



УДК 591.6

С.Е. Саух, д-р техн. наук

Ин-т проблем моделирования в энергетике им. Г.Е. Пухова НАН Украины
(Украина, 03164, Киев-164, ул. Генерала Наумова, 15,
тел. 4249164, e-mail: ssaukh@gmail.com)

Методы обновления столбцово-строчных факторных матриц для ускоренного решения нелинейных систем алгебраических уравнений большой размерности

Предложены методы столбцового, строчного, столбцово-строчного и многогрангового обновления факторных матриц, получаемых в результате *CR*-факторизации. Даны оценки вычислительной сложности методов *CR*-обновления матриц и определены условия эффективного их применения в составе итерационного метода Ньютона решения нелинейных систем алгебраических уравнений. Представлены экспериментальные результаты использования методов *CR*-обновления матриц для решения тестовых систем нелинейных уравнений большой размерности.

Запропоновано методи стовпцевого, рядкового, стовпцево-рядкового і багаторангового оновлення факторних матриць, що утворюються в результаті *CR*-факторізації. Дано оцінки обчислювальної складності методів *CR*-оновлення матриць і визначено умови ефективного їх застосування у складі ітераційного методу Ньютона розв'язку нелінійних систем алгебраїчних рівнянь. Представлено експериментальні результати застосування методів *CR*-оновлення матриць для розв'язування тестових систем нелінійних рівнянь великої розмірності.

Ключевые слова: разреженные матрицы, обновление матриц, столбцово-строчная факторизация.

При использовании вычислительных методов для решения различных математических задач часто возникает необходимость многократного решения систем линейных алгебраических уравнений. Так, в симплекс методе Данцига решения задач линейного программирования на каждом итерационном шаге необходимо решать системы линейных уравнений дважды [1]. Метод Лемке решения задач дополнительности [2], а также методы Ньютона решения нелинейных систем алгебраических уравнений и задач оптимизации требуют решения линейных систем уравнений на каждой итерации [3]. В задачах большой размерности затраты вычислительных ресурсов на многократное решение систем линейных уравнений становятся домини-

© С.Е. Саух, 2013

ирующими, что обусловлено трудоемкостью алгоритмов факторизации матриц. Для сокращения вычислительных затрат каждый раз при обновлении базиса переменных в симплекс методе, а также в методе Лемке используются алгоритмы LU-обновления факторных матриц [4, 5]. Алгоритмы ускоренного блочного или частичного LU-обновления матриц Якоби применяются в итерационных методах Ньютона решения систем нелинейных алгебраических уравнений большой размерности [6].

Существующие алгоритмы обновления факторных матриц основаны на методе LU-факторизации, более трудоемком по сравнению с методом столбцово-строчной (CR) факторизации [7]. Обладая свойством адаптивности к динамически выбираемым ведущим элементам, метод CR-факторизации матриц не требует перестановок строк и столбцов, что позволяет существенно, в среднем более чем на треть, сократить время вычислений. Поэтому разработка методов CR-обновления факторных матриц является актуальной.

Метод CR-факторизации матриц. Основой метода CR-факторизации является последовательность действий, выполняемых над заданной $n \times n$ матрицей A в соответствии с формулами

$$\begin{aligned} A_{(-i_1, -j_1)} &= A - C_{j_1} R_{i_1}, \quad A_{(-i_1, -j_1)}(i_1, :) = 0, \quad A_{(-i_1, -j_1)}(:, j_1) = 0, \\ i_1 &\in \{1, 2, \dots, n\}, \quad j_1 \in \{1, 2, \dots, n\}; \\ A_{(-i_2, -j_2)} &= A_{(-i_1, -j_1)} - C_{j_2} R_{i_2}, \quad A_{(-i_2, -j_2)}(i_2, :) = 0, \quad A_{(-i_2, -j_2)}(:, j_2) = 0, \\ i_2 &\in \{1, 2, \dots, n\}, \quad i_2 \neq i_1, \quad j_2 \in \{1, 2, \dots, n\}, \quad j_2 \neq j_1; \end{aligned} \quad (1)$$

$$\begin{aligned} A_{(-i_n, -j_n)} &= A_{(-i_{n-1}, -j_{n-1})} - C_{j_n} R_{i_n} = 0, \quad i_n \in \{1, 2, \dots, n\}, \quad i_n \notin \{i_k \mid k = 1, 2, \dots, n-1\}, \\ j_n &\in \{1, 2, \dots, n\}, \quad j_n \notin \{j_k \mid k = 1, 2, \dots, n-1\}. \end{aligned}$$

Если множества вектор-столбцов и вектор-строк объединить в матрицы соответственно $C = |C_{j_1} \ C_{j_2} \ \cdots \ C_{j_n}|$ и $R = |R_{i_1} \ R_{i_2} \ \cdots \ R_{i_n}|^T$, то получим разложение $A = CR = \sum_{s=1}^n C_{j_s} R_{i_s}$.

Методы CR-обновления матриц. Изменение значений элементов p -го столбца исходной матрицы. Пусть в соответствии с (1) для исходной матрицы A получены факторные множители C и R :

$$A = CR. \quad (2)$$

Найдем факторные множители матрицы \mathcal{A} , образованной из исходной матрицы A заменой ее p -го столбца $A_{:,p} = |a_{1p} \ a_{2p} \ \dots \ a_{np}|^T$ на столбец $\mathcal{A}_{:,p} = |a_{1p} \ a_{2p} \ \dots \ a_{np}|^T$:

$$\mathcal{A} = A + (\mathcal{A}_{:,p} - A_{:,p}) \mathbf{e}_p^T, \quad (3)$$

где $\mathbf{e}_p^T = |0 \ \dots \ 1 \ 0 \ \dots \ 1|$ — вектор-строка, все элементы которого равны нулю, за исключением p -го, равного единице. Для этого выражение (3) запишем в тождественном виде

$$\mathcal{A} = A + AA^{-1}\mathcal{A}_{:,p} \mathbf{e}_p^T - A \mathbf{e}_p \mathbf{e}_p^T, \quad (4)$$

или

$$\mathcal{A} = A(E + \mathbf{v}_p \mathbf{e}_p^T - \mathbf{e}_p \mathbf{e}_p^T), \quad (5)$$

где вектор $\mathbf{v}_p = A^{-1}\mathcal{A}_{:,p}$, очевидно, является решением уравнения

$$A\mathbf{v}_p = \mathcal{A}_{:,p} \text{ или } CR\mathbf{v}_p = \mathcal{A}_{:,p}. \quad (6)$$

