# یادگیری ماشین با پایتون

فرزاد حسینیان

دانشگاه محقق اردبیلی

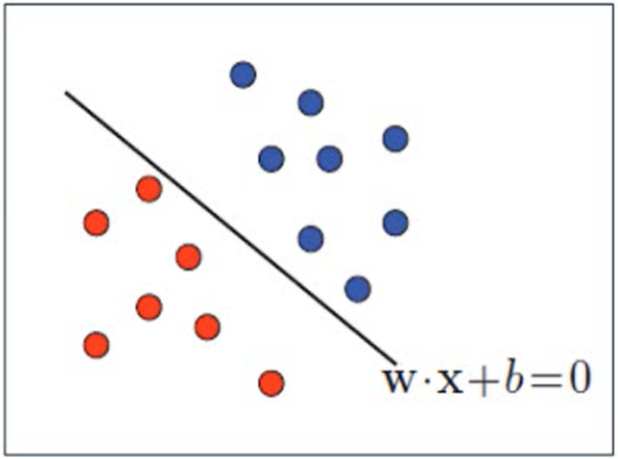
دانشکده علوم کامپیوتر

آبان 98

# جلسه ششم

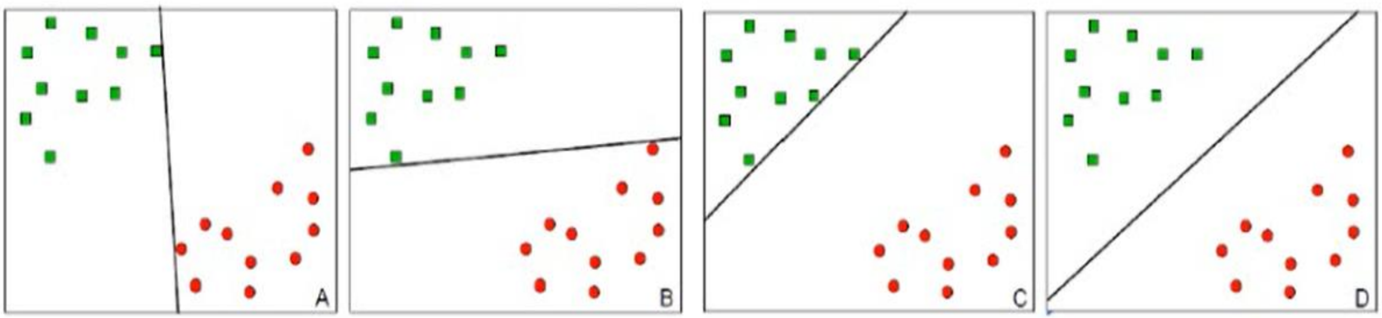
# ماشین بردار پشتیبان

# ماشین بردار پشتیبان



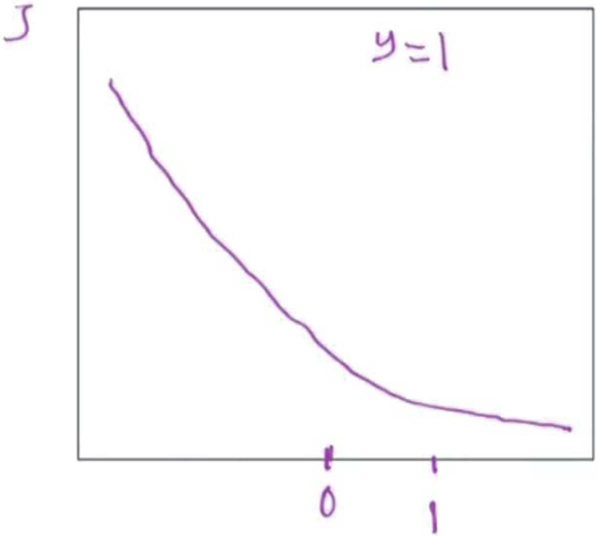

- روشی برای طبقه‌بندی (یادگیری با نظارت)
- در سال ۱۹۶۳ توسط Vapnik معرفی شد.
- مرزی است که به بهترین شکل، دسته‌های داده‌ها را از یکدیگر جدا می‌کند.

