【算法】Prediction1

张楚珩
清华大学 交叉信息院博士在读

之前导师就推荐了 *Prediction, learning, and games* 这本书，最近的项目可能涉及相关的东西，所以来看一下。这一篇讲第一章：Introduction。

原文传送门

Cesa-Bianchi, Nicolo, and Gabor Lugosi.*Prediction, learning, and games*. Cambridge university press, 2006.

由于是图书，就不放链接了，直接 google 就能搜得到 PDF。

特色

第一章分别讲了 prediction、learning、games 之间的联系，然后举了两个有趣的例子，在这里记录一下。这本书里面研究的 expert problem、multi-armed bandit problem 等问题在强化学习和金融中都有广泛的应用和启发。

考虑项目的相关性，我只读右半子树对应的章节。

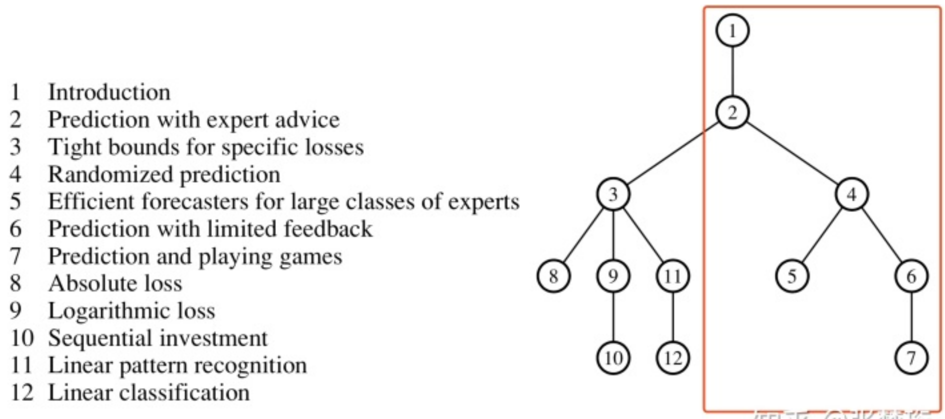

Figure 1.1. The dependence structure of the chapters.

过程

1. Prediction, learning and games

这本书里面讲的 prediction 主要是序列的预测，大致 setting 如下：考虑预测一个时间离散的序列，每个时刻 t，有若干个专家（expert）来给出对于该时刻可能出现状态的预测，玩家结合这些『建议』来决定一个预测，然后该时刻的实际序列被揭晓，根据预测的准确性，玩家会遭受一个损失。预测的目标就是使得玩家积累起来遭受的损失最小。具体地，希望最小化 regret，regret 定义为玩家遭受的损失和某一个专家（基准）遭受损失的差值。

Prediction 问题通常不对序列的分布或者变化规律做任何假设，甚至可能是对抗性的（adversarial），然而即使在这样的前提下，很多策略仍然能够保证玩家遭受的损失几乎和所有专家中最好的专家差不多。

Learning 问题则在序列预测的基础上认为存在一些 side information 能够暗示一些序列未来的信息，因此我们需要统计学习到其中的规律，以便更为准确地做预测。

由于本书中的设定是回合制的，在玩家和环境中相互交替，即玩家做预测，环境给出该时刻的实际序列。因此，环境和玩家之间或者多玩家之间可能形成一定的博弈。

Ps. 个人感觉，learning 和 prediction 的关系有点像被动投资（目标是跟随大盘）和主动投资（目标是绝对盈利）。

2. 一个例子

前面说到，prediction 问题研究的是：不管环境给出什么样的序列，是否存在一个基于专家的预测策略，使得玩家遭受的损失几乎和最好的专家遭受的损失差不多。注意到，每一轮都是 1）专家给出建议；2）玩家做出预测；3）环境给出结果；4）玩家遭受损失。这里不对专家和环境做出任何假设，甚至环境可以在玩家做出预测之后，给出专门针对玩家的结果。即使在这样的假设下，都有一个方案使得玩家和最优的专家水平差不多，这件事情其实比较反直觉，因此，下面考虑两个简单的例子来好让大家有个大致的印象。

考虑 binary 的序列 $y_t \in \{0,1\}$ ，每个专家给出建议 $f_{i,t} \in \{0,1\}$ ，然后玩家基于专家的意见作出决定 $\hat{p}_t \in \{0,1\}$ 。如果预测正确，损失为0；如果预测错误，损失为 1。

下面考虑一个简化，即已知存在一个专家一直不会犯错误，但是不会告诉你是哪一个专家。假设专家的数目为 N 。下面要展示的结论显示，存在一个策略，使得玩家最多犯错 $\lfloor \log_2 N \rfloor$ 次。策略也很简单，维护一个『可信』的专家池，每次预测结果都是『可信』专家池中 majority vote 的结果。一旦有专家预测出错，就把它移出这个专家池。

可以看出，玩家每次犯错都说明至少有超过一半的专家出错，因此至少能够剔除掉『可信』专家池中一半以上的专家。前提条件规定了，至少有一个专家是一直预测正确的，即至少有这样一个专家一直在专家池中。一旦专家池只剩下这一个专家的时候，玩家只需要每次都听从该专家的建议就不会再出错了。

3. 更进一步的例子

前面假设存在一个永不犯错的专家，这个假设显然比较强，那么把这个假设去掉还能得到一个比较好的策略么？能！

前面的专家池算是比较『硬性』的，一旦专家犯错就把它移除；而这里考虑一个比较『软性』的方案，即对于每个专家维护一个权重 w_i ，一旦专家犯错，就减小它的权重 $w_i \leftarrow \beta w_i$ ，其中 $0<\beta<1$ 是一个超参数。每个专家的初始权重都为 1。玩家做决策的方案就是如果投票给 0 的专家的总权重比投票给 1 的专家的总权重更大，就预测 0，反之预测 1。假定若干轮之后，最好的专家犯错的数目为 m^* ，那么玩家犯错的数目不会超过 $Am^*+B\log N$ ，其中 N 为专家的总数目。

证明过程如下：假定玩家在犯第 m 个错误之后的专家总权重为 W_m 。如果玩家不犯错，专家的总权重可能不变、可能减小；而玩家犯错过后则至少有权重超过一半的专家权重会减小 β ，即 $W_m \le (1/2+1/2\beta)W_{m-1}$ 。而初始的时候 $W_0=N$ ，有 $W_m \le (1/2+1/2\beta)^m W_0$ 。最好的那个专家犯错 m^* 次，那么它的权重应该为 β^{m^*} 。由于所有专家的权重都为正数，因此 $\beta^{m^*} \le (1/2+1/2\beta)^m N$ 。最后得到玩家犯错数 m 的上界。

$$m \le \left\lfloor \frac{\log_2 N + m^* \log_2(1/\beta)}{\log_2 \frac{2}{1+\beta}} \right\rfloor .$$

编辑于 2019-07-23

强化学习 (Reinforcement Learning) 博弈论 金融