Impact Factor:

ISRA (India) = **1.344**	**SIS** (USA) = **0.912**	**ICV** (Poland) = **6.630**
ISI (Dubai, UAE) = **0.829**	**РИНЦ** (Russia) = **0.207**	**PIF** (India) = **1.940**
GIF (Australia) = **0.564**	**ESJI** (KZ) = **4.102**	**IBI** (India) = **4.260**
JIF = **1.500**	**SJIF** (Morocco) = **2.031**	

SOI: 1.1/TAS DOI: 10.15863/TAS

International Scientific Journal
Theoretical & Applied Science

p-ISSN: 2308-4944 (print) **e-ISSN:** 2409-0085 (online)

Year: 2018 **Issue:** 04 **Volume:** 60

Published: 30.04.2018 http://T-Science.org

S. U. Zhanatauov
candidate of physics and mathematical sciences,
Department «Information technologies, mathematics, physics»,
Associate professor,
Noncommercial joint-stock company "Kazakh national agrarian university"
Kazakhstan
sapagtu@mail.ru

SECTION 2. Applied mathematics. Mathematical modeling.

INVERSE MODEL OF MULTIPLE LINEAR REGRESSION ANALYSIS

Abstract: *The inverse model of multiple linear regression analysis (IM MLRA) is worked up in the article. Theorem is proved about of n z-variables of multidimensional (IM MLRA)-sample. The values of n z-variables located in the n columns of matrix Z_{mn} The rank of this matrix is equal to n-1. The new inverse problem of multiple linear regression analysis (IP MLRA) is solved with the use of Inverse Model of Principal Component Analysis (IM PCA[4]) statistical modeling of the n correlated z-variables satisfying IM MLRA to all equations and mathematical relations. A new inverse problem of multiple linear regression analysis (IP MLRA) of statistical modeling of n correlated z-variables (n-1 independent, 1 dependent) that satisfy all the equations and relations of the IM MLRA is solved. The input parameter of the MZPM is the vector $\beta=(\beta_1,..., \beta_{n-1})^T$ of the regression coefficients for n-1 independent z-variables. A theorem on n z-variables of a multidimensional sample of IM MLRA is proved with values located in n columns of the matrix Z_{mn} of rank n-1. Numerical algorithms have been tested using the example of modeling a multidimensional Λ-sample of z-variables from OM MLRA (for n=4). The known vector $\beta=(\beta_1, \beta_2, \beta_3)^T$ of the regression coefficients from the article [1] is used. The obtained model data are adequate for the values of given statistics of a real multidimensional sample.*

Key words: *inverse model of multiple linear regression analysis*
Language: *Russian*
Citation: Zhanatauov SU (2018) INVERSE MODEL OF MULTIPLE LINEAR REGRESSION ANALYSIS. ISJ Theoretical & Applied Science, 04 (60): 201-212.
Soi: http://s-o-i.org/1.1/TAS-04-60-38 ***Doi***: crossref https://dx.doi.org/10.15863/TAS.2018.04.60.38

ОБРАТНАЯ МОДЕЛЬ МНОЖЕСТВЕННОГО ЛИНЕЙНОГО РЕГРЕССИОННОГО АНАЛИЗА

Аннотация: *В статье разработана обратная модель множественного линейного регрессионного анализа (ОМ МЛРА). С применением равенств из обратной модели главных компонент (ОМ ГК [4]) решена новая обратная задача множественного линейного регрессионного анализа (ОЗ МЛРА) статистического моделирования значений n коррелированных z–переменных, удовлетворяющих всем уравнениям и соотношениям ПМ МЛРА. При наличии разбиения множества z–переменных на 2 части: n-1 независимых плюс 1 зависимая z–переменная. Входным параметром ОЗ МЛРА является вектор $\beta=(\beta_1,...,\beta_{n-1})^T$ коэффициентов регрессии при n-1 независимых z–переменных. Доказана Теорема об n z–переменных многомерной выборки ОМ МЛРА. с значениями, расположенными в n столбцах матрицы Z_{mn} ранга n-1. Численные алгоритмы апробированы на примере моделирования многомерной Λ-выборки z–переменных из ОМ МЛРА (при n=4). Использован известный вектор $\beta=(\beta_1,\beta_2,\beta_3)^T$ регрессионных коэффициентов из статьи [1]. Полученные модельные данные адекватны по значениям заданных статистик реальной многомерной выборки.*

Ключевые слова: *обратная модель множественного линейного регрессионного анализа*

Введение

Регрессионные модели из подмножества линейных моделей, являются одними из важных инструментов статистического анализа реальных данных. Рассмотрим регрессионную модель вида $z_n=\beta_1z_1+\beta_2z_2+...+\beta_{n-1}z_{n-1}+\alpha$, где $z_1,z_2,...,z_{n-1}$ - набор объясняющих (независимых) переменных («регрессоров»), z_n-переменная отклика (зависимая переменная), $\beta_1,...,\beta_{n-1}$-регрессионные коэффициенты, α-свободный член. Эта модель отражает взаимосвязь между двумя или более *объясняющими* переменными и одной переменной *отклика* путем подгонки вышеприведенного линейного уравнения к стандартизованным значениям z-переменных $z_{ij}=(x^0_{ij}-x^{cp}_j)/s_j$. Законы распределения одномерных случайных величин $\xi_1,\xi_2,\ldots,\xi_n$, соответствующих z-переменным $z_1,z_2,\ldots,z_n$ не известны. Здесь x^0_{ij}-i-ое значение j-го признака реального объекта, $x^{cp}_j=(x^0_{1j}+\ldots+x^0_{mj})/m$ – среднее арифметическое, $s^2_j=(x^2_{1j}+\ldots+ x^2_{mj})/m$-стандартное отклонение, $x_{ij}=x^0_{ij}-x^{cp}_j$-отклонение от среднего значения x^{cp}_j. Стандартизованные значения z_n изменяются относительно значений $z_1,z_2,\ldots,z_{n-1}$ с одинаковыми стандартными отклонениями, равными 1. В соответствии с этим разбиением z-переменных m значений всех n z-переменных образует 2 подматрицы Z_1,Z_2 матрицы $Z_{mn}=[Z_1 \rfloor Z_2]$ для m-на-n матрицы Z_{mn}. Элементы столбцов (с номерами j=1,…,n) матрицы Z_{mn} центрированы выборочными средними и нормированы стандартными отклонениями: $z_{ij}=(x^0_{ij}-x^{cp}_j)/s_j$. Элементы $z_{ij}=(x^0_{ij}-x^{cp}_j)/s_j$ матрицы стандартизованных отклонений не имеют размерности, и все ее столбцы имеют одинаковые дисперсии, равные единице. Это – одно из удобств для наших задач. Шаги при моделировании значений $z_{ij}=(x^0_{ij}-x^{cp}_j)/s_j$ n z-переменных отделены от шагов при вычислении выборочных средних x^{cp}_j и дисперсий s_j^2 для реальных данных x^0_{ij} i=1,…,m, j=1,…,n. Векторы выборочных средних и дисперсий должны определяться из матрицы реальных данных $X^0_{mn}=\{x^0_{ij}\}$. Моделирование значений z-переменных проводится отдельно, а при преобразовании их в x^0- переменные $x^0_{ij}=z_{ij}s_j+x^{cp}_j$, i=1,…,m, j=1,…,n, можно использовать любые векторы выборочных средних и стандартных отклонений (дисперсий). При этом в полученной модельной матрице $Z_{mn}=[Z_1 \rfloor Z_2]$ можно переставлять местами строки – эти действия не влияют на значения элементов корреляционных матриц, вычисляемых ниже.

Регрессионные модели разработаны в многочисленных вариантах, применялись во многих предметных областях. Многочисленные статьи, монографии на разных языках, программы для ЭВМ общеизвестны, стали повседневным инструментарием.

Однако, в связи с с кризисными событиями, например, в финансовой сфере, актуальны задачи управления вычисленными значениями коэффициентов регрессии при n-1 независимых z–переменных, линейная комбинация которых влияет на зависимую переменную.

Например, прибыль (зависимая переменная или отклик) в зависимости от прироста ресурсов, вложений (независимых переменных) в прибыльные активные операции. Предварительная фиксация значений коэффициентов регрессии $ß_1,\ldots,ß_{n-1}$ и моделирование значений независимых и зависимой переменных позволит иметь ряды матриц $Z_{mn}=[Z_1 \rfloor Z_2]$, достигающих по построению такой матрицы $Z_{mn}=[Z_1 \rfloor Z_2]$, которая будет иметь заданные целевые значения коэффициентов регрессии $ß_1,\ldots,ß_{n-1}$. Тогда возможно проектирование рядов векторов значений коэффициентов регрессии $ß=(ß_1,\ldots,ß_{n-1})^T$ и соответствующих им рядов матриц $\{Z_{mn}=[Z_1 \rfloor Z_2]\}$ с заданны ми свойствами.

