第二章　随机微积分

实际问题的认识只是可能性的,这种论点是当代科学方法论的中心论点之一。

Nagel (1969, P.5)

1. 引言

本章介绍随机微积分的各种数学方法,它们对于分析不确定情形下的经济和金融现象并建立模型是有用的。这些方法包含 $It\hat{o}$ 引理、随机微分方程、随机稳定性和随机控制。

学习这些方法需要懂得各种数学概念和结果，例如，随机积分的概念，随机微分方程解的性质，随机稳定性的各种判定方法，等等。为了使读者习惯随机分析的理论基础，对这些概念和结果在本章中也进行了介绍。

2. 建立不确定性模型

在经济现象的建模、分析和预测中，研究人员越来越把重点放在随机方法上。人们希望这种方法能抓住经济现实的纷繁复杂性、测量误差和不确定性。自然会产生的问题是：在创立经济理论过程中所带来的复杂性、不确定性和未知性，如何才能结合到动态分析当中？这个问题的答案就在本节。在本节中我们首先讨论离散时间的情况，然后讨论连续时间的情况。材料来源于 Astrom(1970)。

2.1 离散时间

设离散时间指标集 T 是由正整数组成的集合 $\{0, 1, 2, 3, \cdots\}$，用 $x(t)$ 表示在时间 t，$t \in T$，的实状态变量。一个动态确定性系统可以用下列差分方程来描述：

$$x(t+1) = f(t, x(t)),\ t \in T \tag{2.1}$$

初始条件为 $x(0) = x_0$。不用说，这种确定性差分方程模型在经济理论中相当多。例如：三部门宏观经济模型：

$$\begin{aligned} Y_t &= C_t + I_t + G_t \\ C_t &= a_0 + a_1 Y_{t-1} \end{aligned}$$

对于给定的 I_t 和 G_t 的值，分别用 $\bar{I}_t$ 和 $\bar{G}_t$ 表示，这样，就得到一个差分方程：

$$Y_t = (a_0 + \bar{I}_t + \bar{G}_t) + a_1 Y_{t-1} \qquad (2.2)$$

方程（2.2）是方程（2.1）的特例，叫做线性自治差分方程。

像计量经济学家建立不确定模型一样，我们开始把确定性模型方程（2.1）变成随机模型回忆一下，在方程（2.2）中引入不确定因素的标准方式是添加一个随机变量U_t，通常假定U_t是一个相互独立的随机变量序列，且服从均值为 0、方差有限的正态分布。于是（2.2）变成：

$$Y_t = (a_0 + \bar{I}_t + \bar{G}_t) + a_1 Y_{t-1} + U_t \qquad (2.3)$$

方程（2.3）是一个线性随机差分方程的例子。

现在回到方程（2.1)，一种能把不确定性具体化的形式是假定$x(t+1)$是一个随机变量且有如下表达式：

$$x(t+1) = f(t, x(t)) + v(t, x(t)), \quad t \in T \qquad (2.4)$$

其中 f 是随机变量 $x(t+1)$ 关于 $x(t)$ 的条件均值，v 是一个均值为 0、方差有限（记为 $\sigma^2(t,x)$）的随机变量。我们假定随机变量v在给定$x(t)$下的条件分布与$x(s)$，$s<t$，相对独立。于是，方程（2.4）是一个随机差分方程，且在独立性的假设下，我们得到过程$\{x(t), t \in T\}$是一个马尔可夫过程。

其次，如果假定 v 在给定 $x(t)$ 下的条件分布是正态分布，那么随机变量

$$u_t = v(t)/\sigma(t,x)$$

服从均值为 0、方差为 1 的正态分布。注意到u_t与x无关，于是$\{u(t), t \in T\}$可以取成一个相互独立同服从均值为 0、方差为 1 的正态分布的随机变量序列。因此，方程（2.4）变成：

$$x(t+1) = f(t, x(t)) + \sigma(t, x(t))u(t), \quad t \in T \qquad (2.5)$$

方程（2.5）是一个随机差分方程。在许多书中，例如 Astrom(1970)和 Chow(1975),都有这方面的研究。

2.2 连续时间

考虑由微分方程描述的动态确定性模型：

$$\frac{dx}{dt} = f(t, x(t)), \ t \in T \qquad (2.6)$$

其中$T=[0,T]$或者$T=[0,\infty)$，初始条件$x(0)=x_0$。有许多经济模型描述成形如（2.6）

的例子。例如，在 Solow（1956）中的新古典增长模型：

$$\frac{dk}{dt} = sf(k(t)) - nk(t), \quad k(0) = k_0 \tag{2.7}$$

注意（2.7）是（2.6）的特例，因为它是自控的，或换句话说，它是时间相关的。

为了把不确定因素引入（2.6），我们按如下方式进行。考虑在时刻 t 和时刻 $t+\Delta t$ 时的模型；从方程

$$x(t+\Delta t) - x(t) = f(t, x(t))\Delta t + o(\Delta t) \tag{2.8}$$

中，通过两边除以 Δt 和并在 $\Delta t \to 0$ 取极限，得到确定性方程（2.6）。注意，（2.8）中的 $o(\Delta t)$ 意思是当 $\Delta t \to 0$ 时，$o(\Delta t)/\Delta t \to 0$。

现在，（2.8）是一个差分方程，我们可以重复离散情形的步骤。设 $\{v(t), t \in T\}$ 是一个独立增量过程，即如第 1 章第 7 节所示，如果 $H_i \in \Re$ 且 $0 \le t_0 \le t_1 \le \cdots \le t_k$，那么：

$$P[v(t_i) - v(t_{i-1}) \in H_i, i \le k] = \prod_{i \le k} P[v(t_i) - v(t_{i-1}) \in H_i \text{。}$$

这时，方程（2.8）变成：

$$x(t+\Delta t) - x(t) = f(t, x(t))\Delta t + v(t+\Delta t) - v(t) + o(\Delta t) \tag{2.9}$$

假定 v 的增量在给定 $x(t)$ 下的条件分布是正态分布且记作：

$$v(t+\Delta t) - v(t) = \sigma(t, x(t))[z(t+\Delta t) - z(t)] \tag{2.10}$$

其中 $\{z(t), t \in T\}$ 是一个均值为 0、方差系数为 1 的维纳（Wiener）过程。把（2.10）代入（2.9），得：

$$x(t+\Delta t) - x(t) = f(t, x(t))\Delta t + \sigma(t, x(t))[z(t+\Delta t) - z(t)] + o(\Delta t) \tag{2.11}$$

在上述方程中，$o(\Delta t)$ 是一个随机变量，通常意思是当 $\Delta t \to 0$ 时，$E|o(\Delta t)|^2/\Delta t \to 0$。注意，在（2.11）中，我们不能像（2.8）一样，两边除以 Δt 再取极限，这是因为，极限

$$\lim \frac{z(t+\Delta t) - z(t)}{\Delta t}$$

不存在。回忆一下第一章定理 7.2，维纳过程在一个概率为 0 的集合之外处处不可微。因此，代替用 Δt 除（2.11）再取通常意义的极限，我们只取另一意义下的极限，它将在下一节中解释。总之，（2.11）可以写成一个形式表达式：

$$dx = f(t, x)dt + \sigma(t, x)dz \tag{2.12}$$

它叫做 $It\hat{o}$ 随机微分方程，初始条件为 $x(0) = x_0$，$t \in T$。

方程（2.12）在应用数学中起重要作用。可参看书籍 Schuss(1980)、Soong(1973)以及 Tsokos 和 Padegett(1974)。方程(2.12)始自 1908 年的历史发展概要可参看 Wonham(1970)。方程（2.12）在金融和经济理论中也起重要作用，这一点下两节将会阐明。由于方程（2.12）在经济学和金融中的连续随机状态模型里的重要性，我们开始给出其形式表达式的意义。

3. 随机积分

考虑概率空间 (Ω, F, P) 上定义的一个随机过程 $x(t,\omega)$ 和一个维纳过程 $z(t,\omega)$，$\omega \in T$ 且 $t \in T$。方程（2.12）是方程

$$dx(t,\omega) = f(t, x(t,\omega))dt + \sigma(t, x(t,\omega))dz(t,\omega) \tag{3.1}$$

的简略记法。方程（3.1）也可以写成：

$$dx(t) = f(t, x(t))dt + \sigma(t, x(t))dz(t) \tag{3.2}$$

或者

$$dx_t = f(t, x_t)dt + \sigma(t, x_t)dz_t \tag{3.3}$$

其中 ω 在（3.2）和（3.3）中被隐藏了。

把方程（3.2）变换成积分方程，得到：

$$x(t) = x(0) + \int_0^t f(s, x(s))ds + \int_0^t \sigma(s, x(s))dz(s) \tag{3.4}$$

注意，$x(0)$ 是一个随机变量，因为 $x(0) = x(0,\omega)$，对一切 $\omega \in \Omega$。$x(0)$ 也可能退化成一个常数。一般来说，（3.4）式右边的第一个积分可理解成 Riemann 积分，由于 $dz(t)$ 不存在，所以第二个积分有一点问题。换句话说，尽管 $z(t)$ 连续，但却是一个无界变差函数，于是第二个积分不能理解成 Riemann-Stieljes 积分。因此，我们需要提出一个理论，使得第二个积分有意义。

随机积分由维纳在 1923 年首先引进，Ito(1944)将其发展。Ito 的部分原始工作最初由 Doob(1953)介绍，后来，俄国数学家 Gihman 和 Skorohod(1969)也介绍过。然而，直到最近，由于 Arnold(1974)、Astrom(1970)、Gihman 和 Skorohod(1972)、Bharucha-Reid(1972)、Ladde 和 Lakshmikantham(1980)等书的出版，随机积分的想法和一般随机微积分理论才被应用研究人员所接受。

在下面的讨论中，我们先给出直观分析结果，然后再详细说明随机积分理论。

3. 1 直观分析

把常微分方程（2.6）写成下列形式：

$$dx = f(t,x)dt;\ x(0) = x_0;\ t \in T = [0,T] \tag{3.5}$$

讨论其意义。如果对每一个 $s \in T$，有：

$$\frac{dy(s)}{ds} = f(s, y(s)) \tag{3.6}$$

那么，称 $y(s)$ 是（3.5）的解。相同地，积分（3.6）式的两边，如果对所有时刻 t，s，$s \le t$，有：

$$y(t) - y(s) = \int_s^t f(r, y(r))dr \tag{3.7}$$

那么，称 $y(s)$ 是（3.5）的解。

注意到（3.7）式的右边是一个积分，它可以通过把时间区间 $[s,t] \subset [0,T]$ 分成小子区间的方法来逼近。设 $\varepsilon > 0$，给定划分：

$$\begin{aligned} & s = t_0 < t_1 < t_2 < \ldots < t_n = t \\ & \max_{0 \le i \le n-1} \left| t_{i+1} - t_i \right| \le \varepsilon, [t_i, t_{i+1}) \subset [s,t] \end{aligned} \tag{3.8}$$

那么，我们有：

$$y(t) - y(s) = \sum_{i=0}^{n-1} f(t_i, y(t_i))(t_{i+1} - t_i) + 0(\varepsilon)\text{。}$$

其中当 $\varepsilon \to 0$ 时，$0(\varepsilon) \to 0$。这是在微积分中计算积分的标准逼近过程。现在，对于随机微积分，我们采取类似的方法。称随机过程 $y(t)$ 是（3.2）的解，如果对所有的 t，s，$s \le t$，有：

$$y(t) - y(s) = \int_s^t f(r, y(r))dr + \int_s^t \sigma(r, y(r))dz(r) \tag{3.9}$$

其中（3.9）式的右边定义为：

$$\begin{aligned} & \int_s^t f(r, y(r))dr + \int_s^t \sigma(r, y(r))dz(r) = \sum_{i=0}^{n-1} f(t_i, y(t_i))(t_{i+1} - t_i) \\ & + \sum_0^{n-1} \sigma(t_i, y(t_i))[z(t_{i+1}) - z(t_i)] + 0(\in) \end{aligned} \tag{3.10}$$

在（3.10）中，$0(\varepsilon)$ 是一个随机变量，且当 $\varepsilon \to 0$，$0(\varepsilon) \to 0$ 的意义为当 $\varepsilon \to 0$ 时，$E\left|0(\varepsilon)\right|^2 \to 0$，这里 E 表示无条件数学期望。但在说明（3.10）的合理性和证明对于形如（3.8）的划分，满足条件

$$y(t)-y(s)=\sum_{i=0}^{n-1} f(t_i,y(t_i))(t_{i+1}-t_i)$$
$$+\sum_{i=0}^{n-1}\sigma(t_i,y(t_i))[z(t_{i+1})-z(t_i)]+0(\varepsilon)$$

的随机过程 $y(t)$ 的存在性时，有一些数学上的技术问题。它们将在下一节较详细地处理。

让我们继续说明一个可能的困难。给定函数 $g(t,z)$ 和 $z(t)$，当 $\max|t_{i+1}-t_i|\to 0$ 时，两个极限值：

$$A=\lim\sum_i g(t_i,z(t_i))[z(t_{i+1})-z(t_i)]$$

和

$$B=\lim\sum_i g(t_{i+1},z(t_{i+1}))[z(t_{i+1})-z(t_i)]$$

在确定性情形下是相同的。但对于随机情形，就不是这样。顺便介绍一个简单例子。设 $g(t,z)=z(t)$，$z(t)$ 是一个具有单位方差的维纳过程，当 $\varepsilon\to 0$ 时，记：

$$A=\lim\sum_i z(t_i)[z(t_{i+1})-z(t_i)]$$

和

$$B=\lim\sum_i z(t_{i+1})[z(t_{i+1})-z(t_i)]$$

在确定性情形中，即 $z(t)$ 是确定性的而不是随机的，A 和 B 都是通常的 Riemann-Stieljes 积分。如果 $z(t)$ 可积，那么当 $\varepsilon\to 0$ 时，$B-A\to 0$。因此，在确定性情形中，不存在有关收敛性的歧义。而在随机情形中，这些极限的意义为：当 $\varepsilon\to 0$ 时，

$$E\left|A-\sum_i z(t_i)[z(t_{i+1})-z(t_i)]\right|^2\to 0$$
$$E\left|B-\sum_i z(t_{i+1})[z(t_{i+1})-z(t_i)]\right|^2\to 0 \qquad (3.11)$$

其中 E 表示数学期望。由（3.11）定义的极限叫做均方极限，它描述了除在第一章第 3 节中定义的三种收敛概念以外的一种收敛概念。对（3.11）中定义的 A 和 B，利用维纳过程有独立平稳增量 $z(t_{i+1})-z(t_i)$ 且其方差为 $t_{i+1}-t_i$，可证 $B-A=t-s\neq 0$。因此，我们得出随机积分的大小取决于我们采取 A 型、B 型或者它们的凸组合型作为极限表达式。$It\hat{o}$ 分析采用 A 型作为极限表达式，这种随机积分下面我们将讨论。另一种类型的随机积分参看本章最后的第 15 节。

3．2 严格的阐述

下面的分析内容是根据Arnold(1974,pp.57-75)、Bharucha-Reid(1972,pp.221-228)、Gihman和Skorohod(1972,pp.11-32)以及Ladde和Lakshmikantham(1980,pp.114-122)给出的，目的在于严格地定义$It\hat{o}$随机积分并介绍其基本性质。

设z_t，$t \ge 0$，是一个定义在某一概率空间(Ω, F, P)上的维纳过程，称$\sigma-$域族F_t，$t \ge 0$，关于z_t是非预期的，如果

（1）$F_{t_1} \subset F_{t_2}$，对于$0 \le t_1 \le t_2$；

（2）F_t包含由$z_s, 0 \le s \le t$,生成的$\sigma-$域,即$F_t \supset \sigma$（$z_s, 0 \le s \le t$）$\equiv Z_t$;

（3）F_t与增量$z_u - z_t, t \le u < \infty$,生成的$\sigma$（$z_u - z_t, t \le u < \infty$）相互独立。

注意，条件（2）意味着对于$t \ge 0$，z_t关于F_t可测。条件（3）意味着对于$h = u - t, t \le u < \infty$，过程$z(t+h) - z(h)$与$\sigma-$域$F_h$中的任一事件相互独立。特别地，条件（3）意味着$F_0$只能含有那些与整个过程$z_t$，$t \ge 0$，相互独立的事件。

作为一个例子，考虑$F_t = Z_t$，它是最小可能的非预期$\sigma-$域族。通常根据需要用其他与$\sigma(z_u - z_t, t \le u < \infty)$相互独立的事件扩大$Z_t$，例如描述初始条件的事件。在随机微分方程的情形里，我们通常取$F_t = \sigma(Z_t, c)$，其中c是一个与$\sigma(z_u - z_t)$相互独立的随机变量。

现在考虑$\sigma(t,\omega):[0,T] \times \Omega \to R$，假设它关于$(t,\omega)$可测。称它关于$\sigma-$域族$F_t$是非预期的，如果满足两个条件：

(1) 对所有$t \in [0,T]$，样本轨道$\sigma(t,\cdot)$是F_t-可测的；

(2) 积分

$$\int_0^T |\sigma(t,\omega)|^2 dt \qquad (3.12)$$

以概率1 （w.p.1）有限。

当函数$\sigma(t,\omega)$以概率1具有连续样本轨道时，最后的这个积分是通常的Remann积分。更一般地，(3.12)取作Lebesgue积分。上述定义可推广到矩阵值函数，这时，(3.12)中的范数定义为：

$$\left|\sigma(t,\omega)\right| = (\sum_i \sum_j \sigma_{ij}^2)^{1/2} = (tr\sigma \cdot \sigma')^{1/2},$$

其中tr表示迹，“ ′ ”表示转置。

一个特殊的非预期函数类是非预期阶梯函数类。非预期函数$\sigma(t,\omega)$叫做阶梯函数，如果存在区间$[0,T]$的一个划分，比方说：

$$0 = t_0 < t_1 < t_2 < \ldots < t_n = T \tag{3.13}$$

使得对于$t \in [t_i, t_{i+1}), i = 0,1,2,\ldots,n-1$，有$\sigma(t,\omega) = \sigma(t_i,\omega)$。对于这样的阶梯函数，现在我们来定义$It\hat{o}$随机积分。

设(Ω, F, P)是一个概率空间，$\sigma(t,\omega):[0,T]\times\Omega \to R$是一个具有划分形式（3.13）的非预期阶梯函数，$z(t,\omega):[0,T]\times\Omega \to R$是一个维纳过程。$\sigma$关于$z$在区间$[0,T]$上的随机积分，用$I(\sigma)$表示，是一个实值随机变量，定义为：

$$\begin{aligned} I(\sigma) &= I(\sigma(\omega)) = \int_0^T \sigma(t,\omega)dz(t,\omega) \\ &= \int_0^T \sigma dz = \sum_{i=1}^n \sigma(t_{i-1},\omega)[z(t_i,\omega) - z(t_{i-1},\omega)] \\ &= \sum_1^n \sigma(t_{i-1})[z(t_i) - z(t_{i-1})] \\ &= \sum_0^{n-1} \sigma(t_i)[z(t_{i+1}) - z(t_i)] \end{aligned} \tag{3.14}$$

$\omega \in \Omega$的出现目的在于强调$It\hat{o}$积分是一个随机变量，为了记号方便起见，有时把ω省略掉。需要着重指出的是，在阶梯函数的定义中，进而在随机积分的定义中，采用了子区间左端点的值。

其次，考虑任意的非预期函数$\sigma(t,\omega)$。假设$\{\sigma_n\}$是一个非预期阶梯函数序列，自然要问，是否能用$\{\sigma_n\}$在某种意义下逼近函数σ。下面的引理给出肯定的回答。

引理 3.1 设$\sigma(t,\omega)$是一个非预期函数。那么存在一个非预期阶梯函数序列$\{\sigma_n\}$，使得

$$\int_0^T \left|\sigma_n(t) - \sigma(t)\right|^2 dt \to 0,\ w.p.1 \tag{3.15}$$

这个引理的证明参看 Arnold(1974,pp.67)。引理 3.1 的结论很强，它蕴含更弱的结果，即

$$P\left[\omega:\int_0^T\left|\sigma_n(t,\omega)-\sigma(t,\omega)\right|^2dt>\varepsilon\right]\to 0 \tag{3.16}$$

