Министерство образования и науки Российской Федерации

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ

ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ

ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

ИМЕНИ Н.Г. ЧЕРНЫШЕВСКОГО»

Кафедра теории функций и стохастического анализа

Управляемые марковские процессы и их применение к задаче о распределении ставок в игре

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

студентки 4 курса 412 группы

направления 01.03.02 Прикладная математика и информатика

механико-математического факультета

Бунчуковой Анастасии Александровны

Научный руководитель
к. ф.-м. н., доцент И.А. Кузнецова
должность, уч. степень, уч. звание подпись, дата инициалы, фамилия

Зав. кафедрой
д. ф.-м. н., доцент С.П. Сидоров
должность, уч. степень, уч. звание подпись, дата инициалы, фамилия

Саратов 2017

ВВЕДЕНИЕ

Актуальность темы исследования.

Управляемые марковские процессы возникают в самых разнообразных областях. Обратимся, например, к экономическому планированию. Планировать можно работу отдельного предприятия, отрасли или всего народного хозяйства. В начале каждого периода, исходя из достигнутого состояния, намечается план на следующий период. Развитие системы можно описывать математически как управляемый детерминированный процесс, если считать, что состояние системы в конце каждого периода однозначно определяется состоянием в начале периода и планом на этот период. Однако не всегда можно пренебрегать влиянием таких факторов, как метеорологические условия, демографические сдвиги, колебания спроса, несовершенство координации сложных производственных процессов, научные открытия и изобретения. Эти факторы лучше учитываются стохастическими моделями, в которых, зная состояние в начале периода и план, можно вычислить лишь распределение вероятностей для состояния в конце периода. Таким образом приходим к управляемому марковскому процессу.

Управляемые марковские процессы рассматриваются в задачах таких как: распределение ресурса между производством и потреблением и между различными отраслями производства, замена оборудования, стабилизация линейной систем, находящейся под влиянием случайных возмущений, распределение ставок в игре и т.д.

Актуальность определила выбор **темы** данной работы: «Управляемые марковские процессы и их применение к задаче о распределении ставок в игре».

Целью работы является изучение теории управляемых марковских процессов и решение задач о распределении ставок в игре для конечного случая.

Объект и предмет исследования - управляемые марковские процессы и задача о распределении ставок в игре.

Исследование имеет **практическую значимость**. Вместо двух игр можно рассматривать два способа помещения денег (например, в сбербанк или коммерческий банк), или две производственные отрасли с различными коэффициентами отдачи.

Основное содержание работы

Выпускная квалификационная работа состоит из введения, трех теоретических и одной практической главы, заключения, списка использованных источников и приложения.

Введение содержит основные положения: статически подкрепленную актуальность темы исследования; цель, объект, предмет, и практическую значимость исследования.

В первой главе «Конечные управляемые марковские модели на промежутке времени [m,n]» приведены основные компоненты управляемых марковских моделей, рассмотрены стратегии и оценки, фундаментальное уравнение, уравнения оптимальности. Теоретической основой были источники литературы [1]-[2].

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

1. Множества $X_m, ..., X_t, ..., X_n$. $X_t(t = m, ..., n)$ - множество состояний в момент времени t,

X_m - множество начальных 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 $A_m, ..., A_t, ..., A_{n-1}$. $A_t(t = m, ..., n-1)$ - 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 $\forall x \in X_t$ $\alpha(x) \subset A_t$.

4. Отображение P, ставящее в соответствие каждому управлению $a \in A_t$ распределение вероятностей P_a на X_{t+1}. P - переходная функция.

Распределения P_a мы будем 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 $x = j(a)$, 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.

Путем называется следующая последовательность состояний и управлений

$$l = x_m a_m ... x_t a_t ... x_{n-1} a_{n-1} x_n.$$

L - *множество всех путей.*

Оценкой пути l называется число

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

Вероятностью пути называется число

$$P(l) = \mu(x_m) \prod_{t=m}^{n-1} P(x_{t+1} \mid a_t).$$

В главе также рассмотрена простая стратегия. С помощью нее в каждом нефинальном состоянии выбирается конкретное допустимое управление. Существует более сложный способ управления, чем простые стратегии, а именно способ, когда в каждом состоянии выбирается не конкретное управление, а распределение вероятностей на множестве допустимых управлений, и оно зависит не только от состояния, в котором находится процесс, но и от всего его развития. Таким способом является стратегия π.

Далее в главе даны определения оптимальности стратегий и приведены теоремы, в которых доказывается, что в конечной управляемой марковской модели существует равномерно оптимальная стратегия.

