

Санкт–Петербургский государственный университет

Третьякова Софья Станиславовна

Выпускная квалификационная работа

О влиянии уровня предельно допустимых выбросов на решение дифференциальной игры управления объемами производства

Уровень образования: бакалавриат

Направление 01.03.02 «Прикладная математика и информатика»

Основная образовательная программа СВ.5005.2017 «Прикладная математика, фундаментальная информатика и программирование»

Профиль «Системный анализ, исследование операций и управление»

Научный руководитель:

профессор, кафедра математической теории игр
и статистических решений, д.ф. - м.н.

Громова Екатерина Викторовна

Рецензент:

ассистент, кафедра дифференциальных уравнений
математико-механического факультета СПбГУ.

Родионова Анастасия Александровна

Санкт-Петербург

2021 г.

Содержание

Введение	3
Обзор литературы	6
Принцип оптимальности Беллмана	8
Глава 1. Модель с выигрышем в линейно-квадратичной форме	12
1.1. Постановка задачи	12
1.2. Общий подход к решению	12
1.3. Случай со сменой верхнего предела	14
1.3.1 Владелец предприятия не знает о грядущем изменении	15
1.3.2 Владелец предприятия осведомлён об изменении	16
1.4. Оценка ценности информации	19
Глава 2. Модель с выигрышем в степенном виде	21
2.1. Постановка задачи	21
2.2. Общий подход к решению	21
2.3. Случай со сменой верхнего предела ($\bar{b} < b$)	24
2.3.1 Владелец предприятия не знает о грядущем изменении	24
2.3.2 Владелец предприятия осведомлён об изменении	27
2.3.3 Оценка ценности информации	30
2.4. Случай со сменой верхнего предела ($\bar{b} > b$)	30
2.4.1 Владелец предприятия не знает о грядущем изменении	30
2.4.2 Владелец предприятия осведомлён об изменении	33
2.4.3 Оценка ценности информации	33
Выводы	34
Заключение	35
Список литературы	36

Введение

Хорошо известно, что экологическая проблема в современном мире стоит довольно остро и учёные, занимающиеся этой проблемой придерживаются единого мнения, что если сейчас люди не изменят свое отношение к этой проблеме, то всему человечеству будет грозить опасность. Она состоит не только в том, что в чистый воздух, воду и почву попадают вредные вещества, губительные для живых организмов, но и в вызываемом загрязнением изменении климата Земли. В частности загрязнение воздуха является одним из наиболее серьезных факторов риска, угрожающих здоровью людей. Каждый год от заболеваний, связанных с загрязнением воздуха, умирают миллионы жителей нашей планеты. Наиболее высокая смертность в результате загрязнения воздуха в ресурсодобывающих странах, например, в России или Китае. Именно поэтому многие развитые страны, понимая последствия неэкологичной политики, пытаются активно бороться за охрану природы. Однако очевидно, что если лишь несколько стран будут поддерживать баланс взаимодействия человека и природы, то из этого ничего не выйдет. К примеру, если в одной стране все будут ездить на электромобилях, а производственные загрязнения сведутся к минимуму, то все равно воздух не станет достаточно чистым, так как в нем будут содержаться загрязняющие вещества, производимые соседней страной.

Для того чтобы предпринимать согласованные совместные усилия по сохранению окружающей среды и отдельных ее компонентов, к настоящему моменту подписано несколько десятков конвенций и протоколов. Первым международным соглашением в области охраны окружающей среды стала Конвенция о трансграничном загрязнении воздуха на большие расстояния. Она была подписана в Женеве 13 ноября 1979 г. После этого в дополнение к Конвенции были подписаны 8 протоколов, которые предназначены

для снижения загрязнения воздуха и не только. Из 34 международных природоохранных соглашений 11 касаются вопросов охраны атмосферы, 6 – биологического разнообразия, 5 – обращения с веществами, 3 – водных ресурсов. В области охраны атмосферы подписано также самое многостороннее соглашение – Венская конвенция об охране озонового слоя и Монреальский протокол по веществам, разрушающим озоновый слой к ней. Это соглашение подписали 197 стран.

Несмотря на стремление экологов как можно быстрее ввести новые ограничения на производственные процессы, государство заинтересовано также в поддержке стабильного развития экономики, поэтому оно не стремится принимать быстрых решений. Однако, наличие вышеупомянутых протоколов накладывает на правительство некоторые обязательства. Чтобы урегулировать соблюдение порогового уровня выбросов, вышестоящие, в том числе наднациональные органы, вводят штрафную политику. Если владелец предприятия не будет вовремя осведомлен об изменении максимально возможного уровня загрязнений, то ему придется понести потери в виде штрафа за превышение порога по загрязнениям.

В работе рассмотрена дифференциальная игра, в которой для упрощения рассуждений будет исследовано поведение одного игрока. Таким образом решение дифференциальной игры сведётся к задаче оптимизации выигрыша. В работе рассмотрены две модели, в которых в некоторый момент времени происходит изменение верхнего предела допустимого уровня загрязнений, создаваемых производством. Для данной задачи изучено поведение оптимальных управлений и оптимальных траекторий, с помощью подхода, основанного на расширении классического принципа максимума на случай задач с изменением структуры. Такой подход носит название гибридного принципа максимума [11].

В первой главе рассмотрена модель с выигрышем в линейно-квадратичной форме. Сначала был применен общий подход к решению, без условия смены верхнего уровня выбросов. Затем были рассмотрены два случая решения модифицированной задачи и в конце посчитан коэффициент значимости информации. Во второй главе рассмотрена модель с выигрышем содержащим функцию полезности вида квадратный корень.

Обзор литературы

Одним из основоположников теории дифференциальных игр является американский математик Р. Айзекс [1].

Нельзя не упомянуть значительный вклад русских математиков Петросяна Л. А. [2], Красовского Н. Н. [3], Субботина А. И. [4] и Понтрягина Л. С. [5] в развитие данной теории.

Физические процессы, имеющие место в технике, как правило могут быть управляемы. Из-за этого возникает вопрос о нахождении наилучшего в одном или другом смысле управления. У нас в стране значительный вклад в развитие оптимизации процессов внёс Понтрягин Л. В [6], сформулировал классический принцип максимума Понтрягина.

В классической теории дифференциальных игр [7, 8] и оптимального управления, как правило, рассматриваются задачи без смены режима. Недавно в [9] был описан класс дифференциальных игр, описывающих процесс управления объемом вредных выбросов для случая, когда характеристики игры изменяются непрерывно. Нам будет интересна постановка задачи, в которой структура игры изменяется в некоторые моменты времени, определяемые заранее заданными условиями [11]. В динамических задачах с множественными режимами или так называемыми гибридных дифференциальных играх возникают новые особенности, которых нет в стандартных постановках. Для таких задач применим гибридный принцип максимума, сформулированный S. Shaikh и P. Caines [18] и развитый далее в работах Громовым Д. В и Громовой Е. В. в [10, 11]. Необходимо отметить, что спектр работ посвященных гибриднему принципу максимума очень широк. Упомянутые работы содержат обширный обзор литературы, посвященной этой теме.

Применительно к решению задач, поставленных в рамках квалифи-

кационной работы, был использован подход к решению задач гибридного оптимального управления, описанный в [11] и основанный на использовании принципа оптимальности Беллмана, сформулированным Р. Беллманом [13].

В настоящее время активно исследуется задача определения ценности информации. Впервые эта характеристика была предложена в [14], однако только последнее время она была использована в контексте теории дифференциальных игр. В предложенной работе будет рассмотрена задача определения ценности информации в контексте двух дифференциальных игр управления объемом загрязнений с различными видами функционала выигрыша.

Принцип оптимальности Беллмана

Мы начнем с рассмотрения следующей задачи оптимального управления. Пусть динамика системы задается следующим уравнением:

$$\dot{x} = f(x(t), u(t), t), \quad x(0) = x_0. \quad (1)$$

Выигрыш, который мы минимизируем может быть записан как:

$$J(x, u) = \int_{t_0}^T f_0(x(\tau), u(\tau)) d\tau + F_0(x(T)). \quad (2)$$

Также мы не накладываем жесткое ограничение на терминальное состояние, позволяя свободно выбирать $x(T)$.

Решение описанной задачи оптимального управления основывается на следующем общем принципе, который был сформулирован Р. Беллманом и поэтому называется Принципом оптимальности Беллмана:

Оптимальная стратегия имеет свойство, что какими бы ни были начальное состояние и начальное решение, последующие решения должны составлять оптимальный курс действий по отношению к состоянию, полученному в результате первого решения.

Давайте рассмотрим этот принцип более подробно. Предположим, что мы решаем задачу оптимизации (1)-(2) и получим оптимальные управление и траекторию на отрезке $[t_0, T]$.