这是正如第一章第 3 节所指出的那样，以概率 1 收敛蕴涵依概率收敛。$It\hat{o}$ 随机积分的一般存在性将借助于下述过程建立。具体来说，就是证明：对于满足（3.16）的非预期阶梯函数序列$\{\sigma_n\}$，积分序列$\int_0^T\sigma_n dz$必依概率收敛到$\int_0^T\sigma dz$，即：

$$\int_0^T\sigma_n dz\xrightarrow{P}\int_0^T\sigma dz \tag{3.17}$$

假设$\{\sigma_n\}$是一个如（3.16）所示依概率逼近非预期函数σ的非预期阶梯函数序列。根据

$$\int_0^T\left|\sigma_n-\sigma_m\right|^2dt\le 2\int_0^T\left|\sigma-\sigma_n\right|^2dt+2\int_0^T\left|\sigma-\sigma_m\right|^2dt$$

和$\{\sigma_n\}$满足（3.16），我们得到当$n,m\to\infty$时：

$$\int_0^T\left|\sigma_n-\sigma_m\right|^2dt\xrightarrow{P}0 \tag{3.18}$$

如果我们证明了$\int_0^T\sigma_n dz$是一个依概率收敛的 cauchy 列，那么我们将得出：存在一个随机变量，使$\int_0^T\sigma_n dz$依概率收敛于它。在证明$\int_0^T\sigma_n dz$是随机 Cauchy 序列的过程中需要一个方法性引理。

引理 3.2　　设$\sigma(t,\omega)$是一个非预期阶梯函数。那么对所有$\varepsilon>0$和$\delta>0$，有：

$$P\left[\left|\int_0^T\sigma(t)dz(t)\right|>\delta\right]\le\frac{\varepsilon}{\delta^2}+P\left[\int_0^T\left|\sigma(t)\right|^2dt>\varepsilon\right]$$

该引理的详细证明，参看 Arnold(1974,pp.68-69)

对$\sigma_n-\sigma_m$应用这个引理，得：

$$\begin{aligned}&\limsup_{n,m\to\infty}P\left[\left|\int_0^T\sigma_n(t)dz(t)-\int_0^T\sigma_m(t)dz(t)\right|>\delta\right]\\&\le\frac{\varepsilon}{\delta^2}+\limsup_{n,m\to\infty}P\left[\int_0^T\left|\sigma_n(t)-\sigma_m(t)\right|^2dt>\varepsilon\right]\end{aligned} \tag{3.19}$$

从（3.18）和（3.19），我们得出结论，$\int_0^T \sigma_n dz$ 是一个随机 Cauchy 序列，即：

$$P\left[\left|\int_0^T \sigma_n(t)dz(t)-\int_0^T \sigma_m(t)dz(t)\right|>\varepsilon\right]\to 0$$

当 $n,m\to\infty$ 时。因此，（3.17）成立，极限 $\int_0^T \sigma dz$ 以概率 1 唯一且不依赖于满足（3.16）的非预期阶梯函数序列 $\{\sigma_n\}$ 的选取。综合这些结果，有：

引理 3.3 设 $\sigma(t,\omega)$ 是一个非预期函数，且 $\{\sigma_n\}$ 是 σ 的在（3.16）意义下的逼近非预期阶梯函数序列。对于每一个 n，假定 $I(\sigma_n)$ 由（3.14）定义，那么，存在一个随机变量 $\int_0^T \sigma(t,\omega)dz(t,\omega)$ 使得当 $n\to\infty$ 时有：

$$P\left[\omega:\left|\int_0^T \sigma_n(t,\omega)dz(t,\omega)-\int_0^T \sigma(t,\omega)dz(t,\omega)\right|>\varepsilon\right]\to 0 \qquad (3.20)$$

上述分析结果导致 $It\hat{o}$ 随机积分的一种定义。如（3.20）所示，它是一种基于依概率收敛的定义。基于不同收敛概念的其他定义也能给出。我们将简短地给出基于均方收敛的随机积分的定义。关于随机积分主题的详细讨论，读者可参看进一步的资料。

定义　设 (Ω,F,P) 是一个概率空间。$z(t,\omega)$ 是一个维纳过程，$\sigma(t,\omega)$ 是一个非预期函数，两者都定义在 $[0,T]\times\Omega$ 上。σ 关于 z 在区间 $[0,T]$ 上的 $It\hat{o}$ 随机积分：

$$\int_0^T \sigma(t,\omega)dz(t,\omega)$$

用 $I(\sigma)$ 表示，是一个随机变量，以概率 1 唯一且定义为随机 Cauchy 序列 $\int_0^T \sigma_n(t,\omega)dz(t,\omega)$ 依概率收敛的极限，即如引理 3.3 所示，

$$\int_0^T \sigma_n(t,\omega)dz(t,\omega)\xrightarrow{P}\int_0^T \sigma(t,\omega)dz(t,\omega)\equiv \mathrm{I}(\sigma)$$

这里 $\{\sigma_n\}$ 是 σ 的在依概率收敛意义下的逼近非预期阶梯函数序列，即：

$$\int_o^T \left|\sigma(t,\omega)-\sigma_n(t,\omega)\right|^2 dt\xrightarrow{P}0$$

再重复一遍，$I(\sigma)$ 以概率 1 唯一，且不依赖于序列 $\{\sigma_n\}$ 的选取。唯一性的证明见 Friedman(1975,pp.66)。

有一个可替换的定义。正如在本节前面所提到，$It\hat{o}$ 随机积分的一个可替换的定义是用利用均方收敛代替依概率收敛。因为均方收敛蕴涵依概率收敛，所以即将给出的定义更一般，上述给出的定义可作为其特殊情况。定义基于下面的引理。

引理 3.4　设 $\sigma(t,\omega)$ 是非预期函数，且：

$$\int_0^T E\left|\sigma(t,\omega)\right|^2 dt < \infty \tag{3.21}$$

则存在一个非预期阶梯函数序列 $\{\sigma_n\}$，对每一个 n 满足

$$\int_0^T E\left|\sigma_n(t,\omega)\right|^2 dt < \infty$$

且在均方极限下逼近 σ，即当 $n\to\infty$ 时，有：

$$\int_0^T E\left|\sigma_n(t)-\sigma(t)\right|^2 dt \to 0$$

进而存在一个随机变量 $\int_0^T \sigma(t)dz(t)$，使得当 $n\to\infty$ 时，有：

$$E\left|\int_0^T \sigma_n(t)dz(t)-\int_0^T \sigma(t)dz(t)\right|^2 \to 0 \tag{3.22}$$

引理的证明参看 Arnold(1974,pp.71-73)与 Gihman 和 Skorohod(1972,pp.11-15)。注意，为了得到均方收敛，我们必须有比引理 3.1 更多的假设（3.21）。基于引理 3.4 的新定义类似于早先的定义，唯一的不同之处是收敛概念。特别地，我们把 $\int_0^T \sigma(t,\omega)dz(t,\omega)$ 定义为 Cauchy 序列 $\int_0^T \sigma_n(t,\omega)dz(t,\omega)$ 的均方极限，如（3.22）所示。这样的极限唯一存在且不依赖于 $\{\sigma_n\}$ 的选取，这是因为 $\int_0^T \sigma_n dz$ 是满足条件：

$$E\left|\int_0^T \sigma_n(t)dz(t)-\int_0^T \sigma_m(t)dz(t)\right|^2 \to 0 \text{ 当} m,n\to\infty\text{时}$$

的 Cauchy 序列。

现在结束有关随机积分概念的讨论。在叙述随机积分性质之前，我们给对这个领域有兴

趣的读者提供额外的参考资料。它们是 Meyer(19670,Mckean(1969)、Metivier 和 Pellaumail(1980)以及 Friedman(1975)

给出了$\int \sigma dz$的意义之后，我们叙述它的一些性质。

3.3 Itô 随机积分的性质

Itô 随机积分有许多有用的性质。在下列定理中，我们概述了一些。

定理 3.1 设(Ω, F, P)是一个概率空间，$z(t,\omega):[0,T]\times\Omega \to R$是一个维纳过程。

（1）如果$\sigma_1(t,\omega)$和$\sigma_2(t,\omega)$是定义在$[0,T]\times\Omega$上的实值非预期函数，$a_1 \in R$，$a_2 \in R$，那么：

$$\int_0^T (a_1\sigma_1 + a_2\sigma_2)dz = a_1\int_0^T \sigma_1 dz + a_2\int_0^T \sigma_2 dz$$

（2）假设σ_1和σ_2同(1)一样，$a_1(\omega)$和$a_2(\omega)$是实值随机变量且$a_1\sigma_1 + a_2\sigma_2$是$[0,T]$上的非预期函数。那么

$$\int_0^T (a_1\sigma_1 + a_2\sigma_2)dz = a_1\int_0^T \sigma_1 dz + a_2\int_0^T \sigma_2 dz$$

（3）设$[s,u]$是$[0,T]$的子集，$\chi_{[s,u]}$表示区间$[s,u]$的特征函数。那么：

$$\int_0^T \chi_{[s,u]} dz = z(u) - z(s)$$

（4）假设$\sigma(t,\omega)$是$[0,T]$上的非预期函数，且$\int_0^T E|\sigma(s)|^2 ds < \infty$，那么$E\int_0^T \sigma(t)dz(t) = 0$，且：

$$E\left|\int_0^T \sigma(t)dz(t)\right|^2 = \int_0^T E|\sigma(t)|^2 dt$$

这些性质的证明参看 Gihmna 和 Skorohod(1972,pp.11-13)，注意，性质（1）-（3）与通常的 Riemann-Steiljes 积分的有关性质类似。

对于任意的（而不是阶梯函数）依概率逼近σ的非预期函数序列$\{\sigma_n\}$，下面建立其依概率收敛性。

定理 3.2 设(Ω, F, P)是一个概率空间，$z(t,\omega):[0,T]\times\Omega\to R$为维纳过程，$\{\sigma_n\}$是定义在$[0,T]\times\Omega$上的任意实值非预期函数，且$\sigma(t,\omega):[0,T]\times\Omega\to R$为非预期函数。如果，当$n\to\infty$时：

$$\int_0^T |\sigma_n(t)-\sigma(t)|^2 ds \xrightarrow{P} 0 \tag{3.23}$$

那么：

$$\int_0^T \sigma_n dz \xrightarrow{P} \int_0^T \sigma dz \tag{3.24}$$

其中（3.23）和（3.24）都是依概率收敛。证明见 Arnold(1974, pp. 74)。

我们要叙述的最后一个定理是关于$It\hat{o}$随机积分作为积分上限函数的问题。假设$\chi_{[0,t]}$是区间$[0,t]\subset[0,T]$的特征函数。对于每一个t $(0\le t\le T,\text{其中}T\text{任意大})$，$\sigma$是$[0,t]$上的非预期函数。定义过程：

$$x(t)=x(t,\omega)=\int_0^t \sigma(s,\omega)dz(s,\omega)=\int_0^T \sigma(s)\chi_{[0,t]}dz(s) \tag{3.25}$$

需要指出，$x(t)$是一个实值随机过程，对于$t\in[0,T]$，（以随机等价）唯一地定义，且$x(0)=0 \quad w.p.1$。下面的定理叙述了$x(t)$的一些性质。这里，$x(t)$假定为一个可分表示（见第一章第 7 节）。

定理 3.3 设(Ω, F, P)是一个概率区间，$z(t,\omega):[0,T]\times\Omega\to R$为维纳过程，$\sigma(t,\omega):[0,T]\times\Omega\to R$是非预期函数，且按（3.25）定义$x(t)$。那么：

（1）$x(t,\omega)$是F_t-可测的，因此它是非预期的。

（2）如果$\int_0^t E|\sigma(s,\omega)|^2 ds<\infty$，对所有$t\in[0,T]$，那么$(x_t, F_t)$是实值鞅，同时，

$$Ex_t=0\,,\ E|x_t|^2=\int_0^t E|\sigma(s,\omega)|^2 ds \tag{3.26}$$

（3）$x(t,\omega)$以概率 1 有连续样本轨道。

证明见 Arnold(1974, pp. 81-84)。既然(x_t, F_t)是鞅，那么$(x_t - x_r, F_t)$，$0\le r\le t\le T$也是鞅，$(|x_t - x_r|, F_t)$是下鞅。因此，我们可运用各种鞅和下鞅的定理，得到关于(x_t, F_t)和

$(|x_t - x_r|, F_t)$ 的有用结果。

至此，我们结束有关随机积分的讨论，要了解该主题的更详细的展开，读者可参考 Mckean(1969)或者更新一点的书 Metivier 和 Pellaumail(1980)。

读者也应该被告知，一种叫做 Strasbourg 方法的新的更高级随机积分理论已经建立起来。这里，原始的参考资料是 Meyer(1976)。Strasbourg 方法在经济分析中的迷人的应用可参考 Harrison 和 Pliska(1981)。它们利用 Strasbourg 方法发展了具有连续交易的无磨擦证券市场理论。

4. $It\hat{o}$ 引理

现在，我们已做好准备，先叙述随机微分的定义，然后叙述和证明 $It\hat{o}$ 引理，它是计算复合随机函数的随机微分的基本定理。

设 (Ω, F, P) 是一个概率空间，随机过程 $x(t,\omega):[0,T]\times\Omega\to R$，对于每一个 $t\in[0,T]$，关于 F_t- 可测，$z(t,\omega):[0,T]\times\Omega\to R$ 为维纳过程。假定 $\sigma(t,\omega):[0,T]\times\Omega\to R$ 是 $[0,T]$ 上的非预期函数，同时，$f(t,\omega):[0,T]\times\Omega\to R$，对每一个 $t\in[0,T]$，关于 F_t- 可测且 $\int_0^T |f(t,\omega)|dt<\infty\ w.p.1$。设 $0\le r\le s\le T$，如果：

$$x(s)-x(r)=\int_r^s f(t,\omega)dt+\int_r^s \sigma(t,\omega)dz(t,\omega)$$

那么随机过程 $x(t)$ 的随机微分定义为 $f(t)dt+\sigma(t)dz(t)$，并用 $dx(t)$ 表示，即：

$$dx(t)=f(t)dt+\sigma(t)dz(t)$$

下面，我们陈述和证明 $It\hat{o}$ 引理，它最初出现在 $It\hat{o}$ (1951a)和后来的 $It\hat{o}$ (1961)。我们取材于 Arnold(1974,pp.96-99)、Gihman 和 Skorohod(1969,99.387-389)，以及 Ladde 和 Lakshmikantham(1980,pp.122-126)。

引理 4.1（$It\hat{o}$） 设 $u(t,x):[0,T]\times R\to R$ 是一个连续非随机函数，且具有连续的偏导数 u_t, u_x 和 u_{xx}，假定随机过程 $x(t,\omega):[0,T]\times\Omega\to R$ 的随机微分为：

$$dx(t)=f(t)dt+\sigma(t)dz(t)\text{。}$$

令 $y(t)=u(t,x(t))$。那么，过程 $y(t)$ 在 $[0,T]$ 上也有随机微分，且：

$$dy(t) = [u_t(t, x(t)) + u_x(t, x(t)) f(t) + \frac{1}{2} u_{xx}(t, x(t)) \sigma^2(t)]dt + u_x(t, x(t)) \sigma(t) dz(t)$$。

证明： 只需就 f 和 σ 为阶梯函数证明引理就足够了，因为一般情形可通过取极限而得到。

同时，只需在 f 和 σ 作为 t 的函数为常数，即 $f(t,\omega) = f(\omega)$ 和 $\sigma(t,\omega) = \sigma(\omega)$ 的子区间上给出引理的证明。考虑一个划分：

$$0 < t_1 < t_2 < \ldots < t_n = t \le T$$

因为 $y(t) = u(t, x(t)),\ y(0) = u(0, x(0))$，所以有：

$$y(t) - y(0) = \sum_{k=1}^{n} [u(t_k, x(t_k)) - u(t_{k-1}, x(t_{k-1}))] \tag{4.1}$$

利用 Taylor 定理，得：

$$\begin{aligned} u(t_k, x(t_k)) - u(t_{k-1}, x(t_{k-1})) = {} & u(t_{k-1} + \theta(t_k - t_{k-1}), x(t_{k-1}))(t_k - t_{k-1}) \\ & + u_x(t_{k-1}, x(t_{k-1}))(x(t_k) - x(t_{k-1})) \\ & + \frac{1}{2} u_{xx}(t_{k-1}, x(t_{k-1}) + \bar{\theta}_k [x(t_k) - \\ & - x(t_{k-1})])(x(t_k) - x(t_{k-1}))^2 \end{aligned} \tag{4.2}$$

其中 $0 < \theta_k < 1$，$0 < \bar{\theta}_k < 1$。

首先证明，当 $\delta_n = \max(t_k - t_{k-1}) \to 0$ 时，$y(t) - y(0)$ 的极限不变。从 $x(t)$，u_t 和 u_{xx} 的连续性假设和 $x(t)$ 的有界性可知，存在随机变量 α_n, β_n 满足

$$\left| u_t(t_{k-1} + \theta_k(t_k - t_{k-1}), x(t_{k-1})) - u_t(t_{k-1}, x(t_{k-1})) \right| \le \alpha_n$$

$$\left| u_{xx}(t_{k-1}, x(t_{k-1}) + \bar{\theta}_k [x(t_k) - x(t_{k-1})]) - u_{xx}(t_{k-1}, x(t_{k-1})) \right| \le \beta_n$$

且当 $\delta_n \to 0$ 时，α_n, β_n 以概率 1 趋于 0 。

在上面的第一个不等式中求和，得到：

$$\begin{gathered} \left| \sum u_t(t_{k-1} + \theta_k(t_k - t_{k-1}), x(t_{k-1}))(t_k - t_{k-1}) - \right. \\ \left. - \sum u_t(t_{k-1}, x(t_{k-1}))(t_k - t_{k-1}) \right| \le \sum \alpha_n (t_k - t_{k-1}) \end{gathered}$$。

对于第二个不等式，我们注意到：

$$\sum [x(t_k) - x(t_{k-1})]^2 \le 2 f^2 \sum (t_k - t_{k-1})^2 + 2\sigma^2 \sum [z(t_k) - z(t_{k-1})]^2$$

和 $E \sum [z(t_k) - z(t_{k-1})]^2 = t - 0 = t$

因此，利用这两个结果，在上面的第二个不等式中求和，得到：

$$
\begin{aligned}
&|\sum u_{xx}(t_{k-1},x(t_{k-1})+\bar{\theta}_k[x(t_k)-x(t_{k-1})])[x(t_k)-x(t_{k-1})]^2 \\
&\qquad -\sum u_{xx}(t_{k-1},x(t_{k-1}))[x(t_k)-x(t_{k-1})]^2 | \\
&\qquad \le \beta_n \sum [x(t_k)-x(t_{k-1})]^2 \xrightarrow{P} 0
\end{aligned}
$$

当$\delta_n \to 0$时，因此，当$\delta_n \to 0$时，（4.1）中$y(t)-y(0)$的极限不变。

其次，把（4.2）代入（4.1），如果能证明：

$$
\begin{aligned}
&\sum \{u_t(t_{k-1},x(t_{k-1}))(t_k-t_{k-1})+u_x(t_{k-1},x(t_{k-1}))[x(t_k)-x(t_{k-1})] \\
&\quad +\frac{1}{2}u_{xx}(t_{k-1},x(t_{k-1}))[x(t_k)-x(t_{k-1})]^2\}\xrightarrow{P} \\
&\int_0^t [u_t(s,x(s))+u_x(s,x(s))f+\frac{1}{2}u_{xx}(s,x(s))\sigma^2]ds \\
&+\int_0^t u_x(s,x(s))\sigma dz(s)
\end{aligned}
\tag{4.3}
$$

那么，定理的证明将完成。

注意，根据连续性假设，当$\delta_n \to 0$时，有：

$$
\sum u_t(t_{k-1},x(t_{k-1}))(t_k-t_{k-1}) \to \int_0^t u_t(s,x(s))ds \qquad w.p.1 \tag{4.4}
$$

和

$$
\begin{aligned}
&\sum u_x(t_{k-1},x(t_{k-1}))[x(t_k)-x(t_{k-1})]\xrightarrow{P} \\
&\int_0^t u_x(s,x(s)fds+\int_0^t u_x(s,x(s)\sigma dz(s)
\end{aligned}
\tag{4.5}
$$

