高级算法/ 2026 年秋季

线性规划与优化专题

乘法权重更新

专家学习与近似优化

第6 讲：100 分钟

本讲义仅覆盖栗师老师负责的线性规划专题。本专题共八讲，讲次编号与课件一致。第一、第二部分是本讲内部的阅读划分；选读附录供课后拓展。

授课教师：栗师
南京大学

以授课教师的课件为依据，借助AI 将讲解、证明和课后练习展开为文字。中英文版本的数学内容相互对应。

学习目标与阅读路线

第一部分通过势函数分析建立在线学习的一个基本保证；第二部分将同一个不等式用于计算零和博弈的近似最优策略，以及线性规划的近似可行解。附录介绍用来引出该方法的确定性预测算法。

- 区分学习率η、目标误差δ 与宽度界ρ。
- 分别从上方估计总权重、从下方估计某个专家的权重，从而证明遗憾界。
- 将最优回应或聚合约束的求解过程转化为近似优化算法。

目录

1 第一部分：专家学习 1
1.1 在线决策问题 . . . 1
1.2 乘法权重更新规则 . . . 1
1.3 势函数分析 . . . 1
1.4 缩放损失范围 . . . 2

2 第二部分：博弈与线性规划近似可行性 2
2.1 零和博弈：明确双方的优化方向 . . . 2
2.2 通过一个聚合约束求近似可行解 . . . 3
2.3 课后练习 . . . 4

A 选读：折半法与加权多数法 4
A.1 存在完美专家时 . . . 4
A.2 不要求完美专家 . . . 5
A.3 为什么确定性预测无法突破因子二 . . . 5

1 第一部分：专家学习

1.1 在线决策问题

有$m \geq 1$ 个专家，决策过程共进行$T \geq 1$ 轮。在第t 轮损失揭晓之前，学习者选择一个分布$p^t \in \Delta_m$，其中$\Delta_m = \{p \geq 0 : \sum_i p_i = 1\}$。随后揭晓损失向量$\ell^t \in [-1,1]^m$。专家$i$ 承受损失ℓ_i^t，学习者承受分数损失$L_t = \langle p^t, \ell^t \rangle$。负损失表示收益。比较基准是事后表现最好的一个固定专家，而不是每一轮分别选择表现最好的专家。

定义1.1 (遗憾)**.**

$$R_T = \sum_{t=1}^{T} L_t - \min_{i \in [m]} \sum_{t=1}^{T} \ell_i^t.$$

若$R_T = o(T)$，则学习者的平均损失逐渐接近最佳固定专家的平均损失。

如果按照p^t 随机抽取专家，分数损失也可以解释为期望损失；这种解释要求当前损失向量不能依赖本轮抽到的专家。下面的代数不等式对每一条实际产生的损失序列都成立，即使该序列依赖过去的历史或已公布的分布。对分数损失的界，并不自动成为对随机抽样实际损失的逐条样本路径保证。

1.2 乘法权重更新规则

指数形式的乘法权重更新

选择$0 < \eta \leq 1$，并令所有$w_i^0 = 1$。对$t = 1, \ldots, T$：

1. 令$\Phi^{t-1} = \sum_i w_i^{t-1}$，选择$p_i^t = w_i^{t-1} / \Phi^{t-1}$。
2. 观察ℓ^t，承受损失$L_t = \sum_i p_i^t \ell_i^t$，并更新

$$w_i^t = w_i^{t-1} \exp(-\eta \ell_i^t).$$

损失越小，专家的相对权重越大。尤其是，算法不会因为某一轮表现不好就永久丢弃该专家。闭式表达$w_i^t = \exp(-\eta \sum_{s=1}^{t} \ell_i^s)$ 表明，算法以一种平滑的方式偏好累计损失较小的专家。

例1.2. 对于两个专家，取$\eta = \ln 2$，初始分布为$p^1 = (1/2, 1/2)$。若$\ell^1 = (1, 0)$，新权重为$(1/2, 1)$，于是$p^2 = (1/3, 2/3)$。若下一轮损失为$(0, 1)$，权重变为$(1/2, 1/2)$，分布恢复均匀。此时两位专家的累计损失恰好相同。

