ПРИМЕНЕНИЕ МЕТОДОВ ПСЕВДООБРАЩЕНИЯ К ЗАДАЧАМ ОПТИМИЗАЦИИ СТРУКТУР

Григорий Кудин

Abstract: *We present analytical expressions for the perturbations of pseudoinverse and projection matrices pozvolyayuschienahodit solutions of optimal synthesis of linear systems with fuzzy parameters of the task. В работе приводятся аналитические выражения для возмущений псевдообратных и проекционных матриц позволяющиенаходить решений задач оптимального синтеза линейных систем с нечётко задаными параметрами.*

Keywords: *возмущения псевдообратных и проекционных матриц оптимальный синтез линейных систем.*

ACM Classification Keywords: *Computing Methodologies. Simulation and modeling. optimum synthesis*

Введение

Многочисленные задачи, возникающие при решении современных проблем цифровой обработки информации: распознавание образов, кластеризации, построения ассоциативной памяти, обучения нейрокомпьютеров, анализа и решения систем уравнений, сводятся к исследованию линейных моделей линейные. Современные требования к математическим методам, применяемым для этих целей, существенно повысились: необходимо учитывать неопределённость, нечёткость, противоречивость в задании параметров моделей, возможность учитывать многофакторность и многокритериальность оптимизационных постановок задач, другие усложняющие процесс решения задач обстоятельства. В часности, современный взгляд на проблему решения системы линейных алгебраических уравнений (СЛАУ) порождает проблему анализа точности представления исходной информации, учёта вычислительных погрешностей и т. п. [Волошин, Кудин, 2010]. Одним из направлений исследования СЛАУ – анализ конечномалых возмущений линейных моделей в процессе её численного решения (анализ, исследование) [Кудин, 2010]. Альтернативным подходом является получение аналитических зависимостей влияния малых возмущений на решение с дальнейшей численной реализацией. Последний подход основан на математическом аппарате теории псевдообратных матриц [Алберт], дополненный новыми результатами: обратные формулы Гревиля, аналитическое представление возмущений псевдообратных матриц при возмущении элементов исходных матриц или изменении самой структуры матриц [Кириченко, 1997 , 2001, 2009]. В предлагаемой работе обобщены результаты по методам нахождения производных от псевдообратных матриц [Кудин Г] и получены аналитические представления возмущений псевдообратных и проекционных матриц, вызванных возмущениями исходных матриц. Приведены примеры вычислительной полезности их для оптимизации решений СЛАУ, синтеза линейных систем.

Будут приняты следующие способы представления матрицы

$$A \in R^{m\times n}:\ A = (a(1) \vdots \ldots \vdots a(n)) \equiv (a_{(1)} \vdots \ldots \vdots a_{(m)})^T \equiv (a_{ij})_{\substack{i=\overline{1,m}\\ j=\overline{1,n}}} \in R^{m\times n}.$$

где $a(j) = (\alpha_{1j},...,\alpha_{ij},...\alpha_{mj})^T \in R^m$, $a_{(i)} = (\alpha_{i1},...,\alpha_{ij},...\alpha_{in}) \in R^n$, $i = \overline{1,m}$, $j = \overline{1,n}$.

Для матрицы $A \in R^{m\times n}$ псевдообратная по Пенроузу [8] матрица $A^+ \in R^{m\times n}$ определяется соотношением:

$$\forall b \in R^m \ \ A^+ b = \arg \min_{x \in \Omega_A(b)} \|x\|^2, \text{ где } \Omega_A(b) = Arg \min_{x \in R^n} \|Ax - b\|^2 .$$

В практике применения псевдообращения важными являются проекционные матрицы, которые определяются и вычисляются с использованием матриц A и A^+:

1. проекционная матрица $P(A) = A^+ A$ - ортогональный проектор на подпространство L_{A^T}, порождённое векторами – строками матрицы A, т.е. на подпространство значений A^T;
2. проекционная матрица $P(A^T) = AA^+$ - ортогональный проектор на подпространство L_A, порождённое векторами - столбцами матрицы A, т.е. на подпространство значений A;
3. проекционная матрица $Z(A) = I_n - P(A)$ - ортогональный проектор на подпространство, ортогональное подпространству L_{A^T};
4. проекционная матрица $Z(A^T) = I_m - P(A^T)$ - ортогональный проектор на подпространство, ортогональное подпространству L_A;
5. матрица $R(A) = A^+(A^+)^T$, матрица $R(A^T) = (A^+)^T A^+$ - R - операторы.

