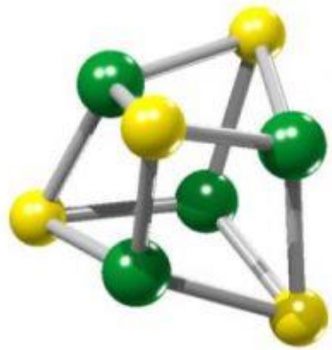
МЕТОД НЕРАВНОМЕРНЫХ ПОКРЫТИЙ ДЛЯ ЗАДАЧ ОПТИМИЗАЦИИ С ОДНИМ И НЕСКОЛЬКИМИ КРИТЕРИЯМИ

Ю.Г. Евтушенко, А.И. Голиков,

М.А. Посыпкин

Вычислительный центр им. А.А. Дородницына РАН

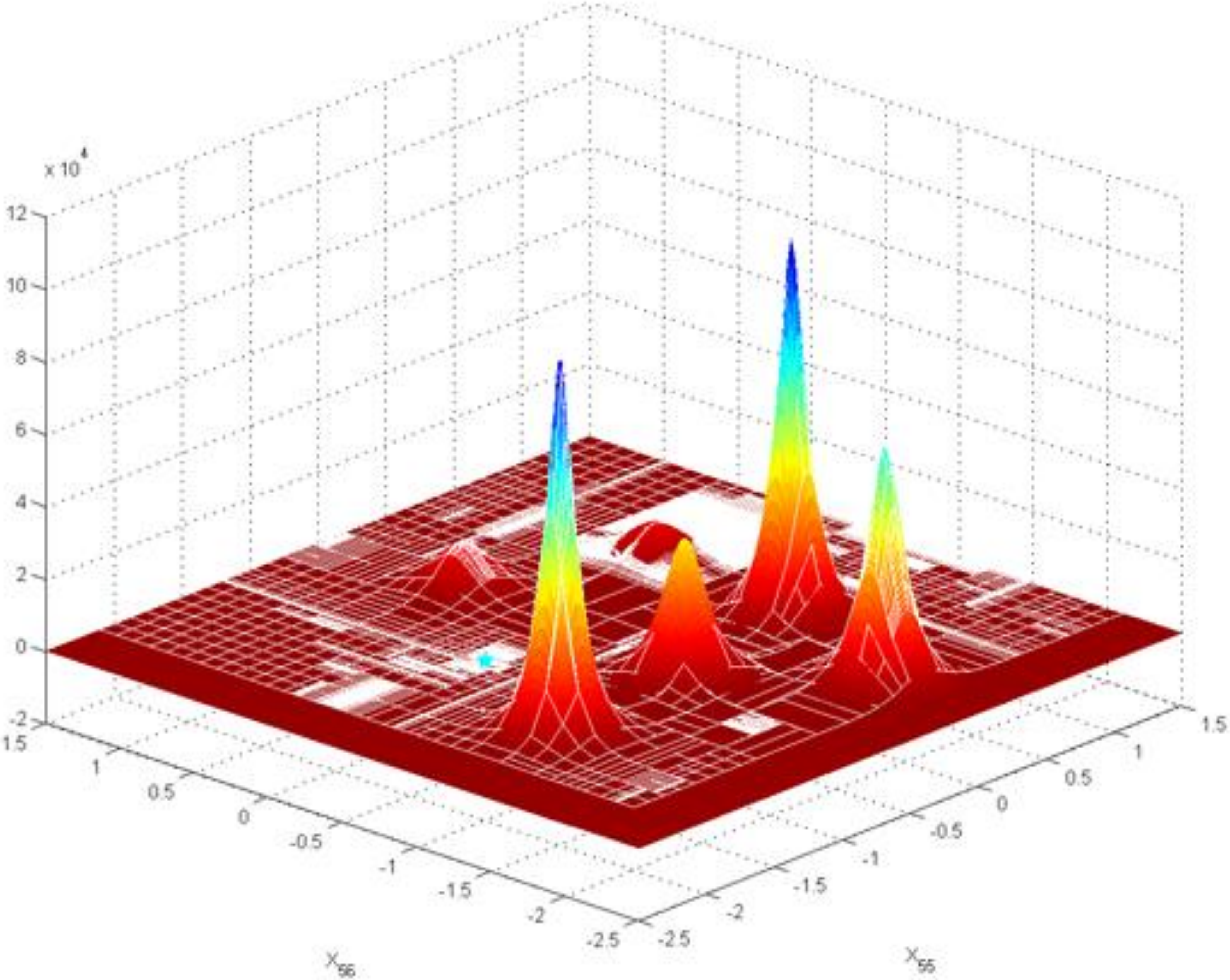
Москва 2014



Функция энергии молекулярного кластера (потенциал Морзе):

$$F(\rho)=\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left(\left(e^{\rho\left(1-\left\|X_i-X_j\right\|\right)}-1\right)^2-1\right)$$

где $\boldsymbol{\rho}$ – скалярный параметр,
$\boldsymbol{x}_i$ и $\boldsymbol{x}_j$ – трехмерные векторы координат центров аминокислот *i* и *j*, соответственно.

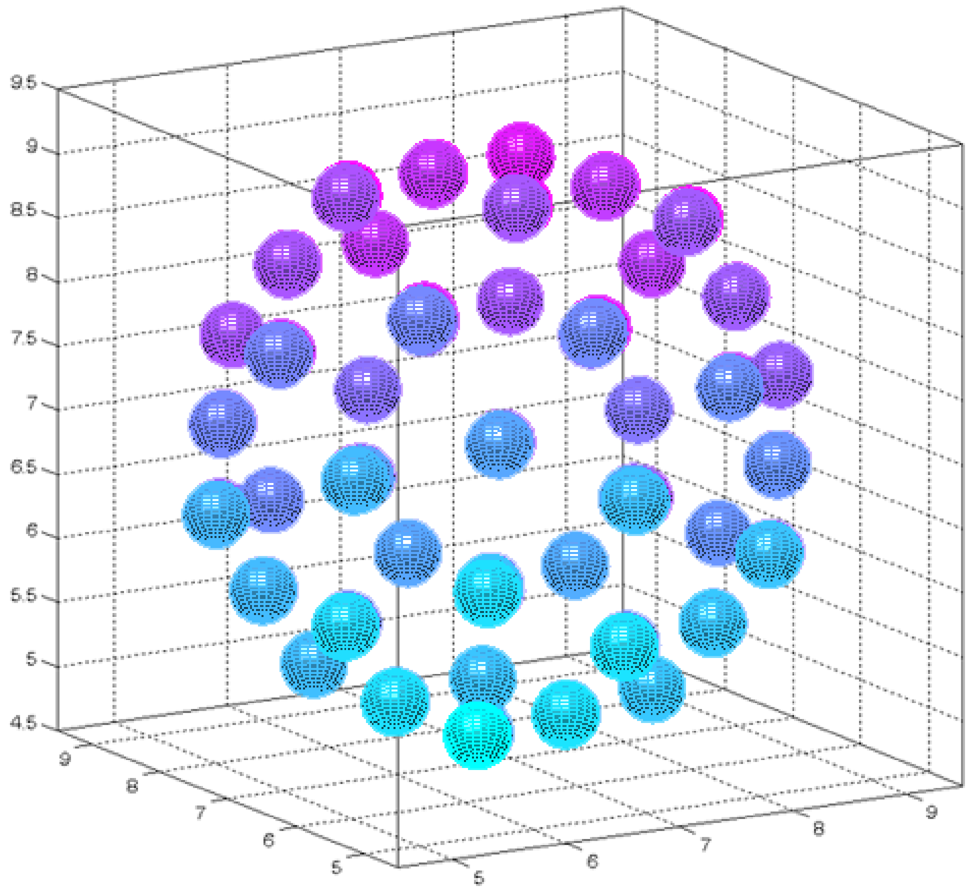


x 10^4
12
10
8
6
4
2
0
-2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
X_{56}
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
X_{55}

ПРИМЕР ОПТИМАЛЬНОЙ КОНФИГУРАЦИИ

Молекулярный кластер *85* атомов, потенциал Морзе,

r = *6*, $f_* = -405.25$ *(А.А. Станевичюс, В.У. Малкова)*



ГЛОБАЛЬНАЯ МИНИМИЗАЦИЯ ФУНКЦИИ

$$f_* = f(x_*) = glob\min_{x \in X} f(x)$$

$X_* = \{x_* \in X : f(x_*) = f_*\}$ - множество решений

$X_\varepsilon = \{x_\varepsilon \in X : f_* \le f(x_\varepsilon) \le f_* + \varepsilon\}$ - множество ε - решений

$$X_* \subseteq X_\varepsilon \subseteq X$$

$$f_* \ge f(x_\varepsilon) - \varepsilon$$

Условия глобальной оптимальности

1. Критерий глобальной оптимальности

$$x_* \in X_* \Leftrightarrow L(f(x), X, f_*) = X$$

$$f(x) \geq f_* \quad \forall x \in X$$

2. Критерий глобальной ε -оптимальности

$$x_\varepsilon \in X_\varepsilon \Leftrightarrow L(f(x), X, f(x_\varepsilon) - \varepsilon) = X$$

$$f(x) \geq f(x_\varepsilon) - \varepsilon \ \forall x \in X \longrightarrow f_* \geq f(x_\varepsilon) - \varepsilon \longrightarrow x_\varepsilon \in X_\varepsilon$$

3. Для любого набора множеств $\{X_i\}$, $X_i \subseteq X$, $\bigcup_{i=1}^{k} X_i = X$

справедливо $$x_\varepsilon \in X_\varepsilon \Leftrightarrow \bigcup_{i=1}^{k} L(f(x), X_i, f(x_\varepsilon) - \varepsilon) = X$$

Набор точек $x_1, x_2, \ldots, x_k \in X$

набор множеств $X_1, X_2, \ldots, X_k \subset X, x_i \in X_i$

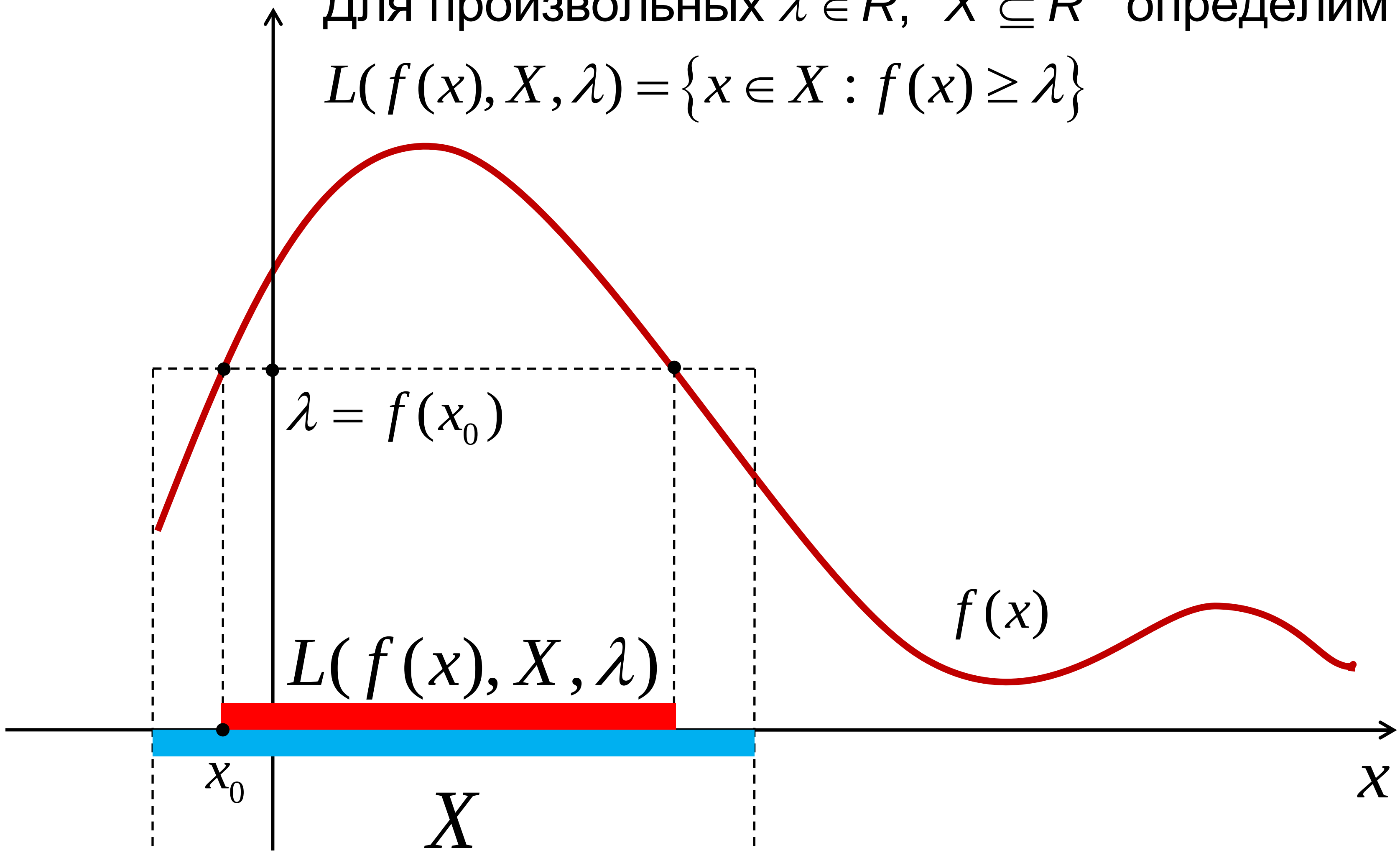
рекордная точка $x_{rk} = \arg\min_{1 \le i \le k} f(x_i)$

