Труды Карельского научного центра РАН
№ 3. 2010. С. 83–86

УДК 519.2

ПОЛИНОМИАЛЬНЫЕ ИНВАРИАНТЫ ДЛЯ МОДЕЛЕЙ С ЛАТЕНТНЫМИ ПЕРЕМЕННЫМИ

С. В. Стафеев

Институт прикладных математических исследований Карельского научного центра РАН

В статье рассматривается модель факторного анализа с зависимыми остатками. Стандартный подход к получению статистических выводов с помощью этой модели основан на параметрическом представлении и на методе максимального правдоподобия. К сожалению, такой подход не всегда применим. Например, данная ситуация возникает тогда, когда выборочная матрица ковариаций вырождена. Один из подходов для преодоления данных трудностей состоит в использовании полиномиальных инвариантов, представляющих из себя полиномы, которые зависят от элементов матрицы ковариаций наблюдаемых случайных величин. В статье разработан метод получения таких инвариантов для произвольного числа латентных переменных.

Ключевые слова: факторный анализ, латентные переменные, полиномиальные инварианты.

S. V. Stafeev. POLYNOMIAL INVARIANTS FOR MODELS WITH LATENT VARIABLES

Statistical inference in factor analysis models is usually based on parametric representation and on maximum likelihood estimates. However , such approach is not always applicable. For example, the case when a sample covariance matrix is singular the likelihood ratio test cannot be used. One of the approaches for overcoming these difficulties consists in using model invariants, that is, polynomial equality relations that the model imposes on the entries of the covariance matrix of the observed variables. In this paper a simple method for finding model invariants for any number of latent variables is developed.

Key words: factor analysis, latent variables, polynomial invariants.

Постановка задачи

Рассмотрим следующую модель факторного анализа:

$$X_i = \sum_{j=1}^{k} a_{ij} H_j + Y_i, \quad i = 1,...,n, \tag{1}$$

где $\mathbf{X} = \{X_1,...,X_n\}^t$ – вектор наблюдаемых случайных величин с матрицей ковариаций $\Sigma = (\sigma_{ij})$; $\mathbf{H} = \{H_1,...,H_k\}^t$ – множество независимых в совокупности нормально распределенных латентных (ненаблюдаемых) случайных

величин (факторов), причем $\mathrm{M}H_i = 0$, $\mathrm{D}H_i = 1$, $i = 1,...,k$; $A = (a_{ij})$ – матрица факторных нагрузок. Вектор остатков $\mathbf{Y} = \{Y_1,...,Y_k\}^t$ имеет невырожденное нормальное распределение с нулевым вектором математических ожиданий и матрицей ковариаций $\Theta = (\theta_{ij})$. Векторы $\mathbf{Y}$ и $\mathbf{H}$ являются независимыми. Взаимосвязь компонент вектора $\mathbf{Y}$ представляется ковариационной графовой моделью [Стафеев, 2005] со структурой $G = (V, E)$, где $V = \{1,...,n\}$, а $E = \{(i, j) \mid i \neq j, \theta_{ij} \neq 0\}$.

Стандартный подход к получению статистических выводов с помощью этой модели основан на параметрическом представлении модели и на методе максимального правдоподобия [Harman, 1976]. К сожалению, такой подход не всегда применим. Например, данная ситуация возникает тогда, когда выборочная матрица ковариаций вырождена. Заметим, что когда число наблюдений меньше числа переменных, это всегда так. Один из подходов для преодоления данных трудностей состоит в использовании полиномиальных инвариантов модели. Инвариант модели (1) представляет из себя полином, который зависит от элементов матрицы ковариаций наблюдаемых случайных величин. Заметим, что проблема нахождения инвариантов модели (1) связана с проблемой параметрической идентифицируемости модели (1) [Стафеев, 2006, 2009]. Основная проблема данного подхода состоит в том, что нахождение инвариантов для $k > 2$ очень трудно с вычислительной точки зрения [Drton et al., 2007]. В статье разработан простой метод получения таких инвариантов для произвольного числа латентных переменных.

Полиномиальные инварианты

Легко видеть, что матрица ковариаций наблюдаемых случайных величин Σ допускает следующее представление: $\Sigma = AA^t + \Theta$. (2)

Пусть P_G^k – множество всех положительно определенных матриц, допускающих разложение (2), и пусть $f(\Sigma)$ – полином, зависящий от элементов σ_{ij} матрицы Σ. Полином $f(\Sigma)$ называется полиномиальным инвариантом модели (1), если $f(\Sigma) = 0$ для всех матриц $\Sigma \in \mathrm{P}_G^k$. Легко видеть, что множество всех инвариантов модели (1) образует идеал $I_{k,n}^G$ кольца многочленов $\mathbf{R}[\sigma_{ij}, i < j, (i, j) \notin E]$. Заметим, что согласно известной теореме Гильберта, любой идеал имеет конечный базис.

