

Part IV: Problem Set 2

Due on May 4 Monday

Use the cross-sectional data on happiness, age and income provided in the Canvas to answer the following questions. Try the two bandwidth parameters chosen by the rule-of-thumb method and the cross-validation approach and the four kernel functions used for Problem Set 1. Obtain your answers using your own code written in Matlab. You are encouraged to work with your classmates, but I expect that you write your solutions independently.

1. Let y_i and x_i , $i = 1, \dots, n$, denote iid observations on happiness and income. Define $m(x_i) = \mathbb{E}(y_i|x_i)$, and consider the nonparametric regression model

$$y_i = m(x_i) + u_i$$

where u_i is the regression error satisfying $\mathbb{E}(u_i|x_i) = 0$. Answer the following:

- (a) Obtain the local constant estimator $\hat{m}(x)$ and the local linear estimator $\tilde{m}(x)$ of the regression function $m(x)$. Discuss the results obtained from $\hat{m}(x)$ and $\tilde{m}(x)$.
- (b) Compare the results based on the nonparametric regression function estimates $\hat{m}(x)$ and $\tilde{m}(x)$ with the results obtained from estimating the linear model $y_i = \alpha + \beta x_i + \varepsilon_i$.
- (c) Finally, compare the results in parts (a) and (b) with those obtained from the nonparametric regression of happiness on age in Problem Set 1. Be explicit about the similarities and differences.

2. Now let z_i denote age, and assume (y_i, x_i, z_i) , $i = 1, \dots, n$, are iid observations. Consider the functional coefficient model

$$y_i = \alpha(z_i) + x_i\beta(z_i) + u_i$$

where $\alpha(z)$ and $\beta(z)$ are unspecified smooth functions.

- (a) Obtain the local constant estimators, $\hat{\alpha}(z)$ and $\hat{\beta}(z)$, and the local linear estimators, $\tilde{\alpha}(z)$ and $\tilde{\beta}(z)$, of the unknown functional coefficients $\alpha(z)$ and $\beta(z)$. Discuss the results.
- (b) To allow for possible nonlinear relationship between income and happiness, add the square of income as an additional regressor to capture this nonlinearity as

$$y_i = \alpha(z_i) + x_i\beta_1(z_i) + x_i^2\beta_2(z_i) + u_i$$

and estimate the functional coefficients $\alpha(z)$, $\beta_1(z)$ and $\beta_2(z)$ by both the local constant and local linear estimators. Discuss the results. How these results differ from those obtained in part (a).

3. Split the sample in the following ways, and do Questions 1 and 2 for each of the subsamples.

- (a) Male and female subsamples.
- (b) Employed and unemployed subsamples.
- (c) With college education and without college education subsamples.