Стратегия π_0 называется *оптимальной для начального распределения* μ, если справедливо равенство

$$\omega(\mu, \pi_0) = \max_{\pi} \omega(\mu, \pi).$$

Стратегия π_0 называется *оптимальной для начального состояния* $x \in X_m$, если справедливо равенство

$$\omega(\mu, \pi_0) = \max_{\pi} \omega(x, \pi).$$

После приведены понятия фундаментального уравнения и уравнений оптимальности.

Оценка стратегии π при начальном состоянии x в исходной модели связана с оценкой ω' в производной модели равенством:

$$\omega(x, \pi) = \sum_{a \in \alpha(x)} \pi(a/x)[q(a) + \omega'(P_a, \pi'_a)],$$

называемым *фундаментальным уравнением*.

Если ν - оценка состояний в исходной модели Z, ν' - оценка состояний в производной модели Z', то ν и ν' связаны соотношением

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

где $u(a) = q(a) + \nu'(P_a) = q(a) + \sum\limits_{y \in X_{m+1}} P(y/a)\nu'(y)$.

Соотношение называется *уравнением оптимальности*.

Во второй главе «Задача о распределении ставок в игре» описана постановка задачи о распределении ставок в игре и приведены результаты решения этой задачи для конечного случая.

Задача состоит в следующем:

Пусть имеющийся капитал x можно распределить между двумя вариантами игры. При ставке z выигрыш в первой игре равен σz, а во второй τz, где σ и

τ- случайные величины с различными распределениями вероятностей. Игра повторяется многократно. Пусть x_t - общая сумма, которой играющий располагает в момент времени t. Тогда, если a_t - капитал, вкладываемый на шаге t в первую игру, $x_t - a_t$ - капитал, вкладываемый во вторую игру, то капитал на шаге $t+1$ вычисляется по формуле:

$$x_{t+1} = \sigma_t a_t + \tau_t (x_t - a_t).$$

Вместо двух игр можно рассматривать два способа помещения денег (например, в сбербанк или коммерческий банк), или две производственные отрасли с различными коэффициентами отдачи.

Целью является получение максимально возможного окончательного выигрыша, который можно оценить с помощью неубывающей функции $r(x)$. Оптимальное поведение зависит от вида функции r. Может случиться, что требуется определенная сумма c и целью является выиграть эту сумму с максимальной вероятностью. В этом случае надо положить:

$$r(x) = \begin{cases} 1, x \geq c, \\ 0, x < c. \end{cases}$$

В таблице 2.2 приведены результаты выполнения алгоритма решения задачи о распределении ставок в игре для конечного случая при вероятностях: $p = 0.2$ и $p = 0.8$.

Таблица 2.2 — Расчет начальных состояний и оптимальных стратегий $p = 0.8$

x	a	$\nu_t(x) = \max u_t(x, a)$	$\psi_t^0(x)$
0	0	0	0
0.1	0, 0.1	0.0056	0.1
0.2	0, 0.1, 0.2	0.0144	0.1
0.3	0, 0.1, 0.2, 0.3	0.0418	0.1
0.4	0, 0.1, 0.2, 0.3, 0.4	0.072	0.1
0.5	0, 0.1, 0.2, 0.3, 0.4, 0.5	0.1108	0.1
0.6	0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6	0.2364	0.1
0.7	0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7	0.2807	0.1
0.8	0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8	0.3215	0.1
0.9	0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9	0.3852	0.1
1	0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1	1.0	0.1

Продолжение таблицы 2.2

$p = 0.2$

x	a	$\nu_t(x) = \max u_t(x, a)$	$\psi_t^0(x)$
0	0	0	0
0.1	0, 0.1	0.2016	0.1
0.2	0, 0.1, 0.2	0.2498	0.2
0.3	0, 0.1, 0.2, 0.3	0.3124	0.3
0.4	0, 0.1, 0.2, 0.3, 0.4	0.3637	0.4
0.5	0, 0.1, 0.2, 0.3, 0.4, 0.5	0.4218	0.5
0.6	0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6	0.4762	0.4
0.7	0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7	0.5122	0.3
0.8	0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8	0.5618	0.2
0.9	0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9	0.6235	0.1
1	0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1	1.0	0

В третьей главе «Полунепрерывные модели» рассмотрены основные определения и свойства полунепрерывной модели, и также описано нахождение оптимальной стратегии в задаче о распределении ставок в игре. Теоретической основой были источники литературы [3]-[10].

Пусть E - произвольное метрическое пространство. Напомним, что множество E называется *метрическим пространством*, если любым x, $y \in E$ сопоставлено неотрицательное число $\rho(x, y)$ (расстояние между x и y), такое что:

1) $\rho(x, y) = \rho(y, x)$;

2) $\rho(x, y) = 0$ тогда и только тогда, когда $x = y$;

3) $\rho(x, y) \leq \rho(x, z) + \rho(y, z) \; \forall \; x, y, z \in E$.

