Глава 4. Стохастические модели эволюции

Опираясь на сказанное о случайности в п. 2.2.1, коротко остановимся на стохастических моделях эволюции во времени (случайных процессах). Они могут быть заданы либо непосредственно указанием в явном виде их статистических свойств (плотностей распределения вероятностей, корреляционных функций и т.д., пп. 4.1-4.3), либо стохастическими разностными или дифференциальными уравнениями. Некоторые наиболее известные уравнения, их свойства и приложения обсуждаются в пп. 4.4 и 4.5.

4.1. Элементы теории случайных процессов

Если при фиксированных с заданной точностью начальных условиях $\mathbf{x}(t_0)$ и параметрах процесс демонстрирует одну и ту же временную реализацию, то естественно его динамическое описание (глава 3). Однако на практике эта ситуация часто не реализуется: при каждом новом испытании «в одних и тех же условиях» реализация процесса отличается от предыдущих. Такую неоднозначность связывают с влиянием многочисленных неконтролируемых факторов, которое всегда имеет место в реальном мире. При этом целесообразно отказаться от детерминированного описания и использовать аппарат теории вероятностей и теории случайных процессов [55, 53, 113, 150, 87].

4.1.1. Понятие случайного процесса

Случайный процесс (случайная функция времени) является обобщением понятия случайной величины, уместным для описания упомянутой ситуации. Более точно, его определение следующее.

Во-первых, *случайная функция* – это случайная величина, зависящая не только от случайного события ω, но и от какого-либо параметра. Если этот параметр – время, то случайная функция называется случайным процессом и обозначается $\xi(t,\omega)$. Величина ξ может быть как скалярной (скалярный случайный процесс), так и векторной (векторный или многомерный случайный процесс). Она может принимать как дискретные значения (процесс с дискретными состояниями), так и пробегать непрерывный ряд значений. Будем пока говорить для определенности только о последнем случае. Изучением и разработкой таких моделей занимается теория случайных процессов [53, 55]. В случае дискретного времени $t = 0,1,2,...$ случайный процесс называют также *случайной последовательностью*.

Для случайного процесса результатом одного испытания является не одно число (как для случайной величины), а функция $\xi(t,\omega_1)$, где ω_1 – это реализовавшееся в данном испытании случайное событие. Случайное

событие можно интерпретировать как совокупность случайных факторов, действовавших во время протекания процесса и влиявших на него. Функция $\xi(t,\omega_1)$ называется *реализацией* случайного процесса. Это уже детерминированная (неслучайная) функция времени, поскольку случайное событие фиксировано $\omega = \omega_1$. В результате различных испытаний получаются, вообще говоря, различные реализации. Множество реализаций, полученных в результате разных испытаний (при различных ω) называют *ансамблем* реализаций (рис.4.1).

4.1.2. Характеристики случайного процесса

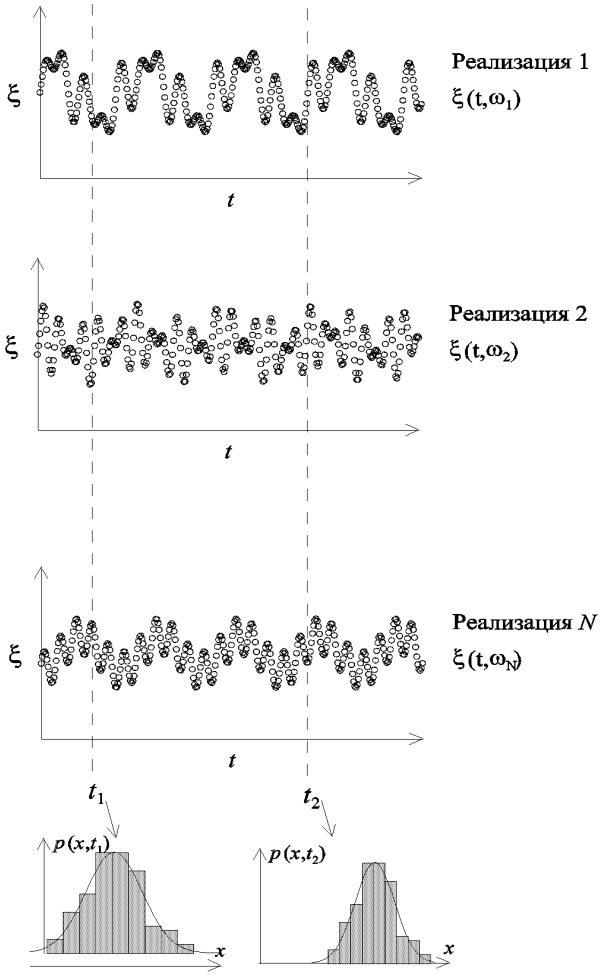

Рис.4.1. Ансамбль из N реализаций (три показаны) и два сечения случайного процесса

В любой фиксированный момент времени t значение случайного процесса $\xi(t,\omega)$ является случайной величиной. Она называется *сечением* случайного процесса в момент времени t и характеризуется плотностью распределения вероятностей, которую обозначим $p(x,t)$. Этот закон распределения называется *одномерным распределением* случайного процесса. Оно зависит от параметра – времени t – и может быть различным для двух различных моментов времени. Зная одномерный закон распределения процесса $p(x,t)$, можно вычислить его математическое ожидание и дисперсию в момент времени t. Если закон распределения меняется в зависимости от времени, то могут меняться и математическое ожидание процесса

$$m(t) = M\left[\xi(t,\omega)\right] = \int_{-\infty}^{\infty} x p(x,t)\,dx, \qquad (4.1)$$

и его дисперсия:

$$\sigma_\xi^2(t) = M\left[\xi(t,\omega) - m(t)\right]^2 = \int_{-\infty}^{\infty}\left[x - m(t)\right]^2 p(x,t)dx\,. \qquad (4.2)$$

Это детерминированные (неслучайные) функции времени, поскольку зависимость от случайных событий исключена при усреднении.

