УДК 519.833

Н.М. Оскорбин

Математические модели систем с латентными переменными*

N.M. Oskorbin

Mathematical Models of Systems with Latent Variables

В статье рассмотрены проблемы математического моделирования трех классов систем с латентными переменными, для которых предложены структуры моделей и методы идентификации параметров. Результаты исследования носят теоретический характер, а предложенные модели и методы требуют детализации с учетом особенностей систем с латентными переменными и имеющейся априорной информации.

Ключевые слова: системы с латентными переменными; регрессионный анализ, метод главных компонент, теоретико-игровые модели, декомпозиция игры.

In the paper author consider problems of mathematical modeling of three classes of systems with latent variables. For these classes of systems structures of models and methods for identification of parameters are proposed. Results of research have theoretical character, and the offered models and methods demand on specification taking into account features of systems with latent variables and available a priori information.

Key word: latent variable models, regression analysis, principal component analysis, game-theoretic models, game decomposition.

Рассматриваются проблемы математического моделирования систем, в которых в процессе их функционирования отдельные существенные переменные не наблюдаются. В литературе этот тип переменных называют латентными [1]. Введем обозначения: $v \in R^m$ – латентные переменные; $z \in R^n$ – наблюдаемые (контролируемые) переменные; $y \in R^p$ – выходные переменные моделируемого объекта. Тогда система с латентными переменными может быть представлена в следующем виде:

$$y = f^r(z,v), \;\; z \in Z \subset R^n, \quad v \in V \subset R^m. \quad (1)$$

В данном случае без потери общности мы рассматриваем частный случай реальных объектов, которые характеризуются только количественными переменными, а их функционирование не зависит от времени.

Системы типа (1) изучаются, например, в психологии и социологии [2], когда переменные z – результаты тестирования индивида, v – параметры состояния личности. При моделировании экономических процессов исходной информацией выступает статистическая, бухгалтерская отчетность. При этом скрытыми остаются параметры поведения производственного персонала, технические и технологические факторы.

Проведем исследование трех классов систем (1), для которых возможно построение работоспособных математических моделей. Первый класс систем (1) рассматривается в классическом регрессионном анализе, в рамках которого искомая математическая модель представляется в следующем виде $(p = 1)$:

$$y = f(z,\beta) + \varepsilon(v), \quad (2)$$

где β – идентифицируемые параметры заданной структуры функциональной зависимости; ε – агрегированная переменная, которая предполагается нормально распределенной случайной величиной, имеющей нулевое математическое ожидание и постоянную дисперсию.

Предположим, что переменные $\{y_i, z_i\}_{i=1}^N$ для N независимых объектов аналогов измеряются точно. Тогда идентификации подлежит функция регрессии $f(z,\beta) = M_\varepsilon[y / z]$ – условное математическое ожидание выходной переменной системы при фиксированном значении z, а задача решается методом наименьших квадратов. Качество моделирования (показатель работоспособности полученной математической модели) оценивается, например, коэффициентом детерминации, нормативные значения которого приведены в [3]. Заметим, что при наличии ошибок измерения переменной $y_i, i = 1,\dots,N$ в моделях прогноза необходима коррекция этого показателя [4], которая способствует более точной идентификации дисперсии переменной ε в модели (2). В результате математического моделирования систем

* Работа выполнена в соответствии с проектом РФФИ №12-01-98008 (конкурс р_сибирь_а).

с латентными переменными получается модель (2), которая широко используется в теоретических и прикладных исследованиях.

Второй класс систем (1) имеет структуру следующего вида:

$$y = f(z, \tilde{v}(z)), \; z \in Z \subset R^n. \qquad (3)$$

В данном случае идентификации подлежат вектор-функция $\tilde{v}(z): Z \to V$ и системная вектор-функция $f(z,v): Z \times V \to Y$, где Y – множество значений выходных переменных моделируемого объекта. В этом классе систем (1) значения латентных переменных полностью определяются наблюдаемыми переменными. Как будет показано ниже, в частном случае структуры модели (3) задача идентификации решается методом главных компонент (МГК), который эффективно применяется не только для оценки латентных структур, но и как метод многомерного анализа экспериментальных данных [2]. В практике применения МГК задача моделирования решается успешно, если достаточно высоко значение коэффициента детерминации, а латентные переменные допускают непротиворечивую интерпретацию в рамках теории рассматриваемой предметной области. Модель (3) может быть рассмотрена в контексте модели (2), т.е. с включением в ее структуру аддитивной или мультипликативной составляющей ошибки описания системы (1).

