

Научная статья

УДК 531.01+004.94

DOI: 10.17072/1993-0550-2023-1-30-46

Итерационный метод решения систем линейных алгебраических уравнений с положительно полуопределенными матрицами системы

Владимир Николаевич Иванов

Пермский государственный национальный исследовательский университет, Пермь, Россия
precol@psu.ru, <https://orcid.org/0000-0002-2634-4296>

Аннотация. Рассматривается задача нахождения нормального псевдорешения возмущенных систем линейных алгебраических уравнений с положительно полуопределенными матрицами системы. Данная задача возникает при численном моделировании динамики систем связанных твердых тел, на движение которых наложены дополнительные связи. Задача решается итерационным квазиньютоновским методом, основанным на SR1-формуле поиска минимума квадратичных функций. Доказаны основные свойства алгоритма. Приведен пример исследования динамики одной механической системы, на котором показана работоспособность метода.

Ключевые слова: системы линейных алгебраических уравнений; итерационные методы; квазиньютоновские методы; SR1-формула; системы абсолютно твердых тел; уравнения движения; дополнительные связи; численное интегрирование

Для цитирования: Иванов В. Н. Итерационный метод решения систем линейных алгебраических уравнений с положительно полуопределенными матрицами системы // Вестник Пермского университета. Математика. Механика. Информатика. 2023. № 1(60). С. 30–46. DOI: 10.17072/1993-0550-2023-1-30-46.

Статья поступила в редакцию 06.11.2022; одобрена после рецензирования 26.01.2023; принята к публикации 15.03.2023.

Research article

An Iterative Method for Linear Algebraic Equations Systems Solving With Positive Semidefinite Matrices

Vladimir N. Ivanov

Perm State University, Perm, Russia
precol@psu.ru, <https://orcid.org/0000-0002-2634-4296>

Abstract. The normal pseudo-solution for perturbed linear algebraic equations systems with positive semidefinite matrices finding problem is considered. This problem is actual for numerical simulation process of a multibody systems dynamics taking into account additional constraints on its motion. The problem is solved by the iterative quasi-Newtonian method based on the SR1-formula for quadratic functions minimum finding. The main algorithm properties are proved. An one mechanical system dynamics study example is given, which shows the method efficiency.

Keywords: systems of linear algebraic equations; iterative methods; Quasi-Newton methods; SR1-formula; multibody system; equations of motion; additional constraints; numerical integration

For citation: Ivanov V. N. An Iterative Method for Linear Algebraic Equations Systems Solving With Positive Semidefinite Matrices. Bulletin of Perm University. Mathematics. Mechanics. Computer Science. 2023;1(60):30-46. (In Russ.). DOI: 10.17072/1993-0550-2023-1-30-46.

The article was submitted 06.11.2022; approved after reviewing 26.01.2023; accepted for publication 15.03.2023.



Эта работа © 2023 Иванов В. Н. лицензируется под CC BY 4.0. Чтобы просмотреть копию этой лицензии, посетите <http://creativecommons.org/licenses/by/4.0/>.

Введение

Рассматривается задача определения нормальных псевдорешений возмущенных систем линейных уравнений (СЛАУ) с положительно полуопределенными матрицами системы. Задача возникает при численном моделировании динамики систем абсолютно твердых тел (САТТ), на движение которых наложены дополнительные связи.

Будем считать, что в САТТ все связи голономны, склерономны и идеальны. Введем обозначения: $\mathbf{q} = \mathbf{q}(t) = (q_1, \dots, q_m)^T$ – матрица-столбец обобщенных координат, $\dot{\mathbf{q}} = \frac{d\mathbf{q}}{dt}$, t – время, N – число тел в системе, $\mathbf{v} = \text{col}(v_1, \dots, v_{6N}) = B\dot{\mathbf{q}}$ – матрица-столбец абсолютных скоростей всех тел системы, $B = B_{6N \times m} = B(\mathbf{q})$ – матрица касательного локального базиса многообразия $\{\mathbf{q}\}$. Пусть на механическую систему наложено n дополнительных связей:

$$\mathbf{g}(\mathbf{q}) = 0. \quad (1)$$

Будем считать, что их избыточное число, т.е. часть из них либо является линейно зависимыми, либо тождественно равными 0. Такая ситуация возникает при образовании замкнутых кинематических цепей с помощью замыкания части вершин первичного графа механической системы со структурой дерева или, другими словами, наложения дополнительных шарниров между телами.

В этом случае уравнения движения механической системы можно выписать в обобщенных координатах в форме уравнений Лагранжа 1 рода [1–3], учитывая дополнительные связи с помощью множителей Лагранжа

$$\boldsymbol{\mu} = (\mu_1, \dots, \mu_n):$$

$$\begin{cases} M^* \ddot{\mathbf{q}} + G^T \boldsymbol{\mu} = \mathbf{f}(\mathbf{q}, \dot{\mathbf{q}}, t), \\ \ddot{\mathbf{g}} = G \ddot{\mathbf{q}} + \mathbf{h}(\mathbf{q}, \dot{\mathbf{q}}) = 0, \end{cases} \quad (2)$$

где $M^* = M^*(\mathbf{q}) = B^T M B$ – симметричная, положительно определенная инерционная матрица, $M = \text{diag}(M_1, \dots, M_N)$ – матрица масс и моментов инерции тел системы, $\mathbf{f} = \mathbf{f}(\mathbf{q}, \dot{\mathbf{q}}, t)$ – вектор слагаемых уравнений движения, не зависящих от ускорений, включающий актив-

ные, потенциальные, гироскопические и пр.

силы, $G = G(\mathbf{q}) = \frac{\partial \mathbf{g}(\mathbf{q})}{\partial \mathbf{q}^T}$ – $n \times m$ матрица, ор-

тогональная к дополнительным связям (1),

$$\mathbf{h} = \mathbf{h}(\mathbf{q}, \dot{\mathbf{q}}) = \frac{\partial}{\partial \mathbf{q}^T} \left(\frac{\partial \mathbf{g}(\mathbf{q})}{\partial \mathbf{q}^T} \dot{\mathbf{q}} \right) \dot{\mathbf{q}} \text{ – вектор-столбец}$$

правых частей дважды продифференцированных по времени уравнений связей.

Пусть $\text{rank}(G) = r < n$ – ранг матрицы G в силу избыточности связей (1).

Уравнения (2) представляют собой систему $m + n$ дифференциально-алгебраических уравнений (ДАЕ), линейных относительно обобщенных ускорений $\ddot{\mathbf{q}}$ и множителей Лагранжа $\boldsymbol{\mu}$. При этом матрицы $M^*(\mathbf{q})$ и $G(\mathbf{q})$, входящие в систему, являются переменными во времени. Для численного интегрирования подобных ДАЕ на каждом шаге алгоритма необходимо разрешать их относительно ускорений $\ddot{\mathbf{q}}$.

Для этого из (2) необходимо исключить множители Лагранжа. Выражая из первого уравнения системы (2) вектор $\ddot{\mathbf{q}}$ и подставляя его во второе уравнение системы, можно получить явное уравнение для вектора $\boldsymbol{\mu}$ в виде

$$A \boldsymbol{\mu} = \mathbf{b}, \quad (3)$$

где $A = G(M^*)^{-1} G^T = G(B^T M B)^{-1} G^T$ – симметричная положительно полуопределенная $n \times n$ матрица ранга r , совпадающего с рангом матрицы G [4], $\mathbf{b} = G(B^T M B)^{-1} \mathbf{f} + \mathbf{h}$ – вектор правых частей системы уравнений (3), причем ранг расширенной матрицы $(A \mid \mathbf{b})$ так же равен r .

Векторное уравнение (3) при указанных выше условиях имеет бесконечно много решений. Однако в силу исходной механической постановки задачи, нас интересует только нормальное псевдорешение [4] этой системы уравнений, т.е. решение минимальной длины, доставляющее минимум квадратичной функции $(A \boldsymbol{\mu} - \mathbf{b})^2$. Это следует из того, что множители Лагранжа $\boldsymbol{\mu}$ однозначно определяют реакции дополнительных связей (1), и любая механическая система удовлетворяет принципу наименьшего принуждения Гаусса [3], в соответствии с которым уравнения дви-

жения (2) получаются в результате минимизации по вектору \ddot{q} квадратичного функционала

$$\frac{1}{2} (M^* \ddot{q} - f)^2 - \mu^T (G \ddot{q} + h).$$

Нормальное псевдорешение СЛАУ обычно ищут методом сингулярного разложения матрицы системы [4], что требует больших вычислительных затрат по сравнению с методами Гаусса или Холецкого, которые могут быть применены в случае положительной определенности матрицы системы. Таким образом, при интегрировании системы уравнений (2) необходимо сначала сформировать и решить систему уравнений (3), а затем, найденные множители Лагранжа подставить в первую группу уравнений (2) для определения вектора ускорений \ddot{q} . Видно, что решение уравнений (2) требует множества матричных операций, число которых, в общем случае, возрастает по кубическому закону с ростом размерности задачи.

Заметим, что при малой величине шага интегрирования изменения матриц систем (2) и (3) малы. Это позволяет для снижения вычислительных затрат применить итерационные алгоритмы решения системы уравнений (3), основанные на использовании решений, полученных на предыдущих шагах интегрирования.

В настоящей статье для решения этой задачи предлагается использовать итерационный квазиньютоновский метод, основанный на SR1-формуле поиска минимума квадратичных функций. Особенность метода – возможность находить решение СЛАУ с положительно полуопределенными матрицами системы. Ниже доказываются основные свойства предлагаемого метода.

1. Постановка задачи

Сформулируем следующую задачу: требуется определить вектор $\mu = \mu(t)$ из СЛАУ (3) в момент времени $t_j = t_{j-1} + \Delta t$ с использованием решения в предыдущий момент времени t_{j-1} , где Δt – малое приращение по времени.

Введем стандартные обозначения СЛАУ. Для этого переобозначим вектор $\mu = \mu(t)$ через n -мерный вектор-столбец неизвестных x . Тогда задачу уточнения решения системы уравнений (3) при малых изменениях коэффициентов уравнений можно переформулировать следующим образом.

Пусть задана СЛАУ

$$Ax = b, \quad (4)$$

где $A = A_0 + \varepsilon F$, A_0 , F – известные симметричные и положительно полуопределенные матрицы n -го порядка; $b = b_0 + \varepsilon f_0 \in R^n$ – известный вектор-столбец правых частей; εF , εf_0 – матрица и вектор-столбец возмущений; $x \in R^n$ – вектор-столбец неизвестных; ε – малый параметр; $\|F\| = O(\|A_0\|)$, $\|f_0\| = O\|b_0\|$. Система уравнений (4) совместна. Кроме того, будем считать, что, если матрица возмущений $\varepsilon F = \mu A_0$, то $\varepsilon f_0 = \mu b_0$. В общем случае положим, что матрица возмущений εF непропорциональна матрице A_0 .

