VẤN ĐỀ XẤP XỈ NGẪU NHIÊN VÀ ỨNG DỤNG

Nguyễn Văn Thu [(1)], Hoàng Văn Bắc [(2)]

(1) Trường Đại học Quốc tế, ĐHQG-HCM

(2) Trường THPT Đức Trọng, tỉnh Lâm Đồng

(Bài nhận ngày 08 tháng 11 năm 2009, hoàn chỉnh sửa chữa ngày 22 tháng 11 năm 2010)

TÓM TẮT: *Xấp xỉ ngẫu nhiên là một công cụ vô cùng quan trọng của giải tích số. Trong bài này chúng tôi sẽ trình bày tổng quát về xấp xỉ ngẫu nhiên đồng thời cũng nêu ra một phương pháp đặc biệt của xấp xỉ ngẫu nhiên, đó là phương pháp Robbins - Monro.*

Từ khóa: *xấp xỉ ngẫu nhiên, phương pháp Robbins – Monro.*

1. CÁC VÍ DỤ THỰC TẾ

a. Để biết độ cứng của hợp kim đồng - sắt ở nhiệt độ 500^0C người ta thường xét khoảng thời gian x và $Y(x)$ là độ cứng tương ứng của hợp kim. Vấn đề được đặt ra là tìm các giá trị của x sao cho hợp kim có độ cứng trung bình α. Biết rằng các loại hợp kim khác nhau ứng với độ cứng khác nhau.

b. Xét độ nhạy của một chất nổ khi bị va chạm. Thông thường người ta thả cho nó rơi tự do từ một độ cao xác định. Biết rằng mỗi một loại chất nổ ứng với một độ cao nhất định, gọi là độ cao phát nổ.

c. Tương tự, trong việc kiểm tra thuốc trừ sâu, người ta cũng phải xác định liều lượng của các loại thuốc đối với từng loại côn trùng và mức độ sử dụng sao cho phù hợp để đạt kết quả cao trong việc bảo vệ cây trồng.

2. XẤP XỈ NGẪU NHIÊN

Các tình huống trong ví dụ thực tế nêu trên có thể vận dụng toán học để giải quyết như sau. Chọn ngẫu nhiên một giá trị x_1, sau đó quan sát giá trị $y(x_1)$ của biến ngẫu nhiên $Y(x_1)$ với kỳ vọng $M(x) = E\{Y(x_1)\}$. trong đó E kí hiệu kỳ vọng toán học và M là một hàm tăng chưa biết dạng chính xác. Ta chọn một dãy các số dương a_n giảm dần theo n, ví dụ chọn $a_n = \frac{c}{n}$, trong đó c là hằng số dương tuỳ ý. Vấn đề đặt ra là xác định giá trị của θ sao cho $M(\theta) = \alpha$. Ta thiết lập hệ thức đệ quy để tìm các giá trị x_n cho các thí nghiệm liên tiếp:

$$x_{n+1} = x_n - \frac{c}{n}\left[y(x_n) - \alpha\right]. \quad (1)$$

Giả sử ta làm được thí nghiệm thứ n và đã biết được x_n cũng như giá trị $y(x_n)$. Từ (1) ta có thể xác định giá trị cụ thể của x_n để sử dụng cho lần thí nghiệm thứ $n+1$. Ta thử kiểm tra hệ thức đệ quy này. Trường hợp đơn giản nhất, nếu $\alpha = 0$ thì (1) có dạng

$$x_{n+1} = x_n - \frac{c}{n} y(x_n) \quad (2)$$

Nếu $y(x_n) > 0$ thì $x_{n+1} < x_n$ và nếu $y(x_n) < 0$ thì $x_{n+1} > x_n$. Nếu $y(x_n)$ là dương thì giảm giá trị của x cho lần thí nghiệm thứ $n+1$ và ngược lại.

Ta sẽ nghiên cứu một ứng dụng của xấp xỉ ngẫu nhiên, đó là phương pháp Robbins - Monro được trình bày sau đây.

