МЕТОД ВУЛФА НАХОЖДЕНИЯ ТОЧКИ МНОГОГРАННИКА, БЛИЖАЙШЕЙ К НАЧАЛУ КООРДИНАТ*

М. В. Долгополик

6 декабря 2023 г.

В докладе даётся подробное теоретическое описание и доказательство конечной сходимости метода Вулфа нахождения точки выпуклого многогранника, ближайшей к началу координат, из статьи [1].

1°. При построении и анализе метода Вулфа важную роль играют понятия аффинной оболочки и аффиной независимости векторов. Поэтому для полноты изложения приведём их определения.

Напомним, что множество $M \subseteq \mathbb{R}^n$ называется *аффинным*, если

$$(1-\alpha)x + \alpha y \in M \quad \forall x, y \in M,\ \alpha \in \mathbb{R},$$

то есть если для любых точек $x, y \in M$ множество M содержит прямую $\{x + \alpha(y - x) \in \mathbb{R}^n \mid \alpha \in \mathbb{R}\}$, проходящую через эти точки. *Аффинной оболочкой* $\operatorname{aff} M$ множества $M \subset \mathbb{R}^n$ называется наименьшее аффинное множество, содержащее множество M. Как нетрудно видеть, пересечение любого семейства аффинных множеств является аффинным множеством. Поэтому аффинная оболочка множества M корректно определена и совпадает с пересечением всех аффинных множеств, содержащих множество M. Заметим, что если $M = \{p_1, \ldots, p_m\}$, то

$$\operatorname{aff} M = \Big\{ x = \sum_{i=1}^m \alpha[i] p_i \in \mathbb{R}^n \Bigm| \alpha \in \mathbb{R}^m \colon \sum_{i=1}^m \alpha[i] = 1 \Big\}.$$

В частности, при $m = 1$ будет $\operatorname{aff}\{p_1\} = \{p_1\}$. Для того чтобы дать определение аффинной независимости векторов нам потребуется следующее вспомогательное утверждение.

*Семинар по оптимизации, машинному обучению и искусственному интеллекту «O&ML» http://oml.cmlaboratory.com/

ЛЕММА 1. *Пусть $p_1, \dots, p_m \in \mathbb{R}^n$ — произвольный набор векторов и $M = \operatorname{aff}\{p_1, \dots, p_m\}$. Следующие утверждения эквивалентны:*

1) *ни один из векторов p_j не принадлежит аффинной оболочке всех остальных векторов p_i, то есть $p_j \notin \operatorname{aff}\{p_i \mid i \in 1\colon m \setminus \{j\}\}$ для всех $j \in 1\colon m$;*

2) *для всех $j \in 1\colon m$ векторы $p_i - p_j$, $i \in 1\colon m \setminus \{j\}$, линейно независимы;*

3) *существует $j \in 1\colon m$ такое, что векторы $p_i - p_j$, $i \in 1\colon m \setminus \{j\}$, линейно независимы;*

4) *существует вектор $x \in M$ такой, что размерность линейного подпространства $\operatorname{lin}\{p_1 - x, \dots, p_m - x\}$ равна $m - 1$;*

5) *для любого $x \in M$ размерность $\operatorname{lin}\{p_1 - x, \dots, p_m - x\}$ равна $m - 1$;*

6) *равенства*
$$\sum_{i=1}^{m} \alpha[i] p_i = 0, \quad \sum_{i=1}^{m} \alpha[i] = 0$$
выполняются тогда и только тогда, когда $\alpha = 0$;

7) *векторы $\binom{p_1}{1}, \dots, \binom{p_m}{1}$ линейно независимы.*

Доказательство. $1 \implies 2$. Предположим, что векторы $p_i - p_j$, $i \neq j$, линейно зависимы для некоторого $j \in 1\colon m$. Тогда найдутся такие $\alpha[i]$, $i \neq j$, не равные одновременно нулю, что
$$\sum_{i \neq j} \alpha[i](p_i - p_j) = 0.$$
Если $\beta_1 := \sum_{i \neq j} \alpha[i] \neq 0$, то из данного равенства следует, что
$$p_j = \sum_{i \neq j} \frac{\alpha[i]}{\beta_1} p_i, \quad \sum_{i \neq j} \frac{\alpha[i]}{\beta_1} = 1,$$
то есть $p_j \in \operatorname{aff}\{p_i \mid i \neq j\}$, что противоречит утверждению 1. Если же $\beta_1 = 0$, то для произвольного $k \neq j$ справедливо равенство
$$0 = \sum_{i \neq j} \alpha[i] p_i = \sum_{i \neq j} \alpha[i] p_i - \Big(\sum_{i \neq j} \alpha[i]\Big) p_k = \sum_{i \in I_k} \alpha[i](p_i - p_k), \tag{1}$$
где $I_k := 1\colon m \setminus \{j, k\}$.