Современные тенденции в теории и практике финансового анализа связаны с проблемой модификации системы финансовых коэффициентов, с приведением этой системы к форме, удобной для принятия адекватных управленческих решений в области финансового менеджмента. В этом направлении существует несколько подходов. Предпочтителен подход, когда выбирают из всех существующих финансовых показателей и коэффициентов незначительное количество тех, которые наиболее полно и всесторонне характеризуют финансовое состояние банка.

Здесь мы остановимся на статистическом подходе к коэффициентному методу финансового анализа. Суть нашего подхода может быть сведена к анализу выборочных коэффициентов корреляции и коэффициентов регрессии. Последние имеют практически важные смысл и интерпретацию: «если банк увеличит на 1 тысячу тенге свои кредитные вложения, то банк потерпит убыток в 347,87 тенге, а если банк увеличит на 1 тысячу тенге свои вложения в ценные бумаги, то банк потерпит убыток в 225,42 тенге. т. е. банку в это время нельзя заниматься традиционными операциями» [1].

Для иллюстрации статистического подхода к финансовому экспресс-анализу нужны модельные данные, адекватные по значениям статистик многомерной выборки. Перечень этих статистик (векторы, матрицы) будет выявлен по мере изложения текста. При этом закон распределения значений 1-мерных переменных для финансовых показателей бывает неопределенным. , что достигается применением обратной модели главных компонент (ОМ ГК) [2,3], для 1-мерных z-переменных из R-, Λ-выборок не определены законы распределений.

Модели и задачи

Исходной гипотезой для рассматриваемой ниже обратной задачи множественной линейной регрессии (ОЗ МЛРА) является существование уравнения регрессии вида $z_n=\beta_1 z_1+\beta_2 z_2 + \ldots+\beta_{n-1}z_{n-1}$, где, в отличие от прямой задачи множественной линейной регрессии (ПЗ МЛРА) известны значения $ß_1,\ldots,ß_{n-1}$ вектора $ß=(ß_1,\ldots,ß_{n-1})^T$ регрессионных коэффициентов, значение свободного члена α. Модель множественной линейной регрессии, где вычисляется единственный вектор $ß=(ß_1,\ldots,ß_{n-1})^T$ регрессионных коэффициентов, назовем (при $\alpha=0$) прямой моделью множественной линейной регрессии (ПМ МЛРА) и обозначим так: $Z_{mn}=[Z_1\rfloor Z_2]=>(R^{-1}{}_{11},R_{12},\beta)$. В ПМ МЛРА решена ПЗ МЛРА, ее решение β единственно и равно $\beta_R=R^{-1}{}_{11}R_{12}$. Ранг матрицы R_{11} равен n-1. Для каждого значения z_n из реальной выборки и оценки ее значения из ПМ МЛА разность этих величин не равна нулю. В нашей ОМ МЛРА аналогичная разность равна нулю. В этом состоит преимущество ОМ МЛРА, где модельные значения n z–переменных точно удовлетворяют формуле $z_n=\beta_1z_1+\beta_2z_2 + \ldots+\beta_{n-1}z_{n-1}$. Аддитивное случайное приращение α_i, i=1,…,m, к значениям z_{in} (присущее ПМ МЛРА) в ОМ МЛРА придает вектору-решению $(z_{i1}, z_{i2},\ldots,z_{i(n-1)},z_{in})^T$ нашей модельной выборки свойство ошибки предсказанного значения. Значениями этих ошибок в ОМ МЛРА можно управлять, что невозможно в ПМ МЛРА. Следовательно теоретическое решение ПМ МЛРА является одним из бесконечного множества теоретических решений ОМ МЛРА. Математического доказательства не приводим.

Трудным местом ПЗ МЛРА$Z_{mn}=[Z_1\rfloor Z_2]=>(R^{-1}{}_{11}, R_{12},\beta)$ является вычисление обратной матрицы для симметрической корреляционной матрицы «регрессоров» R_{11}, которая может быть неполного ранга – тогда не существует для нее обратной матрицы. Если она «плохо обусловлена», то уменьшение числа обусловленности матрицы показывает насколько матрица близка к матрице неполного ранга (для квадратных матриц - к вырожденности). В работах [2-9] число обусловленности корреляционной матрицы измеряется набором ее f-параметров $f_1(\Lambda_{nn})=\lambda_1+\ldots+\lambda_n)=n$, $f_2(\Lambda_{nn})=(\lambda^2{}_1+\ldots+\lambda^2{}_n)$, $f_3(\Lambda_{nn})=\lambda_1/\lambda_n$, $f_4(\Lambda_{nn})=(\lambda_1+\ldots+\lambda_\ell)/n<1$, $f_5(\Lambda_{nn})=\lambda_1\times\lambda_2\times\lambda_3\times\ldots\times\lambda_n$, $f_6(\Lambda_{nn})= \lambda_1/\lambda_2+\ldots+\lambda_{n-1}/\lambda_n)$. Значение f-параметра $f_3(\Lambda_{nn})=\lambda_1/\lambda_n$ измеряет значение числа обусловленности, а остальные – близость (удаленность) от вырожденности кореляционной матрицы R_{nn}. Для нахождения значений f-параметров необходимо решить прямую спектральную задачу (ПСЗ):$R_{nn}=>(C_{nn},\Lambda_{nn})$, где квадратная ортонормированная матрица C_{nn} - матрица собственных векторов $\mathbf{c_j}=(c_{1j},c_{2j}\ldots c_{nj})^T$,j=1,…,n. Они образуют ортогональную матрицу $C_{nn}=[\mathbf{c_1}|\mathbf{c_2}|\ldots|\mathbf{c_n}]$, согласованную с матрицей собственных чисел (со спектром) $\Lambda_{nn}=diag(\lambda_1,\lambda_2,\ldots,\lambda_n)$, $\lambda_1>\ldots>\lambda_n>0$, таким образом, что выполняются равенства $R_{nn}C_{nn}=C_{nn}\Lambda_{nn}$, $C^T{}_{nn}C_{nn}=C_{nn}C^T{}_{nn}=I_{nn}$, где $diag(R_{nn})=(1,\ldots,1)$, $tr(R_{nn})=1+1+1 + 1+1+1=tr(\Lambda_{nn})=\lambda_1+\ldots+\lambda_n=n$. Матрицы C_{nn} и Λ_{nn} вычислятся одновременно по известной корреляционной матрице R_{nn}. Матрица R_{nn} вычисляется по стандартизованной выборке Z_{mn}:$R_{nn}=(1/m)Z^T{}_{mn}Z_{mn}$. Элементы спектра $\Lambda_{nn}=diag(\lambda_1,\ldots,\lambda_n)$, n>2, являются вышеприведенными измерителями как степени вырожденности, так и других свойств.

Решаемая ниже ОЗ МЛРА, как показано ниже в «Теореме о z–переменных в ОМ МЛРА», имеет бесконечное множество решений $(R^{(\ell)}{}_{11},R^{(\ell)}{}_{12},Z^{(\ell)}{}_1, Z^{(\ell,t)}{}_2)$, где корреляционные матрицы $R^{(\ell)}{}_{11}$ модели руются в модели вида:$(n,\varphi_{11})=>(R^{(\ell)}{}_{11})$, подматрицы $R^{(\ell)}{}_{12}$ вычисляются: $R^{(\ell)}{}_{12}=R^{(\ell)}{}_{11}ß$, подматрицы $Z^{(\ell,t)}{}_1$ являются решением ОЗ АГК:$R^{(\ell)}{}_{11}=>(C^{(\ell)}{}_{11},\Lambda^{(\ell)}{}_{11}, Y^{(t)}{}_{m(n-1)},Z^{(\ell,t)}{}_{m(n-1)})$, подматрица $Z^{(\ell,t)}{}_2$-решением Оптимизационной задачи №5, $t=1,\ldots\kappa_t<\infty$, $\ell=1,\ldots,k_\ell<\infty$.

Выборки $Z^{(t)}{}_1,Z^{(\ell,t)}{}_2$ ОМ ГК удовлетворяют соотношениям: $(1\backslash m)Z^{(t)T}{}_1Z^{(t)}{}_1=R^{(\ell)}{}_{11}$, $(1\backslash m)Z^{(t)T}{}_1Z^{(\ell,t)}{}_2=R^{(\ell)}{}_{12}$, $(1\backslash m)Z^{(\ell,t)T}{}_2Z^{(\ell,t)}{}_2=R_{22}=1$. Матрицы$C^{(\ell)}{}_{11},\Lambda^{(\ell)}{}_{11},Y^{(t)}{}_{m(n-1)},Z^{(\ell,t)}{}_{m(n-1)})$ из решаемых задач используются для достижения требуемых равенств, а также удовлетворяют соотношениям ОМ ГК, доказанным в Теореме о Λ–выборках [9].

Обратная модель множественной линейной регрессии

Необходимо иметь данные, демонстрирующие всевозможные динамики рассматриваемых нами агрегированных показателей. Динамики этих показателей покажут оптимистические или неблагоприятные тенденции в периоды времени, наличие которых мы будем определять по значениям показателей, по коэффициентам корреляции, по значениям коэффициентов регрессии $\beta_1,\ldots, \beta_{n-1}$, по коэффициентам эластичности переменной z_n по объясняющей переменной z_j с номером j, где j может принимать одно из значений 1,2 …, n-1.

