Поперечники Колмогорова, многообразия Грассмана и развертка временных рядов

В. М. Бухштабер

МИАН им. В.А. Стеклова, МГУ им. М.В.Ломоносова

Конференция Мехмат–24
“Актуальные проблемы математики и механики”,
МГУ, 02–04 декабря 2024 г.

Понятие n-поперечника было введено А.Н. Колмогоровым в 1936 году.

В алгебраической топологии одним из самых известных гладких многообразий является многообразие Грассмана $G(N, n)$ всех n-мерных подпространств $L \subset \mathbb{R}^N$.

Доклад посвящён теории и приложениям многомерных разверток временных рядов. В центре внимания будут методы, основанные на теории операторов сдвига и функциональных уравнениях.

Будут представлены алгоритмы решения экстремальных задач на многообразиях Грассмана.

Мы обсудим связи с теорией поперечников в задачах об аппроксимации функций и восстановлении параметров семейств функций.

Работы А.Н. Колмогорова

Доклад связан с работами Андрея Николаевича Колмогорова:

№ 28. “О наилучшем приближении функций, заданного функционального класса”, 1936 г.
№ 42. “Кривые в гильбертовом пространстве, инвариантные по отношению к однопараметрической группе движений”, 1940 г.
(см. А.Н. Колмогоров, Избранные труды, Математика и механика, Составитель В.М. Тихомиров, М.: Наука 1985).

Понятие n-поперечника $D_n(F)$ класса функций F относительно n-мерных подпространств пространства аппроксимирующих функций введено в работе № 28. Таким образом, классу F была сопоставлена убывающая последовательность $\{D_n(F),\ n = 1, 2, \ldots\}$ неотрицательных чисел и поставлена задача:

Однозначно ли определяется n-мерное подпространство функций, доставляющее значение $D_n(F)$?

Работы В.М. Тихомирова

Теория и приложения n-поперечников находятся в центре внимания научной деятельности Владимира Михайловича Тихомирова. Его результаты в этом направление получили мировое признание.

В Math–Net.Ru список его основных публикаций начинается с работы:

В. М. Тихомиров, “Поперечники множеств в функциональных пространствах и теория наилучших приближений”, УМН, 15:3(93) (1960), 81–120.

В работе В.М. Тихомиров, “Три этапа развития теории приближений”, Труды МИАН, 319, 2022, 7–19, дано изложение и анализ достижений по теории и приложениям n-поперечников до 2022 года.

n-поперечники Колмогорова

Пусть выбрано некоторое пространство функций $\mathfrak{F}$ с фиксированным расстоянием ρ. Пусть заданы множество $F \subset \mathfrak{F}$ и линейное пространство аппроксимирующих функций $\Phi \subset \mathfrak{F}$.
Рассмотрим задачу о приближении функции $f \in F$ линейными формами вида $\varphi = c_1\varphi_1 + \ldots + c_n\varphi_n$, где $\varphi_1, \ldots, \varphi_n$ – данный набор функций из Φ. Обозначим через $E_n(f)$ расстояние ρ от f до n-мерного линейного пространства, порождённого функциями φ.
А.Н. Колмогоров назвал n-поперечником множества F нижнюю грань $D_n(F)$ величин $E_n(f)$, где $f \in F$.
В обзоре В.М. Тихомирова (1960) отмечается, что значения отклонений множества *периодических* функций F от пространств Φ *тригонометрических и обыкновенных полиномов* были подсчитаны в работах: Ж. Фавара; Н.И. Ахиезера и М.Г. Крейна; Н.И. Ахиезера; М.Г. Крейна; Б. Надя и К.И. Бабенко.

n-поперечники Колмогорова и многообразия Грассмана

Понятие n-поперечника связало теорию приближений функций с рядом областей теоретической и прикладной математики, в которых решаются экстремальные задачи на многообразиях Грассмана $G(N, n)$.

