

Математическое моделирование
Mathematical modeling

УДК 004.023, 519.677
<https://doi.org/10.32362/2500-316X-2023-11-6-57-67>



НАУЧНАЯ СТАТЬЯ

Исследование влияния степени овражности целевой функции на погрешность определения координат ее минимума

А.В. Смирнов [®]

МИРЭА – Российский технологический университет, Москва, 119454 Россия

[®] Автор для переписки, e-mail: av_smirnov@mirea.ru

Резюме

Цели. Целью работы было исследование зависимостей, связывающих характеристики оврагов, т.е. участков рельефа минимизируемой функции, на которых ее изменение по одному из направлений значительно медленнее, чем по другим направлениям, с погрешностью определения координат ее минимума.

Методы. В экспериментах использовалась специально разработанная тестовая функция с изменяемыми в широких пределах параметрами овражности. В сериях опытов случайно задавались положение и параметры оврага и координаты стартовой точки поиска. Размерность и степень овражности оценивались по собственным числам аппроксимированного гессиана функции в точке окончания поиска минимума. Погрешность определялась как евклидово расстояние между заданным положением минимума функции и конечной точкой поиска. Для статистической обработки результатов применены линейный регрессионный анализ и аппроксимация с помощью модели искусственной нейронной сети (ИНС).

Результаты. Установлено наличие линейной зависимости между логарифмами степени овражности и погрешности определения координат минимума функции. Коэффициент детерминации $R^2 \sim 0.88$. Дополнительный учет евклидовой нормы градиента функции в точке окончания поиска позволил повысить коэффициент детерминации до $R^2 \sim 0.95$, а при использовании модели ИНС – до $R^2 \sim 0.97$.

Выводы. Найденные зависимости можно использовать для оценки ожидаемой погрешности определения координат экстремумов оптимизируемых функций. В дальнейшем необходимо расширить методику на функции с размерностью оврагов более единицы и на другие типы сложных для алгоритмов оптимизации участков рельефа.

Ключевые слова: рельеф целевой функции, овражность рельефа, степень овражности, размерность овражности, собственные значения гессиана, линейная регрессия, аппроксимация, искусственная нейронная сеть

• Поступила: 10.04.2023 • Доработана: 19.05.2023 • Принята к опубликованию: 07.09.2023

Для цитирования: Смирнов А.В. Исследование влияния степени овражности целевой функции на погрешность определения координат ее минимума. *Russ. Technol. J.* 2023;11(6):57–67. <https://doi.org/10.32362/2500-316X-2023-11-6-57-67>

Прозрачность финансовой деятельности: Автор не имеет финансовой заинтересованности в представленных материалах или методах.

Автор заявляет об отсутствии конфликта интересов.

RESEARCH ARTICLE

Investigation of influence of objective function valley ratio on the determination error of its minimum coordinates

Alexander V. Smirnov [®]

MIREA – Russian Technological University, Moscow, 119454 Russia

[®] Corresponding author, e-mail: av_smirnov@mirea.ru

Abstract

Objectives. A valley is a region of an objective function landscape in which the function varies along one direction more slowly than along other directions. In order to determine the error of the objective function minimum location in such regions, it is necessary to analyze relations of valley parameters.

Methods. A special test function was used in numerical experiments to model valleys with variables across wide ranges of parameters. The position and other valley parameters were defined randomly. Valley dimensionality and ratio were estimated from eigenvalues of the approximated Hessian of objective function in the termination point of minimum search. The error was defined as the Euclidian distance between the known minimum position and the minimum search termination point. Linear regression analysis and approximation with an artificial neural network model were used for statistical processing of experimental data.

Results. A linear relation of logarithm of valley ratio to logarithm of minimum position error was obtained. Here, the determination coefficient R^2 was ~ 0.88 . By additionally taking into account the Euclidian norm of the objective function gradient in the termination point, R^2 can be augmented to ~ 0.95 . However, by using the artificial neural network model, an approximation $R^2 \sim 0.97$ was achieved.

Conclusions. The obtained relations may be used for estimating the expected error of extremum coordinates in optimization problems. The described method can be extended to functions having a valley dimensionality of more than one and to other types of hard-to-optimize algorithms regions of objective function landscapes.

Keywords: objective function landscape, valley landscape, valley ratio, valley dimensionality, Hessian eigenvalues, linear regression, approximation, artificial neural network

• Submitted: 10.04.2023 • Revised: 19.05.2023 • Accepted: 07.09.2023

For citation: Smirnov A.V. Investigation of influence of objective function valley ratio on the determination error of its minimum coordinates. *Russ. Technol. J.* 2023;11(6):57–67. <https://doi.org/10.32362/2500-316X-2023-11-6-57-67>

Financial disclosure: The author has no a financial or property interest in any material or method mentioned.

The author declares no conflicts of interest.

ВВЕДЕНИЕ

Задача поиска оптимального решения \mathbf{x}_{opt} формулируется как

$$\mathbf{x}_{\text{opt}} = \arg \min_{\mathbf{x} \in X} f(\mathbf{x}), \quad (1)$$

где X – область поиска, $f(\mathbf{x})$ – целевая функция (ЦФ). Известны многочисленные методы решения задачи (1), как имеющие достаточно строгое математическое обоснование и применимые в случаях, когда ЦФ удовлетворяет определенным условиям (выпуклость, гладкость и т.д.) [1, 2], так и эвристические,

не накладывающие жестких требований на свойства ЦФ, но и не гарантирующие нахождение оптимального решения [3, 4].

Возможность и точность решения задачи (1) определяются свойствами и ЦФ, и алгоритма поиска. В связи с этим значительное внимание исследователей привлекает анализ рельефа ЦФ, где под рельефом понимается множество пар $\{\mathbf{x} \in X, f(\mathbf{x})\}$. При этом аналитическое выражение функции $f(\mathbf{x})$ отсутствует, и ее значения приходится находить путем моделирования оптимизируемой системы (задачи black box optimization). Данное направление получило название exploratory landscape analysis (ELA).

