Московский государственный университет имени М. В. Ломоносова
Факультет вычислительной математики и кибернетики
Кафедра системного анализа

Егоров Кирилл Юлианович

Математическое моделирование движений руки, держащей предмет

выпускная квалификационная работа

Научный руководитель
к.ф.-м.н., доцент И. В. Востриков

Москва, 2023

Содержание

1 Математическое моделирование

1.1 Планарная модель руки, держащей предмет

Рассмотрим руку человека, держащего стержень. В некотором приближении можно считать, что мы имеем трехсекционный математический маятник. Для каждого из 3-х сочленений нам известны:

1. Масса сочленения m_i, $i = 1, 2, 3$;
2. Линейная плотность сочленения $\rho_i = \rho_i(x)$, $0 \leqslant x \leqslant l_i$, $i = 1, 2, 3$;
3. Длина сочленения l_i, $i = 1, 2, 3$;
4. Угол поворота сочленения θ_i, $i = 1, 2, 3$ относительно оси абсцисс Oe_1.

Также считаем, что положение плечевого сустава фиксировано для определенности в точке $O = (0, 0)$. На Рис. 1 приведена схема с примером данного маятника и соответствующая позиция человека.

Рис. 1: Иллюстрация предложенной модели. Рисунок слева сгенерирован нейросетью *Lexica Aperture* по текстовому запросу и приведен для визуального соответствия сочленений маятника на схеме с частями тела человека.

1.2 Уравнение динамики

Для получения уравнения динамики рассматриваемой физической системы воспользуемся методом Эйлера–Лагранжа [11]. Идея метода состоит в проведении следующих последовательных шагов:

1. Выбор обобщенных координат;
2. Получение выражения для кинетической K и потенциальной Π энергий системы, записанных в обобщенных координатах;
3. Получение выражения для лагражиана системы $\mathcal{L}$;
4. Составление системы уравнений движения, соответствующих каждой обобщенной координате.

Обобщенными координатами для нашей системы выберем углы поворота сочленений θ_i, $i = 1, 2, 3$. Далее перейдем к выражению энергий через обобщенные координаты.

Для подсчета кинетической энергии воспользуемся теоремой Кёнинга [ссылка куда-то].

Теорема 1 (Кёнинг). *Кинетическая энергия тела есть энергия поступательного движения центра масс плюс энергия вращательного движения относительно центра масс*

$$K = \frac{1}{2}m\|v_c\|^2 + \frac{1}{2}\omega^{\mathrm{T}} I\omega, \tag{1.1}$$

где m — полная масса тела, I — тензор инерции тела, v_c — линейная скорость центра масс, ω — скорость вращения тела относительно центра масс.

Далее в работе мы будем полагать, что каждое из сочленений представляет собой однородный стержень длины l_i массы m_i. В таком случае получаем следующие значения для положения центра масс $c^i \in \mathbb{R}^2$ i-ого сочленения:

$$c^i = \begin{bmatrix} \sum\limits_{j=1}^{i-1} l_j cos\theta_j + \frac{l_i}{2} cos\theta_i \\ \sum\limits_{j=1}^{i-1} l_j sin\theta_j + \frac{l_i}{2} sin\theta_i \end{bmatrix}.$$

Выражения для момента инерции и скорости вращательного движения относительно центра масс для стержня получаются соответственно:

$$I_i = \int\limits_{(m_i)} r^2\, dm = \rho_i \int\limits_{(l_i)} r^2\, dl = \frac{m_i l_i^2}{12},$$

$$\omega_i = 2\dot{\theta}_i.$$

Потенциальная энергия i-ого сочленения рассчитывается по формуле

$$\Pi_i = m_i g c_2^i,$$

где $g \approx 9{,}8$ — ускорение свободного падения.

Общая кинетическая и потенциальная энергии системы рассчитываются как сумма энергий каждого из сочленений:

$$K = \sum_{i=1}^{3} K_i = \sum_{i=1}^{3} \left(\frac{m_i \|\dot{c}^i\|^2}{2} + \frac{m_i l_i^2 |\dot{\theta}_i|^2}{6} \right),$$

$$\Pi = \sum_{i=1}^{3} \Pi_i = \sum_{i=1}^{3} m_i g c_2^i.$$

Теперь введём лагражиан системы

$$\mathcal{L} = K - \Pi$$

и построим систему уравнений Эйлера–Лагранжа [ссылка]:

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{\theta}_i} - \frac{\partial \mathcal{L}}{\partial \theta_i} = \tau_i,\ i = \overline{1,3}, \tag{1.2}$$

где τ_i — момент силы, действующий на i-ое сочленение, который доступен для управления.

Продифференцировав члены из левой части уравнения (1.2), получим уравнение динамики для рассматриваемой системы:

$$M(\theta)\ddot{\theta} + L(\theta, \dot{\theta}) = \tau, \tag{1.3}$$

где $M(\theta) = M^{\mathrm{T}}(\theta) \in \mathbb{R}^{3\times 3}$ — матрица инерции системы, $L(\theta, \dot{\theta}) \in \mathbb{R}^3$ — вектор центростремительных и кореолисовых сил. На Рис. 2 приведено численное моделирование свободного падения в соответствии с полученным уравнением динамики (1.3).

