



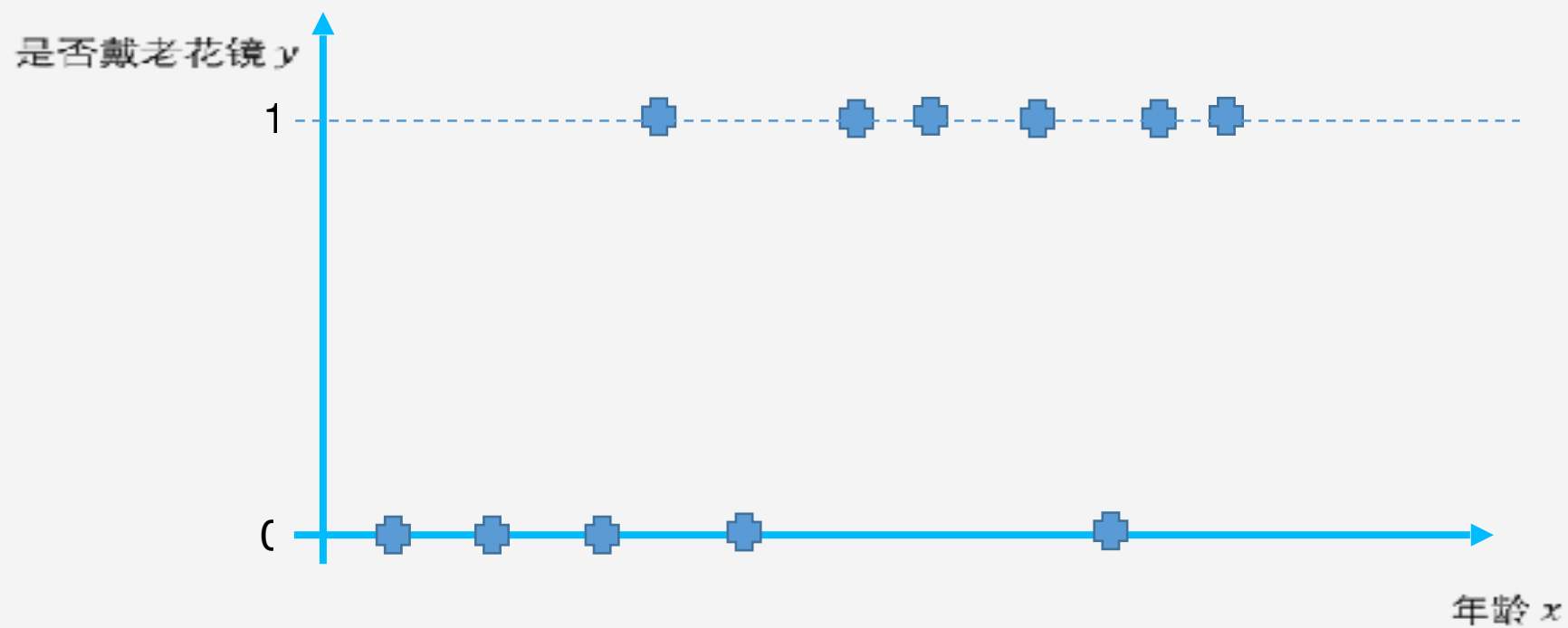
逻辑回归

贪心学院讲师：袁源

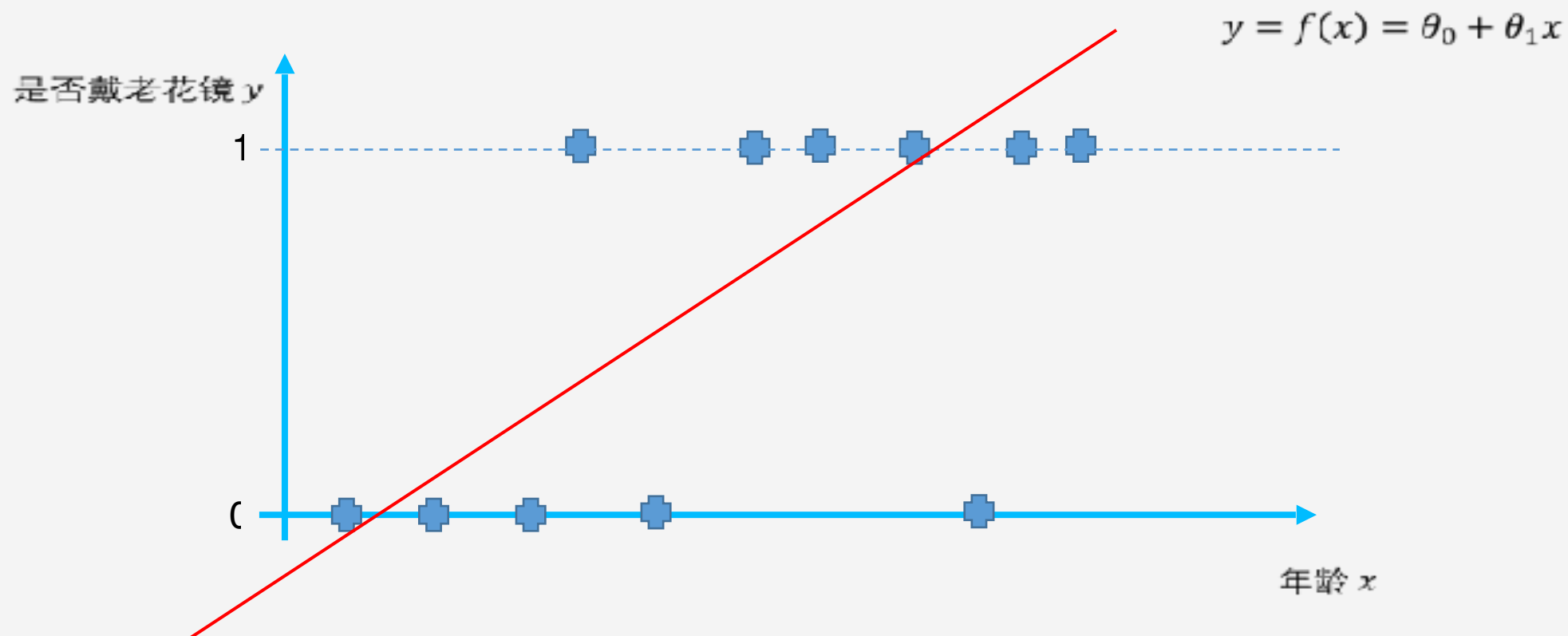
讲师简介： 贪心科技技术负责人，博士毕业于美国新泽西理工，拥有14年人工智能、推荐系统、自然语言处理、数字图像和视频处理项目经验；现任职于美国微软总部，曾任职于美国亚马逊总部、主导多款核心推荐系统的研发，是人工智能、分布式系统、云计算方面的专家。