В [5] и других работах той же группы авторов дана классификация свойств рельефов ЦФ высокого уровня, определяемых качественно методом экспертной оценки, и свойств низкого уровня, оцениваемых количественно путем обработки результатов расчетов ЦФ в пробных точках и результатов поиска экстремумов ЦФ из стартовых точек. К свойствам высокого уровня относятся мульти-модальность, т.е. наличие многих локальных экстремумов, регулярность и равномерность свойств рельефа ЦФ в области поиска, наличие плато и др. Среди свойств низкого уровня – статистика значений ЦФ, показатели кривизны и выпуклости, показатели корреляции различий значений ЦФ и расстояний между пробными точками и многие другие. В [6] рассматриваются более 300 свойств низкого уровня и приведен список публикаций по данной тематике. Для поиска статистических зависимостей между свойствами низкого и высокого уровней, а также между свойствами рельефа и характеристиками работы различных алгоритмов оптимизации на этом рельефе применяются технологии машинного обучения [5, 7].

Однако в указанных выше и других известных нам работах практически не рассматривались такие объекты рельефа ЦФ, как овраги (valley), т.е. области, в которых ЦФ по одному или нескольким направлениям изменяется существенно медленнее, чем по другим направлениям [8], и стены, представляющие собой резкие изменения ЦФ по какому-либо направлению [9]. Между тем, при наличии таких объектов поиск может заканчиваться не в экстремуме, а в какой-либо другой точке на дне оврага или у подножья стены. В таком случае алгоритмы оптимизации будут находить неправильные решения. Проблемы, связанные с обнаружением указанных объектов в рельефе ЦФ и оценкой их количественных характеристик, исследованы мало.

Теоретические аспекты возникновения оврагов в рельефе и методов поиска решений при их наличии рассмотрены в [8]. Из нескольких определений овражности, предложенных в этой монографии, далее будем использовать наиболее удобное в приложениях, которое приведем с некоторым упрощением, опустив дополнительные условия.

Пусть D – некоторая область n -мерного пространства R^n ; $J(x) \in C^2(D)$ – функционал, имеющий в D непрерывные вторые производные; $H(x)$ – матрица вторых производных (гессиан) функционала $J(x)$ в точке x ; $\lambda_i[H(x)]$, $i = 1, n$ – собственные числа гессиана $H(x)$ в точке x , упорядоченные по убыванию.

Функционал называется овражным, т.е. содержащим овраг, если найдутся такое число $\sigma \gg 1$ и множество $Q \subset D$, что

$$\forall x \in Q \quad \lambda_1[H(x)] \geq \dots \geq \lambda_{n-r}[H(x)] \geq \sigma \lambda_{n-r+1}[H(x)] \geq \dots \geq \sigma \lambda_n[H(x)]. \quad (2)$$

Это означает, что наибольшие $(n - r)$ собственных чисел гессиана существенно больше, чем остальные r собственных чисел во всех точках x , принадлежащих области оврага Q . При этом число r называется размерностью оврага, а число σ – степенью овражности. Степень овражности показывает, насколько в данном овраге скорость изменения ЦФ вдоль его дна меньше, чем по направлениям, ортогональным дну. Эти показатели обобщенно можно назвать характеристиками или показателями овражности рельефа.

Результатом наличия оврагов в рельефе ЦФ, как уже отмечено выше, является погрешность определения координат экстремумов ЦФ. Именно эта погрешность может служить объективной характеристикой овражности. Но непосредственно оценивать погрешность при поиске минимума ЦФ на практике невозможно, т.к. истинное положение минимума неизвестно. В то же время степень овражности может быть оценена на основе определения (2). В связи с этим представляет интерес задача исследования зависимости, связывающей указанную погрешность со степенью овражности. Такая задача в [8] и других известных нам работах не рассматривалась, как и практическая реализация оценки степени овражности при выполнении поиска минимума ЦФ.

Отметим также работу [10], в которой вводится определение оврага как одномерного множества с использованием понятия топологического гооморфизма. Там же представлен метод определения положения и направления оврага, основанный на выделении из множества пробных точек подмножества точек с наименьшими значениями ЦФ и применении к этому подмножеству метода анализа главных компонент. Количественные характеристики овражности в данной работе не рассматривались. В [11] исследуются свойства рельефа ЦФ известной комбинаторной задачи «странствующего коммивояжера» и показано, что этот рельеф содержит группы близко расположенных впадин, также названные valley, но эти результаты неприменимы к задачам оптимизации функций непрерывных переменных.

Цель данной работы – исследовать зависимость погрешности определения координат искомого минимума ЦФ от характеристик овражности в окрестностях точек окончания поиска.

Для достижения поставленной цели требуется выполнить цикл экспериментов по поиску из различных стартовых точек минимума ЦФ с варьируемыми параметрами оврага, включая его положение,

ориентацию относительно осей координат, крутизну и кривизну склонов и т.д. При этом в конце каждого поиска необходимо определять собственные числа гессиана ЦФ, а по ним размерность оврага и степень овражности в соответствии с определением (2), а также вычислять погрешность, равную расстоянию конечной точки поиска до истинного положения минимума ЦФ. Затем следует провести анализ статистических зависимостей, связывающих характеристики овражности и погрешность определения координат минимума ЦФ.

МАТЕРИАЛЫ И МЕТОДЫ

В первую очередь надо выбрать метод поиска минимума ЦФ. Погрешность координат минимума при наличии оврагов может существенно различаться для разных алгоритмов оптимизации. Выбрав алгоритм и выполнив экспериментальные исследования, можно будет построить шкалу оценки овражности, с которой затем будут сравниваться результаты других алгоритмов.

В данной работе использовался квазиньютоновский алгоритм локального поиска, реализуемый в программной среде *MATLAB* функцией *fmincon*(..). Алгоритм такого типа рекомендован в [6] как образцовый для поиска локальных экстремумов тестовых функций при оценке свойств рельефа ЦФ. Кроме того, в квазиньютоновских методах аппроксимация гессиана является неотъемлемой частью алгоритма на каждой итерации и, следовательно, автоматически получается в конечной точке поиска [1].

