Online Convex Optimization

UNIVERSITY *of* WASHINGTON

Convex surrogate loss functions

Previous section for the adversarial case suggested using multiplicative weights over the |H| hypotheses, which is completely intractable in practice.

And in the stochastic case we used $h_t \in \arg\min_{h \in \mathcal{H}} \sum_{s=1}^{t-1} \mathbf{1}\{h(x_s) \neq y_s\}$ which is also intractable to compute!

So it seems we have no practical algorithm! Solution: relax the objective.

Convex surrogate loss functions

Previous section for the adversarial case suggested using multiplicative weights over the |H| hypotheses, which is completely intractable in practice.

And in the stochastic case we used $h_t \in \arg\min_{h \in \mathcal{H}} \sum_{s=1}^{t-1} \mathbf{1}\{h(x_s) \neq y_s\}$ which is also intractable to compute!

So it seems we have no practical algorithm! Solution: relax the objective.

Instead of $$\max_{h \in \mathcal{H}} \sum_{t=1}^{T} \mathbf{1}\{h_t(x_t) \neq y_t\} - \mathbf{1}\{h(x_t) \neq y_t\}$$

We use $$\max_{h \in \mathcal{H}} \sum_{t=1}^{T} \ell(h_t, (x_t, y_t)) - \ell(h, (x_t, y_t))$$ with $\mathcal{H}$ convex

Example: Linear classification takes $\mathcal{H} \subset \mathbb{R}^d$ and $\ell(h, (x_t, y_t)) = \log(1 + \exp(-y_t h^\top x_t))$

Convex surrogate loss functions

Goal: $\max_{h \in \mathcal{H}} \sum_{t=1}^{T} \ell(h_t, (x_t, y_t)) - \ell(h, (x_t, y_t))$ with $\mathcal{H}$ convex

Online gradient descent

`Input:` $\mathcal{H} \subset \mathbb{R}^d$, convex loss function ℓ, step size $\eta > 0$

`Initialize:` Choose any $h_1 \in \mathcal{H}$

`for` $t = 1, 2, \ldots$

Player plays $h_t \in \mathcal{H}$

Adversary simultaneously reveals (x_t, y_t)

Player pays loss $\ell_t(h_t) := \ell(h_t, (x_t, y_t))$

Player updates $w_{t+1} = \Pi_{\mathcal{H}}(w_t - \eta \nabla_h \ell_t(h_t))$

Theorem Online gradient descent satisfies for any $h_* \in \mathcal{H}$

$$\sum_{t=1}^{T} \ell(h_t, (x_t, y_t)) - \ell(h_*, (x_t, y_t)) \leq \frac{\|h_*\|_2^2}{2\eta} + \frac{\eta}{2} \sum_{t=1}^{T} \|\nabla_h \ell_t(h_t)\|_2^2$$

if $\max_{h \in \mathcal{H}} \|h_*\|_2 \leq R$ and $\ell(\cdot)$ is G-Lipschitz then regret$\leq RB\sqrt{T}$

Proof

Theorem Online gradient descent satisfies for any $h_* \in \mathcal{H}$

$$\sum_{t=1}^T \ell(h_t, (x_t, y_t)) - \ell(h_*, (x_t, y_t)) \leq \frac{\|h_*\|_2^2}{2\eta} + \frac{\eta}{2} \sum_{t=1}^T \|\nabla_h \ell_t(h_t)\|_2^2$$

Proof

Theorem Online gradient descent satisfies for any $h_* \in \mathcal{H}$
$\sum_{t=1}^T \ell(h_t, (x_t, y_t)) - \ell(h_*, (x_t, y_t)) \leq \frac{\|h_*\|_2^2}{2\eta} + \frac{\eta}{2} \sum_{t=1}^T \|\nabla_h \ell_t(h_t)\|_2^2$

$$
\begin{aligned}
\|h_{t+1} - h_*\|_2^2 &= \|\Pi_{\mathcal{H}}(h_{t+1}) - \Pi_{\mathcal{H}}(h_*)\|_2^2 \\
&= \|\Pi_{\mathcal{H}}(h_t - \eta \nabla \ell_t(h_t)) - \Pi_{\mathcal{H}}(h_*)\|_2^2 \\
&\leq \|h_t - \eta \nabla \ell_t(h_t) - h_*\|_2^2 \\
&= \|h_t - h_*\|_2^2 - 2\eta \nabla \ell_t(h_t)^\top (h_t - h_*) + \eta^2 \|\nabla \ell_t(h_t)\|_2^2 \\
&\leq \|h_t - h_*\|_2^2 - 2\eta \big(\ell_t(h_t) - \ell_t(h_*)\big) + \eta^2 \|\nabla \ell_t(h_t)\|_2^2
\end{aligned}
$$