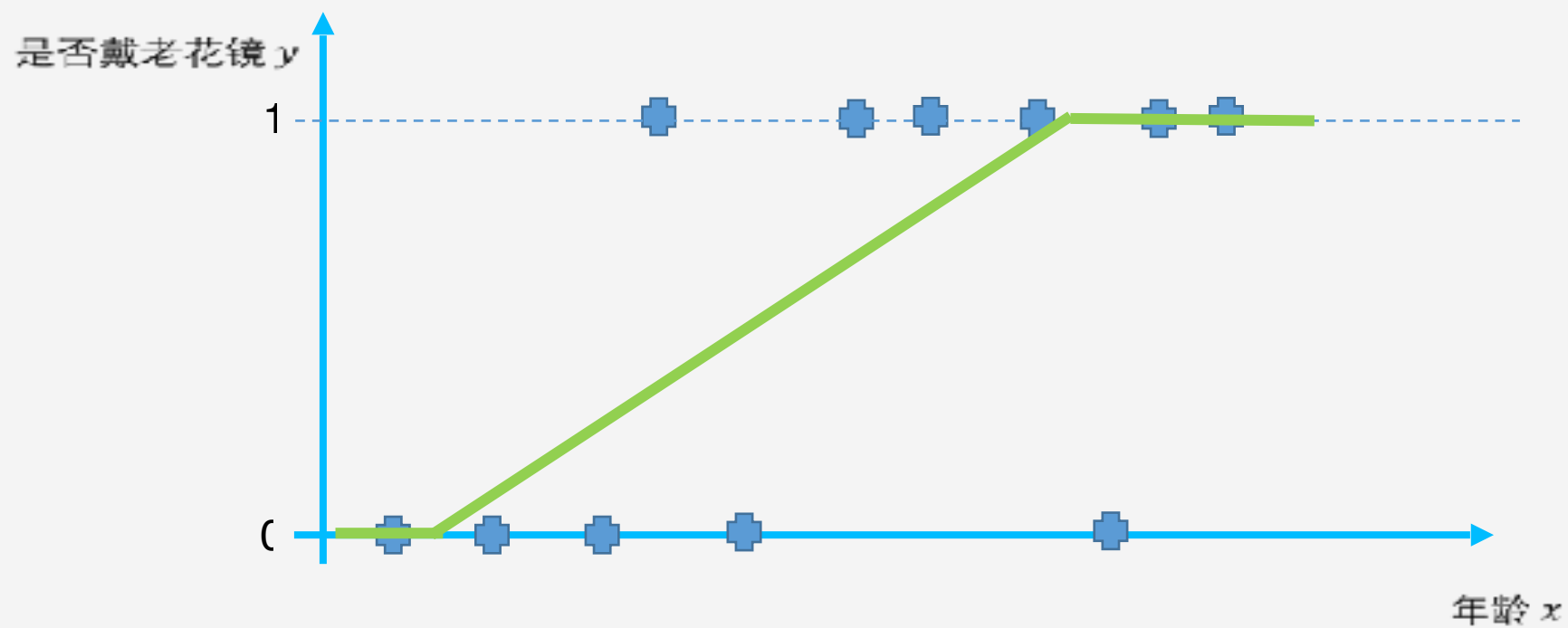
二元分类问题



二元分类问题



二元分类问题

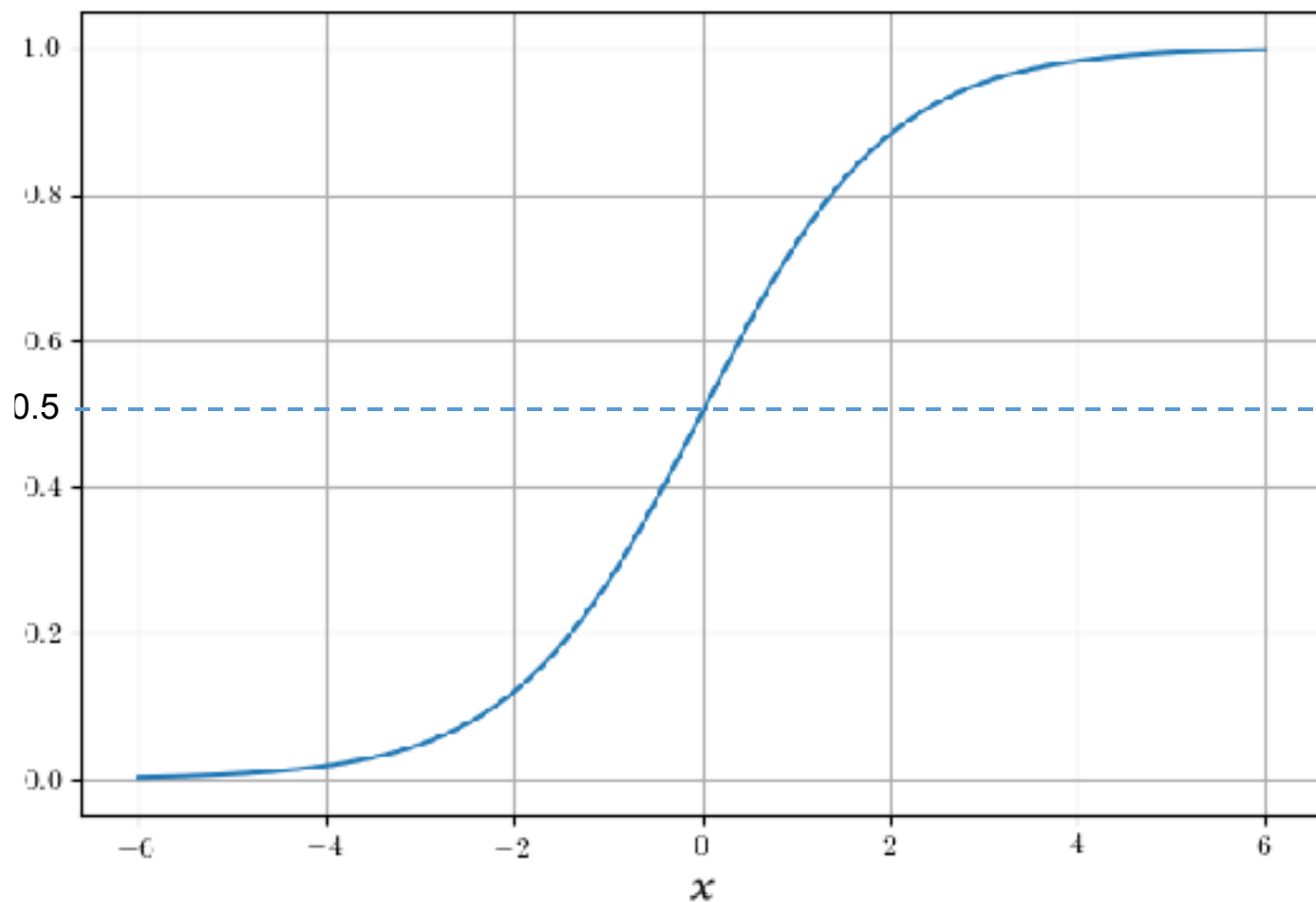


