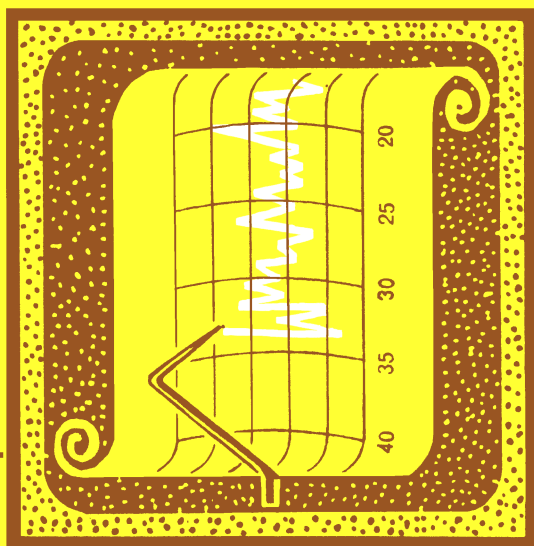


НОВОЕ  
В ЖИЗНИ, НАУКЕ,  
ТЕХНИКЕ

ЗНАНИЕ

М. И. Фрейдлин  
СЛУЧАЙНЫЕ  
ПРОЦЕССЫ  
И ОПТИМАЛЬНОЕ  
УПРАВЛЕНИЕ



7/1972

СЕРИЯ  
МАТЕМАТИКА,  
КИБЕРНЕТИКА

**М. И. Фрейдлин,**  
доктор физико-математических наук

**СЛУЧАЙНЫЕ  
ПРОЦЕССЫ  
И ОПТИМАЛЬНОЕ  
УПРАВЛЕНИЕ**

ИЗДАТЕЛЬСТВО «ЗНАНИЕ»  
МОСКВА 1972

517.8  
Ф 86

**Фрейдлин Марк Иосифович**

Ф86      Случайные процессы и оптимальное управление. М., «Знание», 1972.  
48 с. (Новое в жизни, науке, технике. Серия «Математика, кибернетика», 7)

В брошюре рассматриваются вероятностные задачи оптимального управления, излагаются основные факты из теории случайных процессов, необходимые для построения и исследования стохастических моделей систем управления, и вводится ряд понятий теории оптимального управления.

Брошюра рассчитана на широкий круг читателей, интересующихся проблемами современной математики и практическими приложениями теории вероятностей.

2—2—3

517.8

## ВВЕДЕНИЕ

Трудно назвать область человеческой деятельности, где не возникали бы задачи управления. Это — управление станочной линией и полетом космической ракеты, управление экономикой в масштабе отрасли или даже страны и семейным бюджетом, управление в живых организмах и большими коллективами людей. Задачи такого рода имеют много различных аспектов, и каждая из них при своем решении требует привлечения специальных знаний из соответствующего раздела науки или техники. Управление космической ракетой немислимо, например, без учета законов баллистики и газовой динамики; изучение процессов регулирования в живых организмах основывается на достижениях биологии. Имеются, однако, некоторые черты, общие во всех этих задачах, позволяющие рассматривать существенные стороны таких задач с единой точки зрения. Эта общность становится зримой, когда мы заменяем реальную задачу математической моделью. Здесь мы сталкиваемся с положением вещей, которое нередко встречается в науке. Достаточно вспомнить, например, математическую теорию колебаний, которая хорошо описывает процессы, имеющие совершенно различную физическую природу. Рассмотрению некоторого класса математических моделей процессов управления и посвящена эта брошюра. Точнее, речь будет идти о моделях, параметры которых эволюционируют по законам случайного процесса.

Необходимость рассмотрения стохастических моделей систем управления возникает в большом количестве задач и объясняется рядом причин. Если, скажем, речь идет об управлении созданием запасов, то, обычно, необходимо учитывать случайные флуктуации спроса. К рассмотрению стохастических моделей приводят также учет метеорологических условий, неоднородность сырья, поступающего на обработку, случайный характер возникновения неисправностей в сложных системах, «шум» в электронных устройствах и другие причины.

Построение математической модели системы управления, достаточно хорошо описывающей изучаемый реальный объект,

с одной стороны, и допускающей математический расчет, с другой, требует знания возможностей теории случайных процессов. В связи с этим сначала мы охарактеризуем различные классы случайных процессов и те качественные условия, которые к ним приводят, рассмотрим вопрос о прохождении случайных процессов через линейные и некоторые нелинейные системы, обсудим различные подходы к понятию устойчивости при случайных возмущениях, затем введем основные понятия теории оптимального управления, такие, как функция риска, оптимальная и  $\epsilon$ -оптимальная стратегия, достаточные статистики. Мы рассмотрим различные конкретные схемы управления случайными процессами: управление созданием запасов, задачу об оптимальной остановке и обнаружении разладки производственного процесса, оптимальную стабилизацию при случайных возмущениях. В заключение коротко обсудим вопрос о построении вероятностных моделей процессов управления.

## § 1. Вероятности и случайные величины

Различные разделы математики с разных сторон описывают объекты реального мира и отношения между этими объектами. Так, точки, треугольники, сферы, которые рассматриваются в геометрии, являются математическими моделями реальных фигур; аксиомы геометрии отражают закономерности взаимного расположения этих реальных объектов в пространстве. Понятия и аксиомы математической теории вероятностей также связаны с некоторыми реальными объектами. Само понятие вероятности является абстрактным аналогом понятия частоты. Аксиомы теории вероятностей — абстракция свойств частот. Применение теории вероятностей в практических задачах предполагает, что изучаемые явления носят массовый характер и при этом частоты тех или иных событий (исходов экспериментов) обладают свойствами устойчивости. Это предположение следует иметь в виду и при интерпретации получающихся результатов. Если мы говорим, например, что в некотором опыте определенным исходом  $A$  имеет вероятность  $1/4$ , то это означает, что при проведении серии таких опытов исход  $A$  произойдет примерно в  $1/4$  случаев; высказывание «исход  $B$  маловероятен» означает, что частота исхода  $B$  в серии опытов мала.

Напомним основные понятия и аксиомы теории вероятностей. Пусть некоторый эксперимент может иметь  $N$  различных исходов  $\omega_1, \omega_2, \dots, \omega_N$ . Совокупность всех этих исходов обозначим  $\Omega$ ; множество  $\Omega$  называют пространством элементарных исходов или элементарных событий. Всевозможные подмножества пространства  $\Omega$  называются событиями. Говорят, что произошло событие  $A$ , если в результате эксперимента осуществился элементарный исход  $\omega_k$ , принадлежащий множеству  $A$ .

Если есть два события  $A_1$  и  $A_2$ , то их объединением называется событие, состоящее из исходов, принадлежащих или  $A_1$  или  $A_2$ ; пересечением событий  $A_1$  и  $A_2$  называют совокупность исходов, принадлежащих как  $A_1$ , так и  $A_2$ . Говорят, что события  $A_1$  и  $A_2$  не пересекаются, если они не имеют общих элементов.

Пусть для каждого события  $A$  определено число  $P(A)$ , и эти числа удовлетворяют условиям:

- 1)  $0 \leq P(A) \leq 1, P(\Omega) = 1$ ;
- 2) если множество  $A$  есть объединение непересекающихся событий  $A_1$  и  $A_2$ , то  $P(A) = P(A_1) + P(A_2)$ .

В таком случае число  $P(A)$  называется вероятностью события  $A$ .

Если иметь в виду, что вероятность является абстрактным аналогом понятия частоты, то аксиомы 1) и 2) отражают основные свойства частот. В самом деле, если мы повторим наш эксперимент  $n$  раз, то частота события  $A$  есть  $\frac{\nu^{(n)}(A)}{n}$ , где  $\nu^{(n)}(A)$  — число опытов, в которых  $A$  произошло. Так как  $0 \leq \nu^{(n)}(A) \leq n$ , то, очевидно, частота заключена между 0 и 1, а так как в результате каждого эксперимента мы обязательно получаем один из исходов  $\omega_1, \omega_2, \dots, \omega_N$ , то  $\nu^{(n)}(\Omega) = n$  и, стало быть, частота события  $\Omega$  равна 1. Вторая аксиома тоже отражает некоторое свойство частоты. Если  $A_1$  и  $A_2$  — непересекающиеся подмножества из  $\Omega$ , то события  $A_1$  и  $A_2$  не могут произойти в одном эксперименте, поэтому  $\nu^{(n)}(A) = \nu^{(n)}(A_1) + \nu^{(n)}(A_2)$ , и, следовательно, частота события  $A = A_1 + A_2$  равна сумме частот событий  $A_1$  и  $A_2$ .

Если  $A$  и  $B$  два события, принадлежащие пространству элементарных исходов  $\Omega$ , то условная вероятность  $P(A/B)$  события  $A$  при условии  $B$  определяется равенством

$$P(A/B) = \frac{P(AB)}{P(B)},$$

где через  $AB$  обозначено пересечение событий  $A$  и  $B$ . Естественность такого понятия также может быть оправдана рассуждениями с частотами.

На основе понятия условной вероятности определяется независимость событий. Событие  $A$ , естественно, считать независимым от  $B$ , если условная вероятность  $P(A/B)$  совпадает с безусловной вероятностью  $P(A)$  события  $A$ . Если воспользоваться определением условной вероятности, то последнее высказывание можно записать так:  $P(AB) = P(A) \cdot P(B)$ . Это равенство и служит определением независимости событий.

Наряду с событиями — подмножествами пространства  $\Omega$ , для которых определена вероятность, рассматриваются также числовые функции на пространстве  $\Omega$ . Такие функции называются случайными величинами. Если  $\xi(\omega)$  случайная величина,

то с ней можно связать семейство событий  $\{\omega : \xi(\omega) < x\}$  — подмножество пространства  $\Omega$ , на котором  $\xi(\omega)$  принимает значения меньше  $x$ ,  $x$  — любое вещественное число. Если одновременно рассматривается несколько случайных величин  $\xi^1(\omega), \xi^2(\omega), \dots, \xi^{(n)}(\omega)$  на пространстве  $\Omega$ , то говорят, что задана многомерная случайная величина  $\xi(\omega) = (\xi^1(\omega), \xi^2(\omega), \dots, \xi^{(n)}(\omega))$ .

Две величины  $\xi(\omega)$  и  $\eta(\omega)$  называются независимыми, если события, связанные с одной случайной величиной и с другой величиной, независимы, то есть если при любых  $x$  и  $y$  выполняется равенство

$$P\{\xi(\omega) < x, \eta(\omega) < y\} = P\{\xi(\omega) < x\} \cdot P\{\eta(\omega) < y\}.$$

Основной характеристикой случайной величины  $\xi(\omega)$  является функция распределения  $F_\xi(x) = \{P\xi(\omega) < x\}$ . В многомерном случае тоже вводится функция распределения. Если, скажем,  $\xi(\omega) = (\xi^1, \xi^2)$  — двумерная случайная величина, то функция распределения  $F_\xi(x, y)$  определяется равенством  $F_\xi(x, y) = P\{\xi^1 < x, \xi^2 < y\}$ .

Рассмотрим пример. Предположим, что проводится серия из  $n$  испытаний. В каждом испытании два исхода, которые мы условимся называть «успех» (У) и «неуспех» (Н). Тогда в результате испытаний мы получим последовательность из «успехов» и «неуспехов». Всего различных таких последовательностей длины  $n$  может быть  $2^n$ . Таким образом, пространство элементарных событий  $\Omega$  в нашем эксперименте состоит из  $2^n$  точек. Чтобы задать вероятности на подмножествах пространства  $\Omega$ , следует уточнить, как происходят испытания. Во-первых, будем считать, что вероятность успеха равна  $p$  и не меняется от испытания к испытанию. Тогда вероятность неуспеха есть  $q = 1 - p$ . Во-вторых, будем считать, что испытания независимые. Это значит, что для вычисления вероятности одной точки пространства  $\Omega$  последовательности из  $n$  испытаний, следует перемножить вероятности исходов в первом, во втором, ..., в  $n$ -ом испытании. Так, например, если  $n = 4$ , то вероятность последовательности УННН равна  $pqqq = pq^3$ . Таким образом, мы вводим вероятности на пространстве  $\Omega$ . Такие независимые испытания с двумя исходами и неизменной вероятностью успеха называются испытаниями Бернулли. Испытания Бернулли это одна из самых простых и в то же время вполне содержательных схем, на которой были замечены и доказаны многие теоремы теории вероятностей.

Рассмотрим на введенном пространстве  $\Omega$  случайную величину  $S_n = S_n(\omega)$  — число успехов в  $n$  испытаниях. Величина  $S_n(\omega)$  принимает значения  $0, 1, 2, 3, \dots, n$ ; нетрудно проверить, что

$$P\{S_n(\omega) = k\} = C_n^k p^k q^{n-k} = \frac{n! p^k q^{n-k}}{k!(n-k)!}.$$

Функция распределения  $F(x)$  величины  $S_n$  представляет собой

ступенчатую функцию на числовой прямой, возрастающую от 0 до 1. Скачки функции  $F(x)$  сосредоточены в целых точках  $k = 0, 1, 2, \dots, n$ . Величина скачка в точке  $k$  равна  $P\{S_n = k\}$ . Случайная величина  $S_n(\omega)$  представляет собой пример так называемой дискретной случайной величины. Дискретными называют случайные величины, принимающие только конечное или счетное число значений. Функция распределения у таких величин кусочно-постоянная. Другой пример дискретной случайной величины дает случайная величина, имеющая распределение Пуассона. Такая величина  $\nu$  принимает все возможные целые неотрицательные значения и

$$P\{\nu = k\} = e^{-\lambda} \frac{\lambda^k}{k!},$$

где  $\lambda$  — положительный параметр. Это распределение весьма часто встречается в различных прикладных задачах.

В теории вероятностей случайные величины рассматриваются, как правило, с точностью до функции распределения. Поэтому обычно говорят: «рассмотрим случайную величину с такой-то функцией распределения», не указывая пространства элементарных событий, на котором она определена. Так, например, мы ввели пуассоновскую величину, не упоминая о соответствующем пространстве элементарных событий.

Если функция распределения  $F_\xi(x)$  величины  $\xi$  представима в виде  $F_\xi(x) = \int_{-\infty}^x p_\xi(y) dy$ , то  $p_\xi(y)$  называется плотностью случайной величины  $\xi$ . Плотность всюду неотрицательна и  $\int_{-\infty}^{\infty} p_\xi(y) dy = 1$ .

Если плотность случайной величины равна нулю вне некоторого отрезка  $[a, b]$ , а на этом отрезке равна константе, то говорят, что эта случайная величина распределена равномерно на  $[a, b]$ . Другой пример случайных величин, имеющих плотность, доставляют показательно распределенные величины. В этом случае плотность зависит от одного параметра и имеет вид:

$$p(x) = ae^{-ax} \quad \text{при } x \geq 0 \quad \text{и} \quad p(x) = 0 \quad \text{при } x < 0.$$

Такие распределения часто возникают, например, в задачах массового обслуживания и надежности.

Важный класс случайных величин составляют так называемые нормальные или гауссовские случайные величины. Это величины, плотность которых задается формулой

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}},$$



где  $a$  и  $\sigma$  — параметры. Оказывается, что, если некоторая случайная величина является суммой большого числа независимых, небольших по сравнению со всей суммой случайных слагаемых, то распределение этой величины близко к нормальному. Соответствующее точное утверждение составляет содержание центральной предельной теоремы теории вероятностей. Именно этим утверждением и объясняется широкое распространение случайных величин, распределение которых с большой точностью можно считать нормальным.