Замечание 1. Матрица инерции $M(\theta)$ является положительно-определённой, поскольку кинетическая энергия системы K всегда неотрицательна:

$$K(\theta, \dot{\theta}) = \frac{1}{2} \left\langle \dot{\theta}, M(\theta)\dot{\theta} \right\rangle > 0 \text{ для любого } \dot{\theta} \neq 0.$$

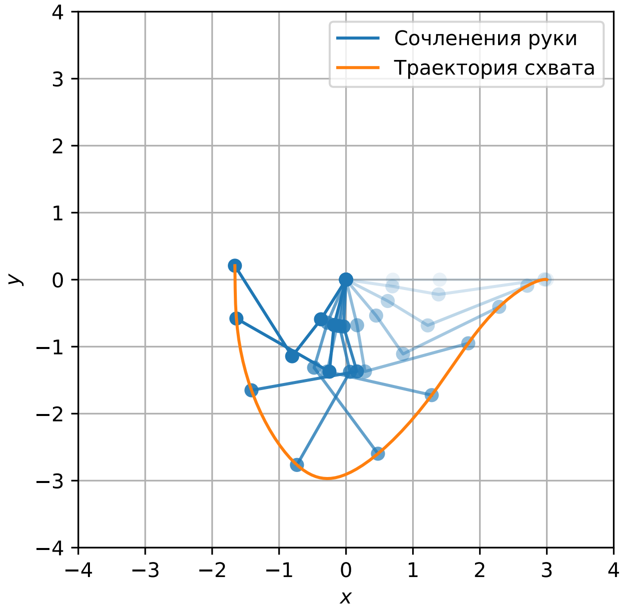

Рис. 2: Траектория руки $l = [0{,}7\ 0{,}7\ 1{,}6]^{\mathrm{T}}$, $m = [0{,}8\ 0{,}8\ 1{,}2]^{\mathrm{T}}$ в свободном падении из начального положения $\theta^{\text{start}} = [0\ 0\ 0]^{\mathrm{T}}$, $\dot{\theta}^{\text{start}} = [0\ 0\ 0]^{\mathrm{T}}$ на временном интервале $0 \leqslant t \leqslant 1$. Положения, соответствующие более раннему времени, показаны бледнее. Всюду далее работе для численного моделирования используется данная конфигурация маятника l, m.

1.3 Учёт энергетических затрат

Для моделирования биологического движения необходимо выяснить, какими принципами руководствуется мозг при выборе траектории для некоторого целевого движения. Существует бесконечное число возможных путей и профилей скорости для перемещения руки из одной точки в другую, и каждая траектория может достигнута несколькими возможными комбинациями углов между сочленениями. При этом нервная и моторно-двигательные системы человека для выбора одной конкретной траектории анализируют большой объем информации, поступающий от всех органов чувств.

В силу того, что нервная система человека есть результат оптимизационных процессов: эволюции, адаптации к условиям среды, обучения, мы постулируем следующий биологический принцип оптимальности.

Утверждение 1 (Биологический принцип оптимальности)**.** *Выбираемые нервной системой схемы движения являются оптимальными для поставленной задачи.*

Применение данного принципа позволяет не только моделировать движения методами оптимального управления, но и анализировать их причины.

В работе [2] было показано, что оптимизации проводятся с целью уменьшения затрат энергии. Однако общего подхода к формализации энергетических затрат пока не выработано. Так, например, в работе [4] предлагается минимизировать *рывок* схвата, то есть

$$\int\limits_{t_{\text{start}}}^{t_{\text{final}}} \left\| \frac{d^3 e^3}{dt^3} \right\|^2 dt \longrightarrow \min,$$

а в работе [10] — изменение крутящего момента

$$\int\limits_{t_{\text{start}}}^{t_{\text{final}}} \left\| \frac{d\tau}{dt} \right\|^2 dt \longrightarrow \min. \tag{1.4}$$

Причём существуют и другие менее популярные варианты, например, [3].

Мы будем использовать для формализации энергетических затрат выражение (1.4), поскольку данный критерий напрямую зависит от динамики руки и лучше согласуется с эмпирическими данными, чем модель рывка.

2 Постановка задачи

2.1 Непрерывная постановка задачи

Поставим задачу целевого управления для модели, построенной в Разделе 1. Для этого рассмотрим расширенное фазовое пространство с состоянием

$$x = [\theta \; \dot{\theta} \; \tau]^{\mathrm{T}} \in \mathbb{R}^9.$$

Тогда уравнение динамики системы (1.3) можно переписать в виде системы однородных дифференциальных уравнений

$$\dot{x} = A(x) + Bu, \tag{2.1}$$

где

$$A(x) = \begin{bmatrix} x_2 \\ M^{-1}(x_1)(x_3 - L(x_1, x_2)) \\ O \end{bmatrix}, \quad B = \begin{bmatrix} O \\ O \\ I \end{bmatrix}.$$

Считаем, что для данной системы поставлена задача Коши, то есть нам известно начальное состояние системы

$$x(t_0) = x^{\mathrm{start}}. \tag{2.2}$$

Замечание 2. Отметим, что для выполнения достаточных условий существования и единственности решения Каратеодори для задачи Коши (2.1)-(2.2) управление u достаточно брать из класса измеримых на рассматриваемом отрезке $t_{\mathrm{start}} \leqslant t \leqslant t_{\mathrm{final}}$ функций.

Для задачи Коши (2.1)-(2.2) поставим задачу поиска управления $u \in U[t_{\mathrm{start}}, t_{\mathrm{final}}]$, минимизирующего функционал вида:

$$J = q^{\mathrm{final}}(x(t_{\mathrm{final}})) + w_1 \int\limits_{t_{\mathrm{start}}}^{t_{\mathrm{final}}} q(x(t))\, dt + w_2 \int\limits_{t_{\mathrm{start}}}^{t_{\mathrm{final}}} r(u(t))\, dt, \tag{2.3}$$

где q^{final}, q отвечают за терминальное и фазовые ограничения соответственно и выбираются в зависимости от конкретной постановки задачи, а r отвечает за энергетические затраты и в соответствии с (1.4) равна:

$$r(u) = \|u\|^2,$$

а w_1, w_2 — веса соответствующих критериев для данной многокритериальной задачи.

Для дальнейших рассуждений потребуем, чтобы функции q^{final}, q были дважды непрерывно дифференцируемыми. Полученные из модели функции A и r заведомо удовлетворяют этому требованию.