Следует заметить, что используемое в выражениях (4) и (5) произведение векторов $\mathbf{e}_p \mathbf{e}_p^T$ представляет собой $n \times n$ матрицу, все элементы которой равны нулю за исключением p -го диагонального элемента, равного единице. Поэтому, содержащееся в (5) матричное выражение вида $E - \mathbf{e}_p \mathbf{e}_p^T$ в дальнейшем будем представлять матрицей

$$\mathcal{E}_p = \begin{vmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & 0 & \\ & & & & 1 \\ & & & & & \ddots \\ & & & & & & 1 \end{vmatrix},$$

подчеркивая ее отличие от единичной матрицы E только p -м нулевым диагональным элементом. Следовательно,

$$E - \mathbf{e}_p \mathbf{e}_p^T = \mathcal{E}_p. \quad (7)$$

Учитывая соотношения (5)–(7), выражение (3) представим в виде произведения трех матричных множителей:

$$\mathcal{A} = CR(\mathcal{E}_p + \mathbf{v}_p \mathbf{e}_p^T) = CR \mathcal{V}_p. \quad (8)$$

Третий обновляющий множитель вида

$$\mathcal{V}_p = E_p + \mathbf{v}_p \mathbf{e}_p^T = \begin{vmatrix} 1 & v_{1p} & & \\ & 1 & v_{2p} & \\ & & \ddots & \vdots \\ & & & v_{pp} \\ & & & \vdots \\ & & & v_{np} \\ & & & & \ddots \\ & & & & & 1 \end{vmatrix}$$

имеет простую структуру, что позволяет непосредственно использовать разложение (8) матрицы \mathcal{A} в различных вычислительных алгоритмах.

Так, осуществляя поиск вектора неизвестных \mathbf{X} нелинейной системы уравнений

$$\mathbf{F}(\mathbf{X}) = 0 \quad (9)$$

итерационным методом Ньютона, по известным значениям элементов вектора \mathbf{X}^{k-1} на k -й итерации формируем систему линеаризованных уравнений

$$\mathbf{F}(\mathbf{X}^k) \approx \mathbf{F}(\mathbf{X}^{k-1}) + \frac{\partial \mathbf{F}}{\partial \mathbf{X}} \Big|_{\mathbf{X}=\mathbf{X}^{k-1}} (\mathbf{X}^k - \mathbf{X}^{k-1}) = 0 \quad (10)$$

относительно искомого вектора \mathbf{X}^k .

Пусть матрица Якоби $\partial \mathbf{F} / \partial \mathbf{X}$ обладает особенностью такой, что элементы лишь одного ее p -го столбца функционально зависят от неизвестной x_p , т.е. все функции-элементы $|f_1(\mathbf{X}) f_2(\mathbf{X}) \cdots f_n(\mathbf{X})|^T$ вектор-функции $\mathbf{F}(\mathbf{X})$ являются линейными по всем переменным, кроме переменной x_p , относительно которой они являются нелинейными. Например, для системы нелинейных уравнений

$$\begin{aligned} x_1 + 3x_2 + 5x_3 &= 3, \\ 4x_1 + x_2 + x_2^2 + 2x_3 &= 2, \\ 3x_1 + (2+x_2^2)x_2 + x_3 &= 1 \end{aligned}$$

матрица Якоби имеет вид

$$\frac{\partial \mathbf{F}}{\partial \mathbf{X}} = \begin{vmatrix} 1 & 3 & 5 \\ 4 & 1+2x_2 & 2 \\ 3 & 2+3x_2^2 & 1 \end{vmatrix}.$$

Тогда, следуя (6) и (8), множество трудоемких вычислительных процедур факторизации матриц $\partial \mathbf{F} / \partial \mathbf{X} \Big|_{\mathbf{X}=\mathbf{X}^{k-1}}$, входящих в уравнение (10), можно

заменить одноразовой процедурой вычисления на первой итерации факторных множителей C и R матрицы $A = \partial \mathbf{F} / \partial \mathbf{X} \Big|_{X=X^0}$ и использовать их для последовательного решения систем уравнений вида

$$CR(\mathbf{X}^1 - \mathbf{X}^0) = -\mathbf{F}(\mathbf{X}^0) \text{ при } k=1,$$

$$\left\{ \begin{array}{l} CR(E_p + \mathbf{v}_p \mathbf{e}_p^T)(\mathbf{X}^k - \mathbf{X}^{k-1}) = -\mathbf{F}(\mathbf{X}^{k-1}), \\ CR \mathbf{v}_p = \langle \partial \mathbf{F} / \partial \mathbf{X} \Big|_{X=X^{k-1}} \rangle_{:p} \end{array} \right\} \text{при } k > 1, \quad (11)$$

где $\langle \partial \mathbf{F} / \partial \mathbf{X} \Big|_{X=X^{k-1}} \rangle_{:p}$ — p -й столбец матрицы Якоби, элементы которого зависят от значения x_p^{k-1} .

На каждой итерации $k=2, 3, \dots$ алгоритм решения систем уравнений (11) состоит из следующих трех шагов.

Алгоритм 1.

1. Решение системы уравнений $CR \mathbf{v}_p = \langle \partial \mathbf{F} / \partial \mathbf{X} \Big|_{X=X^{k-1}} \rangle_{:p}$ относительно вектора \mathbf{v}_p .
2. Решение системы уравнений $CR \mathbf{Y} = -\mathbf{F}(\mathbf{X}^{k-1})$ относительно вектора \mathbf{Y} .
3. Решение системы уравнений $(E_p + \mathbf{v}_p \mathbf{e}_p^T)(\mathbf{X}^k - \mathbf{X}^{k-1}) = \mathbf{Y}$ относительно искомого вектора \mathbf{X}^k .

Решение каждой из представленных систем уравнений не требует громоздких вычислений, что значительно ускоряет выполнение итерационных процедур метода Ньютона.