逻辑函数



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

$$\frac{1}{1+e^{\infty}} = \frac{1}{1+\infty} = 0$$

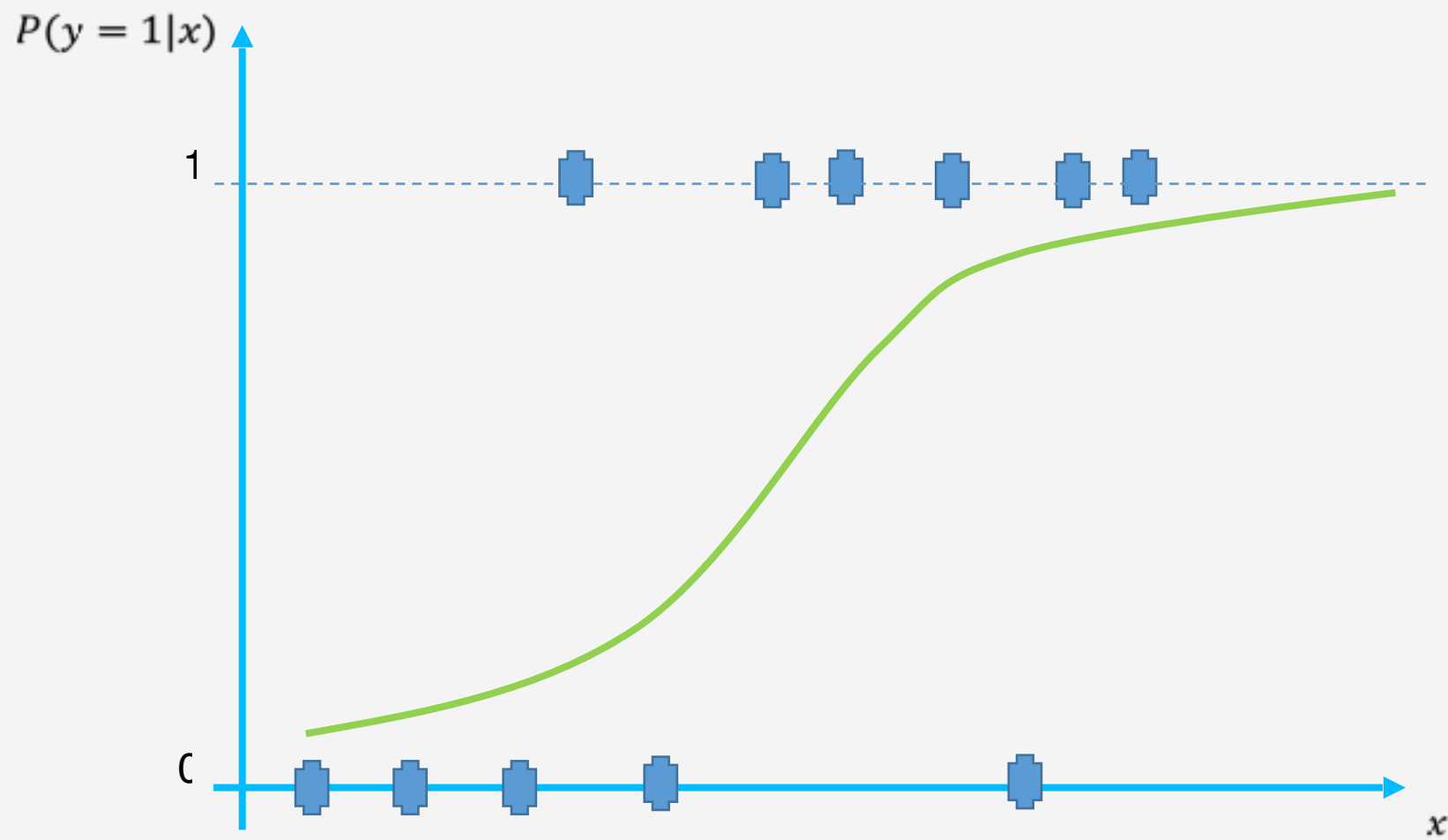


$$\frac{1}{1+e^{-\infty}} = \frac{1}{1+0} = 1$$

逻辑回归



