

УДК 519.63

doi 10.26089/NumMet.v16r115

О ТЕХНОЛОГИЯХ УСКОРЕНИЯ ПАРАЛЛЕЛЬНЫХ МЕТОДОВ ДЕКОМПОЗИЦИИ**Я. Л. Гурьева¹, В. П. Ильин²**

Одним из главных препятствий масштабированному распараллеливанию алгебраических методов декомпозиции для решения сверхбольших разреженных систем линейных алгебраических уравнений (СЛАУ) является замедление скорости сходимости аддитивного итерационного алгоритма Шварца в подпространствах Крылова при увеличении количества подобластей. Целью настоящей статьи является сравнительный экспериментальный анализ различных приемов ускорения итераций: параметризованное пересечение подобластей, использование специальных интерфейсных условий на границах смежных подобластей, а также применение грубосеточной коррекции (агрегации, или редукции) исходной СЛАУ для построения дополнительного предобусловливателя. Распараллеливание алгоритмов осуществляется на двух уровнях программными средствами для распределенной и общей памяти. Тестовые СЛАУ получаются при помощи конечно-разностных аппроксимаций задачи Дирихле для диффузионно-конвективного уравнения с различными значениями конвективных коэффициентов на последовательности сгущающихся сеток.

Ключевые слова: декомпозиция областей, аддитивный метод Шварца, алгоритмы редукции, предобусловленные крыловские процессы, масштабируемое распараллеливание, распределенная и общая память, вычислительный эксперимент.

1. Введение. Настоящая статья посвящена ключевой проблеме обеспечения масштабированного параллелизма при решении сверхбольших систем линейных алгебраических уравнений с разреженными матрицами, возникающими при решении многомерных сеточных краевых задач. Эта проблема связана с двумя антагонистическими обстоятельствами. С одной стороны, ускорение вычислений на многопроцессорных вычислительных системах требует применения декомпозиции расчетной области на большое количество подобластей, для каждой из которой соответствующая подзадача решается синхронно с остальными путем формирования “своего” MPI-процесса и применения “внутреннего” ускорения за счет многопоточковых технологий. С другой стороны, рост числа подобластей неизбежно ведет к замедлению внешнего итерационного процесса. Для устранения этого недостатка активно ведутся исследования по повышению эффективности двухуровневых итерационных методов декомпозиции [1–5].

Наибольшие успехи в этом направлении достигаются за счет применения алгоритмов ускорения, которые в литературе называются методами редукции, грубосеточной коррекции, агрегации и дефляции, имеющими общую идеологию в том смысле, что все они используют построение предобусловливателей на основе малоранговых аппроксимаций для матриц, обратных к исходной матрице. В некоторых вариантах эти подходы перекликаются с алгебраическими многосеточными принципами. Теоретические исследования с получением оценок скорости сходимости итераций и оптимизации алгоритмов здесь удаются далеко не всегда, и поэтому большую роль имеют систематические численные эксперименты со сравнительным анализом результативности различных приемов.

Целью данной работы являются экспериментальные исследования параллельных итерационных методов декомпозиции областей (аддитивных алгоритмов Шварца или блочных методов Якоби) при помощи построения предобусловливающих матриц на основе агрегирования СЛАУ с использованием интерполяционных финитных базисных функций лагранжевого типа. В разделе 2 настоящей статьи дается описание постановки проблемы и методов ее решения, в разделе 3 приводятся особенности параллельной реализации предлагаемых алгоритмов, а в разделе 4 анализируются результаты модельных расчетов для достаточно представительной серии задач, получаемых из конечно-разностных аппроксимаций диффузионно-конвективного уравнения с различными значениями диффузионных коэффициентов на разных сгущающихся сетках. В заключении обсуждаются некоторые нерешенные проблемы и указываются направления дальнейших исследований.

¹ Институт вычислительной математики и математической геофизики СО РАН, просп. Лаврентьева, 6, 6630090, Новосибирск; ст. науч. сотр., e-mail: yana@lapasrv.sccc.ru

² Институт вычислительной математики и математической геофизики СО РАН, просп. Лаврентьева, 6, 6630090, Новосибирск; гл. науч. сотр., e-mail: ilin@sccc.ru

2. Двухуровневые итерационные процессы с редукцией. Рассмотрим вещественную СЛАУ вида

$$Au = f, \quad A = \{a_{l,l'}\} \in \mathcal{R}^{N,N}, \quad u = \{u_l\}, \quad f = \{f_l\} \in \mathcal{R}^N, \quad (1)$$

с разреженной матрицей A , которую будем предполагать полученной из аппроксимации некоторой многомерной краевой задачи на сетке Ω^h . Методической конкретизацией нам будет служить двумерная задача Дирихле в ограниченной области Ω для диффузионно-конвективного уравнения

$$-\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} + p \frac{\partial u}{\partial x} + q \frac{\partial u}{\partial y} = f(x, y), \quad (x, y) \in \Omega, \quad u|_{\Gamma} = g(x, y), \quad (2)$$

где Γ — граница области Ω , а коэффициенты p и q — заданные функции координат. Иллюстрация вычислительных подходов далее проводится на СЛАУ вида (1), полученной из сеточной аппроксимации краевой задачи (2).

