XII LIETUVOS MATEMATIKOS RINKINYS
ЛИТОВСКИЙ МАТЕМАТИЧЕСКИЙ СБОРНИК 2

1972

УДК 518:512.39

О ДОСТАТОЧНЫХ УСЛОВИЯХ ПРИМЕНИМОСТИ МЕТОДА МОНТЕ-КАРЛО ДЛЯ РЕШЕНИЯ СИСТЕМ НЕЛИНЕЙНЫХ УРАВНЕНИЙ

В. Клейза

Решение систем линейных уравнений методом Монте-Карло является одной из плодотворнейших областей его применения. Эффективность метода при большом числе уравнений делает его применимым для численного решения дифференциальных и интегральных уравнений. Известные методы решения систем линейных уравнений с матрицей частного вида (собственные числа матрицы по модулю меньше единицы) берут начало в работе [1]. Существует также метод, допускающий системы с матрицей общего вида [2].

В области решения систем нелинейных уравнений возможности метода мало изучены. Применение метода в этой области было рассмотрено в работе [3], где использованы результаты статистической физики, относящиеся к блужданию группы частиц в поле, потенциал которого зависит от решаемой нелинейной системы.

В настоящей статье изучается применимость метода [2] к решению систем нелинейных уравнений. В работе [2] рассматривается система линейных уравнений:

$$\sum_{k=1}^{n} a_{ik}\, x_k = b_i, \quad i \in S = \{\ 1,\ 2,\ \ldots,\ n\ \}, \tag{1}$$

решение которой равносильно нахождению минимума квадратичной формы

$$V(\mathbf{x}) = V(x_1,\ x_2,\ \ldots,\ x_n) = \sum_{i=1}^{n} \alpha_i \Big(\sum_{k=1}^{n} a_{ik}\, x_k - b_i\Big)^2, \quad \alpha_i > 0. \tag{2}$$

Далее показывается, что математические ожидания координат случайного вектора с функцией плотности вероятностей

$$p_\alpha(\mathbf{x}) = \delta \exp\{-V(\mathbf{x})\}, \quad \delta = 1 \Big/ \int \cdots \int \exp\{-V(\mathbf{x})\}\, d\mathbf{x} \tag{3}$$

равны координатам решения системы (1). Действительно, при замене переменных $x_k = x_k^0 + y_k$ $\left(x_k^0,\ k \in S\right.$ — координаты решения системы $\left.(1)\right)$, получаем

$$\begin{aligned} \mathrm{E}x_l &= \delta \int \cdots \int x_l \exp\{-V(\mathbf{x})\}\, d\mathbf{x} = \\ &= x_l^0\, \delta \int \cdots \int \exp\{-V(\mathbf{y})\}\, d\mathbf{y} + \delta \int \cdots \int y_l \exp\{-V(\mathbf{x}^0 + \mathbf{y})\}\, d\mathbf{y} = \\ &= x_l^0, \quad l \in S, \end{aligned} \tag{4}$$

так как

$$\int \cdots \int y_l \exp \{ -V(\mathbf{x}^0+\mathbf{y}) \} \, d\mathbf{y}=0, \quad l \in S. \tag{5}$$

Расчет эмпирических $\mathrm{E}x_l$, $l \in S$ при помощи моделированных случайных векторов, завершает метод [2].

Рассмотрим теперь совместную систему нелинейных уравнений

$$F_i(x_1, x_2, \ldots, x_n)=0, \quad i \in S, \tag{6}$$

имеющую единственное решение, где F_i – непрерывные функции. Применим к системе (6) метод, описанный выше. Если применять его непосредственно, то возникает вопрос о дополнительных ограничениях на функции F_i. Можно также исследовать влияние на применимость метода значений параметров α_i, $i \in S$. Эти два вопроса и исследуются в настоящей статье.

