

Об алгоритме Козинца*

В. Н. Малоземов¹, Н. А. Соловьёва², Г. Ш. Тамасян^{3,4}

¹ Санкт-Петербургский государственный университет,

Российская Федерация, 199034, Санкт-Петербург, Университетская наб., 7–9

² Санкт-Петербургский государственный экономический университет,

Российская Федерация, 191023, Санкт-Петербург, наб. кан. Грибоедова, 30/32

³ Военно-космическая академия им. А. Ф. Можайского,

Российская Федерация, 197198, Санкт-Петербург, ул. Ждановская, 13

⁴ Институт проблем машиноведения РАН,

Российская Федерация, 199178, Санкт-Петербург, Большой пр. В. О., 61

Для цитирования: Малоземов В. Н., Соловьёва Н. А., Тамасян Г. Ш. Об алгоритме Козинца // Вестник Санкт-Петербургского университета. Математика. Механика. Астрономия. 2025. Т. 12 (70). Вып. 1. С. 48–63. <https://doi.org/10.21638/spbu01.2025.104>

В 1964 г., на заре машинного обучения, Б. Н. Козинец предложил простой численный метод (алгоритм) для решения следующей экстремальной задачи. «В n -мерном евклидовом пространстве заданы два конечных множества P_1 и P_2 . Предполагается, что соответствующие выпуклые оболочки C_1 и C_2 этих множеств не имеют общих точек. Требуется построить гиперплоскость, разделяющую множества P_1 и P_2 , т. е. такую гиперплоскость, которая не имеет общих точек со множествами C_1 и C_2 и, кроме того, множества C_1 и C_2 лежат по разные стороны этой гиперплоскости. На самом деле, желательнее среди всех гиперплоскостей, разделяющих множества P_1 и P_2 , найти такую гиперплоскость, расстояние от которой до множества $P_1 \cup P_2$ имеет максимальную величину. Очевидно, что такой гиперплоскостью будет гиперплоскость, проходящая через середину отрезка, соединяющего какие-нибудь две ближайшие точки множеств C_1 и C_2 , перпендикулярно к нему». В дальнейшем эту задачу стали называть задачей жесткого SVM-отделения (Support Vector Machine) двух конечных множеств. В алгоритме Козинца используется естественный геометрический вариант критерия оптимальности для рассматриваемой задачи. В данной статье приводится детальный анализ алгоритма Козинца с современных позиций. В частности, дано корректное доказательство его сходимости. Предложена рабочая схема алгоритма. Рассмотрены два примера, в которых сравнивается эффективность принципиальной и рабочей схем.

Ключевые слова: машинное обучение, жесткое SVM-отделение, алгоритм Козинца, усеченные итерации.

1. Формализация задачи. Критерий оптимальности. Пусть в пространстве \mathbb{R}^n с евклидовой нормой заданы два конечных множества

$$P_1 = \{p_i\}_{i=1}^s \quad \text{и} \quad P_2 = \{p_j\}_{j=s+1}^m,$$

состоящие из попарно различных точек. Здесь $s \in 1 : m - 1$. Будем предполагать, что выпуклые оболочки C_1 и C_2 этих множеств не имеют общих точек. Задачу, поставленную Б. Н. Козинцом [1], с учетом указанной им идеи ее решения можно

*Результаты раздела 8 получены в Институте проблем машиноведения РАН при финансовой поддержке Российского научного фонда (проект № 23-41-00060).

© Санкт-Петербургский государственный университет, 2025

формализовать следующим образом:

$$\|x - y\| \rightarrow \min_{x \in C_1, y \in C_2}. \quad (1)$$

Задача (1) имеет решение, вообще говоря, не единственное. Вместе с тем вектор $w^* = x - y$, где (x, y) — некоторое решение задачи (1), является единственным.

Теорема 1 (Б. Н. Козинец). *Для того чтобы план (x, y) задачи (1) был оптимальным, необходимо и достаточно, чтобы выполнялись неравенства*

$$\langle p_i - x, y - x \rangle \leq 0 \quad \text{для всех } i \in 1 : s, \quad (2)$$

$$\langle p_j - y, x - y \rangle \leq 0 \quad \text{для всех } j \in (s + 1) : m. \quad (3)$$

Рис. 1 поясняет содержание теоремы.

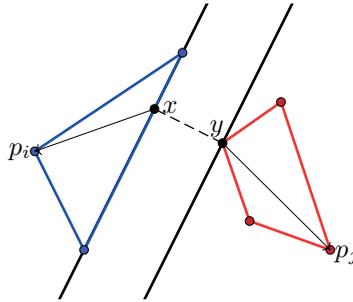


Рис. 1. Пояснение к теореме 1.

ДОКАЗАТЕЛЬСТВО. **Достаточность.** Пусть для плана (x, y) выполнены условия (2) и (3). Зафиксируем индексы $i \in 1 : s$ и $j \in (s + 1) : m$. Сложив (2) и (3), получим

$$\langle x - y - (p_i - p_j), x - y \rangle \leq 0.$$

Отсюда следует, что

$$\|x - y\|^2 \leq \langle p_i, x - y \rangle - \langle p_j, x - y \rangle. \quad (4)$$

Покажем, что для произвольного плана (\tilde{x}, \tilde{y}) выполняется неравенство

$$\|x - y\| \leq \|\tilde{x} - \tilde{y}\|. \quad (5)$$

По определению выпуклой оболочки имеем

$$\tilde{x} = \sum_{i=1}^s u[i] p_i, \quad \tilde{y} = \sum_{j=s+1}^m u[j] p_j,$$

где все $u[i]$, $i \in 1 : m$ неотрицательны и

$$\sum_{i=1}^s u[i] = 1, \quad \sum_{j=s+1}^m u[j] = 1.$$

Умножим (4) на $u[i]$ и просуммируем по $i \in 1 : s$. Получим

$$\|x - y\|^2 \leq \langle \tilde{x}, x - y \rangle - \langle p_j, x - y \rangle, \quad j \in (s + 1) : m. \quad (6)$$

Умножим (6) на $u[j]$ и просуммируем по $j \in (s+1) : m$. Получим

$$\|x - y\|^2 \leq \langle \tilde{x}, x - y \rangle - \langle \tilde{y}, x - y \rangle = \langle \tilde{x} - \tilde{y}, x - y \rangle \leq \|\tilde{x} - \tilde{y}\| \cdot \|x - y\|. \quad (7)$$

Если $x = y$, то неравенство (5) тривиально. В случае $x \neq y$ к неравенству (5) придем, поделив (7) на $\|x - y\|$.

