

УДК 51:155.001.57:681.3.06

АЛГОРИТМ "ZET-75" ЗАПОЛНЕНИЯ ПРОБЕЛОВ В
ЭМПИРИЧЕСКИХ ТАБЛИЦАХ

Н.Г.Загоруйко, В.Н.Ёлкина, В.С.Тимеркаев

В работе [I] был предложен алгоритм заполнения пробелов в эмпирических таблицах (алгоритм " ZET "). Имеющийся к настоящему времени опыт эксплуатации программы, реализующей этот алгоритм, позволяет сделать следующие выводы:

I. В различных прикладных областях имеется большое число эмпирических таблиц (протоколов), содержащих пробелы, которые было бы желательнее заполнить правдоподобными ("как можно более правильными") значениями. Такие таблицы встречаются в медицине, геологии, почвоведении, океанологии, характеристиках технических устройств и т.д.

II. Помимо заполнения пробелов (прогнозирования пропущенных значений) алгоритм " ZET " можно использовать для автоматического обнаружения ошибок в таблицах (редактирование таблиц). С этой целью поочередно "прогнозируются" все (или выборочные) элементы таблицы, и для дальнейшей проверки отбираются те элементы, для которых предсказанные значения сильно отличаются от фактически имеющихся в таблице.

III. Естественно использовать алгоритм " ZET " и для распознавания образов, т.е. определения принадлежности объекта (строки таблицы) к одному из известных образов, - прогнозирования целевого признака.

IV. Избыточность информации в реальных таблицах такова, что ее, как правило, оказывается достаточно для хорошего прогнозирования даже относительно большого числа пропущенных элемен-

тов. Хорошие результаты были получены для всех реальных таблиц, для которых применялся алгоритм "ЗЕТ". Таких таблиц различной природы было более 10, и количество пробелов в некоторых из них достигало 30%.

Однако программа, реализующая предыдущий вариант алгоритма "ЗЕТ", требовала больших затрат времени и памяти. Поскольку задача заполнения пробелов актуальна и основная идея алгоритма "ЗЕТ" хорошо согласована с особенностями реальных задач, нами была разработана новая версия алгоритма заполнения пробелов в эмпирических таблицах - алгоритм "ЗЕТ-75". Данный алгоритм оказался эффективнее с точки зрения сокращения затрат машинного времени и памяти при той же "прогнозирующей способности", что и алгоритм "ЗЕТ".

Основные отличия этой версии алгоритма состоят в следующем:

1. Как и в алгоритме "ЗЕТ", в прогнозировании пропущенного элемента участвуют только "компетентные" строки и столбцы. Но если в прежнем варианте "компетентные" группы строк и столбцов отбирались с помощью алгоритмов таксономии, то теперь их "компетентность" определяется как функция некоторой меры близости между строками (столбцами) матрицы и строкой (столбцом), содержащей искомый пробел *).

2. "Компетентность" строк (столбцов) зависит также и от того, много ли пропусков содержат сами эти строки (столбцы), т.е. от степени их заполненности. Естественно, что компетентность прогнозирующих строк (столбцов) тем выше, чем ближе они к предсказываемым и чем больше взаимно непустых элементов они содержат.

3. В алгоритме "ЗЕТ-75" в качестве меры близости как между столбцами, так и между строками матрицы принято значение модуля коэффициента парной корреляции, вычисляемого после нормировки всех столбцов матрицы к интервалу [0,1].

4. При вычислении прогнозируемого элемента учитывается "вклад" каждой строки (столбца), зависящий от степени ее запол-

* Ясно, что в "компетентные" группы отбираются только строки (столбцы), по которым можно прогнозировать искомый пробел, т.е. не имеющие пробела на месте, соответствующем искомому пробелу.

ненности, близости к строке (столбцу) с искомым пробелом и от некоторого параметра α . Значение параметра α подбирается в процессе решения для каждого пробела из некоторого заданного диапазона значений. Критерием для выбора α служит минимум средней ошибки предсказания всех известных элементов строки (столбца), содержащей искомый пробел.

5. В алгоритме "ЗЕТ-75" оценка "качества прогнозирования" делается по точности восстановления всех известных элементов строки (столбца), содержащей этот пробел.

Процесс работы алгоритма "ЗЕТ-75" заключается в следующем. Пусть дана матрица размерности $n \times m$, где n - число столбцов (признаков), m - количество строк (объектов).

0. Производится нормировка всех столбцов матрицы исходных данных к интервалу [0,1].

1. Выделяется очередной пробел a_{ij} , который нужно прогнозировать.

	1	...	k	j	...	p
1						
...						
i				a_{ij}		
...						
l						
...						
m						

2. Для каждого k -го столбца, не имеющего пробелов в i -й строке, вычисляется его мера заполненности L_{jk} ($j \neq k$) по отношению к j -у столбцу, равная числу взаимно непустых пар элементов j -го и k -го столбцов.

3. В этом совместном пространстве вычисляется мера близости r_{jk} столбцов j и k .