因此，在（4.3）中，还需讨论：

$$
\begin{aligned}
&\sum u_{xx}(t_{k-1},x(t_{k-1}))[x(t_k)-x(t_{k-1})]^2 \\
&= f^2\sum u_{xx}(t_{k-1},x(t_{k-1}))(t_k-t_{k-1})^2 \\
&\quad +2f\sigma\sum u_{xx}(t_{k-1},x(t_{k-1}))(t_k-t_{k-1})[z(t_k)-z(t_{k-1})] \\
&\quad +\sigma^2\sum u_{xx}(t_{k-1},x(t_{k-1}))[z(t_k)-z(t_{k-1})]^2
\end{aligned}
\tag{4.6}
$$

根据u_{xx}和$z(t)$的连续性，可知(4.6)中的前面两项以概率 1 收敛到 0，故只需证明当$\delta_n \to 0$时，有：

$$
\sum u_{xx}(t_{k-1},x(t_{k-1}))[z(t_k)-z(t_{k-1})]^2 \xrightarrow{P} \int_0^t u_{xx}(s,x(s))ds \tag{4.7}
$$

为了证明（4.7），首先指出当$\delta_n \to 0$时，有：

$\sum u_{xx}(t_{k-1}, x(t_{k-1}))(t_k - t_{k-1}) \to \int_0^t u_{xx}(s, x(s))ds, \text{w.p.1}$ 。于是，只需证明当$\delta_n \to 0$时，有：

$$\sum u_{xx}(t_{k-1}, x(t_{k-1}))([z(t_k) - z(t_{k-1})]^2 - (t_k - t_{k-1})) \xrightarrow{P} 0 \quad （4.8）$$

（4.7）就会得证。为方便起见，我们用S_n表示（4.8）式的左边。于是（4.8）可写成当$\delta_n \to 0$时，

$$S_n \xrightarrow{P} 0 \quad （4.9）$$

为证明（4.9），我们使用截去法。
定义

$$I_k^N(\omega) = \begin{cases} 1 & 如果 |x(t_i)| \le N, 对所有 i \le k, \\ 0 & 否则, \end{cases}$$

再令：

$$e_k = [z(t_k) - z(t_{k-1})]^2 - (t_k - t_{k-1})$$

最后：

$$S_n^N = \sum_1^n u_{xx}(t_{k-1}, x(t_{k-1})) I_{k-1}^N e_k$$

因为e_k之间相互独立，e_k与$u_{xx}(t_{k-1}, x(t_{k-1})) I_{k-1}^N$相互独立，所以$ES_n^N = 0$且$E(S_n^N)^2 \to 0$，当$\delta_n \to 0$时。进一步，蕴含着对每一个$N$，当$\delta_n \to 0$时，有$S_n^N \xrightarrow{P} 0$，同时，截去的误差为：

$$P[S_n \ne S_n^N] = P[\max_s |x(s)| > N]$$

由于$x(s)$以概率 1 有限，所以只要选取充分大的N，可使误差任意小。最后，得出结论：

$$P[|S_n| > \varepsilon] \le P[|S_n^N| > \varepsilon] + P[S_n \ne S_n^N] \quad （4.10）$$

可以任意小。（4.9）式得证。于是，对于特殊情况，定理证毕。

最后，转到一般情况。可以选取阶梯函数序列$\{f_n\}$和$\{\sigma_n\}$，使以概率 1 地有：

$$\int_0^t |f_n(s) - f(s)| ds \to 0$$

$$\int_0^t \left|\sigma_n(s)-\sigma(s)\right|^2 ds \to 0$$

且随机过程序列：

$$x_n(t)=x_n(0)+\int_0^t f_n(s)ds+\int_0^t \sigma_n(t)dz(t)$$

以概率 1 一致收敛到 $x(t)$，那么随机过程序列 $y_n(t)=u(t,x_n(t))$ 也以概率 1 一致收敛到 $y(t)$。在下式：

$$\begin{aligned}y_n(t)-y_n(0)=&\int_0^t [u_t(s,x_n(s))+u_x(s,x_n(s))f_n(s)\\&+\frac{1}{2}u_{xx}(s,x_n(s))\sigma_n^2(s)]ds\\&+\int_0^t u_x(s,x_n(s))\sigma_n(s)dz(s)\end{aligned}$$

中取极限 $n\to\infty$，就得到 $It\hat{o}$ 引理。

在转到使用 $It\hat{o}$ 引理的例子之前，我们介绍一下 $It\hat{o}$ 引理的直接推广。

引理 4.2（一般 $It\hat{o}$ 引理）设 $u(t,x):[0,T]\times R^d\to R^k$ 是一个连续的非随机函数且有连续的偏导数 u_t，u_{x_i}，$u_{x_ix_j}$，其中：

$$u_t=\frac{\partial}{\partial t}u(t,x),$$

$$u_{x_i}=\frac{\partial}{\partial x_i}u(t,x),\quad i=1,2,...,d$$

$$u_{x_ix_j}=\frac{\partial}{\partial x_i\partial x_j}u(t,x),\quad i,j\le d$$

假定随机过程 $x(t)=x(t,\omega):[0,T]\times\Omega\to R^d$ 的随机微分为：

$$dx(t)=f(t)dt+\sigma(t)dz(t)$$

其中 $f(t)=f(t,\omega):[0,T]\times\Omega\to R^d$ 关于 (t,ω) 可测，即二元可测，且：

$$\sigma(t)=\sigma(t,\omega):[0,T]\times\Omega\to R^d\times R^m$$

这里的 σ 是一 $(d\times m)$ 矩阵值函数，在 $[0,T]$ 上非预期，最后

$z(t)=z(t,\omega):[0,T]\times\Omega\to R^m$ 为 m 维维纳过程。令 $y(t)=u(t,x(t))$，那么，过程 $y(t)$ 在 $[0,T]$ 上有随机微分，且：

$$\begin{aligned} dy(t) &= [u_t(t,x(t))+u_x(t,x(t))f(t) \\ &+\frac{1}{2}\sum_i\sum_j u_{x_ix_j}(t,x(t))[\sigma(t)\sigma'(t)]_{ij}]dt \\ &+u_x(t,x(t))\sigma(t)dz(t) \end{aligned} \tag{4.11}$$

至此，一般 $It\hat{o}$ 引理叙述完毕。注意，双重求和也可记成：

$$\sum_i\sum_j u_{x_ix_j}(\sigma\sigma')_{ij}=tr(u_{xx}\sigma\sigma')=tr(\sigma\sigma' u_{xx})$$

其中 $u_{xx}=u_{x_ix_j}$ 是一个 $(d\times d)$ 矩阵。于是 $dy(t)$ 的另一种表达式为：

$$dy(t)=u_t dt+u_x dx(t)+\frac{1}{2}tr(\sigma\sigma' u_{xx})dt$$

下面，通过强调三点，扼要地重述一下这一节内容。

第一，$It\hat{o}$ 引理是一个有用的结果，因为它使我们能够计算那种以具有随机微分的随机过程作为自变量的任意函数的随机微分。在这一点上，$It\hat{o}$ 公式与普通微积分的连锁规则一样有用。

第二，已知 $x(t)$ 是一个关于维纳过程 $z(t)$ 的 $It\hat{o}$ 随机过程，令 $y(t)=u(t,x(t))$ 为一个新过程，那么 $It\hat{o}$ 公式给出 $y(t)$ 的随机微分，且 $dy(t)$ 由同一个维纳过程给出。

第三， 审查 $It\hat{o}$ 引理的证明，可见它由高等微积分中的 Taylor 定理的应用和为了建立合适的积分收敛性所需的几个概率命题组成。因此，读者可以运用 Taylor 定理得到 $It\hat{o}$ 公式，不必记住特殊的结果。下面以三种情形来说明这一点。

情形 1：它对应一般 $It\hat{o}$ 公式。已知：

$$f(t,\omega):[0,T]\times\Omega\to R^d$$

$$\sigma(t,\omega):[0,T]\times\Omega\to R^d\times R^m$$

$$z(t,\omega):[0,T]\times\Omega\to R^m$$

它们满足 $It\hat{o}$ 引理的假设。过程 $x(t)$ 有随机微分：

$$dx_i(t)=f_i(t)dt+\sigma_{ir}(t)dz_r(t)$$

其中 $i=1,2,...,d$ ，且 $r=1,2,...,m$ 。我们定义：

$$y(t)=u(t,x(t)):[0,T]\times R^d\to R^k$$

注意，三个过程值域的相容性：

$$\begin{array}{cccc} dx(t) & = f(t)dt & + \quad \sigma(t) & dz(t) \\ (R^d \times R^1) & (R^d \times R^1) & (R^d \times R^m) & (R^m \times R^1) \end{array}$$

现在运用 Taylor 定理，作必要的代换，重新组织各项，得（注意，t 被省去，求和从 1 到 d ）：

$$\begin{aligned} dy(t) &= u_t dt + \sum u_{x_i} dx_i + \frac{1}{2}\sum\sum u_{x_i x_j} dx_i dx_j \\ &= u_t dt + \sum u_{x_i} (f_i dt + \sigma_{ir} dz_r) \\ &\quad + \frac{1}{2}\sum\sum u_{x_i x_j} (f_i dt + \sigma_{ir} dz_r)(f_j dt + \sigma_{jr} dz_r) \\ &= u_t dt + \sum u_{x_i} f_i dt + \sum u_{x_i} \sigma_{ir} dz_r + \frac{1}{2}\sum\sum u_{x_i x_j} [\sigma_{ir}\sigma'_{jr}] dz_r dz'_r \\ &= u_t dt + u_x f dt + \frac{1}{2}\sum\sum u_{x_i x_j} [\sigma\sigma']_{ij} dt + u_x \sigma dz \\ &= (u_t + u_x f + \frac{1}{2}\sum\sum u_{x_i x_j} [\sigma\sigma']_{ij}) dt + u_x \sigma dz \text{。} \end{aligned}$$

注，推导过程中用到下列乘法规则：

$$\begin{aligned} & dt \times dt = (dt)^2 = 0 \\ & dt \times dz_r = 0 \\ & dz_r \times dz_q = 0 \qquad \text{当} r \neq q \text{时} \\ & dz_r \times dz_q = dt \qquad \text{当} r = q \text{时} \end{aligned}$$

情形 2：　这里，我们部分地特殊化情形 1 的结果，假设：

$$f(t,\omega):[0,T]\times\Omega \to R^d \text{，}$$

$$\sigma(t,\omega):[0,T]\times\Omega \to R^d \text{，}$$

$$z(t,\omega):[0,T]\times\Omega \to R^1 \text{，}$$

同时，设过程 $x(t)$ 有随机微分：

$$dx_i(t) = f_i(t)dt + \sigma_i(t)dz(t),\ \ i = 1,2,\ldots d \text{，}$$

定义：

$$y(t) = u(t, x(t)) = u(t, x_1(t), \ldots x_d(t)) : [0,T]\times R^d \to R \text{。}$$

注意，维数的相容性：

$$\begin{array}{cccc} dx(t) & = f(t)dt & + \quad \sigma(t) & dz(t) \\ (R^d \times R^1) & (R^d \times R^1) & (R^d \times R^1) & (R^1 \times R^1) \end{array}$$

与情形 1 一样，运用 Taylor 定理，作必要的代换，重排各项，并利用乘法规则：

$$(dt)^2 = 0; (dz)^2 = dt; dt \times dz = 0 \tag{4.12}$$

得（求和从 1 到 d；t 被省略）：

$$\begin{aligned} dy(t) &= u_t dt + \sum u_{x_i} dx_i + \frac{1}{2}\sum\sum u_{x_i x_j} dx_i dx_j \\ &= u_t dt + \sum u_{x_i}(f_i dt + \sigma_i dz) + \frac{1}{2}\sum\sum u_{x_i x_j}(f_i dt + \sigma_i dz)(f_j dt + \sigma_j dz) \\ &= u_t dt + \sum u_{x_i} f_i dt + \sum u_{x_i}\sigma_i dz + \frac{1}{2}\sum\sum u_{x_i x_j}\sigma_i\sigma_j]dt \\ &= (u_t + u_{x_i} f_i + \frac{1}{2}\sum\sum u_{x_i x_j}\sigma_i\sigma_j)dt + (\sum u_{x_i}\sigma_i)dz \text{。} \end{aligned}$$

情形 3：进一步特殊化情形 2，假设 f 和 σ 都是实值函数。这是 $It\hat{o}$ 引理所陈述的一种情形。我们直接计算：

$$\begin{aligned} dy(t) &= u_t dt + u_x dx + \frac{1}{2}u_{xx}(dx)^2 \\ &= u_t dt + u_x(fdt + \sigma dz) + \frac{1}{2}u_{xx}(fdt + \sigma dz)^2 \\ &= u_t dt + u_x f dt + u_x \sigma dz + \frac{1}{2}u_{xx}\sigma^2 dt \\ &= (u_t + u_x f + \frac{1}{2}u_{xx}\sigma^2)dt + u_x \sigma dz \end{aligned}$$

这就是 $It\hat{o}$ 引理的结果。注意，为了产生：

$$(fdt + \sigma dz)^2 = \sigma^2 dt$$

运用了（4.12）中的乘法规则。

5. 例题

第三章和第四章将给出几个来自经济学与金融学的例子，来说明 $It\hat{o}$ 引理的应用。这里给出一些纯数学意义的 $It\hat{o}$ 引理的例子。

例 1：设 $y(t) = u(x(0), t) = x(0)e^{at}$ 且 $dx(t) = ax(t)dt$ 其中 a 为非零实数，那么：

$$\begin{aligned} dy(t) &= u_t dt + u_x dx + \frac{1}{2}u_{xx}(dx)^2 \\ &= ax(0)e^{at}dt \end{aligned}$$

即 $y(t)$ 是方程 $dx(t) = ax(t)dt$ 的解。

例 2：假设 $y(t) = u(x(t)) = e^{x(t)}$ 且 $y(0) = e^{x(0)} > 0$，又设

$dx(t) = -\dfrac{1}{2}\sigma^2(t)dt + \sigma(t)dz(t)$，由 $It\hat{o}$ 引理，得：

$$\begin{aligned} dy(t) &= u_t dt + u_x dx(t) + \tfrac{1}{2}u_{xx}(dx(t))^2 \\ &= e^{x(t)}[-\tfrac{1}{2}\sigma^2(t)dt + \sigma(t)dz(t)] \\ &\quad + \tfrac{1}{2}e^{x(t)}[-\tfrac{1}{2}\sigma^2(t)dt + \sigma(t)dz(t)]^2 \\ &\quad + \tfrac{1}{2}e^{x(t)}[-\tfrac{1}{2}\sigma^2(t)dt + \sigma(t)dz(t)]^2 \\ &= -\tfrac{1}{2}e^{x(t)}\sigma^2(t)dt + e^{x(t)}\sigma(t)dz(t) + \tfrac{1}{2}e^{x(t)}\sigma^2(t)dt \\ &= e^{x(t)}\sigma(t)dz(t) \end{aligned}$$

这里指出：

$$\begin{aligned} y(t) &= y(0)e^{x(t)} \\ &= y(0)\exp\left\{-\tfrac{1}{2}\int_0^t \sigma^2(s)ds + \int_0^t \sigma(s)dz(s)\right\} \end{aligned}$$

满足 $dy(t)$ 的方程，对一切 $t \in [0,T]$。

例 3：假设 $y(t) = u(x(t)) = e^{x(t)}$，其中 $dx(t) = -\tfrac{1}{2}dt + dz(t)$，且 $x(0) = 0$。运用 $It\hat{o}$ 引理，得：

$$\begin{aligned} dy(t) &= u_t dt + u_x dx(t) + \tfrac{1}{2}u_{xx}[dx(t)]^2 \\ &= e^{x(t)}[-\tfrac{1}{2}dt + dz(t)] + \tfrac{1}{2}e^{x(t)}[-\tfrac{1}{2}dt + dz(t)]^2 \\ &= e^{x(t)}dz(t) \end{aligned}$$

这是例 2 的特例。它说明,给定随机微分方程：

$$dy(t) = e^{x(t)}dz(t) = y(t)dz(t) \tag{5.1}$$

其中初始条件 $y(0) = e^{x(0)} = 1$，它的解为：

$$y(t) = 1\exp\left\{-\tfrac{1}{2}t + z(t)\right\} \tag{5.2}$$

这个结果很有用，它揭示了常微分方程与随机微分方程的差别。如果（5.1）是一个常微分方程，那么它的解应是 $y(t) = \exp[z(t)]$，与（5.2）不同。

例 4：假设 $y(t) = u(x_1, x_2) = tz(t)$ 且 $dx_1(t) = dt$，$dx_2(t) = dz(t)$，由 $It\hat{o}$ 引理，计算得：

$$\begin{aligned} dy(t) &= u_t dt + \sum u_{x_i} dx_i(t) + \tfrac{1}{2}\sum\sum u_{x_i x_j} dx_i(t)dx_j(t) \\ &= z(t)dt + tdz(t) + \tfrac{1}{2}[dtdz(t) + dz(t)dt] \\ &= z(t)dt + tdz(t) \end{aligned}$$

例 5：假设 $y(t) = u(x_1, x_2) = tx_1x_2$，且：

$$dx_1(t) = f_1(t)dt + \sigma_1(t)dz(t)$$
$$dx_2(t) = \sigma_2(t)dz(t)$$

运用 $It\hat{o}$ 公式，计算得：

$$\begin{aligned}dy(t) &= u_t dt + \sum u_{x_i} dx_i(t) + \tfrac{1}{2}\sum\sum u_{x_i x_j} dx_i(t)dx_j(t) \\ &= x_1x_2dt + tx_2dx_1(t) + tx_1dx_2(t) \\ &\quad + \tfrac{1}{2}[tdx_1(t)dx_2(t) + tdx_2(t)dx_1(t)] \\ &= x_1x_2dt + tx_2[f_1dt + \sigma_1dz] + tx_1\sigma_2dz \\ &\quad + \tfrac{1}{2}[t\sigma_1\sigma_2dt + t\sigma_1\sigma_2dt] \\ &= (x_1x_2 + tx_2f_1 + t\sigma_1\sigma_2)dt + (tx_2\sigma_1 + tx_1\sigma_2)dz\text{。}\end{aligned}$$

例 6：假设 $y(t) = u(x(t)) = u(z(t))$，其中 u 关于 x 二次连续可微，$x(0) = z(0) = 0$ 且 $dx(t) = dz(t)$。由 $It\hat{o}$ 公式，得：

$$\begin{aligned}dy(t) &= u_t dt + u_x dx(t) + \tfrac{1}{2}u_{xx}[dx(t)]^2 \\ &= u'(z(t))dz(t) + \tfrac{1}{2}u''(z(t))dt\end{aligned} \tag{5.3}$$

这里符号“'”和“"”表示通常的导数。如果把（5.3）写成积分形式，那么它变成：

$$y(t) = u(z(t)) = u(0) + \int_0^t u'(z(s))dz(s) + \tfrac{1}{2}\int_0^t u''(z(s))ds \tag{5.4}$$

从（5.4）中，解出右边的第二项，得：

$$\int_0^t u'(z(s))dz(s) = u(z(t)) - u(z(0)) - \tfrac{1}{2}\int_0^t u''(z(s))ds \tag{5.5}$$

方程（5.5）通常叫做 $It\hat{o}$ 随机积分基本定理。它是非常有用的，因为它把（5.5）左边的随机积分表示成了普通积分。

上面的例子只是在数学领域展示了 $It\hat{o}$ 引理的应用，可能与经济分析中的应用相差甚远。然而，人们已经发现 $It\hat{o}$ 引理在经济学和金融中的许多应用。这里，我们提及三个大的应用领域：第一，现代衍生资产定价理论。主要的模型是著名的 Black & Scholes 期权定价公式。第二，在随机场合中，适当的预算约束的阐述。主要的模型是 Merton 关于最优消费和投资组合的工作。第三，最优随机控制问题的描述和分析。这一点将在本章第 10 节进行举例说明。

6. 随机微分方程

在这一节中，我们研究保证 $It\hat{o}$ 型随机微分方程解存在且唯一的条件。这种方程首先在上面的第 2 节的（2.12）式中被引入。回忆一下，这种方程能转换成如（3.4）式的积分方程，也正是（3.4）式，促使了随机积分的发展。

