THEORETICAL MACHINE LEARNING COS 511

LECTURE #9 MARCH 1, 2016

LECTURER: ELAD HAZAN SCRIBE: VLADIMIR FEINBERG

In the following, let Δ_n be the n-dimensional simplex. Let $\boldsymbol{\delta}_i$ be the i-th standard vector. For vectors $\mathbf{x}$, $x_i = \boldsymbol{\delta}_i^\top \mathbf{x}$ denotes the i-th element. OCO stands for online convex optimization. Let $[n] = \mathbb{Z} \cap [1, n]$.

1 Betting with an Edge

Suppose we play a game: we start with \$1, and we may wager some proportion of it by betting that a fair coin comes up heads - if it does, we win three times our wager, else we lose our bet. Let us work under the assumption that we want to maximize our expected wealth.

1.1 One Iteration

With initial wealth W_0, if we bet a proportion $p \in \Delta_1$ of our wealth, in case of loss we have $(1-p)W_0$ and in case of a win we have $(1-p)W_0 + 3pW_0$. Then the expectation for our wealth after a single game W_1 is given by $\mathbb{E}(W_1) = (1-p)W_0 + 3pW_0/2 = (1+p/2)W_0$. To maximize our expected returns, it is clear that we should wager all our money, $p = 1$.

1.2 Over Time

Betting over several (T) iterations, our intuitions tell us that this fact may no longer be true - the probability that we leave with any positive amount of money if $p = 1$ is 2^{-T}.

Let us suppose at every time t we indeed bet a constant proportion p of our wealth. We may give a heuristic proof justification to this strategy by recursion; suppose an optimal strategy modifies p_t over time; this is independent of the initial amount of wealth since we always make returns as a factor of our starting capital. After one iteration of the game played with p_1, we find ourselves in a situation where we must maximize our capital yet again. It must be that re-applying the optimal strategy from $t = 2$ must also yield the best result at time $T + 1$. For any sensible metric of evaluating our performance f and large T, $f(W_T)$ is maximized when $f(W_{T+1})$ is; in turn we find that playing with probability p_1 every time is sure to be just as good as the time-dependent strategy p_t.

Define W_t to be the rv for our wealth after game t. We set $W_0 = 1$, and have by the same reasoning as before $\mathbb{E}\left(W_t | W_{t-1}\right) = (1+p/2)W_{t-1}$. We see at this point that inductively $\mathbb{E}(W_T) = (1+p/2)^T W_0$. This suggests we should set $p = 1$ as before. However, if we reframe the problem, we find a surprising result.

1.3 Kelly Criterion

Theorem 1.1. *If W_t^p is the rv for the game played with probability p, then $\exists p^*$ such that as $T \to \infty$, $\mathbb{P}\left(W_T^{p^*} > W_T^p\right) = 1$ for any other p.*

We note that $p^* < 1$ since it is less than any other strategy with probability 1 as $T \to \infty$ and $W_T^p > 0$ for any $p < 1$. Also, $p^* > 0$ by Markov's inequality: $\mathbb{E}(W_T^0 / W_T^p) = (1 + p/2)^{-T}$. Kelly (or rather, Bernoulli, earlier on) showed that such a p^* satisfies:

$$p^* = \underset{p\in\Delta_1}{\operatorname{argmax}}\, \mathbb{E}\left(\log \frac{W_T^p}{W_0}\right) = \underset{p\in\Delta_1}{\operatorname{argmax}}\, \mathbb{E}\left(\frac{1}{T}\log \prod_{t=1}^{T} \frac{W_t^p}{W_{t-1}^p}\right) = \underset{p\in\Delta_1}{\operatorname{argmax}}\, \mathbb{E}\left(\log G\right)$$

Where we notice G is the geometric mean of the returns.

2 Portfolio Extension

We can extend this to n independent bets to simulate the behaviour of a portfolio, which contains n assets. We define the return $r_t = p_{t+1}/p_t$, where p_t is the price of the asset at day t. We summarize $\mathbf{r}_t \in \mathbb{R}_+^n$.

We bet a proportion of our wealth in each asset at the beginning of day t, $\mathbf{p}_t \in \Delta_n$, by investing p_i fully liquid and infinitely-sized stock. The rv for our wealth at the beginning of day t is then given by:

$$W_t = \mathbf{r}_t^\top \mathbf{p}_t W_{t-1}$$

One may wonder why we must invest all our money in the assets - perhaps one would like to withdraw money from the markets. By introducing an asset with $r_t = 1$, we may reduce this scenario to the fully-invested option, with money withheld being equivalent to investment in a constant-return asset.

2.1 Intuition for Kelly Criterion

Example 2.1. Let us take an example where $n = 2$. Let $\mathbf{r}_t = \begin{pmatrix} a_t & b_t \end{pmatrix}^\top$, where:

$$a_t = \begin{cases} 2 & t \text{ even} \\ \frac{1}{2} & \text{o/w} \end{cases} \qquad b_t = \begin{cases} 1/3 & t \text{ even} \\ 2 & \text{o/w} \end{cases}$$

Here, we see that investment in only one of the stocks is not helpful. Having $\mathbf{p}_t = \boldsymbol{\delta}_1$ leads to no change in wealth, and $\mathbf{p}_t = \boldsymbol{\delta}_2$ results in a 33% loss of wealth every two days.

