

機械学習を利用した混合整数最適化問題のオンライン高速求解

武井 柊悟[†] 鮎川 矩義[‡] 高野 祐一[§]
 筑波大学[†] 法政大学[‡] 筑波大学[§]

1 はじめに

混合整数凸最適化問題 (mixed-integer convex optimization: MICO) とは, 実数変数と整数変数が混在し, 目的関数と制約関数が凸関数となる最適化問題である. MICO は NP 困難な問題であるため, 短時間で厳密解を求めることは困難であることが一般には予想される.

オンライン最適化とは, 目的変数や制約条件が逐次的に変化する最適化である [4]. 標準的な最適化とは異なり, オンライン最適化では逐次的なパラメータ更新に応じて高速に解を更新することが求められる.

オンラインで MICO を高速に解く方法として, MLOPT [1] や CoCo [2] といったフレームワークが考案されている. MLOPT はニューラルネットワーク (neural network: NN) を利用してオンラインで MICO を高速に解くためのフレームワークである. 事前に, 制約関数の係数などの問題パラメータを様々なパターンで用意し, 各パラメータに対して MICO を求解しておく. 次に, 問題パラメータを説明変数, 事前に求解した結果での整数変数の値を目的変数として NN の学習をしておく. オンラインで MICO を解く際には, NN が予測した整数解を MICO に代入することによって, 実数変数のみの凸最適化問題 (convex optimization: CO) に変形して求解する. CoCo は MLOPT を発展させたフ

レームワークであり, 画像データを対象とした深層学習モデルである畳み込みニューラルネットワーク (convolutional neural network: CNN) を使用している.

CoCo の欠点として NN の学習時間と, 求解成功割合 (実行可能解を出力する割合) が挙げられる. NN を短時間で学習し, 求解成功割合を向上させることを本研究の目的とする.

2 問題設定

MICO の実例として, 2 次元平面上を移動するロボットの経路生成を考える. 初期地点, 目的地点, 長方形の障害物が事前に与えられ, ロボットは障害物を避けながら初期地点から目的地点まで進む. 障害物は 1 つの画像内に N_m 個存在し, 各障害物 m は左下頂点の座標 (x_{\min}^m, y_{\min}^m) , 右上頂点の座標 (x_{\max}^m, y_{\max}^m) で定義される.

ステップ $t = 0, 1, \dots, N_t$ のロボットの位置を $\mathbf{p}_t \in \mathbb{R}^2$, 速度を $\mathbf{v}_t \in \mathbb{R}^2$ とし, 状態を $\mathbf{x}_t = (\mathbf{p}_t, \mathbf{v}_t)$ とする. また, 状態の最大値と最小値をそれぞれ $\mathbf{x}_{\max}, \mathbf{x}_{\min}$ とする. エンジンの噴出力などロボットへの入力を $\mathbf{u}_t \in \mathbb{R}^2$, 入力の上限を u_{\max} , ロボットの初期状態を \mathbf{x}_{init} , 目的状態を \mathbf{x}_{goal} とすると, 経路生成の MICO は以下のように定式化できる.

目的関数 (1) は最終的な状態, 移動中の状態それぞれを目的状態に近づけ, ロボットへの入力を小さくする. ただし, $\lambda_1, \lambda_2, \lambda_3$ は目的関数の各項の重みとする. 式 (2) はロボットの状態遷移を表し, 行列 $\mathbf{A} \in \mathbb{R}^{4 \times 4}$, $\mathbf{B} \in \mathbb{R}^{4 \times 2}$ には問題に応じた定数が格納されている. 式 (3) は入

Fast online solution of mixed-integer optimization problems using machine learning

[†] Shugo Takei, University of Tsukuba

[‡] Noriyoshi Sukegawa, Hosei University

[§] Yuichi Takano, University of Tsukuba

力の上限, 式 (4) は状態の範囲を表す. 式 (5) は初期状態を定義する. 式 (6), (7) は $\delta_t^{m,i} = 0$ となる時, ステップ t においてロボットが障害物 m の面 i 側に存在することを示す. また, ロボットが障害物に衝突しないように式 (8) で制御している.

