Белорусский государственный университет
Механико-математический факультет
Кафедра диф. уравнений и системного анализа

Метод эллипсоидов

Чергинец Дмитрий Николаевич

Положительно определенная матрица

Вещественная матрица D называется положительно определенной,

если она симметричная $(D = D^T)$ и все ее собственные значения положительны

Следующие утверждения эквивалентны:

1) D – положительно определенная;
2) D^{-1} – положительно определенная;
3) $\exists M$ – невырожденная матрица, что $D = MM^T$;
4) D – симметричная и

$$\forall x \in \mathbb{R}^n \; (x \neq 0 \implies x^T D x > 0).$$

Эллипсоид

Эллипсоидом называется

$$E(z, D) := \{x \in \mathbb{R}^n \mid (x - z)^T D^{-1}(x - z) \leq 1\} =$$

$$= \{x \in \mathbb{R}^n \mid x = z + My, ||y|| \leq 1\}.$$

$z \in \mathbb{R}^n$ – центр эллипсоида,
M – $n \times n$ невырожденная матрица,
$D := MM^T$.

Полупространство

Полупространством называется

$$\Gamma(a,\beta) := \{x \in \mathbb{R}^n \mid a^T x \le a^T z + \beta\sqrt{a^T D a}\}.$$

$a \in \mathbb{R}^n$ – нормальный вектор плоскости.
$\beta \in \mathbb{R}$.

Если $\beta \geq 1$, то

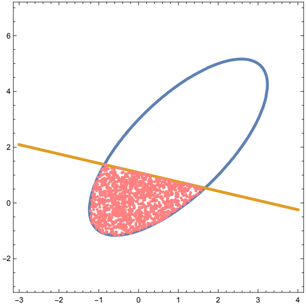
$$E(z,D) \cap \Gamma(a,\beta) = E(z,D).$$

Если $\beta < -1$, то

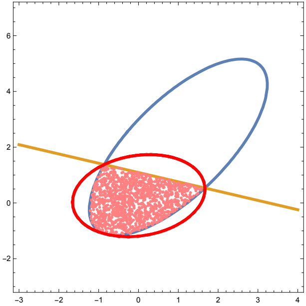
$$E(z,D) \cap \Gamma(a,\beta) = \emptyset.$$

Отсечение эллипсоида

Вход



Выход



Найти эллипсоид $E(z_k, D_k)$ наименьшего объема, описанный вокруг эллипсоида $E(z, D)$, отсеченного полупространством $\Gamma(a, \beta)$:

$$E(z, D) \cap \Gamma(a, \beta) \subset E(z_k, D_k),$$

$$V(E(z_k, D_k)) \to \min .$$

Искомый эллипсоид

$$E(z, D) \cap \Gamma(a, \beta) \subset E(z_k, D_k),$$

$$z_k := z + \frac{\beta n - 1}{n+1} \frac{Da}{\sqrt{a^T D a}},$$

$$D_k := \frac{n^2(1-\beta^2)}{n^2-1} \left(D - \frac{2(1-n\beta)}{(1-\beta)(n+1)} \frac{Daa^T D}{a^T D a} \right).$$

Здесь вектор-столбец a умножается на вектор-строку a^T, в результате чего получается aa^T – матрица $n \times n$.
Формулы действительны при $-1 < \beta \leq \frac{1}{n}$.

Полиэдр

Полиэдром называется

$$P = P(A, b) := \{x \in \mathbb{R}^n \mid Ax \leq b\}.$$

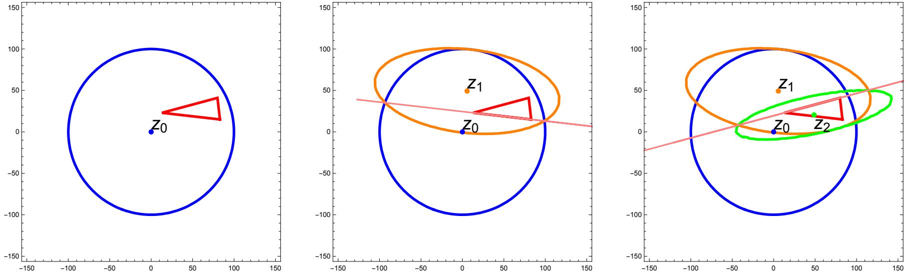
Здесь A – $m \times n$ матрица,
$b \in \mathbb{R}^m$ – вектор-столбец.
Полиэдр P это пересечение полупространств:

$$P = \bigcap_{i=1}^{m} \{x \in \mathbb{R}^n \mid A_i x \leq b_i\}.$$

Необходимо найти какую-нибудь точку, принадлежащую ограниченному полиэдру.

