РОССИЙСКАЯ АКАДЕМИЯ НАУК
СИБИРСКОЕ ОТДЕЛЕНИЕ
АВТОМЕТРИЯ
2003, том 39, № 1

УДК 62-504

**Б. А. Беседин**
*(Омск)*

# ПРОЦЕДУРА КИФЕРА – ВОЛЬФОВИЦА, ОПТИМАЛЬНАЯ ПО ИНТЕГРАЛЬНОМУ КРИТЕРИЮ РИСКА

Процедура Кифера – Вольфовица поиска экстремума непрерывной выпуклой функции находится решением задачи стохастически оптимального управления с экспериментами при непрерывных измерениях с аддитивными погрешностями. Устанавливаются свойства аналитического решения.

**Введение.** Многочисленные адаптивные процессы поиска экстремума непрерывной функции регрессии представляют собой различные варианты и обобщения процедуры стохастической аппроксимации [1–3]. Базовой и исторически первой является градиентная в конечных разностях процедура Кифера – Вольфовица [1].

Основной целью исследований в теории стохастической аппроксимации является математический анализ поведения рекуррентно-генерируемой последовательности в зависимости от той или иной априорной информации о минимизируемой функции. В конечном счете в заданных условиях задачи устанавливаются ограничения на коэффициенты рекуррентности, обеспечивающие некоторую асимптотическую вероятностную сходимость последовательности к точке экстремума и соответствующую среднюю скорость сходимости [1–4].

Однако практическое применение процедуры Кифера – Вольфовица для оптимизации технологических объектов оставляет открытым вопрос о потерях, связанных с «раскачиванием» объекта при оценках градиентов в конечных разностях. Естественно возникает задача выбора поисковой последовательности как функции оценок градиентов в конечных разностях, которая минимизировала бы интегральный риск – математическое ожидание интегральных потерь на конечном временном интервале $[0,T]$. Формализация такого подхода приводит к задаче синтеза стохастически оптимального управления с интегральным критерием [5, 6]. Его иллюстрация дана автором в [7]. Наиболее полные результаты в этом направлении получены в [8]. Данная работа развивает этот подход далее.

**Постановка задачи в байесовском случае.** В ряде публикаций отмечается (см., например, [3]), что если непрерывная функция имеет единственный минимум, то ее значение вблизи него можно описать некоторой выпуклой квадратичной формой. Исходя из этого, будем для начала предполагать, что минимизируемая функция $x(u)$ имеет наименьшую неопределенность:

является выпуклой вниз квадратичной формой, известной с точностью до вектор-столбца сдвига $\theta$ ее минимума относительно начала координат в $m$-мерном евклидовом пространстве $R^m$: $x(u)=(u-\theta)^+A(u-\theta)$. Здесь вектор-столбец $(u-\theta)\in R^m$; $\theta$ – неизвестный случайный вектор-столбец с нормальным априорным распределением вероятностей с математическим ожиданием $\bar{\theta}$ и ковариационной матрицей $D$; «+» – знак транспонирования векторов и матриц; известная $m\times m$-матрица $A$ – симметричная положительно-определенная.

В качестве процедуры поиска минимума выберем процесс со структурой градиентного в конечных разностях поиска при непрерывном времени измерений $t$ на отрезке $[0,T]$. Вследствие этого поисковый вектор $u$ является кусочно-постоянной функцией на $[0,T]$. Для ее описания разобьем временной отрезок $[0,T]$ на $n$ полуотрезков $\tau_i$ отсчетами $t_0=0,\ t_1,\ t_2,\ldots,t_n=T$, так что $\tau_i=t_i-t_{i-1}$. Далее каждый отрезок $\tau_i$ разобьем на $m$ пар отрезков $\pi_{ij}$ и $\pi'_{ij}$ равной длины и таких, что $\tau_i=\bigcup\limits_{j=1}^{m}(\pi_{ij}\cup\pi'_{ij})$. В этих обозначениях $j$-я компонента $u(t)$ по определению будет иметь вид

$$u_i^j(t)=\begin{cases}u_{i-1}+\delta_{ij}I_j, & t\in\pi_{ij},\\ u_{i-1}-\delta_{ij}I_j, & t\in\pi'_{ij},\end{cases} \tag{1}$$

где $u_i$ – постоянный $m$-мерный вектор-столбец в течение $i$-го цикла длительности $\tau_i$; $I_1,I_2,\ldots,I_m$ – ортонормированный базис вектор-столбцов в $R^m$; $\delta_{ij}$ – неотрицательные числа, характеризующие покомпонентные амплитуды «возмущений» функции $x(u)$ при вычислении компонент ее градиента в точке $u_i$.

