Исследование алгоритмов поиска оптимальных планов эксперимента

Старосельский Евгений Михайлович, 522-я группа

Санкт-Петербургский Государственный Университет
Математико-механический факультет
Кафедра статистического моделирования

Научный руководитель — к.ф.-м.н., доцент **Пепелышев Андрей Николаевич**
Рецензент — д.ф.-м.н., профессор **Мелас Вячеслав Борисович**

Санкт-Петербург
2007г.

- Пусть результаты эксперимента проведенного в точках $x_1 \dots x_N$, представляются в виде:

$$y_j = \eta(x_j, \theta) + \varepsilon_j, \quad j = 1, \dots, N,$$

$\theta \in \mathbf{R}^k$ — вектор оцениваемых параметров.
Пусть предположения относительно ошибок имеют вид

$$E\varepsilon_j = 0,\ E\varepsilon_j\varepsilon_i = 0 \text{ для } j \neq i \text{ и } D\varepsilon_j = \sigma^2.$$

Таблица: Примеры векторов функций регрессии

Название	**Вектор регрессии** $f(x) = \partial\eta/\partial\theta$
2-degree polynomial	$f(x) = (1,\ x,\ x^2)$
3-degree polynomial	$f(x) = (1,\ x,\ x^2,\ x^3)$
mixed poly/trig	$f(x) = (x,\ x^2,\ \sin(2\pi x),\ \cos(2\pi x))$
exponential	$f(x) = (e^{-x},\ xe^{-x}, e^{-2x},\ xe^{-2x})$
fractional	$f(x) = (1,\ 1/(1+x),\ 1/(1+x)^2)$

- *Непрерывный план эксперимента* — дискретная вероятностная мера задаваемая таблицей:

$$\xi = \begin{pmatrix} x_1 & \dots & x_n \\ \omega_1 & \dots & \omega_n \end{pmatrix},$$

где $x_1, \dots, x_n \in \mathcal{X}$ — различные моменты наблюдений, а $\omega_1, \dots, \omega_n > 0$ $(\sum \omega_j = 1)$ — весовые коэффициенты.

- *Информационная матрица* плана ξ:

$$M(\xi, \theta) = \left(\sum_{l=1}^{n} \frac{\partial}{\partial \theta_i} \eta(x_l, \theta) \frac{\partial}{\partial \theta_j} \eta(x_l, \theta) \omega_l \right)_{i,j=1}^{k}.$$

Информационная матрица характеризует качество МНК оценки

$$\mathrm{Cov}(\hat{\theta}) \approx \frac{\sigma^2}{N} M^{-1}(\xi).$$

- План ξ^* будем называть локально D-оптимальным, если он является решением задачи

$$\det M(\xi) \longrightarrow \sup_{\xi}.$$

- План ξ^* назовем локально A-оптимальным, если он является решением задачи

$$\mathrm{trace} M^{-1}(\xi) \longrightarrow \inf_{\xi}.$$

- План ξ^* назовем локально E-оптимальным, если он является решением задачи

$$\lambda_{\min} M(\xi) \longrightarrow \sup_{\xi}.$$

Рекуррентные алгоритмы Титтерингтона (1976) для D-критерия

Будем считать, что точки плана фиксированы.

- Алгоритм 1.

$$\omega_i^{(r+1)} = \omega_i^{(r)} \frac{d(x_i, \xi^r)}{k},$$

- Алгоритм 2.

$$\omega_i^{(r+1)} = \frac{1}{b} \left(\omega_i^{(r)} + \frac{1}{k} - \frac{1}{d(x_i, \xi)} \right)_+,$$

- Алгоритм 3.

$$\omega_i^{(r+1)} = \omega_i^{(r)} \frac{d(x_i, \xi^r) - 1}{k - 1},$$

где r — номер итерации,

$$d(x_i, \xi) = \frac{\partial}{\partial \omega_i} \ln \det M(\xi) = f^T(x_i) M^{-1} f(x_i)$$

есть двойственная функция из теоремы эквивалентности Кифера-Вольфовица.

