КОМПЬЮТЕРНЫЕ ИССЛЕДОВАНИЯ И МОДЕЛИРОВАНИЕ 2017 Т. 9 № 5 С. 679−703
DOI: 10.20537/2076-7633-2017-9-5-679-703

УДК: 519.85

Прямые мультипликативные методы для разреженных матриц. Ньютоновские методы

А. Б. Свириденко

ФГБОУ ВПО «Кубанский государственный университет», филиал в г. Новороссийске,
Россия, 353922, г. Новороссийск, ул. Героев Десантников, д. 87

E-mail: roshechka@gmail.com

Получено 20.02.2017, после доработки — 22.09.2017.
Принято к публикации 29.09.2017.

Рассматривается численно устойчивый прямой мультипликативный алгоритм решения систем линейных уравнений, учитывающий разреженность матриц, представленных в упакованном виде. Преимущество алгоритма состоит в возможности минимизации заполнения главных строк мультипликаторов без потери точности результатов, причем изменения в позиции очередной обрабатываемой строки матрицы не вносятся, что позволяет использовать статические форматы хранения данных. Решение системы линейных уравнений прямым мультипликативным алгоритмом — это, как и решение с помощью LU-разложения, просто другая схема реализации метода исключения Гаусса.

В данной работе этот алгоритм лежит в основе решения следующих задач.

Задача 1. Задание направления спуска в ньютоновских методах безусловной оптимизации путем интеграции одной из известных техник построения существенно положительно определенной матрицы. Такой подход позволяет ослабить или снять дополнительные специфические трудности, обусловленные необходимостью решения больших систем уравнений с разреженными матрицами, представленных в упакованном виде.

Задача 2. Построение новой математической формулировки задачи квадратичного программирования и новой формы задания необходимых и достаточных условий оптимальности. Они достаточно просты и могут быть использованы для построения методов математического программирования, например для поиска минимума квадратичной функции на многогранном множестве ограничений, основанного на решениях систем линейных уравнений, размерность которых не выше числа переменных целевой функции.

Задача 3. Построение непрерывного аналога задачи минимизации вещественного квадратичного многочлена от булевых переменных и новой формы задания необходимых и достаточных условий оптимальности для разработки методов их решения за полиномиальное время. В результате исходная задача сводится к задаче поиска минимального расстояния между началом координат и угловой точкой выпуклого многогранника (полиэдра), который является возмущением n-мерного куба и описывается системой двойных линейных неравенств с верхней треугольной матрицей коэффициентов с единицами на главной диагонали. Исследованию подлежат только две грани, одна из которых или обе содержат вершины, ближайшие к началу координат. Для их вычисления достаточно решить $4n - 4$ систем линейных уравнений и выбрать среди них все ближайшие равноудаленные вершины за полиномиальное время. Задача минимизации квадратичного полинома является NP-трудной, поскольку к ней сводится NP-трудная задача о вершинном покрытии для произвольного графа. Отсюда следует вывод, что $P = NP$, в основе построения которого лежит выход за пределы целочисленных методов оптимизации.

Ключевые слова: NP-трудные задачи, разреженные матрицы, ньютоновские методы, прямой мультипликативный алгоритм, направление спуска, новые математические формулировки, необходимые и достаточные условия оптимальности, минимизация псевдобулевой функции, псевдобулево программирование, линейное программирование

COMPUTER RESEARCH AND MODELING
2017 VOL. 9 NO. 5 P. 679−703
DOI: 10.20537/2076-7633-2017-9-5-679-703

MATHEMATICAL MODELING AND NUMERICAL SIMULATION

UDC: 519.85

Direct multiplicative methods for sparse matrices. Newton methods

A. B. Sviridenko

FSEI of HPE "Kuban State University", branch in Novorossiysk,
Geroev Desantnikov st. 87, Novorossiysk, 353922, Russia

E-mail: roshechka@gmail.com

Received 20.02.2017, after completion — 22.09.2017.
Accepted for publication 29.09.2017.

We consider a numerically stable direct multiplicative algorithm of solving linear equations systems, which takes into account the sparseness of matrices presented in a packed form. The advantage of the algorithm is the ability to minimize the filling of the main rows of multipliers without losing the accuracy of the results. Moreover, changes in the position of the next processed row of the matrix are not made, what allows using static data storage formats. Linear system solving by a direct multiplicative algorithm is, like the solving with *LU*-decomposition, just another scheme of the Gaussian elimination method implementation.

In this paper, this algorithm is the basis for solving the following problems:

Problem 1. Setting the descent direction in Newtonian methods of unconditional optimization by integrating one of the known techniques of constructing an essentially positive definite matrix. This approach allows us to weaken or remove additional specific difficulties caused by the need to solve large equation systems with sparse matrices presented in a packed form.

Problem 2. Construction of a new mathematical formulation of the problem of quadratic programming and a new form of specifying necessary and sufficient optimality conditions. They are quite simple and can be used to construct mathematical programming methods, for example, to find the minimum of a quadratic function on a polyhedral set of constraints, based on solving linear equations systems, which dimension is not higher than the number of variables of the objective function.

Problem 3. Construction of a continuous analogue of the problem of minimizing a real quadratic polynomial in Boolean variables and a new form of defining necessary and sufficient conditions of optimality for the development of methods for solving them in polynomial time. As a result, the original problem is reduced to the problem of finding the minimum distance between the origin and the angular point of a convex polyhedron, which is a perturbation of the n-dimensional cube and is described by a system of double linear inequalities with an upper triangular matrix of coefficients with units on the main diagonal. Only two faces are subject to investigation, one of which or both contains the vertices closest to the origin. To calculate them, it is sufficient to solve $4n - 4$ linear equations systems and choose among them all the nearest equidistant vertices in polynomial time. The problem of minimizing a quadratic polynomial is *NP*-hard, since an *NP*-hard problem about a vertex covering for an arbitrary graph comes down to it. It follows therefrom that $P = NP$, which is based on the development beyond the limits of integer optimization methods.

Keywords: *NP*-hard problem, sparse matrices, Newton methods, direct multiplication algorithm, the direction of descent, a new mathematical formulation, necessary and sufficient conditions of optimality, minimization pseudo Boolean functions, pseudo Boolean programming, linear programming

Citation: *Computer Research and Modeling*, 2017, vol. 9, no. 5, pp. 679–703 (Russian).

Общая постановка проблемы

Ньютоновские методы безусловной оптимизации «относятся к средствам численного моделирования (моделирования с применением вычислительных машин)» [Гилл и др., 1985] и являются фундаментальными инструментами численного анализа, исследования операций, оптимизации и управления [Поляк, 2006]. Например, большинство наиболее эффективных методов в линейном и нелинейном программировании строятся на их основе [Поляк, 2006].

Общий принцип построения большинства ньютоновских методов можно описать следующим образом [Свириденко, 2015]. Рассмотрим задачу безусловной минимизации:

$$\min_{x \in R^n} F(x), \tag{1}$$

где $F: R^n \to R^1$ — гладкая функция, а R^n обозначает n-мерное евклидово пространство. Пусть h^k, H^k — градиент и гессиан, вычисленные на итерации k в точке x^k процесса безусловной минимизации $F = F(x)$. Тогда, в приведенных обозначениях, общий принцип построения большинства ньютоновских методов безусловной оптимизации с регулировкой шага состоит в следующем. На каждой итерации сначала строится некоторая «связанная» с H^k существенно положительно определенная матрица $\bar{H}^k$, а затем направление спуска p^k вычисляется как решение системы

$$\bar{H}^k p^k = -h^k \tag{2}$$

и определяется новая точка:

$$x^{k+1} = x^k + \alpha_k p^k, \tag{3}$$

где α_k — длина шага, для которого $F^{k+1} < F^k$. При этом процедуру построения $\bar{H}^k$ организуют так, чтобы $\bar{H}^k$ совпадала с исходной матрицей H^k, если последняя сама является положительно определенной, причем выяснение определенности H^k и построение $\bar{H}^k$ осуществляются параллельно в рамках одной процедуры на основе некоторых матричных разложений, которые позволяют выявить знаки собственных чисел H^k и приспособиться для генерации $\bar{H}^k$ [Gill et al., 1974; Гилл и др., 1977; Гилл и др., 1985].

Цели и задачи исследования. Данная работа является продолжением исследований [Свириденко, 2015; Свириденко, 2016; Свириденко, 2017], в которых рассматривалось построение прямого мультипликативного алгоритма расчета элементов факторов Холесского D, U для положительно определенной матрицы H и решения p^* системы линейных уравнений

$$Hp + h = 0 \tag{4}$$

в рамках одной процедуры, где $U \in R^{n \times n}$ — верхняя треугольная матрица с диагональными элементами, равными единице, $D \in R^{n \times n}$ — диагональная матрица со строго положительными элементами на диагонали. Здесь и далее, в случаях, когда это возможно, будем опускать номер итерации k.

Цель данного исследования — положить прямой мультипликативный алгоритм решения системы линейных уравнений в основу:

- расчета направления спуска в ньютоновских методах безусловной оптимизации для задач большой размерности;
- построения новой математической формулировки задачи квадратичного программирования (здесь и далее — КП) и новой формы задания необходимых и достаточных условий оптимальности;

- построения непрерывного аналога задачи минимизации квадратичного полинома от булевых переменных и разработки методов их решения за полиномиальное время.

Инструментальными средствами достижения поставленных целей являются три варианта вычислительной схемы прямого мультипликативного алгоритма расчета элементов D, U модифицированной факторизации Холесского для преобразования

$$Hp = -h \Rightarrow \bar{H}p = -h \tag{5}$$

и решения p^* системы уравнений в рамках одной процедуры.

Таким образом, задачей исследования является построение следующих вариантов вычислительных схем.

- Первый вариант вычислительной схемы представляет собой алгоритм расчета направления спуска p^*, реализацию которого задает техника построения матрицы $\bar{H}$ (обсуждается в § 1). В этом случае выделение объема памяти для построения факторов Холесского не является обязательным требованием. Данная вычислительная схема обсуждается в § 2.
- Второй вариант вычислительной схемы является модификацией первого и определяет построение новой математической формулировки задачи КП, а также новой формы задания необходимых и достаточных условий оптимальности, которые достаточно просты и могут быть использованы для построения методов математического программирования. Представляет собой алгоритм расчета элементов p^*, D, U. В этом случае необходимо выделение объема памяти для задания факторов Холесского. Данная вычислительная схема обсуждается в § 3.
- Третий вариант вычислительной схемы является модификацией второго и определяет построение непрерывного аналога задачи минимизации квадратичного полинома от булевых переменных и новой формы задания необходимых и достаточных условий оптимальности для разработки методов их решения за полиномиальное время. В этом случае требуется новая техника аппроксимации квадратичного многочлена от булевых переменных строго выпуклым квадратичным многочленом от вещественных переменных таким образом, чтобы их значения совпадали на множестве вершин единичного n-мерного куба. Данная вычислительная схема и алгоритмы минимизации полиномов от булевых переменных обсуждаются в § 3.

В работе принято:

- определение разреженной матрицы, которое утверждает, что приписывание матрице свойства разреженности эквивалентно утверждению о существовании алгоритма, который, используя ее разреженность, приводит к сокращению временной и емкостной сложности реализации по сравнению со стандартными алгоритмами [Писсанецки, 1988; Григорьева и др., 2011; Дмитриева, 2014];
- использование формата хранения разреженных матриц, допускающего возможность параллельного выполнения любых матричных или матрично-векторных операций без распаковывания [Свириденко, 2016].

Замечания. В основе построения прямых мультипликативных методов решения системы (4), учитывающих разреженность матриц, представленных в упакованном виде, лежит интеграция техники исключения Гаусса и факторизации Холесского.

