

## 0.1 Описание метода

Существует два базовых метода для решения задач нелинейных задач оптимального управления, к классу которых относится задача (??)-(??):

1. Метод дифференциального динамического программирования (DDP) [?], [?];
2. Метод итеративного линейно-квадратичного регулятора (iLQR<sup>1</sup>) [?].

Методы идейно схожи:

1. Методы итеративны и используют полную информацию о состоянии системы<sup>2</sup>;
2. На каждой итерации методы используют некоторое *референсное* управление  $\bar{u}$  и соответствующую ему референсную траекторию  $\bar{x}$ ;
3. Вдоль референсной траектории задача полиномиально аппроксимируется;
4. На основании аппроксимированной системы строится некоторая поправка на исходное референсное управление.

**Определение 1.** Под *референсным управлением*  $\bar{u}$  мы будем понимать управление, которое подается на вход каждой итерации соответствующего алгоритма. Под *референсной траекторией* — соответствующую референсному управлению траекторию системы  $\bar{x}$ , либо иногда пару  $(\bar{u}, \bar{x})$ .

Отличие методов заключается в способе получения оптимальной поправки: метод DDP строит поправку как градиент гамильтониана аппроксимированной задачи

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

метод iLQR — как её линейно-квадратичный регулятор.

Считается, что метод iLQR более надежный, так как в меньшей степени подвержен проблемам, присущим градиентным методам, таким как остановка в локальном минимуме, но сходится за большее число итераций, чем метод

---

<sup>1</sup>В некоторых источниках, например [?], используется аббревиатура SLQ.

<sup>2</sup>Наличие уравнения наблюдения предполагало бы, что каждая итерация алгоритма проводится не на компьютере, а не реальном объекте. Применительно к нашей модели это означает, что человек достигнет цели движения только с некоторой попытки.

DDP. Однако при проведении численного эксперимента для сравнения скорости сходимости на конкретных задачах авторы приходят к противоположным результатам. Лучше всего резюмирует это положение вещей работа [?], в которой проведено сравнение двух методов для трёх классических задач механики, и в каждой задаче методы показывают разную асимптотику сходимости.

В данной работе для построения управления был выбран метод iLQR. Выпишем его основные шаги:

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  $\delta u$  как линейно-квадратичный регулятор аппроксимированной задачи;
4. Если не выполнено терминальное условие

$$|J(\bar{u}) - J(\bar{u} + \delta u)| < \varepsilon, \quad (0.1)$$

то используем поправленное управление  $\bar{u} + \delta u$  в качестве референсного на следующей итерации алгоритма.