FINAL EXAM ECON 7800 FALL 2003

ECONOMICS DEPARTMENT, UNIVERSITY OF UTAH

Problem 49. *3 points Let* $x_1, \dots, x_n$ *and* $y_1, \dots, y_n$ *be two arbitrary vectors, and* α *and* β *two arbitrary scalars. As usual, we use the notation* $\bar{x} = \frac{1}{n}\sum_{j=1}^{n} x_j$ *and* $\bar{y} = \frac{1}{n}\sum_{k=1}^{n} y_k$. *Show that*

$$\sum_{i=1}^{n}(x_i - \alpha)(y_i - \beta) = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) + n(\bar{x} - \alpha)(\bar{y} - \beta) \tag{1}$$

Date of exam Wednesday December 10, 10:30–12:30 in BUC 207.

Answer. Here is the answer for which you need to know a trick:

$$\sum_{i=1}^{n}(x_i-\alpha)(y_i-\beta) = \sum_{i=1}^{n}\Big((x_i-\bar{x})+(\bar{x}-\alpha)\Big)\Big((y_i-\bar{y})+(\bar{y}-\beta)\Big) \tag{2}$$

$$= \sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y}) + \sum_{i=1}^{n}(x_i-\bar{x})(\bar{y}-\beta)$$

$$+ \sum_{i=1}^{n}(\bar{x}-\alpha)(y_i-\bar{y}) + \sum_{i=1}^{n}(\bar{x}-\alpha)(\bar{y}-\beta) \tag{3}$$

$$= \sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y}) + (\bar{y}-\beta)\sum_{i=1}^{n}(x_i-\bar{x})$$

$$+ (\bar{x}-\alpha)\sum_{i=1}^{n}(y_i-\bar{y}) + n(\bar{x}-\alpha)(\bar{y}-\beta) \tag{4}$$

but here the two middle summands are zero.

A different successful strategy would be to multiply out both sides and verify that they are equal. The lefthand side is

$$\sum_{i=1}^{n}(x_i-\alpha)(y_i-\beta)=\sum_{i=1}^{n}x_iy_i-\sum_{i=1}^{n}x_i\beta-\sum_{i=1}^{n}\alpha y_i+\sum_{i=1}^{n}\alpha\beta \tag{5}$$

$$=\sum_{i=1}^{n}x_iy_i-\beta\sum_{i=1}^{n}x_i-\alpha\sum_{i=1}^{n}y_i+n\alpha\beta \tag{6}$$

$$=\sum_{i=1}^{n}x_iy_i-\beta n\bar{x}-\alpha n\bar{y}+n\alpha\beta \tag{7}$$

Now the righthand side is

$$\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})+n(\bar{x}-\alpha)(\bar{y}-\beta)=\sum_{i=1}^{n}x_iy_i-\sum_{i=1}^{n}x_i\bar{y}-\sum_{i=1}^{n}\bar{x}y_i+n\bar{x}\bar{y}+n\bar{x}\bar{y}-n\bar{x}\beta-n\alpha\bar{y}+ \tag{8}$$

$$=\sum_{i=1}^{n}x_iy_i-\bar{y}\sum_{i=1}^{n}x_i-\bar{x}\sum_{i=1}^{n}y_i+n\bar{x}\bar{y}+n\bar{x}\bar{y}-n\bar{x}\beta-n\alpha\bar{y}+ \tag{9}$$

$$=\sum_{i=1}^{n}x_iy_i-n\bar{x}\bar{y}-n\bar{x}\bar{y}+n\bar{x}\bar{y}+n\bar{x}\bar{y}-n\bar{x}\beta-n\alpha\bar{y}+n\alpha\beta \tag{10}$$

but this is the same as the lefthand side. □

Problem 50.

- **a.** *2 points Verify that the matrix* $\boldsymbol{D} = \boldsymbol{I} - \frac{1}{n}\boldsymbol{\iota}\boldsymbol{\iota}^\top$ *is symmetric and idempotent.*

- **b.** *1 point Compute the trace* $\operatorname{tr}\boldsymbol{D}$.

Answer. $\operatorname{tr}\boldsymbol{D} = n - 1$. One can see this either by writing down the matrix element by element, or use the linearity of the trace plus the rule that $\operatorname{tr}(\boldsymbol{A}\boldsymbol{B}) = \operatorname{tr}(\boldsymbol{B}\boldsymbol{A})$. $\operatorname{tr}\boldsymbol{I} = n$ and $\operatorname{tr}(\boldsymbol{\iota}\boldsymbol{\iota}^\top) = \operatorname{tr}(\boldsymbol{\iota}^\top\boldsymbol{\iota}) = \operatorname{tr} n = n$. □

- **c.** *1 point For any vector of observations* $\boldsymbol{y}$ *compute* $\boldsymbol{D}\boldsymbol{y}$.

Answer. Element by element one can write

$$\boldsymbol{D}\boldsymbol{y} = \begin{bmatrix} 1-\frac{1}{n} & -\frac{1}{n} & \dots & -\frac{1}{n} \\ -\frac{1}{n} & 1-\frac{1}{n} & \dots & -\frac{1}{n} \\ \vdots & & \ddots & \vdots \\ -\frac{1}{n} & -\frac{1}{n} & \dots & 1-\frac{1}{n} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} y_1 - \bar{y} \\ y_2 - \bar{y} \\ \vdots \\ y_n - \bar{y} \end{bmatrix} \tag{11}$$

There is also a more elegant matrix theoretical proof available □

- **d.** *1 point Is there a vector* $\boldsymbol{a} \neq \boldsymbol{o}$ *for which* $\boldsymbol{D}\boldsymbol{a} = \boldsymbol{o}$? *If so, give an example of such a vector.*

Answer. $\boldsymbol{\iota}$ is, up to a scalar factor, the only nonzero vector with $\boldsymbol{D}\boldsymbol{\iota} = \boldsymbol{o}$. □

