Saddle Property of Optimal $\mathcal{RO}$ Solution

As the $\mathcal{RO}$ problem has two levels, we have different saddle property related to the bi-level optimization problem. For the $\mathcal{RO}$ problem as a whole, there is a saddle property based on the Lagrangian function (16), which we call the $\mathcal{RO}$ saddle property. In the context of Lagrangian duality, a saddle point of the Lagrangian function is a point where the function is minimized with respect to the convex variables and maximized with respect to the concave variables.

Remark 9 (Non-standard Saddle Point Property)**.** *Classical results such as Sion's minimax theorem [35] or Rockafellar's saddle point theorem [36] require joint concavity in the maximization variables. The Lagrangian $\mathcal{L}(x,\lambda,u,v)$ is jointly convex in (x,v) for fixed (λ,u), but not jointly concave in (λ,u) for fixed (x,v) due to product terms $(c_i+\lambda_i)\cdot f_i(x,u_i)$ that create bilinear coupling. This violation of joint concavity renders existing primal-dual methods [27], [28] inapplicable and requires a proof that exploits the min-max-max-min structure of robust optimization.*

Lemma 1. *[Saddle property] Let $z^\star=(x^\star,\lambda^\star,u^\star,v^\star)$ be an optimal $\mathcal{RO}$ solution. Then, for all $x,\lambda\geq 0,u,v\geq 0$, $z^\star$ satisfies the $\mathcal{RO}$ saddle property, namely,*

$$\mathcal{L}(x^\star,\lambda,u,v^\star)\leq\mathcal{L}(x^\star,\lambda^\star,u^\star,v^\star)\leq\mathcal{L}(x,\lambda^\star,u^\star,v)\ . \tag{22}$$

V. Dynamical system solving $\mathcal{RO}$

Motivation and Challenges

So far, we have characterized the "optimization" properties of the Robust Optimization problem under study. In this paper, we are interested in understanding if there is a continuous-time dynamical system that can solve $\mathcal{RO}$ and how it would operate. Our main motivations are two: (1) Understanding from a dynamical system perspective how $\mathcal{RO}$ can be solved. (2) Studying how physically interacting systems with very simple capabilities and or "intelligence" can cooperate to solve complex optimization problems (learning, estimation, and decision) well outside the single element capabilities. While there are now answers to these questions for large classes of convex optimization problems [28], [37], this is the first work, to the best of our knowledge, that addresses $\mathcal{RO}$ problems. Our task turned out to be quite non-trivial, as explained below.

The basic method to obtain a continuous-time dynamics that solves a constrained convex optimization problem goes back to [27]. The main idea is quite intuitive. The primal dynamics evolves with the negative gradient of the problem's Lagrangian function, w.r.t, the primal variable, that is x, while the dual dynamics evolves with the positive gradient of the Lagrangian w.r.t to the dual variables, that is, λ_is. The primal descent and dual ascent dynamics is globally converging to the optimal solution under minor assumptions. The proof is based on a simple quadratic Lyapunov function. However, from a system point of view, it is the passivity of the gradient of a convex function that provides the convergence mechanism [38], [39].

For the $\mathcal{RO}$ problem, the standard approach does not work due to the nested structure of optimization. Thus, the natural match between intuitive primal-descent dual-ascent dynamics via a traditional quadratic Lyapunov function is broken. It turns out that it is not easy to find the right combination of dynamics and Lyapunov function that show global convergence.

$\mathcal{RO}$ Dynamics

This section presents the continuous-time dynamical system ($\mathcal{RO}$ dynamics) whose solutions globally converge to robust optimal solutions. Let

$$M=\sum_{i=0}^{N}m_i\ ,\ K=\sum_{i=0}^{N}K_i\ .$$

Consider the following $\mathcal{RO}$ dynamics defined on $\mathbb{S}:=\mathbb{R}^n\times\mathbb{R}_+^N\times\mathbb{R}^M\times\mathbb{R}_+^K$

$$\begin{cases}\dot{x}=-\nabla_x f_0(x,u_0)-\sum_{i=1}^{N}(c_i+\lambda_i)\nabla_x f_i(x,u_i)\\ \dot{\lambda}_i=[f_i(x,u_i)-v_i^\top h_i(u_i)]^+_{\lambda_i},\ i\in[N]^+\\ \dot{u}_i=\nabla_{u_i}f_i(x,u_i)-\sum_{j=1}^{K_i}v_{ij}\nabla_{u_i}h_{ij}(u_i),\ i\in[N]\\ \dot{v}_0=[h_0(u_0)]^+_{v_0}\\ \dot{v}_i=[(c_i+\lambda_i)\ h_i(u_i)]^+_{v_i},\ i\in[N]^+\end{cases}. \tag{23}$$

The $\mathcal{RO}$ dynamics has state vector $z:=(x,\lambda,u,v)\in\mathbb{S}$ with the following structure:

- The x dynamics: gradient descent on the Lagrangian with respect to primal variables.
- The λ_i dynamics: gradient ascent on dual variables for robust constraints, with projection to maintain $\lambda_i\geq 0$.
- The u_i dynamics: gradient ascent on uncertainty variables to find worst-case realizations.
- The v_i dynamics: gradient ascent on dual variables for uncertainty set constraints with projection to maintain $v_i\geq 0$.

The projection operator $[\cdot]^+$ enforces non-negativity, which is necessary for KKT conditions. The coupling through $(c_i+\lambda_i)$ balances constraint satisfaction and objective minimization.

Illustrative Examples

It is interesting to describe the structure using special examples.

1) Min-max problem: Consider the following $\mathcal{RO}$ problem with no constraints

$$\mu=\min_x\max_{u_0:h_0(u_0)\leq 0}f_0(x,u_0)\ .$$

Such problems are popular in machine learning. In this case, the continuous-time dynamical system is given by

$$\begin{cases}\dot{x}=-\nabla_x f_0(x,u_0)\\ \dot{u}_0=\nabla_{u_0}f_0(x,u_0)-v_0^\top\nabla_{u_0}h_0(u_0)\\ \dot{v}_0=[h_0(u_0)]^+_v\end{cases}.$$