

УДК 681.513.8

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ВЫЧИСЛИТЕЛЬНОЙ СЛОЖНОСТИ АЛГОРИТМОВ РЕЛАКСАЦИОННО-ИТЕРАЦИОННОГО ТИПА

Н.В. Кондрашова

*Международный научно-учебный центр информационных технологий и систем
НАНУ и МОН Украины,**NKondrashova@ukr.net*

Незважаючи на те, що узагальнений релаксаційний ітераційний алгоритм (УРІА) на сьогодні найшвидший і точний ітераційний алгоритм МГУА, для якого доведена збіжність, його аналоги: багаторядний спрощений алгоритм (БСА) і багаторядний алгоритм з комбінаторикою і селекцією узагальнених змінних (БАКСУ) також мають свою «нішу» застосовності. У площині двох параметрів: розміру вибірки (числа спостережень) і складності моделі (числа аргументів) показані області перевищення обчислювальної складності (швидкодії) одного алгоритму по відношенню до іншого. Проведено порівняльний аналіз швидкодії нерекуррентного та рекуррентних варіантів УРІА між собою і кожного з них у порівнянні з БСА.

Ключові слова: метод групового урахування аргументів (МГУА), узагальнений релаксаційний ітераційний алгоритм (УРІА), багаторядний спрощений алгоритм (БСА), багаторядний алгоритм з комбінаторикою і селекцією узагальнених змінних (БАКСУ), рекуррентні обчислення.

Despite the fact that the generalized relaxation iterative algorithm (GRIA) for today is most fast and precise iterative algorithm GMDH for which the convergence is proved, its analogues: a multi-layered simplified algorithm (MSA) and multi-layered algorithm with combinatorics and selection of generalized variables (MACSG) also have their "niche" applicability. Areas of exceedance computational complexity (running speed) of an algorithm with respect to another in the plane of two parameters such as sample size (number of observations) and model complexity (number of arguments) are shown. A comparative analysis of the running speed of nonrecurrent and recurrent variants GRIA between themselves and each of them compared to the MSA is provided.

Keywords: Group Method of Data Handling (GMDH), Generalized Relaxation Iterative Algorithm (GRIA), Multilayered Simplified Algorithm (MSA), Multilayered Algorithm with Combinatorics and Selection of Generalized variables (MACSG), recurrent calculation.

Несмотря на то, что обобщенный релаксационный итерационный алгоритм (ОРИА) на сегодня самый быстрый и точный итерационный алгоритм МГУА, для которого доказана сходимость, его аналоги: многорядный упрощенный алгоритм (МУА) и многорядный алгоритм с комбінаторикой и селекцией обобщенных переменных (МАКСО) также имеют свою «нишу» применимости. В плоскости двух параметров: размера выборки (числа наблюдений) и сложности модели (числа аргументов) показаны области превышения вычислительной сложности (быстродействия) одного алгоритма по отношению к другому. Проведен сравнительный анализ быстродействия нерекуррентного и рекуррентных вариантов ОРИА между собой и каждого из них в сравнении с МУА.

Ключевые слова: метод группового учета аргументов (МГУА), обобщенный релаксационный итерационный алгоритм (ОРИА), многорядный упрощенный алгоритм (МУА), многорядный алгоритм с комбінаторикой и селекцией обобщенных переменных (МАКСО), рекуррентные вычисления.

Вступление

Данная работа выполнена в рамках темы «Разработка и исследование интеллектуальных технологий индуктивного моделирования для задач поддержки решений в сложных системах». Созданная технология построения регрессионных моделей на основе ОРИА МГУА [1] с рекуррентными вычислениями является ее важным результатом. Прообразом этого алгоритма является МУА [2], а предшественником – модификация МУА – МАКСО [3]. Если первый алгоритм – МУА – имеет исторический интерес и теоретическое значение, то МАКСО, имеющий современный интерфейс и расширенные возможности, является широко используемым программным продуктом. Решение практических задач моделирования по выборкам данных даже средней размерности для этого алгоритма является трудной задачей. Например, на получение одной МАКСО-модели по выборке, имеющей около десятка переменных и несколько десятков точек наблюдения, было затрачено несколько суток расчетного времени компьютера с процессором Intel Core i3 M350 2.27 GHz. При этом использовался «веерный» вариант алгоритма, описанный в [4]. МУА и МАКСО по вычислительной сложности являются тождественными, поэтому для сравнения рассмотрим первый многорядный упрощенный алгоритм.

Для решения задач большой размерности с сотнями тысячами наблюдений и тысячами переменных был создан ОРИА. Время решения задач этим алгоритмом по сравнению с аналогом снижается в десятки, сотни, тысячи и т.д. раз в зависимости от числа аргументов и наблюдений. В данной статье исследуется вопрос, насколько по быстродействию отличаются релаксационно-итерационные алгоритмы при небольшом объеме выборки. Целью данной статьи является исследование быстродействия указанных алгоритмов для граничных соотношений объемов обучающей и проверочной выборок, в зависимости от двух параметров: сложности модели и размера выборки.

1. Постановка задачи

Алгоритмы, о которых речь пойдет ниже, предназначены для решения задачи моделирования по таблицам или выборкам числовых данных. В задаче идентификации методом группового учёта аргументов известна матрица измерений исходных данных \mathbf{W} , $\dim \mathbf{W} = n_W \times (m+1)$, m – число входных переменных (число столбцов). Размерность выходного вектора $\dim \mathbf{y} = n_W \times 1$, n_W – количество наблюдений. Неизвестны характеристики шума ξ в объекте. Предполагается, что он аддитивен, некоррелирован с полезными сигналами, имеет нулевое математическое ожидание ($M[\xi] = 0$), диагональную матрицу ковариаций ($\xi \xi^T = \sigma^2 E_{n_W}$) и конечную дисперсию ($\sigma^2 < \infty$). Где E_{n_W} – единичная матрица $n_W \times n_W$. Необходимо найти структуру и параметры θ_i , $i = \overline{0, m}$ модели

$$y = \sum_{i=0}^m \theta_i z_i, \quad (1)$$

удовлетворяющие максимальной точности заданного критерия. Выражение (1) является скалярным видом линейной модели. Под структурой линейной модели объекта понимается информативный набор (множество) аргументов z_i , который его описывает с максимальной точностью заданного критерия. Пусть матрица входов \mathbf{X} имеет размерность $n_W \times m$. Векторный, нелинейный вид модели:

$$\mathbf{y} = f(\Theta, \mathbf{X}). \quad (2)$$

ОРИА МГУА строит линейные по параметрам модели. Выражение (2) путем различных преобразований (линеаризацией, разложением в ряд, логарифмированием и прочее) и последующим переобозначением преобразованных переменных приводится к виду (1). Преобразованные и переобозначенные переменные $z_i = f_i(\mathbf{X})$ называются аргументами. Структурой модели в этом случае называется нелинейная функция входных переменных, заданная с точностью до значений, линейно входящих в неё параметров $\hat{\Theta} = (\hat{\theta}_1, \dots, \hat{\theta}_M)$. Как правило, $M \geq m$. В дальнейшем будем считать, что $M = m$.

