УДК 004.942 + 623.454.862

ПОДХОД РАНДОМИЗАЦИИ В ЗАДАЧАХ ВОССТАНОВЛЕНИЯ СИГНАЛОВ ПО РЕЗУЛЬТАТАМ НЕПРЯМЫХ ИЗМЕРЕНИЙ

Ревунова Е.Г.
Международный научно-учебный центр информационных технологий и систем НАН и МОН Украины

Введение

Задача восстановления истинного сигнала объекта по результатам косвенных дистанционных измерений часто возникает в технических системах. Сигнал объекта претерпевает искажения при взаимодействии со средой распространения и системой детектирования. Результаты косвенных измерений объектов внешней среды формируют сигнальный образ объекта наблюдения, по которому необходимо восстановить истинный сигнал. Такая необходимость существует в практике обработки сигналов в различных областях - геофизическая разведка, спектрометрия, выявления присутствия отдельных химических элементов в объектах, медицинская диагностика (визуализация, томография), неразрушающий контроль, дистанционный мониторинг. При восстановлении сигналов объекта наблюдения по результатам косвенных дистанционных измерений, возникает необходимость решения обратной задачи. При этом матрица, характеризующая взаимодействие сигнала со средой, имеет большое число обусловленности, ряд ее сингулярных значений плавно спадает до нуля, а вектор правой части уравнения, как правило, искажен аддитивным шумом. При таких условиях это дискретная некорректная обратная задача. Известно, что ее решение, как задачи наименьших квадратов на основе псевдообращения матрицы, является неустойчивым. Для преодоления неустойчивости и повышения точности решения дискретных некорректных задач были предложены методы регуляризации [1, 2]. Регуляризация накладывает на искомое решение ограничения, обеспечивающие устойчивость. Недостатками методов решения дискретных некорректных задач на основе регуляризации Тихонова является высокая вычислительная сложность и трудность выбора надлежащего параметра регуляризации, влияющего на устойчивость решения. Это определяет **цель** работы – развитие альтернативных подходов к решению дискретных некорректных задач, которые должны обеспечивать точность, сравнимую с точностью регуляризации Тихонова, но при более низких вычислительных затратах.

Рассмотрим линейные задачи наименьших квадратов и выделим класс задач – дискретные некорректные задачи.

Линейные задачи наименьших квадратов и дискретные некорректные задачи

Многие приложения математики, физики, анализа данных и др. требуют нахождения приближенного решения системы линейных уравнений:

$$\mathbf{Ax} \approx \mathbf{y}, \qquad (1)$$

где матрица $\mathbf{A} \in \Re^{N \times N}$ и вектор $\mathbf{y} \in \Re^{N}$, искаженный аддитивным шумом $\boldsymbol{\varepsilon} \in \Re^{N}$ $\mathbf{y}=\mathbf{y}_0+\boldsymbol{\varepsilon}$, известны и требуется оценить вектор сигнала $\mathbf{x} \in \Re^{N}$.

Если матрица $\mathbf{A}$ имеет полный ранг и хорошо обусловлена, оценку $\mathbf{x}'$ для метода наименьших квадратов можно получить путем решения системы нормальных уравнений [2], например, с помощью разложения Холецкого. В случае, когда матрица $\mathbf{A}$ имеет неполный ранг или плохо обусловлена, для получения оценки $\mathbf{x}'$ может использоваться решение на основе QR-разложения. В случае, когда матрица $\mathbf{A}$ очень плохо обусловлена, для получения устойчивой оценки $\mathbf{x}'$ используют разложение по сингулярным значениям (SVD). Идея получения устойчивой оценки $\mathbf{x}'$ на основе сингулярного разложения состоит в следующем: если в спектре сингулярных значений имеется резкий перепад, а сингулярные значения после него очень малы, они могут рассматриваться как шумовые и устраняются применением порога.

В случае, когда $\mathbf{y}$ содержит шум, ряд сингулярных чисел σ_i матрицы $\mathbf{A}$ плавно спадает к нулю, $\mathbf{A}$ имеет высокое число обусловленности $\sigma_{max}/\sigma_{min}$, задачу оценки $\mathbf{x}$ называют дискретной некорректной обратной задачей [1]. Прием «усечения» сингулярного разложения не работает в дискретных некорректных задачах, поскольку, как отмечалось выше, здесь нет разрывов в ряду сингулярных значений и численный ранг не определен.

Для преодоления неустойчивости и, соответственно, повышения точности решения используют регуляризацию [1,2]. 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, использующий проекционную версию рандомизированных алгоритмов приближения матриц [3, 4].