Функция, заданная на E называется *полунепрерывной*, если все множества $\{x : f(x) \geq c\}$, где c - действительное число, замкнуты. Обычно такие функции называют полунепрерывными сверху.

Пусть даны два метрических пространства $(E, B(E))$ и $(E', B(E'))$. Соответствие из E в E' сопоставляет каждой точке x пространства E непустое множество $\Phi(x)$ в другом пространстве E'. Говорят, что соответствие $\Phi(x)$ *квазинепрерывно по* x, если при $x_k \to x \in X (x_k \in X)$ и $a_k \in \Phi(x_k)$ последовательность a_k имеет предельную точку, принадлежащую $\Phi(x)$.

Функция $\delta : E \to E'$ называется *селектором соответствия* Φ, если $\delta(x)$ принадлежит множеству $\Phi(x)$ при всех x из E.

Соответствие допускает измеримый выбор, если для него существует измеримый селектор.

Модель Z назовем полунепрерывной, если:

1) Множество состояний X - сепарабельное метрическое пространство, при этом $X_m, X_{m+1}, ..., X_n$ - замкнутые подмножества X.

2) Множество управлений A - сепарабельное метрическое пространство и $A_m, ..., A_{n-1}$ - замкнутые подмножества A.

3) Соответствие $\alpha(x)$ квазинепрерывно по x.

4) Если $f \in L(X_t)$ и $g(a) = \int_{X_t} p(dx|a) f(x)$, где $a \in A_t$. Тогда $g \in L(A_t)$ $(t = m+1, ..., n)$.

5) Плата q на множестве A_t принадлежит $L(A_t)$.

6) Плата r принадлежит $L(X_n)$.

Далее в главе приведено понятие дерзкой стратегии и с помощью индукции по n доказано, что дерзкая стратегия оптимальна.

Суть дерзкой стратегии заключается в том, что нужно делать возможно большие ставки, совместимые с наличными средствами, т.е избегая бесцельного риска. Это значит, что при $x \leq \frac{1}{2}$ следует ставить на игру весь имеющийся капитал x, при $\frac{1}{2} \leq x \leq 1$ - ставить недостающую сумму $1 - x$, при $1 \leq x$ - вообще ничего не ставить. Дерзкая стратегия задается на всех шагах одним и тем же селектором

$$\psi_0(x) = \begin{cases} x,\ 0 \le x \le \frac{1}{2}, \\ 1 - x,\ \frac{1}{2} \le x \le 1, \\ 0,\ x \ge 1. \end{cases}$$

В четвертой главе «Разработка алгоритма решения задачи о распределении ставок в игре для конечного случая» приведен алгоритм решения задачи о распределении ставок в игре для конечного случая, написанный на языке программирования Java.

В приложении представлены получившийся программный код и результаты его выполнения.

ЗАКЛЮЧЕНИЕ

В данной работе была решена задача о распределении ставок в игре с помощью теории управляемых марковских процессов.
Были рассмотрены теоретические сведения теории управляемых марковских процессов, был приведен алгоритм решения задачи для конечного случая, результаты выполнения которого доказывают, что нужно делать возможно большие ставки, совместимые с наличными средствами, т.е избегая бесцельного риска. При капитале $x \leq \frac{1}{2}$ следует ставить на игру весь имеющийся капитал x, при $\frac{1}{2} \leq x \leq 1$ - ставить недостающую сумму $1 - x$, при $1 \leq x$ - вообще ничего не ставить.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Кузнецова, И.А. Управляемые марковские процессы и их приложения. СГУ, 1999. – С. 4-10, 17-19.

2. Дынкин, Е. Б., Юшкевич, А.А. Управляемые марковские процессы и их приложения. – М. Наука, 1975. – С. 185.

3. Аоки, М. Оптимизация стохастических систем. – М. Физматлит, 1971. С. 150.

4. Губенко, Л.Т., Штатланд, Э.С. Об управляемых полумарковских процессах. – М. Кибернетика, 1972. – С. 26-29.

5. Крылов, Н.В. О существовании ε-оптимальных однородных марковских стратегий для управляемой цепи. – М. ДАН СССР, 1964. – С. 747-750.

6. Юшкевич, А.А. Об одном классе стратегий в общих управляемых марковских моделях. – Теор.вер.и ее примен., 1972. – С. 815-817.

7. Дынкин, Е.Б. Основания теории марковских процессов. М. Физматлит, 1959. – С. 25-26.

8. Дынкин, Е.Б. Управляемые случайные последовательности. – Теор.вер.и ее примен., 1965. – С. 3-18.

9. Дынкин, Е.Б. Пространство выходов марковского процесса. – Успехи мат. наук., 1969. – С. 89-152.

10. Гнеденко, Б.В. Курс теории вероятностей. – М. Наука, 1988. – С. 88-90.