ЧИСЛЕННЫЕ ЭКСПЕРИМЕНТЫ ПО СТРОГОЙ h-ОТДЕЛИМОСТИ*

Е. К. Чернэуцану

katerinache@yandex.ru

17 декабря 2011 г.

В задачу о строгом h-отделении двух множеств вводится параметр. Изучается влияние этого параметра.

1°. Пусть в пространстве $\mathbb{R}^n$ заданы два конечных множества

$$A = \{a_i\}_{i=1}^m \quad \text{и} \quad B = \{b_j\}_{j=1}^k.$$

Задачу строгого отделения выпуклой оболочки множества A от множества B с помощью h гиперплоскостей H_s, определяемых уравнениями

$$\langle w^s, x\rangle = \gamma_s, \quad s \in 1:h,$$

где $w^s \neq \mathbb{O}$ при всех s, можно формализовать так:

$$\begin{aligned} F(G) := \frac{1}{m}\sum_{i=1}^m \max_{s\in 1:h}\big[\langle w^s, a_i\rangle - \gamma_s + c\big]_+ &+ \\ + \frac{1}{k}\sum_{j=1}^k \min_{s\in 1:h}\big[-\langle w^s, b_j\rangle + \gamma_s + c\big]_+ &\to \min_G, \end{aligned} \tag{1}$$

Здесь G — матрица размера $h \times (n+1)$ со строками

$$g^s = (w^s, \gamma_s), \quad s \in 1:h;$$

$c > 0$ — параметр. В работах [1, 2] задача (1) исследовалась только при $c = 1$. Матрицу G указанных размеров будем называть *подходящей*, если у неё все w^s отличны от нулевого вектора.

Справедливо следующее утверждение.

---

*Семинар по дискретному гармоническому анализу и геометрическому моделированию «DHA & CAGD»: `http://www.dha.spb.ru/`

ТЕОРЕМА 1. *Выпуклая оболочка множества A и множество B строго h-отделимы тогда и только тогда, когда существует подходящая матрица G_*, такая, что $F(G_*) = 0$.*

Доказательство. Достаточность практически очевидна. Действительно, пусть $F(G_*) = 0$ для некоторой подходящей матрицы G_* со строками (w_*^s, γ_s^*). Тогда

$$\begin{aligned} \langle w_*^s, a_i\rangle - \gamma_s^* \leqslant -c < 0 \quad &\text{при всех} \quad i \in 1:m \quad \text{и всех} \quad s \in 1:h, \\ \langle w_*^s, b_j\rangle - \gamma_s^* \geqslant c > 0 \quad &\text{при каждом} \quad j \in 1:k \quad \text{и некотором} \quad s \in 1:h. \end{aligned} \tag{2}$$

Эти соотношения и определяют строгую h-отделимость выпуклой оболочки множества A от множества B.

Необходимость доказывается по той же схеме, что и в работах [1, 2]. □

2°. В задаче (1) функция $F(G)$ минимизируется по всем матрицам G размера $h \times (n+1)$. Как показано в [3], задача (1) всегда имеет решение, однако оптимальная матрица, вообще говоря, может не быть подходящей. В этой связи представляет интерес такое утверждение (см. [1, 2]).

ТЕОРЕМА 2. *Пусть $F(G_*) = 0$. В этом случае*

1) *у G_* не все w_*^s равны нулевому вектору;*

2) *если $w_*^s = \mathbb{O}$ на множестве $S \subset 1:h$, то выпуклая оболочка множества A и множество B строго $\big(h - |S|\big)$-отделимы.*

3°. По определению $F(G) \geqslant 0$ на всех матрицах G. Особенность задачи (1) заключается в том, что может существовать подходящая матрица G_*, которая строго отделяет $co(A)$ от B в смысле (2), но при этом $F(G_*) > 0$.

ПРИМЕР. Рассмотрим на плоскости $\mathbb{R}^2$ множества A и B, содержащие по одной точке $a = (0,0)$ и $b = (0,2)$ соответственно. Речь пойдёт, естественно, об 1-отделении. Положим $w^1 = (0,1)$, а в качестве γ_1 возьмём любое число из интервала $(0,2)$. В этом случае прямая $x_2 = \gamma_1$ строго отделяет точку a от b (см. рис. 1).

