# Изучение байесовских экспериментальных планов для нелинейных регрессионных моделей

Колесник Ирина Владимировна, гр. 522

Санкт-Петербургский государственный университет
Математико-механический факультет
Кафедра статистического моделирования

Научный руководитель: д.ф.-м.н., Мелас В.Б.
Рецензент: Старосельский Ю.М.

Санкт-Петербург
2007г.

Одним из приложений теории планирования является планирование экспериментов в микробиологии и химической кинетике.
В работе рассмотрены две модели, использующиеся для описания кинетики ферментов:

- модель ферментативной кинетики первого порядка описывается уравнением Михаэлиса-Ментена:

$$\eta(x) = \theta_0 x/(\theta_1 + x),$$

  где $x$ - концентрация, выбираемая экспериментатором, $\theta_0$ и $\theta_1$ неизвестные параметры.

- модель ферментативной кинетики второго порядка:

$$\eta(x) = \theta_0 x(\theta_1 + x)/(\theta_2 + \theta_3 x + x^2).$$

# Основные понятия

- *Множество планирования* $\chi$: интервал возможных значений аргумента $[d_1, d_2]$.
- Виды планов:
  - *точный* или *дискретный план* эксперимента
    $\xi = \{x_1, \ldots, x_N\}, \; x_i \in \chi;$
  - *дискретный нормированный план эксперимента*
    $\xi = \{x_1, \ldots, x_n; \mu_1 \ldots, \mu_n\}$, где $x_i \in \chi, \; \mu_i > 0,$
    $\sum_{i=1}^n \mu_i = 1, \; \mu_i = r_i/N, r_i$ – целые числа, $i = 1, \ldots, n;$
  - *непрерывный план эксперимента* (дискретная вероятностная мера)
    $\xi = \{x_1, \ldots, x_n; \mu_1 \ldots, \mu_n\}$, где $x_i \in \chi, \; \mu_i > 0, \; \sum_{i=1}^n \mu_i = 1.$
- МНК: $\hat{\theta}; \quad V(\hat{\theta}|\theta, \xi)$

$$\hat{\theta} = \arg\min_{\theta \in \Theta} \sum_{j=1}^{N} \left(y_j - \eta(x_j, \theta)\right)^2$$

- Критерии оптимальности
  - $D$ критерий: $\det V(\hat{\theta}|\theta,\xi) \to \inf_{\xi}$;
  - $L$ критерий : $\operatorname{tr}\left\{WV(\hat{\theta}|\theta,X)(\xi)\right\} \to \inf_{\xi}$,

    $W$ - симметричная неотрицательно определенная матрица;
  - $A$ критерий (частный случай $L$ критерия):
    $W$ - диагональная, $W_{ii} > 0$
    (*взвешенная сумма дисперсий оценок параметров*).
- Планирование эксперимента для нелинейной регрессии:
  - локальный подход,
  - минимаксный подход,
  - байесовский подход:
    - априорное распределение $P(d\theta)$;
    - байесовский $L$-оптимальный план:

$$\xi_B = \arg\min_{\xi} \int_{\Theta} \operatorname{tr}\left\{WV(\hat{\theta}|\theta,X)(\xi)\right\} P(d\theta).$$

- Критерии оптимальности
  - $D$ критерий: $\det V(\hat{\theta}|\theta, \xi) \to \inf\limits_{\xi}$;
  - $L$ критерий : $\operatorname{tr}\left\{WV(\hat{\theta}|\theta, X)(\xi)\right\} \to \inf\limits_{\xi}$,

    $W$ - симметричная неотрицательно определенная матрица;
  - $A$ критерий (частный случай $L$ критерия):
    $W$ - диагональная, $W_{ii} > 0$
    (*взвешенная сумма дисперсий оценок параметров*).

