Санкт-Петербургский государственный университет
Прикладная математика и информатика
Вычислительная стохастика и статистические модели

Барсуков Егор Вячеславович

Оптимальные планы для оценивания производных в полиномиальной регрессионной модели без свободного члена

Отчет о научно-исследовательской работе

Научный руководитель:
д. ф.-м. н., профессор В. Б. Мелас

Санкт-Петербург
2019

Оглавление

Введение

Рассмотрим регрессионную модель

$$y_j = \theta^\top f_j(x_j) + \varepsilon_j, \quad j = 1 \dots N,\, x_j \in \mathcal{X}, \tag{1}$$

где N — количество экспериментов, $\mathcal{X} \subset \mathbb{R}$, $f(x) = (f_1(x), \dots, f_n(x))^\top$ — регрессионная функция, $\theta = (\theta_1, \dots, \theta_n)^\top$ — неизвестные параметры, ε_i — некоррелированные ошибки наблюдения. При этом $\mathrm{E}[\varepsilon_i] = 0$, $\mathrm{D}[\varepsilon_i] = \sigma^2$.

Для того, чтобы при фиксированном количестве наблюдений N получить наилучшую в каком-либо смысле оценку θ строят *планы эксперимента* т.е. наборы точек $x_i \in \mathcal{X}$ в каждой из которых должно быть произведено n_i экспериментов так, что $\sum_i^N n_i = N$. В каком именно смысле будет улучшена оценка параметров модели зависит от выбора критерия одного из нескольких критериев оптимальности.

В работе будут рассматриваться полиномиальные регрессионные модели, т.е. $f_i(x) = x^{k_i}$. Для таких моделей во многих случаях явным образом были описаны оптимальные планы. Несколько работ были посвящены нахождению D-оптимальных планов [1–4]. Также существуют для такой модели явные решения для нахождения E-оптимальных планов [5–8].

В этой работе будут рассматриваться c-оптимальные планы эксперимента. Ими являются планы минимизирующие дисперсию значения скалярного произведения θ и c для заданного $c \in \mathbb{R}^n$ [8]. В общем случае нахождение c-оптимальных планов может быть достаточно сложно: для случаев малой размерности решение можно найти используя теорему Элвинга [9], однако явного решения для произвольного c не существует.

В практических приложениях важны несколько частных случаев: $c = f(z)$ для некоторого $z \notin \mathcal{X}$ — задача экстраполяции в точке z и $c = f'(z)$ — задача оценки производной в точке z. Для обычной полиномиальной модели оптимальный план экстраполяции был описан достаточно давно [10], также существует несколько явных решений для задачи оценки производной в некоторых случаях [11, 12].

В этой работе рассмотрен случай нахождения нахождения планов для оценки производной при полиномиальной модели без свободного члена при $\mathcal{X} = [0, d]$. Такая модель, например, может быть использована в тех случаях, когда исходя из практической задачи значения z могут принимать только положительные значения и существует априорное знание о значении функции в точке 0. Простым примером такой функции является зависимости расстояния до начальной точки от времени.

Глава 1

С-оптимальные планы эксперимента

1.1. Определения

Определение 1. Согласно [13] непрерывным планом в регрессионной модели (1) будем называть дискретную вероятностную меру

$$\xi = \begin{pmatrix} x_1 & \dots & x_m \\ \omega_1 & \dots & \omega_m \end{pmatrix}, \quad x_i \in \mathcal{X},$$

где $\omega_i \geqslant 0$, $\sum_i^n \omega_i = 1$, $i = 1, \dots, m$.

Для проведения N измерений с мерой ξ необходимо провести $n_i \approx N\omega_i$ измерений в точке x_i таким образом, чтобы $\sum_i^m n_i = N$.

Определение 2. Информационной матрицей для непрерывного плана эксперимента заданного мерой ξ является

$$M(\xi) = \int_{\mathcal{X}} f(x) f^\top(x) \xi(dt).$$

Определение 3. *с-оптимальным планом* для некоторого вектора c называется план эксперимента ξ минимизирующий следующую функцию

$$\Phi(\xi) = \begin{cases} c^\top M(\xi)^- c, & \text{если существует } v\text{, такой, что } \ c = M(\xi)v \\ +\infty, & \text{иначе} \end{cases},$$

где $M(\xi)^-$ — матрица, обобщенно обратная к информационной матрице плана ξ. План называется допустимым, если существует такое v, что $c = M(\xi)v$.