Будем предполагать, что компоненты u_l и f_l соответствуют l -му узлу сетки Ω^h , общее число которых (участвующих в алгебраической системе после исключения из рассмотрения известных в силу краевых условий значений искомого решения) равно N .

Проведем декомпозицию сеточной расчетной области Ω^h , т.е. представим ее сначала в виде объединения P непересекающихся примерно одинаковых подобластей $\Omega^h = \bigcup_{s=1}^P \Omega_s^h$, каждая из которых содержит

“в среднем” число узлов сетки $N_s \cong N/P$, $\sum_{s=1}^P N_s = N$.

Перейдем далее к нумерации узлов, соответствующей их упорядоченности по подобластям — сначала пронумеруем все узлы в Ω_1^h , затем — в Ω_2^h и т.д. Аналогичным образом переупорядочиваем и компоненты векторов u и f , в результате чего исходная СЛАУ (1) принимает следующую блочно-матричную форму:

$$A_{s,s} \bar{u}_s + \sum_{s' \in Q_s} A_{s,s'} \bar{u}_{s'} = f_s, \quad s = 1, \dots, P.$$

Здесь $\bar{u}_s \in \mathcal{R}^{N_s}$ — подвектор вектора u , компоненты которого соответствуют узлам из Ω_s , а Q_s — совокупность номеров соседних к ней подобластей.

Отметим, что приведенная выше формалистика относится к декомпозиции сеточной области без пересечения подобластей и без применения узлов-разделителей, общих для соседних подобластей. Однако для повышения общности и эффективности рассматриваемых далее алгоритмов необходимо перейти к построению “расширенных” подобластей $\bar{\Omega}_s^h$ с пересечениями (индексы “ h ” далее опускаем).

Пусть ω_l — сеточный шаблон, или совокупность соседних узлов для l -го узла, т.е. набор номеров компонент искомого сеточного решения, участвующих в соответствующем l -м уравнении СЛАУ (1). Для сеточной подобласти Ω_s через $\Gamma_s = \Gamma_s^0$ обозначим ее границу, т.е. совокупность узлов, внешних по отношению к Ω_s , но у которых хотя бы один из соседних узлов лежит в Ω_s . Через $\bar{\Omega}_s^0 = \Omega_s \cup \Gamma_s^0$ обозначим замыкание исходной сеточной подобласти Ω_s . Пусть, далее, Γ_s^1 — первая расширенная граница, или первый внешний фронт $\bar{\Omega}_s$, т.е. множество узлов, не лежащих в $\bar{\Omega}_s$, но имеющих хотя бы один соседний узел из $\bar{\Omega}_s^0$. Пусть $\bar{\Omega}_s^1$ — первое расширение $\bar{\Omega}_s^0$. Аналогично определим последующие стадии расширения сеточной подобласти, а количество таких стадий Δ будем называть параметром расширенной подобласти $\bar{\Omega}_s = \bar{\Omega}_s^\Delta = \Omega_s^\Delta \cup \Gamma_s^\Delta$, где узлы из Γ_s^Δ уже не принадлежат подобласти Ω_s^Δ , число узлов которой обозначаем через N_s^Δ . Иллюстрация расширенной подобласти с параметром $\Delta = 3$ приведена на рис. 1.

При построении итерационного процесса Шварца по сеточным подобластям можно различным образом учитывать интерфейсные связи между соседними подобластями. Пусть l -й узел является околосеточным для подобласти Ω_s^Δ , т.е. $l \in \Gamma_s^{\Delta-1}$. Соответствующее уравнение алгебраической системы запишем в виде

$$\left(a_{l,l} + \theta_l \sum_{l' \notin \Omega_s^\Delta} a_{l,l'} \right) u_l^n + \sum_{l \in \Omega_s^\Delta} a_{l,l'} u_{l'}^n = f_l + \sum_{l' \notin \Omega_s^\Delta} a_{l,l'} (\theta_l u_{l'}^{n-1} - u_{l'}^{n-1}). \quad (3)$$

Здесь n — номер итерации, а к правой и левой частям соотношения (4) добавлены члены с одинаковыми коэффициентами, содержащие множитель θ_l , который является параметром итерационного процесса (рис. 2). Отметим, что случай $\theta_l = 0$ можно интерпретировать как использование граничного условия Дирихле при решении вспомогательной подзадачи в области Ω_s . Случай $\theta_l = 1$ соответствует граничному условию Неймана, а значение $\theta_l \in (0, 1)$ — граничному условию 3-го рода, или Робина.

