Equations for Gradient Descent in Linear Regression

by
Kalyan Cheerla

Input:

$$\hat{y}_i = \vec{w}\vec{x} + b = w_1x_1 + w_2x_2 + \cdots + w_nx_n + b \tag{1}$$

$$J(\vec{w}, b) = \frac{1}{2n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2 \tag{2}$$

$$\begin{aligned} w_i &= w_i - \frac{\partial J(\vec{w}, b)}{\partial w_i} \\ b &= b - \frac{\partial J(\vec{w}, b)}{\partial b} \end{aligned} \tag{3}$$

Solution:

Lets try to solve for $\frac{\partial J(\vec{w},b)}{\partial w_i}$,

$$\begin{aligned} \frac{\partial J(\vec{w}, b)}{\partial w_i} &= \frac{\partial}{\partial w_i}\left(\frac{1}{2n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2\right) \\ &= \frac{1}{2n}\sum_{i=1}^{n}\left(\frac{\partial}{\partial w_i}(\hat{y}_i - y_i)^2\right) \end{aligned}$$

$$\begin{aligned} \because\ \frac{dy^2}{dx} &= 2y * \frac{dy}{dx}\ \& \\ \frac{dk}{dx} &= 0 \text{ where k is constant.} \end{aligned} \tag{4}$$

$$\begin{aligned} &= \frac{1}{2n}\sum_{i=1}^{n}\left(2(\hat{y}_i - y_i) * \frac{\partial(\hat{y}_i - y_i)}{\partial w_i}\right) \\ &= \frac{1}{n}\sum_{i=1}^{n}\left((\hat{y}_i - y_i) * \frac{\partial \hat{y}_i}{\partial w_i}\right) \end{aligned}$$

Now substitute the eq (1),

$$\begin{aligned} &= \frac{1}{n}\sum_{i=1}^{n}\left((\hat{y}_i - y_i) * \frac{\partial(\vec{w}\vec{x} + b)}{\partial w_i}\right) \\ &= \frac{1}{n}\sum_{i=1}^{n}\left((\hat{y}_i - y_i) * \frac{\partial(w_1x_1 + w_2x_2 + \cdots + w_nx_n + b)}{\partial w_i}\right) \end{aligned}$$

Based on the eq (4),

$$
\begin{aligned}
&= \frac{1}{n}\sum_{i=1}^{n}\left((\hat{y}_i - y_i) * \left(0 + 0 + \cdots + \frac{\partial(w_i x_n)}{\partial w_i} + \cdots + 0 + 0\right)\right) \\
&= \frac{1}{n}\sum_{i=1}^{n}[(\hat{y}_i - y_i) * x_i]
\end{aligned}
$$

Finally the equations are,

$$
\begin{aligned}
w_i &= w_i - \frac{1}{n}\sum_{i=1}^{n}[(\hat{y}_i - y_i) * x_i] \qquad (5) \\
&= w_i - \frac{1}{n}\sum_{i=1}^{n}[(\vec{w}\vec{x} + b - y_i) * x_i] \\
b &= b - \frac{1}{n}\sum_{i=1}^{n}[(\hat{y}_i - y_i) * 1]
\end{aligned}
$$

Equations for Gradient Descent in Logistic Regression

by
Kalyan Cheerla

Input:

$$\begin{aligned} \hat{y}_i &= \frac{1}{1+e^{-z}} = \frac{1}{1+e^{-(\vec{w}\vec{x}+b)}} \\ z &= \vec{w}\vec{x}+b = w_1x_1 + w_2x_2 + \cdots + w_nx_n + b \end{aligned} \tag{1}$$

$$J(\vec{w}, b) = -\frac{1}{n}\sum_{i=1}^{n}\left[y_i \log\left(\hat{y}_i\right) + (1-y_i)\log\left(1-\hat{y}_i\right)\right] \tag{2}$$

$$\begin{aligned} w_i &= w_i - \frac{\partial J(\vec{w}, b)}{\partial w_i} \\ b &= b - \frac{\partial J(\vec{w}, b)}{\partial b} \end{aligned} \tag{3}$$

Solution:

Lets try to solve for $\frac{\partial J(\vec{w},b)}{\partial w_i}$,

$$\begin{aligned} \frac{\partial J(\vec{w}, b)}{\partial w_i} &= \frac{\partial}{\partial w_i}\left(-\frac{1}{n}\sum_{i=1}^{n}[y_i \log\left(\hat{y}_i\right) + (1-y_i)\log\left(1-\hat{y}_i\right)]\right) \\ &= -\frac{1}{n}\sum_{i=1}^{n}\left(\frac{\partial}{\partial w_i}[y_i \log\left(\hat{y}_i\right) + (1-y_i)\log\left(1-\hat{y}_i\right)]\right) \\ &= -\frac{1}{n}\sum_{i=1}^{n}\left(y_i\frac{\partial}{\partial w_i}(\log\left(\hat{y}_i\right)) + (1-y_i)\frac{\partial}{\partial w_i}(\log\left(1-\hat{y}_i\right))\right) \end{aligned}$$

$$\begin{aligned} \because\ \frac{d(\log_a y)}{dx} &= \frac{1}{y\log_e a} * \frac{dy}{dx}\ \& \\ \frac{dk}{dx} &= 0 \text{ where k is constant.} \end{aligned} \tag{4}$$