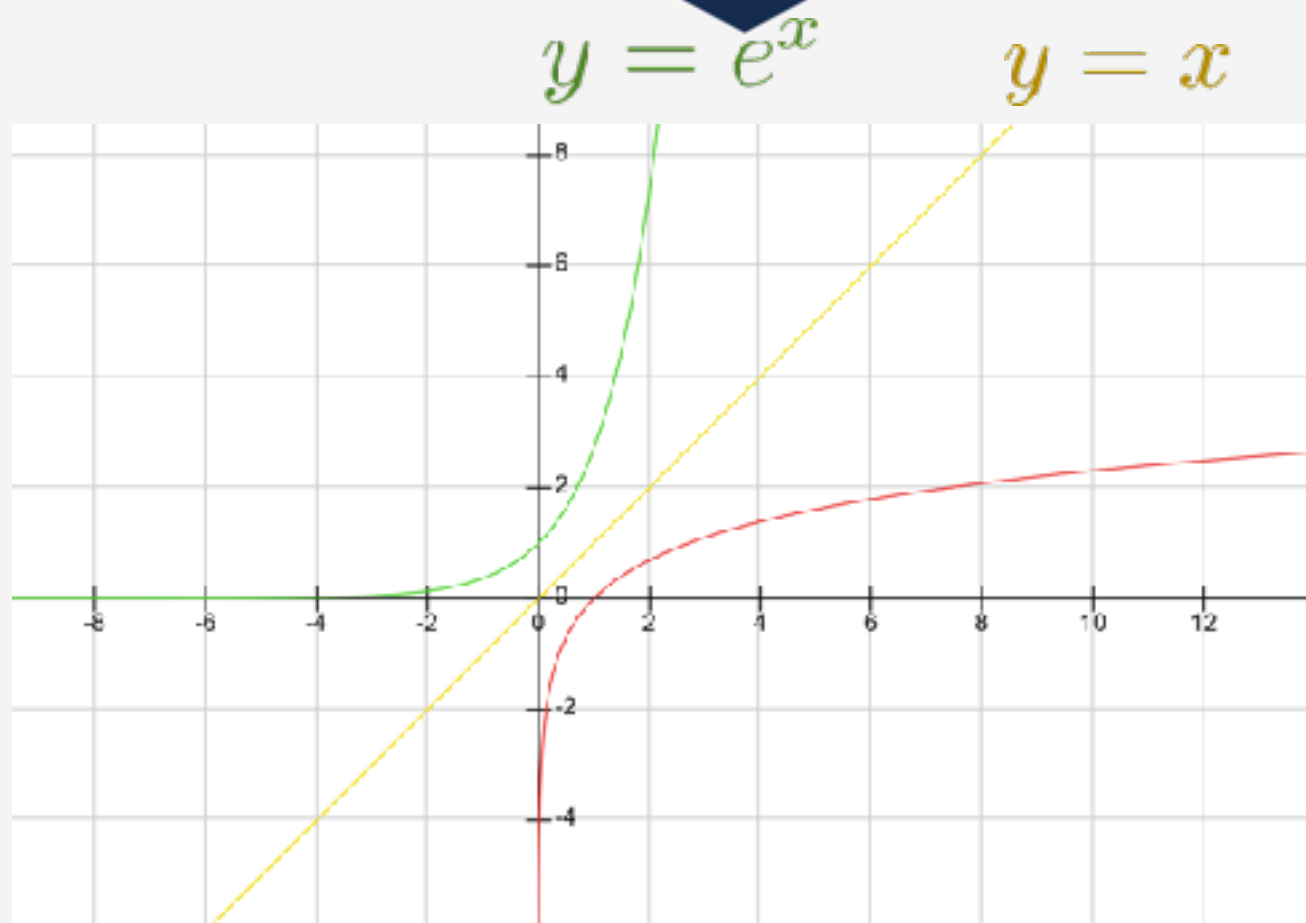
$$P(y = 1|x) = f(x) = \frac{1}{1 + e^{-(\theta_0 + \theta_1 x)}}$$



