$$\underset{n\times 1}{\underset{\sim}{Y}} = X\underset{\sim}{b} + \underset{\sim}{e}$$

$$\mathbb{E}\,\underset{\sim}{e} = \underset{\sim}{0}$$

$$\text{Cov}\,\underset{\sim}{e} = \sigma^2 \underset{\sim}{I}$$

uncorrelated

STAT 714 fa 2025 Lec 03

Gauss-Markov model, BLUE, Aitken model, generalized least-squares

Want to estimate $\underset{\sim}{c}^T\underset{\sim}{b}$ when $\underset{\sim}{c}$ is estimable.

Karl B. Gregory

University of South Carolina

$$\underset{\sim}{x}_i = \begin{bmatrix} x_{i1} - \bar{x}_{i\cdot} \\ \vdots \\ x_{in} - \bar{x}_{i\cdot} \end{bmatrix}$$

$$\underset{\sim}{1}^T \underset{\sim}{x}_i = 0$$

$$\sum_{j=1}^{n} \left(x_{ij} - \bar{x}_{i\cdot}\right) = 0$$

$$\text{Var}\left(\underset{\sim}{c}^T \hat{\underset{\sim}{b}}\right) = ?$$

These slides are an instructional aid; their sole purpose is to display, during the lecture, definitions, plots, results, etc. which take too much time to write by hand on the blackboard. They are not intended to explain or expound on any material.

Gauss-Markov model

The *Gauss-Markov model* assumes $\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}$, where $\mathbb{E}\mathbf{e} = \mathbf{0}$ and $\operatorname{Cov}\mathbf{e} = \sigma^2\mathbf{I}_n$.

We consider the mean and variance of the LS estimator $\mathbf{c}^T\hat{\mathbf{b}}$ in this model.

We also consider how to estimate the error term variance σ^2.

Result (Must know)

For a random vector $\mathbf{y}$*, vectors* $\mathbf{a}$ *and* $\mathbf{b}$*, and a matrices* $\mathbf{A}$ *and* $\mathbf{B}$*, we have*

1. $\mathbb{E}\mathbf{a}^T\mathbf{y} = \mathbf{a}^T\mathbb{E}\mathbf{y}$.
2. $\operatorname{Var}\mathbf{a}^T\mathbf{y} = \mathbf{a}^T(\operatorname{Cov}\mathbf{y})\,\mathbf{a}$.
3. $\operatorname{Cov}(\mathbf{a}^T\mathbf{y}, \mathbf{b}^T\mathbf{y}) = \mathbf{a}^T(\operatorname{Cov}\mathbf{y})\,\mathbf{b}$.
4. $\operatorname{Cov}\mathbf{A}\mathbf{y} = \mathbf{A}(\operatorname{Cov}\mathbf{y})\mathbf{A}^T$.
5. $\operatorname{Cov}(\mathbf{A}\mathbf{y}, \mathbf{B}\mathbf{y}) = \mathbf{A}(\operatorname{Cov}\mathbf{y})\mathbf{B}^T$.

Result (Variance of estimable contrast)

Let $\mathbf{c}^T\mathbf{b}$ *be an estimable contrast in* $\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}$ *with LS estimator* $\mathbf{c}^T\hat{\mathbf{b}}$*. Then*

$$\operatorname{Var}\mathbf{c}^T\hat{\mathbf{b}} = \sigma^2\mathbf{c}^T(\mathbf{X}^T\mathbf{X})^{-}\mathbf{c}.$$

Exercise: Derive and show invariance to the choice of generalized inverse.

$$\text{Var } X = \mathbb{E}\left[(X - \mathbb{E}X)^2\right]$$

$$\text{Var } \underset{\sim}{a}^T \underset{\sim}{Y} = \text{Var}\left(\sum_{i=1}^n a_i Y_i\right)$$

$$= \mathbb{E}\left(\left(\sum_{i=1}^n a_i Y_i - \sum_{i=1}^n a_i \mathbb{E} Y_i\right)^2\right)$$

$$= \mathbb{E}\left(\sum_{i=1}^n a_i (Y_i - \mathbb{E} Y_i)\right)^2$$

$$= \mathbb{E}\left[\sum_{i=1}^n \sum_{j=1}^n a_i a_j (Y_i - \mathbb{E} Y_i)(Y_j - \mathbb{E} Y_j)\right]$$

$$= \sum_{i=1}^n \sum_{j=1}^n a_i a_j \underbrace{\mathbb{E}(Y_i - \mathbb{E} Y_i)(Y_j - \mathbb{E} Y_j)}_{\text{Cov}(Y_i, Y_j)}$$

$$= \sum_{i=1}^n \sum_{j=1}^n a_i a_j \text{Cov}(Y_i, Y_j)$$

$$= \underline{\underline{\underset{\sim}{a}^T (\text{Cov } \underset{\sim}{Y}) \underset{\sim}{a}}}$$

Let $\underset{\sim}{c}^T \underset{\sim}{b}$ be estimable. Means $\underset{\sim}{c} \in C(X^T)$.

Let $\hat{\underset{\sim}{b}}$ satisfy $X^T X \hat{\underset{\sim}{b}} = X^T y$.

Then $\text{Var}\left(\underset{\sim}{c}^T \hat{\underset{\sim}{b}}\right) = \underset{\sim}{c}^T (X^T X)^- \underset{\sim}{c}\, \sigma^2$

PROOF

Define all solutions to Normal equations as

$$\underset{\sim}{b}(\underset{\sim}{z}) = (X^T X)^- X^T \underset{\sim}{y} + \left(I - (X^T X)^- X^T X\right) \underset{\sim}{z}$$

$$\forall\ \underset{\sim}{z} \in \mathbb{R}^p$$

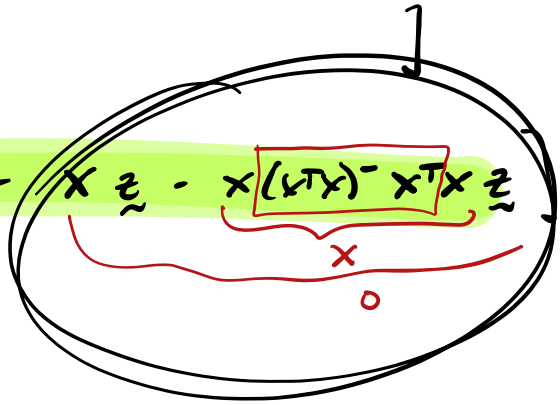
$$c^T \hat{b}(z) = c^T \left[(X^TX)^- X^T y + \left(I - (X^TX)^- X^TX\right) z \right]$$

$$= d^T X \left[\quad \cdots \right]$$

write $c = X^T d$ for some d.

