

Оптимизация, системный анализ и исследование операций

© 2024 г. П.В. ПАКШИН, д-р физ.-мат. наук (pakshinpv@gmail.com),
Ю.П. ЕМЕЛЬЯНОВА, канд. физ.-мат. наук (EmelianovaJulia@gmail.com)
(Арзамасский политехнический институт (филиал)

Нижегородского государственного технического университета им. Р.Е. Алексеева),
Э. РОДЖЕРС, доктор философии (etar@ecs.soton.ac.uk)
(Университет Саутгемптона, Саутгемптон, Великобритания)

СИНТЕЗ УПРАВЛЕНИЯ С ИТЕРАТИВНЫМ ОБУЧЕНИЕМ ДЛЯ ДИСКРЕТНЫХ СИСТЕМ НА ОСНОВЕ НАБЛЮДАТЕЛЯ СОСТОЯНИЯ С ИСПОЛЬЗОВАНИЕМ МЕТОДА ТЯЖЕЛОГО ШАРИКА¹

Рассматривается задача синтеза управления с итеративным обучением на основе наблюдателя состояния для дискретных линейных систем. С целью ускорения сходимости ошибки обучения предлагается новый подход к ее решению, использующий комбинацию метода тяжелого шарика из теории оптимизации и метода векторной функции Ляпунова для класса двумерных систем, известных как повторяющиеся процессы. Приведен пример, включающий сравнение с известными результатами.

Ключевые слова: управление с итеративным обучением, повторяющиеся процессы, устойчивость, сходимость, наблюдатель состояния, метод тяжелого шарика, векторная функция Ляпунова, линейные матричные неравенства.

DOI: 10.31857/S0005231024080074, EDN: WPDYEU

1. Введение

Метод управления с итеративным обучением (УИО) появился в результате исследований, направленных на повышение точности повторяющихся операций, выполняемых роботами [1], таких как операции выбора и размещения. Ранние результаты в этой активно развивающейся области можно найти, например, в обзорных статьях [2, 3], а более поздние результаты, например, в [4]. В настоящее время УИО эффективно применяется в аддитивном производстве, в частности в установках высокоточного многослойного лазерного напыления [5, 6], в медицинских роботах для реабилитации пациентов, перенесших инсульт [7, 8], в устройствах поддержки желудочка сердца [9] и во многих других приложениях.

¹ Работа выполнена при частичной поддержке Российского научного фонда, грант № 23-71-01044, <https://rscf.ru/project/23-71-01044>.

Класс систем, к которым может применяться УИО, составляют системы, выполняющие одну и ту же задачу конечной продолжительности в повторяющемся режиме. Характерным примером является робот, работающий в режиме захвата и размещения; он забирает полезную нагрузку из указанного места, перемещает ее в течение ограниченного периода времени, устанавливает ее в определенную позицию, а затем возвращается в исходное положение, чтобы повторить эту последовательность задач. Каждое выполнение, обычно называемое в публикациях по УИО пробным шагом, имеет конечную продолжительность. После завершения очередного пробного шага вся информация, генерируемая на его протяжении, становится доступной для обновления управляющего сигнала последующего пробного шага, что подчеркивает итеративный характер УИО. Например, в производстве химикатов, фармацевтических препаратов, потребительских товаров и биопродуктов широко используется пакетная обработка, где одна и та же операция конечной продолжительности применяется к последовательности объектов с ограничением по времени между окончанием обработки одного и началом обработки следующего. Поскольку пакетные процессы имеют повторяющийся характер, то применение УИО для таких процессов вполне обосновано [10].

В рассматриваемых приложениях обычно задается эталонная траектория движения. В операции захвата и размещения эта траектория представляет собой идеальный путь, которому робот должен следовать на каждом пробном шаге. Разности между выборочными значениями на этой траектории и результатами пробных шагов образуют последовательность ошибок обучения. Задача проектирования управления включает в себя построение последовательности пробных входных сигналов. Эти сигналы предназначены для управления движением робота таким образом, чтобы последовательность ошибок сходилась (относительно пробных шагов) либо к нулю, либо к значению в пределах заданного допуска. Переходный процесс в течение пробного шага можно оценить на основании знания конкретной задачи.

Как хорошо известно, двухшаговые методы, такие как метод тяжелого шарика и метод сопряженных градиентов, позволяют существенно ускорить сходимость градиентного метода в задачах оптимизации [11–13]. В этой статье исследуется задача ускорения сходимости ошибки обучения относительно пробных шагов с помощью управления, значение которого на любом пробном шаге является явной функцией входных данных текущего и предыдущих пробных шагов. Анализ основан на использовании метода тяжелого шарика [11], и ключевой идеей здесь является подобие структур градиентного метода и управления с итеративным обучением.

В литературе рассматривались альтернативные стратегии использования данных предыдущих пробных шагов, чтобы ускорить сходимость ошибки обучения. Был проведен значительный объем исследований по синтезу законов УИО, где для создания входных данных для последующего пробного повторения используется более одной ошибки обучения предыдущих повторений, например, [14–20]. Законы УИО, использующие управляющие входные данные

из двух и более предыдущих пробных повторений для вычисления входных данных, используемых на последующем пробном повторении, получили название законов высшего порядка. В общих чертах в этих подходах сначала на основе эвристических соображений определялась структура закона обновления УИО. Далее параметры этого закона находились из условия сходимости ошибок относительно пробных повторений. Выбор количества предыдущих пробных испытаний для достижения наиболее значимого эффекта пока остается открытой исследовательской проблемой.

В [21] описан закон УИО, основанный на методе Нестерова [22], но достигнутая скорость сходимости оказалась медленнее, чем установленная в [22]. По видимому, это связано с использованием так называемого причинного закона УИО. В системах, где применяется УИО, все данные, полученные на текущем пробном шаге, доступны для использования при вычислении закона управления для последующего пробного шага; для дискретной динамики в момент выборки p можно использовать данные, например, из выборки $p + \lambda$ с предыдущего пробного шага. Известно [2], что неиспользование такой информации может серьезно ограничить преимущества закона УИО. В заключение этого краткого анализа отметим, что на текущий момент не получено строгого теоретического доказательства преимуществ использования законов высшего порядка.

На протяжении статьи обозначения переменных имеют вид $h_k(p)$, $0 \leq p \leq N - 1$, где h – рассматриваемый вектор или скалярная переменная, целое число $k \geq 0$ обозначает номер пробного шага, а N – (конечное) количество выборок на текущем пробном шаге (N , умноженное на период выборки, дает длительность пробного шага). Символы $\succ 0$, $\prec 0$, $\succeq 0$ и $\preceq 0$ обозначают соответственно симметричную положительно определенную матрицу, симметричную отрицательно определенную матрицу, симметричную положительно полуопределенную матрицу и симметричную отрицательно полуопределенную матрицу.

2. Постановка задачи

Рассмотрим дискретную систему в повторяющемся режиме, описываемую уравнением состояния

$$(2.1) \quad \begin{aligned} x_k(p+1) &= Ax_k(p) + Bu_k(p), \\ y_k(p) &= Cx_k(p), \quad 0 \leq p \leq N-1, \quad k \geq 0, \end{aligned}$$

где $x_k(p) \in \mathbb{R}^n$ – вектор состояния, $u_k(p) \in \mathbb{R}$ и $y_k(p) \in \mathbb{R}$ – соответственно входная и выходная скалярные переменные, в рассматриваемом классе задач переменная k обычно называется пробным шагом, а $y_k(p)$ – профилем пробного шага. Матрицы A , B и C известны и предполагается, что $CB \neq 0$, это предположение сделано только для компактности изложения и может быть снято, например, с использованием результата [23]. Граничные условия имеют вид $x_k(0) = x_0$, $k \geq 0$, x_0 известен и задан (после завершения пробного

шага система возвращается в исходное состояние), $u_k(p)$ для $k = 0$ задана, ограничена и для $k \geq 1$ вычисляется в соответствии с описанным далее законом обновления. Заметим, что дальнейшее изложение без принципиальных затруднений распространяется на многомерные системы с одинаковыми размерностями входа и выхода.

