

# Analýza a klasifikace dat – přednáška 5



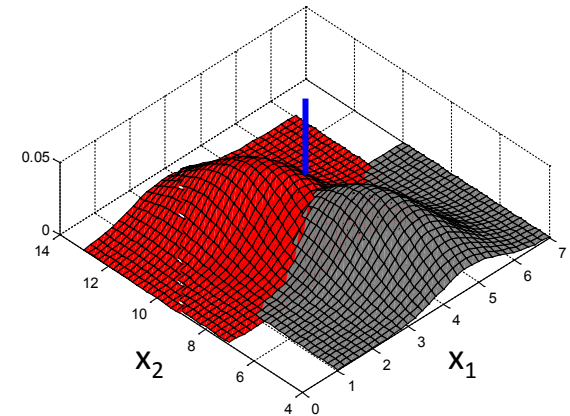
RNDr. Eva Janoušová

Podzim 2015

# Typy klasifikátorů – podle principu klasifikace

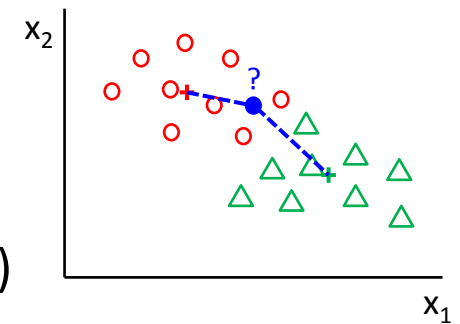
- **klasifikace pomocí diskriminačních funkcí:**

- diskriminační funkce určují míru příslušnosti k dané klasifikační třídě
- pro danou třídu má daná diskriminační funkce nejvyšší hodnotu



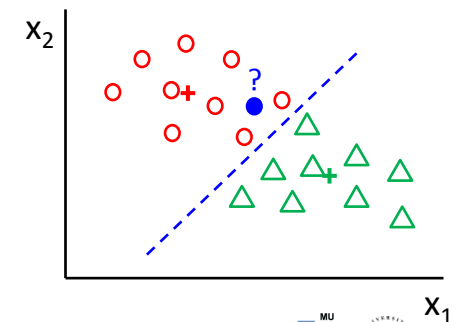
- **klasifikace pomocí vzdálenosti od etalonů klasif. tříd:**

- etalon = reprezentativní objekt(y) klasifikační třídy
- počet etalonů klasif. třídy různý – od jednoho vzorku (např. centroidu) po úplný výčet všech objektů dané třídy (např. u klasif. pomocí metody průměrné vazby)



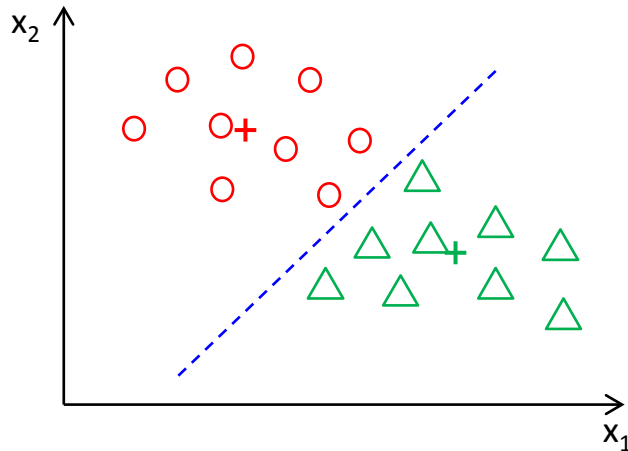
- **klasifikace pomocí hranic v obrazovém prostoru:**

- stanovení hranic (hraničních ploch) oddělujících klasifikační třídy

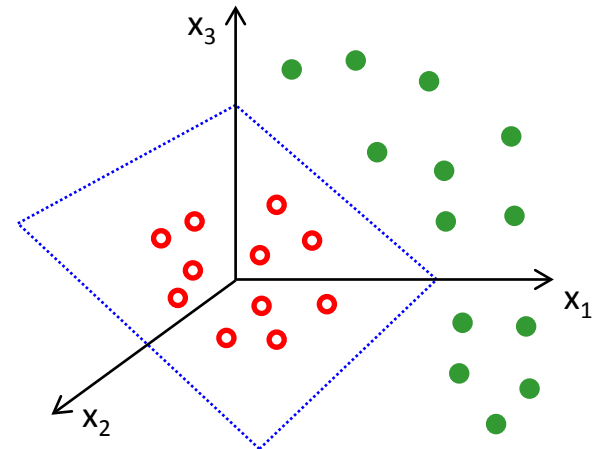


# Motivace

2-rozměrný prostor



3-rozměrný prostor



Hranice je nadplocha o rozměru o jedna menší než je rozměr prostoru

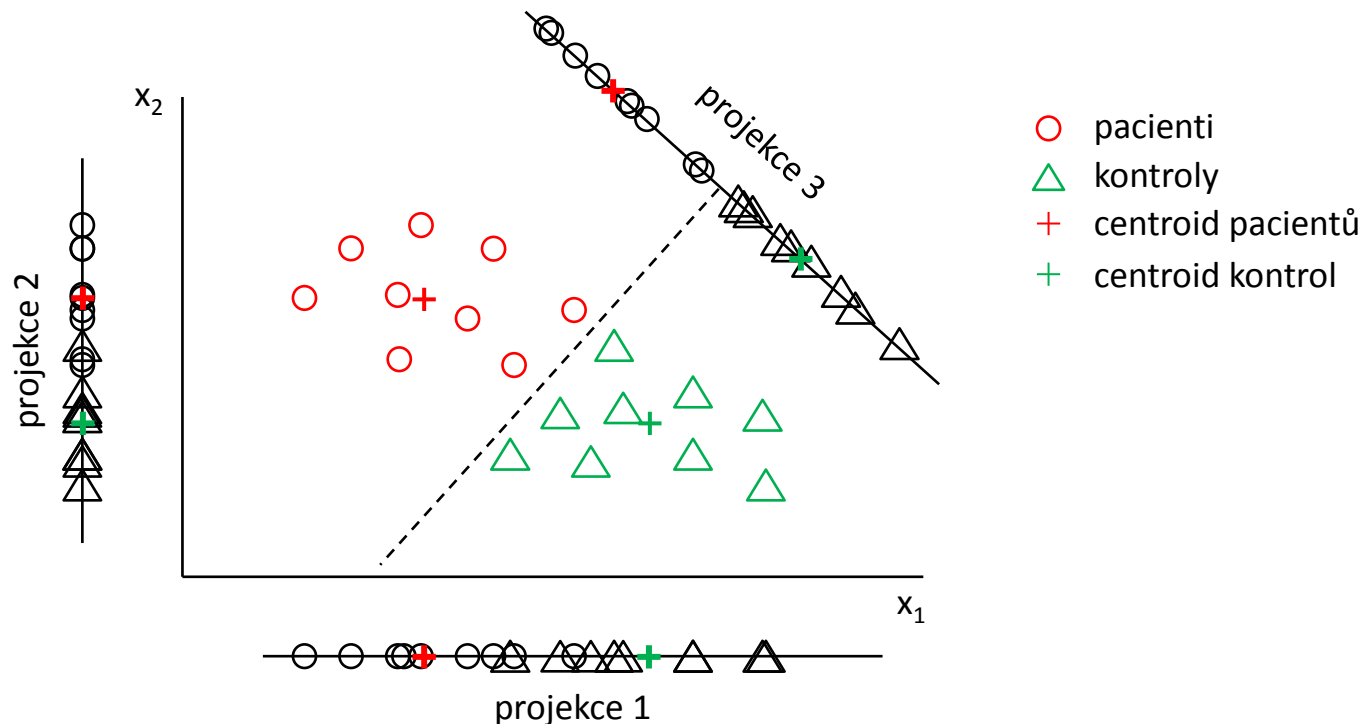
- ve 2-rozměrném prostoru je hranicí křivka (v lineárním případě přímka)
- v 3-rozměrném prostoru plocha (v lineárním případě rovina)

Hranice je tedy dána rovnicí:  $h(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + w_0 = 0$

Výpočet hranice různými metodami (např. Fisherova LDA, SVM apod. – viz dále)

# Fisherova lineární diskriminace (FLDA)

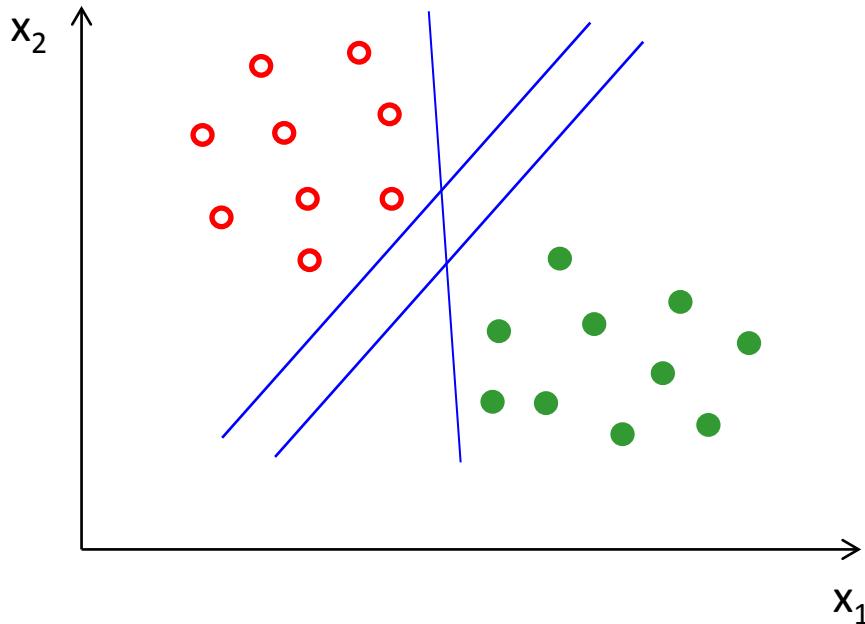
- použití pro lineární klasifikaci
- **princip:** transformace do jednorozměrného prostoru tak, aby se třídy od sebe maximálně oddělily (maximalizace vzdálenosti skupin a minimalizace variability uvnitř skupin)



