STAT0041: Stochastic Calculus

Lecture 13 - Diffusion Process

Lecturer: Weichen Zhao Fall 2025

Key concepts:

- *Itô*扩散过程;
- 扩散算子;
- *Fokker-Planck*方程。

13.1 扩散过程

扩散是一种物理现象，指任何物质（如原子、离子、分子、能量）从浓度较高的区域向浓度较低的区域的净移动。例如, 颜料滴进水中，在水里扩散开来的运动过程。扩散过程并没有统一的数学定义, 但其核心是**轨道连续的马氏过程**.

历史上，研究扩散过程主要有两种方法：

1. 通过建立**转移概率**满足的方程来刻画其宏观性质的演化, 例如, 将墨水滴入水中, 墨水浓度在不同时刻和位置遵从的规律，这是 Kolmogorov 的**分析**方法。

2. 利用追踪每个花粉粒子或墨水粒子的**轨迹**, 构造概率空间, 建立**随机微分方程**从微观角度描述其服从的运动规律，这是 Itô 的随机分析方法。

经典扩散过程是随机过程中的内容，可以参考：Pavliotis G A. Stochastic processes and applications[J]. Texts in applied mathematics, 2014, 60. Chapter 2. Diffusion Processes

本节课我们从随机分析角度，研究扩散过程。

Definition 13.1 (Itô扩散过程) *Itô* 扩散过程为一个随机过程

$$X_t(\omega) = X(t,\omega) : [0,\infty) \times \Omega \to \mathbb{R}^n$$

满足如下形式的随机微分方程

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t, \quad t \geq s; \quad X_s = x$$

其中B_t为一个m维布朗运动，并且 $b : [0,\infty) \times \mathbb{R}^n \to \mathbb{R}^n$，$\sigma : [0,\infty) \times \mathbb{R}^n \to \mathbb{R}^{m\times n}$满足随机微分方程的强解存在唯一性条件：

(1) 存在常数C, 使得

$$|b(t,x)|+|\sigma(t,x)| \leqslant C(1+|x|), \quad \forall x \in \mathbb{R}^n, t \in [0,T],$$

(2) 存在常数K, 使得

$$|b(t,x) - b(t,y)|+|\sigma(t,x) - \sigma(t,y)| \leqslant K|x-y|, \quad \forall x,y \in \mathbb{R}^n, t \in [0,T].$$

如果b, σ只依赖于X_t，并且满足

$$|b(x) - b(y)|+|\sigma(x) - \sigma(y)| \leqslant D|x-y|; \quad x,y \in \mathbb{R}^n,$$

我们称X_t为一个时齐的 *Itô* 扩散过程。

下面我们证明 Itô 扩散过程是一个Markov过程。

简单回忆Markov过程：设$\{X_t : t \in \mathcal{T}\}$为一个$(\Omega, \mathscr{F}, (\mathscr{F}_t), \mathrm{P})$上$\mathscr{F}_t$适应过程，状态空间为$(E, \mathscr{E})$，称$X$ 为一个Markov过程，如果对所有E上有界可测函数f， $s < t \in \mathcal{T}$

$$\mathbb{E}[f(X_t)|\mathscr{F}_s] = \mathbb{E}[f(X_t)|X_s] \quad a.s.$$

Theorem 13.2 *Itô* 扩散过程是一个 *Markov* 过程。

Proof: 考虑Itô扩散过程的积分形式：

$$X_t = X_0 + \int_0^t b(u, X_u)\mathrm{d}u + \int_0^t \sigma(u, X_u)\mathrm{d}B_u,$$

记s时刻从x位置出发的解为

$$X_t^{s,x} \triangleq x + \int_s^t b(u, X_u^{s,x})\mathrm{d}u + \int_s^t \sigma(u, X_u^{s,x})\mathrm{d}B_u, \quad s < t.$$

