

В. П. Ильин

О ПРОЕКЦИОННЫХ МЕТОДАХ В ПОДПРОСТРАНСТВАХ КРЫЛОВА

§1. ВВЕДЕНИЕ

В 1937 г. польский математик С. Качмаж опубликовал в [1] итерационный метод решения систем линейных алгебраических уравнений (СЛАУ), имеющий уникальную общность и изящную геометрическую интерпретацию. Каждое из уравнений системы

$$(Au)_i \equiv \sum_{j=1}^N a_{i,j} u_j = f_i, \quad i = 1, 2, \dots, N, \quad (1)$$

может ассоциироваться с гиперплоскостью в N -мерном пространстве \mathcal{R}^N . Для произвольного начального приближения u^0 , представляющего собой радиус-вектор некоторой точки в \mathcal{R}^N , определяем ортогональную проекцию $u_1^1 = P_1(u^0)$ на первую гиперплоскость, где P_1 есть соответствующий оператор проектирования. Далее проецируем полученную точку на вторую гиперплоскость и находим $u_2^1 = P_2(u_1^1)$. Продолжая этот процесс до конца, вычисляем значения

$$u_i^1 = P_i(u_{i-1}^1), \quad i = 2, \dots, N, \quad (2)$$

а полученный вектор $u^1 = (u_1^1, \dots, u_N^1)^T$ называем первым итерационным приближением алгоритма Качмажа, который можно записать в виде

$$\begin{aligned} u_i^{n+1} &= P_1(u^n), \quad u_i^{n+1} = P_i(u_{i-1}^n), \quad i = 2, \dots, N, \\ u^{n+1} &= (u_1^{n+1}, u_2^{n+1}, \dots, u_N^{n+1}), \quad n = 0, 1, \dots \end{aligned} \quad (3)$$

Если СЛАУ (1) имеет единственное решение, то все соответствующие гиперплоскости пересекаются в одной точке u , и именно к ней очевидным образом сходится последовательность точек u_i^{n+1} .

Ключевые слова: несимметричные разреженные матрицы, весовые методы Чиммино, мультипредобусловленные алгоритмы полусопряженных невязок, методы наименьших квадратов, алгоритмы дефляции.

Работа поддержана грантами РФФИ 16-29-15122 офи-м и РФФИ 18-01-00295.

В 1938 г. итальянский математик Дж. Чиммино в работе [2] предложил итерационный алгоритм, также основанный на ортогональном проектировании на гиперплоскости, но только не последовательно, а одновременным образом. То есть, проецируя точку u^0 на все гиперплоскости, новое приближение находим с помощью усреднения всех значений $u_i^1 = P_i(u^0)$. В итоге метод Чиммино определяется следующим образом:

$$u^{n+1} = \frac{1}{N} \sum_{i=1}^N P_i(u^n), \quad i = 0, 1, \dots \quad (4)$$

Долгие годы проекционные методы Качмажа и Чиммино оставались малозамеченными в профессиональной литературе по вычислительной математике, особенно в монографиях и учебниках (одним из исключений является их краткое изложение в [3], отметим также, что сама работа [1] была повторно опубликована на английском языке в 1993 г.). В 1985 г. В. Хакбушем в книге [4] было показано, что алгоритм Качмажа совпадает с итерационным методом Зейделя, примененным к уравнению, получаемому из системы (1) с помощью левой трансформации Гаусса, т.е. умножения слева на транспонированную матрицу T :

$$A^T A u = A^T f. \quad (5)$$

Отсюда следует, что скорость сходимости итераций метода Качмажа даже меньше, чем в “медленном” алгоритме Зейделя, поскольку число обусловленности СЛАУ (5) квадратичным образом ухудшается, в сравнении с исходной системой (1).

Если алгоритмы Качмажа и Зейделя относятся к классу итерационных методов последовательных смещений [3], то алгоритм Чиммино можно сопоставить с методом одновременных смещений Якоби, в котором все компоненты нового приближения u^{n+1} явно зависят только от компонент вектора предыдущей итерации u^n . С одной стороны, методы одновременных смещений, как правило, медленней соответствующих методов последовательных смещений, однако они (первая группа) обладают преимуществом естественного паралелизма при реализации на многопроцессорных вычислительных системах (МВС).

Что касается метода Качмажа, то несколько направлений его развития были предложены в [5]: введение оптимизирующего параметра ω по аналогии с переходом от алгоритма Зейделя к методу исследовательной верхней релаксации (SOR – Successive Over Relaxation),

переход к альтернирующему варианту с обратным симметризирующим ходом итерации (аналогия с SSOR – Symmetric Successive Over Relaxation), а также вариационное ускорение получаемого итерационного процесса в подпространствах Крылова. Очевидно, что последний подход можно применить и для ускорения метода Чиммино.

Отметим еще, что методы Качмажа и Чиммино легко обобщаются на блочный случай, когда матрица A представляется в виде совокупности m блочных строк

$$A_q \in R^{N_q \times N}, \quad q = 1, \dots, m, \quad N_1 + \dots + N_m = N,$$

причем количество матричных строк N_q в каждом блоке может быть разным. Для простоты далее мы будем считать блоки одинаковыми, т.е. $N_q = M = N/m$. Естественно, для записи блочных вариантов данных методов в формулах (2)–(4) достаточно заменить индексы i, N на q, M соответственно.

