

K. II. 1130

IX/16

ALGORYTMY

VOL. IX • No. 16 • 1972



INSTYTUT MASZYN MATEMATYCZNYCH



ALGORYTMY
Vol. IX, N° 16 1972

[Faint, illegible text, likely bleed-through from the reverse side of the page]

[Faint, illegible text at the bottom right corner]

Copyright © 1972 - by Instytut Maszyn Matematycznych
Poland

Wszelkie prawa zastrzeżone



K o m i t e t R e d a k c y j n y

Antoni MAZURKIEWICZ /red. nacz./, Krzysztof MOSZYŃSKI, Zdzisław PAWLAK
Jan WIERZBOWSKI, Andrzej WIŚNIEWSKI, Ryszard ZIELIŃSKI

Romana NITKOWSKA /sekr. red./

Adres Redakcji: Warszawa, ul. Krzywickiego 34, tel. 28-37-29

W/ (40,-)
Ak. dn. 17 XI 72

TREŚĆ
СОДЕРЖАНИЕ
CONTENTS

Анатолий Герасимович

str.

СРАВНИТЕЛЬНЫЙ АНАЛИЗ СТАТИСТИЧЕСКИХ КРИТЕРИЕВ ПРЕДЛОЖЕННЫХ ДЛЯ ВЫЯВЛЕНИЯ ИЗМЕРЕНИЙ, СОДЕРЖАЩИХ ГРУБЫЕ ОШИБКИ 5

ANALIZA PORÓWNAWCZA TESTÓW STATYSTYCZNYCH SŁUŻĄCYCH DO WYKRYWANIA BŁĘDÓW DUŻYCH

A COMPARATIVE ANALYSIS OF STATISTIC CRITERIA PRESENTING MEASUREMENTS INCLUDING BIG ERRORS

Jan Goliński

PRZEGLĄD KILKU METOD OPARTYCH NA KONCEPCJI GRADIENTU EKSPERYMENTALNEGO I PORÓWNANIE ICH EFEKTYWNOŚCI 27

ПРОСМОТР НЕКОЛЬКИХ МЕТОДОВ ОСНОВАННЫХ НА КОНЦЕПЦИИ ЭКСПЕРИМЕНТАЛЬНОГО ГРАДИЕНТА И ИХ СРАВНЕНИЕ

A REVIEW OF SOME METHODS BASED ON THE CONCEPTION OF AN EXPERIMENTAL GRADIENT AND THEIR EFFECTIVITY COMPARISON

Piotr Krzystek

PEWNA METODA NUMERYCZNEGO ROZWIĄZANIA ZAGADNIENIA BRZEGOWEGO DLA NIELINIOWEGO RÓWNANIA RÓŻNICZKOWEGO ZWYCZAJNEGO IV RZĘDU 53

НЕКИЙ МЕТОД ЧИСЛЕННОГО РЕШЕНИЯ БЕРЕГОВОЙ ПРОБЛЕММЫ ДЛЯ НЕЛИНЕЙНОГО-ДИФФЕРЕНЦИАЛЬНОГО ПРЯМОГО УРАВНЕНИЯ IV ПЯДА

A CERTAIN METHOD OF A NUMERICAL SOLUTION OF A BORDER PROBLEM FOR A SIMPLE NONLINEAR DIFFERENTIAL EQUATION OF THE IV RANGE

Jerzy Leszczyński

JĘZYK MODELWANIA SYMULACYJNEGO PROCESÓW DYSKRETNYCH
SYM 69 85

ЯЗЫК СУМ 69 ДЛЯ МОДЕЛИРОВАНИЯ ДИСКРЕТНЫХ ПРОЦЕССОВ
SYM 69 LANGUAGE FOR DISCRETE PROCESSES SIMULATION

СРАВНИТЕЛЬНЫЙ АНАЛИЗ СТАТИСТИЧЕСКИХ
КРИТЕРИЕВ ПРЕДЛОЖЕННЫХ ДЛЯ ВЫЯВЛЕНИЯ
ИЗМЕРЕНИЙ, СОДЕРЖАЩИХ ГРУБЫЕ ОШИБКИ

Анатолий ГЕРАСИМОВИЧ

Белорусский политехнический институт

В данной статье методом Монте-Карло на ЭВМ Минск-22 моделируются кривые мощностей некоторых статистических критериев, предложенных для отбраковки сомнительных наблюдений.

ВВЕДЕНИЕ

В настоящее время в последних работах по теории ошибок [2], метрологии [5] намечается тенденция к подразделению грубых ошибок по происхождению на два подкласса: а/ явно грубые или промахи, б/ неявно грубые или "большие" по терминологии [5].

К явно грубым ошибкам обычно относят промахи в наблюдениях, вызываемые невнимательностью оператора, неисправностью инструмента, плохой организацией работы и т.д.

Измерения, содержащие "неявно грубые" ошибки, возникают из-за внезапных кратковременных изменений условий измерений, т.е. возникают в том случае, когда основная масса ($n-k$) измерений проведена при стабильных условиях, а " k " измерений при изменившихся. Измерения, проведенные при изменившихся условиях, в со-

ветской литературе называют "аномальными", "резко выделяющимися", "сомнительными", в английской и американской - "outliers", "struggles", "blunders", "wildshots".

Будем считать, что неявно грубые ошибки распределены по нормальному закону, т.е. выборка, содержащая "сомнительные" измерения, соответствует распределению вида

$$(1-\gamma) N(a, \delta) + \gamma N(a+\lambda_1, \delta\lambda_2); \quad 0 < \gamma < 1.$$

В литературе [13] имеются сведения, что даже тщательно поставленные эксперименты содержат до 10% измерений с неявно грубыми ошибками. Однако эти 10% могут дать сильное смещение при оценке параметров нормального закона. Ввиду этого вопросы выявления таких измерений остаются по-прежнему актуальными. Об этом свидетельствует тот факт, что вплоть до последнего времени продолжают появляться новые методы и критерии [9, 13, 14, 15]. Однако, как существующие ранее, так и вновь предложенные критерии, как правило, не классифицированы, их достоинства и недостатки слабо изучены, ввиду чего трудно отдать предпочтение какому-либо из них. Обычно автор, предлагая свой критерий, не объясняет, чем он лучше других критериев.

Поэтому в данной статье сделана попытка классифицировать методы и критерии, наиболее часто применяемые в практических расчетах (см. рис.1), а также сравнить эффективность некоторых статистических и вероятностных критериев выявления неявно грубых ошибок (на рис.1 исследуемые критерии обведены пунктирной линией).

В противовес методу Гаусса, основанному на испытании условий и отбраковке измерений, полученных при условиях существенно отличающихся от стабильных, все применяемые в настоящее

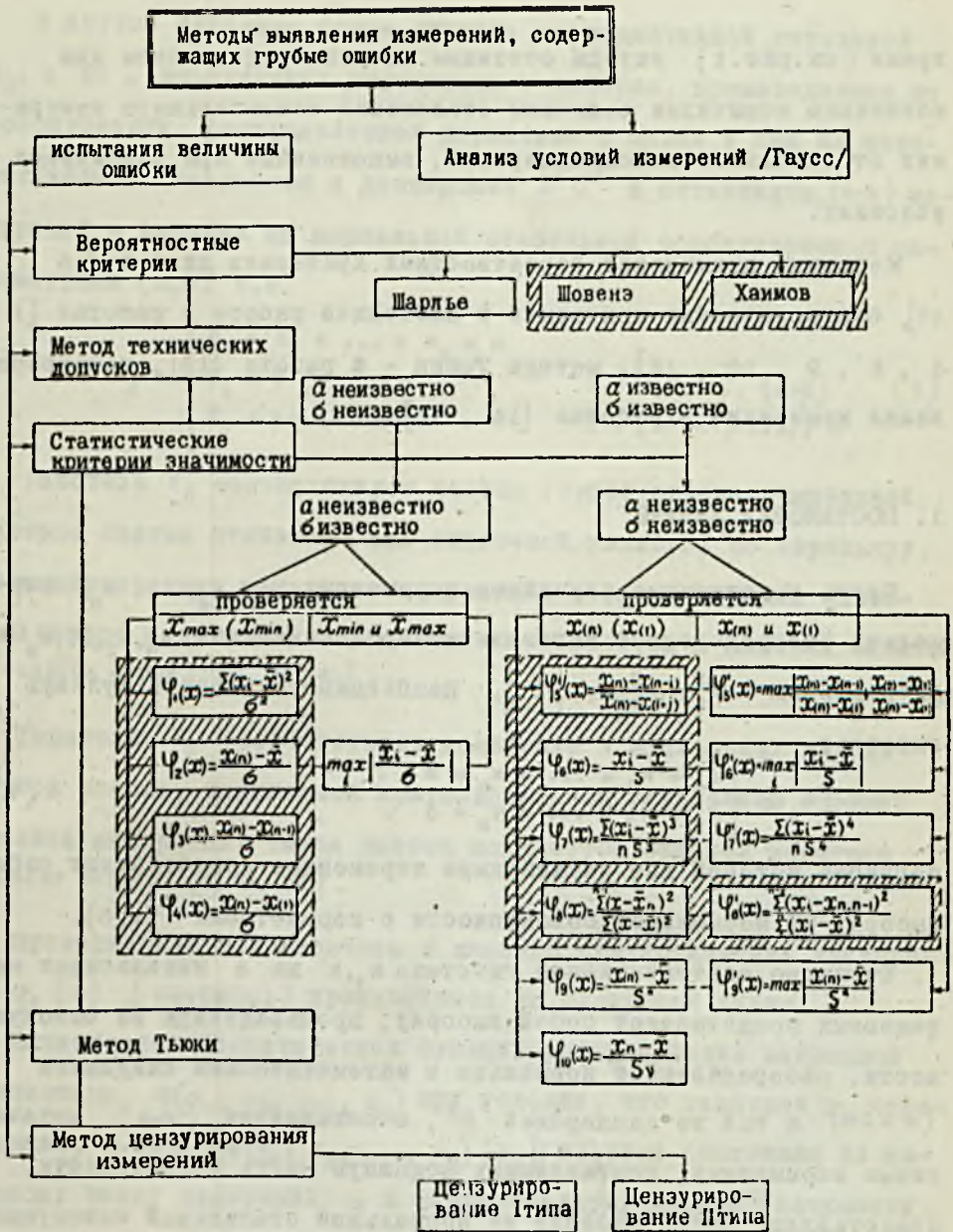


Рис.1

время (см.рис.1) методы основаны на непосредственном или косвенном испытании величины отклонения сомнительного измерения от остальной массы измерений, выполненных при стабильных условиях.

Методика применения вероятностных критериев дана в [16, 17], статистических критериев в настоящей работе и работах [1, 4, 8, 9, 10, 12], метода Тьюки - в работе [13], цензурирования измерений в работах [14, 15].

1. ПОСТАНОВКА ЗАДАЧИ

Пусть независимые случайные переменные x_1, x_2, \dots, x_n нормально распределены с математическими ожиданиями a_1, a_2, \dots, a_n и дисперсиями $\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2$. Необходимо проверить нулевую гипотезу

$$H_0 : \begin{cases} a_1 = a_2 = \dots = a_n = a \\ \sigma_1 = \sigma_2 = \dots = \sigma_n = \sigma \end{cases} \quad (1)$$

согласно которой все независимые переменные представляют собой выборку из нормальной совокупности с параметрами (a, σ) .

Согласно альтернативной гипотезе $H_{1,k}$ из n независимых переменных представляет собой выборку, произведенную из совокупности, распределенной нормально с математическим ожиданием $(a + \lambda\sigma)$ и той же дисперсией σ^2 , а оставшиеся $(n-k)$ независимых переменных, составляющих основную часть совокупности, представляют собой выборку из нормальной стабильной совокупности с параметрами (a, σ) т.е.

$$H_{1,k} : \begin{cases} a_i = a & i = 1, 2, \dots, (n-k) \\ a_j = a + \lambda\sigma & j = (n-k+1), \dots, n \\ \sigma_1 = \sigma_2 = \dots = \sigma_n \end{cases} \quad (2)$$

В другом варианте будем считать альтернативной гипотезой H_2 : k из n независимых переменных - выборка, произведенная из совокупности, распределенной нормально с одним и тем же математическим ожиданием и дисперсией $\lambda^2 \sigma^2$ а оставшиеся $(n-k)$ измерений - выборка из нормальной стабильной совокупности с параметрами (a, σ) т.е.

$$H_2 : \begin{cases} a_1 = a_2 = \dots = a_n = a \\ \sigma_1 = \sigma \\ \sigma_j = \lambda \sigma \end{cases} \quad \begin{matrix} i = 1, 2, \dots, (n-k) \\ j = (n-k+1), \dots, n \end{matrix} \quad (3)$$

Гипотеза H_1 соответствует неявно грубой ошибке, вызванной неверно снятым отсчетом (или ошибочной записью) по вереньеру, кратковременному изменению напряжения в блоке питания светодиода и др. Такого рода ошибки относятся к ошибкам сдвига (location error) [8, 9, 10].

Гипотеза H_2 соответствует появлению в измерениях неявно грубой ошибки, вызываемой кратковременным ухудшением внешних условий измерений. Такая ошибка называется ошибкой масштаба (scalar error) [8, 9, 10].

Проверка нулевой гипотезы с помощью статистических критериев $\varphi_1(x)$ (см. рис. 1) производится по следующей схеме:

Вычисляется теоретическая функция распределения выбранной статистики $\varphi(x_1, x_2, \dots, x_n)$ при условии, что гипотеза H_0 справедлива, т.е. $F[\varphi(x_1, x_2, \dots, x_n) | H_0]$; на этом основании по заданному числу измерений n и фиксированному уровню значимости α определяются предельные значения этой статистики $\gamma(\alpha, n)$; по результатам измерений вычисляется эмпирическое значение статистики $\varphi_{\text{эмп}}(x_1, x_2, \dots, x_n)$; сравнивает вычисленное эмпирическое значение статистики $\varphi_{\text{эмп}}(x_1, x_2, \dots, x_n)$ с предельным зна-

чением $\gamma(\alpha, n)$. Если $\Phi_{\text{ЭМП}}(x_1, x_2, \dots, x_n) > \gamma(\alpha, n)$ то на основании принципа практической невозможности маловероятных событий считают, что проверяемое измерение содержит неявно грубую ошибку, и не включают в обработку. В противном случае в обработку включают все измерения.

Практически проверка нулевой гипотезы сводится к вычислению эмпирического значения статистики $\Phi(x_1, x_2, \dots, x_n)$ и сравнение вычисленного значения с предельным /табличным/ значением $\gamma(\alpha, n)$.

Рассмотрим некоторые критерии проверки нулевой гипотезы (1), основанные на выборе различных характеристик $\varphi_1(x)$. В зависимости от того, известны или неизвестны параметры a и b , их можно разбить [1] на четыре класса: a неизвестно, b известно; a, b неизвестны; a, b известны; a известно, b неизвестно. Рассмотрим два первых.

Критерии, основанные на априорном знании параметра a , не рассматриваем, так как большинство измерений, как правило, ставятся с целью определения этого параметра.

Пусть x_1, x_2, \dots, x_n - результаты измерения некоторой величины - случайная выборка объема n из нормальной совокупности;

$x_{(1)}, x_{(2)}, x_{(3)}, \dots, x_{(n)}$ - эта же выборка, расположенная в неубывающий ряд $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$; $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ - средняя арифметическая из всех результатов измерений; $\bar{x}_{n-1} = \frac{1}{(n-1)} \sum_{i=1}^{n-1} x_i$

- средняя арифметическая, полученная после исключения из выборки измерения x_n ; $s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n}}$ - средняя квадратическая ошибка измерений.

Для проверки одного максимального резко отклоняющегося измерения $x_{(n)}$ в случае, если σ известно, исследованы критерии основанные на следующих статистиках:

$$\varphi_1(x) = \frac{\sum (x_i - \bar{x})^2}{\sigma^2}; \quad \varphi_2(x) = \frac{x_{(n)} - \bar{x}}{\sigma}; \quad \varphi_3(x) = \frac{x_{(n)} - x_{(n-1)}}{\sigma}; \quad \varphi_4(x) = \frac{x_{(n)} - x_{(n)}}{\sigma}$$

Для проверки одного максимального резко отклоняющегося измерения $x_{(n)}$ в случае, если σ неизвестно, исследованы критерии основанные на тестовых статистиках:

$$\varphi_5(x) = \gamma_n = \frac{x_{(n)} - x_{(n-1)}}{x_{(n)} - x_{(2)}}; \quad \varphi_5'(x) = \gamma_{22} = \frac{x_{(n)} - x_{(n-2)}}{x_{(n)} - x_{(3)}};$$

$$\varphi_6(x) = \frac{x_{(n)} - \bar{x}}{S}; \quad \varphi_7(x) = \frac{\sum_{i=1}^n (x_i - \bar{x})^3}{nS^3};$$

$$\varphi_8(x) = \frac{\sum_{i=1}^{n-1} (x_i - \bar{x}_n)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}; \quad \varphi_8'(x) = \frac{\sum_{i=1}^{n-2} (x_i - \bar{x}_{n,n-1})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}.$$

Таблицы предельных значений статистик $\varphi_1(x)$, $\varphi_2(x)$, $\varphi_3(x)$, $\varphi_4(x)$, $\varphi_5(x)$, $\varphi_5'(x)$, соответствующих данному числу измерений n и фиксированному уровню значимости α , имеются в [1], статистик $\varphi_6(x)$, $\varphi_8(x)$ и $\varphi_8'(x)$ - в [10], статистики $\varphi_7(x)$ - в [9].

Выявим среди них наиболее эффективные, а также оценим потерю эффективности, возникающую при применении менее эффективных критериев, но основанных на более простых тестовых статистиках $\varphi_1(x)$.

2. МЕТОДИКА ИССЛЕДОВАНИЯ

Проверка нулевой гипотезы (H_0) с помощью статистических критериев $\varphi_1(x)$ сопровождается вероятностями принятия ложных ре-

шений, т.е. браковки H_0 , когда она верна (ошибка I рода), или принятия H_0 , когда она неверна (ошибка II рода). Вероятность совершения ошибки первого рода определяется из условия

$$P \left\{ \varphi_1(x) > \tau(\alpha, n) \mid H_0 \right\} = \alpha; \quad (4)$$

вероятность ошибки II-го рода — из условия

$$P \left\{ \varphi_1(x) \leq \tau(\alpha, n) \mid H_1 \right\} = \beta \quad (5)$$

Из двух критериев K_1 и K_2 заданного уровня α критерий K_1 разумно считать более эффективным, чем критерий K_2 , если $\beta_1 < \beta_2$.

Отношение

$$R = \frac{1 - \beta(K_2)}{1 - \beta(K_1)} \quad (6)$$

характеризует эффективность критерия K_2 по сравнению с критерием K_1 при проверке гипотезы H_0 по сравнению с альтернативой H_1 .

Необходимые сведения с мощности различных статистических критериев можно получить, применяя метод цифрового статистического моделирования (метод Монте-Карло).

Процесс сравнения эффективности статистических критериев значимости методом Монте-Карло можно разделить на ряд этапов:

- а) Получение n последовательностей случайных чисел объема n имитирующих истинные ошибки измерений, распределенные по нормальному закону в соответствии с нулевой гипотезой H_0 .
- б) Расположение полученных в последовательностях чисел в вариационный ряд

$$x_{(1)i} \leq x_{(2)i} \leq x_{(3)i} \leq \dots \leq x_{(n)i}, \quad (7)$$

где $i = 1, 2, \dots, N$ — номер последовательности.

- в) Вычисление уровня значимости $\alpha = \frac{m_{\varphi_1 > \tau}}{N}$ по результатам всех случайных чисел, превосходящих квантиль τ .
- г) Введение в k крайних максимальных (минимальных) членов вариационного ряда неявно грубой ошибки прибавлением постоян-

ного числа λb (согласно гипотезе H_1) или умножением на постоянное число λ (согласно гипотезе H_2)

д) Вычисление оценки вероятностей совершить ошибку П-рода

$$\beta = \frac{m \Phi_1 \epsilon \Gamma}{N}$$

е) Оценка точности вычисления вероятностей α и β .

Наиболее часто нормально распределенные числа формируются на основании некоторой вспомогательной совокупности чисел, которая может быть получена по возможности с наименьшими затратами машинного времени и которая обеспечивает удобство дальнейших преобразований. В качестве такой вспомогательной совокупности нами была моделирована совокупность чисел, равномерно распределенных в интервале $(0, 1)$.

Равномерно распределенные числа $[\alpha_i]$ моделировались нами по формуле

$$\alpha_{i+1} = [\alpha_i + \alpha_{i-1}] \pmod{4} \quad (8)$$

где $\alpha_0 = \pi$; $\alpha_1 = 5^{17} \cdot 2^{-42}$; $[\alpha_i + \alpha_{i-1}] \pmod{4}$

- остаток от деления суммы чисел на 4, предложенной Дэвисом в [7].

Нормально распределенные числа получались суммированием 15 равномерно распределенных в интервале $(0, 1)$ чисел.

Основанием для этого служит центральная предельная теорема теории вероятностей для одинаково распределенных случайных величин: "если независимые случайные величины $\alpha_1, \alpha_2, \dots, \alpha_n$ имеют одно и то же распределение вероятностей и если каждое α_i имеет математическое ожидание a_i и среднее квадратическое отклонение b_i , то сумма асимптотически нормальна с математическим ожиданием na_1 и средним квадратическим отклонением $b_i \sqrt{n}$ "

Полученные таким образом нормально распределенные числа нормировались, т.е. преобразовывались в числа, нормально распределенные с параметрами $(0,1)$. Дополнительно для улучшения распределения на краях нормальной кривой применено преобразование Корниша-Фишера (см. [3, стр.100]). Проверка качества нормально распределенных чисел проводилась по методу, предложенному А.А.Петровым, суть которого состоит в том, что все исследуемые в данной работе статистики $\varphi_1(x)$ могут применяться для проверки гипотезы нормальности по выборкам малого объема (при отсутствии в выборке "загрязненных" членов). Так как уровень значимости исследуемых критериев, определенный по методу Монте-Карло, не превосходит номинальный, то можно считать, что качество нормально распределенных чисел довольно высокое. Вероятность совершения ошибки II-рода β оценивалась частотой $\frac{m}{N}$ с точностью не более 0,04 и надежностью 0,95, т.е.

$$P \left\{ \left| \beta - \frac{m}{N} \right| \leq 0.04 \right\} = 0.95$$

3. РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

Исследование мощностей $M = 1 - \beta$ статистических критериев произведено в трех вариантах. В 1,2,3 варианте исследовались восемь статистических критериев, основанных на тестовых статистиках $\varphi_1(x)$, $\varphi_2(x)$, $\varphi_3(x)$, $\varphi_4(x)$, $\varphi_5(x)$, $\varphi_6(x)$, $\varphi_7(x)$, $\varphi_8(x)$ при $\alpha = 0,005$; 0,01; 0,05; 0,10 и объеме выборки $n = 10, 15, 20$, при наличии в выборке только одного измерения, содержащего явно грубую ошибку, образованную кратковременным сдвигом центра распределения (согласно гипотезе H_1) или изменением мас-

штаба (согласно гипотезе H_2). Число выборок N объема n , на основании которых производилось оценивание, принято равным 1000, 750, 500 для 1, 2 и 3-го варианта соответственно.

Результаты моделирования представлены в виде графиков мощностей исследуемых критериев. Так, на рис. 2 и 3 изображены кривые мощностей исследуемых критериев, для $n=10$, $\alpha=0,01$. Кривые мощностей для других значений уровня значимости α и объема выборки n здесь не приведены, так как они носят аналогичный характер.

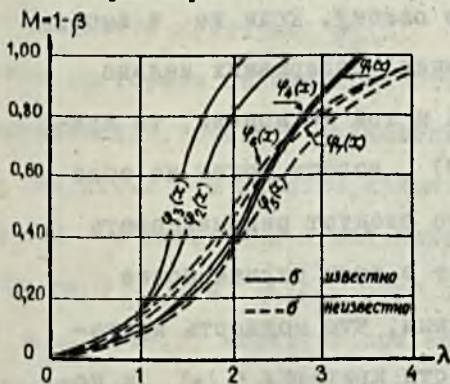


Рис. 2. Неявно грубая ошибка $x_{(n)}$ образована сдвигом центра распределения

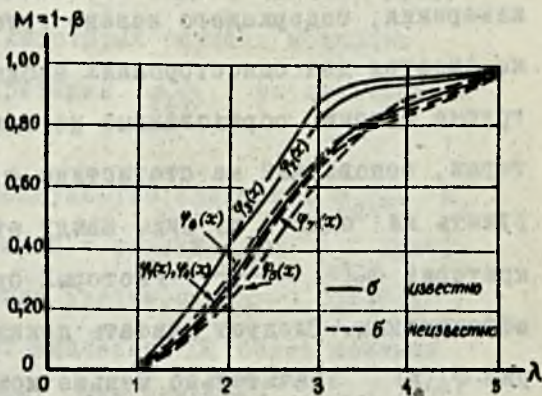


Рис. 3. Неявно грубая ошибка $x_{(n)}$ образована изменением масштаба

Анализируя рис. 2 и 3, приходим к выводу, что наиболее мощным, а следовательно, и наиболее эффективным критерием среди рассмотренных, при известной дисперсии измерений, является критерий, основанный на статистике $\varphi_3(x) = \frac{x_{(n)} - x_{(n-1)}}{6}$. Преимущества этого критерия сохраняются при любом уровне значимости α для $n=10, 15, 20$ и при двух моделях образования неявно грубой ошибки.

Мощность критерия, основанного на статистике $\varphi_2(x) = \frac{x^{(n)} - \bar{x}}{\sigma}$, близка к мощности критерия $\varphi_3(x) = \frac{x^{(n)} - x^{(n-1)}}{\sigma}$, а мощности критериев, основанных на статистиках $\varphi_1(x) = \frac{\sum (x_i - \bar{x})^2}{\sigma^2}$ и $\varphi_4(x) = \frac{x^{(n)} - x^{(1)}}{\sigma}$ одинаковы (в пределах точности вычислений), но значительно отличаются в худшую сторону от мощностей критериев $\varphi_3(x)$ и $\varphi_2(x)$.

Заметим, что при известном σ здесь исследовались мощности критериев при наличии в выборке только одного (максимального) измерения, содержащего неявно грубую ошибку. Если же в выборке имеются два односторонних наблюдения, содержащих неявно грубые ошибки, образованные по одной и той же модели, то критерий, основанный на статистике $\varphi_3(x)$, вообще может не обнаружить ни одного из них. Ввиду этого следует рекомендовать критерий $\varphi_2(x) = \frac{x^{(n)} - \bar{x}}{\sigma}$, который будет в этом случае более эффективным. Следует считать доказанным, что мощность критерия $\varphi_4(x)$ значительно меньше мощности критерия $\varphi_2(x)$ и потому критерий $\varphi_2(x)$ рекомендуется применять для установления технических допусков.

