PHBS WORKING PAPER SERIES

Adjustable Regret for Continuous Control of Conservatism and Competitive Ratio Analysis

Yingjie Lan
Peking University

September 2021

Working Paper 20210907

Abstract

A major issue of the increasingly popular robust optimization is the tendency to produce overly conservative solutions. This paper deals with this by proposing a new parameterized robust criterion that is flexible enough to offer fine-tuned control of conservatism. The proposed criterion also leads to a new approach for competitive ratio analysis, which can reduce the complexity of analysis to the level for the minimax regret analysis. The properties of this new criterion are studied, which facilitates its applications, and validates the new approach for competitive ratio analysis. Finally, the criterion is applied to the well studied robust one-way trading problem to demonstrate its potential in controlling conservatism and reducing the complexity of competitive ratio analysis.

Keywords: Robust optimization, decision criteria, over-conservatism, online optimization, competitive ratio analysis, one-way trading

Peking University HSBC Business School
University Town, Nanshan District
Shenzhen 518055, China

Adjustable Regret for Continuous Control of Conservatism and Competitive Ratio Analysis

Yingjie Lan
HSBC Business School, Peking University
ylan@pku.edu.cn

May 12, 2021

Abstract

A major issue of the increasingly popular robust optimization is the tendency to produce overly conservative solutions. This paper deals with this by proposing a new parameterized robust criterion that is flexible enough to offer fine-tuned control of conservatism. The proposed criterion also leads to a new approach for competitive ratio analysis, which can reduce the complexity of analysis to the level for the minimax regret analysis. The properties of this new criterion are studied, which facilitates its applications, and validates the new approach for competitive ratio analysis. Finally, the criterion is applied to the well studied robust one-way trading problem to demonstrate its potential in controlling conservatism and reducing the complexity of competitive ratio analysis.

Keywords: robust optimization; decision criteria; over-conservatism; online optimization; competitive ratio analysis; one-way trading

1 Introduction

Robust optimization is a popular method for decision making under uncertainty, which changes the values of parameters in the problem. Another well known method for such task is stochastic programming, which demands a probability distribution to describe the uncertainty. In practice, however, it is often difficult to correctly estimate this distribution, due to reasons such as lack of data, inaccurate data, or nonstationarity of the stochastic process. A probability distribution is not needed in robust optimization, which only requires an uncertainty set that contains the possible scenarios. Thus robust optimization does not assume risk neutrality, as no expectations can be calculated, and it requires less information or historic data to apply, which makes it applicable in more situations. Successful applications of robust optimization are widely seen in many fields, the interested reader is referred to [BTEGN09] and [BBC11].

However, a major issue in robust optimization is that of over-conservatism, meaning that too much performance is sacrificed for the sake of robustness. The robustness of a solution pertains to maintaining its feasibility or a certain level of performance amid the uncertain realizations of scenarios, where the former is feasibility robustness and the latter performance robustness. Over-conservatism has been a very important topic in piorneering works in robut optimization such as [BTEGN09]. The literature on mitigating over-conservatism can be grouped by the component of the robust optimization model being worked upon, namely, (i) the uncertainty set of scenarios, (ii) the robust criterion to evaluate solutions, and (iii) the set of robust feasible solutions.

The method of uncertainty budget of [BS04] belongs to group (i) as it puts a special budget constraint on the uncertainty set so that some extreme scenarios may be excluded from consideration. This way it allows a level of flexibility in choosing the tradeoff between robustness and performance without significantly increasing the computational complexity. However, the recommended solution with a tight budget can be infeasible under an extreme scenario that is excluded by the budget constraint. Therefore, it is more suitable for applications where it is tolerable for an implemented solution to violate constraints in some extreme scenarios, but not for mission critical applications, such as spaceship design, where all extreme scenarios must be fully considered to avoid disasters.

In group (ii), criteria with different levels of conservatism have been proposed. For clean presentation, reward maximization is assumed, as it is similar for cost minimization. The maximin criterion is initially developed by [Neu28] for game theory. It looks for a robust solution by maximizing the worst scenario reward. This criterion certainly appeals to conservative decision makers who love solutions with a worst case performance guarantee. It is usually highly tractable and can produce acceptible solutions similar to those by stochastic programming in many applications. However, it only focuses on worst scenarios and sometimes fails to take advantage of favorable scenarios, and is likely to sacrifice too much performance for robustness. Trying to alieviate such extreme converatism, [Hur51] proposes a criterion that evaluates a solution by "a weighted sum of its worst and best possible outcomes". Although the Hurwicz criterion often gives reasonable results, it could also lead to quite illogical answers, see [GW14] for a detailed analysis with examples. Such inconsistency makes it difficult to manage and control the degree of conservatism in a reliable way.

The minimax regret criterion by [Sav51] is based on the idea of regret or opportunity loss. The regret of a solution is the difference between its reward realized in a scenario and the best possible reward in that same scenario. The maximum regret across all possible scenarios is a guarantee of worst regret by a solution, and a solution with the best or minimal regret guarantee is chosen. In the competitive ratio criterion, the ratio instead of the difference is considered, therefore it is also known as the relative minimax regret criterion, whereas the regret by difference is known as absolute. Both criteria are still considered as conservative, but they are less than the maximin criterion. For other less known criteria, the reader is referred to [CSCS$^+$14].

The methods in group (iii) screens out overly conservative solutions from the robust feasible set beforehand, such as the p-robustness method by [Sny06]. This method excludes solutions whose worst regret exceeds an upper limit, where both relative and absolute regret may be used. The choice of the limit must be careful: too tight a limit may disqualify all possible solutions, while too loose a limit can not screen out overly conservative ones. This method is not so popular in practice as it may add a constraint for each scenario, which soon becomes intractable. It is also noted by [Sny06] that it can be very difficult to determine if a limit renders a problem feasible or not in some applications.

This paper introduces a new robust criterion to mitigate the conservatism without tampering with the uncertainty set, avoiding the disadvantage of the uncertainty budget method. Unlike the Hurwicz criterion, this new criterion introduces a parameter to recalibrate the benchmark against which the regret is defined, thus it may be called the adjustable regret criterion (ARC). As no artificial constraints are introduced, it is free from the problems faced by the p-robustness method. Though many criteria for robust optimization with different levels of conservatism have been proposed in the literature, the choices are very limited, which can not provide fine control of conservatism. By adjusting the introduced parameter continuously, the ARC can provide a continuum of choices that includes some of the most widely used robust criteria. Such a continuum of choices provides a basis for a framework of analysis that enables fine control of conservatism.

The contribution of this paper is threefold. First, a new criterion named ARC is proposed with its framework of analysis that offers continuous control of conservatism of robust optimization, which is of both theoretical and practical interest. Second, the properties of ARC are investigated, deepening our theoretical understanding and facilitating its practical applications. And finally, a new approach for competitive ratio analysis is developed based on ARC. The competitive ratio analysis originates from algorithm designs in computer science, and is now widely applied in many fields. However, for some problems the competitive ratio criterion is much more difficult to analyze than the minimax regret criterion. See [LGBK08] for an example with both criteria on the same problem where the minimax regret analysis turns out to be the much easier one. This new approach for competitive ratio analysis can significantly reduce the complexity of analysis to a comparable level as that of the minimax regret analysis.

