1975 г (Конец года)

К ВОПРОСУ О ФИКСАЦИИ ОБОБЩЕННЫХ ОБРАТНЫХ МАТРИЦ.

Альт А.М.

г. Одесса

$f: X \to Y$

$f: E \to F$

$y = f^{-1}: F \to E$

Широкий круг технологических и экономических задач основан на решении систем несовместных уравнений с использованием условия минимума остаточной суммы квадратов.

В последеннее время в литературе все чаще появляются работы, в которых решения, удовлетворяющие условию минимума остаточной суммы квадратов производятся с помощью различных обратных матриц.

Так, в (I) обобщенные обратные матрицы применяются в методе группового учета аргументов (МГУА) для решения систем "частных" описаний.

Целью данной работы является систематизация существующих определений и обозначений обобщенных обратных матриц, а также получение всех необходимых теоретических результатов для построения алгоритма МГУА на обобщенныъх обратных матрицах.

Введем обозначения :

Пусть A – матрица, тогда $r(A)$ – ранг матрицы A; $\rho_1(A)$ – число строк матрицы A; $\rho_2(A)$ – число столбцов матрицы A; $\rho(A) = (\rho_1(A), \rho_2(A))$ – размер матрицы A

A^* – сопряженная (здесь транспонированная) к A матрица, т.е.

$$a^*_{ij} = a_{ji}$$

В этих обозначениях имеем следующие соотношения:

$$r(A,B) \le \min(r(A), r(B))$$

$$r(A) \le \min(\rho_1(A), \rho_2(A))$$

- единичная матрица, размер которой либо будет ясен из текста, либо отмечен следующим образом: - единичная матрица размера (n,n).

<u>Определение I</u>. Всякую матрицу B такую, что (I) $BA=I$ будем называть левой обратной к A.

<u>Определение 2.</u> Всякую матрицу C такую, что (2) $AC=I$ будем называть правой обратной к A.

Для матрицы A существует левая обратная тогда и только тогда, когда $\rho_1(A) \geq \rho_2(A) = r(A)$ (т.е. когда A столбцово-невырождена матрица).

Для матрицы A существует правая обратная тогда и только тогда, когда $\rho_2(A) \geq \rho_1(A) = r(A)$ (т.е когда A строчно-невырождена).

Если для матрицы A существует левая обратная B и правая обратная C, то $B=C$, и тем самым определена едиництвенная двусторонне-обратная матрица $A^{-1}=B=C$ и называемая обратной, и в этом случае $\rho_1(A)=\rho_2(A)=r(A)$ (т.е. A - квадратная невырожденная матрица).

в случае, когда $r(A) < \min(\rho_1(A), \rho_2(A))$, для матрицы A не существует обратной, левой обратной и правой обратной.

Тем не менее, для произвольных матриц A имеет смысл говорить о так называемых обобщенных обратных матрицах, которые в невырожденных случаях совпадают либо с левой обратной, либо с просто обратной, в зависимости от характера невырожденности.

Определение 3. Всякую матрицу $\mathcal{D}$ такую, что

(3) $$A\mathcal{D}A = A$$

(4) $$\mathcal{D}A\mathcal{D} = \mathcal{D}$$

будем называть обобщенной обратной к матрице A.

Из определения 3 непосредственно следует, что если матрица $\mathcal{D}$ будет обобщенной обратной к A, то матрица A, в свою очередь, будет обобщенной обратной к $\mathcal{D}$, то есть обобщенные обратные матрицы наследуют важное свойство обратных матриц – симметрию отношения "быть обратной".