Необходимость. Пусть (x, y) — оптимальный план задачи (1). Покажем, что выполняются неравенства (2) и (3).

Разберемся сначала с неравенством (2). Допустим противное, что найдется индекс $i \in 1 : s$, на котором

$$\langle p_i - x, y - x \rangle > 0. \quad (8)$$

Отсюда, в частности, следует, что $x \neq p_i$. Обозначим через \tilde{x}_i ближайшую к y точку отрезка $[x, p_i]$ и покажем, что

$$\|\tilde{x}_i - y\| < \|x - y\|. \quad (9)$$

Это будет противоречить оптимальности плана (x, y) .

Точки отрезка $[x, p_i]$ допускают представление $x + \lambda(p_i - x)$, где $\lambda \in [0, 1]$. Нахождение точки \tilde{x}_i сводится к решению задачи

$$\varphi_i(\lambda) := \|(x + \lambda(p_i - x)) - y\|^2 \rightarrow \min_{\lambda \in [0, 1]}.$$

Имеем

$$\varphi_i(\lambda) = \|x - y\|^2 - 2\lambda \langle p_i - x, y - x \rangle + \lambda^2 \|p_i - x\|^2,$$

так что

$$\varphi_i'(\lambda) = -2 \langle p_i - x, y - x \rangle + 2\lambda \|p_i - x\|^2.$$

Так как, в силу (8), $\varphi_i'(0) < 0$ и $\varphi_i'(\tilde{\lambda}_i) = 0$ при

$$\tilde{\lambda}_i = \frac{\langle p_i - x, y - x \rangle}{\|p_i - x\|^2} > 0,$$

то $\tilde{x}_i = x + \lambda_i(p_i - x)$, где

$$\lambda_i = \begin{cases} \tilde{\lambda}_i, & \text{если } \tilde{\lambda}_i \leq 1, \\ 1, & \text{если } \tilde{\lambda}_i > 1. \end{cases}$$

Очевидно, что $\tilde{x}_i \in C_1$. По определению $\varphi_i(\lambda)$ получаем

$$\|\tilde{x}_i - y\|^2 = \|x - y\|^2 - 2\lambda_i \langle p_i - x, y - x \rangle + \lambda_i^2 \|p_i - x\|^2. \quad (10)$$

В случае $\tilde{\lambda}_i \leq 1$, т. е. при $\lambda_i = \tilde{\lambda}_i$, из (10) следует, что

$$\|\tilde{x}_i - y\|^2 = \|x - y\|^2 - \left(\frac{\langle p_i - x, y - x \rangle}{\|p_i - x\|} \right)^2 < \|x - y\|^2.$$

Условие $\tilde{\lambda}_i > 1$ можно переписать в эквивалентном виде:

$$\langle p_i - x, y - x \rangle > \|p_i - x\|^2. \quad (11)$$

Поэтому при $\tilde{\lambda}_i > 1$, т. е. при $\lambda_i = 1$, из (10) и (11) следует, что

$$\begin{aligned} \|\tilde{x}_i - y\|^2 &= \|x - y\|^2 - 2\langle p_i - x, y - x \rangle + \|p_i - x\|^2 < \\ &< \|x - y\|^2 - \|p_i - x\|^2 < \|x - y\|^2. \end{aligned}$$

В обоих случаях установлено, что выполняется неравенство (9).

Аналогично показывается, что для оптимального плана (x, y) справедливы неравенства (3). Действительно, допустим вопреки утверждению, что найдется индекс $j \in (s + 1) : m$ такой, что

$$\langle p_j - y, x - y \rangle > 0. \quad (12)$$

Отсюда, в частности, следует, что $y \neq p_j$. Обозначим через \tilde{y}_j ближайшую к x точку отрезка $[y, p_j]$ и покажем, что

$$\|x - \tilde{y}_j\| < \|x - y\|. \quad (13)$$

Нахождение \tilde{y}_j сводится к решению задачи

$$\varphi_j(\lambda) := \|(y + \lambda(p_j - y)) - x\|^2 \rightarrow \min_{\lambda \in [0,1]}.$$

Имеем

$$\varphi_j(\lambda) = \|x - y\|^2 - 2\lambda\langle p_j - y, x - y \rangle + \lambda^2\|p_j - y\|^2,$$

так что

$$\varphi'_j(\lambda) = -2\langle p_j - y, x - y \rangle + 2\lambda\|p_j - y\|^2.$$

Так как, в силу (12), $\varphi'_j(0) < 0$ и $\varphi'_j(\tilde{\lambda}_j) = 0$ при

$$\tilde{\lambda}_j = \frac{\langle p_j - y, x - y \rangle}{\|p_j - y\|^2} > 0, \quad (14)$$

то $\tilde{y}_j = y + \lambda_j(p_j - y)$, где

$$\lambda_j = \begin{cases} \tilde{\lambda}_j, & \text{если } \tilde{\lambda}_j \leq 1, \\ 1, & \text{если } \tilde{\lambda}_j > 1. \end{cases}$$

Очевидно, что $\tilde{y}_j \in C_2$. По определению $\varphi_j(\lambda)$ получаем

$$\|x - \tilde{y}_j\|^2 = \|x - y\|^2 - 2\lambda_j\langle p_j - y, x - y \rangle + \lambda_j^2\|p_j - y\|^2. \quad (15)$$

В случае $\tilde{\lambda}_j \leq 1$, т. е. при $\lambda_j = \tilde{\lambda}_j$, из (15) следует, что

$$\|x - \tilde{y}_j\|^2 = \|x - y\|^2 - \left(\frac{\langle p_j - y, x - y \rangle}{\|p_j - y\|} \right)^2 < \|x - y\|^2.$$

Условие $\tilde{\lambda}_j > 1$ равносильно неравенству

$$\langle p_j - y, x - y \rangle > \|p_j - y\|^2. \quad (16)$$

Поэтому при $\tilde{\lambda}_j > 1$, т. е. при $\lambda_j = 1$, из (15) и (16) следует, что

$$\|x - \tilde{y}_j\|^2 < \|x - y\|^2 - \|p_j - y\|^2 < \|x - y\|^2.$$

В обоих случаях выполняется неравенство (13), что противоречит оптимальности плана (x, y) . \square