Пусть $M_{k,n} \subseteq \mathbf{R}[\sigma_{ij}, i < j]$ – идеал, который генерируется всеми $(k+1)\times(k+1)$ минорами симметричной матрицы Σ. (Другими словами, базис идеала $M_{k,n}$ состоит из всех $(k+1)\times(k+1)$ миноров матрицы Σ.) В работе [Drton, Sullivant, 2007] показано, что

$$I_{k,n} = M_{k,n} \bigcap \mathbf{R}[\sigma_{ij}, i < j],$$

где $I_{k,n}$ – идеал $I_{k,n}^G$, в котором граф G не содержит ребер. Аналогичным образом, можно показать:

$$I_{k,n}^G = I_{k,n} \bigcap \mathbf{R}[\sigma_{ij}, i < j, (i, j) \notin E] =$$
$$M_{k,n} \bigcap \mathbf{R}[\sigma_{ij}, i < j, (i, j) \notin E].$$

В работе [Loera et al., 1995] найден базис идеала $I_{k,n}$. Он состоит из полиномов вида

$$\sigma_{ij}\sigma_{lm} - \sigma_{li}\sigma_{jm}, \quad i \neq j \neq l \neq m. \qquad (3)$$

Полиномы вида (3) называются тетрадами.

Заметим, что простейшими инвариантами, принадлежащими $I_{k,n}$, являются не содержащие диагональных элементов $(k+1)\times(k+1)$ миноры матрицы Σ. Примером нетривиального инварианта для $k = 2$ и $n = 5$ является пентада [Harman, 1976], имеющая вид:

$$\sigma_{12}\sigma_{23}\sigma_{34}\sigma_{45}\sigma_{51} - \sigma_{12}\sigma_{23}\sigma_{35}\sigma_{14}\sigma_{45} -$$
$$\sigma_{12}\sigma_{24}\sigma_{35}\sigma_{34}\sigma_{15} + \sigma_{12}\sigma_{24}\sigma_{13}\sigma_{45}\sigma_{35} +$$
$$\sigma_{12}\sigma_{24}\sigma_{13}\sigma_{45}\sigma_{35} - \sigma_{12}\sigma_{24}\sigma_{13}\sigma_{45}\sigma_{35} -$$
$$\sigma_{13}\sigma_{24}\sigma_{35}\sigma_{14}\sigma_{25} + \sigma_{13}\sigma_{25}\sigma_{34}\sigma_{24}\sigma_{15} +$$
$$\sigma_{14}\sigma_{23}\sigma_{13}\sigma_{45}\sigma_{25} - \sigma_{14}\sigma_{25}\sigma_{23}\sigma_{45}\sigma_{15} -$$
$$\sigma_{15}\sigma_{23}\sigma_{13}\sigma_{24}\sigma_{45} + \sigma_{15}\sigma_{24}\sigma_{23}\sigma_{14}\sigma_{35}.$$

Определим основной объект данной работы. Пусть $\overline{G} = (V, \overline{E})$ – дополнительный для G граф. Определим граф $\mathbf{G} = (\mathbf{V}, \mathbf{E})$, где $\mathbf{V} = \{\mathbf{i} = (i_1,...,i_k), 1 \leq i_1 < i_2 < ... < i_k \leq k\}$,

$$\mathbf{E} = \{(\mathbf{i}, \mathbf{j}) \mid \mathbf{i} = (i_1,...,i_k), \mathbf{j} = (j_1,...,j_k),$$
$$(i_s, j_l) \in \overline{E}, s, l = 1,...,k\}.$$

Заметим, что граф $\mathbf{G}$ был впервые определен в работе [Stafeev, 2007] в контексте проблемы параметрической идентифицируемости модели (1).

