Optimistic Safety for Online Convex Optimization with Unknown Linear Constraints

Spencer Hutchinson
UC Santa Barbara

Tianyi Chen
Rensselaer Polytechnic Institute

Mahnoosh Alizadeh
UC Santa Barbara

CONTRIBUTIONS

We study *online convex optimization with unknown linear constraints* and give algorithms that enjoy $\tilde{O}(\sqrt{T})$ regret and no constraint violation. Our approach leverages *optimism* to maintain low regret while ensuring no violation of the constraint.

Reference	Constraint Variation	Constraint Type	Regret	Violation
[1,2]	static	linear	$T^{2/3}$	no violation
this work	**static**	**linear**	$\sqrt{T}$	**no violation**
[3,4,5,6]	stochastic	convex	$\sqrt{T}$	$\sqrt{T}$ cumulative
this work	**stochastic**	**linear**	$\sqrt{T}$	**no violation**

PROBLEM SETUP

Interaction Model

At each round $t \in [T]$:

1. Choose action $x_t \in \mathcal{X}$.
2. Observe cost function f_t.
3. Observe noisy constraint feedback $y_t = Ax_t + \varepsilon_t$.

Learning Goals

Satisfy constraint in all rounds: $Ax_t \le b \quad \forall t \in [T]$

Minimize regret with respect to feasible set:

$$R_T = \sum_{t=1}^{T} f_t(x_t) - \min_{x \in \mathcal{X}: Ax \le b} \sum_{t=1}^{T} f_t(x)$$

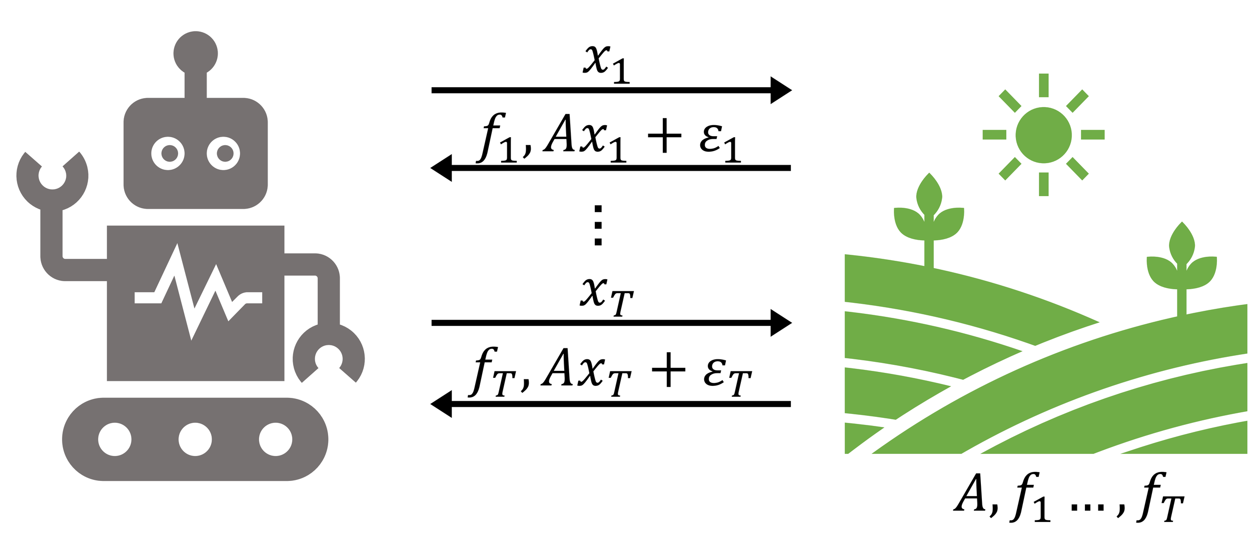
Assumptions

Cost gradients ∇f_t and action set $\mathcal{X}$ are bounded.

Origin is strictly feasible, i.e. $\mathbf{0} \in \mathcal{X}, b > 0$.

Noise ε_t is conditionally-subgaussian.

Constraint matrix A is bounded.



