УДК 004.942 + 623.454.862

ОПРЕДЕЛЕНИЕ МИНИМУМА ОШИБКИ С ИСПОЛЬЗОВАНИЕМ КРИТЕРИЕВ ВЫБОРА МОДЕЛИ ПРИ РЕШЕНИИ ОБРАТНОЙ ЗАДАЧИ МЕТОДОМ СЛУЧАЙНЫХ ПРОЕКЦИЙ

Е.Г. Ревунова

Международный научно-учебный центр информационных технологий и систем НАН и МОН Украины

helab@i.com.ua

Проведений експериментальний аналіз точності рішень дискретної некоректної зворотної задачі методом випадкових проекцій. Визначення мінімуму помилки здійснювалося з використанням критеріїв вибору моделі. Метод випадкових проекцій з використанням критеріїв вибору моделі забезпечує точність на рівні регуляризації Тихонова.

Ключові слова: дискретна некоректна зворотна задача, псевдозвернення, регуляризація, проектування, критерії вибору моделі.

The analysis of decisions exactness of discrete ill-posed inverse problem by random projections method has been performed. Random projections method with model selection criteria has shown accuracy similar to Tikhonov regularization.

Keywords: discrete ill-posed inverse problems, pseudo-inverse, regularization, projection, model selection.

Проведен экспериментальный анализ точности решений дискретной некорректной обратной задачи методом случайных проекций. Определение минимума ошибки осуществлялось с использованием критериев выбора модели. Метод случайных проекций с использованием критериев выбора модели обеспечивает точность на уровне регуляризации Тихонова.

Ключевые слова: дискретная некорректная обратная задача, псевдообращение, регуляризация, проецирование, критерии выбора модели.

Введение

Во многих практических приложениях требуется решать дискретную обратную задачу [1, 2] вида:

$$\mathbf{\Phi x}=\mathbf{y}, \tag{1}$$

где матрица $\mathbf{\Phi}\in\Re^{N\times N}$ и вектор $\mathbf{y}\in\Re^{N}$, искаженный аддитивным шумом $\boldsymbol{\varepsilon}\in\Re^{N}$ $\mathbf{y}=\mathbf{y}_0+\boldsymbol{\varepsilon}$, известны и получены в результате оцифровки интегрального уравнения первого рода. Требуется оценить вектор сигнала $\mathbf{x}\in\Re^{N}$.

В случае, когда $\mathbf{y}$ содержит шум, ряд сингулярных чисел σ_i матрицы $\mathbf{\Phi}$ плавно спадает к нулю, $\mathbf{\Phi}$ имеет высокое число обусловленности $\sigma_{max}/\sigma_{min}$, задачу оценки $\mathbf{x}$ называют дискретной некорректной обратной задачей [1]. Такие свойства $\mathbf{\Phi}$ характерны для задач спектрометрии [3], гравиметрии [4], электроразведки [5].

Решение дискретной некорректной обратной задачи как задачи наименьших квадратов [2]

$$\mathbf{x}'=\mathrm{argmin}_{\mathbf{x}}\ \|\mathbf{y}-\mathbf{\Phi x}\|_2 \tag{2}$$

на основе псевдообращения

$$\mathbf{x}'=\mathbf{\Phi}^{+}\mathbf{y} \tag{3}$$

является неустойчивым [1, 2]. Признаком неустойчивости является то, что малым изменениям вектора $\mathbf{y}$ соответствуют большие изменения решения $\mathbf{x}'$; при этом велико значение ошибки решения.

Для преодоления неустойчивости и, соответственно, повышения точности решения используют регуляризацию [1, 2]. Регуляризация накладывает некоторые ограничения на решения, которые позволяют повысить их устойчивость – например, малость l_2-нормы решения $\|\mathbf{x}'\|_2$. Классическим методом регуляризации является регуляризация Тихонова [2]. Задачу регуляризации Тихонова формулируют следующим образом

$$\mathbf{x}'=\mathrm{argmin}_{\mathbf{x}}\ (\ \|\mathbf{y}-\mathbf{\Phi x}\|_2+\lambda\|\mathbf{x}\|_2\ ), \tag{4}$$

где λ – параметр регуляризации.

