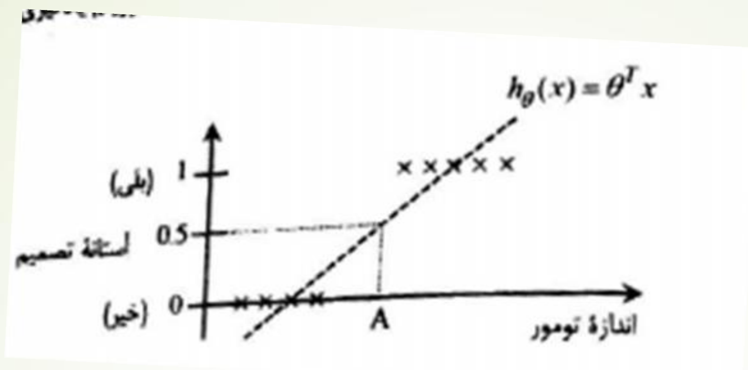
الگوریتم لجستیک جهت دسته بندی

ایمیل های واقعی از اسپم ها
تومورهای خوش خیم از بدخیم
صداها از منابع مختلف
و....

$$y \in \{0,1\} \qquad \begin{array}{l} 0: \text{"Negative Class"} \\ 1: \text{"Positive Class"} \end{array}$$

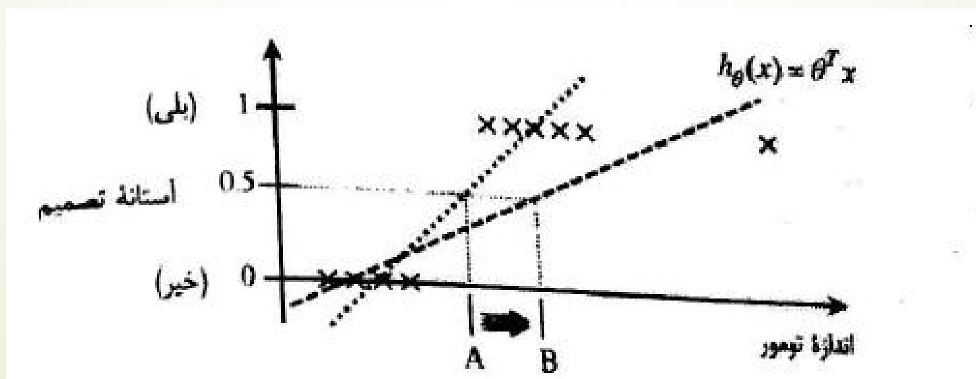
نمونه داده های اندازه تومور و نوع آن را نشان می دهد. (۰ خوش خیم و ۱ بدخیم)
مقادیر y گسسته است.

استفاده از رگرسیون خطی جهت دسته بندی



$$\text{if } h_\theta(x) < ۰٫۵,\ \text{Predict } "y = ۰"$$
$$\text{if } h_\theta(x) \geq ۰٫۵,\ \text{Predict } "y = ۱"$$

- اگر یک تومور بزرگ بدخیم داشته باشیم استانه از A به B جابجا می شود.



- می خواهیم مقدار محاسبه شده بین صفر و یک باشد.

رگرسیون خطی گزینه مناسبی جهت دسته بندی نیست

چون با اضافه شدن اطلاعات جدید اعتباردسته بندی از بین می رود.

برای رفع این موانع از الگوریتم لجستیک استفاده می شود.

$$X_{(n+1)*1} = \begin{matrix} x_0 \\ x_1 \\ x_2 \\ . \\ x_n \end{matrix} \qquad \theta_{(n+1)*1} = \begin{matrix} \theta_0 \\ \theta_1 \\ \theta_2 \\ . \\ \theta_n \end{matrix}$$

