

ББК 22.19
УДК 624.012.03+519.612

Фиалко С.Ю.

Прямые методы решения систем линейных уравнений в современных МКЭ комплексах. –М.:

Издательство СКАД СОФТ, 2009. – 161 с.
ISBN 978–5–903683–06–2

Книга посвящена описанию прямых методов решения систем линейных алгебраических уравнений с симметричными разреженными матрицами, возникающих при применении метода конечных элементов к задачам строительной механики и механики деформированного твердого тела, ориентированных на доступные широкому кругу пользователей многоядерные настольные компьютеры. Особое внимание уделяется достижению высокой производительности на каждом процессоре при использовании симметричной схемы хранения матриц, особенностям распараллеливания на основе многопоточности и реализации блочного многофронтального метода подконструкций, позволяющего создать эффективный и экономичный с точки зрения использования оперативной и дисковой памяти решатель, используемый в течение ряда лет в программном комплексе SCAD.

Книга предназначена для специалистов в области численных методов и метода конечных элементов, для расчетчиков и пользователей программных конечноэлементных комплексов, желающих глубже понять принципы работы современных решателей, для разработчиков вычислительных программных комплексов, для научных работников и исследователей, а также может служить учебным пособием для студентов технических специальностей, связанных с прочностным расчетом инженерных сооружений и для студентов-информатиков, изучающих численное моделирование в технических задачах.

ББК 22.19
УДК 624.012.03+519.612

ISBN 978–5–903683–06–2

©СКАД СОФТ, 2009

ОГЛАВЛЕНИЕ

Предисловие	6
1. Введение в прямые методы решения систем конечноэлементных уравнений	
1.1. Введение	11
1.2. Ленточные и профильные методы	12
1.3. Фронтальный метод	19
1.4. Введение в упорядочение	27
Литература к разделу 1	30
2. Упорядочение	
2.1. Представление структуры разреженной матрицы в виде графа	31
2.2. Цели и задачи упорядочения	35
2.3. Метод вложенных сечений	36
2.4. Алгоритм минимальной степени	41
2.5. Сравнение метода вложенных сечений и алгоритма минимальной степени. Гибридные методы упорядочения	45
Литература к разделу 2	52
3. Высокопроизводительные вычисления для многопроцессорных систем с разделяемой памятью	
3.1. Алгоритмы высокопроизводительные и «наивные»	53
3.2. Производительность алгоритмов линейной алгебры	56
3.3. Разбиение на блоки на уровне регистров	64
3.4. Конвейерные технологии процессоров. Разворачивание циклов	68
3.5. Упаковка данных	73
3.6. Дополнительные возможности 64-разрядной архитектуры	80
3.7. Особенности распараллеливания вычислительных алгоритмов для многоядерных (многопроцессорных) компьютеров с разделяемой оперативной памятью на основе многопоточности	81
Литература к разделу 3	93
4. Прямые методы для разреженных матриц	
4.1. Методы, основанные на компактном формате данных	94
4.2. Классический многофронтальный метод	101
4.3. PARDISO солвер	106
4.4. Метод подконструкций (суперэлементов)	114
4.5. Метод вложенных подконструкций	119
4.6. Блочный многофронтальный метод подконструкций	123
4.6.1. Упорядочение	125
4.6.2. Анализ	126

4.6.3. Численная факторизация	132
4.6.4. Факторизация фронтальной матрицы	136
4.6.5. Параллельный алгоритм блочной факторизации для систем с разделяемой памятью	141
4.6.6. Прямая и обратная подстановки	144
4.6.7. Числовые примеры	145
4.7. Заключение	153
Литература к разделу 4	155
Краткий словарь терминов	157

Предисловие

Назначение этой книги – ввести читателя в круг проблем, связанных с современными методами решения систем линейных алгебраических уравнений с симметричными разреженными матрицами, которые возникают при применении метода конечных элементов (МКЭ) к задачам строительной механики и механики деформированного твердого тела.

Существующие на сегодняшний день методы решения таких систем уравнений можно разделить на две больших группы. К первой группе относятся прямые методы, а ко второй – итерационные.

Данная книга посвящается описанию прямых методов. Основное внимание уделяется особенностям, связанным с разреженной структурой матрицы коэффициентов, с достижением высокой производительности арифметических операций и с организацией параллельных вычислений для многоядерных (multi-core) и многопроцессорных компьютеров с общей (shared-memory) оперативной памятью, прочно захвативших лидерство в вычислительной практике малых и средних конструкторских бюро.