Заметим, что система линейных уравнений

$$\sum_{i\in I_k}\alpha[i]=0,\quad k\in 1\colon m\setminus\{j\}$$

имеет только нулевое решение, поскольку, вычитая из уравнения с произвольным индексом k_1 уравнение с произвольным индексом $k_2\neq k_1$, получим $\alpha[k_1]=\alpha[k_2]$ для всех $k_1,k_2\in 1\colon m\setminus\{j\}$. Так как по нашему предположению $\alpha[i]$ не равны нулю одновременно, найдётся $k\neq j$ такое, что $\beta_2:=\sum_{i\in I_k}\alpha[i]\neq 0$. Поэтому, воспользовавшись равенствами (1), получим

$$p_k=\sum_{i\in I_k}\frac{\alpha[i]}{\beta_2}p_i,\quad \sum_{i\in I_k}\frac{\alpha[i]}{\beta_2}=1,$$

то есть $p_k\in\operatorname{aff}\{p_i\mid i\neq j\}$, что противоречит утверждению 1.

Справедливость утверждения 2 $\implies$ 3 очевидна.

3 $\implies$ 4. Поскольку векторы p_i-p_j, $i\in 1\colon m\setminus\{j\}$, линейно независимы, размерность линейного пространства $\operatorname{lin}\{p_1-x,\dots,p_m-x\}$ равна $m-1$ для $x=p_j\in M$.

Справедливость утверждения 4 $\implies$ 5 следует из того факта, что подпространство $\operatorname{lin}\{p_1-x,\dots,p_m-x\}$ не зависит от выбора $x\in M$ (см., например, [2, Теорема 1.2]).

5 $\implies$ 6. Предположим, что для некоторого $\alpha\in\mathbb{R}^m$, $\alpha\neq 0$, выполняются равенства:

$$\sum_{i=1}^m\alpha[i]p_i=0,\quad \sum_{i=1}^m\alpha[i]=0.$$

Тогда $\alpha[1]=-\sum_{i=2}^m\alpha[i]$, по крайней мере один из коэффициентов $\alpha[i]$, $i\in 2\colon m$, отличен от нуля и

$$\sum_{i=2}^m\alpha[i](p_i-p_1)=0,$$

то есть векторы p_i-p_1, $i\in 2\colon m$, линейно зависимы. С другой стороны, полагая $x=p_1$ в утверждении 5, получаем, что размерность линейного подпространства $\operatorname{lin}\{0,p_2-p_1,\dots,p_m-p_1\}$ равна $m-1$, что противоречит линейной зависимости векторов p_i-p_1, $i\in 2\colon m$.

Справедливость утверждения 6 $\implies$ 7 вытекает непосредственно из определения линейной независимости.

7 $\implies$ 1. Предположим, что $p_j\in\operatorname{aff}\{p_i\mid i\neq j\}$ для некоторого $j\in 1\colon m$. Тогда найдутся $\alpha[i]$, $i\neq j$, такие, что

$$p_j=\sum_{i\neq j}\alpha[i]p_i,\quad \sum_{i\neq j}\alpha[i]=1.$$

Полагая $\alpha[j]=-1$, получаем

$$\sum_{i=1}^m\alpha[i]p_i=0,\quad \sum_{i=1}^m\alpha[i]=0,$$

что противоречит линейной независимости векторов $\binom{p_1}{1},\dots\binom{p_m}{1}$. $\square$

ОПРЕДЕЛЕНИЕ 1. Набор векторов $p_1, \ldots, p_m \in \mathbb{R}^n$ называется *аффинно независимым*, если он удовлетворяет любому из эквивалентных утверждений леммы 1.

2°. Рассмотрим задачу проецирования начала координат на выпуклый многогранник, заданный в виде выпуклой оболочки конечного набора точек $p_1, \ldots, p_m \in \mathbb{R}^n$. Данную задачу можно формализовать следующим образом:

$$\|x\| \to \min, \quad x \in G := \operatorname{co}\{p_1, \ldots, p_m\}, \tag{2}$$

где $\|\cdot\|$ — евклидова норма. Метод Вулфа решения этой задачи основан на поиске некоторых специальных наборов векторов, которые мы будем называть базисными.[1]

ОПРЕДЕЛЕНИЕ 2. Набор векторов $\{p_{i_1}, \ldots, p_{i_\ell}\}$ называется *базисным* для задачи (2), если он удовлетворяет двум следующим условиям:

1) векторы $p_{i_1}, \ldots, p_{i_\ell}$ аффинно независимы;

2) проекция начала координат на выпуклую оболочку $\operatorname{co}\{p_{i_1}, \ldots, p_{i_\ell}\}$ этих векторов совпадает с проекцией начала координат на их аффинную оболочку $\operatorname{aff}\{p_{i_1}, \ldots, p_{i_\ell}\}$.

Приведём пример, поясняющий данное определение.

ПРИМЕР. Пусть $n = 2$, $m = 4$,

$$p_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad p_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad p_3 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad p_4 = \begin{pmatrix} -1 \\ 2 \end{pmatrix}.$$

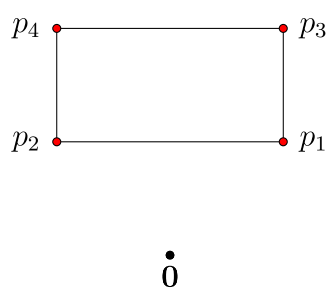

Рис.

