北京航空航天大學
COLLEGE OF SOFTWARE BEIHANG UNIVERSITY 软件学院

人工智能

第2讲：机器学习-有监督学习I

线性回归与线性分类

张晶

2025春季

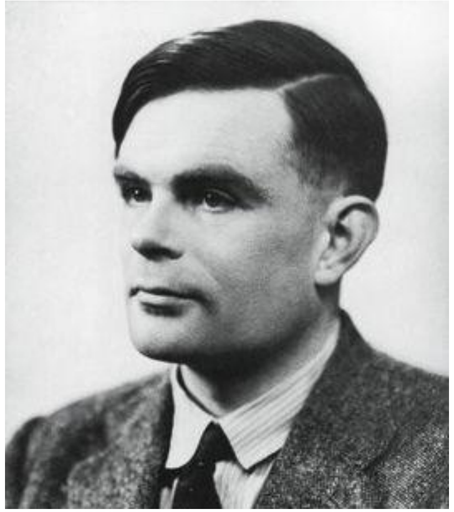
- 参考资料： 吴飞，《人工智能导论：模型与算法》，高等教育出版社
- 在线课程： https://www.icourse163.org/course/ZJU-1003377027?from=searchPage

回顾：什么是人工智能

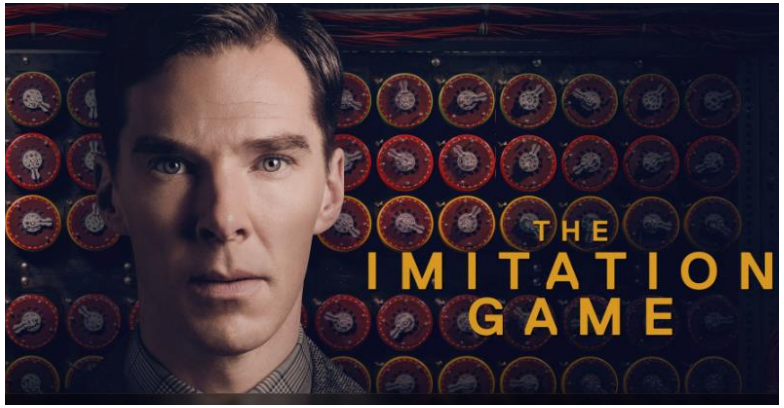
- 从Turing的Imitation Game，讲到Turing测试
- 把人工智能的核心领域与大脑功能做初步的对应

像人一样思考的系统
…… 有头脑的机器
……模拟人类的思维

像人一样行动的系统
…执行人需要智能完成的功能



Alan Turing
阿兰·图灵



根据图灵故事拍摄的电影

回顾：人工智能的发展历史与脉络：**人工智能关键历史节点**

- 1950 图灵发表“计算与智能”
- 1956 达特茅斯会议
- 1957 罗森布拉特发明“感知机”
- 1969 明斯基和佩珀特发表《感知机：计算机几何学》
- 1974 -1980 第一次低谷
- 1980 专家配置系统XCON
- 1982 霍普菲尔德网络
- 1987 -1993 第二次低谷
- 1997 深蓝击败人类棋手
- 2006 辛顿开辟了深度学习领域
- 2011 IBM的沃森在电视智力竞赛Jeopardy!夺冠
- 2016 AlphaGo击败李世石
- 2020 AlphaFold2蛋白质结构预测精度达到实验水平
- 2022 OpenAI发布ChatGPT

回顾：人工智能三种主流方法区别

学习模式	优势	不足
符号主义（用规则教）	与人类逻辑推理相似，解释性强	难以构建完备的知识规则库
联结主义（用数据学）	直接从数据中学	以深度学习为例：依赖于数据、解释性不强
行为主义（用问题引导）	从经验中进行能力的持续学习	非穷举式搜索而需更好策略

从数据到知识与能力，能力增强是最终目标
三种学习方法的综合利用值得关注！

提纲

一、机器学习基本概念

二、线性回归与线性分类

三、线性判别分析

四、支持向量机

五、决策树

六、集成学习

提纲

主观题 10分

设置

请思考，如何实现对这些包含数字图像的自动识别/分类?

0	1	2	3	4	5	6	7	8	9

[0,0,1,1,0,0]

注：图片可转换为矩阵，空白区域用0表示，字迹区域用1表示：

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

作答

传统编程

0	1	2	3	4	5	6	7	8	9

[0,0,1,1,0,0]

输入数据x

程序规则f

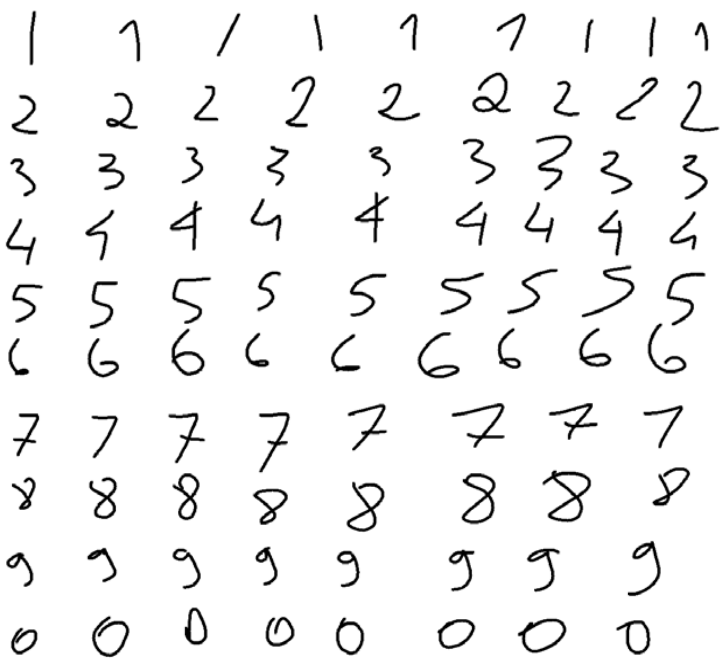
传统编程

输出结果y

```
def classification(x):
  if x==[0,0,1,1,0,0]:
    y=0
  elif x==[0,0,1,0,0,0]:
    y=1
  elif x ==[0,1,1,1,0,0]:
    y=2
  return y
```

下图为经典手写数字识别数据集。

请思考，如何实现对这些**手写数字图像**的自动识别/分类?



Yann LeCun (杨立昆)

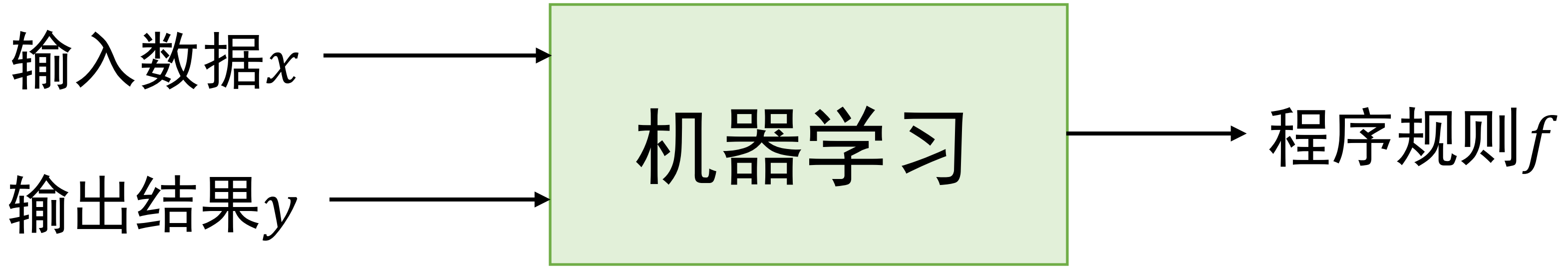
MNIST dataset

http://yann.lecun.com/exdb/mnist/

作答

机器学习 v.s. 传统编程

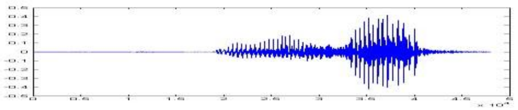

```
def classification(x):
  if x==[0,0,1,1,0,0]:
    y=0
  elif x==[0,0,1,0,0,0]:
    y=1
  elif x ==[0,1,1,1,0,0]:
    y=2
  return y
```




```
x=[[0,0,1,1,0,0],
    [0,0,1,0,0,0],
    [0,1,1,1,0,0]]
y=[0,1,2]
def classification(x,y):
  model=linear()
  model.fit(x,y)
  return model
```