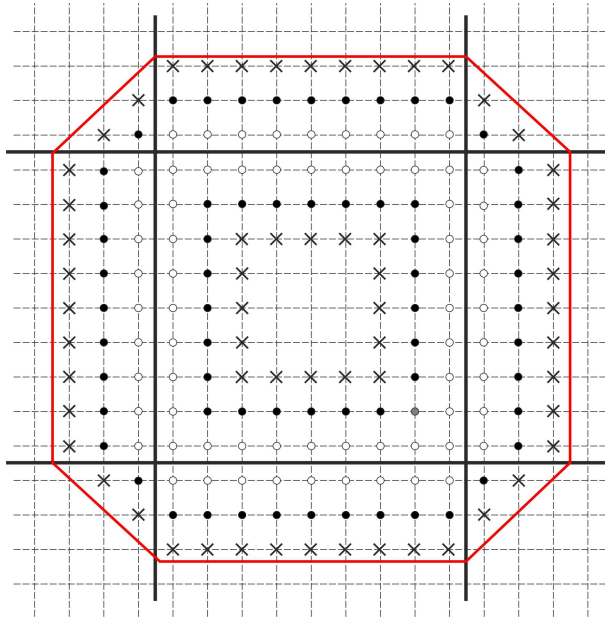


Рис. 1. Расширение сеточной подобласти для $\Delta = 3$

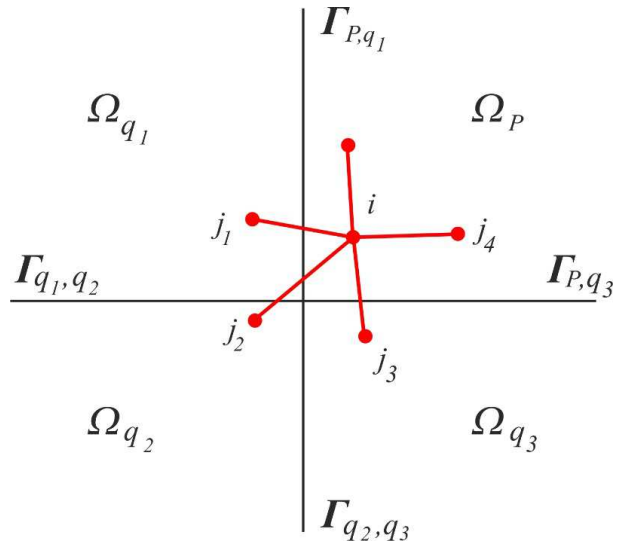


Рис. 2. Сеточный шаблон для околограничного узла

В матричной форме рассмотренный алгоритм можно представить как блочный метод Якоби

$$\bar{B}_s (\tilde{u}_s^{n+1} - \tilde{u}_s^n) = \tilde{f}_s^n - (\bar{A} \tilde{u}_s^n)_s \equiv \tilde{r}_s^n, \tag{4}$$

в котором подвекторы \tilde{u}_s^n и \tilde{f}_s^n относятся к расширенным подобластям и имеют размерности \bar{N}_s , а $\bar{B}_s \in \mathcal{R}^{\bar{N}_s, \bar{N}_s}$ — предобуславливающие матрицы, диагональные элементы которых зависят от параметров θ_l .

Итерационный процесс в форме (4) является недоопределенным, поскольку величины \tilde{u}_s^{n+1} имеют неоднозначные значения в пересечениях подобластей. Мы будем использовать “ограничительный” аддитивный метод Шварца (RAS — Restricted Additive Schwarz), в котором очередное итерационное приближение однозначно определяется в виде $u^{n+1} = \bigcup_s u_s^{n+1}$, где $u_s^{n+1} \in \Omega_s$ — совокупность значений подвектора \tilde{u}_s^{n+1} , определенного в расширенной подобласти $\bar{\Omega}_s$, но соответствующие узлы которого принадлежат Ω_s (для s -й подобласти можно определить оператор сужения $R_s : \bar{\Omega}_s \rightarrow \Omega_s$). Метод RAS можно записать в форме

$$u^{n+1} = u^n + B_{\text{ras}}^{-1} r^n = u^n + p^n, \quad r^{n+1} = r^n - AB_{\text{ras}}^{-1} r^n = r^n - Ap^n, \quad p^n = B_{\text{ras}}^{-1} r^n, \tag{5}$$

$$B_{\text{ras}}^{-1} = R\hat{A}^{-1}W^T, \quad \hat{A} = W^T A W = \text{block-diag} \left\{ A_s \in \mathcal{R}^{\bar{N}_s, \bar{N}_s} \right\},$$

где $W = [w_1, \dots, w_P] \in \mathcal{R}^{N, P}$ — прямоугольная матрица, каждый столбец w_s которой имеет единичные компоненты в узлах из $\bar{\Omega}_s$ и нулевые — в остальных. Отметим, что даже при симметричности исходной СЛАУ предобуславливающая матрица B_{ras} из (5) в общем случае симметричной не является. Обращение блоков A_s матрицы \hat{A} сводится фактически к решению независимых подсистем в соответствующих подобластях, что и является основой распараллеливания аддитивного метода Шварца, или блочного алгоритма Якоби.

Данный итерационный процесс можно записать в форме

$$u^{n+1} = T u^n + \bar{f}, \quad \bar{f} = B_{\text{ras}}^{-1} f, \tag{6}$$

где матрица перехода имеет вид $T = I - B_{\text{ras}}^{-1} A$. Очевидно, что при сходимости итераций $u^n \rightarrow u$ в пределе $n \rightarrow \infty$ мы из (6) получаем решение СЛАУ:

$$\bar{A} u \equiv (I - T) u = \bar{f}, \tag{7}$$

где \bar{A} — предобусловленная матрица.