Метод эллипсоидов

Графическая интерпретация поиска точки полиэдра методом эллипсоидов.



Норма матрицы

Пусть $x \in \mathbb{R}^n$ – вектор-столбец, тогда

$$\|x\| := \sqrt{x^T x}.$$

Спектральной нормой матрицы A наз-ся

$$\|A\| := \max_{\|x\| \neq 0} \frac{\|Ax\|}{\|x\|} = \sqrt{\Lambda(A^T A)},$$

$\Lambda(A^T A)$ – наибольшее собственное значение $A^T A$.

Согласованность норм

Неравенство Коши-Шварца

$$\forall x, y \in \mathbb{R}^n \ (|x^T y| \leq \|x\| \|y\|).$$

Согласованность нормы матрицы и вектора

$$\|Ax\| \leq \|A\| \|x\|.$$

Норма полож. опр. матрицы

Если A – симметричная матрица, то

$$\|A\| = \max\{|\lambda| \mid \lambda - \text{собств. знач. } A\} =$$
$$= \max\{x^T A x \mid \|x\| = 1\}.$$

Если A – симметричная матрица,
λ_1 – максимальное собств. значение A,
λ_n – минимальное собств.значение A, то

для всех $x \in \mathbb{R}^n$ справедливы неравенства

$$\lambda_n \|x\|^2 \le x^T A x \le \lambda_1 \|x\|^2.$$

Норма обратной матрицы

A – невырожденная

λ – собств. знач. $A \quad \Leftrightarrow \quad \lambda^{-1}$ – собств. знач. A^{-1}.

Если A – симметр. невыр. матрица, то

$$\|A^{-1}\| = \max\{|\lambda|^{-1} \mid \lambda - \text{собств. знач. } A\} =$$

$$= \max_{\|x\| \neq 0} \frac{x^T x}{x^T A x}.$$

Норма Фробениуса

Норма Фробениуса

$$\|A\|_F := \sqrt{\sum_{i,j=1}^{n} a_{i,j}^2}.$$

Эквивалентность норм

$$\frac{1}{\sqrt{n}}\|A\|_F \leq \|A\| \leq \|A\|_F.$$

Округление эллипсоида

Пусть $\alpha \in \mathbb{N}$.
Найдем такие $p \in \mathbb{Z}$, $\rho \in \mathbb{R}$, что округлив элементы матриц D, z с точностью до 10^{-p} :

$$\tilde{D} := \frac{Round(10^p D)}{10^p}, \qquad \tilde{z} := \frac{Round(10^p z)}{10^p},$$

будем иметь:
1) $\tilde{D}$ – положит. определенная (после округления у $\tilde{D}$ могут появиться отриц. собств. значения).
2) справедливое включение $E(z, D) \subset E(\tilde{z}, \rho\tilde{D})$.
3) ограниченный рост эллипсоида $E(\tilde{z}, \rho\tilde{D})$

$$\frac{V(E(\tilde{z}, \rho\tilde{D}))}{V(E(z, D))} \leq \exp(10^{-\alpha}).$$

$\tilde{D}$ – положительно опред.?

Обозначим через
$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n > 0$ – собств. значения D.
Пусть $\|x\| = 1$. Тогда $\lambda_n \leq x^T D x \leq \lambda_1$.
Справедлива оценка

$$x^T \tilde{D} x = x^T \tilde{D} x - x^T D x + x^T D x \geq$$

$$\geq \lambda_n - |x^T (\tilde{D} - D) x| \geq \lambda_n - 1\|\tilde{D} - D\|1 \geq$$
$$\geq \lambda_n - \|\tilde{D} - D\|_F \geq \lambda_n - n10^{-p}$$

Если $\lambda_n - n10^{-p} > 0$, то все собственные значения матрицы $\tilde{D}$ больше нуля и она положительно определенная.

