

ԵՐԵՎԱՆԻ ՊԵՏԱԿԱՆ ՀԱՍՏԱՆՍԱՐԱՆ

Մարտիրոսյան Վահրամ Արթուրի

ԵՐԿԱՆԿՅՈՒՆԱԳԾԱՅԻՆ ՄԱՏՐԻՑՆԵՐԻ
ՊՍԵՎԴԻՈՂԱԿԱԳԱՐՁՄԱՆ ԱԼԳՈՐԻԹՄԻ ԿԱՌՈՒՑՈՒՄԸ
ԵՎ ԴՐԱ ՉՈՒԳԱՀԵՌԱՑՄԱՆ ՀԱՐՅԵՐԻ ՀԵՏԱԶՈՏՈՒՄԸ

Ա.01.07 - "Հաշվողական մաթեմատիկա"
մասնագիտությամբ

Ֆիզիկամաթեմատիկական գիտությունների թեկնածուի
գիտական աստիճանի հայցման ատենախոսության

ՍԵՂՍԱԳԻՐ

Երևան–2009

ЕРЕВАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Мартиросян Ваграм Артурович

ПОСТРОЕНИЕ АЛГОРИТМА ПСЕВДООБРАЩЕНИЯ
ДВУХДИАГОНАЛЬНЫХ МАТРИЦ И ИССЛЕДОВАНИЕ
ВОПРОСОВ ЕГО РАСПАРАЛЛЕЛИВАНИЯ

АВТОРЕФЕРАТ

диссертации на соискание ученой степени кандидата
физико-математических наук по специальности

01.01.07 - "Вычислительная математика"

Ереван – 2009

Ատենախոսության թեման հաստատվել է Երևանի պետական համալսարանում

Գիտական ղեկավար՝ ֆիզ.-մաթ. գիտ. դոկտոր Յու.Ռ. Հակոբյան

Պաշտոնական ընդդիմախոսներ՝ ֆիզ.-մաթ. գիտ. դոկտոր Հ.Ա. Հակոբյան
ֆիզ.-մաթ. գիտ. թեկնածու Վ.Ս. Բոնդարենկո

Առաջատար կազմակերպություն՝ ՀՀ ԳԱԱ ինֆորմատիկայի և ավտոմատացման
սրբաբանների ինստիտուտ

Պաշտպանությունը կայանալու է 2009 թ. հունիսի 17-ին, ժ. 14.00 ԵՊՀ-ում գործող
ԲՈՀ-ի 044 "Մաթեմատիկական կիրառական և մաթեմատիկական տրամաբանություն"
մասնագիտական խորհրդի նիստում հետևյալ հասցեով՝ 0025, ք.Երևան, Ալ.Մանուկյան 1:

Ատենախոսությանը կարելի է ծանոթանալ Երևանի պետական համալսարանի
գրադարանում:

Սեղմագիրը առաքված է 2009 թ. մայիսի 16-ին:

Մասնագիտական խորհրդի գիտական քարտուղար,
ֆիզ.-մաթ. գիտ. թեկնածու

Վ.Ժ. Դումանյան

Тема диссертации утверждена в Ереванском государственном университете

Научный руководитель: доктор физ.-мат. наук Ю.Р. Акопян

Официальные оппоненты: доктор физ.-мат. наук А.А. Акопян
кандидат физ.-мат. наук В.С. Бондаренко

Ведущая организация: Институт проблем информатики и
автоматизации НАН РА

Защита состоится 17 июня 2009 г. в 14.00 часов на заседании действующего
в ЕГУ специализированного совета ВАК 044 "Математическая кибернетика и
математическая логика" по адресу: 0025, г.Ереван, ул. А.Манукяна, 1.

С диссертацией можно ознакомиться в библиотеке Ереванского государственного
университета.

Автореферат разослан 16 мая 2009 г.

Ученый секретарь специализированного совета,
кандидат физ.-мат. наук

В.Ж. Думанян

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы. Многие практические задачи сводятся к решению систем линейных алгебраических уравнений, которые часто оказываются несовместными. Это задачи, возникающие в методе наименьших квадратов, математической статистике, линейном программировании, при исследовании марковских процессов др. В таких случаях система может быть "решена" только путем выбора некоторого компромисса – уравнения системы могут быть удовлетворены не полностью, а лишь до некоторой степени. Наиболее эффективным подходом к такому решению несовместных систем является *псевдообращение* (или обобщенное обращение) матрицы системы. Оказывается, что для произвольной матрицы A существует *псевдообратная* матрица A^+ , которая обладает некоторыми свойствами обратной матрицы. В случае, когда A – квадратная невырожденная матрица, псевдообратная матрица A^+ совпадает с обратной A^{-1} .

Впервые понятие псевдообращения упоминается в работах Э.И. Фредгольма (1903 г.) и А. Гурвица (1912 г.), в которых исследовались обобщенные обращения интегральных операторов. Понятие псевдообратной матрицы было введено Э. Муром в 1920 г. Интерес к псевдообращению матриц резко возрос в 50-х годах прошлого столетия – в связи с активизацией исследований по методу наименьших квадратов. В 1955 г. Р. Пенроуз показал, что для любой $m \times n$ матрицы A обратная по Э. Муру $n \times m$ матрица A^+ однозначно определяется следующими четырьмя условиями:

$$AA^+A = A, \quad A^+AA^+ = A^+, \quad (A^+A)^* = A^+A, \quad (AA^+)^* = AA^+.$$

Этот фундаментальный результат оказался толчком к интенсивному развитию теории *обращения Мура-Пенроуза* или, иначе, *псевдообращения*, которая в настоящее время является одним из актуальных направлений как теории матриц, так и вычислительной математики.

