KKT-Informed Neural Network

A parallel solver for parametric convex optimization problem

A Preprint

Carmine Delle Femine
Department of Data Intelligence for Energy and Industrial Processes
Vicomtech Foundation
Donostia-San Sebastián, 20009
cdellefemine@vicomtech.org

September 4, 2024

Abstract

This is the abstract

***K*eywords**
Optimization

1 Introduction

2 Background

Consider a parametric convex optimization problem in the standard form:

$$\begin{aligned} \min_{x \in \mathcal{D} \subseteq \mathbb{R}^n} \quad & f(x, \theta) \\ \text{s.t.} \quad & g_i(x, \theta) \leq 0 \quad i = 1, \dots, m \\ & A(\theta)x - b(\theta) = 0 \end{aligned}$$

where $x \in \mathcal{D} \subseteq \mathbb{R}^n$ is the optimization variable; $\theta \in \mathcal{D}_\theta \subseteq \mathbb{R}^k$ are the parameters defining the problem; $f : \mathcal{D}_f \subseteq \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}$ is the convex cost function; $g_i : \mathcal{D}_{g_i} \subseteq \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}$ are the convex inequality constraints, $A : \mathcal{D}_\theta \to \mathbb{R}^{p \times n}$ and $b : \mathcal{D}_\theta \to \mathbb{R}^p$ defines the affine equality constraints and $\mathcal{D} = \bigcap_{i=1}^m \mathcal{D}_{g_i} \cap \mathcal{D}_f$ is the domain of the optimization problem.

Assume differentiable cost and constraints functions and that g_i satisfies Slater's condition. Given a set of parameters θ, $x^* \in \mathcal{D}$ is optimal if and only if there are λ^* and ν^* that, with x^*, satisy the Karush-Kuhn-Tucker conditions (KKT):

$$A(\theta)x^* - b(\theta) = 0 \tag{1}$$

$$g_i(x^*, \theta) \leq 0 \quad i = 1, \dots, m \tag{2}$$

$$\lambda_i^* \geq 0 \quad i = 1, \dots, m \tag{3}$$

$$\lambda_i^* g_i(x^*, \theta) = 0 \quad i = 1, \dots, m \tag{4}$$

$$\nabla_{x^*} f(x^*, \theta) + \sum_{i=1}^m \lambda_i^* \nabla_{x^*} g_i(x^*, \theta) + A^T \nu^* = 0 \tag{5}$$

3 Proposed method

KKT-Informed Neural Network (KINN) builds upon the principles of Physics-Informed Neural Networks (PINNs), incorporating mathematical conditions of the Karush-Kuhn-Tucker (KKT) conditions directly into the neural architecture. This integration facilitates a disciplined learning process where the network not only predicts optimization variables but also ensures these predictions are compliant with KKT conditions, essential for guaranteeing the optimality of solutions in convex optimization under exam.

Network architecture is a MLP designed to take problem parameters θ as input and predict x^*, λ^*, ν^*. A ReLU function is placed at the end of the branch predicting λ^* to ensure its feasability.

$$\hat{x}, \hat{\lambda}, \hat{\nu} = \text{KINN}(\theta) \tag{6}$$
$$\hat{\lambda} \in \mathbb{R}^0_+ \tag{7}$$

Loss function is so defined:

$$\mathcal{L} = \mathcal{L}_S + \sum_{i=1}^{m} \mathcal{L}_{I,i} + \mathcal{L}_E + \sum_{i=1}^{m} \mathcal{L}_{C,i}$$

where:

$$\mathcal{L}_S = \|\nabla_{\hat{x}} f(\hat{x}, \theta) + \textstyle\sum_{i=1}^{m} \hat{\lambda}_i \nabla_{\hat{x}} g_i(\hat{x}, \theta) + A^T \hat{\nu}\| \tag{8}$$
$$\mathcal{L}_{I,i} = \|\max(0, g_i(\hat{x}, \theta))\| \tag{9}$$
$$\mathcal{L}_E = \|A(\theta)\hat{x} - b(\theta)\| \tag{10}$$
$$\mathcal{L}_{C,i} = \|\hat{\lambda}_i g_i(\hat{x}, \theta)\| \tag{11}$$
$$\tag{12}$$

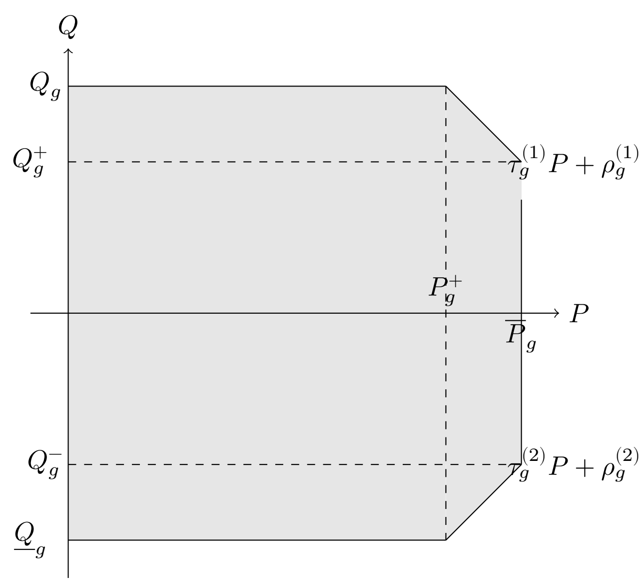
4 Case study

Let us take such a problem as a test case for this approach:

You have a renewable energy generator in a power grid, whose active and reactive power injections are controllable. The set of injection points (P, Q) is limited by physical constraints, however, so the set-points (a_P, a_Q) must be projected onto that set.

The feasibile set $\mathcal{D}$ is:

$$\mathcal{D} = \{(P, Q) \in \mathbb{R}^2 | \underline{P}_g \leq P \leq P^{(\max)}_{g,t}\} \tag{13}$$



4.1 Problem description

4.2 Experimental results

5 Conclusions