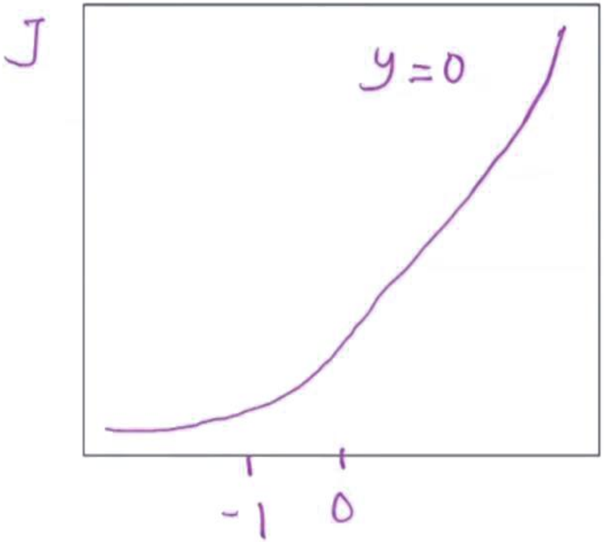
# کدام مرز جدا سازی؟



# تابع خطای SVM

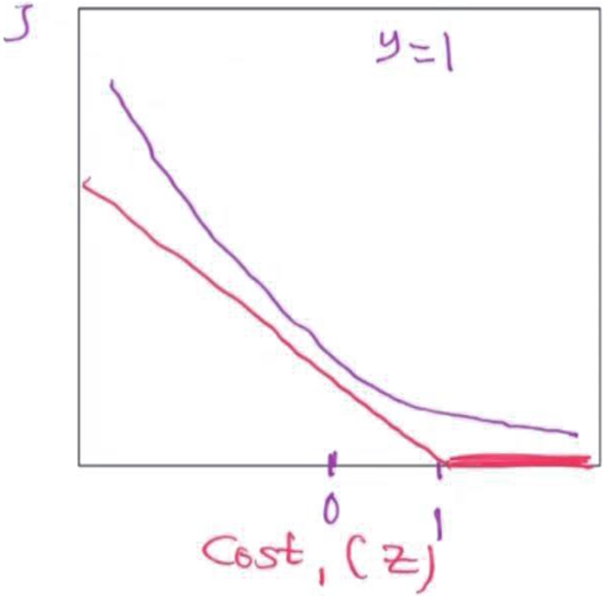
$$J(\theta) = -\frac{1}{m}\sum_1^m \underbrace{y^{(i)}log(h_\theta(x^{(i)}))}_{} + \underbrace{(1-y^{(i)})log(1-h_\theta(x^{(i)}))}_{}$$

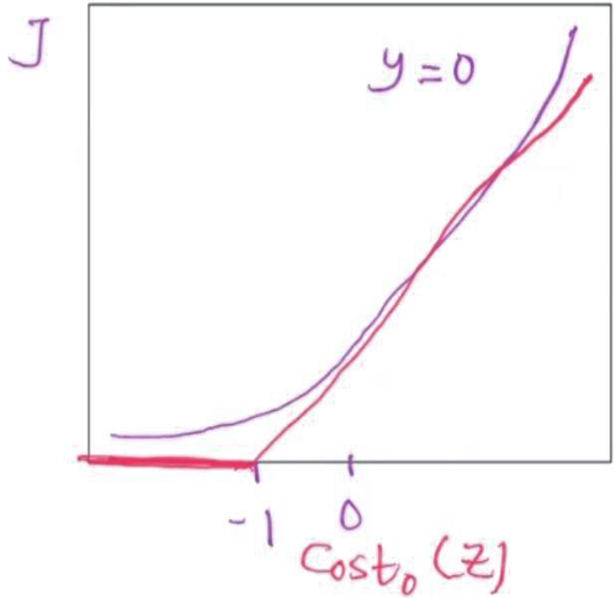


# تابع خطای SVM

$$J(\theta) = -\frac{1}{m}\sum_1^m \underbrace{y^{(i)}log(h_\theta(x^{(i)}))} + \underbrace{(1-y^{(i)})log(1-h_\theta(x^{(i)}))}$$



## تابع هزینه Svm

$$J(\theta) = \frac{1}{m}\sum_{i=1}^{m}[y^{(i)}cost_1(\theta^t(x^{(i)})) + (1-y^{(i)})cost_0(\theta^t(x^{(i)}))] + \frac{\lambda}{2m}\sum_{j=1}^{n}\theta_j^2$$

