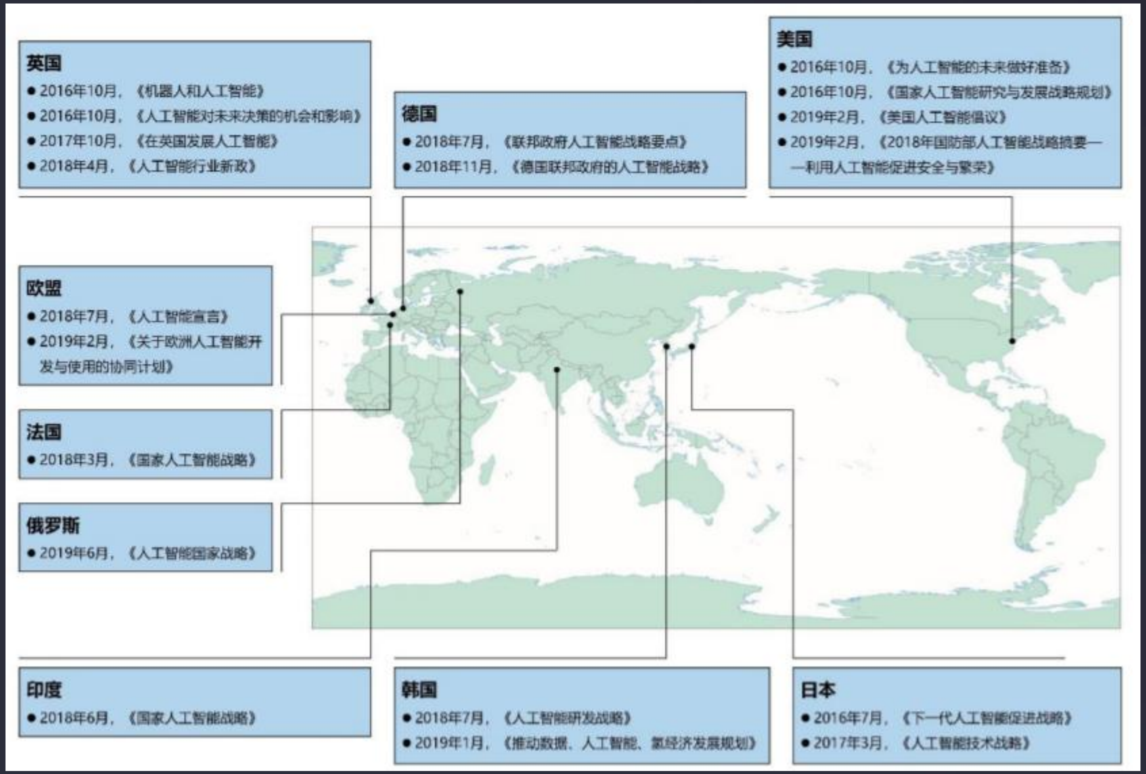
EDU

CSDN学院 IT实战派

精炼数学小课

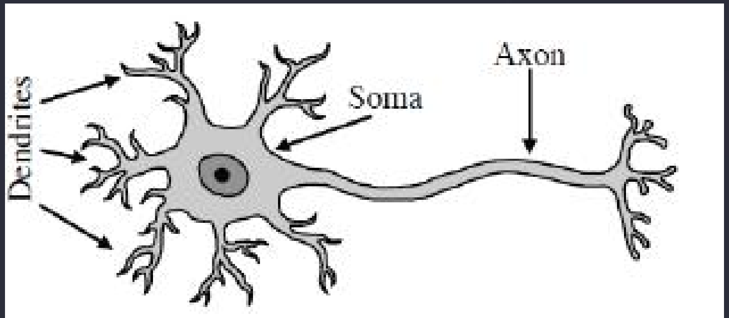
一星期理解人工智能的底层逻辑

人工智能的时代已经来临

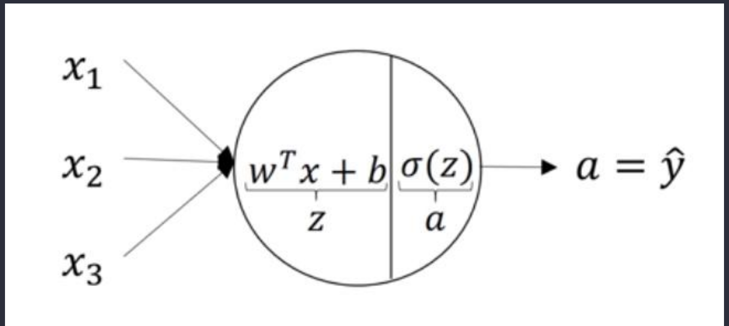


人工智能的底层技术：深度学习神经网络

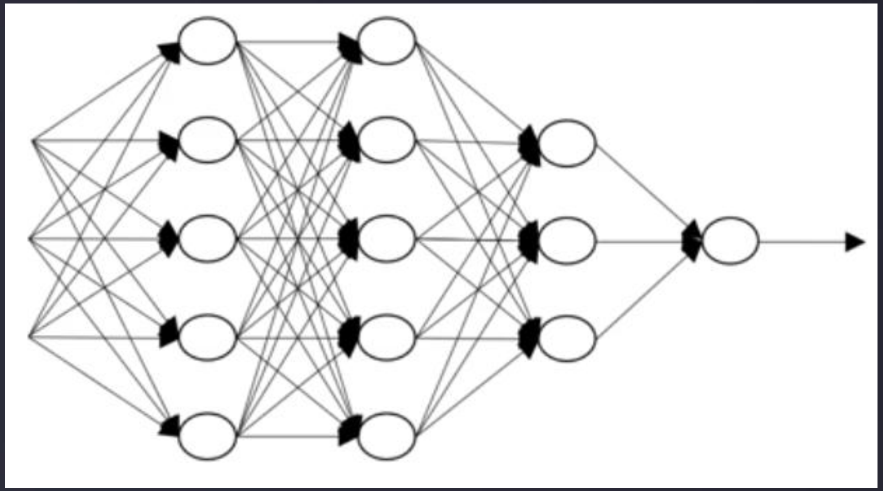
生物神经元



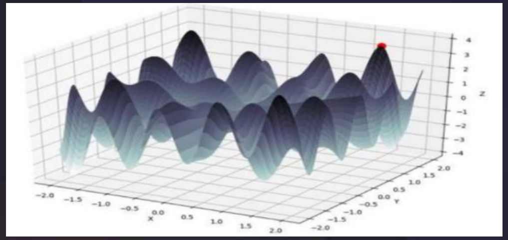
人工神经元



神经网络（MLP,CNN,RNN）



大数据分析



计算视觉

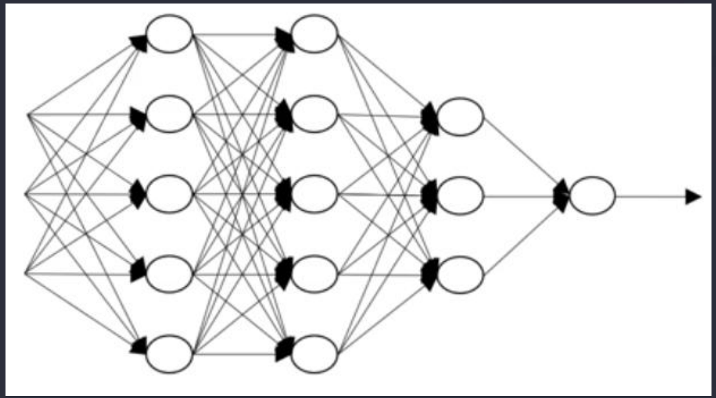
自然语言处理

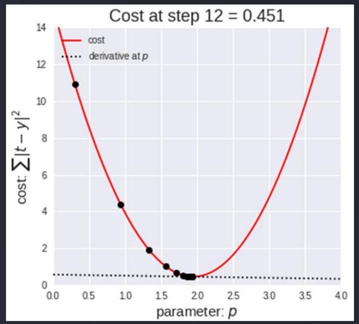
人工智能的底层是数学！

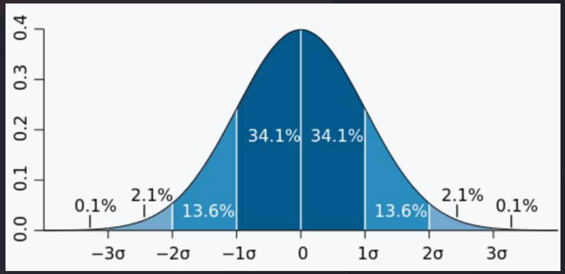
$$y = f(x)$$

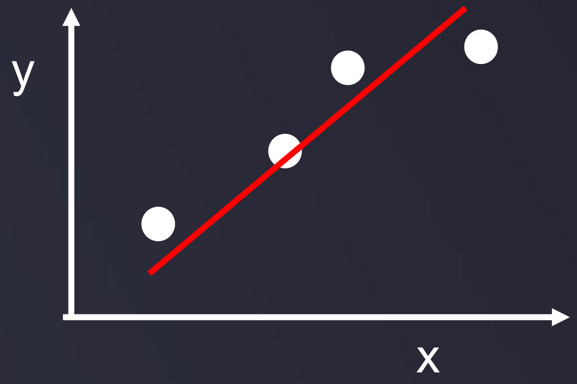
$$\begin{bmatrix} 3 & 5 & 2 \\ 1 & 1 & 7 \\ 1 & 1 & 9 \end{bmatrix} \quad \begin{bmatrix} 5 & 2 \\ 4 & 3 \\ 1 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 3 & 5 & 4 \\ 2 & 4 & 2 & 1 \end{bmatrix}$$





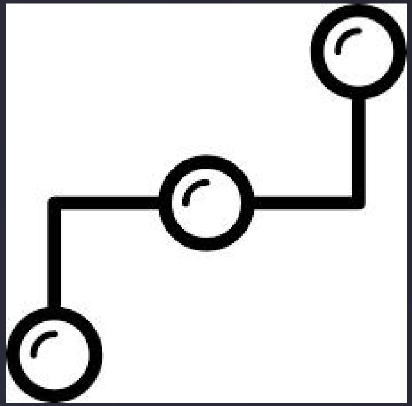




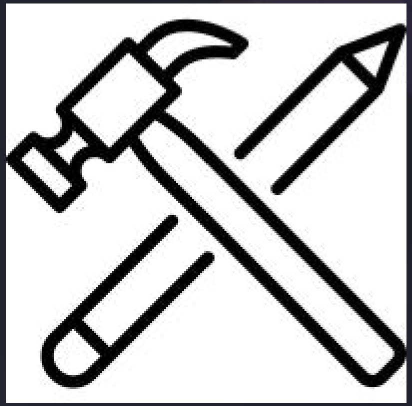
$$P(S_{t+1}|S_t, S_{t-1}, \dots, S_0) = P(S_{t+1}|S_t)$$

课程的预期目标：掌握人工智能所必备的数学基础知识

学习人工智能深度
学习底层运行最精
华的数学知识



分析数学知识与
人工智能领域的
关联和应用案例



结合PYTHON编
程实战，手把手
应用学到的知识

讲师介绍-Charlie老师

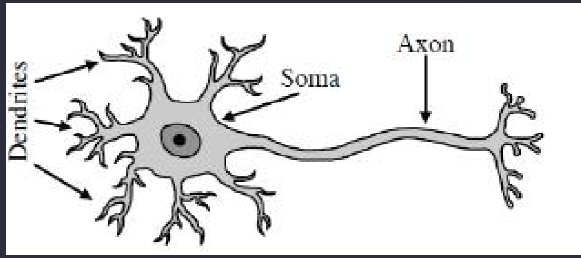
- 人工智能算法科学家-曾服务于某世界500强中国AI Lab，后自主创业
- 2019国家培训计划讲师-面向高职院校老师的培训课程
- 深圳市海外高层次人才认定(孔雀人才)
- 美国圣地亚哥国家超算中心博士后
- 加利福尼亚大学圣地亚哥全奖博士
- 参与美国自然科学基金(NSF)及加州能源局 (CEC)资助的392MW IVANPAH等智慧电网项目
- 21篇国际期刊文章(sci收录17篇)，总引用接近1000
- 第一作者发明专利11份

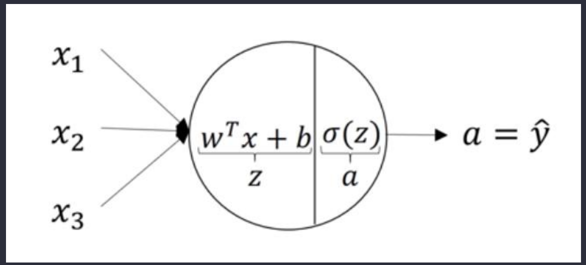
课程相关资料

学习资料

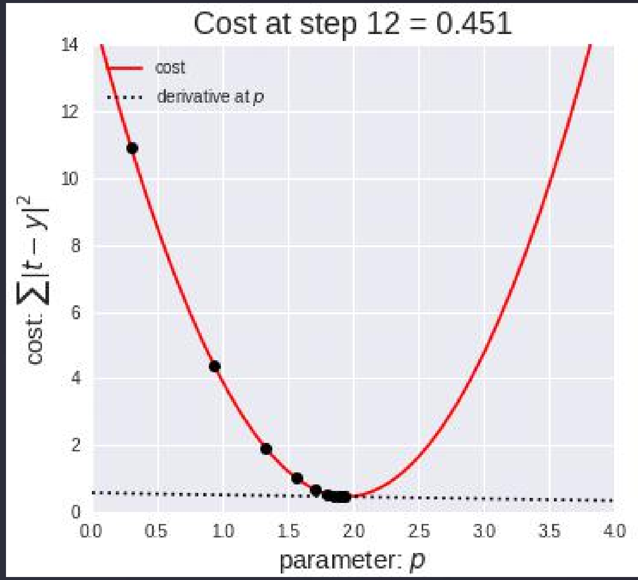
欢迎大家扫码或者添加微信好友ai_flare（学习小助手），加入学习群，老师会在群里和大家进行交流和答疑（名额有限、人满即止）

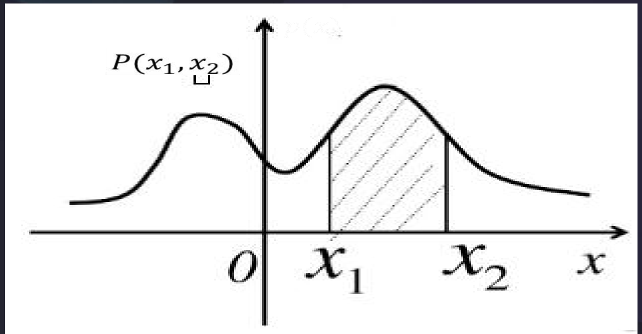
讲解基础数学与人工智能底层逻辑的关联





$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{12} \\ a_{21} & a_{22} & \cdots & a_{22} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$





概率预测

$$p \rightarrow \bar{p}\,, \sigma(p)$$

模型参数初始化

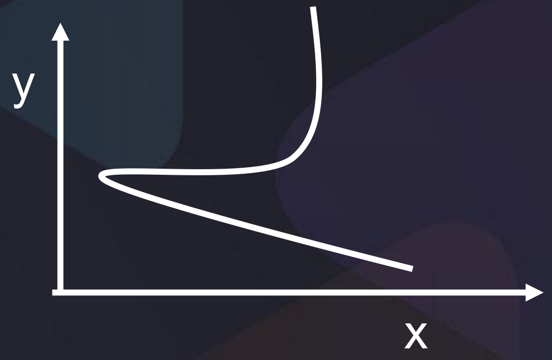
$$init.\,(w, b)$$

课程大纲

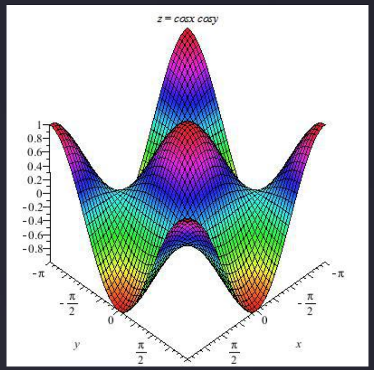
- 人工智能的本质-函数
- 神经网络的底层机理-矩阵运算
- 深度学习的学习原理-微积分
- AI模型的评估和优化-概率与统计
- 入门经典的机器学习-线性回归
- 补充小章节1：人工智能模型的应用与配套-逻辑运算
- 补充小章节2：强化学习的基础-马尔科夫链

课程大纲：函数

- 函数的定义
- 常见函数类型
- 复合函数
- 分段函数
- 函数与反函数
- 函数的单调性，奇偶性，周期性
- 多元函数与人工智能



$$f\left(\frac{1}{3x}\right) = \frac{1 - \frac{3}{x^2}}{3x + 2}$$



课程大纲：矩阵运算

- 矩阵的定义和应用
- 同型矩阵
- 矩阵的加减，数乘，乘法等运算
- 矩阵的转置
- 向量的定义和运算

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{12} \\ a_{21} & a_{22} & \cdots & a_{22} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

$$AB \neq BA$$

$$(AB)C = A(BC)$$

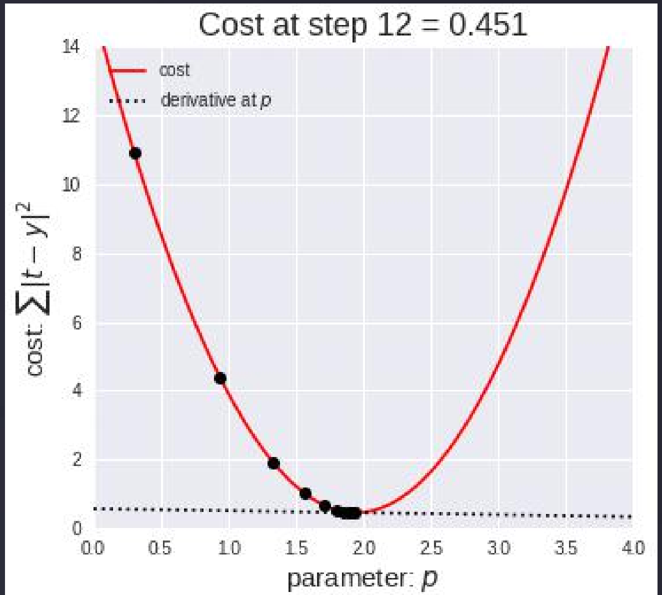
$$A\,(B + C) = AB + AC$$

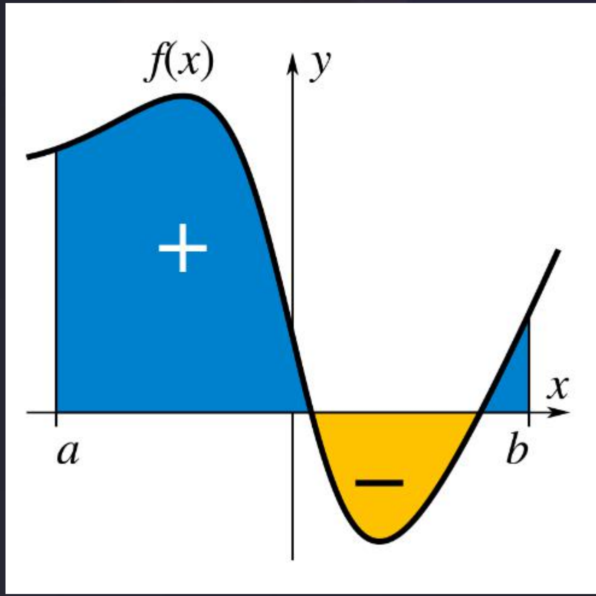
$$A = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \qquad A = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}$$

课程大纲：微积分

$$A = \lim_{x \to x_0} f(x)$$

- 极限
- 导数
- 极值点
- 梯度下降
- 积分

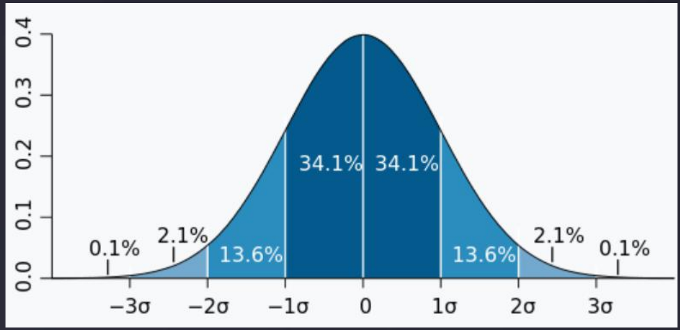




课程大纲：统计与概率

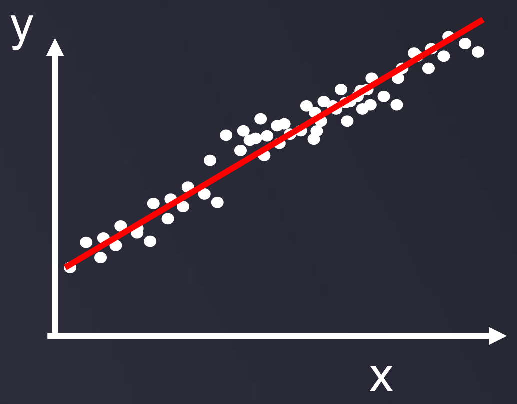
- 概率
- 期望
- 统计
- 正态分布

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \qquad \sigma = \sqrt{\sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n}}$$