机器学习：从数据中学习知识 $f(x) = y$

- 语音识别

$f($  $)=$ “你好”

- 图像分类

$f($  $)=$ “猫”

- 围棋游戏

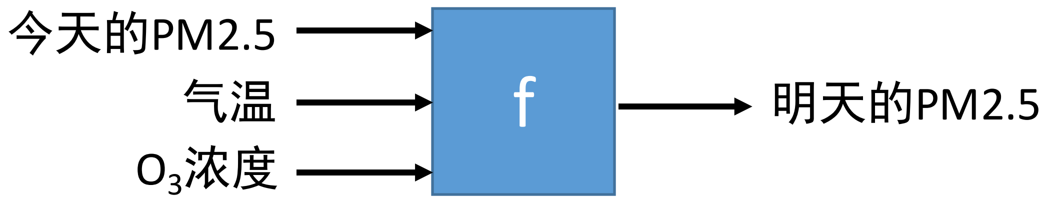
$f($  $)=$ “5–5”

（下一步落子位置）

机器学习的分类（按问题分类）

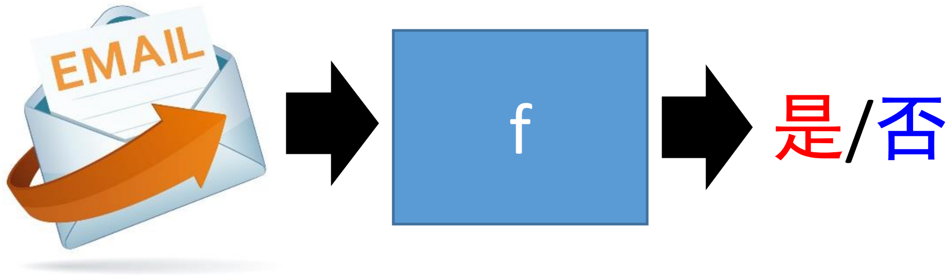
- 回归问题：函数输出标量（连续）

预测PM2.5

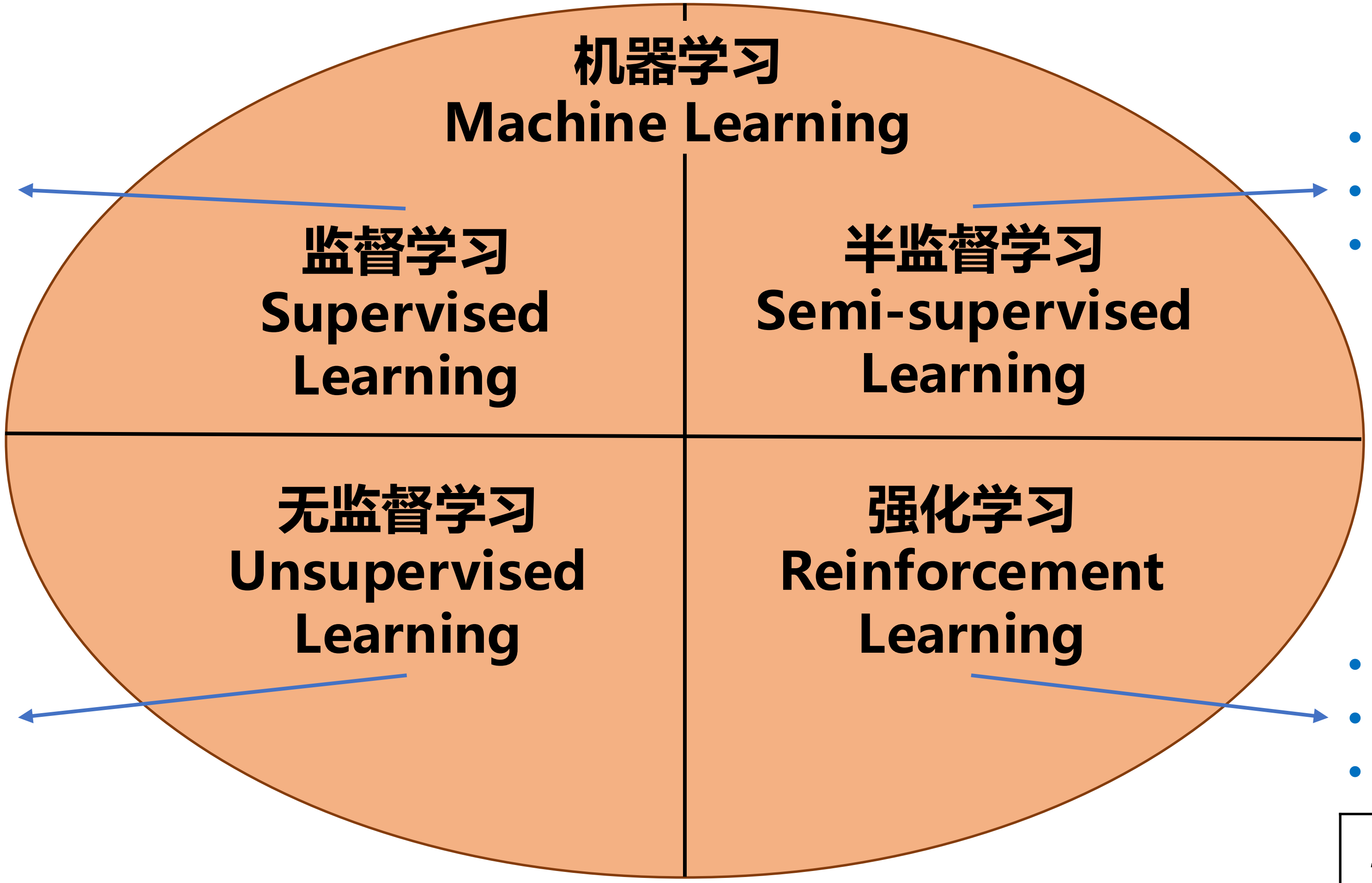


- 分类问题：给定选项（类别），函数输出其中一个正确选项（离散）

垃圾邮件过滤



机器学习的分类（按数据标注情况分类）



- 数据有标签
- 直接反馈
- 预测结果

x_1	x_2	x_3	x_4
y_1	y_2	y_3	y_4

- 部分数据有标签
- 部分反馈
- 预测结果

x_1	x_2	x_3	x_4
y_1	y_2		

- 数据无标签
- 无反馈
- 寻找数据规律

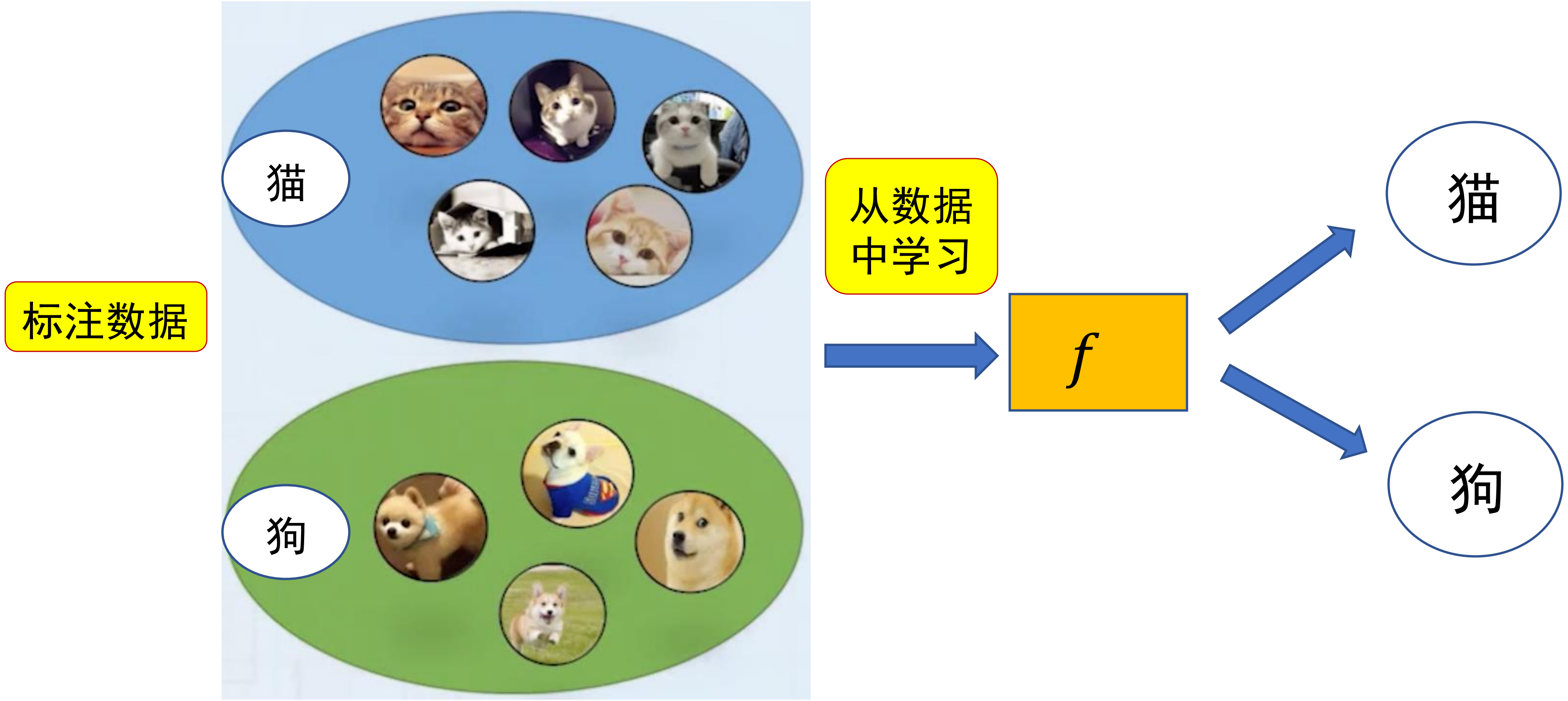
x_1	x_2	x_3	x_4

- 稀疏标签（奖励）
- 稀疏反馈
- 序列决策

$x_1 \rightarrow x_2 \rightarrow x_3 \rightarrow x_4$
y

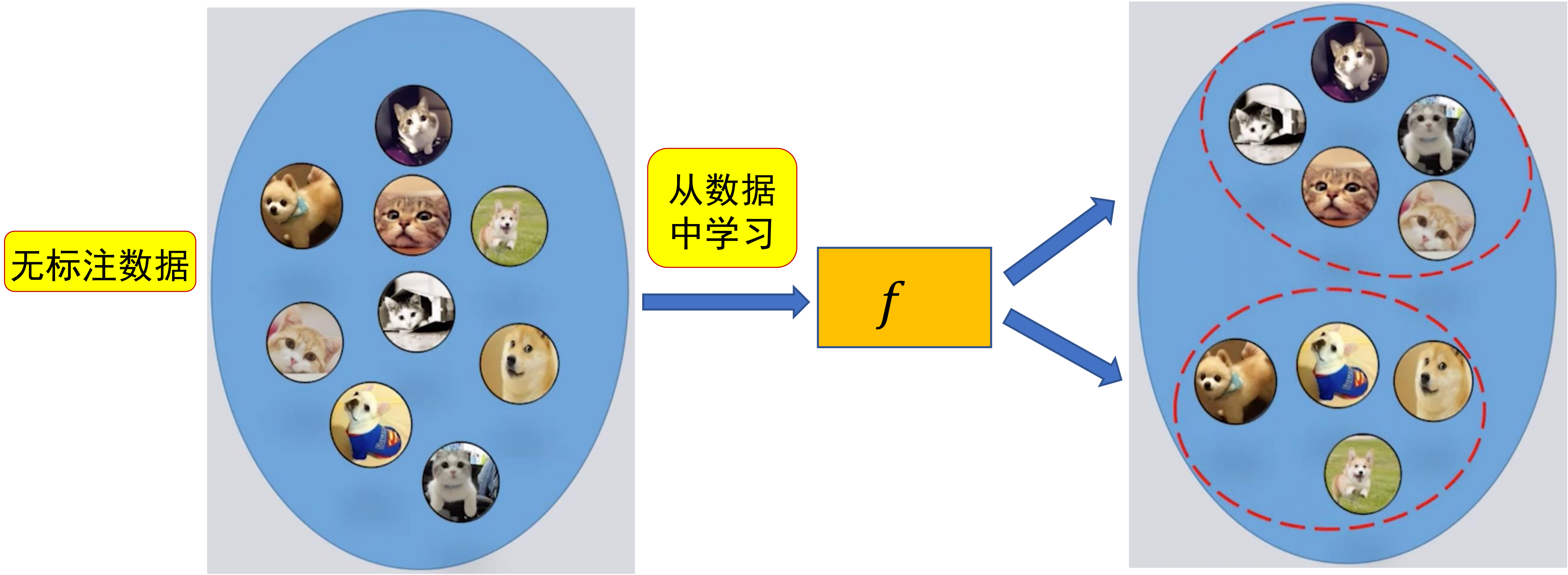