Сечения $\xi(t,\omega)$ в различные моменты времени t_1 и t_2 имеют, вообще говоря, разные функции плотности распределения $p(x,t_1)$ и $p(x,t_2)$, см. рис.4.1. Совместное поведение сечений описывается двухмерной плотностью распределения $p_2(x_1,t_1,x_2,t_2)$. Таким же образом можно задать n-мерные законы распределения p_n для всевозможных наборов t_1, t_2, …, t_n. Они образуют совокупность конечномерных распределений случайного процесса $\xi(t,\omega)$. Вероятностные свойства процесса полностью определены, только если задана вся эта совокупность. Однако, поскольку она представляет собой бесконечное множество законов распределения, то в общем случае полностью описать случайный процесс практически невозможно.

Реально приходится ограничиваться использованием только некоторых характеристик, например, одно- и двухмерных распределений, моментов невысоких порядков – математического ожидания, дисперсии, автоковариационной функции – и т.д. Так, автоковариационная функция в общем случае зависит от двух аргументов:

$$\begin{aligned} K(t_1,t_2) &= M\big(\xi(t_1,\omega) - m(t_1)\big)\big(\xi(t_2,\omega) - m(t_2)\big) = \\ &= \iint \big(x_1 - m(t_1)\big)\big(x_2 - m(t_2)\big)p_2(x_1,t_1,x_2,t_2)dx_1dx_2\,. \end{aligned} \qquad (4.3)$$

Для фиксированных t_1 и t_2 выражение (4.3) определяет *ковариацию* случайных величин $\xi(t_1,\omega)$ и $\xi(t_2,\omega)$. Если нормировать ее на среднеквадратичные отклонения, получим *автокорреляционную функцию* $\rho(t_1,t_2) = K(t_1,t_2)/\big(\sigma_\xi(t_1)\sigma_\xi(t_2)\big)$, т.е. коэффициент корреляции между случайными величинами $\xi(t_1,\omega)$ и $\xi(t_2,\omega)$.[1] Автокорреляционная функция принимает значения в интервале от –1 до 1; $|\rho(t_1,t_2)| = 1$ соответствует детерминированной линейной зависимости $\xi(t_1,\omega) = const \cdot \xi(t_2,\omega)$.

Для характеристики процессов часто используются условные одномерные распределения $p_1(x,t\,|\,x_1,t_1)$, т.е. распределения сечения $\xi(t)$, при условии, что в момент времени t_1 величина ξ приняла значение $\xi(t_1) = x_1$. Функцию $p_1(x,t\,|\,x_1,t_1)$ называют плотностью вероятностей

[1] Есть и другая терминология, согласно которой (4.3) называется автокорреляционной функцией, а после нормировки – нормированной автокорреляционной функцией [172]. Мы ее не используем.

перехода из состояния x_1 (в момент времени t_1) в состояние x (в момент времени t).

4.1.3. Стационарность и эргодичность случайных процессов

Важным свойством процесса является его *стационарность* (или нестационарность). *Стационарным в узком смысле* называют такой процесс, все конечномерные распределения которого не меняются при сдвиге по времени, т.е. $\forall n, t_1, ..., t_n, \tau$ $p_n(x_1, t_1, ..., x_n, t_n) = p_n(x_1, t_1 + \tau, ..., x_n, t_n + \tau)$. Другими словами, при сдвиге по времени не меняется ни одна из характеристик процесса. *Стационарным в широком смысле* называют такой случайный процесс, для которого при сдвиге по времени не меняются математическое ожидание, дисперсия и корреляционная функция (т.е. моменты до второго порядка включительно). Для стационарных (как в узком, так и в широком смысле) процессов имеет место $m(t) = const$, $\sigma_\xi^2(t) = const$, $K(t_1, t_2) = k(\tau), \tau = t_2 - t_1$ Из стационарности в узком смысле следует стационарность в широком смысле.[2]

Вообще говоря, термин «стационарность» означает неизменность какого-то свойства во времени. Если интересующее нас свойство (например, момент одномерного распределения порядка n) не меняется, то процесс называют стационарным относительно этого свойства.

Эргодическим называют процесс, все характеристики которого могут быть получены по одной его реализации (бесконечно длинной). Например, математическое ожидание определится как $m = \lim\limits_{T \to \infty} \frac{1}{T} \int\limits_0^T \xi(t, \omega_1) dt$ для почти любого ω_1, т.е. усреднение по времени и усреднение по пространству дают одинаковые результаты. Если по одной реализации можно получить *все* характеристики процесса, то его называют эргодическим *в строгом смысле*. Если можно получить только некоторые характеристики, то его называют эргодическим относительно этих характеристик: вводят понятия эргодичности первого порядка – относительно первых моментов, и т.п.

Эргодические процессы представляют собой важный класс, поскольку на практике зачастую имеется только одна реализация, а не большой ансамбль. По этой единственной реализации только для эргодического процесса можно восстановить его свойства. Поэтому часто при анализе отдельного временного ряда принимают гипотезу об эргодичности

[2] Есть и несколько иная трактовка, связанная с дополнительным требованием *конечной* дисперсии для стационарного в широком смысле процесса, при этом такое следование не будет иметь места.

исследуемого процесса. Эргодический процесс является стационарным, но не всякий стационарный процесс является эргодическим.