• **e.** *1 point Show that the sample variance of a vector of observations $\boldsymbol{y}$ can be written in matrix notation as*

$$\text{The sample variance of } \boldsymbol{y} \text{ is } \frac{1}{n}\sum(y_i - \bar{y})^2 = \frac{1}{n}\boldsymbol{y}^\top \boldsymbol{D}\boldsymbol{y} \tag{12}$$

Answer. Let's get rid of the factor $\frac{1}{n}$ which appears on both sides: we have to show that

$$\sum(y_i - \bar{y})^2 = \boldsymbol{y}^\top \boldsymbol{D}\boldsymbol{y} = \boldsymbol{y}^\top \boldsymbol{D}^\top \boldsymbol{D}\boldsymbol{y} \tag{13}$$

This is the squared length of the vector $\boldsymbol{D}\boldsymbol{y}$ which we computed in part c. □

Problem 51. *The fitted values $\hat{\boldsymbol{y}}$ and the residuals $\hat{\boldsymbol{\varepsilon}}$ are "orthogonal" in two different ways.*

• **a.** *2 points Show that the inner product $\hat{\boldsymbol{y}}^\top \hat{\boldsymbol{\varepsilon}} = 0$. Why should you expect this from the geometric intuition of Least Squares?*

Answer. Use $\hat{\boldsymbol{\varepsilon}} = \boldsymbol{M}\boldsymbol{y}$ and $\hat{\boldsymbol{y}} = (\boldsymbol{I} - \boldsymbol{M})\boldsymbol{y}$: $\hat{\boldsymbol{y}}^\top \hat{\boldsymbol{\varepsilon}} = \boldsymbol{y}^\top(\boldsymbol{I} - \boldsymbol{M})\boldsymbol{M}\boldsymbol{y} = 0$ because $\boldsymbol{M}(\boldsymbol{I} - \boldsymbol{M}) = \boldsymbol{O}$. This is a consequence of the more general result given in problem **??**.

□

• **b.** *2 points Sometimes two random variables are called "orthogonal" if their covariance is zero. Show that $\hat{\boldsymbol{y}}$ and $\hat{\boldsymbol{\varepsilon}}$ are orthogonal also in this sense, i.e., show that for every i and j, $\operatorname{cov}[\hat{y}_i, \hat{\varepsilon}_j] = 0$. In matrix notation this can also be written $\mathcal{C}[\hat{\boldsymbol{y}}, \hat{\boldsymbol{\varepsilon}}] = \boldsymbol{O}$. Here the covariance matrix $\mathcal{C}[\boldsymbol{x}, \boldsymbol{z}]$ is that matrix whose (i, j) element* *is* $\operatorname{cov}[x_i, z_j]$. *The covariance matrix satisfies the rules* $\mathcal{C}[\boldsymbol{B}\boldsymbol{y}, \boldsymbol{T}\boldsymbol{z}] = \boldsymbol{B}\,\mathcal{C}[\boldsymbol{y}, \boldsymbol{z}]\boldsymbol{T}^\top$, $\mathcal{C}[\boldsymbol{y}, \boldsymbol{y}] = \mathcal{V}[\boldsymbol{y}]$, $\mathcal{C}[\boldsymbol{z}, \boldsymbol{y}] = (\mathcal{C}[\boldsymbol{y}, \boldsymbol{z}])^\top$, $\mathcal{C}[\boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z}] = \mathcal{C}[\boldsymbol{x}, \boldsymbol{z}] + \mathcal{C}[\boldsymbol{y}, \boldsymbol{z}]$, *and* $\mathcal{C}[\boldsymbol{x}, \boldsymbol{c}] = \boldsymbol{O}$ *if* $\boldsymbol{c}$ *is a vector of constants.*

Answer. $\mathcal{C}[\hat{\boldsymbol{y}}, \hat{\boldsymbol{\varepsilon}}] = \mathcal{C}[(\boldsymbol{I} - \boldsymbol{M})\boldsymbol{y}, \boldsymbol{M}\boldsymbol{y}] = (\boldsymbol{I} - \boldsymbol{M})\,\mathcal{V}[\boldsymbol{y}]\boldsymbol{M}^\top = (\boldsymbol{I} - \boldsymbol{M})(\sigma^2\boldsymbol{I})\boldsymbol{M} = \sigma^2(\boldsymbol{I} - \boldsymbol{M})\boldsymbol{M} = \boldsymbol{O}$. This is a consequence of the more general result given in question **??**. □

Problem 52. *2 points The Cobb-Douglas production function postulates the following relationship between annual output* q_t *and the inputs of labor* ℓ_t *and capital* k_t*:*

$$q_t = \mu \ell_t^\beta k_t^\gamma \exp(\varepsilon_t). \tag{14}$$

q_t, ℓ_t, *and* k_t *are observed, and* μ, β, γ, *and the* ε_t *are to be estimated. By the variable transformation* $x_t = \log q_t$, $y_t = \log \ell_t$, $z_t = \log k_t$, *and* $\alpha = \log \mu$, *one obtains the linear regression*

$$x_t = \alpha + \beta y_t + \gamma z_t + \varepsilon_t \tag{15}$$

Sometimes the following alternative variable transformation is made: $u_t = \log(q_t/\ell_t)$, $v_t = \log(k_t/\ell_t)$, *and the regression*

$$u_t = \alpha + \gamma v_t + \varepsilon_t \tag{16}$$

is estimated. How are the regressions (15) and (16) related to each other?

