САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ
УНИВЕРСИТЕТ

На правах рукописи

Старосельский Юрий Михайлович

ИССЛЕДОВАНИЕ ОПТИМАЛЬНЫХ ПЛАНОВ ЭКСПЕРИМЕНТА ДЛЯ НЕЛИНЕЙНЫХ ПО ПАРАМЕТРАМ РЕГРЕССИОННЫХ МОДЕЛЕЙ

05.13.18 – Математическое моделирование, численные
методы и комплексы программ

АВТОРЕФЕРАТ

диссертации на соискание ученой степени

кандидата физико-математических наук

Санкт-Петербург
2008

Работа выполнена на кафедре статистического моделирования математико - механического факультета Санкт-Петербургского государственного университета

Научный руководитель: доктор физико-математических наук, профессор Мелас Вячеслав Борисович

Официальные оппоненты: доктор физико-математических наук, профессор Седунов Евгений Витальевич (Санкт-Петербургский государственный университет технологии и дизайна)

кандидат физико-математических наук, доцент Буре Владимир Мансурович (Санкт-Петербургский государственный университет)

Ведущая организация: Санкт-Петербургский государственный политехнический университет

Защита состоится "____"____________ 200__ г. в ____ часов на заседании диссертационного совета Д 212.232.51 по защите докторских и кандидатских диссертаций при Санкт-Петербургском государственном университете по адресу: 198504, Санкт-Петербург, Старый Петергоф, Университетский пр., д. 28, математико-механический факультет, ауд. 405.

С диссертацией можно ознакомиться в Научной библиотеке имени М. Горького Санкт-Петербургского государственного университета по адресу: 199034, Санкт-Петербург, Университетская набережная, дом 7/9.

Автореферат разослан "____"____________ 200__ г.

Ученый секретарь диссертационного совета — Мартыненко — Б.К. Мартыненко

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы. Одной из важнейших методологий исследования объектов и систем различной природы является построение и анализ регрессионных моделей на основе экспериментальных данных. При этом организация эксперимента играет существенную роль для повышения точности результатов при ограниченных ресурсах. Оптимальная организация эксперимента изучается теорией планирования регрессионных экспериментов, которая интенсивно развивается с середины 60-х годов 20-го века.

Эта теория была инициирована работами Дж. Элвинга, Г.Чернова, Дж. Кифера и Дж. Вольфовица, а в ее развитие значительный вклад внесли российские ученые В.В.Налимов, Г.К.Круг, В.В.Федоров, С.М. Ермаков, М.Б.Малютов, В.П.Козлов, Е.В.Седунов и многие другие. В рамках этой теории весьма полно изучены линейные по параметрам регрессионные модели. Однако во многих экспериментальных исследованиях существенную роль играют нелинейные по параметрам модели. Трудность исследования этих моделей заключается в зависимости асимптотической ковариационной матрицы оценок от истинных значений параметров. Для преодоления этой трудности используются известные в статистике подходы: локально оптимальный, минимаксный и байесовский. Возникающие в этих подходах задачи построения оптимальных планов являются весьма сложными экстремальными задачами. До недавнего времени аналитическое решение таких задач оказывалось возможным лишь для простейших моделей с одним или двумя параметрами.

В работах В.Б.Меласа [2], [4] был предложен подход, который открывает перспективы значительно более полного исследования таких задач. Этот подход назван функциональным подходом. Его основная идея состоит в исследовании опорных точек оптимальных планов как неявно заданных функций вспомогательных величин. В работах В.Б.Меласа и А.Н.Пепелышева (см. монографию [5]) функциональный подход был развит для построения локально оптимальных и максиминно эффективных планов. Настоящая работа посвящена развитию функционального подхода для байесовских планов и сравнительному исследованию всех основных типов планов (локально оптимальных, максиминно эффективных, байесовских и планов в равноотстоящих точках).