Радиус ρ

Задача

Пусть $p \in \mathbb{Z}$,

$$\tilde{D} := \frac{Round(10^p D)}{10^p}, \qquad \tilde{z} := \frac{Round(10^p z)}{10^p},$$

Необходимо найти такое $\rho = \rho(p)$, что будет справедливо включение

$$E(z, D) \subset E(\tilde{z}, \rho\tilde{D}).$$

Нахождение радиуса ρ

Пусть

$$D = B^2,$$

где B – положительно определенная матрица.

$B^{-1}\tilde{D}B^{-1}$ – симметричная

B^{-1}, $\tilde{D}$ – симметричные, поэтому

$$(B^{-1}\tilde{D}B^{-1})^T = B^{-1}\tilde{D}B^{-1}.$$

Следовательно, $B^{-1}\tilde{D}B^{-1}$ – симметричная.

Оценка собств. знач. $B^{-1}\tilde{D}B^{-1}$

Пусть $\|x\| = 1$.

$$|x^T B^{-1}\tilde{D}B^{-1}x - 1| = |x^T B^{-1}\tilde{D}B^{-1}x - x^T B^{-1}DB^{-1}x| =$$

$$= |x^T B^{-1}(\tilde{D} - D)B^{-1}x| \leq \|x\|^2\|B^{-1}\|^2\|\tilde{D} - D\| =$$

$$= \|D^{-1}\|\|\tilde{D} - D\| \leq \lambda_n^{-1}\|\tilde{D} - D\|_F \leq \lambda_n^{-1}n10^{-p}.$$

Получили неравенство

$$1 - \lambda_n^{-1}n10^{-p} \leq x^T B^{-1}\tilde{D}B^{-1}x \leq 1 + \lambda_n^{-1}n10^{-p}.$$

Оценка нормы $\|B\tilde{D}^{-1}B\|$

Из неравенства

$$1 - \lambda_n^{-1} n 10^{-p} \leq x^T B^{-1} \tilde{D} B^{-1} x \leq 1 + \lambda_n^{-1} n 10^{-p}$$

следует, что все собственные значения матрицы $B^{-1}\tilde{D}B^{-1}$ лежат в промежутке

$$[1 - \lambda_n^{-1} n 10^{-p}, 1 + \lambda_n^{-1} n 10^{-p}].$$

Собственные значения взаимно обратных матриц взаимнообратны, поэтому собственные значения матрицы $B\tilde{D}^{-1}B$ принадлежат интервалу

$$\left[\frac{1}{1 + \lambda_n^{-1} n 10^{-p}}, \frac{1}{1 - \lambda_n^{-1} n 10^{-p}}\right].$$

Нахождение радиуса ρ

Пусть $x \in E(z, D)$, т.е. $(x-z)^T D^{-1}(x-z) \le 1$.
Поэтому $\|B^{-1}(x-z)\|^2 = (x-z)^T D^{-1}(x-z) \le 1$.
Оценим

$$(x-\tilde{z})^T(\rho\tilde{D})^{-1}(x-\tilde{z}) =$$

$$= \rho^{-1}(x-\tilde{z})^T B^{-1}(B\tilde{D}^{-1}B)B^{-1}(x-\tilde{z}) \le$$

$$\le \rho^{-1}\|B\tilde{D}^{-1}B\|\|B^{-1}(x-\tilde{z})\|^2 \le$$

$$\le \frac{\rho^{-1}}{1-\lambda_n^{-1}n10^{-p}}\left(\|B^{-1}(x-z)\| + \|B^{-1}(z-\tilde{z})\|\right)^2 \le$$

$$\le \frac{(1+\lambda_n^{-1/2}\sqrt{n}10^{-p})^2}{\rho(1-\lambda_n^{-1}n10^{-p})}.$$

Формула для радиуса ρ

Потребуем, чтобы

$$\frac{(1+\lambda_n^{-1/2}\sqrt{n}10^{-p})^2}{\rho(1-\lambda_n^{-1}n10^{-p})} \le 1.$$

Тогда

$$(x-\tilde{z})^T(\rho\tilde{D})^{-1}(x-\tilde{z}) \le 1,$$

и $x \in E(\tilde{z}, \rho\tilde{D})$.