Answer. Write (16) as

$$x_t - y_t = \alpha + \gamma(z_t - y_t) + \varepsilon_t \tag{17}$$

and collect terms to get

$$x_t = \alpha + (1-\gamma)y_t + \gamma z_t + \varepsilon_t \tag{18}$$

From this follows that running the regression (16) is equivalent to running the regression (15) with the constraint $\beta + \gamma = 1$ imposed. □

Problem 53. *4 points Let $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ be a regression model with intercept, in which the first column of $\boldsymbol{X}$ is the vector $\boldsymbol{\iota}$, and let $\hat{\boldsymbol{\beta}}$ the least squares estimator of $\boldsymbol{\beta}$. Show that the covariance matrix between $\bar{y}$ and $\hat{\boldsymbol{\beta}}$, which is defined as the matrix (here consisting of one row only) that contains all the covariances*

$$\mathcal{C}[\bar{y}, \hat{\boldsymbol{\beta}}] \equiv \begin{bmatrix} \operatorname{cov}[\bar{y}, \hat{\beta}_1] & \operatorname{cov}[\bar{y}, \hat{\beta}_2] & \cdots & \operatorname{cov}[\bar{y}, \hat{\beta}_k] \end{bmatrix} \tag{19}$$

has the following form: $\mathcal{C}[\bar{y}, \hat{\boldsymbol{\beta}}] = \frac{\sigma^2}{n}\begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}$ where n is the number of observations. Hint: *That the regression has an intercept term as first column of the $\boldsymbol{X}$-matrix means that $\boldsymbol{X}\boldsymbol{e}^{(1)} = \boldsymbol{\iota}$, where $\boldsymbol{e}^{(1)}$ is the unit vector having 1 in the first place and zeros elsewhere, and $\boldsymbol{\iota}$ is the vector which has ones everywhere.* Second Hint: *The covariance matrix satisfies the rules $\mathcal{C}[\boldsymbol{B}\boldsymbol{y}, \boldsymbol{T}\boldsymbol{z}] = \boldsymbol{B}\,\mathcal{C}[\boldsymbol{y}, \boldsymbol{z}]\boldsymbol{T}^\top$ and $\mathcal{C}[\boldsymbol{y}, \boldsymbol{y}] = \mathcal{V}[\boldsymbol{y}]$. (Other rules for the covariance matrix, which will not be needed here,* *are* $\mathcal{C}[\boldsymbol{z}, \boldsymbol{y}] = (\mathcal{C}[\boldsymbol{y}, \boldsymbol{z}])^\top$, $\mathcal{C}[\boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z}] = \mathcal{C}[\boldsymbol{x}, \boldsymbol{z}] + \mathcal{C}[\boldsymbol{y}, \boldsymbol{z}]$, *and* $\mathcal{C}[\boldsymbol{x}, \boldsymbol{c}] = \boldsymbol{O}$ *if* $\boldsymbol{c}$ *is a vector of constants.)*

Answer. Write both $\bar{y}$ and $\hat{\boldsymbol{\beta}}$ in terms of $\boldsymbol{y}$, i.e., $\bar{y} = \frac{1}{n}\boldsymbol{\iota}^\top \boldsymbol{y}$ and $\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \boldsymbol{y}$. Therefore

$$\mathcal{C}[\bar{y}, \hat{\boldsymbol{\beta}}] = \frac{1}{n}\boldsymbol{\iota}^\top \mathcal{V}[\boldsymbol{y}]\boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{X})^{-1} = \frac{\sigma^2}{n}\boldsymbol{\iota}^\top \boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{X})^{-1} = \frac{\sigma^2}{n}\boldsymbol{e}^{(1)\top}\boldsymbol{X}^\top \boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{X})^{-1} = \frac{\sigma^2}{n}\boldsymbol{e}^{(1)\top}. \tag{20}$$

□

Problem 54. *You have two data vectors* x_i *and* y_i $(i = 1, \ldots, n)$, *and the true model is*

$$y_i = \beta x_i + \varepsilon_i \tag{21}$$

where x_i *and* ε_i *satisfy the basic assumptions of the linear regression model. The least squares estimator for this model is*

$$\tilde{\beta} = (\boldsymbol{x}^\top \boldsymbol{x})^{-1}\boldsymbol{x}^\top \boldsymbol{y} = \frac{\sum x_i y_i}{\sum x_i^2} \tag{22}$$

- **a.** *1 point Is* $\tilde{\beta}$ *an unbiased estimator of* β*? (Proof is required.)*

Answer. First derive a nice expression for $\tilde\beta - \beta$:

$$\begin{aligned}
\tilde\beta - \beta &= \frac{\sum x_i y_i}{\sum x_i^2} - \frac{\sum x_i^2 \beta}{\sum x_i^2} \\
&= \frac{\sum x_i (y_i - x_i\beta)}{\sum x_i^2} \\
&= \frac{\sum x_i \varepsilon_i}{\sum x_i^2} \qquad \text{since} \quad y_i = \beta x_i + \varepsilon_i \\
\mathrm{E}[\tilde\beta - \beta] &= \mathrm{E}\left[\frac{\sum x_i \varepsilon_i}{\sum x_i^2}\right] \\
&= \frac{\sum \mathrm{E}[x_i \varepsilon_i]}{\sum x_i^2} \\
&= \frac{\sum x_i \mathrm{E}[\varepsilon_i]}{\sum x_i^2} = 0 \qquad \text{since} \quad \mathrm{E}\,\varepsilon_i = 0.
\end{aligned}$$

□

- **b.** *2 points Derive the variance of $\tilde\beta$. (Show your work.)*

Answer.

$$
\begin{aligned}
\operatorname{var} \tilde{\beta} &= \mathrm{E}[\tilde{\beta} - \beta]^2 \\
&= \mathrm{E}\left(\frac{\sum x_i \varepsilon_i}{\sum x_i^2}\right)^2 \\
&= \frac{1}{(\sum x_i^2)^2} \mathrm{E}[\sum x_i \varepsilon_i]^2 \\
&= \frac{1}{(\sum x_i^2)^2} \left(\mathrm{E} \sum (x_i \varepsilon_i)^2 + 2\, \mathrm{E} \sum_{i<j} (x_i \varepsilon_i)(x_j \varepsilon_j) \right) \\
&= \frac{1}{(\sum x_i^2)^2} \sum \mathrm{E}[x_i \varepsilon_i]^2 \qquad \text{since the } \varepsilon_i\text{'s are uncorrelated, i.e., } \operatorname{cov}[\varepsilon_i, \varepsilon_j] = 0 \text{ for } i \neq j \\
&= \frac{1}{(\sum x_i^2)^2} \sigma^2 \sum x_i^2 \qquad \text{since all } \varepsilon_i \text{ have equal variance } \sigma^2 \\
&= \frac{\sigma^2}{\sum x_i^2}.
\end{aligned}
$$

