2024-Bernasconi-2

No-Regret is not enough! Bandits with General Constraints through Adaptive Regret Minimization

- **Date**: 2024-05-10
- Link
- **Authors:**
  - Bernasconi, Martino
  - Castiglioni, Matteo
  - Celli, Andrea
- **Cites:**
- **Cited by:**
- **Keywords:** #bandits #bandits-with-knapsacks #contextual-bandits-with-linear-constraints #primal-dual-constraints
- **Collections:**
- **Status:** #in-progress

0. Abstract

- Problem description:
  - In the standard bandits with knapsacks (BwK) problem, the learner has m resource-consumption constraints.
  - In this paper, a general BwK problem is considered where the learner has a set of general long-term constraints.
  - **Goal** - maximize cumulative reward while limiting cumulative constraint violations.
- **Problem** - conventional methods fail to yield sublinear violations of constraints.
- **Solution** - circumvent the issue by requiring the primal and dual algorithm to be *weakly adaptive*.
  - **Key Idea** - Without information on Slater's parameter ρ , the interplay between weakly adaptive primal and dual regret minimizers yields a "self-bounding" property of two variables.
- **Results**
  - Stochastic constraints - show that the algorithm guarantees sublinear regret.
  - Adversarial constraints - establish a tight competitive ratio of $\rho/(1+\rho)$.
  - Both cases - constraint violations are guaranteed to be sublinear in time.
  - Contextual bandits with linear constraints - provides the first no-α-regret guarantees for adversarial contexts.