Как было отмечено во введении, c-оптимальный план минимизирует дисперсию несмещенной оценки по методу наименьших квадратов $c^\top\hat{\theta}$ линейной комбинации $c^\top\theta$ [8].

Определение 4. Если $c = f(z)$ для некоторого $z \in \mathbb{R}$, то соответствующий c-оптимальный план называется *оптимальном планом экстраполяции* в точке z.

Определение 5. Если $c = f'(z)$ для некоторого $z \in \mathbb{R}$, то соответствующий c-оптимальный план называется *оптимальным планом для оценки производной* в точке z.

1.2. Теорема Элвинга

Для решения задачи нахождения c-оптимальных планов во множестве случаев (в том числе в данной работе) используется теорема Элвинга, являющаяся геометрически интерпретируемым критерием c-оптимальности плана эксперимента.

Теорема 1. (Элвинга) [11] Допустимый план $\xi^\star$ с носителем $x_1, \ldots, x_m \in \mathcal{X}$ и весами $\omega_1, \ldots, \omega_m$ является c-оптимальным тогда и только тогда, когда существует $p \in \mathbb{R}^k$ и константа h такие, что выполняются следующие условия:

$$\left|p^\top f(x_i)\right| = 1 \qquad i = 1..m \leqslant n \tag{1.1a}$$

$$\left|p^\top f(x)\right| \leqslant 1 \qquad x \in \mathcal{X} \tag{1.1b}$$

$$c = h \sum_{i=1}^{m} \omega_i f(x_i) p^\top f(x_i). \tag{1.1c}$$

Кроме того

$$h^2 = c^\top M^-(\xi^*) c$$

Функцию $p^\top f(x_i)$ в определениях теоремы Элвинга в этой работе также будет называться *экстремальным многочленом.*

1.3. Явная формула для весов оптимального плана

Теорема 2. Оптимальный план для оценивания производной полиномиальной модели без свободного члена с опорными точками $t_1^*, \ldots, t_m^*$, где $m = n$ или $m = n - 1$ имеет веса вычисленные по следующей формуле:

$$\omega_i = \frac{|L_i'(z)|}{\sum_{j=1}^{m} \left|L_j'(z)\right|}, \tag{1.2}$$

где L_i задается следующим образом

$$L_i(x) = \frac{x \prod_{l=1}^{n} (x - t_l^*)}{t_i^* \prod_{l \neq i}^{n} (t_i^* - t_l^*)}, \tag{1.3}$$

то есть является i-ым базисным многочленом Лагранжа без нулевого члена построенным по точкам $t_1^*, \ldots, t_m^*$.

Глава 2

План для нахождения производной на промежутке с началом в нуле

Здесь будут описаны оптимальные планы для нахождения производной на промежутке вида $[0, d]$ в модели без нулевого члена. Для промежутков вида $[-1, 1]$ такие планы были описаны в [12], однако в таком случае полученные решения значительно отличаются от полученных в этом разделе. Здесь, не умаляя общности, будет доказана теорема для случая $[0, 1]$, а в конце главы будет показано, как он переносится в общий вид.

Везде в этой главе считается, что $f(x) = (x, x^2, \ldots, x^n)^\top$, то есть рассматривается модель без нулеого члена.

2.1. План для нахождения производной на промежутке $[0, 1]$

При доказательстве основной теоремы этой главы будет использована следующая лемма.

Лемма 1. Пусть $P_1(x)$ и $P_2(x)$ — многочлены степени n с корнями $t_1^1 < \ldots < t_n^1$ и $t_1^2 < \ldots < t_n^2$ соответственно. При этом корни располагаются следующим образом:

$$t_1^1 \leqslant t_1^2 < t_2^1 \leqslant t_2^2 < \ldots < t_n^1 \leqslant t_n^2,$$

где хотя бы одно из неравенств $t_l^1 \leqslant t_l^2$ $(l = 1, \ldots, n)$ является строгим. Также обозначим корни многочленов $P_1'(x)$ и $P_2'(x)$ как $v_1^1, \ldots, v_{n-1}^1$ и $v_1^2, \ldots, v_{n-1}^2$. Тогда справедливо следующее выражение

$$v_1^1 < v_1^2 < v_2^1 < v_2^2 < \ldots < v_n^1 < v_n^2.$$

Доказательство этого утверждения можно найти в [14] или в приложении к [12].

