

Е.Г. Ревунова

Определение минимальной ошибки с использованием критериев выбора модели для задачи преобразования выхода линейной системы в выход системы с заданным базисом

Рассмотрен подход к определению минимума ошибки для метода устойчивого преобразования выхода линейной системы в выход системы с заданным базисом. Подход основан на использовании критериев выбора модели. Приведены результаты эксперимента.

An approach is considered to finding the error minimum for the method of a stable transformation of a linear system output to the output of a system with a given basis is suggested. The approach is based on the model selection criteria. The results of the experimental investigation are presented.

Розглянуто підхід до визначення мінімуму помилки для методу стійкого перетворення виходу лінійної системи в вихід системи із заданим базисом. Підхід базується на використанні критеріїв вибору моделі. Наведено результати експерименту.

Введение. Анализ существующих подходов к решению задачи преобразования выхода показал, что существующие методы ориентированы на случай, когда матрица линейного преобразования исходной системы и ковариационная матрица шума – не вырождены.

Пусть имеется линейная система, вектор выхода которой формируется путем линейного преобразования входного вектора (входа) и добавления аддитивного шума. Это может быть измерительная система, для которой известны вектор наблюдений выхода и матрица (линейного преобразования вход–выход), описывающая взаимодействие сигнала со средой и (или) особенности измерительных средств (детектора, преобразователя сигнала и др.) и необходимо восстановить дискретно заданный сигнал объекта измерения. Столбцы матрицы линейного преобразования можно рассматривать как отсчеты дискретно заданных базисных функций линейной системы. Набор базисных функций отражает свойства конкретной измерительной системы, т.е. не может быть произвольным. Соответственно, наблюдаемый выход, определяемый базисными функциями, может: не отвечать требованиям пользователя или быть не совместимым с методами дальнейшей обработки. В случае, если известен набор базисных функций, которые дали бы выход с необходимыми свойствами, можно поставить задачу поиска преобразования наблюдаемого выхода в выход системы с необходимым базисом.

В работах [1–3] предложено получать искомое преобразование с использованием обраще-

ния матрицы линейного преобразования исходной системы. Однако если матрица линейного преобразования исходной системы имеет высокое число обусловленности и ряд ее сингулярных чисел плавно спадает до нуля, получаемое с использованием обратной матрицы решение будет неустойчивым и ошибка решения велика, то требуется регуляризация [4].

Для преодоления данного недостатка автором предложен подход к устойчивому решению задачи преобразования выхода, основанный на усеченном сингулярном разложении [5–7]. Обеспечение устойчивости к шуму в векторе измерений осуществляется путем нахождения решения, близкого к оптимальному (решения, для которого ошибка восстановления вектора измерений минимальна) путем балансирования составляющих ошибки в зависимости от числа компонент сингулярного разложения. Оптимальным числом компонент сингулярного разложения предложено считать такое, при котором среднеквадратичная ошибка преобразования выхода минимальна. Чтобы выбрать число компонент сингулярного разложения (k), при котором ошибка решения близка к минимальной в реальных условиях, т.е. когда точное решение неизвестно, предлагается [5] использовать критерии выбора модели.

Экспериментальное исследование [5] показало, что тенденция к снижению сложности модели при нарастании уровня шума зачастую приводит к тому, что сложность модели (число компонент сингулярного разложения), полученная по критерию выбора модели, оказывается значи-

тельно меньше оптимальной. Использование модели с неоптимальным числом компонент сингулярного разложения приводит к нарастанию ошибки преобразования выхода, поэтому актуальным остается поиск подходов, позволяющих находить k , близкое к оптимальному. В данной статье предложены подход к поиску оптимального k и его экспериментальное исследование.

