Problem statement and estimators

stats545.1 Theory of linear models

Gunnar Stefansson

September 18, 2018

Multiple linear regression problem

For y-observations, we want descriptive and predictive linear model of several variables

$y = \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p$

Formulate with matrices...

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{e}$$

Note that intercept is implicit...

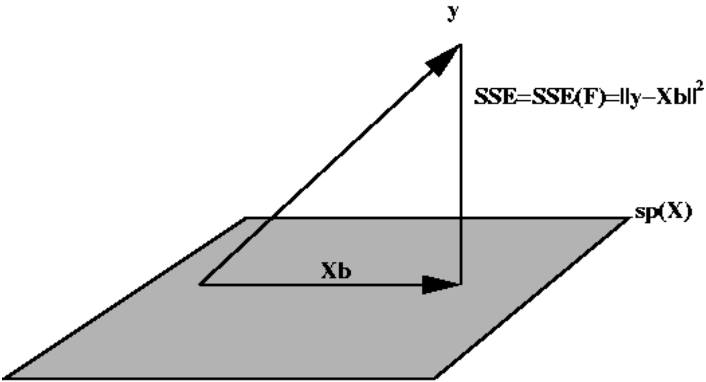
Statistical assumptions will be handled later!

Example: When a straight line is not an appropriate model for explaining the relationship between pairs of measurements, (x_i, y_i), it is possible to consider a quadratic response function, i.e. define the model $EY_i = \alpha + \beta x_i + \gamma x_i^2, \quad i = 1, \ldots, n$.

Example: Consider the data set (from Stefansson, Skuladottir and Petursson) of indices from Icelandic waters. Here T=temperature, U=catch per unit effort of (adult) shrimp, I=index of juvenile shrimp abundance, Y=catch of shrimp, B=biomass of capelin, G=measure of growth of cod from age 4 to 5, S=biomass of spawning cod, J=biomass of juvenile (immature) cod. Defining $x_{i1} = 1,\ x_{i2} = x_i,\ x_{i3} = x_i^2$, this becomes a multiple linear regression model.

This example illustrates clearly how the multiple linear regression model refers to **linearity in the unknown parameters**, not in the independent variables.

Geometric visualization of the multiple regression problem



Want β to give

$$min_{b_1,\ldots,b_p} \sum_{i=1}^{n} \left(y_i - (b_1 x_{i1} + b_2 x_{i2} + \ldots + b_p x_{ip})\right)^2 .$$

i.e. minimize

$$\min ||\mathbf{y} - \mathbf{Xb}||^2$$

Normal equations

Have

$$\mathbf{X}'\mathbf{X}\hat{\beta} = \mathbf{X}'\mathbf{y}$$

The solution

Solution:

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

Prediction:

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

Estimated residuals:

$$\hat{\mathbf{e}} = \mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \mathbf{X}\hat{\beta} = \left(I - \mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}^T\right)\mathbf{y}.$$

... When the matrix is of full rank!

Sums of squares and norms

Sum of squared errors

$$SSE = ||\hat{\mathbf{e}}||^2 = \sum_i (y_i - \hat{y}_i)^2 .$$

Denote SSE by $SSE(F)$ or $SSE(R)$ when comparing models.

Projection matrices

Projecton, "hat", matrix onto $\mathbf{V} = sp(\mathbf{X})$:

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

and onto $\mathbf{V}^{\perp} = sp(\mathbf{X})^{\perp}$:

$$\mathbf{I} - \mathbf{H} = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$$

References Neter, J., Kutner, M. H., Nachtsheim, C. J. and Wasserman, W. 1996. Applied linear statistical models. McGraw-Hill, Boston. 1408pp.