Developing Data-driven Methods for Mixed-integer Nonlinear Programs

Xuan Lin and Dennis W. Hong

***Abstract*—Robot motion planning and control problems tend to include discrete decision-making such as contact scheduling and continuous nonlinear (non-convex) programs such as kinematics and dynamics constraints. Simultaneously planning discrete and continuous non-convex variables are known as mixed-integer nonlinear programs (MINLPs). MINLPs of practical sizes are typically challenging. In this work, we compare the performance of several methods including the Alternating Direction Method of Multipliers (ADMM), data-driven mathematical programs with complementary constraints, and data-driven mixed-integer programs with envelopes, to solve a 2D bookshelf organization problem formulated as a mixed-integer bilinear program. The final goal is to develop a data-driven pipeline for solving MINLPs.**

I. Introduction

The problems we are interested in solving are mixed-integer bilinear programs parametrized by Θ that is drawn from a distribution $D(\Theta)$. For each Θ, we seek a solution to the optimization problem:

$$\begin{aligned} &\underset{\mathbf{x},\ \mathbf{z}}{\text{minimize}}\ f_{obj}(\mathbf{x},\mathbf{z};\Theta) \\ \text{s. t.}\ \ & f_i(\mathbf{x},\mathbf{z};\Theta) \leq 0, \quad i=1,...,m_f \\ & b_j(\mathbf{x},\mathbf{z};\Theta) \leq 0, \quad j=1,...,m_b \end{aligned} \tag{1}$$

Where $\mathbf{x}$ denotes continuous variables and $\mathbf{z}$ binary variables with $z_i \in \{0,1\}$ for $i = 1,...,dim(\mathbf{z})$. Constraints f_i are mixed-integer convex, meaning if the binary variables $\mathbf{z}$ are relaxed into continuous variables $\mathbf{z} \in [0,1]$, f_i becomes convex. Constraints b_j are mixed-integer bilinear, meaning that relaxing the binary variables gives bilinear constraints.

There are two directions to convert this problem to either mixed-integer convex programs (MICPs) or mathematical programs with complementary constraints (MPCCs). The MICP approach converts the bilinear constraint b into mixed-integer linear constraints using McCormick envelopes. The MPCC approach converts integer variables into continuous variables with additional complementary constraints. While MICPs can guarantee a global optimal solution, McCormick envelope constraints make them very slow. On the other hand, MPCCs only seek local optimal solutions, but almost impossible to solve without a good initial guess. For both approaches, presolved data becomes very helpful. Data can uncover important regions to reduce the size of MICPs. In the best case, it can be reduced to a convex optimization problem. A learner can be trained offline to map the problem parameter Θ to a good candidate solution. The solution will be modified online for solving new problem instances. Data can also provide good initial guesses that dramatically increase the chance of success to solve MPCCs.

However, it is difficult to collect enough data for warm-start MICP or MPCC solvers, especially when the problem becomes larger. We also demonstrate using ADMM to solve the mixed-integer bilinear programs. This approach does not depend on data, and have decent chance of convergence thus can be used to collect the initial dataset.

II. Bookshelf Problem

Fig. 1 shows the complete formulation of 2D bookshelf organization problem. The main goal of this problem is to use robot manipulator to insert one book into the shelf such that minimal movement of existing books is required. The complete formulation of this problem is given in 1. This problem contains integer variables representing different contact states of the books, and continuous variables related to rotation matrices and collision avoidance. In particular, this formulation formulates collision avoidance using separating planes into bilinear inequality constraints. This approach is applicable to multi-agent planning problems. This problem may have a solving time limitation and optimal cost requirement, as a less optimal motion may take significantly longer for the manipulator to finish. We demonstrate the aforementioned approaches on this problem and compare their performances.

III. Approaches

A. ADMM Formulation

We implemented the non-data-driven algorithm ADMM in [1]. The basic idea is to separate the MINLP into a MIP formulation and an NLP formulation with the exact copies of variables but different constraints, and iterate between them until the consensus is reached. The MIP formulation contains constraints $A, B, D, G, H1, I1, J1, K2, K3, L2, L3$, while missing bilinear constraints $C, E, F, K1, L1$. The NLP formulation contains constraints $A, B, C, D, E, F, H, I, J, K, L$, while missing the mode constraint G which says z_is are integers. The ADMM iterates the following 3 steps: first it solves the MIP formulation, then solves the NLP formulation. Finally it updates the dual variables. The process continues till convergence. The readers are referred to [1] for details. The results are shown in II column 1, 2.

B. MPCC Formulation

An equivalent formulation of the bookshelf organization problem in Fig. 1 is to turn all the integer variables z_i into continuous variables with complementary constraints, such that the complete formulation can be solved through NLP