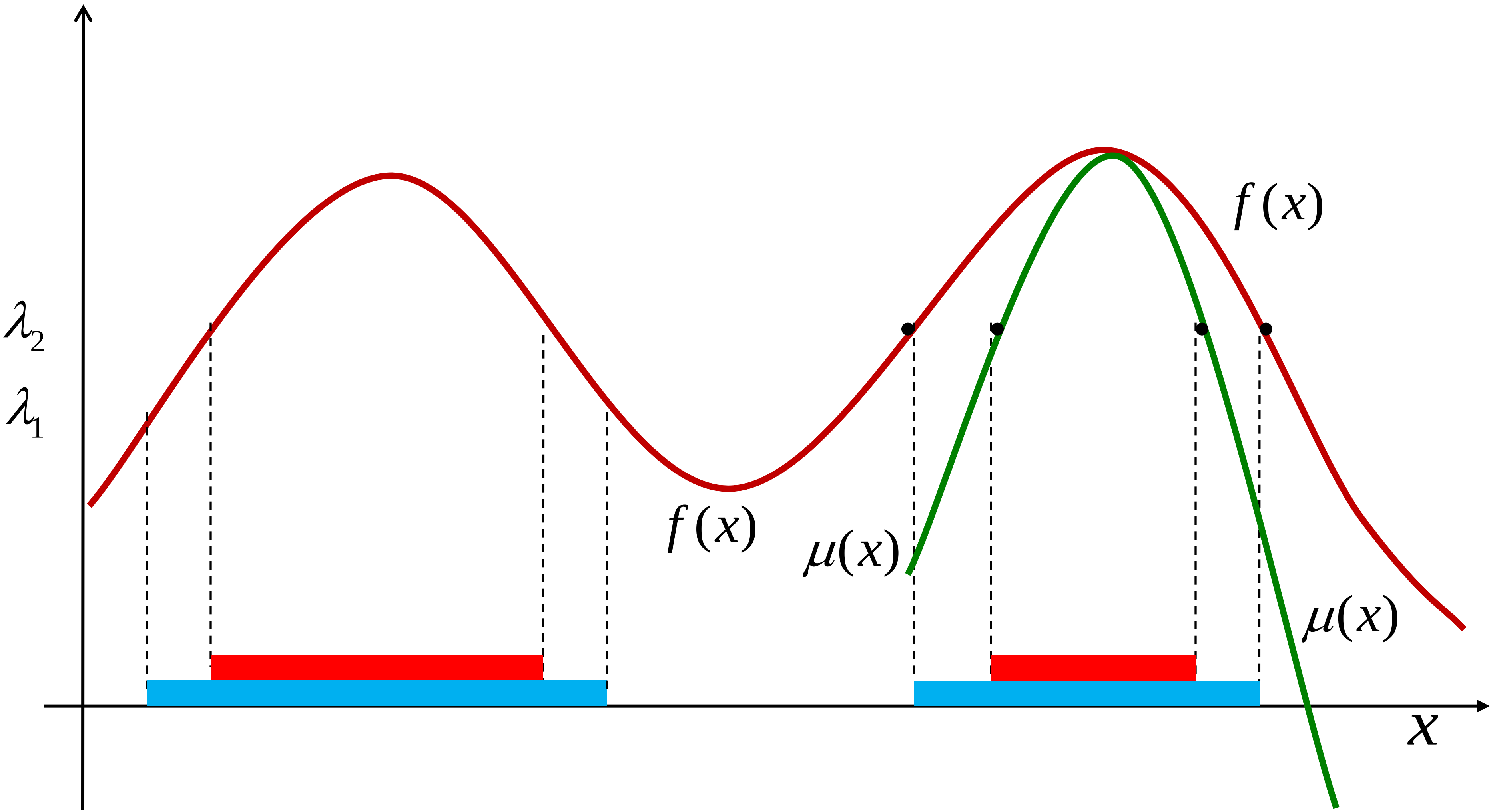
$$f(x_{rk}) \ge f(x_{r(k+1)})$$

Лебеговское множество

Для произвольных $\lambda \in R$, $X \subseteq R^n$ определим

$$L(f(x), X, \lambda) = \{x \in X : f(x) \geq \lambda\}$$



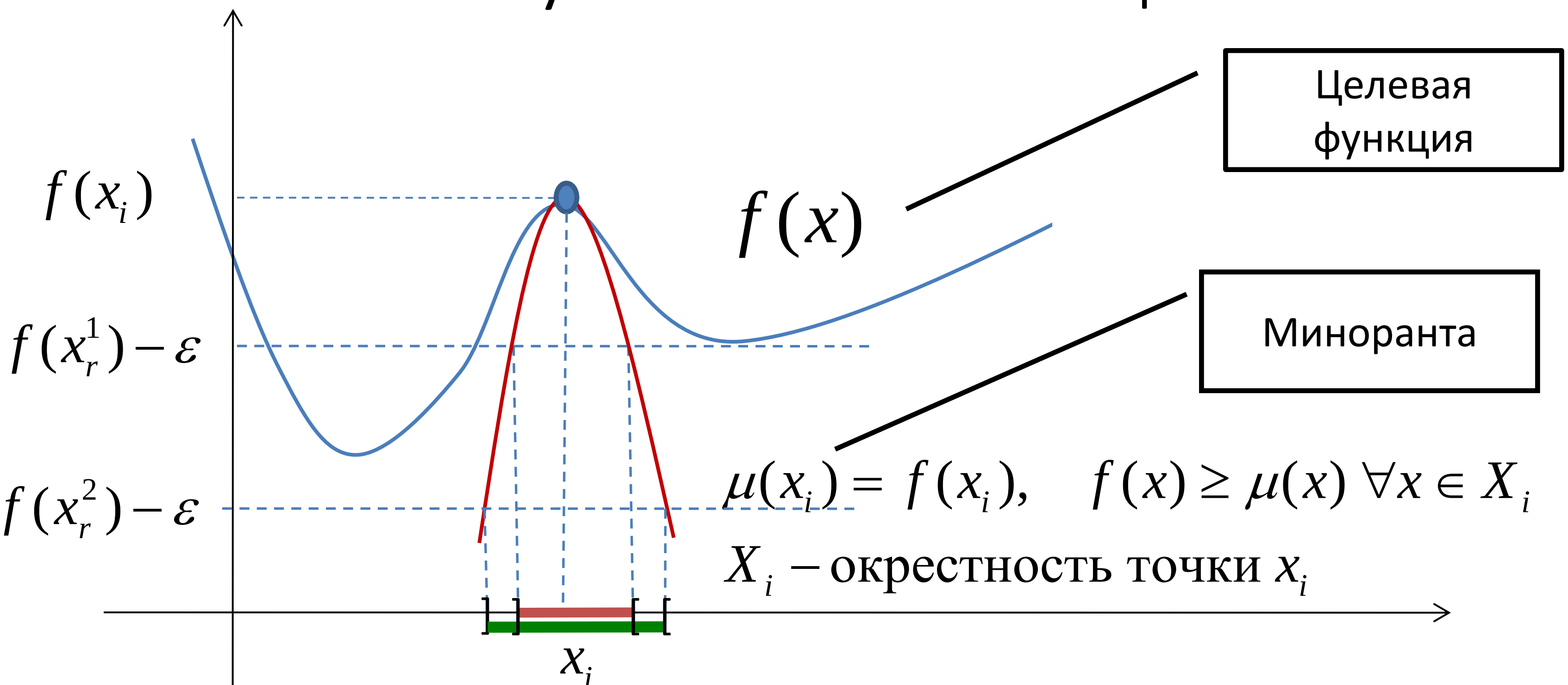


$$L(\ldots\lambda_2) \subseteq L(\ldots\lambda_1) \qquad L(\mu(x)\ldots) \subseteq L(f(x)\ldots)$$

Если X покрыто с уровнем λ_1,

то оно будет покрыто при всяком $\lambda_2 \le \lambda_1$

Метод неравномерных покрытий для одномерной безусловной оптимизации



$$\mu(x_i) = f(x_i), \quad f(x) \geq \mu(x)\ \forall x \in X_i$$

X_i – окрестность точки x_i

$$f(x_r^1) > f(x_r^2)$$

ОСНОВНАЯ ТЕОРЕМА

$$X_1,\dots,X_k \subseteq R^n, \qquad x_i \in X_i, \qquad x_{rk} = \arg\min_{1\le i\le k} f(x_i)$$

$$L(\mu_i(x), X_i, f(x_{ri}) - \varepsilon) = L_{ri}$$

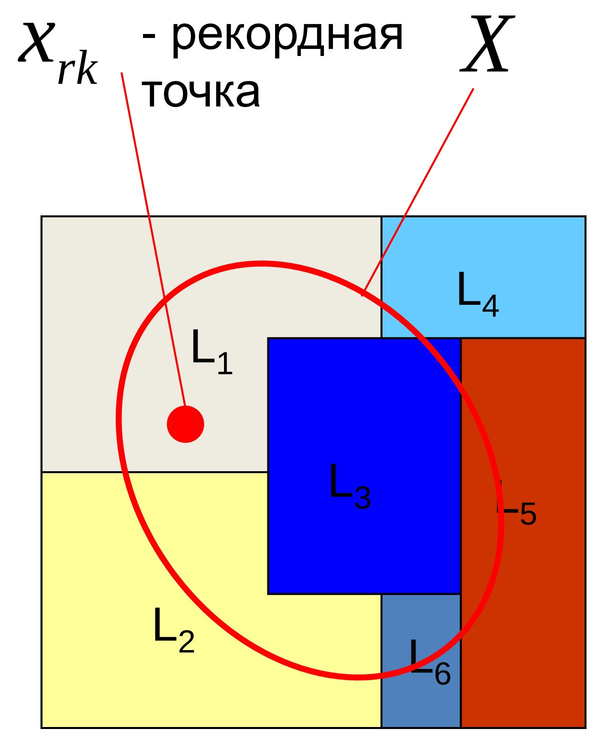
Теорема 1. Если выполнено

$$f(x) \ge \mu_i(x),\ \forall x \in X_i,\ X \subseteq \bigcup_{i=1}^{k} L_{ri},$$

то $\exists s : x_* \in L_{rs}$

$$f(x_s) \ge f_* \ge f(x_s) - \varepsilon,$$

поэтому $x_s \in X_\varepsilon$



МИНОРАНТА 1

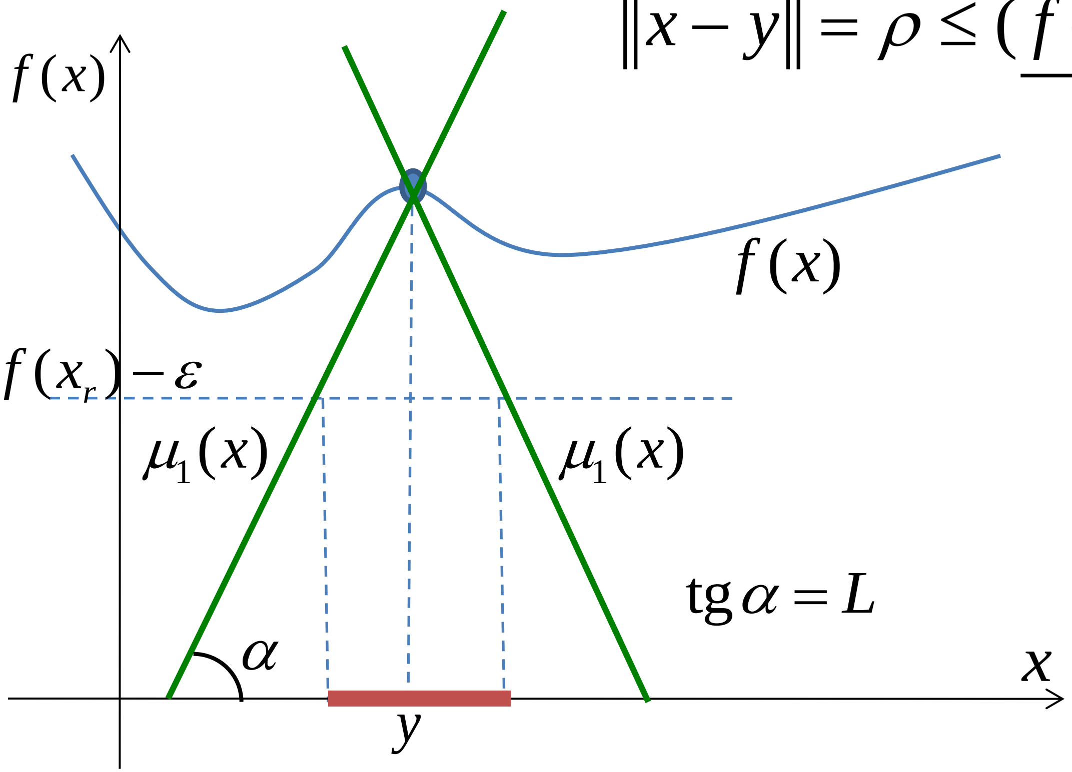
Условие Липшица: $\left| f(x) - f(y) \right| \le L \| x - y \|, x, y \in X_i$

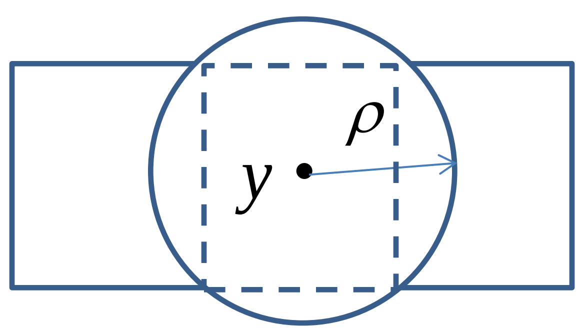
Миноранта: $f(x) \ge f(y) - L\|x-y\| = \mu_1(x) \ge f(x_r) - \varepsilon$

Можно исключить из рассмотрения шар радиуса

$$\|x-y\| = \rho \le (\underbrace{f(y) - f(x_r)}_{\ge 0} + \varepsilon)/L \ge \varepsilon / L$$

с центром в точке y.





