in \mathcal{P}, \pi = \{\pi_1, \pi_2, ..., \pi_n\}$ let

$$Y_{\pi} = (Y_{\pi_1}, Y_{\pi_2}, ..., Y_{\pi_n})', X_{\pi} = (X_{\pi_1}, X_{\pi_2}, ..., X_{\pi_n})' \text{ and } \varepsilon_{\pi} = (\varepsilon_{\pi_1}, \varepsilon_{\pi_2}, ..., \varepsilon_{\pi_n})'.$$

Put

 $\hat{\beta}^{(WLS,n,\pi)} = (X'_{\pi}WX_{\pi})^{-1}X'_{\pi}WY_{\pi} \text{ and } S^{2}_{\pi} = \sum_{i=1}^{n} w_{i}\left(Y_{\pi_{i}} - X'_{\pi_{i}}\hat{\beta}^{(WLS,n,\pi)}\right)^{2}.$

... for LWS-estimation ... simulations of *p*-values

Deriving existence of $\hat{\beta}^{(LWS,n)}$

Then for any $\pi \in \mathcal{P}$

$$S_{\pi}^{2} = \sum_{j=1}^{n} w_{j} \left(Y_{\pi_{j}} - X_{\pi_{j}}^{\prime} \hat{\beta}^{(WLS,n,\pi)} \right)^{2} \leq \min_{\beta \in \mathbb{R}^{p}} \sum_{j=1}^{n} w_{j} \left(Y_{\pi_{j}} - X_{\pi_{j}}^{\prime} \beta \right)^{2}.$$
 (3)

... for LWS-estimation ... simulations of *p*-values

Deriving existence of $\hat{\beta}^{(LWS,n)}$

Then for any $\pi \in \mathcal{P}$

$$S_{\pi}^{2} = \sum_{j=1}^{n} w_{j} \left(Y_{\pi_{j}} - X_{\pi_{j}}^{\prime} \hat{\beta}^{(WLS,n,\pi)} \right)^{2} \leq \min_{\beta \in R^{p}} \sum_{j=1}^{n} w_{j} \left(Y_{\pi_{j}} - X_{\pi_{j}}^{\prime} \beta \right)^{2}.$$
 (3)

Let (Ω, \mathcal{A}, P) be the basic *Probability Space*

and denote for any r.v. Z by $Z(\omega)$ its value at point $\omega \in \Omega$.

... for LWS-estimation ... simulations of *p*-values

Deriving existence of $\hat{\beta}^{(LWS,n)}$

Then for any $\pi \in \mathcal{P}$

$$S_{\pi}^{2} = \sum_{j=1}^{n} w_{j} \left(Y_{\pi_{j}} - X_{\pi_{j}}^{\prime} \hat{\beta}^{(WLS,n,\pi)} \right)^{2} \leq \min_{\beta \in R^{p}} \sum_{j=1}^{n} w_{j} \left(Y_{\pi_{j}} - X_{\pi_{j}}^{\prime} \beta \right)^{2}.$$
 (3)

Let (Ω, \mathcal{A}, P) be the basic <u>Probability Space</u> and denote for any r. v. Z by $Z(\omega)$ its value at point $\omega \in \Omega$. Finally, fix ω_0 and put

 $\hat{\pi}(\omega_0) = \underset{\pi \in \mathcal{P}}{\operatorname{arg\,min}} S^2_{\pi}(\omega_0).$

... for LWS-estimation ... simulations of *p*-values

Deriving existence of $\hat{\beta}^{(LWS,n)}$

Then for any $\pi \in \mathcal{P}$

$$S_{\pi}^{2} = \sum_{j=1}^{n} w_{j} \left(Y_{\pi_{j}} - X_{\pi_{j}}^{\prime} \hat{\beta}^{(WLS,n,\pi)} \right)^{2} \leq \min_{\beta \in R^{p}} \sum_{j=1}^{n} w_{j} \left(Y_{\pi_{j}} - X_{\pi_{j}}^{\prime} \beta \right)^{2}.$$
 (3)

Let (Ω, \mathcal{A}, P) be the basic <u>Probability Space</u> and denote for any r. v. Z by $Z(\omega)$ its value at point $\omega \in \Omega$. Finally, fix ω_0 and put

$$\hat{\pi}(\omega_0) = \underset{\pi \in \mathcal{P}}{\operatorname{arg\,min}} S_{\pi}^2(\omega_0).$$

Then - due to (3) $S^{2}_{\hat{\pi}(\omega_{0})}(\omega_{0}) \leq \min_{\pi \in \mathcal{P}} \min_{\beta \in \mathbb{R}^{p}} \sum_{j=1}^{n} w_{j} \left(Y_{\pi_{j}} - X'_{\pi_{j}}\beta \right)^{2} = \min_{\beta \in \mathbb{R}^{p}} \min_{\pi \in \mathcal{P}} \sum_{j=1}^{n} w_{j} \left(Y_{\pi_{j}} - X'_{\pi_{j}}\beta \right)^{2}.$

... for LWS-estimation ... simulations of *p*-values

Deriving existence and form of $\hat{\beta}^{(LWS,n)}$

It means

$$\hat{\beta}^{(WLS,n,\hat{\pi}(\omega_0))}(\omega_0) = \hat{\beta}^{(WLS,n)}\left(Y_{\pi(\omega_0)}(\omega_0), X_{\pi(\omega_0)}(\omega_0)\right) = \hat{\beta}^{(LWS,n,w)}(\omega_0)$$



... for LWS-estimation ... simulations of *p*-values

Deriving existence and form of $\hat{\beta}^{(LWS,n)}$

It means

$$\hat{\beta}^{(\mathsf{WLS},n,\hat{\pi}(\omega_0))}(\omega_0) = \hat{\beta}^{(\mathsf{WLS},n)}\left(Y_{\pi(\omega_0)}(\omega_0), X_{\pi(\omega_0)}(\omega_0)\right) = \hat{\beta}^{(\mathsf{LWS},n,w)}(\omega_0)$$

Repeating it for all $\omega \in \Omega$, we prove the existence of $\hat{\beta}^{(LWS,n,w)}$.



... for LWS-estimation ... simulations of *p*-values

Deriving existence and form of $\hat{\beta}^{(LWS,n)}$

It means

$$\hat{\beta}^{(WLS,n,\hat{\pi}(\omega_0))}(\omega_0) = \hat{\beta}^{(WLS,n)}\left(Y_{\pi(\omega_0)}(\omega_0), X_{\pi(\omega_0)}(\omega_0)\right) = \hat{\beta}^{(LWS,n,w)}(\omega_0)$$

Repeating it for all $\omega \in \Omega$, we prove the existence of $\hat{\beta}^{(LWS,n,w)}$.

Fix $\pi \in \mathcal{P}$ and put $B(\pi) = \{\omega \in \Omega : \pi = \hat{\pi}(\omega)\}.$

... for LWS-estimation ... simulations of *p*-values

Deriving existence and form of $\hat{\beta}^{(LWS,n)}$

It means

$$\hat{\beta}^{(WLS,n,\hat{\pi}(\omega_0))}(\omega_0) = \hat{\beta}^{(WLS,n)}\left(Y_{\pi(\omega_0)}(\omega_0), X_{\pi(\omega_0)}(\omega_0)\right) = \hat{\beta}^{(LWS,n,w)}(\omega_0)$$

Repeating it for all $\omega \in \Omega$, we prove the existence of $\hat{\beta}^{(LWS,n,w)}$.

Fix $\pi \in \mathcal{P}$ and put $B(\pi) = \{\omega \in \Omega : \pi = \hat{\pi}(\omega)\}.$ Then $\omega \in B(\pi) \Rightarrow \hat{\beta}^{(LWS,n,W)}(\omega) = \hat{\beta}^{(WLS,n,\pi)}(\omega)$ i. e. $\hat{\beta}^{(LWS,n,W)}(\omega) = \hat{\beta}^{(WLS,n,\pi)}(\omega) = \arg\min_{\beta \in R^{p}} \sum_{j=1}^{n} w \left(\frac{j-1}{n}\right) \left(Y_{\pi_{j}} - X'_{\pi_{j}}\beta\right)^{2}.$

... for LWS-estimation ... simulations of *p*-values

Deriving existence and form of $\hat{\beta}^{(LWS,n)}$

It means

$$\hat{\beta}^{(WLS,n,\hat{\pi}(\omega_0))}(\omega_0) = \hat{\beta}^{(WLS,n)}\left(Y_{\pi(\omega_0)}(\omega_0), X_{\pi(\omega_0)}(\omega_0)\right) = \hat{\beta}^{(LWS,n,w)}(\omega_0)$$

Repeating it for all $\omega \in \Omega$, we prove the existence of $\hat{\beta}^{(LWS,n,w)}$.

