Online Learning with Knapsacks: the Best of Both Worlds

ICML 2022



Matteo Castiglioni
(Politecnico di Milano)

Andrea Celli
(Bocconi University)

Christian Kroer
(Columbia University)

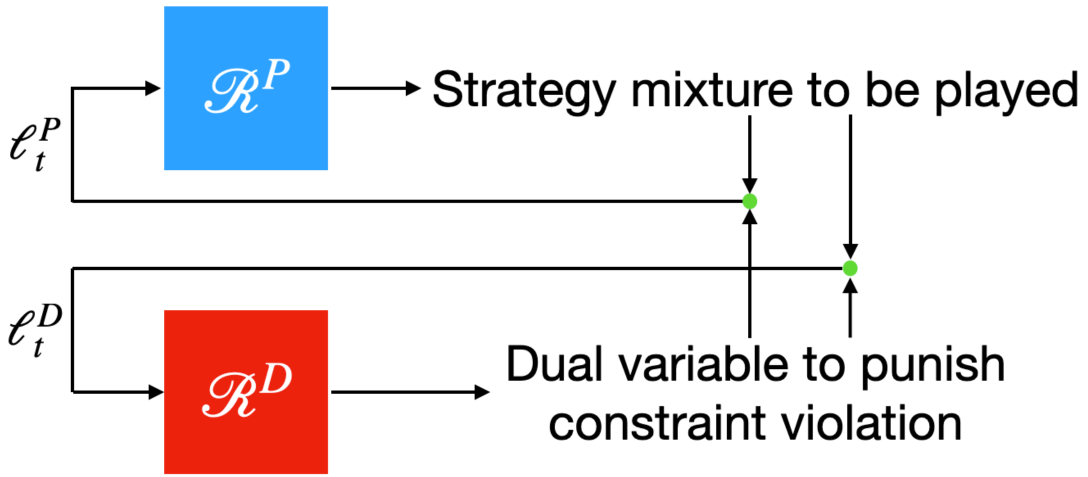
Our setting

- There are T rounds and m resources
- At each round
  1. Pick an action $x_t \in X$
  2. Observe reward $f_t: X \to [0,1]$ and resource consumption $c_t: X \to [0,1]^m$
- For each resource there's a budget B
- **Goal: maximize reward subject to $\sum_{t=1}^{T} \boldsymbol{c}_t(\boldsymbol{x}_t)\,[\boldsymbol{i}] < \boldsymbol{B}$ for each resource $\boldsymbol{i} \in [\boldsymbol{m}]$**
- There exist a void action
- Regime $B = \Omega(T)$
- Framework similar to BwK / ABwK (Badanidiyuru et al., 2013; Agrawal & Devanur, 2014; Immorlica et al., 2019)

The algorithm

- We allow the decision maker to play randomized strategies over their action set X. We call such strategies strategy mixtures Ξ
- Primal/dual approach

- Two regret minimizers
- Primal regret minimizer
  - Outputs the strategy mixture to be played
  - Full or bandit feedback depending on the setting
- Dual regret minimizer
  - Penalizes the primal objective when resource-constraints are violated
  - Full feedback by construction



Main results

- Reward and cost functions **need not be convex** (if there exists a suitable primal regret minimizer)
- **Constant-factor competitive ratio in the adversarial case**
- **Best-of-both-worlds algorithm**: the same algorithm can be applied in the stochastic setting with a regret upper-bound matching that of previous work
- Application to budget-management in **repeated first-price auctions**