УДК 502.51 (282.257.5):504.5

Кузнецов И.В., Дальневосточный государственный
университет путей сообщений, г. Хабаровск

Рукавишников А.В., Институт прикладной
математики ДВО РАН, г. Хабаровск

ОБОБЩЕННЫЙ МЕТОД МИНИМАЛЬНЫХ НЕВЯЗОК

В работе рассмотрен обобщенный метод минимальных невязок и его принципы реализации с использованием параллельных вычислений.

Ключевые слова: итерационные методы решения СЛАУ, подпространство Крылова.

Kuznetcov I.V., Far Eastern State
Transport University, Khabarovsk

Rukavishnikov A.V., Institute of Applied
Mathematics FEB RAS, Khabarovsk

GENERALIZED MINIMUM RESIDUAL METHOD

In the paper the generalized minimum residual method and its implementation principles using parallel computations are considered.

Key words: iterative methods for linear systems, Krylov subspace.

Введение. Рассмотрим задачу, связанную с решением СЛАУ больших размерностей. Они возникают, при приближенном решении дифференциальных краевых задач, например, течений вязкой несжимаемой жидкости, обтекания крыла самолёта, климата и многие другие. При решении таких задач, как правило, ставится вопрос о нахождении приближенного решения, а не точного ввиду невозможности его отыскания.

Для нахождения приближенного решения (вектора неизвестных СЛАУ) используются итерационные методы решения. Современные методы основаны на разложении вектора невязки в подпространстве Крылова и обладают высокими характеристиками решения – точность, минимальность количества необходимых итераций.

$$K_m = span\{v, Av, A^2v, ..., A^{m-1}v\}$$

Одним из таких методов является обобщенный метод минимальных невязок (GMRES). Впервые данный метод получил описание в [1] в виде геометрической реализации. Алгебраическая реализация предложена в [2].

Метод минимизации невязок на подпространствах Крылова за конечное число итераций приводит к точному решению и по этой причине является прямым методом. Но чаще его рассматривают как итерационный метод – с типичными для итерационных методов оценками сходимости ввиду того, что спустя некоторое небольшое количество итераций n, полученное приближение уже является хорошим для точного решения.

Структура алгоритма. Рассмотрим алгоритм метода и остановимся подробно на его отдельных элементах. Алгоритм, представленный на рисунке, можно условно поделить на три части: вычисление нормы невязки, алгоритм Арнольди и нахождение решения.

1. *Compute* $r_0 = b - Ax_0$, $\beta := \|r_0\|_2$, *and* $v_1 := r_0/\beta$
2. *For* $j = 1, 2, \ldots, m$ *Do:*
3.   *Compute* $w_j := Av_j$
4.   *For* $i = 1, \ldots, j$ *Do:*
5. $h_{ij} := (w_j, v_i)$
6. $w_j := w_j - h_{ij}v_i$
7.   *EndDo*
8. $h_{j+1,j} = \|w_j\|_2$. *If* $h_{j+1,j} = 0$ *set* $m := j$ *and go to 11*
9. $v_{j+1} = w_j/h_{j+1,j}$
10. *EndDo*
11. *Define the* $(m+1) \times m$ *Hessenberg matrix* $\bar{H}_m = \{h_{ij}\}_{1 \le i \le m+1, 1 \le j \le m}$.
12. *Compute* y_m *the minimizer of* $\|\beta e_1 - \bar{H}_m y\|_2$ *and* $x_m = x_0 + V_m y_m$.