Предполагается, что система является полностью управляемой и наблюдаемой и на каждом пробном шаге k измерению доступна переменная $y_k(p)$. Для оценки вектора состояния используем асимптотический наблюдатель полного порядка

$$(2.2) \quad \begin{aligned} \hat{x}_k(p+1) &= A\hat{x}_k(p) + Bu_k(p) + G(y_k(p) - C\hat{x}_k(p)), \\ \hat{x}_k(0) &= x_0, \quad 0 \leq p \leq N-1, \quad k \geq 0, \end{aligned}$$

где матрица G выбирается любым известным методом из условия, что ошибка оценивания

$$\tilde{x}_k(p) = x_k(p) - \hat{x}_k(p) \rightarrow 0$$

при $p \rightarrow \infty$.

Замечание 1. При указанных начальных условиях наблюдатель точно восстанавливает вектор состояния. Если x_0 недоступен для наблюдателя и его начальные условия будут другими, переходный процесс по ошибке оценивания приведет к снижению точности. Этот случай требует самостоятельного изучения, выходящего за рамки данной статьи.

Обозначим через $y_{ref}(p)$, $0 \leq p \leq N-1$, $y_{ref}(0) = Cx_0$, заданную желаемую траекторию выходной переменной системы (2.1). Тогда на пробном шаге k переменная

$$(2.3) \quad e_k(p) = y_{ref}(p) - y_k(p)$$

будет означать ошибку воспроизведения желаемой траектории. Далее эту переменную будем называть ошибкой обучения. Задача управления с итеративным обучением состоит в построении последовательности входных переменных $\{u_k\}$ такой, чтобы при всех $0 \leq p \leq N-1$ для $e_k(p)$ и $u_k(p)$ выполнялись условия

$$(2.4) \quad |e_k(p)| \leq \alpha \rho^k, \quad \alpha > 0, \quad 0 < \rho < 1,$$

$$(2.5) \quad \lim_{k \rightarrow \infty} |u_k(p)| = |u_\infty(p)| < \infty.$$

Переменную $u_\infty(p)$ в литературе часто называют обученным управлением. В приложениях УИО продолжительность пробного шага конечна; следовательно, можно добиться сходимости ошибки относительно пробных шагов (по k) для неустойчивых по переменной p траекторий. Тогда существуют два варианта: а) разработать закон управления со стабилизирующей обратной связью, а затем применить УИО к результирующей управляемой системе,

б) разработать закон управления, который одновременно обеспечивает сходимость ошибки относительно пробных шагов к нулю либо к значению в пределах заданного допуска и регулирует динамику вдоль траектории пробного шага. В последнем случае систему с УИО следует рассматривать как двумерную систему, в которой информация распространяется как относительно пробных шагов (k), так и вдоль траектории пробного шага (p).

3. Анализ сходимости и синтез управления

В теории УИО получили распространение два подхода. Подход с использованием расширенной модели для систем с дискретным временем предполагает использование так называемых супервекторов [14]. Поскольку длина профиля конечна, значения переменной на этой длине можно объединить в вектор-столбец, например, для входных данных $\mathbf{U}_k = [u_k(0) \dots u_k(N-1)]$. В результате динамика обновления переменных относительно пробных шагов может быть выражена в виде стандартных разностных уравнений. Такой подход может привести к моделям очень больших размерностей. Более того, во многих случаях необходимо применять управление с обратной связью для регулирования динамики вдоль профиля пробного шага. Альтернативой является представление динамики в виде особой формы 2D-систем – дискретных повторяющихся процессов [24], что позволяет избежать повышения размерности. Этот подход используется в данной статье. Для построения модели в виде повторяющегося процесса в соответствии с [24] введем приращения переменных:

$$(3.1) \quad \begin{aligned} \eta_{k+1}(p+1) &= x_{k+1}(p) - x_k(p), \quad \hat{\eta}_{k+1}(p+1) = \hat{x}_{k+1}(p) - \hat{x}_k(p), \\ \tilde{\eta}_{k+1}(p+1) &= \tilde{x}_{k+1}(p) - \tilde{x}_k(p), \quad 0 \leq p \leq N-1, \quad k \geq 0. \end{aligned}$$

Используя (3.1) и (2.1), (2.2), (2.3), получим

$$(3.2) \quad \begin{aligned} \tilde{\eta}_{k+1}(p+1) &= (A - GC)\tilde{\eta}_{k+1}(p), \\ \hat{\eta}_{k+1}(p+1) &= GC\tilde{\eta}_{k+1}(p) + A\hat{\eta}_{k+1}(p) + B\Delta u_{k+1}(p-1), \\ e_{k+1}(p) &= -CGC\tilde{\eta}_{k+1}(p) - CA\hat{\eta}_{k+1}(p) - CB\Delta u_{k+1}(p-1) + e_k(p), \end{aligned}$$

где $\Delta u_{k+1}(p-1) = u_{k+1}(p-1) - u_k(p-1)$. Закон управления с итеративным обучением обычно формируется как сумма входных данных предыдущего испытания плюс поправочный член (закон обновления). При завершении очередного пробного шага вся информация с предыдущих пробных шагов доступна для формирования управления на следующем пробном шаге. Во многих случаях используются только данные предыдущего пробного шага. Возникает естественный вопрос: “Будет ли полезным использование конечного числа предыдущих пробных шагов, большего одного?” В этой статье рассматривается закон управления, который использует сигналы управления двух предыдущих пробных шагов и процедуру оптимизации для ускорения сходимости ошибки обучения относительно пробных шагов.

Закон управления для (3.2) имеет следующую структуру:

$$(3.3) \quad u_{k+1}(p-1) = u_k(p-1) + \Delta u_{k+1}(p-1)$$

и задача сводится к нахождению обновляющей поправки $\Delta u_{k+1}(p-1)$, обеспечивающей выполнение условий сходимости (2.4) и (2.5). Дополнительно потребуем, чтобы обновляющая поправка удовлетворяла соотношению

$$(3.4) \quad \begin{aligned} \Delta u_{k+1}(p-1) = & \alpha \Delta u_k(p-1) + \beta \nabla \left(\frac{1}{2} e_{k+1}(p)^2 \right) + \\ & + \gamma [\Delta u_k(p-1) - \Delta u_{k-1}(p-1)], \end{aligned}$$

где

$$(3.5) \quad \nabla \left(\frac{1}{2} e_{k+1}(p)^2 \right) = \frac{\partial}{\partial \Delta u_{k+1}(p-1)} \left(\frac{1}{2} e_{k+1}(p)^2 \right).$$

Соотношение (3.4) сформировано по аналогии с алгоритмом метода тяжелого шарика в теории оптимизации. Этот простейший двухшаговый метод ускоряет сходимость градиентного метода. Как будет показано, такой же эффект достигается при управлении с итеративным обучением с обновляющей поправкой, удовлетворяющей (3.4). Заметим, что это лишь один из вариантов использования прошлых пробных шагов в законе УИО.

Замечание 2. В оригинальной работе [11] коэффициент при первом слагаемом в правой части (3.4) равен единице, но в рассматриваемом приложении закон обновления (3.4) будет погружен в динамику системы (3.2) и, в отличие от [11], аргумент $\Delta u_{k+1}(p-1)$ параметризован по переменной p . Следовательно, параметры α, β и γ , гарантирующие сходимость, могут отличаться от рекомендованных в [11].

