0.1 Описание метода

Есть два базовых метода для решения задач типа (??)-(??):

1. Метод дифференциального динамического программирования (DDP) [8], [9];
2. Метод итеративного линейно-квадратичного регулятора (iLQR) [6].

Оба метода итеративны и требуют на каждой итерации некоторое *референсное* управление $\bar{u}$ и соответствующую ему референсную траекторию $\bar{x}$. Далее в работе под референсной траекторией понимается пара $(\bar{u}, \bar{x})$. Вдоль данной траектории задача Коши и функционал качества полиномиально аппроксимируются. После чего к аппроксимированной системе применяется соответствующий метод. Результатом итерации является поправка на референсное управление δu.

Метод DDP строит поправку как градиент гамильтониана аппроксимированной задачи

$$\delta u^k = \alpha \nabla_u H(\bar{u}^k),$$

метод iLQR — как линейно-квадратичный регулятор.

Считается, что метод iLQR более надежный, так как не подвержен проблемам, присущим градиентным методам, таким как остановка в локальном минимуме, но сходится за большее число итераций, чем метод DDP. Однако при проведении сравнения скорости сходимости на конкретных примерах выясняется, что нельзя заранее предсказать, какой метод покажет себя лучше [7].

В данной работе применяется метод iLQR. Его основная идея:

1. На каждой итерации имеем референсную траекторию $(\bar{u}, \bar{x})$;
2. Вдоль референсной траектории линеаризуем задачу Коши и аппроксимируем функционал качества до второго порядка;
3. Строим поправку на управление δu как линейно-квадратичный регулятор аппроксимированной задачи;
4. Если не выполнено терминальное условие $|J(\bar{u}) - J(\bar{u} + \delta u)| < \varepsilon$ используем поправленное управление $\bar{u} + \delta u$ в качестве референсного на следующей итерации алгоритма.

Список литературы

[1] Ernesto G. Birgin and J. M. Martínez. *Practical augmented Lagrangian methods.* Springer US, Boston, MA, 2009.

[2] M. Jordan E. Todorov. Optimal feedback control as a theory of motor coordination. *Nature Neuroscience*, 5(11):1226–1235, 2002.

[3] C. M. Harris and D. M. Wolpert. Signal-dependent noise determines motor planning. *Nature*, 394, August 1998.

[4] N. Hogan. An organizing principle for a class of voluntary movement. *Journal of Neuroscience*, 4(11):2745–2754, 1984.

[5] Donald E. Knuth. *The TEX Book.* Addison-Wesley Professional, 1986.

[6] Weiwei Li and Emanuel Todorov. Iterative linear quadratic regulator design for nonlinear biological movement systems. In *ICINCO (1)*, pages 222–229. Citeseer, 2004.

[7] Z. Manchester and S. Kuindersma. Derivative-free trajectory optimization with unscented dynamic programming. In *2016 IEEE 55th Conference on Decision and Control (CDC)*, pages 3642–3647, 2016.

[8] D. Mayne. A second-order gradient method for determining optimal trajectories of non-linear discrete-time systems. *International Journal of Control*, 3(1):85–95, 1966.

[9] D. Murray and S. Yakowitz. Differential dynamic programming and newton's method for discrete optimal control problems. *Journal of Optimization Theory and Applications*, 43(3):395–414, 1984.

[10] M. Kawato Y. Uno and R. Suzuki. Formation and control of optimal trajectory in human multijoint arm movement - minimum torque-change model. *Biological Cybernetics*, 61(2):89–101, 1989.

[11] С. А. Колюбин. *Динамика робототехнических систем.* Редакционно-издательский отдел Университета ИТМО, Санкт-Петербург, 2017.