Module 11:
GLM Estimation

GLM

A standard GLM can be written:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \qquad \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \mathbf{V})$$

where

$$\underbrace{\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}}_{\text{fMRI Data}} = \underbrace{\begin{bmatrix} 1 & X_{11} & \cdots & X_{1p} \\ 1 & X_{21} & \cdots & X_{2p} \\ \vdots & \vdots & & \vdots \\ 1 & X_{np} & \cdots & X_{np} \end{bmatrix}}_{\text{Design matrix}} \times \underbrace{\begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}}_{\text{Regression coefficients}} + \underbrace{\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}}_{\text{Noise}}$$

V is the covariance matrix whose format depends on the noise model.

Problem Formulation

- Assume the model:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \qquad \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \mathbf{I}\sigma^2)$$

- The matrices **X** and **Y** are assumed to be known, and the noise is considered to be uncorrelated.

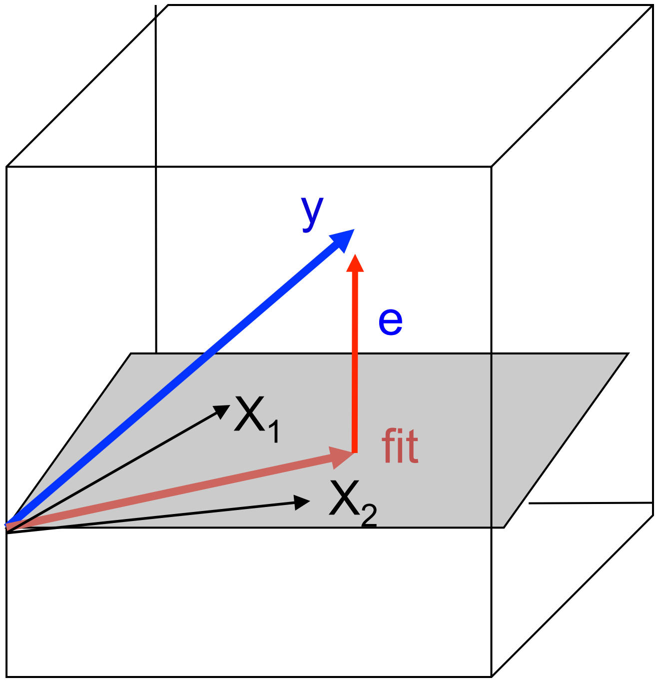
- Our goal is to find the value of $\boldsymbol{\beta}$ that minimizes:

$$\left(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\right)^T \left(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\right)$$

OLS Solution

Ordinary least squares solution

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$$



<u>Properties:</u>

Maximum likelihood estimate

$$E\left(\hat{\beta}\right) = \beta$$

$$Var(\hat{\beta}) = \sigma^2\left(\mathbf{X}^T\mathbf{X}\right)^{-1}$$

Can we do better?

Gauss Markov Theorem

- The Gauss-Markov Theorem states that any other unbiased estimator of β will have a larger variance than the OLS solution.
- Assume $\tilde{\beta}$ is an unbiased estimator of β.
- Then according to G-M Theorem,

$$Var(\tilde{\beta}) \geq Var(\hat{\beta})$$

- $\hat{\beta}$ is the best linear unbiased estimator (BLUE) of **β.**

Estimation

- If ε is i.i.d., then Ordinary Least Square (OLS) estimate is optimal

model | estimate

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \longrightarrow \hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$$

- If Var(ε) $= \mathbf{V}\sigma^2 \neq \mathbf{I}\sigma^2$, then Generalized Least Squares (GLS) estimate is optimal

model | estimate

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \longrightarrow \hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{Y}$$

GLM Summary

model

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

estimate

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{Y}$$

fitted values

$$\hat{\mathbf{Y}} = \mathbf{X}\hat{\beta}$$

residuals

$$\begin{aligned}\mathbf{r} &= \mathbf{Y} - \hat{\mathbf{Y}} \\ &= (\mathbf{I} - (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1})\mathbf{Y} \\ &= \mathbf{R}\mathbf{Y}\end{aligned}$$

Estimating the Variance

- Even if we assume $\boldsymbol{\varepsilon}$ is i.i.d., we still need to estimate the residual variance, σ^2.

- Our estimate:

$$\hat{\sigma}^2 = \frac{\mathbf{r}^T \mathbf{r}}{tr(\mathbf{RV})}$$

- For OLS:

$$\hat{\sigma}^2 = \frac{\mathbf{r}^T \mathbf{r}}{N - p}$$

- Estimating $\mathbf{V} \neq \mathbf{I}$ more difficult.

End of Module