

Data modeling: CSCI E-106

Applied Linear Statistical Models

Chapter 5 - Matrix Approach to Simple Linear Regression Analysis

Matrices

- Definition of Matrix
- Square Matrix
- Vector
- Transpose
- Equality of Matrices
- Matrix operation: $+$, $-$, \times , \div etc.

Matrices, cont'd

- A column with all elements 1 and Zero Vector: **1** and **0**

$$\mathbf{1}_{r \times 1} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \quad \mathbf{0}_{r \times 1} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

- A square matrix with all elements 1: **J**

$$\mathbf{J}_{r \times r} = \begin{bmatrix} 1 & \dots & 1 \\ \vdots & & \vdots \\ 1 & \dots & 1 \end{bmatrix} = \mathbf{1}\mathbf{1}'$$

- **$\mathbf{1}'\mathbf{1} = n$**

Matrices, cont'd

- Linear Dependence
- Rank of Matrix
- Inverse of a Matrix: $A^{-1}A = AA^{-1} = I$ (identity matrix)

Matrices Form for regression analysis

- Y : consisting of the n observations on the response variable

$$\underset{n \times 1}{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}$$

- X matrix:

$$\underset{n \times 2}{X} = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_n \end{bmatrix}$$

often referred to as the **design matrix**

Matrices Form for regression analysis, cont'd

- The regression model:

$$Y_i = E\{Y_i\} + \varepsilon_i, \quad i = 1, \dots, n$$

$$\underset{n \times 1}{E\{\mathbf{Y}\}} = \begin{bmatrix} E\{Y_1\} \\ E\{Y_2\} \\ \vdots \\ E\{Y_n\} \end{bmatrix}, \quad \underset{n \times 1}{\boldsymbol{\varepsilon}} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

- The regression model:

$$\underset{n \times 1}{\mathbf{Y}} = \underset{n \times 1}{E\{\mathbf{Y}\}} + \underset{n \times 1}{\boldsymbol{\varepsilon}}, \quad E\{\boldsymbol{\varepsilon}\} = \mathbf{0}$$

Matrices Form for regression analysis, cont'd

- Product:

$$\mathbf{Y}'\mathbf{Y} = [Y_1 \quad Y_2 \quad \cdots \quad Y_n] \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = [Y_1^2 + Y_2^2 + \cdots + Y_n^2] = [\sum Y_i^2]$$

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ X_1 & X_2 & \cdots & X_n \end{bmatrix} \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_n \end{bmatrix} = \begin{bmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{bmatrix}$$

$$\Rightarrow (\mathbf{X}'\mathbf{X})^{-1} = \begin{bmatrix} \frac{1}{n} + \frac{\bar{X}^2}{\sum (X_i - \bar{X})^2} & -\frac{\bar{X}}{\sum (X_i - \bar{X})^2} \\ -\frac{\bar{X}}{\sum (X_i - \bar{X})^2} & \frac{1}{\sum (X_i - \bar{X})^2} \end{bmatrix}$$

$$\mathbf{X}'\mathbf{Y} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ X_1 & X_2 & \cdots & X_n \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} \sum Y_i \\ \sum X_i Y_i \end{bmatrix}$$

Matrices Form for regression analysis, cont'd

- Variance-Covariance Matrix of \mathbf{Y} :

$$\sigma^2\{\mathbf{Y}\} = \begin{bmatrix} \sigma^2\{Y_1\} & \sigma\{Y_1, Y_2\} & \sigma\{Y_1, Y_3\} \\ \sigma\{Y_2, Y_1\} & \sigma^2\{Y_2\} & \sigma\{Y_2, Y_3\} \\ \sigma\{Y_3, Y_1\} & \sigma\{Y_3, Y_2\} & \sigma^2\{Y_3\} \end{bmatrix}$$

$$= E\{[\mathbf{Y} - E\{\mathbf{Y}\}][\mathbf{Y} - E\{\mathbf{Y}\}]'\}$$

$$\sigma^2\{\mathbf{Y}\} = E \left\{ \begin{bmatrix} Y_1 - E\{Y_1\} \\ Y_2 - E\{Y_2\} \\ Y_3 - E\{Y_3\} \end{bmatrix} \begin{bmatrix} Y_1 - E\{Y_1\} & Y_2 - E\{Y_2\} & Y_3 - E\{Y_3\} \end{bmatrix} \right\}$$

Multiplying the two matrices and then taking expectations, we obtain:

Location in Product	Term	Expected Value
Row 1, column 1	$(Y_1 - E\{Y_1\})^2$	$\sigma^2\{Y_1\}$
Row 1, column 2	$(Y_1 - E\{Y_1\})(Y_2 - E\{Y_2\})$	$\sigma\{Y_1, Y_2\}$
Row 1, column 3	$(Y_1 - E\{Y_1\})(Y_3 - E\{Y_3\})$	$\sigma\{Y_1, Y_3\}$
Row 2, column 1	$(Y_2 - E\{Y_2\})(Y_1 - E\{Y_1\})$	$\sigma\{Y_2, Y_1\}$
etc.	etc.	etc.

Matrices Form for regression analysis, cont'd

To **generalize**, the variance-covariance matrix for an $n \times 1$ random vector \mathbf{Y} is:

$$\sigma^2\{\mathbf{Y}\}_{n \times n} = \begin{bmatrix} \sigma^2\{Y_1\} & \sigma\{Y_1, Y_2\} & \cdots & \sigma\{Y_1, Y_n\} \\ \sigma\{Y_2, Y_1\} & \sigma^2\{Y_2\} & \cdots & \sigma\{Y_2, Y_n\} \\ \vdots & \vdots & & \vdots \\ \sigma\{Y_n, Y_1\} & \sigma\{Y_n, Y_2\} & \cdots & \sigma^2\{Y_n\} \end{bmatrix}$$

Note again that $\sigma^2\{\mathbf{Y}\}$ is a symmetric matrix.

- variance-covariance for ε , $\sigma(\varepsilon) = \sigma^2 \mathbf{I}$ 

Matrices Form for regression analysis, cont'd

- Reminder for Matrix Algebra:

\mathbf{W} , \mathbf{Y} : random vectors

\mathbf{A} : a constant matrix

$$E\{\mathbf{A}\} = \mathbf{A}$$

$$E\{\mathbf{W}\} = E\{\mathbf{A}\mathbf{Y}\} = \mathbf{A}E\{\mathbf{Y}\}$$

$$\sigma^2\{\mathbf{W}\} = \sigma^2\{\mathbf{A}\mathbf{Y}\} = \mathbf{A}\sigma^2\{\mathbf{Y}\}\mathbf{A}'$$

Multivariate Normal Distribution

$$\mathbf{Y}_{p \times 1} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_p \end{bmatrix} \quad \boldsymbol{\mu}_{p \times 1} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix} \quad \boldsymbol{\Sigma}_{p \times p} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2p} \\ \vdots & \vdots & & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_p^2 \end{bmatrix}$$

- Density Function:

$$f(\mathbf{Y}) = \frac{1}{(2\pi)^{p/2} |\boldsymbol{\Sigma}|^{1/2}} \exp \left[-\frac{1}{2} (\mathbf{Y} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{Y} - \boldsymbol{\mu}) \right]$$

Simple Linear Regression Model in Matrix Terms

- The normal error regression model (2.1):

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \quad i = 1, \dots, n$$

$$\mathbf{Y}_{n \times 1} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} \quad \mathbf{X}_{n \times 2} = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_n \end{bmatrix} \quad \boldsymbol{\beta}_{2 \times 1} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} \quad \boldsymbol{\varepsilon}_{n \times 1} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_n \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \Rightarrow \mathbf{Y}_{n \times 1} = \mathbf{X}_{n \times 2} \boldsymbol{\beta}_{2 \times 1} + \boldsymbol{\varepsilon}_{n \times 1} \Rightarrow \mathbb{E}\{\mathbf{Y}\}_{n \times 1} = \mathbf{X}_{n \times 2} \boldsymbol{\beta}_{2 \times 1}$$

$$\mathbb{E}\{\boldsymbol{\varepsilon}\} = \mathbf{0} \text{ and } \sigma^2\{\boldsymbol{\varepsilon}\} = \sigma^2 \mathbf{I}$$

Simple Linear Regression Model in Matrix Terms

- The normal error regression model (2.1):

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

- $\boldsymbol{\varepsilon}$: a vector of independent normal r.v.
- $E\{\boldsymbol{\varepsilon}\} = \mathbf{0}$
- $\sigma(\boldsymbol{\varepsilon}) = \sigma^2 \mathbf{I}$

$$\sigma^2_{\mathbf{\varepsilon}} = \begin{bmatrix} \sigma^2 & 0 & 0 & \dots & 0 \\ 0 & \sigma^2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & \sigma^2 \end{bmatrix}$$