Применение псевдообратных матриц позволяет получать обобщения классических решений для многих задач, наглядно представить структуру полученных результатов, уяснить смысл часто возникающих некорректностей решений и находить пути их регуляризации.

Цель и задачи диссертационной работы. На сегодняшний день наиболее эффективный метод нахождения обратной матрицы Мура-Пенроуза заключается в следующем. Этот метод включает в себя два основных этапа. На первом этапе исходная матрица с помощью отражений Хаусхолдера приводится к двухдиагональному виду. Тем самым задача сводится к псевдообращению двухдиагональной матрицы. На втором этапе строится итерационная процедура, известная как алгоритм Голуба-Райнша¹, которая с помощью матриц плоских вращений генерирует последовательность матриц, сходящихся к диагональной матрице. В итоге получают приближение к сингулярному разложению двухдиагональной матрицы, с помощью которого и находят обращение Мура-Пенроуза этой матрицы. Целью диссертационной работы является построение

¹Golub G.H. and Ch.F. van Loan. Matrix Computations, 3rd ed.- The Johns Hopkins University Press, 1996.

эффективного прямого метода псевдообращения двухдиагональных матриц, альтернативного алгоритму Голуба-Райнша. При этом ставились следующие основные задачи:

- построение оптимального по порядку объема вычислительной работы алгоритма обращения Мура-Пенроуза двухдиагональных матриц;
- получение в явной форме выражений для элементов псевдообратной матрицы через элементы исходной двухдиагональной матрицы;
- исследование вопросов, связанных с распараллеливанием построенного алгоритма.

Методы исследования. Для решения поставленных задач были использованы методы линейной алгебры, теории матриц и численного анализа, а также методы параллельного программирования в среде MPI.

Научная новизна. Построен новый прямой метод вычисления обращения Мура-Пенроуза для двухдиагональных матриц, требующий затраты $O(4n^2)$ флопов, где n – порядок матрицы (для сравнения укажем, что алгоритм Голуба-Райнша требует затраты $O(n^3)$ флопов). Получены формулы, выражающие в явной форме элементы псевдообратной матрицы через элементы исходной двухдиагональной матрицы. Проведен детальный анализ распараллеливания вычислительных процессов и дана подробная схема их реализации в среде MPI.

Практическая значимость. Результаты диссертационной работы носят как теоретический, так и практический характер. Полученные формулы для элементов псевдообратной матрицы могут быть использованы при теоретических исследованиях. Предлагаемый алгоритм псевдообращения двухдиагональных матриц может быть использован при построении программ псевдообращения матриц как на однопроцессорных компьютерах, так и на кластерах.

Основные положения, выносимые на защиту:

- построение и теоретическое обоснование оптимального по порядку объема вычислительной работы алгоритма обращения Мура-Пенроуза двухдиагональных матриц;
- получение в явной форме выражений для элементов псевдообратной матрицы через элементы исходной двухдиагональной матрицы;
- разработка методов распараллеливания построенных вычислительных процессов, их анализ и реализация на кластере.

Апробация полученных результатов. Результаты диссертационной работы докладывались на годовых научных конференциях Российско-Армянского (Славянского) университета (3-7.12.2007 и 5-10.12.2008), на семинарах кафедры системного программирования РАУ, на общем семинаре факультета информатики и прикладной математики ЕГУ.

Публикации. Основные результаты диссертации опубликованы в трех научных статьях.

Структура и объем работы. Диссертация состоит из введения, трех глав, заключения, списка цитируемой литературы (36 наименований) и приложения. Общий объем работы – 127 стр. (без приложения – 87 стр.).

СОДЕРЖАНИЕ РАБОТЫ

Во **Введении** обсуждается актуальность темы исследования, формулируются цели и задачи диссертационной работы. Дается аннотация содержания основных разделов диссертации.

Глава 1 носит обзорный характер. В ней дается определение обращения Мура-Пенроуза (псевдообращения), указываются области применения этого понятия. Далее кратко описываются и характеризуются существующие на сегодняшний день основные методы псевдообращения матриц: методы, основанные на использовании скелетного разложения, QR -разложения, сингулярного разложения, метод Гревилля. Особое внимание уделено методу, основанному на предварительной bidiagonalизации матрицы, так как диссертационная работа посвящена псевдообращению именно двухдиагональных матриц.

Глава 2 диссертации посвящена построению алгоритма обращения Мура-Пенроуза двухдиагональных матриц, а также выводу формул, явно выражающих элементы псевдообратной матрицы через элементы исходной двухдиагональной матрицы.

Широко используемый на практике метод псевдообращения двухдиагональных матриц основан на получении сингулярного разложения этой матрицы путем построения итерационной процедуры Голуба-Райнша². Эта итерационная процедура требует затраты $O(n^3)$ флопов для матриц порядка n . Построенный в диссертации алгоритм требует затраты $O(4n^2)$ флопов. Тем самым этот алгоритм является оптимальным по порядку.

В параграфе 2.1 обсуждается идея построения алгоритма обращения Мура-Пенроуза двухдиагональной матрицы

$$A = \begin{bmatrix} d_1 & b_1 & & & \\ & d_2 & b_2 & & 0 \\ & & \ddots & \ddots & \\ & 0 & & d_{n-1} & b_{n-1} \\ & & & & d_n \end{bmatrix} \quad (1)$$

в предположении, что $d_1 d_2 \dots d_n = 0$ и

$$b_1, b_2, \dots, b_{n-1} \neq 0 \quad (2)$$

(в противном случае, если какие-либо из наддиагональных элементов b_1, b_2, \dots, b_{n-1} равны нулю, задача псевдообращения сводится к аналогичной задаче для двухдиагональных матриц более низкого порядка).

