

یادگیری ماشین با پایتون

فرزاد حسینیان

دانشگاه محقق اردبیلی

دانشکده علوم کامپیوتر

آبان 98

جلسه سوم

رگريسيون لجسټيک

Binary Classification

هدف از کلاسبندی: نسبت دادن یک کلاس (لیبل، تگ) به داده ها

$$y \in \{0, 1\}$$



Multi-Class Classification

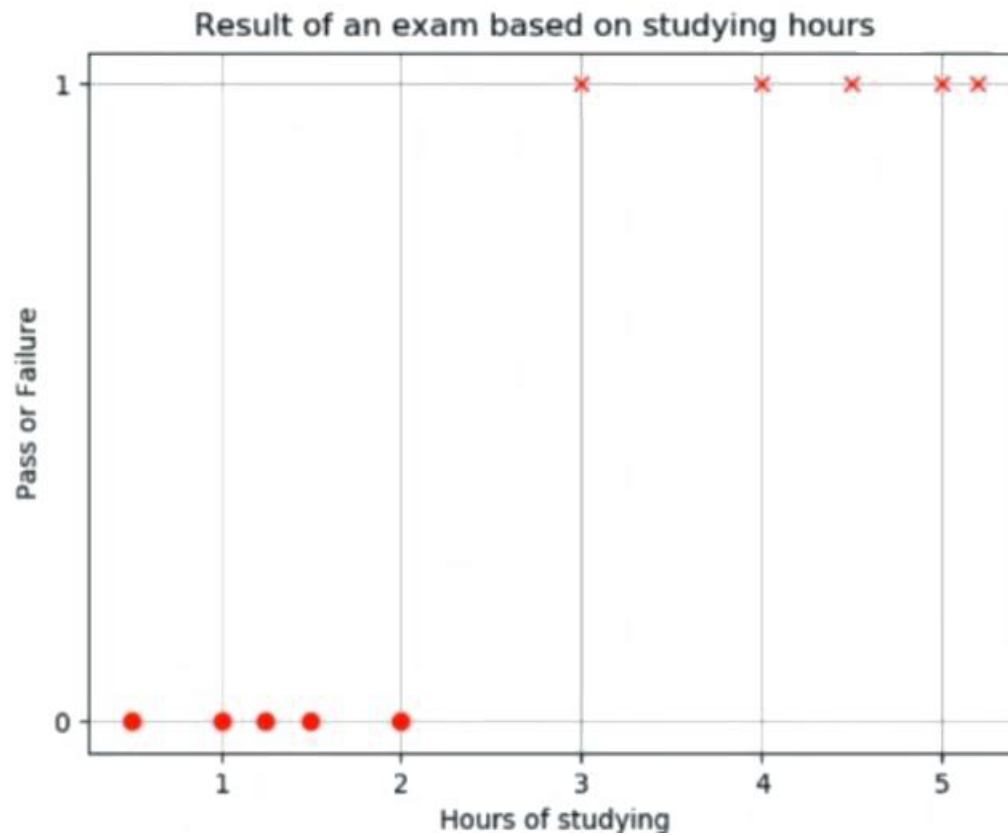
$$y \in \{0, 1, 2, 3, 4, \dots\}$$

- در کلاسبندی نتایج گسسته هستند

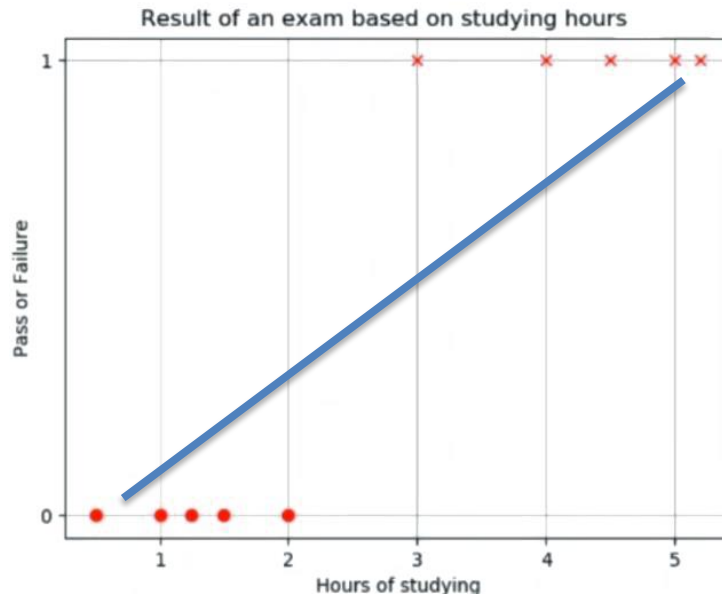