Псевдообращение при возмущении элементов исходной матрицы.

Формулы возмущения псевдообратной и проекционных матриц в условиях матричного возмущения предоставляют возможность получить формулы псевдообращения при условии возмущения отдельных элементов матрицы A. Ниже подаются и исследуются такие формулы.

Пусть для матрицы $A \in R^{m\times n}$ вида (1), известна её псевдообратная матрица $A^+ \in R^{n\times m}$, для которой введены обозначения:

$$A^+ = (p(1) \vdots \ldots \vdots p(m)) = (p_{(1)} \vdots \ldots \vdots p_{(n)})^T \equiv (p_{ij})_{\substack{i=\overline{1,n}\\ j=\overline{1,m}}} \in R^{n\times m},$$

$$p(j) = (p_{1j},...,p_{ij},...p_{nj})^T \in R^n, \quad j = \overline{1,m}, \quad p_{(i)} = (p_{i1},...,p_{ij},...p_{im})^T \in R^m, \quad i = \overline{1,n}.$$

Если элемент α_{ij} матрицы $A \in R^{m\times n}$ изменяется на некоторую величину $\delta_{ij} \in R$, т.е. она приобретает возмущённый вид $\widetilde{A} = A + \Delta A(\delta_{ij})$.

Возмущение $\Delta A(\delta_{ij})$ можно представить в матричном виде $\Delta A(\delta_{ij}) = \delta_{ij}\, e_m(i)\ e_n^T(j)$, где векторы $e_n(j) \in R^n$, $e_m(i) \in R^m$ единичные векторы.

Возникает задача - определить влияние возмущение такого вида на элементы псевдообратной матрицы $\Delta A^{+}(\delta_{ij}) = \widetilde{A}^{+} - A^{+}$ на представление возмущений Z - проекционных матриц

$\Delta Z(A,\delta_{ij}) = Z(\widetilde{A}) - Z(A)$, $\Delta Z(A^T,\delta_{ij}) = Z(\widetilde{A}^T) - Z(A^T)$, а также возмущений для R - операторов $\Delta R(A,\delta_{ij}) = R(\widetilde{A}) - R(A)$, $\Delta R(A^T,\delta_{ij}) = R(\widetilde{A}^T) - R(A^T)$.

Очевидно, что составляющие матрицы возмущений $\Delta A(\delta_{ij})$ - векторы $e_m(i)$ и $e_n^T(j)$ - это векторы, соответствующие векторам a и b^T матричных возмущений к матрице A в работе [Кириченко, 1997].

Таким образом, задача описания псевдообращения для поэлементно возмущённой матрицы сводится к рассмотрению четырёх вариантов линейной зависимости или независимости от векторов – столбцов и векторов - строк матрицы A векторов $e_m(i)$ и $e_n^T(j)$ соответственно. Условия линейной зависимости векторов $e_m(i)$ и $e_n^T(j)$ от векторов – столбцов и векторов - строк матрицы A имеют вид $e_m{}^T(i)\ Z(A^T)e_m(i) = 0$, $e_n^T(j) Z(A) e_n(j) = 0$ соответственно.

С учётом представлений для Z - проекционных матриц $Z(A^T) = (I_m - AA^{+}) \in R^{m\times m}$, $Z(A) = (I_n - A^{+}A) \in R^{n\times n}$, эти условия переписываются в виде:

$$e_i^T(m)\ Z(A^T)e_i(m) = 0 \quad \Leftrightarrow \quad a_{(i)}^T p(i) = 1,\ e_j^T(n) Z(A) e_j(n) = 0 \quad \Leftrightarrow \quad p_{(j)}^T a(j) = 1.$$

При этом условия линейной независимости векторов $e_m(i)$ и $e_n^T(j)$ от векторов – столбцов и векторов - строк матрицы A принимают вид $a_{(i)}^T p(i) < 1$.

