DC3: A Learning Method for Optimization with Hard Constraints

Priya L. Donti[1,*], David Rolnick[2,*], J. Zico Kolter[1,3]

[1]Carnegie Mellon University, [2]McGill University and Mila, [3]Bosch Center for AI

Motivation

Large **nonconvex optimization problems with hard constraints** occur in many real-world applications, e.g. electrical grids

Traditional solvers are slow - **fast approximate solvers** needed

Ideal solvers are **differentiable** for integration into larger systems

Naive deep learning methods (e.g. soft loss for constraint violations) do not **guarantee feasibility**

Method

DC3: a **general DL framework** for optimization under hard constraints

1. Neural net predicts **partial set** of variables
2. **"Completion"** to full set of variables, by solving equality constraints (e.g. Newton's method)
3. **"Correction"** to satisfy inequality constraints, by gradient descent on output variables

Trained end-to-end via implicit function thm to optimize "soft loss"

Synthetic experiments

Comparison against traditional opt on (1) convex QPs, (2) nonconvex optimization with linear constraints

Feasibility: Satisfies constraints, unlike other DL-based methods tested

Optimality: Good objective value, 8-11% optimality gap

Speed: (1) Convex QP: 78× faster than DL-based convex optimizer qpth, (2) Non-convex task: 9× faster than traditional optimizer IPOPT

AC Optimal Power Flow

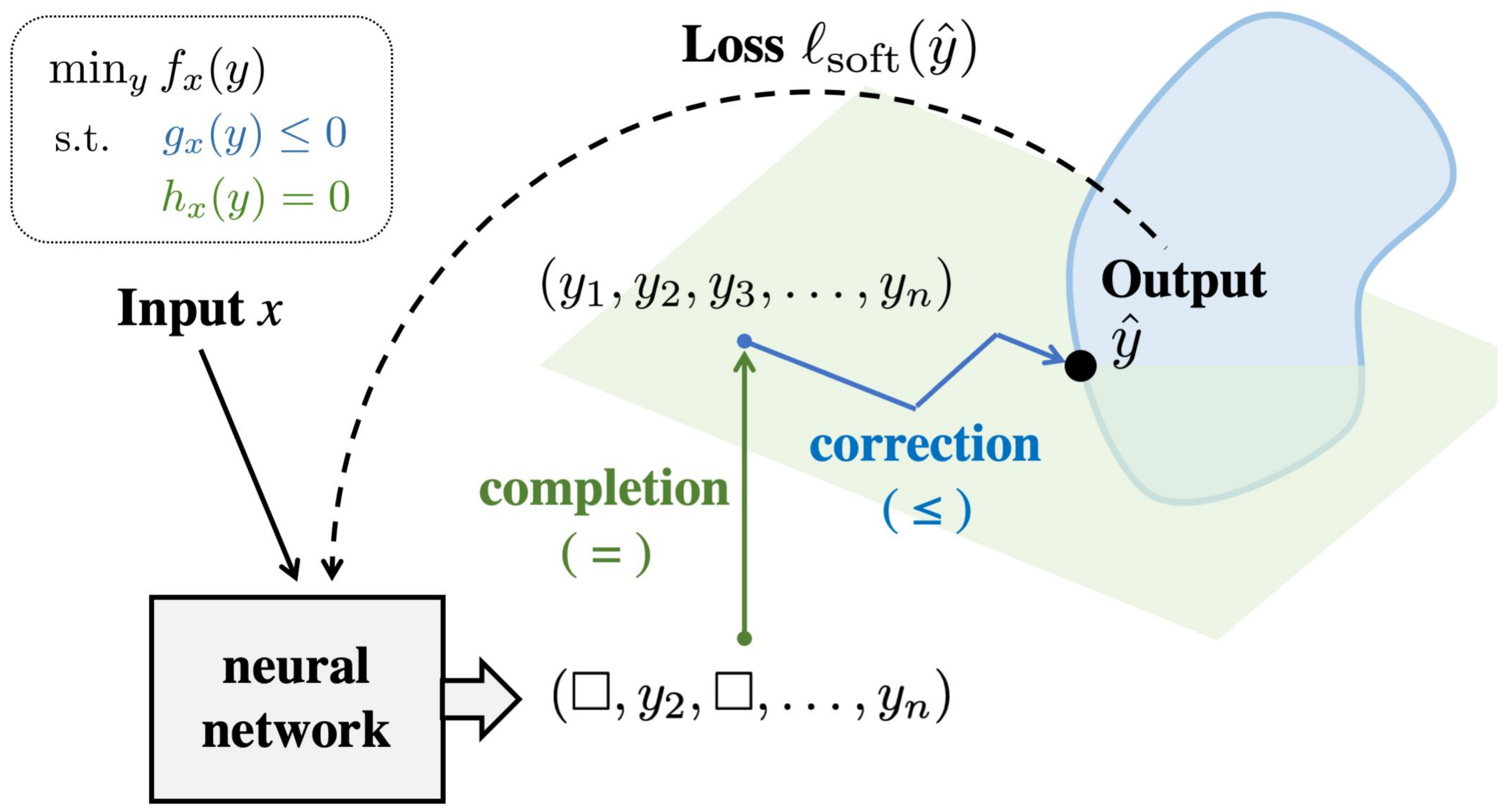
Nonconvex QP (with quadratic constraints)