Идея построения ортогональных или косоугольных проекций лежит в основе построения многих методов решения симметричных или несимметричных СЛАУ. К ним относятся различные итерационные процессы в подпространствах Крылова, см. [6, 7], зачастую связанные с вариационными свойствами функционалов, а также алгоритмы дефляции, агрегации, грубосеточной коррекции и некоторые другие, которые фактически базируются на малоранговых аппроксимациях матриц, замечательные и перспективные свойства которых представлены в [8, 9].

Отметим, что вследствие проекционных особенностей алгоритмов типа Качмажа и Чиммино, они иногда называются методами “строковых проекций” или “строкового действия” (Row Action), а также методами проективной агрегации (РАМ, Projected Aggregation Method), см. обзоры в [10, 11].

Целью данной работы является конструирование и исследование весовых алгоритмов Чиммино для решения симметричных и несимметричных СЛАУ на основе минимизации невязки, чему посвящен §2. В следующем параграфе проводится развитие предложенных подходов как предобусловленных итерационных алгоритмов в подпространствах Крылова. §4 посвящен методам ускорения полученных итерационных процессов, а в заключении обсуждаются полученные результаты и перспективы дальнейших исследований в данных направлениях. Все рассуждения проводятся для простоты на примере систем с

вещественными квадратными матрицами, хотя они могут быть перенесены и на комплексные СЛАУ с прямоугольными матрицами.

§2. ВЕСОВЫЕ АЛГОРИТМЫ ЧИММИНО

Исходную систему уравнений (1) перепишем в следующем блочном виде:

$$A_q u = f_q, \quad q = 1, \dots, m, \quad f_q \in \mathcal{R}^{N_q}, \quad u \in \mathcal{R}^N, \quad (6)$$

где предполагается, что $m \ll N$, а f_q суть подвекторы полного вектора правой части $f = (f_1^T, \dots, f_m^T)^T$. Рассмотрим векторы $\hat{u}_q \in \mathcal{R}^N$ как возможные решения m недоопределенных и невязанных систем

$$A_q \hat{u}_q = f_q, \quad q = 1, \dots, m. \quad (7)$$

Векторы \hat{u}_q будем искать в виде

$$\hat{u}_q^0 = u^0 + A_q \check{u}_q^0, \quad \check{u}_q^0 \in \mathcal{R}^{N_q}, \quad q = 1, \dots, m, \quad (8)$$

где u^0 – некоторый произвольный вектор. Подставляя представление (8) в СЛАУ (7) (фактически при этом осуществляется правая трансформация Гаусса), для определения “маленьких” векторов \check{u}_q^0 получаем симметричную положительно полуопределенную систему

$$B_q \check{u}_q^0 = r_q^0, \quad B_q = A_q A_q^T \in \mathcal{R}^{N_q}, \quad r_q^0 = f_q - A_q u^0, \quad (9)$$

решение которой с помощью обобщенной обратной матрицы B_q^+ записываем в виде

$$\check{u}_q^0 = B_q^+ r_q^0, \quad q = 1, \dots, m. \quad (10)$$

Отметим, что если прямоугольная матрица A_q имеет полный ранг, то матрица B_q^+ является невырожденной и $B_q = B_q^{-1}$. В противном случае обобщенную обратную матрицу можно вычислять, например, с использованием формулы Гревилля [12]. С помощью (8), (10) “большие” векторы \hat{u}_q записываются как

$$\hat{u}_q^0 = u^0 + A_q^T B_q^+ r_q^0 = u^0 + A_q^T (A_q A_q^T)^+ r_q^0. \quad (11)$$

Отметим, что q -й блочный подвектор невязки, соответствующий вектору \hat{u}_q^0 из (11), равен нулю, так как

$$\hat{r}_q^0 = (f - A u_q^0 - A u_q^0)_q = r_q^0 - A_q A_q^T (A_q A_q^T)^+ r_q^0 = 0.$$

Построим теперь итерационный процесс для решения СЛАУ (6), используя произвольный вектор $u^0 \in \mathcal{R}^N$ в качестве начального приближения. Определим первое итерационное приближение u^1 в следующей форме:

$$\begin{aligned} u^1 &= u^0 + c_1^0 \hat{u}_1^0 + \dots + c_m^0 \hat{u}_m^0 = u^0 + V_0 c^0, \\ c^0 &= (c_1^0, \dots, c_m^0)^T \in \mathcal{R}^m, \quad V_0 = (\hat{u}_1^0, \dots, \hat{u}_m^0) \in \mathcal{R}^{N,m}. \end{aligned} \quad (12)$$

Неизвестные коэффициенты c_q^0 в (12) будем определять из условия минимума невязки r^1 , соответствующей вектору u^1 :

$$r^1 = f - Au^1 = r^0 - W_0 c^0, \quad W_0 = AV_0 \in \mathcal{R}^{N,m}.$$

Полагая формально в (13) $r^1 = 0$, получаем для вектора c^0 переопределенную систему уравнений

$$W_0 c^0 = r^0, \quad (13)$$

для построения обобщенного решения которой можно применить, например, сингулярное разложение (SVD) или QR-алгоритм [7]. Мы будем искать обобщенное нормальное решение СЛАУ (13), имеющее наименьшую норму и минимизирующее норму невязки r^1 , применяя левую трансформацию Гаусса:

$$W_0^T W_0 c^0 = W^T r^0, \quad c^0 = (W_0^T W_0)^+ W_0^T r^0. \quad (14)$$

Отсюда на первой итерации находим векторы

$$\begin{aligned} u^1 &= u^0 + V_0 (W_0^T W_0)^+ W_0^T r^0, \\ r^1 &= T_0 r^0, \quad T_0 = I - P_0, \quad P_0 = W_0 (W_0^T W_0)^+ W_0^T, \end{aligned} \quad (15)$$

