



МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ М. В. ЛОМОНОСОВА

Факультет вычислительной математики и кибернетики

Кафедра системного анализа

Егоров Кирилл Юлианович

О задаче целевого управления по результатам наблюдений, поступающих с запаздыванием

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА

Научный руководитель

к.ф.-м.н., доцент И. В. Востриков

Москва, 2020

Содержание

1	Введение	3
2	Постановка задачи	4
3	Непрерывная линейно-квадратичная задача	6
4	Численный синтез управления	9
5	Пример работы программы	10
6	Система с интервалом между наблюдениями.	16
7	Дискретная задача	19
8	Дискретный пример.	21
9	Заключение.	24

1 Введение

Работа посвящена задаче построения линейно-квадратичного регулятора динамической системы в форме стратегии с запаздывающей обратной связью.

Данная задача прежде всего возникает при передаче сигнала на большие расстояния, либо при помощи компьютерных и прочих сетей. В таких случаях информация о состоянии фазовых переменных поступает с некоторой задержкой. В работе допускается, что эта задержка детерминированна и фазовые координаты передаются без помех.

Способом решения поставленной задачи предлагается метод динамического программирования, разработанный Р. Беллманом [1]. Данный метод предполагает синтезировать управление как минимизатор в уравнении Гамильтона–Якоби–Беллмана.

В работе представлен алгоритм построения оптимальной стратегии, а также редукция к дискретной системе в случае интервалов между замерами фазовых переменных больших, чем величина запаздывания. Представлены примеры работы программы.

2 Постановка задачи

Рассмотрим управляемый объект, положение которого задаётся динамической системой дифференциальных уравнений

$$\frac{dx}{dt} = f(t, x, v) \quad (1)$$

на промежутке времени $t_0 \leq t \leq t_1$ с заданным начальным состоянием

$$x(t_0) = x^0.$$

Информация о состоянии объекта передаётся центру управления с некоторой известной задержкой h_{tc} , а само управление передается объекту с другой известной задержкой h_{fc} . То есть необходимо синтезировать оптимальную стратегию в форме

$$v(t) = v(t, x(t - h_{tc})),$$

на промежутке $t_0 - h_{fc} \leq t \leq t_1 - h_{fc}$. На момент времени t центру известно управление, переданное объекту за все предыдущее время $\tau : t_0 - h_{fc} \leq \tau < t$. При этом система (1) преобразуется к виду

$$\frac{d}{dt}x(t) = f(t, x(t), v(t - h_{fc}, x(t - h_{tc} - h_{fc}))).$$

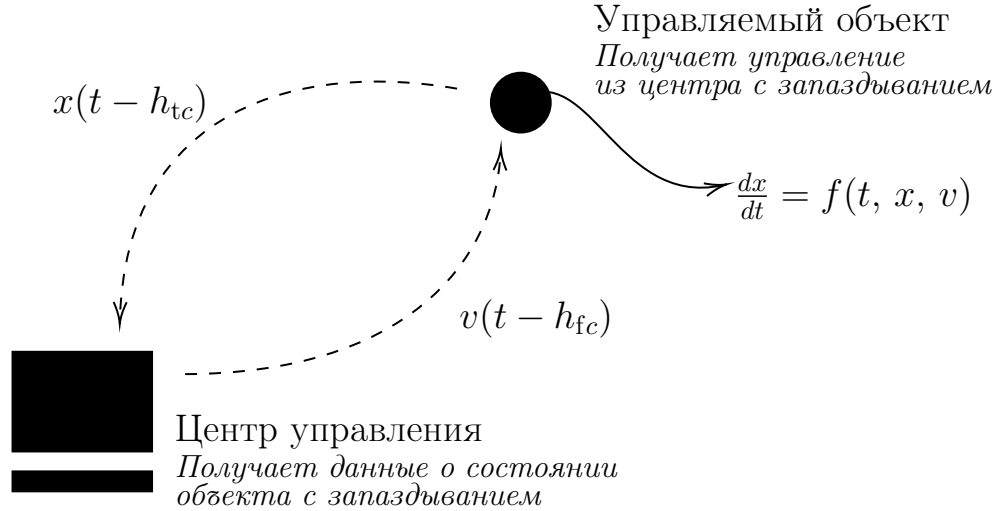


Рис. 1: Иллюстрация поставленной задачи.

Обозначим за h общую величину задержки $h = h_{tc} + h_{fc}$, а за u такое управление, что $u(t) = v(t - h_{fc})$ для всех $t \in [t_0, t_1]$. Тогда мы можем сформулировать общую постановку задачи следующим образом: необходимо найти

оптимальное управление u в форме стратегии $u = u(t, x(t - h))$ для следующей динамической системы

$$\frac{d}{dt}x(t) = f(t, x(t), u(t)), \quad t_0 \leq t \leq t_1 \quad (2)$$

с начальным условием

$$x(t_0) = x^0.$$

3 Непрерывная линейно-квадратичная задача

Рассмотрим линейную динамическую систему с непрерывными ограниченными коэффициентами

$$\frac{d}{dt}x(t) = A(t)x(t) + B(t)u(t) \quad (3)$$

и поставим для него задачу минимизации следующего квадратичного функционала

$$J(u) = \int_{t_0}^{t_1} [\langle x, M(s)x \rangle + \langle u, N(s)u \rangle] ds + \langle x(t_1), Tx(t_1) \rangle \rightarrow \min_u, \quad (4)$$

где $M(s) = M^T(s) \geq 0$, $N(s) = N^T(s) > 0$, $T > 0$.

Обозначим за $\hat{x}(t, x^0, t_0)$ решение системы (3) с заданным начальным условием $x(t_0) = x^0$ и введём в рассмотрение функцию цены

$$V(t, x) = \inf_u \left\{ \int_t^{t_1} [\langle \hat{x}(s, t, x), M(s)\hat{x}(s, t, x) \rangle + \langle u(s), N(s)u(s) \rangle] ds + \right. \\ \left. + \langle \hat{x}(t_1, t, x), T\hat{x}(t_1, t, x) \rangle \right\}.$$

При этом $V(t_0, x^0) = \inf_u J(u)$. Тогда согласно [2] такая функция цены удовлетворяет уравнению Гамильтона–Якоби–Беллмана:

$$\left\{ \begin{array}{l} \frac{\partial V}{\partial t} + \min_u \left\{ \underbrace{\left\langle \frac{\partial V}{\partial x}, Ax + Bu \right\rangle + \langle x, Mx \rangle + \langle u, Nu \rangle}_{\Psi} \right\} = 0, \\ V(t_1, x) = \langle x, Tx \rangle. \end{array} \right. \quad (5)$$

Как следствие получаем, что управление $u^* = \arg \min_u \Psi$, если такое существует, будет оптимальным.

