linear and spline methods

Meilin Yan

June 10, 2016

Packages Loading

```
library(dlnm)
library(splines)
library(ggplot2)
library(ggthemes)
library(lubridate)
```

Data prep

```
bj <- read.csv("0812bj.csv")
bj$date <- as.Date(bj$date)
bj$year <- year(bj$date)
bj$month <- month(bj$date)
bj$month <- as.factor(bj$month)
bj$tot <- bj$A00toR99
bj$cir <- bj$I00toI99
bj$resp <- bj$J00toJ99
```

Holidays

Set up holiday variable.

Prediction and imputation

```
# Observed PM from Beijing (Non-embassy) monitor
bj$pm_bj <- bj$pm25hdbl

# Observed PM from Embassy monitor
bj$pm_us <- bj$pm25ussg

# predict pm_bj with pm_us
mod_bj <- lm(pm_bj ~ pm_us, na.action = na.exclude, data = bj)
bj$pre_bj <- predict(mod_bj, newdata = bj)
bj$pm_bj <- ifelse(is.na(bj$pm_bj), bj$pre_bj, bj$pm_bj)
# After replacing NA with predicted values, it has only 15 NA.
# The observed data from non-Embassy monitor has 156 NA.

# predict pm_us with pm_bj
```

```r
mod_us <- lm(pm_us ~ pm_bj, data = bj)
bj$pre_us <- predict(mod_us, newdata = bj)
bj$pm_us <- ifelse(is.na(bj$pm_us), bj$pre_us, bj$pm_us)
# After replacing NA with predicted values, it has only 15 NA.
# The observed data from US Embassy has 254 NA.

# Get the average value of PM2.5
bj$ave_pm <- (bj$pm_bj + bj$pm_us)/2

# Create lag01 PM2.5
bj$bj_pm01 <- filter(bj$pm25hdbl, c(1,1)/2, sides = 1)
bj$bj_pm01 <- as.numeric(bj$bj_pm01)
# bj$bj_pm01 <- round(bj$bj_pm01, 2)

bj$us_pm01 <- filter(bj$pm25ussg, c(1,1)/2, sides = 1)
bj$us_pm01 <- as.numeric(bj$us_pm01)
# bj$us_pm01 <- round(bj$us_pm01, 2)

bj$pm01 <- filter(bj$ave_pm, c(1,1)/2, sides = 1)
bj$pm01 <- as.numeric(bj$pm01)
summary(bj$pm01)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   13.86   49.93   76.33   85.25  108.60  365.00      28
```

```r
# 28 NA
```

ave_pm is the aveargе PM levels, pm01 is the lag01 PM levels.

Subset cold season and warm season

```r
bj.warm <- subset(bj, quarters(date) %in% c("Q2", "Q3"))
bj.cold <- subset(bj, quarters(date) %in% c("Q1", "Q4"))
```

Model Analysis

Functions

Crossbasis of linear PM at lag01

```r
cb.lin <- function(data = c()){
  cb <- crossbasis(data$ave_pm,
                   lag = 1,
                   argvar = list(fun = "lin",
                                 cen = FALSE),
                   arglag = list(fun = "strata",
                                 df = 1))
  return(cb)
}
```

Crossbasis of spline PM at lag01

```
cb.spl <- function(data = c(), knots = c()){
  cb <- crossbasis(data$ave_pm,
                   lag = 1,
                   argvar = list(fun = "ns",
                                 knots = knots,
                                 cen = 0),
                   arglag = list(fun = "strata",
                                 df = 1))
  return(cb)
}
```

GLM Model

```
mod <- function(data = c(), cause = c(), cb = c(), df=c()) {
  model <- glm(data[, cause] ~ cb + ns(tmean,3) +
                 splines::ns(rh, 3) + ns(time, 5*df) +
                 dow + holiday,
               family = quasipoisson(link = "log"),
               data = data,
               na.action = na.exclude,
               control = glm.control(epsilon = 10E-8,
                                     maxit = 10000))
  return(model)
}
```

Predict function

```
pred <- function(cb = c(), model = c(), at = c()) {
  pred.pm <- crosspred(cb, model, at = at, cumul = FALSE)
  return(pred.pm)
}
```

Two types of Analyses

1. **Knots at 75 and 150**

- Knots

```
a.knots <- c(75, 150)
w.knots <- c(75, 150)
c.knots <- c(75, 150)
```