Обозначим через $|\Sigma_{\mathbf{i},\mathbf{j}}|$ детерминант матрицы $\Sigma_{\mathbf{i},\mathbf{j}} = (\sigma_{i_m, j_l})_{m,l=1}^k$. Пусть $C = (V_C, E_C)$, где $V_C = \{\alpha_i, i = 1,...,m\}$, а $E_C = \{(\alpha_i, \alpha_{i+1}), i = 1,...,m-1, (\alpha_1, \alpha_m)\}$, – четный простой цикл. Определим следующий полином:

$$f_C(\Sigma) = \prod_{(\mathbf{i},\mathbf{j})\in E_C^{'}} |\Sigma_{\mathbf{i},\mathbf{j}}| - \prod_{(\mathbf{i},\mathbf{j})\in E_C^{''}} |\Sigma_{\mathbf{i},\mathbf{j}}|,$$

где $E_C^{'} = \{(\alpha_{2j+1}, \alpha_{2j+2}),\ j = 0,...,m/2-1\}$ и $E_C^{''} = E_C \setminus E_C^{'}$.

Пусть $T = (V_T, E_T)$ – граф, состоящий из простого нечетного цикла с множеством ребер $E_1 = \{(\alpha_1,\alpha_2),...,(\alpha_{m_1-1},\alpha_{m_1})\}$, простой цепи с множеством ребер $E_2 = \{(\alpha_{m_1},\alpha_{m_1+1}),...,(\alpha_{m_2-1},\alpha_{m_2})\}$ и простого нечетного цикла с множеством ребер $E_3 = \{(\alpha_{m_2},\alpha_{m_2+1}),...,(\alpha_{m-1},\alpha_m),(\alpha_m,\alpha_{m_2})\}$. Определим следующий полином:

$$f_T(\Sigma) = \prod_{(\mathbf{i},\mathbf{j})\in E_T^{'}\cap(E_1\cup E_3)} |\Sigma_{\mathbf{i},\mathbf{j}}| \prod_{(\mathbf{i},\mathbf{j})\in E_T^{'}\cap E_2} |\Sigma_{\mathbf{i},\mathbf{j}}|^2 -$$

$$\prod_{(\mathbf{i},\mathbf{j})\in E_T^{''}\cap(E_1\cup E_3)} |\Sigma_{\mathbf{i},\mathbf{j}}| \prod_{(\mathbf{i},\mathbf{j})\in E_T^{''}\cap E_2} |\Sigma_{\mathbf{i},\mathbf{j}}|^2,$$

где $E_T^{'} = \{(\alpha_{2j+1},\alpha_{2j+2}),\ j = 0,...,[m/2]-1\}$, $E_T^{''} = E_T \setminus E_T^{'}$.

Следующая теорема описывает большой класс полиномиальных инвариантов.

Теорема 1. *Пусть* $\mathbf{C}(\mathbf{G})$ *– множество всех четных простых циклов графа* $\mathbf{G}$*, а* $\mathbf{T}(\mathbf{G})$ *– множество всех подграфов графа* $\mathbf{G}$*, состоящих из двух нечетных простых циклов, соединенных простой цепью. Тогда*

$$\{f_C(\Sigma),\ f_T(\Sigma),\ C\in \mathbf{C}(\mathbf{G}),\ T\in \mathbf{T}(\mathbf{G})\} \subset I_{k,n}^G.$$

Доказательство. Рассмотрим случай $k=1$. Модель (1) в этом случае имеет вид:

$$X_i = a_i H + Y_i,$$

где $A = \{a_1,...,a_n\}^t$. Используя независимость H и $\mathbf{Y}$, получаем для $(i,j)\notin E$

$$\sigma_{ij} = \mathrm{M}(X_i X_j) = \mathrm{M}[(a_i \mathrm{H} + Y_i)(a_j H + Y_j)] = a_i a_j. \quad (4)$$

Пусть $C = (V_C, E_C)$ – четный простой цикл графа $\mathbf{G}$, причем $V_C = \{\alpha_i,\ i = 1,...,m\}$, а

$E_C = \{(\alpha_i,\alpha_{i+1}),\ i = 1,...,m-1,\ (\alpha_1,\alpha_m)\}$. Получаем:

$$f_C(\Sigma) = \prod_{(i,j)\in E_C^{'}} \sigma_{ij} - \prod_{(i,j)\in E_C^{''}} \sigma_{ij} = \prod_{i=1}^{m} a_{\alpha_i} - \prod_{j=1}^{m} a_{\alpha_j} = 0.$$