Так как получившаяся функция Ψ строго выпукла по переменной u , то минимум будет достигнут в единственной точке. Запишем необходимое условие экстремума:

$$\begin{aligned} \text{grad}_u \Psi(u^*) &= 0, \\ B^T \frac{\partial V}{\partial x} + 2Nu^* &= 0. \end{aligned}$$

Получается, что оптимальное управление может быть задано как

$$u^* = -\frac{1}{2}N^{-1}B^T \frac{\partial V}{\partial x}. \quad (6)$$

Вернёмся к уравнению (5), подставив в него получившийся результат:

$$\frac{\partial V}{\partial t} + \left\langle \frac{\partial V}{\partial x}, Ax \right\rangle - \frac{1}{4} \left\langle \frac{\partial V}{\partial x}, BN^{-1}B^T \frac{\partial V}{\partial x} \right\rangle + \langle x, Mx \rangle = 0.$$

Будем искать функцию цены как квадратичную форму $V(t, x) = \langle x, P(t)x \rangle$, где $P(t) = P^T(t) > 0$. Тогда

$$\begin{aligned} \left\langle x, \frac{dP}{dt}x \right\rangle + 2\langle Px, Ax \rangle - \langle Px, BN^{-1}B^TPx \rangle + \langle x, Mx \rangle &= 0, \\ \left\langle x, \frac{dP}{dt}x \right\rangle + \langle x, PAx \rangle + \langle A^TPx, x \rangle - \langle Px, BN^{-1}B^TPx \rangle + \langle x, Mx \rangle &= 0, \\ \left\langle x, \frac{dP}{dt}x \right\rangle + \langle x, PAx \rangle + \langle x, A^TPx \rangle - \langle x, PBN^{-1}B^TPx \rangle + \langle x, Mx \rangle &= 0. \end{aligned}$$

Так как в правой части скалярных произведений стоят симметричные матрицы, мы можем записать следующее матричное дифференциальное уравнение, которому должна удовлетворять матрица P :

$$\begin{cases} \frac{dP}{dt} + PA + A^TP - PBN^{-1}B^TP + M = 0, \\ P(t_1) = T. \end{cases} \quad (7)$$

Уравнение (7) называют *уравнением Риккати*. Подробно о свойствах решений уравнения Риккати можно прочитать в работе [3]. Нам достаточно того, что для симметричных матриц данное уравнение имеет единственное решение на рассматриваемом интервале $t_0 \leq t \leq t_1$.

Теперь вспомним, что информация о состоянии системы в текущий момент времени недоступна наблюдению. Воспользуемся формулой Коши:

$$x(t) = X(t, t-h)\xi + \int_{t-h}^t X(t, s)B(s)u(s)ds,$$

где ξ — состояние системы в момент времени $t-h$, а $X(t, s)$ — матрица Коши, определяющаяся следующими соотношениями:

$$\begin{cases} \frac{\partial}{\partial t}X(t, s) = A(t)X(t, s), \\ X(s, s) = I. \end{cases}$$

Таким образом, мы можем записать оптимальную стратегию

$$u^*[t] = -N^{-1}(t)B^T(t)P(t)x[t], \quad (8)$$

где матрица $P(t)$ является решением системы (7), а

$$x[t] = \begin{cases} x_0, & \text{при } t = t_0 \\ X(t, t_0)x_0 + \int_{t_0}^t X(t, s)B(s)u^*[s] ds, & \text{при } t_0 < t < t_0 + h, \\ X(t, t-h)x[t-h] + \int_{t-h}^t X(t, s)B(s)u^*[s] ds, & \text{при } t_0 + h < t \leq t_1. \end{cases}$$

Отметим также, что с точки зрения поставленной задачи интегрирование управления до момента t нам доступно, так как центр может хранить информацию о посланном ранее управлении.

4 Численный синтез управления

Приведем алгоритм для построения управления. Мы будем строить кусочно-постоянное управление на сетке с мелким шагом, как минимум меньшим, чем величина запаздывания. Такой способ не идеален: основная ошибка будет накапливаться не только за счет интегрирования, а в первую очередь из-за выбора класса управления, никак не учтенного в теории. С другой стороны кусочно-постоянные управления чаще встречаются в существующих реальных системах и, при стремлении шага разбиения к нулю, численный алгоритм приближен к теоретическому результату.

1. Введём на отрезке времени $[t_0, t_1]$ равномерное разбиение с шагом ε :

$$\{t^k\}_{k=1}^N, \text{ где } t_0 = t^1 < t^2 < \dots < t^N \leq t_1, \quad t^{k+1} - t^k = \varepsilon, \quad t_1 - t^N < \varepsilon.$$

2. На каждом участке $[t^k, t^{k+1}]$ будем численно строить постоянное управление по формуле (8). Причем запаздывание наблюдения увеличим до величины $\tilde{h} = \lceil \frac{h}{\varepsilon} \rceil \varepsilon$.
3. Решим систему (3) для построенного управления на отрезке $[t^k, t^{k+1}]$ и запомним финальное состояние $x(t^{k+1})$.

Такое разбиение также моделирует невозможность реальных систем непрерывно передавать данные о своем состоянии. Если мы рассматриваем такие системы, то величину ε можно интерпретировать как время между двумя отправками данных. Тогда мы можем провести численные эксперименты поведения посторенного решения при увеличении этого времени.

5 Пример работы программы

Рассмотрим в качестве примера работы программы модель управления скоростью электродвигателя постоянного тока. Данная модель была предложена и исследована на управляемость в [4]. Мы сравним полученные в работе [4] управления без запаздывания с построенными нами.

Итак, модель представляет собой систему:

$$\begin{cases} \frac{d}{dt}i(t) = -\frac{R}{L}i(t) - \frac{K_b}{L}\omega(t) + \frac{1}{L}u(t), \\ \frac{d}{dt}\omega(t) = -\frac{K_T}{J}i(t) - \frac{B}{J}\omega(t). \end{cases} \quad (9)$$

Здесь $i, [\text{A}]$ — сила тока на соответствующем участке цепи; $\omega, \left[\frac{\text{рад}}{\text{с}}\right]$ — угловая скорость вращения; $u, [\text{В}]$ — управляемое нами напряжение на концах цепи.

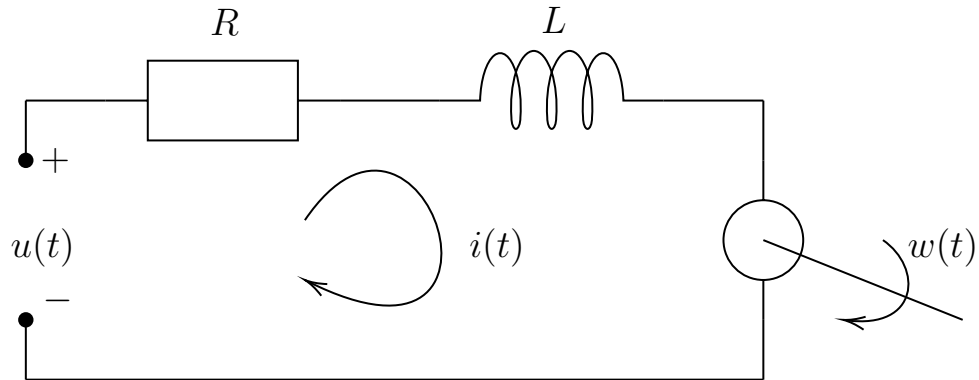


Рис. 2: Схема электрической цепи электродвигателя постоянного тока.

