**КОМПЬЮТЕРНЫЕ ИССЛЕДОВАНИЯ
И МОДЕЛИРОВАНИЕ 2013 Т. 5 № 2 С. 123−130**

Ки&М

МАТЕМАТИЧЕСКИЕ ОСНОВЫ И ЧИСЛЕННЫЕ МЕТОДЫ МОДЕЛИРОВАНИЯ

УДК: 519.237.07

Поиск косоугольной факторной структуры методом «облимакс»

В. А. Шовин

Омский филиал Учреждения Российской академии наук
Института математики им. С. Л. Соболева Сибирского отделения РАН,
Россия, 644099, г. Омск, ул. Певцова, д. 13

E-mail: v.shovin@mail.ru

*Получено 8 декабря 2012 г.,
после доработки 8 марта 2013 г.*

Предлагается усовершенствованный метод косоугольного вращения «облимакс». Рассматривается проблема выбора пар факторов. Предлагается оригинальная последовательность выбора пар факторов, приводящая к наилучшей факторной структуре по критерию «облимакс».

Ключевые слова: факторный анализ, косоугольное вращение «облимакс»

Search oblique factor structure by the "oblimax" method

V. A. Shovin[1]

[1] Omsk Branch of the Institution of the Russian Academy of Sciences Institute of Mathematics. S. Siberian Branch of RAS, 13 Pevtsov st., Omsk, 644099, Russia

Abstract. — An improved method of oblique rotation “oblimax” is suggested. The problem of choosing the pairs of factors is discussed. A novel sequence to select pairs of factors that lead to the best factor structure by “oblimax” is suggested.

Keywords: factor analysis, oblique “oblimax” rotation

Citation: *Computer Research and Modeling*, 2013, vol. 5, no. 2, pp. 123–130 (Russian).

© 2013 Владимир Алексеевич Шовин

Введение

Построение линейных факторных моделей широко используется в психологии, в нейрофизиологии, социологии, политологии, в экономике, статистике и других науках.

Поэтому представляется актуальным построение эффективных алгоритмов в этой области. Одной из главных проблем факторного моделирования является проблема вращения.

Проблема вращения возникает ввиду неоднозначности факторного решения. Существует бесконечное множество равноценных с точки зрения критерия минимальных невязок факторных решений, но при этом имеющих различное распределение факторных нагрузок переменных. Новое факторное решение можно получить путем вращения старого факторного решения не изменяя при этом значение критерия минимальных невязок.

Возможны два вида вращения: ортогональное и косоугольное.

В случае ортогональных методов вращения результирующая матрица вращения должна быть ортогональной, это значит, что произведение транспонированной матрицы вращения и исходной матрицы вращения должно равняться единичной матрице вращения.

В случае косоугольных методов вращения на результирующую матрицу вращение накладывается лишь следующее ограничение: у матрицы, являющейся результатом произведения транспонированной косоугольной матрицы вращения и исходной косоугольной матрицы вращения, на диагонали должны стоять единицы, тогда как внедиагональные элементы могут быть отличны от нуля.

Смысл этих условий заключается в том, что в случае ортогонального факторного вращения результирующие факторы остаются ортогональными, тогда как в случае косоугольного факторного вращения конечные факторы могут не быть друг другу ортогональны, т. е. геометрически это означает, что угол между косоугольными факторами может отличаться от прямого.

Матрицу вращения находят с помощью специального критерия. Смысл всех критериев сводится к максимизации доли больших и маленьких элементов в конечной факторной структуре, что позволяет получить максимально «интерпретабельное» факторное решение [Иберла, 1980].

Косоугольное вращение дает большую свободу при выборе положений каждой оси, благодаря чему можно получить большое число нулевых факторных нагрузок и достичь простой структуры. Но если между результирующими факторами, выделенными при косоугольном вращении, будут отсутствовать корреляции, то можно использовать ортогональное вращение.