考虑概率空间(Ω, F, P)和区间$[0,T]$，当然也可以是$[0,\infty)$，把$It\hat{o}$随机微分方程写成：

$$dx(t) = f(t, x(t))dt + \sigma(t, x(t))dz(t) \tag{6.1}$$

初始条件为：

$$x(0,\omega) = c(\omega) = c \tag{6.2}$$

方程（6.1）和（6.2）定义了一个随机初值问题，它能表示成一个随机积分方程：

$$x(t) = c + \int_0^t f(s, x(s)ds + \int_0^t \sigma(s, x(s))dz(s) \tag{6.3}$$

同前面一样，$x(t)$是一个实值随机过程，f和σ是定义在$[0,T]\times R$上实值可测函数，$z(t)$是一个维纳过程。我们假设$x(t)$为F_t-可测，且对一切$t \geq 0$，F_t必须与由$z_u - z_t$，$u \geq t$生成的$\sigma-$域相互独立。这里，选择F_t为使得初始随机变量$c(\omega)$和随机变量z_s，$s \leq t$可测的最小$\sigma-$域就足够了。现在，我们准备给出定义。

称随机过程$x(t)$是（6.1）和（6.2）的解，即一个$It\hat{o}$随机微分方程在$[0,T]$上的解。如果它满足下面的三个性质：

（1）$x(t)$是F_t-可测的，即对于$t \in [0,T]$非预期。

（2）函数f和σ使得以概率 1 有：

$$\int_0^T \left| f(t, x(t)) \right| dt < \infty \text{ 且 } \int_0^T \left| \sigma(t, x(t)) \right|^2 dt < \infty$$

（3）对一切$t \in [0,T]$，以概率 1 地有（6.3）式成立。

在随机微分方程的理论和应用中，从分析的角度来看，解的存在性和唯一性问题很重要。通常，在研究随机微分方程模型解的各种概率性质之前，研究者的兴趣放在建立解的存在性和唯一性。现在， 我们叙述随机微分方程解的存在性和唯一性定理。解的各种基本性质放在下一节讨论。

定理 6.1　考虑方程（6.1），初始条件由（6.2）给出且假定下列条件满足：

（1）　函数$f(t,x)$和$\sigma(t,x)$对于$t \in [0,T]$，$x \in R$有定义，且二元可测。

（2）　存在常数$K > 0$，使得对于$t \in [0,T]$，$x \in R$，$y \in R$，有：

$$\left| f(t,x) - f(t,y) \right| + \left| \sigma(t,x) - \sigma(t,y) \right| \leq K \left| x - y \right| \tag{6.4}$$

$$\left| f(t,x) \right|^2 + \left| \sigma(t,x) \right|^2 \leq K^2 (1 + \left| x \right|^2) \tag{6.5}$$

（3） （6.2）式中的初始条件 $x(0,\omega)$ 不依赖于 $z(t)$，且 $Ex(0,\omega)^2 < \infty$。 那么，方程（6.1）存在满足初始条件（6.2）的解 $x(t)$，它以概率 1 唯一，以概率 1 地有连续样本轨道，且 $\sup\limits_t Ex(t)^2 < \infty$。

定理的证明过程参看 Gihman 和 Skorohod(1972,pp.40-43)。证明方法是利用 Picard-Lindelof 迭代（这一点与常微分方程所使用的相类似）和 Borel-Cantelli 引理。

很明显，该定理对于向量值函数 $x(t)$，$z(t)$，$f(t)$ 和合适维数的矩阵值过程 $\sigma(t)$ 也是成立的。

方程（6.4）叫做 Lipschitz 条件，而（6.5）叫做增长限制条件。解的唯一性是指，如果 $x_1(t)$ 和 $x_2(t)$ 是两个解，那么：

$$P[\sup_t |x_1(t) - x_2(t)| = 0] = 1 \tag{6.6}$$

注意，如果函数 $f(t,x)$ 和 $\sigma(t,x)$ 在 $[0,\infty)\times R$ 上有意义，且在 $[0,\infty)$ 的每一个有限区间 $[0,T]$ 上，定理 6.1 的假设成立，那么我们说随机微分方程有一个定义在 $t\in[0,\infty)$ 上的全局解。自治随机微分方程，即 $f(t,x) = f(x), \sigma(t,x) = \sigma(x)$，就是这种情形。

方程（6.4）成立的充分条件之一是函数 $f(t,x)$ 和 $\sigma(t,x)$ 对每一个 $t\in[0,T]$，两者都有关于 x 的一阶连续偏导数，且它们在 $[0,T]\times R$ 上有界。

作为(6.1)的特殊情况，考虑自治随机微分方程，它在 Gihman & Skorohod(1972,pp.106)中处理过：

$$dx(t) = f(x(t))dt + \sigma(x(t))dz(t) \tag{6.7}$$

其中函数 $f(x)$ 和 $\sigma(x)$ 对所有 $x\in R$ 有定义。假定 $f(x)$ 和 $\sigma(x)$ 满足下列条件：

存在某一常数 $K>0$，有

$$|f(x)| + |\sigma(x)| \le K(1+|x|) \tag{6.8}$$

成立，同时，对于每一个 $C>0$，存在一个正实数 L_C 使得当 $|x|\le C, |y|\le C$ 时，有：

$$|f(x)-f(y)| + |\sigma(x)-\sigma(y)| \le L_C|x-y| \tag{6.9}$$

那么，对于每一个不依赖于 $z(t)$ 的随机初始条件 $c(\omega)$，方程（6.7）存在满足初始条件的唯一解。这个结果在许多产生自治随机微分方程的应用问题中被用到。在这个情况中，方程（6.8）和（6.9）被用于建立解的存在性和唯一性。

最后，我们考虑一类重要的随机微分方程，即线性方程。设 $x(t)$ 一个 d 维过程，$a(t)$和$\sigma(t)$是$(\mathrm{d}\times\mathrm{d})$ 矩阵值函数，$z(t)$ 是 d 维维纳过程。考虑：

$$dx(t)=a(t)x(t)dt+\sigma(t)dz(t) \tag{6.10}$$

初始条件 $x(0)=c$ 。假定解的存在性和唯一性条件满足，那么问题是：在这种特殊情况下，解是什么样子？我们说（6.10）在$[0,T]$上的解有下列形式：

$$x(t)=\phi(t)[c+\int_0^t\phi^{-1}(s)\sigma(s)dz(s)] \tag{6.11}$$

其中 $\phi(t)$ 是方程

$$\dot{x}(t)=a(t)x(t) \tag{6.12}$$

的基本解矩阵。让我们证明一下这一论断。从方程

$$y(t)=c+\int_0^t\phi^{-1}(s)\sigma(s)dz(s) \tag{6.13}$$

开始，把它写成微分形式：

$$dy(t)=\phi^{-1}(t)\sigma(t)dz(t) \tag{6.14}$$

从（6.11）和（6.13）可得：

$$x(t)=\phi(t)y(t) \tag{6.15}$$

现在，运用 $It\hat{o}$ 引理，证明 $dx(t)$ 有（6.10）的形式。为了这一点，要用到（6.15）和（6.14）两个方程。

$$\begin{aligned}dx(t)&=\frac{\partial}{\partial t}[\phi(t)y(t)]dt+\frac{\partial}{\partial y}[\phi(t)y(t)]dy(t)\\&\quad+\tfrac{1}{2}\frac{\partial^2}{\partial y^2}[\phi(t)y(t)][dy(t)]^2\\&=\dot{\phi}(t)y(t)dt+\phi(t)dy(t)+0\\&=a(t)\phi(t)y(t)dt+\sigma(t)dz(t)\\&=a(t)x(t)dt+\sigma(t)dz(t)\text{。}\end{aligned}$$

为了进一步阐明（6.11），考虑线性情况：

$$dx(t)=ax(t)dt+\sigma(t)dz(t) \tag{6.16}$$

初值 $x(0)=c$ 。这里，由于矩阵 a 是自治的，那么，从常微分方程理论得知，系统 $\dot{x}(t)=ax(t)$

有基本解矩阵：

$$\phi(t)=e^{at}，\ y\in[0,T]，$$

其中 $\phi^{-1}(t)=e^{-at}$。运用（6.11），得到：

$$\begin{aligned}x(t)&=\phi(t)[c+\int_0^t\phi^{-1}(s)\sigma(s)dz(s)\\&=e^{at}[c+\int_0^t e^{-as}\sigma(s)dz(s)]\\&=ce^{at}+\int_0^t e^{a(t-s)}\sigma(s)dz(s)\end{aligned}$$

上述问题的更详细讨论以及另外的相关课题可参看 Arnold(1974)和 Gihman & Skorohod(1972)。

7. 解的性质

在一定的条件下，常微分方程的解满足两个有名的性质：解关于参数和初值的依赖性和解的可微性。自然地要问随机微分方程的解是否也满足类似的性质。

在本节中，我们将展示解关于参数和初值的依赖性成立，但解的可微性不成立。后者的原因是：如果 $x(t)$ 是 $It\hat{o}$ 随机微分方程的唯一解，那么它依赖于维纳过程 $z(t)$，$z(t)$ 的处处不可微性意味着 $x(t)$ 的不可微性。然而，如果我们定义均方可微性，那么在一定的假设条件下，可以证明 $x(t)$ 是均方可微的。除了这两个性质外，随机微分方程另外满足一些概率方面的性质，如：解是 Markov 过程，在一定的假设条件下解是扩散过程。现在，我们明确地建立各种性质。

考虑依赖于参数 p 的随机微分方程：

$$dx(p,t)=f(p,t,x)dt+\sigma(p,t,x)dz(t) \tag{7.1}$$

$t\in[0,T],\ p\in P$，其初始条件为 $x(p,0)=c(p)$，其中 P 为参数集。用 $x(p,t)$ 表示（7.1）的解，设 $p_0\in P$，下述定理确立了解关于参数的依赖性。

定理 7.1 设 $x(p,t)$ 是（7.1）的解。假设对所有 p，$f(p,t,x)$ 和 $\sigma(p,t,x)$ 满足存在性和唯一性条件，即满足（6.4）与（6.5）。同时还假定：

（1）当 $p\to p_0$ 时，$c(p)\xrightarrow{P}c(p_0)$； （7.2）

（2）对于每一个 $N>0,|x|\le N,\ t\in[0,T]$，当 $p\to p_0$ 时，

$$\lim \sup_{x}(\left| f(p,t,x) - f(p_0,t,x) \right| + \left| \sigma(p,t,x) - \sigma(p_0,t,x) \right|) = 0\text{；}$$

（3）存在一个与 p 无关的常数 K，使得对于 $t \in [0,T]$ 有：

$$\left| f(p,t,x) \right|^2 + \left| \sigma(p,t,x) \right|^2 \le K^2(1 + |x|)^2$$

那么，当 $p \to p_0$ 时，有：

$$\sup_{t} \left| x(p,t) - x(p_0,t) \right| \xrightarrow{P} 0 \qquad (7.3)$$

定理的证明过程参见 Gihman & Skorohod(1972, pp. 54-55)。 作为定理 7.1 的特例，假定函数 f 和 σ 不依赖于参数 p，即：

$$dx(p,t) = f(t,x)dt + \sigma(t,x)dz(t), t \in [0,T] \qquad (7.4)$$

初始条件 $x(p,0) = c(p)$。这里，我们仍然让初始条件依赖于参数 p。作为定理 7.1 的推论，我们得到解对初值的随机连续依赖性，即对于方程（7.4），如果假定解的存在性和唯一性，那么（7.2）意味着（7.3）。

现在，我们定义随机过程的均方微分。设 $x(t), t \in [0,T]$，为随机过程。令 $u \in [0,T]$ 是一特定的点。我们称 $x(t)$ 在 $t = u$ 处均方可微，随机变量 $y(u)$ 是它的导数，如果当 $h \to 0$ 时有：

$$E\left| [x(u+h) - x(u)]/h - y(u) \right|^2 \to 0$$

我们把上述定义运用于带参数的随机过程。过程由下式给出：

$$x(s,t,c) = c + \int_s^t f(v, x(s,v,c))dv + \int_s^t \sigma(v, x(s,v,c))dz(v) \qquad (7.5)$$

其中 $v \in [s,T], 0 \le s \le t \le T,$ 且 $x(s,s,c) = c \in R$。一个自然会产生的问题是：在什么条件下，（7.5）中的 $x(s,t,c)$ 关于初值 c 均方可微？下列定理给出了答案。

定理 7.2 假定（7.5）中的 f 和 σ 关于 (t,x) 连续，关于 x 的一阶和二阶偏导数有界，那么，对于给定的 $t = u, u \in [s,T]$，解 $x(s,u,c)$ 关于 c 二次均方可微，且：

$$\frac{\partial}{\partial c} x(s,u,c) = 1 + \int_s^u f_x(v, x(s,v,c)) \frac{\partial}{\partial c} x(s,v,c)dv$$
$$+ \int_s^u \sigma_x(v, x(s,v,c)) \frac{\partial}{\partial c} x(s,v,c)dz(v)\text{。}$$

本定理的证明及其推广参看 Gihman & Skorohod(1972), pp. 59-62)。下面我们回忆 Markov 过程的定义并证明随机微分方程解是 Markov 过程。

称定义在概率空间 (Ω, F, P) 上的实值随机过程 $x(t), t \in [0,T]$，是一个 Markov 过程，如果对于 $0 \le s \le t \le T$ ，每一个 $B \in \Re$ ，以概率 1 使下列方程成立：

$$P[x(t) \in B \mid F_s] = P[x(t) \in B \mid \sigma(x(s))]。$$

这里，$\Re$ 表示 R 的 Borel $-\sigma$ 域，F_s 表示由 $x(s)$ 在 $s \in [0,T]$ 时生成的 $\sigma-$ 域, $\sigma(x(s))$ 表示由单个随机变量 $x(s)$ 生成的 $\sigma-$ 域。简言之，上述定义表明对于 Markov 过程，当现在已知时，过去与将来在统计上相互独立。对于一个 Markov 过程 $x(t)$ ，从条件概率理论，我们得到转移概率函数的存在性，用 $P(s,x,t,B)$ 表示。它的意义是在给定 $x(s)=s$ 的条件下，$x(t) \in B$ 的概率，其中， $0 \le s \le t \le T$ 。

定理 7.3 考虑随机微分方程（6.1）和初始条件（6.2），假定存在唯一解 $x(t), t \in [0,T]$。那么 $x(t)$ 是一个 Markov 过程，其在 $t=0$ 处的初始概率分布为 c ，转移概率由

$$P(s,x,t,B) = P[x(t) \in B \mid x(s) = x]$$

给定。

定理的证明过程参看 Arnold(1974, pp. 147)。扩散过程是一种特殊的 Markov 过程，我们称具有几乎必然连续样本轨道的实值 Markov 过程 $x(t), t \in [0,T]$ 为扩散过程。如果对于每一个 $s \in [0,T], x \in R$ 和 $\varepsilon > 0$ ，它的转移概率函数 $P(s,x,t,B)$ 满足下列三个条件：

（1） $$\lim_{t \downarrow s} \frac{1}{t-s} \int_{|y-x|>\varepsilon} P(s,x,t,dy) = 0 \qquad (7.6)$$

（2）存在一个实值函数 $f(s,x)$ ，使得：

$$\lim_{t \downarrow s} \frac{1}{t-s} \int_{|y-x|\le\varepsilon} (y-x) P(s,x,t,dy) = f(s,x) \qquad (7.7)$$

（3）存在一个实值函数 $h(s,x)$ 使得：

$$\lim_{t \downarrow s} \frac{1}{t-s} \int_{|y-x|\le\varepsilon} (y-x)^2 P(s,x,t,dy) = h(s,x) \qquad (7.8)$$

注意， f 叫做趋势系数，h 叫做扩散系数，它们从条件（2）和（3）得到。条件（1）是说在短期内 $x(t)$ 发生大的变化是不可能的。

定理 7.4 考虑带有初始条件（6.2）的随机微分方程（6.1）。假定存在唯一解 $x(t), t \in [0,T]$，还假定函数 f 和 σ 关于 t 连续，那么解 $x(t)$ 为扩散过程，趋势系数为 $f(t,x)$，扩散系数为 $h(t,x) = \sigma^2(t,x)$。

这个结果的证明见 Arnold(1974, pp. 153)。

定理 7.4 的作用可以解释如下。注意，从扩散过程的定义可见，对于定义趋势系数 $f(t,x)$ 和扩散系数 $h(t,x)$，转移概率系数 $P(s,x,t,B)$ 起决定性的作用。随之而来的问题是：假定一个扩散过程是由系数 f 和 h 给定，能否从 f 和 h 得到 $P(s,x,t,B)$？答案是肯定的。实际上，扩散过程的确定性是：在一定的条件下，转移概率 $P(s,x,t,B)$ 由系数 f 和 h 唯一决定。这个事实令人吃惊，因为 f 和 h 只涉及一阶和二阶矩。而一阶和二阶矩，通常不足以定义分布。因此，给定一个满足定理 7.4 的假设的随机微分方程，它的解是一个扩散过程，在一定的条件下，我们能得到其转移概率，而不需要真正地求出显式解。如何做到这一点呢？要回答这个问题，我们需要以下一些定义并陈述有关的定理。

对于每一个具有系数 $f = (f_i), i = 1,2,...,d$ 和 $h = (h_{ij}), i, j = 1,2,...,d$ 的扩散过程，我们指定一个二阶微分算子：

$$D = \sum_i f_i(s,x)\frac{\partial}{\partial x_i} + \tfrac{1}{2}\sum_i \sum_j h_{ij}(s,x)\frac{\partial^2}{\partial x_i \partial x_j} \tag{7.9}$$

对于每一个二次偏可微函数 $g(x)$, 可以形式地写下 Dg。如果 f 和 h 为实值，那么（7.9）变成：

$$D = f(s,x)\frac{d}{dx} + \tfrac{1}{2}h(s,x)\frac{d^2}{dx^2} \tag{7.10}$$

用 $E_{s,x}$ 表示在时刻 s, x 处的条件数学期望。在我们的分析中，下列定理是重要的。

定理 7.5 假定 $x(t), t \in [0,T]$ 是一个 d 维扩散过程，具有连续的系数 $f(s,x)$ 和 $h(s,x)$，且（7.6），（7.7）和（7.8）中的条件（1）、（2）、（3）对于 $s \in [0,T]$ 一致成立。设 $g(x)$ 是一个连续、有界的实值函数，使得对于 $s < t$，t 固定，$x \in R^d$，有：

$$u(s,x) = E_{s,x}g(x(t)) = \int_{R^d} g(y)P(s,x,t,dy) \tag{7.11}$$

假定 u 有连续有界的偏导数：

$$\frac{\partial u}{\partial x_i};\frac{\partial^2 u}{\partial x_i \partial x_j};1 \le i,j \le d,$$

那么，$u(s,x)$关于s可微，且满足偏微分方程：

$$\frac{\partial u}{\partial s}+Du=0 \tag{7.12}$$

边界条件为当$s \to t$时，$u(s,x)=g(x)$。

定理的证明见 Gihman & Skorohod(1969, pp. 373)。方程（7.12）叫做 Kolmogorov 向后方程，因为微分是关于后向变量s和x求的，其中$s<t,x<y$。Kolmogorov 向后方程使我们能确定$P(s,x,t,y)$。首先，如果知道了（7.11），那么$P(s,x,t,y)$唯一决定。其次，如果对于连续有界实值函数g，（7.12）有唯一解，那么，对于给定的f和h，我们能计算出$u(s,x)$，然后再从它之中计算出$P(s,x,t,y)$。更特别地，对于随机微分方程，下列结果是定理 7.5 的推论。

定理 7.6　假定带有初始条件（6.2）的方程（6.1）满足定理 7.4 的条件，因此它有唯一解$x(t)$，且$x(t)$是一个扩散过程。仍假定f和σ具有连续有界的关于x的一阶和二阶偏导数。又设$g(x)$为连续有界的实值函数，且具有连续有界的一阶和二阶导数。对于$0 \le s \le t \le T$，$x \in R$，令：

$$u(s,x)=E_{s,x}g(x(s,t,x))$$

那么u和u关于x的一阶、二阶偏导数以及u关于s的偏导数连续有界且满足方程：

$$\frac{\partial}{\partial s}u(s,x)+Du(s,x)=0, \text{当}s \to t\text{时}\ u(s,x) \to g(x)$$

