УДК 681.513.8

ПАВЛОВ А.В.,
ПАВЛОВ В.А.

МЕТОДИКА ЭКСПЕРИМЕНТАЛЬНЫХ ИССЛЕДОВАНИЙ СХОДИМОСТИ ИТЕРАЦИОННЫХ АЛГОРИТМОВ МЕТОДА ГРУППОВОГО УЧЁТА АРГУМЕНТОВ

В работе предложена новая методика численного исследования сходимости итерационных алгоритмов. Основываясь на методике, экспериментально исследована скорость сходимости Обобщенного Релаксационного Итерационного Алгоритма (ОРИА).

The paper offers a new numerical investigation method for iterative algorithms convergence study. The method has been applied for convergence rate numerical investigation of generalized relaxational iterative algorithm of GMDH.

Введение

Данная работа посвящена описанию методики численного исследования сходимости. Экспериментально показана сходимость к решению и по структуре модели [1]. Поскольку любой алгоритм из ОРИА даёт один и тот же результат оценки параметров модели, проведем исследование для РИА, генерирующего Полное Дерево Структур (ПДС) [2].

Численное исследование скорости сходимости итерационных алгоритмов

Скорость сходимости итерационного алгоритма определяется количеством итераций r^*, необходимых для получения решения с заданной точностью ε. Однако, при проведении численных экспериментов, как правило, r^* существенно зависит от свойств (характеристик) данных, подаваемых на вход алгоритма. Такими характеристиками могут быть коррелированность столбцов или строк входной матрицы, число её обусловленности и др. Поэтому к ряду исследуемых параметров алгоритма (например, свобода выбора F на каждой итерации, количество и вид одночленов модели и т.д.) следует до бавить характеристику данных, влияющую на скорость его сходимости. Не ограничивая общ ности, пусть это будет одна характеристика Pr.

Для осуществления численных экспериментов используют Процедуру Генерации Матрицы (ПГМ), в основе которого лежит Генератор Псевдослучайных Чисел (ГПЧ). В экспериентах необходимо генерировать матрицы со свойствами, не влияющими на скорость сходимости алгоритма. Однако даже при настройке ПГМ на генерацию матрицы с определённым значением параметра Pr, этот параметр является случайной величиной. Поэтому необходимо настроить ПГМ так, чтоб математическое ожидание параметра Pr генерируемых матриц было близко к заданному значению. Для получения объективных результатов сходимости алгоритма при разных значениях его параметров и параметра Pr, предлагается следующая методика проведения экспериментальных исследований.

Методика проведения экспериментальных исследований

Идея состоит в осуществлении серии генераций входных матриц и получении гистограммы распределения значений параметра r^* при заданных значениях параметров алгоритма и параметра Pr. Для определения необходимого числа генераций матриц (количества реализаций) RN^* задаётся значение экспериментальной погрешности δ в процентах.

Методика численного исследования сходимости алгоритма моделирования состоит из трёх этапов:

Этап 1. Определить параметры ПГМ и их значения, позволяющие генерировать матрицы с математическим ожиданием параметра Pr, близким к заданному значению.

Этап 2. Исследовать скорость сходимости алгоритма (число итераций, необходимых для получения решения с заданной точностью) при изменении значения параметра Pr и неизменных параметрах алгоритма.

Этап 3. Для каждого из параметров алгоритма исследовать скорость сходимости при из менении его значения и неизменных значениях остальных параметрах (включая параметр Pr).

На каждом из этапов выполняется алгоритм *Explorer*, входным параметром которого является исследуемый параметр $par \in \{Pr, F, \ldots\}$.

Алгоритм *Explorer*

Шаг 1. Определить количество реализаций RN^*, удовлетворяющее заданному значению δ для гистограмм распределения параметра r^* при заданном значении параметра *par*.

1.1 Сгенерировать матрицу с помощью ПГМ.

1.2 Выполнить исследуемый алгоритм построения модели, подав на его вход сгенерированную матрицу.

