Optimal bidding: a dual approach

Carlos Pita
jampp.com

ABSTRACT

We propose a practical yet rigorous near-optimal bidding strategy for demand-side platforms (DSPs) that scales to thousands of advertising campaigns programmatically bidding in real-time (RTB) in billions of auctions per day. The strategy is logically derived from two different —stylized but realistic— constrained global profit maximization problems, so there are no *ad hoc* rules for pacing, pricing, etc. The approach relies on a few assumptions that we identify and analyze, providing a basis for further extensions to other kinds of problem/contract. It is expected to find a near-optimal solution by solving a convex relaxation of the original hard combinatorial problem. It is based on Lagrange duality so it has a sound, well-known, theoretical foundation. Optimal bids for first/second-price auctions can be cheaply computed in real-time given the shadow prices of the problem constraints; on the other hand, shadow prices are daily updated by a simple subgradient descent algorithm that converges logarithmically. The algorithm should be robust in the face of noisy real-life environments and of market seasonal oscillations and structural breaks. For a special case, we also offer an alternative derivation based on an intuitive continuous relaxation argument that reinforces our confidence in the general solution proposed here.

CCS CONCEPTS

• **Theory of computation** → *Convex optimization.*

KEYWORDS

Convex Optimization, Combinatorial Optimization, Duality, Subgradient Descent, RTB, Real Time Bidding, DSP, Demand-Side Platform

ACM Reference Format:
Carlos Pita. 2019. Optimal bidding: a dual approach. In *25th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, August 05–08, 2019, Anchorage, Alaska–USA.* ACM, New York, NY, USA, 6 pages.

1 INTRODUCTION AND PRELIMINARIES

Jampp is a demand-side platform (DSP) that manages thousands of advertising campaigns for hundreds of clients (advertisers). Advertising inventory coming from multiple sources is programmatically bought by real-time bidding (RTB) in billions of (mostly second-price) auctions per day. With each RTB auction certain information about a user (audience) currently visiting a web site or mobile app is made available for our bidder to infer[1], in a few milliseconds, the probabilities of getting different events of interest (clicks, app installs, in-app events, return-on-ad-spend, etc.) and to choose a campaign, an ad (banner) and a bid price on behalf of clients running campaigns that target the auction profile. As a firm operating in a market economy, Jampp maximizes profit subject to constraints imposed by the goals of its clients (*e.g.* restrictions on the volume and cost of purchases). An auction has to be weighted not only by the expected profit of assigning it to each campaign, but also by the expected impact of that assignment on the constraint set of each campaign; *i.e.* , the "shadow prices" of constraints have to be taken into account. Given the huge number of non-divisible auctions to be considered, this is a hard combinatorial optimization problem that we can only expect to approximately solve in a relaxed form. Finding a good relaxation of the constraints is our main task in what follows, but first we need to establish a formal setting.

Definition 1 (Campaign). *We have n advertising campaigns ($i, j \in [n] = \{1, \cdots, n\}$) competing for the RTB market.*

Definition 2 (Market). *The RTB market consists of m auctions ($k \in [m] = \{1, \cdots, m\}$), each one characterized by:*

- *$w_k(b)$: the probability of winning the auction by bidding the amount b.*
- *$c_k(b)$: the expected cost (or clearing price) of winning the auction by bidding the amount b.* [2]
- *$e_k(i, a)$: the probability of getting an event of interest given that we won the auction and displayed ad a in behalf of campaign i.*

Bids belong to the interval $(0, \overline{b}]$ while ads belong to the finite set $[\overline{a}] = \{1, \cdots, \overline{a}\}$. [3] For each auction, we have to decide whether we ignore it or, otherwise, we assign it to some campaign and choose a bid and an ad for it. This policy constitutes our bidding strategy:

Definition 3 (Bidding strategy). *A bidding strategy $x\colon [m] \to [n] \times [\overline{a}] \times (0, \overline{b}] \cup \{(0,0,0)\}$ is a mapping from auctions to vectors (i, a, b), where i is a campaign, a is an ad and b is a bid. As a convention, ignored auctions are mapped to the vector $(0, 0, 0)$. We call X the set of all possible strategies.*

We also define some functions that aggregate auctions across campaigns in order to simplify the exposition of optimization problems in the following sections.

Definition 4 (Aggregate functions). *Given a strategy x and a campaign i:*

- *$C_i(x) = \sum_{(k,(j,a,b))\in x \mid j=i} w_k(b) \cdot c_k(b)$ is the aggregate cost function.*
- *$E_i(x) = \sum_{(k,(j,a,b))\in x \mid j=i} w_k(b) \cdot e_k(j, a)$ is the aggregate event function.*

2 MBFP AND FBMP BIDDING PROBLEMS

With the above definitions in place, we can now state a first version of the bidding problem. In fact, Jampp daily deals with two different types of problem:

Definition 5 (MBFP: Maximum Budget, Fixed Price). *The client sets an upper bound $\bar{B}_i$ to the amount of money he wants to spend*

[1] The inference problem is interesting *per se* and far from trivial, but here we assume we already have proper estimates at hand and take those estimates as a given from now on in order to focus on the economical optimization problem. Our post "Learning from the RTB market" in `https://geeks.jampp.com/` gives further details about our approach to the learning problem.

[2] Which is always b for first-price auctions.

[3] Not all ads are available to every campaign, but here we can disregard this fact without loss of generality.

(the campaign budget) and a price $\bar{p}_i$ he will pay for each goal event. Formally:

$$\max_{x \in X} \sum_{i \in [n]} \bar{p}_i E_i(x) - C_i(x) \quad s.t. \bigwedge_{i \in [n]} \bar{p}_i E_i(x) \leq \bar{B}_i$$

or, equivalently $\max_{x \in X | g(x) \leq \vec{0}} f(x)$ *where:*

- $f : X \to \mathbb{R} \mid f(x) = \sum_{i \in [n]} \bar{p}_i E_i(x) - C_i(x)$.
- $g : X \to \mathbb{R}^n \mid g(x) = (\bar{p}_1 E_1(x) - \bar{B}_1, \ldots, \bar{p}_n E_n(x) - \bar{B}_n)$.