证明从定理 7.5 得出。这个结果显示随机微分方程及其解的研究与形如（7.12）的二阶偏微分方程的研究有密切关系。

对于形如（6.7）那样的特殊自治随机微分方程，由定理 7.3 和 7.4 的推论我们立即得出，它们的解是 Markov 过程，而且是扩散过程。

本节的最后一个结果是关于解的矩的。

定理 7.7　考虑带有初始条件（6.2）的随机微分方程（6.1），假定它有唯一解。同时假定：

$$E|c|^{2n}<\infty,$$

其中n是正整数。那么（6.1）的解$x(t),t \in [0,T]$，满足：

$$E|x(t)|^{2n} \le (1+E|c|^{2n})e^{Ct},$$

$$E|x(t)-c|^{2n} \le D(1+E|c|^{2n})te^{Ct}$$

其中$C=2n(2n+1)K^2$, D为只依赖于n,T和（6.4）中的K的常数。

证明参见 Arhold（1974，pp.116-118）。

8. 点均衡和稳定性

平衡解是随机微分方程的一个特解，它在许多应用问题中起重要作用。

考虑方程（6.1），对于$t\in[0,T]$有定义。如果存在一个非随机的常数c使得对所有$t\in[0,T]$，有：

$$f(t,c)=\sigma(t,c)=0 \tag{8.1}$$

而对于所有的$t\in[0,T]$，以概率 1 有$x(t,\omega)=c$，那么，我们说c是一个点平衡解或者简称平衡解。这个概念的另一个术语是平稳点或稳态点。注意，在定义中，指定c（如果它存在）为非随机常数，或等价地称为单点分布。

存在有许多其他的平衡解概念的定义，有关这些定义的综述，读者可查阅 Majumdar(1975)。一个与上面所述的不同的定义在下一节里给出。

假定$It\hat{o}$型自治随机微分方程

$$dx(t)=f(x(t))dt+\sigma(x(t))dz(t) \quad t\in[0,T] \tag{8.2}$$

存在平衡解。在本节中，我们将研究形如（8.2）的自治随机微分方程平衡解的稳定性。为了方便和不失一般性，我们假定0是一个平衡解。为了使读者熟悉两种方法，我们使用两种方法讨论随机稳定性，尽管它们产生相同的结果。两种方法分别叫做 Gihman-Skorohod 方法和 Liapunov-Kushner 方法。

8.1 Gihman-Skorohod 方法

对该方法的陈述在 Gihman-Skorohod(1972, pp.145-151)中。

考虑方程（8.2），假定对于$t\in[0,\infty)$，0是唯一的平衡解，且：

$$f(0)=\sigma(0)=0 \tag{8.3}$$

我们说$0-$平衡是稳定的，如果对于任意的$\varepsilon>0$，存在一个$\eta>0$，使得如果$|x|<\eta$，那么，当$t\to\infty$时，有：

$$P[\lim x(t)=0]\ge 1-\varepsilon \tag{8.4}$$

这里 $x = x(0,\omega)$，即 x 是初始条件，且在（8.4）中，其概率是关于 x 的条件概率。如果对于任意的 $\varepsilon > 0$，存在一个 $\eta > 0$，使得（8.4）只对 $x \in (0,\eta)$ 或者 $x \in (-\eta,0)$ 成立，则我们相应地称 $0-$ 平衡是右稳定或者左稳定。

下面的定理给出了稳定性的条件。

定理 8.1 假定对于方程（8.2）条件（8.3）成立，且存在一个 $\eta > 0$，使得 $\sigma(x) > 0, 0 < |x| < \eta$。那么，（1）右稳定、（2）左稳定和（3）两边稳定的充分条件分别为：对于某个 $\eta > 0$，有（1）$I_1 < \infty$；（2）$I_2 < \infty$ 和（3）$I_1 + I_2 < \infty$。其中：

$$I_1 = \int_0^\eta \exp\left\{\int_u^\eta \frac{2f(y)}{\sigma^2(y)}dy\right\}du \tag{8.5}$$

$$I_2 = \int_{-\eta}^0 \exp\left\{-\int_{-\eta}^u \frac{2f(y)}{\sigma^2(y)}dy\right\}du \tag{8.6}$$

定理的证明见 Gihman & Skorohod(1972, pp. 146-148)，本定理是有用的，因为它把稳定性的研究简化成（8.5）和(8.6)的计算。

有时候，要证实一个平衡解的稳定性很困难，因为给定的模型可能没有这种性质。在这种情况下，研究人员也许会研究方程的解是否有界。在许多应用中，解的有界性是仅次于随机稳定性的最需要性质。因此，正如下一个定理所述，利用类似（8.5）和（8.6）的方程来建立有界性是不足为奇的。

定理 8.2 假定方程（8.2）和初始条件 $x(0) = c$ 有唯一解，用 $x(t)$ 表示，而且对于 $x \in R$ 有 $\sigma(x) > 0$。令：

$$I_1 = \int_{-\infty}^x \exp\left\{-\int_0^u \frac{2f(y)}{\sigma^2(y)}dy\right\}du \tag{8.7}$$

$$I_2 = \int_x^\infty \exp\left\{-\int_0^u \frac{2f(y)}{\sigma^2(y)}dy\right\}du \tag{8.8}$$

如果 $I_1(x) < \infty$ 和 $I_2(x) = \infty$，那么对于 $t \in [0,T]$，$P[\omega : \sup_t x(t,\omega) < \infty] = 1$。

这个定理的证明见 Gihman & Skorohod(1972, pp. 119)。

现在，我们陈述 Liapunov-Kushner 方法的基本内容。

8.2 Liapunov-Kushner 方法

该方法是由 Kushner(1967, pp. 27-76)发展的，它是确定性的 Liapunov 间接稳定性方法

的推广。叙述主要定理之前，我们先介绍一些预备概念。

第一，我们引进一个函数$V(x)$，它起 Liapunov 函数的作用，其性质在定理 8.3 中描述。

第二，定义集合$Q_m \equiv \{x : V(x) < m < \infty\}$，且假定$Q_m$是有界集。

第三，$D_m V(x)$是由（7.10）给出的微分算子。它相当于在确定稳定性中的轨道导数，代表过程$V(x(t))$在t时刻给定$x(t) = x$的平均时间变化率。于是，对于形如（8.2）那样的方程，我们有：

$$D_m V(x) = f(x)\frac{dV(x)}{dx} + \tfrac{1}{2}\sigma^2(x)\frac{d^2V(s)}{dx^2} \tag{8.9}$$

定理 8.3 设$x(t)$是（8.2）的唯一解，$V(x)$是一个有界连续非负函数，且在Q_m中有有界连续的一阶和二阶导数。如果在Q_m中有$D_m V(x) \le 0$，那么$x(t) \to \bar{x}$，其中$\bar{x}$至少以$1 - V(x)/m$的概率属于集合$\{x : D_m V(x) = 0\} \cap Q_m$。

这个定理的证明见 Kusher(1967, pp.42)为了演示定理 8.3，我们来看由 Kusher(1967, pp.55-56)给出的一个简单例子。设$x(t)$是形如

$$dx = axdt + \sigma xdz \tag{8.10}$$

的$It\hat{o}$方程的解。取$V(x) = x^2$，那么，对每一个集合$Q_m = \{x : x^2 < m < \infty\}$，$V(x)$非负，有界且在$Q_m$中具有有界的一阶、二阶导数。运用（8.9），得到

$$D_m V(x) = ax2x + \tfrac{1}{2}2(\sigma x)^2 = x^2(2a + \sigma^2)$$

如果$2a + \sigma^2 < 0$，那么$D_m V(x) \le 0$。于是根据定理 8.3，有$x(t) \to 0,\ w.p.1$，因为m可以任意大。

最后，作两点注释是适宜的。第一，两种稳定性方法得到的结果表明，$It\hat{o}$方程的解不只是逗留在$0-$平衡解的小邻域内，而是真正趋于$0-$平衡解。在专业技术上，我们把这种稳定刻画为局部渐近随机稳定。如果$0-$平衡是稳定的，且对于任意的非随机的初始条件$x \in R$，有$x(t) \to 0,\ w.p.1$，那么我们说该$0-$平衡是全局渐近稳定的，或者说大范围随机稳定性成立。

第二，如同确定性情况一样，在求证稳定性的过程中，找一个合适的 Liapunov 函数通常很困难。对于形如（8.2）的$It\hat{o}$方程，Kushner(1967, pp. 60-61)曾叙述过一种将会在下面介绍的方法。类似的方法 Feller(1954)也作过陈述。在叙述这种方法时，注意它与

Gihman-Skorohod 方法的相似之处，特别是与方程（8.5）式的相似。

假定 Liapunov 函数$V(x)$存在，且有$V(0)=0$，$V(x)>0$对于$x\neq 0$。设$V(x)$为连续、有界函数，且有有界、连续的一阶和二阶导数，使得$D_mV(x)$在某一个有界开集Q_m中有定义，且（8.9）式成立。设：

$$D_mV(x)=f(x)V_x(x)+\tfrac{1}{2}\sigma^2(x)V_{xx}(x)\leq 0 \tag{8.11}$$

由（8.11）我们得到：

$$\frac{V_{xx}(x)}{V_x(x)}\leq -\frac{2f(x)}{\sigma^2(x)} \tag{8.12}$$

当等式严格成立时，满足（8.12）的函数是：

$$V(x)=\int_0^x \exp\left\{-\int_0^u \frac{2f(y)}{\sigma^2(y)}dy\right\}du \tag{8.13}$$

只要（8.13）中的积分存在。因此，在 Gihman-Skorohod 方法中使用的方程（8.5）和（8.6）起着 Liapunov 函数的作用。稳定性的进一步研究参考 Ladde & Lakshmikanthan(1980,pp.56-91)。

9. 平稳分布的存在性

平稳分布（或者稳态分布，均衡分布）的概念是较前一节讨论的平衡解更广泛的概念。这样的平稳分布可从随机微分方程的解中得到，它是一个不依赖于时间的随机变量，即$x(t,\omega)=x(\omega)$。

随机微分方程的平稳分布的存在性问题，在数学文献中只解决了某些特殊情况。该领域中的一些先驱性工作包括 Feller（1954）、Tanaka（1957）和 Khas'minskii（1962）等等。他们获得的结果总结在 Mandl（1968）的书中，且 Bourguignon（1974）和 Merton（1975a）把它应用于经济学中。下面，我们将陈述形如（8.2）的自治$It\hat{o}$随机微分方程的平稳分布的存在性条件。下一章，我们将利用该结果得到不确定情况的新古典主义经济增长模型的平稳分布。

考虑方程(8.2)，假设$f(x)$和$\sigma(x)$在$[0,\infty)$中连续可微，在$(0,\infty]$上$\sigma(x)>0$，且(8.3)式成立。从定理 7.4 我们知道，(8.2）的解$x(t)$是一个取值于$[0,\infty]$的扩散过程。进一步，区间$[0,\infty]$的端点是吸收状态，即如果$x(t)=0$，那么对于$u>t$ $x(u)=0$，类似地，如果$x(t)=\infty$，那么对于$u>t$，$x(u)=\infty$。

一般地，在下列意义下，平稳分布总是存在的：(1）在边界点之一，x被吸收。(2）在

区间$(0,\infty)$上它有有限密度函数。（3）它有（1）和（2）的离散的概率混合。可能性（2）是一个非平凡的情况，在已叙述的假设下，它对于（8.2）十分重要。然而，要注意在情况（2）中，边界点是不可及的，即当$\varepsilon \to 0$时：

$$P[x(t) \le \varepsilon] \to 0 \tag{9.1}$$

同时：

$$P[x(t) \ge 1/\varepsilon] \to 0 \tag{9.2}$$

（9.1）和（9.2）成立的充要条件是（9.3）—（9.5）式成立。

$$\int_0^x I_2(u)du = \infty \tag{9.3}$$

$$\int_x^\infty I_2(u)du = \infty \tag{9.4}$$

$$\int_0^\infty I_1(u)du < \infty \tag{9.5}$$

其中$I_1(u)$和$I_2(u)$定义如下：

$$I_1(u) = \frac{1}{\sigma^2(u)} \exp\left\{\int_0^u \frac{2f(y)}{\sigma^2(y)}dy\right\}$$

$$I_2(u) = \frac{1}{\sigma^2(u)} \int^u \exp\left\{\int_y^u \frac{2f(v)}{\sigma^2(v)}dv\right\} du$$

如果条件（9.3）—（9.5）满足，那么平稳分布存在，用$\pi(x)$表示，它由下式给出：

$$\pi(x) = \frac{m}{\sigma^2(x)} \exp\left\{\int^x \frac{2f(y)}{\sigma^2(y)}dy\right\} \tag{9.6}$$

选择常数m使得$\int_0^\infty \pi(x)dx = 1$。

10. 随机控制

在这一节里，我们建立几个有关随机控制的命题，它们在经济学中已得到应用。这里的分析是直观的，严格的分析参看 Fleming 和 Rishel（1975）。

考虑问题：

$$J(k(t),t,\infty)=\max_{v(\cdot)} E_t\int_t^{\infty} e^{-\rho s}u(k(s),v(s))ds \tag{10.1}$$

S.t.

$$dk(t)=T(k(t),v(t))dt+\sigma(k(t),v(t))dz(t)\text{，}k(t)\text{给定} \tag{10.2}$$

这里，$v=v(t)=v(t,\omega)$ 是控制变量，$k=k(t)=k(t,\omega)$ 是状态变量，$\rho\geq 0$ 是关于将来效用的折扣，u 表示效用函数，T 是技术的趋势成份，σ 是扩散成份。注意，E_t 表示关于 $k(t)$ 和 $v(t)$ 的条件数学期望。（10.1）和（10.2）描述的问题类似于 Arrow 和 Kurz（1970,pp.27-51）所研究的随机问题。解决问题的标准方法，如同 Arrow 和 Kurz 一样，是 Bellman 最优原理。即，一个最优策略应有这样的性质：无论初始决策所产生的状态是什么，剩下的决策对于由最初决策所产生的状态而言，必须构成最优决策。参见 Bellman（1957,pp.83）。这里讨论无折扣、有限水平的情况，即 $\rho=0$ 和 $N<\infty$。如下面指出的一样，这种情况将进一步细分成某些子情况。

10.1 一维情况的最大值原理

考虑（10.1）和（10.2）的特殊情况：

$$J(k(t),t,N)=\max_{v} E_t\int_t^{N} u(k,v)ds \tag{10.3}$$

$$\text{S.t. } dk(t)=T(k,v)dt+\sigma(k,v)dz\text{，}k(t)\text{给定} \tag{10.4}$$

利用 Bellman 的动态规划方法，问题（10.3）和（10.4）可以如下分析：

$$\begin{aligned}
J(k(t),t,N)&=\max_{v} E_t\int_t^{N} u(k,v)ds\\
&=\max_{v} E_t\int_t^{t+\Delta t} u(k,v)ds+\max_{v} E_{t+\Delta t}\int_{t+\Delta t}^{N} u(k,v)ds\\
&=\max_{v} E_t\int_t^{t+\Delta t} u(k,v)ds+J(k(t+\Delta t),t+\Delta t,N)\\
&=\max_{v} E_t[\int_t^{t+\Delta t} u(k,v)ds+J(k(t+\Delta t),t+\Delta t,N)] \qquad (10.5)\\
&=\max_{v} E_t[u(k(t),v(t))\Delta t+J(k(t),t,N)\\
&\quad +J_k\Delta k+J_t\Delta t+\tfrac{1}{2}J_{kk}(\Delta k)^2\\
&\quad +J_{kt}(\Delta k)(\Delta t)+\tfrac{1}{2}J_{tt}(\Delta t)^2+o(\Delta t)]
\end{aligned}$$

注意，为了得到（10.5）式使用了 Taylor 定理，并且假设 J 在某一个含有连接两点 $(k(t),t)$ 和 $(k(t+\Delta t),t+\Delta t)$ 的线段的开集中，有小于 3 阶的连续偏导数。把（10.4）近似写为：

$$\Delta k = T(k,v)\Delta t + \sigma(k,v)\Delta z + o(\Delta t) \tag{10.6}$$

把（10.6）代入（10.5）。回顾第二章方程（4.12）中的乘法原则，得：

$$\begin{aligned}0 = \max_{v} E_t[u(k(t),v(t))\Delta t + (J_k T + J_t + \tfrac{1}{2} J_{kk}\sigma^2)\Delta t \\ + J_k \sigma\Delta z + o(\Delta t)]\end{aligned} \text{ 。} \tag{10.7}$$

为了符号方便，令：

$$\Delta J = [J_t + J_k T + \tfrac{1}{2} J_{kk}\sigma^2]\Delta t + J_k \sigma \Delta z \tag{10.8}$$

利用（10.8），方程（10.7）变为：

$$0 = \max_{v} E_t[u(k(t),v(t))\Delta t + \Delta J + o(\Delta t)] \tag{10.9}$$

方程（10.9）是一个偏微分方程，其边界条件如下：

$$\frac{\partial J}{\partial k}(k(N),N,N) = 0 \tag{10.10}$$

把（10.9）中的 E_t 求到括号里面去，然后用 Δt 除两边，令 $\Delta t \to 0$，得：

$$0 = \max_{v}[u(k(t),v(t))\Delta t + J_t + J_k T(k(t),v(t)) + \tfrac{1}{2} J_{kk}\sigma^2(k(t),v(t))] \tag{10.11}$$

方程（10.11）通常写成：

$$-J_t = \max_{v}[u(k(t),v(t))\Delta t + J_k T(k(t),v(t)) + \tfrac{1}{2} J_{kk}\sigma^2(k(t),v(t))] \tag{10.12}$$

它是随机控制理论中的有名的 Hamilton-Jacobi-Bellman 方程。

让我们进一步继续分析，定义共态变量 $p(t)$ 如下：

$$p(t) = J_k(k(t),t,N) \tag{10.13}$$

从（10.13），立刻得出：

$$p_k = \frac{\partial p}{\partial k} = J_{kk} \tag{10.14}$$

利用（10.13）和（10.14），我们可以把（10.12）重写为：

$$-J_t = \max_{v} H(k,v,p,\frac{\partial p}{\partial k}) \tag{10.15}$$

其中 H 是（10.12）中括号内部的表达式的泛涵记号。接下来假定，（10.15）中的最大值问题的解函数 v 存在，表示为：

$$v^0 = v^0(k,p,\frac{\partial p}{\partial k}) \tag{10.16}$$

注意，v^0 只是沿最优轨道 $k(t)$ 和 t 的函数，这是因为 J_k 只是 $k(t)$ 和 t 的函数。在应用控制

论文献中，特别是在经济应用中，v^0 叫做策略函数。假定策略函数 v^0 存在，那么（10.15）可以重写为：

$$
\begin{aligned}
-J_t &= \max_v H(k, v, p, \frac{\partial p}{\partial k}) \\
&= H\left(k, v^0\left(k, p, \frac{\partial p}{\partial k}\right), p, \frac{\partial p}{\partial k}\right) \\
&= H^0\left(k, p, \frac{\partial p}{\partial k}\right)
\end{aligned} \quad (10.17)
$$

最后这个方程（10.17）是（10.12）右边内部表达式在存在最优控制 v^0 的假设下的泛涵记号，即：

$$H^0(k, p, \frac{\partial p}{\partial k}) = u(k, v^0) + pT(k, v^0) + \frac{1}{2}\frac{\partial p}{\partial k}\sigma^2(k, v^0) \quad (10.18)$$

根据上述分析，在本节中，我们的最终目的是推导、刻画状态变量和共态变量的随机微分方程组，即找出 dk 和 dp 的表达式。dk 的表达是可以从（10.4）、（10.16）和（10.18）得到，特别地：

$$
\begin{aligned}
dk &= T(k, v^0)dt + \sigma(k, v^0)dz \\
&= H_p^0(k, p, \frac{\partial p}{\partial k})dt + \sigma(k, p, \frac{\partial p}{\partial k})dz \\
&= H_p^0 dt + \sigma dz
\end{aligned} \quad (10.19)
$$

注意，在（10.19）的推导过程中用到的 $\frac{\partial H^0}{\partial p} = H_p^0 = T$，可从（10.18）中得到。接下来我们推导 dP 的表达式。利用在（10.13）中给出的 $p(t)$ 的定义、方程（10.4）和 $It\hat{o}$ 引理，得：

$$
\begin{aligned}
dp &= J_{kt}dt + J_{kk}dk + \tfrac{1}{2}J_{kkk}(dk)^2 \\
&= J_{kt}dt + J_{kk}(Tdt + \sigma dz) + \tfrac{1}{2}J_{kkk}(Tdt + \sigma dz)^2 \\
&= [J_{kt} + J_{kk}T + \tfrac{1}{2}J_{kkk}\sigma^2]dt + J_{kk}\sigma dz \text{。}
\end{aligned} \quad (10.20)
$$

