

Вывод некоторых уравнений математической физики

В.Н. Тутубалин

(редакция 25 августа 2012 г.)

Аннотация

Данный текст представляет собой учебный материал по курсам «Математическая статистика и случайные процессы» для студентов-механиков 4-го курса (8-ой семестр) и по курсам естественно-научного содержания «Финансовая статистика» и «Приложения вероятностных методов». Первая задача этого материала состоит в том, чтобы на примерах разъяснить студентам, что происходит при так называемом «выводе» уравнений математической физики. Вторая задача – познакомить студентов с уравнением теплопроводности, уравнением Колмогорова для диффузионных случайных процессов и с уравнением Блэка-Шоулса для цены опциона (которое отличается от уравнения Колмогорова лишь несущественными деталями). Приводится также теорема Колмогорова о предельном переходе от последовательности марковских цепей к диффузионному процессу (в смысле сходимости переходных вероятностей).

1. Уравнение теплопроводности

Уравнения математической физики выражают, вообще говоря, физические законы, однако физический закон – это не то, что представляет себе большинство студентов-математиков. В самом деле, математики обычно думают, что то или иное положение физики должно быть чем-то вроде, например, определения или аксиомы, или теоремы, или обобщения экспериментальных фактов. При знакомстве с конкретными физическими законами они начинают искать эти элементы, но их не находят. Возьмем для примера второй закон Ньютона в виде $F = ma$. Ну, во-первых, сам Ньютон не мог написать такую формулу, поскольку она верна лишь в том случае, когда сила, масса и ускорение измеряются в рамках определенной системы единиц, а во времена Ньютона понятия системы единиц еще не было. Во-вторых, эта формула сама употребляется для создания системы единиц (например, с помощью нее устанавливается единица силы, если единицы длины и времени – тем самым и ускорения, а также массы уже установлены). Математик с неудовольствием увидел бы в этом логический круг. Наконец, в физике нет общего понятия силы: силы бывают только конкретные – механические, электростатические, электромагнитные, сила всемирного тяготения и т.д. Поэтому неясно, что бы мог означать второй закон Ньютона в общей форме (а формулируется он именно в общей форме). Наконец, экспериментальная проверка второго закона Ньютона в земных условиях невозможна, так как нельзя учесть все силы, действующие на то или иное тело. Для проверки этого закона нужно выйти в космос, в котором существенна лишь сила всемирного тяготения.

Короче говоря, формулу $F = ma$ нельзя подвести под привычную для математика рубрику: этот физический закон (как и все прочие) представляет собой некоторое мистическое откровение, суть которого состоит в следующем. Если мы правильно выделим некое тело, на которое действуют некие другие тела; правильно определим равнодействующую всех сил; правильно определим массу и ускорение тела – ну тогда и будет выполняться это равенство. А если оно не выполняется, то сами виноваты: что-то определили неправильно.

Теперь займемся другим мистическим откровением, которое называется уравнением теплопроводности. При выводе уравнений математической физики используется весьма разнообразная математика, но в данном тексте из математики

применяется лишь формула Тейлора. Техника состоит в том, что нечто раскладывается по формуле Тейлора, но в этом занятии надо вовремя остановиться. Формально момент остановки определяется математическими предположениями, налагаемыми на рассматриваемые функции, но сами эти предположения нередко определяются тем, что с помощью них удастся получить нечто, доступное дальнейшему анализу.

Если мистическими сущностями механики являются сила, масса и ускорение, то в термодинамике появляются новые сущности – количество тепловой энергии Q и температура T . Что касается температуры, то мы привыкли к тому, что она измеряется в разных ситуациях, так что это понятие можно не пояснять: все равно яснее не станет. А вот тепловую энергию желательно было бы пояснить. В эпоху, когда наука признавала флогистон, Q было просто количеством флогистона (в данном объеме и в данный момент времени). Но потом наука решила, что флогистона не бывает, а бывает тепловое движение неких мелких частиц. Однако если мы станем разбираться, какие именно мелкие частицы и каким образом движутся в железной кочерге, которой мы мешаем угли в печке, то навеки погрязнем в этих деталях и никакого уравнения теплопроводности не получим. Примем, что тепловая энергия перетекает из более горячих мест тела в более холодные таким образом, какой был бы вполне естественным для флогистона. Это перетекание выражается следующим законом *Фурье*. Представим себе маленькую площадку внутри тела, расположенную перпендикулярно к градиенту температуры. Количество ΔQ тепловой энергии, которое за время Δt перетекает через эту площадку (против градиента) дается формулой $\Delta Q = |\text{grad}T| q(t, x) \Delta S \Delta t$, где ΔS - площадь площадки, а $q(t, x)$ - некая величина, называемая *теплопроводностью* тела в момент t и в точке x .

(Вообще говоря, эта величина зависит от температуры, но этой зависимостью мы пренебрежем, иначе не получится желаемого уравнения теплопроводности. Пренебрежем также ее зависимостью от времени.)

Переток тепловой энергии должен менять температуру. Это предлагается учесть с помощью понятия *теплоемкости*. Можно говорить о теплоемкости на единицу массы, но в данном тексте удобней – на единицу объема. Тогда это понятие можно определить следующим образом. Пусть в некотором маленьком объеме ΔV количество тепловой энергии изменилось на ΔQ . Тогда температура получит приращение $\Delta T = \frac{\Delta Q}{c(t, x) \Delta V}$, где $c(t, x)$ - (объемная) теплоемкость (ее зависимостью от времени t и температуры T снова пренебрегаем).