Цель работы. Целью работы является разработка методологии построения и исследования байесовских оптимальных планов для ряда нелинейных по параметрам моделей дробно-рационального и экспоненциального вида, а также сравнение этих планов с альтернативными планами.

Общая методика работы. В работе применяются методы функционального анализа (теорема о неявном отображении), линейной алгебры, теории планирования регрессионных экспериментов (теоремы эквивалентности, рекуррентные формулы для вычисления разложений точек и весов плана), используются математические пакеты символьной обработки данных (Maple, Matlab) для построения графиков и для вычисления коэффициентов разложений неявно заданных функций в ряды Тейлора.

Научная новизна. В данной работе впервые разработана методология построения и исследования байесовских планов эксперимента на основе функционального подхода. Для наиболее распространенных критериев оптимальности впервые дано строгое математическое обоснование применимости подхода для нелинейных моделей с несколькими неизвестными параметрами. Впервые для ряда нелинейных моделей, представляющих интерес в микробиологии и химии, проведено параллельное исследование всех основных типов оптимальных планов. Показано, что байесовские планы являются менее чувствительными к изменению параметров и позволяют значительно уменьшить количество наблюдений, необходимое для достижения заданной точности, по сравнению с планами в равноотстоящих точках.

Теоретическая и практическая ценность. Теоретическая ценность предложенной в диссертации методологии заключается в том, что она предоставляет новый аппарат построения и исследования оптимальных планов эксперимента для нелинейных по параметрам регрессионных моделей. Кроме того, модели, изученные в диссертации, часто используются при решении различных задач микробиологии и аналитической химии. Поэтому планы, построенные в диссертации, имеют практическую ценность.

Апробация работы. Основные результаты диссертации докладывались и обсуждались на семинаре кафедры статистического моделирования математико-механического факультета СПбГУ, а также на XX-ой Международной научной конференции по прикладной математике (ММТТ-20) (Ярославль, 28 мая - 1 июня 2007 г.) и на XVI Международной конференции по междисциплинарным математическим и статистическим методам (Мемфис, США, 16 - 18 мая 2008 г.).

Публикации. По теме диссертации опубликовано три работы, перечисленные в конце автореферата. Статья [7] опубликована в журнале, входящем в перечень ВАК по специальности 05.13.18. В статье [8] автору диссертации принадлежат численные результаты, а в статье [9] – все результаты, за исключением теоремы о том, что опорные точки E-оптимальных планов являются аналитическими функциями.

Структура и объем диссертации. Диссертация состоит из введения, четырех глав и заключения. Библиография содержит 37 наименований. Общий объем работы 109 страниц.

СОДЕРЖАНИЕ РАБОТЫ

Диссертация посвящена построению и исследованию оптимальных планов эксперимента для нелинейных по параметрам регрессионных моделей.

Планом эксперимента называется дискретная вероятностная мера ξ, которая определяется таблицей

$$\xi = \begin{pmatrix} t_1 & \dots & t_n \\ \omega_1 & \dots & \omega_n \end{pmatrix}, \quad t_i \in \chi, \quad i = 1, 2, \dots, n.$$

Носитель плана состоит из точек, принадлежащих множеству планирования χ, а весовые коэффициенты ω_i удовлетворяют условиям $\omega_i \geq 0,\ \sum_{i=1}^{n} \omega_i = 1$.

Задача построения оптимального плана состоит в нахождении экстремума (минимума или максимума) величины

$$\Phi(M(\xi)), \tag{1}$$

где Φ – некоторая заданная функция (критерий оптимальности), $M(\xi)$ – информационная матрица плана ξ. Для нелинейной по параметрам функции регрессии $\eta(x, \theta)$, информационная матрица имеет вид

$$M(\xi, \theta^*) = \left(\sum_{s=1}^{n} f_i(x_s, \theta^*) f_j(x_s, \theta^*) \omega_s \right)_{i,j=1}^{m},$$

где

$$f_i(x, \theta) = \frac{\partial}{\partial \theta_i} \eta(x, \theta),\ i = 1, \dots, m, \quad \theta = (\theta_1, \dots, \theta_m)^T,$$

θ^*– истинное значение вектора параметров (см. [1]).

