TMA4300Ex2

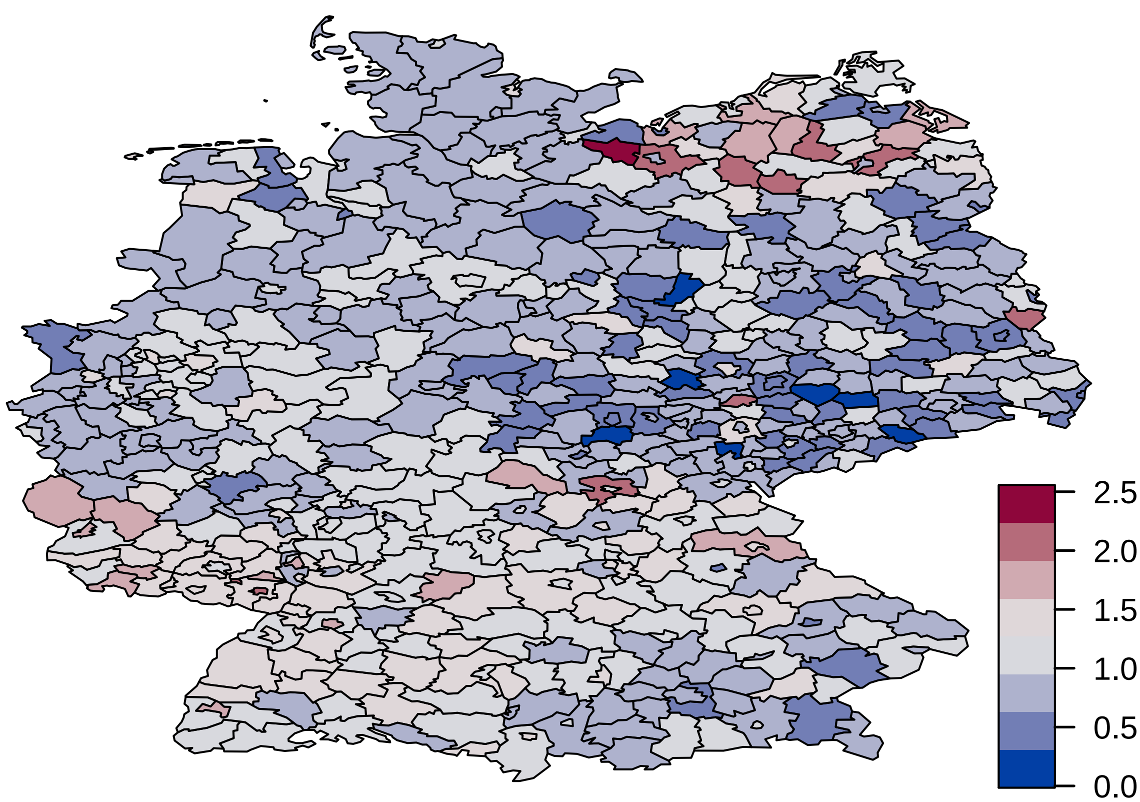
Karine Foss

18 2 2019

```r
# Loading libraries
library(ggplot2)
library(spam)  # load the data
str(Oral)  # see structure of data
```

```
## 'data.frame':    544 obs. of  3 variables:
##  $ Y  : int  18 62 44 12 18 27 20 29 39 21 ...
##  $ E  : num  16.4 45.9 44.7 16.3 26.9 ...
##  $ SMR: num  1.101 1.351 0.985 0.735 0.668 ...
```

```r
# 'data.frame': 544 obs. of 3 variables: $ Y : int 18 62 44 12 18 27
# 20 29 39 21 . . .  $ E : num 16.4 45.9 44.7 16.3 26.9 . . .  $ SMR:
# num 1.101 1.351 0.985 0.735 0.668 . . .
attach(Oral)  # allow direct referencing to Y and E
# load some libraries to generate nice map plots
library(fields, warn.conflict = FALSE)
library(colorspace)
col <- diverge_hcl(8)  # blue - red
# use a function provided by spam to plot the map together with the
# mortality rates
germany.plot(Oral$Y/Oral$E, col = col, legend = TRUE)
```



```r
# Set seed so that the task can be reproduced
set.seed(42)
```

Exercise 1: Derivations

a)