Замечание 3. Учёт фазовых ограничений в интегральной части функционала качества J, представленный в работе, позволяет лишь приближенно уписать условия вида

$$g_i(x) \leqslant 0,$$

которые часто встречаются в задачах, например, для обхода препятствия. Для этого функция q выбирается таким образом, чтобы штрафовать за приближение траектории к препятствию. Для строго формального решения задачи с подобным условием, необходимо пользоваться методами расширенного лангранжиана [1], которые предполагают решение серии задач типа (2.1)-(2.2)-(2.3). Это приводит к существенному ухудшению асимптотики алгоритмов и тем самым увеличению времени работы программного решения.

2.2 Дискретизация задачи

Для удобства дальнейших рассуждений дискретизируем задачу (2.1)-(2.2)-(2.3) по времени $t_{\text{start}} \leqslant t \leqslant t_{\text{final}}$. Для этого введем равномерную сетку с шагом Δt:

$$\{t_i\}_{i=1}^{N+1}, \quad t_1 = t_{\text{start}}, \quad t_{N+1} = t_{\text{final}}, \quad t_{i+1} - t_i = \Delta t.$$

Тогда, сузив класс допустимых управлений до кусочно-постоянных, получаем дискретный вариант рассматриваемой задачи Коши (2.1)-(2.2):

$$\begin{cases} x^{k+1} = f(x^k, u^k),\ k = \overline{1, N}, \\ x^1 = x^{\text{start}}, \end{cases} \tag{2.4}$$

где

$$f(x^k, u^k) = \Delta t \left(A(x^k) + Bu^k \right) + x^k.$$

При этом функционал (2.3) для дискретной задачи приобретет вид

$$J = q^{N+1}(x^{N+1}) + \sum_{k=1}^{N} q^k(x^k) + \sum_{k=1}^{N} r^k(u^k), \tag{2.5}$$

где

$$q^{N+1} = q^{\text{final}}, \qquad q^k = w_1 q \Delta t, \qquad r^k = w_2 r \Delta t \qquad k = \overline{1, N}.$$

3 Синтез оптимального управления

3.1 Описание метода

Есть два базовых метода для решения задач типа (2.4)-(2.5):

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.

3.2 Синтез оптимальной поправки

Допустим мы имеем некоторое референсное управление $\bar{u} = \{\bar{u}^k\}_{k=1}^{N}$ и соответствующую ему референсную траекторию $\bar{x} = \{\bar{x}^k\}_{k=1}^{N+1}$. Введем обозначения:

$$f_x^k = \left.\frac{\partial f}{\partial x}\right|_{(\bar{x}^k, \bar{u}^k)}, \quad f_u^k = \left.\frac{\partial f}{\partial u}\right|_{(\bar{x}^k, \bar{u}^k)},$$

$$q^k = q(\bar{x}^k), \quad q_x^k = \left.\frac{\partial q}{\partial x}\right|_{\bar{x}^k}, \quad q_{xx}^k = \left.\frac{\partial^2 q}{\partial x^2}\right|_{\bar{x}^k},$$

$$r^k = r(\bar{u}^k), \quad r_x^k = \left.\frac{\partial r}{\partial u}\right|_{\bar{u}^k}, \quad r_{xx}^k = \left.\frac{\partial^2 r}{\partial u^2}\right|_{\bar{u}^k}.$$

Тогда, линеаризуя вдоль референсной траектории задачу Коши (2.4) и строя квадратичную аппроксимацию вдоль той же траектории функционала качества (2.5), получаем следующую задачу:

$$\begin{cases} \delta x^{k+1} = f_x^k \delta x + f_u^k \delta u, \ k = \overline{1, N}, \\ \delta x^1 = 0. \end{cases} \tag{3.1}$$

$$J = q^{N+1} + q_x^{N+1} \tilde{x}^{N+1} + \frac{1}{2} \langle \tilde{x}^{N+1}, q_{xx}^{N+1} \tilde{x}^{N+1} \rangle + \\ + \sum_{k=1}^{N} \left[q^k + q_x^k \tilde{x}^k + \frac{1}{2} \langle \tilde{x}^k, q_{xx}^k \tilde{x}^k \rangle \right] + \\ + \sum_{k=1}^{N} \left[r^k + r_u^k \tilde{u}^k + \frac{1}{2} \langle \tilde{u}^k, r_{uu}^k \tilde{u}^k \rangle \right], \tag{3.2}$$

где

$$\tilde{x}^k = \bar{x}^k + \delta x^k, \qquad \tilde{u}^k = \bar{u}^k + \delta u^k.$$

Построим гамильтониан для задачи (3.1)-(3.2):

$$H_k = q^k + q_x^k \tilde{x}^k + \frac{1}{2} \langle \tilde{x}^k, q_{xx}^k \tilde{x}^k \rangle + \\ + r^k + r_u^k \tilde{u}^k + \frac{1}{2} \langle \tilde{u}^k, r_{uu}^k \tilde{u}^k \rangle + \\ + (\lambda^{k+1})^{\mathrm{T}} (f_x^k \delta x^k + f_u^k \delta u^k), \tag{3.3}$$

где λ^{k+1} — мультипликаторы Лагранжа.