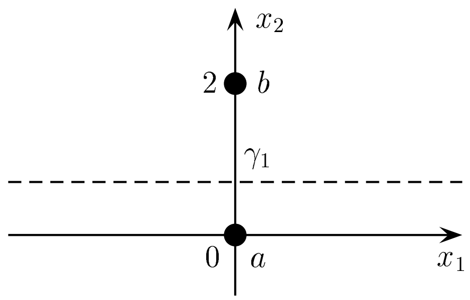

Рис. 1

Вместе с тем,

$$F(G) = [-\gamma_1 + c]_+ + [-2 + \gamma_1 + c]_+.$$

На рис. 2 представлен график $F(G)$ как функции от γ_1. Видим, что $F(G)$ до-

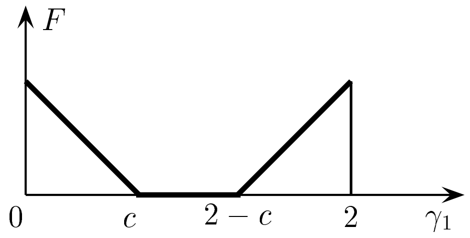

Рис. 2

стигает своего минимального значения, равного нулю, на парах (w^1, γ_1) при $\gamma_1 \in [c, 2-c]$. При $\gamma_1 \in (0, c) \cup (2-c, 2)$ пара (w^1, γ_1) по-прежнему строго отделяют точку a от b, но в этом случае $F(G) > 0$. Наконец, только при $c = 1$ условие $F(G) = 0$ выполняется на единственной паре вида (w^1, γ_1), у которой $\gamma_1 = 1$.

4°. Как показано в [3], задача (1) сводится к конечному числу задач линейного программирования. Количество таких задач равно k^h. Оно может быть достаточно большим. Поэтому представляют интерес приближённые методы и прежде всего градиентный метод, предложенный в [1, 4]. Напомним его описание.

Введём обозначения

$$f(v, u) = \langle v, u \rangle + c,$$

$$\hat{a}_i = \begin{pmatrix} a_i \\ -1 \end{pmatrix}, \quad \check{b}_j = \begin{pmatrix} -b_j \\ 1 \end{pmatrix},$$

$$\varphi_i(G) = \max_{s \in 1:h} \big[f(g^s, \hat{a}_i) \big]_+, \quad \psi_j(G) = \min_{s \in 1:h} \big[f(g^s, \check{b}_j) \big]_+.$$

Тогда задача (1) перепишется в более компактном виде

$$F(G) := \frac{1}{m}\sum_{i=1}^{m}\varphi_i(G) + \frac{1}{k}\sum_{j=1}^{k}\psi_j(G) \to \min_G. \tag{3}$$

Функции от матриц $\varphi_i(G)$ и $\psi_j(G)$ дифференцируемы по направлениям (в качестве которых выступают также матрицы). Чтобы записать формулы для производных по направлению $\varphi_i'(G,V)$ и $\psi_j'(G,V)$, введём дополнительные обозначения

$$\hat{\varphi}_i(G) = \max_{s\in 1:h} f(g^s, \hat{a}_i), \quad \check{\psi}_j(G) = \min_{s\in 1:h} f(g^s, \check{b}_j),$$

так что

$$\varphi_i(G) = \big[\hat{\varphi}_i(G)\big]_+, \quad \psi_j(G) = \big[\check{\psi}_j(G)\big]_+.$$

Положим

$$\hat{R}_i(G) = \big\{s \in 1:h \mid f(g^s, \hat{a}_i) = \hat{\varphi}_i(G)\big\},$$
$$\check{R}_j(G) = \big\{s \in 1:h \mid f(g^s, \check{b}_j) = \check{\psi}_j(G)\big\}.$$

Тогда (см. [1, 4])

$$\varphi_i'(G,V) = \begin{cases} \max_{s\in \hat{R}_i(G)} \langle v^s, \hat{a}_i\rangle, & \text{если } \hat{\varphi}_i(G) > 0; \\ \max_{s\in \hat{R}_i(G)} \big[\langle v^s, \hat{a}_i\rangle\big]_+, & \text{если } \hat{\varphi}_i(G) = 0; \\ 0, & \text{если } \hat{\varphi}_i(G) < 0; \end{cases}$$

и

$$\psi_j'(G,V) = \begin{cases} \min_{s\in \check{R}_j(G)} \langle v^s, \check{b}_j\rangle, & \text{если } \check{\psi}_j(G) > 0; \\ \min_{s\in \check{R}_j(G)} \big[\langle v^s, \check{b}_j\rangle\big]_+, & \text{если } \check{\psi}_j(G) = 0; \\ 0, & \text{если } \check{\psi}_j(G) < 0. \end{cases}$$

Согласно (3)

$$F'(G,V) = \frac{1}{m}\sum_{i=1}^{m}\varphi_i'(G,V) + \frac{1}{k}\sum_{j=1}^{k}\psi_j'(G,V).$$

Отметим, что $F'(G,tV) = tF'(G,V)$ при $t > 0$.