Сначала выясним, насколько существенна в методе [2] линейность функций F_i. В нелинейном случае можно непосредственно получить (2), (3) и частично (4), то есть

$$\mathrm{E}x_l = x_l^0 + \delta \int \cdots \int y_l \exp\left\{ -\sum_{i=1}^{n} \alpha_i F_i^2(\mathbf{x}^0+\mathbf{y}) \right\} d\mathbf{y},$$

где $\mathbf{x}^0$ – решение системы (6). Это получается потому, что (5) необязательно выполняется для произвольных нелинейных функций системы (6). Заметим также, что в линейном случае интегралы соотношения (5) обращаются в нуль на множестве ограниченных областей типа

$$D=\left\{ \begin{matrix} |y_i| \leqslant R_i, \; R_i>0 \\ i \in S \end{matrix} \right\},$$

а функция $\varphi(\mathbf{y})=\exp\left\{ -\sum_{i=1}^{n} \alpha_i \left(\sum_{k=1}^{n} a_{ik} y_k \right)^2 \right\}$ обладает свойством

$$\varphi(\mathbf{y}) \equiv \varphi(-\mathbf{y}), \quad \mathbf{y} \in D. \tag{7}$$

Потребуем от нелинейных функций F_i выполнения условия

$$\int_D \cdots \int y_l \exp\left\{ -\sum_{i=1}^{n} F_i^2(\mathbf{x}^0+\mathbf{y}) \right\} d\mathbf{y}=0, \quad l \in S. \tag{8}$$

Нетрудно установить достаточные условия выполнения равенств (8).

Теорема 1. *Пусть $\Phi(\mathbf{y})$ непрерывна на замкнутой, измеримой (по Жордану) и ограниченной области G, и пусть G такова, что из $\mathbf{y} \in G$ следует $-\mathbf{y} \in G$, а $\Phi(\mathbf{y})$ обладает свойством $\Phi(\mathbf{y}) \equiv \Phi(-\mathbf{y})$, $\mathbf{y} \in G$, тогда*

$$\int_G \cdots \int y_l \, \Phi(\mathbf{y}) \, d\mathbf{y}=0, \quad l \in S.$$

Доказательство. Пусть $K_q^{(P)}$ и $K_q^{(N)}$ $(q=1, 2, \ldots, 2^{n-1})$ – квадранты n-мерной системы координат, пронумерованные так, что при замене переменных $y_k = -z_k$, $k \in S$ множество $K_q^{(P)}$ переводится в $K_q^{(N)}$. Тогда множество G можно представить в виде суммы

$$G = \bigcup_{q=1}^{2^{n-1}} D_q^{(P)} \cup \bigcup_{q=1}^{2^{n-1}} D_q^{(N)}, \quad D_q^{(P)} = K_q^{(P)} \cap G, \quad D_q^{(N)} = K_q^{(N)} \cap G.$$

Так как область G разделена гиперповерхностями нулевой меры, получаем

$$\int\limits_G \cdots \int y_l \Phi(\mathbf{y})\, d\mathbf{y} = \sum_{q=1}^{2^{n-1}} \left\{ \int\limits_{D_q^{(P)}} \cdots \int y_l \Phi(\mathbf{y})\, d\mathbf{y} + \int\limits_{D_q^{(N)}} \cdots \int y_l \Phi(\mathbf{y})\, d\mathbf{y} \right\}$$

$$= \sum_{q=1}^{2^{n-1}} \left\{ \int\limits_{D_q^{(P)}} \cdots \int y_l \Phi(\mathbf{y})\, d\mathbf{y} - \int\limits_{D_q^{(P)}} \cdots \int y_l \Phi(-\mathbf{y})\, d\mathbf{y} \right\}$$

$$= \sum_{q=1}^{2^{n-1}} \int\limits_{D_q^{(P)}} \cdots \int y_l [\Phi(\mathbf{y}) - \Phi(-\mathbf{y})]\, d\mathbf{y} = 0, \quad l \in S.$$

Следствие 1. *Если функции* F_i *системы* (6) *непрерывны в области* $D = \left\{ \begin{array}{c} |y_l| \leqslant R_l, \ l \in S \\ R_l > 0 \end{array} \right\}$, $\mathbf{x}^0$ *— решение системы* (6), *а* F_i *обладают свойством*

$$|F_i(\mathbf{x}^0 + \mathbf{y})| = |F_i(\mathbf{x}^0 - \mathbf{y})|, \quad i \in S, \tag{9}$$

то

$$\int\limits_D \cdots \int y_l \exp\left\{ -\sum_{i=1}^{n} \alpha_i F_i^2(\mathbf{x}^0 + \mathbf{y}) \right\} d\mathbf{y} = 0.$$

Доказательство. Подынтегральная функция является четной (в смысле (7)) в области D. Тогда по теореме 1

$$\int\limits_D \cdots \int y_l \exp\left\{ -\sum_{i=1}^{n} \alpha_i F_i^2 \right\} d\mathbf{y} = 0,$$

так как D является областью типа G.