Общая схема построения алгоритма такова. Сначала, в качестве базового алгоритма, строится алгоритм для частного случая, когда

$$d_1, d_2, \dots, d_{n-1} \neq 0, d_n = 0. \quad (3)$$

Затем, на основе этого алгоритма строится алгоритм для общего случая, то есть для любого расположения одного или более нулей на главной диагонали

²Golub G.H. and Ch.F. van Loan. Matrix Computations, 3rd ed.- The Johns Hopkins University Press, 1996.

матрицы A из (1). Вывод алгоритмов основан на использовании известного равенства³

$$A^+ = \lim_{\varepsilon \rightarrow +0} (A^T A + \varepsilon I)^{-1} A^T, \quad (4)$$

где I – единичная матрица, справедливо для любой матрицы $A \in R^{m \times n}$. Матрица $(A^T A + \varepsilon I)$ при любом $\varepsilon > 0$ является невырожденной трехдиагональной матрицей.

В параграфе 2.2 получены выражения для элементов обратной матрицы $(A^T A + \varepsilon I)^{-1}$.

В параграфе 2.3, на основе полученных в предыдущем параграфе формул для элементов обратной матрицы $(A^T A + \varepsilon I)^{-1}$, выводятся выражения для элементов матрицы $(A^T A + \varepsilon I)^{-1} A^T$, зависящие от ε . Путем определения характера этой зависимости, с последующим устремлением ε к нулю, получены, согласно (4), элементы псевдообратной матрицы A^+ . Подчеркнем, что равенство (4) используется здесь как *инструмент получения аналитических выражений* для элементов матрицы A^+ .

Запишем алгоритм вычисления элементов псевдообратной матрицы.

Алгоритм 2d/pinv ($A, n; A^+$)

1. *Вход*: двухдиагональная матрица A вида (1) при условии (2) и (3).
2. Вычисляются величины

$$\overset{\circ}{\mu}_n = 1, \quad \overset{\circ}{\mu}_i = -\frac{b_i}{d_i} \overset{\circ}{\mu}_{i+1}, \quad i = n-1, n-2, \dots, 1. \quad (5)$$

3. Вычисляются величины

$$w_{n-1} = -\frac{1}{b_{n-1}}, \quad w_i = -\frac{d_{i+1} w_{i+1} + \overset{\circ}{\mu}_{i+1}}{b_i}, \quad i = n-2, n-3, \dots, 1. \quad (6)$$

4. Вычисляется величина

$$T_1 = \overset{\circ}{\mu}_1 + d_1 w_1. \quad (7)$$

5. Вычисляются величины

$$v_1 = \frac{1}{T_1}, \quad v_i = -\frac{d_{i-1}}{b_{i-1}} v_{i-1}, \quad i = 2, 3, \dots, n. \quad (8)$$

6. Вычисляются элементы матрицы $A^+ = [\overset{\circ}{y}_{ij}]_{n \times n}$:

$$\left. \begin{aligned} \overset{\circ}{y}_{ij} &= v_i w_j, \quad i = 1, 2, \dots, j \\ \overset{\circ}{y}_{ij} &= \frac{\overset{\circ}{\mu}_i}{\overset{\circ}{\mu}_j} \left(v_j w_j - \frac{1}{d_j} \right), \quad i = j+1, j+2, \dots, n \end{aligned} \right\} j = 1, 2, \dots, n-1, \quad (9)$$

$$\overset{\circ}{y}_{in} = 0, \quad i = 1, 2, \dots, n.$$

7. *Выход*: псевдообратная матрица $A^+ = [\overset{\circ}{y}_{ij}]_{n \times n}$.

конец алгоритма

³Ben-Israel A. and T.N.E. Greville. Generalized Inverses. Theory and Applications.-Springer, New York, 2003.

Непосредственный подсчет показывает, что реализация алгоритма **2d/pinv** требует затраты $O(n^2)$ флопов.

Суммируя вышеизложенное, приходим к следующему утверждению.

Теорема 2.3.1⁴ Для любой двухдиагональной матрицы

$$A = \begin{bmatrix} d_1 & b_1 & & & \\ & d_2 & b_2 & & 0 \\ & & \ddots & \ddots & \\ & 0 & & d_{n-1} & b_{n-1} \\ & & & & 0 \end{bmatrix},$$

где $b_i, d_i \neq 0$, $i = 1, 2, \dots, n-1$, элементы псевдообратной матрицы $A^+ = [\overset{\circ}{y}_{ij}]_{n \times n}$ вычисляются по формулам (5)–(9) с затратой $O(n^2)$ флопов.

В результате проведенных исследований получено также следующее утверждение.

Теорема 2.3.2 Для любой двухдиагональной матрицы

$$A = \begin{bmatrix} d_1 & b_1 & & & \\ & d_2 & b_2 & & 0 \\ & & \ddots & \ddots & \\ & 0 & & d_{n-1} & b_{n-1} \\ & & & & 0 \end{bmatrix},$$

