МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования
**«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ Н. Г. ЧЕРНЫШЕВСКОГО»**

Кафедра теории функций и стохастического анализа

ЗАДАЧА О ЗАМЕНЕ ОБОРУДОВАНИЯ

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

студента 4 курса 412 группы
направления 01.03.02 — Прикладная математика и информатика

механико-математического факультета
Куцевалова Данилы Вячеславовича

Научный руководитель
доцент к. ф.-м. н., доцент ______________ И. А. Кузнецова

Заведующий кафедрой
д. ф.-м. н., доцент ______________ С. П. Сидоров

Саратов 2023

Введение. Актуальность темы работы. Многие предприятия по всему миру используют машинное оборудование при производстве товаров. Однако оно со временем стареет и теряет производительность, что приводит к увеличению расходов на текущий ремонт. Это заставляет предприятия заменять старое оборудование. Оптимальный план замены можно составить с помощью методов решения задачи о замене оборудования, что позволит максимизировать прибыль или минимизировать затраты. Задача заключается в определении оптимальных сроков замены оборудования, критерием является либо прибыль, либо суммарные затраты на эксплуатацию в течение планируемого периода.

Целью данной работы является построение алгоритма для решения поставленной задачи и его реализация методами динамического программирования, для этого были поставлены и выполнены следующие задачи.

Основные задачи бакалаврской работы:

1. Изучение теории управляемых марковских процессов.
2. Моделирование задачи о замене оборудования с помощью теории управляемых марковских процессов.
3. Реализация алгоритма решения задачи о замене оборудования с использованием уравнений оптимальности.
4. Реализация разработанного алгоритма на языке программирования Python в виде программного продукта.
5. Оптимизация разработанного алгоритма и кода с помощью применения сторонних библиотек и метода оптимизации.

Практическая значимость выпускной квалификационной работы заключается в создании программы-продукта, который бы позволял решать задачу о замене оборудования.

Структура и содержание бакалаврской работы. Данная работа включает в себя 2 раздела. В первом разделе задаются компоненты конечной управляемой марковской модели и вводятся основные понятия, связанные с ними, кроме этого исследуется фундаментальное уравнение, рассматривается и обосновывается задача оптимального уравнения, после вводятся уравнения оптимальности, с помощью которых происходит вычисление оценок начальных состояний и нахождение оптимальной стратегии. Во втором разделе приводится постановка и моделирование задачи, а также её решение методами динамического программирования, кроме этого там же описаны методы оптимизации данного алгоритма.

Основное содержание работы.

Для задания управляемой марковской модели на промежутке времени $[m, n]$ необходимо определить следующие компоненты:

1. Множества $X_m, \ldots, X_t, \ldots, X_n$. $X_t(t = m, \ldots, n)$ – множество состояний в момент времени t, X_m – множество начальный состояний; X_n – множество конечных или финальных состояний,
   $X = \bigcup\limits_{t=m}^{n} X_t$ – множество всех состояний,
   $X^* = \bigcup\limits_{t=m}^{n-1} X_t$ – множество всех нефинальных состояний.
2. Множества $A_m, \ldots, A_t, \ldots, A_{n-1}$. $A_t(t = m, m+1, \ldots, n-1)$ – множество управлений в момент времени t,
   $A = \bigcup\limits_{t=m}^{n-1} A_t$ – множество всех управлений.
3. Отображение $\alpha : X' \to 2^A$ (2^A – множество всех подмножеств A), обладающее свойством $\forall x \in X_t$, $\alpha(x) \subset A_t$.
   Для любого $x \in X'\alpha(x)$ интерпретируется как множество управлений, допустимых в состоянии . Предполагаем, что при всех $x' \neq x''$,
   $\alpha(x') \cap \alpha(x'') = \varnothing$. Данное свойство позволяет определить отображение $j : A \to X$ следующим образом: $j(a) = x \iff a \in \alpha(x)$ $(a \in A, x \in X)$.
4. Отображение P, ставящее в соответствии каждому управлению $a \in A_t$ распределение вероятностей P_a на X_{t+1}. P – переходная функция. Распределения P_n будет также обозначать $P(\cdot|a)$, а значение
   $P_a(y) = P(y|a)$. $P_a(y)$ – это вероятность, применив управление $a \in A_t$ в состоянии $x = j(a)$, можно попасть на следующем шаге в состояние $y \in X_{t+1}$.
5. Функция $q : A \to R$ – текущая плата.
6. Функция $r : X_n \to R$ – финальная плата.