$$= d^T \left[X(X^TX)^- X^T y + Xz - X(X^TX)^- X^TXz \right]$$

$$= d^T P_X y$$



$$\Rightarrow \quad \mathrm{Var}\left(c^T \hat{b}(z)\right) = \mathrm{Var}\left(d^T P_X y\right)$$

$y = Xb + e$

$$= \mathrm{Var}\left((P_X d)^T y\right)$$

$\mathrm{Var}(a^T y) = a^T (\mathrm{Cov}\, y)\, a$

$$= (P_X d)^T (\mathrm{Cov}\, y) P_X d$$

$$= d^T P_X (\sigma^2 I) P_X d$$

$$= \sigma^2 d^T P_X d$$

$$= \sigma^2 d^T X (X^TX)^- X^T d$$

$$= \sigma^2 c^T (X^TX)^- c$$

For $c \in \mathcal{C}(X^T)$,

$$\mathrm{Var}(c^T\hat{b}) = c^T(X^TX)^- c\,\sigma^2$$

$$\mathrm{E}\, c^T\hat{b} = c^T b \quad \forall \; b$$

Best linear unbiased estimator (BLUE)

A *best linear unbiased estimator* (BLUE) of an estimable contrast $\mathbf{c}^T\mathbf{b}$ is an estimator of the form $a_0 + \mathbf{a}^T\mathbf{y}$ such that $\mathbb{E}[a_0 + \mathbf{a}^T\mathbf{y}] = \mathbf{c}^T\mathbf{b}$ for all $\mathbf{b}$ and such that no other such estimator has smaller variance for any $\mathbf{b}$.

A BLUE achieves the smallest possible variance among linear unbiased estimators.

Theorem (Gauss-Markov Theorem)

Let $\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}$ *with* $\mathbb{E}\mathbf{e} = \mathbf{0}$ *and* $\operatorname{Cov}\mathbf{e} = \sigma^2\mathbf{I}_n$ *and let* $\mathbf{c}^T\mathbf{b}$ *be an estimable contrast. If* $\hat{\mathbf{b}}$ *satisfies* $\mathbf{X}^T\mathbf{X}\hat{\mathbf{b}} = \mathbf{X}^T\mathbf{y}$ *then* $\mathbf{c}^T\hat{\mathbf{b}}$ *is the BLUE for* $\mathbf{c}^T\mathbf{b}$.

See Thm 4.1 of Monahan (2008).

Prove the result.

Let $c^T b$ be estimable, so $c \in \text{Col } X^T$

$\Rightarrow$ $c = X^T d$ for some d.

Suppose $\exists$ a_0, a such that

$$E[a_0 + a^T Y] = c^T b \quad \forall \; b.$$

$$\Rightarrow \quad a_0 + a^T X b = c^T b \quad \forall \; b.$$

$$\Rightarrow \quad a_0 = 0, \quad c^T = a^T X.$$

$$\Leftrightarrow \quad \boxed{c = X^T a}$$

$$\text{Var}(a^T Y) = \text{Var}(a^T Y - c^T \hat{b} + c^T \hat{b})$$

$$= \text{Var}(a^T Y - c^T \hat{b}) + \text{Var}(c^T \hat{b}) + 2\,\text{Cov}(a^T Y - c^T \hat{b}, c^T \hat{b})$$

$$\text{Cov}(a^T Y - c^T \hat{b}, c^T \hat{b})$$

$$= \text{Cov}(a^T Y - a^T X \hat{b}, a^T X \hat{b})$$

$$= \text{Cov}(a^T Y - a^T P_X Y, a^T P_X Y)$$

$$= \text{Cov}([(I - P_X)a]^T Y, (P_X a)^T Y)$$

$X(X^TX)^- X^T Y$

$\text{Cov}(a^T Y, b^T Y) = a^T (\text{Cov } Y) b$

$$= \left[(I - P_X)\underset{\sim}{a}\right]^T \left(\mathrm{Cov}\ \underset{\sim}{Y}\right) P_X\ \underset{\sim}{a}$$

$$= \sigma^2\ \underset{\sim}{a}^T (I - P_X) P_X\ \underset{\sim}{a}$$

$$= 0$$

$\Rightarrow$

$$\mathrm{Var}\left(\underset{\sim}{a}^T \underset{\sim}{Y}\right) = \underbrace{\mathrm{Var}\left(\underset{\sim}{a}^T \underset{\sim}{Y} - \underset{\sim}{c}^T \hat{\underset{\sim}{b}}\right)}_{\geq 0} + \mathrm{Var}\left(\underset{\sim}{c}^T \hat{\underset{\sim}{b}}\right)$$

$$\Rightarrow \mathrm{Var}\left(\underset{\sim}{a}^T \underset{\sim}{Y}\right) \geq \mathrm{Var}\left(\underset{\sim}{c}^T \hat{\underset{\sim}{b}}\right).$$

$\hat{\sigma}^2 =$

Result (Expected value of a quadratic form)

Let $\mathbf{z}$ *be a random vector with* $\mathbb{E}\mathbf{z} = \boldsymbol{\mu}$ *and* $\operatorname{Cov}\mathbf{z} = \boldsymbol{\Sigma}$. *Then*

$$\mathbb{E}\mathbf{z}^T\mathbf{A}\mathbf{z} = \boldsymbol{\mu}^T\mathbf{A}\boldsymbol{\mu} + \operatorname{tr}(\mathbf{A}\boldsymbol{\Sigma}).$$

See Lem 4.1 of Monahan (2008).

Can be used to prove the following result.

Result (Unbiased estimator of the variance)

Let $\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}$ *with* $\mathbb{E}\mathbf{e} = \mathbf{0}$ *and* $\operatorname{Cov}\mathbf{e} = \sigma^2\mathbf{I}_n$ *and let* $\mathbf{X}$ *have rank* r. *Then*

$$\hat{\sigma}^2 = \frac{\|\hat{\mathbf{e}}\|^2}{n-r}, \quad \text{where } \hat{\mathbf{e}} = (\mathbf{I} - \mathbf{P}_\mathbf{X})\mathbf{y},$$

is an unbiased estimator of σ^2.

$\underset{\sim}{y} = \hat{\underset{\sim}{y}} + \hat{\underset{\sim}{e}}$

$\hat{\underset{\sim}{y}} = P_x \underset{\sim}{y} = X\hat{\underset{\sim}{b}}$

$\hat{\underset{\sim}{e}} = (I - P_x)\underset{\sim}{y}$

See Res 4.2 of Monahan (2008).

Prove both results.

