Записки научных
семинаров ПОМИ
Том 334, 2006 г.

М. Дана, Х. Д. Икрамов

ЕЩЕ РАЗ О РЕШЕНИИ СИСТЕМ ЛИНЕЙНЫХ УРАВНЕНИЙ, МАТРИЦЫ КОТОРЫХ ЯВЛЯЮТСЯ МАЛОРАНГОВЫМИ ВОЗМУЩЕНИЯМИ ЭРМИТОВЫХ МАТРИЦ

1. Введение

Пусть нужно решить систему линейных уравнений

$$Ax = b \tag{1}$$

с нормальной $n \times n$-матрицей A. Нормальность в данном контексте означает выполнение условия

$$AA^* = A^*A. \tag{2}$$

Для разреженной неэрмитовой матрицы A стандартным выбором алгоритма для решения системы (1) является обобщенный метод минимальных невязок GMRES. Однако в этом методе не удается использовать специфику матрицы A, выражаемую равенством (2): алгоритм Арнольди, лежащий в основе GMRES, приводит нормальную матрицу A к той же форме Хессенберга, что и матрицу общего вида. В результате сохраняются два главных недостатка, присущих GMRES по сравнению с методами типа Ланцоша: большой расход памяти и зависимость количества арифметической работы, выполняемой на отдельном шаге, от его номера k.

При дополнительном предположении, что спектр матрицы A принадлежит алгебраической кривой Γ невысокого порядка (примером реалистической ситуации, где это предположение выполнено, могут служить нормальные матрицы с почти-вещественным спектром), в [1] предложен иной подход к решению системы (1), который вкратце состоит в следующем. Методы типа Ланцоша и GMRES ищут приближенные решения в крыловских подпространствах возрастающей размерности, т.е. в линейных оболочках конечных отрезков степенной последовательности

$$x, Bx, B^2x, \ldots . \tag{3}$$

Здесь B – это или сама матрица A, или некоторая матрица, тесно с A связанная; в качестве x обычно (при отсутствии хорошего начального приближения) берут правую часть b системы (1) или полученной из нее "предобусловленной" системы.

Метод, предложенный в [1], является алгоритмом минимальных невязок для *обобщенных крыловских подпространств*, натянутых на конечные отрезки последовательности

$$x, Ax, A^*x, A^2x, AA^*x, A^{*2}x, A^3x, A^2A^*x, \dots . \tag{4}$$

В отличие от GMRES, этот метод, называемый MINRES-N, обладает фиксированной длиной рекурсии, зависящей от порядка m кривой Γ. Так, для случая $m = 2$, наиболее подробно рассмотренного в [1], длина рекурсии равна 6. То же самое можно сказать иначе: вместо унитарного приведения матрицы A к форме Хессенберга, неявно выполняемого в GMRES, A теперь приводится к специальной компактной форме, которую можно считать блочно-трехдиагональной или ленточной матрицей с шириной ленты, зависящей от m. Тем самым устраняются оба отмеченных выше недостатка GMRES.

MINRES-N2 представляет собой специализацию метода MINRES-N для нормальных матриц со спектрами, лежащими на кривых второго порядка. В [2] было показано, что MINRES-N2 примени́м и к системам с анормальными матрицами, которые могут быть описаны следующим образом. Как известно, всякая квадратная комплексная матрица A может быть единственным образом представлена в виде

$$A = B + iC, \quad B = B^*, \quad C = C^*. \tag{5}$$

Эрмитовы матрицы

$$B = \frac{1}{2}(A + A^*), \qquad C = \frac{1}{2i}(A - A^*) \tag{6}$$

называются соответственно вещественной и мнимой частями матрицы A. Нормальные матрицы A характеризуются тем свойством, что соответствующие им матрицы (6) перестановочны. Анормальные матрицы, поддающиеся обработке методом MINRES-N2, могут теперь быть описаны условием

$$\operatorname{rank} C = 1.$$

Целью данной статьи является обобщение этого факта на случай

$$\operatorname{rank} C = k \geqslant 1. \tag{7}$$

В разделе 2 мы показываем, что матрица A, удовлетворяющая условию (7), может быть приведена посредством унитарного подобия к блочно-трехдиагональному виду, где порядок каждого диагонального блока не превосходит $1+k$. Это означает, что систему (1) с такой матрицей A можно решать методом MINRES-N(1+k), т.е. вариантом метода MINRES-N, первоначально предназначенным для нормальных систем со спектрами, лежащими на кривых порядка $1+k$. В разделе 3 мы обсуждаем примеры, в которых производительность MINRES-N(1+k) для систем, подчиняющихся условию (7), сравнивалась с производительностью GMRES.

