

# Neuronové sítě-Delta učení

Filip Habr

České vysoké učení technické v Praze  
Fakulta jaderná a fyzikálně inženýrská

30. března 2009

# Obsah prezentace

## Delta učení

- 1 Teorie k delta učení
- 2 Zadání příkladu
- 3 Vyhledání  $m_j$  a  $M_j \rightarrow$  výpočet standardizovaných hodnot
- 4 Standardizovaná tabulka
- 5 Výpočet prvních dvou iterací  $\bar{w}_1$  a  $\bar{w}_2$  a výsledného  $\bar{w}$
- 6 Grafické znázornění

# Příprava dat

Data  $\mathbf{X} \in \mathbb{R}^{m \times n}$  je vhodné před spuštěním perceptronu standardizovat  $\rightarrow$  např. rozpětím:

$$\mathbf{x}_{ij}^{new} = \frac{\mathbf{x}_{ij} - m_j}{M_j - m_j} \in \langle 0, 1 \rangle$$

$$m_j = \min(\mathbf{X}_{kj})$$

$$M_j = \max(\mathbf{X}_{kj})$$

# Delta učení

- $y = f(\bar{x}, \bar{w})$
- $y_k = f(\bar{x}_k, \bar{w})$
- $\Delta_k = y_k^* - y_k$
- $p_k(\bar{w}) = \frac{1}{2} \Delta_k^2$
- $P(\bar{w}) = \sum_{k=1}^m p_k(\bar{w}) = \min$

$$P(\bar{w}) = \sum_{k=1}^m \frac{1}{2} \Delta_k^2 = \sum_{k=1}^m \frac{1}{2} (y_k^* - y_k)^2$$

$$= \sum_{k=1}^m \frac{1}{2} (y_k^* - f(\bar{x}_k, \bar{w}))^2 = \min$$

# Delta učení

## Algoritmus

- $\bar{w}_0 \in \mathbb{R}^{n+1}$  (malé),  $\lambda_0 > 0$
- náhodná volba k. vzoru,  $\lambda_j = \frac{\lambda_0}{1+j}$
- výpočet  $y_k$  a  $\Delta_k$
- výpočet citlivosti sítě na váhy
- adaptace vah

## Učení hladkého perceptronu

- $y = f(\bar{x}, \bar{w}) = \tanh(\sum_{i=0}^n w_i x_i)$
- $\frac{\partial \tanh(\bar{x}, \bar{w})}{\partial w_i} = (1 - y^2)x_i$
- $\bar{w}_{j+1} = \bar{w}_j + \lambda_j \Delta_k (1 - y_k^2) \bar{x}_k$

# Zadání příkladu

Tabulka hodnot:

$x_0$	$x_1$	$x_2$	$y^*$
1	1	5	-1
1	2	3,5	-1
1	2	7	1
1	3	3,5	-1
1	3	4,5	1
1	4	4	1
1	4	6	1
1	5	5	1
1	6	5,5	1
1	8	2	1

## Zadání příkladu

Tabulka hodnot:

$x_0$	$x_1$	$x_2$	$y^*$
1	1	5	-1
1	2	3,5	-1
1	2	7	1
1	3	3,5	-1
1	3	4,5	1
1	4	4	1
1	4	6	1
1	5	5	1
1	6	5,5	1
1	8	2	1

Výpočet  $m_j$  a  $M_j$  :

- $m_1 = 1$
- $m_2 = 2$
- $M_1 = 8$
- $M_2 = 7$

## Zadání příkladu

Tabulka hodnot:

$x_0$	$x_1$	$x_2$	$y^*$
1	1	5	-1
1	2	3,5	-1
1	2	7	1
1	3	3,5	-1
1	3	4,5	1
1	4	4	1
1	4	6	1
1	5	5	1
1	6	5,5	1
1	8	2	1

Výpočet  $m_j$  a  $M_j$  :

- $m_1 = 1$
- $m_2 = 2$
- $M_1 = 8$
- $M_2 = 7$

Standardizované  
hodnoty:

$x_0$	$x_1$	$x_2$	$y^*$
1	0	0,6	-1
1	0,14	0,3	-1
1	0,14	1	1
1	0,28	0,3	-1
1	0,28	0,5	1
1	0,43	0,4	1
1	0,43	0,8	1
1	0,57	0,6	1
1	0,71	0,7	1
1	1	0	1



# První iterace

- Zvolili jsme  $\lambda_0 = 1$

# První iterace

- Zvolili jsme  $\lambda_0 = 1$
- Jako první vzor jsme náhodně vybrali:  $k = 2$

# První iterace

- Zvolili jsme  $\lambda_0 = 1$
- Jako první vzor jsme náhodně vybrali:  $k = 2$
- Váhový vektor jsme zvolili  $\bar{w}_0 = (0; 0; 0)$

# První iterace

- Zvolili jsme  $\lambda_0 = 1$
- Jako první vzor jsme náhodně vybrali:  $k = 2$
- Váhový vektor jsme zvolili  $\bar{w}_0 = (0; 0; 0)$
- $y_2 = \tanh(0 \cdot 1 + 0 \cdot 0,14 + 0 \cdot 0,3) = 0$   
 $\Rightarrow y = \tanh(\sum_{i=0}^n w_i x_i)$

# První iterace

- Zvolili jsme  $\lambda_0 = 1$
- Jako první vzor jsme náhodně vybrali:  $k = 2$
- Váhový vektor jsme zvolili  $\bar{w}_0 = (0; 0; 0)$
- $y_2 = \tanh(0 \cdot 1 + 0 \cdot 0,14 + 0 \cdot 0,3) = 0$   
 $\Rightarrow y = \tanh(\sum_{i=0}^n w_i x_i)$
- $\Delta_2 = -1 - 0 = -1$   $\Rightarrow \Delta_k = y_k^* - y_k$

# První iterace

- Zvolili jsme  $\lambda_0 = 1$
- Jako první vzor jsme náhodně vybrali:  $k = 2$
- Váhový vektor jsme zvolili  $\bar{w}_0 = (0; 0; 0)$
- $y_2 = \tanh(0 \cdot 1 + 0 \cdot 0,14 + 0 \cdot 0,3) = 0$   
 $\Rightarrow y = \tanh(\sum_{i=0}^n w_i x_i)$
- $\Delta_2 = -1 - 0 = -1 \qquad \Rightarrow \Delta_k = y_k^* - y_k$
- $\bar{w}_1 = (0; 0; 0) + 1 \cdot (-1) \cdot (1 - 0^2) \cdot (1; 0,14; 0,3) =$   
 $(-1; -0,14; -0,3)$   
 $\Rightarrow \bar{w}_{j+1} = \bar{w}_j + \lambda_j \Delta_k (1 - y_k^2) \bar{x}_k$

# První iterace

- Zvolili jsme  $\lambda_0 = 1$
- Jako první vzor jsme náhodně vybrali:  $k = 2$
- Váhový vektor jsme zvolili  $\bar{w}_0 = (0; 0; 0)$
- $y_2 = \tanh(0 \cdot 1 + 0 \cdot 0,14 + 0 \cdot 0,3) = 0$   
 $\Rightarrow y = \tanh(\sum_{i=0}^n w_i x_i)$
- $\Delta_2 = -1 - 0 = -1 \qquad \Rightarrow \Delta_k = y_k^* - y_k$
- $\bar{w}_1 = (0; 0; 0) + 1 \cdot (-1) \cdot (1 - 0^2) \cdot (1; 0,14; 0,3) =$   
 $(-1; -0,14; -0,3)$   
 $\Rightarrow \bar{w}_{j+1} = \bar{w}_j + \lambda_j \Delta_k (1 - y_k^2) \bar{x}_k$
- Rovnice přímky tedy vypadá takto:  $-0,3x_1 - 1x_2 - 0,14 = 0$

