MTH210: Assignment 3 Solutions

Assignment 3 Solutions

First, we write the closed form expression of the determinant of the variance.

$$
\begin{aligned}
\det\left(N\text{Var}(\hat{\theta}_g)\right) &= \det\left(|\Sigma|^{1/2}|(2I_2 - \Sigma^{-1})^{-1}|^{1/2}(2I_2 - \Sigma^{-1})^{-1}\right)\\
&= \frac{|\Sigma|}{|(2I_2 - \Sigma^{-1})|}\det\left((2I_2 - \Sigma^{-1})^{-1}\right)\\
&= \frac{|\Sigma|}{|(2I_2 - \Sigma^{-1})|^2}
\end{aligned}
$$

Further, we note that

$$
|\Sigma| = \sigma^4(1-\rho^2)\,,
$$

and

$$
\begin{aligned}
\Sigma^{-1} = \frac{1}{|\Sigma|}\begin{bmatrix} \sigma^2 & -\rho\sigma^2 \\ -\rho\sigma^2 & \sigma^2 \end{bmatrix} &= \begin{bmatrix} \dfrac{1}{\sigma^2(1-\rho^2)} & -\dfrac{\rho}{\sigma^2(1-\rho^2)} \\ -\dfrac{\rho}{\sigma^2(1-\rho^2)} & \dfrac{1}{\sigma^2(1-\rho^2)} \end{bmatrix}\\
\Rightarrow 2I_p - \Sigma^{-1} &= \begin{bmatrix} 2 - \dfrac{1}{\sigma^2(1-\rho^2)} & \dfrac{\rho}{\sigma^2(1-\rho^2)} \\ \dfrac{\rho}{\sigma^2(1-\rho^2)} & 2 - \dfrac{1}{\sigma^2(1-\rho^2)} \end{bmatrix}\\
\Rightarrow |2I_p - \Sigma^{-1}| &= \frac{1}{1-\rho^2}\left[\left(2 - \frac{1}{\sigma^2}\right)^2 - 4\rho^2\right].
\end{aligned}
$$

Putting these two together, and letting $t = \sigma^2$

$$
f = \det\left(N\mathrm{Var}(\hat{\theta}_g)\right) = \frac{|\Sigma|}{|2I_p - \Sigma^{-1}|^2} = \frac{\sigma^4(1-\rho^2)^3}{\left[\left(2 - \dfrac{1}{\sigma^2}\right)^2 - 4\rho^2\right]^2} = \frac{t^2(1-\rho^2)^3}{\left[\left(2 - \dfrac{1}{t}\right)^2 - 4\rho^2\right]^2}
$$

Taking log and derivatives

$$
\begin{aligned}
\log f &= 2\log t + 3\log(1-\rho^2) - 2\log\left[\left(2 - \frac{1}{t}\right)^2 - 4\rho^2\right] \\
\frac{\partial \log f}{\partial t} &= \frac{2}{t} - \frac{4}{t^2}\frac{2 - \dfrac{1}{t}}{\left(2 - \dfrac{1}{t}\right)^2 - 4\rho^2} \\
\frac{\partial \log f}{\partial \rho} &= \frac{16\rho}{\left(2 - \dfrac{1}{t}\right)^2 - 4\rho^2} - \frac{6\rho}{1-\rho^2}\,.
\end{aligned}
$$

With these gradients, we implement a gradient descent algorithm.

```
obj <- function(sig2, rho)
{
  2*log(sig2) + 3*log(1 - rho^2) - 2*log( (2 - 1/sig2)^2 - 4*rho^2)
}

grad <- function(sig2, rho)
{
  g1 <- 2/sig2 - 4/(sig2^2) * (2 - 1/sig2)/( (2 - 1/sig2)^2 - 4*rho^2)
  g2 <- 16*rho/( (2 - 1/sig2)^2 - 4*rho^2) - 6*rho/(1 - rho^2)
  return(-c(g1, g2))
}

curr <- c(1, .1)  # unit variance and close to no correlation
store <- curr
t <- .1           # tuning this is tricky, so play around with multiple ts
tol <- 1e-10
iter <- 0
diff <- 100
b <- 1
while((diff > tol) && iter < 100)  #not too many iterations
{
  iter <- iter+1
  old <- curr
  curr = old +  t*grad(old[1], old[2])
```

```
  store <- rbind(store, curr)
  diff <- sum( (curr - old)^2)
}
curr
```

```
[1]  1.499936e+00 -2.843979e-12
```