Под информативным понимается при отсутствии шума набор «истинных» аргументов, при наличии шума – набор «адекватных» шуму аргументов. «Адекватное» упрощение модели происходит при превышении дисперсии реального шума в объекте значений, так называемых, критических дисперсий [5]. Следующее определение «истинной» структуры является обобщением определения, данного в [6].

Определение 1. Назовем «истинной» структуру f^0 , которая является несмещенной оценкой функции f , зависящей от вектора параметров Θ и матрицы входов \mathbf{X} , если $f^0(\mathbf{X}, M[\hat{\Theta}]) = f(\mathbf{X})$. При этом структура f^0 имеет сложность:

$$s_0 = m - \sum_{i=1}^m Equal(M[\hat{\theta}_i], 0),$$

где $M[\cdot]$ – оператор математического ожидания, которое осуществляется по всевозможным реализациям шума $\xi_i, i = 1, 2, \dots$. Функция сравнения $Equal(M[\hat{\theta}_i], 0)$ равна 1, если математическое ожидание какого-либо параметра $M[\hat{\theta}_i]$ равно 0. Истинная модель при этом имеет вид:

$$y_j = f(x_{j,1}, \dots, x_{j,m}) + \xi_j, \quad j = \overline{1, n_W},$$

где x_{ji} измерены точно. Неважно, какие значения имеют линейные параметры. Полагается, что они не равны нулю и сложность истинной модели равна s_0 .

Определение 2. Назовем «адекватной» идентифицируемую модель, которая упрощается, становится соответствующей неизвестному уровню шума при его увеличении на выходе объекта.

2. Краткие сведения об ОРИА МГУА

ОРИА известен в двух своих модификациях с нерекуррентными и рекуррентными вычислениями оценок параметров модели и значений критериев. При каких условиях их скорость различается на порядки, было доказано и показано в [7]. Резонный вопрос: зачем говорить о двух модификациях, если и так ясно, что модификация с рекуррентными вычислениями должна быть лучше. Покажем, что при малом числе аргументов это не выполняется. Проведем также аналитическое сравнение двух указанных модификаций ОРИА с его аналогом МУА для того, чтобы стал очевидным тот факт, что для безусловной независимости от числа наблюдений переборной стадии недостаточно выделить отдельный блок формирования ковариационных матриц, как это сделано в РИА с нерекуррентным методом (РИАНМ) вычислений параметров и критерия. Ограничимся рассмотрением случая, когда исходная выборка W разбита на две подвыборки: обучающую A и проверочную B , $W = A \cup B$, $A \cap B = \emptyset$.

В работе речь пойдет об одном и том же алгоритме – ОРИА, нерекуррентная и рекуррентная модификации которого совпадают в двух своих следующих стадиях:

1) преобразование матрицы входных переменных в аргументы, для получения моделей заданного класса (например, полиномиальных, линейных или нелинейных разностных моделей);

2) подготовительных вычислений, где осуществляется: преобразование вектор-столбцов матриц аргументов выборок A и B путем вычитания средних значений элементов A выборки, а также расчет «ковариационных» матриц для указанных выборок Σ_A, Σ_B .

Модификации ОРИА отличаются своей самой вычислительно затратной третьей стадией, на которой модели строятся итеративно, с использованием элементов ковариационных матриц указанных выборок. Рассмотрим кратко отличительную третью стадию этих модификаций: одну с нерекуррентными вычислениями и две – с рекуррентными.

3. Релаксационный итерационный алгоритм с использованием нерекуррентного метода

Этот алгоритм является вариантом ОРИА с нерекуррентным методом вычисления параметров и критерия селекции. Алгоритм начинает работу полным перебором из m аргументов по два:

$$\hat{y}_{lA} = \hat{\omega}_l \mathbf{x}_{iA} + \hat{\omega}_l \mathbf{x}_{jA}, \quad l = \overline{1, C_m^2}. \quad (3)$$

Оценки параметров находятся методом наименьших квадратов (МНК):

$$\Omega_\ell = (\mathbf{X}_l^T \mathbf{X}_l)^{-1} \mathbf{X}_l^T \mathbf{y}, \quad (4)$$

из условия минимизации остаточной суммы квадратов RSS_l . При оценке параметров модели (3) $\Omega_\ell = (\widehat{\omega}_l^1, \widehat{\omega}_l^2)^T$, $\dim(\mathbf{X}_l^T \mathbf{X}_l) = 2 \times 2$, $l = \overline{1, C_m^2}$, $\mathbf{X}_l = (\mathbf{x}_i, \mathbf{x}_j)$, $i, j = \overline{1, m}$, $j > i$. На самом деле оценки параметров модели (3) получаются МНК для вложенных структур, сначала для моделей одного аргумента,

$$\widehat{\omega}_l^1 = \frac{\mathbf{x}_{lA}^T \mathbf{y}_{lA}}{\mathbf{x}_{lA}^T \mathbf{x}_{lA}}, \quad \widehat{\mathbf{y}}_{lA}^1 = \widehat{\omega}_l^1 \mathbf{x}_{lA}, \quad l = \overline{1, m},$$

потом для двух аргументов

$$\widehat{\mathbf{y}}_{lA}^2 = \widehat{\omega}_l^2 \widehat{\mathbf{y}}_{lA}^1 + \widehat{\omega}_j^2 \mathbf{x}_{jA}, \quad l = \overline{1, m}, \quad j = \overline{(l+1), m},$$

как результат минимизации $RSS_l^r = (\mathbf{y}_A - \widehat{\mathbf{y}}_{lA}^r)^T (\mathbf{y}_A - \widehat{\mathbf{y}}_{lA}^r) \rightarrow \min$ при $r = 1, 2$. Число перебранных моделей, для которых они вычисляются равно $m(m-1)/2$. Параметры модели определяются только по данным выборки A , поэтому индекс A для параметров опускаем.

Опишем кратко алгоритм при свободе выбора моделей $F=1$ на каждой итерации, для номера итерации $r \geq 2$. Свобода выбора $F > 1$ на факт сходимости алгоритма к «истинной» модели не влияет, а влияет только на скорость сходимости, при условии что в переборе участвуют все «истинные» аргументы и $n \geq m$ [7].

Пусть необходимо найти лучшую линейную модель $(r+1)$ -й итерации:

$$\widehat{\mathbf{y}}_A^{r+1} = \widehat{\omega}^{r+1} \widehat{\mathbf{y}}_A^r + \widehat{\omega}^{r+1} \mathbf{x}_A^{r+1}, \quad r \geq 2. \quad (5)$$

Для этого строятся все возможные модели $(r+1)$ -й итерации вида (5)

$$\widehat{\mathbf{y}}_{l,A}^{r+1} = \widehat{\omega}_l^{r+1} \widehat{\mathbf{y}}_A^r + \widehat{\omega}_l^{r+1} \mathbf{x}_{l,A}^{r+1}, \quad r \geq 2, \quad l = \overline{1, m}. \quad (6)$$

где $\widehat{\omega}_l^{r+1}$, $\widehat{\omega}_j^{r+1}$ – МНК-оценки, получаемые из условия равенства правой части (6) исходному выходному вектору \mathbf{y}_A . Можно показать, что оценки вложенных структур являются точно соответствующими МНК-оценкам (4) $\widehat{\omega}_l^1 \widehat{\omega}_l^2 = \widehat{\omega}_l$, $\widehat{\omega}_j^2 = \widehat{\omega}_j$ только при $r \leq 2$. Если в матрице \mathbf{X}_l больше, чем два столбца, то оценки

параметров при $\mathbf{x}_{jA}, j=1,2,\dots,r$, получаемые по формуле (4), а также итерационно по формулам (6) не равны между собой для одного и того же r .

