

Chương 6

LÝ THUYẾT KHUẾCH TÁN

§1. Quá trình khuếch tán

Các quá trình khuếch tán thực chất là các mô hình toán học của các chuyển động của các hạt riêng biệt trong quá trình này sinh ra một số chất từ chất khác và các phân tử của chúng chuyển động hỗn loạn.

Các quá trình khuếch tán cũng mô tả các hiện tượng sinh học như sự biến đổi theo thời gian của số các tế bào trong một cơ thể sinh vật, hoặc sự tập trung của các gen trong một quần thể.

Lý thuyết các quá trình khuếch tán gắn bó tự nhiên với Lý thuyết Phương trình đạo hàm riêng. Thực ra, theo quan điểm tất định, một quá trình khuếch tán là lời giải của bài toán Cauchy cho một loại phương trình đạo hàm riêng parabolic.

Theo quan điểm ngẫu nhiên, thì quá trình khuếch tán thực chất là một họ các quá trình ngẫu nhiên và là các quá trình Markov. Các quá trình này thoả mãn một phương trình vi phân ngẫu nhiên mà ta cũng gọi là phương trình khuếch tán.

Tất nhiên có một sự tương quan mật thiết giữa phương trình khuếch tán tất định và phương trình khuếch tán ngẫu nhiên.

1.1. Định nghĩa

Một họ các quá trình Markov (X_t, P_x) trên không gian $(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})$ được gọi là quá trình khuếch tán trên \mathbb{R}^n , nếu:

a/ Toán tử sinh cực vi A của quá trình Markov X xác định trên mọi hàm hữu hạn khả vi liên tục hai lần, và tồn tại hàm vector liên tục $b^i(x)$ và ma trận vector liên tục $(a^{ij}(x))$ đối xứng và xác định không âm với mọi x sao cho

$$\forall f \in C^2 : Af(x) = Lf(x) \equiv \frac{1}{2} \sum_{i,j=1}^n a^{ij}(x) \frac{\partial^2 f}{\partial x^i \partial x^j} + \sum_{i=1}^n b^i(x) \frac{\partial f}{\partial x^i}$$

b/ Toàn bộ quỹ đạo của các X_t đều là liên tục

1.2. Chú ý

a/ Từ định nghĩa trên, ta có thể nói rằng quá trình khuếch tán là một họ các quá trình Markov sao cho toán tử sinh cực vi A của nó trùng với toán tử vi phân cấp hai L .

b/ Nhắc lại toán tử sinh cực vi (infinitesimal generator) của một quá trình Markov.

Một quá trình Markov X tương ứng với một bán nhóm (P_t) xác định trên các hàm thuộc lớp C^2 bởi

$$(P_t f)(x) = \int P_t(x, dy) f(y) \text{ với } P_t(x, A) \text{ là xác suất chuyển.}$$

Khi đó toán tử sinh cực vi tương ứng A được xác định bởi

$$A = \lim_{h \downarrow 0} \frac{P_h - I}{h},$$

trong đó I là toán tử đồng nhất.

c/ Ta cũng có thể chứng minh rằng, một quá trình khuếch tán X trên \mathbb{R}^n là một quá trình với quỹ đạo liên tục $X = (X^1, X^2, \dots, X^n)$ sao cho với mọi $t \geq 0$ và $h > 0$ thì:

$$E[X_{t+h}^i - X_t^i | X_s, s \leq t] = b^i(X_t)h + o(h)$$

$$\text{và } E\{[X_{t+h}^i - X_t^i - b^i(X_t)h][X_{t+h}^j - X_t^j - b^j(X_t)h]\} = a^{ij}(X_t)h + o(h)$$

với những hàm b^i ($1 \leq i \leq n$) nào đó trên \mathbb{R}^n mà ta gọi là *hệ số dịch chuyển* và những hàm a^{ij} ($1 \leq i, j \leq n$) nào đó trên \mathbb{R}^n mà ta gọi là các *hệ số khuếch tán*.

