Online Machine Learning Homework

Assignment 1

Student: Lei Feng, `please-help@each.other`
Student ID: 123456
Lecturer:

In the online optimization setting, the player plays a point $w_t \in \mathcal{W}$, the adversary responds with a non-negative function f_t, and the player suffers a loss of $f_t(w_t)$. Assume that $\mathcal{W}$ is a bounded set and that each f_t is lower bounded on $\mathcal{W}$. The player's *regret* after T rounds is defined as

$$\sum_{t=1}^{T} f_t(w_t) - \min_{w_t \in W} \sum_{t=1}^{T} f_t(w)$$

A *regret* bound is a function $R(T)$ such that for any sequence f_1,..., f_T it holds that

$$\forall\, T,\ \sum_{t=1}^{T} f_t(w_t) - \min_{w \in W} \sum_{t=1}^{T} f_t(w) \leq R(T)$$

Problem 1: Understand conservative online optimization alogorithm

First, prove that we can assume, without loss of generality, that $\min f_t(x) = 0$ for each t. An online optimization algorithm is *conservative* if

$$f_t(w_t) = 0 \Rightarrow w_{t+1} = w_t$$

In other words, a conservative algorithm keeps playing the same point as long as it doesn't suffer any loss. Let $\mathcal{A}$ be an online optimization algorithm with a regret bound of $R(T)$. Use $\mathcal{A}$ to define a conservative online optimization algorithm $\mathcal{A}^{'}$ with the same regret bound.

Proof. Note that regardless of the actual value of $\min f_t(x)$, we can always define a new loss function $f_t'(w_t)$ as

$$f_t'(w_t) = f_t(w_t) - \min f_t(x).$$

The first term represents the original loss, and the second term is the minimal loss with which we are interested. Bu subtracting the minimal loss, we have now achieved zero-minimal for the new loss function $f_t'(w_t)$.

Thus, the new regret is now defines as

$$\sum_{t=1}^{T} f_t'(w_t) - \min_{w_t \in W} \sum_{t=1}^{T} f_t'(w)$$

Since the second term already proved to be 0, the new regret can be re-written as

$$\begin{aligned}
\sum_{t=1}^{T} f_t'(w_t) - \min_{w_t \in W} \sum_{t=1}^{T} f_t'(w) &= \sum_{t=1}^{T} f_t(w_t) - \sum_{t=1}^{T} \min_{x \in \mathcal{W}} f_t(x) - 0 \\
&= \sum_{t=1}^{T} f_t(w_t) - \min_{w_t \in W} \sum_{t=1}^{T} f_t(w)
\end{aligned}$$

which is exactly the same as the original regret.

Hence, the new loss function $f_t'(w_t)$ serves as a zero-minimal alternative to the original loss without compromising generality. $\square$

Next, we define a conservative online optimization algorithm $\mathcal{A}^{'}$ with the same regret bound. Intuitively, define $\mathcal{A}^{'}$ as follows.

Algorithm 1 conservative online optimization algorithm $\mathcal{A}^{'}$

$w_1 \leftarrow w_1^{\mathcal{A}}$
for rounds $t = 1, 2, \ldots, T$ **do**
  if $f_t(w_t) = 0$ **then**
    $w_{t+1} \leftarrow w_t$
  else
    $w_{t+1} \leftarrow w_{t+1}^{\mathcal{A}}$
  end if
end for

This means that $\mathcal{A}'$ starts with the same move as $\mathcal{A}$, and then check the loss of each round. If the loss is 0, then $\mathcal{A}'$ keeps playing the same point in the next round, if not, then $\mathcal{A}'$ follows whichever point that $\mathcal{A}$ would play in the next round. The former condition ensures the conservative attribute of $\mathcal{A}'$, while the latter guarantees that $\mathcal{A}'$ shares the same regret bound as $\mathcal{A}$. In actuality, the conservative rule makes sure that the player doesn't incur any extra loss, which means that the regret of $\mathcal{A}'$ is *identical* to that of $\mathcal{A}$.

> **Problem 2: Understand Convexification**
>
> Recall that a function f is *convex* if
>
> $$f(\alpha x \ + (1-\alpha)x^{'}) \ \leq \ \alpha f(x) + (1-\alpha) f(x^{'})$$
>
> for any $\alpha \in [0, 1]$ and any x and $x^{'}$ in f's domain.
>
> Let $f : \mathbb{R} \to \mathbb{R}$ be a convex function and let $g : \mathbb{R} \to \mathbb{R}$ be a convex monotonically non-decreasing function. Prove that the composition $g \circ f$ is convex($g \circ f \equiv g(f(x))$).

