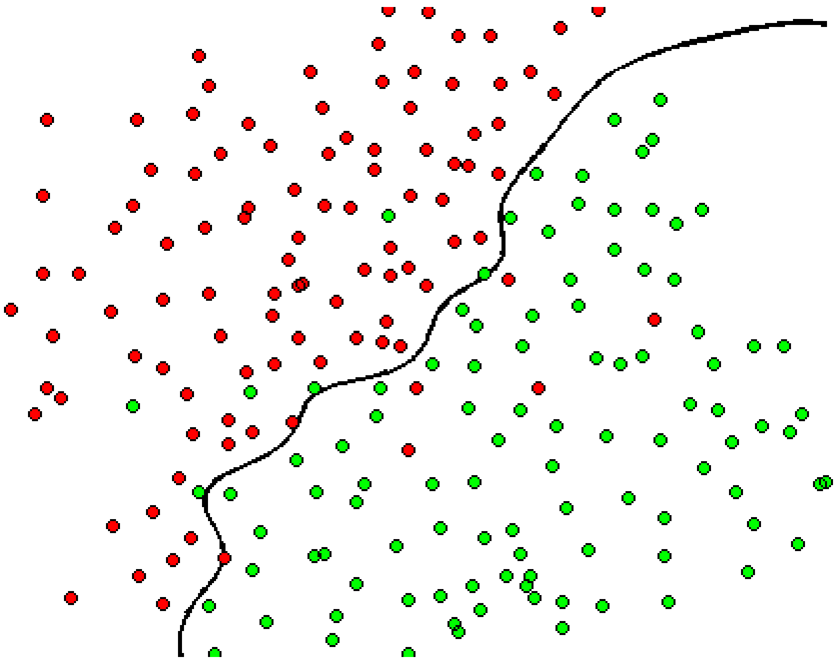
法律声明

- 本课件包括：演示文稿，示例，代码，题库，视频和声音等，小象学院拥有完全知识产权的权利；只限于善意学习者在本课程使用，不得在课程范围外向任何第三方散播。任何其他人或机构不得盗版、复制、仿造其中的创意，我们将保留一切通过法律手段追究违反者的权利。

关注 **小象学院**



分类模型(2)