Скорость сходимости описанного итерационного процесса зависит от количества подобластей, а точнее говоря — от диаметра графа, представляющего образуемую декомпозицией макросеть. Это можно

наглядно объяснить тем, что за одну итерацию возмущение решения в одной подобласти передается только соседним, или смежным, подобластям. Для повышения скорости сходимости итераций естественно использовать на каждом шаге не только ближние, но и дальние связи подобластей. С этой целью в алгоритмах декомпозиции используются различные ускоряющие подходы: методы дефляции, грубосеточной коррекции, агрегации и др. Эти подходы в той или иной степени перекликаются с многосеточными принципами, а также с малоранговыми аппроксимациями матриц, см. обширную цитируемую литературу на специальном сайте [5]. Общим в этих случаях является то, что каким-то специальным образом строится вспомогательная СЛАУ малого порядка (редуцированная система), решение которой в определенном смысле приближает исходное решение. Поэтому всем алгоритмам такого типа можно дать общее название — методы редукции.

Мы рассмотрим основанный на интерполяционном принципе следующий подход. Пусть Ω_c — некоторая грубая сетка с числом узлов $N_c \ll N$ в расчетной области Ω , причем узлы исходной и грубой сеток могут и не совпадать.

Обозначим через $\varphi_1, \dots, \varphi_{N_c}$ совокупность базисных интерполяционных (лагранжевого типа) многочленов некоторой степени M на сетке Ω_c , которые без ограничения общности будем предполагать финитными и образующими разложение единицы, т.е. $\sum_{k=1}^{N_c} \varphi_k(x, y) = 1$. Тогда вектор искомого решения СЛАУ (1) можно представить в виде разложения по данному базису:

$$u = \left\{ u_l \approx u_l^c = \sum_{k=1}^{N_c} c_k \varphi_k(x_l, y_l) \right\} = \Phi \hat{u} + \psi, \quad (8)$$

где $\hat{u} = \{c_k\} \in \mathcal{R}^{N_c}$ — вектор коэффициентов разложения по базисным функциям, ψ — погрешность аппроксимации, а $\Phi = [\varphi_1, \dots, \varphi_{N_c}] \in \mathcal{R}^{N, N_c}$ — прямоугольная матрица, каждый k -й столбец которой состоит из значений базисной функции $\varphi_k(x_i, y_j)$ в узлах исходной сетки Ω (в силу финитности базиса большинство элементов Φ — нулевые). Столбцы, или функции φ_k , можно считать ортонормированными, но это не обязательно. Если в k -м узле P_k грубой сетки Ω_c отлична от нуля только одна базисная функция ($\varphi_k(P_{k'}) = \delta_{k, k'}$), то $\hat{u}_k = c_k$ — точное значение искомого решения в точке P_k . После подстановки (8) в исходную СЛАУ (1) получаем систему $A\Phi\hat{u} = f - A\psi$, после умножения которой на Φ^T получаем (индекс “ T ” означает транспонирование)

$$\hat{A}\hat{u} \equiv \Phi^T A\Phi\hat{u} = \Phi^T f - \Phi^T A\psi \equiv \hat{f} \in \mathcal{R}^{N_c}. \quad (9)$$

Предполагая далее погрешность ψ в (8) достаточно малой, после ее отбрасывания получаем систему для приближенного “грубосеточного” решения \hat{u} :

$$\hat{A}\hat{u} = \Phi^T f \equiv \hat{f}. \quad (10)$$

Если матрица A невырожденная, а матрица Φ имеет полный ранг (намного меньший N), что мы и предполагаем в дальнейшем, то из (10) получим

$$u \approx \tilde{u} = \Phi\hat{u} = \Phi\hat{A}^{-1}\hat{f} = B_c^{-1}f, \quad B_c^{-1} = \Phi(\Phi^T A\Phi)^{-1}\Phi^T,$$

причем для погрешности полученного приближенного решения имеем $u - \tilde{u} = (A^{-1} - B_c^{-1})f$.

Ошибку приближенного решения можно выразить и через погрешность аппроксимации ψ . Вычитая почленно уравнения (9) и (10), имеем $\hat{A}(\hat{u} - \hat{u}) = -\Phi^T A\psi$. Отсюда далее получаем необходимое соотношение: $u - \tilde{u} = \Phi\hat{u} + \psi - \Phi\hat{u} = \psi - B_c^{-1}A\psi$.

Введенная выше матрица B_c^{-1} может рассматриваться как малоранговое приближение к матрице A^{-1} и использоваться в качестве предобусловливателя для построения итерационного процесса. В частности, для произвольного вектора u^{-1} мы можем выбирать начальное приближение по формуле

$$u^0 = u^{-1} + B_c^{-1}r^{-1}, \quad r^{-1} = f - Au^{-1}. \quad (11)$$

При этом соответствующая начальная невязка $r^0 = f - Au^0$ будет ортогональна “грубосеточному” подпространству $\Phi = \text{span}\{\varphi_1, \dots, \varphi_{N_c}\}$ в смысле выполнения соотношения

$$\Phi^T r^0 \equiv \Phi^T (r^{-1} - A\Phi\hat{A}^{-1}\Phi^T r^{-1}) = 0.$$

Приведенные соотношения в работе [7] положены в основу метода сопряженных градиентов с дефляцией, при этом начальный направляющий вектор выбирается по формуле

$$p^0 = (I - B_c^{-1}A)r^0 = Cr^0, \quad C = I - B_c^{-1}A, \quad (12)$$

что обеспечивает выполнение условия ортогональности $\Phi^T Ap^0 = 0$.

