

ПАРАЛЛЕЛЬНЫЙ АЛГОРИТМ РЕШЕНИЯ ЗАДАЧИ СГЛАЖИВАНИЯ ЭКСПЕРИМЕНТАЛЬНЫХ ЗАВИСИМОСТЕЙ ПО МЕТОДУ НАИМЕНЬШИХ КВАДРАТОВ

Е.А. Добрынина, Е.В. Ларкин, М.Б. Цудиков

Предложен метод решения задачи сглаживания экспериментальных зависимостей по методу наименьших квадратов. В рамках предложенного метода производится разработка алгоритма для минимизации случайной составляющей, которая появляется при погрешности измерений сигналов. Описан разработанный параллельный алгоритм, основанный на распараллеливании вычислительных операций для обеспечения быстродействия при обработке экспериментальных данных. Произведен анализ производительности метода.

Ключевые слова: параллельный алгоритм, экспериментальные данные, сеть Петри, кластерный способ распараллеливания, многопроцессорная обработка.

Производственные испытания объектов являются одним из важных этапов разработки изделий и поставки их на производство [1]. Одна из задач эксперимента – это определение реальных характеристик объекта, то есть анализ их значений в зависимости от различных параметров и сравнение их с требованием технического задания. Для определения характеристик в процессе проведения эксперимента задаются значения сигналов, подаваемых на объект, и измеряются их величины, которые характеризуют состояния объекта. Вследствие погрешности измерений из-за наличия шумов в измерительных каналах во входном и выходном сигнале кроме регулярной составляющей, определяемой физикой процесса в изделии, присутствует случайная составляющая, влияние которой на результат эксперимента, необходимо уменьшить. Для уменьшения влияния случайных факторов на результаты эксперимента осуществляется набор экспериментальных данных как по входным, так и выходным сигналам и сглаживание статистических данных по методу наименьших квадратов [2,9].

Статистические данные, которые используются для сглаживания экспериментальных прямых, представляют собой массивы значительных объемов: массив входных воздействий $X = \{x_1, \dots, x_n, \dots, x_N\}$ и массив значений состояния объекта $Y = \{y_1, \dots, y_n, \dots, y_N\}$ [3]. Значения x_i и y_i ложатся на некоторую параметрическую кривую $y = \varphi(x_i)$ с ошибкой [6]:

$$\varepsilon = \sum_{i=1}^n (y_i - \varphi(x_i))^2. \quad (1)$$

При минимизации производные от ошибки по параметрам принимают вид:

$$\sum_{i=1}^n (y_i - \varphi(x_i, a_1, \dots, a_j, \dots, a_k)) \left[\frac{\partial \varphi(x_i, a_1, \dots, a_j, \dots, a_k)}{\partial a_j} \right] = 0, \quad j = 1, \dots, k. \quad (2)$$

Выражение (2) представляет собой систему из k уравнений с k неизвестными, решение которой дает параметрическую кривую, наилучшим образом соответствующую статистическим данным [4].

В частном случае $\varphi(x_i)$ представляет собой прямую линию

$$y = ax + b, \quad (3)$$

где a и b параметры, которые необходимо определить.

Дифференцирование по параметрам a и b дает следующую систему линейных уравнений:

$$\begin{aligned} \sum_{i=1}^n [y_i - (ax_i + b)]x_i &= 0; \\ \sum_{i=1}^n [y_i - (ax_i + b)] &= 0. \end{aligned} \quad (4)$$

Решение системы имеет вид:

$$a = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2}; \quad (5)$$

$$b = \frac{\sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i - \sum_{i=1}^n x_i \sum_{i=1}^n x_i y_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2}. \quad (6)$$

Алгоритм расчета параметров a и b показан на рисунке 1.

Алгоритм состоит из двух частей. Первая часть циклическая. Она реализует накопление сумм z_1, z_2, z_3, z_4 с использованием счетчика i . Вторая часть линейная [10]. Она реализует расчет числителей (z_5, z_6) в зависимостях (5), (6), знаменателя (z_7) и расчет собственно параметров a и b . Два последовательных цикла представляются сетью Петри, показанной на рисунке 2.