机器学习的分类（按数据标注情况分类）

- 有监督学习（supervised learning）



机器学习的分类（按数据标注情况分类）

- 无监督学习（unsupervised learning）



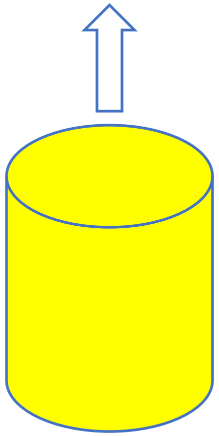
监督学习的重要元素

标注数据	标识了类别信息的数据
学习模型	定义映射模型类型
损失函数	对映射结果进行度量
优化方法	对映射模型进行学习

监督学习：主要要素

训练映射函数f

使得$f(x_i)$预测结果尽量等于y_i



训练数据集

$(\boldsymbol{x_i}, \boldsymbol{y_i}), \boldsymbol{i} = \mathbf{1}, \dots, \boldsymbol{n}$

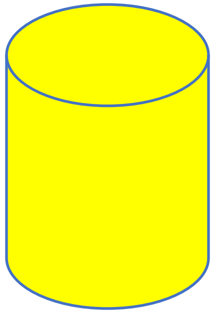
- 训练集中一共有n个标注数据，第i个标注数据记为(x_i, y_i)，
  - x_i：第i个样本数据
  - y_i：x_i的标注信息（真值，groundtruth）。
- 定义一个映射函数f（也被称为模型），f对x_i的预测结果记为$f(x_i)$。
- 损失函数$Loss$就是用来计算x_i真值y_i与预测值$f(x_i)$之间差值的函数。
- 在训练过程中希望优化映射函数，使得在训练数据集上得到“损失”之和最小，即$\min \sum_{i=1}^{n} Loss(f(x_i), y_i)$。