2. Малоранговые возмущения эрмитовых матриц

Рассмотрим матричную последовательность $I_n, A, A^*, \dots,$ соответствующую векторной последовательности (4). Учитывая, что A – анормальная матрица и, следовательно, не коммутирует с A^*, эта матричная последовательность в более подробной записи имеет вид

$$I_n, A, A^*, A^2, AA^*, A^*A, A^{*2}, A^3, A^2A^*, AA^*A, \dots, A^{*3}, \dots. \tag{8}$$

Члены последовательности (8), называемые словами из A и A^*, получаются из одночленов от двух некоммутирующих переменных

$$1, x, y, x^2, xy, yx, y^2, x^3, x^2y, xyx, \dots, y^3, \dots$$

подстановкой

$$x \to A, \quad y \to A^*.$$

Если в слове $w(A, A^*)$ символ A встречается в качестве множителя i раз, а символ A^* – j раз, то число $i+j$ называется степенью этого слова. Совокупность 2^m слов одной и той же степени m образует m-й слой последовательности (8); при этом считается, что единичная матрица I_n составляет нулевой слой.

Обозначим через $\mathcal{L}_m$ линейную оболочку матриц, входящих в слои последовательности (8) с номерами $0, 1, 2, \dots, m$. Положим

$$\ell_m = \dim \mathcal{L}_m \tag{9}$$

и

$$\omega_m = \ell_m - \ell_{m-1} \qquad (m \geq 1). \tag{10}$$

Число ω_m назовем *шириной* m-го слоя. По определению, $\omega_0 = 1$. Оказывается, что размеры блоков в блочно-трехдиагональной форме, к которой A можно привести посредством унитарного подобия, определяются как раз числами ω_m (см. [1, раздел 3]).

Наша задача – оценить числа ω_m для матриц A, удовлетворяющих условию (7). С этой целью заметим, что указанные числа (равно как и числа ℓ_m) можно найти, используя вместо (8) последовательность

$$I, B, C, B^2, BC, CB, C^2, B^3, B^2C, BCB, CB^2, BC^2, \ldots, C^3, \ldots \ . \tag{11}$$

Она эквивалентна последовательности (8) в том смысле, что, будучи разделена на слои, генерирует ту же последовательность подпространств $\mathcal{L}_m$. Это утверждение есть несложное следствие разложения (5).

Теорема 1. *Если $n \times n$-матрица A удовлетворяет условию* (7), *то*

$$\omega_m \leq 1 + k, \qquad m \geqslant 1. \tag{12}$$

Доказательство. Согласно определению (10), ω_m имеет смысл приращения размерности подпространства $\mathcal{L}_m$ по сравнению с $\mathcal{L}_{m-1}$. Обозначим через r_m (соответственно s_m) составляющую этого приращения, полученную за счет матриц из m-го слоя последовательности (11), в выражениях которых как слов из B и C первым слева множителем является B (соответственно C). Очевидно,

$$\omega_m = r_m + s_m. \tag{13}$$

Пусть j – максимальное среди значений индекса m, таких, что $s_m \neq 0$. Тогда

$$\sum_{m=1}^{j} s_m \leqslant k. \tag{14}$$

Слой m последовательности (11) может быть получен умножением матриц из слоя $m-1$ слева сначала на матрицу B, а затем на C. Отсюда вытекает, что

$$r_m \leq \omega_{m-1} = r_{m-1} + s_{m-1}. \tag{15}$$

Используя соотношения (13)–(15), находим

$$\omega_j = r_j + s_j \le r_{j-1} + s_{j-1} + s_j \le \cdots \le r_1 + s_1 + \cdots + s_m$$
$$\le r_1 + k \le 1 + k.$$

Здесь учтено неравенство

$$r_1 \le \omega_0 = 1.$$

Итак,

$$\omega_j \le 1 + k.$$

Из тех же выкладок следует, что

$$\omega_m \le 1 + k, \quad \text{если} \quad m < j.$$

Пусть теперь $m > j$. Поскольку $s_m = 0$ при $m > j$, имеем

$$r_m = \omega_m \le \omega_{m-1} \le \cdots \le \omega_j \le 1 + k.$$

Тем самым $\omega_m \le 1 + k$ для всех m, и теорема доказана.

