

深度强化学习课后练习

Day 1

主要内容：线性代数、微积分、概率论、机器学习概念、Python

练习题：

1. 了解数据的经验分布，推导经验分布的期望、方差与数据的均值、方差之间的关系
2. 搜集资料学习 `numpy`，完成以下任务（1）对称矩阵正交分解（2）计算矩阵特征值（3）绘制正态分布概率密度函数图（4）从给定协方差矩阵的多元正态分布中采样
3. 了解 ROC 曲线，以两个不同均值一元正态分布为例，绘制分类器的 ROC 曲线并计算 AUC
4. 编写求线性回归的最小二乘法的程序，在给定数据上和自生成数据上测试
- *5. 了解大数定律及在统计学上的意义
6. 推导 n 个 iid 随机变量均值的期望和方差

Day 2

主要内容：优化方法、信息论、监督学习方法

练习题：

1. 编写最小二乘的梯度下降法，绘制误差、梯度范数 vs 迭代步数的曲线
2. 证明 KL 散度是凸函数
3. 将多项式拟合问题转化为线性回归问题
4. 熟悉 `scikit-learn`，分别用无惩罚项、L1 范数惩罚、L2 范数惩罚解决线性回归问题，比较最优解的范数大小，以及泛化误差
- *5. 了解 SVM 的优化方法，及拉格朗日乘子法

Day 3

主要内容：神经网络、CNN、TensorFlow

练习题：

1. 独立推导反向传播算法与 Delta 规则
2. 使用 `tf` 搭建 LeNet，使用 MNIST 数据集进行分类

Day 4

主要内容：RL 基本设定、MDP

练习题：

1. 关于 Multi-armed Bandit 问题的各种算法实现
2. 推导 action-value function 的 Bellman 方程和 Bellman 最优性方程
3. 赌博游戏值迭代法实现

Day 5

主要内容：Monte Carlo 方法、TD 方法以及 n-step 方法

练习题：

1. 用蒙特卡洛方法求复杂函数的积分
2. 推导式 14、15 的更新规则
3. 推导 Sarsa 算法的 TD error 与 MC 误差的关系

4. Double Q-learning 是否能推广到 Triple Q-learning 甚至更多？效果会不会更好？试编程回答

Day 6

主要内容：Dyna-Q、参数方法、PG

*1. 推导 PG 定理