课程大纲：线性回归

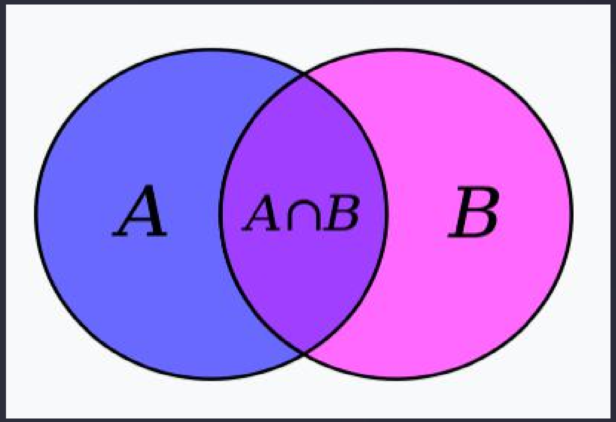
- 人工智能与机器学习的基本概念
- 回归Regression与线性回归
- 最小二乘法求解
- 梯度下降求解



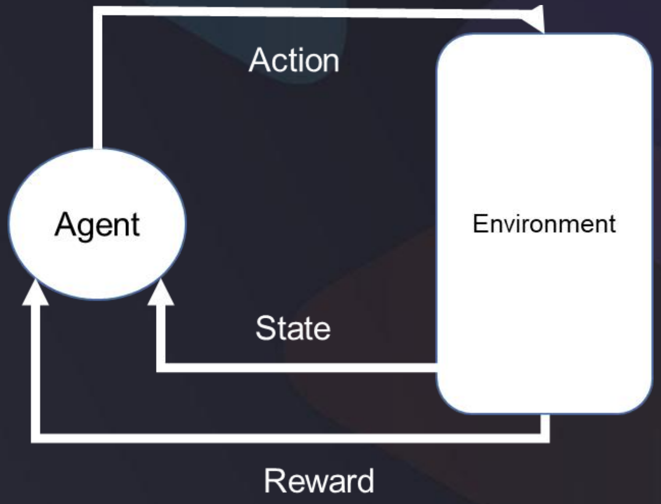
$$a = \frac{\sum_{i=1}^{n}\left(x^{(i)} - \bar{x}\right)\left(y^{(i)} - \bar{y}\right)}{\sum_{i=1}^{n}\left(x^{(i)} - \bar{x}\right)^2}$$

$$b = \bar{y} - a\bar{x}$$

课程大纲：逻辑运算与马尔科夫链小章节



- 逻辑运算与布尔值
- 逻辑运算符
- 比较运算符
- 逻辑运算的优先顺序
- 逻辑运算与人工智能的应用案例



- 马尔科夫链的定义
- 状态转移矩阵
- 马尔科夫奖励过程
- 马尔科夫决策过程与强化学习

结合应用的代码实战案例

1.某城市的居民用电的分段函数收费应用实战：通过python代码实战理解函数的应用

$$y = \begin{cases} ? \\ ? \\ ? \end{cases}$$

2.总店往分店发货的矩阵运算实战应用实战：理解矩阵运算的便利和优势

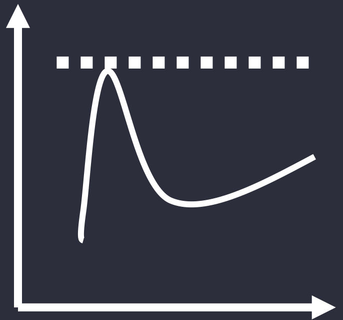
	商店	产品		
		苹果	猪肉	纸巾
数量	1	100	120	800
	2	200	150	1000
单价		5	50	1

$$\begin{bmatrix} 3 & 10 & 6 \end{bmatrix}$$

$$\begin{bmatrix} 5 & 2 \\ 4 & 3 \\ 1 & 1 \end{bmatrix}$$

结合应用的代码实战案例

3.流水线生产优化实战应用实战：考虑该生产线设备启动的固定成本以及用料用工用电用水等因素，通过微分极值求解最优运营策略



4.蒙特卡罗近似算法实战：通过概率方法来近似求解圆周率

5.用某小区抽样调查数据应用实战：搭建家庭食品消费为因变量的线性回归模型

样本	家庭收入	人数	食品消费
1	7900	3	3680.61
2	17200	3	4681.39
3	10000	2	3192.53
4	20900	3	4764.64
5	22000	3	5003.52
6	5500	4	4079.39
7	5300	2	2550.58
8	8300	2	2830.24

人工智能模型的本质：函数

函数的定义

（1）非空数集D

（2）某运算/对应法则f

（3）当$\forall$x $\in D$: 存在唯一对应的 $y = f(x)$

x:自变量　y:因变量

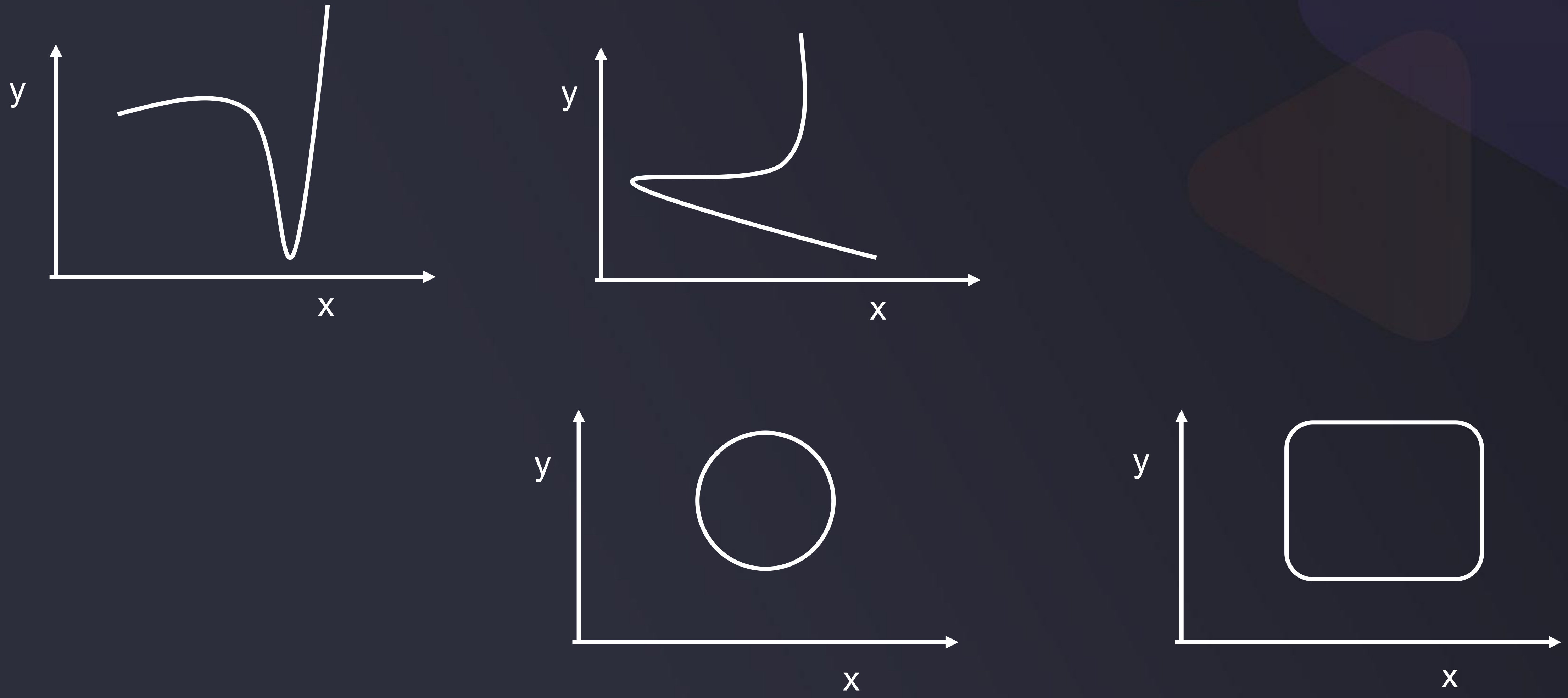
函数三要素：

$\{x \mid x \in D\}$ 函数的定义域

$\{y \mid y = f(x),\ x \in D\}$ 函数的值域

f 对应法则

符合定义的函数案例



定义域与值域

函数的定义域和值域指自变量x和因变量y的取值范围，为非空数集

$$y = \sqrt{x}$$

$$y = \frac{1}{\mathrm{x}}$$

相同函数的定义

定义域和对应法则相同，则两个函数相同

$$y = 10 \qquad y = 10 + x - x$$

$$y = \sqrt{x} * \sqrt{x} \qquad y = x$$

$$y = |\sqrt{x^2}| \qquad y = |x|$$

几种常见函数类型

幂函数 $y = x^k$

指数函数 $y = a^x$

对数函数 $y = \log_a x$

三角函数 $y = \sin(x)$ $y = \cos(x)$ $y = \tan(x)$

复合函数

$y = f(x)$ $u = f(\mathrm{t})$

$y = f(\mathrm{t})$ $\mathrm{t} = g(x)$ $y = f(g(x))$

$y = t + 10$ $\mathrm{t} = 2\mathrm{x}$ $y = 2x + 10$

$y = f(3\mathrm{x} + 5)$ $y = f(\mathrm{t})$ $\mathrm{t} = 3x + 5$

复合函数

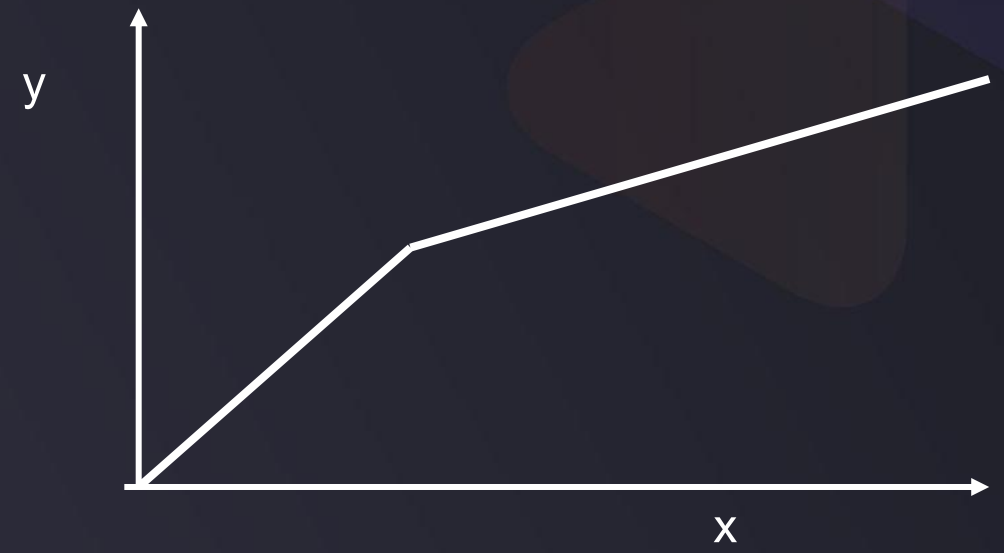
$$f\left(\frac{1}{3x}\right)=\frac{1-\frac{3}{x^2}}{3x+2}$$

求$f(x)$

分段函数

一个函数在不同的定义域的对应/表达式不同

$$y = \begin{cases} x & 0 \le x \le 10 \\ 0.5x & x \ge 10 \end{cases}$$



分段函数

$$y = \begin{cases} 3x^2 & x < 0 \\ 6x - 3 & x \geq 0 \end{cases}$$

1. 如果y=3, 求x

2. 求$f(f(-10))$

函数与反函数

$$y = f(x), x \in D, y \in Z$$

$$\Downarrow$$

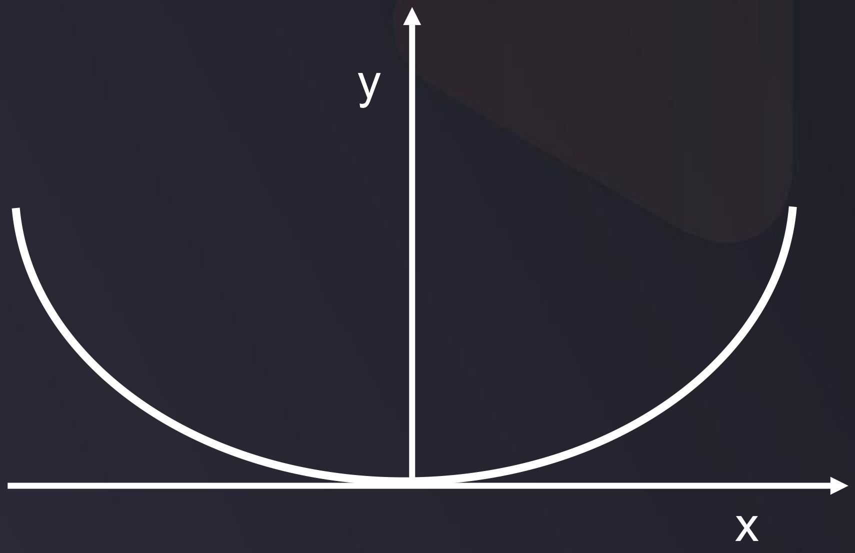
$$y = f^{-1}(x), x \in Z, y \in D$$

$$y = 5x + 10$$

函数与反函数

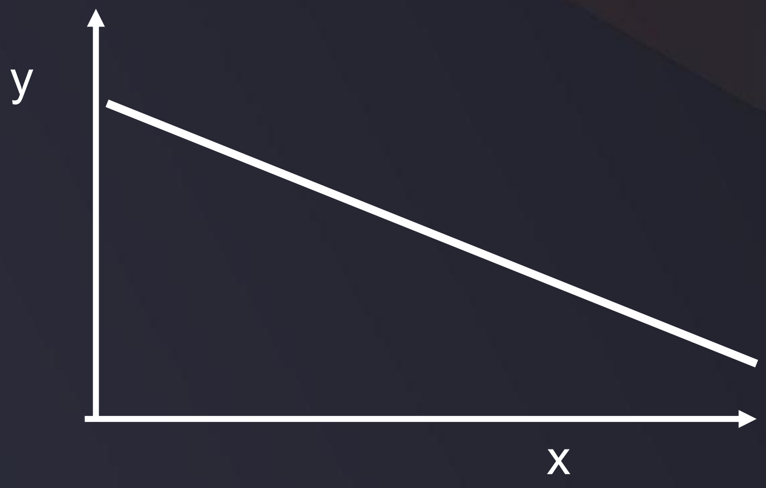
一个函数如果存在反函数，则其必为一个单调函数

$$y = x^2$$



函数的单调性

$y=f(x)$ 在定义域的区间I上面随着x变大而持续变大或变小，那么称$f(x)$ 在I区间上单调递增或递减



函数的单调性

$$y = 2x + 10$$

$$y = x^2$$

函数的单调性

递增函数+递增函数=递增函数

递减函数+递减函数=递减函数

递增函数-递减函数=递增函数

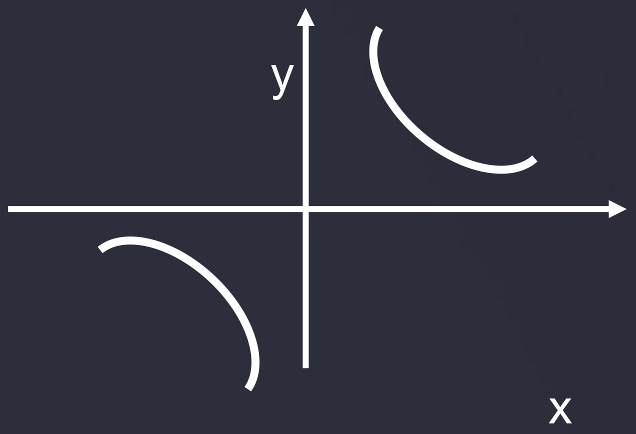
递减函数-递增函数=递减函数

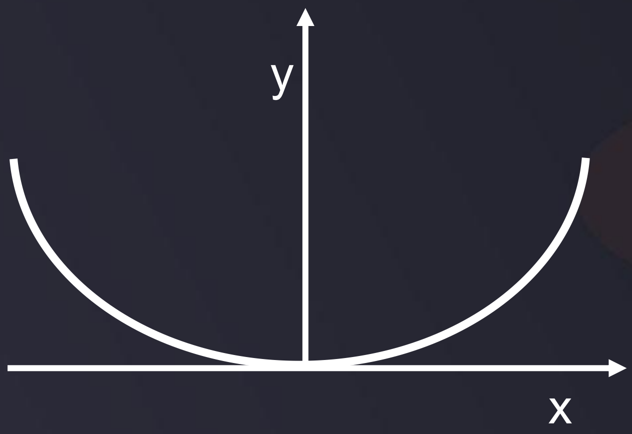
递增（减）函数 x 正数=递增（减）函数

递增（减）函数 x 负数=递减（增）函数

奇函数与偶函数

奇函数对于原点对称，偶函数对于y坐标轴对称





$$y = x$$
$$y = x^3$$
$$y = \sin(x)$$

$$y = |x|$$
$$y = x^2$$
$$y = \cos(x)$$

奇函数与偶函数

奇函数±奇函数=奇函数

偶函数±偶函数=偶函数

奇函数×奇函数=偶函数

偶函数×偶函数=偶函数

奇函数×偶函数=奇函数

函数的周期性

$f(x+\mathrm{T})=f(x)$ T为函数一个周期，最小的正数T为最小正周期

$$y = \sin(x)$$

$$y = \cos(x)$$

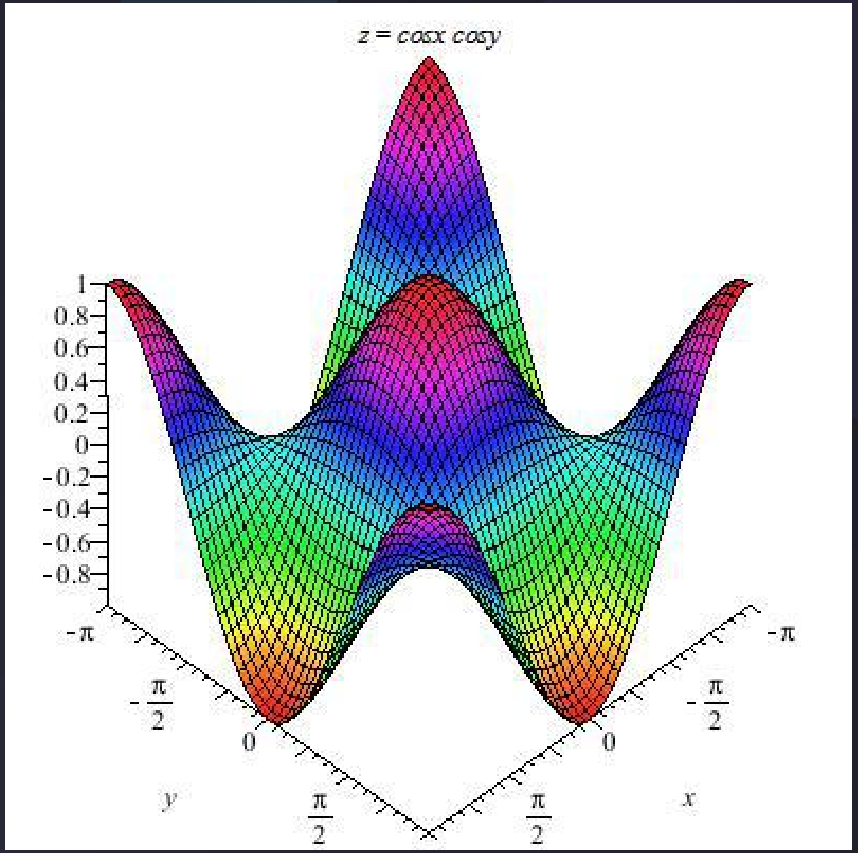
多元函数

（1）非空的n元数集D

（2）某运算/对应法则f

（3）当$(x_1, x_2, \dots x_n) \in D$: 存在唯一对应的 $y = f(x)$

$(x_1, x_2, \dots x_n)$:自变量　y:因变量

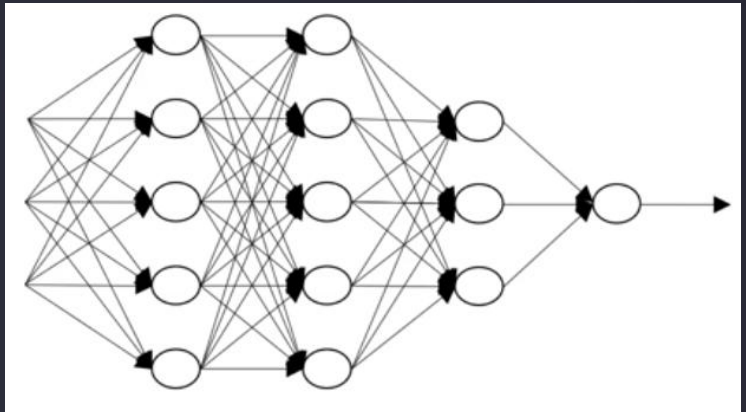


$$y = \cos(x_1)\cos(x_2)$$

人工智能模型可以理解成一个多元函数

输入X → 模型 $f(x)$ → 输出Y

多维度的结构化
金融大数据



股票价格

猫（99.8%）

Hello World!