Definition 6 (FBMP: Fixed Budget, Maximum Price). *The client commits herself to pay a total amount $\bar{b}_i$ in case we deliver enough events to put her unitary price below $\bar{P}_i$. Moreover, we are not allowed to get a margin over costs above $\bar{M}_i$[4]. Formally:*

$$\max_{x \in X} \sum_{i \in [n]} \bar{b}_i - C_i(x) \quad s.t. \bigwedge_{i \in [n]} \frac{\bar{b}_i}{E_i(x)} \leq \bar{P}_i \ \wedge \ 1 \leq \frac{\bar{b}_i}{C_i(x)} \leq 1 + \bar{M}_i$$

or, equivalently $\max_{x \in X | g'(x) \leq \vec{0}} f'(x)$ *where:*

- $f' : X \to \mathbb{R} \mid f'(x) = \sum_{i \in [n]} \bar{b}_i - C_i(x)$.
- $g' : X \to \mathbb{R}^{3n} \mid$
  $g'(x) = (\bar{b}_1 - \bar{P}_1 E_1(x), C_1(x) - \bar{b}_1, \bar{b}_1 - C_1(x)(1 + \bar{M}_1), \ldots, \bar{b}_n - \bar{P}_n E_n(x), C_n(x) - \bar{b}_n, \bar{b}_n - C_n(x)(1 + \bar{M}_n))$

Notice that, in both cases, the objective is our profit. This means we can additively aggregate MBFP and FBMP problems in a way that makes economical sense, *i.e.* by using a monetary measure (profit) that works the same no matter the kind of contract. Hence, our entire strategy can be derived from a single company-wise optimization problem.

3 MBFP CONTINUOUS RELAXATION

Consider a simpler version of MBFP with just one campaign i and both bid b and ad a fixed in advance. A strategy can then be represented as a $\{0, 1\}^m$ vector indicating which auctions will be selected for bidding. Let's refer to $p_i w_k(b) e_k(i, a) = p w_k e_k = \omega_k$ as the *weight* of auction k and to $w_k(b)(p_i e_k(i, a) - c_k(b)) = w_k(pe_k - c_k) = v_k$ as the *value* of auction k. We can also think of the budget $\bar{B}_i = \bar{B} = \bar{W}$ as the *weight capacity* of the campaign. Thus we have reformulated a simplified version of MBFP as an instance of the famous 0-1 knapsack problem, which is a combinatorial optimization problem known to be NP-hard. The problem is hard because we are packing *indivisible* chunks of arbitrary weight and value. One might think of packing these chunks in decreasing "specific value" $\rho_k = v_k / \omega_k$ order but, in general, there will be a next chunk k^* to be packed that does not fit the knapsack anymore, leaving wasted space that could have been put to better use by some other permutation of chunks. But if chunks were divisible, chunk k^* could have been split in a way that the allotted portion exactly filled the knapsack. Indeed, this is Dantzig's solution for the continuous relaxation of the 0-1 knapsack problem (see [5] and [8]).

Still, knapsack is just a simplified instance of MBFP. In our way back to recover full MBFP we start by reintroducing an arbitrary number of campaigns. Value v_{ik}, weight ω_{ik} and specific value ρ_{ik} will again depend on both auction k and campaign i. Now, consider campaigns i and j and an amount of infinitely divisible auctions assigned to each one of them in a way that exactly exhausts budgets B_i and B_j and is consistent with our ρ-ranking when looking at each campaign on its own. Consider also an auction k that was totally or partially assigned to campaign i. Was this assignment the right one? If a fraction α of k with weight $\alpha\omega_{ik}$ is transferred to campaign j, campaign i will have to compensate its loss by buying the same weight at its "margin" $\rho_{ik_i^*}$; conversely, campaign j will have to forgo weight $\alpha\omega_{jk}$ at its own "margin" $\rho_{jk_j^*}$. All in all, the net value of the transfer will be $\alpha\omega_{ik}(\rho_{ik_i^*} - \rho_{ik}) + \alpha\omega_{jk}(\rho_{jk} - \rho_{jk_j^*})$. If this net value is above zero, then the transfer is profitable and the original assignment was suboptimal. The condition is readily reformulated as $\omega_{jk}(\rho_{jk} - \rho_{jk_j^*}) > \omega_{ik}(\rho_{ik} - \rho_{ik_i^*})$. The resulting rule is: pick the campaign i with highest positive $\omega_{ik}(\rho_{ik} - \rho_{ik_i^*})$, in case there is one; pick no campaign otherwise.

We are still supposing divisibility but, because each auction is so small and the market is composed of so many billions of auctions, this is a natural way of modelling the RTB market. We are also assuming that campaign i can always extend its "margin" of purchase by buying some fraction of an auction with specific value quite close to $\rho_{ik_i^*}$; again, given the sheer volume and variety of RTB markets, this seems a relatively mild assumption. Finally, throwing bids and ads again into the mix adds nothing conceptually new: we simply pick the bid and ad pair that maximizes ρ_{ik}. So, given a vector $\rho = (\rho_{1k_1^*}, \ldots, \rho_{nk_n^*})$ of "worst specific values" for each campaign, so to speak, we have designed a rule to optimally assign auctions to campaigns. Of course, the focus has now changed to finding the optimal ρ that exactly exhausts budgets, but this is easier than permuting billions of auctions. In practice, we adjust ρ a bit every day in the direction that most promotes the completion of budgets, as will be detailed in sections 7 and 8.