Для многомерной случайной величины плотность определяется аналогично. Скажем, плотность  $p(x^1, x^2)$  двумерной величины  $\xi = (\xi^1, \xi^2)$  определяется равенством  $F_\xi(y^1, y^2) =$

$$\int_{-\infty}^{y^1} \int_{-\infty}^{y^2} p_\xi(x^1, x^2) dx^1 dx^2.$$

Если случайные величины  $\xi^1, \xi^2$ , образующие вектор  $\xi$ , независимы и нормальны, то  $\xi$  называют двумерной нормальной величиной. Двумерная случайная величина  $\eta = (\eta^1, \eta^2)$ , которая получается из нормальной величины  $\xi = (\xi^1, \xi^2)$  с помощью линейного преобразования, также называется нормальной или гауссовской. Таким образом, при линейных преобразованиях гауссовские величины переходят в гауссовские. Аналогично определяются многомерные гауссовские величины при  $n > 2$ .

Описание случайных величин с помощью функции распределения или плотности не всегда удобно. Это объясняется тем, что такие характеристики трудно определить по наблюдениям, а если они даже определены, то их нелегко использовать. Поэтому вводят более грубые характеристики случайных величин.

Самой простой характеристикой распределения является математическое ожидание. В случае дискретной величины  $\xi$ , принимающей значения  $a_1, a_2, \dots, a_n, \dots$  с вероятностями  $p_1, p_2, \dots, p_n, \dots$  соответственно математическое ожидание  $M\xi$  вычисляется по формуле  $M\xi = \sum_k a_k p_k$ . Если величина  $\xi$  имеет плотность  $p_\xi(x)$ , то  $M\xi = \int_{-\infty}^{\infty} x p_\xi(x) dx$ .

Отклонение случайной величины  $\xi$  от своего математического ожидания  $M\xi$  характеризуется дисперсией  $D\xi = M|\xi - M\xi|^2$ .

В многомерном случае также вводятся числовые характеристики распределения. Математическое ожидание случайной величины  $\xi = (\xi^1, \dots, \xi^n)$  определяется как вектор, состав-

ленный из математических ожиданий компонент. Отклонения от математического ожидания характеризуются матрицей ковариаций  $b^{ij} = M(\xi^i - M\xi^i)(\xi^j - M\xi^j)$ . Если  $\xi$  — многомерная гауссовская величина, то ее плотность однозначно восстанавливается по вектору математических ожиданий и матрице ковариаций.

Для независимых случайных величин  $\xi^i$  и  $\xi^j$  коэффициент  $b^{ij}$  в матрице дисперсий равен нулю. Из равенства нулю  $b^{ij}$ , вообще говоря, не следует независимость  $\xi^i$  и  $\xi^j$ . Если, однако, дополнительно известно, что двумерная величина  $(\xi^i, \xi^j)$  — гауссовская, то обращение  $b^{ij}$  в ноль влечет за собой независимость случайных величин  $\xi^i$  и  $\xi^j$ . На этих свойствах  $b^{ij}$  основывается использование отношения  $r_{ij} = \frac{b^{ij}}{\sqrt{b^{ii}b^{jj}}}$  в качестве меры зависимости величин  $\xi^i$  и  $\xi^j$ . Это отношение  $r_{ij}$  называется коэффициентом корреляции. Можно доказать, что коэффициент корреляции заключен между 0 и 1, и для независимых величин он равен 0. Если  $r_{ij} = 1$ , то между величинами  $\xi^i$  и  $\xi^j$  имеется линейная связь.

## § 2. Гауссовские случайные процессы

Семейство случайных величин  $\xi_t(\omega)$ , зависящих от параметра  $t$ , называется случайным процессом. Нетрудно привести примеры реально наблюдаемых величин, которые в некоторых ситуациях естественно считать случайным процессом. Так, величину тока в электрической сети, измеренную как функцию времени  $t$ , иногда удобно рассматривать как случайный процесс. Случайным процессом можно считать толщину ткацкой нити, температуру в данной точке как функцию времени, положение в момент  $t$  диффундирующей частицы. В этих примерах  $t$  пробегает все точки некоторого отрезка действительной оси. В других случаях параметр  $t$  может принимать только целые значения; например, размер детали с номером  $t$ , число опечаток на первых  $t$  страницах книги. Если  $t$  принимает только целочисленные значения, то процесс  $\xi_t(\omega)$  иногда называют случайной последовательностью. Когда параметр  $t$  изменяется на действительной прямой, его обычно называют временем, хотя он может иметь совсем иную физическую природу. Если  $t$  — многомерный параметр, то процесс  $\xi_t(\omega)$  называют случайным полем.

Таким образом, случайный процесс  $\xi_t(\omega)$  это функция от двух переменных: от момента времени  $t$  и от элементарного исхода  $\omega$ . При фиксированном  $t = t_0$  мы получаем случайную величину  $\xi_{t_0}(\omega)$ . Если зафиксировать  $\omega = \omega_0$ , то мы получим функцию времени  $\xi_t(\omega_0)$ , которую называют траекторией или реализацией процесса  $\xi_t(\omega)$ .

Для описания случайного процесса недостаточно задать функции распределения случайных величин  $\xi_t(\omega)$  при всевозможных  $t$ . Необходимо охарактеризовать взаимную связь случайных величин в различные моменты времени. Чтобы описать случайный процесс, можно задать его конечномерные распределения, то есть распределения многомерных случайных величин  $(\xi_{t_1}, \xi_{t_2}, \dots, \xi_{t_n})$  при всевозможных  $t_1, t_2, \dots, t_n$  и любых натуральных  $n$ . Такое описание случайного процесса весьма громоздко. Поэтому обычно ограничиваются более грубыми характеристиками случайных процессов или рассматривают классы случайных процессов, которые допускают более удобное описание.

Наиболее простыми характеристиками случайного процесса являются его математическое ожидание  $m(t) = M\xi_t(\omega)$  и корреляционная функция  $B(s, t) = M(\xi_s - m(s))(\xi_t - m(t))$ . Корреляционная функция характеризует как отклонение случайного процесса от своего среднего (вследствие того, что  $D\xi_t = B(t, t)$ ), так и в некотором смысле, степень зависимости случайных величин  $\xi_s$  и  $\xi_t$ . Коэффициент корреляции между этими величинами  $r_{st} = \frac{B(s, t)}{\sqrt{B(s, s)B(t, t)}}$ . Та часть теории случайных процессов, которая описывает свойства процесса через его математическое ожидание и корреляционную функцию, называется теорией второго порядка или корреляционной теорией.

Простейшим примером случайного процесса служит последовательность независимых случайных величин  $\xi_1, \xi_2, \dots, \xi_n, \dots$ . Уже этот объект является весьма содержательным, и долгое время изучение последовательностей независимых случайных величин составляло большую часть теории вероятностей. Следует заметить, что рассмотрение случайных процессов более общего вида, в значительной степени основано на представлении таких процессов через независимые случайные величины.

Одним из классов случайных процессов, для которого целый ряд задач может быть решен в сравнительно обозримом виде, является класс гауссовских случайных процессов. Гауссовским называется такой случайный процесс  $x_{t_n}(\omega)$ , у которого все конечномерные распределения (то есть распределения многомерных случайных величин  $(x_{t_1}(\omega), x_{t_2}(\omega), \dots, x_{t_n}(\omega))$ ) гауссовские.

Реальные процессы, по-видимому, довольно часто можно считать гауссовскими. Это объясняется, по существу, центральной предельной теоремой, о которой упоминалось выше. Класс гауссовских процессов замечателен еще в том отношении, что он инвариантен относительно линейных преобразований и предельного перехода. Однако при прохождении через нелинейные фильтры гауссовость нарушается. Поэтому предположение о гауссовском характере наблюдаемого процесса требует анализа

того, как этот процесс формируется, и проверки согласия такой гипотезы с опытом.

Все конечномерные распределения гауссовского процесса  $x_t$  могут быть восстановлены по его математическому ожиданию  $m(t)$  и корреляционной функции  $B(s, t)$ . Таким образом, функции  $m(t)$  и  $B(s, t)$  по существу однозначно определяют гауссовский процесс. Посмотрим, как изменяются эти характеристики при преобразованиях процесса.

Пусть у нас имеется линейная система  $G$  с импульсной характеристикой  $G(s, t)$ . В случае целочисленного параметра  $t$  это значит, что если на вход системы  $G$  поступает сигнал  $\varphi(t)$ , равный 1 в момент  $s$  и 0 в остальные моменты времени, то на выходе получится сигнал  $\psi(t) = G\varphi = G(s, t)$ . Линейность системы означает, что реакция  $h(t)$  системы  $G$  на произвольный входной сигнал  $g(t)$  имеет вид

$$h(t) = \sum_s G(s, t)g(s).$$

Аналогично вводится импульсная характеристика  $G(s, t)$  в случае непрерывного параметра  $t$ : сигнал на выходе  $h(t)$  выражается через сигнал  $g(t)$  на входе по формуле

$$h(t) = \int_{-\infty}^{\infty} G(s, t)g(s)ds.$$

Если на вход линейной системы  $G$  поступает случайная последовательность  $x_t$  со средним  $m(t)$  и корреляционной функцией  $B(s, t)$ , то на выходе мы получим последовательность  $y_t$ . Легко видеть, что  $My_t = \bar{m}(t) = \sum_s G(s, t)m(s)$ . Корреляционная функция  $B(s, t)$

процесса  $y_t$  выражается равенством  $\bar{B}(s, t) = \sum_{u_1, u_2} G(u_1, s)G(u_2, t)$

$B(u_1, u_2)$ . Таким образом, характеристики случайного процесса на выходе линейной системы сравнительно просто выражаются через характеристики входа и импульсную характеристику системы. Это особенно удобно в гауссовском случае, когда  $\bar{m}(t)$  и  $\bar{B}(s, t)$  полностью определяют процесс  $y_t$ .

Значительно сложнее обстоит дело в нелинейном случае. Здесь математическое ожидание и корреляционная функция процесса на выходе системы обычно не могут быть вычислены по этим же характеристикам на входе; необходимо иметь более подробное описание процесса на входе. В таких задачах обычно применяются различные приближенные методы. Выбор приближенного метода существенно зависит от того, какую характеристику процесса на выходе мы хотим вычислить, и от длины отрезка времени, на котором рассматривается процесс.

Среди гауссовских процессов с непрерывным параметром  $t$  особое место занимает так называемый винеровский процесс  $w_t(\omega)$ . Винеровским процессом называют гауссовский случайный процесс  $w_t(\omega)$ , определенный при  $0 \leq t < \infty$ , имеющий нулевое математическое ожидание и корреляционную функцию  $B(s, t) = \min(s, t)$ .

Этот процесс обладает рядом замечательных свойств. Если моменты времени  $t_1, t_2, t_3, t_4$  выбраны в порядке возрастания, то приращения  $w_{t_2} - w_{t_1}$  и  $w_{t_4} - w_{t_3}$  — независимые случайные величины. Траектории винеровского процесса представляют собой непрерывные функции, однако эти функции сильно «изломаны»: у винеровской траектории нигде не существует касательной.

Легко подсчитать, что  $M|w_t - w_s|^2 = Mw_t^2 + Mw_s^2 - 2Mw_s w_t = |t - s|$ . Таким образом, квадрат приращения процесса за время  $t - s$  равен  $|t - s|$ . Отсюда можно заключить, что математическое ожидание мгновенной скорости равно бесконечности. Наоборот, средняя скорость за большой промежуток времени приближается к нулю с ростом этого промежутка. Такое поведение траектории объясняется тем, что траектория «очень часто» меняет направление своего движения. Винеровский процесс иногда называют процессом броуновского движения. Это связано с тем, что винеровский процесс используют для описания движения твердой частицы в жидкости под влиянием соударений с молекулами. Такое описание является вполне удовлетворительным, если рассматриваются перемещения броуновской частицы за достаточно большие отрезки времени. Перемещения за очень малые промежутки времени не соответствуют изменению винеровской траектории (мгновенная скорость броуновской частицы может быть очень большой, но всегда конечна). Здесь мы сталкиваемся с ситуацией, которая обычно возникает при рассмотрении стохастических моделей реальных процессов. Каждая такая модель сравнительно хорошо действует только в определенных масштабах, на не очень маленьких, но и не очень больших отрезках времени.

Как уже говорилось, винеровские траектории  $w_t(\omega)$  нигде не дифференцируемы. Однако в литературе по случайным процессам, особенно прикладной, рассматривается «производная»  $\dot{w}_t$  винеровского процесса, которая называется процессом белого шума. Хотя  $\dot{w}_t$  не является случайным процессом в обычном смысле, многим выражениям, содержащим  $\dot{w}_t$  можно придать точный смысл. Так, если  $\varphi_t$  — гладкая функция, то можно ввести  $\int_0^T \varphi_t \dot{w}_t dt$  так, что этот интеграл будет обладать обычными свойствами. В частности, можно рассматривать реакцию  $x_t$  линейной системы с импульсной характеристикой

$G(s, t)$  на белый шум:  $x_t = \int_{-\infty}^{\infty} G(s, t) \dot{w}_s ds$ . Заметим, что на выходе системы мы получим уже настоящий гауссовский случайный процесс.

Процесс белого шума можно рассматривать как предел последовательности гауссовских процессов, имеющих нулевое среднее и корреляционные функции  $B_n(s, t)$ , сходящиеся при  $n \rightarrow \infty$  к  $\delta$ -функции  $\delta(t - s)$ . Именно так процесс белого шума обычно возникает в прикладных задачах.

### § 3. Стационарные процессы

Рассмотрим случайную последовательность  $\xi_1, \xi_2, \dots, \xi_n, \dots$ . Вообще говоря, все величины  $\xi_k$  имеют различные распределения, совместное распределение пары случайных величин  $(\xi_{k_1}, \xi_{k_2})$  зависит от  $k_1$  и  $k_2$ , распределение тройки  $(\xi_{k_1}, \xi_{k_2}, \xi_{k_3})$  — от чисел  $k_1, k_2, k_3$  и так далее. К определенным упрощениям приводит предположение стационарности.

Случайная последовательность  $\xi_1, \xi_2, \dots, \xi_n, \dots$  называется стационарной (точнее стационарной в узком смысле), если ее конечномерные распределения не меняются при изменении начала отсчета временного параметра. Иными словами, стационарность означает что при любом целом  $h$  многомерные величины  $(\xi_{k_1}, \dots, \xi_{k_n})$  и  $(\xi_{k_1+h}, \dots, \xi_{k_n+h})$  одинаково распределены. Здесь  $k_1, k_2, \dots, k_n$  — произвольные целочисленные моменты времени,  $n$  — любое натуральное число. В частности, если последовательность стационарна, то все величины  $\xi_k$  одинаково распределены; двумерная случайная величина  $(\xi_{k_1}, \xi_{k_2})$  имеет распределение, зависящее только от разности  $k_2 - k_1$ .