От матричной последовательности (8) вернемся к векторной последовательности

$$x, Ax, A^*x, A^2x, AA^*x, A^*Ax, A^{*2}x, A^3x, \dots . \tag{16}$$

Если к системе (1) не применяется предобусловливание и не имеется хорошего начального приближения, то в качестве x берут правую часть b. Базис, в котором A приобретает искомую блочно-трехдиагональную форму, получается посредством ортонормализации последовательности (16). Она проводится аналогично процессу ортонормализации в алгоритмах Ланцоша и Арнольди, но использует на первых шагах матрично-векторные произведения как с матрицей A, так и с A^*. После того как достигнут слой j (см. доказательство теоремы), можно использовать произведения лишь одного типа (только с матрицей A либо только с A^*). Блочно-трехдиагональную форму H матрицы A можно рассматривать и как ленточную матрицу. Из особенностей процесса ортогонализации следует, что нижняя ширина ленты в H равна $1 + k$, а верхняя ширина ленты есть $1 + 2k$. Таким образом, H – ленточная матрица с общей шириной ленты $3(1 + k)$.

3. Численные эксперименты

В этом разделе мы описываем результаты численных экспериментов, в которых метод MINRES-N(1+k) сравнивался с GMRES на системах линейных уравнений с матрицами, удовлетворяющими условию (7). Все системы имели порядок 2000, и их матрицы $A = B + iC$ конструировались следующим образом: B – это вещественная симметричная трехдиагональная матрица, диагональные элементы которой суть псевдослучайные числа, равномерно распределенные на паре сегментов [-350, -275] и [275, 350], а элементы первой наддиагонали (и первой поддиагонали) равномерно распределены на паре сегментов [-70, -40] и [50, 80]. Матрица C также была вещественной симметричной и выбиралась в виде

$$C = u_1 u_1^T + u_2 u_2^T + \ldots + u_k u_k^T,$$

где элементы каждого вектора u_i $(1 \leq i \leq k)$ равномерно распределены в [-1, 1]. Правые части систем строились одним из трех способов: 1) как псевдослучайные векторы порядка 2000 с компонентами, равномерно распределенными в [0,1]; 2) как произведение Ae, где e – вектор порядка 2000, составленный из единиц; в этом случае решение системы (1) известно и равно e; 3) как произведение $A^* e_i$, где e_i – координатный вектор пространства $\mathbf{R}^{2000}$ со случайно выбранным индексом i.

Для экспериментов использовался персональный компьютер Pentium IV с тактовой частотой 256 Мгц. Метод GMRES был представлен библиотечной процедурой gmres системы Matlab; для различных вариантов метода MINRES-N авторами были составлены специальные Matlab-процедуры. Выход из обоих итерационных методов происходил при выполнении условия

$$\|r_m\|_2 < \varepsilon = 10^{-8},$$

где $r_m = b - Ax_m$ – невязка m-го приближения x_m. Величина $\|r_m\|_2$ в алгоритме MINRES-N2 вычислялась, как это описано в [1, раздел 6]. Сходным образом длина невязки вычислялась в остальных вариантах метода MINRES-N.

Результаты экспериментов сведены в четыре таблицы, отвечающие соответственно значениям $k = 2, 3, 4, 5$. Столбцы этих таблиц имеют следующий смысл: b указывает тип правой части (см. выше; при этом rand означает псевдослучайный вектор);

cond — число обусловленности матрицы A; n_1 и t_1 – число итераций и время (в секундах) для MINRES-N(1+k), а n_2 и t_2 дают ту же информацию для GMRES.

Изучение таблиц приводит к трем основным выводам.

1. Для данного класса линейных систем метод MINRES-N(1+k) требует для своей сходимости значительно большего числа итераций, чем GMRES. При этом отношение n_1/n_2 растет с ростом k. Как правило,

$$\frac{n_1}{n_2} \in [k, k+1], \qquad k = 2, 3, 4, 5.$$

2. Несмотря на большее число итераций, MINRES-N(1+k) (почти) всегда затрачивает меньшее время по сравнению с GMRES. При этом для $k = 2, 3$ различие во времени счета весьма внушительно: $t_1 \ll t_2$.

3. Различия в типе правой части b не ведут к различиям в характере полученных результатов.