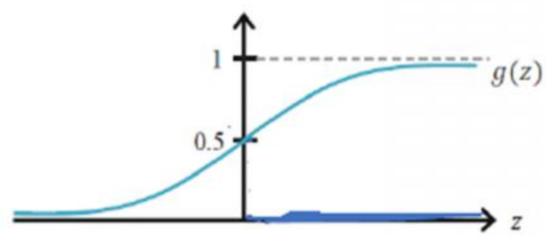
$$h(x) = \theta_0 x_0 + \theta_1 x_1 + \ldots\ldots + \theta_n x_n$$

$$h(x) = \theta^T X = [\theta_0\ \theta_1\ \ \ \theta_n] * \begin{matrix} x_0 \\ x_1 \\ x_2 \\ . \\ x_n \end{matrix}$$

ساخت تابع فرضیه

$$h_\theta(x) = g(\theta^T x) = \frac{1}{1+e^{-\theta^T x}} \qquad 0 \le h_\theta(x) \le 1$$

$$g(z) = \frac{1}{1+e^{-z}}$$



Sigmoid function

تابع زیگموید یا لجستیک

$$if \;\; h_\theta(x) \ge 0.5 \quad : y = 1$$
$$if \;\; h_\theta(x) < 0.5 \quad : y = 0$$

$$if \;\; \theta^T x \ge 0 \; : y = 1$$
$$if \;\; \theta^T x < 0 \; : y = 0$$

نوعی محاسبه احتمال برای داده ورودی است

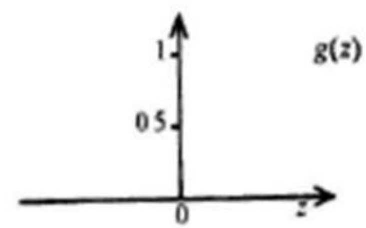
$h(x)$ = P(y=1|x;θ)=0.7

دسته بندی و مرز تفکیک

$\square(x) = g(\theta^T x) = P(y=1|x;\theta)$

$$h_\theta(x) = g(\theta^T x),$$

$$g(z) = \frac{۱}{۱+e^{-z}};\ z \in \mathbb{R}$$



"$y = ۱$" if $h_\theta(x) \geq ۰٫۵$

$g(z) \geq ۰٫۵$ whenever $z \geq ۰$;

Since $h_\theta(x) = g(\theta^T x) \rightarrow$

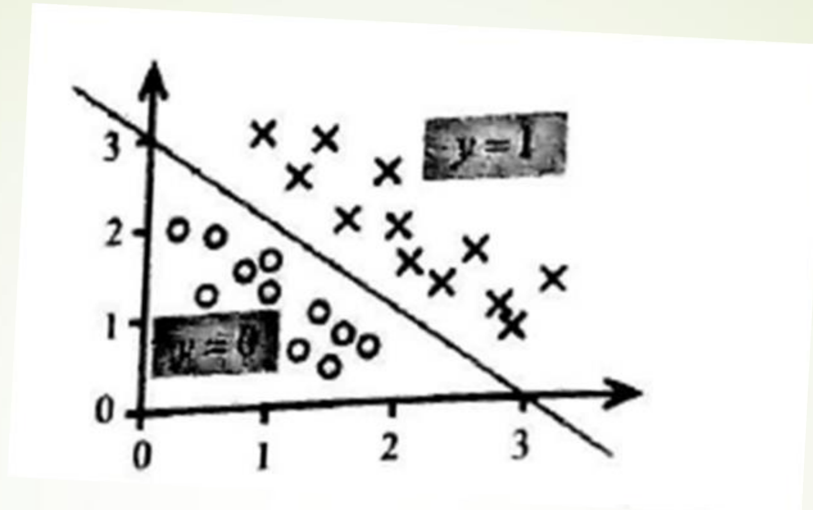
$h_\theta(x) \geq ۰٫۵$ whenever $\theta^T x \geq ۰$

"$y = ۰$" if $h_\theta(x) < ۰٫۵$

$g(z) < ۰٫۵$ whenever $z < ۰$;

