tieli/notes/appl_stoch/lect*.pdf 部分习题参考答案

吴大维

1582 年 10 月 10 日

1 Lecture 2

5. 为了计算 $\mathbb{E}(X_1 \cdots X_k)$, 只需求出矩母函数 $M(t)$ 在原点处对 $t_1, \dots, t_k$ 的混合偏导数

$$\frac{\partial^k M(t)}{\partial t_1 \partial t_2 \cdots \partial t_k}\bigg|_{t=0}.$$

又因为对于正态变量 $X \sim \mathcal{N}(0, \Sigma)$ 而言, $M(t) = \exp\left(\frac{1}{2}t^T \Sigma t\right)$ 是解析的, 所以又只需求出它在原点 Taylor 展开式中 $t_1 \cdots t_k$ 项系数的 $1!1! \cdots 1! = 1$ 倍.

用 e^x 在原点的 Taylor 展开, 得到

$$M(t) = \sum_{k=0}^{\infty} \frac{1}{k!} \left(\frac{1}{2} \sum_{i,j=1}^{n} \sigma_{ij} t_i t_j \right)^k.$$

显然展开式中只有偶数次项, 所以 k 为奇数时 $\mathbb{E}(X_1 \cdots X_k) = 0$.

k 为偶数时, $t_1 \cdots t_k$ 一项只可能来自和式中的 $\frac{k}{2}$ 次项

$$\frac{1}{(k/2)!} \left(\frac{1}{2} \sum_{i=1}^{n} \sigma_{ii} t_i^2 + \sum_{i<j} \sigma_{ij} t_i t_j \right)^{\frac{k}{2}}.$$

又因为 $t_1, \dots, t_k$ 是两两不同的变元, 所以 $t_1 \cdots t_k$ 这一项只可能由交叉项 $\sigma_{ij} t_i t_j$ $(i \neq j)$ 产生, 而且每个括号提供两个因子. 任何一种 $\{1, 2, \dots, k\}$ 的分划 $\{i_1, i_2\}, \{i_3, i_4\}, \dots, \{i_{k-1}, i_k\}$, 对应了一种由每个括号提供因子的方式 (即第 j 个括号提供 $t_{i_{2j-1}}, t_{i_{2j}}$ 的因子); 对于同一种分划方式, 它的 $\frac{k}{2}$ 个因子的二元组有 $(\frac{k}{2})!$ 种不同的排列方式, 所以它产生的系数为 $(\frac{k}{2})! \sigma_{i_1 i_2} \cdots \sigma_{i_{k-1} i_k}$. 因此求和后该项的系数为

$$\frac{1}{(k/2)!} \cdot \sum_{\text{pairing}} (k/2)! \prod \sigma_{ij} = \sum_{\text{pairing}} \prod \sigma_{ij}.$$

而根据二阶矩的性质, $\mathbb{E}(X_i X_j) = \sigma_{ij}$, 证明完毕.

2 Lecture 3

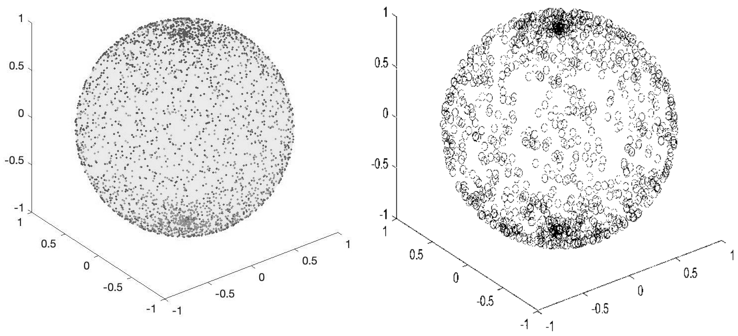
2. 首先, $\mathbb{S}^2$ 指的是三维空间中的二维球面.

$$\mathbb{S}^2 = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\}.$$

其次, 很多同学选择如下的方法生成球面均匀分布.

取 $\theta \sim \mathcal{U}(0, \pi), \varphi \sim \mathcal{U}(0, 2\pi)$, 则 $(\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta) \sim \mathcal{U}(\mathbb{S}^2)$.

部分同学生成的分布如下图所示.



可以明显看到这些样本点在两极较为集中, 这是因为球面上纬度角 θ 并非均匀分布. 球面上的面元等于 $\sin\theta \mathrm{d}\theta \mathrm{d}\varphi$, 所以 θ 的分布的密度函数正比于 $\sin\theta$, 或者说 $\cos\theta$ 是均匀分布. (亦可通过 “球冠面积正比于高度” 的事实推出) 因此, 若采用球坐标方式生成, 生成 θ 的方式应为

$$U \sim \mathcal{U}[-1,1],$$
$$\theta \sim \arccos(U).$$

4. 题中算法的第 2、3 步的总效果是在 $U < \frac{p(X)}{Mg_m(X)}$ 时接受, 所以和算法 2.6 等价 (在算法 2.6 中, 把 $f(x)$ 换成 $Mg_m(x)$, A 换成 M), 是正确的采样算法.