Принцип оптимальности Беллмана гласит, что если мы рассмотрим одну и ту же оптимизационную задачу (1)-(2), но с той лишь разницей, что она начинается в момент времени τ_s из точки $x(\tau_s)$, то оптимальное решение будет совпадать с оптимальной траекторией $x^*(t)$ на интервале $[\tau_s, T]$. Другими словами, "хвост" любой оптимальной траектории также

является оптимальным.

Покажем корректность данного результата. Предположим, что существует другое решение $\tilde{x}(t)$ при $t \in [\tau_s; T)$ и соответствующее управление $\tilde{u}(t)$, такое, что $J(\tilde{x}, \tilde{u}) < J(x^*, u^*)$ на интервале $[\tau_s, T)$. В этом случае пара $(\hat{x}(t), \hat{u}(t))$ определяется как

$$(\hat{x}(t), \hat{u}(t)) = \begin{cases} (x^*(t), u^*(t)), & t \in [t_0, \tau_s], \\ (\tilde{x}(t), \tilde{u}(t)), & t \in [\tau_s, T]. \end{cases}$$

и приводит к уменьшению значения функционала $J(\hat{x}(t), \hat{u}(t))$, т. е. $J(\hat{x}(t), \hat{u}(t)) < J(x(t), u(t))$, что противоречит предположению, следовательно $(x^*(t), u^*(t))$ является оптимальным решением.

Принцип оптимальности Беллмана подразумевает, что любая задача оптимального управления может быть разбита на несколько подзадач оптимального управления путем разбиения соответствующего временного интервала на несколько подынтервалов. Для примера разделим временной промежуток $[t_0, T)$ точкой τ_s на отрезки $[t_0, \tau_s)$ и $[\tau_s, T)$.

Общий выигрыш можно записать в виде суммы:

$$J(x, u) = J_1(x, u) + J_2(x, u) = \int_{t_0}^{\tau_s} f_0(x(\tau), u(\tau))d\tau + \int_{\tau_s}^T f_0(x(\tau), u(\tau))d\tau. \quad (3)$$

Принцип оптимальности Беллмана подразумевает, что эту проблему необходимо решать начиная со второго интервала. Однако мы не знаем начальное значение $x(t)$ в точке $t = \tau_s$, поэтому мы обозначим это значение с помощью новой переменной $x_s = x(\tau_s)$.

Теперь мы можем решить задачу оптимального управления на втором интервале:

$$\dot{x} = f(x(t), u(t), t), \quad x(\tau_s) = x_s,$$

$$J_2(x, u) = \int_{\tau_s}^T f_0(x(\tau), u(\tau)) d\tau. \quad (4)$$

Пусть $(x_2^*(t), u_2^*(t)), t \in [\tau_s, T]$ - оптимальное решение (4). Тогда оптимальное значение второго интеграла в выражении (3) выглядит следующим образом:

$$W_2(x_s) = \min_u J_2(x, u) = \int_{\tau_s}^T f_0(x_2^*(\tau), u_2^*(\tau)) d\tau.$$

Заметим, что значение функции Беллмана на интервале $[\tau_s, T]$, т.е. $W_2(\tau_s, x_s)$, так же как и $x_2^*(t)$ и $u_2^*(t)$ параметризованы переменной x_s тогда как момент времени τ_s фиксирован.

Теперь можно решить проблему на первом временном интервале. Для этого заменим второй интеграл в (3) на $W_2(x_s)$ и решим задачу оптимального управления на участке $[t_0, \tau_s]$:

$$\begin{aligned} \dot{x} &= f(x(t), u(t), t), \quad x(t_0) = x_0, \quad x(\tau_s) = x_s, \\ J(x, u) &= \int_{t_0}^{\tau_s} f_0(x(\tau), u(\tau)) d\tau + W_2(x_s). \end{aligned} \quad (5)$$

Заметим, что задача (5) эквивалентна ранее сформулированной задаче (3) с тем лишь отличием, что вместо второго интеграла выигрыша мы подставляем полученное ранее решение оптимизационной задачи (4). Таким образом, исходная задача с двумя режимами сводится к задаче оптимального управления с одним режимом и терминальной функцией выигрыша $W(x(\tau_s))$.

Задача (5) решается с помощью принципа максимума Понтрягина. При решении поставленной задачи полезно использовать то свойство, что в случае переключения, зависящего только от времени (т. н. автономное переключение), значение сопряженной переменной не претерпевает разрыва в точке $t = \tau_s$.

Обозначим найденную пару оптимальное состояние, оптимальное управление как $(x_1^*(t), u_1^*(t))$. Однако полученное решение определено только на интервале $[0, \tau_s]$. Для того, чтобы найти выражение для оптимального решения на интервале $[\tau_s, T]$, необходимо определить значение x_s как $x_s = x_1^*(\tau_s)$. Используя это значение, можно получить выражения для $x_2^*(t)$ и $u_2^*(t)$.

Окончательное выражение для оптимального управления будет иметь вид

$$u^*(t) = \begin{cases} u_1^*(t), & t \in [0, \tau_s) \\ u_2^*(t), & t \in [\tau_s, T]. \end{cases}$$

Выражение для оптимальной траектории $x^*(t)$ имеет аналогичный вид.

Глава 1. Модель с выигрышем в линейно-квадратичной форме

1.1 Постановка задачи

Пусть имеется некоторое промышленное производство. Предполагается, что объем производства пропорционален загрязнениям. Количество выбросов в атмосферу, в свою очередь, ограничено сверху. Стратегия заключается в выборе управления для максимизации выигрыша.

Динамика изменения объема загрязнений описывается следующим дифференциальным уравнением:

$$\dot{x} = u, \quad u \in [0, b], \quad x(0) = x_0. \quad (6)$$

Интегральный выигрыш игрока имеет вид

$$J(x_0, T, u) = \int_0^T \left(\left(b - \frac{1}{2}u \right) u - dx \right) dt \rightarrow \max_u.$$

Предположим, что в момент времени $\tau_s \in [0, T]$ верхняя граница на управление меняется на \bar{b} , т. е. для $t \in [\tau_s, T]$ управление удовлетворяет ограничению $u(t) \in [0, \bar{b}]$.

1.2 Общий подход к решению

Рассмотрим базовый случай, когда верхняя граница множества допустимых управлений не изменяется. Оптимальное управление и соответствующая ему траектория находятся с помощью принципа максимума Понтрягина.

Гамильтониан имеет вид

$$H = \left(b - \frac{u}{2}\right) u - dx(t) + \psi u,$$

где ψ – сопряженная переменная. Тогда

$$\frac{\partial H}{\partial u}(x, u, \psi) = b - u + \psi, \quad (7)$$

$$\frac{\partial^2 H}{\partial u^2}(x, u, \psi) = -1.$$

Это значит, что гамильтониан H выпуклый, следовательно, достигает своего максимума. Приравнивая равенство (7) к нулю, получаем выражение для оптимального управления

$$u^*(t) = b + \psi(t). \quad (8)$$

Каноническая система имеет вид

$$\begin{cases} \dot{x} = u^*, \\ \dot{\psi} = d. \end{cases} \quad (9)$$

Поскольку терминальное состояние $x(T)$ свободно, выполняется условие $\psi(T) = 0$. Решая второе уравнение системы (9), получаем $\psi = d(t - T)$, а следовательно, оптимальное управление (8) выражается как $u^*(t) = b + d(t - T)$. Очевидно, что оптимальное управление монотонно возрастает как функция от времени, принимая максимальное значение в точке $t = T$, $u^*(T) = b$, а для $t < T$, $u^*(t) < b$. С учетом нижнего ограничения окончательное выражение для оптимального управления принимает вид

$$u^*(t) = \begin{cases} 0, & t \in [0, T - \frac{b}{d}], \\ b + d(t - T), & t \in [T - \frac{b}{d}, T]. \end{cases}$$

Посчитаем, соответствующую этим управлениям траекторию по формуле (6):

$$\dot{x}(t) = b + d(t - T).$$

Получим составное уравнение оптимальной траектории:

$$x^*(t) = \begin{cases} x_0, & t \in [0, T - \frac{b}{d}], \\ (b - dT)t + \frac{dt^2}{2} + C, & t \in [T - \frac{b}{d}, T]. \end{cases}$$

где $C = x_0 - (b - dT)(T - \frac{b}{d}) - \frac{d}{2}(T - \frac{b}{d})^2$.

На рис. 1 изображены оптимальные управление и траектория для значений параметров $T = 30, b = 20, d = 1, x_0 = 10$.