实战习题

某城市的居民用电的收费方式如下

- 10度以内：固定10元
- 10-50度的部分：每度0.538元
- >50度的部分：每度0.752元

请用代码搭建函数，计算用电5.56度、30.21度、100.82度时候的成本

$$y = \begin{cases} ? \\ ? \\ ? \end{cases}$$

实战开发环境准备

- 下载、安装ANACONDA和python3
- 新建开发环境，安装jupyter-notebook
- 安装最常用的python模块numpy和matplotlib

本章回顾

- 函数的定义
- 常见函数类型
- 复合函数
- 分段函数
- 函数与反函数
- 函数的单调性，奇偶性，周期性
- 多元函数与人工智能

课程相关资料

学习资料

欢迎大家扫码或者添加微信好友ai_flare（学习小助手），加入学习群，老师会在群里和大家进行交流和答疑（名额有限、人满即止）

EDU

CSDN学院　IT实战派

神经网络的底层机理-矩阵运算

上一章回顾

- 函数的定义
- 常见函数类型
- 复合函数
- 分段函数
- 函数与反函数
- 函数的单调性，奇偶性，周期性
- 多元函数与人工智能

什么是矩阵？

Matrix

$$\begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

a_{ij}

$$\begin{array}{c} n \\ m\downarrow \begin{matrix} a_{11} & a_{12} & \cdots & a_{12} \\ a_{21} & a_{22} & \cdots & a_{22} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{matrix} \end{array}$$

k

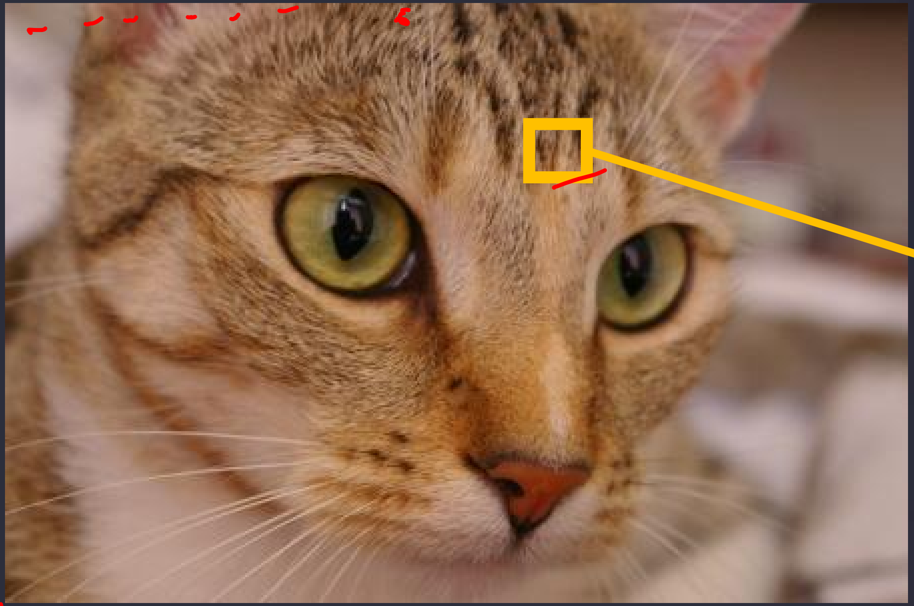
$$A_{m\times n} \quad A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{12} \\ a_{21} & a_{22} & \cdots & a_{22} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

m n

由 m * n个数 a_{ij}（$i=1,2,\ldots m; j=1,2,\ldots,n$）排成成的m行n列的数表

矩阵的应用

CNN



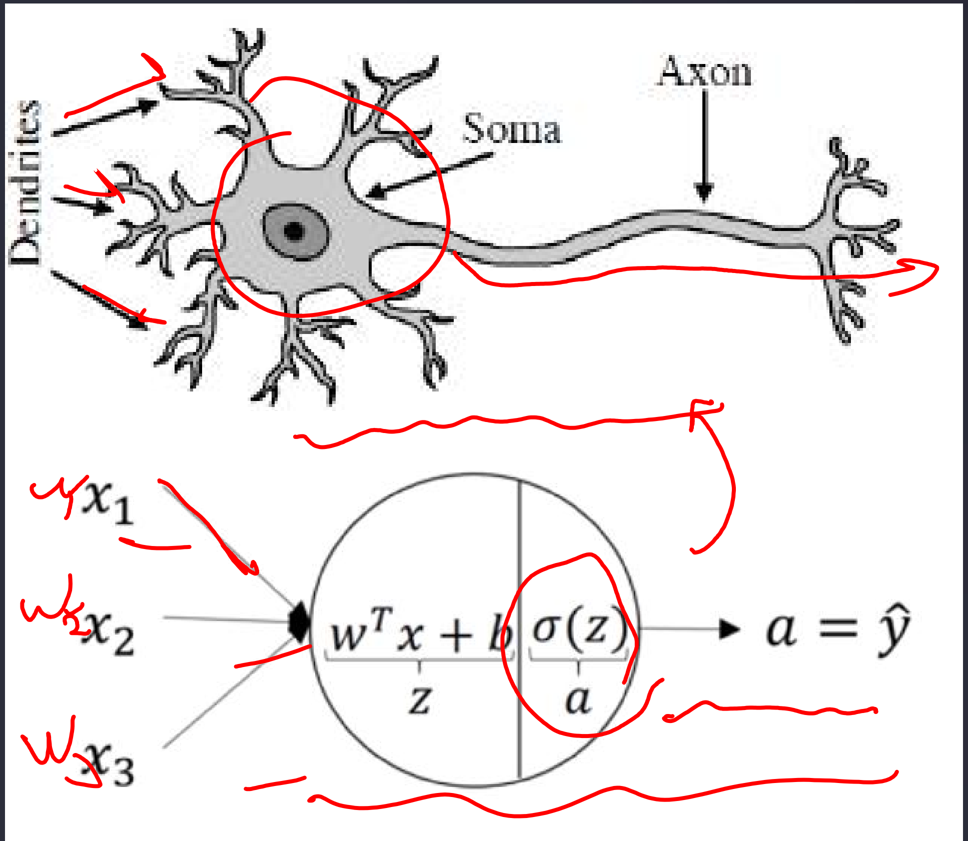
图片尺寸（300 X 400 X 3）

0	255	0	0	0
0	255	0	0	0
0	255	255	0	0
0	0	255	255	0
0	0	0	255	0

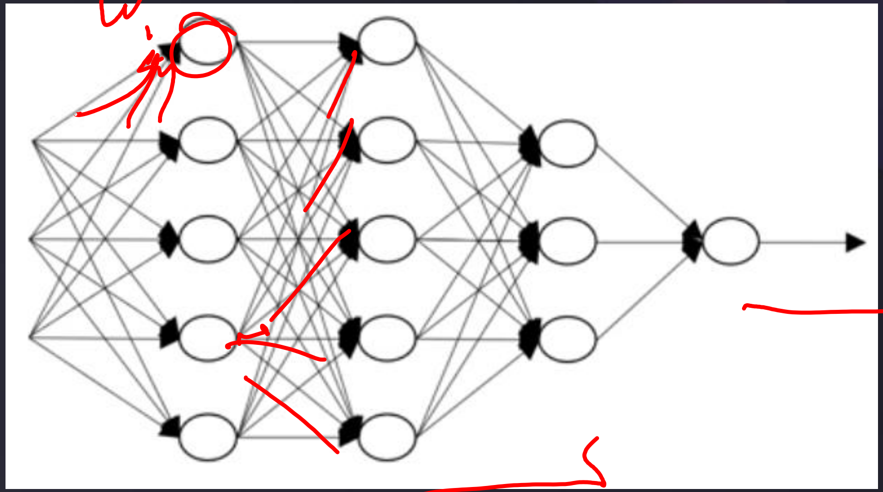
$A_{300,400}$

R G B

矩阵运算与人工智能



Dendrites
Soma
Axon
x_1
x_2
x_3
$\underbrace{w^T x + b}_{z}$
$\underbrace{\sigma(z)}_{a}$
$a = \hat{y}$



矩阵运算应用案例-房价预测

房屋面积、房间数、区域人口密度、房龄等因子预测 合理售价。

房屋号	面积	房间数	区域人口密度	房龄	价格
1	70	2	50	15	742,000
2	60	2	60	10	662,000
3	110	4	70	20	1,174,000
4	80	3	40	15	843,000
5	70	3	30	10	743,000
m	90	3	60	10	???

$$y_m = f(x_{m1}, x_{m2}, x_{m3}, x_{m4}, x_{m5})$$

$$y = f(x_1, x_2 \cdots x_n) = \theta_1 x_1 + \cdots + \theta_n x_n + b$$

$$\begin{cases} y_1 = f(x_{11}, x_{12}, x_{13}, x_{14}, x_{15}) \\ y_2 = f(x_{21}, x_{22}, x_{23}, x_{24}, x_{25}) \\ \vdots \\ y_5 = f(x_{51}, x_{52}, x_{53}, x_{54}, x_{55}) \end{cases}$$

$$y = \begin{bmatrix} x_{11} & x_{12} & x_{13} & x_{14} & x_{15} \\ x_{21} & x_{22} & x_{23} & x_{24} & x_{25} \\ & & \vdots & & \\ x_{51} & x_{52} & x_{53} & x_{54} & x_{55} \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \\ \theta_5 \end{bmatrix} + b$$

同型矩阵：行数、列数分别相同的矩阵

$$1 \rightarrow \begin{bmatrix} 1 & 3 \end{bmatrix}$$

2

$$2 \begin{bmatrix} 1 & 3 & 5 & 4 \\ 2 & 4 & 2 & 1 \end{bmatrix} \quad = \quad 2 \begin{bmatrix} 2 & 2 & 6 & 1 \\ 3 & 1 & 3 & 2 \end{bmatrix}$$

4 4

$$\begin{bmatrix} 3 & 5 & 2 \\ 1 & 1 & 7 \\ 1 & 1 & 9 \end{bmatrix}$$

3 3

$$\begin{bmatrix} 5 & 2 \\ 4 & 3 \\ 1 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \quad \begin{bmatrix} 2 & 7 \\ 9 & 3 \\ 1 & 5 \end{bmatrix}$$

3 2 ≠ 2, 3

3 2

矩阵的基本运算

矩阵的加法：矩阵元素分别相加

$$4\quad\begin{bmatrix}1 & 3 & 2\\7 & 1 & 5\\0 & 4 & 2\\3 & 1 & 2\end{bmatrix}_{3} + 4\begin{bmatrix}0 & 1 & 2\\1 & 3 & 5\\0 & 4 & 4\\5 & 0 & 2\end{bmatrix}_{3} = 4\begin{bmatrix}1 & 4 & 4\\8 & 4 & 10\\0 & 8 & 6\\8 & 1 & 4\end{bmatrix}_{3}$$

互为同型矩阵才能进行加法运算

矩阵的基本运算

矩阵的加法满足交换律、结合律，即：

$$A + B = B + A$$

$$A + B + C = A + (B + C)$$

矩阵的减法可以理解为对负矩阵的加法，即：

$$A + B = A + (-B)$$

$1+3=3+1$

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

$1+2+3 = 1+(2+3)$

$= 6$

$3-1=2$

$3+(-1)=2$

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} - \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} -1 & -1 \\ -1 & -1 \end{bmatrix}$$

矩阵的基本运算

$$\begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 0 & 0 & 1 \\ 4 & 5 & 6 \end{bmatrix} \cdot 2 = \begin{bmatrix} 1\times 2 & 2\times 2 & 3\times 2 \\ 3\times 2 & 2\times 2 & 1\times 2 \\ 0\times 2 & 0\times 2 & 1\times 2 \\ 4\times 2 & 5\times 2 & 6\times 2 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 6 \\ 6 & 4 & 2 \\ 0 & 0 & 2 \\ 8 & 10 & 12 \end{bmatrix}$$

矩阵的数乘：数与矩阵元素分别相乘

矩阵的基本运算

矩阵的数乘满足交换律、结合律、分配律，即：

$$\lambda A = A\lambda$$

$$\lambda A\mu = \lambda(A\mu)$$

$$\lambda(A + B) = \lambda A + \lambda B$$

$$4\times\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix} = \begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix}\times 4$$

$$2\times\left(\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix}+\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}\right) = 2\times\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix}+2\times\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}$$

$$\begin{bmatrix}2 & 3\\ 4 & 5\end{bmatrix} = \begin{bmatrix}4 & 6\\ 8 & 10\end{bmatrix} = \begin{bmatrix}2 & 4\\ 6 & 8\end{bmatrix}+\begin{bmatrix}2 & 2\\ 2 & 2\end{bmatrix}$$

矩阵的基本运算

矩阵与矩阵相乘：行列元素依次相乘并求和

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11}+a_{12}b_{21}+a_{13}b_{31}, & a_{11}b_{12}+a_{12}b_{22}+a_{13}b_{32} \\ a_{21}b_{11}+a_{22}b_{21}+a_{23}b_{31}, & a_{21}b_{12}+a_{22}b_{22}+a_{23}b_{32} \end{bmatrix}$$

$m \times n \quad n \times k$

$= m \times k$

$$\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$$

注意：矩阵相乘时候第一个矩阵的列数等于**第二个矩阵的行数**

$[2 \times 3] \times [3 \times 2] = [2, 2]$

矩阵的基本运算

矩阵与矩阵相乘**不满足交换律**，满足结合律、分配律，即：

$$AB \neq BA$$

$$(AB)C = A(BC)$$

$$A\,(B + C) = AB + AC$$

3, 2 × 2, 3 → 3×3

[3, 2] → 2×2

3×2, (2×4, 4×1)

3×4

→ 3×1

2×1 → 3×1

3×2 2×4 2×4

2×4 → 3×4

3×4 + 3×4 → 3×4

矩阵的基本运算

矩阵的转置运算A^{T}

A^{T}

$a_{ij} \to a_{ji}$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \Rightarrow \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{23} \end{bmatrix}$$

2

3

$m \times n \to n \times m$

$$\begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}^{T} \to \begin{bmatrix} 1 & 1 & 0 \\ 2 & 1 & 1 \\ 3 & 0 & 1 \end{bmatrix}$$

矩阵的基本运算

矩阵转置的相关运算律：

$$(A+B)^T = A^T + B^T$$

$$(A \times B)^T = B^T \times A^T$$

$$(A^T)^T = A$$

$$(kA)^T = k(A^T)$$

$$\underbrace{\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}}_{B^T} \times \underbrace{\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}}_{A^T} = \begin{bmatrix} 3 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}^T + \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix}^T = \begin{bmatrix} 2 & 1 \\ 2 & 0 \end{bmatrix}^T = \begin{bmatrix} 2 & 2 \\ 1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 3 & 0 \\ 0 & 0 \end{bmatrix}^T = \begin{bmatrix} 3 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}^T = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$$

$$a_{ij} \to a_{ji} \to a_{ij}$$

向量

$1 \times n$

$$A = [a_1 \quad a_2 \quad \cdots \quad a_n]$$

只有一行的矩阵

又称为行向量

$n \times 1$

$$A = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

只有一列的矩阵

又称为列向量

向量的基本运算

- ➢遵循矩阵基本运算原则
- ➢矩阵与向量相乘，结果仍为向量

$n\times 1 + 1\times n$

$n\times 1 \quad n\times 1$

$5\times 1 \quad 5\times 1$

$5\times 1 + 1\times 5$

$A\times B \neq B\times A$

$5\times 1 \cdot 1\times 5 = 5\times 5$

$1\times 5 \cdot 5\times 1 = 1\times 1$

$1\times n$

$m\times n \cdot n\times 1 = m\times 1$

$1\times m \times m\times n = 1\times n$

实战习题

A $\begin{bmatrix} 5 & 2 \\ 4 & 3 \\ 1 & 1 \end{bmatrix}$　　B $\begin{bmatrix} 2 & 7 \\ 9 & 3 \\ 1 & 5 \end{bmatrix}$

请计算　$A^{\mathrm{T}} \cdot \mathrm{B}$ 和 $\mathrm{B}^{\mathrm{T}} \cdot A$

实战习题：应用题

总店往2家分店（编号1、2）发送3种商品（苹果，猪肉和纸巾），用矩阵运算来计算向每家分店收取的费用。

	商店	产品		
		苹果	猪肉	纸巾
数量	1	100	120	800
	2	200	150	1000
	3			
单价		5	50	1

4

m 100 $\underbrace{2 \times 3 \quad 3 \times 1}$ $= 2 \times 1$

$\rightarrow m \times 1$

EDU

本章回顾

- 矩阵的定义和应用
- 同型矩阵
- 矩阵的加减
- 矩阵的数乘
- 矩阵的乘法
- 矩阵的转置
- 向量

$m \times n \quad n \times k \rightarrow m \times k$

$a_{ij} \rightarrow a_{ji} \quad m \times n \rightarrow n \times m$

课程相关资料

学习资料

欢迎大家扫码或者添加微信好友ai_flare（学习小助手），加入学习群，老师会在群里和大家进行交流和答疑（名额有限、人满即止）

CSDN学院 IT实战派

机器学习的学习原理-微积分

上一章回顾

- 矩阵的定义和应用
- 同型矩阵
- 矩阵的加减
- 矩阵的数乘
- 矩阵的乘法
- 矩阵的转置
- 向量

极限

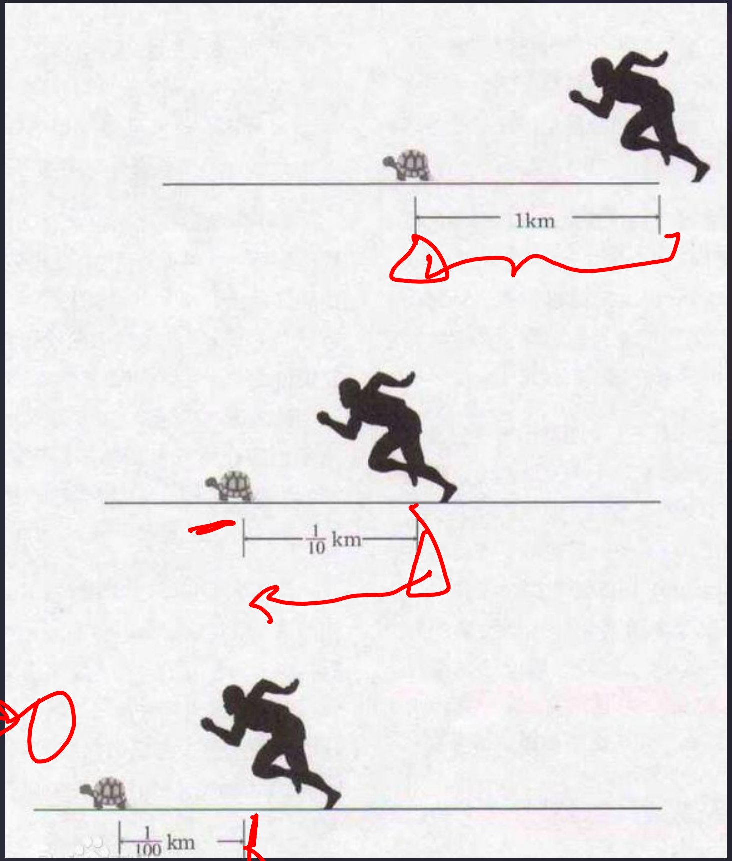
一根长为 1 的木棍，每次从中间切断，切n次后，剩余部分长度为多少?

次数	长度
0	1
1	$\frac{1}{2}$
2	
….	….
1000	
$n \to \infty$	???