Пример (реальный аналог) случайного процесса дает практически любое физическое измерение. Это может быть измерение силы тока в какой-либо нелинейной цепи, демонстрирующей автоколебания. От испытания к испытанию измеренные реализации (временные ряды, см. главу 5) отличаются из-за наличия тепловых шумов, наводок, и т.п. Причем, получив реализацию на каком-то отрезке времени, нельзя однозначно и точно предсказать ее дальнейшее поведение, поскольку оно определяется случайными факторами, которые будут влиять на процесс и в будущем.

Более простой пример случайного процесса – сильно упрощенное модельное представление об испускании фотона возбужденным атомом. Момент испускания, начальная фаза, направление и поляризация непредсказуемы. Но как только фотон испущен и его начальное поведение стало известно, то все последующее уже однозначно предсказуемо. Случайный процесс здесь описывается гармонической функцией времени со случайной начальной фазой (гармонический шум, см. ниже). Случайные процессы такого типа называют *квазидетерминированными* [172], поскольку случайные факторы определяют только начальные условия, а затем поведение подчиняется детерминированной закономерности.

4.1.4. Оценки характеристик случайных процессов

Для того чтобы получить статистические оценки одномерного закона распределения $p(x,t)$ и его моментов, можно провести много испытаний и получить набор реализаций процесса $\xi(t,\omega_1)$, $\xi(t,\omega_2)$, …, $\xi(t,\omega_n)$. Для фиксированного момента времени $t = t^*$ набор значений этих реализаций составляет выборку значений случайной величины $\xi(t^*,\omega)$ объема n (рис.4.1). По этой выборке можно оценить закон распределения $p(x,t^*)$ и т.п. То же самое нужно проделать для всех других моментов времени.

Многомерные законы распределения можно оценивать по ансамблю реализаций аналогично. Но число реализаций для их надежной оценки должно быть существенно больше, чем для оценки одномерного распределения и, тем более, для оценки моментов.

Сложнее ситуация, когда имеется только одна реализация. Только для эргодического процесса по достаточно длинной реализации можно оценить нужные характеристики, заменив усреднение по ансамблю усреднением по времени, см. п.4.1.3.

4.2. Базовые модели случайных процессов

Случайный процесс можно задать, определив явно конечномерные законы распределения вероятностей. Так вводятся базовые (самые простые) модели теории случайных процессов. Рассмотрим некоторые из них [55].

1) Одним из важнейших в теории случайных процессов является *нормальный* (*гауссовский*) случайный процесс. Это процесс, для которого все конечномерные законы распределения являются нормальными. А именно, *n*-мерный закон распределения процесса имеет вид:

$$p_n(x_1,t_1,...,x_n,t_n)=\frac{1}{\sqrt{(2\pi)^n|\mathbf{V}_n|}}\exp\left(-\frac{1}{2}(\mathbf{x}_n-\mathbf{m}_n)^T\cdot\mathbf{V}_n^{-1}\cdot(\mathbf{x}_n-\mathbf{m}_n)\right) \quad (4.4)$$

для любого *n*, где введены обозначения:

$$\mathbf{x}_n=\begin{bmatrix}x_1\\x_2\\...\\x_n\end{bmatrix},\mathbf{m}_n=\begin{bmatrix}m(t_1)\\m(t_2)\\...\\m(t_n)\end{bmatrix},\mathbf{V}_n=\begin{bmatrix}K(t_1,t_1) & K(t_1,t_2) & ... & K(t_1,t_n)\\K(t_2,t_1) & K(t_2,t_2) & ... & K(t_2,t_n)\\... & ... & ... & ...\\K(t_n,t_1) & K(t_n,t_2) & ... & K(t_n,t_n)\end{bmatrix}, \quad (4.5)$$

$m(t)$ – математическое ожидание процесса, $K(t_1,t_2)$ – автокорреляционная функция, *T* означает транспонирование, $|\mathbf{V}_n|$ – определитель матрицы $\mathbf{V}_n$. В данном примере все конечномерные распределения известны (процесс определен полностью), если заданы математическое ожидание и корреляционная функция. При любых линейных преобразованиях нормального процесса он остается нормальным.

2) Процесс *с независимыми приращениями*. Так называется процесс, для которого величины $\xi(t_1,\omega),\xi(t_2,\omega)-\xi(t_1,\omega)$, …, $\xi(t_n,\omega)-\xi(t_{n-1},\omega)$ (т.е. приращения) являются статистически независимыми для любых *n*, $t_1,...,t_n$, таких, что $n>1$ и $t_1<t_2<...<t_n$.

3) *Винеровский* процесс. Это *N*-мерный случайный процесс с независимыми приращениями, для которого при любых $t_1<t_2$ случайный вектор $\boldsymbol{\xi}(t_2,\omega)-\boldsymbol{\xi}(t_1,\omega)$ распределен по нормальному закону с нулевым математическим ожиданием и ковариационной матрицей вида $(t_2-t_1)s^2I_n$, где I_n – единичная матрица порядка *n*. Это нестационарный процесс, в одномерном случае его дисперсия растет линейно во времени как $\sigma^2(t)=\sigma^2(t_0)+s^2\cdot(t-t_0)$.

Винеровский процесс описывает, например, броуновское движение – перемещение броуновской частицы под действием случайных независимых соударений с молекулами окружающей среды.

Можно показать, что винеровский процесс является нормальным процессом. Винеровский процесс с $s=1$ называется стандартным.

