0.1 Описание метода

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

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

Методы идейно схожи:

1. Методы итеративны и используют полную информацию о состоянии системы[2];
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 траекторей* — 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 антиградиент гамильтониана аппроксимированной задачи

$$\delta u^k = -\alpha \nabla_u H(\bar{u}^k), \ \ 0 < \alpha \leqslant 1,$$

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

[1]В некоторых источниках, например [?], используется аббревиатура SLQ.

[2]Наличие уравнения наблюдения предполагало бы, что каждая итерация алгоритма проводится не на компьютере, а на реальном объекте. Применительно к нашей модели это означало бы, 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 δu как линейно-квадратичный регулятор аппроксимированной задачи;
4. Если не выполнено терминальное условие
$$|J(\bar{u}) - J(\bar{u} + \delta u)| < \varepsilon, \tag{0.1}$$
то используем поправленное управление $\bar{u} + \delta u$ в качестве референсного на следующей итерации алгоритма.

Для возможности применения построенного метода важно понимать порядок числа итераций, требующихся для построения оптимального управления. Поэтому далее при построении численных решений конкретных постановок задач будет дополнительно указано, за какое число итераций *сошёлся* приведённый метод.

Определение 2. Будем говорить, что метод (алгоритм) *сошёлся на i-ой итерации*, если построенное на i-ой итерации управление $\bar{u} + \delta u$ удовлетворяет терминальному условию (0.1).