Since $h_\theta(x) = g(\theta^T x) \rightarrow$

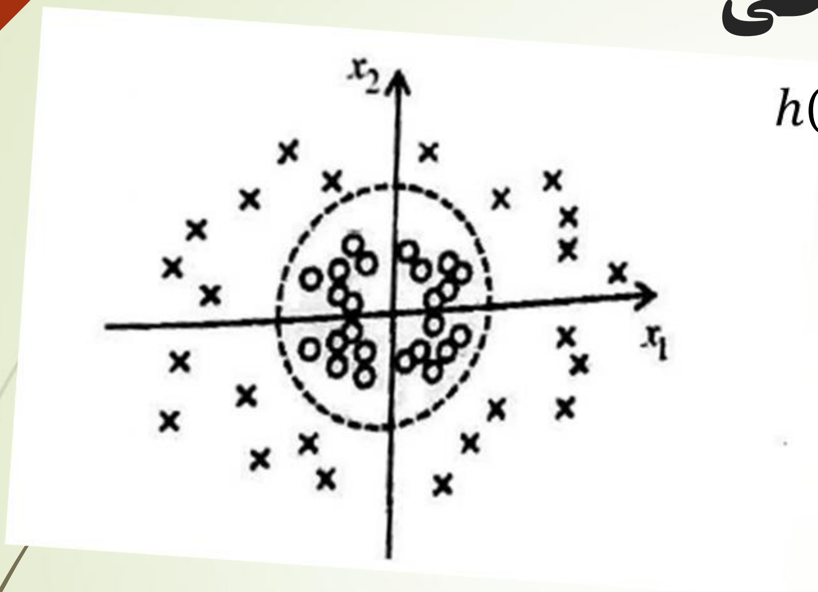
$h_\theta(x) < ۰٫۵$ whenever $\theta^T x < ۰$



$$h(x) = g(\theta_0 x_0 + \theta_1 x_1 + \theta_2 x_2)$$

$$\theta = \begin{matrix} -3 \\ +1 \\ +1 \end{matrix} \qquad -3 + x_1 + x_2 \geq 0 \qquad x_1 + x_2 \geq 3$$

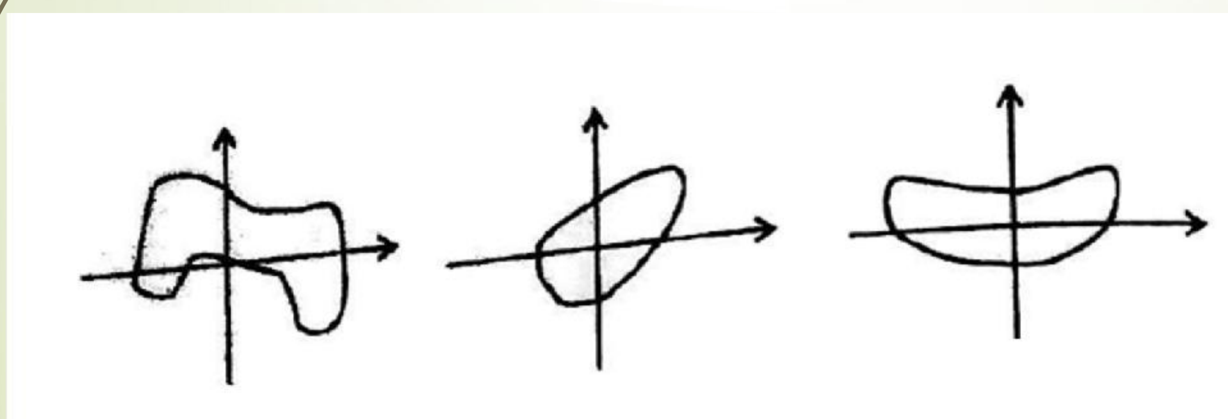
مرزبندی غیرخطی



$$h(x) = g(\theta_0 x_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_1^2 + \theta_4 x_2^2)$$

