Inverse Optimization through Online Learning

Friedrich-Alexander Universität Erlangen-Nürnberg
Georgia Tech

Andreas Bärmann, Alexander Martin, Sebastian Pokutta, Oskar Schneider

Learning Objective Functions

We demonstrate how to learn the (linear) objective function of a decision maker from observing the problem input data and her decisions over multiple rounds. In each round $t = 1, \dots, T$, the decision maker solves the optimization problem

$$\max\{c_{\text{true}}^T x \mid x \in X(p_t)\},$$

where the objective function $c_{\text{true}} \in \mathbb{R}^n$ is only known to her, while the input data $p_t \in \mathbb{R}^k$ in that round and the feasible set $X(p_t)$ they induce are observable. Our aim is to learn enough about the unknown objective function c_{true} over time to be able to find solutions to the above optimization problem that are essentially as good as those of the decision maker. To this end, we present a specialized learning algorithm and prove strong convergence bounds on solution quality. We show their effectiveness in two computational examples.

Algorithmic Approach

Assuming w.l.o.g. that the true objective function c_{true} lies within the unit simplex $\Delta_n \subset \mathbb{R}^n$, we can use online learning to find a strategy $(c_1, c_2, \dots, c_T)$ of objective functions in each round, whose error with respect to the optimal solution vanishes on average. This particular setup allows us to use the *multiplicative weights update (MWU)* method [1], more general setups can be treated via *online gradient descent* [2]. All that is required is an optimization oracle that is able to solve the T instances of the maximization subproblem

$$\max\{c^T x \mid x \in X(p_t)\} \qquad (1)$$

for any guess of linear objective function c, producing an optimal solution $\bar{x}_t$. Our algorithm then proceeds as follows:

input observations (p_t, x_t) for $t = 1, \dots, T$
output sequence of objectives $c_1, c_2, \dots, c_T$
$\eta \leftarrow \sqrt{\frac{\ln n}{T}}$ {set learning rate}
$w_1 \leftarrow \mathbb{1}^n$ {initialize weights}
for $t = 1, \dots, T$ **do**
  $c_t \leftarrow \frac{w_t}{\|w_t\|_1}$ {normalize weights}
  $\bar{x}_t \leftarrow \operatorname{argmax}\{c_t^T x \mid X(p_t)\}$ {solve Subproblem (1)}
  if $\bar{x}_t = x_t$ **then**
    $y_t \leftarrow 0$
  else
    $y_t \leftarrow \frac{\bar{x}_t - x_t}{\|\bar{x}_t - x_t\|_\infty}$
  end if
  $w_{t+1}(i) \leftarrow w_t(i)(1 - \eta y_t(i))$ {update weights}
end for
return $(c_1, c_2, \dots, c_T)$.

Convergence Bounds

The algorithm stated on the left produces a sequence of solutions $(x_t)_t$ which are practically as good as those of the decision maker on average:

Theorem 1. *Let $K \geq 0$ with $\max_{x_1, x_2 \in X(p_t)} \|x_1 - x_2\|_\infty \leq K$ for all $t = 1, \dots, T$. Then we have*

$$0 \leq \frac{1}{T} \sum_{t=1}^T (c_t - c_{true})^T (\bar{x}_t - x_t) \leq 2K \sqrt{\frac{\ln n}{T}},$$

which in particular implies:

1. $0 \leq \frac{1}{T} \sum_{t=1}^T c_t^T (\bar{x}_t - x_t) \leq 2K \sqrt{\frac{\ln n}{T}}$,
2. $0 \leq \frac{1}{T} \sum_{t=1}^T c_{true}^T (x_t - \bar{x}_t) \leq 2K \sqrt{\frac{\ln n}{T}}$.