Задача преобразования выхода

Пусть сигнал \mathbf{b} получен с выхода линейной системы A , выполняющей преобразование

$$Ax + \varepsilon = \mathbf{b}, \quad (1)$$

где $A \in \mathbb{R}^{m \times n}$, $x \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$, и

$$Ax = \mathbf{b}_0. \quad (2)$$

Столбцы матрицы линейного преобразования A можно рассматривать как отсчет дискретно заданных базисных функций линейной системы. Например, как дискретно заданный набор функций отклика детектора измерительной системы A . Набор базисных функций A отражает свойства конкретной измерительной системы и не может быть произвольным. Соответственно наблюдаемый выход \mathbf{b} , определяемый базисными функциями A , может не соответствовать требованиям пользователя. Например, обеспечивать недостаточную разрешающую способность системы A .

С другой стороны, если известен набор базисных функций C , которые дали бы выход с требуемыми свойствами, можно поставить задачу нахождения преобразования наблюдаемого выхода \mathbf{b} в выход системы C с базисом C .

Обозначим как \mathbf{d}_0 выход линейной системы C , выполняющей преобразование

$$Cx = \mathbf{d}_0. \quad (3)$$

Будем искать преобразование выхода как линейное преобразование.

Для получения решения – оценки выхода системы C по \mathbf{b} – сначала получим оценку x' входа x , решив обратную задачу:

$$x' = \mathbf{P}\mathbf{b}, \quad (4)$$

где \mathbf{P} – оператор (матрица), преобразующий выход \mathbf{b} в x' .

Затем получим оценку выхода системы C :

$$\mathbf{d}' = Cx' = CP\mathbf{b} = \mathbf{T}\mathbf{b}. \quad (5)$$

Таким образом, оператор CP преобразует \mathbf{b} в \mathbf{d}' . Матрицу преобразования $T = CP$ в [1] называют матрицей редукции.

Конкретный вид P зависит от свойств матрицы A . Если ряд собственных чисел A спадает монотонно и число обусловленности велико, то задачу относят к классу дискретных некорректных задач [8]. Приближенные решения дискретных некорректных задач как задач наименьших квадратов с использованием численных методов линейной алгебры, таких как разложения LU , Холецкого, QR неустойчивы. Это означает, что малые возмущения во входных данных приводят к большим возмущениям в решении.

Преобразование вектора выхода на основе сингулярного разложения

Разрабатываемый автором подход к устойчивому решению задачи преобразования выхода основан на усеченном сингулярном разложении. Для устойчивого регуляризованного решения при вычислении оператора CP (5) в качестве P используем матрицу, получаемую как

$$P = A_k^+ = V \text{diag} \left(\frac{\varphi_i}{\sigma_i} \right) U^T, \text{ при } i \leq k \quad \varphi_i = 1, \\ \text{иначе } \varphi_i = 0. \quad (6)$$

Здесь $A_k = USV^T$ – приближение матрицы A ($m \times n$) A , полученное по k ($k < n$) компонентам сингулярного разложения, $U = (\mathbf{u}_1, \dots, \mathbf{u}_k)$ – матрица левых сингулярных векторов, $V = (\mathbf{v}_1, \dots, \mathbf{v}_k)$ – матрица правых сингулярных векторов, $S = \text{diag}(\sigma_1, \dots, \sigma_k)$ – матрица сингулярных чисел.

Оценка выхода системы C , полученная с использованием k компонент сингулярного разложения A , такова:

$$\mathbf{d}'_k = CA_k^+ \mathbf{b} = T_k \mathbf{b}, \quad (7)$$

$$T_k = CA_k^+ = CV \text{diag} \left(\frac{\varphi_i}{\sigma_i} \right) U^T \mathbf{d}'_k. \quad (8)$$

Оптимальным числом компонент сингулярного разложения будем считать такое, при котором минимальна среднеквадратичная ошибка преобразования выхода:

$$e = E_\varepsilon \| \mathbf{d}'_k - \mathbf{d}_0 \|^2 = E_\varepsilon \| T_k \mathbf{b} - \mathbf{d}_0 \|^2. \quad (9)$$