For a random vector $\underset{\sim}{z}$ with $\mathbb{E}\,\underset{\sim}{z} = \underset{\sim}{\mu}$, $\mathrm{Cov}\,\underset{\sim}{z} = \Sigma$,

we have

$$\mathbb{E}\,\underset{\sim}{z}^T A \underset{\sim}{z} = \underset{\sim}{\mu}^T A \underset{\sim}{\mu} + \mathrm{tr}(A\Sigma).$$

Proof:

Write

$$\mathbb{E}\left[(\underset{\sim}{z}-\underset{\sim}{\mu})^T A (\underset{\sim}{z}-\underset{\sim}{\mu})\right] = \mathbb{E}\left[\underset{\sim}{z}^T A \underset{\sim}{z} + \underset{\sim}{\mu}^T A \underset{\sim}{\mu} - 2\,\underset{\sim}{z}^T A \underset{\sim}{\mu}\right]$$

$$= \mathbb{E}\left[\underset{\sim}{z}^T A \underset{\sim}{z}\right] - \underset{\sim}{\mu}^T A \underset{\sim}{\mu}.$$

$\Rightarrow$

$$\mathbb{E}\,\underset{\sim}{z}^T A \underset{\sim}{z} = \mathbb{E}\,(\underset{\sim}{z}-\underset{\sim}{\mu})^T A (\underset{\sim}{z}-\underset{\sim}{\mu}) + \underset{\sim}{\mu}^T A \underset{\sim}{\mu}.$$

$$= \mathbb{E}\,\mathrm{tr}(\underset{\sim}{z}-\underset{\sim}{\mu})^T A (\underset{\sim}{z}-\underset{\sim}{\mu}) + \underset{\sim}{\mu}^T A \underset{\sim}{\mu}.$$

$$= \mathbb{E}\,\mathrm{tr}\left[(\underset{\sim}{z}-\underset{\sim}{\mu})(\underset{\sim}{z}-\underset{\sim}{\mu})^T A\right] + \underset{\sim}{\mu}^T A \underset{\sim}{\mu}$$

$$= \mathrm{tr}\left[\mathbb{E}\,(\underset{\sim}{z}-\underset{\sim}{\mu})(\underset{\sim}{z}-\underset{\sim}{\mu})^T A\right] + \underset{\sim}{\mu}^T A \underset{\sim}{\mu}$$

($\mathbb{E}\,(\underset{\sim}{z}-\underset{\sim}{\mu})(\underset{\sim}{z}-\underset{\sim}{\mu})^T$ is $n \times n$)

$$= \mathrm{tr}(\Sigma A) + \underset{\sim}{\mu}^T A \underset{\sim}{\mu}.$$

$\mathrm{tr}(AB) = \mathrm{tr}(BA)$

$$\mathbb{E}\,\hat{\sigma}^2 = \mathbb{E}\left[\frac{\|\hat{\underset{\sim}{e}}\|^2}{n-r}\right] \qquad \|\underset{\sim}{a}\|^2 = \underset{\sim}{a}^T\underset{\sim}{a}$$

$$= \mathbb{E}\ \|(I-P_X)\underset{\sim}{y}\|^2 \,/\,(n-r)$$

$$= \mathbb{E}\left[\underset{\sim}{y}^T(I-P_X)\underset{\sim}{y}\right] / (n-r)$$

$$= \left[(\mathbb{E}\underset{\sim}{y})^T(I-P_X)\,\mathbb{E}\underset{\sim}{y} + tr\left[(Cov\,\underset{\sim}{y})(I-P_X)\right]\right] / (n-r) \qquad (Cov\,\underset{\sim}{y} = \sigma^2 I)$$

$$= \left[\underbrace{(X\underset{\sim}{b})^T(I-P_X)X\underset{\sim}{b}}_{=0} + \sigma^2\, tr\,(I-P_X)\right] / (n-r)$$

$P_X X\underset{\sim}{b} =$

$$= \frac{\sigma^2}{n-r}\, tr\,(I-P_X)$$

$$= \sigma^2.$$

$$tr\,(I-P_X) = tr\,I - tr\,P_X = n - tr\,P_X = n-r.$$

P_X: idempotent, so eigenvalues = 0 or 1

symm., so rank = # nonzero eigenvalues

$tr\,P_X$ = sum of eigenvalues = # nonzero

P_X projects onto its <u>own</u> column space, and this is

$\mathrm{Col}\, X$. & $\mathrm{Col}\, P_X = \mathrm{Col}\, X$.

$\Rightarrow$ $\dim \mathrm{Col}\, P_X = \dim \mathrm{Col}\, X$

$\mathrm{rank}\, P_X = \mathrm{rank}\, X = r$.

A setup for "underfitting" or omitting important covariates

Suppose $\mathbf{y}$ is truly given by

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \boldsymbol{\eta} + \mathbf{e}, \quad \mathbb{E}\mathbf{e} = 0, \quad \operatorname{Cov}\mathbf{e} = \sigma^2 \mathbf{I}_n,$$

but one assumes the model $\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}$ and finds $\hat{\mathbf{b}}$ satisfying $\mathbf{X}^T\mathbf{X}\hat{\mathbf{b}} = \mathbf{X}^T\mathbf{y}$.

So $\boldsymbol{\eta}$ represents the effects of omitted covariates.

Exercise: In the above setup find the expected value of

1. $\mathbf{c}^T\hat{\mathbf{b}}$, where $\mathbf{c} \in \operatorname{Col}\mathbf{X}^T$
2. $\hat{\sigma}^2 = \|\mathbf{y} - \mathbf{X}\hat{\mathbf{b}}\|^2/(n-r)$, where $r = \operatorname{rank}\mathbf{X}$.

In each case give a condition on $\boldsymbol{\eta}$ under which the estimator will be unbiased.

LOOK AT 0.4.0. on your own.

0.4.0.

A setup for "overfitting" or including unimportant covariates

Suppose $\mathbf{y}$ is truly given by

$$\mathbf{y} = \mathbf{X}_1\mathbf{b}_1 + \mathbf{e}, \quad \mathbb{E}\mathbf{e} = 0, \quad \operatorname{Cov}\mathbf{e} = \sigma^2\mathbf{I}_n,$$

but one assumes $\mathbf{y} = \mathbf{X}_1\mathbf{b}_1 + \mathbf{X}_2\mathbf{b}_2 + \mathbf{e}$ and finds $\hat{\mathbf{b}} = [\mathbf{b}_1^T \ \mathbf{b}_2^T]^T$ satisfying

$$\mathbf{X}^T\mathbf{X}\hat{\mathbf{b}} = \mathbf{X}^T\mathbf{y}, \quad \text{where} \quad \mathbf{X} = [\mathbf{X}_1 \ \mathbf{X}_2].$$

In truth $\mathbf{b}_2 = \mathbf{0}$, so $\mathbf{X}_2$ contains unimportant (extra) covariate information.