Значения минимизируемой функции $x(u)$ измеряются за $[0,T]$ непрерывно с аддитивными погрешностями:

$$y(t)=(u(t)-\theta)^+A(u(t)-\theta)+\xi(t), \tag{2}$$

где вектор $u(t)$ определен условиями (1), а $\xi(t)$ определим белым стационарным шумом с нулевым средним и спектральной плотностью $8p$.

После $i$-го эксперимента с планом (1) вычисляются оценки $m$ компонент вектор-столбца градиента $g_i(u_{i-1})$:

$$g_{ij}(u_{i-1})=\frac{1}{2\delta_{ij}\tau_{ij}}\left[\int\limits_{\pi_{ij}}y(t)dt-\int\limits_{\pi'_{ij}}y(t)dt\right],\quad j=1,\ldots,m. \tag{3}$$

Учитывая (1) и (2), из (3) получаем следующее выражение оценки вектор-столбца градиента в $i$-м цикле в точке $u_{i-1}$:

$$g_i(u_{i-1})=A(u_{i-1}-\theta)+\eta_i,\qquad i=1,\ldots,n, \tag{4}$$

где $\eta_i$ (как интеграл от белого шума $\xi(t)$ [9]) – последовательность статистически независимых реализаций гауссовских вектор-столбцов с нулевыми средними и ковариационными матрицами $M[\eta_i\eta_i^+] = pE_i^{-1}$; диагональная матрица

$$E_i = \mathrm{diag}(e_{i1}, e_{i2}, \ldots, e_{im}), \qquad e_{ij} = \delta_{ij}^2 \tau_{ij}. \tag{5}$$

Эту матрицу естественно назвать энергетической характеристикой $i$-го плана эксперимента (1) при оценке $i$-го градиента $g_i(u_{i-1})$, так как $\delta_{ij}^2$ и $e_{ij} = \delta_{ij}^2 \tau_{ij}$ есть соответственно мощность и энергия сигнала, используемого при оценке $j$-й компоненты градиента $g_i(u_{i-1})$.

Пусть $\varepsilon_n$ обозначает совокупность параметров всех $n$ планов экспериментов (1): $\varepsilon_n = (\tau_{ij}, \delta_{ij}, j = 1, \ldots, m, \ i = 1, \ldots, n)$.

Задача синтеза стохастически оптимальной процедуры поиска $\min x(u)$ формулируется следующим образом.

1. При фиксированном плане $\varepsilon_n$ поиска найти последовательность $u_0, u_1, \ldots, u_n \in R^m$, которая минимизирует интегральный риск за $[0, T]$:

$$r^*(\varepsilon_n) = \min M\left[\int_0^T (u(t) - \theta)^+ A(u(t) - \theta) dt\right],$$

где $M[\cdot]$ – знак математического ожидания. При решении этой задачи значения $u_i$ выбираются как функции предшествующих $u_k$ и оценок градиентов. Таким образом находится оптимальная энергопараметрическая процедура Кифера – Вольфовица.

2. Решением задачи минимизации $r^*(\varepsilon_n)$ по допустимым планам экспериментов $\varepsilon_n$ и числу циклов $n$ завершается исследование структуры, оптимальной в оговоренной байесовской постановке процедуры Кифера – Вольфовица за время $[0, T]$.

3. Оптимальная энергопараметрическая процедура Кифера – Вольфовица $u_i^*$ представляет собой последовательность байесовских апостериорных оценок вектора $\theta$. Известным образом отсюда получаем иную версию процедуры Кифера – Вольфовица как последовательность максимально правдоподобных оценок вектора $\theta$, эта версия не требует знания апостериорного распределения $\theta$.