4) *Марковский* процесс (первого порядка) – это случайный процесс, для которого условная функция плотности вероятностей имеет вид: $p_1(x_n,t_n \mid x_{n-1},t_{n-1},...,x_1,t_1)=p_1(x_n,t_n \mid x_{n-1},t_{n-1})$, для любых $n,t_1,...,t_n$, $t_1<t_2<...<t_n$. Это выражают фразой «будущее зависит от прошлого только через настоящее». Любой конечномерный закон распределения такого процесса выражается через его одномерный и двухмерный законы. Можно показать, что винеровский процесс является марковским.

Важным частным случаем является марковский процесс с конечным числом K возможных состояний. Обозначим их $S_1,...,S_K$. Из-за дискретности состояний, он описывается в терминах самих вероятностей, а не плотностей их распределения. Условная вероятность $\mathrm{P}\{\xi(t+\Delta t)=S_j \mid \xi(t)=S_i\}$ называется в этом случае переходной вероятностью, поскольку описывает переход из состояния i в состояние j. А величина $\lambda_{i,j}(t)=\lim\limits_{\Delta t\to+0}\mathrm{P}\{\xi(t+\Delta t)=S_j \mid \xi(t)=S_i\}/\Delta t$ называется плотностью вероятности соответствующего перехода.

Марковские процессы занимают особое место в теории случайных процессов, им посвящено множество исследований.

5) *Пуассоновским* процессом с параметром $\lambda>0$ называют скалярный случайный процесс с дискретными состояниями, обладающий следующими свойствами: а) $\xi(0,\omega)=0$; б) приращения процесса независимы; в) при любых $0\le t_1<t_2$ величина $\xi(t_2,\omega)-\xi(t_1,\omega)$ распределена по закону Пуассона с параметром $\lambda(t_2-t_1)$, т.е. по закону $\mathrm{P}\{\xi(t_2,\omega)-\xi(t_1,\omega)=k\}=\dfrac{\lambda(t_2-t_1)^k}{k!}\exp(-\lambda(t_2-t_1))$, где k – целое неотрицательное число. Пуассоновский процесс часто используют в приложениях, например, в теории массового обслуживания.

6) *Белый шум*. Это стационарный в широком смысле (согласно одной из трактовок, см. п.4.1.3) случайный процесс, значения которого в различные моменты времени некоррелированы, т.е. его автоковариационная функция имеет вид: $k(\tau)=const\cdot\delta(\tau)$. Его называют «белым», потому что спектр мощности этого процесса представляет собой константу, т.е. в нем равноправно представлены все частоты. Здесь проведена аналогия с белым светом, в котором представлены все частоты (все «цвета») видимой части спектра. Дисперсия белого шума бесконечна: $\sigma_\xi^2=k(0)=\infty$.

Распространенной моделью является нормальный белый шум (НБШ). Это стационарный процесс, имеющий нормальный одномерный закон

распределения и корреляционную функцию вида $k(\tau) = const \cdot \delta(\tau)$. Строго говоря, это сочетание требований противоречиво, поскольку белый шум имеет бесконечную дисперсию, а нормальный случайный процесс – конечную дисперсию σ_ξ^2. Тем не менее, несколько противоречивое понятие НБШ полезно на практике и при исследовании стохастических дифференциальных уравнений, см. ниже п. 4.5. Для практической интерпретации можно считать, что НБШ – это процесс, имеющий очень большую дисперсию, а интервал времени, на котором его корреляционная функция спадает почти до нуля, очень мал по сравнению с другими характерными масштабами рассматриваемой задачи.

7) Аналог белого шума в случае дискретного времени – последовательность *независимых одинаково распределенных* случайных величин. Этот процесс тоже часто называют белым шумом. Чаще всего рассматривают нормальное одномерное распределение, хотя возможно и любое другое. В случае дискретного времени дисперсия процесса σ_ξ^2 конечна, так что он является стационарным в широком смысле, независимо от используемого определения этой стационарности.

Белый шум – это «самый непредсказуемый» процесс, т.к. отсутствует зависимость между последовательными значениями. Последовательность независимых нормально распределенных величин используется как базовая модель при конструировании стохастических моделей с дискретным временем – стохастических разностных уравнений, см. п. 4.4.

8) *Цепь Маркова* – это марковский процесс с дискретными состояниями и с дискретным временем. Эта простая модель очень широко используется на практике. Ее главными характеристиками служат вероятности перехода из одного состояния в другое. Для анализа и наглядного представления таких моделей используется аппарат теории графов.

4.3. Уравнения эволюции распределения вероятностей

Выше перечислены эталонные случайные процессы, полученные из общих соображений. Так, нормальный случайный процесс можно получить, исходя из соображений о наличии большого числа независимых факторов, белый шум – из представлений о независимости последовательных значений, пуассоновский процесс – с использованием предположения о редких событиях [55]. Об этих трех процессах известно «все существенное»: конечномерные распределения, моменты и т.д.

Что касается марковских процессов, то в них заложены представления о связи будущих состояний с предыдущими. В общем случае марковский процесс нестационарен. Таким образом, возникает вопрос о том, как будет меняться во времени начальное распределение вероятностей – будет ли

оно сходится к какому-либо стационарному и каково будет это предельное распределение? В определении марковского процесса это непосредственно не сформулировано. Однако, опираясь на это определение, можно вывести уравнения эволюции для закона распределения вероятностей. Для процесса с конечным числом состояний они запишутся в виде системы обыкновенных дифференциальных уравнений (уравнений Колмогорова):

$$\begin{bmatrix} dp_1/dt \\ dp_2/dt \\ \dots \\ dp_K/dt \end{bmatrix} = \begin{bmatrix} -(\lambda_{12}+\dots+\lambda_{1K}) & \lambda_{21} & \dots & \lambda_{K,1} \\ \lambda_{12} & -(\lambda_{21}+\lambda_{23}+\dots+\lambda_{2K}) & \dots & \lambda_{K2} \\ \dots & \dots & \dots & \dots \\ \lambda_{1K} & \lambda_{2K} & \dots & -(\lambda_{K1}+\dots+\lambda_{1K,K-1}) \end{bmatrix} \cdot \begin{bmatrix} p_1 \\ p_2 \\ \dots \\ p_K \end{bmatrix}, \quad (4.6)$$