该算法的优势在于, 第 2 步计算的函数 $g_l(x)$ 非常简单, 所以有一定概率减少计算 $f(x)$ 的次数, 从而减少计算复杂度.

$$\mathbb{P}(\text{不用计算} f(x)) = \int g_m(x) \frac{g_l(x)}{Mg_m(x)} \mathrm{d}x = \frac{1}{M} \int g_l(x) \mathrm{d}x.$$

3 Lecture 4

1. 我们证明对任何单调增函数 f, g, 都有 $\mathrm{Cov}(f(X)g(X)) \geq 0$.

$$\begin{aligned}
\mathrm{Cov}(f(X)g(X)) &= \mathbb{E}(f(X)g(X)) - \mathbb{E}f(X)\mathbb{E}g(X) \\
&= \frac{1}{2}\left[\int_0^1 (f(x)g(x) + f(y)g(y))\mathrm{d}x - \int_0^1 f(x)\mathrm{d}x \int_0^1 g(y)\mathrm{d}y - \int_0^1 g(x)\mathrm{d}x \int_0^1 f(y)\mathrm{d}y\right] \\
&= \frac{1}{2}\int_{[0,1]^2} [f(x)g(x) + f(y)g(y) - f(x)g(y) - f(y)g(x)]\mathrm{d}x\mathrm{d}y \\
&= \frac{1}{2}\int_{[0,1]^2} [f(x) - f(y)][g(x) - g(y)]\mathrm{d}x\mathrm{d}y.
\end{aligned}$$

由于 f, g 单调性相同, 括号内恒非负, 所以积分恒非负. 证明完毕.

3. 由熟知结论 $\log x \leq x - 1, \forall x > 0$. 所以

$$\mathrm{D}(f\|g) = -\int f(x) \log \frac{g(x)}{f(x)} \mathrm{d}x \geq -\int f(x) \left(\frac{g(x)}{f(x)} - 1\right) \mathrm{d}x = \int [f(x) - g(x)]\mathrm{d}x = 0.$$

等号成立当且仅当 $\frac{g(x)}{f(x)} = 1$ 处处成立, 即 $f \equiv g$.

4 Lecture 5

2. (a) 注意到

$$Z_{2n} = \frac{Z_n + \widetilde{Z}_n}{\sqrt{2}},$$

其中 $\widetilde{Z}_n$ 是与 Z_n 独立同分布的随机变量. 记 Z_n 的特征函数为 f_n, 则根据特征函数的性质,

$$f_{2n}(t) = f_n^2\left(\frac{t}{\sqrt{2}}\right).$$

令 $n \to \infty$, 那么根据假设 $f_n \to f$, 所以

$$f(t) = f^2\left(\frac{t}{\sqrt{2}}\right).$$

(b) 由递推式可知

$$f(t) = \left[f(2^{-\frac{n}{2}}t)\right]^{2^n}.$$

因为 $f \in C^2$, 所以在原点处有二阶 Taylor 展开 (条件 $\mathbb{E}X_n = 0 \Rightarrow f'(0) = 0$)

$$f(x) = 1 + \frac{1}{2}f''(0)x^2 + o(x^2).$$

所以

$$f(t) = \left[1 + \frac{1}{2}f''(0)2^{-n}t^2 + o(2^{-n})\right]^{2^n}.$$

令 $n \to \infty$, 即得

$$f(t) = \mathrm{e}^{\frac{1}{2}f''(0)t^2},$$

就是 Gauss 函数的特征函数.

(c) 一般地, 此时成立

$$Z_{m+n} = \frac{mZ_m + n\widetilde{Z}_n}{m+n},$$

从而

$$f_{m+n}(t) = f_m\left(\frac{m}{m+n}t\right)f_n\left(\frac{n}{m+n}t\right).$$

对任意 $s, t \in \mathbb{R}_+$, 令 $m, n \to \infty$ 并且 $\frac{m}{n} \to \frac{s}{t}$, 可得

$$f(s+t) = f(s)f(t).$$

由 $|f| \le 1$ 以及连续性, 根据 Cauchy 方程的性质, 必有 $f(t) = \mathrm{e}^{-kt}, t \ge 0, (k \ge 0)$. 同理 $t \le 0$ 时有 $f(t) = \mathrm{e}^{k't}$ $(k' \ge 0)$.

若加上对称性假设 $f(t) = f(-t)$, 则有 $k = k'$, 所以 $f(t) = \mathrm{e}^{-k|t|}$. $k = 0$ 时, 这是单点分布的特征函数; $k > 0$ 时, 这是 Cauchy 分布的特征函数.

(d) 用 (c) 中的类似推导可以得到此时 $f(t)$ 满足的函数方程

$$f(t) = f(p^\alpha t)f((1-p)^\alpha t), \quad \forall t \in \mathbb{R}, p \in (0,1).$$

$t \ge 0$ 时, 换元 $t = s^\alpha, g(s) = f(t)$, 可得

$$g(s) = g(ps)g((1-p)s),$$

所以由 Cauchy 方程的性质, $g(s) = \mathrm{e}^{-ks}, \forall s \geq 0$. 同理可对 $t < 0$ 讨论, 再由对称性 $f(t) = f(-t)$ 得

$$f(t) = \exp\left(-k|t|^{\alpha^{-1}}\right), \forall t \in \mathbb{R}.$$

由 Durrett 的例 3.3.23 (Durrett, R. Probability: Theory and Examples, 5th ed., pp. 138–139) 即得 α 的范围为

$$0 < \alpha^{-1} \leq 2 \Leftrightarrow \alpha \geq \frac{1}{2}.$$

4. (单边 Laplace 引理) 首先不妨设 $h(0) = 0$, 否则考虑 $h(x) - h(0)$. 根据题意, $h(x)$ 在 $x = 0$ 处可微, 所以对任意 $\varepsilon > 0$, 存在 $\delta > 0$ 使得 $0 < x < \delta$ 时,

$$|h(x) - h'(0)x| < \varepsilon x.$$

我们将积分分为 $(0, \delta)$ 和 (δ, ∞) 两段.

$$I(t) = \left(\int_0^\delta + \int_\delta^\infty\right) \mathrm{e}^{-th(x)} \mathrm{d}x = I_1 + I_2.$$

首先,

$$\frac{1 - \mathrm{e}^{-\delta(-h'(0)+\varepsilon)t}}{(-h'(0) + \varepsilon)t} < \int_0^\delta \mathrm{e}^{t(h'(0)-\varepsilon)x} \mathrm{d}x < I_1 < \int_0^\delta \mathrm{e}^{t(h'(0)+\varepsilon)x} \mathrm{d}x = \frac{1 - \mathrm{e}^{-\delta(-h'(0)-\varepsilon)t}}{(-h'(0) - \varepsilon)t}.$$