Поскольку рассматриваемые нами СЛАУ в общем случае несимметричны, мы будем использовать стабилизированный метод бисопряженных градиентов (BiCGStab — BiConjugate Gradient Stabilized method [11, 12]). Точнее говоря, мы применяем этот метод к решению предобусловленной системы (7). В нижеследующих формулах следует иметь в виду, что умножение на матрицу \bar{A} фактически означает реализацию итерационного шага (6), включающего в себя решение вспомогательных систем в подобластях:

$$\begin{aligned} r^0 &= \bar{f} - \bar{A}u^0, \quad p^0 = r^0, \quad n = 0, 1, \dots : \\ s^n &= r^n - \alpha_n \bar{A}p^n, \quad \alpha_n = \frac{(r^n, r^0)}{(\bar{A}p^n, r^0)}, \quad u^{n+1} = u^n + \alpha_n p^n + \omega_n s^n, \quad \omega_n = \frac{(\bar{A}s^n, s^n)}{(\bar{A}s^n, \bar{A}s^n)}, \\ r^{n+1} &= s^n - \omega_n \bar{A}s^n, \quad p^{n+1} = r^{n+1} + \beta_n (p^n - \omega_n \bar{A}p^n), \quad \beta_n = \alpha_n \frac{(r^{n+1}, r^0)}{(\omega_n (r^n, r^0))}. \end{aligned}$$

Отметим, что реализация данного алгоритма требует двух умножений на матрицу \bar{A} на каждой итерации.

3. Методы и технологии распараллеливания алгоритмов. Реализация параллельных методов декомпозиции областей в целом может быть представлена следующими этапами:

- 1) осуществление предварительных “доитерационных” операций, включающих в себя сбалансированную декомпозицию расчетной сеточной области; основные действия здесь выполняются на корневом процессоре, однако главным результатом является формирование вспомогательных СЛАУ для подобластей, распределенных по соответствующим MPI-процессорам;
- 2) выполнение внешнего итерационного процесса в подпространствах Крылова с организацией обменов между процессами;
- 3) параллельное (синхронное) решение алгебраических подсистем на многоядерных процессорах с общей памятью с использованием “внутреннего” распараллеливания многопоточковых вычислений средствами OpenMP.

Задачами экспериментальных исследований являются верификация, тестирование и сравнительный анализ эффективности различных алгоритмов решения больших разреженных СЛАУ с целью их оптимизации и включения в состав библиотеки параллельных алгебраических решателей KRYLOV [8]. При этом главные требования к разрабатываемому программному коду заключаются в высокой производительности и масштабируемом быстродействии, а также в отсутствии формальных ограничений на порядки решаемых СЛАУ и на количество используемых процессоров или вычислительных ядер. Отметим, что в соответствии с [4] можно различать сильную и слабую масштабируемость. Первое понятие характеризует уменьшение времени выполнения одной большой задачи с увеличением числа вычислительных устройств, а второе — примерное сохранение времени решения при одновременном росте размерности (числа степеней свободы) задачи и количества вычислителей.

Реализация алгоритмов осуществлялась с учетом типовой архитектуры кластера ССКЦ СО РАН [9], на котором функционирует библиотека KRYLOV, однако без использования графических ускорителей общего назначения GPGPU, поскольку их эффективное применение в рассматриваемых методах декомпозиции областей имеет свои технологические вычислительные сложности и требует специального изучения.

Организация вычислений производится следующим естественным образом: если расчетная область разбивается на P подобластей, то решение СЛАУ осуществляется на $P + 1$ процессоре CPU (один — “корневой”, а каждый из остальных соответствует “своей” подобласти), и при этом формируется такое же количество MPI-процессов. Решение вспомогательных алгебраических систем в подобластях получается одновременно на многоядерных CPU с применением многопоточковых вычислений средствами OpenMP.

Поскольку алгоритмы библиотеки KRYLOV ориентированы на решение больших разреженных СЛАУ, возникающих из аппроксимации многомерных краевых задач на неструктурированных сетках, то для хранения ненулевых элементов матриц используется общераспространенный разреженный строчный формат

CSR (Compressed Sparse Row). При этом в проведенных экспериментах на предварительном этапе формируется “глобальная” матрица A на корневом процессоре, а затем организуется распределенное хранение блочных строк \bar{A}_s матрицы A для s -х расширенных подобластей в памяти соответствующих процессоров. Необходимо иметь в виду, что при реализации “производственного” кода для решения сверхбольших СЛАУ требуется “распределенный” режим формирования глобальной матрицы в разных MPI-процессах на соответствующих узлах вычислительной системы без ограничительной перегрузки памяти узла корневого процесса.

