О ЛИНЕЙНОМ ОТДЕЛЕНИИ ДВУХ КОНЕЧНЫХ МНОЖЕСТВ ПРИ ИНТЕРВАЛЬНОЙ НЕОПРЕДЕЛЕННОСТИ ДАННЫХ*

В. И. Ерохин

erohin_v_i@mail.ru

27 апреля 2022 г.

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

$$\mathbf{P} = \{\mathfrak{p}_i\}_{i=1}^{m_1} \qquad \text{и} \qquad \mathbf{Q} = \{\mathfrak{q}_i\}_{i=1}^{m_2}.$$

Указанным множествам соответствуют матрицы

$$P = (p_{ij}) \in \mathbb{R}^{m_1 \times n} \qquad \text{и} \qquad Q = (q_{ij}) \in \mathbb{R}^{m_2 \times n}.$$

Элементы матриц P и Q заданы с интервальной неопределенностью. Вместо точных значений элементов p_{ij} и q_{ij} известны их нижние и верхние границы: $\underline{P} = (\underline{p}_{ij}) \in \mathbb{R}^{m_1 \times n}$, $\overline{P} = (\overline{p}_{ij}) \in \mathbb{R}^{m_1 \times n}$, $\underline{Q} = (\underline{q}_{ij}) \in \mathbb{R}^{m_2 \times n}$, $\overline{Q} = (\overline{q}_{ij}) \in \mathbb{R}^{m_2 \times n}$. Приведенные ниже неравенства выполнены поэлементно:

$$\underline{P} \leqslant P \leqslant \overline{P}, \qquad \underline{Q} \leqslant Q \leqslant \overline{Q}. \tag{1}$$

Хорошо известно (см., например, [1]), что при отсутствии интервальной неопределенности задачу *нестрогого* отделения множеств $\mathbf{P}$, $\mathbf{Q}$ гиперплоскостью, не проходящей через начало координат, можно формализовать как задачу решения системы линейных неравенств вида

$$Pw \leqslant 1, \quad Qw \geqslant 1 \qquad \Leftrightarrow \qquad \begin{bmatrix} P \\ -Q \end{bmatrix} w \leqslant \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \tag{2}$$

где $w \in \mathbb{R}^n$ — вектор неизвестных коэффициентов левой части уравнения гиперплоскости (правая часть уравнения нормирована к 1). Очевидно, что для

*Семинар по оптимизации, машинному обучению и искусственному интеллекту «O&ML» http://www.apmath.spbu.ru/oml/

решения задач машинного обучения используются более тонкие методы, чем система (2) (см., например, [3–5]), поэтому будем её рассматривать в качестве удобного модельного примера и отправной точки для дальнейших рассуждений.

Пусть $\mathbf{W}$ — допустимое множество системы (2). Заметим, что при $\mathbf{W} \neq \varnothing$ в общем случае система (2) имеет бесконечное множество решений, и все они, в смысле задачи нестрогого отделения множеств $\mathbf{P}$ и $\mathbf{Q}$ гиперплоскостью, являются *эквивалентными*.

Если $\mathbf{W} = \varnothing$, при определенных условиях (в контексте задач машинного обучения) может быть полезно *псевдорешение* системы (2), формализованное, например, как решение задачи выпуклой негладкой безусловной минимизации

$$\left\| \left[\left[\begin{array}{c} P \\ -Q \end{array} \right] w - \left[\begin{array}{c} 1 \\ -1 \end{array} \right] \right]_+ \right\| \longrightarrow \min_w, \tag{3}$$

где $|| \cdot ||$ — некоторая векторная норма, $[\ \cdot\ ]_+$ — операция положительной срезки, применяемая к векторному аргументу поэлементно.

Некоторым парадоксом является то факт, что вне контекста машинного обучения задача (3) является более привлекательной, чем задача решения системы (2), поскольку при выборе строго выпуклой нормы (например, евклидовой), она имеет единственное решение.

2°. Постановки и методы решения задач отделения множеств с интервальной неопределенностью. Рассмотрим следующие задачи:

Задача 1. Найти вектор $w \in \mathbb{R}^n$, являющийся решением системы линейных неравенств (2) для *любых* матриц P, Q, удовлетворяющих системе линейных неравенств (1).