Here if we consider logarithmic with base e,

$$\begin{aligned} &= -\frac{1}{n}\sum_{i=1}^{n}\left(\frac{y_i}{\hat{y}_i} * \frac{\partial \hat{y}_i}{\partial w_i} + \frac{1-y_i}{1-\hat{y}_i} * \frac{\partial(1-\hat{y}_i)}{\partial w_i}\right) \\ &= -\frac{1}{n}\sum_{i=1}^{n}\left(\frac{y_i}{\hat{y}_i} - \frac{1-y_i}{1-\hat{y}_i}\right) * \frac{\partial \hat{y}_i}{\partial w_i} \\ &= -\frac{1}{n}\sum_{i=1}^{n}\left(\frac{y_i - y_i\hat{y}_i - \hat{y}_i + y_i\hat{y}_i}{\hat{y}_i(1-\hat{y}_i)}\right) * \frac{\partial \hat{y}_i}{\partial w_i} \\ &= \frac{1}{n}\sum_{i=1}^{n}\left(\frac{\hat{y}_i - y_i}{\hat{y}_i(1-\hat{y}_i)}\right) * \frac{\partial \hat{y}_i}{\partial w_i} \end{aligned} \tag{5}$$

Lets try to reduce $\frac{\partial \hat{y}_i}{\partial w_i}$ by substituting eq (1),

$$\begin{aligned}
\frac{\partial \hat{y}_i}{\partial w_i} &= \frac{\partial}{\partial w_i}\left(\frac{1}{1+e^{-z}}\right) \\
&= \frac{\partial}{\partial w_i}\left(1+e^{-z}\right)^{-1} \\
&\because \ \frac{d(y^k)}{dx} = k * y^{(k-1)} * \frac{dy}{dx} \ \& \\
&\quad\ \frac{d(e^y)}{dx} = e^y * \frac{dy}{dx} \\
&= -1 * (1+e^{-z})^{-2} * \frac{\partial}{\partial w_i}\left(1+e^{-z}\right) \\
&= -\frac{1}{(1+e^{-z})^2} * \left(0 + \frac{\partial (e^{-z})}{\partial w_i}\right) \\
&= -\frac{1}{(1+e^{-z})^2} * e^{-z} * \frac{\partial (-z)}{\partial w_i} \\
&= \frac{e^{-z}}{(1+e^{-z})^2} * \frac{\partial z}{\partial w_i} \\
&= \frac{e^{-z}}{(1+e^{-z})^2} * \frac{\partial (w_1 x_1 + w_2 x_2 + \cdots + w_i x_i + \cdots + w_n x_n)}{\partial w_i} \\
&= \frac{e^{-z}}{(1+e^{-z})^2} * x_i
\end{aligned} \tag{6}$$

Lets try to reduce $\hat{y}_i(1-\hat{y}_i)$ by substituting eq (1),

$$\begin{aligned}
\hat{y}_i(1-\hat{y}_i) &= \frac{1}{1+e^{-z}}\left(1 - \frac{1}{1+e^{-z}}\right) \\
&= \frac{1}{1+e^{-z}}\left(\frac{1+e^{-z}-1}{1+e^{-z}}\right) \\
&= \frac{1}{1+e^{-z}}\left(\frac{e^{-z}}{1+e^{-z}}\right) \\
&= \frac{e^{-z}}{(1+e^{-z})^2} \\
\frac{1}{\hat{y}_i(1-\hat{y}_i)} &= \frac{(1+e^{-z})^2}{e^{-z}}
\end{aligned}$$

Now if we continue on our eq (5),

$$\begin{aligned}
\frac{\partial J(\vec{w}, b)}{\partial w_i} &= \frac{1}{n}\sum_{i=1}^{n}\left(\frac{\hat{y}_i - y_i}{\hat{y}_i(1-\hat{y}_i)}\right) * \frac{\partial \hat{y}_i}{\partial w_i} \\
&= \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i) * \left(\frac{(1+e^{-z})^2}{e^{-z}} * \frac{x_i e^{-z}}{(1+e^{-z})^2}\right) \\
&= \frac{1}{n}\sum_{i=1}^{n}[(\hat{y}_i - y_i) * x_i]
\end{aligned}$$

Finally the equations are,

$$
\begin{aligned}
w_i &= w_i - \frac{1}{n} \sum_{i=1}^{n} [(\hat{y}_i - y_i) * x_i] \\
&= w_i - \frac{1}{n} \sum_{i=1}^{n} [(\vec{w}\vec{x} + b - y_i) * x_i] \\
b &= b - \frac{1}{n} \sum_{i=1}^{n} [(\hat{y}_i - y_i) * 1]
\end{aligned}
\tag{7}
$$