Integrated Nested Laplace Approximations (INLA)

Dr. Jarad Niemi

STAT 615 - Iowa State University

December 7, 2021

Resources

Original manuscript: Rue et al. (2009)
This development based on: Martino and Riebler (2019)
R package: R-INLA (which is actually "INLA")

Latent Gaussian Models

Data level model:

$$y|\phi, \theta \sim \prod_{i=1}^{n} p(y_i|\phi_i, \theta)$$

Latent Gaussian hierarchical structure:

$$\phi|\theta \sim N(0, \Omega^{-1}(\theta))$$

Prior:

$$\theta \sim p(\theta)$$

INLA Applicability

1. Observations are conditionally independent
2. Hyperparameter vector θ is small (< 15)
3. Ω is sparse
4. linear predictor depends linearly on unknown smooth function of covariates
5. inferential interest lies in univariate posterior marginals, e.g. $p(\phi_i|y)$ and $p(\theta_j|y)$ rather than the joint posterior

Generalized mixed effect model structure

Data model:

$$y_i|\phi_i, \theta \sim p(y_i|\phi_i, \theta) \quad \text{with} \quad E[Y_i|\phi_i, \theta] = g^{-1}(\phi)$$

$$\phi_i = \mu + \sum_j \beta_j x_{ij} + \sum_k w_k f^k(u_{ik})$$

with known x (covariates) and w (weights) and Gaussian "prior" on μ, β, and f^k for $k = 1, \ldots, K$.

INLA Objective

True posterior:

$$p(\phi, \theta|y) \propto \exp\left(-\frac{1}{2}\phi^\top \Omega(\theta)\phi + \sum_{i=1}^{n} \log(p(y_i|\phi_i, \theta)) + \log p(\theta)\right)$$

INLA objective is to approximate

$$\pi(\theta_j|y) = \int \int p(\phi, \theta|y) d\phi d\theta_{-j} = \int p(\theta|y) d\theta_{-j}$$

and

$$\pi(\phi_i|y) = \int \int p(\phi, \theta|y) d\phi_{-i} d\theta = \int p(\phi_i|\theta, y) p(\theta|y) d\theta$$

Hyperparameter approximation

Target posterior

$$p(\theta|y) = \frac{p(\phi,\theta|y)}{p(\phi|\theta,y)} \propto \frac{p(y|\phi,\theta)p(\phi,\theta)p(\theta)}{p(\phi|\theta,y)}$$

Approximate denominator at a specific value θ^k

$$\tilde{p}(\theta^k|y) \propto \frac{p(y|\phi,\theta^k)p(\phi,\theta^k)p(\theta^k)}{\tilde{p}_G(\phi|\theta^k,y)}$$

where $\tilde{p}_G$ indicates a Gaussian approximation constructed by matching the mode and curvature, i.e. Laplace approximation (Tierney and Kadane, 1986).
Expensive part is Cholesky decomposition of $Q(\theta^k)$ to obtain denominator approximation.

Latent Field Posterior

Target (approximate) posterior

$$p(\phi_i|\theta^k, y) = \frac{p(\phi|\theta^k, y)}{p(\phi_{-i}|\phi_i, \theta^k, y)} \propto \frac{p(y|\phi, \theta^k)p(\phi, \theta^k)p(\theta^k)}{p(\phi_{-i}|\phi_i, \theta^k, y)}$$

(Does $p(\theta^k)$ need to be in this expression? Does θ need to have k superscript?)
Denominator approximation (maybe?):

$$\tilde{p}(\phi_i|\theta^k, y) \propto \frac{p(y|\phi_{-i}, \phi_i^k, \theta^k)p(\phi_{-i}, \phi_i^k, \theta^k)p(\theta^k)}{\tilde{p}_G(\phi_{-i}|\phi_i^k, \theta^k, y)}$$

using Laplace approximation around each individual ϕ_i^k.
(Add Simplified Laplace approximation to reduce computational requirements.)

