ТЕСТОВАЯ ОЦЕНКА ТОЧНОСТИ РЕШЕНИЯ ЗАДАЧИ КОМПАРАТОРНОЙ ИДЕНТИФИКАЦИИ ПАРАМЕТРОВ МОДЕЛИ МНОГОФАКТОРНОГО ОЦЕНИВАНИЯ

Петров К.Э.
Харьковский национальный университет радиоэлектроники,
61166, г. Харьков, пр. Ленина, 14, кафедра системотехники, тел (057) 702-10-06,
E-mail: kept@mail.ru.

This work is devoted the analysis of results of test modeling with the purpose of estimation of exactness, adequacy and calculable complication of decision of problem of parametrical identification of the model of multifactor estimation. Exactness of identification substantially depends on the concrete features of problem, such as a quantity of local descriptions of alternatives, power of the set of alternatives, complication of polynomial and method of identification. As a class of possible structures of the model of estimation is suggested to use the polynomial of Kolmogorov-Gabor.

Введение

Параметрическая идентификация является обязательным этапом структурно-параметрического синтеза модели многофакторного оценивания, а также имеет самостоятельное значение в случаях, когда структура модели априорно известна. Кроме того, в связи с тем, что коэффициенты определяют относительную важность различных локальных характеристик оцениваемых альтернатив, их численные значения необходимы для решения многих практических задач. Поэтому оценка точности их вычисления представляет большой теоретический и практический интерес.

Постановка задачи компараторной идентификации модели многофакторного оценивания, а также соответствующая математическая модель описаны в [1, 2] и в данной работе рассматриваться не будут. Отметим лишь, что многофакторные оценки определяются в классе полинома Колмогорова-Габора, а все соотношения, входящие в модель компараторной идентификации, являются линейными, относительно искомых коэффициентов (параметров).

Эта задача является некорректной по Адамару (не имеет единственного решения) и для определения единственного точечного решения ее необходимо регуляризировать. Для этого в работе используются методы определения чебышевского решения [1], средней точки [1], а также методы, основанные на применении генетических алгоритмов [3]. Естественно, метод регуляризации специфически влияет на точность определения точечных значений параметров $A=\langle a_1,a_2,...,a_L\rangle$. Поэтому анализ точности проведен для каждого метода регуляризации в отдельности.

В каждом конкретном случае модель представляет собой некоторый фрагмент полинома Колмогорова-Габора, вид которого (структура и численные значения параметров) определяются в результате решения задачи компараторной идентификации. Эта процедура представляет собой последовательность синтеза вариантов структуры и оценки ее параметров, циклически повторяющихся до тех пор, пока будет обеспечено требуемое ”качество” модели оценивания.

Целью работы является сравнительный анализ точности определения числовых значений коэффициентов полинома (параметров модели оценивания), при заданной его структуре, различными методами по результатам тестового моделирования.

Оценка точности решения задачи параметрической идентификации модели многофакторного оценивания

Для достижения сформулированной цели примем следующую методологию анализа. Для каждого анализируемого случая эвристически задается мощность n множества альтернатив $X=\{x_1,x_2,...,x_n\}$, на котором решается задача компараторной параметрической идентификации, размерность R кортежа локальных характеристик

$K = \langle k_1, k_2, ..., k_R \rangle$ и вид (структура) полинома оценивания $P(x_i)$, $x_i \in X$, что в свою очередь однозначно определяет размерность L кортежа коэффициентов $A^{ý} = \left\langle a_1^{ý}, a_2^{ý}, ..., a_L^{ý} \right\rangle$, которая равна числу слагаемых полинома. Затем случайным образом, генерируем реализацию численных значений кортежа параметров (относительных весовых коэффициентов) модели, для которых выполняются ограничения

$$0 \le a_l^{ý} \le 1, \ l = \overline{1, L}; \ \sum_{l=1}^{L} a_l^{ý} = 1 \tag{1}$$

и численные значения кортежа локальных характеристик $K(x_i) = \langle k_1(x_i), k_2(x_i), ..., k_R(x_i) \rangle$ для каждой альтернативы $x_i \in X$, $i = \overline{1, n}$.