Теорема 3. Пусть для $i = 1, \ldots, n$

$$x_i^* = \frac{\cos\frac{(n-i)\pi}{n} + cos\frac{\pi}{2n}}{1 + \cos\frac{\pi}{2n}}, \tag{2.1}$$

и корни производной многочлена L_i из формулы (1.3) построенного по точкам $\{x_j^*\}_j^n$ равны $u_1^i, \ldots, u_{n-1}^i$, причем $u_k^i \leqslant u_l^i$ при $k < l$. Тогда при $z \in \big(-\infty, u_1^1) \cup (u_{n-1}^n, +\infty\big)$

или $z \in \left(u_{n-1}^{j+1}, u_1^j\right)$ для $j = 1, \ldots, n-1$ оптимальный план для нахождения производной в точке z в полиномиальной модели степени n без свободного члена на промежутке $[0, 1]$ имеет опорные точки $\{x_j\}_j^n$ и веса вычисленные по формуле (1.2). Также при z не лежащих в этих промежутках не существует оптимального плана состоящего из n точек.

Доказательство. Полученное здесь решение основывается на применении варианта теоремы Элвинга описанного на странице 6, которая, в том числе, утверждает, что для любого c-оптимального плана должен существовать соответствующий многочлен $p^\top f(x)$, который не превосходит по модулю единицу на соответствующем промежутке и достигает её в опорных точках плана. Поэтому первым шагом будет нахождение таких многочленов степени n.

Построим полином без свободного члена $S_n(x)$ степени n, не превосходящий по модули единицу на промежутке $[0, 1]$, и достигающий её в n точках. Пусть $T_n(x)$ — многочлен Чебышёва степени n. Тогда по свойствам многочленов Чебышёва T_n не превосходит по модулю единицу на промежутке $[-1, 1]$ и достигает её в $n+1$ точках, в том числе в точках -1 и 1. Известно, что корни T_n имеют следующий вид

$$t_i = \cos\left(\frac{\pi(i + 1/2)}{n}\right), \quad i = 0, \ldots, n-1.$$

Если мы возьмём самый маленький корень $t_{\min} = -\cos\frac{\pi}{2n}$ и положим $\widehat{S_n}(x) = T_n(x+t_{\min})$, то $\widehat{S_n}$ будет являться полиномом степени n с нулевым свободным членом, так как $\widehat{S_n}(0) = 0$ по построению. При этом $\left|\widehat{S_n}(x)\right| \leqslant 1$ для x на промежутке $[0, 1+\cos\left(\frac{\pi}{2n}\right)]$ и в этом промежутке абсолютная величина достигает единицы n раз в силу того, что левый край не равен 1.

Для того, чтобы привести промежуток к виду $[0, 1]$ достаточно добавить множитель $1 + \cos\frac{\pi}{2n}$ к x в левой части определения $\widehat{S_n}$. После этого получается многочлен, удовлетворяющий всем требуемым свойствам

$$S_n(x) = T_n\left(x\left(1 + \cos\frac{\pi}{2n}\right) - \cos\frac{\pi}{2n}\right).$$

Экстремальные точки $T_n(x)$ расположены в точках

$$\widehat{x_i^*} = \cos\frac{(n-i)\pi}{n}, \quad i = 0, \ldots, n,$$

поэтому экстремальные точки $S(x)$ на промежутке $[0,1]$ расположены в точках

$$x_i^* = \frac{\cos\frac{(n-i)\pi}{n} + \cos\frac{\pi}{2n}}{1+\cos\frac{\pi}{2n}}, \quad i = 1,\ldots,n,$$

при этом важно, что индексы у x_i^* начинаются с 1, в то время как у $\widehat{x_i^*}$ с 0. Так происходит из-за того, что наименьшая экстремальная точка S_n на всем промежутке оказывается меньше нуля и не попадает в требуемый промежуток.

