Prof. Friedrich Eisenbrand
Location: MA A3 31

Question session: 30.09.10
Discussion: 07.10.10

Exercises

Optimization Methods in Finance

Fall 2010

Sheet 1

Note: This is just one way, a solution could look like. We do not guarantee correctness. It is your task to find and report mistakes.

Exercise 1.1
Let X be a finite set of N elements (assume that N is a power of 2). At least one of these elements is *interesting*. You may ask questions of the form: *Does $H \subseteq X$ contain an interesting element?* Design an algorithm that identifies an interesting element by asking at most $\log_2(N)$ questions of this kind.

Can there exist a deterministic algorithm that identifies an interesting element by asking fewer questions in the worst case?

Solution:

Take a set $H \subseteq X$ of size $|H| = |X|/2$. Ask whether H constains an interesting element. If yes recurse on H, otherwise recurse on $X \backslash H$. We obtain a sequence $X_0 := X, X_1, X_2, \ldots$ with the property that X_t still contains at least one interesting element and $|X_t| = N/2^t$. We can stop, when $|X_t| = 1$ (which happens for $t = \log_2(N)$.

However, this strategy is optimal. Suppose for contradiction that there is a deterministic algorithm asking at most $\log_2(N) - 1$ questions. We can describe the questions of the algorithm with a *decision tree*. In any node, we write the set H, which the algorithm is going to ask. Every interiour node has one outgoing "yes" and one outgoing "no" edge. We label leaves with the output (a number from $1, \ldots, N$) of the algorithm.

Suppose the algorithm asks less than $\log_2(N)$ questions, then the tree has at most $2^{\log_2(N)-1} = \frac{N}{2}$ many leaves. In other words, there is at least one element $i \in \{1, \ldots, N\}$ which is *never* the output of the algorithm. But if i is the only interesting element and we answer all questions correctly, still the algorithm would not yield i as output.

Exercise 1.2
Show the following inequalities for $0 \leq \varepsilon \leq 1/2$:

1. $(1-\varepsilon)^x \leq (1-\varepsilon x)$ for $x \in [0,1]$.
2. $(1+\varepsilon)^{-x} \leq (1-\varepsilon x)$ for $x \in [-1,0]$.
3. $\ln\left(\frac{1}{1-\varepsilon}\right) \leq \varepsilon + \varepsilon^2$.
4. $\ln(1+\varepsilon) \geq \varepsilon - \varepsilon^2$.

Solution:

1. Let $\varepsilon > 0$, $0 \leq x \leq 1$. Define

$$f(\varepsilon) = (1-\varepsilon x) - (1-\varepsilon)^x$$

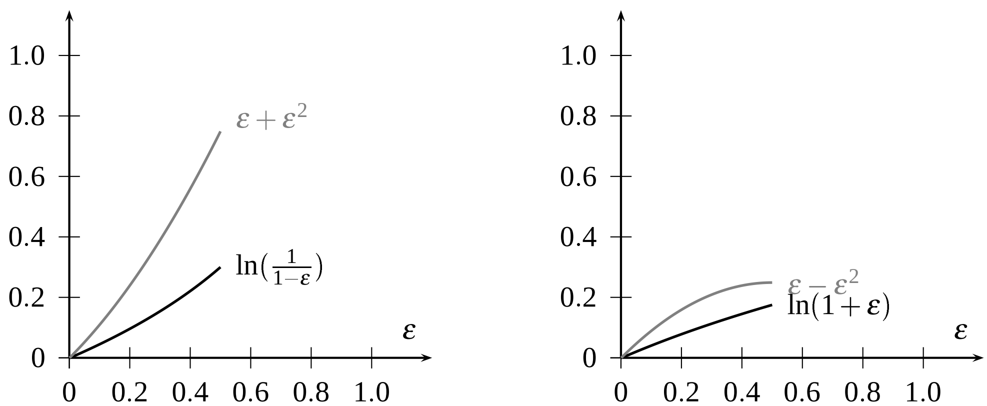
then it suffices to show that $f(\varepsilon) \geq 0$ for all $\varepsilon > 0$. Note that $f(0) = 0$. If there would be an $\varepsilon_0 > 0$ with $f(\varepsilon_0) < 0$, then by the *Mean value Theorem* there would be a $\varepsilon_1 > 0$ with $f'(\varepsilon_1) < 0$. Hence we will show that $f'(\varepsilon) \geq 0$ for all $\varepsilon \geq 0$. Then

$$f'(\varepsilon) = -x + x\cdot(1-\varepsilon)^{x-1} = \underbrace{x}_{\geq 0} \cdot (\underbrace{(1-\varepsilon)^{x-1}}_{\geq 1} - 1) \geq 0$$