$X_t^{s,x}$ 只与$s, x, t, (B_u - B_s)_{s<u\le t}$有关，可以将$X_t^{s,x}$看成$s, x, t$，$(B_u - B_s)_{s<u\le t}$的函数，记为

$$G(s, x, t, (B_u - B_s)_{s<u\le t})$$

另一方面，由积分的线性性质，

$$X_t = X_0 + \int_0^s b(u, X_u)\mathrm{d}u + \int_0^s \sigma(u, X_u)\mathrm{d}B_u + \int_s^t b(u, X_u)\mathrm{d}u + \int_s^t \sigma(u, X_u)\mathrm{d}B_u.$$

由于解的唯一性，

$$\begin{aligned} X_t &= X_s + \int_s^t b(u, X_u)\mathrm{d}u + \int_s^t \sigma(u, X_u)\mathrm{d}B_u \\ &= G(s, X_s, t, (B_u - B_s)_{s<u\le t}) \end{aligned}$$

由布朗运动的增量独立性，$(B_u - B_s)_{s<u\le t}$ 与 $\mathscr{F}_s \triangleq \sigma(B_v, v \le s)$独立，而且$X_s \in \mathscr{F}_s$，这意味着，对于任意的有界可测函数$f$

$$\begin{aligned} \mathbb{E}[f(X_t)|\mathscr{F}_s] &= \mathbb{E}[f(G(s, X_s, t, (B_u - B_s)_{s<u\le t}))|\mathscr{F}_s] \\ &= \mathbb{E}[f(G(s, x, t, (B_u - B_s)_{s<u\le t}))]|_{x=X_s} \end{aligned}$$

那么有

$$\begin{aligned} \mathbb{E}[f(X_t)|X_s] &= \mathbb{E}[\mathbb{E}[f(X_t)|\mathscr{F}_s]|X_s] \\ &= \mathbb{E}[\mathbb{E}[f(G(s, x, t, (B_u - B_s)_{s<u\le t}))]|_{x=X_s}|X_s] \\ &= \mathbb{E}[f(G(s, x, t, (B_u - B_s)_{s<u\le t}))]|_{x=X_s} \\ &= \mathbb{E}[f(G(s, X_s, t, (B_u - B_s)_{s<u\le t}))|\mathscr{F}_s] \\ &= \mathbb{E}[f(X_t)|\mathscr{F}_s]. \end{aligned}$$

■

13.2 Markov半群理论

本节简要介绍Markov半群理论的相关内容。

回顾定义：对状态空间为$(E,\mathscr{E})$的 Markov 过程X_t，定义:

$$(P_t f)(x) \triangleq \mathbb{E}[f(X_t)|X_0 = x] = \int f(y)p(t,x,y)dy.$$

其中

$$P_t : \mathcal{M}_b(E) \to \mathcal{M}_b(E),\ f \mapsto P_t f$$

为$\mathcal{M}_b(E)$上的算子，$\mathcal{M}_b(E)$ 为E上所有有界可测函数的集合。那么由Kolmogorov-Chapman方程，P_t满足半群(semigroup)性质

$$P_{t+s}f = P_t \circ P_s f.$$

称P_t为一个*Markov*半群。

设P_t为一个Markov半群，定义

$$\mathcal{L}f := \lim_{t\downarrow 0} \frac{P_t f - f}{t}.$$

我们称$\mathcal{L}$ 为P_t的无穷小生成元，简称生成元。

对于作用在有界连续函数半群P_t，我们可以定义它的伴随半群*(Adjoint semigroup)* P_t^*，作用在概率测度上。

Definition 13.3 (伴随半群) *设μ为X_0的分布，定义算子:*

$$(P_t^*\mu)(A) \triangleq \int P(X_t \in A|X_0 = x)d\mu(x) = \int p(t,x,A)d\mu(x).$$

那么P_t^*满足

$$\int (P_t f)(x)d\mu(x) = \int f(x)d(P_t^*\mu)(x)$$

我们称算子P_t 和 P_t^*是伴随的，P_t^*为Markov半群P_t的伴随半群。记 $\mu_t \triangleq P_t^*\mu$，那么$\mu_t$为 Markov 过程 X_t 的分布。