В список выходных переменных функции *fmincon*(..) входят вектор координат точки окончания поиска \mathbf{x}_{fin} , значение ЦФ f_{fin} в точке \mathbf{x}_{fin} , индикатор причины окончания поиска *ExitFlag*, вектор градиента ЦФ и аппроксимация гессиана в точке \mathbf{x}_{fin} в виде матрицы действительных чисел. Во входных переменных задавались границы области поиска

$$-5 \leq x_i \leq 5, \quad i = \overline{1, ND}, \quad (3)$$

вариант алгоритма поиска SQP (sequential quadratic programming), максимальное число итераций в каждом поиске равно 1000, остальные настройки – по умолчанию.

Далее рассмотрим использованную в экспериментах ЦФ. Для испытаний и сравнения алгоритмов поиска экстремумов применяются наборы тестовых функций [12, 13]. Среди них есть обладающие свойством овражности, но нет функции, в которой параметры оврага можно было бы изменять в широких пределах. В связи с этим была разработана функция *TestValley*(..), текст которой на языке *MATLAB* приведен на рис. 1.

Здесь \mathbf{x} – координаты точки, в которой рассчитывается значение функции, \mathbf{x}_{opt} и f_{opt} – задаваемые координаты точки минимума и значение функции в ней, \mathbf{R} – ортонормированная матрица, задающая поворот осей координат. Эти параметры позволяют получать различные положения и ориентации оврага в пространстве поиска. Параметр N задает размерность оврага. Параметр W определяет кривизну склонов оврага. При $W = 1$ ЦФ – квадратичная, т.е.

```

%%TestValley - моделирование овражности с изменяемыми параметрами
function f = TestValley(x, fopt, xopt, R, P, W, K, N)
%%P - тип масштабирования по разным координатам.
%%P=0 одинаковая скорость нарастания по всем x(i), i>N.
%%P>0 - при j>i нарастание по x(j) быстрее, чем по x(i).
%%W - тип зависимости от расстояния от дна оврага.
%%W=0.5 - линейная функция,
%%W<0.5 - вогнутая функция, W>0.5 - выпуклая функция.
%%K - общий коэффициент масштаба по направлениям от оси.
%%N - размерность оврага.
n=length(x);
z0=(x-xopt)*R;
L=eye(n);
for n1=1:n
    L(n1,n1)=10^(P*(n1-1)/2/(n-1));
end
z1=z0*L;
z3=z1.^2;
f=sum(z3(1:N),2)+K*(sum(z3(N+1:n),2))^W+fopt;
end

```

Рис. 1. Текст программы, реализующей функцию *TestValley*(..)

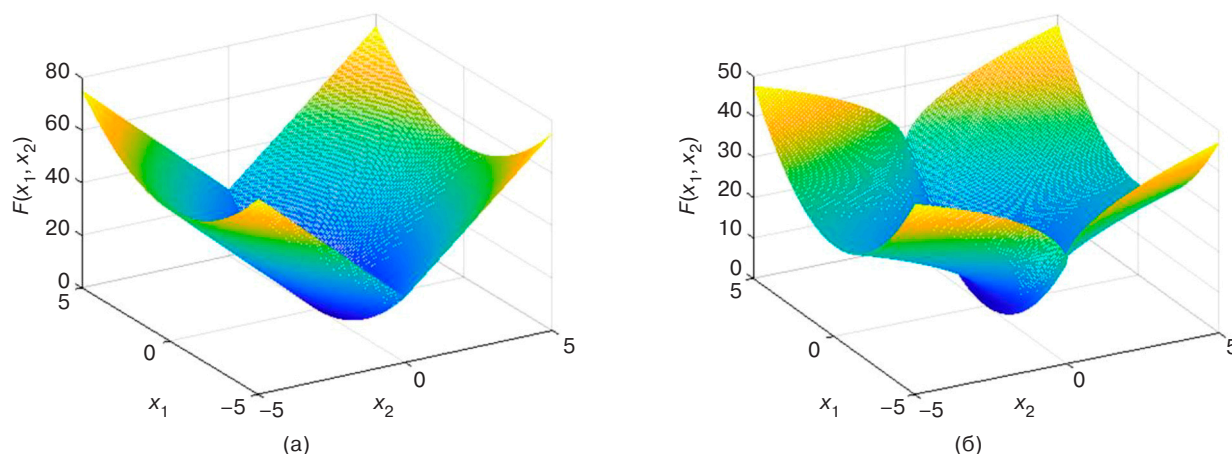


Рис. 2. Графики функции *TestValley*(..): (а) $W = 0.5$; (б) $W = 0.25$

выпуклая и гладкая. При $1 < W < 0.5$ ЦФ выпуклая, но не гладкая. При $W = 0.5$ ЦФ нарастает линейно. Наконец, при $W < 0.5$ ЦФ вогнутая. Параметр K задает единый масштаб скорости нарастания ЦФ по всем направлениям. Параметр P влияет на анизотропию ЦФ. При $P = 0$ ЦФ нарастает с одинаковой скоростью по всем направлениям, а при $P > 0$ скорость нарастания по разным направлениям разная, и эти различия тем больше, чем больше P .

Примеры графиков функции при размерности пространства поиска $ND = 2$, значениях параметров $P = 0$, $K = 10$, $N = 1$ и разных значениях параметра W приведены на рис. 2.

Координаты стартовых точек в количестве N_{Pnt} в пределах границ области поиска (3) задавались с применением алгоритма формирования латинского гиперквадрата, реализуемого в *MATLAB* функцией *lhsdesign*(..). Значение функции f_{opt} в минимуме устанавливалось равным 0. В данной работе исследовались только одномерные овраги при $N = 1$. Значения остальных параметров задавались случайными числами с равномерными распределениями в следующих диапазонах

$$\begin{aligned} -3 \leq x_{opti} \leq 3, \quad i = \overline{1, ND}; \\ 0 \leq P \leq 1; \quad 0.25 \leq W \leq 1.25; \\ 0 \leq \lg K \leq 4. \end{aligned} \quad (4)$$

Матрица поворота \mathbf{R} формировалась в виде квадратной матрицы $ND \times ND$ равномерно распределенных в интервале $(0; 1)$ случайных чисел

```
[X1,GF1,ExitFlag,~,~,grad,hess]=fmincon(@FEval,Xin,[],[],[],[],[],Lb,Ub,[],MIOptions);
Ehess=eig(hess);
HessEV=sort(abs(Ehess));
S0hess=HessEV(2:end)./HessEV(1:end-1);
[SValley,NValley]=max(S0hess);    %степень и размерность овражности
```

Рис. 3. Вызов функции поиска и оценка параметров овражности в точке его окончания

с последующей ортогонализацией с применением функции *MATLAB orth*(..). Задание перечисленных параметров возможно как отдельно для каждого старта, так и один раз для всей серии из N_{Pnt} стартов.

