3. prednáška

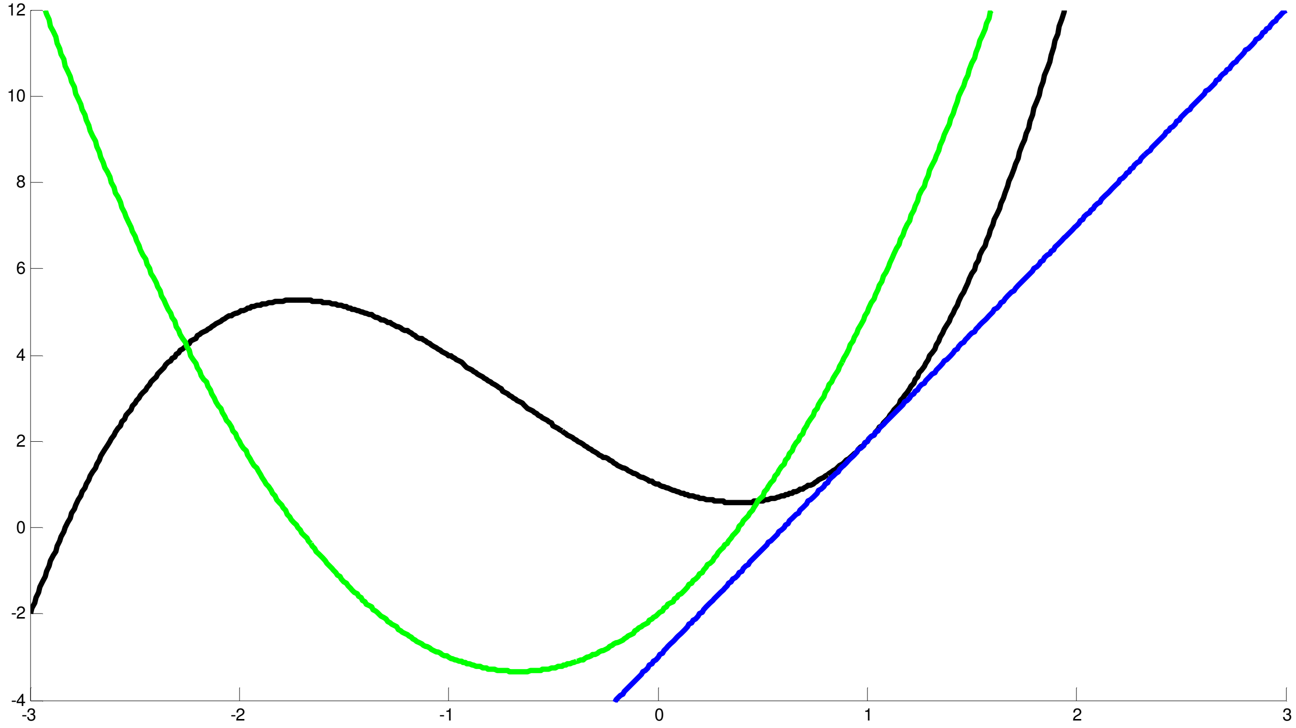
Backpropagation odvodenie

The Steepest Descent Gradient Method

$$f(x) = x^n \qquad f'(x) = nx^{(n-1)}$$

$$f(x) = e^x \qquad f'(x) = e^x$$

$$f(x) = x^3 + 2x^2 - 2x + 1 \qquad f'(x) = 3x^2 + 4x - 2$$



Chain rule

$$\varphi(x) = f(x) + g(x) \qquad \varphi'(x) = f'(x) + g'(x)$$

$$\varphi(x) = (f \circ g)(x) \qquad \varphi'(x) = (f' \circ g)(x) g'(x)$$

$$\varphi(x) = f(g(x)) \qquad \varphi'(x) = f'(g(x)) g'(x)$$

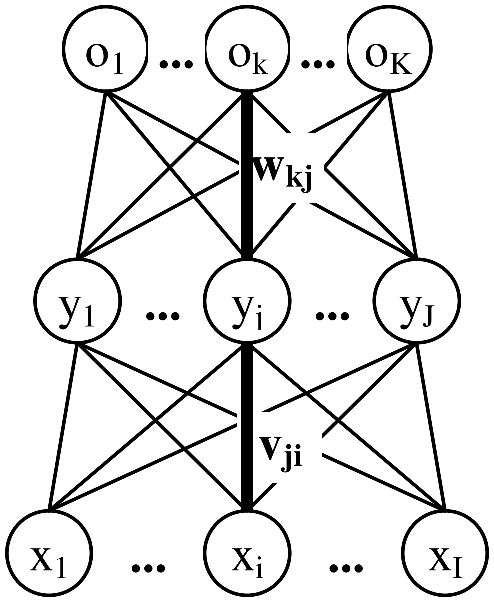
$$\varphi(x) = f(g(h(l(x)))) \qquad \varphi'(x) = f'(g(h(l(x)))) g'(h(l(x))) h'(l(x)) l'(x)$$