Выпишем все базисные наборы для данного примера. Нетрудно видеть, что любой набор, состоящий из одного вектора, является базисным. Непосредственно проверяется, что следующие наборы из двух векторов являются базисными:

$$\{p_1, p_2\}, \quad \{p_3, p_4\}, \quad \{p_2, p_3\}, \quad \{p_4, p_1\}.$$

[1] В оригинальной статье Вулфа [1] такие наборы векторов называются «corral», то есть «загон», «загон для скота». Поскольку данный термин кажется не совсем подходящим для русскоязычной научной литературы, вместо него мы будем использовать термин «базисный набор», опираясь на определённую схожесть метода Вулфа с симплекс-методом и алгоритмом Валле-Пуссена.

Набор $\{p_1, p_3\}$ не является базисным, поскольку проекция нуля на выпуклую оболочку этого набора совпадает с p_1, в то время как проекция нуля на его аффинную оболочку равна $(1,0)^T \in \mathbb{R}^2$. Аналогичным образом доказывается, что набор $\{p_2, p_4\}$ не является базисным. Ни один из наборов, состоящих из трёх векторов, в данном случае не является базисным, поскольку аффинная оболочка любого такого набора совпадает со всем пространством и, следовательно, проекция нуля на аффинную оболочку есть ноль, но $0 \notin G$. Наконец, набор $\{p_1, p_2, p_3, p_4\}$ не является базисным, так как векторы из этого набора аффинно зависимы.

Следующий вспомогательный результат является отправной точкой для построения метода Вулфа.

ЛЕММА 2. *Оптимальный план задачи* (2) *принадлежит выпуклой оболочке некоторого базисного набора для этой задачи.*

Доказательство. Пусть x_* — оптимальный план задачи (2). По теореме Каратеодори (см. [2, Следствие 17.1.1]) найдутся аффинно независимые векторы $p_{i_1}, \dots, p_{i_\ell}$ и вектор $\alpha \in \mathbb{R}^\ell$ такие, что

$$x_* = \sum_{k=1}^{\ell} \alpha[k] p_{i_k}, \quad \sum_{k=1}^{\ell} \alpha[k] = 1, \quad \alpha[k] > 0 \quad \forall k \in 1 : m.$$

Докажем, что набор векторов $p_{i_1}, \dots, p_{i_\ell}$ является базисным.

Заметим, что вектор x_* — оптимальный план задачи

$$\|x\| \to \min, \quad x \in \mathrm{co}\{p_{i_1}, \dots, p_{i_\ell}\}, \tag{3}$$

поскольку он является планом этой задачи и множество планов задачи (3) содержится в множестве планов задачи (2).

Пусть y_* — оптимальный план задачи

$$\|y\| \to \min, \quad y \in \mathrm{aff}\{p_{i_1}, \dots, p_{i_\ell}\}. \tag{4}$$

По определению аффинной оболочки найдутся коэффициенты $\beta \in \mathbb{R}^\ell$ такие, что

$$y_* = \sum_{k=1}^{\ell} \beta[k] p_{i_k}, \quad \sum_{k=1}^{\ell} \beta[k] = 1.$$

Предположим, что набор векторов $p_{i_1}, \dots, p_{i_\ell}$ не является базисным, то есть $y_* \neq x_*$. Тогда $\|y_*\| < \|x_*\|$, так как вектор x_* очевидно является планом задачи (4) и если бы $\|y_*\| = \|x_*\|$, то и вектор y_*, и вектор x_* были бы проекцией нуля на $\mathrm{aff}\{p_{i_1}, \dots, p_{i_\ell}\}$, что невозможно в силу единственности проекции любой точки на аффинное множество.

Определим величину

$$\theta = \min\left\{1, \frac{\alpha[k]}{\alpha[k]-\beta[k]} \;\middle|\; k \in 1\colon m,\ \ \alpha[k]-\beta[k] > 0\right\}.$$

Ясно, что $\theta \in (0,1]$. Положим $z = \theta y_* + (1-\theta)x_*$. Покажем, что вектор z является планом задачи (3).

Действительно, по определению

$$z = \sum_{k=1}^{\ell} \gamma[k] p_{i_k}, \quad \gamma[k] := \theta\beta[k] + (1-\theta)\alpha[k].$$

Заметим, что

$$\sum_{k=1}^{\ell} \gamma[k] = \theta \sum_{k=1}^{\ell} \beta[k] + (1-\theta) \sum_{k=1}^{\ell} \alpha[k] = \theta + (1-\theta) = 1.$$

Кроме того, $\gamma[k] = \alpha[k] - \theta(\alpha[k]-\beta[k])$. Поэтому, если $\alpha[k]-\beta[k] \leqslant 0$, то $\gamma[k] \geqslant 0$. Если же $\alpha[k]-\beta[k] > 0$, то

$$\gamma[k] = (\alpha[k]-\beta[k])\left(\frac{\alpha[k]}{\alpha[k]-\beta[k]} - \theta\right) \geqslant 0$$

