STATS 305A : HW5

Problem 2

(a) We want to fit the model:

$$f(x, t) = \beta_0(t) + \beta_1(t)x + \varepsilon_t$$

where t is a value of the variable T, and x a value of the predictor X

In order to estimate (β_0, β_1) with respect to T, we will use a polynomial splines with a truncated power basis.

By choosing $M-4$ knots evenly spread over the values of T we would have the basis:

$$1, x, x^2, x^3, (x-k_1)_+^3, \dots, (x-k_{M-4})_+^3$$

and also we could define the basis matrix of splines in T: $\quad H \quad (n \times M \text{ matrix})$

s.t.:

$$\left\{ \begin{aligned} \underline{\beta_0} &= H \underline{\theta_0} \\ \underline{\beta_1} &= H \underline{\theta_1} \end{aligned} \right.$$

Finally our model prediction writes:

for all i, $i \in [\![1, n]\!]$:

$$\widehat{f(x_i, t_i)} = \underbrace{(H\hat{\underline{\theta}}_0)[i]}_{*} + \underbrace{(H\hat{\underline{\theta}}_1)[i]\, x_i}_{**}$$

$$(*) = \sum_{m=1}^{M} \hat{\theta}_{0m} h_m(t_i)$$

$$(**) : \sum_{m=1}^{M} \hat{\theta}_{1m} h_m(t_i)$$

Hence, there are $\boxed{2 \times M}$ parameters to estimate which are exactly:

$$\theta_{01}, \dots, \theta_{0M}, \theta_{11}, \dots, \theta_{1M}.$$

For $i \in [\![1, n]\!]$, as the model writes:

$$y_i = \sum_{m=1}^{n} \theta_{0m} h_m(t_i) + \sum_{m=1}^{p} \theta_{1m} h_m(t_i) x_i$$

we can see it as a regression model with covariates:

$$h_1(T), \dots, h_p(T), h_1(T)X, \dots, h_M(T)X$$

So by writing:

$$\tilde{X} = \left(h_1(T), \dots, h_p(T), X h_1(T), \dots, X h_p(T) \right)$$

<u>we can fit the linear regression:</u>

$$y = \tilde{X}\alpha + \varepsilon$$

where α is a vector of size $2n$.