



Жуков В.Т., Новикова Н.Д.,
Феодоритова О.Б.

Сдвиговая стратегия в
обобщенном методе
минимальных невязок

Рекомендуемая форма библиографической ссылки: Жуков В.Т., Новикова Н.Д., Феодоритова О.Б. Сдвиговая стратегия в обобщенном методе минимальных невязок // Препринты ИПМ им. М.В.Келдыша. 2009. № 71. 28 с. URL: <http://library.keldysh.ru/preprint.asp?id=2009-71>

УЧРЕЖДЕНИЕ РОССИЙСКОЙ АКАДЕМИИ НАУК
ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ
им.М.В.Келдыша РАН

В.Т. Жуков, Н.Д. Новикова, О.Б. Феодоритова

СДВИГОВАЯ СТРАТЕГИЯ
В ОБОБЩЕННОМ МЕТОДЕ
МИНИМАЛЬНЫХ НЕВЯЗОК

Москва, 2009

УДК: 519.6

Сдвиговая стратегия в обобщенном методе минимальных невязок

В.Т. Жуков, Н.Д. Новикова, О.Б. Феодоритова

Email: zhukov@kiam.ru, nn@kiam.ru, feodor@kiam.ru

Препринт Института прикладной математики им. М.В.Келдыша РАН.

Предложен подход к решению больших разреженных систем линейных алгебраических уравнений. Он является вариантом популярного обобщенного метода минимальных невязок GMRES и использует чередующиеся этапы коррекции и явных рестартов. Базисы для коррекции формируются с помощью части базисных векторов построенного подпространства Крылова. Правила выбора этой части и есть исследуемая стратегия сдвигов. Приведены численные эксперименты для линейных систем, возникающих в результате стабилизированных конечно-элементных аппроксимаций задач конвекции-диффузии на структурных и неструктурных сетках.

Работа выполнена при частичной финансовой поддержке Российского Фонда Фундаментальных Исследований (код проекта 08-01-00299)

Стр. 28, рис. 19, библиогр. назв. 10

A shift strategy for the generalized method of minimal residuals

O.B. Feodoritova, N.D. Novikova, V.T. Zhukov

Preprint of Keldysh Institute of Applied Mathematics RAS.

An approach to solution of large sparse linear systems of equations is proposed. This approach is a variant of popular generalized method of minimal residuals GMRES and based on alternate correction stages and explicit restarts. The subspace for correction is constructed partly from basis vectors of Krylov subspace. A rule for selection of the desired vectors is an investigated shift strategy. The numerical experiments for linear systems, that appear as a result of stabilized finite-element approximations of convection-diffusion problems on structured and unstructured grids, are represented.

The work is supported by Russian Foundation for Basic Research (project 08-01-00299).

Содержание

1	Введение	3
2	Постановка задачи	4
2.1	Циклический вариант GMRES(m) с явным рестартом	5
2.2	GMRES_DR(m, k) с неявным рестартом	6
2.3	Комбинированный вариант GMRES_Proj с явным рестартом	10
3	Сравнение вариантов. Поиск усовершенствований	11
4	Численные эксперименты	15
4.1	Уравнение Пуассона	16
4.2	Уравнение конвекции-диффузии.	22
5	Заключение	27
	Список литературы	28

1 Введение

Многие задачи науки и техники описываются системами дифференциальных уравнений, в результате аппроксимаций которых возникают системы линейных алгебраических уравнений $Au = b$, где $A \in \mathbb{R}^{n \times n}$ – разреженная и, как правило, несимметричная матрица. Современные практические задачи апеллируют к порядкам $n \simeq 10^7$ и выше для размера этих систем. В области развития методов решения таких систем в последние годы наблюдается заметная активность. По-прежнему, как и десять лет назад, методы подпространств Крылова являются одними из наиболее эффективных инструментов. С их помощью решения линейных систем получаются в ходе итерационного процесса, на каждом этапе которого строятся подпространства меньшей размерности. Однако, и к крыловским методам сейчас предъявляют серьезные претензии вычислители-практики, что приводит к появлению новых направлений развития этой группы методов [1] – [3] .

В [4] – [5] мы имели дело с решением линейных систем, возникающих при решении краевых задач для стационарных уравнений Эйлера и Навье-Стокса. Обычно решение таких задач ищется установлением по времени, и системы линейных уравнений возникают в процессе линеаризации нелинейных разностных уравнений на каждом временном слое. В [5] построен и

исследован многосеточный метод, который может быть использован и как "решатель", и как предобуславливатель для итерационного метода обобщенных минимальных невязок, известного по аббревиатуре как GMRES [6]. Предложенный в [5] многосеточный метод существенно использует геометрическую, сеточную информацию о задаче, как и классический многосеточный метод Р.П. Федоренко. Альтернативный в этом смысле алгебраический многосеточный метод работает только с матрицей системы линейных уравнений. Но этот метод работоспособен для дискретизаций скалярных дифференциальных уравнений, а для систем уравнений он практически не применяется. Предложенные в этой работе модификации GMRES используют идею многосеточного метода коррекции решения в подпространствах меньшей размерности.

Желание опереться исключительно на алгебраическую структуру сеточной задачи привело нас к попыткам улучшить сходимость вышеупомянутого метода GMRES, который является одним из наиболее эффективных представителей группы итерационных методов подпространств Крылова и входит практически во все большие пакеты программ, PETSc, AZTEK и другие. Широко используется алгоритм GMRES(m), осуществляющий перезапуск (рестарт) GMRES через каждые m шагов с затратами $O(mn)$ слов оперативной памяти. Его использование связано с естественным ограничением на память, имеющим место даже для суперкомпьютеров, когда речь заходит о системах очень больших размеров. Кроме того, с вычислительной точки зрения критичной является процедура ортогонализации большого числа векторов, особенно при использовании параллельных компьютеров.

Метод GMRES(m) с рестартами имеет недостатки, например, может, приводить к стагнации итерационного процесса, просто в силу того, что рассчитываемые подпространства Крылова небольшой размерности $m \ll n$ не могут захватить векторы, лежащие в подпространствах, отвечающих минимальным по модулю собственным значениям исходного разностного оператора.

Есть несколько рецептов, как улучшить GMRES(m). В этой работе обсуждаются некоторые из них, приводятся результаты сравнения различных подходов, и строится алгоритм, основанный на выделении подпространств небольшой размерности $k < m$, улучшающий сходимость GMRES(m).

2 Постановка задачи

Рассмотрим систему линейных уравнений

$$Au = b,$$

где $A \in \mathbb{R}^{n \times n}$ – заданная невырожденная матрица, $b \in \mathbb{R}^n$ – заданный вектор правой части, $u \in \mathbb{R}^n$ – вектор неизвестных. Матрица A является разреженной, и как правило, несимметричной матрицей большого размера.

Опустим пока вопрос о предобуславливании и будем предполагать, что начальное приближение u_0 всегда нулевое.

Полный вариант GMRES для больших задач является практически неприемлемым в первую очередь из-за проблем с памятью компьютера и вычислительной неустойчивостью процедуры ортогонализации большого числа базисных векторов. Переход к алгоритму GMRES(m) с явным рестартом по тем же причинам не может быть использован с большими значениями m , а выбор небольших, практически приемлемых значений m может приводить к стагнации итерационного процесса.

Вариант, который мы предлагаем, использует фильтрацию и явные рестарты. Но сначала рассмотрим варианты метода минимальных невязок, с которыми в дальнейшем мы будем проводить сравнение.

Введем некоторые обозначения. Скалярное произведение векторов x и y в пространстве $\mathbb{R}^n$ над полем вещественных чисел будем обозначать как (x, y) или $x^T y$, где

$$x^T y = \sum_{i=1}^n x_i y_i,$$

а норму вектора x как $\|x\| = \sqrt{(x, x)}$. Линейную оболочку набора векторов $Y = \{y_1, \dots, y_m\}$ обозначаем как $\text{span}\{Y\}$. Условие ортонормированности двух базисов X и Y подпространства $\mathbb{R}^m$ записываем как $X^T Y = I$, где I – единичная матрица из $\mathbb{R}^{m \times m}$.

2.1 Циклический вариант GMRES(m) с явным рестартом

В алгоритме GMRES(m) строится подпространство Крылова

$$K_m(A, r_0) = \text{span}\{r_0, Ar_0, \dots, A^{m-1}r_0\},$$

r_0 – начальная невязка, вычисленная по начальному приближению u_0 :

$$r_0 = b - Au_0.$$

Приближенное решение u_m , минимизирующее норму невязки $\|b - Au\|$ в $K_m(A, r_0)$, можно записать в виде

$$u_m = u_0 + q_{m-1}(A)r_0,$$

где q_{m-1} — многочлен степени не выше $(m-1)$. Невязка r_m этого приближения есть результат применения многочлена $p_m(A)$ к r_0 :

$$r_m = b - Au_m = r_0 - Aq_{m-1}(A)r_0 = p_m(A)r_0,$$

($p_m(0) = 1$). Отметим, что невязка r_m ортогональна подпространству $\text{span}\{AK_m\}$. Новое итерационное приближение вычисляется как

$$u_m = u_0 + U_m z_m,$$

$z_m \in \mathbb{R}^m$, U_m — ортонормированный базис в $K_m(A, r_0)$, полученный процедурой Арнольди [7].

Переход от u_0 к u_m составляет один шаг GMRES(m), состоящий из m внутренних итераций. Вычислительные затраты такого шага грубо можно оценить в m матрично-векторных произведений. Затем шаги GMRES(m) циклически повторяются — это внешние итерации — строится новое подпространство Крылова $K_m(A, r_m)$, начиная с невязки $r_m = b - Au_m$, и так до достижения заданной точности. При таком подходе на некотором шаге может возникнуть хорошо известный эффект стагнации внешних итераций, и существует несколько рецептов его преодоления.