- Calculate the 99th percentile for prediction

```
a.pat <- quantile(bj$pm01, probs = 0.99, na.rm = T)
w.pat <- quantile(bj.warm$pm01, probs = 0.99, na.rm = T)
c.pat <- quantile(bj.cold$pm01, probs = 0.99, na.rm = T)
```

- Crossbasis for year-round, warm and cold season

```
a.l.cb <- cb.lin(bj)
w.l.cb <- cb.lin(bj.warm)
c.l.cb <- cb.lin(bj.cold)

a.s.cb <- cb.spl(bj, a.knots)
w.s.cb <- cb.spl(bj.warm, w.knots)
c.s.cb <- cb.spl(bj.cold, c.knots)
```

- Total death
  - Linear model

```
library(splines)
a.l.a <- mod(bj, "tot", a.l.cb, 6)
w.l.a <- mod(bj.warm, "tot", w.l.cb, 3)
c.l.a <- mod(bj.cold, "tot", c.l.cb, 3)

# Predict
a.l.p.a <- pred(a.l.cb, a.l.a, at = 0:a.pat)
w.l.p.a <- pred(w.l.cb, w.l.a, at = 0:w.pat)
c.l.p.a <- pred(c.l.cb, c.l.a, at = 0:c.pat)
```

- Total death
  - Spline model

```
a.s.a <- mod(bj, "tot", a.s.cb, 6)
w.s.a <- mod(bj.warm, "tot", w.s.cb, 3)
c.s.a <- mod(bj.cold, "tot", c.s.cb, 3)

# Predict
a.s.p.a <- pred(a.s.cb, a.s.a, at = 0:a.pat)
w.s.p.a <- pred(w.s.cb, w.s.a, at = 0:w.pat)
c.s.p.a <- pred(c.s.cb, c.s.a, at = 0:c.pat)
```

- Plot total death

Set warning=FALSE because some values were clipped in the rug(...).

```
pdf("Plots/total_knots75_150.pdf", width = 6, height = 10)
par(mfrow=c(3,2))
# Year-round
plot(a.l.p.a, "overall", lag = 0, col = 1, ci = "lines",
     ylab = "RR of Total Death",
     xlab = expression(paste(PM[2.5], " concentration",
                             " (", mu,"g/", m^3, ")",
                             sep = " ")),
     main = "Year-round", ylim = c(0.9, 1.1), xlim = c(0, 250))
rug(bj$pm01, ticksize = 0.05, side = 1, lwd = 0.5)

plot(a.s.p.a, "overall", lag = 0, col = 1, ci = "lines",
```

```
     ylab = "RR of Total Death",
     xlab = expression(paste(PM[2.5], " concentration",
                        " (", mu,"g/", m^3, ")", sep = " ")),
     main = "Year-round", ylim = c(0.9, 1.1), xlim = c(0, 250))
rug(bj$pm01, ticksize = 0.05, side = 1, lwd = 0.5)

# Cold season
plot(c.l.p.a, "overall", lag = 0, col = 1, ci = "lines",
     ylab = "RR of Total Death",
     xlab = expression(paste(PM[2.5], " concentration",
                        " (", mu,"g/", m^3, ")", sep = " ")),
     main = "Cold Season", ylim = c(0.9, 1.1), xlim = c(0, 250))
rug(bj.cold$pm01, ticksize = 0.05, side = 1, lwd = 0.5)

plot(c.s.p.a, "overall", lag = 0, col = 1, ci = "lines",
     ylab = "RR of Total Death",
     xlab = expression(paste(PM[2.5], " concentration",
                        " (", mu,"g/", m^3, ")", sep = " ")),
     main = "Cold Season", ylim = c(0.9, 1.1), xlim = c(0, 250))
rug(bj.cold$pm01, ticksize = 0.05, side = 1, lwd = 0.5)

# Warm season
plot(w.l.p.a, "overall", lag = 0, col = 1, ci = "lines",
     ylab = "RR of Total Death",
     xlab = expression(paste(PM[2.5], " concentration",
                        " (", mu,"g/", m^3, ")", sep = " ")),
     main = "Warm Season", ylim = c(0.9, 1.1), xlim = c(0, 250))
rug(bj.warm$pm01, ticksize = 0.05, side = 1, lwd = 0.5)

plot(w.s.p.a, "overall", lag = 0, col = 1, ci = "lines",
     ylab = "RR of Total Death",
     xlab = expression(paste(PM[2.5], " concentration",
                        " (", mu,"g/", m^3, ")", sep = " ")),
     main = "Warm Season", ylim = c(0.9, 1.1), xlim = c(0, 250))
rug(bj.warm$pm01, ticksize = 0.05, side = 1, lwd = 0.5)
dev.off()
```

```
## pdf
##   2
```

