STA 610 Homework 8

Yuren Zhou

November 17, 2024

Question 1.

(a) (3 points) The -2 log likelihood of μ is

$$-2\log p(Y|\mu) = \sum_{j=1}^{m}\Big(n_j\log(2\pi) + \log\det(\Sigma_j) + (y_j-\mu 1)^\top\Sigma_j^{-1}(y_j-\mu 1)\Big)$$
$$= \left(\sum_{j=1}^{m} 1^\top\Sigma_j^{-1}1\right)\mu^2 - 2\left(\sum_{j=1}^{m}1^\top\Sigma_j^{-1}y_j\right)\mu + \text{constant},$$

which is a quadratic function that achieves its minimum (i.e. MLE) at

$$\widehat{\mu} = \frac{\sum_{j=1}^m 1^\top\Sigma_j^{-1}y_j}{\sum_{j=1}^m 1^\top\Sigma_j^{-1}1}.$$

(b) (4 points) With $\Sigma_j = \tau^2 11^\top + \sigma^2 I$, by Sherman–Morrison formula we have

$$\Sigma_j^{-1} = \sigma^{-2}I - \frac{\sigma^{-4}\tau^2}{1+\frac{\tau^2}{\sigma^2}n_j}11^\top = \sigma^{-2}I - \frac{\sigma^{-2}\tau^2}{\sigma^2+\tau^2 n_j}11^\top.$$

Plugging into (a), we obtain

$$\widehat{\mu} = \frac{\sum_{j=1}^m\left(\sigma^{-2}n_j\bar{y}_j - \frac{\sigma^{-2}\tau^2}{\sigma^2+\tau^2n_j}n_j^2\bar{y}_j\right)}{\sum_{j=1}^m\left(\sigma^{-2}n_j - \frac{\sigma^{-2}\tau^2}{\sigma^2+\tau^2n_j}n_j^2\right)} = \frac{\sum_{j=1}^m\frac{n_j}{\sigma^2+\tau^2n_j}\bar{y}_j}{\sum_{j=1}^m\frac{n_j}{\sigma^2+\tau^2n_j}},$$

where $\bar{y}_j := \frac{1}{n_j}\sum_{k=1}^{n_j}y_{j,k}$ is the average of entries in the vector y_j. The weight for group ℓ is

$$w_\ell := \frac{\frac{n_\ell}{\sigma^2+\tau^2n_\ell}}{\sum_{j=1}^m\frac{n_j}{\sigma^2+\tau^2n_j}} = \frac{\left(\frac{\sigma^2}{n_\ell}+\tau^2\right)^{-1}}{\sum_{j=1}^m\left(\frac{\sigma^2}{n_j}+\tau^2\right)^{-1}}.$$

- Holding everything else fixed, as n_ℓ increases, the weight of group ℓ gets larger and the weights of all other groups get smaller.
- Notice that

$$\frac{\partial w_\ell}{\partial(\sigma^2)} = \frac{\left(\frac{\sigma^2}{n_\ell}+\tau^2\right)^{-1}\cdot\sum_{j=1}^m\frac{1}{n_j}\left(\frac{\sigma^2}{n_j}+\tau^2\right)^{-2}}{\left(\sum_{j=1}^m\left(\frac{\sigma^2}{n_j}+\tau^2\right)^{-1}\right)^2} - \frac{\frac{1}{n_\ell}\left(\frac{\sigma^2}{n_\ell}+\tau^2\right)^{-2}}{\sum_{j=1}^m\left(\frac{\sigma^2}{n_j}+\tau^2\right)^{-1}},$$

which is positive if and only if

$$\frac{\left(\frac{\sigma^2}{n_\ell}+\tau^2\right)^{-1}}{\sum_{j=1}^m\left(\frac{\sigma^2}{n_j}+\tau^2\right)^{-1}} > \frac{\frac{1}{n_\ell}\left(\frac{\sigma^2}{n_\ell}+\tau^2\right)^{-2}}{\sum_{j=1}^m\frac{1}{n_j}\left(\frac{\sigma^2}{n_j}+\tau^2\right)^{-2}} = \frac{\frac{1}{1+\frac{\tau^2}{\sigma^2}n_\ell}\cdot\left(\frac{\sigma^2}{n_\ell}+\tau^2\right)^{-1}}{\sum_{j=1}^m\frac{1}{1+\frac{\tau^2}{\sigma^2}n_j}\cdot\left(\frac{\sigma^2}{n_j}+\tau^2\right)^{-1}},$$

which holds if n_ℓ is relatively large among all n_j's, and does not hold if n_ℓ is relatively small among all n_j's. Hence, holding everything else fixed, as σ^2 increases, the weights of larger groups get larger and the weights of smaller groups get smaller.

- Similarly to analyzing the effect of σ^2, we can find the effect of τ^2, with the effect being reversed. Holding everything else fixed, as τ^2 increases, the weights of larger groups get smaller and the weights of smaller groups get larger.

Note: Due to the complexity, incorrect answers for the effects of σ^2, τ^2 will not result in point deduction.

(c) (3 points) With equal sample sizes, we have simplification

$$\widehat{\mu} = \frac{1}{m}\sum_{j=1}^m \bar{y}_j.$$

Since the expression doesn't depend on σ^2, τ^2, this is the MLE of μ no matter σ^2, τ^2 is known or unknown.