Zhao_Ping_HW3

Ping Zhao

2/15/2018

(a)

(i)

First, find elements for gradient $\frac{\partial l}{\partial \mu_i}$. Note that from lecture,

$$d(l(d\mu)) = \Sigma_{i=1}^{n}(\vec{x_i} - \vec{\mu})^T \Sigma^{-1} d\vec{\mu} = V d\vec{\mu}$$

Then, denote $\Lambda = \Sigma^{-1}$

$$\frac{\partial l}{\partial \mu} = (\Sigma_{i=1}^{n} x_i - n\mu)^T \Lambda$$

Let's denote Λ_{ij} as the element of $\Lambda = \Sigma^{-1}$ in the ith row and jth column. Then, the vector V is

$$V = (\Sigma_{j=1}^{p} \Lambda_{j1}(\Sigma_{i=1}^{n} x_{ij} - n\mu_j), \Sigma_{j=1}^{p} \Lambda_{j2}(\Sigma_{i=1}^{n} x_{ij} - n\mu_j), ..., \Sigma_{j=1}^{p} \Lambda_{jp}(\Sigma_{i=1}^{n} x_{ij} - n\mu_j))$$

As $d\vec{\mu}$ will only have 1 for the ith element and the rest are 0, then $\frac{\partial l}{\partial \mu_i}$ is the ith element of $\frac{\partial l}{\partial \mu}$ So we have:

$$\frac{\partial l}{\partial \mu_i} = V_i = \Sigma_{j=1}^{p} \Lambda_{ji}(\Sigma_{i=1}^{n} x_{ij} - n\mu_j)$$

Second, for $dl(d\Sigma)$, denote $A = \Lambda[n\Sigma - \Sigma_i^n (x_i - \mu)(x_i - \mu)^T]\Lambda$

$$dl(d\Sigma) = -\frac{1}{2} tr(\Lambda[n\Sigma - \Sigma_i^n (x_i - \mu)(x_i - \mu)^T]\Lambda\Sigma) = -\frac{1}{2} tr(A d\Sigma)$$

Thus, for elements for gradient $\frac{\partial l}{\partial \sigma_{ij}}$ as matrix A as $\Lambda[n\Sigma - \Sigma_i^n (x_i - \mu)(x_i - \mu)^T]\Lambda$ From lecture, we have for diagonal term

$$\frac{\partial l}{\partial \sigma_{ii}} = -\frac{1}{2} A_{ii}$$

And 0 for off-diagonal item

$$\frac{\partial l}{\partial \sigma_{ij}} = -A_{ij}$$

(ii) Hessian

As $\Sigma^{-1} = \Lambda$ is symetric,

$$ddl(d\mu d\mu) = -n d\vec{\mu}^T (\Sigma^{-1})^T d\vec{\mu} = -n d\vec{\mu}^T \Lambda d\vec{\mu}$$

For $\vec{\mu}$

$$\frac{\partial^2 l}{\partial \mu^2} = -n\Lambda$$

So the element is

$$\frac{\partial^2 l}{\partial \mu_i \partial \mu_j} = -n\Lambda_{ij}$$

The the element for information matrix for both i=j and i!=j cases:

$$-E(\frac{\partial^2 l}{\partial \mu_i \partial \mu_j}) = n\Lambda_{ij}$$

For $ddl(d\mu d\Sigma)$ as $E(\Sigma_i^n(\vec{x_i} - \mu)) = \vec{0}$, the elements of the information matrix are all zeros.

$$ddl(d\mu d\Sigma) = (d\mu)^T(-\Lambda d\Sigma\Lambda)\Sigma_i^n(\vec{x_i} - \mu)$$

For $ddl(d\Sigma d\Sigma)$

$$ddl(d\Sigma d\Sigma) = \frac{1}{2}tr[\Lambda d\Sigma(nI - 2\Lambda C)\Lambda d\Sigma]$$

For the elements, let $\lambda_{ij} = [\Lambda]_{ij}$ and $c_{ij} = [C]_{ij}$ Case 1: i=j=k=l, second order differencial of diagonal term:

$$\frac{\partial^2 l}{\partial \sigma_i \partial \sigma_{ii}} = \frac{1}{2}(n\lambda_{ii}^2 - 2\lambda_{ii}^3 c_{ii})$$

Take expectation to get information matrix, it will be $-\frac{n}{2}E(\lambda_{ii}^2)$. Case 2: i=k, j=l, i!=j let $D = \Lambda C$,

$$\frac{\partial^2 l}{\partial \sigma_{ij} \partial \sigma_{ij}} = \frac{1}{2}(n\lambda_{ij}^2 - 2\Sigma_i d_{ij}\Sigma_j \lambda_{ij}\lambda_{ji})$$

Take expectation, $-\frac{n}{2}E(\lambda_{ij}^2)$

Other cases will all be 0.

(2)

```r
library(pracma)
gen <- function(n,p,mu,sig,seed = 22013){
  #---- Generate data from a p-variate normal with mean mu and covariance sigma
  # mu should be a p by 1 vector
  # sigma should be a positive definite p by matrix
  # Seed can be optionally set for the random number generator
  set.seed(seed)
  # generate data from normal mu sigma
  x = matrix(rnorm(n*p),n,p)
  datan = x %*% sqrtm(sig)$B + matrix(mu,n,p, byrow = TRUE)
  datan
}

#set mu and sig

mu<-c(-1,1,2)
sig<-matrix(c(1,0.7,0.7,0.7,1,0.7,0.7,0.7,1), 3,3)
```

```r
df<-gen(200,3,mu,sig)

head(df, 5)
```

```
##            [,1]      [,2]      [,3]
## [1,] -0.1065871 1.4852948 2.6244473
## [2,] -1.0153403 2.4358252 3.1963763
## [3,] -1.1833315 0.6469248 2.3291214
## [4,] -1.7931938 0.2476836 0.6460097
## [5,] -1.6639518 0.4307266 0.3277569
```