где P_0 и T_0 являются операторами ортогонального проектирования в силу легко проверяющих соотношений

$$P_0 = P_0^T = P_0^2, \quad T_0 = T_0^T = T_0^2.$$

Таким образом, формулы (15) определяют итерационный процесс

$$\begin{aligned} r^0 &= f - Au^0, \quad r^{n+1} = T_n r^n, \quad T_n = I - P_n, \\ P_n &= W_n (W_n W_n)^+ W_n^T, \quad u^{n+1} = u^n + V_n (W_n^T W_n)^+ W_n^T r^n, \end{aligned} \quad (16)$$

который обладает следующими свойствами ортогональности:

$$W_n r^{n+1} = 0, \quad n = 0, \dots, \quad W_n = AV_n, \quad V_n = (\hat{u}_1^n, \dots, \hat{u}_m^n)^T, \quad (17)$$

где $\hat{u}_q^n \in \mathcal{R}^N$, а матрицы ортогонального проектирования P_n, T_n представляются по аналогичным (11) и (15) формулам соответственно при

замене в них индекса “0” на “n”. Итерационная формула из (16) может быть переписана следующим образом:

$$u^{n+1} = u^n - B_n^{(1)}A(f - Au^n), \quad B_n^{(1)} = V_n(V_n^T A^T AV_n)^+ V_n^T, \quad (18)$$

откуда видно, что это есть алгоритм с динамически меняющимися предобуславливающими матрицами $B_n^{(1)}$, примененный к уравнению, получаемому из исходного путем левой трансформации Гаусса. Очевидно, что последний факт приводит к квадратичному возрастанию числа обусловленных СЛАУ. Во избежании этого эффекта можно рассмотреть другой алгоритм построения обобщенного решения переопределенной системы (1) для вектора коэффициентов c^0 . А именно, мы умножим обе части данного равенства на матрицу V_0^T (вместо W_0^T), в результате чего получим уравнение

$$V_0^T AV_0 c^0 = V_0^T r^0. \quad (19)$$

Отсюда мы имеем решение

$$c^0 = (V_0^T AV_0)^+ V_0^T r^0, \quad (20)$$

а вместо (15) приходим к соотношениям

$$u^1 = u^0 + B_0^{(2)} r^0, \quad B_0^{(2)} = V_0(V_0^T AV_0)^+ V_0^T. \quad (21)$$

Здесь $B_0^{(2)}$ — отличная от $B_0^{(1)}$ предобуславливающая матрица, которая фактически является малоранговым приближением к матрице, являющейся обратной (или обобщенно обратной) к A . Она будет симметричной, если только таким свойством обладает матрица исходной СЛАУ. Если прямоугольная матрица $V_0 \in \mathcal{R}^{N,m}$ имеет ранг m , то $B_0^{(2)}$ является вырожденной или невырожденной одновременно с матрицей A .

Отметим при этом, что матрицы $B_n^{(1)}$ из (18) суть малоранговые приближения (по соответствующим базисам V_n) к матрице $(A^T A)^+$. Таким образом, при определении вектора c^0 в соответствии с (21) вместо (16) мы получаем итерационный процесс

$$\begin{aligned} r^0 &= f - Au^0, \quad r^{n+1} = T_n^{(2)} r^n, \quad T_n^{(2)} = I - P_n^{(2)}, \quad P_n^{(2)} = AB_n^{(2)}, \\ u^{n+1} &= u^n + B_n^{(2)} r^n, \quad B_n^{(2)} = V_n(V_n^T AV_n)^+ V_n^T. \end{aligned} \quad (22)$$

Нетрудно понять, что здесь матрицы $P_n^{(2)}, T_n^{(2)}$ являются проекторами, как и P_n, T_n из (16), однако свойство симметричности у них имеется, только если им обладает матрица A . В данном случае, векторы невязки обладают также и свойством ортогональности (17).

§3. МЕТОДЫ ТИПА ЧИММИНО В ПОДПРОСТРАНСТВАХ
КРЫЛОВА

На основе проекционных алгоритмов (16) или (22) могут быть построены различные мульти-предобусловленные итерационные методы в подпространствах Крылова, обладающие определенными вариационными свойствами. Мы представим их в единообразной форме с предобуславливающими матрицами B_n , которые принимают вид $B_n^{(1)}$, или $B_n^{(2)}$, или какой-то другой в каждом конкретном случае соответственно. Предполагая для общности матрицы A и $B_{n,l}$ несимметричными, приведем формулы динамических мульти-предобусловленных методов полусопряженных направлений (MPSCD), см. [13], которые по скорости сходимости итераций эквивалентны “гибкому” обобщенному методу минимальных невязок FGMRES [6]:

$$\begin{aligned} r^0 &= f - Au^0, \quad n = 0, 1, \dots : u^{n+1} = u^n + P_n \bar{\alpha}_n, \\ r^{n+1} &= r^n - AP_n \bar{\alpha}_n = r^q - AP_q \bar{\alpha}_q - \dots - AP_n \bar{\alpha}_n, \quad 0 \leq q \leq n, \\ P_n &= (p_1^n \dots p_{M_n}^n) \in \mathcal{R}^{N, M_n}, \quad \bar{\alpha}_n = (\alpha_{n,1} \dots \alpha_{n, M_n})^T \in \mathcal{R}^{M_n}, \quad (23) \\ (Ap_k^n, A^\gamma p_{k'}^{n'}) &= \rho_{n,k}^{(\gamma)} \delta_{n,n'}^{k,k'}, \quad \rho_{n,k}^{(\gamma)} = (Ap_k^n, A^\gamma p_k^n), \\ \gamma &= 0, 1; \quad n' = 0, 1, \dots, n-1; \quad k, k' = 1, 2, \dots, M_n. \end{aligned}$$

