Linear Models I

April 22, 2008

1 The basic set-up

For linear models, we begin with a general stucture

$$y = X\beta + \epsilon$$

- y is a matrix whose rows form a series of multivariate measurements, the **response variables**,
- X is a matrix of **explanatory variables**,
- β is a matrix of parameters, and
- ϵ is a matrix containing **residuals** (i.e., errors or noise).

If the residuals have a multivariate normal distribution, then least squares estimation is a maximum likelihood estimation procedure for the β..

Example 1. *For multiple linear regression:*

- Y *is a vector,*
- X *is a matrix of quantitative variables,*
- β *is a vector of parameters, and*
- ϵ *is a vector of independent* $N(0, \sigma^2)$ *random variables.*

Example 2. *For* **(one way) analysis of variance (ANOVA)***:*
The i*-th observation is*

$$y_i = \mu + \beta_j x_{ij} + U_i.$$

- x_{ij} *is 1 if the* i*-th observation belongs to group* j *and 0 otherwise.*
- *The matrix* X *is called a* **design matrix**.
- ϵ_i *are independent* $N(0, \sigma^2)$ *random variables.*

For these models the parameter space is Θ has a vector of parameters β and and perhaps a matrix Σ indicating the covariance structure of the residuals ϵ. In the cases we shall consider here, we will limit ourselves to situations in which the residuals are independent normal random variables, mean 0 and variance σ^2.

Consequently, in matrix form, the likelihood takes the same form as that seen in multiple linear regression. This gives a log-likelihood of

$$\ln L(\beta, \sigma^2|\mathbf{x}, \mathbf{y}) = -\frac{n}{2}(\ln 2\pi + \ln \sigma^2) - \frac{1}{2\sigma^2}(\mathbf{y} - X\beta)^T(\mathbf{y} - X\beta)$$

and estimation of the parameters β is again a least square problem. The maximum likelihood estimators are thus,

$$\hat{\beta} = (X^TX)^{-1}X^T\mathbf{y}.$$

This estimator will have the properties given for the case of multiple linear regression.

The hypothesis test we shall investigate is whether or not some linear combination of the β_i is equal to zero. In other words, for some matrix A,

$$H_0 : A\beta = 0 \quad \text{and} \quad H_1 : A\beta \neq 0.$$

2 Examples

Example 3. *We could consider a model with two explanatory variables*

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \epsilon.$$

To test whether or not the second explanatory variable contributed to the response $\mathbf{y}$*, we have the hypothesis*

$$H_0 : \beta_2 = 0 \quad \text{and} \quad H_1 : \beta_2 \neq 0.$$

In this case, the matrix

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

A special case in one in which we take $x_{i1} = x_i$ *and* $x_{i2} = x_i^2$*. Then,*

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \epsilon$$

and the hypothesis asks whether or not a quadratic relationship between $\mathbf{x}$ *and* $\mathbf{y}$ *is better than a linear relationship.*

Example 4. *A second model with two explanatory variables*

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{i1}^2 + \beta_4 x_{i2}^2 + \beta_5 x_{i1} x_{i2} \epsilon.$$

To test whether or not the two explanatory variables act together to affect to the response $\mathbf{y}$*, we have the hypothesis*

$$H_0 : \beta_5 = 0 \quad \text{and} \quad H_1 : \beta_5 \neq 0.$$

Example 5. *For one way analysis of variance, we could ask whether or not all the groups are the same. The hypothesis in this case is*

$$H_0 : \beta_i = 0 \text{ for all } i \quad \text{and} \quad H_1 : \beta_i \neq 0 \text{ for some } i.$$