МАССИВНО-ПАРАЛЛЕЛЬНЫЕ ПРЕДОБУСЛОВЛЕННЫЕ ИТЕРАЦИОННЫЕ МЕТОДЫ РЕШЕНИЯ СТАНДАРТНЫХ ЗАДАЧ ЛИНЕЙНОЙ АЛГЕБРЫ С РАЗРЕЖЕННЫМИ МАТРИЦАМИ БОЛЬШОГО РАЗМЕРА

И.Е. Капорин[1], О.Ю.Милюкова[2], Ю.Г. Бартенев[3]

[1]Вычислительный центр им. А.А.Дородницына РАН, Москва
[2]Институт прикладной математики им. М.В.Келдыша РАН, Москва
[3]Российский федеральный ядерный центр – Всероссийский НИИ экспериментальной физики, Саров

Для систем линейных алгебраических уравнений и обобщенных частичных задач на собственные значения (порядка до десятков миллионов), возникающих, в частности, при дискретизации краевых задач для эллиптических уравнений в частных производных, описаны массивно-параллельные алгоритмы приближенного решения. Продемонстрирована возможность их эффективной параллельной реализации на многопроцессорных вычислительных системах с распределенной памятью при числе процессов MPI до 3000.

Введение

Одной из наиболее часто встречающихся трудоемких стандартных вычислительных задач является решение невырожденных систем линейных алгебраических уравнений (СЛАУ) $Ax = b$ с разреженной $n \times n$-матрицей A большого размера n. Еще больших затрат ресурсов требует решение частичной обобщенной задачи на собственные значения $Ax = \lambda Bx$ с большими разреженными симметричными положительно определенными матрицами. При этом возможность экономии памяти дает итерационным методам серьезное преимущество по сравнению с методами, использующими точную разреженную треугольную факторизацию матрицы A, если требуется решать задачи большого размера. Кроме того, итерационные методы обладают более гибкой структурой вычислений, что облегчает их параллельную реализацию. Однако для многих важных классов задач (например, СЛАУ с невырожденными матрицами общего вида) известные массивно-параллельные итерационные методы (напр., методы проектирования на крыловские подпространства с диагональным или блочно-диагональным предобусловливанием) обладают недостаточной надежностью и эффективностью.

Таким образом, решение стандартных задач линейной алгебры с большими разреженными матрицами на высокопроизводительных многопроцессорных системах все еще является открытой проблемой вычислительной математики и требует дальнейших исследований. Вместе с тем, достаточно широко распространилось мнение, что в настоящее время следует разрабатывать только такие методы решения прикладных задач большого размера, которые вообще не приводят к необходимости решения разреженных СЛАУ (т.наз. “matrix-free methods”). Представляется, однако, что дальнейшая разработка массивно-параллельных решателей СЛАУ достаточно перспективна ввиду наличия больших теоретических и практических заделов (как в области построения решателей, так и в способах их эффективного использования), а также ясных перспектив их развития, связанных, например, с оптимизацией используемых предобусловливателей. Кроме того, разработка массивно-параллельных предобусловливаний позволяет, например, построить на их основе эффективные решатели для рассматриваемых спектральных задач.

Ниже описано достаточно общее, эффективное и надежное предобусловливание [1-3, 8], пригодное для ускорения метода сопряженных градиентов и других итерационных методов, использующих проекцию на крыловские подпространства (BiCGStab, GMRes), а также для построения итерационных методов решения спектральных задач. Также приводятся результаты тестирования этих решателей на реальных задачах с применением современных многоядерных мультипоцессоров с эффективным использованием до 3000 вычислительных ядер.

1. Предобусловливание посредством оптимального согласования приближенных треугольных разложений перекрывающихся подматриц

Алгоритмической основой реализованных решателей является предобусловливание, использующее специальный метод согласования приближенных треугольных факторизаций перекрывающихся главных подматриц исходной матрицы коэффициентов [3]. Количество используемых главных подматриц принимается равным доступному числу процессов MPI, что обеспечивает хорошую параллельную масштабируемость.