Модифицированный алгоритм 1

- Алгоритм 4, содержащий параметр, контролирующий сходимость

$$\omega_i^{(r+1)} = \omega_i^{(r)} \frac{d(x_i, \xi^r) - \alpha c(\xi^{(\tau)})}{b},$$

где $c(\xi) = \min_i d(x_i, \xi), \alpha \in [0, 1]$. Оптимальным значением параметра α является:

$$\alpha^* = \arg \max_{\alpha \in [0,1]} \det M(\xi_\alpha^{r+1}).$$

Таблица: Число шагов и время работы алгоритма с параметром сходимости $\alpha = 0.6$

	Алгоритм 4		Алгоритм 1	
Модель регрессии	**Число шагов**	**Время**	**Число шагов**	**Время**
2-degree polynomial	64	8.91	103	13.69
3-degree polynomial	78	10.98	129	17.33
mixed poly/trig	93	12.69	131	18.40
exponential	144	19.60	220	28.93
fractional	65	8.59	104	13.75

- Композиционный алгоритм 5
  1. Вычисляем
$$\omega_i^{(r+1)} = \frac{1}{b}\left(\omega_i^{(r)} + \frac{1}{k} - \frac{1}{d(x_i,\xi)}\right).$$
  2. Если $\det M(\xi^{r+1}) \leq \det M(\xi^r)$, то пересчитываем веса по формуле
$$\omega_i^{(r+1)} = \omega_i^{(r)} \frac{d(x_i,\xi^r)}{k}.$$

Таблица: Число шагов и время работы Композиционного алгоритма

	Алгоритм 5		Алгоритм 1	
Модель регрессии	**Число шагов**	**Время**	**Число шагов**	**Время**
2-degree polynomial	20	1.51	103	13.69
3-degree polynomial	62	4.75	129	17.33
mixed poly/trig	53	4.07	131	18.40
exponential	106	7.85	220	28.93
fractional	31	2.39	104	13.75

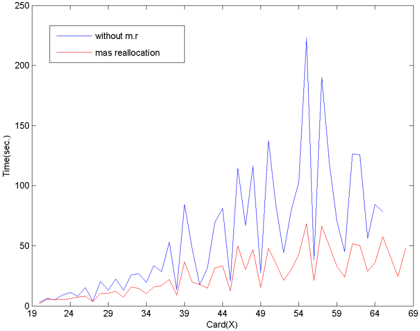
Модифицированный алгоритм с учетом теоремы Пронзато (2003)

Алгоритм 6, который минимизирует зависимость от числа точек в плане.

1. Вычисляем:
$$\omega_i^{(r+1)} = \frac{1}{b}\left(\omega_i^{(r)} + \frac{1}{k} - \frac{1}{d(x_i,\xi)}\right). \tag{1}$$
2. Если det $M(\xi^{r+1}) \le$ det $M(\xi^r)$, то
$$\omega_i^{(r+1)} = \omega_i^{(r)} \frac{d(x_i,\xi^r)}{k}.$$
3. Вычисляем значение: $\epsilon_r = \max\limits_{x_i \in \mathcal{X}} f^T(x_i)M^{-1}(\xi^r)f(x_i) - k$. Если
$$f^T(x_i)M^{-1}(\xi^k)f(x_i) < kr(\epsilon_k),$$
где (Pronzato, 2003)
$$r(\epsilon) = 1 + \frac{\epsilon}{2} - \frac{\sqrt{\epsilon(4+\epsilon)}}{2},$$
то индекс i включаем в множество J_r для выбрасывания точек из плана. Если J_r не пусто, производим перераспределение весов по формуле:
$$\omega_i^{r+1} = \omega_i^r + \beta_r \frac{f^T(x_i)M^{-1}(\xi^r)f(x_i) - k}{\sum_{j \not\subset J_r}[f^T(x_j)M^{-1}(\xi^r)f(x_j) - k]},\ x_i \in \mathrm{supp}\{\xi^r\},\ i \not\subset J_r,$$
$$\beta_r = \sum_{i \in J_r} \omega_i^r.$$

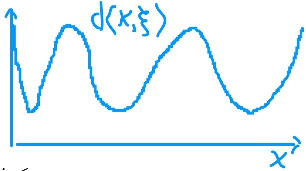
График 1.

- Время работы (Time) в зависимости от размера плана (Card(X)) для полиномиальной модели 3-ей степени.



Кластерный подход (Mandal and Torsney, 2006)

1. Перед основной частью производим 10 итераций Алгоритма 1.
2. Производим кластеризацию весов: разбиваем множество весов на (непересекающиеся) интервалы по локальным минимумам весовой функции. Предположим, что в результате у нас получилось m кластеров.
3. Пусть n_l — размер l-ого кластера, $l = \{1, \dots, m\}$. Пусть p_l^i, $1 \leq i \leq n_l$ — веса точек в l-ом кластере. Перейдем к новым весам:



$$q_l = \sum_{i=1}^{n_l} p_l^i, \; 1 \leq l \leq m,$$

суть суммарные веса каждого из кластеров.

$$r_{l,i} = p_{l,i}/q_l, \; 1 \leq l \leq m, \; 1 \leq i \leq n_l,$$

суть веса точек внутри кластеров.

