

Untitled

Ángel

24 de septiembre de 2017

CAR.normal

La distribución a priori intrínseca Gaussiana CAR.

Llamada

$S[1:N] \sim \text{car.normal}(\text{adj}[], \text{weights}[], \text{num}[], \text{tau})$

Siendo `adj[]` vector. Lista los vecinos de cada región (regiones adyacentes).
`weights[]` vector. Tiene la misma longitud que `adj[]`. Lista los pesos de cada región. $w_{ij}=1$ si las regiones i, j son vecinas y 0 si no lo son.
`num[]` vector. Longitud N , Indica el número de vecinos de cada región (n_i)
`tau` escalar. Indica el parámetro de precisión de la prior CAR Gaussiana.

Con los parámetros `adj[]` y `num[]` se tiene la matriz de adyacencia en forma dispersa, al evitar trabajar con matrices densas el método es más rápido. Los tres primeros términos deben ser fijos al hacer la llamada, el parámetro `tau` normalmente se trata como un parámetro desconocido asignándole una distribución prior.

Anotaciones

- Los pesos han de estar no normalizados.
- Una región no puede ser vecina de sí misma. Geobugs no chequea esto, es responsabilidad del usuario.
- Los pesos han de ser simétricos. En este caso Geobugs sí lo comprueba.
- Las distribuciones `car.normal` y `car.l1` están parametrizadas de modo que la media es 0. Esto quiere decir que tendrás que añadir un intercepto separado asignado a una prior uniforme impropia.
- Como `car.normal` y `car.l1` son distribuciones impropias, solo pueden ser usadas como distribuciones prior y no como verosimilitud.

CAR.l1

Es la versión robusta de `CAR.normal`, donde se asume una distribución doble exponencial (Laplace) en lugar de una Gaussiana.

$S[1:N] \sim \text{car.l1}(\text{adj}[], \text{weights}[], \text{num}[], \text{tau})$

Siendo `adj[]` vector. Lista los vecinos de cada región(regiones adyacentes).
`weights[]` vector. Tiene la misma longitud que `adj[]`. Lista los pesos de cada región. $w_{ij}=1$ si las regiones i,j son vecinas y 0 si no lo son.
`num[]` vector. Longitud N, Indica el número de vecinos de cada región (n_i)
`tau` escalar. Indica la inversa del parámetro escala de la prior de Laplace

La llamada que se hace es practicamente la misma.

CAR.proper

La distribución a priori proper Gaussian CAR.

$S[1:N] \sim \text{car.proper}(\mu[], C[], \text{adj}[], \text{num}[], M[], \text{tau}, \text{gamma})$

Siendo `mu[]` vector. Indica la media de cada región. Esta puede ser introducida como dato o asociada a una distribución.
`C[]` vector. Tiene la misma longitud que `adj[]`. Lista los pesos normalizados de cada región. $w_{ij}=1/n_i$ si las regiones i,j son vecinas y 0 si no lo son.
`adj[]` vector. Lista los vecinos de cada región(regiones adyacentes).
`num[]` vector. Longitud N, Indica el número de vecinos de cada región (n_i)
`M[]` vector. Longitud N, que es la diagonal de la matriz de varianzas condicional.
`tau` escalar. Representa la precisión general (inverso de la varianza).
`gamma` escalar. Representa el grado de dependencia espacial.

`mu[]` puede darse como dato o como parámetro desconocido asociado a una distribución a priori.

El parametro gamma debe pertenecer al intervalo $\gamma_{min}, \gamma_{max}$ donde estos dos extremos son los inversos del menor y del mayor autovalor de la matriz $M^{-1/2}CM^{1/2}$ por lo que primeramente habra que calcular dichos autovalores. Se podría hacer en R de manera externa.

