О.К.Писаренко, Р.Р.Русаков

(ОАО «НИИВК им. М.А. Карцева)

O.K.Pisarenko, R.R.Rusakov

Скоростное решение системы линейных алгебраических уравнений большой размерности методом Холецкого

High-speed large-dimension linear algebraic equations system solutions by Cholesky method

Ключевые слова: СИСТЕМА ЛИНЕЙНЫХ АЛГЕБРАИЧЕСКИХ УРАВНЕНИЙ МЕТОД ХОЛЕЦКОГО

Keywords: SOLUTIONS OF LINEAR ALGEBRAIC EQUATIONS CHOLESKY METHOD

В статье рассматривается аппаратная реализация решения системы линейных алгебраических уравнений размерности 32. Оригинальность разработки заключается в удовлетворении жестких требований технического задания за счет использования специализированного метода решения системы.

The article describes a hardware implementation of solution 32-dimension linear algebraic equations system. Design originality consists in satisfying strict technical conditions by using special solution method.

1. Введение

Существует множество различных методов решения систем линейных алгебраических уравнений (СЛАУ). Выбор оптимального метода решения строится на анализе особенностей конкретной задачи и имеющихся аппаратно-программных средств.

Существует целый ряд прикладных математических программ для решения СЛАУ. Соответствующие функции имеются в таких пакетах, как

MATLAB, Maple, Mathematica и т. п. Однако для успешного применения того или иного метода нужно понимать его особенности и условия применимости.

Все методы решения СЛАУ можно разделить на два класса: прямые (точные) и итерационные (приближенные).

К прямым методам относятся, например, метод Гаусса, прогонки [1]. Все эти методы хороши для систем малой размерности. Если размерность системы больше определенной величины, их применение становится неэффективным - требуется выполнение слишком большого числа арифметических операций.

В числе итерационных методов можно назвать метод простой итерации метод Гаусса-Зейделя и ряд других [1]. При использовании подобных методов решение системы получается как результат последовательных приближений к аналитическому решению с заранее заданной точностью. Число выполняемых при этом арифметических операций может быть гораздо меньше, чем в прямых методах, однако могут возникнуть проблемы со «сходимостью» итерационного. Кроме того, количество итераций может сильно зависеть от выбора начального приближения.

Также существует ряд методов, основанных на том или ином разложении исходной матрицы системы линейных уравнений. Чаще всего в этих методах матрица представляется в виде произведения треугольных и/или диагональных матриц. Затем исходная система уравнений разбивается на более простые части, каждая из которых решается, как правило, обратным ходом метода Гаусса.

2. Разложение Холецкого

Поставленная задача – аппаратное решение СЛАУ размерности 32 с комплекснозначной эрмитовой матрицей не более чем за 25 мкс. Разрядность входных данных: порядок числа – 6, мантиссы реальной и мнимой частей – 24. Точность элементов вектора-решения – не менее 16 разрядов мантиссы.

По результатам проведенного исследования, в ходе которого были опробованы математические модели метода Гаусса, метода сопряженных градиентов и т. п., наиболее эффективным в данных условиях показал себя метод Холецкого (другое название – метод квадратного корня) [1].

В основе метода лежит представление матрицы A (она должна быть симметричной и положительно-определённой) в виде произведения двух матриц: $A = LL^T$, где L — нижняя треугольная матрица со строго положительными элементами на диагонали.

Иногда такое разложение записывается в эквивалентной форме: $A = U^TU$, где $U = L^T$ — верхняя треугольная матрица. Известно, что разложение Холецкого всегда существует и единственно для любой симметричной положительно-определённой матрицы [2].

Существует также обобщение этого разложения на случай комплекснозначных матриц. Если A — положительно-определённая эрмитова матрица, то существует разложение $A = LL^*$, где L — нижняя треугольная матрица с положительными действительными элементами на диагонали, а L^*— эрмитово-сопряжённая к ней матрица.