- **předpoklad:** vícerozměrné normální rozdělení u jednotlivých skupin

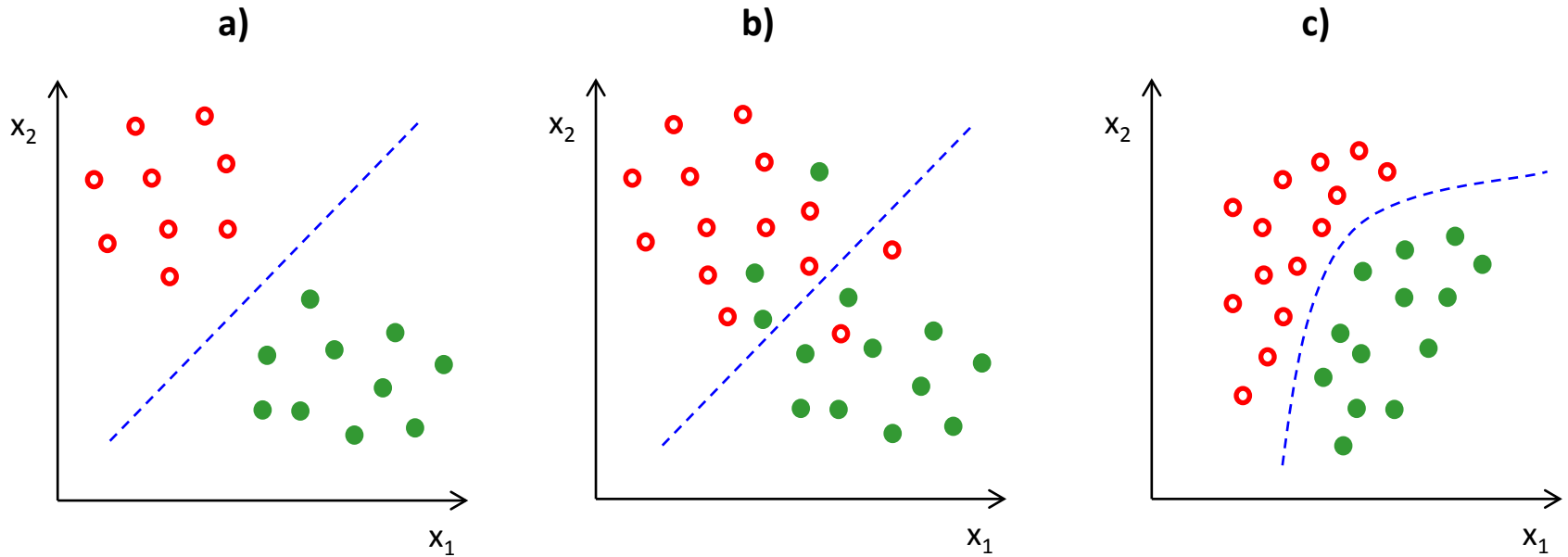
# Metoda podpůrných vektorů (SVM)

- použití pro lineární i nelineární klasifikaci
- **princip:** proložení klasifikační hranice (nadroviny) tak, aby byla v co největší vzdálenosti od subjektů z obou tříd



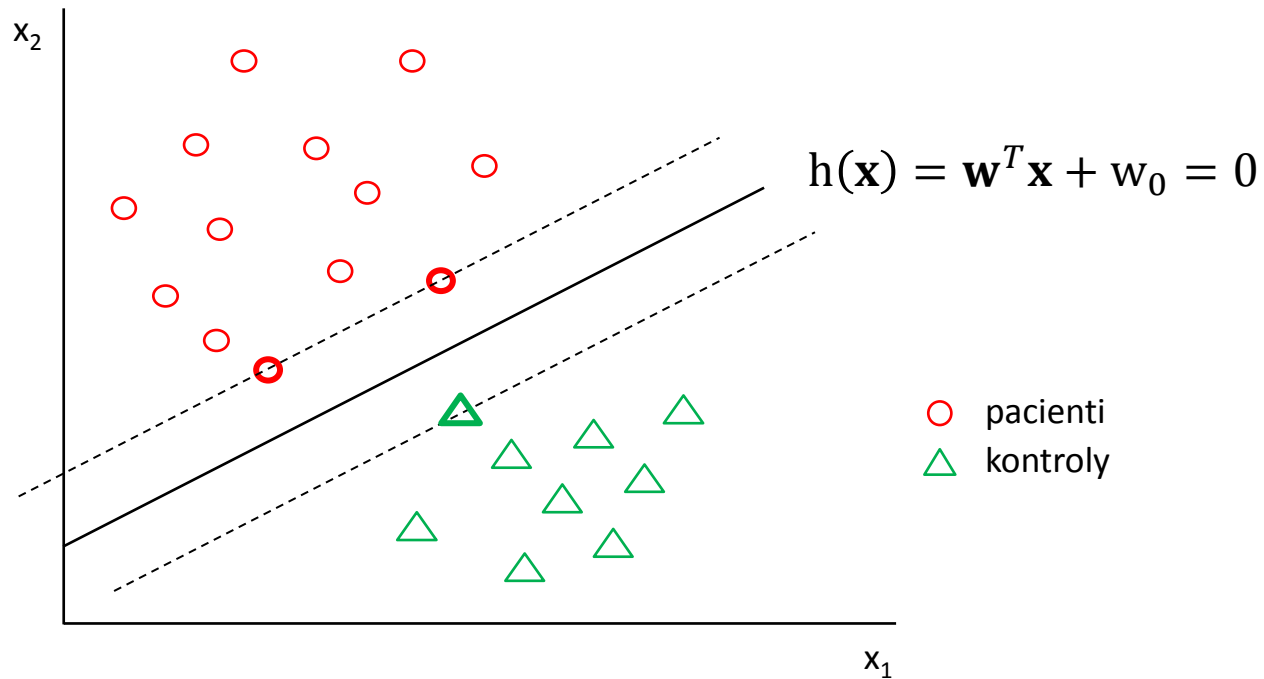
- **výhoda oproti FLDA:** nemá předpoklady o rozdělení dat
- **nevýhoda:** vyžaduje stanovení parametrů (např. C) a případně i typu jádra

# Metoda podpůrných vektorů (SVM) - varianty



- varianty SVM dle typu vstupních dat:
  - a) lineární verze metody podpůrných vektorů pro lineárně separabilní třídy (anglicky *maximal margin classifier*)
  - b) lineární verze metody podpůrných vektorů pro lineárně neseparabilní třídy (anglicky *support vector classifier*)
  - c) nelineární verze metody podpůrných vektorů (anglicky *support vector machine*)

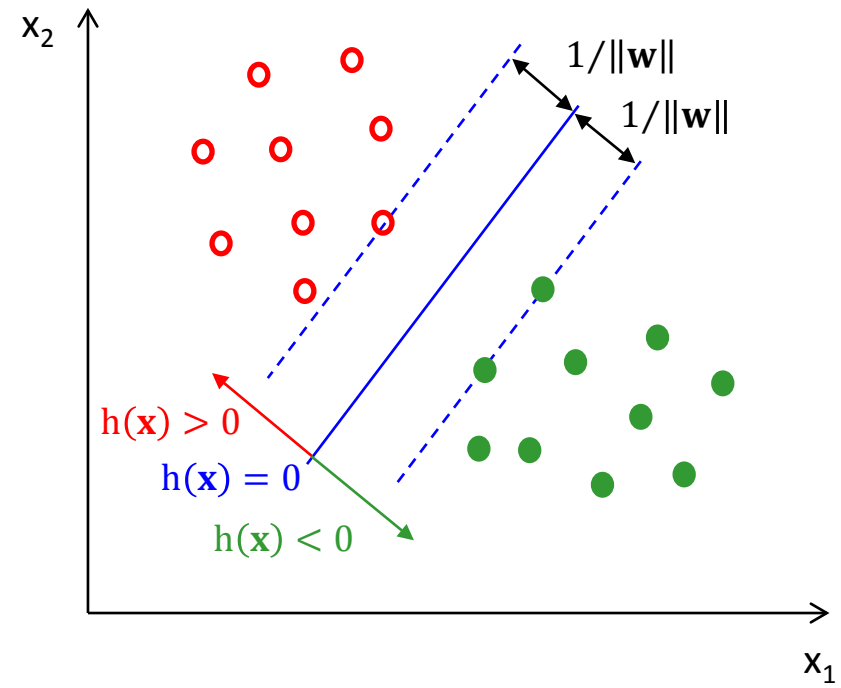
# Lineární SVM – lineárně separabilní třídy