Вызов функции поиска минимума и обработка возвращаемых ей результатов поясняются фрагментом программы на рис. 3. В переменных \mathbf{X}_1 и GF_1 возвращаются, соответственно, координаты точки окончания поиска \mathbf{x}_{fin} и значение ЦФ f_{fin} в ней. Массивы **grad** и **hess** содержат, соответственно, вектор градиента и матрицу аппроксимированного гессиана. Указатель **@FEval** содержит ссылку для вызова функции *TestValley*(..), при выполнении которого задаются ее входные параметры, как это было описано выше.

Программа находит вектор собственных значений гессиана **Ehess** и упорядочивает их по возрастанию абсолютного значения в массиве **HessEV**. Затем рассчитываются отношения смежных значений, сохраняемые в массиве **S0hess**. Наконец определяется оценка степени овражности *SValley* как максимальное из значений этих отношений и размерность оврага *NValley* как номер максимального значения в массиве. Такое определение степени и размерности овражности соответствует приведенному выше определению (2) с тем принципиальным отличием, что упорядочивание собственных значений гессиана выполняется не по убыванию, а по возрастанию. Погрешность определения координат минимума $D\mathbf{X}_{opt}$ рассчитывается как евклидово расстояние между точками \mathbf{x}_{opt} и \mathbf{x}_{fin} .

Анализ зависимостей между переменными осуществлялся двумя методами. Первый из них – линейный регрессионный анализ [14]. Для его реализации применялась функция *MATLAB fitlm(...)*, аппроксимирующая по исходным данным линейную модель. Второй метод – обучение модели искусственной нейронной сети (ИНС), аппроксимирующей искомую зависимость [15]. Для этого использовались функция *fitnet(...)*, создающая модель ИНС с заданной структурой, и функция *train(...)*, выполняющая обучение и тестирование модели.

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЙ

Сначала рассмотрим представленные на рис. 4 результаты предварительных экспериментов, дающие представление о влиянии степени овражности на погрешность нахождения точки минимума функции *TestValley(...)*. Размерность задачи $ND = 4$, число стартов $NPnt = 40$, положение минимума и поворот оврага заданы один раз для всей серии. Входные параметры функции показаны над диаграммами, отображающими перемещение от стартовой точки (красные маркеры) до конечной точки x_{fin} (синие маркеры). На одной диаграмме показаны изменения всех четырех координат, при этом кружками помечены координаты x_1 и x_2 , а треугольниками – x_3 и x_4 .

В случае слабой овражности (рис. 4а) поиск из всех стартовых точек приходит в окрестность точки минимума тестовой функции. Среднее значение погрешности DX_{opt} при этом равно 0.019. В случае же сильной овражности (рис. 4б) поиски, стартовавшие из разных точек, заканчиваются в разных точках, рассредоточенных вдоль дна оврага. При этом среднее значение DX_{opt} достигает 9.2.

Далее была выполнена серия экспериментов, имевших целью выявление зависимости между оценкой степени овражности S_{Valley} и погрешностью нахождения точки минимума тестовой функции DX_{opt} . Размерность пространства ND менялась в диапазоне от 2 до 12. Каждый эксперимент включал $12 \cdot 10^3$ стартов, в каждом из которых задавались случайные положение стартовой точки поиска, положение минимума x_{opt} и поворот R функции *TestValley(...)*, а также параметры овражности P, W, K в диапазонах, определенных неравенствами (4).

Результаты экспериментов для всех значений ND оказались сходными. В качестве примера на рис. 5 приведены гистограммы значений результатов эксперимента для $ND = 4$. В связи с широкими диапазонами значений S_{Valley} и DX_{opt} анализируются и отображаются на графиках их логарифмы. Из диаграммы рассеяния этих переменных, показанной на рис. 6, можно сделать вывод о наличии стохастической зависимости между ними.

Отметим, что значение $ExitFlag = 1$ соответствует окончанию поиска, когда модуль градиента в достигнутой точке не превышает заданного значения *OptimalityTolerance* (по умолчанию 10^{-6}), а значение $ExitFlag = 2$ – когда последнее перемещение при поиске не превышает заданного значения *StepTolerance* (по умолчанию – также 10^{-6}). Во втором случае модуль градиента ЦФ в точке окончания поиска может оказаться значительно больше *OptimalityTolerance*, т.к. на оси оврага нарушаются условия гладкости ЦФ, что было проиллюстрировано рис. 2.

При этом не оказалось ни одного случая, когда поиск закончился из-за превышения заданного числа итераций ($MaxIteration = 1000$) или из-за того,