Fix $\pi \in \mathcal{P}$ and put $B(\pi) = \{\omega \in \Omega : \pi = \hat{\pi}(\omega)\}.$ Then $\omega \in B(\pi) \Rightarrow \hat{\beta}^{(LWS,n,w)}(\omega) = \hat{\beta}^{(WLS,n,\pi)}(\omega)$ i.e. $\hat{\beta}^{(LWS,n,w)}(\omega) = \hat{\beta}^{(WLS,n,\pi)}(\omega) = \arg\min_{\beta \in R^{\rho}} \sum_{j=1}^{n} w \left(\frac{j-1}{n}\right) \left(Y_{\pi_{j}} - X'_{\pi_{j}}\beta\right)^{2}.$ Due to i.i.d. framework, $\forall \pi \in \mathcal{P}$ $P(B(\pi)) = (n!)^{-1}.$

... for LWS-estimation ... simulations of *p*-values

Deriving form of $\hat{\beta}^{(LWS,n)}$

Let's repeat: We have shown that fixing $\pi \in \mathcal{P}$ and putting $B(\pi) = \{\omega \in \Omega : \pi = \pi(\omega)\}$, we have

 $P(B(\pi)) = (n!)^{-1}$

... for LWS-estimation ... simulations of *p*-values

Deriving form of $\hat{\beta}^{(LWS,n)}$

Let's repeat: We have shown that fixing $\pi \in \mathcal{P}$ and putting $B(\pi) = \{\omega \in \Omega : \pi = \pi(\omega)\}$, we have

$$P(B(\pi)) = (n!)^{-1}$$

and one can easy verify that

 $\cup_{\pi\in\mathcal{P}} B(\pi) = \Omega.$

... for LWS-estimation ... simulations of *p*-values

Deriving form of $\hat{\beta}^{(LWS,n)}$

Let's repeat: We have shown that fixing $\pi \in \mathcal{P}$ and putting $B(\pi) = \{\omega \in \Omega : \pi = \pi(\omega)\}$, we have

$$P(B(\pi)) = (n!)^{-1}$$

and one can easy verify that

 $\cup_{\pi\in\mathcal{P}} B(\pi) = \Omega.$

All further considerations can be done conditionally, on π (or on $B(\pi)$, if You want) for $\hat{\beta}^{(WLS,n,\pi)}$.

... for LWS-estimation ... simulations of *p*-values

Deriving form of $\hat{\beta}^{(LWS,n)}$

Let's repeat: We have shown that fixing $\pi \in \mathcal{P}$ and putting $B(\pi) = \{\omega \in \Omega : \pi = \pi(\omega)\}$, we have

$$P(B(\pi)) = (n!)^{-1}$$

and one can easy verify that

 $\cup_{\pi\in\mathcal{P}} B(\pi) = \Omega.$

All further considerations can be done conditionally, on π (or on $B(\pi)$, if You want) for $\hat{\beta}^{(WLS,n,\pi)}$.

Then we take mean value over all conditions, i. e. over all $\pi \in \mathcal{P}$ but the situation for all π is the same, with the same probabilities, hence an unconditional result is the same as conditional.

... for LWS-estimation ... simulations of *p*-values

Deriving existence and form of $\hat{\beta}^{(LWS,n)}$

Let's rewrite the line (the third from bottom) from the last but one slide, for $\omega \in B(\pi)$

$$\hat{\beta}^{(LWS,n,w)}(\omega) = \hat{\beta}^{(WLS,n,\pi)}(\omega) = \arg\min_{\beta \in \mathbb{R}^{p}} \sum_{j=1}^{\infty} w \left(\frac{I-1}{n}\right) \left(Y_{\pi_{j}} - X'_{\pi_{j}}\beta\right)^{2}.$$
 (4)
... for LWS-estimation ... simulations of *p*-values

Deriving existence and form of $\hat{\beta}^{(LWS,n)}$

Let's rewrite the line (the third from bottom) from the last but one slide, for $\omega \in B(\pi)$

 $\hat{\beta}^{(LWS,n,w)}(\omega) = \hat{\beta}^{(WLS,n,\pi)}(\omega) = \arg\min_{\beta \in R^{\rho}} \sum_{j=1}^{n} w\left(\frac{j-1}{n}\right) \left(Y_{\pi_{j}} - X'_{\pi_{j}}\beta\right)^{2}.$ (4) Putting $\tilde{W} = diag\left\{w^{-\frac{1}{2}}(0), w^{-\frac{1}{2}}\left(\frac{1}{n}\right), ..., w^{-\frac{1}{2}}\left(\frac{n-1}{n}\right)\right\}$ and $\tilde{Y} = \tilde{W}Y, \quad \tilde{X} = \tilde{W}X$, then (4) reads

 $\hat{\beta}^{(LWS,n,w)}(\omega) = \underset{\beta \in \mathbb{R}^{p}}{\operatorname{arg\,min}} \left\{ \left(\tilde{Y} - \tilde{X}\beta \right)' \left(\tilde{Y} - \tilde{X}\beta \right) \right\} \\ = \left(\tilde{X}'\tilde{X} \right)^{-1} \tilde{X}'\tilde{Y} = \hat{\beta}^{(OLS,n)}(\tilde{Y},\tilde{X}).$

... for LWS-estimation ... simulations of *p*-values

Deriving existence and form of $\hat{\beta}^{(LWS,n)}$

Let's rewrite the line (the third from bottom) from the last but one slide, for $\omega \in B(\pi)$

$$\hat{\beta}^{(LWS,n,w)}(\omega) = \hat{\beta}^{(WLS,n,\pi)}(\omega) = \arg\min_{\beta \in \mathbb{R}^{p}} \sum_{j=1}^{n} w\left(\frac{j-1}{n}\right) \left(Y_{\pi_{j}} - X'_{\pi_{j}}\beta\right)^{2}.$$
(4)
Putting $\tilde{W} = diag\left\{w^{-\frac{1}{2}}(0), w^{-\frac{1}{2}}\left(\frac{1}{n}\right), ..., w^{-\frac{1}{2}}\left(\frac{n-1}{n}\right)\right\}$
and $\tilde{Y} = \tilde{W}Y, \quad \tilde{X} = \tilde{W}X$, then (4) reads

$$\hat{\beta}^{(LWS,n,w)}(\omega) = \underset{\beta \in R^{p}}{\operatorname{arg\,min}} \left\{ \left(\tilde{Y} - \tilde{X}\beta \right)' \left(\tilde{Y} - \tilde{X}\beta \right) \right\} \\ = \left(\tilde{X}'\tilde{X} \right)^{-1} \tilde{X}'\tilde{Y} = \hat{\beta}^{(OLS,n)}(\tilde{Y},\tilde{X}).$$

Consider, for a while, the model

 $ilde{Y} = ilde{X} eta^0 + ilde{arepsilon} \quad ext{with} \quad \mathcal{L}\left(ilde{arepsilon}\right) = \mathcal{N}\left(0, \sigma^2 ilde{W}^2\right) \quad ext{(notice the heteroscedascity of ε)}.$

... for LWS-estimation ... simulations of *p*-values

Significance of explanatory variable - classical OLS case

Let's recall the simplest classical framework for finite-sample diagnostics:

Regression model

 $Y_i = X'_i \beta^0 + \varepsilon_i, \quad i = 1, 2, ..., n \quad \text{or} \quad Y = X \beta^0 + \varepsilon$

... for LWS-estimation ... simulations of *p*-values

Significance of explanatory variable - classical OLS case

Let's recall the simplest classical framework for finite-sample diagnostics:

Regression model

 $Y_i = X'_i \beta^0 + \varepsilon_i, \quad i = 1, 2, ..., n \quad \text{or} \quad Y = X \beta^0 + \varepsilon$

Conditions :

 $\{(X'_{i},\varepsilon_{i})'\}_{i=1}^{\infty} \underline{\text{i.i.d.}}, F_{X,\varepsilon}(x,v) = F_{X}(x) \cdot F_{\varepsilon}(v), \underline{F_{\varepsilon}(v)} = \mathcal{N}(0,\sigma^{2}), \underline{Q} = E[X_{1} \cdot X_{1}'] \text{ is regular.}$

... for LWS-estimation ... simulations of *p*-values

Significance of explanatory variable - classical OLS case

Let's recall the simplest classical framework for finite-sample diagnostics:

Regression model

 $Y_i = X'_i \beta^0 + \varepsilon_i, \quad i = 1, 2, ..., n \quad \text{or} \quad Y = X \beta^0 + \varepsilon$

Conditions :

$$\{(X'_i,\varepsilon_i)'\}_{i=1}^{\infty} \text{ i.i.d.}, F_{X,\varepsilon}(X,V) = F_X(X) \cdot F_{\varepsilon}(V), F_{\varepsilon}(V) = \mathcal{N}(0,\sigma^2), \\ Q = E[X_1 \cdot X'_1] \text{ is regular.}$$

Significance of ℓ -th explanatory variable $X_{i\ell} \Leftrightarrow \neg H_0$: $\hat{\beta}_{\ell}^{(OLS,n)} = 0$

Denote

$$c_{\ell,\ell}^2 = |(X'X)^{-1}|_{\ell,\ell}$$
 and

nd
$$S_n^2 = \frac{1}{n-p}$$

$$\sum_{i=1}^{n} \left(Y_i - X_i' \hat{\beta}^{(OLS)} \right)$$

... for LWS-estimation

Significance of explanatory variable - classical OLS case

Let's recall the simplest classical framework for finite-sample diagnostics:

Regression model

 $Y_i = X'_i \beta^0 + \varepsilon_i, \quad i = 1, 2, ..., n \quad \text{or} \quad Y = X \beta^0 + \varepsilon$

Conditions :

$$\{(X'_i,\varepsilon_i)'\}_{i=1}^{\infty} \text{ i.i.d.}, F_{X,\varepsilon}(X,V) = F_X(X) \cdot F_{\varepsilon}(V), \frac{F_{\varepsilon}(V) = \mathcal{N}(0,\sigma^2)}{Q = E[X_1 \cdot X'_1] \text{ is regular.}}$$