Теперь выведем уравнение теплопроводности в одномерном случае (т.е. пространственная переменная x одномерна). Представим себе тонкий стержень, термоизолированный с боков (тонкий в том смысле, что во всех точках любого его поперечного сечения одна и та же температура, а термоизолированный в том смысле, что через боковые стенки тепловая энергия не уходит и не приходит). Рассмотрим отрезок стержня вида $[x, x + \Delta x]$. Обозначим через $T(t, y)$, $x \leq y \leq x + \Delta x$, температуру в момент t и посмотрим, что произойдет с ней в момент $t + \Delta t$. Через левую границу отрезка пройдет количество тепловой энергии, равное $-\frac{\partial T(t, x)}{\partial x} q(x) \Delta t$, а через правую границу, соответственно, $\frac{\partial T(t, x + \Delta x)}{\partial x} q(x + \Delta x) \Delta t$. Для получения суммарного приращения тепловой энергии ΔQ на этом отрезке надо сложить два последних выражения; для получения приращения температуры нужно ΔQ разделить на $c(x) \Delta x$. А для получения производной от температуры нужно еще разделить на

Δt . Отсюда вытекает, что ΔQ достаточно вычислить в порядке величины $\Delta x \Delta t$.
Имеем

$$\begin{aligned}\Delta Q &= \frac{\partial T(t, x + \Delta x)}{\partial x} q(x + \Delta x) \Delta t - \frac{\partial T(t, x)}{\partial x} q(x) \Delta t \approx \\ &\approx \left(\frac{\partial T(t, x)}{\partial x} + \frac{\partial^2 T}{\partial x^2} \Delta x \right) \left(q(x) + \frac{\partial q}{\partial x} \Delta x \right) \Delta t - \frac{\partial T(t, x)}{\partial x} q(x) \Delta x \Delta t = \\ &= \left(q(x) \frac{\partial^2 T}{\partial x^2} + \frac{\partial q}{\partial x} \frac{\partial T}{\partial x} \right) \Delta t \Delta x\end{aligned}$$

Следовательно,

$$\frac{\partial T(t, x)}{\partial t} = \lim \frac{\Delta Q}{c(x) \Delta t \Delta x} = \frac{q(x)}{c(x)} \frac{\partial^2 T}{\partial x^2} + \frac{1}{c(x)} \frac{\partial q}{\partial x} \frac{\partial T}{\partial x}.$$

Таким образом, если теплопроводность $q(x)$ не есть постоянная величина, то уравнение теплопроводности содержит в правой части не только вторую, но и первую производную от температуры по пространственной переменной x . Если же как теплоемкость, так и теплопроводность постоянны, можно ввести обозначение $a^2 = q(x) / c(x)$ и получить классическую форму уравнения теплопроводности

$$\frac{\partial T}{\partial t} = a^2 \frac{\partial^2 T}{\partial x^2}.$$

Конечно, при выводе этого уравнения предполагалось, что существуют (и непрерывны) участвовавшие в выводе производные. Это и определило использование приближения нужного порядка малости к формуле Тейлора.

Интересно отметить, что в классической форме уравнения (когда вместо отдельно теплоемкости и теплопроводности рассматривается их отношение a^2) кое-что теряется в смысле прикладных возможностей. Одномерная задача (как ни странно) может иметь определенный практический смысл. Например, будем мысленно строить небольшой дом, каждая стена которого состоит из слоя бетона и слоя утеплителя. Как выгоднее (в смысле уменьшения теплопотерь через стены) расположить утеплитель – с внутренней или наружной стороны стены? Температура внутри дома и снаружи дома примерно одинакова во всех точках стены (внутри дома у верхней части стены несколько теплее, чем у нижней части, но не намного). Поэтому и получается одномерная задача: тепловая энергия будет (как мы приближенно считаем) распространяться только вглубь стены, но не в стороны.

Однако в данном случае теплопроводность меняется скачком на границе слоев бетона и утеплителя, а при выводе уравнения мы предположили гладкость всех функций. Надо как-то разобраться с этой потерей гладкости. Здесь и будет существенным отдельное рассмотрение теплопроводности и теплоемкости (а не их отношения a^2).

Направим ось x -ов внутрь стены так, чтобы точка $x=0$ соответствовала наружной поверхности стены. Примем для простоты обозначений одинаковой и равной L толщину слоя бетона и утеплителя. Тогда на полуинтервале $0 \leq x < L$ теплопроводность принимает одно значение q_1 , а на полуинтервале $L < x \leq 2L$ – другое значение q_2 . Так же меняются и теплоемкости (но они нам сейчас не понадобятся). Следовательно, на этих полуинтервалах выполняются уравнения теплопроводности с различными коэффициентами, а что делается в самой точке $x = L$ пока не известно. Однако ясно, что в этой точке тепло никуда не девается, а следовательно, поток тепла является непрерывной функцией от x . Учет этого условия и заменяет отсутствие уравнения в точке $x = L$. Имеются, конечно, еще и граничные условия: можно наружную температуру принять за 0, а температуру внутри дома

обозначить через T_0 . Спрашивается, при каком порядке следования слоев бетона и утеплителя данная величина T_0 обеспечивается при меньшем потоке тепла.

