Э. М. Браверман, Методы экстремальной группировки параметров и задача выделения существенных факторов, *Автомат. и телемех.*, 1970, выпуск 1, 123–132

УДК 62-50

МЕТОДЫ ЭКСТРЕМАЛЬНОЙ ГРУППИРОВКИ ПАРАМЕТРОВ И ЗАДАЧА ВЫДЕЛЕНИЯ СУЩЕСТВЕННЫХ ФАКТОРОВ

Э. М. БРАВЕРМАН

(Москва)

Формируется задача разбиения на группы совокупности взаимосвязанных случайных величин по степени связи между ними, определяемой матрицей ковариации (или корреляции), вводятся функционалы, экстремизация которых обеспечивает наилучшее разбиение величин на группы, и предлагаются алгоритмы, строящие такое разбиение. Одновременно для каждой группы случайных величин эти алгоритмы определяют случайную величину, которую естественно интерпретировать как существенный фактор. Предложенные методы сравниваются с методами факторного анализа.

1. Постановка задачи

При обработке результатов экспериментальных исследований в технике, экономике, социологии, психологии и в других областях часто приходится иметь дело с большим числом измеряемых или определяемых как-либо иначе случайных величин и возникает задача сокращения описания измеряемой информации с целью получения легко интерпретируемых результатов. Решение этой задачи затруднено тем, что, как правило, измеряемые параметры лишь косвенно отражают существенные факторы, характеризующие исследуемые объекты, так что возникает необходимость в разработке специальных методов для извлечения требуемой информации об этих факторах.

Так, в процессе изучения химических реакторов могут быть измерены температуры и давление в различных точках, расходы реагентов и т. п., тогда как существенными факторами, характеризующими состояние реактора, являются качество катализатора, интенсивность массообмена, граница жидкой и паровой фаз и т. д. В психологии «измеряемыми параметрами» являются реакции людей на различные тесты, а существенными факторами такие непосредственно не измеряемые характеристики субъекта, как «степень интеллектуальности», «работоспособность», «тип нервной системы» и т. д. В социологических обследованиях «измеряемыми параметрами» являются ответы на различные вопросы анкет или реакции людей в условиях социологических экспериментов, а существенным фактором — отношение людей к той или иной социальной или демографической группе.

В этих и других многочисленных параметрах такого рода количество измеряемых параметров несоизмеримо больше числа существенных факторов, и задача прежде всего заключается в том, чтобы выявить эти факторы. Коль скоро такие факторы выявлены, естественно возникает задача сокращения числа измеряемых параметров (число датчиков в автоматическом контроле, число тестов в психологических исследованиях, число вопросов в социологических исследованиях и т. д.), при котором не теряется информация, нужная для определения существенных факторов.

Решение обеих указанных выше задач может опираться на следующее обстоятельство, как правило, имеющее место в приложениях, примеры которых были ранее приведены. В этих приложениях изменение какого-либо фактора сказывается неодинаково на всех измеряемых величинах и поэтому среди измеряемых величин могут быть выделены группы, «особо остро» реагирующие на каждый из факторов порознь. Из сказанного вытекает следующее естественное предположение: измеряемые параметры наиболее сильно коррелируют друг с другом в том случае, когда они наиболее сильно зависят от одного и того же фактора. Приняв это предположение как исходную гипотезу, можно строить разбиение всех измеряемых параметров на такие группы, что параметры, принадлежащие одной группе, в некотором смысле сильно коррелируют между собой, а параметры, принадлежащие разным группам, коррелированы относительно слабо. Задачу такого рода будем называть задачей группировки параметров. В настоящее время создано несколько эвристических алгоритмов, решающих задачу группировки параметров (метод плеяд [1, 2], метод B-коэффициентов [3]), однако точной постановки этой задачи не существует.

Если группировка построена, то в каждой группе можно построить случайную величину, которая в некотором смысле наиболее сильно коррелирует со всеми параметрами, входящими в данную группу. Тогда естественно интерпретировать такую случайную величину, как искомый фактор, существенно влияющий на все параметры данной группы. С другой стороны, построенное разбиение и найденные факторы позволяют сократить число измеряемых параметров, сохранив лишь один или несколько параметров из каждой группы, достаточных для полного описания факторов.

