Estimation Continued

The Analysis of Variance Approach

- Partitioning of total variability

$$y_i - \bar{y} = (\hat{y}_i - \bar{y}) + (y_i - \hat{y}_i)$$

Add and subtract yihat

$$\sum (y_i - \bar{y})^2 = \sum (\hat{y}_i - \bar{y})^2 + \sum (y_i - \hat{y}_i)^2 + 2\sum (\hat{y}_i - \bar{y})(y_i - \hat{y}_i)$$

This term is equal to 0

or

$$\underbrace{\sum (y_i - \bar{y})^2}_{SS_T} = \underbrace{\sum (\hat{y}_i - \bar{y})^2}_{SS_R} + \underbrace{\sum (y_i - \hat{y}_i)^2}_{SS_{\mathrm{Res}}}$$

- It can be shown that $SS_R = \hat{\beta}_1 S_{xy}$

The Analysis of Variance

- Degrees of Freedom

$$\underbrace{\sum \left(y_i - \bar{y} \right)^2}_{SS_T} = \underbrace{\sum \left(\hat{y}_i - \bar{y} \right)^2}_{SS_R} + \underbrace{\sum \left(y_i - \hat{y}_i \right)^2}_{SS_{\mathrm{Res}}}$$

$$n-1 \qquad = \qquad 1 \qquad + \quad (n-2)$$

- Mean Squares

$$MS_R = \frac{SS_R}{1} \qquad\qquad MS_{\mathrm{Res}} = \frac{SS_{\mathrm{Res}}}{n-2}$$

The Analysis of Variance

Rely on Cochran's theorem for underlying distribution of MSR/MSRES

- ANOVA procedure for testing H_0: $\beta_1 = 0$

Source of Variation	Sum of Squares	DF	MS	F_0
Regression	SS_R	1	MS_R	MS_R/MS_{Res}
Residual	SS_{Res}	n-2	MS_{Res}	
Total	SS_T	n-1		

- A large value of F_0 indicates that regression is significant; specifically, reject if $F_0 > F_{\alpha,1,n-2}$
- Can also use the *P*-value approach

Compute the probability under the null hypothesis that the random variable F is greater than the computed $F=F^*$

When p is low H0 must go.

Coefficient of Determination

- R^2 - coefficient of determination
- R^2 - SS(Regression)/SS(Total)
- Proportion of variation explained by the regressor, x

Considerations in the Use of Regression

- Extrapolating
- Extreme points will often influence the slope.
- Outliers can disturb the least-squares fit
- Linear relationship does not imply cause-effect relationship

Multiple Regression Models

General Form of the Multiple Regression Model

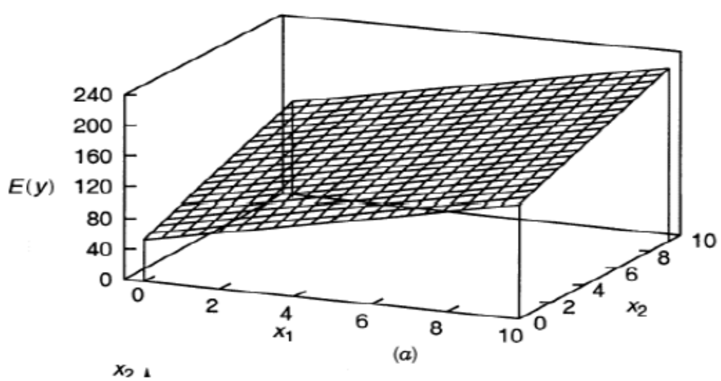
$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon$$

where y is the dependent variable
$x_1, x_2, \ldots, x_k$ are the independent variables
$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k$ is the deterministic portion of the model
β_i determines the contribution of the independent variable x_i

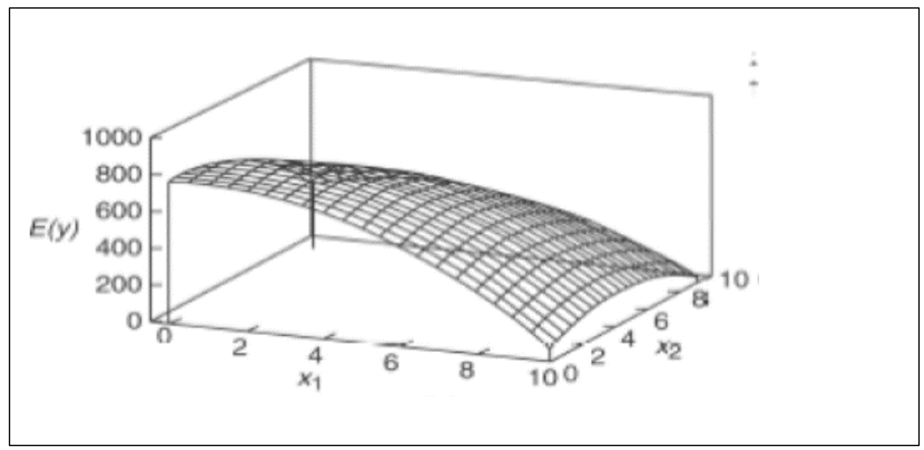
Note: The symbols $x_1, x_2, \ldots, x_k$ may represent higher-order terms for quantitative predictors (e.g., $x_2 = x_1^2$) or terms for qualitative predictors.