Аналогичным образом определяется стационарный в узком смысле процесс  $\xi_t$  с непрерывным параметром  $t$ , пробегающим всю числовую прямую. Только в этом случае естественно предполагать неизменность конечномерных распределений при произвольном, не только целом, сдвиге  $h$  начала отсчета параметра  $t$ .

Предположение стационарности выполняется во многих прикладных задачах. Стационарность означает, что изучаемый реальный процесс уже вышел из переходного режима, и факторы, формирующие этот процесс, остаются на рассматриваемом отрезке времени неизменными.

Если  $\xi_t$  стационарный в узком смысле случайный процесс, то  $M\xi_t = m$  не зависит от  $t$ , так как  $\xi_t$  имеет распределение, независящее от  $t$ . Корреляционная функция такого процесса  $B(s, t) = M(\xi_s - m)(\xi_t - m)$  зависит только от разности  $t - s$ , потому что случайная величина  $(\xi_s, \xi_t)$  распределена так же, как  $(\xi_0, \xi_{t-s})$ .

В корреляционной теории, где случайный процесс характеризуется только своим средним и функцией корреляции, рассматриваются процессы стационарные в широком смысле. Это такие процессы, у которых  $M\xi_t = m = \text{const}$  и  $M(\xi_s - m)(\xi_t - m) = B(t - s)$ , то есть не зависят от выбора начала отсчета параметра  $t$  только математическое ожидание и корреляционная функция. Требования стационарности в широком смысле конечно менее ограничительно, чем стационарность в узком смысле, В гауссовском случае стационарность в широком смысле влечет за собой стационарность в узком смысле, так что для гауссовских процессов оба понятия стационарности совпадают.

В теории стационарных процессов удобно рассматривать комплекснозначные процессы. Комплекснозначная случайная величина  $\xi(\omega)$  определяется равенством  $\xi(\omega) = \eta(\omega) + i\zeta(\omega)$ , где  $\eta(\omega)$  и  $\zeta(\omega)$  — вещественные величины. Математическое ожидание  $M\xi = M\eta + iM\zeta$ ;  $D\xi = M|\xi - M\xi|^2$ . Если  $\xi_t(\omega)$  — комплекснозначный случайный процесс, то его корреляционная функция  $B(s, t) = M(\xi_s - m(s))(\overline{\xi_t - m(t)})$ , где черта означает комплексное сопряжение.

Важный пример стационарного процесса с непрерывным временем можно построить следующим образом. Пусть  $\lambda_1, \lambda_2, \dots, \lambda_n$  — различные вещественные числа, а  $\xi_1, \dots, \xi_n$  — некоррелированные случайные величины с нулевым средним,  $D\xi_k = \sigma_k^2$ . Рассмотрим случайный процесс  $x_t = \sum_k \xi_k e^{i\lambda_k t}$ . Математическое ожидание этого процесса  $Mx_t = \sum_k M\xi_k e^{i\lambda_k t} = 0$ . Вычислим корреляционную функцию:

$$\begin{aligned} B(s, s + \tau) &= Mx_s \overline{x_{s+\tau}} = \sum_{k, l} M\xi_k \overline{\xi_l} e^{i\lambda_k s - i\lambda_l (s + \tau)} = \\ &= \sum_k M|\xi_k|^2 e^{-i\lambda_k \tau} \sum_k \sigma_k^2 e^{-i\lambda_k \tau}. \end{aligned}$$

При этом мы воспользовались тем, что величины  $\xi_k$  и  $\xi_l$  при  $k \neq l$  некоррелированы. Таким образом, корреляционная функция зависит только от разности моментов времени, среднее равно нулю и, следовательно,  $x_t$  — стационарный в широком смысле случайный процесс. Физически процесс  $x_t$  можно интерпретировать как сумму гармонических колебаний с частотами  $\lambda_k$  и случайными амплитудами  $\xi_k$ ;  $\sigma_k^2 = M\xi_k^2$  — величина, характеризующая среднюю энергию колебаний на частоте  $\lambda_k$ . Оказывается, что любой стационарный процесс  $x_t$  в определенном смысле есть сумма гармонических колебаний с различными частотами и некоррелированными на разных частотах случай-

ными амплитудами:  $x_t = \int_{-\infty}^{\infty} e^{i\lambda t} z(d\lambda)$ . Различные частоты, входящие в стационарный процесс  $x_t$ , заполняют, вообще говоря, всю числовую прямую, так что сумму приходится заменить интегралом. Здесь  $z(d\lambda)$  — случайная величина, характеризующая амплитуды гармонических колебаний, входящих в процесс  $x_t$ , на интервале частот от  $\lambda$  до  $\lambda + d\lambda$ . Средняя энергия колебаний на этом интервале  $M|z(d\lambda)|^2 = F(d\lambda)$ . Амплитуды на неперекрывающихся интервалах частот некоррелированы. Если существует функция  $f(\lambda)$  такая, что  $F(d\lambda) = f(\lambda)d\lambda$  то она называется спектральной плотностью случайного процесса. Можно доказать, что спектральная плотность  $f(\lambda)$  связана с корреляционной функцией  $B(\tau)$  соотношением:  $\int_{-\infty}^{\infty} e^{-i\lambda\tau} f(\lambda)d\lambda = B(\tau)$ . В целом ряде задач спектральная плотность оказывается более удобной характеристикой, чем корреляционная функция. Это преимущество особенно заметно при изучении линейных однородных преобразований стационарных процессов.

Линейное однородное преобразование, или, как еще говорят, линейный однородный фильтр, это преобразование, у которого импульсная характеристика зависит только от разности моментов времени:  $G(s, t) = G(t - s)$ . Иными словами, однородность линейного преобразования означает, что реакции на входные сигналы, получающиеся один из другого с помощью сдвига по времени, отличаются только сдвигом по времени. Рассмотрим, как проходит через линейный однородный фильтр  $G$  периодический сигнал  $u_\lambda(t) = e^{i\lambda t}$  ( $\lambda$  — параметр). Если  $G(t - s)$  — импульсная характеристика системы  $G$ , то сигнал на выходе

$$Gu_\lambda(t) = \int_{-\infty}^{\infty} G(t - s)e^{i\lambda s} ds = e^{i\lambda t} \int_{-\infty}^{\infty} G(-v)e^{i\lambda v} dv.$$

Таким образом, сигнал на выходе  $Gu_\lambda(t)$  только множителем  $g(\lambda) = \int_{-\infty}^{\infty} G(-v)e^{i\lambda v} dv$ , не зависящим от  $t$ , отличается от входного сигнала:  $Gu_\lambda(t) = g(\lambda)u_\lambda(t)$ .

Функция  $g(\lambda)$  называется частотной характеристикой системы  $G$ . Она описывает, как преобразуются при прохождении через фильтр  $G$  гармонические колебания с частотой  $\lambda$ .

Если на вход линейного фильтра  $G$  с частотной характеристикой  $g(\lambda)$  поступает сигнал  $h(t) = \sum c_k e^{i\lambda_k t}$ , то в силу линейности на выходе получим  $Gh(t) = \sum c_k g(\lambda_k) e^{i\lambda_k t}$ . Таким образом, за прохождением сигнала удобно наблюдать, если он представлен в виде суммы функций вида  $e^{i\lambda_k t}$ . Спектральное раз-



ложение стационарного процесса как раз и есть представление такого сорта. Так что, если на вход системы  $G$  поступает стационарный процесс  $x_t = \int_{-\infty}^{\infty} e^{i\lambda t} z(d\lambda)$ , то выход  $y_t = Gx_t = \int_{-\infty}^{\infty} e^{i\lambda t} g(\lambda) z(d\lambda)$ .

Если входной процесс имел спектральную плотность  $f(\lambda)$ , то на выходе получится стационарный процесс со спектральной плотностью  $f(\lambda) = |g(\lambda)|^2$ .

Процессы, встречающиеся в прикладных задачах, часто можно представить в виде  $x_t = m(t) + \xi_t$ , где  $m(t)$  — некоторая неслучайная функция, а  $\xi_t$  — стационарный процесс. Такой процесс  $x_t$  превращается в стационарный, если предварительно отфильтровать функцию  $m(t)$  (ее называют трендом процесса  $x_t$ ). Для этого процесс  $x_t$  пропускают через специальным образом подобранные линейные фильтры. Действие этих фильтров основано на том обстоятельстве, что обычно случайная составляющая  $\xi_t$  содержит частоты более высокие, чем те, которые входят в Фурье-разложение функции  $m(t)$ . Поэтому, производя усреднение процесса  $x_t$  по некоторой окрестности точки  $t_0$ , мы получим число, близкое к  $m(t_0)$ . Чем больше отличается спектральный состав тренда  $m(t)$  и случайной составляющей  $\xi_t$ , тем легче выделить тренд из процесса  $x_t$ . После того как отфильтрована низкочастотная составляющая — тренд процесса  $x_t$ , приступают к оценке спектральной плотности  $f(\lambda)$  процесса  $\xi_t$ . Это делается тоже путем пропускания процесса  $\xi_t$  через специальным образом подобранные линейные фильтры. Путем подбора соответствующих линейных фильтров решается задача также об оптимальном (в смысле минимума среднеквадратичной ошибки) прогнозе стационарных гауссовских процессов.

Стационарные случайные процессы давно находят применение в задачах управления. Этому вопросу посвящено большое количество книг, написанных для читателей с самой различной математической подготовкой. Назовем, например, книгу В. С. Пугачева [3]. В связи с этим основное внимание в нашей брошюре мы уделим другим классам случайных процессов, использование которых в задачах управления еще не нашло столь широкого отражения в литературе.

## § 4. Марковские процессы

Простейшим примером случайного процесса является последовательность независимых случайных величин  $\xi_1, \dots, \xi_n, \dots$ . В этом случае значение, которое принимает процесс в момент  $n + 1$ , не зависит от значений в предыдущий момент времени, как, впрочем, и во все другие моменты времени. Для случайных

последовательностей общего вида это, конечно, не так. Если мы знаем, какое значение приняла величина  $\xi_n$ , то мы тем самым имеем некоторую информацию относительно  $\xi_{n+1}$ , так как величины  $\xi_n$  и  $\xi_{n+1}$ , вообще говоря, зависимы. Если, кроме  $\xi_n$ , мы знаем  $\xi_{n-1}$  то информация о  $\xi_{n+1}$  еще увеличивается. Увеличение наших знаний о поведении процесса до момента  $n$  приводит, вообще говоря, к тому, что увеличивается информация о  $\xi_{n+1}$ .

С этой точки зрения наиболее простым и естественным обобщением последовательностей независимых величин являются марковские процессы. Это, грубо говоря, такие процессы, в которых знание значения процесса в момент  $n$  содержит в себе всю информацию о будущем ходе процесса, которую только можно извлечь из поведения процесса до момента  $n$ . Иными словами, для процессов этого класса, при фиксированном положении в настоящий момент, будущий ход процесса и его изменение до настоящего момента независимы.

Для того чтобы уточнить это определение, рассмотрим последовательность случайных величин  $\xi_1, \dots, \xi_n, \dots$ , принимающих лишь конечное число значений  $\{e_1, e_2, \dots, e_N\} = \mathcal{E}$ . Элементы этого множества называют состояниями, а само множество  $\mathcal{E}$ -фазовым пространством. Если  $\xi_k = e_i$ , а  $\xi_{k+1} = e_j$ , то будем говорить, что в момент  $k + 1$  произошел переход из состояния  $e_i$  в состояние  $e_j$ .

Условие независимости прошлого и будущего при фиксированном настоящем означает, что выполняется следующее соотношение между условными вероятностями

$$P\{\xi_{n+1} = e_j / \xi_n = e_i\} = P\{\xi_{n+1} = e_j / \xi_1 = e_{k_1}, \xi_2 = e_{k_2}, \dots, \xi_n = e_i\}.$$

Иными словами, вероятность попасть в момент  $n + 1$  в состояние  $e_j$ , если в момент  $n$  процесс находился в  $e_i$  не зависит от того, как вела себя траектория до момента  $n$ .

Случайные последовательности, для которых выполняется такое условие, называются цепями Маркова. Важнейшей характеристикой такой цепи являются вероятности  $p_{ij} = P\{\xi_{n+1} = e_j / \xi_n = e_i\}$  — вероятности перехода за один шаг из состояния с номером  $i$  в состояние с номером  $j$ . Вообще говоря, эти вероятности могут зависеть еще от момента времени  $n$ . Мы ограничимся рассмотрением однородных цепей Маркова, то есть таких цепей, у которых переходные вероятности  $p_{ij}$  не зависят от  $n$ .

Таким образом, с каждой цепью Маркова можно связать квадратную матрицу  $P = \{p_{ij}\}$  размерности, равной числу  $N$  возможных состояний цепи. Элементы матрицы  $P$  неотрицательны; сумма элементов каждой строки этой матрицы равна 1, так как эта сумма представляет собой вероятность куда-либо перейти за один шаг или остаться на месте. Такая матрица  $P$  называется стохастической.

Предположим, что нам известно распределение случайной величины  $\xi_k : P\{\xi_k = e_i\} = q_i^{(k)}$ ,  $i = 1, 2, \dots, N$ . Обозначим  $q^{(k)}$  вектор-строку  $(q_1^{(k)}, \dots, q_N^{(k)})$ . Вычислим  $q^{(k+1)} = (q_1^{(k+1)}, \dots, q_N^{(k+1)})$  — распределение цепи в следующий момент времени. Учитывая, что в состоянии  $e_j$  из состояния  $e_i$  за один шаг можно попасть с вероятностью  $p_{ij}$ , а  $P\{\xi_k = e_i\} = q_i^{(k)}$ , получим:

$$q_j^{(k+1)} = P\{\xi_{k+1} = e_j\} = \sum_{i=1}^N P\{\xi_k = e_i\} P_{ij} = \sum_{i=1}^N q_i^{(k)} p_{ij}.$$

В векторной форме это равенство можно записать так:  $q^{(k+1)} = q^{(k)}P$ . (Матрицу  $P$  мы применяем к вектору справа, так как  $q^{(k)}$  — вектор-строка). Если мы хотим узнать распределение  $q^{(k+2)}$  в момент  $k+2$ , то, последовательно применяя предыдущую формулу, получим:  $q^{(k+2)} = q^{(k+1)}P = q^{(k)}P^2$ ; распределение через  $l$  шагов дается формулой:  $q^{k+l} = q^{(k)}P^l$ . В частности, если  $q^{(0)}$  — распределение в начальный момент, то распределение через  $n$  шагов  $q^{(n)} = q^{(0)}P^n$ . Таким образом, матрица  $P$  описывает эволюцию распределений величин  $\xi_n$ .