Теорема. Если для матрицы $A \in R^{m\times n}$ известна её псевдообратная матрица $A^{+} \in R^{n\times m}$ и некоторый её элемент α_{ij} изменяется на величину $\delta_{ij} \in R$, то формулы возмущений для псевдообратной и проекционных матриц имеют следующий вид:

1) если $a_{(i)}^T p(i) < 1$, $p_{(j)}^T a(j) < 1$, т.е. когда векторы $e_i(m)$ и $e_j^T(n)$ линейно независимы от векторов – столбцов и векторов - строк матрицы A соответственно:

$$\Delta A^{+}(\delta_{ij}) = -p(i)e_i^T(m)Z(A^T) - Z(A)e_j(n)p_{(j)}^T + Z(A)e_j(n)e_i^T(m)Z(A^T)(p_{ji} + \frac{1}{\delta_{ij}}) ,$$

$$\Delta Z(A,\delta_{ij}) = Z(A)e_j(n)e_j^T(n)Z(A), \qquad \Delta Z(A^T,\delta_{ij}) = Z(A^T)e_i(m)e_i^T(m)Z(A^T),$$

$$\Delta R(A,\delta_{ij}) = -A^{+}p_{(j)}e_j^T(n)Z(A) - Z(A)e_j(n)p_{(j)}^T A^{+T} + p(i)p^T(i) +$$

$$+\left\|p_{(j)}\right\|^2 Z(A)e_j(n)e_j^T(n)Z(A) - (Z(A)e_j(n)p^T(i) + p(i)e_j^T(n)Z(A))(p_{ji} + \frac{1}{\delta_{ij}}) +$$

$$+Z(A)e_j(n)e_j^T(n)Z(A)(p_{ji} + \frac{1}{\delta_{ij}})^2,$$

2) если $a_{(i)}^T p(i) = 1$, $p_{(j)}^T a(j) < 1$, т.е. когда вектор $e_i(m)$ линейно зависим от векторов - столбцов, а вектор $e_j^T(n)$ линейно независим от векторов - строк матрицы A:

$$\Delta A^{+}(\delta_{ij}) = -K(A^{+},\delta_{ij})\,A^{+}\ , \qquad K(A^{+},\delta_{ij}) = \frac{kk^{T}}{\|p(i)\|^{2}\delta_{ij}^{\ 2} - 2p_{ij}\delta_{ij} + 1}\ , \qquad k = p(i)\delta_{ij} - e_{j}(n)\ ;$$

$$\Delta Z(A,\delta_{ij}) = \frac{kk^{T}}{\|p(i)\|^{2}\delta_{ij}^{\ 2} - 2p_{ij}\delta_{ij} + 1} - e_{j}(n)e_{j}^{T}(n)\ ,$$

$$\Delta Z(A^{T},\delta_{ij}) = \frac{k_{T}k_{T}^{T}}{\|p_{(j)}\|^{2}\delta_{ij}^{\ 2} - 2p_{ji}\,\delta_{ij} + 1} - e_{i}(m)e_{i}^{T}(m)\ ,\ k_{T} = \delta_{ij}\,p_{(j)} - e_{i}(m)\ ;$$

$$R(A + \Delta A(\delta_{ij})) = \left(I_{n} - K(A^{+},\delta_{ij})\right)R(A)\left(I_{n} - K(A^{+},\delta_{ij})\right).$$

3) если $a_{(i)}^{T}p(i) = 1$, $p_{(j)}^{T}a(j) = 1$

и при этом $p_{ji}\delta_{ij} = -1$, т.е. когда векторы $e_{i}(m)$ и $e_{j}^{T}(n)$ линейно зависимы от векторов – столбцов и векторов – строк матрицы A соответственно, а ранг возмущённой матрицы $A + \delta_{ij}\,e_{i}(m)\ e_{j}^{T}(n)$ падает:

$$rank(A + \delta_{ij}\,e_{i}(m)\ e_{j}^{T}(n)) = rank(A) - 1.$$

Для этого случая

$$(A + \Delta A(\delta_{ij}))^{+} = A^{+} - \frac{p(i)\,p^{T}(i)}{\|p(i)\|^{2}}A^{+} - A^{+}\frac{p_{(j)}p_{(j)}^{T}}{\|p_{(j)}\|^{2}} + \alpha\,\frac{p(i)}{\|p(i)\|}\frac{p_{(j)}^{T}}{\|p_{(j)}\|}$$