Для всех моделей (6) по данным проверочной выборки B вычисляется критерий селекции $AR_l^{r+1} = (\mathbf{y}_B - \hat{\mathbf{y}}_{l,B}^{r+1})^T (\mathbf{y}_B - \hat{\mathbf{y}}_{l,B}^{r+1})$. По его минимуму находится лучший аргумент и параметры модели (6) на $(r+1)$ -й итерации, т.е. тройка:

$$(\hat{\omega}^{r+1}, \hat{\omega}_l^{r+1}, \mathbf{x}^{r+1}) = \arg \min_{(\hat{\omega}_l^{r+1}, \hat{\omega}_l^{r+1}) \in \mathfrak{R}^2, l=1,m, \mathbf{y} \in \mathfrak{R}^{nW}, \mathbf{x}_l \in \mathfrak{R}^{nW}} AR_l^{r+1}(\hat{\omega}_l^{r+1}, \hat{\omega}_l^{r+1}, \mathbf{x}_l, \mathbf{y}).$$

После этого получается модель в исходных переменных («развернутая» модель): $\hat{\mathbf{y}}^{r+1} = \mathbf{G}^{r+1} \hat{\Omega}^{r+1|r+1-i}$, где число $i = r, (r-1), \dots, 0$ показывает сколько раз корректировался коэффициент при переменной с индексом $(r+1-i)$. Матрица \mathbf{G}^{r+1} составляется из вектор-столбцов аргументов $(\mathbf{x}^1, \dots, \mathbf{x}^r, \mathbf{x}^{r+1})$ матрицы \mathbf{X} , в порядке их включения в модель. Вектор оценок коэффициентов $(r+1)$ -й итерации «развернутой» модели $\hat{\Omega}^{r+1|r+1-i} = (\hat{\omega}^{r+1|1}, \hat{\omega}^{r+1|2}, \dots, \hat{\omega}^{r+1|r}, \hat{\omega}^{r+1})^T$ – при условии, что на предыдущих итерациях были выбраны аргументы $\mathbf{x}^1, \dots, \mathbf{x}^r$, а на $(r+1)$ -й – аргумент $\mathbf{x}_A^{r+1} = \mathbf{x}^{r+1}$ с параметром $\hat{\omega}^{r+1} \stackrel{\Delta}{=} \hat{\omega}^{r+1|r+1}$. Первая часть верхнего индекса $r+1$ определяет номер итерации. Во второй части – символ $|$ означает: при условии выбора на предыдущих итерациях аргументов, верхний индекс которых равен $(r+1-i)$, $i = r, (r-1), \dots, 0$ и совпадает с порядковым номером их включения в модель.

Для того чтобы оценить количество операций, необходимое для вычисления модели данным алгоритмом, выпишем формулы, по которым вычисляются параметры:

$$\hat{\omega}_l^{r+1} = \frac{c_r ((\mathbf{x}_{l,A}^{r+1})^T \mathbf{x}_{l,A}^{r+1}) - b_{r+1} ((\mathbf{x}_{l,A}^{r+1})^T \mathbf{y}_A)}{a_r ((\mathbf{x}_{l,A}^{r+1})^T \mathbf{x}_{l,A}^{r+1}) - b_{r+1}^2}, \quad \hat{\omega}_l^{r+1} = \frac{a_r ((\mathbf{x}_{l,A}^{r+1})^T \mathbf{y}_A) - b_{r+1} c_r}{a_r ((\mathbf{x}_{l,A}^{r+1})^T \mathbf{x}_{l,A}^{r+1}) - b_{r+1}^2} \quad (7)$$

где коэффициенты a_r, b_{r+1}, c_r равны соответственно:

$$a_r = (\hat{\Omega}^{r|r-i})^T (\mathbf{G}^r)^T \mathbf{G}^r \hat{\Omega}^{r|r-i}, \quad b_{r+1} = (\hat{\mathbf{y}}_A^r)^T \mathbf{x}_{l,A}^{r+1} = (\mathbf{x}_{l,A}^{r+1})^T \mathbf{G}^r \hat{\Omega}^{r|r-i}, \\ c_r = (\hat{\mathbf{y}}_A^r)^T \mathbf{y}_A = (\mathbf{y}_A)^T \mathbf{G}^r \hat{\Omega}^{r|r-i}.$$

Критерий селекции AR_l^{r+1} и критерий RSS^{r+1} вычисляются для $l = \overline{1, m}$ по формулам

$$AR_l^{r+1} = \mathbf{y}_B^T \mathbf{y}_B - 2\mathbf{y}_B^T \mathbf{G}_{l,B}^{r+1} \hat{\Omega}_l^{r+1|r+1-i} + (\hat{\Omega}_l^{r+1|r+1-i})^T (\mathbf{G}_{l,B}^{r+1})^T \mathbf{G}_{l,B}^{r+1} \hat{\Omega}_l^{r+1|r+1-i}, \quad (8)$$

$$RSS^{r+1} = \mathbf{y}_A^T \mathbf{y}_A - 2\mathbf{y}_A^T \mathbf{G}_A^{r+1} \hat{\Omega}^{r+1|r+1-i} + (\hat{\Omega}^{r+1|r+1-i})^T (\mathbf{G}_A^{r+1})^T \mathbf{G}_A^{r+1} \hat{\Omega}^{r+1|r+1-i}. \quad (9)$$

При вычислении вектора $\widehat{\Omega}^{r+1|r+1-i}$ первые r оценок параметров лучшей модели $(r+1)$ -й итерации определяются как: $\widehat{\omega}_l^{r+1} \widehat{\Omega}^{r|r-i}$, $i = (r-1), (r-2), \dots, 0$, а оценка $\widehat{\omega}^{r+1}$ – получена по формуле (7).

Из работы [7] известно, что вычислительная сложность построения одной модели РИАНМ на r -й итерации, имеющей r аргументов, равна:

$$Q_{\text{нерек}}(r) = r^2 + 4r + 9. \quad (10)$$

4. Релаксационный итерационный алгоритм с использованием рекуррентного метода (РИАРМ)

Отличие РИАРМ от РИАНМ состоит в том, что коэффициенты a_r, c_r на r -й итерации для модели (6) вычисляются по рекуррентным формулам общего вида $a_r = f(a_{r-1}, b_r, x_A, \widehat{\omega}_r)$, $c_r = \phi(c_{r-1}, b_r, \widehat{\omega}_r)$.