**Оптимальная энергопараметрическая процедура Кифера – Вольфовица.**

**Утверждение 1.** Решением экстремальной задачи 1 является последовательность

$$u_0^* = \overline{\theta}, \qquad u_i^* = u_{i-1}^* - qD_i A E_i g_i(u_{i-1}^*), \tag{6}$$

где $pq = 1$; дисперсионные $m \times m$-матрицы

$$D_0 = D, \qquad D_i = \left(D^{-1} + q \sum_{k=1}^{i} A E_k A\right)^{-1}. \tag{7}$$

При этом оптимальный энергопараметрический риск

$$r^*(\varepsilon_n) = \sum_{i=1}^{n} \mathrm{tr} A E_i + \sum_{i=1}^{n} \mathrm{tr} A (E_i + \tau_i D_{i-1}), \tag{8}$$

где tr означает след матрицы.

Доказательство. Разобьем интегральный риск $r(\varepsilon_n)$ на сумму интегралов по участкам постоянства функции $u(t)$ согласно (1). Получим выражение интегрального риска

$$r(\varepsilon_n) = \sum_{i=1}^{n} \mathrm{tr} A E_i + M \sum_{i=1}^{n} \tau_i [(u_i - \theta)^* A (u_i - \theta)]. \tag{9}$$

Для минимизации $r(\varepsilon_n)$ по $\{u_i\}$ применим метод динамического программирования [5, 6] с учетом линейных уравнений измерений с аддитивными погрешностями (4). Легко показать, что оптимальная последовательность $u_i^*$ суть байесовские апостериорные средние:

$$u_i^* = M[\theta \mid u_0^*, u_1^*, \ldots, u_{i-1}^*, g_1, g_2, \ldots, g_i]$$

(подробно см. в [7]). Так как распределения вероятностей $\theta$ и $\eta_i$ нормальные, то формулы (6), (7) вычисляются по формуле Байеса, а далее – риск (8).

Попутно заметим, что оптимальные параметрические $u_i^*$ не зависят от длительности поиска $T$.

**Оптимальная структура процедуры Кифера – Вольфовица за время [0, *T*].** Перейдем теперь к решению задачи 2, т. е. к минимизации интегрального риска (8) по планам экспериментов $\varepsilon_n$.

Полученное выражение оптимального энергопараметрического риска (8) позволяет дать эквивалентное определение планов $n$ экспериментов по оценке градиентов в конечных разностях:

$$\varepsilon_n = (\tau_i, e_{ij}, \ i = 1, \ldots, n, \ j = 1, \ldots, m),$$

его допустимое множество

$$\gamma = \left\{ \frac{\tau_i, e_{ij}}{i = 1, \ldots, n; \ \ j = 1, \ldots, m} \middle| \frac{\tau_1 = \tau, \ \tau_s \geq 0, \sum \tau_s = T, \ e_{sj} \geq 0}{s = 1, \ldots, n; \ \ j = 1, \ldots, m} \right\}. \tag{10}$$

Далее потребуется следующая

**Лемма.** Пусть даны число $c > 0$; $m \times m$-матрицы $P$, $Q$ – симметричные положительно-определенные; вектор $z = (z_1, z_2, \ldots, z_m)$; диагональная матрица $Z = \mathrm{diag}(z)$. Тогда функция $\varphi(z) = \mathrm{tr} Q (P + cQZQ)^{-1}$ на положительном ортанте $z \in R_+^m$ является положительной, монотонно убывающей и строго выпуклой вниз.