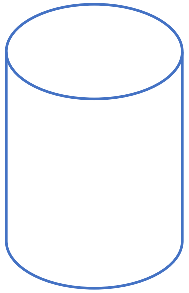
监督学习：训练数据与测试数据

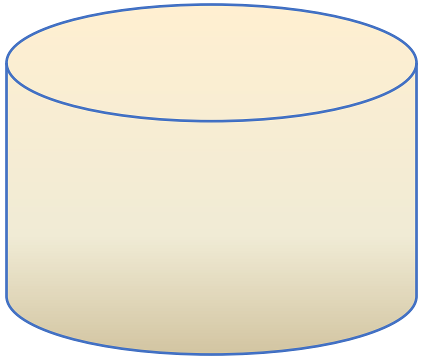
从训练数据集学习
映射函数f的未知参数

在测试数据集
测试映射函数f

未知数据集
上测试映射函数f







训练数据集
$(x_i, y_i), i = 1, \ldots, n$

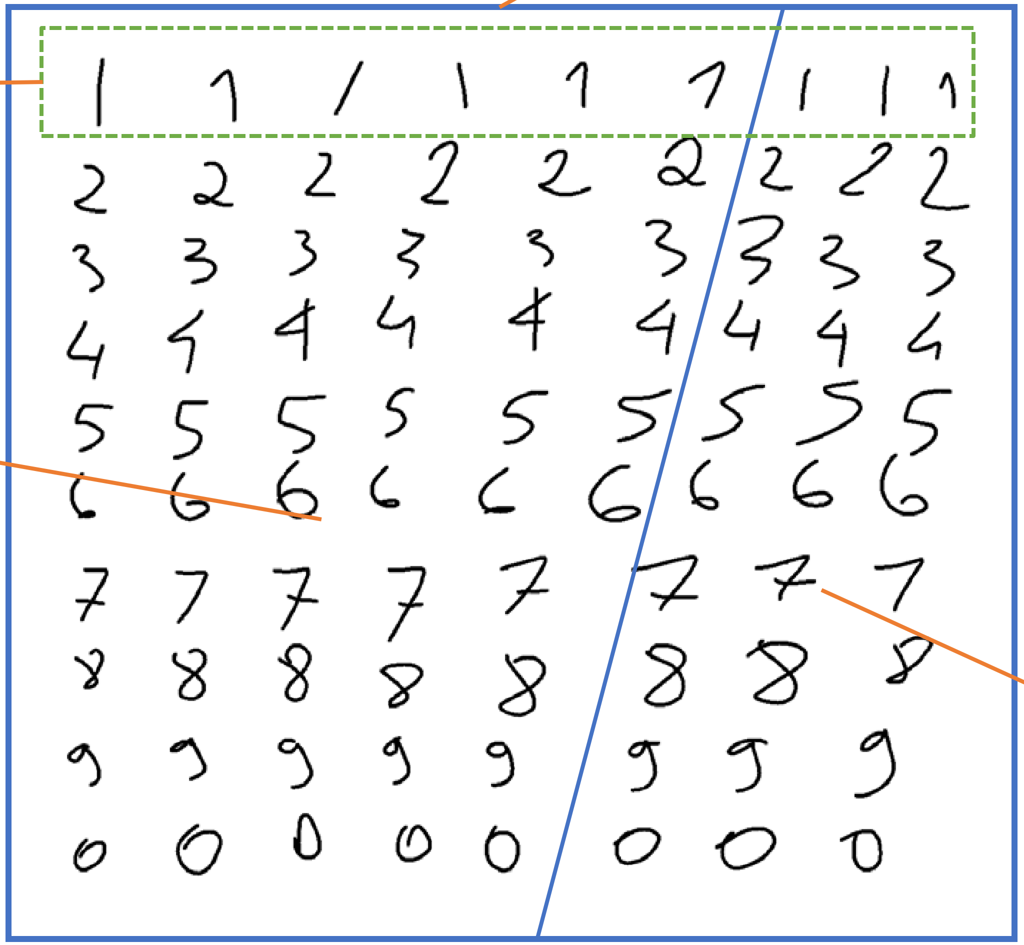
测试数据集
$({x_i}', {y_i}'), i = 1, \ldots, m$

监督学习：训练数据与测试数据

数据样本x

独立同分布
i.i.d

训练集
Training set



1
2
3
4
5
6
7
8
9
0

标注信息y
Labels

测试集
Testing set

监督学习：映射函数f

- 映射函数（模型）f：比如线性模型

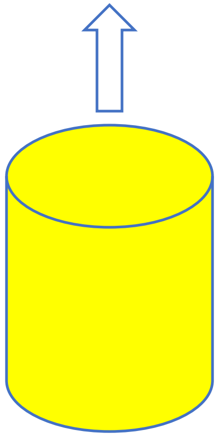
$$\hat{y}_i = f(x_i) = wx_i + b \; (1 \le i \le n)$$

  - x_i：输入样本数据
  - $\hat{y}_i$：模型输出结果
  - w 和 b 为未知参数(需要从数据中学习)
  - 基于领域知识定义模型$f(x_i)$

监督学习：损失函数

训练映射函数f

使得$f(x_i)$预测结果尽量等于y_i



训练数据集

$(\boldsymbol{x_i}, \boldsymbol{y_i}), \boldsymbol{i} = \mathbf{1}, \dots, \boldsymbol{n}$

损失函数名称	损失函数定义
0-1损失函数	$Loss(y_i, f(x_i)) = \begin{cases} 1, f(x_i) \neq y_i \\ 0, f(x_i) = y_i \end{cases}$
平方损失函数	$Loss(y_i, f(x_i)) = (y_i - f(x_i))^2$
绝对损失函数	$Loss(y_i, f(x_i)) = \lvert y_i - f(x_i) \rvert$
对数损失函数/对数似然损失函数	$Loss(y_i, P(y_i \mid x_i)) = -logP((y_i \mid x_i)$