- Circulatory death
  - Linear model

```
a.l.c <- mod(bj, "cir", a.l.cb, 6)
w.l.c <- mod(bj.warm, "cir", w.l.cb, 3)
c.l.c <- mod(bj.cold, "cir", c.l.cb, 3)

# Predict
a.l.p.c <- pred(a.l.cb, a.l.c, at = 0:a.pat)
w.l.p.c <- pred(w.l.cb, w.l.c, at = 0:w.pat)
c.l.p.c <- pred(c.l.cb, c.l.c, at = 0:c.pat)
```

- Circulatory death
  - Spline model

```
a.s.c <- mod(bj, "cir", a.s.cb, 6)
w.s.c <- mod(bj.warm, "cir", w.s.cb, 3)
c.s.c <- mod(bj.cold, "cir", c.s.cb, 3)

# Predict
a.s.p.c <- pred(a.s.cb, a.s.c, at = 0:a.pat)
w.s.p.c <- pred(w.s.cb, w.s.c, at = 0:w.pat)
c.s.p.c <- pred(c.s.cb, c.s.c, at = 0:c.pat)
```

- Plot circulatory death

```
pdf("Plots/cir_knots75_150.pdf", width = 6, height = 10)
par(mfrow=c(3,2))
# Year-round
plot(a.l.p.c, "overall", lag = 0, col = 1, ci = "lines",
     ylab = "RR of Circulatory Death",
     xlab = expression(paste(PM[2.5], " concentration",
                        " (", mu,"g/", m^3, ")",
                        sep = " ")),
     main = "Year-round", ylim = c(0.90, 1.15), xlim = c(0, 250))
rug(bj$pm01, ticksize = 0.05, side = 1, lwd = 0.5)

plot(a.s.p.c, "overall", lag = 0, col = 1, ci = "lines",
     ylab = "RR of Circulatory Death",
     xlab = expression(paste(PM[2.5], " concentration",
                        " (", mu,"g/", m^3, ")", sep = " ")),
     main = "Year-round", ylim = c(0.90, 1.15), xlim = c(0, 250))
rug(bj$pm01, ticksize = 0.05, side = 1, lwd = 0.5)

# Cold season
plot(c.l.p.c, "overall", lag = 0, col = 1, ci = "lines",
     ylab = "RR of Circulatory Death",
     xlab = expression(paste(PM[2.5], " concentration",
                        " (", mu,"g/", m^3, ")", sep = " ")),
     main = "Cold Season", ylim = c(0.95, 1.15), xlim = c(0, 250))
rug(bj.cold$pm01, ticksize = 0.05, side = 1, lwd = 0.5)

plot(c.s.p.c, "overall", lag = 0, col = 1, ci = "lines",
     ylab = "RR of Circulatory Death",
     xlab = expression(paste(PM[2.5], " concentration",
                        " (", mu,"g/", m^3, ")", sep = " ")),
     main = "Cold Season", ylim = c(0.95, 1.15), xlim = c(0, 250))
rug(bj.cold$pm01, ticksize = 0.05, side = 1, lwd = 0.5)

# Warm season
plot(w.l.p.c, "overall", lag = 0, col = 1, ci = "lines",
     ylab = "RR of Circulatory Death",
     xlab = expression(paste(PM[2.5], " concentration",
                        " (", mu,"g/", m^3, ")", sep = " ")),
```

```
     main = "Warm Season", ylim = c(0.90, 1.15), xlim = c(0, 250))
rug(bj.warm$pm01, ticksize = 0.05, side = 1, lwd = 0.5)

plot(w.s.p.c, "overall", lag = 0, col = 1, ci = "lines",
     ylab = "RR of Circulatory Death",
     xlab = expression(paste(PM[2.5], " concentration",
                         " (", mu,"g/", m^3, ")", sep = " ")),
     main = "Warm Season", ylim = c(0.90, 1.15), xlim = c(0, 250))
rug(bj.warm$pm01, ticksize = 0.05, side = 1, lwd = 0.5)
dev.off()
```

```
## pdf
##   2
```