Оптимальное управление δu должно удовлетворять необходимому условию $\frac{\partial H_k}{\partial u^k} = 0$:

$$r_u^k + r_{uu}^k(\bar{u}^k + \delta u^k) + (f_u^k)^{\mathrm{T}}\lambda^{k+1} = 0.$$

что дает следующее выражение для поправки:

$$\delta u^k = -(r_{uu}^k)^{-1}[(f_u^k)^{\mathrm{T}}\lambda^{k+1} + r_u^k] - \bar{u}^k. \tag{3.4}$$

При этом имеет силу сопряженная задача:

$$\begin{cases} \lambda^k = (f_x^k)^{\mathrm{T}}\lambda^{k+1} + q_x^k + q_{xx}^k(\bar{x}^k + \delta x^k) \\ \lambda^{N+1} = q_x^{N+1} + q_{xx}^{N+1}(\bar{x}^{N+1} + \delta x^{N+1}). \end{cases} \tag{3.5}$$

Из (3.4) и (3.5) вытекает

$$\begin{pmatrix} \delta x^{k+1} \\ \lambda^k \end{pmatrix} = \underbrace{\begin{pmatrix} f_x^k & -f_u^k(r_{uu}^k)^{-1}(f_u^k)^{\mathrm{T}} \\ q_{xx}^k & (f_x^k)^{\mathrm{T}} \end{pmatrix}}_{\Phi^k} \begin{pmatrix} \delta x^k \\ \lambda^{k+1} \end{pmatrix} + \underbrace{\begin{pmatrix} -f_u^k(r_{uu}^k)^{-1}r_u^k \\ q_x^k \end{pmatrix}}_{\Gamma^k}. \tag{3.6}$$

Теорема 2. *Оптимальная поправка δu для задачи* (3.1)-(3.2) *вычисляется как*

$$\delta u^k = -(r_{uu}^k + (f_u^k)^{\mathrm{T}}S_{k+1}f_u^k)^{-1}((f_u^k)^{\mathrm{T}}S_{k+1}f_u^k\delta x + v^{k+1} + r_u^k), \tag{3.7}$$

где S_k и v^k высчитываются в обратном времени как

$$\begin{aligned} S_k &= \Phi_{21}^k + \Phi_{22}^k S_{k+1}(I - \Phi_{12}^k S_{k+1})^{-1}\Phi_{11}^k, \\ v^k &= \Phi_{22}^k S_{k+1}(I - \Phi_{12}^k S_{k+1})^{-1}(\Phi_{12}^k v^{k+1} + \Gamma_1^k) + \Phi_{22}^k v^{k+1} + \Gamma_2^k \end{aligned} \tag{3.8}$$

с граничными условиями

$$\begin{aligned} S_{N+1} &= q_{xx}^{N+1}, \\ v^{N+1} &= q_x^{N+1} + q_{xx}^{N+1}\bar{x}^{N+1}. \end{aligned} \tag{3.9}$$

Доказательство. Предположим, что мультипликаторы λ имеют следующую аффинную форму относительно фазовой переменной δx

$$\lambda^k = S_k\delta x^k + v^k \tag{3.10}$$

Тогда из граничного условия (3.5) вытекает граничное условие на S_k, v^k (3.9):

$$\lambda^{N+1} = q_x^{N+1} + q_{xx}^{N+1}\left(\bar{x}^{N+1} + \delta x^{N+1}\right)$$

$$\Downarrow$$

$$S_{N+1} = q_{xx}^{N+1}, v^{N+1} = q_x^{N+1} + q_{xx}^{N+1}\bar{x}^{N+1}.$$

Теперь подставим (3.10) в выражение (3.6) для δx^{k+1}:

$$\delta x^{k+1} = \Phi_{11}^k \delta x^k + \Phi_{12}^k (S_{k+1} \delta x^{k+1} + v^{k+1}) + \Gamma_1^k.$$

Получаем

$$\delta x^{k+1} = \left(\underbrace{I - \Phi_{12}^k S_{k+1}}_{K_k} \right)^{-1} \left(\Phi_{11}^k \delta x^k + \Phi_{12}^k v^{k+1} + \Gamma_1^k \right).$$

Подставим получившееся выражение в (3.6) для λ^k:

$$\begin{aligned} \lambda^k = S_k \delta x^k + v^k = \Phi_{21}^k \delta x^k + \Phi_{22}^k \left(S_{k+1} \delta x^{k+1} + v^{k+1} \right) + \Gamma_2^k = \\ = \Phi_{21}^k \delta x^k + \Phi_{22}^k \left(S_{k+1} K_k^{-1} (\Phi_{11}^k \delta x^k + \Phi_{12}^k v^{k+1} + \Gamma_1^k) + v^{k+1} \right) + \Gamma_2^k. \end{aligned}$$

Таким образом получаем искомые соотношения (3.8):

$$\begin{aligned} S_k &= \Phi_{21}^k + \Phi_{22}^k S_{k+1} K_k^{-1} \Phi_{11}^k, \\ v^k &= \Phi_{22}^k (S_{k+1} K_k^{-1} (\Phi_{12}^k v^{k+1} + \Gamma_1^k) + v^{k+1}) + \Gamma_2^k. \end{aligned}$$

Итоговая формула для оптимальной поправки (3.7) получается прямой подстановкой получившихся соотношений в выражение (3.4).

■

Замечание 4. Приведенная теорема требует существование обратных матриц для $K_k, k = \overline{1, N}$. При этом для нелинейных систем данное условие может не выполняться. Чтобы метод продолжал работать, предлагается в случае нулевого определителя $\det K_k = 0$, заменять в формулах (3.7), (3.8) матрицу K_k на регуляризованную

$$\mathcal{K}_k = K_k + \mu I. \tag{3.11}$$

В данной работе при численном построении управления использовалась следующая константа регуляризации:

$$\mu = 10^{-8}.$$

Замечание 5. Данная теорема не учитывает возможные ограничения на управление, которые естественным образом возникают для данной задачи. Далее мы будем считать, что задано некоторое поточечное ограничение на управление $u^k \in \mathcal{U}^k$. В этом случае, мы будем считать домножать поправку на некоторую величину η, такую что $u^k + \eta \delta u^k \in \partial \mathcal{U}^k$.

Добавить замечание, почему только r_{uu} должна быть > 0 — типа лемма об обращении матриц.