--Robin

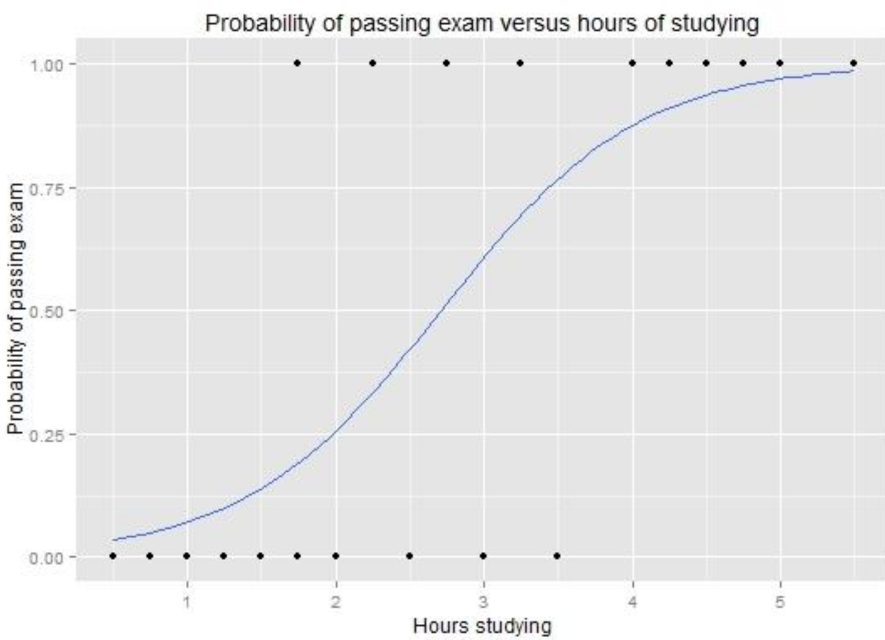
目录

- 逻辑回归
- 正则化
- 支持向量机

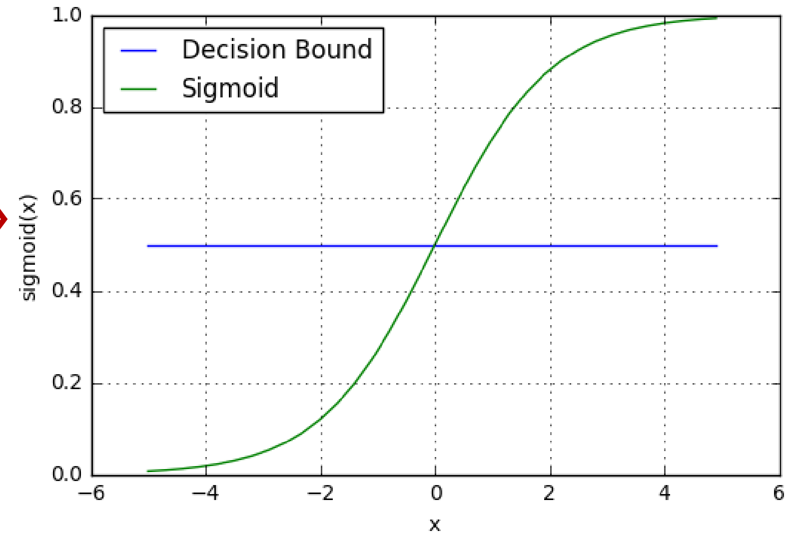
逻辑回归

Q: 现有20个学生投入0-6个小时学习课程的记录，分析投入的时间和是否通过考试的概率的关系。

Hours	0.50	0.75	1.00	1.25	1.50	1.75	1.75	2.00	2.25	2.50	2.75	3.00	3.25	3.50	4.00	4.25	4.50	4.75	5.00	5.50
Pass	0	0	0	0	0	0	1	0	1	0	1	0	1	0	1	1	1	1	1	1