В данной работе управляемым (входным) параметром модели, генерирующей модельные данные, являются значения коэффициентов регрессии $\beta_1,\ldots,\beta_{n-1}$ при переменных $z_1,z_2,\ldots,z_{n-1}$. В общеизвест ной ПМ МЛРА $Z_{mn}=[Z_1\rfloor Z_2]=>(R^{-1}{}_{11},R_{12},\beta)$ входным объектом является реальная многомерная выборка $Z_{mn}=[Z_1\rfloor Z_2]$, а выходным-вектор коэффициентов регресссии $\beta=(\beta_1,\ldots, \beta_{n-1})$.

Как показано ниже для единственного вектора коэффициентов регрессии $\beta=(\beta_1,\ldots,\beta_{n-1})$ существует бесконечное множество стандартизованных многомерных выборок $Z_{mn}=[Z_1\rfloor Z_2]$, являющиеся многомерными Λ-выборками с свойствами, доказанными в теореме о Λ–выборках [9].

Ниже покажем что Обратная модель множественной линейной регрессии (ОМ МЛРА) имеет схематическое изображение вида: ß=> $[R^{(\ell)}_{11},R^{(\ell)}_{12}Z^{(\ell,t)}_1\rfloor Z^{(\ell,t)}_2]$, $t=1,\dots,\kappa_t<\infty$, $\ell=1,\dots,\kappa_\ell<\infty$.

Обратная задача множественного линейного регрессионного анализа

Пусть нам даны перечень наименований и число n-1 независимых переменных, т.е. задано число «регрессоров» $z_{i1},\dots,z_{i,n-1}$ и заданы значения коэффициентов регрессии $\beta_1,\dots,\beta_{n-1}$, функционально определяющих значения зависимой переменной z_n.

Требуется моделировать значения выборочных коэффициентов корреляции между n-1 независимыми z-переменными из подматрицы R_{11} коэффициентов корреляций только независимых переменных («регрессоров»). Требуется моделировать значения выборочных коэффициентов корреляции $r_{1n},\dots,r_{n-1,n}$ между зависимой z-переменной z_n и независимыми z-переменными $z_1,\dots,z_{n-1}$, объединенных в вектор-столбец $R_{12}=(r_{1n},\dots,r_{n-1,n})^T$. Требуется моделировать многомерную m*n-выборку $Z_{mn}=[Z_1|Z_2]$ значений n z–переменных, удовлетворяющих равенствам из ПМ МЛРА $Z_{mn}=[Z_1\rfloor Z_2]$=>$(R^{-1}_{11},R_{12},\beta)$:

$$(1\backslash m)Z^TZ=R_{nn},Z_{mn}=[Z_1|Z_2]=\{(z_{i1},\dots,z_{i,n-1}|z_{in})\},$$

$$(1\backslash m)Z_1^TZ_1=R_{11},$$

$$(1\backslash m)Z_1^TZ_2=R_{12},\ z_n=\mathbf{z}R^{-1}_{11}R_{12}=\mathbf{z}\beta,\ \mathbf{z}=(z_1,\dots,z_{n-1}),$$

$$\boldsymbol{\beta}=R^{-1}_{11}R_{12}.$$

Входным объектом обратной задачи множественной линейной регрессии (ОЗ МЛРА) является вектор $\beta_R=(\beta_1,\dots,\beta_{n-1})$. Выходным объектом ОЗ МЛРА, т.е. решением, является многомерная выборка $Z_{mn}=[Z_1|Z_2]=\{(z_{i1},\dots,z_{i,n-1}|z_{in})\}$, матрица R_{nn} такие, что выполняются соотношения $\beta_R=R^{-1}_{11}R_{12}$, $(1\backslash m)Z_1^TZ_1=R_{11}$, $(1\backslash m)Z_1^TZ_2=R_{12}$. Матрица R_{nn} в соответствии с разбиением R-выборки $Z_{mn}=[Z_1|Z_2]$, где Z_1 – m-на-(n-1) матрица, Z_2 - m-на-1 матрица , матрица R_{nn} разбита на 3 блока : R_{11},R_{12} $R_{21}=R^T_{12}$, и элемент r_{nj} $r_{nn}=1$.

Схематически данную ОЗ МЛРА изобразим пока так:(m,n,φ_{11},β)=>(R_{11},R_{12},Z_1,Z_2). В процессе ее решения схема будет конкретизироваться.

Подзадачу 1 для ОЗ МЛРА обозначим так: (n,φ_{11})=>(R_{11}). Решение подзадачи 1 состоит из 2 шагов. Подзадачу 2 изобразим так: (R_{11},m)=> $(C_{11},\Lambda_{11},Y^{(t)}_{m(n-1)},Z^{(t)}_{m(n-1)})$,$t=1,\dots,\kappa_t<\infty$, а подзадачу 3 (R_{11},β)=>R_{12}. Результат решений трех подзадач после удаления символов сопутствующих матриц обозначается в виде :(m,n,φ_{11},β)=>(R_{11},R_{12},Z_1,Z_2).

Рассмотрим **подзадачу 1**.

Шаг 1. При фиксированном значении параметра φ_{11} моделируем спектр $\Lambda_{n-1)(n-1)}=\mathrm{diag}(\lambda_1,\dots,\lambda_{n-1})$, с значениием φ-параметра спектра $\Lambda_{(n-1)(n-1)}\varphi=[(f_2(\Lambda_{(n-1)(n-1)})-n)/(n(n-1)]^{1/2}$, равного заданному значению φ_{11}. Значение φ_{11} должно принадлежать одному из 5 интервалов изменения коэффициента корреляции по шкале Чеддока (Chaddock scale). По этой шкале количественная мера тесноты связи: абсолютное значение коэффициента корреляции, принадлежащее интервалу от 0 до 0.3 - качественно интерпретируется как «очень слабая», интервалу от 0.3 до 0.5-«слабая, умеренная», интервалу от 0.5 до 0.7-«заметная» (moderate positive), интервалу от 0.7 до 0.9-«высокая», интервалу от 0.9 до 1-«очень высокая».

Схему этого Шага 1 обозначим так:$(n-1,\varphi_{11})$=>$\{\Lambda_{n-1)(n-1)}=\mathrm{diag}(\lambda_1,\dots,\lambda_{n-1})\}$. Фигурные скобки {} обозначают бесконечность множества таких спектров. При этом элементы $\lambda_1,\dots,\lambda_{n-1}$ спектра $\Lambda_{n-1)(n-1)}$ неизвестной корреляционной матрицы $R_{n-1)(n-1)}$ моделируем с применением программы SPEKTR из ППП «Спектр» [10]. Заметим, что программа SPEKTR не моделирует матрицу $R_{n-1)(n-1)}$ (в виде блока R_{11}) а моделирует только спектр $\Lambda_{n-1)(n-1)}=\mathrm{diag}(\lambda_1,\dots,\lambda_{n-1})$ с заданным значением f-параметра $f_2(\Lambda_{n-1)(n-1)})=\lambda^2_1+\dots+\lambda^2_{n-1}=f_2(R_{n-1)(n-1)})$. Элементы спектра упорядочены в порядке убывания:$\lambda_1>\dots>\lambda_{n-1}>0,\lambda_1+\dots+\lambda_{n-1}=n-1$, причем матрица $\Lambda_{n-1)(n-1)}$ может быть как полноранговой, так и, при необходимости, может быть смоделирован и с неполным рангом. Тогда число ее положительных элементов равно k<n-1. Одна из таблиц значений φ_{11}, принадлежащих 5 интервалам изменения параметра φ, приведена в статье [9,стр.187,Таблица 1]. Приведенные «Модельные значения элементов 20 спектров 20 неизвестных корреляционных матриц, имеющих заданные значения f-параметров спектра» получены при значениях φ_{11}=0.20,0.35,0.40,0.45,0.50,0.55,0.60,0.65,-0.70,0.70,0.75,0.80,0.80,0.85,0.85,0.90,0.90,0.95,0.98. Значение φ_{11} $(0<\varphi_{11}<1)$ не зависит от значения n. Оно позволяет моделировать n-1 положительных элементов $\lambda_1,\dots,\lambda_{n-1}$, в сумме равных n-1. В подзадаче 1 эти n-1 положительных чисел позволяют моделировать $[(n-1)^2-(n-1)]/2$ чисел r_{ij}, по абсолютной величине не превосходящих 1. Таким образом одно положительное число φ_{11} (и не только оно) позволяет генерировать множество значений коэффициентов корреляции. Значение φ_{11} регулирует интервал значений, в пределах которого изменяются значения генерируемых коэффициентов корреляции. Примеры фиксации интервалов для значений φ_{11} приведены, например, в [9,стр.187].

