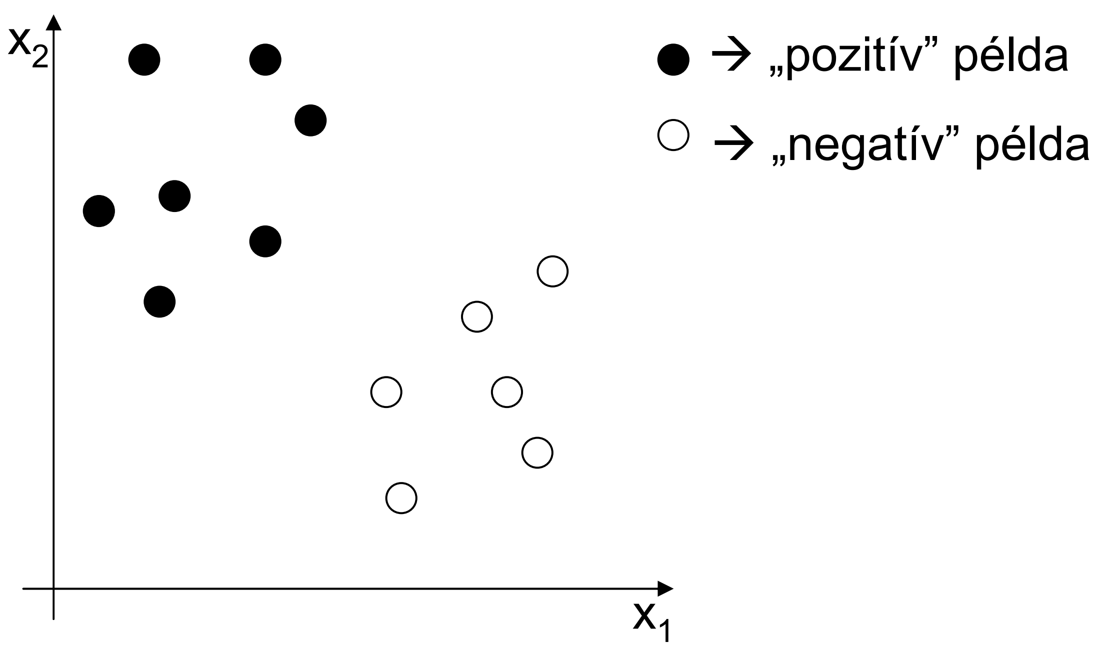
Gépi tanulás a gyakorlatban
SVM

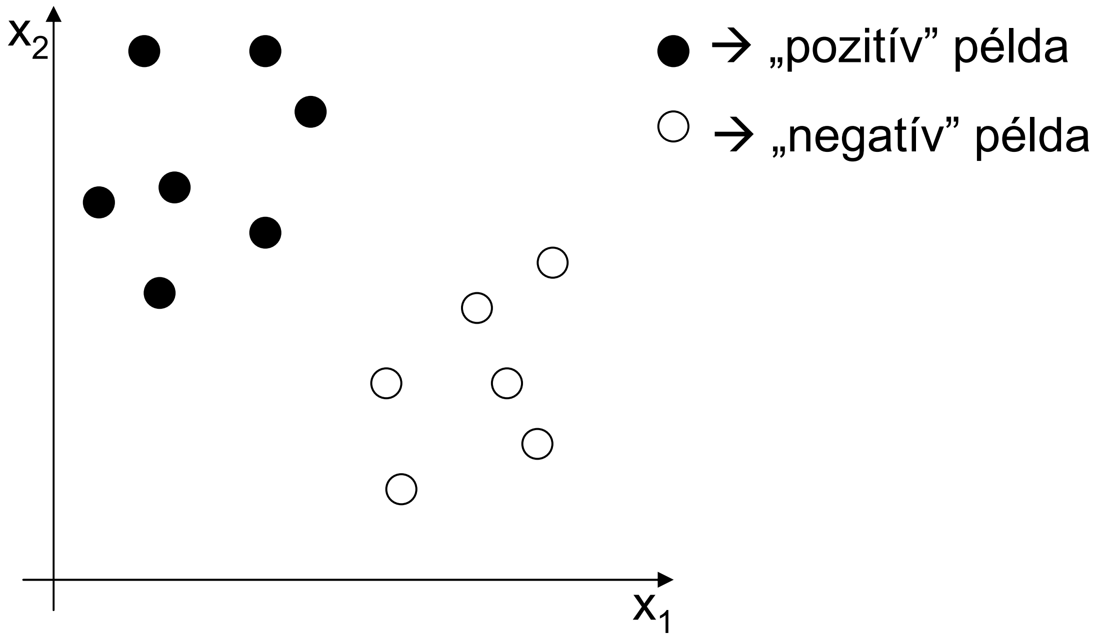
Klasszifikáció

- Feladat: előre meghatározott csoportok elkülönítése egymástól
  - Osztályokat elkülönítő felület
  - Osztályokhoz rendelt döntési függvények



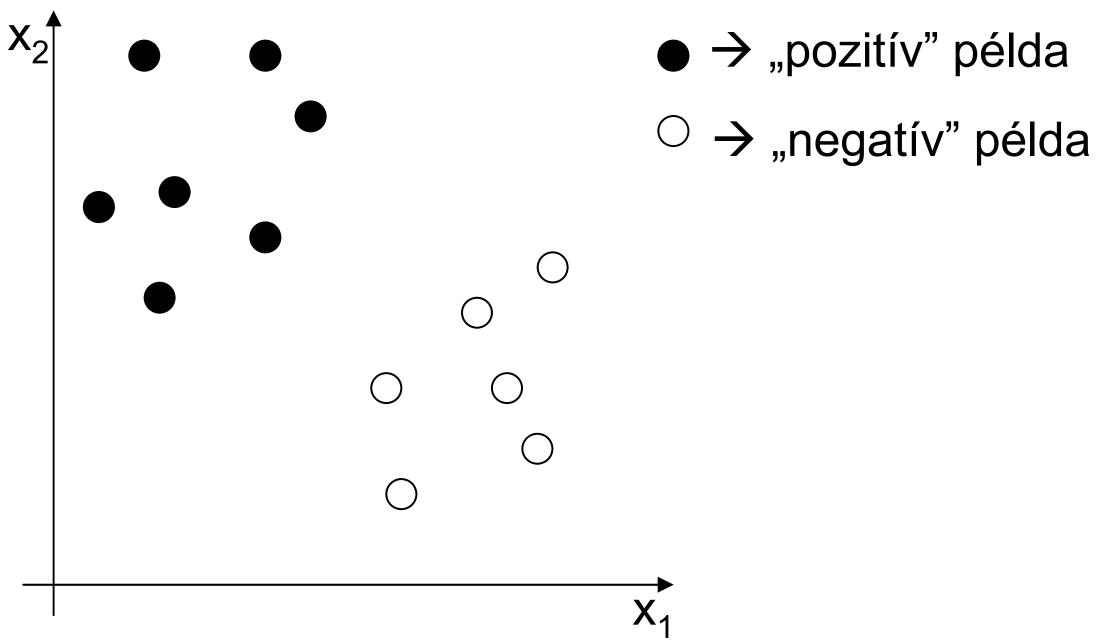
Klasszifikáció

- Feladat: előre meghatározott csoportok elkülönítése egymástól
  - Osztályokat elkülönítő felület
  - Osztályokhoz rendelt döntési függvények