В первом варианте алгоритма (РИАРМ₁) критерий селекции AR_l^{r+1} и критерий RSS^{r+1} для $l = \overline{1, m}$ вместо формул (8) и (9) вычисляются по формулам, имеющим общий вид: $AR_l^{r+1} = \phi(AR^r, \Omega, \mathbf{G}_B, \mathbf{x}_B, \mathbf{y}_B)$, $RSS^{r+1} = \psi(RSS^r, \Omega, \mathbf{G}_A, \mathbf{x}_A, \mathbf{y}_A)$. Из [7] известно, что число операций рекуррентных алгоритмов линейно зависит от числа аргументов r . В РИАРМ₁ число операций одной итерации

$$Q_{\text{рек}_1} = 4r + 29. \quad (11)$$

Второй вариант рекуррентного алгоритма (РИАРМ₂) является более экономным при вычислении критериев по формулам вида: $AR_l^r = \phi_1(a_{r,B}, b_{r,B}, \mathbf{y}_B)$, $RSS^r = \psi_1(a_{r,A}, b_{r,A}, \mathbf{y}_A)$ и его сложность равна:

$$Q_{\text{рек}_2} = 3r + 26. \quad (12)$$

5. Сравнение нерекуррентного и рекуррентных методов вычислений в ОРИА

Рассмотрение предпочтительности применения нерекуррентных по сравнению с рекуррентными вычислениями с позиции здравого смысла – нонсенс. На примере сравнения двух методов вычисления в ОРИА покажем, что нерекуррентные методы «выигрывают» у рекуррентных, если число аргументов в моделях, которые они строят, меньше некоторого количества, и «проигрывают» – в противном случае.

Результаты расчета числа операций при построении одной модели в зависимости от числа итераций r , количество которых равно числу аргументов в «развернутой» модели, для трех вариантов алгоритмов (одного с нерекуррентными и два с рекуррентными вычислениями) представлены графически на рис. 1. Из сравнения графиков видно, что при m меньше четырех, алгоритмы, не имеющие рекуррентных вычислений оценок параметров модели и значений критериев, имеют меньшее число операций, а при $m \geq 5$, наоборот, оба рекуррентных алгоритма являются более «быстрыми».

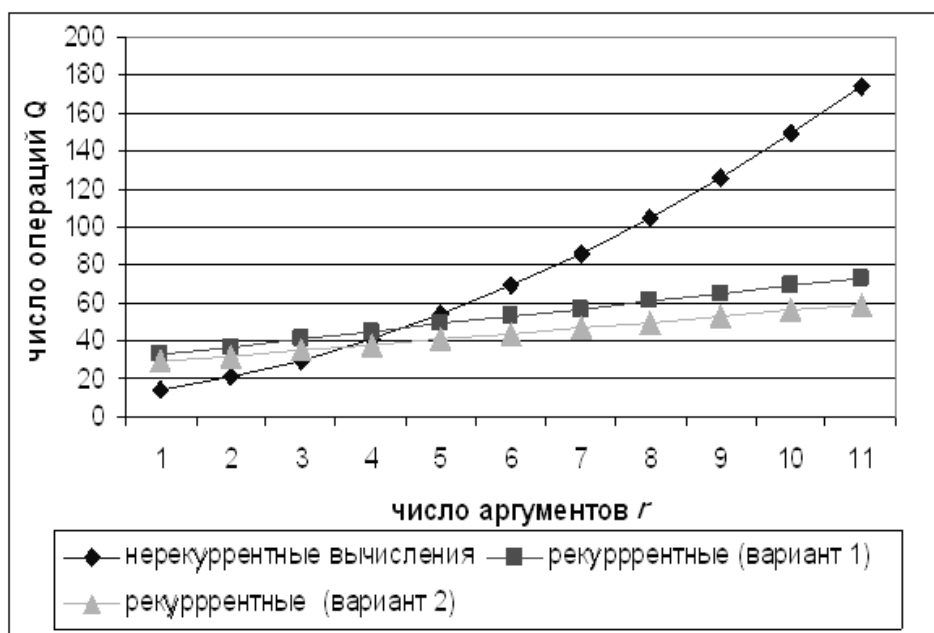


Рис. 1 – Зависимость числа вычислений от числа аргументов (итераций) r .

В целом, результаты, представленные на рис. 1, показывают преимущество второго рекуррентного варианта алгоритма, поскольку число аргументов модели m , при котором нерекуррентный метод является более быстроедействующим по сравнению с обоими рекуррентными, пренебрежимо мало, а малые отличия $Q_{\text{нерек}}$ от $Q_{\text{рек}}$ не могут при решении реальных задач сделать нерекуррентный метод более предпочтительным.

При $m \geq 5$ в зависимости от значений числа аргументов рекуррентные методы быстрее нерекуррентных в 10 раз при $m = 10$, в 100 раз при $m = 100$ и т.д.

6. Сравнительный анализ ОРИА и МУА

Сравним два алгоритма, принадлежащих одному виду релаксационно-итерационных: ОРИА с его прообразом – МУА. Поскольку алгоритм МУА имеет некоторые преимущества (он экономно расходует оперативную память, не хранит ковариационные матрицы Σ_A, Σ_B), важно убедиться, в каких случаях этот алгоритм не хуже в вычислительном плане своего усовершенствованного аналога ОРИА. Для этого сравним число операций,

приходящееся на одну итерацию, этих двух алгоритмов. Известно [7], что число операций при построении МУА одной модели равно

$$Q_{\text{МУА}}(r, n_A, n_B) = n_B r + 7n_A + 2n_B + 3. \quad (13)$$

Для возможности проведения дальнейшего анализа заданных зависимостей (10)-(13) удобно перейти к переменным в области действительных значений, обозначив их $\eta, \rho \in \mathfrak{R}$. Построив известную зависимость в непрерывной области, нужно проверить решение неравенств на корректность в области целочисленных значений. Натуральные числа $n_B, r \in \mathbb{N}$ будем получать в результате округления действительных $\eta, \rho \in \mathfrak{R}$, как $n_B = \lfloor \eta \rfloor$ и $r = \lfloor \rho \rfloor$.

Рассмотрим три основных варианта соотношений при разбиении выборки. Если используется критерий регулярности, то традиционное разбиение равно $n_A : n_B = 2:1$, если усредненный критерий регулярности (УКР) (англ. «leave one out»), то $n_A : n_B = (n_W - 1):1$; если применяется какой либо из критериев несмещенности (согласованности), то выборка разбивается в отношении $n_A : n_B = 1:1$. В задаче прогнозирования с квазиоптимальным разбиением значение отношения $n_A : n_B$, на практике изменяется от 1:1 до 2:1 [8]. Разбиения $n_A : n_B = (n_W - 1):1$ и $n_A : n_B = 1:1$ это – два варианта соотношений размеров подвыборок, которые являются наименьшим и наибольшим по трудоемкости перебора. При поиске лучшего разбиения по УКР осуществляется полный перебор (ПП), имеющий линейную вычислительную сложность по n_W . При $n_A : n_B = 1:1$ ПП, имеющий экспоненциальную сложность $L = 2^{n_W - 1} - 1$, часто заменяется разбиением «по дисперсии» [9]. Традиционно применяемое разбиение $n_A : n_B = 2:1$ является соотношением объемов обучающей и проверочной выборок, полученным опытным путем. Вычислительная сложность поиска лучшего разбиения в данной работе вынесена за рамки исследования. Лучшее разбиение считается полученным, при этом отношение $n_A : n_B$ может быть разным. Рассмотрим случаи А), Б) и В).