$$E(\varepsilon) = 0, \ \mathrm{Var}(\varepsilon) = \sigma^2$$

$E(Y) = 50 + 10x_1 + 7x_2$



$E(Y) = 50 + 10x_1 + 7x_2 + 5\, x_1 x_2$



Example of interaction between two variables.
Time to run a specified obstacle course (min)

Dose of drug (mg)		Low	High
Age	< 65 yo	1.9	1.1
	>=65 yo	2.1	2.0

An Age by Drug Interaction can be observed in the table to the left

Data for Multiple Linear Regression

Observation	y Value	x_1	x_2	$\cdots$	x_k
1	y_1	x_{11}	x_{21}		x_{k1}
2	y_2	x_{12}	x_{22}		x_{k2}
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$
n	y_n	x_{1n}	x_{2n}		x_{kn}

The least squares function is

$$S(\beta_0, \beta_1, \ldots, \beta_k) = \sum_{i=1}^{n} \varepsilon_i^2$$

$$= \sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{j=1}^{k} \beta_j x_{ij} \right)^2$$

The function S must be minimized with respect to the coefficients.

σ^2 is estimated by SSE/n-(k+1)

Matrix notation is typically used.

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

B.2 Matrices and Matrix Multiplication

Although it is very difficult to give the formulas for the multiple regression least squares estimators and for their estimated standard errors in ordinary algebra, it is easy to do so using **matrix algebra**. Thus, by arranging the data in particular rectangular patterns called **matrices** and by performing various operations with them, we can obtain the least squares estimates and their estimated standard errors. In this section and Sections B.3 and B.4, we define what we mean by a matrix and explain various operations that can be performed with matrices. We explain how to use this information to conduct a regression analysis in Section B.5.

Three matrices, **A**, **B**, and **C**, are shown here. Note that each matrix is a rectangular arrangement of numbers with one number in every row–column position.

$$\mathbf{A} = \begin{bmatrix} 2 & 3 \\ 0 & 1 \\ -1 & 6 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} 3 & 0 & 1 \\ -1 & 0 & 1 \\ 4 & 2 & 0 \end{bmatrix} \qquad \mathbf{C} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

Definition B.1 A **matrix** is a rectangular array of numbers.*

The numbers that appear in a matrix are called **elements** of the matrix. If a matrix contains r rows and c columns, there will be an element in each of the row–column positions of the matrix, and the matrix will have $r \times c$ elements. For example, the matrix **A** shown previously contains $r = 3$ rows, $c = 2$ columns, and $rc = (3)(2) = 6$ elements, one in each of the six row–column positions.

Definition B.2 A number in a particular row–column position is called an **element** of the matrix.

Definition B.3 A matrix containing r rows and c columns is said to be an $\boldsymbol{r \times c}$ **matrix** where r and c are the **dimensions** of the matrix.

Definition B.4 If $r = c$, a matrix is said to be a **square matrix**.

Requirement for Matrix Multiplication

$$\mathbf{AB}$$

$$r \times d \qquad d \times c$$

Find the product $\mathbf{AB}$, where

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 4 & -1 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} 2 & 0 & 3 \\ -1 & 4 & 0 \end{bmatrix}$$

Solution

If we represent the product $\mathbf{AB}$ as

$$\mathbf{C} = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \end{bmatrix}$$

The Data Matrices Y and X and the $\hat{\beta}$ Matrix

$$\mathbf{Y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix} \quad \mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{21} & \cdots & x_{k1} \\ 1 & x_{12} & x_{22} & \cdots & x_{k2} \\ 1 & x_{13} & x_{23} & \cdots & x_{k3} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{1n} & x_{2n} & \cdots & x_{kn} \end{bmatrix} \quad \hat{\beta} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_k \end{bmatrix}$$

Special types of matrices:

- Square matrix
- Identity matrix
- Diagonal matrix