Cũng có thể chứng minh rằng quá trình Wiener tiêu chuẩn là một quá trình khuếch tán với hệ số dịch chuyển $b \equiv 0$ và hệ số khuếch tán $a = 1$.

d/ Nếu dịch chuyển b và khuếch tán a là những hàm trơn đến một cấp nào đấy thì hàm mật độ chuyển $p_t(x, y)$ (hoặc viết gọn là $p_t(x, y)$) của quá trình khuếch tán X sẽ thỏa mãn hai phương trình đạo hàm riêng sau đây:

$$(1) \quad \frac{\partial}{\partial t} p_t(x, y) = L_x p_t(x, y) \text{ với } y \text{ cố định và}$$

$$(2) \quad \frac{\partial}{\partial t} p_t(x, y) = L_y^* p_t(x, y) \text{ với } x \text{ cố định}$$

trong đó

$$(L_x p_t)(x, y) = \frac{1}{2} \sum_{i,j} a^{ij}(x) \frac{\partial^2 p(x, y)}{\partial x^i \partial x^j} + \sum_i b^i(x) \frac{\partial p(x, y)}{\partial x^i}$$

$$\text{và } (L_y^* p_t)(x, y) = \frac{1}{2} \sum_{i,j} a^{ij}(y) \frac{\partial^2 p(x, y)}{\partial y^i \partial y^j} - \sum_i b^i(y) \frac{\partial p(x, y)}{\partial y^i}$$

Phương trình (1) được gọi là phương trình Kolmogorov lùi,
Phương trình (2) được gọi là phương trình Kolmogorov tiến.

Trong trường hợp riêng, khi quá trình khuếch tán là một quá trình Wiener, thì vì $b = 0$ và $a = (a^{ij}) = I$, ta có các phương trình lùi và tiến như sau:

$$(3) \quad \frac{\partial p}{\partial t} - \frac{1}{2} \sum_i \frac{\partial^2 p}{\partial x^{i^2}} = 0$$

$$(4) \quad \frac{\partial p}{\partial t} + \frac{1}{2} \sum_i \frac{\partial^2 p}{\partial y^{i^2}} = 0$$

Các phương trình (3) và (4) được gọi là các phương trình Fokker-Planck (lùi và tiến).

§2. Phương trình khuếch tán

Ta vẫn ký hiệu $a(x) = (a_{ij}(x))$ là ma trận các hệ số khuếch tán và $b(x) = (b^1(x), \dots, b^n(x))$ là vector các hệ số chuyển dịch. Gọi σ là ma trận, sao cho $a = \sigma \sigma^T$ với σ^T là chuyển dịch của σ .

Có thể chứng minh rằng quá trình ngẫu nhiên khuếch tán X_t định nghĩa như trên là lời giải của phương trình vi phân ngẫu nhiên Stratonovitch:

$$dX_t = b(X_t)dt + \sigma(X_t) \circ dW_t$$

$X_0 = x$; W_t là chuyển động Brown tiêu chuẩn.

Với mỗi $x \in \mathbf{R}^n$ thì lời giải X_t^x là một quá trình Markov. Cho P_t là một toán tử xác định trên các hàm liên tục giới nội trên \mathbf{R}^1 sao cho:

$$(P_t g)(x) = E[g(X_t^x)], \quad \in C_b(\mathbf{R}^n), \quad t \geq 0.$$

Có thể thấy rằng (P_t) là một nửa nhóm. Ký hiệu \mathcal{G} là toán tử sinh cực vi của (P_t) . Xét một hàm $U : [0, T] \times \mathbf{R}^n \rightarrow \mathbf{R}$ xác định bởi $U(t, x) \equiv (P_t u)(x)$ với một u nào đó $\in C_b(\mathbf{R}^n)$. Giả sử U thuộc lớp C^2 đối với x . Khi đó ta có

$$\begin{aligned} \frac{\partial U}{\partial t} &= \frac{\partial}{\partial t}(P_t u) = \lim_{h \downarrow 0} \frac{P_{t+h} u - P_t u}{h} \\ &= \lim_{h \downarrow 0} P_t \frac{P_h - I}{h} u = \mathcal{G}U \quad (= \mathcal{G}(P_t u)) \end{aligned}$$

Do đó, hàm $U = U(t, x)$ thỏa mãn bài toán Cauchy đối với phương trình đạo hàm riêng

$$\begin{cases} \frac{\partial U}{\partial t} = \mathcal{G}U \\ \lim_{t \downarrow 0} U(t, x) = u(x) \end{cases}$$

(Do tính liên tục h.c.c của X_t^x :

$$U(t, x) = E[u(X_t^x)] = E[u(X_0^x)] = E[u(x)] = u(x))$$

Nếu lưu ý tới nhận xét b. mục 3 nói trên, có thể thấy rằng toán tử P_t xác định ở đây trùng với toán tử bán nhóm của quá trình Markov X , do đó toán tử sinh cực vi của chúng cũng trùng nhau và bằng toán tử vi phân L :