Здесь P_n суть “направляющие” матрицы, состоящие из M_n столбцов – направляющих векторов, количество которых может меняться на различных итерациях. Векторы $\bar{\alpha}_n \in \mathcal{R}^{M_n}$ являются обобщениями соответствующих скалярных коэффициентов для предобусловленных методов полусопряженных направлений (при $M_n = 1$ для всех n), а величины $\gamma = 0, 1$ относятся к алгоритмам полусопряженных градиентов и полусопряженных невязок соответственно. Мульти-предобусловленные направляющие векторы по условиям ортогональности из (23) определяются в общем случае из “длинных” рекурсий

$$\begin{aligned} p_l^0 &= B_{0,l}^{-1} r^0, \quad p_l^{n+1} = B_{n+1,l}^{-1} r^{n+1} - \sum_{k=0}^n \sum_{l=1}^{M_k} \beta_{n,k,l}^{(\gamma)} p_l^k, \quad n = 0, 1, \dots; \\ B_{n,l} &\in \mathcal{R}^{N, N}, \quad l = 1, \dots, M_n; \quad \gamma = 0, 1, \\ \bar{\beta}_{n,k}^{(\gamma)} &= \{\beta_{n,k,l}^{(\gamma)}\} = (\beta_{n,k,1}^{(\gamma)} \dots \beta_{n,k,M_n}^{(\gamma)})^T \in \mathcal{R}^{M_n}. \end{aligned} \quad (24)$$

При этом имеют место следующие соотношения для коэффициентов рекуррентных выражений (24), а также для функционалов невязки при $q = 0, \dots, n$:

$$\begin{aligned} \beta_{n,k,l}^{(\gamma)} &= -(A^\gamma p_l^k, AB_{n+1,l}^{-1} r^{n+1}) / \rho_{n,l}^{(\gamma)}, \quad n = 0, 1, \dots; \\ k &= 0, \dots, n; \quad l = 1, \dots, M_n. \end{aligned} \quad (25)$$

$$\Phi_n^{(\gamma)}(r^{n+1}) \equiv (r^{n+1}, r^{n+1}) = (r^q, r^q) - \sum_{k=q}^n \sum_{l=1}^{M_n} (r^q, A^\gamma p_l^k)^2 / \rho_{k,l}^{(\gamma)}.$$

Заметим, что если $\gamma = 1$ или $A = A^T$, то формулы (23), (24) обеспечивают минимум функционала $\Phi_n^\gamma(r^n)$ в блочном подпространстве Крылова размерности $M = M_0 + \dots + M_n$, см. [13]:

$$\begin{aligned} \mathcal{K}_M(r^0, A) &= \text{Span}\{B_{0,1}^{-1} r^0, \dots, B_{0,M_0}^{-1} r^0, \dots, \\ &AB_{1,1}^{-1} r^0, \dots, AB_{1,M_1}^{-1} r^0, \dots, AB_{n,1}^{-1} r^0, \dots, AB_{n,M_n}^{-1} r^0\}. \end{aligned} \quad (26)$$

Отметим, что если матрица A симметричная, то формулы (24) дают короткие рекурсии для направляющих векторов ($\beta_{n,k,l}^{(\gamma)} = 0$ при $k < n$), а для MPSCD мы получаем мульти-предобусловленные методы сопряженных градиентов или сопряженных невязок (для $\gamma = 0$ или $\gamma = 1$ соответственно).

Если же $A \neq A^T$, то наиболее целесообразным, по-видимому, является применение мульти-предобусловленного метода полусопряженных невязок MPSCR (Multi-Preconditioned Semi-Conjugate Residual). В этом случае при решении плохо обусловленных несимметричных СЛАУ, связанным с большим количеством итераций из-за ресурсоёмкости длинных рекурсий (главным образом вследствие увеличения объема используемой памяти), их приходится принудительно укорачивать, что делается путем или введения так называемой процедуры рестарта, или ограничения количества ортогонализируемых направляющих векторов (или матриц — при мульти-предобуславливании), или использования обоих подходов одновременно. Во всех этих случаях скорость сходимости итераций падает, иногда очень сильно, что является неизбежной платой за экономию памяти.