4 MBFP DUAL PROBLEM

Something along the lines presented in the previous section was our main bidding strategy during the last two years. But the increasing number of FBMP campaigns and the difficulty of reasoning about the FBMP problem in a similar fashion, plus the lack of theoretical convergence results, pressed us for a different foundation that supported both MBFP and FBMP and provided some basic theoretical warranties. Fortunately, the Lagragian relaxation framework satisfies those requirements. Interestingly enough, as we shall soon see, the algorithm derived above is recovered in a much more straightforward —though perhaps less intuitive— manner as the solution to a dual problem, thus reinforcing our confidence in the algorithm.

From definition 5 above it is clear that the Lagrangian of MBFP is:

$$\mathcal{L}(x, \theta^*) = f(x) + \langle \theta^*, g(x) \rangle = \sum_{i \in [n]} \bar{p}_i E_i(x) - C_i(x) + \theta_i^* (\bar{p}_i E_i(x) - \bar{B}_i)$$

Reminding the definition 4 of aggregate functions, we realize $\mathcal{L}(x, \theta^*)$ is a sum over a large number of terms, one term for each auction (plus some terms $\theta_i^* \bar{B}_i$ which don't depend on strategy x):

$$u_k(i, b, a) = w_k(b)(\bar{p}_i e_k(i, a) - c_k(b)) + \theta_i^* w_k(b) \bar{p}_i e_k(i, a)$$

This additive structure implies that the expected contribution of each auction to the aggregate may be computed disregarding other auctions. It should be clear that, if we wanted to maximize $\mathcal{L}$ over X for a given θ^*, we would better play by the following rule:

[4]This clause is necessary because an ensured income $\bar{b}_i$ would induce a profit maximizing agent to reduce costs until barely meeting constraint $\bar{P}_i$.

Definition 7 (MBFP Rule). *Assign auction k to campaign i^* with bid b^* and ad a^* such that $i^*, b^*, a^* = \arg\max_{i,b,a} u_k(i,b,a)$ if and only if $u_k(i^*, b^*, a^*) > 0$.*

In order to find i^*, b^*, a^* we can iterate over relevant (campaign, ad) pairs and compute the optimal bid for each one of them. In practice, we pre-rank these pairs using a cheap ranking metric, inject some noise into the ranking, and then perform the full optimization prescribed by rule 7 just for the first few ones. This is mainly because getting estimates conditional to both the full context of the auction and the history of the device is rather expensive. On the other hand, computing b^* given i and a is cheap, specially in the case of second-price auctions. This becomes evident after re-expressing u_k as $w_k(b)[(1+\theta_i^*)\bar{p}_i E_k(i,a) - C_k(b)]$. Here, we see that finding the optimal bid for a second-price auction is just a matter of taking $b^* = (1+\theta_i^*)\bar{p}_i E_k(i,a)$. For first price auctions the solution has to be numerically obtained except for the simplest statistical models of E_k and C_k. Still, it can be quickly computed in real-time for reasonably complex models. Note that, as a matter of fact, rule 7 codifies a simple algorithm to compute the MBFP dual function q, based on which we can define a dual problem:

Definition 8 (MBFP Dual Function). *Given θ^*, the MBFP dual function $q(\theta^*)$ maximizes the Lagrangian $\mathcal{L}(x,\theta^*)$ over the set of strategies X, i.e. $q(\theta^*) = \sup_{x\in X} \mathcal{L}(x,\theta^*)$.*

Definition 9 (MBFP Dual Problem). *Given the MBFP dual function $q(\theta^*)$, the MBFP dual problem consists in minimizing it over the non-positive orthant, i.e. $\inf_{\theta^*\leq\vec{0}} q(\theta^*)$.*

For now, it suffices to state that by minimizing q over the non-positive orthant we get an approximate solution to MBFP. How good is this approximation and how to effectively carry out that minimization task are questions that will be answered from section 6 onwards. Nevertheless, before concluding this section, we want to show that rule 7 is essentially the "distance from current ρ_i to worst ρ_i" rule that was previously derived by using a —somewhat hand-waved— continuity/divisibility argument:

$$\begin{aligned} u_k(i,b,a) &= w_k(b)(\bar{p}_i e_k(i,a) - c_k(b)) + \theta_i^* w_k(b)\bar{p}_i e_k(i,a) \\ &= w_k(b)\bar{p}_i e_k(i,a)\left(\frac{w_k(b)(\bar{p}_i e_k(i,a) - c_k(b))}{w_k(b)\bar{p}_i e_k(i,a)} + \theta_i^*\right) \\ &= \omega_{ik}(\rho_{ik} + \theta_i^*) \underset{\theta_i^* := -\rho_{ik_i^*}}{=} \omega_{ik}(\rho_{ik} - \rho_{ik_i^*}) \end{aligned}$$

5 FBMP DUAL PROBLEM

Computing the FBMP dual function follows the same pattern as computing the MBFP dual function. As before, we start by writing down the Lagrangian:

$$\begin{aligned} \mathcal{L}'(x,\theta^*) = f'(x) + \langle\theta^*, g'(x)\rangle = \\ \sum_{i\in[n]} \bar{b}_i - C_i(x) + \theta_{1i}^*(\bar{b}_i - \bar{P}_i E_i(x)) + \\ \theta_{2i}^*(C_i(x) - \bar{b}_i) + \theta_{3i}^*(\bar{b}_i - (1+\bar{M}_i)C_i(x)) \end{aligned}$$

Again, because of the additive structure of $\mathcal{L}'$, the expected contribution of auction k to the aggregate when assigned to campaign i with bid b and ad a can be expressed without regarding any other auction:

$$\begin{aligned} u_k'(i,b,a) = &- w_k(b)c_k(b) - \theta_{1i}^*\bar{P}_i w_k(b)e_k(i,a) \\ &+ \theta_{2i}^* w_k(b)c_k(b) - \theta_{3i}^*(1+\bar{M}_i)w_k(b)c_k(b) \end{aligned}$$