Ниже приведем таблицу с описанием констант в системе (9) и их характерными значениями для электродвигателя:

Обозначение	Физ. величина	Хар. значение
J	Момент инерции	$0,01 \frac{\text{кг}\cdot\text{м}^2}{\text{рад}}$
B	Коэффициент вязкого трения	$0,1 \frac{\text{кг}\cdot\text{м}\cdot\text{с}}{\text{рад}}$
K_T	Постоянная кручения	$0,01 \frac{\text{Н}\cdot\text{м}}{\text{А}}$
K_b	Постоянная ЭДС	$0,01 \frac{\text{В}\cdot\text{с}}{\text{рад}}$
R	Сопротивление резистора	1 Ом
L	Индуктивность катушки	0,5 Гн

Таким образом, мы приходим к удобному нам виду уравнения:

$$\frac{dx}{dt} = \begin{pmatrix} -2 & -0,02 \\ -1 & -10 \end{pmatrix} x + \begin{pmatrix} 2 \\ 0 \end{pmatrix} u. \quad (10)$$

Для него мы поставим задачу минимизации на отрезке $1 \leq t \leq 3$ функционала (4) с матрицами:

$$M = \begin{pmatrix} 1 & 0 \\ 0 & 10 \end{pmatrix}, \quad N = (1), \quad T = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Начальными условиями будут:

$$x(1) = \begin{pmatrix} 4 \\ 100 \end{pmatrix}.$$

На рисунках Рис. 3 и Рис. 4 можно посмотреть оптимальное управление и поведение системы (10) для задачи без запаздывания по наблюдению. С этими графиками мы будем сравнивать управления, построенные по нашему алгоритму. Значение функционала при такой постановке задачи равно $J = 4970,8$.

Теперь построим управление по нашему алгоритму. Будем считать, что запаздывание по наблюдению равно $h = 0,5$. Возьмем мелкое разбиение

$\varepsilon = 10^{-2}$. Как видно из рисунков Рис. 5 и Рис. 6 построенное управление практически не отличается от случая без запаздывания, что соответствует приведённой теории. Значение функционала $J = 4971$.

Предложенный алгоритм предполагает вычисление $\left\lceil \frac{h}{\varepsilon} \right\rceil$ интегралов на каждом шаге работы программы. Это значит, что при увеличении времени задержки, время работы программы так же будет увеличиваться. Если на каждом шаге программы запоминать вычисленное положение $x(t - \varepsilon)$, то можно сократить количество интегралов до одного. График сравнения времени работы программы с и без улучшения алгоритма представлен на рисунке Рис. 7.

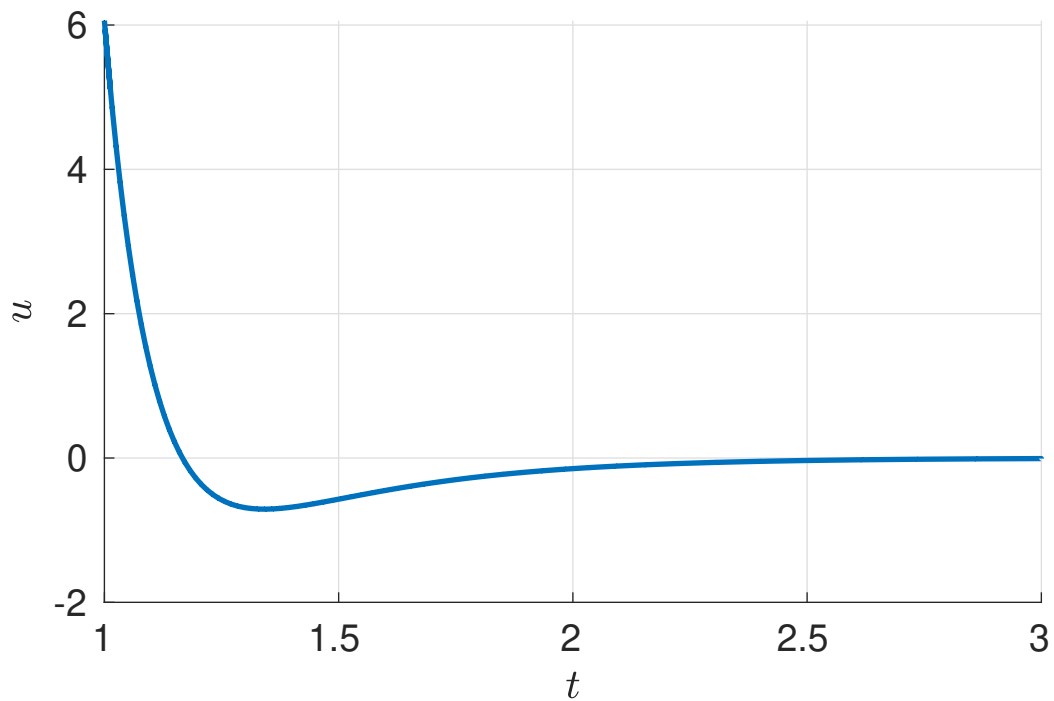


Рис. 3: Оптимальная стратегия без запаздывания наблюдения для системы (10). Значение функционала $J = 4970,8$.

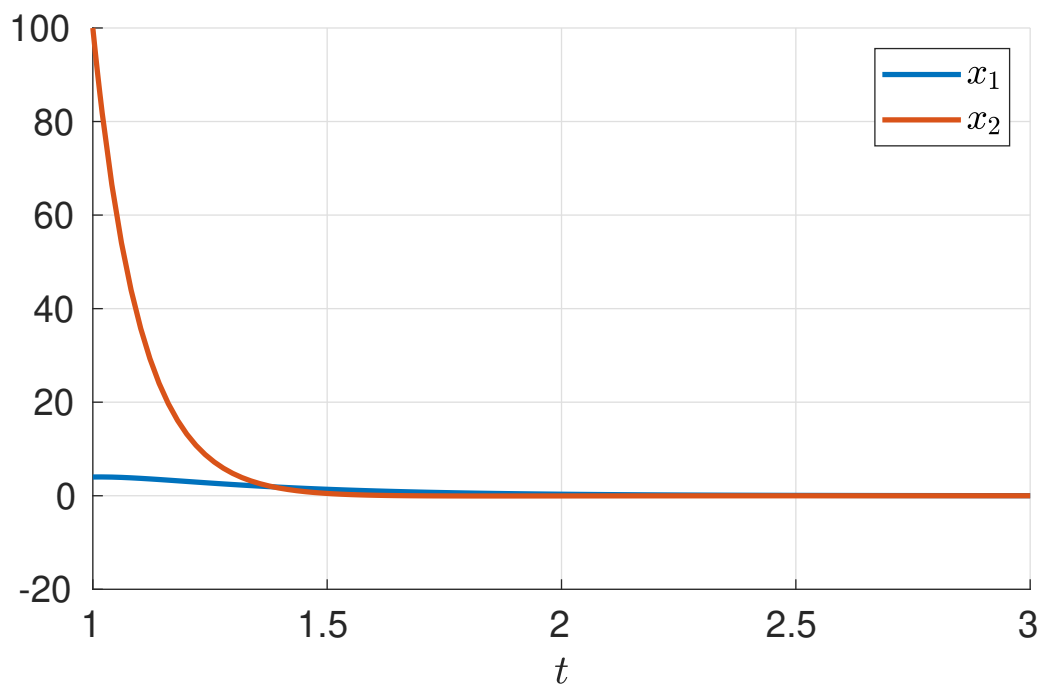


Рис. 4: Поведение системы (10) при использовании оптимальной стратегии без запаздывания наблюдения.