4. Применяем Алгоритм 1 поочередно для весов кластеров и весов точек в каждом кластере.

Сочетание кластерного и композиционного подхода

1. Перед основной частью производим 10 итераций Алгоритма 6.
2. Производим кластеризацию
3. Используя данный подход, модифицируем веса по следующему принципу:

$$q_l^{r+1} = \frac{1}{b}\left(q_l^r + \frac{1}{k} - \frac{1}{d_l^-}\right),$$

$$r_{l,i}^{(r+1)} = \frac{r_{l,i}^{(r)}(d_{l,i}^-)^\delta}{\sum_{h=1}^m r_{l,h}^{(r)}(d_{l,h}^-)^\delta}, 1 \le l \le m.$$

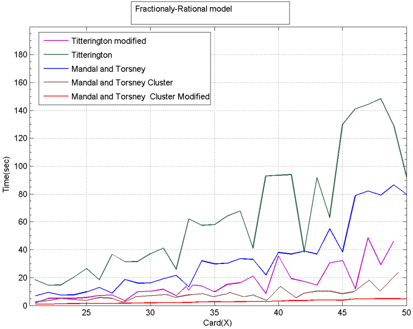
4. Находим истинные значения весов: $\omega^{r+1} = r_{l,i}^{r+1} q_l^{r+1}$.
5. Производим проверку: если $\det(M(x,\omega^{(r+1)})) \le \det(M(x,\omega^{(r)}))$, то пересчитываем суммарные веса: $\omega_i^{(r+1)} = \frac{\omega_i^{(r)} d(x_i,\xi^r)}{\sum d(x_i,\xi^r)\omega_i^{(r)}}$.
6. Вычисляем: d_l^- и $d_{l,i}^-$ ($\frac{\partial\phi}{\partial q_l}, \frac{\partial\phi}{\partial r_{l,i}}$) находим по следующим формулам:

$$d_l^- = \sum_{i=1}^{n_l} d_{l,i} r_{l,i}, \quad d_{l,i}^- = q_l d_{l,i}, \quad d_{l,i} = d(x_i, \xi^r).$$

Таблица: Число шагов и время работы Композиционного алгоритма с кластеризацией весов

Модель регрессии	Число шагов	Время
2-degree polynomial	10+3	1.70
3-degree polynomial	10+3	1.74
mixed poly/trig	10+4	1.90
exponential	10+5	1.94
fractional	10+3	1.71

- Время работы различных алгоритмов в зависимости от размера плана.



Результаты кластерный алгоритма для других критериев.

Таблица: Число шагов и время работы кластерного алгоритма E-критерия.

Модель	Итерации	Время
2-degree polynomial	10+10	1.73
3-degree polynomial	10+8	1.54
exponential	10+10	1.76
fractional	10+9	1.56

Таблица: Число шагов и время работы кластерного алгоритма для A-критерия.

Модель	Итерации	Время
2-degree polynomial	10+6	1.54
3-degree polynomial	10+7	1.79
exponential	10+12	2.08
fractional	10+7	1.68

Таблица: Число шагов и время работы кластерного алгоритма для байесовского D-критерия.

Модель	Итерации	Время
2-degree polynomial	10+3	1.70
3-degree polynomial	10+3	1.74
exponential	10+7	2.23
fractional	10+9	2.46

Заключение.

- В первой части дипломной работы путем комбинирования и улучшения рекуррентных формул Титтерингтона (Titterington, 1976), получен новый алгоритм, который использует уже известный кластерный подход. Основной идеей модификации являлось исправление быстрого алгоритма 2 Титтерингтона, в случае немонотонности, а также удаление из плана точек, которые не могут являться опорными.
- Построенный алгоритм имеет более высокую скорость сходимости, чем уже существующие, при этом сохраняет свойство монотонности и обладает высокой эффективностью.
- Во второй части работы рассматриваются применения построенного модифицированного кластерного алгоритма к критериям E и A оптимальности, а также к задаче нахождения байесовских D-оптимальных планов, которые являются наиболее приближенными к практическим задачам планирования. Наблюдения показали, что построенный модифицированный алгоритм сохраняет свою структуру и для рассмотренных критериев, более того, эффективность планов, посчитанных для различных критериев является достаточно высокой.