В данной работе в основу группировки и выделения факторов положен подход, связанный с экстремизацией некоторого функционала, зависящего как от способа группировки, так и от выбора факторов. Этот функционал выбирается так, что экстремизация его (как по разбиению, так и по выбору факторов) интуитивно соответствует описанной выше содержательной задаче. Разбиение, экстремизирующее этот функционал, будем называть экстремальной группировкой параметров. В силу описанных выше обстоятельств нахождение экстремальной группировки одновременно решает вопрос о выборе факторов.

В настоящей работе вводятся в рассмотрение два функционала указанного выше типа, предлагаются алгоритмы их экстремизации, т. е. определения экстремальной группировки, и доказывается сходимость этих алгоритмов.

Вопрос о том, в какой мере построенные таким образом разбиения адекватны реальным прикладным задачам, не может быть решен в рамках теории, а должны решаться специальными экспериментами. Подобные эксперименты проведены, и они являются предметом специальной публикации [7].

Одна из двух задач, одновременно решенных в этой работе, именно задача о выявлении факторов (без группировки параметров либо в условиях, когда эта группировка предопределена заранее), рассматривается факторным анализом (см., например, [3—6]). В последнем разделе настоящей работы кратко описываются постановки задач и основные алгоритмы определения факторов в факторном анализе и проводится сопоставление идей и методов факторного анализа с идеями и методами экстремальной группировки.

2. Первый алгоритм экстремальной группировки

Будем в дальнейшем коэффициент корреляции (или ковариации) $\rho_{x,y}$ двух случайных величин x и y обозначать так: $\rho_{x,y} = (x, y)$, подчеркивая этим обозначением тот факт, что коэффициент корреляции может пониматься как скалярное произведение случайных величин x и y. Для дисперсии $\rho_{x,x}$ случайной величины x будем применять обозначение

$$\rho_{x,x} = (x, x) = x^2.$$

Пусть множество параметров (случайных величин) $x_1, x_2, \ldots, x_n$ разбито на непересекающиеся группы $A_1, A_2, \ldots, A_k$ и заданы случайные величины $f_1, f_2, \ldots, f_k$ такие, что $f_1^2 = f_2^2 = \ldots = f_k^2 = 1$, которые будем называть факторами. Введем в рассмотрение функционал

$$J_1 = \sum_{x_i \in A_1} (x_i, f_1)^2 + \sum_{x_i \in A_2} (x_i, f_2)^2 + \ldots + \sum_{x_i \in A_k} (x_i, f_k)^2. \qquad (1)$$

Первый алгоритм экстремальной группировки параметров решает задачу максимизации этого функционала как по разбиению параметров на множества $A_1, \ldots, A_k$, так и по выбору случайных величин $f_1, \ldots, f_k$, $f_l^2 = 1$ $(l = 1, 2, \ldots, k)$.

Максимизация функционала (1) соответствует интуитивному требованию такого разбиения параметров, когда в одну группу попадают наиболее «близкие» между собой параметры. Действительно, при максимизации функционала (1) для каждого фиксированного набора случайных величин $f_1, \ldots, f_k$ в одну l-ю группу будут попадать такие параметры, которые наиболее «близки» к величине f_l; в то же время среди всех возможных наборов случайных величин $f_1, \ldots, f_k$ будет отбираться такой набор, что каждая из величин f_l в среднем наиболее «близка» ко всем параметрам из своей группы.

Если заданы группы параметров $A_1, \ldots, A_k$, то максимум функционала J_1 может быть получен, если в качестве факторов $f_1, \ldots, f_k$ выбрать такие случайные величины, что каждая случайная величина f_l $(l = 1, 2, \ldots, k)$ удовлетворяет условию

$$\max_{f_l} \sum_{x_i \in A_l} (x_i, f_l)^2, \quad f_l^2 = 1. \qquad (2)$$

Фактор f_l, удовлетворяющий условию (2) при фиксированном множестве параметров A_l, находится по формуле

$$f_l = \frac{\sum\limits_{x_i \in A_l} \alpha_i x_i}{\sqrt{\left(\sum\limits_{x_i \in A_l} \alpha_i x_i\right)^2}} = \frac{\sum\limits_{x_i \in A_l} \alpha_i x_i}{\sqrt{\sum\limits_{x_i x_j \in A_l} \alpha_i \alpha_j (x_i, x_j)}}, \qquad (3)$$