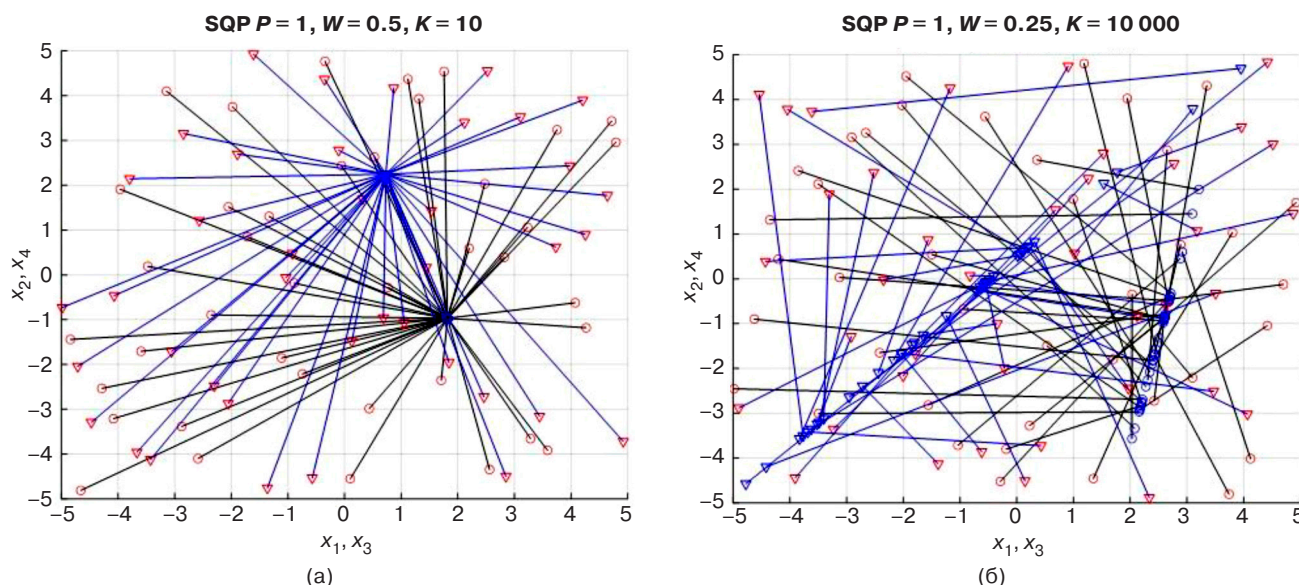


Рис. 4. Результаты поиска минимума *TestValley(...)* при слабой (а) и сильной (б) овражности

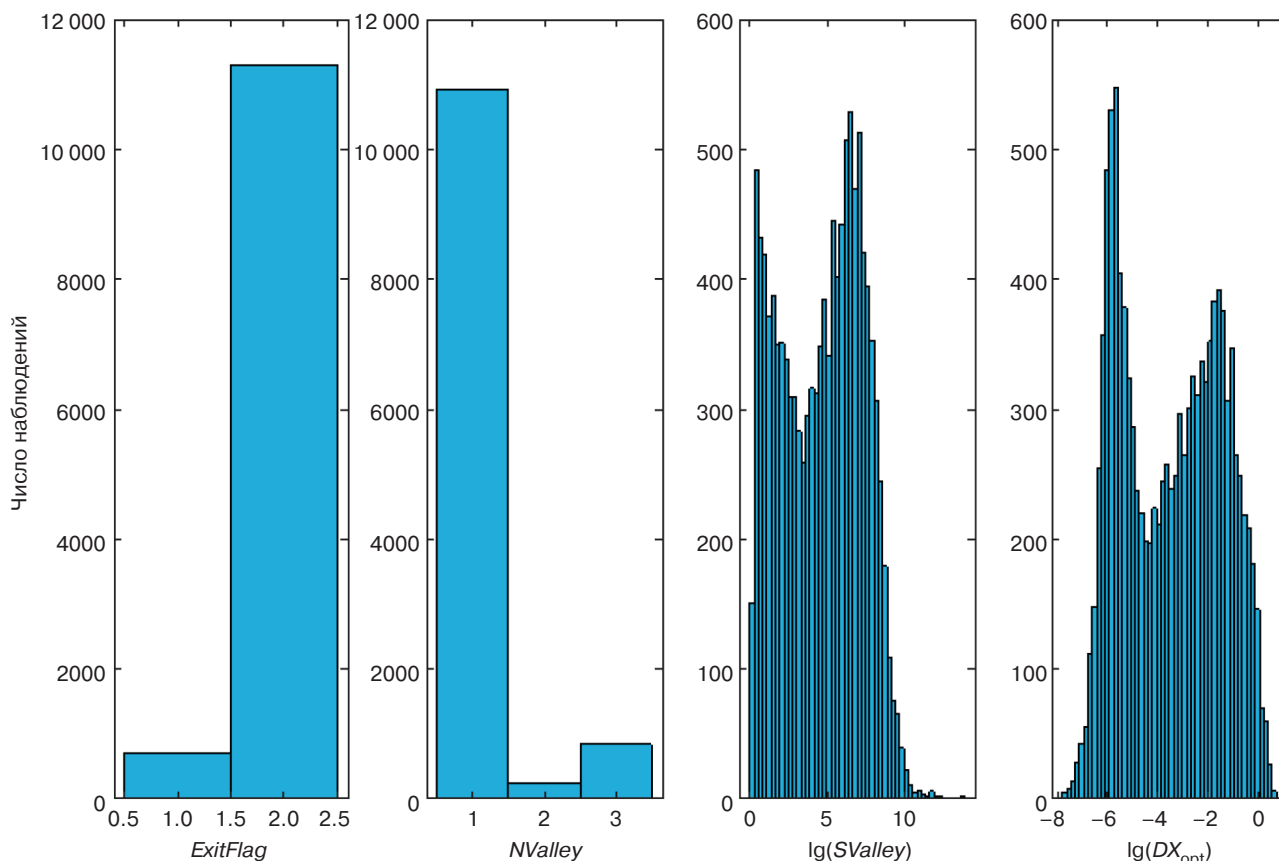


Рис. 5. Гистограммы результатов основного эксперимента. *ExitFlag* – индикатор причины окончания поиска; *NValley* – оценка размерности оврага; *SValley* – оценка степени овражности; *DX_{opt}* – оценка погрешности определения положения минимума ЦФ

что алгоритм не смог найти допустимую точку для дальнейшего движения. Таким образом, все старты поиска закончились в точках, которые алгоритм определил как локальный минимум. Аналогичные результаты зафиксированы для всех значений размерности пространства *ND*.