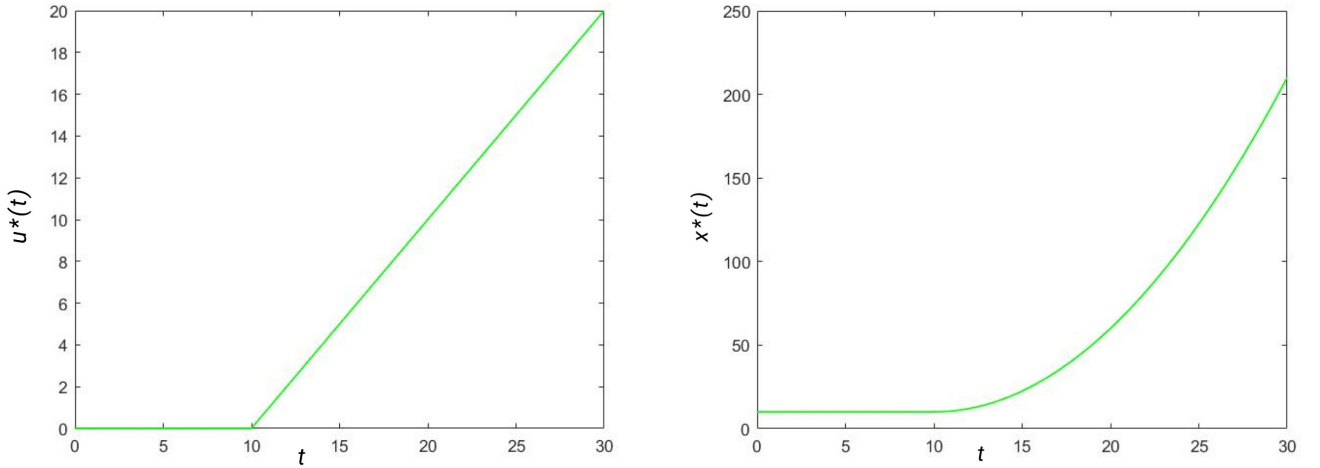


Рис. 1: Оптимальное управление $u^*(t)$ и траектория $x^*(t)$

1.3 Случай со сменой верхнего предела

Поскольку оптимальное управление не превышает b для любых значений $t \in [0, T]$, увеличение верхней границы уровня выбросов $b \rightarrow \bar{b}$ не

будет приводить к изменению оптимального управления. Однако, если на втором интервале ограничение на управление будет меньше, чем на первом, т. е. $(\bar{b} < b)$, то управление примет вид

$$u^*(t) = \begin{cases} b + d(t - T), & t \in [0, \theta), \\ \bar{b}, & t \in [\theta, T], \end{cases}$$

где $\theta \geq \tau_s$ определяется из условия $b + d(\theta - T) = \bar{b}$. Откуда имеем $\theta = T + \frac{\bar{b}-b}{d}$.

Положим, что $b + d(\tau_s - T) > 0$, т. е. на интервале $[\tau_s, T]$ оптимальное управление не достигает нижней границы. Рассмотрим две возможные ситуации.

1.3.1 Владелец предприятия не знает о грядущем изменении

При таком исходе его поведение до момента смены условий ничем не будет отличаться от случая без изменения верхнего предела. Воспользовавшись этой логикой, можно найти составное уравнение оптимального управления. В конце первого участка в точке τ_s мы получим значение управления либо больше, чем новая граница, либо меньше. Исходя из этого, мы введем новую переменную ζ . Если значение $u^*(\tau_s) < \bar{b}$, то $\zeta = \theta = T + \frac{\bar{b}-b}{d}$. В противном случае, $\zeta = \tau_s$.

Таким образом оптимальное управление примет вид

$$u_1^*(t) = \begin{cases} 0, & t \in [0, T - \frac{b}{d}], \\ b + d(t - T), & t \in (T - \frac{b}{d}, \zeta], \\ \bar{b}, & t \in (\zeta, T]. \end{cases} \quad (10)$$

Траектория, соответствующая этому управлению, изображена на рис.

2. Выражение для оптимальной траектории будет выглядеть следующим образом

$$x_1^*(t) = \begin{cases} x_0, & t \in [0, T - \frac{b}{d}], \\ (b - dT)t + \frac{d}{2}t^2 + C_1, & t \in [T - \frac{b}{d}, \zeta], \\ \bar{b}t + C_2, & t \in (\zeta, T], \end{cases}$$

где $C_1 = x_0 - bT + \frac{b^2}{2d} + \frac{dT^2}{2}$ и $C_2 = C_1 + (b - dT)\zeta + \frac{d}{2}\zeta^2 - \bar{b}\zeta$.

Возьмем набор переменных $T = 30$, $b = 20$, $d = 1$, $x_0 = 10$, $\bar{b} = 10$, $\tau_s = 25$ и построим график. На рис. 2 видно, что поскольку в момент τ_s управление $u^*(\tau_s) > \bar{b}$, то управление претерпевает скачок, чтобы соответствовать установленному ограничению.

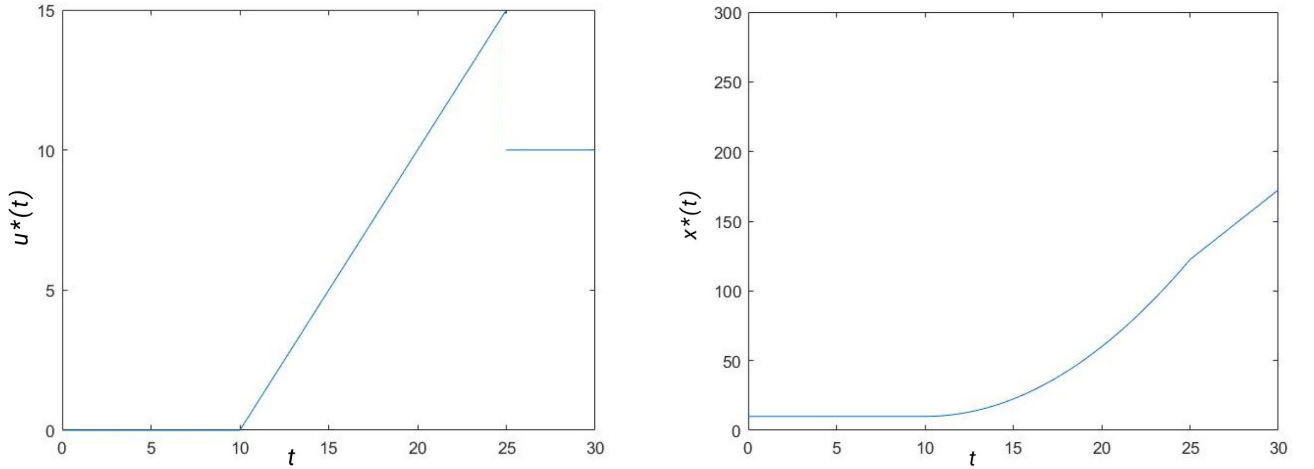


Рис. 2: Оптимальное управление $u_1^*(t)$ и траектория $x_1^*(t)$

1.3.2 Владелец предприятия осведомлён об изменении

Для вычисления оптимального управления необходимо рассмотреть задачу оптимального управления отдельно на двух интервалах, начиная со второго. На интервале $[\tau_s, T]$ определим оптимальное значение функционала

$$J_2(x, u) = \int_{\tau_s}^T \left(\left(b - \frac{1}{2}u \right) u - dx \right) dt,$$

где $u(t)$ определяется как сужение (10) на исследуемый интервал, а $x(t)$ находится из решения (6) с начальным условием $x(\tau_s) = x_s$.

Получим значение выигрыша игрока на втором интервале J_2^* , которое при фиксированном τ_s будет зависеть только от x_s . Теперь рассмотрим первый интервал $[0, \tau_s]$. В момент времени τ_s конец траектории не будет фиксирован, но так как необходимо учитывать выигрыш на втором участке, рассмотрим общий выигрыш в смешанной форме (интегральный и терминальный в момент времени τ_s):

$$J(x_0, T, u) = \int_0^{\tau_s} \left(\left(b - \frac{1}{2}u \right) u - dx \right) dt + J_2^*(x_s) \rightarrow \max.$$

Для того, чтобы записать сопряженные уравнения и условия трансверсальности, найдем производную терминального выигрыша

$$\frac{\partial J_2^*}{\partial x_s} = 2(b - \bar{b}) - 2d(T - \tau_s).$$

Используя условие трансверсальности в задаче с терминальным выигрышем, получаем значение ψ на конце первого интервала: $\psi(\tau_s) = 2(b - \bar{b}) - 2d(T - \tau_s)$.