В основе увеличения численной устойчивости расчета элементов p^*, D, U лежит интеграция прямого мультипликативного метода решения систем уравнений и техники выбора ведущего элемента в алгоритме линейного программирования (здесь и далее — ЛП) [Хакимова и др., 2010; Свириденко, 2015; Свириденко, 2017].

Прямой мультипликативный алгоритм расчета элементов p^*, D, U можно положить в основу задания направления спуска в ньютоновских методах безусловной оптимизации [Свириденко, 2015] путем интеграции одной из предложенных в [Гилл и др., 1985; Зеленков и др., 2013;

Свириденко, 2015; Свириденко и др., 2016] техник построения существенно положительно определенной матрицы $\bar{H}$. Такой подход позволяет ослабить или снять дополнительные специфические трудности [Черноруцкий, 2011], обусловленные необходимостью решения больших систем уравнений с разреженной матрицей H.

1. Задание направления спуска

Техника построения матрицы $\bar{H}^k$ — один из основных моментов, определяющих «лицо» алгоритма, и поэтому требует детального обсуждения подходов, предложенных в [Гилл и др., 1985; Зеленков и др., 2013; Свириденко, 2015; Свириденко и др., 2016], а также обсуждения подхода к заданию направления спуска, более подходящего для задач большой размерности.

Для выяснения определенности H^k и построения $\bar{H}^k$ наиболее приспособленными являются алгоритмы, использующие факторизацию (разложение) Холесского [Гилл и др., 1985]. Разложение Холесского для симметричной положительно определенной матрицы H^k имеет вид

$$H^k = L^k D^k \left(L^k\right)^T = \left(U^k\right)^T D^k U^k, \tag{1.1}$$

где $U^k = \left(L^k\right)^T$. Если столбцы матрицы L^k с номерами от 1-го по (j – 1)-й известны, то ее j-й столбец определяется по формулам

$$d_{jj}^k = h_{jj}^k - \sum_{s=1}^{j-1} d_{ss}^k \left(l_{js}^k\right)^2, \quad l_{ij}^k = \tfrac{1}{d_{jj}^k}\left(h_{ij}^k - \sum_{s=1}^{j-1} d_{ss}^k l_{js}^k l_{is}^k\right). \tag{1.2}$$

Аналогичные формулы существуют и для построчной организации вычислений. В отличие от гауссовых исключений алгоритм Холесского численно устойчив без каких-либо перестановок. Это свойство определяется соотношением

$$\sum_{i=1}^{n} \left(l_{ij}^k\right)^2 d_{jj}^k = h_{jj}^k, \quad j = 1, \ldots, n, \tag{1.3}$$

между элементами H^k и L^k. В силу него существует априорное ограничение сверху на элементы $r_{ij}^k = l_{ij}^k \sqrt{d_{jj}^k}$: каждый из них не превосходит максимальной величины h_{jj}^k. Соответственно, «лавинообразный рост» элементов r_{ij}^k невозможен, независимо от того, будут ведущие элементы малыми или нет.

Факторизация Холесского применима не для любой симметричной матрицы по следующим причинам. Во-первых, для знаконеопределенной матрицы H^k факторизация Холесского может не существовать. Во-вторых, даже если она и существует, то гарантировать численную устойчивость алгоритма уже нельзя, поскольку никаких априорных ограничений на субдиагональные элементы L^k в рассматриваемом случае не будет.

Подход Гилла и Мюррея. Итак, на основе обычной факторизации Холесского построения $\bar{H}^k$ не получить. Для этого нужно воспользоваться модифицированной факторизацией [Gill et al., 1974; Гилл и др., 1985]. Подход Гилла и Мюррея состоит в том, чтобы строить факторы Холесского, подчиняющиеся двум требованиям: все диагональные элементы D^k должны быть существенно положительными; модули всех элементов треугольного фактора L^k должны быть равномерно ограничены сверху. Точнее говоря, требуется, чтобы для всех $j = 1, 2, \ldots, n$

и некоторых заданных положительных δ и β выполнялись неравенства

$$d_{jj}^k \geq \delta, \quad \left|r_{ij}^k\right| \leq \beta, \quad i > j, \tag{1.4}$$

где r_{ij}^k — введенные для удобства изложения вспомогательные величины, по определению равные $u_{ij}^k \sqrt{d_{jj}^k}$. Гринштадт [Гилл и др., 1977] предложил при реализации алгоритма на конкретной ЭВМ, в которой под запись мантиссы отводится t битов, величину δ вычислять по формуле

$$\delta = \max\left\{ 2^{-t} \left\|H^k\right\|_\infty, 2^{-t} \right\}. \tag{1.5}$$

Для сохранения численной устойчивости процедуры построения H^k, а также для совпадения H^k и $\bar{H}^k$ в случае положительно определенной H^k целесообразно величину β вычислять по формуле [Гилл и др., 1985]

$$\beta^2 = \max\left\{ \varsigma, \xi \middle/ \sqrt{n^2 - 1}, \varepsilon_M \right\}. \tag{1.6}$$

Здесь ξ — максимальный модуль недиагонального элемента H^k, ς — значение максимального из диагональных элементов H^k, ε_M — машинная точность, которая вводится для обеспечения устойчивости вычислений, когда норма H^k очень мала [Гилл и др., 1977]. При этом процедура расчета модифицированных факторов L^k, D^k фактически представляет собой обычный алгоритм факторизации Холесского с попутным увеличением (по мере необходимости) диагонали исходной матрицы. Матрицы L^k и D^k, полученные по окончании описанной процедуры, будут факторами Холесского для положительно определенной матрицы $\bar{H}^k$, связанной с H^k следующим образом:

$$L^k D^k \left(L^k\right)^T = H^k + E^k = \bar{H}^k, \tag{1.7}$$

где E^k — неотрицательная диагональная матрица, *j*-й элемент которой равен e_{jj}^k. Таким образом, положительно определенная матрица $\bar{H}^k$ может отличаться от исходной матрицы H^k только диагональными элементами. Направление спуска p^k вычисляется как решение системы $\bar{H}^k p^k = -h^k$, причем $p^k = -\left(\bar{H}^k\right)^{-1} h^k$ определяют последовательным решением двух систем линейных уравнений с треугольными матрицами:

$$L^k y^k = -h^k, \quad D^k \left(L^k\right)^T p^k = y^k. \tag{1.8}$$

Замечание. Построение модифицированного разложения Холесского требует выполнения около $\frac{1}{6}n^3$ арифметических операций, примерно столько же, сколько требует обычное разложение для положительно определенной матрицы.

Увеличение эффективности подхода Гилла и Мюррея. Увеличение эффективности методов, основанных на факторизации Холесского, связано с ослаблением или снятием нерешенных проблем, возникающих при построении методов безусловной минимизации гладкой функции [Гилл и др., 1977; Гилл и др., 1985].

- Выбор масштаба при спуске аналогичен по смыслу требованию подчинения длины шага α_k вдоль выбранного направления спуска p^k неравенствам $0 < \delta \le \left\| \alpha_k p^k \right\| \le \Delta$, где δ есть минимальное расстояние между x^k и x^{k+1}, а Δ — оценка сверху расстояния от x^k до точки минимума $F(x)$ вдоль p^k [Гилл и др., 1985]. Корректные значения параметров δ и Δ, вообще говоря, зависят от x^k и $F(x)$, а появление у матрицы H^k отрицательных собственных значений неизбежно усложняет проблему выбора масштабов при спуске. И, независимо от того, смотреть ли на эту проблему как на проблему нормировки направления спуска или как на проблему выбора начальной оценки шага, универсального решения не видно.
- Ньютоновские методы распадаются на два обширных класса, это «методы с регулировкой шага» и «методы доверительной окрестности». Наряду со сходными чертами между ними имеются и значительные различия, связанные с характером использования информации о вторых производных. Поэтому в настоящее время нет общепринятого определения «метода Ньютона» для расчета направления спуска при знаконеопределенной матрице вторых производных, поскольку среди специалистов нет согласия относительно того, как использовать локальную квадратичную аппроксимацию гладкой функции в этом случае [Гилл и др., 1985]. Такая ситуация сохраняется и поныне [Черноруцкий, 2013].
- Все алгоритмы (не считая некоторых методов прямого поиска) основаны на квадратичной аппроксимации минимизируемой функции $F(x)$, поэтому, если бы удалось построить такой метод аппроксимации, который оказался бы лучше квадратичного, это, по-видимому, могло бы привести к существенной переоценке ценностей [Гилл и др., 1977].

Решение проблемы масштабирования шагов при спуске опишем следующим образом [Зеленков и др., 2013]. Направление спуска p^k в конечном счете вычисляется как решение системы $\left(L^k\right)^T p^k = u^k$, где $u^k = \left(D^k\right)^{-1} y^k$, а y^k есть решение системы $L^k y^k = -h^k$. Величина элементов $\left|u_j^k\right|$ зависит от способа задания направления спуска. Поэтому, следуя подходу Гилла и Мюррея, для численной устойчивости расчета элементов D^k, L^k, u^k достаточно потребовать изменения способа задания p^k так, чтобы для всех $j = 1, 2, \ldots, n$ и некоторых заданных положительных δ, β и ω выполнялись неравенства

$$d_{jj}^k \ge \delta, \quad \left|r_{ij}^k\right| \le \beta, \quad i > j, \quad \left|u_j^k\right| \le \omega. \tag{1.9}$$

Это может стать ключом к решению проблемы масштабирования шагов при спуске [Гилл и др., 1985], но подход линейной алгебры к вычислению направления спуска p^k исключает расчет элементов вектора u^k по ходу построения модифицированного разложения Холесского [Парлетт, 1983]. Нужен альтернативный подход, который опишем так. Равенства $H^k = \left(U^k\right)^T D^k U^k$ достаточно для определения элементов матриц D^k, U^k. В скалярной форме это равенство выглядит следующим образом:

$$h_{ij}^k = \sum_{\mu=1}^{i} u_{\mu i}^k u_{\mu j}^k d_{\mu\mu}^k; \tag{1.10}$$

отсюда, полагая, что $u_{ii}^k = 1$, построим соотношения

$$d_{ii}^k = h_{ii}^k - \sum_{s=1}^{i-1} \left(u_{si}^k\right)^2 d_{ss}^k, \quad u_{ij}^k = \left(h_{ij}^k - \sum_{s=1}^{i-1} u_{si}^k u_{sj}^k d_{ss}^k\right) \Big/ d_{ii}^k, \quad j > i, \tag{1.11}$$

для расчета элементов D^k, U^k. Построение факторов Холесского математически эквивалентно применению метода исключения Гаусса к системе уравнений $H^k p^k = -h^k$ в прямом порядке, при этом техника исключения Гаусса позволяет получить соотношение

$$u_i^k = -\left(h_i^k - \sum_{s=1}^{i-1} u_s^k u_{si}^k d_{ss}^k\right) \Big/ d_{ii}^k \tag{1.12}$$

для расчета элементов вектора u^k и систему уравнений $U^k p^k = u^k$ для вычисления направления спуска p^k.

Интеграция техники исключения Гаусса и факторизации Холесского порождает множество численно устойчивых способов задания направления спуска p^k, в основе которых лежит требование, чтобы на очередном шаге вычислений коэффициентов факторов Холесского соответствующий диагональный элемент матрицы D^k и соответствующий элемент вектора u^k сначала рассчитывались по вычисленным ранее значениям этих коэффициентов. Затем диагональный элемент D^k увеличивается настолько, насколько необходимо, чтобы все диагональные элементы D^k были существенно положительными, модули всех элементов U^k, u^k были равномерно ограничены сверху.