$$\min \lambda_1 \|\mathbf{x}_{N_t} - \mathbf{x}_{\text{goal}}\|_2 + \lambda_2 \sum_{\tau=0}^{N_t-1} \|\mathbf{x}_\tau - \mathbf{x}_{\text{goal}}\|_2 + \lambda_3 \|\mathbf{u}_\tau\|_2 \quad (1)$$

$$\text{s.t. } \mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{B}\mathbf{u}_t, \quad t = 0, 1, \dots, N_t - 1 \quad (2)$$

$$\|\mathbf{u}_t\|_2 \leq u_{\max}, \quad t = 0, 1, \dots, N_t - 1 \quad (3)$$

$$\mathbf{x}_{\min} \leq \mathbf{x}_t \leq \mathbf{x}_{\max}, \quad t = 0, 1, \dots, N_t \quad (4)$$

$$\mathbf{x}_0 = \mathbf{x}_{\text{init}} \quad (5)$$

$$x_{\max}^m - M\delta_t^{m,1} \leq x_t^1 \leq x_{\min}^m + M\delta_t^{m,2}, \quad t = 0, 1, \dots, N_t, m = 1, 2, \dots, N_m \quad (6)$$

$$y_{\max}^m - M\delta_t^{m,3} \leq x_t^2 \leq y_{\min}^m + M\delta_t^{m,4}, \quad t = 0, 1, \dots, N_t, m = 1, 2, \dots, N_m \quad (7)$$

$$\sum_{i=1}^4 \delta_t^{m,i} \leq 3, \quad t = 0, 1, \dots, N_t, m = 1, 2, \dots, N_m \quad (8)$$

$$\delta_t^{m,i} \in \{0, 1\}, \quad t = 0, 1, \dots, N_t, m = 1, 2, \dots, N_m, i = 1, 2, 3, 4 \quad (9)$$

3 Set Transformer の学習

Set Transformer [3] とは集合を扱うことができる深層学習モデルである.

CoCo では 1 つの画像内の障害物ごとに, 回避戦略を CNN (多クラス分類) によって学習していた. 本研究では Set Transformer によって, 1 つの画像内の全障害物を集合とみなして一度に学習する. これにより, CNN と比較して学習の回数を障害物の数だけ減らすことができ, 学習時間が短縮される.

4 回避戦略の選択アルゴリズム

以降, 整数変数の最適解を回避戦略と呼ぶ. NN は各障害物に対応する回避戦略の予測確率

を出力する. 各障害物について最も予測確率が高い回避戦略のみを使用すると, CO を解く際に実行不可能になる可能性がある. 対処法として, いくつかの回避戦略を評価することが考えられる.

効率的に各障害物に回避戦略を割り当てる方法として, 予測確率の高い回避戦略がある障害物にはその回避戦略を使用し, 予測確率の高い回避戦略が存在しない障害物には, 様々な回避戦略を割り当てるアルゴリズムを提案する.

5 数値実験

提案手法の性能を CoCo と比較する. 学習時間は Set Transformer を使用することで CNN よりも短縮され, 求解成功割合は回避戦略の選択アルゴリズムを使用することで向上することを確認する. 実験に使用したデータおよび結果の詳細は当日報告する.

参考文献

- [1] Bertsimas, D., & Stellato, B. (2022). Online mixed-integer optimization in milliseconds. *INFORMS Journal on Computing*, 34(4), 2229–2248.
- [2] Cauligi, A., Culbertson, P., Schmerling, E., Schwager, M., Stellato, B., & Pavone, M. (2021). CoCo: Online mixed-integer control via supervised learning. *IEEE Robotics and Automation Letters*, 7(2), 1447–1454.
- [3] Lee, J., Lee, Y., Kim, J., Kosiorok, A., Choi, S., & Teh, Y. W. (2019, May). Set transformer: A framework for attention-based permutation-invariant neural networks. In *International Conference on Machine Learning* (pp. 3744–3753). PMLR.
- [4] Shalev-Shwartz, S. (2012). Online learning and online convex optimization. *Foundations and Trends® in Machine Learning*, 4(2), 107–194.