Недостатками, присущими методам решения дискретных некорректных обратных задач на основе регуляризации Тихонова, являются высокая вычислительная сложность и сложность подбора правильного параметра регуляризации, от которого в значительной мере зависит устойчивость решения. Поэтому востребованными являются альтернативные подходы к решению дискретной некорректной обратной задачи с точностью на уровне регуляризации Тихонова, но с меньшей вычислительной сложностью.

Нами разрабатывается подход к решению дискретной некорректной обратной задачи с использованием методов псевдообращения и случайных проекций [6, 7]. В данной работе приводятся результаты исследования точности восстановления истинного сигнала при определении минимума ошибки метода случайных проекций с использованием критериев выбора модели.

1. Решение дискретной некорректной обратной задачи методом случайных проекций

Нами разрабатывается подход к решению дискретной некорректной обратной задачи, использующий проекционную версию рандомизированных алгоритмов приближения матриц [8]. В качестве проектора $\mathbf{\Omega}\in\Re^{k?N}$ используется матрица с элементами, сформированными реализациями случайной величины [9, 10].

Для решения обратной задачи с использованием проекционного подхода [9, 6, 7] умножим обе части исходного уравнения (1) на матрицу $\mathbf{\Omega}\in\Re^{k?N}$, $k\leq N$, элементы которой – реализации случайной величины с нормальным распределением, нулевым средним и единичной дисперсией. Число столбцов N матрицы $\mathbf{\Omega}$ определяется размерностью исходной матрицы $\mathbf{\Phi}$, число строк k априори не фиксировано.

Получим уравнение

$$\mathbf{\Omega\Phi x} = \mathbf{\Omega y},\ \text{где}\ \mathbf{\Omega\Phi}\in\Re^{k?N},\ \mathbf{\Omega y}\in\Re^{k}. \tag{5}$$

Тогда задача наименьших квадратов (2) записывается в виде

$$\mathbf{x}' = \mathrm{argmin}_{\mathbf{x}} \|\mathbf{\Omega\Phi x}-\mathbf{\Omega y}\|_2. \quad (6)$$

Восстановление сигнала **x** на основе псевдообращения с использованием случайного проецирования получим, как

$$\mathbf{x}'=(\mathbf{\Omega\Phi})^{+}\mathbf{\Omega y}. \quad (7)$$

Восстановление сигнала на основе псевдообращения с использованием QR разложения матрицы **ΩΦ** получим, как

$$\mathbf{x}'=(\mathbf{Q}^{\mathrm{T}}\mathbf{\Phi})^{+}\mathbf{Q}^{\mathrm{T}}\mathbf{y}. \quad (8)$$

Точность решения обратной задачи будем оценивать с помощью ошибки d восстановления истинного сигнала **x**, вычисляемой как

$$d=\|\mathbf{x}-\mathbf{x}'\|=\|\mathbf{e}\|, \quad (9)$$

где **x**' – вектор восстановленного сигнала, **e** – вектор ошибки решения.

Анализ зависимостей d от k показывает [7], что для всех методов при определенном значении k зависимость имеет минимум, положение которого при возрастании уровня шума смещается в область меньших значений k, а значение ошибки в точке минимума возрастает.

Определение оптимального значения k (соответствующего минимальному значению ошибки) непосредственно по выражению (9) представляет только теоретический интерес, т.к. выражение включает вектор истинного сигнала, который в практических задачах неизвестен.

В данной работе предлагается оценивать k соответствующее минимуму ошибки d путем минимизации критерия выбора модели. Наиболее известными критериями выбора модели являются критерии: C_P Маллоуза [11], *AIC* Акаике [12], *gMDL* Бин Ю [13].

