CMPUT 675: Homework

Shuai Liu

December 2021

1 Part I

1.1 Question 1.1[1]

Since x^* is a local minimizer, $\nabla f(x^*) = 0$. By convexity, we have

$$\forall y \in dom(f), f(y) \geq f(x^*) + (y - x^*)\nabla f(x^*) = f(x^*) \tag{1}$$

x^* is thus a global minimizer.

Define $h(x) = \frac{L}{2}\|x\|_2^2 - f(x)$. $(\nabla h(x) - \nabla h(y))^T(x - y) \geq \|\nabla h(x) - \nabla h(y)\|_2^2 \geq 0$. $h(x)$ is thus convex. By convexity and first-order condition,

$$\begin{aligned}
\forall x, y \in dom(f), h(y) &\geq h(x) + h(x)^T(y - x) \\
\frac{L}{2}y^T y - f(y) &\geq \frac{L}{2}x^T x - f(x) + (Lx - \nabla f(x))^T(y - x)
\end{aligned}$$

Rearranging terms, the following inequality holds:

$$f(y) \leq f(x) + \nabla f(x)^T(x - y) + \frac{L}{2}\|x - y\|_2^2 \tag{2}$$

Also, the sequence $\{F(x^i)\}_{i=1}^k$ is decreasing:

$$\begin{aligned}
f(x^{(i+1)}) &\leq f(x^{(i)}) + \nabla f(x^{(i)})(x^{(i+1)} - x^{(i)}) + \frac{L}{2}\|x^{(i+1)} - x^{(i)}\|_2^2 \\
&\leq f(x^{(i)}) - t\|\nabla f(x^{(i)})\| + \frac{Lt^2}{2}\|\nabla f(x^{(i)})\|_2^2 \\
f(x^{(i)}) - f(x^{(i+1)}) &\geq t(1 - \frac{Lt}{2})\|\nabla f(x^{(i)})\|_2^2 \geq 0
\end{aligned}$$

The last inequality holds since $t(1 - \frac{Lt}{2}) \geq \frac{t}{2}$ by $t \in (0, \frac{1}{L}]$. then,

$$\begin{aligned}
f(x^{(i+1)}) &\leq f(x^{(i)}) - \frac{t}{2}\|\nabla f(x^{(i)})\|_2^2 \\
&\leq f(x^{(*)}) + \nabla f(x^{(*)})^T(x^{(i)} - x^{(*)}) - \frac{t}{2}\|\nabla f(x^{(i)})\|_2^2 \\
&= f(x^{(*)}) + \frac{1}{2t}(\langle x^{(i)} - x^{(i+1)}, x^{(i)} - x^{(*)}\rangle - \langle x^{(i)} - x^{(i+1)}, x^{(i)} - x^{(i+1)}\rangle \\
&= f(x^{(*)}) + \frac{1}{2t}\langle x^{(i)} - x^{(i+1)}, x^{(i)} + x^{(i+1)} - 2x^{(*)}\rangle \\
&= f(x^{(*)}) + \frac{1}{2t}\langle x^{(i)} - x^{(*)} + x^{(*)} - x^{(i+1)}, x^{(i)} - x^{(*)} + x^{(i+1)} - x^{(*)}\rangle \\
&= f(x^{(*)}) + \frac{1}{2t}(\|x^{(i)} - x^{(*)}\|_2^2 - \|x^{(i+1)} - x^{(*)}\|_2^2)
\end{aligned}$$

[1] Referenced https://jliang993.github.io/nsopt/slides/lecture-01.pdf when solving the problem

Summarize over i on both sides of the inequality:

$$\sum_{i=1}^{k} f(x^{(i+1)}) - F(x^{(*)}) \leq \frac{1}{2t}(\|x^{(0)} - x^{(*)}\|_2^2)$$

Finally, since the sequence $\{f(x^i) - f(x^{(*)})\}_{i=1}^k$ is decreasing

$$f(x^k) - f(x^{(*)}) \leq \frac{1}{k}\sum_{i=1}^{k} f(x^{(i+1)}) - f(x^{(*)}) \leq \frac{1}{2tk}(\|x^{(0)} - x^{(*)}\|_2^2) \tag{3}$$

1.2 Question 1.2[2]

Design of the algorithm For simplicity, we assume $B/R \in \mathbb{Z}^+$, if not, we can use $\lfloor \cdot \rfloor$. The algorithm is for $T = B/R - 1$, rent for the first $\min\{T, D\}$ days. If $D = T + 1$, then buy the ski board and finish the game.

Competitive Ratio The competitive ratio of this algorithm is $\frac{TR+B}{\min\{(T+1)R,B\}}$.

Proof of the competitive ratio

Proof. All the deterministic follows that rent for x days and then buy on day $x+1$. The central point thus lies in the design of x. But first, let's consider for any algorithm following the above regime, when does the maximum cost or the worst case cost occur?