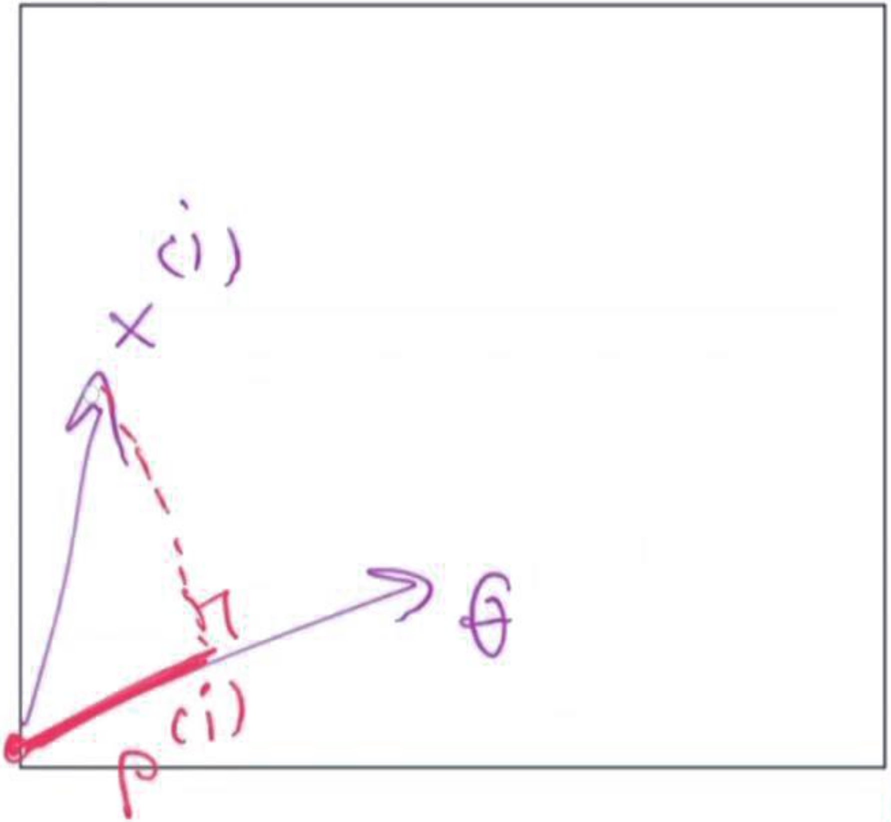
A + Л $B$ => C A + B , C=1/Л

$$J(\theta) = C\sum_{i=1}^{m}[y^{(i)}cost_1(\theta^t(x^{(i)})) + (1-y^{(i)})cost_0(\theta^t(x^{(i)}))] + \frac{1}{2}\sum_{j=1}^{n}\theta_j^2$$

$$\min \frac{1}{2} \underbrace{\sum_{j=1}^{n} \theta_j^2}_{\|\theta\|^2} \rightarrow \min \frac{1}{2} \|\theta\|^2$$

$$\theta = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_n \end{bmatrix}$$

$$\text{s.t.} \begin{cases} \theta^T x^{(i)} \geq 1 & \text{if } y^{(i)} = 1 \\ \theta^T x^{(i)} \leq -1 & \text{if } y^{(i)} = 0 \end{cases}$$