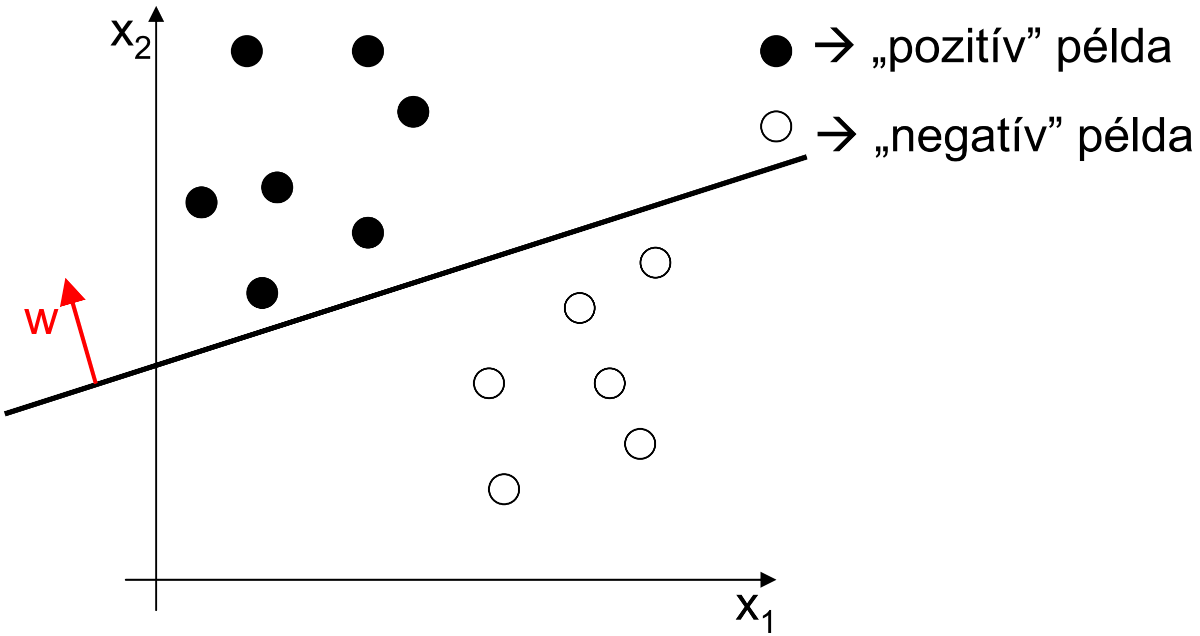
Lineáris klasszifikáció

- Feladat: lineáris diszkriminancia függvény meghatározása



Lineáris klasszifikáció

- Feladat: lineáris diszkriminancia függvény meghatározása → szeparáló hipersík

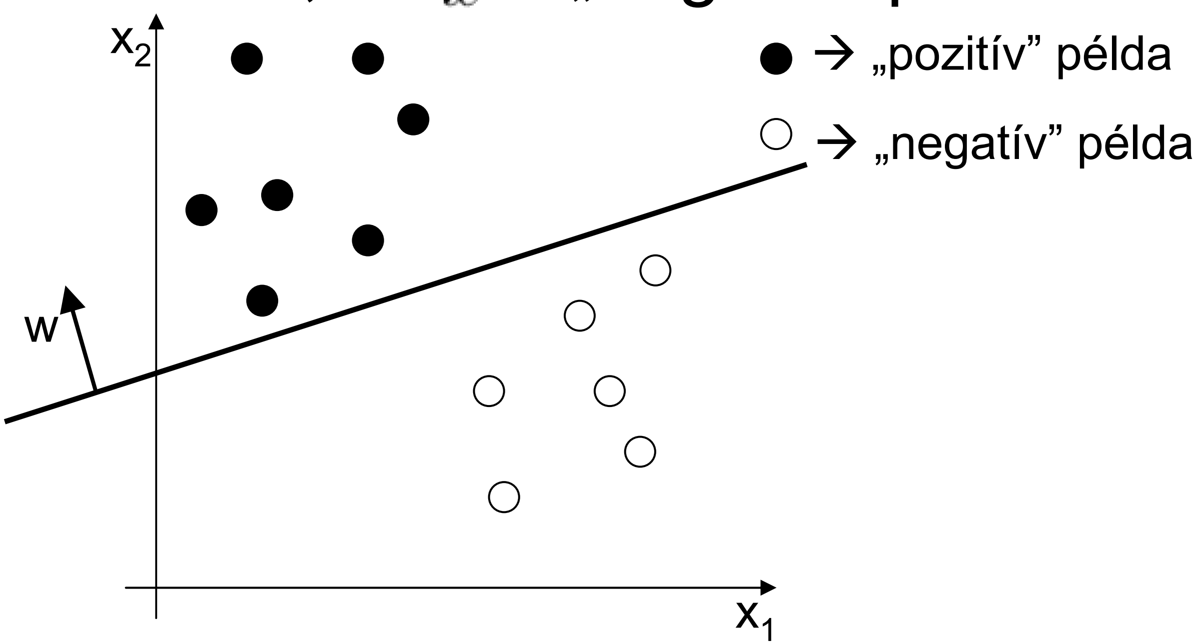


Lineáris klasszifikáció

- Feladat: lineáris diszkriminancia függvény meghatározása → szeparáló hipersík
  - Amelyre teljesül:

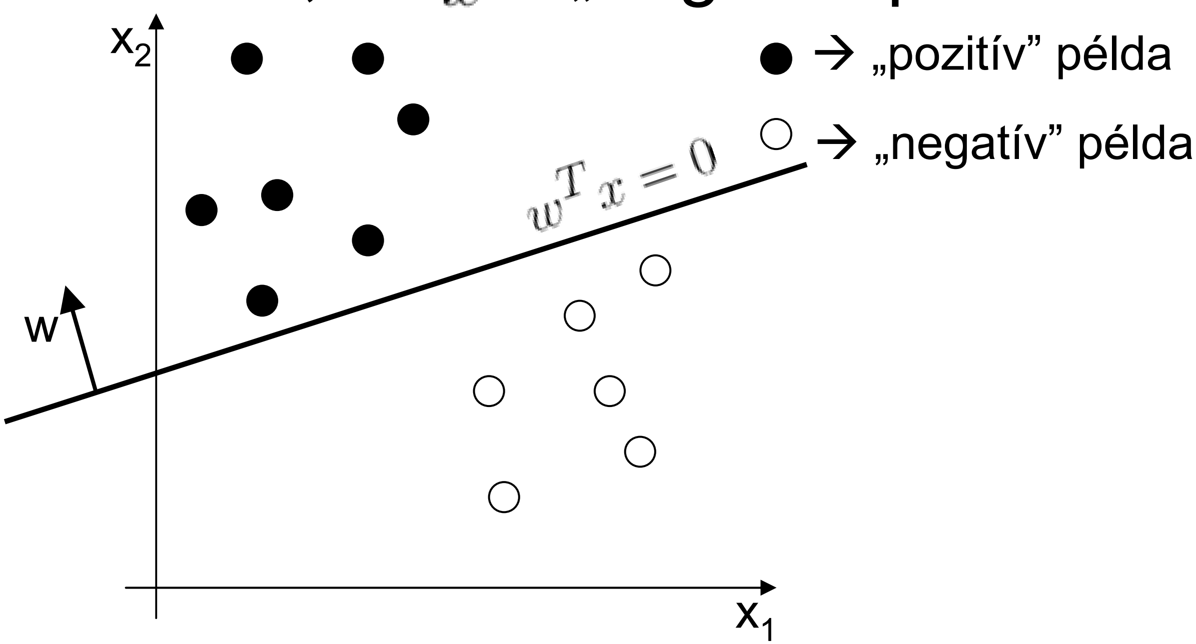
    $w^T x^{(i)} \geq 0$, ha $x^{(i)}$ „pozitív" példa