2. Оценка плана задачи (1). Пусть (x, y) — план задачи (1). Введем обозначения:

$$\begin{aligned}\Delta^{(1)}(x, y) &= \max_{i \in 1:s} \langle p_i - x, y - x \rangle, \\ \Delta^{(2)}(x, y) &= \max_{j \in (s+1):m} \langle p_j - y, x - y \rangle.\end{aligned}$$

Лемма 1. *Справедливы неравенства*

$$\Delta^{(1)}(x, y) \geq 0, \quad \Delta^{(2)}(x, y) \geq 0.$$

ДОКАЗАТЕЛЬСТВО. Проверим первое неравенство. Допустим противное, что $\Delta^{(1)}(x, y) < 0$. Тогда

$$\langle p_i - x, y - x \rangle < 0 \quad \text{при всех } i \in 1 : s. \quad (17)$$

Точка x принадлежит выпуклой оболочке C_1 . Она допускает представление

$$x = \sum_{i=1}^s u[i] p_i, \quad (18)$$

где все $u[i]$, $i \in 1 : s$, неотрицательны и в сумме равны единице. Умножим неравенство (17) на $u[i]$ и сложим по $i \in 1 : s$. Так как среди коэффициентов $u[i]$ есть положительные, то придем к неравенству

$$\left\langle \sum_{i=1}^s u[i] p_i - x, y - x \right\rangle < 0$$

или, в силу (18), $0 < 0$. Получили противоречие.

Аналогично проверяется второе неравенство $\Delta^{(2)}(x, y) \geq 0$. \square

Величину

$$\Delta(x, y) = \max\{\Delta^{(1)}(x, y), \Delta^{(2)}(x, y)\}$$

назовем *оценкой плана* (x, y) задачи 1. Согласно лемме 1, $\Delta(x, y) \geq 0$ на любом плане (x, y) .

Теорема 2. *Для того чтобы план (x, y) задачи (1) был оптимальным, необходимо и достаточно, чтобы выполнялось равенство $\Delta(x, y) = 0$.*

ДОКАЗАТЕЛЬСТВО. *Необходимость.* Пусть (x, y) — оптимальный план. По теореме 1, $\Delta(x, y) \leq 0$. Вместе с тем всегда $\Delta(x, y) \geq 0$. Значит, $\Delta(x, y) = 0$.

Достаточность. Из условия $\Delta(x, y) = 0$ следует, что $\Delta^{(1)}(x, y) \leq 0$ и $\Delta^{(2)}(x, y) \leq 0$. Последние два неравенства эквивалентны неравенствам (2) и (3), а это по теореме 1 гарантирует оптимальность плана (x, y) . \square

3. Алгоритм Козинца (принципиальная схема). Для решения задачи (1) Б. Н. Козинец в 1964 г. предложил простой алгоритм, идейно близкий методу условного градиента. Приведем принципиальную схему алгоритма Козинца.

В качестве начального приближения возьмем произвольный план (x_0, y_0) задачи (1).

Пусть уже имеется k -е приближение — план (x_k, y_k) . Проверим его на оптимальность. Для этого вычислим величины

$$\Delta_k^{(1)} = \max_{i \in 1:s} \langle p_i - x_k, y_k - x_k \rangle, \quad \Delta_k^{(2)} = \max_{j \in (s+1):m} \langle p_j - y_k, x_k - y_k \rangle,$$

$$\Delta_k = \max\{\Delta_k^{(1)}, \Delta_k^{(2)}\}.$$

Если $\Delta_k = 0$, то (x_k, y_k) — оптимальный план. Процесс закончен.

Пусть $\Delta_k > 0$.

В случае $\Delta_k = \Delta_k^{(1)} = \langle p_{i_k} - x_k, y_k - x_k \rangle$, где $i_k \in 1 : s$, вычисляем

$$\tilde{\lambda}_k = \frac{\Delta_k}{\|p_{i_k} - x_k\|^2}.$$

Полагаем $y_{k+1} = y_k$ и $x_{k+1} = x_k + \lambda_k(p_{i_k} - x_k)$, где

$$\lambda_k = \begin{cases} \tilde{\lambda}_k, & \text{если } \tilde{\lambda}_k \leq 1, \\ 1, & \text{если } \tilde{\lambda}_k > 1. \end{cases}$$

На рис. 2 и 3 демонстрируется шаг алгоритма при $\tilde{\lambda}_k \leq 1$ и $\tilde{\lambda}_k > 1$ соответственно.

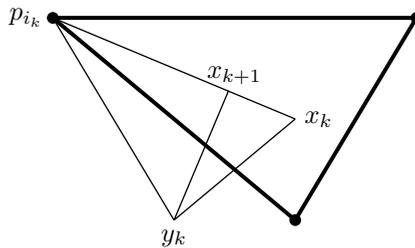


Рис. 2. Шаг алгоритма при $\tilde{\lambda}_k \leq 1$.

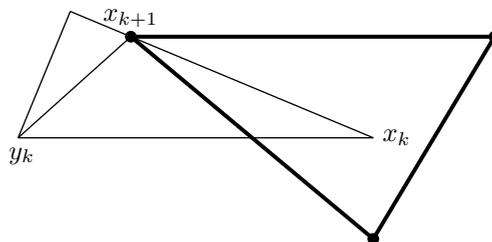


Рис. 3. Шаг алгоритма при $\tilde{\lambda}_k > 1$.

В случае $\Delta_k = \Delta_k^{(2)} = \langle p_{j_k} - y_k, x_k - y_k \rangle$, где $j_k \in (s+1) : m$, вычисляем

$$\tilde{\lambda}_k = \frac{\Delta_k}{\|p_{j_k} - y_k\|^2}.$$

Полагаем $x_{k+1} = x_k$ и $y_{k+1} = y_k + \lambda_k(p_{j_k} - y_k)$, где

$$\lambda_k = \begin{cases} \tilde{\lambda}_k, & \text{если } \tilde{\lambda}_k \leq 1, \\ 1, & \text{если } \tilde{\lambda}_k > 1. \end{cases}$$

Описание алгоритма завершено.

С помощью этого алгоритма строится последовательность планов задачи (1):

$$(x_0, y_0), (x_1, y_1), \dots, (x_k, y_k), \dots \quad (19)$$

Она конечна, если при некотором k_0 выполнится равенство $\Delta_{k_0} = 0$. В этом случае план (x_{k_0}, y_{k_0}) будет оптимальным. Если $\Delta_k > 0$ при всех $k = 0, 1, \dots$, то последовательность (19) бесконечна. Ниже будет доказана ее сходимость по норме к решению задачи (1).