Решение проблемы аппроксимации неквадратичными функциями и интеграции с методом доверительной окрестности опишем следующим образом. В процессе построения D^k, U^k, u^k вычислить элементы $g_i^k = d_{ii}^k / l_i^k$ (если $d_{ii}^k > l_i^k \gamma$, взять $g_i^k = l_i^k / \gamma$), где $l_i^k = \max\left\{\delta, \left|d_{ii}^k\right|\right\}$. Вычислить $\gamma^k = \max\limits_{1 \le i \le n} g_i^k$ и решить систему $\frac{1}{\gamma^k} \overline{H}^k p^k = -h^k$ для определения направления спуска p^k. Здесь параметр γ задается пользователем, число γ^k ($\gamma^k \ge 1$ по построению) характеризует степень однородности и задает простейшую форму аппроксимации, отличную от квадратичной. Отсюда следует, что

$$\overline{H}^k p^k = -\gamma^k h^k = -h^k - \left(\gamma^k - 1\right) h^k = -h^k - G^k p^k \;\Rightarrow\; \left(\overline{H}^k + G^k\right) p^k = -h^k, \tag{1.13}$$

где G^k есть диагональная матрица с элементами $g_{ii}^k = \left(\gamma^k - 1\right) h_i^k / p_i^k$ на диагонали, что означает сдвиг на g_{ii}^k всех i-х собственных значений матрицы $\overline{H}^k$.

В методах с регулировкой шага матрицу H^k модифицируют так, чтобы изменения не затрагивали подпространства, натянутого на ее собственные векторы с положительными собственными значениями [Гилл и др., 1985]. Если же замена осуществляется в методе доверительной окрестности, то она отражается на всех векторах, так как результатом замены H^k на матрицу $\overline{H}^k = H^k + \lambda^k I$ будет сдвиг на λ^k всех собственных значений матрицы H^k [Гилл и др., 1985]. Таким образом, интеграция техники исключения Гаусса и факторизации Холесского определяет и интеграцию с методом доверительной окрестности.

Исследования [Зеленков и др., 2013] продолжены в [Свириденко, 2015]. Доказано, что интеграция техники исключения Гаусса и факторизации Холесского определяет и подход к дальнейшему уменьшению значения $\left\|E^k\right\|$, и подход к заданию направления спуска, более подходящий для задач большой размерности. Изучены два варианта дальнейшего уменьшения значения $\left\|E^k\right\|$. Варианты определяются стратегией выбора ведущего элемента: частичного (здесь и далее — СЧВ) и полного (здесь и далее — СПВ). В основе СЧВ лежит интеграция техники алгоритма ЛП для выбора ведущей строки, а в основе СПВ — для выбора ведущей строки и ведущего столбца.

Интеграция СЧВ в вычислительную схему расчета элементов D^k, U^k, u^k [Зеленков и др., 2013] определяет на шаге 2 следующий вариант дальнейшего уменьшения значения $\|E^k\|$: найти индекс q такой, что $|c_{qq}^k|+|c_q^k| = \max_{i\le j\le n}\left(|c_{jj}^k|+|c_j^k|\right)$, и поменять местами все данные, отвечающие строкам H^k, h^k с номерами q и i, а затем проделать то же самое с данными, отвечающими ее q-му и i-му столбцам H^k. Здесь c_{ii}^k, c_i^k — вспомогательные величины, определяемые следующим образом:

$$c_{ii}^k = h_{ii}^k - \sum_{s=1}^{i-1}(u_{si}^k)^2 d_{ss}^k, \quad c_i^k = -\left(h_i^k - \sum_{s=1}^{i-1} u_s^k u_{si}^k d_{ss}^k\right). \tag{1.14}$$

Такая стратегия дает возможность увеличения численной устойчивости расчета элементов и остается работоспособной и в том случае, когда $\|h^k\|$ не равна нулю, но очень мала.

Интеграция СПВ для дальнейшего уменьшения значения $\|E^k\|$ влечет изменение порядка пересчета вспомогательных величин, потерю фактора U^k и излишнее усложнение вычислительной схемы. Нужен альтернативный подход, который опишем следующим образом. Интеграция правил построения $\bar{H}^k$ [Зеленков и др., 2013] в вычислительную схему мультипликативного алгоритма ЛП, предложенного в [Хакимова и др., 2010; Свириденко, 2015] как метода решения системы $H^k p^k = -h^k$, допускает модификацию, которая позволяет выявить знаки собственных чисел H^k и приспособиться для генерации $\bar{H}^k$. При этом процедуру построения $\bar{H}^k$ можно организовать так, чтобы $\bar{H}^k$ совпадала с исходной матрицей H^k, если последняя сама является положительно определенной, причем выяснение определенности H^k и вычисление направление спуска p^k осуществляется параллельно в рамках одной процедуры. Такой подход к заданию направления спуска является более подходящим для задач большой размерности, однако вычислительная схема подхода не рассматривалась.

Исследования [Зеленков и др., 2013; Свириденко, 2015] развиты и продолжены в [Свириденко и др., 2016]. Изучены взаимосвязь и реализация квазиньютоновских и ньютоновских методов безусловной оптимизации, основанных на факторизации Холесского, с регулировкой шага и с конечно-разностной аппроксимацией первых и вторых производных. Рассмотрен подход к заданию квазиньютоновского поиска, позволяющий ослабить или снять следующие недостатки квазиньютоновских методов.

- Большинство вариантов квазиньютоновских методов (например, одна из наиболее эффективных схем Бройдена–Флетчера–Гольдфарба–Шенно) при минимизации сильно выпуклых квадратичных функционалов приводят к одной и той же траектории спуска, вырождаясь в хорошо изученные методы сопряженных градиентов. В то же время известна особенность методов сопряженных градиентов, существенно ограничивающая область их эффективного применения. Она заключается в понижении скорости сходимости для плохо обусловленных (жестких) задач оптимизации.
- Предположение о невыпуклости вносит дополнительные трудности: в этих условиях метод сопряженных градиентов по характеристикам сходимости эквивалентен градиентному методу наискорейшего спуска со всеми вытекающими отсюда последствиями.
- Кроме отмеченных дефектов, общих для методов сопряженных градиентов и квазиньютоновских методов, последние имеют дополнительные недостатки, связанные с проблемой потери положительной определенности квазиньютоновских матриц из-за накопле-

ния вычислительных погрешностей в рекуррентных процедурах аппроксимации матриц вторых производных.

В основе подхода лежит модифицированное разложение Холесского квазиньютоновской матрицы. Такой подход определяет для квазиньютоновских методов и решение проблемы масштабирования шагов при спуске, и аппроксимацию неквадратичными функциями, и интеграцию с методом доверительной окрестности. Кроме этого, фактическое значение нормы $\left\|E^k\right\|$ можно дополнительно уменьшить, если использовать симметричные перестановки строк и столбцов квазиньютоновской матрицы H^k. Результаты численных исследований [Свириденко и др., 2016] показали, что предложенный квазиньютоновский метод с конечно-разностной аппроксимацией первых производных практически не уступает в точности решения тестовых задач ньютоновскому методу с аналитическим вычислением первых и вторых производных.

Замечания. Подход к вычислению длины шага α_k предложен в [Хакимова и др., 2010; Хакимова, Дикусар и др., 2010]. В основе лежит интеграция подходов Пшеничного [Пшеничный и др., 1975], Полака [Полак, 1974], Карманова [Карманов, 1975], Гилла, Мюррея и Райт [Гилл и др., 1985].

Подход к модификации критериев останова, разработанных Гиллом, Мюрреем и Райт [Гилл и др., 1985], предложен в [Свириденко, 2015].

Подход к оцениванию конечно-разностных интервалов и подход к конечно-разностной аппроксимации первых и вторых производных предложены в [Зеленков и др., 2013]. В основе лежит интеграция подходов Полака [Полак, 1974], Гилла, Мюррея и Райт [Гилл и др., 1985].

2. Задание направления спуска для задач большой размерности

Факторы Холесского для матриц H^k позволяют получать оценки их чисел обусловленности. Например, если максимальный и минимальный элементы фактора D^k равны $d_{\max}^k$ и $d_{\min}^k$, то $cond\left(H^k\right) \geq d_{\min}^k / d_{\max}^k$ [Гилл и др., 1985]. Поэтому предлагается выделить объем памяти для построения D^k с целью задания элементов целевой функции задачи ЛП, чтобы учесть влияние на результат «малых» по модулю собственных чисел матрицы H^k.

На промежуточных итерациях такие оценки требуются для выявления ситуаций, когда матрица H^k обусловлена настолько плохо, что в численном решении p^k системы

$$H^k p^k = -h^k, \tag{2.1}$$

скорее всего, не окажется ни одной правильной цифры. При этом всегда можно перейти к проведению «альтернативной процедуры спуска», как и в случаях, когда регулярная процедура поиска направления спуска не дает существенных результатов [Свириденко, 2016].

Первый вариант вычислительной схемы является основой для разработки второго и третьего вариантов, поэтому необходимо предусмотреть возможность упрощения его реализации и описания вычислительной схемы расчета элементов p^k, D^k, U^k.

Следует отметить, что для ньютоновских методов доказана возможность их вырождения [Черноруцкий, 2013], поэтому предлагается вычислять количество несовместных (n_d) и вырожденных (n_i) систем линейных уравнений (2.1) для исследования частоты их появления при решении задач большой размерности.

Ниже приводится детальное описание всех операций, выполняемых по ходу расчета элементов p^k, D^k, U^k прямым мультипликативным алгоритмом с перестановками. Если в процессе вычислений элементы становятся меньше по абсолютной величине так называемого критического значения ε_0, то их предлагается приравнивать к нулю.

Первый вариант вычислительной схемы прямого мультипликативного метода расчета элементов p^k, D^k, U^k

Шаг 0 (инициализация)**.** Присвоить номеру итерации γ значение 1. Положить

$$p^0 = \begin{pmatrix} 0 & 0 & \cdots & 0 \end{pmatrix}^T,$$

вычислить:

$$\beta^2 = \max\{\varsigma, \xi/\nu, \varepsilon_M\},$$

где $\nu = \max\left\{1, \sqrt{n^2-1}\right\}$, а числа ς и ξ суть максимальные значения модулей диагонального и недиагонального элементов H^k. Если $k = 0$, то положить

$$c^0 = \begin{pmatrix} |h_{11}| & |h_{22}| & \cdots & |h_{nn}| \end{pmatrix};$$

иначе положить

$$c^0 = \begin{pmatrix} d_{11}^{k-1} & d_{22}^{k-1} & \cdots & d_{nn}^{k-1} \end{pmatrix}.$$

Замечание. Для построения положительно определенной матрицы принята техника, предложенная Гиллом и Мюрреем [Гилл и др., 1985] с целью упрощения описания вычислительной схемы расчета элементов p^k, D^k, U^k.