Спектр $\Lambda_{n-1)(n-1)}$ может быть получен и по-другому. Если требуется моделировать спектр с заданными значениями его f-параметров f_1,f_2,f_4, то задача моделирования (f_1,f_2,f_4)-спектров $\Lambda_{n-1)(n-1)}$ $(f_1=n-1)$ по-разному решены в [4-11]. Все модельные (f_1,f_2,f_4)-спектры из их бесконечного множества имеют заданные значения f-параметров f_2,f_4, с заданной погрешностью [1,2] $\|f_2-f_2(\Lambda_{nn})\|\le\epsilon_0$, $\|f_4-f_4(\Lambda_{nn})\|=0$.

Шаг 2. Моделирование блока R_{11} корреляционной матрицы R_{nn}. Здесь нам предстоит воспользоваться одним из решений из бесконечного множества решений Обратной спектральной задачи (ОСЗ): $\Lambda_{n-1)(n-1)}$=> $(C^{(\ell)}{}_{n-1)(n-1)},R^{(\ell)}{}_{n-1)(n-1)})$, $\ell=1,\ldots,k_\ell<\infty$. Здесь мы будем использовать одну из корреляционных матриц $R^{(\ell)}{}_{n-1)(n-1)}$, элементы которой по абсолютным величинам должны удовлетворять нас и соответствовать выбранному значению φ_{11} в предыдущем Шаге 1.

Для моделирования бесконечного множества корреляционных матриц $R_{n-1)(n-1)}$ с одним и тем же спектром $\Lambda_{n-1)(n-1)}=diag(\lambda_1,\ldots,\lambda_{n-1})$ используем программу CORMAP (CORMAT, COMA31) из ППП «Спектр» [10]. В нашей ОЗ МЛРА и в других задачах анализа данных, где применяется ОМ ГК [2-5], для модельного спектра $\Lambda_{(n-1)(n-1)}$ существует бесконечное множество корреляционных матриц $R^{(\ell)}{}_{(n-1)(n-1)},\ell=1,\ldots,\kappa_\ell<\infty$ (лемма C.P.Chalmers [12]). Зафиксируем номер ℓ решения обратной спектральной задачи (ОСЗ) [5], а полученную корреляционную матрицу $R^{(\ell)}{}_{n-1)(n-1)}$ обозначим как искомый блок R_{11} матрицы R_{nn} . Подзадача 1: $(n,\varphi_{11})=>(R^{(\ell)}{}_{11})$, $\ell=1,\ldots,\kappa_\ell<\infty$ решена.

Рассмотрим **Подзадачу 2**: $R^{(\ell)}{}_{11}$, m)=> $(C^{(\ell)}{}_{11},\Lambda_{11},Y^{(t)}{}_{m(n-1)},Z^{(t,\ell)}{}_{m(n-1)})$, $t=1,\ldots,\kappa_t<\infty$, $\ell=1,\ldots,\kappa_\ell<\infty$. Для полученного блока $R_{n-1)(n-1)}$ корреляционной матрицы размерности n-1, располагаемого в первых n-1 столбцах корреляцион ной матрицы размерности n, надо моделировать выборку $Z^{(t,\ell)}{}_{m(n-1)}$. Она будет выступать в роли m-на-(n-1) подматрицы Z_1 из матрицы $Z_{mn}=[Z_1|Z_2]$. Подзадача 2 решается за 2 Шага.

Шаг 1. Реализация варианта 2 обратной модели главных компонент (ОМ ГК) с применением программы IMPC2 из ППП «Спектр» [ ].. Здесь решается обратная задача анализа главных компонент (ОЗ АГК,[2-6]), где входным объектом является не диагональная матрица дисперсий главных компонент $\Lambda_{(n-1)(n-1)}$ = $diag(\lambda_1,\ldots\lambda_{n-1})$., а корреляционная матрица $R^{(\ell)}{}_{11}$ размерности (n-1)-на-(n-1). Свойства решений ОЗ АГК, а именно свойства Λ_{11}-выборок $Y^{((t)}{}_{m(n-1)}$, $Z^{(\ell,t)}{}_{m(n-1)})$, $t=1,\ldots,\kappa_t<\infty$, $\ell=1,\ldots,\kappa_\ell<\infty$, доказаны в [5,67]. Для моделирования Λ_{11}-выборок $Y^{((t)}{}_{m(n-1)}$, $Z^{(\ell,t)}{}_{m(n-1)}$, $\ell=1,\ldots,\kappa_\ell<\infty,t=1,\ldots,\kappa_t<\infty$, используем программу IMPC2 из ППП «Спектр» [10]. Зафикси руем значение номера t и выберем матрицу $Z_{m(n-1)}$ в качестве подматрицы Z_1 из разбиения n-на-n матрицы $Z_{mn}=[Z^{(t,\ell)}{}_1\rfloor Z^{(\ell)}{}_2]$. Здесь подматрица $Z^{(\ell)}{}_2$ пока неизвестна. Теперь имеем матрицы $R^{(\ell)}{}_{11}$ и $Z^{(t,\ell)}{}_1$ таковы, что выполняется требуемое равенство (1\m) $Z^{(t,\ell)T}{}_1Z^{(t,\ell)}{}_1=R^{(\ell)}{}_{11}$.

Шаг 2. В нашей задаче 1 подзадача моделирования бесконечного множества подматриц $R_{12}=(1/m)Z^T{}_1Z_2$ схематично изображена (Шаг 2 п**одзадачи 2)** так: $(R^{(\ell)}{}_{11},\beta_R)=>R^{(\ell)}{}_{12}$, $\ell=1,\ldots,k_\ell<\infty$. Здесь подматрица $R^{(\ell)}{}_{12}$ легко вычисляется. Так как ранее нами смоделирована матрица $R^{(\ell)}{}_{11}$, а вектор коэффициентов регрессии $\beta=(\beta_1,\ldots,\beta_{n-1})^T$ задан по условию нашей ОЗ МЛРА, то подматрица $R^{(\ell)}{}_{12}$ равна $R^{(\ell)}{}_{12}=R^{(\ell)}{}_{11}\beta$. Ее схема: $(R^{(\ell)}{}_{11},\beta)$ $=>Z^{(\ell,t)}{}_2$, $t=1,\ldots,\kappa_t<\infty$, $\ell=1,\ldots,\kappa_\ell<\infty$.

Подзадача 3 состоит в моделировании подматрицы $Z^{(t,\ell)T}{}_2$ из матрицы $Z_{mn}=[Z_1\rfloor Z_2]$. Здесь Z_1 является Λ_{11}-выборкой из их бесконечного множества, обладает свойствами из Теоремы о Λ-выборках [9]. Она смоделирована выше. Задача моделирования элементов подматрицы Z_2 решена в работе [8]. В задаче 4 из работы [8] «матрица $R_{12}=(1/m)Z^T{}_1Z_2$, является вектором длины (n-1) и ее элементы r_{1j},r_{2n}, $r_{3n},\ldots,r_{nn}$ в рамках оптимизационной задачи №4 [8] являются реализациями случайной величины.

В подзадаче 3 моделируем элементы из подматрицы Z_2 из матрицы $Z_{mn}=[Z_1|Z_2]$. По условию ОЗ МЛРА элементы из подматрицы $R^{(\ell)}{}_{12}=R^{(\ell)}{}_{11}\beta$ должны удовлетворять равенству $R^{(\ell)}{}_{12}=(1/m)Z^{(\ell)T}{}_1Z_2$. Это матричное уравнение представляет собой систему, состоящую из n-1 линейных и одного нелинейного уравнений. Решаем эту систему методом Ньютона, применяя надстройку «Поиск решения» в ЭТ. (процедура Solver).

Разработаем программу-таблицу (Таблица 1) и решаем Оптимизационную задачу №5 (Рисунки 1,2). Нажав на кнопку «Выполнить» найдем элементы подматрицы $Z^{(\ell,t)}{}_2$, зависящую от выборки $Z^{(t)}{}_1$ и от матрицы $R^{(\ell)}{}_{12}$. Пара номеров (ℓ,t) у подматрицы $Z^{(\ell,t)}{}_2$ образуется из номера t подматрицы $Z^{(t)}{}_1$ и номера ℓ подматрицы $R^{(\ell)}{}_{12}$.

Отличие нашей задачи №5 от задачи №4 из [8] состоит не только в способе задания значений n коэффициентов корреляции $\mathbf{r_{1j},r_{2j}, r_{3j},\ldots, r_{jj},\ldots, r_{nj}}$. Общим местом в задачах №4 и №5 является неизвестность значений j-ой-в задаче 4 и n–ой - в задаче 5 z-переменной: $z_{1n}, z_{2n},\ldots,z_{mn}$. Каждое из этих неизвестных значений, например $z_{\kappa j}$, умножается на известное значение $z_{\kappa 1}$: $z_{\kappa 1}\times z_{\kappa j}$. Сумма таких произведений (если индекс k изменяется от 1 по m), умноженная на постоянную величину (1/m), равна коэффициенту корреляции r_{1j}: $(1/m)\times(z_{11}\times z_{1j}+z_{21}\times z_{2j}+\ldots+z_{\kappa 1}\times z_{\kappa j}+\ldots+z_{m1}\times z_{mj})$. Аналогичный вид имеют формулы для коэффициентов корреляции r_{2j}, $r_{13j},\ldots,r_{nj}$, $j=1,\ldots,n$.

