ChatGPT

记忆已满

共享

LU

背景

在简单的线性回归模型中，预测值为：

$$\hat{y} = Wx + b$$

$$\hat{y} = Wx + b$$

目标是最小化均方误差 (MSE) 损失函数：

$$L = \frac{1}{n}\sum_{i=1}^{n}(y_i - (Wx_i + b))^2$$

$$L = \frac{1}{n}\sum_{i=1}^{n}(y_i - (Wx_i + b))^2$$

其中：

n

n: 数据点数量

x_i, y_i

x_i, y_i：第

i

i 个样本的输入和真实输出

W, b

W,b: 待优化的模型参数

为了找到最优的

W

W 和

b

b，需要对损失函数

L 对 W 和 b 求偏导，使偏导数等于 0。

推导过程

展开损失函数：

$$L = \frac{1}{n}\sum_{i=1}^{n}(y_i - Wx_i - b)^2$$

对 W 求偏导：

$$\frac{\partial L}{\partial W} = \frac{1}{n}\sum_{i=1}^{n} 2(y_i - Wx_i - b)(-x_i)$$

$$\sum_{i=1}^{n} 2(y_i - Wx_i - b)(-x_i)$$

简化后：

$$\frac{\partial L}{\partial W} = -\frac{2}{n}\sum_{i=1}^{n} x_i (y_i - Wx_i - b)$$

$$\frac{\partial L}{\partial W} = -\frac{2}{n}\sum_{i=1}^{n} x_i (y_i - Wx_i - b)$$

整理得：

$$\frac{\partial L}{\partial W} = -\frac{2}{n}\left(\sum_{i=1}^{n} x_i y_i - W \sum_{i}\right.$$

=
1
n
x
i
2
−
b
∑
i
=
1
n
x
i
)
∂W
∂L

=−
n
2

(
i=1
∑
n

x
i

y
i

−W
i=1
∑
n

x
i
2

−b
i=1
∑
n

x
i

)
对
b
b 求偏导：
∂
L
∂
b
=
1
n
∑
i
=
1
n
2
(
y
i
−
W
x
i
−
b
)
(
−
1
)
∂b
∂L

$$= \frac{1}{n} \sum_{i=1}^{n} 2(y_i - Wx_i - b)(-1)$$

简化后：

$$\frac{\partial L}{\partial b} = -\frac{2}{n} \sum_{i=1}^{n} (y_i - Wx_i - b)$$

$$\frac{\partial L}{\partial b} = -\frac{2}{n} \sum_{i=1}^{n} (y_i - Wx_i - b)$$

整理得：

$$\frac{\partial L}{\partial b} = -\frac{2}{n} \left(\sum_{i=1}^{n} y_i - W \sum_{i=1}^{n} x_i \right.$$

−
n
b
)
∂b
∂L

=−
n
2

(
i=1
∑
n

y
i

−W
i=1
∑
n

x
i

−nb)
偏导数等于 0 的方程
令
∂
L
∂
W
=
0
∂W
∂L

=0 和
∂
L
∂
b
=
0
∂b
∂L

=0，可以得到以下方程组：

对
W
W 的方程：
∑
i
=
1
n
x
i
y
i
=
W
∑
i
=
1
n
x
i
2
+
b
∑
i
=
1
n
x
i
i=1
∑
n

$$x_i y_i = W\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i$$

对 b 的方程：

$$\sum_{i=1}^{n} y_i = W\sum_{i=1}^{n} x_i + nb$$

这两个方程对应于你给出的具体数字：

$$60W + 20b - 34.4 = 0$$

$$20W + 8b - 12.2 = 0$$

其中，系数和常数项来源于你的数据点的计算，例如：

∑
x
i
,
∑
y
i
,
∑
x
i
2
,
∑
x
i
y
i
∑x
i

,∑y
i

,∑x
i
2

,∑x
i

y
i

总结
这些偏导的形式来自于损失函数的梯度公式，方程系数与数据点的统计量直接相关。通过解这组线性方程，能够找到使损失函数最小的
W
W 和
b
b。

在上述推导中，我们使用了线性回归的核心数学工具——最小二乘法 (Ordinary Least Squares, OLS)，通过对均方误差 (MSE) 损失函数求偏导并令其为零，得到最优参数 W 和 b 的解析解。