Шаг 1 (расчет элементов h_γ^γ, $h_{\gamma j}^\gamma$ уравнения связи)**.** Вычислить номер r ведущей строки:

$$\theta_r = \max_i \left| h_i + h_{i\bullet} p^{\gamma-1} \right| \left(i = \gamma, \ldots, n \right),$$

где $h_{i\bullet} = \begin{pmatrix} h_{i1} & h_{i2} & \cdots & h_{in} \end{pmatrix}$. Поменять местами элементы r-й и γ-й строк матриц H^k, h^k. Если $\gamma = 1$, то вычислить:

$$h_\gamma^\gamma + \begin{pmatrix} h_{\gamma\gamma}^\gamma & h_{\gamma\,\gamma+1}^\gamma & \cdots & h_{\gamma n}^\gamma \end{pmatrix} \begin{pmatrix} p_\gamma & p_{\gamma+1} & \cdots & p_n \end{pmatrix}^T = 0 \Rightarrow$$
$$\Rightarrow h_\gamma^\gamma = h_\gamma,\ \begin{pmatrix} h_{\gamma\gamma}^\gamma & h_{\gamma\,\gamma+1}^\gamma & \cdots & h_{\gamma n}^\gamma \end{pmatrix} = \begin{pmatrix} h_{\gamma\gamma} & h_{\gamma\,\gamma+1} & \cdots & h_{\gamma n} \end{pmatrix}$$

и перейти к шагу 2; иначе вычислить:

$$h_\gamma^\gamma + \begin{pmatrix} h_{\gamma\gamma}^\gamma & h_{\gamma\,\gamma+1}^\gamma & \cdots & h_{\gamma n}^\gamma \end{pmatrix} \begin{pmatrix} p_\gamma & p_{\gamma+1} & \cdots & p_n \end{pmatrix}^T = 0 \Rightarrow$$

$$\Rightarrow h_\gamma^\gamma = h_\gamma + \sum_{j=1}^{\gamma-1} h_{\gamma j} p_j^j,\ \begin{pmatrix} h_{\gamma\gamma}^\gamma & h_{\gamma\,\gamma+1}^\gamma & \cdots & h_{\gamma n}^\gamma \end{pmatrix} = h_{\gamma\bullet} \prod_{i=1}^{\gamma-1} E_{1_i}^i = h_{\gamma\bullet} \begin{pmatrix} e_{1_{\gamma-1}\gamma}^1 & \cdots & e_{1_{\gamma-1}n}^1 \\ & \vdots & \\ e_{1_{\gamma-1}\gamma}^{\gamma-1} & \cdots & e_{1_{\gamma-1}n}^{\gamma-1} \\ 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}.$$

Шаг 2 (анализ элементов h_γ^γ, $h_{\gamma j}^\gamma$ уравнения связи)**.** Если

$$\left|h_\gamma^\gamma\right| > \varepsilon_0, \ \left|h_{\gamma\, j}^\gamma\right| \le \varepsilon_0 \left(j = \gamma, \ldots, n\right)$$

(проверка несовместности), то положить

$$h_{\gamma j}^\gamma = \delta \ \ (j = \gamma, \ldots, n), \quad n_d = n_d + 1;$$

иначе, если

$$\left|h_\gamma^\gamma\right| \le \varepsilon_0, \ \left|h_{\gamma\, j}^\gamma\right| \le \varepsilon_0 \left(j = \gamma, \ldots, n\right)$$

(проверка вырожденности), положить

$$h_{\gamma j}^\gamma = \delta \ \ (j = \gamma, \ldots, n), \quad n_i = n_i + 1.$$

Шаг 3 (расчет γ-го диагонального элемента фактора D и главной строки $e_{1_\gamma \bullet}^\gamma$ мультипликатора $E_{1_\gamma}^\gamma$). Вычислить номер q ведущего столбца:

$$\Theta_q = \min_j \left|c_j^{\gamma-1} \big/ h_{r\, j}\right| \left(j = \gamma, \ldots, n\right).$$

Поменять местами элементы q-го и γ-го столбцов матриц H^k, c^0, запомнить порядок неизвестных.

Найти

$$\theta_\gamma = \max_{\gamma+1 \le i \le n} \left|h_{\gamma\, i}^\gamma\right|,$$

вычислить:

$$d_{\gamma\,\gamma} = \max\left\{\delta, \left|h_{\gamma\,\gamma}^\gamma\right|, \theta_\gamma^{\,2} \big/ \beta^2\right\}$$

и пересчитать элементы уравнения связи:

$$p_\gamma = p_\gamma^\gamma + e_{1_\gamma \bullet}^\gamma \begin{pmatrix} p_{\gamma+1} & p_{\gamma+2} & \cdots & p_n \end{pmatrix} \Rightarrow$$
$$\Rightarrow \ p_\gamma^\gamma = -h_\gamma^\gamma \big/ d_{\gamma\,\gamma}, \ e_{1_\gamma \bullet}^\gamma = -\begin{pmatrix} h_{\gamma\,\gamma+1}^\gamma & h_{\gamma\,\gamma+2}^\gamma & \cdots & h_{\gamma\, n}^\gamma \end{pmatrix} \big/ d_{\gamma\,\gamma},$$
$$E_{1_\gamma}^\gamma = \begin{pmatrix} e_{1_\gamma\,\gamma+1}^\gamma & e_{1_\gamma\,\gamma+2}^\gamma & \cdots & e_{1_\gamma\, n}^\gamma \\ 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}.$$

Замечание. Мультипликатор $E_{1_\gamma}^\gamma$ — это матрица размера $\left(n-\gamma\right)\times\left(n-\gamma-1\right)$, первая строка которой произвольная, обозначим ее через $e_{1_\gamma \bullet}^\gamma = \begin{pmatrix} e_{1_\gamma\,\gamma+1}^\gamma & e_{1_\gamma\,\gamma+2}^\gamma & \cdots & e_{1_\gamma\, n}^\gamma \end{pmatrix}$, а остальные — строки единичной матрицы размера $\left(n-\gamma-1\right)\times\left(n-\gamma-1\right)$. Элементарная матрица — мультипликатор $E_{r_q}^\gamma$ — с точностью до знака транспонирования совпадает с мультипликатором в ЛП (для ограничений-неравенств) [Хакимова и др., 2010; Свириденко, 2015], за исключением удаления q-го столбца, все элементы которого равны нулю.

Шаг 4 (расчет элементов p^{γ}, c^{γ}, $e^{i}_{1_{\gamma}\bullet}$)**.** Пересчитать $c^{\gamma-1}$:

$$c^{\gamma} = \begin{pmatrix} c^{\gamma}_{\gamma+1} & c^{\gamma}_{\gamma+2} & \cdots & c^{\gamma}_{n} \end{pmatrix} = c^{\gamma-1} E^{\gamma}_{1_{\gamma}} .$$

Если $\gamma = 1$, то вычислить решение уравнения $h^{k}_{\gamma\bullet} p^{k} = -h^{k}_{\gamma}$:

$$p^{\gamma} = \begin{pmatrix} p^{\gamma}_{1} & 0 & 0 & \cdots & 0 \end{pmatrix}^{T},$$

перейти к шагу 5; иначе вычислить решение системы уравнений $h^{k}_{i\bullet} p^{k} = -h^{k}_{i} \left(i = 1, \ldots, \gamma\right)$:

$$p_i = p^{\gamma}_{i} + e^{i}_{1_{\gamma}\bullet} \begin{pmatrix} p_{\gamma+1} & p_{\gamma+2} & \cdots & p_n \end{pmatrix}^{T},\ p^{\gamma}_{i} = p^{i}_{i} + p^{\gamma}_{\gamma} e^{i}_{1_{\gamma-1}\,\gamma} \quad \left(i = 1, \ldots, \gamma - 1\right) \Rightarrow$$
$$\Rightarrow p^{\gamma} = \begin{pmatrix} p^{\gamma}_{1} & p^{\gamma}_{2} & \cdots & p^{\gamma}_{\gamma} & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Пересчитать $e^{i}_{1_{k-1}\bullet}$ (минимизация заполнения главных строк мультипликаторов):

$$e^{i}_{1_{\gamma}\bullet} = e^{i}_{1_{\gamma-1}\bullet} E^{\gamma}_{1_{\gamma}} \left(i = 1, \ldots, \gamma - 1\right).$$

Шаг 5 (расчет *n*-го диагонального элемента фактора *D* и направления спуска p^{k}). Положить $\gamma = \gamma + 1$. Если $\gamma \neq n$, то перейти к шагу 1; иначе вычислить:

$$h^{n}_{n} + h^{n}_{n\,n} p_n = 0 \Rightarrow$$

$$\Rightarrow h^{n}_{n} = h_n + h_{n\,1} p^{1}_{1} + h_{n\,2} p^{2}_{2} + \cdots + h_{n\,n-1} p^{n-1}_{n-1},\ h^{n}_{n\,n} = h_{n\bullet} \prod_{i=1}^{n-1} E^{i}_{1_i} = h_{n\bullet} \begin{pmatrix} e^{1}_{1_{n-1}\,n} \\ e^{2}_{1_{n-1}\,n} \\ \vdots \\ e^{n-1}_{1_{n-1}\,n} \\ 1 \end{pmatrix},$$

$$d_{n\,n} = \max\left\{\delta, \left|h^{n}_{n\,n}\right|\right\},$$

$$p^{n} = p^{k} = \begin{pmatrix} p^{n-1}_{1} - e^{1}_{1_{n-1}\,n}\, h^{n}_{n} / d_{n\,n} & p^{n-1}_{2} - e^{2}_{1_{n-1}\,n}\, h^{n}_{n} / d_{n\,n} & \cdots & p^{n-1}_{n-1} - e^{n-1}_{1_{n-1}\,n}\, h^{n}_{n} / d_{n\,n} & -h^{n}_{n} / d_{n\,n} \end{pmatrix}.$$

Замечания. Матрица H^{k} симметрична, поэтому в памяти хранятся только ее диагональ и верхний треугольник. Элементы главных строк мультипликаторов строятся непосредственно в отдельно отведенной области памяти, а не записываются на место элементов H^{k}, поскольку строки являются менее разреженными, чем матрица H^{k}, и для их хранения сначала требуется последовательное выделение объема памяти до $O\left(\frac{1}{4} n^{2}\right)$, а затем ее последовательное освобождение. Последнее означает, что выделение объема памяти с целю построения U^{k} не требуется, в то время как подход Писсанецки [Писсанецки, 1988] к практической факторизации Холесского требует выделения объема памяти $O\left(\frac{1}{2} n^{2}\right)$ для построения верхнего треугольного множителя. Здесь расчеты объема памяти проводились для случая плотных симметричных матриц.

3. Моделирование

Прямой мультипликативный алгоритм расчета элементов модифицированной факторизации Холесского обманчиво прост. Однако имеются два аспекта, которые значительно глубже, чем простая техника построения существенно положительно определенной матрицы, и которые (вместе с самим алгоритмом) обсуждаются в этом разделе. Перечислим эти аспекты.

Построение новой математической формулировки задачи КП и новой формы задания необходимых и достаточных условий оптимальности. Интеграция техники исключения Гаусса и факторизации Холесского определяет и новую математическую формулировку задачи КП:

$$\min_{p\in R^n} f(p) = \frac{1}{2} p^T H p + p^T h, \tag{3.1}$$

которую опишем следующим образом. В [Хакимова и др., 2003] доказано, что (3.1) сводится к эквивалентной задаче:

$$\min_{p,x\in R^n} f(p) = -\tfrac{1}{2}\sum_{j=1}^{n}\left(u_j\right)^2 \Big/ d_{jj} + \tfrac{1}{2} x^T D x \,\Big|\, u + Up = Dx. \tag{3.2}$$

Очевидно,

$$\min_{p,x\in R^n} f(p) \Rightarrow \min_{p,x\in R^n} x^T D x \,\Big|\, p = p^* + Px. \tag{3.3}$$

При этом выяснение определенности H и построение $\bar{H}$ осуществляются параллельно в рамках одной из процедур расчета элементов p^*, D, U, описания которых можно найти в [Хакимова и др., 2003; Зеленков и др., 2013; Свириденко, 2015; Свириденко и др., 2016]. Это означает, что (3.2) и (3.3) — это еще одна математическая формулировка рассмотренной ранее задачи задания направления спуска в ньютоновских методах безусловной оптимизации, которая достаточно проста и может быть использована для построения новых математических формулировок задач математического программирования и новых форм задания необходимых и достаточных условий оптимальности.

Ниже приводится детальное описание всех операций, выполняемых по ходу расчета элементов p^*, D, U прямым мультипликативным алгоритмом с перестановками. Если в процессе вычислений элементы становятся меньше по абсолютной величине так называемого критического значения ε_0, то их предлагается приравнивать нулю.