1.3 Добавить полученное значение r^* в гистограмму.

1.4 Выполнить п.п. 1.1-1.3 заданное число раз *RN*.

1.5 Выполнить п. 1.4 для разных значений *RN*, увеличивая его до тех пор, пока мера отличия гистограмм для двух последовательных значений *RN* не станет удовлетворять заданной погрешности δ.

Шаг 2. Построить гистограммы распределения параметра r^* используя найденное количество реализаций RN^* для разных значений параметра *par*.

На первом этапе методики входным параметром алгоритма *Explorer* является параметр *Pr*. При этом в алгоритм *Explorer* добавляется третий шаг:

Шаг 3. Выполняется поиск значений параметров АГВМ, при которых математическое ожидание параметра *Pr* генерируемой матрицы близко к заданному значению.

Численное исследование скорости сходимости РИА ПДС

В экспериментах строятся линейные по аргументам модели. Будем исследовать сходимость алгоритма при условии, что матрица X_A содержит только истинные аргументы. Для описания ПГМ введём обозначения: W – выходная матрица ПГМ, $W = (X_A \vdots y)$, $\dim W = n \times (s_{\text{lin}} + 1)$; n – число наблюдений; s_{lin} – количество линейных истин-ных аргументов. Входными параметрами ПГМ являются: n, s_{lin}.

Процедура генерации матриц

1. Сгенерировать матрицу X_A, $\dim X_A = n \times s_{\text{lin}}$.

2. Для аргументов матрицы X_A сгенерировать вектор коэффициентов Θ со значениями в интервале $[a; b]$.

3. Рассчитать вектор выхода сгенерированной модели.

В ПГМ используется один из широко известных ГПЧ – Mersenne Twister MT19937 [3]. Период генератора равен 2^{19937}. Случайные числа генерируются по равномерному закону распределения на интервале [0; 1]. Применим методику для исследования скорости сходимости РИА ПДС.

Этап 1. Можно показать, что скорость сходимости РИА ПДС зависит от такой характеристики матрицы X_A, как мера ортогональности её вектор-столбцов. Сходимость РИА ПДС отслеживается по разности нормированной остаточной суммы квадратов $NRSS_A$ на двух соседних итерациях: $NRSS_{A,r} - NRSS_{A,r+1} = \Delta_{r+1}$. Останов алгоритма осуществляется при условии $\Delta_{r+1} < \varepsilon$. Формула для расчёта $NRSS_A$ имеет вид:

$$NRSS_A = \sum_{i=1}^{n} (\tilde{y}_{A,i} - \tilde{\hat{y}}_{A,i})^2 \Big/ \sum_{i=1}^{n} \tilde{y}_{A,i}^2 ,$$

где $\tilde{y}_A$, $\tilde{\hat{y}}_A$ – центрированные на выборке A значения векторов исходного и моделируемого выхода y и $\hat{y}$.

Несложно показать, что если вектор-столбцы матрицы X_A образуют ортонормированную систему, корреляционная матрица $\Sigma_X = X_A^T X_A$ является единичной с детерминантом равным 1. Поэтому в качестве меры ортогональности вектор-столбцов матрицы X_A выбрано значение детерминант d корреляционной матрицы Σ_X. Поскольку для меры ортогональности не важен знак парной корреляции между вектор-столбцами матрицы X_A, матрица Σ_X содержит модули значений. Следовательно, $d \in [0; 1]$, причём, если $d = 0$, матрица Σ_X – вырожденная, а при $d = 1$, она – единичная.

Определим параметры ПГМ, влияющие на значение детерминанта d матрицы корреляций. Как показывает анализ результатов генерации двух случайных вектор-столбцов в ПГМ, они обладают следующим свойством: чем больше число точек n в векторах, тем меньше значение их парной корреляции. Это объясняется тем, что чем больше случайных элементов в векторах, тем более они ортогональны между собой в n-мерном пространстве точек. Если это свойство обобщить на всю матрицу Σ_X, то число наблюдений можно использовать для генерации матрицы X_A с заданным значением детерминанта её корреляционной матрицы.