Изменение значений элементов множества столбцов исходной матрицы. Если в исходной матрице $A = CR$ одновременно заменить столбцы $A_{:,p_j}$ столбцами $\mathcal{A}_{:,p_j}$, номера которых p_j образуют множество $P = \{p_1, p_2, \dots, p_J\}$, и таким способом сформировать новую матрицу \mathcal{A} , то формула обновления факторных множителей примет вид

$$\mathcal{A} = CR \left(E_{\{p_1, p_2, \dots, p_J\}} + \sum_{j=1}^J \mathbf{v}_{p_j} \mathbf{e}_{p_j}^T \right) = CR \mathcal{V}_P, \quad (12)$$

где матрица $E_{\{p_1, p_2, \dots, p_J\}}$ отличается от единичной матрицы E нулевыми значениями диагональных элементов в позициях p_1, p_2, \dots, p_J ; для $\forall p_j \in P$ векторы обновления \mathbf{v}_{p_j} удовлетворяют системам уравнений

$$CR \mathbf{v}_{p_j} = \mathcal{A}_{:,p_j} \quad (13)$$

и вместе с матрицей $E_{\{p_1, p_2, \dots, p_J\}}$ образуют матрицу обновления

$$\mathcal{V}_P = E_{\{p_1, p_2, \dots, p_J\}} + \sum_{j=1}^J \mathbf{v}_{p_j} \mathbf{e}_{p_j}^T = \begin{vmatrix} 1 & \mathbf{v}_{1p_1} & \mathbf{v}_{1p_J} \\ \ddots & \vdots & \vdots \\ & \mathbf{v}_{p_1p_1} & \mathbf{v}_{p_1p_J} \\ & \vdots & 1 \\ & \mathbf{v}_{p_Jp_1} & \mathbf{v}_{p_Jp_J} \\ \vdots & \vdots & \ddots \\ \mathbf{v}_{n\,p_1} & \mathbf{v}_{n\,p_1} & 1 \end{vmatrix}. \quad (14)$$

Очевидно, использование формулы (12) позволяет расширить возможности алгоритма 1 ускоренного решения линеаризованных систем уравнений вида (10) на случай, когда элементы множества столбцов $P = \{p_1, p_2, \dots, p_J\}$ матрицы Якоби $\partial \mathbf{F} / \partial \mathbf{X}$ функционально зависят от неизвестных $x_{p_1}, x_{p_2}, \dots, x_{p_J}$, т.е. все функции-элементы $|f_1(\mathbf{X}) f_2(\mathbf{X}) \cdots f_n(\mathbf{X})|^T$ вектор-функции $\mathbf{F}(\mathbf{X})$ являются линейными по всем переменным, кроме переменных $x_{p_1}, x_{p_2}, \dots, x_{p_J}$, относительно которых они являются нелинейными. В этом случае вместо систем уравнений (11) запишем

$$\left. \begin{array}{l} CR(\mathbf{X}^1 - \mathbf{X}^0) = -\mathbf{F}(\mathbf{X}^0) \text{ при } k=1, \\ CR \left(E_{\{p_1, p_2, \dots, p_J\}} + \sum_{j=1}^J \mathbf{v}_{p_j} \mathbf{e}_{p_j}^T \right) (\mathbf{X}^k - \mathbf{X}^{k-1}) = -\mathbf{F}(\mathbf{X}^{k-1}), \\ \forall p_j \in P \quad CR \mathbf{v}_{p_j} = \left\langle \partial \mathbf{F} / \partial \mathbf{X} \Big|_{X=X^{k-1}} \right\rangle_{:p_j} \end{array} \right\} \text{ при } k > 1. \quad (15)$$

Отличие алгоритма решения системы уравнений (15) от алгоритма 1 решения системы уравнений (11) заключается в том, что на шаге 1 необходимо решить J систем уравнений $CR \mathbf{v}_{p_j} = \left\langle \partial \mathbf{F} / \partial \mathbf{X} \Big|_{X=X^{k-1}} \right\rangle_{:p_j}$ для $\forall p_j \in P$, а на шаге 3 — более сложную систему уравнений вида

$$\left(E_{\{p_1, p_2, \dots, p_J\}} + \sum_{j=1}^J \mathbf{v}_{p_j} \mathbf{e}_{p_j}^T \right) (\mathbf{X}^k - \mathbf{X}^{k-1}) = \mathbf{Y}. \quad (16)$$

Принимая во внимание структуру матрицы обновления (14), быстрое решение системы уравнений (16) осуществляется посредством решения подсистемы уравнений размерностью J , которая связывает подмножество элементов правых частей $\mathbf{Y}_{\{p_1, p_2, \dots, p_J\}}$ с подмножеством искомых элементов $(\mathbf{X}^k - \mathbf{X}^{k-1})_{\{p_1, p_2, \dots, p_J\}}$, и затем нахождения остальных элементов вектора \mathbf{Y} .

тока ($\mathbf{X}^k - \mathbf{X}^{k-1}$) по простым явным выражениям с использованием в них подстановок уже найденных значений $(\mathbf{X}^k - \mathbf{X}^{k-1})_{\{p_1, p_2, \dots, p_J\}}$.

Изменение значений элементов q -й строки исходной матрицы. Найдем факторные множители матрицы \mathcal{A} , образованной из исходной матрицы $A = CR$ заменой ее q -й строки $A_{q:} = |a_{q1} a_{q2} \cdots a_{qn}|$ строкой $\mathcal{A}_{q:} = |a_{q1} a_{q2} \cdots a_{qn}|$:

$$\mathcal{A} = A + \mathbf{e}_q (\mathcal{A}_{q:} - A_{q:}), \quad (17)$$

где $\mathbf{e}_q = [0 \cdots 1 \ 0 \cdots 0]^T$ — вектор-столбец, все элементы которого равны нулю, за исключением q -го, равного единице. Для этого выражение (17) запишем в тождественном виде:

$$\mathcal{A} = A + \mathbf{e}_q \mathcal{A}_{q:} A^{-1} A - \mathbf{e}_q \mathbf{e}_q^T A, \quad (18)$$

или

$$\mathcal{A} = (E - \mathbf{e}_q \mathbf{e}_q^T + \mathbf{e}_q \mathbf{u}_q) A, \quad (19)$$

где вектор-строка $\mathbf{u}_q = \mathcal{A}_{q:} A^{-1}$, очевидно, является решением уравнения

$$\mathbf{u}_q A = \mathcal{A}_{q:} \text{ или } \mathbf{u}_q C R = \mathcal{A}_{q:}. \quad (20)$$

Заметим, что используемое в выражениях (18), (19) произведение векторов $\mathbf{e}_q \mathbf{e}_q^T$ представляет собой $n \times n$ матрицу, все элементы которой равны нулю за исключением q -го диагонального элемента, равного единице. Поэтому содержащееся в (19) матричное выражение $E - \mathbf{e}_q \mathbf{e}_q^T$ можно представить в виде матрицы $E - \mathbf{e}_q \mathbf{e}_q^T = E_q$.