Задача 2. Найти вектор $w \in \mathbb{R}^n$, являющийся решением задачи (3) для *любых* матриц P, Q, удовлетворяющих системе линейных неравенств (1).

Заметим, что в множества матриц $\mathbf{P}$, $\mathbf{Q}$, удовлетворяющих системе неравенств (1), являются бесконечными. При этом $\mathbf{P}$ имеет $m_1 \times 2^n$ крайних точек, $\mathbf{Q}$ — $m_2 \times 2^n$ крайних точек, что дает основание опасаться NP-трудности задач 1 и 2. Указанные опасения, к счастью, не оправдываются, поскольку в интервальном анализе известна следующая

ТЕОРЕМА 1. [2] *Система линейных неравенств $Ax \leqslant b$ разрешима для всех A, b таких что $\underline{A} \leqslant A \leqslant \overline{A}$, $\underline{b} \leqslant b \leqslant \overline{b}$ тогда и только тогда, когда разрешима система $\overline{A}x^+ - \underline{A}x^- \leqslant \underline{b}$, $x^+, x^- \geqslant 0$. При этом $x = x^+ - x^-$.*

В силу теоремы 1 легко указать методы решения задач 1 и 2.

УТВЕРЖДЕНИЕ 1. *Задача 1 имеет решение тогда и только тогда, когда имеет решение система линейных неравенств*

$$\begin{bmatrix} \overline{P} \\ -\underline{Q} \end{bmatrix} w^+ - \begin{bmatrix} \underline{P} \\ -\overline{Q} \end{bmatrix} w^- \leqslant \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \qquad w^+, w^- \geqslant 0. \tag{4}$$

При этом $w = w^+ - w^-$.

УТВЕРЖДЕНИЕ 2. *Если* $||\cdot||$ *– абсолютная векторная норма, то задача 2 эквивалентна задаче*

$$\|\Psi(w)\| = \left\| \begin{bmatrix} P_c w + P_r\,|w| - 1 \\ -Q_c w + Q_r\,|w| + 1 \end{bmatrix}_+ \right\| \longrightarrow \min_w, \tag{5}$$

где $P_c = (\underline{P} + \overline{P})/2$, $P_r = (\overline{P} - \underline{P})/2$, $Q_c = (\underline{Q} + \overline{Q})/2$, $Q_r = (\overline{Q} - \underline{Q})/2$, $|\cdot|$ *— операция взятия абсолютной величины, применяемая к векторному аргументу поэлементно.*

Доказательство. В силу теоремы 1 задача 2 может быть записана в виде

$$\left\|\Phi(w^+, w^-)\right\| = \left\| \left[\begin{bmatrix} \overline{P} \\ -\underline{Q} \end{bmatrix} w^+ - \begin{bmatrix} \underline{P} \\ -\overline{Q} \end{bmatrix} w^- - \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right]_+ \right\| \longrightarrow \min_{w^+, w^- \geqslant 0}.$$

Учитывая приведенные выше соотношения для P_c, P_r, Q_c, Q_c и проводя перегруппировку аргументов w^+, w^-, выражение $\Phi(w^+, w^-)$ можно переписать в виде

$$\Phi(w^+, w^-) = \left[\begin{bmatrix} P_c \\ -Q_c \end{bmatrix} \left(w^+ - w^-\right) + \begin{bmatrix} P_r \\ Q_r \end{bmatrix} \left(w^+ + w^-\right) - \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right]_+.$$

Заметим, что $w^+ - w^- = w$ в силу теоремы 1, в то время как $|w| \leqslant w^+ + w^-$. С учетом приведенных выше выкладок и свойств абсолютной векторной нормы (см., например [6]) имеем $0 \leqslant \Psi(w) = \Psi(w^+ - w^-) \leqslant \Phi(w^+, w^-)$, $\|\Psi(w)\| \leqslant \|\Phi(w^+, w^-)\|$, $\min_w \|\Psi(w)\| \leqslant \min_{w^+, w^- \geqslant 0} \|\Phi(w^+, w^-)\|$, откуда и получаем доказываемое утверждение. □

Заметим, что (5), также как и (3), является задачей выпуклой негладкой безусловной минимизации, имеющей при использовании строго выпуклой нормы единственное решение.

