中国科学技术大学
University of Science and Technology of China

Follow The Regularized Leader

彭辉阳

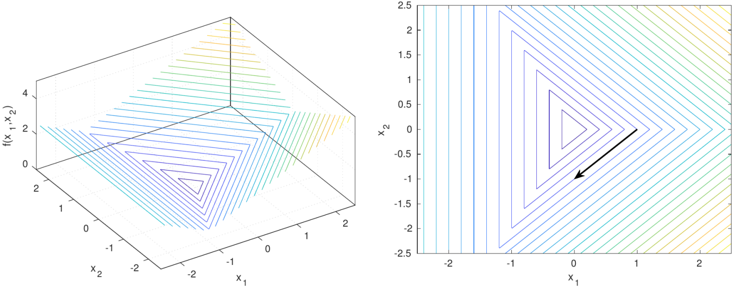
USTC 管理学院

2025 年 5 月 27 日

为什么 OSD 有效？

- 次梯度并不一定一直指向函数值在最小点。
- OSD 有效的依据是，次梯度指向其中一个函数值下降方向：

$$\ell(x_t) - \ell(u) \leq \langle g_t, x_t - u \rangle$$



中国科学技术大学 University of Science and Technology of China

由次梯度的定义，$f(x)$ 可由一个局部线性函数近似：

$$f(x) \geq \tilde{f}(x) = f(x_0) + \langle g, x - x_0 \rangle, \forall x \in V$$

然后我们将他限定在某一个小邻域内：

$$\begin{aligned} x_{t+1} = \quad & \arg\min_{x \in V} f(x_t) + \langle g, x - x_t \rangle \\ & s.t. ||x_t - x||^2 \leq h \end{aligned}$$

类似地，对于某个 $\eta > 0$，我们也能定义一个二次局部近似函数

$$\arg\min_{x \in V} \hat{f}(x) = f(x_0) + \langle g, x - x_0 \rangle + \frac{1}{2\eta} ||x_0 - x||_2^2$$

最后一步，我们可以证明，对二次局部近似函数求最小值的过程，可以等效为我们在 OSD 中进行的迭代：

$$
\begin{aligned}
\arg\min_{x\in V}\langle g_t, x\rangle + \frac{1}{2\eta_t}||x_t - x||_2^2 &= \arg\min_{x\in V}||\eta_t g_t||^2 + 2\eta_t\langle g_t, x - x_t\rangle + ||x_t - x||_2^2 \\
&= \arg\min_{x\in V}||x_t - \eta_t g_t - x||_2^2 \\
&= \Pi_V(x_t - \eta_t g_t)
\end{aligned}
$$

这里 $\Pi_V(x) = \arg\min_{y\in V}||x - y||_2$。

定义 (1)

设 $\psi : X \to \mathbb{R}$ 强凸、连续可微，那么 **Bregman 散度定义为**：$B_\psi : X \times \mathrm{int}(X) \to \mathbb{R}$

$$B_\psi(x; y) = \psi(x) - \psi(y) - \langle \nabla\psi(y), x - y \rangle.$$

我们将 Bregman 散度用作一种相似度的度量，他有如下性质：

- $B_\psi(x; y) \geq 0$，等号成立当且仅当 $x = y$。
- 大多数情况下 $B_\psi(x; y) \neq B_\psi(y; x)$，无对称性。
- 三角不等式：$B_\psi(z; x) + B_\psi(x; y) - B_\psi(z; y) = \langle \nabla\psi(y) - \nabla\psi(x), z - x \rangle$。

由泰勒公式，存在某个 $\alpha \in [0,1]$，且 $z = \alpha x + (1-\alpha)y$:

$$B_\psi(x;y) = \psi(x) - \psi(y) - \langle \nabla \psi(y), x-y \rangle = \frac{1}{2}(x-y)^T \nabla^2 \psi(z)(x-y)$$

我们能看出这种度量依赖于 ψ 的 Hessian 矩阵信息。

另外，如果 ψ 是 $\lambda-$ 强凸的，那么有:

$$B_\psi(x;y) \geq \frac{\lambda}{2}||x-y||^2$$

Algorithm 1 在线镜像下降

Require: $V \subseteq X \subseteq \mathbb{R}^d$, $\psi : X \to \mathbb{R}$ 在 $\mathrm{int}(X)$ 上强凸且连续可微, $x_1 \in V$, ψ 在 x_1 处可微; $\eta_1, \cdots, \eta_T > 0$