- در رگرسیون نتایج پیوسته اند

$$y = 1.23 \mid y = 3422.12 \mid y = 17 \mid y = 17.1 \mid y = 17.11 \mid y = 17.111 \mid \dots$$

مقادیر پیوسته اند



$$Y = \begin{cases} 0, & x < 0.5 \\ 1, & x \geq 0.5 \end{cases}$$



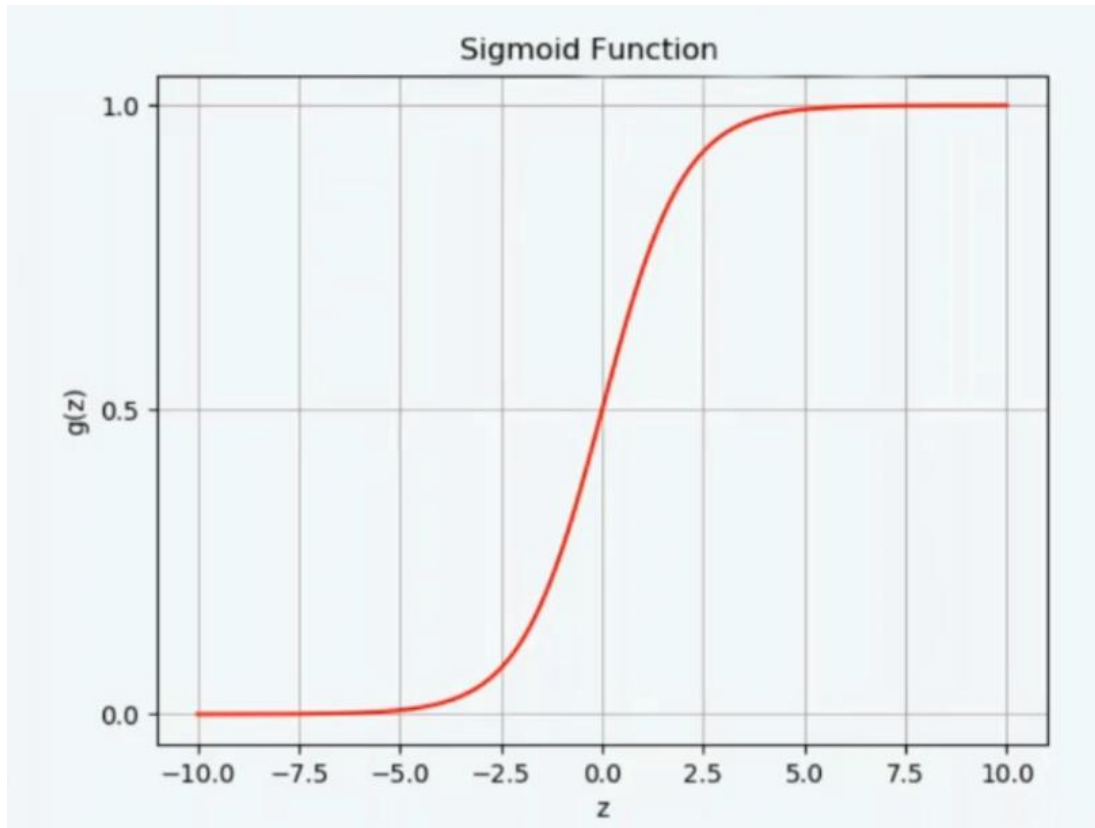
$$h_{\theta}(x) = \theta^T x$$

ایرادات:

- تغییرات شدید در صورت داشتن داده پرت
- محدود نبودن به بازه 0 و 1
- ...