Выражение для критерия *gMDL* с учетом проецирования имеет вид:

$$\begin{aligned} gMDL &= 0.5L\log(S) + 0.5k\log(F) + \log(L) \text{ при } R^2 \geq k/N, \\ gMDL &= 0.5L\log(\mathbf{b}^{\mathrm{T}}\mathbf{b}/L) + 0.5\log(L) \text{ – иначе,} \end{aligned} \quad (10)$$

где $S=RSS/(N-k)$, $F=(\mathbf{b}^{\mathrm{T}}\mathbf{b}-RSS)/kS$, R – множественный коэффициент корреляции, $R^2= 1 - RSS/(\mathbf{b}^{\mathrm{T}}\mathbf{b})$, $RSS = \|\mathbf{AA}^{+}\mathbf{b} - \mathbf{b}\|_2$

Для проецирования случайной матрицей:

$$\mathbf{b}=\mathbf{\Omega y},\ \mathbf{A}=\mathbf{\Omega\Phi}. \quad (11)$$

Для проецирования с использованием QR разложения:

$$\mathbf{b}=\mathbf{Q y},\ \mathbf{A}=\mathbf{Q\Phi}. \quad (12)$$

Проведем экспериментальное исследование точности оценивания оптимального значения k с использованием критериев выбора модели: C_P Маллоуза, *AIC* Акаике, *gMDL* Бин Ю.

2. Экспериментальное исследование

Исследуем экспериментально зависимость ошибки восстановления сигнала d от числа строк k для матриц проекторов **R** и **Q**.

Столбцы матрицы **Ф** были сформированы радиальными базисными функциями вида $f_n(z)=\exp(-k_1(z-c)^2)$; где $c=d?n+b$; (d=5, b=20), k={0.001, 0.01, 0.05, 0.06, 0.1}, z={1,2,…,100}, n – номер базисной функции. Компоненты вектора x следующие: x(60)=100; x(61)=98; x(62)=90; x(63)=70; x(64)=60; x(65)=30; x(65)=20; x(59)=98; x(58)=90; x(57)=70; x(56)=60; x(55)=30; x(54)=20; не указанные компоненты равны нулю. Вектор правой части **y** формировался путем умножения матрицы **Ф** на вектор **x** согласно выражению (1).

В данной задаче матрица **Ф** имела размерность 200?200, высокое число обусловленности ($\sigma_{max}/\sigma_{min}$>>1), и ряд сингулярных чисел, плавно спадающий к нулю. Вектор правой части уравнения (1) искажался аддитивным шумом ε с нормальным распределением.

