1 Założenia i różne oznaczenia

Mamy macierz **Data**, taką że

$$\mathbf{Data} = \begin{bmatrix} x_1(1) & x_2(1) & \dots & x_p(1) \\ x_1(2) & x_2(2) & \dots & x_p(2) \\ \vdots & \vdots & \dots & \vdots \\ x_1(n) & x_2(n) & \dots & x_p(n) \end{bmatrix} \tag{1}$$

gdzie zapis $x_i(j)$ oznacza j-tą obserwację i-tej cechy.

Oraz odpowiadająca jej macierz tar, zawierającą obserwację zmiennej pożądanej.

$$\mathbf{tar} = \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(n) \end{bmatrix} \tag{2}$$

Chcemy uzależnić zmienna tar od zmiennych $x_1, x_2 \dots, x_p$ w sposób liniowy, to znaczy dla dowolnej k-tej obserwacji: $k \in 1, 2, \dots, n$ wartość y(i) musi być równa:

$$y(\hat{k}) = \beta_0 + \sum_{i=1}^{p} \beta_i x_i(k) \tag{3}$$

lub powinno zachodzić macierzowe równanie:

$$t\hat{a}r = X * \beta \tag{4}$$

gdzie:

$$X = \begin{bmatrix} 1 & x_1(1) & x_2(1) & \dots & x_p(1) \\ 1 & x_1(2) & x_2(2) & \dots & x_p(2) \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & x_1(n) & x_2(n) & \dots & x_p(n) \end{bmatrix} \tag{5}$$

oraz

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \beta_p \end{bmatrix} \tag{6}$$

W celu znalezienia wektora , β, minimalizującego model, zdefiniujmy funkcję błędu:

$$Error = t\hat{a}r((\beta) - tar)^T * (t\hat{a}r(\beta) - tar) \tag{7}$$

Przekonajmy się, że równanie (7) daje nam w wyniku macierz kwadratową stopnia 1, która można utozszamić z jedną liczbą rzeczywistą.

Okazuje się, że parametr $\beta = (X^T X)^{-1} X^T Y$ minimalizuje funkcję (7)