Отметим, что в рассматриваемых двумерных сеточных краевых задачах применяется двумерная же декомпозиция сбалансированных подобластей, когда для примерно равного числа узлов $N_S \approx N/P$ в каждой из них диаметр d макросетки из подобластей приблизительно равен \sqrt{P} . Так как количество итераций аддитивного метода Шварца, даже с использованием методов Крылова, пропорционально величине d^γ , $\gamma > 0$, то это дает значительное преимущество перед одномерной декомпозицией, для которой $d \approx P$.

Масштабируемое распараллеливание алгоритмов обеспечивается синхронизацией вычислений по подобластям программными средствами системы MPI (Message Passing Interface) и минимизацией временных потерь на межпроцессорные обмены. Решение автономных СЛАУ в области $\bar{\Omega}_s$ осуществляется прямыми или итерационными методами, что требует на каждом шаге двухуровневого процесса применения $(N/P)^{\gamma_1}$, $\gamma_1 > 0$, объема вычислений. При этом необходим обмен данными, соответствующими только периферийным узлам смежных подобластей, однако объем такой информации гораздо меньше и пропорционален $(N/P)^{\gamma_1/2}$, что позволяет в принципе совмещать во времени арифметические действия и коммуникационные операции.

Высокая производительность реализованных вычислительных методов и технологий поддерживается активным использованием стандартных функций и векторно-матричных операций из систем BLAS и SPARSE BLAS, включенных в состав библиотеки MKL INTEL [10].

Следует отметить, что используемые в данной работе вычислительные технологии являются “полуметодическими”. С одной стороны, используется универсальный формат CSR для хранения СЛАУ, но в то же время мы ограничиваемся достаточно простыми двумерными краевыми задачами на прямоугольных сетках.

4. Примеры численных экспериментов. Приведем результаты методических расчетов по решению пятиточечных СЛАУ, полученных с помощью монотонных аппроксимаций экспоненциального типа [6], для решения задачи Дирихле (2) в квадрате на квадратной сетке с числом узлов $64^2, 128^2$ и 256^2 . Вычисления проводились на $P = 2^2, 4^2, 8^2$ процессорах, каждый из которых соответствовал одной из подобластей, образующих квадратную макросетку. Итерации по подобластям осуществлялись с помощью алгоритма BiCGStab [11] с использованием критерия останова $\|r^n\|_2 \leq \epsilon \|f\|_2$, $\epsilon = 10^{-8}$. Решение вспомогательных подсистем в подобластях осуществлялось с помощью прямого алгоритма, реализуемого в многопоточном режиме программой PARDISO [10]. При этом наиболее трудоемкая часть LU -разложения матрицы выполняется только один раз перед началом итераций.

В приводимых далее результатах численных экспериментов указывается только количество итераций, определяющих скорость сходимости итераций исследуемых методов. Поскольку в рассматриваемых двумерных задачах СЛАУ не являются сверхбольшими, времена расчетов здесь не являются характерными (хотя они примерно пропорциональны числу итераций).

В табл. 1 и 2 мы приводим результаты экспериментов соответственно для значений $p = q = 0$ и $p = q = 4$ в уравнении (2) по исследованию влияния методов редукции и параметра перехлеста Δ подобластей на скорость сходимости итераций. В расчетах использовались значения $\Delta = 0, 1, 2$ и при этом дополнительно применялось ускорение внешних итераций с помощью упрощенного варианта агрегации, или грубосеточной коррекции, когда корректируются только начальное приближение u^0 и начальный корректирующий вектор p^0 по формулам (11) и (12). При этом использовались интерполяционные приближения искомого решения вида (8) нулевого или первого порядка. В простейшем варианте значения базисных функций $\varphi_k(x_i, y_j)$, число которых равно $N_c = P_x \times P_y$, брались равными единице, если узел (x_i, y_j) принадлежит подобласти Ω_k , и нулю, если $(x_i, y_j) \notin \Omega_k$.

Во втором случае использовались $N_c = (P_x + 1)(P_y + 1)$ базисных кусочно-билинейных функций вида

$$\begin{aligned} \varphi_k(x, y) &= \hat{\psi}_I(x)\check{\psi}_J(y), \quad \hat{\psi}_I(x_{I'}) = \delta_{I,I'}, \quad \check{\psi}_J(y_{J'}) = \delta_{J,J'}, \\ I &= 0, 1, \dots, P_x, \quad J = 0, 1, \dots, P_y, \quad k = I + 1 + J(P_x + 1), \end{aligned} \tag{13}$$

определенных на квадратной макросетке из подобластей при значениях $P_x = P_y = \sqrt{P}$ ($\delta_{I,I'}$ и $\delta_{J,J'}$ — символы Кронекера), а линии $x = X_I$ и $y = Y_J$ выбирались проходящими через вершины подобластей. В формуле (8) величина $u(x_I, y_J)$ аппроксимировалась с помощью четырех “своих” базисных функций

$\varphi_k(x, y)$ по условиям принадлежности ячейке макросетки $X_I \leq x_l \leq X_{I+1}$, $Y_J \leq y_l \leq Y_{J+1}$. Соответственно в каждой строке матрицы Φ содержится только 4 ненулевых элемента, а столбцы этой матрицы отличны от нуля только в позициях, соответствующих номерам l узлов сетки, принадлежащих “своей” подобласти.