$$\alpha = \left(\frac{p_{(j)}^{T}}{\|p_{(j)}\|}A^{+T}\frac{p(i)}{\|p(i)\|}\right)\ , \quad \Delta Z(A,\delta_{ij}) = \frac{p(i)}{\|p(i)\|}\frac{p_{(j)}^{T}}{\|p_{(j)}\|}\ , \quad \Delta Z(A^{T},\delta_{ij}) = \frac{p_{(j)}p_{(j)}^{T}}{\|p_{(j)}\|^{2}}\ ,$$

$$\Delta R(A,\delta_{ij}) = A^{+}(\Delta A^{+}(\delta_{ij}))^{T} + (\Delta A^{+}(\delta_{ij}))A^{+^{T}} + (\Delta A^{+}(\delta_{ij}))(\Delta A^{+}(\delta_{ij}))^{T}\ .$$

4) если $a_{(i)}^{T}p(i) = 1$, $p_{(j)}^{T}a(j) = 1$ и при этом $p_{ji}\delta_{ij} \neq -1$, т.е. когда векторы $e_{i}(m)$ и $e_{j}^{T}(n)$ линейно зависимы от векторов - строк и векторов - столбцов матрицы A соответственно, а ранг возмущённой матрицы $A + \delta\,e_{i}(m)\ e_{j}^{T}(n)$ не падает: $rank(A + \delta_{ij}\,e_{i}(m)\ e_{j}^{T}(n)) = rank(A).$

Для этого случая

$$\Delta A^{+}(\delta_{ij}) = -\delta_{ij}\frac{p(i)p_{(j)}^{T}}{1 + p_{ji}\delta_{ij}}\ ,\Delta Z(A,\delta_{ij}) = \Delta Z(A^{T},\delta_{ij}) = 0\ ,$$

$$\Delta R(A,\delta_{ij}) = -\frac{\delta_{ij}}{1 + p_{ji}\delta_{ij}}\left(p(i)p_{(j)}^{T}A^{+^{T}} + A^{+}p_{(j)}p^{T}(i)\right) + \frac{\delta_{ij}^{\ 2}\|p_{(j)}\|^{2}}{(1 + p_{ji}\delta_{ij})^{2}}\ p(i)p^{T}(i),$$

Очевидно, что малые возмущения элемента α_{ij} матрицы A при условиях линейной независимости от векторов – строк и векторов - столбцов матрицы A векторов $e_{i}(m)$ и $e_{j}(n)$ соответственно, могут существенно изменить псевдообратную матрицу A^{+} - в формуле есть слагаемое обратно пропорциональное величине возмущения. Как следствие, могут существенно измениться также R -

операторы - матрицы $R(A),\ R(A^T)$. Проекционные матрицы $Z(A),\ Z(A^T)$ при возмущении элемента α_{ij} исходной матрицы A не изменяются.

Малые возмущения элемента α_{ij} матрицы A , соответствующие второму случаю, вызвают малые возмущения для псевдообратной матрицы A^+, а также для всех проекционных матриц.

Возмущения элемента α_{ij} матрицы A при условиях линейной зависимости от векторов - строк и векторов - столбцов матрицы A векторов $e_i(m)$ и $e_j(n)$ соответственно, обуславливают независимые от δ_{ij} возмущения для псевдообратной матрицы A^+ , а также для проекционных матриц $Z(A),\ Z(A^T)$ и взвешенной проекционной матрицы $R(A)$.

Возмущения элемента α_{ij} матрицы A при условиях линейной зависимости от векторов - строк и векторов - столбцов матрицы A векторов $e_i(m)$ и $e_j(n)$ соответственно, удовлетворяющими условию (64), изменяют элементы псевдообратнщй матрицы A^+ и проекционных матриц $R(A),\ R(A^T)$ пропорционально возмущению элемента α_{ij} матрицы A , проекционные матрицы $Z(A),\ Z(A^T)$ при этом не изменяются.

Численный эксперимент

Полученные зависимости псевдообратной матрицы A^+ от возмущений $\Delta A(\delta_{ij})$ исходной матрицы A , позволяют аналитически, в явном виде, вычислять возмущённые псевдообратные матрицы, используя только невозмущённый её прообраз, величину возмущения и свойства векторов $e_i(m)$ и $e_j(n)$.