4. Оценка погрешности алгоритма. Пусть (x_k, y_k) — план из последовательности (19) и (x^*, y^*) — оптимальный план. Обозначим

$$w_k = x_k - y_k, \quad w^* = x^* - y^*.$$

Отметим, что

$$\langle w_k, w^* \rangle \geq \langle w^*, w^* \rangle. \quad (20)$$

Действительно, в силу выпуклости множеств C_1 и C_2 при $t \in (0, 1)$ имеем

$$\begin{aligned} \|w^* + t(w_k - w^*)\|^2 &= \|tw_k + (1-t)w^*\|^2 = \\ &= \|(tx_k + (1-t)x^*) - (ty_k + (1-t)y^*)\|^2 \geq \|x^* - y^*\|^2 = \|w^*\|^2. \end{aligned}$$

Отсюда следует, что

$$2t\langle w^*, w_k - w^* \rangle + t^2\|w_k - w^*\|^2 \geq 0.$$

Поделим это неравенство на $t \in (0, 1)$, после чего перейдем к пределу при $t \rightarrow +0$. Получим (20).

Теорема 3. *Справедлива следующая оценка погрешности алгоритма Козинца:*

$$\|w_k - w^*\|^2 \leq 2\Delta_k,$$

где $\Delta_k = \Delta(x_k, y_k)$ — оценка плана (x_k, y_k) .

ДОКАЗАТЕЛЬСТВО. В силу (20) имеем

$$\|w_k - w^*\|^2 = (\|w_k\|^2 - \langle w_k, w^* \rangle) + (\|w^*\|^2 - \langle w_k, w^* \rangle) \leq \|w_k\|^2 - \langle w_k, w^* \rangle.$$

Воспользуемся тем, что вектор w^* допускает представление

$$w^* = x^* - y^* = \sum_{i=1}^s u^*[i] p_i - \sum_{j=s+1}^m u^*[j] p_j,$$

где как $u^*[i]$, $i \in 1 : s$, так и $u^*[j]$, $j \in (s + 1) : m$, неотрицательны и в сумме равны единице. Получим

$$\begin{aligned} \|w_k - w^*\|^2 &\leq \|w_k\|^2 + \sum_{i=1}^s u^*[i] (\langle p_i - x_k, -w_k \rangle - \langle x_k, w_k \rangle) + \\ &+ \sum_{j=s+1}^m u^*[j] (\langle p_j - y_k, w_k \rangle + \langle y_k, w_k \rangle) = \|w_k\|^2 - \langle x_k - y_k, w_k \rangle + \\ &+ \sum_{i=1}^s u^*[i] \langle p_i - x_k, y_k - x_k \rangle + \sum_{j=s+1}^m u^*[j] \langle p_j - y_k, x_k - y_k \rangle \leq \\ &\leq \Delta^{(1)}(x_k, y_k) + \Delta^{(2)}(x_k, y_k) \leq 2\Delta_k. \end{aligned}$$

□

5. Неусеченные и усеченные итерации. В дальнейшем будем считать, что выполнено условие $\Delta_k > 0$ при всех $k = 0, 1, \dots$. Это условие гарантирует, в частности, что при всех $k = 0, 1, \dots$ вектор $w_k = x_k - y_k$ отличен от нулевого.

При $\Delta_k > 0$ план (x_k, y_k) задачи (1) порождает k -ю итерацию алгоритма Козинца. Прежде всего вычисляется $\tilde{\lambda}_k$. При $\tilde{\lambda}_k \in (0, 1]$ k -ю итерацию назовем *неусеченной* (см. рис. 2), при $\tilde{\lambda}_k > 1$ — *усеченной* (рис. 3). На неусеченной итерации полагаем $\lambda_k = \tilde{\lambda}_k$, на усеченной — $\lambda_k = 1$.

Обозначим через D_1 диаметр множества C_1 , через D_2 — диаметр множества C_2 . Положим $D = \max\{D_1, D_2\}$.

Лемма 2. *Допустим, что k -я итерация алгоритма Козинца является неусеченной. Тогда справедливо неравенство*

$$\|w_k\|^2 - \|w_{k+1}\|^2 \geq \left(\frac{\Delta_k}{D}\right)^2. \quad (21)$$

ДОКАЗАТЕЛЬСТВО. Имеем $\Delta_k > 0$ и $\lambda_k = \tilde{\lambda}_k$. Возможны два случая.

1) $\Delta_k = \Delta_k^{(1)} = \langle p_{i_k} - x_k, y_k - x_k \rangle$, где $i_k \in 1 : s$. По описанию алгоритма

$$\tilde{\lambda}_k = \frac{\Delta_k}{\|p_{i_k} - x_k\|^2}$$

и $x_{k+1} = x_k + \tilde{\lambda}_k(p_{i_k} - x_k)$, $y_{k+1} = y_k$. Вычислим

$$\begin{aligned} \|x_{k+1} - y_{k+1}\|^2 &= \|x_k - y_k + \tilde{\lambda}_k(p_{i_k} - x_k)\|^2 = \\ &= \|x_k - y_k\|^2 - 2\tilde{\lambda}_k \langle p_{i_k} - x_k, y_k - x_k \rangle + \tilde{\lambda}_k^2 \|p_{i_k} - x_k\|^2 = \\ &= \|x_k - y_k\|^2 - 2\tilde{\lambda}_k \Delta_k + \frac{\Delta_k^2}{\|p_{i_k} - x_k\|^2} = \|x_k - y_k\|^2 - \frac{\Delta_k^2}{\|p_{i_k} - x_k\|^2}. \end{aligned}$$

Отсюда непосредственно следует неравенство (21).