Понятно, что теплоемкости бетона и утеплителя не имеют существенного значения, потому что в стене довольно скоро установится стационарное распределение температуры, а основные теплопотери пойдут на поддержание этого распределения за счет потока тепла, который в конечном счете согревает наружный воздух. Рассмотрим поэтому только стационарное решение, т.е. решение, удовлетворяющее уравнению $\frac{\partial T(t, x)}{\partial t} = 0$. Стационарное решение уравнения

теплопроводности с постоянными теплоемкостью и теплопроводностью есть линейная функция. Таким образом, на полуинтервале $0 \leq x < L$ имеем (с учетом условия $T(t, 0) = 0$) стационарное решение вида $T(x) = a_1 x$, а на полуинтервале $L < x \leq 2L$, соответственно, $T(x) = a_1 L + a_2(x - L)$ (учтена непрерывность температуры в точке $x = L$). Должно также выполняться граничное условие на внутренней поверхности стены $T_0 = T(2L) = (a_1 + a_2)L$, что дает одно уравнение на пока неизвестные величины a_1 и a_2 . Второе уравнение получается из условия равенства потоков тепла слева и справа от границы раздела $x = L$. Поток тепла пропорционален произведению теплопроводности на градиент температуры; следовательно, при равенстве потоков, градиенты a_1 и a_2 обратно пропорциональны теплопроводностям, т.е. $a_1 q_1 = a_2 q_2$. Решая оба уравнения для a_1 и a_2 , получаем

$$a_1 = \frac{T_0 q_2}{L(q_1 + q_2)}, a_2 = \frac{T_0 q_1}{L(q_1 + q_2)}. \quad \text{При этом поток тепла пропорционален}$$

$$a_1 q_1 = a_2 q_2 = \frac{T_0 q_1 q_2}{L(q_1 + q_2)}, \quad \text{т.е. симметрично зависит от } q_1 \text{ и } q_2. \quad \text{Это означает, что все}$$

равно, в каком порядке размещать слои бетона и теплоизолятора (неважно, что снаружи и что внутри).

Замечание. При реальном строительстве учитывается не только перенос тепла, но и перенос паров воды, всегда имеющихся внутри помещения. Не безразлично, где будет конденсироваться (а может быть, и замерзать) эта влага. Ну и конечно, теплый воздух внутри помещения стремится подняться по отношению к холодному воздуху снаружи, и этот ветер, дующий снизу вверх, должен учитываться (особенно в высоких зданиях).

2. Уравнение Колмогорова для диффузионных случайных процессов

2.1. Марковское свойство. Если какой-нибудь объект как-то изменяется во времени, то говорят, что он движется, а сам объект на изящном языке называется «динамической системой». Существует понятие фазового пространства динамической системы: каждая его точка представляет собой такой набор параметров, характеризующих систему, что если значение этого набора задать в какой-то момент времени, то его будущее движение определится однозначно. Например, если движение объекта задается какой-то системой обыкновенных дифференциальных уравнений, то можно (вводя необходимое количество производных в качестве новых параметров) переписать эту систему в виде системы уравнений первого порядка, и тогда задание всех параметров в какой-то момент времени однозначно определит будущее движение системы (если, конечно, будет выполняться теорема существования и единственности решения). Марковское

свойство говорит о понятии фазового пространства, но в ситуации, когда имеется случайность. Указание положения системы в какой-то момент времени определяет не её будущую траекторию, а лишь распределения вероятностей для положения системы в будущие моменты времени. (Здесь мистической сущностью является вероятность.) Таким образом, мы приходим к понятию *переходной функции* (или *переходной вероятности*) $P_s^t(x, A)$, которая обозначает следующее. Пусть в момент $s < t$ положение системы $x(s) = x$, A – некоторое подмножество фазового пространства $X = \{x\}$ (конечно, принадлежащее некоторой сигма-алгебре подмножеств этого пространства). Тогда $P_s^t(x, A)$ есть вероятность того, что в момент t система окажется в множестве A . Постулируется, что переходная функция удовлетворяет так называемому *уравнению Чэпмена-Колмогорова* следующего вида: для моментов времени $s < t < u$

$$P_s^u(x, A) = \int_X P_s^t(x, dy) P_t^u(y, A).$$

Наглядный смысл этого равенства состоит в том, что для того, чтобы система за время от s до u перешла из точки x в множество A , она сначала должна за время от s до t попасть из x в любую точку $y \in X$, а уж потом за время от t до u попасть из y в множество A .

В теории марковских процессов иногда рассматриваются процессы, определенные при всех $t, -\infty < t < \infty$, но чаще имеется некоторый начальный момент времени (обычно $t = 0$), в который система возникает в некоторой случайной точке фазового пространства. Рассматривается, следовательно, мера π на пространстве X , определяемая соотношением $\pi(A) = P\{x(0) \in A\}$. Эта мера называется *начальным распределением вероятностей*.

Имея начальное распределение и переходную функцию, можно построить марковский процесс путем задания конечномерных распределений, как это полагается в аксиоматике Колмогорова. Определим, например, вероятность, отвечающую трем моментам времени $0 = s < t < u$ следующим равенством:

$$P\{x(s) \in A, x(t) \in B, x(u) \in C\} = \int_A \pi(dx) \int_B P_s^t(x, dy) P_t^u(y, C).$$

Согласованность подобных конечномерных распределений вытекает из уравнения Чэпмена-Колмогорова. При таком построении марковского процесса значения переходной функции оказываются условными вероятностями, причем условное распределение $x(t)$ относительно сигма-алгебры, порожденной случайной величиной $x(s)$, совпадает с условным распределением относительно сигма-алгебры, порожденной всеми случайными величинами $x(u)$ при $u \leq s$. Можно написать равенство

$$P(x(t) \in A | x(s) = x) = P(x(t) \in A | x(s) = x, x(u), u < s) = P_s^t(x, A).$$

Традиционное определение марковского свойства дается в виде последнего равенства. Но в приложениях более удобна несколько другая система определений.