Exercise: In the above setup, assume (for simplicity) that $\mathbf{X}$ is full-rank. Then:

1. Find $\mathbb{E}\mathbf{c}_1^T\hat{\mathbf{b}}_1$ and $\operatorname{Var}\mathbf{c}_1^T\hat{\mathbf{b}}_1$
2. In what situation does adding "extra" covariates have no effect on $\operatorname{Var}\mathbf{c}_1^T\hat{\mathbf{b}}_1$?
3. Find $\mathbb{E}\|(\mathbf{I} - \mathbf{P}_{\mathbf{X}})\mathbf{y}\|^2/(N - \operatorname{rank}\mathbf{X})$ and $\mathbb{E}\|(\mathbf{I} - \mathbf{P}_{\mathbf{X}_1})\mathbf{y}\|^2/(N - \operatorname{rank}\mathbf{X}_1)$.

G-M: $\underset{\sim}{y} = x\underset{\sim}{b} + \underset{\sim}{e}$, $\mathbb{E}\,\underset{\sim}{e} = \underset{\sim}{0}$, $\mathrm{Cov}\,\underset{\sim}{e} = \sigma^2 I_n$

Aitken: $\underset{\sim}{y} = x\underset{\sim}{b} + \underset{\sim}{e}$, $\mathbb{E}\,\underset{\sim}{e} = \underset{\sim}{0}$, $\mathrm{Cov}\,\underset{\sim}{e} = V\sigma^2$

Aitken model

The *Aitken model* assumes $\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}$, where $\mathbb{E}\mathbf{e} = \mathbf{0}$ and $\operatorname{Cov}\mathbf{e} = \sigma^2\mathbf{V}$, $\mathbf{V}$ pd.

The ordinary least-squares estimator $\mathbf{c}^T\mathbf{b}$ may not be the BLUE anymore...

Generalized least-squares estimator of an estimable contrast

Under the Aitken model, the *GLS estimator* of a contrast is $\mathbf{c}^T\hat{\mathbf{b}}_{\text{gls}}$, where $\hat{\mathbf{b}}_{\text{gls}}$ is any vector such that $(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})\hat{\mathbf{b}}_{\text{gls}} = \mathbf{X}^T\mathbf{V}^{-1}\mathbf{y}$.

Generalized Normal Eqs.

Exercise: Show that $\hat{\mathbf{b}}_{\text{gls}}$ minimizes $Q_{\mathbf{V}}(\mathbf{b}) = (\mathbf{y} - \mathbf{X}\mathbf{b})^T\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\mathbf{b})$.

$Q(\underset{\sim}{b}) = (\underset{\sim}{y} - X\underset{\sim}{b})^T(\underset{\sim}{y} - X\underset{\sim}{b})$

$$Q_V(\underset{\sim}{b}) = \underset{\sim}{y}^T V^{-1} \underset{\sim}{y} + \underset{\sim}{b}^T X^T V^{-1} X \underset{\sim}{b} - 2\, \underset{\sim}{y}^T V^{-1} X \underset{\sim}{b}$$

$$\frac{\partial}{\partial \underset{\sim}{b}} Q_V(\underset{\sim}{b}) = 2\, X^T V^{-1} X\, \underset{\sim}{b} - 2\, X^T V^{-1} \underset{\sim}{y} \overset{\text{set}}{=} \underset{\sim}{0}$$

$$\Rightarrow \quad \boxed{X^T V^{-1} X\, \underset{\sim}{b} = X^T V^{-1} \underset{\sim}{y}}$$

gen. Norm. eqs.

We will need these results for establishing some properties of the GLS estimator.

Result ("Generalized cool result" and another gen. inverse of $\mathbf{X}$)

1. *If* $\mathbf{V}$ *is positive definite there exists nonsingular* $\mathbf{R}$ *such that* $\mathbf{RR} = \mathbf{V}^{-1}$.
2. $\text{Nul}\, \mathbf{X}^T\mathbf{V}^{-1}\mathbf{X} = \text{Nul}\, \mathbf{X}$.
3. $\mathbf{X}^T\mathbf{V}^{-1}\mathbf{XA} = \mathbf{X}^T\mathbf{V}^{-1}\mathbf{XB} \iff \mathbf{XA} = \mathbf{XB}$
4. $(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^-\mathbf{X}^T\mathbf{V}^{-1}$ *is a generalized inverse of* $\mathbf{X}$.

Prove the results.

(1) V p.d. $\Rightarrow$ $\exists$ R nonsing, symm, s.t. $RR = V^{-1}$.

Proof V p.d. $\Rightarrow$ $V = PDP^T$ where $P^T = P^{-1}$ s. $P^TP = PP^T = I$.

D diag. with positive diagonals.

Why do we want Cov. matrices to be p.d.?

Suppose

$\text{Cov}\, \underset{\sim}{y} = V$.

Then

$\text{Var}(\underset{\sim}{a}^T \underset{\sim}{y}) = \underset{\sim}{a}^T V \underset{\sim}{a} \overset{\text{want}}{>} 0$

$P D^{-1} P^T V = P D^{-1} \overbrace{P^T P}^{I} D P^T$

$= P D^{-1} D P^T$

take reciprocal of diags $= P P^T = I$

$\Rightarrow \quad V^{-1} = P D^{-1} P^T$

$= P D^{-1/2} D^{-1/2} P^T$

$= \underbrace{P D^{-1/2} P^T}_{I} \underbrace{P D^{-1/2} P^T}$

$= R R$

$D = \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{bmatrix}$

$D^{-1} = \begin{bmatrix} \frac{1}{d_1} & & \\ & \ddots & \\ & & \frac{1}{d_n} \end{bmatrix}$

$D^{-1/2} = \begin{bmatrix} \frac{1}{\sqrt{d_1}} & & \\ & \ddots & \\ & & \frac{1}{\sqrt{d_n}} \end{bmatrix}$

$R = P D^{-1/2} P^T$

Symmetric, nonsing, square root matrix for V^{-1}.

$R^{-1} = P D^{1/2} P^T$

because $P D^{1/2} P^T P D^{-1/2} P^T = P D^{1/2} D^{-1/2} P^T$

$= P P^T = I$.

(2) $\text{Nul}\, X^T V^{-1} X = \text{Nul}\, X$

(Prev. $\text{Nul}\, X^T X = \text{Nul}\, X$)

"$\subseteq$" $\quad \underset{\sim}{x} \in \text{Nul}\, X^T V^{-1} X$

$\Rightarrow \quad X^T V^{-1} X \underset{\sim}{x} = \underset{\sim}{0}$

$\Rightarrow \quad \underset{\sim}{x}^T X^T V^{-1} X \underset{\sim}{x} = 0$

R symm

$$\Rightarrow \quad \underset{\sim}{x}^T X^T R R X \underset{\sim}{x} = 0$$

$$\Rightarrow \quad \| R X \underset{\sim}{x} \|^2 = 0$$

$$\Rightarrow \quad R X \underset{\sim}{x} = 0 \qquad (\text{multiply by } R^{-1})$$

$$\Rightarrow \quad X \underset{\sim}{x} = 0$$

$$\Rightarrow \quad \underset{\sim}{x} \in \text{Nul } X.$$

"$\supset$" $\quad \underset{\sim}{x} \in \text{Nul } X \Rightarrow X \underset{\sim}{x} = \underset{\sim}{0}$

$$\Rightarrow \quad X^T V^{-1} X \underset{\sim}{x} = \underset{\sim}{0}.$$

$$\Rightarrow \quad \underset{\sim}{x} \in \text{Nul } X^T V^{-1} X.$$

(3)

$$X^T V^{-1} X A = X^T V^{-1} X B$$

$$\Leftrightarrow \quad X^T V^{-1} X (A - B) = 0$$

$$\Leftrightarrow \quad \text{Cols of } A-B \in \text{Nul } X^T V^{-1} X = \text{Nul } X$$

$$\Leftrightarrow \quad X(A-B) = 0$$

$$\Leftrightarrow \quad XA = XB.$$

4. $(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^-\mathbf{X}^T\mathbf{V}^{-1}$ *is a generalized inverse of* $\mathbf{X}$.