Элементы матрицы L можно вычислить, начиная с верхнего левого угла матрицы, по формулам:

$$L_{ii} = \sqrt{A_{ii} - \sum_{k=1}^{i-1} L_{ik}^2};$$

$$L_{ij} = \frac{1}{L_{jj}} \left(A_{ij} - \sum_{k=1}^{j-1} L_{ik} L_{jk} \right), \text{ если } j < i.$$

Выражение под корнем всегда положительно, если A — действительная положительно-определённая матрица.

Для комплекснозначных эрмитовых матриц используются формулы:

$$L_{ii} = \sqrt{A_{ii} - \sum_{k=1}^{i-1} L_{ik} L_{ik}^*}. \qquad (1)$$

$$L_{ij} = \frac{1}{L_{jj}} \left(A_{ij} - \sum_{k=1}^{j-1} L_{ik} L_{jk}^* \right), \text{ если } j < i. \qquad (2)$$

Для СЛАУ $Ax = b$, когда матрица A уже представлена в виде $A = LL^T$, решение x получается последовательным решением двух треугольных систем уравнений: $Ly = b$ и $L^Tx = y$.

В случае, когда эрмитова матрица A представлена в виде $A = LL^*$, для СЛАУ $Ax = b$ решение x получается последовательным решением двух треугольных систем уравнений: $Ly = b$ (3) и $L^*x = y$ (4).

3. Аппаратная реализация

Для реализации метода на языке описания аппаратуры была предложена схема, упрощенный вариант которой приведен на рисунке 1.

