

Lecture 15: quadratic approximation to ℓ

- Two term Taylor expansion of ℓ about $\hat{\theta}$:

$$2 \left[\ell(\hat{\theta}) - \ell(\theta) \right] \approx (\theta - \hat{\theta})^T \mathcal{I}(\theta_0) (\theta - \hat{\theta})$$

where minus the second derivative matrix has been replaced by the Fisher information matrix.

- Worked on large sample theory of above evaluated at $\theta = \theta_0$.
- Began with $X \sim MVN_p(0, \Sigma)$ and showed

$$X^T \Sigma^{-1} X \sim \chi_p^2.$$

- Then started general theory of quadratic form:

$$X^t B X = (AZ)^T B (AZ) = Z^T Q Z$$

where $Q = A^T B A$ and $A A^T = \Sigma$.

- Write $Q = P \Lambda P^T$ where P is orthogonal ($P P^T = P^T P = I$) and Λ is diagonal. Columns of P are eigenvectors of Q and diagonal entries Λ are eigenvalues.



Quadratic Forms in Multivariate Normal

- Suppose $X \sim MVN_m(\mu, \Sigma)$ and Q is $m \times m$ symmetric.
- Write $\Sigma = AA^T$ and assume Σ is not singular so A is not singular.
- If Σ is not singular then

$$X^T Q X \sim \sum_{i=1}^m \lambda_i (Z_i + \delta_i)^2$$

where

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$$

are the eigenvalues of AQA^T .

- If v_1, \dots, v_n are corresponding orthonormal eigenvectors we can take

$$\delta_i = v_i^T A^{-1} \mu$$



Regression example

- Did $Y = X\beta + \epsilon$ and $\epsilon = \sigma Z$ example.
- Here $Z \sim MVN_n(0, I)$.
- We are regressing Y on columns of X .
- Least squares estimates of β (which are the MLE) are

$$\hat{\beta} = (X^T X)^{-1} X^T Y = \beta + \sigma (X^T X)^{-1} X^T Z.$$

- The fitted values are $\hat{Y} = X\hat{\beta}$.
- The residuals are

$$\hat{\epsilon} = Y - \hat{Y} = \sigma (I - X(X^T X)^{-1} X^T) Z.$$



Regression example

- The error sum of squares is

$$\text{ESS} = \hat{\epsilon}^T \hat{\epsilon} = (Y - \hat{Y})^T (Y - \hat{Y}).$$

- So

$$\frac{\text{ESS}}{\sigma^2} = Z^T M Z$$

where

$$M = I - X(X^T X)^{-1} X^T.$$

- M is symmetric and *idempotent*: $MM^T = MM = M$.
- So all eigenvalues of M are 0 or 1.
- Number of 1s is $\text{tr}(M) = n - p$.
- So

$$\frac{\text{ESS}}{\sigma^2} \sim \chi_{n-p}^2$$



Coverage in the text

- See course slides “Likelihood Ratio Tests” 13-17.
- See “course notes” on web pages 130-131.