Очень важным свойством марковских цепей является свойство эргодичности. Это свойство состоит в том, что с ростом  $n$  распределение  $q^{(n)}$  величин  $\xi_n$  сходится к некоторому распределению  $q = (q_1, q_2, \dots, q_N)$ , не зависящему от начального распределения:  $q = \lim_{n \rightarrow \infty} q^{(n)}$ . Для того чтобы цепь была эргодической, необходимо на ее переходную матрицу наложить некоторые условия. Например, достаточно потребовать, чтобы все элементы матрицы  $P$  были отличны от нуля. Можно дать и менее ограничительные условия. Чтобы найти предельное распределение  $q$ , применим к равенству  $q = \lim_{n \rightarrow \infty} q^{(n)}$  матрицу  $P$ . Тогда получим:

$$qP = \left( \lim_{n \rightarrow \infty} q^{(n)} \right) P = \lim_{n \rightarrow \infty} q^{(n)} P = \lim_{n \rightarrow \infty} q^{(n+1)} = q.$$

Таким образом, предельное распределение  $q$  удовлетворяет системе линейных алгебраических уравнений  $qP = q$ , то есть является собственным вектором матрицы  $P$  с собственным значением 1. Если цепь эргодична, то такой собственный вектор с точностью до множителя единственный. Этот множитель определяется из условия  $\sum_{i=1}^N q_i = 1$  (числа  $q_1, q_2, \dots, q_N$  должны образовывать распределение вероятностей). Если в качестве начального распределения марковской цепи выбрать вектор предельного распределения  $q$ , то  $q^{(n)} = qP^n = q$  т.е. распределение такой цепи не меняется со временем, и последовательность  $\xi_0, \xi_1, \dots, \xi_n, \dots$  в этом случае представляет собой стационарный процесс.

В качестве примера рассмотрим так называемую задачу игрушечных дел мастера. Этот пример заимствован из книги Р. А. Ховарда «Динамическое программирование и марковские процессы». Игрушечных дел мастер может находиться в одном из двух состояний, Первое состояние — если игрушка, которую делает мастер, получает большой спрос. Второе состояние — если игрушка не найдет спроса. Раз в неделю мастер должен принять решение, продолжать ли ему на следующей неделе делать старую игрушку или начать делать новую. Предположим, что если игрушка имеет спрос, т. е. мастер находится в состоянии 1, то с вероятностью  $1/2$  на следующей неделе этот спрос сохранится и с вероятностью  $1/2$  упадет. Будучи в состоянии 2, т. е. выпуская плохую игрушку, мастер на следующей неделе начнет выпускать новую игрушку. Однако придумать и сделать новую игрушку, на которую будет большой спрос, не так-то просто. Поэтому с вероятностью  $3/5$  мастер на следующей неделе останется в состоянии 2 и только с вероятностью  $2/5$  перейдет в первое состояние. Таким образом, в нашем случае имеется два состояния и элементы матрицы  $P$  таковы:  $p_{11} = p_{12} = \frac{1}{2}$ ,  $p_{21} = \frac{2}{5}$ ,  $p_{22} = \frac{3}{5}$ .

Предположим, что в начальный момент мастер выпускал игрушку, имеющую спрос. Это значит, что начальное распределение имеет вид  $q^{(0)} = (1, 0)$ . Тогда, через одну неделю  $q^{(1)} = q^{(0)}P = (\frac{1}{2}, \frac{1}{2})$ . Через две недели распределение будет таково:  $q^{(2)} = q^{(0)}P^2 = (\frac{9}{20}, \frac{11}{20})$ , т. е. уже через две недели более вероятным будет состояние 2. Если мы продолжим вычисление распределений  $q^{(n)}$ , то заметим, что они приближаются к распределению  $q = (\frac{4}{9}, \frac{5}{9})$ . Это и есть предельное распределение для нашей цепи.

Если в начальный момент мастер выпускал неудачную игрушку, то через некоторое время распределение тоже приблизится к вектору  $q = (\frac{4}{9}, \frac{5}{9})$ . Легко проверить, что  $q$  есть собственный вектор матрицы  $P$  с собственным значением 1:  $q = qP$ . Приближение к предельному распределению происходит экспоненциально быстро.

Переходим теперь к рассмотрению марковских процессов с непрерывным временем. Сначала остановимся на процессах  $x_t(\omega)$ , которые принимают конечное или счетное число значений  $e_1, e_2, e_3, \dots, e_N, \dots$  В таких процессах переход из одного состояния в другое происходит уже не в целые моменты времени, а в произвольные, вообще говоря, случайные моменты. Наиболее удобными характеристиками здесь служат уже не

вероятности переходов, а так называемые интенсивности переходов. Интенсивность  $a_{ij}$  перехода из состояния  $e_i$  в состояние  $e_j$ , отличное от  $e_i$ , определяется следующим образом. За малый интервал времени  $\Delta$  процесс, который находится в момент  $t$  в состоянии  $e_i$ , может совершить переход в состояние  $e_j$  с вероятностью  $a_{ij} \cdot \Delta + o(\Delta)$ . Вероятность двух или более переходов за время  $\Delta$  есть  $o(\Delta)$ . (Здесь, как обычно,  $o(\Delta)$  есть такая функция от  $\Delta$ , что  $\lim_{\Delta \rightarrow 0} \frac{o(\Delta)}{\Delta} = 0$ ). Интенсивности  $a_{ij}$  представляют собой неотрицательные числа. Постоянство интенсивностей, также как в случае цепей независимость от времени элементов переходной матрицы, представляет собой условие однородности процесса по времени. Марковский характер процесса состоит в том, что вероятностные характеристики поведения процесса в произвольном состоянии  $e_i$  не зависят от того, каким образом процесс в это состояние попал. Интенсивности  $a_{ij}$  определены при  $i \neq j$ . Положим  $a_{jj} = 1 - \sum_{i, i \neq j} a_{ij}$ , и обозначим  $A$  матрицу с элементами  $a_{ij}$ . Пусть  $\pi(t) = (\pi_1(t), \dots, \pi_N(t))$  — распределение процесса  $x_t$  в момент  $t$ :  $\pi_i(t) = P\{x_t = e_i\}$ . Составим для функций  $\pi_i(t)$  систему дифференциальных уравнений. Для этого заметим, что в момент  $t + \Delta$  процесс может находиться в состоянии  $e_j$  только, если он находился в этом состоянии в момент  $t$  и в течение времени  $\Delta$  после момента  $t$  оставался на месте или если в момент  $t$  он находился в каком-то состоянии  $e_i$  и за время  $\Delta$  перешел из  $e_i$  в  $e_j$ . Все остальные способы попасть в момент  $t + \Delta$  в состояние  $e_j$  имеют вероятность  $o(\Delta)$ . Учитывая эти соображения, для  $\pi_j(t + \Delta)$  можно выписать соотношение

$$\pi_j(t + \Delta) = \pi_j(t) \left( 1 - \sum_{i, i \neq j} a_{ij} \Delta \right) + \sum_{i, i \neq j} \pi_i(t) a_{ij} \Delta + o(\Delta).$$

Отсюда получим

$$\frac{\pi_j(t + \Delta) - \pi_j(t)}{\Delta} = \sum_i \pi_i(t) a_{ij} + \frac{o(\Delta)}{\Delta}.$$

Если перейти теперь к пределу при  $\Delta \rightarrow 0$ , то мы придем к дифференциальному уравнению для  $\pi_j(t)$ . В векторной форме эти уравнения принимают вид

$$\frac{d\pi(t)}{dt} = \pi(t)A.$$

Чтобы выделить единственное решение этой системы уравнений, необходимо еще задать начальное условие:  $\pi(0) = \pi_0$ . Здесь  $\pi_0$  — начальное распределение нашего процесса:  $\pi_0 = (\pi_1^0, \pi_2^0, \dots, \pi_N^0)$ ,  $\pi_k^0 = P\{x_0 = e_k\}$ . Эта система полностью описывает эволюцию распределений марковского процесса  $x_t(\omega)$ . При небольших

дополнительных ограничениях процесс  $x_t$  будет эргодическим, т. е. с ростом  $t$  распределение  $\pi(t)$  будет стремиться к пределу  $\pi$ , не зависящему от начального распределения. Это предельное распределение  $\pi$  можно найти из условия стационарности:  $\pi A = \frac{d\pi}{dt} = 0$ .

Рассмотрим пример. Пусть некоторый станок может находиться в одном из двух состояний — рабочем (состояние 1) или нерабочем (состояние 2). Если в настоящее время станок работает, то он может сломаться за малое время  $\Delta$  с вероятностью, скажем,  $5\Delta + o(\Delta)$ . Если же станок не работает, то с вероятностью  $4\Delta + o(\Delta)$  он будет за время  $\Delta$  отремонтирован. Таким образом, мы получаем процесс с двумя состояниями. Матрица  $A$  в этом случае имеет вид

$$A = \begin{pmatrix} -5 & 5 \\ 4 & -4 \end{pmatrix}.$$

Из условий задачи можно вывести, что время  $\tau$  между двумя поломками имеет показательное распределение:  $P\{\tau > t\} = e^{-5t}$ . Показательное распределение имеет и случайная величина  $\beta$  — время пребывания во втором состоянии, т. е. время, необходимое для ремонта:  $P\{\beta > t\} = e^{-4t}$ . При этом среднее время между поломками  $M\tau = \frac{1}{5}$ , а  $M\beta = \frac{1}{4}$ . Так что практически интенсивности могут назначаться на основе оценки средних значений случайных величин  $\tau$  и  $\beta$ . Предположим, что в начальный момент станок работал. Это значит, что к системе дифференциальных уравнений  $\frac{d\pi(t)}{dt} = \pi(t)A$  нужно приписать начальное условие  $\pi_0 = (1, 0)$ . Решая это уравнение, получим

$$\pi_1(t) = \frac{4}{9} + e^{-9t} \frac{5}{9};$$

$$\pi_2(t) = \frac{5}{9} - e^{-9t} \frac{5}{9}.$$

Если при  $t = 0$  станок не работал, то распределение  $\bar{\pi}(t)$  в момент  $t$  будет иметь вид:  $\bar{\pi}_1(t) = \frac{4}{9} - \frac{4}{9}e^{-9t}$ ,  $\bar{\pi}_2(t) = \frac{5}{9} + \frac{4}{9}e^{-9t}$ . Из этих формул мы видим, что с ростом  $t$  распределения  $\pi(t)$  и  $\bar{\pi}(t)$  экспоненциально быстро приближаются к распределению  $\pi = (\frac{4}{9}, \frac{5}{9})$ . Это и есть не зависящее от начального состояния предельное распределение нашего процесса.

Прежде чем переходить к марковским процессам с непрерывными траекториями, остановимся еще на так называемом пуассоновском процессе. Этот процесс очень часто возникает в приложениях. Пуассоновский процесс  $S(t)$  это марковский про-

цесс с непрерывным временем, пространство состояний которого состоит из всех целых неотрицательных чисел. Матрица интенсивностей этого процесса такова, что  $a_{i,i+1} = \lambda$  для всех  $i = 0, 1, 2, 3, \dots$ ; все другие переходы имеют интенсивность 0. Таким образом, траектории пуассоновского процесса не могут убывать, а могут только претерпевать скачки величины 1 вправо. Для вероятностей  $\pi_k(t)$  можно написать дифференциальные уравнения, аналогичные тем, которые мы писали в случае конечного числа состояний. Эти уравнения можно решить. В результате для  $S(t)$  получим пуассоновское распределение:  $\pi_k(t) = P\{S(t) = k\} = (\lambda t)^k \frac{e^{-\lambda t}}{k!}$ , если в начальный момент  $P\{S(0) = 0\} = 1$ . Можно доказать, что пуассоновский процесс будет процессом с независимыми приращениями. Ранее нам уже встречался пример процесса с независимыми приращениями. Это винеровский процесс, введенный в § 2. Винеровский процесс  $W_t$  — это гауссовский процесс, имеющий нулевое математическое ожидание и корреляционную функцию  $B(s, t) = \min(s, t)$ . Винеровский процесс имеет непрерывные траектории. Приращение винеровского процесса за время  $[t, t + h]$  не зависит от поведения процесса до момента  $t$ , поэтому при фиксированном  $W_t$  случайная величина  $W_{t+h} = W_t + (W_{t+h} - W_t)$  не зависит от течения процесса до момента  $t$ . Таким образом, винеровский процесс, как и все процессы с независимыми приращениями, является марковским процессом: при фиксированном настоящем будущее и прошлое независимы.

Другим примером непрерывного марковского процесса служит детерминированное движение с постоянной скоростью: если мы знаем положение движущейся таким образом частицы в настоящий момент, то для того, чтобы предсказать положение через какое-то время, нам не нужно знать, как проходило движение в прошлом. Оказывается, что, по существу, любой марковский процесс  $x_t$  с непрерывным временем и непрерывными траекториями в малом устроен как линейная комбинация этих двух основных процессов — винеровского процесса и детерминированного движения:

$$dx_t = b(x_t)dt + \sigma(x_t)dw_t.$$

Иными словами, приращение  $dx_t$  процесса  $x_t$  за малое время  $dt$  есть линейная комбинация приращения винеровского процесса  $dw_t$  за время  $dt$  и приращения  $dt$  детерминированного движения со скоростью 1. Коэффициенты этой линейной комбинации зависят от положения процесса  $x_t$  в момент  $t$ . Если разделить обе части последнего соотношения на  $dt$  и ввести в рассмотрение  $\dot{W}_t$  — процесс белого шума (см. § 2), то мы получим следующее дифференциальное уравнение для определения процесса  $x_t$ :  $\dot{x}_t = b(x_t) + \sigma(x_t)\dot{w}_t$ ;  $\dot{x}_t = \frac{dx_t}{dt}$ .

Чтобы выделить единственное решение этого уравнения, следует задать начальное условие  $x_0 = x$ . Так как в правой части этого уравнения стоит случайный процесс, то решение тоже будет представлять собой случайный процесс. Этот процесс будет марковским. Это следует из теоремы единственности для дифференциального уравнения процесса  $x_t$  и независимости приращений винеровского процесса.

Коэффициент  $b(x)$ , характеризующий интенсивность детерминированного движения вблизи точки  $x$ , называется коэффициентом переноса, коэффициент  $a(x) = \sigma^2(x)$  — коэффициентом диффузии. Процессы такого сорта хорошо описывают физическую диффузию, поэтому их называют диффузионными. Однако они нередко возникают в самых различных ситуациях. Непрерывные марковские процессы оказываются полезными при изучении систем автоматического управления, находящихся под действием шумов, в задачах генетики; такие процессы могут служить хорошим приближением в некоторых задачах массового обслуживания.

Знание локальных характеристик марковского процесса  $x_t$  — коэффициентов  $b(x)$  и  $a(x)$  дает возможность вычислить различные вероятности и средние значения некоторых функционалов от траекторий процесса  $x_t$ . Пусть, например,  $f(x)$  — ограниченная гладкая функция на прямой, и мы хотим вычислить функцию  $u(t, x) = M_x f(x_t)$ . Значок  $x$  у математического ожидания означает, что рассматриваются траектории  $x_t$ , начинающиеся в точке  $x : x_0 = x$ . Оказывается, что функция  $u(t, x)$  есть решение следующей задачи

$$\frac{\partial u}{\partial t} = \frac{a(x)}{2} \frac{\partial^2 u}{\partial x^2} + b(x) \frac{\partial u}{\partial x} \quad \text{при } t > 0, \\ u(0, x) = f(x).$$

Этими соотношениями функция  $u(t, x)$  определяется однозначно. Дифференциальное уравнение, которому удовлетворяет функция  $u(t, x)$ , называется обратным уравнением Колмогорова.