МИНОРАНТА 2

Градиент удовлетворяет условию Липшица $\|f_x(x) - f_x(y)\| \le L\|x - y\|$

Миноранта $f(x) \ge \mu_2(x) = f(y) + \langle f_x(y), x - y \rangle - \frac{L}{2}\|x - y\|^2$

Шар радиуса

$$\rho = \sqrt{\frac{2}{L}\left(f(y) + \frac{1}{2L}\|f_x(y)\|^2 - f(x_r) + \varepsilon \right)^3} \sqrt{\frac{2\varepsilon}{L}}$$

с центром в точке $c = y + f_x(y)/L$

может быть исключен из дальнейшего рассмотрения.

МИНОРАНТА 3

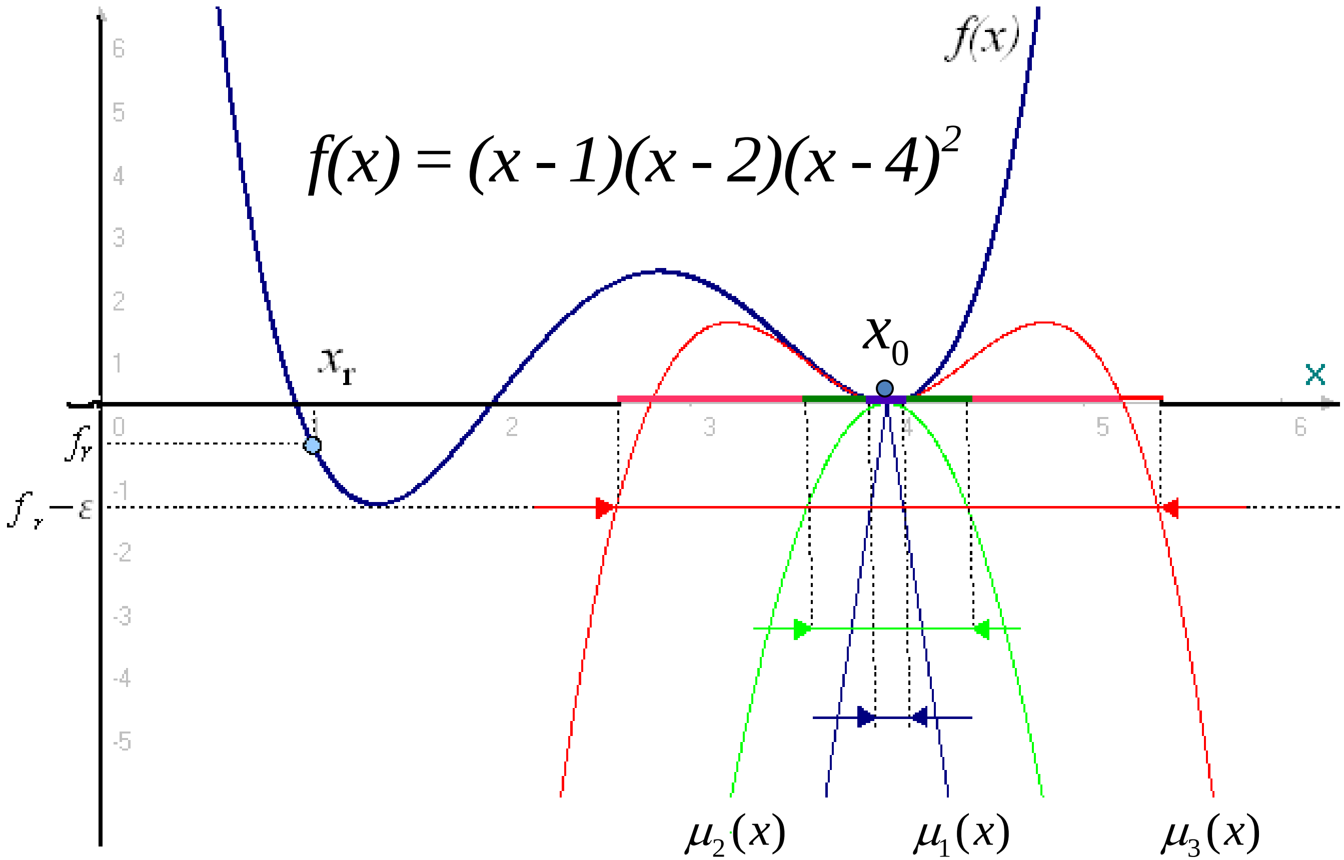
Гессиан удовлетворяет условию Липшица

$$\| f_{xx}(x) - f_{xx}(y) \| \le M \|x - y\|, \; x, y \in X$$

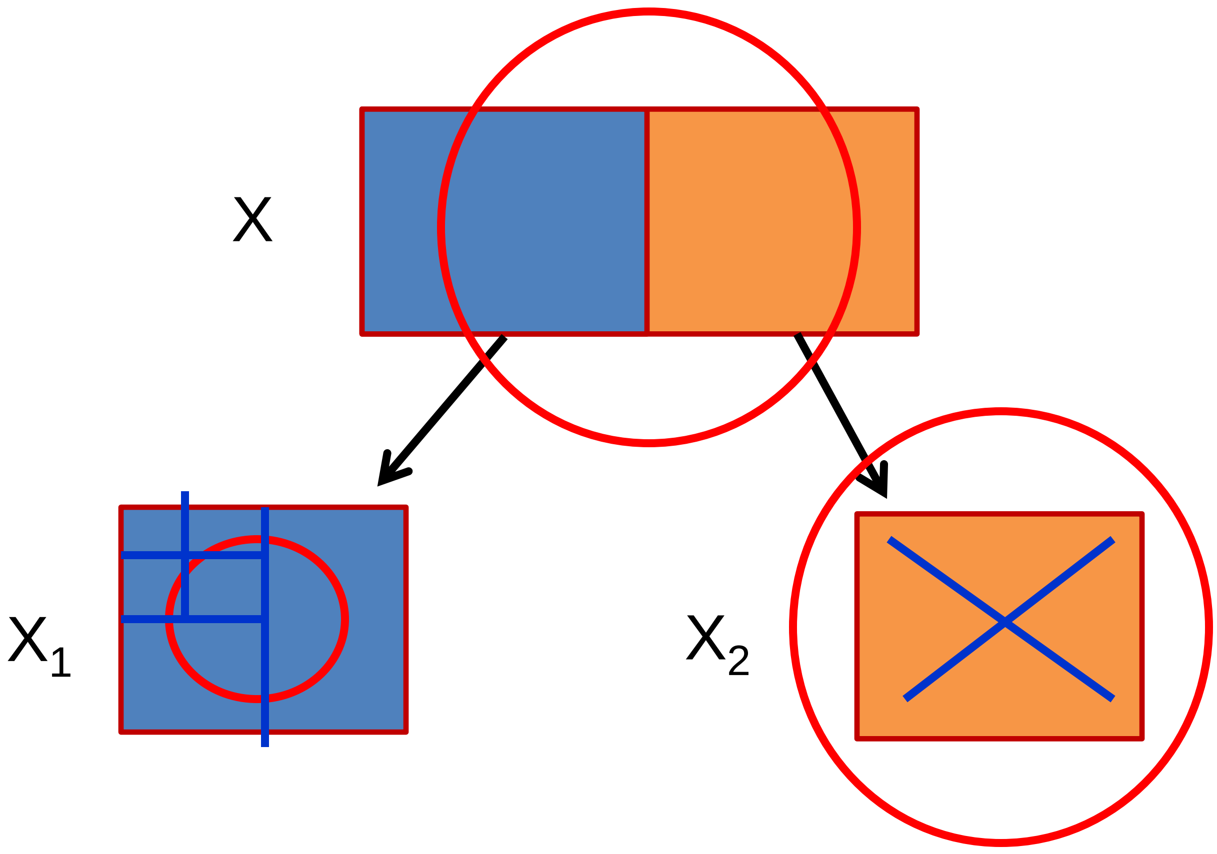
$$\mu_3(x) = f(y) + \langle f_x(y), x - y \rangle + \frac{1}{2}(x - y) f_{xx}(y)(x - y)^T - \frac{M}{6}\|x - y\|^3$$

$$\mu_3(x) = f(y) + \langle f_x(y), x - y \rangle + \frac{\lambda_{\min}}{2}\|x - y\|^2 - \frac{M}{6}\|x - y\|^3$$

СРАВНЕНИЕ МИНОРАНТ



Декомпозиция исходного множестве



УЧЕТ ЦЕЛОЧИСЛЕННОСТИ Mixed integer problem

$$f(x) = x^1 \to \min,$$

$$g^1(x) = (x^1 - 5)^2 + 2(x^2 - 5)^2 + (x^3 - 5)^2 - 18 \le 0,$$

$$g^2(x) = 100 - (x^1 + 7 - 2x^2)^2 -$$

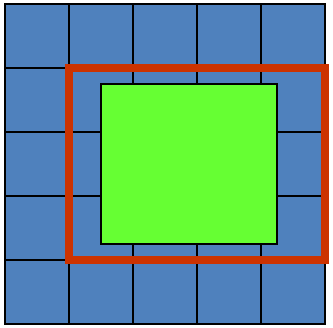
$$4(2x^1 + x^2 - 11)^2 - 5(x^3 - 5)^2 \le 0.$$

$$x \in Z^3$$

$$x_* = (1,4,5)$$

$$f_* = f(x_*) = 1$$

Метод	Число итераций
Без учета целочисленности $\varepsilon = \delta = 0.01$	2671
Липшицева функция	585
Градиент удовлетворяет условию Липшица	121
Градиент + сокращение области поиска	55