Выделим в выражении для среднеквадратичной ошибки преобразования выхода детер-

минированную и стохастическую составляющие. С учетом $\mathbf{b} = \mathbf{b}_0 + \boldsymbol{\varepsilon}$ выражение (9) принимает следующий вид:

$$e = E_{\varepsilon} \|\mathbf{T}_k \mathbf{b} - \mathbf{d}_0\|^2 = E_{\varepsilon} \|\mathbf{T}_k \mathbf{b}_0 - \mathbf{d}_0\|^2 + (10) \\ + E_{\varepsilon} \|\mathbf{T}_k \boldsymbol{\varepsilon}\|^2 + 2E_{\varepsilon} \langle \mathbf{T}_k \mathbf{b}_0 - \mathbf{d}_0, \mathbf{T}_k \boldsymbol{\varepsilon} \rangle.$$

Так как $E_{\varepsilon} \langle \mathbf{T}_k \mathbf{b}_0 - \mathbf{d}_0, \mathbf{T}_k \boldsymbol{\varepsilon} \rangle = 0$,

$$e = E_{\varepsilon} \|\mathbf{T}_k \mathbf{b} - \mathbf{d}_0\|^2 = \|\mathbf{T}_k \mathbf{b}_0 - \mathbf{d}_0\|^2 + E_{\varepsilon} \|\mathbf{T}_k \boldsymbol{\varepsilon}\|^2. \quad (11)$$

Исследование зависимости величины стохастической составляющей ошибки преобразования выхода от числа компонент сингулярного разложения матрицы \mathbf{A} [8] показало, что математическое ожидание $E_{\varepsilon} \|\mathbf{T}_k \boldsymbol{\varepsilon}\|^2$ можно представить в таком виде:

$$E_{\varepsilon} \|\mathbf{T}_k \boldsymbol{\varepsilon}\|^2 = \sigma^2 \text{trace}(\mathbf{T}_k^T \mathbf{T}_k) = \\ = \text{trace}(\mathbf{H} \mathbf{M}_{k-1}) + \sigma_k^{-2} \text{trace}(\mathbf{v}_k^T \mathbf{C}^T \mathbf{C} \mathbf{v}_k), \quad (12)$$

где $\mathbf{H} = \mathbf{C}^T \mathbf{C}$, $\mathbf{M}_k = \mathbf{A}_k^+ \mathbf{A}_k^{+T}$. Значение выражения $\mathbf{v}_k^T \mathbf{C}^T \mathbf{C} \mathbf{v}_k$ в (17) есть положительная величина, поэтому $\sigma_k^{-2} \text{trace}(\mathbf{v}_k^T \mathbf{C}^T \mathbf{C} \mathbf{v}_k) > 0$. Из выражения (17) и положительности $\sigma_k^{-2} \text{trace}(\mathbf{v}_k^T \mathbf{C}^T \mathbf{C} \mathbf{v}_k)$ следует, что стохастическая составляющая ошибки преобразования выхода возрастает с ростом k .

Зависимость величины детерминированной составляющей ошибки $\|\mathbf{T}_k \mathbf{b}_0 - \mathbf{d}_0\|^2$ от числа компонент сингулярного разложения матрицы \mathbf{A} представима в виде

$$\|\mathbf{T}_k \mathbf{b}_0 - \mathbf{d}_0\|^2 = \|\mathbf{C} \Delta \mathbf{A}_k^+ \mathbf{b}_0\|^2 = \|\mathbf{C} \Delta \mathbf{A}_{k-1}^+ \mathbf{b}_0\|^2 - \delta, \quad (13)$$

где

$$\Delta \mathbf{A}_k^+ = \mathbf{A}^+ - \mathbf{A}_k^+, \quad \delta = 2(\mathbf{b}_0^T \mathbf{u}_k \sigma_k^{-1} \mathbf{v}_k^T \mathbf{H} \Delta \mathbf{A}_{k-1}^+ \mathbf{b}_0 - \\ - \mathbf{b}_0^T \mathbf{u}_k \sigma_k^{-1} \mathbf{v}_k^T \mathbf{H} \mathbf{v}_k \sigma_k^{-1} \mathbf{u}_k^T \mathbf{b}_0).$$