Если мы хотим найти плотность распределения  $m(t, x)$  процесса  $x_t$  в момент  $t$ , при условии, что в момент ноль процесс был распределен с известной плотностью  $m_0(x)$ , то нужно решить другое уравнение — прямое уравнение Колмогорова:

$$\frac{\partial m(t, x)}{\partial t} = \frac{1}{2} \frac{\partial^2}{\partial x^2} (a(x)m(t, x)) - \frac{\partial}{\partial x} (b(x)m(t, x))$$

с начальным условием  $m(0, x) = m_0(x)$ . В частности, функция  $m(x)$  будет плотностью стационарного распределения процесса  $x_t$ , если эта функция удовлетворяет прямому уравнению Колмогорова, в котором слева стоит ноль:

$$0 = \frac{1}{2} \frac{d^2}{dx^2} (a(x)m(x)) - \frac{d}{dx} (b(x)m(x)).$$



Многомерные диффузионные процессы определяются аналогичным образом. Если  $W_t^1, W_t^2$  — два независимых винеровских процесса, то векторное дифференциальное уравнение

$$\dot{x}_t = b(x_t) + \sigma(x_t)\dot{W}_t$$

определяет двумерный процесс  $x_t = (x_t^1, x_t^2)$ . Здесь  $x = (x^1, x^2)$ ,  $b(x) = (b^1(x), b^2(x))$ ,  $W_t = (W_t^1, W_t^2)$ ,  $\sigma(x) = \{\sigma^i_j(x)\}$  матрица порядка 2. Вектор  $b(x)$  называется переносом, матрица  $a(x) = \{a^{ij}(x)\} = \sigma(x)\sigma^*(x)$  матрицей диффузии (через  $\sigma^*(x)$  обозначена транспонированная к  $\sigma(x)$  матрица).

С помощью этих локальных характеристик так же, как в одномерном случае, можно написать дифференциальные уравнения для средних значений различных случайных величин, связанных с процессом  $x_t$ . Так, если  $D$  есть область на плоскости с границей  $\Gamma$ , то среднее время  $u(x)$ , необходимое процессу  $x_t$  для достижения границы, исходя из точки  $x$ , лежащей внутри области  $D$ , есть решение следующей задачи

$$\frac{1}{2} \sum_{i,j=1}^2 a^{ij}(x) \frac{\partial^2 u}{\partial x^i \partial x^j} + \sum_{i=1}^2 b^i(x) \frac{\partial u}{\partial x^i} = -1 \quad \text{при } x \text{ из } D,$$

$u(x) = 0$  для  $x$ , принадлежащих  $\Gamma$ . Мы рассмотрим некоторые задачи, приводящие к управлению марковскими процессами в § 7.

## § 5. Основные понятия теории оптимального управления

В этом параграфе мы введем некоторые понятия, которые оказываются полезными при рассмотрении задач управления, независимо от конкретной природы этих задач. Сначала мы рассмотрим детерминированные системы. Затем изучим, какие уточнения и дополнения следует внести, чтобы рассматривать задачи управления стохастическими системами.

Пусть движение некоторого объекта описывается системой дифференциальных уравнений

$$\dot{x} = b(x).$$

Здесь  $x = (x^1, x^2)$  — точка плоскости;  $b(x) = (b^1(x), b^2(x))$ , функции  $b^1(x)$  и  $b^2(x)$  предполагаются ограниченными во всей плоскости и гладкими; точка над функцией  $x_t$  означает дифференцирование по времени. Решения этой системы называют траекториями.

Предположим, что в момент  $t = 0$  траектория  $x_t$  исходит из точки  $x$  и нам хотелось бы, чтобы эта траектория достигла некоторого множества  $G$ , не содержащего  $x$ . Как известно, су-

существует единственное решение нашей системы, начинающееся в точке  $x$ . Это решение при фиксированной функции  $b(x)$ , вообще говоря, может никогда не попасть в множество  $G$ . Если функция  $b(x)$  фиксирована раз и навсегда, то, собственно, никакой задачи управления нет. Предположим, однако, что функция  $b(x)$  зависит еще от параметра  $u$ , скалярного или векторного,  $b = b(x, u)$ , и мы можем изменять правые части системы дифференциальных уравнений, выбирая в каждый момент те или иные значения параметра  $u$ .

Параметр  $u$  можно выбирать, вообще говоря, не произвольно. Физически выбор параметра  $u$  означает выбор того или иного положения «рулей», количества топлива, подающегося в двигатель, или какой-либо другой физической величины, влияющей на движение. На все эти величины обычно накладываются определенные ограничения, поэтому естественно считать, что в каждый момент времени параметр  $u$  можно выбирать из какого-то ограниченного множества. Например, если  $u$  — скалярный параметр, ограничение на  $u$  может иметь вид  $|u| < 1$ . Если  $u$  — вектор, то ограничения могут накладываться отдельно на каждую компоненту или на все компоненты сразу. Множество значений, которые можно придавать параметру  $u$ , мы обозначим  $D$  и будем называть областью управления.

Заданием области  $D$  не исчерпываются, вообще говоря, все ограничения на управление. В каждый момент времени мы выбираем свое значение для параметра  $u$ , так что управлением у нас будет функция  $u(t)$ . Обычно еще накладываются ограничения на управление как на функцию от  $t$ . Эти ограничения могут быть самыми разными. Они могут носить характер требований регулярности функции  $u(t)$ . Например, требование, чтобы  $u(t)$  была кусочно-гладкой функцией. Более ограничительным будет требование, чтобы управление было кусочно-постоянной функцией; иногда дополнительно требуется, чтобы число разрывов (число переключений) не превосходило бы наперед заданного числа.

Ограничения на функцию  $u(t)$  могут носить другой характер. Так они могут иметь вид  $\int_0^T u(s) ds < \text{const}$ . Такие ограничения естественны, например, если  $u(s)$  — это скорость поступления топлива в двигатель, а общее количество топлива ограничено. Функции  $u(t)$ , принимающие в каждый момент значения из области управления  $D$  и удовлетворяющие еще некоторым дополнительным ограничениям, вроде тех, которые приводились выше, называются допустимыми управлениями.

Если при всевозможных допустимых управлениях мы имеем достаточно обширный запас правых частей  $b(x, u)$ , то при некоторых  $u = u(t)$  траектории  $x_t^u$ , исходящие в момент  $t = 0$  из  $x$ , достигнут множества  $G$  (значок  $u$  у траектории означает, что она

является решением дифференциального уравнения с этим самым управлением  $u = u(t) : x_t^u = b(x_t^u, u(t))$ .

Таких допустимых управлений, которые приводят из  $x$  в  $G$ , может быть много. Обычно требования на выбор управления не ограничиваются только тем, что  $x_t^u$  должна достигнуть множества  $G$ . Необходимо выбрать  $u(t)$  так, чтобы множество  $G$  достигалось наиболее выгоднейшим образом. Понятие о том, какой из двух способов лучше, выгоднее, конечно, требует уточнения, исходя из постановки реальной задачи. Например, мы можем предпочесть то из двух управлений  $u_1(t), u_2(t)$ , при котором время, необходимое для достижения области  $G$ , меньше. Такая задача называется задачей на оптимальное быстроедействие.

Более общая постановка такова; для каждой траектории, достигающей  $G$  в момент  $T(u)$ , вычисляется интеграл

$$l|u| = \int_0^{T(u)} f(x_t^u, u(t)) dt,$$

где  $f(x, u)$  — некоторая наперед заданная функция. Требуется выбрать допустимое управление таким образом, чтобы этот интеграл принимал наименьшее значение. В частности, если  $f(x, u) = 1$ , то мы получаем задачу на оптимальное быстроедействие. В общем случае  $f(x, u) dt$  можно интерпретировать как плату, которую мы выплачиваем за пребывание в течение времени  $dt$  вблизи точки  $x$ , если в это время мы используем управление  $u$ . Весь интеграл  $l[u]$  в таком случае показывает наши затраты при использовании управления  $u(t)$ .

Допустимое управление  $\hat{u}(t)$  называется оптимальным, если траектория  $x_t^{\hat{u}}$ , начинающаяся в момент  $t = 0$ , в точке  $x$  приводит в множество  $G$ , и при этом никакое другое управление, приводящее из  $x$  в  $G$ , не доставляет интегралу  $l[u]$  значения меньшего, чем  $l[\hat{u}]$ .

Вообще говоря, оптимальных управлений может быть несколько. С другой стороны, нередко бывают задачи, в которых среди допустимых управлений нет оптимального. В таких случаях полезным оказывается понятие  $\epsilon$ -оптимального управления. Применительно к рассматриваемой задаче управление  $\tilde{u}(t)$  называется  $\epsilon$ -оптимальным, если оно переводит  $x$  в  $G$ , а значение  $l[\tilde{u}]$  не более чем на  $\epsilon$  превосходит значение интеграла  $l[u]$  при любом другом допустимом управлении  $u(t)$ , переводящем  $x$  в  $G$ .  $\epsilon$ -оптимальные управления существуют при очень слабых дополнительных предположениях. С точки зрения приложений вполне можно ограничиться  $\epsilon$ -оптимальными управлениями, так как они дают результат, близкий к оптимальному.

Если допустимые управления выделяются только условиями типа ограниченности и кусочной непрерывности, то оптимальное

управление  $\hat{u}(t)$  в задаче на быстроедействие, сформулированной в этом параграфе, имеет вид  $\hat{u}(t) = v(x_t)$ . Иными словами, оптимальное управление в момент  $t$  зависит только от того, в какой точке пространства в это время находится траектория. Поведение траектории до момента  $t$  не влияет на выбор оптимального управления в момент  $t$ .

Такое положение дел вполне естественно, так как если  $x$  — это положение оптимальной траектории в момент  $t_1$ , а  $y$  — положение в момент  $t_2$ , то оптимальное управление должно доставлять минимум времени, необходимого для перемещения из точки  $x$  в  $y$ . Значит, оптимальное управление не зависит от того, что было до попадания в  $x$ , и от того, что будет после выхода из  $y$ . Отсюда следует, что оптимальное управление в момент  $t$  зависит только от  $x_t$ . Функция  $v(x)$ , через которую выражается оптимальное управление, называется синтезирующей функцией, а задача построения этой функции — задачей синтеза оптимальных управлений. Соображение, которое мы использовали выше, — любая часть оптимальной траектории сама является оптимальной — есть частный случай так называемого принципа оптимальности Беллмана [6]. На основе подобных рассуждений можно решать многие задачи оптимального управления. Мы еще встретимся с этим принципом в дальнейшем. Другой метод решения оптимальных задач, подобных тем, которые рассматривались в этом параграфе, — принцип максимума Л. С. Понтрягина.

Переходим теперь к стохастическому случаю. Для определенности рассмотрим систему, которая получается из рассмотренной выше, при добавлении к правым частям некоторого случайного процесса:

$$\dot{x}_t = b(x_t, u(t)) + \zeta_t.$$

Здесь  $\zeta_t = (\zeta_t^1, \zeta_t^2)$  — двумерный случайный процесс; ниже мы делаем относительно него различные предположения.

Рассмотрение дифференциальных уравнений, правые части которых возмущаются случайными шумами, представляет собой одну из важных задач теории оптимального управления. В случае стохастической системы мы также должны прежде всего определить, какие управления являются допустимыми. Здесь положение сложнее, чем в детерминированном случае. Вместе с теми ограничениями, которые делались раньше, мы должны более четко указать, по каким наблюдениям выбирается управление в момент  $t$ . Так естественно управление в момент  $t$  вычислять только по ходу процесса до момента  $t$ . Иначе управление не будет физически осуществимым. В детерминированном случае оптимальное управление нередко выражается через синтезирующую функцию. В стохастических задачах тоже бывает так, что оптимальное управление в момент  $t$  зависит только от  $t$

и положения процесса  $x_t$  в момент  $t$ . В этом случае говорят, что существует марковское оптимальное управление. Так, если процессы  $\zeta_t^1$  и  $\zeta_t^2$  суть два независимых процесса белого шума (в этом случае процесс  $x_t = (x_t^1, x_t^2)$  — марковский), то в целом ряде задач управления случайным процессом  $x_t$  будут существовать марковские оптимальные стратегии.

В других случаях может быть так, что оптимальное управление в момент  $t$  существенно зависит не только от положения процесса  $x_t$  в момент  $t$ , но и от некоторых других функций, вычисляемых по поведению процесса  $x_t$  до момента  $t$ . Если можно указать конечное число функций  $f_t^1, f_t^2, \dots, f_t^r$ , вычисляемых по поведению процесса до момента  $t$ , так что оптимальное управление в момент  $t$  зависит только от этих функций  $f_t^1, f_t^2, \dots, f_t^r$ , то функции  $f_t^1, f_t^2, \dots, f_t^r$  называются достаточными статистиками. Знание достаточных статистик существенно упрощает задачу, оптимального управления.

Так же, как в детерминированном случае, в задачах управления стохастическими системами необходимо иметь способ, с помощью которого можно было бы судить, какое из двух управлений лучше. Для детерминированной системы мы можем использовать с этой целью, например, интеграл  $l[u]$ . В стохастическом случае  $l[u]$  будет случайной величиной. Здесь в качестве характеристики управления и можно принять среднюю величину платы:  $\bar{l}[u] = Ml[u]$ . Так или иначе, чтобы иметь возможность выбирать из двух управлений лучшее, мы должны задать способ вычислять по заданному управлению некоторое число — среднюю плату за это управление. Это число иногда также называют средним выигрышем, риском, функцией потерь для данного управления.

Следует заметить, что функция риска (плата, выигрыш) часто назначается довольно произвольно, поэтому, как правило, практический интерес представляют стратегии, которые являются оптимальными или близкими к оптимальным не только при заданной функции риска, но и при функциях риска близким к рассматриваемым. Иногда следует поступиться требованием оптимальности в пользу стратегий хороших, но не оптимальных, которые устойчивы относительно малых изменений условий задачи и функции риска.

В последующих параграфах мы разберем примеры, поясняющие понятия, которые мы здесь ввели.

## § 6. Управление созданием запасов

В этом параграфе мы рассмотрим задачи, связанные с созданием запасов. Эти задачи являются характерными представителями широкого круга проблем оптимального управления стохастическими системами.

Итак, пусть в целые моменты времени  $1, 2, \dots, n, \dots$  поступают заявки на некоторый товар. Размеры этих заявок носят случайный характер: спрос в момент  $k$  есть случайная величина  $\xi_k$ . Будем считать, что величины  $\xi_k$  при разных  $k$  независимы и одинаково распределены с плотностью  $p(x)$ , которую мы считаем положительной при  $x > 0$ . Для удовлетворения спроса мы имеем возможность запастись некоторым количеством товара заранее, т. е. в момент 0, и, кроме того, в каждый момент  $k$  можем сделать дополнительный заказ  $u(k)$ , который мы получим в момент  $k + 1$ . Товар, заказанный заранее, который нам доставляют через единицу времени после заказа, обходится в  $h$  денежных единиц за единицу товара. Если в какой-то момент спрос превышает имеющееся количество товаров, то мы имеем возможность получить товар тотчас же и тем самым удовлетворить спрос. Однако в этом случае товар поступает по цене  $H$  за единицу товара. Естественно,  $H > h$ . Наша задача состоит в выборе такой политики предварительных заказов  $u_0, u_1, \dots, u_{N-1}$  в течение  $N$  единиц времени, чтобы удовлетворить спрос с наименьшими общими затратами. Описанная схема — одна из простейших постановок задачи управления запасами.

