STT 465

I. Multiple Linear Regression (MLE/OLS)
II. Multivariate Normal Distribution
II. Bayesian Multiple Linear Regression

Multiple Linear Regression

- Linear Regression Model

$$y_i = \mu + x_{i1}\beta_1 + x_{i1}\beta_2 + \ldots + x_{ip}\beta_p + \varepsilon_i$$

$$= \mu + \sum_{j=1}^{p} x_{ij}\beta_j + \varepsilon_i$$

- Matrix representation

Let $x_i' = (1, x_{i1}, x_{i2}, \ldots, x_{ip}) \quad \beta = (\mu, \beta_1, \beta_2, \ldots, \beta_p)'$

Then $y_i = x_i'\beta + \varepsilon_i$

Stack equations 1 to n to get $y = X\beta + \varepsilon$

Where $y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$ $X = \begin{bmatrix} x_1' \\ \vdots \\ x_n' \end{bmatrix}$ or $X = [x_1, \ldots, x_p]$ and $\varepsilon = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}$

Multiple Linear Regression

- Residual sum of squares

$$RSS = \sum_{i=1}^{n}\left(y_i - \mu - \sum_{j=1}^{p} x_{ij}\beta_j\right)^2 = \left(y - X\beta\right)'\left(y - X\beta\right)$$

- Ordinary-Least Squares (OLS)

- Take derivative of the RSS with respect to one coefficient
- Set the resulting equation equal to zero (FOC)
- Do the same for all coefficients
- This yields as many equations as unknowns, solve for the coefficients.
- We are going to stack all these FOC to get a closed-form matrix representation of the OLS solution.
- The solution will take the following form

$$\left[X'X\right]\hat{\beta} = X'y$$

or, for full-rank systems

$$\hat{\beta} = \left[X'X\right]^{-1} X'y$$

Steps for deriving OLS estimates

$$\frac{dRSS}{d\beta_j} = -2\sum_{i=1}^{n}\left(y_i - \sum_{k=1}^{p} x_{ik}\beta_k\right)x_{ij}$$

$$= -2\sum_{i=1}^{n}\left(x_{ij}y_i - \sum_{k=1}^{p} x_{ij}x_{ik}\beta_k\right)$$

$$= -2\left[\sum_{i=1}^{n} x_{ij}y_i - \sum_{i=1}^{n}\sum_{k=1}^{p} x_{ij}x_{ik}\beta_k\right]$$

$$= -2\left[\sum_{i=1}^{n} x_{ij}y_i - \sum_{k=1}^{p}\sum_{i=1}^{n} x_{ij}x_{ik}\beta_k\right]$$

$$= -2\left[x_j'y - \sum_{k=1}^{p} x_j'x_k\beta_k\right]$$

$$FOC_j : -2\left[x_j'y - \sum_{k=1}^{p} x_j'x_k\hat{\beta}_k\right] = 0 \Leftrightarrow \sum_{k=1}^{p} x_j'x_k\hat{\beta}_k = x_j'y$$

Steps for deriving OLS estimates

$$\frac{dRSS}{d\beta_j} = -2\sum_{i=1}^{n}\left(y_i - \sum_{k=1}^{p} x_{ik}\beta_k\right)x_{ij}$$

$$= -2\sum_{i=1}^{n}\left(x_{ij}y_i - \sum_{k=1}^{p} x_{ij}x_{ik}\beta_k\right)$$

$$= -2\left[\sum_{i=1}^{n} x_{ij}y_i - \sum_{i=1}^{n}\sum_{k=1}^{p} x_{ij}x_{ik}\beta_k\right]$$

$$= -2\left[\sum_{i=1}^{n} x_{ij}y_i - \sum_{k=1}^{p}\sum_{i=1}^{n} x_{ij}x_{ik}\beta_k\right]$$

$$= -2\left[x_j'y - \sum_{k=1}^{p} x_j'x_k\beta_k\right]$$

$$FOC_j : -2\left[x_j'y - \sum_{k=1}^{p} x_j'x_k\hat{\beta}_k\right] = 0 \Leftrightarrow \sum_{k=1}^{p} x_j'x_k\hat{\beta}_k = x_j'y$$

Stack all the FOCs in a system of linear equations

$$FOC_j : \sum_{k=1}^{p} x_j' x_k \hat{\beta}_k = x_j' y$$

$$\begin{bmatrix} x_1' x_1 & \cdots & x_1' x_p \\ \vdots & \ddots & \vdots \\ x_p' x_1 & \cdots & x_p' x_p \end{bmatrix} \begin{pmatrix} \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_p \end{pmatrix} = \begin{bmatrix} x_1' y \\ \vdots \\ x_p' y \end{bmatrix}$$

$$[X'X]\hat{\beta} = X'y$$

Maximum Likelihood Estimation Under Normal Assumptions

Multiple linear regression with normal error terms

$$y_i = \sum_{j=1}^{p} x_{ij}\beta_j + \varepsilon_i \qquad [\mathrm{x}_{1i} = 1; \beta_1 = \mu] \quad \varepsilon_i \overset{iid}{\sim} N\left(0, \sigma_\varepsilon^2\right)$$