Ниже рассмотрены фрагменты решения таких задач:

1. минимизация нормы главного решения системы линейных алгебраических уравнений;
2. приближение псевдорешения системы линейных алгебраических уравнений к желаемому решению.

1. Пусть задана система линейных неоднородных алгебраических уравнений (СЛАУ)

$Ax = y$, где $A \in R^{m\times n}$ **-** действительная матрица вида (1), $x \in R^n$, $y \in R^m$.

Согласно теории псевдообратных и проекционных матриц известно, что

1) при выполнении условия $y^T Z(A^T) y = 0$ система уравнений совместна;

2) при выполнении условия $\det A^T A > 0$ решение (псевдорешение) системи уравнений единственное;

3) множество решений (псевдорешений) системы описывается формулой

$$\hat{x} = A^+ y + Z(A)v, \quad \forall v \in R^n .$$

При этом так называемое главное решение СЛАУ:

$$x_{гл} = A^+ y \equiv (A^T A)^{-1} A^T y \equiv A^T (AA^T)^+ y ,$$

в котором для случая квадратной невырожденной матрицы A имеет место равенство $A^+ = A^{-1}$.

В задачах управления линейными дискретными системами [**Бублик**] возникает необходимость выбора параметров математической модели так, чтобы найти минимальное возможное значение функции, которая определяет норму управления - норму главного решения некоторого СЛАУ, т.е. функции

$$D_{гл}(\widetilde{A}) = y^T R(\widetilde{A}^T) y .$$

Предполагается, что элемент α_{ij} матрицы A изменяется на некоторую величину $\delta_{ij} \in R$, выполняются условия четвёртого случая возможных возмущений, т.е. $\Delta A(\delta_{ij}) = \delta_{ij}\, e_i(m)\; e_j^T(n)$ и при этом векторы $e_i(m)$ и $e_j^T(n)$ линейно зависимы с вектор - столбцами и вектор - строками матрицы A соответственно ($zm(i,i)=0, \quad zn(j,j)=0$), а также $p_{ji}\,\delta_{ij} \neq -1$. Задача сводится к нахождению в окрестности нулевого значения возмущения δ_{ij} такого δ_0, чтобы

$$\Delta D_{гл}(A,\alpha_{ij},\delta_0) = D_{гл}(A+\Delta A(\delta_0)) - D_{гл}(A) < 0 ,$$

т.е. чтобы $\Delta D_{гл}(A,\alpha_{ij},\delta_0) = y^T \Delta R(A^T,\delta_0)\, y < 0$,

Используя представление для возмущённого R - оператора несложно получить

$$\Delta\big(D_{гл}(A,\alpha_{ij},\delta_{ij})\big) = g_1 \frac{\delta_{ij}}{1+p_{ij}\delta_{ij}} + g_2 \left(\frac{\delta_{ij}}{1+p_{ij}\delta_{ij}}\right)^2 ,$$

где $g_1 = -2l_1 q_i,\ l_1 = y^T A^+ p_{(j)}, \qquad g_2 = l_2 q_i,\ l_2 = \|p(i)\|^2 x_{гл\,j},\ q_i = (p^T(i) y), \quad x_{гл\,j} = y^T p_{(j)}$,

$x_{гл\,j}$ - j – ая компонента главного решения системы .

Исследование функции $\Delta D_{гл}(A,\alpha_{ij},\delta_{ij})$ усложнено наличием параметров, поэтому целесообразно ограничиться некоторыми подмножествами их значений. Если предположить, что

$$p_{ij} = p > 0, \quad l_1 > 0, \quad l_2 > 0, \quad l_2 - p l_1 > 0, \quad q_i > 0 ,$$

то будет иметь место неравенство $\Delta D_{гл}(A,\alpha_{ij},\delta_{ij}) < 0$ при $0 < \delta_{ij} \le \delta_0 = 2l_1/(l_2 - 2l_1 p)$. Более того, если $0 < \delta_{ij} \le \delta_1 = l_1/(l_2 - l_1 p)$, то функция $\Delta D_{гл}(A,\alpha_{ij},\delta_{ij})$ монотонно убывает к значению $\Delta D_{гл}(A,\alpha_{ij},\delta_1) = -q_i l_1^2 / l_2$.