По мнению автора в силу целого ряда объективных и субъективных причин оказалось так, что на территории Советского Союза, а затем на постсоветском пространстве данный вопрос освещался слабо. В то же время в западном мире интенсивно велись работы по созданию мощных конечноэлементных решателей, что позволило выйти на значительно более высокий уровень точности и достоверности расчетных моделей. Настало время познакомить отечественных специалистов с основными принципами, лежащими в основе этих достижений.

С 1995 года автор занимался разработкой прямых и итерационных методов решения данного класса задач высокой размерности в фирме RoboBAT (Firma Informatyczna RoboBAT – www.robobat.com), а с 2002 года по сегодняшний день – в фирме SCAD Soft (www.scadsoft.com), разрабатывающих конечноэлементные программные комплексы для расчета строительных конструкций на статику, устойчивость и колебания. Поэтому представленный в книге материал прошел серьезное тестирование многочисленными пользователями разных стран.

Книга предназначена для студентов технических специальностей, связанных с прочностным расчетом инженерных сооружений, для студентов-информатиков, изучающих курсы высокопроизводительных вычислений и моделирования в технических задачах, для специалистов в области численных методов и метода конечных элементов, для расчетчиков и пользователей программных конечноэлементных комплексов, желающих глубже понять принципы работы современных решателей – модулей решения систем линейных алгебраических уравнений, для разработчиков

программных комплексов, а также для научных работников и исследователей, интересующихся подобными вопросами.

Успех при создании современного решателя лежит на пересечении таких областей знаний, как линейная алгебра, матричные вычисления, теория графов, информатика, а в ряде случаев и строительная механика. Поэтому автор старался привести сведения из всех упомянутых областей знаний ровно в том объеме, в каком это необходимо для понимания материала. Автор стремился также сделать материал доступным для широкого круга читателей, поэтому во многих местах вместо строгих доказательств приведены частные примеры, а для желающих познакомиться со строгими доказательствами имеются соответствующие ссылки на литературу.

С другой стороны автор опасался погрузиться в популизм – стиль изложения, при котором сложность вопроса остается "за кадром", а у читателя складывается обманчивое впечатление о том, что все, о чем здесь написано, очень просто и предельно ясно.

Книга состоит из четырех разделов. Первый раздел посвящен введению в прямые методы решения систем линейных алгебраических уравнений с симметричными разреженными матрицами. Рассмотрены основы профильного и фронтального методов.

Во втором разделе приведены сведения об алгоритмах упорядочения – методе вложенных сечений и алгоритме минимальной степени. Рассмотрен также гибридный метод, объединяющий преимущества указанных базовых подходов. Основные идеи иллюстрируются на простых задачах – плоских рамах. На примере типичной расчетной модели показано, насколько различные методы упорядочения уменьшают объем факторизованной разреженной матрицы.

В третьем разделе представлены современные компьютерные методы умножения плотных матриц. Показано, что учет иерархичности структуры памяти современного компьютера, состоящей из регистров процессора, нескольких уровней кэш и оперативной памяти, и тонкое понимание особенностей работы каждого из этих устройств позволяет многократно повысить производительность матричного умножения по сравнению с так называемыми наивными алгоритмами, не учитывающими этих особенностей. Отметим, что каждый из приведенных алгоритмов реализует одно и то же известное алгебраическое правило умножения двух матриц. Однако в зависимости от умения работать с разными уровнями памяти производительность этих алгоритмов для больших матриц отличается в десятки раз.

Ключевыми моментами увеличения производительности решения систем линейных алгебраических уравнений с разреженными матрицами являются сведение всех операций в разреженной матрице большой размерности к последовательности операций в относительно небольших плотных матрицах, применение блочных методов их факторизации и достижение достаточно большого размера блока. В свою очередь перечисленные факторы позволяют

применить высокопроизводительное матричное умножение при вычислении дополнения Шура, трудоемкость которого многократно превосходит все остальные процедуры, выполняющиеся при факторизации плотной матрицы. Именно поэтому такое внимание уделено алгоритмам высокопроизводительного матричного умножения. Рассмотрены подходы, позволяющие достичь высокой производительности вычислений: разбиение матриц на блоки на уровне оперативная память – кэш – оперативная память (блокирование кэш $L1$), на уровне кэш – регистры процессора – кэш (блокирование регистров), векторизации вычислений при применении векторных регистров ХММ, выполняющих одновременно для чисел двойной точности (тип `double`) за один цикл процессора два умножения или два сложения, специальная упаковка матрицы, позволяющая многократное использование памяти кэш (данные, один раз попавшие в кэш, должны оставаться там как можно дольше) и минимизирующая количество кэш-промахов, а также разворачивание циклов, способствующая максимальному использованию конвейерных технологий процессора. Основное внимание при этом уделено подходам, используемым в программе `dgemm` (general matrix multiplication) библиотеки Intel Math Kernel Library (Intel MKL).