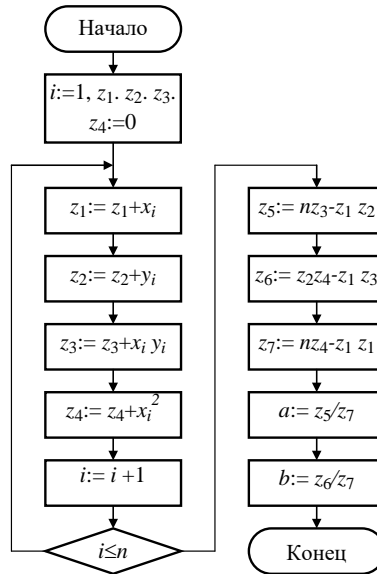


Рис. 1. Последовательный алгоритм сглаживания экспериментальных кривых

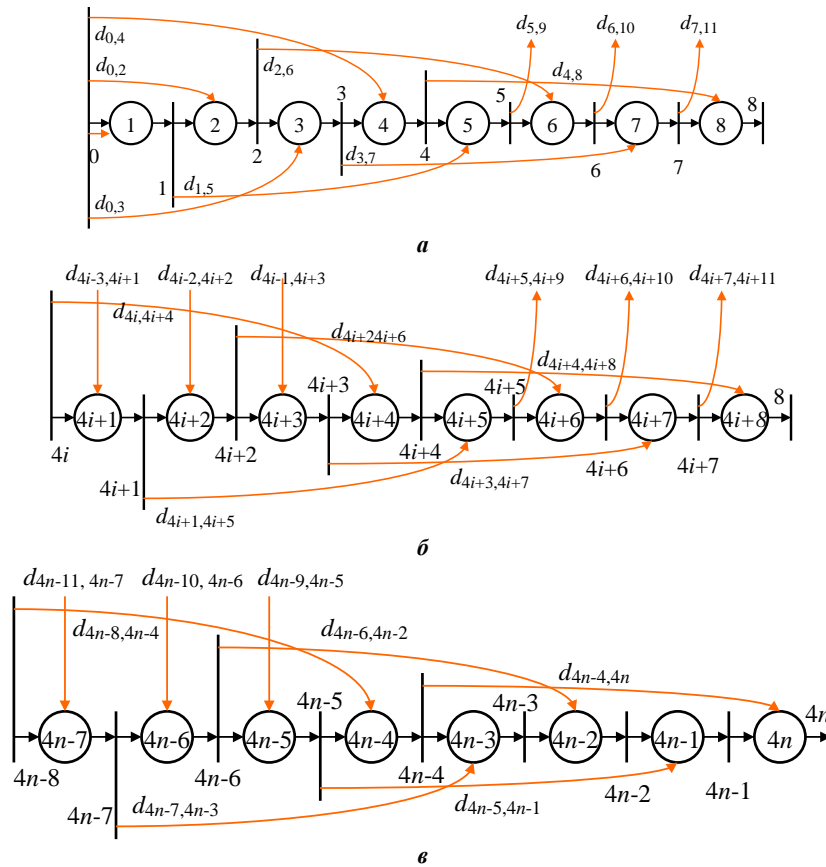


Рис. 2. Сеть Петри, представляющая циклическую часть алгоритма сглаживания экспериментальных кривых: а – начало сети; б – середина сети; в – конец сети

Линейная часть алгоритма представляется сетью, показанной на рисунке 3. Позиции в сети Петри моделируют выполнение операторов алгоритма, а переходы – переключение из предыдущего оператора в следующий. Дуги сети определяют передачу данных между операторами ($d_{...}$) или передачу управления ($c_{...}$).

Последовательное преобразование сети в параллельную структуру показано на рисунке 4.

Аналогичное преобразование циклической части дает параллельный алгоритм, структура которого приведена на рисунке 5.