1: **for** $t = 1$ to T **do**
2: 输出 x_t
3: 接受 ℓ_t 并计算 $\ell_t(x_t)$
4: 计算 $g_t \in \partial \ell_t(x_t)$
5: 更新 $x_{t+1} = \arg\min_{x \in V} \langle g_t, x \rangle + \frac{1}{\eta_t} B_\psi(x; x_t)$
6: **end for**

该算法可能存在问题，如果 $x_{t+1} \in \partial X$，那么在 x_{t+1} 处 Bregman 散度没有定义，更新无法继续下去。
要解决该问题，往往需要下面两个假设之一：

1.
$$\lim_{x \to \partial X} ||\nabla \psi(x)||_2 = +\infty$$

2.
$$V \subseteq \mathrm{int}(X)$$

下面我们要分析研究 OMD 的 regret 性质。

引理 (1)

$\psi : X \to \mathbb{R}$ 在 V 中是关于 $||\cdot||\lambda-$ 强凸的。设 $V \subseteq X$ 是一个凸集，$g_t \in \partial \ell_t(x_t)$。若上述两个条件之一成立，那么 $\forall u \in V$，下面的不等式成立：

$$\eta_t(\ell_t(x_t) - \ell_t(u)) \le \eta_t \langle g_t, x_t - u \rangle \le B_\psi(u; x_t) - B_\psi(u; x_{t+1}) + \frac{\eta_t^2}{2\lambda}||g_t||_*^2$$

证明.

由 OMD 更新的最优条件，有：

$$\langle \eta_t g_t + \nabla\psi(x_{t+1}) - \nabla\psi(x_t), u - x_{t+1} \rangle \ge 0, \forall u \in V$$

左边的不等式是显然的。

$$
\begin{aligned}
& \langle \eta_t g_t, x_t - u \rangle \\
= \quad & \langle \nabla\psi(x_t) - \nabla\psi(x_{t+1}) - \eta_t g_t, u - x_{t+1} \rangle + \langle \nabla\psi(x_{t+1}) - \nabla\psi(x_t), u - x_{t+1} \rangle + \langle \eta_t g_t, x_t - x_{t+1} \rangle \\
\leq \quad & \langle \nabla\psi(x_{t+1}) - \nabla\psi(x_t), u - x_{t+1} \rangle + \langle \eta_t g_t, x_t - x_{t+1} \rangle \\
= \quad & B_\psi(u; x_t) - B_\psi(u, x_{t+1}) - B_\psi(x_{t+1}; x_t) + \langle \eta_t g_t, x_t - x_{t+1} \rangle \\
\leq^{(*)} \quad & B_\psi(u; x_t) - B_\psi(u, x_{t+1}) - B_\psi(x_{t+1}; x_t) + \frac{\eta_t^2}{2\lambda} ||g_t||_*^2 + \frac{\lambda}{2} ||x_t - x_{t+1}||^2 \\
\leq \quad & B_\psi(u; x_t) - B_\psi(u, x_{t+1}) + \frac{\eta_t^2}{2\lambda} ||g_t||_*^2
\end{aligned}
$$

这里 $(*)$ 由 Fenchel-Young 不等式可得，$\langle p, x \rangle \leq f(x) + f^*(p)$，其中取 $f(x) = \frac{\lambda}{2}||x||^2$。 □

定理 (1)

设 $x_1 \in V$ 且 ψ 在 x_1 处可微，且 $\eta_{t+1} \leq \eta_t$，$t = 1, \cdots, T$。那么在引理 *1* 的假设下，下面的不等式成立：

$$\sum_{t=1}^{T}(\ell_t(x_t) - \ell_t(u)) \leq \max_{1 \leq t \leq T} \frac{B_\psi(u; x_t)}{\eta_t} + \frac{1}{2\lambda}\sum_{t=1}^{T}\eta_t ||g_t||_*^2$$

如果 $\eta_t = \eta$ 为常数，那么有：

$$\sum_{t=1}^{T}(\ell_t(x_t) - \ell_t(u)) \leq \frac{B_\psi(u; x_1)}{\eta} + \frac{\eta}{2\lambda}\sum_{t=1}^{T}||g_t||_*^2$$

证明.