4. Прогнозируются при различных α значениях все известные элементы a_{ij} j -го столбца по столбцам, не имеющим пробела в i -й строке.

$$a_{ij} = \frac{\sum_{k=1}^p a_{ij}^k |r_{jk}|^\alpha \cdot L_{jk}}{\sum_{k=1}^p |r_{jk}|^\alpha \cdot L_{jk}} ; \quad (I)$$

$$i = 1, \dots, m; \quad i \neq i; \quad k = 1, \dots, p,$$

где p - количество столбцов с ненулевой дисперсией и без пробела в i -й строке; r_{jk} - коэффициент корреляции j и k столбцов в совместном подпространстве; a_{ij}^k - значение элемента j -го столбца, прогнозируемое по элементам k -го столбца: $a_{ij}^k = b_{jk} \cdot a_{ik} + c_{jk}$; b_{jk} , c_{jk} - коэффициенты уравнения линейной регрессии.

Для случая, когда дисперсия j -го или k -го столбца в совместном подпространстве равна 0, в алгоритме предусмотрены следующие решения:

а) $\sigma_j \neq 0, \sigma_k = 0: r_{jk} = 0$, элемент a_{ij}^k не предсказывается.

б) $\sigma_j = 0, \sigma_k \neq 0, a_{1j} \neq 0, a_{1j} \neq 1: a_{ij}^k = a_{ij}$. В формуле (I) для этого случая $r_{jk} = 1$. При $a_{1j} = 0$ или $a_{1j} = 1$ элемент a_{ij}^k по столбцам не предсказывается. Если же и при проверке строк оказывается, что этот элемент находится в сплошь нулевой (или сплошь единичной) строке, то $a_{ij}^k = 0$ или $a_{ij}^k = 1$ соответственно. Иначе предсказание производится по строкам на основании уравнения регрессии.

в) $\sigma_j = \sigma_k = 0, a_{1j} \neq 0: a_{ij}^k = \frac{a_{1k}}{a_{1j}} \cdot a_{ik}, r_{jk} = 1$.

г) $\sigma_j = \sigma_k = 0, a_{1j} = 0: a_{ij}^k = 0$.

5. Для каждого α вычисляется средняя ошибка $\delta^0(\alpha)$ прогноза известных значений a_{ij} .

6. Выбирается то α , при котором $\delta^0(\alpha) = \min \delta^0(\alpha)$.

7. При $\alpha = \alpha(\delta_{min}^0)$ вычисляется искомое прогнозируемое значение a_{ij}^0 .

8. Выполняется аналогичная пп. 2-7 процедура прогнозирования по строкам. В результате этого находятся значения параметра $\alpha^* = \alpha(\delta_{min}^*)$, при котором получена минимальная ошибка прогнозирования известных элементов i -й строки и предсказано значение a_{ij}^* .

9. Проверяются условия $\delta_{min}^0 \leq \delta_0, \delta_{min}^* \leq \delta_0$, где δ_0 - заданная точность. Если ни одно из значений δ_{min}^0 или δ_{min}^* не удовлетворяют условию заданной точности, то пробел остается незаполненным. Если оба значения удовлетворяют заданной точности, то для прогнозирования выбирается то, которое соответствует

$\delta = \min(\delta^0, \delta^*)$. В случае равенства ($\delta_{min}^0 = \delta_{min}^*$) $\leq \delta_0$ для прогноза a_{ij} вычисляется величина

$$a_{ij} = \frac{a_{ij}^0 + a_{ij}^*}{2}$$

10. После окончания заполнения всех пропущенных элементов в таблице повторяются пп. 2-9.

В алгоритме предусмотрены две возможности заполнения элементов:

а) прогнозируемое значение ставится сразу в таблицу и учитывается при заполнении последующих пробелов;

б) на первой итерации прогнозируются все пропущенные значения без учета уже предсказанных.

II. После каждой очередной итерации производится оценка среднего суммарного отличия от результатов прогнозирования, полученных на предыдущем шаге.

Процесс прекращается, если это отличие становится несущественным.

Предусмотрен также выход из цикла по количеству итераций.

Программа "ZET-75"

Обращение к подпрограмме "ZET-75" производится оператором ФОРТРАНА

CALL ZET 75 (M1, M2, M3, M4, M5, M6, GAP, BEG, FIN, STEP, REST, X1, X2, X3, K2, K3),

где M1 - количество столбцов в таблице;

M2 - количество строк в таблице;

M3 - max { M1, M2 };

M4^{*}) - количество пробелов;

M5^{*}) - const, при |M5| \neq 7 выдается стандартная информация, если |M5| = 7, то на печать выдается и дополнительная информация (см. описание выдачи результатов, стр. 27);

M6^{*}) - количество итераций;

GAP - значение, отмечающее пробел в таблице исходных данных;

*) В программе используются абсолютные величины параметров M4, M5, M6. Знаки этих параметров необходимы для задания режимов работы программы.

BEG - начальное значение параметра α ;
 FIN - конечное значение параметра α ;
 STEP^{*)} - шаг изменения параметра α ;
 REST - заданная точность прогнозируемого значения ;
 X1 - массив исходных данных размерности $M1 \cdot M2$;
 X2 - рабочий массив размерности $10 \cdot M3$;
 X3 - рабочий массив размерности $9 \cdot M4$;
 K2 - рабочий массив размерности $10 \cdot M3$;
 K3 - рабочий массив размерности $9 \cdot M4$.