Из следствия 1 можно заключить, что если функции F_i четны или нечетны (в смысле (7)) относительно решения системы, то (8) выполняется и аналогично линейному случаю можно получить соотношение (4). Это означает, что четность или нечетность функций системы относительно решения системы вообще, а линейность этих функций в частности, обеспечивает выполнение (8).

Выше указывалось, что при выполнении равенств (4) метод [2] завершается расчетом эмпирических $\mathrm{E}x_l$, т.е. методом Монте-Карло рассчитываются интегралы от функции

$$\Psi(\mathbf{y}) = \exp\left\{ -\sum_{i=1}^{n} \alpha_i F_i^2 \right\}. \tag{10}$$

Если функции F_i линейны, то при удалении точки $\mathbf{y}$ от решения системы, функция (10) быстро убывает и стремится к нулю. Поэтому интегралы (5) по симметрическим областям (типа D) мало отличаются от интегралов по асимметрическим областям, если они достаточно большие. Это свойство функции (10) позволяет применить метод Монте-Карло для расчета эмпирических $\mathrm{E}x_l$. Случайные векторы моделируются в n-мерном прямоугольнике, содержащем решение системы, но необязательно симметричном относительно решения.

В нелинейном случае указанное свойство функции (10) необязательно выполняется, а поэтому свойство (9) функций F_i, достаточное для выполнения (4), еще недостаточно для практического расчета по методу [2].

Исследуем теперь влияние значений параметров α_i, $i\in S$ на применимость метода [2] к нелинейным системам (6). Ниже доказывается, что при всех $\alpha_i\to+0$ или $\alpha_i\to+\infty$

$$\delta \int\cdots\int y_l \exp\left\{-\sum_{i=1}^{n}\alpha_i F_i^2(\mathbf{x}^0+\mathbf{y})\right\}\mathrm{d}\mathbf{y}\to 0,\quad D=\left\{\begin{matrix}|y_l|\leqslant R_l,\ R_l>0\\ l\in S\end{matrix}\right\},\tag{11}$$

т.е. выполняется (в пределе) равенство (5) и при этом не требуется четности функций F_i и, конечно, линейности этих функций. Если обозначить

$$p_\alpha(\mathbf{y})=\delta\cdot\exp\left\{-\sum_{i=1}^{n}\alpha_i F_i^2(\mathbf{x}^0+(\mathbf{y})\right\}=$$
$$=\frac{\exp\left\{-\sum\limits_{i=1}^{n}\alpha_i F_i^2\right\}}{\int\limits_D\cdots\int \exp\left\{-\sum\limits_{i=1}^{n}\alpha_i F_i^2(\mathbf{x}^0+\xi)\right\}\mathrm{d}\xi},\tag{12}$$

то будет действительна следующая теорема.

Теорема 2. *Пусть F_i непрерывны в области D, тогда*

$$\lim_{\substack{\alpha_i\to+0\\ \forall i\in S}} p_\alpha(\mathbf{y})=\frac{1}{2^n\prod\limits_{l=1}^{n}R_l}\quad\text{и}\quad p_\alpha(\mathbf{y})\rightrightarrows\frac{1}{2^n\prod\limits_{l=1}^{n}R_l}.$$

Доказательство. Заметим, что $p_\alpha(\mathbf{y})$ непрерывна и положительна в области D. Действительно,

$$\exp\left\{-\sum_{i=1}^{n}\alpha_i F_i^2\right\}>0,\quad\text{а}\quad \int\limits_D\cdots\int\exp\left\{-\sum_{i=1}^{n}\alpha_i F_i^2\right\}\mathrm{d}\xi>0,$$