Доказательство. Положительность следует из того, что обращаемая в $\varphi(z)$ матрица при $z \in R_+^m$ является симметричной положительноопределенной. Для установления убывания рассмотрим пару векторов $z > z'$, иначе говоря, $Z \succ Z'$ ($A \succ B$ означает, что $A - B$ есть положительно-определенная матрица). Имеем

$$P + cQZQ = P + cQZ'Q + cQ(Z - Z')Q \succ cQZ'Q.$$

Отсюда

$$Q(P + cQZQ)^{-1} \prec Q(P + cQZ'Q)^{-1}.$$

Вычисление следа этих матриц переводит матричное неравенство в арифметическое, что и требовалось. Для доказательства строгой выпуклости $\varphi(z)$ образуем выпуклую линейную комбинацию $z = \alpha z' + (1-\alpha)z''$, $0 < \alpha < 1$, которая порождает выпуклую комбинацию матриц: $Z = \alpha Z' + (1-\alpha)Z''$. Вследствие этого

$$Q(P + cQZQ)^{-1} = Q[\alpha(P + cQZ'Q) + (1-\alpha)(P + cQZ''Q)]^{-1} \prec$$

$$\prec \alpha Q(P + cQZ'Q)^{-1} + (1-\alpha)Q(P + cQZ''Q)^{-1}.$$

Вычисление следа матриц дает

$$\varphi(z) = \varphi[\alpha z' + (1-\alpha)z''] < \alpha\varphi(z') + (1-\alpha)\varphi(z''),$$

что означает строгую выпуклость функции $\varphi(z)$ на $R_+^m$.

**Утверждение 2.** Оптимальная структура процедуры (6), (7), минимизирующая параметрический риск (8) по планам $\varepsilon_n$ на временном отрезке $[0, T]$, содержит не более одного эксперимента по оценке градиента в конечных разностях: $n^* \in (1, 2)$, при этом минимальный риск

$$r^{**} = \min_{\varepsilon_1 \in \gamma} \operatorname{tr} A[E_1 + \tau_1 D + (T - \tau_1)(D^{-1} + qAE_1A)^{-1}], \qquad (11)$$

где $\varepsilon_1 = (\tau_1 = \tau,\ e_{1j},\ j = 1, \ldots, m)$.

Доказательство. Введем диагональные матрицы $Z_0 = 0$, $Z_k = \sum_{i=1}^{k} E_i$ и векторы $z_k$ из диагональных элементов $Z_k$, $k = 1, \ldots, n$. Тогда выражение оптимального параметрического риска (8) с учетом (7) запишется как

$$r^*(\varepsilon_n) = \operatorname{tr} AZ_n + \sum_{s=1}^{n} \tau_s \operatorname{tr} A(D^{-1} + qAZ_{s-1}A)^{-1} = \operatorname{tr} AZ_n + \sum_{s=1}^{n} \tau_s \psi(z_{s-1}),$$

где функции $\psi(z_{s-1}) = \operatorname{tr} A(D^{-1} + qAZ_{s-1}A)^{-1}$ того же типа, что $\varphi(z)$. Далее, векторы $z_s$ по определению образуют неубывающую последовательность векторов в $R_+^m$. Согласно леммы $\psi(0) \geq \psi(z_1) \geq \ldots \geq \psi(z_n)$. Учитывая это, получаем цепочку неравенств

$$r^*(\varepsilon_n) = \operatorname{tr} AZ_n + \sum_{s=1}^{n} \tau_s \psi(z_{s-1}) \geq \operatorname{tr} AZ_n + \tau_1\psi(0) + (T - \tau_1)\psi(z_n).$$

Отсюда

$$\min_{\varepsilon_n \in \gamma} r^*(\varepsilon_n) \geq \min_{\{z_s\}} [\mathrm{tr} A Z_n + \tau_1 \psi(0) + (T - \tau_1)\psi(z_n)]. \qquad (12)$$

В (12) равенство может иметь место только при $\psi(z_s) \equiv \mathrm{const}$, так как на $R_+^m$ $\psi(z) > 0$. С учетом монотонного убывания $\psi(z)$ это возможно только при $z_1 = z_2 = \ldots = z_n$, что эквивалентно $e_{2j} = e_{3j} = \ldots = e_{nj} = 0$ или $\delta_{2j} = \delta_{3j} = \ldots = \delta_{nj} = 0$ для всех $j = 1, \ldots, m$. Последнее означает, что оптимальное $n^* \in (1, 2)$.

