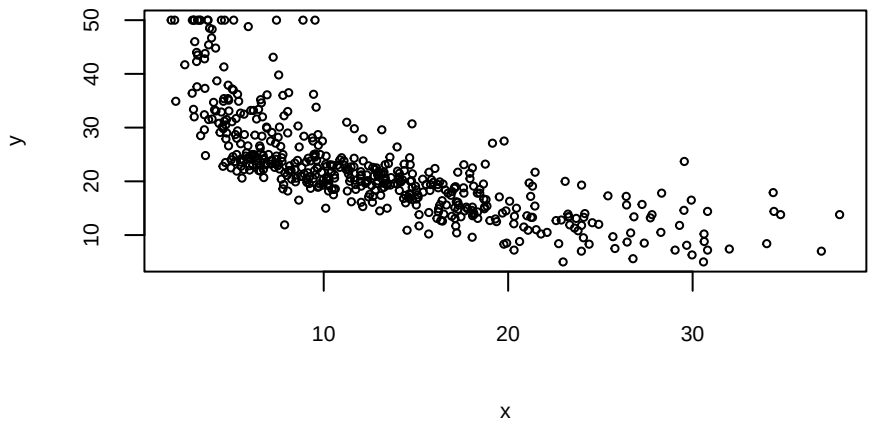
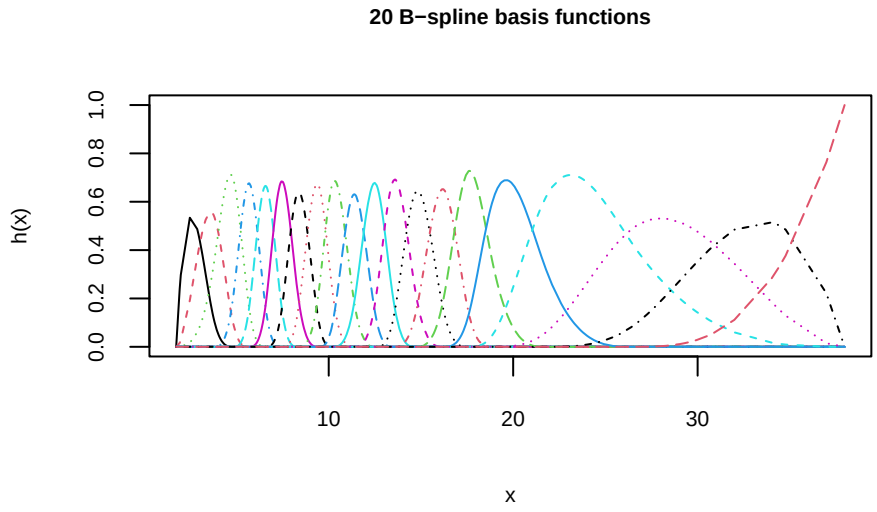
B-spline bases

```
library("MASS")
data(Boston)
plot(Boston$lstat, Boston$medv, xlab = "x", ylab = "y", cex = .5, cex.axis=.7,
     cex.lab = .7)
```




```
library("splines")
bs_x <- bs(Boston$lstat, df = 20)
matplot(Boston$lstat[order(Boston$lstat)], bs_x[order(Boston$lstat), ],
        type = "l", xlab = "x", ylab = "h(x)",
        main = "20 B-spline basis functions",
        cex.main = .7, cex.axis=.7, cex.lab = .7)
```



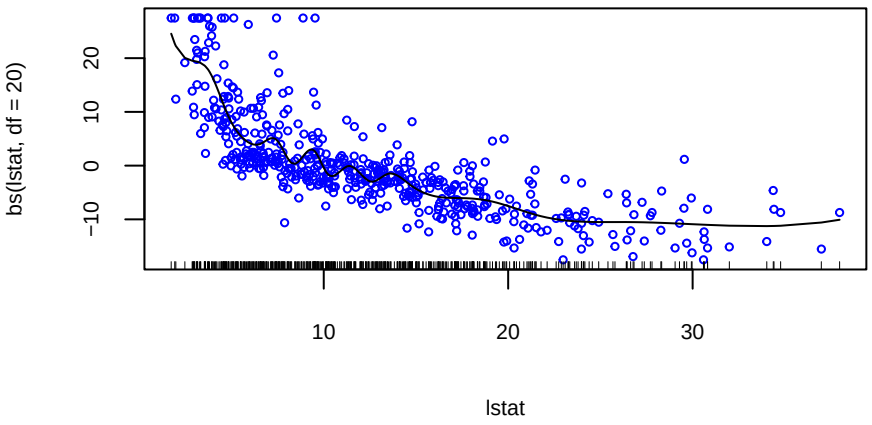
Note that by default, knots are placed based on quantiles (so more knots in areas with higher density).

Estimate coefficients (in an unpenalized manner) and inspect the fitted spline curve:

```
library("gam")
mod_df20 <- gam(medv ~ bs(lstat, df = 20), data = Boston)
coef(mod_df20)
```

```
##          (Intercept)  bs(lstat, df = 20)1  bs(lstat, df = 20)2
##            47.084153            -7.542724            -1.107027
##  bs(lstat, df = 20)3  bs(lstat, df = 20)4  bs(lstat, df = 20)5
##           -15.832894           -20.156822           -21.295745
##  bs(lstat, df = 20)6  bs(lstat, df = 20)7  bs(lstat, df = 20)8
##           -17.657272           -27.771313           -18.138844
##  bs(lstat, df = 20)9 bs(lstat, df = 20)10 bs(lstat, df = 20)11
##           -29.350675           -21.835623           -29.433491
## bs(lstat, df = 20)12 bs(lstat, df = 20)13 bs(lstat, df = 20)14
##           -24.612153           -28.927923           -30.660864
## bs(lstat, df = 20)15 bs(lstat, df = 20)16 bs(lstat, df = 20)17
##           -30.524051           -30.998140           -35.464397
## bs(lstat, df = 20)18 bs(lstat, df = 20)19 bs(lstat, df = 20)20
##           -34.369267           -36.984612           -34.627202
```

```
plot(mod_df20, residuals = TRUE, col = "blue", cex = .5,
     cex.axis=.7, cex.lab = .7)
```


bs(lstat, df = 20)
20
10
0
−10
10
20
30
lstat