Составное уравнение управления в этом случае запишется следующим образом:

$$u_2^*(t) = \begin{cases} 0, & t \in [0, \theta - \frac{\bar{b}}{d}], \\ 3b - 2\bar{b} + d(t - 2T + \tau_s), & t \in (\theta - \frac{\bar{b}}{d}, \theta], \\ \bar{b}, & t \in (\theta, T], \end{cases}$$

где $\theta = 2T - \tau_s + \frac{3\bar{b}-3b}{d}$.

Решая (6), найдем оптимальную траекторию

$$x_2^*(t) = \begin{cases} x_0, & t \in [0, \theta - \frac{\bar{b}}{d}], \\ (3b - 2(\bar{b} + dT) + d\tau_s)t + \frac{d}{2}t^2 + C_3, & t \in (\theta - \frac{\bar{b}}{d}, \theta], \\ \bar{b}t + C_4, & t \in (\theta, T]. \end{cases}$$

Подводя итоги, возьмем набор переменных $T = 30, b = 20, d = 1, x_0 = 10, \bar{b} = 10, \tau_s = 15$ и совместим графики двух случаев, чтобы наглядно увидеть разницу. На рис. 3 пунктирной линией отмечен график случая, в котором грядущие изменения были ожидаемы, а сплошной, соответственно, график случая, когда неизвестно об изменении верхней границы.

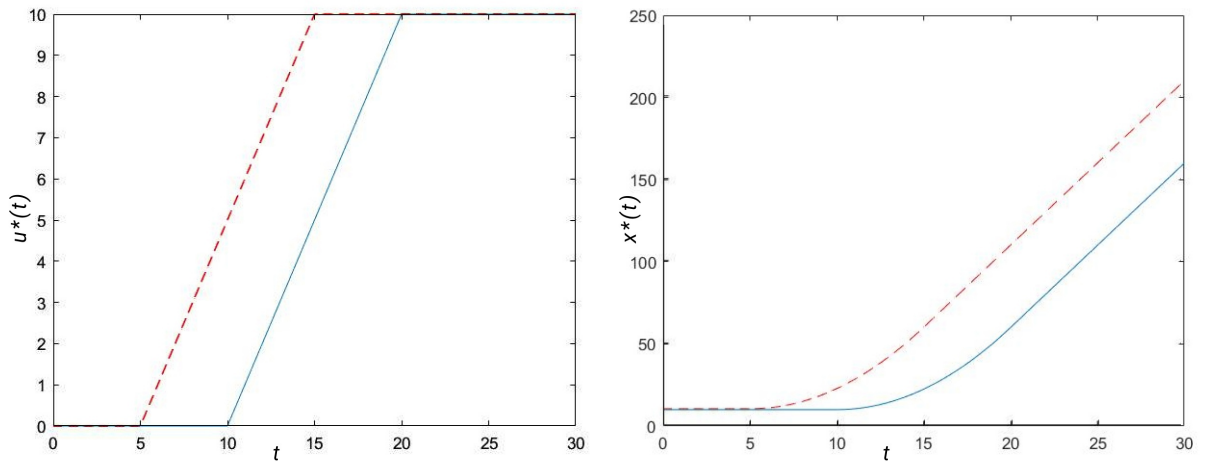


Рис. 3: Сравнение оптимальных управлений $u_i^*(t)$ и оптимальных траекторий $x_i^*(t)$

Заметим, что в ситуации, когда заранее известно о смене верхней границы выбросов, оптимальное управление быстрее выходит на нужный максимально допустимый уровень, а траектория достигает более высокого значения, чем в ситуации, когда ориентируются только на данные, которые имеются в настоящий момент.

1.4 Оценка ценности информации

Для того, чтобы оценить ценность информации мы используем характеристику, которая будет обозначаться NVI. Эта характеристика описывает относительное увеличение выигрыша в процентах от максимального значения. Соответствующее выражение имеет вид

$$NVI = \frac{J_2^* - J_1^*}{J_2^*} \times 100\%.$$

Для полученных нами оптимальных траекторий и управлений J_1^* и J_2^* будут выглядеть следующим образом:

$$\begin{aligned} J_1^* &= \int_0^{T-\frac{b}{d}} ((b-0)0 - dx_0) dt + \int_{\zeta}^T \left((b - \frac{1}{2}\bar{b})\bar{b} - d(\bar{b}t + C_2) \right) dt + \\ &+ \int_{T-\frac{b}{d}}^{\zeta} \left((b - \frac{1}{2}(b + d(t-T)))(b + d(t-T)) - d((b-dT)t + \frac{d}{2}t^2 + C_1) \right) dt = \\ &= -dx_0 \left(T - \frac{b}{d} \right) - \frac{d^2}{3} \left(\zeta^3 - \left(T - \frac{b}{d} \right)^3 \right) + \left(d^2T - \frac{bd}{2} \right) \left(\zeta^2 - \left(T - \frac{b}{d} \right)^2 \right) + \\ &+ \left(\frac{b^2}{2} - \frac{d^2T^2}{2} - dC_1 \right) \left(\zeta - T + \frac{b}{d} \right) + \left(b\bar{b} - \frac{\bar{b}^2}{2} - dC_2 \right) (T-\zeta) - \frac{d\bar{b}}{2}(T^2-\zeta^2). \end{aligned}$$

$$\begin{aligned}
J_2^* &= \int_0^{\theta - \frac{\bar{b}}{d}} ((b - 0)0 - dx_0) dt + \int_{\theta}^T \left((b - \frac{1}{2}\bar{b})\bar{b} - d(\bar{b}t + C_4) \right) dt + \\
&+ \int_{\theta - \frac{\bar{b}}{d}}^{\theta} \left((b - \frac{1}{2}(3b - 2\bar{b} + d(t - 2T + \tau_s))) (3b - 2\bar{b} + d(t - 2T + \tau_s)) - \right. \\
&\quad \left. - d((3b - 2(\bar{b} + dT) + d\tau_s)t + \frac{d}{2}t^2 + C_3) \right) dt = \\
&= -dx_0 \left(\theta - \frac{\bar{b}}{d} \right) - \frac{d^2}{3} \left(\theta^3 - \left(\theta - \frac{\bar{b}}{d} \right)^3 \right) + \\
&\quad + \left(2\bar{b}d - \frac{5bd}{2} + 2d^2T - d^2\tau_s \right) \left(\theta^2 - \left(\theta - \frac{\bar{b}}{d} \right)^2 \right) + \\
&+ \frac{\bar{b}}{d} \left(-\frac{3b^2}{2} + 4b\bar{b} - 2bd\tau_s + 4bdT - 2\bar{b}^2 + 2\bar{b}d\tau_s - 4\bar{b}dT - \frac{d^2\tau_s^2}{2} + 2d^2T\tau_s - 2d^2T^2 - dC_3 \right) + \\
&\quad + \left(b\bar{b} - \frac{\bar{b}^2}{2} - dC_4 \right) (T - \theta) - \frac{d\bar{b}}{2} (T^2 - \theta^2).
\end{aligned}$$

Для параметров, рассмотренных выше, характеристика ценности информации равна 48.3871%, что указывает на то, что в текущих обстоятельствах значение информации довольно важно. Неосведомленность о корректировке верхнего предела может сильно повлиять на бизнес, сократив почти до половины теоретическую прибыль, которая может быть получена, если информация известна.

Глава 2. Модель с выигрышем в степенном виде

2.1 Постановка задачи

Пусть имеется некоторое промышленное производство. Предполагается, что объем производства пропорционален загрязнениям. Количество выбросов в атмосферу, в свою очередь, ограничено сверху. Стратегия заключается в выборе управления для максимизации выигрыша.

Динамика изменения объема загрязнений описывается следующим дифференциальным уравнением:

$$\dot{x} = u, \quad u \in [0, b], \quad x(0) = x_0. \quad (11)$$

Интегральный выигрыш игрока имеет вид

$$J(x_0, t_0, T, u(t)) = \int_{t_0}^T \left(b\sqrt{u(t)} - dx(t) \right) dt, \quad d > 0.$$

Предположим, как и в первой главе, что в момент времени $\tau_s \in [0, T]$ верхняя граница на управление меняется на \bar{b} , т. е. для $t \in [\tau_s, T]$ управление удовлетворяет ограничению $u(t) \in [0, \bar{b}]$.

2.2 Общий подход к решению

Рассмотрим базовый случай, когда верхняя граница множества допустимых управлений не изменяется. Оптимальное управление и соответствующая ему траектория находятся с помощью принципа максимума Понтрягина.

Гамильтониан имеет вид

$$H = \psi(t)u(t) + b\sqrt{u(t)} - dx(t).$$

где ψ – сопряженная переменная. Тогда

$$\frac{\partial H}{\partial u}(x, u, \psi) = \psi + \frac{b}{2\sqrt{u}}, \quad (12)$$

$$\frac{\partial^2 H}{\partial u^2}(x, u, \psi) = -\frac{b}{4\sqrt{u^3}}.$$

Это значит, что гамильтониан H выпуклый, следовательно, достигает своего максимума. Приравнивая равенство (12) к нулю, получаем выражение для оптимального управления

$$u^* = \frac{b^2}{4\psi(t)^2}. \quad (13)$$

Каноническая система имеет такой же вид, как и в предыдущей главе.