By the same token as before, we derive a bidding rule based on this definition of u_k' for FBMP:

Definition 10 (FBMP Rule). *Assign auction k to campaign i^* with bid b^* and ad a^* such that $i^*, b^*, a^* = \arg\max_{i,b,a} u_k'(i,b,a)$ if and only if $u_k'(i^*, b^*, a^*) > 0$.*

Regrouping u_k' we arrive at:

$$u_k'(i,b,r) = w_k(b)[-\theta_{1i}^*\bar{P}_i e_k(i,a) - c_k(b)(1-\theta_{2i}^* + \theta_{3i}^*(1+\bar{M}_i))]$$

Given θ_{2i}^*, θ_{3i}^* and $\bar{M}_i$ such that $1-\theta_{2i}^* + \theta_{3i}^*(1+\bar{M}_i) > 0$ the maximizers of u_k' will be the same than the maximizers of: [5]

$$w_k(b)\left(\frac{-\theta_{1i}^*\bar{P}_i e_k(i,a)}{1-\theta_{2i}^* + \theta_{3i}^*(1+\bar{M}_i)} - c_k(b)\right)$$

Therefore, the optimal amount b^* to bid for a campaign i and an ad a in a second-price auction is easily computed as:

$$b^* = \frac{-\theta_{1i}^*\bar{P}_i g_k(i,a)}{1-\theta_{2i}^* + \theta_{3i}^*(1+\bar{M}_i)}$$

Again, for first-price auctions a moderate increase in computational complexity could be expected, but there is no additional conceptual complexity.

Finally, we give analogous definitions for the FBMP dual function and the FBMP dual problem:

Definition 11 (FBMP Dual Function). *Given θ^*, the FBMP dual function $q'(\theta^*)$ maximizes the Lagrangian $\mathcal{L}'(x,\theta^*)$ over the set of strategies X, i.e. $q'(\theta^*) = \sup_{x\in X} \mathcal{L}'(x,\theta^*)$.*

Definition 12 (FBMP Dual Problem). *Given the FBMP dual function $q'(\theta^*)$, the FBMP dual problem consists in minimizing it over the non-positive orthant, i.e. $\inf_{\theta^*\leq\vec{0}} q'(\theta^*)$.*

As before, the idea is that by minimizing q' over the non-positive orthant we get an approximate solution to FBMP, as we shall see in the following section.

6 CONVEX RELAXATION

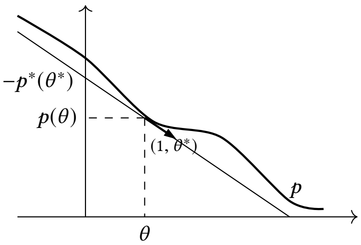
Now consider the more general problem:

$$p(\theta) = \min_{x|g(x)\leq\theta} -f(x) = -\max_{x|g(x)\leq\theta} f(x)$$

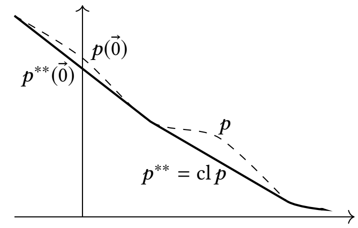
where, for technical reasons and since it is customary, we are re-hashing a maximization problem as a minimization one. Here f and g are arbitrary vector functions. By varying θ one can tighten or relax the constraints. It is evident that p is non-increasing on any component θ_i of θ, since by increasing any θ_i we are enlarging the feasible set of the problem. We are mainly interested in $\theta = \vec{0}$ —or, more precisely, in those x that solve $p(\vec{0})$— but, instead of directly solving $p(\vec{0})$, we are going to work with an approximation $p^{**}(\vec{0})$, where p^{**} is a "convexified" version of p (hence the name "convex relaxation"). In our way to get p^{**} we first take the Fenchel conjugate of p:

$$\begin{aligned} p^*(\theta^*) &= \sup_\theta\{\langle\theta^*,\theta\rangle - p(\theta)\} = \sup_\theta\{\langle\theta^*,\theta\rangle - \min_{x|g(x)\leq\theta} -f(x)\} \\ &= \sup_\theta \max_{x|g(x)\leq\theta}\{\langle\theta^*,\theta\rangle + f(x)\} = \sup_{x,\delta\geq\vec{0}}\{\langle\theta^*, g(x)+\delta\rangle + f(x)\} \\ &= \sup_x\{f(x) + \langle\theta^*, g(x)\rangle\} + \sup_{\delta\geq\vec{0}}\langle\theta^*,\delta\rangle \underset{\theta^*\leq\vec{0}}{=} q(\theta^*) + 0 = q(\theta^*) \end{aligned}$$

[5] In general, we expect this condition to hold for a solution, since otherwise u_k would favor strategies that ignore costs or, worse, that prefer higher costs to lower ones.



(1.a) $-p^*(\theta^*)$ is the intercept of a supporting hyperplane of epi p.



(1.b) p^{**} results from the convex closure of epi p.

As we can see, q is the dual function of problem $\max_{x\,|\,g(x)\leq\theta} f(x)$ as was earlier defined. Some thinking will reveal that $-p^*(\theta^*) = -q(\theta^*)$ gives the intercept of the supporting hyperplane of the epigraph of p with slope θ^*, as shown in figure 1.a. Thus p^* can be seen as encoding an alternative representation of p by mapping slopes to intercepts. Notice also that, since p is non-increasing on every θ_i, we are only interested in slopes θ^* in the non-positive orthant. The encoding will be loseless only when p is convex; in that case, recovering p from p^* is just a matter of conjugating it again, *i.e.* $p = p^{**}$. More generally, p^{**} will be a function $\mathrm{cl}\, p$ related to p by taking the convex closure (or hull) of the epigraph of p, as illustrated in figure 1.b.

