I. Mátyáš, Random Optimization, *Avtomat. i Telemekh.*, 1965, Volume 26, Issue 2, 246–253

1965

УДК 62-505

СЛУЧАЙНАЯ ОПТИМИЗАЦИЯ

И. МАТЫАШ

(*Пардубице, ЧССР*)

Предлагается метод адаптивной случайной оптимизации нескольких параметров. Доказывается сходимость случайной последовательности состояний системы к оптимальному. Описывается схема случайного оптимизатора.

Теория оптимальных систем привлекает в последнее время внимание многих специалистов [1—3]. Однако только простейшие задачи можно решить аналитически, так как при этом требуется проведение громоздких вычислений. Поэтому для решения задач оптимизации в большинстве случаев используются цифровые или моделирующие вычислительные устройства.

В рассматриваемой ниже системе оптимальные параметры можно определить различными методами. Различают методы с неизменной программой и методы, программу которых можно тем или другим способом изменять (методы адаптивной оптимизации). В настоящей статье описывается новый метод адаптивной оптимизации, в котором используются случайные изменения параметров (случайные опыты).

1. Постановка задачи

Пусть S — физическая система или процесс (рис. 1), причем
$\mathbf{u} = [u_1, u_2, \ldots, u_r]$ — вектор ее входных величин,
$\mathbf{v} = [v_1, v_2, \ldots, v_s]$ — вектор выходных величин,

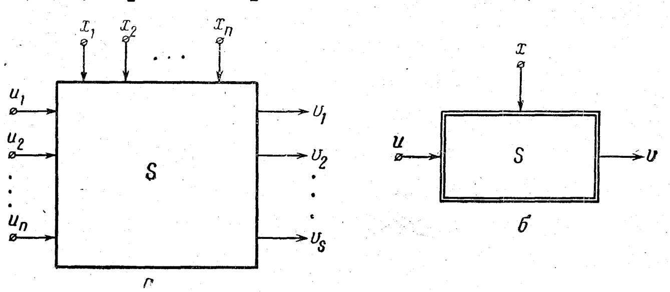

Рис. 1

$\mathbf{x} = [x_1, x_2, \ldots, x_n]$ — состояние системы S (вектор параметров, при помощи которых можно управлять поведением системы S).

Требуется определить вектор $\mathbf{x}$ (т. е. значения параметров x_i), который является оптимальным состоянием системы S. При этом необходимо определить критерий оптимальности Q (т. е. указать, с какой точки зрения состояние системы S считается оптимальным). Критерий Q, вообще говоря, является функционалом от вектора $\mathbf{v}$. Так как поведение S зависит от параметров $\mathbf{x}$, то критерий Q является функцией от вектора $\mathbf{x}$ (от состояния S).

Оптимальным состоянием S будем считать вектор $\mathbf{x} = \mathbf{x}_{\text{опт}}$, который обращает функцию Q в минимум (максимум)

$$Q(\mathbf{x}_{\text{опт}}) \leqslant Q(\mathbf{x}) \quad \text{для всех } x \in \Omega, \tag{1}$$

где Ω — область пространства X, в которой определена функция $Q(\mathbf{x})$. Для простоты предположим, что функция Q имеет единственный минимум.

Задачу сформулируем следующим образом: необходимо определить состояние x системы S, которое минимизировало бы функционал Q. Наибольшее распространение при решении подобных задач получили метод градиента и метод наискорейшего спуска. Настоящая статья посвящена описанию нового метода случайной оптимизации.

2. Основы метода случайной оптимизации

Обозначим через $\mathbf{x}^{(1)}$ начальное состояние системы и вычислим критерий Q в этом состоянии: $Q(\mathbf{x}^{(1)})$.