注. 设 T 为内积空间 H 上的算子，则 T 的伴随算子 T^* 是满足以下关系的唯一算子：

$$\langle T(x), y\rangle = \langle x, T^*(y)\rangle, \quad \forall x, y \in H,$$

把 $d\mu(x)$ 写成密度函数 $g(x)dx$，则 $d(P_t^*\mu)(x) = P_t^* g(x)dx$，伴随性质即为

$$\int P_t f(x)g(x)dx = \int f(x)P_t^* g(x)dx.$$

不变测度也称平稳分布，是Markov过程研究中的重要概念，下面在Markov半群语言下给出不变测度的概念。

Definition 13.4 (不变测度) 给定一个*Markov* 半群 P_t, 我们称测度 π 是P_t不变的，如果对任意正有界可测函数 $f : E \to \mathbb{R}$以及任意$t \geq 0$，有

$$\int_E P_t f d\pi = \int_E f d\pi.$$

或者等价地，

$$P_t^* \pi = \pi$$

13.3 扩散算子

经典扩散过程的研究一般先给出生成元，本节我们考虑扩散过程的生成元算子，一般称为扩散算子。

Definition 13.5 (扩散算子) 一个*Markov*扩散算子*(diffusion operator)* $\mathcal{L}$为一个如下形式的二阶微分算子

$$\begin{aligned}\mathcal{L}f &= \sum_i b_i(x)\frac{\partial}{\partial x_i}f + \frac{1}{2}\sum_{i,j}\Sigma_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j}f \\ &= b(x)\cdot\nabla f + \frac{1}{2}Tr(\Sigma(x)^T\nabla^2 f)\end{aligned} \tag{13.1}$$

其中 $\Sigma(x) = (\Sigma_{ij}(x))_{1\leq i,j\leq n}$ 和 $b(x) = (b_i(x))_{1\leq i\leq n}$分别为$x$的光滑 $n\times n$ 对称矩阵值和 $\mathbb{R}^n$值函数。有时也记 $\Sigma : \nabla^2 f \triangleq Tr(\Sigma(x)^T\nabla^2 f)$为矩阵的*Frobenius*内积。

经典扩散过程由扩散算子$\mathcal{L}$构造，$b(x)$和$\Sigma(x)$具有相应的含义，我们下面给出经典一维扩散过程的定义。

Definition 13.6 (经典一维扩散过程) 一个$\mathbb{R}$上转移函数为$p(s,x;t,A)$ 的 *Markov* 过程 X_t 称为一个扩散过程，如果：

(1) (连续性Continuity). 对于任意x和$\epsilon > 0$,

$$\lim_{t-s\downarrow 0} \frac{1}{t-s}\, p(s,x;t,\{|X_t - x| > \varepsilon\}) = 0.$$

(2) (漂移系数的定义). 存在函数$b(s,x)$使得，对于任意x和$\epsilon > 0$,

$$\lim_{t-s\downarrow 0} \frac{1}{t-s} \int_{|y-x|\le\varepsilon} (y-x)p(s,x;t,\mathrm{d}y) = b(s,x).$$

(3) (扩散系数的定义).存在函数$\Sigma(x,s)$使得，对于任意x和$\epsilon > 0$,

$$\lim_{t-s\downarrow 0} \frac{1}{t-s} \int_{|y-x|\le\varepsilon} (y-x)^2 p(s,x;t,\mathrm{d}y) = \Sigma(x,s).$$

注. 在物理学上，一阶矩主要描述动力学趋势，二阶矩主要给出随机波动。

接下来我们看看Itô用SDE构造的Itô扩散过程和经典的是否一致，也就是说我们下面的定理：它说明了Itô扩散过程的生成元就是扩散算子(13.1)

Theorem 13.7 设X_t为一个生成元为$\mathcal{A}$的时齐*Itô*扩散过程

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t$$

那么对于所有的二次连续可微函数 f，有

$$\mathcal{A}f = \mathcal{L}f,$$

其中

$$\begin{aligned}\mathcal{L}f &= \sum_i b_i(x)\frac{\partial}{\partial x_i}f + \frac{1}{2}\sum_{i,j}\Sigma_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j}f \\ &= b(x)\cdot\nabla f + \frac{1}{2}Tr(\Sigma^T(x)\nabla^2 f),\end{aligned}$$

并且 $\Sigma_{ij}(x) = \sum_{k=1}^n \sigma_{ik}(x)\sigma_{kj}(x) = \sigma(x)\sigma^T(x)$.