$$
\begin{aligned}
\sum_{t=1}^T \big(\ell_t(h_t) - \ell_t(h_*)\big) &\leq \sum_{t=1}^T \frac{\|h_t - h_*\|_2^2 - \|h_{t+1} - h_*\|_2^2}{2\eta} + \sum_{t=1}^T \frac{\eta}{2} \|\nabla \ell_t(h_t)\|_2^2 \\
&\leq \frac{\|h_1 - h_*\|_2^2}{2\eta} + \sum_{t=1}^T \frac{\eta}{2} \|\nabla \ell_t(h_t)\|_2^2
\end{aligned}
$$

Universal Portfolio Optimization

Universal Portfolio Optimization

Given a collection of stocks, let the ith stock have price $S_t(i)$ over time t.
You start with v_1 dollars and fractionally invest it into d stocks according to $p_1 \in \triangle_d$.

Your portfolio at time 2 is worth $v_2 := \sum_{i=1}^{d} v_1 p_1(i) r_1(i) = v_1 \langle p_1, r_1 \rangle$ dollars

where $r_t(i) = \dfrac{S_{t+1}(i)}{S_t(i)} = \dfrac{\text{price of GOOG at time t+1}}{\text{price of GOOG at time t}}$.

Universal Portfolio Optimization

Given a collection of stocks, let the ith stock have price $S_t(i)$ over time t.
You start with v_1 dollars and fractionally invest it into d stocks according to $p_1 \in \triangle_d$.

Your portfolio at time 2 is worth $v_2 := \sum_{i=1}^{d} v_1 p_1(i) r_1(i) = v_1 \langle p_1, r_1 \rangle$ dollars

where $r_t(i) = \frac{S_{t+1}(i)}{S_t(i)} = \frac{\text{price of GOOG at time t+1}}{\text{price of GOOG at time t}}$.

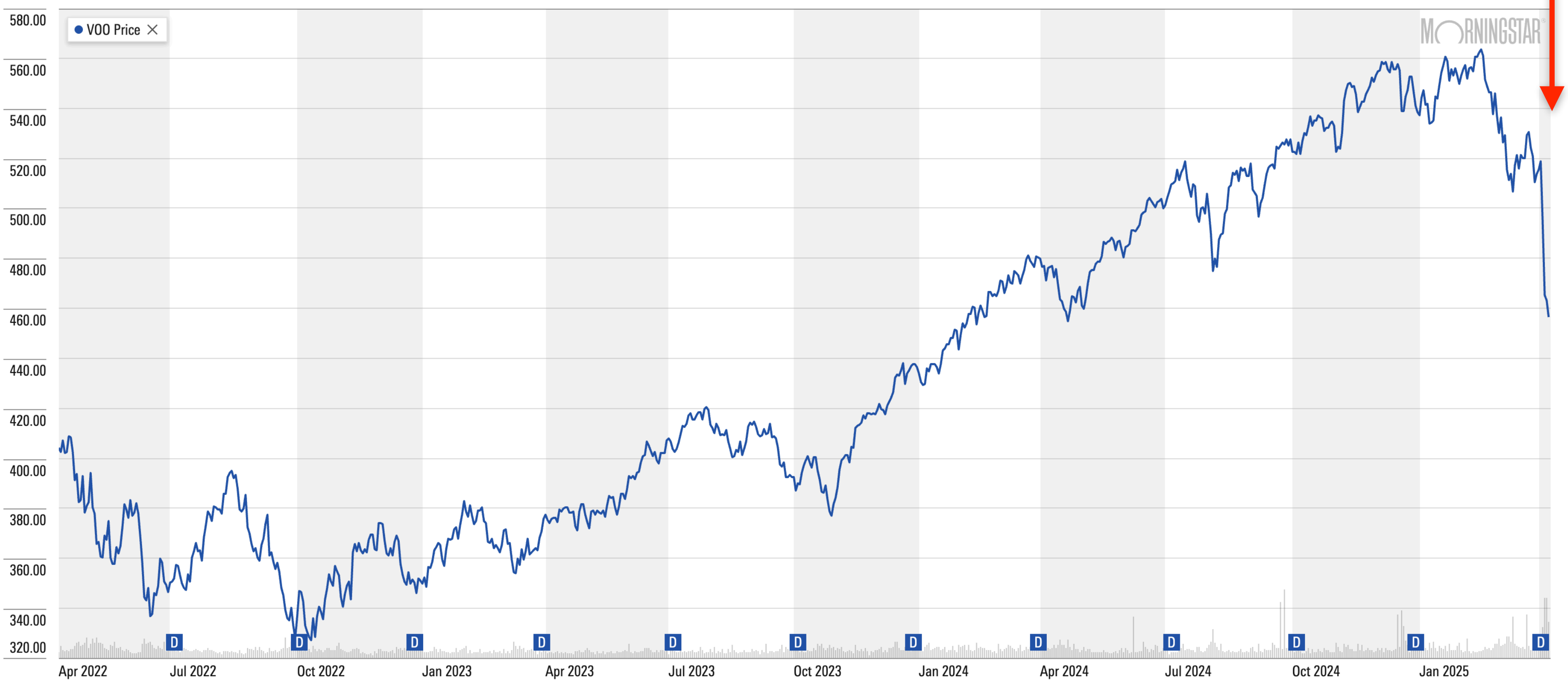
Classical Portfolio Theory (Markowitz 1952): Assume returns $r_t \in \mathbb{R}^n_+$ are IID with mean $\mu = \mathbb{E}[r_t]$ and covariance $\Sigma = \mathbb{E}[(r_t - \mu)(r_t - \mu)^\top]$. The for a return target $\bar{r} \geq 0$ solve