В случае неизвестного σ фактически анализировались мощности только трех статистических критериев значимости, так как тестовая статистика критерия $\varphi_8(x)$ согласно [10], выражается через тестовую статистику $\varphi_6(x)$:

$$\varphi_8(x) = 1 - \frac{\varphi_6^2(x)}{n-1}$$

Поэтому мощности критериев, основанных на этих статистиках, одинаковы.

Анализируя рис. 2 и 3, приходим к выводу, что мощности исследуемых статистических критериев близки между собой. Крите-

рий, основанный на статистике $\varphi_6(x)$, который по теоретическим исследованиям [11] является наилучшим инвариантным критерием для всех значений λ , в большинстве случаев сохраняет свои преимущества перед другими критериями.

Мощность статистического критерия $\varphi_7(x)$ незначительно превышает мощность критерия $\varphi_5(x)$. Результаты вычислений мощностей всех исследуемых статистических критериев приводят к выводу: мощности статистических критериев $\varphi_5(x)$ и $\varphi_7(x)$ очень незначительно отличаются в худшую сторону (не более, чем на 0,06) от мощности критерия $\varphi_6(x)$ в некоторых случаях мощность этих критериев больше мощности критерия $\varphi_6(x)$ на величину, не превышающую точность вычислений.

На рис. 2 и 3 заметно явное преимущество критериев $\varphi_2(x)$ и $\varphi_3(x)$ с известным b над критериями с неизвестным b . Критерии же $\varphi_6(x)$ и $\varphi_7(x)$ при уровне значимости $\alpha = 0,005 + 0,10$ и $n = 10, 15, 20$ являются при малых значениях λ более мощными (несмотря на то, что b неизвестно), чем критерии $\varphi_1(x)$ и $\varphi_4(x)$.

По этим соображениям, при наличии в выборке только одного подозреваемого измерения, следует рекомендовать (учитывая простоту вычисления тестовой статистики) критерий Диксона

$$\varphi_5(x) = \frac{x_{(n)} - x_{(n-1)}}{x_{(n)} - x_{(2)}} \quad (\text{при } n = 8 \div 11)$$

$$\varphi_5'(x) = \frac{x_{(n)} - x_{(n-2)}}{x_{(n)} - x_{(3)}} \quad (\text{при } n = 14 \div 30)$$

Критерий $\varphi_6(x) = \frac{x_{(n)} - \bar{x}}{s}$ следует применять для решения спорных вопросов, когда эмпирическое значение тестовой статистики

$\varphi_5(x)$ близко к табличному значению.

Ввиду сложности вычисления тестовых статистик $\varphi_7(x)$ и $\varphi_8(x)$ при одинаковых мощностях, эти критерии применять не рекомендуется.

Рис. 4 и 5 отображают изменение мощности статистического критерия $\varphi_3(x) = \frac{x_{(n)} - x_{(n-1)}}{\sigma}$, вызванное изменением уровня значимости от 0,5 до 10% при двух моделях образования неявно грубой ошибки.

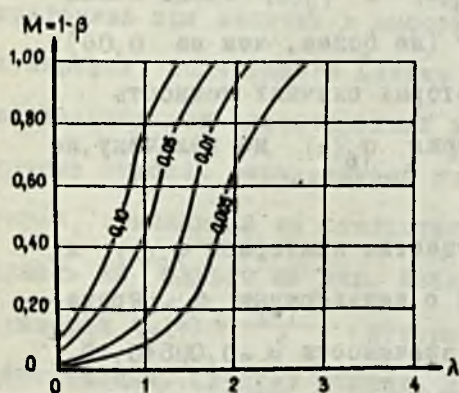


Рис. 4. Неявно грубая ошибка $x_{(n)}$ образована сдвигом центра распределения

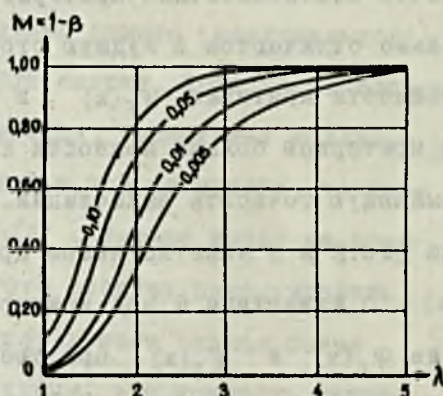


Рис. 5. Неявно грубая ошибка $x_{(n)}$ образована изменением масштаба

Анализируя их, замечаем, что незначительное изменение уровня значимости волизи нуля (от 0,5% до 1%) дает такой же прирост мощности критерия, что и изменение уровня значимости данного критерия от 5 до 10%. Аналогично ведут себя и другие критерии при $n = 10, 15, 20$.

Эти рекомендации справедливы, если в выборке имеется только одно измерение, содержащее грубую ошибку. Если в выборке существует два односторонних измерения, содержащих грубые ошибки, то некоторые критерии обладают маскирующим эффектом,

который обнаружен и описан Пирсоном и Чандра Секар в [12]. Суть эффекта состоит в том, что при наличии в выборке более одного измерения, содержащего неявно грубую ошибку, некоторые критерии вообще не выявляют ни одного из этих измерений. Маскирующим эффектом обладают в некоторой степени все статистические критерии. Нами исследовался маскирующий эффект, возникающий из-за присутствия в выборке двух односторонних максимальных измерений, содержащих неявно грубые ошибки у четырех статистических критериев

$$\varphi_5''(x) = \frac{x_{(n)} - x_{(n-2)}}{x_{(n)} - x_{(1)}} ; \quad \varphi_6(x) = \frac{x_{(n)} - \bar{x}}{s}$$

$$\varphi_7(x) = \frac{\sum (x_i - \bar{x})^3}{ns^3} ; \quad \varphi_8(x) = \frac{\sum_{i=1}^{n-2} (x_i - \bar{x}_{n,n-1})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

Выбор только этих критериев для исследования маскирующего эффекта обусловлен тем, что исследуемые ранее статистические критерии теоретически обоснованы для проверки только одного максимального (минимального) измерения.

Так как в практике может оказаться большое число "загрязненных" измерений, то мы исследовали при неизвестном b наилучшую статистику $\varphi_6(x) \frac{x_{(n)} - \bar{x}}{s}$ и ввели новые, теоретически обоснованные для проверки двух односторонних "загрязненных" измерений, т.е. статистики $\varphi_5''(x)$, $\varphi_7(x)$ и $\varphi_8(x)$. Модели с числом "загрязненных" членов больше двух не рассматривались, так как, на наш взгляд, при числе "загрязненных" членов больше 10% целесообразность применения статистических критериев для отбора измерений вызывает сомнение.

Исследования произведены для $\alpha = 0,10; 0,05; 0,01; 0,005$

при $n = 15$, $N_{15} = 750$.

Для наглядности кривые мощностей исследуемых критериев при двух моделях загрязнения изображены на рис.6 и 7.

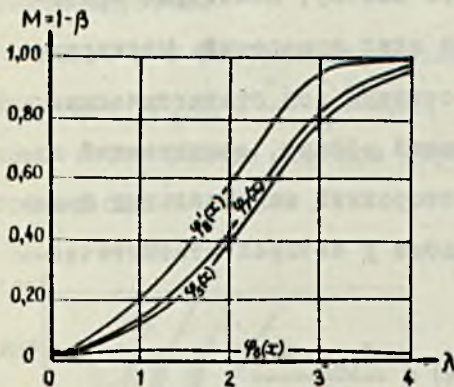


Рис.6. Неявно грубые ошибки $x_{(n)}$ и $x_{(n-1)}$ образованы сдвигом центра распределения

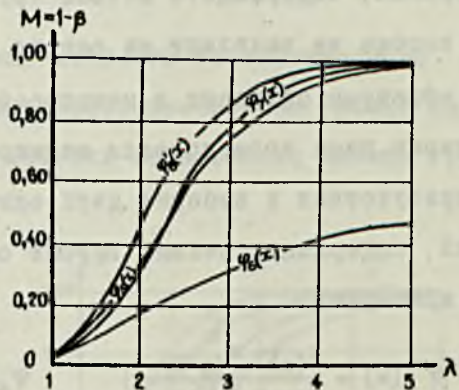


Рис.7. Неявно грубые ошибки $x_{(n)}$ и $x_{(n-1)}$ образованы изменением масштаба

Как видно из рис.6, применение критерия $\varphi_6(x) = \frac{x_{(n)} - \bar{x}}{s}$ для отбора измерений, при наличии в выборках двух односторонних максимальных измерений, содержащих неявно грубые ошибки, образованные сдвигом центра распределения при $\alpha = 0,01$, нецелесообразно, так как он не выявляет таких измерений, как бы не велико было λ . В случае образования ошибки "изменением масштаба" мощность критерия $\varphi_6(x)$ очень мала.

Наиболее мощным в этом случае является критерий $\varphi_8^*(x) = \frac{\sum_{i=1}^{n-1} (x_i - \bar{x}_{n,n-1})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$. Преимущества этого критерия сохраняются также при $\alpha = 0,10; 0,05$ и $0,005$. Мощности критериев $\varphi_5^*(x)$ и $\varphi_7(x)$ примерно одинаковы.

В этой связи отметим преимущество критерия Диксона $\varphi_5(x)$ [8], заключающееся в наличии ряда статистик r_{10}, r_{11}, r_{12} ,

r_{20} , r_{21} , r_{22} • что позволяет выбрать такую, в которую будет входить только одно из k подозреваемых резко отклоняющихся измерений, и таким образом практически свести к нулю маскирующий эффект.

Таким образом, при наличии в выборке более одного, резко отклоняющегося измерения, целесообразно применять критерий Диксона, а для решения спорных вопросов, когда эмпирическое значение тестовой статистики $\varphi_5(x)$ меньше, но близко к табличному значению, можно рекомендовать критерий $\varphi_8(x)$.

Помимо статистических критериев нами в вариантах № 1, 2, 3 определялся уровень значимости α и определялась мощность 4-х вероятностных критериев - критерия Шовене и критериев II, III и IV порядка, предложенных З.С.Хаимовым. Критерий Шарлье и критерий I-го порядка З.С.Хаимова не исследовались ввиду того, что они фактически не дают правила проверки подозреваемого измерения, а указывают только теоретическое число ошибок, превосходящих фиксированный предел, причем первый вообще не ограничивает величину предельной ошибки, а второй ограничивает пределом $\pm 3\sigma$.

В таблице 1 даны уровни значимости указанных критериев, вычисленные по методу Монте-Карло.

Таким образом, критерий Шовене или критерий З.С.Хаимова II порядка при числе измерений $n < 20$ будет браковать в среднем одно измерение из 5 проверяемых, которые в действительности не содержат грубых ошибок.

Результаты моделирования показывают, что уровень значимости исследуемых вероятностных критериев $\alpha > 0,05$, поэтому их не следует применять для отбора измерений при $n < 20$. Как исклю-

чение, ввиду практической распространенности вероятностных критериев, можно рекомендовать применять критерий IV порядка, предложенный З.С.Хаимовым.

Таблица 1

	Критерии	Уровень значимости		
		n = 10	n = 15	n = 20
1	Шовене	0,23	0,18	0,18
2	Хаимова II порядка	0,26	0,21	0,21
3	Хаимова III порядка	0,09	0,14	0,14
4	Хаимова IV порядка	0,07	0,08	0,10

4. ЗАКЛЮЧЕНИЕ

В заключение изложим некоторые соображения о целесообразности применения статистических критериев для анализа распределения измерений. Известно, что устойчивость статистических критериев зависит от выполнения на практике исходных предпосылок, на которых они основаны, а именно: нормальности, независимости и стабильности основной массы измерений. Если нормальность и независимость измерений на практике в большинстве случаев выполняется (частичное невыполнение этих условий приводит к увеличению уровня значимости), то стабильность условий измерений выполняется не всегда. Например, высокоточные угловые измерения планируются на несколько дней, производится исполнителями разной квалификации и опыта, в разные часы су-

ток. Казалось бы, что здесь сводит к нулю целесообразность применения статистических критериев. Однако даже в этом случае их применение целесообразно, так как с их помощью можно произвести анализ условий измерений и, таким образом, проверить целесообразность и эффективность принятого метода измерений.

Помимо этого существует класс лабораторных измерений, производимых при стабильных условиях, например, исследование измерительных приборов, компарирование, где статистические критерии можно с успехом применять. В этом случае они могут быть использованы либо для выявления неявно грубых ошибок, либо для контроля стабильности условий измерений с целью выявления факторов, нарушающих стабильность.

Произведенный сравнительный анализ критериев с известным b позволяет отобрать наиболее эффективные статистические критерии для установления единых технических допусков на размах измерений $R = x_{(n)} - x_{(1)}$ или на отклонение двух крайних членов $\Delta = x_{(n)} - x_{(n-1)} = x_{(1)} - x_{(2)}$ вариационного ряда (7). Сравнительный анализ критериев с неизвестным b позволяет отобрать наилучшие в этом смысле критерии для установления дифференцированных технических допусков, косвенно позволяющих учитывать квалификацию, опыт исполнителя, а также изменение внешних условий измерений.

Литература

- [1] БОЛЬШЕВ Л.Н., СМЕРНОВ Н.В.: Таблицы математической статистики, М., Наука, 1955,

- [2] БОЛЬШАКОВ В.Д.: Теория ошибок наблюдений с основами теории вероятностей, М., Недра, 1965,
- [3] ГОЛЕНКО Д.И.: Моделирование и статистический анализ псевдослучайных чисел на электронных вычислительных машинах, М., Наука, 1965,
- [4] ЛИННИК Д.В.: Метод наименьших квадратов и основы теории обработки наблюдений, М., Физматгиз, 1962,
- [5] МАЛИКОВ С.Ф.; ТЮРИН Н.И.: Введение в метрологию, М., Стандартгиз, 1965,
- [6] ШЕЙНИН О.В.: К истории проблемы отбора (отбраковки) измерений, ВИНТИ, М., 1963,
- [7] DAVIS P., RABINOWITZ P.: Some Monte-Carlo Experiments in Computing Multiple Integrals, Math. Tables other Aids Comp., 10, 1956: 1-8
- [8] DIXON W.J.: Analysis of Extreme Values, Annals of Mathematical Statistics, 1950:21, 488-506
- [9] FERQUSON T.S.: On the Rejection of Outliers. Proceeding of the 4-th Berkeley Symposium on Mathematical Statistical and Probability, 1961:1, 253-281
- [10] GRUBBS F.E.: Sample Criteria for Testing Outlying Observation, Annals of Mathematical Statistics, 1950:21, 27-58
- [11] PAULSON E.: An Optimum Solution to the k-Sample Slippage Problem for the Normal Distribution, Annals of Mathematical Statistics, 1952:23, 610-616
- [12] PEARSON E.S., CHANDRA SECAR C.: The Efficiency of Statistical Tools and Criterion for Rejection of Outlying Observation, Biometrika, 1936:28, 308-320
- [13] TUKEY J.W.: The Future of Data Analysis, Annals of Mathematical Statistics, 1962:33, 1, 1-74
- [14] GUPTA A.K.: Estimation of the Mean and Standard Deviation of a Normal Population from a Censored Sample, Biometrika, 1952:39, 260-273
- [15] SARCHAN A.E., GREENBERG B.G.: Estimation of Location and Scale

Parameters by Order Statistics from Singly and Doubly Censored Samples, Part I, Annals of Mathem. Statistics, 1956:27, 427-451

- [16] ЧЕВОТАРЕВ А.С.: Способ наименьших квадратов с основами теории вероятностей, М., Геодезиздат, 1958,
- [17] ХАИМОВ З.С.: Критерии ничтожных и предельных погрешностей, Известия вузов, Геодезия и аэрофотосъемка, 1963.

ANALIZA PORÓWNAWCZA TESTÓW STATYSTYCZNYCH SŁUŻĄCYCH
DO WYKRYWANIA BŁĘDÓW DUŻYCH

Streszczenie

W pracy rozważa się kilka testów statystycznych służących do weryfikacji jednej i tej samej hipotezy o rozkładzie prawdopodobieństwa i porównuje się te testy metodą Monte Carlo.

A COMPARATIVE ANALYSIS OF STATISTIC CRITERIA PRESENTING
MEASUREMENTS INCLUDING BIG ERRORS

Summary

Curves of power of some statistic criteria presented to reject dubious observations are simulated by the Monte Carlo method on a Minsk-22 computer.

PRZEGLĄD KILKU METOD OPARTYCH NA
KONCEPCJI GRADIENTU EKSPERYMENTAL-
NEGO I PORÓWNANIE ICH EFEKTYWNOŚCI

Jan GOLIŃSKI

Stożeczny Ośrodek Elektronicznej
Techniki Obliczeniowej,
Warszawa, ul. Mollera 6

Przedstawiono kilka zrealizowanych algorytmów programowania nieliniowego na maszyny ZAM 2 oraz ZAM 41. Efektywność ich pokazano na przykładach. Niektóre z tych algorytmów mają szereg stopni swobody, które pozwalają różnie kształtować proces poszukiwania ekstremum. Od właściwego wyboru parametrów determinujących proces zależy czas, a zatem również i koszt poszukiwań. W pracy podano wyniki eksperymentów, które powinny umożliwić użytkownikom właściwy wybór tych stopni swobody.

WSTĘP

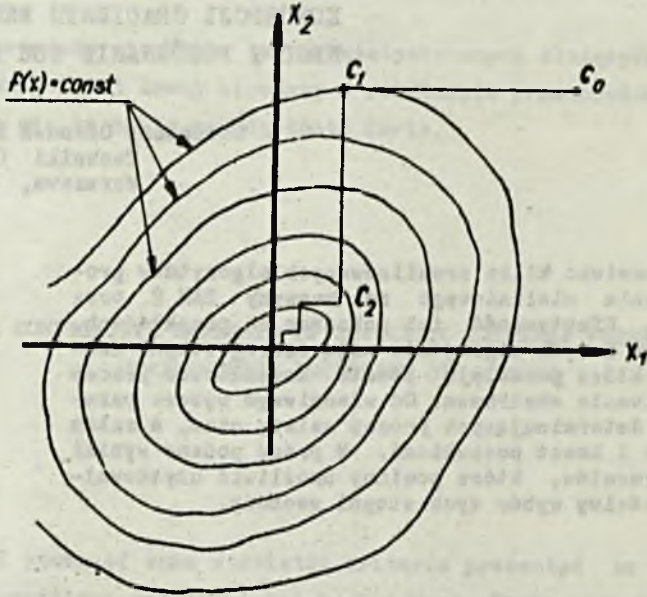
Metody optymalizacji nieliniowej znajdują coraz szersze zastosowanie w różnych dziedzinach nauki i techniki. Różnorodność metod stawia użytkownika przed wyborem właściwej metody, która pozwoliłaby najlepiej rozwiązać sformułowane zadanie.

W artykule przedstawiono kilka metod optymalizacyjnych i wyniki badań ich efektywności.

1. METODA GAUSSA-SEIDLA [1]

Polega na tym, że kierunek zmian szukanych parametrów jest zgodny z układem współrzędnych. Rysunek 1 ilustruje postępowanie dla przykładu 2 zmiennych. Poszukujemy minimum

znajdującego się w punkcie $(0,0)$. Zaczynamy postępowanie z punktu C_0 i kolejno zmieniamy współrzędne. Najpierw np. x_1 (tak długo, jak długo zmienia się wartość funkcji celu (aż do C_1)), a następnie x_2 - aż do C_2 , itd.



Rys. 1. Ilustracja postępowania przy metodzie Gaussa-Seidla

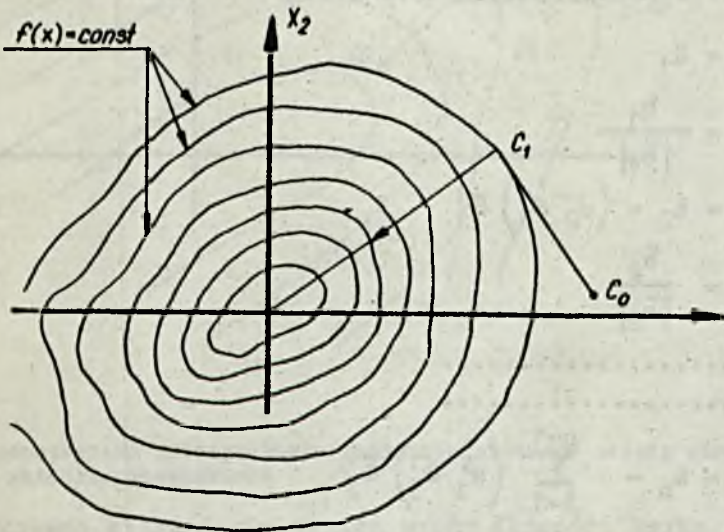
2. METODA ROSENBROCKA [2]

Jest modyfikacją metody Gaussa-Seidla. Modyfikacja polega na tym, że po każdym cyklu iteracyjnym, aby przyspieszyć proces szukania optymalnego punktu dokonuje się obrotu układu współrzędnych.

Korzyści jakie daje tego rodzaju modyfikacja ilustruje rys. 2. W przykładzie tym kierunek zmian poszukiwanych parametrów zgodny jest z kierunkiem głównych półosi elips odpowiadających krzywym o stałej wartości funkcji celu.

Metodę Rosenbrocka charakteryzuje względna prostota oraz dobra zbieżność.

Najbardziej charakterystyczną cechą tej metody jest wybór ortogonalnego układu wersorów \bar{S} , wzdłuż których następują zmiany szukanych parametrów. Przyjmujemy układ współrzędnych naturalnych jako układ wyjściowy. Pierwsza iteracja polega na szukaniu minimum przez zmienianie po kolei poszczególnych parametrów. Po kilku cyklach iteracyjnych wersor \bar{S}_1^k wskaże przybliżony kierunek zwrotu. Tak więc dla przykładu z rys. 2, dla którego poszukuje się minimum wskaże - grad $f(x)$.



Rys. 2. Ilustracja postępowania przy metodzie Rosenbrocka

Obrotu układu dokonujemy po tym, gdy w każdym z ortogonalnych kierunków dokonano przynajmniej po jednym dobrym kroku.

W celu wyznaczenia nowego układu wersorów tworzymy wektory

$$\bar{a}_1 = \alpha_1 s_1^0 + \alpha_2 s_2^0 + \dots + \alpha_n s_n^0$$

$$a_2 = \alpha_2 s_2^0 + \dots + \alpha_n s_n^0$$

.....

.....

$$\bar{a}_n = \dots + \alpha_n s_n^0$$

gdzie $s_1^0, s_2^0, \dots, s_n^0$ - wersory początkowego układu współrzędnych

α_i ($i = 1, 2, \dots, n$) odcinek charakteryzujący przesunięcie wzdłuż s_i

Nowe wersory $\bar{s}_1^1, \bar{s}_2^1, \dots, \bar{s}_n^1$ dla kolejnej iteracji wyznacza się z następujących zależności:

$$\bar{b}_1 = \bar{a}_1$$

$$\bar{s}_1^1 = \frac{\bar{b}_1}{|\bar{b}_1|}$$

$$\bar{b}_2 = \bar{a}_2 - (a_2^1 s_1^1) \bar{s}_1^1$$

$$\bar{s}_2^1 = \frac{\bar{b}_2}{|\bar{b}_2|}$$

.....

.....

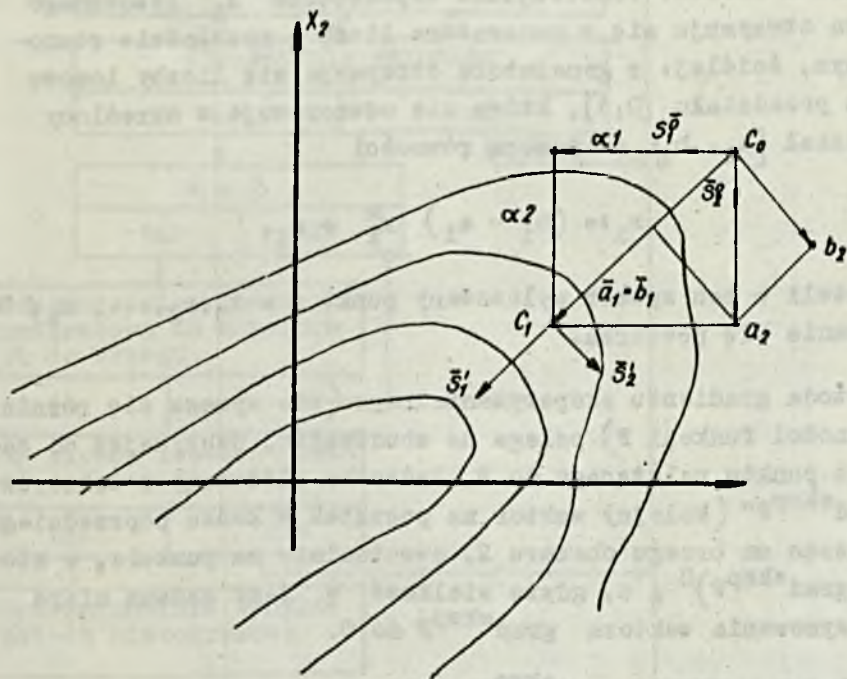
$$\bar{b}_n = \bar{a}_n - \sum_{j=1}^{n-1} (a_j^1 \bar{s}_j^1) \bar{s}_j^1$$

$$\bar{s}_n^1 = \frac{\bar{b}_n}{|\bar{b}_n|}$$

Po kilku cyklach iteracyjnych wersor \bar{s}_1^k wskaże przybliżony kierunek zwrotu ($-\text{grad } f(x)$).

Rys. 3 ilustruje sposób wyznaczania nowego układu wersorów dla przykładu dwuwymiarowego. Z punktu początkowego C_0 osiągamy punkt C_1 zmieniając x_1 o α_1 , a x_2 o α_2 . Wektor

$\bar{a}_1 = \bar{b}_1$ łączy punkt C_0 z C_1 . Jego normalizacja daje wektor \bar{S}_1 . Wektor \bar{a}_2 równy jest $\alpha_2 S_2^0$. Jego rzut na oś prostopadłą do \bar{S}_1 daje wektor \bar{b}_2 , który wyznacza wektor \bar{S}_2 .



Rys. 3. Geometryczna interpretacja znajdowania nowego układu wektorów w metodzie Rosenbrocka

Dodatkowego wyjaśnienia wymaga wybór długości kroku, którym poruszamy się wzdłuż współrzędnych. W metodzie Rosenbrocka rozpoczynamy postępowanie od dowolnej wartości e . Jeżeli krok jest dobrze wybrany, tzn. nowa wartość funkcji celu jest lepsza od poprzedniej, to krok β razy wydłużamy ($\beta > 1$). W przypadku przeciwnym mnożymy e przez γ ($\gamma < 1$).