1.3 势函数分析

引理1.3. 对任意$z \in [-1, 1]$，有$e^z \leq 1 + z + z^2$。

证明. 当$z \in [0, 1]$ 时，幂级数展开给出$e^z - 1 - z \leq z^2 \sum_{k \geq 2} 1/k! = (e-2)z^2 \leq z^2$。当$z \in [-1, 0]$ 时，Taylor 定理给出$e^z = 1 + z + e^{\xi} z^2/2$，其中$\xi \in [z, 0]$，所以余项不超过$z^2/2$。 □

定理1.4 (MWU 遗憾界)**.** 上述算法对每个专家i 都满足

$$\frac{1}{T} \sum_{t=1}^{T} L_t \leq \frac{1}{T} \sum_{t=1}^{T} \ell_i^t + \eta + \frac{\ln m}{\eta T}. \tag{1.1}$$

证明. 由于$|\eta \ell_i^t| \le 1$，前面的引理给出

$$\begin{aligned}\Phi^t &= \Phi^{t-1} \sum_i p_i^t e^{-\eta \ell_i^t} \\ &\le \Phi^{t-1} \sum_i p_i^t (1 - \eta \ell_i^t + \eta^2 (\ell_i^t)^2) \\ &\le \Phi^{t-1} (1 - \eta L_t + \eta^2) \le \Phi^{t-1} \exp(-\eta L_t + \eta^2).\end{aligned}$$

最后一步使用$1 + u \le e^u$。从$\Phi^0 = m$ 迭代得到上界，而保留任意一个加数得到下界：

$$e^{-\eta \sum_t \ell_i^t} = w_i^T \le \Phi^T \le m \exp\Big(-\eta \sum_t L_t + T\eta^2\Big).$$

两边取对数，移项后除以ηT，便得到(1.1)。 □

因此，$T \ge \ln m / \eta^2$ 时，平均遗憾不超过2η。对于$m > 1$，若预先知道轮数$T \ge \ln m$，取$\eta = \sqrt{\ln m / T}$ 可得总遗憾不超过$2\sqrt{T \ln m}$。当$m = 1$ 时，学习者就是唯一的专家，遗憾为零。

1.4 缩放损失范围

推论1.5. 设$\ell_i^t \in [-\rho, \rho]$，其中$\rho > 0$。给定目标误差$0 < \delta \le 2\rho$，取

$$T = \max\left\{1, \left\lceil \frac{4\rho^2 \ln m}{\delta^2} \right\rceil\right\}, \qquad w_i^t = w_i^{t-1} \exp\left(-\frac{\delta \ell_i^t}{2\rho^2}\right).$$

则$T^{-1} \sum_t L_t \le \min_i T^{-1} \sum_t \ell_i^t + \delta$。

证明. 对归一化损失ℓ_i^t / ρ 应用定理，取学习率$\eta = \delta/(2\rho) \le 1$。归一化后的误差不超过$2\eta = \delta/\rho$，再将整个不等式乘以$\rho$ 即可。 □

第一部分的核心结论

同一种更新规则同时处理正、负损失。算法的平均损失可以与每一个固定专家比较。损失范围扩大后，所需轮数按ρ/δ 的平方增长。

2 第二部分：博弈与线性规划近似可行性

2.1 零和博弈：明确双方的优化方向

设$M \in [-1, 1]^{m \times n}$ 表示行玩家向列玩家的支付。行玩家希望最小化，列玩家希望最大化。对于$x \in \Delta_m$、$y \in \Delta_n$，记$M(x, y) = x^{\mathsf{T}} M y$。由第3 讲的极小极大定理，双方有共同的博弈值

$$\lambda^* = \min_x \max_j M(x, j) = \max_y \min_i M(i, y).$$

把每一行看作一个专家。在第t 轮，MWU 给出行分布x^t；选择使$M(x^t, j)$ 最大的列j^t，并令损失$\ell_i^t = M_{i, j^t}$。给定$0 < \delta \le 2$，参数取为：

$$\eta = \frac{\delta}{2}, \qquad T = \max\left\{1, \left\lceil \frac{4 \ln m}{\delta^2} \right\rceil\right\}.$$

提取双方的近似策略

令$\hat{t}$ 为使$M(x^t, j^t)$ 最小的一轮，并设$\hat{x} = x^{\hat{t}}$。令$\hat{y}$ 为所选各列的经验分布：

$$\hat{y}_j = \frac{|\{t : j^t = j\}|}{T}.$$

同一列在不同轮中多次出现时，需要重复计数。

定理2.1.