Так, для решения СЛАУ применяются итерационные методы проекций на крыловские подпространства

$$x_i - x_0 \in K_i(Hr_0; HA) = \mathrm{span}(Hr_0, HAHr_0, \ldots, (HA)^{i-1} Hr_0),$$

где H - матрица-предобусловливатель. Таким образом, невязка i-го приближения к решению имеет вид $r_i = b - Ax_i = \pi_i(AH)r_0$, где $\pi_i(t)$ - подходящий многочлен, удовлетворяющий $\pi_i(0) = 1$ (так, $\pi_i(\cdot)$ в методе GMRes выбирается из условия минимизации евклидовой нормы невязки). При этом должно достигаться достаточно хорошее приближение типа $I \approx AH$ (случай использования метода сопряженных градиентов детально рассмотрен, напр., в [3]).

Опишем теперь конструкцию предобусловливателя, предназначенного для применения к матрицам с достаточно «сильной» диагональю (например, к положительно определенным матрицам), который лишь несущественно сложнее «блочного метода Якоби», однако способен обеспечивать существенно более быструю сходимость итераций.

Пусть матрица A переупорядочена и разбита на блоки, причем на блочной диагонали расположены p квадратных блоков размера n_s, $1 \le s \le p$. Обозначим $k_s = n_1 + \ldots + n_s$, и пусть $m_s \ge n_s$ - размеры расширенных блоков. Определим прямоугольные матрицы

$$V_s = [e_{j_s(1)} | \ldots | e_{j_s(m_s - n_s)} | e_{k_{s-1}+1} | \ldots | e_{k_s}],$$

столбцы которых являются единичными n-векторами, где $k_{s-1}+1, \ldots, k_s$ представляют собой индексы s-го блока, а $j_s(1), \ldots, j_s(m_s - n_s)$ являются индексами перекрытия, причем $j_s(.) \le k_{s-1}$. Для построения индексов перекрытия s-го блока используется множество столбцовых позиций соответствующего блока строк матрицы A^q, где $q>0$ (заметим, что при $q=0$ перекрытие пусто, и получается т.наз. блочный метод Якоби). Применим теперь к каждой расширенной $m_s \times m_s$-подматрице $V_s^T A V_s$ приближенное треугольное разложение 2-го порядка «по значению» ILU2(τ) (the 2nd order LU-factorization), см. [4]:

$$V_s^T A V_s = L_s U_s + L_s R_s + K_s U_s - S_s, \quad \|R_s\| = O(\tau), \quad \|K_s\| = O(\tau), \quad \|S_s\| = O(\tau^2),$$

где R_s и K_s - основные матрицы погрешностей (строго треугольные, верхняя и нижняя, соответственно), S_s - остаточная матрица погрешностей, $0 < \tau \le 1$ - порог отсечения, $s = 1, \ldots, p$. Предлагаемый предобусловливатель, получивший обозначение BIILU(p;q)-ILU2(τ) (от Block Incomplete Inverse LU), запишем в следующем виде [1, 2]:

$$H = \sum_{s=1}^{p} V_s U_s^{-1} \begin{bmatrix} 0 & 0 \\ 0 & I_{n_s} \end{bmatrix} L_s^{-1} V_s^T.$$

Строгое обоснование оптимальности такого метода предобусловливания (при заданной заранее структуре перекрытий) имеется в случае произвольной симметричной положительно определенной матрицы A (когда $L_s = U_s^T$, а для отыскания U_s применяется IC2(τ)-разложение), см. [1, 3]. Отметим, что для решения СЛАУ с несимметричной матрицей, не обладающей "сильной" диагональю, можно использовать другое представление для H, см. [1, 3], а к подматрицам типа $V_s^T A V_s$ применять более общее разложение ML-ILU(τ) [9, 10].

Реализация на FORTRAN /MPI соответствующих матрично-векторных операций аналогична описаннной в [2, 5]. В параллельных алгоритмах, реализующих предобусловленные методы BiCGStab и GMRes, для умножения матрицы коэффициентов на вектор используется ее распределенное представление, отвечающее специальному аддитивному разложению

$$A = \sum_{s=1}^{p} \tilde{V}_s [A_s \mid C_s] \tilde{W}_s^T, \qquad \tilde{V}_s = [e_{k_{s-1}+1} | \dots | e_{k_s}], \qquad \tilde{W}_s = [e_{k_{s-1}+1} | \dots | e_{k_s} | e_{j'_s(1)} | \dots | e_{j'_s(m'_s - n_s)}],$$

где $n_s \times n_s$-матрица A_s является s-м диагональным блоком матрицы A, а C_s - $n_s \times (m'_s - n_s)$-матрица, содержащая ненулевые столбцы подматрицы, расположенные слева и справа от A_s, а индексы $j'(\cdot)$ представляют собой номера соответствующих столбцовых позиций в A. Используются p процессов MPI (по числу слагаемых аддитивных представлений A и H), причем s-й процесс выполняет вычисления, отвечающие умножению на вектор s-го слагаемого матрицы A или предобусловливателя H. Для реализации межпроцессорных обменов разработаны надежные и эффективные алгоритмы на базе операций MPI_Send и MPI_Recv.