典型的损失函数

监督学习

训练映射函数f

训练数据集
$(x_i, y_i), i = 1, \ldots, n$

- 回归
- 分类
- 识别
- 推荐
- 生成
- …

提纲

一、机器学习基本概念

二、线性回归与线性分类

三、线性判别分析

四、支持向量机

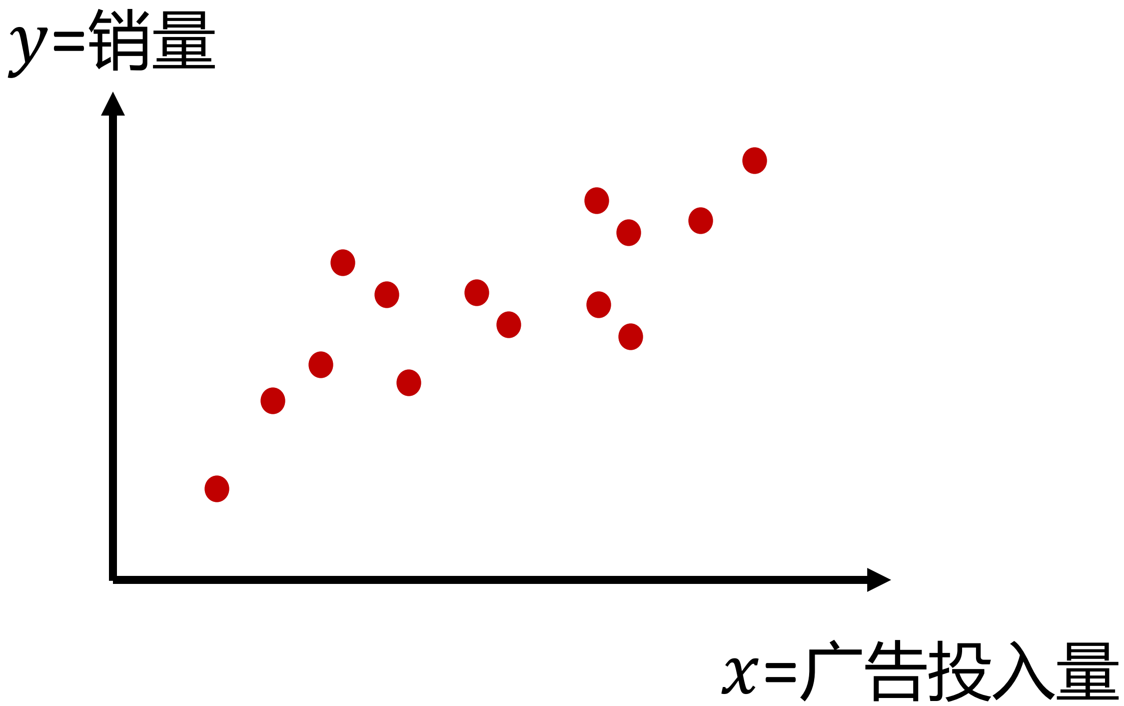
五、决策树

六、集成学习

线性回归
Linear Regression

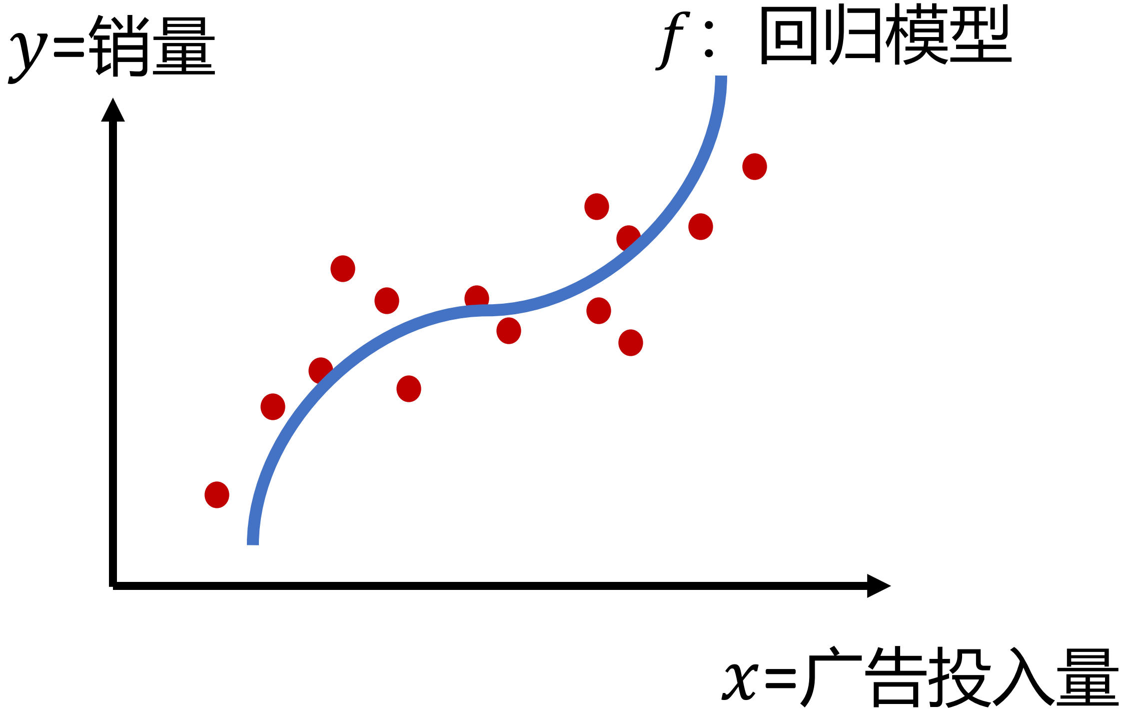
线性回归 (linear regression)

- 在现实生活中，往往需要分析若干变量之间的关系，如碳排放量与气候变暖之间的关系、某一商品广告投入量与该商品销售量之间的关系。



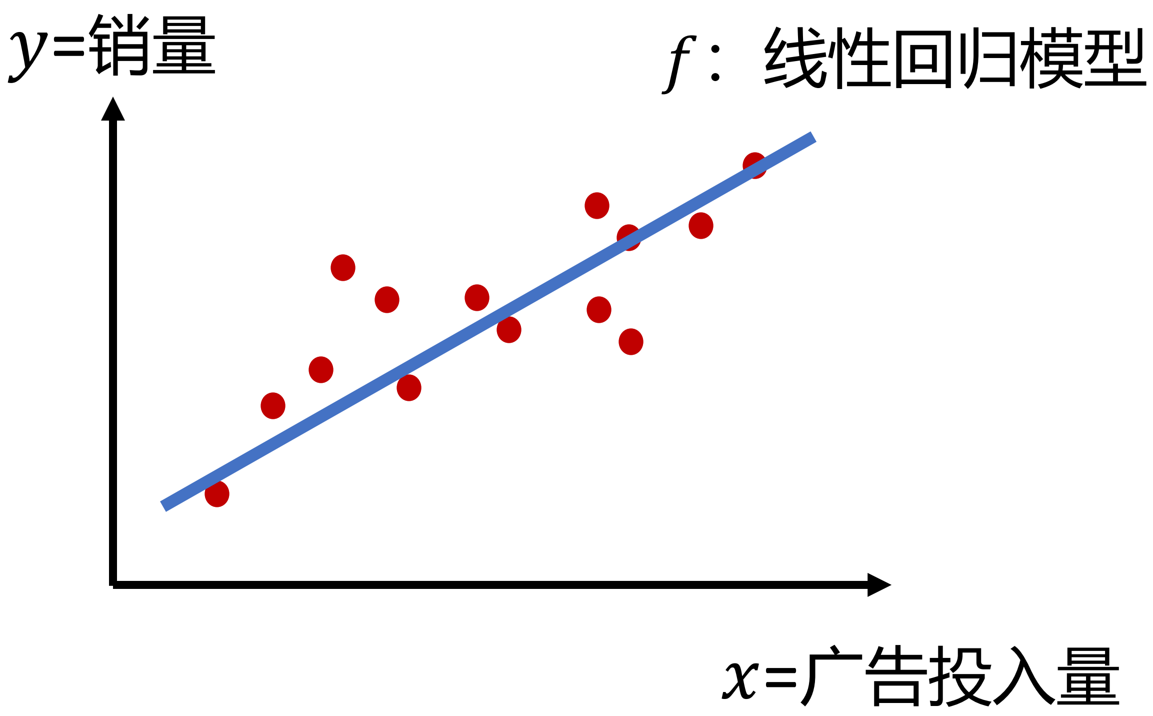
线性回归 (linear regression)

- 这种分析不同变量之间存在关系的研究叫回归分析，刻画不同变量之间关系的模型被称为回归模型。



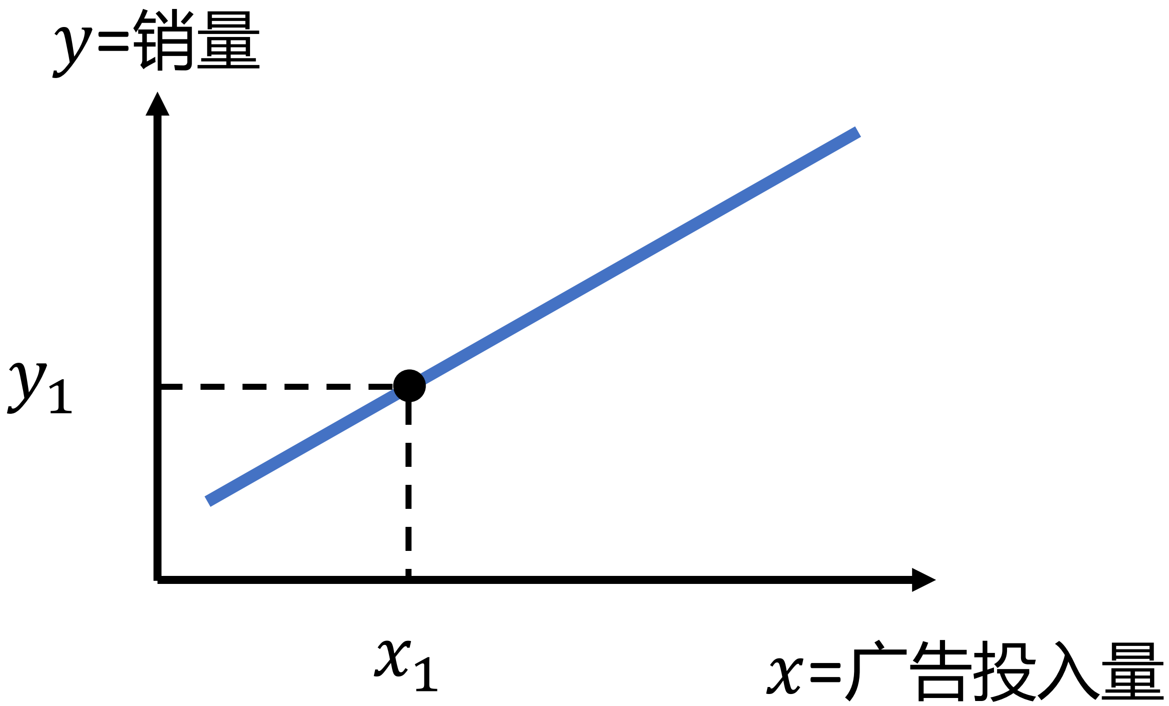
线性回归 (linear regression)