где $p_i(t)$ – вероятность состояния S_i, а $\lambda_{i,j}(t)$ – плотности вероятностей перехода. Если функции $\lambda_{i,j}(t)$ заданы, то, интегрируя уравнения Колмогорова, можно проследить за эволюцией распределения вероятностей из любого начального распределения. В простых частных случаях, например, при постоянных $\lambda_{i,j}$, решение можно найти аналитически.

В случае цепей Маркова задача несколько упрощается (по крайней мере, для численного исследования) – эволюция вектора вероятностей описывается разностным уравнением порядка *K*. Для наглядного представления марковских процессов с дискретными состояниями часто используют графы, на которых кружками указывают различные состояния, а стрелками – возможные переходы между состояниями.

В случае непрерывнозначных марковских процессов состояние нужно описывать уже не вектором вероятностей, а функцией плотности распределения вероятностей. Поэтому вместо обыкновенных дифференциальных уравнений (4.6) для описания эволюции закона распределения вероятностей получают уравнения в частных производных. Это обобщенное уравнение Маркова (другие названия – уравнение Колмогорова – Чепмена, прямое уравнение Колмогорова) для условной плотности распределения:

$$\frac{\partial p(x,t \mid x_0,t_0)}{\partial t} = \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \frac{\partial^k}{\partial x^k} \left[c_k(x,t)\, p(x,t \mid x_0,t_0) \right], \quad (4.7)$$

где величины $c_k(x,t) = \lim\limits_{\Delta t \to 0} \frac{1}{\Delta t} \int\limits_{-\infty}^{\infty} (x'-x)^k\, p(x',t+\Delta t \mid x,t)\, dx'$ – это коэффициенты, связанные с «вероятностями изменения» состояния x и определяющие «гладкость» реализаций процесса.

В важном частном случае *диффузионного марковского процесса* ($c_k = 0$ при любом $k > 2$) уравнение упрощается и приводится к виду:

$$\frac{\partial p(x,t)}{\partial t} = -\frac{\partial}{\partial x}\big(c_1(x,t)p(x,t)\big) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\big(c_2(x,t)p(x,t)\big), \qquad (4.8)$$

где c_1 называют коэффициентом сноса, а c_2 – коэффициентом диффузии. Уравнение (4.8) называют также уравнением Фоккера – Планка. Это уравнение параболического типа, такой же вид имеют уравнения диффузии и теплопроводности в математической физике. Отсюда и названия коэффициентов уравнения. О связи коэффициентов сноса и диффузии в уравнении Фоккера – Планка с параметрами стохастических дифференциальных уравнений, задающих исходный процесс, см. ниже п.4.5.

4.4. Процессы авторегрессии – скользящего среднего

Случайный процесс можно задать с помощью стохастических уравнений. При этом он определяется как решение стохастического уравнения, т.е. при его подстановке в уравнение оно обращается в тождество. Здесь мы рассмотрим стохастические разностные уравнения, определяющие случайный процесс «авторегрессии – скользящего среднего» [45], который является одной из наиболее популярных конструкций при моделировании по наблюдаемым временным рядам.

Линейный фильтр. В качестве базовой модели при описании сложных реальных процессов часто принимают нормальный белый шум $\xi(t)$. Для простоты выкладок будем полагать, что он имеет нулевое среднее и дисперсию σ_ξ^2. При анализе реального сигнала его свойства могут противоречить гипотезе о том, что это – нормальный белый шум (например, оценки значений автокорреляционной функции $\rho(\tau)$ могут быть существенно отличны от нуля для ненулевых τ). В этом случае полезным для многих практических ситуаций оказался следующий подход. Рассмотрим нормальный белый шум, преобразованный *линейным фильтром*. В общем случае это преобразование определяется выражением

$$x_n = \xi_n + \sum_{i=1}^{\infty} \psi_i \xi_{n-i}\,. \qquad (4.9)$$

Чтобы дисперсия процесса x_n была конечна (чтобы он был стационарным), веса ψ_i должны удовлетворять условию $\sum_{i=1}^{\infty}\psi_i^2 \le const$. Линейное преобразование (4.9) сохраняет нормальность процесса и вносит ненулевые автокорреляции $\rho(\tau)$ при ненулевых сдвигах.

Процессы скользящего среднего. Разумеется, на практике использовать модель с бесконечным количеством весов не представляется

возможным. Однако, как правило, при описании свойств реальных процессов разумно предположить, что значения ψ_i быстро убывают с ростом номера i (далекое прошлое слабо влияет на настоящее) и ограничиться моделью (4.9) с конечным числом весов q. Получим процесс скользящего среднего порядка q (обозначается СС(q) или MA(q), от английского «Moving Average» — скользящее среднее), который задается разностным уравнением

$$x_n = a_n - \sum_{i=1}^{q} \theta_i \xi_{n-i} \tag{4.10}$$

и имеет q+1 параметров – веса θ_1, θ_2, …, θ_q и σ_ξ^2.

