Vectorization of Gradient Descent for Logistic Regression ($\theta := \theta - \alpha\delta$)

$$X = \begin{bmatrix} x_0^{(1)} & x_1^{(1)} & x_2^{(1)} & \cdots & x_n^{(1)} \\ x_0^{(2)} & & & & \\ \vdots & & \ddots & & \vdots \\ x_0^{(m)} & x_1^{(m)} & \cdots & & x_n^{(m)} \end{bmatrix} \rightarrow \mathbb{R}^{m \times n+1} \qquad \theta = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_n \end{bmatrix} \Rightarrow \mathbb{R}^{n+1 \times 1}$$

$$h_\theta(x) = \begin{bmatrix} h_\theta(x^{(1)}) \\ h_\theta(x^{(2)}) \\ \vdots \\ h_\theta(x^{(m)}) \end{bmatrix} \Rightarrow \mathbb{R}^m \qquad y = \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(m)} \end{bmatrix} \Rightarrow \mathbb{R}^m$$

$$J(\theta) = \frac{1}{m}\left(-y^T \log(h_\theta(x)) - (1-y)^T \log(1-h)\right)$$

$$h_\theta(x) = X\theta = \begin{bmatrix} \theta_0 x_0^{(1)} & \theta_1 x_1^{(1)} & \cdots & \theta_n x_n^{(1)} \\ \vdots & & & \vdots \\ \theta_0 x_0^{(m)} & & & \theta_n x_n^{(m)} \end{bmatrix}$$

$$\boxed{\theta := \theta - \frac{\alpha}{m}\sum_{i=1}^{m}\left(h_\theta(x^{(i)}) - y^{(i)}\right)x^{(i)}} \quad m \times 1$$

Code ⇒

```
theta = theta - (α/m)* X' * (X*theta - y)
```

Mathematical ⇒

$$\boxed{\theta := \theta - \frac{\alpha}{m} X^T (g(X\theta) - y)}$$