где α_i — компоненты собственного вектора матрицы $R_l = \{(x_i, x_j)\}$, $x_i, x_j \in A_l$, соответствующего ее наибольшему собственному значению. С другой стороны, если величины $f_1, \ldots, f_k$ заданы, то разбиение параметров на группы $A_1, \ldots, A_k$, обеспечивающие максимум функционала J_1, должно удовлетворять условию: для каждого $x_i \in A_l$

$$(x_i, f_l)^2 \geqslant (x_i, f_q)^2 \quad (q = 1, 2, \ldots, k), \qquad (4)$$

так как в противном случае функционал J_1 можно было бы увеличить, перебросив параметр x_i из группы A_l в ту группу A_q, для которой соотношение (4) не выполнено. Соотношения (2) и (4) в совокупности являются необходимыми условиями максимума функционала J_1.

Можно предложить следующий итерационный алгоритм, определяющий одновременно группы $A_1, \ldots, A_k$ и факторы $f_1, \ldots, f_k$, удовлетворяющие этим условиям.

Пусть на p-м шаге итерации построено разбиение параметров на группы $A_1^p, \ldots, A_k^p$. Для каждой такой группы параметров строят факторы f_l^p по формуле (3) и новое, $(p+1)$-е, разбиение параметров $A_1^{p+1}, \ldots, A_k^{p+1}$ в соответствии с правилом: параметр x_i относится к группе A_l^{p+1}, если

$$(x_i, f_l^p)^2 \geqslant (x_i, f_q^p)^2 \quad (q = 1, 2, \ldots, k). \tag{5}$$

В том случае, когда существуют два или более факторов и такой параметр x_i, что для этих факторов и этого параметра в (5) имеет место равенство, параметр x_i относится к одной из соответствующих групп произвольно. Предложенный выше алгоритм сходится к максимуму (возможно, локальному) функционала J_1, поскольку, каковы бы ни были факторы $f_1^{p-1}, \ldots, f_k^{p-1}$, на каждом шаге итерации функционал J_1 не убывает.

3. Второй алгоритм экстремальной группировки

Для построения второго алгоритма экстремальной группировки для каждого разбиения параметров на группы $A_1, \ldots, A_k$ и каждого набора факторов $f_1, \ldots, f_k$ определим функционал

$$J_2 = \sum_{x_i \in A_1} |(x_i, f_1)| + \sum_{x_i \in A_2} |(x_i, f_2)| + \ldots + \sum_{x_i \in A_k} |(x_i, f_k)|. \tag{6}$$

Этот функционал в содержательном смысле похож на функционал (1), и максимизация его также соответствует основному интуитивному требованию к характеру разбиения параметров.

Докажем следующую теорему, устанавливающую необходимые и достаточные условия максимума функционала J_2.

Теорема. Необходимыми и достаточными условиями максимума функционала J_2 является следующая совокупность условий.

1. Разбиение параметров на группы $A_1, \ldots, A_k$ таково, что функционал

$$J_3 = \sum_{l=1}^{k} \sqrt{\Big(\sum_{x_i \in A_l} \sigma_i x_i\Big)^2}, \tag{7}$$

где σ_i может принимать лишь значения $+1$ или -1, достигает максимума как по разбиению на группы, так и по значениям коэффициентов σ_i.

2. Факторы f_l определяются соотношениями

$$f_l = \sum_{i \in A_l} \sigma_i x_i \Big/ \sqrt{\sum_{x_i, x_j \in A_l} \delta_i \sigma_j (x_i, x_j)}. \tag{8}$$

Доказательство теоремы приводится в Приложении. Установленная теорема показывает, что задача максимизации функционала J_2 может быть заменена задачей максимизации функционала J_3. Эта задача более проста, так как при фиксированном разбиении на группы функционал J_3 достигает максимума тогда, когда для каждого l соответствующие коэффициенты σ_i таковы, что они максимизируют величину

$$\Big(\sum_{x_i \in A_l} \sigma_i x_i\Big)^2. \tag{9}$$

Ниже предлагается рекуррентная процедура максимизации функционала J_3. В этой процедуре циклически перебираются параметры $x_1, \ldots, x_n$, на каждом шаге принимается решение об отнесении очередного параметра x_j к одной из групп $A_1, \ldots, A_k$ и определяется знак * σ_j.