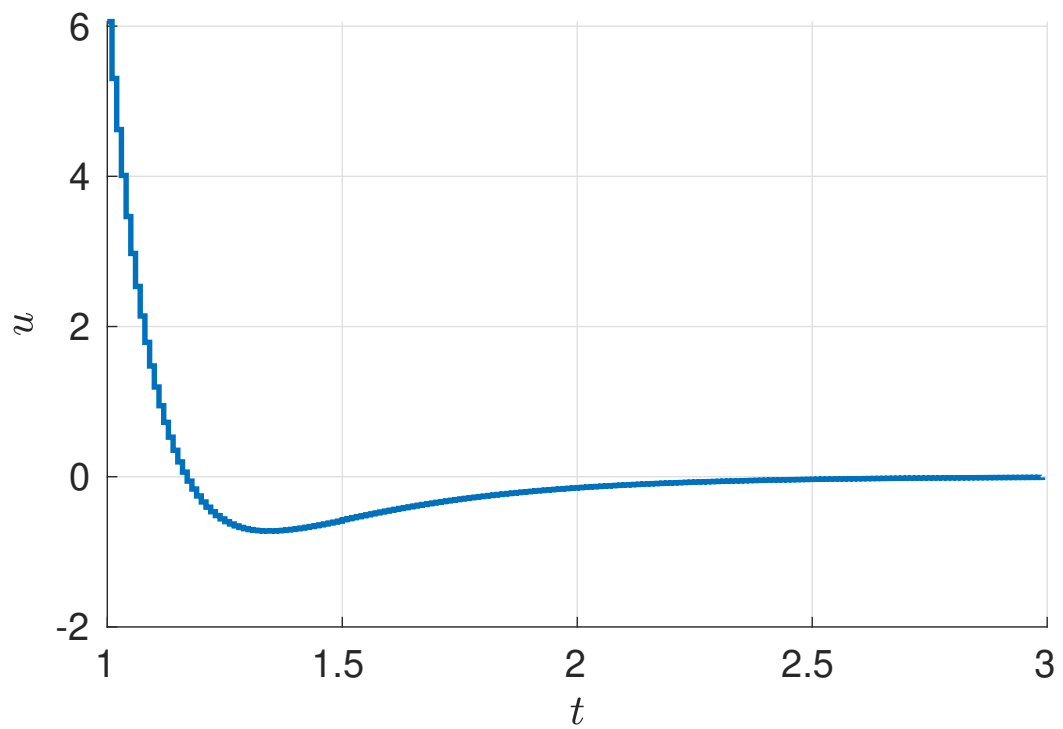


Рис. 5: Оптимальная стратегия с запаздыванием наблюдения $h = 0,5$ для системы (10) с разбиением $\varepsilon = 0,01$. Значение функционала $J = 4971$.

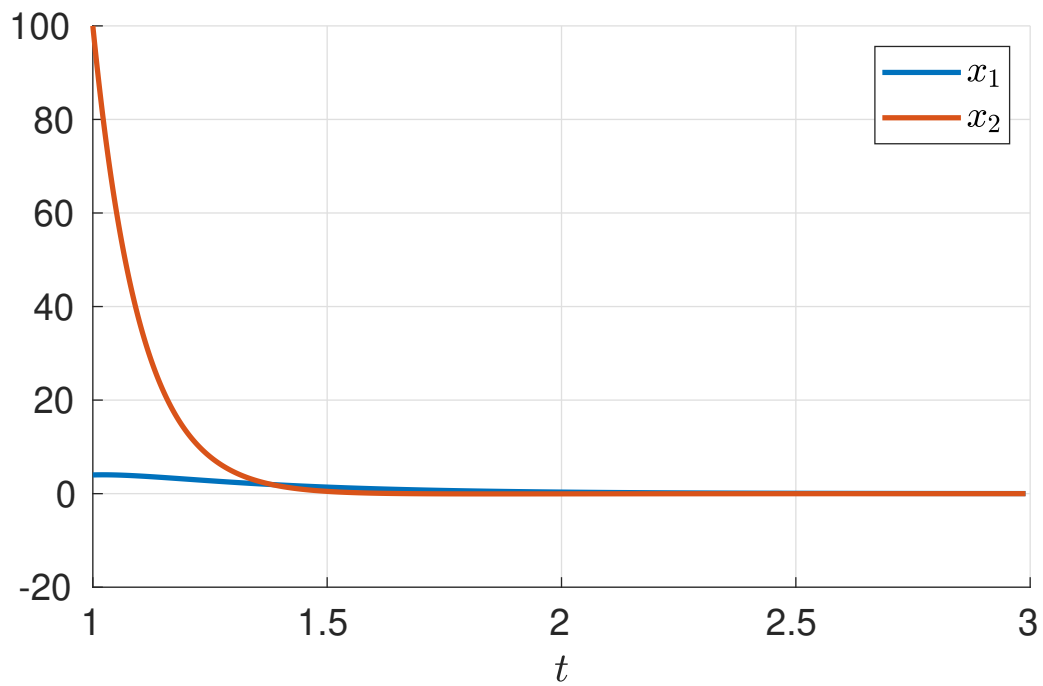


Рис. 6: Поведение системы (10) при использовании оптимальной стратегии с запаздыванием наблюдения (Рис. 5).