Из последнего выражения видно, что при $\delta > 0$ квадрат нормы $\|\mathbf{T}_k \mathbf{b}_0 - \mathbf{d}_0\|^2$ будет убывать с ростом k ; при $\delta < 0$ квадрат нормы $\|\mathbf{T}_k \mathbf{b}_0 - \mathbf{d}_0\|^2$ будет возрастать.

Знак δ зависит от конкретного \mathbf{b}_0 . Рассмотрим случай, когда вектор \mathbf{b}_0 представлен реализацией случайного процесса ξ ($\mathbf{b}_0 = \xi$) с нормальным законом распределения, нулевым средним и дисперсией v^2 . Усреднение

$\|\mathbf{C} \Delta \mathbf{A}_k^+ \xi\|^2$ по реализациям случайного процесса ξ приводит к выражению

$$E_{\varepsilon} \|\mathbf{C} \Delta \mathbf{A}_k^+ \xi\|^2 = v^2 \text{trace}(\Delta \mathbf{A}_k^{+T} \mathbf{H} \Delta \mathbf{A}_k^+) = \\ = v^2 \text{trace}(\mathbf{N}_k), \quad (14)$$

где $\mathbf{N}_k = \mathbf{H} \Delta \mathbf{A}_k^+ \Delta \mathbf{A}_k^{+T}$.

$$E_{\varepsilon} \|\mathbf{C} \Delta \mathbf{A}_k^+ \xi\|^2 = v^2 \text{trace}(\mathbf{N}_{k-1}) - \\ - v^2 \sigma_k^{-2} \text{trace}(\mathbf{v}_k^T \mathbf{C}^T \mathbf{C} \mathbf{v}_k). \quad (15)$$

Из рекурсивного выражения (15) и положительности $v^2 \sigma_k^{-2} \text{trace}(\mathbf{v}_k^T \mathbf{C}^T \mathbf{C} \mathbf{v}_k)$ следует, что (в случае, когда вектор \mathbf{b}_0 представлен реализацией случайного процесса) детерминированная составляющая ошибки преобразования выхода убывает с ростом k . Факт возрастания стохастической и убывания детерминированной составляющей ошибки преобразования выхода с ростом k говорит о том, что при определенном соотношении составляющих ошибки преобразования выхода может иметь минимум при $k < N$. В этом случае число компонент сингулярного разложения в (8) необходимо выбирать таким, чтобы ошибка преобразования выхода была минимальна. Чтобы выбрать k , при котором ошибка решения близка к минимальной в реальных условиях, т.е. когда точное решение неизвестно, предлагаем [5] использовать критерии выбора модели [9–11].

Подход к нахождению числа компонент сингулярного разложения, близкого к оптимальному с использованием критериев выбора модели, рассмотрим на примере критерия минимальной длины описания. Выражение для критерия MDL при $R^2 \geq k/N$ имеет следующий вид:

$$MDL_k = 0,5L \log((\mathbf{r}_k^T \mathbf{r}_k)/(N-k)) + \\ + 0,5k \log((\mathbf{b}^T \mathbf{b} - \mathbf{r}_k^T \mathbf{r}_k)/kS) + \log(L), \quad (16)$$

где $\mathbf{r}_k = \mathbf{A} \mathbf{A}_k^+ \mathbf{b} - \mathbf{b}$, $\mathbf{A}_k^+ = \sum_{i=1}^k \mathbf{v}_i \sigma_i^{-1} \mathbf{u}_i^T$, $R^2 = 1 - (\mathbf{r}_k^T \mathbf{r}_k)/(\mathbf{b}^T \mathbf{b})$.

