Introduction to Laplace approximation with examples in R

Brice Ozenne

July 29, 2019

Ressource: http://www.imm.dtu.dk/~hmad/GLM/Slides_2012/week11/lect11.pdf

1 Principle

Consider a positive function F of two variables x and y. We would like to marginalize F over x:

$$\begin{aligned} F(y) &= \int_x F(x,y)dx \\ &= \int_x \exp(f(x,y))dx \end{aligned}$$

where $f(x,y) = log(F(x,y))$ Assume that $f(x,y)$ admits a global maximum (with respect to x) at $\hat{x}$. Then, under some regularity assumption, we can use a taylor expansion to obtain:

$$\begin{aligned} F(y) &= \int_x F(x,y)dx \\ &= \int_x \exp\left(f(\hat{x},y) + \frac{(\hat{x}-x)^2}{2} f''(\hat{x},y) + o_p\left((\hat{x}-x)^2\right)\right) dx \\ &= F(\hat{x},y) \int_x \exp\left(\frac{(\hat{x}-x)^2}{2} f''(\hat{x},y)\right) dx + o_p\left((\hat{x}-x)^2\right) \\ &= F(\hat{x},y)\sqrt{\frac{2\pi}{|f''(\hat{x},y)|}} + o_p\left((\hat{x}-x)^2\right) \end{aligned}$$

2 Application

2.1 Linear mixed model

2.1.1 Formula

Consider the following linear mixed model:

$$Y_{ij} = X_{ij}\beta + u_i + \varepsilon_{ij}$$

where $\varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$ and $u_i \sim \mathcal{N}(0, \tau)$. Then denoting $\theta = (\beta, \sigma^2, \tau)$:

$$\begin{aligned}
F(u_i, \theta) &= \left(\prod_{j=1}^{m} \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left(-\frac{1}{2\sigma^2}(Y_{ij} - X_{ij}\beta - u_i)^2 \right) \right) \frac{1}{(2\pi\tau)^{1/2}} \exp\left(-\frac{u_i^2}{2\tau} \right) \\
f(u_i, \theta) &= -\sum_{j=1}^{m} \frac{1}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}(Y_{ij} - X_{ij}\beta - u_i)^2 - \frac{1}{2}\log(2\pi\tau) - \frac{u_i^2}{2\tau} \\
&= -\frac{m}{2}\log(2\pi\sigma^2) - \frac{1}{2}\log(2\pi\tau) - \frac{1}{2\sigma^2}\sum_{j=1}^{m}(Y_{ij} - X_{ij}\beta - u_i)^2 - \frac{1}{2\tau}u_i^2
\end{aligned}$$

So

$$f''(u_i, \theta) = -\frac{m}{\sigma^2} - \frac{1}{\tau}$$

and we note that a second order Taylor expansion is enough since $f'''(u_i, \theta) = 0$. Therefore we get for the log-likelihood:

$$\begin{aligned}
f(\theta) = &-\frac{m}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{j=1}^{m}(Y_{ij} - X_{ij}\beta - \hat{u}_i)^2 \\
&- \frac{1}{2\tau}\hat{u}_i^2 - \frac{1}{2}\log(2\pi\tau) \\
&+ \frac{1}{2}\log\left(\frac{2\pi}{m\sigma^{-2} + \tau^{-1}} \right)
\end{aligned}$$

2.1.2 R code

Load packages:

```
library(lava)
library(lavaSearch2)
library(mvtnorm)
library(nlme)
library(data.table)
```

Simulate data

```
mSim <- lvm(c(Y1,Y2,Y3,Y4,Y5)~tau,
         tau ~ X1+X2)
latent(mSim) <- ~tau
transform(mSim, id ~ tau) <- function(x){1:NROW(x)}

m <- lvm(c(Y1,Y2,Y3,Y4,Y5)~1*tau,
         tau ~ 0+X1+X2)
variance(m, ~Y1) <- "sigma"
variance(m, ~Y2) <- "sigma"
variance(m, ~Y3) <- "sigma"
variance(m, ~Y4) <- "sigma"
variance(m, ~Y5) <- "sigma"

set.seed(10)
n <- 100
dW <- as.data.table(lava::sim(mSim, n = n, latent = FALSE))
dL <- melt(dW, id.vars = c("id","X1","X2"), variable.name = "time",
    value.name = "Y")
```

Fit linear mixed effect model:

```
e.lava <- estimate(m, dW)
e.nlme <- lme(Y ~ -1 + X1 + X2 + time,
              random = ~ 1|id, data = dL, method = "ML")

logLik(e.lava)
logLik(e.nlme)
```

```
'log Lik.' -810.9451 (df=9)
'log Lik.' -810.9451 (df=9)
```