The rest of this paper is arranged as follows. Section 2 gives robust formulations with ARC for problems under uncertainty. The properties of ARC are studied in Section 3 to facilitate its applications in robust optimization. The novel approach for competitive ratio analysis based on a clever use of ARC is also discussed. In section 4 the well-studied robust one-way trading problem is employed to demonstrate the application of ARC with its flexibility and potential. Finally, in section 5 conclusions are drawn with future research directions.

2 Formulation

This section formulates ARC in a framework conducive to both theoretical analysis and practical applications. The framework is first introduced in a single stage setting, and then extended to multistage problems.

In a single stage problem, an action x is first taken from the set of all robustly feasible actions X, then a scenario ω is realized from all possible scenarios Ω. The set X can be described as $X = \{x : \forall\omega \in \Omega\ f_{i\in I}(x,\omega) = 0, g_{j\in J}(x,\omega) \geq 0\}$, with index set I and J for the various constraints, whose parameters are affected by ω. Integral constraints can be added here, so that X does not have to be continuous. The final reward depends on both x and ω, and it is given by the reward function $r(x,\omega)$. Let $r^*(\omega) = \max_{x\in X} r(x,\omega)$ denote the ex post optimal reward after ω is realized. For a disasterous scenario ω, if there is a way to avoid the risk entirely, such as clearing all holdings in the stock market, then $r^*(\omega) \geq 0$ can be maintained for all $\omega \in \Omega$. As for applications where even the best action can not avoid a net loss ($r^*(\omega) < 0$) in a terrible scenario, a reference point of performance can be provided by a function $f(\omega)$, and the relative reward $r_+(x,\omega) = r(x,\omega) - f(\omega)$ that indicates the reduced loss above the reference may be employed instead. The function $f(\omega)$ is chosen according to the context, for example, $f(\omega) = \min_{x\in X} r(x,\omega)$ or simply $f(\omega) = \min_{\omega\in\Omega} r^*(\omega)$. As long as $f(\omega) \leq r^*(\omega)$, the $r^*(\omega)$ based on $r_+(x,\omega)$ restores the condition of $r^*(\omega) \geq 0$.

In order to enable continuous control of conservatism, a parameter $\beta \in [0,\infty)$ is introduced to moderate $r^*(\omega)$ and produce a benchmark performance of $\beta r^*(\omega)$. The regret in ARC for action x after the realization of scenario ω is defined as $D(x,\omega;\beta) = \beta r^*(\omega) - r(x,\omega)$. Note that ω is not known at the time of decision on x, the maximal regret $\bar{D}(x;\beta) = \max_{\omega\in\Omega} D(x,\omega;\beta)$ is defined as a way to evaluate x, which can be interpreted as a guarantee of worst case regret. The ARC then chooses an x to have the best guarantee of

$$D(\beta) = \min_{x\in X} \bar{D}(x;\beta) = \min_{x\in X} \max_{\omega\in\Omega} D(x,\omega;\beta).$$

Some intuition is offered on how β helps enable continuous control of conservatism. The maximal regret $\bar{D}(x;\beta)$ can be seen as a measure of how much the performance of x falls behind the β-moderated benchmark performance. Therefore the ARC ends up choosing an x whose performance most closely follows that benchmark. With the assumption of $r^*(\omega) \geq 0$ for all $\omega \in \Omega$, the benchmark gets more aggressive as β increases, hence β may be called the parameter of aggressiveness here. Intuitively, as the benchmark becomes more aggressive as β increases, the solution that most closely follows it will also become so.

This criterion unifies the well-known, seemingly unrelated robust criteria into a continuum as β takes on different values. At $\beta = 0$, it is the maximin criterion. It then becomes the competitive ratio criterion if β takes a special value between 0 and 1 (more details on this later). Then at $\beta = 1$ it is the minimax regret criterion, and finally it transforms into the maximax criterion as $\beta \to \infty$.

The formulation developed so far easily extends to multistage problems, which call for sequential decision making under uncertainty. Let $t = 1, \cdots, T$

label the sequential stages, with smaller t for an earlier stage. The decision variable x now consists of T subverctors $(x_1, \cdots, x_T)$, with x_t (the stage decision) corresponding to the decision in stage t. Likewise a whole scenario now consists of sub-scenarios or stage scenarios for each stage: $\omega = (\omega_1, \cdots, \omega_T)$. Without loss of generality, in a stage of standard formulation, the stage decision x_t is first implemented, then the stage scenario ω_t is always realized afterwards. For applications where a stage scenario is first realized in the very beginning before any actions are taken, a dummy decision with only one choice of action (i.e. to participate in the decision process) can be inserted in the very beginning to transform to the standard formulation. Therefore the standard formulation helps simplify discussions, but the results are general nevertheless.

Just as in multistage stochastic programming (MSP), there is an implicit assumption: the realization of scenarios is independent of decisions, in other words, the decision maker can not influence how the scenario develops. All scenarios in Ω makes up a scenario tree, which has the stage scenarios ω_t as nodes such that a path from the root (representing the state right before the first stage) to a leaf node of some ω_T defines a scenario by sequencing the nodes of stage scenarios in the path to get $(\omega_1, \cdots, \omega_T)$. As the stage scenarios realize themselves, one moves along a particular path in the scenario tree. The possible future scenarios faced by the decision maker at a certain node in the path correspond to the ever shrinking sub-tree of scenarios from the current node. Therefore in stage t, only scenarios that do agree with the partial scenario $\omega_{1:t}$ revealed before stage t, as given by $\omega_{1:t} = (\omega_1, \cdots, \omega_{t-1})$, can be included in the set of future scenarios: $\Omega(\omega_{1:t}) = \{\omega' \in \Omega : \omega'_{1:t} = \omega_{1:t}\}$.

The same notion of nonanticipativity as in MSP, which requires that decisions must occur before observations in the later stages, is treated here by having stage decisions determined by what is already observed in the earlier stages. The stage decision x_t depends on the knowledge of the current history of both stage decisions and stage scenarios in the earlier stages. Let $x_{1:t} = (x_1, \cdots, x_{t-1})$ be the partial sequence of stage decisions before stage t. Note that the set of robustly feasible actions in the current stage t depends not only on $x_{1:t}$, but also on $\omega_{1:t}$, as $\Omega(\omega_{1:t})$ determines the possible values for parameters in the constraints defining X. Therefore let $X_t(h_t)$ denote the set of robustly feasible stage actions x_t for stage t, with $h_t = (x_{1:t}, \omega_{1:t})$ as a shorthand.