В управляющей программе с помощью оператора EQUIVALENCE массивы (X2, K2) и (X3, K3) должны быть объявлены эквивалентными, т.е. EQUIVALENCE (X2, K2), (X3, K3).

Параметры M4, M6, STEP в подпрограмме "ZET-75" переопределяются, так что в управляющей программе им должны соответствовать идентификаторы.

Возможны 8 режимов использования подпрограммы: 4 варианта заполнения пропущенных значений и 4 варианта редактирования значений исходной матрицы.

Режимы прогнозирования пропущенных значений

1. Алгоритм "ZET-75" находит все пропущенные элементы, и на первой итерации прогнозирование каждого значения выполняется независимо от остальных. На последующих итерациях при прогнозировании учитываются результаты заполнения пробелов на предыдущей итерации. Режим выполнения этой процедуры задается условиями:

$$M4 > 1, M5 \geq 0, M6 > 0, STEP \geq 0.$$

2. Алгоритм находит все пропущенные элементы. Начиная с первой итерации, заполнение каждого последующего пробела ведется с учетом всех найденных к этому моменту элементов. Режим выполнения процедуры задается условиями:

$$M4 > 1, M5 \geq 0, M6 < 0, STEP \geq 0.$$

3. Прогнозируются только указанные пропущенные элементы. Для этого в управляющей программе в массивы K3(1, M4), K3(2, M4) должны быть занесены номера столбцов и строк, на пересечении

*) В программе используется абсолютная величина параметра STEP. Знак этого параметра необходим только для задания режимов работы программы.

которых находятся элементы, подлежащие прогнозированию. Элементы, не включаемые в этот список, хотя и имеют значение пробела GAP, восстанавливаться не будут. Восстановление значений ведется так же, как в режиме 1. Условия задания режима:

$$M4 > 1, M5 < 0, M6 > 0, STEP \geq 0.$$

4. Прогнозируются только значения, заданные списком, как в режиме 3. Заполнение каждого из указанных элементов производится аналогично режиму 2. Условия задания режима:

$$M4 > 1, M5 < 0, M6 < 0, STEP \geq 0.$$

Режимы редактирования^{*)}

а) Редактирование каждого из элементов таблицы без внесения "исправленного" значения в исходную матрицу. Условия режима:

$$M4 = 1, M6 < 0, STEP \geq 0, M5 \geq 0.$$

б) Проверка всех элементов с автоматическим "исправлением" таблицы. Условия:

$$M4 = 1, M6 > 0, STEP \geq 0, M5 \geq 0.$$

в) То же, что а), но для отдельных элементов таблицы. Условия:

$$M4 > 1, M5 \geq 0, M6 < 0, STEP < 0.$$

г) То же, что б), но для отдельных элементов таблицы. Условия:

$$M4 > 1, M5 \geq 0, M6 > 0, STEP < 0.$$

В случаях в) и г) список элементов, подлежащих проверке, должен быть сформулирован, как в режиме 3.

Для всех режимов редактирования на печать выдается информация о тех элементах, прогнозное значение которых сильно отличается от фактически имеющихся в таблице.

Общее рабочее поле подпрограммы равно

$$K = (M1 \cdot M2) + (9 \cdot M4) + (10 \cdot M3).$$

Подпрограмма "ZET-75" имеет обращение к следующим подпрограммам:

1) Подпрограмма печати исходных данных MATP. Вызывается оператором

CALL MATP (M1, M2, M3, M, X1, K2),

*) В режиме редактирования, независимо от значения |M6|, выполняется только одна итерация.

где $M = \text{const}$, при $M = 0$ печатается матрица с заголовком "МАТРИЦА ИСХОДНЫХ ДАННЫХ", при $M \neq 0$ - с заголовком "РЕЗУЛЬТАТ".
Остальные параметры, как и в описанных ниже подпрограммах, выполняют ту же функцию, что и в подпрограмме "ZET-75". Подпрограмма MATP в процессе работы использует первый столбец матрицы X2. Параметры не переопределяются.

2) Подпрограмма нормализации и денормализации исходных данных NORM1. Вызывается оператором:

CALL NORM1 (M1, M2, M3, X1, X2, GAP, M),

где $M = \text{const}$. Если $M = 0$, то производится нормализация исходных данных к интервалу $[0, 1]$; при $M \neq 0$ производится денормализация данных - возврат ранее нормализованных значений к исходному виду. Минимальные и максимальные значения ($x_{j, \min}$, $x_{j, \max}$) для столбцов исходной матрицы хранятся соответственно в 6 и 7 столбцах матрицы X2. Остальные столбцы матрицы X2 программой NORM1 не используются. Параметры не переопределяются.