Таким образом, модель параллельных вычислений представляет собой четыре независимых цепочки, начинающиеся в позициях 1, 2, 3, 4 соответственно. По физическому смыслу цепочка позиций 1-5-9-...-(4n-3) моделирует последовательное накопление суммы $\sum_{i=1}^n x_i$, цепочка 2-6-10-...-(4n-2) - накопление суммы $\sum_{i=1}^n y_i$; цепочка 3-

7-11-...-(4n-1) - суммы $\sum_{i=1}^n x_i^2$, а цепочка 4-8-12-...-4n моделирует накопление – суммы $\sum_{i=1}^n x_i y_i$. По временной вычислительной сложности эти цепочки неравнозначны. На накопление сумм $\sum_{i=1}^n x_i$ и $\sum_{i=1}^n y_i$ тратится гораздо меньше времени, чем на накопление сумм $\sum_{i=1}^n x_i^2$ и $\sum_{i=1}^n x_i y_i$, поскольку при накоплении сумм элементов массива тратится τ_Σ единиц машинного времени, а на накопление сумм произведений и возведений в квадрат – $\tau_\Sigma + \tau_\Pi$ единиц машинного времени. Для команд формата регистр-регистр

$$\tau_\Sigma = 4\tau; \tau_\Pi = 20\tau, \quad (7)$$

где τ – период сигнала частоты процессора (физическое время).

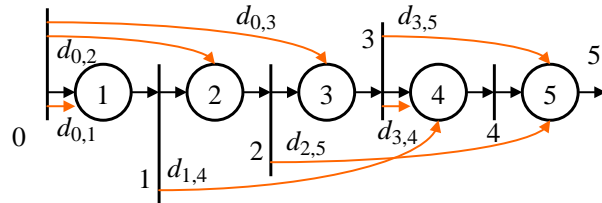


Рис. 3. Сеть Петри, моделирующая линейную часть алгоритма

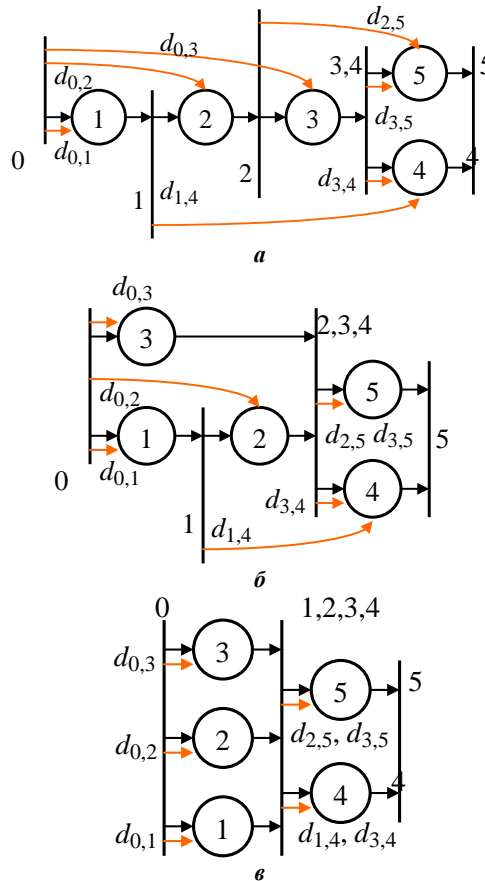


Рис. 4. Последовательное преобразование (а, б) алгоритма в параллельную модель (в)

Пусть обработка результатов эксперимента проводится на четырехпроцессорной MIMD-ЭВМ. При параллельном представлении алгоритма, показанном на рисунке 5 массив объемом n точек обрабатывается за

$$T = (\tau_\Sigma + \tau_\Pi)n = 24m \quad (8)$$

единиц времени, причем $20\tau_\Pi n = \tau_\Pi n$ ед. времени процессоры, производящие расчет $\sum_{i=1}^n x_i$ и $\sum_{i=1}^n y_i$, будут простаивать.