2) $\Delta_k = \Delta_k^{(2)} = \langle p_{j_k} - y_k, x_k - y_k \rangle$, где $j_k \in (s + 1) : m$. По описанию алгоритма

$$\tilde{\lambda}_k = \frac{\Delta_k}{\|p_{j_k} - y_k\|^2}$$

и $x_{k+1} = x_k$, $y_{k+1} = y_k + \tilde{\lambda}_k(p_{j_k} - y_k)$. Вычислим

$$\begin{aligned} \|x_{k+1} - y_{k+1}\|^2 &= \left\| x_k - y_k - \tilde{\lambda}_k(p_{j_k} - y_k) \right\|^2 = \\ &= \|x_k - y_k\|^2 - 2\tilde{\lambda}_k \langle p_{j_k} - y_k, x_k - y_k \rangle + \tilde{\lambda}_k^2 \|p_{j_k} - y_k\|^2 = \\ &= \|x_k - y_k\|^2 - 2\tilde{\lambda}_k \Delta_k + \frac{\Delta_k^2}{\|p_{j_k} - y_k\|^2} = \|x_k - y_k\|^2 - \frac{\Delta_k^2}{\|p_{j_k} - y_k\|^2}. \end{aligned}$$

Отсюда непосредственно следует неравенство (21). \square

Лемма 2 гарантирует, в частности, что на неусеченных итерациях выполняется неравенство $\|w_k\| > \|w_{k+1}\|$. Нетрудно проверить, что оно выполняется и на усеченных итерациях.

Действительно, пусть $\Delta_k = \Delta_k^{(1)} = \langle p_{i_k} - x_k, y_k - x_k \rangle > 0$, где $i_k \in 1 : s$, и

$$\tilde{\lambda}_k := \frac{\langle p_{i_k} - x_k, y_k - x_k \rangle}{\|p_{i_k} - x_k\|^2} > 1.$$

Последнее неравенство перепишем в виде

$$\langle p_{i_k} - x_k, y_k - x_k \rangle > \|p_{i_k} - x_k\|^2. \quad (22)$$

По описанию алгоритма Козинца имеем $\lambda_k = 1$,

$$x_{k+1} = x_k + (p_{i_k} - x_k) = p_{i_k}, \quad y_{k+1} = y_k.$$

Вычислим

$$\begin{aligned} \|x_{k+1} - y_{k+1}\|^2 &= \|x_k - y_k + (p_{i_k} - x_k)\|^2 = \\ &= \|x_k - y_k\|^2 - 2\langle p_{i_k} - x_k, y_k - x_k \rangle + \|p_{i_k} - x_k\|^2. \end{aligned}$$

На основании (22) и условия $x_k \neq p_{i_k}$ получаем

$$\|x_{k+1} - y_{k+1}\|^2 < \|x_k - y_k\|^2 - \|p_{i_k} - x_k\|^2 < \|x_k - y_k\|^2.$$

В случае $\Delta_k = \Delta_k^{(2)} = \langle p_{j_k} - y_k, x_k - y_k \rangle$, где $j_k \in (s+1) : m$, и

$$\tilde{\lambda}_k := \frac{\langle p_{j_k} - y_k, x_k - y_k \rangle}{\|p_{j_k} - y_k\|^2} > 1$$

аналогично предыдущему придем к неравенству

$$\|x_{k+1} - y_{k+1}\|^2 < \|x_k - y_k\|^2 - \|p_{j_k} - y_k\|^2 < \|x_k - y_k\|^2.$$

Таким образом, во всех случаях выполняется неравенство

$$\|w_k\| > \|w_{k+1}\|, \quad k = 0, 1, \dots \quad (23)$$

Это означает, что числовая последовательность $\{\|w_k\|\}$ строго убывает. Кроме того, она ограничена снизу нулем. По теореме из математического анализа приходим к следующему заключению.

Лемма 3. Числовая последовательность $\{\|w_k\|\}$ имеет конечный предел.

Нам потребуется одно специфическое свойство усеченных итераций.

Лемма 4. Количество идущих подряд усеченных итераций конечно.

ДОКАЗАТЕЛЬСТВО. Допустим, что k -я итерация алгоритма Козинца является усеченной и следом за ней идут только усеченные итерации. Если $\Delta_k = \Delta_k^{(1)}$, то по описанию алгоритма имеем $x_{k+1} = p_{i_k}$, $y_{k+1} = y_k$. При этом $\|w_{k+1}\| < \|w_k\|$. В случае, когда и далее $\Delta_{k+\ell} = \Delta_{k+\ell}^{(1)}$, $\ell = 1, 2, \dots$, получим

$$x_{k+\ell+1} = p_{i_{k+\ell}}, \quad y_{k+\ell+1} = y_k, \quad \|w_{k+\ell+1}\| < \|w_{k+\ell}\|.$$

Такой процесс не может быть бесконечным, поскольку точек p_i конечное число ($i \in 1 : s$). Значит, при некотором ℓ выполнится равенство $\Delta_{k+\ell} = \Delta_{k+\ell}^{(2)}$. По описанию алгоритма получим

$$x_{k+\ell+1} = x_{k+\ell} = p_{i_{k+\ell-1}}, \quad y_{k+\ell+1} = p_{j_{k+\ell}}, \quad \|w_{k+\ell+1}\| < \|w_{k+\ell}\|.$$

Следующие пары (x_r, y_r) будут иметь вид (p_i, p_j) , где $i \in 1 : s$, $j \in (s+1) : m$, при строгом убывании величины $\|w_r\|$. И такой процесс не может быть бесконечным ввиду конечности множества пар (p_i, p_j) .

Аналогично рассматривается случай $\Delta_k = \Delta_k^{(2)}$. □

Согласно лемме 4, последовательность (19) может содержать конечные блоки планов (x_k, y_k) , порождающих усеченные итерации. Если эти блоки исключить из последовательности (19), то останется бесконечная подпоследовательность планов (x_{k_j}, y_{k_j}) , порождающих неусеченные итерации.

6. Доказательство сходимости алгоритма Козинца. Имеется в виду следующее утверждение.

Теорема 4. Если $\Delta_k > 0$ при всех $k = 0, 1, \dots$, то

$$\lim_{k \rightarrow \infty} w_k = w^*.$$

ДОКАЗАТЕЛЬСТВО. Рассмотрим бесконечную подпоследовательность (x_{k_j}, y_{k_j}) последовательности (19), состоящую из планов, порождающих неусеченные итерации. По лемме 2 имеем

$$\|w_{k_j}\|^2 - \|w_{k_{j+1}}\|^2 \geq \left(\frac{\Delta_{k_j}}{D}\right)^2. \quad (24)$$