不妨假设 $\varepsilon < |h'(0)|$, 那么上式取极限 $t \to \infty$ 可得

$$\frac{1}{-h'(0) + \varepsilon} \leq \liminf_{t \to \infty} tI_1 \leq \limsup_{t \to \infty} tI_1 \leq \frac{1}{-h'(0) - \varepsilon}$$

然后,

$$I_2 \leq \mathrm{e}^{(t-1)\eta} \int_\delta^\infty \mathrm{e}^{h(x)} \mathrm{d}x = O(\mathrm{e}^{\eta t}),$$

其中 $\eta = \sup_{x>\delta} h(x) < 0$. 因为 $\mathrm{e}^{\eta t} \ll t^{-1}$, 所以 $t \to \infty$ 时将两式相加可得

$$\frac{1}{-h'(0) + \varepsilon} \leq \liminf_{t \to \infty} tI(t) \leq \limsup_{t \to \infty} tI(t) \leq \frac{1}{-h'(0) - \varepsilon}.$$

由 ε 的任意性即得结论.

5 Lecture 6

3. 考虑 Q 过程 (转移速率矩阵 $\mathbf{Q} = ((q_{ij}))$), 则有

$$\begin{aligned} h_i(t + \Delta t) &= \mathbb{E}(f(X_{t+\Delta t}); X_0 = i) \\ &= \mathbb{E}[\mathbb{E}(f(X_{t+\Delta t})|X_{\Delta t}); X_0 = i] \\ &= \sum_j p_{ij}(\Delta t) h_j(t). && \text{(根据独立增量性)} \end{aligned}$$

所以有

$$\begin{aligned} \frac{h_i(t + \Delta t) - h_i(t)}{\Delta t} &= \frac{p_{ii}(\Delta t) - 1}{\Delta t} h_i(t) + \sum_{j \neq i} \frac{p_{ij}(\Delta t)}{\Delta t} h_j(t) \\ &\to q_{ii} h_i(t) + \sum_{j \neq i} q_{ij} h_j(t),\ \Delta t \to 0. \end{aligned}$$

若将 $\{h_i(t)\}$ 写作一个向量 $\boldsymbol{h}(t)$, 那么它满足的方程为

$$\boldsymbol{h}'(t) = \mathbf{Q}\boldsymbol{h}(t),$$

称为后向方程. 注意我们在推导过程中先走 Δt 的时间步再走 t 的时间步, 即时间增量是向后的. 另一种理解方式为: 概率转移矩阵 $\mathbf{P}(t)$ 满足

$$\mathbf{P}(t) = \mathrm{e}^{t\mathbf{Q}} \Rightarrow \mathbf{P}' = \mathbf{QP} = \mathbf{PQ},$$

所以

$$\boldsymbol{h}(t) = \mathbf{P}\boldsymbol{f} \Rightarrow \boldsymbol{h}'(t) = \mathbf{QP}\boldsymbol{f} = \mathbf{Q}\boldsymbol{h}.$$

5. (1) 当转移速率 $\lambda(t)$ 不是常数时, 记 $p_m(t)$ 为 t 时刻的分布, 则由 Poisson 过程条件 (4),

$$p_m(t+h) = p_m(t)(1-\lambda(t)h) + p_{m-1}(t)\lambda(t)h + o(h).$$

移项并取微分得到

$$p_m'(t) = -\lambda(t)p_m(t) + \lambda(t)p_{m-1}(t),$$

以及初值条件

$$p_m(0) = \delta_{m0} = \begin{cases} 1, & m = 0, \\ 0, & m \geq 1. \end{cases}$$

记 $\Lambda(t) = \int_0^t \lambda(s)\mathrm{d}s$, 则经过变换, 可得

$$[\mathrm{e}^{\Lambda(t)}p_m(t)]' = \lambda(t)\mathrm{e}^{\Lambda(t)}p_{m-1}(t),$$

所以

$$p_m(t) = \mathrm{e}^{-\Lambda(t)}\int_0^t \lambda(s)\mathrm{e}^{\Lambda(s)}p_{m-1}(s)\mathrm{d}s, \quad m \geq 1.$$

$m = 0$ 时有解 ($p_{-1}(t) \equiv 0$)

$$p_0(t) = \mathrm{e}^{-\Lambda(t)},$$

代入递推可得

$$p_m(t) = \mathrm{e}^{-\Lambda(t)}\int_0^t \lambda(s)\frac{(\Lambda(s))^{m-1}}{(m-1)!}\mathrm{d}s = \mathrm{e}^{-\Lambda(t)}\frac{(\Lambda(t))^m}{m!}, \quad \forall m,$$

即 $X_t \sim \mathcal{P}(\Lambda(t))$. 也可通过对时间 t 进行重参数化 $s = \Lambda(t)$ 来解释.

(2) 记 T_t 为所求的等待时. 由独立增量性, 从 t 时刻开始的算起的轨道 $\{X_{t+s}-X_t\}_{s\geq 0}$ 和以 $\lambda(t+s)$ 为转移速率的过程同分布. $T_t > s$ 的概率就是过程 $\{X_{t+s} - X_t\}$ 时刻该过程仍然在 0 位置的概率, 即上面小题中的 $p_m(0)$, 故直接代入可得

$$\mathbb{P}(T_t > s) = \exp\left(-\int_0^s \lambda(t+s)\mathrm{d}s\right).$$

6 Lecture 7

2. 因为 $Q(\sigma \to \sigma') = Q(\sigma' \to \sigma)$, 所以为了验证细致平衡条件, 只需验证

$$\pi(\sigma)A(\sigma \to \sigma') = \pi(\sigma')A(\sigma' \to \sigma),$$

对于所有只有一个位点不同的 σ, σ'.

