

ML 课程第一次作业：衡量线性回归的指标

原理与推导

- 一般地，我们对于简单线性回归的要求是：找到一组 a, b 使得

$$\sum_{i=1}^m \left(y^{(i)} - ax^{(i)} - b \right)^2$$

尽可能小。也就是训练集尽可能做到 $\sum_{i=1}^m \left(y_{train}^{(i)} - \hat{y}_{train}^{(i)} \right)^2$ 的条件下，测试集也亦然，做到

$\sum_{i=1}^m \left(y_{test}^{(i)} - \hat{y}_{test}^{(i)} \right)^2$ 尽量小。即训练集**拟合效果好**，同时尽量避免测试集**过拟合**。首先做第一层

改进，引入 MSE，使 loss 值与样本大小无关，便于不同大小测试集效果横向比较。

1. 均方误差 MSE

MSE 全称 Mean Squared Error，当测试集大小为 m 时，计算方法为

$$\frac{1}{m} \sum_{i=1}^m \left(y_{test}^{(i)} - \hat{y}_{test}^{(i)} \right)^2$$

MSE 优点在于**其值均正且可导**，便于计算梯度后做各种 gradient descent；也有其缺点，就是 loss 的量纲与测试集不一致。可以采取 RMSE 的方式使量纲一致，提高可解释性。

2. 均方根误差 RMSE

RMSE 全称 Root Mean Squared Error，计算方法为

$$\sqrt{\frac{1}{m} \sum_{i=1}^m \left(y_{test}^{(i)} - \hat{y}_{test}^{(i)} \right)^2}$$

也就是 $\sqrt{MSE_{test}}$ ，RMSE 是使用最广泛的一种 loss 评估方式。

3. 平均绝对误差 MAE

MAE 全称 Mean Absolute Error，同样也是解决量纲不一致的一种途径。与 MSE 相比在绝对值内正转负的零点不可导，不便后续用 gradient descent 求极值来优化函数。当然作为一种模型评估的手段是完全可以的，计算量很小。其计算方法为

$$\frac{1}{m} \sum_{i=1}^m \left| y_{test}^{(i)} - \hat{y}_{test}^{(i)} \right|$$

与 RMSE 相比，RMSE 扩大了较大误差，对误差更敏感，更准确，且可以用于函数优化；而 MAE 计算简便。

4. R方 R Squared

R Squared 用于提供一个模型评估的上下限，直观了解模型所在的质量定位。其计算方法为

$$R^2 = 1 - \frac{SS_{residual}}{SS_{total}} = 1 - \frac{\sum_i \left(\hat{y}^{(i)} - y^{(i)} \right)^2}{\sum_i \left(\bar{y}^{(i)} - y^{(i)} \right)^2}$$

将 loss 归约到了 $0 \sim 1$ 之间。从其数学意义上看，可以看出减去的部分分子即 MSE，分母即方差

$$R^2 = 1 - \frac{\sum_i \left(\hat{y}^{(i)} - y^{(i)} \right)^2}{\sum_i \left(\bar{y} - y^{(i)} \right)^2} = 1 - \frac{\frac{1}{m} \left(\sum_{i=1}^m \left(\hat{y}^{(i)} - y^{(i)} \right)^2 \right)}{\frac{1}{m} \left(\sum_{i=1}^m \left(\bar{y} - y^{(i)} \right)^2 \right)} = 1 - \frac{MSE(\hat{y}, y)}{Var(y)}$$

代码实现

Python 按照公式构造函数非常简单，此不赘言，只介绍 sklearn 库中的包用法：

```
from sklearn import metrics

# MSE
MSE = metrics.mean_squared_error(y_test, y_predict)

# RMSE
RMSE = np.sqrt(metrics.mean_squared_error(y_test, y_predict))

# MAE
MAE = metrics.mean_absolute_error(y_test, y_predict)

# R Squared
R2 = metrics.r2_score(y_test, y_predict)
```

by 徐思喆