△

Формулы (10)-(12) не изменяют своего вида при подстановке $r = \rho$.

А) Если $n_A : n_B = 2:1$, то в соответствии с новыми обозначениями выражение (13) примет вид:

$$Q_{\text{МУА}}(\rho, \eta) = \rho \eta + 16\eta + 3. \quad (14)$$

Б) Если $n_A : n_B = (n_W - 1):1$, то (13) имеет вид:

$$Q_{\text{МУА}}(\rho, \eta) = \rho + 7(n_W - 1) + 5. \quad (15)$$

В) Если $n_A : n_B = 1:1$, то формула (13) выглядит следующим образом

$$Q_{\text{МУА}}(\rho, \eta) = \rho \eta + 9\eta + 3. \quad (16)$$

Рассмотрим, при каких условиях число вычислений МУА, больше или равно числу вычислений ОРИА

$$Q_{\text{МУА}} \geq Q_{\text{ОРИА}}. \quad (17)$$

7. Сравнительный анализ вычислительной сложности РИАНМ и МУА

Сравним МУА сначала с модификацией ОРИА с нерекуррентным методом вычислений (РИАНМ) сложности (10).

Рассмотрим случай А), когда отношение $n_A : n_B = 2:1$. Подставив в (17) выражения (14) и (10) получим нестрогое неравенство для анализа в области действительных переменных

$$\rho^2 + (4 - \eta)\rho - 16\eta + 6 \leq 0. \quad (18)$$

Найдем корни уравнения (18) ρ_1, ρ_2 .

$$\rho_{1,2} = -(4 - \eta)/2 \pm \sqrt{(16 - 8\eta + \eta^2)/4 + 16\eta - 6} = -(4 - \eta)/2 \pm 0.5\sqrt{\eta^2 + 56\eta - 8}. \quad (19)$$

В результате имеем соотношение, связывающее непрерывные параметры, когда $\eta, \rho \in \mathfrak{R}$:

$$\begin{aligned} \rho_1 &= -(4 - \eta)/2 - 0.5\sqrt{(\eta + 56)\eta - 8}, \\ \rho_2 &= -(4 - \eta)/2 + 0.5\sqrt{(\eta + 56)\eta - 8} \end{aligned} \quad (20)$$

При любых $\eta \in \mathfrak{R}$, $n_B = \lfloor \eta \rfloor$ значения $\rho_1 < 0$, поэтому это решение исключается из рассмотрения. Исследуем при каких значениях n_B , (n_W) и $r = \lfloor \rho \rfloor$ выполняется неравенство (18), и в какой из частей области изменения значений, которую пересекает решение (20), это неравенство не выполняется. Последовательно подставляя значения $\eta = n_B = 1, 2, 3, \dots$ в (20), получим ряд значений ρ_2 , округляя которые с недостатком (ниже кривой), получим выполнение неравенства (17), а с избытком (выше кривой) – выполнение противоположного неравенства

$$Q_{МУА} \leq Q_{ОРИА}. \quad (21)$$

Если $\eta = n_B = 1$, ($n_W = 3$), то $\rho_2 = -(4 - 1)/2 + 0.5\sqrt{1 + 56 - 8} = 2.27$, то в точках $0 < r \leq \lfloor 2.77 \rfloor_- = 2$ выполняется неравенство (17), а в точках $r \leq \lfloor 2.77 \rfloor_+ = 3$ выполняется (21).

Квадратные скобки $\lfloor \cdot \rfloor_-$ обозначают операцию отбрасывания дробной части числа (округление с недостатком), $\lfloor \cdot \rfloor_+$ – округление до ближайшего целого с избытком. Подстановкой $n_B = 1$, $r = 2$ в неравенство

$$r^2 + (4 - n_B)r - 16n_B + 6 \leq 0 \quad (22)$$

можно убедиться, что оно удовлетворяется и, значит, $r = 2$ наибольшее значение r , при котором еще выполняется $Q_{МУА} \geq Q_{РИАНМ}$.

Аналогично, если $\eta = n_B = 2$, $n_W = 6$, то

$$\rho_2 = -(4 - 2)/2 + 0.5\sqrt{(2 + 56)2 - 8} = 4.39, \text{ то } 0 \leq r \leq \lfloor 4.39 \rfloor_- = 4 \text{ и т.д.}$$

Построив график и линейную регрессию, можно получить при малых значениях r и n_W зависимость

$$r(n_W) \approx \frac{2}{3} n_W. \quad (23)$$

Рассмотрим подкоренное выражение (20)

$$\eta^2 + 56\eta - 8 = 0, \quad \eta_{1,2} = -28 \pm \sqrt{676 + 8} = -28 \pm 28.1533, \quad \eta_1 \cong -56, \quad \eta_2 = 0.1533 \cong 0.$$

Поскольку величина $n_B = \eta$ является целочисленной переменной, подкоренное выражение аппроксимируем, как $(\eta + 56)\eta$. Легко заметить, что тангенс угла наклона пограничной кривой (19) асимптотически при $n_B \rightarrow \infty$ стремится к 1, т.е. $r \approx n_B$. Тогда в области значений этих параметров выше указанной границы

($r > n_B$, $r \approx 1/3 n_W$), алгоритм МУА имеет бóльшее быстроедействие, чем

РИАНМ и, наоборот, при моделировании сложных объектов чаще бывает, что $r < n_B$ – в таком случае быстроедействие РИАНМ выше, чем МУА.

В области малых значений $n_B \leq 3$, если зависимость $r = a n_B$ аппроксимировать линейной функцией, где коэффициент $a=2$, то если $0 < r \leq 2 n_B$ ($r \leq 2/3 n_W$) РИАНМ по скорости предпочтительнее, чем МУА и, наоборот, МУА имеет большее быстроедействие, чем РИАНМ при $r > 2 n_B$ ($r > 2/3 n_W$).

При использовании УКР, когда имеется одна точка в проверочной последовательности ($n_B = 1$) неравенство (18) имеет вид $\rho^2 + 2\rho - 16 \leq 0$, решая которое совместно с неравенством (17), получим что при $0 < r < 3$ $Q_{МУА} \geq Q_{РИАНМ}$, а при $r \geq 3$ $Q_{МУА} < Q_{РИАНМ}$. Заметим, что если $n_B=1$, то вышеприведенные условия выполняются для единственного значения $n_W = 3$, в силу $n_A : n_B = 2:1$.

Рассмотрим более общий случай Б) для любого $n_W \geq 2$, при соотношении $n_A : n_B = (n_W - 1) : 1$.

Подстановкой (15) и (10) в (17) получаем неравенство, которое имеет вид

$$r^2 + 3r - 7n_W + 7 \leq 0. \quad (24)$$

Построив параболу (24) (сплошная линия, см. рис. 3), подстановкой $r, n_W \in \mathbb{N}$ определим соответствующие области. В заштрихованной области, ограниченной штрихпунктирной кривой, построенной при целочисленных значениях r, n_W , выполняется условие большего быстроедействия МУА. В области ниже светлой штрихпунктирной кривой, наоборот, РИАНМ делает меньше операций.

Осталось рассмотреть еще один важный для практики случай В) деления выборки пополам $n_A : n_B = 1:1$.