Все множества X_t, $t = m, \ldots, n$, и A_t, $t = m, \ldots, n-1$, считаем конечными. Если к управляемой марковской модели Z добавить распределение вероятностей μ на множестве начальных состояний X_m, то получаем управляемый марковский процесс Z_μ.

Стратегии и оценки. Путем называется следующая последовательность состояний и управлений: $l = x_m a_m \dots x_t a_t \dots x_{n-1} a_{n-1} x_n$, L – множество всех путей. Оценкой пути l называется число:

$$I(l) = \sum_{t=m}^{n} q(a_t) + r(x_n). \tag{1}$$

Вероятностью пути l называется число: $P(l) = \mu(x_m) \prod_{t=m}^{n-1} P(x_{t+1}|a_t)$.

Путем, определяемым простой стратегией φ, называется путь: $l_\varphi = x_m \varphi(x_m) \dots x_t \varphi(x_t) \dots x_{n-1} \varphi(x_{n-1}) x_n$, а его вероятность – по формуле: $P(l_\varphi) = \mu(x_m) \prod_{t=m}^{n-1} P(x_{t+1}|\varphi(x_t))$.

Историей h_t к моменту времени t называется следующая последовательность состояний и управлений: $h_t = x_m a_m \dots x_{t_1} a_{t-1} x_t$,
H_t – множество всех историй к моменту времени t.

Стратегией π называется отображение, ставящее в соответствие при любом $t = m, \dots, n-1$ каждой истории $h_t \in H_t$ распределение вероятностей $\pi(\cdot|h_t)$ на A_t, сосредоточенное на $\alpha(x_t)$, где x_t – конец истории h_t.

Если $l_\pi = x_m a_m \dots x_t a_t \dots x_{n-1} a_{n-1} x_n$ – путь, определяемый стратегией π, то его оценка по-прежнему равна: $I(l_\pi) = \sum_{t=m}^{n-1} q(a_t) + r(x_n)$,

Оценкой стратегии l_π при начальном распределении π называется число:

$$\omega(\mu, \pi) = \sum_{l_\pi} I(l_\pi) P(l_\pi), \tag{2}$$

Производные модели. Фундаментальное уравнение. Процесс управления можно представить себе как ряд последовательных шагов. Первый шаг состоит в выборе распределения вероятностей на A_t, зависящего от начального распределения. Если этот выбор сделан, то каждому начальному распределению μ и X_m соответствует распределение вероятностей μ' на X_{m+1}. Рассмотрим μ' как начальное распределение в момент $m+1$, разделим задачу максимизации на две задачи:

1. при любом начальном распределении на X_{m+1} выбрать оптимальное

поведение в последующие моменты;

2. выбрать первый шаг так, чтобы была максимальная сумма платы за этот шаг и оценки оптимального поведения в последующие моменты при начальном распределении μ'.

Производная управляемая модель Z' получается из управляемой марковской модели Z путем изъятия множеств X_m и A и всего, что с ними связано. Таким образом, в Z' X_{m+1} является множеством начальных состояний, отображение α определено на $\bigcup\limits_{t=m+1}^{n-1} X_t$ и принимает значения в 2^A $(A = \bigcup\limits_{t=m+1}^{n-1} A_t)$, история h_t в Z' имеет вид $h'_t = x_{m+1}, a_{m+1}, \ldots, x_t$ и т.п.

Если ω - оценка стратегии в исходной модели, ω' - оценка стратегии в производной модели, то ω и ω' связаны следующим соотношением, называемым фундаментальным уравнением:

$$\omega(x, \pi) = \sum_{a \in \alpha(x)} \pi(a|x)[q(a) + \omega'(P_a, \pi_a)], \tag{3}$$

где $x \in X_m$ – некоторое начальное состояние,
π – произвольная стратегия в модели Z,
P_a – распределение вероятностей на X_{m+1}, задаваемое переходной функцией и рассматриваемое в производной модели Z' как начальное состояние, т.е.:

$$P_a = P(\cdot|a), \tag{4}$$

π_a – стратегия в производной модели Z', связанная со стратегией π, равенством:

$$\pi_a(\cdot|h') = \pi(\cdot|xah'), \quad (h' \in H', \quad x = j(a)), \tag{5}$$

(h' – история в модели $Z' \cdot \pi_a$ предписывает пользоваться стратегией π, предваряя каждую историю h' предысторией $j(a)a$).

Уравнение оптимальности. Для нахождения оптимальной стратегии воспользуемся принципом оптимальности.

Принцип оптимальности заключается в следующем. Пусть при любом начальном состоянии и принятом в нем решении процесс перешел в некоторое новое состояние. Тогда, если исходная стратегия была оптимальной, то ее оставшаяся часть тоже оптимальна для процесса, начинающегося из нового состояния.