**Утверждение 3.** Решение задачи (11) существует, единственно и ограничено.

**Доказательство.** Минимизируемая в (11) функция представима согласно (12) в виде

$$r^*(e_1) = \tau_1 \varphi(0) + \mathrm{tr} A E_1 + (T - \tau_1)\varphi(e_1),$$

т. е. суммой константы, линейной $\mathrm{tr} A E_1 \geq 0$ и строго выпуклой убывающей $\psi(e_1) > 0$ функции вектора $e_1 \in R_+^m$. При норме $\|e_1\| \to \infty$ имеем $r^*(e_1) \to \infty$, следовательно, точка минимума заведомо принадлежит некоторому ограниченному выпуклому замкнутому подмножеству $R_+^m$, на котором минимум единственный, а его координаты ограничены.

**Утверждение 4.** Существуют условия задачи, при которых оптимальная процедура Кифера – Вольфовица вырождается в априорную оценку экстремума $u^*(t) = \theta$ с интегральным риском $T \mathrm{tr} AD$.

**Доказательство.** Утверждение, очевидно, равнозначно тому, что в пределах условий задачи возможно достижение минимума $r^*(e_1)$ на $R_+^m$ в точке $e_1^* = 0$. Так как минимизируемая функция $r^*(e_1)$ непрерывна и строго выпукла на $R_+^m$, для существования ее минимума в точке $e_1 = 0$ необходимо и достаточно, чтобы градиент функции $r^*(e_1)$ в точке $e_1 = 0$ был неотрицательным. Имеем

$$\mathrm{grad} r^*(e_1) = \left\{ \mathrm{tr} I_k \left( A - \frac{T - \tau_1}{p} L \right),\ k = 1, \ldots, m \right\},$$

где $I_k$ – $m \times m$-матрица из нулей за исключением $k$-го элемента главной диагонали, равного 1; $L = (l_{ij}) = A D_1 A D_1 A$. Условие $\mathrm{grad} r^*(e_1) \geq 0$ порождает область $a_{kk} - p^{-1}(T - \tau_1) l_{kk} \geq 0$, $k = 1, \ldots, m$. Числа $a_{kk}$, $l_{kk}$ положительны и ограничены, как диагональные элементы положительно-определенных матриц, число $p > 0$ и ограничено. Следовательно, может существовать непустая область значений $A$, $p$, $T, \tau_1$, например, при $\min \frac{a_{kk}}{l_{kk}} \geq \frac{T - \tau_1}{p}$ условия неотрицательности компонент градиента совместимы.

Последнее утверждение означает, что при оценке эффективности процедуры Кифера – Вольфовица интегральным риском на $[0, T]$ ее применение не всегда целесообразно.

**Процедуры Кифера – Вольфовица как максимально правдоподобные оценки.** Перейдем к рассмотрению задачи 3.

Может случиться, что априорное распределение вектора $\theta$ неизвестно и процедуру поиска (6), (7) использовать нельзя. Однако в математической статистике известно (см., например, [5, 6]), что если формально обнулить в (7) матрицу $D^{-1}$, то получим последовательность максимально правдоподобных оценок $\theta$:

$$u_0^* = u, \quad u_i^* = u_{i-1}^* - qD_i AE_i g_i(u_{i-1}^*), \tag{13}$$

где $u$ – произвольный ограниченный вектор-столбец, а дисперсионные матрицы оценки $\theta$

$$D_i = p\left(\sum_{k=1}^{i} AE_k A\right)^{-1}, \quad i = 1, 2, \ldots. \tag{14}$$

Как байесовские оценки (6), (7), так и максимально правдоподобные оценки (13), (14) в среднеквадратическом и с вероятностью 1 асимптотически сходятся к истинному $\theta$ при числе циклов (шагов) $n \to \infty$. Иными словами, обе формы синтезированных процедур Кифера – Вольфовица асимптотически сходятся к $\min x(u)$ в среднеквадратическом и с вероятностью 1.

Отметим особенность оптимизации структуры интегрального риска при использовании процедуры максимально правдоподобных оценок (13), (14).