$\frac{1}{2^n} \to 0$

阿基里斯悖论

10x



$\frac{1}{10}$

极限的定义

某一个函数中的某一个变量，此变量在变大（或者变小）的永远变化的过程中，逐渐向某一个确定的数值A不断地逼近的过程中，此变量的变化，被人为规定为“永远靠近而不停止”、其有一个“不断地极为靠近A点的趋势”称作极限

$$A = \lim_{x \to x_0} f(x)$$

当变量x无限趋近于某一个数值时，对应的函数结果

$$\lim_{x \to 2} \frac{1}{x} = \frac{1}{2}$$

$$\lim_{x \to \infty} \frac{1}{x} \to 0$$

求极限案例

$$y = \lim_{x \to \infty} \frac{(2x^3 + 6x^2 + 1)/x^3}{(5x^3 + 7x)/x^3} = \frac{2 + \frac{6}{x} + \frac{1}{x^3}}{5 + \frac{7}{x^2}} = \frac{2}{5}$$

$$\frac{6}{x} \to 0, \quad \frac{1}{x^3} \to 0, \quad \frac{7}{x^2} \to 0$$

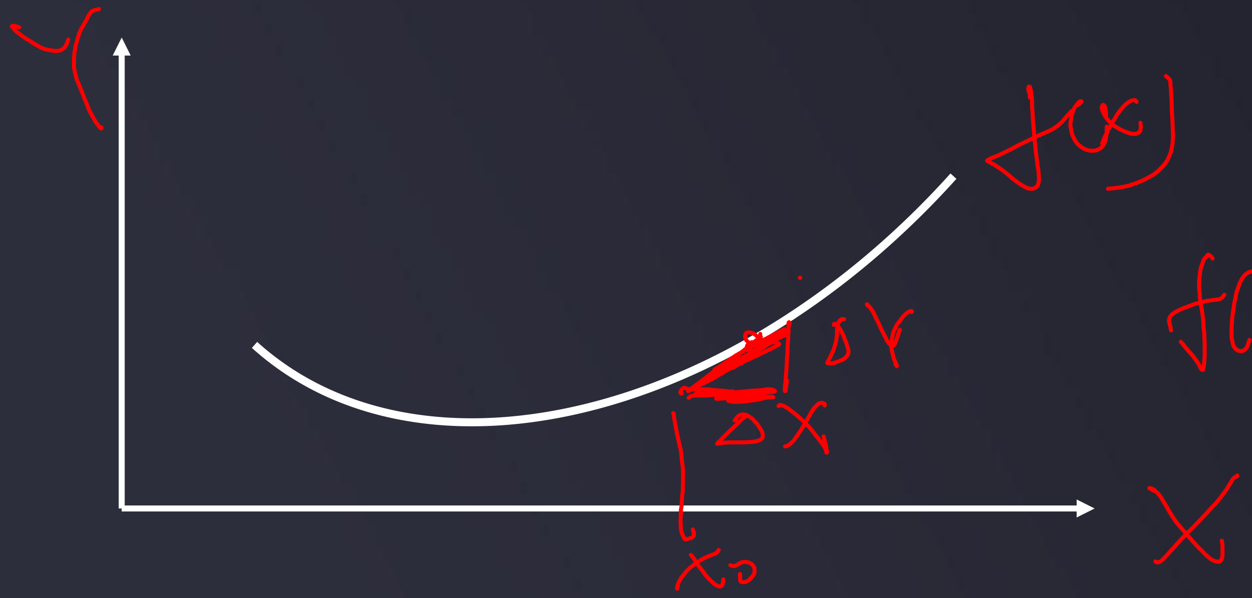
$$y = \lim_{x \to 0} \frac{4x - 1}{2x^3 - x - 1} = \frac{-1}{0 - 0 - 1}$$

$$= 1$$

导数

$f(x)$

导数的定义：一个函数在某一点的导数描述了这个函数在这一点附近的变化率。



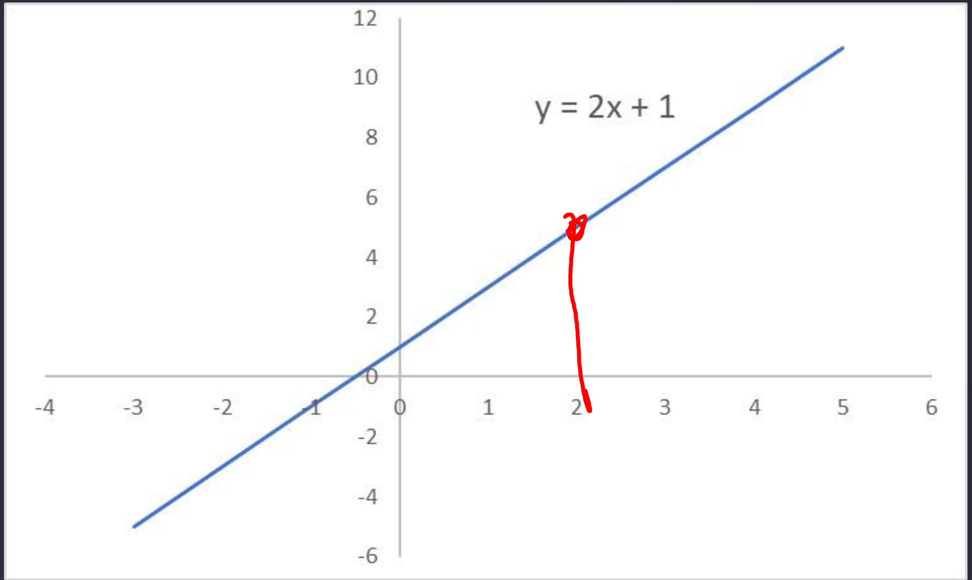
$$f'(x_0)=\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x}$$

对于任意函数$y=f(x)$，自变量x在一点x_0上产生一个增量Δx时，在Δx趋于0时的极限，函数输出值的增量Δy与自变量增量Δx的比值即为函数在x_0处的导数，记作$f'(x_0)$或$\frac{d\,f(x_0)}{d\,x_0}$。

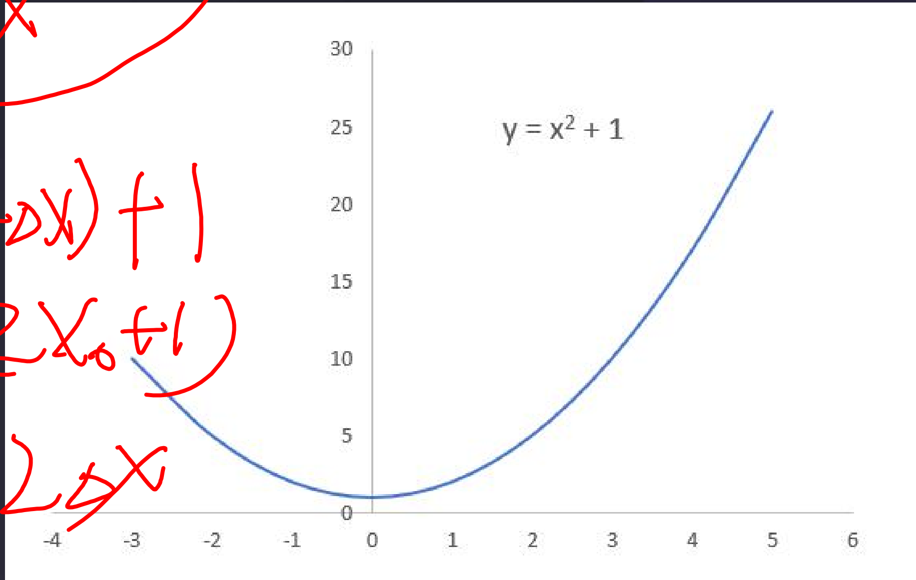
求导数案例

②

$$\frac{\Delta y = f(x+\Delta x) - f(x)}{\Delta x}$$



$2(x_0+\Delta x)+1$
$-(2x_0+1)$
$= 2\Delta x$



$f'(x=2)$

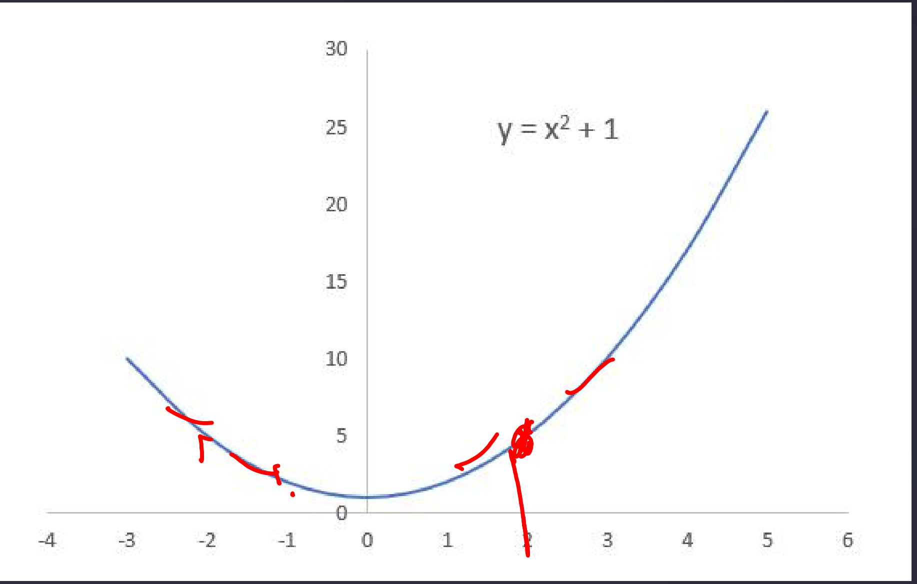
$= \frac{2\Delta x}{\Delta x}$

$= 2$

计算y = 2x+1 在 x=2的导数

计算y = x²+1 在 x=2的导数

求导数案例



计算$y = x^2+1$ 在 $x=2$的导数

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}$$

$$\frac{(x+\Delta x)^2+1-(x^2+1)}{\Delta x}$$

$$\frac{\cancel{x^2}+2x\Delta x+\Delta x^2}{\Delta x}$$

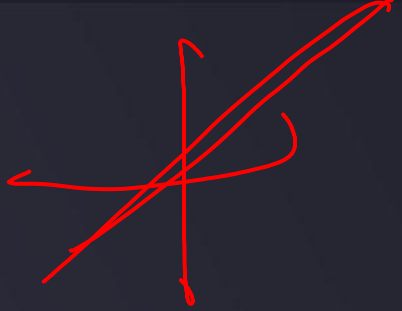
$$= 2x + \Delta x, \quad \Delta x \to 0$$

2

$2x$

$=4$

常用导数公式



$$(C)' = 0 \qquad (Cx)' = C$$

$$(x^u)' = ux^{u-1} \qquad \left(\frac{1}{x}\right)' = -\frac{1}{x^2}$$

$$(e^x)' = e^x \qquad (\ln x)' = \frac{1}{x}$$

$$(\sin x)' = \cos x \qquad (\cos x)' = -\sin x$$

常用导数公式

$$(C)' = 0$$

$$(Cx)' = C$$

$$(x^u)' = ux^{u-1}$$

$$\left(\frac{1}{x}\right)' = -\frac{1}{x^2}$$

$$(e^x)' = e^x$$

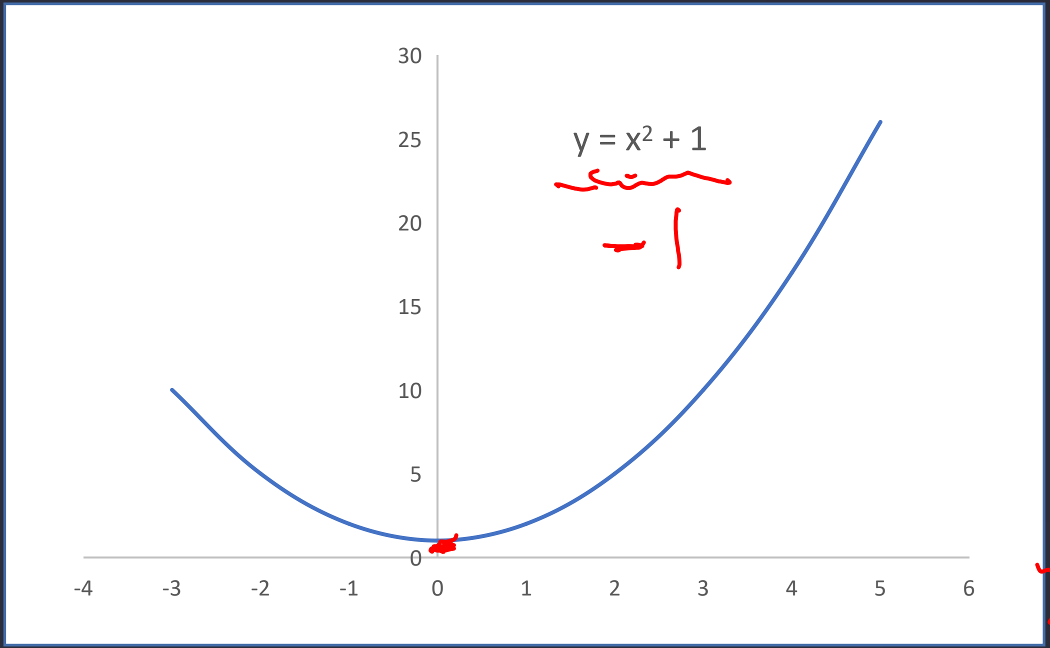
$$(\ln x)' = \frac{1}{x}$$

$$(\sin x)' = \cos x$$

$$(\cos x)' = -\sin x$$

导数与极值点

如果$y = f(x)$在 x = x_0 的导数 $f'(x_0) = 0$，则 x_0 为该函数的一个极值点。



$(x^2+1)'$ $2x$ 0

$x=0$ $2x=0$

y = x^2+1 在 x = 0 的导数为0，该点为极小值点，y = 1。

$f(0)=1$

$y = \frac{x^3}{3} - x$

x^2-1 1 -1

$-\infty$ ∞

导数在机器学习中的应用：梯度下降

机器学习常常通过梯度下降来对模型函数进行优化。通过向函数上当前点对应梯度（导数）的反方向的规定步长距离点进行迭代搜索，直到在极小点收敛。

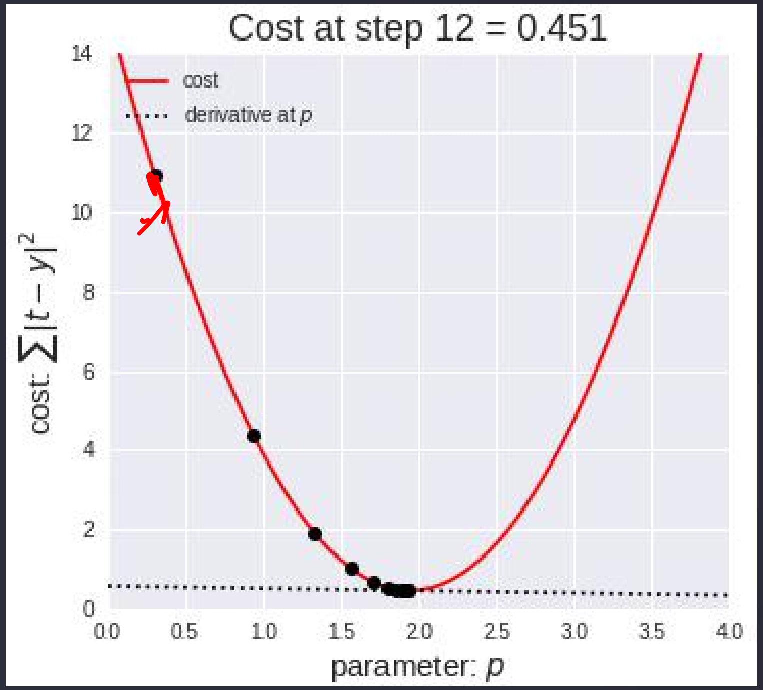
$$J = f(p) \quad \Longrightarrow \quad p_{i+1} = p_i - \alpha \frac{\partial}{\partial p_i} f(p_i)$$

搜索方法

input ⇒ θ_1 ... ⇒ output

$y = \theta_1 x_1 + \theta_2 x_2 + \cdots \theta_n x_n + b$

梯度下降案例



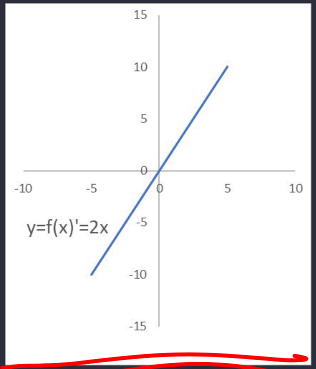
$$f(p) = 3.5p^2 - 14p + 14$$

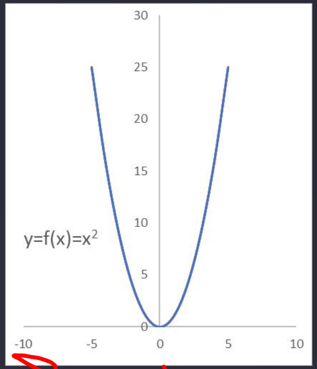
7P − 14

P = 2

积分

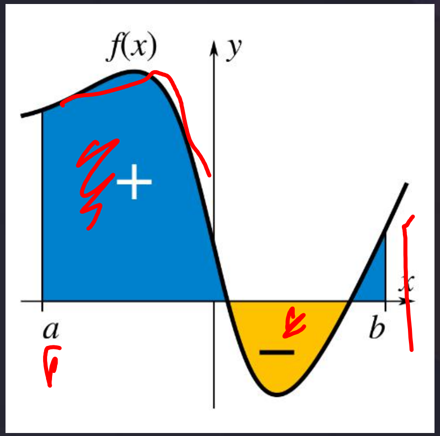
积分通常分为**不定积分和定积分**两种





$y(x)=x^2+C$

2x



不定积分：函数f 的不定积分，是一个可导函数F且其导数等于原来的函数f ，即$F' = f$

定积分：对于一个给定的正实值函数$f(x)$，由曲线、直线(x=a,x=b)以及轴围成的面积值。。

积分

函数的不定积分可以理解为其对应的反导数，有无穷多个

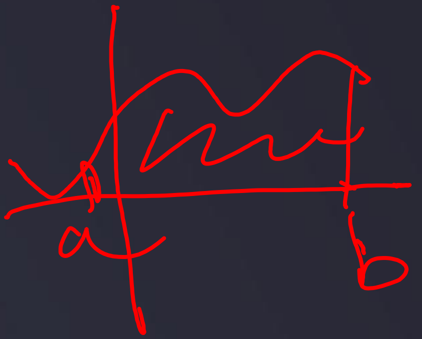
$$f(x) = 2x$$

$$F(x) = x^2 \quad {}' \rightarrow 2x$$

$$F(x) = x^2 + 1 \rightarrow 2x + 0$$

$$F(x) = x^2 + 10$$

$$\int_a^b f(x)dx = F(b) - F(a)$$



$$b^2 + c$$

$$-(a^2 + c)$$

常用积分公式

$C' = 0$

$$\int 0 \, \mathrm{d}x = C$$

$$\int a \, \mathrm{d}x = ax + C$$

$$\int x^a \, \mathrm{d}x = \frac{1}{a+1} x^{a+1} + C$$

$$\int \frac{1}{x} \, \mathrm{d}x = ln|x| + C$$

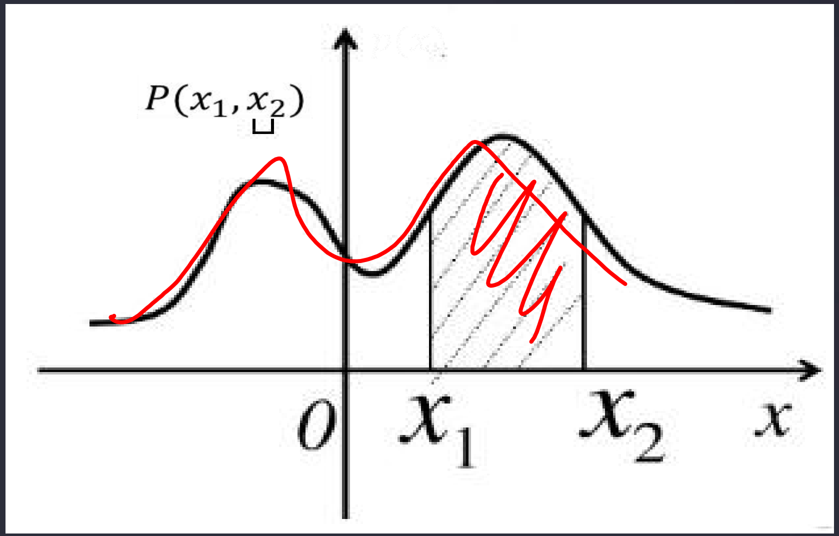
$$\int \sin x \, \mathrm{d}x = -\cos x + C$$