对引理 1 的结论作累加，得：

$$\begin{aligned}
&\sum_{t=1}^{T}(\ell_t(x_t)-\ell_t(u)) \le \sum_{t=1}^{T}\left(\frac{1}{\eta_t}B_\psi(u;x_t)-\frac{1}{\eta_t}B_\psi(u;x_{t+1})\right)+\sum_{t=1}^{T}\frac{\eta_t}{2\lambda}||g_t||_*^2\\
&=\frac{1}{\eta_1}B_\psi(u;x_1)-\frac{1}{\eta_T}B_\psi(u;x_{T+1})+\sum_{t=1}^{T-1}\left(\frac{1}{\eta_{t+1}}-\frac{1}{\eta_t}\right)B_\psi(u;x_{t+1})+\sum_{t=1}^{T}\frac{\eta_t}{2\lambda}||g_t||_*^2\\
&\le\frac{1}{\eta_1}D^2+D^2\sum_{t=1}^{T-1}\left(\frac{1}{\eta_{t+1}}-\frac{1}{\eta_t}\right)+\sum_{t=1}^{T}\frac{\eta_t}{2\lambda}||g_t||_*^2\\
&=\frac{D^2}{\eta_T}+\sum_{t=1}^{T}\frac{\eta_t}{2\lambda}||g_t||_*^2
\end{aligned}$$

这里我们记：$D^2=\max_{1\le t\le T}B_\psi(u;x_t)$。 □

定理 (2)

$\psi : X \to \mathbb{R}$ 是 $\lambda > 0$ 强凸的。$V \subseteq X$ 是一个闭凸集，且 $\mathrm{int}(X) \cap V \neq \emptyset$，那么

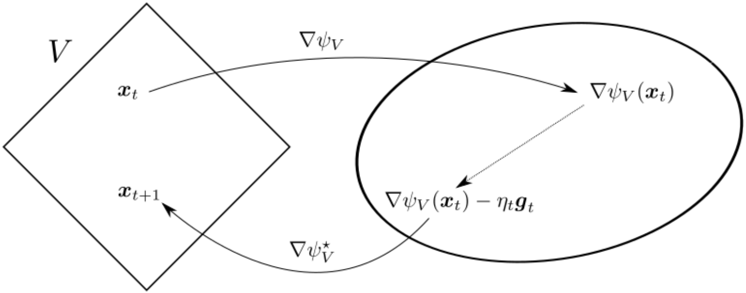
$$\arg\min_{x \in V} \langle g_t, x \rangle + \frac{1}{\eta_t} B_\psi(x; x_t) = \nabla \psi_V^*(\nabla \psi_V(x_t) - \eta_t g_t)$$

其中 ψ_V 为将 ψ 限制到 V 上，$\psi_V = \psi + i_V$，$i_V(x) = \begin{cases} 0, & x \in V \\ +\infty, & x \notin V \end{cases}$。

现在我们有一个新的更新公式

$$x_{t+1} = \nabla \psi_V^*(\nabla \psi_V(x_t) - \eta_t g_t),$$

这意味着什么？



- $\nabla\psi_V$ 和 $\nabla\psi_V^*$ 是从原空间和对偶空间转换的算子。
- 初始点 x_t，转换到对偶空间 $\nabla\psi_V(x_t)$，在对偶空间中次梯度下降。
- 通过 $\nabla\psi_V^*$ 转换回原空间。

Contents

在前面的算法中，我们对每一轮的更新都使用同样的正则项 ψ，在 FTRL 算法中，我们在每一轮中优化过去的损失之和加上正则项作最优化。

Algorithm 2 FTRL

Require: 一系列正则项 $\psi_1, \cdots, \psi_T : \mathbb{R}^d \to (-\infty, +\infty]$

1: **for** $t = 1$ to T **do**

2: 输出 $x_t \in \arg\min_{x \in \mathbb{R}^d} \psi_t(x) + \sum_{i=1}^{t-1} \ell_i(x)$

