CSE599s Spring 2012 - Online Learning
Homework Exercise 1 - due 4/12/12

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

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

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

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

1. 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

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

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

2. 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 $f : \mathbb{R} \mapsto \mathbb{R}$ be a convex function and let $g : \mathbb{R} \mapsto \mathbb{R}$ be a convex monotonically non-decreasing function. Prove that the composition $g \circ f$ is convex ($g \circ f(x) \equiv g(f(x))$).

3. 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.

- Write ϕ_t in terms of $w_1, \ldots, w_t$ and $c_0, c_1, \ldots, c_t$.
- 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 $c_0, c_1, \ldots, c_t$.
- Define the wealth of the best CRP in hindsight after T rounds as $\phi_T^\star = \max_w \phi_T^w$. Define regret after T rounds as $\log(\phi_T^\star/\phi_T)$. 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 managment problem as an online convex optimization problem).

4. 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 $\|x_t\| \leq \rho$ for all t, such that there exists $w^\star$ with $\|w^\star\| = 1$ and $y_t w^\star \cdot x_t \geq \gamma$ for all t, and such that the Perceptron makes exactly $\lfloor \rho^2/\gamma^2 \rfloor$ mistakes.