3. ALGORYTM MIESZANY LOSOWO-GRADIENTOWY

Korzysta jak to z nazwy wynika z dwóch metod.

Metody losowej, polegającej w tym przypadku na kolejnym porównywaniu wartości funkcji $F(x)$ w punktach wybranych z obszaru R losowo. Poszczególne współrzędne x_1 losowanego punktu otrzymuje się z generatora liczb o rozkładzie równomiernym, ściślej: z generatora otrzymuje się liczby losowe x_1^m z przedziału $[0, 1]$, które się odwzorowuje w określony przedział $[a_1, b_1]$ za pomocą równości

$$x_1 := (b_1 - a_1) x_1^m + a_1.$$

Jeżeli w ten sposób wylosowany punkt $x = x_1, x_2, \dots, x_n \notin R$, losowanie się powtarza.

Metoda gradientu eksperymentalnego (nie wymaga się różniczkowalności funkcji F) polega na zbudowaniu, poczynając od danego punktu należącego do R , łańcucha złożonego z wektorów " $-\text{grad}^{\text{eksp}_F}$ " (kolejny wektor ma początek w końcu poprzedniego), a kończąc na brzegu obszaru R , ewentualnie na punkcie, w którym $(\text{grad}^{\text{eksp}_F})^2 \leq \epsilon$, gdzie wielkość ϵ jest zadaną miarą aproksymowania wektora $\text{grad}^{\text{eksp}_F}$ do 0.

Składowe wektora $\text{grad}^{\text{eksp}_F}$ obliczane są za pomocą wzoru

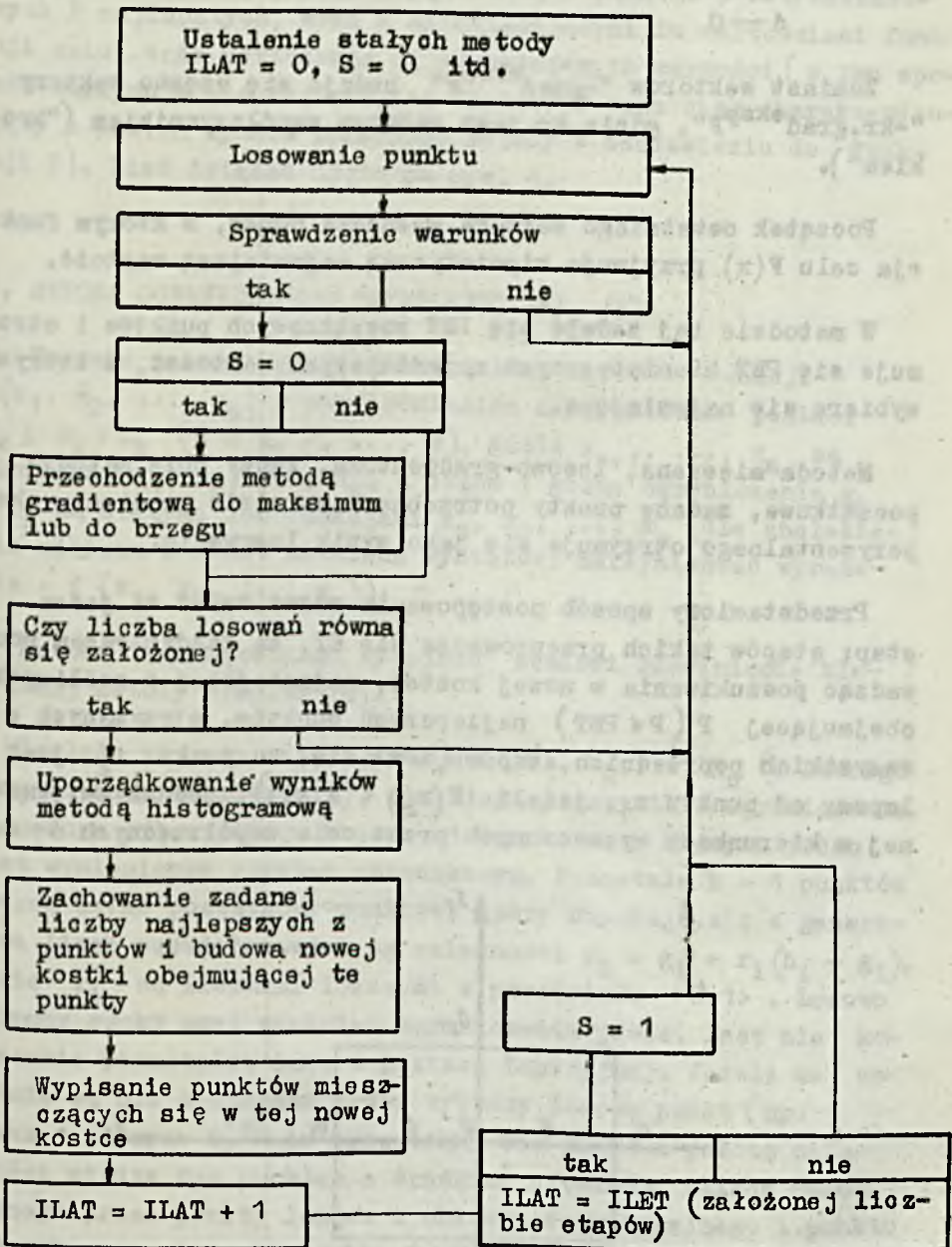
$$\text{grad}_1^{\text{eksp}_F}(x) := \frac{F(x + d \cdot v_1) - F(x - d \cdot v_1)}{2d}$$

gdzie:

$$x = (x_1, x_2, \dots, x_1, \dots, x_n),$$

$$v_1 = [0, 0, \dots, 1, \dots, 0] \text{ jest wektorem } i\text{-tej osi współrzędnych, zaś } d \text{ jest zadaną (małą) liczbą.}$$

Gdyby funkcja celu $F(x)$ była różniczkowalna, wtedy zachodziłaby równość



Rys. 4. Sieć działań dla metody kombinowanej - gradientowo-łosowej

$$\lim_{d \rightarrow 0} \text{grad}^{\text{eksp}} F(x) = \text{grad } F(x)$$

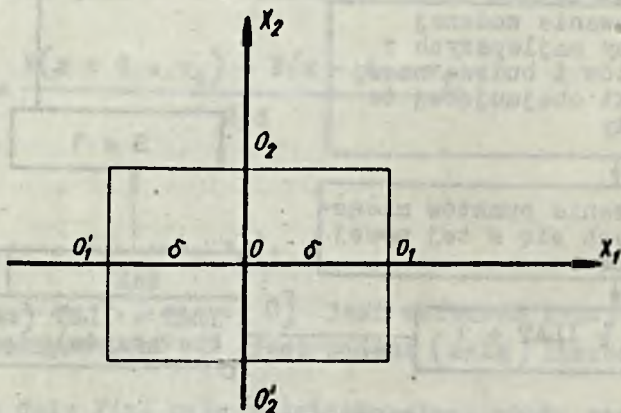
Zamiast wektorów " $-\text{grad}^{\text{eksp}} F$ " buduje się często wektory " $-\text{kr. grad}^{\text{eksp}} F$ ", gdzie kr jest zadany współczynnikiem ("krokiem").

Początek ostatniego wektora wyznacza punkt, w którym funkcja celu $F(x)$ przyjmuje hipotetyczną najmniejszą wartość.

W metodzie tej zadaje się PET początkowych punktów i otrzymuje się PET hipotetycznych najmniejszych wartości, z których wybiera się najmniejszą.

Metoda mieszana, losowo-gradientowa, łączy obie metody: początkowe, zadane punkty potrzebne w metodzie gradientu eksperymentalnego otrzymuje się jako wynik losowania.

Przedstawiony sposób postępowania można uznać za jeden etap; etapów takich przeprowadza się ET, za każdym razem prowadząc poszukiwania w nowej kostce, najmniejszej z możliwych, obejmującej P ($P \in \text{PET}$) najlepszych punktów, otrzymanych z wszystkich poprzednich etapów (mówi się, że punkt x_2 jest lepszy od punktu x_1 , jeżeli $F(x_2) < F(x_1)$), jednak zwiększonej w kierunkach wyznaczonych przez osie współrzędnych o d .



Rys. 5. Ilustracja obliczania gradientu eksperymentalnego

Na końcu wybiera się PAL ($PAL \in P$) z punktów z dotychczasowych P najlepszych, wraz z odpowiadającymi im wartościami funkcji celu, oraz porządkuje wg. wzrastających wartości (w ten sposób otrzymuje się najmniejszą wartość F oraz ciąg charakteryzujący w pewien sposób zbieżność metody w odniesieniu do funkcji F). Sieć działań ujęto na rys. 4.

4. METODA OGRANICZONEGO "SYMPLEXU" [4], [5]

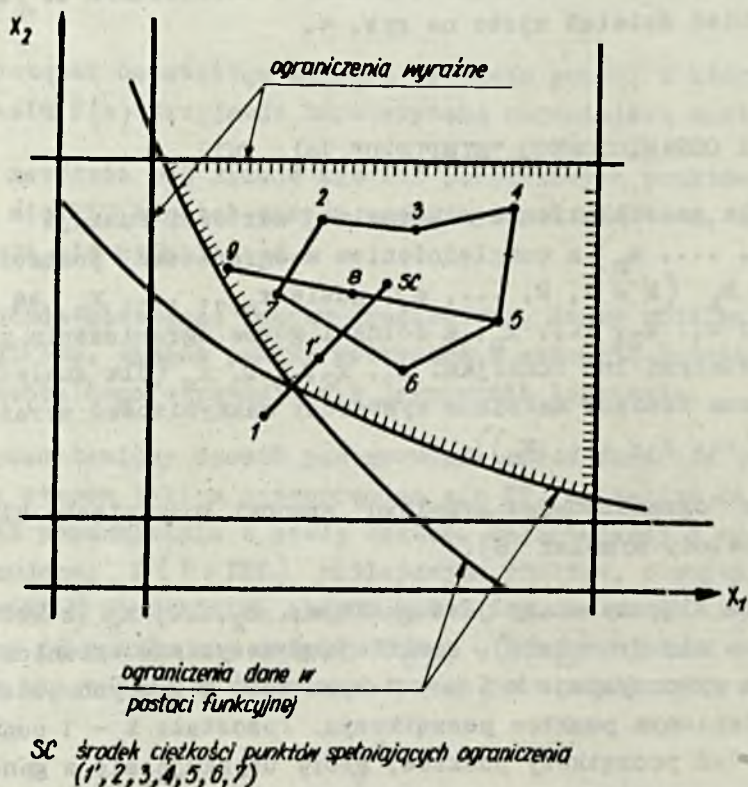
Pozwala na znalezienie maksymalnej wartości funkcji $f(x_1, x_2, \dots, x_n)$ z uwzględnieniem m ograniczeń postaci $g_k \leq x_k \leq h_k$ ($k = 1, 2, \dots, m$), gdzie x_{n+1}, \dots, x_m są funkcjami x_1, x_2, \dots, x_n , a dolne i górne ograniczenia g_k i h_k są stałymi lub funkcjami x_1, x_2, \dots, x_n (dla znalezienia minimum zamiast maksimum wystarczy maksymizować wyrażenie $-f(x_1, x_2, \dots, x_n)$).

Metoda "ograniczonego sympleksu" stanowi rozwinięcie klasycznej metody sympleks [6].

Zakłada się, że znany jest punkt x_1, x_2, \dots, x_n (z którego rozpoczyna się procedurę), spełniający wszystkie ograniczenia. Procedura wykorzystuje $k > n + 1$ punktów, z których jeden jest wymienionym punktem początkowym. Pozostałe $k - 1$ punktów tworzą układ początkowy punktów, który uzyskuje się z generatora liczb pseudolosowych wg. zależności $x_i = g_i + r_i(h_i - g_i)$, gdzie r_i są liczbami losowymi z przedziału $\langle 0, 1 \rangle$. Losowo wybrany punkt musi spełniać ograniczenia jawne, lecz nie koniecznie pozostałe, dane w postaci funkcyjnej. Jeżeli te ostatnie są nie spełnione przez wybrany losowo punkt (np. punkt 1 na rys. 6)^{m)} to przesuwamy ten punkt o połowę odległości między tym punktem a środkiem ciężkości układu utworzonego przez punkty losowe z obszaru dopuszczalnego i punktu początkowego, tzn. do punktu 1. Powtarzamy to postępowanie aż

^{m)} Teraz nastąpi opis postępowania, które ilustruje rys. 6

do trafienia na punkt, który leży w obszarze. Przy założeniu, że obszar dopuszczalny jest wypukły, na pewno trafi się przy tym postępowaniu w końcu na punkt spełniający wszystkie warunki ograniczające.



Rys. 6. Postępowanie przy metodzie "ograniczonego sympleksu" na przykładzie dwóch zmiennych

Funkcja celu jest obliczana w każdym wierzchołku figury geometrycznej opisanej przez wybrane punkty losowe (tzn. w punktach 0, 1, 2, 3, 4, 5, 6).

Wierzchołek tzn. punkt, w którym wartość funkcji celu jest najmniejsza, np. punkt 5, zastępuje się przez inny leżący na prostej przechodzącej przez ten "najgorszy punkt" (5) i środek ciężkości figury utworzonej przez pozostałe wybrane punk-

ty (8). Ten nowy (9) leży po przeciwnej stronie środka ciężkości w odległości α razy większej ($\alpha > 1$).

Jeżeli tak znaleziony punkt nie spełnia funkcyjnych warunków ograniczających, wówczas dzieli się odcinek między tym punktem a środkiem ciężkości na połowę, sprawdza się ponownie warunki, powtarzając postępowanie aż do uzyskania dobrego.

Jeżeli natomiast wylosowany punkt nie spełnia ograniczenia nałożonego na zmienne niezależne x_i , $i = 1, 2, \dots, n$ wówczas jest zastępowany przez punkt leżący na prostej przechodzącej przez wylosowany punkt i równoległy do osi losowanej współrzędnej w odległości 0.000001 od brzegu obszaru dopuszczalnego i wewnątrz tego obszaru.

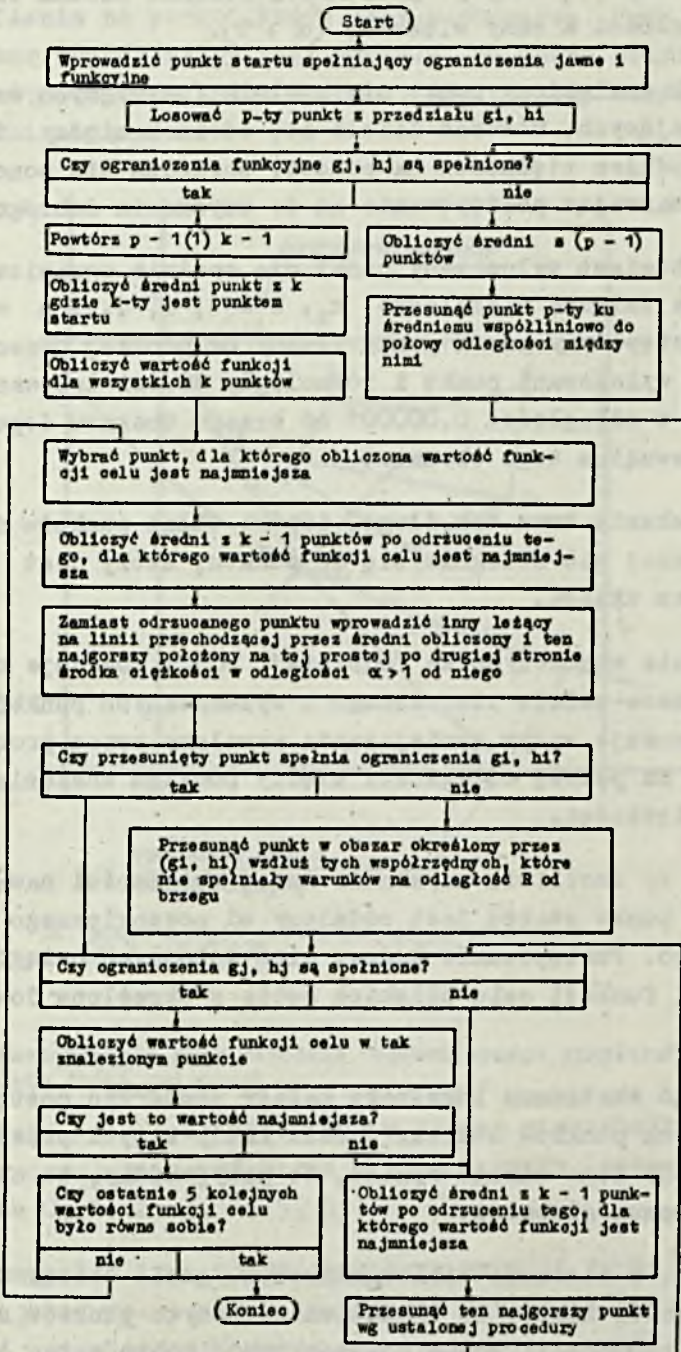
Proces szukania trwa tak długo, dopóki układ punktów figury geometrycznej nie zbliżnie się do punktu, który jest bieżącym środkiem układu.

Wprowadzenie współczynnika "odbicia" $\alpha > 1$ powoduje ciągle powiększanie układu utworzonego z wylosowanych punktów, a zatem kompensuje wpływ zmniejszania wywołany przez procedurę dzielenia na połowę odległości między punktem znalezionym i środkiem ciężkości,

Działanie to umożliwia uzyskanie dużej zbieżności nawet wtedy, kiedy punkt startu jest oddalony od poszukiwanego punktu optymalnego. Postępowanie kończy pięć kolejnych osiągniętych wartości funkcji celu bliskich sobie z określoną dokładnością.

Aby uniknąć ekstremum lokalnego należy powtórzyć postępowanie z różnych punktów startu. Jeżeli każdy z tych przebiegów prowadzi do tego samego wyniku, to przyjmujemy, że osiągnęliśmy ekstremum globalne.

Zdarza się, w obszarze wielowymiarowym, gdzie obszar dopuszczalny jest małej miary, że znalezienie różnych punktów startu następuje z trudnością. Można wówczas pomóc sobie w ten spo-



Rys. 7. Sieć działań dla metody "ograniczonego sympleksu"

sób, że postępowanie zaczynamy z tego samego punktu początkowego, natomiast zmieniamy początek ciągu generowanych liczb pseudolosowych.

Siedź działań "metody ograniczonego sympleksu" podano na rys. 7.

5. PRZYKŁADY

Dla porównania metod przeprowadzono eksperymenty na różnych przykładach. Oprócz metod opisanych w [1, 2, 3, 4] wykorzystano również z surowej metody Monte Carlo [3].

5.1. Przykład 1

Minimalizowano funkcję

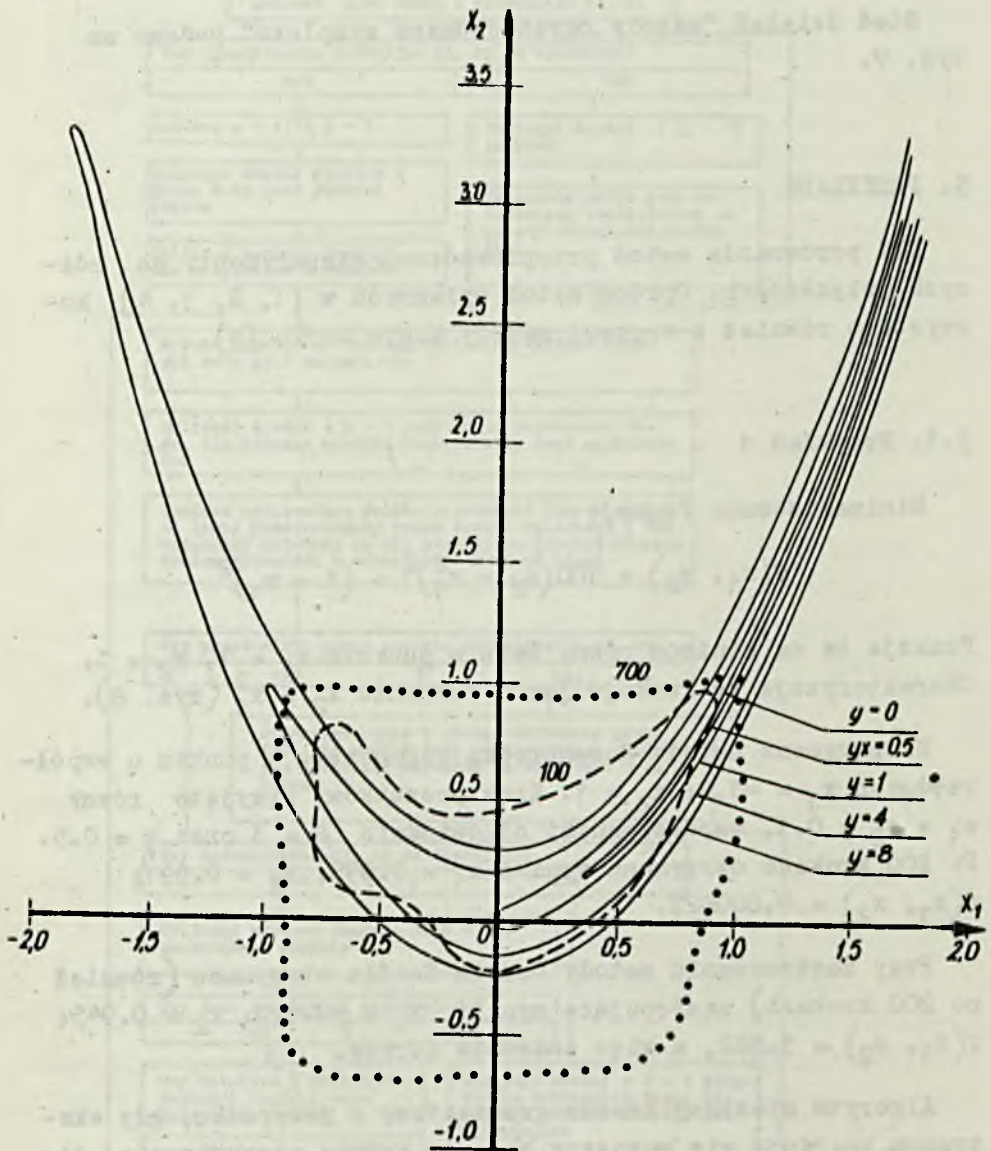
$$f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$$

Funkcja ta ma minimum równe zeru w punkcie $x_1 = 1, x_2 = 1$. Charakteryzuje tę funkcję jar o równaniu $x_2 = x_1^2$ (rys. 8).

Eksperyment metodą Rosenbrocka rozpoczęto z punktu o współrzędnych $x_1 = -1,2; x_2 = 1$. Krok początkowy przyjęto równy $e_1 = e_2 = 0,1$, współczynniki odpowiednio $\beta = 3$ oraz $\gamma = 0,5$. Po 200 krokach otrzymano wyniki $x_1 = 0,995, x_2 = 0,991$; $f(x_1, x_2) = 0,000022$.

Przy zastosowaniu metody Gaussa-Seidla otrzymano (również po 200 krokach) następujące wyniki $x_1 = -0,970; x_2 = 0,945$; $f(x_1, x_2) = 3,882$, a więc znacznie gorsze.

Algorytm mieszany losowo-gradientowy w przypadku, gdy ekstremum znajduje się wewnątrz obszaru kończy poszukiwanie, jeżeli pierwiastek z sumy kwadratów gradientów składowych jest mniejszy od zadanej tolerancji. L. Najlepsze wyniki z próby, która trwała 15 minut to: $x_1 = 0,5433; x_2 = 0,2752$;



Rys. 8. Graficzne przedstawienie przykładu $f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$

$f(x_1, x_2) = 0.2483$. Obszar poszukiwań wyznaczony dla 2 eksperymentów $L = 700$ i $L = 100$ podano na rys. 8 liniami przerywanymi.

Metoda ta charakteryzuje się dużym rozrzutem punktów, dla których funkcja celu jest bliska poszukiwanemu optimum. Okazuje się, że jej wykorzystanie dla zadania, w którym ekstremum znajduje się wewnątrz obszaru następuje pewne trudności, związane z doбором właściwego kroku i oszacowaniem wartości gradientu, który determinuje zakończenie obliczeń.

Surową metodą Monte Carlo osiągnięto wyniki ujęte w tabeli 1. Wskazują one, że dla tego rodzaju zadania zaskakująco efektywną okazuje się właśnie ta metoda. Porównanie wyników uzyskanych różnymi metodami ujęto w tabeli 2, która wskazuje na małą przydatność metod: kombinowanej i Gaussa-Seidla.

Tabela 1

F - minimum	Czas liczenia	Losowanie Nr
0.163965	0.55.02	68
0.052929	1.01.01	79
0.001958	1.19.07	277
0.000990	26.20.00	15 731
0.000061	45.00.00	26 895

Tabela 2

Metoda Współ.	Gauss-Seidel	Rosenbrock	Gradiento- wo-losowa	Losowanie
x_1	- 0.97	0.995	0.5433	0.996
x_2	0.945	0.991	0.2752	0.992
F	3.882	0.000022	0.2483	0.00006

5.2. Przykład 2

Znaleźć maksimum funkcji f , 5 zmiennych przy spełnieniu 8 ograniczeń.

$$f = (a_2 y_1 + a_3 y_2 + a_4 y_3 + a_5 y_4 + 7840a_6 - 100000a_0 + \\ - 50800ba_7 + k_{31} + k_{32}x_2 + k_{33}x_3 + k_{34}x_4 + k_{35}x_5)x_1 - \\ + 24345 + a_1x_6$$

gdzie:

$$b = x_2 + 0.01 x_3$$

$$x_6 = (k_1 + k_2x_2 + k_3x_3 + k_4x_4 + k_5x_5) x_1$$

$$y_1 = k_6 + k_7x_2 + k_8x_3 + k_9x_4 + k_{10}x_5$$

$$y_2 = k_{11} + k_{12}x_2 + k_{13}x_3 + k_{14}x_4 + k_{15}x_5$$

$$y_3 = k_{16} + k_{17}x_2 + k_{18}x_3 + k_{19}x_4 + k_{20}x_5$$

$$y_4 = k_{21} + k_{22}x_2 + k_{23}x_3 + k_{24}x_4 + k_{25}x_5$$

$$x_7 = (y_1 + y_2 + y_3) x_1$$

$$x_8 = (k_{26} + k_{27}x_2 + k_{28}x_3 + k_{29}x_4 + k_{30}x_5) x_1 + x_6 + x_7$$

x_1, x_2, x_3, x_4, x_5 są zmiennymi niezależnymi, a poszukiwane ekstremum musi spełniać warunki:

$$0 < x_1 \\ 1.2 < x_2 < 2.4 \\ 20 < x_3 < 60$$

$$\begin{aligned}
 9 &< x_4 < 9,3 \\
 6.5 &< x_5 < 7.0 \\
 0 &< x_6 < 294000 \\
 0 &< x_7 < 294000 \\
 0 &< x_8 < 277200
 \end{aligned}$$

wartości stałych a_i oraz k_i podano w tabelach 3 i 4.

