Derivation of Linear Regression Loss Function Minimization

The loss function for linear regression with n data points is given by:

$$J(m,c) = \sum_{i=1}^{n}(y_i - mx_i - c)^2$$

Quadratic in m (treating c as fixed)

To express $J(m,c)$ as a quadratic in m, expand the squared term for each i:

$$(y_i - mx_i - c)^2 = (y_i - c - mx_i)^2 = (y_i - c)^2 - 2mx_i(y_i - c) + m^2x_i^2$$

Sum over all i from 1 to n:

$$\sum_{i=1}^{n}(y_i - mx_i - c)^2 = \sum_{i=1}^{n}\left[(y_i - c)^2 - 2mx_i(y_i - c) + m^2x_i^2\right]$$
$$= \sum_{i=1}^{n}(y_i - c)^2 - 2m\sum_{i=1}^{n}x_i(y_i - c) + m^2\sum_{i=1}^{n}x_i^2$$

Expand the middle term:

$$\sum_{i=1}^{n}x_i(y_i - c) = \sum_{i=1}^{n}x_iy_i - c\sum_{i=1}^{n}x_i$$

Thus,

$$J(m,c) = \sum_{i=1}^{n}(y_i - c)^2 - 2m\left(\sum_{i=1}^{n}x_iy_i - c\sum_{i=1}^{n}x_i\right) + m^2\sum_{i=1}^{n}x_i^2$$
$$= \sum_{i=1}^{n}(y_i - c)^2 - 2m\sum_{i=1}^{n}x_iy_i + 2mc\sum_{i=1}^{n}x_i + m^2\sum_{i=1}^{n}x_i^2$$

This is a quadratic in m of the form $am^2 + bm + d$, where:

$$a = \sum_{i=1}^{n}x_i^2, \quad b = -2\sum_{i=1}^{n}x_iy_i + 2c\sum_{i=1}^{n}x_i, \quad d = \sum_{i=1}^{n}(y_i - c)^2$$

The value of m that minimizes J with respect to m occurs at $-\frac{b}{2a}$:

$$m = -\frac{b}{2a} = -\frac{-2\sum_{i=1}^{n}x_iy_i + 2c\sum_{i=1}^{n}x_i}{2\sum_{i=1}^{n}x_i^2} = \frac{2\sum_{i=1}^{n}x_iy_i - 2c\sum_{i=1}^{n}x_i}{2\sum_{i=1}^{n}x_i^2} = \frac{\sum_{i=1}^{n}x_iy_i - c\sum_{i=1}^{n}x_i}{\sum_{i=1}^{n}x_i^2}$$

Quadratic in c (treating m as fixed)

To express $J(m,c)$ as a quadratic in c, expand the squared term:

$$(y_i - mx_i - c)^2 = (y_i - mx_i)^2 - 2c(y_i - mx_i) + c^2$$

Sum over all i:

$$\sum_{i=1}^{n}(y_i - mx_i - c)^2 = \sum_{i=1}^{n}(y_i - mx_i)^2 - 2c\sum_{i=1}^{n}(y_i - mx_i) + \sum_{i=1}^{n} c^2$$

Since c^2 is constant, $\sum_{i=1}^{n} c^2 = nc^2$. Expand the middle term:

$$\sum_{i=1}^{n}(y_i - mx_i) = \sum_{i=1}^{n} y_i - m\sum_{i=1}^{n} x_i$$

Thus,

$$\begin{aligned} J(m,c) &= \sum_{i=1}^{n}(y_i - mx_i)^2 - 2c\left(\sum_{i=1}^{n} y_i - m\sum_{i=1}^{n} x_i\right) + nc^2 \\ &= \sum_{i=1}^{n}(y_i - mx_i)^2 - 2c\sum_{i=1}^{n} y_i + 2cm\sum_{i=1}^{n} x_i + nc^2 \end{aligned}$$

