Стохастический процессор для определения экстремума функции регрессии

Свистунов С. Г.
Петербургский государственный университет путей сообщения Императора Александра I
Санкт-Петербург, Россия,
ssg47@mail.ru

Аннотация. **В измерительной технике, радиолокации, гидроакустике часто требуется решить адаптивную задачу идентификации, т. е. построить модель по имеющимся данным об объекте. Примерами могут служить задачи статистической оценки параметров, регрессии, робастное оценивание, рекуррентное оценивание, анализ данных и т. д. В статье предлагается использовать для решения указанной задачи адаптивный квазиградиентный алгоритм, реализуемый на стохастическом вычислительном устройстве. Предлагаемый подход дает возможность сократить время решения и упростить используемые аппаратные средства. В статье приведен подробный алгоритм функционирования специализированного вычислительного устройства.**

Ключевые слова: **экстремум функции регрессии, адаптивные алгоритмы стохастической оптимизации, стохастические вычислительные устройства, преобразователь код-вероятность, ПКВ, СтВУ.**

Введение

В измерительной технике, радиолокации, гидроакустике часто требуется решить адаптивную задачу идентификации [1], т. е. построить модель по имеющимся данным об объекте. Примерами могут служить задачи статистической оценки параметров, регрессии, робастное оценивание, рекуррентное оценивание, анализ данных и т. д.

Решение этих задач можно свести к алгоритму поиска экстремума функции регрессии следующим образом [1]: для задач оценки параметров x^* плотности распределения вероятностей $w(z, x)$ по реализациям случайной величины z вводится функция

$$f(x) = MQ(z,x) = \int Q(z,x)w(z,x^*)dz,$$

где M – математическое ожидание, $Q(z,x) = -\ln[w(z,x)]$.

Очевидно, что в точке x^* минимум функции $f(x)$:

$$f(x) = -\int \left[\ln w(z,x^*) \right] w(z,x^*)dz -$$
$$-\int \ln\{1+[w(z,x)-w(z,x^*)]/w(z,x^*)\}w(z,x^*)dz \geq$$
$$\geq f(x^*) - \int \left[w(z,x) - w(z,x^*) \right] dz = f(x^*),$$

поскольку $-\ln(1+a) \geq -a$.

Пусть измерения z^1, ..., z^s поступают последовательно, x^s – найденное после обработки s измерений приближение для x^*. Тогда можно считать $\nabla Q(z^{s+1}, x^s)$ приближением для $\nabla f(x^s) = M\nabla_x Q(z,x^*)$ и применять метод градиентного типа для минимизации $f(x)$:

$$x^{s+1} = x^s - \rho_s \nabla_x Q(z^{s+1}, x^s) =$$
$$= x^s - \rho_s \left[\nabla_x w(z^{s+1}, x^s) \right] / w(z^{s+1}, x^s).$$

Скорость сходимости порядка $O(1/s)$, где $O(b)$ – величина одного порядка малости с b.

Оценка дисперсии сигнала с нулевым средним

Для оценки дисперсии сигнала с нулевым средним можно использовать следующую рекуррентную формулу [2]:

$$\sigma_{s+1}^2 = \sigma_s^2 + \rho_s (z_s^2 - \sigma_s^2),$$

обычно ρ_s=1/s [3].

Дисперсия с нулевым средним – это математическое ожидание квадрата амплитуды сигнала. Приведенная выше формула требует вычисления z^2_s, шаговый множитель $\rho_s = 1/s$ является детерминированным. Для измерения доступна величина z_s, не составляет труда определить $|z|= (\text{sign}(z)\sqrt{|z|})^2 = y^2$, где sign(z) – знак z.