Пусть ранг матриц A , A_0 , F , $(A \mid b)$ равен $r \leq n$.

Квадрат нормы матрицы:

$$\|A\|^2 = \|A\|_2^2 = \sup_{\eta \neq 0} \frac{\eta^T (A^T A) \eta}{\eta^T \eta}.$$

В силу симметрии и положительной полуопределенности матрицы A_0 все ее собственные значения являются действительными неотрицательными числами [4, 5]. Будем считать, что $\lambda_1 > \dots > \lambda_r$ ненулевые собственные значения матрицы A_0 . Оставшиеся $n - r$ собственных значений матрицы A_0 равны 0. Матрица A_0 допускает каноническое (спектральное) разложение на произведение 3-х матриц [4, 5]:

$$A_0 = U_{n \times n} \Lambda_{n \times n} U_{n \times n}^T = U_{n \times r} \Lambda_{r \times r} U_{r \times n}^T, \quad (5)$$

где $\Lambda_{n \times n} = \text{diag}(\lambda_1, \dots, \lambda_n)$, $\Lambda_{r \times r} = \text{diag}(\lambda_1, \dots, \lambda_r)$ – диагональные матрицы собственных значений, $U_{n \times n} = (u_1 \mid \dots \mid u_n)$, $U_{n \times r} = (u_1 \mid \dots \mid u_r)$ – ортогональные матрицы нормированных собственных векторов матрицы A_0 , соответствующие собственным значениям, $U_{n \times n}^T U_{n \times n} = E_{n \times n}$ и $U_{r \times n}^T U_{n \times r} = E_{r \times r}$. Заметим, что для положительно полуопределенных матриц 2-нормы и собственные значения связаны соотношением $\|A_0\| = \lambda_1$.

Обозначим через $P^r \subset R^n$ линейное многообразие векторов с базисом $\{u_1, \dots, u_r\}$, а через $P_{\perp}^{n-r} \subset R^n$ подпространство векторов с базисом $\{u_{r+1}, \dots, u_n\}$. P^r и P_{\perp}^{n-r} взаиморто-

гональны в том смысле, что скалярное произведение любого вектора из P^r на любой вектор из P_{\perp}^{n-r} равно 0, причем $P^r \cup P_{\perp}^{n-r} = R^n$.

Из равенства рангов матриц системы A , A_0 и расширенной матрицы системы $(A \mid \mathbf{b})$ следует, что $\mathbf{b}, \mathbf{b}_0, \mathbf{f}_0 \in P^r$, т.е. вектор \mathbf{b} можно представить в виде линейной комбинации базисных векторов

$$\mathbf{b} = U_{n \times r} \boldsymbol{\beta}, \quad (6)$$

где $\boldsymbol{\beta}$ – матрица-столбец координат вектора \mathbf{b} в базисе $\{\mathbf{u}_1, \dots, \mathbf{u}_r\}$.

Будем считать, что все собственные векторы матрицы A , соответствующие ненулевым собственным значениям, принадлежат подпространству P^r .

Требуется найти приближение к нормальному псевдорешению $\mathbf{x}^* \approx \mathbf{x}^0 = A^+ \mathbf{b}$ системы уравнений (4), если известна псевдообратная матрица $H_0 = A_0^+$. Приближение к решению рассматриваем в среднеквадратическом смысле, т.е. считаем, что в точке $\mathbf{x}^* \approx \mathbf{x}^0$ выполняется неравенство

$$(\mathbf{A}\mathbf{x}^* - \mathbf{b})^T (\mathbf{A}\mathbf{x}^* - \mathbf{b}) < \varepsilon_{абс}^2,$$

где $\varepsilon_{абс}$ – допустимое абсолютное отклонение. Особенность поставленной задачи заключается в том, что в алгоритме нахождения вектора \mathbf{x}^* используется матрица H_0 в качестве начального приближения к псевдообратной матрице A^+ .

Заметим, что, если известно разложение (5), то псевдообратную матрицу можно найти по формуле [4]:

$$A_0^+ = U_{n \times r} \Lambda_{r \times r}^{-1} U_{r \times n}^T, \quad (7)$$

где $\Lambda_{r \times r}^{-1} = \text{diag}(\lambda_1^{-1}, \dots, \lambda_r^{-1})$. Кроме того, нормальное псевдорешение $\mathbf{x}^0 \in P^r$, т.е. принадлежит многообразию P^r , что следует из следующей формулы, полученной последовательным применением формул (6) и (7):

$$\mathbf{x}^0 = A^+ \mathbf{b} = U_{n \times r} \Lambda_{r \times r}^{-1} U_{r \times n}^T U_{n \times r} \boldsymbol{\beta} = U_{n \times r} \boldsymbol{\beta}', \quad (8)$$

где $\boldsymbol{\beta}' = \Lambda_{r \times r}^{-1} \boldsymbol{\beta}$ – матрица-столбец координат вектора \mathbf{x}^0 в базисе $\{\mathbf{u}_1, \dots, \mathbf{u}_r\}$.

Дополнительно, легко показать, что

$$\mathbf{x}^0 = A^+ \mathbf{A}\mathbf{x}^0 = \mathbf{A}\mathbf{A}^+ \mathbf{x}^0. \quad (9)$$

Например,

$$A^+ \mathbf{A}\mathbf{x}^0 = U_{n \times r} \Lambda_{r \times r}^{-1} U_{r \times n}^T U_{n \times r} \Lambda_{r \times r} U_{r \times n} U_{n \times r} \boldsymbol{\beta}' = \mathbf{x}^0.$$

2. Метод решения

Как уже было сказано выше, нормальное псевдорешение системы линейных уравнений (4) есть решение минимальной длины задачи минимизации квадратичной функции

$$f(\mathbf{x}) = \frac{1}{2} (\mathbf{A}\mathbf{x} - \mathbf{b})^2.$$

Введем обозначения:

$\nabla f(\mathbf{x}_k) = \mathbf{r}_k = \mathbf{A}\mathbf{x}_k - \mathbf{b}$ – градиент $f(\mathbf{x})$ в точке \mathbf{x}_k , $\mathbf{s}_k = \mathbf{x}_{k+1} - \mathbf{x}_k$ – направление спуска,

$$\begin{aligned} \mathbf{y}_k &= \mathbf{r}_{k+1} - \mathbf{r}_k = \nabla f(\mathbf{x}_{k+1}) - \nabla f(\mathbf{x}_k) = \\ &= \mathbf{A}(\mathbf{x}_{k+1} - \mathbf{x}_k) = \mathbf{A}\mathbf{s}_k. \end{aligned}$$

Пусть матрица A (матрица Гессе функции f) положительно определена. В любом методе переменной метрики (квазиньютоновском методе) последовательность приближений $\{\mathbf{x}_k\}$ к точке минимума задается следующей процедурой:

$$\begin{aligned} \mathbf{s}_k &= -\alpha_k^* A_k^{-1} \mathbf{r}_k = -\alpha_k^* H_k \mathbf{r}_k, \\ \mathbf{x}_{k+1} &= \mathbf{x}_k + \mathbf{s}_k, \end{aligned}$$

где $H_k = A_k^{-1}$ оценка обратной матрицы A^{-1} в точке \mathbf{x}_k . Для сходимости метода длина шага α_k^* в направлении вектора \mathbf{s}_k должна удовлетворять условиям Вульфа [6, 7]. Обычно выбирают $\alpha_k^* = \arg \min_{\alpha \in R} f(\mathbf{x}_k + \alpha H_k \mathbf{r}_k)$, т.е. как решение задачи одномерной минимизации $f(\mathbf{x})$ в направлении \mathbf{s}_k . Разные методы переменной метрики различаются формулами пересчета матриц H_k или A_k . При этом гарантируется при определенных условиях сходимость последовательности $\{H_k\}$ к A^{-1} .

В квазиньютоновских алгоритмах, основанных на SR1-формуле [8–31], обновление матриц H_k выполняется по формуле

$$H_{k+1} = H_k + \frac{(\mathbf{s}_k - H_k \mathbf{y}_k)(\mathbf{s}_k - H_k \mathbf{y}_k)^T}{(\mathbf{s}_k - H_k \mathbf{y}_k)^T \mathbf{y}_k}. \quad (10)$$

Существуют исследования [8–10], которые показывают, что в ряде неквадратичных экстремальных задач SR1-метод демонстрирует более высокие результаты, чем другие, более сложно устроенные методы переменной метрики (такие как BFGS, DFP).

Но формула (10) имеет и свои недостатки, которые отмечаются рядом авторов [11–14]. В частности, к недостаткам относится возможность обращения в ноль знаменателя

[15] и сложность процедур предотвращения подобного явления.

В настоящее время продолжают исследования, направленные на улучшение этого метода. Так, в работах [8–10] описываются свойства локальной и глобальной сходимости метода при решении различных задач. Ряд его модификаций предлагается в [17–29].

Применим формулу (10) для решения поставленной задачи решения систем линейных уравнений с симметричными положительно полуопределенными матрицами. Для этого сделаем в SR1-методе следующие преобразования. Зафиксируем параметр $\alpha_k^* = 1$. Тогда $s_k - H_k y_k = -H_k r_{k+1}$. В соответствии с этим преобразуем формулу (10) пересчета матриц H_k и переставим местами порядок вычисления H_k и s_k . В результате получим следующий алгоритм.

Алгоритм:

```

 $H_0 = A_0^+$ ,  $x_1 = s_0 = H_0 b$ ,  $y_0 = Ax_1$ ,  $r_1 = y_0 - b$ ,  $k = 0$ 
while  $r_{k+1}^T r_{k+1} > \varepsilon_{\text{доп}}$ 
   $s = H_k r_{k+1}$ 
  if  $|s^T y_k| > \varepsilon_{\text{доп}} |s| \cdot |y_k|$ 
     $s_{k+1} = - \left( 1 - \frac{(H_k r_{k+1})^T r_{k+1}}{(H_k r_{k+1})^T y_k} \right) H_k r_{k+1}$ 
     $H_{k+1} = H_k - \frac{(H_k r_{k+1})(H_k r_{k+1})^T}{(H_k r_{k+1})^T y_k}$ 
  else
     $s_{k+1} = s$ ,  $H_{k+1} = H_k$ 
  end if
   $k = k + 1$ 
   $x_{k+1} = x_k + s_k$ 
   $r_{k+1} = Ax_{k+1} - b$ ,  $y_k = r_{k+1} - r_k$ 
end while
 $x^0 = x_{k+1}$ ,  $A^+ = H_k$ .