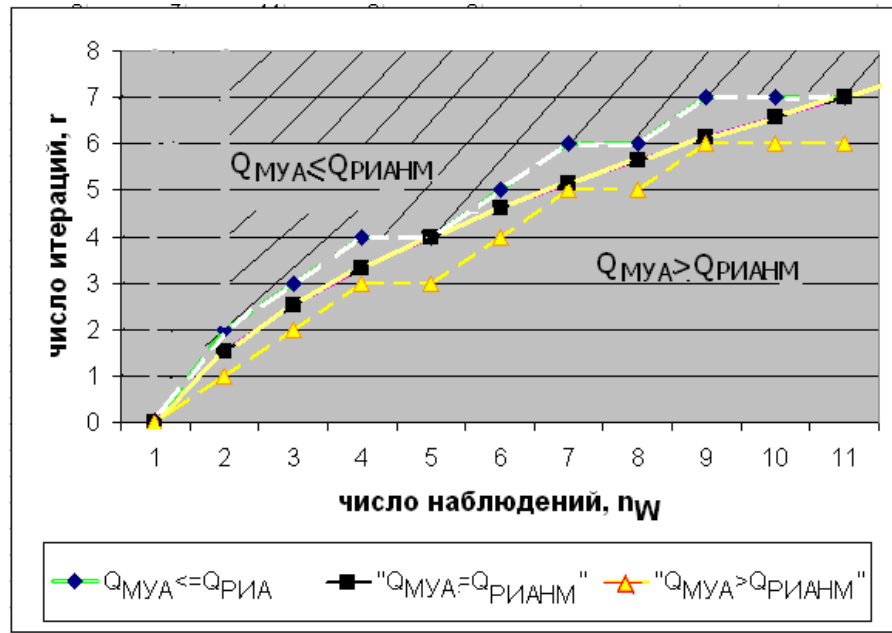


Рис. 2 – Результаты сравнения быстродействия РИАНМ и МУА в плоскости числа аргументов r «развернутой» модели и количества наблюдений n_W при условии $n_B = 1$ и отношении объемов выборок $n_A : n_B = (n_W - 1) : 1$.

Подстановкой (16) и(10) в (17) получаем следующее неравенство:

$$\rho^2 + (4 - \eta)\rho - 9\eta + 6 \leq 0. \tag{25}$$

Выбираем одно из решений равенства (25),

$$\rho_2 = -(4 - \eta) / 2 + 0.5\sqrt{(\eta + 28)\eta - 8} \tag{26}$$

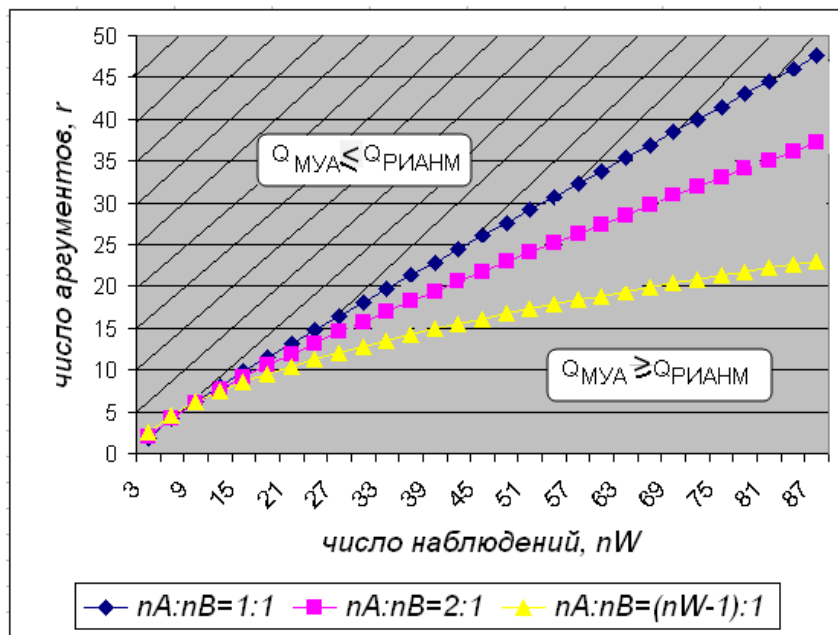


Рис. 3 – Результаты сравнения быстродействия РИАНМ и МУА в плоскости числа аргументов r «развернутой» модели и количества наблюдений n_W в зависимости от трех рассмотренных соотношений $n_A : n_B$.

Используя (26), для положительных значений $\eta > 0$, $\rho > 0$ можно указать области выполнения неравенств (17) и (21). Подстановкой целочисленных значений $(n_B, n_W, r) \in \mathbb{N}$ находятся соответствующие области выполнения неравенств, как и в случае А), округлением с недостатком $r = \lfloor \rho \rfloor_-$ (выполняется (17)) и округлением с избытком $r = \lceil \rho \rceil_+$, где выполняется (21).

В асимптотике при $n_B \rightarrow \infty$, как и в случае А), тангенс угла стремится к 1, т.е. $r \underset{n_B \rightarrow \infty}{\approx} n_B$, но $r \underset{n_W \rightarrow \infty}{\approx} 1/2n_W$.

На рисунке 3 приведены три кривые, отвечающие трем вышерассмотренным соотношениям объемов обучающей и проверочной выборок. Неравенства (выписанные в рамочках) в заштрихованных областях выполняются независимо от отношения $n_A : n_B$, в незаштрихованной – выполнение неравенств зависит от отношения объемов подвыборок. В наибольшей степени преимущество по быстройдействию РИАНМ по сравнению с МУА проявляется в случае деления выборки пополам, в наименьшей – когда в проверочной выборке всего одна точка.

7. Сравнительный анализ вычислительной сложности РИАРМ и МУА

Известны два метода и две формулы (9) и (10) оценки сложности вычислений релаксационных итерационных алгоритмов с рекуррентными вычислениями (РИАРМ), отличающиеся на $\rho + 3$. Сравним количество операций МУА и алгоритм с меньшим числом операций (РИАРМ₂), формулы (10) и (12). Также как и в случае с РИАНМ проверяем условие (17).

Рассмотрим случай А), когда отношение $n_A : n_B = 2:1$. С учетом перехода к непрерывным переменным, подставив в (17) выражения (14) и (12), получим:

$$\rho\eta + 16\eta + 3 - 3\rho - 26 \geq 0. \quad (27)$$

Преобразуем (27) к виду $(\eta - 3)(\rho + 16) + 25 \geq 0$. Откуда в соответствии с условием (17), если $\eta > 3$ должно выполняться неравенство

$$\rho \geq -\frac{25}{(\eta - 3)} - 16, \quad (28)$$

или по отношению к другой переменной – неравенство $\eta \geq -\frac{25}{(\rho + 16)} + 3$.

Построив гиперболу, соответствующую равенству (28), покажем на рисунке 4 область на плоскости двух координат, в которой выполняется условие (17). При подстановке $\eta = n_W \in \mathbb{N}$ в равенство (28) и округлением $r = \lceil \rho \rceil_+$ с избытком найдем область выполнения строгого неравенства (21), а округлением $r = \lfloor \rho \rfloor_-$ с недостатком – область строгого неравенства (17).