С учетом соотношений (19), (20) выражение (17) запишем в виде произведения трех матричных множителей:

$$\mathcal{A} = (E_q + \mathbf{e}_q \mathbf{u}_q) C R = \mathcal{U}_q C R. \quad (21)$$

Здесь первый обновляющий множитель

$$\mathcal{U}_q = E_q + \mathbf{e}_q \mathbf{u}_q = \begin{vmatrix} 1 & & & & & \\ & 1 & & & & \\ & & \ddots & & & \\ u_{q1} & u_{q2} & \cdots & u_{qq} & \cdots & u_{qn} \\ & & & & \ddots & \\ & & & & & 1 \end{vmatrix}$$

имеет простую структуру, что позволяет непосредственно использовать разложение (21) матрицы \mathcal{A} в различных вычислительных алгоритмах.

Вновь обращаясь к итерационному методу Ньютона решения нелинейной системы уравнений (9), рассмотрим последовательность решения линеаризованных систем уравнений (10) для случая, когда элементы лишь одной q -й строки матрицы Якоби $\partial\mathbf{F}/\partial\mathbf{X}$ функционально зависят от значений вектора неизвестных \mathbf{X} , т.е. когда все функции-элементы $|f_1(\mathbf{X}) f_2(\mathbf{X}) \cdots f_n(\mathbf{X})|^T$ вектор-функции $\mathbf{F}(\mathbf{X})$ — линейные по всем переменным, кроме функции $f_q(\mathbf{X})$, которая является нелинейной. Например, построенная для системы нелинейных уравнений

$$\begin{aligned} x_1 + 3x_2 + 5x_3 &= 3, \\ x_1 x_2^2 x_3 &= 200, \\ 3x_1 + 10x_2 + x_3 &= 1 \end{aligned}$$

матрица Якоби имеет вид

$$\frac{\partial\mathbf{F}}{\partial\mathbf{X}} = \begin{vmatrix} 1 & 3 & 5 \\ x_2^2 x_3 & 2x_1 x_2 x_3 & x_1 x_2^2 \\ 3 & 10 & 1 \end{vmatrix}.$$

Тогда, следуя (20) и (21), множество трудоемких вычислительных процедур факторизации матриц $\partial\mathbf{F}/\partial\mathbf{X}|_{X=X^{k-1}}$, входящих в уравнение (10), можно заменить одноразовой процедурой вычисления на первой итерации факторных множителей C и R матрицы $A=\partial\mathbf{F}/\partial\mathbf{X}|_{X=X^0}$ и использовать их для последовательного решения систем уравнений

$$\left. \begin{aligned} CR(\mathbf{X}^1 - \mathbf{X}^0) &= -\mathbf{F}(\mathbf{X}^0) \text{ при } k=1, \\ \left\{ \begin{aligned} (\mathcal{E}_q + \mathbf{e}_q \mathbf{u}_q) CR(\mathbf{X}^k - \mathbf{X}^{k-1}) &= -\mathbf{F}(\mathbf{X}^{k-1}), \\ \mathbf{u}_q CR = \langle \partial\mathbf{F}/\partial\mathbf{X} |_{X=X^{k-1}} \rangle_q: \end{aligned} \right. \end{aligned} \right\} \text{при } k>1, \quad (22)$$

где $\langle \partial\mathbf{F}/\partial\mathbf{X} |_{X=X^{k-1}} \rangle_q$: — q -я строка матрицы Якоби, элементы которой в общем случае зависят от значений элементов вектора \mathbf{X}^{k-1} .

На каждой итерации $k=2,3,\dots$ алгоритм решения систем уравнений (22) реализуется последовательным выполнением следующих трех шагов.

Алгоритм 2.

1. Решение системы уравнений $\mathbf{u}_q CR = \langle \partial\mathbf{F}/\partial\mathbf{X} |_{X=X^{k-1}} \rangle_q$: относительно вектор-строки \mathbf{u}_q .
2. Решение системы уравнений $(\mathcal{E}_q + \mathbf{e}_q \mathbf{u}_q) \mathbf{Y} = -\mathbf{F}(\mathbf{X}^{k-1})$ относительно вектора \mathbf{Y} .

3. Решение системы уравнений $CR(\mathbf{X}^k - \mathbf{X}^{k-1}) = \mathbf{Y}$ относительно искового вектора \mathbf{X}^k .

Заметим, что вычислительные сложности алгоритмов 1 и 2 идентичны.

Изменение значений элементов множества строк исходной матрицы. Если в исходной матрице $A = CR$ одновременно заменить строки A_{q_i} : строками \mathcal{A}_{q_i} , номера которых q_i образуют множество $Q = \{q_1, q_2, \dots, q_I\}$, и таким способом сформировать новую матрицу \mathcal{A} , то формула обновления факторных множителей примет вид

$$\mathcal{A} = \left(E_{\{q_1, q_2, \dots, q_I\}} + \sum_{i=1}^I \mathbf{e}_{q_i} \mathbf{u}_{q_i} \right) CR = \mathcal{V}_Q CR, \quad (23)$$

где матрица $E_{\{q_1, q_2, \dots, q_I\}}$ отличается от единичной матрицы E нулевыми значениями диагональных элементов в позициях $Q = \{q_1, q_2, \dots, q_I\}$; для $\forall q_i \in Q$ вектор-строки обновления \mathbf{u}_{q_i} удовлетворяют системам уравнений

$$\mathbf{u}_{q_i} CR = \mathcal{A}_{q_i}; \quad (24)$$

и вместе с матрицей $E_{\{q_1, q_2, \dots, q_I\}}$ образуют матрицу обновления

$$\mathcal{V}_Q = E_{\{q_1, q_2, \dots, q_I\}} + \sum_{i=1}^I \mathbf{e}_{q_i} \mathbf{u}_{q_i} = \begin{vmatrix} 1 & & & & & \\ & \ddots & & & & \\ u_{q_1 1} & \cdots & u_{q_1 q_1} & \cdots & u_{q_1 q_I} & \cdots & u_{q_1 n} \\ & & & & & & \\ u_{q_I 1} & \cdots & u_{q_I q_1} & \cdots & u_{q_I q_I} & \cdots & u_{q_I n} \\ & & & & & \ddots & \\ & & & & & & 1 \end{vmatrix}. \quad (25)$$