Significance of ℓ -th explanatory variable $X_{i\ell} \Leftrightarrow \neg H_0$: $\hat{\beta}_{\ell}^{(OLS,n)} = 0$

Denote

Denote
$$c_{\ell,\ell}^2 = \left[(X'X)^{-1} \right]_{\ell,\ell}$$
 and $s_n^2 = \frac{1}{n-p} \sum_{i=1}^n \left(Y_i - X'_i \hat{\beta}^{(OLS,n)} \right)^2$.
Test is based on
 $\mathcal{L} \left(\frac{\hat{\beta}_{\ell}^{(OLS,n)} - \beta_{\ell}^0}{\mathbf{s}_n \cdot \mathbf{c}_{\ell,\ell}} \right) = \mathcal{L} \left(t_{\ell} \right) = t_{n-p}$ (Fisher-Cochran theorem)

... for LWS-estimation ... simulations of *p*-values

Recalling the derivation of *t*-statistics in the classical regression

First of all, let's recall that in any regression model and hence also in our model for $\left(\, \tilde{Y}, ilde{X} \,
ight)$ $\tilde{r}\left(\hat{\beta}^{(OLS,n)}\left(\tilde{Y},\tilde{X}\right)\right) = \tilde{Y} - \tilde{X}\hat{\beta}^{(OLS,n)}\left(\tilde{Y},\tilde{X}\right) \perp \mathcal{M}\left(\tilde{X}\right)$ $(\mathcal{M}(\tilde{X})$ is the set of all linear combinations of the columns of \tilde{X}).

... for LWS-estimation ... simulations of *p*-values

Recalling the derivation of *t*-statistics in the classical regression

First of all, let's recall that in any regression model and hence also in our model for $ig(ilde{Y}, ilde{X}ig)$ $\tilde{r}\left(\hat{\beta}^{(OLS,n)}\left(\tilde{Y},\tilde{X}\right)\right) = \tilde{Y} - \tilde{X}\hat{\beta}^{(OLS,n)}\left(\tilde{Y},\tilde{X}\right) \perp \mathcal{M}\left(\tilde{X}\right)$ $(\mathcal{M}(\tilde{X})$ is the set of all linear combinations of the columns of \tilde{X}). Secondly, $\widehat{\widetilde{Y}} = \widetilde{X}\widehat{eta}^{(OLS,n)} = \widetilde{X}\left(\widetilde{X}'\widetilde{X}\right)^{-1}\widetilde{X}'\widetilde{Y}$
$$\begin{split} & = \tilde{X} \left(\tilde{X}' \tilde{X} \right)^{-1} \tilde{X}' \left(\tilde{X} \beta^{0} + \tilde{\varepsilon} \right) = \tilde{X} \beta^{0} + \tilde{X} \left(\tilde{X}' \tilde{X} \right)^{-1} \tilde{X}' \tilde{\varepsilon}, \\ & \mathcal{L} \left(\widehat{Y} \right) = \mathcal{N} \left(\tilde{X} \beta^{0}, \sigma^{2} \tilde{X} \left(\tilde{X}' \tilde{X} \right)^{-1} \tilde{X}' \tilde{W}^{2} \tilde{X} \left(\tilde{X}' \tilde{X} \right)^{-1} \tilde{X}' \right). \end{split}$$
i. e.

... for LWS-estimation ... simulations of *p*-values

Recalling the derivation of *t*-statistics in the classical regression

As $\hat{\tilde{Y}} = \tilde{X}\hat{\beta}^{(OLS,n)}$, i.e. $\hat{\tilde{Y}}$ is a linear combination of the columns of \tilde{X} $\Rightarrow \hat{\tilde{Y}} \perp \tilde{r}\left(\hat{\beta}^{(OLS,n)}\left(\tilde{Y},\tilde{X}\right)\right) = \tilde{Y} - \tilde{X}\hat{\beta}^{(OLS,n)}\left(\tilde{Y},\tilde{X}\right)$

and due to the normality of $\hat{\tilde{Y}}$, it is independent with $r\left(\hat{\beta}^{(OLS,n)}\left(\tilde{Y},\tilde{X}\right)\right)$.

... for LWS-estimation ... simulations of *p*-values

Recalling the derivation of *t*-statistics in the classical regression

As $\hat{\tilde{Y}} = \tilde{X}\hat{\beta}^{(OLS,n)}$, i. e. $\hat{\tilde{Y}}$ is a linear combination of the columns of \tilde{X} $\Rightarrow \hat{\tilde{Y}} \perp \tilde{r}\left(\hat{\beta}^{(OLS,n)}\left(\tilde{Y},\tilde{X}\right)\right) = \tilde{Y} - \tilde{X}\hat{\beta}^{(OLS,n)}\left(\tilde{Y},\tilde{X}\right)$ and due to the normality of $\hat{\tilde{Y}}$, it is independent with $r\left(\hat{\beta}^{(OLS,n)}\left(\tilde{Y},\tilde{X}\right)\right)$. By standard way

$$\hat{\beta}^{(OLS,n)}(\tilde{Y},\tilde{X})' = (\tilde{X}'\tilde{X})^{-1}\tilde{X}\tilde{Y} = \left[(\tilde{X}'\tilde{X})^{-1}\tilde{X}'\tilde{X} \right] (\tilde{X}'\tilde{X})^{-1}\tilde{X}\tilde{Y}$$
$$= (\tilde{X}'\tilde{X})^{-1}\tilde{X}' \left[\tilde{X} (\tilde{X}'\tilde{X})^{-1}\tilde{X}\tilde{Y} \right] = (\tilde{X}'\tilde{X})^{-1}\tilde{X}'\hat{\tilde{Y}}$$
(5)

... for LWS-estimation ... simulations of *p*-values

Recalling the derivation of *t*-statistics in the classical regression

As $\hat{\tilde{Y}} = \tilde{X}\hat{\beta}^{(OLS,n)}$, i. e. $\hat{\tilde{Y}}$ is a linear combination of the columns of \tilde{X} $\Rightarrow \hat{\tilde{Y}} \perp \tilde{r}\left(\hat{\beta}^{(OLS,n)}\left(\tilde{Y},\tilde{X}\right)\right) = \tilde{Y} - \tilde{X}\hat{\beta}^{(OLS,n)}\left(\tilde{Y},\tilde{X}\right)$ and due to the normality of $\hat{\tilde{Y}}$, it is independent with $r\left(\hat{\beta}^{(OLS,n)}\left(\tilde{Y},\tilde{X}\right)\right)$.

By standard way

$$\hat{\beta}^{(OLS,n)}(\tilde{Y},\tilde{X})' = \left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}\tilde{Y} = \left[\left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}'\tilde{X}\right]\left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}\tilde{Y}$$
$$= \left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}'\left[\tilde{X}\left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}\tilde{Y}\right] = \left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}'\hat{\tilde{Y}}$$
(5)

Then due to (5) $\hat{\beta}^{(OLS,n)}(\tilde{Y},\tilde{X})$ is independent with $\tilde{r}(\hat{\beta}^{(OLS,n)}(\tilde{Y},\tilde{X}))$ (we shall need it in a few minutes).

... for LWS-estimation ... simulations of *p*-values

Recalling the derivation of *t*-statistics in the classical regression

Let's recall

 $\hat{\beta}^{(OLS,n)}\left(\tilde{Y},\tilde{X}\right) - \beta^{0} = \hat{\beta}^{(WLS,n,W,\pi)} - \beta^{0} = \left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}'\tilde{\varepsilon},$

i. e. due to normality of disturbances, also $\hat{\beta}^{(WLS,n,W,\pi)} - \beta^0$ is normally distributed (with heteroscadasticity).

... for LWS-estimation ... simulations of *p*-values

Recalling the derivation of *t*-statistics in the classical regression

Let's recall

$$\hat{\beta}^{(OLS,n)}\left(\tilde{Y},\tilde{X}\right) - \beta^{0} = \hat{\beta}^{(WLS,n,W,\pi)} - \beta^{0} = \left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}'\tilde{\varepsilon},$$

i. e. due to normality of disturbances, also $\hat{\beta}^{(WLS,n,W,\pi)} - \beta^0$ is normally distributed (with heteroscadasticity). Let's recall (as it is well-known from the classical regression)

$$\tilde{r}\left(\hat{\beta}^{(OLS,n)}\left(\tilde{Y},\tilde{X}\right)\right) = \tilde{r}\left(\hat{\beta}^{(WLS,n,W,\pi)}\right) = \left(\mathbf{I} - \tilde{X}\left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}'\right)\tilde{\varepsilon},$$

i. e. due to normality of disturbances, $\tilde{r}\left(\hat{\beta}^{(WLS,n,W,\pi)}\right)$ is normally distributed (with heteroscadasticity).