Пусть ξ — n-мерный нормальный случайный вектор с нулевым средним значением и единичной корреляционной матрицей. Реализации этого вектора обозначим через $\xi^{(i)}$: $\xi^{(1)}, \xi^{(2)}, \xi^{(3)}, \ldots$

Вычислим Q в случайном состоянии $\mathbf{x}^{(1)} + \xi^{(1)}$ и назовем соответствующий случайный шаг (случайный опыт) удачным, если для заранее избранного $\varepsilon > 0$

$$Q(\mathbf{x}^{(k)} + \xi^{(k)}) < Q(\mathbf{x}^{(k)}) - \varepsilon \qquad (k = 1, 2, \ldots), \tag{2}$$

что обозначим

$$y^{(k)} = 1. \tag{3}$$

Аналогично случайный шаг назовем неудачным, если

$$Q(\mathbf{x}^{(k)} + \xi^{(k)}) \geqslant Q(\mathbf{x}^{(k)}) - \varepsilon \qquad (k = 1, 2, \ldots), \tag{4}$$

что обозначим

$$y^{(k)} = 0. \tag{5}$$

В зависимости от того, каким был первый случайный шаг — удачным или нет, определим состояние $\mathbf{x}^{(2)}$ системы S после первого случайного шага следующим образом:

$\mathbf{x}^{(2)} = \mathbf{x}^{(1)}$ при $y^{(1)} = 0$ (при неудачном шаге),

$\mathbf{x}^{(2)} = \mathbf{x}^{(1)} + \xi^{(1)}$ при $y^{(1)} = 1$ (при удачном шаге).

Затем вычислим критерий Q в случайном состоянии $\mathbf{x}^{(2)} + \xi^{(2)}$ и т. д. При k-м случайном шаге вычислим прежде всего $Q(\mathbf{x}^{(k)} + \xi^{(k)})$ и по (2)—(5) установим, был ли проведенный случайный шаг удачным. Далее определим следующее состояние $\mathbf{x}^{(k+1)}$ системы S

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} \text{ при } y^{(k)} = 0, \tag{6}$$

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \xi^{(k)} \text{ при } y = 1. \tag{7}$$

Продолжая эти шаги, далее получим случайную последовательность состояний $\{\mathbf{x}^{(k)}\}$.

Основным является вопрос о сходимости этой случайной последовательности $\{\mathbf{x}^{(k)}\}$ к оптимальному состоянию $\mathbf{x}_{\text{опт}}$. На этот вопрос дает ответ теорема сходимости, сформулированная и доказанная в Приложении.

Описанным способом продолжаем случайные шаги до тех пор, пока оптимальное состояние $\mathbf{x}_{\text{опт}}$ не будет определено с достаточной точностью. По доказанной теореме это всегда возможно.

В большинстве простых случаев вероятность удачного шага близка к 0,5 всюду за исключением малой окрестности состояния $\mathbf{x}_{\text{опт}}$. Следовательно, в этих случаях в среднем каждый второй случайный шаг будет удачным (см. [4, 5]).

Описанный метод будем называть простой случайной оптимизацией.

3. Адаптивная случайная оптимизация

Процесс сходимости последовательности $\{\mathbf{x}^{(k)}\}$ к точке $\mathbf{x}_{\text{опт}}$ минимума функции Q можно ускорить следующим образом: заменим случайный вектор $\boldsymbol{\xi}$ дискретным стохастическим векторным процессом $\boldsymbol{\zeta}(k)$ с переменным средним значением и переменной корреляционной матрицей.

Процесс $\boldsymbol{\zeta}(k)$ определим, например, при помощи случайного вектора $\boldsymbol{\xi}$ следующим образом:

$$\boldsymbol{\zeta}(k) = \mathbf{d}^{(k)} + \mathbf{T}^{(k)}\boldsymbol{\xi}, \tag{8}$$

где $\mathbf{d}^{(k)}$ есть среднее значение $\boldsymbol{\xi}^{(k)}$

$$M\{\boldsymbol{\zeta}^{(k)}\} = \mathbf{d}^{(k)}, \tag{9}$$

$\mathbf{T}^{(k)}$ — матрица преобразования.