Почти везде (т.к. если $n_B \geq 3$ при любом количестве итераций (17) выполняется) РИАРМ₂ превосходит по быстродействию МУА, и только при $n_B = 2$ число итераций ограничено сверху, т.е. $0 < r \leq 9$. МУА имеет преимущество в небольшой области значений $n_B \in [1; 2]$, при $n_B = 2$, есть ограничение $r > 9$, а при $n_B = 1$ нет ограничений по r . Тот же характер имеет граница для РИАРМ₁

$$\rho \geq -\frac{12.5}{(\eta - 3)} - 13, \tag{29}$$

но график сдвинут правее при пересечении оси абсцисс ($\rho=0$) из точки $\eta = 1.44$ в точку $\eta = 2.03$. С учетом целочисленности значений n_B эти две гиперболы (28) и (29) при анализе существенно не отличаются (сливаются в одну на участке $3 > \eta > 2$), только в точке при $n_B = 2$ отсутствует ограничение на число итераций (МУА имеет преимущество при любом числе итераций). В области параметров $(n_W, r) \in \mathbb{N}$ поскольку $n_B = 1/3 n_W$, то при $n_W = 9$ число итераций $r \rightarrow \infty$, а значит РИАРМ без ограничений быстрее, чем МУА, $n_W \geq 9$.

При использовании УКР при $n_B = 1$ и $n_A : n_B = 2:1$ условия (22) и (23) всегда выполняются, а значит, оба варианта РИАРМ при любых ρ имеют большее быстродействие, чем МУА.

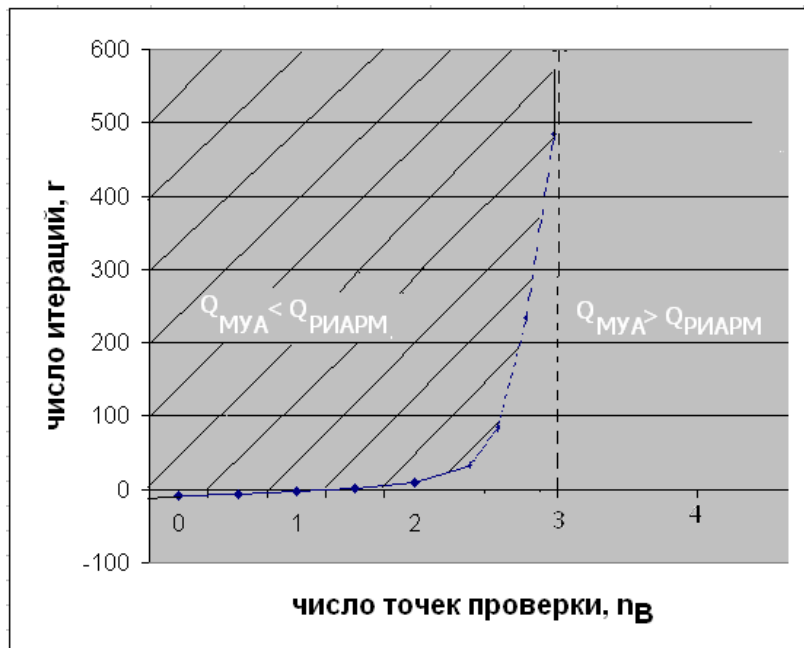


Рис. 4 – Результаты сравнения скорости РИАРМ₂ и МУА в плоскости значений аргументов r и объема проверочной выборки n_B при условии $n_A : n_B = 2:1$.

Рассмотрим случай Б) для любого $n_W \geq 2$ и соотношения $n_A : n_B = (n_W - 1) : 1$. При подстановке (15) и (11) в (17) получим, что для выполнения неравенства (17) следует провести для РИАРМ₁ анализ неравенства $4r + 29 \leq r + 7(n_W - 1) + 5$. При подстановке (15) и (12) в (17) получим $3r + 26 \leq r + 7(n_W - 1) + 5$ для РИАРМ₂. После преобразований имеем две линейные зависимости и убеждаемся, что в области значений $r \in \mathbb{N}$ выше пря-

мых $r = \frac{7}{3}n_W - \frac{31}{3}$ для РИАРМ₁ и $r = 3,5n_W - 14$ для РИАРМ₂ МУА имеет большее быстродействие, и наоборот, под прямыми находится область $r \in \mathbb{N}$ большего быстродействия РИАРМ, причем под прямой РИАРМ₂ область больше.

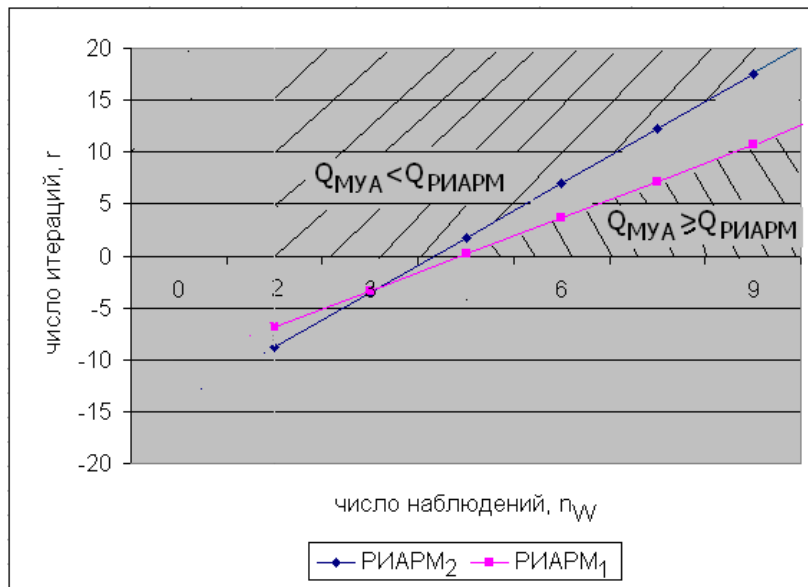


Рис. 5 – Результаты сравнения быстродействия РИАРМ и МУА в плоскости числа аргументов r «развернутой» модели и количества наблюдений n_W при условии $n_B = 1, n_A : n_B = (n_W - 1) : 1$.

Заштрихованные области соответствуют случаям безусловного выполнения неравенств (17) и (21), незаштрихованная – зависит от нерекуррентного метода вычисления параметров и критериев (варианта РИАРМ).

Рассмотрим случай В) деления исходной выборки пополам, $n_A : n_B = 1 : 1$. Подстановкой (16) и (11) в (17) получаем следующее неравенство для РИАРМ₁:

$$\rho\eta + 9\eta - 4\rho - 26 \geq 0. \tag{30}$$

Подстановкой (16) и (12) в (17) получаем неравенство для РИАРМ₂:

$$\rho\eta + 9\eta - 3\rho - 23 \geq 0. \tag{31}$$

Преобразуем (31) к виду $(\eta - 3)(\rho + 9) \geq -4$. Откуда для РИАРМ₂ получим отношение $\rho \geq -\frac{4}{\eta - 3} - 9$, а для РИАРМ₁ $\rho \geq -\frac{2}{\eta - 3} - 6$. Две полученные гиперболы, соответствующие равенствам (30) и (31), в области параметров $(n_B, r) \in \mathbb{N}$ ничем не отличаются от гиперболы, полученной решением равенства (29), только ось абсцисс они пересекают в разных точках: $\eta = 2,56$ и $\eta = 2,67$. В области целочисленных значений $(n_B, r) \in \mathbb{N}$ эти 2 гиперболы сливаются в одну. Их отличия можно пронаблюдать, построив их в области $(n_W, r) \in \mathbb{N}$. Поскольку $n_B = 0,5 n_W$ при $n_A : n_B = 1 : 1$, то $r \rightarrow \infty$ при $n_W = 6$.

