$$A_{ij} \mapsto A_{ij} - \eta \nabla_{ij} \mathcal{L}$$

learning rate

back-propagation: weights updated backwards and gradients are recycled.