Proof. According to the definition of convex functions, $f : \mathbb{R} \to \mathbb{R}$ being convex implies that,

$$f(\alpha x \ + (1-\alpha)x^{'}) \ \leq \ \alpha f(x) + (1-\alpha) f(x^{'}) \tag{1}$$

for any $\alpha \in [0, 1]$ and any $x, x' \in \mathbb{R}$.

Also, $g : \mathbb{R} \to \mathbb{R}$ being monotonically non-decreasing means that

$$\forall x_1, x_2 \in \mathbb{R} \ , x_1 \leq x_2 \Rightarrow g(x_1) \leq g(x_2).$$

By replacing x_1 and x_2 with the left and right side of inequality in Eq.1 respectively, we see that

$$g(f(\alpha x \ + (1-\alpha)x^{'})) \leq g(\alpha f(x) + (1-\alpha) f(x^{'}))$$

that is,

$$g \circ f(\alpha x \ + (1-\alpha)x^{'}) \leq g(\alpha f(x) + (1-\alpha) f(x^{'})) \tag{2}$$

Since g is also convex, apply the definition of convex functions on g and we get

$$g(\alpha f(x) + (1-\alpha)f(x^{'})) \leq \alpha g(f(x)) + (1-\alpha)g(f(x^{'})),$$

that is,

$$g(\alpha f(x) + (1-\alpha)f(x^{'})) \leq \alpha\, g \circ f(x) + (1-\alpha) g \circ f(x^{'}) \tag{3}$$

Thus, combining Eq.2 and Eq.3, we have

$$g \circ f(\alpha x\ + (1-\alpha)x^{'}) \leq \alpha\, g \circ f(x) + (1-\alpha) g \circ f(x^{'})$$

for all $x_1, x_2 \in \mathbb{R}$ and $\alpha \in [0, 1]$. This proves that the composition $g \circ f$ is convex. □

Problem 3: Online Stock Portfolio Management

Consider the problem of managing an online stock portfolio in a market with no transaction costs. Assume that the market has n different stocks, we can change our investment portfolio at the end of each trading day, and the prices of the n stocks at the end of day t are denoted by the vector c_t. Our initial wealth is ϕ_0 and our wealth after round t is ϕ_t. On each round, we play a distribution vector $w_t \in \mathcal{W}$ ($\mathcal{W}$ is the set of non-negative vectors that sum to 1). Namely, on round t we invest $\phi_{t-1}w_{t,i}$ dollars in stock i.

(a) Write ϕ_t in terms of $w_1, ..., w_t$ and $c_0, c_1, ..., c_t$.

(b) A *constantly rebalancing portfolio* (CRP) defined by a fixed probability vector w is an investment strategy that rebalances every day so that exactly w_i of our wealth is invested in stock i on each day. Let ϕ_t^w denote the wealth of the CRP defined by w on day t. Write ϕ_t^w in terms of w and $c_0, c_1, ..., c_t$.

(c) Define the wealth of the best CRP in hindsight after T rounds as $\phi_T^* = \max_w \phi_T^w$. Define regret after T rounds as $log(\phi_T^*/\phi_T)$. Show that minimizing this definition of regret is a special case of the online convex optimization framework discussed in class (Hint: use Problem 2 to show that $-\log(u \cdot v)$ is convex and write the portfolio management problem as an online convex optimization problem).

(a) Consider the first day(Day 1): we start with wealth ϕ_0, and invest $\phi_0 w_{1,i}$ dollars in stock i. By the end of the day, the price of stock i went from $c_{0,i}$ to $c_{1,i}$. Hence, the money invested in stock i, went from $\phi_0 w_{1,i}$ to $\phi_0 w_{1,i}\frac{c_{1,i}}{c_{0,i}}$. Thus, the wealth we now possess is

$$\phi_1 = \sum_{i=1}^{n} \phi_0 \frac{w_{1,i}c_{1,i}}{c_{0,i}} = \phi_0 w_1/c_0 \times c_1^T.$$

Here, / denotes element-wise division between two vectors of the same size.