# První iterace

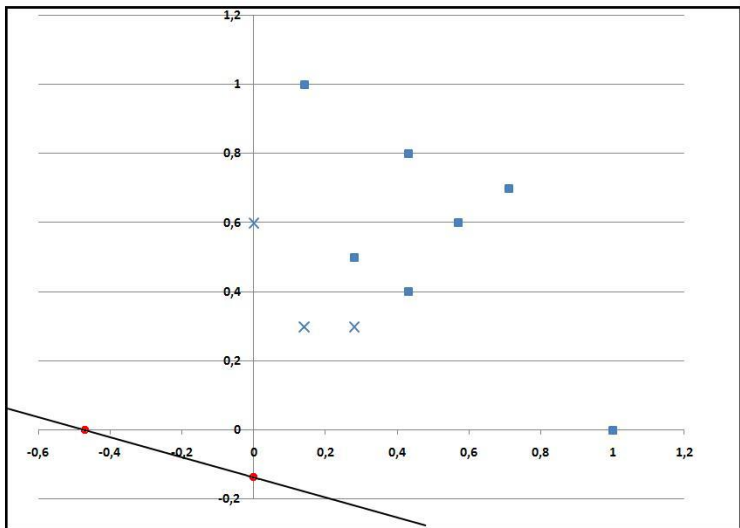
- Zvolili jsme  $\lambda_0 = 1$
- Jako první vzor jsme náhodně vybrali:  $k = 2$
- Váhový vektor jsme zvolili  $\bar{w}_0 = (0; 0; 0)$
- $y_2 = \tanh(0 \cdot 1 + 0 \cdot 0,14 + 0 \cdot 0,3) = 0$   
 $\Rightarrow y = \tanh(\sum_{i=0}^n w_i x_i)$
- $\Delta_2 = -1 - 0 = -1 \quad \Rightarrow \Delta_k = y_k^* - y_k$
- $\bar{w}_1 = (0; 0; 0) + 1 \cdot (-1) \cdot (1 - 0^2) \cdot (1; 0,14; 0,3) =$   
 $(-1; -0,14; -0,3)$   
 $\Rightarrow \bar{w}_{j+1} = \bar{w}_j + \lambda_j \Delta_k (1 - y_k^2) \bar{x}_k$
- Rovnice přímky tedy vypadá takto:  $-0,3x_1 - 1x_2 - 0,14 = 0$
- Průsečík s osou  $x_1$  je:  $-0,47$



# První iterace

- Zvolili jsme  $\lambda_0 = 1$
- Jako první vzor jsme náhodně vybrali:  $k = 2$
- Váhový vektor jsme zvolili  $\bar{w}_0 = (0; 0; 0)$
- $y_2 = \tanh(0 \cdot 1 + 0 \cdot 0,14 + 0 \cdot 0,3) = 0$   
 $\Rightarrow y = \tanh(\sum_{i=0}^n w_i x_i)$
- $\Delta_2 = -1 - 0 = -1 \quad \Rightarrow \Delta_k = y_k^* - y_k$
- $\bar{w}_1 = (0; 0; 0) + 1 \cdot (-1) \cdot (1 - 0^2) \cdot (1; 0,14; 0,3) =$   
 $(-1; -0,14; -0,3)$   
 $\Rightarrow \bar{w}_{j+1} = \bar{w}_j + \lambda_j \Delta_k (1 - y_k^2) \bar{x}_k$
- Rovnice přímky tedy vypadá takto:  $-0,3x_1 - 1x_2 - 0,14 = 0$
- Průsečík s osou  $x_1$  je:  $-0,47$
- Průsečík s osou  $x_2$  je:  $-0,14$

# První iterace



# Druhá iterace

- Jako druhý vzor jsme náhodně vybrali:  $k = 7$

# Druhá iterace

- Jako druhý vzor jsme náhodně vybrali:  $k = 7$
- $y_7 = \tanh((-1) \cdot 1 + (-0,14) \cdot 0,43 + (-0,3) \cdot 0,8) = -0,86$