Многокритериальная оптимизация

$$F(x) \to \min,$$

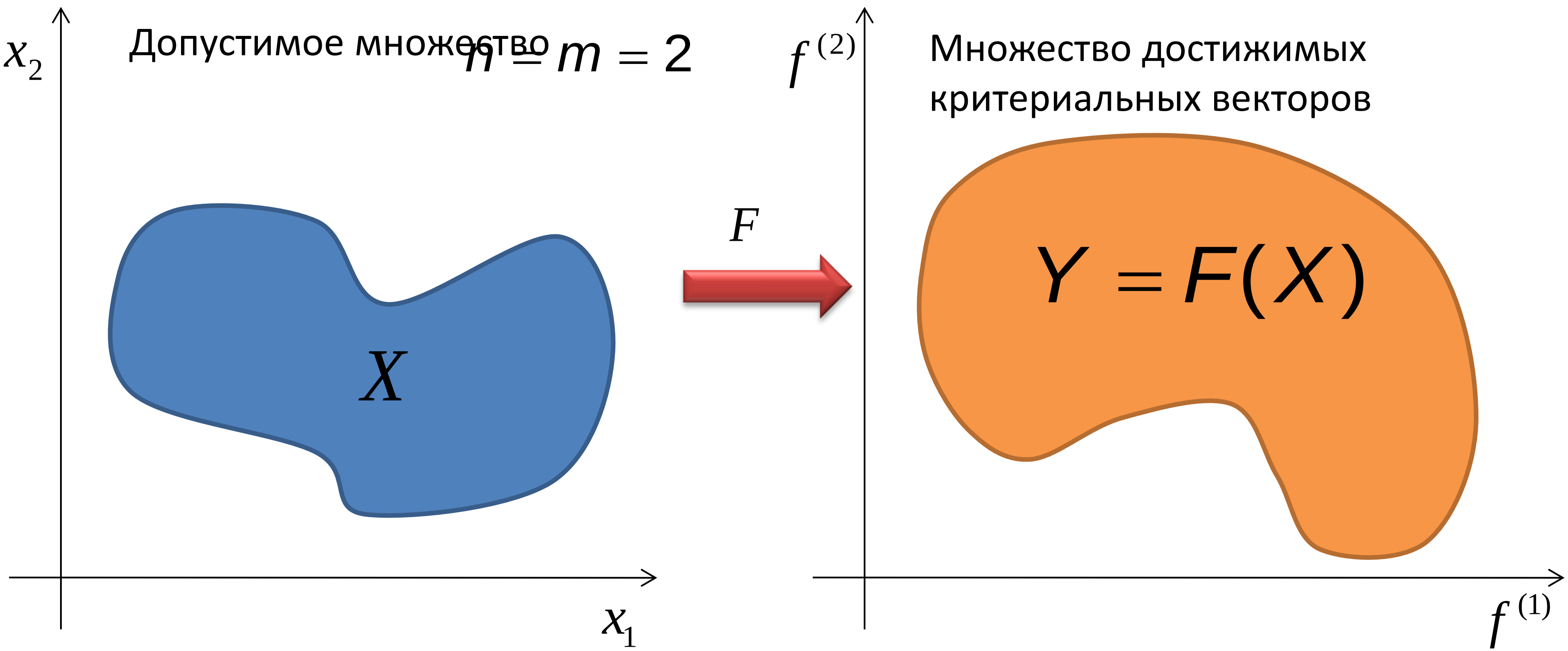
$$x \in X,$$

$$F(x) : R^n \to R^m,$$

$$F(x) = \left[f^{(1)}(x), \dots, f^{(m)}(x) \right]$$

$F(x)$ - непрерывная вектор-функция

Многокритериальная оптимизация



R^n - пространство параметров

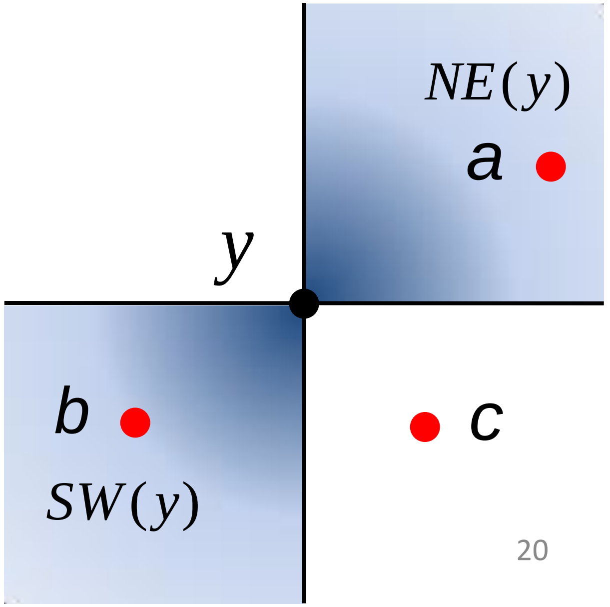
R^m - пространство критериев

Полезные обозначения
случай минимизации

1. $u \le y \Leftrightarrow u_i \le y_i$ *для* $i = 1,\dots,m$

2. $SW(y) = \{u \in R^m : u \le y\}$

3. $NE(y) = \{u \in R^m : u \ge y\}$

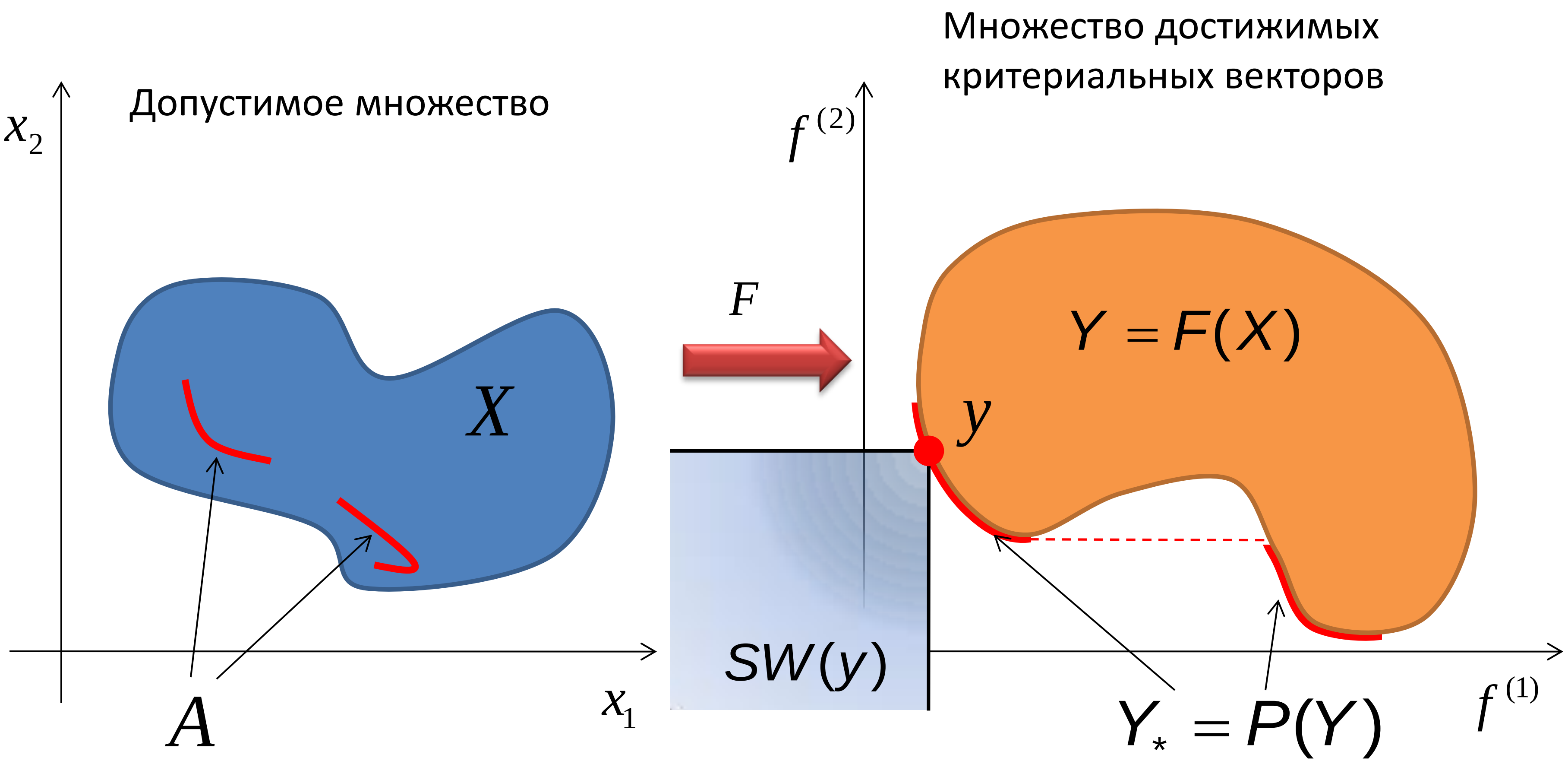


Точка y лучше, чем a,
если $a \ge y$, т.е. все $a^i \ge y^i$

Точка y хуже, чем b,
если $b \le y$, т.е. все $b^i \le y^i$

Точки a и c несравнимые.

Y_* - множество (граница) Парето

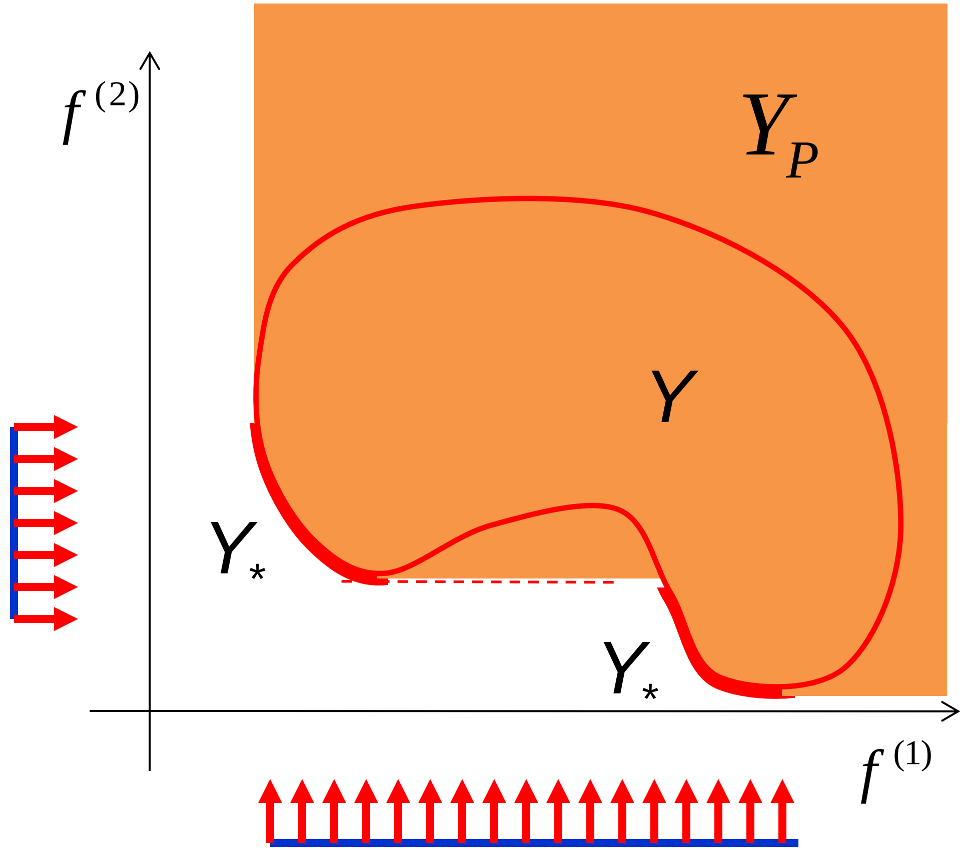


$$P(Y) = \{y \in Y : SW(y) \cap Y = y\} = Y_*$$

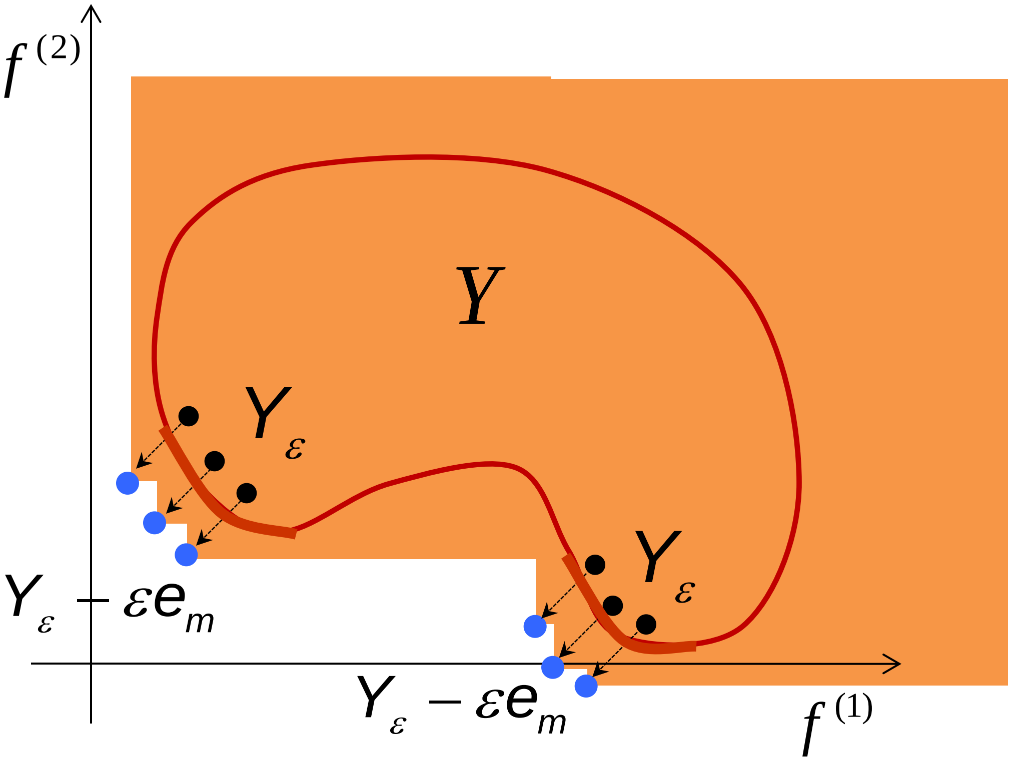
$$Y_* = F(A)$$

$NE(Y_*)$ - оболочка Эджворта-Парето

$$Y_* \subseteq NE(Y_*)$$



Y_ε - множество ε -Парето



Y_ε множество $\varepsilon - Парето$

если

1. $Y_\varepsilon \subseteq Y$

2. $P(Y_\varepsilon) = Y_\varepsilon$

3. $Y_* \subseteq NE(Y_\varepsilon - \varepsilon \cdot e_m)$

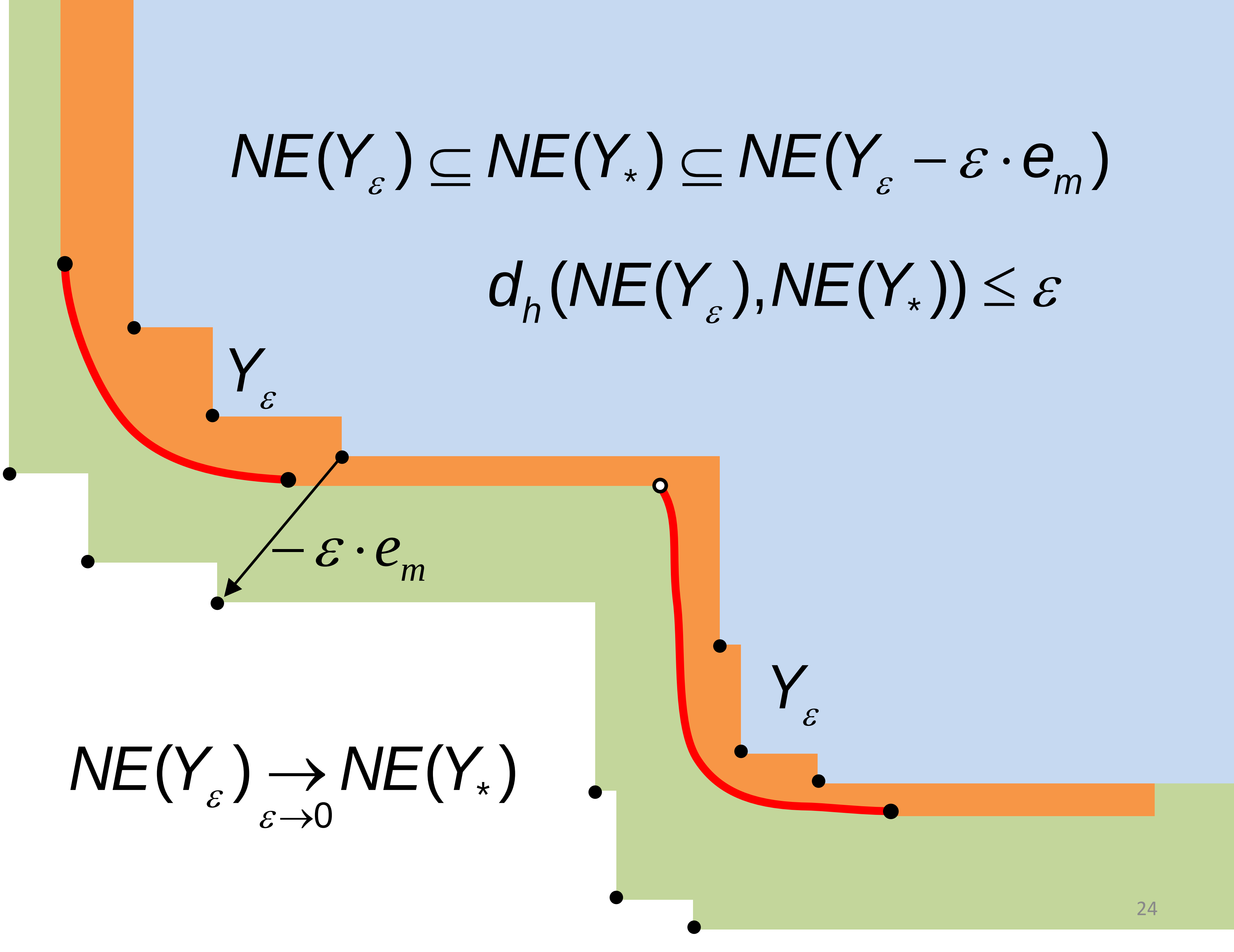
$$|Y_\varepsilon| = 6$$

$$e_m = \underbrace{(1,\ldots,1)}_{m}$$

Ю. Г. Евтушенко, М. А. Потапов.
Методы решения многокритериальных задач.
Доклады Академии наук СССР, Т. 291, N 1,
С. 25-39, 1986