2.2 GMRES_DR(m, k) с неявным рестартом

В [2] предложен алгоритм GMRES_DR(m, k), в котором на каждом внутреннем цикле на основе построенного крыловского подпространства находятся так называемые гармонические числа и векторы Ритца. Для симметричной матрицы гармонические векторы Ритца максимизируют отношение Рэлея обратной матрицы. Поскольку A несимметрична, то ее собственные значения могут быть плохо обусловлены. Тем не менее получаемая спектральная информация оказывается полезной. Среди гармонических векторов Ритца выбирается подмножество из k векторов, как правило, отвечающих наименьшим по модулю гармоническим числам Ритца. Тогда можно надеяться, что выбранные векторы будут приближать базисы подпространств, отвечающих наименьшим по модулю собственными значениями исходной матрицы. Конечно, если известны причины, по которым должно быть выбрано иное подмножество векторов, это легко может быть сделано. При следующем неявном рестарте выбранные k векторов Ритца используются для формирования нового базиса в подпространстве Крылова.

Пунктирно приведем здесь алгоритм GMRES_DR, который отличается от предложенного в [2] в основном тем, что использует вместо собственного разложения более устойчивую для несимметричных матриц процедуру разложения Шура. Описание алгоритма позволит нам одновременно ввести необходимые для дальнейшего обозначения.

1. Выбираем параметры процесса:

m – максимальный размер подпространства Крылова $K_m = K_m(A, r_0)$;
 k – размер подпространства, которое интерпретируется как приближение подпространства, отвечающего малым по модулю собственным значениям матрицы A ;
 $r_0 = b - A u_0$ – начальный вектор невязки, u_0 – заданное начальное приближение к решению u .

2. Используя начальный вектор r_0 и процедуру Арнольди, строим последовательность U_{m+1} из $m + 1$ векторов, $U_{m+1}^T U_{m+1} = I$. Первые m векторов из них образуют ортонормированный базис U_m m -мерного подпространства Крылова K_m . Дополнительно рассчитывается верхняя хессенбергова $(m + 1) \times m$ матрица $\bar{H}_m = U_{m+1}^T A U_m$. Через H_m обозначим матрицу, полученную из матрицы $\bar{H}_m$ удалением последней строки. Матрица H_m является проекцией матрицы A на подпространство Крылова K_m : $H_m = U_m^T A U_m$.

Новое приближение к решению $u = u_0 + U_m d$ находится в результате решения задачи минимизации нормы $\|\beta e_1 - \bar{H}_m d\|$ для d , где e_1 – первый координатный вектор, а $\beta = \|r_0\|$.

Если требуемая точность достигнута, т.е. норма невязки $r = b - A u$ мала, $\|r\|_2 < \delta \|r_0\|_2$, то итерационный процесс завершается.

Пункты 1 и 2 полностью совпадают со стандартным вариантом GMRES(m).

Если точность не достигнута, то переходим к процессу фильтрации.

3. Рассчитываем матрицу $B = H_m + h_{m+1,m}^2 H_m^{-T} e_m e_m^T$ и берем к ней обратную $C = B^{-1}$. Здесь $e_m \in \mathbb{R}^m$ – m -й координатный вектор, $h_{m+1,m}$ – соответствующий элемент матрицы $\bar{H}_m$. Ниже приведено обоснование такого выбора матрицы B . Собственные числа матрицы B называются гармоническими числами Ритца и используются в качестве приближения к соответствующим собственным числам матрицы A .

4. К матрице C применяем QR -алгоритм для получения разложения Шура $C = Q D Q^T$, где Q – вещественная ортогональная матрица, $Q^T Q =$

I, D – верхняя блочно-треугольная $(m \times m)$ матрица с диагональными блоками размеров 1×1 и 2×2 . Собственными значениями матрицы являются собственные значения ее диагональных блоков. Беря обратные к ним значения, получаем гармонические числа Ритца исходной матрицы A . Разложение Шура выполняем с помощью процедуры DGEES пакета LAPACK .

5. Упорядочиваем векторы Шура по возрастанию собственных значений, выбираем из них k первых векторов $p_1, \dots, p_k$, и формируем $m \times k$ матрицу P_k . Заметим, что число k может быть изменено для того, чтобы комплексно-сопряженная пара собственных значений не была разделена.
6. Дополняем матрицу P_k нулевой строкой, оставляя за ее столбцами обозначение $p_1, \dots, p_k$, только теперь векторы p_k имеют длину $m + 1$. Полученный в п.2 вектор невязки $\hat{r} = \beta e_1 - \bar{H}_m d$ длины $m + 1$ ортонормализуем относительно векторов $p_1, \dots, p_k$, получая вектор p_{k+1} и $(m + 1) \times (k + 1)$ матрицу P_{k+1} .
7. Формируем для очередного рестарта обновленные части матрицы H_m и базиса U_m , полагая

$$U_{k+1}^{new} = U_{m+1} P_{k+1}, \quad \bar{H}_k^{new} = P_{k+1}^T \bar{H}_m P_k.$$

После этой процедуры оказывается полезно провести ортонормализацию последнего вектора базиса U_{k+1}^{new} относительно его предыдущих векторов.

8. Переходим к п. 2 построения нового m -мерного базиса, используя в качестве начальных данных для рестарта не вектор невязки, а построенные в предыдущем пункте векторы U_{k+1}^{new} и матрицу $\bar{H}_k^{new}$.

Приведенная процедура итерационно повторяется до сходимости или до исчерпания заданного числа рестартов.

Замечание 1. Оригинальный алгоритм GMRES_DR работает с матрицей B , а не с обратной к ней. Работа с обратной матрицей позволяет более устойчиво рассчитывать подпространства, соответствующие именно минимальным собственным значениям исходного оператора.

Как уже отмечалось, числа, обратные к максимальным по модулю собственным значениям матрицы

$$B = H_m + h_{m+1,m}^2 H_m^{-T} e_m e_m^T,$$

приближают минимальные по модулю собственные значения исходной матрицы A (что нам и нужно). Разберемся в смысле этого высказывания. Рассмотрим задачу $A^{-1}y = \lambda y$ на подпространстве $W = AK_m$. Галеркинская формулировка этой задачи имеет вид: найти вектор $y \in W$, и комплексное число $\lambda \in \mathbb{C}$, такие, что

$$(A^{-1}y - \lambda y, w) = 0, \quad \forall w \in W. \quad (1)$$

Общий вид векторов w и y известен:

$$w = AU_m z_2, \quad y = AU_m z_1, \quad z_1, z_2 \in \mathbb{R}^m.$$

После подстановки w и y в (1) имеем

$$(A^{-1}AQ_m z_1 - \lambda AQ_m z_1, AQ_m z_2) = 0.$$

Нетрудно получить следующие равенства:

$$(Q_m^T A^T (Q_m z_1 - \lambda AQ_m z_1), z_2) = 0,$$

$$((Q_m^T A^T Q_m - \lambda Q_m^T A^T A Q_m) z_1, z_2) = 0,$$

$$((H_m^T - \lambda Q_m^T A^T Q_{m+1} \bar{H}_m) z_1, z_2) = 0,$$

$$((H_m^T - \lambda \bar{H}_m^T \bar{H}_m) z_1, z_2) = 0.$$

Так как это выполнено для произвольных векторов z_1, z_2 из $\mathbb{R}^m$, то поставленная на подпространстве W задача (1) сведена к задаче на нахождения собственных значений λ и отвечающих им собственных векторов g обобщенной задачи

$$H_m^T g = \lambda \bar{H}_m^T \bar{H}_m g, \quad g \in \mathbb{R}^m. \quad (2)$$

Так как

$$\bar{H}_m^T \bar{H}_m = [H_m^T; \beta e_m] \begin{bmatrix} H_m \\ \dots \\ \beta e_m^T \end{bmatrix} = H_m^T H_m + \beta^2 e_m e_m^T,$$

где

$$e_m = (0, \dots, 0, 1) \in \mathbb{R}^m, \quad \beta = h_{m+1, m},$$

то вместо (2) имеем

$$H_m^T g = \lambda (H_m^T H_m + \beta^2 e_m e_m^T) g,$$

или

$$g = \lambda(H_m + \beta^2 H_m^{-T} e_m e_m^T)g.$$

Итак, мы получили задачу

$$g = \lambda B g,$$

которую запишем в привычном виде

$$B^{-1}g = \lambda g. \quad (3)$$

Решая задачу (3) с помощью QR-алгоритма, наиболее надежно мы можем найти инвариантное подпространство, отвечающее группе максимальных по модулю собственных значений матрицы B^{-1} . А значит, обратные к ним значения $\beta = 1/\lambda$ приближают минимальные по модулю собственные значения исходной матрицы A . Собственная пара (β, g) с вектором $g \in \mathbb{R}^m$ соответствует паре (β, y) с вектором $y \in \mathbb{R}^n$, где $y = U_m g$, представляющей гармонические числа и векторы Ритца [2].

Вне рамок данной работы остается проверка утверждения о надежности расчета головной части спектра. Такая проверка может быть основана на вычислении количественных характеристик дихотомии спектра матрицы B^{-1} .

2.3 Комбинированный вариант GMRES_Proj с явным рестартом

Рассмотрим вариант с явными рестартами – алгоритм GMRES_Proj, который состоит из двух фаз – стандартный цикл GMRES(m) чередуется с фазой проектирования, на которой находится поправка к решению в подпространстве небольшой размерности $k < m$. На начальной стадии алгоритм GMRES_Proj совпадает с алгоритмом GMRES_DR. В результате рассчитываются k векторов U_k , линейная оболочка которых является приближением подпространства, соответствующего k наименьшим по модулю собственным значениям исходной матрицы.

На втором этапе мы добавляем к U_k вектор невязки r , и в подпространстве $\text{span}\{U_{k+1}\} \equiv \text{span}\{U_k, r\}$ полученное приближенное решение u корректируется, т.е. находится поправка к решению, лежащая в $\text{span}\{U_{k+1}\}$. После этого выполняется следующий стандартный цикл GMRES(m) и так далее, до сходимости или исчерпания числа повторений. Этап коррекции выглядит следующим образом:

1. Текущую невязку r_m ортонормализуем по отношению к построенному подмножеству векторов U_k , т.е. присоединяем к этому набору вектор

u_{k+1} , полученный в процессе ортогонализации и нормировки:

$$u_{k+1} = r_m - \sum_{j=1}^k h_{j,k} u_j, \quad h_{j,k} = u_j^T r_m, \quad j = 1, \dots, k,$$

$$u_{k+1} := u_{k+1} / \|u_{k+1}\|.$$

Базис $U = U_{k+1}$ является ортонормированным, $U^T U = I_{k+1}$.