Numerical integration for hyperparameters θ

Options:

1. Grid (if θ is 1 or 2 dimensions)
2. Central composite design (if θ is more than 2 dimensions)
3. Empirical Bayes, i.e. plug in $\hat{\theta}_{EB} = \text{argmax}_\theta \tilde{p}(\theta|y)$

INLA Computing Scheme

1. Explore $\tilde{p}(\theta|y)$: find mode of $\tilde{p}(\theta|y)$ and locate $\{\theta^1, \ldots, \theta^K\}$ with high density of $\tilde{p}(\theta|y)$.
2. Compute $\tilde{p}(\theta^1|y), \ldots, \tilde{p}(\theta^K|y)$.
3. For each θ^k, find $\pi(\phi_i|\theta^k, y)$ using one of three (?) possible approximations: Laplace, Simplified Laplace, or Gaussian.
4. Via numerical integration, find

$$\tilde{p}(\phi_i|y) = \sum_{k=1}^{K} \tilde{p}(x_i|\theta^k, y)\tilde{p}(\theta^k|y)\delta_k$$

where δ_k are appropriate weights.

Salmonella data

```r
library("INLA")
head(Salm)
```

```
##    y dose rand
## 1 15    0    1
## 2 21    0    2
## 3 29    0    3
## 4 16   10    4
## 5 18   10    5
## 6 21   10    6
```

Salmonella model

Let Y_{ij} be the number of salmonella bacteria for the jth observation under dose i.

Data model:

$$Y_{ij} \stackrel{ind}{\sim} Po(\lambda_{ij})$$

Latent Gaussian structure:

$$\log(\lambda_{ij}) = \beta_0 + \beta_1 \log(dose[i] + 10) + \beta_2 dose[i] + u_{ij}, \quad u_{ij} \stackrel{ind}{\sim} N(0, \sigma^2)$$

Prior:

$$\sigma^2 \sim \text{PC-prior?}$$

Run INLA

```
# specify the prior for the log precision parameter
my.hyper <- list(theta = list(prior="pc.prec", param=c(1,0.01)))

# specify the linear predictor
formula <- y ~ log(dose + 10) + dose + f(rand, model = "iid", hyper = my.hyper)

# run INLA
result <- inla(formula=formula, data=Salm, family="Poisson")
```

```
summary(result)
```

```
##
## Call:
##    "inla(formula = formula, family = \"Poisson\", data = Salm)"
## Time used:
##     Pre = 0.556, Running = 0.218, Post = 0.245, Total = 1.02
## Fixed effects:
##                  mean    sd 0.025quant 0.5quant 0.975quant   mode kld
## (Intercept)     2.168 0.359      1.451    2.170      2.874  2.174   0
## log(dose + 10)  0.313 0.098      0.119    0.313      0.506  0.313   0
## dose           -0.001 0.000     -0.002   -0.001      0.000 -0.001   0
##
## Random effects:
##   Name   Model
##     rand IID model
##
## Model hyperparameters:
##                     mean    sd 0.025quant 0.5quant 0.975quant  mode
## Precision for rand 20.84 18.28       5.72    16.46      61.71 11.92
##
## Expected number of effective parameters(stdev): 12.05(2.08)
## Number of equivalent replicates : 1.49
##
## Marginal log-Likelihood:  -83.68
```

References I

Sara Martino and Andrea Riebler. Integrated nested Laplace approximations (INLA), 2019.

Håvard Rue, Sara Martino, and Nicolas Chopin. Approximate Bayesian inference for latent gaussian models by using integrated nested Laplace approximations. *Journal of the royal statistical society: Series b (statistical methodology)*, 71(2):319–392, 2009.

Luke Tierney and Joseph B Kadane. Accurate approximations for posterior moments and marginal densities. *Journal of the American Statistical Association*, 81(393):82–86, 1986.