... for LWS-estimation ... simulations of *p*-values

Recalling the derivation of *t*-statistics in the classical regression

Recalling that $\tilde{W} = diag \left\{ w_1^{\frac{1}{2}}, w_2^{\frac{1}{2}}, ..., w_n^{\frac{1}{2}} \right\}$ and

$$\mathcal{L}(\tilde{\varepsilon}) = \mathcal{N}\left(\mathbf{0}, \sigma^2 \tilde{W}^2\right)$$

we have from

$$\hat{\beta}^{(OLS,n)}\left(\tilde{Y},\tilde{X}\right) - \beta^{0} = \hat{\beta}^{(WLS,n,W,\pi)} - \beta^{0} = \tilde{X}\left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}'\tilde{\varepsilon}.$$

$$\boldsymbol{E}\left\{\hat{\beta}^{(OLS,n)}\left(\tilde{Y},\tilde{X}\right) - \beta^{0}\right\} = \boldsymbol{E}\left\{\hat{\beta}^{(WLS,n,W,\pi)} - \beta^{0}\right\} = 0$$

$$\operatorname{cov}\left\{\hat{\beta}^{(OLS,n)}\left(\tilde{Y},\tilde{X}\right) - \beta^{0}\right\} = \left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}'\tilde{W}^{2}\tilde{X}\left(\tilde{X}'\tilde{X}\right)^{-1}.$$

$$= \left(X'\tilde{W}^{2}X\right)^{-1}\cdot\sum_{i}^{n}w_{i}^{2}\cdot X_{i}\cdot X_{i}'\left(X'\tilde{W}^{2}X\right)^{-1}.$$

i=1

and

... for LWS-estimation ... simulations of *p*-values

Recalling the derivation of *t*-statistics in the classical regression

Denote

$$\begin{bmatrix} \operatorname{cov} \left\{ \hat{\beta}^{(OLS,n)} \left(\tilde{Y}, \tilde{X} \right) - \beta^{0} \right\} \end{bmatrix}_{\ell\ell} = \begin{bmatrix} \operatorname{cov} \left\{ \hat{\beta}^{(WLS,n,W,\pi)} \left(Y, X \right) - \beta^{0} \right\} \end{bmatrix}_{\ell\ell} \\ = \begin{bmatrix} \left(X' \tilde{W}^{2} X \right)^{-1} \cdot \sum_{i=1}^{n} w_{i}^{2} \cdot X_{i} \cdot X_{i}' \left(X' \tilde{W}^{2} X \right)^{-1} \end{bmatrix}_{\ell\ell} \underbrace{=}_{\ell\ell \ (denote)} d_{n,\ell}(W, X).$$

... for LWS-estimation ... simulations of *p*-values

Recalling the derivation of *t*-statistics in the classical regression

Denote

$$\begin{bmatrix} \operatorname{cov}\left\{\hat{\beta}^{(OLS,n)}\left(\tilde{Y},\tilde{X}\right)-\beta^{0}\right\} \end{bmatrix}_{\ell\ell} = \begin{bmatrix} \operatorname{cov}\left\{\hat{\beta}^{(WLS,n,W,\pi)}\left(Y,X\right)-\beta^{0}\right\} \end{bmatrix}_{\ell\ell} \\ = \begin{bmatrix} \left(X'\tilde{W}^{2}X\right)^{-1} \cdot \sum_{i=1}^{n} w_{i}^{2} \cdot X_{i} \cdot X_{i}'\left(X'\tilde{W}^{2}X\right)^{-1} \end{bmatrix}_{\ell\ell} \underbrace{=}_{\ell\ell \text{ (denote)}} d_{n,\ell}(W,X).$$

Then

$$\mathcal{L}\left(\frac{\hat{\beta}_{\ell}^{(OLS,n)}\left(\tilde{Y},\tilde{X}\right)-\beta^{0}}{\sigma d_{n,\ell}(W,X)}\right)=\mathcal{L}\left(\frac{\hat{\beta}_{\ell}^{(LWS,n,W)}-\beta^{0}}{\sigma d_{n,\ell}(W,X)}\right)=\mathcal{N}(0,1).$$

... for LWS-estimation ... simulations of *p*-values

Establishing the result

We can show (similarly as in the OLS-regression),

$$\mathcal{L}\left(\sigma^{-2}\tilde{r}'\left(\hat{\beta}^{(\mathsf{WLS},n,\mathsf{W})}\right)\cdot\tilde{r}\left(\hat{\beta}^{(\mathsf{WLS},n,\mathsf{W})}\right)\right)=\mathcal{L}\left(\sigma^{-2}\cdot\mathsf{RSS}\right)=\chi^{2}_{generalized}\left(n-p\right)$$

in the sense that $\chi^2_{generalized} (n - p)$ is distribution of the sum of squares of n - p independent r. v.'s normally distributed with zero mean but variance not equal one, but w_i .



... for LWS-estimation ... simulations of *p*-values

Establishing the result

We can show (similarly as in the OLS-regression),

$$\mathcal{L}\left(\sigma^{-2}\tilde{r}'\left(\hat{\beta}^{(WLS,n,W)}\right)\cdot\tilde{r}\left(\hat{\beta}^{(WLS,n,W)}\right)\right)=\mathcal{L}\left(\sigma^{-2}\cdot RSS\right)=\chi^{2}_{generalized}\left(n-p\right)$$

in the sense that $\chi^2_{generalized} (n-p)$ is distribution of the sum of squares of n-p independent r. v.'s normally distributed with zero mean but variance not equal one, but w_i .

Let us recall that for the "classical $\chi^2(n-p)$ we have (written symbolically) $E\chi^2(n-p) = n-p$

... for LWS-estimation ... simulations of *p*-values

Establishing the result

We can show (similarly as in the OLS-regression),

$$\mathcal{L}\left(\sigma^{-2}\tilde{r}'\left(\hat{\beta}^{(WLS,n,W)}\right)\cdot\tilde{r}\left(\hat{\beta}^{(WLS,n,W)}\right)\right) = \mathcal{L}\left(\sigma^{-2}\cdot RSS\right) = \chi^{2}_{generalized}\left(n-p\right)$$

in the sense that $\chi^2_{generalized} (n-p)$ is distribution of the sum of squares of n-p independent r. v.'s normally distributed with zero mean but variance not equal one, but w_i .

Let us recall that for the "classical $\chi^2(n-p)$ we have (written symbolically) $E\chi^2(n-p) = n-p$

and hence in the denominatot of *t*-statistics we put

 $\left\{\frac{\tilde{r}'\left(\hat{\beta}^{(WLS,n,W)}\right)\cdot\tilde{r}\left(\hat{\beta}^{(WLS,n,W)}\right)}{(n-p)\cdot\sigma^2}\right\}^{\frac{1}{2}}.$

... for LWS-estimation ... simulations of *p*-values

Establishing the result

We can show (similarly as in the OLS-regression),

$$\mathcal{L}\left(\sigma^{-2}\tilde{r}'\left(\hat{\beta}^{(WLS,n,W)}\right)\cdot\tilde{r}\left(\hat{\beta}^{(WLS,n,W)}\right)\right) = \mathcal{L}\left(\sigma^{-2}\cdot RSS\right) = \chi^{2}_{generalized}\left(n-p\right)$$

in the sense that $\chi^2_{generalized} (n-p)$ is distribution of the sum of squares of n-p independent r. v.'s normally distributed with zero mean but variance not equal one, but w_i .

Let us recall that for the "classical $\chi^2(n-p)$ we have (written symbolically) $E\chi^2(n-p) = n-p$

and hence in the denominatot of *t*-statistics we put

$$\left\{\frac{\tilde{r}'\left(\hat{\beta}^{(WLS,n,W)}\right)\cdot\tilde{r}\left(\hat{\beta}^{(WLS,n,W)}\right)}{(n-p)\cdot\sigma^{2}}\right\}$$

So, to conclude the derivation we need to calculate $E\chi^2_{\text{generalized}}(n-p)$.

... for LWS-estimation ... simulations of *p*-values

Establishing the result

Let us recall that $W^{\frac{1}{2}} = \tilde{W} = \text{diag}\left\{w_1^{\frac{1}{2}}, w_2^{\frac{1}{2}}, ..., w_n^{\frac{1}{2}}\right\}$ and $\tilde{\varepsilon} = \tilde{W}\varepsilon$

and that (as it is well-known from the classical regression)

$$\tilde{r}\left(\hat{\beta}^{(WLS,n)}
ight) = \left(\mathbf{I} - \tilde{X}\left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}'
ight)\tilde{\varepsilon} = \tilde{M}\tilde{W}\varepsilon.$$

denote it by M

... for LWS-estimation ... simulations of *p*-values

Establishing the result

v

Let us recall that
$$W_{2}^{\frac{1}{2}} = \tilde{W} = \text{diag}\left\{w_{1}^{\frac{1}{2}}, w_{2}^{\frac{1}{2}}, ..., w_{n}^{\frac{1}{2}}\right\}$$
 and $\tilde{\varepsilon} = \tilde{W}\varepsilon$

and that (as it is well-known from the classical regression)

$$\tilde{r}\left(\hat{\beta}^{(WLS,n)}\right) = \left(I - \tilde{X}\left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}'\right)\tilde{\varepsilon} = \tilde{M}\tilde{W}\varepsilon.$$

$$\mathbf{E}\left[\tilde{r}'\left(\hat{\beta}^{(WLS,n)}\right) \cdot \tilde{r}\left(\hat{\beta}^{(WLS,n)}\right)\right] = \sigma^{2} \cdot tr\left[W\left(\mathbf{I} - \tilde{X}\left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}'\right)\right]$$
$$= \sigma^{2} \cdot \left(\sum_{i=1}^{n} w_{i}(1-d_{ii})\right)$$
where $d_{ii} = \left[\tilde{X}\left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}'\right]_{ii}$. So, $\frac{\tilde{r}'(\hat{\beta}^{(WLS,n)}) \cdot \tilde{r}(\hat{\beta}^{(WLS,n)})}{\sigma^{2}\sum_{i=1}^{n} w_{i}(1-d_{ii})} \to 1.$