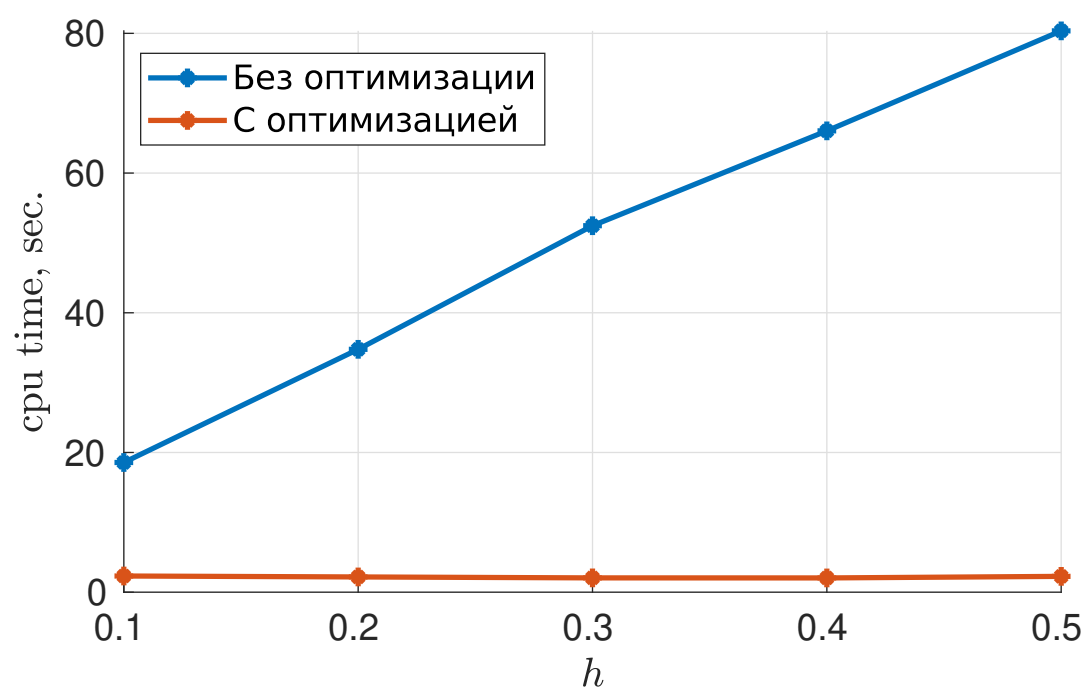


Рис. 7: Время работы программы в зависимости от величины задержки h для изначального и оптимизированного алгоритмов.

6 Система с интервалом между наблюдениями

Мы уже говорили, что реальные системы не могут передавать данные о своём положении непрерывно. В таком случае будем считать, что данные о состоянии передаются с некоторым известным (будем предполагать, детерминированным) интервалом времени. В таком случае можно рассматривать шаг сетки построенного алгоритма ε как время между отправкой наблюдений. При этом мы не можем уменьшать шаг — он становится фиксированным для рассматриваемой системы.

Как видно из рисунков Рис. 9 и Рис. 10, построенные алгоритмом управления приближают исходное в таком случае, но точность становится меньше. Зависимость среднеквадратичного отклонения функционала J_ε от посчитанного без запаздывания J приведена на рисунке Рис. 11.

Особым случаем можно считать ситуацию, при которой время, между двумя отправками данных о состоянии ε становится больше времени запаздывания h . В таком случае кажется естественным провести редукцию системы (3) к дискретному виду.

Мы по прежнему строим кусочно-постоянное управление $u(t) = u^k$ на отрезке $t^{k+1} - h \leq t < t^{k+1}$. Обозначим за $x^k = x(t^k)$. Тогда применим формулу Коши:

$$x^{k+1} = X(t^{k+1}, t^k)x^k + \int_{t^k}^{t^k+h} X(t^{k+1} - h, s)B(s) ds \cdot u^{k-1} + \\ + \int_{t^k+h}^{t^{k+1}} X(t^{k+1}, s)B(s) ds \cdot u^k.$$

Из этого получаем следующую дискретную систему с запаздыванием по управлению:

$$x^{k+1} = \Phi^k x^k + \Gamma_1^k u^k + \Gamma_2^k u^{k-1}. \quad (11)$$

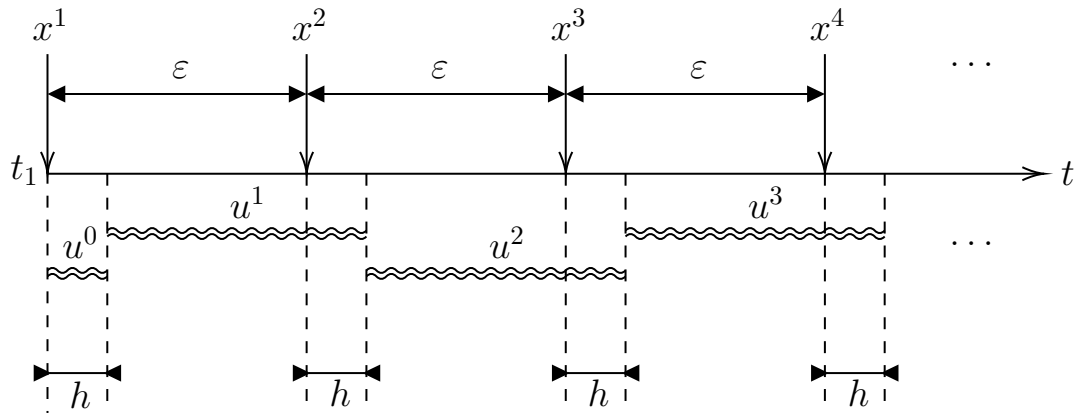


Рис. 8: Иллюстрация системы, где величина запаздывания h меньше интервала времени между наблюдениями ε . Здесь показано в какие моменты времени наблюдается состояние системы и на каких промежутках действуют постоянные управления.

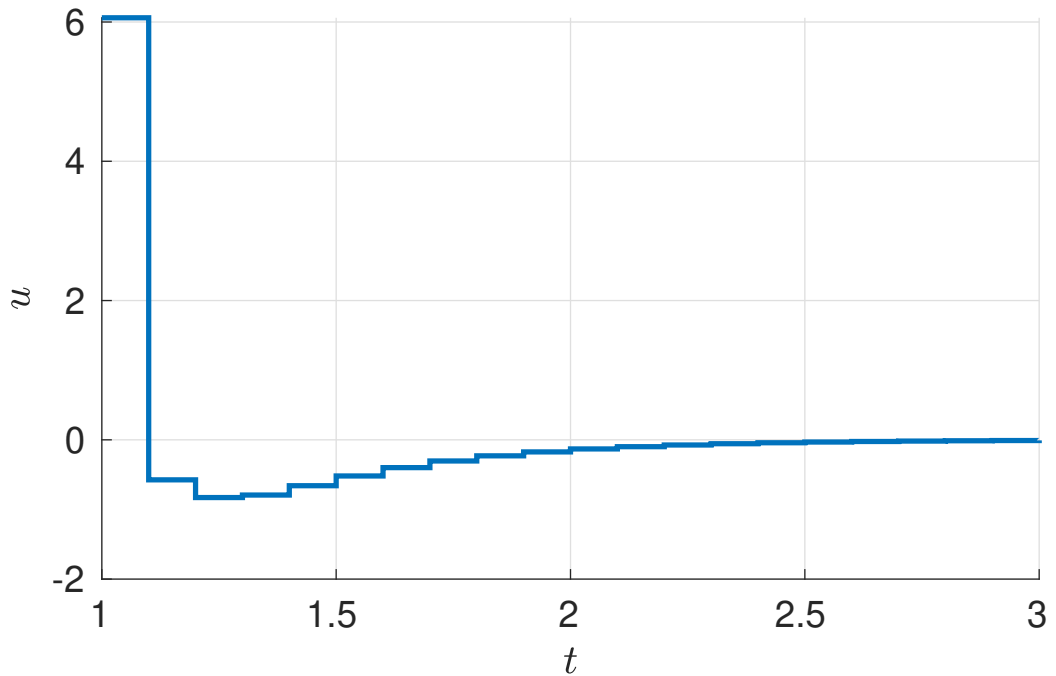


Рис. 9: Оптимальная стратегия с запаздыванием наблюдения $h = 0,5$ для системы (10) с разбиением $\varepsilon = 0,1$.