Близким к оптимальному будем считать такое число компонент сингулярного разложения, для которого значение MDL_k минимально.

Экспериментальное исследование [5] показало, что тенденция к понижению сложности модели при нарастании уровня шума зачастую

приводит к тому, что сложность модели (число компонент сингулярного разложения), полученная по критерию выбора модели, оказывается значительно меньше оптимальной. Это приводит к нарастанию ошибки и стимулирует поиск подходов к поиску оптимального k , свободных от данного недостатка.

Рассмотрим другой подход к нахождению числа компонент сингулярного разложения, близкого к оптимальному. Ошибка решения задачи преобразования выхода зависит от ошибки оценивания \mathbf{x} . Для оценки числа компонент сингулярного разложения, обеспечивающего ошибку, близкую к минимальной, предлагаем использовать условие Пикара.

Условие состоит в том, что коэффициенты $|\mathbf{u}_i^T \mathbf{b}|$ должны убывать быстрее, чем сингулярные значения σ_i , в противном случае оценка \mathbf{x} (17) неустойчива.

$$\mathbf{x}' = \sum (\mathbf{u}_i^T \mathbf{b}) / \sigma_i \mathbf{v}_i. \quad (17)$$

В случае, когда $|\mathbf{u}_i^T \mathbf{b}|$ больше, чем σ_i , вклад в оценку \mathbf{x}' определяется преимущественно членами суммы, соответствующими большим сингулярным значениям (гладким сингулярным векторам), чем обеспечивается гладкость решения \mathbf{x}' . В случае, когда $|\mathbf{u}_i^T \mathbf{b}|$ меньше, чем σ_i , вклад в оценку \mathbf{x}' определяется членами суммы, соответствующими малым сингулярным значениям, относящимся к сильно знакопеременным сингулярным векторам (рис. 2), что приводит к увеличению ошибки решения. Поэтому предлагается считать близким к оптимальному такое число компонент сингулярного разложения, для следующего за которым значения коэффициентов $|\mathbf{u}_i^T \mathbf{b}|$ становятся меньше, чем соответствующие σ_i .

Далее исследуем подходы к нахождению близкого к оптимальному числа компонент сингулярного разложения, основанные на критериях выбора модели и по условию Пикара.

Исследование

Была изучена зависимость ошибки преобразования выхода от числа компонент сингулярного разложения. Столбцы матриц \mathbf{A} и \mathbf{C} содержат m отсчетов радиальных базисных функций: $f_n(z) = \exp(-g(z-c)^2)$, $c = dn + b$, ($d=5$, $b=20$), $z = \{1, 5, 10, \dots, 100\}$, n – номер базисной функ-

ции. Для исходной линейной системы $g=0,5$. На рис. 1 приведены ряды сингулярных чисел матриц. Компоненты вектора \mathbf{x} в экспериментах были следующими: $x_5 = 1$, $x_6 = 0,5$, $x_{10} = 1$, $x_{11} = 0,26$, $x_{12} = 0,26$, другие компоненты вектора \mathbf{x} – нулевые. Векторы \mathbf{b}_0 , \mathbf{d}_0 получены как \mathbf{Ax} , \mathbf{Cx} соответственно. В качестве шума использована случайная величина с гауссовым распределением и среднеквадратичным отклонением $\{0,1; 0,3; 0,6\}$. На рис. 3–5 для трех уровней шума приведена зависимость ошибки e от k для системы C при $g = \{0,3; 0,5; 0,7\}$. Экспериментально исследованы зависимости: ошибки e при различных матрицах \mathbf{C} ($g = \{0,5; 0,2; 0,7\}$); значения критерии Акаике, Маллоуза и Бин Ю от k ; коэффициентов Пикара от k .

Зависимость ошибки e при различных матрицах \mathbf{C} ($g = \{0,5; 0,2; 0,7\}$) и значений критерия Акаике от k приведены на рис. 2–4. Видно, что зависимость e от k имеет минимум при $k < n$. С ростом уровня шума положение минимума смещается в область меньших значений k , а ошибка в точке минимума растет, что происходит также с ростом различия между базисными функциями линейных систем A и C .