Для борьбы с данным эффектом мы рассмотрим применение метода наименьших квадратов (МНК, [14]), ограничиваясь при этом использованием рестартов “в чистом виде”. Предположим для простоты, что рестарты повторяются периодически через одинаковое количество

итераций m . Это означает, что на каждой итерации с номером

$$n_t = mt, \quad t = 0, 1, \dots,$$

вектор невязки вычисляется не из рекуррентных соотношений (13), а из исходного уравнения, т.е.

$$r^{n_t} = f - Au^{n_t}, \quad (27)$$

после чего рекурсии опять используются обычным образом. Более точно такой итерационный процесс удобно описать в двухиндексных обозначениях соответствующих номеров последовательных приближенных векторов

$$u^{n_t} = u^{t,0}, \quad u^n = u^{t,k}, \quad k = n - n_t, \quad n \in [n_t, n_{t+1}].$$

При этом связь между соседними рестартовыми приближениями можно формально установить посредством соответствующих операторов перехода, или матриц и векторов:

$$u^{n_{t+1}} = B_t u^{n_t} + g_t, \quad (28)$$

где конкретный вид B_t и g_t можно выразить с помощью формулы и применяемого итерационного метода.

§4. МЕТОДЫ ВНУТРЕННЕГО И ВНЕШНЕГО УСКОРЕНИЯ

Пусть нам уже известны значения “рестартовых” приближений $u^{n_0}, u^{n_1}, \dots, u^{n_t}, n_0 = 0$. Для коррекции последнего итерационного приближения сформируем линейную комбинацию векторов

$$\begin{aligned} \hat{u}^{n_t} &= u^{n_t} + b_1 v_1 + \dots + b_t v_t = u^{n_t} + V_t \bar{b}, \quad \bar{b} = (b_1, \dots, b_t)^T, \\ V_t &= \{v_k = u^{n_k} - u^{n_{k-1}}; k = 1, \dots, t\} \in \mathcal{R}^{N,t}, \end{aligned} \quad (29)$$

коэффициенты b_n которой будем находить из обобщенного решения переопределенной системы, получаемой после умножения уравнений (29) на исходную матрицу A :

$$W_t \bar{b} = r^{n_t} = f - Au^{n_t}, \quad W_t = AV_t. \quad (30)$$

Как и в формулах (14)–(20) для весовых методов Чиммино, здесь есть несколько путей для решения СЛАУ (30): использование разложения SVD или QR для матрицы W_t , нахождение обобщенной обратной матрицы W_t^+ по формулам Гревилля, вычисление нормального решения методом наименьших квадратов, или применения “облегченного”

(по числу обусловленности) преобразования путем сведения к квадратной системе

$$V_t^T AV_t \bar{b} = V_t^T r^{n_t}. \quad (31)$$

Во всех этих случаях после нахождения вектора \bar{b} и проведения коррекции итерационного приближения по формуле (29) следующий “рестарт” начинается с вычисления невязки

$$r^{n_t+1} = f - Au^{n_t+1}, \quad u^{n_t+1} = \hat{u}^{n_t}. \quad (32)$$

Нетрудно заметить, что поскольку векторы u^{n_t} в (28) вычисляются последовательно, то вместо рассмотренного подхода с хранением всех рестартовых приближений для их оптимизации можно применить другой алгоритм, который фактически является в некотором смысле обобщением метода полусопряженных невязок [17]. Мы его рассмотрим как частный случай формул (23)–(24) для $\gamma = 1$ и $M_n = 1$ т.е. без мульти-предобуславливания. Если векторы u^{n_t} и r^{n_t} известны, то начало следующего “рестарта” определяем по “стандартным” формулам для крыловских процессов:

$$\begin{aligned} u^{n_t+1} &= u^{n_t} + \alpha_t p^t = u^0 + \alpha_0 p^0 + \dots + \alpha_t p^t, \\ r^{n_t+1} &= r^{n_t} + \alpha_t A p^t = r^0 - \alpha_0 A p^0 - \dots - \alpha_t A p^t, \end{aligned} \quad (33)$$

где α_t, p^t суть некоторые коэффициенты и направляющие векторы (обычно полагается $p^0 = r^0$). Соотношения (33) обладают тем замечательным свойством, что если векторы p^t удовлетворяют условиям ортогональности

$$(A p^k, A p^t) = \rho_t \delta_{k,t}, \quad \rho_t = (A p^t, A p^t), \quad (34)$$

где $\delta_{k,t}$ есть символ Кронекера, то формулы (33) обеспечивают для любого значения t минимум нормы невязки r^{n_t+1} , если только коэффициенты α_t определяются соотношениями

$$\alpha_t = \sigma_t / \rho_t, \quad \sigma_t = (A r^{n_t}, r^{n_t}). \quad (35)$$

В свою очередь, направляющие “рестартовые” векторы будут обладать свойствами (34), если их вычислять с помощью длинных рекурсий

$$p^{t+1} = r^{n_t+1} - \sum_{k=0}^t \beta_{t,k} p^k, \quad \beta_{t,k} = (A p^k, A r^{n_t+1}) / \rho_t, \quad (36)$$

которые фактически представляют собой реализацию процесса ортогонализации Грама–Шмидта. Отметим, что для обеспечения устойчивости вычислений при решении плохо обусловленных матриц формулы для $\beta_{t,k}$ в (36) следует видоизменить, перейдя к модифицированному методу Грама–Шмидта (MGS, см. [17] и цитируемую там литературу).

Поскольку сами рестартовые приближения u^{nt} связаны между собой соотношением (28), реализующим m “межрестартовых” внутренних итераций, то формулы (33)–(36) можно рассматривать как внешний итерационный процесс общего двухуровневого метода в подпространстве Крылова с “псевдополиномиальным” предобуславливанием, поскольку оператор B_t включает в общем случае предобуславливатели $B_{n,l}$ из (24).