3: 接受 ℓ_t，计算 $\ell_t(x_t)$

4: **end for**

FTRL 是**一族**算法，就像 OMD 一样，随着正则项 $\psi_1, \cdots, \psi_T$ 变化而变化。

- 对于 OMD 而言，模型当前的“状态”保存在 x_t 中，更新 x_{t+1} 需要依赖于 x_t 和 ℓ_t 的信息。
- 对于 FTRL 而言，x_{t+1} 的更新依赖于过去的所有历史损失函数 $\ell_1, \cdots, \ell_t$。
- 因为要保留更多的信息，我们可能认为 FTRL 在计算和内存方面更为昂贵。事实也确实如此。
- 但我们也将看到，通过采用近似损失的方法，可以使该算法的成本与 OMD 相当，同时仍能保留比 OMD 更严格的信息。

引理 (2)

记 $F_t(x) = \psi_t(x) + \sum_{i=1}^{t-1} \ell_i(x)$。设 $\arg\min_{x\in\mathbb{R}^d} F_t(x) \neq \emptyset$，且 $x_t \in \arg\min_{x\in\mathbb{R}^d} F_t(x)$，那么对于 $\forall u$，有

$$\sum_{t=1}^{T}(\ell_t(x_t) - \ell_t(u)) = \psi_{T+1}(u) - \min_x \psi_1(x) + \sum_{t=1}^{T}[F_t(x_t) - F_{t+1}(x_{t+1}) + \ell_t(x_t)] + F_{T+1}(x_{T+1}) - F_{T+1}(u)$$

该引理的证明如下：

证明.

因为 $\ell_t(x_t)$ 在两边同时出现，所以实际上只需要验证：

$$-\sum_{t=1}^{T}\ell_t(u)=\psi_{T+1}(u)-\min_x \psi_1(x)+\sum_{t=1}^{T}[F_t(x_t)-F_{t+1}(x_{t+1})]+F_{t+1}(x_{t+1})-F_{t+1}(u)$$

因为 $F_1(x_1)=\min_{x\in V}\psi_1(x)$，所以

$$\begin{aligned}&-\textstyle\sum_{t=1}^{T}\ell_t(u)=\psi_{T+1}(u)-F_1(x_1)+F_1(x_1)-F_{T+1}(x_{T+1})+F_{T+1}(x_{T+1})-F_{T+1}(u)\\&=\psi_{T+1}(u)-F_{T+1}(u)\end{aligned}$$

□

然而 Regret 等式并不是一个 regret 上界的保证，因为 ℓ_t 项在等式的两边出现。在给出强凸情况下的 regret 上界之前，我们给出两个关于强凸函数的引理。

引理 (3)

$f : \mathbb{R}^d \to (-\infty, +\infty]$ 关于范数 $||\cdot||$ 是 $\mu-$ 强凸的，那么对于 $x, y \in \mathrm{dom}(f), g \in \partial f(y), g' \in \partial f(x)$，有

$$f(x) - f(y) \leq \langle g, x - y \rangle + \frac{1}{2\mu}||g - g'||_*^2$$

证明.

定义 $\varphi(z) = f(z) - \langle g, z \rangle$，因为 $0 \in \partial \varphi(y)$，因此 $y \in \arg\min_{z \in \mathrm{dom}(\varphi)} \varphi(z)$，另外，因为 $g' - g \in \partial \varphi(x)$，所以

$$
\begin{aligned}
\varphi(y) &= \min_{z \in \mathrm{dom}(\varphi)} \varphi(z) \\
&\geq \min_{z \in \mathrm{dom}(\varphi)} (\varphi(x) + \langle g' - g, z - x \rangle + \frac{\mu}{2} ||z - x||^2) \\
&\geq \inf_{z \in \mathbb{R}^d} (\varphi(x) + \langle g' - g, z - x \rangle + \frac{\mu}{2} ||z - x||^2) \\
&=^{(*)} \varphi(x) - \frac{1}{2\mu} ||g' - g||_*^2
\end{aligned}
$$

其中 $(*)$ 来源于共轭函数的定义。 □

取 $y = x^*$，那么 $0 \in \partial f(y)$，于是有如下显然的推论：

推论 (1)

$f : \mathbb{R}^d \to (-\infty, +\infty]$ 关于范数 $||\cdot||$ 是 $\mu-$ 强凸的，设 $x^* = \arg\min_x f(x)$，那么 $\forall x \in \mathrm{dom}(f), g \in \partial f(x)$，有

$$f(x) - f(x^*) \leq \frac{1}{2\mu}||g||_*^2$$

下面我们给出一个引理，衡量 FTRL 算法的稳定性——即相邻两次预测函数值之差的上界：

引理 (4)