- proložení klasifikační hranice (nadroviny) tak, aby byla v co největší vzdálenosti od subjektů z obou tříd → tzn. aby byl okolo hranice co nejširší pruh bez bodů
- na popis hranice (nadroviny) stačí pouze nejbližší body, kterých je obvykle málo a nazývají se **podpůrné vektory** (support vectors)

# Lineární SVM – lineárně separabilní třídy

- hranice:  $h(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + w_0 = 0$  (kde  $\mathbf{w}$  a  $w_0$  je orientace a poloha hranice)
- klasifikace subjektu  $\mathbf{x}$  do třídy  $\omega_D$ , resp.  $\omega_H$ , bude dána tím, jestli je výraz  $\mathbf{w}^T \mathbf{x} + w_0$  větší, resp. menší, než 0
- vzdálenost jakéhokoliv bodu od klasifikační hranice je:  $d = \frac{|\mathbf{w}^T \mathbf{x} + w_0|}{\|\mathbf{w}\|}$ , kde  $\|\mathbf{w}\|$  je velikost vektoru  $\mathbf{w}$
- pro nejbližší bod  $\mathbf{x}_i$  ze třídy  $\omega_D$  zvolíme hodnotu výrazu  $\mathbf{w}^T \mathbf{x}_i + w_0$  rovnu +1
- pro nejbližší bod  $\mathbf{x}_j$  ze třídy  $\omega_H$  zvolíme hodnota výrazu  $\mathbf{w}^T \mathbf{x}_j + w_0$  rovnu -1
- pak na každé straně od dělicí přímky máme toleranční pásmo o šířce  $\frac{1}{\|\mathbf{w}\|}$ , ve kterém se nenachází žádný bod





# Lineární SVM – lineárně separabilní třídy

- Pro všechny body z trénovací množiny platí:

$$\mathbf{w}^T \mathbf{x} + w_0 \geq 1 \quad \text{pro všechna } \mathbf{x} \text{ z } \omega_D,$$

$$\mathbf{w}^T \mathbf{x} + w_0 \leq -1 \quad \text{pro všechna } \mathbf{x} \text{ z } \omega_H,$$

- což můžeme stručněji zapsat jako

$$\delta_{x_k} (\mathbf{w}^T \mathbf{x}_k + w_0) \geq 1, \text{ pro } k=1, \dots, N,$$

- kde  $\delta_{x_k} = 1$  pro  $\mathbf{x}_k$  ze třídy  $\omega_D$  a  $\delta_{x_k} = -1$  pro  $\mathbf{x}_k$  ze třídy  $\omega_H$

- hledáme takové hodnoty  $\mathbf{w}$  a  $w_0$ , aby byla celková šířka tolerančního pásma

$$\frac{1}{\|\mathbf{w}\|} + \frac{1}{\|\mathbf{w}\|} = \frac{2}{\|\mathbf{w}\|} \text{ co největší}$$

- hledat maximum funkce  $\frac{2}{\|\mathbf{w}\|}$  je to stejné, jako hledat minimum funkce  $\frac{\|\mathbf{w}\|}{2}$  a toto minimum se nezmění, když kladnou hodnotu v čitateli umocníme na druhou (což nám zjednoduší výpočty), takže dostáváme následující kritériální funkci, jejíž hodnotu se snažíme minimalizovat:

$$J(\mathbf{w}, w_0) = \frac{\|\mathbf{w}\|^2}{2}$$

→ řešení pomocí metody Lagrangeových součinitelů

# Lineární SVM – metoda Lagrangeova součinitele

- Chceme minimalizovat výraz  $J(\mathbf{w}, w_0) = \frac{\|\mathbf{w}\|^2}{2}$  za podmínek  $\delta_{x_k}(\mathbf{w}^T \mathbf{x}_k + w_0) \geq 1$
- Zavedeme vektor Lagrangeových součinitelů  $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, \dots, \lambda_N]$ , kde  $\lambda_k \geq 0$  a pomocí nich vyjádříme optimalizovanou funkci jako:

$$L(\mathbf{w}, w_0, \boldsymbol{\lambda}) = \frac{\|\mathbf{w}\|^2}{2} - \sum_{k=1}^N \lambda_k [\delta_{x_k}(\mathbf{w}^T \mathbf{x}_k + w_0) - 1]$$

za podmínek  $\lambda_k [\delta_{x_k}(\mathbf{w}^T \mathbf{x}_k + w_0) - 1] = 0, k = 1, 2, \dots, N$

- tuto Lagrangeovu funkci zderivujeme podle proměnných  $\mathbf{w}$  a  $w_0$  a derivace položíme rovny nule  $\rightarrow$  po dalších úpravách dostaneme soustavu nelineárních rovnic:

$$\mathbf{w} = \sum_{k=1}^N \lambda_k \delta_{x_k} \mathbf{x}_k$$

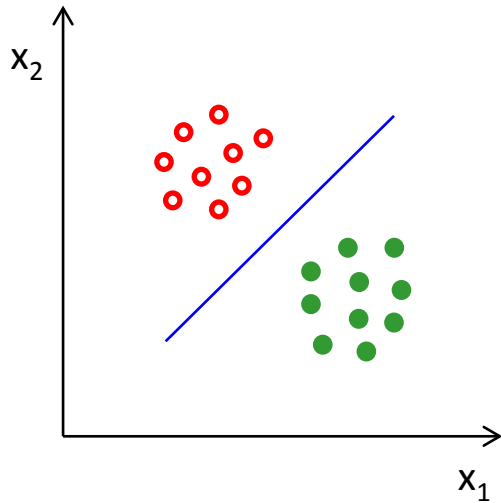
$$\sum_{k=1}^N \lambda_k \delta_{x_k} = 0$$

$\rightarrow$  patrné, že pro výpočet orientace hranice důležité jen ty body, pro které platí  $\lambda_k > 0$

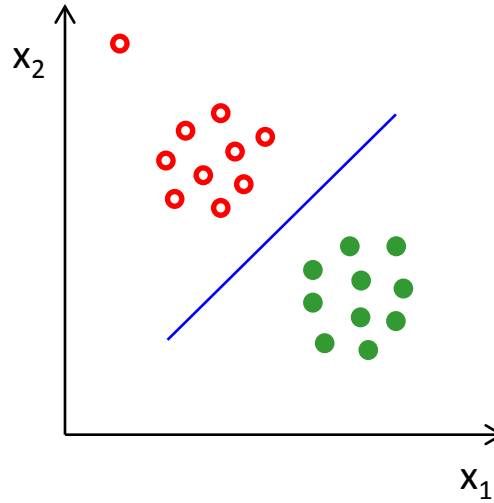
$\rightarrow$  každý takový bod musí splňovat podmínku výše, tedy  $\delta_{x_k}(\mathbf{w}^T \mathbf{x}_k + w_0) - 1 = 0 \rightarrow$  tedy musí ležet přesně na hranici tolerančního pásma

$\rightarrow$  takovým bodům říkáme podpůrné vektory a jen na nich závisí umístění a orientace dělící přímky

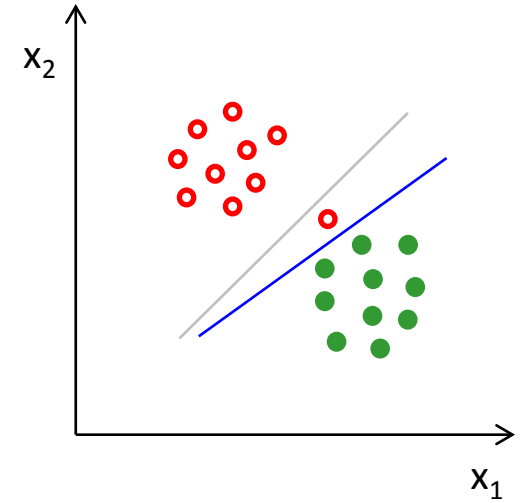
# Lineární SVM – vliv odlehlých hodnot



klasifikace v případě dat  
neobsahujících odlehlé  
hodnoty



klasifikace v případě  
odlehlé hodnoty, která  
není podpurným  
vektorem (poloha  
klasifikační hranice se  
nezmění)