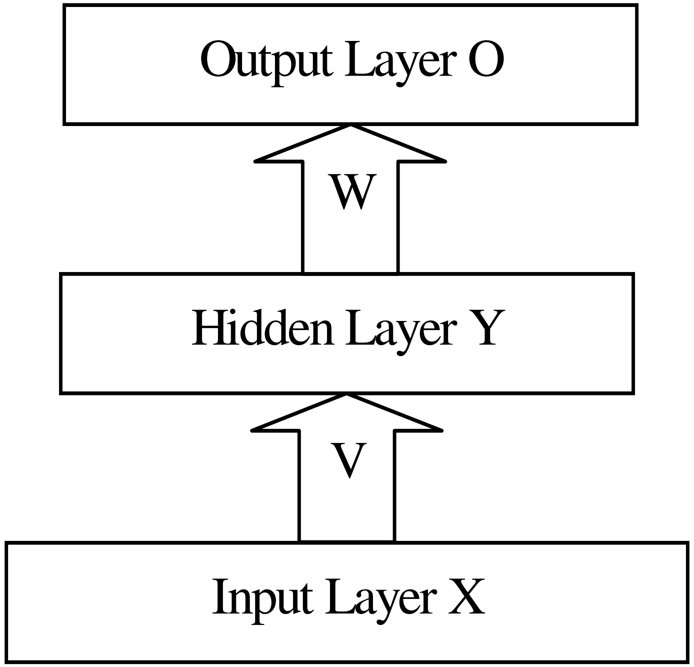
$$f(x) = \frac{1}{1+e^{-x}}$$

$$f'(x) = -\left(1+e^{-x}\right)^{-2} e^{-x}(-1) = \frac{e^{-x}}{\left(1+e^{-x}\right)^2} = \frac{1}{\left(1+e^{-x}\right)} \frac{e^{-x}}{\left(1+e^{-x}\right)} =$$

$$\frac{1}{\left(1+e^{-x}\right)} \frac{1+e^{-x}-1}{\left(1+e^{-x}\right)} = \frac{1}{\left(1+e^{-x}\right)}\left(1-\frac{1}{\left(1+e^{-x}\right)}\right) = f(x)(1-f(x))$$

Feedforward Neural Network – Multilayer Perceptron





- dvojvrstvová dopredná neurónová sieť
- dopredné šírenie signálu
- spätné šírenie chybového signálu

Dopredné šírenie

- vstupný vektor $\bar{x} = (x_1 \ldots x_I)$,
- vypočítaný výstupný vektor $\bar{o} = (o_1 \ldots o_K)$

- pre skrytý neurón j sa vypočíta jeho „net výstup“ $\tilde{y}_j$ ako $\tilde{y}_j = \sum_{i=1}^{I} v_{ji} x_i$ a vyprodukuje výstup $y_j = f(\tilde{y}_j)$.

- pre výstupný neurón k sa vypočíta jeho „net výstup“ $\tilde{o}_k$ ako $\tilde{o}_k = \sum_{j=1}^{J} w_{kj} y_j$ a vyprodukuje výstup $o_k = f(\tilde{o}_k)$.

- v_{ji} je váhové prepojenie spájajúce skrytý neurón j so vstupom i
- w_{kj} je váhové prepojenie spájajúce výstupný neurón k so skrytým neurónom j
- f je aktivačná funkcia
- α rýchlosť učenia

Odvodenie spätného šírenia chyby

- trénovacia množina – P vzorov $\bar{x}^p = \left(x_1^p \ldots x_I^p\right),\ \bar{d}^p = \left(d_1^p \ldots d_K^p\right)$
- minimalizácia celkovej chyby: $E = \sum_{p=1}^{P} E_p$
- čiastkové chyby: $E^p = \frac{1}{2}\sum_{k=1}^{K}(d_k^p - o_k^p)^2$
- aktivita výstupných neurónov: $o_k^p = f\left(\tilde{o}_k^p\right)$
- „net“ aktivita výstupných neurónov: $\tilde{o}_k^p = \sum_{j=1}^{J} w_{kj} y_j^p$
- aktivita skrytých neurónov: $y_j^p = f\left(\tilde{y}_j^p\right)$
- „net“ aktivita skrytých neurónov: $\tilde{y}_j^p = \sum_{i=1}^{I} v_{ji} x_i^p$
- zmena váh výstupných neuronov: $\Delta w_{kj} = -\alpha \frac{\partial E}{\partial w_{kj}}$
- zmena váh skrytých neuronov: $\Delta v_{ji} = -\alpha \frac{\partial E}{\partial v_{ji}}$