□

Problem 55. *3 points The model is $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ but all rows of the $\boldsymbol{X}$-matrix are exactly equal. What can you do? Can you estimate $\boldsymbol{\beta}$? If not, are there any linear combinations of the components of $\boldsymbol{\beta}$ which you can estimate? Can you estimate σ^2?*

Answer. If all rows are equal, then each column is a multiple of $\boldsymbol{\iota}$. Therefore, if there are more than one column, none of the individual components of $\boldsymbol{\beta}$ can be estimated. But you can estimate $\boldsymbol{x}^\top\boldsymbol{\beta}$ (if $\boldsymbol{x}$ is one of the row vectors of $\boldsymbol{X}$) and you can estimate σ^2. □

Problem 56.

- **a.** *2 points Show that*

$$SSE = \boldsymbol{\varepsilon}^\top \boldsymbol{M}\boldsymbol{\varepsilon} \qquad \text{where} \qquad \boldsymbol{M} = \boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^-\boldsymbol{X}^\top \tag{23}$$

Answer. $SSE = \hat{\varepsilon}^\top\hat{\varepsilon}$, where $\hat{\varepsilon} = \boldsymbol{y} - \boldsymbol{X}\hat{\beta} = \boldsymbol{y} - \boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^-\boldsymbol{X}^\top\boldsymbol{y} = \boldsymbol{M}\boldsymbol{y}$ where $\boldsymbol{M} = \boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^-\boldsymbol{X}^\top$. From $\boldsymbol{M}\boldsymbol{X} = \boldsymbol{O}$ follows $\hat{\varepsilon} = \boldsymbol{M}(\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}) = \boldsymbol{M}\boldsymbol{\varepsilon}$. Since $\boldsymbol{M}$ is idempotent and symmetric, it follows $\hat{\varepsilon}^\top\hat{\varepsilon} = \boldsymbol{\varepsilon}^\top\boldsymbol{M}\boldsymbol{\varepsilon}$. □

- **b.** *1 point Is SSE observed? Is $\boldsymbol{\varepsilon}$ observed? Is $\boldsymbol{M}$ observed?*

- **c.** *3 points Under the usual assumption that $\boldsymbol{X}$ has full column rank, show that*

$$\mathrm{E}[SSE] = \sigma^2(n-k) \tag{24}$$

Answer. $\mathrm{E}[\hat{\varepsilon}^\top\hat{\varepsilon}] = \mathrm{E}[\operatorname{tr}\boldsymbol{\varepsilon}^\top\boldsymbol{M}\boldsymbol{\varepsilon}] = \mathrm{E}[\operatorname{tr}\boldsymbol{M}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^\top] = \sigma^2\operatorname{tr}\boldsymbol{M} = \sigma^2\operatorname{tr}(\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^-\boldsymbol{X}^\top) = \sigma^2(n - \operatorname{tr}(\boldsymbol{X}^\top\boldsymbol{X})^-\boldsymbol{X}^\top\boldsymbol{X}) = \sigma^2(n-k)$. □

Problem 57. *4 points Show the following: If $\boldsymbol{\beta} = \hat{\hat{\beta}}$ is the unconstrained minimum argument of the Lagrange function*

$$L(\boldsymbol{\beta}, \boldsymbol{\lambda}^*) = (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})^\top(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}) + (\boldsymbol{R}\boldsymbol{\beta} - \boldsymbol{u})^\top\boldsymbol{\lambda}^* \tag{25}$$

for some fixed value $\boldsymbol{\lambda}^*$*, and if at the same time* $\hat{\hat{\beta}}$ *satisfies* $\boldsymbol{R}\hat{\hat{\beta}} = \boldsymbol{u}$*, then* $\boldsymbol{\beta} = \hat{\hat{\beta}}$ *minimizes* $(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})^\top(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})$ *subject to the constraint* $\boldsymbol{R}\boldsymbol{\beta} = \boldsymbol{u}$*.*

Answer. Since $\hat{\hat{\beta}}$ minimizes the Lagrange function, we know that

$$(\boldsymbol{y} - \boldsymbol{X}\tilde{\boldsymbol{\beta}})^\top(\boldsymbol{y} - \boldsymbol{X}\tilde{\boldsymbol{\beta}}) + (\boldsymbol{R}\tilde{\boldsymbol{\beta}} - \boldsymbol{u})^\top\boldsymbol{\lambda}^* \ge (\boldsymbol{y} - \boldsymbol{X}\hat{\hat{\beta}})^\top(\boldsymbol{y} - \boldsymbol{X}\hat{\hat{\beta}}) + (\boldsymbol{R}\hat{\hat{\beta}} - \boldsymbol{u})^\top\boldsymbol{\lambda}^* \tag{26}$$

for all $\tilde{\boldsymbol{\beta}}$. Since by assumption, $\hat{\hat{\beta}}$ also satisfies the constraint, this simplifies to:

$$(\boldsymbol{y} - \boldsymbol{X}\tilde{\boldsymbol{\beta}})^\top(\boldsymbol{y} - \boldsymbol{X}\tilde{\boldsymbol{\beta}}) + (\boldsymbol{R}\tilde{\boldsymbol{\beta}} - \boldsymbol{u})^\top\boldsymbol{\lambda}^* \ge (\boldsymbol{y} - \boldsymbol{X}\hat{\hat{\beta}})^\top(\boldsymbol{y} - \boldsymbol{X}\hat{\hat{\beta}}). \tag{27}$$

This is still true for all $\tilde{\boldsymbol{\beta}}$. If we only look at those $\tilde{\boldsymbol{\beta}}$ which satisfy the constraint, we get

$$(\boldsymbol{y} - \boldsymbol{X}\tilde{\boldsymbol{\beta}})^\top(\boldsymbol{y} - \boldsymbol{X}\tilde{\boldsymbol{\beta}}) \ge (\boldsymbol{y} - \boldsymbol{X}\hat{\hat{\beta}})^\top(\boldsymbol{y} - \boldsymbol{X}\hat{\hat{\beta}}). \tag{28}$$

This means, $\hat{\hat{\beta}}$ is the constrained minimum argument.