по определению θ. Поэтому $z \in \operatorname{co}\{p_{i_1}, \ldots, p_{i_\ell}\}$. С другой стороны, заметим, что

$$\|z\| = \|\theta y_* + (1-\theta)x_*\| \leqslant \theta\|y_*\| + (1-\theta)\|x_*\| < \|x_*\|,$$

поскольку $\theta \in (0,1]$ и $\|y_*\| < \|x_*\|$, что противоречит оптимальности x_* для задачи (3). Таким образом, набор векторов $p_{i_1}, \ldots, p_{i_\ell}$ является базисным. □

З а м е ч а н и е 1. Как было отмечено в доказательстве леммы, если оптимальный план x_* задачи (2) принадлежит выпуклой оболочке некоторого набора векторов $p_{i_1}, \ldots, p_{i_\ell}$, то он является оптимальным планом соответствующей задачи (3). Если данный набор векторов является базисным, то по определению вектор x_* является также оптимальным планом задачи

$$\|x\| \to \min, \quad x \in \operatorname{aff}\{p_{i_1}, \ldots, p_{i_\ell}\}.$$

Таким образом, оптимальный план x_* задачи (2) является проекцией начала координат на аффинную оболочку любого базисного набора, выпуклой оболочке которого он принадлежит. Предыдущая лемма утверждает, что существует по крайней мере один такой базисный набор.

3°. Принципиальная идея метода Вулфа основана на поиске базисного набора векторов такого, что проекция начала координат на аффинную оболочку этого набора совпадает с оптимальным планом задачи (2). В качестве первого базисного набора может быть выбран вектор с минимальной нормой.

Если уже построен некоторый базисный набор p_i, $i \in I_k \subset 1:m$, $k \in \mathbb{N}$, то для оптимального плана x_k задачи

$$\|x\| \to \min, \quad x \in \operatorname{aff}\{p_i \mid i \in I_k\}$$

проверяется хорошо известное необходимое и достаточно условие оптимальности для задачи (2):

$$\langle x_k, p_i \rangle \geqslant \|x_k\|^2 \quad \forall i \in 1:m$$

(см. [1, 3]). Если данное условие выполнено, то x_k — оптимальный план задачи (2) и метод прекращает свою работу. В противном случае, к текущему набору векторов добавляется тот вектор, на котором достигается минимум скалярных произведений $\langle x_k, p_i \rangle$, $i \in 1:m$. В расширенном таким образом наборе векторов осуществляется поиск нового базисного набора p_i, $i \in I_{k+1} \subset 1:m$, расстояние от начала координат до аффинной оболочки которого строго меньше, чем расстояние от начала координат до аффинной оболочки предыдущего базисного набора. Перед тем как дать подробное описание алгоритма поиска нового базисного набора, докажем, что добавление вектора, на котором достигается минимум скалярных произведений $\langle x_k, p_i \rangle$, не нарушает аффинную независимость рассматриваемого набора векторов.

ЛЕММА 3. *Пусть p_i, $i \in I_k \subset 1:m$, — базисный набор для задачи* (2), *x_k — оптимальный план задачи*

$$\|x\| \to \min, \quad x \in \operatorname{aff}\{p_i \mid i \in I_k\} \tag{5}$$

и предположим, что для некоторого $j \in 1:m$ будет

$$\langle x_k, p_j \rangle = \min_{i \in 1:m} \langle x_k, p_i \rangle < \|x_k\|^2. \tag{6}$$

Тогда векторы p_i, $i \in I_k \cup \{j\}$, аффинно независимы.

Доказательство. Предположим, что векторы p_i, $i \in I_k \cup \{j\}$, аффинно зависимы. Тогда по лемме 1 найдутся такие $\alpha[i]$, $i \in I_k \cup \{j\}$, не равные одновременно нулю, что

$$\sum_{i \in I_k \cup \{j\}} \alpha[i] p_i = 0, \quad \sum_{i \in I_k \cup \{j\}} \alpha[i] = 0. \tag{7}$$

Покажем, что из этих соотношений следует, что $\alpha[i] = 0$ для всех i.

Действительно, необходимое и достаточное условие оптимальности для задачи (5) имеет вид:

$$\langle x_k, x - x_k \rangle \geqslant 0 \quad \forall x \in \operatorname{aff}\{p_i \mid i \in I_k\}.$$

Как хорошо известно,

$$\operatorname{aff}\{p_i \mid i \in I_k\} = x_k + \operatorname{lin}\{p_i - x_k \mid i \in I_k\}.$$

Поэтому

$$\langle x_k, x \rangle = 0 \quad \forall x \in \operatorname{lin}\{p_i - x_k \mid i \in I_k\}. \tag{8}$$

С другой стороны, из соотношений (6) следует, что $\langle x_k, p_j - x_k \rangle < 0$.