Реализацию $\boldsymbol{\zeta}^{(1)}(k)$ этого процессса определим в виде

$$\boldsymbol{\zeta}^{(1)}(k) = \mathbf{d}^{(k)} + \mathbf{T}^{(k)}\boldsymbol{\xi}^{(k)}. \tag{10}$$

Вектор средних значений $\mathbf{d}^{(k)}$ и матрицу преобразования $\mathbf{T}^{(k)}$ можно определить различными способами. Для ускорения процесса нахождения минимума функции Q определим $\mathbf{d}^{(k)}$ и $\mathbf{T}^{(k)}$, принимая во внимание результаты предыдущих шагов:

$$\mathbf{d}^{(k)} = F_d(\boldsymbol{\zeta}^{(1)}, y^{(1)}; \boldsymbol{\zeta}^{(2)}, y^{(2)}; \ldots; \boldsymbol{\zeta}^{(k-1)}, y^{(k-1)}), \tag{11}$$

$$\mathbf{T}^{(k)} = F_T(\boldsymbol{\zeta}^{(1)}, y^{(1)}; \boldsymbol{\zeta}^{(2)}, y^{(2)}; \ldots; \boldsymbol{\zeta}^{(k-1)}, y^{(k-1)}). \tag{12}$$

4. Определение вектора средних значений $\mathbf{d}^{(k)}$

Из указанного выше следует, что на первом случайном шаге, сделанном из начального состояния $x^{(1)}$, лучше всего принять вектор средних значений $\mathbf{d}^{(1)}$ нулевым:

$$\mathbf{d}^{(1)} = 0. \tag{13}$$

Это следует также и из формулы (11).

Принимая во внимание результат первого случайного шага, определим вектор $\mathbf{d}^{(2)}$ следующим образом:

$$\mathbf{d}^{(2)} = c_1\boldsymbol{\zeta}^{(1)}(1),$$

где $c_1 > 0$ при $y^{(1)} = 1$, $c_1 \leqslant 0$ при $y^{(1)} = 0$.

Другими словами, если шаг $\boldsymbol{\zeta}^{(1)}(1)$ был удачным, то среднее значение второго шага выбирается таким, чтобы движение происходило в том же направлении, а если $\boldsymbol{\xi}^{(1)}(1)$ был неудачным, то среднее значение второго шага или остается нулевым, или выбирается таким, чтобы обеспечить движение в обратном направлении. В результате выбираем именно то направление, в котором ожидается уменьшение Q.

Аналогично определим вектор средних значений $\mathbf{d}^{(k+1)}$ для дальнейших случайных шагов

$$\mathbf{d}^{(k+1)} = c_0\mathbf{d}^{(k)} + c_1\boldsymbol{\xi}^{(1)}(k), \tag{14}$$

где c_0, c_1 удовлетворяют следующим условиям:

$$0 \leqslant c_0 < 1, \quad c_1 > 0, \quad c_0 + c_1 > 1 \text{ при } y^{(k)} = 1, \tag{15}$$

$$0 \leqslant c_0 < 1, \quad c_1 \leqslant 0, \quad |c_0 + c_1| < 1 \text{ при } y^{(k)} = 0. \tag{16}$$

Найденное описанным выше способом среднее значение повышает вероятность удачного шага, так как оно совпадает с ожидаемым направлением уменьшения Q. Оно также удлиняет случайные шаги и тем самым значительно ускоряет процесс нахождения оптимального состояния $\mathbf{x}_{\text{опт}}$.

В процессе определения средних значений $\mathbf{d}^{(k)}$ приобретается опыт.

В выражении для среднего значения следующего шага последний случайный шаг входит с коэффициентом c_1. Назовем c_1 коэффициентом облучения.

При неудачных шагах можно положить $c_1 = 0$. В этом случае последовательность $\mathbf{d}^{(k)}$ представляет собою геометрическую прогрессию с отношением c_0 ($|c_0| < 1$), общий член которой стремится к нулю (накопленный опыт забывается, так как он оказался плохим). Поэтому назовем c_0 коэффициентом забывания.

Формулы (14)—(16) определяют последовательность $\mathbf{d}^{(k)}$ средних значений, и, таким образом, они эквивалентны функции F_d.

5. Ограничение величины случайного шага

Если случайные шаги оказываются удачными, то их длина растет в среднем по геометрической прогрессии с коэффициентом $c_0 + c_1 > 1$. Такой быстрый рост представляет известное неудобство и поэтому его желательно ограничить. Введем ограничение.

Пусть D — тот отрезок, длиннее которого случайный шаг не должен быть. Далее определим описанным выше способом все случайные шаги длиной меньше D, т. е. $\|\boldsymbol{\zeta}^{(1)}(k)\| \leqslant D$. Если длина некоторого шага превышает D ($\|\boldsymbol{\zeta}^{(1)}(k)\| > D$), то шаг $\boldsymbol{\zeta}^{(1)}(k)$ заменяется шагом того же направления, но длины D. Иными словами вектору $\boldsymbol{\zeta}^{(1)}(k)$ приписывается то же направление, что и раньше, но длина его ограничивается отрезком D. Величину D можно сделать переменной, определив тот или иной закон ее изменения.

