Métodos computacionais para inferência estatística - modelos dinâmicos via INLA

Elias Teixeira Krainski

LEG: Laboratório de Estatística e Geoinformação
Universidade Federal do Paraná

31 de julho de 2012

Laboratório de Estatística e Geoinformação

Sumário

1. Modelo Simples
2. Regressão dinâmica
3. Modelo espaço temporal

O modelo dinâmico mais simples

Idéia: distribuição conjunta.

Vetorialmente temos:

$$\mathbf{y} = \boldsymbol{\theta} + \boldsymbol{\nu} \tag{1}$$

$$\boldsymbol{\theta}_{-1} = \phi \boldsymbol{\theta}_{-n} - \boldsymbol{\omega} \tag{2}$$

$\boldsymbol{\theta}_{-i}$ é o vetor $\boldsymbol{\theta}$ sem a entrada i

Distribuição de $\boldsymbol{\theta}$

$$\boldsymbol{\theta}|W \propto exp\{-\frac{W}{2}\boldsymbol{\theta}'\mathbf{R}\boldsymbol{\theta}\}$$

$$\mathbf{R} = \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 & 0 & 0 \\ -1 & 2 & -1 & \cdots & 0 & 0 & 0 \\ 0 & -1 & 2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ 0 & 0 & 0 & \cdots & 2 & -1 & 0 \\ 0 & 0 & 0 & \cdots & -1 & 2 & -1 \\ 0 & 0 & 0 & \cdots & 0 & -1 & 1 \end{pmatrix}.$$

Laboratório de Estatística e Geoinformação

O modelo dinâmico mais simples

Simula dados do modelo simples

```
> n <- 100; V <- 2; W <- 3; rho <- 0.9
> w <- rnorm(n, 0, sqrt(W))
> v <- rnorm(n, 0, sqrt(V))
> x <- rep(0, n); x[1] <- w[1]
> for (i in 2:n)
+   x[i] <- rho*x[i-1] + w[i]
> y <- x + v
```

Encontra as marginais a posteriori

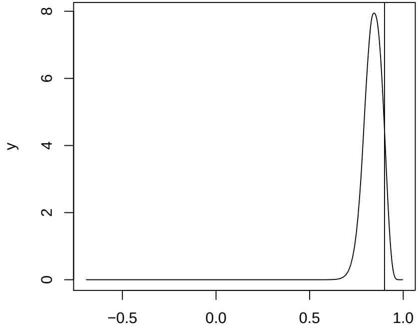
```
> res1 <- inla(y ~ f(t, model="ar1"),
+              data=data.frame(y=y,t=1:n))
> res1$cpu
 Pre-processing    Running inla Post-processing           Total
        0.14148         0.16697         0.06512         0.37357
```

Sumários

```
> res1$summary.fixed
              mean    sd 0.025quant 0.5quant 0.975quant kld
(Intercept) 0.3656 1.637     -3.108   0.4016      3.595   0
> res1$summary.hyperpar
                                    mean    sd 0.025qua
Precision for the Gaussian observations 0.29994 0.091    0.163
Precision for t                     0.07082 0.027    0.029
Rho for t                           0.89278 0.047    0.783
                                    0.5quant 0.975quant
Precision for the Gaussian observations  0.28559     0.5179
Precision for t                      0.06755     0.1318
Rho for t                            0.89905     0.9639
```

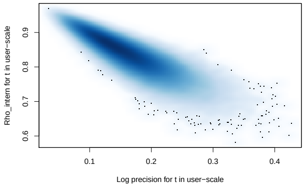
Marginais e HPD

```
> inla.hpdmarginal(0.95, res1$marginals.hyperpar[[3]])
                    low      high
level:0.95 0.8014188 0.972016
> plot(res1$marginals.hyperpar[[3]], type="l")
```



Simula da conjunta dos hyperparâmetros

```
> system.time(shyp <- inla.hyperpar.sampler(1e5, res1))
   user  system elapsed
  2.668   0.012   2.687
> cor(unname(shyp))
           [,1]       [,2]       [,3]
[1,]  1.0000000 -0.2617281 -0.4022859
[2,] -0.2617281  1.0000000 -0.4276081
[3,] -0.4022859 -0.4276081  1.0000000
> smoothScatter(shyp[,2:3])
```



Regressão dinâmica

```
dad <- data.frame(y=y, i0=1:n, i1=1:n,
                  i2=1:n, x1=x[2,], x2=x[2,])
hyp <- list(theta1=list(param=c(.5,.1), initial=0),
            theta2=list(param=c(.5,.3)))
mod <- inla(y ~ 0 + f(i0, model="ar1", hyper=hyp) +
            f(i1, x1, model="ar1", hyper=hyp) +
            f(i2, x2, model="ar1", hyper=hyp),
            data=dad,
            control.data=list(hyper=list(theta=list(initial=0,
                                               param=c(.5,.1)))))
```

Modelo espaço temporal

- $$o_{i,t} \sim Poisson(\psi_{i,t} E_{i,t})$$
  - o_i número de óbitos infantis no município i no ano t
  - E_i número esperado de óbitos
  - $\psi_{i,t}$ é o risco relativo
- o preditor linear $\eta_{i,t}$:
  $$log(\psi_{i,t}) = \eta_{i,t} = \alpha_{i,t} + \beta_{i,t} X_{i,t}$$
- $$\alpha_{i,t} = \phi_\alpha \alpha_{i,t-1} + w1_{i,t} \qquad \beta_{i,t} = \phi_\beta \beta_{i,t-1} + w2_{i,t}$$
- $$w1_{i,t} | w1_{-i,t} \sim N(\sum_{j \sim i} w1_{j,t}/d_i, \sigma^2_{w1}/d_i)$$
- Precisão a priori de $w1_{,t}$ é Q_S
- Precisão de evolução temporal Q_T
- Matriz de precisão de $w1$ é $Q = Q_S \otimes Q_T$

Diferentes modelos para $\eta_{i,t}$

m_0 :	α_0
m_1 :	$\alpha_0 + \alpha_{0,t}$
m_2 :	$\alpha_0 + \alpha_{i,0}$
m_3 :	$\alpha_0 + \alpha_{0,t} + \alpha_{i,0}$
m_4 :	$\alpha_0 + \alpha_{i,t}$

Fórmulas para os cinco modelos.

```
y ~ 1
y ~ 1 + f(i, model="ar1"),
y ~ 1 + f(i, model="besag", graph="dados/mesoc.g"),
y ~ 1 + f(t, model="ar1") +
    f(i, model="besag", graph="dados/mesoc.g"),
y ~ 1 + f(i, model="besag", graph="dados/mesoc.g",
          group=t, control.group=list(model="ar1",
            hyper=list(theta=list(param=c(0,1))))))
```

Laboratório de Estatística e Geoinformação