На основе перечисленных исходных данных вычисляем эталонные значения скалярных многофакторных оценок всех альтернатив $P^{ý}(x_i)$, $x_i \in X$, $i = \overline{1, n}$ и устанавливаем на множестве X отношение порядка по убыванию значений $P^{ý}(x_i)$. На этом заканчивается формирование эталонной ситуации.

Для моделирования решения задачи компараторной параметрической идентификации в качестве исходных данных задаются следующие эталонные данные:

- множество альтернатив X;
- численные значения кортежа локальных характеристик $K(x_i)$, $x_i \in X$, $i = \overline{1, n}$;
- вид (структура) полинома оценивания;
- информация о качественном отношении порядка на множестве альтернатив X.

Информация об отношении порядка на X задается в виде единственной лучшей альтернативы или отношения порядка на всем множестве альтернатив X.

По перечисленным данным решается задача компараторной параметрической идентификации и, в результате, определяются численные ”модельные” значения параметров полинома оценивания $A^{ì} = \left\langle a_1^{ì}, a_2^{ì}, ..., a_L^{ì} \right\rangle$, удовлетворяющие ограничениям (1).

С учетом значений $A^{ì}$, характеристик $K(x_i)$ и заданной структуры полинома для каждой альтернативы $x_i \in X$ вычисляем значения ”модельных” скалярных оценок $P^{ì}(x_i)$, $x_i \in X$, $i = \overline{1, n}$ и устанавливаем ”модельное” отношение порядка на множестве X.

Точность идентификации определяется путем сравнения эталонных и модельных значений скалярных оценок $P(x_i)$, кортежей параметров A отношений порядка на X и A. Для формализации такого сравнения введем следующие критерии.

1. Абсолютная погрешность определения оценок $P(x_i)$

$$\Delta P(x_i) = \left| P^{ì}(x_i) - P^{ý}(x_i) \right|, \ i = \overline{1, n}. \tag{2}$$

2. Экстремальные абсолютные погрешности определения оценок $P(x_i)$

$$\begin{aligned} \Delta P_{max} &= \max_i \Delta P(x_i); \\ \Delta P_{min} &= \min_i \Delta P(x_i). \end{aligned} \tag{3}$$

3. Средняя абсолютная погрешность определения оценок

$$\Delta P_{\text{ñð}} = \frac{\sum\limits_{i=1}^{n} \Delta P(x_i)}{n}. \qquad (4)$$

4. Средняя квадратичная погрешность определения оценок

$$\Delta P_{\text{êâ}} = \frac{\sum\limits_{i=1}^{n} \left[\Delta P(x_i)\right]^2}{n}. \qquad (5)$$

5. Коэффициент ранговой корреляции Спирмена, учитывающий отклонения модельного отношения порядка значений оценок от эталонного

$$r_P = 1 - \frac{6 \cdot \sum\limits_{i=1}^{n} (R_i^{\text{ì}} - R_i^{\text{ý}})^2}{n \cdot (n^2 - 1)}, \qquad (6)$$

где $R_i^{\text{ì}}$ и $R_i^{\text{ý}}$ – ранги (места) модельных и эталонных значений оценок в установленном отношении порядка на множестве X.

Аналогичные по смыслу критерии, предлагается использовать и для анализа различий "модельных" и эталонных значений параметров A.

Результаты исследования и выводы

В качестве примера приведем сводную таблицу значений критерия (4) полученных для различных методов.

Таблица 1.

Значения критериев оценки точности идентификации параметров $\Delta a_{\text{ñð}}$ и функции полезности $\Delta P_{\text{ñð}}$, полученных с использованием различных методов (задано отношение порядка на множестве альтернатив X)