Пусть в задаче Колмогорова выбрано некоторое пространство $\mathfrak{F}$ непрерывных функций на отрезке $[0, 1]$. Выберем шаг дискретизации h и сопоставим каждой функции $f \in \mathfrak{F}$ временной ряд $f = (f_1, \dots, f_N)$, где $f_s = f((s-1)h)$, $s = 1, \dots, N$. Таким образом, мы приходим к задаче:

Пусть задано компактное множество $K \subset \mathbb{R}^N$. Стандартным способом определяется средняя точка x_K множества K, которая выбирается в качестве $0 \in \mathbb{R}^N$. Для любого $n < N$ определено расстояние $E_K(L)$ от множества K до данного n-мерного подпространства $L \subset \mathbb{R}^N$ и, следовательно, определена функция $E_K(L)$ на многообразии $G(N, n)$.

Вводится n-поперечник множества $K \subset \mathbb{R}^N$ как минимальное значение функции $E_K(L)$.

Многообразия Грассмана

Определение 1

Многообразие Грассмана $G_{\mathbb{R}^n}(n,q)$ (далее $G(n,q)$) – это множество всех q-мерных линейных подпространств пространства $\mathbb{R}^n$.

Каждое q-мерное линейное подпространство L в $\mathbb{R}^n$ натянуто на q ортонормированных n мерных векторов, которые определены с точностью до ортогонального преобразования пространства L.
Пусть $\mathrm{Sym}(n,n)$ — линейное пространство всех симметрических $(n \times n)$-матриц M. На $\mathrm{Sym}(n,n)$ действует ортогональная группа $O(n)$ сопряжением по формуле AMA', где A' — транспонированная матрица.
Определено $O(n)$ — эквивариантное вложение $G(n,q)$ в $\mathrm{Sym}(n,n)$, при котором подпространству L в $\mathbb{R}^n$ сопоставляется матрица $\Pi\Pi'$, где Π $(n \times n)$-матрица, образованная набором из q ортонормированных n-мерных векторов.
Таким образом, многообразие $G(n,q)$ реализуется как гладкое подмногообразие в линейном пространстве размерности $(n+1)n/2$.

Пусть дан временной ряд $f = (f_1, \ldots, f_N)$.

Определение 2

Кусочно линейная кривая $\mathfrak{X}_f$ в n-мерном евклидовом пространстве $\mathbb{R}^n$, полученная последовательным соединением векторов $X_1, \ldots, X_p$ ($p = N - n + 1$), где $X_q' = (f_q, \ldots, f_{q+n-1})$, называется *$n$-мерной разверткой* $\mathfrak{X}_f = \mathfrak{X}_f(N; n)$ временного ряда $f = (f_1, \ldots, f_N)$.
Векторы $X_1, \ldots, X_p$ ($p = N - n + 1$) в этом случае называются узлами кривой $\mathfrak{X}_f$.

Пространство разверток временных рядов

Далее мы рассматриваем ортогональные проекции пространства $\mathbb{R}^n$ в его q-мерные подпространства L.

Пусть $X_f \subset \mathbb{R}^n$ – развёртка временного ряда f.

1. Задача: найти подпространство $L \subset \mathbb{R}^n$, для которого проекция кривой X_f будет наиболее выразительна.
2. Задача: найти подпространства L_1 и $L_2 \subset \mathbb{R}^n$, такие, что проекции $X_{f,1}$ и $X_{f,2}$ кривой X_f на эти подпространства будут показывать разные выразительные свойства кривой X_f.

Пространство $M(n,p)$ всех кусочно-линейных кривых с p-узлами в $\mathbb{R}^n$ отождествим с $\mathbb{R}^{np}$.
Множество T^N всех временных рядов $f_1, \dots, f_N$ отождествим с $\mathbb{R}^N$.
Обозначим через $T_u^N \subset M(n,p)$ подпространство разверток временных рядов длины N, $\dim T_u^N = N$.

Восстановление ряда по проекциям его развёртки

Рассмотрим развёртку X_f в $\mathbb{R}^n$ временного ряда f с N отсчетами, выберем подпространство L в $\mathbb{R}^n$ и обозначим через $\pi(X_f)$ проекцию кривой X_f в L.

3. Задача: Найти проекцию $\pi_T(X_f)$ кривой $\pi(X_f) \subset M(n,p)$ на линейное пространство T_u^N.

Определение 3

$\pi(X_f)$-аппроксимацией временного ряда f называется временной ряд, соответствующий развёртке $\pi_T(X_f)$.

4. Задача: Найти подпространство L для которого $\pi(X_f)$-аппроксимация временного ряда f является наилучшей, в равномерном или статистическом смысле.