若采用 Metropolis 算法, 左边等于

$$\frac{1}{Z}\mathrm{e}^{-\beta H(\sigma)}\min\{1,\mathrm{e}^{-\beta(H(\sigma')-H(\sigma))}\}=\frac{1}{Z}\min\{\mathrm{e}^{-\beta H(\sigma)},\mathrm{e}^{-\beta H(\sigma')}\},$$

由对称性, 右边显然也等于这个值, 所以结论成立.

若采用 Glauber 算法, 左边等于

$$\frac{1}{Z}\mathrm{e}^{-\beta H(\sigma)}\frac{1}{1+\mathrm{e}^{\beta(H(\sigma')-H(\sigma))}}=\frac{1}{Z}\frac{1}{\mathrm{e}^{\beta H(\sigma)}+\mathrm{e}^{\beta H(\sigma')}},$$

由对称性, 右边显然也等于这个值, 所以结论成立.

7 Lecture 9

解决此题时我们需要考虑到 $V(x)$ 极小点处可能存在的不同的阶数. 记 $V(x)$ 全局极小点的集合为 M. 假设 M 有限且 $V(x)$ 光滑, 且不妨设 $\min V(x)=0$, 那么在其中一点 $x_i\in M$ 附近 $V(x)$ 的 Taylor 展开必定为

$$V(x)=\frac{V^{(2k)}(x_i)}{(2k)!}(x-x_i)^{2k}+O((x-x_i)^{2k+1}),$$

这是因为极小值点处最低阶的一个非零导数不可能为奇数阶. 在上述 x_i 的一个邻域内, 可以仿照 Laplace 渐近公式的证法证明

$$\int_{x_i-\varepsilon}^{x_i+\varepsilon}\mathrm{e}^{-\beta V(x)}\mathrm{d}x=[C_k+o(1)](\beta V^{(2k)}(x_i))^{-1/2k},\quad \beta\to\infty,$$

其中 C_k 是由 k 决定的常数.

考虑 M 中所有点的 ε 邻域, 其中 ε 取得充分小使得它们两两不交, 则 $\mathrm{e}^{-\beta V(x)}$ 在这些邻域之并以外的积分指数下降; 在 x_i 点, 若阶数为 k_i, 那么 $\int_{B_\varepsilon(x_i)}\mathrm{e}^{-\beta V(x)}$ 的衰减阶数为 $O(\beta^{-1/2k_i})$, 所以那些最大的 k_i 附近的那些概率是主导项. 记这些点组成的集合为 $M^*\subset M, k^*=\max_i k_i$, 用渐近展开公式可以决定这些主导项之间的比例关系, 再归一化得到

$$\Pi_\beta\left(\bigcup_{x\in M^*}B_\varepsilon(x)\right)\to 1,$$

$$\Pi_\beta(B_\varepsilon(x_i))\to\frac{[V^{(2k^*)}(x_i)]^{-1/2k^*}}{\sum_{x\in M^*}[V^{(2k^*)}(x)]^{-1/2k^*}}.$$

由 ε 的任意性, Π_β 的极限为

$$\frac{\sum_{x\in M^*}[V^{(2k^*)}(x)]^{-1/2k^*}\delta_x}{\sum_{x\in M^*}[V^{(2k^*)}(x)]^{-1/2k^*}}.$$

8 Lecture 10

4. Arcsine Law 有两种形式, 分别关于最后一个零点和大于 0 的时长. 以下两个定理是来自 Durrett 的 Thm. 4.9.5 和 Thm. 4.9.6 (pp. 261–267), 请参考那里获取证明。

Theorem (关于最后一个零点的反正弦律)**.** 设 $\{S_n\}$ 是 $\mathbb{Z}$ 上的对称随机游走, 令

$$L_{2n}=\sup\{m\le 2n: S_m=0\},$$

则

$$\lim_{n\to\infty}\mathbb{P}\left(a\le\frac{L_{2n}}{2n}\le b\right)=\int_a^b\frac{1}{\pi\sqrt{x(1-x)}}\mathrm{d}x=\frac{2}{\pi}(\arcsin\sqrt{b}-\arcsin\sqrt{a}).$$

Theorem (关于大于 0 时长的反正弦律). 设 $\{S_n\}$ 是 $\mathbb{Z}$ 上的对称随机游走, 令 π_{2n} 为 $[0,2n]$ 上 x 轴上方的线段 $(k-1,S_{k-1}) \to (k,S_k)$ 的数目, 并且令 $u_m = \mathbb{P}(S_m = 0)$, 则

$$\mathbb{P}(\pi_{2n} = 2k) = u_{2k}u_{2n-2k},$$

从而

$$\lim_{n\to\infty} \mathbb{P}\left(a \le \frac{\pi_{2n}}{2n} \le b\right) = \int_a^b \frac{1}{\pi\sqrt{x(1-x)}}\mathrm{d}x = \frac{2}{\pi}(\arcsin\sqrt{b} - \arcsin\sqrt{a}).$$

还有结论 $\pi_{2n} \overset{d}{=} L_{2n}$.

9 Lecture 12

1. 根据上一讲的等价定义, 为说明某个随机过程是 Brownian motion, 只需说明:

(i) 它是 Gauss 过程.

(ii) 均值 $m(t) = 0$, 协方差 $K(s,t) = s \wedge t$.

(iii) 轨道 a.s. 连续.

对于本题的三个过程, 前两点都是容易验证的, 但很多人忘了验证第三点.