This is a quadratic in c of the form $ac^2 + bc + d$, where:

$$a = n, \quad b = -2\sum_{i=1}^{n} y_i + 2m\sum_{i=1}^{n} x_i, \quad d = \sum_{i=1}^{n}(y_i - mx_i)^2$$

The value of c that minimizes J with respect to c occurs at $-\frac{b}{2a}$:

$$c = -\frac{b}{2a} = -\frac{-2\sum_{i=1}^{n} y_i + 2m\sum_{i=1}^{n} x_i}{2n} = \frac{2\sum_{i=1}^{n} y_i - 2m\sum_{i=1}^{n} x_i}{2n} = \frac{\sum_{i=1}^{n} y_i - m\sum_{i=1}^{n} x_i}{n}$$

Solving for m and c by Substitution

To find the values of m and c that jointly minimize J, use the expressions derived:

$$c = \frac{\sum_{i=1}^{n} y_i - m\sum_{i=1}^{n} x_i}{n} = \frac{1}{n}\sum_{i=1}^{n} y_i - \frac{m}{n}\sum_{i=1}^{n} x_i$$

$$m = \frac{\sum_{i=1}^{n} x_i y_i - c\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i^2}$$

Substitute the expression for c into the expression for m:

$$m = \frac{\sum_{i=1}^{n} x_i y_i - \left(\frac{1}{n}\sum_{i=1}^{n} y_i - \frac{m}{n}\sum_{i=1}^{n} x_i\right)\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i^2}$$

Expand the numerator:

$$\sum_{i=1}^{n} x_i y_i - \left(\frac{1}{n}\sum_{i=1}^{n} y_i\right)\left(\sum_{i=1}^{n} x_i\right) + \left(\frac{m}{n}\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} x_i\right)$$

$$= \sum_{i=1}^{n} x_i y_i - \frac{1}{n} \left(\sum_{i=1}^{n} y_i \right) \left(\sum_{i=1}^{n} x_i \right) + m \cdot \frac{1}{n} \left(\sum_{i=1}^{n} x_i \right) \left(\sum_{i=1}^{n} x_i \right)$$

So,

$$m = \frac{\sum_{i=1}^{n} x_i y_i - \frac{1}{n} \left(\sum_{i=1}^{n} y_i \right) \left(\sum_{i=1}^{n} x_i \right) + m \cdot \frac{1}{n} \left(\sum_{i=1}^{n} x_i \right) \left(\sum_{i=1}^{n} x_i \right)}{\sum_{i=1}^{n} x_i^2}$$

Multiply both sides by $\sum_{i=1}^{n} x_i^2$:

$$m \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i - \frac{1}{n} \left(\sum_{i=1}^{n} y_i \right) \left(\sum_{i=1}^{n} x_i \right) + m \cdot \frac{1}{n} \left(\sum_{i=1}^{n} x_i \right) \left(\sum_{i=1}^{n} x_i \right)$$

Move terms involving m to the left:

$$m \sum_{i=1}^{n} x_i^2 - m \cdot \frac{1}{n} \left(\sum_{i=1}^{n} x_i \right) \left(\sum_{i=1}^{n} x_i \right) = \sum_{i=1}^{n} x_i y_i - \frac{1}{n} \left(\sum_{i=1}^{n} y_i \right) \left(\sum_{i=1}^{n} x_i \right)$$

Solve for m:

$$m = \frac{\sum_{i=1}^{n} x_i y_i - \frac{1}{n} \left(\sum_{i=1}^{n} y_i \right) \left(\sum_{i=1}^{n} x_i \right)}{\sum_{i=1}^{n} x_i^2 - \frac{1}{n} \left(\sum_{i=1}^{n} x_i \right) \left(\sum_{i=1}^{n} x_i \right)}$$

The final formula for c in terms of m is:

$$c = \frac{1}{n} \sum_{i=1}^{n} y_i - \frac{m}{n} \sum_{i=1}^{n} x_i$$