Lecture 15

Agenda:
* Universal portfolio formulation and constantly balanced portfolio
* Solution using probability assignment
* Lower bounds

[A] Universal Portfolio Formulation

Consider two stocks A and B.

Suppose over a duration of n days, stock A is valued at Rs 100, and stock B is rather volatile and follows the price pattern

100, 200, 100, 200, ...

An investor with Rs 1 lakh at the outset splits his total wealth between the two stocks (in appropriate ratio). If the investor puts all his money on any one stock, she will not make any money at the end.

On the other hand, with the crystal clarity of hindsight, an investor may choose to invest the 1 lakh amount in two stocks in equal proportion. Then, each day, the stock either increases by a ratio of \((\frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 2) = \frac{3}{2}\) or decreases by \((\frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 1) = \frac{3}{4}\). Thus, on any two consecutive days, the portfolio will grow by \(\frac{9}{8}\) and double every 12 days!
Both strategies above are instances of a "constant
ly balanced (or rebalanced) portfolio" where every day we split
our wealth in the same proportion b/w the two stocks.
Note that the actual value of the stock does not
enter our calculation since we allow the purchase of any
fraction of the share. The performance of a given
constant balanced portfolio depends on the actual sequence
of gains, which is \((1, 2), (1, 3), (1, 2), (1, 3), \ldots\) in the example
above. Our goal is to compete with the best constantly
balanced portfolio in hindsight.
Formally, consider the sequence \(x_1, \ldots, x_n\) where each \(x_t \in [0, 1]^2\).
At time \(t\), we are allowed to see the performance
\(x_{t-1}\) and decide on what fraction \(\theta_t \in [0, 1]\) of our wealth
should we invest on stock \(A\). Thus, the gain at time \(t\)
is
\[
Y_t = \theta_{t1} x_{t1} + (1 - \theta_{t}) x_{t2}.
\]
The overall gain after \(n\) days is given by
\[
T_n (x, \hat{\theta}) = \prod_{t=1}^{n} Y_t = \prod_{t=1}^{n} (\theta_{t1} x_{t1} + \theta_{t2} x_{t2})
\]
For a constantly rebalanced portfolio (CRP),
\[
\theta_t = \Theta \neq t.
\]
Our goal is to remain multiplicatively competitive with the best CRP in hindsight. To that end, for a \( \hat{\theta}_t : x^{t-1} \mapsto \hat{\theta}_t = \hat{\theta}_t(x^{t-1}) \),

denote

\[
\pi_n(x, \hat{\theta}) = \left( \max_{\theta} \log T_n(x, \theta) \right) - T_n(x, \hat{\theta})
\]

\[
= \max_{\theta} \log \frac{T_n(x, \theta)}{T_n(x, \hat{\theta})},
\]

where \( T_n(x, \theta) \) denotes the performance of CRP and equals \( \prod_{t=1}^n (x_{t1} \theta + (1-\theta)x_{t2}) \).

Denote

\[
\pi_n = \min_{\hat{\theta}} \max_{x \in \{0,1\}^n} \pi_n(x, \hat{\theta})
\]

We will see that \( \pi_n = \frac{1}{2} \log n + O(1) \) and is attained by a simple Bayesian strategy.

B. Probability assignment and optimal portfolios

Note that for \( k = 2 \),

\[
T_n(x, \theta) = \prod_{t=1}^n (\theta_{t1} x_{t1} + \theta_{t2} x_{t2})
\]

\[
= \sum_{j \in \{1,2\}^n} \prod_{t=1}^n \theta_{tj} x_{tj} = \prod_{t=1}^n \theta_{tj_t} = \prod_{t=1}^n x_{tj_t}
\]
\[ \log \frac{T_n(x, \theta)}{\widehat{T}_n(x, \hat{\theta})} = \log \frac{\sum_{j \in \{1, 2, \ldots, n\}} \theta(j) \lambda(j)}{\sum_{j \in \{1, 2, \ldots, n\}} \theta'(j) \lambda(j)} \leq \max_{j \in \{1, 2, \ldots, n\}} \log \frac{\theta(j)}{\theta'(j)} \]