The rewards may be accrued over time or may be received at once in the end, so let $r(x, \omega)$ denote the total reward over all stages. Let $r^*(\omega) = \max_{x \in X(\omega)} r(x, \omega)$ be the ex post optimal reward, where $X(\omega) = \{x \in X | x_t \in X_t(x_{1:t}, \omega_{1:t}), t = 1, \cdots, T\}$ is the set of all actions compatible with scenario ω. At the end of the last stage, the complete history $h_{T+1} = (x, \omega)$ is known, and the regret is

$$D_T(h_{T+1}; \beta) = \beta r^*(\omega) - r(x, \omega). \tag{1}$$

To evaluate x_t in the context of h_t at $t = T$, a guarantee of worst regret is

$$\bar{D}_t(x_t, h_t; \beta) = \max_{\omega_t \in \Omega_t(\omega_{1:t})} D_t(h_{t+1}; \beta), \tag{2}$$

where h_{t+1} is formed by appending x_t and ω_t to $x_{1:t}$ and $\omega_{1:t}$ respectively. An optimal stage action x_t is chosen to minimize the guarantee of regret

$$\begin{aligned} D_{t-1}(h_t;\beta) &= \min_{x_t \in X_t(h_t)} \bar{D}_t(x_t, h_t; \beta), \\ &= \min_{x_t \in X_t(h_t)} \max_{\omega_t \in \Omega_t(\omega_{1:t})} D_t(h_{t+1}; \beta). \end{aligned} \tag{3}$$

The definition in (3) can be applied recursively for $t = T, \cdots, 1$ backwards, which gives a plain formulation that incorporates nonanticipativity. Note that when $t = 1$, there is no history in h_1, so let $D(\beta) = D_0(h_1; \beta)$, which is the best guarantee of regret for the entire problem.

An alternative formulation is based on policies. A policy π can be described as a sequence of functions $\pi = \{\pi_t : h_t \to X_t(h_t), t = 1, 2, \cdots, T\}$ to make stage decisions according to $x_t = \pi_t(h_t)$, which automatically takes care of nonanticipativity. Note that $X_t(h_t)$ can be replaced by the set of probabilistic mixtures of elements in $X_t(h_t)$ to allow for random policies, but the discussions here focus on deterministic policies for the sake of clarity. The regret under a policy π is defined as follows. It is possible to start applying a policy from stage t on with an arbitrary history h_t. As there is nothing left to do, the regret in the end with a full history $h_{T+1} = (x, \omega)$ is simply

$$D_T^\pi(h_{T+1}; \beta) = \beta r^*(\omega) - r(x, \omega). \tag{4}$$

For $t = T, \cdots, 1$ the regret is defined backwards and recursively by

$$D_{t-1}^\pi(h_t; \beta) = \max_{\omega_t \in \Omega_t(\omega_{1:t})} D_t^\pi(h_{t+1}^\pi; \beta), \tag{5}$$

where $h_{t+1}^\pi = ((x_{1:t}, \pi_t(h_t)), (\omega_{1:t}, \omega_t))$ denote the history evolution under π. To compute the overall regret $D^\pi(\beta) = D_0^\pi(h_1, \beta)$ (since h_1 is empty), simply apply (5) recursively to have

$$\begin{aligned} D^\pi(\beta) &= \max_{\omega_1 \in \Omega_1(\omega_{1:1})} D_1^\pi(h_2^\pi; \beta) \\ &= \max_{\omega_1 \in \Omega_1(\omega_{1:1})} \max_{\omega_2 \in \Omega_2(\omega_{1:2})} D_2^\pi(h_3^\pi; \beta) \\ &= \max_{\omega_1 \in \Omega_1(\omega_{1:1})} \cdots \max_{\omega_T \in \Omega_T(\omega_{1:T})} D_T^\pi(h_{T+1}^\pi; \beta) \\ &= \max_{\omega \in \Omega} D_T^\pi(h_{T+1}^\pi; \beta) \end{aligned} \tag{6}$$

Let Π be the set of all policies, and $r^\pi(\omega) = r(\pi(\omega), \omega)$, where $\pi(\omega)$ is the action by policy π in scenario ω from start to end. The policy-based formulation is given by

$$\min_{\pi \in \Pi} D^\pi(\beta) = \min_{\pi \in \Pi} \max_{\omega \in \Omega} \beta r^*(\omega) - r^\pi(\omega). \tag{7}$$

3 Properties

The analysis starts by establishing the correspondence and equivalence between the two formulations. It is assumed that the min and max operators in the formulations are well defined so that there is always an optimal solution.

Theorem 3.1 (Correspondence)**.** *The plain formulation and the policy-based formulation are equivalent in that (i) with an arbitrary history h_t there is*

$$D_{t-1}(h_t;\beta) = D^{\pi^*}_{t-1}(h_t;\beta), \ \textit{for } t = 1, \cdots, T+1, \tag{8}$$

for any optimal policy π^*, and (ii) an optimal policy π^* is constructed by

$$\pi^*_t(h_t) = \operatorname*{argmin}_{x_t \in X_t(h_t)} \max_{\omega_t \in \Omega_t(h^\omega_t)} D_t(h_{t+1};\beta), t = 1, \cdots, T, \tag{9}$$

where the argmin *operator arbitrarily takes one when there are many minimizers.*

Proof. It is clear that (8) trivially holds for $t = T+1$. For $t \le T$, recall (3) and proceed as follows

$$\begin{aligned} D_{t-1}(h_t;\beta) &= \min_{x_t \in X_t(h_t)} \max_{\omega_t \in \Omega_t(\omega_{1:t})} D_t(h_{t+1};\beta) \\ &= \max_{\omega_t \in \Omega_t(\omega_{1:t})} D^{\pi^*}_t(h^{\pi^*}_{t+1};\beta) \\ &= D^{\pi^*}_{t-1}(h_t;\beta), \end{aligned}$$

where the second equality comes by (9), and the last equality comes by (5).

It remains to prove that π^* is optimal to (7) by showing for an arbitrary $\pi \in \Pi$ there is

$$D_{t-1}(h_t;\beta) \le D^{\pi}_{t-1}(h_t;\beta), \tag{10}$$

for $t = 1, 2, \cdots, T+1$ via backward induction on t. As the initial step, it trivially holds for $t = T+1$. For the induction step, assume that (10) holds for $t+1$: $D_t(h_{t+1};\beta) \le D^\pi_t(h_{t+1};\beta)$, then show (10) also holds for t. Recall (3) and replace $D_t(h_{t+1};\beta)$ with $D^\pi_t(h_{t+1};\beta)$ to have

$$\begin{aligned} D_{t-1}(h_t;\beta) &\le \min_{x_t \in X_t(h_t)} \max_{\omega_t \in \Omega_t(\omega_{1:t})} D^\pi_t(h_{t+1};\beta) \\ &\le \max_{\omega_t \in \Omega_t(\omega_{1:t})} D^\pi_t(h^\pi_{t+1};\beta) \\ &= D^\pi_{t-1}(h_t;\beta), \end{aligned}$$

where the second inequality comes by having $x_t = \pi_t(h_t)$, and the last equality comes by (5). Therefore (10) holds for all t by backward induction, and π^* is indeed an optimal policy. □

It is handy to have both formulations: the plain formulation is more useful in solving problems for practical applications, while the policy-based formulation facilitates theoretical analysis. Also note that there is $D^{\pi^*}(\beta) = D(\beta)$ by (8) with $t = 1$.

