Table Of Contents:

今日工作

1、阅读《计算机视觉 模型、学习和推理》第八章1、2节——线性回归和贝叶斯线性回归。

2、加深对后验分布与先验分布的理解。

3、编程实现书中的线性回归拟合算法。

线性回归

待求的模型：

$$Pr(w_i|\mathbf{x}_i,\theta) = \mathrm{Norm}_{w_i}[\phi_0 + \boldsymbol{\phi}^T X_i, \sigma^2]$$

X 是数据集合，维度为N * D，N是数据的个数，D是数据*x* 的维度。

$$\boldsymbol{\phi} = [\phi_1, \phi_2, \cdots, \phi_D]^T$$ 通过矩阵乘法，*Φ_i* 可以理解为第*i* 维的的梯度。

通过变换：

$$\mathbf{x}_i \leftarrow (1 \quad \mathbf{x}_i^T)^T$$ 和

原分布可以写成:

例如1维的*X* :

变成：(50, 2)

Φ 为：(50, 1)

1	5.9818
2	6.4648
3	5.6083
4	5.5870
5	6.1694
6	7.0114
7	5.5484
8	6.1514
9	5.7116
10	6.4357
11	5.0302

参数：

对数化不改变极值点位置。由于函数是线性的，可以通过求偏导求得极值点位置:

$$\hat{\phi} = (XX^{T})^{-1}Xw$$

$$\hat{\sigma}^2 = \frac{(w - X^{T}\phi)^{T}(w - X^{T}\phi)}{I}$$

贝叶斯线性回归

这个方法中引入了Φ 的先验$Pr(\Phi)$ 。

Φ 是多元连续的。

将$Pr(\Phi)$ 建模为均值为0，协方差为球形的多元正态分布。

根据训练样本，Φ 的后验分布为：

似然函数为前一节线性回归 的：

根据正态分布的运算特性，后验分布可以化为以下式子：

其中：

最终，我们对ω 的预测为:

后验与先验

例子：

一个班有50个人，其中49个男的，1个女的。

在公园碰到一个同学，这个同学性别是男的概率为49/50。

如果知道这个同学穿裙子，这个同学性别是男的概率是 p(穿裙子|男生)* 49 / 50

这个同学是男生的**先验概率**是基于客观情况的估计即 P(男生)=49/50

这个同学是男生的**后验概率**是 P(男生|穿裙子) = p(穿裙子|男生)* 49 / 50，即观测到一定条件后之后，一个事件的概率，同时这个条件是否能够发生本身是有一个概率。

在先验分布$Pr(\Phi)$ 中，Φ 的分布本来是非常广的：

但是由于存在一个观测到的事实：x ，那么Φ 就会变窄，因为在Φ 这个区域内时，我们观测到x 的概率是更高的。

线性回归拟合算法

这个算法比较简单，关键步骤是：

1、根据书中的偏导公式，可以直接计算极值点处的参数。

```
def fit_lr(X, w):
    phi = np.dot(np.dot(np.linalg.inv(np.dot(X, X.T)), X), w)
    # 正态分布均值 : phi[1, :]
    batch_size = X.shape[1]
    temp = w - np.dot(X.T, phi)
    sigma = np.dot(temp.T, temp) / batch_size
    return phi, sigma
```

2、建立一个自定义的状态ω 与x 的线性关系。

```
phi = np.array([[7], [-0.5]])
X_phi = np.dot(X, phi)
r = np.random.randn(batch_size, 1)
for i in range(batch_size):
    mu = X_phi[i]
    w[i] = mu + r[i]
```

第一维梯度是-0.5，7是*Φ_0*，r[i]为干扰信息。

3、最后是验证，验证思路与之前的算法相同，从区域内均匀提取大量的点，计算这些点的概率密度。

结果分析：

这是概率密度用热力图可视化以及数据的散点图可视化结果。

可以看出，数据集中的地方概率密度值是更大的。