Для начала рассматривается *марковская цепь*, т.е. случай, когда время принимает дискретные значения (их можно считать целочисленными $n=0, 1, 2, \dots$). В начальный момент $n=0$ возникает значение $x(0)$ с вероятностями, задаваемыми начальным распределением π . Затем рассматривается последовательность *независимых* случайных объектов (т.е., например, случайных величин или случайных процессов и т.д.) $\xi_1, \xi_2, \dots, \xi_n, \dots$ такая, что эволюция цепи задается по правилу $x(n) = F(n, x(n-1), \xi_n)$, где в правой части стоит некоторая функция от указанных

величин. В словесной форме: *следующее состояние цепи есть функция от предыдущего и от новой независимой случайности.*

Например, движение системы может задаваться системой обыкновенных дифференциальных уравнений, в правые части которых входят случайные процессы. В качественной форме мы предполагаем, что зависимость между значениями этих случайных процессов в различные моменты времени становится несущественной, если расстояние между этими моментами делается большим. В математической форме это выражается как ослабление зависимости между сигма-алгебрами событий, наблюдаемых по течению процессов на далеко удаленных друг от друга интервалах времени. Но эта математическая форма, кажется, не нашла приложений. Поэтому в приложениях обычно применяют модель *процессов с обновлением*. Именно, считается, что в целочисленные моменты времени одна реализация процесса прекращается и начинается новая, статистически независимая. В этом случае последовательность положений системы в целочисленные моменты автоматически оказывается цепью Маркова. Переходная функция этой цепи, в принципе, может быть рассчитана, если заданы распределения случайных объектов, функцией от которых является цепь. Однако некоторые недавние опыты в этом направлении выявили большие вычислительные трудности даже в случае, который априори представлялся несложным. Поэтому желателен не полный расчет переходных вероятностей, а расчет лишь их каких-то простейших параметров. Одно из направлений такого упрощения – это использование диффузионного приближения, о котором речь идет ниже.

2.2. Марковские диффузионные процессы.

Диффузионные марковские процессы выделяются из совокупности всех марковских процессов некоторыми условиями, налагаемыми на переходную функцию. Эти условия формулируются с точностью до $o(\Delta t)$, где Δt - стремящееся к нулю приращение времени.

Замечание. В математике все величины по умолчанию считаются безразмерными (в простейшем случае – разделенными каждая на некую единицу измерения). В физике, наоборот, по умолчанию величины размерные. Например, в математике вполне возможны функции e^x или $\sin t$, где x изображает расстояние по оси абсцисс от начала координат, а t – время, протекшее от некоторого условного нуля. Но в физике может быть только e^{ax} или $\sin \omega t$, где величины a и ω таковы, что ax и ωt безразмерны. В данном тексте используется математическое умолчание.

В терминах этого умолчания диффузионные процессы выделяются тем условием, что приращение $\Delta x(t) = x(t + \Delta t) - x(t)$ имеет порядок величины $\sqrt{\Delta t}$ (это относится и к случаю, когда $x(t)$ многомерно). Например, так будет для винеровского процесса, потому что для него $\mathbf{E}(\Delta w(t))^2 = \Delta t$, а следовательно, сама величина $\Delta w(t) = \xi \sqrt{\Delta t}$, где ξ - стандартная нормальная величина. Более точно, накладываются следующие условия.

1. Для любого $\varepsilon > 0$ справедливо соотношение $P(|\Delta x(t)| > \varepsilon | x(t) = x) = P_t^{t+\Delta t}(\overline{O_\varepsilon(x)}) = o(\Delta t)$, где $\overline{O_\varepsilon(x)}$ обозначает дополнение (не замыкание!) ε -окрестности точки x .

Замечание. Конечно, математик должен спросить, как именно понимается термин o : равномерно или неравномерно по переменным t и x , от которых зависит левая часть выписанного соотношения. Для вывода уравнения Колмогорова достаточно предположить, что равномерно в некоторой окрестности точки (t, x) .

Далее вместо приращений $\Delta x(t)$ рассмотрим «урезанные» приращения $\tilde{\Delta}x(t)$, определяемые формулами $\tilde{\Delta}x(t) = \Delta x(t)$, если $|\Delta x(t)| \leq \varepsilon$, и $\tilde{\Delta}x(t) = 0$ в противном случае. Это связано с тем, что в теории вероятностей даже близкие с большой вероятностью к нулю случайные величины могут все-таки не иметь математического ожидания за счет очень больших, хотя и маловероятных значений.

$$2. \mathbf{E}\{\tilde{\Delta}x(t) \mid x(t) = x\} = \int_{|y-x| \leq \varepsilon} (y-x) P_t^{t+\Delta t}(x, dy) = a(t, x)\Delta t + o(\Delta t).$$

Коэффициент $a(t, x)$ называется *коэффициентом сноса*. В случае многомерного x это соотношение понимается покомпонентно. Величина $a(t, x)$ не зависит от выбора ε . Действительно, если значение $\varepsilon = \varepsilon_1$ заменяется, допустим, на меньшее значение $\varepsilon_2 < \varepsilon_1$, то сокращается область интегрирования по y : из нее выбрасываются такие y , которые удовлетворяют неравенству $\varepsilon_2 < |y-x| \leq \varepsilon_1$. Но в силу условия 1. вероятность попадания y в эту область есть $o(\Delta t)$, а следовательно, интеграл изменится также на величину $o(\Delta t)$ (поскольку в выбрасываемой области подынтегральная функция ограничена).

$$3. \mathbf{E}\{[\tilde{\Delta}x(t)]^2 \mid x(t) = x\} = \int_{|y-x| \leq \varepsilon} (y-x)^2 P_t^{t+\Delta t}(x, dy) = \sigma^2(t, x)\Delta t + o(\Delta t).$$

Коэффициент $\sigma^2(t, x)$ называется *коэффициентом диффузии*. В многомерном случае учитываются дисперсии и ковариации для приращений различных координат процесса $x(t)$, так что соответствующее предположение приобретает вид

$$\mathbf{E}\{\tilde{\Delta}x_i(t)\tilde{\Delta}x_j(t) \mid x(t) = x\} = \int_{|y-x| \leq \varepsilon} (y_i - x_i)(y_j - x_j) P_t^{t+\Delta t}(x, dy) = \sigma_{ij}^{(2)}(t, x)\Delta t + o(\Delta t).$$

Квадратная симметричная неотрицательно определенная матрица с элементами $\sigma_{ij}^{(2)}(t, x)$ называется *матрицей коэффициентов диффузии*. В порядке величины Δt её элементы совпадают с дисперсиями и ковариациями элементов вектора $\tilde{\Delta}x(t)$. Элементы корня квадратного из этой матрицы можно обозначить $\sigma_{ij}(t, x)$.