Главной причиной, побудившей автора заняться разработкой алгоритма многопоточного высокопроизводительного матричного умножения, явилось отсутствие в библиотеке Intel MKL аналогичной программы, использующей симметричную схему хранения матриц. Программы `dgemm`, `dsyrk` (Intel MKL) демонстрируют прекрасные результаты, однако требуют для хранения симметричной матрицы \mathbf{C} при вычислении дополнения Шура $\mathbf{C} = \mathbf{C} - \mathbf{W} \cdot \mathbf{W}^T$ массив из N^2 элементов, где размерности матриц \mathbf{C} , \mathbf{W} составляют соответственно $N \times N$ и $N \times M$. Однако для хранения симметричной матрицы \mathbf{C} достаточно $N(N+1)/2$ элементов, что дает возможность почти двукратной экономии оперативной памяти.

В конце данного раздела приводятся некоторые сведения из области параллельных вычислительных систем, а также особенности ускорения алгоритмов линейной алгебры при увеличении количества процессоров на многоядерных настольных компьютерах, слабым местом которых является низкая пропускная способность системной шины. Показано, что на компьютерах такой архитектуры хорошему ускорению подлежат алгоритмы типа высокопроизводительного матричного умножения, имеющие оценку производительности не хуже чем $q = f / m \sim \sqrt{M}$, где f , m , M – соответственно количество арифметических операций алгоритма, количество пересылок данных оперативная память – кэш – оперативная память и размер кэш $L1$, выраженный в словах типа `double`.

В разделе 4 представлены прямые методы для разреженных матриц. Описаны алгоритмы `left looking` и `right looking`, использующие компактный формат для хранения только ненулевых элементов матрицы. На простом

примере представлены основы классического многофронтального метода, а также метода PARDISO, включенного в библиотеку Intel MKL. Приведены результаты нескольких тестов PARDISO.

Представлен метод подконструкций (суперэлементов), использующий автоматическое разбиение исходной конструкции на подконструкции методом параллельных сечений, а также метод вложенных подконструкций – метод суперэлементов иерархического вложения, в котором разбиение на подконструкции выполняется методом вложенных сечений.

Главное внимание уделено разработанному автором блочному многофронтальному методу подконструкций, который с 2002 года применяется в программном комплексе SCAD (www.scadsoft.com).

От других известных реализаций метода суперэлементов его отличает полная автоматизация разделения исходной конструкции на подконструкции, возможность использования произвольного алгоритма упорядочения и тот факт, что каждой подконструкции соответствует плотная матрица, которая по аналогии с классическим многофронтальным методом названа фронтальной.

От классического многофронтального метода, являющегося чисто алгебраическим, блочный многофронтальный метод подконструкций отличается четко выраженной суперэлементной трактовкой, которая на этапе анализа позволяет оперировать не столбцами разреженной матрицы, а узлами расчетной модели, что приводит к существенному уменьшению объема информации графа смежности и естественному объединению уравнений в группы, ассоциированные с узлами расчетной модели. Первое значительно ускоряет процесс поиска оптимального для данной задачи упорядочения, позволяя в считанные секунды "проиграть" несколько различных алгоритмов и выбрать из них наиболее оптимальный, а второе облегчает объединение уравнений в блоки для блочной факторизации.

Разработка высокопроизводительного алгоритма матричного умножения с симметричной схемой хранения позволила создать экономный с точки зрения требований к оперативной памяти конечноэлементный решатель для настольных многоядерных компьютеров.

В конце раздела приведены тесты, демонстрирующие возможности блочного многофронтального метода подконструкций.

Автор выражает глубокую благодарность академику НАН Украины, проф. Ярославу Михайловичу Григоренко за ценные советы и замечания, проф. Виктору Андреевичу Баженову и всему коллективу кафедры строительной механики и научно-исследовательского института строительной механики КНУБА за обсуждение на семинарах и конференциях различных частей материала, представленного в данной книге, проф. Александру Сергеевичу Сахарову, замечания и высказывания которого способствовали формированию у автора трактовки данного метода как метода подконструкций.

Автор глубоко признателен проф. Анатолию Викторовичу Перельмутеру, многие советы которого оказали неоценимую пользу при написании этой книги, к. т. н. Виктору Семеновичу Карпиловскому, к. т. н. Эдуарду Зиновьевичу Криксунову за поддержку при внедрении решателя в вычислительный программный комплекс SCAD, а также всему коллективу SCAD Soft, проявившему незаурядное терпение и выдержку при устранении ошибок, обнаруженных в процессе эксплуатации.

Март 2009 г.

С. Ю. Фиалко