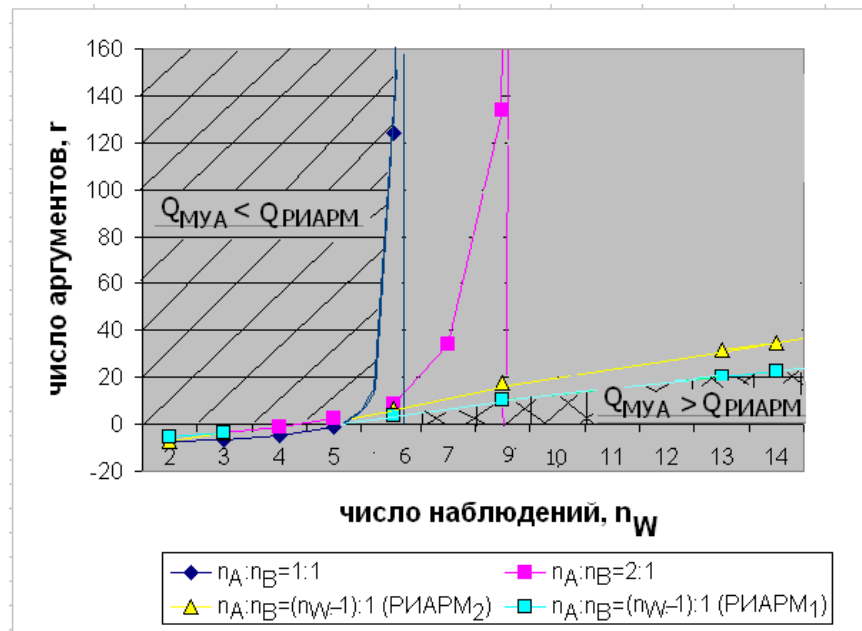


Рис. 6 – Результаты сравнения быстродействия РИАРМ и МУА в плоскости числа аргументов r «развернутой» модели и количества наблюдений n_W в зависимости от трех рассмотренных соотношений $n_A : n_B$

На рис. 6 приведены графики четырех функций $r = f(n_W)$, отвечающие трем рассмотренным соотношениям $n_A : n_B$ (четвертая функция для того, чтобы показать отличие РИАРМ₁ и РИАРМ₂). В заштрихованных областях указанные на рисунках неравенства выполняются независимо от отношения $n_A : n_B$, в незаштрихованной – выполнение неравенств зависит от отношения точек в подвыборках. В наибольшей степени преимущество по быстродействию РИАРМ по сравнению с МУА, как и при сравнении пары РИАНМ и МУА, проявляется в случае деления выборки пополам, а в наименьшей – когда в проверке имеется одна точка.

8. Выводы

Сравнение вычислительной сложности рекуррентных и нерекуррентного методов вычислений в ОРИА показало, что область значений числа аргументов m , при которых проявляются преимущества нерекуррентных алгоритмов, невелика и поэтому этот факт не может изменить общепринятого мнения о предпочтительности рекуррентных вычислений.

Сравнительный анализ вычислительной сложности МУА с РИАНМ и РИАРМ показал, что если в алгоритме МУА сделано одно усовершенствование – отделена стадия вычисления ковариационной матрицы от переборной стадии, то быстродействие ОРИА с нерекуррентными вычислениями зависит от числа точек n_B при любых возможных отношениях $n_A : n_B$.

Сравнение РИАРМ и МУА выявило обширную область изменения числа наблюдений $n_B \geq 3$ ($n_W \geq 9$) при отсутствии ограничений на число аргументов (итераций), когда РИАРМ превосходит МУА по быстродействию в двух традиционно применяемых вариантах разбиений: $n_A : n_B = 1:1$ и $n_A : n_B = 2:1$. Результа-

ты сравнения вычислительной сложности РИАРМ и МУА в случае $n_A : n_B = (n_W - 1) : 1$ показали, что между параметрами r и n_W имеется линейная зависимость. Выше этой линии МУА является более быстрым, по сравнению с РИАРМ, ниже РИАРМ имеет большее быстродействие. Поэтому использование УКР ($n_B = 1$), несколько ограничивает применение РИАРМ₂ в сравнении с МУА, при этом важным является то, что $r > n_W$, $r = 3,5n_W - 14$, а при некоррелированных аргументах для внутренней сходимости «по структуре» достаточно $r \approx 3n_W$ [7].

Литература

1. Павлов А. В. Обобщённый релаксационный итерационный алгоритм МГУА. Индуктивне моделювання складних систем. Збірник наук. праць. – К.: МННЦІТС, 2011. вип.3. – С. 95-108.
2. Шелудько О.И., Патереу С.Г. Упрощенный алгоритм идентификации характеристик сложных объектов по МГУА. В кн.: программы прямого синтеза моделей (по принципу самоорганизации). Киев: РФАП, 1975, вып.1.
3. Павлов А. В. Модифицированный алгоритм с комбинаторной селекцией и ортогонализацией переменных и его анализ. Индуктивне моделювання складних систем. Збірник наук. праць. К.:МННЦІТС, 2010, вип.2. – С.130-139.
4. Кондрашова Н.В., Павлов В. А., Павлов А.В. Многорядный алгоритм веерных решений // Вісник Національного технічного університету України „КПІ”. Інформатика, управління та обчислювальна техніка. К., «Век+», 2006. – 45. – С. 218-227.
5. Степашко В. С. Метод критических дисперсий как аналитический аппарат теории индуктивного моделирования / Степашко В. С. // Проблемы управления и информатики. – № 2. – 2008. – С. 8-26.
6. Ивахненко А.Г. Моделирование сложных систем по экспериментальным данным / Ивахненко А.Г., Юрачковский Ю.П. – М.: Радио и связь, 1987.– 120 с.
7. Павлов А. В. Технологія побудови регресійних моделей на основі релаксаційного ітераційного алгоритму з рекурентними обчисленнями об'єктів [Текст] : Дис. ... канд. техн. наук : 05.13.06 / Павлов Андрій Володимирович; Міжнародний науково-навчальний центр інформаційних технологій та систем МОН молоді та спорту України. – К., 2012. – 251 с.
8. Kondrashova N.V., Pavlov A.V., Pavlov Ya.V. Sample division with sliding interval as a method of increase of accuracy for time series forecasting – // Proc. of ICIM 2008. 2nd International conference on inductive modeling, Sept.15-19, 2008. Kyiv. - P. 250-254, ISBN 978-966-02-4889.
9. Ивахненко А.Г. Системы эвристической самоорганизации в технической кибернетике. – Киев: «Техніка», 1971. – 372 с.