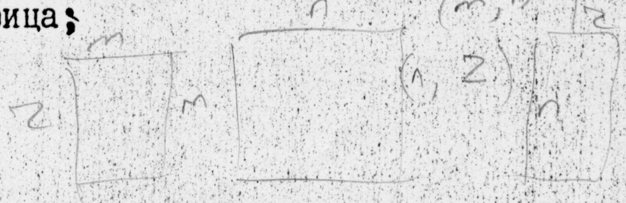
$\rho(A) = (m,n)$ $r(A) = r$

Определение 4. Всякую пару матриц (R_1, R_2) таких, что выполняются условия :

$\rho(R_1) = (n, r)$, $\rho(R_2) = (r, \ldots$

1. $\rho(R_1) = (\rho_2(A), r(A))$, $\rho(R_2) = (r(A), \rho_1(A))$ и
$r(R_1) = r(R_2) = r(A)$;

2. $\det(R_2 A R_1) \neq 0$, то есть R_2AR_1 – квадратная невырожденная матрица;



назовём A – фиксацией.

Имеют место следующие теоремы (доказательства см. в приложении), которые в совокупности дают исчерпывающее конструктивное описание обобщенных обратных матриц.

Теорема I. Для всякой матрицы $\mathcal{D}$, обобщенно-обратной к A, существует A – фиксация $(R_1 R_2)$, что

$$\mathcal{D} = R_1 (R_2 A R_1)^{-1} R_2$$

Теорема 2. Пусть $(R_1, R_2)-A$ - фиксация, тогда матрица $R_1(R_2AR_1)^{-1}R_2 \stackrel{def}{=} A^+_{R_1R_2}$ будет обобщенной обратной к A матрицей.

На множестве A- фиксаций введем отношение эквивалентности, определяемое следующим образом: A - фиксация (R_1R_2) эквивалентна A- фиксации (R_1', R_2') (обозначается $(R_1R_2)\approx(R_1'R_2')$, если существуют такие квадратные невырожденные матрицы T_1 и T_2 $(\rho_2(T_1)=\rho(T_2)=(r(A); r(A)))$, что $R_1'=R_1T_1$ и $R_2'=T_2R_2$

Теорема 3. Пусть $(R_1 R_2)$ и (R_1', R_2') - две A - фиксации, тогда $A^+_{R_1R_2}=A^+_{R_1'R_2'}$ тогда и только тогда, когда $(R_1R_2)\approx(R_1^1, R_2^1)$

Пусть A - строчно-невырожденная матрица, и пусть $(R_1, R_2)-A$ - фиксация, тогда $\rho(R_2)=(r(A), r(A))$ и $r(R_2)=r(A)$ то есть R_2 - квадратная невырожденная матрица, и значит

$$(R_1, R_2)\approx(R_1R_2^{-1}, I)$$

Таким образом, все A - фиксации строчно-невырожденной матрицы A определяются парами вида (R, I), где R произвольная матрица такая, что $\rho(R)=(\rho_2(A), r(A))$ и $r(A)=r(R)$ и AR - невырожденная квадратная матрица, то есть $det=(AR)\neq 0$

В этом случае $A^+_{RI}=R(AR)^{-1}$ имеет специальное обозначение $A^{\bar{R}}$.

Аналогично, если A - столбцово-невырожденная матрица, то все A - фиксации задаются парами вида (I, R), где R - произвольная матрица, удовлетворяющая условиям

$$\rho(R)=(r(A), \rho_1(A)),\quad r(R)=r(A)$$

$$det(RA)\neq 0$$

В этом случае $A^+_{IR}=(RA)^{-1}R$ имеет специальное обозначение $_R^-A$

Пусть $(R_1R_2)-A$ - фиксация и $A=BC$ скелетное разложение, тогда $(R_1, I)-C$ -фиксация и $(I, R_2)-B$ - фиксация и $A^+_{R_1R_2}=C^-_{R_1}\times R^-_2 B$ (см. приложение)

Пусть A - произвольная матрица и $A=BC$ - ее некоторое скелетное разложение.

