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**

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

- A square matrix with all elements 1: J

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

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

Matrices, cont'd

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

Matrices Form for regression analysis

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

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

- ***X matrix:***

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

often referred to as the **design matrix**

Matrices Form for regression analysis, cont'd

- The regression model:

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

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

- The regression model:

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

Matrices Form for regression analysis, cont'd

- Product:

$$\underset{1\times 1}{\mathbf{Y}'\mathbf{Y}} = [Y_1 \quad Y_2 \quad \cdots \quad Y_n]\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = [Y_1^2 + Y_2^2 + \cdots + Y_n^2] = \left[\sum Y_i^2\right]$$

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

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

$$\underset{2\times 1}{\mathbf{X}'\mathbf{Y}} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ X_1 & X_2 & \cdots & X_n \end{bmatrix}\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} \sum Y_i \\ \sum X_i Y_i \end{bmatrix}$$

Matrices Form for regression analysis, cont'd

- Variance-Covariance Matrix of ***Y***:

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

$$= \mathrm{E}\{[\boldsymbol{Y} - \mathrm{E}\{\boldsymbol{Y}\}][\boldsymbol{Y} - \mathrm{E}\{\boldsymbol{Y}\}]'\}$$

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

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:

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

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

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

Matrices Form for regression analysis, cont'd

- Reminder for Matrix Algebra:

W**, **Y: random vectors

A: a constant matrix

$$\mathrm{E}\{\boldsymbol{A}\} = \boldsymbol{A}$$

$$\mathrm{E}\{\boldsymbol{W}\} = \mathrm{E}\{\boldsymbol{A}\boldsymbol{Y}\} = \boldsymbol{A}\mathrm{E}\{\boldsymbol{Y}\}$$

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

Multivariate Normal Distribution

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

- Density Function:

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

Simple Linear Regression Model in Matrix Terms

- The normal error regression model (2.1):

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \qquad i = 1, \ldots, n$$

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

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

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

Simple Linear Regression Model in Matrix Terms

- The normal error regression model (2.1):

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

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

$$\underset{n\times n}{\boldsymbol{\sigma}^2\{\boldsymbol{\varepsilon}\}} = \begin{bmatrix} \sigma^2 & 0 & 0 & \cdots & 0 \\ 0 & \sigma^2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \sigma^2 \end{bmatrix}$$

Least Squares Estimation of Regression Parameters

- Normal Equations

$$nb_0 + b_1 \sum X_i = \sum Y_i$$

$$b_0 \sum X_i + b_1 \sum X_i^2 = \sum X_i Y_i$$

$$\Rightarrow \underset{2\times 2}{\boldsymbol{X}'\boldsymbol{X}}\ \underset{2\times 1}{\boldsymbol{b}} = \underset{2\times 1}{\boldsymbol{X}'\boldsymbol{Y}}$$

- The vector of the least squares regression coefficients:

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

Least Squares Estimation of Regression Parameters, cont'd

- Minimize the quantity:

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

- To find the value of β that minimizes *Q*:

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

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

Fitted Values and Residuals

Fitted values:

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

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

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

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

Variance-Covariance Matrix of Residuals:

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

$$\begin{aligned} \boldsymbol{\sigma}^2\{\boldsymbol{e}\} &= (\boldsymbol{I} - \boldsymbol{H})\boldsymbol{\sigma}^2\{\boldsymbol{Y}\}(\boldsymbol{I} - \boldsymbol{H})' \\ &= \sigma^2(\boldsymbol{I} - \boldsymbol{H})(\boldsymbol{I} - \boldsymbol{H}) \end{aligned}$$

Analysis of Variance Results

A square matrix with all elements 1 will be denoted by J

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

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

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

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

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

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

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

$$SSE = \boldsymbol{Y}'\left[\boldsymbol{I} - \boldsymbol{H}\right]\boldsymbol{Y}$$

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

Inferences in Regression Analysis

The variance-covariance matrix of **b**:

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

is:

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

or, from (5.24a):

$$\underset{2\times 2}{\boldsymbol{\sigma}^2\{\mathbf{b}\}} = \begin{bmatrix} \dfrac{\sigma^2}{n} + \dfrac{\sigma^2\bar{X}^2}{\sum(X_i - \bar{X})^2} & \dfrac{-\bar{X}\sigma^2}{\sum(X_i - \bar{X})^2} \\ \dfrac{-\bar{X}\sigma^2}{\sum(X_i - \bar{X})^2} & \dfrac{\sigma^2}{\sum(X_i - \bar{X})^2} \end{bmatrix}$$

Inferences in Regression Analysis, cont'd

The estimated variance-covariance matrix of ***b***:

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

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

$$\Rightarrow \boldsymbol{\sigma}^2\{\boldsymbol{b}\} = \boldsymbol{A}\boldsymbol{\sigma}^2\{\boldsymbol{Y}\}\boldsymbol{A}' = \sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}$$

Mean Response

The mean response at X_h:

$$\underset{2\times 1}{\boldsymbol{X}_h} = \begin{bmatrix} 1 \\ X_h \end{bmatrix}; \quad \hat{Y}_h = \boldsymbol{X}'_h \boldsymbol{b}$$

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

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

Prediction of New Observation:

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