$$\int \cos x \, \mathrm{d}x = \sin x + C$$

$$\int e^x \, \mathrm{d}x = e^x + C$$

积分的应用案例

通过积分求概率



区间(x_1,x_2)的概率为：

1 10

$$P(x_1,x_2)=\int_{x_1}^{x_2}p(x)dx$$

实战习题

计算函数 $y = \frac{1}{3}x^3 - x$ 的极值

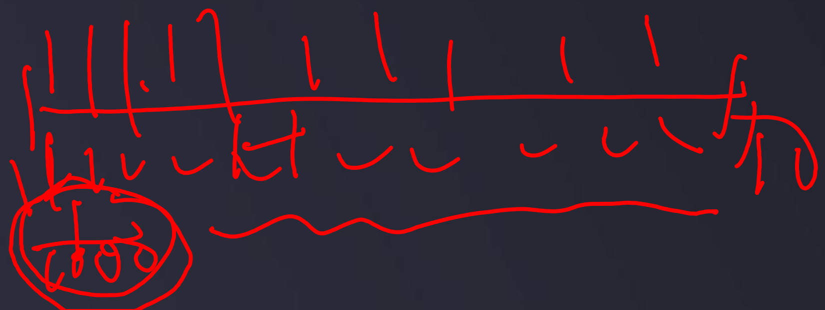
$+(-1x)$

$x^2 - 1$

-1 1

计算函数$y = \frac{1}{3}x^3 - x$ 从x = 1 到 x = 10与x轴形成的区域的面积。

Sympy



实战习题：应用题

某玩具厂的一条生产线生产某种玩具A，考虑该生产线设备启动的固定成本以及用料用工用电用水等因素，生产x个产品的总成本函数为

$C(x) = 2000 + 0.2x^2 + 5x$

(1)当x=10时的总成本、平均成本和边际成本？

(2)当生量x为多少时,平均成本最小？

$$\frac{2070}{10} = 207$$

$$2000 + 0.2\times 10^2 + 5\times 10 = 2070$$

20　　50

$$2000 + 0.2\times 121 + 55 = 2079.2$$

24.2

9.2

实战习题：应用题

某玩具厂的一条生产线生产某种玩具A，考虑该生产线设备启动的固定成本以及用料用工用电用水等因素，生产x个产品的总成本函数为

$C(x) = 2000 + 0.2x^2 + 5x$

(1)当x=10时的总成本、平均成本和边际成本？

(2)当生产量x为多少时,平均成本最小？

$\frac{C(x)}{x} = \left(\frac{2000}{x} + 0.2x + 5\right)'$

$= -\frac{2000}{x^2} + 0.2 \Rightarrow \frac{2000}{x^2} = 0.2$

$x^2 = 10000$

$x = 100, \cancel{-100}$

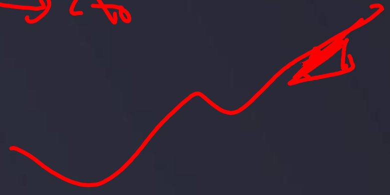
$20 + 20 + 5$

$= 45$

本章回顾

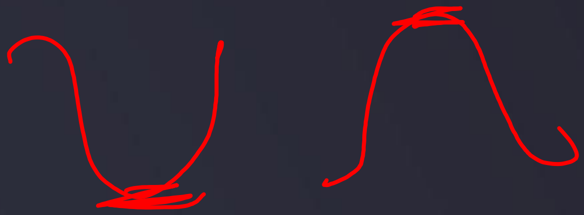
- 极限
- 导数
- 极值点
- 梯度下降
- 积分

$\lim_{x \to x_0} f(x)$



$x^n \to nx^{n-1}$

$c \to 0$



$F(x)' = f(x)$

$f(x) \to F(x)$

$F(b) - F(a)$

课程相关资料

学习资料

欢迎大家扫码或者添加微信好友ai_flare（学习小助手），加入学习群，老师会在群里和大家进行交流和答疑（名额有限、人满即止）

CSDN学院 IT实战派

评估和优化AI模型的表现-概率与统计

上一章回顾

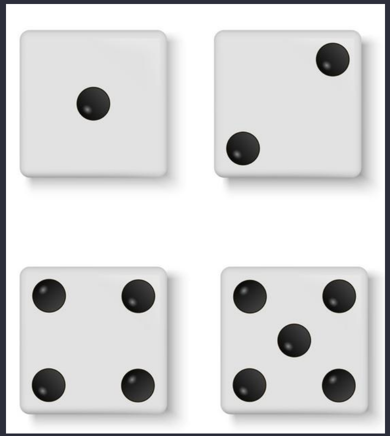
- 极限
- 导数
- 极值点
- 梯度下降
- 积分

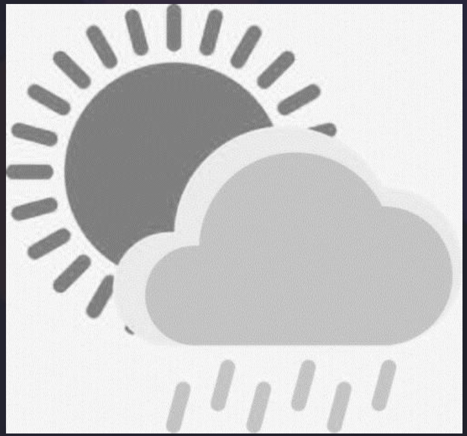
概率的定义

抛100次硬币，问正面朝上的情况有多少次？

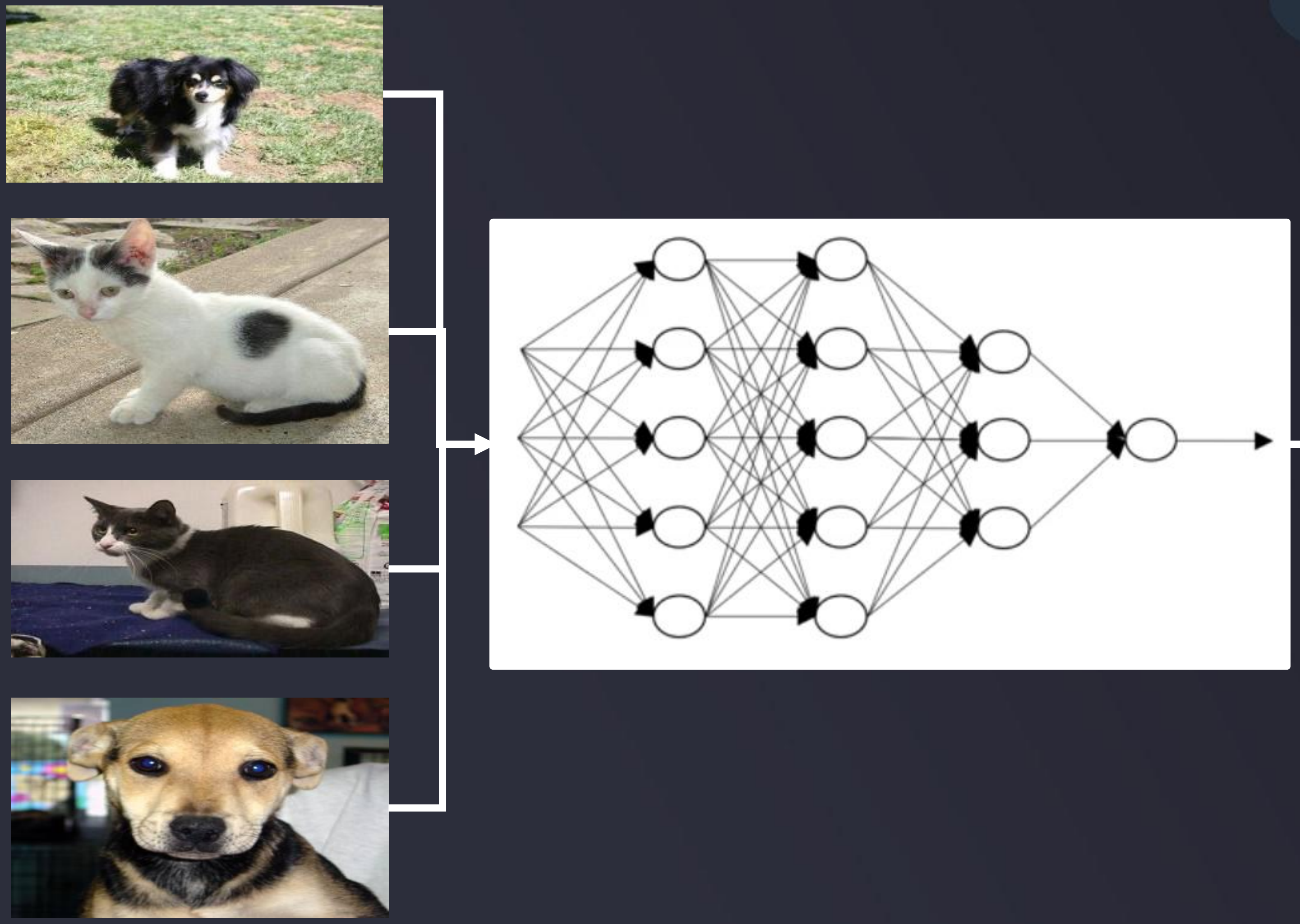
概率的定义：是一个在0到1之间的实数，是对随机事件发生之可能性的度量

这个世界是概率的！





深度学习中的概率



Dog: 0.99
Cat: 0.01

Dog: 0.09
Cat: 0.91

Dog: 0.01
Cat: 0.09

Dog: 0.98
Cat: 0.02

金融量化选股中的概率

	名称	操作时间	操作类型	操作价格	数量
1	中兴通讯	2018-06-28 15:0	买入	12.92	77300
2	中兴通讯	2018-11-23 15:0	卖出	19.77	77300
3	中国神华	2018-12-21 15:0	买入	18.91	20100
4	中国平安	2018-12-21 15:0	买入	58.34	6500
5	建设银行	2018-12-21 15:0	买入	6.30	60600
6	大族激光	2018-12-21 15:0	买入	30.09	12600
7	中国神华	2019-03-26 15:0	卖出	19.42	20100
8	中国平安	2019-03-26 15:0	卖出	72.69	6500
9	建设银行	2019-03-26 15:0	卖出	6.81	60600
10	大族激光	2019-03-26 15:0	卖出	39.64	12600
11	隆基股份	2019-05-27 15:0	买入	23.90	12400
12	中兴通讯	2019-05-27 15:0	买入	29.29	10100
13	五 粮 液	2019-05-27 15:0	买入	102.00	2900
14	新 希 望	2019-05-27 15:0	买入	17.61	16800
15	贵州百灵	2019-05-27 15:0	买入	11.62	25500
16	同花顺	2019-05-27 15:0	买入	[illegible]	[illegible]

预测为1 (涨幅>4%)概率：0.9
预测为0 (涨幅<=4%)概率：0.1

预测为1 (涨幅>4%)概率：0.2
预测为0 (涨幅<=4%)概率：0.8

期望值

期望值是指在一个离散性随机变量试验中每次可能结果的概率乘以其结果的总和。期望值也可以理解为随机试验在同样的机会下重复多次的结果计算出的的平均值。

$0.5 \times 3 + 0.5 \times -1$

$= 1$

一个抛硬币游戏，正面朝上收益为3，反面朝上收益为-1，问抛100次硬币的预期收益为多少？1000次，10000次呢？

100

10000

统计

某商场到访人员消费金额记录

样本	年龄	消费金额
1	25	1200
2	21	500
3	28	610
4	35	920
5	39	740
6	51	810
7	18	1048
8	22	360
9	41	130

平均值 $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$

中间值

标准差 $\sigma = \sqrt{\sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n}}$

15 15 16 15 16

10 60 30 45 21

统计

某商场到访人员消费金额记录

样本	年龄	消费金额
1	25	1200
2	21	500
3	28	610
4	35	920
5	39	740
6	51	810
7	18	1048
8	22	360
9	41	130

$$\frac{(25+21+28+35+\cdots+41)}{9}$$

平均值 $$\bar{x}=\frac{\sum_{i=1}^{n} x_i}{n}$$

中间值

标准差 $$\sigma=\sqrt{\sum_{i=1}^{n}\frac{(x_i-\bar{x})^2}{n}}$$

平均值和标准差在机器学习中的应用-评估模型表现

预测结果 \ 真实结果	狗	猫
狗	98	4
猫	2	96

$\frac{98}{100} = 98\%$

$\frac{96}{96+4} = 96\%$

$\frac{98\times1 + 96\times1 + 2\times0 + 4\times0}{98+96+2+4} = 97\%$

$\frac{98\times1 + 2\times0}{98+2} = 98\%$

1 0

平均值和标准差在机器学习中的应用-优化输入数据

区域人口密度
1000-10000

房屋面积
20-100

模型

房价

输入数据归一化

$$\frac{x - \bar{x}}{\sigma}$$

原始输入

区域人口密度
1000-10000

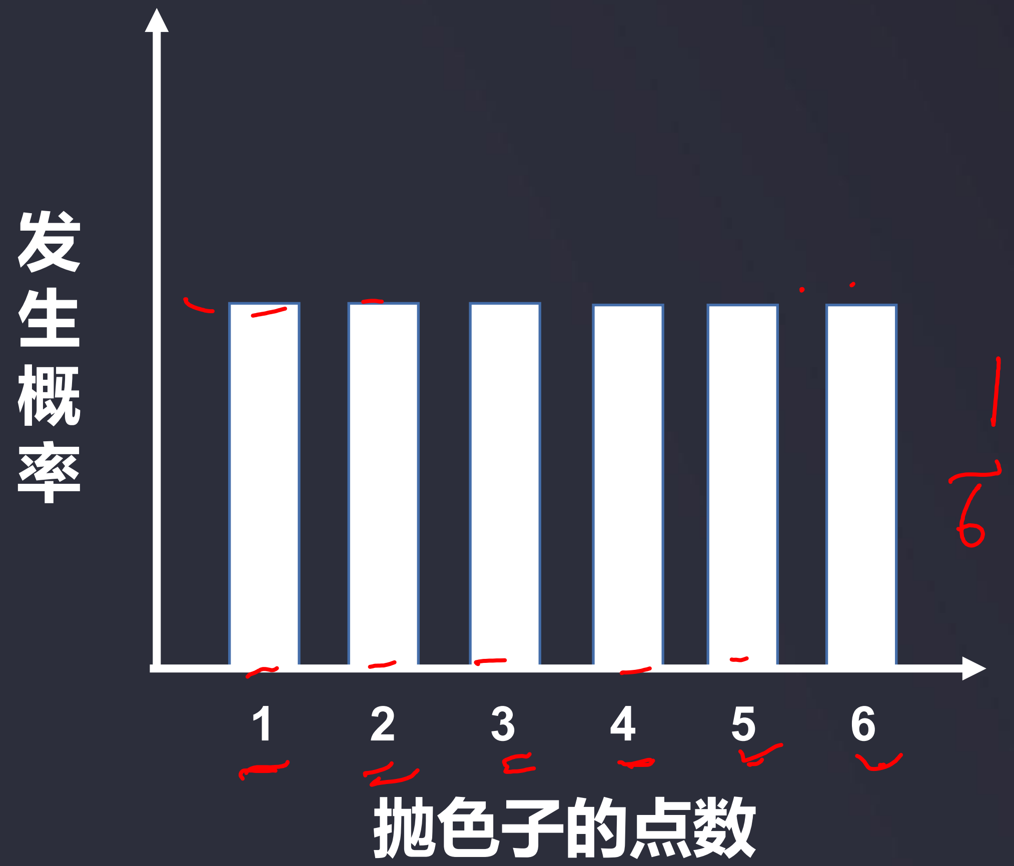
房屋面积
20-100

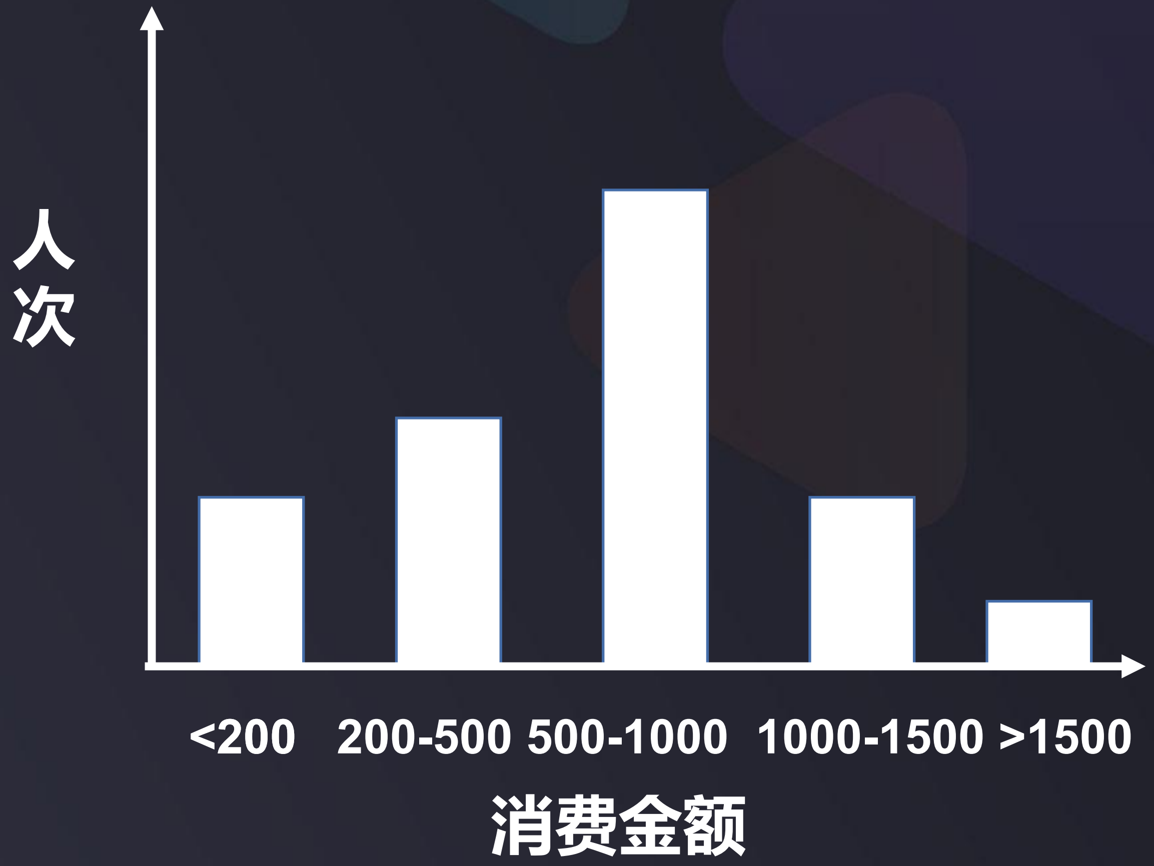
归一化输入

区域人口密度因子
-10 - 10

房屋面积因子
-5 - 5

统计与分布



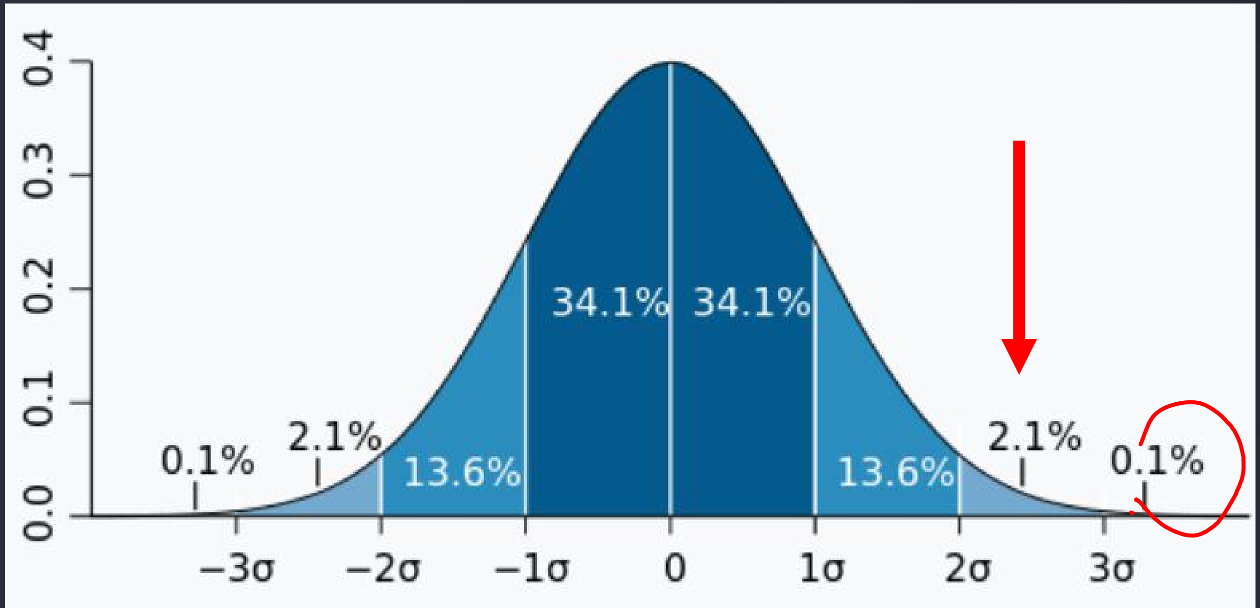


$P_1+P_2+P_3+P_4+P_5+P_6=1$

$\sum P_i = 1$

EDU

正态分布（高斯分布）



$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(x-\overline{x})^2}{2\sigma^2}\right)$$

68.2%

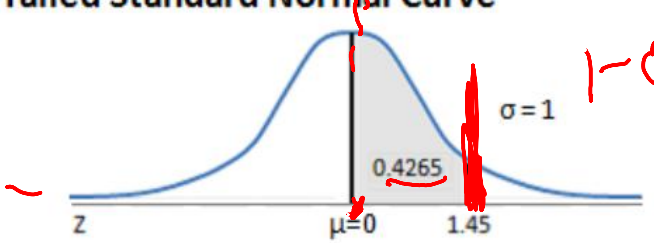
正态分布表的使用

0.5　0.5　0.9265

$Z = \frac{X - \bar{X}}{\sigma}$

Areas Under the One-Tailed Standard Normal Curve

This table provides the area between the mean and some Z score.
For example, when Z score = 1.45 the area = 0.4265.



