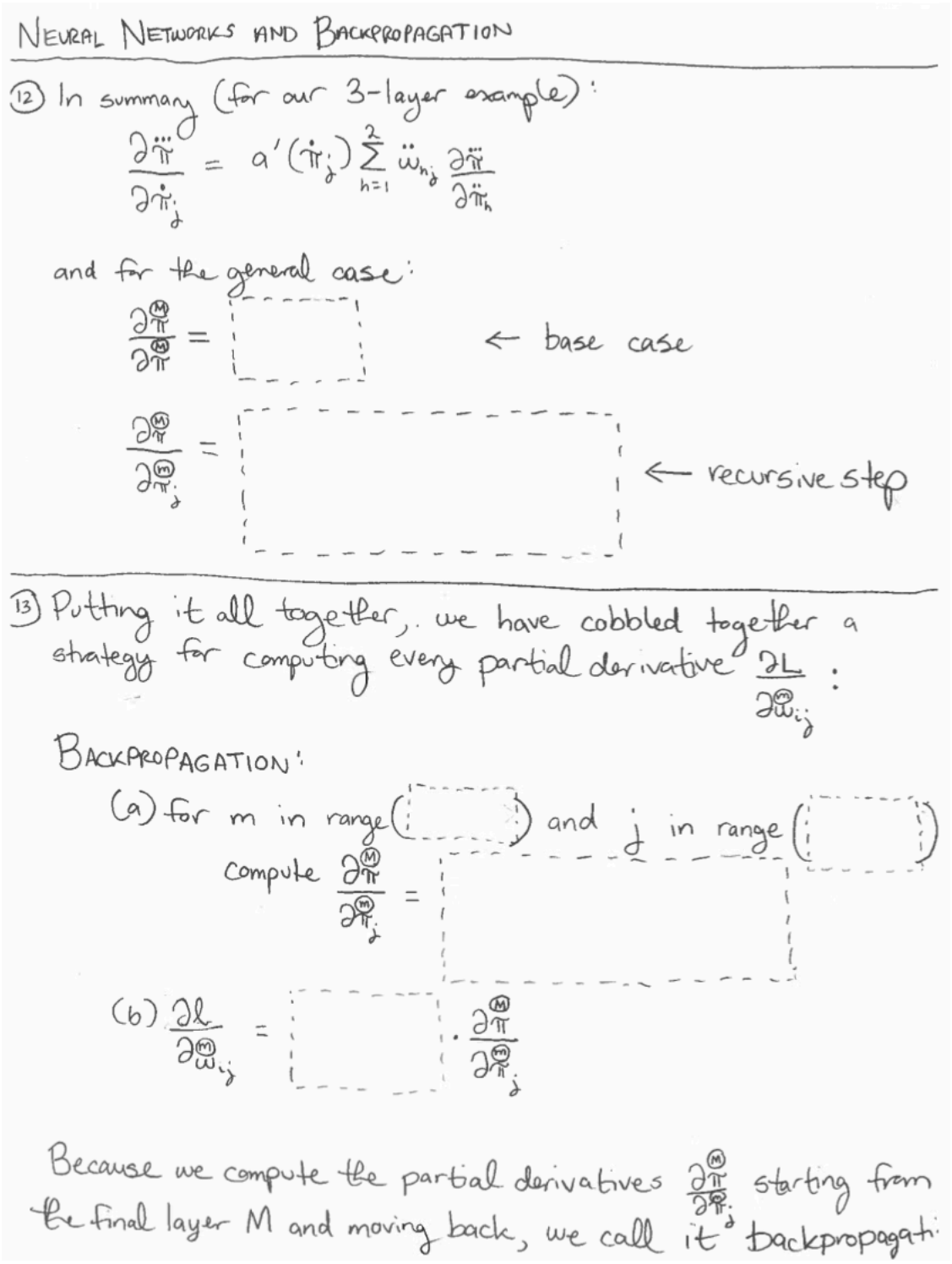
HW15: Backpropagation

Complete the following page from the "Neural Networks and Backpropagation" lecture notes:



NEURAL NETWORKS AND BACKPROPAGATION

(12) In summary (for our 3-layer example):

$$\frac{\partial \ddot{\pi}}{\partial \dot{\pi}_j} = a'(\dot{\pi}_j) \sum_{h=1}^{2} \ddot{w}_{hj} \frac{\partial \ddot{\pi}}{\partial \ddot{\pi}_h}$$

and for the general case:

$$\frac{\partial \overset{(M)}{\pi}}{\partial \overset{(M)}{\pi}} = \square \quad \leftarrow \text{base case}$$

$$\frac{\partial \overset{(M)}{\pi}}{\partial \overset{(m)}{\pi}_j} = \square \quad \leftarrow \text{recursive step}$$

(13) Putting it all together, we have cobbled together a strategy for computing every partial derivative $\frac{\partial L}{\partial \overset{(m)}{w}_{ij}}$:

BACKPROPAGATION:

(a) for m in range(□) and j in range(□)

compute $\frac{\partial \overset{(m)}{\pi}}{\partial \overset{(m)}{\pi}_j} =$ □

(b) $\frac{\partial L}{\partial \overset{(m)}{w}_{ij}} =$ □ $\cdot \frac{\partial \overset{(m)}{\pi}}{\partial \overset{(m)}{\pi}_j}$

Because we compute the partial derivatives $\frac{\partial \overset{(m)}{\pi}}{\partial \overset{(m)}{\pi}_j}$ starting from the final layer M and moving back, we call it backpropagati