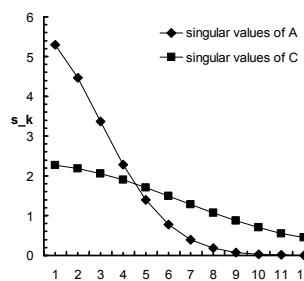


Рис. 1. Ряды сингулярных чисел матриц \mathbf{A} и \mathbf{C}

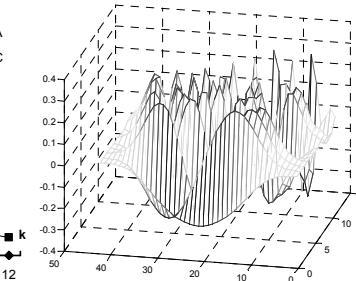
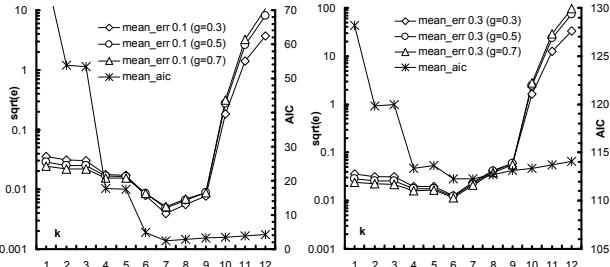


Рис. 2. Сингулярные вектора $\mathbf{u}_1, \dots, \mathbf{u}_{15}$ для матрицы \mathbf{A}

Минимум критерия Акаике в данной задаче совпадает с минимумом ошибки e , чем может быть обеспечена достаточно высокая точность редукции измерений. Значения ошибки преобразования выхода при определении k по критериям Акаике, Маллоуза, Бин Ю и условию Пикара приведены в таблице.

Таблица

nl	e	AIC	Cp	MDL	$Picard Cond.$
0,1	0,003914	0,003914	0,017816	0,003914	0,005577
0,3	0,012754	0,012754	0,019335	0,019335	0,023751
0,6	0,023583	0,023583	0,032667	0,032667	0,027731



Зависимости величины ошибки e от числа компонент сингулярного разложения k при разных уровнях шума
Рис. 3. При уровне шума 0,2 Рис. 4. При уровне шума 0,3

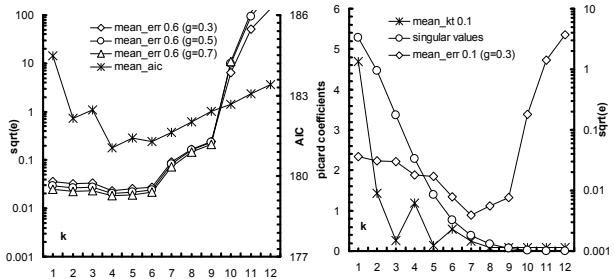


Рис. 5. Зависимость величины ошибки e от числа компонент сингулярного разложения k при уровне шума 0,6

Рис. 6. Зависимость коэффициентов Пикара от k при уровне шума 0,1

Зависимость коэффициентов Пикара от k для различных уровней шума показана на рис. 6–8.

Оптимальное значение k , полученное в соответствии с условием Пикара, приближается к минимуму ошибки e , чем может быть обеспечена достаточно высокая точность преобразования выхода линейной системы в выход системы с заданным базисом.

