

Tutorial 8

1. suppose we have sample $(\mathbf{x}_i, y_i), i = 1, \dots, n$. we estimate the regression model $y_i = g(\mathbf{x}_i) + \varepsilon_i$, where $g(x)$ is a spline function of the form

$$g(x) = \sum_{j=1}^{J+4} \beta_j B_j(x)$$

- (a) Estimate the derivative $g'(x)$ of $g(x)$
 - (b) find the 95% confidence band for $g'(x)$.
2. suppose we have sample $(\mathbf{x}_i, \mathbf{z}_i, y_i), i = 1, \dots, n$. we estimate the regression model $y_i = \beta_0 + \beta_1 \mathbf{z}_i + g(\mathbf{x}_i) + \varepsilon_i$ and $g(x)$ is a spline function of the form

$$g(x) = \sum_{j=1}^{J+4} \beta_j B_j(x)$$

- (a) write the expression for the estimator of $g(x)$
 - (b) find the 95% confidence band for $g(x)$.
3. In the first question, if \mathbf{x}_i just takes $m (< n)$ values. what is the maximum number of knots you can use to estimate the regression function g .
 4. Consider [data A](#) (the first column is Y and the others X). fit a “best” generalized additive model to the data and plot the estimated functions with 2 times standard error confidence bands. predict [data B](#).