In particular,

\[ \log \frac{T_n(x, \theta)}{\widehat{T}_n(x, \hat{\theta})} \leq \max_{j \in \{1, 2, \ldots, n\}} \log \theta^{n(1-x^0)}(1-\theta)^{n-n(1-x^0)} \theta'(j) \lambda(j) \]

where

\[ \hat{\theta}(j \mid x) = \theta_1(j_1) \theta_2(j_2 \mid x), \theta_3(j_3 \mid x^2), \ldots = \prod_{t=1}^{k} \theta_t(j_t \mid x^{t-1}) \]

where \( x^0 = \phi \). Note that

\[ \sum_{j} \hat{\theta}(j \mid x) = 1. \]

Thus,

\[ \tau_n \leq \min \max \max \max \log \theta^{n(1-x^0)}(1-\theta)^{n-n(1-x^0)} \theta'(j) \lambda(j) \hat{\theta}(j \mid x) \]

We note that the cost function on the right-hand side appears very similar to that for the probability
assignment problem seen in the previous lecture. However, the observables are $x^{t-1}$ and not $i^{t-1}$.

Nevertheless, there is a very simple way to convert one problem to the other. Consider a distribution $Q$ on $\Xi_1, \Xi_2$.

Define $\hat{\Theta} = \Theta_{\theta}$ by

$$\Theta_{\theta} = \hat{\Theta}(1 | x^{t-1}) = \frac{\sum_{d : j_t = 1} Q(d) \prod_{i=1}^{t-1} x_{ij_i}}{\sum_{d} Q(d) \prod_{i=1}^{t-1} x_{ij_i}}$$

Then,

$$\hat{\Theta}(d | x) = \prod_{t=1}^{n} \hat{\Theta}(d_t | x^{t-1})$$

$$= \prod_{t=1}^{n} \left[ \frac{\sum_{d' : j_t = d_t} Q(d') \prod_{i=1}^{t-1} x_{ij_i}'}{\sum_{d'} Q(d') \prod_{i=1}^{t-1} x_{ij_i}'} \right]$$

$$= \prod_{t=1}^{n} \left[ \frac{\sum_{j_{t-1}} Q(j_{t-1} | j_t^{t-1}) Q(d' | j_{t-1}) \prod_{i=1}^{t-1} x_{ij_i}}{\sum_{j_{t-1}} Q(j_{t-1} | j_t^{t-1}) \prod_{i=1}^{t-1} x_{ij_i}} \right]$$

$$= Q(d) \cdot \frac{Q(d_t | x_n) + Q(d_{t-1} | x_{n-1})}{Q(d_t | x_n) + Q(d_{t-1} | x_{n-1})}.$$
Note that
\[
(\hat{\theta}(1)x_n + \hat{\theta}(2)x_{n+1})(\hat{\theta}(1)x_{n+1} + \hat{\theta}(2)x_{n+2})
\]
\[
= \sum_{j \in \{1,2\}} q(j) \prod_{t=1}^n x_{t+j},
\]
and similarly for larger time horizons. Therefore, restricting the min. on the right-side of (1) to scheme \(\hat{\theta}\) given by (2), we have
\[
\tau_n \leq \min_{Q \in \mathcal{P}(\mathbb{Z})} \max_{j \in \{1,2\}} \max_{\theta \in [0,1]} \log \frac{\Theta^{\theta(j)}}{(1-\theta)^{\theta(2j)}}.
\]

where the optimal \(Q\) for the right-side was obtained (using KT estimator) in the previous class.

\[
\Rightarrow \tau_n \leq \mathcal{R}(2,n) = \frac{1}{2} \log n + O(1).
\]

\[\square\ \text{Lower bound}\]

We have

\[
\max_{x \in [0,1]^n} \log \frac{T_n(x, \theta)}{T_n(x, \hat{\theta})} = \max_{x \in \Delta(0,1 \setminus \{0,1\})^n} \log \frac{T_n(x, \theta)}{T_n(x, \hat{\theta})}
\]

For any \(x \in \Delta(0,1 \setminus \{0,1\})^n\) and \(\theta\), we have
\[
T_n(x, \theta) = \prod_{t=1}^n (\theta + I_{x_{2t+1}=1} + (1-\theta)I_{x_{2t}=0})
\]
\[ = \Theta(x,)
\]

where \( \Theta(x_t = 1 \mid x_{t-1}) = \Theta(x_t = 1 \mid x_{t-1}) = \Theta(x_t = 1 \mid x_{t-1}) \).

Therefore, when restricted to these inputs, the problem gets reduced to the probability assignment problem, and so,

\[ r(n) \geq r(2, n) \]

Overall, we have

\[ r(n) = r(2, n), \]

and the portfolio (2) corresponding to \( \Phi(x_t) \) is optimal.