指数与对数



$$\begin{aligned}e^0 &= 1 \\e^1 &= e \approx 2.72 \\e^{-1} &= \frac{1}{e} \\e^{0.693} &\approx 2 \\e^{\ln x} &= x\end{aligned}$$



$$y = \ln x$$

$$\begin{aligned}\ln 1 &= 0 \\ \ln e &= 1 \\ \ln \frac{1}{e} &= -1 \\ \ln 2 &\approx 0.693 \\ \ln e^x &= x\end{aligned}$$

逻辑回归

- 解决二元 (0/1) 分类的问题
- $P(y = 1|x; \theta) = f(x; \theta) = \frac{1.0}{1.0 + e^{-\theta^T x}}$
- $\theta^T x = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 + \dots$
- $\theta = [\theta_0, \theta_1, \theta_2, \theta_3, \dots]$
- $x = [1, x_1, x_2, x_3, \dots]$
- 当 $P(y = 1|x)$ 的值大于 0.5, 输出 1 ; 否则输出 0

逻辑回归知识点

类别1的概率

$$P = \frac{1}{1 + e^{-(\theta^T x)}}$$

类别0的概率

$$1 - P = \frac{1 + e^{-\theta^T x} - 1}{1 + e^{-\theta^T x}} = \frac{1}{1 + e^{\theta^T x}}$$

类别1与0概率比值

$$\frac{P}{1 - P} = e^{\theta^T x}$$

类别1与0概率比值的
自然对数

$$\ln \frac{P}{1 - P} = \theta^T x$$

示例



年龄 (岁)	年收入 (万元/年)	是否买车 (1表示是, 0表示否)	年龄 (岁)	年收入 (万元/年)	是否买车 (1表示是, 0表示否)	是否买车 (1表示是, 0表示否)
20	3	0	20	3	0	是否买车 (1表示是, 0表示否)
23	7	1	23	7	1	
31	10	1	31	10	1	
42	13	1	42	13	1	
50	7	0	50	7	0	
60	5	0	60	5	0	
20	3	0				
23	7	1				
31	10	1				
42	13	1				
50	7	0				
60	5	0				
28	8	?				