(4) $X^T V^{-1} X \underbrace{(X^T V^{-1} X)^- X^T V^{-1} X}_{A} = X^T V^{-1} X \underbrace{I}_{B}$

$\Rightarrow \quad X \underbrace{(X^T V^{-1} X)^- X^T V^{-1}}_{\text{gen-inv of } X} X = X$

ALSO:

$N(1 \;\; X^T V^{-1} X) = N(1 \;\; X)$

$\Leftrightarrow \quad (N(1 \;\; X^T V^{-1} X))^\perp = (N(1 \;\; X))^\perp$

$C(1 \;\; X^T V^{-1} X) = C(1 \;\; X^T)$

$(C(1 \;\; X))^\perp = N(1 \;\; X^T)$

$(C(1 \;\; X^T))^\perp = N(1 \;\; X)$

$C(1 \;\; X^T) = (N(1 \;\; X))^\perp$

$C(1 \;\; X^T V^{-1} X) = C(1 \;\; X^T)$

Note: Estimability of $\mathbf{c}^T\mathbf{b}$ is still equivalent to $\mathbf{c} \in \text{Col}\,\mathbf{X}^T$.

Result (Properties of the GLS estimator of an estimable contrast)

Let $\mathbf{c}^T\mathbf{b}$ be an estimable contrast. Then the GLS estimator $\mathbf{c}^T\hat{\mathbf{b}}_{\text{gls}}$

1. *is invariant to the choice of $\hat{\mathbf{b}}_{\text{gls}}$ which satisfies $(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})\hat{\mathbf{b}}_{\text{gls}} = \mathbf{X}^T\mathbf{V}^{-1}\mathbf{y}$.*
2. *has expected value equal to $\mathbf{c}^T\mathbf{b}$ for all $\mathbf{b}$.*
3. *has variance $\text{Var}\,\mathbf{c}^T\hat{\mathbf{b}}_{\text{gls}} = \mathbf{c}^T(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^-\mathbf{c}\sigma^2$*

Prove the results.

(1) $\underset{\sim}{c}^T\underset{\sim}{b}$ estimable $\Rightarrow$ $\underset{\sim}{c} \in \text{Col}\,X^T \Rightarrow$ $\underset{\sim}{c} = X^T\underset{\sim}{d}$ for some $\underset{\sim}{d}$.

Let $\hat{\underset{\sim}{b}}_1$ and $\hat{\underset{\sim}{b}}_2$ solve the gen. norm. eqs.

I.e.

$$(X^T V^{-1} X)\,\hat{b}_1 = X^T V^{-1} y$$

$$(X^T V^{-1} X)\,\hat{b}_2 = X^T V^{-1} y$$

$$\Rightarrow \quad (X^T V^{-1} X)(\hat{b}_1 - \hat{b}_2) = 0$$

$$\Rightarrow \quad \hat{b}_1 - \hat{b}_2 \in \text{Nul } X^T V^{-1} X = \text{Nul } X$$

$$\Rightarrow \quad X(\hat{b}_1 - \hat{b}_2) = 0$$

$$\Rightarrow \quad X\hat{b}_1 = X\hat{b}_2 .$$

Now $\quad c^T \hat{b}_1 = d^T X \hat{b}_1 = d^T X \hat{b}_2 = c^T \hat{b}_2 .$

② $\quad E\, c^T \hat{b}_{gls} = E\, d^T X \hat{b}_{gls}$

$$= E\, d^T X (X^T V^{-1} X)^- X^T V^{-1} X \hat{b}_{gls}$$

$$= E\, d^T X (X^T V^{-1} X)^- X^T V^{-1} y$$

$$= d^T X (X^T V^{-1} X)^- X^T V^{-1} X b$$

$$= d^T X b$$

$= \underset{\sim}{c}^T \underset{\sim}{b}$

(3)

$$\mathrm{Var}\left(\underset{\sim}{c}^T \hat{\underset{\sim}{b}}_{gls}\right) = \mathrm{Var}\left(\underset{\sim}{d}^T X (X^T V^{-1} X)^- X^T V^{-1} \underset{\sim}{y}\right)$$

assume symmetry

$$= \underset{\sim}{d}^T X (X^T V^{-1} X)^- X^T V^{-1} (V\sigma^2) V^{-1} X (X^T V^{-1} X)^- X^T \underset{\sim}{d}$$

$$= \sigma^2 \underset{\sim}{d}^T X (X^T V^{-1} X)^- X^T V^{-1} X (X^T V^{-1} X)^- X^T \underset{\sim}{d}$$

$$= \sigma^2 \underset{\sim}{d}^T X (X^T V^{-1} X)^- X^T \underset{\sim}{d}$$

$$= \sigma^2 \underset{\sim}{c}^T (X^T V^{-1} X)^- \underset{\sim}{c} .$$

Theorem (Aitken's Theorem)

Let $\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}$, *where* $\mathbb{E}\mathbf{e} = \mathbf{0}$ *and* $\operatorname{Cov}\mathbf{e} = \sigma^2\mathbf{V}$, $\mathbf{V}$ *pd, and let* $\mathbf{c}^T\mathbf{b}$ *be an estimable contrast. Then* $\mathbf{c}^T\hat{\mathbf{b}}_{\text{gls}}$ *is the BLUE for* $\mathbf{c}^T\mathbf{b}$.

See Thm 4.2 of Monahan (2008).

Prove the result.