Процессы авторегрессии. Заметим теперь, что общее выражение (4.9) можно эквивалентно переписать в виде:

$$x_n = \xi_n + \sum_{i=1}^{\infty} \pi_i x_{n-i}, \tag{4.11}$$

где веса π_i выражаются через ψ_i. Подробнее, переход от (4.9) к (4.11) можно осуществить следующим образом: нужно последовательно исключать из выражения (4.9) величины ξ_{n-1}, ξ_{n-2} и т.д. Для этого сначала выразим значение шума ξ_{n-1} через x_{n-1} и предыдущие значения ξ с помощью формулы $\xi_{n-1} = x_{n-1} - \sum_{i=1}^{\infty} \psi_i \xi_{n-1-i}$. Подставим это выражение в формулу (4.9), исключив из нее, таким образом, ξ_{n-1}. Далее аналогично исключим ξ_{n-2}, и т.д. Процесс (4.11) также содержит бесконечное количество параметров π_i. Но часто имеет место быстрое убывание весов $\pi_i \to 0$ при $i \to \infty$. Т.е. далекое прошлое слабо влияет на настоящее, но уже в терминах зависимости текущего значения процесса от предыдущих значений самого процесса. Тогда достаточно ограничиться конечным числом слагаемых в (4.11). В результате приходим к процессу авторегрессии порядка p (обозначается АР(p) или AR(p), от английского «AutoRegressive» — авторегрессионный), который задается разностным уравнением

$$x_n = \xi_n + \sum_{i=1}^{p} \phi_i x_{n-i}\,. \tag{4.12}$$

Эта конструкция содержит p+1 параметров – веса ϕ_1, ϕ_2, …, ϕ_p и дисперсию σ_ξ^2, причем значения весов должны удовлетворять определенным соотношениям [45], чтобы процесс был стационарным. Так,

в случае $p=1$ дисперсия процесса (4.12) выражается через дисперсию σ_ξ^2 как $\sigma_x^2=\sigma_\xi^2/(1-\phi_1^2)$ и для стационарности необходимо $|\phi_1|<1$. Термин «авторегрессия» появился из-за того, что сумма в (4.12) определяет регрессию текущего значения процесса x_n на предыдущие значения самого же процесса – отсюда «авто». О термине «регрессия» см. п.7.2.

Процессы авторегрессии – скользящего среднего. Более эффективную для описания широкого класса процессов конструкцию, можно получить, объединив (4.10) и (4.12). Целесообразность такого объединения вместо использования только одного из представлений (4.10) или (4.12) вызвана следующими обстоятельствами. Предположим, что наблюдаемый временной ряд генерируется процессом авторегрессии порядка 1. Если попытаться описать его процессом скользящего среднего, то потребуется модель (4.10) с бесконечным числом параметров θ_i (по крайней мере, с очень большим). При моделировании по экспериментальным данным оценки значений большого числа параметров менее надежны, и это обязательно приведет к существенному снижению качества модели. И обратно, если ряд генерируется процессом скользящего среднего порядка 1, то для его описания потребовался бы процесс авторегрессии очень высокого порядка. Поэтому разумно объединить в модели выражения (4.10) и (4.12), чтобы можно было экономично (при помощи небольшого числа параметров) описать наблюдаемый процесс и вида (4.10), и вида (4.12), и смешанный. Получаем процесс авторегрессии и скользящего среднего порядка (p, q) (обозначается АРСС(p, q) или ARMA(p, q)):

$$x_n=\xi_n+\sum_{i=1}^{p}\phi_i x_{n-i}-\sum_{i=1}^{q}\theta_i\xi_{n-i}\ . \tag{4.13}$$

Он зависит от p+q+1 параметров.

Процессы авторегрессии – проинтегрированного скользящего среднего. Стационарный процесс (4.13) не может быть адекватной моделью для описания нестационарных процессов с детерминированным трендом или стохастическим трендом (*нерегулярные* чередования интервалов, на которых процесс следует почти детерминированной зависимости). Однако для некоторого класса трендов (а именно, полиномиальных) адекватной моделью является процесс, конечная разность которого является стационарным АРСС-процессом. Конечная разность порядка d определяется как $y_n=\nabla^d x_n$, где $\nabla x_n=x_n-x_{n-1}$ — первая разность (аналог дифференцирования), а ∇^d означает последовательное применение d раз оператора ∇. Таким образом, получаем процесс авторегрессии и проинтегрированного скользящего среднего порядка (p, d, q) (обозначается АРПСС(p, d, q) или ARIMA(p, d,

q) от английского — «AutoRegressive Integrated Moving Average»), который определяется разностными уравнениями

$$\begin{aligned} &y_n = \xi_n + \mu + \sum_{i=1}^{p} \phi_i y_{n-i} - \sum_{i=1}^{q} \theta_i \xi_{n-i}, \\ &\nabla^d x_n = y_n. \end{aligned} \tag{4.14}$$

Постоянный член μ определяет наличие детерминированного тренда. Чтобы выразить значения процесса x_n через значения АРСС-процесса y_n, нужно использовать оператор суммирования (аналог интегрирования), обратный оператору ∇. Этим объясняется наличие слова «проинтегрированный» в названии процесса.

АРСС- и АРПСС-процессы более полувека (1920-1970-е гг.) были основным аппаратом для моделирования и прогноза сложных процессов на практике. Они широко использовались для решения задач управления техническими процессами [45, ч.2]. Развивались и их всевозможные модификации, в частности, сезонные АРПСС-модели, которые определяются как АРПСС-процессы для сезонной разности $\nabla_s x_n = x_n - x_{n-s}$ вида

$$\begin{aligned} &y_n = \xi_n + \sum_{i=1}^{P} \Phi_i y_{n-is} - \sum_{i=1}^{Q} \Theta_i \xi_{n-is}, \\ &\nabla_s^D x_n = y_n, \end{aligned} \tag{4.15}$$

где ξ_n– процесс АРПСС(p,d,q). Процесс (4.15) называется сезонным АРПСС-процессом порядка $(P,D,Q)\times(p,d,q)$. Такие модели уместны для описания процессов с сезонными трендами (характерными масштабами порядка s).