3) Подпрограмма вычисления коэффициентов корреляции r_{jk} и линейной регрессии (b_{jk}, c_{jk}) - MEPA1. Вызывается оператором

CALL MEPA1 (MK, M3, XA, XB, XC, X2),

где MK - количество взаимно ненулевых элементов (a_{1k}, a_{1j}) в столбцах j, k ; XA, XB, XC - константы; в XA, XB, XC заносятся результаты работы подпрограммы - значения r_{jk}, b_{jk}, c_{jk} соответственно. Подпрограмма MEPA1 использует в процессе работы первый и второй столбцы матрицы X2, в которых находятся a_{1j} и a_{1k} .

4) Подпрограмма вычисления прогнозируемого значения по строкам и столбцам PRSGN. Вызывается оператором
CALL PRSGN(M1, M2, M3, M4, MM, GAP, BEG, FIN, STER, REST, X1, X2, X3, K2, K3),
где MM - const, хранящая порядковый номер прогнозируемого элемента. При редактировании MM не учитывается.

5) Подпрограмма вывода информации на печать EXPT. Вызывается оператором

CALL EXPT (M1, M2, M3, M4, M5, MM, GAP, REST, FU, F, X1, X2, X3, K2, K3),

где MM - const та же, что и в предыдущей подпрограмме; FU - const, в нее заносится прогнозируемое значение, если $FU \leq REST$; F - const, в ней должно быть значение исходной матрицы, подлежащее проверке.

Подготовка исходных данных. Исходные данные готовятся согласно инструкции подготовки данных ФОРТРАНА IV и вводятся построчно. Вместо пробела ставится элемент GAP, отличный от элементов таблицы, например, $GAP = 8 \cdot 10^{17}$.

Вывод информации

1. Выводится на печать матрица исходных данных.
 2. Печатается номер итерации и информация о прогнозировании пропущенного элемента по строке, столбцу и по средней взвешенной величине. При редактировании параллельно с прогнозируемым значением печатается соответствующее значение исходной матрицы.
 3. Если $|M5| = 7$, то дополнительно печатается список строк (столбцов), внесших наибольший вклад в прогнозирование элемента.
 4. Если процесс не окончен, то повторяется пп. 2-3, в противном случае печатается матрица исправленных данных.
- В приложении приводится текст комплекса программ "ZET-75" и результаты решения для контрольного примера.

Литература

1. ЗАГОРУЙКО Н.Г., ЕВКИНА В.Н., ТИМЕРКАЕВ В.С. Алгоритмы заполнения пропусков в эмпирических таблицах (алгоритм "ZET"). - В кн.: Вычислительные системы. Вып. 61, Новосибирск, 1975, с. 3-27.

Поступила в ред.-изд. отд.
27 апреля 1976 года

ПРИЛОЖЕНИЕ I

ПРОГРАММА

```

SUBROUTINE ZET 75
*(M1,M2,M3,M4,M5,M6,GAP,BEG, FIN,STEP,REST,X1,X2,X3,K2,K3)
DIMENSION X1(M1,M2),X2(10,M3),X3(9,M4),K2(10,M3),K3(9,M4)
CALL MATP(M1,M2,M3,0,X1,K2)
CALL HOPM1(M1,M2,M3,X1,X2,GAP,0)
MX=M6
M6=IABS(M6)
MP=0
MM=0
XT=-7.
F=0.
MK=STEP
STEP=ABS(STEP)
M=M4
IF(M4.GT.1.AND.M5.LT.0) GO TO 2
IF(M4.BQ.1) GO TO 3
IF(MK.LT.0) GO TO 2
M=0
DO 1 J=1,M2
DO 1 I=1,M1
IF(X1(I,J).NE.GAP) GO TO 1
M=M+1
K3(1,M)=I
K3(2,M)=J
1 CONTINUE
2 MP=MP+1
PRINT 50,MP
IF(MK.LT.0) GO TO 80
PRINT 60
DO 99 MM=1,M
I=K3(1,MM)
J=K3(2,MM)
X1(I,J)=GAP
CALL PROGN(M1,M2,M3,M4,MM,GAP,BEG,FIN,STEP,REST,X1,X2,X3,
*K2,K3)
CALL EXPT(M1,M2,M3,M4,M5,MM,GAP,REST,FU,F,X1,X2,X3,K2,K3)

```

```

IF(FU.BQ.XT) GO TO 99
IF(MX.GT.C.AND.MP.BQ.1) GO TO 98
X1(I,J)=FU
GO TO 99
98 X3(8,MM)=FU
99 CONTINUE
I=(MP.NE.1.OR.MX.LT.0) GO TO 97
DO 96 MM=1,M
IF(X3(8,MM).BQ.XT) GO TO 96
I=K3(1,MM)
J=K3(2,MM)
X1(I,J)=X3(8,MM)
96 CONTINUE
97 PRINT 62
IF(MP.LT.M6) GO TO 2
GO TO 7
3 IF(MX.LT.0) GO TO 91
4 MP=MP+1
PRINT 50,MP
91 PRINT 70
MM=1
DO 93 J=1,M2
DO 93 I=1,M1
MT=1
K3(1,1)=I
K3(2,1)=J
F=X1(I,J)
X1(I,J)=GAP
CALL PROGN(M1,M2,M3,M4,MT,GAP,BEG,FIN,STEP,REST,X1,X2,X3,
*K2,K3)
CALL EXPT(M1,M2,M3,M4,M5,MM,GAP,REST,FU,F,X1,X2,X3,K2,K3)
MM=MM+1
IF(MX.GT.0) GO TO 94
X1(I,J)=F
GO TO 93
94 IF(FU.NE.XT) X1(I,J)=FU
93 CONTINUE
IF(MX.LT.0) GO TO 7