Следует объяснить, каким образом математические ожидания самих величин $\tilde{\Delta}x(t)$ (которые предполагаются малыми: порядка $\sqrt{\Delta t}$) могут оказаться величинами того же порядка малости, как и их квадраты. Дело просто в том, что положительные и отрицательные значения этих приращений могут компенсировать друг друга при вычислении математических ожиданий.

Уравнение Колмогорова

Введем новое обозначение, которое позволяет более коротко записывать интегралы от функций по переходной вероятности: рассмотрим оператор T_s^t , действие которого на функцию f определяется формулой

$$T_s^t f(x) = \mathbf{E}\{f(x(t)) \mid x(s) = x\} = \int_x P_s^t(x, dy) f(y).$$

У этого оператора есть «теоретико-игровой», или «экономический» смысл, который состоит в следующем. Пусть в момент s игрок делает некоторую ставку на поведение процесса $x(\bullet)$ в некоторый будущий момент времени t . Какова должна быть справедливая ставка, если известно, что в момент t игрок получит сумму

$f(x(t))$? Классическая точка зрения состоит в том, что справедливая ставка равна математическому ожиданию выигрыша. Но поскольку процесс $x(\cdot)$ марковский, вся информация о его будущем поведении содержится в значении $x(s)$. Следовательно, математическое ожидание выигрыша должно вычисляться при условии $x(s) = x$, а это и есть $T'_s f(x)$. Примером сокращения формульных записей при использовании введенных операторов может быть следующий. Условие Чэпмена-Колмогорова, записанное в п.2.1 с помощью интегралов, может быть переписано в следующем виде: для моментов времени $s < t < u$ выполняется равенство $T'_s = T'_s T'_t$.

Операторы T'_s заведомо применимы к ограниченным функциям f (измеримым в смысле той сигма-алгебры, которая считается заданной на фазовом пространстве X). Но можно их рассматривать на более узких классах ограниченных функций, вводя дополнительные требования непрерывности или гладкости. Понятно, что задание этих операторов, скажем, на гладких финитных функциях однозначно определяет переходную функцию.

В предыдущем подпункте были введены коэффициенты сноса и диффузии, которые определяются по течению процесса $x(t)$ за сколь угодно малое время от t до $t + \Delta t$. В этом смысле введенные характеристики могут быть названы *инфинитезимальными* характеристиками $x(t)$. Сейчас мы выведем так называемое *уравнение Колмогорова на левом конце*, которое показывает, в частности, что по инфинитезимальным характеристикам диффузионного процесса, в принципе, восстанавливаются его переходные вероятности за любое время от s до t .

С этой целью фиксируем момент t , оставляя момент $s < t$ переменным, и введем функцию $g(s, x) = T'_s f(x)$. Предположим, что эта функция обладает достаточной для дальнейших вычислений гладкостью. Дадим s малое отрицательное приращение $(-\Delta s)$ и рассмотрим разность

$$g(s - \Delta s, x) - g(s, x) = T'_{s-\Delta s} g(s, x) - g(s, x) = \int_X P'_{s-\Delta s}(x, dy) g(s, y) - g(s, x).$$

Функцию $g(s, x)$ в правой части последнего равенства можно подвести под знак интеграла, поскольку она не зависит от переменной интегрирования y и верно равенство $\int_X P'_s(x, dy) = 1$. Полученный интеграл преобразуем следующим образом:

$$\int_X P'_{s-\Delta s}(x, dy) [g(s, y) - g(s, x)] = \left\{ \int_{|y-x| \leq \varepsilon} + \int_{|y-x| > \varepsilon} \right\} P'_{s-\Delta s}(x, dy) [g(s, y) - g(s, x)].$$

Интеграл по области $|y - x| > \varepsilon$ представляет собой величину порядка $o(\Delta s)$, а интеграл по области $|y - x| \leq \varepsilon$ мы сейчас вычислим с некоторой достаточной точностью. Для этого разложим при $|y - x| \leq \varepsilon$ разность $g(s, y) - g(s, x)$ по формуле Тейлора в точке x , используя две производных и предположив непрерывность второй производной. Точная формула Тейлора получится, если вторую производную взять в некоторой промежуточной между x и y точке, но мы возьмем её в точке x , сделав при этом ошибку не более некоторой величины $\delta(\varepsilon)$, стремящейся к нулю вместе с ε . Получим следующую формулу:

$$g(s, y) - g(s, x) = \frac{\partial g}{\partial x}(y - x) + \frac{1}{2} \frac{\partial^2 g}{\partial x^2}(y - x)^2 + \frac{1}{2} \delta(\varepsilon)(y - x)^2.$$

(Производные в этой формуле берутся в точке x и тем самым не зависят от y ; величина $\delta(\varepsilon)$ зависит, вообще говоря, от y , но равномерно по y в пределах окрестности $|y - x| \leq \varepsilon$ стремится к нулю.)