2. Алгоритм решателя частичной задачи на собственные значения

Рассмотрим симметричную частичную обобщенную задачу на собственные значения

$$Ax = \lambda Bx, \qquad A = A^T \geq 0, \qquad B = B^T \geq 0,$$

с разреженными матрицами A и B большого размера n, где требуется найти порядка 10^2 собственных пар (x, λ), отвечающих наименьшим собственным значениям λ.

Используется следующий метод, аналогичный известному локально-оптимальному блочному методу сопряженных градиентов (LOBPCG) для спектральных задач [6]:

$$V_0 = x_0, \qquad (V_0^T A V_0) Q_0 = (V_0^T B V_0) Q_0 \Lambda_0;$$

$$x_1 = (I - V_c V_c^T B) V_0 Q_0, \qquad r_1 = Ax_1 - Bx_1 \Lambda_0, \qquad V_1 = [x_1 \mid (I - V_c V_c^T B) H_d r_1];$$

для $k = 1,2,\dots$:

$$(V_k^T A V_k) Q_k = (V_k^T B V_k) Q_k \Lambda_k, \qquad Q_k = [q_k^{(1)} \mid q_k^{(2)} \mid *], \qquad \Lambda_k = Diag(\Lambda_k^{(1)}, \Lambda_k^{(2)}, *),$$

$$x_{k+1} = V_k q_k^{(1)}, \qquad r_{k+1} = Ax_{k+1} - Bx_{k+1} \Lambda_k^{(1)}, \qquad V_{k+1} = [x_{k+1} \mid (I - V_c V_c^T B) H_d r_{k+1} \mid x_k],$$

где $n \times m$-матрица x_k содержит итерируемые приближения к искомым собственным векторам, а $n \times m_c$-матрица V_c составлена из m_c уже вычисленных приближений. Предобусловливатель задается матрицей $H_d = p_{d-1}(H_1 A) H_1$, где H_1 - предобусловливатель для матрицы A, построенный по методу BIIC-IC2 (вариант указанной выше конструкции при $L_s = U_s^T$), а $p_{d-1}(\cdot)$ - специальным способом выбранный многочлен четной степени [7]. Построена численно устойчивая реализация метода, а также предлагаются достаточно надежные методы рестарта итераций, позволяющие при фиксированном размере блока m порядка 10 находить требуемое количество собственных пар. Спектральные подзадачи малого размера решаются методом Фалька-Лангемейера [8].

3. Программная реализация решателей СЛАУ и их тестирование

На основе описанного выше подхода были разработаны два комплекса программных модулей (библиотеки VC_RAN_SLAU и VC_RAN_EIG) предназначенные для приближенного решения больших разреженных систем линейных алгебраических уравнений с симметричными положительно-определёнными и несимметричными матрицами и частичной симметричной задачи на собственные значения, соответственно. Созданный программный код, написанный на языке FORTRAN-90 с использованием интерфейса MPI, удовлетворяет достаточно высоким требованиям надежности и эффективности функционирования, отличаясь в то же время небольшим объемом (менее 200Kbytes текста) и полной замкнутостью. Библиотеки создавалась для пакетов программ инженерного анализа, разработанных ФГУП "РФЯЦ-ВНИИЭФ".

Расчеты задач и замеры времени производились на вычислительной системе ЭВМ ВЦКП ВНИИЭФ. Вычислительные узлы связаны между собой быстрой коммуникационной сетью. В каждом узле многопроцессорной системы (возможно, за исключением последнего узла) во всех случаях использовались все вычислительные ядра.