Согласно лемме 3 последовательность $\{\|w_k\|^2\}$ сходится. Значит, левая часть неравенства (24) стремится к нулю при $j \rightarrow \infty$. Как следствие, получаем

$$\lim_{j \rightarrow \infty} \Delta_{k_j} = 0.$$

По теореме 3

$$\lim_{j \rightarrow \infty} w_{k_j} = w^*.$$

Очевидно, что и

$$\lim_{j \rightarrow \infty} \|w_{k_j}\| = \|w^*\|. \quad (25)$$

Вместе с тем известно, что вся последовательность $\{\|w_k\|\}$ строго убывает и имеет конечный предел (см. (23) и лемму 3). Учитывая (25), приходим к предельному соотношению

$$\lim_{k \rightarrow \infty} \|w_k\| = \|w^*\|. \quad (26)$$

Теперь отметим, что в силу (20) справедливо неравенство

$$\|w_k - w^*\|^2 = \|w_k\|^2 - 2\langle w_k, w^* \rangle + \|w^*\|^2 \leq \|w_k\|^2 - \|w^*\|^2.$$

Согласно (26), правая часть этого неравенства стремится к нулю. Значит,

$$\lim_{k \rightarrow \infty} \|w_k - w^*\|^2 = 0,$$

что равносильно заключению теоремы. \square

7. Алгоритм Козинца (рабочая схема). Рассмотрим возможную вариацию алгоритма Козинца.

Пусть уже имеется k -е приближение (x_k, y_k) . Вычислим оценку

$$\Delta^{(1)}(x_k, y_k) := \max_{i \in 1:s} \langle p_i - x_k, y_k - x_k \rangle = \langle p_{i_k} - x_k, y_k - x_k \rangle.$$

- Если $\Delta^{(1)}(x_k, y_k) > 0$, то находим $x_{k+1} = x_k + \lambda_k(p_{i_k} - x_k)$, где

$$\lambda_k = \min \left\{ 1, \frac{\Delta^{(1)}(x_k, y_k)}{\|p_{i_k} - x_k\|^2} \right\}.$$

- Если $\Delta^{(1)}(x_k, y_k) = 0$, то полагаем $x_{k+1} = x_k$.

Далее вычислим вторую оценку:

$$\Delta^{(2)}(x_{k+1}, y_k) := \max_{j \in (s+1):m} \langle p_j - y_k, x_{k+1} - y_k \rangle = \langle p_{j_k} - y_k, x_{k+1} - y_k \rangle.$$

- Если $\Delta^{(2)}(x_{k+1}, y_k) > 0$, то находим $y_{k+1} = y_k + \lambda_k(p_{j_k} - y_k)$, где

$$\lambda_k = \min \left\{ 1, \frac{\Delta^{(2)}(x_{k+1}, y_k)}{\|p_{j_k} - y_k\|^2} \right\}.$$

- Если $\Delta^{(2)}(x_{k+1}, y_k) = 0$, то полагаем $y_{k+1} = y_k$.

При выполнении условий $\Delta^{(1)}(x_k, y_k) = 0$ и $\Delta^{(2)}(x_{k+1}, y_k) = 0$ план (x_k, y_k) будет оптимальным, так как в этом случае $x_{k+1} = x_k$. Иначе (x_{k+1}, y_{k+1}) — очередное приближение.

Описание рабочей схемы завершено.

Замечание 1. Основное отличие рабочей схемы от принципиальной состоит в том, что очередное приближение в рабочей схеме строится сразу по обоим переменным, при условии положительности соответствующих оценок.

8. Примеры. В этом разделе приведены два примера, иллюстрирующих работу предложенных вычислительных схем алгоритма Козинца. Вычисления заканчивались, когда на очередном приближении (x_k, y_k) соответствующие им оценки удовлетворяли условиям:

$$\begin{cases} \Delta^{(1)}(x_k, y_k) < \varepsilon, \\ \Delta^{(2)}(x_{k+1}, y_k) < \varepsilon, \end{cases}$$

где $\varepsilon > 0$ — параметр метода.

В качестве начального приближения (x_0, y_0) брались центры масс множеств P_1 и P_2 . Результаты расчетов представлены на графиках.

Пример 1. На рис. 4, 5 проиллюстрированы результаты расчетов по принципиальной и рабочей схемам при различных значениях параметра ε .

При $\varepsilon = 0.1$ вычисления по принципиальной схеме закончились за 3 итерации, в рабочей схеме — за 2 итерации.

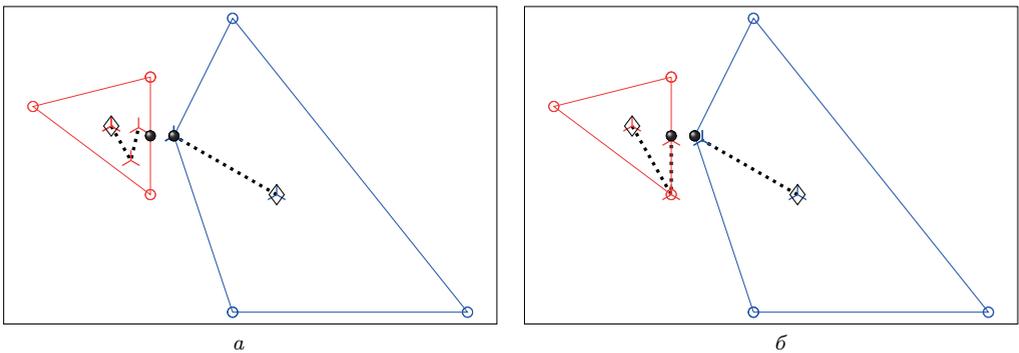


Рис. 4. Результаты расчетов при $\varepsilon = 0.1$ для принципиальной (а) и рабочей схем (б).

Символ \circ — точки множеств P_1 и P_2 ; \blacktriangle — точки приближения x_k и y_k ; \blacklozenge — центры масс множеств P_1 и P_2 ; \bullet — оптимальное решение.

Результаты расчетов по обеим схемам при $\varepsilon = 0.01$ изображены на рис. 5. В принципиальной схеме вычисления закончились за 21 итерацию, в рабочей — за 3 итерации. Обе предельные точки совпали с оптимальными только для рабочей схемы.