Не забыть, что так как метод итеративный, нет надобности рассматривать большую систему с уравнение наблюдаемости, так как это предполагало бы, что каждая итерация алгоритма производится не на компьютере, где с наблюдаемостью все хорошо, а на настоящем человеке.

3.3 Регуляризация оптимальной поправки

Согласно Теореме 1 оптимальная поправка имеет следующую аффинную форму

$$\delta u^{k*} = L_k \delta x^k + d^k,$$

где L_k — коэффициент управления с обратной связью, d^k — коэффициент управления без обратной связи, возникающий по причине того, что мы имеем дело с отклонениями от заданного состояния.

Данная форма не налагает никаких ограничений на поправку δu. На практике это означает, что на начальных итерациях алгоритма, когда референсная траектория далека от оптимальной, поправка зачастую выводит систему за область действия аппроксимации. Визуально это выражается в том, что на каждой итерации алгоритм выдаёт некоторую случайную траекторию и в конечном итоге не сходится к оптимальной траектории. Чтобы избежать такого эффекта, необходимо регуляризовать коэффициент управления без обратной связи d^k:

$$\delta u^{k*}(\eta) = L_k \delta x^k + \eta d^k.$$

Теперь остается ответить на вопрос, как выбрать подходящий коэффициент регулярицации η. Это можно сделать двумя способами:

1. Дополнительно поточечно ограничить область допустимых управлений

   $$u^k \in \mathcal{U}^k.$$

   Тогда можно выбирать коэффициент следующим способом

   $$\eta = \min\{\eta \,|\, \bar{u}^k + \delta u^{k*}(\eta) \in \mathcal{U}^k\}.$$

   Такой способ предполагает один дополнительный проход алгоритма для поиска минимума, однако существенно замедляет скорость сходимости и накладывает ограничения, не предусмотренные исходной задачей.

2. Использовать ожидаемое отклонение от функции цены

$$\xi_1 \leqslant \frac{J(\bar{u}) - J(\bar{u} + \delta u)}{J_\delta(0) - J_\delta(\delta u^*(\eta))} \leqslant \xi_2.$$

Данные способ не накладывает никаких дополнительных ограничений. При этом для нахождения правильного коэффициента может потребоваться несколько итераций. Изначально выбирается $\eta = 1$, считается оптимальная поправка, если условие не выполнено, но коэффициент $\eta = \gamma\eta$.

3.4 Алгоритм

С учетом вышесказанного можно построить следующий алгоритм. Алгоритмы 1, 2 демонстрируют обратный и прямой проход для получения оптимальной поправки. Алгоритм 3 демонстрирует общий алгоритм построения управления.

Алгоритм 1: Обратный проход

```
function BackwardPass(ū, x̄)
begin
    S_{N+1}, v^{N+1} ← (3.9)
    for k ← N to 1 do
        S_k, v^k ← (3.8)
    end

    return S, v
end
```

Продемонстрируем работу алгоритма для классической задачи перехода в целевое состояние x^{final} без фазовых ограничений. Для этого будем использовать следующие компоненты функции цены:

$$q^{\text{final}}(x) = \|x - x^{\text{final}}\|^2, \qquad q(x) \equiv 0. \tag{3.12}$$

Начальным референсным управлением выберем

$$\bar{u}^k = 0, \ k = \overline{1, N}. \tag{3.13}$$

Алгоритм 2: Прямой проход

function *ForwardPass($\bar{u}$, $\bar{x}$)*

begin

  $\eta, \gamma \leftarrow 1, \frac{1}{2}$

  do

    for $k \leftarrow 1$ **to** N **do**

      $\delta u^k, \delta x^{k+1} \leftarrow$ (3.7), (3.1)

      $u^k \leftarrow \bar{u}^k + \delta u^k$

    end

    $\eta \leftarrow \gamma\eta$

  while $\xi_1 \leqslant \frac{J_{\mathrm{prev}} - J}{J_\delta(0) - J_\delta(\delta u)} \leqslant \xi_2$

  return u

end

Алгоритм 3: Синтез управления

function *Synthesis($\bar{u}$)*

begin

  $J \leftarrow$ (формула)

  do

    $J_{\mathrm{prev}} \leftarrow J$

    $\bar{x} \leftarrow$ (формула)

    $S, v \leftarrow \mathrm{BackwardPass}(\bar{u}, \bar{x})$

    $u \leftarrow \mathrm{ForwardPass}(S, v, J_{\mathrm{prev}})$

    $J \leftarrow$ (формула)