6. Определение матрицы $\mathbf{T}^{(k)}$

Выбирая соответствующим образом матрицу $\mathbf{T}^{(k)}$, можно еще больше увеличить вероятность появления удачных шагов. Определение этой матрицы по результатам предыдущих случайных шагов в общем случае является довольно сложной проблемой. Рассмотрим ее при некоторых упрощениях.

Предположим, что $\mathbf{T}^{(k)}$ представляет собою диагональную матрицу. Это соответствует случаю, когда отклонения $\boldsymbol{\zeta}(k)$ от его среднего значения, т. е. слагаемые вектора $\boldsymbol{\eta} = \mathbf{T}^{(k)}\boldsymbol{\xi}$ не коррелированы. Тогда вместо n^2 коэффициентов матрицы $\mathbf{T}^{(k)}$ надо определить только n ее диагональных коэффициентов. Эти диагональные коэффициенты соответствуют стандартным отклонениям слагаемых вектора $\boldsymbol{\eta} = \mathbf{T}^{(k)}\boldsymbol{\xi}$.

Дальнейшее упрощение связано с представлением матрицы $\mathbf{T}^{(k)}$ в виде

$$\mathbf{T}^{(k)} = b_k \mathbf{E}, \tag{17}$$

где $b_k > 0$, а $\mathbf{E}$ — единичная матрица. При помощи коэффициента b_k управляем в этом случае стандартными отклонениями всех слагаемых вектора $\boldsymbol{\eta} = \mathbf{T}^{(k)}\boldsymbol{\xi}$ одновременно. Для более точного нахождения $\mathbf{x}_{\text{опт}}$ коэффициент b_k уменьшают по мере приближения $\mathbf{x}^{(k)}$ к $\mathbf{x}_{\text{опт}}$.

7. Блок-схема случайного оптимизатора

Все описанные выше операции, производимые в процессе адаптивной случайной оптимизации, осуществляются автоматически при помощи автоматического случайного оптимизатора.

Блок-схема этого оптимизатора приведена на рис. 2. Оптимизатор работает следующим образом. В запоминающих устройствах M_3 запоминается состояние системы $\mathbf{x}^{(k)}$, а в запоминающем устройстве M_4 — соответствующее значение критерия $Q(\mathbf{x}^{(k)})$. На вход M_3 подается случайный шаг $\boldsymbol{\zeta}^{(1)}(k)$. На выходе устройства M_3 появляется случайное состояние $\mathbf{x}^{(k)} + \boldsymbol{\zeta}^{(1)}(k)$ и в соответствии с этим система S переводится в новое состояние

$\mathbf{x}^{(k)} + \zeta^{(1)}(k)$. Блок K вычисляет значение $Q(\mathbf{x}^{(k)} + \zeta^{(1)}(k))$, которое сравнивается в управляющем логическом блоке L с предыдущим значением $Q(\mathbf{x}^{(k)})$. Далее определяется $y^{(k)}$ — т. е. устанавливается, случайный шаг $\zeta^{(1)}(k)$ был удачным или нет. При помощи величины $y^{(k)}$ логический блок L управляет блоками L_1, L_2, M_3, M_4 так, чтобы они осуществляли требуемые операции на следующем шаге. В конце случайного шага генератор