$X_t = tW_{1/t}$ 在 $t = 0$ 处的轨道连续性是不那么平凡的, 需要说明

$$\lim_{t\to\infty} \frac{W_t}{t} = 0, \quad \text{a.s..}$$

用下述关于轨道最大值点的结论.

$$\mathbb{P}\left(\max_{[0,t]} W_t \ge a\right) = \mathbb{P}(\tau_a \le t) = 2\mathbb{P}(W_t > a), \quad \forall a > 0.$$

用这个更强的上界限制 W_t/t 的取值.

$$\begin{aligned}\mathbb{P}\left(\sup_{t\in[T,2T]} \frac{|W_t|}{t} \ge \varepsilon\right) &\le \mathbb{P}\left(\max_{t\in[0,2T]} |W_t| \ge 2\varepsilon T\right)\\ &\le 4\mathbb{P}(W_{2T} > 2\varepsilon T)\\ &= C\frac{\mathrm{e}^{-\varepsilon T}}{\sqrt{T}}.\end{aligned}$$

上式对 $T_n = 2^nT$ 求和后收敛, 由 Borel-Cantelli 引理, a.s. 只出现有限次, 所以 a.s. 成立

$$\lim_{t\to\infty} \frac{|W_t|}{t} \le \varepsilon.$$

因此 $W_t/t \to 0$, a.s.. 事实上有更强的结论 (Law of the Iterated Logarithm):

$$\limsup_{t\to\infty} \frac{W_t}{\sqrt{2t\log\log t}} = 1, \quad a.s..$$

2. 序列 $Y_n(t,\omega)$ 不是一个 IID 序列的部分和, 因为 Y_{n+1} 的前 2^n 项并非来自 Y_n, 而只是同分布. 但还是可以通过 Borel-Cantelli 引理证明其 a.s. 收敛性.

$$\begin{aligned}&\mathbb{P}(|Y_n(t,\omega) - t| \ge \varepsilon) \le \varepsilon^{-2}\mathrm{Var}(Y_n(t,\omega)) = \frac{t^2}{2^n\varepsilon},\\ &\Rightarrow \sum_n \mathbb{P}(|Y_n(t,\omega) - t| \ge \varepsilon) < \infty,\\ &\Rightarrow \mathbb{P}\left(\{|Y_n(t,\omega) - t| \ge \varepsilon\} \text{ i.o.}\right) = 0,\\ &\Rightarrow Y_n(t,\omega) \to t, \text{a.s..}\end{aligned}$$

Y_n 似乎并不是单调的, 因为 $(a+b)^2 \le a^2+b^2$ 并不总是对任意 $a,b\in\mathbb{R}$ 成立.

10 Lecture 13

1. 右端点积分可直接用二次变差的性质推出. 中点积分会遇到形如

$$\sum_{j=0}^{N-1}\left(W_{t_{j+1/2}}-W_{t_j}\right)^2$$

的项, 可用证明二次变差等于 $T/2$ 的相同方法证明其极限为 $T/4$.

2. (a) 考虑函数 $f(x,u)=\varphi(x-u)=\exp\left(\frac{1}{2}(x-u)^2\right)$, 它是两变量的解析函数. 在 $(x,0)$ 点对 u 进行 Taylor 展开, 得到

$$f(x,u)=\sum_{n=0}^{\infty}\frac{u^n}{n!}\frac{\partial^n f}{\partial u^n}\bigg|_{(x,0)}.$$

另一方面, 根据 $f(x,u)=\phi(x-u)$ 的形式可知,

$$\frac{\partial^n f}{\partial u^n}\bigg|_{(x,0)}=(-1)^n\frac{\partial^n f}{\partial x^n}\bigg|_{(x,0)}=(-1)^n\frac{\partial^n(\mathrm{e}^{-x^2/2})}{\partial x^n}\bigg|_{(x,0)}=\mathrm{e}^{-x^2/2}h_n(x).$$

两边乘 $\mathrm{e}^{-x^2/2}$ 即得结论.

$a>0$ 时, 作变量替换 $u\leftarrow a^{\frac{1}{2}}u, x\leftarrow a^{-\frac{1}{2}}x$, 第二个关系式是显然的. $a\to 0$ 时, $H_n(x,a)\to H_n(x,0)$, 所以取极限可以得到对 $a=0$ 也成立 (或者直接验证).

(b) (a) 中第二个式子对 x 求二阶导数, 得到

$$\sum_{n=0}^{\infty}\frac{u^n}{n!}\partial_x^2 H_n(x,a)=u^2\mathrm{e}^{ux-\frac{au^2}{2}}.$$

对 a 求一阶导数, 得到

$$\sum_{n=0}^{\infty}\frac{u^n}{n!}\partial_a H_n(x,a)=-\frac{u^2}{2}\mathrm{e}^{ux-\frac{u^2}{2}}.$$

相加得到

$$\sum_{n=0}^{\infty}\frac{u^n}{n!}\left(\frac{1}{2}\partial_x^2+\partial_a\right)H_n(x,a)=0.$$

由于该式对任意 u 都成立, 所以左边级数中的每一项系数都是 0, 即第一个结论成立.

对 x 求一阶导数, 得到

$$\sum_{n=0}^{\infty}\frac{u^n}{n!}\partial_x H_n(x,a)=u\mathrm{e}^{ux-\frac{au^2}{2}}=u\sum_{n=0}^{\infty}\frac{u^n}{n!}H_n(x,a).$$

对比两边对应项系数, 即得第二个结论.

(c) 对 n 归纳. $n=1$ 时, $h_1(x)=x$, 所以

$$W_t=\int_0^t \mathrm{d}W_{t_1}=\frac{1}{1!}t^{\frac{1}{2}}\frac{W_t}{t^{\frac{1}{2}}},$$

结论成立.