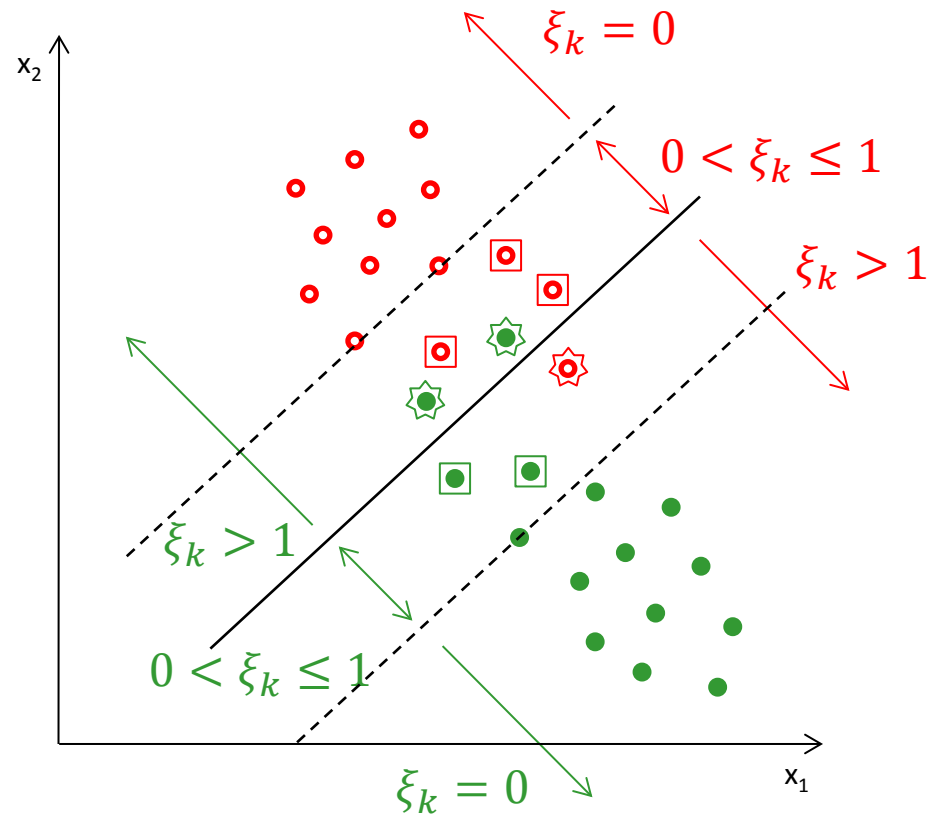
klasifikace v případě  
odlehlé hodnoty, která je  
podpurným vektorem  
(poloha hranice se změní)

→ lepší použít lineární SVM  
pro lineárně neseparabilní  
třídy, kterou tato odlehlá  
hodnota téměř neovlivní

# Lineární SVM – lineárně neseparabilní třídy

- zavedeme relaxační proměnné  $\xi_k \geq 0$  vyjadřující, jak moc každý bod porušuje podmínku  $\delta_{x_k}(\mathbf{w}^T \mathbf{x}_k + w_0) \geq 1$
- 3 situace:
  1. objekt leží **vně** pásma a je **správně** klasifikován:  $\xi_k = 0$
  2. objekt leží **uvnitř** pásma a je **správně** klasifikován (body s čtverečky):  $0 < \xi_k \leq 1$
  3. objekt leží na **opačné straně** hranice a je **chybně** klasifikován (body s hvězdičkami):  $\xi_k > 1$
- podmínky jsou pak ve tvaru:

$$\delta_{x_k}(\mathbf{w}^T \mathbf{x}_k + w_0) \geq 1 - \xi_k$$



# Lineární SVM – lineárně neseparabilní třídy

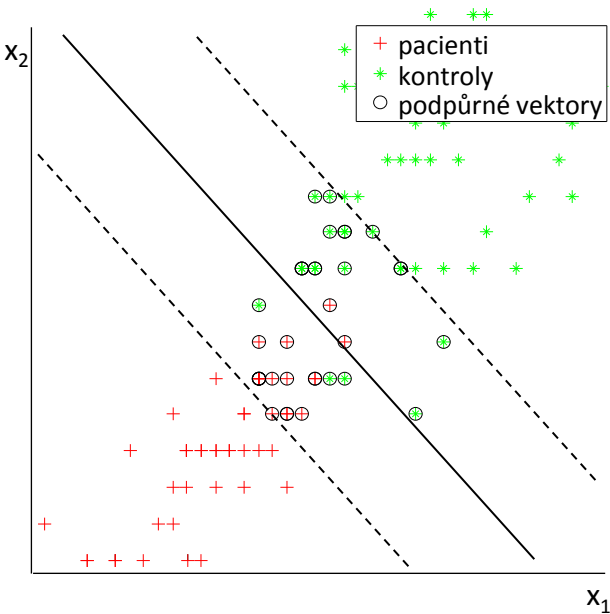
- když chceme najít hranici poskytující co nejrobustnější klasifikaci, musíme se snažit:
  - maximalizovat šířku tolerančního pásma
  - minimalizovat počet subjektů z trénovací množiny, které leží v tolerančním pásmu nebo jsou dokonce špatně klasifikovány (tj. těch, pro které  $\xi_k > 0$ )
- to můžeme vyjádřit jako minimalizaci kriteriální funkce:

$$J(\mathbf{w}, w_0, \xi) = \frac{\|\mathbf{w}\|^2}{2} + C \sum_{k=1}^N \xi_k$$

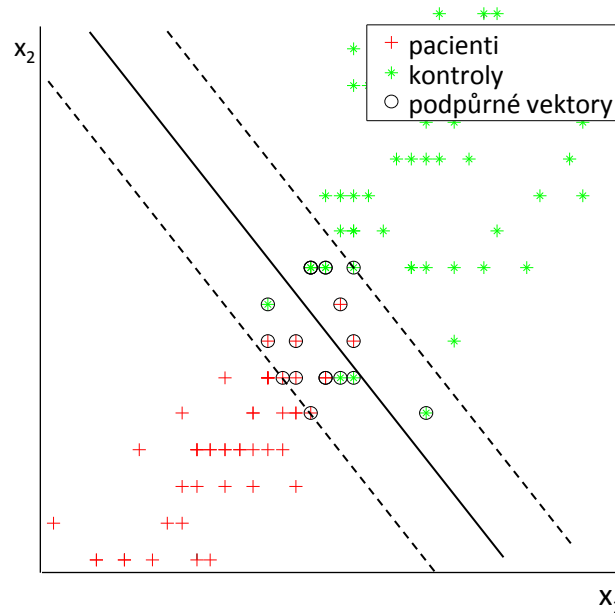
- kde  $C$  vyjadřuje poměr vlivu obou členů kriteriální funkce:
  - **pro nízké hodnoty  $C$**  bude toleranční pásmo širší a počet trénovaných subjektů v tolerančním pásmu a počet chybně klasifikovaných trénovacích subjektů bude vyšší
  - **pro vysoké hodnoty  $C$**  bude toleranční pásmo užší, ale počet trénovaných subjektů v tolerančním pásmu a počet chybně klasifikovaných trénovacích subjektů bude nižší
- řešíme opět pomocí metody Lagrangeova součinitele

# SVM – vliv parametru C

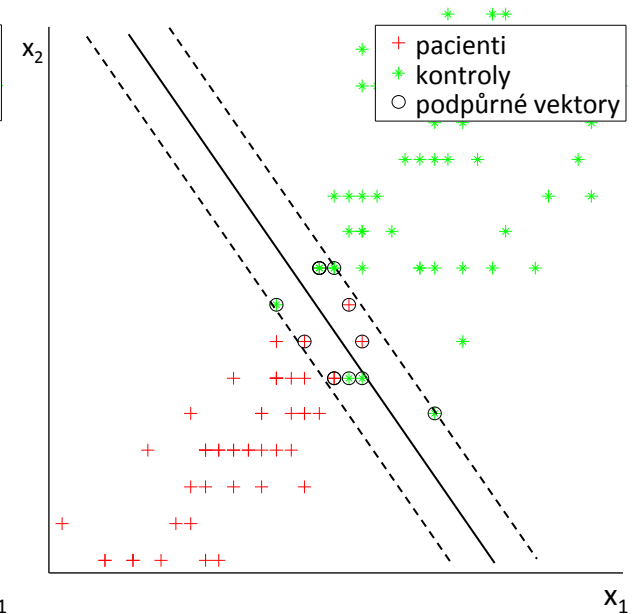
**C = 0.1**



**C = 1**



**C = 10**



- **pro nízké hodnoty C** – toleranční pásmo širší, ale počet trénovaných subjektů v tolerančním pásmu a počet chybně klasifikovaných trénovacích subjektů vyšší
- **pro vysoké hodnoty C** – toleranční pásmo užší, ale počet trénovaných subjektů v tolerančním pásmu a počet chybně klasifikovaných trénovacích subjektů nižší
- zpravidla nevíme, jaká hodnota parametru C pro data nejvhodnější → volba C podle křížové validace

# Lineární SVM – metoda Lagrangeova součinitele

- Chceme minimalizovat  $J(\mathbf{w}, w_0) = \frac{\|\mathbf{w}\|^2}{2} + C \sum_{k=1}^N \xi_k$  za podmínek  $\delta_{x_k}(\mathbf{w}^T \mathbf{x}_k + w_0) \geq 1 - \xi_k$
- Zavedeme vektor Lagrangeových součinitelů  $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, \dots, \lambda_N]$ , kde  $\lambda_k \geq 0$ , a pomocí nich vyjádříme optimalizovanou funkci jako:

$$L(\mathbf{w}, w_0, \boldsymbol{\xi}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \frac{\|\mathbf{w}\|^2}{2} + C \sum_{k=1}^N \xi_k - \sum_{k=1}^N \lambda_k [\delta_{x_k}(\mathbf{w}^T \mathbf{x}_k + w_0) - 1 + \xi_k] - \sum_{k=1}^N \mu_k \xi_k$$