Anotaciones

- `C`, `adj`, `num` y `M` deben ser dados como datos, no pueden tener asociada una distribución a priori.
- Los pesos `C` son normalizados (por filas)
- Un area no puede ser vecina de sí misma. Geobugs no comprueba esto, es responsabilidad del usuario.
- Se cumple la simetria de modo que $C_{ij}M_{jj} = C_{ji}M_{ii}$. Geobugs sí comprueba esto.
- Hay que tener cuidado con la prior de gamma, ya que tiene que estar entre las bandas apropiadas.

Expresión como normal multivariante

$$\Sigma = (I - \gamma W)^{-1} M$$

siendo W la matriz de pesos y M la matriz diagonal cuya diagonal es proporcional a la varianza condicional. Se establece $M_{ii} = 1/n_i$ $W_{ij} = 1/n_i$ si son vecinos y 0 en caso contrario. En el mapeo de enfermedades Cressie and Chan escogieron una parametrización alternativa:

$$M_{ii} = 1/E_i \quad W_{ij} = (E_j/E_i)^{1/2} \text{ Siendo } E_i \text{ la población de la región } i.$$

spatial.exp

Los modelos de kriging Gaussiano bayesiano.

$S[1:N] \sim \text{spatial.exp}(\mu[], x[], y[], \text{tau}, \text{phi}, \text{kappa})$

Siendo $\mu[]$ vector. Indica la media de cada región. Esta puede ser introducida como dato o asociada a una distribución.
 $x[]$ vector. Longitud N que indica la coordenada x de las regiones(centroides).
 $y[]$ vector. Longitud N que indica la coordenada y de las regiones(centroides).
 tau escalar. Representa la precisión general (inverso de la varianza).
 phi escalar. Es la tasa de disminucion de la correlación con la distancia
 kappa escalar. Controla el suavizado espacial

La medida de phi dependera de las unidades unidades en las que se tengan las posiciones x, y (metros, kilometros,...). El parámetro kappa debe estar en el intervalo $[0, 2]$

Anotaciones

Este algoritmo puede ser muy lento con bases de datos de un tamaño considerable (el algoritmo es de orden N^3). Puede ser mejor centralizar jerarquicamente este modelo. Por ejemplo:

- No centralizado jerarquicamente

```
for (i in 1:N){
y[i] ~ dnorm(S[i], gamma)
mu[i] <- alpha+beta*z[i]
}
S[1:N] ~ spatial.exp(mu[], x[], y[], tau, phi,1)
```

- Centralizado jerarquicamente

```
for (i in 1:N){
y[i] ~ dnorm(S[i], gamma)
S[i] <- alpha+beta*z[i] + W[i]
mu[i]<-0
}
W[1:N] ~ spatial.exp(mu[], x[], y[], tau, phi,1)
```

A menudo hay muy poca información en los datos sobre los parametros phi y kappa , por lo que se recomienda usar una prior razonablemente informativa o que sean fijados a priori basados en los conocimientos de los expertos o por un análisis exploratorio usando por ejemplo variogramas.

Expresión como normal multivariante

Es trata de una normal multivariante con

$$\Sigma_{ij} = e^{-(\phi d_{ij})^\kappa} \quad \phi > 0, \quad \kappa \in [0, 2]$$

Si ϕ es grande, la correlación decae rápidamente. Este parámetro se suele especificar como una uniforme entre ϕ_{min}, ϕ_{max} de modo que este rango sea adecuado.

spatial.disc

Los modelos de kriging Gaussiano bayesiano.

`S[1:N] ~ spatial.disc(mu[], x[], y[], tau, alpha)`

Siendo mu[]	vector. Indica la media de cada región. Esta puede ser introducida como dato o a
x[]	vector. Longitud N que indica la coordenada x de las regiones(centroides).
y[]	vector. Longitud N que indica la coordenada y de las regiones(centroides).
tau	escalar. Representa la precisión general (inverso de la varianza).
alpha	escalar. Representa el radio del disco centrado en cada centroide

Este modelo es exactamente igual que el anterior, pero en lugar de especificar la matriz de covarianzas como una exponencial, lo hace mediante una función “disco”.