Тогда пара $(C^*, B^*)-A$ - фиксация. Пусть $A=B_1C_1$ другое скелетное разложение, тогда очевидно $(C_1^*, B_1^*)\approx(C^*B^*)$

Обобщенная обратная матрица $A^+_{C^*B^*}$, определенная скелетным разложением $A=BC$, не зависит от выбора скелетного разложение (см. теорему 3), в этом случае обозначается через A^+ и носит специальное название псевдообратной. Псевдообратная к A матрица A^+ представляет особый интерес, поскольку обладает рядом замечательных, хорошо изученных свойств, из которых мы отметим одно (2) свойство, наиболлее важное для нас.

Пусть дано уравнение $Ax=y$, где A - произвольная матрица, а y - произвольный вектор-столбец.

Обозначим множество $\{\hat{x}/\hat{x}=A^+y+(I-A^+A)t, t\in E_1^{\rho_2(A)}\}$ через E, тогда для всякого $x\in E\ \|A\hat{x}-y\|=\min\limits_{x\in E}\|Ax-y\|$ при этом $\|A^+y\|=\min\limits_{\hat{x}\in E}\|\hat{x}\|$, то есть E есть множество решений уравнения $Ax=y$ как задачи наилучшего приближения вектора y векторами Ax, где $x\in E_1^{\rho_2(A)}$ и при этом вектор $\hat{x}_0=A^+y$ является самым коротким по норме вектором-решением.

Если $r(A)=\min(\rho_1(A)\rho_2(A))$, то есть A невырождена, то A^+ обозначается через ${}^-A$, если $r(A)=\rho_2(A)$ и через A^- , если $r(A)=\rho_1(A)$, тогда ${}^-A=(A^*A)^{-1}A^*$ $A^-=A^*(AA^*)^{-1}$ и называются соответственно левой и правой псевдообратной матрицей к А .

Пусть $A=BC$ - скелетное разложение $C={}^-BA$ Тогда $A^+=C^-{}^-B=({}^-(C^*))^*\,{}^-B$, т.е. вычисленное A^+ можно свести к нахождению двух левых псевдообратных матриц.

Рассмотрим теперь вопрос об обобщенном обращении произвольных матриц, разбитых на блоки.

Пусть A -произвольная матрица, разбитая на два вертикальных блока, т.е. $A(A_1, A_2)$ и пусть $A=BC$ - скелетное разложение, тогда разбиение $A=A_1A_2$ индуцирует разбиения $B=(B_1, B_2)$ и $C=\begin{pmatrix} C_1 \\ C_2 \end{pmatrix}$

За счет этого вопрос об обращении матрицы A , разбитой на блоки, редуцируется к обращению вертикальных и горизонтальных невырожденных матриц, разбитых на блоки.

Пусть A - столбцово-невырожденная матрица, пусть R - некоторая фиксирующая матрица (т.е. (I, R)-A - фиксация) и пусть $A=(A_1, A_2)$ - разбиение A на два вертикальных блока.

Рассмотрим следующую задачу фиксации блоков.

Найти такие фиксирующие матрицы R_1 и R_2 для A_1 и A_2 такие, что $R^-A=\begin{pmatrix} R_1 A_1 \\ R_2 A_2 \end{pmatrix}$

Напомним сначала следующую теорему Бархамера [3]

Теорема. Пусть A - квадратная невырожденная матр.. и $A=(A_1, A_2)$, тогда

$$A^{-1}=\begin{pmatrix} R_1^- A_1 \\ R_2^- A_2 \end{pmatrix}, \text{ где } R_1^* \text{ и } R_2^*$$

матрицы фундаментальных решений уравнений $A_2^* X=0$ и $A_1^* X=0$ соответственно.

Теорема 4. Пусть $A=(A_1 A_2)$ - произвольная вертикальная невырожденная матрица и ${}_R^-A$ - обратная слева к A матрица, тогда

$${}_R^-A=\begin{pmatrix} R_1^- A_1 \\ R_2^- A_2 \end{pmatrix}, \text{ где } R_1=X_1 R \text{ и } R_2=X_2 R$$

и X_1^* и X_2^* - матрицы ~~фун~~ фундаментальных решений уравнений $(RA_2)^* X=0$ и $(RA_1)^* X=0$ соответственно.

