Linear Models 1

Isfahan University of Technology
Fall Semester, 2014

References:

[1] G. A. F., Seber and A. J. Lee (2003). Linear Regression Analysis (2nd ed.). Hoboken, NJ: Wiley.

[2] A. C. Rencher and G. B. Schaalje (2008). Linear Models in Statistics (2nd ed.). John Wiley & Sons, Inc.

[3] R. B., Bapat (2000), Linear Algebra and Linear Models, (2^{nd} ed.). Springer-Verlag.

[4] S. R., Searle (1971). Linear Models. New York: Wiley.

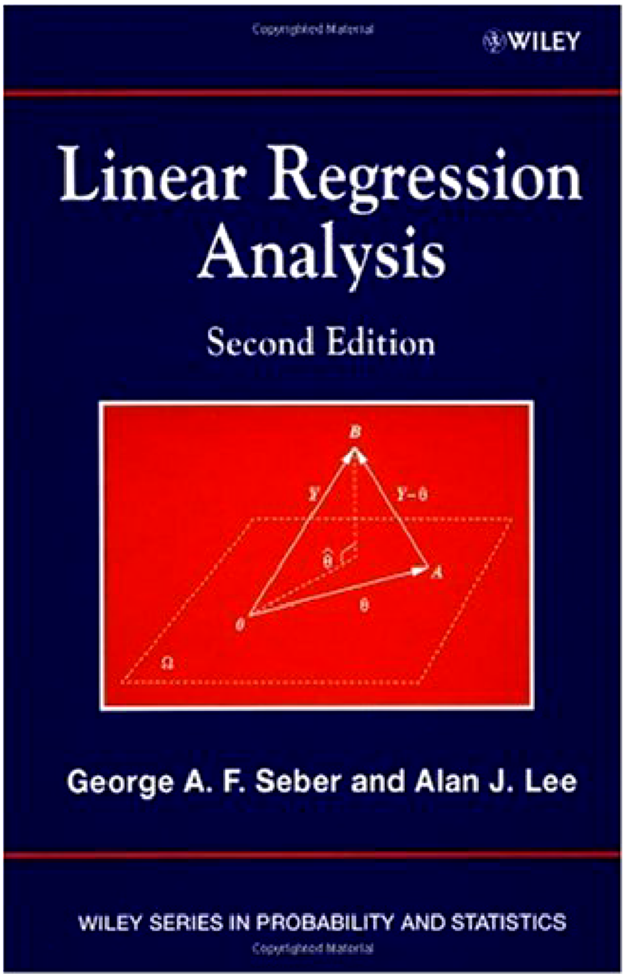
The PDF versions of references are available in the link http://rikhtehgaran.iut.ac.ir/

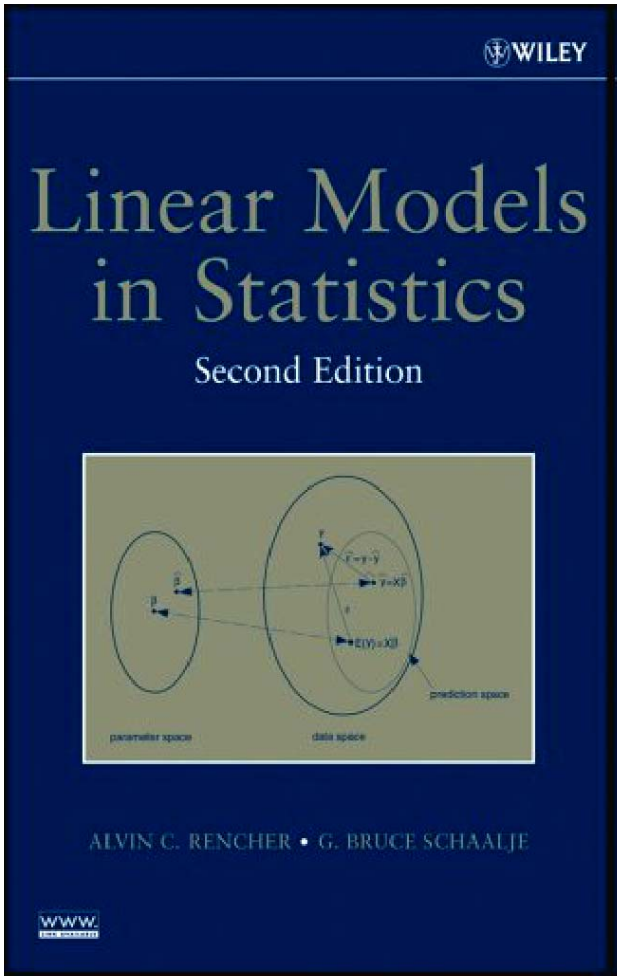
Prerequisites:

- . Linear Algebra
- Basic Statistical Inference.

Grading Policy:

- **Assignments & Quizes:** *3-5 points.*
- **Midterm Exam:** *5-7 points .*
- **Final Exam:** *10 points.*


Copyrighted Material
WILEY
Linear Regression Analysis
Second Edition
George A. F. Seber and Alan J. Lee
WILEY SERIES IN PROBABILITY AND STATISTICS
Copyrighted Material


WILEY
Linear Models in Statistics
Second Edition
prediction space
parameter space
data space
ALVIN C. RENCHER • G. BRUCE SCHAALJE

Contents

Ch. 2. Matrix Algebra. **(Rencher and Schaalje)**

y'Ay

Why statistical models?

It is in human nature to try and understand the physical and natural phenomena that occur around us. When observations on a phenomenon can be quantified, such an attempt at understanding often takes the form of building a **mathematical model**, even if it is only a simplistic attempt to capture the essentials. Either because of our ignorance or in order to keep it simple, many relevant factors may be left out. Also models need to be validated through measurement, and such measurements often come with error. In order to account for the measurement or observational errors as well as the factors that may have been left out, one needs a ***statistical* model** which incorporates some amount of uncertainty.

Why a linear model?

Some of the reasons why we undertake a detailed study of the linear model are as follows.

(a) Because of its simplicity, the linear model is better understood and easier to interpret than most of the other competing models, and the methods of analysis and inference are better developed Therefore, if there is no particular reason to presuppose another model, the linear model may be used at least as a first step.

(b) The linear model formulation is useful even for certain nonlinear models which can be reduced to the linear form by means of a transformation.

(c) Results obtained for the linear model serve as a stepping stone for the analysis of a much wider class of related models such as mixed effects model, state-space and other time series models.

(d) Suppose that the response is modelled as a nonlinear function of the explanatory variables plus error. In many practical situations only a part of the domain of this function is of interest. For example, in a manufacturing process, one is interested in a narrow region centered around the operating point. If the above function is reasonably smooth in this region, a linear model serves as a good first order approximation to what is globally a nonlinear model.

Many statistical concepts can be viewed in the framework of linear models

suppose that we wish to compare the means of two populations, say, $\mu_i = E[U_i]$ $(i = 1, 2)$. Then we can combine the data into the single model

$$\begin{aligned} E(Y) &= \mu_1 + (\mu_2 - \mu_1)x \\ &= \beta_0 + \beta_1 x, \end{aligned}$$

where $x = 0$ when Y is a U_1 observation and $x = 1$ when Y is a U_2 observation. Here $\mu_1 = \beta_0$ and $\mu_2 = \beta_0 + \beta_1$, the difference being β_1. We can extend this idea to the case of comparing m means using m -1 dummy variables.

In a similar fashion we can combine two straight lines,

$$U_j = \alpha_j + \gamma_j x_1 \qquad (j = 1, 2),$$

and 1 otherwise. The combined model is

$$\begin{aligned} E(Y) &= \alpha_1 + \gamma_1 x_1 + (\alpha_2 - \alpha_1)x_2 + (\gamma_2 - \gamma_1)x_1 x_2 \\ &= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3, \end{aligned}$$