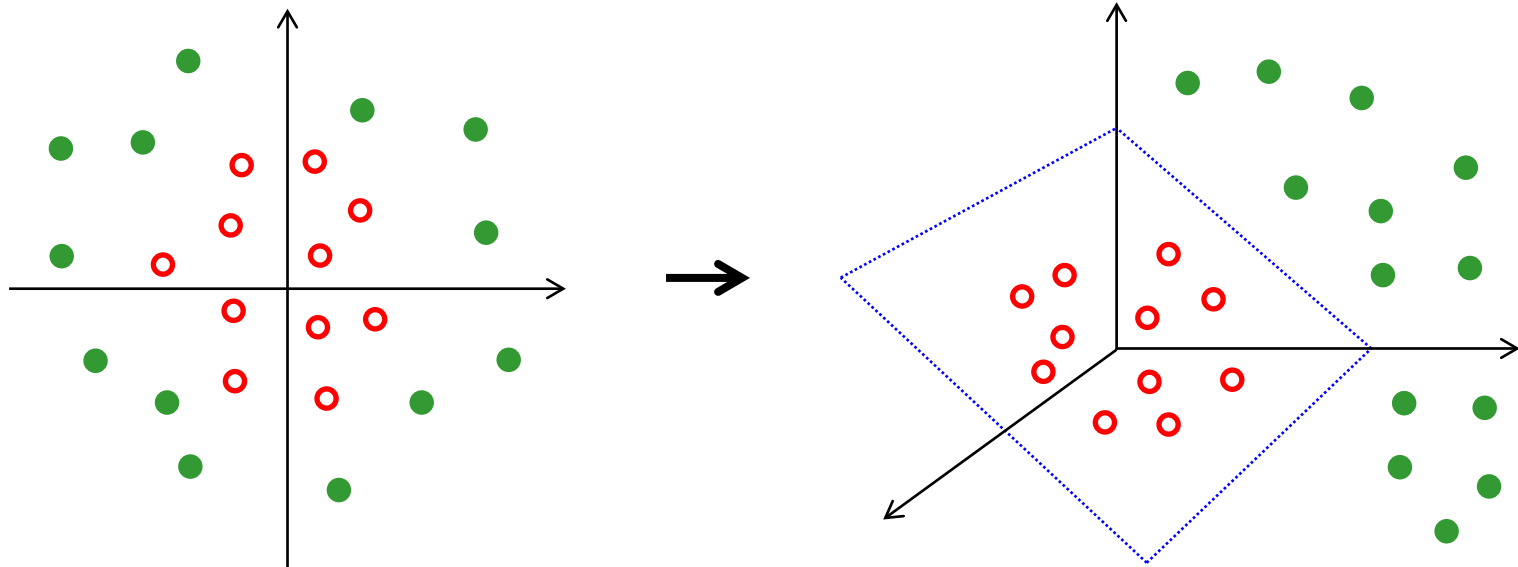
- za podmínek  $\lambda_k [\delta_{x_k}(\mathbf{w}^T \mathbf{x}_k + w_0) - 1 + \xi_k] = 0$  a  $\mu_k \xi_k \geq 0$ , pro  $k = 1, \dots, N$ .
- tuto Lagrangeovu funkci zderivujeme podle proměnných  $\mathbf{w}$ ,  $w_0$  a  $\boldsymbol{\xi}$  a derivace položíme rovny nule  $\rightarrow$  po dalších úpravách dostaneme soustavu nelin. rovnic:

$$\left. \begin{aligned} \mathbf{w} &= \sum_{k=1}^N \lambda_k \delta_{x_k} \mathbf{x}_k & \lambda_k + \mu_k &= C, \\ & & \lambda_k [\delta_{x_k}(\mathbf{w}^T \mathbf{x}_k + w_0) - 1 + \xi_k] &= 0, \\ & & \mu_k \xi_k &= 0. \end{aligned} \right\} \text{ pro } k = 1, \dots, N$$

$$\sum_{k=1}^N \lambda_k \delta_{x_k} = 0$$

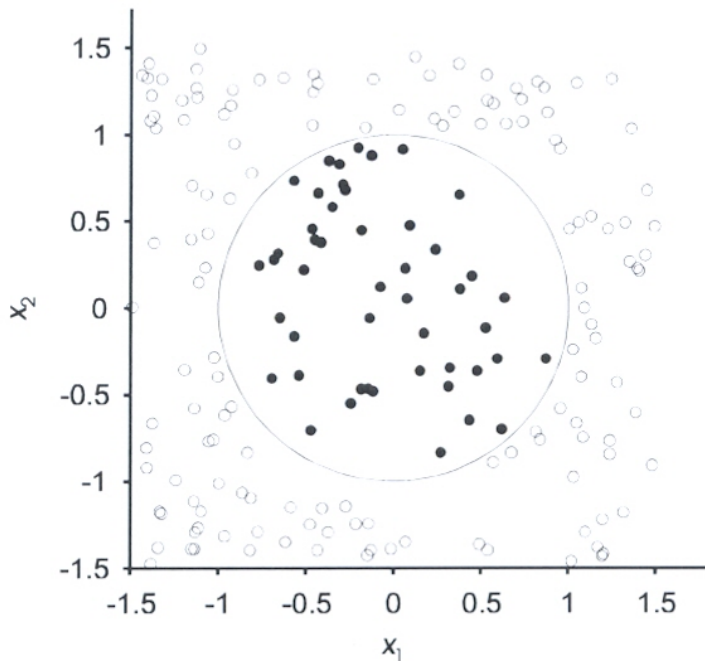
# Nelineární SVM

- zobrazíme původní  $p$ -rozměrný obrazový prostor nelineární transformací do nového  $m$ -rozměrného prostoru tak, aby v novém prostoru byly klasifikační třídy lineárně separabilní

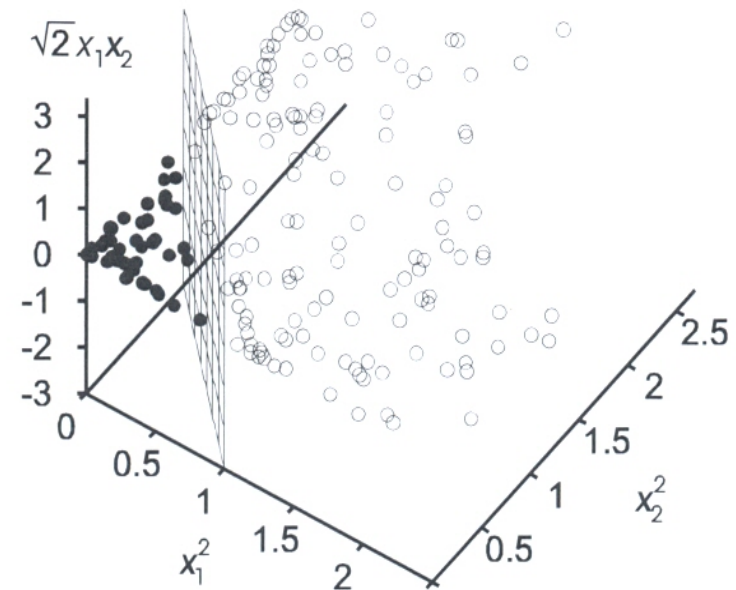




# Nelineární SVM – ukázka 2

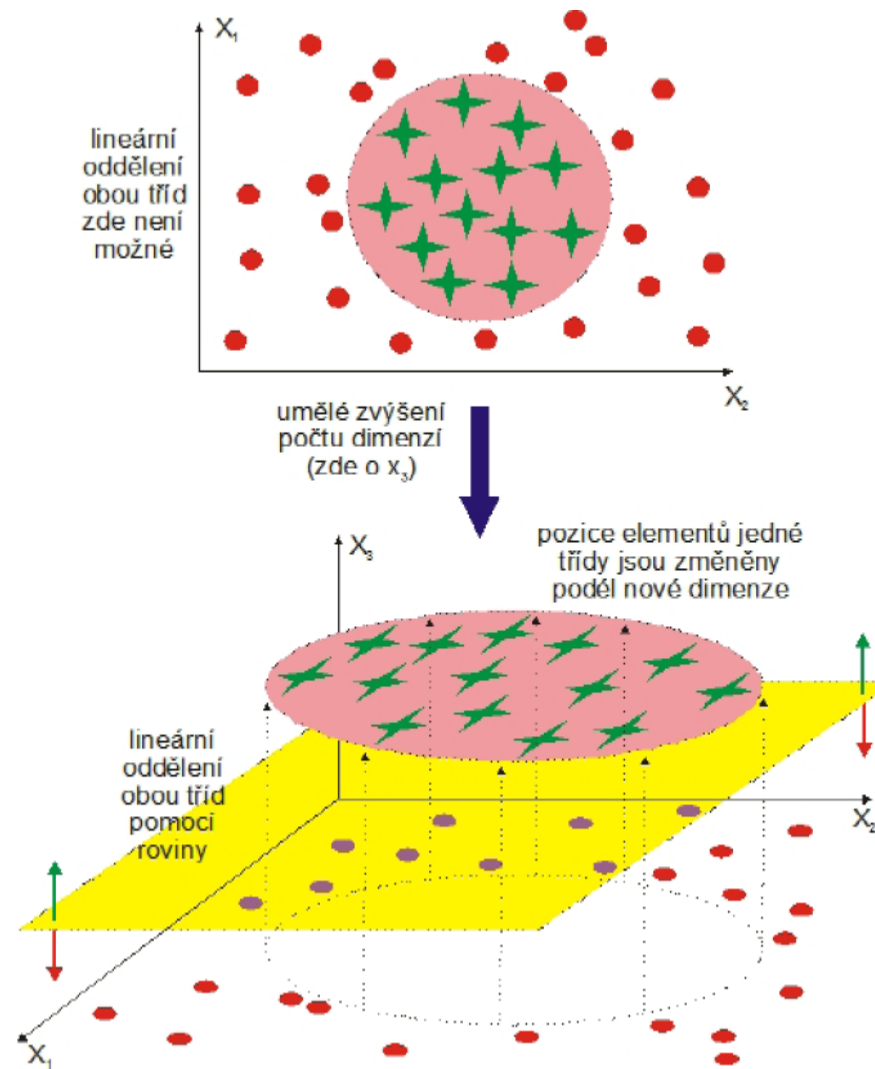


dvourozměrný prostor s  
oddělovací hranicí ve tvaru  
 $x_1^2 + x_2^2 \leq 1$



tatáž situace zobrazená do  
trojrozměrného prostoru  
( $x_1^2, x_2^2, \sqrt{2}x_1x_2$ ) – **kruhová  
hranice se stane lineární**