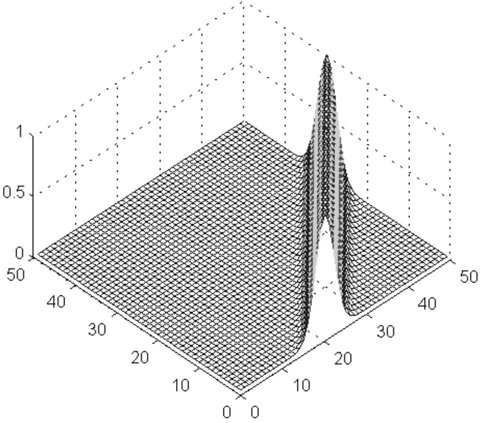

Рис.1 Подматрица (50×50) матрицы Ф

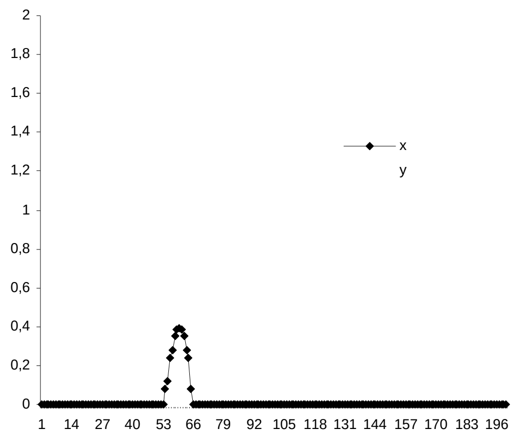

Рис.2 Вектор сигнала x и правой части y

На рис. 3 приведены зависимости ошибки восстановления сигнала d от числа строк k матрицы проектора при уровнях шума 10^{-6}, 10^{-4}, 10^{-2} для проецирования матрицами: **R**(k?N) Dx Q, и **Q**(k?N) Dx R.

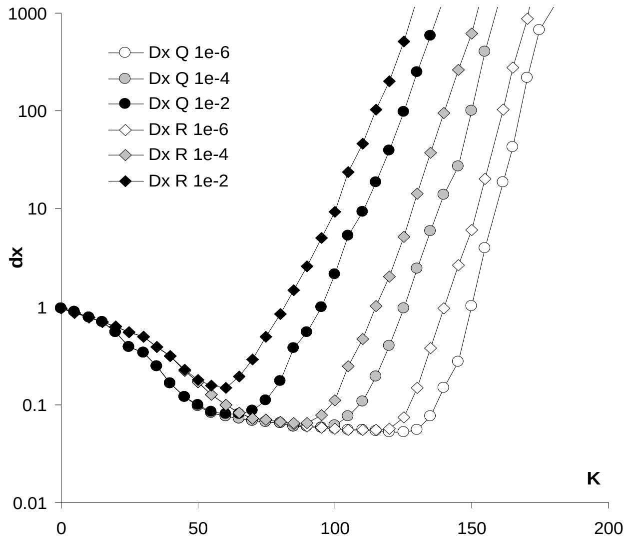

Рис.3 Зависимость ошибки восстановления сигнала d от числа строк k матриц R и Q

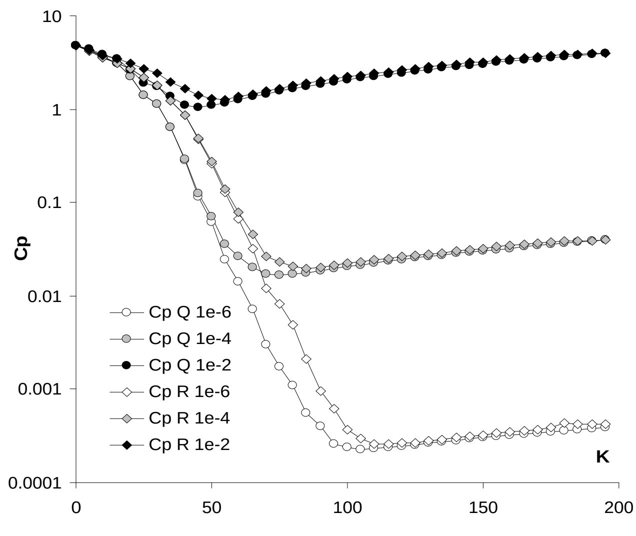

Рис.4 Зависимость значения критерия Маллоуза от числа строк k матриц R и Q

Анализ зависимостей d от k показывает, что для всех методов при определенном значении k зависимость имеет минимум, положение которого при возрастании уровня шума смещается в область меньших значений k, а значение ошибки в точке минимума возрастает.

При использовании в качестве проектора матрицы **Q** значение ошибки в точке минимума для всех уровней шума меньше чем при использовании матрицы **R**.

На рис. 4-6 приведены зависимости значений критериев выбора модели (Акаике *AIC*, Маллоуза *Cp*, Бин Ю *MDL*) от числа строк k матрицы проектора при уровнях шума 10^{-6}, 10^{-4}, 10^{-2} для проецирования матрицами **R** и **Q**.