3. PHƯƠNG PHÁP ROBBINS - MONRO

3.1. Giải tích thích ứng - Không thích ứng

Trong việc kiểm tra thuốc trừ sâu, ta nhận thấy hiện tượng là có hoặc không có loại côn trùng mà thuốc có tác dụng. Do đó, vấn đề là để xác định phù hợp chủng loại và liều lượng thích ứng cho từng loại côn trùng. Về mặt toán học, các vấn đề này được giải quyết như sau. Xét Z là biến ngẫu nhiên với hàm phân bố M. Nếu x là số thực và $Y(x)$ là biến ngẫu nhiên sao cho:

$$Y(x) = 1 \text{ nếu } Z \le x$$
$$= 0 \text{ nếu } Z > x$$

Khi đó,

$$P[Y(x)=1] = P[Z \le x] = M(x),$$
$$P[Y(x)=0] = P[Z > x] = 1 - M(x),$$
$$E[Y(x)] = 1.M(x) + 0.(1 - M(x)) = M(x).$$

Bây giờ $Y(x)$ là một quan sát thích ứng đối với số lượng x (khối lượng thuốc trừ sâu chẳng hạn). Vấn đề là để xác định giá trị của x cho sự thích ứng α. Ta có định lí sau:

Định lí 1. *Giả sử M là một hàm phân phối và α là một số thực ứng với một số thực θ thoả mãn $M(\theta) = \alpha$. Giả sử M khả vi tại θ và $M'(\theta) > 0$. Cho x_1 là một số thực và một số nguyên dương n. Nếu*

$$X_{n+1} = X_n - \frac{1}{n}(Y_n - \alpha) \quad (3)$$

trong đó Y_n là một nghiệm ngẫu nhiên sao cho

$$P[Y_n = 1 | X_1 X_2, ..., X_n, Y_1, ..., Y_{n-1}] = M(X_n)$$
$$P[Y_n = 0 | X_1 X_2, ..., X_n, Y_1, ..., Y_{n-1}] = 1 - M(X_n)$$

Khi đó,

$$\lim_{n\to\infty} E(X - \theta)^2 = 0.$$

Do đó, dãy các biến ngẫu nhiên $\{X_n\}$ hội tụ đến θ theo bình phương trung bình và theo xác suất.

Gợi ý chứng minh:

Đặt

$$\xi_n = E(X_n - \theta)^2,$$

ta chỉ cần chứng minh $\lim_{n\to\infty} \xi_n = 0$.

3.2. Xấp xỉ ngẫu nhiên một chiều

Bây giờ ta xét câu hỏi của tình huống tổng quát trong đó Y không bị hạn chế nhận giá trị 1 hoặc 0 mà có thể nhận bất kì giá trị nào.

Định lí 2 (Dvoretzky).

Giả sử $\{\alpha_n\}, \{\beta_n\}, \{\gamma_n\}$ là các dãy số thực không âm sao cho

$$\lim_{n\to\infty} \alpha_n = 0 \quad (4)$$

$$\sum_{1}^{\infty} \beta_n < \infty \qquad (5)$$

$$\sum_{1}^{\infty} \gamma_n = \infty \qquad (6)$$

Xét θ là một số thực và T_n là các phép biến đổi đo được sao cho

$$|T_n(X_1,...,X_n) - \theta| \le \max\left[\alpha_n, (1+\beta_n)|X_n - \theta| - \gamma_n\right] \qquad (7)$$

với mọi $X_1,...,X_n$.

Xét X_1 và Y_n $(n = 1, 2, ...)$ là các biến ngẫu nhiên. Định nghĩa

$$X_{n+1} = T_n(X_1,...,X_n) - Y_n(X_1,...,X_n), \forall n \ge 1. \quad (8)$$

Theo giả thiết $E\left\{X_1^2\right\} < \infty$.

$$\sum_{n=1}^{\infty} E\left\{Y_n^2\right\} < \infty \qquad (9)$$

và

$$E\left\{Y_n \mid X_1,...,X_n\right\} = 0 \qquad (10)$$

với xác suất 1 với mọi n, suy ra

$$P\left[\lim_{n\to\infty} X_n = 0\right] = 1 \qquad (11)$$

Chứng minh: Không mất tính tổng quát, ta có thể chọn $\theta = 0$.