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

với

$$Lf(x) \equiv \frac{1}{2} \sum_{i,j=1}^n a^{ij}(x) \frac{\partial^2 f}{\partial x^i \partial x^j} + \sum_{i=1}^n b^i(x) \frac{\partial f}{\partial x^i}$$

trong đó $a(x) = (a^{ij}(x))$ đối xứng và xác định không âm với mọi x sao cho $a = \sigma \sigma^T$ và vector $b(x) = (b^1(x), \dots, b^n(x))$ là liên tục

Nói tóm lại, với những giả thiết đó về ma trận khuếch tán a và vector dịch chuyển b thì ta có

• Phương trình khuếch tán tất định là

$$\begin{cases} \frac{\partial}{\partial t} U(t, x) = \frac{1}{2} \sum_{i,j=1}^n a^{ij}(x) \frac{\partial^2 U}{\partial x^i \partial x^j} + \sum_{i=1}^n b^i(x) \frac{\partial U}{\partial x^i} \\ U(0, x) = u(x) \end{cases}$$

với $U(t, x)$ biểu thị sự khuếch tán của vật chất theo thời gian và không gian

• Phương trình khuếch tán ngẫu nhiên là

$$\begin{cases} dX_t = b(X_t)dt + \sigma(X_t) \circ dW_t; & a = \sigma\sigma^T \\ X_0 = x \end{cases}$$

Lời giải của phương trình này là một quá trình Markov và là quá trình ngẫu nhiên khuếch tán.

§3. Phương pháp Nung luyện mô phỏng và Phương trình khuếch tán

3.1. Nung luyện mô phỏng là gì?

Nung luyện là một quá trình nhiệt, vốn được nghiên cứu trong vật lý các vật liệu cô đặc: Một vật liệu rắn được đặt trong một bể nhiệt cho đến khi bị nung chảy, sau đó người ta hạ từ từ nhiệt độ T trong bể nhiệt xuống theo một cách "khéo léo" nào đó để cho vật liệu cô đặc lại một cách mịn màng nhất, có cấu trúc cao nhất, tại đó năng lượng U của vật liệu đạt giá trị nhỏ nhất. Tại mỗi thời điểm và mỗi nhiệt độ, năng lượng U phụ thuộc vào trạng thái X của toàn bộ hệ thống nung luyện; trạng thái này là ngẫu nhiên do nhiều yếu tố, trong đó có sự chuyển động hỗn loạn của các phân tử vật chất dưới tác động của các lực ngẫu nhiên do việc cấp nhiệt và hạ nhiệt trong bể nhiệt. Trạng thái X của hệ thống là ngẫu nhiên và theo nghiên cứu của nhiệt động học, X là một biến ngẫu nhiên có phân bố Boltzmann với hàm mật độ xác suất cho bởi

$$\pi(x) = \frac{1}{Z} e^{-\frac{1}{kT} U(x)}, \quad x \in D \subset \mathbb{R}^n,$$

trong đó $U(x)$ là thế năng của hệ thống tại trạng thái x của miền D , T là một tham số (nhiệt độ của hệ thống), còn Z là một hằng số chuẩn hóa đảm bảo cho hàm π đúng là một hàm mật độ phân bố xác suất, tức là sao cho

$$\int_{\mathbb{R}^n} \pi(x) dx = 1, \text{ còn } k \text{ là hằng số Planck.}$$

Trong toán học, phân bố xác suất này còn được gọi là các *phân bố Gibbs*.

Từ biểu thức của $\pi(x)$ ta thấy rằng, $U(x)$ càng nhỏ thì mật độ $\pi(x)$ càng lớn, tức là phân bố Gibbs tập trung tại những điểm cực tiểu toàn cục của $U(x)$.

Từ nhận xét đó, người ta muốn mô phỏng quá trình nung luyện trong Vật lý, để tìm ra những thuật toán tìm cực trị toàn cục của một hàm U cho trước và những thuật toán này được gọi chung là *thuật toán Nung luyện Mô phỏng* (NLMP); thuật ngữ tiếng Anh là *Simulated Annealing* (S.A.).