Tabela 3

Symbol	Wartość
a_0	9
a_1	15
a_2	50
a_3	9.583
a_4	20
a_5	15
a_6	6
a_7	0.75

Optymalna wartość funkcji celu jest dokładnie znana $f = 5\ 280\ 334$, co odpowiada punktowi o współrzędnych:

$$x_1 = 4.53743$$

$$x_2 = 2.4$$

$$x_3 = 60$$

$$x_4 = 9.3$$

$$x_5 = 7.0$$

Metodą Rosenbrocka [2] przeprowadzono eksperyment przyjmując za punkt początkowy:

$$x_1 = 2.52$$

$$x_2 = 2.$$

Tabela 4

Symbol	Wartość	Symbol	Wartość	Symbol	Wartość	Symbol	Wartość
k ₁	- 145 421.402	k ₂	- 2 931.1506	k ₃	- 40.42793	k ₄	5 106.192
k ₅	+ 15711.36	k ₆	- 161622.577	k ₇	4.176.1533	k ₈	2.825008
k ₉	+ 9200.476	k ₁₀	13160.295	k ₁₁	- 21686.9194	k ₁₂	123.5693
k ₁₃	- 21.11889	k ₁₄	706.834	k ₁₅	2898.573	k ₁₆	28 298.388
k ₁₇	60.81096	k ₁₈	31.24212	k ₁₉	329.574	k ₂₀	- 2882.082
k ₂₁	74095.385	k ₂₂	- 306.2625	k ₂₃	16.24365	k ₂₄	- 3094.252
k ₂₅	- 5566.2628	k ₂₆	- 26.237	k ₂₇	99.	k ₂₈	- 0.42
k ₂₉	1 300.	k ₃₀	2 100.	k ₃₁	925 548.252	k ₃₂	- 61968.8432
k ₃₃	23.3088196	k ₃₄	- 27097.648	k ₃₅	- 50843.766		

$$x_3 = 37.5$$

$$x_4 = 9.25$$

$$x_5 = 6.8$$

dla którego wartość funkcji celu $f = 2351244$.

Najlepszy wynik osiągnięto w 689 próbie. Mimo, że prowadzono eksperyment aż do próby 1388 nie uzyskano już dalszej poprawy wyniku.

Otrzymane wyniki

$$x_1 = 4.58353$$

$$x_2 = 2.21545$$

$$x_3 = 31.77327$$

$$x_4 = 9.29997$$

$$x_5 = 6.99863$$

dla którego $f(x_1, \dots, x_8) = 5\,222\,459$.

Surową metodą Monte Carlo uzyskano $f = 4\,659\,433$ po 954 próbach, z których tylko niewiele było w obszarze dopuszczalnym.

Metodą "ograniczonego sympleksu" [4] osiągnięto wartość 5 280 327 i współrzędne punktu równe

$$x_1 = 4.537459$$

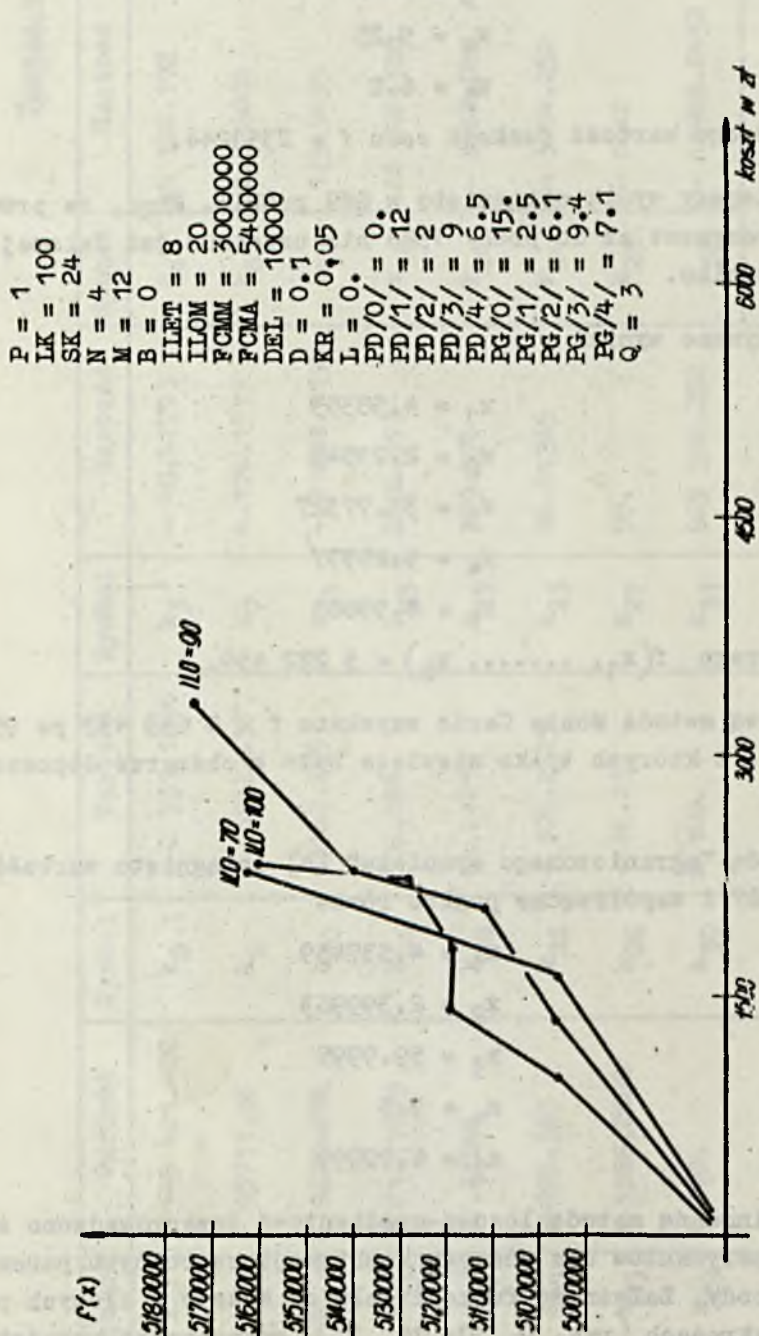
$$x_2 = 2.399963$$

$$x_3 = 59.9995$$

$$x_4 = 9.3$$

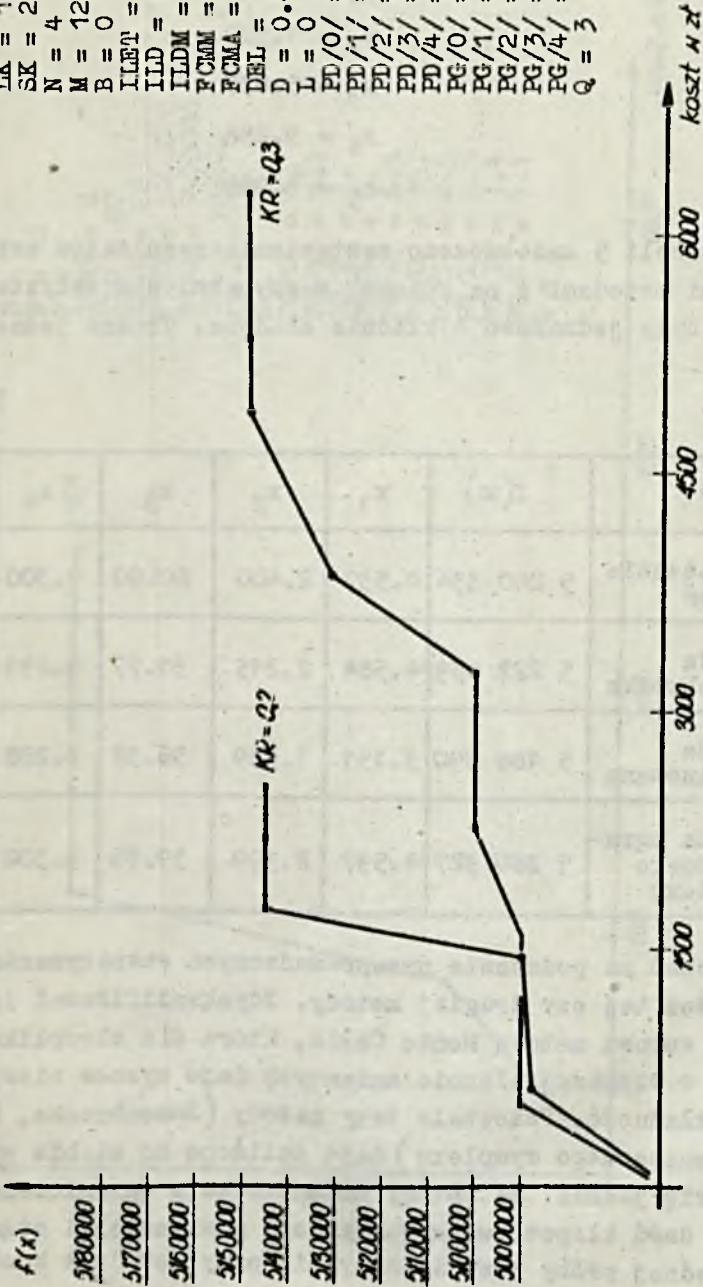
$$x_5 = 6.99999$$

Kombinowaną metodą losowo-gradientową przeprowadzono szereg eksperymentów dla różnych kombinacji swobodnych parametrów metody. Zależność funkcji celu od kosztu w złotych podano na wykresach (rys. 9, 10, 11, 12), zaznaczając przyjęte wartości stopni swobody algorytmu.



Rys. 9. Zależność osiągniętej funkcji celu od kosztu liczenia dla przykładu 5 zmiennych dla różnych wartości parametru ILO

$P = 1$
 $LK = 100$
 $SK = 24$
 $N = 4$
 $M = 12$
 $B = 0$
 $ILFT = 8$
 $ILD = 80$
 $ILDM = 20$
 $FCM = 3000000$
 $FCMA = 6400000$
 $DEL = 100000$
 $D = 0.4$
 $L = 0$
 $PD/0/ = 0$
 $PD/1/ = 1.2$
 $PD/2/ = 2$
 $PD/3/ = 9$
 $PD/4/ = 6.5$
 $PG/0/ = 15$
 $PG/1/ = 2.5$
 $PG/2/ = 6.1$
 $PG/3/ = 9.4$
 $PG/4/ = 7.1$
 $Q = 3$



Rys. 10. Zależność osiągniętej funkcji celu od kosztu liczenia dla przykładu 5 zmiennych dla różnych wartości parametru KR

Najlepszy osiągnięty wynik to:

$$f = 51868 \text{ } 90.385$$

$$x_1 = 5.151$$

$$x_2 = 1.589$$

$$x_3 = 5.637$$

$$x_4 = 9.288$$

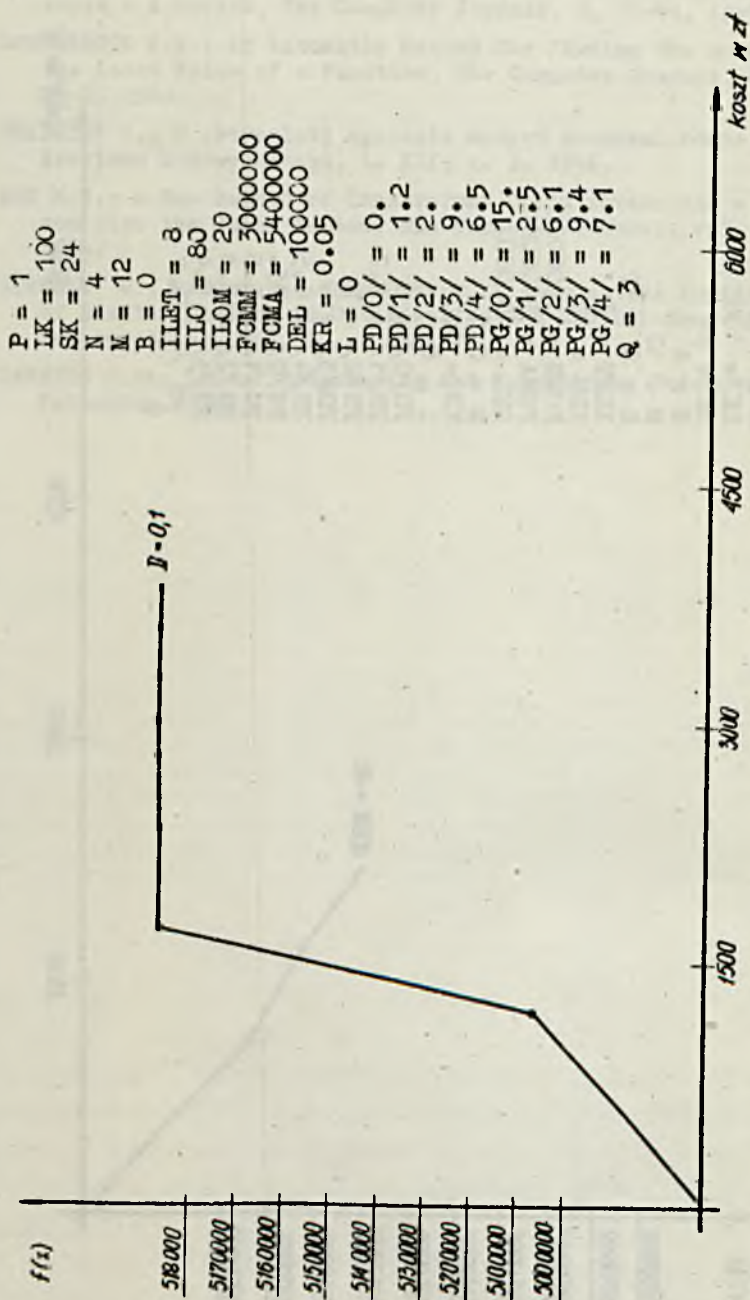
$$x_5 = 6.959$$

W tabeli 5 umieszczono zestawienie rezultatów osiągniętych różnymi metodami i na różnych maszynach. Nie wszystkie algorytmy były jednakowo dokładnie zbadane. Trudno jednoznacznie

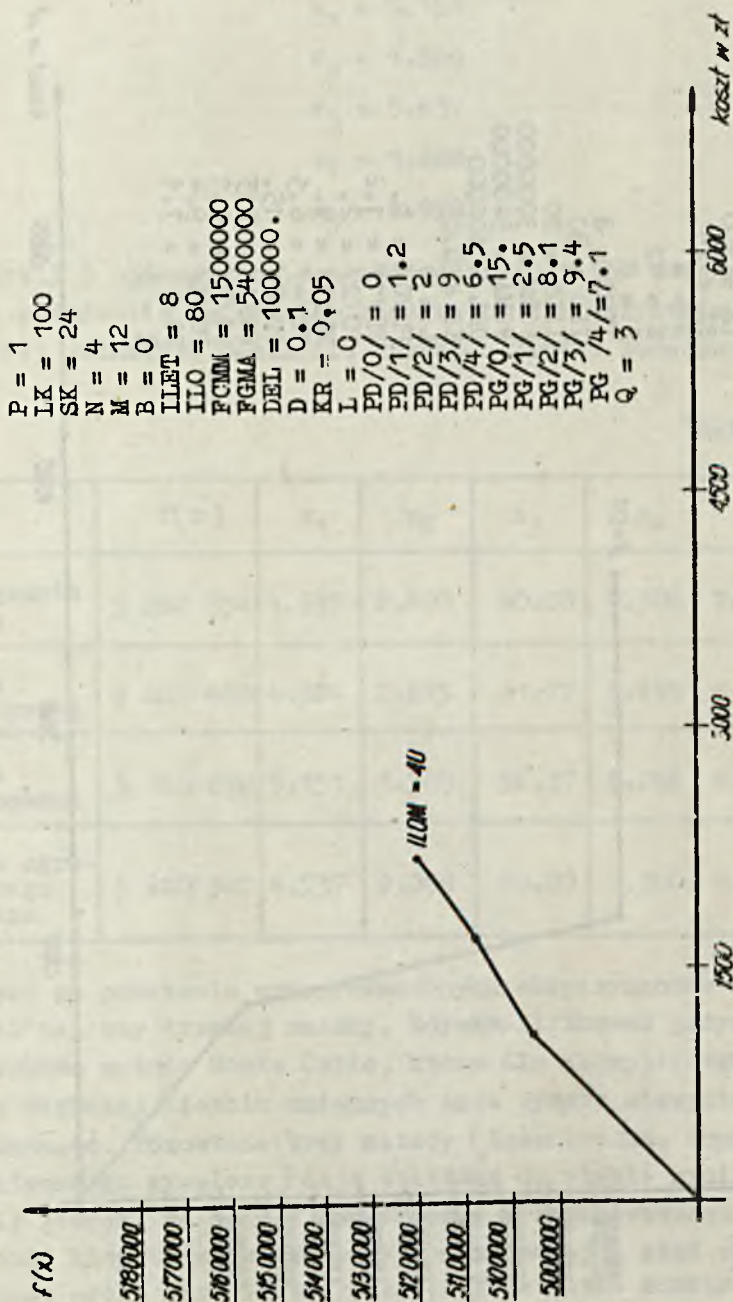
Tabela 5

W	$f(x)$	x_1	x_2	x_3	x_4	x_5
Rozwiązanie ścisłe	5 280 334	4.537	2.400	60.00	9.300	7.000
Metoda Rosenbrocka	5 222 459	4.584	2.215	31.77	9.299	6.999
Metoda kombinowana	5 186 890	5.151	1.589	56.37	9.288	6.959
Metoda ograniczonego sympleksu	5 280 327	4.537	2.399	59.99	9.300	6.999

wyrokować na podstawie przeprowadzonych eksperymentów o przydatności tej czy drugiej metody. Zdyskwalifikować jedynie można surową metodę Monte Carlo, która dla skomplikowanego zadania o większej liczbie zmiennych daje wysoce niewystarczającą dokładność. Pozostałe trzy metody (Rosenbrocka, kombinowana i ograniczonego sympleksu) dają zbliżone do siebie wyniki. Wydaje się jednak, że metody Rosenbrocka i ograniczonego sympleksu są dość kłopotliwe w realizacji maszynowej i stąd czas trwania jednej próby jest dłuższy niż przy metodzie kombinowanej gradientowo-losowej.



Rys. 11. Zależność osiągniętej funkcji celu od kosztu liczenia dla przykładu 5 zmiennych dla parametru δ



Rys. 12. Zależność osiągniętej funkcji celu od kosztu liczenia dla przykładu 5 zmiennych dla zadanego parametru ILOM

Literatura

- [1] FLETCHER R.: Function Minimization without Evaluating Derivatives - A Review, The Computer Journal, 8, 33-41, April 1965.
- [2] ROSENBROCK H.H.: An Automatic Method for Finding the Greatest or the Least Value of a Function, The Computer Journal, vol. 3, No 2, 1960.
- [3] GOLIŃSKI J.: O optymalnej syntezie maszyn metodami Monte Carlo, Archiwum Budowy Maszyn, t. XII, z. 3, 1966.
- [4] BOX M.J.: A New Method of Constrained Optimization and a Comparison with the Other Method, The Computer Journal, vol. 7, No. 3, 1964.
- [5] SPENDLEY W., HEXT G.R., HIMSWORTH F.R.: Sequential Applications of Simplex Designs in Optimization and Evolutionary Operation, Technometrics, vol. 4, p. 441.
- [6] DANTZIG G.B.: Linear Programming and Extensions, Princeton Press, Princeton, 1963.

ПРОСМОТР НЕСКОЛЬКИХ МЕТОДОВ ОСНОВАННЫХ НА КОНЦЕПЦИИ ЭКСПЕРИМЕНТАЛЬНОГО ГРАДИЕНТА И ИХ СРАВНЕНИЕ

Резюме

Представлено несколько реализованных алгоритмов нелинейного программирования для электронных вычислительных машин ZAM 2 и ZAM 41. Их эффективность показана на примерах. Некоторые из этих алгоритмов имеют разную степень свобод, которые позволяют по разному образовать процесс поиска экстремума. Время, а затем и стоимость поиска зависит от соответственного выбора параметров, определяющих процесс. В работе поданы результаты экспериментов, которые должны помочь потребителю сделать соответствующий выбор этой степени свободы.

A REVIEW OF SOME METHODS BASED ON THE CONCEPTION OF AN EXPERIMENTAL GRADIENT AND THEIR EFFECTIVITY COMPARISON

Summary

Presented are some realized algorithms of nonlinear programming for ZAM 2 and ZAM 41 computers. Then effectivity is shown in examples. Some of these algorithms have a range of freedom degrees which permit to form the process of extremum searching in a different way, The time, and thus, the cost of searching depend on a proper choice parameters determining the process. Given the results of experiments which should enable the users to make a proper choice of the above mentioned freedom degrees.

PEWNA METODA NUMERYCZNEGO ROZWIĄZANIA
ZAGADNIENIA BRZEGOWEGO DLA NIELINIO-
WEGO RÓWNANIA RÓŻNICZKOWEGO ZWYCZAJ-
NEGO IV RZĘDU

Piotr KRZYSZEK

Politechnika Warszawska, Instytut Matematyczny
Warszawa, ul. Waszyngtona 41-47

Praca ta napisana została jako praca doktorska pod kierunkiem doc. dr J. Wojtowicza.

W pracy rozwiązano zagadnienia brzegowe dla nieliniowego równania różniczkowego zwyczajnego IV rzędu metodą iteracji prostej. Iterację tę otrzymano poprzez aproksymację z rzędem h^2 pochodnych występujących w problemie różniczkowym ilorazami różnicowymi. Różnicowa analogia funkcji Greena wyznaczona w postaci wzoru (1.13) pozwala uzyskać otwarty schemat iteracyjny. Uzyskano jednoznaczne rozwiązanie postawionego problemu, wykazano zbieżność przyjętej metody iteracyjnej, która badana jest w normie C^2 . Następnie oszacowano błąd dla rozwiązania numerycznego i podano go w postaci wzoru (3.23).

WSTĘP

L. Collatz ([2] str. 48-50 i 197-199) podaje pewną metodę ogólną rozwiązania zagadnienia brzegowego w przypadku ciągłym. Metoda ta wykorzystuje funkcje Greena. Opierając się na odsyłaczu podanym przez Collatza można wnioskować, że zagadnieniem brzegowym II-go rzędu zajmował się F. Lettenmeyer. Praca opublikowana w "Deutsche Mathematik" 7 (1944) jest w Polsce niedostępna i o wynikach w niej zawartych nie powiedzieć nie można.

W niniejszej pracy rozwiązanie zagadnienia brzegowego dla nieliniowego równania różniczkowego oparte jest na metodzie podanej przez J.S. Bierzina i H.P. Ziedkova [8] dla równania różniczkowego II-go rzędu. Wydaje się, że różnicowa realizacja wspomnianej wyżej idei (Collatz-Lettenmeyer) jest znacznie bardziej wygodna w użyciu dla wyznaczenia rozwiązań przybliżonych zagadnienia brzegowego.

I.

Niech dane będzie równanie różniczkowe postaci:

$$y^{(IV)} = f(x, y, y', y'', y''') \quad (1.1)$$

z warunkami brzegowymi na końcach przedziału $\langle 0, 1 \rangle$

$$\left. \begin{aligned} \alpha_0 y_0 - \alpha_1 y'_0 &= a_1 \\ \alpha_2 y'_0 - \alpha_3 y''_0 &= a_2 \\ \beta_0 y_n + \beta_1 y'_n &= b_1 \\ \beta_2 y'_n + \beta_3 y''_n &= b_2 \end{aligned} \right\} \quad (1.2)$$

gdzie: $y_0 = y(0)$; $y'_0 = y'(0)$; $y''_0 = y''(0)$

$y_n = y(1)$; $y'_n = y'(1)$; $y''_n = y''(1)$

Zakładamy, że w obszarze:

$$G \stackrel{\text{def}}{=} [(x, y, z, u, v) : 0 \leq x \leq 1; -\infty < y, z, u, v < +\infty]$$

funkcja $f(x, y, z, u, v)$ spełnia warunki^{*)}:

A. $f \in C^2(G)$, jest ograniczona wraz z pochodnymi do drugiego rzędu włącznie.

B. Liczby α_i, β_i ($i = 0, 1, 2, 3$) są dodatnie.

^{*)} Przyjęto oznaczenia $z = y'$, $u = y''$, $v = y'''$

To, że warunki brzegowe rozpatrujemy na końcach przedziału $\langle 0, 1 \rangle$ nie zmniejsza ogólności, gdyż wprowadzając nową zmienną $t = \frac{x-a}{b-a}$ dla $a < b$ przekształcimy przedział $\langle a, b \rangle$ na przedział $\langle 0, 1 \rangle$.

W przedziale $\langle 0, 1 \rangle$ wprowadzamy siatkę o węzłach:

$$x_i = h i \quad (i = 0, 1, \dots, n), \text{ gdzie: } h = \frac{1}{n}$$

Równanie (1.1) w każdym punkcie $x_k \in \langle 0, 1 \rangle$ aproksymujemy równaniem różnicowym postaci:

$$\begin{aligned} & y_{k+2} - 4 y_{k+1} + 6 y_k - 4 y_{k-1} + y_{k-2} = \\ & = h^4 f \left(x_k, \frac{y_{k+1} - y_{k-1}}{2h}, \frac{y_{k+1} - 2y_k + y_{k-1}}{h^2}, \frac{y_{k+2} - 2y_{k+1} + 2y_{k-1} - y_{k-2}}{2h^3} \right) \end{aligned} \quad (1.3)$$

zaś warunki brzegowe (1.2) zastąpimy związkami:

$$R_{10}(y) = \alpha_0 y_0 - \alpha_1 \frac{-3y_0 + 4y_1 - y_2}{2h} = a_1$$

$$R_{20}(y) = \alpha_2 \frac{-3y_0 + 4y_1 - y_2}{2h} - \alpha_3 \frac{2y_0 - 5y_1 + 4y_2 - y_3}{h^2} = a_2 \quad (1.4)$$

$$R_{1n}(y) = \beta_0 y_n + \beta_1 \frac{3y_n - 4y_{n-1} + y_{n-2}}{2h} = b_1$$

$$R_{2n}(y) = \beta_2 \frac{3y_n - 4y_{n-1} + y_{n-2}}{2h} + \beta_3 \frac{2y_n - 5y_{n-1} + 4y_{n-2} - y_{n-3}}{h^2} = b_2$$

gdzie: $y_i = y(x_i)$ ($i = 0, 1, \dots, n$) oraz $y(x)$ jest rozwiązaniem problemu (1.1) i (1.2).