$$NE(Y_\varepsilon) \subseteq NE(Y_*) \subseteq NE(Y_\varepsilon - \varepsilon \cdot e_m)$$

$$d_h(NE(Y_\varepsilon), NE(Y_*)) \leq \varepsilon$$



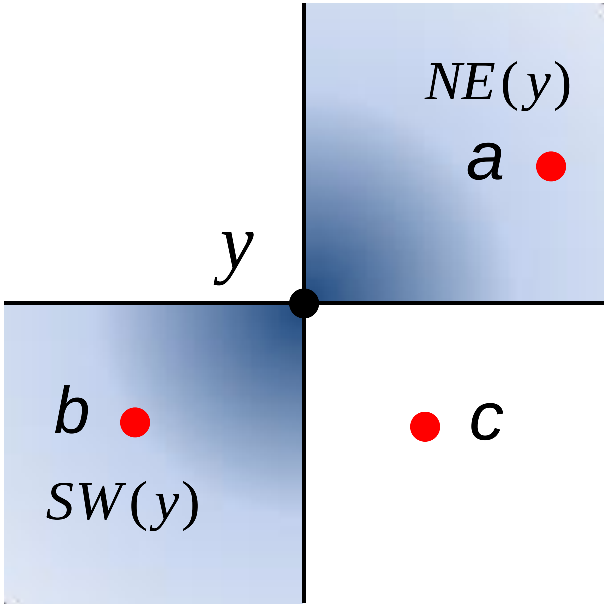
$$NE(Y_\varepsilon) \xrightarrow[\varepsilon \to 0]{} NE(Y_*)$$

Mixing max/min

$\lambda \in R^m,$

$\lambda^i = 1,$ *если ищется* max *по* y_i,

$\lambda^i = -1,$ *если ищется* min *по* y_i



Точка y **лучше**, чем a,
если

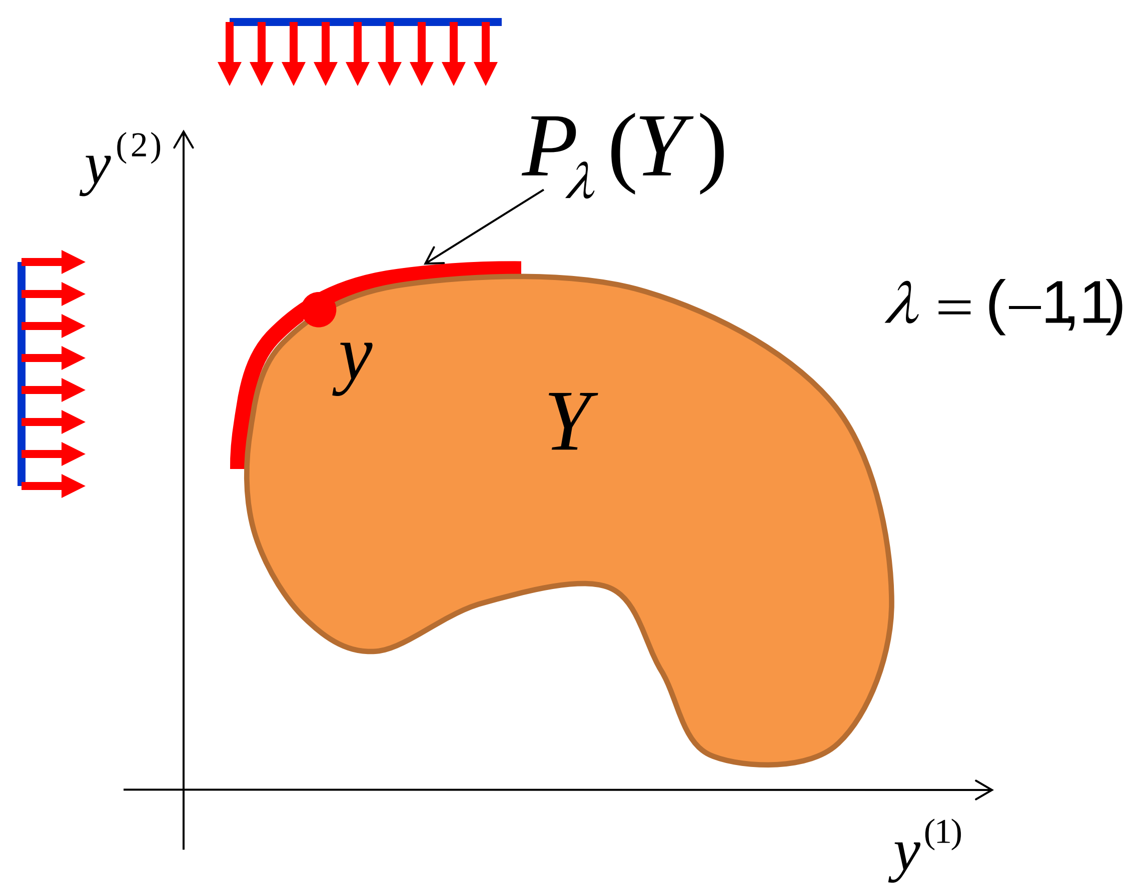
$\lambda^i y^i \geq \lambda^i a^i$ *для всех* $i = 1,\ldots,m$

Точка y **хуже**, чем b,
если

$\lambda^i y^i \leq \lambda^i a^i$ *для всех* $i = 1,\ldots,m$

Точки a и c несравнимые.

Mixing max/min

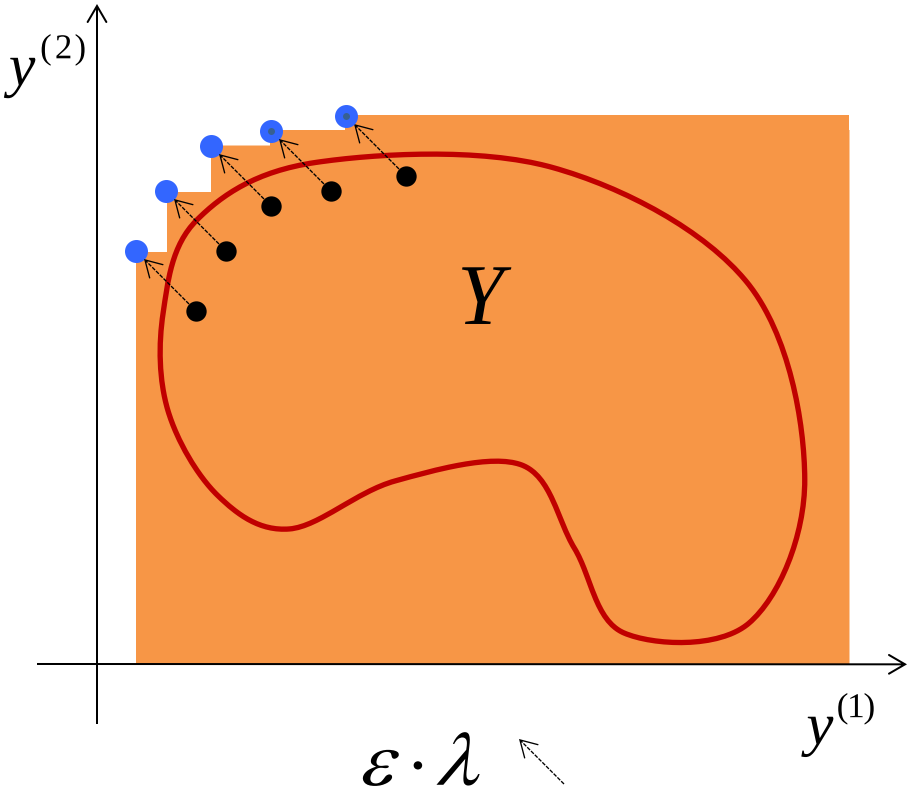


$f^{(1)}(x) \to \min$

$f^{(2)}(x) \to \max$

$$\max_x f(x) = -\min_x(-f(x))$$

ε -Pareto Set



Y_ε is $\varepsilon-$Pareto set if

1. $Y_\varepsilon \subseteq Y$
2. $P_\lambda(Y_\varepsilon) = Y_\varepsilon$
3. $Y \subseteq Y_\varepsilon + \varepsilon \cdot \lambda + R^m_{-\lambda}$

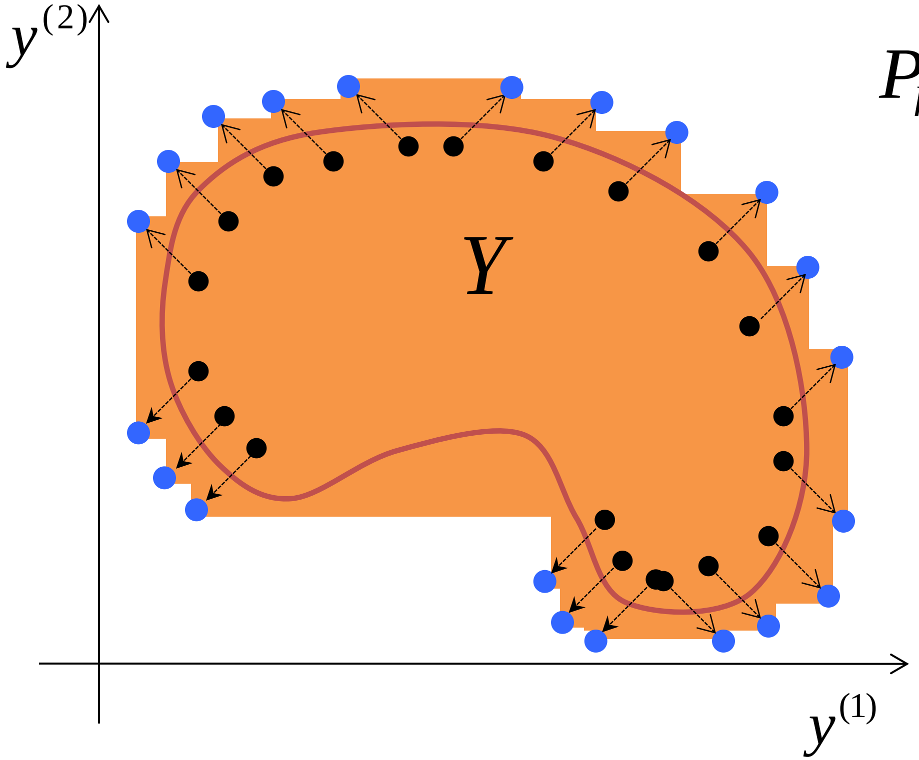
●●● Y_ε

●●● $Y_\varepsilon + \varepsilon \cdot \lambda$

$f^{(1)}(x) \to \min$

$f^{(2)}(x) \to \max$

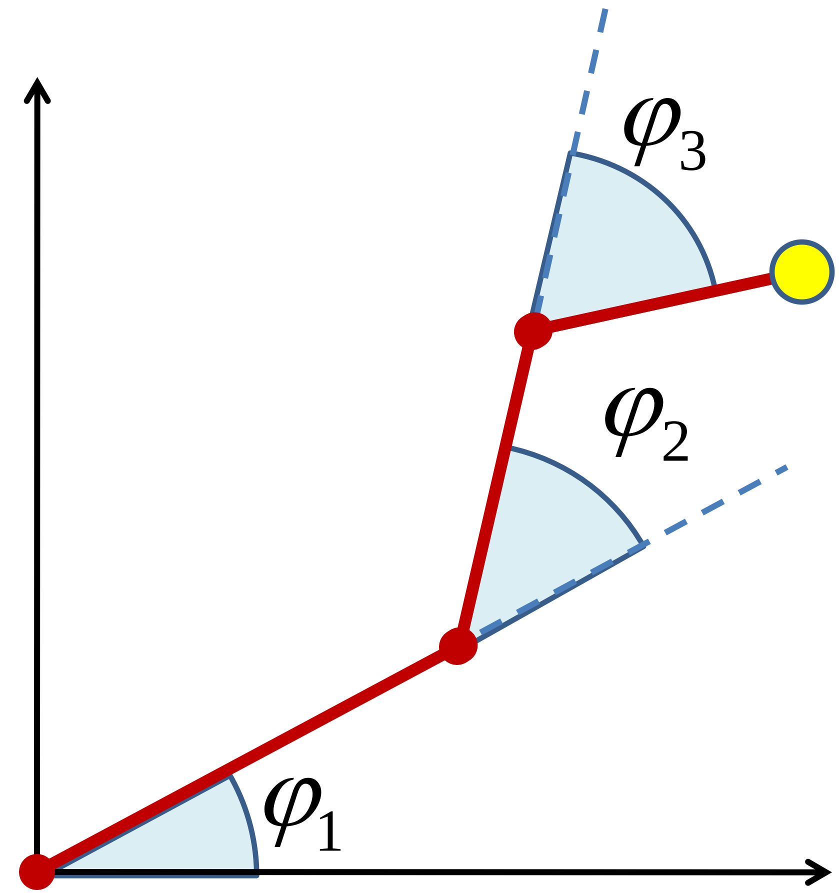
ε-оболочка множества Y



$$P_h^{\varepsilon}(Y) = \bigcap_{\lambda \in \{-1,1\}^m} (Y_{\varepsilon}^{\lambda} + \varepsilon \cdot \lambda + R_{-\lambda}^m)$$

$$Y \subseteq P_h^{\varepsilon}(Y)$$

Практический пример: построение области достижимости многосекционного робота-манипулятора