Let $a_0 + \underset{\sim}{a}^T\underset{\sim}{y}$ be a L.U.E. so

$$\mathbb{E}\left[a_0 + \underset{\sim}{a}^T\underset{\sim}{y}\right] = \underset{\sim}{c}^T\underset{\sim}{b} \quad \forall\ \underset{\sim}{b}.$$

$$\Rightarrow \quad a_0 + \underset{\sim}{a}^T\mathbf{X}\underset{\sim}{b} = \underset{\sim}{c}^T\underset{\sim}{b} \quad \forall\ \underset{\sim}{b}$$

$\Rightarrow \quad a_0 = 0 \;, \quad \underset{\sim}{a}^T X = \underset{\sim}{c}^T$

$$\mathrm{Var}(X+Y) = \mathrm{Var}\,X + \mathrm{Var}\,Y + 2\,\mathrm{Cov}(X,Y)$$

Now, write

$$\mathrm{Var}\left(\underset{\sim}{a}^T \underset{\sim}{Y}\right) = \mathrm{Var}\left(\underset{\sim}{a}^T \underset{\sim}{Y} - \underset{\sim}{c}^T \hat{\underset{\sim}{b}}_{gls} + \underset{\sim}{c}^T \hat{\underset{\sim}{b}}_{gls}\right)$$

$$= \underbrace{\mathrm{Var}\left(\underset{\sim}{a}^T \underset{\sim}{Y} - \underset{\sim}{c}^T \hat{\underset{\sim}{b}}_{gls}\right)}_{\geq 0} + \mathrm{Var}\left(\underset{\sim}{c}^T \hat{\underset{\sim}{b}}_{gls}\right)$$

$$+ 2\,\mathrm{Cov}\left(\underset{\sim}{a}^T \underset{\sim}{Y} - \underset{\sim}{c}^T \hat{\underset{\sim}{b}}_{gls},\; \underset{\sim}{c}^T \hat{\underset{\sim}{b}}_{gls}\right)$$

Still to show $\mathrm{Cov}\left(\underset{\sim}{a}^T \underset{\sim}{Y} - \underset{\sim}{c}^T \hat{\underset{\sim}{b}}_{gls},\; \underset{\sim}{c}^T \hat{\underset{\sim}{b}}_{gls}\right) = 0.$

$$\mathrm{Cov}\left(\underset{\sim}{a}^T \underset{\sim}{Y} - \underset{\sim}{c}^T \hat{\underset{\sim}{b}}_{gls},\; \underset{\sim}{c}^T \hat{\underset{\sim}{b}}_{gls}\right) \qquad \underset{\sim}{c}^T = \underset{\sim}{a}^T X$$

$$= \mathrm{Cov}\left(\underset{\sim}{a}^T \underset{\sim}{Y} - \underset{\sim}{a}^T X \hat{\underset{\sim}{b}}_{gls},\; \underset{\sim}{a}^T X \hat{\underset{\sim}{b}}_{gls}\right)$$

$$= \mathrm{Cov}\left(\underset{\sim}{a}^T \underset{\sim}{Y} - \underset{\sim}{a}^T X \underbrace{(X^T V^{-1} X)^- X^T V^{-1}}_{X^T V^{-1} \underset{\sim}{Y}} X \hat{\underset{\sim}{b}}_{gls},\; \underset{\sim}{a}^T X \underbrace{(X^T V^{-1} X)^- X^T V^{-1}}_{X^T V^{-1} \underset{\sim}{Y}} X \hat{\underset{\sim}{b}}_{gls}\right)$$

$$= \mathrm{Cov}\left(\underset{\sim}{a}^T \underset{\sim}{Y} - \underset{\sim}{a}^T X (X^T V^{-1} X)^- X^T V^{-1} \underset{\sim}{Y},\; \underset{\sim}{a}^T X (X^T V^{-1} X)^- X^T V^{-1} \underset{\sim}{Y}\right)$$

$$= \mathrm{Cov}\left(\underset{\sim}{a}^T R^{-1} R \underset{\sim}{Y} - \underset{\sim}{a}^T R^{-1} R X \left(X^T R R X\right)^- X^T R R \underset{\sim}{Y},\right.$$

$$\left. \underset{\sim}{a}^T R^{-1} R X \left(X^T R R X\right)^- X^T R R \underset{\sim}{Y}\right)$$

$$= Cov\left(a^T R^{-1} z - a^T R^{-1} U (U^T U)^{-} U^T z, \right.$$

$$\left. a^T R^{-1} U (U^T U)^{-} U^T z \right)$$

$$= Cov\left(a^T R^{-1} (I - P_U) z, \; a^T R^{-1} P_U z \right)$$

$$= a^T R^{-1} (I - P_U) \underbrace{Cov(z)}_{\sigma^2 I} P_U R^{-1} a$$

$$= a^T R^{-1} \underbrace{(I - P_U) P_U}_{0} R^{-1} a \; \sigma^2$$

$$= 0.$$

$$V^{-1} = R R$$

$$y = X b + e$$

$$Ry = R X b + R e$$

$$z = U b + g$$

$$z = Ry, \; U = RX, \; g = Re.$$

$$y = R^{-1} z \qquad X = R^{-1} U$$

$$P_U = U (U^T U)^{-} U^T$$

$$Cov(z) = Cov(Ry)$$

$$= R \, Cov\, y \, R$$

$$= R \, V \, R \, \sigma^2$$

$$= R \, (V^{-1})^{-1} R \, \sigma^2$$

$$= R \, (RR)^{-1} R \, \sigma^2$$

$$= I \, \sigma^2$$

Result (Unbiased estimator of the variance in the Aitken model)

Let $\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}$ *with* $\mathbb{E}\mathbf{e} = \mathbf{0}$ *and* $\operatorname{Cov}\mathbf{e} = \sigma^2\mathbf{V}$ *and let* $\mathbf{X}$ *have rank* r. *Then*

$$\hat{\sigma}^2_{\text{gls}} = \frac{1}{n-r}(\mathbf{y} - \mathbf{X}\hat{\mathbf{b}}_{\text{gls}})^T\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\hat{\mathbf{b}}_{\text{gls}})$$

is an unbiased estimator of σ^2.

Prove the result.

$$\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}\hat{\mathbf{b}}_{\text{gls}} = \mathbf{X}^T\mathbf{V}^{-1}\mathbf{y}$$

$$\mathbb{E}\,\hat{\sigma}^2_{\text{gls}}$$

$$= \frac{1}{n-r}\,\mathbb{E}\left[\left(\mathbf{y} - \mathbf{X}\left(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}\hat{\mathbf{b}}_{\text{gls}}\right)^T\mathbf{V}^{-1}\left(\mathbf{y} - \mathbf{X}(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}\hat{\mathbf{b}}_{\text{gls}}\right)\right]$$

$$= \frac{1}{n-r} \mathbb{E}\left[\left(y - X(X^T V^{-1} X)^{-} X^T V^{-1} y \right)^T V^{-1} \left(y - X(X^T V^{-1} X)^{-} X^T V^{-1} y \right) \right]$$

(RR) (RR) (RR)

$$= \frac{1}{n-r} \mathbb{E}\left[\left(R^{-1} z - R^{-1} U (U^T U)^{-} U^T z \right)^T RR \left(R^{-1} z - R^{-1} U (U^T U)^{-} U^T z \right) \right]$$