Based on well known properties of the Fenchel conjugate p^* we can easily infer analogous properties of q that will be used in the next section: [6]

Proposition 1. *q is convex.*

Proof. $q = p^*$ and p^* is always convex (even if p is not). □

Proposition 2. *$g(x^*) \in \partial q(\theta^*)$, i.e. $g(x^*)$ is a subgradient of q at point θ^*, where x^* is a solution for $q(\theta^*)$.*

Proof. We use Fenchel-Young inequality, which states that $p(\theta) + p^*(\theta^*) \geq \langle\theta^*, \theta\rangle$. We also use that, when Fenchel-Young result holds with equality, it implies $\theta \in \partial p^*(\theta^*)$ (as well as $\theta^* \in \partial p(\theta)$, which we don't need here). Take $\theta = g(x^*)$, then:

(1) Because of Fenchel-Young inequality:
$p^*(\theta^*) \geq -p(\theta) + \langle\theta^*, g(x^*)\rangle$
(2) And because of the definition of x^* and p:
$p^*(\theta^*) = q(\theta^*) = f(x^*) + \langle\theta^*, g(x^*)\rangle \leq \max_{x\,|\,g(x)\leq g(x^*)} \{f(x)\} + \langle\theta^*, g(x^*)\rangle = -p(\theta) + \langle\theta^*, g(x^*)\rangle$

But 1 and 2 imply that $p^*(\theta^*) + p(\theta) = \langle\theta^*, g(x^*)\rangle$. Therefore $\theta = g(x^*) \in \partial p^*(\theta^*)$. □

Finally, let's conjugate p^* again in order to (maybe approximately) recover our initial p:

$$
\begin{aligned}
p^{**}(\theta) &= \sup_{\theta^*\leq\vec{0}} \{\langle\theta^*, \theta\rangle - p^*(\theta^*)\} \\
&= \sup_{\theta^*\leq\vec{0}} \{\langle\theta^*, \theta\rangle - \sup_x\{f(x) + \langle\theta^*, g(x)\rangle\}\} \\
&= -\inf_{\theta^*\leq\vec{0}} \sup_x f(x)\langle\theta^*, g(x) - \theta\rangle \underset{\theta=\vec{0}}{=} -\inf_{\theta^*\leq\vec{0}} q(\theta^*)
\end{aligned}
$$

What we have here is that, while $-p(\vec{0})$ is the maximum we were initially looking for, *i.e.* $\max_{x\,|\,g(x)\leq\vec{0}} f(x)$, $-p^{**}(\vec{0})$ is instead the result of our dual rehashing of the problem, *i.e.* $\inf_{\theta^*\leq\vec{0}} q(\theta^*)$. So by studying the relationship between $-p$ and $-p^{**}$ we can learn about the relationship between primal and dual solutions. As before, we rely on well known facts about Fenchel conjugacy in what follows.

Proposition 3. *If $-p$ is concave, then both primal and dual optima for $\max_{x\,|\,g(x)\leq\vec{0}} f(x)$ are the same: $-p(\vec{0})$. If not, the dual optimum is instead $-(\mathrm{cl}\, p)(\vec{0}) \geq -p(\vec{0})$, where $\mathrm{cl}\, p$ is the function with epigraph equal to the convex closure of the epigraph of p (equivalently, $-(\mathrm{cl}\, p)$ is the function with hypograph equal to the convex closure of the hypograph of $-p$). The difference between both optima is called the duality gap.*

Proof. We use the aforementioned fact that $p^{**} = \mathrm{cl}\, p$. This implies:

(1) If p is convex, then $p^{**} = p$, since $\mathrm{cl}\, p = p$.
(2) $p^{**}(\theta) \leq p(\theta)$ for all θ, since $\mathrm{cl}\, p \leq p$.

(see figure 1.b for an illustration of these properties). Finally, substitute concave for convex and $\geq$ for $\leq$, given that we are interested in $-p^{**}(\vec{0})$ and $-p(\vec{0})$. □

By solving the dual it is as if we were maximizing a concavified version of our original problem. More precisely, of a function $-p(\theta) = \max_{x\,|\,g(x)\leq\theta} f(x)$ that relates a perturbation θ of the constraints to a solution of the perturbed problem. Hence, if $-p$ is concave enough it is expected that $-p^{**}$ will be close to it. Notice that every constraint in MBFP and FBMP imposes limits on the sum of billions of small auctions showing a great degree of substitutability between them. Therefore, if we progressively relaxed one of these constraints by changing component θ_i, the optimizer would pick opportunities with high "specific value" for constraint i first, thus yielding decreasing marginal returns as the constraint is relaxed. So, as a function of θ_i, we expect $-p$ to be approximately concave. Analogously, we could consider relaxations along any line $\theta + tv$ for some slope v to the same effect. In consequence, we expect $-p$ to be approximately concave.

7 SUBGRADIENT DESCENT FOR THE DUAL

Recall that, according to propositions 1 and 2 above, the dual function q is convex and there is a cheap method for computing a subgradient of it at point θ^*: simply evaluate the constraint function g at a maximizer x^* such that $\mathcal{L}(x^*, \theta^*) = \sup_{x\in\mathcal{X}} \mathcal{L}(x, \theta^*)$. Recall also that, given θ^*, we compute this optimal strategy x^* by picking $\arg\max_{i,b,a} u_k(i, b, a)$ for each auction k. Thus, we could iteratively descend along a sequence of multipliers $\theta_1^*(x_0^*), \theta_2^*(x_1^*), \ldots, \theta_t^*(x_{t-1}^*), \ldots$ in an "outer loop", so to speak, while implementing a sequence of optimal strategies $x_1^*(\theta_1^*), x_2^*(\theta_2^*), \ldots, x_t^*(\theta_t^*), \ldots$ in the "inner loop", this until convergence, which —as we shall soon see— is guaranteed under rather mild assumptions. The assertion

[6] If you want to dig deeper into this subject, we strongly recommend [9].