Очевидно, использование формулы (23) позволяет расширить возможности алгоритма 2 ускоренного решения линеаризованных систем уравнений (10) на случай, когда элементы множества строк $Q = \{q_1, q_2, \dots, q_I\}$ матрицы Якоби $\partial \mathbf{F} / \partial \mathbf{X}$ функционально зависят от \mathbf{X} , т.е. когда все функции-элементы $|f_1(\mathbf{X}) f_2(\mathbf{X}) \cdots f_n(\mathbf{X})|^T$ вектор-функции $\mathbf{F}(\mathbf{X})$ — линейные по всем переменным, кроме функций $f_{q_1}(\mathbf{X}) f_{q_2}(\mathbf{X}) \cdots f_{q_I}(\mathbf{X})$, которые являются нелинейными. В этом случае вместо системы уравнений (22) получаем

$$\left\{ \begin{array}{l} CR(\mathbf{X}^1 - \mathbf{X}^0) = -\mathbf{F}(\mathbf{X}^0) \text{ при } k=1, \\ \left(\mathbf{E}_{\{q_1, q_2, \dots, q_I\}} + \sum_{i=1}^I \mathbf{e}_{q_i} \mathbf{u}_{q_i} \right) CR(\mathbf{X}^k - \mathbf{X}^{k-1}) = -\mathbf{F}(\mathbf{X}^{k-1}), \\ \forall q_i \in Q \quad \mathbf{u}_{q_i} CR = \left\langle \partial \mathbf{F} / \partial \mathbf{X} \Big|_{\mathbf{X}=\mathbf{X}^{k-1}} \right\rangle_{q_i}; \end{array} \right\} \text{при } k > 1. \quad (26)$$

Отличие алгоритма решения системы (26) от алгоритма 2 решения системы (22) заключается в том, что на шаге 1 необходимо решить I систем уравнений $\mathbf{u}_{q_i} CR = \left\langle \partial \mathbf{F} / \partial \mathbf{X} \Big|_{\mathbf{X}=\mathbf{X}^{k-1}} \right\rangle_{q_i}$: для $\forall q_i \in Q$, а на шаге 2 — более сложную систему уравнений вида

$$\left(\mathbf{E}_{\{q_1, q_2, \dots, q_I\}} + \sum_{i=1}^I \mathbf{e}_{q_i} \mathbf{u}_{q_i} \right) \mathbf{Y} = -\mathbf{F}(\mathbf{X}^{k-1}). \quad (27)$$

Принимая во внимание структуру матрицы обновления (25), быстрое решение системы уравнений (26) осуществляется посредством решения простейшей подсистемы уравнений размерностью $n-I$ с единичной диагональной матрицей, которая связывает подмножество элементов правых частей $\mathbf{F}_{\notin \{q_1, q_2, \dots, q_I\}}(\mathbf{X}^{k-1})$ с подмножеством искомых элементов $\mathbf{Y}_{\notin \{q_1, q_2, \dots, q_I\}}$, и вычисления значений остальных элементов вектора $\mathbf{Y}_{\{q_1, q_2, \dots, q_I\}}$ в результате решения редуцированной системы уравнений размерности $I \times I$.

Одновременное изменение значений элементов множества строк и столбцов исходной матрицы. Если в исходной матрице $A = CR$ одновременно заменить столбцы $A_{:, p_j}$ столбцами $\mathcal{A}_{:, p_j}$, номера которых p_j образуют множество $P = \{p_1, p_2, \dots, p_J\}$, а строки $A_{q_i,:}$ — строками $\mathcal{A}_{q_i,:}$, номера которых образуют множество $Q = \{q_1, q_2, \dots, q_I\}$, и таким способом сформировать новую матрицу \mathcal{A} , то формула обновления факторных множителей примет вид

$$\mathcal{A} = \left(\mathbf{E}_{\{q_1, q_2, \dots, q_I\}} + \sum_{i=1}^I \mathbf{e}_{q_i} \mathbf{u}_{q_i} \right) CR \left(\mathbf{E}_{\{p_1, p_2, \dots, p_J\}} + \sum_{j=1}^J \mathbf{v}_{p_j} \mathbf{e}_{p_j}^T \right) = \mathcal{U}_Q CR \mathcal{V}_P. \quad (28)$$

Использование формулы (28) позволяет расширить возможности алгоритмов 1 и 2 ускоренного решения линеаризованных систем уравнений вида (10) на случай, когда элементы множества строк $Q = \{q_1, q_2, \dots, q_I\}$ и столбцов $P = \{p_1, p_2, \dots, p_J\}$ матрицы Якоби $\partial \mathbf{F} / \partial \mathbf{X}$ функционально зависят от \mathbf{X} . В этом случае получаем

$$CR(\mathbf{X}^1 - \mathbf{X}^0) = -\mathbf{F}(\mathbf{X}^0) \text{ при } k=1,$$

$$\left\{ \begin{array}{l} \mathcal{U}_Q CR \mathcal{V}_P (\mathbf{X}^k - \mathbf{X}^{k-1}) = -\mathbf{F}(\mathbf{X}^{k-1}), \\ \forall q_i \in Q \quad \mathbf{u}_{q_i} CR = \langle \partial \mathbf{F} / \partial \mathbf{X} \Big|_{X=X^{k-1}} \rangle_{q_i}, \\ \forall p_j \in P \quad CR \mathbf{v}_{p_j} = \langle \partial \mathbf{F} / \partial \mathbf{X} \Big|_{X=X^{k-1}} \rangle_{p_j} \end{array} \right\} \text{при } k > 1. \quad (29)$$

На каждой итерации $k=2, 3, \dots$ алгоритм решения системы уравнений (29) реализуется посредством последовательного выполнения следующих пяти шагов.

Алгоритм 3.

1. Решение I систем уравнений $\mathbf{u}_{q_i} CR = \langle \partial \mathbf{F} / \partial \mathbf{X} \Big|_{X=X^{k-1}} \rangle_{q_i}$: относительно вектор-строк \mathbf{u}_{q_i} .
2. Решение J систем уравнений $CR \mathbf{v}_{p_j} = \langle \partial \mathbf{F} / \partial \mathbf{X} \Big|_{X=X^{k-1}} \rangle_{p_j}$: относительно векторов \mathbf{v}_{p_j} .
3. Решение системы уравнений $\mathcal{U}_Q \mathbf{Y} = -\mathbf{F}(\mathbf{X}^{k-1})$ относительно вектора \mathbf{Y} , где матрица $\mathcal{U}_Q = E_{\{q_1, q_2, \dots, q_I\}} + \sum_{i=1}^I \mathbf{e}_{q_i} \mathbf{u}_{q_i}$ имеет вид (25).
4. Решение системы уравнений $CR \mathbf{Z} = \mathbf{Y}$ относительно вектора \mathbf{Z} .
5. Решение системы уравнений $\mathcal{V}_P (\mathbf{X}^k - \mathbf{X}^{k-1}) = \mathbf{Z}$ относительно искомого вектора \mathbf{X}^k , где матрица $\mathcal{V}_P = E_{\{p_1, p_2, \dots, p_J\}} + \sum_{j=1}^J \mathbf{v}_{p_j} \mathbf{e}_{p_j}^T$ имеет вид (14).