```

В дальнейшем покажем, что указанная модификация позволяет сохранить для СЛАУ с симметричными положительно полуопределенными матрицами системы основное свойство методов переменной метрики – сходимость к решению за конечное число шагов. Причем, если известно начальное приближение $H_0 = A_0^+$, то число шагов в данном методе не превосходит ранга матрицы возмущений F системы линейных уравнений (4). В результате этого, в применении к поставленной

во введении задачи многократного разрешения уравнений движения механических систем относительно ускорений, возникающей в процессе их численного интегрирования, данный метод сходится быстрее других методов, в которых не вычисляется псевдообратная матрица системы или не используется начальная оценка этой матрицы.

Кроме того, на каждой итерации алгоритма (11) выполняется только две операции умножения матрицы на вектор, требующие порядка n^2 арифметических операций. В SR1-методе вычисляются три таких произведения. Также в SR1-методе требуются дополнительные операции для вычисления длины шага α_k^* .

Конечно, сокращение числа арифметических операций не прошло бесследно. Ниже показано, что это привело к добавлению еще одной итерации для достижения псевдонормального решения СЛАУ.

Свойства алгоритма (11) в применении к решению СЛАУ с симметричными положительно определенными матрицами системы рассмотрены в работе [30].

Пусть теперь A – симметричная и положительно полуопределенная матрица ранга $r \leq n$. В первую очередь заметим, что вектор s_k является нормальным псевдорешением системы уравнений $As_k = y_k$, т.е. $s_k = A^+ y_k$.

Покажем, что предлагаемая в (11) схема обеспечивает линейную независимость векторов s_k и сходимость последовательности $\{x_k\}$ к решению СЛАУ (4) за конечное число итераций.

Докажем пять вспомогательных лемм.

Лемма 1. Если $x_k \in P^r$, $s_k \in P^r$, то $y_k \in P^r$.

Доказательство

$x_{k+1} = x_k + s_k \in P^r$ в силу линейности многообразия P^r . С использованием формул (5), (6) и (8) построим следующие последовательности равенств ($i = k$ или $k + 1$):

$$r_i = Ax_i - b = (U_{n \times r} \Lambda_{r \times r} U_{r \times n}^T) U_{n \times r} \beta'_i - U_{n \times r} \beta = U_{n \times r} (\Lambda_{r \times r} \beta'_i - \beta) = U_{n \times r} \beta_i,$$

где $\beta_i = \Lambda_{r \times r} \beta'_i - \beta$ – матрица-столбец координат вектора r_i в базисе $\{u_1, \dots, u_r\}$.

Таким образом, $r_k, r_{k+1} \in P^r$, $y_k = r_{k+1} - r_k \in P^r$. Лемма доказана.

Дополнительно из леммы 1 следует, что все векторы s_k, y_k , генерируемые алгоритмом (11), принадлежат линейному многообразию P^r . Это можно обосновать следующей последовательностью рассуждений.

Разложение (7) представим в виде

$$\begin{aligned} H_0 &= A_0^+ = U_{n \times r} \Lambda_{r \times r}^{-1} U_{r \times n}^T = \\ &= \sum_{i=1}^r \lambda_i^{-1} u_i u_i^T + 0 \sum_{i=r+1}^n u_i u_i^T. \end{aligned}$$

Отсюда, для любого $x \in P^r$ следует, что $H_0 x \in P^r$ и для любого $x \in P_{\perp}^{n-r}$ следует, что $H_0 x = 0$. Из условия задачи $b \in P^r$. Тогда, проводя вычисления по алгоритму (11), получим, что $x_1, s_0, y_0, s_1, r_1, H_0 r_1 \in P^r$. Так как матрица H_1 формируется из матриц H_0 прибавлением к ней матрицы диадного произведения вектора $H_0 r$ самого на себя, то матрица H_1 обладает теми же проективными свойствами, что и матрица H_0 . Можно продолжить аналогичные рассуждения на последующих шагах алгоритма (11).

Таким образом, получаем, что все векторы $s_k, y_k \in P^r$.

Следующие леммы 2–5 в работе [30] доказаны для положительно определенных СЛАУ. Здесь доказательства лемм обобщены на случай положительно полуопределенных СЛАУ.

Лемма 2. Матрицы H_{k+1} ($k = \overline{0, r-1}$), вычисляемые по формулам (11), удовлетворяют квазиньютоновским условиям

$$H_{k+1} y_i = A^+ y_i = s_i, \quad i = \overline{0, k}. \quad (12)$$

Доказательство

При $i = k$ ($k = \overline{0, r-1}$) выполнение условия (12) можно проверить непосредственной подстановкой (11) в (12):

$$H_{k+1} y_k = H_k y_k - \frac{(H_k r_{k+1})(H_k r_{k+1})^T y_k}{(H_k r_{k+1})^T y_k} = \quad (13)$$

$$= H_k y_k - H_k r_{k+1} = H_k y_k + s_k - H_k y_k = A^+ y_k.$$

Дальнейшее доказательство проводим методом математической индукции. Пусть

$$H_k y_i = A^+ y_i, \quad i = \overline{0, k-1}.$$

Тогда из формулы (11) после умножения H_{k+1} на y_i получим

$$\begin{aligned} (H_k r_{k+1})^T y_i &= (s_k - H_k y_k)^T y_i = \\ &= s_k^T y_i - y_k^T H_k y_i = (A^+ y_k)^T y_i - y_k^T A^+ y_i = 0. \end{aligned}$$

Следовательно, $H_{k+1} y_i = H_k y_i = A^+ y_i$, $i = \overline{0, k-1}$.

Лемма доказана.

Лемма 3. Последовательность направлений поиска $\{s_k\}$ удовлетворяет условиям

$$s_{k+1} = \alpha_k (E - H_k A) s_k, \quad (k = 0, 1, \dots, r-2). \quad (14)$$

Доказательство

Для начала напомним, что $y_k = A s_k$.

Тогда из формул (11) следует цепочка равенств:

$$\begin{aligned} s_{k+1} &= - \left(1 - \frac{(H_k r_{k+1})^T r_{k+1}}{(H_k r_{k+1})^T y_k} \right) H_k r_{k+1} = \\ &= \alpha_k (s_k - H_k y_k) = \alpha_k (E - H_k A) s_k, \end{aligned}$$

$$\text{где } \alpha_k = - \left(1 - \frac{(H_k r_{k+1})^T r_{k+1}}{(H_k r_{k+1})^T y_k} \right).$$

Лемма доказана.

Лемма 4. Векторы s_k и s_{k+1} ($k = 0, 1, \dots, r-2$) линейно независимы.

Доказательство

Сначала покажем, что вектор $s_{k+1}^* = -H_{k+1} r_{k+1}^*$, где $r_{k+1}^* = A x_{k+1}^* - b$, $x_{k+1}^* = x_k + s_k^*$, $s_k^* = \alpha^* H_k r_k = \alpha^* s_k = x_{k+1}^* - x_k$, $\alpha^* = \arg \min_{\alpha \in R} f(x_k + \alpha H_k r_k)$, можно представить в виде линейной комбинации векторов s_k и s_{k+1} :

$$\begin{aligned} s_{k+1}^* &= -H_{k+1} (A x_{k+1}^* - b) = -H_{k+1} (A s_k^* + r_k) = \\ &= -H_{k+1} (A s_k^* + r_{k+1} - y_k) = \\ &= -\alpha^* H_{k+1} A s_k - H_{k+1} r_{k+1} + H_{k+1} y_k = \\ &= (1 - \alpha^*) H_{k+1} y_k + s_{k+1} = (1 - \alpha^*) s_k + s_{k+1}. \end{aligned}$$

Но вектор s_{k+1}^* A -сопряжен с s_k , т.к.

$$\begin{aligned} s_k^T A s_{k+1}^* &= -s_k^T A H_{k+1} r_{k+1}^* = -(H_{k+1} A s_k)^T r_{k+1}^* = \\ &= -(H_{k+1} y_k)^T r_{k+1}^* = -s_k^T \nabla f(x_{k+1}^*) = 0. \end{aligned}$$

Равенство нулю имеет место в силу того, что в точке x_{k+1}^* $\nabla f(x_{k+1}^*) \perp s_k$. Сопряженные направления линейно независимы [4].

Следовательно, используя доказанное выше свойство, что вектор s_{k+1}^* представим в виде линейной комбинации векторов s_k и s_{k+1} , приходим к заключению, что $E - H_k A \neq \lambda E$ и векторы s_k и s_{k+1} линейно независимы. Лемма доказана.

Лемма 5. Векторы s_k , $k = \overline{2, r-1}$ удовлетворяют условиям A -сопряженности

$$s_i^T A s_k = 0, \quad i = \overline{0, k-2}. \quad (15)$$

Доказательство

Равенства (15) доказываются последовательным применением соотношений (14) и (12) с учетом равенств (9):

$$\begin{aligned} s_i^T A s_k &= y_i^T s_k = \alpha_{k-1} y_i^T (E - H_{k-1} A) s_{k-1} = \\ &= \alpha_{k-1} [(E - A H_{k-1}) y_i]^T s_{k-1} = \\ &= \alpha_{k-1} (y_i - A A^+ y_i)^T s_{k-1} = 0. \end{aligned}$$

Лемма доказана.

Из лемм 4 и 5 следует, что все направления поиска s_k , $k = \overline{0, r-1}$ линейно независимы, следовательно, векторы y_k , $k = \overline{0, r-1}$ также линейно независимы.

Из леммы 2 следует, что

$$H_r y_i = A^+ y_i, \quad \forall i = \overline{0, r-1}. \quad (16)$$

Из рассуждений, выполненных после леммы 1, следует, что, с одной стороны y_i , $H_r y_i \in P^r$, а с другой, для любого собственного вектора $u_i \in P_{\perp}^{n-r}$ $H_r u_i = A^+ u_i = 0$. Построим матрицу $X_{n \times n} = (y_0 \dots y_{r-1} \ u_{r+1} \dots u_n)$. Тогда равенства (16) можно записать в виде

$$H_r X_{n \times n} = A^+ X_{n \times n}.$$

Матрица $X_{n \times n}$ по построению имеет полный ранг. Отсюда следует, что $H_r = A^+$.

Таким образом, если все вычисления проводились точно, последний шаг алгоритма

$$x_{r+1} = x_r - H_r r_r = x_r - A^{-1} r_r$$

приведет нас в точку x^0 , являющуюся нормальным псевдорешением системы уравнений (4). Заметим, что в точной арифметике для достижения точки x^0 в поставленной задаче решения СЛАУ с симметричной положительно полуопределенной матрицей системы достаточно, чтобы $H_0 = A_0^+$.

Мы доказали, что рассматриваемый метод является конечно сходящимся. Однако для того чтобы использовать его в качестве итерационного метода решения СЛАУ, необходимо обеспечить монотонность убывания последовательности погрешностей решения $\{\|x_k - x^*\|\}$. В этом случае решение в среднеквадратическом смысле может быть достигнуто за меньше, чем $r+1$ итерацию k ($k < r+1$).