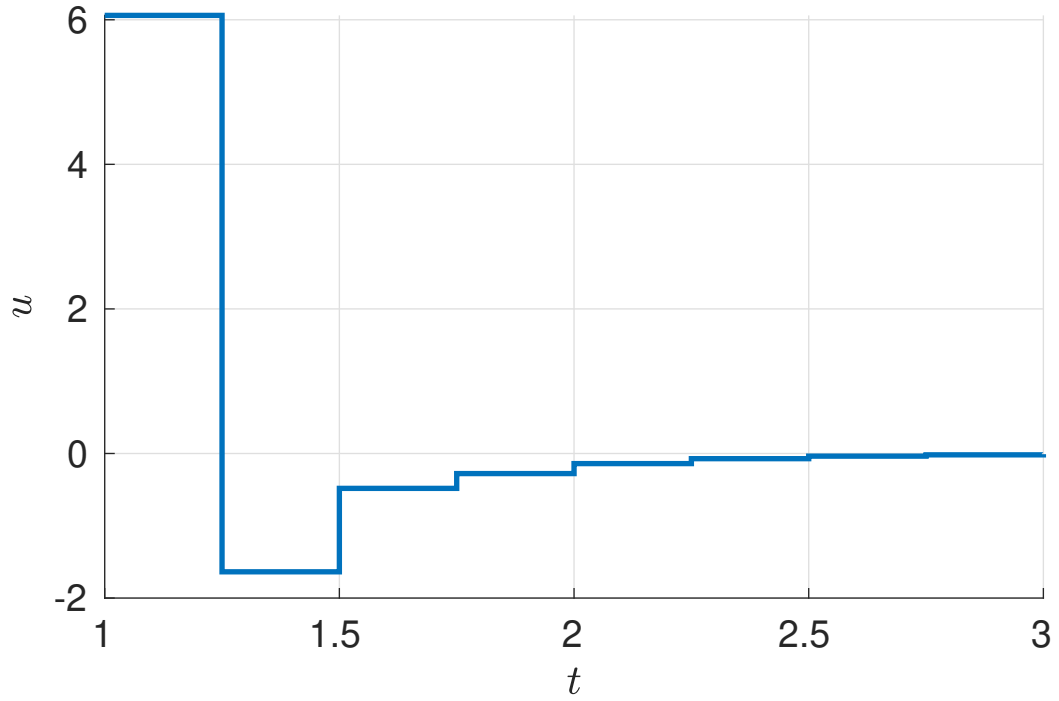


Рис. 10: Оптимальная стратегия с запаздыванием наблюдения $h = 0,5$ для системы (10) с разбиением $\varepsilon = 0,25$.

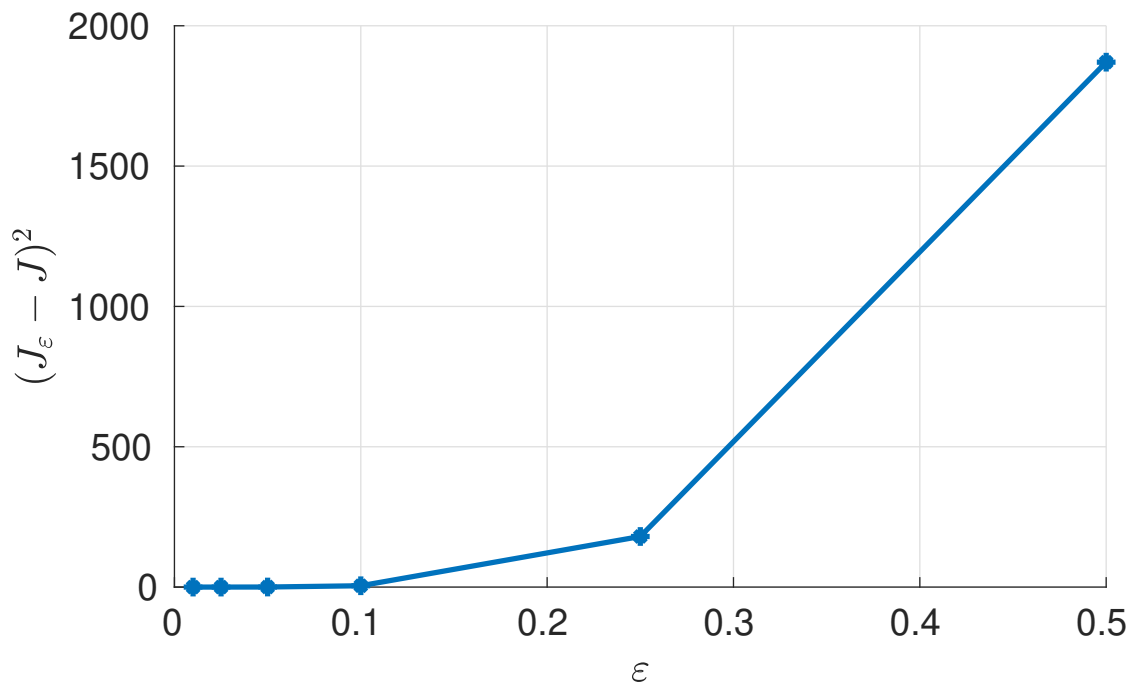


Рис. 11: Среднеквадратичное отклонение функционала качества построенного алгоритма в зависимости от шага сетки ε . Здесь время запаздывания выбрано $h = 0,5$.

7 Дискретная задача

Итак, рассмотрим дискретную динамическую систему

$$x^{k+1} = \Phi^k x^k + \Gamma_1^k u^k + \Gamma_2^k u^{k-1}$$

и составим для нее задачу минимизации функционала

$$J(u) = \sum_{k=1}^N [\langle x^k, M^k x^k \rangle + \langle u^k, N^k u^k \rangle] + \langle x^{N+1}, T x^{N+1} \rangle \rightarrow \min_u,$$

где $M^k = (M^k)^T \geq 0$, $N^k = (N^k)^T > 0$ для всех $k = \overline{1, N}$ и $T = T^T > 0$.

Сделаем замену фазовой переменной, теперь

$$z^k = \begin{pmatrix} x^k \\ u^{k-1} \end{pmatrix},$$

Теперь предыдущее уравнение можно переписать в виде:

$$z^{k+1} = \underbrace{\begin{pmatrix} \Phi^k & \Gamma_2^k \\ O & O \end{pmatrix}}_{\tilde{\Phi}^k} z^k + \underbrace{\begin{pmatrix} \Gamma_1^k \\ I \end{pmatrix}}_{\tilde{\Gamma}^k} u^k,$$

$$J = \sum_{k=1}^N \left\langle z^k, \underbrace{\begin{pmatrix} M^k & O \\ O & O \end{pmatrix}}_{\tilde{M}^k} z^k \right\rangle + \sum_{k=1}^N \langle u^k, N^k u^k \rangle + \left\langle z^{N+1}, \underbrace{\begin{pmatrix} T & O \\ O & O \end{pmatrix}}_{\tilde{T}} z^{N+1} \right\rangle.$$

Утверждение. Матрицы \tilde{M}^k и \tilde{T} являются положительно определенными.

