

---

## 8.5.4 Extensions and Issues

In practice, the models that we use might be more complicated than the simple linear regression model. We almost always use computers to do regression analysis (MATLAB, EXCEL, etc). Therefore, when we understand the basics of the simple linear regression, we can simply use computers to do more complicated analyses. A natural extension to the simple linear regression is *multiple linear regression*.

### Multiple Linear Regression:

In the above discussion, our model had only one predictor (explanatory variable),  $x$ . We can consider models with more than one explanatory variable. For example, suppose that you would like to have a model to predict house prices based on square footage, age, number of bedrooms, etc. Here, your response variable  $y$  is the house price. Your goal is to have a linear model

$$y = \beta_0 + \beta_1 x + \beta_2 z + \cdots + \beta_k w + \epsilon,$$

where  $x, z, \dots, w$  are the explanatory variables (square footage, age, number of bedrooms, etc). Such a model is an example of a **multiple linear regression** model. It is possible to extend the method of least squares to this case to compute estimates of  $\beta_0, \beta_1, \dots, \beta_k$ . In MATLAB, the command `regress` can be used for multiple linear regression.

It is worth noting that when we say linear regression, we mean linear in the unknown parameters  $\beta_i$ . For example, the model

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \epsilon$$

is a linear regression model since it is linear in  $\beta_0, \beta_1$ , and  $\beta_2$ . We would like to end this section by mentioning that when running regression algorithms, one needs to be mindful about some practical considerations. Issues such as overfitting [\[21\]](#), heteroscedasticity [\[22\]](#), and multicollinearity [\[23\]](#) might cause problems in regression analysis.