В общем случае ξ^* – решение задачи (1) – зависит от θ^*, которое априорно не известно. Чтобы преодолеть это противоречие, существует несколько подходов: локально оптимальный, байесовский и максиминный.

План называется локально оптимальным относительно вектора параметров θ', если он является решением задачи $\Phi(M(\xi, \theta')) \to \sup\limits_{\xi}$ (см. [3]).

Однако часто локально оптимальный подход оказывается неудовлетворительным, так как оптимальный план для θ' может быть неудачным для θ^*. Для характеристики качества плана вводится понятие эффективности

$$eff(\xi, \theta) = \left(\frac{\Phi(M(\xi, \theta))}{\Phi(M(\xi^*(\theta), \theta))} \right).$$

Чтобы в результате получить план, который не зависел бы от фиксированных значений параметров и обеспечивал хорошие результаты планирования при любом θ из некоторого заданного множества, исходную задачу заменяют на максиминный или байесовский аналог (см. [6], [5]).

- Максиминный:
$$\max_{\xi} \min_{\theta \in \Omega} \left(\frac{\Phi(M(\xi, \theta))}{\Phi(M(\xi^*(\theta), \theta))} \right) \tag{2}$$
- Байесовский:
$$\min_{\xi} \int \left(\frac{\Phi(M(\xi^*(\theta), \theta))}{\Phi(M(\xi, \theta))} \right) P(d\theta), \tag{3}$$
где Ω – некоторый компакт в $\mathbb{R}^m$, априорно задано распределение $P(d\theta)$, $\xi^*(\theta)$ – локально оптимальный план.

Во **введении** приведен обзор современного состояния предметной области, сформулированы цели исследования и кратко изложены результаты.

В **первой главе** дается краткий обзор основных понятий и методов теории оптимального планирования регрессионных экспериментов, описываются модели, рассмотренные в диссертации, проводится обзор литературы.

В первом параграфе этой главы формулируются базовые определения и описываются основные методы построения оптимальных планов.

Во втором параграфе формулируются теоремы эквивалентности для байесовских и максиминных планов, которые являются основным аппаратом проверки оптимальности планов.

В третьем параграфе описываются модели, которые рассмотрены в диссертации:

- Модель Михаэлиса – Ментена:

$$\eta(x,\Theta)=\frac{\theta_1 x}{\theta_2+x},\ \ x\in[0,d_2],\ \theta_2>0.$$

  На практике такая модель используется, в частности, для описания скорости протекания реакции: θ_1 – максимальная скорость протекания реакции, x – концентрация вещества.

- Модель кинетики второго порядка.

  Более сложная модель, применяемая в экспериментальной химии – модель кинетики второго порядка:

$$\eta(x,\Theta)=\frac{\theta_0 x(\theta_1+x)}{\theta_2+\theta_3 x+x^2},\ x\in[0,d_2],\ \theta_2,\theta_3>0.$$

- Модель в виде алгебраической суммы двух экспонент:

$$\eta(x,\Theta)=\theta_1 e^{-\theta_3 x}+\theta_2 e^{-\theta_4 x},\ x\in[0,\infty),\ \theta_4>\theta_3>0.$$

- Модель экспоненциального роста:

$$\eta(x,\Theta)=\theta_1+\theta_2 e^{-\theta_3 x},\ x\in[0,\infty),\ \theta_3>0.$$

Вторая глава посвящена описанию методологии фукнкционального подхода и его обоснованию для D-критерия.

В первом параграфе даются основные определения и предположения метода.