Рассмотрим теперь возможности ускорения описанных блочных итерационных методов крыловского типа на основе дефляционного подхода, предложенного в 1987 году Р. А. Николайдесом [16], а впоследствии развиваемого многими авторами, см. обзор в [17]. В данном случае, помимо традиционных вариационных и/или ортогональных свойств вычислительных последовательных приближений, на них накладываются дополнительные условия ортогональности специально вводимому m -мерному фиксированному дефляционному подпространству, ассоциированному с прямоугольной матрицей

$$V = (v_1 \dots v_m) \in \mathcal{R}^{N,m}. \quad (37)$$

Мы обсудим предварительно использование метода дефляции в применении к алгоритму сопряженных невязок для решения СЛАУ с симметричной матрицей A . Во-первых, здесь существует подход к оптимизации, в определенном смысле, вектора начального итерационного приближения u^0 . Пусть задан произвольный вектор u^{-1} , тогда определим векторы

$$u^0 = u^{-1} + Vc, \quad r^0 = r^{-1} - AVc. \quad (38)$$

Вектор неизвестных коэффициентов $c = (c_1, \dots, c_m)^T$ в (38) будем определять (полагая формально $r^0 = 0$) из решения переопределенной системы

$$Wc = AVc = r^{-1}. \quad (39)$$

Применяя к (39) метод наименьших квадратов (МНК), находим нормальное решение

$$c = (W^T W)^+ W^T r^{-1}, \quad (40)$$

обеспечивающее минимальную норму вектора невязки:

$$\begin{aligned} r^0 &= T_0 r^{-1}, \quad T_0 = I - W(W^T W)^{-1} W^T, \\ (r^0, r^0) &= (r^{-1}, r^{-1}) - (W(W^T W)^{-1} W^T r^{-1}, r^{-1}) \\ &= (W^T W z, z) - ((W^T W)^{-1} z, z), \quad z = W^T r^{-1}, \end{aligned} \quad (41)$$

где матрица T_0 является симметричным (ортогональным) проектором, обладающим следующими свойствами:

$$T_0 = T_0^T = T_0^2, \quad W^T T_0 = T_0 W = 0,$$

то есть пространство $\text{Span}(W)$ принадлежит его ядру $\mathcal{N}(T_0)$. Определяя далее начальный вектор направлений в форме

$$p^0 = r^0 - V(W^T W)^{-1} W^T A r^0 = B r^0, \quad B = I - V(W^T W)^{-1} W^T A, \quad (42)$$

для векторов r^0 и p^0 получаем дефляционные условия ортогональности

$$W^T r^0 = 0, \quad W^T A p^0 = 0. \quad (43)$$

При этом введенная матрица B обладает следующими легко проверяемыми ортогональными свойствами:

$$W^T A B = 0, \quad B V = 0.$$

Дальнейшие итерации получаемого дефляционного алгоритма DCR (Deflated Conjugate Residual) осуществляются по “стандартным” формулам метода сопряженных невязок с предобуславливающей матрицей, роль которой в данном случае играет B из (42):

$$\begin{aligned} u^{n+1} &= u^n + \alpha_n p^n, \quad \alpha_n = \sigma_n / \rho_n, \quad \rho_n = (A p^n, A p^n), \\ r^{n+1} &= r^n - \alpha_n A p^n, \quad \sigma_n = (A B r^n, r^n), \\ p^{n+1} &= B r^{n+1} + \beta_n p^n, \quad \beta_n = \sigma_{n+1} / \sigma_n. \end{aligned} \quad (44)$$

При этом на каждой итерации обеспечивается минимум нормы невязки $\|r^n\|$ в предобусловленном подпространстве Крылова

$$\mathcal{K}_n(A, r^0, B) = \text{Span}(r^0, A B r^0, \dots, (A B)^{n-1} r^0), \quad (45)$$

а вычисляемые векторы удовлетворяют условиям ортогональности

$$\begin{aligned} (A B r^k, r^n) &= \sigma_n \delta_{k,n}, \\ (A p^k, A p^n) &= \rho_n \delta_{k,n}, \quad k = 0, 1, \dots, n-1, \\ W^T r^n &= 0, \quad W^T A p^n = 0, \quad n = 0, 1, \dots \end{aligned} \quad (46)$$

Здесь оказываются справедливы следующие соотношения для функционалов:

$$\begin{aligned} (r^{n+1}, r^{n+1}) &\leq (r^n, r^n) - \frac{(ABr^n, r^n)^2}{(Ap^n, Ap^n)} \\ &\leq (r^k, r^k) - \frac{(ABr^k, r^k)^2}{(Ap^k, Ap^k)} - \dots - \frac{(ABr^n, r^n)^2}{(Ap^n, Ap^n)}, \quad k = 0, 1, \dots, n. \end{aligned} \quad (47)$$

Отметим, что введенная предобуславливающая матрица B является вырожденной, поскольку $BW = 0$, однако соотношения (43) и (46) обеспечивают ортогональность всех векторов невязок ядру матрицы $\bar{A} = AB$, $\mathcal{N}(\bar{A}) = \text{Span}(W)$, что обеспечивает сходимость итерационного процесса (41)–(44).