Поскольку терминальное состояние $x(T)$ свободно, выполняется условие $\psi(T) = 0$. Решая второе уравнение системы (9), получаем

$$\psi(t) = d(t - T),$$

а следовательно, оптимальное управление (13) выражается как $u^* = \frac{b^2}{4d^2(t-T)^2}$.

Теперь нужно найти управления, которые удовлетворяют начальному условию $u^*(t) \in [0, b]$.

Рассмотрим неравенство:

$$0 \leq \frac{b^2}{4d^2(t - T)^2} \leq b.$$

Видно, что оптимальное управление явно не меньше 0, поэтому возьмём только вторую часть неравенства

$$\frac{b^2}{4d^2(t - T)^2} \leq b.$$

Таким образом получим составное уравнение оптимального управления

$$u^*(t) = \begin{cases} \frac{b^2}{4d^2(t-T)^2}, & t \in [0, T - \frac{\sqrt{b}}{2d}], \\ b, & t \in (T - \frac{\sqrt{b}}{2d}, T]. \end{cases}$$

Посчитаем, соответствующую этим управлениям траекторию по формуле (11):

$$\dot{x}(t) = \frac{b^2}{4d^2(t-T)^2}.$$

Получим составное уравнение оптимальной траектории:

$$x^*(t) = \begin{cases} -\frac{b^2}{4d^2(t-T)} + C_1, & t \in [0, T - \frac{\sqrt{b}}{2d}], \\ bt + C_2, & t \in (T - \frac{\sqrt{b}}{2d}, T]. \end{cases}$$

где $C_1 = x_0 - \frac{b^2}{4d^2T}$ и $C_2 = C_1 - bT + \frac{b\sqrt{b}}{d}$.

На рис. 4 показан график для значений параметров $T = 30$, $b = 100$, $d = 1$, $x_0 = 100$.

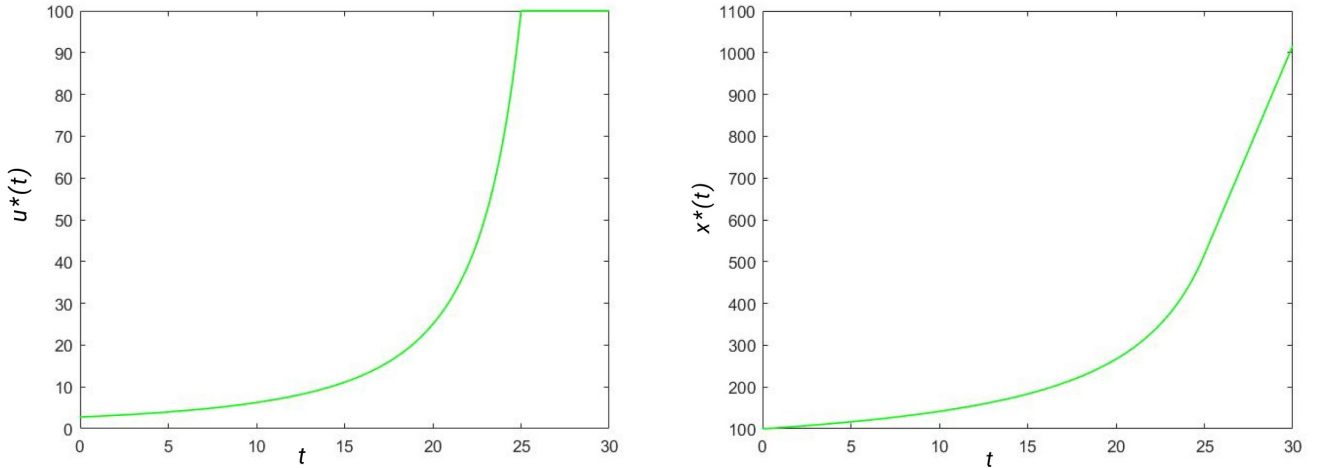


Рис. 4: Оптимальное управление $u^*(t)$ и траектория $x^*(t)$

2.3 Случай со сменой верхнего предела ($\bar{b} < b$)

Предположим, что в момент времени $\tau_s \in [0, T]$ верхняя граница на управление меняется на \bar{b} , т.е. для $t \in [\tau_s, T]$ управление удовлетворяет ограничению $u(t) \in [0, \bar{b}]$.

2.3.1 Владелец предприятия не знает о грядущем изменении

Тогда нужно рассмотреть два варианта:

1. $\tau_s \leq T - \frac{\sqrt{b}}{2d}$;
2. $\tau_s > T - \frac{\sqrt{b}}{2d}$.

Первый вариант предполагает, что владелец узнает об изменениях на этапе, когда управление еще не достигло своей максимальной точки b . Следовательно мы получаем уравнение оптимального управления состоящие из двух частей.

Таким образом оптимальное управление запишется в виде:

$$u_1^*(t) = \begin{cases} \frac{b^2}{4d^2(t-T)^2}, & t \in [0, \theta], \\ \bar{b}, & t \in (\theta, T]. \end{cases}$$

где $\theta \geq \tau_s$ определяется из условия $\frac{b^2}{4d^2(\theta-T)^2} = \bar{b}$, откуда имеем $\theta = T - \frac{b}{2d\sqrt{\bar{b}}}$ при $u_1^*(\tau_s) < \bar{b}$, в противном случае $\theta = \tau_s$.

Выражение для оптимальной траектории имеет вид:

$$x_1^*(t) = \begin{cases} -\frac{b^2}{4d^2(t-T)} + C_1, & t \in [0, \theta], \\ \bar{b}t + C_2, & t \in (\theta, T]. \end{cases}$$

где $C_1 = x_0 - \frac{b^2}{4d^2T}$ и $C_2 = C_1 - \frac{b^2}{4d^2(\theta-T)} - \bar{b}\theta$.

Графики для этого варианта в случае, когда переменные принимают значения $T = 30, b = 100, d = 1, x_0 = 100, \bar{b} = 20, \tau_s = 10$ изображены на Рис. 5.

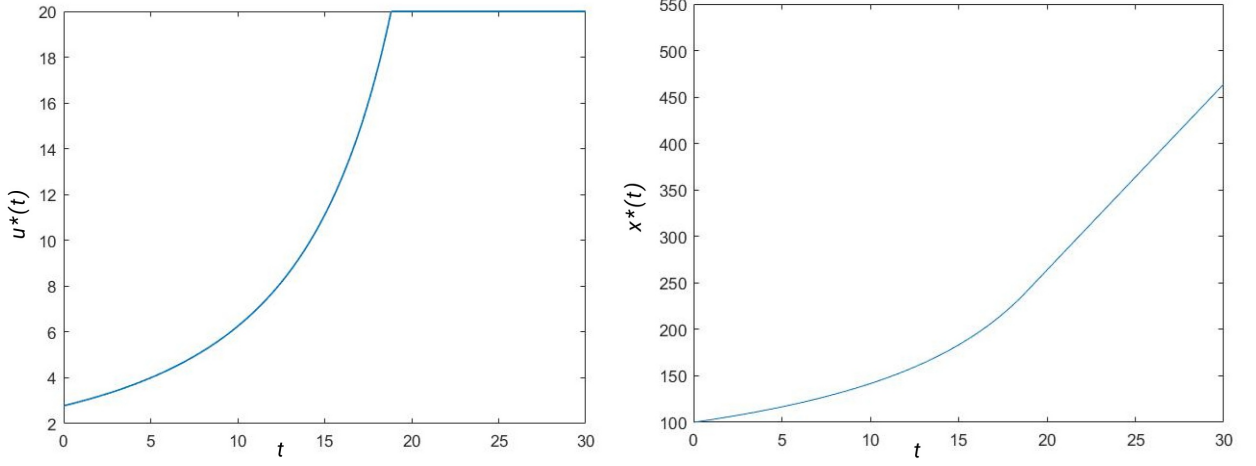


Рис. 5: Оптимальное управление $u_1^*(t)$ и траектория $x_1^*(t)$

Выигрыш в этом случае имеет вид:

$$\begin{aligned}
 J_{11}^* &= \int_0^\theta \left(b \sqrt{\frac{b^2}{4d^2(t-T)^2}} - d \left(-\frac{b^2}{4d^2(t-T)} + C_1 \right) \right) dt + \\
 &\quad + \int_\theta^T \left(b\sqrt{\bar{b}} - d(\bar{b}t + C_2) \right) dt = \\
 &= -\frac{b^2}{4d} \ln \left(\frac{T-\theta}{T} \right) - dC_1\theta + (b\sqrt{\bar{b}} - dC_2)(T-\theta) - \frac{d\bar{b}}{2}(T^2 - \theta^2).
 \end{aligned}$$

Второй вариант же означает, что изменения были введены после достижения управлением своей максимальной точки b . Следовательно мы получаем уравнение оптимального управления состоящие из трёх частей.