Подтвердим этот вывод экспериментально. Исследуем, как изменяется гистограмма рас-

пределения r^* и соответствующее значение усреднённого детерминанта d_{aver} при изменении параметра n. В соответствии с методикой вначале необходимо определить количество реализаций RN^* из некоторого множества значений.

Гистограммы строились для следующих значений $RN \in \psi$, $\psi = \{10^3, 10^4, 10^5, 10^6\}$. Частота (вероятность), указанная на рисунках ниже – это отношение количества матриц, при которых алгоритм находит модель с заданной точностью ε за количество итераций из соответствующего интервала r^* к общему количеству матриц, равному RN.

Обычно в качестве меры отличия гистограмм используют статистический крите-рий Пирсона χ^2 распределения параметра r^*. При этом минимальной мере отклонения соответствует оптимальное значение RN^*. Однако, учитывая известный произвол при выборе допустимой величины уровня значимости для χ^2-критерия, и необходимость разработки простой процедуры для автоматической многократной проверки гистограмм предлагается следующая мера M.

Процедура вычисления меры отличия гистограмм

Пусть количество интервалов параметра r^* гистограмм равно $IntNum$. Мера M_i, (i = 1, 2, 3) пары соседних значений из множества ψ рассчитывается следующим образом:

1. Для каждого интервала значений r^* гистограмм меньшее значение частоты делится на большее. Получаем $IntNum$ значений точности $\gamma_j \in [0; 1]$, $j=\overline{1, IntNum}$.

2. $M_i = \frac{1}{IntNum}\sum_j \gamma_j$.

Выполнив данную процедуру для всех пар, получаем множество значений мер отличия гистограмм: {82.1%, 91.5%, 98%}, которые можно интерпретировать как точность моделирования.

Гистограммы для последней пары показаны на рис. 5. Как видно из рисунка, если $RN = 10^6$ взять за «эталон» полученная погрешность $\tilde{\delta}$ = =2% (M_3 = 98%) удовлетворяет значению заданной погрешности δ = 5%, то значение RN^*= 10^5. Именно это значение используется в дальнейших экспериментах, поскольку, как было установлено в численных экспериментах для ряда исследуемых значений параметров алгоритма моделирования, на первом шаге выполнения алгоритма *Explorer* были получены аналогичные гистограммы.

В данном эксперименте строилась модель:

$$\breve{y} = 6.29447 + 8.11584\cdot x_1 - 7.46026\cdot x_2 - 7.29046\cdot x_3 + 6.70017\cdot x_4 + 9.37736\cdot x_5. \quad (5)$$

Свобода выбора алгоритма F = 5, ошибка моделирования $\varepsilon = 10^{-12}$. Гистограммы r^* при изменении количества наблюдений n показаны на рис. 6. Над столбиками гистограмм представлены тысячные доли значений усреднённого детерминанта d_{aver}.

Усредненное значение детерминанта рассчитывается следующим образом. Вычисляются значения детерминантов всех матриц, для которых алгоритм сошёлся к решению за r^* итераций, попадающих в соответствующий интервал для r^*. Значение d_{aver} есть среднее от рассчитанных значений.