Добавление к исходной матрице матрицы более низкого ранга.

Если к исходной матрице $A = CR$ добавляется матрица $\Delta A = \sum_{s=1}^m \mathbf{c}_s \mathbf{r}_s$ ранга $m \ll n$, то такой способ формирования новой матрицы $\mathcal{A} = A + \Delta A$ называется многоранговым обновлением матрицы A . В этом случае матрица ΔA представляется суммой m парных произведений n -мерных вектор-столбцов $\{\mathbf{c}_s \mid s=1, m\}$ и вектор-строк $\{\mathbf{r}_s \mid s=1, m\}$.

Используя факторные множители C и R матрицы A , представим матрицу \mathcal{A} в виде

$$\mathcal{A} = A + \Delta A = CR + \sum_{s=1}^m \mathbf{c}_s \mathbf{r}_s = C \mathcal{W}_m R,$$

где \mathcal{W}_m — матрица обновления,

$$\mathcal{W}_m = \left(E_{\{1, 2, \dots, m\}} + \sum_{s=1}^m \bar{\mathbf{c}}_s \mathbf{e}_s^T \right) \left(E_{\{1, 2, \dots, m\}} + \sum_{s=1}^m \mathbf{e}_s \bar{\mathbf{r}}_s \right)^T + E - E_{\{1, 2, \dots, m\}}.$$

Здесь n -мерные множества вектор-столбцов $\{\bar{\mathbf{c}}_s \mid s = \overline{1, m}\}$ и вектор-строк $\{\bar{\mathbf{r}}_s \mid s = \overline{1, m}\}$ определяются решением систем уравнений вида

$$\{C\bar{\mathbf{c}}_s = \mathbf{c}_s \mid s = \overline{1, m}\}, \quad \{\bar{\mathbf{r}}_s R = \mathbf{r}_s \mid s = \overline{1, m}\}, \quad (30)$$

а матрица $E_{\{1, 2, \dots, m\}}$ отличается от единичной матрицы E нулевыми значениями первых m ее диагональных элементов.

Если учесть особенности слагаемых матрицы W_m , то можно построить достаточно простой способ ее параметрического обращения при решении соответствующих систем линейных алгебраических уравнений. Покажем это на следующем примере.

Пусть обновление матрицы Якоби является многограновым, т.е.

$$\left. \frac{\partial \mathbf{F}}{\partial \mathbf{X}} \right|_{X=X^{k-1}} = \left. \frac{\partial \mathbf{F}}{\partial \mathbf{X}} \right|_{X=X^0} + \sum_{s=1}^m \mathbf{c}_s^{k-1} \mathbf{r}_s^{k-1} = CR + \sum_{s=1}^m \mathbf{c}_s^{k-1} \mathbf{r}_s^{k-1}. \quad (31)$$

Тогда уравнение (10) для $k > 1$ можно представить в виде расширенной системы уравнений

$$\begin{aligned} CR(\mathbf{X}^k - \mathbf{X}^{k-1}) + \sum_{s=1}^m \alpha_s \mathbf{c}_s^{k-1} &= -\mathbf{F}(\mathbf{X}^{k-1}), \\ \{\alpha_s = \mathbf{r}_s^{k-1}(\mathbf{X}^k - \mathbf{X}^{k-1})\}_{s=\overline{1, m}}, \end{aligned} \quad (32)$$

содержащей не только искомый вектор переменных \mathbf{X}^k , но и множество неизвестных параметров $\{\alpha_s\}_{s=\overline{1, m}}$. Для нахождения значений параметров $\{\alpha_s\}_{s=\overline{1, m}}$ разрешим верхнюю часть системы (32) относительно вектора $(\mathbf{X}^k - \mathbf{X}^{k-1})$ и подставим полученное для него выражение в нижнюю часть этой системы. В результате найдем

$$\left\{ \alpha_s = \mathbf{r}_s^{k-1} (CR)^{-1} \left(-\mathbf{F}(\mathbf{X}^{k-1}) - \sum_{t=1}^m \alpha_t \mathbf{c}_t^{k-1} \right) \right\}_{s=\overline{1, m}},$$

или

$$\left\{ \begin{array}{l} \alpha_s + \sum_{t=1}^m \alpha_t \mathbf{w}_s \mathbf{c}_t^{k-1} = -\mathbf{w}_s \mathbf{F}(\mathbf{X}^{k-1}) \\ \mathbf{w}_s CR = \mathbf{r}_s^{k-1} \end{array} \right\}_{s=\overline{1, m}}. \quad (33)$$

Таким образом, на каждой итерации $k = 2, 3, \dots$ алгоритм решения системы уравнений (32) реализуется посредством последовательного выполнения трех шагов.

Алгоритм 4.

1. Решение подсистем уравнений $\{\mathbf{w}_s CR = \mathbf{r}_s^{k-1}\}_{s=1, \dots, m}$ в системе (32) относительно вектора-строк $\{\mathbf{w}_s\}_{s=1, \dots, m}$.
2. Решение подсистемы уравнений $\left\{\alpha_s + \sum_{t=1}^m \alpha_t \mathbf{w}_s \mathbf{c}_t^{k-1} = -\mathbf{w}_s \mathbf{F}(\mathbf{X}^{k-1})\right\}_{s=1, \dots, m}$ в системе (33) относительно множества искомых параметров $\{\alpha_s\}_{s=1, \dots, m}$.
3. Решение подсистемы уравнений $CR(\mathbf{X}^k - \mathbf{X}^{k-1}) = -\mathbf{F}(\mathbf{X}^{k-1}) - \sum_{s=1}^m \alpha_s \mathbf{c}_s^{k-1}$ в (32) относительно искомого вектора \mathbf{X}^k .