- Используем для уточнения решения подпространство $\text{span}\{U\}$. Нам известны текущее итерационное приближенное u_0 и соответствующая невязка r_0 . Найдем поправку $\delta u = u^{new} - u_0$, решая систему $A\delta u = r_0$ на подпространстве $\text{span}\{U\}$, то есть, решая систему линейных уравнений

$$U^T A U d = U^T r_0$$

размера $k < m$ относительно вектора коэффициентов $d \in \mathbb{R}^k$.

Новое уточненное решение есть $u^{new} = u_0 + U d$, а новая невязка $r^{new} = b - A u^{new}$ соответственно.

- Этот новый вектор r^{new} используется для начала нового итерационного цикла GMRES(m): $r_0 = r^{new}$.

3 Сравнение вариантов. Поиск усовершенствований

Сравним поведение описанных вариантов на задаче конвекции-диффузии.

Задача с разрывными краевыми условиями [8],[9]. Рассмотрим уравнение конвекции - диффузии

$$-\varepsilon(u_{xx} + u_{yy}) + p_1 u_x + p_2 u_y = f, \quad \varepsilon = 10^{-6}, \quad p_1 = p_2 = \frac{\sqrt{2}}{2}, \quad f = 0$$

в единичном квадрате (см. рис. 1 а).

Поле скоростей p_1, p_2 однородно и направлено под углом 45° градусов к оси ОХ. Разрыв переносится внутрь области характеристикой $y = x$ и вдоль нее в результате диффузии образуется тонкий пограничный слой.

Для аппроксимации используем метод стабилизированных конечных элементов [10], реализованный на треугольной сетке, которая получена из прямоугольной сетки делением каждой ячейки диагональю на два треугольника (см. рис. 1 б).

Сравним сходимость представленных выше вариантов GMRES для данной задачи, см. рис. 2. Расчетная сетка состоит из 16641 узлов, что соот-

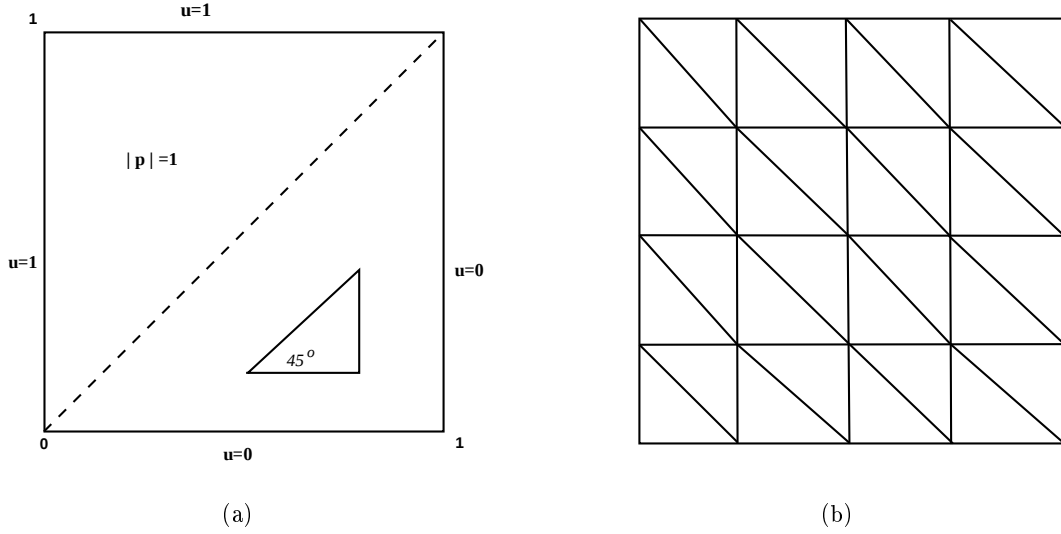


Рис. 1: Задача конвекции-диффузии с разрывными краевыми условиями: (a) постановка задачи; (b) фрагмент сетки.

ветствует разбиению квадратной сетки на 128×128 ячеек.

Исучаемая задача – это задача с доминирующей конвекцией и полученная после аппроксимации матрица A не является симметричной, ее спектр комплексный. Наихудшую сходимость из представленных вариантов демонстрирует GMRES_DR, а лучше всего себя показывает GMRES_Proj. Для понимания ситуации на комплексной плоскости нарисуем гармонические числа Ритца, получаемые в ходе итерационного процесса. На рис. 3– 5 показано распределение гармонических чисел Ритца для варианта с рестартом GMRES(30) (рис. 3), варианта с неявным рестартом GMRES_DR(30,10) (рис. 4) и GMRES_Proj (рис. 5). Легко видеть, что GMRES_DR лучше других алгоритмов аппроксимирует "голову" спектра (набор минимальных по модулю собственных значений), однако усилия эти являются чрезмерными. Оптимальным оказывается вариант GMRES_Proj, в котором гармонические числа Ритца концентрируются около некоторых точек, равномерно расставленных вдоль контура, образованного на комплексной плоскости числами Ритца классического GMRES.

Представленная картина распределения собственных чисел типична для задач конвекции-диффузии с доминирующей конвекцией, а значит и сравнительные характеристики сходимости аналогичны.

Мы предлагаем рассмотреть группу методов, содержащих этап коррекции, представителем которых является рассмотренный выше вариант GMRES_Proj. Для этого расширим вариант GMRES_Proj, оснастив его сдвигами. Выбор корректирующего подпространства $\text{span}(U_{k+1})$ является

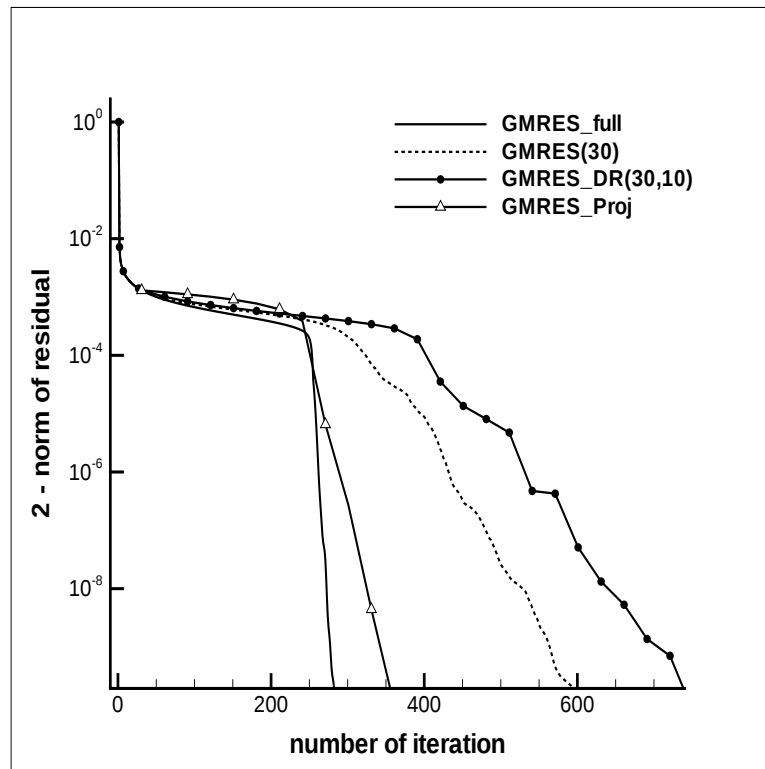


Рис. 2: Сходимость вариантов GMRES для задачи с разрывными краевыми условиями. Сетка 128×128

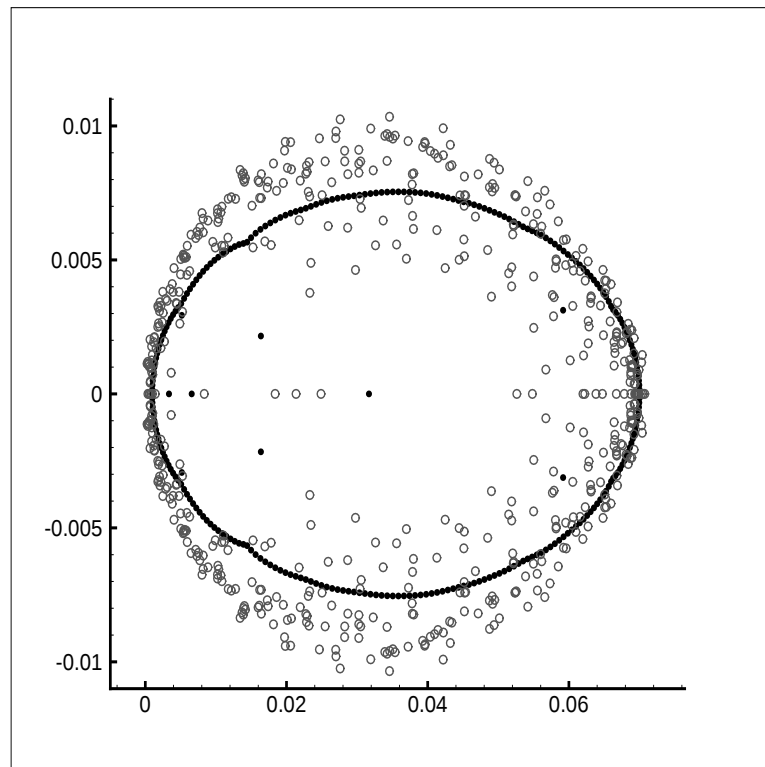


Рис. 3: Гармонические числа Ритца GMRES_full (кружки) и GMRES(30) (кольца) для задачи с разрывными граничными условиями