Во втором параграфе описывается функциональный подход. Основной идеей является изучение вектора опорных точек и весов τ оптимального плана ξ_τ как неявной функции некоторого вспомогательного параметра

z, которая задается уравнением $g(\tau, z) = 0$, где $g(\tau, z)$ – вектор функция состоящая из частных производных $\varphi(\tau, z)'_{\tau_i}$, $\varphi(\tau, z)$ – критерий оптимальности. Это уравнение называется основным уравнением.

Для произвольной (скалярной, векторной или матричной) вещественно аналитической функции $\Psi(z)$ обозначим

$$\Psi_{(0)} = \Psi(z_0), \quad \Psi_{(n)} = \frac{1}{n!}\frac{d^n}{dz^n}\Psi(z_0), \quad n = 1, 2, \ldots$$

$$\Psi_{<n>}(z) = \Psi_{(0)} + \sum_{k=1}^{n} \Psi_{(k)} z^k.$$

Следующая теорема является инструментом обоснования функционального подхода:

Теорема 1 *Пусть H открытое множество $\mathbb{R}^{m-1}$, $\tau \in H, z \in [z_1, z_2]$, $z_1 < z_2$ произвольные вещественные числа, а $\varphi(\tau, z)$ вещественная аналитическая функция на множестве $H \times [z_1, z_2]$. Введем обозначения:*

$$g(\tau, z) = \left(\frac{\partial}{\partial \tau_i}\varphi(\tau, z)\right)_{i=1}^{m-1},$$

$$V(\tau, z) = \left(\frac{\partial^2}{\partial \tau_i \partial \tau_j}\varphi(\tau, z)\right)_{i,j=1}^{m-1}.$$

При этих условиях если для некоторой точки $(\tau_{(0)}, z_0) \in H \times [z_1, z_2]$ выполняется равенство

$$g(\tau_{(0)}, z_0) = 0$$

и $\det V(\tau_{(0)}, z_0) \neq 0$, то:

I *В окрестности z_0 существует единственная непрерывная функция $\widehat{\tau}(z)$ такая, что $\widehat{\tau}(z_0) = \tau_0$ и $g(\widehat{\tau}(z), z) = 0$. Эта функция является вещественной аналитической в окрестности z_0.*

II *Имеют место следующие рекуррентные формулы:*

$$\tau_{(n+1)} = -[V^{-1}](g(\tau_{<n>}(z), z))_{(n+1)},\ n = 0, 1, \ldots, \tag{4}$$

где $V = V(\tau_{(0)}, z_0)$.

Первая часть этой теоремы есть известная в функциональном анализе теорема о неявном отображении, вторая часть доказана в работе ([5]). Формулы (4) называются основными рекуррентными формулами, они лежат в основе вычисления коэффициентов разложения опорных точек и весов оптимальных планов в ряд Тейлора.

Методология построения оптимальных планов включает в себя выбор класса планов, в котором ищется оптимальный план, выбор семейства априорных вероятностных мер и применение теоремы 1.

Обоснование использования этой теоремы проводится для наиболее важных D- и L-критериев оптимальности. Первым рассмотрен D- критерий, который требует максимизации величины $\det M(\xi, \theta)$.

Решение задач оптимизации (2) и (3) ищется в классе планов с m опорными точками, одна из которых совпадает с границей множества χ, а весовые коэффициенты равны между собой

$$\xi_\tau = \begin{pmatrix} \tau_1 & \ldots & \tau_{m-1} & d \\ 1/m & \ldots & 1/m & 1/m \end{pmatrix},$$

где d – граничная точка χ. Такие планы называются насыщенными планами. Они представляют интерес, поскольку для многих моделей локально D-оптимальные планы имеют такой вид (см. [5]).

Одна из трудностей байесовского подхода состоит в выборе априорной меры. Мы исходим из того, что этот выбор можно рассматривать как эвристический прием, позволяющий построить планы, эффективные для оценивания векторов θ, истинное значение которых принадлежит некоторому множеству.