В силу последнего уравнения из (3.2) выражение под знаком производной в (3.5) можно записать в виде

$$\begin{aligned} \frac{1}{2} e_{k+1}(p)^2 = & [-CGC\tilde{\eta}_{k+1}(p) - CA\hat{\eta}_{k+1}(p) - CB\Delta u_{k+1}(p-1) + e_k(p)]^T \times \\ & \times [-CGC\tilde{\eta}_{k+1}(p) - CA\hat{\eta}_{k+1}(p) - CB\Delta u_{k+1}(p-1) + e_k(p)]. \end{aligned}$$

Вычисляя производную правой части по $\Delta u_{k+1}(p-1)$, получим

$$\begin{aligned} \nabla \left(\frac{1}{2} e_{k+1}(p)^2 \right) = & (CB)^2 \Delta u_{k+1}(p-1) + CBCA\hat{\eta}_{k+1}(p) + \\ & + CBCGC\tilde{\eta}_{k+1}(p) - CBe_k(p). \end{aligned}$$

Подставив это выражение в (3.4), после несложных преобразований имеем

$$(3.6) \quad \begin{aligned} \Delta u_{k+1}(p-1) = & K_1 \Delta u_k(p-1) + K_2 CGC\tilde{\eta}_{k+1}(p) + K_2 CA\hat{\eta}_{k+1}(p) - \\ & - K_2 e_k(p) - K_3 \Delta u_{k-1}(p-1), \end{aligned}$$

где

$$K_1 = \frac{\alpha + \gamma}{1 - \beta(CB)^2}, \quad K_2 = \frac{\beta CB}{1 - \beta(CB)^2}, \quad K_3 = \frac{\gamma}{1 - \beta(CB)^2}.$$

Далее с целью применения техники линейных матричных неравенств (ЛМН) вместо исходных параметров α, β и γ будут использоваться K_1, K_2 и K_3 . Обозначим:

$$\zeta_k(p) = \tilde{\eta}_{k+1}(p), \quad \xi_k(p) = \hat{\eta}_{k+1}(p), \quad \Delta \bar{u}_k(p) = \Delta u_k(p-1).$$

Подставляя (3.6) в (3.2), получим

$$\begin{aligned} \zeta_k(p+1) &= (A - GC)\zeta_k(p), \\ \xi_k(p+1) &= (I + BK_2C)GC\zeta_k(p) + (I + BK_2C)A\xi_k(p) + \\ &\quad + BK_1\Delta \bar{u}_k(p) - BK_3\Delta \bar{v}_k(p) - BK_2e_k(p), \\ (3.7) \quad \Delta \bar{u}_{k+1}(p) &= K_2CA\xi_k(p) + K_1\Delta \bar{u}_k(p) - K_3\Delta \bar{v}_k(p) - K_2e_k(p), \\ \Delta \bar{v}_{k+1}(p) &= \Delta \bar{u}_k(p), \\ e_{k+1}(p) &= -C(I + BK_2C)GC\zeta_k(p) - C(I + BK_2C)A\xi_k(p) - \\ &\quad - CBK_1\Delta \bar{u}_k(p) + CBK_3\Delta \bar{v}_k(p) + (1 + CBK_2)e_k(p). \end{aligned}$$

В модели (3.7) векторы $\zeta_k(p)$ и $\xi_k(p)$ изменяются вдоль траектории повторения (относительно p), а переменные $\Delta \bar{u}_k(p)$, $\Delta \bar{v}_k(p)$ и $e_k(p)$ изменяются от повторения к повторению (относительно k). Системы с таким характером динамики известны как дискретные линейные повторяющиеся процессы [24]. На каждом повторении (пробный шаг в задачах УИО) в таких системах протекает динамический процесс конечной продолжительности. После завершения очередного повторения система возвращается в исходное положение и становится готовой к следующему повторению. Уникальной особенностью является то, что результат любого повторения влияет на динамику следующего, а это может привести к колебаниям, которые увеличиваются от повторения к повторению и не могут управляться стандартными методами. Для таких систем была разработана теория устойчивости в рамках линейной динамики [24], которая была применена к синтезу УИО с экспериментальным подтверждением [25].

Как и при анализе устойчивости и разработке законов управления в стандартной теории линейных систем, одним из подходов к решению может быть построение подходящей функции Ляпунова. В рассматриваемом случае она должна быть построена из суммы членов, изменяющихся относительно пробных шагов и вдоль траекторий пробных шагов. Но уравнения в (3.7) связаны, следовательно, градиент функции Ляпунова (в рассматриваемом случае его дискретный аналог) можно найти только в том случае, если решения этих уравнений известны, что является очень сильным ограничением, и вместо этого необходимо использовать векторную функцию Ляпунова и заменять градиент дивергенцией [26]. (В случае линейной динамики подходы [24, 26] приводят к эквивалентным результатам.)

В данной работе векторная функция Ляпунова для системы (3.7) имеет вид

$$(3.8) \quad V(\bar{\xi}_k(p), \bar{e}_k(p)) = \begin{bmatrix} V_1(\bar{\xi}_k(p)) \\ V_2(\bar{e}_k(p)) \end{bmatrix},$$

где $\bar{\xi}_k(p) = [\zeta_k(p)^T \xi_k(p)^T]^T$, $\bar{e}_k(p) = [\Delta \bar{u}_k(p) \ \Delta \bar{v}_k(p) \ e_k(p)]^T$, $V_1(\bar{\xi}_k(p)) > 0$, $\bar{\xi}_k(p) \neq 0$, $V_2(\bar{e}_k(p)) > 0$, $\bar{e}_k(p) \neq 0$, $V_1(0) = 0$, $V_2(0) = 0$. Определим аналог оператора дивергенции этой функции на траекториях системы (3.2):

$$(3.9) \quad DV(\bar{\xi}_k(p), \bar{e}_k(p)) = V_1(\bar{\xi}_k(p+1)) - V_1(\bar{\xi}_k(p)) + V_2(\bar{e}_{k+1}(p)) - V_2(\bar{e}_k(p)).$$

Справедливо следующее утверждение.

Теорема 1. *Предположим, что существуют векторная функция Ляпунова (3.8) и положительные скаляры c_1, c_2, c_3 такие, что на траекториях системы (3.7) справедливы неравенства*

$$(3.10) \quad c_1 \|\bar{\xi}_k(p)\|^2 \leq V_1(\bar{\xi}_k(p)) \leq c_2 \|\bar{\xi}_k(p)\|^2,$$

$$(3.11) \quad c_1 \|\bar{e}_k(p)\|^2 \leq V_2(\bar{e}_k(p)) \leq c_2 \|\bar{e}_k(p)\|^2,$$

$$(3.12) \quad DV(\bar{\xi}_k(p), \bar{e}_k(p)) \leq -c_3(\|\bar{\xi}_k(p)\|^2 + \|\bar{e}_k(p)\|^2).$$

Тогда выполняются условия сходимости ошибки обучения (2.4) и ограниченности управления (2.5).

Доказательство. В [26, теорема 1] доказано, что при выполнении (3.10)–(3.12) справедлива оценка

$$(3.13) \quad \|\bar{\xi}_k(p)\|^2 + \|\bar{e}_k(p)\|^2 \leq \bar{\varkappa} \varrho^k, \quad 0 \leq p \leq N-1, \quad k \geq 0,$$

где $\bar{\varkappa}$ и $0 < \varrho < 1$ зависят от c_1, c_2 и c_3 . Из (3.13) с учетом определения $\bar{e}_k(p)$ следует

$$(3.14) \quad |e_k(p)|^2 \leq \bar{\varkappa} \varrho^k, \quad |\Delta \bar{u}_k(p)|^2 \leq \bar{\varkappa} \varrho^k.$$

Из первого неравенства в (3.14) следует (2.4) при $\varkappa = \bar{\varkappa}^{\frac{1}{2}}$ и $\rho = \varrho^{\frac{1}{2}}$. Ограниченность управления (2.5) следует из очевидного соотношения

$$|u_k(p)| \leq |u_{k-1}(p)| + |\Delta \bar{u}_k(p)|, \quad 0 \leq p \leq N-1, \quad k = 1, 2, \dots$$

с учетом ограниченности $u_0(p)$ и второго неравенства из (3.14).

Определим следующие матрицы и векторы:

$$\bar{A} = \begin{bmatrix} A - GC & 0 & 0 & 0 & 0 \\ GC & A & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ -CGC & -CA & 0 & 0 & 1 \end{bmatrix}, \quad \bar{B} = \begin{bmatrix} 0 \\ B \\ 1 \\ 0 \\ -CB \end{bmatrix}, \quad \bar{C} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ CGC & CA & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 & 0 \end{bmatrix}.$$

Следующая теорема позволяет найти параметры закона обновления с использованием техники линейных матричных неравенств.