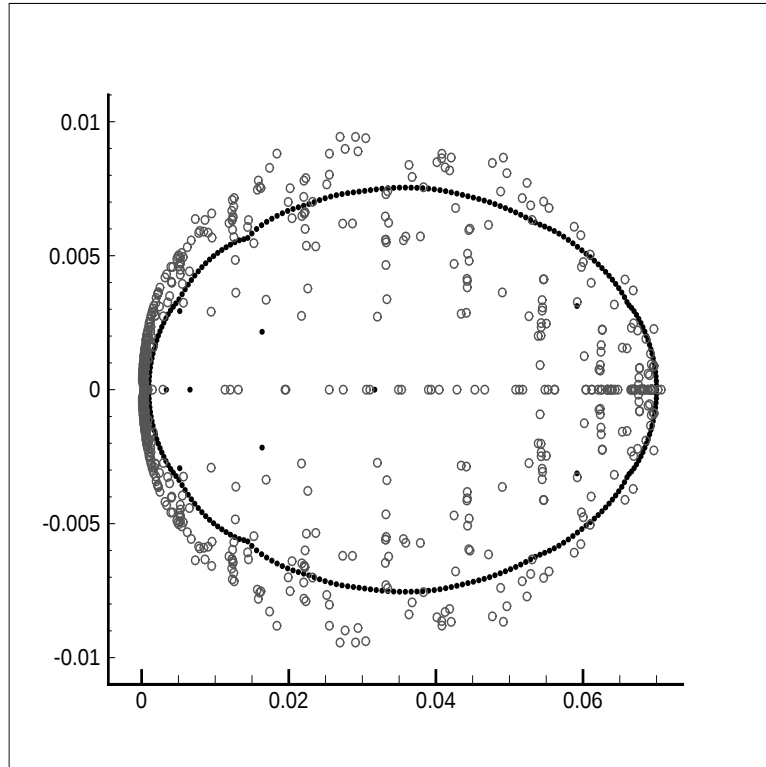


Рис. 4: Гармонические числа Ритца GMRES_full (кружки) и GMRES_DR(30,10)(кольца) для задачи с разрывными граничными условиями

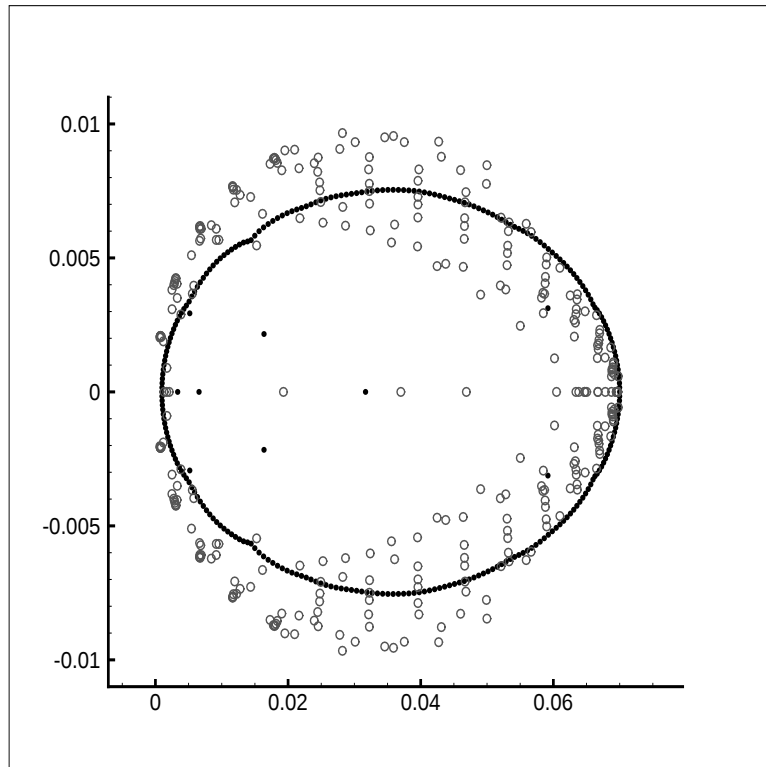


Рис. 5: Гармонические числа Ритца GMRES_full (кружки) и GMRES_Proj (кольца) для задачи с разрывными граничными условиями

ключевым для сходимости проекционных методов. Будем выбирать базис U_k , меняя правило выбора k векторов при каждом рестарте $\text{GMRES}(m)$, опираясь не только на вектора Шура, соответствующие минимальным по модулю собственным значениям матрицы A . Стратегия выбора (ниже мы будем называть ее сдвиговой стратегией) может быть самой разной и, возможно, зависеть от специфики рассматриваемой задачи. В настоящий момент мы не можем отдать предпочтение какой-либо из них. Ниже при описании численных экспериментов мы предлагаем несколько сдвиговых стратегий и приводим результаты, к которым приводит их применение.

4 Численные эксперименты

Сравним поведение различных вариантов GMRES на некоторых модельных примерах, возникающих в результате конечно-элементных аппроксимаций 2D уравнений конвекции-диффузии. Выберем для сравнения классический вариант GMRES_full , вариант с рестартами $\text{GMRES}(m)$, $\text{GMRES_DR}(m, k)$ и, наконец, варианты $\text{GMRES}(m, k)_\text{Proj}$ со сдвигами, которые будем обозначать как $\text{GMRES}(m, k)_\text{Proj}(L)$, где L – параметр, определяющий правило выбора k векторов, с помощью которых строится корректирующее подпространство U_{k+1} , см. §2.3.

Сходимость контролируется обычным образом. На каждом цикле внешних итераций i рассчитывается невязка $\|b - Au_i\|$. Начальная невязка равняется $\|b\|$, что соответствует выбору нулевого начального приближения. Итерации продолжаются до тех пор, пока относительная невязка не становится меньше указанной точности,

$$\frac{\|b - Au_i\|}{\|b\|} < \delta.$$

В приведенных ниже экспериментах мы использовали различные стратегии сдвига для варианта $\text{GMRES}(m, k)_\text{Proj}(L)$. Всюду ниже мы будем опускать из названия значения величин m, k , если из контекста эти значения понятны. Значение $L = 1$ на представленных рисунках означает использование стратегии без сдвигов, т.е. начальный базис для построения корректирующего пространства $\text{span}(U_{k+1})$ всегда состоит из векторов Шура, соответствующих k первым гармоническим числам Ритца. Мы предполагаем, что вектора Шура упорядочены по возрастанию модуля соответствующих гармонических чисел Ритца. Таким образом, для случая $L = 1$ всегда выбираем k первых векторов Шура.

Если $L \neq 1$, то корректирующее пространство на каждой внешней итерации меняется. Нас интересовали простые стратегии сдвига. Если $L = 2$ выбираются следующие друг за другом группы из k векторов Шура, т.е. сначала первые k векторов Шура, на следующей внешней итерации вторые k векторов и т.д. При $L = 3$ дважды выбираются первые k векторов, затем дважды вторые k векторов; в конце дважды выбираются последние k векторов Шура. И, наконец, мы рассматриваем вариант (ему соответствует параметр $L = 4$), когда сдвиг осуществляется только после троекратного повтора каждой группы векторов.

4.1 Уравнение Пуассона

Рассмотрим уравнение Пуассона

$$-\varepsilon_x u_{xx} - \varepsilon_y u_{yy} = f$$

в единичном квадрате $\Omega = [0, 1] \times [0, 1]$ с нулевыми краевыми условиями Дирихле. Коэффициенты диффузии ε_x , ε_y предполагаем равными 1. Для аппроксимации используем равномерную квадратную сетку, состоящую из $N \times N$ ячеек. Получающаяся матрица имеет в этом случае размер $(N + 1)^2 \times (N + 1)^2$.

Правая часть f есть результат умножения $f = A u_{exact}$ при заданном решении $u_{exact}(x, y)$, в качестве которого выберем линейную комбинацию гладких и негладких собственных функций оператора Лапласа

$$u_{exact}(x, y) = u^{smooth} + u^{non-smooth},$$

$$u^{smooth} = \sum_{i=i_0}^{i_0+I} \sin(\pi i x) \sum_{j=j_0}^{j_0+J} \sin(\pi j y), \quad i_0 = j_0 = 1, \quad I = J = 10,$$

$$u^{non-smooth} = \sum_{i=i_0}^{i_0+I} \sin(\pi i x) \sum_{j=j_0}^{j_0+J} \sin(\pi j y), \quad i_0 = j_0 = \frac{N}{2}, \quad I = J = 10.$$

Сравнивая сходимость представленных вариантов GMRES для уравнения Пуассона на квадратной сетке 256×256 (рис. 6) при заданной относительной точности $\delta = 10^{-7}$, мы наблюдаем улучшение сходимости при применение сдвигового варианта GMRES(30,10)_Proj(2) по сравнению с вариантом GMRES(30,10)_Proj(1) и GMRES_DR. Однако при задании точ-

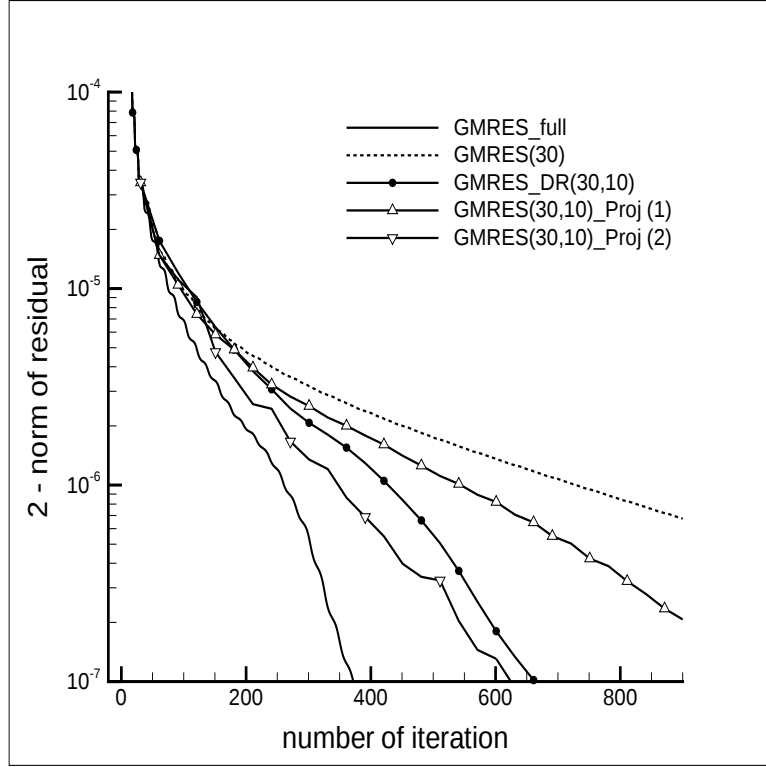


Рис. 6: Сходимость вариантов GMRES при $\delta = 10^{-7}$. Уравнение Пуассона. Сетка 256×256 .