Далее, в большинстве стартов правильно определена размерность оврага в конечной точке поиска *NValley* = 1. При *ND* = 4 в 1092 точках размерность оврага была определена ошибочно. Эти точки на рис. 6 отмечены черным цветом. Все они находятся в области, где погрешность поиска *DX_{opt}* пренебрежимо мала, а степень овражности *SValley* < 1000. При *ND* > 4 встречались оценки *NValley* > 3, но также только в области *lg(DX_{opt})* < -4. При *ND* = 2 значения *NValley* > 1, очевидно, невозможны.

Перейдем к статистической обработке собранных в экспериментах данных. Предварительно был выполнен анализ влияния параметров *P*, *W*, *K* функции *TestValley(..)* на степень овражности *SValley*. Исследовалась линейная модель

$$\lg(SValley) = k_1P + k_2W + k_3\lg(K) + b. \quad (5)$$

Регрессионный анализ этой модели показал, что на степень овражности сильнее всего влияет

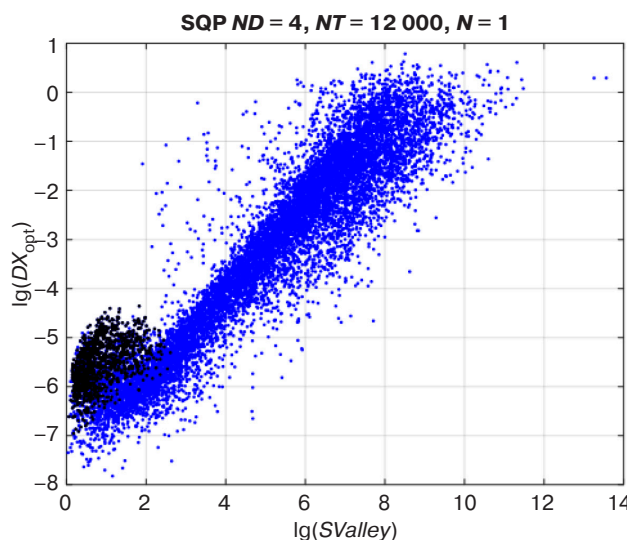


Рис. 6. Диаграмма рассеяния значений логарифмов степени овражности *SValley* и погрешности определения координат минимума *DX_{opt}*

параметр *W*, определяющий кривизну и выпуклость или вогнутость склонов оврага. Следующим по вкладу в результат идет параметр *K*, а влияние параметра *P* наименее существенно, но все же пренебречь им нельзя. Значения коэффициента детерминации *R*²,

определяющего адекватность модели [14], для разных ND находились в диапазоне 0.88–0.90.

Далее была рассмотрена линейная модель, связывающая погрешность нахождения точки минимума ЦФ с оценкой степени овражности

$$\lg(DX_{\text{opt}}) = k \lg(S\text{Valley}) + b. \quad (6)$$

Результаты регрессионного анализа приведены в табл. 1, где первая группа столбцов соответствует учету всех точек, а вторая группа – с исключением точек с оценкой $N\text{Valley} > 1$. Этот отбор данных приводит к некоторому улучшению точности модели, выражающемуся в увеличении коэффициента

детерминации R^2 и уменьшению среднеквадратической ошибки (остатка) регрессии $StdErr$.

Далее была исследована возможность повышения точности модели путем учета евклидовой нормы (длины) вектора градиента ЦФ $\|\mathbf{grad}\|$ в точке окончания поиска. Рассматривалась следующая линейная модель

$$\lg(DX_{\text{opt}}) = k_1 \lg(S\text{Valley}) + k_2 \lg(\|\mathbf{grad}\|) + b. \quad (7)$$

Результаты регрессионного анализа этой модели даны в табл. 2. Как и для предыдущей модели, выделены случаи с учетом всех точек и с исключением точек с $N\text{Valley} > 1$. Заметим, что коэффициент

Таблица 1. Результаты регрессионного анализа модели (6)

ND	Все точки				Исключены точки с $N\text{Valley} > 1$			
	b	k	R^2	$StdErr$	b	k	R^2	$StdErr$
2	−6.97	0.564	0.846	0.825	–	–	–	–
3	−6.81	0.675	0.855	0.755	−6.98	0.702	0.856	0.747
4	−6.80	0.714	0.881	0.686	−7.08	0.758	0.885	0.666
5	−6.82	0.744	0.882	0.685	−7.15	0.797	0.884	0.667
6	−6.79	0.756	0.884	0.674	−7.09	0.806	0.885	0.660
7	−6.72	0.757	0.883	0.673	−7.03	0.808	0.886	0.656
8	−6.69	0.766	0.876	0.700	−7.00	0.817	0.878	0.687
9	−6.63	0.764	0.874	0.700	−6.94	0.816	0.877	0.684
10	−6.60	0.772	0.868	0.712	−6.93	0.828	0.869	0.699
11	−6.59	0.777	0.865	0.716	−6.90	0.830	0.868	0.702
12	−6.54	0.779	0.857	0.740	−6.86	0.834	0.859	0.729

Таблица 2. Результаты регрессионного анализа модели (7)

ND	Все точки					Исключены точки с $N\text{Valley} > 1$				
	b	k_1	k_2	R^2	$StdErr$	b	k_1	k_2	R^2	$StdErr$
2	−3.53	0.041	0.526	0.936	0.534	–	–	–	–	–
3	−3.84	0.125	0.464	0.929	0.529	−3.97	0.145	0.469	0.937	0.495
4	−3.91	0.163	0.453	0.935	0.507	−4.08	0.184	0.475	0.950	0.439
5	−4.01	0.196	0.438	0.936	0.502	−4.15	0.209	0.479	0.956	0.411
6	−3.98	0.198	0.442	0.937	0.495	−4.07	0.201	0.490	0.956	0.407
7	−3.95	0.199	0.444	0.938	0.490	−4.06	0.204	0.491	0.959	0.395
8	−3.86	0.184	0.460	0.939	0.492	−3.99	0.196	0.501	0.959	0.398
9	−3.89	0.195	0.451	0.937	0.493	−4.02	0.208	0.492	0.959	0.394
10	−3.84	0.190	0.456	0.937	0.492	−4.02	0.211	0.493	0.958	0.397
11	−3.84	0.194	0.458	0.936	0.494	−4.00	0.212	0.496	0.958	0.397
12	−3.84	0.195	0.454	0.934	0.501	−4.03	0.218	0.490	0.957	0.401