Получили формулу для радиуса

$$\rho := \frac{(1+\lambda_n^{-1/2}\sqrt{n}10^{-p})^2}{1-\lambda_n^{-1}n10^{-p}}.$$

Где $p \in \mathbb{N}$,

$$\lambda_n^{-1}n10^{-p} < 1/2.$$

Сравнение объемов

$$\frac{V(E(\tilde{z}, \rho\tilde{D}))}{V(E(z, D))} = \sqrt{\frac{\det(\rho\tilde{D})}{\det D}} = \sqrt{\det(\rho B^{-1}\tilde{D}B^{-1})} \leq$$

$\leq$ (определитель матрицы равен произведению собственных значений) $\leq$

$$\leq \rho^{1/2}(1 + \lambda_n^{-1} n 10^{-p})^{n/2}.$$

Оценка $\rho^{1/2}$

$$\rho^{1/2} = \frac{1 + {\lambda_n^{-1/2}}\sqrt{n}10^{-p}}{(1 - \lambda_n^{-1} n 10^{-p})^{1/2}}.$$

При $\lambda_n^{-1} n 10^{-p} < 1/2$ справедлива оценка

$$\frac{1}{(1 - \lambda_n^{-1} n 10^{-p})^{1/2}} <$$

$$< [\frac{1}{1-x} = 1 + \frac{x}{1-x} < 1 + 2x \text{ при } x < 1/2] <$$

$$< (1 + 2\lambda_n^{-1} n 10^{-p})^{1/2} \leq [1 + x \leq \exp(x)] \leq$$

$$\leq \exp(\lambda_n^{-1} n 10^{-p}).$$

Сравнение объемов

$$\frac{V(E(\tilde{z},\rho\tilde{D}))}{V(E(z,D))} \leq$$

$$\leq \rho^{1/2}(1+\lambda_n^{-1}n10^{-p})^{n/2} =$$

$$= \frac{(1+\lambda_n^{-1/2}\sqrt{n}10^{-p})(1+\lambda_n^{-1}n10^{-p})^{n/2}}{(1-\lambda_n^{-1}n10^{-p})^{1/2}} <$$

$$< \exp(\lambda_n^{-1}n10^{-p}+\lambda_n^{-1/2}\sqrt{n}10^{-p}+1/2\lambda_n^{-1}n^2 10^{-p}).$$

Выберем $p \in \mathbb{Z}$ таким, чтобы

$$\frac{V(E(\tilde{z},\rho\tilde{D}))}{V(E(z,D))} \leq \exp(10^{-\alpha}), \qquad \alpha \in \mathbb{N}.$$

Формула для p

p должно удовлетворять следующим неравенствам:

$$10^{-p} < \lambda_n / n,$$

$$10^{-p} < \lambda_n / (2n),$$

$$10^{-p} < 10^{-\alpha} \lambda_n / (n + \lambda_n^{1/2} \sqrt{n} + 1/2 n^2).$$

Если выполняется третье неравенство, то выполняются и первые два. Выражая из третьего неравенства p, получаем

$$p > \log_{10}(\lambda_n^{-1} n + \lambda_n^{-1/2} \sqrt{n} + 1/2 \lambda_n^{-1} n^2) + \alpha.$$

$$p := Ceiling[\log_{10}(\lambda_n^{-1} n + \lambda_n^{-1/2} \sqrt{n} + 1/2 \lambda_n^{-1} n^2) + \alpha].$$

Выбор α

Имеем оценку

$$\frac{V(E(\tilde{z}_k, \rho \tilde{D}_k))}{V(E(z_{k-1}, D_{k-1}))} =$$

$$= \frac{V(E(\tilde{z}_k, \rho \tilde{D}_k))}{V(E(z_k, D_k))} \frac{V(E(z_k, D_k))}{V(E(z_{k-1}, D_{k-1}))} <$$

$$< \exp\left(10^{-\alpha} - \frac{(1-n\beta)^2}{2(n+1)}\right).$$

Выберем α таким, чтобы

$$\frac{V(E(\tilde{z}_k, \rho \tilde{D}_k))}{V(E(z_{k-1}, D_{k-1}))} < M_0 < 1.$$

Формула для α

$$10^{-\alpha} - \frac{(1-n\beta)^2}{2(n+1)} < 0.$$

Если $\beta \le \frac{1}{n+1}$, то

$$-\frac{(1-n\beta)^2}{2(n+1)} < -\frac{(1-n/(n+1))^2}{2(n+1)} = -\frac{1}{2(n+1)^3}.$$