ности $\delta = 10^{-10}$ вариант GMRES(30,10)_Proj(2) уступает GMRES_DR (см. рис. 7). По построению, и это подтверждается в расчетах, GMRES_DR лучше, чем остальные из представленных вариантов приближает группу минимальных собственных значений исходного оператора. Можно предположить, что именно это свойство позволяет GMRES_DR показывать лучшую асимптотическую сходимость.

Рассмотрим вопрос о влиянии выбора стратегии построения корректирующего базиса U_{k+1} на сходимость всего итерационного процесса в целом (см. рис. 8). В представленных на рисунке графиках две стратегии $L = 2$ и $L = 3$ (суть каждой из них описана выше) показывают близкий характер сходимости, сравнимый на данной задаче с вариантом GMRES_DR. Каждое скачкообразное уменьшение нормы невязки на графике соответствует переходу от последнего набора векторов U_k , соответствующего максимальным гармоническим числам Ритца, к первому набору, связанному с их минимальными значениями. Близкий характер сходимости при двух выбранных стратегиях сдвига не исключает существования оптимальной сдвиговой тактики, отличной от рассмотренных в данной работе.

Смоделируем ситуацию с практической точки зрения труднореализуемую, но позволяющую нам обозначить цель и, как следствие, возможный путь ее достижения.

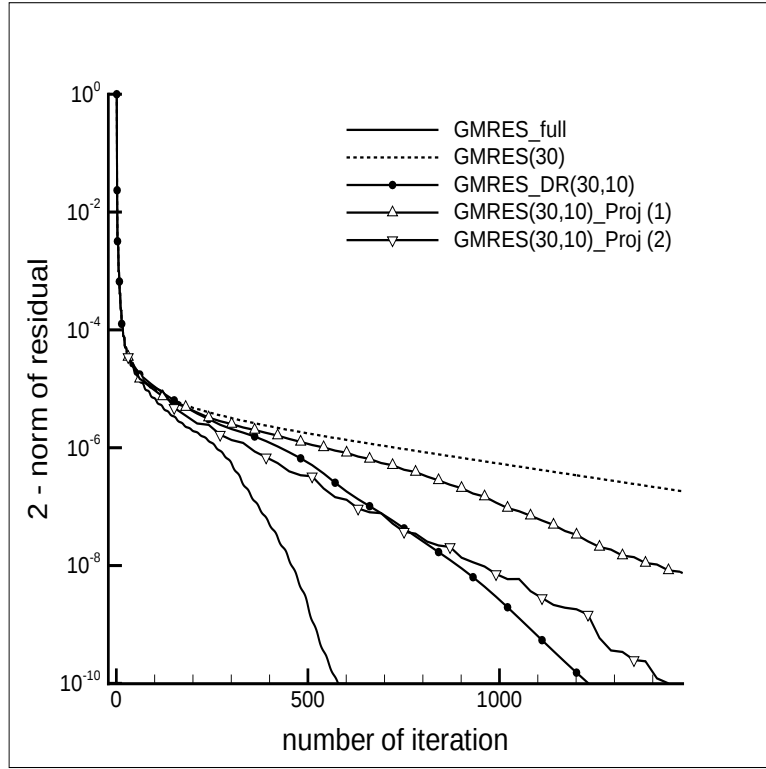


Рис. 7: Сходимость вариантов GMRES при $\delta = 10^{-10}$. Уравнение Пуассона. Сетка 256×256 .

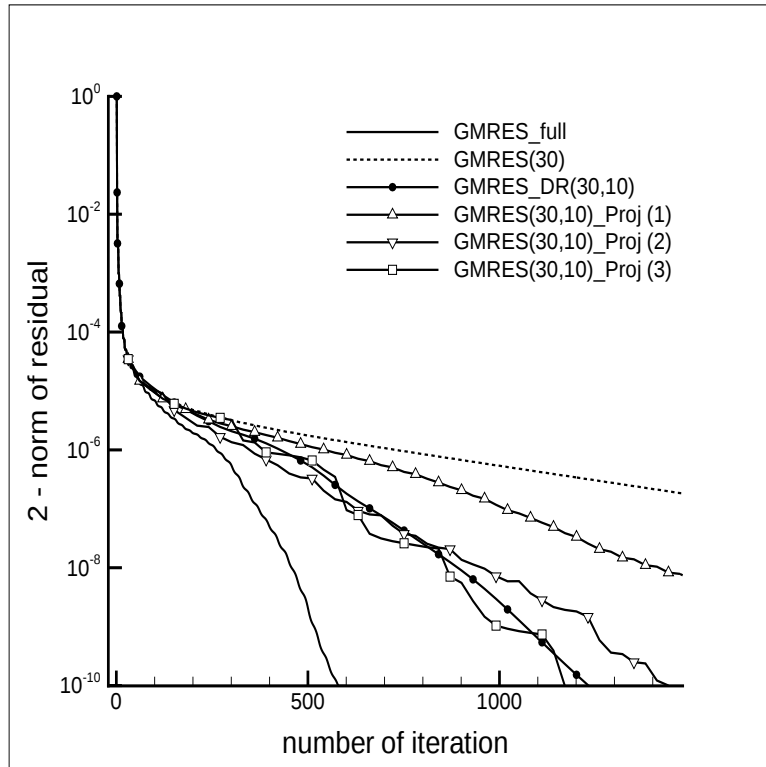


Рис. 8: Зависимость скорости сходимости от стратегии выбора корректирующего базиса U_{k+1} . Уравнение Пуассона. Сетка 256×256 .

Прежде, чем корректировать невязку, модифицируем уже выбранное подпространство $\text{span}\{U_{k+1}\}$. Для этого применим к ортонормированному базису U_{k+1} сглаживающую процедуру

$$S = \prod_{q=k+1}^{q=m} (I - \frac{1}{\lambda_q} A).$$

Здесь λ_q – соответствующие гармонические числа Ритца. С практической точки зрения процедура эта неприемлема по двум причинам. Для нормальных матриц с вещественным спектром она определена, но является дорогостоящей. Для задач с комплексным спектром такой многочлен (или его аналог) неизвестен.

Предлагаемая модификация алгоритмически выглядит следующим образом.

1. Берем построенный корректирующий базис U_{k+1} , $U_{k+1}^T U_{k+1} = I_{k+1}$ и p раз применяем к нему сглаживающую процедуру S . Здесь p – заданный параметр.
2. Ортонормализуем модифицированный базис $U = U_{k+1}^{new} = S(p)U_{k+1}$. Для этого формируем симметричную матрицу $G = U^T U$ размера $(k+1) \times (k+1)$ и делаем ее собственное разложение

$$G = V \text{diag}(g_1, g_2, \dots, g_{k+1}) V^T, \quad V V^T = I_{k+1},$$

$$g_1 \geq g_2 \geq \dots \geq g_{k+1} \geq 0.$$

Искомый ортонормированный базис W имеет вид

$$W = U V_{k+1} \text{diag} \left(\frac{1}{\sqrt{g_1}}, \frac{1}{\sqrt{g_2}}, \dots, \frac{1}{\sqrt{g_{k+1}}} \right).$$

3. Окончательно корректирующий базис имеет вид $U_{k+1} = W$.

Влияние сглаживающей процедуры на сходимость итерационного процесса представлено на рис. 9. Варианты, в которых применялась сглаживающая процедура, обозначены GMRES(m, k)_Proj(L)_S(p). Число сглаживающих итераций намеренно выбрано большим, равным 4. Как видно из графиков, сглаживающая процедура ускоряет итерационный процесс. Скорость сходимости делает этот вариант (в выбранной системе координат) сравнимым с классическим методом GMRES. Как уже было сказано вы-

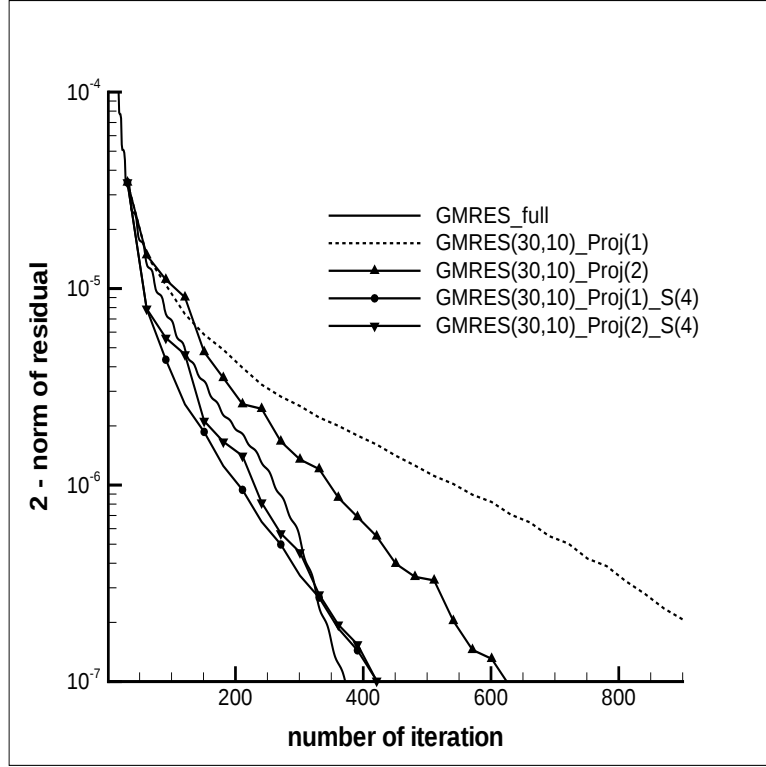


Рис. 9: Сходимость при использовании сглаживающей процедуры. Уравнение Пуассона. Сетка 256×256 .

ше, мы не рассматриваем этот подход с практической точки зрения, но нас интересует, в чем заключается суть проведенной модификации. Для выяснения этого рассчитаем и сравним гармонические числа Ритца, присущие корректирующему подпространству $\text{span}\{U_{k+1}\}$.

На рис. 10 обозначены все гармонические числа Ритца, рассчитанные для всех корректирующих подпространств, получаемых в ходе итерационного процесса. Видно, что модификация базиса приводит к кластеризации гармонических чисел Ритца, причем количество кластеров совпадает с числом сдвигов.