Найдем, при каких условиях метод обеспечивает монотонную сходимость последовательностей $\{H_k\}$ и $\{x_k\}$ к A^+ и x^0 соответственно.

Введем обозначения:

$$A_k = H_k^+, \quad H_k = A_k^+,$$

$$\varepsilon F_k = A - A_k, \quad \varepsilon \Delta H_k = H_k - H,$$

где ε – малый параметр, εF_k – симметричная матрица возмущений ($\varepsilon F_0 = \varepsilon F$).

Обозначим через $\kappa(A_k) = \|A_k\| \cdot \|A_k^+\| = \|A_k\| \cdot \|H_k\| = \frac{\lambda_1}{\lambda_r}$ – эффективное число обусловленности матрицы A_k .

Сделаем предварительные выкладки. Используя свойство псевдообратных матриц, зафиксированное в равенстве (9), и матричный ряд Тейлора [4] для матрицы полного ранга $(E + \varepsilon A_k^+ F_k)^{-1}$, можно получить следующую последовательность равенств:

$$\begin{aligned} H &= A^+ = (A_k + \varepsilon F_k)^+ = \\ &= (A_k (E + \varepsilon A_k^+ F_k))^{+} = (E + \varepsilon A_k^+ F_k)^{-1} A_k^+ = \\ &= (E - \varepsilon H_k F_k + \varepsilon^2 (H_k F_k)^2 - \dots) H_k = \\ &= H_k - (E - \varepsilon H_k F_k + \dots) \varepsilon H_k F_k H_k = \\ &= H_k - (E + \varepsilon H_k F_k)^{-1} \varepsilon H_k F_k H_k. \end{aligned}$$

Таким образом, возмущения прямой и обратной матрицы связаны следующим матричным соотношением:

$$\varepsilon \Delta H_k = H_k - H = (E + \varepsilon H_k F_k)^{-1} \varepsilon H_k F_k H_k. \quad (17)$$

Из равенства (17) следует, что, если матрицы H_k , F_k положительно полуопределены и имеют ранг $r \leq n$, а $E + \varepsilon H_k F_k$ – положительно определенная матрица, то ΔH_k также положительно полуопределенная ранга $r \leq n$ матрица как произведение симметричных определенной и полуопределенной матриц.

Это свойство доказывает следующая лемма.

Лемма 6. Пусть $C = C_{n \times n}$ – симметричная определенная матрица ранга n и $D = D_{n \times n}$ – симметричная полуопределенная матрица ранга $r \leq n$. Тогда CD – симметричная полуопределенная матрица ранга $r \leq n$.

Доказательство

В соответствии с (5) матрицы C и D допускают канонические разложения на произведение 3-х матриц:

$$C = U_{n \times n} \Lambda_{n \times n}^C U_{n \times n}^T, \quad D = V_{r \times r} \Lambda_{r \times r}^D V_{r \times r}^T,$$

где $\Lambda^C = \Lambda_{n \times n}^C$, $\Lambda^D = \Lambda_{r \times r}^D$ – диагональные матрицы положительных собственных значений, $U = U_{n \times n}$, $V = V_{r \times r}$ – ортогональные матрицы нормированных собственных векторов u_i , v_i . Для простоты изложения будем считать, что первые r векторов u_i , v_i принадлежат линейному многообразию P^r . Пусть $W = U^T V$ – ортогональная матрица координат w_i векторов v_i в базисе $\{u_i\}$. Тогда матрица

$$\begin{aligned} CD &= U \Lambda^C U^T V \Lambda^D V^T = \\ &= U \Lambda^C (W \Lambda^D W^T) U^T \end{aligned}$$

положительно полуопределена, если $\Lambda^C (W \Lambda^D W^T)$ – положительно полуопределенная матрица. Разложим $\Lambda^C (W \Lambda^D W^T)$ на сумму матриц ранга 1:

$$\Lambda^C (W \Lambda^D W^T) = \sum_{i=1}^r (\Lambda^C \lambda_i^D) w_i w_i^T.$$

Матрицы $\Lambda^C \lambda_i^D$ положительно определены. Легко видеть, что

$$w_k^T \Lambda^C (W \Lambda^D W^T) w_k = w_k^T (\Lambda^C \lambda_k^D) w_k > 0$$

для всех $k = \overline{1, r}$ и

$$u_k^T \Lambda^C (W \Lambda^D W^T) u_k = u_k^T \sum_{i=1}^r (\Lambda^C \lambda_i^D) w_i (w_i^T u_k) = 0$$

для всех $k = \overline{r+1, n}$. Отсюда следует положительная полуопределенность матрицы CD в пространстве R^n и положительная определенность на линейном многообразии P^r . Лемма доказана.

Найдем явную зависимость матрицы H_{k+1} от матриц возмущений εF_k , $\varepsilon \Delta H_k$.

Для этого заметим, что

$$\begin{aligned} H_k r_{k+1} &= H_k (r_k + y_k) = -s_k + H_k A s_k = (H_k A - E) s_k = \\ &= ((H_k + \varepsilon \Delta H_k) A - E) s_k = \varepsilon \Delta H_k A s_k, \\ H_k r_{k+1} &= (H_k A - E) s_k = (H_k (A_k + \varepsilon F_k) - E) s_k = \\ &= -H_k A_k H_k r_k + \varepsilon H_k F_k s_k + H_k r_k = \varepsilon H_k F_k s_k. \end{aligned}$$

Здесь использовано свойство псевдообратных матриц $H_k A_k H_k = H_k$.

Дополнительно заметим, что $r_k \in P^r$, и используя свойство (9), можно получить обратное равенство:

$$r_{k+1} = \varepsilon F_k s_k. \quad (18)$$

Кроме того,

$$(H_k r_{k+1})^T y_k = s_k^T (\varepsilon A \Delta H_k A) s_k.$$

Подставим полученные соотношения в формулу (11) для H_{k+1} . Окончательно придем к следующему выражению

$$H_{k+1} = H_k - \varepsilon \frac{(\Delta H_k A s_k)(\Delta H_k A s_k)^T}{s_k^T (A \Delta H_k A) s_k}. \quad (19)$$

Кроме того, из теории известно [31, 32], что любой метод переменной метрики может быть сформулирован как для обратной матрицы $H = A^{-1}$, так и для прямой исходной матрицы системы $A = A_0 + \varepsilon F_0$. Алгоритм пересчета матрицы приближений A_k к прямой матрице A можно получить из формул (10) заменой: $H_k \rightarrow A_k$, $s_k \rightarrow y_k$, $y_k \rightarrow s_k$:

$$\begin{aligned} A_{k+1} &= A_k + \frac{(y_k - A_k s_k)(y_k - A_k s_k)^T}{(y_k - A_k s_k)^T s_k} = \\ &= A_k + \frac{r_{k+1} r_{k+1}^T}{r_{k+1}^T s_k} = A_k + \varepsilon \frac{(F_k s_k)(F_k s_k)^T}{s_k^T F_k s_k}. \end{aligned} \quad (20)$$

Здесь дополнительно использованы равенства (18). Легко проверить непосредственной подстановкой, что, если $H_0 = A_0^+$, то последовательности приближений удовлетворяют условию взаимной обратимости матриц H_{k+1} и A_{k+1} :

$$A_{k+1} H_{k+1} A_{k+1} = A_{k+1}, \quad H_{k+1} A_{k+1} H_{k+1} = H_{k+1}$$

и квазинытоновским условиям (см. лемму 1 и формулы (12) и (9)):

$$H_{k+1} y_i = s_i = A^+ y_i, \quad (21)$$

$$A_{k+1} s_i = y_i = A s_i, \quad i = \overline{0, k}.$$

Теперь сформулируем и докажем достаточное условие монотонности убывания последовательности погрешностей решения $\{\|x_k - x^0\|\}$.

Теорема. Если выполнены все условия, зафиксированные в постановке задачи, и

$$\|\varepsilon F_0\| < \frac{1}{\kappa(A_0)} \|A_0\| = \frac{1}{\|H_0\|}, \quad (22)$$

то алгоритм (11) обеспечивает выполнение следующих свойств:

1. $\{H_k\}$ – последовательность положительно определенных матриц по отношению к любым векторам $\eta \in P^r$, т.е. $\eta^T H_k \eta > 0$ для всех $\eta \in P^r$, $k = \overline{1, r}$.

2. Последовательности $\{\|\varepsilon F_k\|\}$, $\{\|H - H_k\|\}$ и $\{\|x_k - x^0\|\}$ монотонно убывают.

Доказательство

Из условия (22) следует, что

$$\|\varepsilon F_0\| < \frac{1}{\kappa(A_0)} \|A_0\| = \min_{\lambda(A_0)} \lambda = \lambda_r.$$

То есть спектральный радиус матрицы εF_0 меньше, чем минимальное собственное значение матрицы A_0 . Используя это неравенство и каноническое (спектральное) разложение матриц легко показать (аналогичными рассуждениями, приведенными в лемме б), что, если матрицы F и A_0 имеют ранг r и положительно определены на линейном многообразии P^r , то минимальное ненулевое собственное значение матрицы $A_0 + \varepsilon F_0$ положительно и

$$\eta^T A \eta = \eta^T (A_0 + \varepsilon F_0) \eta > 0$$

для всех $\eta \neq 0 \in P^r$ и $\eta^T A \eta = 0$ для всех $\eta \in R^n \setminus P^r$, т.е. условие (22) является достаточным условием положительной определенности на P^r матрицы A ранга r при любом знаке малого параметра ε . Кроме того, в силу неравенства (22), положительной полуопределенности матрицы $H_0 F_0$ (по лемме б) и свойств 2-норм матриц, можно выписать следующую цепочку неравенств:

$$1 > \|H_0\| \cdot \|\varepsilon F_0\| \geq \|\varepsilon H_0 F_0\| = |\varepsilon| \cdot \max_{\lambda(H_0 F_0)} \lambda = |\varepsilon| \cdot \lambda_{\max},$$

где $\lambda_{\max} > 0$ – максимальное собственное значение матрицы $H_0 F_0$.

Следовательно, повторяя рассуждения, приведенные при доказательстве положительной определенности матрицы $A_0 + \varepsilon F_0$ на линейном многообразии P^r , заключаем, что $\eta^T (E - |\varepsilon| H_0 F_0) \eta > 0$ для всех $\eta \neq 0 \in P^r$.

Это означает, что матрица $E + \varepsilon H_0 F_0$ положительно определена на P^r при любом знаке малого параметра ε .

Отсюда следует, используя равенство (17), что матрица ΔH_0 также является положительно определенной на P^r .

Дальнейшее доказательство можно провести методом математической индукции. Сначала проверим справедливость теоремы при $k = 1$.