$$\min_{p \in \triangle_d} p^\top \Sigma p \quad \text{subject to} \quad p^\top \mu \geq \bar{r}$$

In practice, estimate μ, Σ from data. What could possibly go wrong?

Universal Portfolio Optimization

Trump administration announces Tariffs



Returns are not an IID stochastic random walk!

Can we model the stock market as an online learning problem and develop an algorithm that is robust to even adversarial returns?

Universal Portfolio Optimization

You start with v_1 dollars and fractionally invest it into d stocks according to $p_1 \in \triangle_d$

Your portfolio at time 2 is worth $v_2 := \sum_{i=1}^{d} v_1 p_1(i) r_1(i) = v_1 \langle p_1, r_1 \rangle$ dollars

where $r_t(i) = \dfrac{\text{price of GOOG at time t+1}}{\text{price of GOOG at time t}}$.

After T times your portfolio is worth $v_T = v_1 \prod_{t=1}^{T-1} \langle p_t, r_t \rangle$.

Universal Portfolio Optimization

You start with v_1 dollars and fractionally invest it into d stocks according to $p_1 \in \triangle_d$

Your portfolio at time 2 is worth $v_2 := \sum_{i=1}^{d} v_1 p_1(i) r_1(i) = v_1 \langle p_1, r_1 \rangle$ dollars

where $r_t(i) = \dfrac{\text{price of GOOG at time t+1}}{\text{price of GOOG at time t}}$.

After T times your portfolio is worth $v_T = v_1 \prod_{t=1}^{T-1} \langle p_t, r_t \rangle$.