После интегрирования разности $(y-x)$ получим $a(s-\Delta s, x)\Delta s + o(\Delta s)$, где $a(s, x)$ - коэффициент сноса, после интегрирования $(y-x)^2$ получим $\sigma^2(s-\Delta s, x)\Delta s + o(\Delta s)$, где $\sigma^2(s, x)$ - коэффициент диффузии. Наконец, после интегрирования последнего члена $\frac{1}{2}\delta(\varepsilon)(y-x)^2$ получим остаток, который по абсолютной величине не больше, чем $\gamma(\varepsilon)\sigma^2(s, x)\Delta s$, где $\gamma(\varepsilon) \rightarrow 0$ вместе с $\varepsilon \rightarrow 0$. Собирая вместе члены, выражающие разность $g(s-\Delta s, x) - g(s, x)$, деля их сумму на Δs и переходя к пределу при $\Delta s \rightarrow 0$, но пока фиксированном ε , получаем (дополнительно предполагая непрерывность коэффициентов сноса и диффузии по s)

$$-\frac{\partial g}{\partial s} = a(s, x)\frac{\partial g}{\partial x} + \frac{1}{2}\sigma^2(s, x)\frac{\partial^2 g}{\partial x^2} + \gamma(\varepsilon)\sigma^2(s, x).$$

Устремляя в последнем равенстве ε к нулю, получаем окончательно уравнение

$$-\frac{\partial g}{\partial s} = a(s, x)\frac{\partial g}{\partial x} + \frac{1}{2}\sigma^2(s, x)\frac{\partial^2 g}{\partial x^2}.$$

В этом уравнении производные и коэффициенты сноса и диффузии берутся в точке (s, x) , т.е., так сказать, на левом конце интервала времени $[s, t]$. Поэтому уравнение называется *уравнением Колмогорова на левом конце*.

Замечание. Существует ещё уравнение Колмогорова *на правом конце*, но здесь оно не рассматривается.

Кроме самого уравнения, следует указать, какая рассматривается задача и почему она корректна. Если в качестве X берется в одномерном случае вся прямая (или в многомерном случае – все пространство), то не возникает границ и граничных условий. Возникает, однако, следующее начальное условие. Представим себе, что $s \uparrow t$, т.е. промежуток времени между s и t делается малым. Поскольку диффузионный процесс за малое время далеко уйти не может (а функцию f мы теперь предположим непрерывной), должно выполняться предельное соотношение $g(s, x) \rightarrow f(x)$ при $s \uparrow t$. Это и есть начальное условие. Известно, что уравнение теплопроводности, в котором перед производной $\frac{\partial}{\partial t}$ стоит знак $+$, корректно решается в сторону увеличения значений времени. Понятно, что уравнение, в левой части которого стоит $(-\frac{\partial g}{\partial t})$, будет корректно решаться в сторону уменьшения значений времени. Именно так и поставлена задача.

Дополнительные замечания к уравнению Колмогорова

1. В многомерном случае (т.е. когда $X = R^n$) формула Тейлора применяется к функции $g(s, x)$, в которой x многомерно. Возникает уравнение

$$-\frac{\partial g}{\partial s} = \sum_{i=1}^n a_i(s, x)\frac{\partial g}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^n \sigma_{ij}^{(2)}(s, x)\frac{\partial^2 g}{\partial x_i \partial x_j}.$$

Вывод этого уравнения ничем не отличается от одномерного случая.

2. Интересным частным случаем является однородный во времени диффузионный процесс, для которого коэффициенты сноса и диффузии не зависят от времени. В этом случае переходная функция зависит лишь от разности моментов времени $t-s$:

$$P'_s(x, A) = P(t-s, x, A).$$

Поэтому функция $g(s, x) = \mathbf{E}\{f(x(t)) | x(s) = x\}$ также зависит лишь от разности $t - s$. Сделаем замену $t - s = u$, тогда $s = t - u$ и $-\frac{\partial}{\partial s} = \frac{\partial}{\partial u}$, а начальное условие при $s \uparrow t$ переходит в условие при $u \downarrow 0$. Таким образом, функция $h(u, x) = g(t - u, x)$ удовлетворяет уравнению $\frac{\partial h}{\partial u} = a(x) \frac{\partial h}{\partial x} + \frac{1}{2} \sigma^2(x) \frac{\partial^2 h}{\partial x^2}$ и начальному условию $h(0, x) = f(x)$. Её вероятностный смысл следующий: $h(u, x) = \mathbf{E}\{f(x(u)) | x(0) = x\}$. В теории однородных марковских процессов вместо u пишут опять t и обозначают через A оператор $A = a(x) \frac{d}{dx} + \frac{1}{2} \sigma^2(x) \frac{d^2}{dx^2}$. Тогда для функции $h(t, x) = \mathbf{E}\{f(x(t)) | x(0) = x\}$ получается уравнение $\frac{\partial h}{\partial t} = Ah$, в котором оператор A применяется по переменной x . Оператор A называется *инфинитезимальным оператором* однородного марковского процесса.

2.5. Уравнение Колмогорова с учетом дисконтирования выплат

Напомним, что «экономический» смысл функции $g(s, x)$ - это математическое ожидание выплаты $f(x(t))$. Но будущие выплаты принято дисконтировать банковским процентом r , так что при непрерывном начислении процентов нужно рассматривать в качестве выплаты величину $e^{-r(t-s)} f(x(t))$. Рассмотрим соответствующее математическое ожидание, т.е. функцию $V(s, x) = \mathbf{E}\{e^{-r(t-s)} f(x(t)) | x(s) = x\} = e^{-r(t-s)} g(s, x)$. Выведем для нее уравнение на левом конце. Имеем

$$V(s - \Delta s, x) = e^{-r(t-s+\Delta s)} g(s - \Delta s, x) = e^{-r(t-s)} e^{-r\Delta s} g(s - \Delta s, x).$$

В правой части последнего равенства можно с ошибкой порядка $o(\Delta s)$ заменить $e^{-r\Delta s}$ на $1 - r\Delta s$ и затем произведение $(-r\Delta s g(s - \Delta s, x))$ на $(-r\Delta s g(s, x))$. В результате получится, что

$$V(s - \Delta s, x) \approx e^{-r(t-s)} \{g(s - \Delta s, x) - r\Delta s g(s, x)\}.$$