3°. Сведение к безусловной минимизации в случае отделимых множеств. Рассмотрим задачи

$$\left\| \left[\begin{bmatrix} P \\ -Q \end{bmatrix} w - \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \gamma \cdot 1 \right]_+ \right\| \longrightarrow \min_w (= \delta_\gamma), \qquad (6)$$

$$\left\| \left[\begin{bmatrix} P_c \\ -Q_c \end{bmatrix} w + \begin{bmatrix} P_r \\ Q_r \end{bmatrix} |w| - \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \gamma \cdot 1 \right]_+ \right\| \longrightarrow \min_w (= \delta_\gamma), \qquad (7)$$

где $\gamma > 0$ — скалярный параметр.

Задачи (6) и (7) актуальны в случае линейно отделимых множеств (соответственно точных или обладающих интервальной неопределенностью). Очевидно, что $\delta_\gamma = 0$ при $\gamma = 0$ и задачи (6) и (7) имеют бесконечное множество решений. В то же время можно показать, что начиная с некоторого значения $\gamma > 0$ будет выполняться условие $\delta_\gamma > 0$ и упомянутые выше задачи при использовании строго выпуклой нормы будут иметь единственное решение.

4°. Иллюстративные примеры. На рисунках 1 и 2 представлены иллюстрации результатов решения модельных задач 7 и 5 (с интервальной неопределенностью данных) при $n = 2$ и использовании евклидовой нормы. Расчеты проводились в среде `Mathcad`®`15.0`. Матрицы, содержащие численные значения координат элементов разделяемых множеств и границы соответствующих интервалов, вычислялись с помощью генератора псевдослучайных чисел. Решение задач безусловной минимизации (7) и (5) (вектор w^*) вычислялось с помощью встроенного в систему решателя `Minimize` с использованием метода сопряженных градиентов и конечно-разностной центральной аппроксимацией производных. Минимальное значение γ, гарантирующее выполнение условия $\delta_\gamma > 0$ в задаче (7) выбиралось с помощью одномерного поиска. Уравнение разделяющей прямой подвергалось *дополнительной коррекции*, в результате которой принимало вид

$$w_1^* x_1 + w_1^* x_2 = 1 + \frac{\beta - \alpha}{2},$$

где

$$\alpha = \min_i \left(1 - P_c w^* - P_r |w^*|\right)_i, \ \beta = \min_i \left(-1 + Q_c w^* - Q_r |w^*|\right)_i.$$

Абсолютные значения параметров α и β пропорциональны расстояниям от разделяющей прямой до интервальной границы ближайшего к указанной прямой представителя множества $\mathbf{P}$ и множества $\mathbf{Q}$ соответственно. При этом α, $\beta > 0$ при строгом отделении $\mathbf{P}$ от $\mathbf{Q}$, и α, $\beta < 0$ для неразделимых множеств. Учитывая данную информацию, оказывается также возможным построить две дополнительные прямые ($w_1^* x_1 + w_1^* x_2 = 1 - \alpha$ — граница множества $\mathbf{P}$, $w_1^* x_1 + w_1^* x_2 = 1 + \beta$ — граница множества $\mathbf{Q}$), образующими в случае линейно разделимых множеств «разделительную полосу» (возможно, максимальной ширины), а в случае неразделимых множеств — полосу минимальной ширины, в которой находятся представители обеих разделяемых множеств с учетом их интервальных границ.