Рисунок. Алгоритм GMRES

Вычисление нормы невязки представлено в первой его части не требует подробного описания. Вторая часть алгоритма представляет собой вычисление матрицы Хессенберга, и ортонормированного базиса подпространства Крылова. Матрица Хессенберга представляет собой неквадратную матрицу размера (m+1) m

$$H_m = \begin{bmatrix} h_{11} & \ldots & \ldots & h_{m,1} \\ h_{21} & \ldots & \ldots & \ldots \\ 0 & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & h_{mm} \\ 0 & \ldots & 0 & h_{m+1,m} \end{bmatrix}.$$

Для использования её в алгоритме, над данной матрицей проводят m операций вращения Гивенса. Вращение представляет собой операцию поворота матрицы на некоторый заданный угол. На плоскости такой поворот определяется матрицей линейного оператора:

$$G(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}.$$

Для вращения матрицы Хессенберга строится матрица вида:

$$\Omega_i = \begin{pmatrix} 1 & & & & & \\ & \dots & & & & \\ & & c_i & s_i & & \\ & & -s_i & c_i & & \\ & & & & \dots & \\ & & & & & 1 \end{pmatrix},$$

где

$$c_i = \frac{h_{ii}}{\sqrt{h_{ii}^2 + h_{i,i+1}^2}}, s_i = \frac{h_{i+1,i}}{\sqrt{h_{ii}^2 + h_{i,i+1}^2}}$$

Проведем m умножений матрицы Хессенберга на матрицу поворота, тогда матрица Хессенберга к верхнетреугольному виду (последняя строка нулевая).

Третья часть алгоритма одной итерации заключается в минимизации нормы невязки, которая задана уравнением:

$$r_m = H_m y_m - \beta e_1.$$

Минимизация евклидовой нормы вектора невязки представляет собой решение линейной задачи минимальных квадратов размерности m.

Для этого необходимо решить уравнение

$$H_m y_m = \beta e_1,$$

после чего можно найти значение решения

$$x_m = x_0 + V y_m.$$

Данный алгоритм описывается с помощью операций скалярного умножения векторов, разности и произведений векторов и матриц, а также умножения матрицы или вектора на скаляр. Самая трудоёмкая и самая часто вызываемая операция в алгоритме – умножение матрицы на вектор. В первую очередь это связано с вычислением матрицы Хессенберга и ортонормированного базиса подпространства Крылова. Число умножений матрицы на вектор: n–1 в алгоритме Арнольди, K в выбранном алгоритме решения задачи минимальных квадратов и 1 в подсчёте результата. Также в алгоритм входят $\frac{n(n-1)}{2}$ операций скалярного произведения.

Основным недостатком метода GMRES является увеличение объема требуемой памяти с каждой итерацией. В связи с этим, для некоторых вычислительных систем выполнение программы метода может быть невозможным. Для некоторых систем используют модифицированный обобщенный метод минимальных невязок – GMRES(m), который подразумевает использование перезапуска алгоритма.

В данном случае, в качестве начального приближения передаётся уже полученное значение приближенного решения. Реализация такого алгоритма будет иметь рекурсивный характер. Перезапуск алгоритма позволяет существенно снизить требования к занимаемой памяти виду того, что в такой реализации требуется постоянно хранить только данные требуемые для перезапуска.

Ресурс распараллеливания. Рассмотрим вопрос проведения параллельных вычислений в алгоритме. GMRES состоит из операций, которые целесообразно запускать в режиме параллельных вычислений. Это обусловлено тем, что результат выполнения элементарных операций алгоритма не зависит и не влияет на входные данные.

Скалярное произведение векторов, умножение вектора или матрицы на скаляр и разность матриц, применяемые в алгоритме не требуют дополнительных операций для возможности запуска их в параллельном режиме.

Умножение матрицы на вектор также можно представить в параллельном виде, если описать каждое значение матрицы-результата как скалярное произведение строки из матрицы и вектора. Подобным образом можно распараллелить и умножение двух матриц. В обоих случаях для этого требуется разбиение матриц на совокупности векторов, представляющих собой строки или столбцы матрицы.

Реализация параллельных вычислений в этих операциях влияет на каждый шаг алгоритма, что обуславливает высоких потенциал распараллеливания.