Muốn vậy, người ta tìm cách thiết kế các quá trình ngẫu nhiên Markov X phụ thuộc vào thời gian t và tham số nhiệt độ T có phân bố xác suất là $\pi(x)$ nói trên sao cho:

$$\lim_{T \downarrow 0} \lim_{t \uparrow \infty} \{\text{Xác suất để cho } X \text{ đạt giá trị tối ưu}\} = 1$$

Nói cách khác, các quá trình ngẫu nhiên X đó hội tụ đến các điểm cực trị toàn cục X^* về mặt xác suất:

$$X = X^* \text{ và } U(X) = \min_{x \in D} U(x).$$

Nếu thiết kế các quá trình rời rạc, ta có xích Markov. Trong trường hợp liên tục, ta có các quá trình khuếch tán thỏa mãn một phương trình vi phân ngẫu nhiên.

Các quá trình ngẫu nhiên được thiết kế như vậy gọi là *quá trình Nung luyện Mô phỏng*.

3.2. Phương trình Langevin-Smoluchowski

Cho U là một hàm lấy giá trị thực xác định trên một miền $D \subset \mathbb{R}^n$, và cho trước một phân bố π trên $\Omega = \mathbb{R}$ cho bởi

$$\pi(x) = \frac{1}{Z_T} \exp\left\{-\frac{1}{T}U(x)\right\}$$

với T là một tham số không âm.

Để thiết kế một quá trình khuếch tán n chiều X_t , $t \geq 0$ trên Ω có π là một độ đo bất biến duy nhất, ta xét phương trình khuếch tán ngẫu nhiên sau đây, gọi là phương trình Langevin-Smoluchowski:

$$dX_t = -\nabla U(X_t)dt + \epsilon_0 dW_t,$$

trong đó ∇ là toán tử gradient, còn $\epsilon_0 = \sqrt{2T}$.

Phương trình này do Langevin đề xuất năm 1908 như là một sự mở rộng của lý thuyết Einstein về chuyển động Brown; nó mô tả chuyển động của một hạt trong một chất lỏng nhớt. (Trong chuyển động Brown; giả thiết độ nhớt là bằng 0). Phương trình Langevin-Smoluchowski là phương trình toán học đầu tiên mô tả các lực nhiệt động không cân bằng. Số hạng $\epsilon_0 dW_t$ mô tả những nhiễu loạn vi mô gây nên bởi các lực Brown, còn số hạng $-\nabla U(X_t)dt$ mô tả các lực cản sinh ra bởi độ nhớt của chất lỏng.

Trong quá trình nung luyện, vật chất chảy ra, hóa lỏng và các phân tử vật chất khuếch tán liên tục theo mô hình của phương trình Langevin-Smoluchowski.

Về mặt toán học, ta thấy phương trình Langevin-Smoluchowski là một sự mở rộng của thuật toán gradient tắt định: trong thuật toán gradient tắt định để tìm những giá trị của x làm cho $U(x)$ đạt cực tiểu, người ta không vượt qua được cực trị địa phương. Nhưng điều đó có thể làm được với phương pháp gradient ngẫu nhiên: việc thêm số hạng $\epsilon_0 dW_t$ cho phép quá trình X vượt qua được cực tiểu địa phương của hàm U .

Viết lại phương trình Langevin:

$$dX_t = -\nabla U(X_t)dt + \sqrt{2T}dW_t$$

Giả thử X_t là một lời giải của phương trình này: đó là một quá trình ngẫu nhiên khuếch tán mà ta ký hiệu mật độ phân phối là $p(t, x)$.

Khi đó, dưới một số giả thiết về tính chính quy của hàm U , người ta chứng minh được rằng $p(t, x)$ hội tụ đến phân phối Gibbs $\pi(x) = \frac{1}{Z_T} \exp \left\{ -\frac{U(x)}{T} \right\}$ khi $t \rightarrow \infty$.

Điều đó có nghĩa là lời giải X_t của phương trình Langevin-Smoluchowski hội tụ yếu (hội tụ theo luật) đến một biến ngẫu nhiên X^* là giá trị làm cho U đạt cực tiểu toàn cục, mà biến ngẫu nhiên này có phân phối Gibbs.

Chương 7

ĐIỀU KHIỂN TỐI ƯU CÁC HỆ VI PHÂN NGẪU NHIÊN

Trong 30 năm gần đây, việc nghiên cứu các hệ điều khiển có tác động ngẫu nhiên đã trở thành một hướng nghiên cứu đầy triển vọng và thu hút được nhiều nhà toán học làm việc trong lĩnh vực xác suất ngẫu nhiên.