Экстремальные задачи на $G(n, q)$

В задачах 1-4 используются следующие понятия:
"наиболее выразительная проекция", "минимум потери информации", "наиболее представительная проекция", "наилучшая аппроксимация временного ряда".

Все эти понятия формализуются в терминах критериев.

Критерии задают функции на $G(n, q)$, а решения задач 1-4 сводятся к алгоритмам нахождения экстремальных значений этих функций.

Градиентная оптимизация на $G(n, q)$

Пусть $1 \leq k \leq q, \quad q+1 \leq l \leq n, \quad 1 \leq i, j \leq n.$

Рассмотрим следующий набор ортогональных $(n \times n)$-матриц $A(t; k, l) = (a_{i,j}(t; k, l))$ с матричными элементами $a_{i,i}(t; k, l) = 1$, если $i \neq k$ или $i \neq l$, $\quad a_{k,k}(t; k, l) = a_{l,l}(t; k, l) = \cos t$, $a_{k,l}(t; k, l) = -a_{l,k}(t; k, l) = \sin t$, при $l \neq k$, и $a_{i,j}(t; k, l) = 0$ в остальных случаях.

Пример: $n = 5, q = 2, t = [0, 2\pi], k = \{1, 2\}, l = \{3, 4, 5\}$.

$$A(t; 1, 3) = \begin{bmatrix} \cos t & 0 & \sin t & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ -\sin t & 0 & \cos t & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{1}$$

Градиентная оптимизация на $G(n, q)$

Отметим, что $\dim(G(n, q)) = q(n - q)$.
Для каждой функции Ψ на многообразии $G(n, q)$ и каждой точки $L \in G(n, q)$ мы получаем набор из $q(n - q)$ функций $\Psi_L(t; k, l) = \Psi(A(t; k, l)L)$, $0 \leq t \leq 2\pi$, и $\Psi_L(0; k, l) = \Psi(L)$.

Теорема 1

Набора функций $\Psi_L(t; k, l)$ необходимо и достаточно для организации градиентной оптимизация любой гладкой функции $\Psi(L)$ на $G(n, q)$.

Метод главных компонент

Пусть $X = \{X_1, \dots, X_p\}$, где $X_k \in \mathbb{R}^n$ и $\bar{X} = \frac{1}{p}(X_1 + \dots + X_p)$, $\widetilde{X_k} = X_k - \bar{X}$.

Определение 4

Матрицей рассеивания $W = (w_{ij})$ совокупности векторов X называется матрица Грамма совокупности $\widetilde{X} = \{\widetilde{X_k}\}$, где $w_{ij} =< \widetilde{X}_i, \widetilde{X}_j >$ – стандартное скалярное произведение.

Матрица W симметрическая и неотрицательно определённая. Пусть $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n \geq 0$ – набор собственных чисел матрицы W и $v_1, ..., v_n$ – ортонормированный набор собственных векторов этой матрицы, где v_k – собственный вектор, соответствующий λ_k, который называется k-ой главной компонентой совокупности X.

Вектора v_k – решение экстремальных задач на $G(n,q)$

Пусть $X = \{X_1, \ldots, X_p\}$, где $X_k \in \mathbb{R}^n$, как и выше.
Введем функцию $\Psi_X(L) = ||L(\widetilde{X_1})||^2 + \ldots + ||L(\widetilde{X_p})||^2$ на многообразии $G(n,q)$.

Теорема 2

Пусть $A_* \in G(n,q)$ – оператор ортогональной проекции на подпространство L_* в $\mathbb{R}^n$, натянутое на первые q главные компоненты $v_1, \ldots, v_q$. Тогда

$$\Psi_X(L_*) = \max\{\Psi_X(L), L \in G(n,q)\}. \tag{2}$$

Доказательство использует явное описание градиента функции $\Psi_X(L)$ на многообразии Грассмана.

Метод многомерной развертки

Определение 5

Кусочно-линейная кривая $K \in M(n,p)$ имеет ранг r, если она лежит в r-мерном аффинном подпространстве L пространства R^n.