Факторная модель

Факторный анализ основывается на следующей линейной модели, связывающей исходные показатели Z и факторы P:

$$Z = AP,$$

где $\underset{m\times n}{Z} \leftrightarrow z_{ij}$ — матрица размерности $m\times n$ значений m параметров у n объектов, $\underset{g\times n}{P} \leftrightarrow p_{ij}$ — матрица размерности $g\times n$ значений g факторов у n объектов, $\underset{m\times g}{A} \leftrightarrow a_{ij}$ — матрица факторной структуры размерности $m\times g$ весовых коэффициентов.

В ортогональной факторной модели предполагается, что факторы ортогональны, т. е. $PP^T = E$ — единичная матрица.

Матрица $R = ZZ^T$ корреляций переменных определяется согласно теореме Терстоуна [Иберла, 1980]:

$$R = AA^{\mathrm{T}}.$$

Для поиска факторного отображения можно использовать метод главных компонент. Следующим этапом факторного моделирования является проблема вращения, состоящая в поиске факторной структуры, обладающей лучшей «интерпретабельностью».

Метод косоугольного вращения «облимакс»

В случае косоугольного вращения на результирующую матрицу вращения накладывается следующее ограничение: у матрицы, являющейся результатом произведения транспонированной косоугольной матрицы вращения и исходной косоугольной матрицы вращения, на диагонали должны стоять единицы, тогда как внедиагональные элементы могут быть отличны от нуля.

Т. е. в случае косоугольного факторного вращения конечные факторы могут не быть друг другу ортогональны, геометрически это означает, что угол между косоугольными факторами может отличаться от прямого.

С помощью метода главных факторов было построено прямое решение, т. е. получено факторное отображение:

$$A = \begin{pmatrix} a_{11} & \dots & a_{1g} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mg} \end{pmatrix},$$

где m — число изучаемых параметров, g — число общих факторов.

Вращение заключается в следующей матричной операции:

$$V = A\Lambda,$$

$\underset{m\times g}{V} \leftrightarrow v_{ij}$ — косоугольная факторная структура; $\underset{g\times g}{\Lambda} \leftrightarrow \lambda_{ij}$ — матрица вращения.

«Облимакс» критерий заключается в максимизации следующей функции:

$$K = \frac{\sum_{i=1}^{m}\sum_{p=1}^{g} v_{ip}^4}{\left(\sum_{i=1}^{m}\sum_{p=1}^{g} v_{ip}^2\right)^2}.$$

Максимизация K эквивалентна максимизации эксцесса случайной величины ξ, представленной выборкой v_{ij} и $-v_{ij}$:

$$\gamma = \frac{\frac{2}{mg}\sum_{i=1}^{m}\sum_{j=1}^{g}\left(v_{ij}\right)^4}{\left(\frac{2}{mg}\sum_{i=1}^{m}\sum_{j=1}^{g}\left(v_{ij}\right)^2\right)^2} - 3 = \frac{mg\sum_{i=1}^{m}\sum_{j=1}^{g}\left(v_{ij}\right)^4}{2\left(\sum_{i=1}^{m}\sum_{j=1}^{g}\left(v_{ij}\right)^2\right)^2} - 3.$$

В результате максимизируется доля больших и маленьких (близких к нулю) элементов факторной структуры.

«Облимакс» критерий для двух факторов выглядит следующим образом:

$$K_p = \frac{\sum_{i=1}^{m} v_{ip}^4}{\left(\sum_{i=1}^{m} v_{ip}^2\right)^2} = \max.$$

Вращение осуществляют в плоскости двух фиксированных факторов $\mathbf{P}_i$ и $\mathbf{P}_j$.

$$v_{sp} = a_{si}\lambda_{ip} + a_{sj}\lambda_{jp}, \ p = i, j.$$

Вводят новую переменную $x = \dfrac{\lambda_{jp}}{\lambda_{ip}}$, при этом $\lambda_{ip}^2 + \lambda_{jp}^2 = (\cos\alpha)^2 + (\sin\alpha)^2 = 1.$

Следовательно,

$$\lambda_{ip} = \frac{1}{\sqrt{1+x^2}}; \tag{1}$$

$$\lambda_{jp} = \frac{x}{\sqrt{1+x^2}}; \tag{2}$$

$$v_{sp} = \frac{a_{si}}{\sqrt{1+x^2}} + \frac{a_{sj}x}{\sqrt{1+x^2}} = \frac{a_{si} + a_{sj}x}{\sqrt{1+x^2}}.$$

