Tutorial 5

1. comparing the estimators of β on the top of page 2 (chapter 2 part 1) and (2.3) on page 3 of part 2. Explain why there is no weight function in the estimator of page 2.

2. Based on the notation in Lecture note (Chapter 2, part 2), eqns between (2.2) and (2.3), prove (2.3)

3. Consider the air-pollution and health **data** in Hong Kong. Suppose we consider the effect of pollutants (NO_2, SO_2, O_3, Particulate matters (PM)) on the number of hospital admission suffering respiratory diseases Y. We fit each pollutant a linear regression model to find the dependence. For example

$$Y = a + b * NO_2 + \varepsilon. \tag{1}$$

If we want to see how the weather conditions (temperature or humidity) change the dependence between Y and the pollutants. We may further consider a varying coefficient regression model. For example

$$Y = a(Temperature) + b(Temperature) * NO_2 + \varepsilon. \tag{2}$$

estimate model (1) and (2) (for all different combinations of pollutants and weather conditions).

   (a) What is your conclusion based on the estimated models?

   (b) For temperature = 20, NO_2 = 80, predict Y based on models (1) and (2) respectively.

4. Write a code (function) for the least squares estimation of a linear regression model. In the code, the input data are X and Y. The output includes the estimate of the coefficient parameters, R^2 and estimate of the variance of regression error.

5. Suppose there is a model

$$Y = a_0 + a_1(Z)X + \varepsilon$$

where a_0 is an unknown parameter and $a_1(.)$ is an unknown function. Suppose $(X_i, Y_i, Z_i), i = 1, ..., n$ are observations. Propose a method to estimate the model, estimate a_0 and function $a_1(z)$.