"we could iteratively descend" demands some qualification: a subgradient, like a gradient, gives an affine global underestimator of its function but, unlike a (non-zero) gradient, its negative does not necessarily point towards a descent direction, as you can see in figure 2.a. Nevertheless, a negative subgradient still is a descent direction in a different sense. Consider points θ^* and $\underaccent{\bar}{\theta}^*$ such that $q(\underaccent{\bar}{\theta}^*) < q(\theta^*)$; that is, q descends from θ^* to $\underaccent{\bar}{\theta}^*$. Every time there is such a $\underaccent{\bar}{\theta}^*$, it is true that $-g$ is a descent direction for $||\theta^* - \underaccent{\bar}{\theta}^*||_2$ if $g \in \partial q(\theta^*)$. In other words, we are not necessarily moving down along q, but at least we are getting closer to a point where q is lower if both *(i)* there is such a point and *(ii)* we move slowly enough, *i.e.* $||\theta^* - \alpha g - \underaccent{\bar}{\theta}^*||_2 < ||\theta^* - \underaccent{\bar}{\theta}^*||_2$ for some small learning rate α. Intuitively, since the subgradient gives an affine global underestimator, the corresponding hyperplane passes under q at $\underaccent{\bar}{\theta}^*$, thus it has a negative slope in the trajectory from θ^* to $\underaccent{\bar}{\theta}^*$, hence the vector $\underaccent{\bar}{\theta}^* - \theta^*$ and the subgradient form an angle with negative cosine or, in other words, $\underaccent{\bar}{\theta}^* - \theta^*$ and the negative subgradient form an angle with positive cosine, therefore the negative subgradient direction is somewhat aligned with $\underaccent{\bar}{\theta}^* - \theta^*$ and, in this direction, we will be moving closer (in L^2-norm) to $\underaccent{\bar}{\theta}^*$ if the step is appropriately chosen in relation to the angle (see figure 2.b). The next proposition states and proves this fact; it also provides a condition on α that ensures descent in the sense described above.

Proposition 4. *Take $\delta = q(\theta^*) - q(\underaccent{\bar}{\theta}^*)$ and assume G upper bounds in L^2-norm the set of all subgradients of q. Then if $\delta > 0$, $g \in \partial q(\theta^*)$ and $0 < \alpha < \frac{2\delta}{G^2}$, it is the case that $||\theta^* - \alpha g - \underaccent{\bar}{\theta}^*||_2 < ||\theta^* - \underaccent{\bar}{\theta}^*||_2$. Moreover, $\theta^* - \alpha g$ is at least $2\alpha\delta - \alpha^2 G^2$ closer to $\underaccent{\bar}{\theta}^*$ than θ^*.*

Proof. We first apply the law of cosines to get:

$$||\theta^* - \alpha g - \underaccent{\bar}{\theta}^*||_2^2 = ||\theta^* - \underaccent{\bar}{\theta}^*||_2^2 - 2\alpha\langle g, \theta^* - \underaccent{\bar}{\theta}^*\rangle + \alpha^2||g||_2^2$$

Then we note that, since $q(\theta^*) + \langle g, \_ - \theta^*\rangle$ is a global underestimator of q, it must be that $-\langle g, \theta^* - \underaccent{\bar}{\theta}^*\rangle \le -(q(\theta^*) - q(\underaccent{\bar}{\theta}^*)) = -\delta$. Hence, the expression above is not greater than $||\theta^* - \underaccent{\bar}{\theta}^*||_2^2 - 2\alpha\delta + \alpha^2||g||_2^2$. Therefore:

$$||\theta^* - \underaccent{\bar}{\theta}^*||_2^2 - ||\theta^* - \alpha g - \underaccent{\bar}{\theta}^*||_2^2 \ge 2\alpha\delta - \alpha^2||g||_2^2 \ge 2\alpha\delta - \alpha^2 G^2$$

For $\alpha > 0$, this expression is positive if and only if $2\delta - \alpha G^2 > 0$, thus we get the condition $0 < \alpha < \frac{2\delta}{G^2}$. □

Proposition 4 means that when we are far enough from the minimum value, which is to say that δ is large, we will be moving towards the minimizer on each step if neither α nor the subgradients found along the way are too large. In this case, each step leaves us $2\alpha\delta - \alpha^2 G^2$ closer to the minimizer; notice here that a small $\alpha \ll 1$ more strongly penalizes G^2 than 2δ. It is not difficult to infer from this simple result that a small and constant α will leave us pretty close to the minimizer, moving faster towards it the furthest the current value is from the minimum value. This is an encouraging conclusion, but more useful convergence results can be read from the next, slightly more complex, proposition.