# Nelineární SVM – ukázka 3



# Nelineární SVM

- transformace do nového prostoru může proběhnout navýšením počtu proměnných (např. přidáním kvadratických forem původních proměnných, tedy soubor pak bude obsahovat proměnné  $\mathbf{x}_1, \mathbf{x}_1^2, \mathbf{x}_2, \mathbf{x}_2^2, \dots, \mathbf{x}_p, \mathbf{x}_p^2$ )
- tento přístup však výpočetně náročný  $\rightarrow$  použití jader (*kernels*)
- u lineárního SVM pro lineárně separabilní i neseperabilní třídy lze Lagrangeovu funkci přepsat do podoby

$$L(\mathbf{w}, w_0, \boldsymbol{\lambda}) = \sum_{k=1}^N \lambda_k - \frac{1}{2} \sum_{i,j} \lambda_i \lambda_j \delta_{x_i} \delta_{x_j} \mathbf{x}_i^T \mathbf{x}_j$$

- kde si skalární součin  $\mathbf{x}_i^T \mathbf{x}_j$  můžeme zapsat obecně jako  $K(\mathbf{x}_i, \mathbf{x}_j)$ , kde  $K$  je nějaká funkce, kterou nazveme jádro

- typy jader:

- lineární jádro:  $K(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i^T \mathbf{x}_j \rightarrow$  lineární SVM

- polynomiální jádro stupně  $d$ :  $K(\mathbf{x}_i, \mathbf{x}_j) = (1 + \mathbf{x}_i^T \mathbf{x}_j)^d = (1 + \sum_{l=1}^p x_{il} x_{jl})^d$

- radiální bázové jádro:  $K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\gamma \sum_{l=1}^p (x_{il} - x_{jl})^2\right)$

- atd.

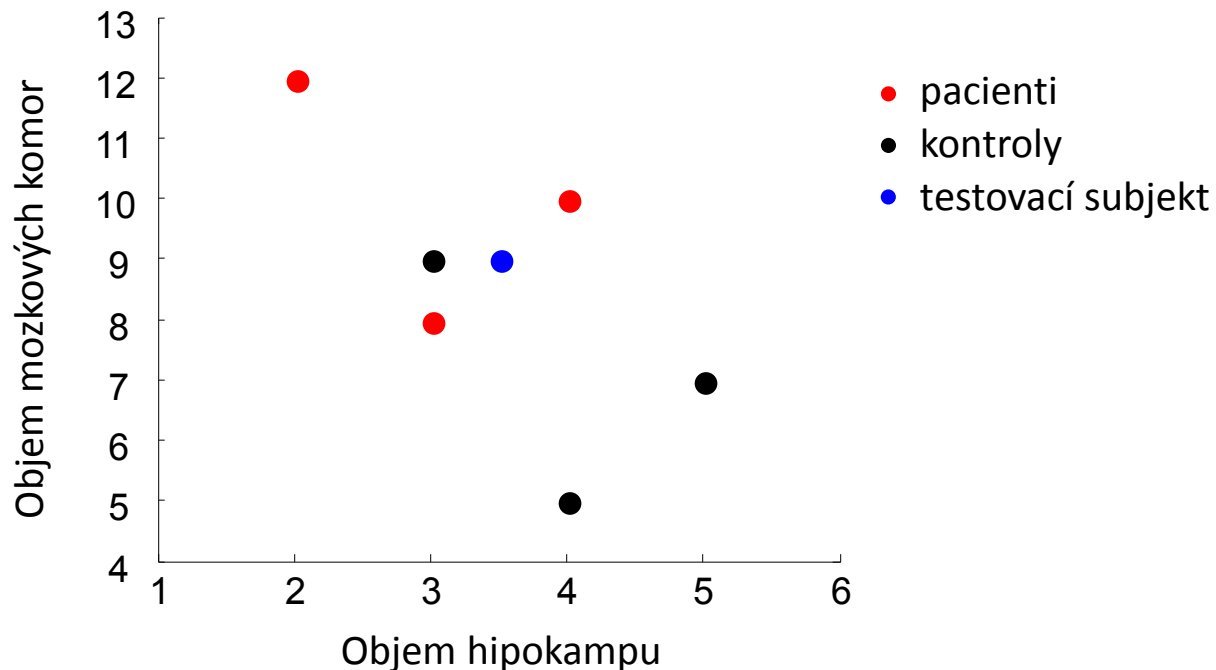
nelineární  
SVM

# Příklad

**Příklad:** Bylo provedeno měření objemu hipokampu a mozkových komor

(v  $\text{cm}^3$ ) u 3 pacientů se schizofrenií a 3 kontrol:  $\mathbf{X}_D = \begin{bmatrix} 2 & 12 \\ 4 & 10 \\ 3 & 8 \end{bmatrix}$ ,  $\mathbf{X}_H = \begin{bmatrix} 5 & 7 \\ 3 & 9 \\ 4 & 5 \end{bmatrix}$ .

Určete, zda testovací subjekt  $\mathbf{x}_0 = [3,5 \quad 9]$  patří do skupiny pacientů či kontrolních subjektů pomocí metody podpůrných vektorů.



# Příklad – řešení pro parametr $C = 1$

- výsledkem jsou hodnoty  $\lambda = \left[0, \frac{1}{2}, 1, \frac{1}{2}, 1, 0\right]$
- podpůrnými vektory jsou tedy body  $\mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4$  a  $\mathbf{x}_5$ , protože jim příslušející  $\lambda_2, \lambda_3, \lambda_4$  a  $\lambda_5$  jsou nenulové. Vypočítáme orientaci hranice:

$$\begin{aligned}\mathbf{w} &= \sum_{k=1}^6 \lambda_k \delta_{x_k} \mathbf{x}_k = \frac{1}{2} \mathbf{x}_2 + \mathbf{x}_3 - \frac{1}{2} \mathbf{x}_4 - \mathbf{x}_5 = \frac{1}{2} \begin{bmatrix} 4 \\ 10 \end{bmatrix} + \begin{bmatrix} 3 \\ 8 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 5 \\ 7 \end{bmatrix} - \begin{bmatrix} 3 \\ 9 \end{bmatrix} \\ &= \begin{bmatrix} -1/2 \\ 1/2 \end{bmatrix}\end{aligned}$$

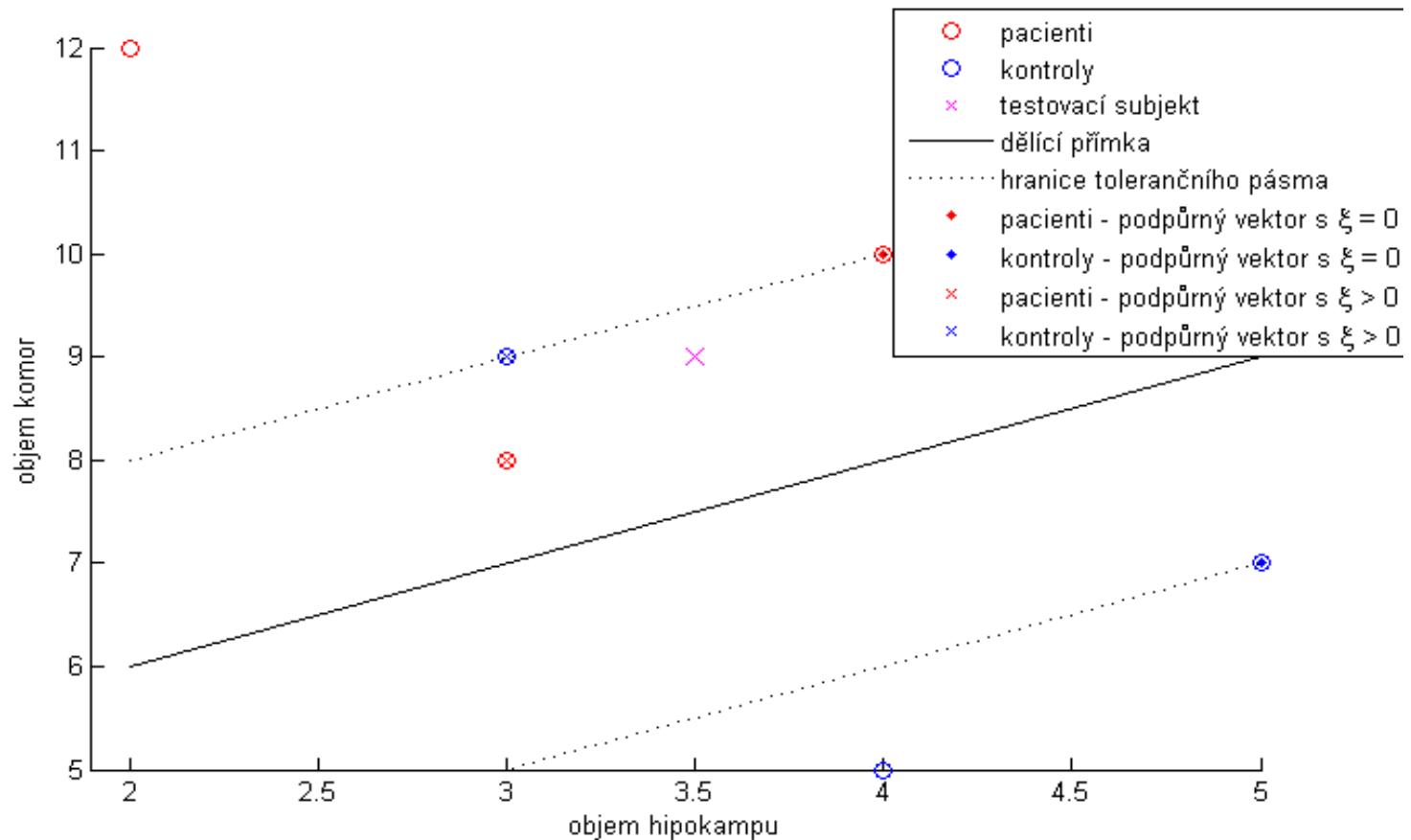
- Pokud zvolíme takové  $\mathbf{x}_k$ , pro které platí  $0 < \lambda_k < C$ , tak podle vztahu  $\lambda_k + \mu_k = C$  musí být  $\mu_k > 0$  a odtud podle vztahu  $\mu_k \xi_k = 0$  plyne, že  $\xi_k = 0$ . Vzorec se tak zjednoduší na  $\delta_{x_k} (\mathbf{w}^T \mathbf{x}_k + w_0) = 1$ . Tedy například pro  $\mathbf{x}_2$  ( $0 < \lambda_2 = \frac{1}{2} < C = 1$ ):