Обозначим  $x_n$  количество товаров, оставшееся после удовлетворения  $n$ -й заявки  $\xi_n$ . Тогда  $x_{n+1} = x_n + u_n$ , если только правая часть этого равенства положительна. Если спрос  $\xi_{n+1}$  в момент  $n+1$  превышает количество товаров  $x_n + u_n$ , имеющихся к моменту  $n+1$ , то недостаток товаров покрывается за счет немедленной закупки по штрафной цене, и в этом случае  $x_{n+1} = 0$ . Если условиться обозначать  $f^+(x)$  функцию, равную  $f(x)$  при  $f(x) > 0$  и нулю при  $f(x) \leq 0$ , то связь  $x_{n+1}$  и  $x_n$  можно записать так:

$$x_{n+1} = (x_n + u_n - \xi_{n+1})^+. \quad (*)$$

При фиксированной последовательности  $u = (u_0, u_1, \dots, u_{N-1})$  предварительных закупок наши затраты составят следующую сумму:

$$h \sum_{k=0}^{N-1} u_k + H \sum_{k=0}^{N-1} (\xi_{k+1} - x_k - u_k)^+ = l_N[u].$$

Первое слагаемое этой суммы есть плата за предварительно заказанные товары, второе слагаемое — стоимость штрафных закупок. Вся сумма  $l_N[u]$  — случайная величина. В качестве характеристики последовательности  $u$  естественно принять среднее значение этой случайной величины  $\overline{l_N[u]} = Ml_N[u]$ . Наша цель состоит в выборе такого допустимого управления  $u = (u_0, u_1, \dots, u_{N-1})$ , чтобы минимизировать среднюю стоимость товаров за время  $[0, N]$ .

Поясним, какие управления считаются допустимыми. Прежде всего, все компоненты вектора  $u$  — неотрицательные величины,

так как заказ не может быть отрицательным. Далее, величина заказа в момент  $k - u_k$  вычисляется только по ходу процесса до момента  $k$  включительно, то есть по  $x_1, x_2, \dots, x_k$ . Это последнее требование очень существенно, ведь ход процесса после момента  $k$  определяется величинами  $\xi_{k+1}, \xi_{k+2}, \dots$ , которые до момента  $k$  невозможно определить, поэтому управление, которое вычисляется с учетом будущего, не может быть физически осуществлено.

Используя независимость величин  $\xi_1, \xi_2, \dots, \xi_n, \dots$ , нетрудно доказать, что в нашей задаче существует оптимальная марковская стратегия  $\hat{u} = (u_0, u_1, \dots, u_{N-1})$ . Это значит, что оптимальная величина заказа в  $k$ -й момент —  $u_k$  может быть найдена среди функций, зависящих только от  $k$  и  $x_k$ :  $\hat{u}_k = v(k, x_k)$  (конечно,  $\hat{u}_k$  зависит также от величин  $h, H, p(x)$  и числа этапов  $N$ ). Если мы используем управление такого вида, то последовательность  $x_n$ , определяемая равенством (\*), является марковским процессом: при фиксированном настоящем прошлое и будущее независимы. Эти соображения существенно упрощают задачу.

Переходим к вычислению оптимального управления. Для этого рассмотрим чуть более общую задачу. Будем допускать, что в начальный момент уже было некоторое количество  $x$  товаров. Обозначим  $f_n(x)$  ожидаемую стоимость товаров за  $n$  этапов при начальном количестве  $x$ , если используется оптимальная  $n$ -этапная стратегия. Таким образом,  $l[\hat{u}] = f_N(0)$ . Используя принцип оптимальности, для функций  $f_n(x)$  можно получить следующие соотношения:

$$f_1(x) = \min_{y \geq x} \left[ h(y - x) + H \int_y^\infty (s - y)p(s)ds \right],$$

$$f_{n+1}(x) = \min_{y \geq x} \left[ h(y - x) + H \int_y^\infty (s - y)p(s)ds + f_n(0) \int_y^\infty p(s)ds + \right. \\ \left. + \int_0^y f_n(y - s)p(s)ds \right]. \quad (**)$$

Поясним эти равенства. Предположим, что заказ в момент 0 есть  $u_0 = y - x$ . Его стоимость  $h(y - x)$  — первое слагаемое в квадратных скобках обеих равенств (\*\*). Если спрос  $\xi_1$  превзойдет  $y$ , то мы должны еще приобрести  $\xi_1 - y$  единиц товара по штрафной стоимости; на это в среднем мы тратим

$$H \int_y^\infty (s - y)p(s)ds.$$

Таким образом, в квадратных скобках первого равенства стоит средняя стоимость товаров, если  $u_0 = y - x$ . Чтобы получить

минимальную среднюю стоимость в одноэтапном процессе  $f_1(x)$ , необходимо выбрать  $y - x$  или  $y$ , так как  $x$  фиксировано, из условия минимума средних расходов за первый этап. Это и записано в первом из равенств (\*\*).

Если  $n > 1$ , то необходимо учесть расходы на дальнейших  $n - 1$  этапах. После выполнения первой заявки у нас с вероятностью  $P\{\xi_1 > y\} = \int_y^\infty p(s)ds$  не останется товаров, и тогда, при использовании оптимальной  $(n - 1)$ -этапной стратегии в течение оставшихся  $n - 1$  единицы времени, мы в среднем истратим  $f_{n-1}(0)$  денежных единиц. Если же  $y - \xi_1 > 0$ , то средние расходы за оставшиеся  $n - 1$  этапов при оптимальной стратегии составят  $\int_0^y f_{n-1}(y - s)p(s)ds$ . Значит, в квадратных скобках второго из равенств (\*\*) стоит средняя стоимость закупок, если в момент 0 мы заказали  $y - x$  единиц товара, а в дальнейшем пользовались оптимальной стратегией. Ясно, что для получения  $f_n(x)$  нам остается выбрать оптимальным образом начальную закупку  $y - x$  (принцип оптимальности) или, так как  $x$  фиксировано, выбрать величину  $y$ . Именно это требование и составляет содержание второго соотношения (\*\*).

Итак, мы получили для  $f_n(x)$  функциональные уравнения. Прежде чем приступить к их исследованию, сделаем простое замечание относительно соотношений вида

$$u(x) = \min_{y: y \geq x} G(x, y).$$

Если минимум достигается во внутренней точке полупрямой  $y \geq x$  и  $G(x, y)$  — гладкая функция, то при минимизирующем значении  $y$  частная производная  $G_y(x, y) = 0$ . Это последнее равенство определяет функцию  $y^* = y^*(x)$ , которая доставляет  $\min G(x, y)$  по второму аргументу на множестве  $y \geq x$ . При этом  $u(x) = G(x, y^*)$  и  $u'(x) = G'_x + G'_y \frac{dy^*}{dx} = G'_x(x, y^*)$ , так как  $G'_y(x, y^*) = 0$ . Итак, функция  $y^* = y^*(x)$ , доставляющая минимум  $G(x, y)$  удовлетворяет уравнению  $G'_y(x, y^*) = 0$ , а  $u'(x) = G'_x(x, y^*)$ .

Применим это замечание к нашей задаче. Покажем, что существует убывающая последовательность  $a_0, a_1, \dots, a_{N-1}$  такая, что оптимальная стратегия на  $k$ -м этапе  $\hat{u}_k$  устроена так: если количество товаров  $x$ , оставшееся после удовлетворения  $(k - 1)$ -го требования, больше  $a_k$ , то ничего заказывать не надо ( $\hat{u}_k = 0$ ), если же  $a_k - x > 0$ , то нужно сделать заказ  $\hat{u}_k = a_k - x$ .

Начнем рассмотрение с последнего этапа. Если после удовлетворения  $N - 1$  требования остается  $x$  единиц товара, то при оптимальном заказе  $\hat{u}_{N-1}$  наши средние затраты на последнем этапе составят  $f_1(x)$ . Функция  $f_1(x)$  удовлетворяет первому из



равенств (\*\*). Применяя к этому равенству приведенное выше замечание, получим, что минимум достигается или при  $y = x$ , или же при некотором  $y$ , для которого частная производная по  $y$  от выражения, стоящего в квадратных скобках, обращается в нуль. Таким образом, для минимизирующего значения  $y$  получим уравнение

$$\frac{\partial}{\partial y} \left[ h(y - x) + H \int_y^{\infty} (s - y)p(s)ds \right] = h - H \int_y^{\infty} p(s)ds = 0.$$

Легко проверить, что это уравнение имеет единственное решение. Обозначим его  $y = a_{N-1}$ . Это число и есть то критическое количество товаров, которое нужно иметь перед последним этапом, чтобы минимизировать затраты на последнем этапе. Поэтому  $\hat{u}_{N-1} = (a_{N-1} - x)^+$ , где  $x$  — количество товаров, оставшееся после удовлетворения предпоследней заявки. Используя наше замечание, можно выписать выражение для  $f'_1(x)$ .

Для определения  $a_{N-2}$  нужно рассмотреть последние два этапа. Оптимальные средние затраты на последних двух этапах  $f_2(x)$  удовлетворяют соотношению (\*\*) при  $n = 2$ . Минимум в этом соотношении достигается или при  $y = x$ , или при  $y = a_{N-2}$ , где  $a_{N-2}$  — корень уравнения

$$h - H \int_y^{\infty} p(s)ds - \int_0^y f'_1(y - s)p(s)ds = 0.$$

Можно показать, что это уравнение также имеет единственное решение  $y = a_{N-2}$ ; при этом  $a_{N-2} > a_{N-1}$ . Остальные числа  $a_{N-k}$  определяются последовательно из уравнений:

$$h = H \int_y^{\infty} p(s)ds - \int_0^y f'_{N-k+1}(y - s)p(s)ds.$$

С увеличением  $k$  эти числа возрастают.

Таким образом, оказалось, что в нашей задаче существует простая оптимальная стратегия предварительных заказов. На основе интуитивных соображений довольно просто прийти к выводу, что оптимальное управление должно иметь ту структуру, которую мы получили путем вычислений — существует свой на каждом этапе критический уровень запасов, и наша политика заказов должна сводиться к поддержанию этого уровня. Нетрудно также сообразить, что этот критический уровень тем ниже, чем ближе к заключительному этапу. Однако величины оптимальных запасов  $a_k$ , конечно, нельзя найти без вычислений.

Скажем несколько слов относительно некоторых других постановок задачи управления запасами.

Никаких принципиальных трудностей не возникает, если в условиях предыдущей задачи речь идет о запасании не одного

вида товаров, а нескольких. В этом случае оптимальная политика также состоит в поддержании критического уровня запасов каждого товара. Аналогично рассматривается задача о создании запасов, когда заказанные товары поступают с запаздыванием на фиксированное число этапов. Новые эффекты возникают, если время запаздывания — случайная величина.

В ряде задач более естественна другая структура функции штрафов. Может случиться так, — что для нас нежелательно, — что запас превзойдет некоторую величину — вместимость имеющихся складов, а дефицит товаров — неудовлетворенный спрос, не превзойдет некоторой другой константы. Тогда оптимальная политика создания запасов должна состоять в том, чтобы случайный процесс  $x_n$  — количество имеющихся или недостающих товаров в момент  $n$ , в течение рассматриваемого отрезка времени не выходил за эти границы по возможности с большей вероятностью. Или, другая постановка, — чтобы среднее время до выхода за эти границы было бы максимальным. Такие задачи близки к вопросам, рассматриваемым в следующем параграфе для процессов с непрерывным временем. Отметим, что если величины заказов на каждом этапе невелики, а этапов много, то в задачах управления запасами можно с успехом применять модели с непрерывным временем.

Все примеры этого параграфа относились к случаю, когда спрос есть последовательность независимых одинаково распределенных величин. Такое допущение, конечно, далеко не всегда выполняется в конкретных прикладных задачах. Иногда спрос в момент  $n$  удобно представить в виде  $\xi_n = f(n) + \zeta_n$ , где  $f(n)$  — некоторая известная функция — детерминированная составляющая спроса,  $\zeta_n$  — случайные отклонения, независимые при разных  $n$ . Для такого спроса возможны рассмотрения, аналогичные тем, которые приводились в этом параграфе. Если случайные отклонения  $\zeta_n$  можно считать маленькими  $\zeta_n = \varepsilon \tilde{\zeta}_n$ , ( $\varepsilon$  — малый параметр), то к задаче управления запасами, даже в случае зависимых  $\zeta_n$ , может быть применен метод, близкий к тому, который рассматривается в следующем параграфе. Имеются и некоторые другие случаи зависимых величин  $\zeta_n$ , в которых задача управления запасами допускает более или менее окончательное теоретическое решение. В других случаях лучшие результаты могут быть получены путем моделирования задач на электронных вычислительных машинах.

## § 7. Управляемые марковские процессы. Оптимальная стабилизация

Пусть состояние некоторой физической системы описывается точкой плоскости  $x = (x^1, x^2)$ , а его изменение во времени задается дифференциальным уравнением

$$x = b(x, u), \quad b(x, u) = (b^1(x, u), b^2(x, u)). \quad (*)$$

Здесь  $u = u(t) = (u^1(t), u^2(t))$  — управление, которое выбирается из некоторого множества допустимых управлений. Структуру этого множества мы опишем несколько позже. Одна из важных задач изучения таких систем — исследование их устойчивости относительно малых возмущений правых частей и начальных условий. Интерес к таким задачам объясняется тем, что фактически реализуются только устойчивые решения уравнения (\*).

В этом параграфе мы рассмотрим задачи, связанные с устойчивостью положения равновесия системы (\*). Точка 0 называется положением равновесия этой системы (при фиксированном управлении  $u(t)$ ), если  $b(0, u(t)) \equiv 0$  при всех  $t$ . Очевидно, что траектория системы (\*), начинающаяся в такой точке 0, никогда из нее не выходит. Положение равновесия 0 называется устойчивым (точнее асимптотически устойчивым), если любые решения системы (\*), начинающиеся вблизи 0, приближаются к 0 при  $t \rightarrow \infty$ . Асимптотическая устойчивость означает, что небольшие возмущения положения системы со временем нивелируются.

Предположим теперь, что на правые части системы (\*) наложены некоторые случайные возмущения:

$$\tilde{x}_t = b(\tilde{x}_t, u) + \zeta_t.$$

Здесь  $\zeta_t = (\zeta_t^1, \zeta_t^2)$  — случайный процесс;  $\tilde{x}_t = \tilde{x}_t^u$  в этом случае тоже случайный процесс. Его траектории с ростом  $t$ , вообще говоря, уже не будут приближаться к положению равновесия 0, даже если оно было устойчивым для системы (\*); они могут рано или поздно выйти из любой окрестности точки 0. Однако если возмущения не очень велики, то траектории долгое время будут находиться вблизи 0.

Обычно можно считать, что физическая система, описываемая уравнением (\*), «не разрушается», если ее состояние характеризуется точкой плоскости, не слишком далеко отстоящей от точки 0. Иными словами, можно указать такую критическую область  $D$ , что пока  $x_t$  находится внутри  $D$ , система не разрушается. Как уже упоминалось, процесс  $x_t$ , вообще говоря, выходит из области  $D$ . Обозначим  $\tau^u$  — случайную величину — время, необходимое  $x_t^u$  для достижения границы области  $D$ . В течение времени  $\tau^u$  физическая система не разрушается.