Воспользовавшись равенствами (7), получим

$$0 = \sum_{i \in I_k \cup \{j\}} \alpha[i] p_i - \Big(\sum_{i \in I_k \cup \{j\}} \alpha[i] \Big) x_k = \sum_{i \in I_k \cup \{j\}} \alpha[i](p_i - x_k).$$

Домножая данное равенство скалярно на x_k, имеем

$$0 = \sum_{i \in I_k \cup \{j\}} \alpha[i] \langle x_k, p_i - x_k \rangle.$$

Как было замечено выше (см. равенство (8)), $\langle x_k, p_i - x_k \rangle = 0$ для всех $i \in I_k$ и $\langle x_k, p_j - x_k \rangle < 0$. Поэтому $\alpha[j] = 0$ и, следовательно, $\alpha[i] = 0$ для всех $i \in I_k$ в силу аффинной независимости векторов p_i, что противоречит предположению о том, что $\alpha[i]$, $i \in I_k \cup \{j\}$, не равны нулю одновременно. Значит, векторы p_i, $i \in I_k \cup \{j\}$, аффинно независимы. $\square$

Предыдущая лемма гарантирует, что добавление нового вектора к текущему базисному набору p_i, $i \in I_k$, не нарушает аффинную независимость рассматриваемой системы векторов. Поэтому для перехода к следующему базисному набору p_i, $i \in I_{k+1}$, требуется только найти такое семейство векторов, для которого проекция начала координат на аффинную оболочку совпадает с проекцией начала координат на выпуклую оболочку. Поиск такого семейства может быть основан на процедуре, неявно описанной в доказательстве леммы 2. Таким образом, мы приходим к следующей теоретической схеме метода Вулфа решения задачи (2):

- **Шаг 0. Построение начального приближения.** Найдём индекс $j \in 1:m$, для которого
$$\|p_j\| = \min_{i \in 1:m} \|p_i\|.$$
Положим $I_0 = \{j\}$, $x_0 = p_j$, $\alpha_0[j] = 1$, $\alpha_0[i] = 0$ для всех $i \neq j$ и $k = 0$.

- **Шаг 1. Проверка условия оптимальности.** Найдём индекс $j_k \in 1\colon m$, для которого
$$\langle x_k, p_{j_k}\rangle = \min_{i\in 1\colon m} \langle x_k, p_i\rangle.$$
Если $\langle x_k, p_{j_k}\rangle \geqslant \|x_k\|^2$, то Стоп. В противном случае, положим $S_k = I_k \cup \{j_k\}$.

- **Шаг 2. Поиск нового базисного набора (внутренний цикл):**
  - **Шаг 2.0.** Положим $S_{k0} = S_k$, $x_{k0} = x_k$, $\alpha_{k0}[i] = \alpha_k[i]$ для всех $i \in I_k$, $\alpha_{k0}[i] = 0$ для всех $i \notin I_k$ и $t = 0$.
  - **Шаг 2.1.** Найдём оптимальный план $y_{kt} = \sum_{i\in S_{kt}} \beta_{kt}[i]p_i$, где $\sum_{i\in S_{kt}} \beta_{kt}[i] = 1$, задачи
$$\|y\| \to \min, \quad y \in \operatorname{aff}\{p_i \mid i \in S_{kt}\}.$$
Если $\beta_{kt}[i] > 0$ для всех $i \in S_{kt}$, то положим $x_{k+1} = y_{kt}$, $\alpha_{k+1} = \beta_{kt}$, $I_{k+1} = S_{kt}$, $k \leftarrow k+1$ и перейдём на **Шаг 1**. Иначе перейдём на следующий шаг.
  - **Шаг 2.2.** Найдём
$$\theta_{kt} = \min\left\{ \frac{\alpha_{kt}[i]}{\alpha_{kt}[i] - \beta_{kt}[i]} \;\middle|\; i \in S_{kt}\colon \alpha_{kt}[i] - \beta_{kt}[i] > 0 \right\},$$
положим
$$x_{k(t+1)} = \theta_{kt}y_{kt} + (1-\theta_{kt})x_{kt}, \quad \alpha_{k(t+1)} = \theta_{kt}\beta_{kt} + (1-\theta_{kt})\alpha_{kt},$$
и $S_{k(t+1)} = S_{kt}$. Удалим из множества $S_{k(t+1)}$ по крайней мере один индекс i, для которого $\alpha_{k(t+1)}[i] = 0$. Положим $t \leftarrow t+1$ и перейдём на **Шаг 2.1**.

Исследуем сходимость данного метода. Для доказательства его сходимости нам потребуется два вспомогательных утверждения.

ЛЕММА 4. *Пусть для некоторых $k, t \in \{0, 1, 2, \dots\}$ метод Вулфа выполняет Шаг 2.2. Тогда справедливы следующие утверждения:*

1) $\theta_{kt} \in [0, 1]$;

2) $\alpha_{k(t+1)}[i] \geqslant 0$ для всех $i \in S_{k(t+1)}$;

3) $\sum_{i\in S_{k(t+1)}} \alpha_{kt}[i] = 1$ (т.е. $x_{k(t+1)} \in \operatorname{co}\{p_i \mid i \in S_{k(t+1)}\}$);

4) существует индекс $i \in S_{kt}$ такой, что $\alpha_{k(t+1)}[i] = 0$.