Доказательство.

Имеем ${}_R^-A=(RA)^{-1}R$, но $RA=(RA_1, RA_2)$ и по теореме Берхамера

$$(RA)^{-1}=\begin{pmatrix} X_1^-(RA_1) \\ X_2^-(RA_2) \end{pmatrix},$$

где X_1^* и X_2^* - матрицы фундаментальных решений уравнений $(RA_2)^* X=0$ и $(RA_1)^* X=0$ соответственно.

Отсюда ${}_R^-A=(RA)^{-1}R=\begin{pmatrix} X_1^-(RA_1)R \\ X_2^-(RA_2)R \end{pmatrix}=$

$$\begin{pmatrix} (X_1 RA_1)^{-1} X_1 R \\ (X_2 RA_2)^{-1} X_2 R \end{pmatrix}=\begin{pmatrix} R_1^- A_1 \\ R_2^- A_2 \end{pmatrix}, \text{ где}$$

$$R_1=X_1 R \text{ и } R_2=X_2 R$$

Пусть ~~теперь~~ теперь A - строчно-невырожденная матрица. Имеет место

Теорема 5. Пусть $A=\begin{pmatrix} A_1 \\ A_2 \end{pmatrix}$ и A_R^- - правая обратная к A, тогда $A_R^-=\left((A_1)_{R_1}^-, (A_2)_{R_2}^-\right)$, где $R_1 = RX_1$ и $R_2 = RX_2$ и X_1 и X_2 - матрицы фундаментальных решение уравнений $A_2 R x = 0$ и $A_1 R x = 0$ соответственно.

Доказательство получается транспонированием теоремы 4.

Следствие I. Пусть A столбцово невырождена и $A = (A_1, A_2)$ тогда $\bar{A} = \begin{pmatrix} \bar{R}_1 A_1 \\ R_2 A_2 \end{pmatrix}$, где $R_1 A^*$ и $R_2 = X_2 A^*$ и X_1^* и X_2^* - матрицы фундаментальных решений уравнений $A_2^* A x = 0$ и $A_1^* A x = 0$ соответственно.

Следствие 2. Пусть A - горизонтальная невырожденная матрица и $A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}$, тогда $\bar{A} = ((A_1)\bar{R}_1, (A_2)\bar{R}_2)$

где $R_1 = A^* X_1$; $R_2 = A^* X_2$ и X_1, X_2 - матрицы фундаментальных решний уравнений $A_2 A^* x = 0$ и $A_1 A^* x = 0$ соответственно.

ПРИЛОЖЕНИЕ

Доказательство теоремы I.

Пусть D - обобщенная обратная матрица к A и $D = ST$ - скелетное разложение, тогда $r(D) = r(A) = r(S) = r(T)$ и $\rho(S) = (\rho_2(A), r(A))$ и $\rho(T) = (r(A), \rho_1(A))$.

Покажем, что TAS невырождена. Имеем $DAD = D$, но $D = ST$, поэтому $STAST = ST$; S - столбцово невырождена и, значит, существует S_1 такая, что $S_1 S = I$; T - строчно вырождена и, значит, существует T_1 такая, что $TT_1 = I$, поэтому имеем

$$S_1 STASTT_1 = S_1 STT_1 = I$$

и, значит, $TAS = I$, то есть невырождена.

Но мы доказали нечто большее.

I. В каждом классе эквивалентных A-фиксаций есть A-фиксации $R_1 R_2$ такие, что $R_2 A R_1 = I$, $A^+_{R_1 R_2} = R_1 R_2$ то есть дают скелетное разложение.

2. Пусть $A = BC$ - скелетное разложение и D - обобщенная обратная к A, то $CDB = I$.

3. Если $D = ST$ - скелетное разложение, то $C = TA$; $B = AS$ задают скелетное разложение $A = BC$. Действительно, $r(A) \le \min(r(AS), r(TA))$ и $r(TA) \le r(A)$, то есть $r(AS) = r(TA) = r(A)$.