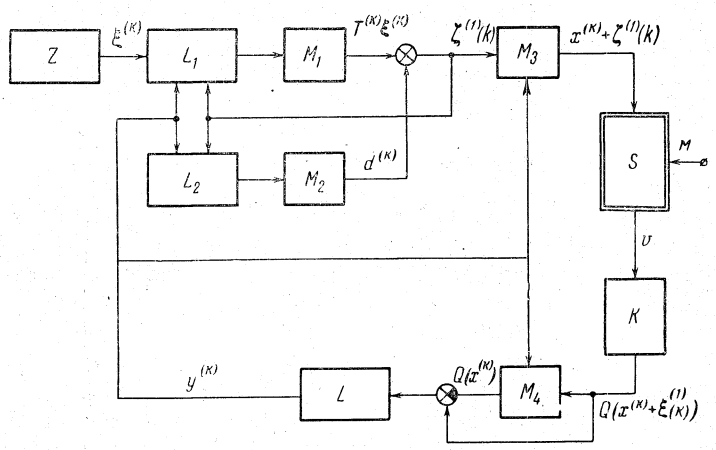

Рис. 2

случайных величин Z выдает случайный вектор $\xi^{(k+1)}$. В блоке L_1 этот вектор умножается на матрицу преобразования $\mathbf{T}^{(k)}$, а результат умножения подается на вход запоминающего устройства M_1. В блоке L_2 образуется среднее значение $\mathbf{d}^{(k+1)}$ для следующего шага; оно сохраняется в запоминающем устройстве M_2. Следующий случайный шаг находится как результат сложения выходных величин блоков M_1, M_2. Затем описанный цикл повторяется.

Логический блок L осуществляет управление другими блоками оптимизатора с тем, чтобы они производили операции, требуемые для адаптивной случайной оптимизации.

8. Реализация случайного оптимизатора

В предыдущих разделах описан метод адаптивной случайной оптимизации. Сравнивая этот метод с другими методами оптимизации нескольких параметров, видим, что он имеет более простую логику, и его можно легко реализовать при помощи вычислительных машин.

Для реализации автоматического случайного оптимизатора требуются следующие вычислительные блоки: генератор случайных величин, запоминающие устройства, логические цепи для выделения удачных шагов от неудачных, релейные и управляющие цепи. Все эти блоки можно относительно просто реализовать при помощи цифровых или аналоговых вычислительных машин.

9. Пример

Для экспериментального исследования сходимости случайной последовательности $\{\mathbf{x}^{(k)}\}$ к точке $\mathbf{x}_{\text{опт}}$ на моделирующем устройстве была собрана схема адаптивного случайного оптимизатора. Опишем один из нескольких примеров, решенных на этой модели.

Критерием оптимальности является эллиптический параболоид с отношением длин полуосей 1 : 5

$$Q(x_1, x_2) = 0{,}26x_1^2 + 0{,}26x_2^2 - 0{,}48x_1x_2.$$

Следовательно, $\mathbf{x}_{\text{опт}} = 0$, $Q(0, 0) = 0$.

Процесс адаптивной случайной оптимизации был осуществлен несколько раз из начального состояния $\mathbf{x}^{(1)} = (15, 30)$, для которого $Q(15, 30) = 76{,}5$. Модель закончила процесс оптимизации при $Q < 0{,}2$, т. е. точка $\mathbf{x}_{\text{опт}}$ была определена с достаточной степенью точности. Среднее число (из 10 процессов оптимизации) всех требуемых случайных шагов равно 93, а среднее число удачных шагов равно 55.

Нетрудно видеть, что использование предложенного метода значительно ускорило процесс оптимизации.

10. Заключение

В описанном методе адаптивной случайной оптимизации нескольких параметров с целью повышения вероятности следующих удачных шагов используются предыдущие результаты для формирования средних значений $\mathbf{d}^{(k)}$ и матрицы преобразования $\mathbf{T}^{(k)}$ вектора случайных шагов. За счет увеличения вероятности удачных шагов процесс оптимизации ускоряется.

Сходимость случайной последовательности состояний $\{\mathbf{x}^{(k)}\}$ к оптимальному состоянию $\mathbf{x}_{\text{опт}}$ для метода простой случайной оптимизации доказана в Приложении. Новые операции, применяемые в методе адаптивной случайной оптимизации, могут ускорить или замедлить процесс оптимизации, но не могут сделать процесс расходящимся. Таким образом, теорема в Приложении доказывает также сходимость по вероятности метода адаптивной случайной оптимизации к оптимальному состоянию $\mathbf{x}_{\text{опт}}$.

Автоматический оптимизатор, использующий этот метод, накапливает опыт прошлых случайных шагов — он обучается, приспосабливается к локальным свойствам критерия Q в соответствующем состоянии $\mathbf{x}^{(k)}$.