为了简化方程（10.20），我们由（10.17）来计算 J_{kt}，并假设混合偏导数有如下等式：

$$
\begin{aligned}
-J_{tk} &= H_k^0 + H_p^0\frac{\partial p}{\partial k} + H_{pk}^0\frac{\partial^2 p}{\partial k^2} \\
&= H_k^0 + TJ_{kk} + \tfrac{1}{2}\sigma^2 J_{kkk}
\end{aligned} \quad (10.21)
$$

注意，为得到（10.21），首先，我们使用了事实 $H_p^0 = T$ 和 $H_{pk}^0 = \frac{1}{2}\sigma^2$，两者均来源于（10.18）的偏微分。其次，使用了（10.14）中的定义。把（10.21）代入（10.20），得到我们所需要的结果：

$$\begin{aligned} dp &= [-H_k^0 - J_{kk}T - \tfrac{1}{2}\sigma^2 J_{kkk} + J_{kk}T + \tfrac{1}{2}\sigma^2 J_{kkk}]dt + J_{kk}\sigma dz \\ &= -H_k^0 dt + \sigma J_{kk} dz \end{aligned} \tag{10.22}$$

把上述分析总结起来，得到一个命题。

命题 10.1（Pontryagin 随机最大值原理）假设 $k(t), v^0(t)$ 是下列问题的解：

$$\max_v E_0 \int_0^N u(k(t), v(t))dt$$

S.t. $dk = T(k(t), v(t))dt + \sigma(k(t), v(t))dz$， $k(t)$ 给定

那么，存在一个共轭状态变量 $p(t)$ 使得对每一个 t， $t \in [0, N]$：

（1） v^0 使函数 $H(k, v, p, \partial p / \partial k)$ 最大化，其中：

$$H(k, v, p, \frac{\partial p}{\partial k}) = u(k, v) + pT(k, v) + \frac{1}{2}\sigma^2 \frac{\partial p}{\partial k}$$

（2）共轭状态函数 $p(t)$ 满足随机微分方程：

$$dp = -H_k^0 dt + \sigma(k, v^0) J_{kk} dz$$

（3）横截条件成立

$$p(k(N), N) = \frac{\partial J}{\partial k}(k(N), N, N) \geq 0$$

$$p(N)k(N) = 0$$

下面，我们进行一些推广。

10.2　一般化

第一，从数学上来说，从 Pontryagin 随机最大值原理得到的最优轨道是满足某些条件的两个随机微分方程的解。更明确地，我们把这些方程重新写在一起：

$$dk = H_p^0 dt + \sigma dz \tag{10.19}$$

$$dp = -H_k^0 dt + \sigma J_{kk} dz \tag{10.22}$$

$$k(0) \text{给定} \tag{10.23}$$

$$p(k(N), N) = 0 \tag{10.24}$$

第二，通过引进遗产函数 $B(k(t), t)$，容易把命题 10.1 推广。如果是这样的话，那么，

最大值问题将变成：

$$\max_v E_0[\int_0^N u(k(t),v(t))dt + B(k(N),N)]$$

约束于前面相同的条件（10.4），命题 10.1 成立，只是带有新的横截条件：

$$p(k(N),N) = \frac{\partial B}{\partial k}(k(N),N)$$

第三，推广（10.3）和（10.4）的另一个方向是引进折扣，我们不必重复前面的分析。设：

$$J(k(t),t,N) = \max_v E_t \int_t^N e^{-\rho s} u(k,v)ds$$

满足（10.4）且 $k(t)$ 给定。记：

$$W(k(t),t,N) = e^{\rho t} J(k(t),t,N)$$

和

$$J_t = \frac{d}{dt}\left\{e^{-\rho t}W\right\} = -\rho e^{-\rho t}W$$

因此，Hamilton-Jacobi-Bellman 方程转换成：

$$\rho W = \max_v [u(k(t),v(t)) + W_k T(k(t),v(t)) + \frac{1}{2}\sigma^2(k(t),v(t))W_{kk}]\text{。}$$

剩下的分析按照上面已作过的模式进行。

最后，考虑一般的多维依赖于时间的情形。它被描述为：

$$J(k(t),t,N) = \max_v E_t[\int_t^N u(k(s),v(s),s)ds + B(k(N),N)]$$

S.t. $dk_i(t) = T_i(k(t),v(t),t)dt + \sigma_i(k(t),v(t),t)dz_i(t), i = 1,2,...,n$， $k(t)$ 给定。

对于这个问题，Hamilton-Jacobi-Bellman 方程为：

$$\begin{aligned}
-J(k(t),t,N) &= \max_v [u(k(t),v(t),t) + J_k' T(k(t),v(t),t) \\
&\quad + \tfrac{1}{2} tr(J_{kk}\sigma(k(t),v(t),t)\sigma'(k(t),v(t),t))] \\
&= \max_v H(k(t),v(t),p(t),p_k(t),t) \\
&= H^0(k(t),p(t),p_k(t),t)
\end{aligned}$$

同前面一样，“ ′ ”表示转置。多维共轭状态的随机微分方程变成：

$$dp_i = H_{k_i}^0 dt + \sum_{j=1}^n J_{k_i k_j}\sigma_j dz_j$$

其中 dz_i 与 dz_j 不相关，对于 $i \neq j$ 。详细的分析可在 Bismut（1973）中找到。

10.3 带有约束的控制问题

现在，我们研究随机控制理论中更一般的问题。模型如下：

$$J(k(t),t,N)=\max_{v}E_t[\int_t^N u(k(s),v(s),s)ds+B(k(N),N)] \tag{10.25}$$

S.t.

$$dk_i(t)=T_i(k(t),v(t),t)dt+\sum_{j_i=1}^{n_i}\sigma_{ij_i}(k(t),v(t),t)dz_{ij_i}, \quad i=1,2,\ldots,n \tag{10.26}$$

$k(t)$给定。

$$g_Q(k(t),v(t),t)\geq 0,\ Q=1,2,\ldots,L. \tag{10.27}$$

换句话说，控制随机过程$v(t)$约束于一个由L个不等式组成的集合。注意，假设$d_{z_{ij_i}}$是Wiener过程，满足条件：$\mathrm{cov}(dz_{rj_r},dz_{sj_s})=\rho_{rj_rsj_s}dt$，其中$\rho_{rj_rsj_s}$为相关系数，它与$k(t)$和$v(t)$独立。

我们继续这个问题的分析，最后，如前面做的一样，把所得到的结果总结成一个命题。

根据Bellman最优原理，写出递归方程：

$$J(k(t),t,N)=\max_{v}E_t[\int_t^{t+\Delta t}u(k(s),v(s),s)ds+J(k(t+\Delta t),t+\Delta t,N)\text{。}$$

于是，如果逼近

$$E_t[\int_t^{t+\Delta t}u(k(s),v(s),s)ds=u(k(t),v(t),t)\Delta t+o(\Delta t)$$

是有效的，那么有：

$$J(k(t),t,N)=\max_{v}[u(k(t),v(t),t)\Delta t+E_tJ(k(t+\Delta t),t+\Delta t,N)+o(\Delta t)]\text{。}$$

令：

$$\Delta J(t)=J(k(t+\Delta t),t+\Delta t,N)-J(k(t),v(t),N)$$

因此，得到：

$$0=\max_{v}[u(k(t),v(t),t)\Delta t+E_t\Delta J(t)+o(\Delta t)] \tag{10.28}$$

假设J二次连续可微，利用Taylor定理，在$(k(t),t)$处展开$\Delta J(t)$：

$$\Delta J(t)=J_t\Delta t+J_k'\Delta k+\tfrac{1}{2}\Delta k' J_{kk}\Delta k+o(\Delta t) \tag{10.29}$$

在（10.29）中，取 $\Delta J(t)$ 的条件数学期望，我们有：

$$E_t\Delta J(t)=J_t\Delta t+E_t[J_k'\Delta k]+E_t[\tfrac{1}{2}\Delta k' J_{kk}\Delta k]+o(\Delta t) \tag{10.30}$$

因此，我们必须计算 $E_t[J_k'\Delta k]$ 和 $E_t[\frac{1}{2}\Delta k' J_{kk}\Delta k]$。下面，我们进行计算。从（10.26）有：

$$\Delta k_i=T_i\Delta t+\sum_{j_i=1}^{n_i}\sigma_{ij_i}\Delta z_{ij_i}+o(\Delta t) \tag{10.31}$$

取（10.31）的条件数学期望，得：

$$E_t\Delta k_i=E_t[T_i\Delta t+\sum_{j_i=1}^{n_i}\sigma_{ij_i}\Delta z_{ij_i}+o(\Delta t)_J=T_i\Delta t+o(\Delta t)$$

因为对所有的 i 和 j_i 有 $E_t\Delta z_{ij_i}=0$，利用以上结果，我们计算出：

$$\begin{aligned}E_t\{J_k'\Delta k\}&=E_t\left\{\sum_{i=1}^{n}J_{k_i}'\Delta k_i\right\}=\sum_{i=1}^{n}J_{k_i}'E_t(\Delta k_i)\\&=\sum_{i=1}^{n}J_{k_i}'T_i\Delta t+o(\Delta t)\end{aligned}$$

进一步：

$$\begin{aligned}E_t\{(\Delta k)'J_{kk}(\Delta k)\}&=E_t\left\{\sum_{r=1}^{n}\sum_{s=1}^{n}(\Delta k_r)J_{k_rk_s}(\Delta k_s)\right\}\\&=E_t\left\{\sum_{r=1}^{n}\sum_{s=1}^{n}\sum_{j_r=1}^{n_r}\sum_{j_s=1}^{n_s}J_{k_rk_s}\sigma_{rj_r}\sigma_{sj_s}\Delta z_{rj_r}\Delta z_{sj_s}+o(\Delta t)\right\}\\&=\sum_{r=1}^{n}\sum_{s=1}^{n}\sum_{j_r=1}^{n_r}\sum_{j_s=1}^{n_s}J_{k_rk_s}\sigma_{rj_r}\sigma_{sj_s}\rho_{rj_rsj_s}\Delta t+o(\Delta t)\end{aligned}$$

上式推导中用到了 $E_t\{\Delta z_{rj_r}\Delta z_{sj_s}\}=\rho_{rj_rsj_s}\Delta t+o(\Delta t)$。

收集上面的结果，（10.30）变为：

$$\begin{aligned}E_t\Delta J(t)=&J_t\Delta t+\sum_{i=1}^{n}J_{k_i}T_i\Delta t\\&+\frac{1}{2}\sum_{r=1}^{n}\sum_{s=1}^{n}\sum_{j_r=1}^{n_r}\sum_{j_s=1}^{n_s}J_{k_rk_s}\sigma_{rj_r}\sigma_{sj_s}\rho_{rj_rsj_s}\Delta t+o(\Delta t),\end{aligned}$$

同时有：

$$\begin{aligned}-J_t(k(t),v(t),N)\Delta t=\max_v\{&u(k(t),v(t),t)\Delta t\\&+\phi(k(t),v(t),t,N)\Delta t+o(\Delta t)\},\end{aligned}$$

其中ϕ定义为：

$$\phi(k(t),v(t),t,N)=\sum_{i=1}^{n}J_{k_i}T_i+\frac{1}{2}\sum_{r,s,j_r,j_s}J_{k_rk_s}\sigma_{rj_r}\sigma_{sj_s}\rho_{rj_rsj_s} \qquad (10.32)$$

但是要注意，现在ϕ是(k,v,t,N)的函数，因为每一个T_i和σ_{ij_i}都是(k,v,t,N)的函数。因此，我们可以用更简单的形式来代替上述最大值问题，即：

$$-J_t(k(t),t,N)\Delta t=\max_{v\in C_g(k(t),t)}[u(k(t),v(t),t)\Delta t + \phi(k(t),v(t),t,N)\Delta t+o(\Delta t)],$$

其约束集定义为：

$$C_g(k(t),t)=\{v\in R^m : g(k(t),v,t)\ge 0\}。$$

下面，把这些结果总结成一个命题。

命题 10.2（带有约束的随机最大值原理）设$J(k,t,N)$关于(k,t)二次连续可微，最优量$k(t)$和$v(t)$使得对所有r有：

$$E_r\int_r^{r+\Delta r}u(k(s),v(s),s)ds=u(k(r),v(r),r)\Delta r+o(\Delta r),$$

那么，在每一时刻t，最优控制$v(t)$是问题

$$\max_{v\in C_g(k(t),t)}[u(k(t),v,t)+\phi(k(t),v,t,N)]$$

的解，其中：

$$C_g(k(t),t)=\{v\in R^m : g_Q(k(t),v,t)\ge 0, Q=1,2,\dots,L\}$$

和$J(k,t,N)$必须满足下列偏微分方程：

$$-J_t(k,t,N)=\max_{v\in C_g}[u(k,v,t)+\phi(k,v,t,N)],$$

其中ϕ由（10.32）定义，边界条件为：

$J(k,N,N)=B(k,N,N)$，对所有k成立。

11. Bismut 方法

Bismut（1973）把由 Rockafellar（1970）发展的一般凸分析方法运用于最优随机控制问题。在这一节里，我们直观地介绍取材于 Bismut（1975）的 Bismut 方法。尽管 Bismut 方法的特殊应用要推迟到下一章，但在下面的最优随机控制的分析中，还是试图按照经济概念来

阐明其数学理论的方法和结果。特别要强调的是，最优随机控制中有两个重要概念在经济运用中是很有用的，它们是风险爱好和信息加工。

如前一样，问题的模型为：

$$\max E_0\int_0^T u(k,v,t,\omega)dt \tag{11.1}$$

S.t.

$$dk = f(k,v,t,\omega)dt + \sigma(k,v,t,\omega)dz \tag{11.2}$$

$k(0) = k_0$ 给定。

这里，u 可以代表瞬时效用或利润函数，k 表示资本存量，v 为投资决策，ω 为环境因素。为了简便，我们只研究一维情况。假设一个递增的信息系统可以从 $\sigma -$ 代数族 $\{F_t, t\in[0,T]\}$ 中得到，且信息中包含 k 和 z 的过去值，也包含 f 和 σ 的过去值。这里，同前面一样，z 为 Wiener 过程。在这些假设下，对于任意的决策 v，资本增量的期望均值和方差是已知的。设 p_t 表示在 t 时刻资本的边际值，由下式给出：

$$p_t = \frac{\partial E_t}{\partial k}\int_t^T u(k,v,s,\omega)ds \tag{11.3}$$

即 p_t 是效用函数的条件数学期望关于 k 的偏导数，式中的 v 为最优策略。Bismut（1973,p.387）和 Bismut（1975,p.242）假设 p_t 可以写成：

$$p_t = p_0 + \int_0^t \dot{p}_s ds + \int_0^t H_s dz_s + M_t \tag{11.4}$$

其中 $\dot{p}_s$ 是 p 的无穷小期望增长率，H_s 是 p 关于 z 的无穷小条件方差。M 为可预测项，$M_0 = 0$。它是一个给定随机变量在 t 时刻的最好估计，且它与 z 独立。因此，p_t 的无穷小增量在每一个时刻的条件期望值都是 0。

这个分解式的意思是 p_t 可以分解成四项之和。第一项为 p_0，第二项给出 p_t 在每一时刻的期望无穷小增量，第三项把不确定因素并入一个积累过程中，最后一项收集了关于环境因素方面的综合信息 M_t。

现在，定义 $\widetilde{H}$ 为：

$$\widetilde{H} = u(k,v,t,\omega) + pf(k,v,s,\omega) + H\sigma(k,v,s,\omega) \tag{11.5}$$

Bismut（1973,p.401）证明，对于最优控制v，下列关系成立：

$$\frac{\partial \tilde{H}}{\partial v}=0 \tag{11.6}$$

$$dp=-\frac{\partial \tilde{H}}{\partial k}dt+Hdz+dM \tag{11.7}$$

$$p_T=0 \tag{11.8}$$

如前节一样，dk 由（11.2）给出，其中v是最优控制。需要指出，方程（11.6）—（11.8）和（11.2）非常类似于命题 10.1 中所述的方程。更明确的是，方程（11.6）来自v的最大性，而方程（10.19）和代入最大元v的（11.2）是相同的。方程（10.24）和（11.8）一样，它们表示横截条件。最后，（10.22）和（11.7）是唯一一对不同的方程。注意，（11.7）中的 Bismut 随机变量H对应于（10.22）中的随机变量$J_{kk}\sigma$，而（11.7）中的dM 项在（10.22）中没有对应的项。粗略来说，原变量与对偶变量的 Bismut 对应如下：

$$f\rightarrow p \tag{11.9}$$

$$\sigma\rightarrow H \tag{11.10}$$

$$F_t\rightarrow M_t \tag{11.11}$$

为了揭示它们的经济意义，让我们来解释上面的各个变量。首先（11.5）中的$\tilde{H}$等于瞬时效用或（利润）加上（以边际期望值计的）期望资本无穷小增量的和，减去相联系的投资策略的风险（以费用计算）。对于风险的瞬时态度用H表示。如果它是一个风险爱好者，那么H是正的。如果它是一个风险厌恶者，那么H是负的。下面，我们解释dp或者$-dp$，即资本边际值的条件期望贬值率。从（11.7），我们看到$-dp$等于对效用或（利润）的资本贡献加上能提高资本存量增量的期望值的资本贡献之和，减去对增加资本存量增量的条件标准差（以风险价值计算）的资本贡献，再减去其他两项：Hdz 和dM 。为了解释Hdz，从（11.2）得

$$dz=(1/\sigma)(dk-fdt) \tag{11.12}$$

乘以H，得：

$$dz=(H/\sigma)(dk-fdt)\text{。} \tag{11.13}$$

于是，（11.7）中的Hdz是在资本边际值的演变过程中的修正项，它以p来估计dk和$E(dk)$的差，其中$E(dk)=fdt$。最后一项dM 表示在长期不确定性预测中的变化，它可能增加也可能减少资本值。直观地说，M 体现不含在z的过去值中的信息。z包含所有出现在积累

过程中的短期不确定性，而 M 是长期不确定性的预测。

至此，我们结束 Bismut 方法的讨论。

12. 跳跃过程

在这一节中，我们将建立跳跃过程的一般 $It\hat{o}$ 公式。跳跃过程是由一个 Possion 过程所建立的模型，用来描述随机事件的发生。当 Possion 事件发生时，状态变量就出现一个跳跃，这个跳跃按预先指定的密度函数分布。在一般 $It\hat{o}$ 公式建立后，我们将按与扩散过程情形一样的方法发展最大值原理。这里的讨论仍然是直观性的。

12.1 一般 $It\hat{o}$ 公式

我们将建立含有 Possion 过程和 Wiener 过程的混合情况下的一般 $It\hat{o}$ 公式。

考虑：

$$dx(t) = f(t,x)dt + \sigma(x,t)dz(t) + g(t,x)dq(t) \tag{12.1}$$

这里，对于 $R = (-\infty, \infty)$，

$$f(t,x):[0,\infty)\times R \to R, \sigma(x,t):[0,\infty)\times R \to R，$$

$$g(t,x):[0,\infty)\times R \to R。$$

同时，$\{z(t)\}_{t=0}^{\infty}$ 是标准 Wiener 过程，$\{q(t)\}_{t=0}^{\infty}$ 是 Possion 过程。为简便起见，假设它们相互独立。

用 $\lambda\Delta t + o(\Delta t)$ 表示 $q(t)$ 在区间 $(t, t+\Delta t)$ 内发生一次跳跃的概率，且跳跃的幅度 A 是一个密度函数为 $p(a)$ 的随机变量，即 $p(a)da$ 是跳幅包含在 $(a, a+da)$ 内的概率（忽略 da 的高阶项）。假设 $q(t)$ 在区间 $(t, t+\Delta t)$ 中发生一次以上跳跃的概率为 $o(\Delta t)$，于是 $q(t)$ 在 $(t, t+\Delta t)$ 上为常数的概率是 $1-\lambda\Delta t + o(\Delta t)$，这样一维情形的一般 $It\hat{o}$ 公式就可以叙述了。严格的陈述和证明参见 Kushner（1967,p.18）。Gihman & Skorohod（1972,p.263）有比 Kushner 更完整的一般 $It\hat{o}$ 公式的处理。