$$\max_j M(\hat{x}, j) \le \min_i M(i, \hat{y}) + \delta.$$

因此，$\hat{x}$ 保证支付不超过$\lambda^* + \delta$，而$\hat{y}$ 保证得到至少$\lambda^* - \delta$ 的支付。

证明. 依次使用最优回应性质、$\hat{t}$ 的选择方式和遗憾不等式，得到

$$\begin{aligned}\max_j M(\hat{x}, j) = M(x^{\hat{t}}, j^{\hat{t}}) &\le \frac{1}{T}\sum_t M(x^t, j^t) \\ &\le \min_i \frac{1}{T}\sum_t M(i, j^t) + \delta = \min_i M(i, \hat{y}) + \delta.\end{aligned}$$

任意行策略的最坏支付至少为λ^*，任意列策略所能保证的支付至多为λ^*。因此

$$\lambda^* \le \max_j M(\hat{x}, j) \le \min_i M(i, \hat{y}) + \delta \le \lambda^* + \delta.$$

□

平均行策略$\bar{x} = T^{-1}\sum_t x^t$ 也可行：函数$x \mapsto \max_j M(x, j)$ 的凸性保证，其最坏支付不超过每轮支付的平均值。课件采用最佳迭代点，这样无需存储平均策略。对于显式给出的矩阵，一次最优回应需要$O(mn)$ 次算术运算；当列对应指数多个选择时，关键资源则是一个最优回应预言机。

2.2 通过一个聚合约束求近似可行解

设$K \subseteq \mathbb{R}^n$ 是一个非空紧凸多胞体，并且在K 上容易进行线性优化。目标是寻找满足$Ax \ge b$ 的$x \in K$，其中A 有$m \ge 1$ 行。假设已知宽度界

$$|A_i x - b_i| \le \rho \quad (x \in K,\ i \in [m]), \qquad \rho > 0.$$

例如，当$K = [0,1]^n$ 时，只要$\max_i(\sum_j |a_{ij}| + |b_i|) > 0$，便可将它作为$\rho$；任意更大的正数也适用。给定$y \in \Delta_m$，只需求解一个线性优化问题

$$\max_{x \in K} \langle y, Ax - b\rangle.$$

对于盒子$K = [0,1]^n$，按$A^{\mathsf{T}}y$ 中各坐标的符号选择$x_j = 0$ 或1 即可。在一般情形中，算法假设可以调用这一线性优化预言机。

用于线性规划近似可行性的MWU

选择$0 < \delta \le 2\rho$，并按照推论1.5 设置轮数。初始化$w_i^0 = 1$。第t 轮：

1. 将权重归一化，得到$y^t \in \Delta_m$。
2. 寻找满足$\langle y^t, Ax^t - b\rangle \ge 0$ 的$x^t \in K$。若该表达式在K 上的最大值仍为负，则返回“不可行”。

3. 令$\ell_i^t = A_i x^t - b_i$，并更新

$$w_i^t = w_i^{t-1} \exp\left(-\frac{\delta(A_i x^t - b_i)}{2\rho^2}\right).$$

若所有轮次均成功，则返回$\hat{x} = T^{-1}\sum_t x^t$。

定理2.2. 算法或者正确证明精确可行域为空，或者返回满足$A\hat{x} \geq b - \delta\mathbf{1}$ 的$\hat{x} \in K$。

证明. 若存在精确可行解x^*，则由$y \geq 0$ 可知，每个分布y 都满足$\langle y, Ax^* - b\rangle \geq 0$。因此，预言机返回的负最优值是有效的不可行性证书。若算法没有提前终止，则学习者每轮的损失均非负。对每一行i，缩放后的遗憾界给出

$$0 \leq \frac{1}{T}\sum_t \langle y^t, Ax^t - b\rangle \leq \frac{1}{T}\sum_t (A_i x^t - b_i) + \delta = A_i\hat{x} - b_i + \delta.$$

