

机器学习理论研究导引

作业五

你的姓名 你的学号

2022 年 5 月 16 日

作业提交注意事项

- (1) 本次作业提交截止时间为 **2022/05/31 23:59:59**, 截止时间后不再接收作业, 本次作业记零分;
- (2) 作业提交方式: 使用此 LaTeX 模板书写解答, 只需提交编译生成的 pdf 文件, 将 pdf 文件上传到以下ftp服务器的指定位置:
地址: sftp://210.28.132.67:22, 用户名: mlt2022, 密码: mltspring2022@nju
文件夹位置: /C:/Users/mlt2022/hw_submissions/hw5_submission/ ;
- (3) pdf 文件命名方式: 学号-姓名-作业号-v版本号, 例 MG1900000-张三-5-v1; 如果需要更改已提交的解答, 请在截止时间之前提交新版本的解答, 并将版本号加一;
- (4) 未按照要求提交作业, 或 pdf 命名方式不正确, 将会被扣除部分作业分数.

1 [50pts] Consistent Surrogate Loss

考虑对率函数 $\phi(t) = \log(1 + e^{-t})$, 回答并证明下述问题.

1. [15pts] 试求解最优实值输出函数 $f_{\phi}^*(\mathbf{x})$.
2. [10pts] 试求解最优实值输出函数对应的最优替代泛化风险 R_{ϕ}^* .
3. [25pts] 证明对率函数针对原0/1目标函数具有替代一致性.

Proof.

2 [50pts] Convergence Rate

试分析下述情形下梯度下降算法 (课件第9页, 课本7.2.1节) 的收敛率.

- 可行域 \mathcal{W} 是凸集.
- 目标函数 f 是 \mathcal{W} 上可微的 α -强凸函数.
- 目标函数在 \mathcal{W} 上是 l -Lipschitz连续的, 即

$$\forall \mathbf{u} \in \mathcal{W}, \|\nabla f(\mathbf{u})\| \leq l. \quad (2.1)$$

- 梯度下降算法的步长为:

$$\forall t, \eta_t = \frac{1}{\alpha t}. \quad (2.2)$$

Proof.