Motivation of matrix products

Herman Jaramillo
Universidad de Medellín

April 23, 2024

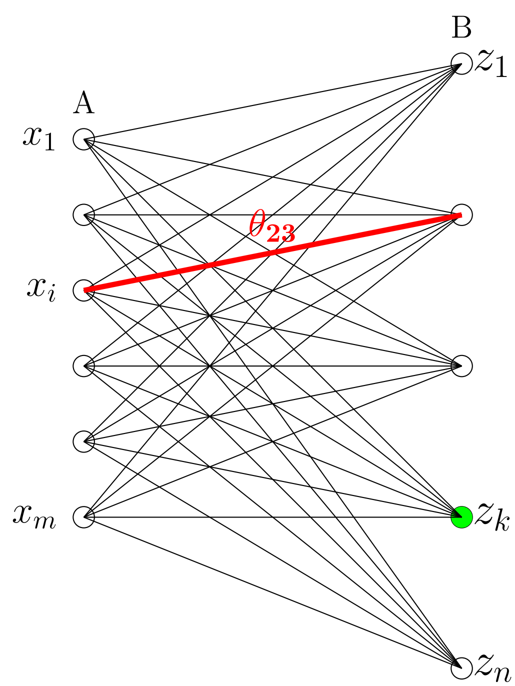

Figure 1: Connection of two consecutive layers in a neural network.

This small document has as a purpose to show how the matrix multiplication, weird as it could be its definition in linear algebra (as compared with the "more" natural Hadamard product), comes natural as an exercise of propagation in artificial neural networks. Thik of an input layer A connected to the following layer B as shown in Figure 1

Choose z_k in the first hidden layer $\ell = 2$. We want to have that node to have all the contributions of the input layer $x_1, x_2, \cdots, x_m$. That is, we want that the value z_k at the output node k. The weights are the values θ_{ij} where i is the node in the output layer and j the node in the input layer. Then, the weighted average is given by

$$z_k = \sum_{i=1}^{m} x_i \theta_{ki} = \sum_{i=1}^{m} \theta_{ki} x_i.$$

This, in matrix form is

$$z = \Theta x$$

where x is the vector $(x_1, x_2, \cdots, x_m)^T$ and Θ is the $n \times m$ matrix of weights. Think of two column vectors x_1, x_2 and compute this twice for each column vector, that is

$$\begin{aligned} z_1 &= \Theta x_1 \\ z_2 &= \Theta x_2 \end{aligned}$$

where z_1, z_2 are now two column vectors. We can think of $x_1, x_2, \cdots, x_p$ column vectors with coordinates for $x_i = (x_{1i}, x_{2i}, \cdots x_{mi})^T$ for $i = 1, 2, \cdots, p$, and get the $z_1, x_2, \cdots, z_p$ outputs as

$$\begin{aligned} z_1 &= \Theta x_1 \\ z_2 &= \Theta x_2 \\ &\vdots \\ z_p &= \Theta x_p \end{aligned}$$

where now $z_k = (z_{1k}, z_{2k}, \cdots, z_{pk})^T$, and

$$z_{ik} = (\Theta x_k)_i = \sum_{j=1}^{m} \theta_{ij} x_{jk} \tag{1}$$

with $i = 1, 2, \cdots, n$ and $k = 1, \cdots, p$, Equation 1 is the product matrix:

$$Z = \Theta X$$

as defined usually in linear algebra.

In our class notes, one of the matrices shows up as transposed, since we use the notation θ_{ij} for i the input node and j the ouput node (that is, the reverse order as the one used here).