EFFICIENT VERSION

The original approach requires an algorithm for online optimization over a non-convex set, which cannot be computed efficiently. To remedy this:

- Relax 2-norm to infinity-norm in the optimistic set:

$$\mathcal{Y}^o = \left\{x \in \mathcal{X}: \hat{A}x - \sqrt{d}\beta \left\| V^{-1/2}x \right\|_\infty \mathbf{1} \le b\right\}$$

- Parameterize the infinity-norm:

$$\mathcal{Y}^o(k,\xi) = \left\{x \in \mathcal{X}: \hat{A}x - \sqrt{d}\xi\beta \left[V^{-1/2}\right]_k x\mathbf{1} \le b\right\}$$

$$\mathcal{Y}^o = \bigcup_{k\in[d],\xi\in\{-1,1\}} \mathcal{Y}^o(k,\xi)$$

- Simultaneously use gradient descent within $\mathcal{Y}^o(k,\xi)$ and Hedge to choose k, ξ.

Algorithm 2 HedgeDescent

```
input {X_m}_{m∈[M]}.
for t = 1 to T do
    Sample Hedge: m_t ~ p_t.
    Play x_t(m_t) and observe f_t.
    Update experts:
        x_{t+1}(m) = Π_{X_m}(x_t(m) − η_t ∇f_t(x_t(m))).
    Update Hedge:
        p_{t+1}(m) ∝ p_t(m) exp(−ζ_t f_t(x_t(m))).
end for
```

ALGORITHM

Algorithm 1 Optimistically Safe OCO (OSOCO)

```
input Online algorithm A.
Initialize: t = 1, V = λI, S = 0, V̄ = det(V).
while t ≤ T do
    Estimate constraints: Â = SV^{-1}.
    Pessimistic set: Y^p := {x ∈ X : Âx + β‖x‖_{V^{-1}} 1 ≤ b}
    Optimistic set: Y^o := {x ∈ X : Âx − β‖x‖_{V^{-1}} 1 ≤ b}
    Initialize A with Y^o as action set.
    while det(V) ≤ 2V̄ do
        A chooses x̃_t ∈ Y^o.
        Safe scaling: γ_t = max{μ ∈ [0,1] : μx̃_t ∈ Y^p}.
        Play x_t = γ_t x̃_t and observe f_t, y_t.
        Send f_t to A.
        V = V + x_t x_t^⊤, S = S + y_t x_t^⊤.
        t = t + 1.
    end while
    V̄ = det(V)
    Terminate A.
end while
```

Constraint Estimation

Least-squares estimator $\hat{A}$ provides high-probability confidence bounds on the constraint:

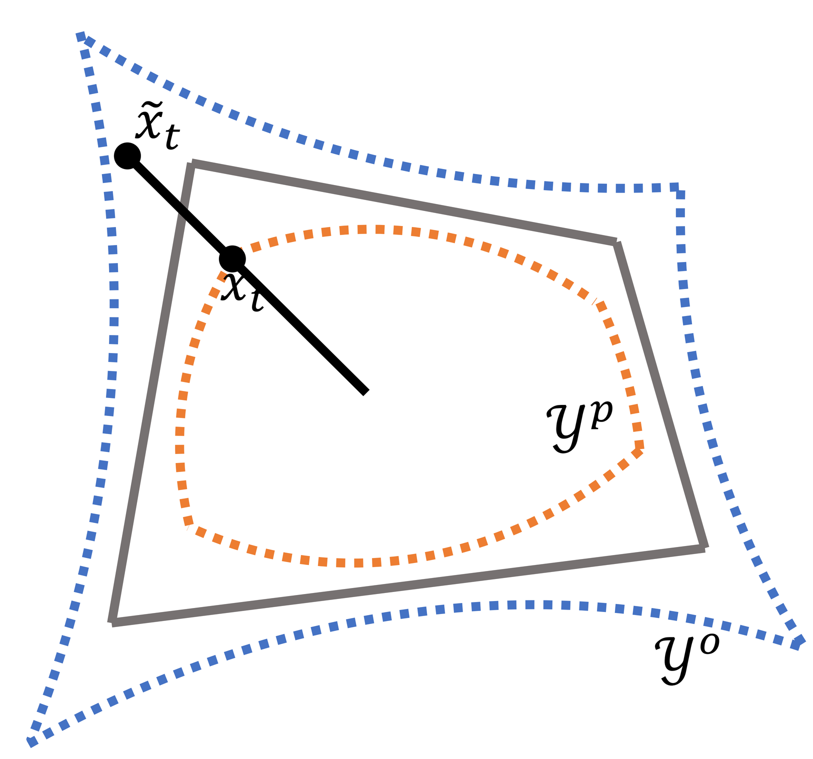
$$\hat{A}x - \beta\|x\|_{V^{-1}}\mathbf{1} \le Ax \le \hat{A}x + \beta\|x\|_{V^{-1}}\mathbf{1}$$