Пусть $a_s \equiv a_{si}$ и $b_s \equiv a_{sj}$, тогда

$$v_{sp} = \frac{a_s + b_s x}{\sqrt{1+x^2}};$$

$$K_p = \frac{\sum\limits_{s=1}^{m} v_{sp}^4}{\left(\sum\limits_{s=1}^{m} v_{sp}^2\right)^2} = \frac{\sum\limits_{s=1}^{m}\left(\dfrac{a_s + b_s x}{\sqrt{1+x^2}}\right)^4}{\left(\sum\limits_{s=1}^{m}\left(\dfrac{a_s + b_s x}{\sqrt{1+x^2}}\right)^2\right)^2} = \frac{\sum\limits_{s=1}^{m}(a_s + b_s x)^4}{\left(\sum\limits_{s=1}^{m}(a_s + b_s x)^2\right)^2}.$$

Обозначим $E \equiv \sum\limits_{s=1}^{m}(a_s + b_s x)^4$ и $F \equiv \sum\limits_{s=1}^{m}(a_s + b_s x)^2$, тогда

$$K_p = \frac{E}{F^2}.$$

Необходимое условие максимума K_p:

$$K'_p = \frac{E'F^2 - E(2FF')}{F^4} = 0, \ E'F^2 - E(2FF') = 0,$$

$$FE' - 2F'E = 0.$$

Проводя следующие алгебраические преобразования:

$$F \equiv \sum_{s=1}^{m}(a_s + b_s x)^2 = \sum_{s=1}^{m}\left(a_s^{\,2} + 2a_s b_s x + b_s^{\,2} x^2\right),$$

$$E \equiv \sum_{s=1}^{m}(a_s + b_s x)^4 = \sum_{s=1}^{m}\left(a_s^{\,4} + 4a_s^{\,2} b_s^{\,2} x^2 + b_s^{\,4} x^4 + 4a_s^{\,3} b_s x + 2a_s^{\,2} b_s^{\,2} x^2 + 4a_s b_s^{\,3} x^3\right),$$

$$E \equiv \sum_{s=1}^{m}\left(b_s^{\,4} x^4 + 4a_s b_s^{\,3} x^3 + 6a_s^{\,2} b_s^{\,2} x^2 + 4a_s^{\,3} b_s x + a_s^{\,4}\right),$$

$$E' \equiv \sum_{s=1}^{m}\left(4b_s{}^4x^3 + 12a_sb_s{}^3x^2 + 12a_s{}^2b_s{}^2x + 4a_s{}^3b_s\right),$$

$$F' = \sum_{s=1}^{m}\left(2a_sb_s + 2b_s{}^2x\right),$$

$$FE' = \sum_{s=1}^{m}\left(a_s{}^2 + 2a_sb_sx + b_s{}^2x^2\right)\sum_{q=1}^{m}\left(4b_q{}^4x^3 + 12a_qb_q{}^3x^2 + 12a_q{}^2b_q{}^2x + 4a_q{}^3b_q\right) =$$

$$= \sum_{s=1}^{m}a_s{}^2\sum_{q=1}^{m}4b_q{}^4x^3 + \sum_{s=1}^{m}a_s{}^2\sum_{q=1}^{m}12a_qb_q{}^3x^2 + \sum_{s=1}^{m}a_s{}^2\sum_{q=1}^{m}12a_q{}^2b_q{}^2x + \sum_{s=1}^{m}a_s{}^2\sum_{q=1}^{m}4a_q{}^3b_q +$$

$$+ \sum_{s=1}^{m}8a_sb_s\sum_{q=1}^{m}b_q{}^4x^4 + \sum_{s=1}^{m}a_sb_s\sum_{q=1}^{m}24a_qb_q{}^3x^3 + \sum_{s=1}^{m}a_sb_s\sum_{q=1}^{m}24a_q{}^2b_q{}^2x^2 + \sum_{s=1}^{m}a_sb_s\sum_{q=1}^{m}8a_q{}^3b_qx +$$