在引理 2 的假设下，如果 F_t 是关于 $||\cdot||$ 的正常 λ_t- 强凸函数，且 ℓ_t 是正常凸函数。另外假设 $\partial\ell_t(x_t) \neq \emptyset$，那么

$$F_t(x_t) - F_{t+1}(x_{t+1}) + \ell_t(x_t) \leq \frac{||g_t||_*^2}{2\lambda_t} + \psi_t(x_{t+1}) - \psi_{t+1}(x_{t+1})$$

对于任意的 $g_t \in \partial\ell_t(x_t)$ 成立。

证明.

$$\begin{aligned} F_t(x_t) - F_{t+1}(x_{t+1}) + \ell_t(x_t) &= (F_t(x_t) + \ell_t(x_t)) - (F_t(x_{t+1}) + \ell_t(x_{t+1})) + \psi_t(x_{t+1}) - \psi_{t+1}(x_{t+1}) \\ &\leq (F_t(x_t) + \ell_t(x_t)) - (F_t(x_t^*) + \ell_t(x_t^*)) + \psi_t(x_{t+1}) - \psi_{t+1}(x_{t+1}) \\ &\leq^{(*)} \frac{||g_t||_*^2}{2\lambda_t} + \psi_t(x_{t+1}) - \psi_{t+1}(x_{t+1}) \end{aligned}$$

其中不等式 $(*)$ 成立是因为 $\partial\ell_t(x_t) \neq \emptyset$，所以我们可以选择 $g_t' \in \partial\ell_t(x_t)$，因为 $x_t = \arg\min_x F_t(x)$，所以 $0 \in \partial F_t(x_t)$，于是 $g_t' \in \partial(F_t(x_t) + \ell_t(x_t))$，且 $x_t^* = \arg\min_x F_t(x) + \ell_t(x)$，应用推论 1 即可。 □

Contents

当 ℓ_i 为线性函数时，FTRL 的更新公式可以写作：

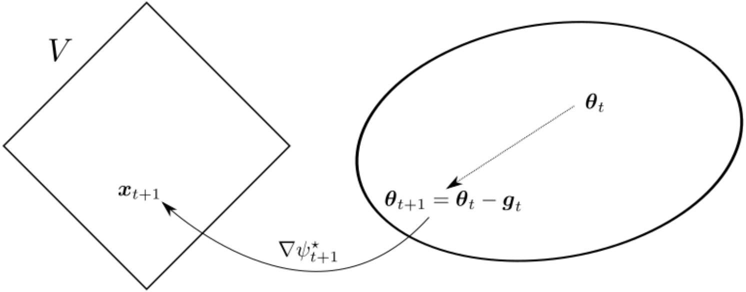
$$x_{t+1} \in \arg\min_x \psi_{t+1}(x) + \sum_{i=1}^{t} \langle g_i, x \rangle = \arg\max_x \langle -\sum_{i=1}^{t} g_i, x \rangle - \psi_t(x)$$

由对偶函数的极大值条件，有 $x_{t+1} \in \partial \psi_{t+1}^*(-\sum_{i=1}^{t} g_i)$
此外，如果 ψ_{t+1} 强凸，则 ψ_{t+1}^* 可微，于是

$$x_{t+1} = \nabla \psi_{t+1}^*(-\sum_{i=1}^{t} g_i)$$

这一更新公式可以写作：

$$\theta_{t+1} = \theta_t - g_t$$
$$x_{t+1} = \nabla\psi^*_{t+1}(\theta_{t+1})$$



OMD 也可以写成类似的形式：

$$\begin{aligned}\theta_{t+1} &= \nabla\psi(x_t) - \eta_t g_t \\ x_{t+1} &= \nabla\psi^*(\theta_{t+1})\end{aligned}$$

- OMD 算法中，状态保存在 x_t 中，所以要用 $\nabla\psi$ 转化到对偶空间来更新，再转回到原空间中。
- 线性损失 FTRL 中，状态就在对偶空间中，所以直接进行更新，原空间只用来做预测，不用来更新。
- OMD 中，学习率 η_t 通常是递减的，样本权重逐渐减小。
- 线性损失 FTRL 中，次梯度的权重都相同，但正则项通常是增加的。