2. Similar to (1).

4. The claim is equivalent to showing $1+\varepsilon \geq e^{\varepsilon-\varepsilon^2}$. Recall that $e^x = \sum_{i=0}^{\infty} \frac{x^i}{i!}$. Hence

$$= e^{\varepsilon-\varepsilon^2} \overset{\text{Tailor series}}{=} \sum_{i=0}^{\infty} \frac{(\varepsilon-\varepsilon^2)^i}{i!} \leq 1 + (\varepsilon-\varepsilon^2) + \sum_{i=2}^{\infty} \frac{\varepsilon^i}{i!} \leq 1 + (\varepsilon-\varepsilon^2) + \frac{1}{2}\varepsilon^2 \underbrace{\sum_{i=0}^{\infty} (1/2)^i}_{=2} = 1+\varepsilon$$



3. Similar to (4).

Exercise 1.3 (*)
Consider the randomized weighted majority algorithm and suppose that the loss-vectors at time t satisfy $\ell^t \in [0,\rho]^N$ for $t = 0,\ldots,T$. Show that the expected loss of the forecaster is bounded by

$$E[L] \leq \frac{\rho \cdot \ln N}{\varepsilon} + (1+\varepsilon)\cdot L^j,$$

if one uses the update rule $w_j := w_j(1-\varepsilon)^{\ell_j^T/\rho}$. As in the lecture, $L^j = \sum_{t=0}^{T} \ell_j^t$ is the loss accumulated by expert j.

Solution:

Let $\overline{\ell^t} = \frac{1}{\rho}\ell^t$ be the scaled loss vector. Then $0 \leq \overline{\ell^t_j} \leq 1$, hence the result from the lecture implies that $E[\overline{L}] \leq \frac{\ln(N)}{\varepsilon} + (1+\varepsilon)\overline{L}^j$. I.e.

$$E[L] \leq \rho\frac{\ln(N)}{\varepsilon} + (1+\varepsilon)\underbrace{\rho\overline{L}^j}_{=L^j}$$

Exercise 1.4 (*)

Suppose that the loss vectors satisfy $\ell^t \in [-1,1]^N$ for $t = 0,\dots,T$ and consider the following updating rule

$$w_j^{t+1} := \begin{cases} w_j^t(1-\varepsilon)^{\ell_j^t} & \text{if } \ell_j^t \geq 0 \\ w_j^t(1+\varepsilon)^{-\ell_j^t} & \text{if } \ell_j^t < 0. \end{cases}$$

Show that

$$E[L] \leq \frac{\ln N}{\varepsilon} + (1+\varepsilon)\cdot\sum_{t:\ell_j^t \geq 0}\ell_j^t + (1-\varepsilon)\cdot\sum_{t:\ell_j^t<0}\ell_j^t$$

holds. Show furthermore that this is also guaranteed if the update rule $w_j^{t+1} := w_j(1-\varepsilon\ell_j^t)$ is used.

Solution:

Let $W^t = \sum_{j=1}^N w_j^t$ the total weight at time t. We abbreviate the expected loss of the forecaster in period t by

$$E[\hat{L}_t] := \sum_{j=1}^N \hat{p}_j\cdot\ell_t^j = \sum_{j=1}^N \frac{w_j^t\cdot\ell_t^j}{W^t}$$

We know that the total weight of the experts behaves as follows:

$$\begin{aligned}
W^{t+1} &= \sum_{j=1}^N w_j^{t+1} \\
&= \sum_{j:\ell_j^t\geq 0} w_j^t\cdot\underbrace{(1-\varepsilon)^{\ell_j^t}}_{\leq 1-\varepsilon\ell_j^t \text{ (Ex 1.2.1)}} + \sum_{j:\ell_j^t<0} w_j^t\cdot\underbrace{(1+\varepsilon)^{-\ell_j^t}}_{\leq 1-\varepsilon\ell_j^t \text{ (Ex 1.2.2)}} \\
&\leq \sum_{j=1}^N w_j^t\cdot(1-\ell_j^t\cdot\varepsilon) \\
&= W^t - \varepsilon\cdot\underbrace{\sum_{j=1}^N w_j^t\ell_j^t}_{=W^t\cdot E[L_t]} \\
&= W^t\cdot(1-\varepsilon\cdot E[L_t]) \\
&\overset{1-x\leq e^x}{\leq} W^t\cdot e^{-\varepsilon E[L_t]}
\end{aligned}$$