- 这种分析不同变量之间存在关系的研究叫回归分析，刻画不同变量之间关系的模型被称为回归模型。如果这个模型是线性的，则称为线性回归模型。



线性回归 (linear regression)

- 一旦确定了回归模型，就可以进行预测等分析工作，如从碳排放量预测气候变化程度、从广告投入量预测商品销售量等。



线性回归：一元线性回归

- 训练数据：$\mathcal{D} = \{(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)\}$
- 学习模型：$f(x_i) = wx_i + b \ (1 \le i \le n)$
  - 记在当前参数下第i个训练样本x_i的预测值为$f(x_i)$
  - x_i的标注值（实际值）y_i与预测值$f(x_i)$之差记为$(y_i - f(x_i))^2$
- 损失函数：$L(w, b) = \sum_{i=1}^{n} loss(x_i, y_i) = \sum_{i=1}^{n}(y_i - (wx_i + b))^2$
- 优化目标：寻找一组w和b，使得损失函数定义的误差总和$L(w, b)$值最小。
- 优化方法：最小二乘法
  - 对$L(w, b)$ 参数w和b分别求偏导，令其导数值为零，再求取参数w和b的取值。

线性回归：一元线性回归

一元线性回归模型例子

- 下表给出了芒提兹尼欧（Montesinho）地区发生森林火灾的部分历史数据，表中列举了每次发生森林火灾时的气温温度取值x和受到火灾影响的森林面积y。

气温温度	5.1	8.2	11.5	13.9	15.1	16.2	19.6	23.3
火灾影响面积	2.14	4.62	8.24	11.24	13.99	16.33	19.23	28.74

- 可否对气温温度与火灾所影响的森林面积之间关系进行建模呢？
- 初步观察之后，可以使用简单的线性模型构建两者之间关系。
- 即气温温度x与火灾所影响的森林面积y之间存在$y = wx + b$形式的关系。

线性回归：一元线性回归

- 训练数据：$\mathcal{D}=\{(x_1,y_1),(x_2,y_2),\cdots,(x_n,y_n)\}$

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
气温温度x_i	5.1	8.2	11.5	13.9	15.1	16.2	19.6	23.3
火灾影响面积y_i	2.14	4.62	8.24	11.24	13.99	16.33	19.23	28.74
	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8

- 学习模型：$f(x_i)=wx_i+b\ (1\le i\le n)$
- 损失函数：$L(b,w)=\sum_{i=1}^{n} loss(x_i,y_i)=\sum_{i=1}^{8}(y_i-(wx_i+b\,))^2$

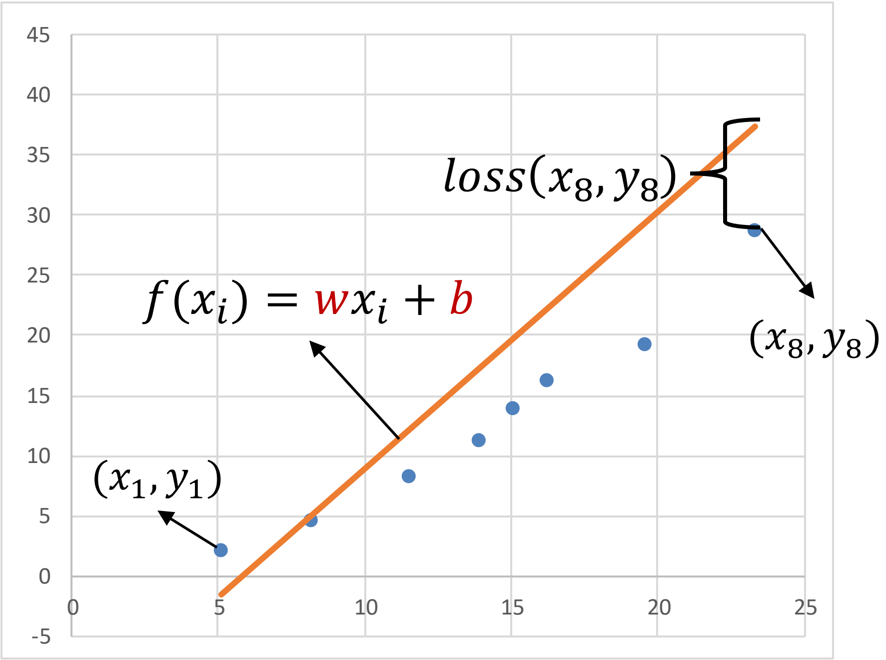
线性回归：一元线性回归

气温温度	5.1	8.2	11.5	13.9	15.1	16.2	19.6	23.3
火灾影响面积	2.14	4.62	8.24	11.24	13.99	16.33	19.23	28.74

- 训练数据：$\mathcal{D}=\{(x_1,y_1),(x_2,y_2),\cdots,(x_n,y_n)\}$
- 学习模型：$f(x_i)=wx_i+b\ (1\le i\le n)$
  - w为直线的斜率；b为直线的截距
- 损失函数：

$$L(w,b)=\sum_{i=1}^{n}loss(x_i,y_i)=\sum_{i=1}^{8}(y_i-(wx_i+b))^2$$

  - 损失函数是关于未知参数w和b的函数。
  - 最优解即是使得损失最小所对应的w和b的值。



气温温度取值和受到火灾影响森林面积之间的一元线性回归模型（橙色实线为回归模型）

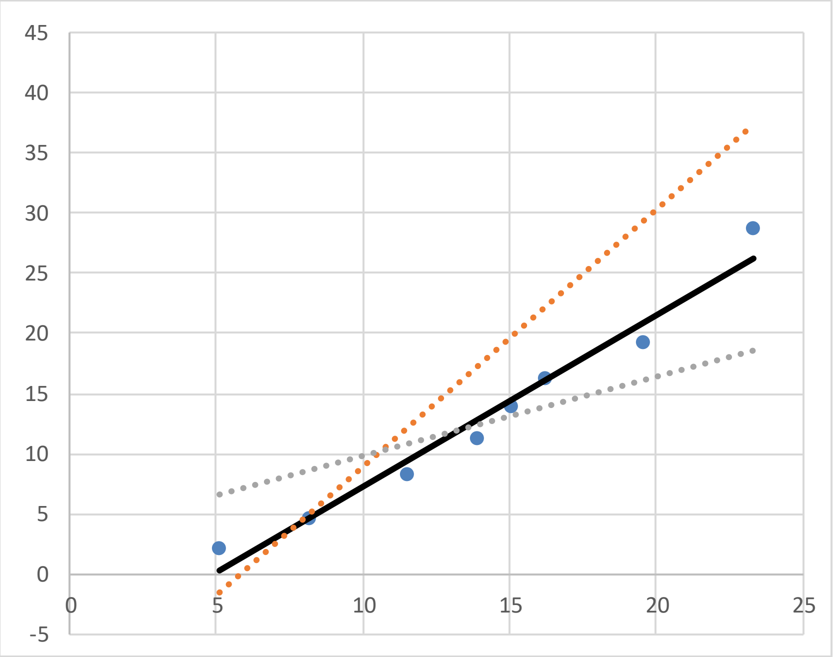
线性回归：一元线性回归

气温温度	5.1	8.2	11.5	13.9	15.1	16.2	19.6	23.3
火灾影响面积	2.14	4.62	8.24	11.24	13.99	16.33	19.23	28.74