Анализ зависимостей значений критериев выбора модели от k показывает, что для всех критериев зависимость имеет единственный минимум, положение которого при возрастании уровня шума смещается в область меньших значений k, а значение ошибки в точке минимума возрастает.

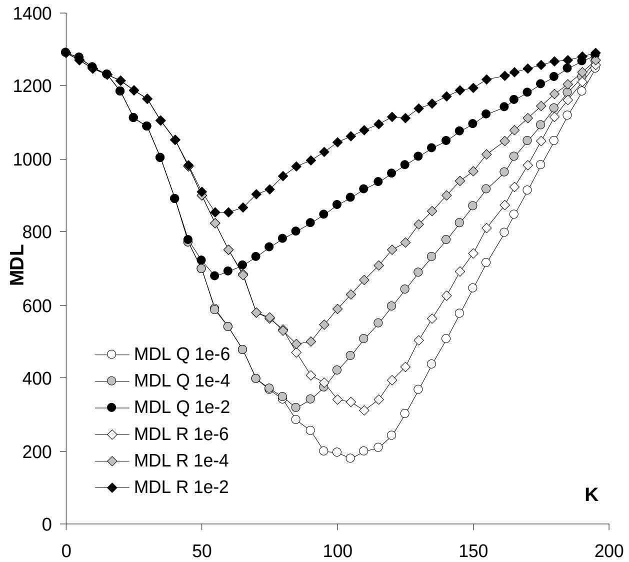

Рис.5 Зависимость значения критерия Бин Ю от числа строк k матриц R и Q

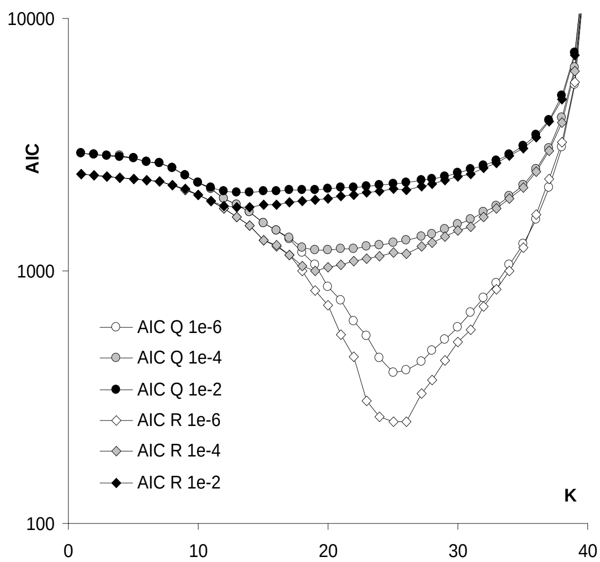

Рис.6 Зависимость значения критерия Акаике от числа строк k матриц R и Q

Отметим, что число строк матрицы проектора k_{minCR}, при котором значение критерия выбора модели минимально не совпадает в точности со значением k_{minDx} при котором минимально значение ошибки восстановления сигнала вычисленной с использованием вектора истинного сигнала.

Несовпадение k_{minCR} и k_{minDx} является причиной приращения ошибки восстановления сигнала относительно теоретической при использовании метода поиска оптимального k на основе критериев выбора модели.

Значения ошибки для методов без проецирования и минимальные значения ошибки для методов с проецированием приведены в таблице 1. Сравнивались значения ошибки для следующих методов, работающих без проецирования: регуляризация Тихонова с подбором параметра регуляризации методом перекрестной проверки *GCV*, методом *L*-кривой *Lcur* и метод на основе псевдообращения *pin*.