... for LWS-estimation ... simulations of *p*-values

Recalling the classical regression for \tilde{Y}, \tilde{X}

We conclude THEOREM $\mathcal{L}\left(\frac{\hat{\beta}_{\ell}^{(LWS,n,W)} - \beta^{0}}{d_{n,\ell}(W,X)} \cdot \left[\frac{\sum_{i=1}^{n} w_{i}(1 - d_{ii})}{RSS}\right]^{\frac{1}{2}}\right) = t_{generalized}\left(n - p\right)$

... for LWS-estimation ... simulations of *p*-values

Recalling the classical regression for \tilde{Y}, \tilde{X}

We conclude THEOREM

$$\mathcal{L}\left(\frac{\hat{\beta}_{\ell}^{(LWS,n,W)} - \beta^{0}}{d_{n,\ell}(W,X)} \cdot \left[\frac{\sum_{i=1}^{n} w_{i}(1 - d_{ii})}{RSS}\right]^{\frac{1}{2}}\right) = t_{generalized} \left(n - p\right)$$

in the sense that $t_{generalized} (n - p)$ is a ratio of standard normal r.v.

and $\chi^2_{generalized}$ (n-p) r.v. and numerator and denominator are independent

(in the appendix which follows after the end of lecture, an alternative derivation with more details is given, see also ICORS2011).

... for LWS-estimation ... simulations of *p*-values

Function generating the weights

Let me recall that we already know that under low contamination, the intuitively optimal (left) and really optimal (right) weight functions are (in the sense of mean square error of the estimates of regression coefficients).

 $W_{\ell} = W\left(\frac{\ell-1}{n}\right)$

Contamination : 4% outliers



... for LWS-estimation ... simulations of *p*-values

Function generating the weights



... for LWS-estimation ... simulations of *p*-values

Content

- Disqualitying classical regression analysis by ignoring ...
 .../ broken orthogonality condition
 collinearity
- 2 Significance of individual explanatory variable
 - ... simulations of p-values

... for LWS-estimation ... simulations of *p*-values

Framework of simulations

For each value of n = 20, 30, ..., 190 we generated 5000 times



... for LWS-estimation ... simulations of *p*-values

Framework of simulations

For each value of n = 20, 30, ..., 190 we generated 5000 times

$$\frac{\hat{\beta}_{\ell}^{(LWS,n,W)} - \beta^{0}}{d_{n,\ell}(W,X)} \cdot \left[\frac{\sum_{i=1}^{n} w_{i}(1-d_{ii})}{RSS}\right]^{\frac{1}{2}}.$$

The 4875 and 4975 order statistics

among these 5000 values were found.

... for LWS-estimation ... simulations of *p*-values

Framework of simulations

For each value of n = 20, 30, ..., 190 we generated 5000 times

 $\frac{\hat{\beta}_{\ell}^{(LWS,n,W)}-\beta^{0}}{d_{n,\ell}(W,X)}\cdot\left[\frac{\sum_{i=1}^{n}w_{i}(1-d_{ii})}{RSS}\right]^{\frac{1}{2}}.$

- The 4875 and 4975 order statistics among these 5000 values were found.
- We repeated it 100 times and have found empirical means and the roots of mean square errors over these 100 repetitions (these roots of mean square errors are in parentheses).

... for LWS-estimation ... simulations of *p*-values

TABLE 1The simulated quantiles for 5%.

n	20	30	40	50	60	70
$\hat{t}_{0.975}^{LWS}(n)_{(var)}$	2.148 (0.047)	2.087 (0.040)	2.056 _(0.046)	2.027 _(0.045)	2.017 _(0.046)	2.012 _(0.045)
t _{0.975} (<i>n</i>)	2.085	2.043	2.022	2.009	2.000	1.995

n	80	90	100	110	120	130
$\hat{t}_{0.975}^{LWS}(n)$	2.008 (0.040)	1.999 _(0.041)	1.992 _(0.040)	1.991 _(0.041)	1.990 (0.040)	1.988 _(0.040)
$t_{0.975}(n)$	1.990	1.987	1.984	1.982	1.980	1.978

n	140	150	160	170	180	190
$\hat{t}_{0.975}^{LWS}(n)$	1.986 _(0.043)	1.989 (0.041)	1.975 _(0.035)	1.974 _(0.035)	1.973 _(0.035)	1.973 _(0.035)
$t_{0.975}(n)$	1.977	1.976	1.975	1.974	1.974	1.973

By the way, the 0.975-upper quantile of the standard normal distribution is equal to 1.959964, i. e. $\Phi(1.959964) = 0.975$.

... for LWS-estimation ... simulations of *p*-values

TABLE 2

The simulated quantiles for 1%.

n	20	30	40	50	60	70	
$\hat{t}_{0.995}^{LWS}(n)$	2.999 _(0.100)	2.825 (0.082)	2.766 _(0.080)	2.702 (0.085)	2.688 _(0.077)	2.678 _(0.079)	
t _{0.995} (n)	2.845	2.748	2.705	2.678	2.661	2.651	
n	80	90	100	110	120	130	

$\hat{t}_{0.995}^{LWS}(n)$	2.659 _(0.067)	2.644 (0.075)	2.633 (0.077)	2.627 _(0.063)	2.629 (0.070)	2.626 _(0.071)
t _{0.995} (n)	2.640	2.632	2.625	2.619	2.614	2.612

n	140	150	160	170	180	190
$\hat{t}_{0.995}^{LWS}(n)$	2.619 _(0.072)	2.621 (0.073)	2.609 (0.079)	2.609 (0.070)	2.620 (0.078)	2.602 (0.078)
$t_{0.995}(n)$	2.611	2.610	2.609	2.608	2.606	2.605

Again, the 0.995-upper quantile of the standard normal distribution is equal to 2.575, i. e. $\Phi(2.575) = 0.995$.

... for LWS-estimation ... simulations of *p*-values



THANKS FOR ATTENTION



... for LWS-estimation ... simulations of *p*-values



... for LWS-estimation ... simulations of *p*-values

Significance of explanatory variable

Let's recall the simplest classical framework for finite-sample diagnostics:

Regression model

 $Y_i = X'_i \beta^0 + \varepsilon_i, \quad i = 1, 2, ..., n \quad \text{or} \quad Y = X \beta^0 + \varepsilon$

... for LWS-estimation ... simulations of *p*-values

Significance of explanatory variable

Let's recall the simplest classical framework for finite-sample diagnostics:

Regression model

 $Y_i = X'_i \beta^0 + \varepsilon_i, \quad i = 1, 2, ..., n \quad \text{or} \quad Y = X \beta^0 + \varepsilon$

Conditions :

 $\{(X'_i,\varepsilon_i)'\}_{i=1}^{\infty} \underline{\text{i.i.d.}}, F_{X,\varepsilon}(x,v) = F_X(x) \cdot F_{\varepsilon}(v), \underline{F_{\varepsilon}(v)} = \mathcal{N}(0,\sigma^2), \underline{Q} = E[X_1 \cdot X'_i] \text{ is regular.}$
$C_{\ell,\ell}^2 = \left[(X'X)^{-1} \right]_{\ell,\ell}$

... for LWS-estimation ... simulations of *p*-values

Significance of explanatory variable

Let's recall the simplest classical framework for finite-sample diagnostics:

Regression model

 $Y_i = X'_i \beta^0 + \varepsilon_i, \quad i = 1, 2, ..., n \quad \text{or} \quad Y = X \beta^0 + \varepsilon$

Conditions :

$$\{(X'_i,\varepsilon_i)'\}_{i=1}^{\infty} \text{ i.i.d., } F_{X,\varepsilon}(x,v) = F_X(x) \cdot F_{\varepsilon}(v), F_{\varepsilon}(v) = \mathcal{N}(0,\sigma^2), \\ Q = E[X_1 \cdot X'_1] \text{ is regular,}$$

Significance of ℓ -th explanatory variable $X_{i\ell} \Leftrightarrow H_0$: $\hat{\beta}_{\ell}^{(LS,n)} = 0$

Denote

$$\boldsymbol{s}_{n}^{2} = \frac{1}{n-p} \sum_{i=1}^{n} \left(\boldsymbol{Y}_{i} - \boldsymbol{X}_{i}^{\prime} \hat{\boldsymbol{\beta}}^{(LS,n)} \right)^{2}$$

... for LWS-estimation ... simulations of *p*-values

Significance of explanatory variable

Let's recall the simplest classical framework for finite-sample diagnostics:

Regression model

 $Y_i = X'_i \beta^0 + \varepsilon_i, \quad i = 1, 2, ..., n \quad \text{or} \quad Y = X \beta^0 + \varepsilon$

Conditions :

$$\{(X'_i,\varepsilon_i)'\}_{i=1}^{\infty} \text{ i.i.d., } F_{X,\varepsilon}(x,v) = F_X(x) \cdot F_{\varepsilon}(v), F_{\varepsilon}(v) = \mathcal{N}(0,\sigma^2), \\ Q = E[X_1 \cdot X'_1] \text{ is regular,}$$