Второй вариант вычислительной схемы прямого мультипликативного метода расчета элементов p^*, D, U

Шаг 0 (инициализация)**.** Присвоить номеру итерации γ значение 1. Положить:

$$p^0 = \begin{pmatrix} 0 & 0 & \cdots & 0 \end{pmatrix}^T,$$

вычислить:

$$\beta^2 = \max\left\{\varsigma, \xi/\nu, \varepsilon_M\right\},$$

где $\nu = \max\left\{1, \sqrt{n^2-1}\right\}$, а числа ς и ξ суть максимальные значения модулей диагонального и недиагонального элементов H. Положить

$$c^0 = \left(\left|h_{11}\right| \quad \left|h_{22}\right| \quad \cdots \quad \left|h_{nn}\right|\right).$$

Шаг 1 (расчет элементов h_γ^γ, $h_{\gamma j}^\gamma$ уравнения связи)**.** Вычислить номер r ведущей строки:

$$\theta_r = \max_i \left|h_i + h_{i\bullet} p^{\gamma-1}\right| \quad (i = \gamma, \ldots, n),$$

поменять местами элементы r-й и γ-й строк матриц H, h. Если $\gamma = 1$, то вычислить:

$$h_\gamma^\gamma + \begin{pmatrix} h_{\gamma\,\gamma}^\gamma & h_{\gamma\,\gamma+1}^\gamma & \cdots & h_{\gamma\,n}^\gamma \end{pmatrix} \begin{pmatrix} p_\gamma & p_{\gamma+1} & \cdots & p_n \end{pmatrix}^T = 0 \Rightarrow$$
$$\Rightarrow h_\gamma^\gamma = h_\gamma,\ \begin{pmatrix} h_{\gamma\,\gamma}^\gamma & h_{\gamma\,\gamma+1}^\gamma & \cdots & h_{\gamma\,n}^\gamma \end{pmatrix} = \begin{pmatrix} h_{\gamma\,\gamma} & h_{\gamma\,\gamma+1} & \cdots & h_{\gamma\,n} \end{pmatrix},$$

перейти к шагу 2; иначе вычислить:

$$h_\gamma^\gamma + \begin{pmatrix} h_{\gamma\,\gamma}^\gamma & h_{\gamma\,\gamma+1}^\gamma & \cdots & h_{\gamma\,n}^\gamma \end{pmatrix} \begin{pmatrix} p_\gamma & p_{\gamma+1} & \cdots & p_n \end{pmatrix}^T = 0 \Rightarrow$$
$$\Rightarrow h_\gamma^\gamma = h_\gamma + \sum_{j=1}^{\gamma-1} h_{\gamma\,j}\, p_j^j,\ \begin{pmatrix} h_{\gamma\,\gamma}^\gamma & h_{\gamma\,\gamma+1}^\gamma & \cdots & h_{\gamma\,n}^\gamma \end{pmatrix} = h_{\gamma\bullet} \prod_{i=1}^{\gamma-1} E_{1_i}^i.$$

Шаг 2 (анализ элементов h_γ^γ, $h_{\gamma\,j}^\gamma$ уравнения связи)**.** Если

$$\left|h_\gamma^\gamma\right| > \varepsilon_0,\ \left|h_{\gamma\,j}^\gamma\right| \le \varepsilon_0 \left(j = \gamma, \ldots, n\right)$$

(проверка несовместности), то положить

$$h_{\gamma j}^\gamma = \delta\ \ (j = \gamma, \ldots, n), \quad n_d = n_d + 1;$$

иначе, если

$$\left|h_\gamma^\gamma\right| \le \varepsilon_0,\ \left|h_{\gamma\,j}^\gamma\right| \le \varepsilon_0 \left(j = \gamma, \ldots, n\right)$$

(проверка вырожденности), положить

$$h_{\gamma j}^\gamma = \delta\ \ (j = \gamma, \ldots, n), \quad n_i = n_i + 1.$$

Шаг 3 (расчет γ-го диагонального элемента фактора D, главной строки $e_{1_\gamma\bullet}^\gamma$ мультипликатора $E_{1_\gamma}^\gamma$ и γ-й строки верхнего треугольного фактора U)**.** Вычислить номер q ведущего столбца:

$$\Theta_q = \min_j \left| c_j^{\gamma-1} \big/ h_{r\,j} \right| \left(j = \gamma, \ldots, n\right).$$

Поменять местами элементы q-го и γ-го столбцов матриц H, c^0, запомнить порядок неизвестных.

Найти

$$\theta_\gamma = \max_{\gamma+1 \le i \le n} \left|h_{\gamma i}^\gamma\right|,$$

вычислить:

$$d_{\gamma\,\gamma} = \max\left\{\delta, \left|h_{\gamma\,\gamma}^\gamma\right|, \theta_\gamma^{\,2} \big/ \beta^2\right\}$$

и пересчитать элементы уравнения связи:

$$p_\gamma = p_\gamma^\gamma + e_{1_\gamma\bullet}^\gamma \begin{pmatrix} p_{\gamma+1} & p_{\gamma+2} & \cdots & p_n \end{pmatrix} \Rightarrow$$
$$\Rightarrow p_\gamma^\gamma = -h_\gamma^\gamma \big/ d_{\gamma\,\gamma},\ e_{1_\gamma\bullet}^\gamma = -\begin{pmatrix} h_{\gamma\,\gamma+1}^\gamma & h_{\gamma\,\gamma+2}^\gamma & \cdots & h_{\gamma\,n}^\gamma \end{pmatrix} \big/ d_{\gamma\,\gamma},$$

$$E_{1_\gamma}^\gamma = \begin{pmatrix} e_{1_\gamma\,\gamma+1}^\gamma & e_{1_\gamma\,\gamma+2}^\gamma & \cdots & e_{1_\gamma\,n}^\gamma \\ 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}.$$

Положить

$$u_{\gamma\bullet}=\begin{pmatrix}0 & 0 & \cdots & 0 & 1 & -e^{\gamma}_{1_\gamma\,\gamma+1} & -e^{\gamma}_{1_\gamma\,\gamma+2} & \cdots & -e^{\gamma}_{1_\gamma\,n}\end{pmatrix}.$$

Шаг 4 (расчет элементов p^γ, c^γ, $e^i_{1_\gamma\bullet}$). Пересчитать $c^{\gamma-1}$:

$$c^\gamma=\begin{pmatrix}c^\gamma_{\gamma+1} & c^\gamma_{\gamma+2} & \cdots & c^\gamma_n\end{pmatrix}=c^{\gamma-1}E^\gamma_{1_\gamma}.$$

Если $\gamma=1,$ то вычислить решение уравнения $h_{\gamma\bullet}p=-h_\gamma$:

$$p^\gamma=\begin{pmatrix}p^\gamma_1 & 0 & 0 & \cdots & 0\end{pmatrix}^T,$$

перейти к шагу 5; иначе вычислить решение системы уравнений $h_{i\bullet}p=-h_i\,(i=1,\ldots,\gamma)$:

$$p_i=p^\gamma_i+e^i_{1_\gamma\bullet}\begin{pmatrix}p_{\gamma+1} & p_{\gamma+2} & \cdots & p_n\end{pmatrix}^T,\ p^\gamma_i=p^i_i+p^\gamma_\gamma e^i_{1_{\gamma-1}\,\gamma}\quad(i=1,\ldots,\gamma-1)\Rightarrow$$
$$\Rightarrow p^\gamma=\begin{pmatrix}p^\gamma_1 & p^\gamma_2 & \cdots & p^\gamma_\gamma & 0 & 0 & \cdots & 0\end{pmatrix}.$$

Пересчитать $e^i_{1_{k-1}\bullet}$ (минимизация заполнения главных строк мультипликаторов):

$$e^i_{1_\gamma\bullet}=e^i_{1_{\gamma-1}\bullet}E^\gamma_{1_\gamma}\ (i=1,\ldots,\gamma-1).$$

Шаг 5 (расчет n-го диагонального элемента фактора D и решения p^* системы (5))**.** Положить $\gamma=\gamma+1$. Если $\gamma\neq n,$ то перейти к шагу 1; иначе вычислить:

$$h^n_n+h^n_{n\,n}p_n=0\Rightarrow h^n_n=h_n+h_{n\,1}p^1_1+h_{n\,2}p^2_2+\cdots+h_{n\,n-1}p^{n-1}_{n-1},\ h^n_{n\,n}=h_{n\bullet}\prod_{i=1}^{n-1}E^i_{1_i},$$
$$d_{n\,n}=\max\left\{\delta,\left|h^n_{n\,n}\right|\right\},$$
$$p^n=p^*=\left(p^{n-1}_1-e^1_{1_{n-1}\,n}\,h^n_n/d_{n\,n}\quad p^{n-1}_2-e^2_{1_{n-1}\,n}\,h^n_n/d_{n\,n}\quad\cdots\quad p^{n-1}_{n-1}-e^{n-1}_{1_{n-1}\,n}\,h^n_n/d_{n\,n}\quad -h^n_n/d_{n\,n}\right).$$

В данной работе модификацию второго варианта предлагается положить в основу построения непрерывного аналога для задач минимизации квадратичного псевдобулева полинома с целью их решения за полиномиальное время, а следовательно, и для решения одной из центральных проблем математики. Следуя Бересневу [Береснев, 2005], опишем эту проблему следующим образом. В теории вычислительной сложности среди задач распознавания принято выделять классы *P* (разрешимых за полиномиальное время) и *NP* (проверяемых за полиномиальное время). В классе *P* можно проверить любой ответ, и, значит, $P\subseteq NP$. Доказать или опровергнуть обратное включение пока никому не удается: на сегодняшний день это одна из центральных проблем математики, связанная с существованием задач, решения которых не могут быть найдены за полиномиальное время, но решения которых могут быть проверены за полиномиальное время. Одним из аргументов в пользу того, что $P\neq NP$, является существование в классе *NP* еще одного сложностного класса *NPC*, называемого классом *NP*-полных задач и включающего в себя самые трудные задачи из класса *NP*.

Задачу из класса *NP* называют *NP*-полной, если существование полиномиального алгоритма для ее решения влечет существование полиномиальных алгоритмов для всех задач из класса *NP*. Задачу называют *NP*-трудной, если к ней полиномиально сводится любая задача из класса *NP*. *NP*-полные задачи разрешимы тогда и только тогда, когда $P=NP$, в то время как все, что можно сказать о *NP*-трудной задаче, — это то, что она неразрешима за полиномиальное время, если $P\neq NP$ [Гэри и др., 1982]. К настоящему времени известно огромное число *NP*-полных задач, однако ни для одной из них так и не удалось разработать точный полиномиальный алгоритм, поэтому общепринятой является гипотеза о том, что $P \neq NP$. Для того чтобы ее опровергнуть, достаточно построить полиномиальный алгоритм решения любой (какой-нибудь одной) NP-трудной задачи. Список таких задач включает десятки тысяч задач из различных областей науки.

Из вышесказанного следует, что разработка второго варианта вычислительной схемы относится к актуальным проблемам вычислительной математики.

Построение непрерывного аналога для задач минимизации квадратичного псевдобулева полинома. Интеграция техники исключения Гаусса и факторизации Холесского определяет и новую математическую формулировку, и новую форму задания необходимых и достаточных условий оптимальности решения задачи минимизации псевдобулевой функции:

$$\min_{p\in B^n} f(p) = \frac{1}{2} p^T Hp + p^T h. \tag{3.4}$$

Здесь B^n — множество вершин единичного n-мерного куба, а псевдобулева функция — это произвольное отображение множества бинарных наборов длины n на вещественную прямую. Такого рода функции являются естественным обобщением классических булевых функций и находят многочисленные применения в разного рода прикладных исследованиях [Crama et al., 2011]. На первый взгляд поставленная задача оптимизации (3.4) не имеет непрерывного аналога. Однако такой аналог становится очевидным, если задачу переформулировать следующим образом.