Тогда

$$10^{-\alpha} < \frac{1}{2(n+1)^3};$$

$$\alpha > \log_{10}(2(n+1)^3);$$

$$\alpha := Ceiling(\log_{10}(2(n+1)^3)) + 5.$$

Сравнение $\tilde{\lambda}_1$ и $\tilde{\lambda}_2$

Определение $\tilde{\lambda}_1$, $\tilde{\lambda}_2$

$$\tilde{\lambda}_1 = \frac{n(\beta+1)}{n+1}, \qquad \tilde{\lambda}_2^2 = \frac{n^2(1-\beta^2)}{n^2-1}.$$

Сравним $\tilde{\lambda}_1$ и $\tilde{\lambda}_2$:

$$\frac{\tilde{\lambda}_1^2}{\tilde{\lambda}_2^2} = \frac{n^2(\beta+1)^2(n^2-1)}{(n+1)^2 n^2(1-\beta^2)} = \frac{(\beta+1)(n-1)}{(n+1)(1-\beta)}.$$

$$(\beta+1)(n-1) - (n+1)(1-\beta) = 2(n\beta-1) < 0,$$

$$\beta < 1/n \quad \Rightarrow \quad \tilde{\lambda}_1 < \tilde{\lambda}_2.$$

Норма матрицы L

Определение L

На прошлой лекции мы получили

$$L = diag(\tilde{\lambda}_1, \tilde{\lambda}_2, \tilde{\lambda}_2, \ldots, \tilde{\lambda}_2).$$

Вычисление $\|L\|$

L – симметричная, поэтому

$$\|L\| = \max\{\tilde{\lambda}_1, \tilde{\lambda}_2\} = \tilde{\lambda}_2,$$

$$\|L^{-1}\| = \max\left\{\frac{1}{\tilde{\lambda}_1}, \frac{1}{\tilde{\lambda}_2}\right\} = \frac{1}{\tilde{\lambda}_1} = \frac{n+1}{n(\beta+1)}.$$

Норма ортогональной матрицы

Матрица U называется ортогональной, если

$$UU^T = U^T U = E.$$

Вычислим $\|U\|$

$$\|U\| := \max_{x \neq 0} \frac{\|Ux\|}{\|x\|} = \max_{x \neq 0} \frac{\sqrt{x^T U^T U x}}{\|x\|} =$$

$$= \max_{x \neq 0} \frac{\sqrt{x^T x}}{\|x\|} = 1.$$

Норма матрицы M^{-1}

Напомним, что D имеет собственные знач.

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n > 0.$$

Поэтому D^{-1} имеет собственные значения

$$\frac{1}{\lambda_n} \geq \frac{1}{\lambda_{n-1}} \geq \dots \geq \frac{1}{\lambda_1} > 0.$$

Пусть $D = MM^T$.

$$\|M^{-1}\| = \sqrt{\Lambda((M^{-1})^T M^{-1})} = \sqrt{\Lambda(D^{-1})} = \frac{1}{\sqrt{\lambda_n}}.$$

Собств. знач. λ_n и $\lambda_{n,k}$

Пусть $\lambda_{n,k}$ наименьшее собственное значение матрицы D_k, которая вычисляется по формуле

$$D_k := \frac{n^2(1-\beta^2)}{n^2-1}\left(D - \frac{2(1-n\beta)}{(1-\beta)(n+1)}\frac{Daa^TD}{a^TDa}\right).$$

λ_n – наименьшее собственное значение D.

Сравним λ_n и $\lambda_{n,k}$. Отметим, что

$$\|D^{-1}\| = \frac{1}{\lambda_n}, \qquad \|D_k^{-1}\| = \frac{1}{\lambda_{n,k}}.$$

Сравнение λ_n и λ_n^k

На прошлой лекции мы получили

$$M_k = MULU^T.$$

Оценим

$$\frac{1}{\sqrt{\lambda_{n,k}}} = \|M_k^{-1}\| = \|UL^{-1}U^TM^{-1}\| \leq$$

$$\leq \|U\|\|L^{-1}\|\|U^T\|\|M^{-1}\| = \frac{n+1}{n(\beta+1)}\frac{1}{\sqrt{\lambda_n}}.$$

Доказали неравенство

$$\lambda_{n,k} \geq \frac{n^2(\beta+1)^2}{(n+1)^2}\lambda_n.$$

Метод глубоких отсечений

Вход

A – матрица $m \times n$ с целыми коэффициентами,
$b \in \mathbb{R}^m$,
$z_0 \in \mathbb{R}^n$ – центр начального эллипсоида,
D_0 – положительно определенная матрица $n \times n$.

