STAT 423

WIN 2016

GLS

$$y = X\beta + \epsilon \quad , \quad Cov(\epsilon) = \Sigma$$

$$\Sigma^{-1/2} y = \Sigma^{-1/2} X\beta + \Sigma^{-1/2}\epsilon$$

$$\tilde{y} = \tilde{X}\beta + \tilde{\epsilon} \quad , \quad Cov(\tilde{\epsilon}) = I \qquad \left(\begin{array}{l} \tilde{X} = \Sigma^{-1/2} X \\ \tilde{y} = \Sigma^{-1/2} y \end{array} \right.$$

β_{GLS}:

$$\hat{\beta}_{GLS} = (\tilde{X}^T \tilde{X})^{-1} (\tilde{X}^T \tilde{y})$$
$$= \underline{\underline{(X^T \Sigma^{-1} X)^{-1} (X^T \Sigma^{-1} y)}}$$

What is $\hat{\beta}_{GLS}$ minimizing?

$\hat{\beta}_{GLS}$ is OLS for $\tilde{y}, \tilde{X}$:

$$\hat{\beta}_{GLS} = \arg\min_{\beta} \| \tilde{y} - \tilde{X}\beta \|^2$$
$$= \arg\min_{\beta} (\tilde{y} - \tilde{X}\beta)^T (\tilde{y} - \tilde{X}\beta)$$
$$= \arg\min \underline{\underline{(y - X\beta)^T \Sigma^{-1} (y - X\beta)}}$$

→ $\hat{\beta}_{GLS}$ minimizes a "weighted" version of the residuals.

WLS: In WLS, $\Sigma = \sigma^2 W^{-1}$, $W^{-1} = \begin{pmatrix} \frac{1}{w_1} & & 0 \\ & \ddots & \\ 0 & & 1/w_n \end{pmatrix}$

$$\hat{\beta}_{WLS} = \hat{\beta}_{GLS} = \arg\min (y - X\beta)^T W (y - X\beta)$$
$$= \arg\min \sum_i (y_i - x_i^T\beta)^2 w_i$$

So $\hat{\beta}_{WLS}$ tries to fit more to obs with $\begin{cases} \text{big weights } w_i \\ \text{small variances } \sigma^2/w_i \end{cases}$

The smaller the variance, the more reliable the measurement.

Is $\hat{\beta}_{GLS}$ unbiased?

$$E[\hat{\beta}_{GLS} | X] = E[(X^T\Sigma^{-1}X)^{-1} X^T\Sigma^{-1} y]$$
$$= (X^T\Sigma^{-1}X)^{-1}X^T\Sigma^{-1} E[y]$$
$$= (X^T\Sigma^{-1}X)^{-1} X^T\Sigma^{-1} X\beta$$
$$= \beta \;!$$

What is the variance?

$$V(\hat{\beta}_{GLS}) = E[(\hat{\beta}_{GLS} - \beta)(\hat{\beta}_{GLS} - \beta)^T]$$

Now
$$\hat{\beta}_{GLS} = (X^T\Sigma^{-1}X)^{-1} X^T\Sigma^{-1} y$$
$$= (X^T\Sigma^{-1}X)^{-1} X^T\Sigma^{-1}(X\beta + \epsilon)$$
$$= \beta + (X^T\Sigma^{-1}X)^{-1} X^T\Sigma^{-1}\epsilon$$

$$\hat{\beta}_{GLS} - \beta = (X^T\Sigma^{-1}X)^{-1}(X^T\Sigma^{-1})\epsilon$$

$$V(\hat{\beta}_{GLS}) = E[(X^T\Sigma^{-1}X)^{-1} X^T\Sigma^{-1}\epsilon\epsilon^T\Sigma^{-1}X(X^T\Sigma^{-1}X)^{-1}]$$
$$= (X^T\Sigma^{-1}X)^{-1} X^T\Sigma^{-1} E[\epsilon\epsilon^T]\Sigma^{-1}X(X^T\Sigma^{-1}X)^{-1}$$
$$= \quad '' \quad \Sigma \quad ''$$
$$= (X^T\Sigma^{-1}X)^{-1} X^T\Sigma^{-1}X \, (X^T\Sigma^{-1}X)^{-1}$$
$$= \underline{\underline{(X^T\Sigma^{-1}X)^{-1}}}$$

Check: $\Sigma = \sigma^2 I$, $\Sigma^{-1} = \sigma^{-2} I$

$$(X^T\Sigma^{-1}X)^{-1} = (X^TX/\sigma^2)^{-1} = \sigma^2(X^TX)^{-1}$$