假定对 $n-1$ 成立, 则只需证明

$$\frac{1}{n!}H_n(W_t,t)=\int_0^t\frac{1}{(n-1)!}H_{n-1}(W_s,s)\mathrm{d}W_s.$$

由 Itō 公式和前两小问的结论,

$$\mathrm{d}H_n(W_t,t) = \partial_x H_n(W_t,t)\mathrm{d}W_t + \underbrace{\frac{1}{2}\partial_x^2 H_n(W_t,t)\mathrm{d}t + \partial_a H_n(W_t,t)\mathrm{d}t}_{\text{(b) 结论 1}}$$
$$= \partial_x H_n(W_t,t)\mathrm{d}W_t = nH_{n-1}(W_t,t)\mathrm{d}W_t.$$

4. 根据 ODE 的熟知结论, $(\mathrm{e}^{t\boldsymbol{A}})' = \boldsymbol{A}\mathrm{e}^{t\boldsymbol{A}} = \mathrm{e}^{t\boldsymbol{A}}\boldsymbol{A}.$, 所以根据 Itō 公式,

$$\mathrm{d}(\mathrm{e}^{-t\boldsymbol{A}}\boldsymbol{X}_t) = -\mathrm{e}^{-t\boldsymbol{A}}\boldsymbol{A}\boldsymbol{X}_t\mathrm{d}t + \mathrm{e}^{-t\boldsymbol{A}}\mathrm{d}\boldsymbol{X}_t = \mathrm{e}^{-t\boldsymbol{A}}(-\boldsymbol{A}\boldsymbol{X}_t\mathrm{d}t + \mathrm{d}\boldsymbol{X}_t),$$

其中因为函数关于 $\boldsymbol{X}_t$ 线性所以二阶导为零. 代入原方程可得

$$\mathrm{d}(\mathrm{e}^{-t\boldsymbol{A}}\boldsymbol{X}_t) = \mathrm{e}^{-t\boldsymbol{A}}\boldsymbol{\sigma}\mathrm{d}\boldsymbol{W}_t,$$

所以它的显式解可以写成

$$\boldsymbol{X}_t = \mathrm{e}^{t\boldsymbol{A}}\boldsymbol{X}_0 + \int_0^t \mathrm{e}^{(t-s)\boldsymbol{A}}\boldsymbol{\sigma}\mathrm{d}\boldsymbol{W}_s.$$

假设 $\lambda(\boldsymbol{A}) < 0$, 那么上面的分布在 $t \to \infty$ 时有极限, 也就是稳态分布. 它同分布于

$$\int_0^\infty \mathrm{e}^{t\boldsymbol{A}}\boldsymbol{\sigma}\mathrm{d}\boldsymbol{W}_t.$$

均值显然为零, 协方差用 Itō 等距性计算.

$$\begin{aligned}\mathbb{E}_\pi(\boldsymbol{X}\boldsymbol{X}^T) &= \int_0^\infty (\mathrm{e}^{t\boldsymbol{A}}\boldsymbol{\sigma})(\mathrm{e}^{t\boldsymbol{A}}\boldsymbol{\sigma})^T\mathrm{d}t \\ &= \int_0^\infty \mathrm{e}^{t\boldsymbol{A}}\boldsymbol{\sigma}\boldsymbol{\sigma}^T\mathrm{e}^{t\boldsymbol{A}^T}\mathrm{d}t.\end{aligned}$$

5. 我们形式地把右端点的函数值写成

$$\sigma(X_{t+\Delta t}, t+\Delta t) = \sigma(X_t,t) + \sigma_x\Delta X_t + \sigma_t\Delta t.$$

展开右端点 SDE. 根据 Itō 公式, 把 $\mathrm{d}W_t$ 看成 $(\mathrm{d}t)^{\frac{1}{2}}$, 且省去所有 $\frac{3}{2}$ 阶以上的项.

$$\mathrm{d}X_t = b(X_t,t)\mathrm{d}t + [\sigma(X_t,t) + \sigma_x\mathrm{d}X_t]\mathrm{d}W_t.$$

扩散项中多出的 $\mathrm{d}X_t\mathrm{d}W_t$ 中, 只有 $\sigma(X_t,t)\mathrm{d}W_t\mathrm{d}W_t$ 的部分非零, 所以最后剩下几项为

$$\mathrm{d}X_t = [b(X_t,t) + \sigma\sigma_x(X_t,t)]\mathrm{d}t + \sigma(X_t,t)\mathrm{d}W_t.$$

11 Lecture 14

1. 使用 Einstein's summation. 首先根据上一章形式展开成 Itō 的 SDE.

$$\mathrm{d}\boldsymbol{X}_t = \boldsymbol{b}\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{W}_t + \frac{1}{2}(\nabla\sigma\cdot\mathrm{d}\boldsymbol{X}_t)\mathrm{d}\boldsymbol{W}_t.$$

写成指标形式.

$$\mathrm{d}X_t^j = b^j\mathrm{d}t + \sigma^{jk}\mathrm{d}W_t^k + \frac{1}{2}\partial_i\sigma^{jk}\mathrm{d}X_t^i\mathrm{d}W_t^k.$$

对于最后一项, 我们递归地展开 $\mathrm{d}X_t^i$, 并只保留 $(\mathrm{d}W_t^k)^2$ 的项, 得到

$$\begin{aligned}\partial_i\sigma^{jk}\mathrm{d}X_t^i\mathrm{d}W_t^k &= \partial_i\sigma^{jk}(b^i\mathrm{d}t + \sigma^{il}\mathrm{d}W_t^l)\mathrm{d}W_t^k \\ &= \partial_i\sigma^{jk}\sigma^{ik}\mathrm{d}t.\end{aligned}$$

然后我们就得到了 Itō 形式的 SDE, 可以由此写出前后向算子

$$
\begin{aligned}
\mathrm{d}X_t^j &= b^j \mathrm{d}t + \frac{1}{2}\partial_i \sigma^{jk}\sigma^{ik}\mathrm{d}t + \sigma^{jk}\mathrm{d}W_t^k \\
\mathcal{L}u &= b^j\partial_j u + \frac{1}{2}\partial_i\sigma^{jk}\sigma^{ik}\partial^j u + \frac{1}{2}\sigma^{ik}\sigma^{jk}\partial_{ij}u \\
&= b^j\partial_j u + \frac{1}{2}\sigma^{ik}\partial_i(\sigma^{jk}\partial_j u) \\
\mathcal{L}^* p &= -\partial_j(b^j p) + \frac{1}{2}\partial_j[\sigma^{jk}\partial_i(\sigma^{ik}p)]
\end{aligned}
$$

再写成矩阵向量形式, 得 Fokker-Planck 方程.

$$
p_t = \mathcal{L}^* p = -\nabla\cdot(p\boldsymbol{b}) + \frac{1}{2}\nabla\cdot(\boldsymbol{\sigma}\nabla\cdot(p\boldsymbol{\sigma}^T)),
$$

其中矩阵的散度按行求. 右端点积分的推导是类似的.