Proof: 对$f(X_t)$应用Ito公式，

$$\begin{aligned}\mathrm{d}f(X_t) &= \sum_i \frac{\partial}{\partial x_i}f(X_t)dX_t^i + \frac{1}{2}\sum_{i,j}\frac{\partial^2}{\partial x_i\partial x_j}f(X_i)\mathrm{d}X_t^i\mathrm{d}X_t^j \\ &= \sum_i \frac{\partial}{\partial x_i}f(X_t)\sum_j \sigma_{ij}(X_t)dB_t^j + \sum_i \frac{\partial}{\partial x_i}f(X_t)b_i(X_t)\mathrm{d}t + \frac{1}{2}\sum_{i,j}\Sigma_{ij}(X_t)\frac{\partial^2}{\partial x_i\partial x_j}f(X_i)\mathrm{d}t\end{aligned}$$

$$
\begin{aligned}
\mathbb{E}[f(X_t)|X_0=x] = f(x) &+ \sum_{i,j}\mathbb{E}\left[\int_0^t \sigma_{ij}(X_s)\frac{\partial f(X_s)}{\partial x_i}dB_s^j \mid X_0=x\right] \\
&+ \mathbb{E}\left[\int_0^t\left(\sum_i b_i(X_s)\frac{\partial f(X_s)}{\partial x_i}+\frac{1}{2}\sum_{i,j}\Sigma_{i,j}(X_s)\frac{\partial^2 f(X_s)}{\partial x_i\partial x_j}\right)ds \mid X_0=x\right] \\
&= f(x)+\mathbb{E}\left[\int_0^t \mathcal{L}f(X_s)ds \mid X_0=x\right]+\sum_{i,j}\mathbb{E}\left[\int_0^t \sigma_{ij}(X_s)\frac{\partial f(X_s)}{\partial x_i}dB_s^j \mid X_0=x\right]
\end{aligned}
$$

由于对于有界可测函数g，$\xi_t \triangleq \int_0^t g(X_s)dB_s$是鞅，

$$
\mathbb{E}\left[\int_0^t g(X_s)dB_s \mid X_0=x\right]=\xi_0=\int_0^0 g(X_s)dB_s=0
$$

所以

$$
\mathbb{E}[f(X_t)|X_0=x]=f(x)+\mathbb{E}\left[\int_0^t \mathcal{L}f(X_s)ds \mid X_0=x\right].
$$

那么

$$
\begin{aligned}
\mathcal{A}f(x) \triangleq& \lim_{t\downarrow 0}\frac{\mathbb{E}[f(X_t)|X_0=x]-f(x)}{t} \\
=& \lim_{t\downarrow 0}\frac{1}{t}\mathbb{E}\left[\int_0^t \mathcal{L}f(X_s)ds \mid X_0=x\right] \\
=& \mathbb{E}\left[\lim_{t\downarrow 0}\frac{1}{t}\int_0^t \mathcal{L}f(X_s)ds \mid X_0=x\right] \qquad \text{(控制收敛定理)} \\
=& \mathbb{E}\left[\mathcal{L}f(X_0) \mid X_0=x\right] \qquad (*) \\
=& \mathcal{L}f(x)
\end{aligned}
$$

为验证$(*)$，只需考虑对某个连续有界函数h

$$
\begin{aligned}
\left|\frac{1}{t}\int_0^t h(X_s)\,ds-h(X_0)\right| &= \frac{1}{t}\left|\int_0^t (h(X_s)-h(X_0))\,ds\right| \\
&\le \frac{1}{t}\int_0^t |h(X_s)-h(X_0)|\,ds.
\end{aligned}
$$