- Планирование эксперимента для нелинейной регрессии:
  - локальный подход,
  - минимаксный подход,
  - байесовский подход:
    - априорное распределение $P(d\theta)$;
    - байесовский $L$-оптимальный план:

$$\xi_B = \arg\min_{\xi} \int_{\Theta} \operatorname{tr}\left\{WV(\hat{\theta}|\theta, X)(\xi)\right\} P(d\theta).$$

# Постановка задачи

Исследование двух методов построения байесовских планов:

- метод численного поиска точного байесовского плана, описанный в статье Steven G. Gilmour, Luzia A. Trinca, *Bayesian optimal exact design of experiments for enzyme kinetic models*;
- метод, использующий функциональный подход, разработанный В.Б. Меласом, осуществляющий построение непрерывного байесовского плана через разложение опорных точек и весов плана в ряд Тейлора по некоторому параметру; коэффиценты разложения вычисляются с помощью реккурентных формул.

Метод основан на применении алгоритма для поиска локального минимума функции на дискретном множестве точек:

- значение критерия $\int\limits_{\Theta} \operatorname{tr}\left\{WV(\hat{\theta}|\theta, X)(\xi)\right\} P(d\theta)$ оценивается методом Монте-Карло;
- алгоритм поиска минимума:
  - выбирается множество из $M$ точек-кандидатов;
  - выбирается стартовый план из $N$ необязательно различных точек этого множества;
  - каждая точка текущего плана последовательно заменяется на остальные точки множества;
  - изменение в плане «принимается», если оно улучшает значение критерия.

Нетрудно понять, что алгоритм находит лишь локальный экстремум.

СтатМод

- Общая схема метода:
  - Пусть $g(\tau, z): \ \mathbb{R}^m \times \mathbb{R} \to \mathbb{R}^m$, $g(\tau_0, z_0) = 0$, где $\tau, \tau_0 \in \mathbb{R}^m, \ z, z_0 \in \mathbb{R}$; и $\det\big(J(\tau_0, z_0)\big) \neq 0$, где $J(\tau, z) = \frac{\partial}{\partial \tau} g(\tau, z)$;
  - $\exists\ \tau^*(z)$ - вектор-функция: $\tau^*(z_0) = \tau_0$ и $g\big(\tau^*(z), z\big) = 0$, $\tau^*(z) = \tau_0 + \sum_{t=1}^{\infty} \tau^*_{(t)}(z - z_0)^t$;
  - реккурентные формулы для вычисления коэффициентов:
  $$\tau^*_{(s)} = -J^{-1}(\tau_0, z_0)\left(g\Big(\sum_{l=0}^{s-1} \tau^*_{(l)}(z - z_0)^l, z\Big)\right)_{(s)},$$
  где $f_{(s)}$ – $s$-тый коэффициент в разложении Тейлора.

- Построение плана разложением:
  - вектор $\tau(z)$ содержит в себе опорные точки и веса плана,
  - $g(\tau, z) = \big(\frac{\partial}{\partial \tau_1}\varphi(\xi, z), \ldots, \frac{\partial}{\partial \tau_m}\varphi(\xi, z)\big)$,
  
    где $\varphi(\xi, z) = \int\limits_{\Theta} \operatorname{tr}\{WV(\xi|\theta, z)\}\, P(d\theta)$.

# Выбор априорных распределений

- Ограничения.
  Параметры моделей, согласно своему физическому смыслу, принимают только положительные значения.

- Форма распределения.
  Равномерные априорные распределения параметров на отрезках вида:

$$\left[\theta_i^*(1-z), \theta_i^*/(1-z)\right], \quad z \in [0,1),$$

  где $\theta_i^*$ центральная точка распределения параметра $\theta_i$.
  При $z = 0$ распределение вырождается в точку $\theta_i^*$;
  $\left[\theta_i^*(1-z), \theta_i^*/(1-z)\right] \to [0, +\infty)$ при $z \to 1$.

- Независимость.
  Параметры $\theta_i$ предполагаются независимыми.

# Количество опорных точек в методе разложения

- Разложение строится для фиксированного числа опорных точек. При изменении $z$ количество точек плана не меняется.
- Разложение построено в окрестности точки $z_0 = 0$. Таким образом, в качестве начального приближения используется локально-оптимальный план.
- Как показывают эксперименты, локально $L$-оптимальные планы имеют минимальное количество опорных точек, равное количеству параметров модели.

Таким образом, в работе проведено изучение байесовских оптимальных планов с минимальной поддержкой.