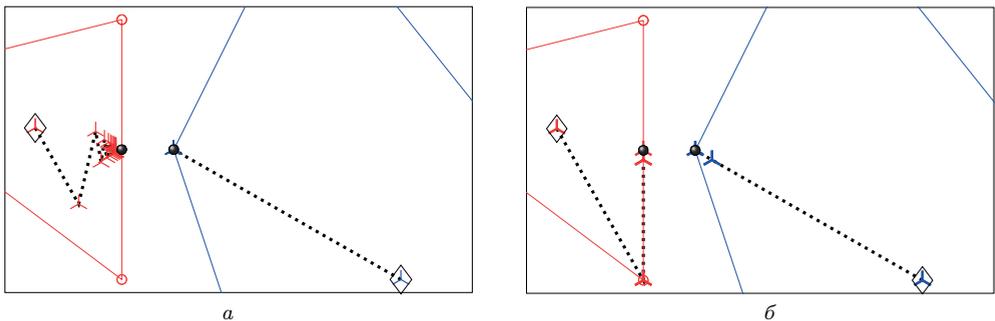


Рис. 5. Результаты расчетов при $\varepsilon = 0.01$ для принципиальной (а) и рабочей схем (б).
Условное обозначение см. на рис. 4.

Аналогичная картина сохранилась для рабочей схемы и при $\varepsilon = 0.001$. Как и ранее, вычисления завершились в оптимальной точке за 3 итерации. Намного увеличилось количество итераций в принципиальной схеме — до 240 итераций. Следовательно, в принципиальной схеме скорость сходимости замедляется при приближении к решению.

Пример 2. На рис. 6 приведены результаты расчетов при $\varepsilon = 0.01$ для множеств, у которых $|P_1| = 30$ и $|P_2| = 50$. К единственному решению задачи последовательность (x_k, y_k) , построенная по рабочей схеме, сошлась на порядок быстрее.

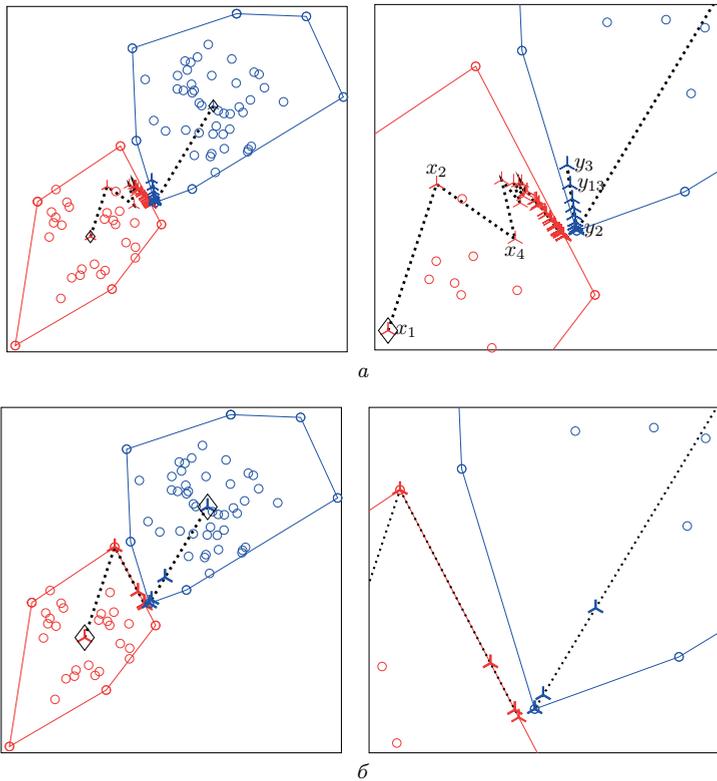


Рис. 6. Результаты расчетов при $\varepsilon = 0.01$ для принципиальной схемы (694 итер.) (а); для рабочей схемы (4 итер.) (б). Условные обозначения см. на рис. 4.

Замечание 2. Трудоемкость одного шага в принципиальной схеме соответствует двум шагам в рабочей схеме.

Замечание 3. Были также проведены массовые вычислительные эксперименты для сравнения эффективности рассматриваемых схем алгоритма Козинца. Замерялось время и вычислялось количество итераций каждой схемы на решении одинаковых задач. В множествах P_1 и P_2 размерность векторов доходила до 200, а их количество — до 3000. В среднем принципиальной схемой задача решается в два раза дольше, чем рабочей схемой. Количество же итераций не намного отличается друг от друга.

9. Заключение. В статье с современных позиций анализируется алгоритм Козинца решения задачи жесткого SVM-отделения двух конечных множеств. Вводится оценка плана рассматриваемой задачи. Оценка неотрицательна на любом плане и обращается в ноль тогда и только тогда, когда план оптимальный. Положительная оценка позволяет улучшить план. Это служит основой для построения алгоритма Козинца.

При доказательстве его сходимости возникло затруднение, связанное с тем, что в алгоритме присутствуют, наряду с неусеченными, усеченные итерации. Это затруднение удалось преодолеть, доказав конечность числа идущих подряд усеченных итераций.

В статье также предложена рабочая схема алгоритма Козинца, с вычислительной точки зрения более эффективная, чем принципиальная схема.

Отметим, что Б. Н. Козинец был сотрудником лаборатории теоретической кибернетики при Ленинградском университете. Его алгоритм активно использовался в лаборатории при решении практических задач распознавания образов [2–5].

В иностранной литературе алгоритм Козинца называют GSK-алгоритмом (Gilbert, Schlezinger, Kozinec) [6–8]).

Результаты настоящей статьи могут оказаться полезными при разработке новых алгоритмов машинного обучения и математической диагностики.

Благодарность. Авторы выражают глубокую благодарность проф. А. Л. Фрадкову за интерес к этой работе и консультации по достижениям лаборатории теоретической кибернетики, руководителем которой был В. А. Якубович.

Литература

1. Козинец Б. Н. Об одном алгоритме обучения линейного персептрона. *Вычислительная техника и вопросы программирования* **3**, 80–83 (1964).
2. Козинец Б. Н., Ланцман Р. М., Якубович В. А. Криминалистическая экспертиза близких почерков при помощи электронно-вычислительных машин. *ДАН СССР* **167**, 1008–1011 (1966).
3. Фомин В. Н. *Математическая теория обучаемых опознающих систем*. Ленинград, Изд-во Ленингр. ун-та (1976).
4. Гелиг А. Х., Матвеев А. С. *Введение в математическую теорию обучаемых распознающих систем и нейронных сетей*. Санкт-Петербург, Изд-во С.-Петерб. ун-та (2014).
5. Матвеев А. С., Фрадков А. Л., Шепелявый А. И. Обзор работ научной школы В. А. Якубовича по искусственному интеллекту и робототехнике. *Вестник Санкт-Петербургского университета. Математика. Механика. Астрономия* **10** (68), вып. 4, 665–685 (2023). <https://doi.org/10.21638/spbu01.2023.406>
6. Gilbert E. G. An iterative procedure for computing the minimum of a quadratic form on a convex Set. *SIAM Journal on Control and Optimization* **4** (1), 61–80 (1966).
7. Шлезингер М. И. *Математические средства обработки изображений*. Киев, Наукова думка (1989).
8. Lopez J., Dorransoro J.R. *A Common framework for the convergence of the GSK, MDM and SMO algorithms*. K. Diamantaras, W. Duch, L. S. Iliadis (eds). Berlin; Heidelberg, Springer-Verlag. ICANN, Part II, LNCS 6353, 82–87 (2010). https://doi.org/10.1007/978-3-642-15822-3_10

Статья поступила в редакцию 10 апреля 2024 г.;

доработана 25 мая 2024 г.;

рекомендована к печати 29 августа 2024 г.