From the definition of conditional probability and the nature of the assumptions in this exercise, we know that

$$
\begin{aligned}
\mathrm{p}(\boldsymbol{\eta}, \boldsymbol{u}, \kappa_u, \kappa_v|\boldsymbol{y}) &\propto \mathrm{p}(\boldsymbol{y}\,|\boldsymbol{\eta}, \boldsymbol{u}, \kappa_u, \kappa_v)\mathrm{p}(\boldsymbol{\eta}\,|\boldsymbol{u}, \kappa_u, \kappa_v)\mathrm{p}(\boldsymbol{u}\,|\kappa_u, \kappa_v)\mathrm{p}(\kappa_u\,|\kappa_v)\mathrm{p}(\kappa_v) \\
&\propto \mathrm{p}(\boldsymbol{y}\,|\boldsymbol{\eta})\mathrm{p}(\boldsymbol{\eta}\,|\boldsymbol{u}, \kappa_v)\mathrm{p}(\boldsymbol{u}\,|\kappa_u)\mathrm{p}(\kappa_u)\mathrm{p}(\kappa_v).
\end{aligned}
$$

By inserting the corresponding probabilities, this becomes

$$
\begin{aligned}
\mathrm{p} &\propto \left(\prod_{i=1}^{n} \left(E_i e^{\eta_i}\right)^{y_i} e^{E_i e^{\eta_i}}\right) |\kappa_v \mathrm{I}|^{\frac{1}{2}}\, e^{-\frac{\kappa_v}{2}(\boldsymbol{\eta}-\boldsymbol{u})^T(\boldsymbol{\eta}-\boldsymbol{u})} \kappa_u^{(n-1)/2} e^{-\frac{\kappa_u}{2}\boldsymbol{u}^T \mathbf{R} \boldsymbol{u}} \kappa_u^{\alpha_u - 1} e^{-\beta_u \kappa_u} \kappa_v^{\alpha_v - 1} e^{-\beta_v \kappa_v} \\
&\propto \kappa_u^{\frac{n-1}{2}+\alpha_u-1} \kappa_v^{\frac{n}{2}+\alpha_v-1} \exp\left\{-\beta_u\kappa_u - \beta_v\kappa_v - \frac{\kappa_v}{2}\left(\boldsymbol{\eta}-\boldsymbol{u}\right)^T\left(\boldsymbol{\eta}-\boldsymbol{u}\right) - \frac{\kappa_u}{2}\boldsymbol{u}^T\mathbf{R}\boldsymbol{u} + \sum_i \left(y_i\eta_i - E_i e^{\eta_i}\right)\right\}.
\end{aligned}
$$

b)

The sum over e^{η_i} in the posterior means that the full conditional of η_i is difficult to sample from. We therefore want to approximate the distribution of $\boldsymbol{\eta}$ it with a multivariate normal in order to use Metropolis-Hastings steps for it. We define the function

$$
f(\eta_i) = y_i\eta_i - E_i e^{\eta_i},
$$

which has derivatives

$$
\begin{aligned}
f'(\eta_i) &= y_i - E_i e^{\eta_i} \\
f''(\eta_i) &= -E_i e^{\eta_i}.
\end{aligned}
$$

This yeilds the following Taylor series expansion of f around z_i,

$$
\begin{aligned}
\tilde{f}(\eta_i) &= y_i z_i - E_i e^{z_i} + (y_i - E_i e^{z_i})(\eta_i - z_i) + \frac{1}{2}(-E_i e^{z_i})(\eta_i - z_i)^2 \\
&= a(z_i) + b(z_i)\eta_i - \frac{1}{2}c(z_i)\eta_i^2,
\end{aligned}
$$

where $a(z_i) = E_i e^{z_i}(z_i - z_i^2/2 - 1)$, $b(z_i) = y_i + E_i e^{z_i}(z_i - 1)$ and $c(z_i) = E_i e^{z_i}$.

c)

As $\mathrm{p}(\theta_i|\boldsymbol{\theta}_{\setminus i}, \boldsymbol{y}) \propto \mathrm{p}(\boldsymbol{\theta}|\boldsymbol{y})$, we can quite easily find the full conditionals from the posterior. Using this, we see that

$$\mathrm{p}(\kappa_u|\boldsymbol{y}, \kappa_v, \boldsymbol{\eta}, \boldsymbol{u}) \propto \kappa_u^{(n-1)/2+\alpha_u-1} e^{-(\beta_u + \frac{1}{2}\boldsymbol{u}^T\mathbf{R}\boldsymbol{u})\kappa_u}.$$