Также нужно отметить, что

$$0 < x_1^* < x_2^* < \ldots < x_n^*, \quad i = 1,\ldots,n$$

и

$$S_n(x_i^*) = (-1)^{n+i}, \quad i = 1,\ldots,n$$

Покажем, что $S_n(x)$ — это единственный многочлен (с точностью до знака) без нулевого члена, удовлетворяющий свойствам свойствам (1.1a) и (1.1b) на промежутке $[0,1]$. Пусть $P_n(x)$ — многочлен степени n и при этом $P_n(0) = 0$, $P_n(1) = 1$ и в промежутке $[0,1)$ этот многочлен имеет ровно $n-1$ экстремум $t_1 < \ldots < t_{n-1}$ в которых он равен ± 1. В силу того, что нам требуется единственность с точностью до знака, тот факт, что $P_n(1)$ был положен положительным, не умаляет общности. Так как $P_n'(x)$ имеет степень $n-1$ и, соответственно, имеет $n-1$ корней, которые были обозначены как $t_1,\ldots,t_{n-1}$, то только в этих точках этот многочлен может менять свою монотонность, но также $|P_n(t_i)| = 1$ для $i = 1,\ldots,n-1$, поэтому в этих точках он меняет свою монотонность и $P_n(t_i) = (-1)^{n-1-i}$. Также поэтому при $x \leqslant t_1$ (ранее было определено, что $t_1 > 0$) $\operatorname{sign}(P_n'(x)) = (-1)^n$, а учитывая, что $P(0) = 0$ и $P(t_1) = (-1)^n$, то должна существовать такая единственная точка $t_0 < 0$, что $P(t_0) = (-1)^{n+1}$. После введения обозначения $t_n = 1$ из факта, что $P_n(t_i) = (-1)^{n-1-i}$ для $i = 0,\ldots,n$ следует, что $P_n(x)$ является чебышевским многочленом на отрезке $[t_0, 1]$, то есть он единственен и $P_n = S_n$.

Из единственности многочлена без нулевого члена степени n, удовлетворяющего свойствам (1.1a) и (1.1b) на отрезке $[0,1]$ следует, что оптимальный план состоящий из n опорных точек сосредоточен в экстремальных точках многочлена $S_n(x)$, которые имеют вид (2.1).

Осталось показать, когда выполняется условие (1.1c) при $c = f'(z)$ и, соответственно, точки (2.1) являются опорными точками оптимального плана для оценки производной в точке z.

Построим базисные полиномы Лагранжа степени n без нулевого члена по точкам $\{x_i^*\}_{i=1}^n$

$$L_i(x) = \frac{x\prod_{l=1}^n (x - x_l^*)}{x_i^8 \prod_{l\neq i}^n (x_i^* - x_l^*)}$$

Так как теорема 4 о весах оптимального плана для оценки производной в случае полиномиальной модели работает для любых промежутков, то для нахождения весов можно использовать её, и тогда, соответственно, веса имеют следующий вид

$$\omega_i = \frac{|L_i'(z)|}{\sum_{j=1}^n \left|L_j'(z)\right|}.$$

Введем обозначения $F = \left((x_j^*)^i\right)_{i,j=1}^n$ и $\beta = (\omega_i(-1)^{i+n})_{i=1}^n$. Так как x_j^* при $i = 1,\ldots,n$ являются экстремальными точками многочлена S_n и при этом $S_n(x_j^*) = (-1)^{j+n}$, то выполнение равенства для некоторого h

$$f'(z) = hF\beta \tag{2.2}$$

эквивалентно выполнению условия (1.1c) теоремы Элвинга для нахождения оптимального плана оценки производной в точке z.

Утверждение $F^{-1}F = I_n$, где I_n — единичная матрица размера n, а i–ый столбец матрицы F на самом деле равен $f(x_i^*)$ $(i = 1,\ldots,n)$ можно переписать, как систему равенств

$$e_i^\top F^{-1} f(x_j^*) = \delta_{ij}\,, \quad i,j = 1,\ldots,n,$$

где δ_{ij} — дельта Кронекера, а e_i — i–ый единичный вектор. Поскольку в левой части равенств (**??**) содержатся многочлены без нулевого коэффициента степени не больше n вычисленные в точках x_j^*, $j = 1,\ldots,n$, а для каждого i существует только одно j, такое, что $\delta_{ij} \neq 0$, то они определяют все базисные многочлены Лагранжа без нулевого члена степени n вычисленные в точках x_j^*, $j = 1,\ldots,n$, таким образом

$$e_i^\top F^{-1} f(z) = L_i(z)\,, \quad i,j = 1,\ldots,n.$$

Если в предыдущем выражении вычислить производную по z и переписать полученное выражение в векторной форме получим

$$f'(z) = F\left(L_1'(z),\ldots,L_n'(z)\right)^\top. \tag{2.3}$$

Приравняв правые части (2.3) и (2.2) и домножив равенство на F^{-1} слева, получаем, что

$$h\beta = \left(L_1'(z),\ldots,L_n'(z)\right)^\top,$$

что с учетом введенных ранее обозначений влечет, что $\mathrm{sign}(L_i'(z)) = \mathrm{sign}((-1)^{i+n})$, $i = 1, \dots, n$ или, вспомнив, что экстремальным многочленом также может быть $-S(x)$, $\mathrm{sign}(L_i'(z)) = \mathrm{sign}((-1)^{i+n+1})$, $i = 1, \dots, n$.