Будем указывать индексом сверху значение величин и множеств, выстраиваемых к соответствующему шагу алгоритма. Пусть к s-му шагу алгоритма построены разбиения параметров на группы $\overset{s}{A_1}, \ldots, \overset{s}{A_k}$ и коэффициенты $\overset{s}{\sigma_1}, \ldots, \overset{s}{\sigma_n}$, которые могут принимать лишь значения $+1$ или -1, и пусть на этом шаге рассматривается параметр $x_i \in A_{l^*}^s$. Тогда строятся k коэффициентов $\overset{s+1}{\sigma_{i,l}}$ $(l = 1, 2, \ldots, k)$ по формуле

$$\overset{s+1}{\sigma_{i,l}} = \operatorname{sign} \sum_{x_j \in A_l} \overset{s}{\sigma_j}(x_i, x_j) \quad (j \neq i) \tag{10}$$

и для всех l вычисляются разности

$$\Delta_l^{s+1} = \sqrt{\Big(\sum_{\substack{x_j \in \overset{s}{A_l} \\ j \neq i}} \sigma_j^s x_j + \sigma_{i,\,l}^{s+1} x_i\Big)^2} - \sqrt{\Big(\sum_{\substack{x_j \in \overset{s}{A_l} \\ j \neq i}} \sigma_j^s x_j\Big)^2} =$$

$$= \sqrt{\sum_{\substack{x_p, x_q \in \overset{s}{A_l} \\ p,\, q \neq i}} \sigma_p^s \sigma_q^s (x_p, x_q) + 2\sigma_{i,\,l}^{s+1} \sum_{\substack{x_j \in \overset{s}{A_l} \\ j \neq i}} \sigma_j^s (x_i, x_j) + x_i^2} -$$

$$- \sqrt{\sum_{\substack{x_p, x_q \in \overset{s}{A_{l^*}} \\ p,\, q \neq i}} \sigma_p^s \sigma_q^s (x_p, x_q)}. \tag{11}$$

Далее выбирается такой номер ** $l = l^{**}$, что

$$\overset{s+1}{\Delta_{l^{**}}} = \max_{1 \leqslant l \leqslant k} \overset{s+1}{\Delta_l} \tag{12}$$

и параметр x_i исключается из группы $A_{l^*}^s$ и присоединяется к группе $A_{l^{**}}$; остальные группы параметров на этом шаге не перестраиваются. В результате оказывается построенным новое разбиение параметров ***

$$\overset{s+1}{A_1}, \ldots, \overset{s+1}{A_k}.$$

Новые значения коэффициентов σ_j^{s+1} определяются по формулам

$$\overset{s+1}{\sigma_j} = \overset{s}{\sigma_j} \ (j \neq i), \ \overset{s+1}{\sigma_i} = \overset{s+1}{\sigma_{i,l^{**}}}. \tag{13}$$

После этого выполняется следующий, $(s+1)$-й шаг алгоритма, на котором рассматривается параметр x_{i+1}, если $i \neq n$, и параметр x_1, если $i = n$.

Если на очередном цикле при рассмотрении всех параметров от первого до последнего как разбиение параметров на группы, так и значения всех коэффициентов не изменились, то процедура заканчивается, а полученное разбиение и значения параметров считаются окончательными. Этот алгоритм гарантирует получение локального максимума функционала J_3. Поэтому при его практическом использовании целесообразно по-

* При решении задачи раздела 2 такая «рекуррентная» структура алгоритма, когда при рассмотрении очередного параметра перестраиваются как сами группы, так и коэффициенты разложения факторов по параметрам, по-видимому, нецелесообразна. Процесс построения коэффициентов как компонент собственного вектора, соответствующего наибольшему собственному значению некоторой матрицы, является трудоемким процессом, и желательно при работе алгоритма применять его возможно реже.

** Если существует несколько таких номеров l, для которых разность Δ_l^{s+1} принимает наибольшее значение, то в качестве l^{**} выбирается один из этих номеров произвольно.

*** В частности, это разбиение может совпадать с предыдущим, если оказывается, что $l^{**} = l^*$.

вторить описанную процедуру несколько раз при различных начальных разбиениях $A_1', \ldots, A_k'$ и начальных значениях коэффициентов $\sigma_1', \ldots, \sigma_n'$ и принять в качестве окончательного разбиения и окончательных значений коэффициентов те, которые обеспечивают максимум функционала J_3.

