

广义线性模型：逻辑回归

学习目标：

- 逻辑回归损失函数推导
- 了解模型欠拟合、过拟合状态，了解L1、L2正则化方法减缓模型过拟合。

知识点

1. [逻辑回归理解（文章第七、第八部分理解即可）](#)

勘误：文章损失函数部分应添加负号，对齐梯度下降方向

2. [无偏估计](#)

理解衡量一个估计好坏的三个方向：无偏（低偏差）、有效（低方差）、一致（数据划分）

3. [\[推荐版本\]方差与偏差的权衡](#)，[\[辅助版本\]方差与偏差的权衡](#)

4. [控制方差与偏差权衡的方法：L1与L2正则化\(1-7部分\)](#)

[稀疏性的图解法](#)

[稀疏性的其他解释](#)

QA

思考、查阅资料，尝试回答：

1. 为什么不能直接用线性回归做分类？
 - 部分理解可参考[为什么不用线性回归解决分类问题](#)；
 - 尝试从损失函数梯度的角度理解；
2. 多重共线性变量会给逻辑回归带来什么问题？为什么？怎么处理？
3. L1与L2正则化会给模型带来什么影响，为什么？
4. 逻辑回归本身只具备线性的表达能力，如何让模型学会非线性关系？

项目

[Titanic: Machine Learning from Disaster](#)

尝试用逻辑回归提交预测结果，代码可参考[Titanic: logistic regression with python](#)。

Refs

- [WIKI](#)
- [逻辑回归 logistics regression 公式推导](#)
- [马同学博客](#)