    $w^T x^{(i)} < 0$, ha $x^{(i)}$ „negatív" példa



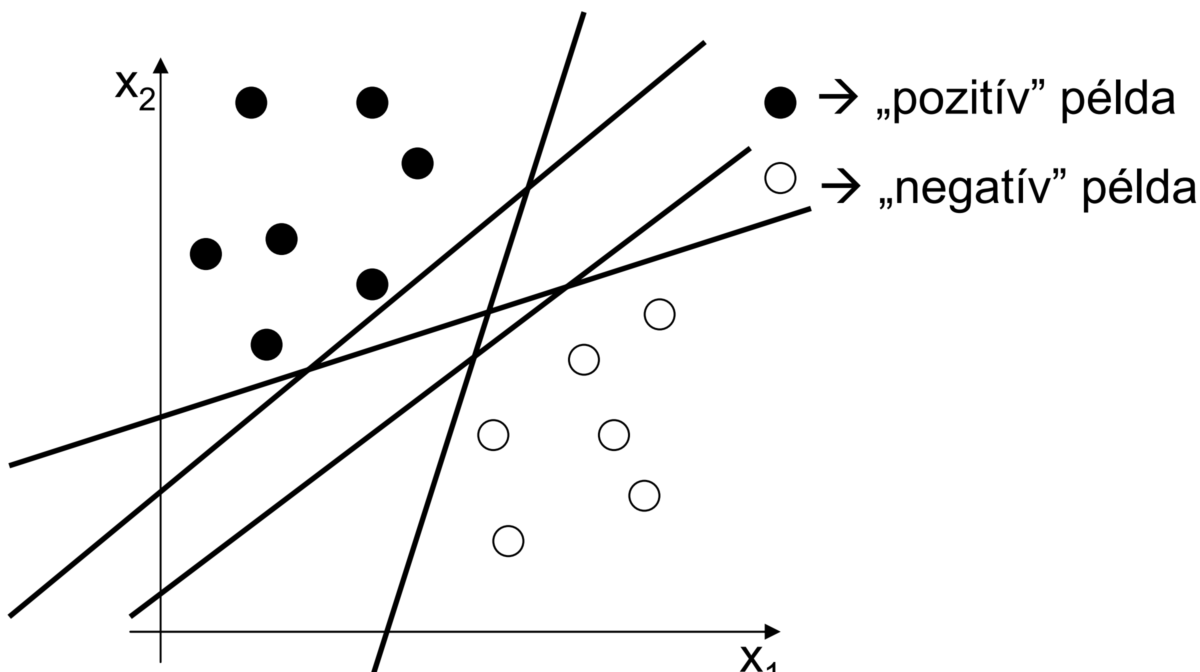
Lineáris klasszifikáció

- Feladat: lineáris diszkriminancia függvény meghatározása → szeparáló hipersík
  - Amelyre teljesül:
    $w^T x^{(i)} \geq 0$,ha $x^{(i)}$ „pozitív" példa
    $w^T x^{(i)} < 0$,ha $x^{(i)}$ „negatív" példa



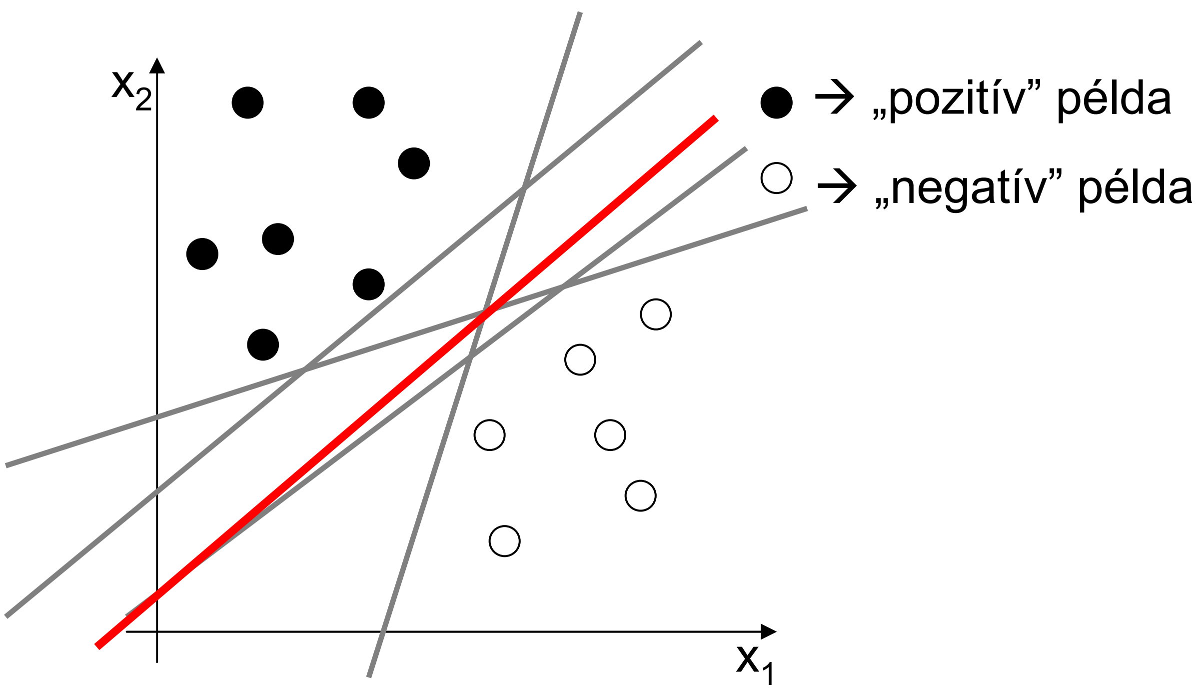
Lineáris klasszifikáció

- Feladat: lineáris diszkriminancia függvény meghatározása → szeparáló hipersík
- De melyiket???