Теорема 2. Предположим, что для некоторой матрицы $Q \succ 0$ и положительного скаляра R система из линейного матричного неравенства и уравнения

$$(3.15) \quad \begin{bmatrix} X & (\bar{A}X + \bar{B}Y\bar{C})^T & X & (Y\bar{C})^T \\ \bar{A}X + \bar{B}Y\bar{C} & X & 0 & 0 \\ X & 0 & Q^{-1} & 0 \\ Y\bar{C} & 0 & 0 & R^{-1} \end{bmatrix} \succeq 0, \quad \bar{C}X = Z\bar{C}$$

имеет решение $X = \text{diag}[X_1 \ X_2]$, Y , Z , где X_1 и X_2 – положительно определенные матрицы размеров $2n \times 2n$ и 3×3 . Тогда закон управления с итеративным обучением

$$(3.16) \quad \begin{aligned} u_{k+1}(p) &= u_k(p) + \Delta u_{k+1}(p), \\ \Delta u_{k+1}(p) &= K_1 \Delta u_k(p) + K_2 CA[\hat{x}_{k+1}(p) - \hat{x}_k(p)] + \\ &+ K_2 CG[(y_{k+1}(p) - y_k(p)) - (\hat{y}_{k+1}(p) - \hat{y}_k(p))] - \\ &- K_2[y_{ref}(p+1) - Cx_k(p+1)] - K_3 \Delta u_{k-1}(p), \end{aligned}$$

где

$$K = [K_1 \quad K_2 \quad K_3] = YZ^{-1},$$

гарантирует выполнение условия сходимости ошибки обучения (2.4) и условия ограниченности управления (2.5).

Доказательство. Применяя к неравенству (3.15) лемму о дополнении Шура, получим

$$X - (\bar{A}X + \bar{B}Y\bar{C})^T X^{-1} (\bar{A}X + \bar{B}Y\bar{C}) - XQX - (Y\bar{C})^T R(Y\bar{C}) \succeq 0.$$

Учитывая, что $Y = KZ$ и $CX = Z\bar{C}$ перепишем это неравенство в виде

$$X - (\bar{A}X + \bar{B}K\bar{C}X)^T X^{-1} (\bar{A}X + \bar{B}K\bar{C}X) - XQX - (K\bar{C}X)^T R(K\bar{C}X) \succeq 0.$$

Умножая последнее неравенство на $P = X^{-1}$ слева и справа и затем представляя слагаемые имеем

$$(3.17) \quad (\bar{A} + \bar{B}K\bar{C})^T P (\bar{A} + \bar{B}K\bar{C}) - P + Q + (K\bar{C})^T R(K\bar{C}) \preceq 0.$$

Выберем компоненты векторной функции (3.8) в виде квадратичных форм

$$(3.18) \quad \begin{aligned} V_1(\bar{\xi}_k(p)) &= \bar{\xi}_k^T P_1 \bar{\xi}_k(p), \\ V_2(\bar{e}_k(p)) &= \bar{e}_k^T P_2 \bar{e}_k(p), \end{aligned}$$

тогда вычисляя дивергенцию (3.9) на траекториях системы (3.7), получим, что условия (3.10)–(3.12) будут выполнены, если

$$(3.19) \quad (\bar{A} + \bar{B}K\bar{C})^T P (\bar{A} + \bar{B}K\bar{C}) - P \prec 0,$$

но последнее неравенство справедливо в силу (3.17). Соотношение (3.16) следует из (3.6) с учетом (3.1) и (2.3).

Замечание 3. Система (3.15) по структуре аналогична соотношениям задачи о линейно-квадратичном регуляторе (ЛКР) в терминах ЛМН. Конкретное значение матрицы $K = [K_1 \ K_2 \ K_3]$ зависит от Q и R , которые выбираются исходя из достижения желаемой скорости сходимости ошибки обучения. Как и в задаче ЛКР, полная формализация такого выбора отсутствует и можно рекомендовать эвристический подход, основанный на знании особенностей динамики рассматриваемой системы.

Для дальнейшего сравнительного анализа рассмотрим наиболее часто применяемый закон обновления (см., например, [27])

$$(3.20) \quad \Delta u_k(p) = \widehat{K}_1 \widehat{\eta}_k(p) + \widehat{K}_2 e_{k-1}(p+1).$$

Подставляя (3.20) в (3.2) и применяя теорему 1 с компонентами функции (3.8)

$$(3.21) \quad \begin{aligned} V_1(\bar{\xi}_k(p)) &= \bar{\xi}_k^T(p) \widehat{P}_1 \bar{\xi}_k(p), \\ V_2(e_k(p)) &= e_k^T(p) \widehat{P}_2 e_k(p), \end{aligned}$$

получим

$$(3.22) \quad \begin{aligned} \mathcal{D}V(\bar{\xi}_k(p), e_k(p)) &= \\ &= [\bar{\xi}_k^T(p) \ e_k(p)] [(\widehat{A} + \widehat{B} \widehat{K} \widehat{C})^T \widehat{P} (\widehat{A} + \widehat{B} \widehat{K} \widehat{C}) - \widehat{P}] [\bar{\xi}_k^T(p) \ e_k(p)]^T, \end{aligned}$$

где $\widehat{P} = \text{diag}[\widehat{P}_1 \ \widehat{P}_2]$,

$$\widehat{A} = \begin{bmatrix} A - GC & 0 & 0 \\ GC & A & 0 \\ -CGC & -CA & 1 \end{bmatrix}, \quad \widehat{B} = \begin{bmatrix} 0 \\ B \\ -CB \end{bmatrix}, \quad \widehat{C} = \begin{bmatrix} 0 & I & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \widehat{K} = [\widehat{K}_1 \ \widehat{K}_2].$$

Повторяя выводы из доказательства теоремы 2, получим, что если система из линейного матричного неравенства и уравнения

$$(3.23) \quad \begin{bmatrix} \widehat{X} & (\widehat{A} \widehat{X} + \widehat{B} \widehat{Y} \widehat{C})^T & \widehat{X} & (\widehat{Y} \widehat{C})^T \\ \widehat{A} \widehat{X} + \widehat{B} \widehat{Y} \widehat{C} & \widehat{X} & 0 & 0 \\ \widehat{X} & 0 & \widehat{Q}^{-1} & 0 \\ \widehat{Y} \widehat{C} & 0 & 0 & \widehat{R}^{-1} \end{bmatrix} \succeq 0, \quad \widehat{C} \widehat{X} = \widehat{Z} \widehat{C}$$

имеет решение $\widehat{X} = \text{diag}[\widehat{X}_1 \ \widehat{X}_2]$, \widehat{Y} , \widehat{Z} , где $\widehat{X}_1 \succ 0$ матрица размера $2n \times 2n$, а \widehat{X}_2 – положительный скаляр, то закон УИО с обновляющей поправкой (3.20) и матрицей $\widehat{K} = \widehat{Y} \widehat{Z}^{-1}$ гарантирует условие сходимости ошибки обучения (2.4) и условие ограниченности управления (2.5).

Замечание 4. Подход на основе применения метода тяжелого шарика приводит к новой структуре закона управления (3.16), отличающейся от используемой в известной литературе. Если при этом значения параметров обеспечивают выполнение условий (2.4) и (2.5), то обновляющая поправка в

соответствии с (3.4) изменяется вдоль направления градиента квадратичной функции этой поправки. В случае, когда обновляющая поправка формируется в соответствии с (3.20), это гарантировать нельзя. По-видимому, этим обеспечивается ускорение сходимости ошибки обучения при использовании закона управления (3.16), но строгого доказательства получить пока не удалось.

4. Пример

В [25] рассмотрен порталный робот, специально спроектированный для имитации операции захвата и размещения, к которому применимо УИО. Этот робот имеет три взаимно перпендикулярные оси, одна из которых X направлена вдоль конвейерной ленты, вторая Y перпендикулярна этой оси в плоскости конвейера. Третья ось Z перпендикулярна плоскости XY (см. рис. 1). Система спроектирована так, что взаимодействием между осями можно пренебречь, а модели для проектирования системы управления были получены на основе частотных характеристик.