Significance of ℓ -th explanatory variable $X_{i\ell} \Leftrightarrow H_0: \hat{\beta}_{\ell}^{(LS,n)} = 0$

Denote $c_{\ell,\ell}^2 = \left[(X'X)^{-1} \right]_{\ell,\ell}$ and $s_n^2 = \frac{1}{n-p} \sum_{i=1}^n \left(Y_i - X'_i \hat{\beta}^{(LS,n)} \right)^2$. <u>Test is based on</u> $\mathcal{L} \left(\frac{\hat{\beta}_{\ell}^{(LS,n)} - \beta_{\ell}^0}{\mathbf{s}_n \cdot \mathbf{c}_{\ell,\ell}} \right) = t_{n-p}$ (Fisher-Cochran theorem)

... for LWS-estimation ... simulations of *p*-values

Significance of explanatory variable

Let's recall the classical "treatment" under heteroscedaticity: (we shall need it)

Conditions :

 $\{(X'_i,\varepsilon_i)'\}_{i=1}^{\infty} \underline{i.d.}, F_{X,\varepsilon_i}(x,v) = F_X(x) \cdot F_{\varepsilon_i}(v), F_{\varepsilon_i}(v) = \mathcal{N}(0,\sigma_i^2),$

 $QE[X_1 \cdot X_1']$ is regular,

... for LWS-estimation ... simulations of *p*-values

Significance of explanatory variable

Let's recall the classical "treatment" under heteroscedaticity: (we shall need it)

Conditions :

 $\{(X'_i,\varepsilon_i)'\}_{i=1}^{\infty} \underline{i.d.}, F_{X,\varepsilon_i}(x,v) = F_X(x) \cdot F_{\varepsilon_i}(v), F_{\varepsilon_i}(v) = \mathcal{N}(0,\sigma_i^2),$

 $QE[X_1 \cdot X_1']$ is regular,

Significance of ℓ -th explanatory variable $X_{i\ell} \Leftrightarrow H_0: \hat{\beta}_{\ell}^{(LS,n)} = 0$

Denote

 $\boldsymbol{d}_{\ell,\ell}^{2} = \left[\left(\boldsymbol{X}^{\prime}\boldsymbol{X}\right)^{-1} \sum_{i=1}^{n} r_{i}^{2} \left(\hat{\beta}^{(LS,n)} \right) \cdot \boldsymbol{X}_{i} \cdot \boldsymbol{X}_{i}^{\prime} \left(\boldsymbol{X}^{\prime}\boldsymbol{X} \right)^{-1} \right]_{\ell,\ell}.$

(Halbert White estimator - 1980)

... for LWS-estimation ... simulations of *p*-values

Significance of explanatory variable

Let's recall the classical "treatment" under heteroscedaticity: (we shall need it)

Conditions :

$$[(X'_i,\varepsilon_i)']_{i=1}^{\infty} \text{ i.d., } F_{X,\varepsilon_i}(x,v) = F_X(x) \cdot F_{\varepsilon_i}(v), F_{\varepsilon_i}(v) = \mathcal{N}(0,\sigma_i^2),$$

 $QE[X_1 \cdot X_1']$ is regular,

Significance of ℓ -th explanatory variable $X_{i\ell} \Leftrightarrow H_0$: $\hat{\beta}_{\ell}^{(LS,n)} = 0$

Denote

$$\mathcal{O}_{\ell,\ell}^{2} = \left[\left(X'X \right)^{-1} \sum_{i=1}^{n} r_{i}^{2} \left(\hat{\beta}^{(LS,n)} \right) \cdot X_{i} \cdot X_{i}' \left(X'X \right)^{-1} \right]_{\ell}$$

(Halbert White estimator - 1980)

Test is based on

$$\frac{dn - \beta_{\ell}^{0}}{d_{\ell,\ell}} \approx t_{n-p} \quad (?!?!) \\ (\mathcal{L}\left((n-p) \cdot d_{\ell,\ell}^{2}\right) \approx \chi_{n-p}^{2} ?!?!)$$

... for LWS-estimation ... simulations of *p*-values

Recalling several facts

Let's $W = diag \{w_1, w_2, ..., w_n\}$ be a weight matrix. Then the <u>classical</u> Weighted Least Squares is given as

 $\hat{\beta}^{(WLS,n)} = (X'WX)^{-1} X'WY.$

... for LWS-estimation ... simulations of *p*-values

Recalling several facts

Let's $W = diag \{w_1, w_2, ..., w_n\}$ be a weight matrix. Then the classical Weighted Least Squares is given as $\hat{\beta}^{(WLS,n)} = (X'WX)^{-1}X'WY.$ Let \mathcal{P} be the set of all permutations of integers $\{1, 2, ..., n\}$. Fix $\pi \in \mathcal{P}, \pi = \{i_1, i_2, \dots, i_n\}$ and put $Y_{\pi} = (Y_{i_{h}}, Y_{i_{h}}, \dots, Y_{i_{h}})', \quad X_{\pi} = (X_{i_{h}}, X_{i_{h}}, \dots, X_{i_{h}})' \quad (\text{and} \quad \varepsilon_{\pi} = (\varepsilon_{i_{h}}, \varepsilon_{i_{h}}, \dots, \varepsilon_{i_{h}})')$ and consider model $Y_{\tau} = X_{\tau}\beta^0 + \varepsilon_{\tau}$

... for LWS-estimation ... simulations of *p*-values

Recalling several facts

Let's $W = diag \{w_1, w_2, ..., w_n\}$ be a weight matrix. Then the classical Weighted Least Squares is given as $\hat{\beta}^{(WLS,n)} = (X'WX)^{-1} X'WY.$ Let \mathcal{P} be the set of all permutations of integers $\{1, 2, ..., n\}$. Fix $\pi \in \mathcal{P}, \pi = \{i_1, i_2, \dots, i_n\}$ and put $Y_{\pi} = (Y_{i_{h}}, Y_{i_{h}}, \dots, Y_{i_{h}})', \quad X_{\pi} = (X_{i_{h}}, X_{i_{h}}, \dots, X_{i_{h}})' \quad (\text{and} \quad \varepsilon_{\pi} = (\varepsilon_{i_{h}}, \varepsilon_{i_{h}}, \dots, \varepsilon_{i_{h}})')$ and consider model $Y_{\pi} = X_{\pi}\beta^0 + \varepsilon_{\pi}.$ Then denote $\hat{\beta}^{(WLS,n,\pi)} = (X'_{\pi}WX_{\pi})^{-1}X'_{\pi}WY_{\pi} \text{ and } S^{2}_{\pi} = \sum_{i}^{n} w_{i} \left(Y_{i_{i}} - X'_{i_{i}}\hat{\beta}^{(WLS,n,\pi)}\right)^{2}.$

... for LWS-estimation ... simulations of *p*-values

Deriving existence of $\hat{\beta}^{(LWS,n)}$

Then for any $\pi \in \mathcal{P}$

$$S_{\pi}^{2} = \sum_{j=1}^{n} w_{i} \left(Y_{i_{j}} - X_{i_{j}}^{\prime} \beta^{(WLS,n,\pi)} \right)^{2} \leq \min_{\beta \in \mathbb{R}^{p}} \sum_{j=1}^{n} w_{i} \left(Y_{i_{j}} - X_{i_{j}}^{\prime} \beta \right)^{2}.$$
 (6)

... for LWS-estimation ... simulations of *p*-values

Deriving existence of $\hat{\beta}^{(LWS,n)}$

Then for any $\pi \in \mathcal{P}$

$$S_{\pi}^{2} = \sum_{j=1}^{n} w_{i} \left(Y_{i_{j}} + X_{i_{j}}^{\prime} \hat{\beta}^{(WLS,n,\pi)} \right)^{2} \leq \min_{\beta \in R^{p}} \sum_{j=1}^{n} w_{i} \left(Y_{i_{j}} - X_{i_{j}}^{\prime} \beta \right)^{2}.$$
 (6)

Let (Ω, \mathcal{A}, P) be the basic <u>Probability Space</u> and write $Z(\omega)$ for the value of r.v. at point $\omega \in \Omega$.

