1 Math

2 Applied

Topics covered up to 31 October

Matrix / Vector differentiation

2.1 Linear Models

Least Squares Principle

$$\arg\ \min_\beta \|Y - X\beta\|^2$$

Normal Model

$$Y = X\beta + \epsilon, \qquad \epsilon \sim N(0, \sigma^2 I)$$

Normal Equations - Any b satistfying this solves the least squares

$$X^T X b = X^T y$$

Gauss Markov Theorem - $\hat{\beta}(X^T X)^{-1} y$ is Best Linear Unbiased Estimator (BLUE) of β.

Estimating the variance: $\frac{\|y - X\hat{\beta}\|^2}{\sigma^2} \sim \chi^2_{n-p}$.

$$\hat{\sigma}^2 = \frac{\left\|y - X\hat{\beta}\right\|^2}{n-p}$$

Use t test for hypothesis testing and confidence intervals for the value of a particular β_j coefficient.

$$\frac{\beta_j - \beta_j^*}{\hat{\sigma}\sqrt{w_{ii}}} \sim t_{n-p}$$

General linear tests. Partition $\beta = (\beta_1, \beta_2)$ where β_1 is an r vector and β_2 is $p - r$. Null hypothesis $H_0 : \beta_2 = \beta_2^*$ (often 0), and $H_a : \beta_2 \neq \beta_2^*$. Then $SSE_r = \left\|y - X_2\beta_2^* - X_1\tilde{\beta}_1\right\|^2$ is the sum of squared error for the reduced model and $SSE_f = \left\|y - X\hat{\beta}\right\|^2$ is the squared sum of error for the full model. Under H_0:

$$\frac{\frac{SSE_r - SSE_f}{p-r}}{\frac{SSE_f}{n-p}} \sim F_{p-r, n-p}$$

2.2 General Techniques

Singular Value Decompostion (SVD) Any matrix X can be written

$$X = UDV^T$$

with U, V orthogonal, and D diagonal.

Moore Penrose Psuedoinverse A^+ exists uniquely for every matrix A.

Projection matrix P are symmetric and idempotent. They have eigenvalues either 0 or 1.

$$P = P^T \qquad P^2 = P$$

2.3 Distributions

Multivariate Normal $y \sim N(\mu, \Sigma)$, A an $m \times n$ matrix, then

$$Ay \sim N(A\mu, A\Sigma A^t)$$

Noncentral χ^2. Let $Y \sim N(\mu, I)$ be an n vector. Then

$$\|Y\|^2 \sim \chi^2_n(\|(\| \mu)^2)$$