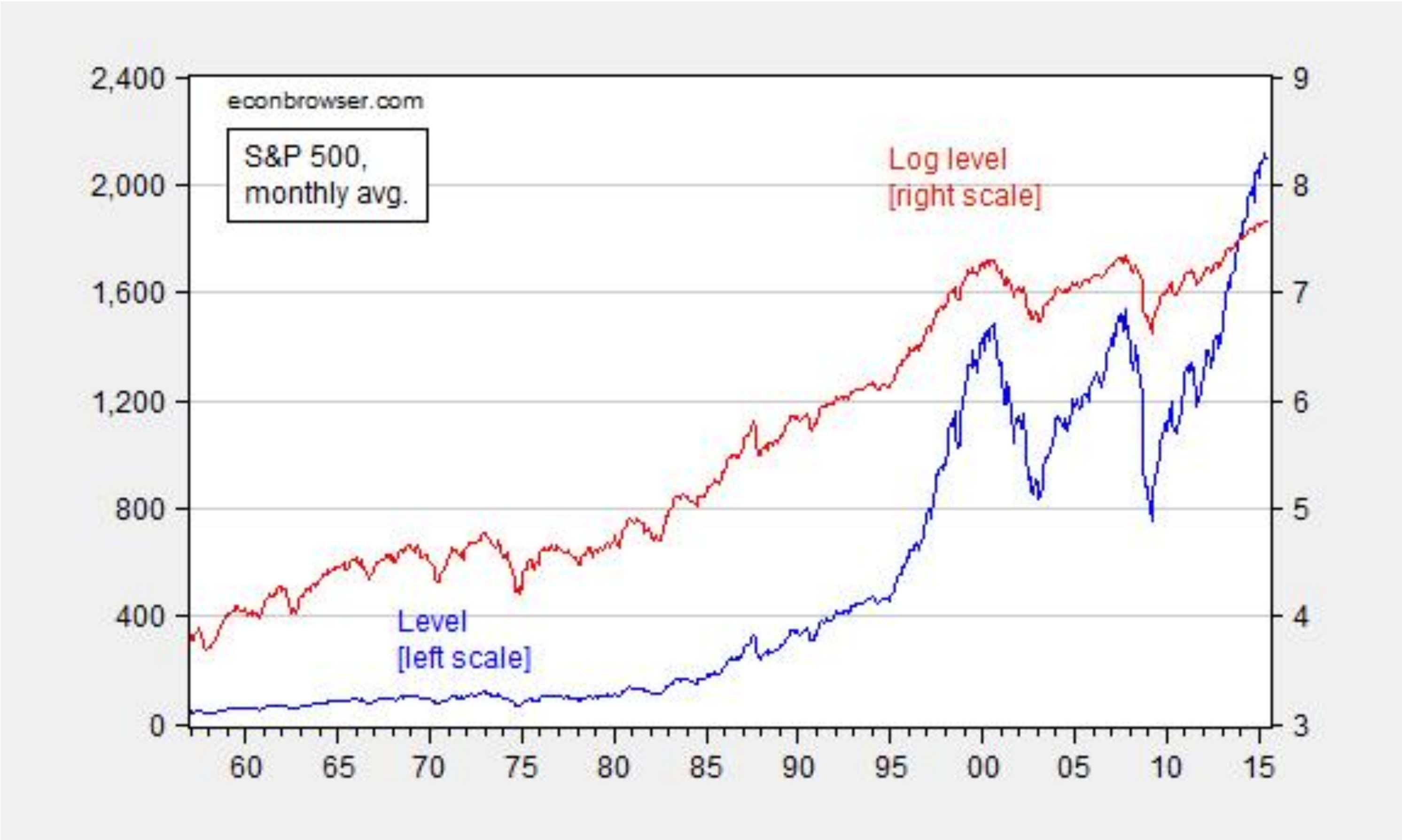
Goal: Maximize your return $\frac{v_T}{v_1}$, equivalent to $\log\left(\frac{v_T}{v_1}\right) = \sum_{t=1}^{T-1} \log \langle p_t, r_t \rangle$

$$\text{Regret} = \max_{p \in \triangle_d} \sum_{t=1}^{T-1} \log \langle p, r_t \rangle - \sum_{t=1}^{T-1} \log \langle p_t, r_t \rangle$$

Universal Portfolio Optimization

$$\text{Regret} = \max_{p \in \triangle_d} \sum_{t=1}^{T-1} \log\langle p, r_t \rangle - \sum_{t=1}^{T-1} \log\langle p_t, r_t \rangle$$

The SP500 (VOO) is an index that weights 500 stocks by their market capitalization.
An alternative index (RSP) weights these 500 stocks uniformly $p = (\frac{1}{500}, \ldots, \frac{1}{500})$.

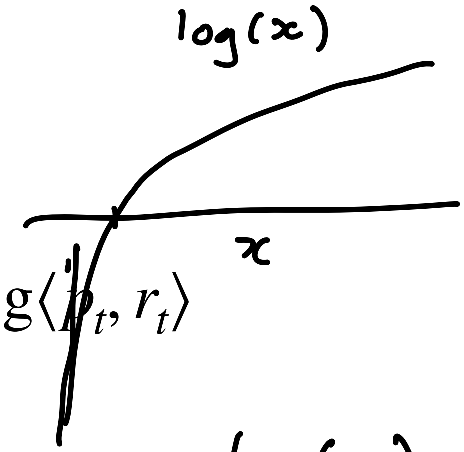


Adaptive Regret Bounds

$$\sum_t \ell(p_t, z_t) - \sum_t \ell(u_t, z_t) \leq \sqrt{\left(\sum_t \|u_t - u_{t-1}\|_1 + 1\right) T}$$

$$\sum_t \ell(p_t, z_t) - \sum_t \ell(p, z_t) \leq \sqrt{\left(\sum_t \ell(p, z_t) + 1\right)}$$

Universal Portfolio Optimization



$$\text{Regret} = \max_{p \in \triangle_d} \sum_{t=1}^{T-1} \log\langle p, r_t \rangle - \sum_{t=1}^{T-1} \log\langle p_t, r_t \rangle$$