1. Từ (7) và (9) suy ra rằng với mọi n

$$E\left(X_n^2\right) < \infty.$$

2. Đặt $s(n)$ là dấu của

$$\left[T_n(X_1,...,X_n)\right]\left[X_n\right]$$

nếu cả 2 thừa số là khác 0, và $s(n) = 1$ nếu một trong hai thừa số bằng 0. Đặt

$$\prod(m,n) = \prod_{j=m}^{n} s(j), \ Y_n' = \prod(1,n) Y_n.$$

Khi đó, $\sum_{1}^{\infty} Y_n'$ hội tụ với xác suất 1 bởi (9) và (10). Đặt

$$Z(m,n) = \sum_{j=m}^{n} Y_j'.$$

Với $\forall \delta > 0, \forall \varepsilon > 0, \exists M_0(\delta, \varepsilon)$ sao cho

$$P\left\{\sup_{\substack{m,n \\ M \le m \le n}} |Z(m,n)| > \frac{\delta}{48}\right\} < \varepsilon / 2. \quad (12)$$

3. Đặt $d(m, m-1) = 1$,

$$d(m,n) = \prod_{j=m}^{n+1} (1+\beta_j) \text{ với } n \ge m.$$

Xét tổng

$$S(m,n) = \sum_{j=m}^{n+1} d(j,n) Y_{j-1}'$$

$$= \sum_{j+m}^{n-1} Z((m-2),(j-1))\left[d(j,n) - d(j+1,n)\right] - Y_{n-2}' d(m,n)$$

$$+ Z((m-2),(n-1)) d(n,n) + Y_n' \quad (13)$$

Khi $d(j,n) \ge d(j+1,n)$ ta thấy rằng giá trị tuyệt đối của (13) không lớn hơn

$$2\left[\sup_{\substack{j \\ m-1 \le j \le n}} |Z((m-2),(j-1))|\right](d(m,n)) + |Y_n|$$

Do đó từ (5) và (12) ta có:

với $\delta > 0, \varepsilon > 0$ tồn tại

$$M_{00}(\delta, \varepsilon) \ge M_0(\delta, \varepsilon)$$

sao cho $d(m,\infty) < 3/2$ với $m \ge M_{00}$ và

$$P\left\{ \sup_{\substack{m,n \\ M_{00} \le m \le n}} |Z(m,n)| < \frac{\delta}{48}, \sup_{\substack{m,n \\ M_{00} \le m \le n}} |S(m,n)| < \frac{\delta}{8} \right\} > 1 - \varepsilon / 2. \quad (14)$$

Ta suy ra điều phải chứng minh.

Định lí 3 (Dvoretzky)

Cho $\{\alpha_n(X_1,...,X_n\}, \{\beta_n(X_1,...,X_n)\}$ và $\{\gamma_n(X_1,...,X_n)\}$ là những dãy hàm không âm của biến số thực $X_1,...,X_n$ sao cho hàm $\alpha_n(X_1,...,X_n)$ là bị chặn đều và $\lim_{n\to\infty} \alpha_n(X_1,...,X_n) = 0$ hội tụ đều với mọi $X_1,...,X_n$; (15)

hàm $\beta_n(X_1,...,X_n)$ là đo được và $\sum_1^\infty \beta_n(X_1,...,X_n)$ là bị chặn đều và hội tụ đều theo $X_1,...,X_n$; (16) và

$\forall L > 0, \exists \gamma_n(X_1,...,X_n) \ge 0 : \sum_1^\infty \gamma_n = 0,$ (17)

$$\text{và } \left|T_n(X_1,...,X_n) - \theta\right| \le \max\{\alpha_n(X_1,...,X_n), [1 + \beta_n(X_1,...,X_n)] \,|\, X_n - \theta \,| - \gamma_n\} \quad (18)$$

cố định đều với mọi dãy $X_1,...,X_n$ thoả mãn $\sup_{n=1,2,...} |X_n| < L$, trong đó L là số dương tuỳ ý, (19) ở đây $T_n(X_1,...,X)$ là phép biến đổi đo được sao cho $X_{n+1} = T_n(X_1,...,X_n) + Y_n(X_n),\ n \ge 1$ (20) và $E(X_1^2) < \infty$ (21)

$$\sum_1^\infty E(Y_n^2) < \infty; \quad (22)$$

với xác suất 1

$$E\{Y_n \,|\, X_1,...,X_n\} = 0. \quad (23)$$

Thì $P\left\{\lim_{n\to\infty} X_n = \theta\right\} = 1$ (24)

và $\lim_{n\to\infty} E(X_n - \theta)^2 = 0.$ (25)