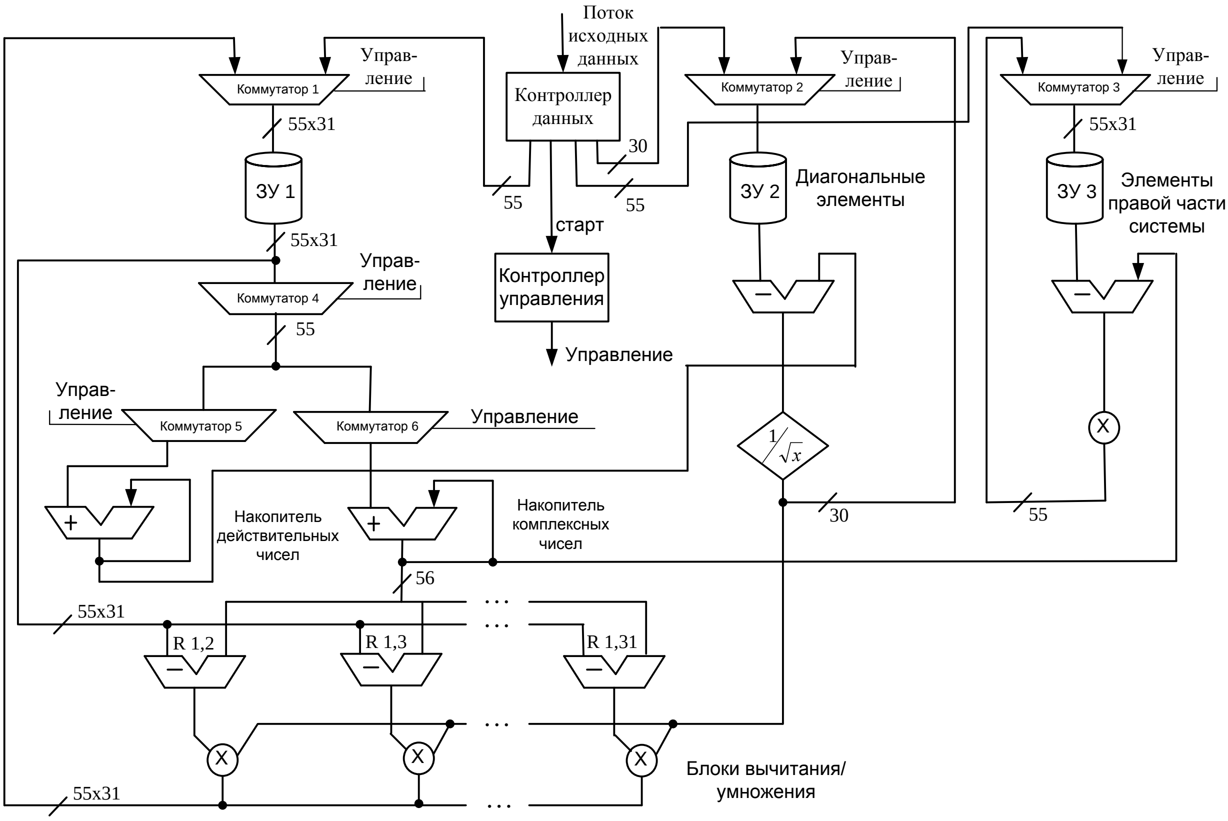

Рисунок 1.

а) загрузка и хранение данных

Исходные данные системы уравнений от различных независимых источников данных подаются на контроллер данных, который распределяет их на три независимых памяти: память элементов матрицы «ЗУ 1» (кроме элементов, стоящих на главной диагонали), память элементов главной диагонали «ЗУ 2» и память правых частей системы уравнений «ЗУ 3». Таким образом, исходные данные разбиваются на блоки по типу обработки.

Поскольку, по условию технического задания, исходная матрица системы является эрмитовой, то нет необходимости хранить все её элементы. Поэтому «ЗУ 1» состоит из FIFO различной емкости – от 1 до 31 слова - элементы, расположенные ниже главной диагонали можно исключить.

б) решение СЛАУ

Наиболее сложно реализуемая операция - «квадратный корень обратной величины», обозначенная на схеме как $\frac{1}{\sqrt{x}}$, реализуется методом линейно-ломанной аппроксимации [3].

Алгоритм решения системы разбивается на три независимых этапа:

Этап 1 - преобразование исходной системы уравнений по формулам (1) и (2). Используется следующий функционал: накопитель действительных чисел и накопитель комплексных чисел, блок вычитания/умножения, вычитатель 1 и блок операции $\frac{1}{\sqrt{x}}$. По завершению операции преобразованные элементы матрицы замещают исходные.

Этап 2 – корректировка правой части СЛАУ. Используется оборудование: накопитель комплексных чисел, вычитатель 2 и умножитель. Аналогичным образом полученные данные замещают исходные.

Этап 3 – решение СЛАУ. Для решения полученных систем (3) и (4) используется обратный ход метода Гаусса, реализованный на накопителе комплексных чисел. Выход адресных счетчиков ЗУ 1 и ЗУ 2 инвертируется для возможности чтения данных в обратном порядке. Полученный вектор-решение располагается в ЗУ 3.

Контроллер управления устройством синхронизирует работу всей схемы, вырабатывая импульсы и стробы разрешения работы для отдельных блоков, получая от них импульсы завершения операций.

4. Результаты

В результате предварительного моделирования отдельных узлов схемы было получены следующие результаты решения:

а) временные затраты (f_{clk} = 200 МГц):

- Загрузка данных: 2.6 мкс;
- Этап 1 – преобразование матрицы: 6.4 мкс;
- Этап 2 – преобразование правых частей: 3.6 мкс;
- Этап 3 – решение СЛАУ: 3.8 мкс.

  Общее время – 16.4 мкс.

б) точность полученного вектора-решения (исходная разрядность мантисс – 24 бита) - 19 бит.

5. Литература

1. Амосов А.А., Дубинский Ю.А., Копченова Н.П. Вычислительные методы для инженеров. – М.: Мир, 1998.

2. Horn Roger A.; Johnson Charles R., Matrix Analysis, Cambridge University Press, 1985, ISBN 0-521-38632-2.

3. Петров Е.Ю., Андреев А.С. Потоковая обработка степенной функции с показателем степени -0.5. Вопросы радиоэлектроники, серия ЭВТ (в наст. выпуске).