```

```

PRINT 62
IF(MP.LT.M5) GO TO 4
GO TO 7
80 PRINT 70
N=K3(1,1)
L=K3(2,1)
MT=1
DO 89 MM=1,M4
I=K3(1,MM)
J=K3(2,MM)
F=X1(I,J)
X1(I,J)=GAP
K3(1,1)=I
K3(2,1)=J
CALL PROG(M1,M2,M3,MT,MT,GAP,BEG,FIN,STEP,REST,X1,X2,X3,
*K2,K3)
CALL EXPT(M1,M2,M3,MT,M5,MM,GAP,REST,FU,F,X1,X2,X3,K2,K3)
IF(MX.GT.0) GO TO 88
X1(I,J)=F
GO TO 89
88 IF(FU.NE.XT) X1(I,J)=FU
89 CONTINUE
IF(MX.LT.0) GO TO 7
PRINT 62
K3(1,1)=N
K3(2,1)=L
IF(MP.LT.M5) GO TO 2
7 CALL НОРМ1(M1,M2,M3,X1,K2,GAP,1)
CALL МАТР(M1,M2,M3,1,X1,K2)
RETURN
50 FORMAT(/50X,11ИТЕРАЦИЯ N=,I4/)
70 FORMAT(/40X,40Н=====/)
13X,'НОМЕР',6X,'НОМЕР',5X,'НОМЕР',4X,'КОЭФФИЦИЕНТ',4X,
*'ПРОЦЕНТ',
23X,'0-СТРОКА',3X,'ПРЕДСКАЗАННОЕ',4X,'ИСХОДНОЕ',3X,'НОМЕР'/
35X,'П/П',4X,'СТОЛБЦА',4X,'СТРОКИ',20X,'ОШИБКИ',3X,'1-
*СТОЛБЕЦ',
48X,'ЗНАЧЕНИЕ',4X,'ЗНАЧЕНИЕ',5X,'П/П'/58X,'2-СОВМЕЩНО'/)

```

```

60 FORMAT(/40X,40Н=====/)
13X,'НОМЕР',6X,'НОМЕР',5X,'НОМЕР',4X,'КОЭФФИЦИЕНТ',4X,
*'ПРОЦЕНТ',
23X,'0-СТРОКА',3X,'ПРЕДСКАЗАННОЕ',3X,'НОМЕР'/5X,'П/П',4X,
3'СТОЛБЦА',4X,'СТРОКИ',20X,'ОШИБКИ',3X,'1-СТОЛБЕЦ',8X,
4'ЗНАЧЕНИЕ',5X,'П/П'/58X,'2-СОВМЕЩНО'/)
62 FORMAT(/40X,40Н-----/)
END

SUBROUTINE EXPT
*(M1,M2,M3,M4,M5,MM,GAP,REST,FU,F,X1,X2,X3,K2,K3)
DIMENSION X1(M1,M2),X2(10,M3),X3(9,M4),K2(10,M3),K3(9,M4)
J=MM
D=GAP
G=GAP
FU=-7.
XT=FU
IF(M4.EQ.1) J=1
N=K3(1,J)
L=K3(2,J)
I=K3(7,J)
X=X2(6,N)-X2(7,N)
P=X2(7,N)
IF(L.GE.0) GO TO 2
PRINT 61,J,N,L,G,G,I,G,J
GO TO 8
2 IF(L.GE.1) GO TO 3
IF(X3(6,J).NE.XT) GO TO 6
A=X3(3,J)
C=X3(5,J)
T=X3(8,J)*X+P
IF(C.LE.REST) FU=X3(8,J)
GO TO 5
3 IF(L.GT.1) GO TO 6
IF(X3(5,J).NE.XT) GO TO 6
A=X3(4,J)
C=X3(6,J)
T=X3(9,J)*X+P
IF(C.LE.REST) FU=X3(9,J)

```



```

M6=M7+1
8 M7=M7+10
  IF(M11) 20,20,7
7 M7=M7-9
  PRINT 12,(K2(1,I),I=M7,M1)
  DO 19 J=1,M2
  PRINT 13,J,(X1(I,J),I=M7,M1)
19 CONTINUE
20 PRINT 14
  RETURN
10 FORMAT(/50X,'МАТРИЦА ИСХОДНЫХ ДАННЫХ '/')
11 FORMAT(/50X,'РЕЗУЛЬТАТ '/')
12 FORMAT(/10X,10I10/10X,'*****
1*****
2*****')
13 FORMAT(2X,14,2X,2H* ,10E10.3)
14 FORMAT(10X,'*****
1*****
2*****')
  END

SUBROUTINE HOPM1
*(M1,M2,M3,X1,X2,CT,MP)
DIMENSION X1(M1,M2),X2(10,M3)
IF(MP.NE.0) GO TO 1
DO 99 I=1,M1
P1=0.
P2=9.E18
DO 98 J=1,M2
IF(X1(I,J).GE.P2.OR.X1(I,J).EQ.CT) GO TO 98
P2=X1(I,J)
98 CONTINUE
DO 96 J=1,M2
IF(X1(I,J).LE.P1.OR.X1(I,J).EQ.CT) GO TO 96
P1=X1(I,J)
96 CONTINUE
IF(P1-P2) 93,92,93
92 DO 91 J=1,M2