Результаты тестирования на несимметричных СЛАУ, возникающих при моделировании задач аэродинамики пакетом ЛОГОС Аэрогидромеханика [12], а также на СЛАУ с симметричными положительно определенными матрицами, возникающих при расчете напряженно-деформированного состояния конструкций пакетом ЛОГОС Прочность [13] и при расчете задач подземной фильтрации пакетом НИМФА [14] (см. табл.1), приведены на Рис.1-2.

Таблица 1. Структурные характеристики тестовых матриц СЛАУ.

Матрица	***n***	***nz(A)***	тип	Источник
C20_A_0420_CFL_50	3 655 200	126 831 200	5x5 несимм.	ЛОГОС-Аэро
ONERA_M6_A_0386_CFL_100	9 494 800	329 975 600	5x5 несимм.	ЛОГОС-Аэро
Duck_A_0486_CFL_100	24 473 160	851 654 700	5x5 несимм.	ЛОГОС-Аэро
49_750	2 615 169	190 101 409	симм.пол.опр.	ЛОГОС-Проч
p4_6	4 216 212	228 406 070	симм.пол.опр.	ЛОГОС-Проч
18mln	18 063 492	1 335 851 712	симм.пол.опр.	ЛОГОС-Проч
Balt12	12 140 928	314 737 280	симм.пол.опр.	НИМФА
32mln	32 022 528	815 096 620	симм.пол.опр.	НИМФА

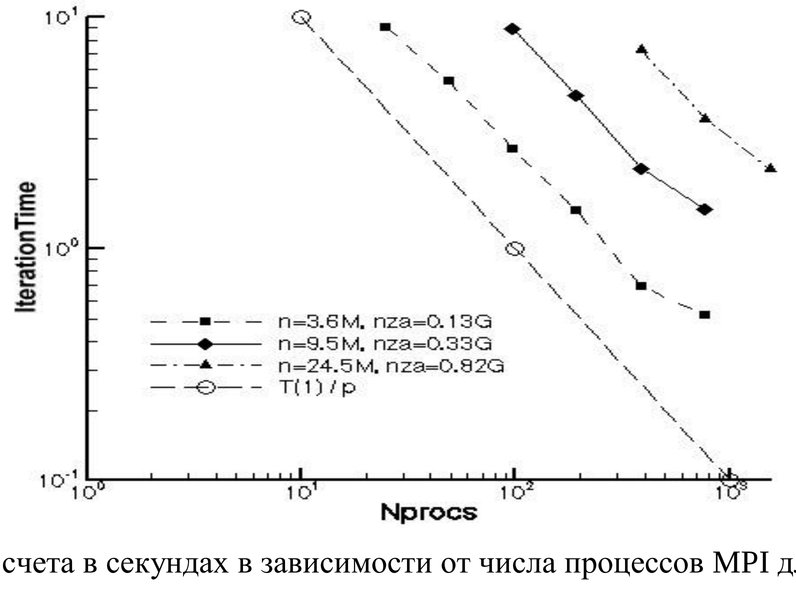

Рис. 1. Время счета в секундах в зависимости от числа процессов MPI для СЛАУ с несимметричными матрицами

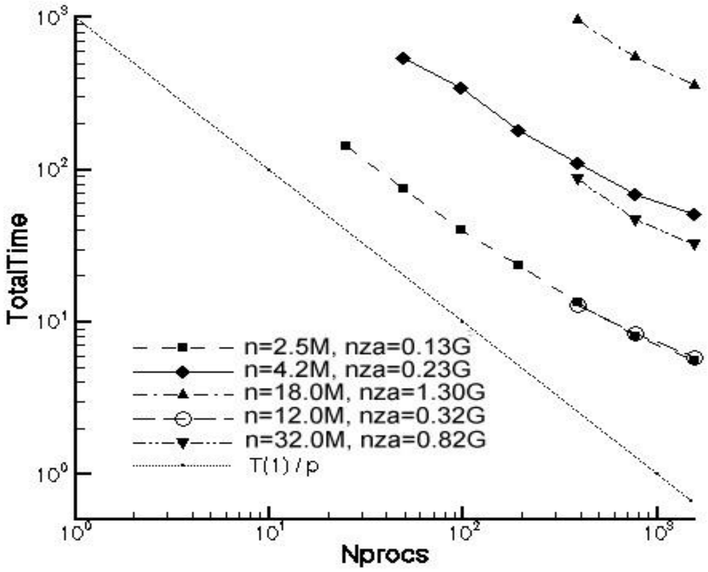

Рис. 2. Время счета в секундах в зависимости от числа процессов MPI для СЛАУ с симметричными положительно определенными матрицами

Результаты тестирования алгоритма п.3 на задачах большого размера, предоставленных РФЯЦ-ВНИИЭФ (см. табл. 2), возникающих при выполнении модального анализа (отыскание 100 наименьших собственных значенияй и соответствующих собственных векторов) программным комплексом ЛОГОС Прочность [13], представлены на Рис.3.