Układ równań różnicowych (1.3) i (1.4) aproksymuje problem różniczkowy (1.1) i (1.2) z rzędem h^2 . W dalszym ciągu będziemy zajmować się rozwiązaniem problemu (1.3) i (1.4). Związki

te dają nam układ $(n+1)$ równań nieliniowych o $(n+1)$ niewiadomych. Do rozwiązania tego układu równań użyjemy metody iteracji, określonej w następujący sposób:

$$y_{k+2}^{r+1} - 4y_{k+1}^{r+1} + 6y_k^{r+1} - 4y_{k-1}^{r+1} + y_{k-2}^{r+1} = h^4 f_k^r \quad (k=2,3,\dots,n-2)$$

$$R_{\nu 0}(y^{r+1}) = a_\nu \quad (\nu = 1, 2) \quad (1.5)$$

$$R_{\nu n}(y^{r+1}) = b_\nu$$

gdzie:

$$f_k^r = f \left(x_k, y_k^r, \frac{y_{k+1}^r - y_{k-1}^r}{2h}, \frac{y_{k+1}^r - 2y_k^r + y_{k-1}^r}{h^2} \right),$$

$$\frac{y_{k+2}^r - 2y_{k+1}^r + 2y_{k-1}^r - y_{k-1}^r - y_{k-2}^r}{2h^3}$$

Górne indeksy przedstawiają numer kolejnego przybliżenia. Układ równań (1.5) jest przy obranym przybliżeniu zerowym $\langle y_k^0 \rangle \in G$ układem równań liniowych niejednorodnych, który należy rozwiązać przy ustalonym r i h . Do rozwiązania układu równań (1.5) wykorzystujemy układy równań (1.6) i (1.7):

$$\left. \begin{aligned} t_{k+2} - 4t_{k+1} + 6t_k - 4t_{k-1} + t_{k-2} &= 0 \\ R_{\nu 0}(t) &= a_\nu \quad (\nu = 1, 2) \\ R_{\nu n}(t) &= b_\nu \quad (k = 2, 3, \dots, n-2) \end{aligned} \right\} \quad (1.6)$$

$$\left. \begin{aligned} z_{k+2}^r - 4z_{k+1}^r + 6z_k^r - 4z_{k-1}^r + z_k^r &= h^4 f_k^r \\ R_{\nu 0}(z^r) &= 0 \quad (\nu = 1, 2) \\ R_{\nu n}(z^r) &= 0 \quad (k = 2, 3, \dots, n-2) \end{aligned} \right\} \quad (1.7)$$

Twierdzenie 1.1. Jeżeli przy ustalonym h i r układy równań 1.6 i (1.7) posiadają jednoznaczne rozwiązanie, to układ równań (1.5) posiada jednoznaczne rozwiązanie, przy czym rozwiązanie to jest postaci:

$$y_k^{r+1} = t_k + z_k^r \quad (1.8)$$

Dowód twierdzenia jest oczywisty i natychmiastowy.

Pokażemy teraz, że zachodzi:

Twierdzenie 2.1. Przy ustalonym h i r układ równań (1.6) posiada jednoznaczne rozwiązanie postaci:

$$t_k = \frac{1}{\Delta} \left(\Delta_0 + \Delta_1 k + \Delta_2 k^2 + \Delta_3 k^3 \right) \quad (1.9)$$

gdzie: $\Delta, \Delta_0, \Delta_1, \Delta_2, \Delta_3$ są to wyznaczniki stopnia 4-go zależne od α_1, β_1 ($i = 0, 1, 2, 3$) oraz od h ,

Dowód. Pierwszy związek w (1.6) wskazuje nam, że czwarta różnica skończona dla funkcji $t_k = t(k)$ jest równa zero. Zatem t_k jest postaci:

$$t_k = C_0 + C_1 k + C_2 k^2 + C_3 k^3. \quad (1.10)$$

gdzie C_0, C_1, C_2, C_3 są stałe, które należy wyznaczyć tak, aby były spełnione pozostałe związki w (1.6).

Po przeliczeniu otrzymamy następujący układ równań ze względu na C_i ($i = 0, 1, 2, 3$):

$$\alpha_0 h C_0 - \alpha_1 C_1 + 2\alpha_1 C_3 = a_1 h$$

$$\alpha_2 h C_1 - 2\alpha_3 C_2 - 2\alpha_2 h C_3 = a_2 h^2$$

$$\beta_0 h C_0 + (\beta_0 + \beta_1) C_1 + (\beta_0 + 2\beta_1) h C_2 + (\beta_0 h^2 + 3\beta_1 h^2 - 2\beta_1) C_3 = b_1 h$$

$$\beta_2 h C_1 + 2(\beta_2 + \beta_3) C_2 + (3\beta_2 h + 6\beta_3 h - 2\beta_2 h) C_3 = b_2 h^2$$

Wyznacznik główny tego układu równań wynosi:

$$\begin{aligned} \Delta = & \alpha_0 \alpha_2 (\beta_0 \beta_2 + 4\beta_0 \beta_3 + 6\beta_1 \beta_3 - 4h^2 \beta_0 \beta_2 - 4h^2 \beta_0 \beta_3) + \\ & + 4\alpha_0 \alpha_3 (\beta_0 \beta_2 + 3\beta_0 \beta_3 + 3\beta_1 \beta_3 - h^2 \beta_0 \beta_2) + \\ & + 6\alpha_1 \alpha_3 \beta_0 (\beta_2 + \beta_3) \geq \alpha_0 \alpha_2 \left(\frac{3}{4} \beta_0 \beta_2 + \frac{3}{4} \beta_0 \beta_3 + 6\beta_1 \beta_3 \right) + \\ & + 4\alpha_0 \alpha_3 \left(\frac{3}{4} \beta_0 \beta_2 + 3\beta_0 \beta_3 + 3\beta_1 \beta_3 \right) + 6\alpha_1 \alpha_3 \beta_0 (\beta_2 + \beta_3) = \delta \end{aligned} \quad (1.10^*)$$

Nierówność

$$\Delta > \delta > 0 \quad (1.11)$$

wynika z założenia (B) oraz z tego, że $n \geq 4$, a zatem i $h^2 \leq \frac{1}{16}$.

Z (1.11) wynika, że stałe C_i ($i = 0, 1, 2, 3$) można wyznaczyć jednoznacznie. Następnie wyznaczamy wyznaczniki

$\Delta_i = (i = 0, 1, 2, 3)$, które mają postać następującą:

$$\begin{aligned} \Delta_0 = & -2(\beta_0 + 3\beta_1 + 2\beta_0 h^2) \left((\beta_2 + \beta_3) (a_1 \alpha_2 + a_2 \alpha_1) + \alpha_3 (b_2 \alpha_1 + a_1 \alpha_2) \right) + \\ & + 3(\beta_2 + 2\beta_3) \left((\beta_0 + 2\beta_1) (a_1 \alpha_2 + a_2 \alpha_1) + 2\alpha_3 (\beta_1 \alpha_1 + a_1 \beta_0 + a_1 \beta_1) \right) \\ \Delta_1 = & \beta_0 h \left(4(\beta_2 + \beta_3) (a_1 \alpha_2 + a_2 \alpha_1) h^2 + 4b_2 \alpha_1 \alpha_3 h^2 - 2a_1 \alpha_3 (3\beta_2 + 6\beta_3 - 2\beta_2 h)^2 \right) \\ & + \alpha_0 h \left(-2(\beta_0 + 3\beta_1 - 2\beta_1 h^2) (b_2 \alpha_3 + a_1 (\beta_2 + \beta_3)) + \right. \\ & + \left. (3\beta_2 + 6\beta_3 - 2\beta_2 h^2) (a_2 (\beta_0 + 2\beta_1) + 2b_1 \alpha_3) \right) + \\ & + 2\alpha_2 h^2 (b_2 (\beta_0 + 2\beta_1) - 2b_1 (\beta_2 + \beta_3)) \end{aligned}$$

$$\Delta_2 = \alpha_0 h^2 (3(\beta_2 + 2\beta_3)(b_1 \alpha_2 - a_2(\beta_0 + \beta_1)) +$$

$$+ (\beta_0 + 3\beta_1 + 2\beta_0 h^2)(a_2 \beta_2 - b_2 \alpha_2)) - 3\beta_0 (\beta_2 + 2\beta_3)(a_2 \alpha_1 + a_1 \alpha_2) h^2$$

$$\Delta_3 = \alpha_0 h^3 ((\beta_0 + 2\beta_1)(b_2 \alpha_2 - a_2 \beta_2) + 2(\beta_0 + \beta_1)(a_2 \beta_2 + a_2 \beta_3 + b_2 \alpha_3) +$$

$$- b_1(\alpha_3 \beta_2 + \alpha_2 \beta_2 + \alpha_2 \beta_3)) + 2\beta_0 h^3 (\alpha_3 (b_2 \alpha_1 + a_1 \beta_2) + (\beta_2 + \beta_3)(a_1 \alpha_2 + a_2 \alpha_1))$$

Funkcja (1.10) przyjmie zatem postać (1.9), która jest wyznaczona w sposób jednoznaczny i która spełnia pierwsze z równań układu (1.6) dla $k = 2, 3, \dots, n-2$ oraz pozostałe cztery równania, czyli układ $(n+1)$ równań o $(n+1)$ niewiadomych.

Ponieważ układ ten jest układem niejednorodnym, a funkcja (1.9) jest jednoznaczny rozwiązaniem tego układu równań, zatem macierz główna tego układu (identyczna z macierzą główną układu (1.7)) jest nieosobliwa. Stąd układ równań (1.7) ma przy ustalonym h i r rozwiązanie jednoznaczne. Na mocy twierdzenia 2.1 mamy:

Wniosek 1. Układy równań (1.7) i (1.6) przy ustalonym h i r posiadają jednoznaczne rozwiązanie.

Twierdzenie 3.1. Przy ustalonym h następujący układ równań:

$$\varepsilon_{1k+2} - \varepsilon_{1k+1} + 6\varepsilon_{1k} - 4\varepsilon_{1k-1} + \varepsilon_{1k-2} = \delta_{1k}$$

$$\alpha_0 \varepsilon_{10} - \alpha_1 \frac{-3\varepsilon_{10} + 4\varepsilon_{11} - \varepsilon_{12}}{2h} = 0$$

$$\alpha_2 \frac{-3\varepsilon_{10} + 4\varepsilon_{11} - \varepsilon_{12}}{2h} - \alpha_3 \frac{2\varepsilon_{10} - 5\varepsilon_{11} + 4\varepsilon_{12} - \varepsilon_{13}}{h^2} = 0 \quad (1.12)$$

$$\beta_0 (\varepsilon_{1n} + \beta_1 \frac{3\varepsilon_{1n} - 4\varepsilon_{1n-1} + \varepsilon_{1n-2}}{2h}) = 0$$

$$\beta_2 \frac{\varepsilon_{1n} - 4\varepsilon_{1n-1} + \varepsilon_{1n-2}}{2h} + \beta_3 \frac{2\varepsilon_{1n} - 5\varepsilon_{1n-1} + 4\varepsilon_{1n-2} - \varepsilon_{1n-3}}{h^2} = 0$$

($i = 0, 1, 2, \dots, n$)
 ($k = 2, 3, 4, \dots, n-2$) δ_{ik} - symbol Kroneckera
 posiada jednoznaczne rozwiązanie postaci:

$$\begin{aligned} \xi_{ik} = & \frac{1}{\Delta} (w_0 + w_1 k + w_2 k^2 + w_3 k^3) + \\ & + \frac{1}{6} \delta(i, k+1) (k^3 - 3ik^2 + (3i^2 - 1)k - i^3 + i) \end{aligned} \quad (1.13)$$

gdzie:

$$\delta(i, k+1) \stackrel{\text{def}}{=} \sum_{q=2}^{k+1} \delta_{iq}, \quad \text{tzn. } \delta(i, k+1) = \begin{cases} 1 & 2 \leq i \leq k+1 \\ 0 & i > k+1 \end{cases} \\ (k = 1, 2, \dots, n)$$

$$\text{oraz } \delta(i, 1) = 0 \quad (i = 0, 1, \dots, n)$$

W_1 , są wyznacznikami stopnia czwartego zależne od α_1 , β_1 ($i = 0, 1, 2, 3$), n , i oraz h .

Dowód. Weźmy pod uwagę równanie postaci:

$$\xi_{ik+4} - 4\xi_{ik+3} + 6\xi_{ik+2} - 4\xi_{ik+1} + \xi_{ik} = \delta_{ik+2} \quad \left(\begin{matrix} i=0, 1, \dots, n \\ k=0, 1, \dots, n-4 \end{matrix} \right) \quad (1.14)$$

Równanie (1.14) jest równoważne pierwszemu z równań (1.12).

Układ funkcji: $1, k, k^2, k^3$ tworzy fundamentalny układ rozwiązań równania jednorodnego odpowiadającego równaniu (1.14). Zatem rozwiązanie ogólne tego równania jednorodnego będzie postaci:

$$\xi_{ik} = a + b k + c k^2 + d k^3 \quad (1.15)$$

Rozwiązanie szczególne równania (1.14) ([11] str. 390-391), jest postaci:

$$\gamma_{ik} = \sum_{q=0}^{k-1} \frac{1}{1.2} \begin{vmatrix} 1 & q+1 & (q+1)^2 & (q+1)^3 \\ 1 & q+2 & (q+2)^2 & (q+2)^3 \\ 1 & q+3 & (q+3)^2 & (q+3)^3 \\ 1 & k & k^2 & k^3 \end{vmatrix} \delta_{ik+2} = \frac{1}{6} \sigma(i, k+1) (k^3 - 3ik^2 + (3i^2 - 1)k - i^3 + i) \quad (1.16)$$

gdzie:

$$\left. \begin{aligned} \sigma(i, k+1) &= \sum_{q=2}^{k+1} \delta_{iq} \text{ oraz } \sigma(i, 1) = 0 \text{ dla } i = 0, 1, \dots, n \\ \sigma(i, k+1) &= \begin{cases} 1 & \text{dla } 2 \leq i \leq k+1 \\ 0 & \text{dla } i > k+1 \end{cases} \quad (k = 1, 2, \dots, n) \end{aligned} \right\} \quad (1.17)$$

Z (1.15) i (1.16) wynika, że rozwiązanie ogólne równania (1.14), a tym samym i pierwszego z równań (1.12) jest postaci:

$$\xi_{ik} = a + bk + ck^2 + dk^3 + \frac{1}{6} \sigma(i, k+1) (k^3 - 3ik^2 + (3i^2 - 1)k - i^3 + i) \quad (1.18)$$

Łatwo sprawdzić, że funkcja (1.18) jest rozwiązaniem ogólnym pierwszego z równań (1.12) dla $i = k$ ($k = 2, 3, \dots, n-2$), zaś dla pozostałych "i" jest rozwiązaniem równania jednorodnego odpowiadającego wyżej wspomnianemu równaniu. Wspecjalizujemy teraz bliżej funkcję ξ_{ik} wyznaczając stałe a, b, c, d. Stałe te wyznaczymy korzystając z faktu, że funkcja ξ_{ik} ma spełniać pozostałe równania (1.12). Szczegółowe przeliczenia podamy tylko dla drugiego z równań w (1.12), zaś dla pozostałych przeliczenia będą przebiegać identycznie.

$$\begin{aligned} & (2\alpha_0 h + 3\alpha_1) \left(a + \frac{\sigma(1,1)}{6} (-i^3 + i) \right) + \\ & - 4\alpha_1 \left(a + b + c + d + \frac{1}{6} \sigma(1,2) (1 - 3i + 3i^2 - 1 - i^3 + i) \right) + \\ & + \alpha_1 \left(a + 2b + 4c + 8d + \frac{1}{6} \sigma(1,3) (8 - 12i + 6i^2 - 2 - i^3 + i) \right) = 0 \end{aligned}$$

Zatem pierwsze równanie jest postaci:

$$2\alpha_0 h a - 2\alpha_1 b + 4\alpha_1 d = \frac{2\alpha_0 h + 3\alpha_1}{6} \sigma(1,1)1(1-1)(1+1) + \\ - \frac{4\alpha_1}{6} \sigma(1,2) 1(1-1)(1-2) + \frac{\alpha_1}{6} \sigma(1,3)(1-1)(1-2)(1-3)$$

Z (1.17) wynika, że prawa strona tego równania jest równa zero. Zatem:

$$\alpha_0 h a - \alpha_1 b + 2\alpha_1 d = 0$$

Dokonując analogicznych przeliczeń dla pozostałych równań otrzymamy układ równań:

$$\left. \begin{aligned} \alpha_0 h a - \alpha_1 b + 2\alpha_1 d &= 0 \\ \alpha_2 h b - 2\alpha_3 o - 2\alpha_2 h d &= 0 \\ \beta_0 h a + (\beta_0 + \beta_1)b + (\beta_0 + 2\beta_1)n o + \\ &+ (\beta_0 n^2 + 3\beta_1 n^2 - 2\beta_1) d = \frac{1}{2} A_3 \\ \beta_2 h b + 2(\beta_2 + \beta_3)o + (3\beta_2 n + 6\beta_3 n - 2\beta_2 h) d &= \frac{1}{2} A_4 \end{aligned} \right\} (1.19)$$

gdzie:

$$A_3 = -\frac{1}{3} \sigma(1, n-2) (\beta_0 h (n^3 - 3in^2 + (3i^2 - 1)n - i^3 + 1) + \\ + 3\beta_1 (n^2 + i^2 - 2ni - 1)) \quad (1.20)$$

$$A_4 = -6(1, n-2) (\beta_2 h (n^2 + i^2 - 2ni - 1) + 2\beta_3 (n - 1))$$

Wyznacznik główny układu równań (1.19) był już policzony dla funkcji t_k i oznaczyliśmy go przez Δ .

Następne wyznaczniki są postaci:

$$\begin{aligned}
 W_0 &= \alpha_1 \alpha_3 \left(3(\beta_2 + 2\beta_3)n A_3 - (\beta_0 n^2 + 3\beta_1 n^2 + 2\beta_0) A_4 \right) \\
 W_1 &= \alpha_0 h \left(\alpha_2 \left((\beta_0 + 2\beta_1) A_4 - 2h(\beta_2 + \beta_3) A_3 \right) - \alpha_3 \left((\beta_0 n^2 + 3\beta_1 n^2 - 2\beta_1) A_4 + \right. \right. \\
 &\quad \left. \left. - (3\beta_2 n + 6\beta_3 n - 2\beta_2 h) A_3 \right) \right) + 2\alpha_1 \alpha_3 \beta_0 h A_4 \quad (1.21) \\
 W_2 &= \frac{\alpha_0 \alpha_2 h}{2} \left(3(\beta_2 + 2\beta_3) A_3 - (\beta_0 n + 3\beta_1 n + 2\beta_0 h) A_4 \right) \\
 W_3 &= \alpha_0 h \left(\frac{\alpha_2}{2} \left((\beta_0 + 2\beta_1) A_4 - 2(\beta_2 + \beta_3) A_3 \right) + \right. \\
 &\quad \left. + \alpha_3 \left((\beta_0 + \beta_1) A_4 - \beta_2 h A_3 \right) \right) + \alpha_1 \alpha_3 \beta_0 h A_4
 \end{aligned}$$

Z (1.19), (1.21) oraz z (1.18) wynika, że ξ_{1k} jest postaci danej w (1.13).

Twierdzenie 4.1. Funkcja $z_k^r = h^4 \sum_{i=2}^{n-2} \xi_{1k} f_1^r$ ($k = 0, 1, \dots, n$) przy ustalonym h i r jest jednoznaczny rozwiązaniem układu równań (1.7).

Dowód. Niech h i r będzie dowolne, ale ustalone.

Weźmy wyrażenie:

$$\begin{aligned}
 z_{k+2}^r - 4z_{k+1}^r + 6z_k^r - 4z_{k-1}^r + z_{k-2}^r &= h^4 \sum_{i=2}^{n-2} \xi_{1k+2} f_1^r - \\
 - 4h^4 \sum_{i=2}^{n-2} \xi_{1k+1} f_1^r + 6h^4 \sum_{i=2}^{n-2} \xi_{1k} f_1^r - 4h^4 \sum_{i=2}^{n-2} \xi_{1k-1} f_1^r + \\
 + h^4 \sum_{i=2}^{n-2} \xi_{1k+2} f_1^r &= h^4 \sum_{i=2}^{n-2} \left(\xi_{1k+2} - 4\xi_{1k+1} + 6\xi_{1k} - 4\xi_{1k-1} + \xi_{1k-2} \right) f_1^r
 \end{aligned}$$

Ponieważ k zmienia się od 2 do $n-2$, więc na pewno istnieje jedno takie " i " równe " k ". Stąd na mocy (1.13) i pierwszego równania w (1.12) wynika, że wyrażenie w nawiasie pod znakiem sumy równa się jedności dla $i = k$, zaś zero dla $i \neq k$.

W takim razie mamy:

$$Z_{k+2}^r - 4Z_{k+1}^r + 6Z_k^r - 4Z_{k-1}^r + Z_{k-2}^r = h^4 f_k^r$$

To oznacza, że funkcja Z_k^r zdefiniowana wyżej spełnia pierwsze z równań (1.12). Podobnie możemy pokazać, że ta funkcja spełnia pozostałe równania (1.12), czyli warunki brzegowe. A więc tak zdefiniowana funkcja Z_k^r jest jednoznacznym rozwiązaniem układu równań (1.7).

Wniosek 2. Na mocy wniosku 1 oraz twierdzenia 1.1 jednoznaczne rozwiązanie układu równań (1.5) jest postaci:

$$y_k^{r+1} = \frac{1}{\Delta} \left(\Delta_0 + \Delta_1 k + \Delta_2 k^2 + \Delta_3 k^3 \right) + h^4 \sum_{i=2}^{n-2} \varepsilon_{1k} f_1^r \quad (k = 0, 1, \dots, n) \quad (1.22)$$

Wzór iteracyjny (1.22) pozwalający obliczać przybliżone wartości funkcji $y = y(x)$ w dowolnym punkcie $x_k \in \langle 0, 1 \rangle$ jest wygodny w użyciu dlatego, że przy obliczaniu wartości $y_k = y(x_k)$ wystarczy dla obranego kroku h policzyć tylko raz funkcje ε_{1k} , a następnie możemy znajdować kolejne przybliżenia rozwiązania równania różniczkowego. Zmieniając następnie krok h możemy znów ten sam proces powtarzać.

II.

Wykażemy teraz zbieżność procesu iteracyjnego. W tym celu dokonamy najpierw oszacowań pewnych wielkości. Z założenia (A) i (B) oraz z (1.13) wynikają nierówności:

$$\left. \begin{aligned} \left| h^4 \sum_{i=2}^{n-2} \varepsilon_{1k} f_1^r \right| &< h^4 \sum_{i=2}^{n-2} \left| \varepsilon_{1k} \right| \cdot \left| f_1^r \right| < \\ \max_1 \left| f_1^r \right| \cdot h^4 \sum_{i=2}^{n-2} \left| \varepsilon_{1k} \right| &< \max_1 \left| f_1^r \right| \cdot B_0 \end{aligned} \right\} \quad (2.1)$$

gdzie:

$$B_0 = \frac{1}{8} \left(\alpha_0 \alpha_2 \left(\frac{\beta_0 \beta_2}{8} + \frac{7}{12} \beta_0 \beta_3 + \frac{7}{6} \beta_1 \beta_3 \right) + \right. \\ \left. + \alpha_0 \alpha_3 \left(\frac{\beta_0 \beta_2}{3} + \frac{3}{2} \beta_0 \beta_3 + 2 \beta_1 \beta_3 \right) + \right. \\ \left. + \alpha_1 \alpha_3 \left(\frac{5}{12} \beta_0 \beta_2 + \frac{3}{2} \beta_0 \beta_3 + \beta_1 \beta_3 \right) \right) + \frac{1}{24}$$

$$\left| h^4 \sum_{i=2}^{n-2} \frac{\varepsilon_{ik+1} - \varepsilon_{ik-1}}{2h} f_1^r \right| \leq \frac{h^3}{2} \sum_{i=2}^{n-2} |\varepsilon_{ik+1} - \varepsilon_{ik-1}| \cdot |f_1^r| \leq$$

$$\max_1 |f_1^r| \left| \frac{h^3}{2} \sum_{i=2}^{n-2} |\varepsilon_{ik+1} - \varepsilon_{ik-1}| \right| < \max_1 |f_1^r| \cdot B_1$$

gdzie:

$$B_1 = \frac{1}{8} \left(\alpha_0 \alpha_2 \left(\frac{\beta_0 \beta_2}{3} + \frac{3}{2} \beta_0 \beta_3 + 3 \beta_1 \beta_3 \right) + \right. \\ \left. + \alpha_0 \alpha_3 \left(\frac{5 \beta_0 \beta_2}{6} + \frac{7}{2} \beta_0 \beta_3 + 4 \beta_1 \beta_3 \right) + \right. \\ \left. + \alpha_1 \alpha_3 \left(\beta_0 \beta_2 + 3 \beta_0 \beta_3 + 3 \beta_1 \beta_3 \right) \right) + \frac{1}{6}$$

$$\left| h^4 \sum_{i=2}^{n-2} \frac{\varepsilon_{ik+1} - 2\varepsilon_{ik} + \varepsilon_{ik-1}}{h^2} f_1^r \right| < \max_1 |f_1^r| \cdot B_2$$

gdzie:

$$B_2 = \frac{1}{8} \left(\alpha_0 \alpha_2 \left(\frac{7}{12} \beta_0 \beta_2 + \frac{3}{2} \beta_0 \beta_3 + 5 \beta_1 \beta_3 \right) + \right. \\ \left. + \alpha_0 \alpha_3 \left(\frac{3}{2} \beta_0 \beta_2 + 6 \beta_0 \beta_3 + 6 \beta_1 \beta_3 \right) + \right. \\ \left. + \alpha_1 \alpha_3 \left(2 \beta_0 \beta_2 + 6 \beta_0 \beta_3 \right) \right) + \frac{1}{2}$$

$$\left| h^4 \sum_{i=2}^{n-2} \frac{\varepsilon_{1k+2} - 2\varepsilon_{1k+1} + 2\varepsilon_{1k-1} - \varepsilon_{1k-2}}{2h^3} f_1^r \right| < \max_1 |f_1^r| \cdot B_3$$

gdzie:

$$B_3 = \frac{1}{\delta} \left(\alpha_0 \alpha_2 \left(\beta_0 \beta_2 + 8\beta_0 \beta_3 + 16\beta_1 \beta_3 \right) + \right. \\ \left. + \alpha_0 \alpha_3 \left(3\beta_0 \beta_2 + 12\beta_0 \beta_3 + 12\beta_1 \beta_3 \right) + \right. \\ \left. + \alpha_1 \alpha_3 \left(4\beta_0 \beta_2 + 12\beta_0 \beta_3 \right) \right) + 1$$

Z (1.10') wynika, że δ nie zależy od h , a zatem B_0, B_1, B_2, B_3 również nie zależą od h . Rozpatrzmy przestrzeń R^{n+1} , której elementami są uporządkowane ciągi liczb rzeczywistych $\langle y_k \rangle = (y_0, y_1, \dots, y_n)$ z metryką określoną w następujący sposób:

$$\rho(\langle y_k \rangle, \langle \bar{y}_k \rangle) = L_0 \max_{0 \leq k \leq n} |y_k - \bar{y}_k| + L_1 \max_{1 \leq k \leq n-1} |z_k - \bar{z}_k| + \\ + L_2 \max_{1 \leq k \leq n-1} |u_k - \bar{u}_k| + L_3 \max_{2 \leq k \leq n-2} |v_k - \bar{v}_k| \quad (2.2)$$

gdzie: L_0, L_1, L_2, L_3 oznaczają nam stałe Lipschitza dla funkcji $f(x, y, z, u, v)$ odpowiadające odpowiednio zmiennym y, z, u, v , zaś

$$z_k = \frac{y_{k+1} - y_{k-1}}{2h}, \quad u_k = \frac{y_{k+1} - 2y_k + y_{k-1}}{h^2},$$

$$v_k = \frac{y_{k+2} - 2y_{k+1} + 2y_{k-1} - y_{k-2}}{2h^3}$$

Przy tak określonej odległości R^{n+1} jest przestrzenią metryczną, gdyż (2.2) spełnia aksjomaty metryki oraz jest przestrzenią zupełną, jako przestrzeń skończenie wymiarowa.