```

```

IF(X1(I,J)-CT) 88,91,88
88 X1(1,J)=1.
91 CONTINUE
GO TO 90
93 DO 89 J=1,M2
  IF(X1(I,J)-CT) 87,89,87
87 X1(I,J)=X(X1(I,J)-P2)/(P1-P2)
89 CONTINUE
90 X2(6,I)=P1
  X2(7,I)=P2
99 CONTINUE
GO TO 2
  1 DO 3 I=1,M1
    DO 3 J=1,M2
      IF(X1(I,J)-CT) 4,3,4
  4 X1(I,J)=X1(I,J)*(X2(6,I)-X2(7,I))+X2(7,I)
  3 CONTINUE
  2 RETURN
  END

SUBROUTINE MEPA1
*(MK,M3,XA,XB,XC,X2)
DIMENSION X2(10,M3)
T=0.
TK=T
CL=T
PK=T
BY=T
EX=T
EY=T
R=T
D=T
XA=T
XB=T
XC=T
Z=T
DO 1 J=1,MK
Y=X2(1,J)

```

```

X=X2(2,J)
TX=TX+X
TY=TY+Y
BX=BX+X*X
T=T+X+Y
1 CONTINUE
CX=TX/MK
CY=TY/MK
PX=TX*TX
P=TX*TY
DO 2 J=1,MK
X=X2(2,J)
Y=X2(1,J)
AX=CX-X
AY=CY-Y
R=R+AX*AY
EX=EX+AX*AX
EY=EY+AY*AY
2 CONTINUE
IF(EX.NE.Z.AND.EY.NE.Z) GO TO 4
IF(EX.EQ.Z.AND.EY.EQ.Z) GO TO 3
IF(EX.EQ.Z.AND.EY.NE.Z) GO TO 5
XA=Z
XB=Z
XC=CY
GO TO 7
5 XA=Z
XB=Z
XC=Z
GO TO 7
3 IF(X2(1,1).NE.Z) GO TO 8
XA=1.
XB=Z
XC=CY
GO TO 7
8 XA=1.
XB=CX/CY
XC=Z

```

20

```

GO TO 7
4 XA=R/SQRT(EX*EY)
P=MK*T-P
T=MK*BX-PX
IF(P.EQ.Z.OR.T.EQ.Z) GO TO 6
XB=P/T
XC=CY-XB*CX
GO TO 7
6 XB=1.
XC=Z
7 CONTINUE
RETURN
END

```

```

SUBROUTINE PROGN
*(M1,M2,M3,M4,MM,GAP,BEG,FIN,STEP,REST,X1,X2,X3,K2,K3)
DIMENSION X1(M1,M2),X2(10,M3),X3(9,M4),K2(10,M3),K3(9,M4)
XT=-7.
X=0.
MN=0
N=0
DO 9 M=3,6
9 X3(M,MM)=XT
DO 1 M=1,2
IF(M.EQ.2) GO TO 99
L1=M1
L2=M2
MA=K3(1,MM)
MB=K3(2,MM)
GO TO 98
99 L2=M1
L1=M2
MB=K3(1,MM)
MA=K3(2,MM)
98 MP=-1
MZ=0
MT=0
MR=-1

```

21

```

MQ=-1
IF(M.EQ.2) GO TO 20
DO 23 K=1,2
XK=K-1
DO 15 J=1,I2
IF(X1(MA,J).EQ.GAP) GO TO 15
IF(X1(MA,J).NE.XK) GO TO 23
MZ=MZ+1
15 CONTINUE
MP=K-1
23 CONTINUE
16 DO 12 K=1,2
XK=K-1
DO 13 I=1,I1
IF(X1(I,MB).EQ.GAP) GO TO 13
IF(X1(I,MB).NE.XK) GO TO 12
MT=MT+1
13 CONTINUE
MR=K-1
12 CONTINUE
IF(MP.EQ.MQ.AND.MR.EQ.MQ) GO TO 20
DO 24 K=1,2
MS=K-1
XK=K-1
IF(MP.EQ.MS.AND.MR.EQ.MS) GO TO 19
IF(MP.EQ.MS.AND.MR.EQ.MQ) GO TO 73
IF(MP.EQ.MQ.AND.MR.EQ.MS) GO TO 28
24 CONTINUE
IF(MZ.GE.MT) GO TO 19
K3(7,MM)=0
X3(3,MM)=STEP
X2(5,MM)=X
X3(8,MM)=MR
GO TO 11
73 MN=MN+1
GO TO 1
19 K3(7,MM)=1
X3(4,MM)=STEP