1−0.9265
0.0735

0.9332 − 0.6915 = 0.2417

$\bar{X} = 35$

$\sigma = 10$

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545

$Z = \frac{50 - 35}{10} = 1.5$

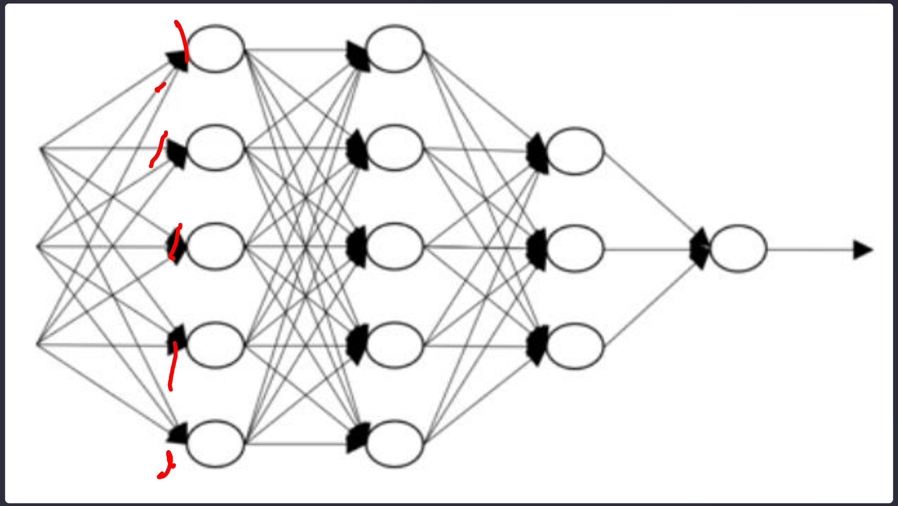
$Z = \frac{40 - 35}{10} = 0.5$

0.5+0.1915 = 0.6915　0.9332

正态分布在机器学习人工智能的应用

15 20 25

15% 30%



0 0 0 0 0 0

0.1 0.15 0.22 0.37

概率预测

$$p \rightarrow \bar{p}, \sigma(p)$$

模型参数初始化

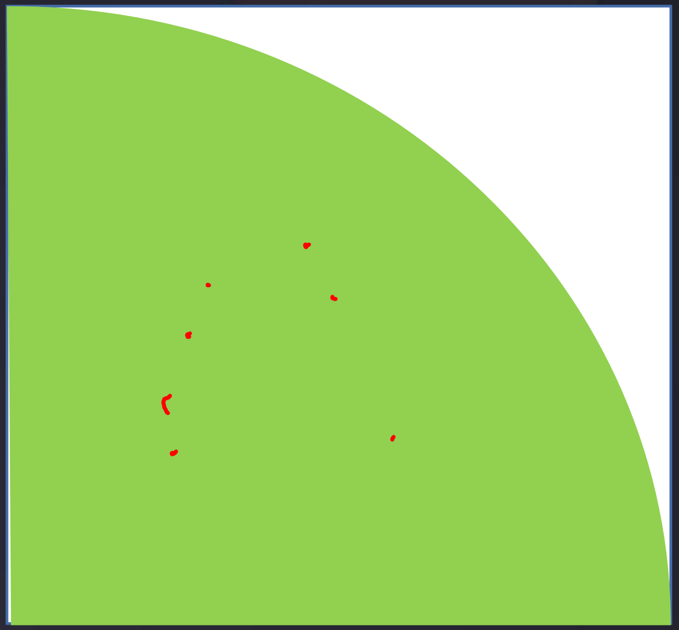
$$init.(w, b)$$

实战习题

1. 生成平均值为0，标准差为1的10000个正态分布的随机数，统计数值>0,>0.5,>1,>2的随机数个数

2. 蒙特卡罗近似算法：在边长为1正方形区间内随机落点，问落点在绿色四分之一圆区间内的概率是多少？

$\frac{\pi r^2}{4}$



EDU

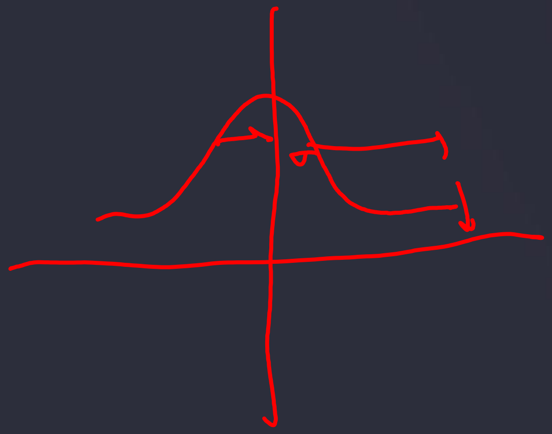
本章回顾

- 概率
- 期望
- 统计
- 正态分布

1 0 10 20 25

ΣP·V

Mean Median Std



课程相关资料

学习资料

欢迎大家扫码或者添加微信好友ai_flare（学习小助手），加入学习群，老师会在群里和大家进行交流和答疑（名额有限、人满即止）

一个典型的机器学习：线性回归

上一章回顾

- 概率
- 期望
- 统计
- 正态分布

人工智能、机器学习、与深度学习

AI

人工智能：专家系统/物理模型等

机器学习：线性回归/kNN/SVM等

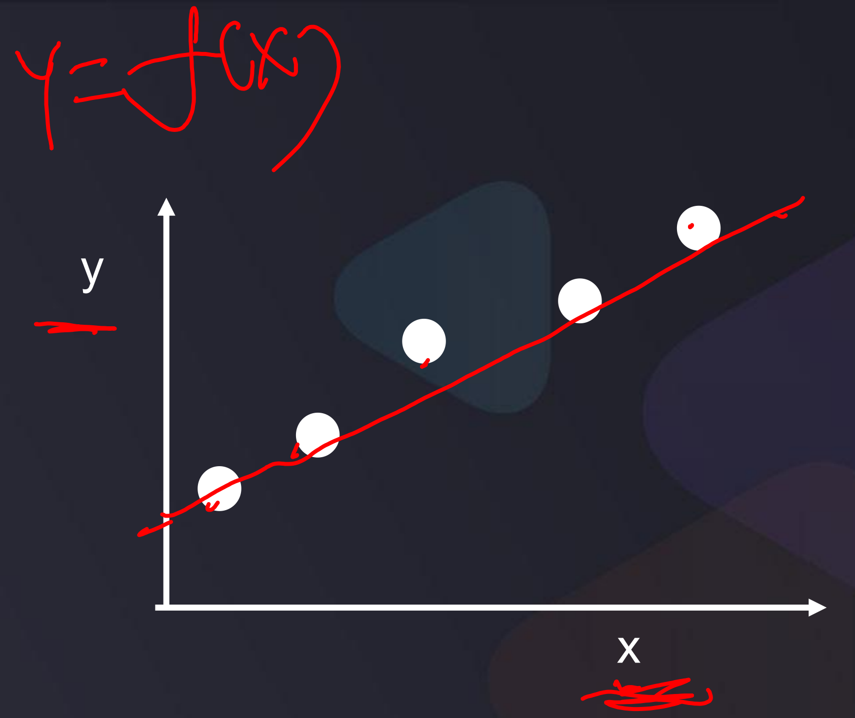
深度学习：
1.全连接神经网络
2.卷积神经网络
3.循环神经网络

ANN
MLP
CNN
RNN

机器学习的本质

以数据为驱动，自主的从数据中寻找规律的方法

- 更少的人为假设和定义
- 更灵活的应用
- 更高的准确度



$$y = 2x + 1$$

$$y = ax + b$$

$$y = f(x)$$

EDU

常见的机器学习方法

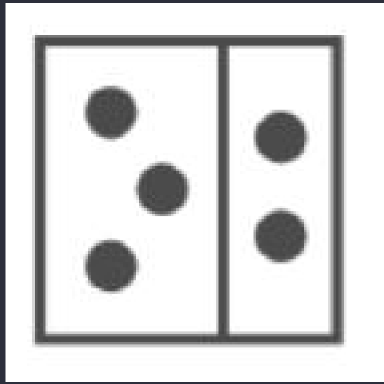
KNN
SVM
CNN
RNN
模型
机器学习
$f(x)$

非监督学习

监督学习

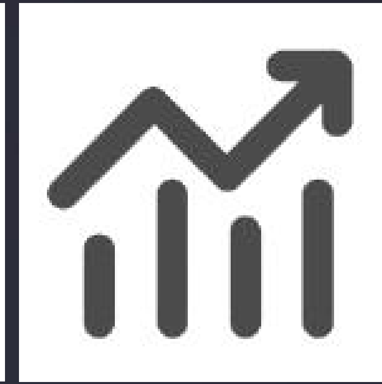
强化学习

优化

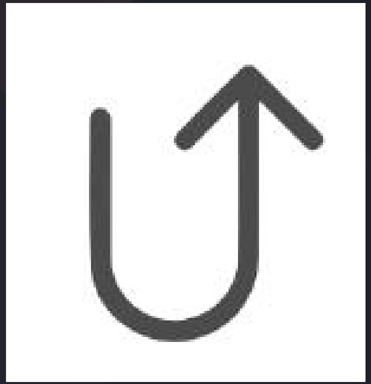


聚类

分类

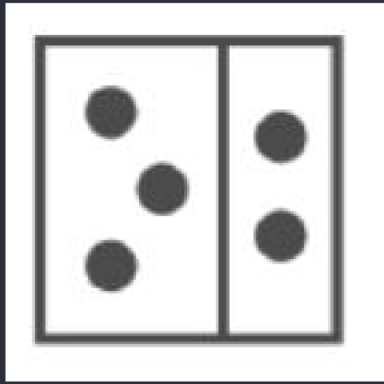


量化预测



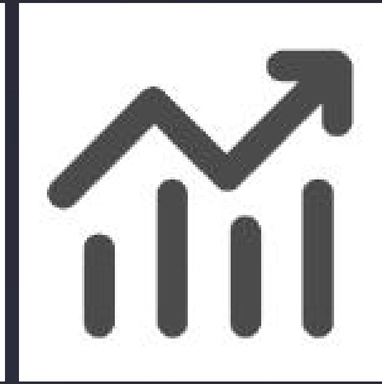
常见的机器学习方法

非监督学习



聚类

监督学习

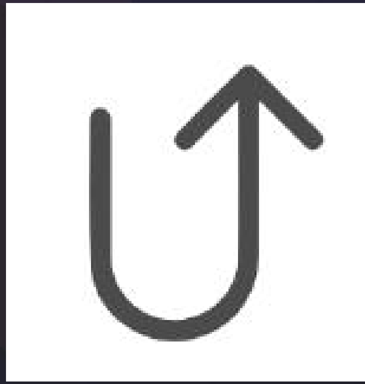


分类 ⟷ 量化预测

回归

强化学习

优化



GA

$Y = f(x)$

回归

Regression

根据已有的数据，建立因变量Y与自变量X的定量（函数）关系的模型

$$y = f(x_1, x_2, \ldots x_n)$$

线性回归

$$y = \mathrm{ax} + \mathrm{b}$$

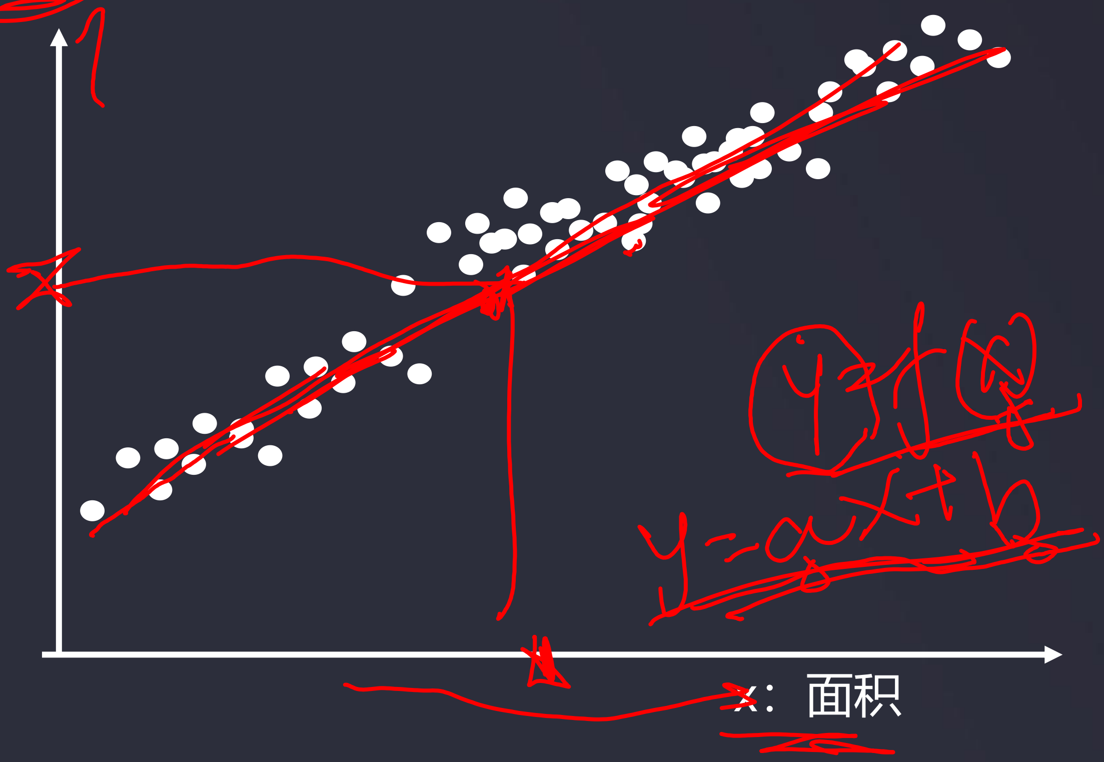
非线性回归

$$y = \mathrm{a}x^2 + \mathrm{bx} + \mathrm{c}$$

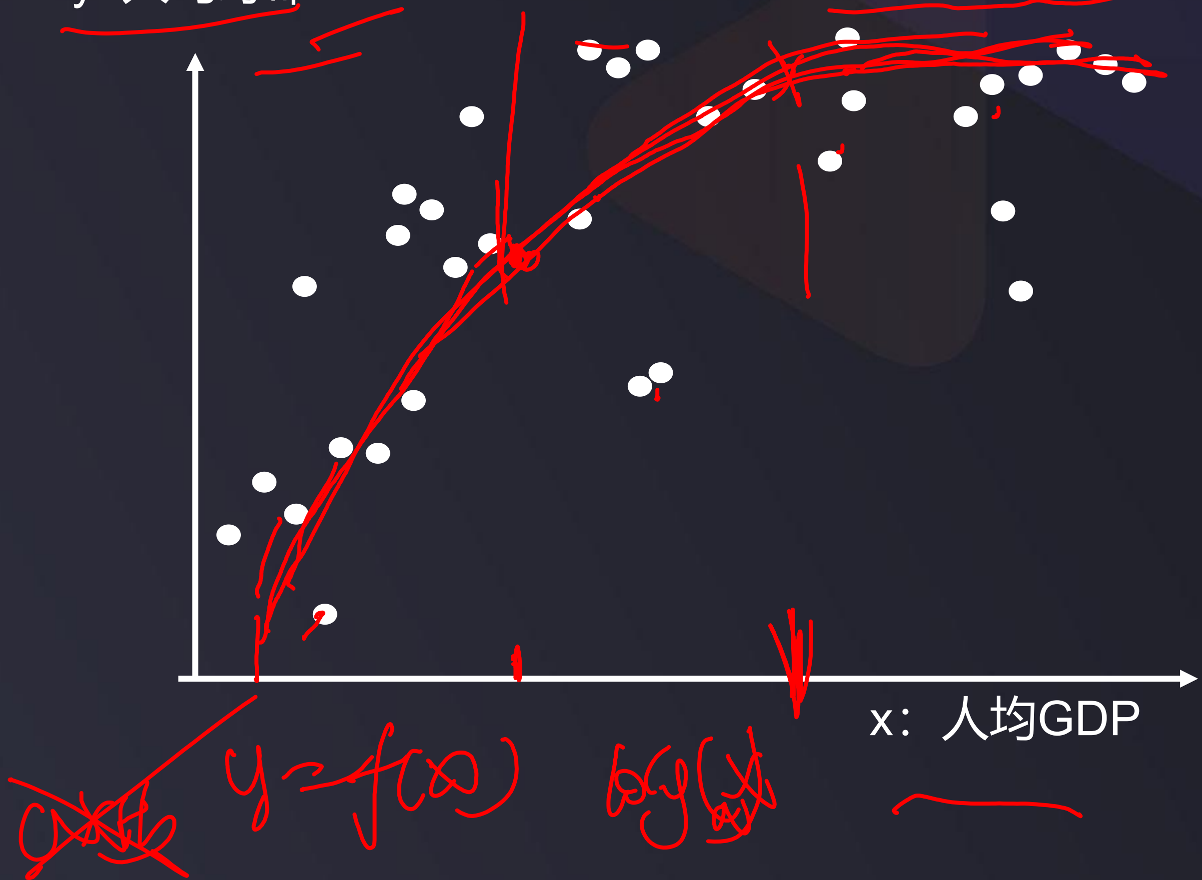
$\sqrt{x}$

回归的应用例子

在同一城区根据房屋面积预测房价



根据人均GDP预测人均寿命



线性回归

因变量Y与自变量X的关系为一次方函数关系

$$y = ax + b$$

$y = a_1x_1 + a_2x_2 + a_3x_3 + \cdots + a_nx_n + b$

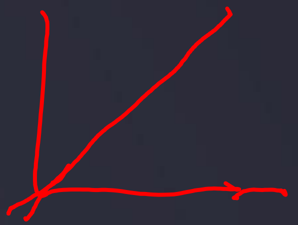
案例:

- 居家用水量（自变量）和水费（因变量）的关系
- 车速（自变量）和行驶距离(因变量)