It occurs on day $x+1$. For all days $D \geq x+1$, the proposed algorithm costs exactly 0 while the optimal cost is at least 0 since it may not have bought the ski board. For days $D < x+1$, the cost is increasing with D. Finally, let's compare $D = x+1$ and $D = x$. If $xR > B$, then the denominators of the ratios are both B while the numerator of $x+1$ is larger. If$(x+1)R \leq B$, then the ratio are both 1. If $xR \leq B$ while $(x+1)R > B$, then the ratio of x is 1 but the ratio of $x+1$ is $\frac{xR+B}{B} > 1$.

So, how to design this x? The best competitive ratio is achieved when $x+1 = B/R$. If $x+1 > B/R$, then decreasing T doesn't change the denominator but decreases the numerator. If$x+1 < B/R$, then ratio is $1 + \frac{B/R-1}{x}$, then increasing x will decrease if x is increased. □

1.3 1.3[3]

first bullet point Denote τ such that X_τ is the first that $X_\tau \geq \lambda$ and $p = \mathbb{P}(\max_i X_i > \lambda)$. With probability $1-p$, the algorithm accepts nothing. So with probability $1-p$, the algorithm accepts nothing at round i. So $\forall x > \lambda$, with probability $(1-p)\mathbb{P}(X_i > x)$, at round i, the

[2] Referenced https://www.cs.princeton.edu/ smattw/Teaching/521fa17lec24.pdf when solving the problem
[3] Referenced [Kleinberg and Weinberg, 2012] and https://livanos3.web.engr.illinois.edu/talks/intro-to-prophet-inequalities.pdf when solving the problem

algorithm accepts nothing in the previous rounds and $X_i > x$. Therefore, we have:

$$\begin{aligned}\mathbb{P}(X_\tau > x) &\geq \sum_i (1-p)\mathbb{P}(X_i > x)\\ &> (1-p)\mathbb{P}(\bigcup_i \{X_i > x\})\\ &= (1-p)\mathbb{P}(\max_i X_i > x)\end{aligned}$$

Finally,

$$\begin{aligned}ALG = \mathbb{E}(X_\tau) &= \int_0^\lambda \mathbb{P}(X_\tau > x)dx + \int_\lambda^\infty \mathbb{P}(X_\tau > x)dx\\ &\geq p\lambda + (1-p)\int_\lambda^\infty \mathbb{P}(\max_i X_i > x)dx\\ &\geq p\lambda + (1-p)\cdot \frac{1}{2}\mathbb{E}[\max_i X_i]\\ &= \lambda\end{aligned}$$

second bullet point Similarly, denote τ such that X_τ is the first that $X_\tau > \eta$, $(A)^+ = 0$ if $A \leq 0$ otherwise A and $\mathcal{E}_i$ as the event first $i-1$ rounds accept nothing.

$$\begin{aligned}ALG = \mathbb{E}[X_\tau] &= \eta\mathbb{P}[\max_i X_i > \eta] + \sum_i \mathbb{P}(\mathcal{E}_i)\mathbb{E}[(X_i - \eta)^+]\\ &= \frac{1}{2}\eta + \sum_i \mathbb{P}(\max_{j<i} X_j < \eta)\cdot \mathbb{E}[(X_i - \eta)^+]\\ &\geq \frac{1}{2}\eta + \mathbb{P}(\max_i X_i < \eta)\cdot \mathbb{E}[(X_i - \eta)^+]\\ &= \frac{1}{2}\eta + \frac{1}{2}\mathbb{E}[(X_i - \eta)^+]\\ &= \frac{1}{2}\mathbb{E}[(X_i - \eta)^+ + \eta]\\ &\geq \mathbb{E}[\max_i X_i]\end{aligned}$$

third bullet point Neither relationship holds necessarily. If $\max_i X_i = 1$ $w.p.1$, then $\lambda = \frac{1}{2} < \eta = 1$. If $\mathbb{P}(\max_i X_i = 1) = 1/9, \mathbb{P}(\max_i X_i = 2) = 2/9, \mathbb{P}(\max_i X_i = 3) = 3/9, \mathbb{P}(\max_i X_i = 36) = 3/9$, then $\eta = 3 < \lambda = 0.5 \cdot 37\frac{5}{9}$

The result should hold for γ between η and λ

2 Part II Question 2.1[4]

2.1 Empirical Result of infinitesimal setting

The empirical result of $OPT(\sigma)/ALG(\sigma) = 1.3565406037480383$.