Таким образом для того, чтобы доказать, что оптимальный план находится в точках $(x_i^*)_{i=1}^n$, $i = 1, \dots, n$ с указными ранее весами, осталось доказать равенство знаков $L_i'(z)$ и $\pm S_n(x_i^*)$. Но так как знаки экстремальных точек многочлена S_n чередуются, достаточно показать при каких z выражения $(-1)^i L_i'(z)$ имеет одинаковый знак для $i = 1, \dots, n$.

Так как корни многочленов L_i и L_j для любых i и j таких, что $i < j$ удовлетворяют условию леммы 1, условие которой было приведено на странице 7, то последовательно её применяя ко всем базисным многочленам получаем, что корни их производных, обозначения для которых были описаны в условии этой теоремы, удовлетворяют следующему соотношению:

$$u_1^n < u_1^{n-1} < \ldots < u_1^1 < u_2^n < u_2^{n-1} < \ldots < u_{n-1}^1.$$

Можно видеть, что, так как все узловые точки больше нуля, знак многочлена $L_i(z)$ при $z \to -\infty$ будет равен $(-1)^{n+i+1}$. В то же время знак $L_i'(z)$ будет противоположным $L_i(z)$, так как меняется четность многочлена и при этом не меняется знак при старшем коэффициенте, то есть $\mathrm{sign}(L_i'(z)) = \mathrm{sign}((-1)^{n+i})$ при $z \to -\infty$. И, следовательно, $\mathrm{sign}((-1)^i L_i'(z)) = \mathrm{sign}((-1)^{n+2i}) = \mathrm{sign}((-1)^n)$ при $z \to -\infty$, то есть $\mathrm{sign}((-1)^i L_i'(z))$ не зависит от i и имеет постоянный знак для любых i, что означает, что при $z \in (-\infty, u_1^n)$ третье условие теоремы Элвинга выполняется и план является оптимальным.

Осталось изучить как ведут себя знаки $\mathrm{sign}((-1)^i L_i'(z))$ на остальных промежутках. На промежутках $[u_j^1, u_j^n]$ $(j = 1, \ldots n-1)$ каждый базисный многочлен меняет свой знак ровно 1 раз и на этих промежутках знаки производных не совпадают со знаками экстремального многочлена, а на промежутках (u_j^n, u_{j+1}^1) $(j = 1, \ldots n-2)$ нет ни одного корня и поэтому $\mathrm{sign}((-1)^i L_i'(z)) = \mathrm{sign}((-1)^{n+j}$ при $z \in (u_j^n, u_{j+1}^1)$ $(j = 1, \ldots n-2)$, что также подтверждает третье условие теоремы Элвинга в этих промежутках и показывает, что показанный план оптимален.

На промежутке $(-\infty, u_{n-1}^n]$ каждый базисный многочлен поменял свой знак одинаковое количество раз, а так как при $z \to -\infty$ условие выполнялось, то при $z \in (u_{1,n-1}, +\infty)$ план также является оптимальным.

Таким образом план эксперимента с опорными точками (2.1) и соответствующими им весами (1.2) является оптимальным планом для оценки производной в точке z в модели без нулевого члена тогда и только тогда, когда $z \in (-\infty, u_1^n) \cup (u_{1,n-1}, +\infty)$ или $z \in (u_{j+1}^n, u_j^1)$ для $j = 1, \dots n-2$. Причем в силу единственности (с точностью до знака) экстремального многочлена с n экстремальными точками, не существует других планов состоящих из n опорных точек, что доказывает теорему.

$\square$

2.2. Промежуток вида $[0, d]$

В общем случае оптимальный для всех промежутков, начинающихся в нуле, существенно не отличается от случая промежутка $[0, 1]$, что будет показано в следующей теореме.