Доказательство. Докажем для матрицы \tilde{T} , для остальных матриц аналогично. По определению:

$$\langle z, \tilde{T}z \rangle = \langle x, Tx \rangle + \langle u, Ou \rangle = \langle x, Tx \rangle > 0.$$

ч.т.д.

Рассмотрим функцию цены $V^k(z)$, она удовлетворяет уравнению Гамильтона–Якоби–Беллмана, которое позволяет определить $V^k(z)$:

$$V^k(z) = \langle z, \tilde{M}z \rangle + \min_u \left\{ \langle u, Nu \rangle + V^{k+1}(\tilde{\Phi}z + \tilde{\Gamma}u) \right\}.$$

Предположим, что $V^{k+1}(z) = \langle z, P^{k+1}z \rangle$, где $P^{k+1} = (P^{k+1})^T \geq 0$ и покажем, что V^k будет иметь ту же форму:

$$V^k(z) = \langle z, \tilde{M}z \rangle + \min_u \left\{ \langle u, Nu \rangle + \langle \tilde{\Phi}z + \tilde{\Gamma}u, P^{k+1}(\tilde{\Phi}z + \tilde{\Gamma}u) \rangle \right\},$$

приравняем производную нулю

$$2u^T N + 2(\tilde{\Phi}z + \tilde{\Gamma}u)^T P^{k+1} B = 0,$$

$$u^* = -(N + \tilde{\Gamma}^T P^{k+1} \tilde{\Gamma})^{-1} \tilde{\Gamma}^T P^{k+1} \tilde{\Phi}z.$$

Тогда функция цены

$$\begin{aligned} V^k(z) &= \langle z, \tilde{M}z \rangle + \langle u^*, Nu^* \rangle + \langle \tilde{\Phi}z + \tilde{\Gamma}u^*, P^{k+1}(\tilde{\Phi}z + \tilde{\Gamma}u^*) \rangle = \\ &= \langle z, [\tilde{M} + \tilde{\Phi}^T P^{k+1} \tilde{\Phi} - \tilde{\Phi}^T P^{k+1} \tilde{\Gamma} (N + \tilde{\Gamma}^T P^{k+1} \tilde{\Gamma})^{-1} \tilde{\Gamma}^T P^{k+1} \tilde{\Phi}]z \rangle = \\ &= \langle z, P^k z \rangle. \end{aligned}$$

Видно, что $P^k = (P^k)^T > 0$.

$$P^{N+1} = T.$$

8 Дискретный пример

Рассмотрим ту же самую модель электродвигателя, что и в позапрошлой секции. Но теперь будем предполагать, что управление запаздывает на $h = 0,05$, а периодичность наблюдений $\varepsilon = 0,2$.

Проведем редукцию той системы к дискретной. Посчитаем все матрицы численными методами.

$$x^{k+1} = \Phi x^k + \Gamma_1 u^k + \Gamma_2 u^{k-1},$$

где

$$\Phi = e^{A\varepsilon} \approx \begin{pmatrix} 0,6705 & -0,0013 \\ -0,0669 & 0,1354 \end{pmatrix},$$

$$\Gamma_1 = \int_h^\varepsilon e^{As} ds \cdot B \approx \begin{pmatrix} 0,2345 \\ -0,0175 \end{pmatrix},$$

$$\Gamma_2 = \int_0^h e^{As} ds \cdot B \approx \begin{pmatrix} 0,0952 \\ -0,0021 \end{pmatrix}.$$

Убедиться в том, что построенная система действительно приближает исходную можно на рисунке Рис. 12.

Оставим все абсолютно так же и посмотрим на графики координат и оптимального управления.

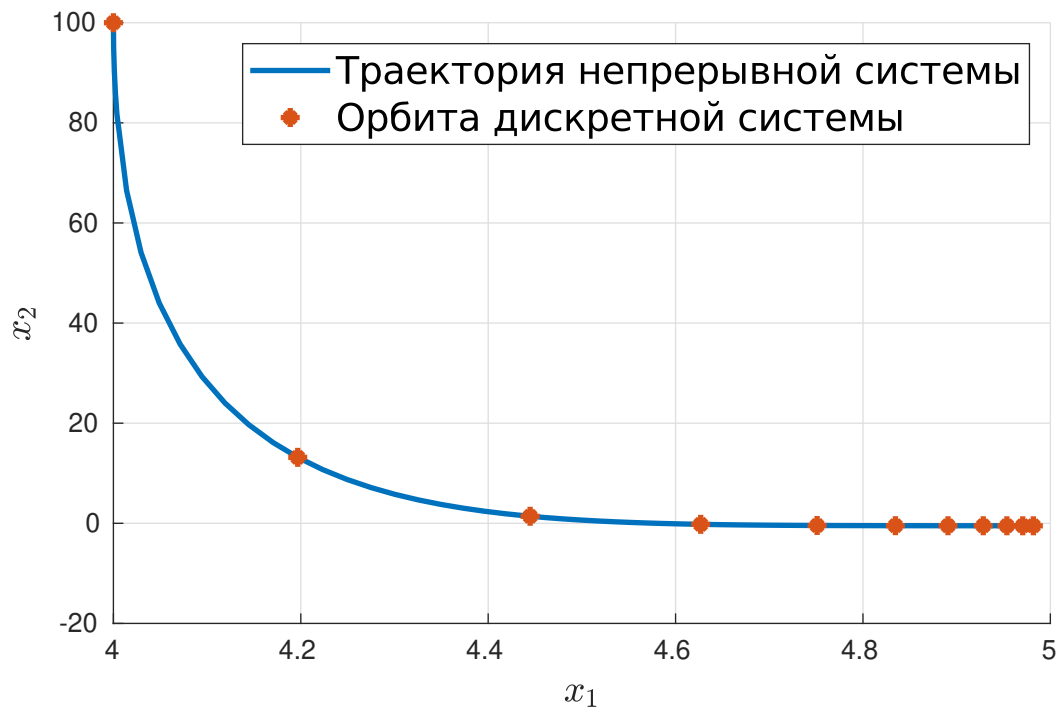


Рис. 12: Траектория непрерывной системы и орбита соответствующей редуцированной дискретной системы выпущенной из начальной точки $x_0 = [4, 100]^T$ с постоянным управлением $u \equiv 1$.