Likelihood Function

$$p\left(y \mid \beta, \sigma_\varepsilon^2\right) = \prod_{i=1}^{n} \frac{Exp\left\{\dfrac{-\left(y_i - \sum_{j=1}^{p} x_{ij}\beta_j\right)^2}{2\sigma_\varepsilon^2}\right\}}{\sqrt{2\pi\sigma_\varepsilon^2}}$$

$$= \left(\frac{1}{2\pi\sigma_\varepsilon^2}\right)^{-n/2} Exp\left\{\frac{-1}{2\sigma_\varepsilon^2}\sum_{i=1}^{n}\left(y_i - \sum_{j=1}^{p} x_{ij}\beta_j\right)^2\right\}$$

Maximum Likelihood Estimation Under Normal Assumptions

Likelihood Function

$$L\left(\beta,\sigma_\varepsilon^2 \mid y\right) = \left(\frac{1}{2\pi\sigma_\varepsilon^2}\right)^{-n/2} Exp\left\{\frac{-RSS\left(y,\beta\right)}{2\sigma_\varepsilon^2}\right\}$$

Log-Likelihood Function

$$l\left(\beta,\sigma_\varepsilon^2 \mid y\right) = -\frac{n}{2}\log\left(2\pi\sigma_\varepsilon^2\right) - \frac{1}{2\sigma_\varepsilon^2} RSS\left(y,\beta\right)$$

MLE of Reg. Coefficients

$$l\left(\beta \mid \sigma_\varepsilon^2, y\right) \propto -\frac{1}{2\sigma_\varepsilon^2} RSS\left(y,\beta\right) \Rightarrow MLE = OLS$$

Sampling Distribution of OLS (& ML) Estimates

Sampling Distribution of OLS estimates

OLS estimator

$$\hat{\beta} = \left[X'X\right]^{-1} X'y$$

Expected value

$$E\left[\hat{\beta}\right] = E\left[\left(X'X\right)^{-1} X'y\right]$$

$$= \left(X'X\right)^{-1} X'E\left[X\beta + \varepsilon\right] \quad \text{[assuming } X'E(\varepsilon)=0\text{]}$$

$$= \left(X'X\right)^{-1} X'X\beta$$

$$= \beta \quad \text{[OLS estimates are unbiased].}$$

Sampling Distribution of OLS estimates

OLS estimator

$$\hat{\beta} = \left[X'X\right]^{-1} X'y$$

Variance

$$Cov\left[\hat{\beta}, \hat{\beta}'\right] = \sigma_{\varepsilon}^{2} \left(X'X\right)^{-1} \quad \text{[derivation presented in class]}$$

Asymptotic distribution

$$\hat{\beta} \sim MVN\left[\beta, \left(X'X\right)^{-1} \sigma_{\varepsilon}^{2}\right]$$

Applications

- **Intercept model.**
- **Two (or more) means model.**
- **Linear regression.**
- **Linear regression with two or more groups.**

Applications

- **Intercept model.**

- **Two (or more) means model.**

- **Linear regression.**

- **Linear regression with two or more groups**

Multivariate Normal Distribution

$$x \sim MVN\left[\mu, \Sigma\right]$$

$$x = \left(x_1, \ldots, x_q\right)' \qquad \mu = \left(\mu_1, \ldots, \mu_q\right)'$$

$$\Sigma = Cov\left(x, x'\right) = \begin{bmatrix} Cov\left(x_1, x_1\right) & \cdots & Cov\left(x_1, x_q\right) \\ \vdots & \ddots & \vdots \\ Cov\left(x_q, x_1\right) & \cdots & Cov\left(x_q, x_q\right) \end{bmatrix}$$

Multivariate Normal Distribution

$$p\left(x\right)\sim\left(2\pi\right)^{-n/2}\left\|\Sigma\right\|^{-1/2}Exp\left\{-\frac{1}{2}\left(x-\mu\right)'\Sigma^{-1}\left(x-\mu\right)\right\}$$

Multivariate Normal Distribution

$$x = \left(x_1', x_2'\right)' \quad \mu = \left(\mu_1', \mu_2'\right)' \quad \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

Important Results

- All marginal are normal
- All conditional distributions are also normal
- The normal distribution is closed under linear transformations (i.e., linear transformations of MVN random variables are also MVN).

Multivariate Normal Distribution

$$x \sim MVN\left[\mu, \Sigma\right] \quad x = \left(x_1', x_2'\right)' \quad \mu = \left(\mu_1', \mu_2'\right)' \quad \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

Marginal distributions are normal: $x_j \sim MVN\left[\mu_j, \Sigma_{jj}\right]$

Conditional distributions are normal: $x_2 \mid x_1 \sim MVN\left[\mu_{2|1}, \Sigma_{2|1}\right]$

where: $\mu_{2|1} = E[x_2 \mid x_1] = \mu_2 + \Sigma_{21}\Sigma_{11}^{-1}\left(x_1 - \mu_1\right)$

and $\Sigma_{2|1} = Cov\left(x_2 \mid x_1\right) = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$