Optimistic and Pessimistic Sets

Use confidence bounds to construct a set that overestimates the constraint (*optimistic set* $\mathcal{Y}^o$) and a set that underestimates the constraint (*pessimistic set* $\mathcal{Y}^p$).

Optimistic Action

Use a generic online optimization algorithm $\mathcal{A}$ to choose an action $\tilde{x}_t$ from optimistic set $\mathcal{Y}^o$.



Safe Scaling

Scale $\tilde{x}_t$ until it is in the pessimistic set, and play the resulting action x_t.

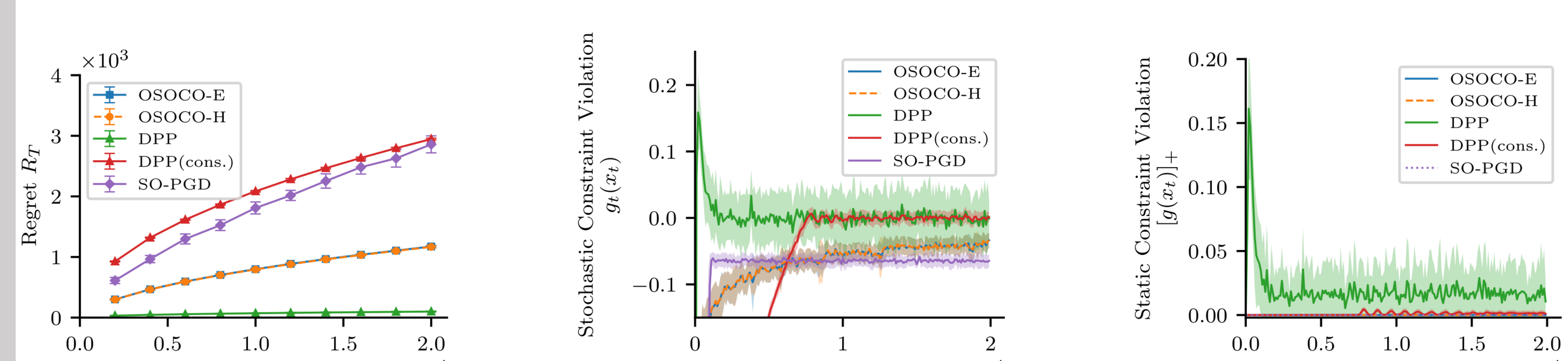
Rare updating

Only need to update optimistic and pessimistic sets when determinant of Gram matrix $\det(V)$ doubles.

STOCHASTIC CONSTRAINTS

Our approach extends to stochastic linear constraints of the form $g_t(x) = A_t x - b_t$. In this case, we guarantee $\tilde{O}(\sqrt{T})$ regret and no expected violation $\mathbb{E}[g(x_t)] \le 0$, assuming that $(A_t, b_t)_t$ is i.i.d., A_t is bounded, b_t is bounded, $\mathbb{E}[b_t]$ is positive, and $\mathbb{E}[b_t]$ is known.

EXPERIMENTS



REFERENCES

[1] Chang et al. Dynamic regret analysis of safe distributed online optimization for convex and non-convex problems. TMLR 2023.
[2] Chaudhary and Kalathil. Safe online convex optimization with unknown linear safety constraints. AAAI 2022.
[3] Yu et al. Online convex optimization with stochastic constraints. NeurIPS 2017.
[4] Wei et al. Online primal-dual mirror descent under stochastic constraints.
[5] Castiglioni et al. A unifying framework for online optimization with long-term constraints. NeurIPS 2022.
[6] Yu et al. Online bandit convex optimisation with stochastic constraints via two-point feedback. IJSS 2023.

This work was supported by NSF grant #2330154.