Algorithm 3 线性损失 FTRL

Require: 一系列正则项 $\psi_1, \cdots, \psi_T : \mathbb{R}^d \to (-\infty, +\infty]$

1: **for** $t = 1$ to T **do**

2: 输出 $x_t \in \arg\min_{x \in \mathbb{R}^d} \psi_t(x) + \sum_{i=1}^{t-1} \langle g_i, x \rangle$

3: 接受 ℓ_t，计算 $\ell_t(x_t)$

4: 计算 $g_t \in \partial \ell_t(x_t)$

5: **end for**

考虑 $V=\{x\in\mathbb{R}^d:||x||_2\leq 1\}$，先考虑 OMD，正则项函数 $\psi(x)=\frac{||x||^2}{2}$，学习率 $\eta_t=\frac{1}{\sqrt{t}}$，$x_1=0$，那么更新公式为：

$$
\begin{aligned}
\tilde{x}_{t+1} &= x_t-\frac{1}{\sqrt{t}}g_t \\
x_{t+1} &= \tilde{x}_{t+1}\min\left(\frac{1}{||\tilde{x}_{t+1}||^2},1\right)
\end{aligned}
$$

对于线性损失 FTRL，使用 $\psi_t(x)=\frac{\sqrt{t}}{2}||x||_2^2+i_V(x)$，那么其更新公式为：

$$
\begin{aligned}
\tilde{x}_{t+1} &= \frac{-\sum_{i=1}^t g_i}{\sqrt{t}} \\
x_{t+1} &= \tilde{x}_{t+1}\min\left(\frac{1}{||\tilde{x}_{t+1}||^2},1\right)
\end{aligned}
$$

直觉上，FTRL 会使用更多的信息，而 OMD 只使用前面一步的结果。

Contents

重复进行以下游戏：

- 初始时，拥有财富 ϵ： $\mathrm{Wealth}_0 = \epsilon$
- 在 $t = 1, \cdots, T$ 的每一轮中：
    - 赌上价值 $|x_t|$ 的财富在 $sign(x_t)$ 代表的硬币一面，你不能赌上比你现在拥有的更多的钱。
    - 抛硬币决定结果： $c_t \in \{-1, 1\}$
    - 获得财富： $c_t x_t$，即： $\mathrm{Wealth}_t = \mathrm{Wealth}_{t-1} + c_t x_t = \epsilon + \sum_{i=1}^{t} c_i x_i$

因为不允许借钱，所以可以记 $x_t = \beta_t \mathrm{Wealth}_{t-1}$，这里 $\beta_t \in [-1, 1]$，$|\beta_t|$ 是赌上的财富占比，$sign(\beta_t)$ 是赌注的硬币朝向。

我们在这个游戏中没法一直赢下去，所以我们目标是尽量跑过一个在正常游戏中采用固定的 β^* 的策略。

$$\begin{aligned}\mathrm{Wealth}_t &= \mathrm{Wealth}_{t-1} + c_t x_t = \mathrm{Wealth}_{t-1} + c_t \beta_t \mathrm{Wealth}_{t-1} = \mathrm{Wealth}_{t-1}(1 + c_t\beta_t) \\ &= \epsilon \prod_{i=1}^{t}(1 + c_t\beta_t)\end{aligned}$$

我们希望最小化下面的 regret：

$$\begin{aligned} & \ln \max_{\beta\in[-1,1]} \epsilon \prod_{t=1}^{T}(1+\beta c_t) - \ln \mathrm{Wealth}_T \\ = & \ln \max_{\beta\in[-1,1]} \epsilon \prod_{t=1}^{T}(1+\beta c_t) - \ln(\epsilon \prod_{t=1}^{T}(1+\beta_t c_t)) \\ = & \max_{\beta\in[-1,1]} \sum_{t=1}^{T} \ln(1+\beta c_t) - \sum_{t=1}^{T} \ln(1+\beta_t c_t) \end{aligned}$$

这实际上和损失函数 $\ell_t(x) = -\ln(1 + xc_t)$, $V = [-1, 1]$ 的在线凸优化 (OCO) 问题是一致的，这里我们允许硬币是连续的，因而 $c_t \in [-1, 1]$。

