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

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

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

- 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

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

or, for full-rank systems

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

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

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