Доказательство. Докажем лемму индукцией по t. Заметим, что $\alpha_{k0}[i] \geqslant 0$ для всех $i \in S_{k0}$ и $\sum_{i\in S_{k0}} \alpha_{k0}[i] = 1$, в силу условия выхода из внутреннего цикла метода Вулфа (см. Шаг 2.1) и определения α_{k0}.

Предположим теперь, что $\sum_{i\in S_{kt}} \alpha_{kt}[i] = 1$ и $\alpha_{kt}[i] \geqslant 0$, $i \in S_{kt}$, для некоторого t и метод Вулфа выполняет Шаг 2.2 для данного t. Так как данный шаг выполняется тогда и только тогда, когда $\beta_{kt}[i_0] \leqslant 0$ для некоторого $i_0 \in S_{kt}$, можно заключить, что $\theta_{kt} \in [0,1]$, поскольку

$$\frac{\alpha_{kt}[i]}{\alpha_{kt}[i]-\beta_{kt}[i]} \begin{cases} \geqslant 0, & \forall i \in S_{kt}\colon \alpha_{kt}[i]-\beta_{kt}[i] > 0, \\ \leqslant 1, & \text{при } i = i_0. \end{cases}$$

По определению $\alpha_{k(t+1)} = \theta_{kt}\beta_{kt} + (1-\theta_{kt})\alpha_{kt}$. Поэтому

$$\begin{aligned} \sum_{i\in S_{k(t+1)}} \alpha_{k(t+1)}[i] &= \sum_{i\in S_{kt}} \alpha_{k(t+1)}[i] \\ &= \theta_{kt}\sum_{i\in S_{kt}} \beta_{kt}[i] + (1-\theta_{kt})\sum_{i\in S_{kt}} \alpha_{kt}[i] = \theta_{kt} + (1-\theta_{kt}) = 1. \end{aligned}$$

Кроме того, заметим, что

$$\alpha_{k(t+1)}[i] = \theta_{kt}\beta_{kt}[i] + (1-\theta_{kt})\alpha_{kt}[i] = \alpha_{kt}[i] - \theta_{kt}(\alpha_{kt}[i]-\beta_{kt}[i]).$$

Поэтому, если $\beta_{kt}[i]-\alpha_{kt}[i] \leqslant 0$, то $\alpha_{k(t+1)}[i] \geqslant 0$. Если же $\beta_{kt}[i]-\alpha_{kt}[i] > 0$, то

$$\alpha_{k(t+1)}[i] = (\alpha_{kt}[i]-\beta_{kt}[i])\left(\frac{\alpha_{kt}[i]}{\alpha_{kt}[i]-\beta_{kt}[i]} - \theta_{kt}\right) \geqslant 0$$

по определению θ_{kt}. Более того, $\alpha_{k(t+1)}[i] = 0$ для любого $i \in S_{kt}$, на котором достигается минимум из определения θ_{kt}. □

ЛЕММА 5. *Для всех $k \in \{0,1,2,\ldots\}$ будет $\|y_{k0}\| < \|x_k\|$ и $\theta_{k0} > 0$.*

Доказательство. Определим вектор $y(\gamma) = \gamma p_{j_k} + (1-\gamma)x_k$. Заметим, что $y(\gamma) \in \operatorname{aff}\{p_i \mid i \in S_{k0}\}$ для любого $\gamma \in \mathbb{R}$ в силу того, что $x_k \in \operatorname{aff}\{p_i \mid i \in I_k\}$ и $S_{k0} = I_k \cup \{j_k\}$.

Имеем

$$f(\gamma) := \|y(\gamma)\|^2 = \|x_k\|^2 + 2\gamma\langle x_k, p_{j_k} - x_k\rangle + \gamma^2\|p_{j_k} - x_k\|^2.$$

По определению метода Вулфа Шаг 2.1 выполняется тогда и только тогда, когда $\langle x_k, p_{j_k}\rangle < \|x_k\|^2$. Поэтому $f'(0) = \langle x_k, p_{j_k} - x_k\rangle < 0$ и для достаточно малых $\gamma > 0$ будет $f(\gamma) < f(0)$. Отсюда следует, что

$$\|y_{k0}\|^2 = \min\left\{\|y\|^2 \;\middle|\; y \in \operatorname{aff}\{p_i \mid i \in S_{k0}\}\right\} \leqslant \min_{\gamma\in\mathbb{R}} f(\gamma) < f(0) = \|x_k\|^2,$$

то есть $\|y_{k0}\| < \|x_k\|$.

Учитывая условие выхода из внутреннего цикла, имеем $\alpha_{k0}[i] > 0$ для всех $i \in I_k$ и $\alpha_{k0}[j_k] = 0$. Поэтому $\theta_{k0} > 0$, если $\beta_{k0}[j_k] \geqslant 0$. Проверим справедливость неравенства $\beta_{k0}[j_k] \geqslant 0$.