$-\log(x)$ is convex

for $t = 1, 2, \ldots$

- Player picks $p_t \in \triangle_d$
- Adversary simultaneously reveals $r_t \in \mathbb{R}_+^d$
- Player pays loss $\ell_t(p_t) = -\log\langle p_t, r_t \rangle$

Exponential weights algorithm

Initialize: $w_1 = (1, \ldots, 1) \in \mathbb{R}^d$

for $t = 1, 2, \ldots$

- Player plays $p_t(i) = w_t(i) / \sum_{j=1}^{d} w_t(j)$
- Adversary simultaneously reveals convex loss $\ell_t(\cdot)$
- Player pays loss $\ell_t(p_t)$
- Player updates weights $w_{t+1}(i) = w_t(i) \exp(-\eta \ell_t(\mathbf{e}_i))$

$i^{th} \rightarrow \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \in \mathbb{R}^d$

Universal Portfolio Optimization

$$\prod_{t=1}^{T} \langle p_t, r_t \rangle \geq \frac{1}{d} \max_i \prod_{t=1}^{T} \langle e_i, r_t \rangle$$

$$\text{Regret} = \max_{p \in \triangle_d} \sum_{t=1}^{T-1} \log\langle p, r_t \rangle - \sum_{t=1}^{T-1} \log\langle p_t, r_t \rangle$$

Competes with the single best stock in hindsight!

Theorem: With $\eta = 1$ and $l_t(p) = -\log\langle p, r_t \rangle$, $\max_{i \in [d]} \sum_{t=1}^{T-1} \log\langle \mathbf{e}_i, r_t \rangle - \log\langle p_t, r_t \rangle \leq \log(d)$

Exponential weights algorithm

Initialize: $w_1 = (1, \ldots, 1) \in \mathbb{R}^d$

for $t = 1, 2, \ldots$

Player plays $p_t(i) = w_t(i) / \sum_{j=1}^{d} w_t(j)$

Adversary simultaneously reveals convex loss $\ell_t(\cdot)$

Player pays loss $\ell_t(p_t)$

Player updates weights $w_{t+1}(i) = w_t(i) \exp(-\eta \ell_t(\mathbf{e}_i))$

Proof

Theorem: With $\eta = 1$ and $l_t(p) = -\log\langle p, r_t\rangle$, $\max_{i\in[d]} \sum_{t=1}^{T-1} \log\langle \mathbf{e}_i, r_t\rangle - \log\langle p_t, r_t\rangle \leq \log(d)$

$$\begin{aligned}
\log \frac{W_{T+1}}{W_1} &= \sum_{t=1}^{T} \log \frac{W_{t+1}}{W_t} \\
&= \sum_{t=1}^{T} \log\left(\sum_{i=1}^{d} \frac{w_{t+1}(i)}{W_t}\right) \\
&= \sum_{t=1}^{T} \log\left(\sum_{i=1}^{d} \frac{w_t(i)\exp(-\eta \ell_t(\mathbf{e}_i))}{W_t}\right) \\
&= \sum_{t=1}^{T} \log\left(\sum_{i=1}^{d} p_t(i)\exp(-\eta \ell_t(\mathbf{e}_i))\right) \\
&= \sum_{t=1}^{T} \log\left(\sum_{i=1}^{d} p_t(i)\exp(\log\langle \mathbf{e}_i, r_t\rangle)\right) \\
&= \sum_{t=1}^{T} \log\langle p_t, r_t\rangle
\end{aligned}$$

$$\begin{aligned}
\log \frac{W_{T+1}}{W_1} &\geq \log \frac{w_{T+1}(i)}{W_1} \\
&= -\log(d) + \log\left(\prod_{t=1}^{T} \exp(-\eta \ell_t(\mathbf{e}_i))\right) \\
&= -\log(d) - \sum_{t=1}^{T} \eta \ell_t(\mathbf{e}_i) \\
&= -\log(d) + \sum_{t=1}^{T} \log\langle \mathbf{e}_i, r_t\rangle
\end{aligned}$$

$$\implies \max_{i\in[d]} \sum_{t=1}^{T} \log\langle \mathbf{e}_i, r_t\rangle - \log\langle p_t, r_t\rangle \leq \log(d)$$

Universal Portfolio Optimization

$$\text{Regret} = \max_{p \in \triangle_d} \sum_{t=1}^{T-1} \log\langle p, r_t \rangle - \sum_{t=1}^{T-1} \log\langle p_t, r_t \rangle$$

Competes with the single best stock in hindsight!