$A = \langle a_l \rangle$, $l = \overline{1, L}$	$X = \{x_i\}$, $i = \overline{1, n}$	$\Delta a_{\text{ñð}}$				$\Delta P_{\text{ñð}}$			
		Чеб.т.	Сред.т. (вар. 1)	Сред.т. (вар. 2)	ГА	Чеб.т.	Сред.т. (вар. 1)	Сред.т. (вар. 2)	ГА
$L = 3$	$n = 5$	0.1000	0.0557	0.0569	0.0829	0.0756	0.0387	0.0434	0.0603
	$n = 10$	0.0107	0.0041	0.0039	0.0111	0.0066	0.0026	0.0024	0.0068
	$n = 15$	0.0107	0.0122	0.0115	0.0035	0.0049	0.0066	0.0062	0.0019
$L = 5$	$n = 5$	0.0280	0.0201	0.0084	0.0465	0.0396	0.0246	0.0089	0.0475
	$n = 10$	0.0564	0.0231	0.0249	0.0137	0.0266	0.0136	0.0132	0.0087
	$n = 15$	0.0364	0.0169	0.0170	0.0340	0.0185	0.0095	0.0087	0.0193
$L = 7$	$n = 5$	0.0560	0.0450	0.0450	0.0411	0.0347	0.0332	0.0332	0.0313
	$n = 10$	0.0206	0.0278	0.0272	0.0255	0.0151	0.0220	0.0216	0.0144
	$n = 15$	0.0123	0.0045	0.0048	0.0119	0.0094	0.0036	0.0039	0.0084
$L = 9$	$n = 5$	0.0707	0.0598	0.0598	0.0334	0.0914	0.0762	0.0762	0.0354
	$n = 10$	0.0380	0.0349	0.0297	0.0197	0.0295	0.0355	0.0349	0.0128
	$n = 15$	0.0211	0.0177	0.0480	0.0116	0.0167	0.0116	0.0167	0.0231

Результаты тестового моделирования подтвердили корректность и работоспособность метода компараторной идентификации для параметрической идентификации модели многофакторного оценивания.

Тестовое моделирование однозначно подтверждает теоретическое утверждение о большей информативности ситуации, когда на множестве альтернатив задано отношение полного порядка по сравнению со случаем, когда задана одна лучшая альтернатива.

Анализ результатов тестового решения задачи параметрической идентификации методами чебышевской и средней точек показывает, что:

− оценки параметров, полученные по методу средней точки (вариант 1) и (вариант 2) очень близки, а в большинстве случаев совпадают, поэтому можно рекомендовать для практического использования более простую формулу (вариант 1);

− на данном этапе исследования невозможно сделать однозначный вывод о предпочтительности по точности методов чебышевской точки или средних значений, так как точность зависит от конфигурации многогранника допустимых решений, определяемого конкретными количественными и качественными характеристиками ограничений модели компараторной идентификации.

Учитывая сказанное можно рекомендовать определять оценки параметров моделей обоими методами и затем проводить эвристический анализ полученных результатов с учетом предметной области.

Время решения задачи параметрической идентификации модели оценивания для задач большой размерности методами чебышевской и средней точек изменяется в соответствии с полиномиальной зависимостью близкой к линейной (от 0.03 сек. при $L=3$ и $n=5$ до 1.14 сек. при $L=9$ и $n=15$).

Использование для параметрической идентификации метода основанного на использовании ГА позволяет получить более высокую точность определения оценок по сравнению с методами чебышевской и средней точек. Следует отметить закономерное повышение точности с ростом мощности множества альтернатив. Например, при $n=15$, во всех случаях методами ГА получены более точные результаты. Это объясняется увеличением обучающей последовательности, на которой адаптируется ГА. Платой за точностную эффективность ГА является существенное увеличение временных затрат на реализацию метода (до 7 мин. при $L=9$ и $n=15$, и количестве популяций параметров равным 30).

Список литературы

1. Овезгельдыев А.О., Петров Э.Г., Петров К.Э. Синтез и идентификация моделей многофакторного оценивания и оптимизации. – К.: Наук. думка, 2002. – 164 с.

2. Петров К.Э. Проблема формализации интеллектуальной деятельности человека// Бионика интеллекта. 2004. – № 61. – С. 86 – 90.

3. Овезгельдыев А.О., Петров К.Э. Построение модели индивидуального многофакторного оценивания с применением элементов МГУА и генетических алгоритмов // Кибернетика и системный анализ. – 2007. – №1. – С. 151 – 159.