```

```

X3(6,MM)=X
X3(9,MM)=MP
GO TO 11
28 N=1
MN=MN+1
20 IF(M.EQ.2.AND.N.EQ.1) GO TO 1
DO 2 I=1,I1
X2(4,I)=X
XA=X
XB=X
XC=X
MK=0
IF(I.EQ.MA) GO TO 40
DO 3 J=1,I2
IF(J.EQ.MB) GO TO 3
IF(M.EQ.2) GO TO 97
IF(X1(MA,J).EQ.GAP.OR.X1(I,J).EQ.GAP) GO TO 3
MK=MK+1
X2(1,MK)=X1(MA,J)
X2(2,MK)=X1(I,J)
GO TO 3
97 IF(X1(J,MA).EQ.GAP.OR.X1(J,I).EQ.GAP) GO TO 3
MK=MK+1
X2(1,MK)=X1(J,MA)
X2(2,MK)=X1(J,I)
3 CONTINUE
IF(MK.LT.2) GO TO 40
CALL MEPA1(MK,M3,XA,XB,XC,X2)
X2(4,I)=XA
XA=MK
X2(5,I)=XA/I2
X2(3,I)=XB
X2(8,I)=XC
GO TO 2
40 X2(4,I)=XT
X2(5,I)=XT
2 CONTINUE
DO 52 I=1,I1

```

```

IF(X2(5,I).NE.XT) GO TO 53
52 CONTINUE
MN=MN+1
GO TO 1
53 XA=BEG
XX=9.B17
7 DO 4 J=1,L2
X2(1,J)=X
X2(2,J)=X
MK=0
IF(J.EQ.MB) GO TO 36
DO 5 I=1,L1
IF(I.EQ.MA) GO TO 5
IF(M.EQ.2) GO TO 96
IF(X1(MA,J).EQ.GAP) GO TO 36
IF(X1(I,J).EQ.GAP.OR.X2(4,I).EQ.XT) GO TO 5
XB=X1(MA,J)
XB=X1(I,J)*X2(3,I)+X2(8,I)
GO TO 95
96 IF(X1(J,MA).EQ.GAP) GO TO 36
IF(X1(J,I).EQ.GAP.OR.X2(4,I).EQ.XT) GO TO 5
XB=X1(J,MA)
XB=X1(J,I)*X2(3,I)+X2(8,I)
95 MK=MK+1
XC=ABS(X2(4,I))*XA
X2(2,J)=X2(2,J)+XC*X2(5,I)
X2(1,J)=X2(1,J)+XB*XC*X2(5,I)
5 CONTINUE
IF(MK.EQ.0) GO TO 36
IF(X2(2,J)) 38,39,38
38 X2(1,J)=X2(1,J)/X2(2,J)
GO TO 37
39 X2(1,J)=X
37 X2(2,J)=XB
GO TO 4
36 X2(1,J)=XT
X2(2,J)=XT
4 CONTINUE

```

```

56 XB=X
XC=X
DO 6 J=1,L2
IF(X2(1,J).EQ.XT) GO TO 6
IF(X2(2,J)) 33,34,33
33 X2(2,J)=ABS((X2(2,J)-X2(1,J))/X2(2,J))*100.
GO TO 32
34 X2(2,J)=ABS(X2(2,J)-X2(1,J))*100.
32 XB=XB+1.
XC=XC+X2(2,J)
6 CONTINUE
XC=XC/XB
IF(XC-XX) 30,31,31
30 XX=XC
XY=XA
31 XA=XA+STEP
IF(XA.LE.FIN) GO TO 7
IF(M.EQ.2) GO TO 94
X3(4,MM)=XY
X3(6,MM)=XX
GO TO 93
94 X3(3,MM)=XY
X3(5,MM)=XX
93 MK=0
XA=X
XX=X
DO 57 I=1,L1
IF(M.EQ.2) GO TO 21
K2(9,I)=0
GO TO 22
21 K2(10,I)=0
22 CONTINUE
IF(I.EQ.MA) GO TO 57
IF(M.EQ.2) GO TO 92
IF(X1(I,MB).EQ.GAP.OR.X2(4,I).EQ.XT) GO TO 57
XB=X1(I,MB)*X2(3,I)+X2(8,I)
GO TO 91
92 IF(X1(MB,I).EQ.GAP.OR.X2(4,I).EQ.XT) GO TO 57

```

Контрольный пример и образец выдачи на печать

```

XB=X1(MB,I)*X2(3,I)+X2(8,I)
91 XC=ABS(X2(4,I))*XY
    XA=XA+XC*X2(5,I)
    XX=XX+XB*XC*X2(5,I)
    IF(M.EQ.2) GO TO 54
    K2(9,I)=XC*X2(5,I)*1000.
    GO TO 57
54 K2(10,I)=XC*X2(5,I)*1000.
57 CONTINUE
    IF(XA) 58,59,58
58 XX=XX/XA
    GO TO 60
59 XX=X
60 IF(M.EQ.2) GO TO 51
    X3(9,MM)=XX
    GO TO 1
51 X3(8,MM)=XX
    1 CONTINUE
    IF(MN.NE.2) GO TO 8
    K3(7,MM)=-1
    X3(8,MM)=XT
    X3(9,MM)=XT
    GO TO 11
8 IF(MN.NE.1) GO TO 10
    IF(X3(5,MM).EQ.XT) K3(7,MM)=1
    IF(X3(6,MM).EQ.XT) K3(7,MM)=0
    GO TO 11
10 IF(X3(5,MM).LT.X3(6,MM)) K3(7,MM)=0
    IF(X3(6,MM).LT.X3(5,MM)) K3(7,MM)=1
    IF(X3(5,MM).EQ.X3(6,MM)) K3(7,MM)=2
11 CONTINUE
    RETURN
    END
    
```