We recognise this as the core of a gamma distribution which means that the full conditional density of κ_u is gamma$(\frac{n-1}{2} + \alpha_u, \beta_u + \frac{1}{2}\boldsymbol{u}^T\mathbf{R}\boldsymbol{u})$. By the same reasoning, we see that gamma$(\frac{n}{2} + \alpha_v, \beta_v + \frac{1}{2}(\boldsymbol{\eta} - \boldsymbol{u})^T(\boldsymbol{\eta} - \boldsymbol{u}))$ is the full conditional density of κ_v. Similarly

$$\begin{aligned}
\mathrm{p}(\boldsymbol{u}|\boldsymbol{y}, \boldsymbol{\eta}.\kappa_u, \kappa_v) &\propto \exp\left\{-\frac{\kappa_v}{2}(\boldsymbol{\eta} - \boldsymbol{u})^T(\boldsymbol{\eta} - \boldsymbol{u}) - \frac{\kappa_u}{2}\boldsymbol{u}^T\mathbf{R}\boldsymbol{u}\right\} \\
&\propto \exp\left\{-\frac{1}{2}\boldsymbol{u}^T(\kappa_u\mathbf{R} + \kappa_v\mathrm{I})\boldsymbol{u} + \kappa_v\boldsymbol{u}^T\boldsymbol{\eta}\right\}.
\end{aligned}$$

We recognise this as the canonical form of a multivariate normal distribution. All these distributions are easy to sample from, and can be used in the Gibbs algorithm directly.

The full conditional distribution for $\boldsymbol{\eta}$ however takes the form

$$\mathrm{p}(\boldsymbol{\eta}|\boldsymbol{y}, \boldsymbol{u}, \kappa_u, \kappa_v) \propto \exp\left\{-\frac{\kappa_v}{2}(\boldsymbol{\eta} - \boldsymbol{u})^T(\boldsymbol{\eta} - \boldsymbol{u}) + \sum_i f(\eta_i)\right\}.$$

This does not correspond to any standard distribution, but by applying the approximation $\tilde{f}(\eta_i)$, we get

$$\begin{aligned}
\mathrm{q}(\boldsymbol{\eta}|\boldsymbol{z}, \boldsymbol{y}, \boldsymbol{u}, \kappa_u, \kappa_v) &\propto \exp\left\{-\frac{\kappa_v}{2}\boldsymbol{\eta}^T\boldsymbol{\eta} + \kappa_v\boldsymbol{\eta}^T\boldsymbol{u} - \frac{1}{2}\boldsymbol{\eta}^T\mathrm{diag}(c(\boldsymbol{z}))\boldsymbol{\eta} + \boldsymbol{\eta}^T b(\boldsymbol{z})\right\} \\
&= \exp\left\{-\frac{1}{2}\boldsymbol{\eta}^T\Big(\kappa_v\mathbf{I} + \mathrm{diag}(c(\boldsymbol{z}))\Big)\boldsymbol{\eta} + \boldsymbol{\eta}^T(\kappa_u\boldsymbol{u} + b(\boldsymbol{z}))\right\},
\end{aligned}$$

where $\boldsymbol{z} = [z_1, ..., z_n]^T$ is the point around which we Taylor expand f, $b(\boldsymbol{z}) = [b(z_1), ..., b(z_n)]^T$ and $c(\boldsymbol{z}) = [c(z_1), ..., c(z_n)]^T$. q is the canonical form of a multivariate normal distribution, and can be used for Metropolis-Hastings steps for $\boldsymbol{\eta}$.

Exercise 2: Implementation of the MCMC sampler

Before we can implement the Metropolis-Hastings part of the sampler, we need to simplfy the expression for the acceptance propability α. If we first consider the ratio between true the full conditionals of $\boldsymbol{\eta}^*$ and $\boldsymbol{\eta}$ we can simply insert into the expression found in 1c)

$$\frac{\mathrm{p}(\boldsymbol{\eta}^*|\boldsymbol{y}, \boldsymbol{u}, \kappa_u, \kappa_v)}{\mathrm{p}(\boldsymbol{\eta}|\boldsymbol{y}, \boldsymbol{u}, \kappa_u, \kappa_v)} = \exp\left\{-\frac{\kappa_v}{2}\boldsymbol{\eta}^{*T}\boldsymbol{\eta}^* + \boldsymbol{\eta}^{*T}(\kappa_v\boldsymbol{u} + \boldsymbol{y}) - \exp(\boldsymbol{\eta}^*)^T\mathbf{E} + \frac{\kappa_v}{2}\boldsymbol{\eta}^T\boldsymbol{\eta} - \boldsymbol{\eta}^T(\kappa_v\boldsymbol{u} + \boldsymbol{y}) + \exp(\boldsymbol{\eta})^T\mathbf{E}\right\}.$$