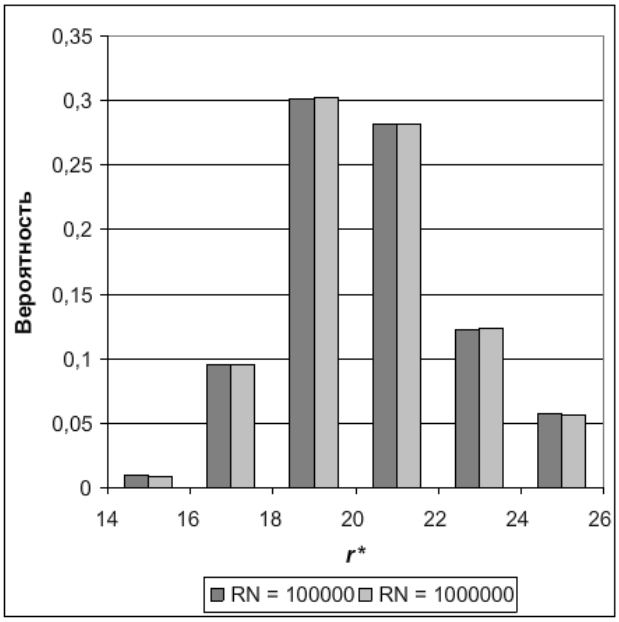

Рис. 5. Гистограмма распределения параметра r^* для последней пары значений из множества RN

Анализируя рис. 6, можно сделать выводы:

1. Параметр n прямо пропорционально влияет на значение детерминанта d, а значит, и на меру ортогональности векторстолбцов матрицы X_A. Следовательно, количество наблюдений можно использовать для получения матрицы корреляций с заданным значением детерминанта.

2. Чем больше значение n, тем меньше дисперсия параметра d генерируемых матриц. Этим объясняется увеличение вероятности генерации матриц с заданным значением d

3. Скорость сходимости алгоритма возрастает пропорционально увеличению значения параметра *d*.

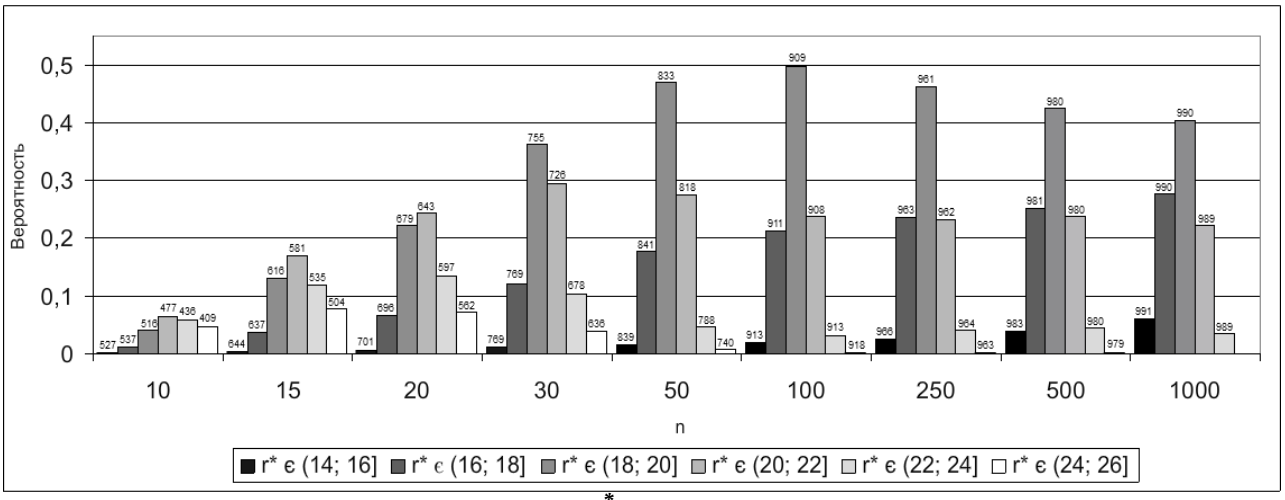

Рис. 6. Гистограммы распределения r^* при изменении количества наблюдений n

Для модели (5) были определены значения *n*, позволяющие генерировать матрицы X_A с заданным значением детерминанта $d \in D$, $D = \{0.5, 0.6, 0.7, 0.8, 0.9, 0.99\}$, см. табл. 1.

Табл. 1. Количество точек и соответствующее значение детерминанта

d	0.5	0.6	0.7	0.8	0.9	0.99
n	12	16	25	41	94	875

Этап 2. Исследуем, как скорость сходимости алгоритма зависит от параметра *d*.