P_U P_U

$$= \frac{1}{n-r} \mathbb{E}\left[(z - P_U z)^T (z - P_U z) \right]$$

$$= \frac{1}{n-r} \mathbb{E}\left[z^T (I - P_U) z \right]$$

$$= \frac{1}{n-r} \left[(Ub)^T (I - P_U) Ub + tr\left((I - P_U) \sigma^2 \right) \right]$$

0

$$= \frac{\sigma^2}{n-r} tr(I - P_U)$$

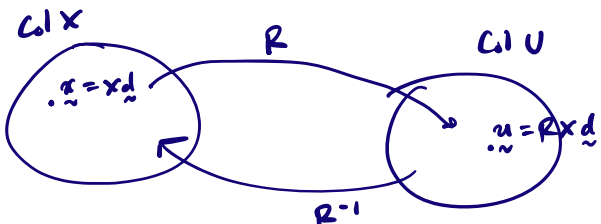
$$= \frac{\sigma^2}{n-r} \left[n - tr(P_U) \right]$$

$= \text{rank } P_U$ $= \text{rank } U$ $= \text{rank } RX$ $= \text{rank } X$

$$= \frac{\sigma^2}{n-r} (n - r)$$

We know rank U = rank X because the transformation R which takes vectors in Col X and returns vectors in Col U is 1:1.

$$= \sigma^2$$



$$\mathbb{E}\, z^T A z = (\mathbb{E} z)^T A\, \mathbb{E} z + tr(A\, Cov\, z)$$

$V^{-1} = RR$

$z = Ry \iff y = R^{-1} z$

$U = RX \iff X = R^{-1} U$

$$X^T V^{-1} X = X^T RRX = (RX)^T RX = U^T U$$

$$\mathbb{E}\, z = \mathbb{E}\, Ry = RXb = Ub$$

$$Cov\, z = Cov(Ry) = R\, Cov\, y\, R = R V R\, \sigma^2 \quad (V^{-1} = RR)$$

$$= R (RR)^{-1} R\, \sigma^2 = I\, \sigma^2$$

Let $\underset{\sim}{x} = X\underset{\sim}{d}$ some $\underset{\sim}{d}$. Then $\underset{\sim}{u} = R\underset{\sim}{x} = RX\underset{\sim}{d} = U\underset{\sim}{d} \in \text{Col } U$.

Also $R^{-1}\underset{\sim}{u} = \underset{\sim}{x}$.

Exercise: Let $Y_i = x_i\beta + \varepsilon_i|x_i|$, $i = 1, \ldots, n$, where $\varepsilon_1, \ldots, \varepsilon_n \overset{\text{ind}}{\sim} \text{Normal}(0, \sigma^2)$.

1. Write the model in matrix form as an Aitken model.
2. Give $\hat{\beta}_{\text{gls}}$ as well as $\hat{\beta}_{\text{ols}}$.
3. [redacted]
4. [redacted]
5. Give $\mathbb{E}\hat{\beta}_{\text{gls}}$ as well as $\mathbb{E}\hat{\beta}_{\text{ols}}$. — both unbiased
6. Give $\text{Var}\,\hat{\beta}_{\text{gls}}$ as well as $\text{Var}\,\hat{\beta}_{\text{ols}}$.
7. Show that $\text{Var}\,\hat{\beta}_{\text{ols}} \geq \text{Var}\,\hat{\beta}_{\text{gls}}$ using the Cauchy-Schwarz inequality.

(1) $\underset{\sim}{Y} = X\underset{\sim}{b} + \underset{\sim}{e}$ $\quad \mathbb{E}\,\underset{\sim}{e} = \underset{\sim}{0}$ $\quad \text{Cov}\,\underset{\sim}{e} = \sigma^2 V$

$$\underset{\sim}{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \qquad \underset{n\times 1}{X} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \qquad \underset{\sim}{b} = \beta$$

$$\underset{\sim}{e} = \begin{bmatrix} |x_1|\,\varepsilon_1 \\ \vdots \\ |x_n|\,\varepsilon_n \end{bmatrix}$$

$$\text{Cov}\,\underset{\sim}{e} = \sigma^2 \underbrace{\begin{bmatrix} x_1^2 & & \\ & x_2^2 & \\ & & \ddots \\ & & & x_n^2 \end{bmatrix}}_{n\times n} = V$$

(Cov($|x_1|\varepsilon_1, |x_1|\varepsilon_1$))

(2)

$$\hat{\beta}_{ols} = \left(X^T X\right)^{-1} X^T \underset{\sim}{y} = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}$$

$$\left(X^T V^{-1} X\right) \underset{\sim}{\hat{b}}_{gls} = X^T V^{-1} \underset{\sim}{y}$$

$$X^T V^{-1} X = [x_1 \cdots x_n] \begin{bmatrix} \frac{1}{x_1^2} & & \\ & \ddots & \\ & & \frac{1}{x_n^2} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = n$$

$$X^T V^{-1} \underset{\sim}{y} = [x_1 \cdots x_n] \begin{pmatrix} \frac{1}{x_1^2} & & \\ & \ddots & \\ & & \frac{1}{x_n^2} \end{pmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \sum_{i=1}^{n} \frac{x_i y_i}{x_i^2} = \sum_{i=1}^{n} \frac{y_i}{x_i}$$