Рассмотрим подход к решению задачи идентификации модели (3). Пусть функции $f(z, \tilde{v}(z)), \tilde{v}(z)$ заданы с точностью до значений векторов β^0, β, т.е.:

$$v = \tilde{v}(z, \beta); \; y = f(z, \beta^0, \tilde{v}(z, \beta)). \qquad (4)$$

Тогда при известной выборке $\{y_i, z_i\}_{i=1}^N$ для N объектов-аналогов задача идентификации модели (4) может быть записана так. Найти оптимальные значения β^{0*}, β^* векторных параметров из условия ($p = 1$):

$$s = \min\left\{ \left\| y^э - y^p \right\| \,|\, \beta^0 \in R^l, \beta \in B, \quad i = 1,...,N \right\}, \; (5)$$

где $\left\| y^э - y^p \right\|$ – норма разности векторов $y^э = (y_1, ..., y_N)^T$ и y^p с элементами $y_i^p = f(z_i, \beta^0, \tilde{v}(z_i, \beta))$; B – множество допустимых значений вектора β; l – размерность вектора β^0.

Условия задачи (5) при моделировании реальных систем необходимо конкретизировать с учетом дополнительной информации о свойствах объекта моделирования, в частности, о характере распределения латентных переменных. В вычислительном отношении эта задача может быть решена с использованием пакетов программ математических вычислений, в том числе с помощью инструмента «Поиск решения» в среде Excel. Возможны ее постановки в рамках прикладного интервального анализа и в рамках блочного программирования [5].

В частном случае линейных вектор-функций $v = \tilde{v}(z, \beta)$, $y = f(\beta^0, v))$, условий ортогональности латентных переменных и использования для β собственных векторов матрицы $(z^{эT} \cdot z^э)$, где $z^э = (z_1, ..., z_N)^T$ решение задачи (5) совпадает с оценками МГК. Пусть $p = 1$, $n = m$ (размерность вектора латентных переменных совпадает с размерностью вектора z), а функции $v = \tilde{v}(z, \beta)$, $y^p = f(\beta^0, v))$ заданы при всех $i = 1, \dots, N$ следующими выражениями:

$$y_i^p = \sum_{j=1}^n \beta_j^0 v_{ij} + \beta_0^0, \quad v_{ij} = \sum_{\kappa=1}^n \beta_{jk} z_{ik}, \; i = 1,...,N, \; (6)$$

где β – квадратная $(n \times n)$ матрица коэффициентов связи латентных переменных с наблюдаемыми переменными системы (1); $v^j = (v_{1j}, ..., v_{Nj})^T$ – вектор значений латентной переменной j для всех объектов-аналогов; $\beta^0 = (\beta_0^0, ..., \beta_n^n)$ – $(n+1)$-мерный вектор коэффициентов связи латентных переменных с выходной переменной y.

Можно показать, что в условиях невырожденности матрицы $(z^{эT} \cdot z^э)$ решение задачи (5) с функциями (6) и евклидовой нормой совпадает с оценками МГК, если множество B задать следующими выражениями:

$$(\beta \cdot \beta^T) = E_n; \; (v^j, v^k) = 0,$$
$$j = 1,...,n; k = (j+1),...,n, \qquad (7)$$

где E_n – единичная матрица размера n; (v^j, v^k) – скалярное произведение соответствующих векторов.