Результаты табл. 1 и 2 получены при использовании коррекции нулевого начального приближения для точного решения $u = x^2 - y^2$. Значения параметров θ_l из формулы (3) везде брались равными нулю, т.е. интерфейсные граничные условия были 1-го рода. В каждой клетке этих таблиц верхняя строка относится к расчетам без агрегации, средняя — к применению агрегации при помощи кусочно-постоянной интерполяции, нижняя — к применению агрегации с кусочно-билинейными базисными функциями вида (13). Три числа в каждой строке соотносятся со значениями перехлеста подобластей $\Delta = 0, 1, 2$ (левая, средняя и правая цифры в клетках табл. 1 и 2). Данные в табл. 1 и 2 приведены для значений конвективных членов уравнения (2), равных 0 и 4 соответственно.

Таблица 1
Количество итераций BiCGStab
для $\Delta = 0, 1, 2$, $p = q = 0$

$N \setminus P$	2^2			4^2			8^2		
64^2	19	11	8	26	15	12	37	20	15
	20	11	8	27	15	11	31	18	13
	16	9	7	21	12	9	27	15	11
128^2	27	15	11	34	22	17	51	31	21
	29	18	13	39	21	13	48	25	21
	22	14	10	25	16	12	32	21	15
256^2	37	21	17	54	31	23	72	43	32
	41	23	19	52	29	22	59	42	28
	33	18	15	35	21	17	41	26	21

Таблица 2
Количество итераций BiCGStab
для $\Delta = 0, 1, 2$, $p = q = 4$

$N \setminus P$	2^2			4^2			8^2		
64^2	19	10	8	22	13	11	33	16	13
	18	11	8	23	14	10	29	16	12
	16	9	8	20	12	9	28	16	13
128^2	26	15	12	37	21	16	41	25	18
	24	15	13	31	19	15	41	24	17
	26	13	10	30	16	13	41	22	17
256^2	36	21	17	55	26	20	55	37	29
	34	20	16	45	28	21	58	35	26
	34	19	14	53	23	19	47	31	23

Как видно из приведенных результатов, количества итераций для всех сеток естественным образом растут с увеличением числа подобластей P и убывают с ростом Δ . Применение агрегации несколько ускоряет итерационный процесс, но для усиления этого эффекта требуются дополнительные исследования.

В табл. 3 мы приводим количество итераций при различных θ из уравнения (3), определяющих интерфейсные граничные условия для смежных подобластей (в каждой клетке левое и правое числа соответствуют сеткам 128^2 и 256^2), а значения параметров брались постоянными для всех узлов, т.е. $\theta_l = \theta$. Рассматривался случай подобластей без пересечения, т.е. $\Delta = 0$, а также без применения ускоряющих методов редукции.

Таблица 3
Количество итераций на сетках 128^2 и 256^2
при различных параметрах θ

$P \setminus \theta$	0		0.5		0.6		0.7		0.9975	
4	18	27	16	26	16	24	14	23	10	12
16	32	41	28	40	27	39	27	40	31	75
64	43	60	42	56	40	55	41	55	93	86

Как видно из этих результатов, в приведенных расчетах для всех сеток и разного количества подобластей обнаруживается оптимальное значение θ , близкое к единице, однако выигрыш оказывается только в пределах 10–40%. Данные вычисления проводились без пересечения подобластей, а при $\Delta \geq 1$ наилучшая величина θ оказывается нулевой, что соответствует условиям Дирихле на смежных границах.

5. Заключение. Результаты данной работы носят предварительный характер в том смысле, что оптимизация математической эффективности исследуемых алгебраических методов, а также производительности их программных реализаций требует дополнительных исследований. Одним из перспективных направлений является использование мультипредобусловливания [3] во внешнем итерационном процессе, что в нашем случае означает оптимизацию совместного использования блочного метода Якоби (по подобластям) и алгоритмов редукции. Недостаточно также изучен и в данной статье не освещается вопрос о динамическом регулировании критерия останова итерационных процессов при их использовании для решения вспомогательных задач в подобластях. Кроме того, серьезной технологической проблемой остается эффективное применение быстродействующих графических ускорителей общего назначения на типовых архитектурах существующих вычислительных систем.

Работа поддержана грантом Российского научного фонда № 14-11-00485, экспериментальная часть поддержана грантом РФФИ № 14-07-00128.

Статья рекомендована к публикации Программным комитетом Международной научной конференции “Параллельные вычислительные технологии 2015” (<http://agora.guru.ru/pavt2015/>).