В заключение рассмотрим как зависит сходимость итерационного процесса от размера сеточной задачи (рис. 11). Для этого возьмем последовательность, состоящую из трех задач, полученных в результате аппроксимации уравнения Пуассона на последовательности сгущающихся сеток 128×128 , 256×256 , 512×512 . Увеличение размера сеточной задачи приводит к двум явлениям: с одной стороны, в характере сходимости проявляется немонотонность, а с другой стороны, вариант $\text{GMRES_Proj}(2)$ в сравнении с GMRES_DR при увеличении размера задачи показывает лучшую скорость сходимости.

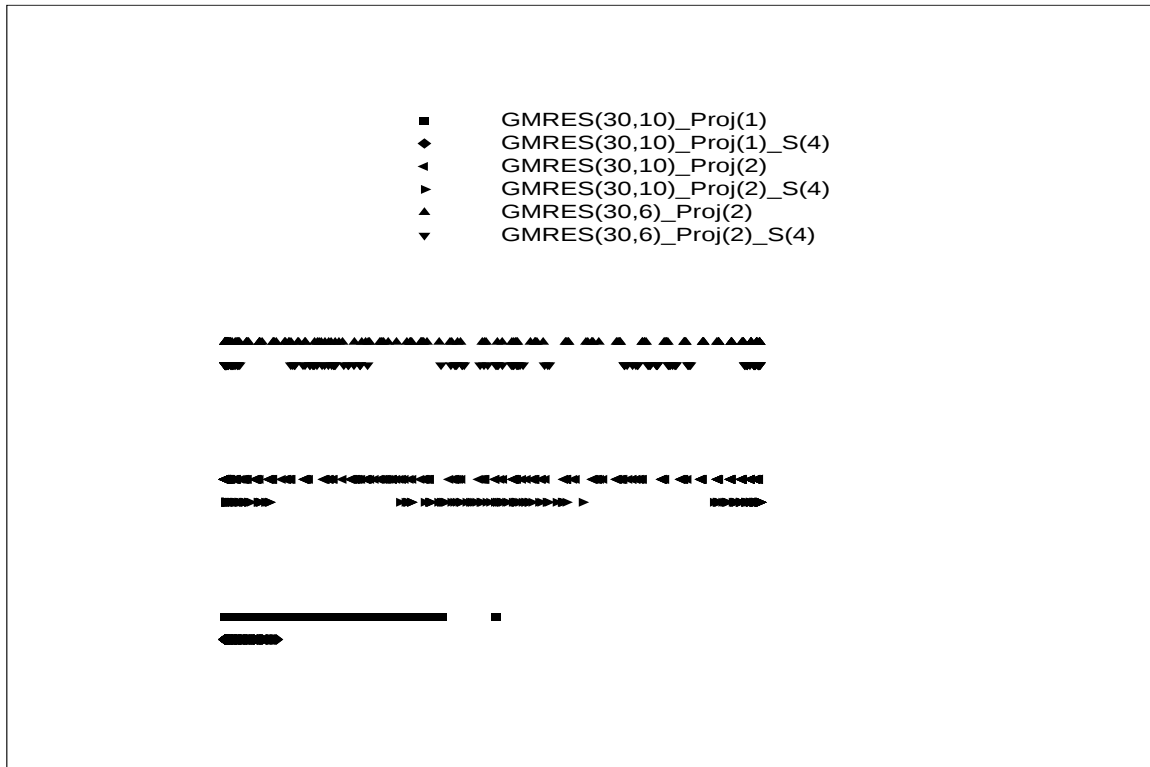


Рис. 10: Гармонические числа Ритца, соответствующие корректирующему пространству $\text{span}\{U_k\}$. Уравнение Пуассона. Сетка 256×256 .

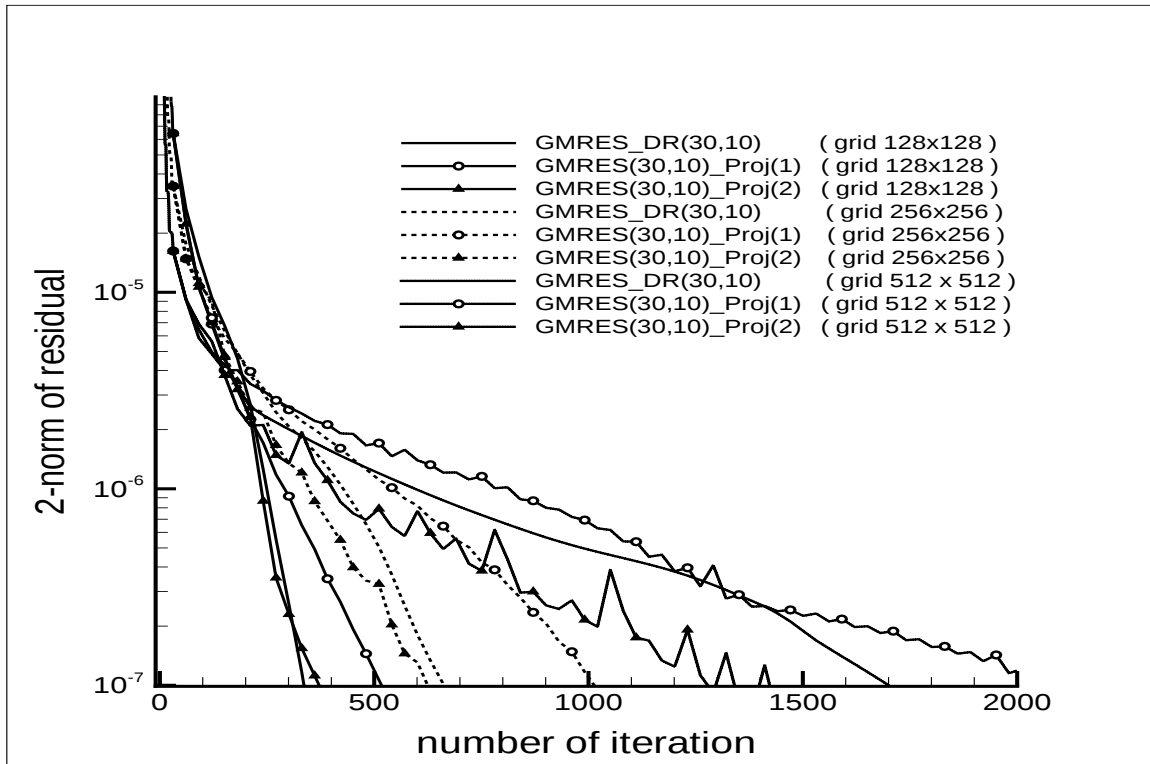


Рис. 11: Сходимость вариантов GMRES на разных сетках. Уравнение Пуассона. Сетки 128×128 , 256×256 , 512×512 .

4.2 Уравнение конвекции-диффузии.

Перейдем к рассмотрению уравнения конвекции-диффузии

$$-\varepsilon(u_{xx} + u_{yy}) + p_1 u_x + p_2 u_y = f.$$

Будем решать уравнение в единичном квадрате $\Omega = [0, 1] \times [0, 1]$ на треугольной сетке, полученной из квадратной. Поступим в той же манере, что и в случае уравнения Пуассона. А именно, возьмем в качестве решения ту же комбинацию собственных функций оператора Лапласа. В результате получим задачу с нулевыми краевыми условиями и с правой частью, рассчитанной после подстановки решения в уравнение.

Зафиксируем поле скоростей $p_1 = \frac{\sqrt{2}}{2}, p_2 = \frac{\sqrt{2}}{2}$ (оно совпадает с полем скоростей представленной выше задачи с разрывными краевыми условиями), а коэффициент диффузии будем варьировать, меняя соотношение между диффузионной и конвективной частями.

Первые эксперименты проведем для задач с доминирующей диффузией, $\varepsilon = 1$. Сеточное число Пекле при таком выборе параметров меньше единицы

$$Pe = \frac{|p|h}{2\varepsilon} = 0.004.$$

Сходимость различных вариантов GMRES (см. рис. 12) аналогична характеру сходимости, показанной для уравнения Пуассона выше: вариант GMRES_Proj(2) показывает лучшую сходимость, чем GMRES_Proj(1) и GMRES_DR, но асимптотически уступает GMRES_DR. При увеличении размера задачи (рис. 13) GMRES_Proj(2) улучшает качества сходимости в большей степени, чем остальные кандидаты.

Влияние размера k на сходимость итерационного процесса в целом можно увидеть на рис. 14. Последние эксперименты (рис. 15) анализируют зависимость скорости сходимости от выбора сдвиговой стратегии. Для задач с доминирующей диффузией наличие сдвига приводит к лучшим параметрам сходимости.

Перейдем к рассмотрению варианта, в котором уравновешены диффузионный и конвективный члены $\varepsilon = 0.01$, $Pe = 0.4$. В этой ситуации (см. рис. 16) применение сдвигов не приводит к улучшению сходимости. В примере, выбранном для анализа, GMRES_DR(30,10) никак не улучшает сходимость классического GMRES(30), а варианты GMRES_Proj(L) дают похожие графики сходимости для обоих вариантов – со сдвигами и без них – хотя и лучшие, чем два других варианта (GMRES_DR, GMRES(m)).

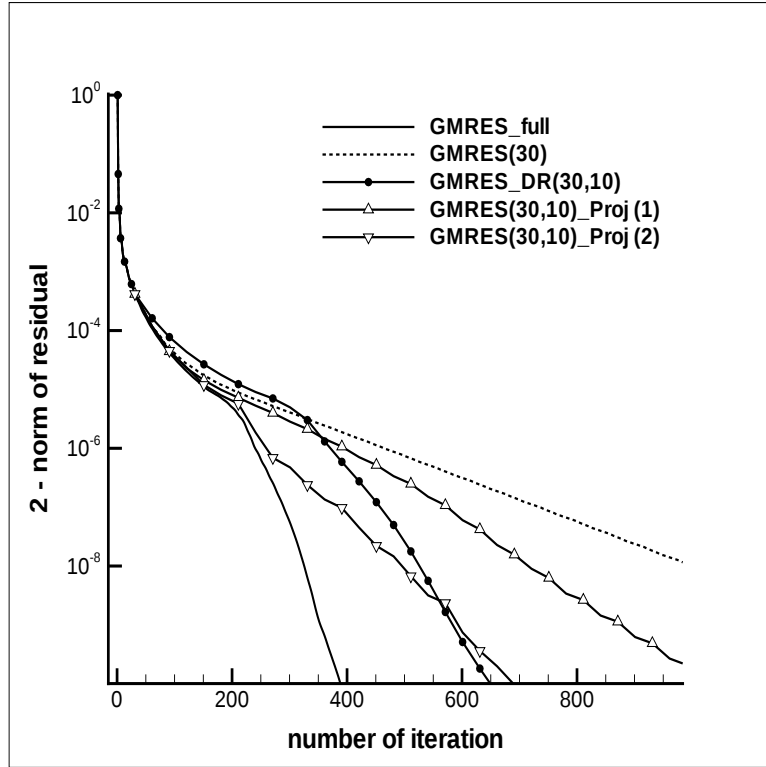


Рис. 12: Сходимость различных вариантов GMRES. Уравнение конвекции-диффузии $-(u_{xx} + u_{yy}) + \frac{\sqrt{2}}{2}u_x + \frac{\sqrt{2}}{2}u_y = f$. Сетка 128×128 .