Только в последние два десятилетия с развитием вычислительной техники и нелинейной динамики АРПСС-модели все больше «отступают» в конкуренции с нелинейными моделями (см. главы 8-11), хотя во многих отраслях знания остаются основным инструментом.

4.5. Стохастические дифференциальные уравнения и белый шум

Для описания случайных процессов в непрерывном времени используются стохастические дифференциальные уравнения (СДУ). Наиболее известно уравнение первого порядка (так называемое уравнение Ланжевена):

$$dx(t)/dt = F(x,t) + G(x,t)\cdot\xi(t), \tag{4.16}$$

где F, G – гладкие функции, $\xi(t)$ – нормальный белый шум с автоковариационной функцией $k(\tau)=\delta(\tau)$.

Введение понятия СДУ совсем не тривиально, поскольку оно включает в себя понятие производной случайного процесса $dx(t)/dt$. Как ее понимать? Наиболее простой путь был бы следующим: считать, что все реализации процесса x непрерывно дифференцируемы, и определить производную как случайную величину, значение которой есть обычная производная реализации x. Но это возможно только для процессов $\xi(t)$ с непрерывными реализациями – для каждой отдельной реализации $\xi(t)$ уравнение (4.16) представляло бы собой ОДУ. Но, например, белый шум не относится к таким процессам, а хорошо описывает многие практические ситуации (серию независимых толчков) и позволяет упростить многие выкладки. Для возможности анализа уравнения (4.16) с процессами $\xi(t)$ такого типа понятие производной случайного процесса $dx(t)/dt$ в точке t обобщили. Производная есть случайная величина $\frac{dx(t)}{dt}=\underset{\Delta t\to 0}{l.i.m.}\frac{x(t+\Delta t)-x(t)}{\Delta t}$, где предел понимается в среднеквадратичном смысле, см., например, [55].

Процесс x в (4.16) может быть и векторным. Тогда и белый шум $\xi(t)$ – многомерный процесс.

Можно показать, что процесс x в (4.16) – марковский. Для него можно записать уравнение Фоккера – Планка, где коэффициент сноса есть F, а коэффициент диффузии – G^2.

Рассмотрим частный случай (4.16): $F=0$, $G=const$, т.е. уравнение

$$dx(t)/dt=\xi(t). \tag{4.17}$$

Решение его формально можно записать в виде

$$x(t)-x(t_0)=\int_{t_0}^{t}\xi(t')dt'. \tag{4.18}$$

Стохастический интеграл также определяется с помощью предела в среднеквадратичном. Причем существуют две наиболее популярных формы: интеграл Ито (интеграл вводится аналогично обычному интегралу Римана – с помощью формулы левых прямоугольников) и интеграл Стратоновича (симметризованный). Вводится также и обобщенный стохастический интеграл, частными случаями которого являются оба упомянутых [55]. Но для (4.18), т.е. при $G=const$, обе формы совпадают.

Можно показать, что процесс (4.17) – винеровский. Его дисперсия линейно зависит от времени и равна $\sigma_x^2(t)=\sigma_x^2(t_0)+G^2\cdot(t-t_0)$.

Дисперсия его приращений на интервале Δt равна $G^2 \cdot \Delta t$. Это хорошо согласуется с результатами наблюдения за движением броуновских частиц: для них средний квадрат смещения тоже растет пропорционально времени.

Пример из геофизики. Уравнение (4.17) позволяет получить известный эмпирический закон Гутенберга – Рихтера повторяемости землетрясений в зависимости от их интенсивности [62]. Пусть x – величина механического напряжения (пропорциональная деформациям) на данном участке земной коры. Предположим, что со временем она может накапливаться из-за различных случайных факторов (толчков и т.п.) – белого шума. В среднем, ее квадрат (т.е. дисперсия) растет как $G^2 \cdot (t - t_0)$ (4.17), начиная с какого-то нулевого момента времени, когда напряжение мало. Землетрясения возникают, когда система набирает (аккумулирует) в течение определенного времени запас упругой энергии и затем сбрасывает ее тем или иным способом. Если сброс происходит по достижении некоторого фиксированного порога E, то время, необходимое, чтобы набрать такую энергию равно

$$\tau = \frac{E}{G^2}. \tag{4.19}$$

Далее, из уравнения (4.19) можно получить, что частота появления землетрясений с энергией не меньше E есть величина $\sim 1/\tau \sim G^2/E$, т.е. частота появления обратно пропорциональна энергии. К подобному виду сводится и закон Гутенберга – Рихтера при некоторых предположениях. Аналогичные законы объясняют появление цунами, оползней и т.п. событий, подробнее см. в [62].

Пример из молекулярной физики. Если предположить, что независимые толчки меняют скачком не смещение частицы x, как в уравнении (4.17), а ее скорость (т.е. белый шум представляет собой действующие на частицу силы), то получим СДУ второго порядка

$$d^2x(t)/dt^2 = \xi(t), \tag{4.20}$$

которое позволяет получить аналитически закон Ричардсона – Обухова, который гласит, что средний квадрат смещения (броуновских) частиц при определенных условиях растет со временем как $(t - t_0)^3$. Этот закон выполняется на поверхности океана (относительная диффузия) в некоторых интервалах масштабов [62].