□

Problem 58. *Prove the following facts about the diagonal elements of the so-called "hat matrix"* $\boldsymbol{H} = \boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top$*, which has its name because* $\boldsymbol{H}\boldsymbol{y} = \hat{\boldsymbol{y}}$*, i.e., it puts the hat on* $\boldsymbol{y}$*.*

- **a.** *1 point* $\boldsymbol{H}$ *is a projection matrix, i.e., it is symmetric and idempotent.*

Answer. Symmetry follows from the laws for the transposes of products: $\boldsymbol{H}^\top = (\boldsymbol{ABC})^\top = \boldsymbol{C}^\top\boldsymbol{B}^\top\boldsymbol{A}^\top = \boldsymbol{H}$ where $\boldsymbol{A} = \boldsymbol{X}$, $\boldsymbol{B} = (\boldsymbol{X}^\top\boldsymbol{X})^{-1}$ which is symmetric, and $\boldsymbol{C} = \boldsymbol{X}^\top$. Idempotency $\boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top = \boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top$. ☐

• **b.** *1 point Prove that a symmetric idempotent matrix is nonnegative definite.*

Answer. If $\boldsymbol{H}$ is symmetric and idempotent, then for arbitrary $\boldsymbol{g}$, $\boldsymbol{g}^\top\boldsymbol{H}\boldsymbol{g} = \boldsymbol{g}^\top\boldsymbol{H}^\top\boldsymbol{H}\boldsymbol{g} = \|\boldsymbol{H}\boldsymbol{g}\|^2 \geq 0$. But $\boldsymbol{g}^\top\boldsymbol{H}\boldsymbol{g} \geq 0$ for all $\boldsymbol{g}$ is the criterion which makes $\boldsymbol{H}$ nonnegative definite. ☐

• **c.** *2 points Show that*

$$0 \leq h_{ii} \leq 1 \tag{29}$$

Answer. If $\boldsymbol{e}_i$ is the vector with a 1 on the ith place and zeros everywhere else, then $\boldsymbol{e}_i^\top\boldsymbol{H}\boldsymbol{e}_i = h_{ii}$. From $\boldsymbol{H}$ nonnegative definite follows therefore that $h_{ii} \geq 0$. $h_{ii} \leq 1$ follows because $\boldsymbol{I} - \boldsymbol{H}$ is symmetric and idempotent (and therefore nonnegative definite) as well: it is the projection on the orthogonal complement. ☐

• **d.** *2 points Show: the average value of the h_{ii} is $\sum h_{ii}/n = k/n$, where k is the number of columns of $\boldsymbol{X}$. (Hint: for this you must compute the trace* $\operatorname{tr}\boldsymbol{H}$*.)*

Answer. The average can be written as

$$\tfrac{1}{n}\operatorname{tr}(\boldsymbol{H}) = \tfrac{1}{n}\operatorname{tr}(\boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top) = \tfrac{1}{n}\operatorname{tr}(\boldsymbol{X}^\top\boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}) = \tfrac{1}{n}\operatorname{tr}(\boldsymbol{I}_k) = \tfrac{k}{n}.$$

Here we used $\operatorname{tr}\boldsymbol{BC} = \operatorname{tr}\boldsymbol{CB}$. ☐

• **e.** *1 point Show that $\frac{1}{n}\boldsymbol{\iota}\boldsymbol{\iota}^\top$ is a projection matrix. Here $\boldsymbol{\iota}$ is the n-vector of ones.*

• **f.** *2 points Show: If the regression has a constant term, then* $\boldsymbol{H} - \frac{1}{n}\boldsymbol{\iota}\boldsymbol{\iota}^\top$ *is a projection matrix.*

Answer. If $\boldsymbol{\iota}$, the vector of ones, is one of the columns of $\boldsymbol{X}$ (or a linear combination of these columns), this means there is a vector $\boldsymbol{a}$ with $\boldsymbol{\iota} = \boldsymbol{X}\boldsymbol{a}$. From this follows $\boldsymbol{H}\boldsymbol{\iota}\boldsymbol{\iota}^\top = \boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{X}\boldsymbol{a}\boldsymbol{\iota}^\top$ $\boldsymbol{X}\boldsymbol{a}\boldsymbol{\iota}^\top = \boldsymbol{\iota}\boldsymbol{\iota}^\top$. One can use this to show that $\boldsymbol{H} - \frac{1}{n}\boldsymbol{\iota}\boldsymbol{\iota}^\top$ is idempotent: $(\boldsymbol{H} - \frac{1}{n}\boldsymbol{\iota}\boldsymbol{\iota}^\top)(\boldsymbol{H} - \frac{1}{n}\boldsymbol{\iota}\boldsymbol{\iota}^\top) = \boldsymbol{H}\boldsymbol{H} - \boldsymbol{H}\frac{1}{n}\boldsymbol{\iota}\boldsymbol{\iota}^\top - \frac{1}{n}\boldsymbol{\iota}\boldsymbol{\iota}^\top\boldsymbol{H} + \frac{1}{n}\boldsymbol{\iota}\boldsymbol{\iota}^\top\frac{1}{n}\boldsymbol{\iota}\boldsymbol{\iota}^\top = \boldsymbol{H} - \frac{1}{n}\boldsymbol{\iota}\boldsymbol{\iota}^\top - \frac{1}{n}\boldsymbol{\iota}\boldsymbol{\iota}^\top + \frac{1}{n}\boldsymbol{\iota}\boldsymbol{\iota}^\top = \boldsymbol{H} - \frac{1}{n}\boldsymbol{\iota}\boldsymbol{\iota}^\top$. □

• **g.** *1 point Show: If the regression has a constant term, then one can sharpen inequality* (29) *to* $1/n \le h_{ii} \le 1$.