(Предложен А.П. Карпенко)

Важность задачи

- Построение области достижимости многосекционного робота-манипулятора является важнейшей задачей
- Гарантия точности аппроксимации является очень важной

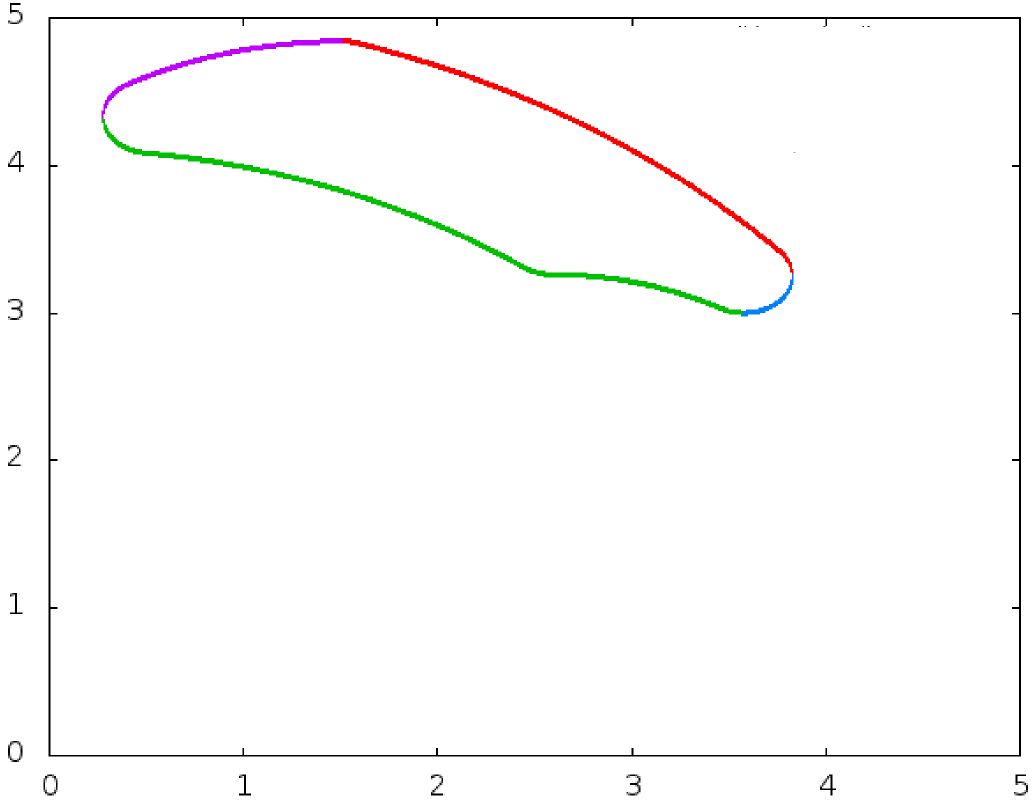
Трехсекционный плоский робот-манипулятор

$$L_1 = 3, L_2 = 2, L_3 = 0.25,$$

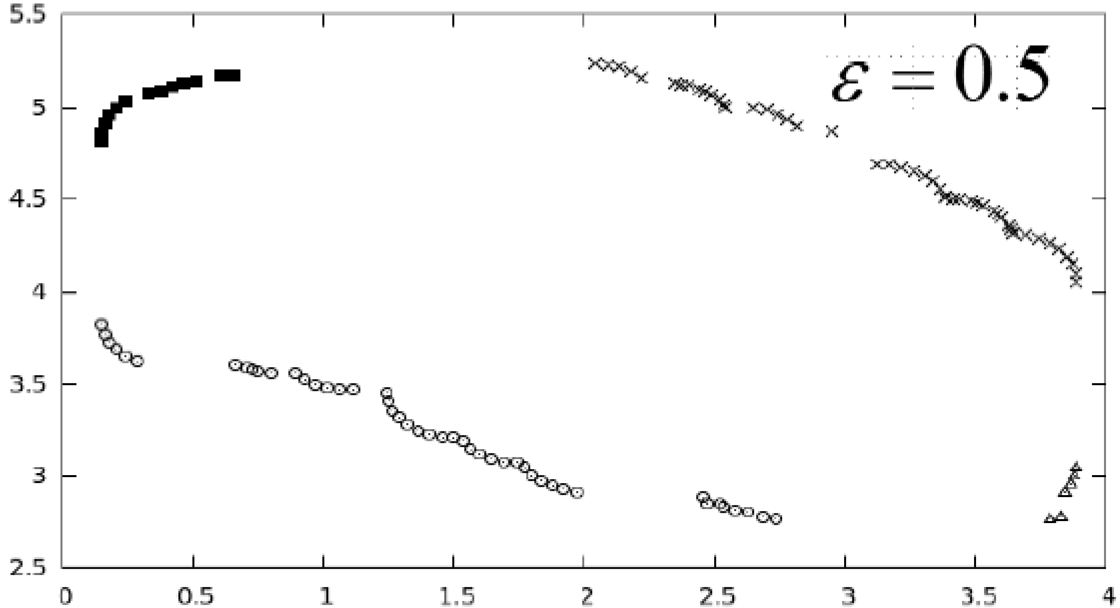
$$\frac{\pi}{6} \le \varphi_1 \le \frac{\pi}{3},$$

$$\frac{\pi}{6} \le \varphi_2 \le \frac{\pi}{3},$$

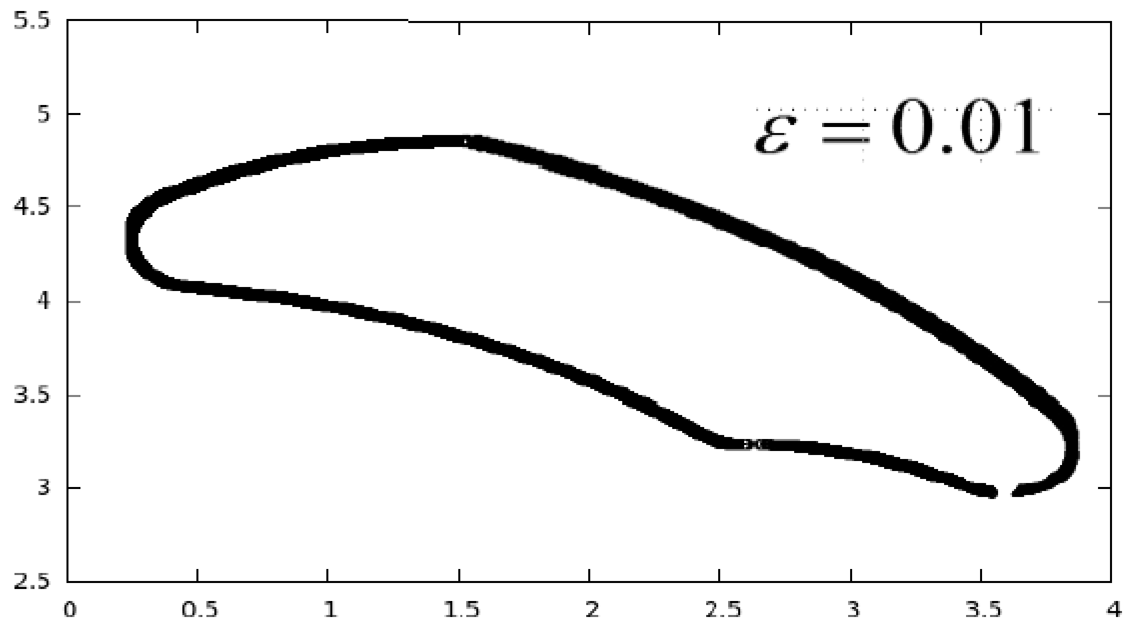
$$-\pi \le \varphi_3 \le \pi$$



0.2 c



47.1 c



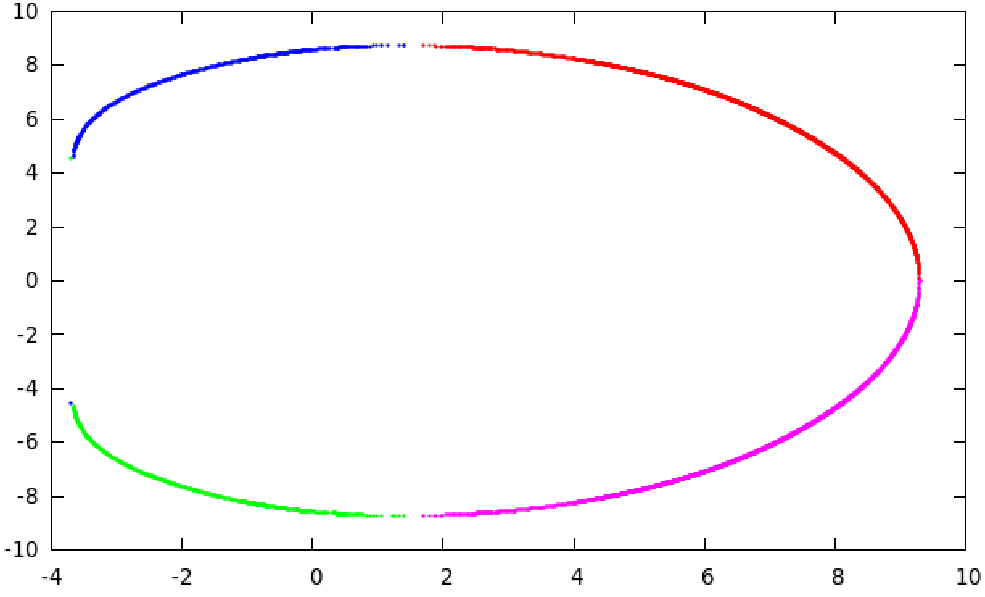
Пятисекционный плоский робот-манипулятор

секции: 1.86, 1.86, 1.86, 1.86, 1.86

углы: $[-\pi/4, \pi/4], [-\pi/4, \pi/4], [-\pi/4, \pi/4], [-\pi/4, \pi/4], [-\pi/4, \pi/4],$

точность аппроксимации 0.1

Расчеты 4-х фрагментов (юго-запад, юго-восток, северо-восток, северо-запад) занял примерно 15 минут с использованием 64х процессоров суперкомпьютера МВС-100К. Для каждого вычисления примерно 1 миллиард вычислений критериев.



Метод неравномерных покрытий

1. Предложен для безусловной оптимизации в 1971
   Ю. Г. Евтушенко, “Численный метод поиска глобального экстремума функций (перебор на неравномерной сетке)”, Ж. вычисл. матем. и матем. физ., 11:6 (1971), 1390–1403

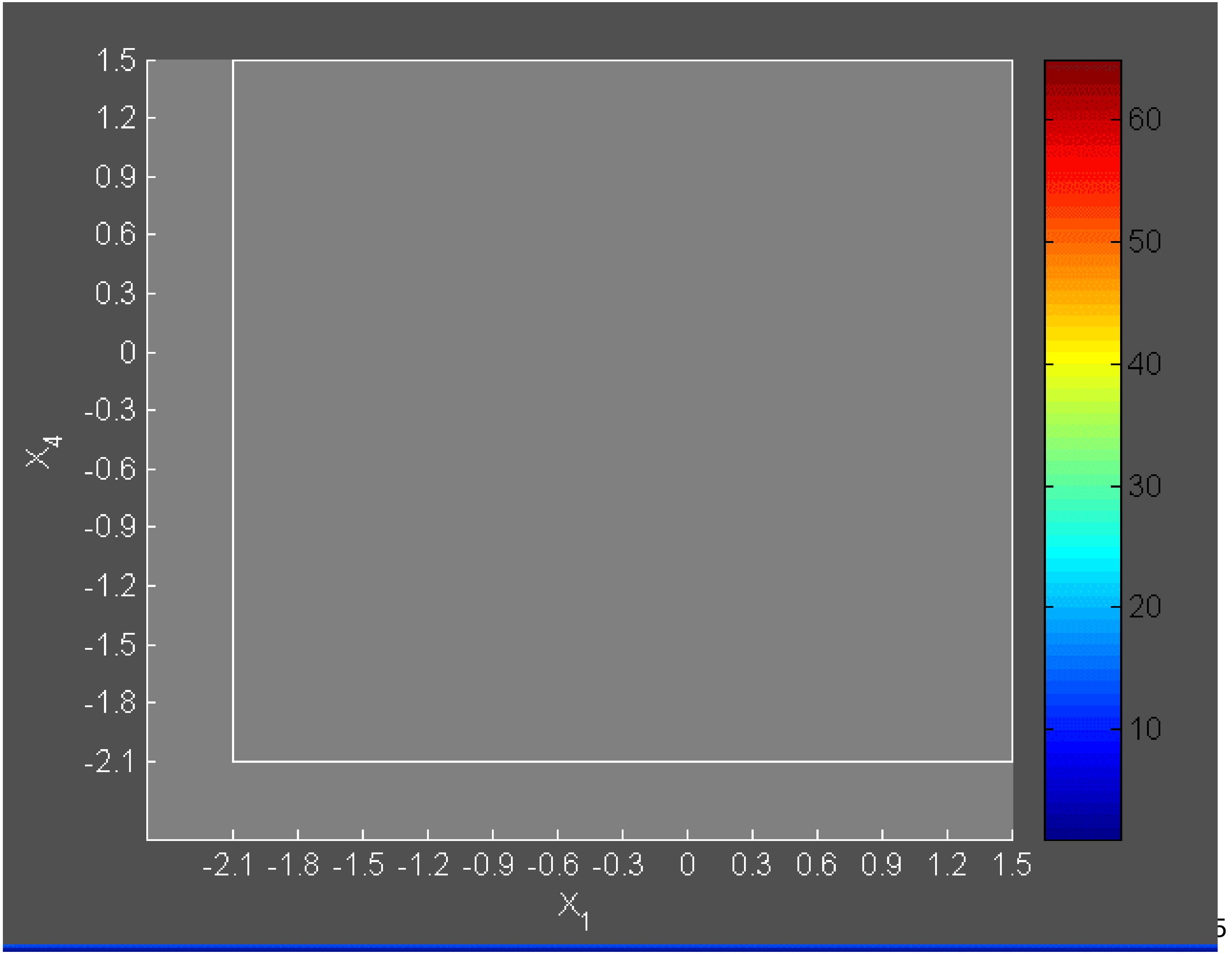
2. Расширен на случай многих критериев в 1986
   Евтушенко Ю. Г., Потапов М. А. «Методы численного решения многокритериальных задач.» ДАН СССР, Т. 291, N 1, 1986, С. 25-39.