Заметим, что по доказанным леммам все направления поиска s_k , $k = \overline{0, r-1}$ принадлежат линейному многообразию P^r и линейно независимы. Следовательно, любое подмножество этих векторов можно рассматривать в качестве базиса некоторого линейного многообразия. Обозначим через P^{r-k} – линейное многообразие с базисом $\{s_r, s_{r-1}, \dots, s_k\}$. Причем $P^1 \subset P^2 \subset \dots \subset P^r$.

Докажем положительную определенность матрицы H_1 по отношению к любому вектору $\eta \in P^r$, используя равенство (19).

Нам необходимо показать, что для всех $\eta \neq 0 \in P^r$ справедливо неравенство

$$\eta^T H_1 \eta = \eta^T H_0 \eta - \varepsilon \frac{\eta^T (\Delta H_0 A s_k) (\Delta H_0 A s_k)^T \eta}{s_0^T (A \Delta H_0 A) s_0} > 0. \quad (23)$$

По доказанным выше условиям знаменатель в формуле (23) положителен.

Тогда, если $\varepsilon < 0$, то

$$\eta^T H_1 \eta = \eta^T H_0 \eta + |\varepsilon| \frac{u^2}{s_0^T (A \Delta H_0 A) s_0} > 0,$$

где $u = s_0^T A \Delta H_0 \eta$.

Следовательно, в этом случае матрица H_1 положительно определена на P^r .

Пусть теперь $\varepsilon > 0$. Приведем неравенство (23) к общему знаменателю и исследуем знак выражения, получаемого в числителе

$$\begin{aligned} f &= (\eta^T H_0 \eta) (s_0^T A \Delta H_0 A s_0) - \varepsilon u^2 = \\ &= (\eta^T H \eta) (s_0^T A \Delta H_0 A s_0) + \\ &+ \varepsilon (\eta^T \Delta H_0 \eta) (s_0^T A \Delta H_0 A s_0) - \\ &\varepsilon \eta^T (\Delta H_0 A s_0) (\Delta H_0 A s_0)^T \eta. \end{aligned}$$

В силу симметрии и того, что $\boldsymbol{\eta}^T \Delta H_0 \boldsymbol{\eta} > 0$ для всех $\boldsymbol{\eta} \in P^r$ и $\boldsymbol{\eta}^T \Delta H_0 \boldsymbol{\eta} = 0$ для всех $\boldsymbol{\eta} \in R^n \setminus P^r$, матрицу ΔH_0 можно разложить (с помощью спектрального разложения) на произведения двух матриц [4, 5]:

$$\begin{aligned} \Delta H_0 &= V_{n \times r} \Lambda_{r \times r} V_{r \times n}^T = (V \Lambda^{1/2}) (\Lambda^{1/2} V^T) = \\ &= C_{n \times r} C_{r \times n}^T = C C^T. \end{aligned}$$

Тогда функцию f можно представить в виде:

$$\begin{aligned} f &= (\boldsymbol{\eta}^T H \boldsymbol{\eta}) (C^T A s_0)^2 + \varepsilon (C^T \boldsymbol{\eta})^2 (C^T A s_0)^2 - \\ &- \varepsilon \left[(C^T \boldsymbol{\eta})^T (C^T A s_0) \right]^2. \end{aligned}$$

Выражение

$$(C^T \boldsymbol{\eta})^2 (C^T A s_0)^2 - \left[(C^T \boldsymbol{\eta})^T (C^T A s_0) \right]^2 \geq 0$$

в силу справедливости для всех $\mathbf{a}, \mathbf{b} \in R^n$ неравенства Коши–Буняковского $\mathbf{a}^2 \mathbf{b}^2 \geq (\mathbf{a}^T \mathbf{b})^2$.

Таким образом, окончательно получаем следующую цепочку неравенств:

$$\begin{aligned} f &\geq (\boldsymbol{\eta}^T H \boldsymbol{\eta}) (C^T A s_0)^2 > 0, \text{ для всех } \boldsymbol{\eta} \neq 0 \in P^r \\ &\text{и } s_0 \neq 0, \text{ доказывающих требуемые неравенства } \boldsymbol{\eta}^T H_1 \boldsymbol{\eta} > 0 \text{ для всех } \boldsymbol{\eta} \in P^r. \end{aligned}$$

Теперь покажем, что $\boldsymbol{\eta}^T F_1 \boldsymbol{\eta} \geq 0$ для всех $\boldsymbol{\eta} \in P^r$ и $\boldsymbol{\eta}^T F_1 \boldsymbol{\eta} > 0$ для всех $\boldsymbol{\eta} \in P^{r-1}$. Используя равенства (20), выведем явную зависимость матрицы возмущений εF_1 от εF_0 :

$$\begin{aligned} \varepsilon F_1 &= A - A_1 = (A - A_0) - (A_1 - A_0) = \\ &= \varepsilon \left(F_0 - \frac{F_0 s_0 (F_0 s_0)^T}{s_0^T F_0 s_0} \right). \end{aligned} \quad (24)$$

Во-первых, $\varepsilon F_1 s_0 = \varepsilon (F_0 s_0 - F_0 s_0) = 0$, т.е. для всех $\boldsymbol{\eta} \in P^r \setminus P^{r-1}$ выполняется $\boldsymbol{\eta}^T F_1 \boldsymbol{\eta} = 0$. Во-вторых, матрица F_0 является симметричной положительно определенной матрицей полного ранга на линейном многообразии P^r и ее можно разложить на произведение двух матриц (аналогично матрице H_0):

$$F_0 = \tilde{C}_{n \times r} \tilde{C}_{r \times n}^T = \tilde{C} \tilde{C}^T.$$

Тогда при любом $\boldsymbol{\eta} \in P^{r-1}$ равенство (24) можно преобразовать к виду:

$$\begin{aligned} \varepsilon \boldsymbol{\eta}^T F_1 \boldsymbol{\eta} &= \varepsilon \left(\boldsymbol{\eta}^T F_0 \boldsymbol{\eta} - \frac{\boldsymbol{\eta}^T F_0 s_0 (F_0 s_0)^T \boldsymbol{\eta}}{s_0^T F_0 s_0} \right) = \\ &= \varepsilon \frac{(\tilde{C}^T \boldsymbol{\eta})^2 (\tilde{C}^T s_0)^2 - \left((\tilde{C}^T \boldsymbol{\eta})^T (\tilde{C}^T s_0) \right)^2}{s_0^T F_0 s_0}. \end{aligned}$$

Так как векторы $\boldsymbol{\eta}$ и s_0 не параллельны, а матрица \tilde{C} имеет полный ранг на линейном многообразии P^r , то скалярное произведение $(\tilde{C}^T \boldsymbol{\eta})^T (\tilde{C}^T s_0)$ строго меньше произведения длин $|\tilde{C}^T \boldsymbol{\eta}|$ и $|\tilde{C}^T s_0|$. Поэтому выражение

$$(C^T \boldsymbol{\eta})^2 (C^T s_0)^2 - \left((C^T \boldsymbol{\eta})^T (C^T s_0) \right)^2 > 0.$$

Следовательно, $\boldsymbol{\eta}^T F_1 \boldsymbol{\eta} > 0$ при любом $\boldsymbol{\eta} \in P^{r-1}$, т.е. матрица положительно определена на линейном многообразии P^{r-1} и положительно полуопределена на P^r . Кроме того, из этого неравенства следует и неравенство для норм $\|\varepsilon F_1\| < \|\varepsilon F_0\|$.

Дополнительно заметим, что из доказанных свойств матриц F_1 , H_1 , соотношений (17) и леммы 6 следует, что матрица возмущений ΔH_1 должна обладать теми же свойствами, что и матрица возмущений F_1 .

Теперь докажем, что $\|\varepsilon H_1 F_1\| < \|\varepsilon H_0 F_0\|$ и $\|H - H_1\| < \|H - H_0\|$.

Из (19) следует, что

$$\begin{aligned} \varepsilon \Delta H_1 &= H_1 - H = (H_1 - H_0) - (H - H_0) = \\ &= \varepsilon \left(\Delta H_0 - \frac{(\Delta H_0 A s_0) (\Delta H_0 A s_0)^T}{s_0^T (\Delta H_0 A) s_0} \right). \end{aligned} \quad (25)$$

Легко заметить совпадение структур равенств (24) и (25) для матриц возмущений εF_1 и $\varepsilon \Delta H_1$. Поэтому, если повторить для матрицы $\varepsilon \Delta H_1$ все рассуждения, выполненные при доказательстве свойств матрицы εF_1 , то можно окончательно доказать, что $\boldsymbol{\eta}^T \Delta H_1 \boldsymbol{\eta} > 0$ при любом $\boldsymbol{\eta} \in P^{r-1}$, т.е. матрица возмущений ΔH_1 также положительно определена на линейном многообразии P^{r-1} и $\boldsymbol{\eta}^T \Delta H_1 \boldsymbol{\eta} = 0$ для всех $\boldsymbol{\eta} \in P^r \setminus P^{r-1}$.

Кроме того, из равенства (25), по аналогии с равенством (24), следует и неравенство для норм $\|\varepsilon\Delta H_1\| < \|\varepsilon\Delta H_0\|$ или $\|H - H_1\| < \|H - H_0\|$.

Вернемся к равенству (17). Его можно преобразовать к виду $\varepsilon\Delta H_k = \varepsilon H F_k H_k = \varepsilon H_k F_k H$. Подставляя последнее в полученное выше неравенство для норм получаем еще одно требуемое неравенство:

$$\|\varepsilon H_1 F_1\| < \|\varepsilon H_0 F_0\| < 1.$$

Теперь осталось показать, что $\|x_1 - x^0\| < \|x_0 - x^0\|$.

Заметим, что

$$\begin{aligned} x_1 - x^0 &= x_0 - H_0 r_0 - x^0 = \\ &= x_0 - x^0 - H_0 (Ax_0 - b) = \\ &= x_0 - x^0 - H_0 (Ax_0 - Ax^0) = \\ &= (E - H_0 A)(x_0 - x^0). \end{aligned}$$

Следовательно, $\|x_1 - x^0\| < \|x_0 - x^0\|$, если $\|E - H_0 A\| < 1$. Последнее неравенство выполняется, так как в силу (9)

$$\begin{aligned} \|E - H_0 A\|^2 &= \sup_{\eta \neq 0} \frac{\eta^T (E - H_0 (A_0 + \varepsilon F_0))^2 \eta}{\eta^T \eta} = \\ &= \sup_{\eta \neq 0} \frac{\eta^T (E - H_0 A_0)^2 \eta - \eta^T (\varepsilon H_0 F_0)^2 \eta}{\eta^T \eta} = \\ &= \|\varepsilon H_0 F_0\|^2 < 1. \end{aligned}$$

Таким образом, утверждения теоремы доказаны при $k = 1$.