Least Squares Estimation of Regression Parameters

- Normal Equations

$$\begin{aligned}nb_0 + b_1 \sum X_i &= \sum Y_i \\b_0 \sum X_i + b_1 \sum X_i^2 &= \sum X_i Y_i \\ \Rightarrow \underset{2 \times 2}{\mathbf{X}' \mathbf{X}} \underset{2 \times 1}{\mathbf{b}} &= \underset{2 \times 1}{\mathbf{X}' \mathbf{Y}}\end{aligned}$$

- The vector of the least squares regression coefficients:

$$\mathbf{b} = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \mathbf{Y}$$

Least Squares Estimation of Regression Parameters, cont'd

- Minimize the quantity:

$$\begin{aligned} Q &= \sum [Y_i - (\beta_0 + \beta_1 X_i)]^2 \\ &= (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) \\ &= \mathbf{Y}'\mathbf{Y} - \boldsymbol{\beta}'\mathbf{X}'\mathbf{Y} - \mathbf{Y}'\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta} \end{aligned}$$

- To find the value of $\boldsymbol{\beta}$ that minimizes Q :

$$\frac{\partial}{\partial \boldsymbol{\beta}}(Q) = \begin{bmatrix} \frac{\partial Q}{\partial \beta_0} \\ \frac{\partial Q}{\partial \beta_1} \end{bmatrix} = -2\mathbf{X}'\mathbf{Y} + 2\mathbf{X}'\mathbf{X}\boldsymbol{\beta}$$

Setting the equal to 0, $\Rightarrow \mathbf{X}'\mathbf{X}\mathbf{b} = \mathbf{X}'\mathbf{Y}$

Fitted Values and Residuals

Fitted values:

$$\hat{\mathbf{Y}}_{n \times 1} = \begin{bmatrix} \hat{Y}_1 \\ \hat{Y}_2 \\ \vdots \\ \hat{Y}_n \end{bmatrix} = \underset{n \times 2 \quad 2 \times 1}{\mathbf{X} \mathbf{b}} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} = \mathbf{H}\mathbf{Y}$$

$$\begin{bmatrix} \hat{Y}_1 \\ \hat{Y}_2 \\ \vdots \\ \hat{Y}_n \end{bmatrix} = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_n \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} b_0 + b_1 X_1 \\ b_0 + b_1 X_2 \\ \vdots \\ b_0 + b_1 X_n \end{bmatrix}$$

Fitted Values and Residuals, cont'd

- Hat Matrix or projection matrix:

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

$$\underset{n \times 1}{\hat{\mathbf{Y}}} = \underset{n \times n}{\mathbf{H}} \underset{n \times 1}{\mathbf{Y}}$$

$$\underset{n \times n}{\mathbf{H}} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$$

- Important in diagnostics for regression analysis
- Symmetric and idempotency:

$$\mathbf{H}\mathbf{H} = \mathbf{H}$$

Residuals

Residuals: $e_i = Y_i - \hat{Y}_i$

$$\underset{n \times 1}{\mathbf{e}} = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} = \underset{n \times 1}{\mathbf{Y}} - \underset{n \times 1}{\hat{\mathbf{Y}}} = \underset{n \times 1}{\mathbf{Y}} - \underset{n \times 1}{\mathbf{X}'\mathbf{b}}$$

Variance-Covariance Matrix of Residuals:

- $\mathbf{e} = \mathbf{Y} - \hat{\mathbf{Y}} = (\mathbf{I} - \mathbf{H})\mathbf{Y}$
- $(\mathbf{I} - \mathbf{H})$: symmetric and idempotent
- $\underset{n \times n}{\sigma^2\{\mathbf{e}\}} = \sigma^2(\mathbf{I} - \mathbf{H}) \Rightarrow$ estimated by $\underset{n \times n}{s^2\{\mathbf{e}\}} = MSE(\mathbf{I} - \mathbf{H})$
$$\underset{n \times n}{\sigma^2\{\mathbf{e}\}} = (\mathbf{I} - \mathbf{H})\underset{n \times n}{\sigma^2\{\mathbf{Y}\}}(\mathbf{I} - \mathbf{H})'$$
$$= \sigma^2(\mathbf{I} - \mathbf{H})(\mathbf{I} - \mathbf{H}) \quad \square$$