Chương này nhằm giới thiệu một vài kết quả cơ bản của điều khiển ngẫu nhiên tối ưu một hệ ngẫu nhiên cho bởi phương trình vi phân ngẫu nhiên Itô.

§1. Xây dựng quá trình ngẫu nhiên Itô có điều khiển

1.1. Mở đầu. Cho một tập hợp mở $F \subset \mathbb{R}^n$, $Q \subset \mathbb{R}^{n+1}$ xác định bởi $Q = [0, 1] \times F$. Trong phần này ta sẽ xây dựng một quá trình Itô có điều khiển dạng

$$dz_t = g(t, z, u_t)dt + \sigma(t, z)dW_t \quad (1.1)$$

$$z(0) = z_0 \in \mathbb{R}^n; \quad z(t) \in \mathbb{R}^n, \quad W_t \in U \text{ đóng} \subset \mathbb{R}^l$$

$\{W_t, 0 \leq t \leq 1\}$ là quá trình chuyển động Brown, khả vi n chiều xác định trên không gian xác suất được lọc $(\Omega, \mathcal{A}_t, \mathcal{A}, \mu)$, $t \in [0, 1]$ và (1.1) có nghiệm trong Q .

Để bắt đầu, ta hãy xét quá trình Wiener $\{B_t, 0 \leq t \leq 1\}$ xác định trên $(\Omega, \mathcal{A}_t, \mathcal{A}, \mu)$ với $\mathcal{A}_t = \sigma(B_s, 0 \leq s \leq t) \subset \mathcal{A}$.

Xét $g \in \mathbb{R}^n$, σ là ma trận vuông cấp n gồm các hàm σ_{ij} mà miền xác định của chúng được xác định dưới đây:

$$g : [0, 1] \times C \times U \rightarrow \mathbb{R}^n; \quad \sigma_{ij} : [0, 1] \times C \rightarrow \mathbb{R}$$

trong đó C là lớp các hàm liên tục trên $[0, 1]$

Xét phương trình vi phân ngẫu nhiên

$$dz_t = \sigma(t, z)dB_t, \quad z(0) = z_0 \quad (1.2)$$

với $(\sigma_{ij}) = \sigma$ thỏa mãn các điều kiện sau:

(i) $\sigma_{ij}(t, x)$ đo được đồng thời theo hai biến và $\sigma_{ij}(t, \cdot)$ là \mathcal{F}_t -đo được với mỗi t , trong đó $\mathcal{F}_t = \sigma\{c(s) : 0 \leq s \leq t, c \in C\}$, $\mathcal{F} = \mathcal{F}_1$.

(ii) Có tồn tại quá trình z_t , $t \in [0, 1]$ thích nghi với \mathcal{F}_t và thỏa mãn phương trình vi phân ngẫu nhiên (1.2) và

$$\sum_{i,j} \int_0^1 \sigma_{ij}^2(t, z) dt < +\infty.$$

(iii) $\sigma(t, z)$ không suy biến.

Trong (ii) quá trình z_t được giả thiết là xác định duy nhất theo nghĩa sau: mọi nghiệm của (1.2) có quỹ đạo liên tục và cảm sinh cùng một độ đo xác suất trên không gian quỹ đạo (C, \mathcal{F}) . Với các điều kiện (i)-(ii) thì (1.2) xác định một độ đo xác suất P trên (C, \mathcal{F}) bởi

$$P(F) = \mu(z^{-1}(F)), \quad F \in \mathcal{F}.$$

khi đó

$$\mathcal{A}_t = z^{-1}(\mathcal{F}_t), \quad \forall t \quad (1.3)$$

bởi vì $B_t = \int_0^t \sigma^{-1} dz_0$ là quá trình chuyển động Brown.

Hàm g có các tính chất sau:

(i) $g : [0, 1] \times C \times U \rightarrow \mathbb{R}^n$ là hàm đo được đồng thời của các biến, trong đó U là tập các điều khiển và là tập Borel xác định của \mathbb{R}^l

(ii) Với mỗi (t, u) xác định $g(t, \cdot, u)$ là phù hợp với \mathcal{F}_t (1.4)

(iii) $\forall (t, z, u) : |\sigma^{-1}(t, z)g(t, z, u)| \leq g^0(\|z\|)$

trong đó $\|\cdot\|$ là chuẩn đều trong C ,

còn g^0 là hàm giá trị thực, tăng

Như vậy

$$\int_0^1 |\sigma^{-1}g|^2 dt \leq g^0(\|z\|)^2 < +\infty \quad P - a.e$$