命题 12.1（一般 $It\hat{o}$ 公式）设 $F(t,x)$ 关于 (t,x) 二次连续可微，存在一个有界闭区间 I，使得 $\{a \mid p(a) > 0\} \subset I$。令 $\Delta F = F(t+\Delta t, x+\Delta x) - F(t,x)$，$E_t\Delta F$ 表示 ΔF 关于 $x(t) = x$ 的条件期望。那么：

$$E_t\Delta F=\{F_t(t,x)+F_x(t,x)f(t,x)+\tfrac{1}{2}F_{xx}(t,x)\sigma^2(t,x)$$
$$+\lambda(\int_{a\in I}F(t,x+g(t,x)a)-F(t,x)]p(a)da)\}\Delta t+o(\Delta t)\text{。}$$

为了建立这个定理，需要指出：

$$E_t\Delta F=(E_t^*\Delta F)\lambda\Delta t+(1-\lambda\Delta t)E_t^{**}\Delta F+o(\Delta t) \qquad (12.2)$$

其中 E_t^* 表示 Possion 事件发生时的条件期望，E_t^{**} 表示 Possion 事件不发生时的条件期望。方程（12.2）可以写成：

$$E_t\Delta F=E_t^{**}\Delta F+\lambda\Delta t(E_t^*\Delta F-E_t^{**}\Delta F)+o(\Delta t) \qquad (12.3)$$

方程（12.3）中的第一项是在 Possion 事件不发生的条件下，函数 F 改变量的条件期望。首先，当 $\Delta t\to 0$ 时，$E_t^{**}\Delta F\to 0$。因此，为了计算方程（12.3）中的右边第二项至 $o(\Delta t)$，我们只需观察 $E_t^*\Delta F$。经计算，得：

$$E_t^{**}\Delta F=F_t\Delta t+F_xE_t^{**}\Delta x+\frac{1}{2}F_{xx}E_t^{**}\Delta x^2$$
$$+\lambda\Delta t(\int_{a\in I}[F(t,x)+g(t,x)a)-F(t,x)]p(a)da)+o(\Delta t) \qquad (12.4)$$

因为当 $\Delta t\to 0$ 时，$\mathrm{E}_t^{**}\Delta\mathrm{F}\to 0$，所以 $\lambda\Delta t\mathrm{E}_t^{**}\Delta\mathrm{F}=\mathrm{o}(\Delta t)$。而且：

$$\lambda\Delta tE_t^*\Delta F=\lambda\Delta tE_t^*\int_{a\in I}[F(t+\Delta t,x+f\Delta t+\sigma\Delta z+ga+o(\Delta t))$$
$$-F(t,x)]p(a)da$$
$$=\lambda\Delta t\int_{a\in I}[F(t,x+ga)-F(t,x)]p(a)da+o(\Delta t))\text{，}$$

这是因为 Δz 是均值为 0 方差为 Δt 的正态分布，所以有当 $\Delta t\to 0$ 时，$f\Delta t+\sigma\Delta z\to 0$。

因此，公式（12.4）可简化为：

$$E_t\Delta F=F_t\Delta t+F_xf\Delta t+\frac{1}{2}F_{xx}\sigma^2\Delta t$$
$$+\lambda\Delta t(\int_{a\in I}(F(t,x+ga)-F(t,x))p(a)da)+o(\Delta t) \qquad (12.5)$$

这就是一般 $It\hat{o}$ 公式。

12.2 跳跃过程的最大值原理

考虑问题：

$$J(k(t),t,N)=\max_{v(\cdot)}E_t[\int_t^N u(k(s),v(s),s)ds+B(k(N),N)] \tag{12.6}$$

S.t.

$$\begin{aligned}dk(s)&=T(k(s),v(s),s)ds+\sigma(k(s),v(s),s)dz(s)\\&\quad+g(k(s),v(s),s)dq(s)\end{aligned} \tag{12.7}$$

其中所有记号同第 10 节一样，除了加上了一个跳跃成分 $g(k(s),v(s),s)dq(s)$ 之外。这里，跳幅 A 按密度 $p(a)$ 分布，且如在一般 $It\hat{o}$ 公式所陈述的，与 $\{z(t)\}_{t=0}^{\infty}$ 相互独立。作这个假设是为了使问题简化。

让我们按照第 10 节的过程，结合一般 $It\hat{o}$ 公式来揭示随机最大值原理。

对于在 t 时刻给定的 (k,v,t)，计算得：

$$\begin{aligned}&E_t[\int_t^{t+\Delta t}uds+J(k(t+\Delta t),t+\Delta t,N)-J(k(t),t,N)]\\&=u(k,v,t)\Delta t+J_t(k,t,N)\Delta t+J_k(k,t,N)T(k,v,t)\Delta t\\&\quad+\frac{1}{2}J_{kk}(k,t,N)\sigma^2(k,v,t)\Delta t\\&\quad+\lambda\Delta t\int_{a\in I}[J(k+g(k,v,t)a,t,N)-J(k,t,N)]p(a)da+o(\Delta t)\end{aligned} \tag{12.8}$$

方程(12.8)是对 $E_t\Delta J$ 应用一般 $It\hat{o}$ 公式得到的。把（12.8）的右边项除以 Δt 后记为 $\phi(k,v,t,N)$。然后，我们有下述命题。

命题 12.2（最大值原理）假设 J 满足一般 $It\hat{o}$ 公式的条件。那么最优控制函数 $v^0(k,t,N)$ 通过求解

$$\max_v \phi(k,v,t,N)$$

得到。且 J 由偏微分方程

$$0=\max_v \phi(k,v,t,N)=\phi(k,v^0(k,t,N),t,N)$$

确定，其边界条件为：$J(k,N,N)=B(k,N)$。

很明显，在这种启发式讨论的难度水平上，最大值原理可以直接推广到多维情况、控制变量 v 受约束的情况以及与 Wiener 过程和 Possion 过程相关的情况，等等。当然，严格的推广将需要许多详细的数学知识。

13. 最优停时和自由边界问题

在第一章第 8 节中，我们给读者介绍了一些基本的最优停时的概念。在那里，我们考虑一个随机变量序列$Y_1, Y_2, \ldots$和它们对应的报酬 $X_1, X_2, \ldots$。在这些随机变量序列上加上一些条件，我们就能够得到最优停时规则的存在性。本节里，我们依靠 Van Moerbeke（1974）来处理连续时间的情形。这里，代替随机变量序列，我们考虑从 $z_0(\omega)=0$ 出发的 Wiener 过程。报酬函数记为 $g(z,t)$，意思是在t时刻，当事态是 $z=z_t(\omega)$ 时，报酬的值为 $g(z,t)$。那么经过随机时间周期 $\tau=\tau(\omega)$ 后的平均报酬为：

$$Eg(z+z_\tau, t+\tau) \tag{13.1}$$

其中 g 的自变量，即 $(z+z_\tau, t+\tau)$，表示从 (z,t) 开始的 Wiener 过程的空间—时间（坐标）。换句话说，假定我们在t时刻，事态为 z 的情况下开始玩一个游戏且玩了一段随机时间 τ，$\tau \le T-t$。对应这段随机时间的事态由所讨论的 Wiener 过程确定，用 z_τ 表示。所以，方程（13.1）估计了这个游戏的期望报酬。我们的主要兴趣是有限区间$[0,T]$，其中$T<\infty$，$t\in[0,T]$。假定 g 和它们的所有偏导数都连续，且当$t\to T$ 时，有极限。在$t=T$ 处，允许不连续，但这时要求$h(z)\equiv g(z,T)-g(z,T-)$除了几个孤立跳跃外，要无穷次可微，其中 $g(z,T-)$ 表示 $g(t,z)$ 在 $t=T$ 处的左极限。函数 h 叫做最终获利。不失一般性，我们假设 $h\ge 0$。这个假设不是限制性的，因为如果在给定的区间中有 $h(z)<0$，那么在碰到终点 $t=T$ 之前，会更早地停止。

定义了函数 g 和 h 后，我们陈述关于它们的假设条件。加在 g 上的条件是一个叫做 Tychonov 条件的增长条件。如果函数：

$$\begin{aligned} & g, \frac{\partial g}{\partial t}, \frac{\partial g}{\partial z}, \frac{\partial^2 g}{\partial z^2}, \frac{\partial^2 g}{\partial z \partial t}, \frac{\partial^3 g}{\partial z^3}, \\ & h, \frac{\partial u}{\partial z}, \frac{\partial^2 h}{\partial z^2}, \frac{\partial^3 h}{\partial z^3}, \end{aligned}$$

当$|z|\to\infty$时，以$e^{o(z^2)}$为界，在任意有限区间$[t,T]$上一致成立。那么称 g 满足 Tychonov 条件。

假定我们从时刻t事态 z 出发开始玩游戏，进行到时刻$T<\infty$。在这种情况下，最优报酬 $\hat{g}(z,t)$ 通过（13.1）式对所有停时 $\tau, \tau\le T-t$，求最大值得到。达到最大值的停时 τ 叫

做最优策略。明显地，最优报酬函数 $\hat{g}$ 对我们很重要，我们将以某种方式刻画它。利用过分函数的概念可以做到。一个下方有界函数 f 称为在开域 $D \subset R^2$ 内是过分的，如果：

（1）对每一个不超过由 D 首次离开的时刻 τ_D 的停时 τ，有 $Ef(z+z_\tau, t+\tau) \le f(z,t)$；

（2）对每一个满足 $\tau_n < \tau_D$，$P[\tau_n \to 0]=1$ 的停时序列 τ_n，有：

$$Ef(z+z_{\tau_n}, t+\tau_n) \to f(z,t)。$$

注意，如果 f 充分可微，那么它在开域 D 内过分与 $\frac{\partial f}{\partial t}+\frac{1}{2}\frac{\partial^2 f}{\partial z^2} \le 0$，对所有 $(z,t)\in D$ 是同一回事。利用过分性定义，我们把 $\hat{g}$ 刻画成超过 g 的最小的过分函数。Tychonov 条件蕴含着 $\hat{g}$ 是有限和连续的。

下面我们区分两个区域：持续区域 C，这里 $\hat{g} > g$，它意味着继续玩会有收益；停止区域 S，这里 $\hat{g} = g$，它意味着退出是最好的。由于 $\hat{g}$ 是连续的，因此 C 为开域。我们假定区域 C 有一个连续可微的边界 $z=s(t)$，除了几个 $|ds/dt|$ 可能爆炸的独立点之外。分离这两个区域的边界叫做最优停时边界。加在 g 上的 Tychonov 条件帮助我们得到结论：最优策略是，只要你保持在 C 内，就继续进行，当你碰到最优边界时就停止。令 τ_0 表示这个碰到最优边界的时刻，那么

$$\hat{g}(z,t) = Eg(z+z_{\tau_0}, t+\tau_0) \qquad （13.2）$$

因此，我们的目的是找到最优停时边界。

此时，指出我们的问题与偏微分方程理论之间存在绝妙的相互作用是适当的。尽管我们马上就到达找到最优停时边界的目标，但仍鼓励有兴趣的读者查看 Van Moerbeke（1974）和他引的许多参考资料。说了这些后，我们指出，寻找 $\hat{g}$ 和最优策略的问题可以通过转成一个热方程的自由边界问题来解决。从（13.2）式得出：在连续区域 C 中，$\hat{g}$ 是抛物的。这意味着：首先，$\hat{g}$ 是过分的；其次，$E\hat{g}(z+z_{\tau_U}, t+\tau_u) = \hat{g}(z,t)$，其中 τ_U 是从任意一个具有紧闭包的开集 U 首次离开的时刻。抛物函数满足向后热方程；反过来，向后热方程的下方有界解都是抛物函数。因此，

$$\frac{\partial \hat{g}}{\partial t}+\frac{1}{2}\frac{\partial^2 \hat{g}}{\partial z^2}=0 \qquad \text{在}C\text{内，} \tag{13.3}$$

$$\hat{g}=g \qquad \text{在}C\text{的边界，} \tag{13.4}$$

$$\hat{g}(z,T)=g(z,T) \tag{13.5}$$

进一步，因为早先做了连续区域C的边界连续可微的假设，所以$\hat{g}$的最优性意味着

$$\frac{\partial \hat{g}}{\partial z}(s(t),t)=\frac{\partial g}{\partial z}(s(t),t) \tag{13.6}$$

同时，对于$(y,u)\in C$，当$(y,u)\to(s(t),t)$时，有

$$\frac{\partial \hat{g}}{\partial z}(y,u)=\frac{\partial \hat{g}}{\partial z}(s(t),t) \tag{13.7}$$

在$|ds/dt|<\infty$的点$(s(t),t)$，（13.6）和（13.7）式均成立。方程（13.5）、（13.6）和（13.7）叫做光滑适合（fit）方程。

让我们思考一下。方程（13.3）—（13.7）描述的是一个具有两个边界条件的初边值问题，这意味着除非我们选择保持边界自由，否则我们的问题是超定的。假定我们保持边界自由，那么，方程（13.3）—（13.7）能决定边界$s(t)$和最优报酬$\hat{g}$似乎是合理的。

我们通过叙述下面的定理结束这一节。

定理 13.1 设C是一个开集，在$t\le T<\infty$中有连续可微的边界曲线$z=s(t)$，除了有限个ds/dt可能爆炸的孤立点之外。又设 Tychonov 型函数u满足：

$$\frac{\partial u}{\partial t}+\frac{1}{2}\frac{\partial^2 u}{\partial z^2}=0 \qquad \text{在}C\text{内，}$$
$$u=g \qquad \text{在}(z,t)=(s(t),t)\text{，}$$
$$u(z,T)=g(z,T),$$

$$\frac{\partial u}{\partial z}=\frac{\partial g}{\partial z} \qquad \text{在}(z,t)=(s(t),t),\text{如果}\left|\frac{ds}{dt}\right|<\infty\text{，}$$
$$u>g\text{在}C\text{内且}u=g\text{在其他地方,}$$
$$H\equiv\frac{\partial g}{\partial t}+\frac{1}{2}\frac{\partial^2 g}{\partial z^2}\le 0 \qquad \text{在}C\text{的余集内。}$$

那么u实际上是$\hat{g}$，$s(t)$是最优停时边界。

这个定理的证明参见 Van Moerbeke（1974），以上的简要分析可以作为连续时间的最优停时问题的引言。

14. 各种应用和习题

（1） 假设$\{X(t),t\in T\}$和$\{Y(t),t\in T\}$是定义在同一概率空间上的实值随机过程，且$E(X(t))<\infty$，$E(Y(t))<\infty$，对所有$t\in T$。这两个过程的协方差函数，用$r_{XY}(s,t)$表示，定义为：

$$\begin{aligned}r_{XY}(s,t)&=\operatorname{cov}(X(s),Y(t))\\&=E([X(s)-E(X(s))][Y(t)-E(Y(t))])\end{aligned}$$

这个定义推广了第一章的（4.12）式。当两个过程相同时，协方差函数叫做自协方差函数，用$r_X(s,t)$表示。证明下列两个简单事实：

$$\begin{aligned}&VarX(t)=r_X(t,t)\quad,\quad 对于t\in T，\\&r_X(s,t)=r_X(t,s)\quad,\quad 对于s,t\in T。\end{aligned}$$

（2） 称使得对所有$t\in T$、$E(X(t))<\infty$的随机过程$\{X(t),t\in T\}$在t时刻均方连续，如果当$h\to 0$时，有$E(X(t+h)-X(t))^2\to 0$。令$\mu_X(t)=E(X(t))$，假设 $\mu_X(t),t\in T$，关于t连续，$r_X(s,t),s,t\in T$关于s,t联合连续。证明在这些假设下，$\{X(t),t\in T\}$均方连续。

（3） 考虑本章的方程（2.11），为了方便，我们把它改写成：

$$x(t+\Delta t)-x(t)=f(t,x(t))\Delta t+\sigma(t,x(t))[z(t+\Delta t)-z(t)]+o(\Delta t)。$$

证明：

$$\begin{aligned}&E(x(t+\Delta t)-x(t))=f(t,x(t))\Delta t+o(\Delta t)\\&Var(x(t+\Delta t)-x(t))=\sigma^2(t,x(t))E(z(t+\Delta t)-z(t))^2+o(\Delta t)\\&\qquad\qquad=\sigma^2(t,x(t))\Delta t+o(\Delta t)。\end{aligned}$$

（4） 考虑具有后向差分的方程（2.11），可以写成：

$$x(t)-x(t-\Delta t)=f(t,x(t))\Delta t+\sigma(t,x(t))(z(t)-z(t-\Delta t)]+o(\Delta t)$$

假定函数f和σ连续。证明：

$$\begin{aligned}E(x(t)-x(t-\Delta t))&=f(t-\Delta t,x(t-\Delta t))\Delta t\\&\quad+\sigma_x(t-\Delta t,x(t-\Delta t))\sigma(t-\Delta t,x(t-\Delta t))\Delta t+o(\Delta t)\\Var(x(t)-x(t-\Delta t))&=\sigma^2(t,x(t))\Delta t+o(\Delta t)\end{aligned}$$

比较这个练习和前一个练习的结果得出：过程增量的均值依赖于所使用的差分类型，而过程增量的方差不受它的影响。

（5） 考虑区间$[s,t]\subset[0,T]$的一个划分：

$$s = t_0 < t_1 < \ldots < t_n = t,$$
$$\max_{0<i<n-1} |t_{i+1} - t_i| \le \varepsilon, [t_i, t_{i+1}) \subset [s,t]\text{,}$$

其中$\varepsilon > 0$且任意小。假设$Z(t,\omega)$是具有单位方差的 Wiener 过程，又当$\varepsilon \to 0$时，A和B满足：

$$E\left|A - \sum_i z(t_i)[z(t_{i+1}) - z(t_i)]\right|^2 \to 0$$
$$E\left|B - \sum_i z(t_{i+1})[z(t_{i+1}) - z(t_i)]\right|^2 \to 0$$

证明：$B - A = t - s$。

（6） 假设$z(t)$是具有单位方差的 Wiener 过程。考虑随机积分$\int_s^t z(u)dz(u)$。利用本章第 3 节和第 15 节的$It\hat{o}$积分和 Stratonovich 积分的定义得到：首先，当积分理解成$It\hat{o}$积分时，有：

$$\int_s^t z(u)dz(u) = \frac{1}{2}[z^2(t) - z^2(s)] - \frac{1}{2}(t-s)\text{,}$$

其次，当积分理解成 Stratonovich 积分时，有：

$$\int_s^t z(u)dz(u) = \frac{1}{2}[z^2(t) - z^2(s)]$$

特别地，Stratonovich 积分满足通常微积分中的分布积分公式，而$It\hat{o}$积分不满足。注意，由分布积分公式得到：

$$\int_s^t z(u)dz(u) = z(t)z(t) - z(s)z(s) - \int_s^t z(u)dz(u)$$

或者等价地：

$$\int_s^t z(u)dz(u) = \frac{1}{2}[z^2(t) - z^2(s)]$$

参看 Stratonovich（1966，p.365）。

（7） 假设：

$$dx_1(t) = x_1^2(t)dt + dz(t)$$
$$dx_2(t) = x_2 dz(t)$$

利用$It\hat{o}$引理，计算下列每一种情况的$dy(t)$：

（a）$y(t) = u(t, x_1, x_1) = x_1(t)x_2(t)$

（b） $y(t)=u(t,x_1,x_2)=t[x_1(t)x_2(t)]$

（c） $y(t)=u(t,x_1,x_2)=z(t)[x_1(t)x_2(t)]$

（8） 考虑带初值的随机微分方程组：

$$\begin{aligned}&x_1(0)=x_2(0)=z(0)=0\\&dx_1(t)=dz(t)\\&dx_2(t)=x_1dz(t)\end{aligned}$$

其中 $z(t)$ 为具有单位方差的 Wiener 过程。利用 $It\hat{o}$ 引理证明上述方程组的解为：

$$\begin{aligned}&x_1(t)=z(t)\\&x_2(t)=\int_0^t z(t)dz(t)\end{aligned}$$

（9） 求解下列两个一阶随机微分方程：

（a） $dx(t)=a_0x(t)dt+\sigma_0dz(t)$

其中 a_0,σ_0 为非负常数。

（b） $dx(t)=a(t)x(t)dt+\sigma(t)dz(t)$

其中 $a(t),\sigma(t)$ 为任意的非随机函数。

（10） 考虑随机微分方程：

$$dx(t)=-x(t)dt+\sigma(t)x(t)dz(t)$$

其中 $\sigma(t)$ 为任意的非随机函数。请找到发现 $0-$ 平衡稳定性的充分条件。

（11） 考虑二阶随机系统：

$$\begin{aligned}&dx_1(t)=x_2(t)dt\\&dx_2(t)=-x_1(t)dt-\sigma x_1(t)dz(t)\end{aligned}$$