... for LWS-estimation ... simulations of *p*-values

Deriving existence of $\hat{\beta}^{(LWS,n)}$

Then for any $\pi \in \mathcal{P}$

$$S_{\pi}^{2} = \sum_{j=1}^{n} w_{i} \left(Y_{i_{j}} + X_{i_{j}}^{\prime} \hat{\beta}^{(WLS,n,\pi)} \right)^{2} \leq \min_{\beta \in R^{p}} \sum_{j=1}^{n} w_{i} \left(Y_{i_{j}} - X_{i_{j}}^{\prime} \beta \right)^{2}.$$
 (6)

Let (Ω, \mathcal{A}, P) be the basic <u>Probability Space</u> and write $Z(\omega)$ for the value of r.v. at point $\omega \in \Omega$. Finally, assume $w_1 \ge w_2 \ge ... \ge w_n$, fix ω_0 and put

 $\hat{\pi}(\omega_0) = \underset{\pi \in \mathcal{P}}{\operatorname{arg\,min}} S^2_{\pi}(\omega_0).$

... for LWS-estimation ... simulations of *p*-values

Deriving existence of $\hat{\beta}^{(LWS,n)}$

Then for any $\pi \in \mathcal{P}$

$$S_{\pi}^{2} = \sum_{j=1}^{n} w_{i} \left(Y_{i_{j}} - X_{i_{j}}^{\prime} \hat{\beta}^{(WLS,n,\pi)} \right)^{2} \leq \min_{\beta \in R^{\rho}} \sum_{j=1}^{n} w_{i} \left(Y_{i_{j}} - X_{i_{j}}^{\prime} \beta \right)^{2}.$$
 (6)

Let (Ω, \mathcal{A}, P) be the basic <u>Probability Space</u> and write $Z(\omega)$ for the value of r.v. at point $\omega \in \Omega$. Finally, assume $w_1 \ge w_2 \ge ... \ge w_n$, fix ω_0 and put

$$\hat{\pi}(\omega_0) = \underset{\pi \in \mathcal{P}}{\operatorname{arg\,min}} S^2_{\pi}(\omega_0).$$

Then - due to (6)

$$S^2_{\hat{\pi}(\omega_0)}(\omega_0) \leq \min_{eta \in R^p} \min_{\pi \in \mathcal{P}} \sum_{j=1}^{\prime\prime} w_j \left(Y_{i_j} - X'_{i_j}eta
ight)^2$$

... for LWS-estimation ... simulations of *p*-values

Deriving form of $\hat{\beta}^{(LWS,n)}$



$$\hat{\beta}^{(WLS,n,\hat{\pi}(\omega_0))}(\omega_0) = \hat{\beta}^{(LWS,n,w)}(\omega_0)$$



... for LWS-estimation ... simulations of *p*-values

Deriving form of $\hat{\beta}^{(LWS,n)}$

It means

$$\hat{\beta}^{(WLS,n,\hat{\pi}(\omega_0))}(\omega_0) = \hat{\beta}^{(LWS,n,w)}(\omega_0)$$

Repeat it for all $\omega \in \Omega$.

... for LWS-estimation ... simulations of *p*-values

Deriving form of $\hat{\beta}^{(LWS,n)}$

It means

$$\hat{\beta}^{(WLS,n,\hat{\pi}(\omega_0))}(\omega_0) = \hat{\beta}^{(LWS,n,w)}(\omega_0)$$

Repeat it for all $\omega \in \Omega$.

Fix $\pi \in \mathcal{P}$ and put

 $oldsymbol{B}(\pi) = \left\{ \omega \in \Omega \ : \ \pi = \hat{\pi}(\omega)
ight\}.$

... for LWS-estimation ... simulations of *p*-values

Deriving form of $\hat{\beta}^{(LWS,n)}$

It means

$$\hat{\beta}^{(WLS,n,\hat{\pi}(\omega_0))}(\omega_0) = \hat{\beta}^{(LWS,n,w)}(\omega_0)$$

Repeat it for all $\omega \in \Omega$.

Fix $\pi \in \mathcal{P}$ and put

$$oldsymbol{B}(\pi) = \left\{ \omega \in \Omega \ : \ \pi = \hat{\pi}(\omega)
ight\}.$$

Then i.e. $\omega \in B(\pi) \implies \hat{\beta}^{(LWS,n,w)}(\omega) = \hat{\beta}^{(WLS,n,\pi)}(\omega)$

$$\hat{\beta}^{(LWS,n,w)}(\omega) = \hat{\beta}^{(WLS,n,\pi)}(\omega) = \underset{\beta \in R^{p}}{\operatorname{arg\,min}} \sum_{i=1} w_{i} \left(Y_{\pi_{1}} - X'_{\pi}\beta\right)^{2}.$$

... for LWS-estimation ... simulations of *p*-values

Deriving form of $\hat{\beta}^{(LWS,n)}$

It means

$$\hat{\beta}^{(WLS,n,\hat{\pi}(\omega_0))}(\omega_0) = \hat{\beta}^{(LWS,n,w)}(\omega_0)$$

Repeat it for all $\omega \in \Omega$.

Fix $\pi \in \mathcal{P}$ and put

$$oldsymbol{B}(\pi) = \left\{ \omega \in \Omega \ : \ \pi = \hat{\pi}(\omega)
ight\}.$$

Then

$$\omega \in B(\pi) \implies \hat{\beta}^{(LWS,n,w)}(\omega) = \hat{\beta}^{(WLS,n,\pi)}(\omega)$$

i.e.

$$\hat{\beta}^{(LWS,n,w)}(\omega) = \hat{\beta}^{(WLS,n,\pi)}(\omega) = \arg\min_{\beta \in R^p} \sum_{i=1}^n w_i \left(Y_{\pi_1} - X'_{\pi}\beta\right)^2.$$

Due to i.i.d. framework, $\forall \pi \in \mathcal{P}$

 $P(B(\pi)) = (n!)^{-1}$

and <u>all B's are the "same"</u>.

... for LWS-estimation ... simulations of *p*-values

Deriving form of $\hat{\beta}^{(LWS,n)}$

Let's rewrite one line of previuos slide, for $\omega_0 \in B(\pi)$

 $\hat{\beta}^{(LWS,n,w)}(\omega_0) = \hat{\beta}^{(WLS,n,\pi)}(\omega_0) = \arg\min_{\beta \in R^p} \sum_{i=1}^n w_i \left(Y_{\pi} - X'_{\pi}\beta\right)^2.$ (7)

and drop for a while ω_0 and π .

... for LWS-estimation ... simulations of *p*-values

Deriving form of $\hat{\beta}^{(LWS,n)}$

Let's rewrite one line of previuos slide, for $\omega_0 \in B(\pi)$

$$\hat{\beta}^{(LWS,n,w)}(\omega_0) = \hat{\beta}^{(WLS,n,\pi)}(\omega_0) = \arg\min_{\beta \in R^p} \sum_{i=1}^n w_i \left(Y_{\pi} - X'_{\pi}\beta\right)^2.$$
(7)

and drop for a while ω_0 and π .

Then putting $\tilde{W} = diag\left\{w_1^{-\frac{1}{2}}, w_2^{-\frac{1}{2}}, ..., w_n^{-\frac{1}{2}}\right\}$ and $\tilde{Y} = \tilde{W}Y, \quad \tilde{X} = \tilde{W}X$, (7) implies

$$\hat{\beta}^{(LWS,n,w)}(\boldsymbol{Y},\boldsymbol{X}) = \left(\tilde{\boldsymbol{X}}'\tilde{\boldsymbol{X}}\right)^{-1}\tilde{\boldsymbol{X}}'\tilde{\boldsymbol{Y}} = \hat{\beta}^{(OLS,n)}(\tilde{\boldsymbol{Y}},\tilde{\boldsymbol{X}})$$

... for LWS-estimation ... simulations of *p*-values

Deriving form of $\hat{\beta}^{(LWS,n)}$

Let's rewrite one line of previuos slide, for $\omega_0 \in B(\pi)$

$$\hat{\beta}^{(LWS,n,w)}(\omega_0) = \hat{\beta}^{(WLS,n,\pi)}(\omega_0) = \arg\min_{\beta \in R^p} \sum_{i=1}^n w_i \left(Y_\pi - X'_\pi\beta\right)^2.$$
(7)

and drop for a while ω_0 and π .

Then putting $\tilde{W} = diag\left\{w_1^{-\frac{1}{2}}, w_2^{-\frac{1}{2}}, ..., w_n^{-\frac{1}{2}}\right\}$ and $\tilde{Y} = \tilde{W}Y, \quad \tilde{X} = \tilde{W}X$, (7) implies

$$\hat{\beta}^{(LWS,n,w)}(\mathbf{Y},\mathbf{X}) = \left(\tilde{\mathbf{X}}'\tilde{\mathbf{X}}\right)^{-1}\tilde{\mathbf{X}}'\tilde{\mathbf{Y}} = \hat{\beta}^{(OLS,n)}(\tilde{\mathbf{Y}},\tilde{\mathbf{X}}).$$

Consider now the model

$$\tilde{Y} = \tilde{X}\beta^{0} + \tilde{\varepsilon}$$
 with $\mathcal{L}(\tilde{\varepsilon}) = \mathcal{N}\left(0, \sigma^{2}\tilde{W}^{2}\right)$

 $(\tilde{W}^2 \text{ is known} \Rightarrow \text{heteroscedasticity is governed by one unknown parameter } \sigma^2)$

... for LWS-estimation ... simulations of *p*-values

Recalling the classical regression for \tilde{Y}, \tilde{X}

By standard way (please read only last but one line)

$$\hat{\beta}^{(OLS,n)}(\tilde{Y},\tilde{X})' = \left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}\tilde{Y} = \left[\left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}'\tilde{X}\right]\left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}\tilde{Y}$$
$$= \left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}'\left[\tilde{X}\left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}\tilde{Y}\right] = \left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}'\hat{\tilde{Y}}$$
(8)

As $\hat{\tilde{Y}} = \tilde{X}\hat{\beta}^{(OLS,n)}$, i.e. $\hat{\tilde{Y}}$ is a linear combimation of the columns of \tilde{X} $\Rightarrow \quad \hat{\tilde{Y}} \perp \tilde{r}\left(\hat{\beta}^{(WLS,n)}\right) = \tilde{Y} - \tilde{X}\hat{\beta}^{(OLS,n)}$

and due to the normality of disturbances \tilde{Y} is independent with \tilde{r} . Then due to (8) $\hat{\beta}^{(LWS,n,w)}$ is independent with $\tilde{r}\left(\hat{\beta}^{(LWS,n,w)}\right)$

(we shall need it in a few minutes).