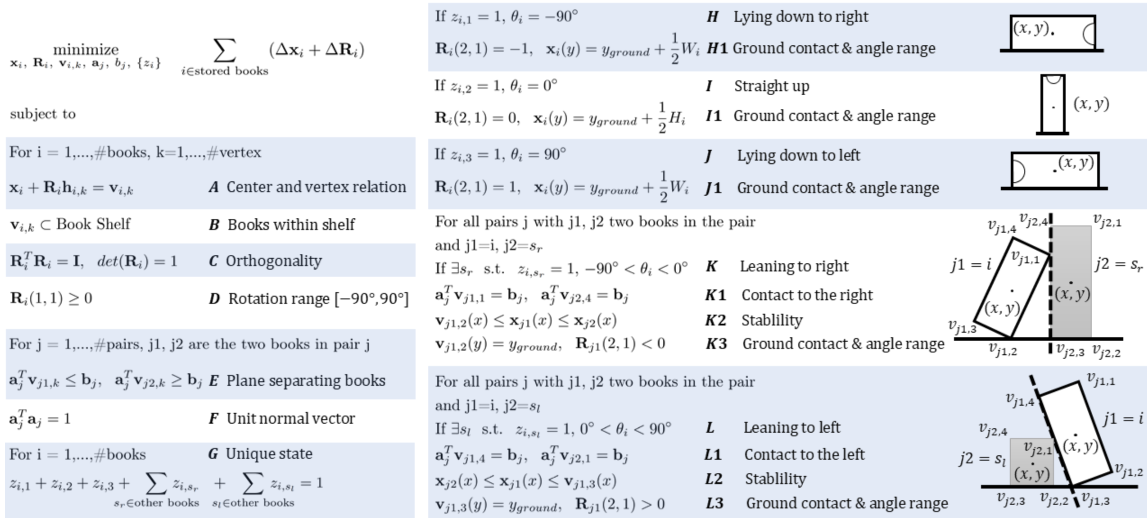

Fig. 1. Complete formulation of the bookshelf organization problem.

	ADMM - Knitro	ADMM - IPOPT	MPCC_default - Knitro	MPCC_default - IPOPT	MPCC_Manual - Knitro	MPCC_Manual - IPOPT	MPCC_KNN3 - Knitro (15000data)	MPCC_KNN3 - IPOPT (15000 data)	MPCC_KNN3 (Rev) - Knitro (36000 data)	ReDUCE - 200 (15000 data)	ReDUCE - 600 (15000 data)	ReDUCE - 200 (78000 data)
Note			Using default initial guess from CasADi		Using manually selected 1 initial guess		KNN selects top 3 guesses from 15000 data		Same to left but select the worst 3 guesses	Max iteration of OSQP=200	Max iteration of OSQP=600	Currently KNN learner
Success Rate	96.5%	94.75%	1.25 %	0 %	78.25%	78.25%	99.25%	98.5%	92.5%	94.5%	98.75%	96%
Avg. Solve Time	260 ms	522 ms	72 ms		29 ms	81 ms	30 ms	96 ms	61 ms	65 ms	80 ms	16 ms
Max. Solve Time	1.29 sec	3.59 sec	82 ms		198 ms	1.96 sec	280 ms	990 ms	327 ms	190 ms	400 ms	174 ms
Avg. Objective	-7250	-7339	-5703		-8620	-8952	-8258	-8618	-6461	-8631	-8687	-8637
Avg. # of iterations	4.73	4.6			1	1	1.05	1.06	1.25	4.85	3.85	2.57
Solver	Gurobi+Knitro	Gurobi+IPOPT	Knitro	IPOPT	Knitro	IPOPT	Knitro	IPOPT	Knitro	OSQP	OSQP	OSQP

solvers. We implemented the formulation from [4]. Specifically $z_i(1 - z_i) = 0$, which is equivalent to $z_i \in \{0, 1\}$.

To get MPCC working, proper initial guesses are required. Using the solver's default initial guess will result in a success rate lower than 5%. We take 2 different approaches to supply better initial guesses. The first approach is choosing from heuristics. For this problem, we choose the initial guess to be the original scene as the problem minimizes the movement of books. The second approach is collecting a dataset offline and using k-nearest neighbor to pick the top 3 candidate solutions online for initial guess, which is data-driven. The results are shown in Table II column $4 - 10$.

C. Convex Formulation

We implemented the approach given by [3], named ReDUCE, which converts bilinear constraints into mixed-integer envelope constraints and uses a data-driven approach to reduce the size of integer problems. In the best case, it can directly learn a mapping to a convex problem in the same fashion as [2]. The results are shown in II column $11 - 13$.

IV. Conclusion

The results show that the non-data-driven ADMM approach has good chance of convergence and decent solving speed. The rate of success is significantly higher than a poorly initialized MPCC and significantly faster than a MICP formulation without warm-start. The MPCC formulation has good performance with data of thousands, mainly due to the strong capability of NLP solvers to find feasible solutions. In most cases, NLP solver takes only 1 trial to find a feasible solution. However, the solving speed seems to be limited. Finally, if the algorithm can learn a mapping to a convex problem well, the convex approach has potential to solve most of the test problems with 1 or 2 trials. One can leverage on the strong performance of convex solvers to achieve fastest speed. However, this approach requires larger amount of data for more precise learning.

The proposed algorithms are being tested on more challenging robotics problems such as modular self-configurable robot systems and hybrid MPC. Some results are under review.

References

[1] Alp Aydinoglu and Michael Posa. Real-time multi-contact model predictive control via admm. *arXiv preprint arXiv:2109.07076*, 2021.

[2] Abhishek Cauligi, Preston Culbertson, Edward Schmerling, Mac Schwager, Bartolomeo Stellato, and Marco Pavone. Coco: Online mixed-integer control via supervised learning. *IEEE Robotics and Automation Letters*, 7 (2):1447–1454, 2021.

[3] Xuan Lin, Gabriel I Fernandez, and Dennis W Hong. Reduce: Reformulation of mixed integer programs using data from unsupervised clusters for learning efficient strategies. *arXiv preprint arXiv:2110.00666*, 2021.

[4] Jaehyun Park and Stephen Boyd. A semidefinite programming method for integer convex quadratic minimization. *Optimization Letters*, 12(3):499–518, 2018.