Для того чтобы доказать, что предложенный алгоритм сходится к максимуму, возможно, локальному, достаточно установить, что на каждом шаге алгоритма функционал J_3 не убывает, т. е.

$$J_3^{s+1} \geqslant J_3^s. \tag{14}$$

Пусть к s-му шагу алгоритма рассматриваемый на этом шаге параметр x_i принадлежал $A_{l^*}^s$. В результате выполнения s-го шага возможны два случая.

1. Параметр x_i остался в той же группе параметров, так что разбиение параметров на группы не изменилось и

$$A_p^{s+1} = A_p^s \quad (p = 1, 2, \ldots, k),$$

но значение коэффициента σ_i^s могло измениться.

2. Параметр x_i перешел из группы с номером l^* в группу с номером $l^{**} \neq l^*$, так что

$$A_p^{s+1} = A_p^s \quad (p = 1, \ldots, k;\ p \neq l^*,\ p \neq l^{**}),$$

$$A_{l^*}^{s+1} \neq A_{l^*}^s,$$

$$A_{l^{**}}^{s+1} \neq A_{l^{**}}^s.$$

Рассмотрим эти случаи порознь.

Случай 1. Имеем

$$J_3^{s+1} = J_3^s + \sqrt{\Big(\sum_{x_j \in A_{l^*}^{s+1}} \sigma_j^{s+1} x_j\Big)^2} - \sqrt{\Big(\sum_{x_j \in A_{l^*}^s} \sigma_j^s x_j\Big)^2} =$$

$$= J_3^s + \sqrt{\Big(\sum_{\substack{x_j \in A_{l^*}^s \\ j \neq i}} \sigma_j^s x_j\Big)^2 + 2\sigma_i^{s+1} \sum_{\substack{x_j \in A_{l^*}^s \\ j \neq i}} \sigma_j^s (x_i, x_j) + x_i^2} -$$

$$- \sqrt{\Big(\sum_{\substack{x_j \in A_{l^*}^s \\ j \neq i}} \sigma_j^s x_j\Big)^2 + 2\sigma_i^s \sum_{\substack{x_j \in A_{l^*}^s \\ j \neq i}} \sigma_j^s (x_i, x_j) + x_i^2}. \tag{15}$$

В силу (10) и (13)

$$\sigma_i^{s+1} \sum_{\substack{x_j \in A_{l^*}^s \\ j \neq i}} \sigma_j^s (x_i, x_j) \geqslant \sigma_i^s \sum_{\substack{x_j \in A_{l^*}^s \\ j \neq i}} \sigma_j^s (x_j, x_i),$$

и поэтому из (15) следует соотношение (14).

Случай 2. Имеем

$$J_3^{s+1} = J_3^s + \sqrt{\Big(\sum_{x_j \in A_{l^*}^{s+1}} \sigma_j^{s+1} x_j\Big)^2} - \sqrt{\Big(\sum_{x_j \in A_{l^*}^s} \sigma_j^s x_j\Big)^2} +$$

$$+ \sqrt{\Big(\sum_{x_j \in A_{l^{**}}^{s+1}} \sigma_j^{s+1} x_j\Big)^2} - \sqrt{\Big(\sum_{x_j \in A_{l^{**}}^s} \sigma_j^s x_j\Big)^2} =$$

$$= J_3^s + \Big(\sqrt{\Big(\sum_{x_j \in A_{l^*}^s} \sigma_j^s x_j + \sigma_{i,\, l^*}^{s+1} x_i\Big)^2} - \sqrt{\Big(\sum_{x_j \in A_{l^*}^s} \sigma_j^s x_j\Big)^2} + (\Delta_{l^{**}}^{s+1} - \Delta_{l^*}^s).$$

В этом выражении первая скобка в правой части больше или равна нулю в силу (10), а вторая скобка больше или равна нулю по определению (12) номера l^{**}. И, следовательно, в случае 2 соотношение (14) также имеет место.

Таким образом, установлено, что при работе алгоритма имеет место соотношение (14), которое гарантирует сходимость предлагаемого алгоритма.