Proposition 3.2 (Continuity)**.** *The best guarantee of worst regret $D(\beta)$ is a continuous function in β.*

Proof. Let $f_\pi^\omega(\beta) = \beta r^*(\omega) - r^\pi(\omega)$, which is a continuous function in β. Thus $f_\pi^*(\beta) = \max_{\omega\in\Omega} f_\pi^\omega(\beta)$, a point-wise max of continuous functions, is also continuous. Likewise, $f^*(\beta) = \min_{\pi\in\Pi} f_\pi^*(\beta)$ is also continuous. By Theorem 3.1 and the policy-based formulation in (7), there is $D(\beta) = f^*(\beta)$. □

Proposition 3.3 (Slope Bounds). *For $0 \le \beta_1 < \beta_2$, let $\pi_i^*, i \in \{1,2\}$ be an optimal policy when $\beta = \beta_i$, and $\omega_{ij}^* = \text{argmax}_{\omega\in\Omega} \beta_i r^*(\omega) - r^{(\pi_j^*}(\omega), i, j \in \{1,2\}$, then there is*

$$r^*(\omega_{21}^*) \ge \frac{D(\beta_2) - D(\beta_1)}{\beta_2 - \beta_1} \ge r^*(\omega_{12}^*). \tag{11}$$

Proof. By the definition of π_2^* and ω_{12}^*, as well as Theorem 3.1, there is

$$\begin{aligned} D(\beta_1) &= \min_{\pi\in\Pi}\max_{\omega\in\Omega} \beta_1 r^*(\omega) - r^\pi(\omega) \\ &\le \max_{\omega\in\Omega} \beta_1 r^*(\omega) - r^{\pi_2^*}(\omega) \\ &= \beta_1 r^*(\omega_{12}^*) - r^{\pi_2^*}(\omega_{12}^*). \end{aligned}$$

And $D(\beta_2) = \max_{\omega\in\Omega} \beta_2 r^*(\omega) - r^{\pi_2^*}(\omega) \ge \beta_2 r^*(\omega_{12}^*) - r^{\pi_2^*}(\omega_{12}^*)$. Therefore $D(\beta_2) - D(\beta_1) \ge (\beta_2 - \beta_1)\ r^*(\omega_{12}^*)$. Similarly,

$$\begin{aligned} D(\beta_2) &= \min_{\pi\in\Pi}\max_{\omega\in\Omega} \beta_2 r^*(\omega) - r^\pi(\omega) \\ &\le \max_{\omega\in\Omega} \beta_2 r^*(\omega) - r^{\pi_1^*}(\omega) \\ &= \beta_2 r^*(\omega_{21}^*) - r^{\pi_1^*}(\omega_{21}^*). \end{aligned}$$

And $D(\beta_1) = \max_{\omega\in\Omega} \beta_1 r^*(\omega) - r^{\pi_1^*}(\omega) \ge \beta_1 r^*(\omega_{21}^*) - r^{\pi_1^*}(\omega_{21}^*)$. Therefore $D(\beta_2) - D(\beta_1) \le (\beta_2 - \beta_1)\ r^*(\omega_{21}^*)$. □

3.1 Convexity

The convexity of $D(\beta)$ is studied in this subsection. Note that $\beta r^*(\omega) - r^\pi(\omega)$ is linear in β, thus function

$$F(\beta;\pi) = \max_{\omega\in\Omega} \beta r^*(\omega) - r^\pi(\omega)$$

is convex in β for a given policy π. However, generally speaking, $D(\beta) = \min_{\pi\in\Pi} F(\beta;\pi)$ is not convex in β. In order for $D(\beta)$ to be convex, some special condition is needed. A weak condition for convexity is introduced first.

Lemma 3.4 (Convexity). *A continuous function $f(y)$ with a convex domain Y is convex if*

$$\forall y_1, y_2 \in Y\ \exists\lambda \in (0,1)\ f(\lambda y_1 + (1-\lambda)y_2) \le \lambda f(y_1) + (1-\lambda) f(y_2). \tag{12}$$

Proof. By contradiction. Assume $f(y)$ is not convex, then there exists $y_1, y_2 \in Y$ and $\lambda \in (0,1)$ such that $g(\lambda) > 0$, where $g(k) = f(y(k)) - (kf(y_1) + (1-k)f(y_2))$ and $y(k) = ky_1 + (1-k)y_2$. As $g(k)$ is continuous with $g(0) = g(1) = 0$, there exists $k_1 = \max\{k \in [0,\lambda) : g(k) = 0\}, k_2 = \min\{k \in (\lambda, 1] : g(k) = 0\}$, such that $0 \le k_1 < \lambda < k_2 \le 1$, $g(k_1) = g(k_2) = 0$ and $\forall k' \in (k_1, k_2)\ g(k') > 0$.

Let $y_1' = y(k_1), y_2' = y(k_2)$, and $k' = \lambda' k_1 + (1-\lambda')k_2$ for a $\lambda' \in (0,1)$, then $y(k') = \lambda' y_1' + (1-\lambda')y_2'$. As $f(y_i') = k_i f(y_1) + (1-k_i)f(y_2), i = 1,2$ from $g(k_1) = g(k_2) = 0$, there is $\lambda' f(y_1') + (1-\lambda')f(y_2') = k' f(y_1) + (1-k')f(y_2)$. By $g(k') > 0$ there is $f(y(k')) > k' f(y_1) + (1-k')f(y_2)$, which implies $f(\lambda' y_1' + (1-\lambda')y_2') > \lambda' f(y_1') + (1-\lambda')f(y_2')$ for any $\lambda' \in (0,1)$, which contradicts (12). □

A policy π dominates another policy π' (denoted as $\pi \succeq \pi'$) if for all $\omega \in \Omega$ there is $r^{\pi}(\omega) \ge r^{\pi'}(\omega)$. Similarly, a scenario ω dominates another scenario ω' (denoted as $\omega \succeq \omega'$) if for all $\pi \in \Pi$ there is $r^{\pi}(\omega) \le r^{\pi}(\omega')$.

Definition 3.1 (Reward Convexity (RC))**.** *The set* Π *has the property of RC if there is*

$$\forall \pi_1, \pi_2 \in \Pi\ \exists \pi \in \Pi\ \exists \lambda \in (0,1)\ \forall \omega \in \Omega\ r^{\pi}(\omega) = \lambda r^{\pi_1}(\omega) + (1-\lambda) r^{\pi_2}(\omega). \quad (13)$$

An example with the RC property is when all randomized policies (a randomized policy uses a probability distribution to choose a deterministic policy before the game starts) are allowed and the reward is given by the expected reward.

Definition 3.2 (Reward Dominance Convexity (RDC))**.** *The set* Π *has the property of RDC if there is*

$$\forall \pi_1, \pi_2 \in \Pi\ \exists \pi \in \Pi\ \exists \lambda \in (0,1)\ \forall \omega \in \Omega\ r^{\pi}(\omega) \ge \lambda r^{\pi_1}(\omega) + (1-\lambda) r^{\pi_2}(\omega). \quad (14)$$

Clearly, if Π has the RC property, then it also has the RDC property. Another more sophisticated example is as follows. If $r(x,\omega)$ is concave in x and $X(\omega)$ is convex for all $\omega \in \Omega$, then (14) is satisfied. To see this, simply let $\pi(\omega) = \pi_1(\omega)/2 + \pi_2(\omega)/2$. By the concavity of $r(x,\omega)$ in x, there is $r(\pi(\omega),\omega) \ge r(\pi_1(\omega),\omega)/2 + r(\pi_2(\omega),\omega)/2$, and hence (14) is satisfied with $\lambda = 1/2$.

Note that both dominated policies and scenarios can be eliminated from consideration to simplify the analysis. Both the RC and RDC properties remain after scenario elimination. However, the RC property can be lost in policy elimination, while the RDC property remains untouched. Let $\hat{\Pi}$ be the set of all non-dominated policies in Π, so that any $\pi \in \Pi$ is dominated by a $\hat{\pi} \in \hat{\Pi}$.