On the other hand, playing $\mathbf{p}_t = \frac{1}{2}(\boldsymbol{\delta}_1 + \boldsymbol{\delta}_2)$ gives a alternating returns of $\frac{7}{6}$ and $\frac{5}{4}$.

From this we see that there may be a way to gain money with uncorrelated asset performance for a market that has average increasing returns.

2.2 Application of Kelly Criterion

Many financial models, such as Black-Scholes, make an assumption about the stochastic process $\{\mathbf{r}_t\}$. Indeed, taking $\{\mathbf{r}_t\}$ to be n independent geometric Brownian motion processes, a similar optimal $\mathbf{p}^*$ exists satisfying the same condition as before.

This $\mathbf{p}$-based strategy, where every day one invests the same proportion of one's wealth into the set of assets, is called a constantly rebalancing portfolio strategy. The reason for this name is that even though $\mathbf{p}_t = \mathbf{p}$ remains constant, the quantity of wealth in each asset every day may change because of the difference in relative pricing of each asset on the next day.

3 OCO Framework for Portfolio Selection

If we instead don't make any assumptions about $\mathbf{r}_t$, we can still aim to minimize W_T/W_0 compared to best single p^* in hindsight.

If we choose to invest $\mathbf{p}_t$ at time t, with $\prod_t = \prod_{t=1}^T$, then

$$\frac{W_T}{W_0} = \prod_t \mathbf{r}_t^\top \mathbf{p}_t$$

In this multiplicative setting, it would seem natural to maximize our performance relative to the optimal portfolio in hindsight, $\mathbf{p}^*$, where $W_T^{\mathbf{p}^*}/W_1 = \prod_t \mathbf{r}_t^\top \mathbf{p}^*$.We seek to make more money than the optimal portfolio, up to some sublinear factors:

$$\frac{W_T/W_0}{W_T^{\mathbf{p}^*}/W_0} \geq e^{-o(T)}$$

The factor on the right may seem peculiar; it implies that the geometric mean of our returns tends to perform at least as well as geometric mean of the optimal portfolio. Recall this matches the objective of the Kelly criterion, which optimizes expected log wealth, an equivalent formulation.

Taking logarithms we are met with a familiar regret construction (let $\sum_t = \sum_{t=1}^T$):

$$\sum_t \log \mathbf{r}_t^\top \mathbf{p}_t - \max_{\mathbf{p} \in \Delta_n} \sum_t \log \mathbf{r}_t^\top \mathbf{p} \geq -o(T)$$

Indeed; multiplying throughout by -1 and observing that the negative logarithm is convex, we have an OCO formulation. Unfortunately, we cannot directly apply online gradient descent to this problem: the gradient is not bounded; take $\mathbf{r} = \mathbf{1}$ and $\mathbf{p} = \epsilon\boldsymbol{\delta}_1 + (1-\epsilon)\boldsymbol{\delta}_2$, with $\epsilon \to 0^+$.

3.1 Cover's Algorithm

Definition 3.1. Let $w_{\mathbf{p}}^t = \prod_{n=1}^t \mathbf{r}_n^\top \mathbf{p}$. Let $\int_{\mathbf{p}} = \int_{\mathbf{p} \in \Delta_n}$ and $w_{\mathbf{p}}^0 = 1$. Then choose our portfolio based on the following (*Cover's algorithm*):

$$\mathbf{p}_t = \int_{\mathbf{p}} w_{\mathbf{p}}^{t-1} \mathbf{p} \Big/ \int_{\mathbf{p}} w_{\mathbf{p}}^{t-1}$$

Note the left term is a vectorized integral. Unfortunately, this integral is not tractable. We perform a discretization to accuracy k^{-1}. Define $\Delta_n^k = \{\mathbf{p} \in \Delta_n \,|\, p_i \in \mathbb{N}/k\}$ to compute integrals. Note $\left|\Delta_n^k\right| = O(k^n)$ (proven by induction on n).

Take $\mathbf{p}^*$ to be $\operatorname{argmax}_{\mathbf{p} \in \Delta_n} \sum_t \log \mathbf{r}_t^\top \mathbf{p}$ (unique by concavity). From here on, let $\sum_{\mathbf{p}} = \sum_{\mathbf{p} \in \Delta_n^k}$.