12 Lecture 15

每个强连续压缩半群 $S(t)$ 对应到一个无界算子 L, 称为它的生成元, 满足:

1. L 是稠定闭算子.
2. 以 $\mathcal{D}(L)$ 内的点为起点, 算子半群的轨道 $S(t)x$ 连续可微, 而且满足

$$
\frac{\mathrm{d}S(t)x}{\mathrm{d}t} = LS(t)x = S(t)Lx.
$$

(a) 由 $\|S(t)\| \le 1$ 知 $\|P\| \le 1$, 是压缩算子. 由半群性质, $S(t)S(t) = S(2t)$. 令 $t \to \infty$, 由 $S(t)$ 的压缩性质, 左边算子复合可以取强极限, 得 $P^2 = P$, 所以是投影.

(b) 由半群性质, $S(s+t) = S(s)S(t) = S(t)S(s)$. 令 $s \to \infty$ 得到 $P = S(t)P = PS(t)$.

由此得到 P 和 L 可交换. 此即对任意 $x \in \mathcal{D}(L)$, $0 = PLx = LPx$. 这是因为

$$
0 = \frac{\mathrm{d}}{\mathrm{d}t}[S(t)Px] = \frac{\mathrm{d}}{\mathrm{d}t}[PS(t)x] = PLx,
$$

从而可推出 $Px \in \mathcal{D}(L)$ 且 $LPx = 0$.

(c) 当 $x \in \mathcal{N}(L)$ 时, $Lx = 0$, 所以根据生成元的性质, $\frac{\mathrm{d}}{\mathrm{d}t}(S(t)x) = S(t)Lx = 0$, 所以 $S(t)x \equiv x, \forall t > 0$. 令 $t \to \infty$ 可得 $Px = x$, 故 $x \in \mathcal{R}(P)$.

当 $x \in \mathcal{R}(P)$ 时, 由 L 的稠定性, 用 $y \in \mathcal{D}(L)$ 逼近 x, 则 $Py \to Px = x$. 由 (b) 的额外结论, $LPy = 0$. 再由 L 的闭性可知 $x \in \mathcal{D}(L)$ 且 $Lx = 0$, 即 $x \in \mathcal{N}(L)$.

(d) 当 $x \in \mathcal{R}(L)$ 时, 设 $x = Ly, y \in \mathcal{D}(L)$, 那么由 (b) 的额外结论, $Px = PLy = 0$, 故 $x \in \mathcal{N}(P)$.

当 $x \in \mathcal{N}(P)$ 时, 假设 $x \notin \overline{\mathcal{R}(L)}$, 则由 Hahn-Banach 定理, 存在连续线性泛函 $f \in \mathcal{B}^*$, 使得 $f\big|_{\overline{\mathcal{R}(L)}} = 0, \langle f, x\rangle \neq 0$. 考察从 x 出发的轨道, 则 $S(t)x \to Px = 0$, 所以 $\langle f, S(t)x\rangle \to 0$. 任取 $y \in \mathcal{D}(L)$, 有

$$
\frac{\mathrm{d}}{\mathrm{d}t}\langle f, S(t)y\rangle = \frac{\mathrm{d}}{\mathrm{d}t}\langle f, LS(t)y\rangle = 0.
$$

这说明 $\langle f, S(t)y\rangle = \langle f, S(0)y\rangle = \langle f.y\rangle$, 对 $y \in \mathcal{D}(L)$. 因为 L 稠定, 所以 $S^*(t)f = f$, 对任意 $t > 0$, 进而 $\langle f, S(t)x\rangle \equiv \langle f, x\rangle \neq 0$, 矛盾!

13 Lecture 16

1. We notice that the variables are jointly Gaussian, so we only need to compute their mean and covariance. W.l.o.g., we let $t_n = 0$, and then the three variables become

$$\Delta W_n = W_{\Delta t} = \int_0^{\Delta t} \mathrm{d}W_\tau,\ \Delta Z_1 = \int_0^{\Delta t} \tau \mathrm{d}W_\tau,\ \Delta Z_2 = \int_0^{\Delta t} W_\tau \mathrm{d}\tau.$$

We also recall that by the Itō integral

$$tW_t = \int_0^t W_s \mathrm{d}s + s\mathrm{d}W_s,$$

the three quantities are correlated, i.e. $\Delta Z_2 = \Delta t \Delta W_n - \Delta Z_1$. We only need to generate $(\Delta W_n, \Delta Z_1)$.