Передаточная функция для оси Y равна

$$(4.1) \quad G_Y(s) = \frac{23,736(s + 661,2)}{s(s^2 + 426,7s + 1,744 \times 10^5)}.$$

Желаемая траектория для этой оси показана на рис. 2, а период дискретизации для дискретной реализации $T_s = 0,01$ с. Для построения дискретной

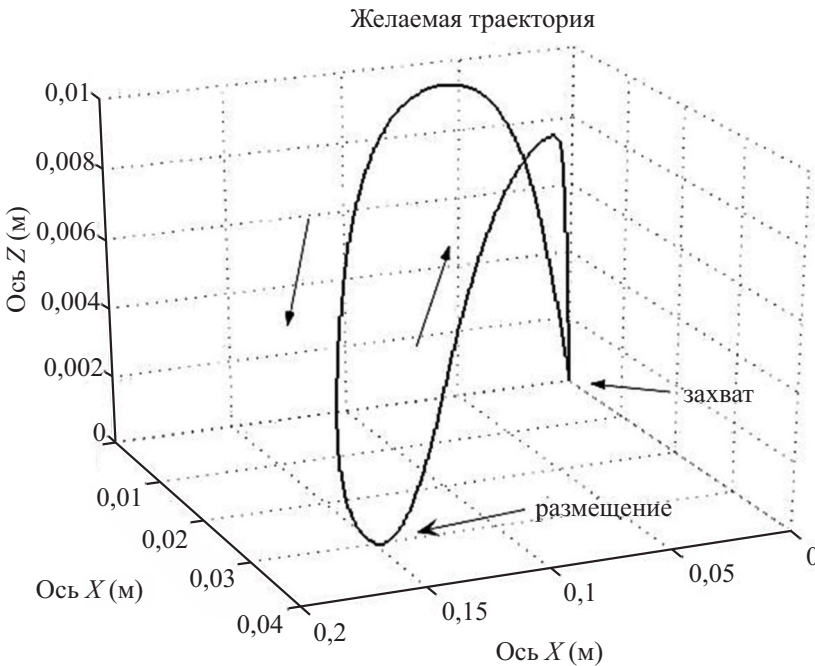


Рис. 1. Желаемая пространственная траектория движения робота.

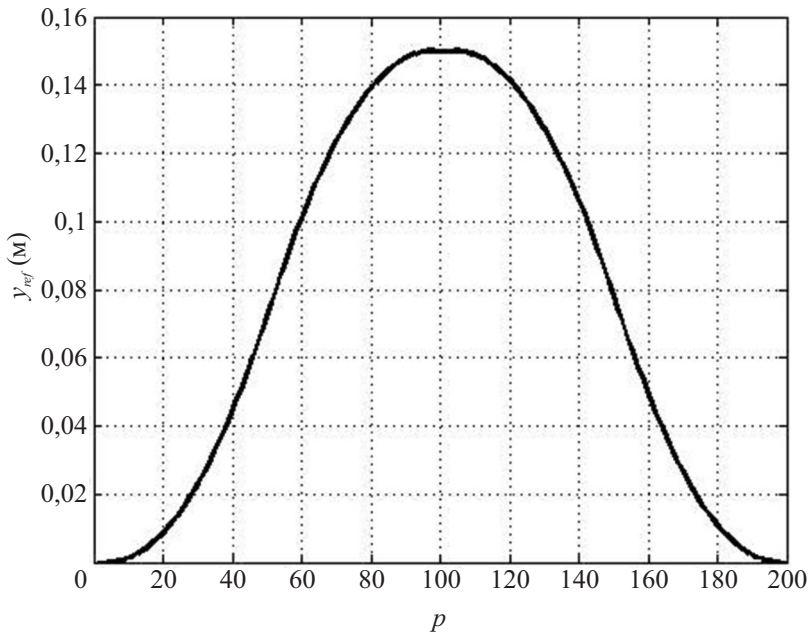


Рис. 2. Желаемая траектория движения робота по оси Y .

модели использованы стандартные функции `ss` и `c2d` пакета MATLAB. Матрицы этой модели имеют следующий вид:

$$A = \begin{bmatrix} -0,0762 & 0,0487 & 0 \\ -0,0732 & -0,1372 & 0 \\ 0,0033 & 0,0026 & 1,0000 \end{bmatrix}, \quad B = \begin{bmatrix} -0,0006 \\ 0,0134 \\ 0,0001 \end{bmatrix}, \quad C = [0 \ 0,0116 \ 7,6631].$$

При оценке эффективности управления необходимо оценивать как сходимость ошибки относительно пробных шагов, так и вид процесса на пробном шаге. На каждом пробном шаге вид процесса может быть оценен с использованием стандартных показателей. Для сходимости ошибки относительно пробных шагов одним из обычно используемых показателей является среднеквадратическая ошибка обучения для каждого пробного шага, отображаемая в зависимости от номера пробного шага, т.е. на шаге k

$$(4.2) \quad E(k) = \sqrt{\frac{1}{N} \sum_{p=0}^{N-1} |e_k(p)|^2}.$$

Для нахождения параметров наблюдателя воспользуемся результатами теории модального управления, реализованными в стандартной функции `place` пакета MATLAB. Задавая собственные значения матрицы $A - GC$ в виде $\lambda_1 = 0,9$, $\lambda_{2,3} = -0,1 \pm 0,01i$, в результате применения этой функции получим

$$G = [-0,3070 \ 2,6427 \ 0,0073]^T.$$

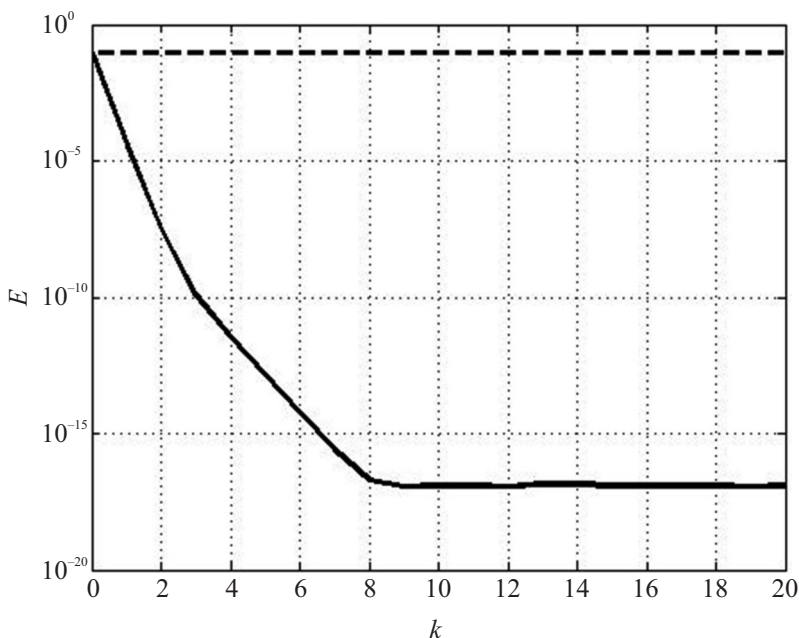


Рис. 3. Изменение $E(k)$ при единичных Q и R при УИО по оси Y с обновляющей поправкой (3.16) (сплошная линия) и при УИО с обновляющей поправкой (3.20) (пунктирная линия).

Как отмечалось при постановке задачи, при сделанных предположениях любая матрица G наблюдателя, при которой собственные значения $A - GC$ располагаются внутри единичного круга, гарантирует точное воспроизведение вектора состояния. Поэтому конкретный выбор собственных значений осуществлялся из условия, что коэффициенты усиления наблюдателя (элементы матрицы G) являются относительно небольшими.