Кроме того, мы попутно доказали следующий ряд утверждений, которые необходимы для дальнейшего доказательства теоремы по индукции. При $k = 1$:

1. $\eta^T F_1 \eta > 0$ при любом $\eta \in P^{r-k}$ и $\eta^T F_1 \eta = 0$ при любом $\eta \in P^r \setminus P^{r-k}$, т.е. матрица положительно определена на линейном многообразии P^{r-k} и положительно полуопределена на P^r .

2. $\|\varepsilon H_1 F_1\| < \|\varepsilon H_0 F_0\| < 1$.

Если теперь принять, что эти два условия справедливы при $j = k$, то очевидно, что мы можем повторить все рассуждения теоремы при $j = k + 1$, $1 \leq k \leq n - 1$.

Таким образом, условие (22) является достаточным условием монотонной сходимости

последовательностей норм $\{\|\varepsilon F_k\|\}$, $\{\|H - H_k\|\}$ и $\{\|x_k - x^0\|\}$. Теорема доказана.

Замечание. Дополнительно из равенства (25) следует то, что в процессе выполнения итераций по формулам (11) знаменатель $(H_k r_{k+1})^T y_k = s_k^T (A \Delta H_k A) s_k$ никогда (кроме последнего шага) не обращается в ноль при соблюдении всех условий, указанных в доказанной теореме и постановке задачи.

Приведенные выше рассуждения позволяют установить еще один факт. Оказывается, что если матрица возмущений εF имеет ранг меньше r , то описываемый итерационный метод сойдется меньше, чем за r итераций.

Пусть матрица возмущений $\varepsilon F = A - A_0$ имеет ранг r_1 ($r_1 \leq r \leq n$). Обозначим через $P^{r_1} \subseteq P^r \subseteq R^n$ линейное многообразие векторов с базисом $\{u_1, \dots, u_{r_1}\}$, состоящим из собственных векторов, соответствующих ненулевым собственным значениям матрицы εF , а через $P_{\perp}^{n-r_1} \subset R^n$ подпространство векторов с базисом $\{u_{r_1+1}, \dots, u_n\}$, включающим собственные векторы, соответствующие нулевым собственным значениям матрицы εF . Тогда $(A - A_0)u_i = \varepsilon F u_i = 0$ и $(A^+ - H_0)u_i = 0$ для всех $u_i \in P_{\perp}^{n-r_1}$.

В этом случае вычисления, выполненные по формулам алгоритма (11), в соответствии с доказанными леммами приведут к тому, что все векторы y_k принадлежат многообразию P^{r_1} , $k = \overline{0, r_1 - 1}$ и линейно независимы. Кроме того, согласно рассмотренной теореме, все матрицы H_k , $k = \overline{1, r_1}$ будут удовлетворять следующим условиям: $\eta^T H_k \eta > 0$ для всех $\eta \in P^{r_1}$, $H_k \eta = H_{k-1} \eta$ для всех $\eta \in P_{\perp}^{n-r_1}$ и квазинытоновским условиям $H_k y_i = A^+ y_i$, $i = \overline{0, k - 1}$.

Отсюда следует, что столбцы матрицы H_{r_1} можно найти из системы линейных уравнений

$$(A^+ - H_{r_1}) X_{n \times n} = 0, \quad (26)$$

с матрицей $X_{n \times n} = (y_0 \dots y_{r_1-1} u_{r_1+1}, \dots, u_n)$.

Матрица системы (26) $X_{n \times n}$ по построению имеет полный ранг. Следовательно, система (26) имеет единственное решение $H_{\dot{r}_1} = A^+$.

Таким образом, можно заключить, что итерационный алгоритм (11) при выполнении всех условий, указанных в постановке задачи, и доказанной теореме, обеспечивает сходимость последовательности приближений $\{x_k\}$ к точному решению x^0 системы уравнений (4) за конечное число шагов, не превышающих ранг матрицы возмущений $\varepsilon F = A - A_0$.

Заметим, что при численном решении уравнений движения механических систем вида (2) обычно на каждом шаге интегрирования эффективный ранг матрицы возмущений много меньше ранга матрицы системы. Поэтому следует ожидать, что приближенное решение системы уравнений (2) рассматриваемым итерационным методом в практических задачах будет достигаться за небольшое число итераций.

4. Пример моделирования механической системы с дополнительными связями

Применение рассмотренного итерационного метода проиллюстрируем на примере моделирования динамики одной механической системы с двумя замкнутыми кинематическими цепями, представляющей собой модель весельной лодки с механическим приводом (рис. 1).



Рис. 1. Модель весельной лодки

Модель включает в себя два весла, закрепленные в уключинах, обеспечивающих угловые перемещения весел в двух плоскостях, и двигатель, задающий вращательное движение, с двумя кривошипами, соединенными сферическими шарнирами с веслами. Система имеет одну степень свободы.

Уравнения движения построим в форме уравнений (2). Для этого расцепим соединения весел с кривошипами. В результате получим расчетную схему системы с пятью степенями свободы, состоящую из трех тел (двух весел и двигателя с кривошипами), представленную на рис. 2. Обобщенные координаты указаны на рис. 1 ($m = 5$).

Порядок обобщенных координат: $q = (\beta, \gamma_2, \alpha_2, \gamma_3, \alpha_3)$.

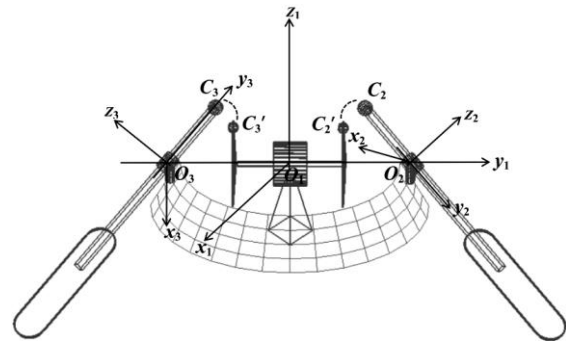


Рис. 2. Расчетная схема лодки

Введем обозначения: r_0 – расстояние между кривошипом и уключиной вдоль Oy_1 и высота кривошипа, $r_h = O_2C_2$ – расстояние вдоль весла от уключины до крепления с кривошипом, $O_1C_2' = (0, r_0, r_0)$, $O_1C_3' = (0, -r_0, r_0)$, $O_1O_2 = 2r_0$, $O_1O_3 = -2r_0$, J_1 – осевой момент инерции механизма вращения с кривошипами, $J_2 = J_3$ – осевой момент инерции весла, $m_2 = m_3$ – масса весла, $r_{2c} = r_{3c}$ – расстояние от уключины до центра тяжести весла.

Вращение кривошипов обеспечивается внешним упруго-демпфирующим моментом:

$$M_\beta = c_\beta \left(\frac{2\pi}{T_\beta} \left(t - k_\beta \frac{T_\beta}{2\pi} \sin \frac{2\pi}{T_\beta} t \right) - \beta \right) + d_\beta \left(\frac{2\pi}{T_\beta} \left(1 - k_\beta \cos \frac{2\pi}{T_\beta} t \right) - \dot{\beta} \right),$$

где T_β – период оборота кривошипа, $2\pi/T_\beta$ – максимальная круговая частота вращения, c_β , d_β – коэффициенты упругости и демпфирования. Сомножители $1 - k_\beta \cos 2\pi t/T_\beta$ и $t - k_\beta T_\beta/2\pi \cdot \sin 2\pi t/T_\beta$ определяют переменную скорость вращения и введены для имита-

ции торможения скорости вращения за счет сил сопротивления при погружении весел в воду, $k_\beta \leq 1$ – коэффициент торможения.

Переход к исходной одностепенной модели будем осуществлять наложением шести дополнительных связей, обеспечивающих совпадение координат точек $C_2 = C_2'$ и $C_3 = C_3'$ соединения весел с кривошипами ($n = 6$). Очевидно, что две из них избыточны.

Уравнения связей имеют вид:

$$\begin{aligned} g_1 &= -r_0 \cdot \sin \beta + r_h \cdot \cos \alpha_2 \cdot \sin \gamma_2 = 0, \\ g_2 &= r_h \cdot \sqrt{1 - \frac{r_0^2}{r_h^2}} - r_h \cdot \cos \alpha_2 \cdot \cos \gamma_2 = 0, \\ g_3 &= -r_0 \cdot \cos \beta - r_h \cdot \sin \alpha_2 = 0, \\ g_4 &= -r_0 \cdot \sin \beta - r_h \cdot \cos \alpha_3 \cdot \sin \gamma_3 = 0, \\ g_5 &= -r_h \cdot \sqrt{1 - \frac{r_0^2}{r_h^2}} + r_h \cdot \cos \alpha_3 \cdot \cos \gamma_3 = 0, \\ g_6 &= -r_0 \cdot \cos \beta + r_h \cdot \sin \alpha_3 = 0. \end{aligned}$$

Компоненты системы дифференциально-алгебраических уравнений вида (2) в нашем случае имеют вид:

$$\begin{aligned} M_{11}^* &= J_1, \quad M_{22}^* = J_2 \cos^2 \alpha_2, \quad M_{33}^* = J_2, \\ M_{44}^* &= J_3 \cos^2 \alpha_3, \quad M_{55}^* = J_3, \quad M_{ij}^* = 0, \quad i \neq j, \\ f_1 &= M_\beta, \quad f_2 = J_2 \cdot \sin 2\alpha_2 \cdot \dot{\gamma}_2 \dot{\alpha}_2, \\ f_3 &= -m_2 g \cdot r_{2c} \cdot \cos \alpha_2 - \frac{1}{2} J_2 \cdot \sin 2\alpha_2 \cdot \dot{\gamma}_2^2, \\ f_4 &= J_4 \cdot \sin 2\alpha_3 \cdot \dot{\gamma}_3 \dot{\alpha}_3, \\ f_5 &= m_3 g \cdot r_{3c} \cdot \cos \alpha_3 - \frac{1}{2} J_3 \cdot \sin 2\alpha_3 \cdot \dot{\gamma}_3^2, \\ G_{11} &= -r_0 \cdot \cos \beta, \quad G_{12} = r_h \cdot \cos \alpha_2 \cdot \cos \gamma_2, \\ G_{13} &= -r_h \cdot \sin \alpha_2 \cdot \sin \gamma_2, \quad G_{14} = G_{15} = 0, \\ G_{22} &= r_h \cdot \cos \alpha_2 \cdot \sin \gamma_2, \quad G_{23} = r_h \cdot \sin \alpha_2 \cdot \cos \gamma_2, \\ G_{21} &= G_{24} = G_{25} = 0, \\ G_{31} &= r_0 \cdot \sin \beta, \quad G_{33} = -r_h \cdot \cos \alpha_2, \\ G_{32} &= G_{34} = G_{35} = 0, \\ G_{41} &= -r_0 \cdot \cos \beta, \quad G_{44} = -r_h \cdot \cos \alpha_3 \cdot \cos \gamma_3, \\ G_{45} &= r_h \cdot \sin \alpha_3 \cdot \sin \gamma_3, \quad G_{42} = G_{43} = 0, \\ G_{54} &= -r_h \cdot \cos \alpha_3 \cdot \sin \gamma_3, \quad G_{55} = -r_h \cdot \sin \alpha_3 \cdot \cos \gamma_3, \\ G_{51} &= G_{52} = G_{53} = 0, \\ G_{61} &= r_0 \cdot \sin \beta, \quad G_{65} = r_h \cdot \cos \alpha_2, \\ G_{62} &= G_{63} = G_{64} = 0, \end{aligned}$$