2. Задача формулируется так: необходимо определить возможное возмущение элемента α_{ij} матрицы A так, чтобы уменьшить расстояние главного решения СЛАУ (70) - $\hat{x} = A^+ y$ от некоторого заданного вектора x_*, т.е. минимизировать функцию $D_\rho(A) = (x_* - A^+ y)^T (x_* - A^+ y)$.

Как и прежде, предполагается, что элемент α_{ij} матрицы A изменяется в соответствии с условиями четвёртого случая возможных возмущений, т.е. $\Delta A(\delta_{ij}) = \delta_{ij}\, e_i(m)\; e_j^T(n)$ и при этом векторы $e_i(m)$ и $e_j^T(n)$ линейно зависимы с вектор - столбцами и вектор - строками матрицы A соответственно ($zm(i,i)=0, \quad zn(j,j)=0$), а также $p_{ji}\delta_{ij} \neq -1$.

Решение задачи можно свести к решению задачи на минимум нормы главного решения соответствуюшей СЛАУ. Действительно, если ввести дополнительно вектор y_*:

$y_* = Ax_*$,то будет иметь место тождество $D_\rho(A) = (x_* - A^+ y)^T (x_* - A^+ y) \equiv (y_* - y)^T R(A^T)(y_* - y)$ Таким образом после введения вектора $\overline{y} = y_* - y$ исследование проводится по схеме исследования функции нормы главного решения системы линейных алгебраических уравнений.

Выводы

Псевдообращение является мощным методом исследования в прикладных задачах, дающим возможность конструктивного описания решений задач, расширяющим возможности при исследовании проблемы оптимизации структуры математической модели. Метод возмущения псевдообратных матриц в настоящей работе распространён на определение аналитической зависимости возмущений элементов псевдообратных и проекционных матриц от возмущений отдельных элементов исходной матрицы. Полученные зависимости найдут в дальнейшем применение при решении задач синтеза линейных моделей, оптимизации алгоритмов цифровой обработки информации, аппроксимации функций, прогноза временных процессов.

Библиография

[Алберт]Алберт А. Регрессия, псевдоинверсия, рекуррентное оценивание. /Пер. с англ. –М.: Наука, 1977.–305 с.

[Бублик] Бублик С.Б , Кудін Г.І., Задачі керування лінійними системами з дискретним аргументом і крайовими умовами// Колективна монографія 'Сучасні методи та інформаційні технології математичного моделювання, аналізу і оптимізації складних систем'. – 2006. - Київ: КНУ.- 198 с. – С. 73-92.

[Волошин, Кудин, 2010] Волошин А, Кудин В., Кудин Г. Методы анализа малых возмущений линейных моделей // Natural and Artificial Intelligence, ЇTHEA, Sofia, 2010. - P. 41-47.

[Кириченко, 1997] Кириченко Н.Ф. Аналитическое представление возмущений псевдообратных матриц. // Кибернетика и системный анализ. – 1997.–№2.–С.98-107.

[Кириченко, 2001] Кириченко Н.Ф.,Лепеха Н.П. Возмущение псевдообратных и проекционных матриц и их применение к идентификации линейных и нелинейных зависимостей. //Проблемы управления и информатики.– 2001.-№1.-С.6-22

[Кириченко, 2009] Кириченко Н.Ф., Кудин Г.И. Анализ и синтез систем классификации сигналов средствами возмущений псевдообратных и проекционных операций // Кибернетика и системный анализ. – 2009. – №3. – С. 47-57

[Кудин Г] Кудін Г.І. Псевдообернені матриці: функції та їх похідні // Журнал обчислювальної та прикладної математики -2003., №2(89),--С. 66-70

[Кудин, 2010] Кудин В., Богаенко В. О принятии решений при анализе малых возмущений линейных моделей // Information Models of Knowledge, ITHEA, Kiev-Sofia, 2010. - P. 226-231.

Сведения об авторе

Григорий Кудин – кандидат физико-математичкских наук, профессор, Киевский национальный университет имени Тараса Шевченка, Украина, 01017 Киев, ул. Владимирская, 64, e-mail: Kudin@unicyb.kiev.ua

Сфера научных интересов: теория линейных систем, псевдообращения матри;, цифровая обработка информации,