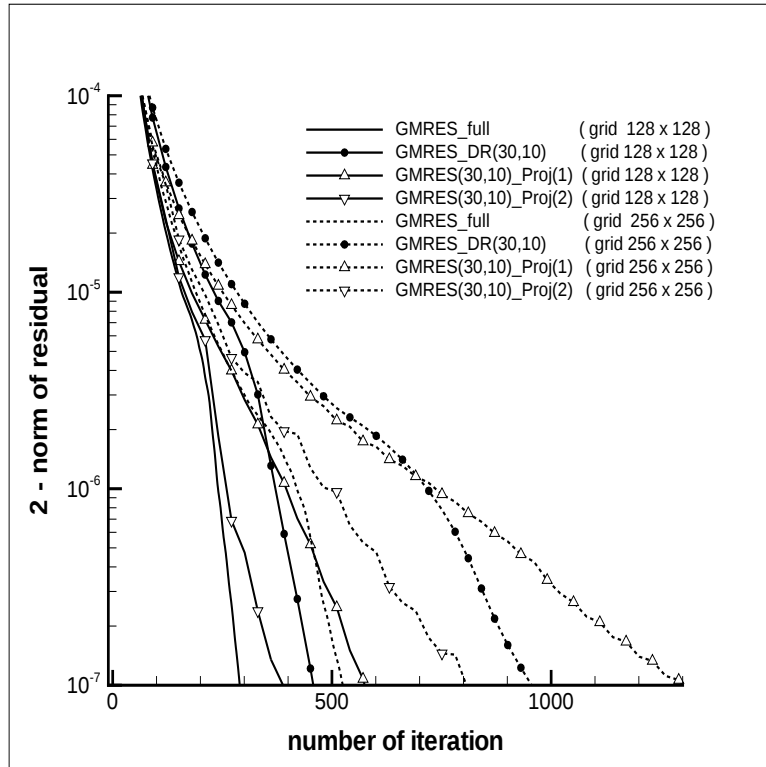


Рис. 13: Сходимость различных вариантов GMRES на разных сетках. Уравнение конвекции-диффузии $-(u_{xx} + u_{yy}) + \frac{\sqrt{2}}{2}u_x + \frac{\sqrt{2}}{2}u_y = f$. Сетки 128×128 , 256×256 .

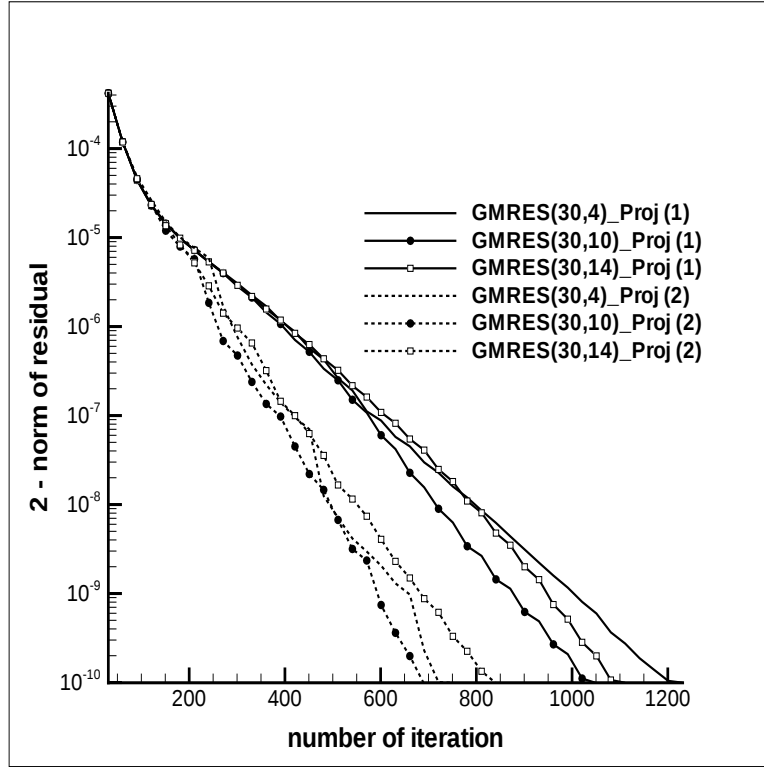


Рис. 14: Зависимость сходимости от размера k корректирующего базиса. Уравнение конвекции-диффузии. $-(u_{xx} + u_{yy}) + \frac{\sqrt{2}}{2}u_x + \frac{\sqrt{2}}{2}u_y = f$. Сетка 128×128 .

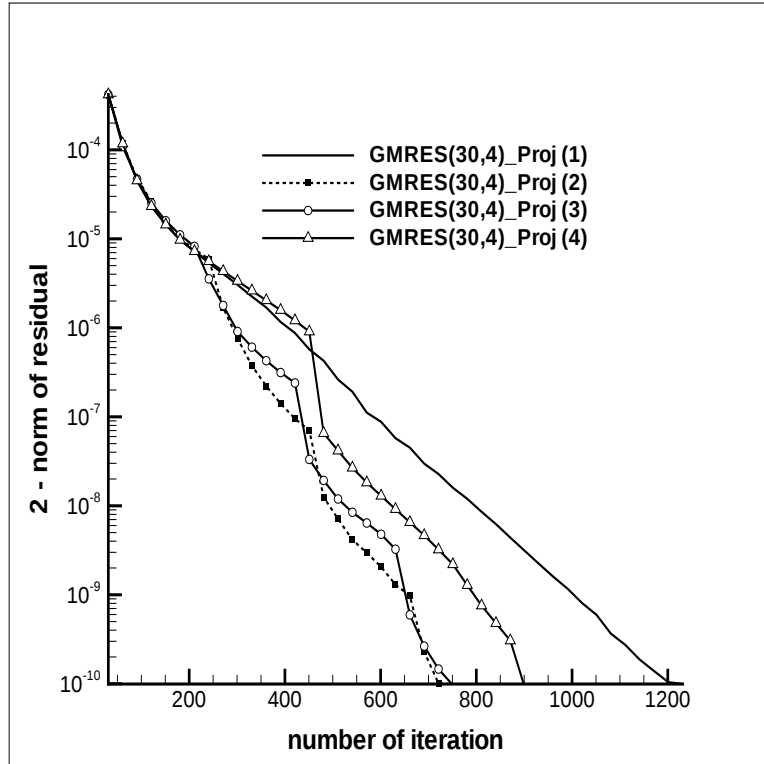


Рис. 15: Зависимость сходимости от стратегии выбора корректирующего базиса. Уравнение конвекции-диффузии $-(u_{xx} + u_{yy}) + \frac{\sqrt{2}}{2}u_x + \frac{\sqrt{2}}{2}u_y = f$. Сетка 128×128 .

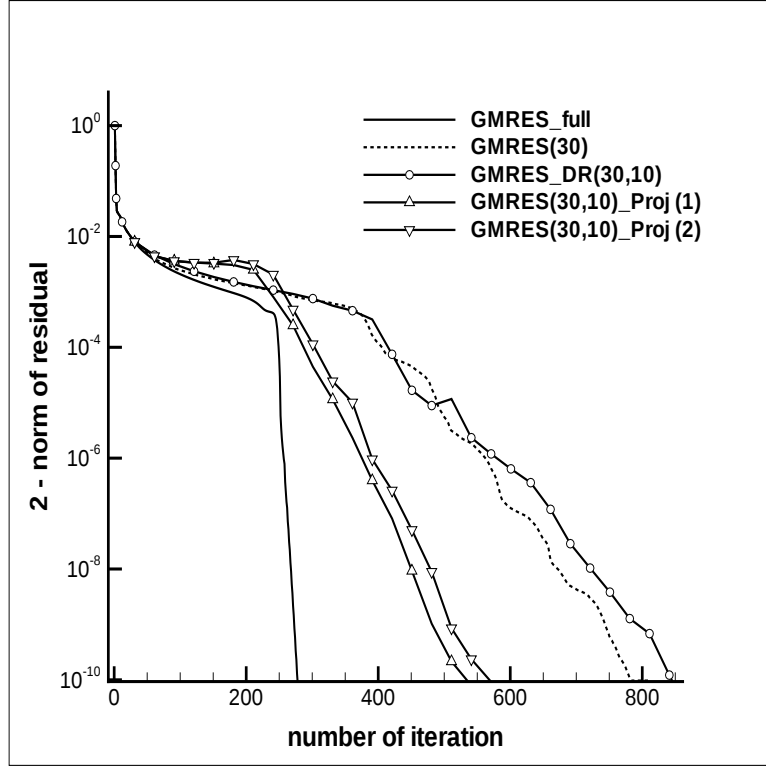


Рис. 16: Сходимость различных вариантов GMRES. Уравнение конвекции-диффузии $-0.01(u_{xx} + u_{yy}) + \frac{\sqrt{2}}{2}u_x + \frac{\sqrt{2}}{2}u_y = f$. Сетка 128×128 .

Зависимость сходимости от размера задачи показана на рис. 17. Увеличение размера задачи только подчеркивает уже сделанный вывод о практически равноправном характере сходимости вариантов с коррекцией $\text{GMRES_Proj}(L)$.