  while $\|J - J_{\mathrm{prev}}\| < \varepsilon$

  return u

end

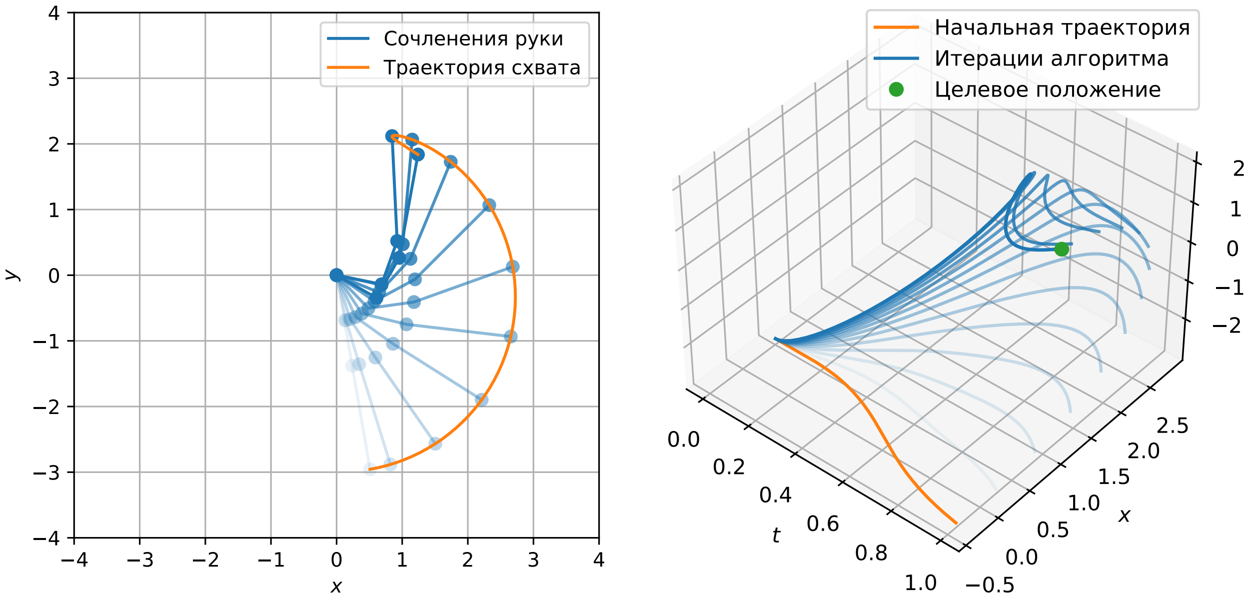

Рис. 3: Решение задачи перехода в целевое состояние (3.12) с начальным референсным управлением (3.13). Слева: поведение системы при полученном управлении. Справа: траектории схвата на каждой итерации алгоритма, более ранние итерации показаны бледнее. Начальное положение $x_1^{\text{start}} = [-1{,}4; -1{,}4; -1{,}4]^{\mathrm{T}}$, $x_2^{\text{start}} = 0$, $x_3^{\text{start}} = 0$. Конечное положение $x_1^{\text{final}} = [-0{,}5; 1{,}1; 1{,}4]^{\mathrm{T}}$, $x_2^{\text{final}} = [-5{,}0; -5{,}0; -5{,}0]^{\mathrm{T}}$, $x_3^{\text{final}} = 0$. Коэффициент значимости энергетического критерия $w_2 = 10^{-2}$. Коэффициент остановки $\varepsilon = 10^{-2}$. Алгоритм сошелся на 14 итерации.

4 Синтез начального референсного управления

4.1 Описание метода

Скорость сходимости метода, приведенного в Разделе 4, зависит от выбора начального референсного управления $\bar{u}$. Считается, что в этом деле можно положиться на мнение эксперта в предметной области, который нарисует траекторию системы $\bar{x}$ и мы потом решим задачу идентификации управления $\bar{u}$, приводящего к такой траектории методом следящего управления.

Для того, чтобы использовать метод необходимо, чтобы он:

1. Строил управление *быстро*. Желательно, чтобы алгоритм имел линейную асимптотику.
2. Получившаяся референсная траектория была близка к оптимальной.
3. Получившаяся референсная траектория была бы возможной для рассматриваемой задачи. Это важно, если в задаче присутствуют строгие ограничения на управление.

В отсутствии эксперта, предлагается использовать следующий метод, удовлетворяющий этим условиям.

1. Необходимо аналитически найти состояние системы, которое минимизирует терминальное условие
$$x^{\text{final}} \in \operatorname{Argmin} q^{\text{final}}(x).$$
2. Привести систему к линейной и поставить для нее задачу минимизации интегрально-квадратичного функционала для перехода в состояние x^{final}.
3. Построить линейно-квадратичный регулятор для полученной задачи. Тем самым мы получим управление, которое минимизирует терминальное условие, но ничего не говорит об энергетическом и фазовом условиях. Тем не менее такой подход будет работать лучше, чем выбор случайного управления.

В случае, если мы можем аналитически найти несколько минимизаторов терминального условия, можно провести перебор с последующим выбором самого подходящего управления.

4.2 Синтез управления

Приведём систему (2.4) к линейному виду заменой управления на

$$\hat{u} = M^{-1}(x_1)[\tau - L(x_1, x_2)]. \tag{4.1}$$

Тогда в фазовом пространстве $\hat{x} = [x_1, x_2]^{\mathrm{T}} \in \mathbb{R}^6$ задача Коши примет вид

$$\begin{cases} \hat{x}^{k+1} = \hat{A}\hat{x}^k + \hat{B}\hat{u}^k \\ \hat{x}^0 = x^{\text{start}}, \end{cases} \tag{4.2}$$

где

$$\hat{A} = \left(\begin{array}{c|c} I & \Delta t I \\ \hline O & I \end{array}\right), \quad \hat{B} = \left(\begin{array}{c} O \\ \hline \Delta t I \end{array}\right).$$

Будем считать, что для исходной задачи задачи (??) известно некоторое состояние x^{final}, минимизирующее терминальное условие q^{final}, то есть

$$x^{\text{final}} \in \operatorname{Argmin} q^{\text{final}}(x).$$

В таком случае для задачи Коши (4.2) поставим задачу минимизации следующего функционала

$$J = \left\|\hat{x}^{N+1} - x^{\text{final}}\right\|^2 + \hat{w}_1 \sum_{k=1}^{N} \left\langle \hat{x}^k, \hat{Q}\hat{x}^k \right\rangle + \hat{w}_2 \sum_{k=1}^{N} \left\|\hat{u}^k\right\|^2 \longrightarrow \min. \tag{4.3}$$

Здесь матрица $\hat{Q} = \hat{Q}^{\mathrm{T}}$ выбирается для исключения возможного *проворачивания* сочленений относительно друг друга следующим образом

$$\hat{Q} = \left(\begin{array}{rrr|c} 2 & -1 & 0 & \\ -1 & 2 & -1 & O \\ 0 & -1 & 1 & \\ \hline & O & & O \end{array}\right). \tag{4.4}$$

Иными словами матрица $\hat{Q}$ является матрицей следующей квадратичной формы

$$\left\langle \hat{x}, \hat{Q}\hat{x} \right\rangle = \theta_1^2 + (\theta_2 - \theta_1)^2 + (\theta_3 - \theta_2)^2.$$

Данное фазовое условие штрафует траекторию в случае большого относительного отклонения между углами сочленений.