Таким образом оптимальное управление запишется в виде:

$$u_1^*(t) = \begin{cases} \frac{b^2}{4d^2(t-T)^2}, & t \in [0, T - \frac{\sqrt{b}}{2d}], \\ b, & t \in (T - \frac{\sqrt{b}}{2d}, \tau_s], \\ \bar{b}, & t \in (\tau_s, T]. \end{cases} \quad (14)$$

Выражение для оптимальной траектории имеет вид:

$$x_1^*(t) = \begin{cases} -\frac{b^2}{4d^2(t-T)} + C_1, & t \in [0, T - \frac{\sqrt{b}}{2d}], \\ bt + C_2, & t \in (T - \frac{\sqrt{b}}{2d}, \tau_s], \\ \bar{b}t + C_3, & t \in (\tau_s, T]. \end{cases} \quad (15)$$

где $C_1 = x_0 - \frac{b^2}{4d^2T}$, $C_2 = C_1 - bT + \frac{b\sqrt{b}}{d}$, а $C_3 = C_2 + (b - \bar{b})\tau_s$.

Графики для этого варианта в случае, когда переменные принимают значения $T = 30, b = 100, d = 1, x_0 = 100, \bar{b} = 20, \tau_s = 27$ изображены на Рис. 6.

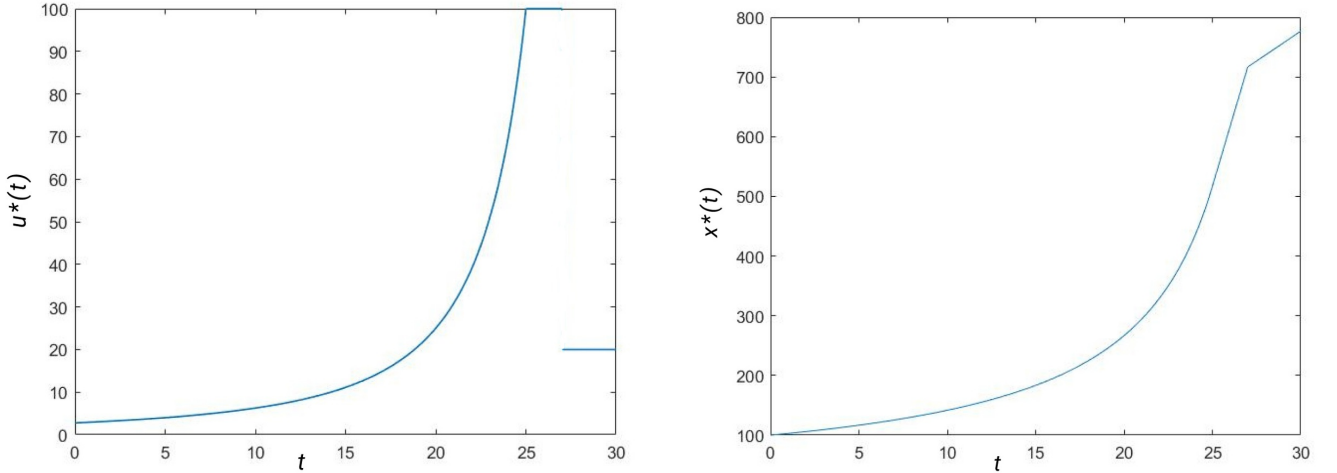


Рис. 6: Оптимальное управление $u_1^*(t)$ и траектория $x_1^*(t)$

Выпишем выигрыш для этого случая.

$$\begin{aligned}
J_{12}^* &= \int_0^{T-\frac{\sqrt{b}}{2d}} \left(b\sqrt{\frac{b^2}{4d^2(t-T)^2}} - d \left(-\frac{b^2}{4d^2(t-T)} + C_1 \right) \right) dt + \\
&+ \int_{T-\frac{\sqrt{b}}{2d}}^{\tau_s} \left(b\sqrt{b} - d(bt + C_2) \right) dt + \int_{\tau_s}^T \left(b\sqrt{\bar{b}} - d(\bar{b}t + C_3) \right) dt = \\
&= -\frac{b^2}{4d} \ln \frac{\sqrt{b}}{2dT} - dC_1 \left(T - \frac{\sqrt{b}}{2d} \right) + (b\sqrt{b} - dC_2) \left(\tau_s - \left(T - \frac{\sqrt{b}}{2d} \right) \right) - \\
&\quad - \frac{db}{2} \left(\tau_s^2 - \left(T - \frac{\sqrt{b}}{2d} \right)^2 \right) + \\
&\quad + (b\sqrt{\bar{b}} - dC_3)(T - \tau_s) - \frac{d\bar{b}}{2}(T^2 - \tau_s^2).
\end{aligned}$$

2.3.2 Владелец предприятия осведомлён об изменении

Для вычисления оптимального управления необходимо рассмотреть задачу оптимального управления отдельно на двух интервалах, начиная со второго. На интервале $[\tau_s, T]$ определим оптимальное значение функционала

$$J_2(x, u) = \int_{\tau_s}^T \left(b\sqrt{u(t)} - dx(t) \right) dt,$$

где $x(t)$ находится из решения (11) с начальным условием $x(\tau_s) = x_s$.

Получим значение выигрыша игрока на втором интервале J_2^* , которое при фиксированном τ_s будет зависеть только от x_s . Теперь рассмотрим первый интервал $[0, \tau_s]$. В момент времени τ_s конец траектории не будет фиксирован, но так как необходимо учитывать выигрыш на втором участке, мы будем рассматривать общий выигрыш в смешанной форме

(интегральный и терминальный в момент времени τ_s):

$$J(x_0, T, u) = \int_0^{\tau_s} \left(b\sqrt{u(t)} - dx(t) \right) dt + J_2^*(x_s) \rightarrow \max.$$

Для того, чтобы записать сопряженные уравнения и условия трансверсальности найдем производную терминального выигрыша по x_s :

$$\frac{\partial J_2^*}{\partial x_s} = \frac{b}{\sqrt{\bar{b}}} - 2d(T - \tau_s), \quad (16)$$

Используя условие трансверсальности в задаче с терминальным выигрышем, получаем значение ψ на конце первого интервала:

$$\psi(\tau_s) = \frac{b}{\sqrt{\bar{b}}} - 2d(T - \tau_s).$$

Составное уравнение управления в этом случае запишется следующим образом:

$$u_2^*(t) = \begin{cases} \frac{b^2\bar{b}}{4(b - d\sqrt{\bar{b}}(2T - \tau_s - t))^2}, & t \in [0, \zeta], \\ \bar{b}, & t \in (\zeta, T]. \end{cases} \quad (17)$$

где $\zeta = 2T - \tau_s - \frac{3b}{2d\sqrt{\bar{b}}}$.

Выражение для оптимальной траектории имеет вид:

$$x_2^*(t) = \begin{cases} \frac{b^2\sqrt{\bar{b}}}{4d(d\sqrt{\bar{b}}(2T - \tau_s - t) - b)} + C_4, & t \in [0, \zeta], \\ \bar{b}t + C_5, & t \in (\zeta, T]. \end{cases} \quad (18)$$

где $C_4 = x_0 - \frac{b^2\sqrt{\bar{b}}}{4d(d\sqrt{\bar{b}}(2T - \tau_s) - b)}$, $C_5 = C_4 + \frac{b^2\sqrt{\bar{b}}}{4d(d\sqrt{\bar{b}}(2T - \tau_s - \zeta) - b)} - \bar{b}\zeta$.

Соответственно, выражение для J_2^* записывается в виде

$$\begin{aligned}
 J_2^* &= \int_0^{\zeta} b \sqrt{\frac{b^2 \bar{b}}{4(b - d\sqrt{\bar{b}}(2T - \tau_s - t))^2}} - d \left(\frac{b^2 \sqrt{\bar{b}}}{4d(d\sqrt{\bar{b}}(2T - \tau_s - t) - b)} + C_4 \right) dt + \\
 &\quad + \int_{\zeta}^T (b\sqrt{\bar{b}} - d(\bar{b}t + C_5)) dt = \\
 &= -\frac{b^2}{4d} \ln \left(\frac{d\sqrt{\bar{b}}(\zeta + \tau_s - 2T) + b}{d\sqrt{\bar{b}}(\tau_s - 2T) + b} \right) - dC_4\zeta + (b\sqrt{\bar{b}} - dC_5)(T - \zeta) - \frac{d\bar{b}}{2}(T^2 - \zeta^2).
 \end{aligned}$$

Совместим графики двух случаев с набором переменных $T = 30$, $b = 100$, $d = 1$, $x_0 = 100$, $\bar{b} = 20$, $\tau_s = 10$. На рис. 7 пунктирной линией отмечен график случая, в котором грядущие изменения были ожидаемы, а сплошной, соответственно, график случая, когда неизвестно об изменении верхней границы.