Theorem: With $\eta = 1$ and $l_t(p) = -\log\langle p, r_t \rangle$, $\max_{i \in [d]} \sum_{t=1}^{T-1} \log\langle \mathbf{e}_i, r_t \rangle - \log\langle p_t, r_t \rangle \leq \log(d)$

Is competing against <u>single best</u> stock a good benchmark? Consider just 2 stocks:

$$r_t(1) = (2, \tfrac{1}{2}, 2, \tfrac{1}{2}, 2, \tfrac{1}{2}, \ldots)$$

$$r_t(2) = (\tfrac{1}{2}, 2, \tfrac{1}{2}, 2, \tfrac{1}{2}, 2\ldots)$$

$$\prod_{t=1}^{T} \langle \mathbf{e}_i, r_t \rangle = 1 \qquad \prod_{t=1}^{T} \langle \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}, r_t \rangle = \left((\tfrac{1}{2})^2 + 1 \right)^{T/2}$$

How do we compete with any $p \in \triangle_d$?

Example due to Elad Hazan's *Introduction to Online Convex Optimization*

Online Gradient Descent

For convex loss ℓ, convex domain $\mathcal{A}$ w/ $\sup_{a \in \mathcal{A}} \|a\|_2 \le R$, $\sup_{a,z} \|\nabla_a \ell(a,z)\|_2 \le G$

$$\sum_{t=1}^{T} \ell(a_t, z_t) - \ell(a, z_t) \le RG\sqrt{T}$$

$$\ell(a,z) = -\log(\langle a, z \rangle)$$

$$\nabla_a \ell(a,z) = -\frac{1}{\langle a, z \rangle} \cdot z$$

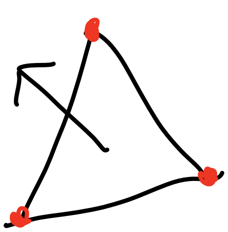
$$\sum_{t=1}^{T} \log(\langle p, r_t \rangle) - \log(\langle p_t, r_t \rangle) \le \sqrt{dT}$$

$$\max_{a^* \in \Delta} \sum_{t=1}^{T} \ell(a_t, z_t) - \ell(a^*, z_t) \le \max_{a^* \in \Delta} \sum_{t=1}^{T} \nabla \ell(a_t, z_t)^T (a_t - a^*)$$

$$= \max_{i \in [d]} \sum_{t=1}^{T} \nabla \ell(a_t, z_t)^T a_t - \nabla \ell(a_t, z_t)^T e_i$$

$$[\tilde{z}_t]_i = \nabla \ell(a_t, z_t)^T e_i$$

$$= \max_{i \in [d]} \sum_{t=1}^{T} \langle a_t, \tilde{z}_t \rangle - \langle e_i, \tilde{z}_t \rangle$$



$$a^* = \arg\min_{a \in \Delta_d} \sum_{t=1}^{T} \nabla \ell(a_t, z_t)^T a$$

$$\Longrightarrow \sum_{t=1}^{T} \log(\langle p, r_t \rangle) - \log(\langle p_t, r_t \rangle) \le \sqrt{\log(d) T}$$

← Expon Weights w/ gradient losses

Continuous Exponential weights

UNIVERSITY of WASHINGTON

Continuous Exponential Weights

Fix a convex set $\mathscr{A}$ and a convex loss function $l(\cdot, z) : \mathscr{A} \to \mathbb{R}$ for each $z \in \mathscr{Z}$.

Theorem:

For any $\eta > 0$ and $l(\cdot, \cdot) \in [0,1]$ we have $\max_{a \in \mathscr{A}} \sum_{t=1}^{T-1} l(a_t, z_t) - l(a, z_t) \leq \frac{d \log(T)}{\eta} + \frac{\eta T}{8} + 1$

Continuous Exponential weights algorithm

$\leq \sqrt{d T \log(T)}$

Initialize: $w_1(a) = 1$ for all $a \in \mathcal{A}$

for $t = 1, 2, \ldots$

Player plays $a_t = \mathbb{E}_{A \sim p_t}[A]$ where $p_t(a) = \frac{w_t(a)}{\int_{a \in \mathcal{A}} w_t(a) da}$