Proposition 5. *Call θ_t^* the θ^* reached at step t. Besides, call $\underaccent{\bar}{\theta}_t^*$ the θ^* with lower $q(\theta^*)$ found up to step t; that is, the best θ^* found until t. Finally, take $\underaccent{\bar}{\theta}^*$ to be a global minimizer. Then, following a learning rate schedule $\alpha_1, \alpha_2, \ldots, \alpha_T$, we have:*

$$q(\underaccent{\bar}{\theta}_T^*) - q(\underaccent{\bar}{\theta}^*) \le \frac{||\theta_1^* - \underaccent{\bar}{\theta}^*||_2^2 + G^2\sum_{t=1}^T \alpha_t^2}{2\sum_{t=1}^T \alpha_t}$$

Proof. As before, we start by applying the law of cosines and the global underestimator property of the subgradient, to get:

$$\begin{aligned}||\theta_{t+1}^* - \underaccent{\bar}{\theta}^*||_2^2 &= ||\theta_t^* - \alpha_t g_t - \underaccent{\bar}{\theta}^*||_2^2 \\ &= ||\theta_t^* - \underaccent{\bar}{\theta}^*||_2^2 - 2\alpha_t\langle g_t, \theta_t^* - \underaccent{\bar}{\theta}^*\rangle + \alpha_t^2||g_t||_2^2 \\ &\le ||\theta_t^* - \underaccent{\bar}{\theta}^*||_2^2 - 2\alpha_t(q(\theta_t^*) - q(\underaccent{\bar}{\theta}^*)) + \alpha_t^2 G^2\end{aligned}$$

Then, by applying the same procedure recursively:

$$||\theta_T^* - \underaccent{\bar}{\theta}^*||_2^2 \le ||\theta_1^* - \underaccent{\bar}{\theta}^*||_2^2 - 2\sum_{t=1}^T \alpha_t(q(\theta_t^*) - q(\underaccent{\bar}{\theta}^*)) + G^2\sum_{t=1}^T \alpha_t^2$$

And since the left hand side of the inequality is non-negative:

$$\begin{aligned}2(q(\underaccent{\bar}{\theta}_T^*) - q(\underaccent{\bar}{\theta}^*))\sum_{t=1}^T \alpha_t &\le 2\sum_{t=1}^T \alpha_t(q(\theta_t^*) - q(\underaccent{\bar}{\theta}^*)) \\ &\le ||\theta_1^* - \underaccent{\bar}{\theta}^*||_2^2 + G^2\sum_{t=1}^T \alpha_t^2\end{aligned}$$

Hence, we conclude:

$$q(\underaccent{\bar}{\theta}_T^*) - q(\underaccent{\bar}{\theta}^*) \le \frac{||\theta_1^* - \underaccent{\bar}{\theta}^*||_2^2 + G^2\sum_{t=1}^T \alpha_t^2}{2\sum_{t=1}^T \alpha_t}$$

□

From proposition 5 we can easily deduce convergence results for different schedules $\alpha_1, \alpha_2, \ldots, \alpha_T$[7]. Here, we are mostly interested in the simple case of a constant learning rate schedule, for which it is easy to give an upper bound to the number of steps required in order to be near the optimal:

Proposition 6. *With constant rate schedule $\alpha_t = \alpha$, A converges sublinearly and logarithmically within $\varepsilon + \alpha G^2/2$ of the optimum in no more than $\lceil ||\theta_1^* - \underaccent{\bar}{\theta}^*||_2^2/2\alpha\varepsilon \rceil$ steps.*

Proof. Using proposition 5 we can show that:

$$\begin{aligned}q(\underaccent{\bar}{\theta}_T^*) - q(\underaccent{\bar}{\theta}^*) &\le \frac{||\theta_1^* - \underaccent{\bar}{\theta}^*||_2^2 + G^2\sum_{t=1}^T \alpha_t^2}{2\sum_{t=1}^T \alpha_t} \\ &= \frac{||\theta_1^* - \underaccent{\bar}{\theta}^*||_2^2 + G^2 T\alpha^2}{2T\alpha} = \frac{||\theta_1^* - \underaccent{\bar}{\theta}^*||_2^2}{2T\alpha} + \frac{\alpha G^2}{2}\end{aligned}$$

It is clear from this expression that A converges within $\alpha G^2/2$ of the optimum sublinearly (in particular, logarithmically). To be no more than $\varepsilon + \alpha G^2/2$ far from the optimum we need the first term not to be larger than ε:

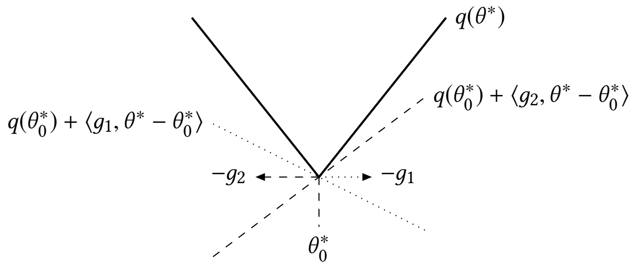
$$\frac{||\theta_1^* - \underaccent{\bar}{\theta}^*||_2^2}{2T\alpha} \le \varepsilon \Leftrightarrow T \ge \frac{||\theta_1^* - \underaccent{\bar}{\theta}^*||_2^2}{2\alpha\varepsilon}$$

□

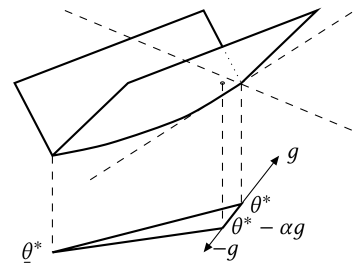
8 SUBGRADIENT DESCENT WITH NOISE

As promised, we present a descent schema adapted to real-life noisy environments. Recall that, at the moment of bidding for auction k, we only have estimates of conditional means w_k, c_k, e_k and not yet the realizations of the respective random variables. Now consider $\mathbf{G}(x^*)$, in which we replace w_k, c_k, e_k in $g(x^*)$ for their random counterparts $\mathbf{W}_k, \mathbf{C}_k, \mathbf{E}_k$; analogously, consider $\tilde{g}(x^*)$, in which we replace the random variables for their realizations $\tilde{w}_k, \tilde{c}_k, \tilde{e}_k$. Unfortunately, at the moment of descent our bidder only reports an aggregate $\tilde{g}(x^*)$ for the last period, which is not the $g(x^*)$ we want. Yet, fortunately, if we assume that $g(x^*)$ is the mean of $\mathbf{G}(x^*)$,

[7] See [2] and [1, §8.2] for further details.