Оценка вычислительной сложности алгоритмов CR-обновления факторных матриц. Как правило, при незначительном количестве столбцов $P = \{p_1, p_2, \dots, p_J\}$, строк $Q = \{q_1, q_2, \dots, q_I\}$ и ранге m обновления исходной матрицы A , т.е. при $J \ll n$, $I \ll n$ и $m \ll n$ число ненулевых элементов $nnz(\mathcal{V}_P)$, $nnz(\mathcal{U}_Q)$ и $nnz(\mathcal{W}_m)$ соответствующих матриц обновления \mathcal{V}_P , \mathcal{U}_Q и \mathcal{W}_m оказывается существенно меньше числа ненулевых элементов $nnz(C^A + R^A)$ факторных матриц C^A и R^A обновленной матрицы A .

Для решения вспомогательных систем уравнений (13), (24) и (30) требуется соответственно $(J+1)n^2$, $(I+1)n^2$ и $2(m+1)n^2$ операций умножения. Для решения систем уравнений с матрицами \mathcal{V}_P , \mathcal{U}_Q , \mathcal{W}_m требуется $\frac{1}{3}J^3 - \frac{1}{3}J + J^2$, $\frac{1}{3}I^3 - \frac{1}{3}I + I^2$, $\frac{2}{3}m^3 - \frac{2}{3}m + 2m^2$ операций умножения на факторизацию и решение соответствующих подсистем уравнений, а также $J(n-J)$, $I(n-I)$, $2m(n-m)$ операций умножения на подстановки найденных решений в остальные уравнения. Таким образом, число операций умножения на решение обновленных систем уравнений будет следующее:

$$(J+1)n^2 + \frac{1}{3}J^3 - \frac{1}{3}J + J^2 + J(n-J), \quad (I+1)n^2 + \frac{1}{3}I^3 - \frac{1}{3}I + I^2 + I(n-I), \\ 2(m+1)n^2 + \frac{2}{3}m^3 - \frac{2}{3}m + m^2 + 2m(n-m).$$

Очевидно, при выполнении условий $n \gg J$, $n \gg I$, $n \gg m$ число операций умножения можно оценивать величинами $(J+1)n^2$, $(I+1)n^2$, $2(m+1)n^2$. Отсюда следует, что при достаточно больших размерностях n исходной матрицы A сокращение числа операций при использовании методов обновления факторных матриц составит

$$\sim \frac{n+3}{3(J+1)}, \quad \sim \frac{n+3}{3(I+1)}, \quad \sim \frac{n+3}{6(m+1)}.$$

Полученные оценки свидетельствуют о многократном сокращении числа выполняемых операций умножения при использовании представленных выше методов обновления факторных матриц вместо метода их повторной факторизации.

Результаты экспериментальных исследований. В основу исследований положен программный код, написанный на языке C++, в котором имплементированы итерационная формула (10) решения тестовых систем нелинейных алгебраических уравнений вида

$$A\mathbf{X} + \mathbf{G}(x_p)\mathbf{e}_p^T - \mathbf{B} = 0, \quad (34)$$

формулы (1) CR-факторизации разреженных $n \times n$ матриц A и алгоритм 1 обновления факторных матриц C и R .

Матрицы A взяты с веб-сайта [8]. Для каждой такой матрицы в системе уравнений (34) вектор \mathbf{B} принят равным вектору $A\mathbf{1} + \mathbf{G}(\mathbf{1})$, где $\mathbf{1}$ — единичный вектор. Вектор $\mathbf{G}(x_p)$ сформирован из совокупности нулевых элементов, располагаемых в тех же позициях, что и нулевые элементы столбца $A_{:,p}$ матриц A , и ненулевых элементов, определяемых полиномиальными функциями вида

$$g_i = x_p^2 \left(1 + \frac{i}{2n} + x_p \left(1 + \frac{i}{3n} + x_p \left(1 + \frac{i}{4n} \right) \right) \right).$$

Из возможных решений тестовых систем уравнений (34) взят вектор $\mathbf{X} = \mathbf{1}$. Для его вычисления по итерационной формуле (10) начальные значения элементов x_i^0 вектора \mathbf{X}^0 для всех индексов $i \neq p$ устанавливаем равными единице, а значение x_p^0 — равным нулю. Признаком завершения итераций метода Ньютона принято выполнение условия $|x_p^k - x_p^{k-1}| \leq 10^{-8}$.

По завершении итерационных вычислений найденные значения элементов искомого вектора \mathbf{X} сопоставлялись с точными единичными значениями и оценивалась норма погрешности решений $e = \sqrt{\sum_{i=1}^n (x_i - 1)^2 / n}$.

Вычисления выполнялись на компьютере Intel P4 (Chipset Intel 865 PE, FSB 800 MHz, CPU 3.0GHz with HT, Dual Channel Memory 1 GB: 2x512MB, DDR400), работающем под управлением операционной системы Microsoft Windows XP.

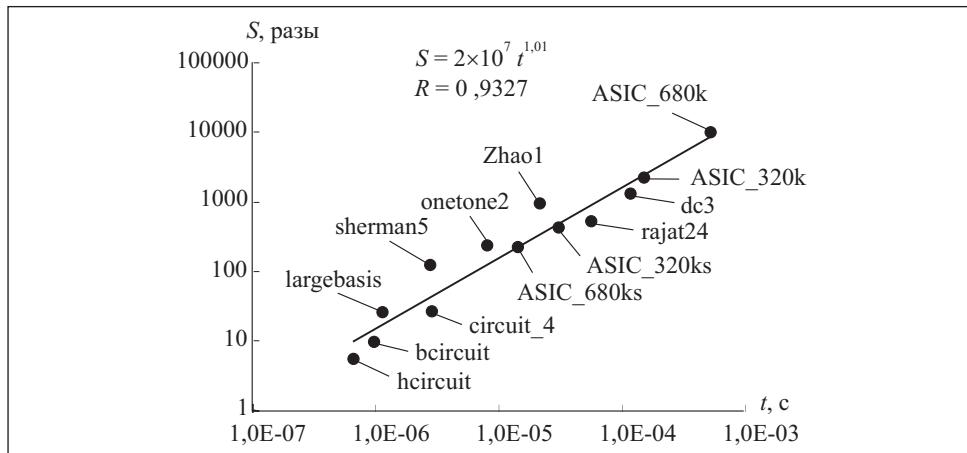
Общие характеристики матриц A и затраты вычислительных ресурсов на их CR-факторизацию представлены в табл. 1, а в табл. 2 — затраты на итерационное решение тестовых систем нелинейных уравнений (34) в

