## **Understanding the dimensions of gradients**

What are we interested in?

1. Consider the backpropagation illustration from the previous section
2. What we are interested in is a(where true output y = 1, L = Layer number, *l* is the index of the correct class-label for the given input, and i is the neuron number)
3. We know that is dependent on a31 and a32
4. Therefore, the derivative at the output layer

|  |  |  |  |
| --- | --- | --- | --- |
| = |  |  |  |
|  |

1. In the above gradient, L = 3 and i ∊ {1, 2}
2. Henceforth, we can use these notations in place of numbers to simplify gradient calculation for all possible gradients.