Формат хранения данных алгоритма. При реализации алгоритма важно уделить внимание формату входных данных. GMRES может находить решение как для плотных, так и для разряженных матриц, что ставит вопрос о том, в каком виде более целесообразно подавать данные в алгоритм ввиду того, что его реализация может существенно отличаться в зависимости от этого. Также стоит отметить, что каждый формат представления СЛАУ рационален лишь для некоторых видом матриц. К примеру, применение форматов хранения CSS и CSR для плотных матриц не рационально из-за того, что по объему занимаемой памяти они будут близки к объему хранения матрицы в виде двумерного вектора. С другой стороны, применение двумерного вектора для хранения разряженных матриц также не выгодно ввиду больших затрат памяти при малом числе уникальных ненулевых элементов.

В качестве решения данной проблемы предлагается реализация конкретного вида алгоритма GMRES и отдельно написание конвертера различных форматов хранения матриц к необходимому. Данный подход может быть в достаточной мере затратным по объему занимаемой памяти, однако он гарантирует универсальность алгоритма, его применяемость для самого широкого спектра матриц. Метод работает даже в случае, когда матрица системы линейных алгебраических уравнений не обладает свойством положительной определенности.

Отметим, что подавляющее число приближенных метод решения задач, решаемых на практике, имеют дело с разреженными матрицами системы. Поэтому при реализации алгоритма имеет смысл пользоваться форматом хранения, предназначенным для них.

Алгоритм с предобуславливанием. Для ускорения выполнения алгоритма необходимо применять предобуславливание матрицы системы. Предобуславливание позволяет сократить число итераций за счет уменьшения числа обусловленности матрицы системы. Такое действие сопряжено с появлением дополнительной погрешности, что компенсируется повышением скорости алгоритма. На практике для построения вышеизложенного метода применяются алгоритмы: ILU(0), ILU(p) с левым, правым и двухсторонним предобуславливанием матрицы системы (см., например, [3–6]) и др.

Заключение. Алгоритм GMRES представляет собой итерационный метод с высокой скоростью сходимости (по числу итераций и времени выполнения), по-сравнению с прямыми методами и большинством итерационных алгоритмов. Эффективность метода можно дополнительно повысить за счёт проведения параллельных вычислений и использования предобуславливания матрицы системы. Данный алгоритм универсален, достаточно точен, а при применении перезапуска алгоритма не слишком требователен к объемам памяти. Предложенный метод позволяет решать различные задачи с приемлемой скоростью, что даёт ему широкий спектр практического применения для задач большой размерности системы линейных алгебраических уравнений, матрица системы которой не обладает симметричности и положительной определенности.

Список литературы

1. Марчук, Г.И. Итерационные методы и квадратичные функционалы / Г.И. Марчук, Ю.А. Кузнецов // Методы вычислительной математики : сборник – Новосибирск : Наука, 1975. – 143 с

2. Saad, Y. Iterative Methods for Sparse Linear Systems, 2nd edition, Society for Industrial and Applied Mathematics, 2003.

3. Рукавишников, А.В. Численный метод решения задачи Стокса с разрывным коэффициентом / А.В. Рукавишников // Вычислительные методы и программирование: новые вычислительные технологии. – Т. 6. – № 1. – 2005. – С. 17–26.

4. Рукавишников, А.В. Неконформный метод конечных элементов для одной задачи гидродинамики с криволинейным интерфейсом / А.В. Рукавишников // Журнал вычислительной математики и математической физики. – Т. 52. – № 6. – 2012. – С. 1072–1092.

5. Рукавишников, А.В. Метод декомпозиции области и численный анализ для одной задачи гидродинамики / А.В. Рукавишников // Журнал вычислительной математики и математической физики. – Т. 54. – № 9. – 2014. – С. 1515–1536.

6. Rukavishnikov, V.A., Rukavishnikov A.V. New approximate method for solving the Stokes problem in a domain with corner singularity // V.A. Rukavishnikov, A.V. Rukavishnikov // Вестник Южно-Уральского государственного университета. Серия: Математическое моделирование и программирование. – Т. 11. – № 1. – 2018. – С. 95–108.