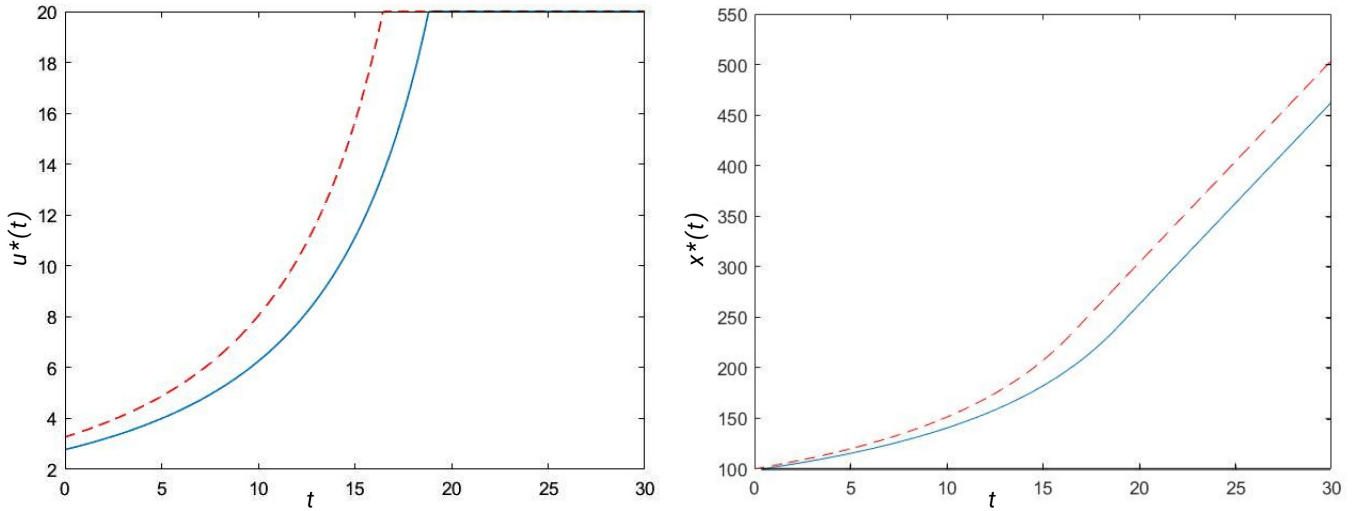


Рис. 7: Сравнение оптимальных управлений $u_i^*(t)$ и оптимальных траекторий $x_i^*(t)$

Как и в предыдущей главе, видно, что траектория достигает более высокого уровня в случае, когда производитель осведомлен об изменении границы.

2.3.3 Оценка ценности информации

Посчитаем характеристику NVI для данной модели:

$$NVI = \frac{J_2^* - J_1^*}{J_2^*} \times 100\%$$

Коэффициент равен $NVI = \frac{J_2^* - J_1^*}{J_2^*} \times 100\% = 0.5334\%$. Это показывает, что знание информации заранее практически не даст выигрыш для владельца предприятия.

2.4 Случай со сменой верхнего предела ($\bar{b} > b$)

2.4.1 Владелец предприятия не знает о грядущем изменении

Рассмотрим теперь ситуацию, в которой вышестоящее руководство ослабляет ограничения и повышает верхнюю границу. В данной ситуации будем рассматривать 2 возможных случая:

1. $\tau_s \leq T - \frac{\sqrt{b}}{2d}$;
2. $\tau_s > T - \frac{\sqrt{b}}{2d}$.

Причем управление и траектория для первого пункта будут совпадать с уравнениями (14), (15). Соответственно выигрыш будет равен выигрышу, посчитанному ранее.

Теперь обратим внимание на второй вариант. При этом условии собственник узнает о смене границы после достижения ранее допустимого верхнего предела, а значит ему нужно будет увеличить темпы производства.

Управление для данной ситуации примет вид:

$$u_1^*(t) = \begin{cases} \frac{b^2}{4d^2(t-T)^2}, & t \in [0, T - \frac{\sqrt{b}}{2d}], \\ b, & t \in (T - \frac{\sqrt{b}}{2d}, \tau_s], \\ \frac{b^2}{4d^2(t-T)^2} + K, & t \in [\tau_s, \nu], \\ \bar{b}, & t \in (\nu, T]. \end{cases}$$

где $K = b - \frac{b^2}{4d^2(\tau_s - T)^2}$, а $\nu \geq \tau_s$ определяется из условия $\frac{b^2}{4d^2(\nu - T)^2} + K = \bar{b}$.

Откуда имеем $\nu = T - \frac{b}{2d\sqrt{\bar{b} - K}}$.

Выражение для оптимальной траектории имеет вид:

$$x_1^*(t) = \begin{cases} -\frac{b^2}{4d^2(t-T)} + C_1, & t \in [0, T - \frac{\sqrt{b}}{2d}], \\ bt + C_2, & t \in (T - \frac{\sqrt{b}}{2d}, \tau_s], \\ -\frac{b^2}{4d^2(t-T)} + Kt + C_3, & t \in [\tau_s, \nu], \\ \bar{b}t + C_4, & t \in (\nu, T]. \end{cases}$$

где $C_1 = x_0 - \frac{b^2}{4d^2T}$, $C_2 = C_1 - bT + \frac{b\sqrt{b}}{d}$, $C_3 = C_2 + b\tau_s + \frac{b^2}{4d^2(\tau_s - T)} - K\tau_s$,
а $C_4 = C_3 - \frac{b^2}{4d^2(\nu - T)} + K\nu - \bar{b}\nu$.

Выпишем выигрыш для этого случая.

$$J_{13}^* = \int_0^{T - \frac{\sqrt{b}}{2d}} \left(b\sqrt{\frac{b^2}{4d^2(t-T)^2}} - d \left(-\frac{b^2}{4d^2(t-T)} + C_1 \right) \right) dt + \\ + \int_{T - \frac{\sqrt{b}}{2d}}^{\tau_s} \left(b\sqrt{b} - d(bt + C_2) \right) dt + \int_{\nu}^T \left(b\sqrt{\bar{b}} - d(\bar{b}t + C_4) \right) dt +$$

$$\begin{aligned}
& + \int_{\tau_s}^{\nu} \left(b \sqrt{\frac{b^2}{4d^2(t-T)^2} + K} - d \left(-\frac{b^2}{4d^2(t-T)} + Kt + C_3 \right) \right) dt = \\
& = -\frac{b^2}{4d} \ln \left(\frac{\sqrt{b}}{2dT} \right) - dC_1 \left(T - \frac{\sqrt{b}}{2d} \right) + (b\sqrt{b} - dC_2) \left(\tau_s - \left(T - \frac{\sqrt{b}}{2d} \right) \right) - \\
& - \frac{db}{2} \left(\tau_s^2 - \left(T - \frac{\sqrt{b}}{2d} \right)^2 \right) + \frac{b^2}{4d} \ln \left(\frac{\frac{2b}{\sqrt{4d^2K(T-\nu)^2+b^2-b}} + 1}{\frac{2b}{\sqrt{4d^2K(T-\tau_s)^2+b^2-b}} + 1} \right) - \\
& - \frac{b}{2d} \left(\sqrt{4d^2K(T-\nu)^2+b^2} - \sqrt{4d^2K(T-\tau_s)^2+b^2} \right) + \\
& + \frac{b^2}{4d} \ln \left(\frac{T-\nu}{T-\tau_s} \right) - \frac{dK}{2} (\nu^2 - \tau_s^2) - dC_3(\nu - \tau_s) + \\
& + (b\sqrt{b} - dC_4)(T - \nu) - \frac{d\bar{b}}{2}(T^2 - \nu^2).
\end{aligned}$$

Возьмем набор переменных $T = 30$, $b = 100$, $d = 1$, $x_0 = 100$, $\bar{b} = 150$, $\tau_s = 26$ и построим график.