Proposition 3.5 (Elimination Invariance)**.** *If the set* Π *has the RCD property, then* $\hat{\Pi}$ *also has the same property, and vice versa.*

Proof. Let $\pi_1, \pi_2 \in \hat{\Pi} \subseteq \Pi$, thus there exists $\pi \in \Pi$ and $\lambda \in (0,1)$ such that $r^{\pi}(\omega) \ge \lambda r^{\pi_1}(\omega) + (1-\lambda) r^{\pi_2}(\omega)$. As there is a $\hat{\pi} \in \hat{\Pi}$ such that $\hat{pi} \succeq \pi$. Hence $\hat{\Pi}$ has the RDC property. To show vice versa, let $\pi_1, \pi_2 \in \Pi$. Clearly there are

$\hat{\pi}_1, \hat{\pi}_2 \in \hat{\Pi}$ such that $\hat{\pi}_1 \succeq \pi_1, \hat{\pi}_2 \succeq \pi_2$. There exists $\hat{\pi} \in \hat{\Pi}$ and $\lambda \in (0,1)$ such that $r^{\hat{\pi}}(\omega) \geq \lambda r^{\hat{\pi}_1}(\omega) + (1-\lambda) r^{\hat{\pi}_2}(\omega) \geq \lambda r^{\pi_1}(\omega) + (1-\lambda) r^{\pi_2}(\omega)$. As $\hat{\pi} \in \Pi$, hence Π also has the RDC property. □

The RDC property even transfers to the set of optimal policies. Let $\Pi^*(\beta)$ be the set of all optimal policies for a given β.

Theorem 3.6. *If Π has the property of RDC, then $\Pi^*(\beta)$ also has the same property.*

Proof. Consider $\forall \pi_1^*, \pi_2^* \in \Pi^*(\beta) \subseteq \Pi$, by (14) there is a $\pi' \in \Pi$ such that

$$\exists \lambda \in (0,1)\ \forall \omega \in \Omega\ r^{\pi'}(\omega) \geq \lambda r^{\pi_1^*}(\omega) + (1-\lambda) r^{\pi_2^*}(\omega).$$

Now show that $\pi' \in \Pi^*(\beta)$ as follows:

$$\begin{aligned} D(\beta) &= \min_{\pi\in\Pi} \max_{\omega\in\Omega} \beta r^*(\omega) - r^{\pi}(\omega) \\ &\leq \max_{\omega\in\Omega} \beta r^*(\omega) - r^{\pi'}(\omega) = D^{\pi'}(\beta) \\ &\leq \max_{\omega\in\Omega} \beta r^*(\omega) - \left(\lambda r^{\pi_1^*}(\omega) + (1-\lambda) r^{\pi_2^*}(\omega)\right) \\ &\leq \lambda \left(\max_{\omega\in\Omega} \beta r^*(\omega) - r^{\pi_1^*}(\omega)\right) + \\ &\quad (1-\lambda)\left(\max_{\omega\in\Omega} \beta r^*(\omega) - r^{\pi_2^*}(\omega)\right) \\ &= \lambda D(\beta) + (1-\lambda) D(\beta) = D(\beta). \end{aligned}$$

Therefore $D^{\pi'}(\beta) = D(\beta)$ and so $\pi' \in \Pi^*(\beta)$. □

The property of RDC is a sufficient condition for $D(\beta)$ to be convex in β, but it is not a necessary one.

Theorem 3.7. *If Π has the property of RDC, then $D(\beta)$ is convex in β.*

Proof. Let π_i^* be an optimal policy for $\beta_i, i = 1, 2$. By (14) there exists $\pi' \in \Pi$ such that $\exists \lambda \in (0,1)\ \forall \omega \in \Omega\ r^{\pi'}(\omega) \geq \lambda r^{\pi_1^*}(\omega) + (1-\lambda) r^{\pi_2^*}(\omega)$. Let $\beta = \lambda\beta_1 + (1-\lambda)\beta_2$ and proceed as follows:

$$\begin{aligned} D(\beta) &= \min_{\pi\in\Pi} \max_{\omega\in\Omega} \beta r^*(\omega) - r^{\pi}(\omega) \\ &\leq \max_{\omega\in\Omega} \beta r^*(\omega) - r^{\pi'}(\omega) \\ &\leq \max_{\omega\in\Omega} \beta r^*(\omega) - \left(\lambda r^{\pi_1^*}(\omega) + (1-\lambda) r^{\pi_2^*}(\omega)\right) \\ &\leq \lambda \left(\max_{\omega\in\Omega} \beta_1 r^*(\omega) - r^{\pi_1^*}(\omega)\right) + \\ &\quad (1-\lambda)\left(\max_{\omega\in\Omega} \beta_2 r^*(\omega) - r^{\pi_2^*}(\omega)\right) \\ &= \lambda D(\beta_1) + (1-\lambda) D(\beta_2). \end{aligned}$$

Therefore $D(\beta)$ is convex in β by Lemma 3.4. □

3.2 Competitive Ratio

In reward maximization applications, the competitive ratio can be defined as

$$\max_{\pi\in\Pi}\min_{\omega\in\Omega} r_\omega(\pi)/r^*_\omega. \tag{15}$$

Theorem 3.8 (CR Equivalence)**.** *Assume $r^*(\omega) > 0$ for all $\omega \in \Omega$, the β_0 that solves $D(\beta) = 0$ is exactly the competitive ratio, and the set of optimal policies for (7) is the same as that for (15).*

Proof. It needs to show for any $\pi^* \in \Pi^*(\beta_0)$ that π^* is an optimal solution to (15), and vice versa. By Theorem 3.1 and (7) there is

$$\begin{aligned}
&\begin{cases} 0 = \min_{\pi\in\Pi}\max_{\omega\in\Omega}\beta_0 r^*(\omega) - r^\pi(\omega) \\ \pi^* \in \operatorname{argmin}_{\pi\in\Pi}\max_{\omega\in\Omega}\beta^* r^*(\omega) - r^\pi(\omega) \end{cases}\\
\Leftrightarrow\ &\begin{cases} 0 = \max_{\omega\in\Omega}\beta_0 r^*(\omega) - r^{\pi^*}(\omega) \\ \forall\pi\in\Pi\ 0 \le \max_{\omega\in\Omega}\beta_0 r^*(\omega) - r^\pi(\omega) \end{cases}\\
\Leftrightarrow\ &\begin{cases} \exists\omega\in\Omega\ 0 = \beta_0 r^*(\omega) - r^{\pi^*}(\omega) \\ \forall\omega\in\Omega\ 0 \ge \beta_0 r^*(\omega) - r^{\pi^*}(\omega) \\ \forall\pi\in\Pi\ \exists\omega\in\Omega\ 0 \le \beta_0 r^*(\omega) - r^\pi(\omega) \end{cases}\\
\Leftrightarrow\ &\begin{cases} \exists\omega\in\Omega : \beta_0 = r^{\pi^*}(\omega)/r^*(\omega) \\ \forall\omega\in\Omega : \beta_0 \le r^{\pi^*}(\omega)/r^*(\omega) \\ \forall\pi\in\Pi\ \exists\omega\in\Omega\ \beta_0 \ge r^\pi(\omega)/r^*(\omega) \end{cases}\\
\Leftrightarrow\ &\begin{cases} \beta_0 = \min_{\omega\in\Omega} r^{\pi^*}(\omega)/r^*(\omega) \\ \forall\pi\in\Pi\ \beta_0 \ge \min_{\omega\in\Omega} r^\pi(\omega)/r^*(\omega) \end{cases}\\
\Leftrightarrow\ &\begin{cases} \beta^* = \max_{\pi\in\Pi}\min_{\omega\in\Omega} r^\pi(\omega)/r^*(\omega) \\ \pi^* \in \operatorname{argmax}_{\pi\in\Pi}\min_{\omega\in\Omega} r^\pi(\omega)/r^*(\omega) \end{cases}
\end{aligned}$$

As the reasoning can go in both directions, the theorem is established. □

Based on the result of Theorem 3.8, the next proposition gives the condition for the existence of a unique competitive ratio.