由于h连续，对任意$\epsilon>0$，存在δ，使得$0<s<\delta$时

$$
|h(X_s)-h(X_0)|<\epsilon
$$

由于$t\downarrow 0$，此时，

$$
\frac{1}{t}\int_0^t |h(X_s)-h(X_0)|\,ds<\frac{1}{t}\int_0^t \epsilon ds<\epsilon
$$

■

注. 之后我们继续使用记号$\mathcal{L}$表示Itô 扩散过程X_t的生成元。记

$$u(t,x) \triangleq \mathbb{E}(f(X_t)|X_0 = x)$$

那么$u(t,x)$满足Kolmogorov向后方程

$$\frac{\partial u(t,x)}{\partial t} = \mathcal{L}u(t,x) = b(x)\cdot \nabla u(t,x) + \frac{1}{2}Tr(\Sigma^T(x)\nabla^2 u(t,x)), \quad t > 0.$$

13.4 Fokker-Planck 方程

本节我们关注Itô扩散过程分布的演化，这是由Fokker-Planck方程描述的，Fokker-Planck方程本质就是Kolmogorov向前方程，我们将在证明中看到。

Theorem 13.8 (Fokker-Planck 方程) *设X_t为一个时齐 Itô扩散过程*

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t,$$

并且 $\mu_t(x) = \mu(x,t)$ 为X_t在t时刻的概率分布，且$\mu_t(x)$关于x和t二次连续可微，设$\mu(x)$为初始概率分布，那么$\mu(x,t)$是下面方程的解

$$\partial_t \mu(x,t) = \mathcal{L}^*\mu(x,t) \quad (t>0), \quad \mu(x,0) = \mu(x),$$

其中算子

$$\begin{aligned}\mathcal{L}^* g &\triangleq -\sum_j \frac{\partial}{\partial x_j}(b_j(x)g) + \frac{1}{2}\sum_{i,j}\frac{\partial^2}{\partial x_i \partial x_j}(\Sigma_{ij}(x)g)\\ &= -\nabla\cdot(b(x)g) + \frac{1}{2}\nabla\cdot\nabla\cdot(\Sigma(x)g)\end{aligned}$$

为$\mathcal{L}$的伴随算子。

Proof: 我们先形式计算算子$\mathcal{L}^*$。由$\mathcal{L}^*$的伴随性，对内积$\langle f,g\rangle := \int fgdx$，有

$$\langle \mathcal{L}f, g\rangle = \langle f, \mathcal{L}^* g\rangle$$

那么

$$\begin{aligned}\langle \mathcal{L}f, g\rangle &= \int \left(b(x)\cdot\nabla f + \frac{1}{2}Tr(\Sigma^T(x)\nabla^2 f)\right) g(x)dx\\ &= \int \left(g(x)b(x)\cdot\nabla f + \frac{1}{2}g(x)Tr(\Sigma^T(x)\nabla^2 f)\right) dx\end{aligned}$$

下面我们分别计算

$$gb \cdot \nabla f, \qquad gTr(\Sigma^T \nabla^2 f)$$

由散度乘积法则，有

$$gb \cdot \nabla f = \nabla \cdot (fgb) - f\nabla \cdot (gb)$$

且

$$\begin{aligned}
&gTr(\Sigma^T \nabla^2 f) \\
&= ((g\Sigma) \cdot \nabla) \cdot (\nabla f) \\
&= \nabla \cdot ((g\Sigma) \cdot (\nabla f)) - (\nabla f) \cdot \nabla \cdot (g\Sigma) && (*) \\
&= \nabla \cdot ((g\Sigma) \cdot (\nabla f)) - [\nabla \cdot (f\nabla \cdot (g\Sigma)) - f\nabla \cdot \nabla \cdot (g\Sigma)] && (\text{散度乘积法则}) \\
&= \nabla \cdot ((g\Sigma) \cdot (\nabla f)) - \nabla \cdot (f\nabla \cdot (g\Sigma)) + f\nabla \cdot \nabla \cdot (g\Sigma)
\end{aligned}$$