Байесовский параметрический риск (8) содержит два слагаемых: первая сумма характеризует потери на проведение экспериментов $\sum_{i=1}^{n} \mathrm{tr} AE_i$, а вторая – ожидаемые потери от самой процедуры (6) за время $[0, T]$. При переходе к процедуре максимально правдоподобных оценок (13), (14) слагаемое $\mathrm{tr}A(E_i + \tau_i D)$ оказывается неограниченным, так как $D^{-1} \to 0$ эквивалентно тому, что диагональные элементы $D$ стремятся к бесконечности. Однако разумно рассматривать первый эксперимент в течение $\tau_1$ как затравочный с вычислением $g_1(u_0)$ и $u_1^*$ по формуле максимально правдоподобной оценки. Последняя будет иметь нормальное распределение.

Далее, пренебрегая потерями в первом, затравочном цикле, естественно оптимизировать структуру поиска (13), (14) по величине конечной составляющей интегрального риска

$$\sum_{i=2}^{n} \mathrm{tr} AE_i + \sum_{i=2}^{n} \mathrm{tr} A(E_i + \tau_i D_{i-1}),$$

где $D_i$ вычисляется по формуле (14). Это выражение аналогично выражению байесовского энергопараметрического риска, начиная со второго цикла. Следовательно, все структурные выводы приложимы и в данном случае с той лишь разницей, что оптимальное число циклов $n \in (1, 2, 3)$ и первый цикл оценки градиента $g_1(u_0)$ имеет произвольно выбранный план $(u_0, \tau_1, \delta_{1j}$, $j = 1, \ldots, m)$, ограниченный условиями: $u_0$ – произвольный ограниченный $m$-мерный вектор-столбец, а также $0 < \tau \le \tau_1 < T$, $\delta_{1j} > 0, j = 1, \ldots, m$.

**Заключение.** В данной работе введены непрерывное время измерений с погрешностями и интегральный риск как характеристика эффективности градиентного в конечных разностях поиска экстремума непрерывной функции. Это позволяет сформулировать задачу оптимизации процедуры поиска в терминах стохастической задачи оптимального управления. Выбор простейшей минимизируемой функции – выпуклой вниз квадратичной формы, известной с точностью до сдвига – позволил в байесовском случае эту задачу решить в явном виде. В синтезируемой таким образом оптимальной процедуре Кифера – Вольфовица ее параметры однозначно определяются планами экспериментов по оценке градиентов в конечных разностях. И далее оптимизировать ее по планам экспериментов на конечном отрезке времени поиска.

Полезна аналогия с теорией статистических оценок непрерывных функций. Наибольшие результаты получены для простейших линейно-параметризованных функций с фиксированными планами измерений и последующей оптимизацией этих планов [10].

## СПИСОК ЛИТЕРАТУРЫ


1. **Вазан М.** Стохастическая аппроксимация. М.: Мир, 1972.
2. **Невельсон М. Б., Хасьминский Р. З.** Стохастическая аппроксимация и рекуррентные оценивания. М.: Наука, 1972.
3. **Ермольев Ю. И.** Методы стохастического программирования. М.: Наука, 1976.
4. **Товстуха Т. И., Томингас К. В.** Оценка сходимости процедуры Кифера – Вольфовица на конечном интервале // Адаптивные системы. Большие системы. М.: Наука, 1971.
5. **Брайсон А., Хо Ю-Ши.** Прикладная теория оптимального управления. М.: Мир, 1972.
6. **Беседин Б. А.** Теория распределенных информационно-измерительных систем. Новосибирск: Изд-во СО РАН, 1999.
7. **Беседин Б. А.** Субоптимальное управление в самонастраивающихся системах с фиксированными экспериментами // Адаптивные системы. Большие системы. М.: Наука, 1971.
8. **Беседин Б. А.** Оптимальная структура адаптивного процесса для квадратичной функции // Стохастические модели сложных систем. Новосибирск: ВЦ СО РАН СССР, 1990.
9. **Лебедев В. А.** Случайные процессы в электрических и механических системах. М.: Физматгиз, 1958.
10. **Математическая** теория планирования экспериментов /Под общ. ред. С. М. Ермакова. М.: Наука, 1983.



*Омский филиал института*
*математики им. С. Л. Соболева СО РАН,*
*E-mail: besedin@iitam.omsk.net.ru*



*Поступила в редакцию*
*27 ноября 2002 г.*