Переходим к градиентному методу. Возьмём начальное приближение G_0. Пусть уже имеется k-е приближение G_k. Решаем вспомогательную задачу

$$F'(G_k, V) \to \min_{V\in\Omega}, \tag{4}$$

где Ω — множество матриц $V = \big\{v^s(\alpha)\big\}$, элементы которых удовлетворяют неравенствам

$$\big|v^s(\alpha)\big| \leqslant K, \quad s \in 1:h,\ \alpha \in 1:n+1.$$

Задача (4) сводится к небольшому числу задач линейного программирования. Она имеет решение (в силу ограниченности множества Ω). Обозначим его V_k. Если

$$F'(G_k, V_k) \geqslant 0, \tag{5}$$

то G_k — точка локального минимума функции $F(G)$. В противном случае матрица V_k является направлением убывания функции $F(G)$ из точки G_k. Находим точку минимума функции $F(G_k + tV_k)$ при $t > 0$. Обозначим её t_k. Полагаем $G_{k+1} = G_k + t_k V_k$, после чего вычисления повторяются. Описание принципиальной схемы градиентного метода для решения задачи (1) завершено.

Отметим, что из процесса имеются два выхода. Вычисления прекращаются, когда выполняется неравенство (5) или когда $F(G_k) = 0$. Для оценки результатов вычислений можно использовать соотношения (2).

С помощью градиентного метода изучим влияние параметра c на качество строгого h-отделения.

5°. Рассмотрим пример строгого 3-отделения при разных значениях параметра c. Особое внимание будем уделять организации вычислений.

Пусть на плоскости заданы множества A и B, состоящие соответственно из точек

$$a_1 = (-3, 0),\ a_2 = (1, 2),\ a_3 = (2, -1),\ a_4 = (0, 1);$$
$$b_1 = (-3, 3),\ b_2 = (4, 1),\ b_3 = (-1, -3).$$

Очевидно, что $\mathrm{conv}(A) \cap B = \varnothing$ (см. рис. 3).

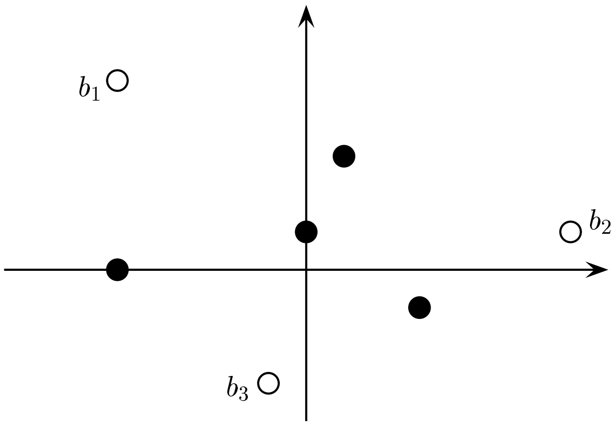

Рис. 3

Сначала будем решать задачу строгого 3-отделения при $c = 0$. Возьмём начальное приближение (см. рис. 4)

$$G_0 = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 4 \\ 0 & -1 & 4 \end{pmatrix}.$$

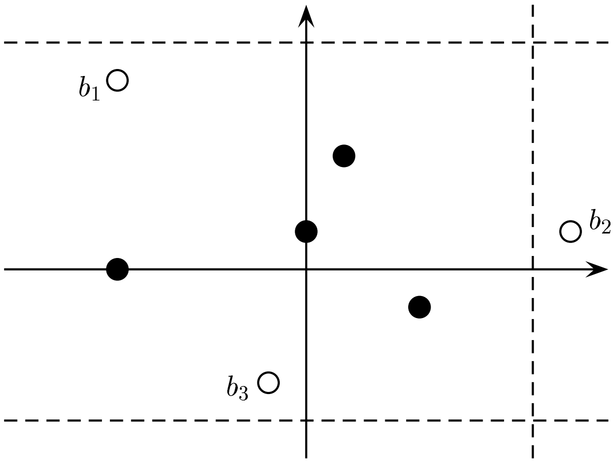

Рис. 4

Заполним две таблицы.