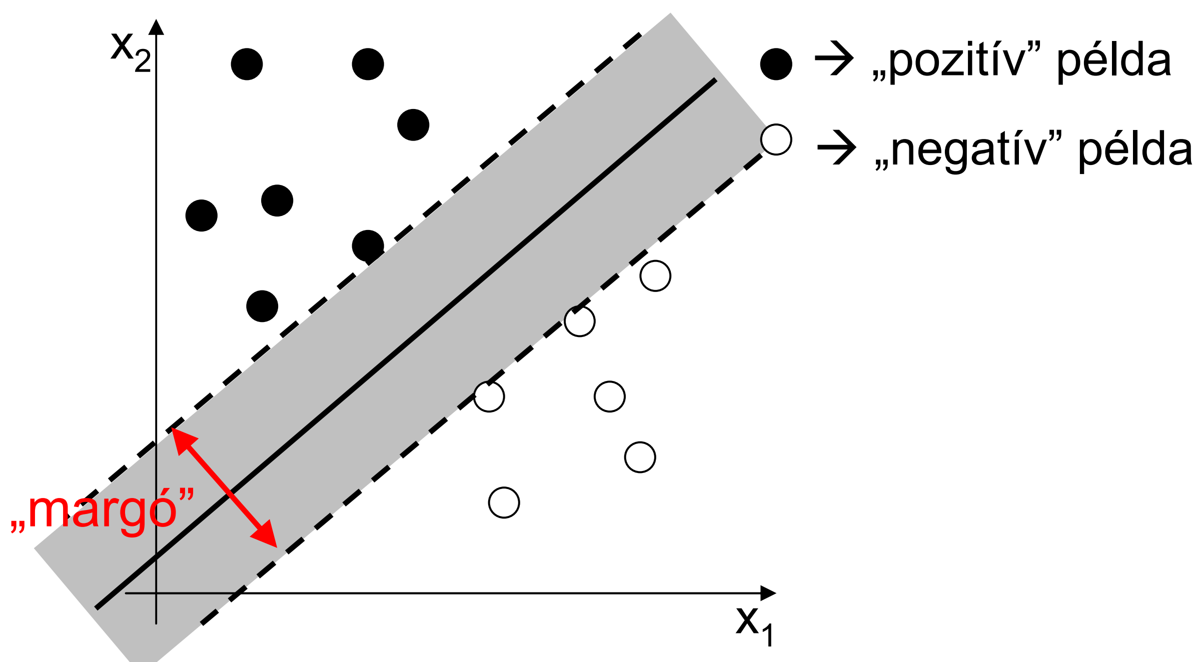
Lineáris klasszifikáció

- Feladat: lineáris diszkriminancia függvény meghatározása → szeparáló hipersík
- De melyiket??? „A legjobbat ⇔ SVM”



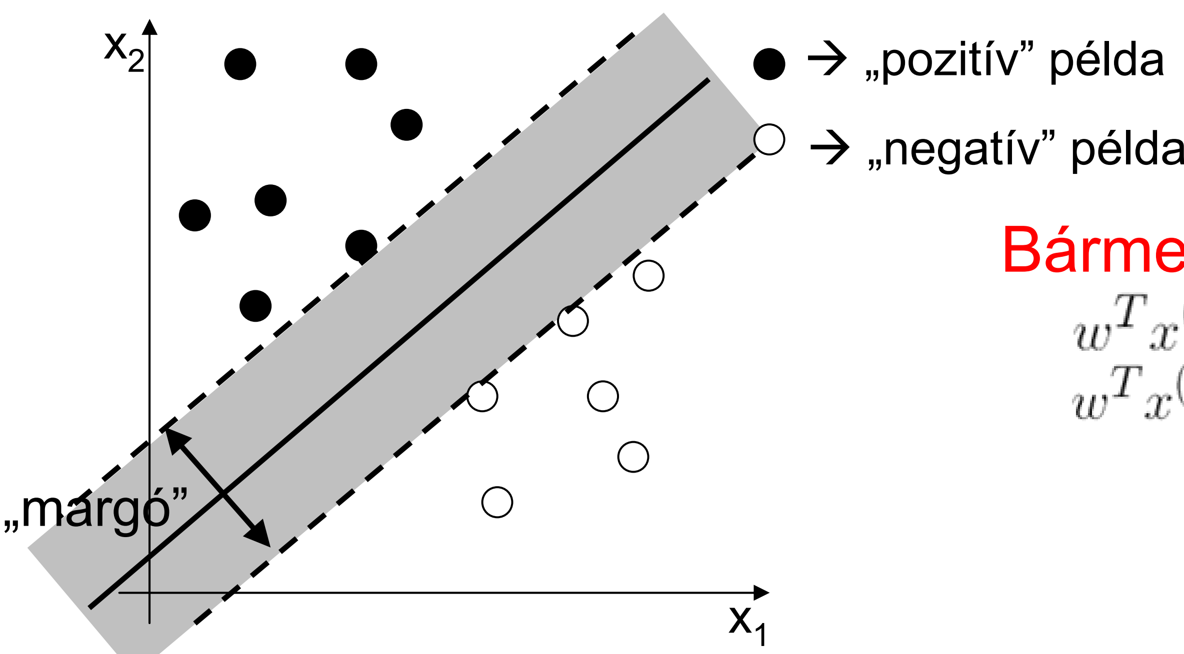
SVM (Support Vector Machine)

- Melyik a legjobb??? → Amelyik maximalizálja a „margó” méretét



SVM (Support Vector Machine)

- Keressük azt a egyenest amelyik maximali-zálja a „margó” méretét. Hogyan???
- Azaz keressük azt a w-t amelyre:



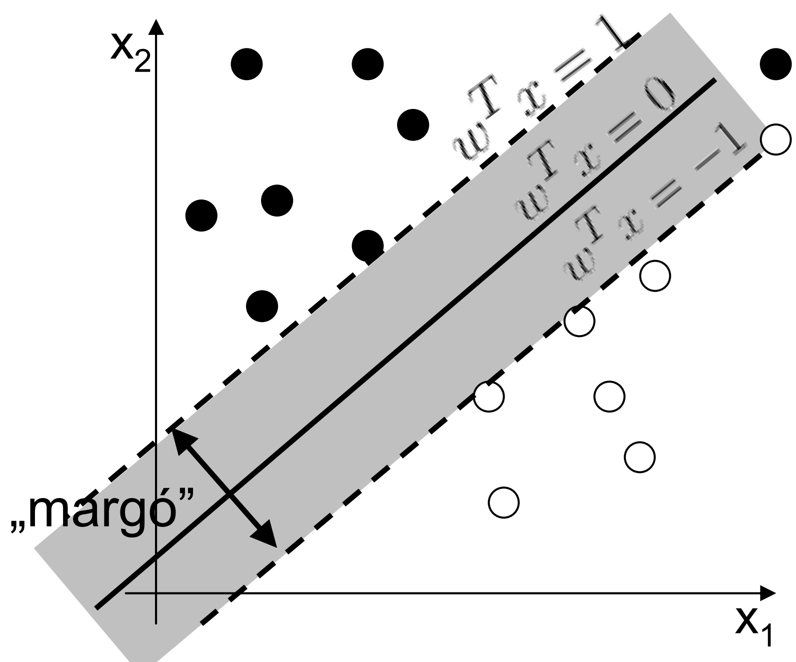
Bármely *i* példára:

$w^T x^{(i)} \geq 1$,ha $x^{(i)}$„pozitív”

$w^T x^{(i)} \leq -1$,ha $x^{(i)}$ „negatív”

SVM (Support Vector Machine)

- Keressük azt a egyenest amelyik maximalizálja a „margó” méretét. Hogyan???
- Azaz keressük azt a w-t amelyre:



→ „pozitív” példa

→ „negatív” példa

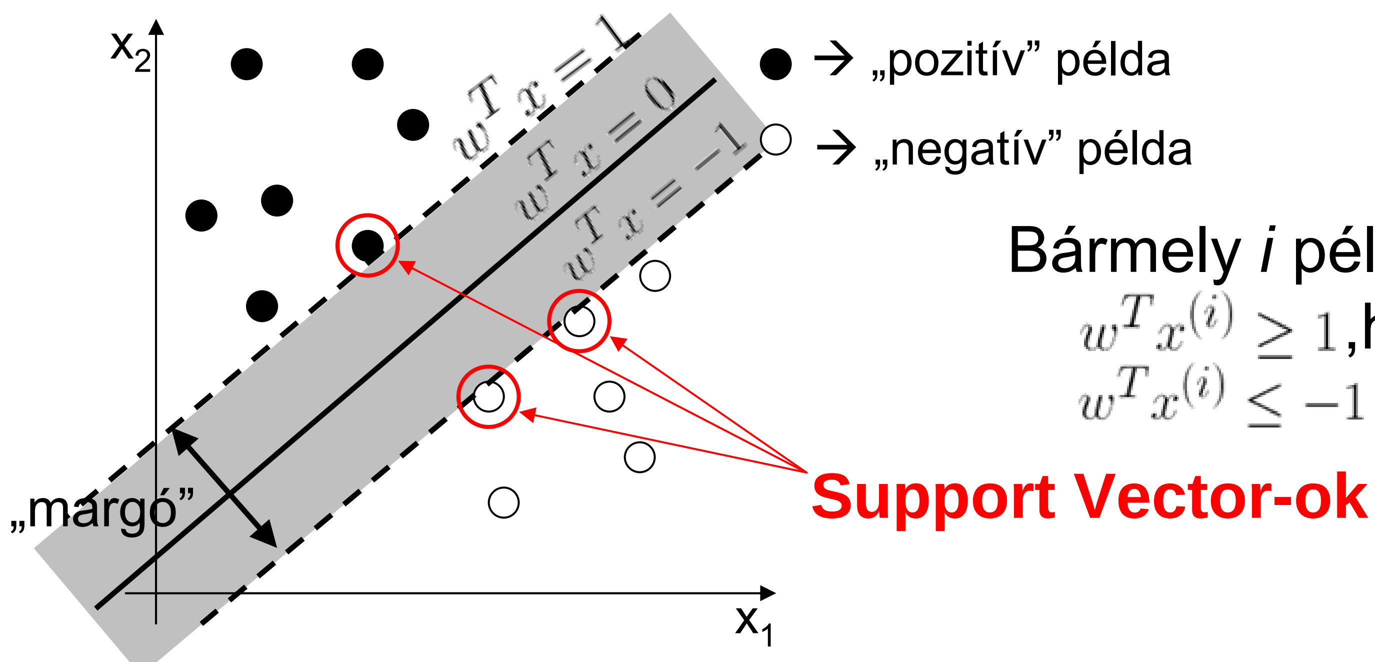
Bármely *i* példára:

$w^T x^{(i)} \geq 1$,ha $x^{(i)}$„pozitív”

$w^T x^{(i)} \leq -1$,ha $x^{(i)}$„negatív”

SVM (Support Vector Machine)

- Keressük azt a egyenest amelyik maximalizálja a „margó” méretét. Hogyan???
- Azaz keressük azt a w-t amelyre:



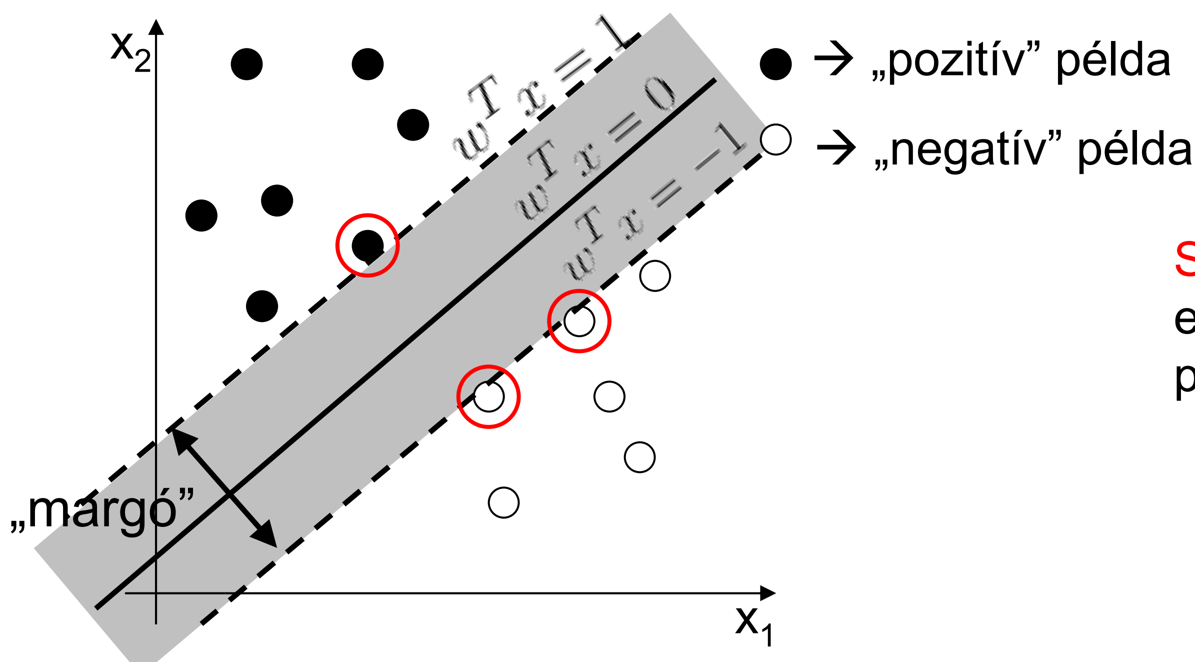
Bármely *i* példára:

$w^T x^{(i)} \geq 1$,ha $x^{(i)}$„pozitív”

$w^T x^{(i)} \leq -1$,ha $x^{(i)}$„negatív”

SVM (Support Vector Machine)

- Keressük azt a egyenest amelyik maximalizálja a „margó” méretét. Hogyan???

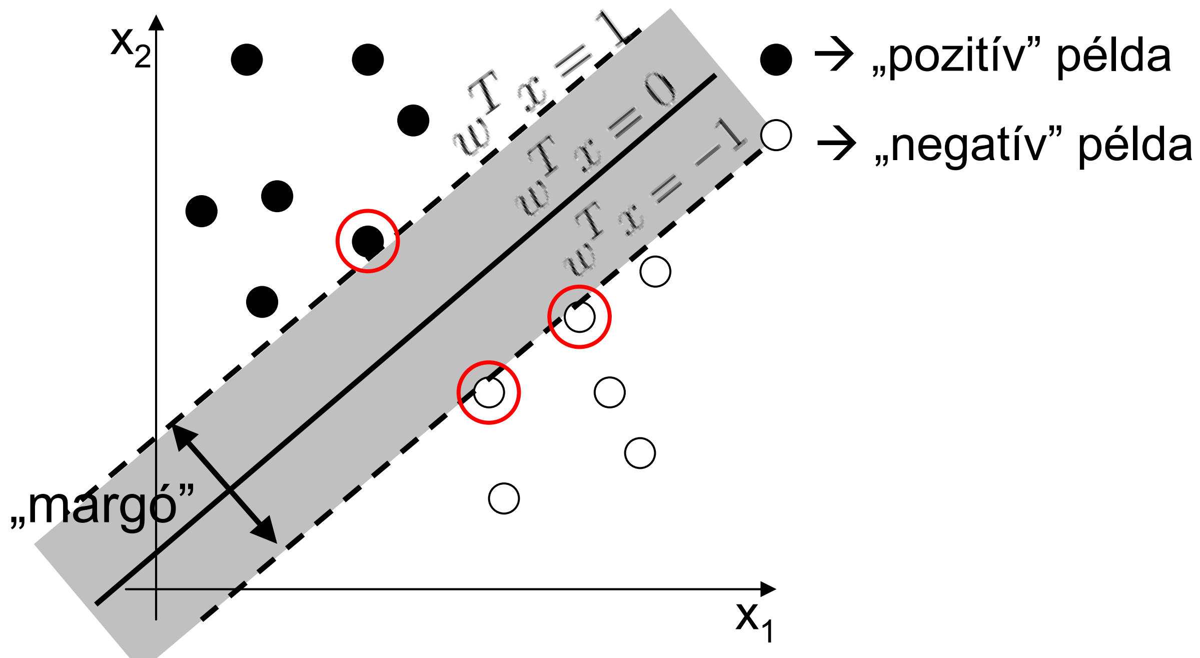


Support Vector (SV): A margó egyeneseire eső példák, azaz i példa SV, ha $|w^T x| = 1$ teljesül rá

SVM (Support Vector Machine)

- Keressük azt a egyenest amelyik maximali-zálja a „margó" méretét. Hogyan???