W dalszym ciągu będziemy rozważać tylko te elementy $\langle y_k \rangle \in R^{n+1}$, dla których punkty:

$$\left(x_k, y_k, z_k, u_k, v_k \right) \in G \quad (2.3)$$

Dla każdego takiego elementu $\langle y_k \rangle \in R^{n+1}$ wzór (1.22) definiuje pewne odwzorowanie nieliniowe postaci:

$$\langle y_k^{r+1} \rangle = A \left(\langle y_k^r \rangle \right)$$

elementu $\langle y_k^r \rangle \in R^{n+1}$ w element $\langle y_k^{r+1} \rangle \in R^{n+1}$, gdzie przez A rozumiemy operacje:

$$A \stackrel{\text{def}}{=} \left(A_0, A_1, \dots, A_n \right)$$

zaś

$$A_p \left(\langle y_k \rangle \right) = S_p + h^4 \sum_{i=2}^{n-2} \varepsilon_{ip} f \left(x_i, y_i, z_i, u_i, v_i \right)$$

oraz

$$S_p = \frac{1}{\Delta} \left(\Delta_0 + \Delta_1 p + \Delta_2 p^2 + \Delta_3 p^3 \right)$$

Aby uniknąć posługiwania się indeksami górnymi definiujemy element $\langle \eta_k \rangle$ jako obraz elementu $\langle y_k \rangle$:

$$\langle \eta_k \rangle = A \left(\langle y_k \rangle \right) \quad (2.4)$$

Twierdzenie 5.1. Jeżeli $\langle y_k \rangle, \langle \bar{y}_k \rangle \in R^{n+1}$, to spełnione są następujące związki:

$$\begin{aligned}
 & \left| \eta_k - \bar{\eta}_k \right| < B_0 \varphi \left(\langle y_k \rangle, \langle \bar{y}_k \rangle \right) \\
 & \left| \frac{\eta_{k+1} - \eta_{k-1}}{2h} - \frac{\bar{\eta}_{k+1} - \bar{\eta}_{k-1}}{2h} \right| < B_1 \varphi \left(\langle y_k \rangle, \langle \bar{y}_k \rangle \right) \\
 & \left| \frac{\eta_{k+1} - 2\eta_k + \eta_{k-1}}{h^2} - \frac{\bar{\eta}_{k+1} - 2\bar{\eta}_k + \bar{\eta}_{k-1}}{h^2} \right| < \\
 & \quad < B_2 \varphi \left(\langle y_k \rangle, \langle \bar{y}_k \rangle \right) \tag{2.5} \\
 & \left| \frac{\eta_{k+2} - 2\eta_{k+1} + 2\eta_{k-2} - \eta_{k-2}}{2h^3} - \frac{\bar{\eta}_{k+2} - 2\bar{\eta}_{k+1} - 2\bar{\eta}_{k-1} - \bar{\eta}_{k-2}}{2h^3} \right| < \\
 & \quad < B_3 \varphi \left(\langle y_k \rangle, \langle \bar{y}_k \rangle \right)
 \end{aligned}$$

oraz

$$\varphi \left(\langle \eta_k \rangle, \langle \bar{\eta}_k \rangle \right) < \gamma \varphi \left(\langle y_k \rangle, \langle \bar{y}_k \rangle \right) \tag{2.6}$$

gdzie:

$$\gamma = L_0 B_0 + L_1 B_1 + L_2 B_2 + L_3 B_3 \quad \text{i nie zależy od } h. \tag{2.6}$$

- L_1 - są stałymi Lipschitza względem zmiennych y, z, u, v ;
 B_i - wielkości zależne od α_i, β_i ($i = 0, 1, 2, 3$) określone uprzednio.

Oszacujemy tylko pierwszą z nierówności (2.5), gdyż pozostałe otrzymane w podobny sposób.

Dowód. Na mocy założenia (A) funkcja $f(x, y, z, u, v)$ spełnia w obszarze G warunek Lipschitza względem zmiennych y, z, u, v , tzn. istnieją takie stałe liczby dodatnie L_0, L_1, L_2, L_3 , że dla każdej pary punktów (x, y, z, u, v) i $(\bar{x}, \bar{y}, \bar{z}, \bar{u}, \bar{v})$ należących do G zachodzi związek:

$$\begin{aligned}
 \left| f(x, y, z, u, v) - f(\bar{x}, \bar{y}, \bar{z}, \bar{u}, \bar{v}) \right| < L_0 |y - \bar{y}| + \\
 + L_1 |z - \bar{z}| + L_2 |u - \bar{u}| + L_3 |v - \bar{v}| \tag{2.7}
 \end{aligned}$$

Jeżeli $\langle y_k \rangle$ i $\langle \bar{y}_k \rangle \in R^{n+1}$, to na mocy (2.2), (2.4) i (2.7) otrzymamy:

$$\left| \eta_k - \bar{\eta}_k \right| = h^4 \left| \sum_{i=2}^{n-2} \varepsilon_{ik} \left(f \left(x_1, y_1, \frac{y_{i+1} - y_{i-1}}{2h}, \frac{y_{i+1} - 2y_i + y_{i-1}}{h^2}, \frac{y_{i+2} - 2y_{i+1} + 2y_{i-1} - y_{i-2}}{2h^3} \right) - f \left(x_1, \bar{y}_1, \frac{\bar{y}_{i+1} - \bar{y}_{i-1}}{2h}, \frac{\bar{y}_{i+1} - 2\bar{y}_i + \bar{y}_{i-1}}{h^2}, \frac{\bar{y}_{i+2} - 2\bar{y}_{i+1} + 2\bar{y}_{i-1} - \bar{y}_{i-2}}{2h^3} \right) \right) \right| <$$

$$< h^4 \sum_{i=2}^{n-2} \left| \varepsilon_{ik} \left(L_0 |y_1 - \bar{y}_1| + L_1 \left| \frac{y_{i+1} - y_{i-1}}{2h} - \frac{\bar{y}_{i+1} - \bar{y}_{i-1}}{2h} \right| + L_2 \left| \frac{y_{i+1} - 2y_i + y_{i-1}}{h^2} - \frac{\bar{y}_{i+1} - 2\bar{y}_i + \bar{y}_{i-1}}{h^2} \right| + L_3 \left| \frac{y_{i+2} - 2y_{i+1} + 2y_{i-1} - y_{i-2}}{2h^3} - \frac{\bar{y}_{i+2} - 2\bar{y}_{i+1} + 2\bar{y}_{i-1} - \bar{y}_{i-2}}{2h^3} \right| \right) \right| <$$

$$< B_0 \left(L_0 \max_{0 \leq i \leq n} |y_1 - \bar{y}_1| + L_1 \max_{1 \leq i \leq n-1} \left| \frac{y_{i+1} - y_{i-1}}{2h} - \frac{\bar{y}_{i+1} - \bar{y}_{i-1}}{2h} \right| + \right.$$

$$+ L_2 \max_{1 \leq i \leq n-1} \left| \frac{y_{i+1} - 2y_i + y_{i-1}}{h^2} - \frac{\bar{y}_{i+1} - 2\bar{y}_i + \bar{y}_{i-1}}{h^2} \right| +$$

$$\left. L_3 \max_{2 \leq i \leq n-2} \left| \frac{y_{i+2} - 2y_{i+1} + 2y_{i-1} - y_{i-2}}{2h^3} - \frac{\bar{y}_{i+2} - 2\bar{y}_{i+1} + 2\bar{y}_{i-1} - \bar{y}_{i-2}}{2h^3} \right| \right)$$

czyli:

$$\left| \eta_k - \bar{\eta}_k \right| < B_0 \varphi(\langle y_k \rangle, \langle \bar{y}_k \rangle) \quad (2.8)$$

W podobny sposób otrzymamy pozostałe nierówności w (2.5). Nierówności (2.5) są prawdziwe dla każdej wartości k , więc będą również prawdziwe dla maksimum na \mathbb{K} . Mnożąc zatem te nierówności odpowiednio przez stałe Lipschitza $L_i > 0$ ($i = 0, 1, 2, 3$) i dodając je stronami otrzymamy nierówność (2.6).

Z warunku (2.6) wynika, że jeżeli γ spełnia warunek:

$$0 < \gamma < 1 \quad (2.9)$$

to operacja A jest zblizająca i prawdziwe są następujące twierdzenia (6.1) i (7.1).

Twierdzenie 6.1. Jeżeli przybliżenie zerowe $\langle y_k \rangle \in G$, funkcja f spełnia warunek Lipschitza oraz A jest operacją zblizającą, to dla każdego punktu $(x_k, y_k, z_k; u_k, v_k)$ prawdziwe są związki:

$$\left. \begin{aligned} |y_k - y_k^1| &\leq \frac{B_0}{1-\gamma} \varphi \left(\langle y_k^0 \rangle, \langle y_k^1 \rangle \right) \\ |z_k - z_k^1| &\leq \frac{B_1}{1-\gamma} \varphi \left(\langle y_k^0 \rangle, \langle y_k^1 \rangle \right) \\ |u_k - u_k^1| &\leq \frac{B_2}{1-\gamma} \varphi \left(\langle y_k^0 \rangle, \langle y_k^1 \rangle \right) \\ |v_k - v_k^1| &\leq \frac{B_3}{1-\gamma} \varphi \left(\langle y_k^0 \rangle, \langle y_k^1 \rangle \right) \end{aligned} \right\} \quad (2.10)$$

Dowód: Nierówności (2.10) wynikają bezpośrednio z własności operacji (2.4) oraz z (2.5) i (2.6). Zbiór G_1 punktów $(x_k, y_k, z_k, u_k, v_k)$, dla których spełnione są związki (2.10) jest zbiorem zwartym oraz $G_1 \subset G$.

Z zasady odwzorowań zblizających wynika twierdzenie:

Twierdzenie 7.1. Jeżeli w obszarze G funkcja $f(x, y, z, u, v)$ jest funkcją ciągłą swoich argumentów i spełnia warunek Lipschitza względem argumentów y, z, u, v oraz operacja A jest zblizającą ze stałą $0 < \gamma < 1$, to istnieje jedno i tylko jedno rozwiązanie $\langle y_k \rangle$ takie, że:

$$\langle y_k \rangle = \lim_{r \rightarrow \infty} \langle y_k^r \rangle \quad \text{dla } (k = 0, 1, \dots, n)$$

które jest rozwiązaniem układu równań (1.3), (1.4) i które można otrzymać metodą kolejnych przybliżeń przy odpowiednio dobranym przybliżeniu zerowym.

Uwaga 1. Warunek (2.9) będzie spełniony wtedy, gdy istnieją stałe Lipschitza $L_1 < 1$. Wielkość γ jest u nas na pewno większa niż jest w rzeczywistości dlatego, że otrzymano ją przez szacowanie z góry odpowiednich wielkości, w których występują zarówno dodatnie, jak i ujemne składniki. γ ulegnie jeszcze zmniejszeniu, gdy prawa strona równania (1.1) nie będzie zawierać np. pochodnej y''' . Wystarczy wtedy położyć w (2.2.) $L_3 = 0$ i rozumowanie nie ulegnie zmianie.

III.

Wykażemy teraz zbieżność rozwiązania różnicowego do rozwiązania zagadnienia brzegowego i oszacujemy błąd rozwiązania numerycznego. W tym celu poczynimy następujące założenia:

1. funkcja f spełnia założenie (A) w obszarze G ,
2. istnieje rozwiązanie $\varphi(x)$ równania (1.1) spełniające warunki brzegowe (1.2), posiadające pochodne ograniczone do 6-go rzędu włącznie,
3. problem różnicowy odpowiadający zagadnieniu (1.1) i (1.2) posiada rozwiązanie, którego wartości nie wychodzą poza obszar G .

Wprowadzimy następujące oznaczenia, Rozwiązanie problemu (1.1) i (1.2) oznaczamy przez $\varphi(x)$, a jego wartości w węzłach x_k przez $\varphi_k = \varphi(x_k)$.

$$f(x, \varphi(x), \varphi'(x), \varphi''(x)) = F(x) \text{ oraz } F(x_k) = F_k \quad (3.1)$$

$$M_k = \max_{x \in \langle 0, 1 \rangle} \left| \varphi \left(\begin{matrix} k \\ x \end{matrix} \right) \right| \quad (3.2)$$

Przybliżone rozwiązanie otrzymane metodą różnicową oznaczamy przez y_k ($k = 0, 1, \dots, n$) oraz:

$$f_k = f \left(x_k, y_k, \frac{y_{k+1} - y_{k-1}}{2h}, \frac{y_{k+1} - 2y_k + y_{k-1}}{h^2} \right) \quad (3.3)$$

$$\left. \frac{y_{k+2} - 2y_{k+1} + 2y_{k-1} - y_{k-2}}{2h^3} \right)$$

$$\varepsilon_k = \varphi_k - y_k \quad (3.4)$$

Weźmy pod uwagę wyrażenie:

$$\varphi_{k+2} - 4\varphi_{k+1} + 6\varphi_k - 4\varphi_{k-1} + \varphi_{k-2} - h^4 \varphi_k^{(IV)}$$

$$(k = 2, 3, \dots, n-2)$$

Wyrażają wartości φ_{k+2} , φ_{k+1} , φ_{k-1} i φ_{k-2} poprzez wzór Taylora dla funkcji $\varphi(x)$ w otoczeniu punktu $x_k \in \langle 0, 1 \rangle$ otrzymujemy:

$$\varphi_{k+2} - 4\varphi_{k+1} + 6\varphi_k - 4\varphi_{k-1} + \varphi_{k-2} = h^4 F_k + h^6 \varrho(k) \quad (3.5)$$

gdzie:

$$\begin{aligned} \varrho(k) = & \frac{64}{6!} \varphi^{(6)}(x_k + \theta_1 h) - \frac{4}{6!} \varphi^{(6)}(x_k + \theta_2 h) + \\ & - \frac{4}{6!} \varphi^{(6)}(x_k - \theta_3 h) + \frac{64}{6!} \varphi^{(6)}(x_k - \theta_4 h) \end{aligned}$$

oraz

$$\left| \varphi(k) \right| \leq \frac{17}{90} M_6 \quad (3.5')$$

W analogiczny sposób otrzymamy układ równań:

$$\begin{aligned} \alpha_0 \varphi_0 - \alpha_1 \frac{-3\varphi_0 + 4\varphi_1 - \varphi_2}{2h} &= a_1 + h^2 \varphi_0 \alpha_1 \\ \alpha_2 \frac{-3\varphi_0 + 4\varphi_1 - \varphi_2}{2h} - \alpha_3 \frac{2\varphi_0 - 5\varphi_1 + 4\varphi_2 - \varphi_3}{h^2} &= \\ &= a_2 + h^2 (\alpha_3 \bar{\varphi}_0 - \alpha_2 \varphi_0) \end{aligned} \quad (3.6)$$

$$\begin{aligned} \beta_0 \varphi_n + \beta_1 \frac{3\varphi_n - 4\varphi_{n-1} + \varphi_{n-2}}{2h} &= b_1 - h^2 \varphi_n \beta_1 \\ \beta_2 \frac{3\varphi_n - 4\varphi_{n-1} + \varphi_{n-2}}{2h} + \beta_3 \frac{2\varphi_n - 5\varphi_{n-1} + 4\varphi_{n-2} - \varphi_{n-3}}{h^2} &= \\ &= b_2 - h^2 (\beta_2 \varphi_n + \beta_3 \bar{\varphi}_n) \end{aligned}$$

gdzie:

$$\left| \varphi_0 \right| \leq M_3; \quad \left| \bar{\varphi}_0 \right| \leq \frac{25}{4} M_4; \quad \left| \varphi_n \right| \leq M_3; \quad \left| \bar{\varphi}_n \right| \leq \frac{25}{4} M_4 \quad (3.6')$$

Ponieważ y_k spełnia układ równań (1.3) i (1.4), więc odejmując odpowiednie równania (1.3) i (1.4) od odpowiednich równań układu (3.5) i (3.6) i korzystając z oznaczenia (3.4) otrzymamy następujący układ równań:

$$\left. \begin{aligned} \varepsilon_{k+2} - 4\varepsilon_{k+1} + 6\varepsilon_k - 4\varepsilon_{k-1} + \varepsilon_{k-2} &= h^4 (F_k - f_k) + h^6 \varphi(k) \\ \alpha_0 \varepsilon_0 - \alpha_1 \frac{-3\varepsilon_0 + 4\varepsilon_1 - \varepsilon_2}{2h} &= h^2 \varphi_0 \alpha_1 = \bar{a}_1 \end{aligned} \right\} (3.7)$$

$$\alpha_2 \frac{-3\varepsilon_0 + 4\varepsilon_1 - \varepsilon_2}{2h} - \alpha_3 \frac{2\varepsilon_0 - 5\varepsilon_1 + 4\varepsilon_2 - \varepsilon_3}{h^2} =$$

$$= h^2(\alpha_3 \bar{\rho}_0 - \alpha_2 \rho_0) = \bar{a}_2$$

$$\beta_0 \varepsilon_n + \beta_1 \frac{3\varepsilon_n - 4\varepsilon_{n-1} + \varepsilon_{n-2}}{2h} = -\beta_1 h^2 \rho_n = \bar{b}_1 \quad (3.7)$$

$$\beta_2 \frac{3\varepsilon_n - 4\varepsilon_{n-1} + \varepsilon_{n-2}}{2h} + \beta_3 \frac{2\varepsilon_n - 5\varepsilon_{n-1} + 4\varepsilon_{n-2} - \varepsilon_{n-3}}{h^2} =$$

$$= -h^2(\beta_2 \rho_n + \beta_3 \bar{\rho}_n) = \bar{b}_2$$

Prowadząc podobne rozumowanie dla układu równań (3.7) jak dla układu równań (1.5) otrzymamy:

$$\varepsilon_k = \frac{1}{\Delta} \left(\bar{\Delta}_0 + \bar{\Delta}_1 k + \bar{\Delta}_2 k^2 + \bar{\Delta}_3 k^3 \right) + \quad (3.8)$$

$$+ h^4 \sum_{i=2}^{n-2} \varepsilon_{ik} (F_i - f_i) + h^6 \sum_{i=2}^{n-2} \varepsilon_{ik} \rho_{(i)} \quad (1)$$

Wyznacznik Δ jest określony wzorem (1.10), ε_{ik} są funkcjami określonymi w (1.18), pozostałe wyznaczniki $\bar{\Delta}_i$ ($i = 0, 1, 2, 3$) są wyznacznikami odpowiadającymi wyznacznikom Δ_i występującym w (1.9), z tym, że zamiast a_1, a_2, b_1, b_2 wchodzi wielkości $\bar{a}_1, \bar{a}_2, \bar{b}_1, \bar{b}_2$ określone w (3.7).

Oszacujemy teraz błąd ε_k :

$$\left| \varepsilon_k \right| = \left| \frac{1}{\Delta} \left(\bar{\Delta}_0 + \bar{\Delta}_1 k + \bar{\Delta}_2 k^2 + \bar{\Delta}_3 k^3 \right) + \quad (3.9)$$

$$+ h^4 \sum_{i=2}^{n-2} \varepsilon_{ik} (F_i - f_i) + h^6 \sum_{i=2}^{n-2} \varepsilon_{ik} \rho_{(i)} \right| \leq$$

$$\leq \frac{1}{\Delta} \left(|\bar{\Delta}_0| + k |\bar{\Delta}_1| + k^2 |\bar{\Delta}_2| + k^3 |\bar{\Delta}_3| \right) +$$

$$+ h^4 \sum_{i=2}^{n-2} |\varepsilon_{1k}| \cdot |F_i - f_i| + h^6 \sum_{i=2}^{n-2} |\varepsilon_{1k}| \cdot |\varphi(1)|$$

Szacując kolejno każdy ze składników prawej strony nierówności (3.9) otrzymamy:

$$|\bar{\Delta}_0| \leq h^2 (E_0 + T_0) \quad (3.10)$$

gdzie:

$$\left. \begin{aligned} E_0 &= 4\alpha_1\alpha_3 \left(4\beta_0\beta_2 + 6\beta_0\beta_3 + 9\beta_1\beta_2 + 12\beta_1\beta_3 + 2\beta_0\beta_2 h^2 \right) M_3 \\ T_0 &= \frac{25}{4} \alpha_1\alpha_3 \left(5\beta_0\beta_2 + 10\beta_0\beta_3 + 12\beta_1\beta_2 + 24\beta_1\beta_3 + \right. \\ &\quad \left. + 4\beta_0(\beta_2 + 2\beta_3)h^2 \right) M_4 \end{aligned} \right\} (3.10')$$

$$k \cdot |\bar{\Delta}_1| \leq h^2 (E_1 + T_1) \quad (3.11)$$

gdzie:

$$\left. \begin{aligned} E_1 &= \left(\alpha_0\alpha_2 \left(5\beta_0\beta_2 + 4\beta_1\beta_2 (3 - 2h^2) + 8\beta_0\beta_3 + 18\beta_1\beta_3 \right) + \right. \\ &\quad \left. + 2\alpha_0\alpha_3 \left(\beta_0\beta_2 + 2\beta_1\beta_2 (3 - 2h^2) + 6\beta_1\beta_3 \right) + \right. \\ &\quad \left. + 6\alpha_1\alpha_3\beta_0(\beta_2 + 2\beta_3) \right) \cdot M_3 \\ T_1 &= \frac{25}{4} \left(2\alpha_0\alpha_2(\beta_0\beta_3 + 2\beta_1\beta_3) + \right. \\ &\quad \left. + \alpha_0\alpha_3(5\beta_0\beta_2 + 12\beta_1\beta_2 + 10\beta_0\beta_2 + 24\beta_1\beta_3 + \right. \\ &\quad \left. + 2(\beta_0\beta_2 + 4\beta_1\beta_2 + 4\beta_1\beta_3)h^2 \right) + \\ &\quad \left. + 4\alpha_1\alpha_3\beta_0(\beta_2 + 2\beta_3)h^2 \right) M_4 \end{aligned} \right\} (3.11')$$

$$k^2 |\bar{\Delta}_2| \leq h^2 (E_2 + T_2) \quad (3.12)$$

gdzie:

$$\left. \begin{aligned} E_2 &= \alpha_0 \alpha_2 (\beta_0 \beta_2 (5 + 4h^2) + 6 \beta_0 \beta_3 + 12 \beta_1 \beta_3) M_3 \\ T_2 &= \frac{25}{4} (\alpha_0 \alpha_2 (\beta_0 \beta_3 (1 + h^2) + 3 \beta_1 \beta_3) + \\ &+ 2 \alpha_0 \alpha_3 (\beta_0 \beta_2 (2 + h^2) + 3 \beta_0 \beta_3 + 3 \beta_1 \beta_2 + 3 \beta_1 \beta_3) + \\ &+ 3 \alpha_1 \alpha_3 \beta_0 (\beta_2 + 2 \beta_3)) M_4 \end{aligned} \right\} (3.12')$$

$$k^3 |\bar{\Delta}_3| \leq h^2 (E_3 + T_3) \quad (3.13)$$

gdzie:

$$\left. \begin{aligned} E_3 &= (\alpha_0 \alpha_2 (4 \beta_0 \beta_2 + 7 \beta_1 \beta_2 + 2 \beta_0 \beta_3 + 2 \beta_1 \beta_3) + \\ &+ \alpha_0 \alpha_3 \beta_2 (2 \beta_0 + 3 \beta_1) + 4 \alpha_1 \alpha_3 \beta_0 \beta_2) M_3 \\ T_3 &= \frac{25}{4} (\alpha_0 \alpha_2 (\beta_0 \beta_2 + 2 \beta_1 \beta_3) + \\ &+ \alpha_0 \alpha_3 (3 \beta_0 \beta_2 + 4 \beta_0 \beta_3 + 4 \beta_1 \beta_2 + 4 \beta_1 \beta_3) + \\ &+ 2 \alpha_1 \alpha_3 \beta_0 (\beta_2 + 2 \beta_3)) M_4 \end{aligned} \right\} (3.13')$$

stąd:

$$\frac{1}{\Delta} (|\bar{\Delta}_0| + k |\bar{\Delta}_1| + k^2 |\bar{\Delta}_2| + k^3 |\bar{\Delta}_3|) \leq \frac{h^2}{\Delta} \sum_{j=0}^3 (E_j + T_j) \quad (3.14)$$