Таблица 1

Название матрицы A	Размерность n	Число ненулевых элементов матрицы		Затраты времени на факторизацию вида $A = CR$, с	Номер столбца p
		A	$C + R$		
ASIC_680ks	682712	1693767	5667160	79,657	50000
ASIC_680k	682862	2638997	6618616	3415,094	50000
ASIC_320ks	321671	1316085	5604966	168,375	5000
ASIC_320k	321821	1931828	5058746	749,672	50000
bcircuit	68902	375558	1049898	1,000	5000
circuit_4	80209	307604	420128	1,172	5000
dc3	116835	766396	1140264	130,516	5000
hcircuit	105676	513072	625005	0,406	50000
largebasis	440020	5240084	22945358	25,750	5000
onetone2	36057	222596	1955857	15,531	500
rajat24	358172	1946979	4105540	227,203	50000
sherman5	3312	20793	178273	0,485	500
Zhao1	33861	166453	12677567	267,235	450

Таблица 2

Название матрицы A	Число итераций метода Ньютона	Затраты времени (с) на итерационные вычисления		Ускорение вычислений S , разы	Среднеквадратическое отклонение полученного решения от точного
		с факторизацией матриц $\partial F / \partial X _{X=X^{k-1}}$	с обновлением матриц C и R		
ASIC_680ks	18	1 360, 513	6, 344	226	4,91e-08
ASIC_680k	9	27 323, 799	3, 047	10 087	6,72e-07
ASIC_320ks	139	23 290, 313	54, 563	429	1,73e-12
ASIC_320k	136	101 252, 283	46, 563	2 190	1,67e-12
bcircuit	24	25, 500	2, 500	10	3,88e-11
circuit_4	35	41, 379	1, 531	27	1,87e-09
dc3	20	2 481, 757	1, 953	1337	3,19e-10
hcircuit	37	17, 304	2, 688	6	1,57e-14
largebasis	31	803, 484	30, 984	26	2,46e-12
onetone2	4	46, 858	0, 265	234	1,81e-09
rajat24	31	6 829, 621	13, 531	521	1,86e-11
sherman5	8	3, 426	0, 031	125	9,14e-14
Zhao1	19	4 815, 511	5, 281	961	4,08e-15



Ускорение вычислений на этапе итерационного решения тестовых систем нелинейных уравнений большой размерности вида (34): t — среднее время вычислений одного элемента матриц C и R

случае CR -факторизации матриц Якоби $\partial\mathbf{F}/\partial\mathbf{X}|_{\mathbf{X}=\mathbf{X}^{k-1}} = \mathbf{A} + \partial\mathbf{G}/\partial\mathbf{X}|_{\mathbf{X}=\mathbf{X}^{k-1}}$ по формулам (1) и в случае обновления факторных матриц C и R по алгоритму 1.

Как видно из табл. 1 и 2, достигнуто значительное сокращение времени вычислений во всех тестовых примерах в случае применения алгоритма обновления. Величины ускорений следует сопоставлять со средними затратами на вычисление одного элемента факторных матриц C и R , поскольку реализация алгоритма CR -факторизации разреженных матриц большой размерности связана с существенными расходами вычислительных ресурсов на выполнение вспомогательных операций [7].

Результаты такого сопоставления представлены на рисунке, где на множестве примеров показана линия тренда ускорений и ее уравнение. Близкий к линейному характер зависимости величины ускорения вычислений от средней величины затрат на вычисление одного элемента факторных матриц свидетельствует о целесообразности применения алгоритмов обновления факторных матриц в случаях, когда затраты на CR -факторизацию матриц Якоби становятся значительными.

Выводы

Предложенные методы столбцового, строчного, столбцово-строчного и многорангового обновления факторных матриц в совокупности с методом CR -факторизации матриц (назовем такую совокупность методов методами CR -обновления матриц) обеспечивают существенное ускорение вы-

числений при решении нелинейных систем алгебраических уравнений итерационным методом Ньютона. Методы *CR*-обновления не оказывают влияния на сходимость итерационных процедур метода Ньютона и обеспечивают высокую эффективность решения систем уравнений большой размерности, содержащих незначительную часть нелинейных функций, определенных на небольшом подмножестве искомых неизвестных.

A column, row, column-row and multi-rank methods of updating the factor matrices derived from *CR*-matrix factorization are proposed. We obtain estimates of the computational complexity of the *CR*-updating methods. The conditions for their effective use in the Newton iterative method for solving nonlinear algebraic equations are determined. The experimental results of using the *CR*-updating methods of test matrices for solution of large scale systems of nonlinear equations are presented.

СПИСОК ЛИТЕРАТУРЫ

1. Dantzig G.B., Thapa M.N. Linear programming. 1: Introduction. — NY. : Springer, 1997. — 436 p.
2. Cottle R.W., Pang J.-S., Stone R.E. The Linear Complementarity Problem. — Philadelphia : SIAM, 1992. — 762 p.
3. [http://www.springerlink.com/content/978-3-540-35447-1#sepcion=424811&page=1&locus=13](http://www.springerlink.com/content/978-3-540-35447-1#seccion=424811&page=1&locus=13)
4. Eldersveld S.K., Saunders M.A. A block-LU update for large-scale linear programming // SIAM J. on Matrix Analysis and Applications. — 1992. — Vol. 13, № 1. — P. 191—201.
5. <http://www.mcs.anl.gov/uploads/cels/papers/P1565.pdf>
6. Tewarson R.P., Zhang Yin Sparse quasi-Newton LU updates // Intern. J. for Numerical Methods in Engineering. — 2005. — Vol. 24, № 6. — P. 1093—1100.
7. Cayuh C.E. Метод *CR*-факторизации матриц большой размерности // Электрон. моделирование. — 2007. — **29**, № 6. — С. 3 — 22.
8. <http://www.cise.ufl.edu/research/sparse/matrices/index.html>

Поступила 20.06.13

САУХ Сергей Евгеньевич, д-р техн. наук, гл. науч. сотр. Ин-та проблем моделирования в энергетике им. Г.Е. Пухова НАН Украины. В 1978 г. окончил Киевский ин-т инженеров гражданской авиации. Область научных исследований — численные операторные методы решения дифференциальных уравнений, декомпозиционные и итерационные методы решения линейных систем большой мерности, математическое моделирование технологических процессов в энергетике и газотранспортных системах, экономико-математические методы моделирования финансовых и макроэкономических процессов.