- 训练数据：$\mathcal{D}=\{(x_1,y_1),(x_2,y_2),\cdots,(x_n,y_n)\}$
- 学习模型： $f(x_i)=wx_i+b\ (1\le i\le n)$
  - w为直线的斜率；b为直线的截距
- 损失函数：

$$L(w,b)=\sum_{i=1}^{n} loss(x_i,y_i)=\sum_{i=1}^{8}(y_i-(wx_i+b\ ))^2$$

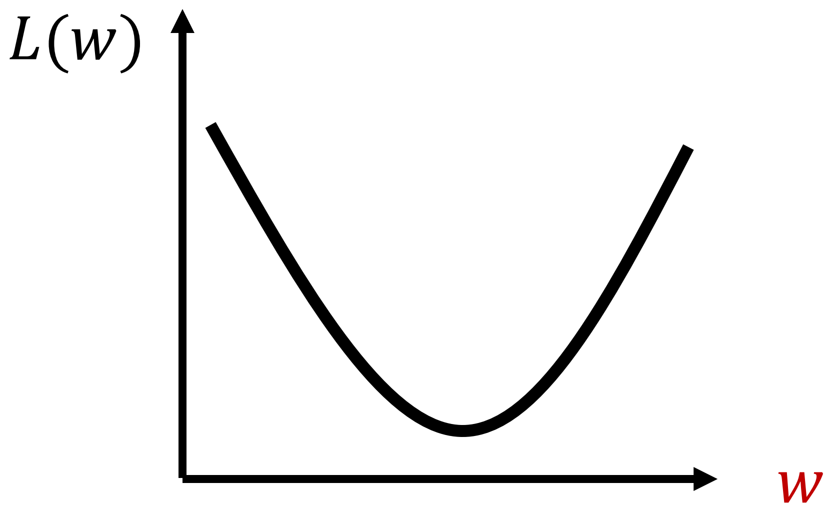
  - 损失函数是关于未知参数w和b的函数。
  - 最优解即是使得损失最小所对应的w和b的值。



气温温度取值和受到火灾影响森林面积之间的一元线性回归模型（黑色实线为最佳回归模型）

线性回归：一元线性回归

- 损失函数： $L(b,w)=\sum_{i=1}^{8}(y_i-(wx_i+b))^2$
- 暂且忽略b： $L(w)=\sum_{i=1}^{8}(y_i-wx_i)^2$



线性回归：一元线性回归

- 损失函数： $L(b,w)=\sum_{i=1}^{8}(y_i-(wx_i+b))^2$
- 优化方法：最小二乘法
  - 对$L(w,b)$ 参数w和b分别求偏导，令其导数值为零，再求取参数w和b的取值。
  - 对b求偏导：

主观题 10分

设置

- 损失函数： $L(b,w)=\sum_{i=1}^{n}(y_i-(wx_i+b))^2$
- 优化方法：最小二乘法
  - 对$L(w,b)$ 参数w和b分别求偏导，令其导数值为零，再求取参数w和b的取值。
  - 对b求偏导：

作答

线性回归：一元线性回归

- 损失函数： $L(b,w)=\sum_{i=1}^{n}(y_i-(wx_i+b))^2$
- 优化方法：最小二乘法
  - 对$L(w,b)$ 参数w和b分别求偏导，令其导数值为零，再求取参数w和b的取值。
  - 对b求偏导：

$$\frac{\partial L(w,b)}{\partial b}=\sum_{i=1}^{n}2(y_i-wx_i-b)(-1)=0$$

$$\to\sum_{i=1}^{n}(y_i-wx_i-b)=0$$

$$\to\sum_{i=1}^{n}y_i-w\sum_{i=1}^{n}x_i-\sum_{i=1}^{n}b=0$$

（令$\bar{y}=\frac{1}{n}\sum_{i=1}^{n}y_i$， $\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$）

$$\to n\bar{y}-nw\bar{x}-nb=0$$

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

主观题 10分

设置

- 损失函数： $L(b,w)=\sum_{i=1}^{n}(y_i-(wx_i+b\ ))^2$
- 优化方法：最小二乘法
  - 对$L(w,b)$ 参数w和b分别求偏导，令其导数值为零，再求取参数w和b的取值。
  - 利用$b=\bar{y}-w\bar{x}$，对w求偏导：

作答

线性回归：一元线性回归

- 对w求偏导：

$$\frac{\partial L(w,b)}{\partial w}=\sum_{i=1}^{n}2(y_i-wx_i-b)(-x_i)=0$$

将 $b=\bar{y}-w\bar{x}$ 代入上式：

$$\rightarrow\sum_{i=1}^{n}(y_i-wx_i-\bar{y}+w\bar{x})(x_i)=0$$

$$\rightarrow\sum_{i=1}^{n}(y_ix_i-wx_ix_i-\bar{y}x_i+w\bar{x}x_i)=0$$

$$\rightarrow\sum_{i=1}^{n}(y_ix_i-\bar{y}x_i)-w\sum_{i=1}^{n}(x_ix_i-\bar{x}x_i)=0$$

$$\rightarrow(\sum_{i=1}^{n}x_iy_i-\bar{x}\bar{y})-w(\sum_{i=1}^{n}x_ix_i-\bar{x}^2)=0$$

$$w=\frac{\sum_{i=1}^{n}x_iy_i-n\bar{x}\,\bar{y}}{\sum_{i=1}^{n}x_ix_i-n\bar{x}^2}$$

线性回归：一元线性回归

- 训练数据：$\mathcal{D}=\{(x_1,y_1),(x_2,y_2),\cdots,(x_n,y_n)\}$

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
气温温度	5.1	8.2	11.5	13.9	15.1	16.2	19.6	23.3
火灾影响面积	2.14	4.62	8.24	11.24	13.99	16.33	19.23	28.74
	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8

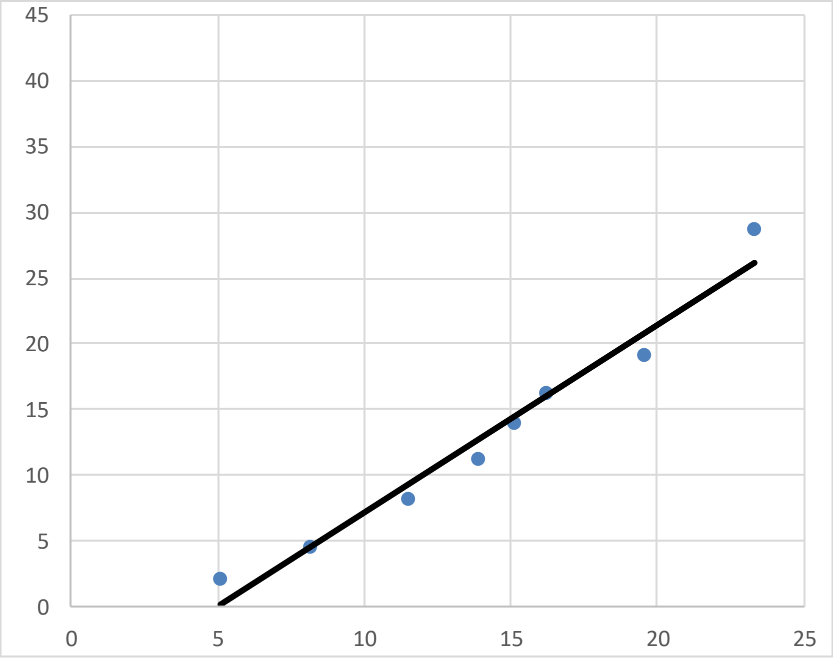
- 优化结果：$b=\bar{y}-w\bar{x}$，$w=\frac{\sum_{i=1}^{n}x_iy_i-\bar{x}\bar{y}}{\sum_{i=1}^{n}x_ix_i-\bar{x}^2}$
- 带入训练数据：$b=\bar{y}-w\bar{x}=-7.09$，$w=\frac{(x_1y_1+x_2y_2+\cdots+x_8y_8)-8\bar{x}\bar{y}}{(x_1^2+x_2^2+\cdots+x_8^2)-8\bar{x}^2}=1.428$
- 预测火灾所影响森林面积与气温温度之间的一元线性回归模型为$y=1.428x-7.09$，即火灾所影响的森林面积 $=1.428\times$ 气温温度 $-\,7.09$