В матрице, приведенной в таблице 1, десять подчеркнутых значений заменены символом пробела GAP = $8 \cdot 10^{17}$. В результате получена таблица 2. Произведено прогнозирование пропущенных значений в режиме I. Выполнены 3 итерации. Заполненная матрица приведена в таблице 3.

В качестве образца приведена выдача для итерации N = 3. На печать выведена следующая информация: порядковый номер пробела и его "координаты" - номер столбца и строки, на пересечении которых он находится; значение коэффициента α и средней ошибки δ (в процентах) при прогнозировании по строке или столбцу.

В зависимости от δ выбирается значение, найденное по строкам или столбцам. В 6-й колонке отмечается, чему отдано предпочтение - предсказанию по строке или столбцу.

Итерация N=3

Номер п/п	Номер столбца	Номер строки	Коэффициент α	Процент ошибки	0-строка 1-столбец 2-ов - место	Предсказанное значение
1	2	1	5.000+00	5.949+00	1	8.927+00
2	3	2	5.000+00	4.002+00	1	2.811+01
3	4	2	5.000+00	3.001+00	1	2.943+01
4	6	2	5.000+00	3.033+00	1	3.502+01
5	4	3	5.000+00	2.809+00	0	4.693+01
6	3	4	4.000+00	6.490-01	0	6.90
7	4	4	2.000+00	8.141-01	0	7.250+01
8	5	4	2.000+00	8.091-01	0	7.471+01
9	1	5	5.000+00	7.174-01	0	8.298+01
10	5	6	5.000+00	1.423+00	0	1.219+02

Образец выдачи результатов

Таблица 1

Матрица исходных данных

	1	2	3	4	5	6	7
1 *	6.940+00	9.010+00	1.081+01	1.201+01	1.400+01	1.600+01	1.900+01
2 *	2.300+01	2.430+01	2.700+01	2.808+01	3.100+01	3.206+01	3.545+01
3 *	3.910+01	4.000+01	4.495+01	4.730+01	5.094+01	5.199+01	5.493+01
4 *	6.355+01	6.537+01	6.972+01	7.239+01	7.492+01	7.896+01	7.990+01
5 *	8.547+01	8.762+01	8.890+01	9.122+01	9.291+01	9.594+01	9.900+01
6 *	1.079+02	1.124+02	1.148+02	1.187+02	1.215+02	1.276+02	1.269+02
7 *	1.329+02	1.373+02	1.389+02	1.785+02	1.810+02	1.840+02	1.862+02
8 *	1.970+02	2.006+02	2.044+02	2.072+02	2.090+02	2.100+02	2.100+02

Таблица 2

Матрица исходных данных

	1	2	3	4	5	6	7
1 *	6.940+01	8.000+17	1.081+01	1.201+01	1.400+01	1.600+01	1.900+01
2 *	2.300+01	2.430+01	8.000+17	8.000+17	3.100+01	8.000+17	3.545+01
3 *	3.910+01	4.000+01	4.495+01	8.000+17	5.094+01	5.199+01	5.493+01
4 *	6.355+01	6.537+01	8.000+17	8.000+17	8.000+17	7.896+01	7.990+01
5 *	8.000+17	8.762+01	8.890+01	9.122+01	9.291+01	9.594+01	9.500+01
6 *	1.079+02	1.124+02	1.148+02	1.187+02	8.000+17	1.276+02	1.269+02
7 *	1.329+02	1.373+02	1.389+02	1.785+02	1.810+02	1.840+02	1.862+02
8 *	1.970+02	2.006+02	2.044+02	2.072+02	2.090+02	2.100+02	2.100+02

Таблица 3

Результат

	1	2	3	4	5	6	7
1 *	6.940+00	8.927+00	1.081+01	1.201+01	1.400+01	1.600+01	1.900+01
2 *	2.300+01	2.430+01	2.811+01	2.943+01	3.100+01	3.502+01	3.545+01
3 *	3.910+01	4.000+01	4.495+01	4.693+01	5.094+01	5.199+01	5.493+01
4 *	6.355+01	6.537+01	6.928+01	7.231+01	7.471+01	7.896+01	7.990+01
5 *	8.298+01	8.762+01	8.890+01	9.122+01	9.291+01	9.592+01	2.900+01
6 *	1.079+02	1.124+02	1.148+02	1.187+02	1.219+02	1.276+02	1.269+02
7 *	1.329+02	1.373+02	1.389+02	1.785+02	1.810+02	1.839+02	1.862+02
8 *	1.970+02	2.006+02	2.044+02	2.072+02	2.090+02	2.100+02	2.100+02