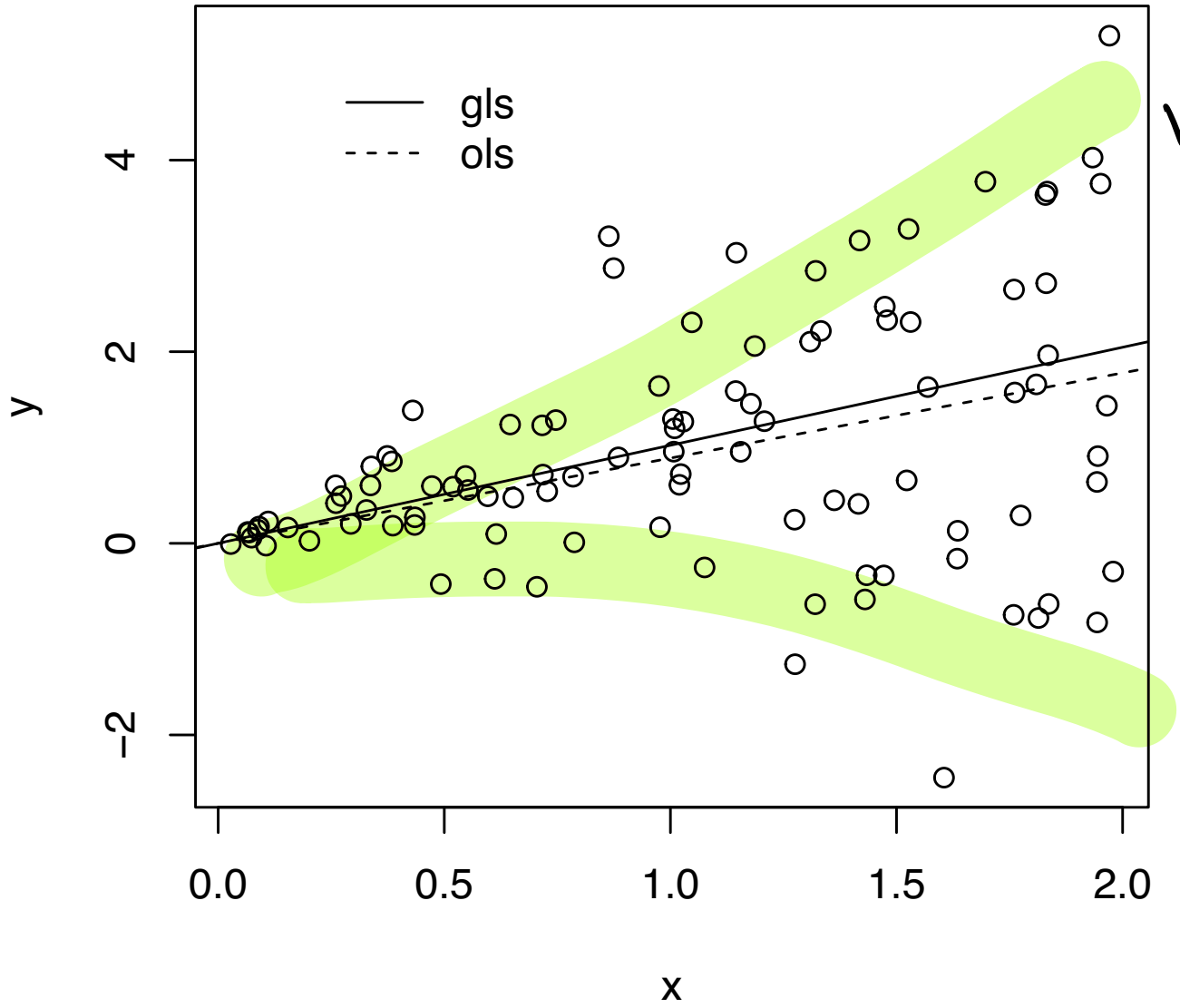
$$\hat{\beta}_{gls} = \left(X^T V^{-1} X\right)^{-1} X^T V^{-1} y = \frac{1}{n} \sum_{i=1}^{n} \frac{y_i}{x_i}$$

$Y_i = x_i \beta + |x_i| \varepsilon_i$

$\text{Cov } e = \sigma^2 I$

$\text{Cov } e_N = \sigma^2 V$

$V = \begin{bmatrix} x_1^2 & & \\ & \ddots & \\ & & x_n^2 \end{bmatrix}$



$\varepsilon_1, \ldots, \varepsilon_n \overset{ind}{\sim} \text{Normal}(0, \sigma^2)$

$$\text{Var}\ \hat{\beta}_{ols} = \text{Var}\left(\frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2} \right)$$

$y_i = x_i \beta + |x_i| \varepsilon_i$

$\text{Var}\ y_i = x_i^2 \sigma^2$

$$= \frac{1}{\left(\sum_{i=1}^n x_i^2\right)^2} \sum_{i=1}^n x_i^2 \text{Var}\ y_i$$

$$= \sigma^2 \frac{\sum_{i=1}^n x_i^4}{\left(\sum_{i=1}^n x_i^2\right)^2}$$

$$\text{Var}\ \hat{\beta}_{gls} = \text{Var}\left(\frac{1}{n} \sum_{i=1}^n \frac{y_i}{x_i} \right)$$

$$= \frac{1}{n^2} \sum_{i=1}^n \frac{1}{x_i^2} \text{Var}\ y_i$$

$$= \frac{1}{n^2} \sum_{i=1}^n \frac{x_i^2}{x_i^2} \sigma^2$$

$$= \frac{\sigma^2}{n}$$

Theory:

$$\text{Var}\ \hat{\beta}_{ols} \geq \text{Var}\ \hat{\beta}_{gls}$$

$\underset{\sim}{a}^T \underset{\sim}{b} \leq \|\underset{\sim}{a}\| \|\underset{\sim}{b}\|$

$$\sum_{i=1}^n x_i^2 = \sum_{i=1}^n 1 \cdot x_i^2$$

$$\leq \sqrt{\sum_{i=1}^n 1^2} \sqrt{\sum_{i=1}^n x_i^4}$$

$$\Rightarrow \left(\sum_{i=1}^n x_i^2\right)^2 \leq n \sum_{i=1}^n x_i^4$$

$$\frac{\sum_{i=1}^n x_i^4}{\left(\sum_{i=1}^n x_i^2\right)^2} \geq \frac{1}{n}$$

Go back to the Gauss-Markov model, impose $\mathbf{P}^T\mathbf{b} = \boldsymbol{\delta}$, and consider BLUE...

Theorem (Best linear unbiased estimator in the restricted G-M model)

Let $\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}$, *where* $\mathbb{E}\mathbf{e} = \mathbf{0}$ *and* $\operatorname{Cov}\mathbf{e} = \sigma^2\mathbf{I}_n$, *with the restriction* $\mathbf{P}^T\mathbf{b} = \boldsymbol{\delta}$.

Let $\hat{\mathbf{b}}_H$ *and* $\hat{\mathbf{u}}$ *satisfy* $\begin{bmatrix} \mathbf{X}^T\mathbf{X} & \mathbf{P} \\ \mathbf{P}^T & \mathbf{0} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{b}}_H \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}^T\mathbf{y} \\ \boldsymbol{\delta} \end{bmatrix}$ *and let* $\mathbf{c}^T\mathbf{b}$ *be estimable.*

Then $\mathbf{c}^T\hat{\mathbf{b}}_H$ *is the BLUE for* $\mathbf{c}^T\mathbf{b}$ *in the restricted model.*

See Res 4.5 of Monahan (2008).

Prove the above in the steps:

1. Show consistency of the equations $\begin{bmatrix} \mathbf{X}^T\mathbf{X} & \mathbf{P} \\ \mathbf{P}^T & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{v}_1 \\ \hat{\mathbf{v}}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{c} \\ \mathbf{0} \end{bmatrix}$. Lemma 4.2 of M.
2. Show $\mathbb{E}\mathbf{c}^T\hat{\mathbf{b}}_H = \mathbf{c}^T\mathbf{b}$ for all $\mathbf{b}$. Lemma 4.3 of M.
3. Show that any unbiased estimator $a_0 + \mathbf{a}^T\mathbf{y}$ has variance $\geq \operatorname{Var}\mathbf{c}^T\hat{\mathbf{b}}_H$.

Monahan, J. F. (2008). *A primer on linear models*. CRC Press.