так как подынтегральная функция строго положительна. Далее

$$\lim_{\substack{\alpha_i\to+0\\ \forall i\in S}}\exp\left\{-\sum_{i=1}^{n}\alpha_i F_i^2\right\}=1,$$

тогда по n-мерному аналогу теоремы Дини

$$\exp\left\{-\sum_{i=1}^{n}\alpha_i F_i^2\right\} \rightrightarrows 1.$$

Отсюда получаем

$$\lim_{\substack{\alpha_i\to+0\\ \forall i\in S}} p_\alpha(\mathbf{y}) = \frac{1}{\int\cdots\int\limits_D d\xi} = \frac{1}{2^n\prod\limits_{l=1}^{n} R_l}.$$

Далее получим

$$p_\alpha(\mathbf{y}) \rightrightarrows 1/2^n \prod_{l=1}^{n} R_l,$$

что завершает доказательство.

Следствие 2. *При обозначениях теоремы* 2 *справедливо*

$$\lim_{\substack{\alpha_i\to+0\\ \forall i\in S}} \int\cdots\int\limits_D y_k p_\alpha(\mathbf{y})\,d\mathbf{y} = 0, \quad k\in S.$$

Доказательство. Из теоремы 2 имеем

$$p_\alpha(\mathbf{y}) \rightrightarrows 1/2^n \prod_{i=1}^{n} R_l,$$

тогда

$$y_k p_\alpha(\mathbf{y}) \rightrightarrows y_k/2^n \prod_{l=1}^{n} R_l.$$

Далее

$$\lim_{\substack{\alpha_i\to+0\\ \forall i\in S}} \int\cdots\int\limits_D y_k p_\alpha(\mathbf{y})\,d\mathbf{y} = \frac{1}{2^n\prod\limits_{l=1}^{n} R_l}\int\cdots\int\limits_D y_k\,d\mathbf{y} = 0.$$

Из доказанного следствия 2 можно заключить, что из $\alpha_i\to+0$ следует выполнение (в пределе) равенств (4) для произвольных непрерывных F_i. С другой стороны, область D симметрична относительно решения системы (6). Заметим здесь, что знание области, симметричной относительно решения, эквивалентно знанию решения этой системы. То есть, при $\alpha_i\to+0$ $\forall i\in S$ выполнение условия (4) еще не является достаточным для применимости метода [2].

Рассмотрим, наконец, случай $\alpha_i\equiv\alpha\to+\infty$ $\forall i\in S$. Если сохранить обозначения теоремы 2, то справедлива следующая теорема.

Теорема 3. *Пусть F_i непрерывны в замкнутой и измеримой области B, содержащей начало координат в качестве внутренней точки. Тогда*

$$\lim_{\substack{\to+\infty\\ \forall i\in S}} \int\cdots\int\limits_B y_l p_\alpha(\mathbf{y})\,d\mathbf{y} = 0, \quad l\in S.$$

Доказательство. Рассмотрим сначала функцию $p_\alpha(\mathbf{y})$ в области $B^* = B \backslash S_\varepsilon$, где S_ε — n-мерный открытый шар с центром в начале координат (единственное решение системы (6)) и радиусом $\varepsilon > 0$, таким, что $S_\varepsilon \subset B$. Докажем, что

$$\lim_{\alpha \to +\infty} p_\alpha(\mathbf{y}) = 0, \quad \mathbf{y} \in B^*.$$

Для этого запишем

$$p_\alpha(\mathbf{y}) = 1 \Big/ \int \cdots \int_B \exp\{-\Delta(\zeta, \mathbf{y})\}\, d\zeta,$$

где

$$\Delta(\zeta, \mathbf{y}) = \alpha_i \sum_{i=1}^{n} [F_i^2(\zeta + \mathbf{x}^0) - F_i^2(\mathbf{y} + \mathbf{x}^0)].$$