$y = 2x$

$a = 2$

$d = vt$

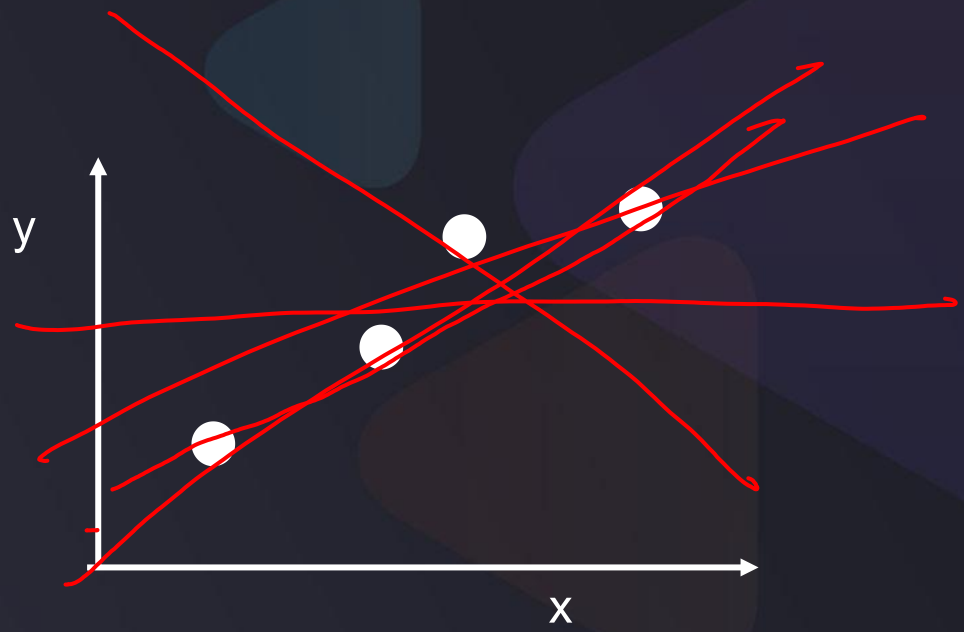


求解线性回归

$y = f(x)$

某流水线生产时长和产品产量的数据如下，求解线性模型

x 时长（分钟）	y 产量（个）
20	195
30	305
50	480
60	580



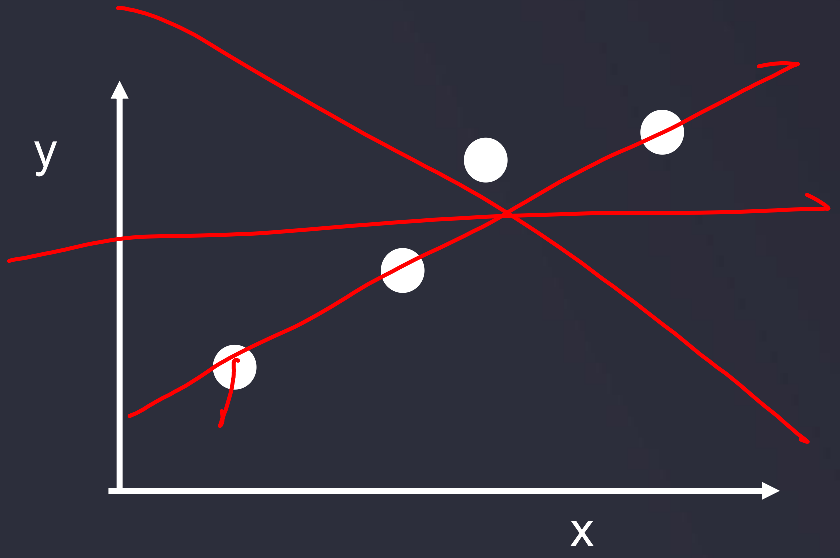
$y = ax + b$

$1x + 10$

$-2x + 5$

求解线性回归

评价回归模型表现的好坏



均方误差

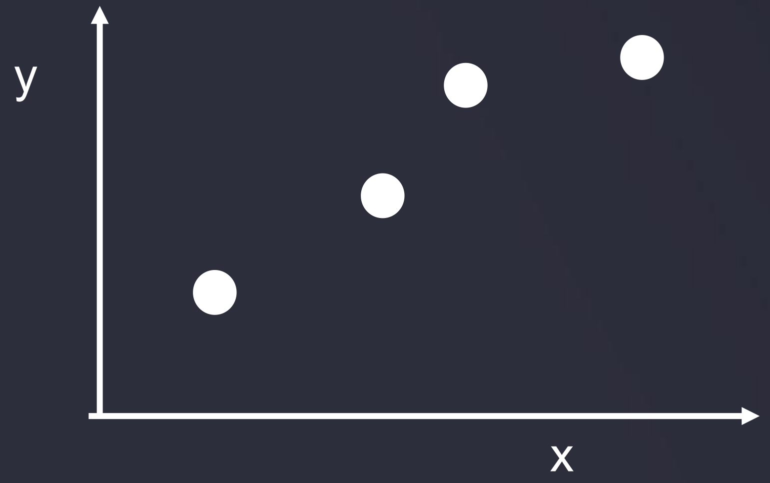
通过回归求解，使得模型预测值$\hat{y}$尽可能接近真实测量y

最小化$\sum(\hat{y}-y)^2$

求解线性回归

EDU

评价回归模型表现的好坏



$y = 10x + 5$

$$\sum(\hat{y} - y)^2 = 1350$$

时长（分钟）	产量（个）
20	195
30	305
50	480
60	580

205 100
305 0
505 625
605 625

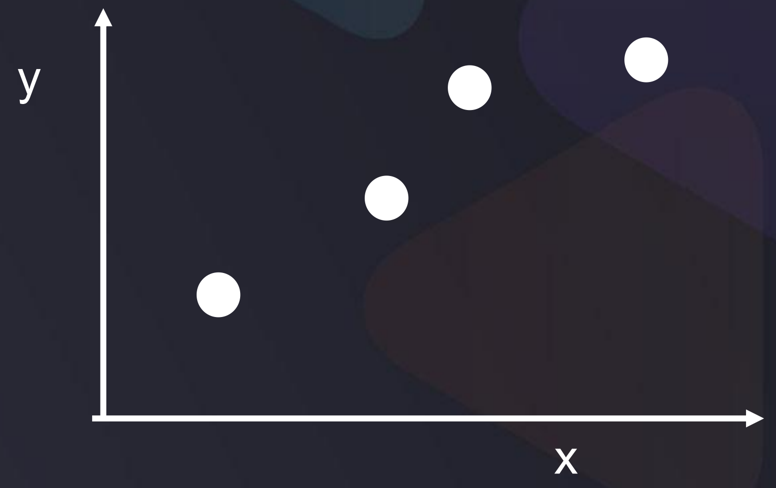
求解线性回归-最小二乘法

某流水线生产时长和产品产量的数据如下，求解线性模型

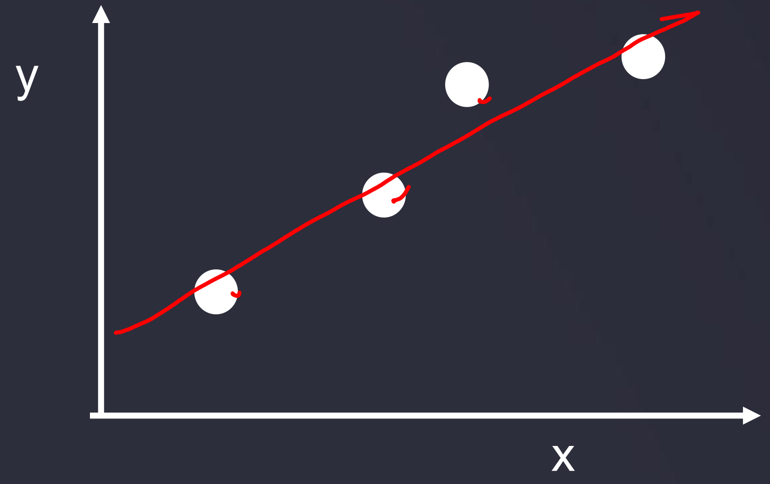
x

时长（分钟）	产量（个）
20	195
30	305
50	480
60	580

y



求解线性回归-最小二乘法



$$y = ax + b$$

$$\sum(\hat{y} - y)^2$$

$$loss = \sum(\hat{Y} - Y)^2$$

$$\frac{\partial loss}{\partial a} = 0 \qquad \frac{\partial loss}{\partial b} = 0$$

$$loss = \sum(ax + b - Y)^2 \qquad (x^2)' = 2x$$

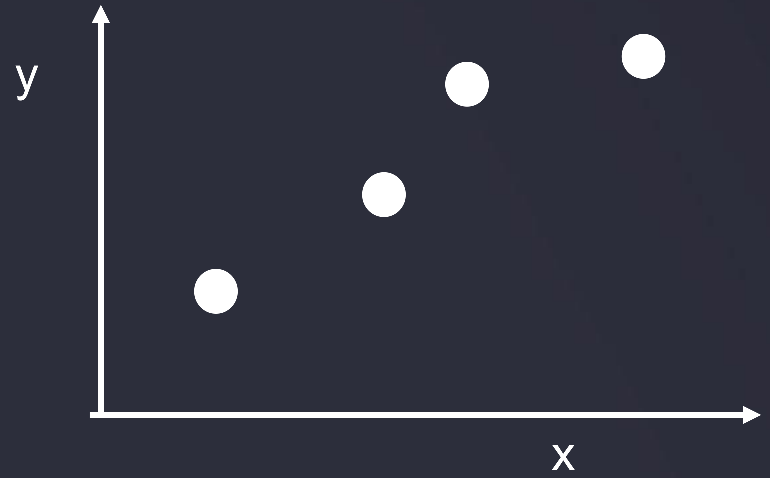
$$\frac{\partial loss}{\partial b} = \sum 2(ax + b - Y) = 0 \qquad \frac{\sum x}{n} = \bar{x}$$

$$= \sum ax + \sum b - \sum Y = 0$$

$$= a\sum X + nb - \sum Y$$

$$= an\bar{x} + nb - n\bar{Y} \rightarrow b = \bar{y} - a\bar{x}$$

求解线性回归-最小二乘法



$$y = ax + b$$

$$\sum (\hat{y} - y)^2$$

$$\frac{\partial \text{loss}}{\partial a} = 0$$

$$b = \bar{y} - a\bar{x}$$

$$\sum (ax + b - y)^2$$

$$= \sum 2(ax + b - y)x$$

$$= \sum 2(ax + \bar{y} - a\bar{x} - y)x$$

$$= \sum \cancel{2}(ax^2 + x\bar{y} - ax\bar{x} - xy) = 0$$

$$\sum (xy - x\bar{y}) = a\sum(x^2 - \bar{x}x)$$

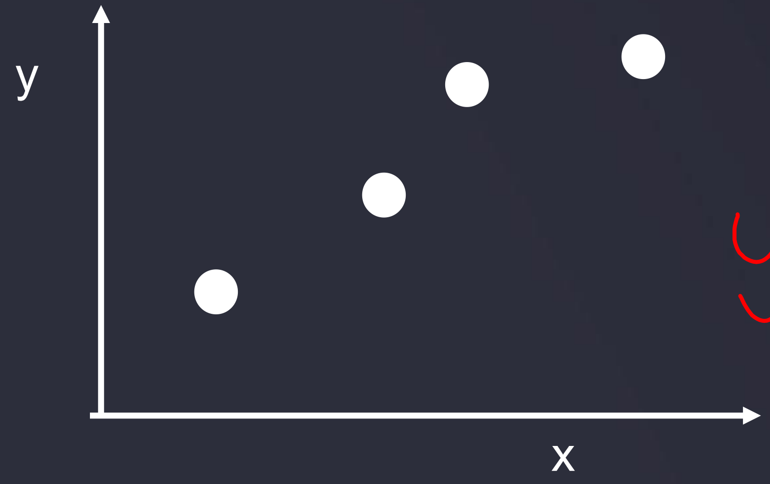
$$a = \frac{\sum(xy - x\bar{y} - \bar{x}y + \bar{x}\bar{y})}{\sum(x^2 - \bar{x}x - \bar{x}x + \bar{x}^2)} = \frac{\sum(x-\bar{x})(y-\bar{y})}{(x-\bar{x})^2}$$

求解线性回归-最小二乘法

$a=9.45$

$b=390-9.45\times40=12$

$$a=\frac{\sum_{i=1}^{n}\left(x^{(i)}-\bar{x}\right)\left(y^{(i)}-\bar{y}\right)}{\sum_{i=1}^{n}\left(x^{(i)}-\bar{x}\right)^{2}}$$



$y=9.45X+12$

$$b=\bar{y}-a\bar{x}$$

$$y=ax+b$$

$$\sum(\hat{y}-y)^2$$

时长（分钟）	产量（个）
20	195
30	305
50	480
60	580

$\bar{X}=160/4=40$　　$1560/4=390$

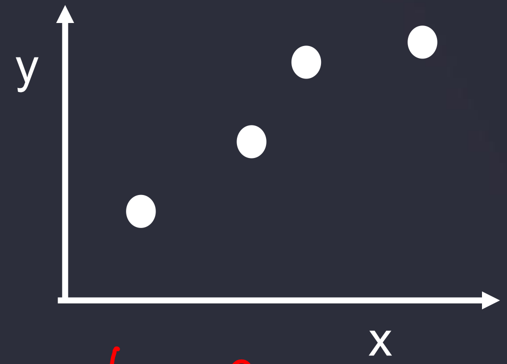
$$\frac{-20\times-195+-10\times-85+10\times90+20\times190}{}$$

$$=\frac{3900+850+900+3800}{400+100+100+400}=\frac{9450}{1000}=9.45$$

EDU

求解线性回归-梯度下降法

$y = ax + b$



最小化 $g(a, b) = \sum(\hat{y} - y)^2$

$J = f(p)$ ⟹ $p_{i+1} = p_i - a\frac{\partial}{\partial p_i}f(p_i)$

搜索方法

$$a_{new} = a_{old} - a\frac{\partial}{\partial a}g(a,b) = a_{old} - a\frac{1}{n}\sum(a_{old}x_i + b_{old} - y_i)x_i$$

$$b_{new} = b_{old} - a\frac{\partial}{\partial b}g(a,b) = b_{old} - a\frac{1}{n}\sum(a_{old}x_i + b_{old} - y_i)$$

$g(a) = \sum(\hat{y} - y)^2$

0.1 - 0.001

100 1000

$a_{new} - a_{old} < \theta$

EDU

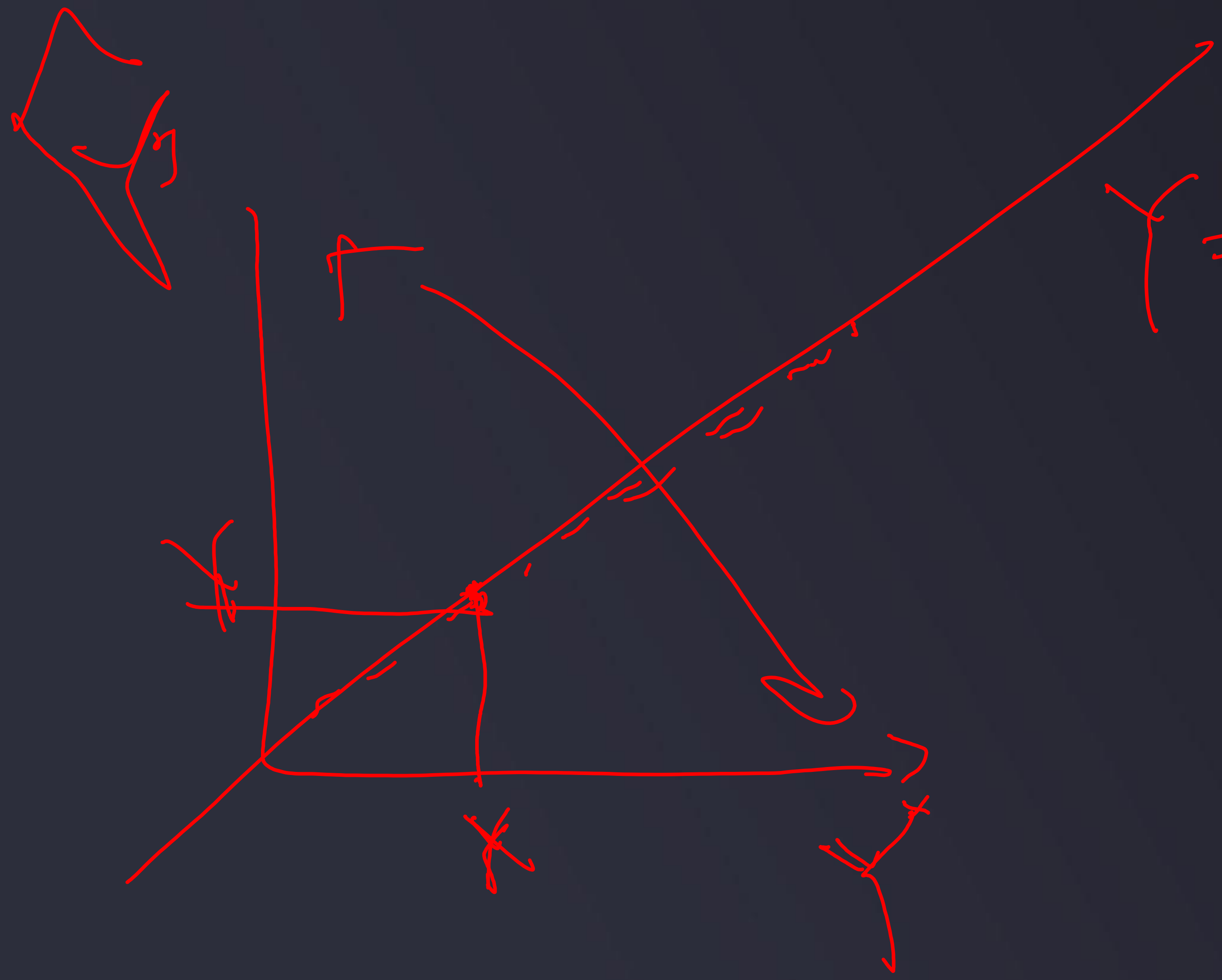
实战习题

自变量X

Y

样本	家庭收入	人数	食品消费
1	7900	3	3680.61
2	17200	3	4681.39
3	10000	2	3192.53
4	20900	3	4764.64
5	22000	3	5003.52
6	5500	4	4079.39
7	5300	2	2550.58
8	8300	2	2830.24

在某小区抽样调查获得家庭月收入，家庭人数，家庭每月食品支出的调查数据，请过通过线性回归建立家庭月收入，家庭人数为自变量，每月食品支出为因变量的多元线性回归模型



$Y = 1X$

本章回顾

- 人工智能与机器学习的基本概念
- 回归Regression与线性回归
- 线性回归-最小二乘法求解
- 线性回归-梯度下降求解

课程相关资料

学习资料

欢迎大家扫码或者添加微信好友ai_flare（学习小助手），加入学习群，老师会在群里和大家进行交流和答疑（名额有限、人满即止）

人工智能模型的应用与配套：逻辑运算

上一章回顾

- 人工智能与机器学习的基本概念
- 回归Regression与线性回归
- 线性回归-最小二乘法求解
- 线性回归-梯度下降求解

什么是逻辑运算？

+ - × ÷

逻辑运算又称布尔运算，是通过数学方法研究逻辑问题来建立了逻辑演算。

布尔值（Bool）：$\begin{cases} \text{是: } True/1/\text{非零} \\ \text{否: } False/0 \end{cases}$

10 进制

2 进制

0 1 10 11 100

101

bit

下雨 晴

否 0

1/0 Statement

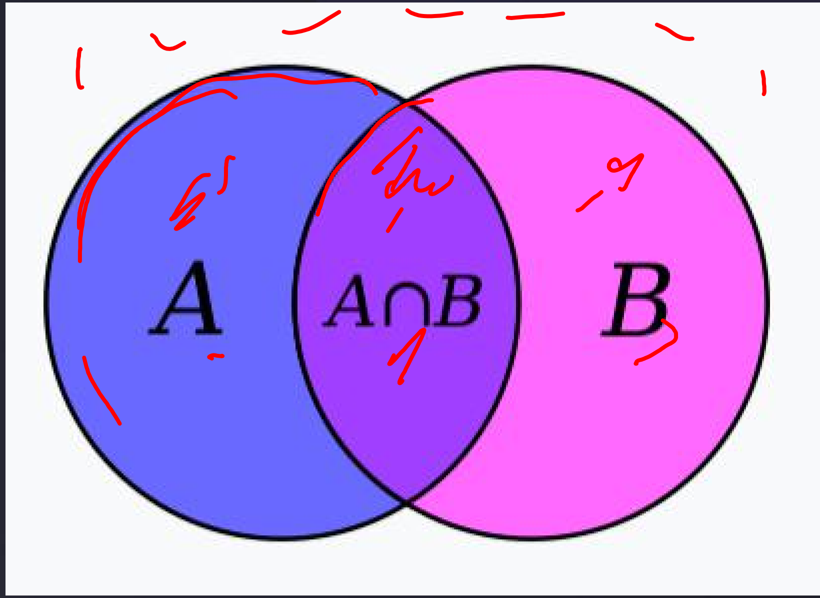
[0 ~ 10]

1 True

False 0

逻辑运算符

- 和/与/and/&：当运算符两边两者同时为True，结果为True，否则为False
- 或/Or/|:　当运算符两边有一个为True，结果为True，否则为False
- 否/非/Not/!:　反转被运算对象的布尔值



EDU

比较运算符

a=3

3==4 → 0 False 假

3!=4 → 1 True

名称	符号	定义
等于	==	比较两个对象是否相等
不等于	! =	比较两个对象是否不相等
大于	>	比较符号前对象是否大于后者
小于	<	比较符号前对象是否小于后者
大于等于	>=	比较符号前对象是否大于或等于后者
小于等于	<=	比较符号前对象是否小于或等于后者

3==5

a>3 and a<5

4 1 True

5>3 → 1

5<3 → 0

5>=5 → 1

5>5 → 0

+ - × ÷

逻辑运算的优先顺序

从高往低排列

括号内的运算（）

指数（** ^）

乘（*）、除（/）、取模（%）、取整（//）

加（+），减（-）

比较运算符（> < >= <=）

比较运算符（== !=）

逻辑运算符（and or not）

EDU

逻辑运算案例

1+5*3 and (3>4 or (not True))

(1+1)/2 > 3 or (1+2)*(4/2)