Пусть $T = (V_T, E_T)$ – подграф графа $\mathbf{G}$, состоящий из простого нечетного цикла с множеством ребер $E_1 = \{(\alpha_1,\alpha_2),...,(\alpha_{m_1-1},\alpha_{m_1})\}$, простой цепи с множеством ребер $E_2 = \{(\alpha_{m_1},\alpha_{m_1+1}),...,(\alpha_{m_2-1},\alpha_{m_2})\}$ и простого нечетного цикла с множеством ребер $E_3 = \{(\alpha_{m_2},\alpha_{m_2+1}),...,(\alpha_{m-1},\alpha_m),(\alpha_m,\alpha_{m_2})\}$. Получаем

$$f_T(\Sigma) = \prod_{(i,j)\in E_T^{'}\cap(E_1\cup E_3)} \sigma_{ij} \prod_{(i,j)\in E_T^{'}\cap E_2} \sigma_{ij}^{\ 2} -$$

$$\prod_{(i,j)\in E_T^{''}\cap(E_1\cup E_3)} \sigma_{ij} \prod_{(i,j)\in E_T^{''}\cap E_2} \sigma_{ij}^{\ 2} =$$

$$\prod_{i=1}^{m_1-1} a_{\alpha_i} \prod_{j=m_1}^{m_2} (a_{\alpha_j})^2 \prod_{l=m_1+1}^{m} a_{\alpha_l} -$$

$$\prod_{i=1}^{m_1-1} a_{\alpha_i} \prod_{j=m_1}^{m_2} (a_{\alpha_j})^2 \prod_{l=m_1+1}^{m} a_{\alpha_l} = 0.$$

Таким образом, для $k=1$ теорема доказана. Рассмотрим общий случай. Обозначим

$$\mathbf{X_i} = \{X_{i_1},...,X_{i_k}\}^t, \quad \mathbf{Y_i} = \{Y_{i_1},...,Y_{i_k}\}^t,$$

$$\mathbf{a_i} = \{a_{i1},...,a_{ik}\}^t, \quad A_{\mathbf{i}} = \{\mathbf{a_{i_1}},...,\mathbf{a_{i_k}}\}^t.$$

Легко видеть, что

$$\mathbf{X_i} = A_{\mathbf{i}}\mathbf{H} + \mathbf{Y_i}, \quad \mathbf{i}\in \mathbf{V}.$$

Для непересекающихся $\mathbf{i}$ и $\mathbf{j}$ получаем

$$\Sigma_{\mathbf{ij}} = \mathrm{M}[\mathbf{X_i}\mathbf{X_j}^{\mathrm{t}}] = \mathrm{M}[(A_{\mathbf{i}}\mathbf{H} + \mathbf{Y_i})(A_{\mathbf{j}}\mathbf{H} + \mathbf{Y_j})^{\mathrm{t}}] = A_{\mathbf{i}}(A_{\mathbf{j}})^t.$$

Таким образом, $\Sigma_{\mathbf{ij}} = A_{\mathbf{i}}(A_{\mathbf{j}})^t$ для непересекающихся $\mathbf{i}$ и $\mathbf{j}$. Получаем, что

$$|\Sigma_{\mathbf{ij}}| = |A_{\mathbf{i}}||(A_{\mathbf{j}})^t| = |A_{\mathbf{i}}||A_{\mathbf{j}}|, \quad (\mathbf{i},\mathbf{j})\in \mathbf{E}. \quad (5)$$

Очевидно, что соотношение (5) аналогично соотношению (4). Таким образом, подставляя в полученных полиномиальных инвариантах для случая $k=1$ вместо ковариаций σ_{ij} соответствующие миноры $|\Sigma_{\mathbf{ij}}|$, получаем требуемый результат.

Пусть $\mathbf{m} = \{m_1,...,m_{k-1}\}$ и $\mathbf{d} = \{d_1,...,d_{k-1}\}$, $\mathbf{m}$ и $\mathbf{d}$ не пересекаются, $i_1 \neq i_2 \neq i_3$, $\{i_1,i_2,i_3\}\notin \mathbf{m}\cup\mathbf{d}$, и пусть граф $\mathbf{G}$ содержит цикл с множеством ребер

$$\{(<i_1,\mathbf{m}>,<i_2,\mathbf{d}>),\ (<i_3,\mathbf{m}>,<i_2,\mathbf{d}>),$$
$$(<i_3,\mathbf{m}>,<i_1,\mathbf{d}>),\ (<i_2,\mathbf{m}>,<i_1,\mathbf{d}>),$$
$$(<i_2,\mathbf{m}>,<i_3,\mathbf{d}>),\ (<i_1,\mathbf{m}>,<i_3,\mathbf{d}>)\}.$$

Тогда, используя теорему 1, получаем следующий инвариант

$$|\Sigma_{<i_1\mathbf{d}><i_2\mathbf{m}>}||\Sigma_{<i_3\mathbf{d}><i_1\mathbf{m}>}||\Sigma_{<i_2\mathbf{d}><i3\mathbf{m}>}| -$$
$$|\Sigma_{<i_3\mathbf{d}><i_2\mathbf{m}>}||\Sigma_{<i2_1\mathbf{d}><i1_2\mathbf{m}>}||\Sigma_{<i_1\mathbf{d}><i3_2\mathbf{m}>}| = 0.$$