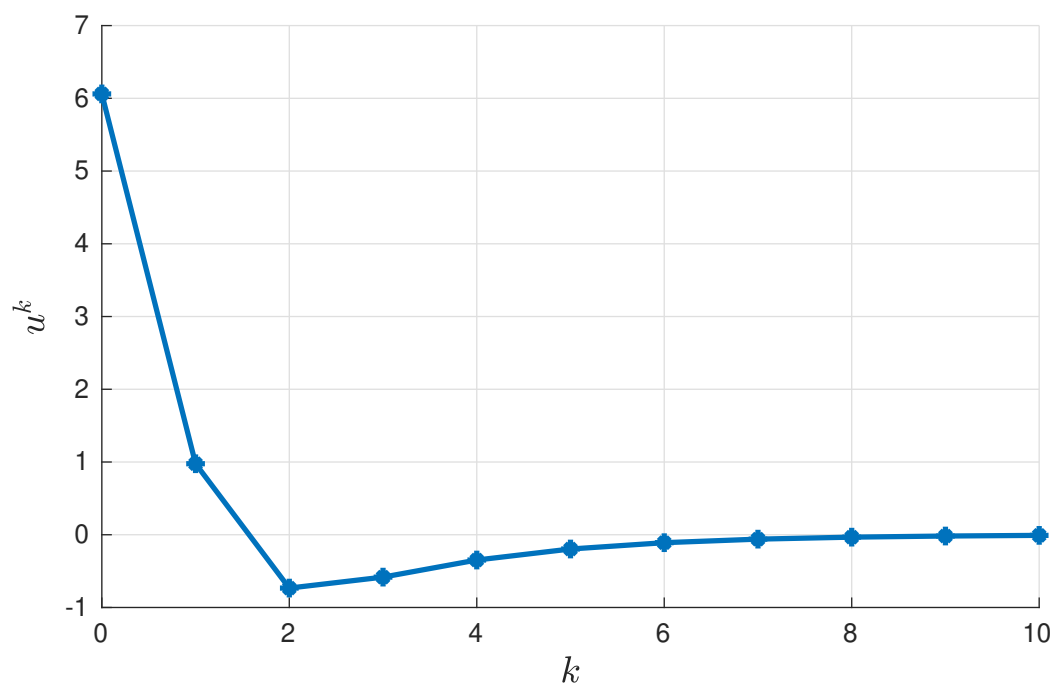


Рис. 13: .

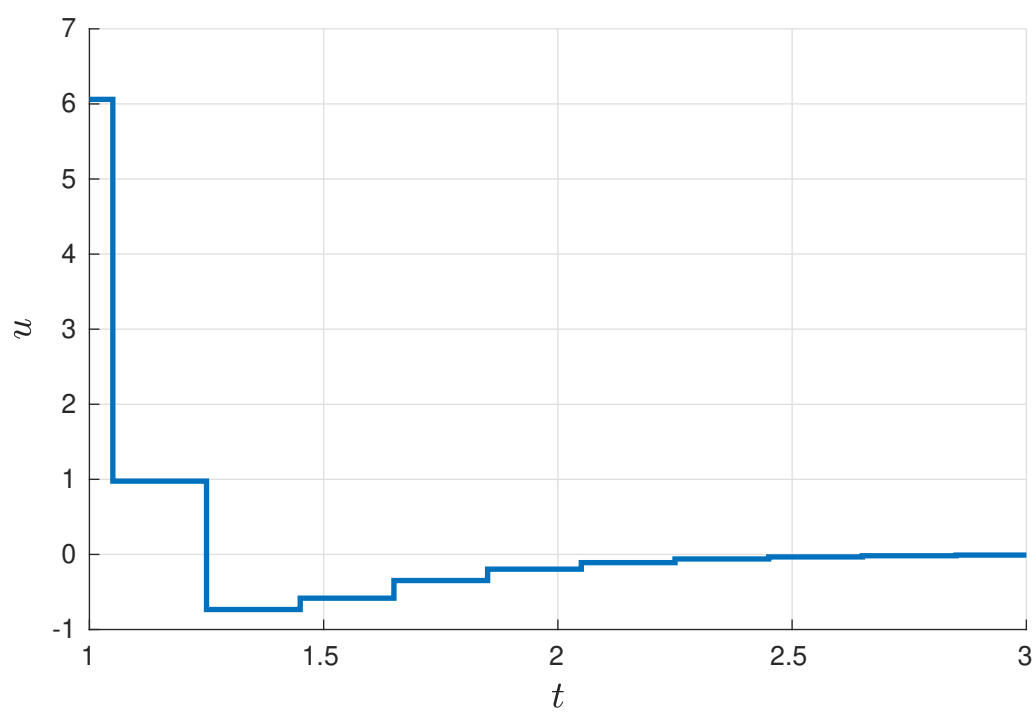


Рис. 14: .

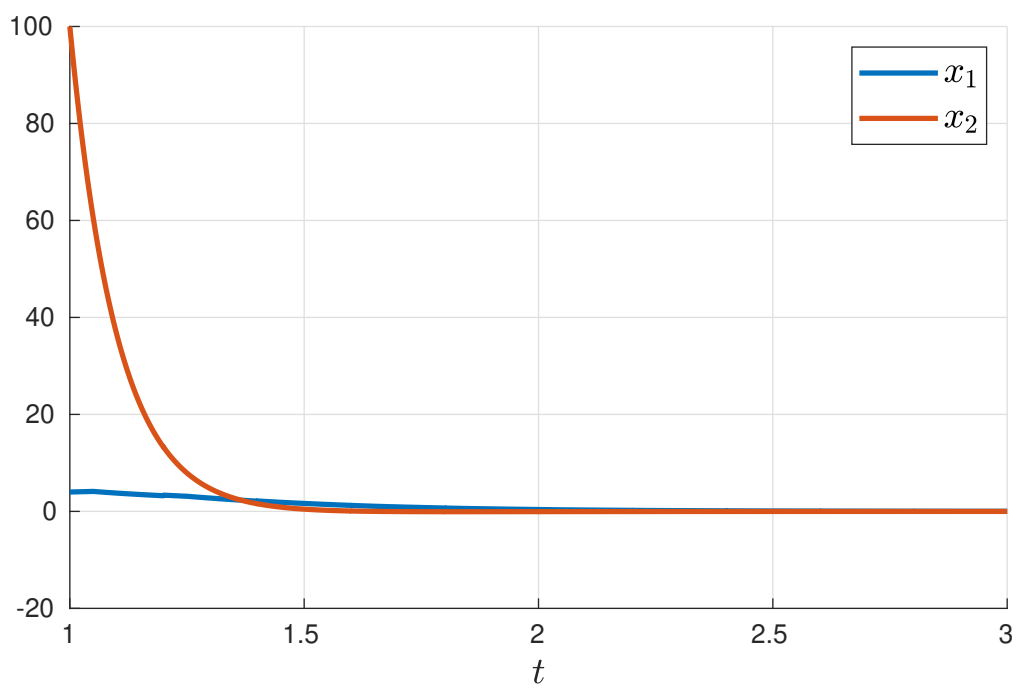


Рис. 15: .

9 Заключение

Список литературы

- [1] Беллман Р. *Динамическое программирование*. М.: Изд-во иностр. лит., 1960, 400с.
- [2] Красовский Н. Н. *Теория управления движением*. М.: Наука, 1968.
- [3] Егоров А. И. *Уравнения Риккати*. М.: Физматлит, 2001, 320с.
- [4] Ruba M. K. Al-Mulla Hummadi *Simulation Of Optimal Speed Control For a Dc Motor Using Linear Quadratic Regulator (LQR)*. Juornal of Engineering, Number 3, Volume 18 march 2012, Baghdad.