Funkcja $F_1 - f_1$ z założenia spełnia warunek Lipschitza.
Zatem:

$$\left. \begin{aligned}
 h^4 \sum_{i=2}^{n-2} \left| \varepsilon_{ik} \right| \left| F_i - f_i \right| &\leq \sum_{i=2}^{n-2} \left| \varepsilon_{ik} \right| \left(L_0 \left| \varphi_i - y_i \right| + \right. \\
 + L_1 \left| \varphi'_i - \frac{y_{i+1} - y_{i-1}}{2h} \right| + L_2 \left| \varphi''_i - \frac{y_{i+1} - 2y_i + y_{i-1}}{h^2} \right| + \\
 \left. + L_3 \left| \varphi'''_i - \frac{y_{i+2} - 2y_{i+1} + 2y_{i-1} - y_{i-2}}{2h^3} \right| \right) & \quad (3.15)
 \end{aligned} \right\}$$

Ponieważ:

$$\begin{aligned}
 \varphi'_i - \frac{y_{i+1} - y_{i-1}}{2h} &= \varphi'_i - \frac{\varphi_{i+1} - \varphi_{i-1}}{2h} + \frac{\varphi_{i+1} - \varphi_{i-1}}{2h} + \\
 - \frac{y_{i+1} - y_{i-1}}{2h} &= \varphi'_i - \frac{\varphi_{i+1} - \varphi_{i-1}}{2h} + \frac{\varepsilon_{i+1} - \varepsilon_{i-1}}{2h} = \\
 &= h^2 \mu_1 + \frac{\varepsilon_{i+1} - \varepsilon_{i-1}}{2h}
 \end{aligned}$$

gdzie:

$$\left| \mu_1 \right| = \frac{1}{2} \left| \frac{\varphi'''(x_i + \theta_1 h)}{3!} + \frac{\varphi'''(x_i - \theta_2 h)}{3!} \right| \leq \frac{1}{6} M_3$$

$$\varphi''_i - \frac{y_{i+1} - 2y_i + y_{i-1}}{h^2} = h^2 \mu_2 + \frac{\varepsilon_{i+1} - 2\varepsilon_i + \varepsilon_{i-1}}{h^2}$$

gdzie:

$$\left| \mu_2 \right| \leq \frac{1}{12} M_4$$

$$\begin{aligned}
 \varphi'''_i - \frac{y_{i+2} - 2y_{i+1} + 2y_{i-1} - y_{i-2}}{2h^3} &= h^2 \mu_3 + \\
 + \frac{\varepsilon_{i+2} - 2\varepsilon_{i+1} + 2\varepsilon_{i-1} - \varepsilon_{i-2}}{2h^3}
 \end{aligned}$$

$$\text{gdzie: } \left| \mu_3 \right| \leq \frac{8}{17} M_5$$

Wielkości μ_i ($i = 1, 2, 3$) są błędami powstałymi z aproksymacji pochodnych odpowiednimi różnicami symetrycznymi. Uwzględniając te rachunki, nierówność (3.15) przyjmie postać:

$$\begin{aligned} h^4 \sum_{i=2}^{n-2} |\varepsilon_{ik}| |F_1 - f_1| &\leq h^4 \sum_{i=2}^{n-2} |\varepsilon_{ik}| \left(L_0 |\varepsilon_i| + L_1 \left| \frac{\varepsilon_{i+1} - \varepsilon_{i-1}}{2h} \right| + \right. \\ &+ L_2 \left| \frac{\varepsilon_{i+1} - 2\varepsilon_i + \varepsilon_{i-1}}{h^2} \right| + L_3 \left| \frac{\varepsilon_{i+2} - 2\varepsilon_{i+1} + 2\varepsilon_{i-1} - \varepsilon_{i-2}}{2h^3} \right| + \\ &+ h^2 \left(L_1 |\mu_1| + L_2 |\mu_2| + L_3 |\mu_3| \right) \Big) \leq B_0 \left(L_0 \max_{0 \leq i \leq n} |\varepsilon_i| + \right. \\ &+ L_1 \max_{1 \leq i \leq n-1} \left| \frac{\varepsilon_{i+1} - \varepsilon_{i-1}}{2h} \right| + L_2 \max_{1 \leq i \leq n-1} \left| \frac{\varepsilon_{i+1} - 2\varepsilon_i + \varepsilon_{i-1}}{h^2} \right| + \\ &+ L_3 \max_{2 \leq i \leq n-2} \left| \frac{\varepsilon_{i+2} - 2\varepsilon_{i+1} + 2\varepsilon_{i-1} - \varepsilon_{i-2}}{2h^3} \right| + \\ &+ h^2 \left(\frac{1}{6} L_1 M_3 + \frac{1}{12} L_2 M_4 + \frac{8}{17} L_3 M_5 \right) \Big) \end{aligned}$$

Na mocy (2.2) mamy:

$$\begin{aligned} h^4 \sum_{i=2}^{n-2} |\varepsilon_{ik}| \cdot |F_1 - f_1| &\leq B_0 \left(\varphi(\varepsilon, 0) + \right. \\ &+ h^2 \left(\frac{1}{6} L_1 M_3 + \frac{1}{12} L_2 M_4 + \frac{8}{17} L_3 M_5 \right) \Big) \quad (3.16) \end{aligned}$$

gdzie: $\varphi(\varepsilon, 0)$ jest odległością określoną w (2.2) punktu $\varepsilon = (\varepsilon_0, \varepsilon_1, \dots, \varepsilon_n)$ od punktu $0 = (\underbrace{0, 0, \dots, 0}_{n+1})$ przestrzeni R^{n+1}

Z (2.1) oraz z (3.5') otrzymujemy:

$$h^6 \sum_{i=2}^{n-2} |\varepsilon_{ik}| \cdot |\varphi(\underline{1})| \leq \frac{17}{90} M_6 B_0 h^2 \quad (3.17)$$

Biorąc pod uwagę związki (3.14), (3.16), i (3.17) nierówność (3.9) przyjmie postać:

$$\begin{aligned} |\varepsilon_k| &\leq \frac{h^2}{\Delta} \sum_{j=0}^3 (E_j + T_j) + B_0 \varphi(\varepsilon, 0) + \\ &+ h^2 B_0 \left(\frac{1}{6} L_1 M_3 + \frac{1}{12} L_2 M_4 + \frac{8}{17} L_3 M_5 + \frac{17}{90} M_6 \right) \\ |\varepsilon_k| &\leq \frac{h^2}{\Delta} \sum_{j=0}^3 (E_j + T_j) + B_0 \varphi(\varepsilon, 0) + h^2 B_0 H \end{aligned} \quad (3.18)$$

gdzie:

$$H = \frac{1}{6} L_1 M_3 + \frac{1}{12} L_2 M_4 + \frac{8}{17} L_3 M_5 + \frac{17}{90} M_6$$

Związek (3.18) nie pozwala jeszcze oszacować błędu ε_k , gdyż nie wiemy jak zachowuje się $\varphi(\varepsilon, 0)$ gdy $h \rightarrow 0$. Należy zatem ocenić $\varphi(\varepsilon, 0)$. W tym celu rozpatrzmy jeszcze następujące trzy wyrażenia. Na mocy związku (3.8) i po dodatkowych przekształceniach związków (3.10) - (3.13) otrzymujemy:

$$\begin{aligned} \left| \frac{\varepsilon_{k+1} - \varepsilon_{k-1}}{2h} \right| &\leq \frac{1}{h\Delta} \left(\left| \bar{\Delta}_1 \right| + 2k \left| \bar{\Delta}_2 \right| + (3k^2 + 1) \left| \bar{\Delta}_3 \right| + \right. \\ &+ h^4 \sum_{i=2}^{n-2} \left| \frac{\varepsilon_{ik+1} - \varepsilon_{ik-1}}{2h} \right| \left| F_1 - f_1 \right| + \\ &+ h^6 \sum_{i=2}^{n-2} \left| \frac{\varepsilon_{ik+1} - \varepsilon_{ik-1}}{2h} \right| \left| \varphi(\underline{1}) \right| \leq \end{aligned} \quad (3.19)$$

$$\leq \frac{h^2}{\delta} \left((E_1 + T_1) + 2(E_2 + T_2) + (3 + h^2)(E_3 + T_3) \right) + B_1 \varphi(\varepsilon, 0) + B_1 H h^2$$

$$\left| \frac{\varepsilon_{k+1} - 2\varepsilon_k + \varepsilon_{k-1}}{h^2} \right| \leq \frac{2}{h^2 \Delta} \left(|\bar{\Delta}_2| + 3k |\bar{\Delta}_3| \right) + h^4 \sum_{i=2}^{n-2} \left| \frac{\varepsilon_{1k+1} - 2\varepsilon_{1k} + \varepsilon_{1k-1}}{h^2} \right| \left(|F_1 - f_1| + h^2 |\varphi(1)| \right) \leq (3.20)$$

$$\leq \frac{2h^2}{\delta} \left((E_2 + T_2) + 3(E_3 + T_3) \right) + B_2 \varphi(\varepsilon, 0) + B_2 H h^2$$

$$\left| \frac{\varepsilon_{k+2} - 2\varepsilon_{k+1} + 2\varepsilon_{k-1} - \varepsilon_{k-2}}{2h^3} \right| \leq \frac{6}{h^3 \Delta} |\bar{\Delta}_3| + h^4 \sum_{i=2}^{n-2} \left| \frac{\varepsilon_{1k+2} - 2\varepsilon_{1k+1} + 2\varepsilon_{1k-1} - \varepsilon_{1k-2}}{2h^3} \right| \left(|F_1 - f_1| + h^2 |\varphi(1)| \right) \leq (3.21)$$

$$\leq \frac{6h^2}{\delta} (E_3 + T_3) + B_3 \varphi(\varepsilon, 0) + B_3 H h^2$$

Nierówności (3.18), (3.19), (3.20) i (3.21) są prawdziwe dla każdego k , więc są prawdziwe również dla maksimum ze względu na k . Mnożąc każdą z tych nierówności odpowiednio przez L_0 , L_1 , L_2 , L_3 i dodając stronami takie nierówności otrzymamy następującą zależność:

$$\varphi(\varepsilon, 0) \leq \gamma \varphi(\varepsilon, 0) + h^2 H \gamma + \frac{h^2}{\delta} \left(L_0 (E_0 + T_0) + (L_0 + L_1)(E_1 + T_1) + (L_0 + 2L_1 + 2L_2)(E_2 + T_2) + (L_0 + 3L_1 + 6L_2 + 6L_3 + h^2 L_1)(E_3 + T_3) \right)$$

Ponieważ $0 < \gamma < 1$, więc:

$$\begin{aligned} \varphi(\varepsilon, 0) &\leq \frac{h^2}{1-\tau} \left(\tau H + \frac{1}{\delta} \left(L_0 (E_0 + T_0) + \right. \right. \\ &+ (L_0 + L_1)(E_1 + T_1) + (L_0 + 2L_1 + 2L_2)(E_2 + T_2) + \\ &\left. \left. + (L_0 + 3L_1 + 6L_2 + 6L_3 + h^2 L_1)(E_3 + T_3) \right) \right) \end{aligned} \quad (3.22)$$

Z (3.22) wynika, że gdy $h \rightarrow 0$, to również $\varphi(\varepsilon, 0) \rightarrow 0$, a stąd i z (3.12) wynika, że $|\varepsilon_k| \rightarrow 0$ i to tak szybko, jak $h^2 \rightarrow 0$.

Oprócz tego z (3.19), (3.20), (3.21) widzimy, że gdy $h \rightarrow 0$, to:

$$\left| \frac{\varepsilon_{i+1} - \varepsilon_{i-1}}{2h} \right| \rightarrow 0, \quad \left| \frac{\varepsilon_{i+1} - 2\varepsilon_i + \varepsilon_{i-1}}{2h} \right| \rightarrow 0,$$

$$\left| \frac{\varepsilon_{k+2} - 2\varepsilon_{k+1} + 2\varepsilon_{k-1} + \varepsilon_{k-2}}{2h^3} \right| \rightarrow 0$$

W ten sposób dla każdego k również:

$$y_k \rightarrow \varphi_k, \quad \frac{y_{k+1} - y_{k-1}}{2h} \rightarrow \varphi'_k, \quad \frac{y_{k+1} - 2y_k + y_{k-1}}{h^2} \rightarrow \varphi''_k$$

$$\frac{y_{k+2} - 2y_{k+1} + 2y_{k-1} - y_{k-2}}{2h^3} \rightarrow \varphi'''_k$$

Jeżeli oprócz tego do prawej strony nierówności (3.18) położymy w miejsce $\varphi(\varepsilon, 0)$ prawą stronę nierówności (3.22), to otrzymamy oszacowanie błędu odchylenia rozwiązania numerycznego od rozwiązania zagadnienia brzegowego.

Oszacowanie błędu jest postaci:

$$\begin{aligned}
|\epsilon_k| = |y_k - \varphi_k| &\leq \frac{h^2}{\delta} \sum_{j=0}^3 (E_j + T_j) + \\
&+ \frac{h^2}{1-\gamma} B_0 \left(\gamma H + \frac{1}{\delta} \left(L_0 (E_0 + T_0) + (L_0 + L_1)(E_0 + T_0) + \right. \right. \\
&\left. \left. + (L_0 + 2L_1 + 2L_2)(E_2 + T_2) + \right. \right. \\
&\left. \left. + (L_0 + 3L_1 + 6L_2 + 6L_3 + h^2 L_1) \cdot (E_3 + T_3) \right) \right) + B_0 h^2 H
\end{aligned} \tag{3.23}$$

Ponieważ δ oraz γ nie zależy od h , zatem z (3.23) widzimy, że błąd ϵ_k dla każdego k wyrażony jest z dokładnością do h^2 , czyli z tą samą dokładnością, z jaką aproksymowane było równanie (1.1) i warunki brzegowe (1.2). Wynika stąd w szczególności, że gdy $h \rightarrow 0$, to oszacowanie błędu w każdym punkcie $x_k \in \langle 0, 1 \rangle$ dąży do zera tak szybko, jak h^2 . Stąd:

Twierdzenie 8.1. Jeżeli spełnione są założenia twierdzenia 7.1 i jeżeli funkcja f posiada ciągle i ograniczone pochodne w obszarze G do drugiego rzędu włącznie, to błąd między rozwiązaniem numerycznym (1.22) i rozwiązaniem zagadnienia brzegowego (1.1) i (1.2) maleje do zera tak szybko jak h^2 .

Literatura

- [1] GELFOND D.O.: Rachunek różnic skończonych, Moskwa 1952.
- [2] COLLATZ L.: Metody numeryczne rozwiązywania równań różniczkowych, Warszawa 1960.
- [3] DEMIDOWICZ B.P.; MARON I.A., SZUWAŁOWA E.I.: Metody numeryczne, Moskwa 1963.
- [4] LUSTERNIK L.A., SOBOLEW W.I.: Elementy analizy funkcjonalnej, Moskwa 1951.
- [5] GODUNOV S.K., RABIENKIJ W.S.: Wstęp do teorii schematów różnicowych, Moskwa 1962.
- [6] KAMKE E.: Równania różniczkowe zwyczajne, Moskwa 1961.
- [7] KANTOROWICZ Ł.W., KRYŁOW W.U.: Przybliżone metody wyższej analizy, Moskwa, Leningrad 1952.
- [8] BEREZIN I.S., ZIEDKOW H.P.: Metody numeryczne, t. I i II, Moskwa 1962
- [9] Referatywny Żurnal, Matematyka (1963-1967).

НЕКИЙ МЕТОД ЧИСЛЕННОГО РЕШЕНИЯ БЕРЕГОВОЙ ПРОБЛЕМЫ ДЛЯ НЕЛИНЕЙНОГО-ДИФФЕРЕНЦИАЛЬНОГО ПРЯМОГО УРАВНЕНИЯ IV РЯДА**Резюме**

В работе решен береговой вопрос для нелинейного дифференциального простого уравнения IV ряда методом простой итерации. Эта итерация получена через аппроксимацию с рядом h^2 производных, выступающих в дифференциальной проблеме, отношениями разностей. Дифференциальная аналогия функции Грина определенная в виде формулы (1.13) позволяет получить открытую итерационную схему. Одновременно была решена поставленная проблема. Доказано совпадение принятого метода итерации, исследованное в норме C_2^1 . Потом была сделана оценка ошибки для численного решения и указана в виде формулы (3.23) (диссертация).

A CERTAIN METHOD OF A NUMERICAL SOLUTION OF A BORDER PROBLEM FOR A SIMPLE NONLINEAR DIFFERENTIAL EQUATION OF THE IV RANGE

Summary

In the paper a border problem for a simple nonlinear differential equation of the IV range is solved by the method of straight iteration. The latter was obtained through approximation with h^2 range of derivatives occurring the differential problem as difference quotients. The differential analogy of Green's function defined by formula 1.13 permits to obtain an open iteration scheme. Simultaneously an explicit solution of the brought forward problem was obtained, the convergence of the accepted iteration method, tested in C_2^1 standard was shown. Then the error for a numerical solution was estimated and given in formula (3.23) (thesis).

JĘZYK MODELWANIA SYMULACYJNEGO
PROCESÓW DYSKRETNÝCH SYM 69

Jerzy LESZCZYŃSKI

Centralny Ośrodek Badań i Rozwoju
Techniki Kolejnictwa
Warszawa, ul. Chłopickiego 50

W pracy przedstawiono język i translator SYM 69 opracowany w Centralnym Ośrodku Badań i Rozwoju Techniki Kolejnictwa do modelowania układów zdarzeń charakterystycznych dla działania obiektów transportu. Język SYM 69 z definicji stanowi nadbudowę na język ALGOL 60. Translator języka został wykonany dla EMC ODRA 1204.

Język SYM 69 służy do opisu zadań symulacji zachowania się w czasie (modelowania symulacyjnego) złożonych obiektów przedstawionych w postaci modeli układów zdarzeń. Potrzeba jego opracowania w Centralnym Ośrodku Badań i Rozwoju Techniki Kolejnictwa wynikała na tle badań sieci transportowych, stacji i linii kolejowych oraz innych obiektów (systemów) o licznych i złożonych powiązaniach pomiędzy elementami, skomplikowanych procedurach działania i hierarchicznych systemach kierowania. Jednakże zakres zastosowań powstałego w wyniku aparatu programowania nie ogranicza się do obiektów transportu, będąc raczej typowym dla języków klasy SIMULA, SOL, SIMSCRIPT.

Język SYM 69 w zasadzie przezręczony jest dla specjalistów w zakresie modelowania. Pozwala on na łatwe sformalizowanie opisu modelu z języka potocznego na opis sformalizowany. Metodycznie język SYM 69 ułatwia kontakt specjalistom z zakresu modelowania ze specjalistami stawiającymi zadanie,

stanowiąc swoistą płaszczyznę dyskusji i umożliwiając formułowanie zadania symulacji metodą kolejnych prób i przybliżeń.

Wykonany został translator języka SYM 69 dla EMC ODRA 1204. Translator ten traktowany początkowo jako doświadczalny okazał się w praktyce efektywnym narzędziem pozwalającym na wykorzystanie EMC ODRA 1204 w badaniach symulacyjnych złożonych systemów. Translator eksploatowany jest od połowy 1969 roku.

Przedstawiony poniżej opis języka SYM 69 zakłada u czytelnika znajomość języka ALGOL 60 w zakresie przynajmniej popularnych podręczników.

Język SYM 69 powstaje z języka ALGOL 60 przez dołączenie pojęć umożliwiających ujęcie opisem wielu czynności zachodzących równolegle oraz dynamiczne operowanie skomplikowanymi strukturami miejsc roboczych. Te cechy języka osiągnięto wprowadzając:

- strukturę process,
- opis set,
- procedury standardowe hold, wait, acti oraz funkcje standardowe curr, time,
- funkcje standardowe trans, occup, free, card.

Ponadto wprowadzono szereg procedur i funkcji standardowych organizujących współpracę z wejściem-wyjściem maszyny oraz ułatwiających programowanie zadań. Elementy te, nie wnoszące nowych cech jakościowych, zostały w niniejszym opisie pominięte.

Zainteresowanego szczegółami czytelnika odsyłamy w tym miejscu do instrukcji programowania w języku SYM 69 [2]. Program zapisany w języku SYM 69 nazywany modelem.

Na opis modelu składa się

- opis czynności realizowanych w modelu,
- opis przedmiotów występujących w modelu.

Zbiór czynności realizowanych w modelu jest podzielony na rozłączne klasy (procesy). Czynności jednej klasy są opisane wspólną procedurą postępowania, stanowiącą opis procesu.

Przedmioty występują w modelu jako przedmioty globalne, to znaczy wspólne dla wszystkich czynności realizowanych w modelu oraz jako przedmioty lokalne, związane z określonymi czynnościami i niedostępne z innych czynności modelu.

Rozpoczęcie lub wznowienie wykonywania czynności jest zdarzeniem zachodzącym w modelu.

Czas zmienia się w modelu dyskretnie od zdarzenia do zdarzenia.

Opis procesu ma budowę podobną do opisu procedury w rozumieniu definicji języka ALGOL 60. Opis procesu nie może zawierać opisu innego procesu, to znaczy nie istnieją procesy lokalne. Parametry formalne występujące w opisie procesu mogą być tylko przedmiotami, a podstawienia mogą się odbywać tylko przez wartość. Z tego powodu pomija się w opisie procesu specyfikacje parametrów formalnych i listę wartości.

Przedmioty globalne opisane są bezpośrednio po symbolu begin otwierającym blok będący programem, a przedmioty lokalne bezpośrednio po symbolach begin otwierających bloki będące ciałami procesów. W innych miejscach programu opisy przedmiotów nie mogą się pojawić.

Wprowadzone wyżej pojęcia w niczym nie ograniczają składni języka ALGOL 60. W odniesieniu do elementów składni języka SYM 69 opis modelu przybiera postać:

```
begin set AA, BB, CC;
```

```
...
```

```
process ALFA;
```

```
  begin set A, B, C;
```

```
  ...
```

```
  end ALFA;
```

```
process BETA;
```

```
  begin set A, B, C;
```

```
...
```


end BETA;

...

end MODEL;

W przedstawionym wyżej schemacie AA, BB, CC są nazwami przedmiotów globalnych, ALFA, BETA - nazwami procesów, a A, B, C - nazwami przedmiotów lokalnych. Miejsca wykropkowane mogą zawierać dowolne elementy języka ALGOL 60 oraz elementy składniowe zawierające instrukcje procesu, nazwy przedmiotów, funkcje i procedury standardowe języka SYM 69. Instrukcja procesu ma postać instrukcji procedury lub funkcji, a nazwa przedmiotu postać nazwy tablicy języka ALGOL 60.

Przedmiot jest elementem lub zbiorem elementów uporządkowanym hierarchicznie w postać swoistego drzewa. Celem zdefiniowania sposobu uporządkowania tego zbioru wprowadza się pojęcie poziomu q_1 rozumiane jako i -ty poziom zbioru X, gdzie X jest nazwą w rozumieniu języka ALGOL 60.

Na poziomie q_0 zbioru X istnieje tylko jeden element, który będziemy nazywać przedmiotem X. Poziom q_1 jest uporządkowanym ciągiem elementów ponumerowanych $0, 1, \dots, n$. Elementy te będziemy nazywać atrybutami przedmiotu X. Każdy z atrybutów poziomu q_1 jest przedmiotem $X[0], X[1], \dots, X[n]$ dla ciągów atrybutów poziomu q_2 . Ciągów atrybutów na poziomie q_2 może być tyle, ile elementów zawiera ciąg na poziomie q_1 , to znaczy n . Każdy z atrybutów poziomu q_2 może być z kolei przedmiotem dla atrybutów poziomu niższego. Zbiór może być rozbudowywany na dowolną głębokość.

Dołączanie atrybutów jest rozbudową, a odłączanie - redukcją przedmiotu X albo inaczej zbioru X. Rozbudowę można rozwijać bez ograniczeń, przy czym zawsze każdemu z ciągów poziomu q_1 odpowiadają przedmioty tworzące jeden z ciągów poziomu q_{1-1} . Redukcję można realizować odłączając przedmioty. Odłączenie przedmiotu będącego jednym z elementów jednego z ciągów poziomu q_1 odłącza również związany z nim ciąg atrybutów poziomu q_{1+1} . Rozbudowę przedmiotu można realizować przez

dołączanie przedmiotów na dowolnym z poziomów. Dołączenie przedmiotu do jednego z ciągów poziomu q_1 dołącza również związany z nim ciąg atrybutów na poziomie q_{1+1} .

Ciągi elementów (przedmiotów, atrybutów) mogą być traktowane jak kolejki, co pozwala na stosowanie określonego zbioru strategii definiujących sposoby dołączania i odłączania elementów w pojęciach teorii kolejek.

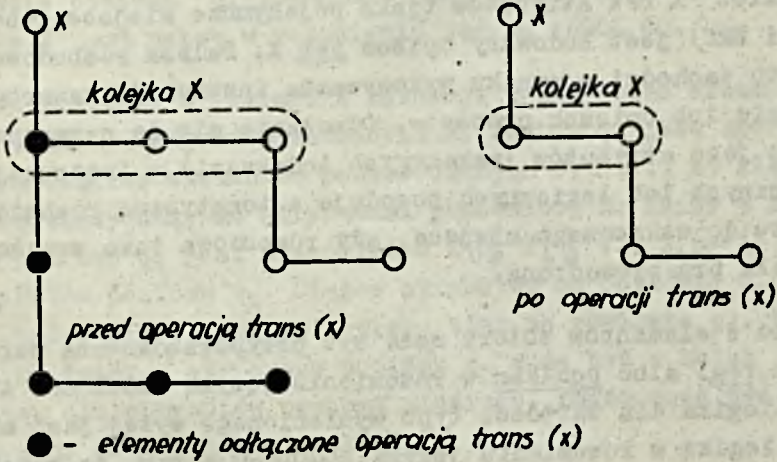
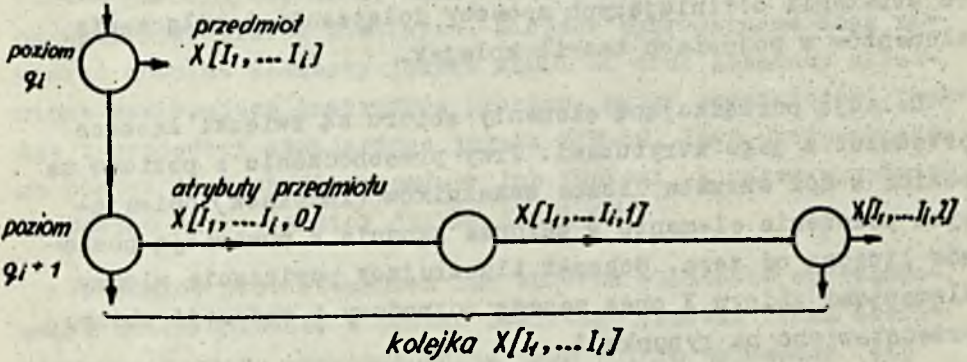
Relacją porządkującą elementy zbioru są związki łączące przedmiot z jego atrybutami. Przy przechodzeniu z poziomu na poziom w dół wzrasta liczba wskaźników (indeksów) opisujących położenie elementu w zbiorze zgodnie z numeracją poziomów liczoną od zera. Schemat ilustrujący powiązania między elementami zbioru X oraz zasadę rozbudowy i redukcji zbioru przedstawiono na rysunku 1.