Таблица 1

i	1	2	3	4
$f(g_0^1, \hat{a}_i)$	-6	-2	-1	-3
$f(g_0^2, \hat{a}_i)$	-4	-2	-5	-3
$f(g_0^3, \hat{a}_i)$	-4	-6	-3	-5
$\hat{\varphi}_i(G_0)$	-4	-2	-1	-3
$\hat{R}_i(G_0)$	$\{2,3\}$	$\{1,2\}$	$\{1\}$	$\{1,2\}$
$\varphi_i(G_0)$	0	0	0	0

Таблица 2

j	1	2	3
$f(g_0^1, \check{b}_j)$	6	-1	4
$f(g_0^2, \check{b}_j)$	1	3	7
$f(g_0^3, \check{b}_j)$	7	5	1
$\check{\psi}_j(G_0)$	1	-1	1
$\check{R}_j(G_0)$	$\{2\}$	$\{1\}$	$\{3\}$
$\psi_j(G_0)$	1	0	1

Последние строки этих таблиц позволяют вычислить значение целевой функции: $F(G_0) = \frac{2}{3}$.

Переходим к вычислению производных по направлению. Имеем

$$\begin{aligned}
&\varphi_i'(G_0, V) \equiv 0 \quad \text{при всех} \quad i \in 1:4;\\
&\psi_1'(G_0, V) = \langle v^2, \check{b}_1 \rangle = 3v^2(1) - 3v^2(2) + v^2(3);\\
&\psi_2'(G_0, V) \equiv 0;\\
&\psi_3'(G_0, V) = \langle v^3, \check{b}_3 \rangle = v^3(1) + 3v^3(2) + v^3(3).
\end{aligned}$$

Производная по направлению V функции $F(G)$ при $G = G_0$ принимает вид

$$F'(G_0, V) = v^2(1) - v^2(2) + \tfrac{1}{3}v^2(3) + \tfrac{1}{3}v^3(1) + v^3(2) + \tfrac{1}{3}v^3(3).$$

Решаем задачу линейного программирования

$$F'(G_0, V) \to \min_{V \in \Omega}, \tag{6}$$

где в качестве Ω возьмём множество (3×3)-матриц V, все элементы которых по модулю не превосходят 10 ($K = 10$). Очевидно, что решением задачи (6) является матрица

$$V_0 = \begin{pmatrix} 0 & 0 & 0 \\ -10 & 10 & -10 \\ -10 & -10 & -10 \end{pmatrix},$$

при этом $F'(G_0, V_0) = -40$. Получили, что V_0 — направление убывания функции $F(G)$ при $G = G_0$. Минимум функции $F(G_0 + tV_0)$ на луче $t > 0$ достигается при $t_0 = 0.1$, поэтому

$$G_1 := G_0 + t_0 V_0 = \begin{pmatrix} 1 & 0 & 3 \\ -1 & 2 & 3 \\ -1 & -2 & 3 \end{pmatrix}.$$

Нетрудно проверить, что $F(G_1) = 0$. Учитывая, что G_1 — подходящая матрица, заключаем, что G_1 — решение рассматриваемой задачи (см. рис. 5).

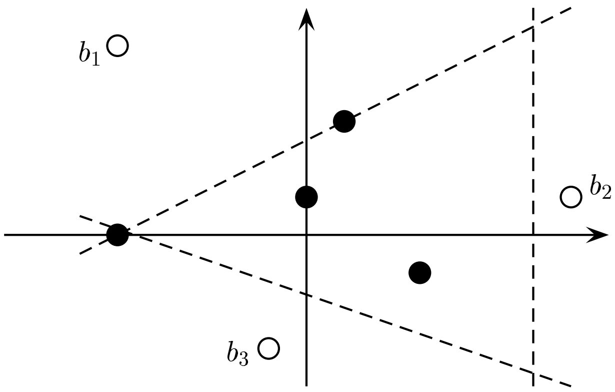

Рис. 5

6°. Изменим в данной задаче значение параметра c. Вместо $c = 0$ положим $c = 1$. В качестве начального приближения возьмём ту же матрицу G_0. Заполним две таблицы.