В эксперименте определялась модель (5). Гистограммы распределения параметра r^* для $d \in D$ показаны на рис. 7. Как видно на рисунке, между значением детерминанта и скоростью сходимости алгоритма имеется прямая пропорциональная зависимость.

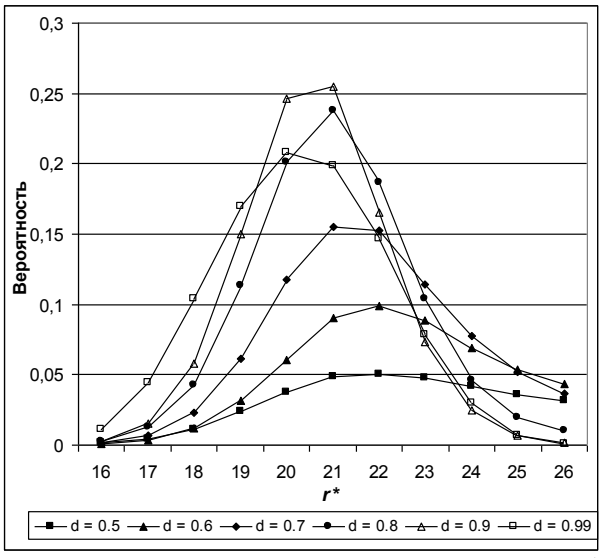

Рис. 7. Гистограммы распределения r^* при разном значении детерминанта

Этап 3. Исследования осуществим для модели (5), варьируя параметром свободы выбора *F* при $d = 0.7$, из табл. 1 соответствующее $n = 25$. Результаты представлены на рис. 8. Из анализа рисунка заключаем, что свобода выбора прямо пропорционально влияет на скорость сходимости алгоритма.

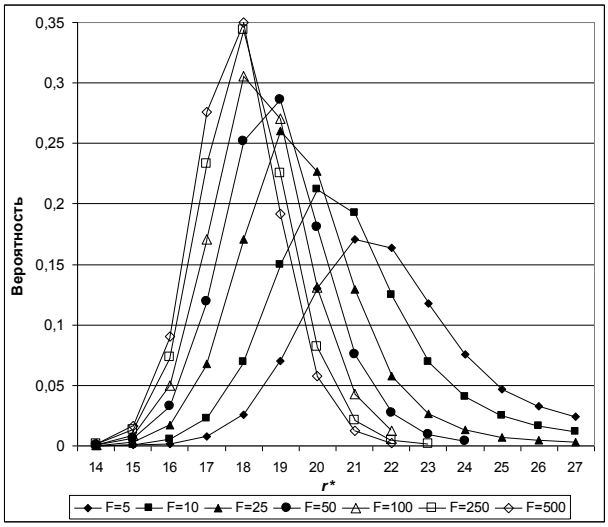

Рис. 8. Гистограммы распределения r^* при изменении параметра F

Исследуем, как изменяется значение критерия $NRSS_A$ на каждой итерации с изменением параметра *F*. В данном эксперименте усреднялись модели, отвечающие пикам гистограмм рисунка 8. Пиковая частота соответствует максимальной частоте на гистограмме. Усреднение осуществлялось по коэффициентам и критерию моделей. Значение r^*, соответствующее пику обозначим через r_{prob}. Строились семейства усреднённых лучших моделей на итерациях с номерами 5, 7, 9, 11, 13, 15, 17 для каждой пары (F, r_{prob}). Результаты в логарифмической шкале представлены на рис. 9