Определение 6

Кусочно-линейная кривая $K \in M(n,p)$ с узлами $\{X_1, ..., X_p\}$ принадлежит ε-окрестности r-мерного аффинного подпространства $L \subset \mathbb{R}^n$, если $\Sigma_{i=1}^{p} ||L - X_i||^2 < \varepsilon$, где $||L - X_i||$ – расстояние от вектора X_i до подпространства L.

Определение 7

Кусочно-линейная кривая $K \in M(n,p)$ имеет ε-ранг r, если она лежит в ε-окрестности r-мерного аффинного подпространства L.

Пусть $X = \{X_1, ..., X_p\}$ – набор узлов кривой $K \in M(n, p)$.

Определение 8

Матрицей рассеивания кривой $K \in M(n, p)$ называется матрица W рассеивания набора X.

Таким образом, определены собственные значения и собственные вектора кривой K.

Определение 9

Матрица рассеивания развёртки X_f ряда f называется матрицей рассеивания этого ряда.

Метод многомерной развертки

Из теории главных компонент следует:

Лемма

Ранг кривой $K \in M(n,p)$ не превосходит r тогда и только тогда, когда $\lambda_i = 0$ для всех $i > r$, где $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n \geq 0$ – собственные значения матрицы W.

ε-ранг кривой X_f не превосходит r тогда и только тогда, когда

$$\sum_{i=r+1}^{n} \lambda_i < \varepsilon. \tag{3}$$

Кривая $K \in M(n,p)$ имеет ранг r тогда и только тогда, когда $K = \bar{X} + V(r)\Big(V(r)^T K\Big)$, где $V(r)$ – $(r \times n)$-матрица, столбцы которой $v_1, ..., v_r$ – собственные вектора матрицы рассеивания кривой K.

Метод многомерной развертки

Анализ геометрии развёртки X_f позволяет получать информацию о свойствах временного ряда.

Значение ранга развёрки даёт оценку количества и характера компонент временного ряда.

Одной из важных задач анализа временного ряда является построение его модели в виде непрерывной функции.

Определение 10

Ранг непрерывной функции $f(t)$ не превосходит r, если для любых $h, n, N, n < N$ ранг развёртки соответствующего временного ряда не превосходит r.

Метод многомерной развертки

Метод многомерной развертки опирается на теорию дифференциальных и функциональных уравнений.

Пусть $g(t)$ – вещественнозначная, достаточное число раз дифференцируемая функция.

Теорема 3

Ранг функции $g(t)$ не превосходит r тогда и только тогда, когда $g(t)$ является решением обыкновенного дифференциального уравнения с постоянными коэффициентами $a_0, \ldots, a_r, C$:

$$\sum_{i=0}^{r} a_i \frac{d^i}{dt^i} g(t) = C, a_r \neq 0. \tag{4}$$

Общее решение уравнения (4) имеет вид:

$$g(t) = \sum a_k(t) e^{\lambda_k t} sin(\omega_k t + \varphi_k), \tag{5}$$

где $a_k(t)$ – полиномы

Метод многомерной развертки

Пусть $g(t)$ – вещественнозначная, достаточное число раз дифференцируемая функция.

Теорема 4

Ранг $g(t)$ не превосходит r тогда и только тогда, когда $g(t)$ является решением функционального уравнения при некоторых $\varphi_k(t)$ и $\psi_k(t)$, $k \in \{1, ..., r\}$, таких что:

$$g(t+\tau) = \sum_{k=1}^{r} \varphi_k(t)\psi_k(\tau) \tag{6}$$

Примеры:

$r = 1 : e^{\lambda t}$;

$r = 2 : \alpha \cos \omega t + \beta \sin \omega t$;

$r \geq 2 : t^{r-1}$.

Рассмотрим пространство $M(n, p) \approx \mathbb{R}^{np}$ пространство всех кусочно-линейных кривых с p узлами в $\mathbb{R}^n$.

Введём в $M(n, p)$ расстояние:

$$||Y_1 - Y_2||^2 = \frac{1}{p} \sum_{i=1}^{p} ||X_{1i} - X_{2i}||^2, \tag{7}$$

где Y_1, Y_2 – кривые из $M(n, p)$, $Y_k = (X_{k1}, ..., X_{kn}), k = \{1, 2\}$, $||.||$ – евклидова норма.