示例

年龄 (岁)	年收入 (万元/年)	是否买车 (1表示是, 0表示否)	年龄 (岁)	年收入 (万元/年)	是否买车 (1表示是, 0表示否)
20	3	0	20	3	0
23	7	1	23	7	1
31	10	1	31	10	1
42	13	1	42	13	1
50	7	0	50	7	0
60	5	0	60	5	0
20	3	0			
23	7	1			
31	10	1			
42	13	1			
50	7	0			
60	5	0			
28	8	?			

$$P(Y = 1|x; \theta) = f(x; \theta) = \frac{1}{1 + e^{-\theta^T x}}$$

$$\theta_0 = -0.04, \theta_1 = -0.20, \theta_2 = 0.92$$

$$\theta^T x = -0.04 - 0.20 * 28 + 0.92 * 8 = 1.75$$

$$P(y = 1|x) = \frac{1.0}{1.0 + e^{-1.75}} = 0.85$$

模型训练

年龄 (岁)	年收入 (万元/年)	是否买车 (1表示是, 0表示否)	年龄 (岁)	年收入 (万元/年)	是否买车 (1表示是, 0表示否)
20	3	0	20	3	0
23	7	1	23	7	1
31	10	1	31	10	1
42	13	1	42	13	1
50	7	0	50	7	0
60	5	0	60	5	0