В целом отношение суммарного времени простоя к общему процессорному времени, затрачиваемому на решение задачи, равно

$$E = \frac{2 \cdot \tau_{\Pi}}{4(\tau_{\Sigma} + \tau_{\Pi})} \quad (9)$$

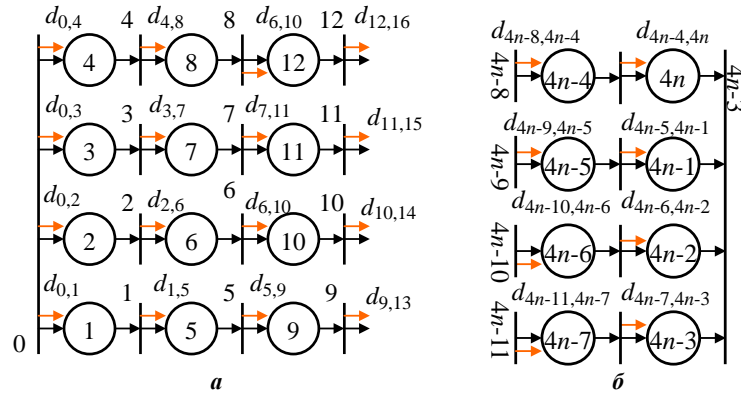


Рис. 5. Параллельная сеть Петри, моделирующая начало (а) и конец (б) циклической части алгоритма

Другой способ решения задачи - это формирование кластеров, в которых вычисляются тетрады $-\sum_{i=1}^n x_i$,

$\sum_{i=1}^n y_i$, $\sum_{i=1}^n x_i^2$, $\sum_{i=1}^n x_i y_i$ (смотри рисунок 6).

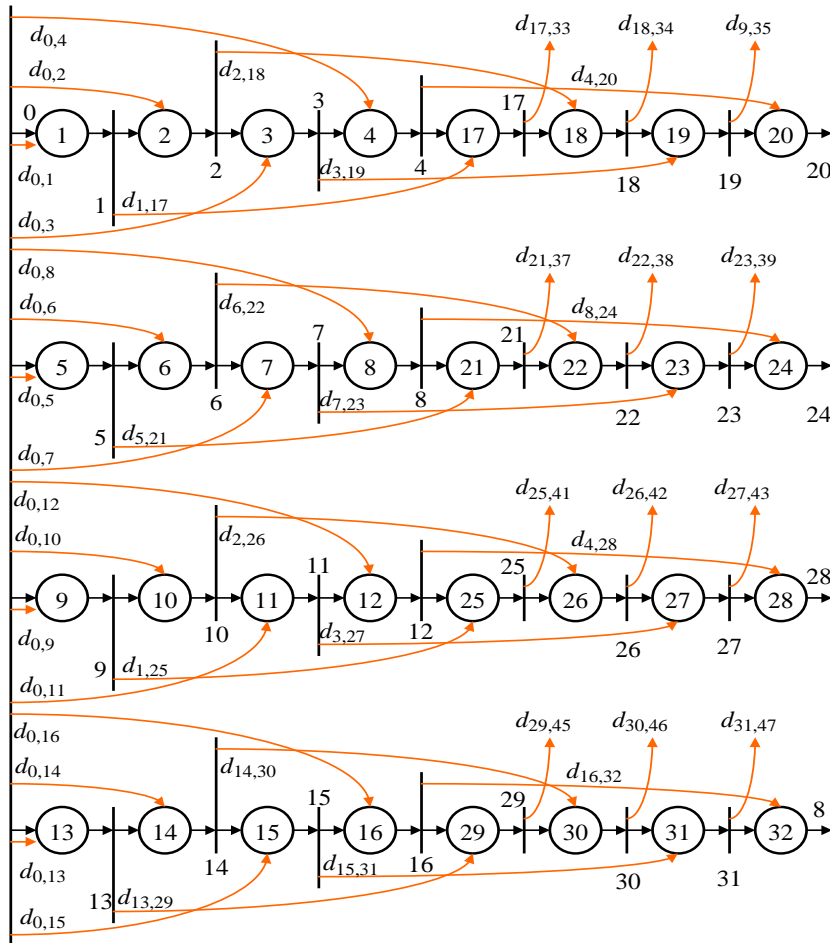


Рис. 6. Кластерный способ распараллеливания

Каждый кластер обрабатывается в течение времени $4\tau_{\Sigma} + 2\tau_{\Pi}$. Количество подобных кластеров равно n . В четырехпроцессорной ЭВМ общее время обработки равно примерно

$$T = \left(\tau_{\Sigma} + \frac{1}{2} \tau_{\Pi} \right) n = 14m \quad (10)$$

Отношение суммарного времени простоя к общему процессорному времени, затрачиваемому на решение задачи, равно нулю.