(2.a) Negative subgradients pointing towards ascent directions.



(2.b) $||\underline{\theta}^* - \theta^*||_2$ decreases along negative subgradient direction.

i.e. that $\tilde{g}(x^*)$ is an unbiased estimate of $g(x^*)$, then the results of propositions 5 and 6 are preserved for $E[q(\underline{\theta}_T^*)]$ (instead of $q(\underline{\theta}_T^*)$), as detailed in proposition 7 below. And the assumption seems tenable because, according to definition 4, $g(x^*)$ has a linear structure in problems MBFP and FBMP: it consists of a linear combination of a large number of mean values[8] $w_k c_k$ and $w_k e_k$, so the expectation operator passes through $\mathbf{G}(x^*)$ down to $\mathbf{W}_k\mathbf{C}_k$ and $\mathbf{W}_k\mathbf{E}_k$.

Proposition 7. *Define everything as in proposition 5. Assume that, at each step t, we have access to an unbiased estimate $\tilde{g}_t \leq \tilde{G}$ of g_t that we use in place of g_t. Then, following a learning rate schedule $\alpha_1, \alpha_2, \ldots, \alpha_T$, we have:*

$$E[q(\underline{\theta}_T^*)] - q(\underline{\theta}^*) \leq \frac{||\theta_1^* - \underline{\theta}^*||_2^2 + \tilde{G}^2 \sum_{t=1}^T \alpha_t^2}{2\sum_{t=1}^T \alpha_t}$$

Proof. The proof proceeds as in 5 but, additionally, uses linearity of $E[\cdot]$ and Jensen's inequality. In particular, Jensen's inequality is used to show that, because of concavity of the minimum:

$$\left(E[q(\underline{\theta}_T^*)] - q(\underline{\theta}^*)\right)\sum_{t=1}^T \alpha_t = \left(E[\min_{t=1}^T q(\theta_t^*)] - q(\underline{\theta}^*)\right)\sum_{t=1}^T \alpha_t$$

$$\leq \sum_{t=1}^T \alpha_t \min_{t=1}^T E[q(\theta_t^*) - q(\underline{\theta}^*)] \leq \sum_{t=1}^T \alpha_t E[q(\theta_t^*) - q(\underline{\theta}^*)]$$

For further details, we refer to [3]. □

Nevertheless, to be able to extrapolate gradient estimates from current to future periods, we still need some kind of stationarity/ergodicity assumption. In practice, we simply consider the realization $\tilde{g}(x^*)$ corresponding to the last day as a good estimate for the following days. Of course, it is true that there are weekly seasonal effects not accounted by this method, but most heterogeneity occurs within a day and not between days. Moreover, by keeping our learning rate above some threshold we ensure the optimizer remains adaptive to longer seasonal cycles and reactive to structural breaks that have to be expected in ever changing RTB markets[9]. Therefore, a reasonable algorithm might consist in computing x^* each day —given the θ^* for that day— by picking $\arg\max_{i,b,a} u_k(i,b,a)$ for every auction k seen and then, at the end of the day, recomputing θ^* by a descent step based on the observed constraints $\tilde{g}(x^*)$ and a not too large constant learning rate α.

[8]Remember that c_k and e_k are conditional to the fact that we won auction k.

[9]On the downside, this prevents exact convergence to the optimum.

9 CONCLUSION AND FUTURE WORK

We think that we have achieved the goal set for this work, *i.e.* to put forward a strong theoretical framework that relies on very few assumptions for the problem of optimal bidding for two —stylized but realistic— kinds of contract. We also believe to have advanced cogent arguments for the assumptions made. Moreover, for one of the contract types we formally linked the proposed solution to a successfully working (for more than two years now) solution.

Although it was suggested to us as a natural extension of what we were already doing in practice for MBFP, we later realized that the convex/Lagrangian relaxation approach is not totally novel in RTB (see [6]). We feel that our modest contributions here are: the extension to different contractual constraints, the identification and analysis of the assumptions behind this extension (hopefully hinting to further extensions), the interesting link with the continuous case, the generalization to first and second-price auctions and the support for a realistic, noisy environment.

Still, we are fully aware that the main weakness of our proposal is its lack of empirical support. We are currently at an advanced stage in the implementation and A/B testing of Gloval, an optimizer module for our bidder along the lines proposed here. Over the next months, we plan to publish a follow-up paper reporting bounds for the duality gap, empirically calibrated values for the learning rate and other hyper-parameters, and overall economical performance of the algorithm.

REFERENCES

[1] Dimitri Bertsekas. 2003. *Convex Analysis and Optimization.* Athena Scientific.
[2] Stephen Boyd. 2003. Subgradient Methods. (2003), 3–4.
[3] Stephen Boyd. 2019. Stochastic Subgradient Method (Lecture Notes). (2019).
[4] Stephen Boyd and Lieven Vandenberghe. 2009. *Convex Optimization.* Cambridge University Press.
[5] George B. Dantzig. 1957. Discrete-Variable Extremum Problems. *Operations Research 5 (2)* (1957), 266–288.
[6] Paul Grigas, Alfonso Lobos, Zheng Wen, and Kuang-chih Lee. 2017. Profit Maximization for Online Advertising Demand-Side Platforms. In *Proceedings of the ADKDD'17.* ACM, 11.
[7] David Luenberger. 1969. *Optimization by Vector Space Methods.* John Wiley & Sons, Inc.
[8] Silvano Martello and Paolo Toth. 1990. *Knapsack Problems - Algorithms and Computer Implementations.* John Wiley & Sons, Inc.
[9] Tyrrell Rockafellar. 1972. *Convex Analysis.* Princeton University Press.