Faculté des Sciences et Techniques Université de Limoges
M2 – Semester 1 – 2024/2025

OPTIMAL CONTROL AND LARGE-SCALE OPTIMIZATION - LAURENT PFEIFFER

Miniconference Report
CHAU Dang Minh

Summary. Dr. Laurent Pfeiffer, a current researcher at Inria working on optimal control, gave a review and extension on reinforcement learning and optimal control. Then he introduced an application in energy where continuous relaxation is used to approximate a discrete solution.

1 Frank-Wolfe Algorithm

Consider the problem

$$\inf_{x \in K} f(x) \tag{$\mathcal{P}$}$$

where f is convex, continuously differentiable, with Lipschitz-continuous gradient and $K \subset \mathbb{R}^n$ is convex and compact. We have the Frank-Wolfe algorithm as describe below. A general convergence result of the algorithm for difference step size selections in Banach space is given in [1].

Algorithm 1 Frank-Wolfe Algorithm

Input: $x_0 \in K$
for $k = 0, 1 \dots$ **do**
$y_k = \underset{y \in \mathbb{R}^n}{\operatorname{argmin}} \langle \nabla f(x_k), y \rangle$
$w_k = \dfrac{2}{k+2}$
$x_{k+1} = w_k y_k + (1 - w_k) x_k$
end for
Output: x_k

Furthermore, when $K = K_1 \times \ldots \times K_N$, we can write

$$\langle \nabla f(x), y \rangle = \sum_{i=1}^{N} \langle \nabla_{y_i} f(x), y_i \rangle \tag{1}$$

and minimize each element independently.

2 Multi-agent Learning Model

Reinforcement learning and optimal control are capable of modeling exactly the same class of problems. In terms of optimal control, we consider the following conventions

- A horizon T and the time interval $\{0, 1, 2, \dots, T\}$.
- The set of states $\mathcal{X}$.
- The set of controls $\mathcal{U}$.
- The set of random outcomes Ω.

Let $f : \mathcal{X} \times \mathcal{U} \times \Omega \to \mathcal{X}$ and X_0 be given. We define the development of the system as

$$X_{t+1} = f(X_t, U_t, \psi_t), \quad t = 0, \dots, T-1, \tag{2}$$

where $\psi_t, t = 0, \dots, T-1$ are random turbulence. We assume that for each t, U_t is independent of $\psi_{t+1}, \dots, \psi_t$. We aim to minimize

$$\min_{U_0,\dots,U_{T-1}} \mathbb{E}\left[\sum_{t=1}^{T-1} \ell(X_t, U_t, \psi_t) + \phi(X_T)\right]. \tag{3}$$

This problem can be solve using dynamic programming, iterative methods, Monte-Carlo methods or a combination of them, which utilizes the cost function (or value function corresponding to a maximization problem) given as a Bellman equation

$$J_t(x) = \mathbb{E}\left[\ell(x, U_t, \psi_t) + J_{t+1}(X_{t+1})\,|X_t = x\right] \tag{4}$$

For a multi-agent problem, each cost function contributes to the total cost. Besides that, there is a social cost. Let $\mathcal{X} = \mathcal{X}_1 \times \ldots \times \mathcal{X}_N$, we define the N-agent problem as

$$\inf_{x\in\mathcal{X}} J(x) = f\left(\frac{1}{N}\sum_{i=1}^{N} J_i(x_i)\right) + \frac{1}{N}\sum_{i=1}^{N} H_i(x_i) \tag{$\mathcal{P}_N$}$$

When the state space $\mathcal{X}$ is discrete, it is difficult to optimize the function iteratively or combinatorially. Fortunately, if we replace each x_i by a probability distribution $\mu_i \in \Delta(\mathcal{X}_i)$, we arrive at a continuous problem whose solution can be used as approximation to the original problem. In particular, we have

$$\mathbb{E}_{\mu_i}[J_i] = \int_{\mathcal{X}_i} J_i(x_i)\,\mathrm{d}\mu_i(x_i), \quad \mathbb{E}_{\mu_i}[H_i] = \int_{\mathcal{X}_i} H_i(x_i)\,\mathrm{d}\mu_i(x_i). \tag{5}$$

Then the so-call relaxed cost function becomes

$$\inf_{x\in\mathcal{X}} \tilde{J}(x) = f\left(\frac{1}{N}\sum_{i=1}^{N} \mathbb{E}_{\mu_i}[J_i]\right) + \frac{1}{N}\sum_{i=1}^{N} \mathbb{E}_{\mu_i}[H_i] \tag{$\tilde{\mathcal{P}}_N$}$$

The function $\tilde{J}$ is convex. Now we can apply Frank-Wolfe algorithm to solve ($\tilde{\mathcal{P}}_N$). The following theorem gives the gap between two problems.

Theorem 1. *The exists $C > 0$ independent of N such that*

$$\mathrm{val}(\tilde{\mathcal{P}}_N) < \mathrm{val}(\mathcal{P}_N) < \mathrm{val}(\tilde{\mathcal{P}}_N) + \frac{C}{N}. \tag{6}$$

We remark that, for larger N, the gap is smaller. With a solution to $\tilde{\mathcal{P}}_N$, we can sample several solutions to $\mathcal{P}_N$ and take the best one.

3 Conclusion

Dynamic system control problems lie in the intersection of reinforcement learning and numerical optimization, and sometimes game theory. We had an opportunity to work with a lifelike problem, hence become more open to real-life problems.

References

[1] Hong-Kun Xu. "Convergence analysis of the Frank-Wolfe algorithm and its generalization in Banach spaces". In: *arXiv preprint arXiv:1710.07367* (2017).