Теорема 4. Пусть для $i = 1, \dots, n$ и для некоторого d

$$\widehat{x_i^*} = dx_i^*, \tag{2.4}$$

где x_i^* из (2.1), и корни производной многочлена L_i из формулы (1.3) построенного по точкам $\{\widehat{x_j^*}\}_j^n$ равны $u_1^i, \dots, u_{n-1}^i$, причем $u_k^i \leqslant u_l^i$ при $k < l$. Тогда при $z \in \big(-\infty, u_1^1) \cup (u_{n-1}^n, +\infty\big)$ или $z \in (u_{j+1}^n, u_j^1)$ для $j = 1, \dots, n-2$ оптимальный план для нахождения производной в точке z в полиномиальной модели степени n без свободного члена на промежутке $[0, 1]$ имеет опорные точки $\{x_j\}_j^n$ и веса вычисленные по формуле (1.2). Также при z не лежащих в этих промежутках не существует оптимального плана состоящего из n точек.

Доказательство. Построим многочлен $S_n^d(x) = S_n(\frac{x}{d})$. Из построения ясно что его экстремальные точки равны (2.4). Он по построению будет являться экстремальным многочленом степени n на отрезке $[0, d]$ и будет повторять все свойства многочлена $S_n(x)$. Дальнейшее доказательство повторяет доказательство теоремы 3 с заменой многочлена S_n на S_n^d. $\square$

Список литературы

1. Hoel P. G. Efficiency problems in polynomial estimation // Annals of Mathematical Statistics. — 1958. — Vol. 29, no. 4. — P. 1134–1145.
2. Studden W. J. D_s-optimal designs for polynomial regression using continued fractions // Annals of Statistics. — 1980. — Vol. 8, no. 5. — P. 1132–1141.
3. Dette H. A generalization of D- and D_1-optimal designs in polynomial regression // Annals of Statistics. — 1990. — Vol. 18. — P. 1784–1805.
4. Dette H., Franke T. Robust designs for polynomial regression by maximizing a minimum of D- and D_1-efficiencies // Annals of Statistics. — 1990. — Vol. 29, no. 4. — P. 1024–1049.
5. Pukelsheim F., Studden W. J. E-optimal designs for polynomial regression // Annals of Statistics. — 1990. — Vol. 21, no. 1. — P. 402–415.
6. Dette H. A note on E-optimal designs for weighted polynomial regression // Annals of Statistics. — 1990. — Vol. 21, no. 2. — P. 767–771.
7. Heiligers B. E-optimal designs in weighted polynomial regression // Annals of Statistics. — 1994. — Vol. 22, no. 2. — P. 917–929.
8. Dette H., Studden W. J. Geometry of E-optimality // Annals of Statistics. — 1993. — Vol. 21, no. 1. — P. 416–433.
9. Elfving G. Optimal allocation in linear regression theory // The Annals of Mathematical Statistics. — 1952. — Vol. 23. — P. 255–262.
10. Hoel P. G., Levine A. Optimal spacing and weighting in polynomial prediction // The Annals of Statistics. — 1964. — Vol. 35, no. 4. — P. 1553–1560.
11. Dette H., Melas V., A. P. Optimal designs for estimating the slope of a regression // Statistics. — 2010. — Vol. 44, no. 6. — P. 617–628.
12. Dette H., Melas V., Shpilev P. Some explicit solutions ofc-optimal design problems for polynomial regression with no intercept // Annals of the Institute of Statistical Mathematics. — 2019.
13. Kiefer J. General equivalence theory for optimum designs (approximate theory) // The Annals of Statistics. — 1974. — Vol. 2, no. 5. — P. 416–433.
14. Sahm M. Optimal designs for estimating individual coefficients in polynomial regression : Ph. D. thesis / M. Sahm ; Fakultät für Mathematik, Ruhr-Universität Bochum, Germany. — 1998.

15. L-bfgs-b - fortran subroutines for large-scale bound constrained optimization : Rep. / ACM Trans. Math. Software ; Executor: Ciyou Zhu, Richard H. Byrd, Peihuang Lu, Jorge Nocedal : 1994.
16. Storn R., Price K. Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces // Journal of Global Optimization. — 1997. — Dec. — Vol. 11, no. 4. — P. 341–359.
17. Lawson C. L., Hanson R. J. Solving least squares problems. — Siam, 1995. — Vol. 15.