Оптимизационная задача №5

Пусть в модели многомерного линейного регрессионного анализа (МЛРА) задано разбиение множества n z-переменных на независимые - $z_1,\ldots,z_{n-1}$ и зависимую - z_n. Предположим, что существует уравнение регрессии вида $z_n=\beta_1 z_1+\beta_2 z_2 + \ldots+\beta_{n-1}z_{n-1}$, где известны значения $ß_1,\ldots,ß_{n-1}$ вектора $ß=(ß_1,\ldots,ß_{n-1})^T$ коэффициентов регрессии.

Рассмотрим задачу моделирования значений z -переменных из подматрицы Z_2 матрицы $Z_{mn}=[Z_1\rfloor Z_2]$ В Шаге 1 решения подзадачи 2 задачи 1 были получены значения элементов подматрицы Z_1. Эти значения будем использовать в качестве параметров уравнений системы (*).

Если ***известны*** значения n-1 z-переменных, расположенных в n-1 столбцах подматрицы Z_1 размерности m×(n-1) и ***известны*** значения $\mathbf{r_{1n},r_{2n},r_{3n},\ldots,r_{nn}}$ n коэффициентов корреляции (между n-1 независимыми $z_1,\ldots,z_{n-1}$ и одной зависимой z_n z-переменными) из подматрицы $R_{12}=(1/m)Z^T{}_1Z_2$, то имеем систему уравнений, состоящую из n-1 линейных уравнений и 1 нелинейного уравнения:

$$(1/m)\times(z_{11}\times z_{1n}+z_{21}\times z_{2n}+\ldots+z_{к1}\times z_{кn}+\ldots+z_{m1}\times z_{mn})=\mathbf{r_{1n}}$$
$$(1/m)\times(z_{21}\times z_{1n}+z_{22}\times z_{2n}+\ldots+z_{к2}\times z_{кn}+\ldots+z_{m2}\times z_{mn})=\mathbf{r_{2n}}$$
……………………………………………………

$$(1/m)\times(z_{1n}\times z_{1n}+z_{2n}\times z_{2n}+\ldots+z_{кn}\times z_{кn}+\ldots+z_{mn}\times z_{mn})=\mathbf{r_{nn}} \quad (*)$$

Требуется найти m-мерное решение–вектор $Z_2=(z_{1n},z_{2n},\ldots,\ z_{mn})^T$ из матрицы $Z_{mn}=[Z_1\rfloor Z_2]$, удовлетворяющее системе уравнений (*). В этой системе известны значения элементов из подматрицы Z_1 и известны значения коэффициентов корреляции из подматрицы R_{12}, удовлетворяющей равенству $R_{12}=(1/m)Z^T{}_1Z_2$ (1/m). Не известны значения $z_{1n},z_{2n},\ldots,\ z_{mn}$.

Для решения данной системы уравнений разработаем программу-таблицу в ЭТ Excel-2003 с применением процедуры Solver. В окне (Рисунок 1) надстройки «Поиск решения» в n-1 ячейках одного столбца на листе Excel Excel поместим n формул левых частей уравнений из системы (*). В правом столбце от этого столбца в ячейках введем значения (числа) коэффициентов $\mathbf{r_{1n},r_{2n},\ r_{3n},\ldots,r_{nn}\ r_{1n},r_{2n},\ r_{3n},\ldots,r_{nn}}$. Одну из ячеек с формулами назначим целевой. Описание примера расчетов при n=4 моделирования ногомерной Λ-выборки z–перемен ных в ОМ МЛРА

В правой части системы уравнений (*) нижние индексы известных коэффициентов корреляции имеют вид: (1,n), (2,n), (3,n),…, (n,n-1), (n,n). Одна пара индексов (n,n) является индексом дисперсии стандартизованной n -ой z-переменной, равной 1: $r_{nn}=1$.

Элементы известной выборки Z_1 размерности m×(n-1) и элементы известной корреляционной матрицей $R_{11}=(1/m)Z^T{}_1Z_1$, размерности (n-1)×(n-1) используются в качестве постоянных коэффициентов системы уравнений (*). Стандартизованная n-ая z-переменная с неизвестными значениями $z_{1n},z_{2n},\ldots,\ z_{mn}$ расположена в n-ом столбце матрицы $Z_{mn}=[Z_1\rfloor Z_2]$, $Z_2=(z_{1n},z_{2n},\ldots,z_{mn})^T$. Матрица $R_{12}=(1/m)Z^T{}_1Z_2$, является вектором длины (n-1) и ее элементы $r_{1j},r_{2n},r_{3n},\ldots,r_{nn}$ в рамках оптимизационной задачи №5 известны тор $R_{12}ß=R^{-1}{}_{11}$ является вектором коэффициентов регрессии.

Система (*) из n уравнений с m>n>2 неизвест ными $z_{1j},\ z_{2j},\ldots,\ z_{mj}$ при фиксированном номере j z-переменной, где $j\leq n$, является недоопределенной системой: имеются n(m-1) известных значений z-переменных, n известных значений $r_{1j},r_{2j},r_{3j},\ldots,r_{jj},\ldots,r_{nj}$, m>n неизвестных $z_{1j},z_{2j},\ldots,z_{mj}$. Имеет бесконечное множество решений $\{z_{1j},z_{2j},\ldots,z_{mj}\}$. Решение Оптимизационной задачи №5 проведем с применением процедуры Solver. Программа-таблица приведена в Таблице 1, где приведены начальные значения и полученные значения $z_{1j},z_{2j},\ldots,z_{mj}$ и формулы целевой функции, формул 5 функций ограничений из (*) программы-таблицы для процедуры Solver. Окно процедуры «Поиск решений» в ЭТ Excel с введенными формулами целевой функции оптимизационной задачи №4 (n=6) и 5 функций ограничений из системы (*) приведены на Рисунке 2, а настроенные параметры процедуры «Поиск решений» в ЭТ Excel для решения оптимизационной задачи №5-на Рисунке 1.

Теорема. Пусть дана модель множественной линейной регрессии с уравнением регрессии вида $z_n=\beta_1 z_1+\beta_2 z_2+\ldots+\beta_{n-1}z_{n-1}$, где $z_1,z_2\ ,\ldots,z_{n-1}$ - набор объясняющих (независимых) переменных, z_n - переменная отклика (зависимая переменная), заданы значения $ß_1,\ldots,ß_{n-1}$ вектора $ß=(ß_1,\ldots,ß_{n-1})^T$ коэффициентов регрессии. Выборка Z_{mn} из m значений z-переменных разбита на 2 части $Z_{mn}=[Z_1\rfloor Z_2]$. Подматрицы Z_1 - размерности m×(n-1),Z_2 – размерности m×1 содержат m значений соответственно n-1 независимых и одной независимой z-переменных. Пусть неизвестны элементы подматриц Z_1 и Z_2 и известен вектор $ß=(ß_1,\ldots,ß_{n-1})^T$ коэффициентов регрессссии, удовлетворяющие равенствам $ß=R^{-1}{}_{11}R_{12}$, $R_{11}=(1\backslash m)Z_1{}^TZ_1$, $R_{12}=(1\backslash m)Z_1{}^TZ_2$, $Z_2=Z_1ß$.

Тогда существуют бесконечные множество решений $z_1,z_2,\ldots,z_{n-1},z_n$ системы уравнений регрессии $Z^{(\ell,t)}{}_2$ $=Z^{(t)}{}_1ß$, $ß$,= $R^{(\ell)}{}_{11}$, $R^{(\ell)}{}_{12}$, $R_{11}=(1/m)Z^{(t)T}{}_1Z^{(t)}{}_1$, $R^{(\ell)}{}_{12}=R^{(\ell)}{}_{11}ß$, $R^{(t)}{}_{12}=(1/m)Z^{(t)T}{}_1Z^{(\ell,t)}{}_2$, $(R^{(\ell)}{}_{11},\ R^{(\ell)}{}_{12},Z^{(t)}{}_1,Z^{(\ell,t)}{}_2)$, где матрицы $R^{(\ell)}{}_{11}$ моделируются в модели вида:$(n,\varphi_{11})\Rightarrow(R^{(\ell)}{}_{11})$, подматрицы $R^{(\ell)}{}_{12}$ вычисляются: $R^{(\ell)}{}_{12}=R^{(\ell)}{}_{11}ß$, подматрицы $Z^{(\ell,t)}{}_1$ являются решениями ОЗ АГК:

$R^{(\ell)}_{11}$=>($C^{(\ell)}_{11}$,$\Lambda^{(\ell)}_{11}$, $Y^{(t)}_{m(n-1)}$,$Z^{(\ell,t)}_{m(n-1)}$), ℓ=1,…,k_ℓ<∞, подматрица $Z^{(\ell,t)}_2$ является решением Оптимизационной задачи №5, t=1,…,k_t<∞, ℓ=1,…,k_ℓ<∞. Выборки $Z^{(t)}_1$,$Z^{(\ell,t)}_2$ ОМ ГК удовлетворяют соотношениям $(1\backslash m)Z^{(t)T}_1Z^{(t)}_1=R^{(\ell)}_{11}$, $(1\backslash m)Z^{(t)T}_1Z^{(\ell,t)}_2=R^{(\ell)}_{12}$, $(1\backslash m)Z^{(\ell,t)T}_2Z^{(\ell,t)}_2=R_{22}=1$. Матрицы $C^{(\ell)}_{11}$, $\Lambda^{(\ell)}_{11}$, $Y^{(t)}_{m(n-1)}$, $Z^{(\ell,t)}_{m(n-1)}$) из решаемых задач используются для достижения требуемых равенств, удовлетворяют соотношениям как ОМ ГК, так и ОМ МЛРА.