Expresión como normal multivariante

$$\Sigma_{ij} = \begin{cases} \frac{2}{\pi} (\cos^{-1}(d_{ij}/\alpha) - [(d_{ij}/\alpha)(1 - (d_{ij}^2/\alpha^2))]) & d_{ij} < \alpha \\ 0 & d_{ij} > \alpha \end{cases}$$

Si α es grande, decae despacio, y si es pequeña decae rápidamente.

spatial.pred

La interpolación espacial o predicción en localizaciones arbitrarias. Esto se usa en conjunción con los modelos `spatial.exp` o `spatial.disc` para el conjunto de datos. `Spatial.pred` realiza una predicción simultanea en el conjunto de localizaciones objetivo mientras que `spatial.unipred` realiza una predicción para un único punto. La diferencia entre ambos es que `spatial.unipred` ignora las correlaciones entre las localizaciones donde predice mientras que `spatial.pred` realiza una predicción conjunta. El problema de la predicción conjunta es que es muy lenta.

- predicción conjunta

`T[1:P] ~ spatial.pred(mu.T[], x.T[], y.T[], S[])`

- predicción en un único lugar

```
for(j in 1:P) {
T[j] ~ spatial.unipred(mu.T[j], x.T[j], y.T[j], S[])
}
```

Siendo P escalar. Indica el número de localizaciones predichas.

mu.T[] vector. Longitud P especifica la media de cada localización de predicción.
(debería de especificarse del mismo modo que la media para los datos observados S)

x.T[] vector. Longitud P coordenada x de cada lugar de predicción

y.T[] vector. Longitud P coordenada y de cada lugar de predicción

mu.T[j] escalar. media del lugar de predicción j

x.T[j] escalar. coordenada x del lugar de predicción j

y.T[j] escalar. coordenada y del lugar de predicción j

S vector de observaciones ajustadas mediante los modelos spatial.exp o spatial.disc.

pois.conv

La distribución conjunta Poisson-gamma espacial de medias moviles puede ser especificado como conteos en lattices (partición discreta geografica)

```
S ~ dpois.conv(mu[])
```

Siendo mu[] vector. Longitud J que representa la influencia de un conjunto de parámetros distribuidos como una gamma para cada j.

#mu[] se define como sigue:

```
for (j in 1 : J) {
mu[j] <- gamma[j] * k[j]
gamma[j] ~ dgamma(a, b)
}
```

Siendo k[j] un peso espacial dependiendo de las distancias entre las localizaciones y a, b parametros.

Normalmente k[] es calculado de forma externa y leida como dato. Normalmente la distribución suele ser usada para cada elemento por separado:

```
for (i in 1:N) {
S[i] ~ dpois.conv(mu[i, ])
for (j in 1 : J) {
mu[i, j] <- gamma[j] * k[i, j]
}
}
```

Expresión

$$S_i \sim \text{Pois}(\lambda_i)$$

$$\lambda_i = \sum_j \gamma_j * k_{ij}$$

$$k_{ij} = \tau / (2\pi\rho^2) \exp(-d_{ij}^2 / 2\rho^2)$$

mv.car

$S[1:p, 1:N] \sim \text{mv.car}(\text{adj}[], \text{weights}[], \text{num}[], \text{omega}[,])$

Siendo `adj[]` vector. Lista los vecinos de cada región(regiones adyacentes).
`weights[]` vector. Misma longitud que `adj[]` son los pesos no normalizados asociados a cada area.
`num[]` vector. Longitud N Indica el número de vecinos de cada región.
`omega[,]` matriz. p x p matriz de precision de la car.normal intrinseca.

Los tres primeros argumentos deben ser dados como datos, la variable omega normalmente es tratada como un parámetro desconocido asociado a una distribución prior

Anotaciones

- Los pesos son no normalizados.
- Un area no puede ser vecina de sí misma.
- Los pesos deben ser simetricos.
- Esta parametrizada de modo que tiene media 0, por lo que es ecesario añadirle un intercepto.
- Como esta distribución es impropia, solo puede ser usada como distribución prior, no como verosimilitud.