- Respiratory death
  - Linear model

```
a.l.r <- mod(bj, "resp", a.l.cb, 6)
w.l.r <- mod(bj.warm, "resp", w.l.cb, 3)
c.l.r <- mod(bj.cold, "resp", c.l.cb, 3)

# Predict
a.l.p.r <- pred(a.l.cb, a.l.r, at = 0:a.pat)
w.l.p.r <- pred(w.l.cb, w.l.r, at = 0:w.pat)
c.l.p.r <- pred(c.l.cb, c.l.r, at = 0:c.pat)
```

- Respiratory death
  - Spline Model

```
a.s.r <- mod(bj, "resp", a.s.cb, 6)
w.s.r <- mod(bj.warm, "resp", w.s.cb, 3)
c.s.r <- mod(bj.cold, "resp", c.s.cb, 3)

# Predict Spline Model
a.s.p.r <- pred(a.s.cb, a.s.r, at = 0:a.pat)
w.s.p.r <- pred(w.s.cb, w.s.r, at = 0:w.pat)
c.s.p.r <- pred(c.s.cb, c.s.r, at = 0:c.pat)
```

- Plot respiratory death

```
pdf("Plots/resp_knots75_150.pdf", width = 6, height = 10)
par(mfrow=c(3,2))
# Year-round
plot(a.l.p.r, "overall", lag = 0, col = 1, ci = "lines",
     ylab = "RR of Respiratory Death",
     xlab = expression(paste(PM[2.5], " concentration",
                         " (", mu,"g/", m^3, ")",
                         sep = " ")),
     main = "Year-round", ylim = c(0.85, 1.30), xlim = c(0, 250))
```

```
rug(bj$pm01, ticksize = 0.05, side = 1, lwd = 0.5)

plot(a.s.p.r, "overall", lag = 0, col = 1, ci = "lines",
     ylab = "RR of Respiratory Death",
     xlab = expression(paste(PM[2.5], " concentration",
                         " (", mu,"g/", m^3, ")", sep = " ")),
     main = "Year-round", ylim = c(0.85, 1.30),
     xlim = c(0, 250))
rug(bj$pm01, ticksize = 0.05, side = 1, lwd = 0.5)
# Cold season
plot(c.l.p.r, "overall", lag = 0, col = 1, ci = "lines",
     ylab = "RR of Respiratory Death",
     xlab = expression(paste(PM[2.5], " concentration",
                         " (", mu,"g/", m^3, ")", sep = " ")),
     main = "Cold Season", ylim = c(0.85, 1.30), xlim = c(0, 250))
rug(bj.cold$pm01, ticksize = 0.05, side = 1, lwd = 0.5)

plot(c.s.p.r, "overall", lag = 0, col = 1, ci = "lines",
     ylab = "RR of Respiratory Death",
     xlab = expression(paste(PM[2.5], " concentration",
                         " (", mu,"g/", m^3, ")", sep = " ")),
     main = "Cold Season", ylim = c(0.85, 1.30), xlim = c(0, 250))
rug(bj.cold$pm01, ticksize = 0.05, side = 1, lwd = 0.5)

# Warm season
plot(w.l.p.r, "overall", lag = 0, col = 1, ci = "lines",
     ylab = "RR of Respiratory Death",
     xlab = expression(paste(PM[2.5], " concentration",
                         " (", mu,"g/", m^3, ")", sep = " ")),
     main = "Warm Season", ylim = c(0.85, 1.30),
     xlim = c(0, 250))
rug(bj.warm$pm01, ticksize = 0.05, side = 1, lwd = 0.5)

plot(w.s.p.r, "overall", lag = 0, col = 1, ci = "lines",
     ylab = "RR of Respiratory Death",
     xlab = expression(paste(PM[2.5], " concentration",
                         " (", mu,"g/", m^3, ")", sep = " ")),
     main = "Warm Season", ylim = c(0.85, 1.30),
     xlim = c(0, 250))
rug(bj.warm$pm01, ticksize = 0.05, side = 1, lwd = 0.5)
dev.off()
```

```
## pdf
##   2
```

2. Knots at the cutoff values in the subset method