4. Факторный анализ и сопоставление его с методами экстремальной группировки

Факторный анализ [3—6] представляет собой совокупность моделей и методов, имеющих целью ввести в рассмотрение относительно небольшое число случайных величин $f_1, f_2, \ldots, f_k$ $(k < n)$ таких, что исходные случайные величины $x_1, x_2, \ldots, x_n$ можно представить в виде

$$x_i = \sum_{s=1}^{k} l_{is} f_s + l_i \qquad (i = 1, 2, \ldots, n), \tag{16}$$

где l_{is} — некоторые числа, называемые факторными нагрузками, а l_i — остаточные случайные величины, «малые» в том или ином смысле. Различные модели факторного анализа отличаются тем, какие гипотезы выдвигаются относительно случайных величин $f_1, \ldots, f_k$, и тем, в каком смысле понимается малость величин l_i. В приложениях случайным величинам $f_1, \ldots, f_k$ придают смысл существенных характеристик или факторов, определяющих значения переменных $x_1, \ldots, x_n$.

Общие схемы факторного анализа не учитывают основное для настоящей работы и достаточно распространенное обстоятельство — то, что параметры могут быть разбиты на группы так, что в каждую группу включаются параметры, «сильно коррелирующие» между собой и «слабо коррелирующие» с параметрами, включенными в другие группы, и что существование таких групп может означать существование некоторого фактора, определяющего в основном поведение каждого параметра из группы.

Приведем здесь некоторые используемые в настоящее время методы факторного анализа, знание которых существенно для понимания данной работы *.

1. Метод главных компонент [5]. В этом методе первый фактор f_1 и нагрузки l_{i1} находятся из условия минимума функционала

$$\sum_{i=1}^{n} (x_i - l_{i1} f_1)^2, \quad |(f_1, f_1)| = f_1^2 = 1. \tag{17}$$

Легко убедиться, что минимум функционала (17) достигается при $l_{i1} = (x_i, f_1)$, а f_1 находится из условия максимума функционала

$$\Sigma (x_i, f_1)^2, \ f_1^2 = 1. \tag{18}$$

Для построения следующего фактора f_2 рассматриваются случайные величины x_i^{II}, остающиеся после учета влияния первого фактора:

$$x_i^{\mathrm{II}} = x_i - (x_i, f_1) f_1 \qquad (i = 1, 2, \ldots, n).$$

Строится матрица ковариации этих случайных величин и находится аналогичным образом следующий фактор f_2. Этот процесс продолжается до тех пор, пока не будет построено желаемое число факторов.

* Существенной проблемой в каждой модели факторного анализа помимо проблемы определения факторов является проблема определения факторных нагрузок l_{is} с тем, чтобы в определенном смысле минимизировать «остатки» l_i. В настоящей работе на этой проблеме останавливаться не будем.

Сравнение формул (18) и (2) показывает, что первый алгоритм экстремальной группировки для каждой группы параметров строит фактор, имеющий смысл первого фактора для этой группы в смысле метода главных компонент.

2. Центроидный метод [3—6]. В этом методе фактор f_1 ищется в виде

$$f = \sum_{i=1}^{n} \sigma_i x_i, \tag{19}$$

где σ_i — числа, принимающие лишь два значения —1 или +1 и выбираемые таким образом, чтобы максимизировать величину

$$f_1^2 = \sum_{i,\, j=1}^{n} \sigma_i \sigma_j \rho_{ij} = \left(\sum_{i=1}^{n} \sigma_i x_i \right)^2. \tag{20}$$

После определения факторных нагрузок* l_{i1} рассматриваются случайные величины

$$x_i^{\mathrm{II}} = x_i - l_{i1} f_1,$$

строится следующий фактор и т. д.

Сравнение выражений (14) и (15) с выражениями (9) и (8) показывает, что максимизация функционала J_2 приводит к построению для каждой группы параметров фактора, отличающегося на некоторый множитель от первого фактора, который был бы построен для этой группы параметров центроидным методом, если бы соответствующее разбиение параметров было известно заранее.