Таблица 2. Структурные характеристики тестовых матриц спектральных задач *(nzu(A)=(nz(A)+n)/2).*

Матрица	***n***	***nzu(A)***	***nzu(B)***	Источник
900000	921 600	32 645 470	11 213 568	ЛОГОС-Проч
49_750	2 595 505	96 540 331	33 076 850	ЛОГОС-Проч
p4_6	4 212 000	144 623 273	49 639 290	ЛОГОС-Проч
18mln	17 999 304	676 893 414	231 630 906	ЛОГОС-Проч

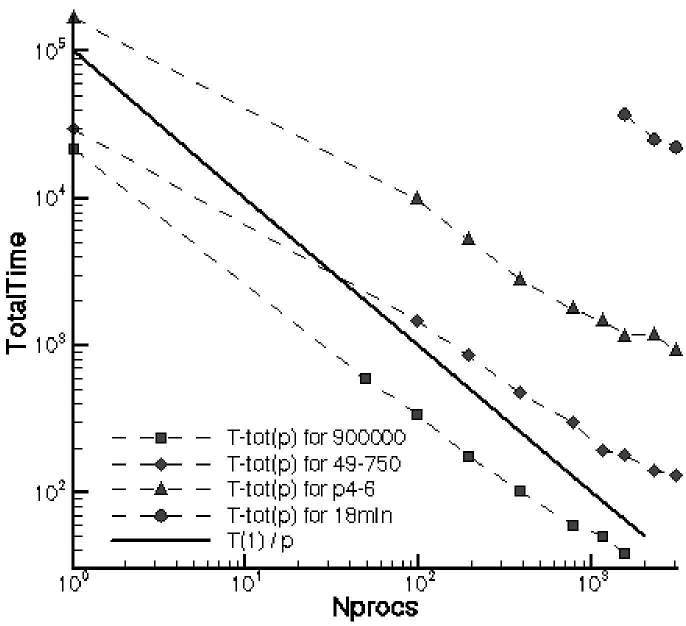

Рис. 3. Время счета в секундах в зависимости от числа процессов MPI для отыскания 100 младших собственных пар спектральных задач с симметричными положительно определенными матрицами.

4. Выводы и направления развития

Использование предлагаемого массивно-параллельного предобусловливания для ускорения итераций методов сопряженных градиентов или BiCGStab оказалось эффективным для решения СЛАУ и частичных обобщенных спектральных задач размера до $n = 3.2 \cdot 10^7$ и числом ненулевых элементов матриц до $nz = 1.3 \cdot 10^9$, возникающих в расчетах по пакетам ЛОГОС и НИМФА. При этом наблюдалась достаточно хорошая параллельная масштабируемость при использовании до 3000 процессов MPI. Производительность построенных алгоритмов предобусловливания может быть улучшена за счет более сбалансированного разбиения матрицы на перекрывающиеся блоки, а также на пути разработки способов предварительной подготовки матриц, например, посредством двустороннего мелкоблочно-диагонального масштабирования. Дальнейшее повышение эффективности решателей СЛАУ (особенно при использовании итераций GMRes) может быть достигнуто за счет применения адаптивного полиномиального предобусловливания. Существенное улучшение надежности и быстродействия спектрального решателя ожидается за счет доработки этапа решения плотных спектральных задач малого размера и использования дополнительных ресурсов параллелизма, связанных с наличием блочной структуры итерационного метода.

Авторы выражают благодарность В.А.Ерзунову, А.Н.Стаканову и Е.Б.Щаниковой за помощь в разработке, тестировании параллельных алгоритмов решения СЛАУ и спектральных задач большого размера и за интеграцию библиотек VC_RAN_SLAU и VC_RAN_EIG в комплекс библиотек параллельных решателей СЛАУ LParSol [15] ВНИИЭФ для “промышленного” применения.