Chứng minh: Lấy $\theta = 0$ và δ, ε là những số dương tuỳ ý. Để chứng minh (24) ta cần chứng minh nó thoả mãn

$$P\{|X_n| < \delta, \forall n\} > 1 - \varepsilon \quad (26)$$

Lấy $M \ge M_0(\delta,\varepsilon)$ là đủ lớn thoả, với $n \ge M$, $\alpha_n < \delta / 8$. Lấy L đủ lớn thoả $L > \delta$ và

$$\underset{1\le j\le m}{Max} E\left\{X_j^2\right\} < \frac{\varepsilon L^2}{32M}. \quad (27)$$

Chúng ta lấy L ở đây là thoả mãn với (17). Nó cũng thoả mãn rằng

$$P\left\{\underset{1\le j\le M}{Max}\left|X_j \le L/4\right|\right\} > 1 - \varepsilon / 2. \quad (28)$$

Giả sử 4 điều kiện sau thoả liên hệ (4). (29)

$\left|X_m\right| \le \delta/4$ với một vài $m \ge M$. (30)

$\left|X_{m+1}\right| \ge \delta/4,\ 1 \le j \le k$. (31)

$\left|X_{m+k+1}\right| \le \delta/4$. (32)

trong đó $1 \le k \le \infty$. Khi $k = \infty$, (31) đúng khi $j \ge 1$ và (32) là rỗng (là rõ ràng khi chứng minh xong mà $k \ne \infty$). Bởi vì $\alpha_n < \delta/8$ khi $n \ge M$ và bởi vì (29), (30), (31) dẫn đến

$\left|T_{m+j}(X_{m+j})\right| > \alpha_{m+j} \quad (0 \le j \le k-1)$ (33)

$sign\left(X_{m+j+1}\right) = sign\left(T_{m+j}(X_{m+j})\right) \ (0 \le j \le k-1)$

(34)

Áp dụng (18) (với $\gamma = 0$) ta có X_{m+1} nằm giữa 0 và

$s(m)(1+\beta_m)X_m + Y_m$ (35)

Lập lại lập luận này, khi $1 \le j \le k$ thì X_{m+j} nằm giữa 0 và

$$s(m+j-1)s(m+j-2)...s(m)d(m,m+j-1)X_m$$
$$+s(m+j-1)...s(m-1)d(m+1,m+j-1)Y_m$$
$$+s(m+j-1)d(m+j-1,m+j-1)Y_{m+j-2} + Y_{m+j-1}$$

(36)

Giá trị tuyệt đối của (36) không lớn hơn

$\left|X_m\right| d(m,m+j-1) + \left|S(m+1,m+j-1)\right|$.

(37)

Vì vậy $\left|X_{m+j}\right|,\ 1 \le j \le k.$ (37)

Để chứng minh (26) ta chỉ cần chỉ ra rằng các kiểu kiện sau không thể xảy ra cả hai.

Liên hệ với (25) và (28) của (29), (39)

$\left|X_m\right| > \delta/4$ với tất cả $m \ge M$. (40)

Để chứng minh (11). Xét $k = \lim\limits_{1 \le j < \infty} \alpha_j$.

Xét N là số nguyên. Bởi vì (24), ta chỉ phải chứng minh rằng

$$\lim_{n\to\infty} E\left\{\left((|X_n| - k)^+\right)^2\right\} = 0$$

ở đây

$$(|X_n| - k)^+ = \max\left\{(|X_n| - k), 0\right\}.$$

3.3. Xấp xỉ cho quá trình tiệm cận chính quy

Trong phần này sẽ nghiên cứu dãy các biến ngẫu nhiên, nhưng chỉ tập trung vào các điều kiện yếu trên các hệ số lặp.

