Parameter estimates and confidence intervals

Hani Nakhoul

22 March 2010

1 The regression model

$$\mathcal{Y} \sim \mathcal{N}(\mathbf{X}\beta, \sigma^2 \mathbf{I}_n)$$

($\dim \mathcal{Y} = n$, $\dim \beta = p$)

- Least squares estimator: $\hat{\beta} = \arg\min_\beta \|\mathbf{y} - \mathbf{X}\beta\|^2$
- $\hat{\beta}$ minimizes sum of n 'residuals', gives 'residual' or 'restricted maximum likelihood' (REML)
- $\hat{\beta}$ is also the maximum likelihood (ML) estimator for β
- Two estimators for σ^2 : $\hat{\sigma}_L^2 = \frac{\|\mathbf{y}-\mathbf{X}\hat{\beta}\|^2}{n}$ (ML), $\hat{\sigma}_R^2 = \frac{\|\mathbf{y}-\mathbf{X}\hat{\beta}\|^2}{n-p}$ (REML)

2 The mixed effects model

$$\begin{aligned}(\mathcal{Y}|\mathcal{B} = \mathbf{b}) &\sim \mathcal{N}(\mathbf{X}\beta + \mathbf{Z}\mathbf{b}, \sigma^2 \mathbf{I}_n) \\ \mathcal{B} &\sim \mathcal{N}(\mathbf{0}, \Sigma_\theta)\end{aligned}$$

$\mathcal{Y}$ response vector ($\dim n$), $\mathcal{B}$ random-effects vector ($\dim q$), β fixed-effects vector ($\dim p$), $\Sigma_\theta = \sigma_1^2 \Lambda_\theta \Lambda_\theta^{\mathrm{T}}$ for $q \times q$ matrix Λ_θ

2.1 ML estimation

Goal: given observation $\mathbf{y}$, minimize deviance

$$d(\theta, \beta, \sigma|\mathbf{y}) = -2\log L(\theta, \beta, \sigma|\mathbf{y}) = n\log(2\pi\sigma^2) + \log(\|\mathbf{L}_\theta\|^2) + \frac{r_{\beta,\theta}^2}{\sigma^2}$$

1. Obtain conditioned estimators $\hat{\beta}_\theta = \arg\min_\beta r_{\beta,\theta}^2$ and $\hat{\sigma}_\theta^2 = r_\theta^2/n$ (where $r_\theta^2 = \min_\beta r_{\beta,\theta}^2$)
2. Profiled deviance $\tilde{d}(\theta|\mathbf{y}) = d(\theta, \hat{\beta}_\theta, \hat{\sigma}_\theta^2)$ is a function of θ alone; numerical optimization then gives the MLE $\hat{\theta}$ of θ
3. Evaluate $\tilde{d}$ at $\hat{\theta}$ to obtain $\hat{\beta}$ and $\hat{\sigma}$

2.2 REML estimation

Goal: given observation $\mathbf{y}$, minimize REML criterion

$$d_R(\theta, \sigma|\mathbf{y}) = -2\log\int_{\mathbb{R}^p} L(\theta, \beta, \sigma|\mathbf{y})d\beta$$

1. Compute integral:

$$d_R(\theta, \sigma|\mathbf{y}) = (n-p)\log(2\pi\sigma^2) + 2\log(\|\mathbf{L}_\theta\|\|\mathbf{R}_X\|) + \frac{r_\theta^2}{\sigma^2}$$

2. Obtain conditioned estimator: $\hat{\sigma}^2_{R,\theta} = r_\theta^2/(n-p)$

3. Profiled REML criterion $\tilde{d}_R(\theta|\mathbf{y}) = d_R(\theta, \hat{\sigma}^2_{R,\theta})$ is a function of θ alone; optimize to obtain $\hat{\theta}_R = \arg\min_\theta \tilde{d}_R(\theta|\mathbf{y})$ and $\hat{\sigma}^2_R = \hat{\sigma}^2_{R,\hat{\theta}_R}$ (by custom $\hat{\beta}_R = \hat{\beta}_{\theta_R}$)

2.3 ML vs. REML: which to choose?

- REML estimation is less biased than ML estimation, although it is not true in general that the REML variance estimator is unbiased
- The linear mixed-effects REML estimator takes the regression estimator as a special case
- ML estimation produces model-fit statistics (e.g. AIC, BIC) that can be used to compare models fitted to the same data
- ML estimators are required to find confidence intervals

3 Confidence intervals

1. Compute ML estimators of parameters to obtain globally optimal fit
2. Fix one parameter at a specific value, find best possible fit, and compute change in deviance from globally optimal fit
3. The change in deviance is the likelihood ratio test (LRT) statistic ($d = -2\log L$); take signed square root ζ as value of $\mathcal{Z} \sim \mathcal{N}(0,1)$
4. Profiled zeta plot illustrates confidence intervals graphically; profiled pairs plot indicates dependence of parameters on each other

4 Hypothesis testing

1. Given null (H_0) and alternative (H_A) hypotheses about parameters, construct:
   - A reference model incorporating both H_0 and H_A
   - A nested (null hypothesis) model satisfying only H_0
2. Take estimates for parameters in each model
3. Evaluate change in deviance (the LRT statistic): $d_{nested} - d_{ref} = -2\log\frac{L_{nested}}{L_{ref}} \sim \chi^2_{df}$ (df difference in number of parameters)
4. Refer to the χ^2_{df} distribution to determine the significance of the choice of model