Таблица 1

nl	$d_{\min}$ методов с проецированием								*d* методов без проецирования		
	dx		*Cp*		*AIC*		*MDL*				
	Q	*R*	*Q*	*R*	*Q*	*R*	*Q*	*R*	*pin*	*GCV*	*Lcur*
10^{-6}	0,052	0,054	0,056	0,055	0,053	0,585	0,055	0,055	10^{3}	0,05	0,11
10^{-4}	0,058	0,064	0,068	0,642	0,059	0,072	0,06	0,064	10^{5}	0,06	0,17
10^{-2}	0,08	0,147	0,129	0,18	0,08	0,162	0,085	0,148	10^{7}	0,08	0,17

Среди исследованных методов, работающих без проецирования, наименьшую ошибку обеспечивает регуляризация Тихонова с подбором параметра регуляризации методом перекрестной проверки. Среди методов с проецированием наименьшую ошибку обеспечивают методы использующие проецирование Q матрицей с критериями выбора модели *AIC* и *MDL*.

3. Выводы

Проецирование снижает ошибку решения регуляризацией Тихонова в случае, когда до проецирования ошибка решения велика. Ошибка решения псевдообращением без проецирования на несколько порядков превышает ошибки остальных методов, и метод псевдообращения без проецирования непригоден для решения рассмотренных обратных задач.

При выборе *k* с использованием критериев выбора модели, точность решения методом на основе псевдообращения матрицы с проецированием находится на уровне лучшего метода регуляризации Тихонова без проецирования. Среди методов с проецированием наименьшую ошибку обеспечивают методы использующие проецирование Q матрицей.

Использование методов на основе псевдообращения с проецированием является перспективным в силу снижения вычислительных затрат. Это снижение происходит из-за уменьшения вычислительной сложности сингулярного разложения матрицы **A** при *k*, составляющих малую долю *N* (что особенно проявляется при увеличении уровня шума), по сравнению с сложностью сингулярного разложения исходной **Ф**.

Литература

1. Hansen P.C. Rank-deficient and discrete ill-posed problems. Numerical Aspects of Linear Inversion. – SIAM, Philadelphia, 1998. – 247 p.

2. Tikhonov A.N., Arsenin V.Y, Solution of ill-posed problems. V.H. Winston, Washington, DC, 1977.

3. Забулонов Ю.Л., Коростиль Ю.М., Ревунова Е.Г. Оптимизация решения обратной задачи по восстановлению функции плотности распределения поверхностных загрязнений // Сборник научных трудов ИПМЭ НАН Украины “Моделирование и информационные технологии”. – 2006. – С. 77 – 83.

4. Булах Е.Г. О новом аппроксимационном подходе к решению обратных задач гравиметрии в классе трехмерных контактных поверхностей // Докл НАНУ. 2006. № 1. С. 108–112.

5. Хмелевский В.К., Бондаренко В.М. Электроразведка. – М.: Недра, 1999. – 438с.

6. Ревунова Е.Г., Рачковский Д.А. Повышение точности решения обратной задачи с использованием случайных проекций, // 16-th Intern. Conf. Knowledge-Dialogue-Solution. – 2009– V.10. – P. 95–102.

7. Revunova E.G. Study of error components for solution of the inverse problem using random projections // Conf. on Inductive Modeling (ICIM-10). – Evpatoria. – 2010. – V.1. – P.1–7.

8. Halko N., Martinsson P.G., Tropp J.A. Finding structure with randomness: Stochastic algorithms for constructing approximate matrix decompositions // ACM Report Caltech. – 2009. – №. 2009-5. – P.1–82.

9. Sarlos T. Improved approximation algorithms for large matrices via random projections // In Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science. – 2006. – P.143–152.

10. Drineas P., Mahoney M.W., Muthukrishnan S., Sarlos T. Faster least squares approximation // Tech. Rep. 0710.1435. – 2007.

11. Mallows C.L. Some comments on Cp // Technometrics. – 1973. – V. 15, № 4. – P. 661-675.

12. Akaike H. A new look at the statistical model identification // IEEE Transactions on Automatic Control. – 1974. – V. 19, № 6. – P. 716-723.

13. Hansen M., Yu B. Model selection and minimum description length principle //J. Amer. Statist. Assoc. – 2001. – V. 96. – P. 746-774.