

Specifying a fixed covariance matrix in INLA

Austin Schumacher

2021-04-13

Simulations

Intercept only

First, notation: let $i = 1, \dots, n$ index observations (regions) and $c = 1, 2$ index “causes” of death.

We will simulate n bivariate normal observations with the following data generating mechanism

$$\begin{aligned}y_{ic} &= \alpha + \epsilon_{ic} \\ \epsilon_i &\sim N(0, \Sigma) \\ \Sigma &= \begin{bmatrix} 1 & 0.5 \\ 0.5 & 2 \end{bmatrix}\end{aligned}$$

We will set $\alpha = 1$ for the following simulations.

In INLA, to specify a fixed covariance matrix for pairs of observations, we will ...

Gamma(1000,1000) prior the precision of the RE used to specify the fixed covariance (so the precision prior has a mean of 1 and a variance of 0.0001) [actually it's a logGamma prior on the log precision], along with fixed Gaussian error with precision 1000000

```
library(INLA); library(mvtnorm);
set.seed(80085)

# simulation settings and results storage
nsim <- 100
alpha.res <- rep(NA, nsim)
coverage.res <- rep(NA, nsim)
covprec.res <- rep(NA, nsim)

# parameters
n <- 1000
ncause <- 2
alpha <- 1
Sigma <- matrix(c(1, 0.5, 0.5, 2), nrow = 2)

# make block diagonal covariance matrix for use in INLA model for fixed cov
bdiag_matlist <- list()
for(i in 1:n) { bdiag_matlist[[i]] <- Sigma }
block_cov <- bdiag(bdiag_matlist)
```

```

# modeling info:
# priors
fe.prec <- list(prec.intercept = 0,
               prec = 0)
cov_prior <- list(prec = list(prior = "loggamma", param = c(10000, 10000)))

# model formula: we have a generic RE with a specified C matrix and a tight prior on the precision cent
m.form <- y ~ 1 + f(obs, model='generic0', Cmatrix = block_cov,
                  hyper = cov_prior)

for (s in 1:nsim) {
  # cat(paste0("Starting sim ", s, "...\n"))
  epsilon <- rmvnorm(n, mean = c(0, 0), sigma = Sigma)
  ymat <- alpha + epsilon
  y <- as.vector(ymat)

  # create dataframe
  dat <- data.frame(reg = rep(1:n, ncause),
                   cause = rep(1:ncause, each = n),
                   y = y)
  dat <- dat[order(dat$reg, dat$cause),]
  dat$obs <- 1:nrow(dat)
  rownames(dat) <- NULL

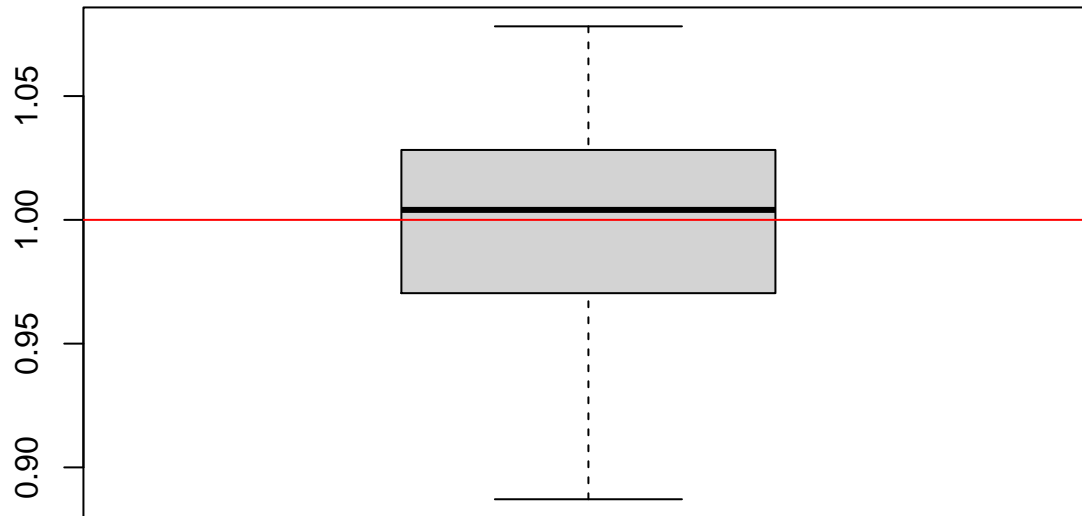
  mod1 <- inla(m.form,
              data = dat,
              family = "gaussian",
              control.fixed = fe.prec,
              control.predictor=list(compute=TRUE),
              control.compute=list(config = TRUE),
              control.family = list(hyper = list(prec = list(initial = log(1), fixed=TRUE))),
              scale = 1000000)

  alpha.res[s] <- mod1$summary.fixed$`0.5quant`
  coverage.res[s] <- alpha > mod1$summary.fixed$`0.025quant` & alpha < mod1$summary.fixed$`0.975quant`
  covprec.res[s] <- mod1$summary.hyperpar$`0.5quant`
  # mod1$summary.fixed
  # mod1$summary.hyperpar
  # plot(as.vector(epsilon), mod1$summary.random$obs$`0.5quant`)
  # abline(0, 1, col = "red")
}

boxplot(alpha.res, main = "Posterior medians of estimated alpha paramter")
abline(h = 1, col = "red")

```

Posterior medians of estimated alpha paramter

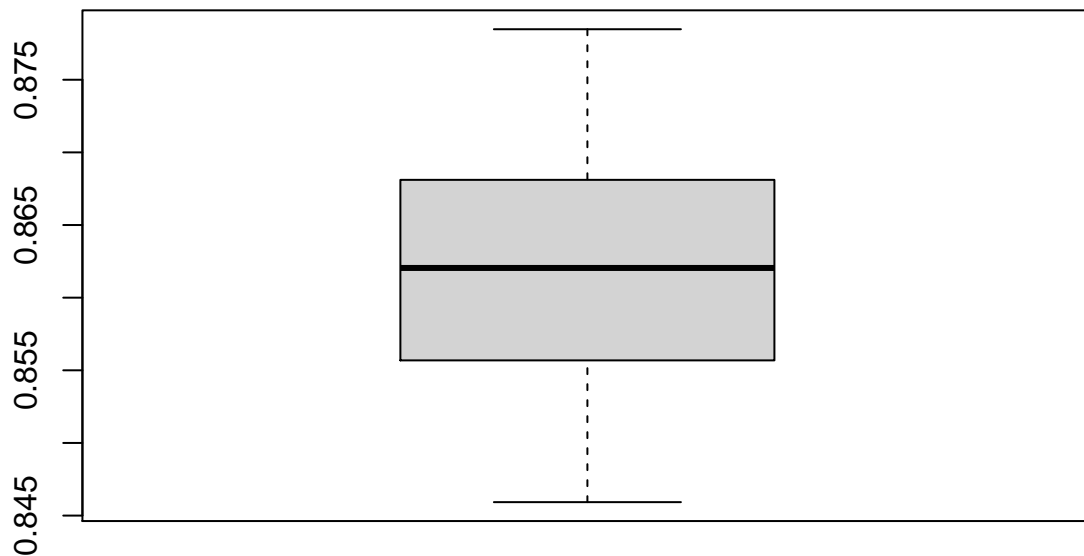


```
mean(coverage.res)
```

```
## [1] 0.58
```

```
boxplot(covprec.res, main = "Posterior medians of estimated precision for RE used to specify fixed covar",
abline(h = 1, col = "red"))
```

Posterior medians of estimated precision for RE used to specify fixed cov



IID REs

Spatial REs

Modeling BGD data