Выигрыш по времени по сравнению с предыдущей схемой равен 41,7%.

Таким образом, приведенные исследования показывают, что при сглаживании экспериментальных кривых целесообразно применение параллельной обработки данных. Из двух возможных реализаций расчета параметров линейного сглаживания более эффективным является кластерный метод [8], который обеспечивает повышенное быстродействие при обработке экспериментальных данных.

Список литературы

1. ГОСТ Р 15.301-2016. Система разработки и постановки продукции на производство. Продукция производственно-технического назначения. Порядок разработки и постановки продукции на производство. М.: Стандартинформ, 2018. 12 с.
2. Задорина Н.А. Обработка экспериментальных данных на ЭВМ. Учебное пособие. Рыбинск: РГАТА, 2009. 100 с.
3. Радченко С.Г. Устойчивые методы оценивания статистических моделей. К.: ПП «Санспарель», 2005. 504 с.
4. Дайнтбегов Д.М., Калмыкова О.В., Черепанов А.И. Программное обеспечение статистической обработки данных: Учебное пособие. М.: Финансы и статистика, 1984. 192 с.
5. Дрейпер Н., Смит Г. Прикладной регрессионный анализ. Множественная регрессия = Applied Regression Analysis. 3-е изд. М.: «Диалектика», 2007. 912 с.
6. Фёрстер Э., Рёнд Б. Методы корреляционного и регрессионного анализа = Methoden der Korrelation - und Regressiolyns analyse. М.: Финансы и статистика, 1981. 302 с.
7. Захаров С.И., Холмская А.Г. Повышение эффективности обработки сигналов вибрации и шума при испытаниях механизмов // Вестник машиностроения. М.: Машиностроение, 2001. № 10. С. 31–32.
8. Радченко С.Г. Методология регрессионного анализа. К.: «Корнийчук», 2011. 376 с.
9. Лоусон Ч., Хенсон Р. Численное решение задач методом наименьших квадратов. М.: Наука, 1986. 232 с.
10. Отнес Р., Эноксон Л. Прикладной анализ временных рядов. М.: Мир, 1982. 427 с.

Добрынина Елизавета Андреевна, студент, dobryninaelizaveta2021@gmail.com, Россия, Тула, Тульский государственный университет,

Ларкин Евгений Васильевич, д-р техн. наук, профессор, elarkin@mail.ru, Россия, Тула, Тульский государственный университет,

Цудиков Михаил Борисович, канд. техн. наук, доцент, tsudickov.mb@yandex.ru, Россия, Тула, Тульский государственный университет

SMOOTHING EXPERIMENTAL DEPENDENCE PROBLEM SOLVING BASED ON ORDINARY LEAST SQUARES METHOD PARALLEL ALGORITHM

E.A. Dobrynina, E.V. Larkin, M.B. Tsudikov

Smoothing experimental dependence problem solving based on ordinary least squares method is offered. Developing an algorithm for random component minimization caused by measurement inaccuracy is executed. Parallel algorithm based on computing operation paralleling for upping the speed of experimental data processing is described. Method efficiency is analyzed.

Key words: parallel algorithm, measured data, Petri net, cluster type, multiprocessing.

Dobrynina Elizaveta Andreeveta, student, dobryninaelizaveta2021@gmail.com, Russia, Tula, Tula State University,

Larkin Eugene Vasylyevich, doctor of technical science, professor, elarkin@mail.ru, Russia, Tula, Tula State University,

Tsudikov Mikhail Borisovich, candidate of technical sciences, docent, tsudickov.mb@yandex.ru, Russia, Tula, Tula State University