تابع فرضیه

$$h_{\theta}(x) = g(\theta^T x)$$



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

عبارت زیر به چه معناست؟

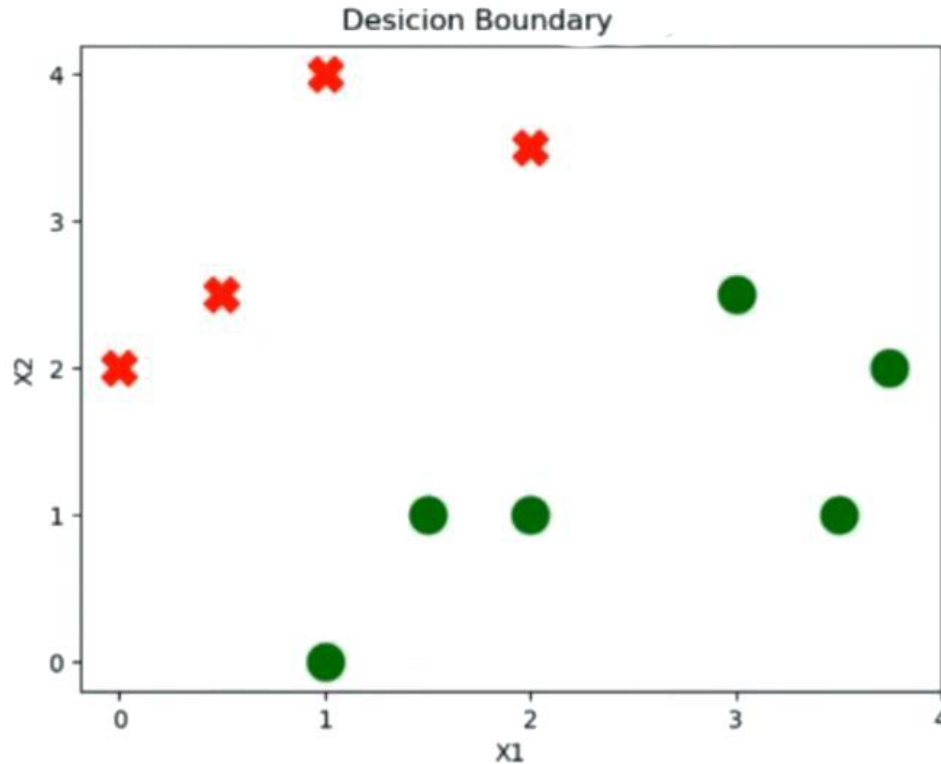
$$h_{\Theta}(x) = 0.87$$

عبارت زیر به چه معناست؟

$$h_{\theta}(x) = 0.87$$

یعنی با احتمال 87% این داده متعلق به دسته 1 است

مرز تصمیم – decision boundary