Контактная информация:

Малоземов Василий Николаевич — д-р физ.-мат. наук, проф.;

<https://orcid.org/0000-0003-0706-0578>, v.malozemov@spbu.ru

Соловьева Наталья Анатольевна — канд. физ.-мат. наук;
<https://orcid.org/0000-0001-8117-7438>, 4vinyo@gmail.com
Тамасян Григорий Шаликович — канд. физ.-мат. наук, доц.;
<https://orcid.org/0000-0002-5138-8610>, grigoriytamasjan@mail.ru

About the Kozinec algorithm*

V. N. Malozemov¹, N. A. Solovyeva², G. Sh. Tamasyan^{3,4}

¹ St. Petersburg State University, 7–9, Universitetskaya nab., St. Petersburg, 199034, Russian Federation

² St. Petersburg State University of Economics,
30/32, nab. kan. Griboedova, St. Petersburg, 191023, Russian Federation

³ Mozhaiskiy Space Military Academy,
13, Zhdanovskaya ul., St. Petersburg, 197082, Russian Federation

⁴ Institute of Problems of Mechanical Engineering of the Russian Academy of Sciences,
61, Bolshoi pr., V. O., St. Petersburg, 199178, Russian Federation

For citation: Malozemov V. N., Solovyeva N. A., Tamasyan G. Sh. About the Kozinec algorithm. *Vestnik of Saint Petersburg University. Mathematics. Mechanics. Astronomy*, 2025, vol. 12 (70), issue 1, pp. 48–63. <https://doi.org/10.21638/spbu01.2025.104> (In Russian)

In 1964 at the dawn of machine learning, B. N. Kozinec proposed a simple numerical method (algorithm) for solving the following extremal problem. “In n -dimensional Euclidean space, two finite sets P_1 and P_2 are given. It is assumed that the corresponding convex hulls C_1 and C_2 of these sets do not have common points. It is required to construct a hyperplane separating the sets P_1 and P_2 , i. e. such a hyperplane that does not have common points with the sets C_1 and C_2 and, in addition, the sets C_1 and C_2 lie on opposite sides of this hyperplane. In fact, it is desirable, among all the hyperplanes separating the sets P_1 and P_2 , to find a hyperplane whose distance to the set $P_1 \cup P_2$ has the maximum value. Obviously, such a hyperplane will be a hyperplane passing through the middle of the vector connecting any two nearest points of the sets C_1 and C_2 , perpendicular to it”. Later this problem was called the problem of hard SVM separation of two finite sets (SVM — abbreviation for Support Vector Machine). The Kozinec algorithm uses a natural geometric version of the optimality criterion for the problem under consideration. This article provides a detailed analysis of the Kozinec algorithm from a modern perspective. In particular, a correct proof of its convergence is given. A working scheme algorithm is proposed. Two examples are considered in which the effectiveness of the conceptual and working schemes is compared.

Keywords: machine learning, hard SVM-separation, Kozinec algorithm, MDM-algorithm, SMO-algorithm, clipped iterations.

References

1. Kozinets B. N. About a linear perceptron learning algorithm. *Vychislitel'naya tekhnika i voprosy programirovaniia* **3**, 80–83 (1964). (In Russian)
2. Kozinets B. N., Lantsman R. M., Yakubovich V. A. Forensic examination of close handwriting using electronic computers. *Doklady Akademii nauk SSSR* **167**, 1008–1011 (1966). (In Russian)
3. Fomin V. N. *Mathematical theory of trainable recognition systems*. Leningrad, Leningrad University Press (1976). (In Russian)
4. Gelig A. Kh., Matveev A. S. *Introduction to the mathematical theory of trainable recognition systems and neural networks*. St. Petersburg, St. Petersburg University Press (2014). (In Russian)
5. Matveev A. S., Fradkov A. L., Shepeljavi A. I. An overview of works of the scientific school created by V. A. Yakubovich on the issues of artificial intelligence and robotics. *Vestnik of*

*The results of Sec. 8 were obtained in the Institute of Problems of Mechanical Engineering, Russian Academy of Sciences with the support of the Russian Science Foundation (project no. 23-41-00060).

Saint Petersburg University. Mathematics. Mechanics. Astronomy **10** (68), iss. 4, 665–685 (2023). <https://doi.org/10.21638/spbu01.2023.406> (In Russian) [Eng. transl.: *Vestnik St. Petersburg University. Mathematics* **56**, 478–492 (2023). <https://doi.org/10.1134/S106345412304009X>].

6. Gilbert E. G. An iterative procedure for computing the minimum of a quadratic form on a convex Set. *SIAM Journal on Control and Optimization* **4** (1), 61–80 (1966).

7. Schlesinger M. *Mathematical image processing tools*. Kyiv, Naukova dumka Publ. (1989). (In Russian)

8. Lopez J., Dorronsoro J. R. *A Common framework for the convergence of the GSK, MDM and SMO algorithms*. Diamantaras K., Duch W., Iliadis L. S. (eds). Berlin; Heidelberg, Springer-Verlag. ICANN, Part II, LNCS 6353, 82–87 (2010). https://doi.org/10.1007/978-3-642-15822-3_10

Received: April 10, 2024

Revised: May 25, 2024

Accepted: August 29, 2024

Authors' information:

Vassili N. Malozemov — <https://orcid.org/0000-0003-0706-0578>, v.malozemov@spbu.ru

Natalya A. Solovyeva — <https://orcid.org/0000-0001-8117-7438>, 4vinyo@gmail.com

Grigoriy Sh. Tamasyan — <https://orcid.org/0000-0002-5138-8610>, grigoriytamasjan@mail.ru