- На практике могут быть использованы лишь дискретные планы, поэтому непрерывные планы требуют округления.

- Для сравнения двух планов используют понятие эффективности:

$$\mathrm{eff}_{\Phi}(\xi,\theta) = \left[\frac{\Phi(\theta,\xi^*)}{\Phi(\theta,\xi)}\right]^{1/m},$$

где $\Phi(\theta,\xi)$ – значение критерия на плане $\xi$ при фиксированном значении параметра $\theta$, $\xi^*$ – локально-оптимальный план.

- Сравнение эффективностей по $L$ критерию
  - точного плана, построенного численно,
  - округленного непрерывного плана, построенного разложением

показывает, точный план незначительно превосходит округленный непрерывный план по эффективности при различных значениях $z$.

| $z$ | 0.1 | 0.3 | 0.5 |
|---|---|---|---|
| $\min \mathrm{eff}(\xi_{exact})$ | 0.9934 | 0.8869 | 0.6238 |
| $\min \mathrm{eff}(\xi_{dec})$ | 0.9924 | 0.8903 | 0.6071 |

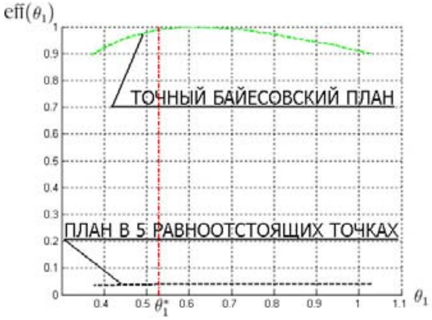
- Время построения разложения значительно меньше времени работы алгоритма для поиска точного плана.

Таким образом можно рекомендовать использовать метод разложения для построения байесовских планов.

Сравнение эффективностей

- точного плана, построенного численно,
- округленного непрерывного плана, построенного разложением
- плана в равноотстоящих точках

показывает, что планы в равноотстоящих точках оказываются крайне неэффективными в случае $L$ критерия.



# Сравнение $D$ и $L$ планов по критерию $e_j M^{-1} e_j^T$

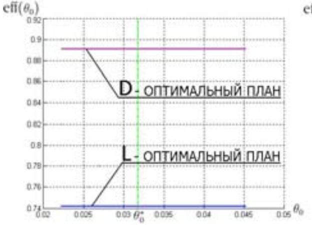
Сравнение эффективностей по критерию индивидуальной оценки параметров

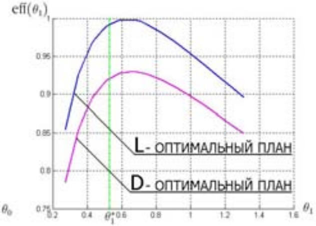
- $D$-оптимального байесовского плана, построенного численно
- $L$-оптимального байесовского плана, построенного разложением

показывает, что выбирая различными весовые матрицы $W$, можно построить $L$-оптимальные планы, которые будут минимизировать дисперсии оценок заранее выбраных параметров.

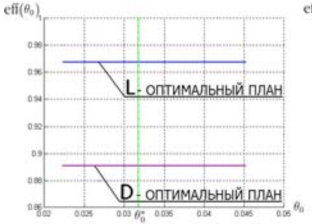
# Выбор весовой матрицы $W$

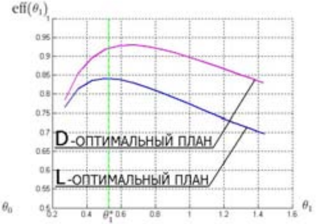
- $W = \mathrm{diag}\{280, 1\}$




- $W = \mathrm{diag}\{28000, 1\}$



В работе получены следующие результаты:

1. реализован метод численного построения точных оптимальных байесовских планов;
2. реализован метод разложения опорных точек и весовых коэффициентов непрерывных байесовских оптимальных планов в ряды Тейлора по степеням параметра, характеризующего размах распределения;
3. с помощью описанных методов построены и исследованы байесовские оптимальные планы для двух нелинейных по параметрам моделей, имеющих практическое применение;
4. в результате проведенных исследований установлены преимущества подхода, основанного на степенных разложениях.