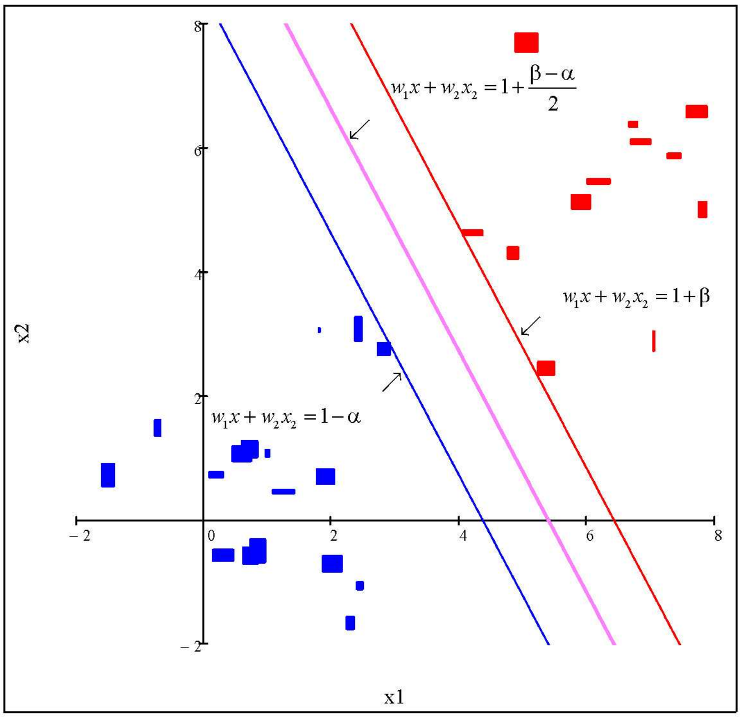

Рис. 1. Строгое линейное отделение двух множеств
с интервальной неопределенностью данных

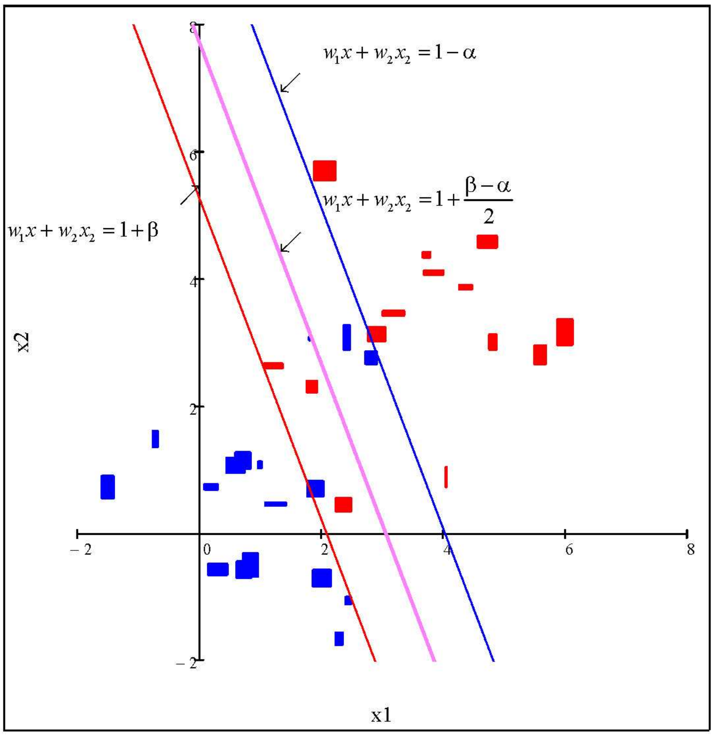

Рис. 2. Нестрогое линейное отделение двух множеств
с интервальной неопределенностью данных

ЛИТЕРАТУРА

1. Вапник В. Н., Червоненкис А. Я. *Теория распознавания образов* (*статистические проблемы обучения*). М.: Наука, 1974.

2. Фидлер М., Недома Й., Рамик Я. и др. *Задачи линейной оптимизации с неточными данными.* М.–Ижевск: НИЦ «Регулярная и хаотическая динамика», Институт компьютерных исследований, 2008.

3. Малоземов В. Н., Плоткин А. В. *Строгое линейное отделение двух конечных множеств и линейное программирование* // Семинар «O&ML». Избранные доклады. 24 февраля 2022 г. (http://www.apmath.spbu.ru/oml/reps22.shtml#0224)

4. Малозёмов В. Н., Плоткин А. В. *SVM-метод строгого линейного отделения двух конечных множеств* // Семинар «O & ML». Избранные доклады. 30 марта 2022 г. (http://www.apmath.spbu.ru/oml/reps22.shtml#0330)

5. Малозёмов В. Н., Плоткин А. В. *SVM-метод: мягкое отделение* // Семинар «O & ML». Избранные доклады. 6 апреля 2022 г. (http://www.apmath.spbu.ru/oml/reps22.shtml#0406)

6. Ланкастер П. *Теория матриц.* М.: Наука, 1973.