This allows us to conclude that high deviations in solution quality between our solutions and those of the expert tend to 0 over time:

Corollary 2. *Under the prerequisites of Theorem 1, we have that for any $\epsilon > 0$ the fraction of rounds t with $c_{true}^T (x_t - \bar{x}_t) \geq 2K \sqrt{\frac{\ln n}{T}} + \epsilon$ is at most*

$$1 - \frac{\epsilon}{2K\sqrt{\frac{\ln n}{T}} + \epsilon}.$$

In the case of unique optima with a common minimum margin of optimality over all rounds, we can even show more:

Corollary 3. *Under the prerequisites of Theorem 1 and if $c_{true}^T (x_t - \bar{x}_t) \geq \Delta$ whenever $x_t \neq \bar{x}_t$ for some fixed $\Delta > 0$, we have*

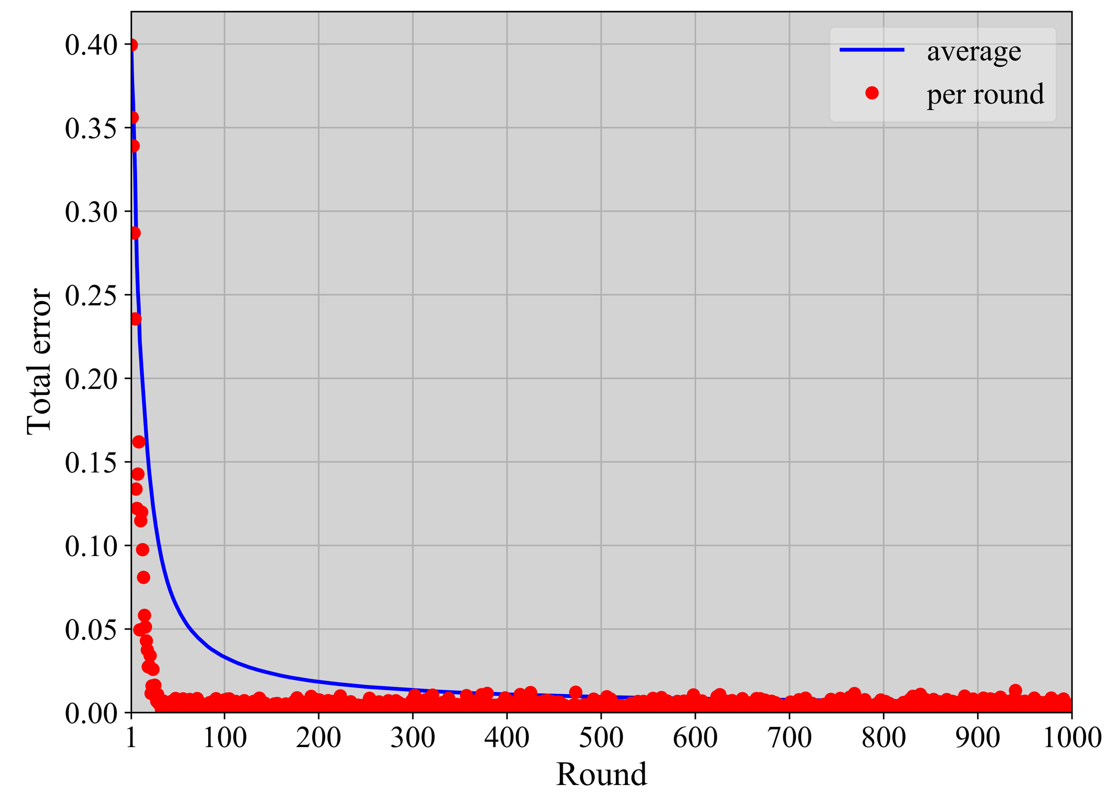
$$|\{t = 1, \dots, T \mid \bar{x}_t \neq x_t\}| \leq 2K \sqrt{\frac{T \ln n}{\Delta}}.$$

Application 1: Customer Preferences

- We consider a market, where customers can buy different goods G. Prices p_{tg} and the budget p_{t0} of the customer, which we can observe, vary over days $t = 1, \dots, T$. The preference u_g for each good $g \in G$ is unknown.
- To maximize his overall utility on each day t, the customer solves the problem

$$\begin{aligned} \max \quad & \sum_{g \in G} u_g x_g \\ \text{s.t.} \quad & \sum_{g \in G} p_{tg} x_g \leq p_{t0} \\ & x \in \{0,1\}^{|G|}, \end{aligned}$$