Предполагается, что полиэдр

$$P = P(A, b) = \{x \mid Ax \leq b\}$$

ограниченный и телесный (содержит внутреннюю точку или, что эквивалентно, $V(P) > 0$).

$E(z_0, D_0)$ – начальный эллипсоид, $P \subset E(z_0, D_0)$.

Дополнительные параметры

Достаточно малый параметр $\varepsilon > 0$

Будем считать, что

$$V(P) > \sqrt{\varepsilon} V(B(0,1)).$$

Если на k-ой итерации мы получили, что

$$\det D_k < \varepsilon,$$

то выдаем результат

$$V(P) < \sqrt{\varepsilon} V(B(0,1))$$

или P – неограничен.

Дополнительные параметры

Параметр $j \in \mathbb{N}$

Вместо того, чтобы на каждой итерации вычислять наименьшее собственное значение $\lambda_{n,k}$ матрицы D_k, можно оценивать $\lambda_{n,k}$ снизу при помощи формулы

$$\lambda_{n,k} = \frac{n^2(\beta+1)^2}{(n+1)^2}\lambda_{n,k-1}.$$

Если $Mod(k,j) = 0$, то вычислять наименьшее собственное значение матрицы D_k непосредственно по определению.

Метод глубоких отсечений

Выход

$z_k \in P$.

Шаг 1. Задаем начальные значения

$k := 0;$
$\alpha := \mathit{Ceiling}(\log_{10}(2(n+1)^3)) + 5.$
Вычисляем наименьшее собственное значение λ_n матрицы D_0.

Метод глубоких отсечений

Шаг 2. Проверка $z_k \in P$?

Находим такое i, что

$$A_i z_k > b_i.$$

Если такого i не существует,

то выдаем результат $z_k \in P$. Конец алгоритма.

Иначе

$k := k + 1;$

$a := A_i^T;$

$\beta := \frac{b_i - a^T z_{k-1}}{\sqrt{a^T D_{k-1} a}}.$

Метод глубоких отсечений

Шаг 3. Вычисляем новый эллипсоид

$$z_k := z_{k-1} + \frac{\beta n-1}{n+1}\frac{D_{k-1}a}{\sqrt{a^T D_{k-1}a}};$$

$$D_k := \frac{n^2(1-\beta^2)}{n^2-1}\left(D_{k-1} - \frac{2(1-n\beta)}{(1-\beta)(n+1)}\frac{D_{k-1}aa^T D_{k-1}}{a^T D_{k-1}a}\right).$$

Шаг 4. Вычисляем наименьшее собственное значение λ_n матрицы D_k

$$\lambda_n = \frac{n^2(\beta+1)^2}{(n+1)^2}\lambda_n. \qquad (*)$$

Если $Mod(k,j)=0$, то вычисляем λ_n матрицы D_k непосредственно по определению, наперед зная, что оно не меньше λ_n, вычисленного по $(*)$.

Метод глубоких отсечений

Шаг 5. Округляем эллипсоид

$$p := Ceiling[\log_{10}(\lambda_n^{-1}n + \lambda_n^{-1/2}\sqrt{n} + 1/2\lambda_n^{-1}n^2) + \alpha];$$

$$\rho := \frac{(1 + \lambda_n^{-1/2}\sqrt{n}10^{-p})^2}{1 - \lambda_n^{-1}n10^{-p}};$$

$$D_k := \frac{Round(10^p \rho D_k)}{10^p};$$

$$z_k := \frac{Round(10^p z_k)}{10^p}.$$

Метод глубоких отсечений

Шаг 6.

Если $Det(D_k) < \varepsilon$,

то выдаем результат

$$V(P) < \sqrt{\varepsilon} V(B(0, 1))$$

или P – неограничен.

Конец алгоритма.

Иначе переходим к шагу 2.