其中$n \times n$矩阵A和n维向量$\mathbf{v}$的数量积为

$$A \cdot \mathbf{v} \triangleq (\mathbf{a}_1 \cdot \mathbf{v}, \ldots, \mathbf{a}_n \cdot \mathbf{v}) = \left(\sum_{j=1}^n a_{1j} v_j, \ldots, \sum_{j=1}^n a_{nj} v_j \right)$$

我们现在验证$(*)$式，写成分量形式

$$((g\Sigma) \cdot \nabla) \cdot (\nabla f) = \sum_{i,j} g\Sigma_{ij} \frac{\partial^2 f}{\partial x_i \partial x_j}$$

$$\nabla \cdot ((g\Sigma) \cdot (\nabla f)) = \sum_i \frac{\partial}{\partial x_i} \left(\sum_j g\Sigma_{ij} \frac{\partial f}{\partial x_j} \right)$$

$$(\nabla f) \cdot \nabla \cdot (g\Sigma) = \sum_i \frac{\partial f}{\partial x_i} \left(\sum_j \frac{\partial g\Sigma_{ij}}{\partial x_j} \right)$$

由于

$$\frac{\partial}{\partial x_i} \left(\sum_j g\Sigma_{ij} \frac{\partial f}{\partial x_j} \right) = \sum_j \left(\frac{\partial g\Sigma_{ij}}{\partial x_j} \frac{\partial f}{\partial x_i} + g\Sigma_{ij} \frac{\partial^2 f}{\partial x_i \partial x_j} \right)$$

所以$(*)$式成立。回到

$$\begin{aligned}
\langle \mathcal{L}f, g \rangle &= \int (\nabla \cdot (fgb) - f\nabla \cdot (gb))\, dx \\
&\quad + \frac{1}{2} \int (\nabla \cdot ((g\Sigma) \cdot (\nabla f)) - \nabla \cdot (f\nabla \cdot (g\Sigma)) + f\nabla \cdot \nabla \cdot (g\Sigma))\, dx \\
&= \int f \left(-\nabla \cdot (bg) + \frac{1}{2} \nabla \cdot \nabla \cdot (\Sigma g) \right) dx \\
&\quad + \int \left(\nabla \cdot (fgb) + \frac{1}{2} \nabla \cdot (g\Sigma \cdot \nabla f) - \frac{1}{2} \nabla \cdot (f\nabla \cdot (g\Sigma)) \right) dx.
\end{aligned}$$

由于f和g无穷远处为0(边界条件)，由散度定理:

$$\langle \mathcal{L}f, g\rangle = \int f\left(-\nabla\cdot(bg) + \frac{1}{2}\nabla\cdot\nabla\cdot(\Sigma g)\right)dx$$

即

$$\mathcal{L}^* g = -\nabla\cdot(bg) + \frac{1}{2}\nabla\cdot\nabla\cdot(\Sigma g)$$

回顾伴随半群的定义

$$(P_t^*\mu)(A) \triangleq \int P(X_t\in A|X_0=x)d\mu(x) = \int p(t,x,A)d\mu(x).$$

满足

$$\mu_t(x) = \mu(x,t) = P_t^*\mu(x), \quad \int (P_tf)(x)d\mu(x) = \int f(x)d(P_t^*\mu)(x)$$

由Kolmogorov向前方程

$$\frac{\partial}{\partial t}P_tf = P_t\mathcal{L}f$$

有

$$\begin{aligned}\partial_t\int f\,\mathrm{d}P_t^*\mu = \partial_t\int P_tf\,\mathrm{d}\mu = \int P_t\mathcal{L}f\,\mathrm{d}\mu \\ = \int \mathcal{L}f\,\mathrm{d}P_t^*\mu = \int f\,\mathrm{d}\mathcal{L}^*P_t^*\mu.\end{aligned}$$

这即是Fokker-Planck 方程

$$\partial_t\mu(x,t) = \mathcal{L}^*\mu(x,t).$$

■

注. 用Fokker-Planck 方程，我们可以快速计算不变测度。由不变测度的定义：

$$\int_E P_tfd\pi = \int_E fd\pi.$$

等价地，

$$P_t^*\pi = \pi$$

那么

$$\mathcal{L}^*\pi = 0$$

由此可以解得不变测度π.