Доказательство. Решение обратной задачи множественной линейной регрессии (ОЗ МЛРА): многомерная выборка $Z_{mn}=[Z_1|Z_2]$, моделируется при решении задачи 1: (m,n,φ_{11},β_R)=> ($R^{(\ell)}_{11}$,$R^{(\ell)}_{12}$,$Z^{(\ell,t)}_1$, $Z^{(\ell,t)}_2$). Подматрица R_{11} моделируется при решении **подзадачи 1**: (n,φ_{11})=>(R_{11}). Для подматрицы R_{11} решается **подзадача 2**: $R^{(\ell)}_{11}$=>($C^{(\ell)}_{11}$,Λ_{11}, $Y^{(t)}_{m(n-1)}$,$Z^{(t,\ell)}_{m(n-1)}$), t=1,…,k_t<∞, ℓ=1,…,k_ℓ<∞., состоящая из двух Шагов. На Шаге 1 Подзадачи 2 моделируем подматрицу $Z_1=Z^{(t,\ell)}_{m(n-1)}$, на Шаге 2 моделируются (вычисляются) подматрицы $R^{(\ell)}_{12}=R^{(\ell)}_{11}\beta$). Подзадача 3 состоит в численном решении Оптимизационной задачи №5 и в получении модельной подматрицы $Z^{(\ell,t)}_2=(z_{1n},z_{2n},\ldots,z_{mn})^{(\ell,t)T}$. Формальная постановка Оптимизационной задачи №5 и процесс ее решения (моделирования подматрицы $Z^{(\ell,t)}_2$, удовлетворяющей равенству $Z^{(\ell,t)}_2=Z^{(t)}_1\beta$), описан в отдельном разделе данной статьи (ее схема: ($Z^{(\ell,t)}_1$,$R^{(\ell)}_{12}$)=>$Z^{(\ell,t)}_2$, t=1,…, k_t<∞,ℓ=1,…, k_ℓ<∞. Получена последовательность $\beta\rightarrow R^{(\ell)}_{11}\rightarrow Z^{(\ell,t)}_1\rightarrow R^{(\ell)}_{12}\rightarrow Z^{(\ell,t)}_2$. Теорема доказана.

ОЗ МЛРА: ß=>[$Z^{(\ell,t)}_1$⌋$Z^{(\ell,t)}_2$], t=1,…,k_t<∞, ℓ=1,…,k_ℓ< ∞, решаема и ее решения существуют. ОЗ МЛРА имеет бесконечное множество решений с номерами t=1,…,k_t<∞,ℓ=1,…,k_ℓ<∞. Выходным объектом ОМ МЛРА является многомерная выборка $Z_{mn}=[Z_1|Z_2]$= {(z_{i1},…,$z_{i,n-1}$|z_{in})}, *Входным* объектом обратной задачи множественной линейной регрессии (ОЗ МЛРА) является вектор $\beta_R=(\beta_1,\ldots, \beta_{n-1})$. *Выходным* объектом ОЗ МЛРА является многомерная выборка $Z_{mn}=[Z_1|Z_2]$= {(z_{i1},…,$z_{i,n-1}$|z_{in})}.

На шаге 1 подзадачи 1 моделируются матрицы $R^{(\ell)}_{11}$ в модели вида: (n,φ_{11})=>($R^{(\ell)}_{11}$), подматрицы $R^{(\ell)}_{12}$ вычисляются: $R^{(\ell)}_{12}=R^{(\ell)}_{11}\beta$, подматрицы $Z^{(\ell,t)}_1$ являются решением ОЗ АГК: $R^{(\ell)}_{11}$=>($C^{(\ell)}_{11}$, $\Lambda^{(\ell)}_{11}$, $Y^{(t)}_{m(n-1)}$, $Z^{(\ell,t)}_{m(n-1)}$), подматрица $Z^{(\ell,t)}_2$-решением Оптимизационной задачи №5, t=1,…,k_t<∞, ℓ=1,…,k_ℓ<∞. Выборки $Z^{(t)}_1$,$Z^{(\ell,t)}_2$ ОМ ГК удовлетворяют соот ношениям $(1\backslash m)Z^{(\ell,t)T}_1Z^{(\ell,t)}_1=R^{(\ell)}_{11}$, $(1\backslash m)Z^{(\ell,t)T}_1Z^{(\ell,t)}_2= R^{(\ell)}_{12}$, $(1\backslash m)Z^{(\ell,t)T}_2Z^{(\ell,t)}_2=R_{22}=1$. При этом матрицы $C^{(\ell)}_{11}$, $\Lambda^{(\ell)}_{11}$,$Y^{(t)}_{m(n-1)}$, $Z^{(\ell,t)}_{m(n-1)}$ из решаемых задач используются для достижения требуемых равенств, удовлетворяют соотношениям ОМ ГК, доказанным в Теореме о Λ–выборках [9].

Стандартизованные значения переменной z_n изменяются относительно значений переменных z_1,z_2,**…**,z_{n-1} с одинаковыми стандартными отклонениями, равными 1. В соответствии с этим разбиением z-переменных совокупность из m значений всех n z-переменных образует 2 подматрицы [Z_1⌋Z_2] для m-на-n матрицы Z_{mn}=[Z_1⌋Z_2].

Конкретные значения полученного решения z_{1j},z_{2j},…,z_{mj} системы уравнениц (*) зависят от начальных значений, назначаемых пользователем процедуры «Поиск решения» и вводимых в соответствующее поле окна этой процедуры. Ниже в примере мы выбрали любой нормированный вектор (z_{1j},z_{2j},…,z_{mj}) такой, что $1=(z_{1j}^2+\ldots+z_{m,j}^2)/m$, где n=6, m=20. Этот вектор в паре с другими векторами из столбцов матрицы Z_{mn} не дает желаемых значений коэффициентов корреляции, но процедура «Поиск решения» легко преобразует этот нормированный вектор (z_{1j},z_{2j},…,z_{mj}) в вектор, являющийся решением системы (*). В настоящее время разрабатываются компьютерные программы, интегрирующие процедуру Solver с модулями и библиотеками ППП «Спектр» [10] . Для Единого Цифрового Объекта (ЕЦО) из Виртуальной Базы Данных [13].

Пример многомерной Λ-выборки z–переменных, моделируемых в ОМ МЛРА

Этапы применения программ из ППП «Спектр» [10] для получения многомерной Λ-выборки z–переменных, точно удовлетворяющих соотношениям, уравнениям ОМ МЛРА состоят из следующих 3 шагов. На Шаге 1 моделируем спектр неизвестной корреляционной матрицы $\Lambda_{n-1)(n-1)}=diag(\lambda_1,\ldots,\lambda_{n-1})$ размерности n-1. Это число равно числу объясняю щих (независимых) переменных («регрессоров»). На Шаге 2 в соответствии с значением n-1 и со значением, характеризующим гипотетические элементы неизвестного спектра $\Lambda_{n-1)(n-1)}=diag(\lambda_1,\ldots,\lambda_{n-1})$ некоторой корреляционной матрицы размерности n-1. В качестве характеристики (параметра) будущей корреляционной матрицы используем параметр φ_{11} - среднеквадратичное коэффициентов корреляции матрицы $R_{n-1)(n-1)}$, имеющей спектр $\Lambda_{n-1)(n-1)}$. Значение параметра φ_{11} принадлежит интервалу (0,1). Схему этого Шага 2 выше мы обозначили так:(n-1,φ_{11})=>{$\Lambda_{n-1)(n-1)}$= diag(λ_1,…,λ_{n-1})}. Значения λ_1,…,λ_{n-1} в рассматриваемом ниже примере не приводим, а 9 значений матрицы $R_{n-1)(n-1)}$ приведены в Таблице 1 (выделены желтым цветом). При этом используем программу SPEKTR из ППП «Спектр» [10] для моделирования спектра $\Lambda_{n-1)(n-1)}$ =diag(λ_1,…,λ_{n-1}). Заметим, что программа SPEKTR не моделирует матрицу $R_{n-1)(n-1)}$ (в виде блока R_{11}) а моделирует только спектр $\Lambda_{n-1)(n-1)}$=diag($\lambda_1,\ldots,\lambda_{n-1}$) с заданным значением f-параметра $f_2(\Lambda_{n-1)(n-1)})=\lambda^2{}_1+\ldots+\lambda^2{}_{n-1}=f_2(R_{n-1)(n-1)})$. На Шаге 2 проводится моделирование блока R_{11} корреляционной матрицы R_{nn} с использованием программы IMPC2 [10]. Одно из решений из бесконечного множества решений Обратной спектральной задачи (ОСЗ):$\Lambda_{n-1)(n-1)}$=>($C^{(\ell)}{}_{n-1)(n-1)}$, $R^{(\ell)}{}_{n-1)(n-1)}$), $\ell=1,\ldots,k_\ell=1000<\infty$ при n-1=3 имеет вид, приведенный в таблице 1. Здесь мы будем использовать одну из корреляционных матриц $R^{(\ell)}{}_{n-1)(n-1)}$ с одним и тем же спектром $\Lambda_{n-1)(n-1)}$= diag($\lambda_1,\ldots,\lambda_{n-1}$), элементы которой по абсолютным величинам должны удовлетворять нас и соответствовать выбранному значению φ_{11} в предыдущем Шаге 1. Используем программу CORMAP (CORMAT, COMA31) из ППП «Спектр» [10]. Она генерирует матрицу R_{11} (Таблица 1). Далее моделируем подматрицу Z_1 (Таблица 2, цифры зеленого цвета) из матрицы $Z_{mn}=[Z_1|Z_2]$. Она выбирается из множества выборок $Z^{(t,\ell)}{}_{m(n-1)}$, t=1,…,$\kappa_t<\infty$, ℓ=1,…, $\kappa_\ell=1000<\infty$. Далее на Шаге 3 вычисляем подматрицу $R_{12}=R_{11}\beta$. Здесь вектор коэффициентов регрессии $ß=(ß_1,\ldots,ß_{n-1})^T$ задан и при n=4 равен β_1= - 0.34787; β_2= -0,22542;β_3=0.63790 [1]. Их значения (компоненты вектора ß и элементы подматрицы R_{12}) приведены в Таблице 1.