Поэтому

$$V(s - \Delta s, x) - V(s, x) \approx e^{-r(t-s)} \{g(s - \Delta s, x) - g(s, x)\} - r\Delta s V(s, x).$$

Учитывая уравнение для функции $g(s, x)$, получаем следующее уравнение:

$$-\frac{\partial V}{\partial s} = a(s, x) \frac{\partial V}{\partial x} + \frac{1}{2} \sigma^2(s, x) \frac{\partial^2 V}{\partial x^2} - rV$$

3. Уравнение Блэка-Шоулса для цены опциона

Это уравнение относится к модели динамики цен акций, в которой логарифм $\ln S(t)$ цены акции совершает броуновское движение со сносом a и коэффициентом диффузии σ^2 , причем выполняется специфическое условие $a + \frac{\sigma^2}{2} = r$. Функция выплаты по опциону имеет вид $f(S(T))$, где T - некоторый заранее оговоренный момент исполнения опциона. Подсчитывается, что для самого процесса $S(t)$ коэффициент сноса равен rS , а коэффициент диффузии есть $\sigma^2 S^2$. Согласно предыдущему, уравнение для цены опциона

$V(t, S)$ в момент $t < T$ имеет вид

$$\frac{\partial V}{\partial t} + rS \frac{\partial V}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - rV = 0.$$

При таком изложении уравнение Блэка-Шоулса выглядит простым частным случаем уравнения Колмогорова. Однако его решение обладает некоторым неожиданным свойством, которое, по-видимому, не возникало ранее ни в какой области математической физики. Выражение для цены опциона имеет форму портфеля ценных бумаг, состоящего из акций и банковского вклада, который является *самофинансируемым*. Самофинансируемость бывает только в рамках указанной выше модели (если моделировать цену акции каким-нибудь другим диффузионным процессом, то самофинансируемого портфеля не получится, однако в рамках принятой модели для самофинансируемости портфеля условие $a + \frac{\sigma^2}{2} = r$ не обязательно). Однако понятие самофинансируемости не обсуждается подробно в данном тексте.

4. Приложение: теорема Колмогорова о предельном переходе

Марковские цепи в математической физике возникают по правилу: *следующее состояние цепи есть функция от предыдущего и от новой независимой случайности*. Если случайные вмешательства за один шаг по времени невелики, то (в случае удачи при замене переменных) задача описания отклонений движения системы от невозмущенной траектории (которая была бы без случайных вмешательств) сводится к марковской цепи с малыми скачками. При этом нередко бывает достаточно ограничиться вычислением суммарных характеристик этих отклонений: среднего значения и среднего квадрата приращения за один шаг по времени, приближенно заменяя распределение вероятностей марковской цепи соответствующим распределением для диффузионного процесса. По традиции, принятой в теории вероятностей, математическим обоснованием такой замены точного распределения на приближенное является предельная теорема о слабой сходимости. В приложениях априори непонятно, хороша ли предлагаемая асимптотика. Однако речь идет о получении хоть какой-нибудь формулы, которая затем сопоставляется с экспериментальными или какими-то иными данными. (Физики убеждены в том, что формулы бывают умнее тех предпосылок, из которых они получены и даже умнее самих своих авторов.)

К сожалению, формулировки предельных теорем громоздки. В случае перехода от распределения цепи к распределению процесса нужно ввести последовательность марковских цепей со все уменьшающимися скачками, но при этом число скачков надо мыслить увеличивающимся. (Это так называемая *схема серий*.) Математическая запись условия теоремы требует громоздких обозначений, так что приходится сокращать обозначения в сравнении с тем, что обычно принимается в математических текстах: часть обозначений приходится подразумевать.

Схема серий выглядит следующим образом. Пусть имеется некий отрезок времени, скажем, $[0, 1]$. Возьмем последовательность разбиений этого отрезка точками $0 = t_0 < t_1 < \dots < t_n = 1$ (эти точки должны бы ещё иметь верхний индекс, отвечающий номеру разбиения n , но он подразумевается). Сопоставим каждому такому разбиению марковскую цепь, переходы в которой происходят в моменты, совпадающие с точками разбиения, и предположим, что эта цепь похожа по своим свойствам на диффузионный процесс на том же отрезке времени $[0, 1]$.

Чтобы записать математически эти свойства, введем переходную вероятность марковской цепи, которую подробно следовало бы обозначить ${}_{(n)}Q_s^t(x, A)$, где числа s и t могут принимать значения только из множества точек n -го разбиения отрезка времени. Но в дальнейшем мы опустим передний индекс (n) , а для перехода за один шаг, т.е. за время от t_i до t_{i+1} , примем обозначение $Q_i(x, A)$. Используем также обозначение $\Delta t_i = t_{i+1} - t_i$.

В этих обозначениях аналог предположения 1. для последовательности цепей можно записать так:

$$1. Q_i(x, \overline{O_\varepsilon(x)}) = o(\Delta t_i).$$

Аналоги предположений о коэффициентах сноса и диффузии запишутся следующим образом:

$$2. \int_{|y-x| \leq \varepsilon} (y-x) Q_i(x, dy) = a(t_i, x) \Delta t_i + o(\Delta t_i).$$

$$3. \int_{|y-x| \leq \varepsilon} (y-x)^2 Q_i(x, dy) = \sigma^2(t_i, x) \Delta t_i + o(\Delta t_i).$$

Далее предположим, что для диффузионного случайного процесса, определенного на отрезке $0 \leq t \leq 1$ и имеющего те же коэффициенты сноса $a(t, x)$ и диффузии $\sigma^2(t, x)$, что и для цепи, «хорошо» решается уравнение Колмогорова на левом конце (хорошо в том смысле, что при гладком ограниченном начальном условии $f(x)$ решение уравнения $g(t, x)$ имеет непрерывные ограниченные производные до второго порядка включительно). Обозначим через T_s^t соответствующий оператор для процесса, т.е. $g(t, x) = T_s^t f(x)$. Обозначим через ${}_n U_s^t$ аналогичный оператор для n -ой цепи.