Приведенные формулы для метода DCR могут быть видоизменены и одновременно упрощены с учетом специальных свойств предобуславливающей матрицы B :

$$B^2 = B, \quad AB = B^T A = B^T AB, \quad (48)$$

аналогично тому, как это было сделано в алгоритме DCG. А именно, определяемый формулами (44) в предобусловленных крыловских подпространствах (45) вектор невязки представим в виде

$$r^{n+1} = P_n(AB)r^n = P_n(B^T AB), \quad (49)$$

где P_n обозначает матричный многочлен степени n . Отсюда следует, что данный итерационный процесс можно представить в следующей экономической форме:

$$\begin{aligned} p^0 &= Br^0, \quad \alpha_n = \sigma_n/\rho_n, \quad q^0 = p^0, \quad s^0 = Ap^0, \\ u^{n+1} &= u^n + \alpha_n p^n, \quad \rho_n = (s^n, s^n), \\ r^{n+1} &= r^n - \alpha_n Ap^n, \quad \sigma_n = (Aq^n, r^n), \quad q^n = Br^n, \\ p^{n+1} &= q^{n+1} + \beta_n p^n, \quad \beta_n = \sigma_{n+1}/\sigma_n, \\ s^{n+1} &\equiv Ap^{n+1} = Aq^{n+1} + \beta_n s^n, \end{aligned} \quad (50)$$

где начальная невязка r^0 определяется из (42). Выполнение формул (50) требует на каждой итерации по одному умножению векторов q^n и r^n на матрицы A и B соответственно.

Рассмотрим теперь формальную модификацию описанного выше метода, которую назовем дефляционным алгоритмом сопряженных

направлений DCD (Deflated Conjugate Direction). Пусть даны два различных семейства векторов v_k и w_k , являющихся столбцами матриц

$$V = (v_1 \dots v_{m_v}) \in \mathcal{R}^{N, m_v}, \quad W = (w_1 \dots w_{m_w}) \in \mathcal{R}^{N, m_w}. \quad (51)$$

При этом введенная матрица W в (51) является обобщением рассмотренного выше случая $W = AV$ и $m_w = m_v$. Выбор начального приближения будем осуществлять, как и ранее:

$$u^0 = u^{-1} + Vc, \quad r^0 = r^{-1} - AVc, \quad c = (c_1, \dots, c_{m_v})^T \in \mathcal{R}^{m_v}.$$

Однако нахождение вектора c проводим из обобщенного решения следующей СЛАУ, которая является совместимой, но имеет прямоугольную матрицу:

$$\begin{aligned} Mc &\equiv W^T AVc = W^T r^{-1}, \quad c = (W^T AV)^+ W^T r^{-1}, \\ M &\in \mathcal{R}^{m_w, m_v}, \quad M^+ \in \mathcal{R}^{m_v, m_w}. \end{aligned} \quad (52)$$

В этом случае вектор начальной невязки определяется как

$$r^0 = \hat{B}r^{-1}, \quad \hat{B} = I - AV(W^T AV)^+ W^T, \quad (53)$$

причем выполняются условия ортогональности

$$W^T r^0 = 0, \quad W^T \hat{B} = 0, \quad \hat{B}AV = 0. \quad (54)$$

Начальный направляющий вектор определим с помощью новой преобуславливающей матрицы:

$$p^0 = \check{B}r^0, \quad \check{B} = I - V(W^T AV)^+ W^T A, \quad (55)$$

обеспечивающей выполнение ортогональных соотношений

$$\check{V} = 0, \quad W^T A\check{B} = 0, \quad W^T Ap^0 = 0. \quad (56)$$

Последующие итерационные приближения в методе DCD выполняем по формулам

$$\begin{aligned} u^{n+1} &= u^n = \alpha_n p^n, \quad \alpha_n = \sigma_n / \rho_n, \quad \rho_n = (Ap^n, Ap^n), \\ r^{n+1} &= r^n - \alpha_n Ap^n, \quad \sigma_n = (A\check{B}r^n, r^n), \\ p^{n+1} &= \check{B}r^{n+1} + \beta_n p^n, \quad \beta_n = \sigma_{n+1} / \sigma_n, \end{aligned} \quad (57)$$

из которых при условиях (52), (53) для всех n следуют такие свойства ортогональности:

$$\begin{aligned} W^T r^{n+1} &= 0, \quad W^T Ap^{n+1} = 0, \\ (A\check{B}r^k, r^n) &= \sigma_n \delta_{k,n}, \quad (Ap^k, Ap^n) = \rho_n \delta_{k,n}, \quad k = 0, 1, \dots, n-1. \end{aligned} \quad (58)$$

Введенные в (51) матрицы V и W при $m_v, m_w < N$ допускают интерпретацию V^T и W как операторов сужения и продолжения, но относительно пространств различной размерности, что позволяет рассматривать многочисленные формальные обобщения представленных дефляционных алгоритмов.

§5. ЗАКЛЮЧЕНИЕ

Предложенные в данной работе весовые блочные варианты итерационного метода Чиммино являются естественным развитием традиционных представлений для проекционных алгоритмов рассматриваемого класса, которые могут быть эффективно применены к решению широкого круга как симметричных, так и в особенности несимметричных СЛАУ. Описанные подходы к построению мульти-предобусловленных итерационных процессов в подпространствах Крылова, а также к их двухуровневому ускорению с помощью методов дефляции и наименьших квадратов, открывают хорошие перспективы для быстрого решения плохо обусловленных алгебраических систем с большими разреженными матрицами. В качестве основы теоретических исследований здесь могут служить результаты по анализу свойств проекционных и дефляционных методов, изложенные в [18–20] и цитируемых там работах.