$$+ \sum_{s=1}^{m}b_s{}^2\sum_{q=1}^{m}4b_q{}^4x^5 + \sum_{s=1}^{m}b_s{}^2\sum_{q=1}^{m}12a_qb_q{}^3x^4 + \sum_{s=1}^{m}b_s{}^2\sum_{q=1}^{m}12a_q{}^2b_q{}^2x^3 + \sum_{s=1}^{m}b_s{}^2\sum_{q=1}^{m}4a_q{}^3b_qx^2,$$

$$2F'E = 2\sum_{s=1}^{m}\left(2a_sb_s + 2b_s{}^2x\right)\sum_{q=1}^{m}\left(b_q{}^4x^4 + 4a_qb_q{}^3x^3 + 6a_q{}^2b_q{}^2x^2 + 4a_q{}^3b_qx + a_q{}^4\right) =$$

$$= \sum_{s=1}^{m}4a_sb_s\sum_{q=1}^{m}b_q{}^4x^4 + \sum_{s=1}^{m}16a_sb_s\sum_{q=1}^{m}a_qb_q{}^3x^3 +$$

$$+ \sum_{s=1}^{m}24a_sb_s\sum_{q=1}^{m}a_q{}^2b_q{}^2x^2 + \sum_{s=1}^{m}16a_sb_s\sum_{q=1}^{m}a_q{}^3b_qx +$$

$$+ \sum_{s=1}^{m}4b_s{}^2\sum_{q=1}^{m}b_q{}^4x^5 + \sum_{s=1}^{m}16b_s{}^2\sum_{q=1}^{m}a_qb_q{}^3x^4 +$$

$$+ \sum_{s=1}^{m}24b_s{}^2\sum_{q=1}^{m}a_q{}^2b_q{}^2x^3 + \sum_{s=1}^{m}16b_s{}^2\sum_{q=1}^{m}a_q{}^3b_qx^2 + \sum_{s=1}^{m}4b_s{}^2\sum_{q=1}^{m}a_q{}^4x.$$

Приходим к уравнению [Харман, 1972]:

$$\alpha_4x^4 + \alpha_3x^3 + \alpha_2x^2 + \alpha_1x + \alpha_0 = 0,$$

где

$$\alpha_4 = \sum_{s=1}^{m}a_sb_s\sum_{q=1}^{m}b_q^4 - \sum_{s=1}^{m}b_s^2\sum_{q=1}^{m}a_qb_q^3,$$

$$\alpha_3 = \sum_{s=1}^{m}a_s^2\sum_{q=1}^{m}b_q^4 + 2\sum_{s=1}^{m}a_sb_s\sum_{q=1}^{m}a_qb_q^3 - 3\sum_{s=1}^{m}b_s^2\sum_{q=1}^{m}a_q^2b_q^2,$$

$$\alpha_2 = 3\sum_{s=1}^{m}a_s^2\sum_{q=1}^{m}a_qb_q^3 - 3\sum_{s=1}^{m}b_s^2\sum_{q=1}^{m}a_q^3b_q,$$

$$\alpha_1 = 3\sum_{s=1}^{m}a_s^2\sum_{q=1}^{m}a_q^2b_q^2 - 2\sum_{s=1}^{m}a_sb_s\sum_{q=1}^{m}a_q^3b_q - \sum_{s=1}^{m}b_s^2\sum_{q=1}^{m}a_q^4,$$

$$\alpha_0 = \sum_{s=1}^{m}a_s^2\sum_{q=1}^{m}a_q^3b_q - \sum_{s=1}^{m}a_sb_s\sum_{q=1}^{m}a_q^4.$$

Это задача нелинейной алгебры на поиск решения трансцендентного уравнения. Находить корни можно, например, методом деления отрезка пополам, методом Ньютона, методом хорд.