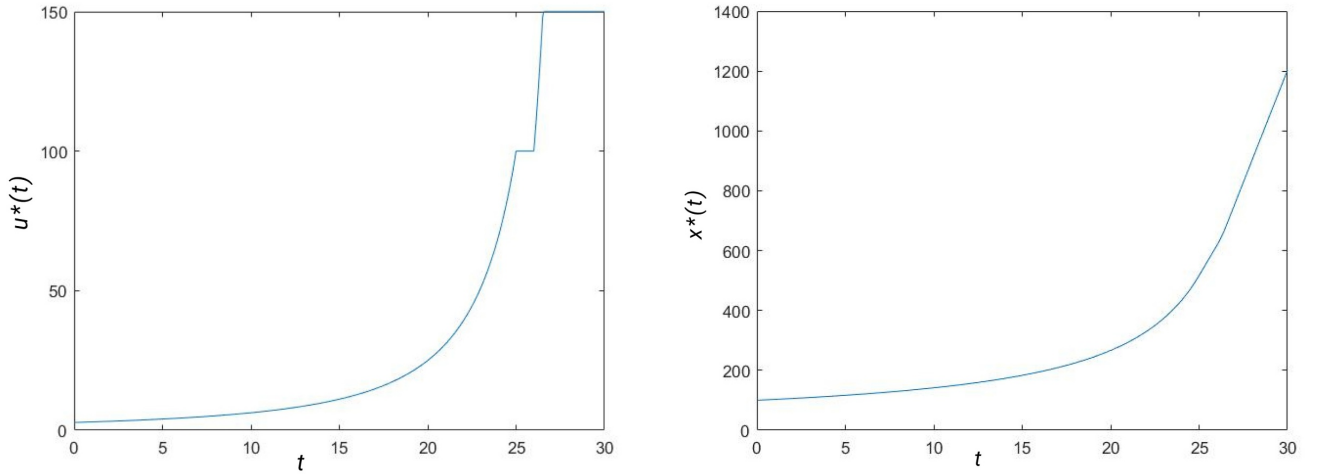


Рис. 8: Оптимальное управление $u_1^*(t)$ и траектория $x_1^*(t)$

2.4.2 Владелец предприятия осведомлён об изменении

Применим тот же подход, что и в пункте **2.3.2** и будем считать, что производная терминального выигрыша в данных обстоятельствах полностью совпадает с найденной ранее и имеет вид 16, а следовательно оптимальные управление и траектория будут представленными составными уравнениями (17), (18). И посмотрим на полученные нами графики и коэффициент информации.

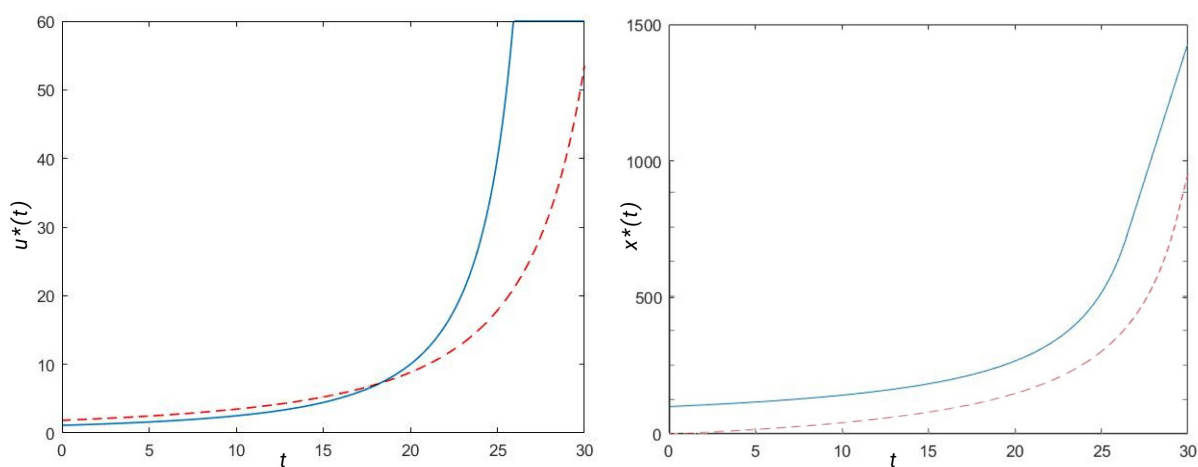


Рис. 9: Оптимальные управления $u_i^*(t)$ и траектории $x_i^*(t)$

2.4.3 Оценка ценности информации

Посчитаем характеристику NVI для данной модели:

$$NVI = \frac{J_2^* - J_1^*}{J_2^*} \times 100\%$$

Коэффициент NVI равен нулю. Это значит, что в данных обстоятельствах осведомленность не принесет никакой выгоды, что также видно на рис. 9, на котором синей линией отмечен график случая с неожиданными изменениями, а красным пунктиром – график, для случая когда изменения были ожидаемы.

Выводы

В работе на примере модели работы производства с выигрышами в разных формах было продемонстрировано решение с помощью подхода, основанного на расширении классического принципа максимума—гибридном принципе максимума. Несмотря на громоздкость вычислений терминальных выигрышей и их производных на втором участке, этот метод показывает хорошие результаты, которые согласуются с интуицией.

Рассмотрев функции выигрыша у обеих моделей, а также характеристику ценности информации, можно заключить, что параметры, характеризующие задачу оптимального управления, а также ее структура очень сильно влияют на ценность информации в рассматриваемой задаче. Так, для одной структуры задачи коэффициент NVI принимает большие значения, тогда как для другой формулировки оказывается пренебрежительно малым.

Дальнейшее исследование может заключаться в рассмотрении задачи с несколькими сменами режимов, а так же в изучении зависимости NVI от параметров.

Заключение

В данной работе была описана модель производственного процесса, которая может быть использована для решения задач оптимизации производственного процесса, с учётом экологической обстановки.

В ходе проделанной работы были найдены оптимальные управления для двух форм выигрышей. Для задачи со сменой верхнего уровня выбросов был рассмотрен подход, основанный на расширении классического принципа максимума Понтрягина–гибридный принцип максимума.

Список литературы

- [1] Айзекс Р. Дифференциальные игры. М.: Мир, 1967. 480 с.
- [2] Петросян Л. А. Дифференциальные игры преследования. Л.: Изд-во Ленингр. ун-та, 1977.
- [3] Красовский Н. Н. Игровые задачи о встрече движений. М.: Наука, 1970. 420 с.
- [4] Красовский Н. Н., Субботин А. И. Позиционные дифференциальные игры. М.: ФИЗМАТЛИТ, 1974. 456 с.
- [5] Понтрягин Л. С. К теории дифференциальных игр // Успехи математических наук, 21, 4 (130), 1966. С. 219—274.
- [6] Понтрягин Л. В. Математическая теория оптимальных процессов. Москва. 1961. С. 392.
- [7] Петросян Л. А., Данилов Н.Н. Кооперативные дифференциальные игры и их приложения. Иркутск: Издательство Иркутского университета. 1985. С. 276.
- [8] Петросян Л. А., Зенкевич Н. А., Шевкопляс Е. В. Теория игр. СПб.: БХВ-Петербург, 2012.
- [9] Gromova E., and Petrosian O. Control of information horizon for cooperative differential game of pollution control. In: International Conference - Stability and Oscillations of Nonlinear Control Systems (Pyatnitskiy's Conference). 2016. P. 1–4, doi: 10.1109/STAB.2016.7541187.
- [10] Bondarev A., Gromov D. On the structure and regularity of optimal solutions in a differential game with regime switching and spillovers.

Dynamic Economic Problems with Regime Switches. *Dynamic Modeling and Econometrics in Economics and Finance*. Springer. 2021. Vol. 25. P. 187–208.

- [11] Gromov D., Gromova E. On a class of hybrid differential games. *Dynamic Games and Applications*. Springer. 2017. Vol. 7. No 2. P. 266–288.
- [12] Gromova E. The Shapley value as a sustainable cooperative solution in differential games of three players. *Recent Advances in Game Theory and Applications, Static and Dynamic Game Theory: Foundations and Applications*. 2016. P. 67–91.
- [13] Bellman R. E. *Dynamic Programming*. Princeton University Press, Princeton, NJ. 1957. (reprinted 2003: Dover).
- [14] Raiffa H. and R. Schlaifer. *Applied statistical decision theory*. New York, Wiley. 1961.
- [15] Dockner E., Jorgensen S., Van Long N., Sorger G. *Differential games in economics and management science*. Cambridge University Press. 2000.
- [16] Boltyansky V. The maximum principle for variable structure systems. *International Journal of Control*. 2004. 77(17). P. 1445–1451.
- [17] Pontryagin L. S. *Mathematical theory of optimal processes*. Routledge, 2018.
- [18] Shaikh, M.S., Caines, P.E.: On the hybrid optimal control problem: Theory and algorithms. *IEEE Transactions on Automatic Control* 52(9), 1587–1603 (2007)