Численное интегрирование СДУ. Выше мы рассмотрели примеры, допускающие аналитическое решение, но при нелинейных F или G приходится пользоваться численными методами, которые в случае СДУ

существенно отличаются от случая ОДУ. Пусть объект описывается уравнением (4.16) с $G = const$ (для простоты иллюстраций):

$$dx/dt = F(x(t)) + \xi(t). \tag{4.21}$$

При фиксированном начальном условии $x(t)$ СДУ определяют не единственную будущую траекторию системы, а целый ансамбль возможных траекторий. Однозначно определяются функцией F только условные плотности распределения вероятностей $p(x(t+\Delta t)|x(t))$. Аналитически получить формулы для условных распределений в случае нелинейной F нельзя, но можно это сделать численно путем получения ансамбля реализаций СДУ, имитируя шумовые воздействия на интервале от t до $t+\Delta t$ с помощью генератора псевдослучайных чисел и численно интегрируя СДУ по шагам.

Численное интегрирование СДУ представляет собой гораздо более сложную задачу, чем численное интегрирование ОДУ, еще и потому, что более сложным понятием является интеграл от случайного процесса $\xi(t)$. Наиболее простой подход – использование метода Эйлера с малым шагом интегрирования h. Разностная схема Эйлера здесь такова:

$$x(t+h) = F(x(t)) \cdot h + \varepsilon_0(t) \cdot \sqrt{h}, \tag{4.22}$$

где $\varepsilon_0(t), \varepsilon_0(t+h), \varepsilon_0(t+2h), ...$ – независимые нормально распределенные величины с нулевым средним и дисперсией G^2. Отметим характерный вид случайного слагаемого с зависимостью интенсивности от шага вида $\sqrt{h}$. Этого нет при интегрировании ОДУ, где вклад правой части уравнения в разностную формулу составляет величину порядка h при малых h. Это имеет место для СДУ из-за интегрирования белого шума $\xi(t)$: в разностную схему входит случайная добавка с дисперсией, пропорциональной времени. При очень малом шаге h доминирует случайное воздействие.

Далее при фиксированном шаге интегрирования h можно сгенерировать ансамбль реализаций шума $\varepsilon_0(t), \varepsilon_0(t+h), \varepsilon_0(t+2h), ...$ и, пользуясь формулой (4.22), рассчитать для каждой реализации значение $x(t+\Delta t)$ в конце интервала интегрирования. По полученному набору значений $x(t+\Delta t)$ можно построить гистограмму, которая и есть оценка условного вероятностного распределения $p(x(t+\Delta t)|x(t))$. При изменении шага h полученная оценка будет меняться и приближаться к истинному распределению только в пределе $h \to 0$, как и в случае ОДУ приближенное решение стремится к истинному при $h \to 0$. Практически же шаг нужно выбирать настолько малым, чтобы при его дальнейшем уменьшении полученное приближенное распределение уже мало менялось.

В качестве примера рассмотрим интегрирование стохастических уравнений осциллятора Ван дер Поля:

$$\begin{aligned} dx_1/dt &= x_2, \\ dx_2/dt &= \mu(1-x_1^2)x_2 - x_1 + \xi(t), \end{aligned} \tag{4.23}$$

где $\mu = 3$. Найдем условные распределения $p(x_1(t+\Delta t)|\mathbf{x}(t))$ для двух начальных условий $\mathbf{x}(t) = (-0.0935, -4.284)$ и $\mathbf{x}(t) = (1.021, -0.9375)$ (значения, взятые произвольно из временной реализации системы) при $\Delta t = 0.5$, что соответствует примерно 1/18 характерного периода [327]. Получим численные оценки при *h*, равных 0.1, 0.01, 0.001 и 0.0001 (рис.4.2). Оценка первого распределения стабилизируется при $h = 0.001$, а второго – при $h = 0.01$. В любом случае для хорошего приближения условного распределения шаг интегрирования должен быть достаточно мал – не более, примерно, одной тысячной от характерного периода. Для сходимости численного метода решения соответствующего ОДУ, т.е. уравнения (4.23) *без шума*, с высокой точностью хватило бы шага 0.01.

Возможно и использование методов Рунге – Кутты более высокого порядка точности, но формулы отличаются от случая ОДУ [132].

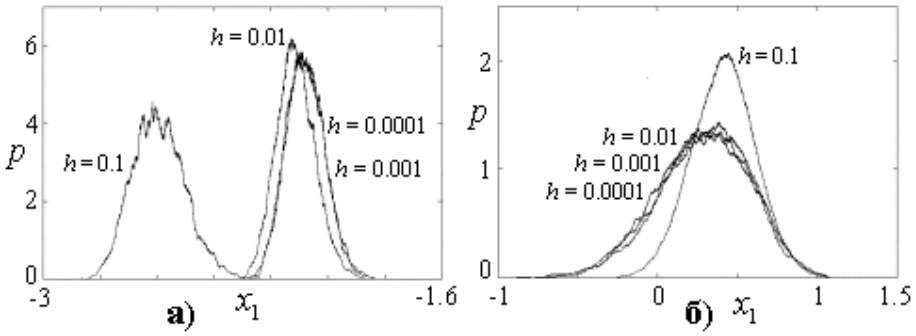

Рис.4.2. Плотности распределения $p(x_1(t+\Delta t)|\mathbf{x}(t))$ для разных шагов интегрирования (0.1, 0.01, 0.001, 0.0001). а) при $\mathbf{x}(t) = (-0.0935, -4.284)$, б) $\mathbf{x}(t) = (1.021, -0.9375)$. Слева – сходимость достигается только при шаге 0.001, а справа – уже при 0.01