$$\theta = \begin{matrix} -1 \\ 0 \\ 0 \\ 1 \\ 1 \end{matrix}$$

$-1 + x_1^2 + x_2^2 \geq 0$ $\quad$ $x_1^2 + x_2^2 \geq 1$ $\quad$ y=1



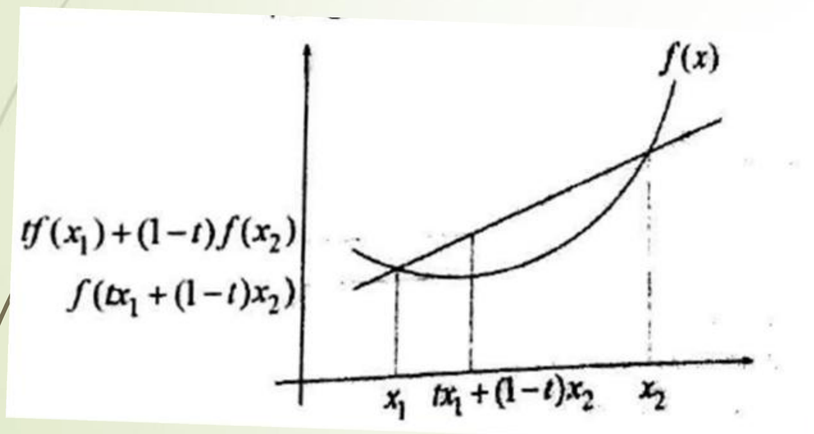
تابع هزینه رگرسیون لجستیک به این صورت غیر محدب است

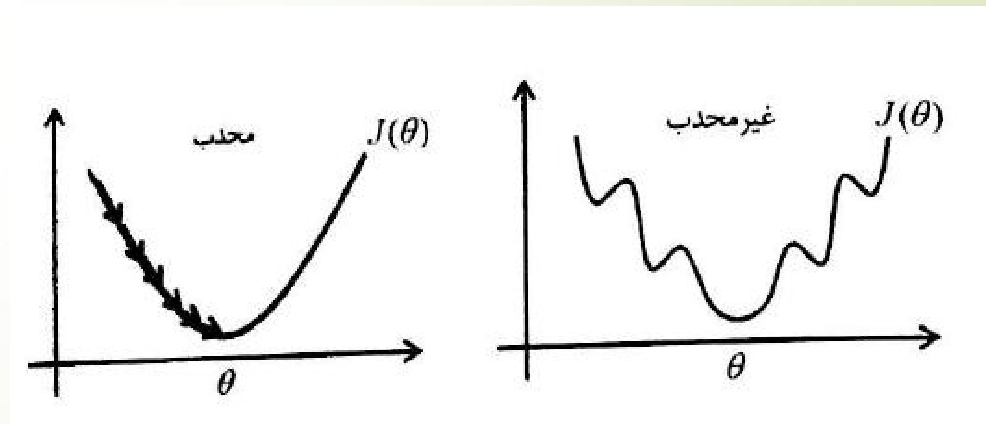
$$\mathrm{Cost}\left(h_\theta(x^{(i)}), y^{(i)}\right) = \frac{1}{2}\left(h_\theta(x^{(i)}) - y^{(i)}\right)^2$$

و به‌صورت ساده‌تر:

$$\mathrm{Cost}\left(h_\theta(x), y\right) = \frac{1}{2}\left(h_\theta(x) - y\right)^2$$

$$J(\theta) = \frac{1}{m}\sum_{i=1}^{m}\frac{1}{2}\left(h_\theta(x^{(i)}) - y^{(i)}\right)^2$$

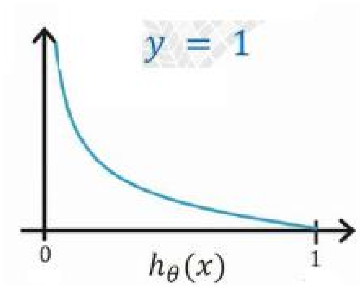




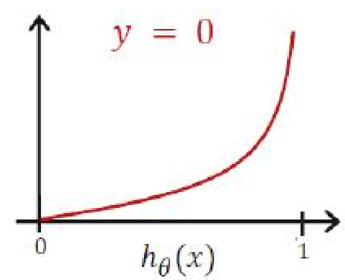
تابع هزینه

$$Cost(h_\theta(x), y) = \begin{cases} -\log(h_\theta(x)) & y = 1 \\ -\log(1 - h_\theta(x)) & y = 0 \end{cases}$$

