

Expectile curves body weight of 100 boys, CV tuned (Dutch boys data)



Expectile smoothing of weights of Dutch boys, with 20 cubic B-spline segments. For each curve λ was tuned with asymmetric cross-validation. R code in `f-exp-curves-100-cv.R`

```
# Expectile curves for body weight of 100 boys tuned with CV
# Paul Eilers and Brian Marx 2019
# A graph in "Practical Smoothing. The Joys of P-splines"
```

```
library(colorspace)
library(AGD)
library(ggplot2)
library(JOPS)
```

```
# Get the data
data(boys7482)
Dat = subset(boys7482, age > 5, select = c(age, wgt))
Dat = na.omit(Dat)
m0 = nrow(Dat)
m = 100
set.seed(2018)
sel = sample(1:m0, m)
x = (Dat$age[sel])
y = Dat$wgt[sel]
```

```
# P-spline parameters
xl = 5
xr = 22
nsegx = 20
bdeg = 3
pp = c(0.05, 0.1, 0.2, 0.3, 0.5, 0.7, 0.8, 0.9, 0.95)
np = length(pp)
```

```
# Compute bases
B = bbase(x, xl, xr, nsegx, bdeg)
nbx = ncol(B)
xg = seq(xl, xr, length = 100)
Bg = bbase(xg, xl, xr, nsegx, bdeg)
```

```
lambdas = 10 ^ seq(-1, 4, by = 0.2)
```

```

# Compute penalty matrix
D = diff(diag(nbx), diff = 2)

# Fitting
Zg = NULL
CV = NULL
for (p in pp) {
  dzmax = 1e-3 * max(y)
  z = 0
  Z = NULL
  cvs = NULL
  for (lambda in lambdas) {
    P = lambda * t(D) %>% D
    for (it in 1:10) {
      r = y - z
      w = ifelse(r > 0, p, 1 - p)
      W = diag(c(w))
      Q = t(B) %>% W %>% B
      a = solve(Q + P, t(B) %>% (w * y))
      znew = B %>% a
      dz = sum(abs(Z - znew))
      if (dz < dzmax) break
      z = znew
    }
    H = solve(Q + P) %>% t(B) %>% W
    h = colSums(t(B) * H)
    h1 = diag(B %>% solve(Q + P, t(B) %>% W))
    rd = (y - z) / (1 - h)
    cv = sqrt(sum(w * rd ^2) / sum(w))
    cvs = c(cvs, cv)
    Z = cbind(Z, Bg %>% a)
  }
  CV = cbind(CV, cvs)
  k = which.min(cvs)
  Zg = cbind(Zg, Z[, k])
  cat(p, k, '\n')
}

cols = pal_rainbow(np)

# Dataframes for ggplot
DF1 = data.frame(x = x, y = y)
np = length(pp)
ng = length(xg)
DF2 = data.frame(x = rep(xg, np), y = as.vector(Zg),
                 p = as.factor(rep(pp, each = ng)))

# Make the graph
plt1 = ggplot(DF1, aes(x = x, y = y)) +
  geom_point(color = col_grey, size = 1.2) +
  xlab('Age (yr)') + ylab('Weight (kg)') +
  ggtitle('Expectile curves for weights of 100 Dutch boys') +
  geom_line(data = DF2, aes(x = x, y = y, group = p, color = p), size = 0.8) +
  JOPS_theme() +
  labs(color = bquote(tau)) +
  scale_color_manual(values = cols)

# Plot it and save it
plot(plt1)

```