Обозначим через $M^r(n, p)$ множество всех кусочно-линейных кривых из $M(n, p)$ ранга, не превышающего r.

Алгоритмы восстановления временного ряда

Определение 11

Кривая $Y_*(r) \in M^r(n,p)$ является проекцией кривой $Y \in M(n,p)$ в $M^r(n,p)$, если:

$$||Y - Y_*(r)||^2 = \min_{\hat{Y} \in M^r(n,p)} ||Y - \hat{Y}||^2. \tag{8}$$

Определение 12

Временной ряд $\hat{f} = (\hat{f}_1, ..., \hat{f}_N)$ называется проекцией кривой $Y \in M^r(n,p)$ на пространство временных рядов T^N, если:

$$||Y - X_{\hat{f}}||^2 = \min_{f \in T^N} ||Y - X_f||^2, \tag{9}$$

где $X_{\hat{f}}, X_f$ – развёртки временных рядов $\hat{f}$ и f.

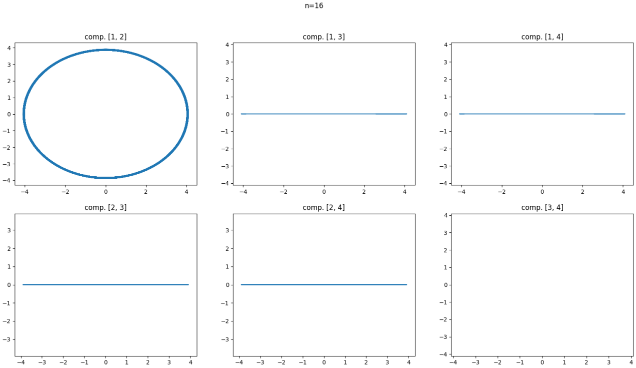
Алгоритмы восстановления временного ряда

Теорема 5

Для кривой $Y \in M(n,p)$, $p = N - n + 1$, *её проекция* $\hat{f}(Y)$ *на пространство временных рядов* T^N *имеет вид* $\hat{f}(Y) = (\hat{f}_1, \ldots, \hat{f}_N)$, *где :*

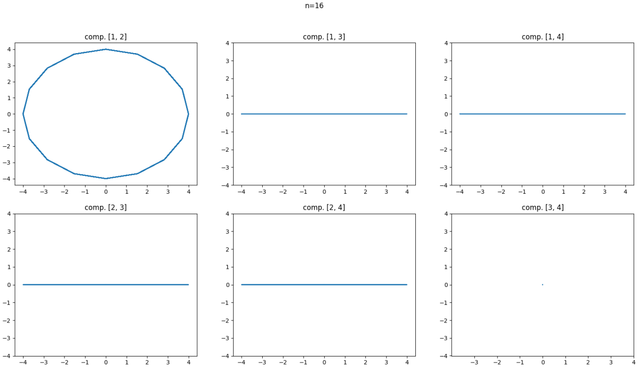
$$\hat{f}_k = \begin{cases} \frac{1}{k} \sum\limits_{i=1}^{k} y_{i,k-i+1}, & 1 \le k \le n, \\ \frac{1}{n} \sum\limits_{i=1}^{n} y_{i,k-i+1}, & n \le k \le p, \\ \frac{1}{N-n+1} \sum\limits_{i=1}^{N-k+1} y_{i+kp,p-i+1}, & p \le k \le N, \end{cases} \tag{10}$$

Анализ моделььных рядов



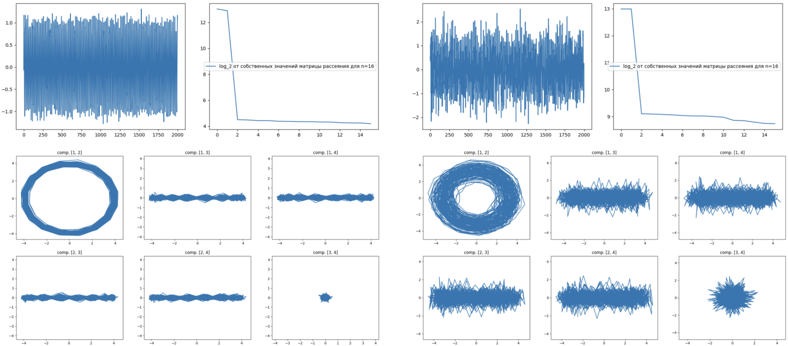
$$sin(\tfrac{3}{8}t), \quad T = \tfrac{16}{3}\pi, \quad n = 16, \quad h = 1, \quad t_1 = 0, \quad N = 2000.$$