Учитывая тождество $p_i^2 \equiv p_i$, для аппроксимации квадратичного многочлена от булевых переменных строго выпуклым квадратичным многочленом от вещественных переменных таким образом, чтобы их значения совпадали на множестве B^n, достаточно во втором варианте вычислительной схемы на каждой итерации γ процедуры расчета элементов p^*, D, U потребовать

$$h_\gamma = h_\gamma + h_{\gamma\gamma} - 1,\ d_{\gamma\gamma} = h_{\gamma\gamma} = 1. \tag{3.5}$$

В результате (3.4) сводится к эквивалентной задаче:

$$\min_{x\in S\subset R^n} x^T x,\ p_i = p_i^* + x_i + \sum_{j=i+1}^{n} p_{ij} x_j,\ p_i \in \{0,1\},\ i = 1,\ldots,n. \tag{3.6}$$

Очевидно,

$$\min_{x\in S\subset R^n} x^T x,\ -p_i^* \le x_i + \sum_{j=i+1}^{n} p_{ij} x_j \le 1 - p_i^*,\ i = 1,\ldots,n. \tag{3.7}$$

Здесь S — множество угловых точек выпуклого многогранника (полиэдра):

$$-p_i^* \le x_i + \sum_{j=i+1}^{n} p_{ij} x_j \le 1 - p_i^*,\ i = 1,\ldots,n, \tag{3.8}$$

граница которого состоит из вершин и граней различных размерностей от 1 до $n - 1$ [Циглер, 2014]. Полиэдр является возмущением n-мерного куба (или n-гиперкуба), имеет $2n$ граней и 2^n вершин [Пападимитриу и др., 1985]. Каждая грань имеет $2n - 2$ соседние грани и одну симметричную грань-антипод [Казанцев, 2012].

Указанные варианты записи задач (3.4), (3.7) преобразуются друг в друга в результате несложных замен переменных, при этом оптимальные значения одних переменных достаточно просто строятся по оптимальным значениям других. Таким образом, такая переформулировка второго варианта вычислительной схемы определяет построение третьего.

Актуальность построения новой математической формулировки и новой формы задания необходимых и достаточных условий оптимальности решения задачи (3.4) для выхода за рамки целочисленных методов оптимизации можно обосновать следующим образом. Задача (3.4) является алгоритмически трудной [Селиверстов, 2013]. Эффективные алгоритмы известны лишь в частных случаях, например, если квадратичные члены составляют трехдиагональную симметричную матрицу, минимум можно найти методом псевдобулева программирования [Береснев, 2005]. Там же отмечено, что задача минимизации полинома с неотрицательными коэффициентами для квадратичных полиномов является *NP*-трудной, поскольку к ней сводится *NP*-трудная задача о вершинном покрытии для произвольного графа. В силу этого попытки построения эффективных универсальных алгоритмов решения данной задачи становятся бесперспективными и для получения точного решения следует обратиться к алгоритмам, основанным на процедурах неявного перебора и, в частности, на вычислительной схеме ветвей и границ.

Решение задач минимизации квадратичного псевдобулева полинома. Исследованию подлежит только две грани полиэдра (3.8):

$$x_n = -p_n, \; x_n = 1 - p_n, \tag{3.9}$$

одна из которых или обе содержат вершины, ближайшие к началу координат. Для их вычисления достаточно решить $4n-4$ систем линейных уравнений с верхней треугольной матрицей

$$P = \begin{pmatrix} 1 & p_{12} & p_{13} & \cdots & p_{1n} \\ & 1 & p_{23} & \cdots & p_{2n} \\ & & \cdots & \cdots & \cdots \\ & & & 1 & p_{n-1\,n} \\ & & & & 1 \end{pmatrix} \tag{3.10}$$

с единицами на главной диагонали и выбрать среди них все ближайшие равноудаленные вершины, что достижимо за полиномиальное время [Schrijver, 1986]. Отсюда следует, что сводимость любой задачи (3.4), в том числе и *NP*-трудных, к эквивалентной задаче (3.7) за полиномиальное время означает доказательство утверждения $P = NP$.

Феноменология построения непрерывного аналога для задач минимизации квадратичного псевдобулева полинома. Феноменологию (описание построения непрерывного аналога) сведения задачи (3.4) к эквивалентной задаче (3.7) рассмотрим на следующем примере.

Пример 3.1. Построить непрерывный аналог для задачи

$$\min_{p \in B^2} f(p) = \frac{1}{2} p^T \begin{pmatrix} 0 & 4 \\ 4 & 0 \end{pmatrix} p + p^T \begin{pmatrix} -5 \\ -5 \end{pmatrix}.$$

Построение проведем выделением полного квадрата следующим образом:

$$f(p) = 4p_1p_2 - 5p_1 - 5p_2 \;\Rightarrow\; p_1^2 + 4p_1p_2 - 6p_1 - 5p_2 \;\Rightarrow$$
$$\Rightarrow\; (p_1 + 2p_2 - 3)^2 + 3p_2 - 9 \;\Rightarrow\; (p_1 + 2p_2 - 3)^2 + (p_2 + 1)^2.$$

Очевидно,

$$x = \begin{pmatrix} -3 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 & 2 \\ & 1 \end{pmatrix} p \;\Rightarrow\; p = \begin{pmatrix} 5 \\ -1 \end{pmatrix} + \begin{pmatrix} 1 & -2 \\ & 1 \end{pmatrix} x.$$

В результате перейдем к эквивалентной задаче:

$$\min_{x \in S \subset R^2} x^T x.$$

Здесь S — множество угловых точек выпуклого многогранника:

$$\begin{pmatrix}-5\\1\end{pmatrix}\le\begin{pmatrix}1&-2\\0&1\end{pmatrix}x\le\begin{pmatrix}-4\\2\end{pmatrix},$$

представленное на рис. 3.1.

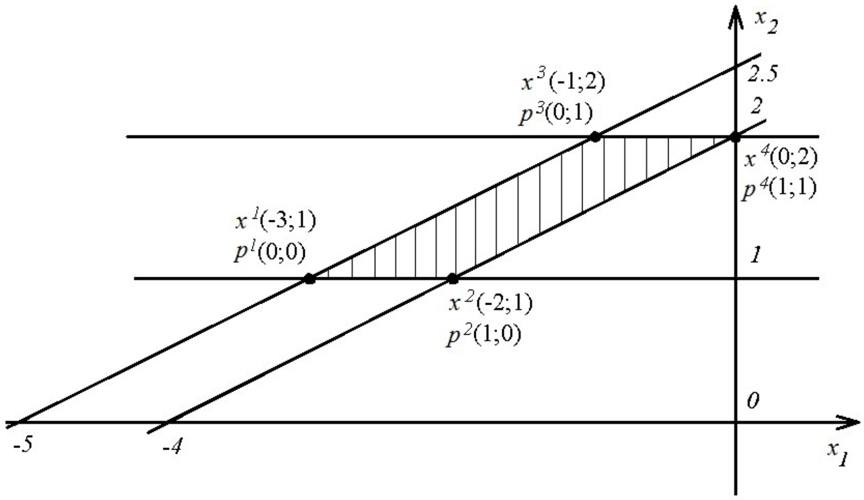

Рис. 3.1. Множество угловых точек выпуклого многогранника

Следует отметить, что:

- оптимальную угловую точку выпуклого многогранника для задач большой размерности можно найти и прямым мультипликативным методом ЛП [Свириденко, 2017]; Стренг в [Стренг, 1980] отметил: «Линейная алгебра, в отличие от анализа, связана с решением только уравнений и не имеет дела с неравенствами. Это всегда казалось очевидным, но в конце концов я понял, что линейное программирование представляет собой контрпример: оно связано с неравенствами, но, безусловно, является частью линейной алгебры»; то же самое справедливо и в отношении ньютоновских методов безусловной оптимизации, и минимизации вещественных квадратичных многочленов от булевых переменных;
- (3.7) определяют и новую математическую формулировку, и новую форму задания необходимых и достаточных условий оптимальности решения задачи (3.4).

Феноменология построения непрерывного аналога для задач минимизации произвольного псевдобулева полинома. В данной работе феноменология построения ограничена рассмотрением задачи минимизации псевдобулева полинома третьей степени.

Пример 3.2. Построить непрерывный аналог для задачи

$$\min_{p\in B^2} f(p)=6p_1p_2p_3+4p_1p_2+2p_1p_3+12p_2p_3-p_1-5p_2+18p_3.$$

Построение проведем выделением полного квадрата следующим образом:

$$f(p)=p_1^2+6p_1p_2p_3+4p_1p_2+2p_1p_3-2p_1+21p_2p_3-5p_2+18p_3 \Rightarrow$$
$$\Rightarrow (p_1+3p_2p_3+2p_2+p_3-1)^2-4p_2p_3-5p_2+19p_3-1 \Rightarrow$$
$$\Rightarrow (p_1+3p_2p_3+2p_2+p_3-1)^2+(p_2-2p_3-3)^2+(p_3+1)^2-11.$$

Очевидно,

$$x=\begin{pmatrix}-1\\-3\\1\end{pmatrix}+\begin{pmatrix}1&2&1\\&1&-2\\&&1\end{pmatrix}\begin{pmatrix}p_1+3p_2p_3\\p_2\\p_3\end{pmatrix}\Rightarrow\begin{pmatrix}p_1+3p_2p_3\\p_2\\p_3\end{pmatrix}=\begin{pmatrix}0\\1\\-1\end{pmatrix}+\begin{pmatrix}1&-2&-5\\&1&2\\&&1\end{pmatrix}x.$$

В результате перейдем к эквивалентной задаче:

$$\min_{x\in S\subset R^2} x^T x.$$

Здесь S — множество угловых точек выпуклого многогранника

$$\begin{pmatrix}3p_2p_3\\-1\\1\end{pmatrix}\le\begin{pmatrix}1&-2&-5\\&1&2\\&&1\end{pmatrix}x\le\begin{pmatrix}1+3p_2p_3\\0\\2\end{pmatrix}.$$

Замечания. Задача минимизации полиномов от булевых переменных относится к слабоизученным задачам, к которым проявляет интерес лишь ограниченный круг специалистов [Береснев, 2005]. Несмотря на отсутствие должного внимания, она имеет большое значение для развития теории дискретной оптимизации: задачу минимизации полиномов от булевых переменных можно рассматривать как некоторую универсальную задачу, в терминах которой могут быть сформулированы многие задачи математического программирования. Например, можно указать три задачи, эквивалентные задаче минимизации полиномов [Береснев, 2005]. Одной из них является двухуровневая задача размещения производства [Береснев и др., 1978], обобщающая задачу размещения. Эта задача достаточно хорошо изучена, что связано с многочисленными возможностями ее практического использования (от размещения предприятий и складов [Михалевич и др., 1986; Barros et al., 1994] до построения оптимальной системы изделий и комплектующих их узлов [Береснев и др., 1978]). Другими тесно связанными задачами являются обобщенная задача выбора множества строк пары матриц [Береснев, 1979], а также обобщенная задача о покрытии множества системой подмножеств. Впервые задачи псевдобулевой оптимизации подробно исследовались в монографии [Hammer et al., 1968], в которой, в частности, были разработаны методы решения аналитически заданных задач псевдобулевой оптимизации. Псевдобулевые функции играют важную роль в оптимизационных моделях в различных областях, таких как проектирование [Юдин и др., 1982; Barahona et al., 1988], теория надежности [Масич, 2010], теория вычислительных систем [Karp et al., 1975], статистика (классификация) [Ranyard, 1976; Rao, 1971], экономика [Hammer et al., 1971], финансы [Hillier, 1969; Laughhunn, 1970; Laughhunn et al., 1971], менеджмент [Weingartner, 1966], дискретная математика (оптимизация на графах) [Hammer, 1977; Ebenegger and th., 1984], промышленность (календарное планирование и составление расписания) [Масич, 2010; Crama, 1997]. Псевдобулевые функции встречаются в комбинаторной теории как функции ранга матроидов [Crama et al., 1989; Welsh, 1976] или как функции, связанные с определенными параметрами графа, такими как число стабильности, хроматическое число и так далее [Fraenkel et al., 1984; Hammer et al., 1981; Nieminen, 1974]. Оптимизация псевдобулевых функций используется в распознавании образов как при отборе информативных признаков [Масич, 2012], так и при построении классификаторов [Масич, 2012; Масич, 2010].