После решения задачи (4.2)-(4.3) мы можем восстановить соответствующее управление исходной задачи. Пусть $\hat{u}^*$, $\hat{x}^*$ — оптимальное управление и соответствующая ему оптимальная траектория задачи (4.2)-(4.3). Тогда соответствующее управление для исходной задачи u, можно получить по формуле

$$u^k = \frac{\hat{\tau}^{k+1} - \hat{\tau}^k}{\Delta t}, \tag{4.5}$$

где

$$\begin{cases} \hat{\tau}^k = M(\hat{x}_1^{k*})\hat{u}^{k*} + L(\hat{x}_1^{k*}, \hat{x}_2^{k*}), \text{ при } k = \overline{1, N} \\ \hat{\tau}^{N+1} = 0. \end{cases}$$

Построим гамильтониан задачи (4.2)-(4.3)

$$\hat{H}_k = \left\langle \hat{x}^k, \hat{w}_1 \hat{Q} \hat{x}^k \right\rangle + \left\langle \hat{u}^k, \hat{w}_2 \hat{u}^k \right\rangle + (\hat{\lambda}^{k+1})^{\mathrm{T}} [\hat{A}\hat{x}^k + \hat{B}\hat{u}^k]. \tag{4.6}$$

Оптимальное управление $\hat{u}^*$ должно удовлетворять необходимому условию оптимальности:

$$\left. \frac{\partial \hat{H}_k}{\partial \hat{u}^k} \right|_{\hat{u}^k = \hat{u}^{k*}} = \hat{w}_2 \hat{u}^{k*} + \hat{B}^{\mathrm{T}} \hat{\lambda}^{k+1} = 0,$$

что дает следующее выражение для управления

$$\hat{u}^{k*} = -\frac{1}{\hat{w}_2} \hat{B}^{\mathrm{T}} \hat{\lambda}^{k+1}. \tag{4.7}$$

И уравнение (4.2) можно переписать в следующем виде:

$$\hat{x}^{k+1} = \hat{A}\hat{x}^k - \frac{1}{\hat{w}_2} \hat{B}\hat{B}^{\mathrm{T}} \hat{\lambda}^{k+1}. \tag{4.8}$$

При этом имеет силу следующая сопряженная система:

$$\begin{cases} \hat{\lambda}^k = \hat{w}_1 \hat{Q} \hat{x}^k + \hat{A}^{\mathrm{T}} \hat{\lambda}^{k+1}, \text{ при } k = \overline{1, N} \\ \hat{\lambda}^{N+1} = \hat{x}^{N+1} - x^{\text{final}} \end{cases} \tag{4.9}$$

Теорема 3. *Оптимальное управление $\hat{u}^*$ задачи* (4.2)-(4.3) *задается формулой*

$$\hat{u}^{k*} = \hat{L}_k \hat{x}^k + \hat{d}^k, \tag{4.10}$$

где

$$\hat{L}_k = -\frac{1}{\hat{w}_2} \hat{B}^{\mathrm{T}} S_{k+1} \left(I + \frac{1}{\hat{w}_2} \hat{B}\hat{B}^{\mathrm{T}} S_{k+1} \right)^{-1} \hat{A},$$

$$\hat{d}^k = -\frac{1}{\hat{w}_2} \hat{B}^{\mathrm{T}} \left(I - S_{k+1} \left(I + \frac{1}{\hat{w}_2} \hat{B}\hat{B}^{\mathrm{T}} S_{k+1} \right)^{-1} \hat{B}\hat{B}^{\mathrm{T}} \right) v^{k+1}.$$

Причем переменные S_k, v^k могут быть найдены в обратном времени из соотношений

$$
\begin{aligned}
S_k &= \hat{w}_1\hat{Q} + \hat{A}^{\mathrm{T}} S_{k+1}\left(I + \frac{1}{\hat{w}_2}\hat{B}\hat{B}^{\mathrm{T}} S_{k+1}\right)^{-1}\hat{A},\\
v_k &= \hat{A}^{\mathrm{T}}\left(I - \frac{1}{\hat{w}_2} S_{k+1}\left(I + \frac{1}{\hat{w}_2}\hat{B}\hat{B}^{\mathrm{T}} S_{k+1}\right)^{-1}\hat{B}\hat{B}^{\mathrm{T}}\right) v^{k+1}
\end{aligned}
\tag{4.11}
$$

с граничными условиями

$$
\begin{aligned}
S_{N+1} &= I,\\
v^{N+1} &= -x^{\mathrm{final}}.
\end{aligned}
\tag{4.12}
$$

Д о к а з а т е л ь с т в о. Аналогично доказательству теоремы (1) будем искать решение сопряженной системы в аффинном виде

$$
\hat{\lambda}^k = S_k \hat{x}^k + v^k. \tag{4.13}
$$

Из сопряженной системы (4.9) получаем граничные условия:

$$
\hat{\lambda}^{N+1} = \hat{x}^{N+1} - x^{\mathrm{final}} \quad \Longrightarrow \quad S_{N+1} = I,\ v^{N+1} = -x^{\mathrm{final}}.
$$

Подставив выражение (4.13) в уравнение (4.8), получим

$$
\hat{x}^{k+1} = \hat{A}\hat{x}^k - \frac{1}{\hat{w}_2}\hat{B}\hat{B}^{\mathrm{T}}(S_{k+1}\hat{x}^{k+1} + v^{k+1}),
$$

откуда выражаем

$$
\hat{x}^{k+1} = \left(\underbrace{I + \frac{1}{\hat{w}_2}\hat{B}\hat{B}^{\mathrm{T}} S_{k+1}}_{K_k}\right)^{-1}\left(\hat{A}\hat{x}^k - \frac{1}{\hat{w}_2}\hat{B}\hat{B}^{\mathrm{T}} v^{k+1}\right).
$$