- If y = 1
  - If h(x) = 0 : costs infinite
  - If h(x) = 1 : costs = 0



- If y = 0
  - If h(x) = 0 : costs = 0
  - If h(x) = 1 : costs infinite



تابع هزینه

$$Cost(h_\theta(x), y) = \begin{cases} -\log(h_\theta(x)) & y = 1 \\ -\log(1 - h_\theta(x)) & y = 0 \end{cases}$$

$$Cost(h_\theta(x), y) = -y\log(h_\theta(x)) - (1-y)\log(1 - h_\theta(x))$$

$$J(\theta) = \frac{1}{m}\sum_{i=1}^{m}\left[-y^{(i)}\log(h_\theta(x^{(i)})) - (1-y^{(i)})\log(1 - h_\theta(x^{(i)}))\right]$$

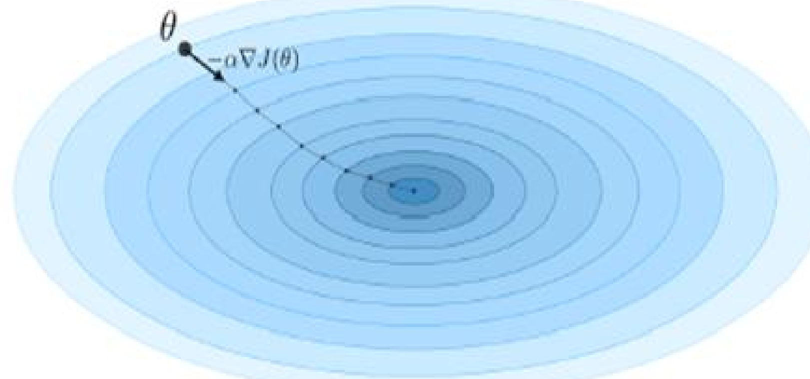
الگوریتم گرادیان کاهشی

تفاوت آن با الگوریتم رگرسیون خطی در تابع فرضیه است.

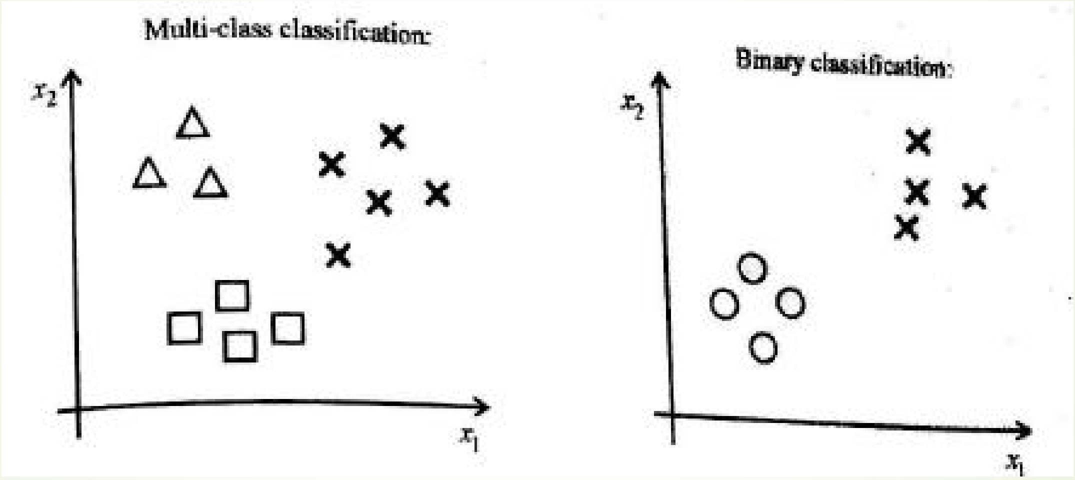
$$\frac{\partial}{\partial\theta_j}J(\theta)=\frac{1}{m}\sum_{I=1}^{m}(h_\theta(x^{(i)})-y^{(i)}).x_j^{(i)}$$

Repeat Until Convergence

$$\theta_j := \theta_j - \alpha\frac{\partial}{\partial\theta_j}J(\theta)$$



Multi-class classification



$$h_\theta^{(i)}(x) = P\left(y = i \mid x;\theta\right), \quad i = ۱,۲,۳$$

Multi-class classification