Выбирая Q и R единичными и применяя теорему 2, получим

$$K_1 = -1,09 \times 10^{-7}, \quad K_2 = -1196,0, \quad K_3 = -1,61 \times 10^{-8}.$$

При этих значениях параметров алгоритм (3.16) уже за один пробный шаг уменьшает ошибку обучения в 10^5 раз, а за 8 повторений эта ошибка достигает компьютерного нуля. Алгоритм УИО с обновляющей поправкой (3.20) дает в этом случае снижение ошибки на 1% за 70 пробных шагов (рис. 3).

Заметим, что коэффициенты K_1 и K_3 получились достаточно малыми. Моделирование показывает, что результат не изменится, если их положить равными нулю. В этом случае приходим к УИО с обновляющей поправкой (3.20), в которой $\hat{K}_1 = K_2 CA$ и $\hat{K}_2 = -K_2$. Таким образом, применение метода тяжелого шарика по сути накладывает ограничения на структуру \hat{K}_1 и \hat{K}_2 , при которой обеспечивается высокая скорость сходимости. Попасть в область этих ограничений за счет эвристического выбора Q и R достаточно сложно, а установление связи скорости сходимости ошибки обучения с Q и R

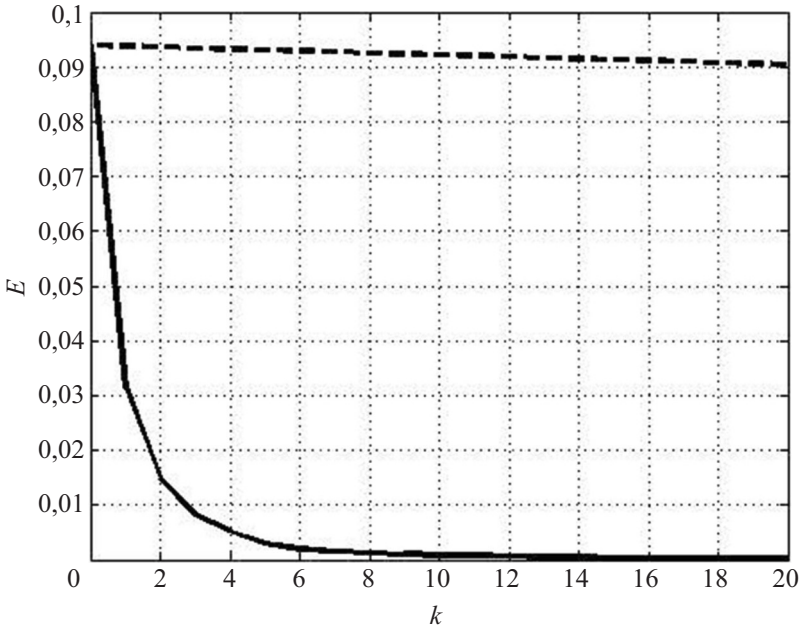


Рис. 4. Изменение $E(k)$ с единичными Q и R при УИО по оси Y с обновляющей поправкой (3.16) при $K_2 = -50$ (сплошная линия) и при УИО с обновляющей поправкой (3.20) при единичных Q и R (пунктирная линия).

является открытой задачей (авторам удалось подобрать матрицы, обеспечивающие примерно в 10 раз меньшую скорость сходимости, и это потребовало существенных затрат времени). С учетом этой особенности можно снижать скорость сходимости уменьшая K_2 с последующей проверкой ЛМН (3.19). На практике это может оказаться важным из-за наличия амплитудных ограничений, приводящих к необходимости компромиссного решения. При $K_2 = -50$ скорость сходимости снижается примерно в 80 раз по отношению к результату, полученному при единичных Q и R в (3.15), тем не менее эта скорость остается выше, чем при единичных Q и R в (3.23) (рис. 4); коэффициенты усиления в случае $K_2 = -50$ и $K_1 = K_3 = 0$ имеют вид

$$(4.3) \quad \hat{K}_1 = [-1,24 \quad -0,93 \quad -383,2], \quad \hat{K}_2 = 50,0,$$

а в случае УИО с обновляющей поправкой (3.20) при единичных Q и R

$$(4.4) \quad \hat{K}_1 = [-21,8 \quad -18,4 \quad -7465,4], \quad \hat{K}_2 = 1,9.$$

Для решения систем линейных матричных уравнений и неравенств использовался программный пакет SeDuMi в комбинации с синтаксическим анализатором YALMIP. Анализируя (4.3) и (4.4), нетрудно заметить, что при отсутствии структурных ограничений на \hat{K}_1 и \hat{K}_2 внутренний алгоритм SeDuMi находит решение в области, обеспечивающей низкую скорость сходимости, в то время

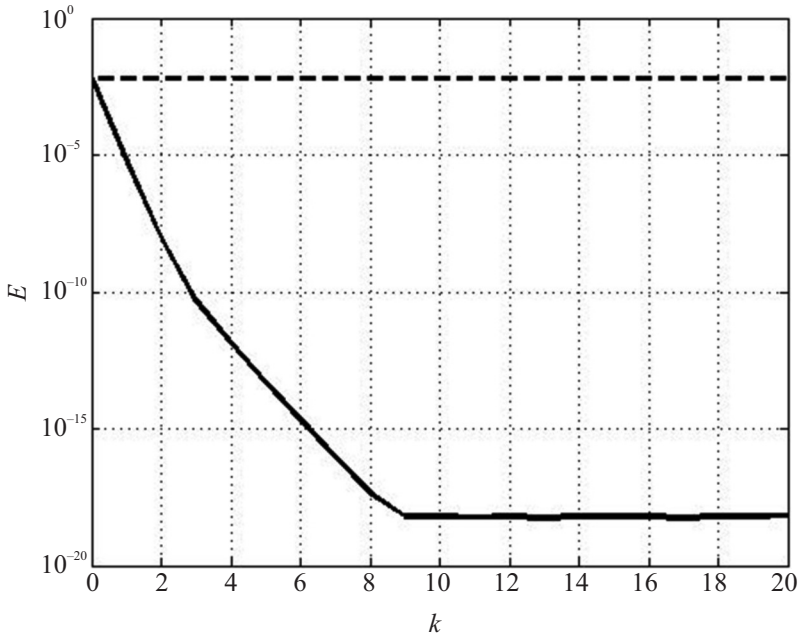


Рис. 5. Изменение $E(k)$ с единичными Q и R при УИО по оси Z с обновляющей поправкой (3.16) (сплошная линия) и с обновляющей поправкой (3.20) (пунктирная линия).

как наличие ограничений, которые дает метод тяжелого шарика, направляет решение в область, обеспечивающую ускорение сходимости.

Рассмотрим построение УИО по осям Z и X , по которым передаточные функции соответственно имеют вид

$$G_Z(s) = \frac{15,8869(s + 661,2)}{s(s^2 + 707,6s + 3,377 \times 10^5)},$$

$$G_X(s) = \frac{13\,077\,183,4436(s + 113,4)(s^2 + 227,9s + 5,467 \times 10^4)}{s(s^2 + 61,57s + 1,125 \times 10^4)(s^2 + 227,9s + 5,467 \times 10^4)(s^2 + 466,1s + 6,142 \times 10^5)}.$$

Для построения наблюдателя по оси Z воспользуемся тем же методом, что и в предыдущем случае задавая полюсы наблюдателя в виде $\lambda_1 = 0,9$, $\lambda_{2,3} = -0,1 \pm 0,01i$. Применяя функцию `place` пакета MATLAB, получим

$$G = [-0,405 \quad -28,01 \quad 0,078]^T.$$

Применение теоремы 2 с единичными Q и R дает

$$K_1 = -6,33 \times 10^{-8}, \quad K_2 = -2,673 \times 10^3, \quad K_3 = -8,33 \times 10^{-9}.$$

Результаты моделирования представлены на рис. 5.