Действительно, по определению

$$y_{k0} = \sum_{i \in I_k} \beta_{k0}[i] p_i + \beta_{k0}[j_k] p_{j_k}, \quad \sum_{i \in I_k} \beta_{k0}[i] + \beta_{k0}[j_k] = 1.$$

Вычитая из обеих частей первого равенства x_k и умножая скалярно получившееся равенство на x_k, получим

$$\langle x_k, y_{k0} - x_k \rangle = \sum_{i \in I_k} \beta_{k0}[i] \langle x_k, p_i - x_k \rangle + \beta_{k0}[j_k] \langle x_k, p_{j_k} - x_k \rangle.$$

Как было замечено в доказательстве леммы 3, из оптимальности x_k для задачи

$$\|y\| \to \min, \quad y \in \operatorname{aff}\{p_i \mid i \in I_k\}$$

(см. Шаг 2.1 метода Вулфа) следует, что $\langle x_k, p_i - x_k \rangle = 0$ для всех $i \in I_k$. Поэтому

$$\langle x_k, y_{k0} - x_k \rangle = \beta_{k0}[j_k] \langle x_k, p_{j_k} - x_k \rangle.$$

Заметим, что $\langle x_k, p_{j_k} - x_k \rangle < 0$ в силу условия перехода на Шаг 2. Кроме того, из первого утверждения леммы следует, что

$$\|y_{k0}\|^2 = \|x_k + (y_{k0} - x_k)\|^2 = \|x_k\|^2 + 2\langle x_k, y_{k0} - x_k \rangle + \|y_{k0} - x_k\|^2 < \|x_k\|^2,$$

то есть $\langle x_k, y_{k0} - x_k \rangle < 0$. Поэтому $\beta_{k0}[j_k] > 0$, что и требовалось доказать. □

ТЕОРЕМА. *Метод Вулфа сходится к оптимальному плану задачи* (2) *за конечное число шагов.*

Доказательство. Заметим, что по лемме 4 Шаг 2.2 метода Вулфа корректно определён и на каждой итерации внутреннего цикла этого метода отбрасывается по крайней мере один индекс из множества $S_{k(t+1)}$, то есть $|S_{k(t+1)}| < |S_{kt}|$ для всех k и t. Кроме того, по лемме 3 для всех $k, t \in \{0\} \cup \mathbb{N}$ векторы p_i, $i \in S_{kt}$, аффинно независимы. Поэтому через конечное число $t+1$, $t \in \{0\} \cup \mathbb{N}$, итераций внутреннего цикла, либо будет выполнено условие выхода из внутреннего цикла на Шаге 2.1, а потому векторы p_i, $i \in S_{kt} = I_{k+1}$, являются базисным набором, либо $|S_{k(t+1)}| = 1$. Если $|S_{k(t+1)}| = 1$, то на $t+2$ итерации внутреннего цикла, как нетрудно видеть, условие выхода будет выполнено автоматически. Таким образом, внутренний цикл корректно определён и конечен.

Рассмотрим последовательность $\{\|x_k\|\}$. Если для некоторого k будет $\beta_{k0}[i] > 0$ для всех $i \in S_{k0}$, то $x_{k+1} = y_{k0}$ и по лемме 5 справедливо неравенство $\|x_{k+1}\| < \|x_k\|$. Если же данное условие не выполнено, то

$$\|x_{k1}\| = \|\theta_{k0}y_{k0} + (1-\theta_{k0})x_k\| \leqslant \theta_{k0}\|y_{k0}\| + (1-\theta_{k0})\|x_k\| < \|x_k\|$$

по лемме 5. Кроме того, по лемме 4 имеем $x_{kt} \in \mathrm{co}\{p_i \mid i \in S_{kt}\}$ для всех t, а потому $\|y_{kt}\| \leqslant \|x_{kt}\|$ и

$$\|x_{k(t+1)}\| = \left\|\theta_{kt}y_{kt} + (1-\theta_{kt})x_{kt}\right\| \leqslant \theta_{kt}\|y_{kt}\| + (1-\theta_{kt})\|x_{kt}\| \leqslant \|x_{kt}\|$$

для всех $t \in \mathbb{N}$. Поэтому $\|x_{k+1}\| < \|x_k\|$ для всех $k \in \{0\} \cup \mathbb{N}$.

По определению вектор x_k является проекцией нуля на аффинную оболочку векторов p_i, $i \in I_k$, для любого $k \in \{0\} \cup \mathbb{N}$. Поэтому

$$\mathrm{dist}\Big(0, \mathrm{aff}\left\{p_i \mid i \in I_{k+1}\right\}\Big) = \|x_{k+1}\| < \|x_k\| = \mathrm{dist}\Big(0, \mathrm{aff}\left\{p_i \mid i \in I_k\right\}\Big).$$

Кроме того, согласно условию выхода из внутреннего цикла на Шаге 2.1 и лемме 3 набор векторов p_i, $i \in I_k$, является базисным для любого k. Значит, все наборы p_i, $i \in I_k$, являются базисными и различными. Так как различных базисных наборов конечное число, через конечное число k итераций метод Вулфа должен найти такой набор, для которого выполняется условие остановки, являющееся необходимым и достаточным условием оптимальности для задачи (2). Соответствующий вектор x_k будет оптимальным планом этой задачи. □

4°. На каждой итерации внутреннего цикла (процедуры поиска нового базисного набора) метода Вулфа необходимо находить проекцию нуля на аффинную оболочку конечного числа аффинно независимых точек. Обсудим возможные методы решения этой задачи подробнее.