Answer. $\boldsymbol{H} - \boldsymbol{\iota}\boldsymbol{\iota}^\top/n$ is a projection matrix, therefore nonnegative definite, therefore its diagonal elements $h_{ii} - 1/n$ are nonnegative. □

Problem 59. *Assume you have* n_1 *observations* $u_j \sim N(\mu_1, \sigma^2)$ *and* n_2 *observations* $v_j \sim N(\mu_2, \sigma^2)$*, all independent of each other, and you want to test whether* $\mu_1 = \mu_2$. *(Note that the variances are known to be equal).*

• **a.** *2 points Write the model in the form* $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$.

Answer.

$$\begin{bmatrix}\boldsymbol{u}\\ \boldsymbol{v}\end{bmatrix} = \begin{bmatrix}\boldsymbol{\iota}_1\mu_1 + \boldsymbol{\varepsilon}_1\\ \boldsymbol{\iota}_2\mu_2 + \boldsymbol{\varepsilon}_2\end{bmatrix} = \begin{bmatrix}\boldsymbol{\iota}_1 & \boldsymbol{o}\\ \boldsymbol{o} & \boldsymbol{\iota}_2\end{bmatrix}\begin{bmatrix}\mu_1\\ \mu_2\end{bmatrix} + \begin{bmatrix}\boldsymbol{\varepsilon}_1\\ \boldsymbol{\varepsilon}_2\end{bmatrix}. \tag{30}$$

here $\boldsymbol{\iota}_1$ and $\boldsymbol{\iota}_2$ are vectors of ones of appropriate lengths. □

• **b.** *2 points Compute $(\boldsymbol{X}^\top\boldsymbol{X})^{-1}$ in this case.*

Answer.

$$\boldsymbol{X}^\top\boldsymbol{X} = \begin{bmatrix} \boldsymbol{\iota}_1^\top & \boldsymbol{o}^\top \\ \boldsymbol{o}^\top & \boldsymbol{\iota}_2^\top \end{bmatrix} \begin{bmatrix} \boldsymbol{\iota}_1 & \boldsymbol{o} \\ \boldsymbol{o} & \boldsymbol{\iota}_2 \end{bmatrix} = \begin{bmatrix} n_1 & 0 \\ 0 & n_2 \end{bmatrix} \tag{31}$$

$$(\boldsymbol{X}^\top\boldsymbol{X})^{-1} = \begin{bmatrix} \frac{1}{n_1} & 0 \\ 0 & \frac{1}{n_2} \end{bmatrix} \tag{32}$$

□

• **c.** *2 points Compute $\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{y}$ in this case.*

Answer.

$$\boldsymbol{X}^\top\boldsymbol{y} = \begin{bmatrix} \boldsymbol{\iota}_1^\top & \boldsymbol{o}^\top \\ \boldsymbol{o}^\top & \boldsymbol{\iota}_2^\top \end{bmatrix} \begin{bmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n_1} u_i \\ \sum_{j=1}^{n_2} v_j \end{bmatrix} \tag{33}$$

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{y} = \begin{bmatrix} \frac{1}{n_1} & 0 \\ 0 & \frac{1}{n_2} \end{bmatrix} \begin{bmatrix} \sum_{i=1}^{n_1} u_i \\ \sum_{j=1}^{n_2} v_j \end{bmatrix} = \begin{bmatrix} \bar{u} \\ \bar{v} \end{bmatrix} \tag{34}$$

□

• **d.** *3 points Compute $SSE = (\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})^\top(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})$ and s^2, the unbiased estimator of σ^2, in this case.*

Answer.

$$\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}} = \begin{bmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{bmatrix} - \begin{bmatrix} \boldsymbol{\iota}_1 & \boldsymbol{o} \\ \boldsymbol{o} & \boldsymbol{\iota}_2 \end{bmatrix} \begin{bmatrix} \bar{u} \\ \bar{v} \end{bmatrix} = \begin{bmatrix} \boldsymbol{u} - \boldsymbol{\iota}_1 \bar{u} \\ \boldsymbol{v} - \boldsymbol{\iota}_2 \bar{u} \end{bmatrix} \tag{35}$$

$$SSE = \sum_{i=1}^{n_1} (u_i - \bar{u})^2 + \sum_{j=1}^{n_2} (v_j - \bar{v})^2 \tag{36}$$

$$s^2 = \frac{\sum_{i=1}^{n_1} (u_i - \bar{u})^2 + \sum_{j=1}^{n_2} (v_j - \bar{v})^2}{n_1 + n_2 - 2} \tag{37}$$

□

• **e.** *1 point Next, the hypothesis $\mu_1 = \mu_2$ must be written in the form $\boldsymbol{R}\boldsymbol{\beta} = u$. Since in the present case $\boldsymbol{R}$ has just has one row, it should be written as a row-vector $\boldsymbol{R} = \boldsymbol{r}^\top$, and since the vector $\boldsymbol{u}$ has only one component, it should be written as a scalar u, i.e., the hypothesis should be written in the form $\boldsymbol{r}^\top\boldsymbol{\beta} = u$. What are $\boldsymbol{r}$ and u in our case?*

