Neuronale Netze (SS 2002), 10.4.

The simple perceptron

Some linear algebra:

- a vector in $\mathbb{R}^n$: $\vec{x} = (x_1, \ldots, x_n)$
- addition $\vec{x} + \vec{y} = (x_1 + y_1, \ldots, x_n + y_n)$
  multiplication with a real number $\lambda\vec{x} = (\lambda x_1, \ldots, \lambda x_n)$
- norm/length of a vector $\|\vec{x}\| = \sqrt{x_1^2 + \ldots + x_n^2}$ where
  - $\|\lambda\vec{x}\| = |\lambda| \cdot \|\vec{x}\|$
  - $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$
- dot product $\vec{x}^t\vec{y} = x_1 y_1 + \ldots + x_n y_n$ where
  - $(\vec{x} + \vec{y})^t\vec{z} = \vec{x}^t\vec{z} + \vec{y}^t\vec{z}$
  - $(\lambda\vec{x})^t\vec{y} = \lambda\vec{x}^t\vec{y}$
  - $\vec{x}^t\vec{y} = \vec{y}^t\vec{x}$
  - $\vec{x}^t\vec{x} = \|\vec{x}\|^2$
  - $\vec{x}^t\vec{y} = \|\vec{x}\|\|\vec{y}\| \cos\varphi$
- The equation $\vec{w}^t\vec{x} - \theta = 0$ determines a hyperplane in $\mathbb{R}^n$.
  $\vec{w}$ is an orthogonal vector, θ determines the offset.

  perceptron computes $\vec{x} \mapsto \mathrm{H}(\vec{w}^t\vec{x} - \theta)$

  $\rightarrow$ perceptron is a linear classifier.

Learning task:

- training set $P = \{(\vec{x}^i, y^i) \in \mathbb{R}^n \times \{0, 1\} \mid i = 1, \ldots, m\}$
  positive and negative patterns
- P linearly separable $\iff$ a perceptron can map all points in P correctly
- task: given a linearly separable training set, find a corresponding perceptron

- idea: examples of an unknown regularity can be observed. The observed patterns allow to recover the regularity.

Perceptron algorithm:

- W.l.o.g. ($\rightarrow$ exercise): Drop the bias.

- error signal

$$\delta(\vec{w},\vec{x}^i) = \begin{cases} 1 & \text{if } \vec{w}^t\vec{x}^i \leq 0, y^i = 1 \\ -1 & \text{if } \vec{w}^t\vec{x}^i \geq 0, y^i = 0 \\ 0 & \text{otherwise} \end{cases}$$

- perceptron algorithm to determine $\vec{w}$:

init $\vec{w}$
repeat while $\vec{x}^i$ with $\delta(\vec{w},\vec{x}^i) \neq 0$
$\quad \vec{w} := \vec{w} + \delta(\vec{w},\vec{x}^i)\vec{x}$

- geometrical interpretation:

  rotate $\vec{w}$ into the direction of positive points and to the opposite direction of negative points

- biological interpretation:

  Hebbian learning (modulo supervision)

- percepton convergence theorem:

  if a solution exists, it will be found by the above algorithm