Lemma 3.2. *Taking the strategy $\boldsymbol{p}_t$ as defined by Cover's algorithm, $W_T/W_1 = \left|\Delta_n^k\right|^{-1} \sum_{\boldsymbol{p}} w_{\boldsymbol{p}}$*

With $t > 1$ we have:

$$\mathbf{r}_t^\top \mathbf{p}_t = \mathbf{r}_t^\top \frac{\sum_{\mathbf{p}} w_{\mathbf{p}}^{t-1} \mathbf{p}}{\sum_{\mathbf{p}} w_{\mathbf{p}}^{t-1}} = \frac{\sum_{\mathbf{p}} w_{\mathbf{p}}^{t-1} \mathbf{r}_t^\top \mathbf{p}}{\sum_{\mathbf{p}} w_{\mathbf{p}}^{t-1}} = \frac{\sum_{\mathbf{p}} w_{\mathbf{p}}^t}{\sum_{\mathbf{p}} w_{\mathbf{p}}^{t-1}}$$

Then it follows by a telescoping product and the definition $w_{\mathbf{p}} = w_{\mathbf{p}}^T$ that:

$$\frac{W_T}{W_1} = \prod_t \mathbf{r}_t^\top \mathbf{p}_t = \prod_t \frac{\sum_{\mathbf{p}} w_{\mathbf{p}}^t}{\sum_{\mathbf{p}} w_{\mathbf{p}}^{t-1}} = \frac{\sum_{\mathbf{p}} w_{\mathbf{p}}^T}{\sum_{\mathbf{p}} w_{\mathbf{p}}^0} = \left|\Delta_n^k\right|^{-1} \sum_{\mathbf{p}} w_{\mathbf{p}}$$

We will also require another observation; that quantifies the error caused by our discretization.

Lemma 3.3. *There exists a single* $\boldsymbol{p}^\dagger \in \Delta_n^k$ *such that for any* $t \in [T]$, $\log \frac{\boldsymbol{r}_t^\top \boldsymbol{p}}{\boldsymbol{r}_t^\top \boldsymbol{p}^*} \geq -k^{-1}$.

By the discretization, there must exist a $\mathbf{p}^\dagger \in \Delta_n^k$ such that $\mathbf{p}^\dagger = \mathbf{p}^* + \boldsymbol{\epsilon}$, where $\|\boldsymbol{\epsilon}\|_\infty < k^{-1}$. Since $\mathbf{r}_t^\top \boldsymbol{\epsilon} \geq -\|\mathbf{r}_t\|\|\boldsymbol{\epsilon}\|$ and $\mathbf{r}_t^\top \mathbf{p}^* \leq \|\mathbf{r}_t\|\|\mathbf{p}^*\|$ by Cauchy-Schwarz, we have:

$$\log \frac{\mathbf{r}_t^\top \mathbf{p}^\dagger}{\mathbf{r}_t^\top \mathbf{p}^*} \geq \log\left(1 + \frac{\mathbf{r}_t^\top \boldsymbol{\epsilon}}{\mathbf{r}_t^\top \mathbf{p}^*}\right) \geq \log\left(1 - \frac{\|\mathbf{r}_t\|\|\boldsymbol{\epsilon}\|}{\|\mathbf{r}_t\|\|\mathbf{p}^*\|}\right)$$

Then, observe $\|\boldsymbol{\epsilon}\|^2 \leq nk^{-2}$ and that $\|\mathbf{p}^*\|^2 \geq \inf_{\mathbf{p}\in\Delta_n} \|\mathbf{p}\|^2 \geq \|\mathbf{1}/n\|^2 = n$. Replacing these inequalities above, we have that the original logarithm is at least $\log(1 - k^{-1}) \geq -k^{-1}$ (we may drop the remaining series since they alternate).

Theorem 3.4. *Cover's algorithm provides a solution to sublinear regret of log returns.*

We start with the definition of regret:

$$\begin{aligned}
\log \frac{W_T}{W_1} - \sum_t \log \mathbf{r}_t^\top \mathbf{p}^* &\geq \log \sum_{\mathbf{p}} w_{\mathbf{p}} - \sum_t \log \mathbf{r}_t^\top \mathbf{p}^* - \log \left|\Delta_n^k\right| && \text{Lemma 3.2} \\
&\geq \max_{\mathbf{p}\in\Delta_n} \log w_{\mathbf{p}} - \sum_t \log \mathbf{r}_t^\top \mathbf{p}^* - \log \left|\Delta_n^k\right| && a, b \in \mathbb{R}_+, \log(a+b) > \log a \\
&\geq \max_{\mathbf{p}} \sum_t \log \mathbf{r}_t^\top \mathbf{p} - \sum_t \log \mathbf{r}_t^\top \mathbf{p}^* - \log \left|\Delta_n^k\right| && \text{def. of } w_{\mathbf{p}} \\
&\geq \max_{\mathbf{p}} \sum_t \log \frac{\mathbf{r}_t^\top \mathbf{p}}{\mathbf{r}_t^\top \mathbf{p}^*} - \log \left|\Delta_n^k\right| \\
&\geq \sum_t \log \frac{\mathbf{r}_t^\top \mathbf{p}^\dagger}{\mathbf{r}_t^\top \mathbf{p}^*} - \log \left|\Delta_n^k\right| \\
&\geq -T/k - O(n \log k) && \text{Lemma 3.3} \\
&\geq -O(n \log T) = o(T) && \text{choice of } k = T^2
\end{aligned}$$