Теперь подставим получившееся выражение в (4.9):

$$
\begin{aligned}
\hat{\lambda}^k = S_k\hat{x}^k + v^k &= \hat{w}_1\hat{Q}\hat{x}^k + \hat{A}^{\mathrm{T}}(S_{k+1}\hat{x}^{k+1} + v^{k+1}) =\\
&= \hat{w}_1\hat{Q}\hat{x}^k + \hat{A}^{\mathrm{T}} S_{k+1} K_k^{-1}\left(\hat{A}\hat{x}^k - \frac{1}{\hat{w}_2}\hat{B}\hat{B}^{\mathrm{T}} v^{k+1}\right) + \hat{A}^{\mathrm{T}} v^{k+1} =\\
&= \left(\hat{w}_1\hat{Q} + \hat{A}^{\mathrm{T}} S_{k+1} K_k^{-1}\hat{A}\right)\hat{x}^k + \hat{A}^{\mathrm{T}}\left(I - \frac{1}{\hat{w}_2} S_{k+1} K_k^{-1}\hat{B}\hat{B}^{\mathrm{T}}\right) v^{k+1}.
\end{aligned}
$$

Откуда получаем искомые соотношения:

$$S_k = \hat{w}_1 \hat{Q} + \hat{A}^{\mathrm{T}} S_{k+1} \left(I + \frac{1}{\hat{w}_2} \hat{B} \hat{B}^{\mathrm{T}} S_{k+1} \right)^{-1} \hat{A},$$
$$v_k = \hat{A}^{\mathrm{T}} \left(I - \frac{1}{\hat{w}_2} S_{k+1} \left(I + \frac{1}{\hat{w}_2} \hat{B} \hat{B}^{\mathrm{T}} S_{k+1} \right)^{-1} \hat{B} \hat{B}^{\mathrm{T}} \right) v^{k+1}.$$

Теперь выражение для оптимального управления (4.10) получается прямой подстановкой получившихся соотношений в выражение (4.7).

■

4.3 Алгоритм

Алгоритм 4 резюмирует метод, предложенный в данном разделе.

На Рис. 4 представлен результат работы алгоритма для задачи (??) с построенным начальным управлением.

Замечание 6. С физической точки зрения решение данной задачи минимизирует угловые ускорения сочленений руки. Таким образом результирующая траектория будет самой плавной из возможных. Кажется естественным, чтобы такая траектория входило в множество допустимых управлений для исходной задачи.

Алгоритм 4: Поиск начальной траектории

```
function InitialControl
begin
    /* Обратный проход                                         */
    S_{N+1}, v^{N+1} ← (4.12)
    for k ← N to 1 do
        S_k, v^k ← (4.11)
    end

    /* Прямой проход                                           */
    x̂^0 ← x^start
    for k ← 1 to N do
        û^k, x̂^{k+1} ← (4.10), (4.2)
    end

    /* Конвертация управления                                  */
    for k ← 1 to N do
        u^k ← (4.5)
    end
    return u
end
```

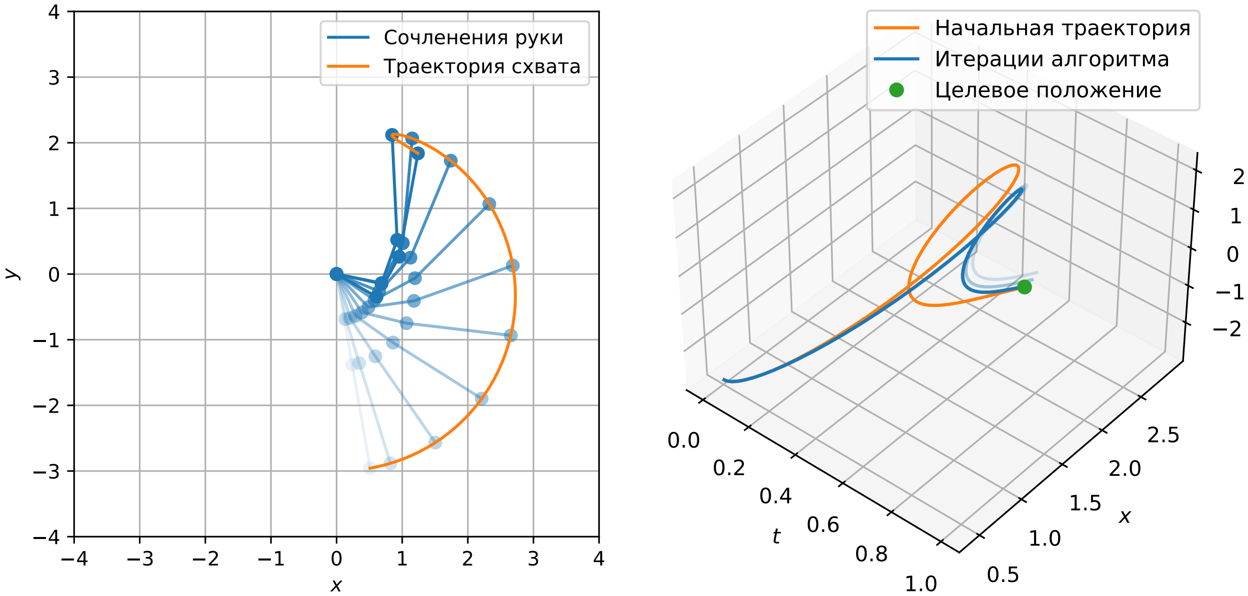

Рис. 4: Оптимальная траектория и траектории схвата на различных итерациях алгоритма при решении задачи (??) с начальным управлением, построеным методом из данного раздела. Алгоритм сошелся на 4 итерации.

[5]

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.