线性回归：一元线性回归

- 训练数据：$\mathcal{D} = \{(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)\}$

气温温度	5.1	8.2	11.5	13.9	15.1	16.2	19.6	23.3
火灾影响面积	2.14	4.62	8.24	11.24	13.99	16.33	19.23	28.74

$$\text{火灾所影响的森林面积} = 1.428 \times \text{气温温度} - 7.09$$



线性回归：多元线性回归

- 训练数据：$\mathcal{D}=\{(\boldsymbol{x}_1,y_1),(\boldsymbol{x}_2,y_2),\cdots,(\boldsymbol{x}_n,y_n)\}$
  - 其中，$\boldsymbol{x_i}=[x_{i,1},x_{i,2},\dots,x_{i,D}]\in\mathbb{R}^D$为代表多元特征的第$i$个数据向量
  - D为数据向量的维度，也就是特征个数
- 学习模型：$f(\boldsymbol{x_i})=\sum_{j=1}^{D}w_j\,x_{i,j}+w_0=\boldsymbol{w^T}\boldsymbol{x_i}+w_0,(1\le i\le n)$
- 损失函数：$L(\boldsymbol{w})=\sum_{i=1}^{n}loss(\boldsymbol{x}_i,y_i)=\sum_{i=1}^{n}(y_i-(\boldsymbol{w^T}\boldsymbol{x_i}+w_0))^2$
- 优化方法：最小二乘法

线性回归：多元线性回归

多元线性回归模型例子

- 扩展到数据特征的维度是多维的情况，在数据中增加一个影响火灾影响面积的因素—风力。

气温	5.1	8.2	11.5	13.9	15.1	16.2	19.6	23.3
风力	4.5	5.8	4	6.3	4	7.2	6.3	8.5
火灾影响面积	2.14	4.62	8.24	11.24	13.99	16.33	19.23	28.74

- 可利用多元线性回归来建模

填空题 2分

设置

该例子中，训练数据样本个数为 [填空1] ，数据特征维度为 [填空2]

作答

线性回归：多元线性回归

- 训练数据：$\mathcal{D} = \{(\boldsymbol{x}_1, y_1), (\boldsymbol{x}_2, y_2), \cdots, (\boldsymbol{x}_n, y_n)\}$
  - 其中， $\boldsymbol{x_i} = [x_{i,1}, x_{i,2}, \dots, x_{i,D}] \in \mathbb{R}^D$为代表多元特征的第$i$个数据向量
  - D为数据向量的维度，也就是特征个数

气温$x_{i,1}$	5.1	8.2	11.5	13.9	15.1	16.2	19.6	23.3
风力$x_{i,2}$	4.5	5.8	4	6.3	4	7.2	6.3	8.5
火灾影响面积y_i	2.14	4.62	8.24	11.24	13.99	16.33	19.23	28.74

- 学习模型：$f(\boldsymbol{x_i}) = \sum_{j=1}^{D} w_j\, x_{i,j} + w_0 = \boldsymbol{w}^{\boldsymbol{T}} \boldsymbol{x_i} + w_0, (1 \le i \le n)$
- 损失函数： $L(\boldsymbol{w}) = \sum_{i=1}^{n} loss(\boldsymbol{x}_i, y_i) = \sum_{i=1}^{n} (y_i - (\boldsymbol{w}^{\boldsymbol{T}} \boldsymbol{x_i} + w_0))^2$
- 优化方法：最小二乘法

填空题 1分

设置

请问多元线性回归中，假设数据向量的维度为D，那么多元线性回归模型的参数量为 [填空1]

作答

线性回归：多元线性回归

- 损失函数：最小化均方误差函数

$$L(\boldsymbol{w}) = \sum_{i=1}^{n} \left(y_i - (\boldsymbol{w}^T \boldsymbol{x}_i + w_0)\right)^2$$

- 为了方便，使用矩阵来表示所有的训练数据和数据标签。

$$X = [\boldsymbol{x}_1, \dots, \boldsymbol{x}_n]^T, \qquad \boldsymbol{y} = [y_1, \dots, y_n]^T$$

- 其中每一个数据$\boldsymbol{x}_i$会扩展一个维度，其值为1，对应参数w_0。

$$X = \begin{bmatrix} \boldsymbol{x}_1, \dots, \boldsymbol{x}_n \\ 1, \dots, 1 \end{bmatrix}^T$$

- 均方误差函数可以表示为：

$$L(\boldsymbol{w}) = (\boldsymbol{y} - X\boldsymbol{w})^T (\boldsymbol{y} - X\boldsymbol{w})$$

线性回归：多元线性回归

- 均方误差函数可以表示为：

$$L(\boldsymbol{w}) = (\boldsymbol{y} - X\boldsymbol{w})^T(\boldsymbol{y} - X\boldsymbol{w})$$

- 均方误差函数$L(\boldsymbol{w})$对所有参数$\boldsymbol{w}$求导可得：

$$\nabla L(\boldsymbol{w}) = -2X^T(\boldsymbol{y} - X\boldsymbol{w})$$

- 因为均方误差函数$L(\boldsymbol{w})$是一个二次的凸函数，所以函数只存在一个极小值点，也同样是最小值点，所以令$\nabla L(\boldsymbol{w}) = 0$可得

$$X^T X\boldsymbol{w} = X^T\boldsymbol{y}$$

$$\boldsymbol{w} = (X^T X)^{-1} X^T\boldsymbol{y}$$

线性回归：多元线性回归

- 对于上面的例子，转化为矩阵的表示形式为：

$$
\begin{array}{cc}
 & \begin{array}{cccccccc} \boldsymbol{x_1} & \boldsymbol{x_2} & \cdots & & & & & \end{array} \\
X = & \begin{bmatrix} 5.1 & 8.2 & 11.5 & 13.9 & 15.1 & 16.2 & 19.6 & 23.3 \\ 4.5 & 5.8 & 4. & 6.3 & 4. & 7.2 & 6.3 & 8.5 \\ 1. & 1. & 1. & 1. & 1. & 1. & 1. & 1. \end{bmatrix}^T \begin{array}{c} x_{i,1} \\ x_{i,2} \\ \end{array}
\end{array}
$$

$$\boldsymbol{y} = [2.14 \quad 4.62 \quad 8.24 \quad 11.24 \quad 13.99 \quad 16.33 \quad 19.23 \quad 28.74]^{\mathrm{T}}$$

- 其中矩阵X多出一行全1，是因为常数项w_0，可以看作是数值为全1的特征的对应系数。计算可得

$$\boldsymbol{w} = [1.312 \quad 0.626 \quad -9.103]^T$$

$$f(x_i) = 1.312x_{i,1} + 0.626x_{i,2} - 9.103$$