Принцип оптимальности был впервые сформулирован Беллманом. Используя этот принцип, получаем рекуррентное соотношение, выведенное ниже в формуле (7).

Пусть имеется некоторая исходная модель Z, имеющая оценку состояний ν, и производная модель Z' со своей оценкой состояний ν'. Тогда оценки ν и ν' связаны соотношением:

$$\nu(x) = \max_{a \in \alpha(x)} u(a), \quad \forall x \in X_m, \tag{6}$$

где

$$u(a) = q(a) + \nu'(P_a) = q(a) + \sum_{y \in X_{m+1}} P(y|a)\nu'(y), \quad \forall a \in A_m.$$

Из (6) следует, что оценки $\nu_m, \nu_{m+1}, \ldots, v_t, \ldots, \nu_{n-1}, \nu_n$ в моделях $Z_m, Z_{m+1}, \ldots, Z_t, \ldots, Z_{n-1}, Z_n$ связаны равенством, при $\forall x \in X_t$:

$$\nu_t(x) = \max_{a \in \alpha(x)} u_t(a), \tag{7}$$

где:

$$u_t(x) = q(a) + \nu_{t+1}(P_a) = q(a) + \sum_{y \in X_{t+1}} P(y|a)\nu_{t+1}(y), \tag{8}$$

полагаем также, что $\nu_n \equiv r$.

Запишем введенное ранее отображение в рекуррентном виде:

$$\begin{gathered} \psi_t^0 : X_t \to A_t \text{ для любого } t = m, \ldots, n-1, \\ \forall x \in X_t, \quad \psi_t^0(x) \in \alpha(x), \quad u_t(\psi_t^0(x)) = \max_{a \in \alpha(x)} u_t(a). \end{gathered} \tag{9}$$

Тогда простая стратегия ψ^0, задаваемая равенством $\psi^0(x) = \psi_t^0(x)$ при $\forall x \in X_m$ и обозначаемая $\psi^0 = \psi_m^0, \ldots, \psi_{m+1}^0, \ldots, \psi_{t+1}^0, \ldots, \psi_{n-1}^0$ является равномерно оптимальной стратегией модели Z.

Задача о замене оборудования. Есть устройство со случайным сроком службы. На каждом периоде нужно решить: заменить устройство или продолжить использовать. Вероятность поломки и доход зависят от времени службы. При замене есть расходы на новое устройство, при поломке - убытки. Цель - максимизировать математическое ожидание прибыли.

Управляемая марковская модель описывает задачу. Состояние – время работы оборудования x. Доступны два управления: c_x (сохранить) и d_x (заменить). При d_x переход в состояние 0. При c_x переход в состояние $x+1$ с вероятностью p_x и в состояние 0 с вероятностью q_x. Вероятность поломки q_x увеличивается с x, принимает значение 1 при $x = k$. x принимает значения $0, 1, 2, \ldots, k$.

На шаге t плата зависит от времени службы прибора, решения и наличия поломки. Доход при переходе из x в $x+1$ при благополучной эксплуатации оборудования обозначим как h_x, а доход за период замены оборудования как α. Для управления d_x доход равен α, а для управления c_x математическое ожидание дохода равно $p_x h_x + q_x \gamma$, где $\gamma < \alpha$. Финальную плату r_x можно положить равной нулю или любой невозрастающей функции от x.

Итак, компоненты управляемой марковской модели на отрезке времени $[0, n]$, соответствующей данной задаче, задаются следующим образом:

$$\begin{gathered}
\forall t = 0, \ldots, n, \quad X_t = \{0, 1, \ldots, x, \ldots, k\} = X, \quad A_t = \{c_x, d_x\}, \\
\forall x \in X_t, \quad \alpha(x) = \{c_x, d_x\}, \quad P(0|d_x) = 1, \quad P(0|c_x) = q_x, \\
P(x+1|c_x) = p_x, \quad (p_x = 1 - q_x), \\
q(d_x) = \alpha, \quad q(c_x) = p_x h_x + q_x \gamma,
\end{gathered}$$

r_x, $x \in X$ – финальная плата, причем справедливы неравенства

$$\begin{gathered}
h_0 \geq h_1 \geq \ldots \geq h_x \geq \ldots \geq h_k > \alpha > \gamma, \\
q_0 \leq q_1 \leq \ldots \leq q_x \leq \ldots \leq q_k = 1, \\
p_0 \geq p_1 \geq \ldots \geq p_x \geq \ldots \geq p_k = 0, \\
r_0 \geq r_1 \geq \ldots \geq r_x \geq \ldots \geq r_k.
\end{gathered} \tag{10}$$

Таким образом, управляемая марковская модель, описывающая задачу о замене оборудования, полностью задана, и для ее решения можно применить уравнение оптимальности.