Kozin & Prodromou（1971）研究了该系统的样本稳定性，我们鼓励有兴趣的读者查看他们的文章。与 $\sigma=0$ 的非随机系统的直接稳定性分析不同，上述系统的随机稳定分析很有技巧性。

（12） 考虑有线性二次目标函数的随机控制问题：

$$-W(x(t))=\min_v E_t\int_{s=t}^{\infty}e^{-\rho(s-t)}\left\{a(x(s))^2+b(v(s))^2\right\}ds$$

$$\text{S.t.}\quad dx(t)=v(t)dt+\sigma x(t)dz(t)$$

其中 $\sigma > 0, a > 0, b > 0, \rho > 0$。请找到一个最优解。注意，这个问题和它的推广之一将在下一章中讨论。

15. 进一步的注释和参考

许多经济学家从研究计量经济学中熟悉了离散时间随机模型的分析。本章，我们简要介绍离散时间的情况，目的是促进连续时间情况的分析。与离散时间随机模型有关的超过计量经济学入门水平的问题，如 Intriligator（1978），在 Chow（1975）、Aoki（1976）和 Bertseka &（1978）中也有较详细的介绍。Sargent（1979）也有一章是讨论线性随机差分方程的。连续时间的不确定模型在 Åström（1970）、Balakrishnan（1973）、Friedman（1975）、Soong（1973）、Tsokos 和 Padegett（1974）及 Gihman 和 Skorohod（1969, 1972）中有介绍。

$It\hat{o}$ 随机微分方程出现在 $It\hat{o}$（1946,1950）中，后来在 $It\hat{o}$（1951b）中有较详细的研究。在他 1964 年的文章中，$It\hat{o}$ 发现某些数学问题，如由 Kolmogorov 提出和后来 Feller 也提出的与扩散过程有关的偏微分方程，能够通过解随机微分方程来研究。后来在他的 1951 研究报告中，$It\hat{o}$ 推广常微分方程的 Picard 迭代法，建立 $It\hat{o}$ 随机微分方程的存在性和唯一性定理。

值得一提的是 $It\hat{o}$ 随机微分方程是许多随机方程的一种。Syski（1967）考虑形如

$$\frac{dx}{dt} = f(x(t), y(t), t) \quad , \ t \in T$$

的基本随机微分方程系统，初始条件为 $x(t_0) = x_0$，其中 f, x, y 可以是合适维数的向量。他把这种方程分成三个基本类型：

（1） 随机初始条件；

（2） 随机非齐次部分；

（3） 随机系数。

$It\hat{o}$ 随机微分方程是一类特殊的随机微分方程，而且有几个理由说明它是重要的一种。首先，条件均值和条件方差作为函数是 $It\hat{o}$ 方程的充分统计量。因此，对 $It\hat{o}$ 方程来说，条件均值和条件方差函数的计算完全决定了整个过程。这与均值和方差是正态分布的充分统计量的统计事实相类似。

第二，$It\hat{o}$ 方程展示了非预期性质，它在建立不确定性的模型中很有用。换句话说，如果系统中的真正不确定源是 dz，且要求微分方程无需敏锐的洞察力，那么下一时刻状态变量 x 的演变应该只依赖于那个时刻的不确定演变过程。$It\hat{o}$ 方程与这种无需洞察力相一致。

第三，$It\hat{o}$ 方程当它们的解存在时有很好的性质。第 7 节中描述了这些性质。这里只要说 Markov 性和扩散性就够了，关于它们的理论已充分发展了。

最后，尽管经济学家刚刚发现 $It\hat{o}$ 方程的用处，但值得注意的是这个方程在工程领域，特别是控制、滤波和通讯理论中，已有重要的应用。

第 3 节的讨论集中在 $It\hat{o}$ 积分。基本的参考资料是 $It\hat{o}$ (1949,1951b)。Doob（1953）较

详细地介绍了 $It\hat{o}$ 原始思想的一些延伸。为了了解关于 Wiener 积分和 $It\hat{o}$ 一般化之间的联系，也可参看 Doob（1966）。近来，Åström（1970）和 Arnold（1974）介绍了 $It\hat{o}$ 积分的简化讨论。随机积分的详细内容在 Mckean（1969）中给出。一种不同的随机积分的方法和一般的随机微积分可在 Mcshane（1974）中找到。Wong & Zakai（1965）讨论了通常积分到随机积分的收敛性，我们注意到第 3 节是随机积分的简单入门，其目的在于激发和提供经典 $It\hat{o}$ 随机积分的定义。对于兴趣在于随机积分研究的读者，我们建议他阅读 Metivier & Pellaumail（1980）、Mckean（1969）和 Kussmaul（1977）。这些书是按照 $It\hat{o}$ 原型发展其理论的。关于平方可积鞅的随机积分由 Doob（1953）中提出，Kunita & Watanabe（1967）中对它进行了推广。Kunita（1970）介绍关于 $Hibert-$ 值过程的随机积分，Meyer（1976）作了进一步的推广，介绍关于一类特殊的 $Banach-$ 值过程的随机积分。自然地，随机积分概念的各种推广也牵扯到 $It\hat{o}$ 引理和随机微分方程的研究。Metivier & Pellaumail（1980）提供了这些推广的详细参考。

在本书中，因为我们选择了分析 $It\hat{o}$ 随机微分方程，所以我们使用 $It\hat{o}$ 积分。然而，应该告知读者另一种随机积分的方法，它由 Stratonovich（1966）中提出。为了展示 $It\hat{o}$ 积分和 Stratonovich 积分之间的差别，我们考虑一个简单特殊的情形：

$$\int_s^t z(u)dz(u) \tag{15.1}$$

其中 $z(t)$ 为具有单位方差的 Wiener 过程。在这种特殊情形下，对于形如（3.8）的划分，这一章的第 3 节，特别是引理 3.4，表明 $It\hat{o}$ 积分有定义，用 A 表示，那么当 $\varepsilon \to 0$ 时，

$$E\left|A-\sum_i z(t_i)[z(t_{i+1})-z(t_i)]\right|^2 \to 0 \tag{15.2}$$

回顾一下，（15.2）来自本章（3.11）得两个方程。在我们这种特殊情形下，Stratonovich（1966,p.363）定义他的随机积分为：

$$E\left|S-\sum_i \frac{z(t_i)+z(t_{i+1})}{2}[z(t_{i+1})-z(t_i)]\right|^2 \to 0 \tag{15.3}$$

其中 S 表示 Stratonovich 随机积分。注意 Stratonovich 随机积分是在（3.11）中定义的积分 A 和 B 的特定线性组合。更明确地，$S=\frac{1}{2}A+\frac{1}{2}B$，其中 A 是 $It\hat{o}$ 积分。A 和 B 与（3.11）中的一样。$It\hat{o}$ 积分和 Stratonovich 积分没有很大的不同。实际上，Stratonovich 发展了利用其他东西来表示一个积分的公式，并且这个公式不复杂。参看 Stratonovich（1966,p.365）。然而，因为随机微积分的本性，这两种积分有不同的性质。$It\hat{o}$ 积分和 $It\hat{o}$ 微分方程保持了状态模型的直观思想。同时，$It\hat{o}$ 积分有有用的性质，如它是鞅，它保持（2.12）中的 dx 的期望是 fdt 和 dx 的条件方差是 $\sigma^2 dt$ 的解释。$It\hat{o}$ 积分的主要缺点是它不能保持通常微积分中的微分规则，而以 $It\hat{o}$ 引理展现。Stratonovich 积分保持通常微积分中的计算规则，但没

有刚才所说的 $It\hat{o}$ 积分的优点。两种积分之间深入详细的关系，参看 Meyer（1976）。

迄今为止，在经济学和金融学的应用中所有随机微积分都使用 $It\hat{o}$ 积分，因为有关的 $It\hat{o}$ 微分方程提供了一个有意义的不确定模型，如早先解释过的。因此，$It\hat{o}$ 引理对于计算复合随机函数的随机微分是重要的。换句话说，对于由 Wiener 过程的随机积分表示的过程，$It\hat{o}$ 引理是一个变量变换公式。注意，$It\hat{o}$ 引理首先出现在 $It\hat{o}$（1951a）中，后来在 $It\hat{o}$（1961）中。这里我们重复我们早先所说的，即 $It\hat{o}$ 引理和 $It\hat{o}$ 积分是 $It\hat{o}$ 在研究偏微分方程和扩散过程的关系时发现的。这个原始的研究也导致随机微分方程的发展。就是 $It\hat{o}$ 和后来的 Doob、Gihman 以及 Skorohod 等人建立了随机微分方程领域。

现在，我们简单地讨论求解一阶随机微分方程的一种方法，正如 Gihman & Skorohod（1972,pp.33-39）中所介绍的一样。考虑 $[0,T]$ 上的方程：

$$dx(t) = f(t,x(t))dt + \sigma(t,x(t))dz(t) \tag{15.4}$$

注意，它可以写成积分形式：

$$x(t) = x(0) + \int_0^t f(s,x(s))ds + \int_0^t \sigma(s,x(s))dz(s) \tag{15.5}$$

其中 $x(0)$ 是初始条件。这种方法的思想是求（15.4）的合适的变换，使得（15.5）中的右边有方便的形式，即未知函数不出现在（15.5）式的右边。我们用一个例子来展示这个方法。读者可参照 Gihman & Skorohod（1972,pp.33-39），那里有详细的数学内容和更多的例子。

考虑方程：

$$dx(t) = a_0 x(t)dt + \sigma_0 x(t)dz \tag{15.6}$$

其中 a_0, σ_0 为常数。为了解这个方程，考虑代换 $y = \log x$ 和使用 $It\hat{o}$ 引理，得：

$$\begin{aligned} dy &= \frac{1}{x}dx + \frac{1}{2}(-\frac{1}{x^2})(dx)^2 = \frac{1}{x}(a_0 x dt + \sigma_0 x dz) - \frac{1}{2}\frac{1}{x^2}\sigma_0^2 x^2 dt \\ &= (a_0 - \frac{1}{2}\sigma_0^2)dt + \sigma_0 dz \end{aligned}$$

积分这个最后的方程，得到：

$$y(t) - y(0) = \int_0^t (a_0 - \frac{1}{2}\sigma_0^2)dt + \int_0^t \sigma_0 dz ,$$

它可以写成：

$$y(t) = y(0) + (a_0 - \frac{1}{2}\sigma_0^2)t + \sigma_0 z(t) \tag{15.7}$$

重新作代换 $y = \log x$，即 $x = e^y$，对于 $x(0) = e^{y(0)}$，我们得到（15.6）的解是：

$$x(t)=e^{y(t)}=x(0)\exp\left\{\left(a_0-\frac{\sigma_0^2}{2}\right)t+\sigma_0 z(t)\right\}$$

因此，找到一个合适的变换和利用积分就可以找到随机微分方程的解。

在第 8 和 9 节中，我们讨论了随机稳定性问题，并分清了点平衡稳定和依分布收敛到一个平稳状态分布意义下的稳定。数学家们已研究了噪声消失的点平衡稳定，而依分布收敛到一个与初始条件无关的平稳状态分布意义下的稳定还没有得到多少注意。即使这样，两个领域仍然需要进一步的研究。第一批点平衡随机稳定的论文之一是俄罗斯数学家 Kats 和 Krasovskii（1960）的工作，他们推广了确定性 Liapunov（1949）方法。Liapunov 确定性的稳定性方法基本参考资料有 Antosiewicz（1958）、Borg（1949）、Hahn（1963,1967）、Casari（1963）、Hartman（1964）、Krasovskii（1965）、La Salle 和 Lefschetz（1961）、La Salle（1964）Massera（1949,1956）及 Yoshizawa（1966）。在美国，Bucy（1965）和 Wonham（1966a,1966b）等在按 Liapunov 方法的思想求解随机稳定问题的工作中做出了贡献。Kushner（1967a）在 Liapunov 方法后，给出了主要随机稳定结果的一个完整的叙述。读者也许会发现 Kushner（1972）的随机稳定性综述对这个主题的入门是有用的。尽管随机稳定性的一般结果还不丰富，但在线性随机系统稳定性方面已取得了一些进展。一些这方面的结果和许多参考资料出现在 Kozin（1972）中。

我们区分点平衡随机稳定和初始条件无关的依分布收敛意义下的稳定并不意味着随机稳定性只有这两种定义方法。在点平衡随机稳定领域就有几种定义和定理存在。我们的第 8 节只给了点平衡随机稳定的一种概念。现在，我们提及两种另外的点平衡随机稳定性的定义，以说明这个重要的数学研究领域的范围。例如，如果我们令 $x(t)$ 表示随机微分方程（8.1）在 $[0,\infty)$ 上的解，$x=0$ 是平衡解，那么我们说 $0-$ 平衡是平均稳定的，若数学期望存在，且对于给定的 $\varepsilon>0$，存在 $\eta>0$ 使得 $|x|\equiv|x(0,\omega)|<\eta$，意味着：

$$E(\sup_t|x(t)|)<\varepsilon$$

另一个定义是：$0-$ 平衡是平均指数稳定的，若数学期望存在，且存在常数 α,β,η，都大于 0，使得 $|x|\equiv|x(0,\omega)|<\eta$，这说明

$$E(|x(t)|)<\beta|x(0,\omega)|e^{-\alpha t},\forall t>0\text{。}$$

几种其他的定义在数学文献中可见，参看 Kozin（1972）和 Kushner（1971）。

为了向读者表明在随机稳定性中所出现的数学奇妙特性，现在举一个简单例证也许是合适的，它取自 Kozin（1972,pp.142-192）。

考虑 $It\hat{o}$ 方程：

$$dx(t)=ax(t)dt+\sigma x(t)dz$$

其中 a 和 σ 为常数，z 为具有单位方差的 Wiener 过程。假定初始条件是 $x(0)=x_0$。上述方程的解以概率 1 为：

$$x(t)=x(0)\exp[(a-\frac{\sigma^2}{2})t+\sigma z(t)]$$

它的 n 阶矩是：

$$E(x^n(t))=x^n(0)\exp[(a-\frac{\sigma^2}{2})nt+\frac{\sigma^2 n^2}{2}t]$$

从最后这个方程，Kozin（1972,p.192）得出：只要 $a<\frac{\sigma^2}{2}(n-n^2)$，那么 n 阶矩是指数稳定的。因此，对于 $n=1$，$a>0$ 意味着一阶矩是指数稳定的，但更高阶矩不稳定。对于 $n=2$，$a<-\sigma^2$ 能保证一阶和二阶矩指数稳定，但更高阶矩不稳定。回顾第 8 节的 Liapunov-Kushner 方法，应用于同一个方程，表明只要 $a<-\frac{\sigma^2}{2}$，$x(t)$ 的样本轨道以概率 1 稳定，但关于矩的稳定性什么都没有说。很难对一个由点平衡稳定、但高阶矩不稳定的由 $It\hat{o}$ 方程所描述的经济模型给出经济解释。因此，我们需要区分样本轨道行为和矩行为，实际上是为了选择对应于这种行为或那种行为的应优先考虑的稳定性性质。

依分布收敛到平稳状态分布意义下的随机稳定性领域是一个有待研究的领域。很难建立一般框架下的均衡分布的存在性，更不用说证明其稳定性。对于特殊情况，我们有均衡分布的存在性定理。这个领域的基本结果在 Mandl（1968）的书中。一些原始工作首先出现在 Feller（1954）、Tanaka（1957）和 Khasminskii（1962）中。Merton（1975a）有一个关于特殊连续增长模型的随机稳定性结果，他通过假设常数储蓄函数是在一个常数储蓄函数集上取最大值而得到的。这是一个非常特殊的稳定性结果。连续随机最优增长中的随机稳定性问题，即证明最优随机过程依分布收敛到均衡状态分布的问题，仍然没有解决。Brock & Mirman（1972）证明了离散随机增长模型的稳定性结果。

关于随机控制这个话题，方程（10.1）和（10.2）所叙述的问题是由 Arrow & Kurz（1970,pp.27-51）所研究的确定性控制问题的随机版本。我们选择研究（10.1）和（10.2）是因为大多数经济学家熟悉由 Arrow-Kurz（1970）分析的确定性控制问题，目前的随机版本将使读者能够把熟悉的东西与新的数学问题和随机推广的结果进行比较。

第 10 节的分析使用 Bellman 最优原理，连同随机分析导出随机最优性条件。Bellman 最优原理首先出现在 Bellman（1957）中，随后，在许多书中普及，例如，Dreyfus（1965）、Hadley（1964）和 Mangasarian（1969）。

Aoki（1967）、Åström（1970）、Kushner（1971）和 Bertsekas（1976）在入门水平上介绍了随机控制理论。然而，与确定性最优控制领域不同，这里有许多书可用，例如 Anderson 和 Moore（1971）、Athans 和 Falb（1966）、Berkovitz（1974）、Bryson 和 Ho（1979）、Kwakernaak

和 Sivan（1972）、Strauss（1968）、Pontryagin 等（1962）、Hestenes（1966）及 Lee 和 Markus（1967），而关于随机最优控制的文献还不多。我们鼓励具有高级水平的读者去查阅 Fleming 和 Rishel（1975）的书和下列文章：Benes（1971），Bismut(1973,1976)，Davis(1973)，Fleming(1969, 1971)，Kushner (1965,1967 b,1975)，Rishel (1970)和 Wonham（1970）。我们注意到 Fleming 和 Rishel（1975,第 5 章）提供了随机最优控制问题的严格分析，它给了我们的启发性方法的一个补充。在他们叫做“核实理论”的定理中，Fleming 和 Rishel（1975,p.159）给出了最优的充分条件：假定带适当边界条件的 Hamiton-Jacobi-Bellman 非线性偏微分方程存在性质非常好的解。

在第 12 节中，我们介绍一般 $It\hat{o}$ 公式和跳跃过程的最大值原理。我们的方法是直观的，其目的是使读者熟悉 Merton（1971）中所用的一些数学知识。严格的分析将需要方程（12.1）中的项 $g(t,x)dq(t)$ 的确切的数学性质和跳过程积分的意义。Kushner（1967,p.18）和 Doob（1953,p.284）讨论了这个问题的某些方面，至于更详细的处理见 Dellacheria（1974）。

注意到这一点是适当的，当经济学家把随机控制理论运用到经济模型时，马上就需要发展这个随机模型的稳定性分析。最近在确定性情况中所作的经济研究，如 Araujo 和 Scheinkman（1977,1976）、Benveniste 和 Scheinkman（1977,1979）、Scheinkman（1976,1978）、Brock 和 Scheinkman（1977,1976）、Brock（1976,1977）、Cass 和 Shell（1976）、Mckenzie（1976）、Magill（1977a）、Samuelson（1972）及 Levhari 和 Liviatan（1972）等，已经证实大量的经济问题由确定性最优控制产生，它的稳定性质对于正确说明这个模型是决定性的。得到确定性控制系统的稳定性方面的数学结果，如 Galperin 和 Krasovskii（1963）、Hale（1969）、Hartman（1961）、Hartman 和 Olech（1962）、Lefschetz（1965）、Mangasarian（1963,1966）、Markus 和 Yamabe（1960）、Rockafellar（1973,1976）及 Roxin（1965,1966），对经济研究工作者是有帮助的。另一方面，没有得到由随机控制产生的随机微分方程系统稳定性方面的足够的数学结果可能推迟这个领域的经济研究。这种经济问题的一个实例见 Brock 和 Magill（1979）及 Magill（1977b）。

在经济学和金融学中，随机方法的许多应用都是用来处理离散时间的随机控制，而不是这一章里所讨论的连续时间的问题。在下面两章中，为了丰富读者的学习，我们选择包括离散时间的随机应用。尽管本章中我们没有明显讨论离散随机方法，但两者存在许多类似，许多应用将说明这一点。这里提及 Kushner（1971）建立离散与连续时间随机模型之间的关系和类似就足够了。离散时间随机问题通常没有对应的连续时间随机问题复杂。后者可以用长度为 h 的时间区间的离散时间过程来逼近。因为 h 可以充分小，于是离散和连续时间模型之间的关系能够建立起来。在金融学中，Merton（1978）只使用基本概率方法逼近连续时间模型推导出连续时间定理。在这样做的过程中，他揭示了埋置在连续时间数学理论中的经济假设。