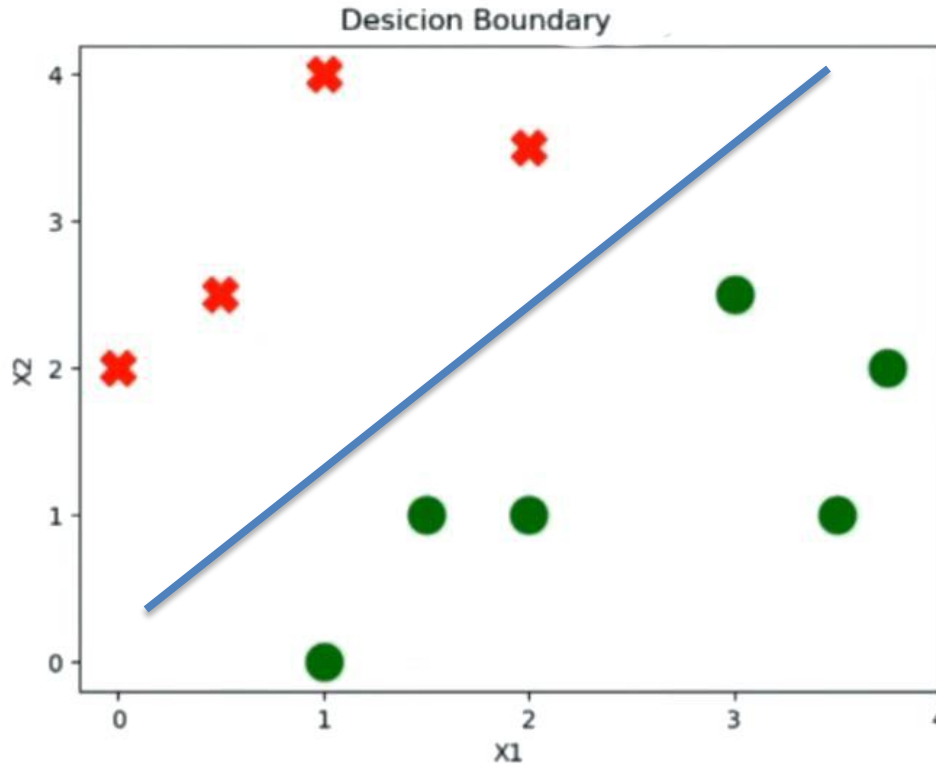
$$h_{\theta}(x) = g(\theta^T x)$$

$$\theta^T x \geq 0 \Rightarrow y = (1)$$

$$\theta^T x < 0 \Rightarrow y = (0)$$

$$\theta_2 x_2 + \theta_1 x_1 + \theta_0$$

مرز تصمیم – decision boundary



$$h_{\theta}(x) = g(\theta^T x)$$

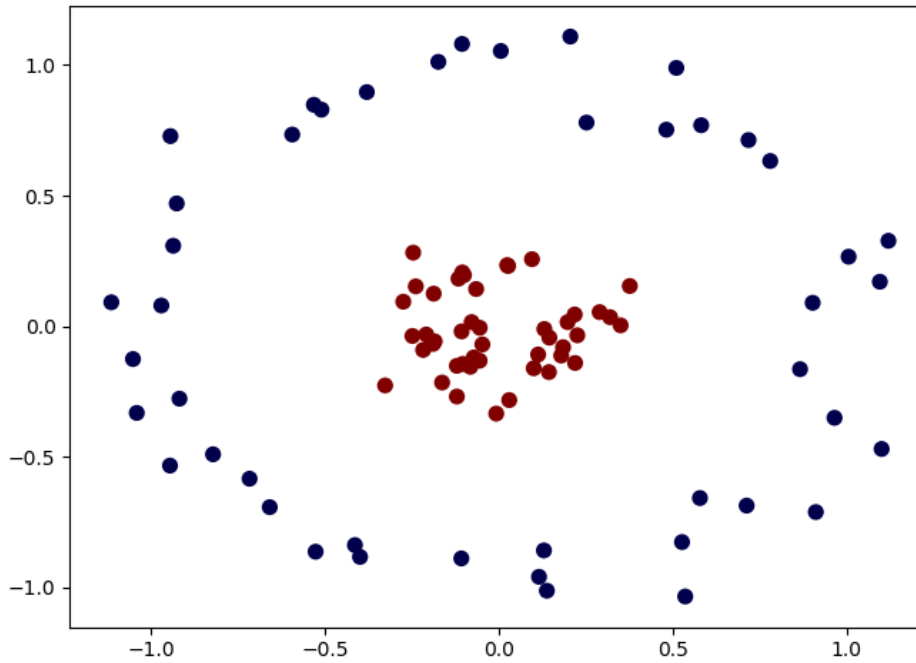
$$\theta^T x \geq 0 \Rightarrow y = (1)$$