Работа выполнялась по договору между ВЦ РАН и ФГУП “РФЯЦ-ВНИИЭФ”, а также поддерживалась Целевыми программами П-15 и П-18 президиума РАН, грантом РФФИ 11-01-00786, грантом НШ-5264.2012.1 и программой ОМН №3 РАН.

Литература

1. Kaporin I.E., Konshin I.N. A parallel block overlap preconditioning with inexact submatrix inversion for linear elasticity problems // Numer. Linear Algebra Appls., 2002, Vol.9, P.141-162.
2. Капорин И.Е., Милюкова О.Ю. Массивно-параллельный алгоритм предобусловленного метода сопряженных градиентов для численного решения систем линейных алгебраических уравнений // Сб. трудов отдела проблем прикладной оптимизации ВЦ РАН (под ред. В.Г.Жадана), М., изд-во ВЦ РАН, 2011. С.32-49.
3. Kaporin I.E. New convergence results and preconditioning strategies for the conjugate gradient method // Numer. Linear Algebra Appls., V.1, N.2, 1994, P.179-210.
4. Kaporin I.E. High quality preconditioning of a general symmetric positive matrix based on its $U^T U + U^T R + R^T U$ -decomposition // Numer. Linear Algebra Appls., 1998, Vol.5, P.484-509.
5. Милюкова О.Ю. Некоторые параллельные итерационные методы с факторизованными матрицами предобусловливания для решения эллиптических уравнений на треугольных сетках // Ж. Вычисл. Матем. и Матем. Физики. 2006. Т.46. № 6. С.1096-1112.
6. Knyazev A.V. Toward the optimal preconditioned eigensolver: locally optimal block preconditioned conjugate gradient method // SIAM J. Sci. Comput. 2001. V.23. N.2. P.517-541.
7. Капорин И.Е. Использование полиномов Чебышева и приближенного обратного треугольного разложения для предобусловливания метода сопряженных градиентов // Журн. Вычисл. Матем. и Матем. Физ. 2012. Т. 52. №2. С.179-204.

8. Slapničar I., Hari V. On the quadratic convergence of the Falk-Langemeyer method // SIAM Journal on Matrix Analysis and Applications, 1991. V.11. N.1. P.84-114.

9. Saad Y. Multilevel ILU with reorderings for diagonal dominance // SIAM J. Sci. Comput. 2005, Vol. 27, P.1032-1057.

10. Kaporin I. Multilevel ILU preconditionings for general unsymmetric matrices // Proc.Int.Conf. NUMGRID/VORONOI-2008, Moscow, 10-13 June 2008, P.150-157.

11. Капорин И.Е., Милюкова О.Ю. Предобусловливание итерационных методов для эффективного массивно-параллельного решения систем линейных алгебраических уравнений // Труды XIII Международного семинара "Супервычисления и математическое моделирование", 3-7 октября 2011г., г.Саров, 2012. – с.243-252.

12. Козелков А.С., Дерюгин Ю.Н., Зеленский Д.К. и др. Многофункциональный пакет программ ЛОГОС для расчета задач гидродинамики и тепломассопереноса на суперЭВМ // Тезисы докладов на XIV Международном семинаре "Супервычисления и математическое моделирование", 1-5 октября 2012г., г.Саров, 2012. – с.108.

13. Циберев К.В., Авдеев П.А., Артамонов М.В. и др. Пакет программ ЛОГОС. Обзор текущих возможностей решения задач прочности. // Тезисы докладов на XIV Международном семинаре "Супервычисления и математическое моделирование", 1-5 октября 2012г., г.Саров, 2012. – с.159-161.

14. Алейников А.Ю., Бардина М.Н., Горев В.В. и др. Параллельная версия комплекса программ НИМФА // Тезисы докладов на XIII Международном семинаре "Супервычисления и математическое моделирование", 3-7 октября 2011г., г.Саров, 2011. – с.26-27.

15. Бартенев Ю.Г., Бондаренко Ю.А., Ерзунов В.А. и др. Комплекс LParSol для решения СЛАУ// Тезисы докладов на XIII Международном семинаре "Супервычисления и математическое моделирование", 3-7 октября 2011г., г.Саров, 2011. – с.34-36.