3. Расширен на задачи математического программирования в 1992
   Yu. G. Evtushenko, M. A. Potapov, V. V. Korotkikh.
   Numerical methods for global optimization.
   In "Recent advances in global optimization”, Princeton University Press, pp. 274-297. 1992.

4. Первая параллельная реализация 2007
   Ю. Г. Евтушенко, В. У. Малкова, А. А. Станевичюс. Распараллеливание процесса поиска глобального экстремума. Автоматика и телемеханика, № 5. С. 46-58, 2007.

5. Новая техника для задач частично-целочисленного программирования 2011
   Ю. Г. Евтушенко, М. А. Посыпкин. Варианты метода неравномерных покрытий для глобальной оптимизации частично-целочисленных нелинейных задач. Доклады Академии наук. Т: 437. № 2. С. 168–172.

6. Обобщенный вариант для задач многокритериальной оптимизации 2012
   Ю.Г. Евтушенко, М.А. Посыпкин. Метод неравномерных покрытий для решения задач многокритериальной оптимизации с гарантированной точностью // ЖВМиМФ (в печати)


1.5
1.2
0.9
0.6
0.3
0
-0.3
-0.6
-0.9
-1.2
-1.5
-1.8
-2.1
X4
-2.1 -1.8 -1.5 -1.2 -0.9 -0.6 -0.3 0 0.3 0.6 0.9 1.2 1.5
X1
60
50
40
30
20
10

Спасибо за внимание!

Прямая задача ЛП: (P)

$$f_* = \min_{x \in X} c^{\mathrm{T}} x, \; X = \{x \in R^n : Ax = b, \;\; x \ge 0_n\}$$

Двойственная задача ЛП: (D)

$$f_* = \max_{u \in U} b^{\mathrm{T}} u, \quad U = \{u \in R^m : A^{\mathrm{T}} u \le c\}$$

$$m << n$$

Задача нахождения проекции точки $\tilde{x}$ на множество решений X^* прямой задачи ЛП имеет вид:

$$\min_{x \in X^*} \tfrac{1}{2} \| x - \tilde{x} \|^2, \qquad (1)$$

$$X^* = \{x \in R^n : Ax = b, c^{\mathrm{T}} x = f_*, x \geq 0_n\}$$

Функция Лагранжа:

$$L(x, p, \beta) = \tfrac{1}{2} \| x - \tilde{x} \|^2 + p^{\mathrm{T}} (b - Ax) + \beta (c^{\mathrm{T}} x - f_*)$$

Двойственная к (1): $\max\limits_{p \in R^m, \beta \in R^1} \min\limits_{x \in R^n_+} L(x, p, \beta)$

Решение внутренней задачи минимизации $x = (\tilde{x} + A^{\mathrm{T}} p - \beta c)_+$

$$\max_{p \in R^m, \beta \in R^1} \{ b^{\mathrm{T}} p - \tfrac{1}{2} \| (\tilde{x} + A^{\mathrm{T}} p - \beta c)_+ \|^2 - \beta f_* + \tfrac{1}{2} \| \tilde{x} \|^2 \}$$

Теорема 1. Пусть множество решений X* прямой задачи (P) непусто. Тогда существует такое β^*, что при любом $\beta \geq \beta^*$ проекция точки на $\tilde{x}$ множество X* задается формулой

$$\tilde{x}^* = (\tilde{x} + A^{\mathrm{T}} p(\beta) - \beta c)_+,$$

где $p(\beta)$ является решением задачи безусловной минимизации

$$\min_{p \in R^m} [-b^{\mathrm{T}} p + \tfrac{1}{2} \| (\tilde{x} + A^{\mathrm{T}} p - \beta c)_+ \|^2]$$

Теорема 2. Пусть множество решений X^* прямой задачи (P) непусто. Тогда для любых $\tilde{x} \in X^*$ и $\beta > 0$ решение задачи (D) дается формулой $u^* = \dfrac{p(\beta)}{\beta}$, где $p(\beta)$ является решением задачи безусловной минимизации

$$\min_{p \in R^m} [-b^{\mathrm{T}} p + \tfrac{1}{2} \| (\tilde{x} + A^{\mathrm{T}} p - \beta c)_+ \|^2].$$

$$p^{k+1} = \arg\min_{p \in R^m} \{-b^{\mathrm{T}}p + \tfrac{1}{2} \| (x^k + A^{\mathrm{T}}p - \beta c)_+ \|^2 \}$$

$$x^{k+1} = (x^k + A^{\mathrm{T}}p^{k+1} - \beta c)_+ \qquad (1)$$

Теорема 3. Пусть множество решений X^* прямой задачи (P) непусто. Для любого $\beta > 0$ и при любом начальном x^0 итерационный процесс (1) сходится к x^* ™ X^* за конечное число итераций K. Решение двойственной задачи (D) дается формулой

$$u^* = \frac{p^{K+1}}{\beta}.$$

P-IV, 2,6 GHz, 1Gb Вычислительный эксперимент

m × n × d	T (Sec.)	Iterat.	$\|\|Ax-b\|\|$	$\|\|(A^T u-c)_+\|\|$	$\|c^T x-b^T u\|$
$100 \times 10^6 \times 0.01$	29.3	17	1.7×10^{-11}	2.0×10^{-13}	9.7×10^{-11}
$300 \times 10^6 \times 0.01$	42.0	13	1.0×10^{-10}	7.0×10^{-13}	2.6×10^{-10}
$600 \times 10^6 \times 0.01$	68.4	12	3.1×10^{-10}	1.5×10^{-12}	2.8×10^{-10}
$1000 \times 10^6 \times 0.01$	95.8	10	9.4×10^{-10}	3.5×10^{-12}	6.9×10^{-10}
$500 \times 10^4 \times 1$	29.3	8	2.9×10^{-8}	3.4×10^{-11}	1.1×10^{-8}
$1000 \times 10^4 \times 1$	117.2	7	1.3×10^{-7}	1.0×10^{-10}	2.9×10^{-7}
$3000 \times 10^4 \times 0.01$	81.5	7	2.0×10^{-9}	9.1×10^{-12}	3.7×10^{-9}
$4000 \times 10^4 \times 0.01$	196.2	8	2.9×10^{-9}	1.2×10^{-11}	2.6×10^{-8}
$1000\times(3\cdot10^6) \times 0.01$	309.1	11	1.2×10^{-9}	4.1×10^{-12}	4.9×10^{-9}
$1000 \times(5\cdot10^6) \times 0.01$	412.8	8	7.3×10^{-9}	7.4×10^{-12}	7.3×10^{-8}
$1000 \times(5\cdot10^7) \times 0.01$	4392.5	6	7.6×10^{-9}	2.1×10^{-12}	1.1×10^{-7}

Компьютер: *Celeron* 2.02 *GHz* , 1.0 *GB* , *Win XP*

№	Размер Задачи $m \times n \times d$	Метод	Время (сек.)	Точности			
				Primal Infeas.	Dual Infeas.	Duality Gap	
1	$500 \times 10000 \times 1$	EGM (MATLAB)	55.0	1.5×10^{-8}	1.8×10^{-12}	1.2×10^{-7}	2
		BPMPD (Interior point)	37.4	2.3×10^{-10}	1.8×10^{-11}	1.1×10^{-10}	1
		MOSEK (Interior point)	87.2	9.7×10^{-8}	3.8×10^{-9}	1.6×10^{-6}	5
		CPLEX (Interior point)	80.3	1.8×10^{-8}	1.1×10^{-7}	0.0	4
		CPLEX (Simplex)	61.8	8.6×10^{-4}	1.9×10^{-10}	7.2×10^{-3}	3
2	$3000 \times 10000 \times 0.01$	EGM (MATLAB)	155.4	6.1×10^{-10}	3.4×10^{-13}	3.6×10^{-8}	3
		BPMPD (Interior point)	223.5	4.6×10^{-9}	2.9×10^{-10}	3.9×10^{-9}	4
		MOSEK (Interior point)	42.6	3.1×10^{-8}	1.2×10^{-8}	3.7×10^{-8}	1
		CPLEX (Interior point)	69.9	1.1×10^{-6}	1.3×10^{-7}	0.0	2
		CPLEX (Simplex)	1764.9	3.0×10^{-3}	8.1×10^{-9}	9.3×10^{-2}	5
3	$1000 \times (3 \cdot 10^{6}) \times 0.01$	EGM (MATLAB)	536.8	6.9×10^{-8}	1.4×10^{-13}	8.4×10^{-7}	3
		BPMPD (Interior point)	-	*Не решил*			-
		MOSEK (Interior point)	-	*Не решил*			-
		CPLEX (Interior point)	340.6	2.4×10^{-2}	1.3×10^{-6}	0.0	1
		CPLEX (Simplex)	370.4	$1.8 \times 10^{+2}$	$3.7 \times 10^{+2}$	1.2×10^{-9}	2
4	$1000 \times (5 \cdot 10^{6}) \times 0.01$	EGM (MATLAB)	1007.5	3.9×10^{-8}	1.4×10^{-13}	6.1×10^{-7}	1

$$\min_{p\in R^m} S(p) = \min_{p\in R^m} \{-b^{\mathrm{T}}p + \tfrac{1}{2}\,\|\,(x^k + A^{\mathrm{T}}p - \beta c)_+\,\|^2\}$$

$S(p)$ выпуклая один раз дифференцируемая кусочно квадратичная функция

Градиент:

$$S_p(p) = -b + A(x^k + A^{\mathrm{T}}p - \beta c)_+$$

Обобщенная матрица Гессе:

$$\partial^2 S(p) = AD^{\#}(x^k + A^{\mathrm{T}}p - \beta c)A^{\mathrm{T}}$$

$D^{\#}(t)$ есть $n\times n$ диагональная матрица с i-м элементом

$$t_i \begin{cases} = 1, & \text{если } t > 0 \\ \in [0,1], & \text{если } t = 0 \\ = 0, & \text{если } t < 0 \end{cases}$$

МЕТОД НЬЮТОНА

1) $$p^{s+1} = p^s - \lambda_s(\partial^2 S(p^s) + \delta I_m)^{-1} S_p(p^s)$$

С регулировкой шага Armijo $\lambda_s = \max\{1, \frac{1}{2}, \frac{1}{4}, \ldots\}$:

$$S(p^s) - S(p^s + \lambda_s d^s) \geq -\frac{\lambda_s}{4} S_p^{\mathrm{T}}(p^s) d^s,$$

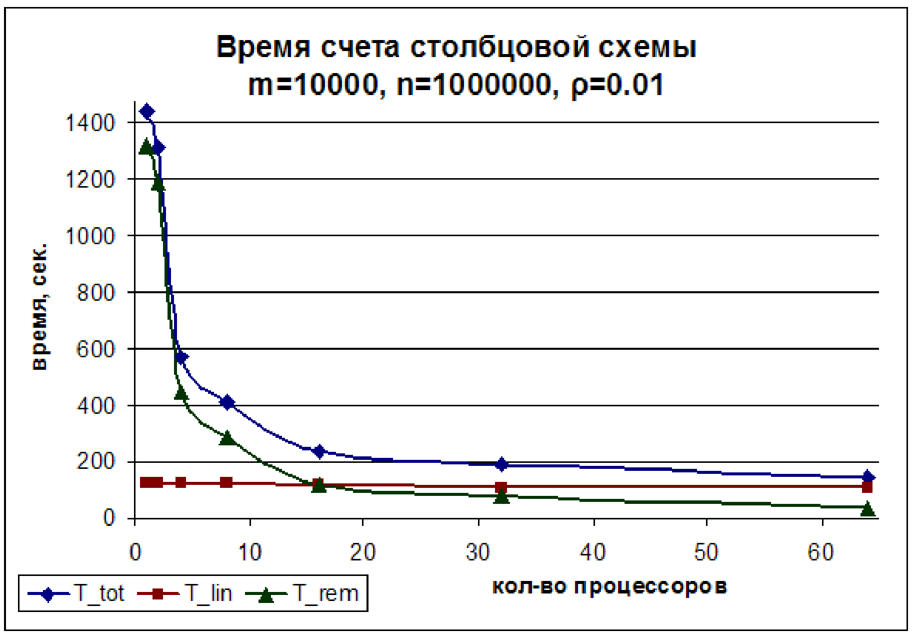
где d^s квазиньютоновское направление:

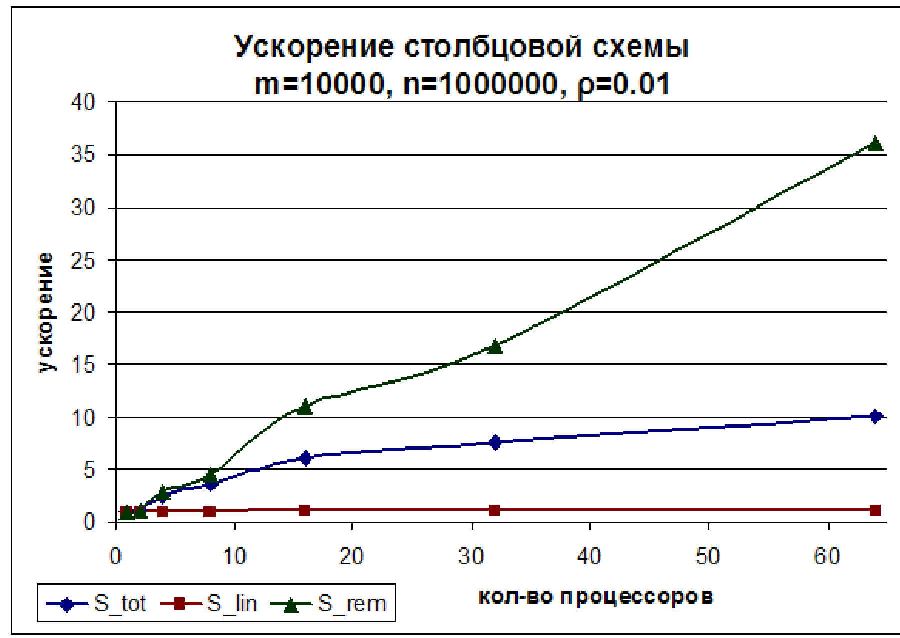
$$d^s = -(\partial^2 S(p^s) + \delta I_m)^{-1} S_p(p^s).$$

2) Стоп, если $\|p^{s+1} - p^s\| < tol$, иначе положить $s = s+1$ и перейти к 1)

Столбцовая схема разбиения данных

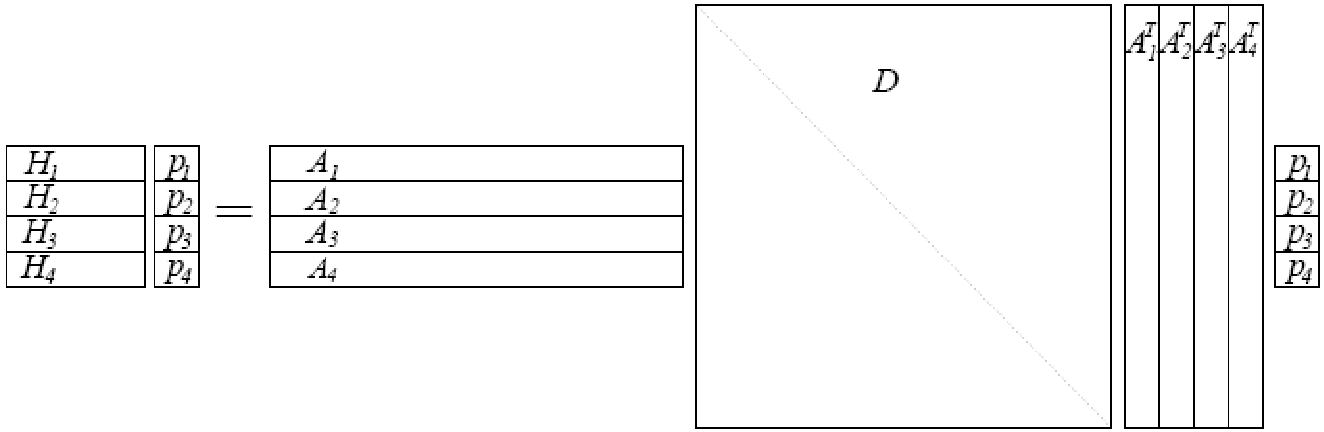
$$\left(H_1 + H_2 + H_3 + H_4 \right) p = \begin{bmatrix} A_1 & A_2 & A_3 & A_4 \end{bmatrix} \begin{bmatrix} D_1 & & & \\ & D_2 & & \\ & & D_3 & \\ & & & D_4 \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ A_4 \end{bmatrix} p$$

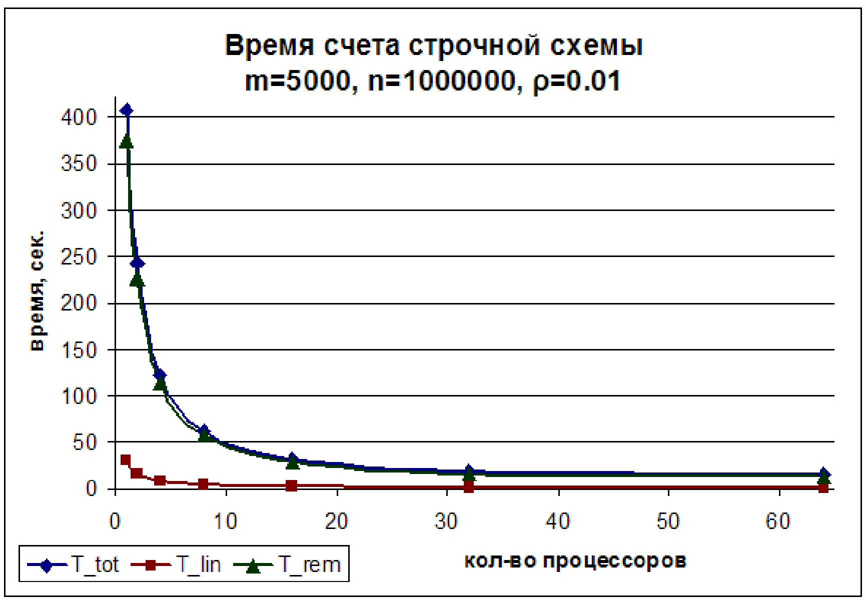


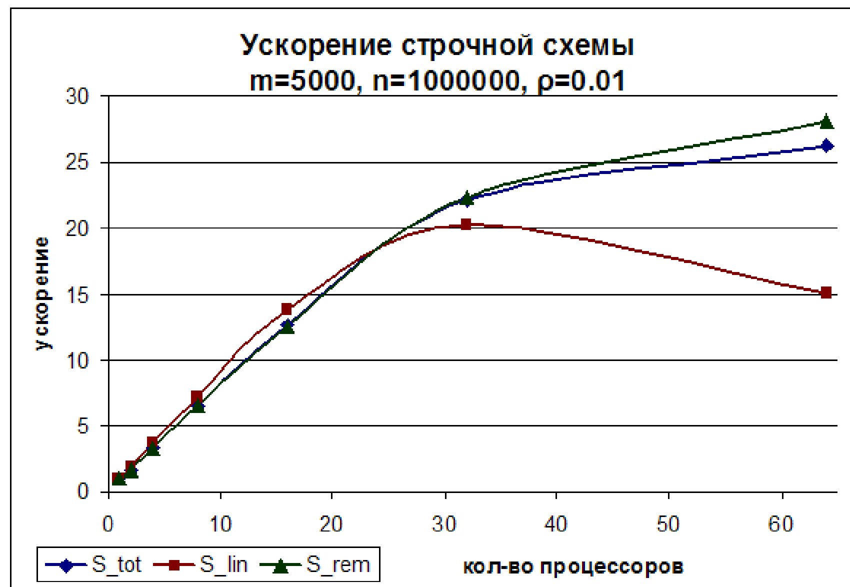


$$\|Ax - b\|_\infty \approx 10^{-8}, \quad \|(A^T u - c)_+\|_\infty \approx 10^{-11}, \quad |c^T x - b^T u| \approx 10^{-8}$$

Строчная схема разбиения данных

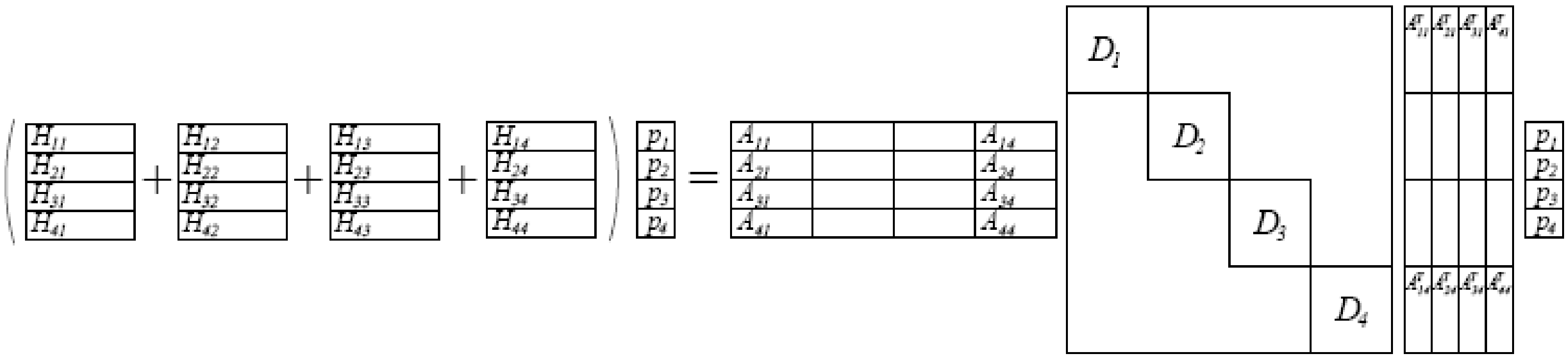


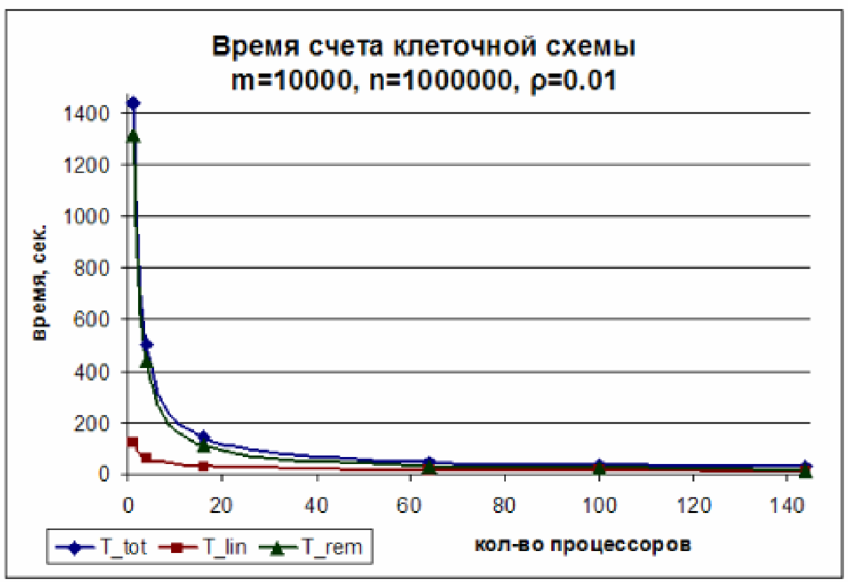


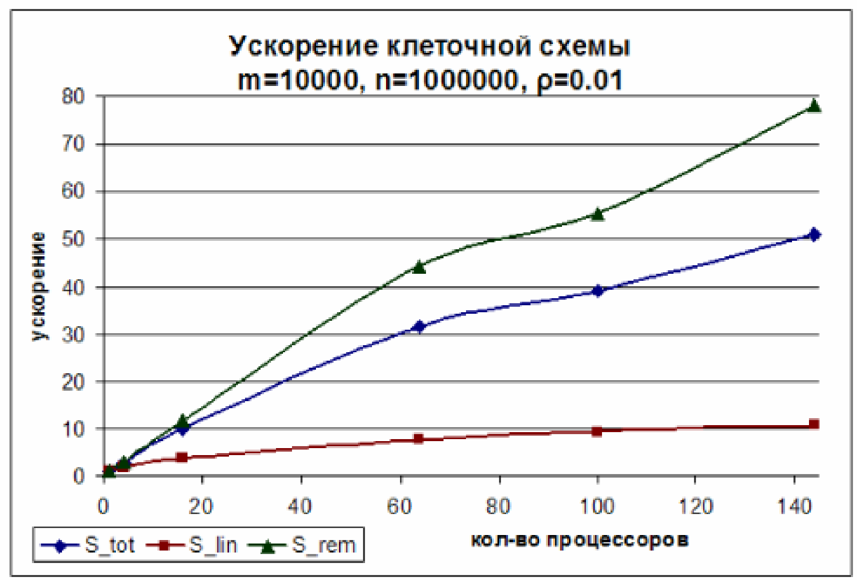


$$\|Ax-b\|_\infty \approx 10^{-8}, \quad \|(A^T u - c)_+\|_\infty \approx 10^{-12}, \quad |c^T x - b^T u| \approx 10^{-8}$$

Клеточная схема разбиения данных







$$\|Ax-b\|_{\infty} \approx 10^{-8}, \quad \|(A^{T}u-c)_{+}\|_{\infty} \approx 10^{-11}, \quad |c^{T}x - b^{T}u| \approx 10^{-8}$$

Вычислительный комплекс МВС-6000IM

Максимальное ускорение 50 на 144 процессорах, клеточная схема, m=10 000, n=1 000 000, t=28 сек.

Максимальное число ограничений m=200 000, n=2 000 000, t=40 мин., безматричная схема на 80 процессорах.

Максимальное число переменных n=60 000 000, m=5000,   t=232 сек., столбцовая схема на 120 процессорах

БЛАГОДАРЮ ЗА ВНИМАНИЕ!