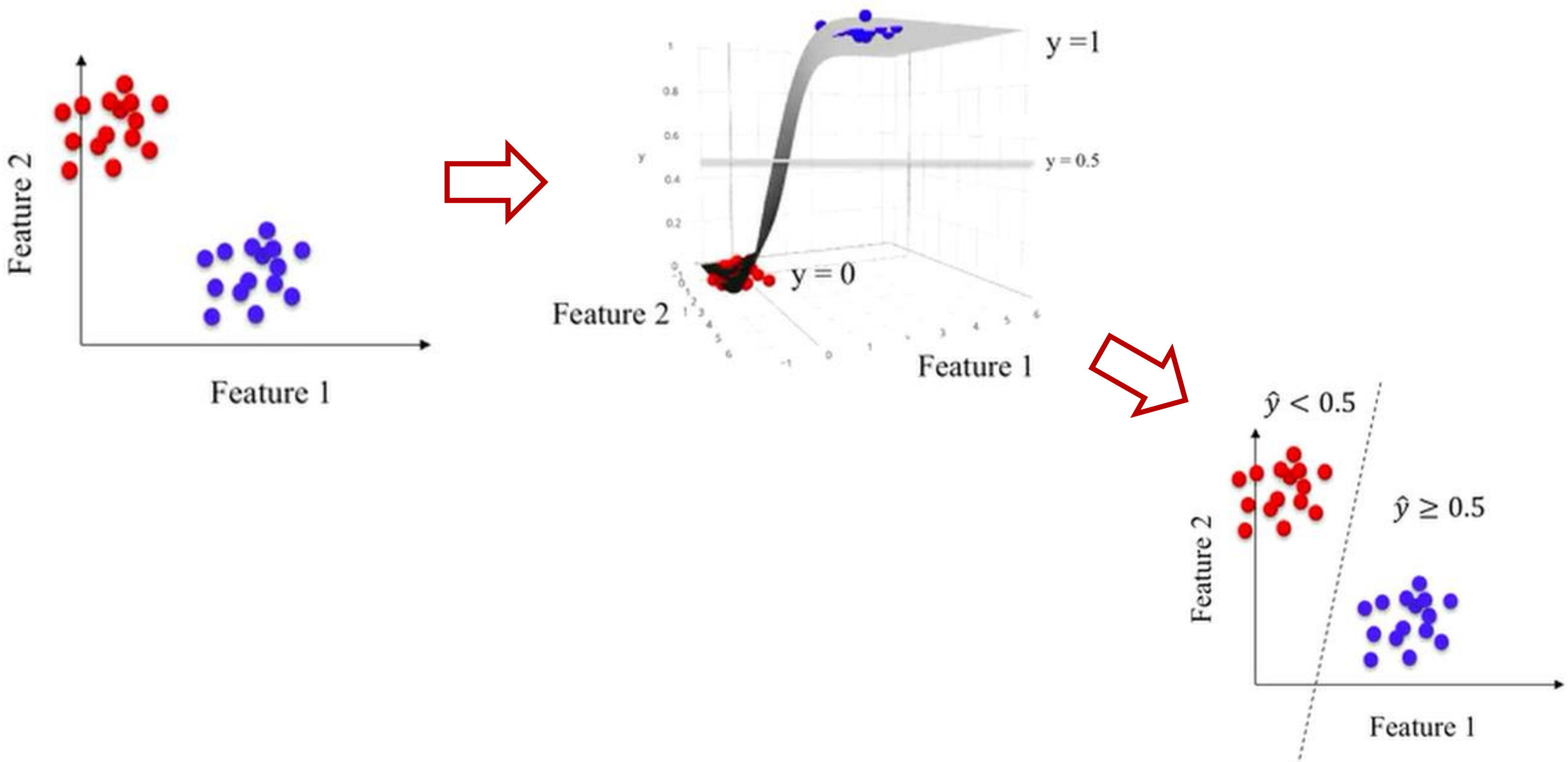


$$f(x) = \frac{1}{1+e^{-x}}$$



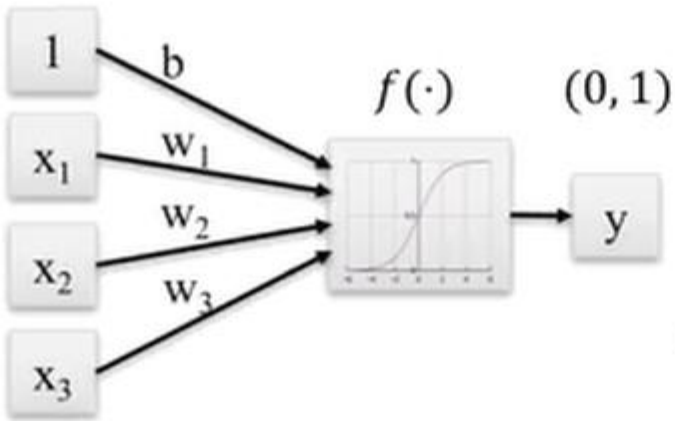
逻辑回归

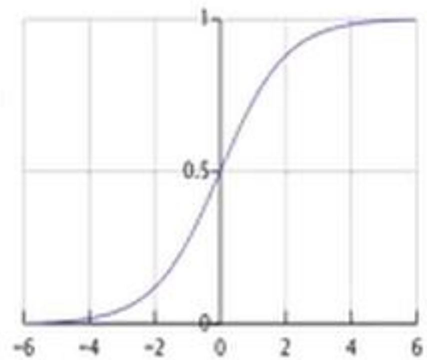
- 例子：样本中包含二维特征



逻辑回归

Input features





The logistic function transforms real-valued input to an output number y between 0 and 1, interpreted as the probability the input object belongs to the positive class, given its input features $(x_0, x_1, \ldots, x_n)$

$$\hat{y} = \text{logistic}(\hat{b} + \hat{w}_1 \cdot x_1 + \cdots \hat{w}_n \cdot x_n) = \frac{1}{1 + \exp\left[-(\hat{b} + \hat{w}_1 \cdot x_1 + \cdots \hat{w}_n \cdot x_n)\right]}$$

