Lec - 15: Standard Line Fitting (Least-Squares)

Prof. K. Madhava Krishna

November 12, 2015 (Thursday)

This lecture will be on the standard *linear least squares* problem, i.e., that of fitting a straight line through a given set of data points.

1 (Linear) Least-Squares

First, we shall briefly look into the problem formulation. Then, we will look at determining the solution to the problem.

1.1 Problem Formulation

Consider a set of points, for now in 2D, Z.

$$Z = \{(x_1, y_1,), ..., (x_n, y_n)\} \tag{1}$$

We are interested in finding the straight line that *fits* the data in Z in the *best* manner. This kind of problem falls under the more general catogroy of *model prediction* problems, where in, given a mathematical model, we need to determine the best parameters for the model that satisfy specified constraints. In our case, the model is that of a straight line, i.e.,

$$\hat{y}_1 = c_0 + c_1 x_i \tag{2}$$

More formally, we need to find values for variables c_0^*, c_1^* such that the *error*, which is to say, the sum of the squared Euclidean distances from each of the points to the line is minimized.

$$c_0^*, c_1^* = \underset{c_0, c_1}{\operatorname{argmin}} \sum_{i=1}^{n} \|z_i - \hat{z}_i\|^2 \tag{3}$$

$$c_0^*, c_1^* = \underset{c_0, c_1}{\operatorname{argmin}} \sum_{i=1}^{n} \|c_0 + c_1 x_i - y_i\|^2 \tag{4}$$

We have

$$\sum_{i=1}^{n} \|c_0 + c_1 x_i - y_i\|^2 = \|c_0 + c_1 x_i - y_i\|^T \|c_0 + c_1 x_i - y_i\| \tag{5}$$

$$\sum_{i=1}^{n} \|c_0 + c_1 x_i - y_i\|^2 = \left(\begin{bmatrix} c_0 & c_1 \end{bmatrix} \begin{bmatrix} 1 & 1 & .... & 1 \\ x_0 & x_1 & .... & x_n \end{bmatrix} - \begin{bmatrix} y_0 & y_1 & .... & y_n \end{bmatrix} \right)$$
$$\left(\begin{bmatrix} 1 & x_0 \\ 1 & x_1 \\ . & . \\ . & . \\ 1 & x_n \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} \begin{bmatrix} y_0 \\ y_1 \\ . \\ . \\ y_n \end{bmatrix} \right)$$

This equation can compactly be written as

$$f = [Au - b]^T [Au - b] = \|Au - b\|^2 \tag{6}$$

where

$$u_{2\times 1} = \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} \tag{7}$$

$$A_{n\times 1} = \begin{bmatrix} 1 & x_0 \\ 1 & x_1 \\ . & . \\ . & . \\ 1 & x_n \end{bmatrix} \tag{8}$$

$$b_{n\times 1} = \begin{bmatrix} y_0 \\ y_1 \\ . \\ . \\ y_n \end{bmatrix} \tag{9}$$

1.2 Finding the Best Line

Let J be the error function we had to minimize. Then,

$$J = [Au - b]^T [Au - b] = [u^T A^T - b^T][Au - b] = u^T A^T Au - u^T A^T u - b^T Au + b^T b \tag{10}$$

We now solve the problem by minimizing J, i.e., by computing the first derivative of J with respect to u and setting it to zero.

$$\frac{dJ}{du} = 2u^T A^T A - 2b^T A = 0 \tag{11}$$

$$u^T A^T A = b^T A \tag{12}$$

$$u^T = b^T A [A^T A]^{-1} \tag{13}$$

We can easily obtain u by transposing the above equation. The above equation is the *right pseudo-inverse formulation*, since the pseudo-inverse of A, i.e., $A^T A$, appears on the right side of the equation. Transposing this equation leads to the left pseudo-inverse formulation.

$$u = [A^T A]^{-1} A^T b \tag{14}$$