Таблица 3

i	1	2	3	4
$f(g_0^1, \hat{a}_i)$	-5	-1	0	-2
$f(g_0^2, \hat{a}_i)$	-3	-1	-4	-2
$f(g_0^3, \hat{a}_i)$	-3	-5	-2	-4
$\hat{\varphi}_i(G_0)$	-3	-1	0	-2
$\hat{R}_i(G_0)$	$\{2,3\}$	$\{1,2\}$	$\{1\}$	$\{1,2\}$
$\varphi_i(G_0)$	0	0	0	0

Таблица 4

j	1	2	3
$f(g_0^1, \check{b}_j)$	7	0	5
$f(g_0^2, \check{b}_j)$	2	4	8
$f(g_0^3, \check{b}_j)$	8	6	2
$\check{\psi}_j(G_0)$	2	0	2
$\check{R}_j(G_0)$	$\{2\}$	$\{1\}$	$\{3\}$
$\psi_j(G_0)$	2	0	2

По последним строкам этих таблиц находим значение целевой функции: $F(G_0) = \frac{4}{3}$. Далее имеем

$$\begin{aligned}
&\varphi_i'(G_0, V) \equiv 0 \quad \text{при} \quad i = 1,\ 2,\ 4;\\
&\varphi_3'(G_0, V) = \big[\langle v^1, \hat{a}_3\rangle\big]_+;\\
&\psi_1'(G_0, V) = \langle v^2, \check{b}_1\rangle;\\
&\psi_2'(G_0, V) = \big[\langle v^1, \check{b}_2\rangle\big]_+;\\
&\psi_3'(G_0, V) = \langle v^3, \check{b}_3\rangle.
\end{aligned}$$

Производная по напрвлению V функции $F(G)$ при $G = G_0$ принимает вид

$$F'(G_0, V) = \tfrac{1}{4}\big[\langle v^1, \hat{a}_3\rangle\big]_+ + \tfrac{1}{3}\Big\{\langle v^2, \check{b}_1\rangle + \big[\langle v^1, \check{b}_2\rangle\big]_+ + \langle v^3, \check{b}_3\rangle\Big\}.$$

Задача (4) при $K = 10$ сводится к задаче линейного программирования

$$\tfrac{1}{4}t_1 + \tfrac{1}{3}\big\{\langle v^2, \check{b}_1\rangle + t_2 + \langle v^3, \check{b}_3\rangle\big\} \to \min$$

при ограничениях

$$\begin{gathered}
\big|v^s(\alpha)\big| \leqslant 10, \quad \alpha \in 1:3,\ s \in 1:3;\\
t_1 \geqslant \langle v^1, \hat{a}_3\rangle,\\
t_2 \geqslant \langle v^1, \check{b}_2\rangle,\\
t_1 \geqslant 0, \quad t_2 \geqslant 0.
\end{gathered}$$

Для решения последней задачи использовалась программа из MATLAB [5]. Получили матрицу

$$V_0 = \begin{pmatrix} 3.06 & 7.8 & 6.3\\ -10 & 10 & -10\\ -10 & -10 & -10 \end{pmatrix},$$

при этом $F'(G_0, V_0) = -40$. Минимум функции $F(G_0 + tV_0)$ на луче $t > 0$ достигается при $t_0 = 0.1$, поэтому

$$G_1 = \begin{pmatrix} 1.3 & 0.8 & 3.6\\ -1 & 2 & 3\\ -1 & -2 & 3 \end{pmatrix}$$

(см. рис. 6).

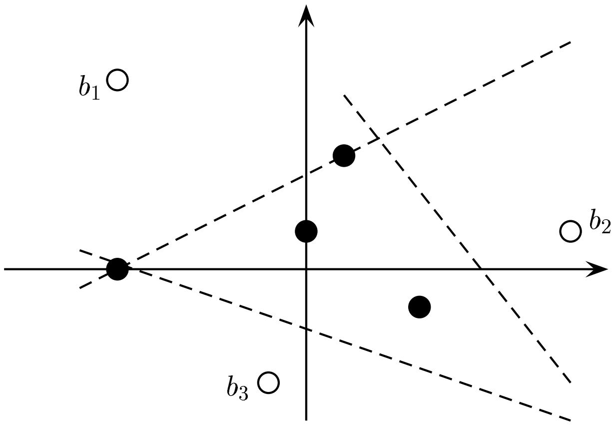

Рис. 6

Непосредственно проверяется, что $F(G_1) = 0.5$.