где $b_i, d_i \neq 0$, $i = 1, 2, \dots, n-1$, элементами псевдообратной матрицы $A^+ = [\overset{\circ}{y}_{ij}]_{n \times n}$ являются:

$$\overset{\circ}{y}_{ij} = \frac{(-1)^{j-i} \left(\prod_{s=1}^{i-1} \frac{d_s}{b_s} \right) \sum_{k=1}^{n-j} \left(\prod_{s=j+1}^{n-k} \frac{d_s}{b_s} \right) \left(\prod_{s=n-k+1}^{n-1} \frac{b_s}{d_s} \right)}{b_j \sum_{k=1}^n \left(\prod_{s=1}^{n-k} \frac{d_s}{b_s} \right) \left(\prod_{s=n-k+1}^{n-1} \frac{b_s}{d_s} \right)}, \quad (10)$$

$$i = 1, 2, \dots, j, \quad j = 1, 2, \dots, n-1;$$

$$\overset{\circ}{y}_{ij} = \frac{(-1)^{i-j} \left(\prod_{s=j}^{i-1} \frac{d_s}{b_s} \right)}{b_j} \left[\frac{\left(\prod_{s=1}^{j-1} \frac{d_s}{b_s} \right) \sum_{k=1}^{n-j} \left(\prod_{s=j+1}^{n-k} \frac{d_s}{b_s} \right) \left(\prod_{s=n-k+1}^{n-1} \frac{b_s}{d_s} \right)}{\sum_{k=1}^n \left(\prod_{s=1}^{n-k} \frac{d_s}{b_s} \right) \left(\prod_{s=n-k+1}^{n-1} \frac{b_s}{d_s} \right)} - \frac{b_j}{d_j} \right], \quad (11)$$

$$i = j+1, j+2, \dots, n, \quad j = 1, 2, \dots, n-1;$$

$$\overset{\circ}{y}_{in} = 0, \quad i = 1, 2, \dots, n. \quad (12)$$

⁴Нумерация теорем дается в соответствии с текстом диссертации.

В параграфе 2.4, на основе алгоритма **2d/pinv**, строится алгоритм псевдообращения для общего случая, то есть для любого расположения одного или более нулей на главной диагонали матрицы A из (1).

Матрица A из (1) представляется в блочном виде

$$A = \begin{bmatrix} A_1 & B_1 & & & \\ & A_2 & B_2 & & \\ & & \ddots & \ddots & \\ & & & A_{m-1} & B_{m-1} \\ & & & & A_m \end{bmatrix} \quad (13)$$

с диагональными блоками A_k , $k = 1, 2, \dots, m$, имеющими размер $n_k \times n_k$ и наддиагональными блоками B_k , $k = 1, 2, \dots, m-1$ размера $n_k \times n_{k+1}$. При этом $n_1 + n_2 + \dots + n_m = n$. Разбиение матрицы осуществляется таким образом, чтобы диагональными блоками A_k являлись блоки одного из следующих трех типов:

тип 1 — все диагональные элементы блока, за исключением последнего, отличны от нуля;

тип 2 — нулевой блок размера 1×1 ;

тип 3 — все диагональные элементы блока отличны от нуля.

На рис. 1 схематически представлены диагональные блоки указанных типов (значком \times обозначены ненулевые элементы).

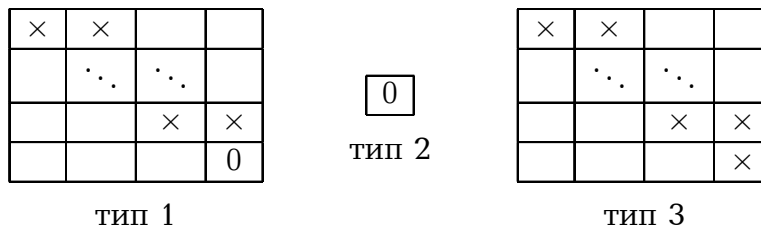


Рис. 1. Типы диагональных блоков.

Заметим, что блоком типа 3 может быть лишь последний диагональный блок A_m в представлении (13). Описанное разбиение матрицы A на блоки является однозначным. Вводятся следующие обозначения для элементов блока A_k , $1 \leq k \leq m$ (при $n_k \geq 2$):

$$A_k = \begin{bmatrix} d_1^{(k)} & b_1^{(k)} & & & \\ & d_2^{(k)} & b_2^{(k)} & & 0 \\ & & \ddots & \ddots & \\ & 0 & & d_{n_k-1}^{(k)} & b_{n_k-1}^{(k)} \\ & & & & d_{n_k}^{(k)} \end{bmatrix}; \quad (14)$$

блоками B_k , $1 \leq k \leq m-1$ являются

$$B_k = \begin{bmatrix} 0 & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 0 \\ \beta_k & 0 & \dots & 0 \end{bmatrix}, \quad \beta_k \equiv b_{n_1+n_2+\dots+n_k}. \quad (15)$$

Применяя равенство (4), показывается, что матрица A^+ имеет блочную структуру

$$A^+ = \begin{bmatrix} Z_1 & & & & & \\ H_2 & Z_2 & & & & \\ & \ddots & \ddots & & & \\ & & & H_{m-1} & Z_{m-1} & \\ & & & & H_m & Z_m \end{bmatrix}, \quad (16)$$

с диагональными блоками Z_k , $k = 1, 2, \dots, m$ размера $n_k \times n_k$ и поддиагональными блоками H_k , $k = 2, 3, \dots, m$ размера $n_k \times n_{k-1}$. Элементы блоков Z_k и H_k вычисляются с помощью следующего алгоритма.

Алгоритм 2d/PINV ($A, n; A^+$)

1. *Вход*: двухдиагональная матрица A вида (1) при условии (2).

2. Осуществляется процедура **2d/decomp** разбиения (13) матрицы A на блоки A_k , $k = 1, 2, \dots, m$ и B_k , $k = 1, 2, \dots, m-1$. Вычисляются величины

$$\beta_k = b_{n_1+n_2+\dots+n_k}, \quad k = 1, 2, \dots, m-1.$$

3. $Z_1 = A_1^+$ (если A_1 – блок типа 1, применяется алгоритм **2d/pinv**, если же A_1 – блок типа 2, то $Z_1 = 0$).

4. Для значений $2 \leq k \leq m$ вычисление блоков Z_k и H_k осуществляется в зависимости от типа блока A_k .