Выбираем два корня x_i, x_j, которые доставляют два максимальных значения K_p. По формулам (1) и (2) находим λ_{ip} и λ_{jp}, где каждому $p = i, j$ соответствует свой корень x_i, x_j. Совершаем вращение $v_{sp} = a_{si}\lambda_{ip} + a_{sj}\lambda_{jp}$ $\left(s = 1\ldots m; p = i, j\right)$.

Проблема выбора пар вращения

Экспериментально установлено, что выбор произвольных пар в общем случае приводит к тому, что на диагонали матрицы корреляций между факторами появляются элементы отличные от 1. Для теоретического подтверждения этого факта докажем следующую теорему.

Введем следующие обозначения и условия:

Матрицы:

$$\underset{m\times m}{A} \leftrightarrow \alpha_{ij} = \delta_{ij}^{\{i_1;\ldots;i_k\}} = \begin{cases} a_{ij}, \text{ если } \{\text{i;j}\} \subset \{i_1;\ldots;i_k\}; \\ \delta_{ij}, \text{ если } \{\text{i;j}\} \not\subset \{i_1;\ldots;i_k\}; \end{cases}$$

$$\underset{m\times m}{B} \leftrightarrow \beta_{ij} = \delta_{ij}^{\{j_1;\ldots;j_t\}} = \begin{cases} b_{ij}, \text{ если } \{\text{i;j}\} \subset \{j_1;\ldots;j_t\}; \\ \delta_{ij}, \text{ если } \{\text{i;j}\} \not\subset \{j_1;\ldots;j_t\}. \end{cases}$$

Символ Кронекера: $\delta_{ij} = \begin{cases} 1, \text{ если } i = j; \\ 0, \text{ если } i \neq j. \end{cases}$

При этом для матриц A и B выполняются следующие условия:

$$\sum_{i=1}^{m} \alpha_{ij}^2 = 1, \ \sum_{i=1}^{m} \beta_{ij}^2 = 1 \text{ для } j=1,\ldots,m.$$

$\underset{m\times m}{AB} \leftrightarrow \lambda_{ij} = \sum_{k=1}^{m} \alpha_{ik}\beta_{kj}$ — матрица равная произведению матриц A и B.

Теорема.

Условие $\sum_{i=1}^{m} \lambda_{ij}^2 = 1$ $(j = 1,\ldots,m)$ *выполняется для матрицы произведения AB только в том случае, когда пересечение* $\{i_1;\ldots;i_k\} \cap \{j_1;\ldots;j_t\}$ *содержит не больше одного элемента.*

Доказательство:

Рассмотрим общий случай, когда пересечение $\{i_1;\ldots;i_k\} \cap \{j_1;\ldots;j_t\} = \{s_1;\ldots;s_n\}$ содержит n элементов.

Рассмотрим следующие множества элементов матрицы AB:

1. $j \in \{j_k\}$, $i \in \{i_k\}$

$$AB = \sum_{l=1}^{m} \delta_{il}^{\{i_k\}} \delta_{lj}^{\{j_k\}} = \sum_{l=1}^{k} a_{ii_l} \delta_{i_l j}^{\{j_k\}} = \sum_{l=1}^{n} a_{is_l} b_{s_l j},$$

2. $j \in \{j_k\}$, $i \in \{j_k\} \setminus \{i_k\}$

$$AB = \sum_{l=1}^{m} \delta_{il} \delta_{lj}^{\{j_k\}} = \delta_{ij}^{\{j_k\}} = b_{ij},$$

3. $j \in \{j_k\}$, $i \notin \{i_k; j_k\}$

$$AB = \delta_{ij}^{\{j_k\}} = \delta_{ij},$$

4. $j \in \{i_k\} \setminus \{j_k\}$, $i \in \{i_k\}$

$$AB = \sum_{l=1}^{k} a_{ii_l} \delta_{i_l j}^{\{j_k\}} = \sum_{l=1}^{k} a_{ii_l} \delta_{i_l j} = a_{ij},$$

5. $j \in \{i_k\} \setminus \{j_k\}$, $i \in \{j_k\} \setminus \{i_k\}$

$$AB = \sum_{l=1}^{m} \delta_{il}^{\{i_k\}} \delta_{lj} = \sum_{l=1}^{m} \delta_{il} \delta_{lj} = \delta_{ij},$$

6. $j \in \{i_k\} \setminus \{j_k\}$, $i \notin \{i_k; j_k\}$

$$AB = \delta_{ij},$$

7. $j \notin \{i_k; j_k\}$

$$AB = \sum_{l=1}^{m} \delta_{il}^{\{i_k\}} \delta_{lj} = \delta_{ij}^{\{i_k\}} = \delta_{ij}.$$

Объединение этих множеств элементов составляет все множество элементов матрицы AB, при этом пересечение этих множеств нулевое.