$$\mathbf{w}^T \mathbf{x}_2 + w_0 = 1 \Rightarrow w_0 = 1 - \begin{bmatrix} -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 4 \\ 10 \end{bmatrix} = 1 - 3 = -2$$

- hranice je tedy dána rovnicí:  $\mathbf{w}^T \mathbf{x} + w_0 = \begin{bmatrix} -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \mathbf{x} - 2$

# Příklad – řešení pro parametr $C = 1$

- hranice je dána rovnicí:  $\mathbf{w}^T \mathbf{x} + w_0 = \left[ -\frac{1}{2} \quad \frac{1}{2} \right] \mathbf{x} - 2$



# Příklad – řešení pro parametr $C = 1$

- můžeme klasifikovat subjekt  $\mathbf{x} = [3,5 \quad 9]$ :

$$\mathbf{w}^T \mathbf{x} + w_0 = \begin{bmatrix} -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 3,5 \\ 9 \end{bmatrix} - 2 = -1,75 + 4,5 - 2 = 0,75$$

- protože  $0,75 > 0$ , testovací subjekt bude zařazen do třídy pacientů
- ověříme, že natrénovaný klasifikátor zařadí subjekty z trénovací množiny tak, jak to odpovídá situaci na obrázku; tj. správně subjekty  $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4$  a chybně subjekt  $\mathbf{x}_5$ :

$$\mathbf{w}^T \mathbf{x}_1 + w_0 = \begin{bmatrix} -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 2 \\ 12 \end{bmatrix} - 2 = -1 + 6 - 2 = 3$$

$$\mathbf{w}^T \mathbf{x}_2 + w_0 = \begin{bmatrix} -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 4 \\ 10 \end{bmatrix} - 2 = -2 + 5 - 2 = 1$$

$$\mathbf{w}^T \mathbf{x}_3 + w_0 = \begin{bmatrix} -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 3 \\ 8 \end{bmatrix} - 2 = -1,5 + 4 - 2 = 0,5$$

$$\mathbf{w}^T \mathbf{x}_4 + w_0 = \begin{bmatrix} -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 5 \\ 7 \end{bmatrix} - 2 = -2,5 + 3,5 - 2 = -1$$

$$\mathbf{w}^T \mathbf{x}_5 + w_0 = \begin{bmatrix} -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 3 \\ 9 \end{bmatrix} - 2 = -1,5 + 4,5 - 2 = 1$$

$$\mathbf{w}^T \mathbf{x}_6 + w_0 = \begin{bmatrix} -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 4 \\ 5 \end{bmatrix} - 2 = -2 + 2,5 - 2 = -1,5$$

# Příklad – řešení pro parametr $C = 10$

- výsledkem jsou hodnoty  
 $\lambda = [0, 3,6371, 8,7629, 2,0371, 10, 0,3629]$
- podpůrnými vektory jsou tedy body  $\mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5$  a  $\mathbf{x}_6$ , protože jim příslušející  $\lambda_2, \lambda_3, \lambda_4, \lambda_5$  a  $\lambda_6$  jsou nenulové. Vypočítáme orientaci hranice:

$$\begin{aligned}\mathbf{w} &= \sum_{k=1}^6 \lambda_k \delta_{x_k} \mathbf{x}_k = 3,6371\mathbf{x}_2 + 8,7629\mathbf{x}_3 - 2,0371\mathbf{x}_4 - 10\mathbf{x}_5 - 0,3629\mathbf{x}_6 \\ &= 3,6371 \begin{bmatrix} 4 \\ 10 \end{bmatrix} + 8,7629 \begin{bmatrix} 3 \\ 8 \end{bmatrix} - 2,0371 \begin{bmatrix} 5 \\ 7 \end{bmatrix} - 10 \begin{bmatrix} 3 \\ 9 \end{bmatrix} - 0,3629 \begin{bmatrix} 4 \\ 5 \end{bmatrix} \\ &= \begin{bmatrix} -4/5 \\ 2/5 \end{bmatrix}\end{aligned}$$

- Polohu dělicí přímky určíme opět ze  $\delta_{x_k} (\mathbf{w}^T \mathbf{x}_k + w_0) = 1$ . Tedy například pro  $\mathbf{x}_2$  ( $0 < \lambda_2 = 3,6371 < C = 10$ ):

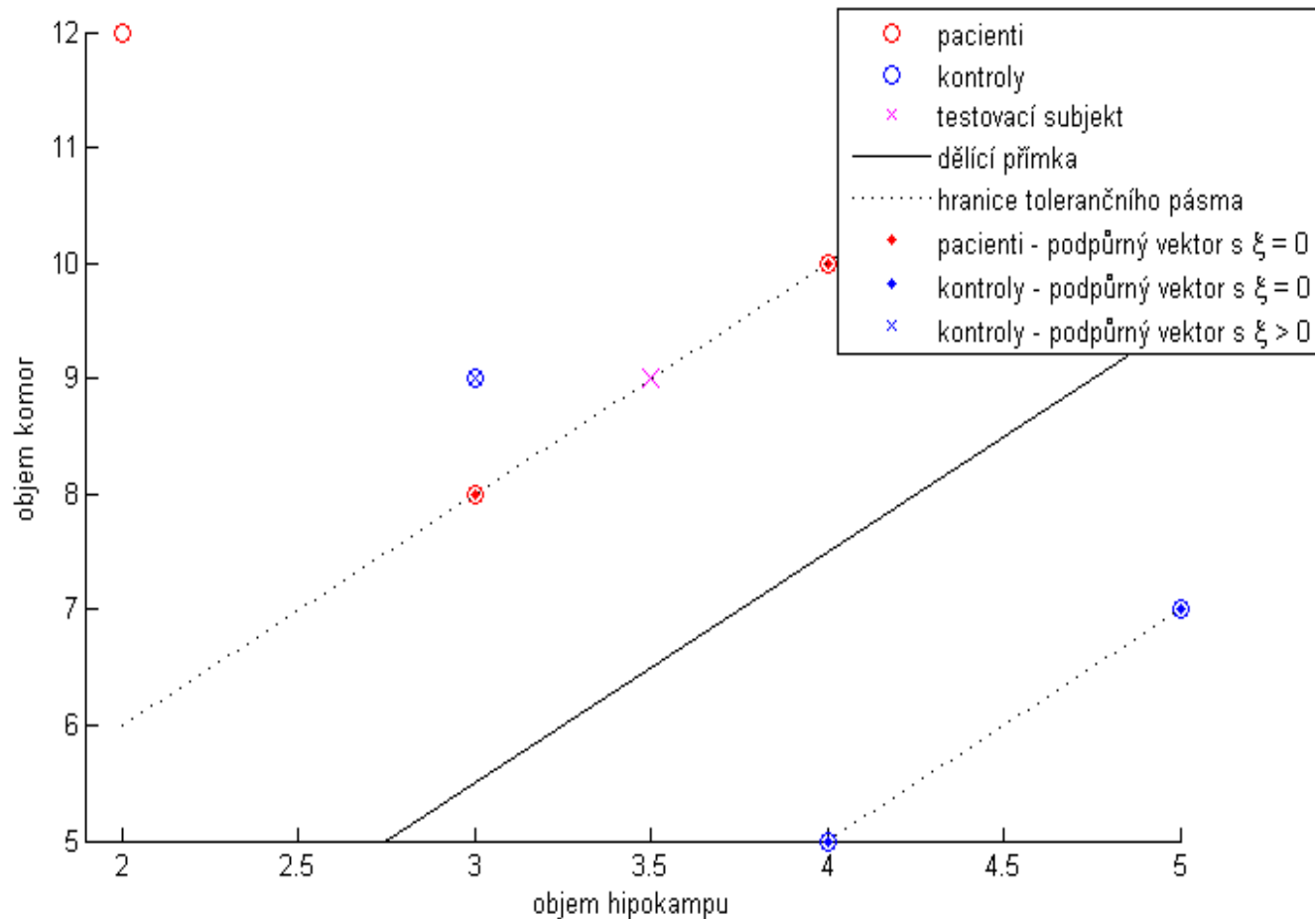
$$\mathbf{w}^T \mathbf{x}_2 + w_0 = 1 \Rightarrow w_0 = 1 - \begin{bmatrix} -\frac{4}{5} & \frac{2}{5} \end{bmatrix} \begin{bmatrix} 4 \\ 10 \end{bmatrix} = 1 - \frac{4}{5} = \frac{1}{5}$$

