Tutorial 6: suggested solutions

1. Suppose $m(x_1, ..., x_p) = \phi(\alpha_1 x_1 + ... + \alpha_p x_p)$. Prove that

$$\begin{pmatrix} \frac{\partial m(x_1,...,x_p)}{\partial x_1} \\ \frac{\partial m(x_1,...,x_p)}{\partial x_2} \\ \cdots \\ \frac{\partial m(x_1,...,x_p)}{\partial x_p} \end{pmatrix} = \phi'(\alpha_1 x_1 + ... + \alpha_p x_p) \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \cdots \\ \alpha_p \end{pmatrix}$$

proof: Because

$$\frac{\partial m(x_1, ..., x_p)}{\partial x_k} = \phi'(\alpha_1 x_1 + ... + \alpha_p x_p)\alpha_k$$

2. For a single-index model $Y = \phi(\alpha^\top X) + \varepsilon$. Suppose (X_i, y_i) are the observations and the estimator for α is $\hat{\alpha}$. For a new point X', predict the expectation of response using local linear kernel smoothing.

the prediction value is

$$\hat{g}(\hat{\alpha}^\top X') = \frac{\sum_{i=1}^n \{s_{n,2}(\hat{\alpha}^\top X')K_h(\hat{\alpha}^\top(X_i - X')) - s_{n,1}(\hat{\alpha}^\top X')K_h(\hat{\alpha}^\top(X_i - X'))\hat{\alpha}^\top(X_i - X')/h\}Y_i}{s_{n,2}(\hat{\alpha}^\top X')s_{n,0}(\hat{\alpha}^\top X') - s^2_{n,1}(\hat{\alpha}^\top X')}$$

where

$$s_{n,k}(\hat{\alpha}^\top X') \sum_{i=1}^n K_h(\hat{\alpha}^\top(X_i - X'))\{\hat{\alpha}^\top(X_i - X')/h\}^k, \quad k = 0, 1, 2$$

3. For the ozone concentration **data**.

   (a) fit a linear regression model and predict the ozone level when radiation=184.8, temperature=77.8, wind =9.9.

   (b) fit a partially linear regression model

$$ozone = \beta_1 * termperature + \beta_2 * Wind + g(Radiation) + \varepsilon.$$

   predict the ozone level when radiation=184.8, temperature=77.8, wind =9.9.

   (c) fit a single-index model and predict the ozone level when radiation=184.8, temperature=77.8, wind =9.9. and what is the prediction confidence interval for the expectation of Y.

(a) The fitted regression model is

$$\hat{Y} = -64.23208116 + 0.05979717 * rdaition + 1.65120780 * temperature - 3.33759763 * wind$$

The predicted value is: 36.26047 (**(cT51a.R)**)

(b) The fitted model is

$$\hat{Y} = 1.590852 * temperature - 3.268622 * wind + \hat{g}(radiation)$$

It is calculated that

$$\hat{g}(184.8) = -46.67959$$

Thus the predicted value is

$$1.590852 * 77.8 - 3.268622 * 9.9 - 46.67959 = 44.72934$$

(**(cT51b.R)**)

(c) The fiteed model is

$$\hat{y} = \hat{\phi}(0.01611658 * radiation + 0.44353068 * temperature - 0.89611427 * wind)$$

The predicted value is

$$\hat{\phi}(28.6135) = 32.95738$$

(**(cT51c.R)**)

4. For **data A** (the first two columns are independent variables, the last one response variable), fit a linear regression model; check whether the model is adequate. If not, propose a new parametric model.

   See code **(t5ca.R)**

5. For **data B**, fit a single-index model

$$Y = g(\beta_1 X_1 + \ldots + \beta_6 X_6) + \varepsilon$$

   and check which covariates can be removed