Пусть задан некоторый набор аффинно независимых векторов $p_1, \ldots, p_m \in \mathbb{R}^n$. Рассмотрим задачу

$$\|x\| \to \min, \quad x \in \mathrm{aff}\left\{p_1, \ldots, p_m\right\}.$$

Эта задача эквивалентная следующей задаче квадратичного программирования:

$$\frac{1}{2}\Big\|\sum_{i=1}^{m} v[i]p_i\Big\|^2 \to \min, \quad \sum_{i=1}^{m} v[i] = 1. \tag{9}$$

Условия Куна-Таккера для данной задачи имеют следующий вид. Существует множитель Лагранжа $\lambda \in \mathbb{R}$ такой, что

$$P^T P v + \lambda e = 0, \quad \langle e, v\rangle = 1, \tag{10}$$

где $P = (p_1, \ldots, p_m) \in \mathbb{R}^{n\times m}$ и $e = (1, \ldots, 1)^T \in \mathbb{R}^m$. Заметим, что эти условия линейны по v и λ. Кроме того, система линейных уравнений (10) имеет единственное решение в силу единственности оптимального плана задачи (9), которая следует из аффинной независимости векторов $p_1, \ldots, p_m$. Поэтому, разрешая данную систему линейных уравнений относительно (v, λ), можно найти оптимальный план задачи (9).

Покажем, что из системы (10) можно исключить множитель Лагранжа λ.

ЛЕММА 6. *Если пара* (v, λ) *является решением системы линейных уравнений* (10)*, то* $\lambda \leqslant 0$ *и вектор* $u = (1-\lambda)^{-1}v$ *является решением системы линейных уравнений*

$$(ee^T + P^T P)u = e. \tag{11}$$

Обратно, если вектор u *— решение системы линейных уравнений* (11)*, то* $\langle e, u\rangle > 0$ *и пара* $(cu, 1-c)$*, где* $c = 1/\langle e, u\rangle$*, является решением системы линейных уравнений* (10)*.*

Доказательство. Пусть пара (v, λ) является решением системы линейных уравнений (10). Тогда

$$\|Pv\|^2 = \langle v, P^T P v\rangle = -\lambda\langle v, e\rangle = -\lambda,$$

откуда следует, что $\lambda \leqslant 0$. Положим $u = (1-\lambda)^{-1}v$. Тогда

$$(ee^T + P^T P)u = \langle e, u\rangle e + P^T P u = \frac{1}{1-\lambda}e - \frac{\lambda}{1-\lambda}e = e,$$

то есть вектор u является решением системы линейных уравнений (11).

Обратно, пусть u — решение системы линейных уравнений (11). Тогда

$$\langle u, e\rangle^2 + \|Pu\|^2 = \langle u, (ee^T + P^T P)u\rangle = \langle u, e\rangle. \tag{12}$$

Поэтому $\langle e, u\rangle \geqslant 0$. Если $\langle e, u\rangle = 0$, то из равенств (12) следует, что $Pu = 0$ и следовательно $(ee^T + P^T P)u = 0$, что противоречит предположению о том, что u — решение системы линейных уравнений (11). Таким образом, $\langle e, u\rangle > 0$.

Определим $v = \langle e, u\rangle^{-1}u$. Тогда

$$P^T P v = \frac{1}{\langle e, u\rangle}P^T P u = \frac{1}{\langle e, u\rangle}\big(e - ee^T u\big) = \left(\frac{1}{\langle e, u\rangle} - 1\right)e,$$

то есть пара (v, λ), где $\lambda = 1 - \langle e, u\rangle^{-1}$, является решением системы линейных уравнений (10). $\square$

Как было замечено выше, решение системы линейных уравнений (10) существует и единственно. Следовательно, решение системы линейных уравнений (11) также существует и единственно, так как между решениями этих систем установлено взаимно однозначное соответствие.

З а м е ч а н и е 2. Вектор u является решением системы линейных уравнений (11) тогда и только тогда, когда он является псевдорешением системы линейных уравнений $Au = b$, то есть оптимальным планом задачи

$$f(u) := \frac{1}{2}\|Au - b\|^2 \to \min,$$

где $A = \binom{e^T}{P}$ и $b = (1, 0, \ldots, 0)^T \in \mathbb{R}^{m+1}$, так как система уравнений (11) совпадает с равенством $\nabla f(u) = 0$.

ЛИТЕРАТУРА

1. Wolfe P. *Finding the nearest point in a polytope* // Mathematical Programming. 1976. Vol. 11, no. 1, pp. 128–149.

2. Рокафеллар Р. Выпуклый анализ. М.: Мир, 1973.

3. Малозёмов В. Н. *МДМ-методу 50 лет* // Семинар «CNSA & NDO». Избранные доклады. 10 ноября 2021 г.