Định lí 4(a) (Comer). *Xét $\{X_n\}$ là một dãy xác định như dưới đây và X_1 là một biến ngẫu nhiên sao cho $E(X_1 - \theta)^2 < V^2 < \infty$, ở đây θ và V là các số thực.* Giả sử rằng

(i) $X_{n+1} = X_n - a'_n\left[Y_n(X_n) - Y_0\right]$, ở đây Y_0 là số thực bất kì và $Y_n(X_n)$ là biến ngẫu nhiên sao cho

$$E\left[Y_n(X) \mid X_n\right] = M(X_n).$$

(ii) $L \le d_n = \dfrac{M(X_n) - Y_0}{X_n - \theta} \le u$

với mọi n, ở đây L và u là những số thực thoả mãn $L < u$. Không mất tính tổng quát giả sử rằng $Y_0 = 0$.

(iii) $\sum_{1}^{\infty} a'_n = \infty$, ở đây $\{a'_n\}$ là dãy số dương.

(iv) $\lim_{n\to\infty} a'_n = a' \geq 0$.

(v) $0 \leq a' \leq \dfrac{1}{u}$.

(vi) $Z_n = Y_n(X_n) - M(X_n)$ và M là liên tục tại θ với $M(0) = 0$.

(vii) $E\left[Z_n^2\right] = k^2$,

(viii) ở đây k là số thực dương bất kì. Khi đó, $\lim_{n\to\infty}\left[E(X_n-\theta)^2\right]^{\frac{1}{2}} \leq k/L$ và $\limsup_{n\to\infty}\left[E\left[Y_n^2\right]\right]^{\frac{1}{2}} \leq \dfrac{uk}{L} + k$.

Chứng minh: Từ (iv) và (v) luôn tồn tại một số N sao cho $a'_n u < 1$ với $n > N$.

Do đó

$$\left.\begin{aligned}
&0 < 1 - a'_n u \leq 1 - a'_n d_n \leq 1 - a_n L < 1, \quad n > N \\
&\quad X_{n+1} - \theta = X_n - \theta - a'_n\left[Z_n\right] - a'_n d_n (X_n - \theta), \\
&\quad X_{n+1} - \theta = (1 - a'_n d_n)(X_n - \theta) - a'_n Z_n \ (n \geq 1), \\
&\quad \left|X_{n+1} - \theta\right| \leq (1 - a'_n d_n)\left|X_n - \theta\right| + a'_n\left|Z_n\right|, \\
&\quad E\left|X_{n+1} - \theta\right|^2 \leq (1 - a'_n L)^2 E\left(X_n - \theta\right)^2 + a'^2_n E\left(Z_n^2\right) \\
&\qquad\qquad + 2(1 - a'_n L) a'_n E\left\{| Z_n \| X_n - \theta |\right\} \\
&\qquad\qquad \leq (1 - a'_n L)^2 E\left(X_n - \theta\right)^2 + a_n^2 E\left(Z_n^2\right) \\
&\qquad\qquad + 2(1 - a'_n L)(a'_n)\left[EZ_n^2\right]^{\frac{1}{2}}\left[E(X_n - \theta)^2\right]^{\frac{1}{2}}, \\
&\left[E(X_{n+1} - \theta)^2\right]^{\frac{1}{2}} \leq (1 - a'_n L)\left[E(X_n - \theta)^2\right]^{\frac{1}{2}} + a'_n k \\
&\qquad\qquad = \left[E(X_n - \theta)^2\right]^{\frac{1}{2}} - a'_n L\left\{\left[E(X_n - \theta)^2\right]^{\frac{1}{2}} - k/L\right\}, n \geq N.
\end{aligned}\right\} \quad (41)$$

Lấy $\varepsilon > 0$ thì ta giả sử trái với giả thiết rằng

$E\left(X_n - \theta\right)^2 \geq k/L + \varepsilon.$ Lấy $n > N.$

Thay vào (41) ta có

$$\left[E\left(X_{n+1} - \theta\right)^2\right]^{\frac{1}{2}} \leq \left[E\left(X_n - \theta\right)^2\right]^{\frac{1}{2}} - a'_n L\varepsilon,$$

vì vậy

$$\left.\begin{aligned} &\sum_{N}^{n(\varepsilon)-1}\left[E(X_{n+1}-\theta)^2\right]^{\frac{1}{2}} \le \sum_{N}^{n(\varepsilon)-1}\left[E(X_n-\theta)^2\right]^{\frac{1}{2}} - L\varepsilon\sum_{N}^{n(\varepsilon)-1} a'_n,\\ &\left[E(X_{n(\varepsilon)}-\theta)^2\right]^{\frac{1}{2}} \le \left[E(X_N-\theta)^2\right]^{\frac{1}{2}} - L\varepsilon\sum_{N}^{n(\varepsilon)-1} a'_n. \end{aligned}\right\} \quad (42)$$