$$P(y = 1|x) = \frac{1.0}{1.0 + e^{-\theta^T x}}$$

如何根据左边的训练数据得到系数的值：

$$\theta_0 = -0.04, \theta_1 = -0.20, \theta_2 = 0.92$$

损失函数

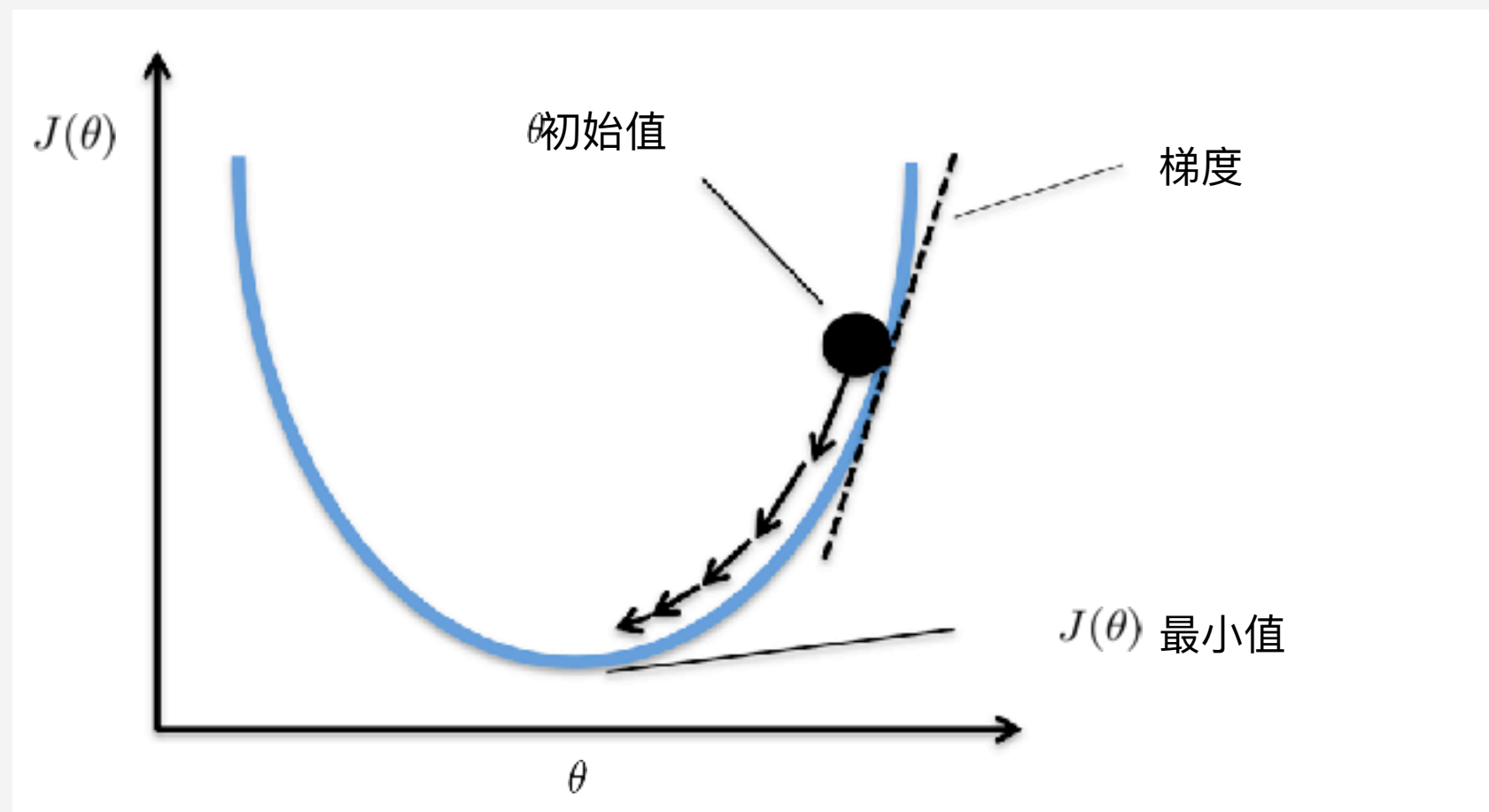


损失函数

$$P(Y = 1|x; \theta) = f(x; \theta) = \frac{1}{1 + e^{-\theta^T x}}$$

$$J(\theta) = - \sum_{i=1}^N y^{(i)} \ln(P(Y = 1|X = x^{(i)}; \theta)) + (1 - y^{(i)}) \ln(1 - P(Y = 1|X = x^{(i)}; \theta))$$

$$\nabla_{\theta} J(\theta) = \sum_i x^{(i)} (f(x^{(i)}; \theta) - y^{(i)})$$



梯度下降法



梯度下降法

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

$$\nabla_{\theta} J(\theta) = \sum_i x^{(i)} (f(x^{(i)}; \theta) - y^{(i)})$$

$$\theta = \theta - \alpha \nabla_{\theta} J(\theta) = \theta - \alpha \sum_i x^{(i)} (f(x^{(i)}; \theta) - y^{(i)})$$

新的值 旧的值 学习率 特征的值 预测类别1的概率 已知类别的值



系数的意义

$$P(Y = 1|x; \theta) = f(x; \theta) = \frac{1}{1 + e^{-\theta^T x}}$$

$$\text{概率比值 } odds = \frac{p}{1-p} = e^{\theta^T x}$$

系数 θ_j 意味着，假设原来的 $odds$ 为 λ_1 ，若对应的特征 x_j 增加1，假设新的 $odds$ 为 λ_2 ，那么 $\frac{\lambda_2}{\lambda_1} \equiv e^{\theta_j}$



模型训练

年龄 (x ₁)	年收入 (x ₂) 万元为単位!	是否买车 (y) 0表示否, 1表示是	年龄 (x ₁)	年收入 (x ₂) 万元为単位!	是否买车 (y) 0表示否, 1表示是
20	3	0	20	3	0
23	7	1	23	7	1
31	10	1	31	10	1
42	13	1	42	13	1
50	7	0	50	7	0
60	5	0	60	5	0

是否买车 (1
表示是, 0表
示否)

$$P(Y = 1|x; \theta) = \frac{1.0}{1.0 + e^{-\theta^T x}}$$

如何根据左边的训练数据得到系数 θ 的值：

$$\theta_0 = -0.04, \theta_1 = -0.20, \theta_2 = 0.92$$

系数 $\theta_2 = 0.92$ 意味着, 如果年收入增加1万, 一个人买车和不买车的概率的比值与之前的比值相比较, 增加 $e^{0.92} = 2.5$ 倍

系数 $\theta_1 = -0.20$ 意味着, 如果年龄增加1岁, 一个人买车和不买车的概率的比值与之前的比值比较降低 $e^{-0.20} = 0.82$ 倍

应用场景



- 垃圾邮件分类
- 广告点击预测
- 医疗效果预测



贪心科技 让每个人享受个性化教育服务

THANKS

贪心学院讲师：袁源