Рассмотрим множество Ω следующего вида:

$$\Omega = \Omega_z = \Omega_z(c) = \{\theta = (\theta_1 \dots, \theta_m)^T : (1-z)c_i \le \theta_i \le c_i/(1-z),\ i = 1, \dots, m\}.$$

Здесь точка $c = (c_1, \dots, c_m)^T$ может рассматриваться как приближение вектора θ^*, а z — как относительная погрешность. В качестве априорного распределения рассмотрим равномерное распределение на Ω_z. Так как $\Omega(0) = \{c\}$, в качестве начального приближения τ_0 при построении разложения используется локально оптимальный план, который находится численно.

В третьем параграфе изучается невырожденность матрицы Якоби основной системы и приводятся условия выполнения теоремы 1:

- Функция $\eta(x, \theta)$ является аналитической на $[d_1, d_2] \times H$, где H окрестность точки $\theta = c$.
- В окрестности точки c существует локально оптимальный план вида
$$\xi_\tau = \begin{pmatrix} \tau_1 & \dots & \tau_{m-1} & d \\ 1/m & \dots & 1/m & 1/m \end{pmatrix}.$$
- Функции $f_i(x, \theta)$, $i = 1, \dots, m$ при любом фиксированном $\theta \in \Omega$ образуют обобщенную систему Чебышева второго порядка на интервале χ.

В **главе 3** исследуются оптимальные планы эксперимента для D-критерия.

В первом параграфе рассматривается модель Михаэлиса – Ментена. Доказано, что эта модель удовлетворяет условиям теоремы 1. Локально оптимальные и максиминные планы для данной модели найдены в явном виде, поэтому разложения в ряд Тейлора опорных точек строятся для байесовского плана. Далее изучается эффективность построенных планов с помощью численных расчетов. Зависимость эффективности от θ_2 представлена на рисунке 1. Как видно из этого рисунка байесовские планы являются более предпочтительными, чем максиминные планы и планы в равноотстоящих точках (рисунок соответствует значению $z = 0.8$).

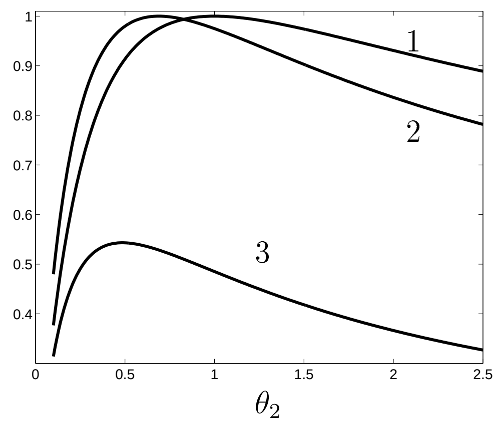

Рис. 1: Модель Михаэлиса – Ментена. Зависимость эффективности байесовского (1), максиминного (2) и плана в равноотстоящих узлах (3) от параметра θ_2

Во втором параграфе исследуются D-оптимальные планы для модели кинетики второго порядка. Доказывается, что опорные точки байесовских оптимальных планов являются аналитическими функциями параметра z, и строятся разложения этих функций в ряд Тейлора.

В третьем и четвертом параграфах исследуются аналитические свойства байесовских оптимальных планов для моделей экспоненциального вида. Эффективность методологии демонстрируется сравнением найденных планов с планами в равноотстоящих узлах и планами, численно найденными в литературе. Оказывается, что байесовские планы требуют приблизительно в два раза меньше наблюдений для достижения заданной точности, чем планы в равноотстоящих узлах.

В **главе 4** на основе функционального подхода исследуются оптимальные планы для L-критерия, требующего минимизации величины $\mathrm{tr} LM^{-1}(\xi, \theta)$, и для E-критерия, требующего максимизации минимального собственного числа информационной матрицы. Эти критерии часто более предпочтительны на практике, чем D-критерий. Планы, построенные на основе различных критериев, сравниваются по величине дисперсии оценок индивидуальных параметров.