$$\begin{aligned} h_1 &= -r_0 \cdot \sin \beta \cdot \dot{\beta}^2 + r_h \cdot \cos \alpha_2 \cdot \sin \gamma_2 \cdot (\dot{\gamma}_2^2 + \dot{\alpha}_2^2) + \\ &+ 2r_h \cdot \sin \alpha_2 \cdot \cos \gamma_2 \cdot \dot{\alpha}_2 \dot{\gamma}_2, \\ h_2 &= -r_h \cdot \cos \alpha_2 \cdot \cos \gamma_2 \cdot (\dot{\gamma}_2^2 + \dot{\alpha}_2^2) + \\ &+ 2r_h \cdot \sin \alpha_2 \cdot \sin \gamma_2 \cdot \dot{\alpha}_2 \dot{\gamma}_2, \\ h_3 &= -r_0 \cdot \cos \beta \cdot \dot{\beta}^2 - r_h \cdot \sin \alpha_2 \cdot \dot{\alpha}_2^2, \\ h_4 &= -r_0 \cdot \sin \beta \cdot \dot{\beta}^2 - r_h \cdot \cos \alpha_3 \cdot \sin \gamma_3 \cdot (\dot{\gamma}_3^2 + \dot{\alpha}_3^2) - \\ &- 2r_h \cdot \sin \alpha_3 \cdot \cos \gamma_3 \cdot \dot{\alpha}_3 \dot{\gamma}_3, \\ h_5 &= r_h \cdot \cos \alpha_3 \cdot \cos \gamma_3 \cdot (\dot{\gamma}_3^2 + \dot{\alpha}_3^2) - \\ &- 2r_h \cdot \sin \alpha_3 \cdot \sin \gamma_3 \cdot \dot{\alpha}_3 \dot{\gamma}_3, \\ h_6 &= -r_0 \cdot \cos \beta \cdot \dot{\beta}^2 + r_h \cdot \sin \alpha_3 \cdot \dot{\alpha}_3^2. \end{aligned}$$

В уравнения движения в форме Лагранжа 1 рода вида (2) для их замыкания используются дважды продифференцированные уравнения дополнительных связей $\ddot{\mathbf{g}}(\mathbf{q}) = 0$. Для предотвращения ухода со связей добавим в эти уравнения стабилизирующие слагаемые по методу Баумгарта [2]:

$$\ddot{\mathbf{g}} + 2\beta_* \dot{\mathbf{g}} + k_* \mathbf{g} = 0,$$

где β_*, k_* – коэффициенты демпфирования и жесткости.

Приведем значения всех конструктивных параметров:

$$\begin{aligned} r_0 &= 0.5, \quad r_h = 1, \quad J_1 = 1, \quad m_2 = m_3 = 4, \\ r_{2c} &= r_{3c} = 1.25, \quad J_2 = J_3 = 13, \quad \beta_* = 5, \\ k_* &= 10000, \quad T_\beta = 5, \quad c_\beta = 10000, \quad d_\beta = 100, \\ k_\beta &= 0.7, \quad g = 9.81. \end{aligned}$$

Начальные условия:

$$\begin{aligned} \mathbf{q}(0) &= \left(0, 0, -\frac{\pi}{6}, 0, 0, \frac{\pi}{6} \right), \\ \dot{\mathbf{q}}(0) &= (0, 0, 0, 0, 0, 0). \end{aligned}$$

Отрезок интегрирования $t \in [0, 12]$.

Уравнения движения в форме (2) интегрировались в САВ *Mathematica*. Система линейных уравнений (3) для определения множителей Лагранжа решалась итерационным методом (11). Начальное приближение псевдообратной матрицы $H_0 = A_0^+$ вычислялось встроенной процедурой *PseudoInverse*.

Результаты численного моделирования представлены на рис. За, б, в.

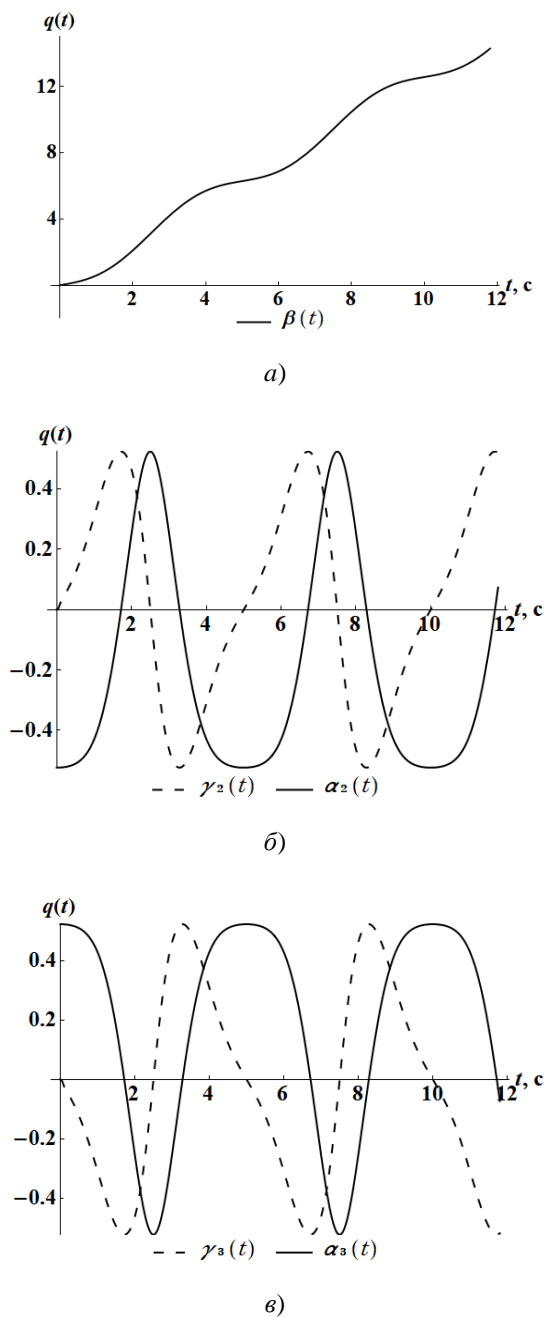


Рис. 3. Результаты моделирования

На графиках угловых перемещений весел хорошо заметно сымитированное замедление их движений при погружении в воду.

На рис. 4, для примера, приведены графики отклонений дополнительных связей между одним веслом и кривошипом.

Графики на рис. 4 показывают, что представленный алгоритм решения задачи обеспечивает хороший уровень точности в удовлетворении дополнительных геометрических связей на длительном промежутке интегрирования.

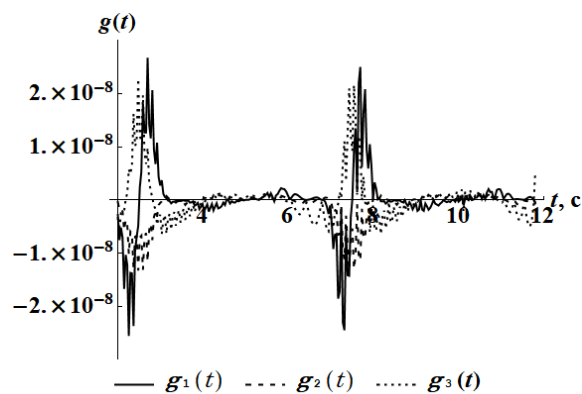


Рис. 4. Точности удовлетворения дополнительных связей

В заключение заметим, что в рассмотренной механической задаче на промежутке моделирования среднее число шагов в итерационном алгоритме (11) составило $k=1.994$, что в два раза меньше ранга матрицы системы уравнений дополнительных связей и в три раза меньше их числа.

Таким образом, результаты моделирования подтверждают описанные в статье свойства итерационного алгоритма решения систем линейных уравнений с положительными полуопределенными матрицами и возможность его применения при моделировании динамики механических систем с дополнительными связями.

Заключение

В статье предложена модификация обратного итерационного метода Пауэлла–Бройдена, основанного на симметричной формуле ранга один (SR1), предназначенная для решения возмущенных СЛАУ с симметричной положительно полуопределенной матрицей системы. Метод используется для решения относительно ускорений уравнений движения механических систем с дополнительными связями при их численном интегрировании. Рассмотрены основные свойства алгоритма.

1. Показано, что все направления поиска решений – векторы s_k , кроме двух последних, являются A -сопряженными. Два последних направления на каждой итерации линейно независимы. Тем самым доказано, что метод глобально сходится за конечное число итераций и точное решение может быть найдено не более чем за $r + 1$ шаг, где r – ранг матрицы системы.

2. Найдены условия, при которых метод локально сходится, т.е. последовательность приближений монотонно стремится к точному решению.

3. Показано, что если матрица возмущений имеет ранг $r_1 < r$, и известно начальное приближение псевдообратной матрицы системы, то алгоритм (11) сходится за $r_1 + 1$ шаг. Таким образом, приближенное решение возмущенной системы n линейных уравнений, может быть найдено с помощью алгоритма (11) много меньше, чем за $n + 1$ шаг.

4. На одном примере проиллюстрировано применение метода при интегрировании уравнений движения механических систем с избыточным числом дополнительных связей. Показано, что количество итераций не превосходит числа линейно независимых связей и остается небольшим на всем промежутке интегрирования.