Proposition 3.9. *If $D(0) < 0$ then there is $r^*(\omega) > 0$ for all $\omega \in \Omega$, and there is a unique $\beta_0 \in (0, 1]$ such that $D(\beta_0) = 0$.*

Proof. Note that at $\beta = 0$ it becomes equivalent to the maximin criterion:

$$D(0) = \min_{\pi\in\Pi}\max_{\omega\in\Omega} -r^\pi(\omega) = -\max_{\pi\in\Pi}\min_{\omega\in\Omega} r^\pi(\omega).$$

Suppose there is a $\dot\omega$ such that $r^*(\dot\omega) \le 0$, then there is

$$-D(0) = \max_{\pi\in\Pi}\min_{\omega\in\Omega} r^\pi(\omega) \le \max_{\pi\in\Pi} r^\pi(\dot\omega) = r^*(\dot\omega) \le 0.$$

Therefore $D(0) \ge 0$, a contradiction! Thus there is $r^*(\omega) > 0$ for all $\omega \in \Omega$, so $D(\beta)$ strictly increases in β. Note that at $\beta = 1$ it is the minimax regret criterion, thus $D(1) \ge 0 > D(0)$, and the conclusion follows by the monotony and continuity of $D(\beta)$. □

4 One-way Trading

The one-way trading problem has been well studied under both the competitive ratio criterion (see [EYFKT01]) and the minimax regret criterion (see [WWLZ16]). Therefore it provides a great opportunity to show the advantages of ARC, which gives a much more general result with its degree of conservativism adjustable by β. The newly proposed approach to competitive ratio analysis involves only plain backward induction, without depending on acute intuitions and special insights as in [EYFKT01].

4.1 Problem Formulation

Consider the one-way trading problem to sell a total amount of fully divisible goods (like gasoline or steel) in a finite time horizon while the price fluctuates in the range of $[m, M]$. For comparable results, the tradition of dividing time into T discrete periods is followed. A fixed price $p_t \in [m, M]$ is revealed in each period $t = 1, \cdots, T$. The trader is a price-taker and must decide on the amount x_t to sell at the current price p_t in each period without knowing the future prices. The goal is to maximize the total sales revenue in the end.

It is helpful to adopt the notations in section 2. A scenario ω here simply corresponds to the prices $p = (p_1, \cdots, p_T)$ revealed over time, with $\omega_t = p_t$. As the prices are independent from each other, there is $\Omega_t(p_{1:t}) = [m, M]$, and $\Omega = [m, M]^T$. Without loss of generality, the total amount of goods is one unit, and the action is $x = (x_1, \cdots, x_T)$ with $X = \{x : \sum_{t=1}^{T} x_t = 1, x \geq 0\}$. For $t < T$ there is $X_t(h_t) = [0, q_t]$ where $h_t = (x_{1:t}, p_{1:t})$ and $q_t = 1 - \sum_{s=1}^{t-1} x_t$ is the remaining amount to sell given h_t, but in the last period there is $X_T(h_T) = [q_T, q_T]$ in order to sell everything. The reward is accumulated over time, so let $r_t = \sum_{s=1}^{t-1} p_s x_s$ be the rewards accumulated in h_t, the reward in the end is $r(x, p) = r_{T+1}$. Let $\hat{p}_t = \max\{p_s : s = 1, \cdots, t-1\}$ denote the highest price seen in h_t, then $r^*(p) = \max\{r(x, p) : \sum_{t=1}^{T} x_t = 1\} = \hat{p}_{T+1}$. At the end of the last stage (1) becomes

$$D_T(h_{T+1}; \beta) = \beta \hat{p}_{T+1} - r_{T+1}. \tag{16}$$

In this multistage problem it is natural to have periods coincide with stages, in which the uncertain price is first revealed, then an action is taken. Therefore it calls for a different formulation from the standard formulation in (3):

$$D_{t-1}(h_t; \beta) \quad = \quad \max_{p_t \in [m,M]} \min_{x_t \in X_t(h_t)} D_t(h_{t+1}; \beta), \tag{17}$$

but the difference is superficial: all of the results in section 3 remain valid.

4.2 Analytic Solution

The analysis starts from the last period T and reasons backwards. In the last period clearly there is $x_T = q_T$, and (17) becomes

$$D_{T-1}(h_T; \beta) = \max_{p_T \in [m,M]} \beta \max(\hat{p}_T, p_T) - (r_T + p_T q_T),$$

which is convex in p_T, and the maximizer is either $p_T = m$ or $p_T = M$. Define auxiliary functions that map a quantity $q \in [0, 1]$ to a price in $[m, M]$,

$$P_j(q) = (M - m)\left(1 - \frac{q}{\beta j}\right)^{+j} + m, j = 1, 2, \cdots,$$

where $y^{+j} = \max(0, y)^j$ denote the positive part of y raised to the j^{th} power. Let $P_n^-(y) = q$ be the inverse of $y = P_n(q)$ for $q \in [0, \beta j]$.

$$\begin{aligned} D_{T-1}(h_T; \beta) &= \max(\beta \hat{p}_T - R_T, \beta M - (r_T + M q_T)) \\ &= \max(\beta \hat{p}_T, \beta M - (M - m) q_T) - R_T \\ &= \beta \max(\hat{p}_T, P_1(q_T)) - R_T, \end{aligned}$$

where $R_t = r_t + m q_t$ for $t = 1, \cdots, T$ is the lower bound on r_{T+1} given h_t. Note that the special but trivial case of $\beta = 0$ is not included here. Continue on with (17) for $t = T - 1, \cdots, 1$, the result is obtained and presented as follows.