Тип 1

4.1.1 Блок $Z_k = [z_{ij}^{(k)}]_{n_k \times n_k}$.

• Вычисляются величины

$$\mu_{n_k}^{(k)} = 1, \quad \mu_i^{(k)} = -\frac{b_i^{(k)}}{d_i^{(k)}} \mu_{i+1}^{(k)}, \quad i = n_k - 1, n_k - 2, \dots, 1.$$

• Вычисляются элементы блока:

$$z_{ij}^{(k)} = 0, \quad j = i, i+1, \dots, n_k, \quad i = 1, 2, \dots, n_k,$$

$$z_{ij}^{(k)} = -\frac{\mu_i^{(k)}}{\mu_j^{(k)}} \frac{1}{d_j^{(k)}}, \quad j = 1, 2, \dots, i-1, \quad i = 2, 3, \dots, n_k.$$

4.1.2 Блок H_k имеет вид

$$H_k = \begin{bmatrix} 0 & \dots & 0 & h_1^{(k)} \\ 0 & \dots & 0 & h_2^{(k)} \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & h_{n_k}^{(k)} \end{bmatrix},$$

где

$$h_i^{(k)} = \frac{1}{\beta_{k-1}} \frac{\mu_i^{(k)}}{\mu_1^{(k)}}, \quad i = 1, 2, \dots, n_k.$$

Тип 2

4.2.1 Блок $Z_k = 0$.

4.2.2 Блок $H_k = [0 \dots 0 h_{n_{k-1}}^{(k)}]$, где

$$h_{n_{k-1}}^{(k)} = \frac{1}{\beta_{k-1}}.$$

Тип 3

4.3.1 Блок $Z_m = [z_{ij}^{(m)}]_{n_m \times n_m}$ (блоком типа 3 может быть лишь блок A_m).

- Вычисляется обратная матрица $A_m^{-1} = [x_{ij}^{(m)}]_{n_m \times n_m}$:

$$x_{jj}^{(m)} = \frac{1}{d_j^{(m)}}, \quad x_{ij}^{(m)} = -\frac{b_i^{(m)}}{d_i^{(m)}} x_{i+1j}, \quad i = j-1, j-2, \dots, 1.$$

$$j = 1, 2, \dots, n_m.$$

- Вычисляются величины

$$\alpha_2 = \frac{d_1^{(m)}}{b_1^{(m)}}, \quad \alpha_i = \frac{d_{i-1}^{(m)}}{b_{i-1}^{(m)}} \alpha_{i-1}, \quad i = 3, 2, \dots, n_m.$$

- Вычисляются величины

$$\gamma_{n_m} = (x_{1n_m}^{(m)})^2, \quad \gamma_i = \gamma_{i+1} + (x_{1i}^{(m)})^2, \quad i = n_m - 1, n_m - 2, \dots, 1.$$

- Вычисляются величины

$$\overset{\circ}{\lambda}_m = 1 + \beta_{m-1}^2 \gamma_1,$$

$$\nu_1^{(m)} = \frac{1}{\lambda_m}, \quad \nu_i^{(m)} = (-1)^i \frac{\beta_{m-1}^2}{\overset{\circ}{\lambda}_m} \alpha_i \gamma_i, \quad i = 2, 3, \dots, n_m.$$

- Вычисляются элементы матрицы Z_m :

$$\left. \begin{aligned} z_{1j}^{(m)} &= \nu_1^{(m)} x_{1j}^{(m)}, \quad j = 1, 2, \dots, n_m; \\ z_{ij}^{(m)} &= \nu_i^{(m)} x_{1j}^{(m)}, \quad j = 1, 2, \dots, i-1, \\ z_{ij}^{(m)} &= \nu_i^{(m)} x_{1j}^{(m)} + x_{ij}^{(m)}, \quad j = i, i+1, \dots, n_m \end{aligned} \right\} i = 2, 3, \dots, n_m.$$

4.3.2 Блок H_m имеет вид

$$H_m = \begin{bmatrix} 0 & \dots & 0 & h_1^{(m)} \\ 0 & \dots & 0 & h_2^{(m)} \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & h_{n_m}^{(m)} \end{bmatrix},$$

элементы которого вычисляются по формулам

$$h_i^{(m)} = \beta_{m-1} \sum_{j=1}^{n_m} z_{ij}^{(m)} x_{1j}^{(m)}, \quad i = 1, 2, \dots, n_m.$$

5. *Выход*: псевдообратная матрица A^+ блочного вида (16).

конец алгоритма

Подсчет числа операций показывает: если A_k – блок типа 1, то для вычисления блоков Z_k и H_k требуется $O(0.5n_k^2)$ флопов; если A_m – блок типа 3, то для вычисления блоков Z_m и H_m требуется $O(4n_m^2)$ флопов;

Таким образом, приходим к следующему утверждению.

Теорема 2.4.1 Для любой вырожденной двухдиагональной матрицы

$$A = \begin{bmatrix} d_1 & b_1 & & & \\ & d_2 & b_2 & & 0 \\ & & \ddots & \ddots & \\ & 0 & & d_{n-1} & b_{n-1} \\ & & & & d_n \end{bmatrix},$$

где $b_i \neq 0$, $i = 1, 2, \dots, n-1$, элементы псевдообратной матрицы A^+ вычисляются с помощью алгоритма **2d/PINV** с затратой $O(4n^2)$ флопов.

Пользуясь полученными в работе соотношениями, можно выписать явные формулы для вычисления блоков матрицы A^+ из (16). Оформим это в виде следующей процедуры.

Процедура 2d/PINV/EXP ($A, n; A^+$)

1. *Вход*: двухдиагональная матрица A вида (1) при условии (2).