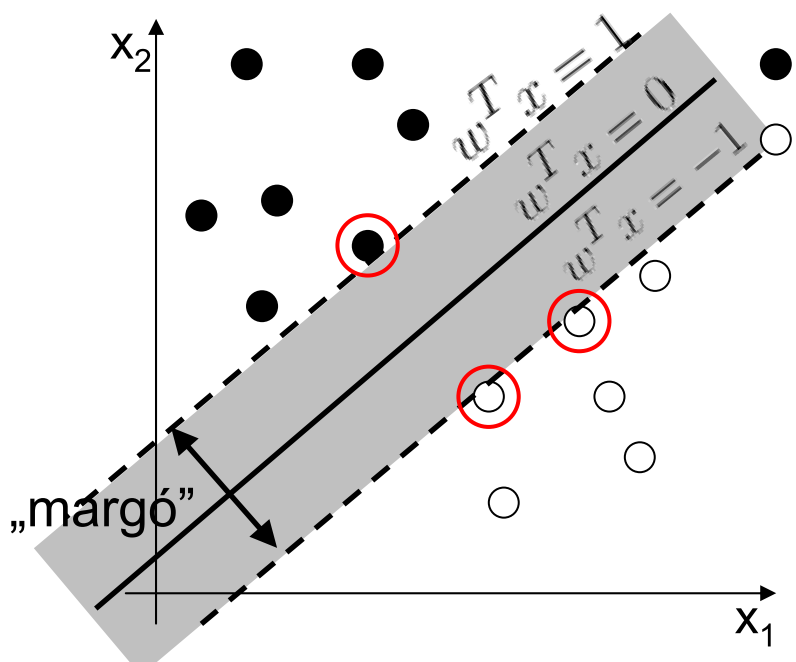
x példa w egyenestől mért (előjeles) távolsága: $\dfrac{w^T x}{\|w\|}$



SVM (Support Vector Machine)

- Keressük azt a egyenest amelyik maximalizálja a „margó" méretét. Hogyan???

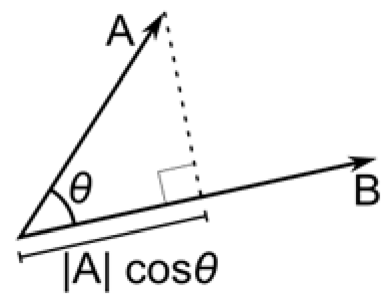
x példa w egyenestől mért (előjeles) távolsága: $$\frac{w^T x}{\|w\|}$$



→ „pozitív" példa

→ „negatív" példa

Mivel:



$$\mathbf{A}_B = \|\mathbf{A}\| \cos\theta$$

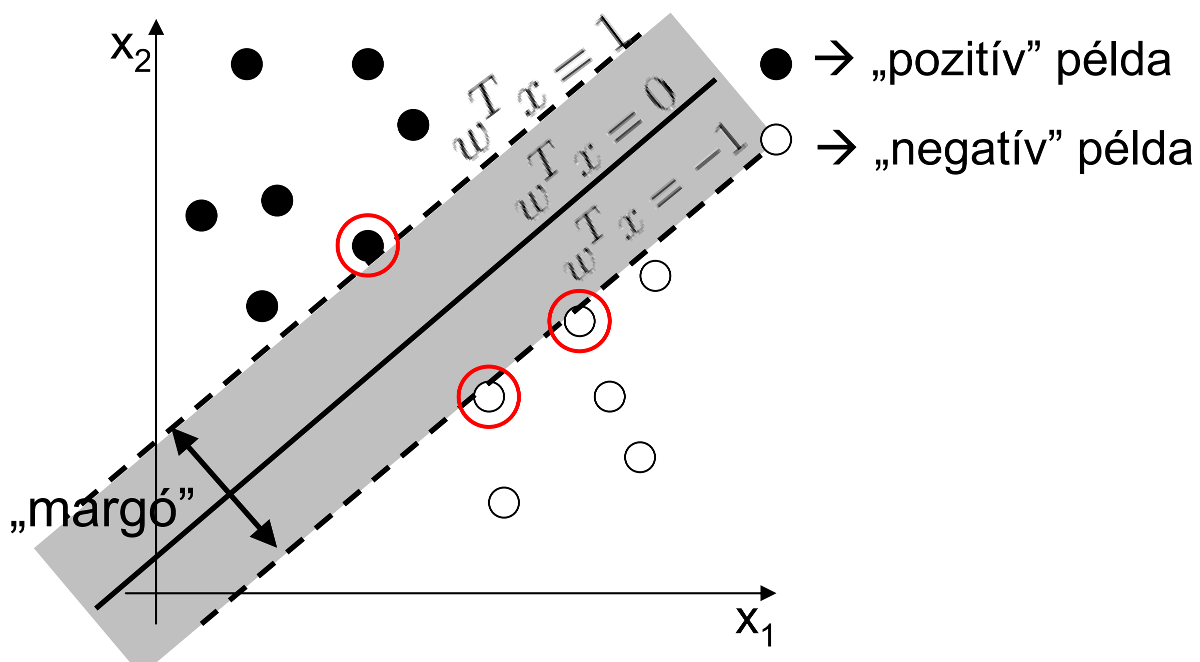
$$\mathbf{A} \cdot \mathbf{B} = \|\mathbf{A}\| \, \|\mathbf{B}\| \cos\theta$$

$$\mathbf{A}_B = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{B}\|}$$

SVM (Support Vector Machine)

- Keressük azt a egyenest amelyik maximalizálja a „margó” méretét. Hogyan???

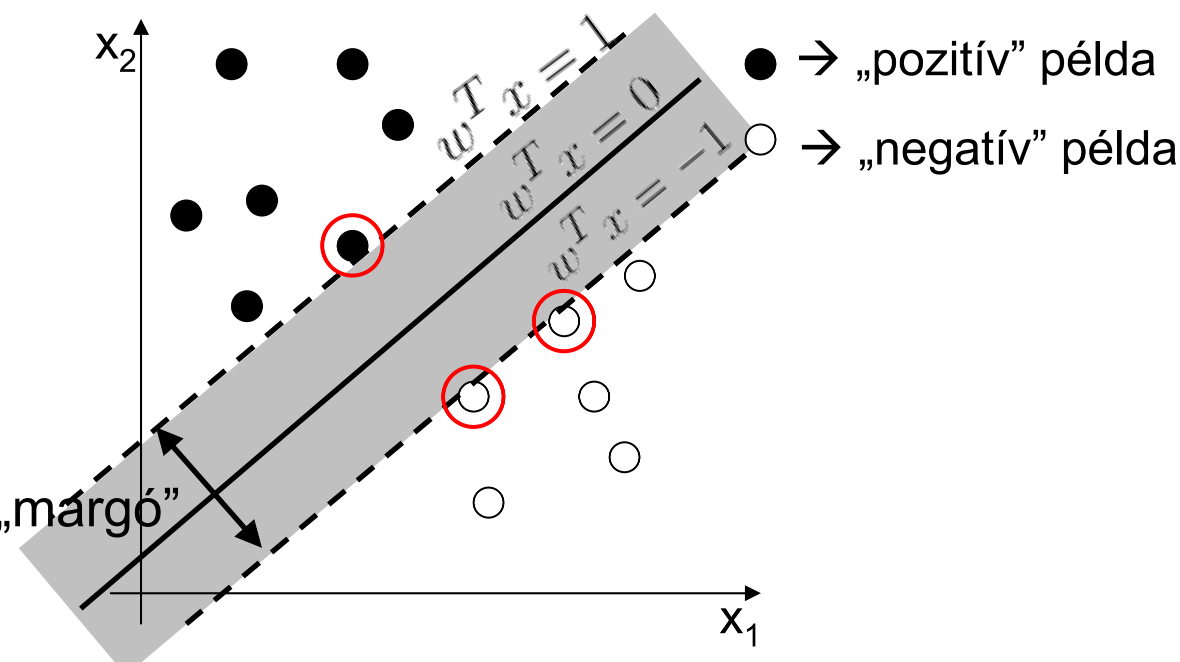
Egy SV példa w egyenestől mért távolsága: $\frac{1}{\|w\|}$



