Homework 3: Part I

Louis Bensard

February 18, 2018

Problem J-2.2

(a)

- $dl(d\mu) = \sum_{i=1}^{n}(x_i - \mu)^T\Sigma^{-1}d\mu = A_{1\times p}\cdot d\mu$. $[A_{1\times p} := \sum_{i=1}^{n}(x_i - \mu)^T\Sigma^{-1}]$. Thus, $\frac{\partial l}{\partial \mu_i} = A_{1i}$
- $dl(d\Sigma) = -\frac{1}{2}tr[(nI - \Sigma^{-1}C)\Sigma^{-1}d\Sigma] = -\frac{1}{2}tr[B_{p\times p}\cdot d\Sigma]$. $[B_{p\times p} := (nI - \Sigma^{-1}C)\Sigma^{-1}]$. Thus,

$$\begin{cases} \frac{\partial l}{\partial \sigma_{ii}} = -\frac{1}{2}B_{ii} \\ \frac{\partial l}{\partial \sigma_{ij}} = -\frac{1}{2}(B_{ij} + B_{ji}), \quad i \neq j \end{cases}$$

- $ddl(d\mu, d\Sigma) = ddl(d\Sigma, d\mu) = -A\cdot d\Sigma\Sigma^{-1}d\mu$. Thus,

$$\frac{\partial^2 l}{\partial \mu_i \partial \sigma_{kl}} = \frac{\partial^2 l}{\partial \sigma_{kl}\partial \mu_i} = \begin{cases} -(A_{1k}[\Sigma^{-1}]_{il} + A_{1l}[\Sigma^{-1}]_{ik}), \quad k \neq l \\ -A_{1k}[\Sigma^{-1}]_{ik}, \quad k = l \end{cases}$$

- $ddl(d\mu, d\mu) = -nd\mu^T\Sigma^{-1}d\mu$. Thus,

$$\begin{cases} \frac{\partial^2 l}{\partial \mu_i^2} = -n[\Sigma^{-1}]_{ii} \\ \frac{\partial^2 l}{\partial \mu_i \partial \mu_j} = -2n[\Sigma^{-1}]_{ij}, \quad i \neq j \end{cases}$$

- $ddl(d\Sigma, d\Sigma) = -\frac{1}{2}tr[(\Sigma^{-1}C\Sigma^{-1} - B_{p\times p})d\Sigma\Sigma^{-1}d\Sigma] = -\frac{1}{2}tr[D_{p\times p}\cdot d\Sigma\Sigma^{-1}d\Sigma]$.
  $[D_{p\times p} := (\Sigma^{-1}C\Sigma^{-1} - B_{p\times p}]$. Thus,

$$\begin{cases} \frac{\partial^2 l}{\partial \sigma_{ii}\partial \sigma_{kk}} = -\frac{1}{2}D_{ki}[\Sigma^{-1}]_{ik} \\ \frac{\partial^2 l}{\partial \sigma_{ii}\partial \sigma_{kl}} = -\frac{1}{2}(D_{li}[\Sigma^{-1}]_{ik} + D_{ki}[\Sigma^{-1}]_{il}), \quad k \neq l \\ \frac{\partial^2 l}{\partial \sigma_{ij}\partial \sigma_{kk}} = -\frac{1}{2}(D_{kj}[\Sigma^{-1}]_{ki} + D_{ki}[\Sigma^{-1}]_{kj}), \quad i \neq j \\ \frac{\partial^2 l}{\partial \sigma_{ij}\partial \sigma_{kl}} = -\frac{1}{2}(D_{ki}[\Sigma^{-1}]_{lj} + D_{li}[\Sigma^{-1}]_{kj} + D_{kj}[\Sigma^{-1}]_{li} + D_{lj}[\Sigma^{-1}]_{ki}), \quad i \neq j, k \neq l \end{cases}$$

(b)

$$I(\mu, \Sigma) = -E[\nabla^2 l(\mu, \Sigma)] = -E\begin{pmatrix} \frac{\partial^2 l}{\partial \mu^2} & \frac{\partial^2 l}{\partial \mu \partial \Sigma} \\ \frac{\partial^2 l}{\partial \Sigma \partial \mu} & \frac{\partial^2 l}{\partial \Sigma^2} \end{pmatrix} = \begin{pmatrix} -E(\frac{\partial^2 l}{\partial \mu^2}) & 0 \\ 0 & -E(\frac{\partial^2 l}{\partial \Sigma^2}) \end{pmatrix}$$

The elements of $-E(\frac{\partial^2 l}{\partial \mu^2})$ are:

$$\begin{cases} -E[\frac{\partial^2 l}{\partial \mu_i^2}] = n[\Sigma^{-1}]_{ii} \\ -E[\frac{\partial^2 l}{\partial \mu_i \partial \mu_j}] = 2n[\Sigma^{-1}]_{ij}, \quad i \neq j \end{cases}$$

Moreover, since $E(C) = \sum_{i=1}^{n} E[(X_i - \mu)(X_i - \mu)^T)] = n\Sigma$, then:

$$E(D) = E(\Sigma^{-1}C\Sigma^{-1} - B) = n\Sigma^{-1} - B := F_{p\times p}$$

Therefore, the elements of $-E(\frac{\partial^2 l}{\partial \Sigma^2})$ are:

$$\begin{cases} -E(\frac{\partial^2 l}{\partial \sigma_{ii} \partial \sigma_{kk}}) = E(\frac{1}{2} D_{ki}[\Sigma^{-1}]_{ik}) = \frac{1}{2} E(D_{ki})[\Sigma^{-1}]_{ik} = \frac{1}{2} F_{ki}[\Sigma^{-1}]_{ik} \\ -E(\frac{\partial^2 l}{\partial \sigma_{ii} \partial \sigma_{kl}}) = \frac{1}{2}(F_{li}[\Sigma^{-1}]_{ik} + F_{ki}[\Sigma^{-1}]_{il}), \quad k \neq l \\ -E(\frac{\partial^2 l}{\partial \sigma_{ij} \partial \sigma_{kk}}) = \frac{1}{2}(F_{kj}[\Sigma^{-1}]_{ki} + F_{ki}[\Sigma^{-1}]_{kj}), \quad i \neq j \\ -E(\frac{\partial^2 l}{\partial \sigma_{ij} \partial \sigma_{kl}}) = \frac{1}{2}(F_{ki}[\Sigma^{-1}]_{lj} + F_{li}[\Sigma^{-1}]_{kj} + F_{kj}[\Sigma^{-1}]_{li} + F_{lj}[\Sigma^{-1}]_{ki}) \quad i \neq j, k \neq l \end{cases}$$

(c)

```r
sqrtm <- function (A) {
  # Obtain matrix square root of a matrix A
  a = eigen(A)
  sqm = a$vectors %*% diag(sqrt(a$values)) %*% t(a$vectors)
  sqm = (sqm+t(sqm))/2
}

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 p 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) + matrix(mu,n,p, byrow = TRUE)
  datan
}

res = gen(200, 3, c(-1,1,2), matrix(c(1,0.7,0.7,0.7,1,0.7,0.7,0.7,1),3))

print(res[1:3,])
```

```
##            [,1]      [,2]     [,3]
## [1,] -0.1065871 1.4852948 2.624447
## [2,] -1.0153403 2.4358252 3.196376
## [3,] -1.1833315 0.6469248 2.329121
```