3. Би-факторный анализ [3, 6] является одной из основных используемых в настоящее время моделей факторного анализа. В этой модели предполагается, что параметры $x_1, \ldots, x_n$ заранее разбиваются на k групп, например $\{x_1, \ldots, x_{n_1}\}$, $\{x_{n_1+1}, \ldots, x_{n_2}\}, \ldots,$ $\{x_{n_{k-1}+1}, \ldots x_{n_k}\}$, и определяются $k+1$ факторов $f_0, f_1, \ldots, f_k$ и соответствующие факторные нагрузки l_{i0}, l_{is} $(i = 1, \ldots, n;\ s = 1, \ldots, k)$ так, чтобы величину x_i, принадлежащую s-й группе, можно было представить в виде

$$x_i = l_{i0} f_0 + l_{is} f_s + l_i. \tag{21}$$

Фактор f_0, называемый общим фактором, находится каким-либо из указанных выше методов с учетом всех переменных x_i $(i = 1, \ldots, n)$, а фактор f_s, называемый s-м групповым фактором, находится с учетом лишь переменных x_i $(i = n_{s-1} + 1, \ldots, n_s)$.

Модель би-факторного анализа представляет собой попытку учесть наличие факторов, влияющих лишь на группу параметров, однако, как указывалось выше, в этой модели построение таких групп не производится — они предполагаются заданными заранее.

Используя результаты предыдущего раздела, можно предложить следующую постановку задачи определения общего и групповых факторов и разбиение на группы, не заданные заранее. Факторы $f_0, f_1, \ldots, f_k$ и разбиение параметров на группы $A_1, \ldots, A_k$ должны быть таковы, чтобы максимизировать функционал

$$J_4 = \sum_{i=1} |(x_i, f_0)| + \sum_{x_i \in A_1} |(x_i, f_1)| + \sum_{x_i \in A_2} |(x_i, f_2)| + \ldots + \sum_{x_i \in A_k} |(x_i, f_k)|. \tag{22}$$

Легко видеть, что выбор общего фактора f_0 в поставленной таким образом задаче может производиться независимо от выбора разбиения и

* См. предыдущую сноску.

групповых факторов $f_1, \dots, f_k$. При этом общий фактор f_0 является для множества параметров $x_i, \dots, x_n$ первым фактором, выстраиваемым центроидным методом, а для получения разбиения на группы $A_1, \dots, A_k$ и факторов $f_1, \dots, f_k$ может быть использован алгоритм раздела 3 настоящей работы.

ПРИЛОЖЕНИЕ

Доказательство теоремы

Покажем сначала, что в точке максимума фактор f_l имеет вид (8). С этой целью проварьируем функционал J_2 по f_l при условии

$$f_l^2 = 1 \tag{П.1}$$

и приравняем вариацию функционала к нулю, в результате получим уравнение

$$\sum_{x_l \in A_l} x_i \operatorname{sign}(x_i, f_l) = 2\lambda_l f_l, \tag{П.2}$$

где λ_l — множитель Лагранжа, определяемый условием (П.1). Полагая

$$\sigma_i = \operatorname{sign}(x_i, f_l) \tag{П.3}$$

и учитывая условие нормировки (П.1), убеждаемся в справедливости соотношения (8).

Покажем теперь, что если f_l имеет вид (8), то при любом наборе коэффициентов $\sigma_i = \pm 1$ и любом разбиении параметров на группы имеет место соотношение

$$J_2 \geqslant J_3, \tag{П.4}$$

если же J_2 достигает максимума, то

$$J_2 = J_3. \tag{П.5}$$

Поскольку из соотношений (П.4) и (П.5) доказательство теоремы следует немедленно, то для завершения доказательства остается установить справедливость этих соотношений. Покажем сначала, что имеет место соотношение (П.4). С этой целью подставим (8) в (6):

$$J_2 = \sum_{x_i \in A_1} |(x_i, f_1)| + \sum_{x_i \in A_2} |(x_i, f_2)| + \dots + \sum_{x_i \in Ak} |(x_i, f_k)| =$$

$$= \sum_{x_i \in A_1} \frac{\left|\left(x_i, \sum\limits_{x_j \in A_1} \sigma_j x_j\right)\right|}{\sqrt{\sum\limits_{x_p, x_s \in A_1} \sigma_p \sigma_s (x_p, x_s)}} + \sum_{x_i \in A_2} \frac{\left|\left(x_i, \sum\limits_{x_j \in A_2} \sigma_j x_j\right)\right|}{\sqrt{\sum\limits_{x_p, x_s \in A_2} \sigma_p \sigma_s (x_p, x_s)}} + \dots$$

$$\dots + \sum_{x_i \in A_k} \frac{\left|\left(x_i, \sum\limits_{x_j \in A_k} \sigma_j x_j\right)\right|}{\sqrt{\sum\limits_{x_p, x_s \in A_k} \sigma_p \sigma_s (x_p, x_s)}}. \tag{П.6}$$