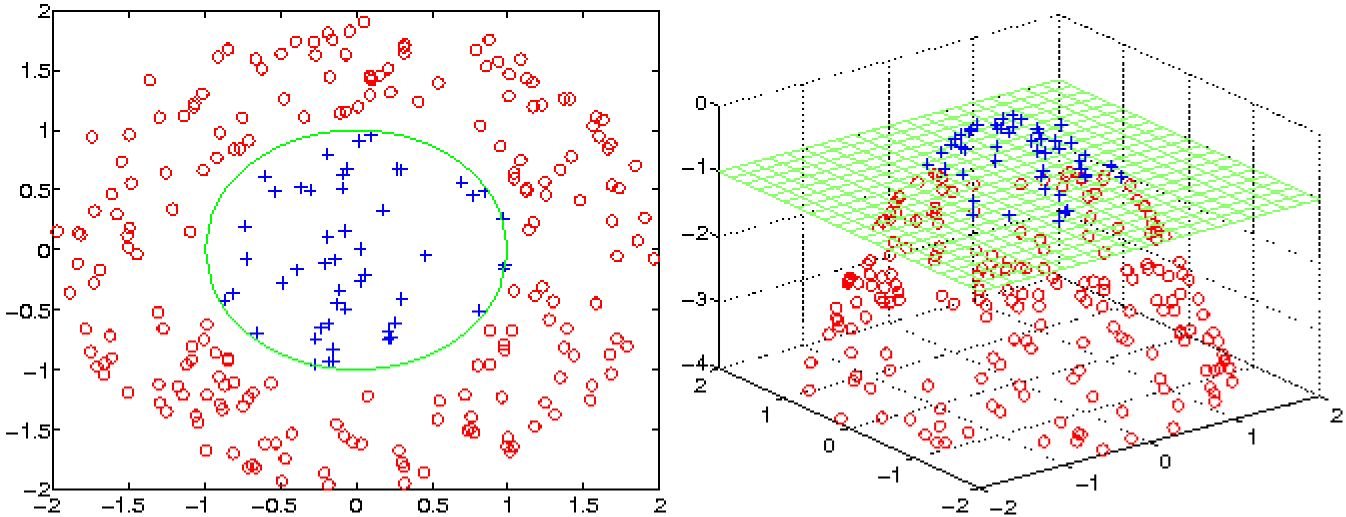



Max

$$\theta^{T} x^{(i)} = \underbrace{p^{(i)}}_{\text{Max}} \; \underbrace{\|\theta\|}_{\text{min}}$$

## Kernel

**ایده:** نگاشت مسئله به یک فضای ویژگی جدید با استفاده از تبدیلات خطی



(a) Input Space (b) Projected Space

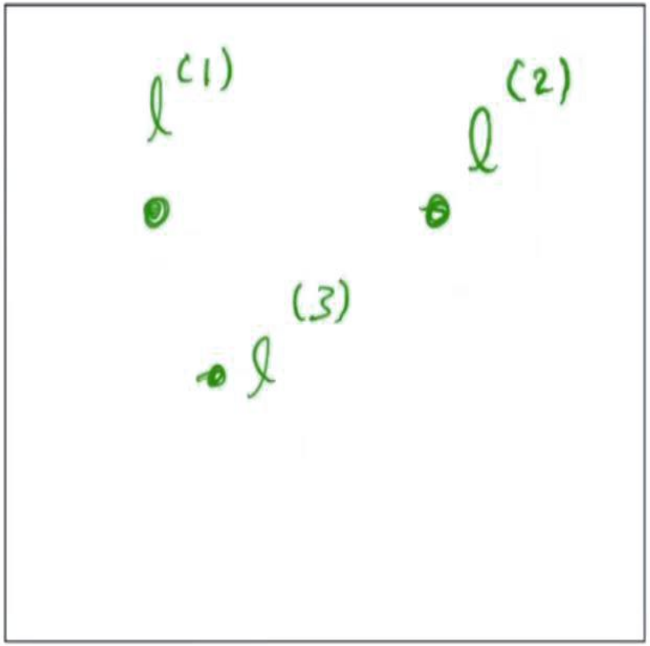
# Kernel

- New features are selected depending on proximity to landmarks

$$f_1 = \text{similarity}(x, l^{(1)})$$

$$f_2 = \text{similarity}(x, l^{(2)})$$

$$\vdots$$




landmark

# Similarity

- A possible choice

$$f_i = \exp\left(-\frac{\|x - l^{(i)}\|^2}{2\sigma^2}\right)$$

- When x is near a landmark

$$x \approx l^{(i)} \longrightarrow f_i \approx \exp(0) = 1$$

- When x is far from a land mark

$$f_i \approx \exp(-\infty) = 0$$

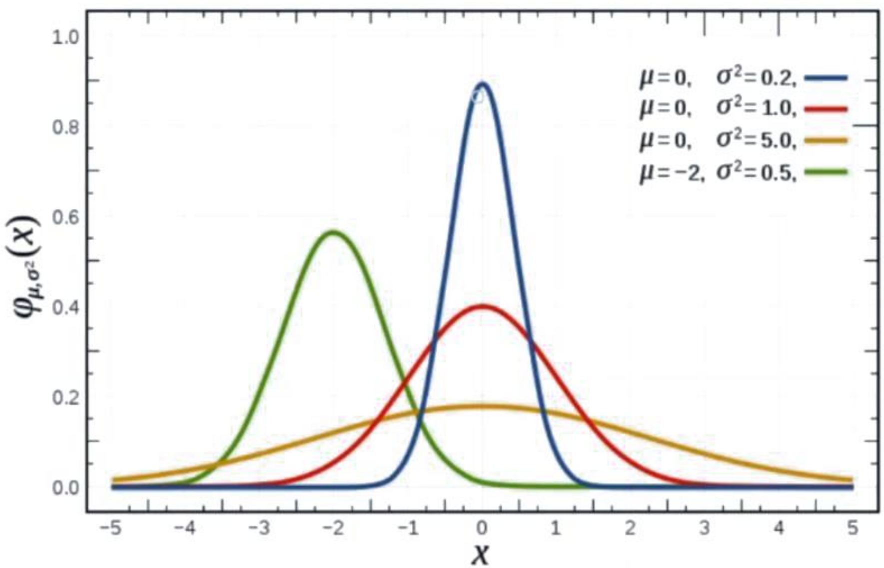
## Gaussian kernels

mean ($l$ )