Przedmiot X bez atrybutów (jako pojedyncze miejsce robocze w pamięci EMC) jest budowany opisem set X. Dalsza rozbudowa przedmiotu zachodzi w wyniku wykonywania instrukcji zawartych w programie lub opisach procesów. Odwołanie się do nazwy przedmiotu lub jego atrybutów (wskazanych indeksami) w instrukcjach arytmetycznych lub logicznych powoduje automatyczną rozbudowę przedmiotu do wskazanego miejsca, gdy rozbudowa taka wcześniej nie została przeprowadzona.

Każdemu z elementów zbioru może być przyporządkowana wartość typu real albo boolean w rozumieniu języka ALGOL 60. Arytmetyka i logika dla wartości typu wymienionego wyżej jest arytmetyką i logiką w rozumieniu języka ALGOL 60 z tym, że każdemu elementowi można przyporządkować wartość dowolnego typu. Typ wartości kontrolowany jest dynamicznie, to znaczy w czasie wykonywania programu (symulacji). Element, któremu nie przyporządkowano żadnej wartości ma, z definicji rozbudowy zbioru, przyporządkowaną wartość zero typu real.

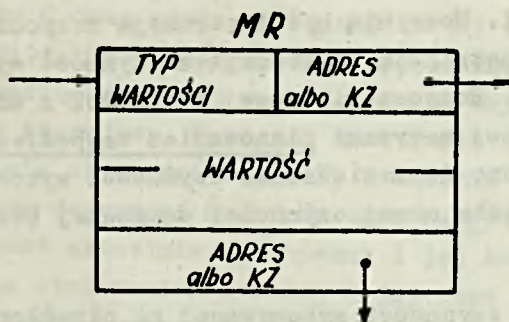
Ponadto wprowadza się wartość typu name, która wykorzystywana jest do pamiętania nazw czynności. W ten sposób uzyskuje się możliwość identyfikowania czynności w klasie (procesie).

Na wartościach typu name można wykonywać tylko operacje pamiętania i porównania. Wartość ta jest generowana i wykorzystywana poprzez instrukcje odwołujące się do specjalnych procedur standardowych nadających wartości tego typu elementom zbiorów.



Rys. 1. Ilustracja powiązań między elementami zbioru

W pamięci maszyny każdy element zbioru jest przedstawiony miejscem roboczym MR zawierającym cztery słowa maszynowe. Schemat miejsca roboczego MR przedstawiono na rysunku 2.



Rys. 2. Schemat miejsca roboczego MR

W pierwszym słowie jest zapamiętany typ wartości i adres prawego sąsiada albo wskaźnik końca zbioru KZ. Następne dwa słowa zawierają wartość przyporządkowaną MR. Ostatnie słowo zawiera adres pierwszego MR w ciągu cech przyporządkowanych przedmiotowi albo wskaźnik KZ, gdy atrybuty przedmiotu nie istnieją.

Czynność jest dynamiczną postacią procesu występującą w czasie wykonywania programu (symulacji). Sterowanie przebiegiem czynności jest automatycznie skojarzone z czasem, który w ten sposób traktowany jest jak zmienna niezależna modelu.

Sterowanie przebiegiem czynności sprowadza się do rozpoczęcia, zawieszania, wznowiania i zakańczania czynności.

Rozpoczęcie czynności następuje w wyniku wykonania instrukcji w innej czynności tej samej lub innej klasy i w rezultacie zawieszania poprzednio wykonywanej czynności. Rozpoczęcie czynności wiąże się z utworzeniem jej miejsc roboczych jako przedmiotów lokalnych oraz rozpoczęciem wykonywania instrukcji klasy (procesu), do której należy rozpoczynana czynność. Kolejne instrukcje arytmetyczne i logiczne są wykonywane (mówimy, że

czynność jest wykonywana) do instrukcji przerywającej wykonywanie czynności. Może nią być instrukcja rozpoczynająca nową czynność lub instrukcja zawieszająca czynność wykonywaną na określony czas. Rozpoczęcie nowej czynności z wnętrza czynności wykonywanej nazywamy planowaniem bezpośrednim, gdyż zawsze prowadzi ono do zawieszenia czynności wykonywanej i rozpoczęcie wykonania nowej czynności wskazanej bezpośrednio klasą.

Zawieszenie czynności wykonywanej na określony czas nazywamy planowaniem pośrednim, gdyż w rezultacie jest wznawiana inna czynność bezpośrednio nie wskazana, a wynikająca z aktualnego stanu wszystkich innych czynności.

Planowanie bezpośrednio realizowane jest przez proste odwołanie się do nazwy procesu. Planowanie pośrednie wymaga odwołania się do specjalnych procedur standardowych SYM 69, których parametrem jest czas.

Czynność może być wielokrotnie zawieszana i wznawiana, zgodnie z procedurą stanowiącą opis procesu. Dojście z wykonywaniem czynności do symbolu end kończącego opis procesu powoduje zakończenie czynności i przejście do wykonywania innej czynności, tak jak przy planowaniu pośrednim. Jednocześnie czynność zakończona jest wymazywana z modelu, a jej miejsca robocze (przedmioty lokalne) i miejsca zajęte przez sterowanie wykonaniem czynności zostają zwolnione i mogą być wykorzystywane przez inne czynności modelu. Sterowanie przebiegiem symulacji w sensie technicznym realizowane jest przez program sterujący interpretacyjnym wykonywaniem instrukcji zawartych w opisach klas czynności (procesów) oraz kontrolujący współbieżność czynności w modelu.

Każda czynność aktualnie istniejąca w modelu ma własny zbiór informacji niezbędnych do sterowania jej wykonaniem zapisany w pamięci maszyny na czterech słowach. Zbiór ten będziemy nazywać kostką sterującą KS. Kostka sterująca zawiera licznik rozkazów LR, adres początku drzewa miejsc roboczych

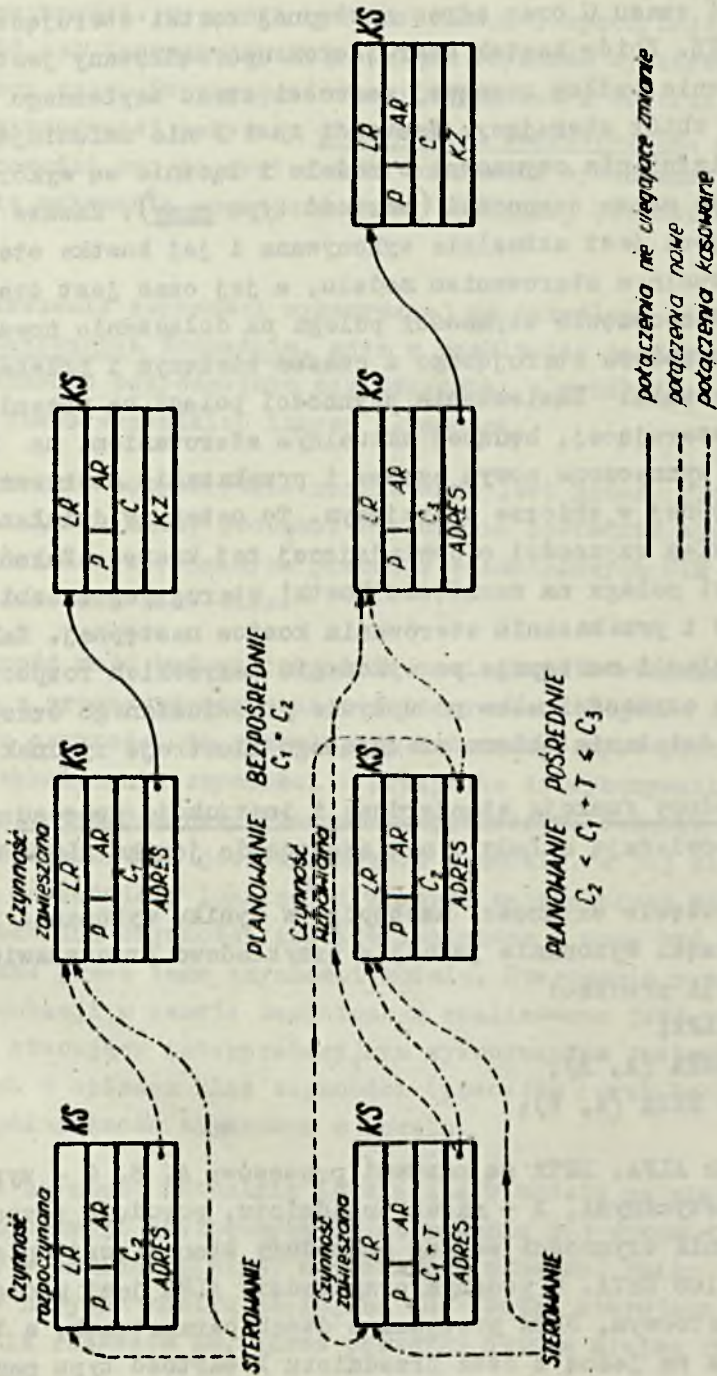
przedmiotów lokalnych AR, symbol klasy czynności P, rejestr wartości czasu C oraz adres następnej kostki sterującej lub symbol KZ. Zbiór kostek sterujących uporządkowany jest chronologicznie według rosnącej wartości czasu zapisanego w C, tworząc zbiór sterujący. Wartości AR i P nie zmieniają się w czasie istnienia czynności w modelu i łącznie są wykorzystywane jako nazwa czynności (wartość typu name). Zawsze jedna z czynności jest aktualnie wykonywana i jej kostka sterująca jest aktualnym sterowaniem modelu, a jej czas jest czasem bieżącym. Rozpoczęcie czynności polega na dołączeniu nowego elementu do zbioru sterującego z czasem bieżącym i przekazanie mu sterowania. Zawieszenie czynności polega na przeniesieniu kostki sterującej, będącej aktualnym sterowaniem, na dalsze miejsce wyznaczone nowym czasem i przekazanie sterowania kostce następnej w zbiorze sterującym. To ostatnie działanie jest wznowieniem czynności odpowiadającej tej kostce. Zakończenie czynności polega na usunięciu kostki sterującej ze zbioru sterującego i przekazaniu sterowania kostce następnej. Zakończenie symulacji następuje po wykonaniu wszystkich rozpoczętych w modelu czynności albo po upływie przewidzianego czasu. Mechanizm działania zbioru sterującego ilustruje rysunek 3.

Procedury funkcje standardowe i instrukcje procesu języka SYM 69 posiadają składnię nawiązującą do języka ALGOL 60.

Rozpoczęcie czynności następuje w wyniku wykonania instrukcji procesu. Wykonanie jednej z przykładowo przedstawionych instrukcji procesu:

```
ALFA;
  BETA (A, B);
X[C] := BETA (A, B);
```

w których ALFA, BETA są nazwami procesów; A, B, C - wyrażeniami arytmetycznymi, X - nazwą przedmiotu, powoduje rozpoczęcie wykonywania czynności według procedury stanowiącej opis procesu ALFA lub BETA. W podanym przykładzie ALFA jest procesem bezparametrowym, BETA procesem o dwóch parametrach, a X [C] podstawia na jedną z cech przedmiotu X wartość typu name określającą nazwę rozpoczynanej czynności klasy BETA.



Rys. 3. Ilustracja działania zbioru sterującego

Rozpoczęcie nowej czynności, jak już o tym mówiono wyżej, wiąże się z zawieszeniem czynności, w której nastąpiło wykonanie instrukcji procesu. Po zawieszeniu lub zakończeniu czynności zainicjowanej omawianą instrukcją procesu następuje powrót do wykonywania czynności zawieszonych, poczynając od instrukcji następnej za instrukcją procesu, jeżeli w inicjowanej czynności nie nastąpiło również zawieszenie lub wykonanie instrukcji procesu.

Zawieszenie czynności następuje w wyniku wykonania jednej z instrukcji:

```
hold (T);
wait (T);
```

w których T jest wyrażeniem arytmetycznym typu real określającym w pierwszym przypadku wartość czasu, do którego czynność należy zawiesić, a w drugim przypadku wartość czasu, na jaki czynność należy zawiesić. Wartość czasu liczona jest w jednostkach całkowitych, a ewentualne zaokrąglenia wykonywane są zgodnie z obowiązującymi w ALGOLu zasadami konwersji wartości z typu real na integer.

Zawieszenie czynności wiąże się z powrotem do wykonywania innej czynności wcześniej zawieszonych. W określonym czasie lub po upływie czasu określonego wartością wyrażenia T następuje powrót do wykonywania czynności zawieszonych, poczynając od instrukcji następnej za instrukcją zawieszenia.

Poza wznowieniem czynności wynikającym z automatyzmu działania zbioru sterującego możliwe jest wznowienie wskazanej nazwą (wartością typu name) czynności.

Wznawianie wskazanej czynności następuje w wyniku wykonania instrukcji:

```
act1 (X);
```

w której X jest nazwą czynności określoną wartością typu name. Wznawienie czynności polega na powrocie do jej wykonywania poczynając od miejsca, w którym wykonanie to zostało za-

wieszzone ale bez zmiany czasu bieżącego. Podobnie jak w poprzednich przypadkach wznowienie czynności wskazanej powoduje zawieszenie czynności ostatnio wykonywanej.

Uzyskanie wartości czasu bieżącego oraz nazwy czynności bieżącej umożliwiają dwie funkcje bezargumentowe:

time

curr

z których pierwsza podaje wartość czasu bieżącego, a druga nazwę bieżącej czynności. Oczywiście wartość pierwszej z nich jest typu real, a drugiej typu name.

Rozbudowa, redukcja i reorganizacja zbioru jest realizowana za pomocą funkcji:

trans (X, Y)

gdzie X jest nazwą przedmiotu, którego ciąg atrybutów jest odłączany, a Y nazwą przedmiotu, do którego ciąg atrybutów jest przyłączany. Wobec względności pojęcia atrybutu w ten sposób realizowany jest transfer elementów (przedmiotów, atrybutów) z jednego zbioru do drugiego. Gdy X i Y należą do tego samego zbioru, transfer ciągu atrybutów reorganizuje ten zbiór. Funkcja może być zapisana również w postaci jednoargumentowej:

trans (X)

gdzie X jest nazwą przedmiotu, którego ciąg atrybutów jest odłączany. W ten sposób realizowana jest redukcja zbioru. Funkcja opisująca transfer cech ma wartość logiczną true, gdy ciąg atrybutów odłączanych zawiera przynajmniej jeden element, oraz wartość logiczną false, gdy ciąg atrybutów odłączanych nie zawiera żadnego elementu. Wykonanie funkcji transferu atrybutów do zbioru nie zawierającego wskazanych miejsc dołączenia powoduje rozbudowę tego zbioru do miejsca dołączenia. Liczność ciągu określa funkcja:

card (X)

gdzie X jest nazwą przedmiotu, którego ciąg atrybutów podlega zliczeniu. Wartością funkcji jest wartość typu real określająca liczbę atrybutów przedmiotu X.

Zajmowanie i zwalnianie przedmiotu jest realizowane przez funkcje:

$$\text{free } (X) \quad \text{occup } (X)$$

gdzie X jest nazwą przedmiotu zajmowanego lub zwalnianego. Zajęty może być tylko przedmiot wolny. Zwolnienie przedmiotu może nastąpić tylko w tej czynności, w której nastąpiło jego zajęcie. Wykonanie funkcji occup (X) na zbiorze nie zawierającym wskazanego przedmiotu powoduje rozbudowę zbioru przez dołączenie wskazanego przedmiotu, a następnie jego zajęcie. Wymienione funkcje mają wartość logiczną true, gdy operacja jest wykonalna, to znaczy może nastąpić zajęcie lub zwolnienie przedmiotu oraz wartość logiczną false, gdy operacja jest niewykonalna.

Zrealizowany praktycznie translator języka SYM 69 dla EMC ODRA 1204 nie zawiera pełnego aparatu języka ALGOL 60, a tylko takie jego fragmenty, które okazały się niezbędne w budowie programów symulacyjnych.

W rezultacie zrealizowano translator dla podzbioru języka SYM 69, w którym:

- jedynym opisem miejsce roboczych jest postać określona opisem set,
- nie wprowadzono opisu procedure z tym, że w pewnym zakresie struktura procedure może być zastąpiona strukturą process,
- nie wprowadzono opisu swith,
- nie wprowadzono warunkowych wyrażeń arytmetycznych, logicznych i mianujących.

Jednakże wymienione wyżej ograniczenia nie są sprzeczne z założeniem generalnym traktowania języka SYM 69 jako swoistej "nabudowy" na język ALGOL 60.

Analiza programu źródłowego odbywa się w 5-ciu pasmach, w których jednolity blok informacji o zmniejszającej się z pasma na pasmo objętości przekształcany jest w program wynikowy. Program wynikowy wykonywany jest interpretacyjnie.

Całość programu wynikowego razem z miejscami roboczymi, częścią interpretującą translatora oraz systemem operacyjnym mieści się w pamięci operacyjnej EMC ODRA 1204 (16k słów 24-bitowych) i tylko ta pamięć jest wykorzystywana w czasie wykonywania programu. Program wynikowy i miejsca robocze mogą zajmować nie więcej niż 12 000 słów, co przeliczone na miejsca robocze zorganizowane w postaci zbiorów wynosi 3000 miejsc.

Poniżej przedstawiony został przykład programu w języku SYM 69 odpowiadającego treścią programowi zamieszczonemu i dokładnie opisanemu w rozdziale 12 publikacji: DAHL O.J., NYGAARD K.: "SIMULA - a Language for Programming and Description of Discrete Event Systems", Borwegian Computing Center, May 1968.

Program wykorzystuje standardowe "zerowanie" miejsc roboczych. Funkcja `rand (a, b)` generuje wartości pseudolosowe typu real o rozkładzie równomiernym w przedziale $\langle a, b \rangle$. Wydruki `printm (...)` realizowane są na konsoli w sposób uzależniony od klucza na pulpicie operatora maszyny, a wydruki `print (...)` na perforatorze lub drukarce wierszowej. Pozostałe symbole użyte w programie nie wymagają komentarzy.

Przykład

```

begin comment   ZARAZA;
set POPULACJA, ZAR, SMIERTELN, ZARAZLIWOSC, MARTWI, OZDROW,
    i,j,k;
process DUR (NUMER);
    begin set DZIEN, PACJENT;
        PACJENT [0] := NUMER;
        for DZIEN := DZIEN + 1 while DZIEN le 8 do
            if SMIERTELN [DZIEN, rand (0,19)] eq 1 then
                begin
                    printm (0, time, '   PACJENT', NUMER, '  ZMARŁ DNIA', DZIEN);
                    PACJENT [0,0] := DZIEN;
                trans (PACJENT, MARTWI);
            end
    end

```

```

    goto END
  end else wait(1);
  printm(0,time,'    PACJENT',NUMER,'    WYZDROWIAŁ');
  trans(PACJENT,OZDROW);
END: end CZYNNOSCI CHOROWANIA NA DUR PACJENTA NUMER;

```

PROGRAM:

```

  read(POPULACJA);
  for i:=i+1 while i le 8 do
    begin j:=-1;
      for j:=j+1 while j le 9 do
        read(SMIERTELN[i,j])
      end CZYTANIA TABLICY OKRESLAJACEJ
        PRAWDOPODOBIENSTWO ZGONU W FUNKCJI
        DNI CHOROBY;

      j:=-1;
      for j:=j+1 while j le 9 do read (ZARAZLIWOSC[j]);
      wait(1);

      printm(40000,time,'  POCZATEK SYMULACJI EPIDEMI');
      for ZAR:=ZAR+1 while ZAR le POPULACJA do
        begin
          DUR(ZAR);
          wait(ZARAZLIWOSC [rand(0,9)] );
        end EPIDEMI;

      wait(9);
      printm(0,time,'    KONIEC SYMULACJI EPIDEMI');
      line(1);
      print(40000, 'POPULACJA',POPULACJA);
      line (1);
      print(0, 'ZMARŁO',card(MARTWI), '    WYZDROWIAŁO',card(OZDROW));

      i:=0;
      for i:=i+1 while i le 8 do
        begin j:=-1; k:=0;
          for j:=j+1 while MARTWI[j,0] ne 0 do
            if MARTWI[j,0] eq i then k:=k+1;

```

```

line (1);
print(0, 'DNIA CHOROBY', 1, ' ZMARŁO PACJENTOW', k)
end;
end SYMULACJI

```

DANE

POPULACJA 100,

PRAWDOPODOBIEŃSTWO ZGONU DLA KOLEJNYCH DNI

1,0,0,0,0,0,0,0,0,0,

1,1,0,0,0,0,0,0,0,0,

1,1,0,0,0,0,0,0,0,0,

1,1,0,0,0,0,0,0,0,0,

1,1,1,0,0,0,0,0,0,0,

1,1,1,1,0,0,0,0,0,0,

1,1,1,0,0,0,0,0,0,0,

1,0,0,0,0,0,0,0,0,0,

ROZKŁAD CZASU W DNIACH POTRZEBNEGO NA ZARAZENIE
 NASTĘPNEGO CZŁONKA POPULACJI

2,3,3,4,4,4,4,5,5,5, |

Literatura

- [1] LESZCZYŃSKI J.: Opracowanie języka i translatora dla symulowania zagadnień transportowych, Prace COBIRTK, 40, Warszawa 1970.
- [2] SYM 69 dla maszyny cyfrowej ODRA 1204, Zakład ETO, COBIRTK, 1969.
- [3] MATYS R.: Studium nad symulacją pracy sieci i węzłów kolejowych, Prace COBIRTK, temat 3001/70, Warszawa 1971.

ЯЗЫК SYM 69 ДЛЯ МОДЕЛИРОВАНИЯ ДИСКРЕТНЫХ ПРОЦЕССОВ

Резюме

Язык SYM 69 служит для описания задач моделирования поведения во времени /имитирующее моделирование/ сложных объектов /систем/ представленных в виде дискретных моделей /моделей схем событий/. Диапазон применения языка является типовым для языков такого класса как SIMULA, SOL, SIMON.

Язык SYM 69 основан на языке ALGOL 60 с тем, что к нему доставлены понятия позволяющие заключить в описании многие события происходящие параллельно, а также динамические оперирование сложными структурами рабочих мест.

Транслятор языка SYM 69 построен для ВЦМ ODRA 1204 и не содержит полного аппарата языка ALGOL 60, лишь только те его фрагменты которые оказались необходимые для программирования задач моделирования. Анализ исходной программы проводится в 5 этапах, в которых однородный блок информации емкостно уменьшающейся с этапа на этап преобразовывался в программу. Программа выполняется интерпретационным образом. Транслятор находится в эксплуатации с 1969 года.

SYM 69 LANGUAGE FOR DISCRETE PROCESSES SIMULATION

Summary

The language SYM 69 serves to describe simulation problems of complex object modelling, presented in the form of discrete models (event oriented models). The scope of the language application is typical for languages like SIMULA, SOL, SIMON.

The SYM 69 language is based on ALGOL 60, to which notions are adjoined permitting to describe many phenomena occurring parallelly, and operate dynamically complex structures of work places.

The SYM 69 translator is constructed for ODRA 1204 computer. It does not comprise the full ALGOL 60 language, but only fragments which appeared necessary for simulation statement programming. The analysis of the source program is made in 5 passes where the uniform information block of capacity diminishing from one pass to the next, is transformed into an object program. The latter is performed by means of an interpretation. The translator has been exploited since 1969.

WYDAWNICTWA IMM

Branżowy Ośrodek Informacji Naukowo-Technicznej i Ekonomicznej Instytutu Maszyn Matematycznych wydaje:

ALGORYTMY - półrocznik, zawiera artykuły na temat teorii programowania i zastosowania elektronicznych maszyn cyfrowych. Do nabycia w księgarni ORWN PAN oraz w Domach Książki. Cena zeszytu 40,- zł.

PRACE IMM - 3 numery w roku, zawierają publikacje naukowe i badawcze pracowników IMM w zakresie projektowania i budowy elektronicznych maszyn cyfrowych oraz systemów przetwarzania informacji. Do nabycia w księgarni ORWN PAN oraz w Domach Książki. Cena zeszytu 60,- zł.

Elektroniczna Technika Obliczeniowa - NOWOŚCI - kwartalnik, zawiera artykuły przeglądowe z dziedziny maszyn matematycznych, opracowane na podstawie najnowszej literatury światowej. Prenumeratę prowadzi Centrala Kolportażu Prasy i Wydawnictw "RUCH". Cena prenumeraty krajowej 240,- zł rocznie.

Automatyzacja Przetwarzania Informacji - INFORMACJA EKSPRESOWA - miesięcznik. Prenumeratę prowadzi Centrala Kolportażu Prasy i Wydawnictw "RUCH". Cena prenumeraty krajowej 240,- zł rocznie.

WYKAZ PRAC

Prace wykonane w okresie od 1 stycznia do 31 grudnia 1972 r. w ramach planu pracy na rok 1972. Wykaz ten zawiera opis prac, ich zakres, wykonawców, datę wykonania oraz kosztorys. Wykaz ten jest załącznikiem do sprawozdania z wykonania planu pracy na rok 1972.

Prace wykonane w okresie od 1 stycznia do 31 grudnia 1972 r. w ramach planu pracy na rok 1972. Wykaz ten zawiera opis prac, ich zakres, wykonawców, datę wykonania oraz kosztorys. Wykaz ten jest załącznikiem do sprawozdania z wykonania planu pracy na rok 1972.

Prace wykonane w okresie od 1 stycznia do 31 grudnia 1972 r. w ramach planu pracy na rok 1972. Wykaz ten zawiera opis prac, ich zakres, wykonawców, datę wykonania oraz kosztorys. Wykaz ten jest załącznikiem do sprawozdania z wykonania planu pracy na rok 1972.

Prace wykonane w okresie od 1 stycznia do 31 grudnia 1972 r. w ramach planu pracy na rok 1972. Wykaz ten zawiera opis prac, ich zakres, wykonawców, datę wykonania oraz kosztorys. Wykaz ten jest załącznikiem do sprawozdania z wykonania planu pracy na rok 1972.

Prace wykonane w okresie od 1 stycznia do 31 grudnia 1972 r. w ramach planu pracy na rok 1972. Wykaz ten zawiera opis prac, ich zakres, wykonawców, datę wykonania oraz kosztorys. Wykaz ten jest załącznikiem do sprawozdania z wykonania planu pracy na rok 1972.

Cena zł 40,-