Theorem 4.1. *The best adjustable regret guarantee for the one-way trading problem for $t = 1, 2, \cdots, T$ is*

$$D_{t-1}(h_t; \beta) = \beta \max(\hat{p}_t, P_{1+T-t}(q_t)) - R_t, \tag{18}$$

and the optimal trading policy is $\pi_t^*(h_t, p_t) = q_t - q_{t+1}^*$, with $q_{T+1}^* = 0$ and

$$q_{t+1}^* = \min(q_t, P_n^-(\hat{p}_{t+1})), \ t = 1, \cdots, T - 1. \tag{19}$$

Proof. By backward induction. It is already verified for $t = T$, which completes the initial step. For the induction step, assume that (18) holds at $t + 1$ with $D_t(h_{t+1}; \beta) = \beta \max(\hat{p}_{t+1}, P_{T-t}(q_{t+1})) - R_{t+1}$, and show that it also holds at $t < T$ using (17). For the nested minimization in (17), let $n = T - t$ and

$$\begin{aligned} \bar{D}_t(h_t, p_t; \beta) &= \min_{x_t \in X_t(h_t)} D_t(h_{t+1}; \beta) \\ &= \min_{q_{t+1} \in [0, q_t]} \beta \max(\hat{p}_{t+1}, P_n(q_{t+1})) - R_{t+1}, \end{aligned} \tag{20}$$

with $q_{t+1} = q_t - x_t$, and $R_{t+1} = r_{t+1} + m q_{t+1}$. The derivative $\partial D_t(h_{t+1}; \beta)/\partial q_{t+1}$ involves the derivative of $P_n(q)$, which is

$$P_n'(q) = -\frac{M - m}{\beta}\left(1 - \frac{q}{\beta n}\right)^{+(n-1)} \leq 0.$$

By the monotony of $P_n(q)$, if $q_{t+1} \leq P_n^-(\hat{p}_{t+1})$ then $P_n(q_{t+1}) \geq \hat{p}_{t+1}$. Similarly $q_{t+1} > P_n^-(\hat{p}_{t+1})$ ensures $P_n(q_{t+1}) \leq \hat{p}_{t+1}$. Thus there is

$$D_t(h_{t+1}; \beta) = \begin{cases} \beta P_n(q_{t+1}) - R_{t+1} & q_{t+1} \leq P_n^-(\hat{p}_{t+1}) \\ \beta \hat{p}_{t+1} - R_{t+1} & q_{t+1} > P_n^-(\hat{p}_{t+1}) \end{cases} \tag{21}$$

$$\frac{\partial D_t(h_{t+1}; \beta)}{\partial q_{t+1}} = \begin{cases} p_t - m + \beta P_n'(q_{t+1}) & q_{t+1} < P_n^-(\hat{p}_{t+1}) \\ p_t - m & q_{t+1} > P_n^-(\hat{p}_{t+1}) \end{cases} \tag{22}$$

Note that in the first branch with $q_{t+1} < P_n^-(\hat{p}_{t+1})$, there is $p_t \le \hat{p}_{t+1} < P_n(q_{t+1}) \le -\beta P_n'(q_{t+1}) + m$, so $p_t - m + \beta P_n'(q_{t+1}) < 0$. And in the second branch with $q_{t+1} > P_n^-(\hat{p}_{t+1})$, there is $p_t - m \ge 0$. Therefore an optimal solution to (20) is (19), which from (21) gives

$$\bar{D}_t(h_t, p_t; \beta) = \beta P_n(q_{t+1}^*) - (r_{t+1} + m q_{t+1}^*). \tag{23}$$

Let $\bar{p}_t = \max(\hat{p}_t, P_n(q_t) \in [m, M]$, and from (17) there is

$$\begin{aligned} D_{t-1}(h_t; \beta) &= \max_{p_t \in [m,M]} \bar{D}_t(h_t, p_t; \beta) \\ &= \max \begin{pmatrix} \max_{p_t \in [m, \bar{p}_t]} \bar{D}_t(h_t, p_t; \beta) \\ \max_{p_t \in [\bar{p}_t, M]} \bar{D}_t(h_t, p_t; \beta) \end{pmatrix} \end{aligned} \tag{24}$$

For the branch with $p_t \in [m, \bar{p}_t]$ in (24), consider two cases: (i) $\bar{p}_t = \hat{p}_t \ge \breve{p}_t$ and (ii) $\bar{p}_t = \breve{p}_t > \hat{p}_t$. In case (i) there is $\hat{p}_{t+1} = \max(\hat{p}_t, p_t) = \hat{p}_t \ge P_n(q_t)$, therefore $P_n^-(\hat{p}_{t+1}) \le q_t$ and (19) simplifies to $q_{t+1}^* = P_n^-(\hat{p}_{t+1})$, thus $P_n(q_{t+1}^*) = \hat{p}_{t+1} = \bar{p}_t$. In case (ii) there is $\hat{p}_{t+1} \le P_n(q_t)$, therefore $P_n^-(\hat{p}_{t+1}) \ge q_t$ and (19) simplifies to $q_{t+1}^* = q_t$, thus $P_n(q_{t+1}^*) = P_n(q_t) = \bar{p}_t$. As in both cases there is $P_n(q_{t+1}^*) = \bar{p}_t$, from (23) there is $\bar{D}_t(h_t, p_t; \beta) = \beta \bar{p}_t - (r_{t+1} + m q_{t+1}^*) = \beta \bar{p}_t - r_t - p_t x_t^* - m q_{t+1}^*$, which is linear in p_t with a slope of $-x_t^* \le 0$ as $x_t^* = q_t - q_{t+1}^* \ge 0$. Thus $p_t^* = m$ is a maximizer, which gives

$$\max_{p_t \in [m, \bar{p}_t]} \bar{D}_t(h_t, p_t; \beta) = \beta \bar{p}_t - r_t - m q_t = \beta \bar{p}_t - R_t.$$

For the branch in (24) with $p_t \in [\bar{p}_t, M]$, as $p_t \ge \bar{p}_t \ge \hat{p}_t$, there is $\hat{p}_{t+1} = p_t \ge \bar{p}_t \ge P_n(q_t)$, thus $P_n^-(\hat{p}_{t+1}) \le q_t$ and (19) simplifies to $q_{t+1}^* = P_n^-(\hat{p}_{t+1})$. Therefore $P_n(q_{t+1}^*) = \hat{p}_{t+1} = p_t$, and (23) simplifies to $\bar{D}_t(h_t, p_t; \beta) = \beta p_t - r_t - p_t x_t^* - m q_{t+1}^* = \beta p_t - r_t - p_t(q_t - q_{t+1}^*) - m q_{t+1}^* = (\beta - q_t + q_{t+1}^*) p_t - m q_{t+1}^* - r_t = (\beta - q_t + q_{t+1}^*) P_n(q_{t+1}^*) - m q_{t+1}^* - r_t$. Now consider function

$$d(z) = (\beta - q_t + z) P_n(z) - mz - r_t, z \in [0, q_t].$$

The derivative is $d'(z) = (\beta - q_t + z) P_n'(z) + P_n(z) - m$. Note that $P_n(z) - m = -(\beta - z/n) P_n'(z)$, thus $d'(z) = (\beta - q_t + z) P_n'(z) - (\beta - z/n) P_n'(z) = (z + z/n - q_t) P_n'(z)$. As $P_n'(z) \le 0$, there is $d'(z) \ge 0$ when $z + z/n - q_t \le 0$, and $d'(z) \le 0$ when $z + z/n - q_t \ge 0$, hence $z^* = n q_t/(n+1)$ is a maximizer of $d(z)$, and

$$\begin{aligned} d(z^*) &= (\beta - q_t + z^*) P_n(z^*) - m z^* - r_t \\ &= (\beta - q_t + z^*)(M - m) \left(1 - \frac{z^*}{\beta n}\right)^{+n} + \beta m - m q_t - r_t \\ &= \beta\left(1 - \frac{q_t}{\beta(n+1)}\right)(M - m) \left(1 - \frac{q_t}{\beta(n+1)}\right)^{+n} + \beta m - R_t \\ &= \beta(M - m) \left(1 - \frac{q_t}{\beta(n+1)}\right)^{+(n+1)} + \beta m - R_t \\ &= \beta P_{n+1}(q_t) - R_t. \end{aligned}$$