Не следует думать, что неравенство

$$n_1 > n_2$$

имеет место для любой линейной системы. Напротив, для произвольной (невырожденной) матрицы A нетрудно указать правые части b, для которых MINRES-N сходится (в смысле числа итераций) заведомо быстрее, чем GMRES. Предположим, например, что

$$A^{-1}b = g(A, A^*)b, \tag{17}$$

где $g(x, y)$ – многочлен невысокой степени k от некоммутирующих переменных x и y, причем g содержит хотя бы один моном, включающий в себя множитель y. Тогда в точной арифметике MINRES-N найдет решение системы (1), т.е. вектор $x_0 = A^{-1}b$, как только будет закончено построение ортонормированного базиса в подпространстве $\mathcal{L}_k$. Метод GMRES, не использующий информацию, предоставляемую сопряженной матрицей A^*, аналогичным свойством не обладает.

Рассмотрим подробней ситуацию, когда в (17) g – линейный многочлен от x и y. В этом случае для некоторых скаляров α, β и γ имеем

$$A^{-1}b = \alpha b + \beta Ab + \gamma A^* b, \tag{18}$$

причем $\gamma \neq 0$. Зададим какие-либо значения β_0 и γ_0 для коэффициентов β и γ (с учетом условия $\gamma_0 \neq 0$) и перепишем (18) в виде

$$(I_n - \beta_0 A^2 - \gamma_0 AA^*)b = \alpha Ab,$$

откуда следует, что в качестве неизвестного коэффициента α и искомой правой части b можно взять произвольное собственное значение и соответствующий собственный вектор задачи

$$Cx = \lambda Ax,$$

где

$$C = I_n - \beta_0 A^2 - \gamma_0 AA^*.$$

MINRES-N3

b	cond	n_1	t_1	n_2	t_2
rand	3.8428	129	5.0680	47	12.4370
rand	5.7831	186	6.1090	67	17.2500
Ae	5.3192	276	10.0780	105	27.8750
Ae	2.9437	141	4.2810	52	13.1880
A^*e_i	194.0375	495	24.7970	239	70.3440
A^*e_i	1.6426	45	1.9690	20	5.0940

MINRES-N4

b	cond	n_1	t_1	n_2	t_2
rand	5.8923	255	9.6100	70	18.1880
rand	2.5477	155	5.6560	46	11.9220
rand	11.6147	295	11.3590	91	23.8590
Ae	2.8855	215	8.4060	52	13.1880
Ae	3.4661	259	9.7500	69	17.7660
A^*e_i	2.9855	87	3.2970	27	6.9220

MINRES-N5

b	cond	n_1	t_1	n_2	t_2
rand	2.7547	178	5.9690	42	11.2810
rand	5.5076	273	10.6250	72	18.6410
rand	4.4511	328	14.0540	82	21.9690
Ae	2.9520	278	12.0470	63	16.6570
Ae	3.5473	253	10.9720	59	15.4370
A^*e_i	3.6827	118	4.2190	30	8.4370

MINRES-N6

b	cond	n_1	t_1	n_2	t_2
rand	4.2080	254	10.6710	57	14.6720
rand	2.7391	266	10.9350	49	12.4530
rand	16.1957	722	49.1410	173	48.4220
Ae	4.0888	452	22.7820	90	23.4840
Ae	3.1573	314	14.2810	59	15.4220
A^*e_i	4.7981	332	14.8120	68	17.5470

Литература

1. М. Дана, А. Г. Зыков, Х. Д. Икрамов, *Метод минимальных невязок для специального класса линейных систем с нормальными матрицами коэффициентов.* — Ж. вычисл. матем. матем. физ. **45** (2005), 1928-1937.
2. М. Дана, Х. Д. Икрамов, *О решении систем линейных уравнений, матрицы которых являются малоранговыми возмущениями эрмитовых матриц.* — Вестн. МГУ. Серия “Вычисл. математика и кибернетика” No. 1 (2005), 15–22.

Dana M., Ikramov Kh. D. Solving systems of linear equations whose matrices are low-rank perturbations of Hermitian matrices, revisited.

MINRES-N is a minimal residual algorithm originally developed by the authors for solving systems of linear equations with normal coefficient matrices whose spectra lie on algebraic curves of low degree. In a previous publication, the authors showed that a variant of MINRES-N called MINRES-N2 is applicable to nonnormal matrices A for which

$$\operatorname{rank}(A - A^*) = 1.$$

This fact is extended to nonnormal matrices A such that

$$\operatorname{rank}(A - A^*) = k, \qquad k \geqslant 1.$$

Университет Курдистана,
г. Санандадж, Исламская
Республика Иран

E-mail: mansourdana2001@yahoo.com

Московский государственный
университет

E-mail: ikramov@cs.msu.su

Поступило 16 января 2005 г.