# Druhá iterace

- Jako druhý vzor jsme náhodně vybrali:  $k = 7$
- $y_7 = \tanh((-1) \cdot 1 + (-0,14) \cdot 0,43 + (-0,3) \cdot 0,8) = -0,86$
- $\Delta_7 = 1 - (-0,86) = 1,86$

# Druhá iterace

- Jako druhý vzor jsme náhodně vybrali:  $k = 7$
- $y_7 = \tanh((-1) \cdot 1 + (-0,14) \cdot 0,43 + (-0,3) \cdot 0,8) = -0,86$
- $\Delta_7 = 1 - (-0,86) = 1,86$
- $\lambda_1 = 0,5 \quad \Rightarrow \lambda_j = \frac{\lambda_0}{1+j}$

# Druhá iterace

- Jako druhý vzor jsme náhodně vybrali:  $k = 7$
- $y_7 = \tanh((-1) \cdot 1 + (-0,14) \cdot 0,43 + (-0,3) \cdot 0,8) = -0,86$
- $\Delta_7 = 1 - (-0,86) = 1,86$
- $\lambda_1 = 0,5 \quad \Rightarrow \lambda_j = \frac{\lambda_0}{1+j}$
- $\bar{w}_2 =$   
 $(-1; -0,14; -0,3) + 0,5 \cdot 1,86 \cdot (1 - (-0,86)^2) \cdot (1; 0,43; 0,8) =$   
 $= (0,73; 0,6; 1,08)$

# Druhá iterace

- Jako druhý vzor jsme náhodně vybrali:  $k = 7$
- $y_7 = \tanh((-1) \cdot 1 + (-0,14) \cdot 0,43 + (-0,3) \cdot 0,8) = -0,86$
- $\Delta_7 = 1 - (-0,86) = 1,86$
- $\lambda_1 = 0,5 \quad \Rightarrow \lambda_j = \frac{\lambda_0}{1+j}$
- $\bar{w}_2 =$   
 $(-1; -0,14; -0,3) + 0,5 \cdot 1,86 \cdot (1 - (-0,86)^2) \cdot (1; 0,43; 0,8) =$   
 $= (0,73; 0,6; 1,08)$
- Rovnice přímky tedy vypadá takto:  $0.6x_1 + 1.08x_2 + 0.73 = 0$



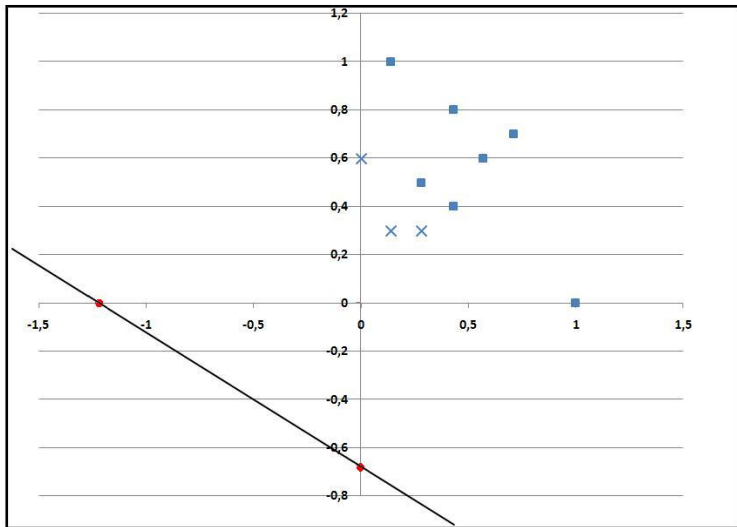
# Druhá iterace

- Jako druhý vzor jsme náhodně vybrali:  $k = 7$
- $y_7 = \tanh((-1) \cdot 1 + (-0,14) \cdot 0,43 + (-0,3) \cdot 0,8) = -0,86$
- $\Delta_7 = 1 - (-0,86) = 1,86$
- $\lambda_1 = 0,5 \quad \Rightarrow \lambda_j = \frac{\lambda_0}{1+j}$
- $\bar{w}_2 =$   
 $(-1; -0,14; -0,3) + 0,5 \cdot 1,86 \cdot (1 - (-0,86)^2) \cdot (1; 0,43; 0,8) =$   
 $= (0,73; 0,6; 1,08)$
- Rovnice přímky tedy vypadá takto:  $0.6x_1 + 1.08x_2 + 0.73 = 0$
- Průsečík s osou  $x_1$  je:  $-1,22$