Odvodenie spätného šírenia chyby – výstupné váhy

$$\frac{\partial E}{\partial w_{kj}} = \sum_{p=1}^{P} \frac{\partial E^p}{\partial w_{kj}}$$

$$\frac{\partial E^p}{\partial w_{kj}} = \frac{\partial E^p}{\partial o_k^p} \frac{\partial o_k^p}{\partial \tilde{o}_k^p} \frac{\partial \tilde{o}_k^p}{\partial w_{kj}}$$

$$\delta^{\text{out}\,p}_{\;k} = -\frac{\partial E^p}{\partial \tilde{o}_k^p} = -\frac{\partial E^p}{\partial o_\kappa^p} \frac{\partial o_k^p}{\partial \tilde{o}_k^p}$$

$$\frac{\partial E^p}{\partial o_k^p} = \tfrac{1}{2} 2(d_k^p - o_k^p)(-1) = -(d_k^p - o_k^p)$$

$$\frac{\partial o_k^p}{\partial \tilde{o}_k^p} = f'\left(\tilde{o}_k^p\right)$$

$$\frac{\partial \tilde{o}_k^p}{\partial w_{kj}} = y_j^p$$

$$\frac{\partial E^p}{\partial w_{kj}} = -(d_k^p - o_k^p) f'\left(\tilde{o}_k^p\right) y_j^p$$

$$\delta^{\text{out}\,p}_{\;k} = -\frac{\partial E^p}{\partial \tilde{o}_k^p} = -\frac{\partial E^p}{\partial o_k^p} \frac{\partial o_k^p}{\partial \tilde{o}_k^p} = (d_k^p - o_k^p) f'\left(\tilde{o}_k^p\right)$$

Batch update:

$$\Delta w_{kj} = -\alpha \frac{\partial E}{\partial w_{kj}} = \alpha \sum_{p=1}^{P} (d_k^p - o_k^p) f'\left(\tilde{o}_k^p\right) y_j^p$$

$$\Delta w_{kj} = -\alpha \frac{\partial E}{\partial w_{kj}} = \alpha \sum_{p=1}^{P} \delta^{\text{out}\,p}_{\;k} y_j^p$$

Iterative update:

$$\Delta w_{kj} = -\alpha \frac{\partial E^p}{\partial w_{kj}} = \alpha (d_k^p - o_k^p) f'\left(\tilde{o}_k^p\right) y_j^p$$

$$\Delta w_{kj} = -\alpha \frac{\partial E^p}{\partial w_{kj}} = \alpha \delta^{\text{out}\,p}_{\;k} y_j^p$$

Odvodenie spätného šírenia chyby – skryté váhy

$$\frac{\partial E}{\partial v_{ji}} = \sum_{p=1}^{P} \frac{\partial E^p}{\partial v_{ji}}$$

$$\frac{\partial E^p}{\partial v_{ji}} = \sum_{k=1}^{K} \left[\frac{\partial E^p}{\partial o_k^p} \frac{\partial o_k^p}{\partial \tilde{o}_k^p} \frac{\partial \tilde{o}_k^p}{\partial y_j^p} \frac{\partial y_j^p}{\partial \tilde{y}_j^p} \frac{\partial \tilde{y}_j^p}{\partial v_{ji}} \right] = \left[\sum_{k=1}^{K} \frac{\partial E^p}{\partial o_k^p} \frac{\partial o_k^p}{\partial \tilde{o}_k^p} \frac{\partial \tilde{o}_k^p}{\partial y_j^p} \right] \frac{\partial y_j^p}{\partial \tilde{y}_j^p} \frac{\partial \tilde{y}_j^p}{\partial v_{ji}}$$

$$\delta^{\text{hid}}{}_j^p = -\frac{\partial E^p}{\partial \tilde{y}_j^p} = -\left[\sum_{k=1}^{K} \frac{\partial E^p}{\partial o_k^p} \frac{\partial o_k^p}{\partial \tilde{o}_k^p} \frac{\partial \tilde{o}_k^p}{\partial y_j^p} \right] \frac{\partial y_j^p}{\partial \tilde{y}_j^p}$$

$$\frac{\partial E^p}{\partial o_\kappa^p} \frac{\partial o_k^p}{\partial \tilde{o}_k^p} = \frac{\partial E^p}{\partial \tilde{o}_k^p} = -\delta^{\text{out}}{}_k^p$$