2. Осуществляется процедура **2d/decomp** разбиения (13) матрицы A на блоки A_k , $k = 1, 2, \dots, m$ и B_k , $k = 1, 2, \dots, m-1$. Вычисляются величины

$$\beta_k = b_{n_1+n_2+\dots+n_k}, \quad k = 1, 2, \dots, m-1.$$

3. Блок $Z_1 = A_1^+$.

Если A_1 – блок типа 2, то $Z_1 = 0$. Если A_1 – блок типа 1, то элементы блока $Z_1 = [z_{ij}^{(1)}]_{n_1 \times n_1}$ вычисляются по следующим формулам:

$$z_{ij}^{(1)} = \frac{(-1)^{j-i} \left(\prod_{s=1}^{i-1} \frac{d_s^{(1)}}{b_s^{(1)}} \right) \sum_{k=1}^{n_1-j} \left(\prod_{s=j+1}^{n_1-k} \frac{d_s^{(1)}}{b_s^{(1)}} \right) \left(\prod_{s=n_1-k+1}^{n_1-1} \frac{b_s^{(1)}}{d_s^{(1)}} \right)}{b_j^{(1)} \sum_{k=1}^{n_1} \left(\prod_{s=1}^{n_1-k} \frac{d_s^{(1)}}{b_s^{(1)}} \right) \left(\prod_{s=n_1-k+1}^{n_1-1} \frac{b_s^{(1)}}{d_s^{(1)}} \right)},$$

$$i = 1, 2, \dots, j, \quad j = 1, 2, \dots, n_1 - 1;$$

$$z_{ij}^{(1)} = \frac{(-1)^{i-j} \left(\prod_{s=j}^{i-1} \frac{d_s^{(1)}}{b_s^{(1)}} \right)}{b_j^{(1)}} \left[\frac{\left(\prod_{s=1}^{j-1} \frac{d_s^{(1)}}{b_s^{(1)}} \right) \sum_{k=1}^{n_1-j} \left(\prod_{s=j+1}^{n_1-k} \frac{d_s^{(1)}}{b_s^{(1)}} \right) \left(\prod_{s=n_1-k+1}^{n_1-1} \frac{b_s^{(1)}}{d_s^{(1)}} \right)}{\sum_{k=1}^{n_1} \left(\prod_{s=1}^{n_1-k} \frac{d_s^{(1)}}{b_s^{(1)}} \right) \left(\prod_{s=n_1-k+1}^{n_1-1} \frac{b_s^{(1)}}{d_s^{(1)}} \right)} - \frac{b_j^{(1)}}{d_j^{(1)}} \right],$$

$$i = j + 1, j + 2, \dots, n_1, \quad j = 1, 2, \dots, n_1 - 1;$$

$$z_{in_1}^{(1)} = 0, \quad i = 1, 2, \dots, n_1.$$

4. Для значений $2 \leq k \leq m$ вычисление блоков Z_k и H_k осуществляется в зависимости от типа блока A_k .

Тип 1

4.1.1 Блок $Z_k = [z_{ij}^{(k)}]_{n_k \times n_k}$.

Элементы блока вычисляются по формулам:

$$z_{ij}^{(k)} = 0, \quad j = i, i+1, \dots, n_k, \quad i = 1, 2, \dots, n_k,$$

$$z_{ij}^{(k)} = \frac{(-1)^{i-j+1}}{b_j^{(k)}} \prod_{s=j+1}^{i-1} \frac{d_s^{(k)}}{b_s^{(k)}}, \quad j = 1, 2, \dots, i-1, \quad i = 2, 3, \dots, n_k.$$

4.1.2 Блок H_k имеет вид

$$H_k = \begin{bmatrix} 0 & \dots & 0 & h_1^{(k)} \\ 0 & \dots & 0 & h_2^{(k)} \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & h_{n_k}^{(k)} \end{bmatrix},$$

где

$$h_i^{(k)} = \frac{1}{\beta_{k-1}} \prod_{s=1}^{i-1} \frac{d_s^{(k)}}{b_s^{(k)}}, \quad i = 1, 2, \dots, n_k.$$

Тип 2

4.2.1 Блок $Z_k = 0$.

4.2.2 Блок $H_k = [0 \dots 0 h_{n_{k-1}}^{(k)}]$, где $h_{n_{k-1}}^{(k)} = \frac{1}{\beta_{k-1}}$.

Тип 3

4.3.1 Блок $Z_m = [z_{ij}^{(m)}]_{n_m \times n_m}$ (блоком типа 3 может быть лишь блок A_m).

- Вычисляется обратная матрица $A_m^{-1} = [x_{ij}^{(m)}]_{n_m \times n_m}$:

$$x_{ij}^{(m)} = (-1)^{i+j} \frac{1}{d_j^{(m)}} \prod_{s=i}^{j-1} \frac{b_s^{(m)}}{d_s^{(m)}}, \quad i = 1, 2, \dots, j-1,$$

$$x_{jj}^{(m)} = \frac{1}{d_j^{(m)}},$$

$$j = 1, 2, \dots, n_m.$$

- Вычисляется величина

$$\overset{\circ}{\lambda}_m = 1 + \beta_{m-1}^2 \sum_{j=1}^{n_m} (x_{1j}^{(m)})^2.$$