Important for managing electrical grid, esp. with **renewable energy**

Feasibility: Satisfies constraints, unlike other DL-based methods tested

Optimality: Near perfect, 0.22% optimality gap

Speed: 10× faster than traditional optimizer PYPOWER



	Obj. value	Max eq.	Mean eq.	Max ineq.	Mean ineq.	Time (s)
Optimizer	3.81 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.949 (0.002)
DC3	3.82 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.089 (0.000)
DC3, $\neq$	3.67 (0.01)	0.14 (0.01)	0.02 (0.00)	0.00 (0.00)	0.00 (0.00)	0.040 (0.000)
DC3, $\not\leq$ train	3.82 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.089 (0.000)
DC3, $\not\leq$ train/test	3.82 (0.00)	0.00 (0.00)	0.00 (0.00)	0.01 (0.00)	0.00 (0.00)	0.039 (0.000)
DC3, no soft loss	3.11 (0.05)	2.60 (0.35)	0.07 (0.00)	2.33 (0.33)	0.03 (0.01)	0.088 (0.000)
NN	3.69 (0.02)	0.19 (0.01)	0.03 (0.00)	0.00 (0.00)	0.00 (0.00)	0.001 (0.000)
NN, $\leq$ test	3.69 (0.02)	0.16 (0.00)	0.02 (0.00)	0.00 (0.00)	0.00 (0.00)	0.040 (0.000)
Eq. NN	3.81 (0.00)	0.00 (0.00)	0.00 (0.00)	0.15 (0.01)	0.00 (0.00)	0.039 (0.000)
Eq. NN, $\leq$ test	3.81 (0.00)	0.00 (0.00)	0.00 (0.00)	0.15 (0.01)	0.00 (0.00)	0.078 (0.000)

$$
\begin{aligned}
\underset{p_g \in \mathbb{R}^b,\ q_g \in \mathbb{R}^b,\ v \in \mathbb{C}^b}{\text{minimize}} \quad & p_g^T A p_g + b^T p_g \\
\text{subject to} \quad & p_g^{\min} \leq p_g \leq p_g^{\max} \\
& q_g^{\min} \leq q_g \leq q_g^{\max} \\
& v^{\min} \leq |v| \leq v^{\max} \\
(p_g - p_d) + (q_g - q_d)i &= \operatorname{diag}(v)\overline{W}\overline{v}.
\end{aligned}
$$