Теперь у нас имеются все данные для решения Оптимизационной задачи №5 (Рисунки 1,2). В разработанной программе-таблице (Таблица 1) вво дим в ее ячейки значения элементов системы линейных и одного нелинейного уравнений. Решаем эту систему методом Ньютона, применяя надстройку «Поиск решения» в ЭТ. Параметры имеют вид Рисунок 4. нажав на кнопку «Выполнить» найдем элементы подматрицы Z_2, зависящую от выборки Z_1 и от матрицы R_{12}. (Таблица 1). Многомерная Λ-выбор ка $Z_{mn}=[Z_1|Z_2]$ (одна из бесконечного множества) z–переменных, полученная нами при решении ОЗ МЛРА имеет вид, приведенный в Таблице 2.

Таблица 1

Модельные значения элементов подматриц коэффициентов корреляции R_{11}, R_{12}, реальные значения коэффициентов регрессии $ß=(ß_1,\ldots,ß_3)^T$

1	2	3	4	ß
1,0000	**0,8963**	**0,9793**	**0,7471**	**-0,347870**
0,8963	**1,0000**	**0,9071**	**0,6797**	**-0,225420**
0,9793	**0,9071**	**1,0000**	**0,8447**	**0,637900**
0,7471	0,6797	0,8447	1,0000	

Таблица 2

Модельные значения элементов подматриц $Z_1|Z_2$ из матрицы $Z_{mn}=[Z_1|Z_2]$, имеющих заданные корреляционные матрицы $R_{11}\rfloor R_{12}$, заданный вектор коэффициентов регрессии $ß=(ß_1,\ldots,ß_3)^T$

Values of the z-variables z_1, z_2, z_3, z_4			
1	2	3	4
0,263397	0,448782	-0,069869	-1,445422
-0,070078	-0,316994	-0,155151	0,407357
0,865076	1,593805	0,796831	0,164729
-0,457251	-0,692492	-0,482185	-0,306252
-0,275975	-0,193384	-0,283746	-0,066682
-0,09438	0,153904	0,002336	-0,180427
-0,620163	-0,664282	-0,498737	0,444775
-0,625477	-0,407686	-0,423976	-0,067232
0,143165	-0,72528	-0,444542	-1,334659
-0,83811	-0,837773	-0,607355	0,031216
-0,712725	-0,996483	-0,67148	-0,781625
0,343139	1,434746	0,466601	0,35025
-0,088094	-0,34599	-0,149825	-0,244908

-0,35595	-0,363648	-0,379696	-0,073036
3,596704	2,812559	3,72064	3,427547
-0,299958	-0,164258	-0,327337	-0,129438
-0,77332	-0,735527	-0,492509	-0,196194
1,000000	1,000000	1,000000	1,000000

Таблица 3

Программа-таблица

Полученное решение			Начальные значения		
1	1	1,00	1	1	1,00
2	0,074801	0,074801	2	12,70094	0,074801
3	0,041463	0,041463	3	11,55418	0,041463
4	0,092738	0,092738	4	14,36062	0,092738
5	1	1	5	1	1

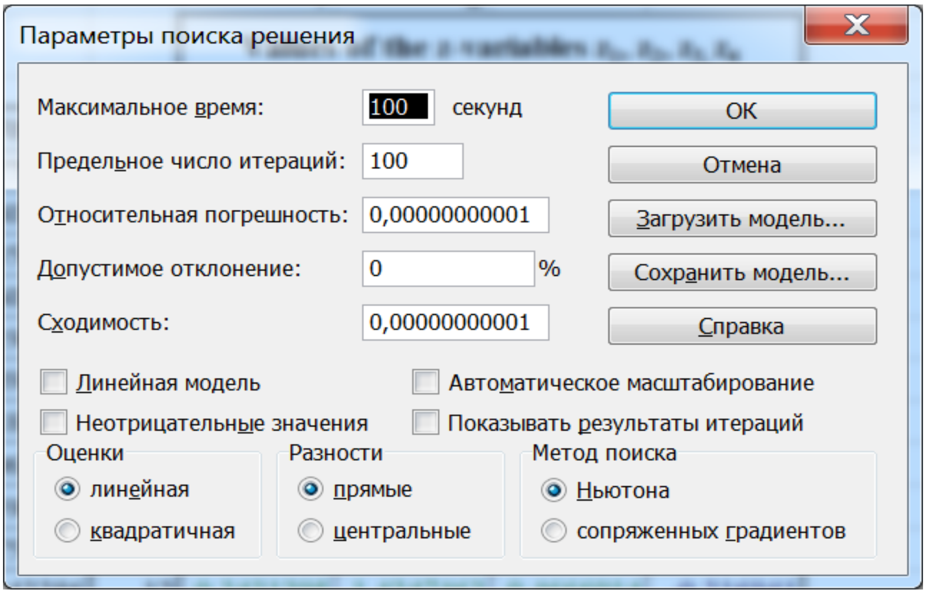

Рисунок 1 Настройки параметров процедуры «Поиск решения» в ЭТ Excel для программы-таблицы в ЭТ Excel с формулами целевой функции оптимизационной задачи №5 (n=4) и 3-х функций ограничений

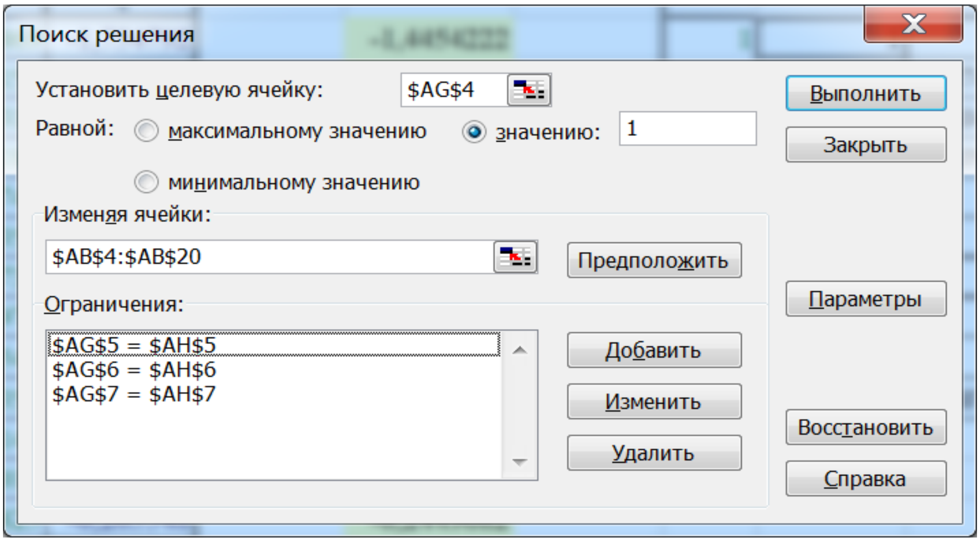

Рисунок 2 Окно процедуры «Поиск решений» в ЭТ Excel с введенными формулами целевой функции оптимизационной задачи №5 (n=4) и 3-х функций ограничений

Заключение

В работе решены задачи моделирования динамики n коррелированных показателей, каждый из которых состоит из m значений, т.е. моделируются m*n–матрицы, интерпретируемые как многомерные выборки объема m. Отличие данного способа моделирования состоит в том, что моделируется не вектор (как принято в методах Монте-Карло), а матрица как единое целое. Второе отличие - многомерная выборка объема m рассматривается как m реализаций n-мерной переменной, при этом для выборки объема m не определен закон распределения генеральной совокупности, из которого извлечена выборка объема m. Зато выборочные средние, выборочные стандартные отклонения, выборочные коэффициенты, выборочные коэффициенты эластичности в точности равны заданным значениям.