Полагая функцию плотности распределения $w(z,x)$ четной, получим [4]

$$My = \int (\text{sign}(z)\sqrt{|z|})w(z,x)dz = 0,$$

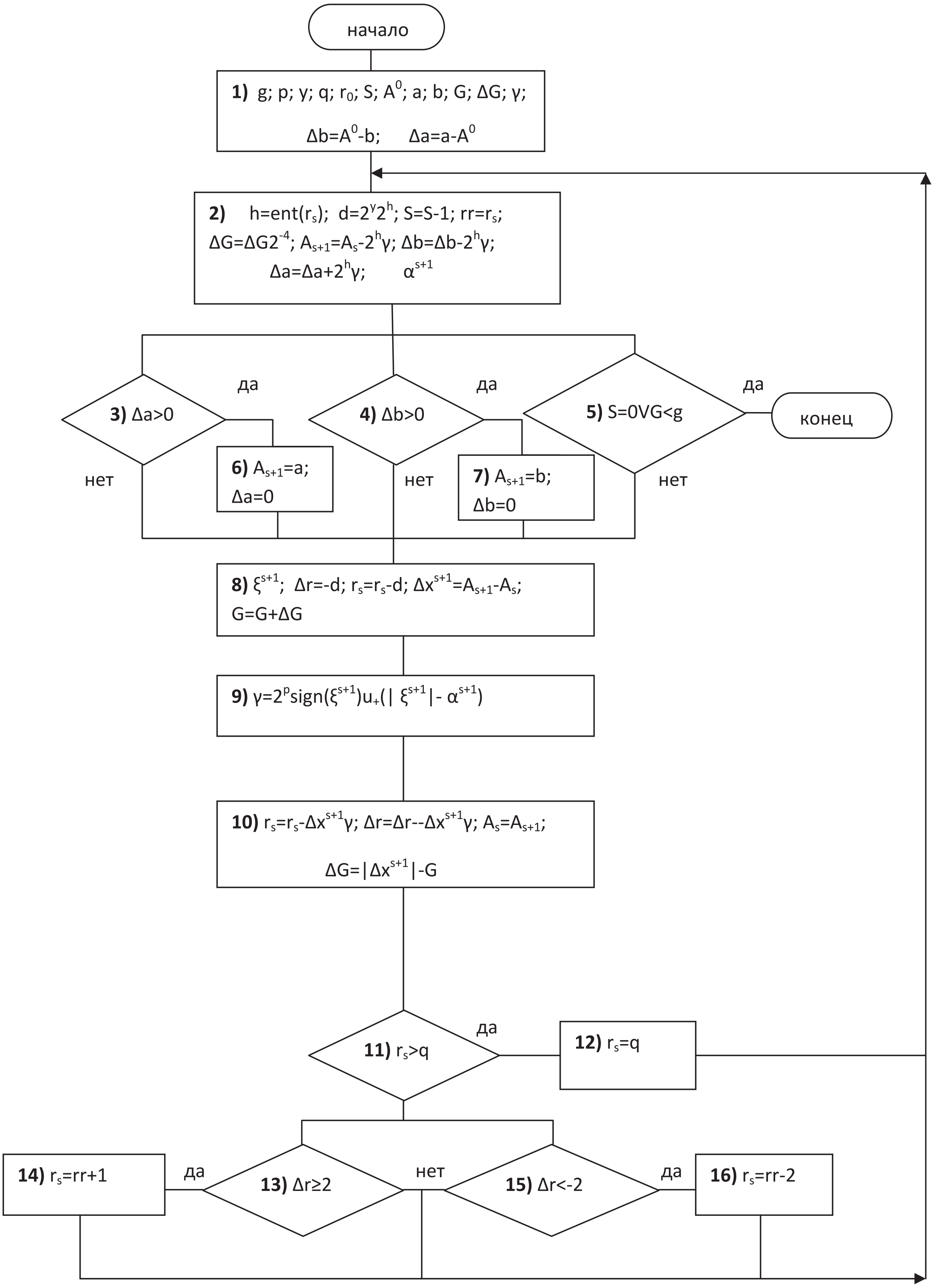
т. к. $\text{sign}(z)\sqrt{|z|}$ – нечетная функция.

Таким образом, измеряя дисперсию величины y, т. е. $My^2 = M|z|$, мы измерим $M|z|$, где $|z|$ – амплитуда сигнала. Обозначим $M|z|=A$, тогда $A_{s+1} = A_s + \rho_s(|z_s| - A_s)$, где A_s – оценка A на шаге s. Осталось определить величину ρ_s. Этот пример показывает, что шаговый множитель $\rho_s = 1/s$ является детерминированным и не учитывает в должной мере расстояния до x^*. Следовательно, для ускорения сходимости вычислительного процесса требуется адаптивная подстройка шагового множителя ρ_s, т. е. необходимо использовать адаптивные алгоритмы стохастической оптимизации [5].