Важной особенностью рассмотренных выше итерационных алгоритмов является их естественный параллелизм. Высокая масштабируемая производительность здесь может быть достигнута на основе декомпозиции областей, реализующей блочные варианты методов Чиммино–Шварца в подпространствах Крылова, при использовании описанных подходов внешнего и внутреннего ускорения итераций. Двухуровневый характер результирующего вычислительного процесса позволяет эффективно применять технологии гибридного программирования на гетерогенных суперкомпьютерах с распределенной и иерархической общей памятью, включая MPI-процессы, многопоточковые вычисления и векторизацию операций. Здесь главная роль принадлежит интенсивным вычислительным экспериментам, которые являются предметом дальнейших исследований по данным направлениям.

СПИСОК ЛИТЕРАТУРЫ

1. S. Kaczmarz, *Angenaherte Auflosung von Systemen lineare Gleichungen*. — Bull. Int. Acad. Polon. Sci. Lett. **35** (1937), 355–357.

2. G. Cimmino, *Calcolo approssimati de soluzioni dei sistemi di equazioni linear.* — Ric. Sci. **11.9** (1938), 326–333.
3. В. П. Ильин, *Методы и технологии конечных элементов*, Новосибирск, изд. ИВМиМГ СО РАН, 2007.
4. W. Hackbush, *Multigrid Methods and Applicatios*, Springer-Verlag, Berlin, 1985.
5. В. П. Ильин, *Об итерационном методе Качмажа и его обобщениях.* — Сиб. ж. индустр. мат. **9**, No. 3 (2006), 39–40.
6. Y. Saad, *Iterative Methods for Sparse Linear Systems*, PWS, 1996.
7. G. H. Golub, Y. C. F. Loan, *Matrix Computations*, John's Hopkins Univ. Press, 1996.
8. M. T. Chu, R. E. Funderlic, G. H. Golub, *A rank-one reduction formula and its application to matrix factorization.* — SIAM Rev. **37** (1999), 512–530.
9. L. Hubert, J. Meulman, W. Heise, *Two purposes for matrix factorization: A historical appraisal.* — SIAM Rev. **42**, No. 1 (2000), 68–82.
10. G. Appleby, D. S. Smolarski, *A linear acceleration row action method for projecting onto subspaces.* — Electr. Trans. Numer. Aanal. Kent State Univ. **20** (2005), 253–275.
11. H. Scolnik, N. Echebest, M. T. Guardarrdarucci, M. C. Vacachino, *A class of optimized row projection methods for solving large nonsymmetric linear systems.* — Appl. Numer. Math. **41** (2002), 499–513.
12. Ф. П. Гантмахер, *Теория матриц*, М., Наука, 1968.
13. В. П. Ильин, *О проблемах решения больших СЛАУ.* — Зап. научн. семин. ПОМИ **439** (2015), 112–127.
14. В. П. Ильин, *Двухуровневые методы наименьших квадратов в подпространствах Крылова.* — Зап. научн. семин. ПОМИ **463** (2017), 224–239.
15. V. P. Il'in, *Methods of semiconjugate directions.* — Russ. J. Numer. Anal. Math. Modelling **23**, No. 4 (2008), 369–387.
16. R. A. Nicolaydes, *Deflation of conjugate gradients with application to boundary value problems.* — SIAM J. Numer. Anal. **24** (1987), 335–365.
17. Y. Gurieva, V. P. Il'in, *On parallel computational technologies of augmented domain decomposition methods*, Parallel Computing Technologies, 13th International Conference PaCT, Petrozavodsk, Russia (2015), 33–46.
18. J. H. Bramble, J. E. Pasciak, J. Wang, J. Xu, *Convergence estimates for product iterative methods with applications to domain decomposition.* — Math. Comp. **57**, No. 195 (1991), 1–21.
19. Л. Ю. Колотилина, *Переобуславливание систем линейных алгебраических уравнений с помощью двойного исчерпывания. I. Теория.* — Зап. научн. семин. ПОМИ **229** (1995), 95–152.
20. J. Liesen, M. Rozloznic, Z. Strakokos, *Least squares residuals and minimal residual methods.* — SIAM J. Sci. Comput. **23**, No. 5 (2002), 1503–1525.

Il'in V. P. Projection methods in Krylov subspaces.

The paper considers preconditioned iterative methods in Krylov subspaces for solving large systems of linear algebraic equations with sparse coefficient matrices arising in solvibg multidimensional boundary-value

problems by finite volume or finite element methods of different orders on unstructured grids. Block versions of the weighted Cimmino methods, based on various orthogonal and/or variational approaches and realizing preconditioning functions for two-level multi-preconditioned semi-conjugate residual algorithms with periodic restarts, are proposed. At the inner iterations between restarts, additional acceleration is achieved by applying deflation methods, providing low-rank approximations of the original matrix and playing the part of an additional preconditioner. At the outer level of the Krylov process, in order to compensate the convergence deceleration caused by restricting the number of the orthogonalized direction vectors, restarted approximations are corrected by using the least squares method. Scalable parallelization of the methods considered, based on domain decomposition, where the commonly used block Jacobi–Schwarz iterative processes is replaced by the block Cimmino–Schwarz algorithm, is discussed. Hybrid programming technologies for implementing different stages of the computational process on heterogeneous multi-processor systems with distributed and hierarchical shared memory are described.

Институт вычислительной математики
и математической геофизики СО РАН,
Новосибирский государственный университет
E-mail: ilin@sccc.ru

Поступило 7 ноября 2018 г.