$$f_i(x) = exp(-\frac{||x - l^{(i)}||^2}{2\sigma^2})$$

Variance

Spread



$y=1$ if $\theta_0+\theta_1 f_1+\theta_2 f_2+\theta_3 f_3 > 0$

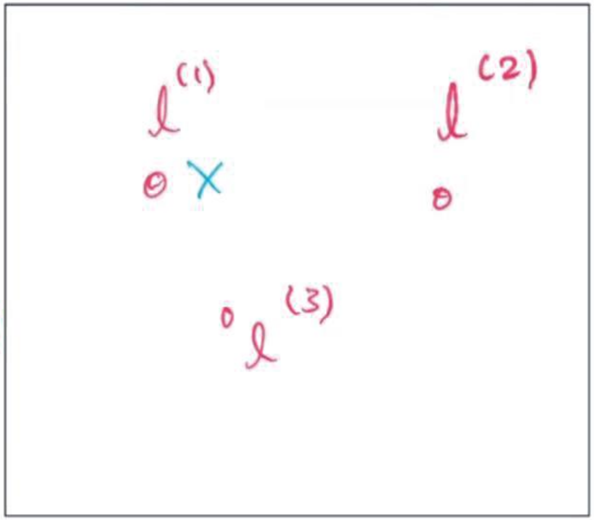
$y=0$ o.w.

$f_1, f_2, f_3$

$$f_1 = \exp\left(-\frac{\|x - l^{(1)}\|^2}{2\sigma^2}\right)$$

$\theta_0 = -0.5, \theta_1 = 1, \theta_2 = 1, \theta_3 = 0$

$$h_\theta(x) = -0.5 + 1(1) + 1(0) + 0(0) = 0.5$$



$y=1$ if $\theta_0+\theta_1 f_1+\theta_2 f_2+\theta_3 f_3 > 0$

$y=0$ o.w.

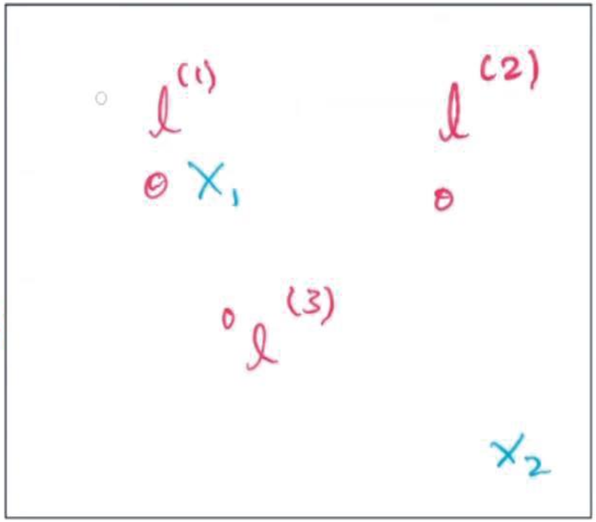
$f_1, f_2, f_3$

$$f_1 = \exp\left(-\frac{\|x - l^{(1)}\|^2}{2\sigma^2}\right)$$

$\theta_0 = -0.5$, $\theta_1 = 1$, $\theta_2 = 1$, $\theta_3 = 0$

$h_\theta(x_1) = -0.5 + 1(1) + 1(0) + 0(0) = 0.5 \longrightarrow$ class 1

$h_\theta(x_2) = -0.5 \longrightarrow$ class 0



$y = 1$ if $\theta_0 + \theta_1 f_1 + \theta_2 f_2 + \theta_3 f_3 > 0$

$y = 0$ o.w.

$f_1, f_2, f_3$

$$f_1 = \exp\left(-\frac{\| x - l^{(1)} \|^2}{2\sigma^2}\right)$$

$\theta_0 = -0.5, \ \theta_1 = 1, \ \theta_2 = 1, \ \theta_3 = 0$

$h_\theta(x_1) = -0.5 + 1(1) + 1(0) + 0(0) = 0.5$

$\longrightarrow$ class 1

$h_\theta(x_2) = -0.5 \rightarrow$ class 0