Analysis of Variance Results

A square matrix with all elements 1 will be denoted by \mathbf{J}

$$\mathbf{J}_{r \times r} = \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{bmatrix}$$

Sums of Squares: $SSTO = \sum (Y_i - \bar{Y})^2 = \sum Y_i^2 - \frac{(\sum Y_i)^2}{n}$

$$SSTO = \sum Y_i^2 - \frac{(\sum Y_i)^2}{n} = \mathbf{Y}' \mathbf{Y} - \left(\frac{1}{n}\right) \mathbf{Y}' \mathbf{J} \mathbf{Y}$$

$$SSE = \mathbf{e}' \mathbf{e} = (\mathbf{Y} - \mathbf{X} \mathbf{b})' (\mathbf{Y} - \mathbf{X} \mathbf{b}) = \mathbf{Y}' \mathbf{Y} - \mathbf{b}' \mathbf{X}' \mathbf{Y}$$

$$SSR = \mathbf{b}' \mathbf{X}' \mathbf{Y} - \left(\frac{1}{n}\right) \mathbf{Y}' \mathbf{J} \mathbf{Y}$$

Sums of Squares as Quadratic Forms: $\mathbf{Y}' \mathbf{A} \mathbf{Y}$

$$SSTO = \mathbf{Y}' \left[\mathbf{I} - \left(\frac{1}{n}\right) \mathbf{J} \right] \mathbf{Y}$$

$$SSE = \mathbf{Y}' [\mathbf{I} - \mathbf{H}] \mathbf{Y}$$



$$SSR = \mathbf{Y}' \left[\mathbf{H} - \left(\frac{1}{n}\right) \mathbf{J} \right] \mathbf{Y}$$

Inferences in Regression Analysis

The variance-covariance matrix of \mathbf{b} :

$$\sigma^2_{\mathbf{b}} = \begin{bmatrix} \sigma^2\{b_0\} & \sigma\{b_0, b_1\} \\ \sigma\{b_1, b_0\} & \sigma^2\{b_1\} \end{bmatrix}$$

is:

$$\sigma^2_{\mathbf{b}} = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$$

or, from (5.24a):

$$\sigma^2_{\mathbf{b}} = \begin{bmatrix} \frac{\sigma^2}{n} + \frac{\sigma^2 \bar{X}^2}{\sum (X_i - \bar{X})^2} & \frac{-\bar{X} \sigma^2}{\sum (X_i - \bar{X})^2} \\ \frac{-\bar{X} \sigma^2}{\sum (X_i - \bar{X})^2} & \frac{\sigma^2}{\sum (X_i - \bar{X})^2} \end{bmatrix}$$

Inferences in Regression Analysis, cont'd

The estimated variance-covariance matrix of \mathbf{b} :

$$\mathbf{s}^2_{2 \times 2}\{\mathbf{b}\} = MSE(\mathbf{X}'\mathbf{X})^{-1} = \begin{bmatrix} \frac{MSE}{n} + \frac{\bar{X}^2 MSE}{\sum (X_i - \bar{X})^2} & \frac{-\bar{X} MSE}{\sum (X_i - \bar{X})^2} \\ \frac{-\bar{X} MSE}{\sum (X_i - \bar{X})^2} & \frac{MSE}{\sum (X_i - \bar{X})^2} \end{bmatrix}$$

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} = \mathbf{A}\mathbf{Y}$$

$$\Rightarrow \sigma^2\{\mathbf{b}\} = \mathbf{A}\sigma^2\{\mathbf{Y}\}\mathbf{A}' = \sigma^2(\mathbf{X}'\mathbf{X})^{-1} \quad \text{🗨️}$$

Mean Response

The mean response at X_h :

$$\mathbf{X}_h = \begin{bmatrix} 1 \\ X_h \end{bmatrix}; \quad \hat{Y}_h = \mathbf{X}_h' \mathbf{b}$$

$$\sigma^2\{\hat{Y}_h\} = \sigma^2 \mathbf{X}_h' (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}_h = \mathbf{X}_h' \sigma^2\{\mathbf{b}\} \mathbf{X}_h = \sigma^2 \left[\frac{1}{n} + \frac{(X_h - \bar{X})^2}{\sum (X_i - \bar{X})^2} \right]$$

$$s^2\{\hat{Y}_h\} = MSE(\mathbf{X}_h' (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}_h)$$

Prediction of New Observation:

$$s^2\{pred\} = MSE(1 + \mathbf{X}_h' (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}_h)$$