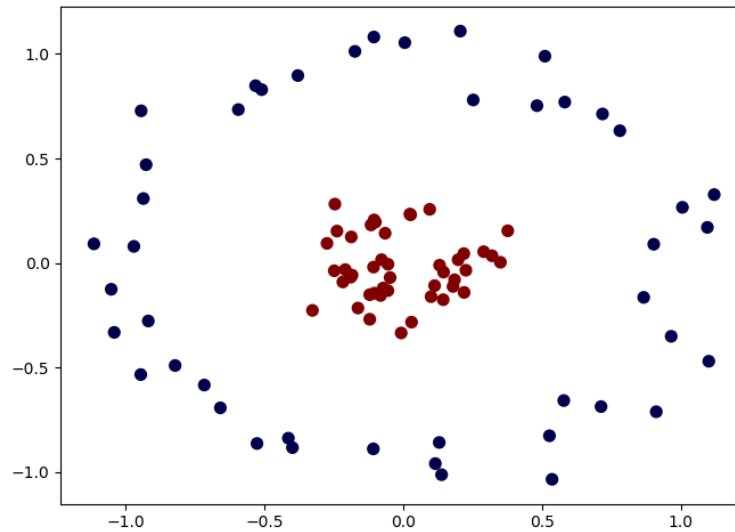
$$\theta^T x < 0 \Rightarrow y = (0)$$

$$\theta_2 x_2 + \theta_1 x_1 + \theta_0$$

مرز تصمیم

$$\theta^T X = 0$$

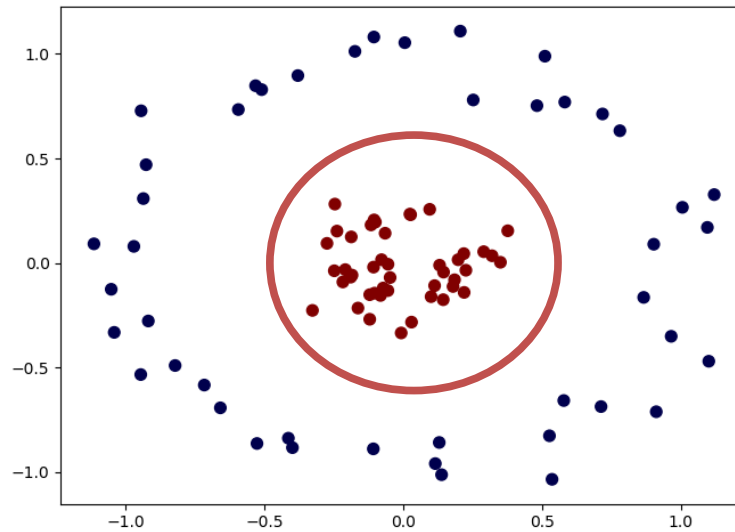




$$h_{\theta}(x) = g(\overset{-1}{\downarrow} \theta_0 + \overset{0}{\downarrow} \theta_1 x_1 + \overset{0}{\downarrow} \theta_2 x_2 + \overset{0}{\downarrow} \theta_3 x_1 x_2 + \overset{1}{\uparrow} \theta_4 x_1^2 + \overset{1}{\uparrow} \theta_5 x_2^2)$$

$$x_1^2 + x_2^2 - 1 = 0$$

$$x_1^2 + x_2^2 = 1$$



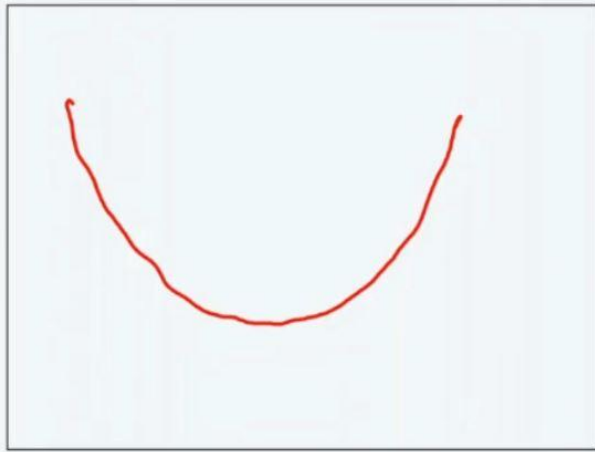
$$h_{\theta}(x) = g(\overset{-1}{\downarrow} \theta_0 + \overset{0}{\downarrow} \theta_1 x_1 + \overset{0}{\downarrow} \theta_2 x_2 + \overset{0}{\uparrow} \theta_3 x_1 x_2 + \overset{1}{\uparrow} \theta_4 x_1^2 + \overset{1}{\uparrow} \theta_5 x_2^2)$$