Предположим, что наша цель состоит в том, чтобы система по возможности дольше не разрушалась. Тогда естественно поставить задачу так: среди всех допустимых управлений выбрать такое управление  $\hat{u}(t)$ , чтобы среднее время жизни системы было максимальным, т. е.  $M_x \tau^{\hat{u}} = \max_u M_x \tau^u$ . Здесь индекс  $x$  у математического ожидания означает, что рассматриваются случайные траектории  $\tilde{x}_t$ , начинающиеся в точке  $x$ . Если мы интересуемся поведением системы только на конечном участке времени

от 0 до  $T$ , то более естественно ставить задачу по-другому: среди всех допустимых управлений выбрать такое, при котором вероятность выхода из области  $D$  за время  $[0, T]$  минимальна.

Заметим, что при различных управлениях  $u = u(t)$  положение точки равновесия может меняться. При некоторых  $u(t)$  система (\*) может вовсе не иметь положений равновесия. Это для нас совершенно неважно; критическая область  $D$  не зависит от управления, и мы интересуемся тем, когда  $\tilde{x}_t^u$  выходит из  $D$ .

Уточним постановку задачи. Случайные процессы  $\zeta_t^1$  и  $\zeta_t^2$  — шум в нашей системе, могут иметь самую разную природу. Один из наиболее простых и вместе с тем важных для приложений случаев мы получим, если предположим, что  $\zeta_t^1$  и  $\zeta_t^2$  — независимые процессы белого шума:  $\zeta_t^1 = \dot{w}_t^1$ ,  $\zeta_t^2 = \dot{w}_t^2$ . Если обозначить  $w_t = (w_t^1, w_t^2)$ , то уравнение для  $x_t^u$  можно записать так:  $\dot{\tilde{x}}_t^u = b(\tilde{x}_t^u, u(t)) + \dot{w}_t$ ,  $x_0^u = x$ . Индекс  $u$  у функции  $x_t^u$  указывает на то, какое управление  $u(t)$  используется.

Опишем теперь, как устроены допустимые управления  $u(t) = u(u^1(t), u^2(t))$ . Будем считать, что  $b^1(x, u) = u^1$ ,  $b^2(x, u) = u^2$ . Таким образом, выбор управления означает выбор векторного поля  $u = b(x, u)$ . Для вычисления управления в момент  $t$  разрешается использовать только информацию о ходе процесса до этого момента. Укажем, наконец, как устроено множество управления: мы считаем, что при всех  $t$  вектор  $u(t)$  имеет фиксированную длину  $c$ .

Среднее время, которое затратит на выход из области  $D$  процесс  $\tilde{x}_t^u$ , начинающийся в точке  $x$ , обозначим  $V^u(x) : V^u(x) = M_x \tau^u$ . Наша задача состоит в выборе допустимого управления, которое максимизирует это среднее время.

Априори могло бы казаться, что это оптимальное управление — свое для каждой начальной точки  $x$ . Но на самом деле это не так. Оказывается, в нашей задаче существует марковское оптимальное управление, т. е. оптимальное управление  $\hat{u}(t)$  можно искать среди функций  $u(t) = b(\tilde{x}_t)$ ,  $b(x) = b(b^1(x), b^2(x))$ : оптимальное управление в момент  $t$  зависит только от положения траектории  $\tilde{x}_t$  в момент  $t$ . Такое оптимальное управление доставляет минимум  $V^u(x)$  сразу для всех точек  $x$ , лежащих внутри  $D$ . Это важное утверждение является следствием того, что винеровский процесс  $W_t$  имеет независимые приращения.

Если управление  $u(t)$  в момент  $t$  имеет вид  $u(t) = b(\tilde{x}_t)$ , то уравнение для  $\tilde{x}_t$  можно переписать в виде:  $\dot{\tilde{x}}_t = b(\tilde{x}_t) + \dot{w}_t$ . Как объяснялось в § 5, решение уравнения представляет собой марковский процесс. В этом случае функция  $V(x) = V^u(x) = M_x \tau^u$  удовлетворяет дифференциальному уравнению

$$\frac{1}{2} \left( \frac{\partial^2 V}{(\partial x^1)^2} + \frac{\partial^2 V}{(\partial x^2)^2} \right) + b^1(x) \frac{\partial V}{\partial x^1} + b^2(x) \frac{\partial V}{\partial x^2} = -1 \quad (**)$$

при  $x$  из  $D$  и граничному условию  $V(x) = 0$  для  $x$ , принадлежащих границе области.

В частности, если  $\hat{b}(x)$  — оптимальное векторное поле, то максимально возможное среднее время выхода —  $W(x)$  удовлетворяет уравнению (\*\*) при  $b(x) = \hat{b}(x)$ . С помощью принципа оптимальности Беллмана можно вывести, что оптимальное поле  $\hat{b}(x)$  должно быть направлено в сторону наискорейшего роста функции  $W(x)$ , то есть вдоль градиента  $\nabla W(x)$  этой функции. Учитывая, что длина вектора  $|\hat{b}(x)| = c$ , получим:

$b(x) = c|\nabla W(x)|^{-1}\nabla W(x)$ , где  $|\nabla W(x)|$  — длина вектора  $\nabla W(x) = \left(\frac{\partial W}{\partial x^1}, \frac{\partial W}{\partial x^2}\right)$ .

Резюмируя все сказанное, получим следующее уравнение для  $W(x)$ :

$$\frac{1}{2} \left( \frac{\partial^2 W}{(\partial x^1)^2} + \frac{\partial^2 W}{(\partial x^2)^2} \right) + c \sqrt{\left( \frac{\partial W}{\partial x^1} \right)^2 + \left( \frac{\partial W}{\partial x^2} \right)^2} = -1.$$

$W(x) = 0$  на границе  $D$ .

По решению  $W(x)$  этого уравнения вычисляется оптимальное управление  $\hat{b}(x)$ .

Решение этого нелинейного дифференциального уравнения, аналитическое или численное, представляет определенные трудности, поэтому интересно иметь и другие способы изучения устойчивости и оптимальной стабилизации при случайных возмущениях. Предположим, что случайные возмущения малы по сравнению с систематической составляющей — длиной  $c$  векторов  $b(x)$ , так что уравнения для процесса  $\tilde{x}_t$  можно переписать так:

$$\dot{x}_t^\varepsilon = b(x_t^\varepsilon) + \varepsilon \dot{W}_t.$$

Здесь  $\varepsilon$  — малый параметр.

При малом  $\varepsilon$  среднее время  $M_x^\varepsilon \tau^u$  выхода процесса  $x_t^\varepsilon$  из критической области  $D$  для любой разумной стратегии будет велико:  $\lim_{\varepsilon \rightarrow 0} M_x^\varepsilon \tau^u = \infty$ . Поэтому естественно поступить следующим образом: вычислить главный член  $M_x^\varepsilon \tau^u$  при  $\varepsilon \rightarrow 0$  и выбирать управление так, чтобы максимизировать этот главный член (а не саму функцию  $M_x^\varepsilon \tau^u$ ). Управление, оптимальное в таком смысле, при достаточно малом  $\varepsilon$ , будет не хуже любого другого, не зависящего от  $\varepsilon$ , управления. Такой подход оправдывается еще и тем, что величина  $\varepsilon$  нам часто бывает неизвестна точно.

Это оптимальное векторное поле  $\hat{b}(x)$ , оказывается, устроено следующим образом. В критическую область  $D$  нужно впи-

сать круг максимального возможного радиуса и выбрать  $\hat{b}(x)$  направленными по радиусам этого круга в сторону центра. Вне максимального круга управление можно выбирать почти произвольно, лишь бы это управление «загоняло» траектории случайного процесса внутрь максимального круга. При этом логарифмический главный член  $M_x^\varepsilon \tau$  при  $\varepsilon \rightarrow 0$  не зависит от  $x$  и от управления вне круга.

В предположении малости возмущений подобным образом можно исследовать устойчивость и для систем более общего вида, например, для систем с запаздыванием, а также для гауссовских возмущений, отличных от белого шума.

Скажем несколько слов еще об одном подходе к устойчивости при случайных возмущениях. В некоторых задачах интенсивность возмущений зависит от точки  $x$ , и в самом положении равновесия возмущения исчезают. В таких случаях можно построить теорию устойчивости стохастических систем, аналогичную классической теории устойчивости. В рамках этой теории устойчивости также рассматриваются задачи оптимальной стабилизации.

В заключение заметим, что задачи, подобные тем, которые рассматривались в этом параграфе, появляются не только в связи с исследованием устойчивости. Управляемые диффузионные процессы возникают в ряде других моделей (задачи погони, оптимальное быстрое действие при наличии шумов, задачи фильтрации и другие). Марковские процессы вообще оказались весьма удобными для описания управляемых стохастических систем. Такое положение в значительной степени объясняется тем, что для исследования марковских процессов можно применять мощный аппарат теории дифференциальных уравнений. Кроме того, в марковском случае обычно бывает легче понять, какая информация о ходе процесса существенна для выработки оптимального управления. Для нахождения этих управлений в марковском случае очень полезным оказывается принцип оптимальности Беллмана. Наконец, важно и то, что многие случайные процессы, которые не являются марковскими, можно приближенно считать компонентами многомерных марковских процессов.

## **§ 8. Задача обнаружения разладки. Оптимальная остановка**

Широкий класс задач, касающихся управления случайными процессами, можно сформулировать как задачи выбора оптимального в том или ином смысле момента времени. Многие такие задачи могут быть рассмотрены с единой точки зрения. Приведем примеры подобных задач.

Предположим, что автоматическая линия выпускает одну деталь в единицу времени, и эта деталь характеризуется каким-то одним параметром, скажем, диаметром некой своей цилиндри-

ческой частью. Если линия хорошо отлажена, то этот диаметр в среднем равен некоторой постоянной величине, но за счет случайных ошибок возможны небольшие отклонения от среднего, так что диаметр детали естественно считать случайной величиной, имеющей некоторую плотность  $p_0(x)$ . Для разных деталей эти величины независимы.

В случайный момент времени  $\theta$  происходит разладка автоматической линии. Это может, например, вызвать систематическое отклонение размеров деталей от стандарта или же при сохранившемся среднем вызвать увеличение разброса. Мы будем считать, что разладка приводит к тому, что диаметры деталей, выпущенных в момент  $\theta$  и позже, уже имеют распределение, отличное от  $p_0(x)$ . Пусть  $p_1(x)$  — плотность нового распределения. Таким образом, диаметр  $n$ -й детали  $\xi_n$  имеет плотность  $p_0(x)$  при  $n < \theta$  и  $p_1(x)$ , если  $n \geq \theta$ .

Наша задача состоит в том, чтобы по наблюдениям за размерами деталей как можно скорее обнаружить разладку. Если мы объявляем, что произошла разладка, то автоматическую линию останавливают и производят переналадку.

Отклонения размеров деталей от стандарта могут объясняться не разладкой линии, а случайными ошибками, поэтому объявление о разладке может быть ложной тревогой; разладки к этому времени на самом деле может не быть. С другой стороны, если мы будем слишком осторожны, то пропустим момент разладки, и линия будет долго работать в разлаженном режиме. Мы сможем выбрать «золотую середину», если будем иметь возможность оценить относительную стоимость переналадки линии и ущерба за единицу времени работы разлаженной линии.

Уточним постановку задачи. Это уточнение можно производить по-разному. Мы приведем так называемую байесовскую постановку. Предположим, что случайный момент  $\theta$  с некоторой вероятностью  $\pi$  равен нулю, то есть с вероятностью  $\pi$  мы с самого начала наблюдаем за разлаженной линией. Если в начальный момент разладки не было, то  $P\{\theta = n/\theta > 0\} = p(1 - p)^{n-1}$ , где  $p > 0$ . Такое распределение, называемое геометрическим, является дискретным аналогом показательного распределения. Оно возникает, по существу, при тех же качественных условиях, которые в непрерывном случае приводят к показательному закону.

Объясним, какие управления в нашей задаче являются допустимыми. Марковским моментом или правилом остановки называется неотрицательная целочисленная случайная величина  $\tau$ , которая не зависит от будущего. Это значит, что при произвольном целом  $t$  мы можем ответить на вопрос  $\tau = t$  или нет по ходу процесса до момента  $t$  включительно, т. е. по значениям, принятым величинами  $\xi_1, \xi_2, \dots, \xi_t$ . Типичным примером марковского момента служит случайная величина, равная номеру первого члена последовательности  $\xi_1, \xi_2, \dots, \xi_n$ , который превосходит некоторый фиксированный уровень  $d$ :  $\tau =$

$= \min\{n : \xi_n > d\}$ . Допустимые управления в нашей задаче это как раз и есть марковские моменты. Такое определение допустимого управления соответствует естественному требованию, что решение об объявлении разладки в момент  $t$  мы должны принять только по результатам измерений деталей, выпущенных в моменты  $1, 2, \dots, t$ .

Теперь поясним, как устроена плата при той или иной допустимой стратегии. Предположим, что стоимость переналадки оценивается в одну денежную единицу, а денежное выражение ущерба от работы разлаженной линии в течение единицы времени равно  $c$ . Если мы выбрали стратегию  $\tau = \tau(\omega)$ , то с вероятностью  $P\{\tau < \theta\}$  будет объявлена ложная тревога, и тогда наши потери равны стоимости переналадки, т. е. 1. Потери от работы разлаженной линии равны 0, если  $\tau \leq \theta$ , и  $c \cdot (\tau - \theta)$ , если  $\tau > \theta$ . Поэтому средние потери от работы разлаженной линии равны  $c \cdot M \max(0, \tau - \theta)$ . Таким образом, общие средние потери при стратегии  $\tau$  составят

$$\rho(\pi, \tau) = 1 \cdot p\{\tau < \theta\} + c \cdot M \max(0, \tau - \theta).$$

Стратегия  $\tau^*$  называется  $\pi$ -оптимальной, если она доставляет минимум  $\rho(\pi, \tau)$  по всем допустимым стратегиям  $\tau$ . Этот минимум обозначим  $\rho(\pi)$ :  $\rho(\pi) = \rho(\pi, \tau^*)$ .

Здесь  $\pi$  — вероятность того, что линия разлажена в момент 0. Обозначим  $\pi_n$  условную вероятность того, что линия разлажена в момент  $n$  при условии, что значения  $\xi_1, \xi_2, \dots, \xi_n$  известны:  $\pi_n = p\{\theta \leq n / \xi_1, \xi_2, \dots, \xi_n\}$ . Вероятность  $\pi_n$  называется апостериорной вероятностью разладки в момент  $n$  в отличие от априорной вероятности  $p\{\theta \leq n\}$ . Апостериорная вероятность  $\pi_n$  разладки в момент  $n$  есть функция от случайных величин  $\xi_1, \xi_2, \dots, \xi_n$  и от числа  $\pi$ , поэтому  $\pi_n$  сама является случайной величиной  $\pi_n = \pi_n(\pi, \omega)$ .