False

True False

False or False

False

1+5X3 and False

16 and False

= False

2/2>3 or 3X2

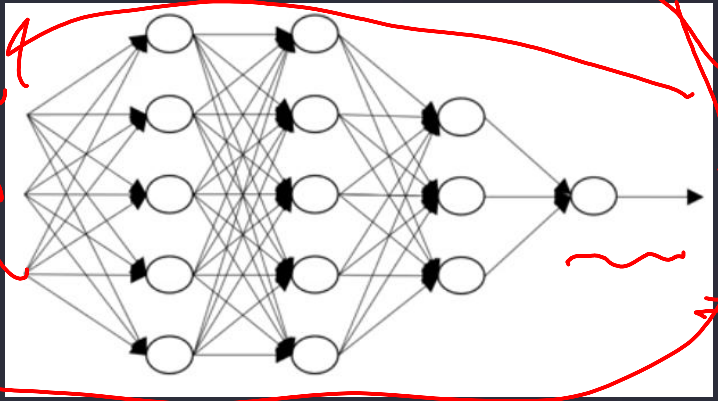
1>3 or 6

False or 6 True

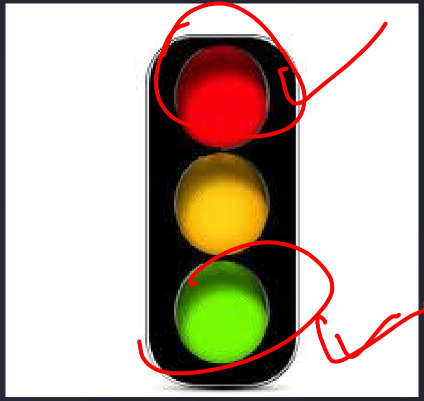
True

逻辑运算常用来搭建AI配套模型来实现具体应用

一个关于贷款违约率预测的模型



自动驾驶路面状态识别模型



本章回顾

False True
0 1
假 真

5+3*2

- 逻辑运算与布尔值
- 逻辑运算符

and or 非

- 比较运算符

== > < >= <=

- 逻辑运算的优先顺序

( ) ** × ÷ + - > < == and or not

- 逻辑运算与人工智能的应用案例

EDU

课程相关资料

学习资料

欢迎大家扫码或者添加微信好友ai_flare（学习小助手），加入学习群，老师会在群里和大家进行交流和答疑（名额有限、人满即止）

CSDN学院 IT实战派

强化学习的基础：马尔科夫链

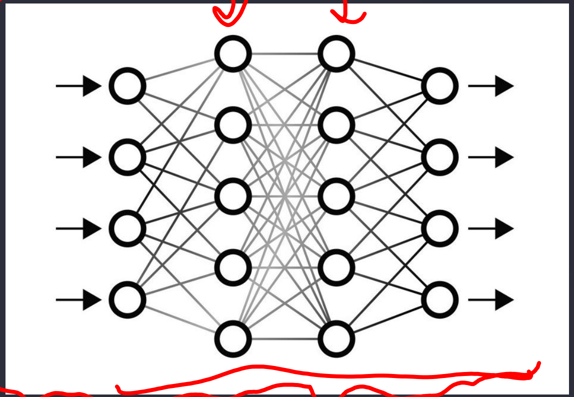
上一章回顾

- 逻辑运算与布尔值
- 逻辑运算符
- 比较运算符
- 逻辑运算的优先顺序
- 逻辑运算与人工智能的应用案例

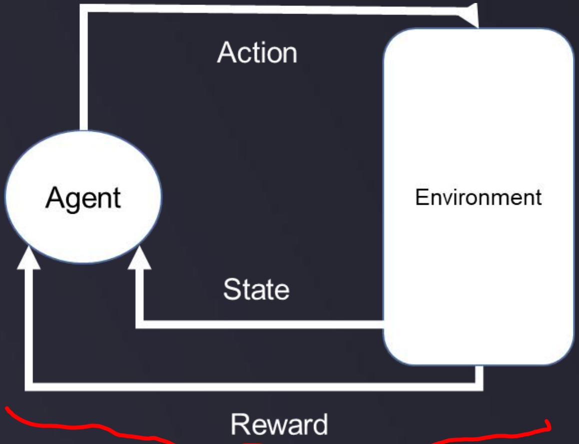
深度强化学习

KNN
SVM

ANN
MLP CNN RNN



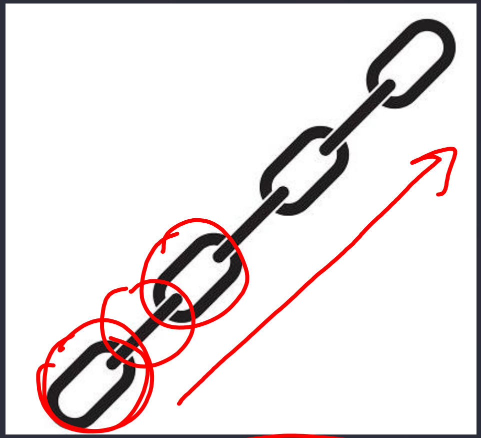
+



Unsupervised
Supervised
reinforcement

马尔科夫链的定义

Markov chain



区块链

状态空间中经过从一个状态到另一个状态的转换的随机过程，下一状态的概率分布只能由当前状态决定，且与它前面的事件均无关。

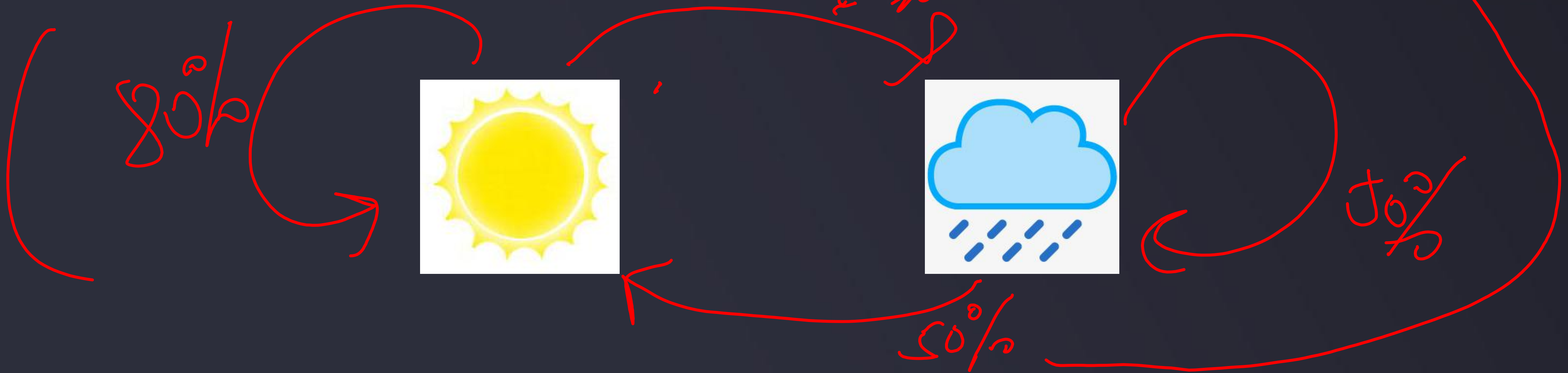
$$P(S_{t+1}|S_t, S_{t-1}, \dots, S_0) = P(S_{t+1}|S_t)$$

一个马尔科夫链的例子：明天的天气

(S, P)

假设有两种天气状态（晴天和下雨），第二天的天气状态只取决前一天的天气状态

- 第一天晴天，第二天晴天（80%），下雨（20%）
- 第一天下雨，第二天晴天（50%），下雨（50%）



状态转移矩阵

假设有三种天气状态（晴天、阴天、下雨），第二天的天气状态只取决前一天的天气状态

- 第一天晴天，第二天晴天（70%），阴天（20%），下雨（10%）
- 第一天阴天，第二天晴天（40%），阴天（40%），下雨（20%）
- 第一天下雨，第二天晴天（20%），阴天（40%），下雨（40%）

$$
\begin{array}{cc}
 & \begin{array}{ccc} \text{晴} & \text{阴} & \text{下雨} \end{array} \\
\begin{array}{c} \text{晴} \\ \text{阴} \\ \text{下雨} \end{array} &
\begin{bmatrix} 0.7 & 0.2 & 0.1 \\ 0.4 & 0.4 & 0.2 \\ 0.2 & 0.4 & 0.4 \end{bmatrix}
\end{array}
$$

$[0.5\ 0.5\ 0] = [0.55, 0.3, 0.15]$

$[1\ 0\ 0] = [0.7, 0.2, 0.1]$

$[0\ 0\ 1] \rightarrow [0.2\ 0.4\ 0.4]$

状态转移矩阵

假设有三种天气状态（晴天、阴天、下雨），第二天的天气状态只取决前一天的天气状态

- 第一天晴天，第二天晴天（70%），阴天（20%），下雨（10%）
- 第一天阴天，第二天晴天（40%），阴天（40%），下雨（20%）
- 第一天下雨，第二天晴天（20%），阴天（40%），下雨（40%）

$[1, 0, 0] \times$

$[0.7, 0.2, 0.1] \times$

$$\begin{bmatrix} 0.7 & 0.2 & 0.1 \\ 0.4 & 0.4 & 0.2 \\ 0.2 & 0.4 & 0.4 \end{bmatrix}$$

$[0.7 \quad 0.2 \quad 0.1]$

$$\begin{array}{ccc} 0.49 & 0.14 & 0.07 \\ +0.08 & 0.08 & 0.04 \\ 0.02 & 0.04 & 0.04 \end{array}$$

$\rightarrow [0.59, \quad 0.26 \quad 0.15]$

马尔可夫链收敛和平稳条件

1. 可能的状态数是有限的。
2. 状态间的转移概率需要固定不变。
3. 从任意状态能够转变到任意状态。
4. 不能是简单的循环，例如全是从x到y再从y到x。

$M=(S\ P)$

马尔科夫奖励过程

马尔科夫过程（Markov Process）主要描述的是状态之间的转移关系，在各个状态的转移过程中赋予不同的奖励值就得到了马尔科夫奖励过程。由一个四元组组成(S, P, R, γ)

S表示状态集合（晴天/阴天/下雨）

P表示状态转移矩阵

R表示奖励函数：$R(s) = E[R_{t+1} | S_t]$

$\gamma \in [0,1]$表示衰减因子

0.1

0.01

不同的的天气状态会给人带来不同的心情状态（奖励）

- 晴天+2
- 阴天+0
- 下雨-1

马尔科夫决策过程：强化学习的基本过程

马尔科夫决策过程(Markov Decision Process, MDP)相比马尔科夫奖励过程多了一个动作 A,它可以用一个五元组(S,A,P,R,γ)

- S表示状态集合
- A表示决策过程的集合
- P表示状态转移矩阵 $P[S_{t+1} \mid S_t, A_t]$
- R表示奖励函数：$R(s)=E[R_{t+1} \mid S_t,\ A_t]$
- $\gamma \in [0,1]$表示衰减因子

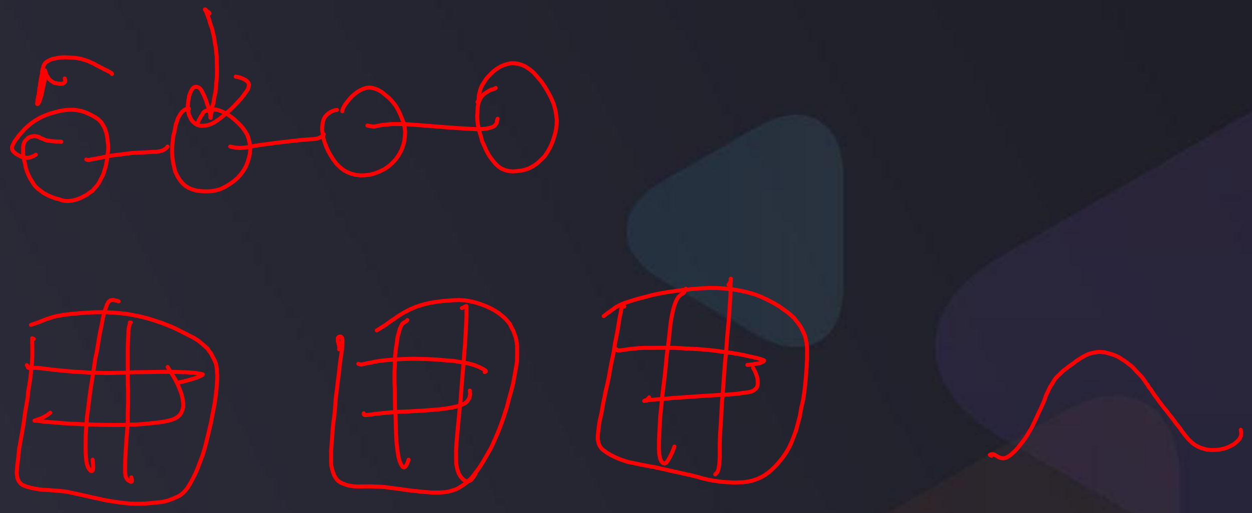
episode

马尔科夫决策过程：强化学习的基本过程

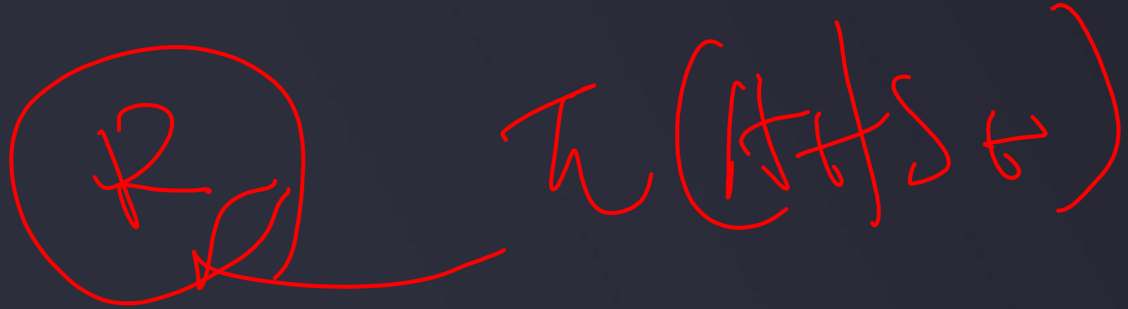
强化学习的目标就是最大化期望回报,相应的结果就是找到从状态空间S映射到动作空间A的最优策略,

$$\pi[A_t|S_t]$$

本章回顾



- 马尔科夫链的定义
- 状态转移矩阵
- 马尔科夫奖励过程
- 马尔科夫决策过程与强化学习

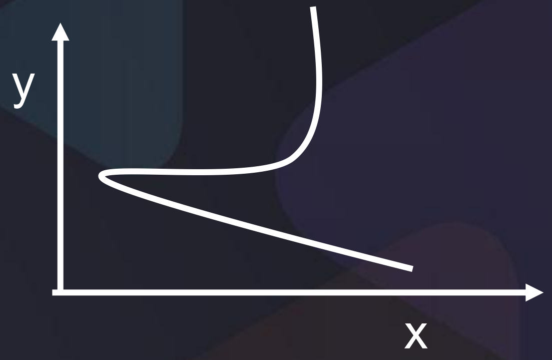


课程回顾

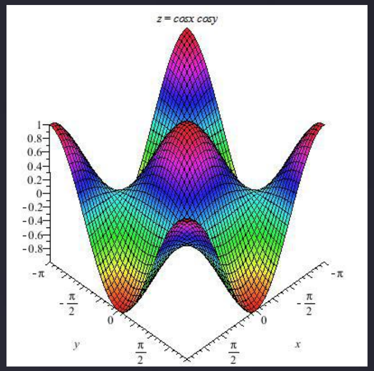
- 人工智能的本质-函数
- 神经网络的底层机理-线性代数
- 深度学习的学习原理-微积分
- AI模型的评估和优化-概率与统计
- 入门经典的机器学习-线性回归
- 补充小章节1：人工智能模型的应用与配套-逻辑运算
- 补充小章节2：强化学习的基础-马尔科夫链

课程回顾：函数

- 函数的定义
- 常见函数类型
- 复合函数
- 分段函数
- 函数与反函数
- 函数的单调性，奇偶性，周期性
- 多元函数与人工智能



$$f\left(\frac{1}{3x}\right)=\frac{1-\frac{3}{x^2}}{3x+2}$$



课程回顾：矩阵运算

- 矩阵的定义和应用
- 同型矩阵
- 矩阵的加减，数乘，乘法等运算
- 矩阵的转置
- 向量的定义和运算

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{12} \\ a_{21} & a_{22} & \cdots & a_{22} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

$$AB \neq BA$$

$$(AB)C = A(BC)$$

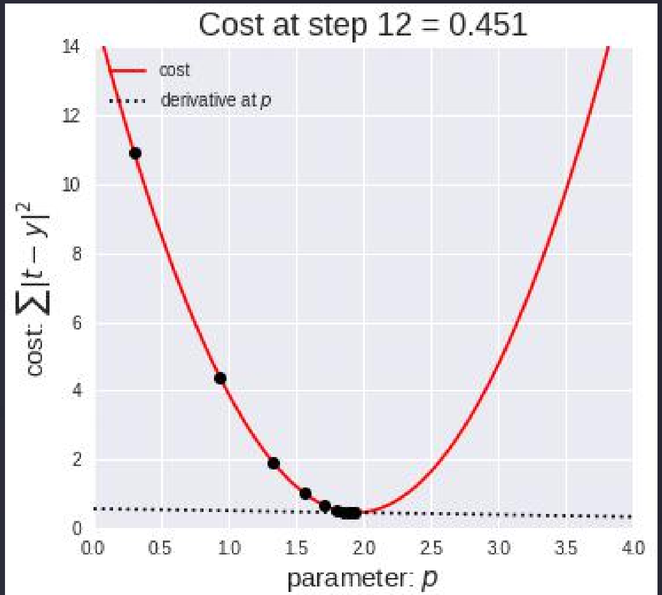
$$A\,(B + C) = AB + AC$$

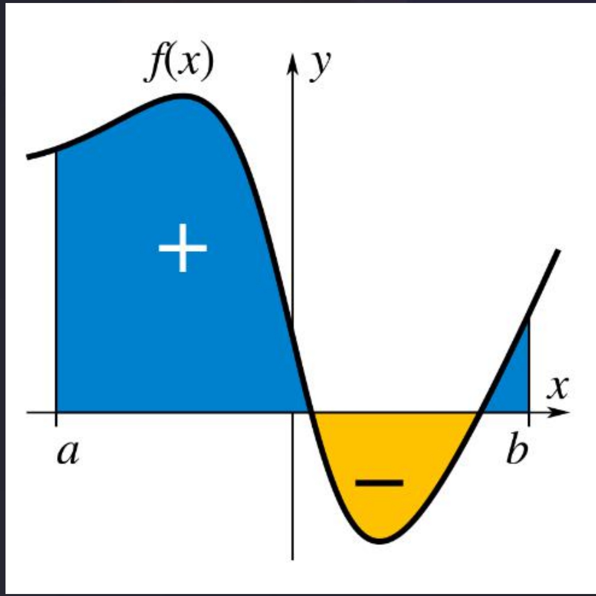
$$A = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \qquad A = [a_1 \quad a_2 \quad \cdots \quad a_n]$$

课程回顾：微积分

$$A = \lim_{x \to x_0} f(x)$$

- 极限
- 导数
- 极值点
- 梯度下降
- 积分

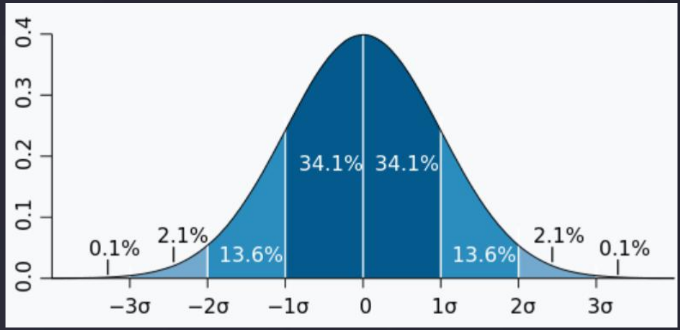




课程回顾：统计与概率

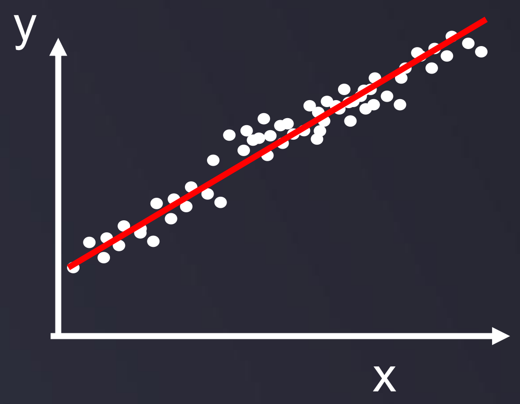
- 概率
- 期望
- 统计
- 正态分布

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \qquad \sigma = \sqrt{\sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n}}$$



课程回顾：线性回归

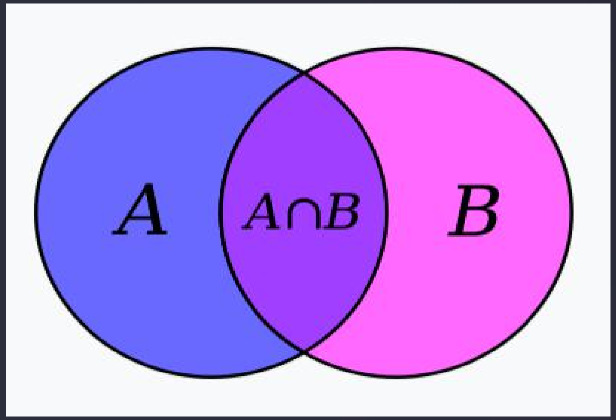
- 人工智能与机器学习的基本概念
- 回归Regression与线性回归
- 最小二乘法求解
- 梯度下降求解



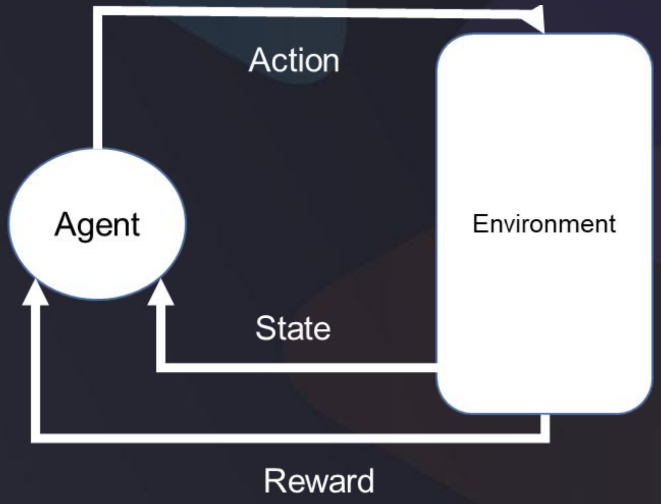
$$a=\frac{\sum_{i=1}^{n}\left(x^{(i)}-\bar{x}\right)\left(y^{(i)}-\bar{y}\right)}{\sum_{i=1}^{n}\left(x^{(i)}-\bar{x}\right)^{2}}$$

$$b=\bar{y}-a\bar{x}$$

课程回顾：逻辑运算与马尔科夫链小章节



- 逻辑运算与布尔值
- 逻辑运算符
- 比较运算符
- 逻辑运算的优先顺序
- 逻辑运算与人工智能的应用案例



- 马尔科夫链的定义
- 状态转移矩阵
- 马尔科夫奖励过程
- 马尔科夫决策过程与强化学习

课程相关资料

学习资料

欢迎大家扫码或者添加微信好友ai_flare（学习小助手），加入学习群，老师会在群里和大家进行交流和答疑（名额有限、人满即止）