Nhưng khi $\sum_{1}^{\infty} a'_n = \infty$ ta có

$$L\varepsilon\sum_{N}^{n(\varepsilon)} a'_n \ge \left[E(X_N-\theta)^2\right]^{\frac{1}{2}} - k/L \quad (43)$$

Thay (43) vào (42) ta được

$$\left[E\left(X_{n(\varepsilon)}-\theta\right)^2\right]^{\frac{1}{2}} \le k/L.$$

Khi $\begin{aligned} &Y_n = d_n(X_n-\theta)+Z_n \quad (n\ge 1),\\ &|Y_n| \le u|X_n-\theta|+|Z_n| \end{aligned}$

và

$$\left[EY_n^2\right]^{\frac{1}{2}} \le u\left[E(X-\theta)^2\right]^{\frac{1}{2}} + k \le uk/L + k.$$

Định lí 4(b)(Comer). Giả sử có các điều kiện (i) đến (iv) của định lí 4(a)

(i) Xét

$\mu_{n,m} = E\left(Z_n \mid Z_{n-m}, Z_{n-m-1},\dots,Z_1\right)$, *ở đây* $m\ge 1$ *và* $n \ge m+1$.

(ii) Tồn tại một dãy các số thực $\{\xi_m\}$ *sao cho*

$$\left[E(\mu_{n,m}^2)\right]^{\frac{1}{2}} \le \xi_m,\ m\ge 1 \text{ và } n\ge m+1$$

và $\lim_{m\to\infty}\xi_m = 0$.

Do đó, tồn tại một hàm g *của* a' *sao cho*

$$\limsup_{n\to\infty} E\left(X_n-\theta\right)^2 \le g(a'),$$

$$\lim_{m\to\infty} E\left[(Y_n-Z_n)\right]^2 \le u^2 g(a')$$

và $\lim_{a'\to 0} g(a') \to 0,\ \ g(0)=0.$

Định lí 4(c). *Giả sử có các điều kiện của định lí 4(a) và 4(b). Thêm vào*

(i) $a'_n \to 0$ *khi* $n\to\infty$

ii) $\left(X_n-\theta\right)M\left(X_n\right) > 0.$

Khi đó,

$$P\left[\lim_{n\to\infty} X_n = \theta\right] = 1.$$

3.4. Lý thuyết mẫu nhỏ

Vì các phương pháp đã trình bày ở các phần trước trong thực tế chỉ áp dụng đối với các vấn đề giải quyết một cỡ mẫu xác định, câu hỏi đặt ra là lí thuyết xấp xỉ tiệm cận tốt như thế nào đối với các trường hợp cụ thể. Trong toán học, khải niệm tuyến tính rất quan trọng, ví dụ trong Kakutani là sự mở rộng của định lí điểm bất động Brauwer có quan hệ gắn với vấn đề đặt ra. Vì vậy nó được giả thiết rằng $M(X) = E\left(y(X)\right)$ là một đường thẳng với phương sai của X được chọn là một hằng số

STOCHASTIC APPROXIMATIONS AND APPLICATIONS

Nguyen Van Thu [(1)], **Hoang Van Bac**[(2)]

(1) International University,VNU-HCM

(2) Secondary School, Duc Trong, Lam Dong

ABSTRACT: *The purpose of this note is to present introductory ideas of stochastic approximation problems which stand for important aspects of numerical analysis. In particular, we illustrate by considering the Robbins – Monro method.*

Từ khóa: *Xấp xỉ ngẫu nhiên, phương pháp Robbins – Monro, biến ngẫu nhiên, hàm phân phối, nghiệm ngẫu nhiên, quá trình tiệm cận chính quy, mẫu nhỏ.*

TÀI LIỆU THAM KHẢO

[1]. M.T. Wasan, *Stochastic Approximation*, Cambridge University Press, (1969).

[2]. Vivek S. Borkar, *Stochastic Approximation,* Cambridge University Press, (2008).