Answer. Since $\boldsymbol{\beta} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$, the constraint can be written as

$$\begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = 0 \qquad \text{i.e.,} \qquad \boldsymbol{r} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \qquad \text{and} \qquad u = 0 \tag{38}$$

□

• **f.** *2 points Compute the standard deviation of $\boldsymbol{r}^\top\hat{\boldsymbol{\beta}}$.*

Answer. First compute the variance and then take the square root.

$$\text{var}[\boldsymbol{r}^\top\hat{\boldsymbol{\beta}}] = \sigma^2\boldsymbol{r}^\top(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{r} = \sigma^2\begin{bmatrix}1 & -1\end{bmatrix}\begin{bmatrix}\frac{1}{n_1} & 0\\ 0 & \frac{1}{n_2}\end{bmatrix}\begin{bmatrix}1\\ -1\end{bmatrix} = \sigma^2\Big(\frac{1}{n_1}+\frac{1}{n_2}\Big) \tag{39}$$

One can also see this without matrix algebra. $\text{var}[\bar{u} = \sigma^2\frac{1}{n_1}$, $\text{var}[\bar{v} = \sigma^2\frac{1}{n_2}$, and since $\bar{u}$ and $\bar{v}$ are independent, the variance of the difference is the sum of the variances. □

• **g.** *2 points Use*

$$\frac{\boldsymbol{r}^\top\hat{\boldsymbol{\beta}} - u}{s\sqrt{\boldsymbol{r}^\top(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{r}}} \sim t_{n-k} \quad \textit{when } H \textit{ is true.} \tag{40}$$

to derive the formula for the t-test.

Answer. The test statistic is $\bar{u}-\bar{v}$ divided by its estimated standard deviation, i.e.,

$$\frac{\bar{u}-\bar{v}}{s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t_{n_1+n_2-2} \quad \text{when } H \text{ is true.} \tag{41}$$

□

Problem 60. [Seb77, exercise 4d-3] *Given $n+1$ observations y_j from a $N(\mu,\sigma^2)$. After the first n observations, it is suspected that a sudden change in the mean of the distribution occurred, i.e., that $y_{n+1} \sim N(\nu,\sigma^2)$ with $\nu \neq \mu$. We will use here three* *different approaches to derive the same test statistic for testing the hypothesis that the $n+1$st observation has the same population mean as the previous observations, i.e., that $\nu = \mu$, against the two-sided alternative. The formulas for this statistic should be given in terms of the observations y_i. It is recommended to use the notation $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ and $\bar{\bar{y}} = \frac{1}{n+1}\sum_{j=1}^{n+1} y_j$.*

- **a.** *3 points First you should derive this statistic by testing whether $\nu - \mu = 0$ (the "Wald principle"). For this you must compute the BLUE of $\nu - \mu$ and its standard deviation and construct the t statistic from this.*

Answer. BLUE of μ is $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$, and that of ν is y_{n+1}. BLUE of $\nu - \mu$ is $\bar{y} - y_{n+1}$. Because of independence $\text{var}[\bar{y} - y_{n+1}] = \text{var}[\bar{y}] + \text{var}[y_{n+1}] = \sigma^2((1/n) + 1) = \sigma^2(n+1)/n$. Standard deviation is $\sigma\sqrt{(n+1)/n}$.

For the denominator in the t-statistic you need the s^2 from the unconstrained regression, which is

$$s^2 = \frac{1}{n-1}\sum_{j=1}^{n}(y_j - \bar{y})^2 \tag{42}$$

What happened to the $(n+1)$st observation here? It always has a zero residual. And the factor $1/(n-1)$ should really be written $1/(n+1-2)$: there are $n+1$ observations and 2 parameters.

Divide $\bar{y} - y_{n+1}$ by its standard deviation and replace σ by s (the square root of s^2) to get the t statistic

$$\frac{\bar{y} - y_{n+1}}{s\sqrt{1 + \frac{1}{n}}} \tag{43}$$

□

• **b.** *2 points One can interpret this same formula also differently (and this is why this test is sometimes called the "predictive" Chow test). Compute the Best Linear Unbiased Predictor of y_{n+1} on the basis of the first n observations, call it $\hat{\hat{y}}(n+1)_{n+1}$. Show that the predictive residual $y_{n+1} - \hat{\hat{y}}(n+1)_{n+1}$, divided by the square root of* $\text{MSE}[\hat{\hat{y}}(n+1)_{n+1}; y_{n+1}]$*, with σ replaced by s (based on the first n observations only), is equal to the above t statistic.*

Answer. BLUP of y_{n+1} based on first n observations is $\bar{y}$ again. Since it is unbiased, $\text{MSE}[\bar{y}; y_{n+1}] = \text{var}[\bar{y} - y_{n+1}] = \sigma^2(n+1)/n$. From now on everything is as in part a. □

• **c.** *6 points Next you should show that the above two formulas are identical to the statistic based on comparing the SSEs of the constrained and unconstrained models (the likelihood ratio principle). Give a formula for the constrained SSE_r, the unconstrained SSE_u, and the F-statistic.*

Answer. According to the Likelihood Ratio principle, one has to compare the residual sums of squares in the regressions under the assumption that the mean did not change with that under the assumption that the mean changed. If the mean did not change (constrained model), then $\bar{\bar{y}}$ is the OLS of μ. In order to make it easier to derive the difference between constrained and unconstrained SSE, we will write the constrained SSE as follows:

$$SSE_r = \sum_{j=1}^{n+1}(y_j - \bar{\bar{y}})^2 = \sum_{j=1}^{n+1} y_j^2 - (n+1)\bar{\bar{y}}^2 = \sum_{j=1}^{n+1} y_j^2 - \frac{1}{n+1}(n\bar{y} + y_{n+1})^2$$

If one allows the mean to change (unconstrained model), then $\bar{y}$ is the BLUE of μ, and y_{n+1} is the BLUE of ν.