Покажем, что последовательность  $\pi_0, \pi_1, \dots, \pi_n, \dots$  — марковская. Из формулы Байеса получим:

$$\pi_{n+1} = \frac{\pi_n p_1(\xi_{n+1}) + (1 - \pi_n) p \cdot p_1(\xi_{n+1})}{\pi_n p_1(\xi_{n+1}) + (1 - \pi_n) p \cdot p_1(\xi_{n+1}) + (1 - \pi_n)(1 - p) p_0(\xi_{n+1})}.$$

Отсюда следует, что для вычисления  $\pi_{n+1}$ , кроме  $\pi_n$ , нужно знать еще только  $\xi_{n+1}$ , так как  $\xi_{n+1}$  не зависит от  $\xi_1, \dots, \xi_n$ , а стало быть, и от  $\pi_1, \pi_2, \dots, \pi_{n-1}$ , которые вычисляются, по величинам  $\xi_k$  при  $k < n$ . Следовательно,  $\pi_0, \pi_1, \dots, \pi_n, \dots$  — марковский процесс. Используя то, что  $\tau$  — марковский момент, нетрудно подсчитать, что

$$\rho(\pi, \tau) = M \left[ (1 - \pi_\tau) + c \sum_{k=1}^{\tau-1} \pi_k \right].$$

Интуитивно ясно, и это можно строго доказать, что для того, чтобы судить о том, произошла разладка в момент  $n$  или нет,



нужно знать  $\pi_0, \pi_1, \dots, \pi_n$ . Значения  $\xi_k$  при  $k \leq n$  не несут ничего нового.

Таким образом, наша задача свелась к следующей. Есть марковская цепь  $\pi_0, \pi_1, \dots, \pi_n, \dots$  и функция  $(1 - \pi_\tau) + c \sum_{k=1}^{\tau-1} \pi_k$  от траектории этой цепи и момента  $\tau$ . Нужно выбрать правило остановки  $\tau$  так, чтобы математическое ожидание указанной функции было бы минимальным.

Позже мы еще вернемся к этой задаче, а сейчас приведем еще один пример, который известен под шуточным названием «задача о разборчивой невесте». Мы покажем, что и эта задача сводится к оптимальной остановке некоторой цепи Маркова.

Предположим, что имеется  $N$  объектов (женихов), причем из любых двух объектов можно указать лучший. Эти объекты рассматриваются в случайном порядке, и мы (невеста) хотим выбрать самый лучший из них. При этом мы можем сравнивать объекты с рассмотренными ранее, но если некоторый объект уже был просмотрен и отвергнут, то возвращаться к нему нельзя. Как нам следует поступить? На каком объекте остановиться, чтобы вероятность того, что этот объект наилучший, была максимальной?

Очевидно, эту задачу можно сформулировать так. Имеется  $N$  вещественных чисел  $b_1, b_2, \dots, b_N$ , среди которых нет одинаковых. Эти числа перемешаны так, что все перестановки равновероятны. Мы последовательно просматриваем эти числа и хотим выбрать самое большое. Возвращаться к числам уже просмотренным и невыбранным запрещается.

Если некоторое число  $b_k$  больше всех предыдущих, то мы будем называть его рекордсменом. Ясно, что наш выбор должен остановиться на каком-то из рекордсменов, поэтому изучим последовательность рекордсменов подробнее. Очевидно, что рекордсменом обязательно будет первое число  $b_1$ . Если  $b_1 = \max_{1 \leq k \leq N} b_k$ , то других рекордсменов вообще не будет. Самое большое из чисел  $b_k$  всегда будет рекордсменом, на каком бы месте оно ни стояло. Обозначим  $r_k$  номер в порядке рассмотрения чисел  $(k+1)$ -го рекордсмена;  $r_0 = 1$ , так как первый рекордсмен  $b_1$ . Если  $r_n$  — номер последнего рекордсмена в последовательности  $b_1, b_2, \dots, b_N$ , то условимся считать, что  $r_m = \infty$  при  $m > n$ . Нетрудно проверить, что  $r_0, r_1, \dots, r_n$  — марковская цепь. Состояния этой цепи суть числа  $1, 2, \dots, N$  и точка  $\infty$ . Найдем ее переходные вероятности:

$$p_{kn} = p\{r_{i+1}=n/r_i=k\} + \frac{p\{r_i=k, r_{i+1}=n\}}{p\{r_i=k\}}.$$

При  $n \leq k$  вероятность, стоящая в числителе, равна нулю. При  $\infty > n > k$  событие  $\{r_i=k, r_{i+1}=n\}$  означает, что среди чисел

$b_1, b_2, \dots, b_n$  правее всех оказались  $b_k$  и  $b_n$ , причем  $b_n > b_k$ . Вероятность этого события равна  $\frac{1}{(n-1)n}$ . Вероятность, стоящая в знаменателе, равна  $\frac{1}{k}$ .

Таким образом, вероятности  $p_{kn}$  для цепи  $r_i$  равны 0 при  $n \leq k < \infty$  и  $\frac{k}{n(n-1)}$  при  $k < n < \infty$ . Вероятность того, что  $k$ -й рекордсмен последний, вычисляется по формуле

$$p_{k\infty} = 1 - \sum_{n=k+1}^N p_{kn} = 1 - \sum_{n=k+1}^N \frac{k}{n(n-1)} = \frac{k}{N}.$$

Если  $b_k$  — последний рекордсмен, то  $b_k$  — абсолютный чемпион, наибольшее число. Следовательно, если мы остановимся на рекордсмене  $b_k$ , то мы выиграем, т. е. выберем наибольшее число с вероятностью  $p_{k\infty} = \frac{k}{N}$ . Посмотрим, какова будет вероятность  $q_k$  выигрыша, если мы остановимся на следующем после  $b_k$  рекордсмене. С вероятностью  $p_{kn}$  следующим рекордсменом будет  $b_n$  и с вероятностью  $p_{n\infty}$  этот рекордсмен будет последним, поэтому

$$q_k = \sum_{n=k+1}^N p_{kn} p_{n\infty} = \frac{k}{N} \left( \frac{1}{k} + \frac{1}{k+1} + \dots + \frac{1}{N} \right).$$

Отношение  $\frac{q_k}{p_{k\infty}} = \frac{1}{k} + \frac{1}{k+1} + \dots + \frac{1}{N}$  с ростом  $k$  монотонно убывает. Пока оно больше единицы, останавливаться не стоит: вероятность выигрыша на следующем рекордсмене больше, чем на нынешнем. Интуитивно ясно, что надо ждать, пока впервые не появится рекордсмен  $b_k$  с таким  $k$ , что  $\frac{q_k}{p_{k\infty}} = \frac{1}{k} + \dots + \frac{1}{N}$  меньше 1. Именно на нем и нужно останавливаться. Все предыдущие рекордсмены имеют меньшую вероятность быть последними, а других рекордсменов с большей вероятностью не будет.

При больших  $N$  и  $k$

$$\frac{1}{k} + \frac{1}{k+1} + \dots + \frac{1}{N} \approx \int_k^N \frac{dx}{x} = \ln \frac{N}{k}.$$

Значит, впервые  $\frac{1}{k} + \dots + \frac{1}{N}$  станет меньше 1 при  $k \approx \frac{N}{e}$ . Таким образом, оптимальная стратегия, во всяком случае при

больших  $N$  состоит в том, что нужно пропустить приблизительно  $\frac{N}{e} = \frac{N}{2,71\dots}$  чисел, а затем выбрать то число, которое больше всех предыдущих. При малых  $N$  нужно пропустить  $k_N$  чисел, а затем выбрать первое число, большее предыдущих, где  $k_N$  — наименьшее решение неравенства

$$\frac{1}{k_N} + \frac{1}{k_N + 1} + \dots + \frac{1}{N} < 1.$$

Так, при  $N = 5$  нужно пропустить три первых числа; при  $N = 10$  оптимальная стратегия состоит в выборе наибольшего после четырех пропущенных.

Можно подсчитать вероятность выигрыша при использовании оптимальной стратегии. Оказывается, при больших  $N$  эта вероятность стремится к  $\frac{1}{e} \approx 0,37$ .

Задача о разборчивой невесте, так же как задача об обнаружении разладки, приводит к отысканию оптимального момента остановки марковской цепи. В задаче о невесте эта цепь — последовательность номеров рекорсменов, и наша задача свелась к такому выбору правила остановки  $\tau$ , чтобы с наибольшей вероятностью остановиться в момент, непосредственно предшествующий скачку в  $\infty$ . Из состояния  $k$  совершается скачок в  $\infty$  с вероятностью  $\frac{k}{N}$ . Поэтому вероятность успеха при стратегии  $\tau$  равна

$$\sum p\{r_\tau = k\} \frac{k}{N} = \frac{1}{N} \sum kp\{r_\tau = k\} = \frac{1}{N} Mr_\tau.$$

Таким образом, задача о разборчивой невесте эквивалентна выбору такого правила остановки  $\tau$  для цепи  $r_0, r_1, \dots, r_n, \dots$ , которое бы максимизировало  $Mr_\tau$ .

Общую задачу об оптимальной остановке марковской цепи  $x_0, x_1, \dots, x_n, \dots$  можно сформулировать так. Пусть есть две функции  $f(x)$  и  $g(x)$  на состояниях цепи. Требуется выбрать такой марковский момент  $\tau$ , чтобы  $M \left[ f(x_\tau) + \sum_{k=0}^{\tau-1} g(x_k) \right]$  приняло бы максимальное значение. Обе рассмотренные в этом параграфе задачи включаются в такую схему. В задаче о разладке  $f(x) = 1 - x$ ,  $g(x) = x$ . Во второй задаче  $f(x) = x$ ,  $g(x) = 0$ . Весьма замечательным является то обстоятельство, что существует изящный и довольно, общий метод для решения задач об оптимальной остановке марковской цепи. Этот метод со многими примерами изложен в книге Е. Б. Дыштша и А. А. Юшкевича [5]. Оптимальной остановке посвящена также книга А. Н. Ширяева

«Статистический последовательный анализ» (М., 1969). В этой последней книге, в частности, рассматривается задача обнаружения разладки и ее обобщения. Там доказывается, что оптимальное правило остановки  $\tau^*$  в задаче обнаружения разладки устроено следующим образом:  $\tau^*(\pi, \omega) = \min\{n : \pi_n(\pi, \omega) \geq A^*\}$ . То есть объявлять о разладке следует в тот момент  $n$ , когда апостериорная вероятность  $\pi_n$  впервые превзойдет некоторую константу  $A^*$ .

## ЗАКЛЮЧЕНИЕ

В этой брошюре мы рассмотрели некоторые вероятностные модели процессов управления. Построение и исследование математических моделей — один из основных методов изучения реальных систем управления. Здесь мы сделаем несколько общих замечаний относительно построения таких моделей.

Обычно математическая модель строится на основе теоретических представлений о ходе реального процесса. Из качественных соображений или по аналогии делаются те или иные предположения о статистической структуре случайных процессов, которые учитываются в системе. Так, часто предполагается, что случайные процессы являются гауссовскими, стационарными, марковскими и так далее. Для подобных предположений бывает два типа причин. Кроме причин, связанных со статистическими закономерностями, формирующими наблюдаемый процесс, немалую роль играет то обстоятельство, что не придерживаясь каких-то гипотез о характере процесса, мы не будем в состоянии вести его математическое исследование. Эти гипотезы обычно бывают в большей или меньшей степени произвольными. Чтобы сделать их по возможности обоснованными, необходимо произвести проверку с помощью соответствующих статистических критериев.

Даже в случае хорошего согласования математической модели с реальным процессом нужно помнить, что эта адекватность имеет место только в определенных границах изменения параметров, при определенных масштабах явлений. Наибольшую достоверность имеют сравнительно грубые, качественные выводы, которые мало изменяются при небольших изменениях параметров и структуры модели. Чтобы практическую ценность приобрели более точные выводы, необходимо повысить требования к адекватности модели и реального процесса. Уточнение модели обычно приводит к ее усложнению. Именно желание дать более адекватное описание некоторых реальных процессов, чем это можно сделать с помощью детерминированных моделей, и привело к стохастическим моделям процессов управления. Здесь уместно отметить, что если система не допускает простого детер-

минированного описания, то это еще не означает, что она может быть охарактеризована с помощью вероятностной модели. Адекватное вероятностное описание реальной системы предполагает, что в работе этой системы наблюдаются определенные устойчивые статистические закономерности.

Другая трудность, которая возникает при построении математических моделей систем управления, связана с назначением функции платы. Не всегда просто оценить ущерб (или выигрыш) от того или иного хода процесса. В таких случаях наибольший интерес представляют стратегии, которые являются хорошими, хотя, быть может, и не оптимальными, для целого семейства возможных функций платы. Построение хорошей математической модели возможно только на основе знания специфических свойств изучаемой конкретной системы управления и учета возможностей математической теории.

В этой брошюре мы стремились дать описание основных классов случайных процессов, которые используются для построения вероятностных моделей систем управления и привести примеры таких систем.

Наконец, скажем несколько слов о литературе по обсуждаемым здесь вопросам. С началами теории вероятностей можно познакомиться по популярной книге американских математиков Ф. Мостеллера, Р. Рурке и Дж. Томаса [1]. Из серьезных курсов теории вероятностей назовем учебник Б. В. Гнеденко [2]. По теории случайных процессов, кроме книг, названных в тексте, можно рекомендовать еще книгу С. Карлина [4], в которой рассматриваются различные классы марковских процессов.

## ЛИТЕРАТУРА

1. **Ф. Мостеллер, Р. Рурке, Дж. Томас.** Вероятность. М., «Мир», 1969.
2. **Б. В. Гнеденко.** Курс теории вероятностей. М., Физматгиз, 1961.
3. **В. С. Пугачев.** Теория случайных функций и ее применение к задачам автоматического управления. М., Госуд. изд-во техн.-теоретич. лит-ры, 1957.
4. **С. Карлин.** Основы теории случайных процессов. М., «Мир», 1971.
5. **Е. Б. Дынкин, А. А. Юшкевич.** Теоремы и задачи о процессах Маркова. М., «Наука», 1967.
6. **Р. Беллман.** Динамическое программирование. М., ИЛ, 1960.
7. **Р. А. Ховард.** Динамическое программирование и марковские процессы. М., «Сов. радио», 1964.

## СОДЕРЖАНИЕ

Введение .....	3
§ 1. Вероятности и случайные величины .....	4
§ 2. Гауссовские случайные процессы .....	9
§ 3. Стационарные процессы .....	13
§ 4. Марковские процессы .....	16
§ 5. Основные понятия теории оптимального управления .....	24
§ 6. Управление созданием запасов .....	28
§ 7. Управляемые марковские процессы. Оптимальная стабилизация .....	33
§ 8. Задача обнаружения разладки. Оптимальная остановка .....	37
Заключение .....	44
Литература .....	45



**ФРЕЙДЛИН Марк Иосифович**

**СЛУЧАЙНЫЕ ПРОЦЕССЫ И ОПТИМАЛЬНОЕ УПРАВЛЕНИЕ**

Редактор **В. Ю. Иваницкий**

Художник **Л. П. Ромасенко**

Худож. редактор **В. И. Конюхов**

Техн. редактор **Т. В. Самсонова**

Корректор **А. А. Пузакова**

А09302 Сдано в набор 30/IV 1972 г. Подписано к печати 20/VI 1972 г.  
Формат бумаги 60 × 90<sup>1/16</sup> Бумага типографская № 2 Бум. л. 1,5  
Печ. л. 3 Уч.-изд. л. 2,58 Тираж 46 340 Издательство «Знание». Москва,  
Центр, Новая пл., д. 3/4. Заказ 783 Цена 9 коп.

Чеховский полиграфкомбинат Главполиграфпрома  
Комитета по печати при Совете Министров СССР  
г. Чехов Московской области

9 коп.

Индекс 70096