Clearly, there is $\bar{D}_t(h_t, p_t; \beta) = d(P_n^-(p_t))$ for $p_t \in [\bar{p}_t, M]$, consider two cases. Case (i) $P_n(z^*) \geq \bar{p}_t$. As $P_n^-(M) = 0 \leq z^* \leq P_n^-(\bar{p}_t)$, there is $\max_{p_t \in [\bar{p}_t, M]} \bar{D}_t(h_t, p_t; \beta) = d(z^*)$. Therefore according to (24) there is

$$D_{t-1}(h_t; \beta) = \max(\beta \bar{p}_t - R_t, d(z^*)).$$

Case (ii) $P_n(z^*) < \bar{p}_t$. As $q_t \geq z^* \geq P_n^-(\bar{p}_t)$, there is

$$\begin{aligned} \max_{p_t \in [\bar{p}_t, M]} \bar{D}_t(h_t, p_t; \beta) &= \max_{p_t \in [\bar{p}_t, M]} \bar{D}_t(h_t, p_t; \beta) d(P_n^-(p_t)) \\ &= \max_{z \in [0, P_n^-(\bar{p}_t)]} d(z) \\ &\leq \max_{z \in [0, q_t]} d(z) = d(z^*). \end{aligned}$$

Direct comparison finds $P_n(z^*) \geq P_{n+1}(q_t)$, thus $\bar{p}_t \geq P_{n+1}(q_t)$. Therefore $d(z^*) = \beta P_{n+1}(q_t) - R_t \leq \beta \bar{p}_t - R_t$, and according to (24) there is

$$D_{t-1}(h_t; \beta) = \beta \bar{p}_t - R_t = \max(\beta \bar{p}_t - R_t, d(z^*)).$$

So in both cases there is $D_{t-1}(h_t; \beta) = \max(\beta \bar{p}_t - R_t, d(z^*))$. Note that $\bar{p}_t = \max(\hat{p}_t, P_n(q_t))$, and $P_n(q_t) \leq P_{n+1}(q_t)$, thus

$$\begin{aligned} D_{t-1}(h_t; \beta) &= \max(\beta \bar{p}_t - R_t, d(z^*)) \\ &= \max(\beta \bar{p}_t - R_t, \beta P_{n+1}(q_t) - R_t) \\ &= \beta \max(\bar{p}_t, P_{n+1}(q_t)) - R_t \\ &= \beta \max(\hat{p}_t, P_n(q_t), P_{n+1}(q_t)) - R_t \\ &= \beta \max(\hat{p}_t, P_{n+1}(q_t)) - R_t \end{aligned}$$

Therefore, as $n = T - t$, it is clear that (18) also holds for t. □

Corollary 4.2. *The adjustable regret guarantee for the one-way trading problem throughout all periods or stages is a convex function of β:*

$$D(\beta) = \beta(M - m)\left(1 - \frac{1}{\beta T}\right)^{+T} - (1 - \beta)m, \tag{25}$$

Proof. In the first period there is $q_1 = 1, r_1 = 0, \hat{p}_1 = m$. Use these in (18) and simplify to have the result. The convexity of $D(\beta)$ is a consequence of the reward convexity in the one-way trading problem and Theorem 3.7. □

Note that the result of [WWLZ16] is a special case of Theorem 4.1 and Corollary 4.2 with $\beta = 1$. Note that this general result is not appreciably harder to obtain than their special results.

Corollary 4.3 (Competitive Ratio)**.** *The competitive ratio defined in (15) for the one-way trading problem is the unique root β_0 of $D(\beta)$ as defined in (25).*

Proof. As $r^*(\omega) \geq m > 0$, it follows from Theorem 3.8 and Proposition 3.9. □

This is in perfect agreement with [EYFKT01], except that they define competitive ratio as the inverse of β_0. Their analysis heavily depends on insights of the worst case price paths, and is much more complicated than the plain analysis here that can deduce the insights as results.

5 Conclusion

This paper tackles the issue of over-conservatism in robust optimization by proposing the new ARC criterion based on the idea of adjustable regret. The ARC chooses a solution that most closely follows an adjustable benchmark, therefore as the benchmark becomes more aggressive, so is the chosen solution. The theoretical analysis of ARC is built upon a general framework that allows both a plain formulation and a policy-based formulation, both of which are equivalent and can deal with both single stage and multistage problems. Various theoretical properties of the adjustable regret, such as continuity, monotony, and convexity are studied. Based on ARC, a new approach for the competitive ratio analysis is discovered. Finally, the ARC framework is applied to the well studied robust one-way trading problem, which produces the results for both the competitive ratio and the minimax regret analysis.

The ARC is a new criterion, this paper only provides an initial theoretical study of its properties, and only presents one application for the purpose of illustration. It is believed that there are still many unknown properties for future study, and ARC can be used in many other applications as well. How to choose an appropriate β in real applications depends on the actual context, and is another worthy topic for future study. Note that in the current formulation of ARC, the benchmark is adjusted along a straight line through the origin. It is conceivable that this line of adjustment can be further customized for special applications, and it could even be a parameterized curve motivated by some practical rationale.

References

[BBC11] Dimitris Bertsimas, David B Brown, and Constantine Caramanis. Theory and applications of robust optimization. *SIAM review*, 53(3):464–501, 2011.

[BS04] Dimitris Bertsimas and Melvyn Sim. The price of robustness. *Operations research*, 52(1):35–53, 2004.

[BTEGN09] Aharon Ben-Tal, Laurent El Ghaoui, and Arkadi Nemirovski. *Robust optimization*. Princeton university press, 2009.

[CSCS+14] Amadeu Coco, Elyn Solano-Charris, Andréa Cynthia Santos, Christian Prins, and Thiago Ferreira de Noronha. *Robust optimization criteria: state-of-the-art and new issues*. PhD thesis, Université de Technologie de Troyes-UTT, 2014.

[EYFKT01] Ran El-Yaniv, Amos Fiat, Richard M Karp, and Gordon Turpin. Optimal search and one-way trading online algorithms. *Algorithmica*, 30(1):101–139, 2001.

[GW14] Helena Gaspars-Wieloch. Modifications of the hurwiczs decision rule. *Central European Journal of Operations Research*, 22(4):779–794, 2014.

[Hur51] Leonid Hurwicz. The generalized bayes minimax principle: a criterion for decision making under uncertainty. *Cowles Commission Discussion Paper, Statistics 335*, 1951.

[LGBK08] Yingjie Lan, Huina Gao, Michael O Ball, and Itir Karaesmen. Revenue management with limited demand information. *Management Science*, 54(9):1594–1609, 2008.

[Neu28] J v Neumann. Zur theorie der gesellschaftsspiele. *Mathematische annalen*, 100(1):295–320, 1928.

[Sav51] Leonard J Savage. The theory of statistical decision. *Journal of the American Statistical association*, 46(253):55–67, 1951.

[Sny06] Lawrence V Snyder. Facility location under uncertainty: a review. *IIE transactions*, 38(7):547–564, 2006.

[WWLZ16] Wei Wang, Liying Wang, Yingjie Lan, and Jean X Zhang. Competitive difference analysis of the one-way trading problem with limited information. *European Journal of Operational Research*, 252(3):879–887, 2016.