Таким образом, метод адаптивной случайной оптимизации определяет принцип обучения (накопление опыта) при помощи вычислительных машин. Новый и хороший опыт закрепляется и сохраняется, а старый и плохой опыт забывается. Этот принцип можно использовать и для других целей, например, для конструирования адаптивных и других сложных логических систем, для решения различных сложных технических задач при помощи вычислительных машин и т. д.

ПРИЛОЖЕНИЕ

Теорема сходимости и ее доказательство

При заданном $\varepsilon > 0$ [см. (2), (4)] вероятность удачного шага зависит только от вида функции Q и от состояния x, из которого делается случайный шаг.

Уравнение

$$Q(\mathbf{x}) = k \qquad (k = \text{const}) \tag{18}$$

является уравнением гиперповерхности, которая является границей области

$$G[k] = \{\mathbf{x} : Q(\mathbf{x}) < k\}. \tag{19}$$

Обозначим через $f(\mathbf{u})$ плотность распределения вероятности случайного вектора ξ и предположим, что $f(\mathbf{u}) \neq 0$ для всех u.

Вероятность удачного шага $P_Q(\mathbf{x})$ из состояния x [см. (2)—(5)] можно выразить в следующем виде:

$$P_Q(\mathbf{x}) = \int\limits_{G[Q(\mathbf{x}) - \varepsilon]} f(\mathbf{u} - \mathbf{x})\, d\mathbf{u}. \tag{20}$$

Область $G[k]$ зависит от вида функционала Q. Предположим, что Q таково, что $G[k]$ имеет следующее свойство.

Свойство А. Для любого $k > Q(\mathbf{x}_{\text{опт}})$ $G[k]$ является непустой областью и, следовательно, имеет размерность n.

Теперь сформулируем теорему.

Теорема сходимости. Пусть функция $Q(\mathbf{x})$ имеет в Ω единственный минимум, пусть $G[k]$ обладает свойством А и пусть плотность вероятности $f(\mathbf{u}) \neq 0$ для всех $\mathbf{u}$; последовательность состояний $\{\mathbf{x}^{(k)}\}$, которая получается в процессе простой случайной оптимизации [см. (6), (7)] стремится по вероятности к оптимальному состоянию $\mathbf{x}_{\text{опт}}$, т. е. к точке минимума Q.

Доказательство. Символом $\delta(\mathbf{x})$ обозначим δ-окрестность точки $\mathbf{x}$, т. е. совокупность точек $\mathbf{u}$, расстояние которых от точки x меньше δ $(\delta > 0)$

$$\delta(\mathbf{x}) = \{\mathbf{u} : \rho(\mathbf{u}, \mathbf{x}) < \delta\}. \tag{21}$$

Требуется доказать, что для любого $\delta > 0$ вероятность того, что $\rho(\mathbf{x}^{(k)}, \mathbf{x}_{\text{опт}}) > \delta$, т. е. что $\mathbf{x}^{(k)} \notin \delta(\mathbf{x}_{\text{опт}})$ стремится к нулю, т. е.

$$P\{\rho(\mathbf{x}^{(k)}, \mathbf{x}_{\text{опт}}) > \delta\} = P\{\mathbf{x}^{(k)} \notin \delta(\mathbf{x}_{\text{опт}})\} \to 0 \tag{22}$$

для $k \to \infty$.

Легко убедиться в том, что для любого $\delta > 0$ существует $\varepsilon(\delta) > 0$ такое, что

$$P_Q(\mathbf{x}) \geqslant \alpha, \quad \alpha > 0, \quad \mathbf{x} \notin \delta(\mathbf{x}_{\text{опт}}). \tag{23}$$

Если обозначить через Q_δ минимальное значение Q, то на границе $\delta(\mathbf{x}_{\text{опт}})$ будет

$$Q_\delta > Q(\mathbf{x}_{\text{опт}}).$$

Определим $\varepsilon(\delta)$ следующим образом

$$0 < \varepsilon(\delta) < Q_\delta - Q(\mathbf{x}_{\text{опт}}). \tag{24}$$

Для всех точек $x \notin \delta(\mathbf{x}_{\text{опт}})$ справедливо

$$Q(\mathbf{x}) - \varepsilon > Q(\mathbf{x}_{\text{опт}}).$$

Из свойства А следует, что $G[Q(\mathbf{x}) - \varepsilon]$ является непустой областью размерности n. Так как $f(\mathbf{u}) > 0$ для всех $\mathbf{u}$, то видим, что существует $\alpha > 0$ такое, что верно соотношение (23).