В общем случае шаговый множитель определяется по формулам:

$$h = -(A_s - |z_s|)(A_s - A_{s-1}) - \delta\rho_{s-1};$$
$$\rho_s = \min\left[\rho_0, \rho_{s-1}a^h\right],$$

где $\rho_0 > 0$, $\delta > 0$, $a > 1$. Вычисление ρ_s требует таких «трудоемких» операций, как умножение и возведение в степень вещественных чисел. Наличие случайных сигналов на входе позволяет отказаться от традиционных вычислительных устройств, где информация представлена многоразрядными двоичными числами, и перейти к использованию стохастических вычислительных устройств (СтВУ), где входная информация преобразуется в импульсы, вероятностные характеристики которых зависят от входного сигнала. В результате преобразования код-вероятность (ПКВ) значительно упрощается аппаратура и уменьшается время выполнения отдельной итерации [6].



Алгоритм адаптивной квазиградиентной оптимизации

Пусть $\xi^s = A_s - |z_s|$ и $|\xi^s| < c = 2^p$ почти наверное (п. н.), тогда на выходе ПКВ имеем $\mathrm{sign}(\xi s)u + (|\xi s| - \alpha S)$, где u_+ – функция Хевисайда; α – равномерно распределенная случайная величина в диапазоне [0, c].

Используя общие теоремы о сходимости рекуррентных стохастических алгоритмов [5], можно доказать, что если выполняется $|\xi^s| < c = 2^p$ п. н., p – целое, $\delta > 0$, $q > 0$, $r_0 = 0$, $\gamma^s = c\,\mathrm{sign}(\xi^s)u_+(|\xi^s| - \alpha^S)$, то последовательность $A_{s+1} = A_s - 2^k\gamma^s$, $s = 0, 1, \ldots$, где $k = \mathrm{ent}(r_s)$ – целая часть числа r_s, $r_{s+1} = \min(q, r_s - \gamma^{s+1}(A_{s+1} - A_s) - \delta 2^k)$, сходится к A.

Адаптивность параметра ρ_s позволяет быстро достичь значения A, а применение СтВУ дает возможность использовать только короткие операции: пересылки, сдвиг, алгебраическое сложение, сравнение чисел, так как

$$\gamma^s \in \{-2p, 0, 2p\},$$

где p – целое.

Алгоритм

Алгоритм численного метода приведен на рисунке. В блоке 1 задаются начальные значения величин:

- p, где $2^p = c > |\xi^s|$ п. н.;
- y, где $\delta = 2^y > 0$;
- $[a, b]$ – границы области для A_s;
- q – ограничение на степень шагового множителя, обычно $q = 2$;
- r_0 – начальное значение степени шагового множителя, обычно $r_0 = 0$;
- A^0 – начальная координата;
- g – условие окончания счета;
- G – средняя величина модуля сдвига, начальное значение $G = 10g$;
- ΔG – приращение модуля сдвига, обычно начальное значение $\Delta G = 0$;
- γ – значение на выходе линейного преобразователя код-вероятность (ЛПКВ), начальное значение $\gamma^0 = 0$;
- S – общее количество итераций;
- $\Delta b = A - b$, $\Delta a = a - A$ – отклонение A от границы области.

В блоке 2 вычисляются значения:

- случайной величины α^{s+1}, равномерно распределенной в диапазоне [0, c];
- добавки к степени шага итерации d;
- новое значение координаты A_{s+1};
- новое значение модуля сдвига ΔG;
- из содержимого счетчика числа итераций S вычитается 1;
- величина отклонения координат (Δa и Δb) от границы области;
- запоминается текущее значение r_s.

В блоках 3, 4, 6, 7 определяется проекция точки на область для A.

В блоке 5 проверяется условие окончания вычислений.

В блоке 8 определяются величины:

- $\xi^{s+1} = A_{s+1} - |z_{s+1}|$;
- приращения координат Δx^{s+1};
- приращение шага степени Δr и новое значение степени шага итерации r_s, новое значение среднего модуля сдвига G.

В блоке 9 осуществляется линейное преобразование код-вероятность (ЛПКВ) и определяется значение γ.