Таблица 3. Результаты аппроксимации с применением моделей ИНС

ND	Аппроксимация по $\lg(SValley)$				По $\lg(SValley)$ и $\lg(\mathbf{grad})$			
	Все точки		Без $NValley > 1$		Все точки		Без $NValley > 1$	
	R^2	$StdErr$	R^2	$StdErr$	R^2	$StdErr$	R^2	$StdErr$
2	0.863	0.777	–	–	0.954	0.451	–	–
3	0.881	0.685	0.877	0.691	0.952	0.437	0.953	0.425
4	0.906	0.609	0.904	0.608	0.959	0.404	0.963	0.378
5	0.904	0.617	0.900	0.618	0.960	0.399	0.966	0.360
6	0.907	0.602	0.904	0.603	0.959	0.402	0.965	0.364
7	0.908	0.598	0.906	0.595	0.961	0.388	0.969	0.341
8	0.903	0.621	0.900	0.622	0.961	0.391	0.970	0.340
9	0.903	0.613	0.901	0.613	0.964	0.376	0.972	0.329
10	0.895	0.634	0.891	0.640	0.960	0.391	0.969	0.339
11	0.891	0.644	0.888	0.648	0.961	0.384	0.970	0.336
12	0.883	0.668	0.879	0.676	0.959	0.398	0.970	0.337

корреляции значений $\lg(||\mathbf{grad}||)$ и $\lg(SValley)$ при разных размерностях пространства ND меняется в пределах от 0.91 до 0.95, т.е. корреляция существенная. Тем не менее, учет нормы градиента дает дополнительную информацию для оценки погрешности DX_{opt} . По сравнению с моделью (6) коэффициент детерминации R^2 стал ближе к единице, а среднеквадратическая погрешность $StdErr$ уменьшилась. При этом исключение точек с ошибочно определенной размерностью оврага, как и в предыдущем случае, улучшает характеристики модели.

Другой подход к аппроксимации зависимостей между данными, собранными в экспериментах, основан на обучении моделей ИНС. Известно, что с помощью ИНС со скрытыми слоями и достаточным количеством нейронов в них можно аппроксимировать любую непрерывную функцию нескольких переменных [15]. Использовалась модель ИНС с одним скрытым слоем, содержащим 5 нейронов. В качестве обучающей выборки использовались те же данные, что и для регрессионного анализа моделей (6) и (7). Использовались настройки функции *MATLAB train(..)*, установленные по умолчанию. Результаты аппроксимации представлены в табл. 3.

Сравнение с результатами из табл. 1 и табл. 2 показывает, что модели ИНС обеспечивают более точную аппроксимацию искомой зависимости

по тем же исходным данным, чем модели линейной регрессии. Отметим, что увеличение числа нейронов до 10 не дало заметного улучшения результатов.

ЗАКЛЮЧЕНИЕ

Проведенное исследование показало, что существует объективная стохастическая зависимость между оценкой степени овражности рельефа ЦФ в окрестности точки окончания поиска минимума и погрешностью определения координат истинного положения минимума ЦФ. Данная зависимость может быть выявлена и зафиксирована в виде уравнения линейной регрессии или в виде обученной модели ИНС, а затем использоваться для оценки ожидаемой погрешности определения координат точки минимума.

Установлено, что модели ИНС обеспечивают более высокую точность предсказания величины погрешности по сравнению с моделями линейной регрессии. Также точность моделей обоих типов повышает учет не только оценки степени овражности, но и эвклидовой нормы градиента ЦФ в точке окончания поиска.

В дальнейшем предполагается расширить методику на функции с размерностью оврагов более единицы и на другие типы сложных для алгоритмов оптимизации участков рельефа.

СПИСОК ЛИТЕРАТУРЫ

1. Гилл Ф., Мюррей У., Райт М. *Практическая оптимизация*; пер. с англ. М.: Мир; 1985. 509 с.
2. Нинул А.С. *Оптимизация целевых функций. Аналитика. Численные методы. Планирование эксперимента*. М.: Изд-во Физико-математической литературы; 2009. 336 с.
3. Карпенко А.П. *Современные алгоритмы поисковой оптимизации. Алгоритмы, вдохновленные природой: учебное пособие*. М.: Изд-во МГТУ им. Н.Э. Баумана; 2017. 446 с.
4. Jasbir S.A. *Introduction to Optimum Design*. 4th edition. Elsevier; 2017. 670 p. <https://doi.org/10.1016/C2013-0-15344-5>
5. Mersmann O., Bischl B., Trautmann H., Preuss M., Weihs C., Rudolf G. Exploratory Landscape Analysis. In: *Proceedings of the 13th Annual Conference on Genetic and Evolutionary Computation (GECCO'11)*. 2011. P. 829–836. <https://doi.org/10.1145/2001576.2001690>
6. Kerschke P., Trautmann H. Comprehensive Feature-Based Landscape Analysis of Continuous and Constrained Optimization Problems Using the R-package flacco. In: *Applications in Statistical Computing—From Music Data Analysis to Industrial Quality Improvement. Series: Studies in Classification, Data Analysis, and Knowledge Organization*. Bauer N., Ickstadt K., Lübke K., Szepannek G., Trautmann H., Vichi M. (Eds.). Berlin/Heidelberg, Germany: Springer; 2019. P. 93–123. https://doi.org/10.1007/978-3-030-25147-5_7
7. Trajanov R., Dimeski S., Popovski M., Korosec P., Eftimov T. *Explainable Landscape-Aware Optimization Performance Prediction*. Preprint. 2021. <https://arxiv.org/pdf/2110.11633v1>
8. Чернолучский И.Г. *Методы оптимизации в теории управления: учебное пособие*. СПб.: Питер; 2004. 256 с.
9. Смирнов А.В. Свойства целевых функций и алгоритмов поиска в задачах многокритериальной оптимизации. *Russ. Technol. J.* 2022;10(4):75–85. <https://doi.org/10.32362/2500-316X-2022-10-4-75-85>
10. He J., Xu T. *New Methods of Studying Valley Fitness Landscapes*. Preprint. 2018. <https://arxiv.org/pdf/1805.00092.pdf>
11. Ochoa G., Veerapen N. Mapping the Global Structure of TSP Fitness Landscapes. *J. Heuristics*. 2018;24(4):265–294. <https://link.springer.com/article/10.1007/s10732-017-9334-0>
12. Hansen N., Finck S., Ros R., Auger A. *Real-Parameter Black-Box Optimization Benchmarking 2009: Noiseless Functions Definitions*. [Research Report] RR-6829, INRIA. 2009. <https://hal.inria.fr/inria-00362633v2>
13. Jamil M., Yang Xin-She. A literature survey of benchmark functions for global optimization problems. *Int. J. Math. Model. Numer. Opt.* 2013;4(2):150–194. <https://doi.org/10.1504/IJMMNO.2013.055204>
14. Большаков А.А., Каримов Р.Н. *Методы обработки многомерных данных и временных рядов*. М.: Горячая линия–Телеком; 2007. 522 с.
15. Хайкин С. *Нейронные сети: полный курс*; пер. с англ. М.: ООО «И.Д. Вильямс»; 2006. 1104 с.

