**EECS598: Prediction and Learning: It's Only a Game** Fall 2013

# Lecture 13: Universal Portfolios Review and Online Convex Optimization

*Prof. Jacob Abernethy* *Scribe: Huanan Zhang*

## Announcements

- Homework 2 due in one week.
- It is time to think about the projects. There will be discussion on early November.

## 13.1 Universal Portfolios Review

Consider the problem with $n$ stocks and $\mathbf{b}^t \in (0, \infty)^n$ as the "Price relatives" where $b_i^t = \frac{\text{Price}_{t+i}(i)}{\text{Price}_t(i)}$. Our decision is the portfolio $\mathbf{w} \in \Delta_n$, which means we have $w_i$ fraction of wealth invest in stock $i$. From time $t$ to time $t+1$, the wealth grows by $\mathbf{w}\mathbf{b}^t = \sum_i w_i b_i^t$. Without loss of generality, we can consider the starting wealth $c = 1$.

### 13.1.1 Constant Rebalanced Portfolio

We focus on a certain class of policies called the constant rebalanced portfolio(CRP), where the $\mathbf{w}$ is constant for each day. Then the total wealth of a CRP policy after $T$ days would be $V_T(w) = \prod_{t=1}^{T}(\mathbf{w} \cdot \mathbf{b}^t)$.

### 13.1.2 Cover's Algorithm - UCRP

To compete with the best CRP under all scenarios, we consider a Universal CRP(UCRP), where the money is invested in *all* CRP's evenly. In another word, all CRP got an infinitesimal amount of weight. We try to develop a regret bound of the UCRP to the best CRP.

**Proof ideas:** Define the ball around $w_0$ as $\text{Ball}_\epsilon(w_0) = \{w : (1-\epsilon)w_0 + \epsilon v, \text{ where } v \in \Delta_n\}$. By this definition, we can see the following two properties:

- $\text{Vol}[\text{Ball}_\epsilon(w_0)] = \epsilon^{n-1}\text{Vol}(\Delta_n)$
- $V_T(w) \geq V_T(w_0)(1-\epsilon)^T$

With these, we are ready to show the regret bound of the UCRP to the best CRP using these two properities.

**Proof:**

$$
\begin{aligned}
\text{Wealth(UCRP)} &= \frac{1}{\text{Vol}(\Delta_n)} \int_{w \in \Delta n} V_T(w) d\mu(w) \\
&\geq \frac{1}{\text{Vol}(\Delta_n)} \int_{w \in \text{Ball}_\epsilon(w^*)} V_T(w) d\mu(w) \\
&\geq \frac{1}{\text{Vol}(\Delta_n)} \int_{w \in \text{Ball}_\epsilon(w^*)} d\mu(w) V_T(w^*)(1-\epsilon)^T \\
&= \epsilon^{n-1}(1-\epsilon)^T V_T(w^*)
\end{aligned}
$$

It is clear that we need to tune the parameter $\epsilon$ to achieve a good bound. The best bound could be a little bit complex to achieve and we just let $\epsilon = \frac{1}{T}$ and it can give a good enough result.
To derive the bound, we let $\epsilon = \frac{1}{T}$ and look at the log-wealth of UCRP and CRP:

$$
\begin{aligned}
\text{Logwealth(UCRP)} &\geq -(n-1)\log\frac{1}{\epsilon} + T\log(1-\epsilon) + \log V_T(w^*) \\
&= -(n-1)\log T + O(1) + \log V_T(w^*)
\end{aligned}
$$

This gives the bound of UCRP. □

The properties of UCRP:

- The regret is $O(n \log T)$in log space.
- UCRP only very mild assumption on price relatives(could be very large or small).
- UCRP requires no tuning(although the analysis part need some tuning).
- Here we ignore the transition cost, but it is easy to be added.([BK99])
- UCRP is not efficient to implement. (See [KV03] for details)
- Lots of work within Online Convex Optimization(OCO) framework on efficient algorithm with the same regret bound.

## 13.2 Online Convex Optimization(OCO)

The general framework in OCO looks like this:

- We have a decision space $X \in \mathbb{R}^n for t = 1, ..., T$. Assume that $X$ is convex, closed and bounded.
- Player chooses $x_t \in X$, then nature chooses $l_t : X \xrightarrow{\text{convex}} \mathbb{R}$,
- Player suffers $l_t(x_t)$ and observe $l_t$, (we will go to bandit setting where player only observe $l_t(x_t)$ soon!)

We want to minimize the regret $\sum_{t=1}^T l_t(x_t) - \min_{x^* \in X} \sum_{t=1}^T l_t(x^*)$.

There are a variety of problems fall into the category of OCO, here are some examples:

| Setting | Action | Data | Loss |
|---|---|---|---|
| Prediction with expert advice | $\mathbf{w} \in \Delta_n$ | $(\mathbf{f}, y) \in (0,1)^{n+1}$ | $l(\mathbf{w}, \mathbf{f}, y)$ |
| Action setting | $\mathbf{P} \in \Delta_n$ | $\mathbf{l}_t \in [0,1]^n$ | $\mathbf{pl}^t$ |
| Linear Prediction (pattern recognition) | $\mathbf{w} \in \mathbb{R}_n$ | $(\mathbf{x}, y) \in \mathbb{R} \times \{-1, 1\}$ | $max(0, y\mathbf{wx})$ |
| Portfolio Selection | $\mathbf{w} \in \Delta_n$ | $\mathbf{b}^t \in (0, \infty)^n$ | $-\log \mathbf{wb}$ |

### 13.2.1 Online Gradient Decent (OGD)[Zin03]

First we define the projection function $\mathrm{Proj}_x(z) = \mathrm{argmin}_{x \in X} \|x - z\|_2$. Then we know: 1) $\mathrm{Proj}_x(z) = z$ iff $z \in X$. 2) $\|x - z\|_2 \geq \left\|x - \mathrm{Proj}_x(z)\right\|_2$.

The main idea of OGD method is the following:

- Choose an initial $x_0$ arbitrarily in $X$.
- For each $t = 1, \ldots, T$
  - $z_{t+1} = x_t - \eta \nabla l_t(x_t)$,
  - $x_{t+1} = \mathrm{Proj}_X(z_{t+1})$.

It can be shown that if $\nabla_t := \nabla l_t(x_t)$ has norm $\leq G$, and $D := \max_{x,y \in X} \left\|x - y\right\|_2$. Then regret(ODG)$\leq \eta \frac{G^2 T}{2} + \frac{D^2}{2\eta}$. Apply tuning to $\eta$ and let $\eta = \frac{D}{G\sqrt{T}}$, we can have regret(ODG)$\leq GD\sqrt{T}$.

## References

[BK99] Avrim Blum and Adam Kalai. Universal portfolios with and without transaction costs. *Mach. Learn.*, 35(3):193–205, June 1999.

[KV03] Adam Kalai and Santosh Vempala. Efficient algorithms for universal portfolios. *J. Mach. Learn. Res.*, 3:423–440, March 2003.

[Zin03] Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. 2003.