Рис. 18 демонстрирует слабую зависимость скорости сходимости вариантов GMRES_Proj как от сдвиговой стратегии, так и от размера k корректирующего базиса U_{k+1} .

В конце вернемся к уже рассмотренной выше задаче с разрывными краевыми условиями. Это – задача с доминирующей конвекцией $\varepsilon = 10^{-6}$, $Pe \approx 4 \cdot 10^4$. Сдвиговой вариант $\text{GMRES_Proj}(2)$ хотя и показывает сходимость лучше, чем GMRES_DR , однако уступает варианту без сдвигов $\text{GMRES_Proj}(1)$, рис. 19.

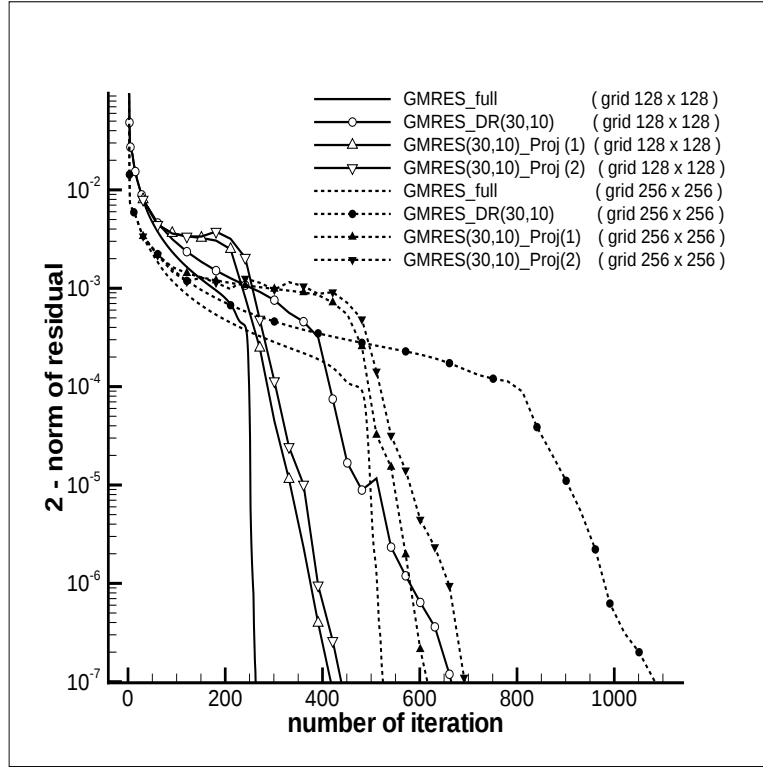


Рис. 17: Зависимость сходимости от размера задачи. Уравнение конвекции-диффузии $-0.01(u_{xx} + u_{yy}) + \frac{\sqrt{2}}{2}u_x + \frac{\sqrt{2}}{2}u_y = f$. Сетки 128×128 , 256×256 .

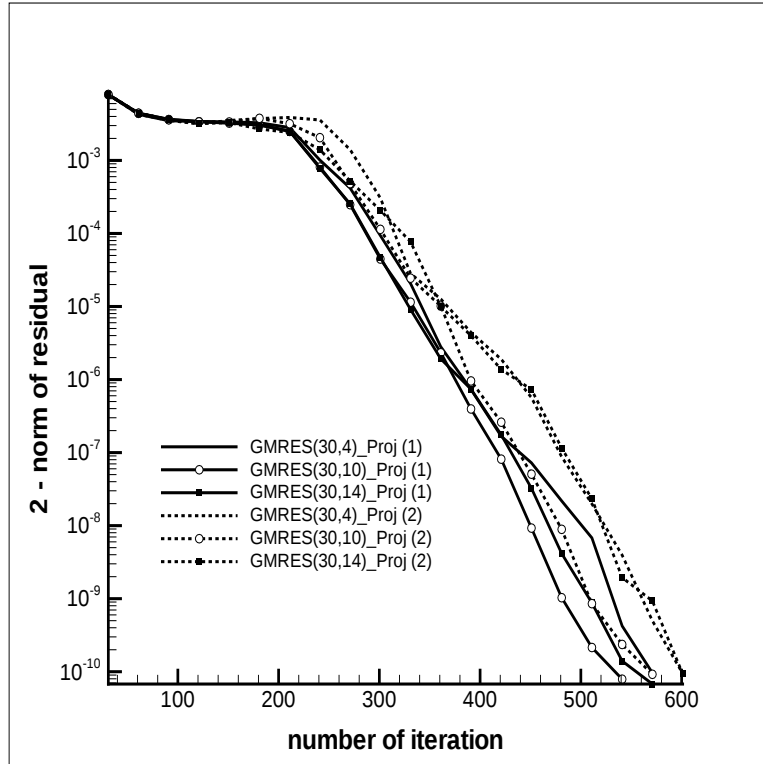


Рис. 18: Зависимость сходимости от размера k корректирующего базиса. Уравнение конвекции-диффузии $-0.01(u_{xx} + u_{yy}) + \frac{\sqrt{2}}{2}u_x + \frac{\sqrt{2}}{2}u_y = f$. Сетка 128×128 .

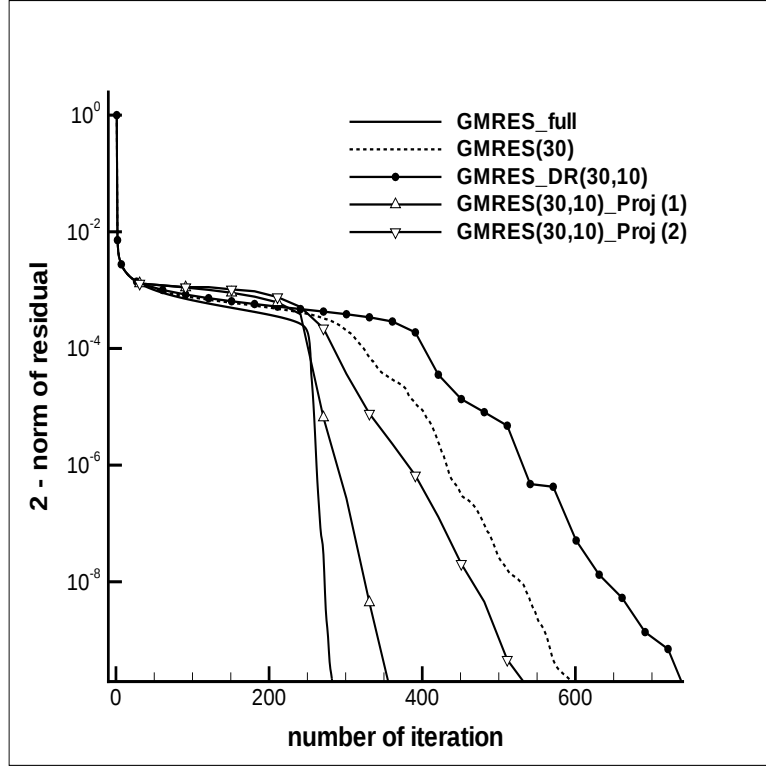


Рис. 19: Сходимость различных вариантов GMRES. Задача с разрывными краевыми условиями. Уравнение конвекции-диффузии $-10^{-6}(u_{xx} + u_{yy}) + \frac{\sqrt{2}}{2}u_x + \frac{\sqrt{2}}{2}u_y = f$. Сетка 128×128 .

5 Заключение

Для приведенных задач (уравнение Пуассона и задачи конвекции-диффузии с доминирующей диффузией) вариант GMRES_Proj со сдвигами показывает лучшие свойства сходимости в сравнении с остальными из предложенных вариантов. В выбранном модельном примере решение состоит из гладкой и негладкой частей, взятых с равными коэффициентами. Однако, если первые гладкие функции выделить существенно большими коэффициентами, GMRES_DR покажет лучшую сходимость по сравнению с любыми из сдвиговых вариантов, по крайней мере в их текущем варианте.

Применительно к задачам с доминирующей конвекцией, а также к задачам с уравновешенными диффузионным и конвективными членами, проекционный вариант без сдвигом показывает лучшую скорость сходимости по сравнению с другими подходами.

Полученные результаты расчетов показывают перспективность выбранного подхода с чередованием корректирующих подпространств небольшой размерности $k < m$ и циклов GMRES(m), и требуют дополнительных исследований.

Список литературы

- [1] *Morgan R.B.* Implicitly restarted GMRES and Arnoldi methods for nonsymmetric systems of equations.— SIAM, V.21, No. 4, pp.1112–1135, 2000.
- [2] *Morgan R.B.* GMRES with deflated restarting. SIAM, V. 24, pp.20–37, 2002.
- [3] *Simoncini V, Szyld D.B.* Recent computational developments in Krylov subspace methods for linear systems. – Numerical Linear Algebra w/Appl., V. 14, N.1, pp.1–59, 2007.
- [4] *Жуков В.Т., Феодоритова О.Б., Янг Д.П.* Итерационные алгоритмы для схем конечных элементов высокого порядка. – Математическое моделирование, т. 16, N 7, 2004.
- [5] *Жуков В.Т., Феодоритова О.Б.* Многосеточный метод для неструктурных конечно-элементных дискретизаций уравнений аэродинамики // Препринт ИПМ им. М.В. Келдыша РАН, 5, Москва, 2008
- [6] *Saad Y.* Iterative methods for sparse linear systems.— The PWS Publishing Company, Boston, 1996.
- [7] *Деммель Дж.* Вычислительная линейная алгебра. Теория и приложения. Пер. с англ. – М.: Мир, 2001. 430 с.
- [8] *В.Т. Жуков, Н.Д. Новикова, Л.Г. Страховская, Р.П. Федоренко, О.Б. Феодоритова.* Применение метода конечных суперэлементов для решения задач конвекции-диффузии. – Математическое моделирование, т. 14, № 11, 2002, с. 78–92.
- [9] *Brezzi F., Franca L.P., Russo A.* Further consideration on residual free bubbles for advective-diffusive equation. Comput. Methods Appl. Mech. Engrg., 166(1998), pp.25–33.
- [10] *Hughes T.J.R., Franca L.P., Hilbert G.M.* A new finite element formulation for computational fluid dynamics: VIII. The Galerkin/least squares method for advective-diffusive equations. Comput. Methods Appl. Mech. Engrg., 73(2): 173–189, 1989.