<u>Доказательство теоремы 2</u>.

Пусть $A = BC$ - скелетное разложение, тогда $R_2 B$ и CR_1 - невырожденные квадратные матрицы. Действительно, $\rho(R_2 B) = \rho(CR_1) = (r(A), r(A))$. Покажем, что $r(R_2 B) = r(CR_1) = r(A)$. Имеем: $r(A) \le \min(r(R_2 B), r(CR_1))$ так как $r(A) = r(R_2 A R_1) = r(R_2 B C R_1)$ с другой стороны, $r(R_2 B) \le \min(r(R_2), r(B)) = r(A)$ и $r(CR_1) \le r(A)$, то есть $r(R_2 B) = r(CR_1) = r(A)$.

Отсюда $(R_2 A R_1)^{-1} (CR_1)^{-1} (R_2 B)^{-1}$

Покажем, что $A^+_{R_1 R_2}$ удовлетворяет условиям (3) и (4) Определения 3.

Имеем (3) $AR_1(R_2AR_1)^{-1}R_2A =$
$= BCR_1(CR_1)^{-1}(R_2B)^{-1}R_2BC = A$

и (4)
$$A^+_{R_1R_2}AA^+_{R_1R_2} = R_1(R_2AR_1)^{-1}R_2AR_1(R_2AR_1)^{-1}R_2 =$$
$$= A^+_{R_1R_2}$$

Доказательство теоремы 3.

I. $(R_1, R_2) \approx (R_1^1, R_2^1) \Rightarrow A^+_{R_1R_2} = A^+_{R_1'R_2'}$ действительно $A^+_{R_1'R_2'} = R_1T_1(T_2R_2AR_1T_1)^{-1}T_2R_2 =$
$R_1T_1T_1^{-1}(RAR_1)T_2^{-1}T_2R_2 = R_1(R_2AR_2)^{-1}R_2 = A^+_{R_1R_2}$

2. $A^+_{R_1R_2} = A^+_{R_1'R_2'} \Rightarrow (R_1, R_2) \approx (R_1'R_2')$

Имеем $R_1(R_2AR_1)^{-1}R_2 = R_1^1(R_2^1AR_1^1)^{-1}R_2^1$

Но R_1^1 – столбцово-невырожденная матрица и R_2^1 – строчно-невырожденная, поэтому существуют матрицы K и F, что $KR_1^1 = I$ и $R_2^1F = I$, отсюда: $R_1^1 = R_1 \cdot (R_2AR_1)^{-1}R_2F \times$
$\times (R_2^1AR_1^1)$ и $R_2^1 = (R_2^1AR_1^1)KR_1(R_2AR_1)^{-1}R_2$
$\rho(R_2F) = \rho(KR_1) = (r(A), r(A))$. С другой стороны
$r(R_2F) = r(KR_1) = r(A)$, так как $(R_2^1AR_1^1) =$
$= (KR_1)(R_2AR_1)^{-1}R_2F$ $r(R_2^1AR_1^1) = r(A) \leq$
$\leq \min(r(KR_1), r(R_2F))$, а $r(KR_1) \leq r(R_1) = r(A)$
и $r(R_2F) \leq r(R_2) = r(A)$. Поэтому матрицы
$T_1 = (R_2AR_1)^{-1}R_2F(R_2^1AR_1^1)$ и
$T_2 = (R_2^1AR_1^1)KR_1(R_2AR_1)^{-1}R_2$

квадратные невырожденные матрицы, и т.д.

ЛИТЕРАТУРА

1. Браверман В.Я. Альт А.М. Применение обобщенных обратных матриц в методе группового учета аргументов.

2. Гантмахер Ф.Р. Теория матриц. Москва, издательство "Наука", 1967г.

3. Bjerhammar A. Estimation with singular inverses. Tellus XXIII 1971, N6

10.12.1975

165