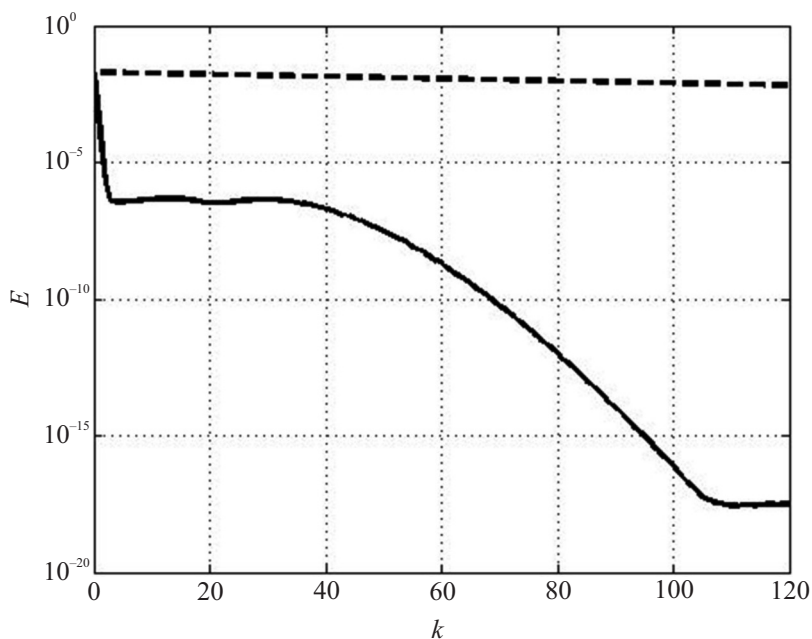


Рис. 6. Изменение $E(k)$ с Q и R вида (4.5) при УИО по оси X с обновляющей поправкой (3.16) (сплошная линия) и с обновляющей поправкой (3.20) (пунктирная линия).

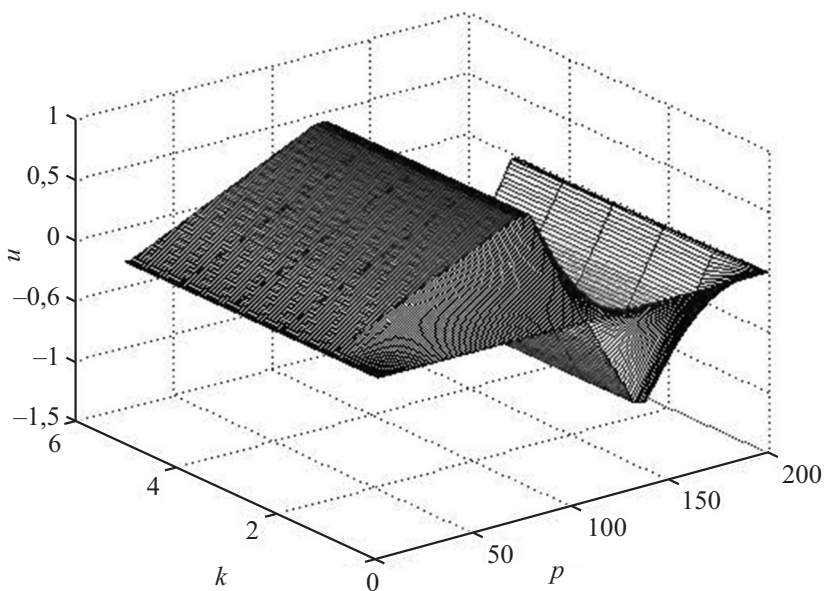


Рис. 7. Изменение управления по оси X при законе обновления (3.16).

Полюсы наблюдателя по оси X зададим в виде

$$\lambda_1 = 0,95, \quad \lambda_2 = -0,4, \quad \lambda_3 = 0,4,$$

$$\lambda_{4,5} = 0,4 \left(\cos \frac{\pi}{3} \pm \sin \frac{\pi}{3} \right), \quad \lambda_{6,7} = 0,4 \left(\cos \frac{\pi}{3} \pm \sin \frac{\pi}{3} \right),$$

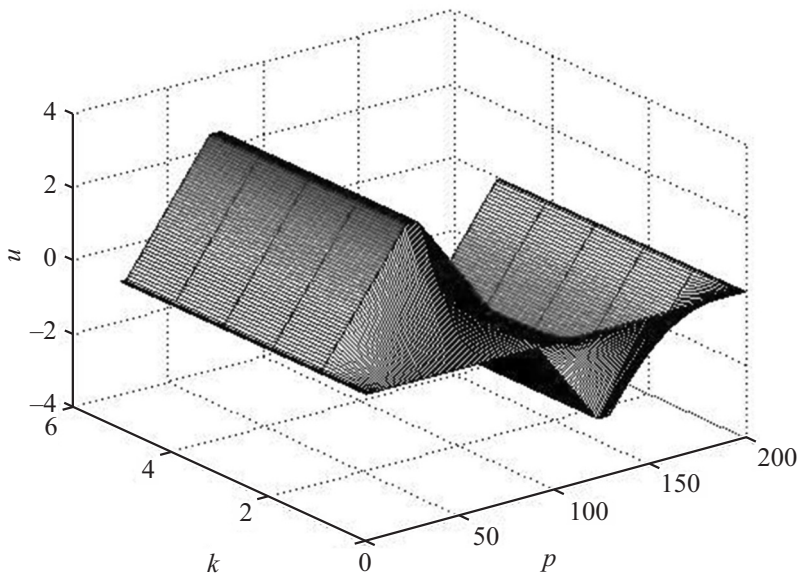


Рис. 8. Изменение управления по оси Y при законе обновления (3.16).

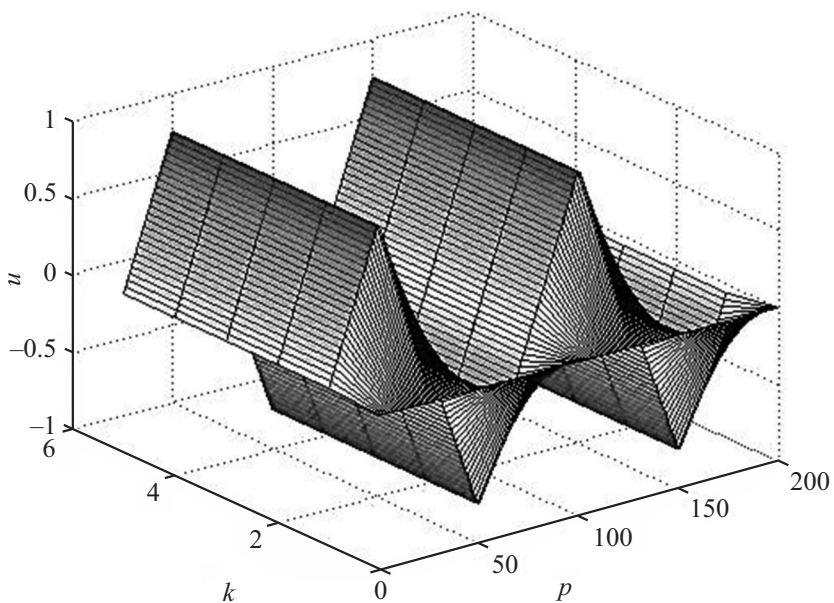


Рис. 9. Изменение управления по оси Z при законе обновления (3.16).

и применяя функцию `place` пакета MATLAB, получим

$$G = [38,1 \quad 40,161 \quad -44,093 \quad 15,913 \quad -16,428 \quad -12,67 \quad 0,309]^T.$$

В данном случае зададим Q и R в виде

$$(4.5) \quad Q = \text{diag}[I_n \quad 10I_n \quad 1 \quad 1 \quad 1], \quad R = 1,$$

тогда, применяя теорему 2, имеем

$$K_1 = -5,715 \times 10^{-9}, \quad K_2 = -1308,2, \quad K_3 = -6,868 \times 10^{-10}.$$

Результаты моделирования представлены на рис. 6. Процессы управления по осям X , Y и Z с законом обновления, заданным (3.16), представлены на рис. 7, 8 и 9. Заметим, что обученное управление по всем осям достигается достаточно быстро (после трех-четырех пробных шагов управления по осям не изменяются в зависимости от k).