Obviously, their means are both zero. Their variances and covariances equal

$$\begin{aligned}
\mathbb{E}(\Delta W_n)^2 &= \Delta t, \\
\mathbb{E}(\Delta Z_1)^2 &= \int_0^{\Delta t} \tau^2 \mathrm{d}\tau = \frac{1}{3}(\Delta t)^3, \\
\mathbb{E}(\Delta W_n \Delta Z_1) &= \int_0^{\Delta t} \tau \mathrm{d}\tau = \frac{1}{2}(\Delta t)^3.
\end{aligned}$$

Therefore, we only need to generate the two-dimensional Gaussian variable subject to mean 0 and covariance matrix

$$\mathbf{A} = \begin{bmatrix} \Delta t & \frac{1}{2}(\Delta t)^2 \\ \frac{1}{2}(\Delta t)^2 & \frac{1}{3}(\Delta t)^3 \end{bmatrix}.$$

2. Consider the diffusion process

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \mathrm{d}W_t,$$

i.e. the process with *constant additive noise*. Use $\{X_n^h\}$ to denote the numerical solution, and $\{X_t\}$ the exact one. We are going to prove that the Euler-Maruyama scheme

$$X_{n+1}^h = X_n^h + b(X_n^h)h + \Delta W_n^h$$

is actually of order 1 strong convergence. Here the assumption on b must be very strong, and we will show them during our proof. The goal is to prove that the L^2 error of the n-th step with $nh \leq T$ is of order $O(h^2)$, which requires the truncation error at each step to be $O(h^3)$.

As is usually done in numerical ODEs, we separate the error of each step into *propagation and accumulation*.

$$X_{n+1}^h - X_{(n+1)h} = (X_{n+1}^h - \bar{X}_{n+1}) + (\bar{X}_{n+1} - X_{(n+1)h}),$$

where $\bar{X}_{n+1}$ denotes the solution obtained by taking one forward time step from the exact solution X_{nh}.

Suppose b is Lipschitz with constant L, and then we have

$$|X_{n+1}^h - \bar{X}_{n+1}| = |(X_n^h - X_{nh}) + h(b(X_n^h) - b(X_{nh}))| \leq (1 + Lh)|X_n^h - X_{nh}|.$$

This is because we use the *same Wiener trajectory* no matter what our precision is. Next, the propagation error at the n-th step caused by discretization.

$$\begin{aligned}
\bar{X}_{n+1} - X_{(n+1)h} &= \int_{nh}^{(n+1)h} [b(X_t) - b(X_{nh})]\mathrm{d}t \\
&= \int_{nh}^{(n+1)h} \mathrm{d}t \int_{nh}^t b'(X_s)[b(X_s)\mathrm{d}s + \mathrm{d}W_s] + b''(X_s)\mathrm{d}s \\
&\triangleq Y_n + Z_n,
\end{aligned}$$

where we used Itô's formula, and denoted by

$$Y_n := \int_{nh}^{(n+1)h} \mathrm{d}t \int_{nh}^{t} (bb' + b'')|_{X_s} \mathrm{d}s,$$
$$Z_n := \int_{nh}^{(n+1)h} \mathrm{d}t \int_{nh}^{t} b'(X_s) \mathrm{d}W_s.$$

By estimating their orders we find that $\mathbb{E}(Y_n^2) = O(h^4)$, and that $\mathbb{E}Z_n^2 = O(h^3)$.

Take the conditional expectation with respect to the n-th step.

$$\begin{aligned}
&\mathbb{E}\left[(X_{n+1}^h - X_{(n+1)h})^2 \mid X_n^h, X_{nh}\right] = \mathbb{E}\left[\left(X_{n+1}^h - \bar{X}_{n+1} + Y_n + Z_n\right)^2 \mid X_n^h, X_{nh}\right] \\
&\le \mathbb{E}\left[(1+\varepsilon)\left(X_{n+1}^h - \bar{X}_{n+1} + Z_n\right)^2 + (1+\varepsilon^{-1})Y_n^2 \mid X_n^h, X_{nh}\right] \\
&\le \mathbb{E}\left[(1+\varepsilon)\left(X_{n+1}^h - \bar{X}_{n+1} + Z_n\right)^2 \mid X_n^h, X_{nh}\right] + (1+\varepsilon^{-1})C_1 h^4.
\end{aligned}$$

Now comes the critical part. Notice that

$$X_{n+1}^h - \bar{X}_{n+1} = [b(X_n^h) - b(X_{nh})]h$$

is determined by X_n^h and X_{nh}, which is considered constant in the conditional expectation. Meanwhile, Z_n is an Itō integral over W_t on $(nh, (n+1)h]$, so it has conditional expectation 0 given X_{nh}, X_n^h because of the Markovian proeprty. Therefore, we can break down the first quadratic term $(X_{n+1}^h - \bar{X}_{n+1} + Z_n)^2$ and eliminate the crossing term in between. The overall estimate then becomes

$$\begin{aligned}
&\mathbb{E}\left[(X_{n+1}^h - X_{(n+1)h})^2 \mid X_n^h, X_{nh}\right] \\
&\le (1+\varepsilon)(X_{n+1}^h - \bar{X}_{n+1})^2 + (1+\varepsilon)\mathbb{E}[Z_n^2 \mid X_n^h, X_{nh}] + (1+\varepsilon^{-1})C_1 h^4 \\
&\le \underbrace{(1+\varepsilon)(1+Lh)|X_n^h - X_{nh}|^2}_{\text{propagation}} + \underbrace{(1+\varepsilon)C_2 h^3 + (1+\varepsilon)^{-1} C_1 h^4}_{\text{truncation}}.
\end{aligned}$$

Choosing $\varepsilon = h$ makes the propagation factor

$$(1+h)(1+Lh) = 1 + O(h) = O(1),$$

and truncation error

$$C_2(1+h)h^3 + C_1\left(1 + \frac{1}{h}\right)h^4 = O(h^3),$$

so the final estimate of $\mathbb{E}|X_{nh} - X_n^h|^2$ is $O(h^2)$ as desired.

14 Lecture 18

1. Discretizing the problem into finite case, we get

$$\mathbb{E}\exp\left(\mathrm{i}\sum_j \xi(t_j)\Delta W_{t_j}\right) = \exp\left(\frac{1}{2}\sum_j (t_j - t_{j-1})\xi^2(t_j)\right)$$

since W_t has independent increments and $\mathbb{E}(\Delta W_{t_j})^2 = t_j - t_{j-1}$. Taking the formal limit $\max_j \Delta t_j \to 0$ gives us the desired result.