- Вычисляются величины

$$\nu_1^{(m)} = \frac{1}{\overset{\circ}{\lambda}_m}, \quad \nu_i^{(m)} = (-1)^i \frac{\beta_{m-1}^2}{\overset{\circ}{\lambda}_m} \left(\prod_{s=1}^{i-1} \frac{d_s^{(m)}}{b_s^{(m)}} \right) \sum_{j=i}^{n_m} (x_{1j}^{(m)})^2,$$

$$i = 2, 3, \dots, n_m.$$

- Вычисляются элементы блока:

$$\left. \begin{aligned} z_{1j}^{(m)} &= \nu_1^{(m)} x_{1j}^{(m)}, \quad j = 1, 2, \dots, n_m; \\ z_{ij}^{(m)} &= \nu_i^{(m)} x_{1j}^{(m)}, \quad j = 1, 2, \dots, i-1, \\ z_{ij}^{(m)} &= \nu_i^{(m)} x_{1j}^{(m)} + x_{ij}^{(m)}, \quad j = i, i+1, \dots, n_m \end{aligned} \right\} i = 2, 3, \dots, n_m.$$

4.3.2 Блок H_m имеет вид

$$H_m = \begin{bmatrix} 0 & \dots & 0 & h_1^{(m)} \\ 0 & \dots & 0 & h_2^{(m)} \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & h_{n_m}^{(m)} \end{bmatrix},$$

где

$$h_i^{(m)} = \beta_{m-1} \sum_{j=1}^{n_m} z_{ij}^{(m)} x_{1j}^{(m)}, \quad i = 1, 2, \dots, n_m.$$

5. *Выход*: псевдообратная матрица A^+ блочного вида (16).

конец процедуры

Таким образом, получено следующее утверждение.

Теорема 2.4.2 Для любой вырожденной двухдиагональной матрицы

$$A = \begin{bmatrix} d_1 & b_1 & & & \\ & d_2 & b_2 & & 0 \\ & & \ddots & \ddots & \\ & 0 & & d_{n-1} & b_{n-1} \\ & & & & d_n \end{bmatrix},$$

где $b_i \neq 0$, $i = 1, 2, \dots, n-1$, элементы псевдообратной матрицы A^+ находятся по формулам, объединенным в процедуру **2d/PINV/EXP**.

Отметим следующую важную особенность алгоритма **2d/PINV** (а также и процедуры **2d/PINV/EXP**). Каждая пара блоков Z_k, H_k в блочном представлении (16), для разных значений k , вычисляется независимо друг от друга.

Глава 3 посвящена разработке методов распараллеливания вычислительных процессов, описанных в главе 2. Рассмотрены несколько альтернативных подходов к процессу распараллеливания, отмечены достоинства и недостатки каждого из них, а также даны решения, оптимальные с точки зрения производительности.

В параграфе 3.1 представлены типы распределенных вычислительных систем — суперкомпьютеры и кластеры. Если сравнивать эти системы, то адаптивность к изменениям и сравнительно низкая стоимость делают кластеры наиболее распространенным типом распределенной вычислительной техники. Более восьмидесяти из ста самых высокопроизводительных распределенных систем в мире являются кластерами. В области параллельного программирования на кластерных системах одним из самых распространенных стандартов организации передачи сообщений между

узлами является стандарт MPI (Message Passing Interface). Поэтому вопросы, связанные с распараллеливанием построенных в главе 2 алгоритмов, обсуждаются именно с точки зрения кластерных систем и стандарта MPI.

В параграфе 3.2 дан краткий обзор стандарта MPI, который является одним из первых стандартов в области организации мужузлового сообщения.

В параграфе 3.3 приведены принципы распараллеливания вычислений, обсуждаются детали реализации классов вектора и матрицы. Построен метод распараллеливания алгоритма обращения невырожденных двухдиагональных матриц. Так как кластер не является системой с разделенной памятью, то для достижения наилучшей производительности векторы и матрицы следует реализовывать как непрерывные участки памяти (одномерный динамический массив). Матрицы, имеющие специальную структуру (диагональные, двухдиагональные, трехдиагональные и др.), перед сетевыми операциями следует привести к компактной форме, где хранятся лишь ненулевые элементы этой матрицы.

В параграфе 3.4 рассмотрены вопросы, связанные с распараллеливанием алгоритма псевдообращения $2d/pinv$, построенного в параграфе 2.3.

В параграфе 3.5 обсуждается вопрос распараллеливания алгоритма $2d/PINV$. Рассматриваются два альтернативных варианта — в зависимости от соотношения количества нулей на главной диагонали матрицы и числа узлов кластера. В случае, когда на главной диагонали матрицы есть достаточное (однако не более, чем число узлов кластера) количество нулей, то распараллеливание можно организовать на уровне блоков матрицы (16). В противном случае проводится более детальный анализ каждого из шагов алгоритма $2d/PINV$ для достижения наилучшей производительности.

В параграфе 3.6 проведен сравнительный анализ работы алгоритма по отношению к уже существующим решениям. Приводятся результаты реализации алгоритма на кластере. В главе 2 была оценена вычислительная сложность построенных алгоритмов, и было показано, что они являются оптимальными по порядку. Это качество алгоритмов выгодно проявляется во время их практического сравнения с другими вычислительными системами. Для сравнения была выбрана система Mathematica 7.0, которая на сегодняшний день является одним из самых популярных инструментов для проведения математических расчетов. Были проведены эксперименты с матрицами, порядок которых изменялся диапазоне от 10 до 2000. Ниже приведен скрипт на языке Mathematica который реализует требуемую функциональность:

```
dims = {10,100,200,500,1000,2000};
count = Length[dims];
For[i = 1, i = count, ++i,
  n = dims[[i]];
  A = Table[
    Table[If[i == j || i == j + 1, RandomReal[1000], 0], {j, 1, n}],
    {i, 1, n}];
  Print[n, "x", n, " ", Timing[PseudoInverse[A]]][[1]];
];
```

Этот скрипт, наряду с однопроцессорным вариантом реализации алгоритма **2d/PINV**, были запущены на машине с процессором Pentium IV 2.9Ghz и с оперативной памятью в 256mb. В таблице 1 приведены длительности работы скрипта и алгоритма **2d/PINV** в секундах.