SVM (Support Vector Machine)

- Keressük azt a egyenest amelyik maximalizálja a „margó” méretét. Hogyan???

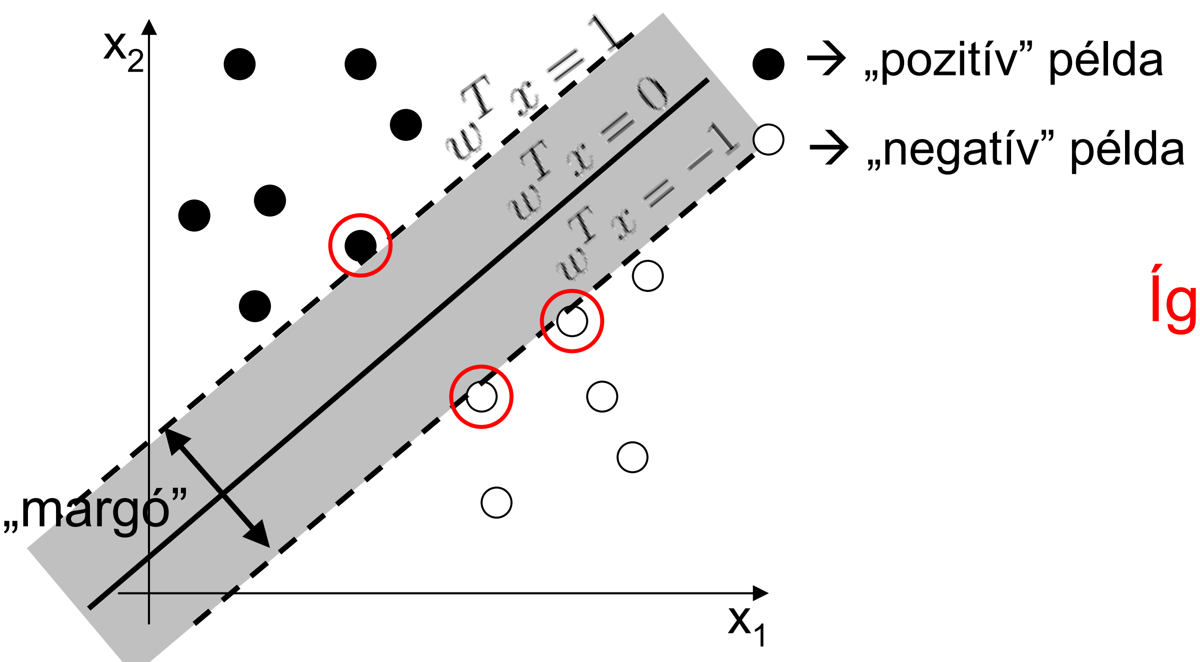
Egy SV példa w egyenestől mért távolsága: $\frac{1}{||w||}$



Mivel:
$|w^Tx| = 1$, ha x SV

SVM (Support Vector Machine)

- Keressük azt a egyenest amelyik maximalizálja a „margó" méretét. Hogyan???



Egy SV példa w egyenestől mért távolsága: $\frac{1}{||w||}$

Mivel:

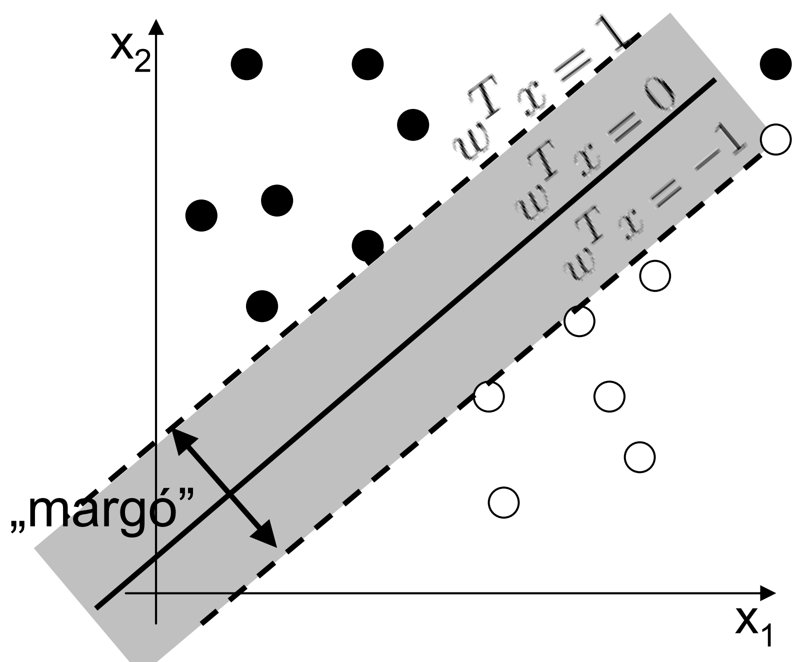
$|w^T x| = 1$, ha x SV

Így a margó mérete:

$$\frac{2}{||w||}$$

SVM (Support Vector Machine)

- Tehát az SVM feladata: maximalizálni a margó, amellett hogy jól osztályozza a példákat:



→ „pozitív” példa

→ „negatív” példa

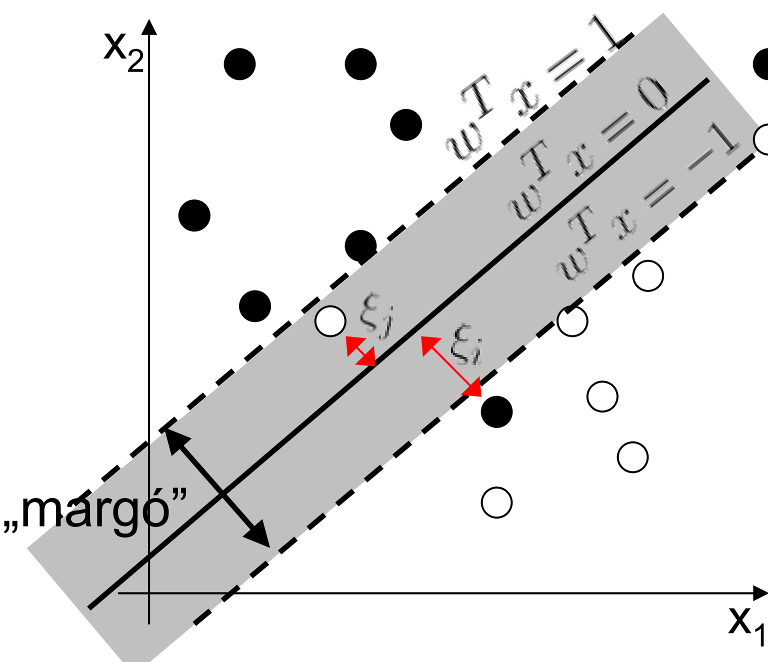
A margó mérete: $\frac{2}{\|w\|}$

$$\min_{w} \frac{1}{2} \|w\|^2$$

$$s.t. \quad y^{(i)}(w^T x^{(i)}) \geq 1, \quad i = 1 \ldots m$$

SVM (Support Vector Machine)