$$\frac{\partial \tilde{o}_k^p}{\partial y_j^p} = w_{kj}$$

$$\frac{\partial y_j^p}{\partial \tilde{y}_j^p} = f'\left(\tilde{y}_j^p\right)$$

$$\frac{\partial \tilde{y}_j^p}{\partial v_{ji}} = x_i^p$$

$$\frac{\partial E^p}{\partial v_{ji}} = \sum_{k=1}^{K} \left[\frac{\partial E^p}{\partial \tilde{o}_k^p} w_{kj} \right] f'\left(\tilde{y}_j^p\right) x_i^p$$

$$\delta^{\text{hid}}{}_j^p = -\frac{\partial E^p}{\partial \tilde{y}_j^p} = -\left[\sum_{k=1}^{K} \frac{\partial E^p}{\partial o_k^p} \frac{\partial o_k^p}{\partial \tilde{o}_k^p} \frac{\partial \tilde{o}_k^p}{\partial y_j^p} \right] \frac{\partial y_j^p}{\partial \tilde{y}_j^p} = \left[\sum_{k=1}^{K} \delta^{\text{out}}{}_k^p w_{kj} \right] f'\left(\tilde{y}_j^p\right)$$

Batch update:

$$\Delta v_{ji} = -\alpha \frac{\partial E}{\partial v_{ji}} = -\alpha \sum_{p=1}^{P} \left[\sum_{k=1}^{K} \frac{\partial E^p}{\partial \tilde{o}_k^p} w_{kj} \right] f'\left(\tilde{y}_j^p\right) x_i^p$$

$$\Delta v_{ji} = -\alpha \frac{\partial E}{\partial v_{ji}} = \alpha \sum_{p=1}^{P} \delta^{\text{hid}}{}_j^p x_i^p$$

Iterative update:

$$\Delta v_{ji} = -\alpha \frac{\partial E^p}{\partial v_{ji}} = -\alpha \sum_{k=1}^{K} \left[\frac{\partial E^p}{\partial \tilde{o}_k^p} w_{kj} \right] f'\left(\tilde{y}_j^p\right) x_i^p$$

$$\Delta v_{ji} = -\alpha \frac{\partial E^p}{\partial v_{ji}} = \alpha \delta^{\text{hid}}{}_j^p x_i^p$$

Spätné šírenie

- očakávaný = želaný výstupný vektor, $\bar{d} = (d_1 \ldots d_K)$
- minimalizujeme chybu $E = \frac{1}{2}\sum_{k=1}^{K}(d_k - o_k)^2$

- zmeny výstupných váh určíme podľa $\Delta w_{kj} = -\alpha \frac{\partial E}{\partial w_{kj}} = \alpha \delta^{\text{out}}{}_k y_j$
- chybový signal $\delta^{\text{out}}{}_k$ výstupného neurónu k je definovaný ako $\delta^{\text{out}}{}_k = f'(\tilde{o}_k)(d_k - o_k)$

- zmeny skrytých váh určíme podľa $\Delta v_{ji} = -\alpha \frac{\partial E}{\partial v_{ji}} = \alpha \delta^{\text{hid}}{}_j x_i$
- chybový signal $\delta^{\text{hid}}{}_j$ skrytého neurónu j je definovaný ako $\delta^{\text{hid}}{}_j = f'(\tilde{y}_j)\sum_{k=1}^{K}\delta^{\text{out}}{}_k w_{kj}$

- úprava váhových prepojení $v_{ji} = v_{ji} + \Delta v_{ji}$ a $w_{kj} = w_{kj} + \Delta w_{kj}$

Sigmoidálna aktivačná funkcia

- často používaná aktivačná funkcia $f(x)=\frac{1}{1+e^{-x}}$
- jej derivácia môže byť jednoducho určená vzťahom $f'(x)=f(x)(1-f(x))$
- chybové signály môžu byť vyjadrené ako $\delta^{\text{out}}{}_k=(d_k-o_k)o_k(1-o_k)$ a $\delta^{\text{hid}}{}_j=y_j(1-y_j)\sum_{k=1}^{K}w_{kj}\delta^{\text{out}}{}_k$

Momentum

- úprava váh podľa vzťahov $\Delta v_{ji}(t)=\alpha\delta^{\text{hid}}{}_j x_i+\beta\Delta v_{ji}(t-1)$ a $\Delta w_{kj}(t)=\alpha\delta^{\text{out}}{}_k y_j+\beta\Delta w_{kj}(t-1)$
- β je momentum