Таблица 1. Сравнение с системой Mathematica 7.0

| размер матрицы | Mathematica 7.0 | 2d/PINV |
|----------------|-------------------------|---------------------------|
| 10 × 10 | 2.498×10^{-16} | 2.59876×10^{-05} |
| 100 × 100 | 0.016 | 0.0110872 |
| 200 × 200 | 0.047 | 0.0430238 |
| 500 × 500 | 0.766 | 0.0356271 |
| 1000 × 1000 | 5.844 | 0.14364 |
| 2000 × 2000 | 44.203 | 0.571514 |

Как видно из таблицы, начиная уже с размеров 100 × 100, алгоритм **2d/PINV** значительно выигрывает во времени по сравнению с Mathematica 7.0.

Результаты работы параллельной версии алгоритма получены на кластерной системе ArmCluster (<http://www.cluster.am>). Использовалось до восьми узлов кластера. В таблице 2 показаны длительности работы алгоритма **2d/PINV** в секундах, соответственно на одном, двух, четырех и восьми узлах, для матриц размеров в диапазоне от 10 × 10 до 4000 × 4000. Рядом, в скобках, показано соотношение длительности сетевых операций и вычислений в процентах.

Таблица 2. Данные о длительности работы алгоритма **2d/PINV**

| размер матрицы | число узлов кластера | | | |
|----------------|----------------------|------------------|------------------|-----------------|
| | 1 | 2 | 4 | 8 |
| 10 × 10 | 0.00050 | 0.02923 (0.1%) | 0.00083 (3.7%) | 0.00390 (82.8%) |
| 100 × 100 | 0.00225 | 0.00283 (47.0%) | 0.00277 (6.9%) | 0.00525 (84.1%) |
| 200 × 200 | 0.00736 | 0.00750 (48.6%) | 0.00828 (22.4%) | 0.01058 (86.7%) |
| 500 × 500 | 0.04365 | 0.02307 (22.3%) | 0.03097 (72.6%) | 0.03239 (87.9%) |
| 1000 × 1000 | 0.17406 | 0.12059 (43.0%) | 0.10531 (68.5%) | 0.08579 (81.0%) |
| 2000 × 2000 | 0.71825 | 0.64598 (58.8%) | 0.55812 (77.6%) | 0.30322 (83.0%) |
| 4000 × 4000 | 3.40232 | 3.05314 (64.2%) | 2.83962 (82.9%) | 1.47887 (86.7%) |

Таким образом, проведенные численные эксперименты подтверждают эффективность построенного алгоритма псевдообращения двухдиагональных матриц.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ РАБОТЫ

В результате проведенных в диссертации исследований получены следующие основные результаты:

1. Построен и теоретически обоснован алгоритм обращения Мура-Пенроуза двухдиагональных матриц, требующий затраты $O(4n^2)$ флопов для матриц порядка n . Тем самым алгоритм является оптимальным по порядку объема вычислительной работы.
2. Получены явные выражения для элементов псевдообратной матрицы через элементы исходной двухдиагональной матрицы.
3. Разработана методика распараллеливания построенных вычислительных процессов, проведен детальный анализ параллельных вычислительных процедур и осуществлена их реализация на кластере.

ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

1. Мартиросян В.А. О псевдообращении двухдиагональных матриц.- Сб. научных статей годичной научной конференции Российско-Армянского (Славянского) университета (03.12-07.12, 2007), физико-математические и естественные науки, изд-во РАУ, Ереван, 2008, 59-64.
2. Мартиросян В.А. Алгоритм псевдообращения двухдиагональных матриц.- Вестник Российско-Армянского (Славянского) университета, серия: физико-математические и естественные науки, Ереван, No.1, 2008, 71-79.
3. Акопян Ю.Р., Мартиросян В.А. О распараллеливании алгоритма псевдообращения матриц.- Вестник Государственного Инженерного Университета Армении, серия: моделирование, оптимизация, управление, Ереван, вып.12, т.1, 2009, 19-27.

ԱՄՓՈՓԱԳԻՐ

Վահրամ Արթուրի Մարտիրոսյան

”Երկանկյունագծային մատրիցների պսևդոհակադարձման
ալգորիթմի կառուցումը և դրա զուգահեռացման
հարցերի հետազոտումը”

Ատենախոսությունում կառուցված է երկանկյունագծային մատրիցների պսևդո-
հակադարձման ըստ հաշվողական աշխատանքի ծավալի կարգի օպտիմալ ալգորիթմ,
և հետազոտված են դրա զուգահեռացման եղանակները: Ստացված են հետևյալ
հիմնական արդյունքները:

1. Կառուցված և տեսականորեն հիմնավորված է երկանկյունագծային մատրից-
ների Մուր-Պենրոուզի հակադարձման ալգորիթմ, որը պահանջում է $O(4n^2)$
թվաբանական գործողություն n -րդ կարգի մատրիցների համար: Այսպիսով,
ալգորիթմը ըստ կարգի օպտիմալ է:
2. Պսևդոհակադարձ մատրիցի տարրերի համար արտածված են բացահայտ
բանաձևեր՝ արտահայտված նախնական երկանկյունագծային մատրիցի
տարրերով:
3. Կառուցված հաշվողական ալգորիթմների համար մշակված է զուգահեռացման
մեթոդ, կատարված է զուգահեռ հաշվողական ընթացքների մանրամասն
վերլուծությունը և իրականացումը կլաստերի վրա:

Vahram A. Martirosyan

”Constructing an algorithm for pseudoinversion of bidiagonal matrices
and research on its parallel implementations”