- Tehát az SVM feladata: maximalizálni a margó, amellett hogy jól osztályozza a példákat (nem szeparálható eset):



→ „pozitív" példa

→ „negatív" példa

A margó mérete: $\dfrac{2}{\|w\|}$

$$\min_{w,\xi} \frac{1}{2}\|w\|^2 + C\sum_i \xi_i$$

$$s.t. \quad y^{(i)}(w^T x^{(i)})) \geq 1 - \xi_i, \quad i = 1 \ldots m$$

Pegasos SVM

- A feladatot a primál alakban oldja meg

INPUT: S, λ, T
INITIALIZE: Set $\mathbf{w}_1 = 0$
FOR $t = 1, 2, \ldots, T$
  Choose $i_t \in \{1, \ldots, |S|\}$ uniformly at random.
  Set $\eta_t = \frac{1}{\lambda t}$
  If $y_{i_t} \langle \mathbf{w}_t, \mathbf{x}_{i_t} \rangle < 1$, then:
    Set $\mathbf{w}_{t+1} \leftarrow (1 - \eta_t \lambda)\mathbf{w}_t + \eta_t y_{i_t} \mathbf{x}_{i_t}$
  Else (if $y_{i_t} \langle \mathbf{w}_t, \mathbf{x}_{i_t} \rangle \geq 1$):
    Set $\mathbf{w}_{t+1} \leftarrow (1 - \eta_t \lambda)\mathbf{w}_t$
  [Optional: $\mathbf{w}_{t+1} \leftarrow \min\left\{1, \frac{1/\sqrt{\lambda}}{\|\mathbf{w}_{t+1}\|}\right\} \mathbf{w}_{t+1}$]
OUTPUT: $\mathbf{w}_{T+1}$

$$\min_{w,\xi} \frac{1}{2} \|w\|^2 + C \sum_i \xi_i$$

$$s.t. \quad y^{(i)}(w^T x^{(i)})) \geq 1 - \xi_i, \quad i = 1 \ldots m$$

Duális SVM

- A primál feladat átírható duális alakra a Lagrange-multiplikátorok segítségével:

$$
\begin{aligned}
\max_{\alpha} \quad & W(\alpha) = \sum_{i=1}^{m} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{m} y^{(i)} y^{(j)} \alpha_i \alpha_j \langle x^{(i)}, x^{(j)} \rangle. \\
\text{s.t.} \quad & \alpha_i \geq 0, \quad i = 1, \ldots, m \\
& \sum_{i=1}^{m} \alpha_i y^{(i)} = 0,
\end{aligned}
$$

Duális SVM

- Megmutatható: $\alpha_i(y^{(i)}(w^T x^{(i)}) - 1) = 0$ teljesül minden példára, azaz

  $y^{(i)}(w^T x^{(i)}) - 1 = 0$ vagy

  $\alpha_i = 0$

- → csak a SV-ok esetén $\alpha_i \neq 0$
- Tehát a hipersík megadható:

$$w = \sum_{\forall i} \alpha_i y^{(i)} x^{(i)} = \sum_{i \in SV} \alpha_i y^{(i)} x^{(i)}$$

Duális SVM

- A hipersík megadható:

$$w = \sum_{\forall i} \alpha_i y^{(i)} x^{(i)} = \sum_{i \in SV} \alpha_i y^{(i)} x^{(i)}$$

- Azaz az egyenes egyenlete:

$$w^T x = \sum_{i \in SV} \alpha_i y^{(i)} x^{(i)T} x$$

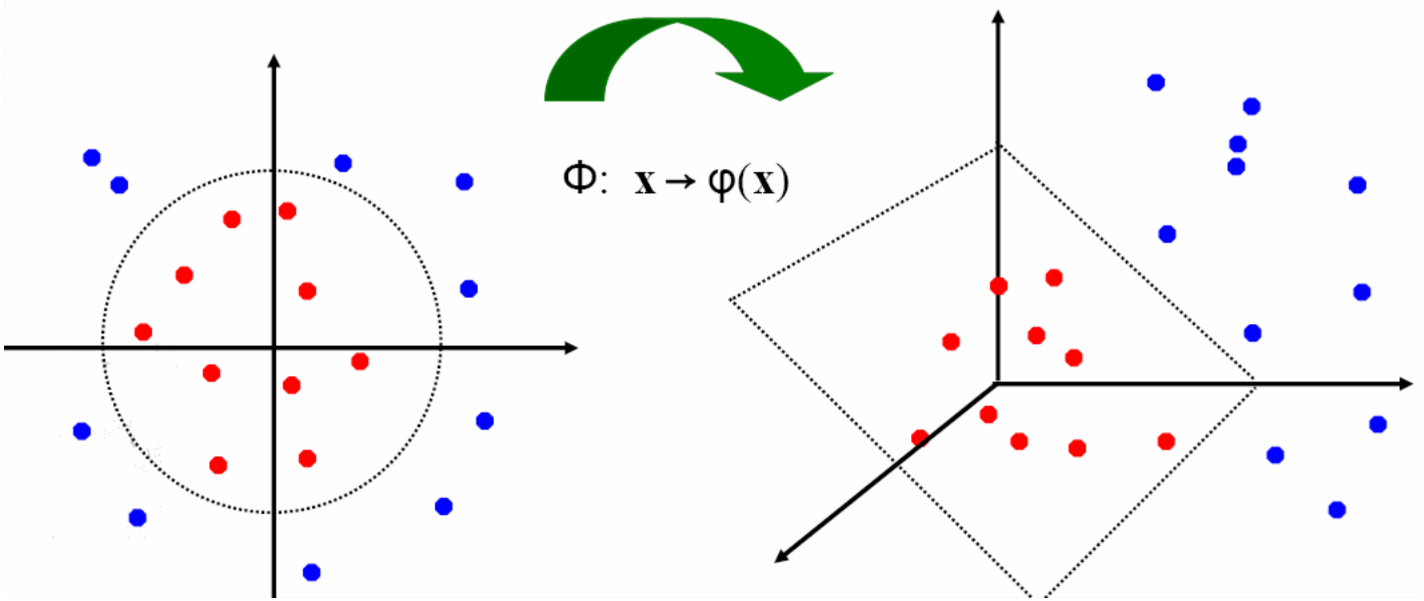
- Kernel trükk:

$$w^T x = \sum_{i \in SV} \alpha_i y^{(i)} K(x^{(i)}, x)$$

Duális SVM kernel trükk

- Ha lineárisan nem szeparálhatóak a példák, transzformáljuk őket olyan térbe ahol igen!

$$K(x_i, x_j) \equiv \Phi(x_i)^T \Phi(x_j)$$



http://www.youtube.com/watch?v=3liCbRZPrZA