Запишем уравнение оптимальности для задачи о замене оборудования:

$$\begin{gathered}
\nu_{t-1}(x) = \max[u_t(c_x), u_t(d_x)], \\
u_t(c_x) = p_x h_x + P\{0|c_x\}\nu_t(0) + P\{x+1|c_x\}\nu_t(x+1) = \\
= p_x h_x + q_x \nu_t(x_0) + p_x \nu_t(x+1), \\
\text{где } P\{0|c_x\} = q_x \text{ и } P\{x+1|c_x\} = p_x, \\
u_t(d_x) = \alpha + \nu_t(0)P\{0|d_x\} = \alpha + \nu_t(0), \text{ где } P\{0|d_x\} = 1, \\
\forall x = 0, \ldots, k, \quad t = 1, \ldots, n, \\
\text{причём } \nu_n(x) \equiv r_x.
\end{gathered} \tag{11}$$

Используя эти формулы находим оптимальную стратегию $\psi = \psi_1 \psi_2 \ldots \psi_n$, где:

$$\psi_t(x) = \begin{cases} c, \text{ при } u_1(c_x) \geq u_t(d_x), \\ d, \text{ при } u_1(c_x) < u_t(d_x). \end{cases} \tag{12}$$

Если $u_t(c_x) = u_t(d_x)$, то в качестве значения $\psi_t(x)$ годятся как c, так и d; для определенности выбираем c.

Явное вычисление ν_t и ψ_t в общем случае затруднительно, но можно дать качественное описание ответа. Естественно ожидать, что оборудование тем выгоднее в эксплуатации, чем оно новее, так что при любом t:

$$\nu_t(0) \geq \nu_t(1) \geq \ldots \geq \nu_t(K). \tag{13}$$

Реализация алгоритма динамического программирования и его оптимизация. Реализация алгоритма динамического программирования для решения задачи о замене оборудования была произведена с использованием сторонних библиотек, таких как *numpy* и *pandas*.

Исходя из того, что при решении задачи о замене оборудования применяется формула Беллмана, которая при вычислении требует рекурсивного расчёта значений функции $\nu_t(x)$ при многократном рассчёте для одних и тех же тех же t и x. Данная особенность подталкивает к самому частому методу оптимизации алгоритмов динамического программирования – мемоизации.

В программировании мемоизация – пример использования кэша при разработке программного обеспечения, в программировании сохранение результатов выполнения функций для предотвращения повторных вычислений. Перед вызовом функции проверяется, вызывалась ли функция ранее:

1. если не вызывалась, то функция вызывается, и результат её выполнения сохраняется;
2. если вызывалась, то используется сохранённый результат.

Кэш – это промежуточный буфер с быстрым доступом хранящий информацию, которая часто запрашивается.

Преимущество использования мемоизации для ускорения данного алгоритма неоспоримо и подтверждается проведёнными испытаниями:

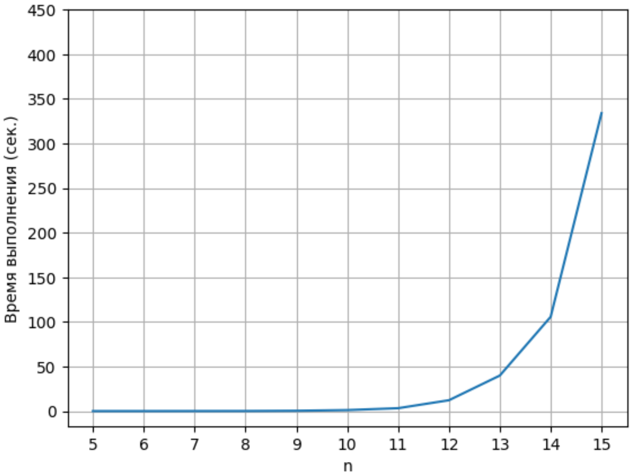

Рисунок 1 – Результат работы программы.

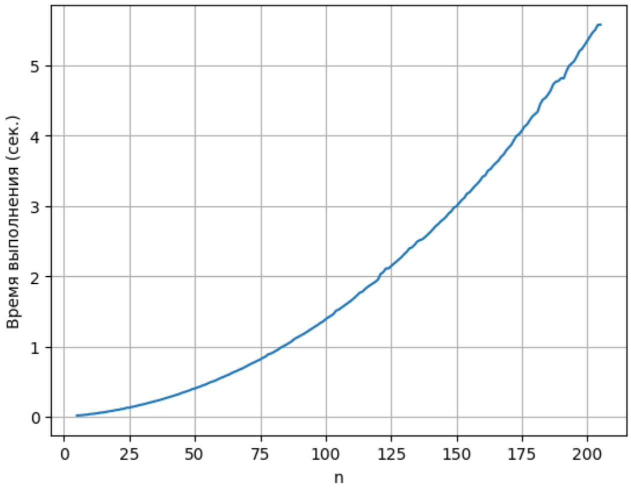

Рисунок 2 – Результат работы программы с применением мемоизации.