Iterating this, yields

$$W^T \leq \underbrace{W^0}_{=N}\cdot\prod_{t=1}^T e^{-\varepsilon E[L^t]} = N\cdot e^{-\sum_{t=1}^T E[L^t]} = N\cdot e^{-\varepsilon\cdot E[L]}$$

Still

$$w_j^T = \prod_{t:\ell_j^t \geq 0}^{T} (1-\varepsilon)^{\ell_j^t} \cdot \prod_{t:\ell_j^t < 0} (1+\varepsilon)^{-\ell_j^t}$$

Again $N \cdot e^{-\varepsilon \cdot E[L]} \geq w_j^T$, hence

$$\begin{aligned} E[L] &\leq \frac{\ln(N)}{\varepsilon} + \frac{1}{\varepsilon}\ln(\frac{1}{w_j^T}) \\ &= \frac{\ln(N)}{\varepsilon} + \sum_{t:\ell_j^t \geq 0} \frac{1}{\varepsilon}\ln(\frac{1}{(1-\varepsilon)^{\ell_j^t}}) + \sum_{t:\ell_j^t < 0} \ln(\frac{1}{(1+\varepsilon)^{-\ell_j^t}}) \\ &= \frac{\ln(N)}{\varepsilon} + \sum_{t:\ell_j^t \geq 0} \ell_j^t \underbrace{\frac{1}{\varepsilon}\ln(\frac{1}{(1-\varepsilon)})}_{\leq 1+\varepsilon} + \sum_{t:\ell_j^t < 0} \ell_j^t \underbrace{\frac{1}{\varepsilon}\ln(1+\varepsilon))}_{\leq 1-\varepsilon} \end{aligned}$$

Exercise 1.5 (*)
Suppose you have some initial belief about the quality of the experts. This belief is represented by a probability distribution on the experts p_j, $j = 1, \ldots, N$ with $p_j > 0$ and $\sum_{j=1}^n p_j = 1$. We modify the weighted majority algorithm by setting the initial weights $w_j := p_j$. Show that this modification results in a guarantee

$$E[L] \leq \frac{\ln(1/p_j)}{\varepsilon} + (1+\varepsilon) \cdot L^j.$$

Suppose now that we have a countably infinite number of experts. Use the result above to argue that one can guarantee

$$E[L] \leq \frac{2 \cdot \ln(j) + 10}{\varepsilon} + (1+\varepsilon) \cdot L^j.$$

by choosing a suitable probability distribution on the experts.

Solution:

When setting $w_j^0 := p_j$, the total inital weight is 1 instead of N. Furthermore the weight of expert j at the end is $p_j \cdot (1-\varepsilon)^{L_j}$ hence

$$E[L] \leq \frac{1}{\varepsilon} \cdot \ln\left(\frac{W^0}{p_j(1-\varepsilon)^{L_j}}\right) = \frac{\ln(1/p_j)}{\varepsilon} + L_j \underbrace{\frac{1}{\varepsilon}\ln(\frac{1}{1-\varepsilon})}_{\leq 1+\varepsilon} \leq \frac{\ln(1/p_j)}{\varepsilon} + (1+\varepsilon) \cdot L^j.$$

Next, choose probability distribution $p_j = \frac{6}{\pi^2 \cdot j^2}$. Then indeed $\sum_{j=1}^\infty p_j = 1$ and

$$E[L] \leq \frac{\ln(1/p_j)}{\varepsilon} + (1+\varepsilon) \cdot L^j = \frac{\ln(\frac{\pi^2 j^2}{6})}{\varepsilon} + (1+\varepsilon) \cdot L^j \leq \frac{2\ln(j) + 1.2}{\varepsilon} + (1+\varepsilon) \cdot L^j$$