Разделим B на три непересекающиеся части $B = B_1 \cup B_2 \cup B_3$, где $B_1 = \{\zeta \in B, \Delta(\zeta, \zeta, \mathbf{y}) > 0\}$, $B_2 = \{\zeta \in B, \Delta(\zeta, \mathbf{y}) = 0\}$, $B_3 = \{\zeta \in B, \Delta(\zeta, \mathbf{y}) < 0\}$. Тогда, поскольку Δ на множестве B^* не достигает своей точной нижней грани, $\operatorname{mes} B_3 > 0$, $\forall \mathbf{y} \in B^*$. Из этого следует

$$\lim_{\alpha \to +\infty} \int \cdots \int_{B_1} \exp\{-\Delta(\zeta, \mathbf{y})\}\, d = \infty,$$

$$\lim_{\alpha \to +\infty} \int \cdots \int_{B_2} \exp\{-\Delta(\zeta, \mathbf{y})\}\, d\zeta \geqslant 0,$$

$$\lim_{\alpha \to +\infty} \int \cdots \int_{B_3} \exp\{-\Delta(\zeta, \mathbf{y})\}\, d\zeta \geqslant 0,$$

т. е. $\lim\limits_{\alpha \to +\infty} p_\alpha(\mathbf{y}) = 0$, $\mathbf{y} \in B^*$. Учитывая, что $p_\alpha(\mathbf{y}) > 0$ непрерывна, и применяя n-мерный аналог теоремы Дини, получаем

$$p_\alpha(\mathbf{y}) \rightrightarrows 0, \quad \mathbf{y} \in B^*.$$

Покажем, что $\lim\limits_{\substack{\alpha_i \to +\infty \\ \forall i \in S}} p_\alpha(0) = \infty$. Получаем

$$\lim_{\substack{\alpha_i \to +\infty \\ \forall i \in S}} p_\alpha(0) = 1 \Big/ \lim_{\substack{\alpha_i \to +\infty \\ \forall i \in S}} \int \cdots \int_B \exp\left\{-\sum_{i=1}^{n} \alpha_i F_i^2(\zeta + \mathbf{x}^0)\right\} d\zeta = \infty.$$

Таким образом, функция $p_\alpha(\mathbf{y})$ при $\alpha \to +\infty$, $\forall i \in S$ образует δ-последовательность [4], а предельная функция является обобщенной δ-функцией. Тогда по свойству δ-функции

$$\lim_{\alpha \to +\infty} \int \cdots \int_B y_i\, p_\alpha(\mathbf{y})\, d\mathbf{y} = 0.$$

Из доказанной теоремы можно заключить, что при $\alpha_i \to +\infty$ $\forall i \in S$, соотношение (5) выполняется (в пределе) независимо от свойств функций F_i. Следует отметить, что интегрирование здесь можно производить по любой области, содержащей решение системы. Это обстоятельство делает теорему 3 существенной в исследуемом направлении.

Институт физики и математики
Академии наук Литовской ССР

Поступило в редакцию
14.VI.1971

Литература

1. G. E. Forsythe, R. A. Leibler, Matrix inversion by a Monte-Carlo method, Mathem. tables and other aids to comp., IV, 31 (1950).
2. Ю. А. Шрейдер, Решение систем линейных алгебраических уравнений по методу Монте-Карло, Вопросы теории математических машин, Сб. I, 1958.
3. T. Tsuda, T. Kiyono, Application of the Monte-Carlo method to systems of nonlinear algebraic equations, Numerische Mathematik, **6** (1964).
4. В. Я. Арсенин, Математическая физика, М., „Наука“, 1966.

APIE PAKANKAMAS NETIESINIŲ LYGČIŲ SISTEMŲ SPRENDIMO MONTE-KARLO METODU SĄLYGAS

V. Kleiza

(*Reziumė*)

Straipsnyje nagrinėjamas vieno Monte-Karlo metodo taikymas netiesinių lygčių sistemų sprendimui. Nagrinėjamas Monte-Karlo metodas, kurį galima taikyti tiesinių algebrinių lygčių sistemų sprendimui. Įrodomos kelios pakankamos sąlygos šio metodo taikymui netiesinių lygčių sistemų sprendimui.

OF SUFFICIENT CONDITIONS OF THE APPLICATION OF THE MONTE-CARLO METHOD TO SYSTEMS OF NONLINEAR EQUATIONS

V. Kleiza

(*Summary*)

The paper deals with the solution of the systems of nonlinear equations by a Monte-Carlo method. The Monte-Carlo method applicable to systems of linear algebraic equations is considered. Some sufficient conditions for the application of this method to systems of nonlinear equations have been proved.