В первом параграфе формулируется теорема эквивалентности, изучается невырожденность матрицы Якоби основного уравнения, формулируются достаточные условия для применения теоремы 1 к изучению L-оптимальных планов.

Во втором и третьем параграфах рассматриваются модель Михаэлиса – Ментена и модель экспоненциального роста. Проверяются условия выполнения теоремы 1. Обосновывается выбор весовой матрицы L. Строятся разложения опорных точек и весов локально оптимальных, байесовских и максиминных планов в ряд Тейлора. Показывается, что байесовские планы менее чувствительны к изменению z и более эффективны, чем планы в равноотстоящих узлах. Численно оценивается максимальное значение z, при котором насыщенный план, построенный на основе разложения, является оптимальным в классе планов с произвольным числом точек.

В четвертом параграфе изучены E-оптимальные планы для модели Михаэлиса – Ментена. Доказана аналитичность опорных точек и весов оптимальных планов как функций величины z и построено их разложение в ряд Тейлора

В **заключении** перечислены основные результаты диссертации.

1. Разработана методология построения и исследования байесовских оптимальных планов для нелинейных по параметрам регрессионных экспе-

риментов на основе функционального подхода.

2. Для случая D- и L-критериев оптимальности дано строгое математическое обоснование применения методологии для ряда нелинейных по параметрам моделей дробно-рационального и экспоненциального вида. Для этих моделей впервые построены разложения опорных точек и весов байесовских оптимальных планов в ряд Тейлора по величине относительной ошибки в задании параметров.

3. Для модели экспоненциального роста и модели кинетики второго порядка впервые изучены аналитические свойства опорных точек и весов максиминно эффективных планов.

4. Впервые проведено сравнительное исследование эффективности всех основных типов планов для нелинейных по параметрам моделей. Показано, что байесовские планы являются менее чувствительными к изменению параметров и позволяют значительно уменьшить количество экспериментов, необходимое для достижения заданной точности, по сравнению с планами в равноотстоящих точках.

Список литературы

[1] *Ермаков С.М., Жиглявский А.А.* Математическая теория оптимального эксперимента. М., Изд-во Наука, 1987.

[2] *Мелас В.Б.* Общая теория функционального подхода к оптимальному планированию эксперимента. Изд-во СПбГУ, 1999.

[3] *Chernoff H.* Locally optimal designs for estimating parameters. Ann. Math. Statist., 1953. Vol. 24. P. 586–602.

[4] *Melas V.B.* Optimal designs for exponential regression. Math. Operations forsh. Statist. Ser. Statisics, 1978. Vol. 9. P. 45–59.

[5] *Melas V.B.* Functional approach to experimental optimal design. Lecture Notes in Statistics, vol. 184. Heidelber. Springer, 2006.

[6] *Müller, C.H.* Maximin efficient designs for estimating nonlinear aspects in linear models. J. Statist. Plann. Inf., 1994. Vol. 44. P. 117–132.

Список публикаций автора

[7] *Старосельский Ю.М.* Исследование байесовских D-оптимальных планов для дробно–рациональных моделей. Вестн. С.-Петербург. ун-та, СПб., Изд-во СПбГУ, серия 10, вып. 3, с. 98–105, 2008.

[8] *Мелас В.Б. Старосельский Ю.М.* О методологии построения и исследования оптимальных планов экспериментов. В сб.трудов XX Междунар. науч. конф. В 10 т. Т. 1. Секция 1 / под общ. ред. В.С. Балакирева. - Ярославль: Изд-во Яросл. гос. техн. ун-та, 2007, с. 66–69.

[9] *Мелас В.Б. Старосельский Ю.М.* Исследование максиминно-эффективных планов для модели Михаэлиса – Ментена. Вестн. С.-Петербург. ун-та, СПб., Изд-во СПбГУ, серия 1, вып. 2, с. 41–50, 2007.