При высказанных условиях имеет место следующая теорема о предельном переходе А.Н. Колмогорова.

Теорема. Для любой гладкой ограниченной функции $f(x)$ имеет место предельное соотношение

$$\lim_{n \rightarrow \infty} \sup_x |{}_n U_0^1 f(x) - T_0^1 f(x)| = 0.$$

Доказательство. Обозначим через T_i оператор для диффузионного процесса, отвечающий переходу за время от t_{i-1} до t_i , а через U_i - аналогичный оператор для цепи. В таком случае $T_0^1 = T_1 T_2 \dots T_n$ и ${}_n U_0^1 = U_1 U_2 \dots U_n$. Для удобства проведения рассуждения по индукции умножим эти произведения операторов справа на тождественный оператор, отвечающий переходу за время от 1 до 1. Теперь будем дописывать к этому тождественному оператору и далее по одному оператору слева и оценивать, как при этом меняется чебышевская норма (т.е. супремум модуля) для разности функций, отвечающих цепи и процессу. Положим $U_n^k = U_{k+1} \dots U_n$ и $T_n^k = T_{k+1} \dots T_n$ и посмотрим, насколько увеличится чебышевская норма разности при дописывании слева операторов U_k и T_k . Имеем

$$U_k U_n^k f - T_k T_n^k f = U_k (U_n^k f - T_n^k f) + (U_k - T_k) T_n^k f.$$

Оценим теперь норму левой части последнего равенства. Поскольку оператор U_k представляет собой интегрирование по марковской переходной функции, он не увеличивает чебышевской нормы; следовательно, норма первого слагаемого правой части не превосходит нормы выражения $U_n^k f - T_n^k f$ (которое было до дописывания добавочных множителей слева). Что касается второго слагаемого, то функция $T_n^k f = g(t_k, x)$ гладкая (как решение уравнения Колмогорова). Это

второе слагаемое выразим через интегралы от переходных функций. Переходная функция за время от t_i до t_{i+1} для цепи у нас обозначена $Q_i(x, A)$; обозначим ту же переходную функцию для процесса через $P_i(x, A)$. В этих обозначениях

$$(U_k - T_k)T_k^n f(x) = \int_X \{Q_{k-1}(x, dy) - P_{k-1}(x, dy)\} g(t_k, y)$$

Под знаком интеграла можно функцию $g(t_k, y)$ заменить на разность $g(t_k, y) - g(t_k, x)$, поскольку интегралы по y от обеих переходных функций равны 1. Напишем равенство

$$\begin{aligned} (U_k - T_k)T_k^n f(x) &= \int_X \{Q_{k-1}(x, dy) - P_{k-1}(x, dy)\} \{g(t_k, y) - g(t_k, x)\} = \\ &= \left\{ \int_{|y-x|>\varepsilon} + \int_{|y-x|\leq\varepsilon} \right\} \{Q_{k-1}(x, dy) - P_{k-1}(x, dy)\} \{g(t_k, y) - g(t_k, x)\} \end{aligned}$$

Интеграл по области $|y-x| > \varepsilon$ даст (в силу свойства 1) величину порядка $o(t_k - t_{k-1})$. В интеграле по области $|y-x| \leq \varepsilon$ воспользуемся формулой Тейлора

$$g(t_k, y) - g(t_k, x) = \frac{\partial g}{\partial x}(y-x) + \frac{1}{2} \frac{\partial^2 g}{\partial x^2}(y-x)^2 + \delta(\varepsilon)(y-x)^2,$$

где $\delta(\varepsilon) \rightarrow 0$ при $\varepsilon \rightarrow 0$. В силу свойств 2 и 3 для цепи и для процесса, интегралы от $y-x$ и $(y-x)^2$ дадут $o(t_k - t_{k-1})$ (производные от функции g берутся в точке (t_k, x) , т.е. от переменной интегрирования y не зависят). Интеграл от последнего члена $\delta(\varepsilon)(y-x)^2$ (в котором $\delta(\varepsilon)$ также может зависеть от y) даст величину, оцениваемую как $\gamma(\varepsilon)\sigma^2(t_k, x)(t_k - t_{k-1})$, где $\gamma(\varepsilon) \rightarrow 0$ при $\varepsilon \rightarrow 0$.

Допуская ограниченность коэффициента диффузии, получим, что увеличение чебышевской нормы разности от прибавления одного сомножителя слева составит не более $o(t_k - t_{k-1}) + C\gamma(\varepsilon)(t_k - t_{k-1})$. Суммирование всех этих увеличений и переход к пределу при $\varepsilon \rightarrow 0$ доказывают теорему (если предположить, что сходимость к нулю всех использованных бесконечно малых величин равномерна по всем переменным, от которых эти величины зависят).

Замечание. Таким образом, задача о движении динамической системы под действием случайных возмущений сводится (если возможно) путем замены переменных к марковской цепи с малыми скачками. Затем для этой цепи вычисляются коэффициенты сноса и диффузии путем учета действия случайных возмущений в первой и второй степени за один шаг марковской цепи. Если получаются величины примерно одного порядка, то можно сделать еще замену времени так, чтобы новое («медленное») время менялось на величину того же порядка за один шаг цепи. После этого возникает ситуация, описываемая в изложенной выше теореме Колмогорова. Учет малых случайных возмущений в первой и второй степени иногда называется *диффузионным приближением*.