Например, в случае $k=2$ и $n=5$ этот инвариант имеет вид:

$$f_C^{2,5} = \begin{vmatrix} \sigma_{12} & \sigma_{32} \\ \sigma_{14} & \sigma_{34} \end{vmatrix} \begin{vmatrix} \sigma_{12} & \sigma_{52} \\ \sigma_{13} & \sigma_{53} \end{vmatrix} \begin{vmatrix} \sigma_{12} & \sigma_{42} \\ \sigma_{15} & \sigma_{45} \end{vmatrix} -$$

$$\begin{vmatrix} \sigma_{12} & \sigma_{52} \\ \sigma_{14} & \sigma_{54} \end{vmatrix} \begin{vmatrix} \sigma_{12} & \sigma_{42} \\ \sigma_{13} & \sigma_{43} \end{vmatrix} \begin{vmatrix} \sigma_{12} & \sigma_{32} \\ \sigma_{15} & \sigma_{35} \end{vmatrix}.$$

Инвариант $f_C^{2,5}$ является произведением σ_{12} и пентады.

Пусть S – выборочная матрица ковариаций. Для проверки нулевой гипотезы $H_f : f(\Sigma) = 0$ можно использовать статистику $f(S)$ [Drton et al., 2008]. Обозначим через $Var_\Sigma(f(S))$ дисперсию $f(S)$. Дисперсия $Var_\Sigma(f(S))$ является полиномом, который зависит от элементов матрицы Σ. Заменяя матрицу Σ оценкой S, получаем оценку для дисперсии статистики $Var_S(f(S))$. В условиях нулевой гипотезы стандартизованная статистика

$$\frac{f(S)}{\sqrt{Var_S(f(S))}}$$

асимптотически имеет стандартное нормальное распределение.

Литература

Стафеев С. В. Факторный анализ с зависимыми остатками: проблема идентифицируемости и оценка параметров // Труды ИПМИ КарНЦ РАН. 2005. Вып. 6. С. 119–130.

Стафеев С. В. О параметрической идентифицируемости модели факторного анализа с зависимыми факторами и остатками // Труды ИПМИ КарНЦ РАН. 2006. Вып. 7. С. 63–80.

Стафеев С. В. О полиномиальных инвариантах для графовых моделей с латентными переменными // Обозрение прикладной и промышленной математики. 2009. Т. 14. Вып. 6. С. 1058–1064.

Drton M., Massam H., Olkin I. Moments of minors of Wishart Matrices // Annals of Statistics. 2008. Vol. 36. P. 2261–2283.

Drton M., Sturmfels B., Sullivant S. Algebraic factor analysis: tetrads, pentads and beyond // Probability Theory and Related Fields. 2007. Vol. 138. P. 463–493.

Drton M., Sullivant S. Algebraic statistical models // Statistica Sinica. 2007. Vol. 17. P. 1273–1297.

Harman H. Modern Factor Analysis. University of Chicago Press, third edition, 1976. 450 p.

Loera J. A., Sturmfels B., Thomas R. R. Grobner bases and triangulations of the second hypersimplex // Combinatorica. 1995. Vol. 15. P. 409–424.

Stafeev S. V. On factor analysis models with correlated residuals // Extended Abstracts of Russian – Scandinavian Symposium «Probability theory and applied probability». Petrozavodsk, 2006. P. 59–61.

Stafeev S. V. On the parameter estimation of recursive «bow – free» models with latent variables // Proceedings of the Eighth International Minsk Conference Computer Data Analysis and Modeling: Complex Stochastic Data and Systems. Minsk, 2007. Vol. 1. P. 178–181.

СВЕДЕНИЯ ОБ АВТОРЕ:

Стафеев Сергей Вячеславович
научный сотрудник, к. ф.-м. н.
Институт прикладных математических исследований КарНЦ РАН
ул. Пушкинская, 11, Петрозаводск, Республика Карелия, Россия, 185910
эл. почта: stafeev@krc.karelia.ru
тел.: (8142) 763370

Stafeev, Sergey
Institute of Applied Mathematical Research, Karelian Research Centre, Russian Academy of Science
11 Pushkinskaya St., 185910 Petrozavodsk, Karelia, Russia
e-mail: stafeev@krc.karelia.ru
tel.: (8142) 763370