以下是完整的计算过程和对关键点的详细说明：

---

1. 损失函数定义

目标是最小化均方误差：

$$L = \frac{1}{n} \sum_{i=1}^{n} (y_i - (Wx_i + b))^2$$

其中：

- n: 数据点数量
- x_i, y_i: 第 i 个样本的输入和真实输出
- W, b: 待优化的模型参数

---

2. 对参数 W 求偏导

展开 L：

L = \frac{1}{n} \sum_{i=1

把我上面的重新写出来就可以

简单线性回归模型

模型定义

在简单的线性回归模型中，预测值为：

$$\hat{y} = Wx + b$$

目标是最小化均方误差 (MSE) 损失函数：

$$L = \frac{1}{n}\sum_{i=1}^{n}(y_i - (Wx_i + b))^2$$

其中：

- n：数据点数量
- x_i, y_i：第 i 个样本的输入和真实输出
- W, b：待优化的模型参数

为了找到最优的 W 和 b，需要对损失函数 L 分别对 W 和 b 求偏导，并令偏导数等于 0。

---

1. 展开损失函数

$$L = \frac{1}{n}\sum_{i=1}^{n}(y_i - Wx_i - b)^2$$

---

2. 对 W 求偏导

$$\frac{\partial L}{\partial W} = \frac{1}{n}\sum_{i=1}^{n}2(y_i - Wx_i - b)(-x_i)$$

简化后：

$$\frac{\partial L}{\partial W} = -\frac{2}{n}\sum_{i=1}^{n}x_i(y_i - Wx_i - b)$$

进一步整理为：

$$\frac{\partial L}{\partial W} = -\frac{2}{n}\left(\sum_{i=1}^{n}x_iy_i - W\sum_{i=1}^{n}x_i^2 - b\sum_{i=1}^{n}x_i\right)$$

---

3. 对 b 求偏导

$$\frac{\partial L}{\partial b} = \frac{1}{n}\sum_{i=1}^{n}2(y_i - Wx_i - b)(-1)$$

简化后：

$$\frac{\partial L}{\partial b} = -\frac{2}{n}\sum_{i=1}^{n}(y_i - Wx_i - b)$$

进一步整理为：

$$\frac{\partial L}{\partial b} = -\frac{2}{n}\left(\sum_{i=1}^{n}y_i - W\sum_{i=1}^{n}x_i - nb\right)$$

---

4. 偏导数等于 0 的方程组

令：

$$\frac{\partial L}{\partial W} = 0$$
$$\frac{\partial L}{\partial b} = 0$$

得到以下方程组：

1. 关于 W 的方程：

$$\sum_{i=1}^{n}x_iy_i = W\sum_{i=1}^{n}x_i^2 + b\sum_{i=1}^{n}x_i$$

2. 关于 b 的方程：

$$\sum_{i=1}^{n}y_i = W\sum_{i=1}^{n}x_i + nb$$

---

5. 具体方程

根据给出的数据点计算得到以下方程：

1. 关于 W 和 b 的方程：

$$60W + 20b - 34.4 = 0$$

2. 关于 b 的方程：

$$20W + 8b - 12.2 = 0$$

这些方程的系数来源于数据点的统计量，例如：

- $\sum x_i$
- $\sum y_i$
- $\sum x_i^2$
- $\sum x_i y_i$

总结

通过解这组线性方程，可以找到使损失函数 L 最小的参数 W 和 b。这些推导基于均方误差的梯度公式，并且系数与数据点的统计量直接相关。