COMS30017 Computational Neuroscience

Week 7 / Video 1 / Supervised learning using the delta rule

Dr. Laurence Aitchison
laurence.aitchison@bristol.ac.uk

Intended Learning Outcomes

- Definition and intuition behind delta-rule
- Formal derivation as error-minimization

The delta rule: definitions

- inputs: x_i
- (synaptic) weights: w_i
- output: $y = \sum_i w_i x_i$
- "target" or optimal output: y^*
- delta, or error $\delta = y^* - y$
- learning rate, η
- weight update:

$$\Delta w_\alpha = \eta \delta x_\alpha$$

The delta rule: intuitions

- $y = \sum_i w_i x_i$
- $\delta = y^* - y$
- $\Delta w_i = \eta \delta x_i$
- If output is too low, $y < y^*$, then $0 < \delta$ and (assuming inputs, x_i, are positive), $0 < \Delta w_i$, so next time output, y, is higher.
- Note that we only update the weights corresponding to inputs that were on. If the inputs were off, $x_i = 0$, then $\Delta w_i = 0$

The delta rule: example

- two inputs, $x_1 = 0$, $x_2 = 1$.
- initialize weights to one, $w_1 = 1$, $w_2 = 1$
- output $y = w_1 \times x_1 + w_2 \times x_2 = 1$
- target, $y^* = 3$
- delta, $\delta = y^* - y = 3 - 1 = 2$
- learning rate, $\eta = 0.1$
- weight updates:

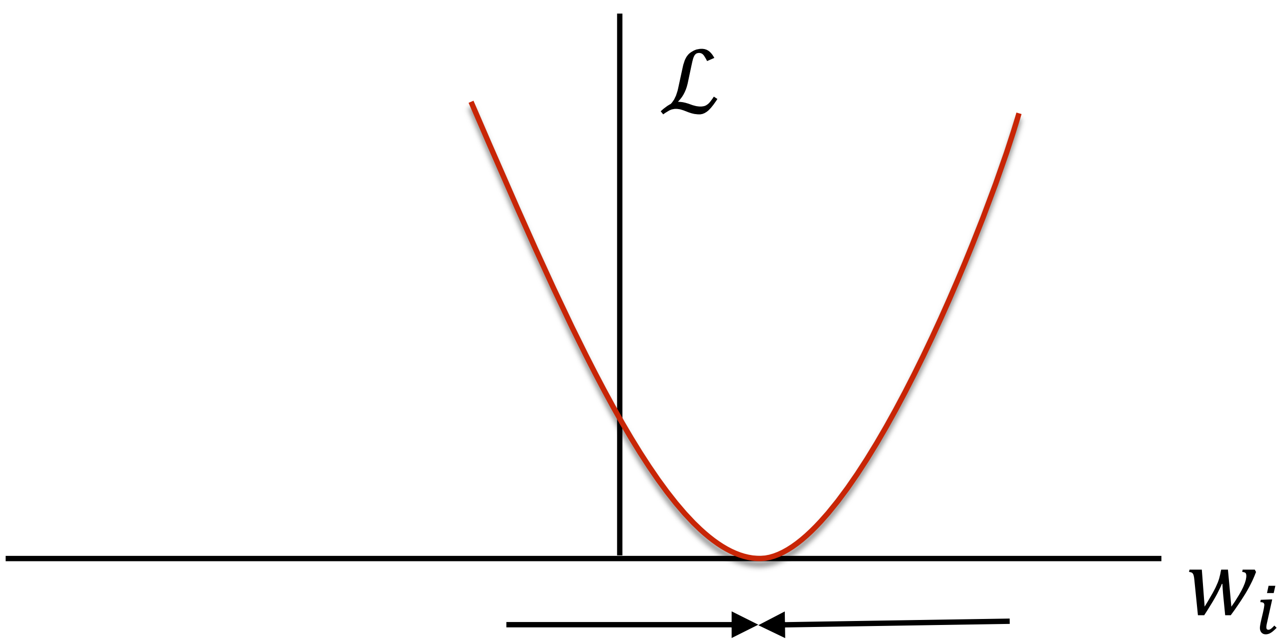
$$\Delta w_1 = \eta \delta x_1 = 0.1 \times 2 \times 0 = 0$$
$$\Delta w_2 = \eta \delta x_2 = 0.1 \times 2 \times 1 = 0.2$$

- final weights, $w_1 = 1$, $w_2 = 1.2$
- so next time, $y = w_1 \times x_1 + w_2 \times x_2 = 1.2$ is closer to the target

The delta rule: derivation (1/3)

- define a "loss" function, the squared error between the target and output,

$$\mathcal{L} = \tfrac{1}{2}(y^* - y)^2 = \tfrac{1}{2}\left(y^* - \sum_i w_i x_i\right)^2 .$$



Learning rule pushes weights "down hill"

The delta rule: derivation (2/3)

- By the definition of the loss,

$$\frac{\partial \mathcal{L}}{\partial w_\alpha} = \tfrac{1}{2}\frac{\partial}{\partial w_\alpha}\left(y^* - \sum_i w_i x_i\right)^2$$

- Applying the chain rule,

$$\frac{\partial \mathcal{L}}{\partial w_\alpha} = \left(y^* - \sum_i w_i x_i\right)\frac{\partial}{\partial w_\alpha}\left(y^* - \sum_i w_i x_i\right)$$

Noting that the gradient of y, x_i and $w_{\{i \neq \alpha\}}$ are all zero,

$$\frac{\partial \mathcal{L}}{\partial w_\alpha} = \left(y^* - \sum_i w_i x_i\right)\frac{\partial}{\partial w_\alpha}(-w_\alpha x_\alpha)$$

$$\frac{\partial \mathcal{L}}{\partial w_\alpha} = \left(y^* - \sum_i w_i x_i\right)(-x_\alpha)$$

$$\frac{\partial \mathcal{L}}{\partial w_\alpha} = -\delta x_\alpha$$

The delta rule: derivation (3/3)

To go downhill a bit, we need:

$$\Delta w_\alpha = -\eta \frac{\partial \mathcal{L}}{\partial w_\alpha}$$

(note the gradient points up hill, so we use minus the gradient to go downhill).

Substituting the gradient from the previous slide, we get back the delta rule,

$$\Delta w_\alpha = \eta \delta x_\alpha$$

End