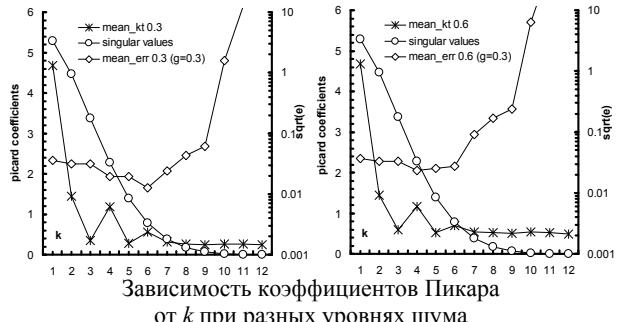


Рис. 7. При уровне шума 0,3 Рис. 8. При уровне шума 0,6

Заключение. Итак, предложенный метод, позволяет для линейной системы, вектор выхода которой \mathbf{b} формируется путем линейного преобразования входа \mathbf{x} и добавления аддитивного шума ($\mathbf{A}\mathbf{x} + \boldsymbol{\varepsilon} = \mathbf{b}$), отыскать преобразование наблюдаемого выхода \mathbf{b} в выход линейной системы с известным набором базисных функций \mathbf{C} , дающих выход с требуемыми свойствами.

Метод обеспечивает устойчивое преобразование выхода линейной системы в выход системы с заданным базисом для случая, когда матрица \mathbf{A} базисных функций

исходной линейной системы имеет высокое число обусловленности и ряд ее сингулярных чисел плавно спадает к нулю.

Проведенное исследование подходов к определению k с использованием критериев выбора модели и использующего условие Пикара, позволяет определить число компонент сингулярного разложения k , при котором ошибка метода преобразования выхода близка к минимуму.

Минимум критерия Акайке в данной задаче совпадает с минимумом ошибки e , что может обеспечить достаточно высокую точность преобразования выхода. Минимумы критериев Маллоуза и Бин Ю близки, но не совпадают с минимумом ошибки e . Соответственно точность преобразования выхода для критериев Маллоуза и Бин Ю меньше, чем для критерия Акайке. Значения ошибки преобразования выхода при определении k по условию Пикара близки к оптимальным.

1. Пытьев Ю.П. Математические методы интерпретации эксперимента. – М.: Высшая школа., 1989. – 351 с.
2. Пытьев Ю.П. Псевдообратный оператор. Свойства и применения // Математический сборник. – 1982. – **118**(160), № 1 (5). – С. 19–49.
3. Касьянюк С.А. К задаче интерпретации данных, полученных наблюдениями с помощью дискретного множества датчиков // ЖВМиМФ. – 1981. – **18**, № 3. – С. 553–560.
4. Tikhonov A.N., Arsenin V.Y. Solution of ill-posed problems. V.H. Winston, Washington, DC, 1977. – 285 p.
5. Ревунова Е.Г. Повышение точности оценки вектора параметров путем преобразования линейной системы к системе с заданными свойствами // Индуктивное моделирование сложных систем. – 2011. – 3. – С. 165–173.
6. Revunova E.G., Rachkovskij D.A. Stable Transformation of a Linear System Output to the Output of System with a Given Basis by Random Projections // The 5th Int. Workshop on Inductive Modelling (IWIM'2012), Kyiv. – 2012. – 1. – P. 37–41.
7. Забулонов Ю.Л., Ревунова Е.Г. Исследование составляющих ошибки для решения задачи редукции измерений с использованием сингулярного разложения // Сб. научн. тр. ИПМЭ НАН Украины «Моделирование и информационные технологии». – 2012. – № 63. – С. 175–185.
8. Hansen P.C. Rank-deficient and discrete ill-posed problems. Numerical aspects of linear inversion. – Philadelphia: SIAM, 1998. – 247 p.
9. Mallows C.L. Some comments on Cp // Technometrics. – 1973. – **15**, № 4. – P. 661–675.
10. Akaike H. A new look at the statistical model identification // IEEE Transactions on Automatic Control. – 1974. – **19**, N 6. – P. 716–723.
11. Hansen M., Yu B. Model selection and minimum description length principle // J. Amer. Statist. Assoc. – 2001. – **96**. – P. 746–774.

E-mail: helab@i.com.ua
© Е.Г. Ревунова, 2013