我们可以对这个问题去应用 OMD 或者 FTRL 算法，但事实上对这个问题存在着一个特定的策略，叫做 Krichevsky-Trofimov (KT) 投注。

Algorithm 4 Krichevsky-Trofimov (KT) 投注

Require: 初始财富 $\mathrm{Wealth}_0 = \epsilon > 0$

for $t = 1$ to T **do**
  计算投注比例 $\beta_t = \frac{\sum_{i=1}^{t-1} c_i}{t}$
  进行投注：$x_t = \beta_t \mathrm{Wealth}_{t-1}$
  获取硬币的结果：$c_i \in [-1, 1]$
  得到或失去对应的财富：$\mathrm{Wealth}_t = \mathrm{Wealth}_{t-1} + c_t x_t$
end for

对于这一策略，其经过 T 轮后的财富有一个下界：

定理 (3)

[1](Cesa-Bianchi, N. and Lugosi, G. , 2006, Theorem 9.4) 如果 $c_t \in \{-1, 1\}$ 对于 $t = 1, 2, \cdots, T$ 都成立，那么 KT 投注策略有下界：

$$\ln \mathrm{Wealth}_T \geq \ln \max_{\beta \in [-1,1]} \prod_{t=1}^{T} (1 + \beta c_t) - \frac{1}{2} \ln T - K$$

最优的固定 β 的策略中，财富变化的倍数有一个指数下界：

引理 (5)

如果 $g_t \in \{-1, 1\}$，$t = 1, \cdots, T$，那么我们有：

$$\max_{\beta \in [-1,1]} \exp\left(\sum_{t=1}^{T} \ln(1 - \beta g_t)\right) \geq \exp\left(\sum_{t=1}^{T} \frac{(\sum_{t=1}^{T} g_t)^2}{4T}\right)$$

证明.

$\max_{\beta \in [-1,1]} \exp(\sum_{t=1}^{T} \ln(1 - \beta g_t)) \geq \max_{\beta \in [-1/2,1/2]} \exp(\sum_{t=1}^{T} \ln(1 - \beta g_t))$

$\geq \max_{\beta \in [-1/2,1/2]} \exp(\beta \sum_{t=1}^{T} g_t - \beta \sum_{t=1}^{T} g_t^2)$ $(\ln(1+x) \geq x - x^2)$ $\square$

$\geq \max_{\beta \in [-1/2,1/2]} \exp(\beta \sum_{t=1}^{T} g_t - \beta^2 T) = \exp(\frac{(\sum_{t=1}^{T} g_t)^2}{4T})$

KT 投注策略保证了一个指数下界，而仅仅付出 $\sqrt{T}$ 的代价。注意到 KT 投注策略不需要设定任何额外的超参数，如学习率、正则项等等，也就是说，KT 投注策略是 Parameter-free 的。
另外，我们也可以将 KT 投注策略拓展到连续硬币的情况：

定理 (4)

(Orabona, F. and Pal, D., 2016, Lemma 14) 如果 $c_t \in [-1, 1]$ 对于 $t = 1, 2, \cdots, T$ 都成立，那么 *KT* 投注策略有下界：

$$\ln \mathrm{Wealth}_T \geq \sum_{t=1}^{T} \frac{(\sum_{t=1}^{T} c_t)^2}{4T} - \frac{1}{2} \ln T - K$$

首先，Coin-Betting 问题等价于，设计一个算法来最小化下面的带线性损失的一维 regret：

$$\mathrm{Regret}_T(u) = \sum_{t=1}^{T} g_t x_t - \sum_{t=1}^{T} g_t u$$

另外，OCO 问题可以被简化为在线线性优化 (OLO) 问题。
于是，我们可以联想到，可以将 OCO 问题用 Coin-Betting 算法的形式来解决。

下面的定理允许我们在 OLO 和 Coin-Betting 问题之间做转换：

定理 (4)
$\phi : \mathbb{R}^d \to (-\infty, +\infty]$ 是一个合理闭凸函数，$\phi^* : \mathbb{R}^d \to (-\infty, +\infty]$ 是它的 *Fenchel* 共轭。某个算法产生某个序列 $x_1, x_2, \cdots, x_T$，它保证

$$\forall g_1, \cdots, g_T \in \mathbb{R}^d, \epsilon - \sum_{t=1}^{T} \langle g_t, x_t \rangle \geq \phi(-\sum_{t=1}^{T} g_t)$$

当且仅当它保证：

$$\forall u \in \mathbb{R}^d, \mathrm{Regret}_T(u) = \sum_{t=1}^{T} \langle g_t, x_t - u \rangle \leq \phi^*(u) + \epsilon$$

证明.