Как видно из рис. 5, сходимость ошибки обучения до компьютерного нуля достигается за несколько шагов. По оси X наблюдается иная картина, рис. 6: $E(k)$ уменьшается в 10^7 раз за четыре пробных шага, затем скорость сходимости уменьшается и уровень компьютерного нуля достигается лишь за 107 пробных шагов, в то же время эта скорость существенно превышает скорость сходимости альтернативного алгоритма.

Таким образом, в данном примере алгоритм УИО с использованием метода тяжелого шарика обеспечивает высокую скорость сходимости, которую не удалось достичь альтернативным методом. Заметим, что для всех трех осей коэффициенты K_1 и K_3 получились достаточно малыми. Является ли это общим свойством предлагаемого подхода или это связано с особенностями динамики рассматриваемой системы, пока остается неясным и требует дополнительных исследований.

5. Заключение

Закон УИО (3.16) отличается от известных в литературе специальной структурой матриц усиления, которая, судя по результатам моделирования, является основным фактором, обеспечивающим ускорение сходимости. В случае неструктурированных матриц, как в (3.20), наверно, можно добиться высоких скоростей сходимости, но это будет связано с длительной эвристической процедурой выбора Q и R . Таким образом, можно сделать предварительный вывод, что применение метода тяжелого шарика дает структуру УИО, обеспечивающую высокую скорость сходимости. Окончательный вывод требует строгого обоснования, которое в рамках данной статьи получить не удалось и будет составлять предмет дальнейшего исследования. В ближайшей перспективе будет рассмотрено применение при синтезе УИО простого градиентного метода и метода Ньютона, возможно, полученные структуры УИО дадут ключ к общему доказательству ускорения сходимости. В более отдаленной перспективе предполагается рассмотреть применение метода сопряженных градиентов и метода Нестерова. В целом, существует большой простор для дальнейших исследований в этой обширной области.

Благодарность

Авторы посвящают эту статью памяти Бориса Теодоровича Поляка, автора метода тяжелого шарика и ряда других выдающихся результатов в теории оптимизации и теории управления.

СПИСОК ЛИТЕРАТУРЫ

1. *Arimoto S., Kawamura S., Miyazaki F.* Bettering Operation of Robots by Learning // *J. Robot. Syst.* 1984. V. 1. P. 123–140.
2. *Bristow D.A., Tharayil M., Alleyne A.G.* A Survey of Iterative Learning Control: A Learning-Based Method for High-Performance Tracking Control // *IEEE Control Syst. Magaz.* 2006. V. 26. No. 3. P. 96–114.
3. *Ahn H-S., Chen Y.Q., Moore K.L.* Iterative Learning Control: Survey and Categorization // *IEEE Trans. Syst. Man Cybern. Part C: Appl. Rev.* 2007. V. 37. No. 6. P. 1099–1121.
4. *Rogers E., Chu B., Freeman C., Lewin P.* Iterative Learning Control Algorithms and Experimental Benchmarking Chichester.: John Wiley & Sons, 2023.
5. *Lim I., Hoelzle D.J., Barton K.L.* A multi-objective iterative learning control approach for additive manufacturing applications // *Control Engineer. Practice.* 2017. V. 64. P. 74–87.
6. *Sammons P.M., Gegel M.L., Bristow D.A., Landers R.G.* Repetitive Process Control of Additive Manufacturing with Application to Laser Metal Deposition // *IEEE Transact. Control Syst. Technol.* 2019. V. 27. No. 2. P. 566–575.
7. *Freeman C.T., Rogers E., Hughes A.-M., Burrige J.H., Meadmore K.L.* Iterative learning control in health care: electrical stimulation and robotic-assisted upper-limb stroke rehabilitation // *IEEE Control Syst. Magaz.* 2012. V. 47. P. 70–80.
8. *Meadmore K.L., Exell T.A., Hallewell E., Hughes A.-M., Freeman C.T., Kutlu M., Benson V., Rogers E., Burrige J.H.* The application of precisely controlled functional electrical stimulation to the shoulder, elbow and wrist for upper limb stroke rehabilitation: a feasibility study // *J. NeuroEngin. Rehabil.* 2014. P. 11–105.
9. *Ketelhut M., Stemmler S., Gesenhues J., Hein M., Abel D.* Iterative learning control of ventricular assist devices with variable cycle durations // *Control Engin. Practice.* 2019. V. 83. P. 33–44.
10. *Wang Y., Gao F., Doyle F.J.* Survey on iterative learning control, repetitive control, and run-to-run control // *J. Proc. Control* 2009. V. 19. P. 1589–1600.
11. *Поляк Б.Т.* О некоторых способах ускорения сходимости итерационных методов // *Журн. вычисл. матем. и матем. физ.* 1964. Т. 4. № 5. С. 791–803.
12. *Поляк Б.Т.* Введение в оптимизацию. М.: Наука, 1983.
13. *d’Aspremont A., Scieur D., Taylor A.* Acceleration Methods. arXiv:2101.09545v3. 2021.
14. *Ahn H.S., Moore K.L., Chen Y.Q.* Iterative Learning Control. Robustness and Monotonic Convergence for Interval Systems / *Lecture Notes in Control and Information Sciences.* London: Springer-Verlag, 2007.
15. *Bien Z., Huh K.M.* Higher-order iterative learning control algorithm // *IEE Proc. D-Control Theory Appl.* 1989. V. 136. P. 105–112.
16. *Chen Y., Gong Z., Wen C.* Analysis of a high-order iterative learning control algorithm for uncertain nonlinear systems with state delays // *Automatica.* 1998. V. 34. P. 345–353.
17. *Norrlof M., Gunnarsson S.* A frequency domain analysis of a second order iterative learning control algorithm // *Proc. 38th IEEE Conf. Decis. Control.* 1999. V. 2. P. 1587–1592.

18. *Bu X., Yu F., Fu Z., Wang F.* Stability analysis of high-order iterative learning control for a class of nonlinear switched systems // *Abstract Appl. Anal.* 2013. V. 2013, P. 1–13.
19. *Wei Y.-S., Li X.-D.* Robust higher-order ILC for non-linear discrete-time systems with varying trail lengths and random initial state shifts // *IET Control Theory Appl.* 2017. V. 11. P. 2440–2447.
20. *Wang X., Chu B., Rogers E.* Higher-order Iterative Learning Control Law Design using Linear Repetitive Process Theory: Convergence and Robustness // *IFAC PapersOnLine.* 2017. V. 50-1. P. 3123–3128.
21. *Gu P., Tian S., Y. Chen Y.* Iterative learning control based on Nesterov accelerated gradient method // *IEEE Access.* 2019. V. 7. P. 115 836–115 842.
22. *Нестеров Ю.Е.* Метод решения задачи выпуклого программирования со скоростью сходимости $O(\frac{1}{k^2})$ // *Докл. АН СССР.* 1983. Т. 269. № 3. С. 543–547.
23. *Paszke W., Rogers E., Galkowski K.* Experimentally verified generalized KYP lemma based iterative learning control design // *Control Engin. Practic.* 2016. V. 33. P. 57–67.
24. *Rogers E., Galkowski K., Owens D.H.* Control Systems Theory and Applications for Linear Repetitive Processes / *Lecture Notes in Control and Information Sciences.* V. 349. Berlin: Springer-Verlag, 2007.
25. *Hladowski L., Galkowski K., Cai Z., Rogers E., Freeman C.T., Lewin P.L.* Experimentally supported 2D systems based iterative learning control law design for error convergence and performance // *Control Engin. Practic.* 2010. V. 18. P. 339–348.
26. *Pakshin P., Emelianova J., Emelianov M., Galkowski K., Rogers E.* Dissipativity and stabilization of nonlinear repetitive processes // *Syst. Control Lett.* 2016. V. 91, P. 14–20.
27. *Пакилин П.В., Емельянова Ю.П.* Управление с итеративным обучением дискретными стохастическими системами с переключениями // *А и Т.* 2020. № 11. С. 93–111.

Статья представлена к публикации членом редколлегии П.С. Щербаковым.

Поступила в редакцию 06.04.2024

После доработки 27.06.2024

Принята к публикации 01.07.2024