Используя значение Q в начальном состоянии, обозначим

$$\frac{Q(\mathbf{x}^{(1)}) - Q_\delta}{\varepsilon} = t \tag{25}$$

и выберем первое натуральное число, меньшее t

$$m = [t]. \tag{26}$$

Если в процессе случайной оптимизации хотя бы $m+1$ случайных шагов окажутся удачными, то все следующие точки последовательности $\{x^{(k)}\}$ принадлежат окрестности $\delta(\mathbf{x}_{\text{опт}})$.

Следовательно, вероятность $P\{\mathbf{x}^{(k)} \notin \delta(\mathbf{x}_{\text{опт}})\}$ меньше или равна вероятности того, что число удачных шагов не превосходит m

$$P\{\mathbf{x}^{(k)} \notin \delta(\mathbf{x}_{\text{опт}})\} \leqslant P\left\{\sum_{i=1}^{k} y^{(i)} \leqslant m\right\}. \tag{27}$$

Эта последняя вероятность увеличивается с уменьшением вероятности удачных шагов и, так как $P_Q(\mathbf{x}) \geqslant \alpha$, то $\mathbf{x} \notin \delta(\mathbf{x}_{\text{опт}})$ отвечает известной теореме Ньютона (о вероятности биномиального распределения)

$$P\left\{\sum_{i=1}^{k} y^{(i)} \leqslant m\right\} \leqslant \sum_{i=0}^{m} \binom{k}{i} \alpha^i (1-\alpha)^{k-i}, \tag{28}$$

где k — число сделанных случайных шагов. Далее при $k > 2m$ и $\alpha < 0{,}5$

$$\sum_{i=0}^{m} \binom{k}{i} \alpha^i (1-\alpha)^{k-i} < (m+1)\binom{k}{m}(1-\alpha)^k =$$

$$= \frac{m+1}{m!} k(k-1)(k-2)\ldots(k-m+1)(1-\alpha)^k < \frac{m+1}{m!} k^m (1-\alpha)^k. \tag{29}$$

Следовательно,

$$P\{\rho(\mathbf{x}^{(k)}, \mathbf{x}_{\text{опт}}) > \delta\} < \frac{m+1}{m!} k^m (1-\alpha)^k. \tag{30}$$

Для $\alpha > 0$ очевидно

$$\lim_{k \to \infty} k^m (1-\alpha)^k = 0, \tag{31}$$

и тем самым теорема доказана.

Для простоты доказательства здесь ограничились функциями Q с единственным минимумом. Утверждение теоремы остается верным и в случае Q с несколькими локальными минимумами. Ограничение $f(\mathbf{u}) > 0$ для всех $\mathbf{u}$ также можно отбросить. В соответствии с видом функции f утверждение теоремы будет справедливым либо для отдельных локальных минимумов, либо для абсолютного минимума. Функция $f(\mathbf{u})$ должна быть ненулевой только в некоторой окрестности точки $\mathbf{u} = 0$.

Нетрудно убедиться в том, что теорема остается верной и при отыскании максимума.

Поступила в редакцию
13 сентября 1963 г.

Цитированная литература

1. Фельдбаум А. А. Вычислительные устройства в автоматических системах. Физматгиз, 1959.
2. Фельдбаум А. А. Автоматический оптимизатор. Автоматика и телемеханика, т. XIX, № 8, 1958.
3. Стаховский Р. И. Многоканальный автоматический оптимизатор. Автоматика и телемеханика, т. XIX, № 8, 1958.
4. Matyáš J. Metody vyšetřování spojitých fysikálních systému a jejich optimální regulace. SNTL, Praha, 1963.
5. Šilhánek J., Barvíř M., Beneš J., Matyáš J. Základní metody optimalizace. Výzkumná zpráva Tesly Pardubice, ÚVR Opočínek, č. 18 L ACG 500.

RANDOM OPTIMIZATION

J. MÁTYÁŠ

An adaptive random optimization method for several variables is suggested. Convergence of random sequence of states to an optimal state is proved and a circuit of a random optimization is described.