Установленная связь задачи (5) с МГК позволяет предложить пути ее конкретизации и рассмотреть проблемные аспекты. Основная проблема постановки и решения задачи (5) возникает в связи с тем, что область значений функции $v = \tilde{v}(z, \beta)$ не задана. В задаче (5) эта область косвенно задается множеством B, которое в МГК задано ограничениями (7). Недостатки данного способа связаны с тем, что предполагается зависимость латентных переменных от всех наблюдаемых переменных и ортогональность соответствующих векторов, т.е. их взаимная независимость в моделируемой системе. По нашему мнению, это требование существенно ограничивает сферу применения МГК, приводит к ошибкам интерпретации латентных переменных и к усложнению вычислительных процессов при моделировании систем (1) в случае их большой размерности. Для задачи (5) можно предложить достаточно универсальную систему ограничений, аналогичную (7), которая используется при декомпозиции блочных задач стохастического программирования [5]. Суть ее состоит в задании функции распределения значений латентных переменных с последующей настройкой параметров этой функции в процессе моделирования. На практике можно ограничиться заданием начальных моментов этого многомерного распределения вероятностей. Учитывая избыточность искомых параметров, в задаче (5) практически приемлемым способом является поиск нормированных значений латентных переменных с нулевыми средними и единичными разбросами. В качестве начального приближения для множества B можно предложить следующие условия:

$$\|\beta\| \le R; \quad \frac{1}{N}\sum_{i=1}^{N} v_{ij} = 0; \quad \frac{1}{N}\sum_{i=1}^{N} v_{ij}^2 = 1,\ j = 1,\dots,m. \quad (8)$$

В данном случае не регламентируется число латентных переменных и структура функции $v = \tilde{v}(z,\beta)$. Внешний параметр R в (8) настраивается при решении задачи (5), а в случае, когда выполняется равенство $\|\beta\| = R$, система ограничений (8) или структура функции $v = \tilde{v}(z,\beta)$ корректируются с использованием дополнительной информации о свойствах исследуемой системы.

Рассмотрим системы (1), которые имеют следующее описание:

$$y = f(\tilde{z}(v)),\ v \in V \subset R^m;\ \tilde{z}(v): V \to Z. \quad (9)$$

В данном случае латентные переменные детерминируют контролируемые переменные z, но непосредственно не оказывают влияния на значения выходных переменных. Такие системы возникают, например, при моделировании социально-экономических процессов, в которых переменные z контролируются верхним уровнем системы (центром), а значения латентных (скрытых для центра) переменных выбирают активные элементы системы. Тогда модель (9) может быть представлена игрой двух лиц в следующем виде:

$$y_ц = f_ц(z^0, \tilde{z}(v)) \to \max_{z^0 \in Z^0}; \quad (10)$$

$$y_{АЭ} = f_{АЭ}(z^0, \tilde{z}(v)) \to \max_{v \in V}. \quad (11)$$

Решение игры (10), (11) базируется на эквивалентном ее преобразовании:

Задача центра:

$$y_ц = f_ц(z^0, \tilde{z}(v)) \to \max_{z^0 \in Z^0}. \quad (12)$$

Задача активных элементов системы:

$$y_{АЭ} = f_{АЭ}(z^0, z) \to \max_{z \in Z}. \quad (13)$$

Вспомогательная задача активных элементов системы:

$$\delta = \min_{v \in V} \|z^* - \tilde{z}(v)\| = 0. \quad (14)$$

Экономический смысл полученных выражений состоит в следующем. Центр и активные элементы системы «играют» на множествах значений контролируемых переменных системы, поскольку при известном игрокам множестве Z выражения (12), (13) не зависят от латентных переменных. При нахождении компромиссного решения (z^{0*}, z^*) активный элемент системы выбирает оптимальные значения латентных переменных из задачи (14), обеспечивая тем самым реализацию своего игрового решения $z^* = \tilde{z}(v^*)$, где v^* – оптимальное для системы поведение активных элементов. Описанный подход моделирования можно применять при произвольном числе активных элементов.