2.2 Non-infinitesimal setting

The empirical result of non-infinitesimal setting is $OPT(\sigma)/ALG(\sigma) = 1.9222752122833546$. If non-infinitesimal setting, then we cannot use integral to approximate the summation in the denominator of $\frac{\Psi(Z)}{\sum_j \psi(z_j)w_j}$ where these notations follow from the original paper[Chakrabarty et al., 2008]. Instead, we can upper bound $\frac{\Psi(Z)}{\sum_j \psi(z_j)w_j}$ with $\frac{\Psi(Z)}{w_{\min}\sum_j \psi(z_j)}$. Then set a threshold of minimum accepted weight.

2.3 multi-dimensional setting

Define $p_{\min} = \left[p^1_{\min}, ..., p^n_{\min}\right]$ and $p_{\max} = \left[p^1_{\max}, ..., p^n_{\max}\right]$ and define the threshold function $\Psi(z) \doteq (p_{\max}e/p_{\min})^z(p_{\max}/e)$. The algorithm is whenever an item j arrives with a weight vector $w_i = \left[w^1_i, ..., w^n_i\right]$ and a value vector $v_j = \left[v^1_j, ..., v^n_j\right]$, denote the filled fraction of bin $i : B_i$ as z_i and $E_j = \{i \in [n] : \frac{v_i}{w_i} \geq \Psi(z_i)\}$ then assign the item to bin $\max_{i\in E_j} \frac{v_i}{w_i}$ as long as it doesn't overfill the capacity. If $E_j = \emptyset$, then discard it.

Proposition 2.1. *The competitive ratio of this algorithm is* $\frac{OPT(\sigma)}{ALG(\sigma)} \leq 2 + \frac{1}{\log(p_{\max}/p_{\min})}$

Proof. Assume the maximum fraction filled for bin i is $Z_i \in (0, 1]$ and $W_i = B_iZ_i$. Let S_i, S^*_i denote the choice of the online algorithm and optimum algorithm respectively. Let $S = \bigcup_i S_i, S^* = \bigcup_i S^*_i$. Then $ALG(\sigma) = \sum_i v_i(S_i), OPT(\sigma) = \sum_i v_i(S^*_i))$. Separate $S^*\S_i$ into X^*_i and Y^*_i where $X^*_i = \{j : E_j = \emptyset\}$ and Y^*_i are those assigned to other bins by the online algorithm. Since the density of Y^*_i are larger in other bins which means the online algorithm could get more by assigning items in Y^*_i to other bins, the following inequality holds:

$$\sum_i v_i(Y^*_i) \leq \sum_i v_i(S_i) = ALG(\sigma)$$

[4] code is available at https://github.com/pseudoinvchou/CMPUT675Assignment

Therefore, we have:

$$
\begin{aligned}
OPT(\sigma) &= \sum_i v_i(S_i^*) \\
&= \sum_i v_i(Y_i^*) + (v_i(S_i^*) - v_i(Y_i^*)) \\
&\leq \sum_i v_i(S_i) + (v_i(S_i^*)) \\
&= \sum_i v_i(S_i) + v_i(S^* \cap S_i) + v_i(S_i^* \setminus S_i) \\
&\leq \sum_i v_i(S_i) + v_i(S^* \cap S_i) \\
&\leq \sum_i (v_i(S_i) + \sum_{j \in S \cap S^*} \Psi(z_{i,j}) w_{i,j} + \Psi(Z_i)(B_i - W_i)) \\
&\leq \sum_i v_i(S_i) + \Psi(Z_i) B_i
\end{aligned}
$$

Also, $ALG(\sigma) = \sum_i v_i(S_i) \geq \sum_i \sum_j \Psi(z_{ij}) w_{ij} = \sum_i B_i \sum_j \Psi(z_{ij}) \Delta z_{ij}$ where $\Delta z_{ij} = z_{i,j+1} - z_{ij}$. Putting them together, we have

$$
\frac{OPT(\sigma)}{ALG(\sigma)} \leq 1 + \frac{\sum_i \Psi(Z_i) B_i}{\sum_i B_i \sum_j \Psi(z_{ij}) \Delta z_{ij}}
$$

For each ratio $\frac{\Psi(Z_i) B_i}{B_i \sum_j \Psi(z_{ij}) \Delta z_{ij}}$, similar to the 1-dimensional setting, it is upper bounded by $\frac{1}{\log(p_{\max}/p_{\min})+1}$. Therefore, the whole competitive ratio is upper bounded by $2 + \frac{1}{\log(p_{\max}/p_{\min})}$ $\square$

The implementation is in multi-dim.ipynb

References

D. Chakrabarty, Y. Zhou, and R. Lukose. Online knapsack problems. In *Workshop on internet and network economics (WINE)*, 2008.

R. Kleinberg and S. M. Weinberg. Matroid prophet inequalities. In *Proceedings of the forty-fourth annual ACM symposium on Theory of computing*, pages 123–136, 2012.