$$h_\theta(X) = \frac{1}{1 + e^{-\theta^T X}} = Pr(Y = 1|X;\theta)$$

$$Pr(Y = 0|X;\theta) = 1 - h_\theta(X)$$

逻辑回归

Logistic Regression中的损失函数

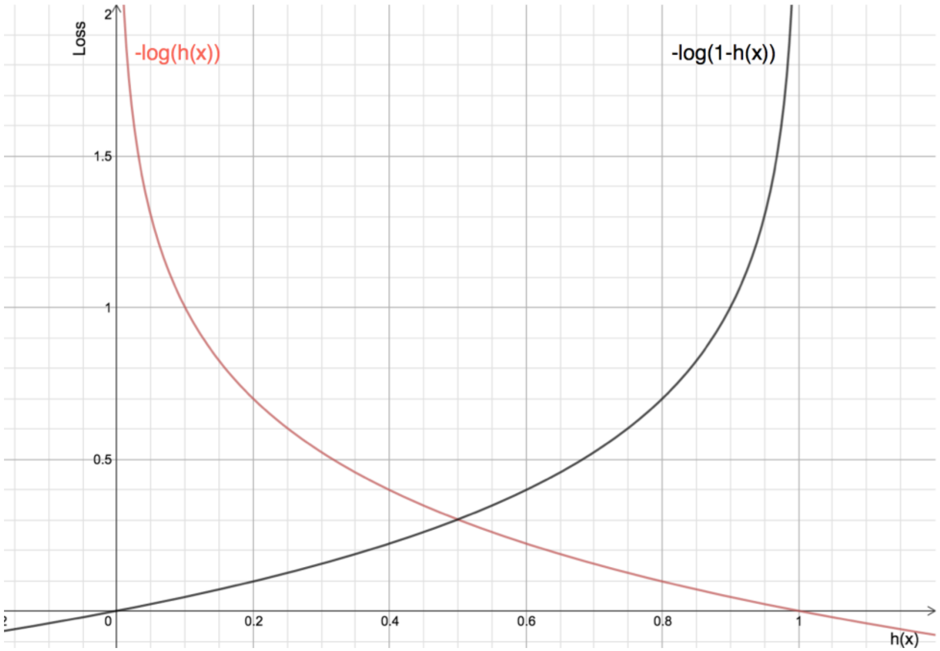
$$J(\theta)=\frac{1}{m}\sum_{i=1}^{m}\mathrm{Cost}(h_\theta(x^{(i)}),y^{(i)})$$

$$\mathrm{Cost}(h_\theta(x),y)=-\log(h_\theta(x)) \qquad \text{if } y=1$$
$$\mathrm{Cost}(h_\theta(x),y)=-\log(1-h_\theta(x)) \qquad \text{if } y=0$$

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

梯度下降求解参数

cross-entropy loss或logloss

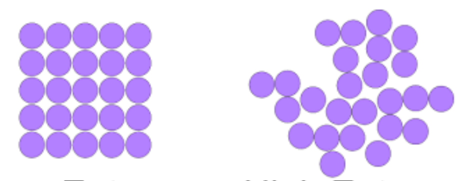


逻辑回归

- 熵（Entropy）

在信息论中，设离散随机变量 X 的概率分布为 ，则概率分布 $P(X=x^{(i)})=p_i,\ i=1,2,\cdots,n$ 的熵(Entropy)的定义为：

$$H(p)=-\sum_{i=1}^{n}p_i\log p_i$$



- 交叉熵（Cross Entropy）

关于同一组事件 $x^{(1)},\cdots,x^{(n)}$ 的两个分布 p,q 其交叉熵(Cross-Entropy)的定义如下：

$$H(p,q)=-\sum_{i=1}^{n}p_i\log q_i$$

当两个分布完全相同时，交叉熵取最小值。

交叉熵可以衡量两个分布之间的相似度，交叉熵越小两个分布越相似。

Softmax Regression

二分类问题

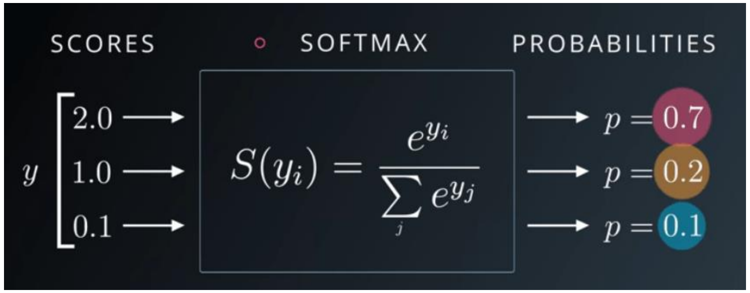
- sigmoid function $$h_\theta(x) = \frac{1}{1+\exp(-\theta^T x)},$$

$$h_\theta(y=1|x;\theta)=\frac{1}{1+e^{-\theta^T x}}=\frac{e^{\theta^T x}}{e^{\theta^T x}+1}=\frac{e^{y=1}}{e^{y=0}+e^{y=1}}$$

$$h_\theta(y=0|x;\theta)=1-h_\theta(y=1|x;\theta)=\frac{1}{e^{\theta^T x}+1}=\frac{e^{y=0}}{e^{y=1}+e^{y=0}}$$

多分类问题

- softmax function



联系我们

小象学院：互联网新技术在线教育领航者

– 微信公众号：**小象学院**