Повторяем процесс, и на следующем шаге получаем матрицу

$$G_2 = \begin{pmatrix} 1.3 & 0.8 & 3.6 \\ -0.6 & 1.6 & 3.4 \\ -0.6 & -2 & 3.4 \end{pmatrix},$$

при этом $F(G_2) = 0.075$. Проверим для G_2 выполнение строгих неравенств (2). Они выполняются (см. рис. 7). Значит, подходящая матрица G_2 осуществляет строгое 3-отделение (хотя $F(G_2) > 0$).

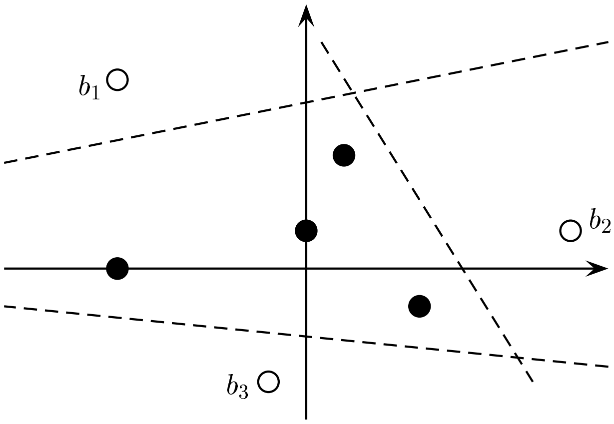

Рис. 7

7°. Еще раз изменим значение параметра c. Положим $c = \frac{1}{2}$. Нетрудно проверить, что $F(G_0) = 1$. Выражение для $F'(G_0, V)$ оказывается таким же, как в п. 5°, поэтому такой же будет матрица G_1:

$$G_1 = \begin{pmatrix} 1 & 0 & 3 \\ -1 & 2 & 3 \\ -1 & -2 & 3 \end{pmatrix}.$$

Вычисляем $F(G_1) = \frac{1}{4}$.

Повторяем процесс. На следующем шаге получаем матрицу

$$G_2 = \begin{pmatrix} 1 & 0 & 3 \\ -0.7 & 1.7 & 3.3 \\ -0.7 & -2 & 3.3 \end{pmatrix},$$

на которой $F(G_2) = 0$. Значит, подходящая матрица G_2 является решением рассматриваемой задачи (см. рис. 8).

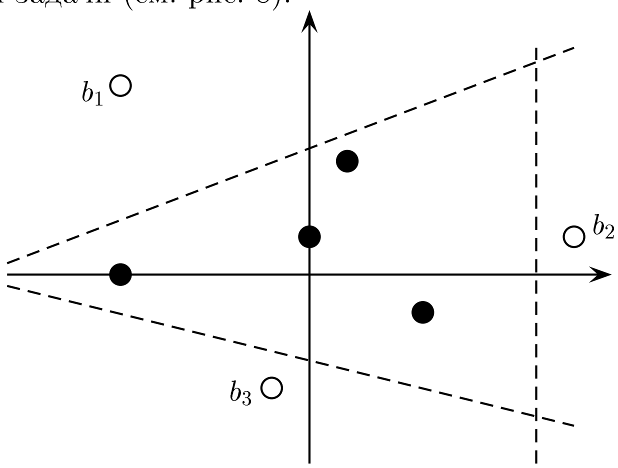

Рис. 8

ЛИТЕРАТУРА

1. Astorino A., Gaudioso M. *Polyhedral separability througt successive LP* // JOTA. 2002. Vol. 112. No. 2. P. 265–293.

2. Малозёмов В. Н., Чернэуцану Е. К. *Строгая h-отделимость двух множеств* // Семинар «DHA & CAGD». Избранные доклады. 18 декабря 2010 г. (`http://dha.spb.ru/reps10.shtml#1218`)

3. Чернэуцану Е. К. *Строгая h-отделимость и линейное программирование* // Семинар «DHA & CAGD». Избранные доклады. 29 января 2011 г. (`http://dha.spb.ru/reps11.shtml#0129`)

4. Малозёмов В. Н., Чернэуцану Е. К. *Численный метод строгого h-отделения* // Семинар «DHA & CAGD». Избранные доклады. 15 октября 2011 г. (`http://dha.spb.ru/reps11.shtml#1015`)

5. Сергеев А. Н., Соловьёва Н. А., Чернэуцану Е. К. *Решение задач линейного программирования в среде MATLAB* // Семинар «DHA & CAGD». Программные реализации. 12 февраля 2011 г. (`http://dha.spb.ru/programs.shtml`)