所以每个约束的违反量不超过δ。由K 的凸性又有$\hat{x} \in K$。 □

更新的符号乍看有些反直觉：一个满足得较好的约束得到正损失，权重反而减小。这恰好符合需要。被违反的约束具有负松弛量，因此权重增大，下一轮的聚合约束会更加重视它。

保证的含义与边界

若精确系统不可行，但存在误差为δ 的近似可行解，则两种输出均被允许。这是加性误差保证，依赖各行约束的缩放方式。迭代次数依赖$(\rho/\delta)^2$，所以该方法本身并不是关于二进制输入长度的多项式时间精确LP 算法。分析采用精确预言机与理想指数更新；实现时需另外控制数值误差和预言机误差。

2.3 课后练习

1. 用一般的学习率η 重新计算双专家例子。若两轮损失向量都是$(1, 0)$，结果如何?
2. 证明：在某一轮所有专家的损失上加同一个标量，不会改变下一轮的归一化分布。解释为什么宽度界仍可能改变。
3. 证明平均行策略的近似博弈保证。任意一个迭代点是否也有该保证?
4. 若每轮所选列的支付与最优回应相差至多γ，证明最终的鞍点间隙不超过$\delta + \gamma$。
5. 设$K = [0, 1]^2$，要求$x_1 \geq 3/4$、$x_2 \geq 3/4$ 和$x_1 + x_2 \leq 1$。寻找这些约束的非负线性组合，以证明不可行。
6. 若成功的预言机调用仅保证$\langle y^t, Ax^t - b\rangle \geq -\gamma$，证明返回的平均点具有$\delta + \gamma$ 的可行性误差。指出在报告精确不可行之前还需要什么证据。

A 选读：折半法与加权多数法

A.1 存在完美专家时

在二元预测问题中，假设至少有一个专家从不犯错。只保留至今一直正确的专家，并按它们的多数意见预测。学习者每犯一次错误，至少一半的剩余专家就会被淘汰。犯错M 次之后，剩余专家数至多为$m/2^M$，而完美专家始终存活。因此$M \leq \log_2 m$，当然也有$M \leq \lceil \log_2 m \rceil$。

A.2 不要求完美专家

从单位权重开始，按加权多数意见预测。每轮结束后，将预测错误的专家权重减半。当学习者预测错误时，至少一半总权重对应的专家也预测错误，所以总权重至少下降四分之一。若专家i 犯错M_i 次，学习者犯错M 次，则

$$2^{-M_i} \le \Phi^T \le m(3/4)^M, \qquad M \le \frac{\ln 2}{\ln(4/3)} M_i + \frac{\ln m}{\ln(4/3)} \le 2.41 M_i + 3.48 \ln m.$$

更一般地，对犯错专家的权重乘以$1-\eta$，其中$0 < \eta < 1$。学习者每犯一次错误，势函数至多变为原来的$1-\eta/2$，因而

$$M \le \frac{\ln(1-\eta)}{\ln(1-\eta/2)} M_i - \frac{\ln m}{\ln(1-\eta/2)} = (2 + O(\eta)) M_i + O(\eta^{-1} \ln m).$$

A.3 为什么确定性预测无法突破因子二

考虑两个专家，一个总预测零，另一个总预测一。对任意确定性学习者，每轮都令结果与其预测相反。学习者犯错T 次，而至少一位专家犯错不超过$T/2$ 次。因此，对于固定$\gamma > 0$ 以及不依赖T 的有限加性项，不可能对所有T 都有$M \le (2-\gamma)\min_i M_i + f(m)$。在结果不能针对学习者本轮随机样本作出反应的通常假设下，随机选择会改变结论。对$\{0,1\}$ 损失，使用更精确的恒等式$e^{-\eta\ell} = 1 - (1-e^{-\eta})\ell$ 可得

$$\sum_t L_t \le \frac{\eta}{1-e^{-\eta}} \min_i \sum_t \ell_i^t + \frac{\ln m}{1-e^{-\eta}} = (1 + O(\eta)) \min_i \sum_t \ell_i^t + O(\eta^{-1} \ln m).$$

这就是课件中提到的乘法因子趋近于$1+\varepsilon$ 的保证。它针对分数损失或在适当模型下的期望损失，因此与确定性下界并不矛盾。

来源与阅读说明

主要依据：授课教师的课件文件夹

`6.MWU`

`Topics.tex` 与`Optional.tex`.

选读部分对应课件中的选讲内容（如有）。补充的例子和练习服务于课后阅读，不占用新的授课讲次。

课程网站. 网站上前段课程的讲义用于参考组织方式与展开程度，本专题的内容仍以线性规划课件为准。