Список источников

1. *Величенко В.В.* Матрично-геометрические методы в механике с приложениями к задачам робототехники. М.: Наука, 1988. 280 с.
2. *Лиллов Л.К.* Моделирование систем связанных тел. М.: Наука, 1993. 272 с.
3. *Суслов Г.К.* Теоретическая механика. М.: Гостехиздат, 1946. 656с.
4. *Голуб Дж., Ван Лоун Ч.* Матричные вычисления. М.: Мир, 1999. 548с.
5. *Шевцов Г.С.* Численные методы линейной алгебры: учеб. пособие / Г.С. Шевцов, О.Г. Крюкова, Б.И. Мызникова. М.: Финансы и статистика: ИНФРА-М, 2008. 480 с.
6. *Wolfe P.* Convergence conditions for ascent methods // SIAM review, Apr. 1969, Vol. 11, № 2. P. 226–235.
7. *Wolfe P.* Convergence conditions for ascent methods. II: Some corrections // SIAM Review, Apr. 1971, Vol. 13, № 2. P. 185–188.
8. *Conn A.R., Gould N.I.M., Toint Ph.L.* Convergence of quasi-Newton matrices generated by the symmetric rank-one update // Mathematical Programming. 1991. 50, № 1. P. 177–195.
9. *Byrd R.H., Khalfan H.F., Schnabel R.B.* Analysis of a symmetric rank-one trust region method // SIAM J. on Optimization. 1996. 6. P. 1025–1039.
10. *Khalfan H.F., Byrd R.H., Schnabel R.B.* A theoretical and experimental study of the symmetric rank-one update // SIAM J. on Optimization. 1993. 3. P. 1–24.
11. *Аоки М.* Введение в методы оптимизации. М.: Наука, 1977. 344 с.
12. *Дэннис Дж., Шнабель Р.* Численные методы безусловной оптимизации и решения нелинейных уравнений. М.: Мир, 1988. 440 с.
13. *Поляк Б.Т.* Введение в оптимизацию. М.: Наука, 1983. 384 с.
14. *Химмельблау Д.* Прикладное нелинейное программирование. М.: Мир, 1975. 536 с.
15. *Nocedal J., Wright S.J.* Numerical Optimization. Berlin: Springer, 2006. 664 p.
16. *Фиакко А., Мак-Кормик Г.* Нелинейное программирование. Методы последовательной безусловной минимизации. М.: Мир, 1972. 240 с.
17. *Farzin M., Malik Abu H., Wah J.L.* Multi-steps symmetric rank-one update for unconstrained optimization // World Applied Sci. J. 2009. 7, № 5. P. 610–615.
18. *Farzin M., Malik Abu H., Wah J.L.* Memoryless modified symmetric rank-one method for large-scale unconstrained optimization // American J. of Applied Sciences. 2009. 6, № 12. P.2054–2059.
19. *Farzin M., Malik Abu H., Wah J.L.* Convergence of symmetric rank-one method based on modified quasi-Newton equation // J. of Math. Res. 2010. 2, № 3. P. 97–102.
20. *Yueting Y., Chengxian X.* A switching algorithm based on modified quasi-Newton equation // Numer. Math. A J. of Chinese Universities. 2006. 15, N 3. 257–267.
21. *Wah J.L., Malik Abu H.* Convergence of a positive definite symmetric rank-one method with restart // Advanced Modeling and Optimization. 2009. 11, № 4. P. 423–433.
22. *Wah J.L., Malik Abu H.* A restarting approach for the symmetric rank one update for unconstrained optimization // Comput. Optim. Appl. 2009. 42. P. 327–334.
23. *Al-Baali M., Fuduli A., Musmanno R.* On the performance of switching BFGS/SR1 algorithms for unconstrained optimization // Optimization Methods and Software Vol. 19, № 2, April 2004, P. 153–164.
24. *Öztoprak F. & Ş., İlker Birbil S.I.* A Symmetric Rank-One Quasi-Newton Line-Search Method Using Negative Curvature Directions // Optimization Methods and Software. Vol. 26, 2011. P. 455–486.
25. *Farzin M., Malik Abu H., Wah J.L.* A symmetric rank-one method based on extra updating techniques for unconstrained optimization // Computers and Mathematics with Applications. 62 (2011). P. 392–400.

26. Farzin M., Wah J.L. On the performance of a new symmetric rank-one method with restart for solving unconstrained optimization problems // Computers and Mathematics with Applications. 64 (2012). P. 2141–2152.
27. Shu-Zhen Lai, Hou-Biao Li, Zu-Tao Zhang. A Symmetric Rank-One Quasi-Newton Method for Nonnegative Matrix Factorization // ISRN Applied Mathematics. Vol. 2014, Article ID 846483. 11 p.
28. Benson H.Y., Shanno D.F. Cubic regularization in symmetric rank-1 quasi-Newton methods // Math. Prog. Comp. 2018. 10. P. 457–486.
29. Indrapriyadarsini S., Mahboubi S., Ninomiya H., Kamio T., Asai H. Accelerating Symmetric Rank-1 Quasi-Newton Method with Nesterov's Gradient for Training Neural Networks // Algorithms 2022, 15, 6. 1–16p.
30. Иванов В.Н. Основные свойства обратного итерационного алгоритма решения систем линейных уравнений с положительно определенными матрицами // Вычислительные методы и программирование. 2012. Т. 13, № 2. С. 366–376.
31. Гулл Ф., Мюррей У., Райт М. Практическая оптимизация. М.: Мир, 1985. 509 с.
32. Мину М. Математическое программирование. Теория и алгоритмы. М.: Наука, 1990. 488 с.
1. Velichenko V.V. Matrichno-geometricheskie metody v mekhanike s prilozheniyami k zadacham robototekhniki. М.: Nauka, 1988. 280 s. (In Russ.).
2. Lilov L.K. Modelirovanie sistem svyazannyh tel. М.: Nauka; 1993. 272 s. (In Russ.).
3. Suslov G.K. Teoreticheskaya mekhanika. М.: Gostekhizdat; 1946. 656 s. (In Russ.).
4. Golub Dzh., Van Loun Ch. Matrichnye vychisleniya. М.: Mir; 1999. 548 s. (In Russ.).
5. Shevcov G.S. Chislennye metody linejnoy algebrы: ucheb. posobie / G.S. SHEvcov, O.G. Kryukova, B.I. Myznikova. М.: Finansy i statistika: INFRA-M; 2008. 480 s. (In Russ.).
6. Wolfe P. Convergence conditions for ascent methods. SIAM review, Apr. 1969;(11(2)):226–235.
7. Wolfe P. Convergence conditions for ascent methods. II: Some corrections. SIAM Review, Apr. 1971;(13(2)):185–188.
8. Conn A.R., Gould N.I.M., Toint Ph.L. Convergence of quasi-Newton matrices generated by the symmetric rank-one update. Mathematical Programming. 1991;(50(1)):177–195.
9. Byrd R.H., Khalfan H.F., Schnabel R.B. Analysis of a symmetric rank-one trust region method. SIAM J. on Optimization. 1996;(6):1025–1039.
10. Khalfan H.F., Byrd R.H., Schnabel R.B. A theoretical and experimental study of the symmetric rank-one update. SIAM J. on Optimization. 1993;(3):1–24.
11. Aoki M. Vvedenie v metody optimizacii. М.: Nauka; 1977. 344 s. (In Russ.).
12. Dennis Dzh., Shnabel' R. Chislennye metody bezuslovnoj optimizacii i resheniya nelinejnyh uravnenij. М.: Mir; 1988. 440 s. (In Russ.).
13. Polyak B.T. Vvedenie v optimizaciyu. М.: Nauka; 1983. 384 s. (In Russ.).
14. Himmel'blau D. Prikladnoe nelinejnoe programmirovaniye. М.: Mir; 1975. 536 s. (In Russ.).
15. Nocedal J., Wright S.J. Numerical Optimization. Berlin: Springer; 2006. 664 p.
16. Fiakko A., Mak-Kormik G. Nelinejnoe programmirovaniye. Metody posledovatel'noj bezuslovnoj minimizacii. М.: Mir; 1972. 240 s. (In Russ.).
17. Farzin M., Malik Abu H., Wah J.L. Multi-steps symmetric rank-one update for unconstrained optimization. World Applied Sci. J. 2009;(7(5)):610–615.
18. Farzin M., Malik Abu H., Wah J.L. Memoryless modified symmetric rank-one method for large-scale unconstrained optimization. American J. of Applied Sciences. 2009;(6(12)):2054–2059.
19. Farzin M., Malik Abu H., Wah J.L. Convergence of symmetric rank-one method based on modified quasi-Newton equation. J. of Math. Res. 2010;(2(3)):97–102.
20. Yueting Y., Chengxian X. A switching algorithm based on modified quasi-Newton equation. Numer. Math. A J. of Chinese Universities. 2006;(15(3)):257–267.
21. Wah J.L., Malik Abu H. Convergence of a positive definite symmetric rank-one method with restart. Advanced Modeling and Optimization. 2009;(11(4)):423–433.
22. Wah J.L., Malik Abu H. A restarting approach for the symmetric rank one update for unconstrained optimization. Comput. Optim. Appl. 2009;(42):327–334.
23. Al-Baali M., Fuduli A., Musmanno R. On the performance of switching BFGS/SR1 algorithms for unconstrained optimization. Optimization Methods and Software. 2004; (19(2)): 153–164.

24. Öztoprak F. & Ş., İlker Birbil S.I. A Symmetric Rank-One Quasi-Newton Line-Search Method Using Negative Curvature Directions. *Optimization Methods and Software*. 2011;(26):455–486.
25. Farzin M., Malik Abu H., Wah J.L. A symmetric rank-one method based on extra updating techniques for unconstrained optimization. *Computers and Mathematics with Applications*. 2011;(62):392–400.
26. Farzin M., Wah J. L. On the performance of a new symmetric rank-one method with restart for solving unconstrained optimization problems. *Computers and Mathematics with Applications*. 2012;(64):2141–2152.
27. Shu-Zhen Lai, Hou-Biao Li, Zu-Tao Zhang. A Symmetric Rank-One Quasi-Newton Method for Nonnegative Matrix Factorization. *ISRN Applied Mathematics*. 2014;(2014):11. Article ID 846483.
28. Benson H. Y., Shanno D. F. Cubic regularization in symmetric rank-1 quasi-Newton methods. *Math. Prog. Comp.* 2018;(10):457–486.
29. Indrapriyadarsini S., Mahboubi S., Ninomiya H., Kamio T., Asai H. Accelerating Symmetric Rank-1 Quasi-Newton Method with Nesterov's Gradient for Training Neural Networks. *Algorithms*. 2022;(15(6)):1–16.
30. Ivanov V.N. Osnovnye svoystva obratnogo iteracionnogo algoritma resheniya sistem linejnyh uravnenij s polozhitel'no opredelennymi matricami. *Vychislitel'nye metody i programmirovaniye*. 2012;(13(2)):366–376. (In Russ.).
31. Gill F., Myurrej U., Rajt M. *Prakticheskaya optimizaciya*. M.: Mir; 1985. 509 s. (In Russ.).
32. Minu M. *Matematicheskoe programmirovaniye. Teoriya i algoritmy*. M.: Nauka; 1990. 488 s. (In Russ.).

Информация об авторе:

В. Н. Иванов – кандидат физико-математических наук, доцент, доцент кафедры высшей математики механико-математического факультета Пермского государственного национального исследовательского университета (614068, г. Пермь, ул. Букирева, 15), AuthorID 13318.

Information about the author:

V. N. Ivanov – Candidate of Physical and Mathematical Sciences, Associate Professor, Associate Professor of the Department of Higher Mathematics, Faculty of Mechanics and Mathematics, Perm State University (15, Bukireva Street, Perm, Russia, 614068), AuthorID 13318.