# Neural Networks theory video notation

The videos on Neural Networks use a slightly different notation than we use in the course, which changes a few things:

1. The parameter vector $w$ is called $\theta$ (theta) instead.
2. The bias term is included in the parameter vector as $\theta_0$.
3. The input vector $x$ is expanded to also contain $x_0$, which is a constant input always set to 1. This constant "feature" get multiplied with $\theta_0$, and adds the bias term that way, which removes the need to add the bias separately after multiplying the inputs and the parameter vector.

This means you might see the model function for a Logistic Regression using just two input features and a linear decision boundary written as:

$$g(\theta^T x)$$

where

$$\theta = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \end{bmatrix}, \; x = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \end{bmatrix}$$

If you write out the vector multiplication, this is just the same as

$$g(\theta_0 x_0 + \theta_1 x_1 + \theta_2 x_2)$$

Next, if we apply the $3^{rd}$ rule stating $x_0$ is always just 1, this becomes

$$g(\theta_0 + \theta_1 x_1 + \theta_2 x_2)$$

Then, applying the $2^{nd}$ rule, we can replace $\theta_0$ with the bias term $b$

$$g(b + \theta_1 x_1 + \theta_2 x_2)$$

Finally, we can apply the $1^{st}$ rule, replacing $\theta$ with $w$

$$g(b + w_1 x_1 + w_2 x_2)$$

which should be the version of the equation you are familiar with.

Note that, conceptually, these notations do exactly the same thing, as this really just a slightly different way of writing the same Logistic Regression.