从左到右:

$$
\begin{aligned}
\mathrm{Regret}_T(u) &= \sum_{t=1}^T \langle g_t, x_t - u\rangle \le -\sum_{t=1}^T \langle g_t, u\rangle - \phi(-\sum_{t=1}^T g_t) + \epsilon \\
&\le \sup_{\theta\in\mathbb{R}^d} \langle \theta, u\rangle - \phi(\theta) + \epsilon = \phi^*(u) + \epsilon
\end{aligned}
$$

从右到左:

$$
\begin{aligned}
-\sum_{t=1}^T \langle g_t, x_t\rangle &= -\sum_{t=1}^T \langle g_t, u\rangle - \sum_{t=1}^T \langle g_t, x_t - u\rangle \\
&= \sup_{u\in\mathbb{R}^d} -\sum_{t=1}^T \langle g_t, u\rangle - \sum_{t=1}^T \langle g_t, x_t - u\rangle \\
&\ge \sup_{u\in\mathbb{R}^d} -\sum_{t=1}^T \langle g_t, u\rangle - \phi^*(u) - \epsilon = \phi(-\sum_{t=1}^T g_t) - \epsilon
\end{aligned}
$$

$\square$

首先我们可以在一维 OLO 和 Coin-betting 问题之间进行转化：令 $c_t = -g_t$，于是就有 Coin-betting 算法给出：

$$\text{Wealth}_T = \epsilon + \sum_{t=1}^{T} x_t c_t = \epsilon - \sum_{t=1}^{T} x_t g_t$$

根据定理 4，如果 Wealth_T 能有一个被 $\sum_{t=1}^{T} c_t$ 的函数表示的下界，则对应 OLO 问题也有对应的下界，这个过程中我们只需要计算出 ϕ 的 Fenchel 共轭。
最后将一维 OCO 问题简化为一维 OLO 问题：

$$\text{Regret}_T(u) = \sum_{t=1}^{T} \ell_t(x_t) - \sum_{t=1}^{T} \ell_t(u) \leq \sum_{t=1}^{T} x_t g_t - \sum_{t=1}^{T} g_t u, g_t \in \partial \ell_t(x_t)$$

回忆 KT 算法，初始财富为 ϵ，每次赌注时 KT 算法赌上 $x_t = \beta_t \mathrm{Wealth}_{t-1}$ 的财富。令 $c_t = -g_t$，$g_t \in \partial \ell_t(x_t)$，且假设 ℓ_t 是 1-Lipschitz 的，所以我们得到

$$x_t = -\frac{\sum_{i=1}^{t-1} g_i}{t}(\epsilon - \sum_{i=1}^{t-1} g_i x_i)$$

Algorithm 5 KT OCO

Require: $\epsilon > 0$(1 到 $\sqrt{T}$ 之间的任何数)

1: **for** $t = 1$ to T **do**
2: 输出 $x_t = -\frac{\sum_{i=1}^{t-1} g_i}{t}(\epsilon - \sum_{i=1}^{t-1} g_i x_i)$
3: 接受 ℓ_t，支付 $\ell_t(x_t)$
4: 令 $g_t \in \partial \ell_t(x_t)$
5: **end for**

Contents

1. Cesa-Bianchi, Nicolo, and Gábor Lugosi. Prediction, learning, and games. Cambridge university press, 2006.
2. Orabona, Francesco, and Dávid Pál. "Coin betting and parameter-free online learning." Advances in Neural Information Processing Systems 29 (2016).
3. Francesco Orabona. "Introduction to Online Learning" https://parameterfree.com/lecture-notes-on-online-learning/
    - Online Mirror Descent I: Bregman version
    - Online Mirror Descent II: Regret and Mirror Version
    - Follow-The-Regularized-Leader I: Regret Equality
    - Parameter-free Online Learning I: Coin-Betting and 1-d OCO

谢谢!