AE 102: Data Analysis and Interpretation

Regression
January - April 2016
Department of Aerospace Engineering

Regression Basics

- Dependent variable or response variables
- Independent variables
- Regression: Finding relationship, based on data
- For example a linear relationship:
- $Y = \beta_0 + \beta_1 x_1 + ... + \beta_r x_r$

Linear Regression

- In actuality to account for error:
- $Y = \beta_0 + \beta_1 x_1 + ... + \beta_r x_r + e$
- Let e be an RV with mean 0, then

$$E[Y|\mathbf{x}] = \beta_0 + \beta_1 x_1 + ... + \beta_r x_r$$

- *Linear regression equation*
- β_i are *regression coefficients*

Linear Regression

- If $r = 1$: simple regression equation
- $r > 1$ multiple regression equation

Least squares estimators for parameters

- Consider: $Y = \alpha + \beta x$
- Let A, B be estimators for the parameters
- Measure two variables, Y_i, x_i
- Minimize sum of squared differences:

$$SS = \sum_{i=1}^{n} (Y_i - A - Bx_i)^2$$

Normal equations

- $\sum Y_i = nA + B\sum x_i$
- $\sum x_i Y_i = A\sum x_i + B\sum x_i^2$
- Solve for A and B.

Solution

- $B = \frac{\sum x_i Y_i - n\bar{x}\bar{Y}}{\sum x_i^2 - n\bar{x}^2}$
- $A = \bar{Y} - B\bar{x}$

Transforming for linearity

- Can transform equation so it is linear
- Example: $W(t) \approx ce^{-at}$
- Example: $1 - P(x) \approx c(1-d)^x$

Polynomial regression

- $Y = \beta_0 + \beta_1 x + \ldots + \beta_r x^r$
- Again minimize error
- Get a set of normal equations

Multiple linear regression

- Need to solve matrices to find the normal equations
- Derive the general equations