В блоке 10 вычисляются величины:

- степень шага итерации r_s;
- приращение степени шага итерации Δr;
- новое значение A_s;
- новое значение приращения модуля сдвига ΔG.

В блоках 11 и 12 происходит ограничение степени шага итерации и переход на блок 2.

В блоках 13, 14, 15, 16 осуществляется ограничение прироста степени шага итерации с учетом того, что степень r_s не должна меняться за один шаг слишком быстро, введено ограничение $|\Delta r| < 2$. На основании алгоритма строится функциональная схема СтВУ.

Выводы

Использование стохастического вычислительного устройства позволяет исключить «медленные» операции: возведение в вещественную степень, умножение и деление многоразрядных чисел. Это позволило сократить время выполнения отдельных итераций, что при случайном сигнале на входе устройства уменьшает общее время определения экстремума функции регрессии.

Литература

1. Поляк Б. Т. Введение в оптимизацию / Б. Т. Поляк. – М. : Наука, 1983. – 384 с.

2. Левин Б. Р. Теоретические основы статистической радиотехники. Кн. 2 / Б. Р. Левин. – М. : Сов. радио, 1968. – 503 с.

3. Невельсон М. Б. Аппроксимация стохастическая и рекуррентное оценивание / М. Б. Невельсон, Р. З. Хасминский. – М. : Наука, 1972. – 304 с.

4. Королюк И. С. Справочник по теории вероятностей и математической статистике / И. С. Королюк, А. В. Скороход. – М. : Наука, 1985. – 640 с.

5. Урясьев С. П. Адаптивные алгоритмы стохастической оптимизации и теории игр / С. П. Урясьев. – М. : Наука, 1990. – 183 с.

6. Федоров Р. Ф. Стохастические преобразователи информации / Р. Ф. Федоров, В. В. Яковлев, Г. В. Добрис. – Л. : Машиностроение, 1978. – 304 с.

The stochastic processor for definition of an extremum of function of regress

Svistunov S. G.
The Petersburg State University
means of communication of Emperor Alexander I
St. Petersburg, Russia,
ssg47@mail.ru

Abstract. **The summary. In the measuring technique, a radar-location, hydro acoustics often it is required to solve an adaptive problem of identification, i. e. to construct model on available data about object. As examples problems of a statistical estimation of parameters can serve, to regress, the analysis of data etc. In article is offered to be used for the decision of the specified problem by the adaptive quasigradient algorithm realized on a stochastic computer. The offered approach gives the chance to reduce time of the decision and to simplify used hardware. In the article the detailed algorithm of a specialized computer is resulted.**

Use of a stochastic computer allows to exclude "slow" operations: erection in material degree, multiplication and division of multidigit numbers. **It has allowed to reduce time of performance of separate iterations, that at a casual signal on a device input reduces general time of definition of an extremum of function of regress.**

Keywords: **an extremum of function of regress, adaptive algorithms of stochastic optimization, stochastic computers, the converter a code-probability, CCP, StC.**

References

1. Poliak B. T. *Vvedeniye v optimizatsiyu* [Introduction in optimization]. Moscow, Nauka, 1983, 384 p.
2. Levin B. R. *Teoreticheskiye osnjvy statisticheskoy radiotekhniki* [Theoretical of a basis of a statistical radio engineering], Book 2. Moscow, Sovetskoye Radio, 1968, 503 p.
3. Nevelson M. B., Hasminsky R. Z. *Approksimatsiya stokhasticheskaya i rekurrentnoye otsenivaniye* [Stochastic approximation and recurrent evaluation]. Moscow, Nauka, 1972, 304 p.
4. Koroljuk I. S., Skorokhod A. V. *Spravochnik po teorii veroyatnostey i matematicheskoy statistike* [Fastwalker director on probability theory and the mathematical statistics]. Moscow, Nauka, 1985, 640 p.
5. Uryas'ev S. P. *Adaptivnyye algoritmy stokhasticheskoy optimizatsii i teorii igr* [Adaptive algorithms of stochastic optimization and the theory of games]. Moscow, Nauka, 1990, 183 p.
6. Fedorov R. F., Jakovlev V. V., Dobris G. V. *Stokhasticheskiye preobrazovateli informatsii* [Stochastic information converters]. Leningrad, Mashinostroyeniye, 1978, 304 p.