Отличия от обычного алгоритма динамического программирования. Под обычным алгоритмом динамического программирования для решения задачи о замене оборудования, подразумевается алгоритм, работа которого не основывается на модели управляемых марковских процессов. Из этого следует, что данный алгоритм не учитывает вероятность поломки оборудования, подразумевая что она равна 0, т.е. $q_x = 0, 0 \leq x \leq k$. Помимо этого рассчёт прибыли в подобном алгоритме производится из нескольких ценовых факторов, зависящих от срока службы оборудования — x, а именно:

1. $R(x)$ – доходность оборудования возрастом x;
2. $S(x)$ – остаточная стоимость оборудования, которая возвращается при замене;

3. $\tilde{P}$ – цена нового оборудования

Уравнение оптимальности Беллмана для такого алгоритма выглядит следующим образом:

$$F_{t-1}(x) = \max \begin{cases} u_t(c_x) = R(x) + F_t(x+1), \\ u_t(d_x) = S(x) - \tilde{P} + R(x_0) + F_t(x_0). \end{cases} \tag{14}$$

$$\text{для } \forall x = 0, \ldots, k, \quad t = 1, \ldots, n,$$

$$\text{причём } F_n(x) \equiv r = 0$$

Рассмотрим уравнение оптимальности Беллмана, для задачи о замене оборудования с применением марковских процессов, заного выразим $u_t(c_x)$ и $u_t(d_x)$, используя начальные условия $R(x)$, $S(x)$ и $\tilde{P}$, тогда получим:

$$\begin{gathered} \tilde{\nu}_{t-1}(x) = \max[\tilde{u}_t(c_x), \tilde{u}_t(d_x)], \\ \tilde{u}_t(c_x) = p_x R(x) + q_x \tilde{\nu}_t(x_0) + p_x \tilde{\nu}_t(x+1), \\ \tilde{u}_t(d_x) = p_x S(x) - \tilde{P} + R(x_0) + \tilde{\nu}_t(0), \\ \forall x = 0, \ldots, k, \quad t = 1, \ldots, n, \\ \text{причём } \tilde{\nu}_n(x) \equiv r = 0 \end{gathered} \tag{15}$$

Таким образом простой алгоритм динамического программирования был расширен путём внесения его в модель управляемых марковских процессов.

Заключение. В ходе работы была достигнута основая цель – построение алгоритма для решения задачи о замене оборудования и его реализация методами динамического программирования. Для достижения её была выполнена каждая из поставленных задач, а именно:

1. Была изучена теория управляемых марковских процессов.
2. Была посторена математическая модель с использованием управляемых марковских процессов для задачи о замене оборудования.
3. Был разработан алгоритм на основе уравнения оптимальности Беллмана.
4. Данный алгоритм был реализован на языке программирования Python.

5. Была произведена оптимизация реализованного алгоритма посредством использования сторонних библиотек и с применением метода для оптимизации программ динамического программирования.

Данная работа включает в себя 2 раздела. В подразделах 1.1 и 1.2 задаются компоненты конечной управляемой марковской модели и вводятся основные понятия, связанные с ними. В подразделе 1.3 исследуется фундаментальное уравнение, рассматривается и обосновывается задача оптимального уравнения. В подразделе 1.4 вводятся уравнения оптимальности, с помощью которых происходит вычисление оценок начальных состояний и нахождение оптимальной стратегии. Во втором разделе приводится постановка и моделирование задачи о замене оборудования, а также её решение методами динамического программирования. Кроме этого во втором разделе, в пунктах 2.2 и 2.3 описаны методы оптимизации данного алгоритма, а в 2.4 описаны его отличия относительно простого алгоритма динамического программирования.

В заключение можно сказать, что марковские процессы нашли широкое применение в различных областях наук. В экономике эти процессы используются для решения многих задач, встречающихся при исследовании операций, управления запасами, прогнозирования, при чём применение оптимальных стратегий управления может дать весьма значительный экономический эффект.

Несомненно, что дальнейшее развитие теории управляемых марковских процессов будет сопровождаться расширением областей их применения в различных науках.