Рассмотрим в качестве примера экономическую систему (9), в которой активный элемент выбором латентных для центра переменных способен получить дополнительную прибыль $\Delta \in [0, \bar{\Delta}],\ \bar{\Delta} > 0$. Центр часть прибыли $\sigma\Delta,\ \sigma \in [0,1]$ передает активный элемент в качестве вознаграждения. Модельное описание данной системы (выражения (10), (11)) примет вид:

$$y_ц = f_ц(\sigma, \tilde{\Delta}(v)) = (1-\sigma)\,\tilde{\Delta}(v) \to \max_{\sigma \in [0,1]}; \quad (15)$$

$$y_{АЭ} = f_{АЭ}(\sigma, \tilde{\Delta}(v), \tilde{\varphi}(\tilde{\Delta}(v))) =$$

$$= \sigma\,\tilde{\Delta}(v) - \tilde{\varphi}(\tilde{\Delta}(v)) \to \max_{v \in V}, \quad (16)$$

где $\tilde{\varphi}(\tilde{\Delta}(v))$ – функция трудозатрат АЭ, которая, по предположению, зависит только от контролируемой центром переменной $\Delta = \tilde{\Delta}(v)$.

Установив соответствия переменных моделей (10), (11) и (15), (16): $z^0 \Leftrightarrow \sigma$; $z \Leftrightarrow (\Delta, \varphi)$, запишем выражения, аналогичные (12)–(14):

Задача центра:

$$y_ц = f_ц(\sigma, \Delta) = (1-\sigma)\,\Delta \to \max_{\sigma \in [0,1]}. \quad (17)$$

Задача активных элементов системы:

$$y_{АЭ} = f_{АЭ}(\sigma, \Delta) = \sigma\,\Delta - \tilde{\varphi}(\Delta) \to \max_{\Delta \in [0, \bar{\Delta}]}. \quad (18)$$

Вспомогательная задача активных элементов системы:

$$\delta = \min_{v \in V} \|\Delta^* - \tilde{\Delta}(v)\| = \|\Delta^* - \tilde{\Delta}(v^*)\| = 0, \quad (19)$$

где Δ^* – оптимальная стратегия активных элементов системы в игре (17), (18); v^* – решение задачи (19).

Игра (17), (18) исследована в литературе как одна из базовых моделей стимулирования [6], как модель «Государство–Предприниматели» [7] в предположении, что игрокам известны все ее функции и значения параметров. Выбор латентных переменных осуществляет активный элемент, как «чистый исполнитель» решений, найденных на множествах контролируемых переменных моделируемой системы. Модель (17)–(19) полностью определяет поведение центра и активных элементов, значения системных (выходных) переменных и адекватна реальному объекту в том случае, когда в (19) имеет место равенство $\delta = 0$. Если $\delta > 0$, то полученная модель не адекватна реальной системе. Приведенные условия контролируются после реализации решения в рассматриваемой системе или на объектах-аналогах.

Заключение. На основании проведенных исследований можно утверждать, что проблема математического моделирования систем (1) в общем случае не разрешима. Выделены частные случаи систем (1), для которых возможно построение работоспособных моделей:

1. Системы с оцениваемой неопределенностью (модель (2)).

2. Системы с латентными агрегатами (модель (3)), которые назовем LC-системами.

3. Системы с управляемыми латентными переменными (модель (9)), которые назовем CL-системами.

Для моделирования LC-систем предложен метод идентификации латентных переменных, который является обобщением МГК, и рассмотрены вычислительные проблемы решения соответствующих математических задач.

Записан вариант математической модели управляемых CL-систем, в которых задача выбора латентных переменных сформулирована как задача реализации согласованных значений контролируемых переменных. Результаты исследования носят теоретический характер, а предложенные модели требуют детализации с учетом особенностей систем с латентными переменными и имеющейся априорной информации.

Библиографический список

1. Латентные переменные (Latent variable) [Электронный ресурс]. – URL: http://en.wikipedia.org/wiki/Latent_variable.

2. Гуц А.К., Фролова Ю.В. Математические методы в социологии. – М., 2007.

3. Ехлаков Ю.П. Теоретические основы автоматизированного управления. – Томск, 2001.

4. Максимов А.В., Оскорбин Н.М. Многопользовательские информационные системы: основы теории и методы исследования. – Барнаул, 2005.

5. Мамченко О.П., Оскорбин Н.М. Моделирование иерархических систем. – Барнаул, 2007.

6. Новиков Д.А., Губко М.В. Теория игр в управлении организационными системами. – М., 2002.

7. Денисенко В.В., Оскорбин Н.М. Сравнение стратегий Г1 и Г2 в игре «Государство–Предприниматели» // Известия АлтГУ. – 2010. – №1.