СПИСОК ЛИТЕРАТУРЫ

1. Ильин В.П. Параллельные методы и технологии декомпозиции областей // Вестник ЮУрГУ. Серия “Вычислительная математика и информатика”. 2012. № 46. 31–44.
2. Toselli A., Widlund O.B. Domain decomposition methods — algorithms and theory. Heidelberg: Springer, 2005.
3. Bridson R., Greif C. A multipreconditioned conjugate gradient algorithm // SIAM J. Matrix Anal. Appl. 2006. **27**, N 4. 1056–1068.
4. Dubois O., Gander M.J., Loisel S., St-Cyr A., Szyld D.B. The optimized Schwarz method with a coarse grid correction // SIAM Journal on Scientific Computing. 2012. **34**, N 1. 421–458.
5. URL: <http://www.ddm.org>
6. Ильин В.П. Методы конечных разностей и конечных объемов для эллиптических уравнений. Новосибирск: Изд-во ИВМиМГ СО РАН, 2000.
7. Chapman A., Saad Y. Deflated and augmented Krylov subspace techniques // Numer. Linear Algebra Appl. 1997. **4**, N 1. 43–66.
8. Бутюгин Д.С., Гурьева Я.Л., Ильин В.П., Перевозкин Д.В., Петухов А.В., Скопин И.Н. Функциональность и технологии алгебраических решателей в библиотеке Krylov // Вестник ЮУрГУ. Серия “Вычислительная математика и информатика”. 2013. **2**, № 3. 92–105.
9. URL: <http://www2.sccc.ru>
10. URL: <https://software.intel.com/en-us/intel-mkl>
11. Saad Y. Iterative methods for sparse linear systems. Philadelphia: SIAM Press, 2003.
12. Ильин В.П. Методы и технологии конечных элементов. Новосибирск: Изд-во ИВМиМГ СО РАН, 2007.

Поступила в редакцию
18.02.2015

On Acceleration Technologies of Parallel Decomposition Methods

Y. L. Gurieva¹ and V. P. Il'in²

¹ *Institute of Computational Mathematics and Mathematical Geophysics, Siberian Branch of Russian Academy of Sciences; prospekt Lavrentyeva 6, Novosibirsk, 630090, Russia; Ph.D., Senior Scientist, e-mail: yana@lapasrv.sccc.ru*

² *Institute of Computational Mathematics and Mathematical Geophysics, Siberian Branch of Russian Academy of Sciences; prospekt Lavrentyeva 6, Novosibirsk, 630090, Russia; Dr. Sci., Professor, Principal Scientist, e-mail: ilin@sccc.ru*

Received February 18, 2015

Abstract: One of the main obstacles to the scalable parallelization of the algebraic decomposition methods for solving large sparse systems of linear algebraic equations consists in slowing the convergence rate of the additive iterative Schwarz algorithm in the Krylov subspaces when the number of subdomains increases. The aim of this paper is a comparative experimental analysis of various ways to accelerate the iterations: a parametrized intersection of subdomains, the usage of interface conditions at the boundaries of adjacent subdomains, and the application of a coarse grid correction (aggregation, or reduction) for the original linear system to build an additional preconditioner. The parallelization of algorithms is performed on two levels by programming tools for the distributed and shared memory. The benchmark linear systems under study are formed using the finite difference approximations of the Dirichlet problem for the diffusion–convection equation with various values of the convection coefficients and on a sequence of condensing grids.

Keywords: domain decomposition, additive Schwarz method, reduction algorithms, preconditioned Krylov processes, scalable parallelization, distributed and shared memory, numerical experiments.

References

1. V. P. Il'in, "Parallel Methods and Technologies of Domain Decomposition," Vestn. South Ural Univ. No. 46, 31–44 (2012).
2. A. Toselli and O. B. Widlund, *Domain Decomposition Methods – Algorithms and Theory* (Springer, Heidelberg, 2005).
3. R. Bridson and C. Greif, "A Multipreconditioned Conjugate Gradient Algorithm," SIAM J. Matrix Anal. Appl. **27** (4), 1056–1068 (2006).
4. O. Dubois, M. J. Gander, S. Loisel, et al., "The Optimized Schwarz Method with a Coarse Grid Correction," SIAM J. Sci. Comput. **34** (1), 421–458 (2012).
5. Domain Decomposition Methods. <http://www.ddm.org>. Cited February 15, 2015.
6. V. P. Il'in, *Methods of Finite Differences and Finite Volumes for Elliptic Equations* (Inst. Comput. Math. Math. Geophys., Novosibirsk, 2000) [in Russian].
7. A. Chapman and Y. Saad, "Deflated and Augmented Krylov Subspace Techniques," Numer. Linear Algebra Appl. **4** (1), 43–66 (1997).
8. D. S. Butyugin, Ya. L. Guryeva, V. P. Il'in, et al., "Parallel Algebraic Solvers Library Krylov," Vestn. South Ural Univ. **2** (3), 92–105 (2013).
9. Siberian Supercomputing Center. <http://www2.sccc.ru>. Cited February 15, 2015.
10. Intel Math Kernel Library. <https://software.intel.com/en-us/intel-mkl>. Cited February 15, 2015.
11. Y. Saad, *Iterative Methods for Sparse Linear Systems* (SIAM Press, Philadelphia, 2003).
12. V. P. Il'in, *Methods and Technologies of Finite Elements* (Inst. Comput. Math. Math. Geophys., Novosibirsk, 2007) [in Russian].