... for LWS-estimation ... simulations of *p*-values

Recalling the classical regression for \tilde{Y}, \tilde{X}

Please read again only the last line

Regress ℓ -th column of matrix \tilde{X} on all other columns of this matrix and denote the residuals by $u^{(\ell)}$.

Recall that we consider $\tilde{Y} = \tilde{X}\beta^0 + \tilde{\varepsilon}$ with $\tilde{Y} = \tilde{W}Y$, etc. Then

(see e.g. Wooldridge (2003))

$$\hat{\beta}_{\ell}^{(LWS,n,w)}(\boldsymbol{Y},\boldsymbol{X}) - \beta_{\ell}^{0} = \hat{\beta}_{\ell}^{(OLS,n)}(\tilde{\boldsymbol{Y}},\tilde{\boldsymbol{X}}) - \beta_{\ell}^{0} = \left[\boldsymbol{u}^{(\ell)}\right]' \tilde{\varepsilon} \cdot \left\|\boldsymbol{u}^{(\ell)}\right\|^{-2}$$

with

$$\operatorname{var}\left(\hat{\beta}_{\ell}^{(LWS,n,W)}\right) = \sigma^{2} \sum_{i=1}^{n} W_{i} \left[u_{i}^{(\ell)}\right]^{2} \left\|u^{(\ell)}\right\|^{-4} \underbrace{=}_{denote} \sigma^{2} D_{n}^{2}(W,X).$$

Then

$$\mathcal{L}\left(\frac{\hat{\beta}_{\ell}^{(LWS,n,w)} - \beta_{\ell}^{0}}{\sigma D_{n}(W,X)}\right) = \mathcal{N}(0,1).$$
(9)

... for LWS-estimation ... simulations of *p*-values

Recalling the classical regression for \tilde{Y}, \tilde{X}

Let us recall that $W_{\frac{1}{2}}^{\frac{1}{2}} = \tilde{W} = \text{diag}\left\{w_{1}^{\frac{1}{2}}, w_{2}^{\frac{1}{2}}, ..., w_{n}^{\frac{1}{2}}\right\}$ and that it is well-known from the classical regression

 $\tilde{r}\left(\hat{\beta}^{(WLS,n)}\right) = \underbrace{\left(\mathbf{I} - \tilde{X}\left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}'\right)}_{\varepsilon} \tilde{\varepsilon} = \tilde{M}\tilde{W}\varepsilon.$

denote it by M

... for LWS-estimation ... simulations of *p*-values

Recalling the classical regression for \tilde{Y}, \tilde{X}

Let us recall that $W^{\frac{1}{2}} = \tilde{W} = \text{diag}\left\{w_1^{\frac{1}{2}}, w_2^{\frac{1}{2}}, ..., w_n^{\frac{1}{2}}\right\}$ and that it is well-known from the classical regression

$$\widetilde{r}\left(\widehat{\beta}^{(WLS,n)}\right) = \underbrace{\left(\mathbf{I} - \widetilde{X}\left(\widetilde{X}'\widetilde{X}\right)^{-1}\widetilde{X}'\right)}_{\text{denote it by }\widetilde{M}} \widetilde{\varepsilon} = \widetilde{M}\widetilde{W}\varepsilon.$$

$$\mathbf{E}\left[\widetilde{r}'\left(\widehat{\beta}^{(WLS,n)}\right) \cdot \widetilde{r}\left(\widehat{\beta}^{(WLS,n)}\right)\right] = \sigma^{2} \cdot tr\left[W\left(\mathbf{I} - \widetilde{X}\left(\widetilde{X}'\widetilde{X}\right)^{-1}\widetilde{X}'\right)\right]$$

$$= \sigma^{2} \cdot \left(\sum_{i=1}^{n} w_{i}(1 - d_{ii})\right)$$

where $d_{ii} = \left[\tilde{X} \left(\tilde{X}' \tilde{X} \right)^{-1} \tilde{X}' \right]_{ii}$. So, $s_n^2 = \frac{\tilde{r}' \left(\hat{\beta}^{(WLS,n)} \right) \cdot \tilde{r} \left(\hat{\beta}^{(WLS,n)} \right)}{\sum_{i=1}^n w_i (1-d_{ii})} \rightarrow \mathbf{\textit{E}} s_n^2 = \sigma^2$.

... for LWS-estimation ... simulations of *p*-values

Recalling the classical regression for \tilde{Y}, \tilde{X}

Let us recall that $W_2^{\frac{1}{2}} = \tilde{W} = \text{diag}\left\{w_1^{\frac{1}{2}}, w_2^{\frac{1}{2}}, ..., w_n^{\frac{1}{2}}\right\}$ and that it is well-known from the classical regression

 $\tilde{r}\left(\hat{\beta}^{(WLS,n)}\right) = \left(\mathbf{I} - \tilde{X}\left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}'\right)\tilde{\varepsilon} = \tilde{M}\tilde{W}\varepsilon.$

denote it by M

Then (recall also that $E\tilde{\varepsilon} = 0$ and $\operatorname{cov}(\tilde{\varepsilon}) = \sigma^2 W$ where $W = \operatorname{diag} \{w_1, w_2, ..., w_n\}$) $\tilde{r}'\left(\hat{\beta}^{(WLS,n)}\right) \cdot \tilde{r}\left(\hat{\beta}^{(WLS,n)}\right) = [\tilde{\varepsilon}]' \tilde{M}\tilde{\varepsilon} = [\tilde{\varepsilon}]' W^{-\frac{1}{2}} \underbrace{W^{\frac{1}{2}}\tilde{M}W^{\frac{1}{2}}}_{M} \underbrace{W^{-\frac{1}{2}}\tilde{\varepsilon}}_{\eta} = \eta' M\eta$

with $\mathcal{L}(\eta) = \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}).$

... for LWS-estimation ... simulations of *p*-values

Recalling the classical regression for \tilde{Y}, \tilde{X}

Let $Q'MQ = \Lambda$ where $Q = [q_1, q, ..., q_n]$ and $\Lambda = \text{diag} \{\lambda_1, \lambda_2, ..., \lambda_n\}$ with $M \cdot q_i = \lambda_i \cdot q_i$, i. e. $\lambda_i > 0$, i = 1, 2, ..., n - p, $\lambda_i = 0$ otherwise and Q'Q = QQ' = I. Then $M = Q\Lambda Q'$ and $\tilde{r}'\left(\hat{\beta}^{(WLS,n)}\right) \cdot \tilde{r}\left(\hat{\beta}^{(WLS,n)}\right) = [\tilde{\varepsilon}]'Q\Lambda Q'\tilde{\varepsilon} = \xi'\xi$ with $\xi = \Lambda^{\frac{1}{2}}Q\eta$ (notice that the last *p* coordinates of $\xi = 0$), i.e. $E\xi = 0$ and $\operatorname{cov}(\xi) = \sigma\Lambda$.

... for LWS-estimation ... simulations of *p*-values

Recalling the classical regression for \tilde{Y}, \tilde{X}

Then

$\mathcal{L}\left(\xi\right) = \mathcal{N}\left(\mathbf{0}, \sigma^{2} \cdot \mathbf{\Lambda}\right).$

... for LWS-estimation ... simulations of *p*-values

Recalling the classical regression for \tilde{Y}, \tilde{X}

Then

Finally,

$$\mathcal{L}(\xi) = \mathcal{N}(\mathbf{0}, \sigma^2 \cdot \mathbf{\Lambda}).$$

$$\mathcal{L}\left(\sigma^{-2}\tilde{r}'\left(\hat{\beta}^{(\mathsf{WLS},\mathsf{n})}\right)\cdot\tilde{r}\left(\hat{\beta}^{(\mathsf{WLS},\mathsf{n})}\right)\right)=\chi^{2}_{generalized}\left(\mathsf{n}-\mathsf{p}\right)$$

in the sense that

 $\chi^2_{generalized} (n - p)$ is distribution

of the sum of squares of n - p independent r. v.'s normally distributed

with zero mean but variance not equal one.

... for LWS-estimation ... simulations of *p*-values

Recalling the classical regression for \tilde{Y}, \tilde{X}

Then

$$\mathcal{L}(\xi) = \mathcal{N}(\mathbf{0}, \sigma^2 \cdot \mathbf{\Lambda}).$$

Finally,

$$\mathcal{L}\left(\sigma^{-2}\tilde{r}'\left(\hat{\beta}^{(\mathsf{WLS},n)}\right)\cdot\tilde{r}\left(\hat{\beta}^{(\mathsf{WLS},n)}\right)\right)=\chi^{2}_{generalized}\left(n-p\right)$$

in the sense that $\chi^2_{generalized} (n-p)$ is distribution of the sum of squares of n-p independent r. v.'s normally distributed with zero mean but variance not equal one.

We conclude,

$$\mathcal{L}\left(\frac{\hat{\beta}^{(LWS,n,w)} - \beta^{0}}{D_{n}(W,X)} \cdot \left[\frac{\sum_{i=1}^{n} w_{i} (1 - d_{ii})}{\tilde{r}'\left(\hat{\beta}^{(WLS,n)}\right) \cdot \tilde{r}\left(\hat{\beta}^{(WLS,n)}\right)}\right]^{\frac{1}{2}}\right) = t_{generalized} (n - p).$$

... for LWS-estimation ... simulations of *p*-values



THANKS FOR ATTENTION