REFERENCES

1. Gill Ph.E., Murray W., Wright M.H. *Prakticheskaya optimizatsiya (Practical Optimization)*. Transl. from Engl. Moscow: Mir; 1985. 509 p. (in Russ.). [Gill Ph.E., Murray W., Wright M.H. *Practical Optimization*. London: Academic Press; 1981. 402 p.]
2. Ninul A.S. *Optimizatsiya tselevykh funktsii. Analitika. Chislennye metody. Planirovanie eksperimenta (Optimization of Objective Functions. Analytics. Numerical Methods. Desing of Experiments)*. Moscow: Fizmatizdat; 2009. 336 p. (in Russ.).
3. Karpenko A.P. *Sovremennye algoritmy poiskovoi optimizatsii. Algoritmy, vdokhnovlennye prirodo: uchebnoe posobie (Modern Search Optimization Algorithms. Nature-Inspired Optimization Algorithms)*. Moscow: Izd. MG TU im. Bauman; 2017. 446 p. (in Russ.).
4. Jasbir S.A. *Introduction to optimum design*. 4th edition. Elsevier; 2017. 670 p. <https://doi.org/10.1016/C2013-0-15344-5>
5. Mersmann O., Bischl B., Trautmann H., Preuss M., Weihs C., Rudolf G. Exploratory Landscape Analysis. In: *Proceedings of the 13th Annual Conference on Genetic and Evolutionary Computation (GECCO'11)*. 2011. P. 829–836. <https://doi.org/10.1145/2001576.2001690>
6. Kerschke P., Trautmann H. Comprehensive Feature-Based Landscape Analysis of Continuous and Constrained Optimization Problems Using the R-package flacco. In: *Applications in Statistical Computing—From Music Data Analysis to Industrial Quality Improvement. Series: Studies in Classification, Data Analysis, and Knowledge Organization*. Bauer N., Ickstadt K., Lübke K., Szepannek G., Trautmann H., Vichi M. (Eds.). Berlin/Heidelberg, Germany: Springer; 2019. P. 93–123. https://doi.org/10.1007/978-3-030-25147-5_7
7. Trajanov R., Dimeski S., Popovski M., Korosec P., Eftimov T. *Explainable Landscape-Aware Optimization Performance Prediction*. Preprint. 2021. <https://arxiv.org/pdf/2110.11633v1>
8. Chernorutskii I.G. *Metody optimizatsii v teorii upravleniya: uchebnoe posobie (Optimization Methods in the Control Theory)*. St. Petersburg: Piter; 2004. 256 p. (in Russ.).
9. Smirnov A.V. Properties of objective functions and search algorithms in multi-objective optimization problems. *Russ. Technol. J.* 2022;10(4):75–85 (in Russ.). <https://doi.org/10.32362/2500-316X-2022-10-4-75-85>
10. He J., Xu T. *New Methods of Studying Valley Fitness Landscapes*. Preprint. 2018. <https://arxiv.org/pdf/1805.00092.pdf>
11. Ochoa G., Veerapen N. Mapping the Global Structure of TSP Fitness Landscapes. *J. Heuristics*. 2018;24(4):265–294. <https://link.springer.com/article/10.1007/s10732-017-9334-0>
12. Hansen N., Finck S., Ros R., Auger A. *Real-Parameter Black-Box Optimization Benchmarking 2009: Noiseless Functions Definitions*. [Research Report] RR-6829, INRIA. 2009. <https://hal.inria.fr/inria-00362633v2>
13. Jamil M., Yang Xin-She. A literature survey of benchmark functions for global optimization problems. *Int. J. Math. Model. Numer. Opt.* 2013;4(2):150–194. <https://doi.org/10.1504/IJMMNO.2013.055204>

14. Bol'shakov A.A., Karimov R.N. *Metody obrabotki mnogomernykh dannykh i vremennykh ryadov (Methods for Processing of Multidimensional Data and Time Series)*. Moscow; Goryachaya Liniya–Telekom; 2007. 522 p. (in Russ.).
15. Haykin S. *Neironnye seti: polnyi kurs (Neural networks. A Comprehensive Foundation)*. Transl. from Engl. Moscow: Vil'yams; 2006. 1104 p. (in Russ.).
[Haykin S. *Neural networks. A Comprehensive Foundation*. Prentice Hall; 1999. 1104 p.]

Об авторе

Смирнов Александр Витальевич, к.т.н., доцент, профессор кафедры телекоммуникаций Института радиоэлектроники и информатики ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: av_smirnov@mirea.ru. Scopus Author ID 56380930700, <https://orcid.org/0000-0002-2696-8592>

About the author

Alexander V. Smirnov, Cand. Sci. (Eng.), Professor, Department of Telecommunications, Institute of Radio Electronics and Informatics, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: av_smirnov@mirea.ru. Scopus Author ID 56380930700, <https://orcid.org/0000-0002-2696-8592>