Here $\boldsymbol{\eta}^*$ is the proposed m'th step, $\boldsymbol{\eta}$ the value of the $(m-1)$'th step while $\boldsymbol{u}$, κ_u and κ_v are the m'th step values. The ratio between the proposal distributions can also be found by insertion

$$
\begin{aligned}
\frac{\mathrm{q}(\boldsymbol{\eta}|\boldsymbol{\eta}^*,\boldsymbol{y},\boldsymbol{u},\kappa_u,\kappa_v)}{\mathrm{q}(\boldsymbol{\eta}^*|\boldsymbol{\eta},\boldsymbol{y},\boldsymbol{u},\kappa_u,\kappa_v)} =& \frac{\Big|\kappa_v \mathrm{I} + \mathrm{diag}(c(\boldsymbol{\eta}^*))\Big|^{\frac{1}{2}}}{\Big|\kappa_v \mathrm{I} + \mathrm{diag}(c(\boldsymbol{\eta}))\Big|^{\frac{1}{2}}} \cdot \\
& \exp\Big\{ -\frac{1}{2}\boldsymbol{\eta}^T \Big(\kappa_v \mathrm{I} + \mathrm{diag}(c(\boldsymbol{\eta}^*))\Big)\boldsymbol{\eta} + \boldsymbol{\eta}^T(\kappa_u\boldsymbol{u} + b(\boldsymbol{\eta}^*)) + \\
& \frac{1}{2}\boldsymbol{\eta}^{*T}\Big(\kappa_v \mathrm{I} + \mathrm{diag}(c(\boldsymbol{\eta}))\Big)\boldsymbol{\eta}^* - \boldsymbol{\eta}^{*T}(\kappa_u\boldsymbol{u} + b(\boldsymbol{\eta}))\Big\}.
\end{aligned}
$$

(dette er stygt, men jeg vet ikke hvordan det kan bli penere). Multiplying these two ratios and using the fact that $b(\boldsymbol{z}) = \boldsymbol{y} + \mathrm{diag}(c(\boldsymbol{z}))\boldsymbol{z} - c(\boldsymbol{z})$ and that $\Sigma c(\boldsymbol{z}) = \exp(\boldsymbol{z})^T\mathbf{E}$, we get

$$
\alpha = \min\left\{1, \frac{\prod_i(\kappa_v + c(\eta_i^*))}{\prod_i(\kappa_v + c(\eta_i))} \exp\left[c(\boldsymbol{\eta})^T\Big(\mathrm{diag}(\boldsymbol{\eta}^*)(\frac{1}{2}\boldsymbol{\eta}^* - \boldsymbol{\eta}) + \vec{1}\Big) - c(\boldsymbol{\eta}^*)^T\Big(\mathrm{diag}(\boldsymbol{\eta})(\frac{1}{2}\boldsymbol{\eta} - \boldsymbol{\eta}^*) + \vec{1}\Big)\right]\right\}
$$

```
c = function(z, E) {
    z * E
}

b = function(z, y, E) {
    y + c(z, E) * (z - 1)
}

drawKappaU = function(n, alpha_u, beta_u, u, R) {
    rgamma(1, (n - 1)/2 + alpha_u, beta_u + t(u) %*% R %*% u)
}

drawKappaV = function(n, alpha_v, beta_v, eta, u) {
    rgamma(1, n/2 + alpha_v, beta_v + t(eta - u) %*% (eta - u)/2)
}

drawU = function(n, kappa_u, kappa_v, eta, R) {
    rmvnorm.canonical(1, kappa_v/2 * eta, kappa_u * R + diag.spam(kappa_v,
        n, n))
}

drawEta = function(n, z, y, kappa_v, kappa_u, u, E) {
    rmvnorm.canonical(1, kappa_u * u + b(z, y, E), diag.spam(kappa_v,
        n, n))
}
```