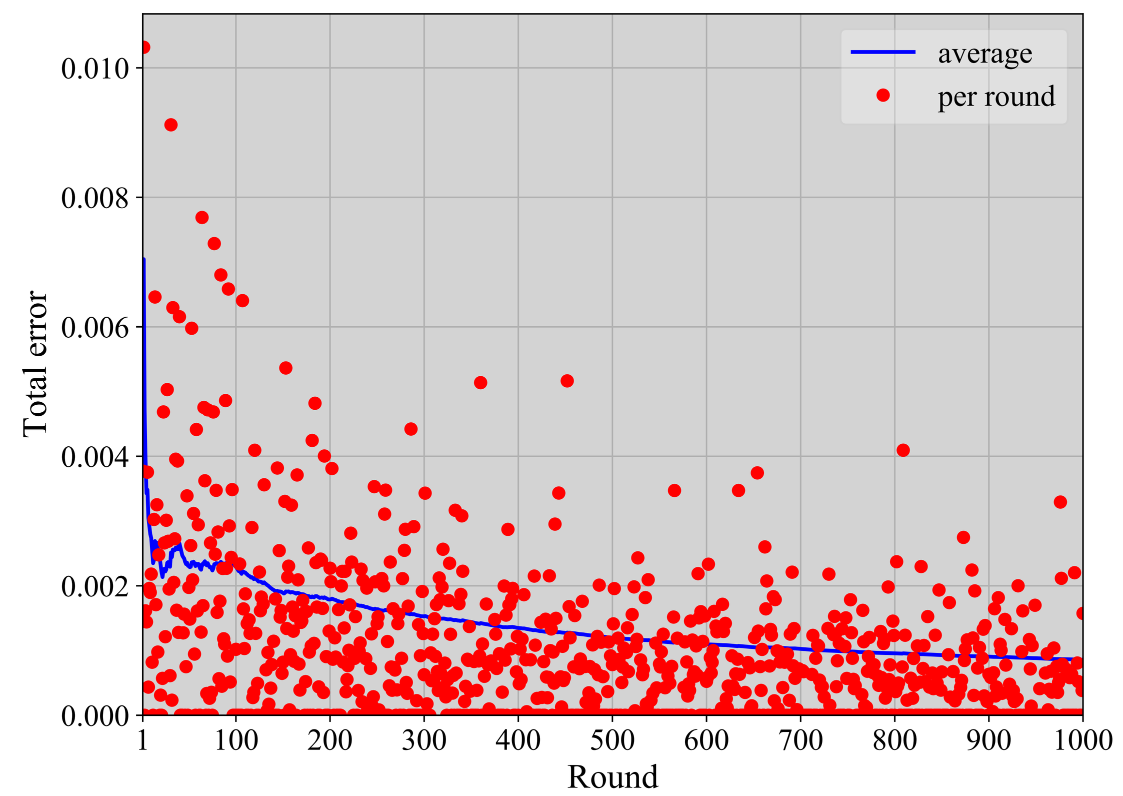
  whose optimal solution we can observe.
- We generated random instances with $T = 1000$ observations and $n = 100$ goods. The preferences and the prices are drawn such that all goods have similar benefit-cost ratios, making the problem hard to solve.



The figure shows the error $\frac{1}{T} \sum_{t=1}^T (c_t - c_{\text{true}})^T (\bar{x}_t - x_t)$ over T on the x-axis in blue. In red we plot the cost $(c_t - c_{\text{true}})^T (\bar{x}_t - x_T)$ of round t. As we can see, after few iterations most solutions reside on the x-axis and only few deviate beyond the average.

Application 2: Travel Times

- In this example, we consider drivers in a road network $G = (V, E)$ who want to minimize their travel times, while not driving overly large detours. For the network arcs $e \in E$, we assume the travel times c_e to be unknown, while their length a_e is known. We observe the path taken by each driver over time $t = 1, \dots, T$.
- Each driver solves a resource-constrained shortest-path problem with the above data to find the optimal path.
- Our test instances use the same model for the random data as in Application 1, with 1000 drivers optimizing over a grid graph with 15 rows and 30 columns, and a maximum detour of 25%.



We show the total error as above. Convergence is slower here as the problem is much more complex. Still, in most rounds we have an error close to 0.

References

[1] Y. Freund and R. E. Schapire. Adaptive game playing using multiplicative weights, 1997.

[2] M. Zinkevich. Online convex programming and generalized infinitesimal gradient ascent, 2003.

Contact Information

Andreas.Baermann@fau.de
Alexander.Martin@fau.de
Sebastian.Pokutta@isye.gatech.edu
Oskar.Schneider@fau.de