Вычислим условия $\sum_{i=1}^{m} \lambda_{ij}^2 = 1$ для всех столбцов.

Возможны три случая:

1. $j \in \{j_k\}$

$$\sum_{i=1}^{m} \lambda_{ij}^2 = \sum_{i \in \{i_k\}} \left(\sum_{k=1}^{n} a_{is_k} b_{s_k j} \right)^2 + \sum_{i \in \{j_k\} \setminus \{i_k\}} \left(b_{ij} \right)^2 + \sum_{i \notin \{i_k; j_k\}} \left(\delta_{ij} \right)^2.$$

Это выражение тождественно равно 1 только в том случае, если $n \le 1$.

2. $j \in \{i_k\} \setminus \{j_k\}$

$$\sum_{i=1}^{m} \lambda_{ij}^2 = \sum_{i \in \{i_k\}} \left(a_{ij} \right)^2 + \sum_{i \in \{j_k\} \setminus \{i_k\}} \left(\delta_{ij} \right)^2 + \sum_{i \notin \{i_k; j_k\}} \left(\delta_{ij} \right)^2 = 1.$$

3. $j \notin \{i_k; j_k\}$

$$\sum_{i=1}^{m} \lambda_{ij}^2 = \sum_{i=1}^{m} \left(\delta_{ij} \right)^2 = 1.$$

Теорема доказана. ■

Авторский алгоритм выбора пар факторов

Из формулировки теоремы непосредственно следует, что выбирая новую пару факторов подлежащих косоугольному вращению, необходимо, чтобы хотя бы один из новых факторов раньше не участвовал во вращении. В этом случае конечная матрица вращения получается всегда правильной, т. е. корреляция любого фактора с самим собою всегда равна 1.

В качестве алгоритма выбора пар факторов подлежащих вращению предлагается использовать следующий:

Пусть V — множество индексов факторов $(1, 2, \ldots, g)$.

U — пустое множество.

1. Выбрать пару (i, j), в которой хотя бы один из индексов принадлежит множеству $V \setminus U$. При этом выбирается пара, которая соответствует максимальному значению главного критерия $K = \frac{\sum_{i=1}^{m}\sum_{p=1}^{g} v_{ip}^4}{\left(\sum_{i=1}^{m}\sum_{p=1}^{g} v_{ip}^2\right)^2} \to \max$.
2. Индексы i, j добавляются в множество U.
3. Пока $V \setminus U$ не пусто, совершается переход к пункту 1.

Заключение

Рассмотрена проблема допустимой последовательности выбора пар факторов.

Для новой пары факторов, подлежащих косоугольному вращению, необходимо, чтобы хотя бы один из факторов раньше не участвовал во вращении. Только в этом случае конечная матрица вращения получается всегда правильной, т. е. корреляция любого фактора с самим собою всегда равна 1.

Из этого класса допустимых последовательностей выбора пар факторов предложена оригинальная последовательность выбора пар факторов, приводящая к наилучшей факторной структуре по критерию «облимакс». Для этого новую пару факторов выбирают таким образом, чтобы она приносила максимум критерию «облимакс» для всей факторной структуры.

Список литературы

Иберла К. Факторный анализ / пер. с нем. В. М. Ивановой; Предисл. А. М. Дуброва. — М.: Статистика, 1980.

Харман Г. Современный факторный анализ / пер. с англ. В. Я. Лумельского; Научное редактирование и вступительная статья Э. М. Бравермана. — М.: Статистика, 1972.