Adversary simultaneously reveals z_t and convex loss $\ell(\cdot, z_t)$

Player pays loss $\ell(a_t, z_t)$

Player updates weights $w_{t+1}(a) = w_t(a) \exp(-\eta \ell(a, z_t))$

Continuous Exponential Weights

Let $\gamma = \frac{1}{T}$, $a^* \in \arg\min_{a\in\mathcal{A}} \sum_{t=1}^{T} \ell(a, z_t)$, $\mathcal{N}_\gamma\{(1-\gamma)a^* + \gamma a, a \in \mathcal{A}\}$

$$\begin{aligned}
\log \frac{W_{T+1}}{W_1} &= \log\left(\frac{\int_{a\in\mathcal{A}} w_{T+1}(a)da}{\int_{a\in\mathcal{A}} 1da}\right) \\
&\geq \log\left(\frac{\int_{a\in\mathcal{N}_\gamma} w_{T+1}(a)da}{\int_{a\in\mathcal{A}} 1da}\right) \\
&= \log\left(\frac{\int_{a\in\mathcal{N}_\gamma} \exp\left(-\eta\sum_{t=1}^{T}\ell(a,z_t)\right)da}{\int_{a\in\mathcal{A}} 1da}\right) \\
&= \log\left(\frac{\int_{a\in\gamma\mathcal{A}} \exp\left(-\eta\sum_{t=1}^{T}\ell((1-\gamma)a^*+a,z_t)\right)da}{\int_{a\in\mathcal{A}} 1da}\right) \\
&\geq \log\left(\frac{\int_{a\in\mathcal{A}} \exp\left(-\eta\sum_{t=1}^{T}\ell\left((1-\gamma)a^*+\gamma a,z_t\right)\right)\gamma^d da}{\int_{a\in\mathcal{A}} 1da}\right) \\
&\geq \log\left(\frac{\int_{a\in\mathcal{A}} \exp\left(-\eta\sum_{t=1}^{T}\left(\ell(a^*,z_t)+\gamma\ell(a,z_t)\right)\right)\gamma^d da}{\int_{a\in\mathcal{A}} 1da}\right) \\
&= d\log\gamma - \eta\sum_{t=1}^{T}\ell(a^*,z_t) - \eta\gamma T
\end{aligned}$$

$$\begin{aligned}
\log \frac{W_{t+1}}{W_t} &= \log\left(\int_{\mathcal{A}} \frac{w_t(a)}{W_t}\exp(-\eta\ell(a,z_t))da\right) \\
&= \log\left(\mathbb{E}\exp(-\eta\ell(A,z_t))\right) \text{ where } \mathbb{P}(A=a) = \frac{w_t(a)}{W_t} \\
&\leq -\eta\mathbb{E}\ell(A,z_t) + \frac{\eta^2}{8} \text{ (Hoeffding's lemma)} \\
&\leq -\eta\ell(\mathbb{E}A,z_t) + \frac{\eta^2}{8} \text{ (Jensen's inequality)} \\
&= -\eta\ell(a_t,z_t) + \frac{\eta^2}{8}.
\end{aligned}$$

Proof due to Sebastien Bubeck's *Introduction to Online Optimization*

Continuous Exponential Weights

Fix a convex set $\mathscr{A}$ and a convex loss function $l(\cdot\,,z) : \mathscr{A} \to \mathbb{R}$ for each $z \in \mathscr{Z}$.

Theorem:

With $\eta = 1$ and $l(a,z) = -\log\langle a, z\rangle$, $\max_{p \in \triangle_d} \sum_{t=1}^{T-1} \log\langle p, r_t\rangle - \log\langle p_t, r_t\rangle \leq 1 + d\log(T)$

Continuous Exponential weights algorithm

Initialize: $w_1(a) = 1$ for all $a \in \mathcal{A}$

for $t = 1, 2, \ldots$

Player plays $a_t = \mathbb{E}_{A \sim p_t}[A]$ where $p_t(a) = \frac{w_t(a)}{\int_{a \in \mathcal{A}} w_t(a)da}$

Adversary simultaneously reveals z_t and convex loss $\ell(\cdot, z_t)$

Player pays loss $\ell(a_t, z_t)$

Player updates weights $w_{t+1}(a) = w_t(a)\exp(-\eta\ell(a, z_t))$

Continuous Exponential Weights

Let $\gamma = \frac{1}{T}$, $a^* \in \arg\min_{a \in \mathcal{A}} \sum_{t=1}^T \ell(a, z_t)$, $\mathcal{N}_\gamma\{(1-\gamma)a^* + \gamma a, a \in \mathcal{A}\}$