$$SSE_u = \sum_{j=1}^{n}(y_j - \bar{y})^2 + (y_{n+1} - y_{n+1})^2 = \sum_{j=1}^{n} y_j^2 - n\bar{y}^2.$$

Now subtract:

$$\begin{aligned}
SSE_r - SSE_u &= y_{n+1}^2 + n\bar{y}^2 - \frac{1}{n+1}(n\bar{y} + y_{n+1})^2 \\
&= y_{n+1}^2 + n\bar{y}^2 - \frac{1}{n+1}(n^2\bar{y}^2 + 2n\bar{y}y_{n+1} + y_{n+1}^2) \\
&= (1 - \frac{1}{n+1})y_{n+1}^2 + (n - \frac{n^2}{n+1})\bar{y}^2 - \frac{n}{n+1}2\bar{y}y_{n+1} \\
&= \frac{n}{n+1}(y_{n+1} - \bar{y})^2.
\end{aligned}$$

Interestingly, this depends on the first n observations only through $\bar{y}$.

Since the unconstrained model has $n+1$ observations and 2 parameters, the test statistic is

$$\frac{SSE_r - SSE_u}{SSE_u/(n+1-2)} = \frac{\frac{n}{n+1}(y_{n+1}-\bar{y})^2}{\sum_1^n (y_j-\bar{y})^2/(n-1)} = \frac{(y_{n+1}-\bar{y})^2\, n(n-1)}{\sum_1^n (y_j-\bar{y})^2\,(n+1)} \sim F_{1,n-1} \tag{44}$$

This is the square of the t statistic (43). □

Problem 61. *3 points We are in the model $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ and we have a matrix $\boldsymbol{W}$ of "instrumental variables" which satisfies the following three conditions:* $\operatorname{plim} \frac{1}{n}\boldsymbol{W}^\top\boldsymbol{\varepsilon} = \boldsymbol{o}$, $\operatorname{plim} \frac{1}{n}\boldsymbol{W}^\top\boldsymbol{W} = \boldsymbol{Q}$ *exists, is nonrandom and positive definite, and* $\operatorname{plim} \frac{1}{n}\boldsymbol{W}^\top\boldsymbol{X} = \boldsymbol{D}$ *exists, is nonrandom and has full column rank. Show that the instrumental variables estimator*

$$\tilde{\boldsymbol{\beta}} = \left(\boldsymbol{X}^\top\boldsymbol{W}(\boldsymbol{W}^\top\boldsymbol{W})^{-1}\boldsymbol{W}^\top\boldsymbol{X}\right)^{-1}\boldsymbol{X}^\top\boldsymbol{W}(\boldsymbol{W}^\top\boldsymbol{W})^{-1}\boldsymbol{W}^\top\boldsymbol{y} \tag{45}$$

is consistent. Hint: Write $\tilde{\boldsymbol{\beta}}_n - \boldsymbol{\beta} = \boldsymbol{B}_n \cdot \frac{1}{n}\boldsymbol{W}^\top\boldsymbol{\varepsilon}$ *and show that the sequence of matrices* $\boldsymbol{B}_n$ *has a* plim.

Answer. Write it as

$$\begin{aligned}\tilde{\beta}_n &= \Big(\boldsymbol{X}^\top \boldsymbol{W}(\boldsymbol{W}^\top \boldsymbol{W})^{-1}\boldsymbol{W}^\top \boldsymbol{X}\Big)^{-1} \boldsymbol{X}^\top \boldsymbol{W}(\boldsymbol{W}^\top \boldsymbol{W})^{-1}\boldsymbol{W}^\top (\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}) \\ &= \boldsymbol{\beta} + \Big(\boldsymbol{X}^\top \boldsymbol{W}(\boldsymbol{W}^\top \boldsymbol{W})^{-1}\boldsymbol{W}^\top \boldsymbol{X}\Big)^{-1} \boldsymbol{X}^\top \boldsymbol{W}(\boldsymbol{W}^\top \boldsymbol{W})^{-1}\boldsymbol{W}^\top \boldsymbol{\varepsilon} \\ &= \boldsymbol{\beta} + \Big((\frac{1}{n}\boldsymbol{X}^\top \boldsymbol{W})(\frac{1}{n}\boldsymbol{W}^\top \boldsymbol{W})^{-1}(\frac{1}{n}\boldsymbol{W}^\top \boldsymbol{X})\Big)^{-1} (\frac{1}{n}\boldsymbol{X}^\top \boldsymbol{W})(\frac{1}{n}\boldsymbol{W}^\top \boldsymbol{W})^{-1}\frac{1}{n}\boldsymbol{W}^\top \boldsymbol{\varepsilon},\end{aligned}$$

i.e., the $\boldsymbol{B}_n$ and $\boldsymbol{B}$ of the hint are as follows:

$$\begin{aligned}\boldsymbol{B}_n &= \Big((\frac{1}{n}\boldsymbol{X}^\top \boldsymbol{W})(\frac{1}{n}\boldsymbol{W}^\top \boldsymbol{W})^{-1}(\frac{1}{n}\boldsymbol{W}^\top \boldsymbol{X})\Big)^{-1} (\frac{1}{n}\boldsymbol{X}^\top \boldsymbol{W})(\frac{1}{n}\boldsymbol{W}^\top \boldsymbol{W})^{-1} \\ \boldsymbol{B} &= \operatorname{plim} \boldsymbol{B}_n = (\boldsymbol{D}^\top \boldsymbol{Q}^{-1}\boldsymbol{D})^{-1}\boldsymbol{D}^\top \boldsymbol{Q}^{-1}\end{aligned}$$

□

Maximum number of points: 73.

References

[Seb77] G. A. F. Seber, *Linear regression analysis*, Wiley, New York, 1977. 17

Economics Department, University of Utah, 1645 Campus Center Drive, Salt Lake City, UT 84112-9300, U.S.A

E-mail address: ehrbar@econ.utah.edu

URL: http://www.econ.utah.edu/ehrbar