Мы конструировали ОМ МЛРА, обозначили постоянные величины (целые, вещественные) и переменные величины (детерминированные, случайные одномерные), воспользовались известными из ПМ МЛРА уравнениями, равенствами. Неизвестный вектор $\beta=(\beta_1,\dots,\beta_{n-1})^T$ коэффициентов регрессии из ПЗ МЛРА объявили в ОЗ МЛРА известным параметром. Постановки обратных задач (ОЗ АГК, ОЗ МЛРА) являются новыми. Известные популярные задачи многомерной прикладной статистики – анализ главных компонент и множественный линейный регрессионный анализ, мы назвали прямыми задачами,, а исследования с их применением: теоремы, задачи, приложения в предметных областях, алгоритмы, компьютерные программы, результаты расчетов по реальным данным, выводы, цифровые знания, новые ключеве слова,- составляют новому научному направлению в прикладной статистике. В рамках этого направления в данной статье разработана сложная математическая модель: ОМ МЛРА. Предстоит проводить глубокий анализ связанных с ОЗ МЛРА, с ОЗ АГК проблем. Исследовать обнаруженные новые явления, объекты, проявления их свойств и разработать алгоритмические методы работы с ними, выходить на новые знания и технологии - виртуальные базы данных [ ], виртуальные лаборатории. Чтобы успешно и всесторонне осмыслить существующие в этих моделях объекты, явления, процессы необходимо рассмотреть вопросы при переходе от безразмерных значений z-переменных к значениям x^0–переменных (исходных переменных). Здесь имеем дело с значениями средних арифметических независимых и одной n-1 зависимой переменной $\mathbf{x}_{ср}=(x_1^{ср},\dots,x_{n-1}^{ср},x_n^{ср})$,, с их стандартными отклонениями, с n-1 эластичностями переменной x_n по x-переменным $x_1,\dots,\mathbf{x}_n$., $x_{ij}= x^0_{ij}–x_j$ $i=1,\dots,m,n=1,\dots,n$. Это позволит оценивать приращения значения $\mathbf{x}_n$ при заданном приращении одной из независимых x-переменных. выходить на новые рубежи знаний и технологий.

Выше мы конструировали ОМ МЛРА - обозначили постоянные величины (целые, вещественные) и переменные величины (детерминированные, случайные одномерные), воспользовались известными из ПМ МЛРА уравнениями, равенствами. Неизвестный вектор $\beta=(\beta_1,\dots,\beta_{n-1})^T$ коэффициентов регрессии из ПЗ

МЛРА объявили в ОЗ МЛРА известным параметром. Постановки обратных задач (ОЗ АГК, ОЗ МЛРА) для известных популярных задач многомерной прикладной статистики – анализ главных компонент и множественный линейный регрессионный анализ -, их задачи мы назвали прямыми задачами, являются новыми, а исследования с их применением: теоремы, задачи, приложения в других предметных областях, алгоритмы, компьютерные программы , результаты расчетов по реальным данным, выводы, цифровые знания, новые ключеве слова,- составляют новое научное направление в современной прикладной статистике. В этом направлении разработанная в данной статье математическая модель ОМ МЛРА найдет свое применение. Предстоит проводить глубокий анализ связанных с ОЗ МЛРА, с ОЗ АГК проблем.

Предыдущий и последующий члены ряда векторов значений коэффициентов регрессии нужно подбирать вручную и в соответствии с реальными производственными, маркетинговыми, административными мероприятиями, обеспе чивающих достижение планируемых финансовых приростов значений ресурсов, вложений. Если нет реального менеджмента в банке по достижению планируемых значений $ß_1,\ldots,ß_{n-1}$, то нет пользы от модельных значений элементов матрицы $Z_{mn}=[Z_1 \rfloor Z_2]$. Матрица $Z_{mn}=[Z_1 \rfloor Z_2]$ при известных значениях средних и стандартных отклонений преобразуются в «реальные» значения с единицами измерения показателей с номерами $j=1,\ldots,n$. Каждая цифра в «реальной» матрице адекватна реальному значению, если значения n z–переменных удовлетворяют системе уравнений (*). При этом практические решения по принятию предыдущего значения и последующего значения должны быть подвергнуты всестороннему анализу. Чтобы осмыслить изучаемые в этих моделях объекты, явления, процессы необходимо рассмотреть вопросы при переходе от безразмерных значений z-переменных к значениям x^0–переменных (исходных переменных) с размерностями. Здесь предстоит нам имееть дело с значениями средних арифметических независимых и одной n-1 зависимой переменной $\mathbf{x}_{cp}=(x_1^{cp},\ldots,x_{n-1}^{cp},x_n^{cp})$, с их стандартными отклонениями, с n-1 эластичностями переменной x_n по x-переменным $x_1,\ldots,\mathbf{x}_n$., $x_{ij}=x^0_{ij}–x_j$ $i=1,\ldots,m$, $j=1,\ldots,n$. Это позволит оценивать приращение значения $\mathbf{x}^0_n$ при заданном приращении значения одной независимой x^0-переменной.

Предстоит исследовать обнаруженные новые явления, объекты, проявления их свойств и разработать алгоритмические методы работы с ними, выходить на новые рубежи знаний и технологий- виртуальные базы данных [14], виртуальные лаборатории [13]. Практически важной задачей является исследование связей коэффициентов эластичности $e_j=(x^{cp}_n/x^{cp}_j)ß_j$ с коэффициентами регрессии $ß_1,\ldots,ß_{n-1}$.

References:

1. Zhanatauov S.U. (2018) Primenenie metoda lineynoy regressii dlya samootsenki riska bankrotstva banka. Finansovaya analitika: problemy i resheniya.
2. Zhanatauov S.U. (1987) Obratnaya model' glavnykh komponent i ee primenenie. Diss. na soiskanie uch. step.. kand. fiz.-mat. nauk:05.13.11:zashchishchena 8.12.1987: utv. 1.06.1988/Zhanatauov Sapargali Utepovich-Vychislitel'nyy tsentr Sibirs kogo otdeleniya AN SSSR, Novosibirsk, 1987g.,302 p.
3. Zhanatauov S.U. (1980) Metod polucheniya vyborki s zadannymi sobstvennymi chislami ee korrelyatsionnoy matritsy. - V kn. Matematicheskie voprosy analiza dannykh. Novosibirsk, 1980, p.62-76.
4. Zhanatauov S.U. (1989) Modelirovanie odnoy zamechatel'noy ekstremal'noy sovokupnosti //Sistem noe modelirovanie-14, -Novosibirsk. 1989. p.27- 33.
5. Zhanatauov S.U. (2013) Obratnaya model' glavnykh komponent:-monografiya. -Almaty: Kazstatin form, 2013.- 201 p.
6. Zhanatauov SU (2017) Optimization problem of modeling missing elements of the spectrum of the correlation matrix. International scientific journal Theoretical&Applied Science.2017, №10, vol.54, p.189-198. http://www.T-Science.org.
7. Zhanatauov SU. (2017) The optimization problem with linearized equations f-parameters (f1,f2,f3,f4,f5,f6)-spectrum. International scientific journal Theoretical &Applied

Science. 2017,№11,vol.55, p.251-267. http://www.T-Science.org.

8. Zhanatauov SU (2017) Block-diagonal correlation matrices of λ-samples. International scientific journal Theoretical &Applied Science. 2017,№12,vol.56, p.101-111. http://www.T-Science.Org
9. Zhanatauov SU (2017) Theorem on the Λ-samples. International scientific journal Theoretical &Applied Science.2017,№ 9, vol.53, p.177-192. http://www.T-Science.Org .
10. Zhanatauov S,U. (1988) O funktsional'nom napolnenii PPP "Spektr". Sistemnoe modelirovanie -13 .-Novosibirsk, 1988, p.3-11.
11. Zhanatauov S.U. (1987) Dialogovyy paket programm modelirovaniya spektra neizvestnoy korrelyatsionnoy matritsy.//Dialogovye sistemy v zadachakh upravle niya. Novosibirsk, 1987.-p.157-163.
12. Chalmers C.P. (1975) Generation of correlation matrices with a given eigen – structure.-J. Stat. Comp. Simul., 1975, vol.4, p.133-139.
13. Zhanatauov SU (2018) Virtual database. International scientific journal Theoretical &Applied Science. 2018,№ 2,vol.58, p.187-198. http://www.T-Science.Org .
14. Zhanatauov S.U. (2011) Virtual'naya laboratoriya. Mater. Vseross. nauchno-praktich. konf. «Innovatsii v nauke–puti razvitiya» – Cheboksary: 2011. -p.33-44. https://ru.wikipedia.org