Заключение

Рассмотрены три варианта вычислительной схемы прямого мультипликативного алгоритма расчета элементов D, U модифицированной факторизации Холесского для преобразования и решения p^* системы (5) в рамках одной процедуры, учитывающие разреженность матрицы H, представленной в упакованном виде. Алгоритм основан на модификации прямого мультипликативного алгоритма ЛП, как метода решения системы линейных уравнений большой размерности, путем интеграции одной из техник построения существенно положительно определенной матрицы $\overline{H}$.

Первый вариант вычислительной схемы представляет собой алгоритм расчета направления спуска в ньютоновских методах безусловной оптимизации. Преимущество первого варианта состоит в минимизации заполнения главных строк мультипликаторов без потери точности результатов, причем изменения в позиции очередной обрабатываемой строки матрицы не вносятся, что позволяет использовать статические форматы хранения данных.

Второй вариант вычислительной схемы является модификацией первого и определяет построение новой математической формулировки задачи КП, а также новой формы задания необходимых и достаточных условий оптимальности, которые достаточно просты и могут быть использованы для построения методов математического программирования. Преимуществом второго варианта является возможность расчета альтернативного направления спуска в тех случаях, когда регулярная процедура поиска направления спуска не дает существенных результатов.

Третий вариант вычислительной схемы является модификацией второго и определяет построение непрерывного аналога задачи минимизации вещественного квадратичного многочлена от булевых переменных и новой формы задания необходимых и достаточных условий оптимальности для разработки методов их решения за полиномиальное время. В результате исходная задача сводится к задаче поиска минимального расстояния между началом координат и угловой точкой выпуклого многогранника (полиэдра), который является возмущением n-мерного куба и описывается системой двойных линейных неравенств с верхней треугольной матрицей коэффициентов с единицами на главной диагонали. Исследованию подлежат только две грани, одна из которых или обе содержат вершины, ближайшие к началу координат. Для их вычисления достаточно решить $4n - 4$ систем линейных уравнений и выбрать среди них все ближайшие равноудаленные вершины, что достижимо за полиномиальное время.

Прямым продолжением данной работы является построение метода поиска минимума квадратичной функции на многогранном множестве ограничений, основанного на решениях систем линейных уравнений, размерность которых не выше числа переменных целевой функции.

Список литературы (References)

Береснев В. Л. Дискретные задачи размещения и полиномы от булевых переменных. — Новосибирск: Изд-во ин-та математики, 2005. — 408 с.

Beresnev V. L. Diskretnye zadachi razmeshhenija i polinomy ot bulevyh peremennyh [Discrete location problems and polynomials of Boolean variables]. — Novosibirsk: The publisher of the Institute of mathematics, 2005. — 408 p. (in Russian).

Береснев В. Л. Алгоритмы минимизации полиномов от булевых переменных // Проблемы кибернетики. — Вып. 36. — М.: Наука, 1979. — С. 225–246.

Beresnev V. L. Algoritmy minimizacii polinomov ot bulevyh peremennyh [Algorithms for minimization of polynomials in Boolean variables] // Problems of Cybernetics. — Vol. 36. — Moscow: Nauka, 1979. — P. 225–246 (in Russian).

Береснев В. Л., Гимади Э. Х., Дементьев В. Т. Экстремальные задачи стандартизации. — Новосибирск: Наука, 1978.

Beresnev V. L., Gimadi Je. H., Dement'ev V. T. Jekstremal'nye zadachi standartizacii [Extremal problems of standardization]. — Novosibirsk: Nauka, 1978 (in Russian).

Гилл Ф., Мюррей У. Численные методы условной оптимизации. — М.: Мир, 1977.

Gill Ph. E., Murray W. Numerical methods for constrained optimization. — National physical laboratory Teddington, Middlesex. // Academic Press, 1974. (Russ. ed.: Gill F., Mjurrej U. Chislennye metody uslovnoj optimizacii. — Moscow: Mir, 1977.)

Гилл Ф., Мюррей У., Райт М. Практическая оптимизация. — М.: Мир, 1985.

Gill Ph. E., Murray W., Wright M. H. Practical optimization. — System Optimization Laboratory Department of Operations Research Stanford University California, USA // Academic Press, 1981. (Russ. ed.: Gill F., Mjurrej U., Rajt M. Prakticheskaja optimizacija. — Moscow: Mir, 1985.)

Григорьева О. Н., Дмитриева О. А. Моделирование линейных динамических систем большой размерности с разреженными матрицами коэффициентов // Информатика и компьютерные технологии-2011. — Донецк: Донецкий национальный технический университет, 2011. — С. 199–203.

Grigor'eva O. N., Dmitrieva O. A. Modelirovanie linejnyh dinamicheskih sistem bol'shoj razmernosti s razrezhennymi matricami kojefficientov [Modeling linear dynamical systems of high dimension with sparse matrices of coefficients] // Informatics and computer technologies-2011. — Donetsk: Donetsk national technical University, 2011. — P. 199–203 (in Russian).

Гэри М., Джонсон Д. Вычислительные машины и труднорешаемые задачи. — М.: Мир, 1982.

Gjeri M., Dzhonson D. Vychislitel'nye mashiny i trudnoreshaemye zadachi [Computers and trudnoreshaemyh tasks]. — Moscow: Mir, 1982 (in Russian).

Джордж А., Лю Дж. Численное решение больших разреженных систем уравнений. — М.: Мир, 1984.

George A., Liu J. W.-H. Computer solution of large sparse positive definite systems. — Prentice-Hall, Inc, Englewood Cliffs, Ne Jersey, 1981. (Russ. ed.: Dzhordzh A., Lju Dzh. Chislennoe reshenie bol'shih razrezhennyh sistem uravnenij. — Moscow: Mir, 1984.)

Дмитриева О. А. Оптимизация выполнения матрично-векторных операций при параллельном моделировании динамических процессов // Наукові праці ДонНТУ. Сер. Обчислювальна техніка та автоматизація. — 2014. — № 1 (26). — С. 94–100.

Dmitrieva O. A. Optimizacija vypolnenija matrichno-vektornyh operacij pri parallel'nom modelirovanii dinamicheskih processov [Optimization of performance of matrix and vector operations at parallel simulation of dynamic processes] // Donetsk National Technical University. — 2014. — No. 1 (26). — P. 94–100 (in Russian).

Зеленков Г. А., Хакимова А. Б. Подход к разработке алгоритмов ньютоновских методов оптимизации, программная реализация и сравнение эффективности // Компьютерные исследования и моделирование. — 2013. — Т. 5, № 3. — С. 367–377.

Zelenkov G. A., Hakimova A. B. Podhod k razrabotke algoritmov n'jutonovskih metodov optimizacii, programmnaja realizacija i sravnenie jeffektivnosti [Approach to development of algorithms of Newtonian methods of unconstrained optimization, their software implementation and benchmarking] // Computer Research and Modeling. — 2014. — Vol. 5, no. 3. — P. 367–377 (in Russian).

Казанцев И. Г. О конструктивном вычислении сечений многомерного куба // Интерэкспо Гео-Сибирь. — 2012. — Т. 1, № 4. — С. 168–171.

Kazancev I. G. O konstruktivnom vychislenii sechenij mnogomernogo kuba [On a constructive computation of the section of hypercube] // Interjekspo Geo-Sibir'. — 2012. — Vol. 1, no. 4. — P. 168–171.

Карманов В. Г. Математическое программирование. — М.: Мир, 1975.

Karmanov V. G. Matematicheskoe programmirovanie [Mathematical programming]. — Moscow: Mir, 1975 (in Russian).

Кочетов Ю. А., Плясунов А. В. Локальный поиск в комбинаторной оптимизации. Нужна ли производная? // Труды XIII Байкальской международной школы-семинара «Методы оптимизации и их приложения», 2–8.07.2005. — Том 1: Иркутск – ИСЭМ СО РАН, 2005. — С. 65–76.

Kochetov Yu. A., Pljasunov A. V. Lokal'nyj poisk v kombinatornoj optimizacii. Nuzhna li proizvodnaja? [Local search in combinatorial optimization. What about the derivative?] // Mathematical programming: Proceedings of XIII Baikal International School-seminar “Optimization methods and their applications”, July, 2–8, Irkutsk, Baikal, 2005. — Vol. 1. Irkutsk: Melentiev Energy Systems Institute SB RAS, 2005. — P. 65–76 (in Russian).

Масич И. С. Задачи оптимизации и их свойства в логических алгоритмах распознавания // Проблемы оптимизации и экономические приложения: материалы V Всероссийской конференции. — Омск: Изд-во Ом. гос. ун-та, 2012.

Masich I. S. Zadachi optimizacii i ih svojstva v logicheskih algoritmah raspoznavanija [Optimization problems and their properties in logical recognition algorithms] // Optimization problems and economic applications: proceedings of the V all-Russian conference. — Omsk: Publishing house of Om. state University, 2012 (in Russian)

Масич И. С. Комбинаторная оптимизация и логические алгоритмы классификации в задаче прогнозирования осложнений инфаркта миокарда // Инновационные тенденции развития российской науки: материалы III Международной научно-практической конференции. — Красноярск, 2010. — С. 276–280.

Masich I. S. Kombinatornaja optimizacija i logicheskie algoritmy klassifikacii v zadache prognoziro-vanija oslozhnenij infarkta miokarda [Combinatorial optimization and Boolean classification algorithms in the task of predicting complications of myocardial infarction] // Innovative trends in the development of Russian science: materials of III International scientific-practical conference. — Krasnoyarsk, 2010. — P. 276–280 (in Russian)

Михалевич В. С., Трубин В. А., Шор Н. З. Оптимизационные задачи производственно-транспортного планирования. — М.: Наука, 1986.
Mihalevich V. S., Trubin V. A., Shor N. Z. Optimizacionnye zadachi proizvodstvenno-transportnogo planirovanija [Optimization problems of production-transportation planning]. — Moscow: Nauka, 1986 (in Russian).

Пападимитриу Х., Стайглиц К. Комбинаторная оптимизация. Алгоритмы и сложность. — М.: Мир, 1985.
Papadimitriou Christos H. Combinatorial optimization: Algorithms and Complexity. — Massachusetts Institute of Technology. National Technical University of Athens // Prentice-Hall Inc. Englewood Cliffs — New Jersey, 1982. (Russ. ed.: Papadimitriu H., Stajglic K. Kombinatornaja optimizacija. Algoritmy i slozhnost'. — Moscow: Mir, 1985.)

Парлетт Б. Симметричная проблема собственных значений. Численные методы. — М.: Мир, 1983.
Parlett B. N. The symmetric eigenvalue problem. — University of California Berkeley, California, 1980. (Russ. ed.: Parlett B. Simmetrichnaja problema sobstvennyh znachenij. Chislennye metody. — Moscow: Mir, 1983.)

Писсанецки С. Технология разреженных матриц: Пер. с англ. — М.: Мир, 1988. — 410 с.
Pissanetzky S. Sparse matrix technology / Centro Atamico Bariloche, Bariloche, Argentina. — Academic Press Inc., 1984. (Russ. ed.: Pissanecki S. Tehnologija razrezhennyh matric. — Moscow: Mir, 1988.)

