



链滴

深入理解 L1,L2 正则化

作者: [zeekling](#)

原文链接: <https://ld246.com/article/1569160005769>

来源网站: 链滴

许可协议: [署名-相同方式共享 4.0 国际 \(CC BY-SA 4.0\)](#)

概念

正则化(Regularization) 是机器学习中对原始损失函数引入额外信息，以防止过拟合和提高模型泛化性能的一类方法的统称。正则化是在经验风险上面加了一个正则化项或者罚项,正则化函数一般是模型法则度的单调增函数,模型越负责,正则化值就越大.

正则化的一般形式:

$$\min_{f \in F} \frac{1}{N} \sum_{i=1}^N L(y_i, f(x_i)) + \lambda J(f)$$

第一项是经验风险,第二项就是正则化项, $\lambda \geq 0$ 为调整两者之间的关系.

L1 正则化和 L2 正则化可以看做是**损失函数的惩罚项**。所谓**惩罚**是指对损失函数中的**某些参数做一些限制**。对于线性回归模型, 使用 L1 正则化的模型叫做 Lasso 回归, 使用 L2 正则化的模型叫做 Ridge 回归 (岭回归)。

线性回归 L1 正则化损失函数:

$$\min_w [\sum_{i=1}^N (w^T x_i - y_i)^2 + \lambda \sum |w_j|] \dots\dots 1)$$

线性回归 L2 正则化损失函数:

$$\min_w [\sum_{i=1}^N (w^T x_i - y_i)^2 + \lambda \sum w_j^2] \dots\dots (2)$$

可以看到正则化项是对系数做了限制。L1 正则化和 L2 正则化的说明如下:

- L1 正则化是指权值向量 w 中各个元素的绝对值之和通常表示为 $\|w\|_1$

- L2 正则化是指权值向量 w 中各个元素的平方和然后求平方根 (可以看到 Ridge 回归的 L2 正则化项有平方符号), 通常表示为 $\|w\|_2$ 。

- 一般都会在正则化项之前添加一个系数 λ 。

未完待续....