$$x_1^2 + x_2^2 - 1 = 0$$

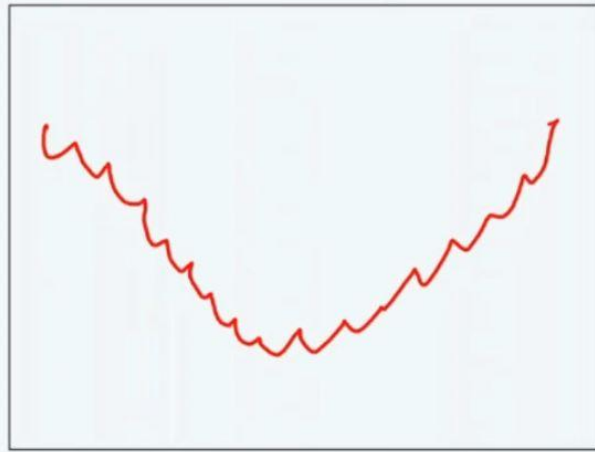
$$x_1^2 + x_2^2 = 1$$

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

تابع هزینه



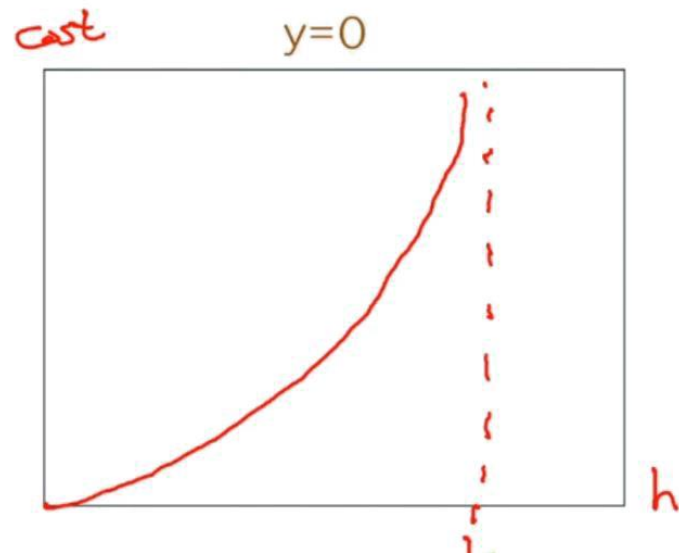
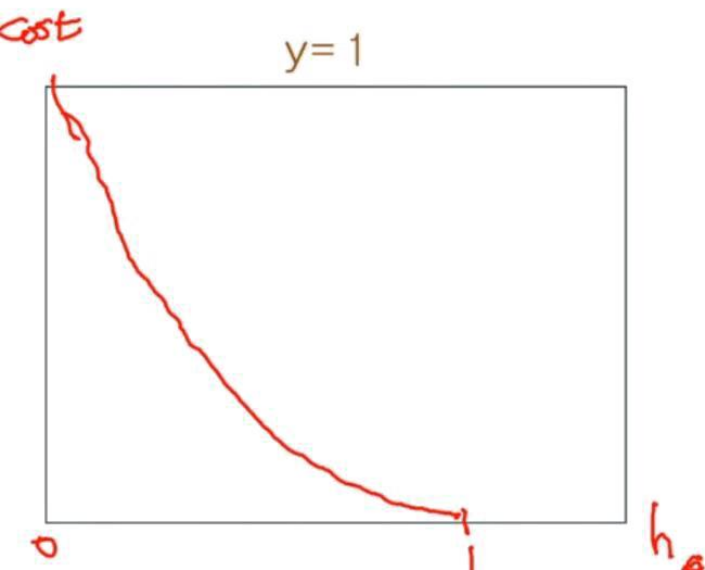
Convex



Non-Convex

تابع هزینه

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



تابع هزینه

$$J(\theta) = \frac{1}{m} \sum_1^m \text{Cost}(h_{\theta}(x^{(i)}), y^{(i)})$$

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

Gradient descent algorithm

repeat until convergence {
 $\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1)$
 (for $j = 1$ and $j = 0$)
}

$$\theta_j := \theta_j - \alpha \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)}) x_j^{(i)}$$