Заметим, что для любого σ_i

$$\left|\left(x_i, \sum_{x_j \in A_l} \sigma_j x_j\right)\right| \geqslant \left(\sigma_i x_i, \sum_{x_j \in A_l} \sigma_j x_j\right) = \sum_{x_j \in A_l} \sigma_i \sigma_j (x_i, x_j) \tag{П.7}$$

Поэтому из (П.6) имеем

$$J_2 \geqslant \sum_{x_i \in A_1} \frac{\sum\limits_{x_j \in A} \sigma_i \sigma_j (x_i, x_j)}{\sqrt{\sum\limits_{x_p, x_s \in A_1} \sigma_p \sigma_s (x_p, x_s)}} + \dots + \sum_{x_i \in A_k} \frac{\sum\limits_{x_j \in A_k} \sigma_i \sigma_j (x_i, x_j)}{\sqrt{\sum\limits_{x_p x_s \in A_k} \sigma_p \sigma_s (x_p, x_s)}} =$$

$$= \sqrt{\sum_{x_i, x_j \in A_1} \sigma_i \sigma_j (x_i, x_j)} + \dots + \sqrt{\sum_{x_i, x_j \in A_k} \sigma_i \sigma_j (x_i, x_j)} = J_3. \tag{П.8}$$

Таким образом, неравенство (П.4) доказано. Для того чтобы доказать соотношение (П.5), следует показать, что если J_2 достигает максимума и $x_i \in A_l$, то

$$\sigma_i = \operatorname{sign} \sum_{x_j \in A_l}{}' \sigma_j (x_i, x_j). \tag{П.9}$$

Действительно, если соотношение (П.9) выполнено, то в (П.7) имеет место точное равенство и, следовательно, имеет место точное равенство (П.8). Для того чтобы установить (П.9), заметим сначала, что если J_2 достигает максимума, то для всех l выполнено соотношение (П.2), причем

$$\lambda_l = \frac{1}{2} J_2 > 0. \tag{П.10}$$

Соотношение (П.10) легко установить, умножив обе части уравнения (П.2) на f_l и воспользовавшись соотношением

$$z \operatorname{sign} z = |z|.$$

Умножив теперь (П.2) на некоторое $x_j \in A_l$ и воспользовавшись обозначением (П.3) и положительностью λ_l, убеждаемся в справедливости (П.9). Таким образом, доказано соотношение (П.5), и следовательно, теорема полностью доказана.

Поступила в редакцию
20 марта 1969 г.

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

1. Терентьев П. В. Метод корреляционных плеяд. Вестник ЛГУ, № 9, 1959.
2. Выханду Л. К. Об исследовании многопризнаковых биологических систем. В сб. «Применение математических методов в биологии», III. Изд-во ЛГУ, 1964.
3. Holzinger K., Harman H. Factor analysis. Univ. Chicago Press, 1941.
4. Лоули Д., Максвелл А. Факторный анализ как статистический метод. «Мир», 1967.
5. Harman H. Modern factor analysis. Univ. Chicago Press, 1960.
6. Horst P. Factor analysis of data matrices. Holt, Rinehart and Winston, New York, 1965.
7. Лумельский В. Я. Группировка параметров на основе квадратной матрицы связи. Автоматика и телемеханика, № 1, 1970.

METHODS OF EXTREMAL GROUPING OF PARAMETERS AND PROBLEM OF APPORTIONMENT OF ESSENTIAL FACTORS

E. M. BRAVERMAN

In the article presented there is formulated the problem of the partition into groups of the ensemble of interconnected random values in accordance with the connection among them, the ensemble defined by the matrix of covariance (or correlation), there are introduced the functionals the extremization of which provides for the best partition of the values into groups; the algorithms consructing such a partition is suggested. These algorithms for each group of random values simultaneously construct a random value which it is quite natural to interprete as an essential factor. The methods suggested are compared with the methods of component analysis.