$$\log \frac{W_{T+1}}{W_1} = \log\left(\frac{\int_{a\in\mathcal{A}} w_{T+1}(a)da}{\int_{a\in\mathcal{A}} 1da}\right)$$
$$= \log\left(\frac{\int_{a\in\mathcal{A}} \exp(-\eta\sum_{t=1}^T \ell(a, z_t))da}{\int_{a\in\mathcal{A}} 1da}\right)$$
$$= \log\left(\frac{\int_{a\in\mathcal{A}} \prod_{t=1}^T \langle a, z_t\rangle da}{\int_{a\in\mathcal{A}} 1da}\right)$$
$$= \log\left(\frac{\int_{a\in\mathcal{A}} \prod_{t=1}^T \langle a, z_t\rangle da}{\int_{a\in\mathcal{A}} 1da}\right)$$
$$\geq \log\left(\frac{\int_{a\in\mathcal{N}_\gamma} \prod_{t=1}^T \langle a, z_t\rangle da}{\int_{a\in\mathcal{A}} 1da}\right)$$
$$\geq \log\left(\frac{\int_{a\in\gamma\mathcal{A}} \prod_{t=1}^T \langle (1-\gamma)a^* + a, z_t\rangle da}{\int_{a\in\mathcal{A}} 1da}\right)$$
$$= \log\left(\frac{\int_{a\in\mathcal{A}} \gamma^d \prod_{t=1}^T \langle (1-\gamma)a^* + \gamma a, z_t\rangle da}{\int_{a\in\mathcal{A}} 1da}\right)$$
$$\geq \log\left(\frac{\int_{a\in\mathcal{A}} \gamma^d \left((1-\gamma)\prod_{t=1}^T \langle a^*, z_t\rangle + \gamma \prod_{t=1}^T \langle a, z_t\rangle\right) da}{\int_{a\in\mathcal{A}} 1da}\right)$$
$$\geq -d\log(1/\gamma) + \log(1-\gamma) + \sum_{t=1}^T \langle a^*, z_t\rangle$$

$$\log \frac{W_{T+1}}{W_1} = \sum_{t=1}^T \log \frac{W_{t+1}}{W_t}$$
$$= \sum_{t=1}^T \log\left(\int_{\mathcal{A}} \frac{w_t(a)}{W_t} \exp(-\eta\ell(a, z_t))da\right)$$
$$= \sum_{t=1}^T \log\left(\mathbb{E}_{A\sim p_t}[\exp(-\eta\ell(A, z_t))]\right)$$
$$= \sum_{t=1}^T \log\langle a_t, z_t\rangle$$

Proof due to Sebastien Bubeck's *Introduction to Online Optimization*

Minimax Analysis

Minimax Analysis

Consider a two-player zero-sum game, like matching pennies:

A \ Z	0	1	2
0	0	1	-1
1	-1	0	1
	1	-1	0

Minimax Analysis

The first player wants to minimize, the second player wants to maximize.

They achieve $\min_{a \in \mathcal{A}} \max_{z \in \mathcal{Z}} L(a, z)$

But this fails to capture the simultaneous play. Rewrite as

$$\min_{q \in \triangle_{\mathcal{A}}} \max_{z \in \mathcal{Z}} \mathbb{E}_{a \sim q}[L(a, z)]$$

Minimax Analysis

The first player wants to minimize, the second player wants to maximize.

$$\min_{q_1, q_2 \in \triangle_{\mathcal{A}}} \max_{z_1, z_2 \in \mathcal{Z}} \mathbb{E}_{a_1 \sim q(\cdot)} \mathbb{E}_{a_2 \sim q(z_1, \cdot)} [L(a_1, z_1, a_2, z_2)]$$

We can think of online learning as a two-player game

$$L(a_1, z_1, a_2, z_2) = \sum_{t=1}^{2} \ell(a_t, z_t) - \inf_{a^* \in \mathcal{A}} \sum_{t=1}^{2} \ell(a^*, z_t)$$

Minimax Analysis

The first player wants to minimize, the second player wants to maximize.

$$\mathscr{V}^{obliv}(\mathscr{A}, \mathscr{Z}) := \underbrace{\min_{q_1,\ldots,q_T}} \max_{z_1,\ldots,z_T \in \mathscr{Z}} \mathbb{E}_{a_1\sim q_1,\ldots,a_T\sim q_T}\left[\sum_{t=1}^{T} \ell(a_t, z_t) - \inf_{a^*} \sum_{t=1}^{T} \ell(a^*, z_t)\right] \lesssim 2R(A,Z)$$

$$\mathfrak{R}(\mathscr{A}, \mathscr{Z}) = \sup_{z_1,\ldots,z_T \in \mathscr{Z}} \mathbb{E}_\epsilon \left[\sup_{a\in\mathscr{A}} \sum_{t=1}^{T} \ell(a, z_t)\epsilon_t\right]$$

$\epsilon_t \in \{-1, 1\}$ with equal probability

If $\ell(a,z) = \langle a, z \rangle$

$$R(A,Z) \leq V \leq 2R(A,Z)$$