Compute marginal likelihood:

```
logLik_marginal <- function(model, data = NULL, param = NULL){
    ## initialize
    if(is.null(data)){
        data <- as.data.frame(model.frame(model))
    }
    if(is.null(param)){
        param <- coef(model)
    }

    ## find sufficient statisitcs
    Sigma <- getVarCov2(model, data = data, param = param)
    epsilon <- residuals(model, newdata = data, p = param)
    m <- NCOL(epsilon)

    ## compute log likelihood
    out <- dmvnorm(x = epsilon, mean = rep(0, m), sigma = Sigma, log =
   TRUE)
    ## n <- NROW(epsilon)
    ## out <- -(n*m/2)*log(2*pi) - (n/2)*log(det(Sigma)) - 0.5*sum((
   epsilon %*% solve(Sigma)) * epsilon)
    return(out)
}
sum(logLik_marginal(e.lava))
```

```
[1] -810.9451
```

Compute conditional likelihood:

```
logLik_conditional <- function(model, Zb = NULL,
                               data = NULL, param = NULL){

    ## initialize
    if(is.null(data)){
        data <- as.data.frame(model.frame(model))
    }
    if(is.null(param)){
        param <- coef(model)
    }

    ## identify variance of the random effect
    df.type <- coefType(model, as.lava=FALSE)
    df.type <- df.type[!is.na(df.type$detail),]
    tau <- param[df.type[df.type$detail=="Psi_var","param"]]

    ## estimate sufficient statistics
    Sigma.m <- getVarCov2(model, data = data, param = param)
    Sigma.c <- Sigma.m - tau
    YmXB <- residuals(model, newdata = data, p = param)

    ## compute random effects
    m <- NCOL(YmXB)
    if(is.null(Zb)){
        Z <- matrix(1, nrow = 1, ncol = m)
        Omega <- solve(Z %*% solve(Sigma.c) %*% t(Z) + 1/tau) %*% Z %*%
 solve(Sigma.c)
        Zb <- as.double(Omega %*% t(YmXB)) ## cbind(ranef(e.nlme),Zb)
    }
    epsilon <- YmXB - Zb

    out1 <- dmvnorm(x = epsilon, mean = rep(0, m), sigma = Sigma.c, log
    = TRUE)
    out2 <- dnorm(x = Zb, mean = 0, sd = sqrt(tau), log = TRUE)
    return(out1 + out2)
}
```

Laplace approximation

```
d2.f <- 5/coef(e.lava)["Y1~~Y1"]+1/coef(e.lava)["tau~~tau"]
sum(logLik_conditional(e.lava) + (1/2)*log(2*pi/d2.f))
```

```
[1] -810.9451
```

2.2 General gaussian model

Consider the following gaussian mixed model:

$$Y_i \sim \mathcal{N}\left(\mu(X_i, \beta, u_i), \Sigma\right)$$

Denoting by m the number of observations per individual we have:

$$\begin{aligned}
f(u_i, \theta) &= -\frac{1}{2}\log\left((2\pi)^m|\Sigma|\right) - \frac{1}{2}(Y_i - \mu(X_i, \beta, u_i))\Sigma^{-1}(Y_{ij} - \mu(X_i, \beta, u_i))^{\intercal} - \frac{1}{2}\log(2\pi\tau) - \frac{u_i^2}{2\tau} \\
&\propto -\frac{1}{2}\log|\Sigma| - \frac{1}{2}\log(\tau) - \frac{1}{2}(Y_i - \mu(X_i, \beta, u_i))\Sigma^{-1}(Y_{ij} - \mu(X_i, \beta, u_i))^{\intercal} - \frac{1}{2\tau}u_i^2
\end{aligned}$$

Since

$$f''(u_i, \theta) = -\mu'(X_i, \beta, u_i)\Sigma^{-1}\mu'(X_i, \beta, u_i)^{\intercal} - \frac{1}{\tau}$$

we get:

$$\begin{aligned}
f(\theta) = &-\frac{1}{2}\log\left((2\pi)^m|\Sigma|\right) - \frac{1}{2}(Y_i - \mu(X_i, \beta, u_i))\Sigma^{-1}(Y_{ij} - \mu(X_i, \beta, u_i))^{\intercal} \\
&- \frac{1}{2}\log(2\pi\tau) - \frac{u_i^2}{2\tau} \\
&+ \frac{1}{2}\log\left(\frac{2\pi}{\mu'(X_i, \beta, u_i)\Sigma^{-1}\mu'(X_i, \beta, u_i)^{\intercal} + \tau^{-1}}\right)
\end{aligned}$$