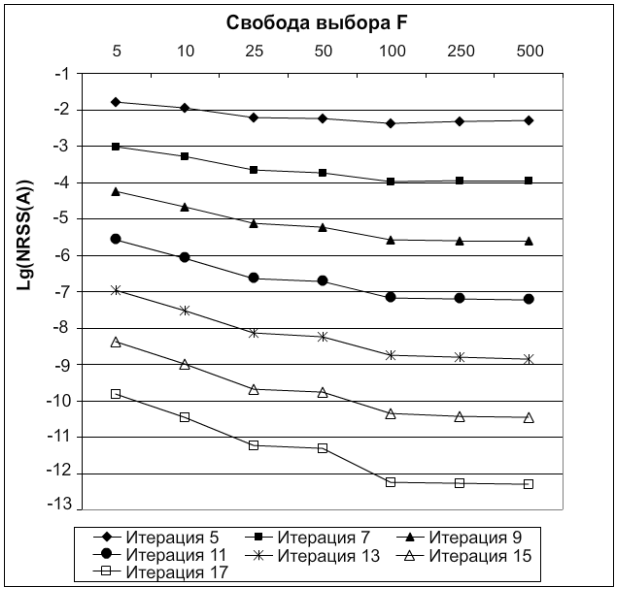

Рис. 9. Изменения минимума критерия $NRSS_A$ при изменении параметра F для выбранных номеров итераций

Анализ результатов, представленных на рис. 9, приводит к следующим выводам:

1. Свобода выбора существенно влияет на изменение минимума критерия $NRSS_A$, причём, чем больше номер итерации, тем это влияние сильнее.

2. Существует значение F, при превышении которого, влияние этого параметра значительно снижается. В данном примере это значение равно 100.

В табл. 3 проиллюстрирована сходимость по структуре и параметрам лучшей усреднённой модели для варианта $F = 50$. В ячейках таблицы содержатся коэффициенты соответствующих аргументов, $\hat{\theta}_0$ – оценка свободного члена модели.

Как видим из таблицы 3 алгоритм обладает достаточно быстрой сходимостью: для получения решения с погрешностью $\varepsilon = 10^{-12}$ ему необходимо осуществить $3s_{lin}$ итераций.

Выводы

С целью исследования скорости сходимости ОРИА предложена новая методика численного исследования итерационных алгоритмов. Основываясь на методике, экспериментально исследована скорость сходимость РИА ПДС. При этом выявлены следующие особенности:

– количество наблюдений можно использовать для генерации входной матрицы с заданым значением детерминанта её корреляционной матрицы;

– значение детерминанта (характеристика исходных данных) и свобода выбора (характеристика алгоритма) связаны со скоростью сходимости к точному решению прямо пропорционально;

– алгоритм обладает быстрой сходимостью при построении линейных моделей.

Табл. 3. Результаты сходимости РИА ПДС к решению

№ итерации	$NRSS_A$	$\hat{\theta}_0$	x_3	x_1	x_4	x_2	x_5
5	0,005473	6.35397	-7.16523	7.99056	6.57297	-7.46026	9.37736
7	0,000184	6.29582	-7.28527	8.11217	6.69561	-7.33251	9.25849
9	$5{,}82\cdot10^{-6}$	6.29464	-7.2905	8.11572	6.69998	-7.45551	9.37293
11	$1{,}8\cdot10^{-7}$	6.2945	-7.29044	8.11584	6.70012	-7.46026	9.37737
13	$5{,}61\cdot10^{-9}$	6.29448	-7.29046	8.11583	6.70017	-7.46028	9.37735
15	$1{,}69\cdot10^{-10}$	6.29447	-7.29046	8.11584	6.70017	-7.46026	9.37736
Истинная модель:		6.29447	-7.29046	8.11584	6.70017	-7.46026	9.37736

Список литературы

1. Юрачковский Ю.П. Сходимость многорядных алгоритмов МГУА // Автоматика, 1981. – № 3. – С. 36-43.
2. Павлов А.В. Обобщённый релаксационный итерационный алгоритм МГУА // Індуктивне моделювання складних систем. Збірник наук. праць. – К.: МННЦІТС, 2011. – С. 130-143.
3. Доступно на сайте http://www.boost.org/.