Stacking the equation $w.r.t\ i$ from 1 through t, thus we have

$$\phi_t = \phi_0 \prod_{i=1}^{n} w_i/c_{i-1} \times c_i^T \tag{4}$$

(b) Alter the w_i term in Eq.4, we then have

$$\phi_t^w = \phi_0 \prod_{i=1}^{n} w/c_{i-1} \times c_i^T$$

(c) First, we provide another expression of ϕ_t^w. In (b), we established that in the context of CRP, $\phi_t^w = \phi_0 \prod_{i=1}^n w/c_{i-1} \times c_i^T$. An alternative way to express ϕ_t^w is by using element-wise multiplication instead of matrix multiplication

$$\phi_t^w = \phi_0 \prod_{i=1}^n w \cdot c_i/c_{i-1} \tag{5}$$

where / still stands for element-wise division.

Then

$$\begin{aligned} Regret &= \log\left(\phi_T^*/\phi_T\right) \\ &= \log \phi_T^* - \log \phi_T \\ &= -\log \phi_T - (-\log \phi_T^*) \end{aligned}$$

Intuitively, we define the loss function $f_t(w_t) = -\log \phi_t$. In this sense, to minimize the loss is to maximize the wealth, which aligns with our ultimate goal.

Thus introducing Eq.5, and we have

$$f_t(w_t) = -\log \phi_0 \prod_{i=1}^n w \cdot c_i/c_{i-1}$$

Denote c_i/c_{i-1} as x_i, then equivalently

$$f_t(w_t) = -\log \phi_0 \prod_{i=1}^n w \cdot x_i$$

which corresponds with the form $-\log(u \cdot v)$.

Finally, we know from Problem 2 that since $g(x) = \log x$ and $f(v) = \frac{1}{u \cdot v}$ are both convex, their composition $g \circ f$ is also convex, hence the convexificaion of the loss function $f_t(w_t)$. Therefore, the portfolio management problem can be interpreted as an online convex optimization problem.

Problem 4: Understand Tight Mistake Bound

Prove that the mistake bound that we proved for the Perceptron algorithm is *tight*. In other words, for any $\gamma > 0$ and any $\rho > 0$ find a sequence $\{(x_t, y_t)\}_{t=1}^\infty$ such that $\|x\| \leq \rho$ for all t , such that there exists w^* with $\|w^*\| = 1$ and $y_t w^* \cdot x_t \geq \gamma$ for all t, and such that the Perceptron makes exactly $\lfloor \rho^2/\gamma^2 \rfloor$ mistakes.

Choose x_t and y_t according to the following pipeline:

Algorithm 2 Sequence Construction Procedure

$\forall\, \gamma, \rho > 0$
$w^* \leftarrow e_1 = (1, 0, \dots)$
$w_1 \leftarrow 0$
for step $t = 1, 2, \dots$ **do**
  $v_t \leftarrow -w_t + \frac{\gamma}{\rho^2} w^*$
  $x_t \leftarrow$
  $y_t \leftarrow sign(w^* \cdot x_t)$
end for

By choosing $x_t = \rho \frac{v_t}{\|v_t\|}$, where $v_t = -w_t + \frac{\gamma}{\rho^2} w^*$, we make sure that $\|x\| = \rho$, and that $y_t w^* \cdot x_t = \gamma$, thus satisfying the requirement.

Each time the Perceptron makes a mistake, the weight vector is updated as:

$$w_{t+1} = w_t + y_t x_t$$

Hence,

$$\|w_{t+1}\|^2 = \|w_t\|^2 + 2 y_t w_t \cdot x_t + \|x_t\|^2.$$

Since $y_t w_t \cdot x_t \leq 0$, and $\|x\| = \rho$, we then have

$$\|w_{t+1}\|^2 \leq \|w_t\|^2 + \rho^2.$$

After M mistakes, the norm satisfies:

$$\|w_{M+1}\|^2 \leq M\rho^2. \tag{6}$$

Additionally, $y_t w^* \cdot x_t \geq \gamma$, which then leads to

$$w_{M+1} \cdot w^* \geq M\gamma \tag{7}$$

after making M mistakes.

From Eq.7, we derive that

$$\|w_M + 1\| \geq w_{M+1} \cdot w^* \geq M\gamma \tag{8}$$

as the norm of a vector is always greater than all of its projections.

Therefore, combining Eq.6 and Eq.8, we have

$$M^2\gamma^2 \leq M\rho^2$$

Hence,

$$M \leq \frac{\rho^2}{\gamma^2}$$