- hranice je tedy dána rovnicí:  $\mathbf{w}^T \mathbf{x} + w_0 = \begin{bmatrix} -\frac{4}{5} & \frac{2}{5} \end{bmatrix} \mathbf{x} + \frac{1}{5}$



# Příklad – řešení pro parametr $C = 10$

- hranice je tedy dána rovnicí:  $\mathbf{w}^T \mathbf{x} + w_0 = \left[ -\frac{4}{5} \quad \frac{2}{5} \right] \mathbf{x} + \frac{1}{5}$



# Příklad – řešení pro parametr $C = 10$

- můžeme klasifikovat subjekt  $\mathbf{x} = [3,5 \quad 9]$ :

$$\mathbf{w}^T \mathbf{x} + w_0 = \begin{bmatrix} -\frac{4}{5} & \frac{2}{5} \end{bmatrix} \begin{bmatrix} 3,5 \\ 9 \end{bmatrix} + \frac{1}{5} = -2,8 + 3,6 + 0,2 = 1$$

- protože  $1 > 0$ , testovací subjekt bude zařazen do třídy pacientů
- ověříme, že natrénovaný klasifikátor zařadí subjekty z trénovací množiny tak, jak to odpovídá situaci na obrázku; tj. správně subjekty  $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4$  a  $\mathbf{x}_6$  a chybně subjekt  $\mathbf{x}_5$ :

$$\mathbf{w}^T \mathbf{x}_1 + w_0 = \begin{bmatrix} -\frac{4}{5} & \frac{2}{5} \end{bmatrix} \begin{bmatrix} 2 \\ 12 \end{bmatrix} + \frac{1}{5} = -\frac{8}{5} + \frac{24}{5} + \frac{1}{5} = \frac{17}{5} = 3,4$$

$$\mathbf{w}^T \mathbf{x}_2 + w_0 = \begin{bmatrix} -\frac{4}{5} & \frac{2}{5} \end{bmatrix} \begin{bmatrix} 4 \\ 10 \end{bmatrix} + \frac{1}{5} = -\frac{16}{5} + \frac{20}{5} + \frac{1}{5} = \frac{5}{5} = 1$$

$$\mathbf{w}^T \mathbf{x}_3 + w_0 = \begin{bmatrix} -\frac{4}{5} & \frac{2}{5} \end{bmatrix} \begin{bmatrix} 3 \\ 8 \end{bmatrix} + \frac{1}{5} = -\frac{12}{5} + \frac{16}{5} + \frac{1}{5} = \frac{5}{5} = 1$$

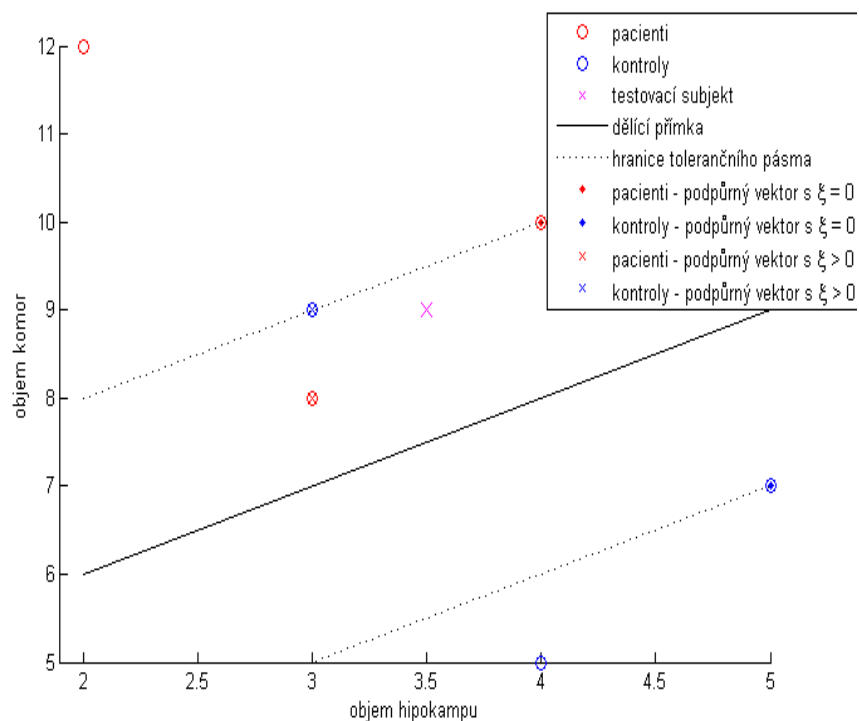
$$\mathbf{w}^T \mathbf{x}_4 + w_0 = \begin{bmatrix} -\frac{4}{5} & \frac{2}{5} \end{bmatrix} \begin{bmatrix} 5 \\ 7 \end{bmatrix} + \frac{1}{5} = -\frac{20}{5} + \frac{14}{5} + \frac{1}{5} = -\frac{5}{5} = -1$$

$$\mathbf{w}^T \mathbf{x}_5 + w_0 = \begin{bmatrix} -\frac{4}{5} & \frac{2}{5} \end{bmatrix} \begin{bmatrix} 3 \\ 9 \end{bmatrix} + \frac{1}{5} = -\frac{12}{5} + \frac{18}{5} + \frac{1}{5} = \frac{7}{5} = 1,4$$

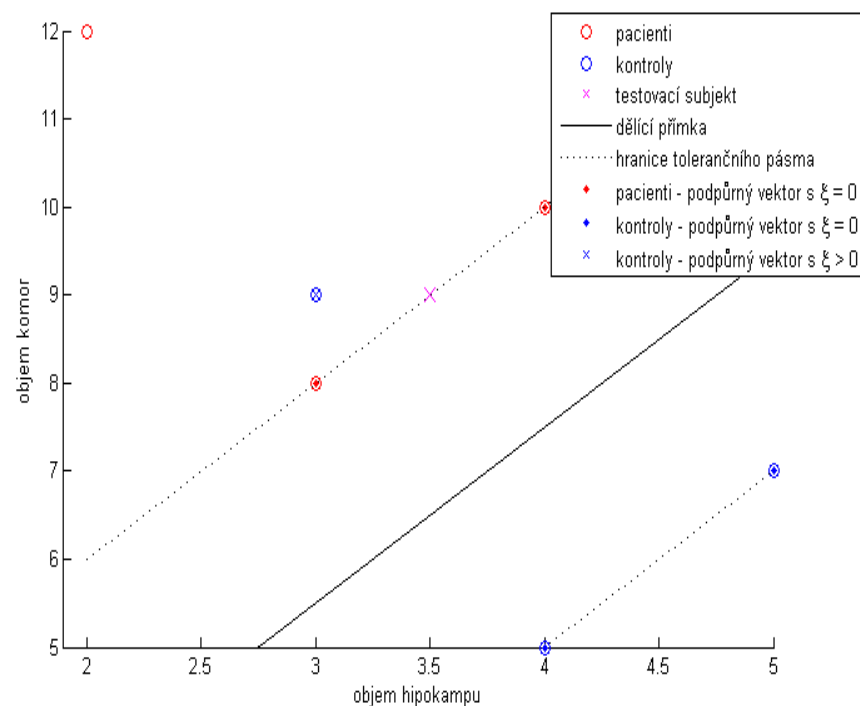
$$\mathbf{w}^T \mathbf{x}_6 + w_0 = \begin{bmatrix} -\frac{4}{5} & \frac{2}{5} \end{bmatrix} \begin{bmatrix} 4 \\ 5 \end{bmatrix} + \frac{1}{5} = -\frac{16}{5} + \frac{10}{5} + \frac{1}{5} = -\frac{5}{5} = -1$$

# Příklad – srovnání výsledků pro $C = 1$ a $C = 10$

$C = 1$ :



$C = 10$ :



# Příprava nových učebních materiálů pro obor Matematická biologie

je podporována projektem OPVK

č. CZ.1.07/2.2.00/28.0043

„Interdisciplinární rozvoj studijního  
oboru Matematická biologie“



MINISTERSTVO ŠKOLSTVÍ,  
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání  
pro konkurenceschopnost

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