# Druhá iterace

- Jako druhý vzor jsme náhodně vybrali:  $k = 7$
- $y_7 = \tanh((-1) \cdot 1 + (-0,14) \cdot 0,43 + (-0,3) \cdot 0,8) = -0,86$
- $\Delta_7 = 1 - (-0,86) = 1,86$
- $\lambda_1 = 0,5 \quad \Rightarrow \lambda_j = \frac{\lambda_0}{1+j}$
- $\bar{w}_2 =$   
 $(-1; -0,14; -0,3) + 0,5 \cdot 1,86 \cdot (1 - (-0,86)^2) \cdot (1; 0,43; 0,8) =$   
 $= (0,73; 0,6; 1,08)$
- Rovnice přímky tedy vypadá takto:  $0.6x_1 + 1.08x_2 + 0.73 = 0$
- Průsečík s osou  $x_1$  je:  $-1,22$
- Průsečík s osou  $x_2$  je:  $-0,68$

# Druhá iterace



# Konečná iterace-výsledek

- V posledním kroku jsme vybrali:  $k = 7$

# Konečná iterace-výsledek

- V posledním kroku jsme vybrali:  $k = 7$
- $y_7 = -0,86$

# Konečná iterace-výsledek

- V posledním kroku jsme vybrali:  $k = 7$
- $y_7 = -0,86$
- $\Delta_7 = 0,36$

# Konečná iterace-výsledek

- V posledním kroku jsme vybrali:  $k = 7$
- $y_7 = -0,86$
- $\Delta_7 = 0,36$
- $\lambda_n = 2 \cdot 10^{-8}$

# Konečná iterace-výsledek

- V posledním kroku jsme vybrali:  $k = 7$
- $y_7 = -0,86$
- $\Delta_7 = 0,36$
- $\lambda_n = 2 \cdot 10^{-8}$
- $\bar{w}_n = (-0,73; 1,49; 0,94)$



# Konečná iterace-výsledek

- V posledním kroku jsme vybrali:  $k = 7$
- $y_7 = -0,86$
- $\Delta_7 = 0,36$
- $\lambda_n = 2 \cdot 10^{-8}$
- $\bar{w}_n = (-0,73; 1,49; 0,94)$
- Rovnice přímky tedy vypadá takto:  
 $1,49x_1 + 0,94x_2 + -0,73 = 0$

# Konečná iterace-výsledek

- V posledním kroku jsme vybrali:  $k = 7$
- $y_7 = -0,86$
- $\Delta_7 = 0,36$
- $\lambda_n = 2 \cdot 10^{-8}$
- $\bar{w}_n = (-0,73; 1,49; 0,94)$
- Rovnice přímky tedy vypadá takto:  
 $1,49x_1 + 0,94x_2 + -0,73 = 0$
- Průsečík s osou  $x_1$  je:  $0,49$

# Konečná iterace-výsledek

- V posledním kroku jsme vybrali:  $k = 7$
- $y_7 = -0,86$
- $\Delta_7 = 0,36$
- $\lambda_n = 2 \cdot 10^{-8}$
- $\bar{w}_n = (-0,73; 1,49; 0,94)$
- Rovnice přímky tedy vypadá takto:  
 $1,49x_1 + 0,94x_2 + -0,73 = 0$
- Průsečík s osou  $x_1$  je: 0,49
- Průsečík s osou  $x_2$  je: 0,78

# Konečná iterace