Анализ модельных рядов



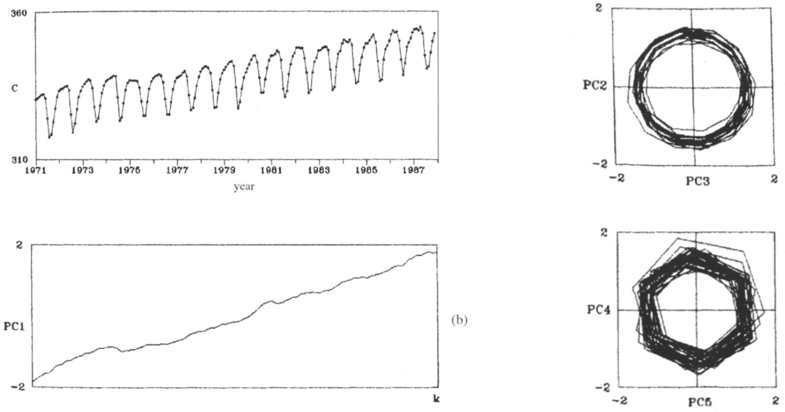
$$sin(\tfrac{\pi}{8}t), \quad T = 16, \quad n = 16, \quad h = 1, \quad t_1 = 0, \quad N = 2000.$$

Анализ модельных рядов



$sin(\frac{3}{8}t)$ + шум, $\quad T = \frac{16}{3}\pi, \quad n = 16, \quad h = 1, \quad t_1 = 0, \quad N = 2000.$
(1) D = 0.01. (2) D = 0.05.

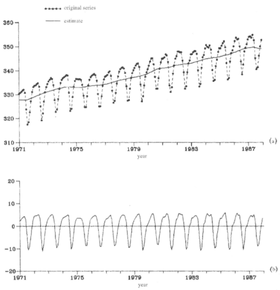
Анализ рядов CO_2, данные мониторинга

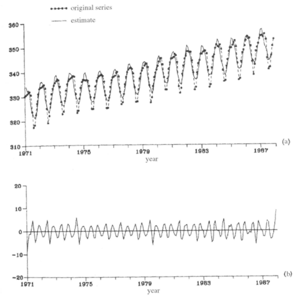


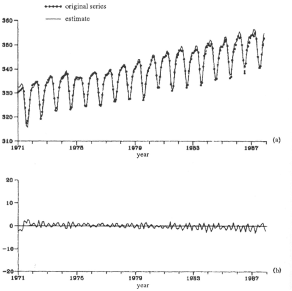
$n = 12, \quad h = 1, \quad t_1 = 0.$

Анализ показал тренд, наличие годового и сезонного циклов.

Анализ рядов CO_2, данные мониторинга







Список литературы

1. В. М. Бухштабер, В. К. Маслов, *Факторный анализ на многообразиях и проблема выделения признаков в распознавании образов.*, Изв. АН СССР, сер. Техн. Киберн., № 6,1975, 194–201.

2. В. М. Бухштабер, В. К. Маслов, *Факторный анализ и экстремальные задачи на многообразиях Грассмана.*, Мат. методы реш. Эконом. Задач, № 7, (Прил. к “Эконом. и матем. методы”), Наука, М., 1977, 85–102.

3. В. М. Бухштабер, С. А. Айвазян, И. С. Енюков, Л. Д. Мешалкин, *Прикладная статистика. Классификация и снижение размерности.*, Финансы и статистика, М., 1989, 608 стр.

4. V. M. Buchstaber, *Time series analysis and Grassmannians.*, in Applied problems of Radon transform, Amer. Math. Sos. Transl., 162:2, AMS, Providence, RI, 1994, 1–17.

5. В. М. Бухштабер, *Многомерные развертки временных рядов. Теоретические основы и алгоритмы.*, Обозрение прикладной и промышл. математики, Серия “Вероятность и статистика”, 4:4, ТВП, М., 1997, 629–645.