Полак Э. Численные методы оптимизации. Единый подход. — М.: Мир, 1974.
Polak E. Computation methods in optimization // Mathematics in Science and Engineering. — Vol. 77. — Academic Press: New York, London, 1971. (Russ. ed.: Polak Je. Chislennye metody optimizacii. Edinyj podhod. — Moscow: Mir, 1974.)

Поляк Б. Т. Метод Ньютона и его роль в оптимизации и вычислительной математике // Труды ИСА РАН 2006. — Т. 28. — С. 48–66.
Poljak B. T. Metod N'jutona i ego rol' v optimizacii i vychislitel'noj matematike [Newton's method and its role in optimization and computational mathematics] // Proceedings of ISA RAS 2006. — Vol. 28. — P. 48–66 (in Russian).

Пшеничный Б. Н., Данилин Ю. М. Численные методы в экстремальных задачах. — М.: Наука, 1975.
Pshenichnyj B. N., Danilin Yu. M. Chislennye metody v jekstremal'nyh zadachah [Numerical methods in extremal problems]. — Moscow: Nauka, 1975 (in Russian).

Свириденко А. Б. Априорная поправка в ньютоновских методах оптимизации // Компьютерные исследования и моделирование. — 2015. — Т. 7, № 4. — С. 835–863.
Sviridenko A. B. Apriornaja popravka v n'jutonovskih metodah optimizacii [The correction to Newton's methods of optimization] // Computer Research and Modeling. — 2015. — Vol. 7, no. 4. — P. 835–863 (in Russian).

Свириденко А. Б. Прямые мультипликативные методы для разреженных матриц. Несимметричные линейные системы // Компьютерные исследования и моделирование. — 2016. — Т. 8, № 6. — С. 833–860.
Sviridenko A. B. Prjamye mul'tiplikativnye metody dlja razrezhennyh matric. Nesimmetrichnye linejnye sistemy [Direct multiplicative methods for sparse matrices. Unbalanced linear systems] // Computer research and modeling. — 2016. — Vol. 8, no. 6. — P. 833–860 (in Russian).

Свириденко А. Б. Прямые мультипликативные методы для разреженных матриц. Линейное программирование // Компьютерные исследования и моделирование. — 2017. — Т. 9, № 2. — С. 143–165.
Sviridenko A. B. Prjamye mul'tiplikativnye metody dlja razrezhennyh matric. Nesimmetrichnye linejnye sistemy [Direct multiplicative methods for sparse matrices. Linear Programming] // Computer research and modeling. — 2017. — Vol. 9, no. 2. — P. 143–165 (in Russian).

Свириденко А. Б., Зеленков Г. А. Взаимосвязь и реализация квазиньютоновских и ньютоновских методов безусловной оптимизации // Компьютерные исследования и моделирование. — 2016. — Т. 8, № 1. — С. 55–78.
Sviridenko A. B., Zelenkov G. A. Vzaimosvjaz' i realizacija kvazin'jutonovskih i n'jutonovskih metodov bezuslovnoj optimizacii [Correlation and realization of quasi-Newton methods of absolute optimization] // Computer Research and Modeling. — 2016. — Vol. 8, no. 1. — P. 55–78 (in Russian).

Селиверстов А. В. О мономах квадратичных форм // Дискрет. анализ и исслед. операций. — 2013. — Т. 20, № 3. — С. 65–70.

Seliverstov A. V. O monomah kvadratichnyh form [On monomials quadratic forms] // Discrete analysis and operations research. — 2013. — Vol. 20, no. 3. — P. 65–70 (in Russian).

Стренг Г. Линейная алгебра и ее применения. — М.: Мир, 1980.

Strang G. Linear algebra and its applications / Massachusetts Institute of Technology. — Academic Press, New York, San Francisco, London, 1976. (Russ. ed.: Streng G. Linejnaja algebra i ee primenenija. — Moscow: Mir, 1980.)

Хакимова А. Б., Дикусар В. В., Зеленков Г. А. Увеличение эффективности ньютоновских методов оптимизации. Информодинамический подход // Труды ИСА РАН «Динамика неоднородных систем». — Вып. 14-А, Т. 53. — М.: Книжный дом «ЛИБРОКОМ», 2010. — С. 97–114.

Khakimova A. B., Dikusar V. V., Zelenkov G. A. Uvelichenie jeffektivnosti n'jutonovskih me-todov optimizacii. Informodinamicheskij podhod [Increasing the efficiency of the Newtonian methods of optimization. Informationmicardis approach] // The works of ISA Russian Academy of Sciences "Dynamics of heterogeneous systems". — Issue 14-A, Vol. 53. — Moscow: Book house "LIBROKOM", 2010. — P. 97–114 (in Russian).

Хакимова А. Б., Зеленков Г. А., Рзун И. Г. Подход к увеличению эффективности мультипликативного алгоритма симплекс-метода // Труды ИСА РАН «Динамика неоднородных систем». — Вып. 14, Т. 53 (2). — М.: Книжный дом «ЛИБРОКОМ», 2010. — С. 245–251.

Khakimova A. B., Zelenkov G. A., Rzun I. G. Podhod k uvelicheniju jeffektivnosti mul'tiplikativnogo algoritma simpleks-metoda [Approach to increase the efficiency of the multiplicative algorithm of the simplex method] // The works of ISA Russian Academy of Sciences "Dynamics of heterogeneous systems". — 2010. — Issue 14, Vol. 53 (2). — P. 245–251 (in Russian).

Хакимова А. Б., Хакимов Б. Б. Единый подход к решению задач математического программирования гуманитарной компьютерной клиники // Сборник статей I-й Международной конференции «Системные, информационные и технические средства и технологии в профессиональной деятельности, образовании, оздоровлении и профилактике». — Санкт-Петербург, 2003. — С. 88–92.

Khakimova A. B., Khakimov B. B. Edinyj podhod k resheniju zadach matematicheskogo programmirovanija gumanitarnoj komp'juternoj kliniki [A unified approach to the solution of problems of mathematical programming Humanities computer clinic] // A collection of articles I international conference "System, information and technical tools and technologies in their professional activities, education, rehabilitation and prevention". — Saint-Petersburg, 2003. — P. 88–92 (in Russian).

Циглер Г. М. Теория многогранников. — М.: Изд-во МЦНМО, 2014.

Ziegler G. M. Lectures on Polytopes. — New York: Springer-Verlag Inc., 1995. (Russ. ed.: Cigler G. M. Teorija mnogogrannikov. — Moscow: Izd-vo MCNMO, 2014.)

Черноруцкий И. Г. Методы оптимизации. Компьютерные технологии. — СПб.: БХВ-Петербург, 2011. — 384 с.

Chernoruckij I. G. Metody optimizacii. Komp'juternye tehnologii [Methods of optimization. Computer technology]. — St. Petersburg: BHV-Petersburg, 2011. — 384 p. (in Russian).

Черноруцкий И. Г. Практическая оптимизация и невыпуклые задачи // Научно-технические ведомости СПбГПУ. Информатика. Телекоммуникации. Управление. — СПб.: Федеральное государственное автономное образовательное учреждение высшего образования «Санкт-Петербургский политехнический университет Петра Великого». — 2013. — № 4 (176). — С. 79–86.

Chernoruckij I. G. Prakticheskaja optimizacija i nevypuklye zadachi [Practical optimization and nonconvex problems] // Nauchno-tekhnicheskie Vedomosti SPbGPU. Informatics. Telecommunications. Management. — St. Petersburg: Federal state Autonomous educational institution of higher professional education "Saint-Petersburg Polytechnic University Peter the Great". — 2013. — No. 4 (176). — P. 79–86 (in Russian).

Юдин Д. Б., Горяшко А. П., Немировский А. С. Математические методы оптимизации устройств и алгоритмов АСУ. — М.: Радио и связь, 1982. — 288 с.

Judin D. B., Gorjashko A. P., Nemirovskij A. S. Matematicheskie metody optimizacii ustrojstv i algoritmov ASU [Mathematical methods for optimization of devices and algorithms ACS]. — Moscow: Radio and communication, 1982 (in Russian).

Barahona F., Grotschel M., Junger M., Reinelt G. An application of combinatorial optimization to statistical physics and circuit layout design // Operation Research. — 1988. — No. 36. — P. 493–513.

Barros A. I., Labbe M. A general model for the uncapacitated facility and depot location problem // Location Science. — 1994. — Vol. 2, no. 3. — P. 173–191.

Crama Y. Combinatorial Optimization Models for Production Scheduling in Automated Manufacturing Systems. // European Journal of Operational Research. — 1997. — No. 99. — P. 136–153.

Crama Y., Hammer P. L. Bimatroidal independence systems // Mathematical Methods of Operations Research. — 1989. — Vol. 33, no. 3. — P. 149–165.

Crama Y., Hammer P. L. Boolean Functions: Theory, Algorithms, and Applications. — New York: Cambridge University Press, 2011. — 687 p.

Ebenegger Ch., Hammer P. L., de Werra D. Pseudo-Boolean Functions and Stability of Graphs // Annals of Discrete Mathematics. — 1984. — No. 19. — P. 83–97.

Fraenkel A. S., Hammer P. L. Pseudo-Boolean functions and their graphs // Annals of Discrete Mathematics. — 1984. — No. 20. — P. 137–146.

Gill P. E., Murray W. Newton-type methods for unconstrained and linearly constrained optimization // Math. Prog. — 1974. — Vol. 7. — P. 311–350.

Hammer P. L. Pseudo-Boolean Remarks on Balanced Graphs // International Series of Numerical Mathematics. — 1977. — No. 36. — P. 69–78.

Hammer P. L., Hansen P., Simeone B. Upper planes of quadratic 0-1 functions and stability in graphs // Nonlinear Programming. — 1981. — No. 4. — P. 395–414.

Hammer P. L., Rudeanu S. Boolean Methods in Operations Research and Related Areas. — Berlin: Springer-Verlag; New York: Heidelberg, 1968. — 310 p.

Hammer P. L., Shliffer E. Applications of Pseudo-Boolean Methods to Economic Problems // Theory and decision. — 1971. — No. 1. — P. 296–308.

Hillier F. S. The Evaluation of Risky Interrelated Investments. — Amsterdam: North-Holland Publishing, 1969.

Karp R. M., Miller R. G., Thatcher J. W. Reducibility Among Combinatorial Problems // Journal of Symbolic Logic. — 1975. — No. 40 (4). — P. 618–619.

Laughhunn D. J. Quadratic Binary Programming with Applications to Capital Budgeting Problems // Operations Research. — 1970. — No. 18. — P. 454–461.

Laughhunn D. J., Peterson D. E. Computational Experience with Capital Expenditure Programming Models under Risk // Business Finance. — 1971. — No. 3. — P. 43–48.

Nieminen J. A linear pseudo-Boolean viewpoint on matching and other central concepts in graph theory // Zastosowania Matematyki. — 1974. — No. 14. — P. 365–369.

Ranyard R. H. An Algorithm for Maximum Likelihood Ranking and Slater's i From Paired Comparisions // British Journal of Mathematical and Statistical Psychology. — 1976. — No. 29. — P. 242–248.

Rao M. R. Cluster Analysis and Mathematical Programming // Journal of the American Statistical Association. — 1971. — Vol. 66. — P. 622–626.

Schrijver A. Theory of linear and integer programming. // Wiley-interscience series in discrete mathematics. Department of econometrics, Tilburg University and Centrum voor Wiskunde en informatica, Amsterdam: A Wiley-interscience publication, 1986.

Weingartner H. M. Capital Budgeting of Interrelated Projects: Survey and Synthesis // Management Science. — 1966. — No. 12. — P. 485–516.

Welsh D. J. A. Matroid theory. — London: Academic Press, 1976.