Восстановление истинного сигнала с использованием случайного проектора

Для решения на основе проекционного подхода обе части исходного уравнения (1) умножим на матрицу $\mathbf{R}_k \in \Re^{k \times N}$, получив уравнение:

$$\mathbf{F}_k\mathbf{x} = \mathbf{b}_k, \text{ где } \mathbf{F}_k = \mathbf{R}_k\mathbf{A},\ \mathbf{F}_k \in \Re^{k \times N},\ \mathbf{b}_k = \mathbf{R}_k\,\mathbf{y},\ \mathbf{b}_k \in \Re^{k}. \qquad (2)$$

Зададим матрицу $\mathbf{G}\in\Re^{N\times N}$, элементы которой – реализации случайной величины с нормальным распределением, нулевым средним и единичной дисперсией. Матрицу $\mathbf{R}$ получим как результат собственного разложения $\mathbf{G=R\Sigma R}^{\mathrm{T}}$, где $\mathbf{R}\in\Re^{N\times N}$ – ортонормированная матрица, $\mathbf{\Sigma}$ – диагональная. Матрицу $\mathbf{R}_k$ получим отбором k строк матрицы $\mathbf{R}$.
Оценку сигнала $\mathbf{x}$ на основе псевдообращения получим, как:

$$\mathbf{x}'_k = \mathbf{F}_k^+\mathbf{b}_k. \tag{3}$$

Запишем выражение для среднеквадратичной ошибки восстановления истинного сигнала при проецировании случайной матрицей $\mathbf{R}_k$:

$$e_k = \mathrm{E}\,\|(\mathbf{F}_k^+\mathbf{F}_k - \mathbf{I})\mathbf{x}\|^2 + \mathrm{E}\,\|\mathbf{F}_k^+\mathbf{R}_k\boldsymbol{\varepsilon}\|^2 + 2\,\mathrm{E}\,\langle(\mathbf{F}_k^+\mathbf{F}_k - \mathbf{I})\mathbf{x},\, \mathbf{F}_k^+\mathbf{R}_k\boldsymbol{\varepsilon}\rangle. \tag{4}$$

Так как $2\,\mathrm{E}\,\langle(\mathbf{F}_k^+\mathbf{F}_k - \mathbf{I})\mathbf{x},\, \mathbf{F}_k^+\mathbf{R}_k\boldsymbol{\varepsilon}\rangle = 0$, и $\mathrm{E}\,\|\mathbf{F}_k^+\mathbf{R}_k\boldsymbol{\varepsilon}\|^2 = \sigma^2\,\mathrm{trace}\,(\mathbf{R}_k^{\mathrm{T}}\mathbf{F}_k^{+\mathrm{T}}\mathbf{F}_k^+\mathbf{R}_k)$

$$e_k = \|(\mathbf{F}_k^+\mathbf{F}_k - \mathbf{I})\mathbf{x}\|^2 + \sigma^2\,\mathrm{trace}\,(\mathbf{R}_k\mathbf{R}_k^{\mathrm{T}}\,\mathbf{F}_k^{+\mathrm{T}}\mathbf{F}_k^+). \tag{5}$$

Рассмотрим зависимость величины стохастической составляющей ошибки решения дискретной некорректной задачи от числа строк матрицы проектора $\mathbf{R}_k$. С учетом того факта, что $\mathbf{R}_k\mathbf{R}_k^{\mathrm{T}} = \mathbf{I}$, стохастическая составляющая ошибки есть:

$$e_{1k} = \mathrm{E}\,\|\mathbf{F}_k^+\mathbf{R}_k\boldsymbol{\varepsilon}\|^2 = \sigma^2\,\mathrm{trace}\,(\mathbf{F}_k^{+\mathrm{T}}\mathbf{F}_k^+). \tag{6}$$

Для исследования зависимости стохастической составляющей ошибки от k, необходимо записать выражение для e_{1k} в рекурсивном виде. Воспользовавшись представлением возмущения псевдообратной матрицы ($\mathbf{F}_k^+$) через возмущение исходной матрицы ($\mathbf{F}_k=\mathbf{F}_{k-1}+\mathbf{E}_k$), предложенным в [5], мы получили рекурсивное выражение для $\mathbf{F}_k^{+\mathrm{T}}\mathbf{F}_k^+$ следующего вида:

$$\mathbf{F}_k^{+\mathrm{T}}\mathbf{F}_k^+ = \begin{bmatrix} \mathbf{B}_k & \mathbf{c}_k \\ \mathbf{c}_k^{\mathrm{T}} & d_k \end{bmatrix},\ \mathbf{B}_k = \mathbf{F}_{k-1}^{+\mathrm{T}}\mathbf{F}_{k-1}^+ + \mathbf{M}_k^{\mathrm{T}}\mathbf{M}_k\,, \tag{7}$$

где $\mathbf{M}_k = \mathbf{F}_k^+\mathbf{P}_k\,\mathbf{E}_k\,\mathbf{Z}_k\,\mathbf{F}_{k-1}^+$, $\mathbf{P}_k = \mathbf{F}_k\mathbf{F}_k^+$, $\mathbf{Z}_k = \mathbf{F}_{k-1}^+\mathbf{F}_{k-1}$. Соответственно:

$$e_{1k} ==\sigma^2\,\mathrm{trace}\begin{bmatrix} \mathbf{F}_{k-1}^{+\mathrm{T}}\mathbf{F}_{k-1}^+ + \mathbf{M}_k^{\mathrm{T}}\mathbf{M}_k & \mathbf{c}_k \\ \mathbf{c}_k^{\mathrm{T}} & d_k \end{bmatrix},\ d_k = \mathbf{t}_k^{\mathrm{T}}\mathbf{t}_k,\ \mathbf{t}_k = \mathbf{F}_k^+\mathbf{P}_k\,\mathbf{P}_{k-1}^{\perp}. \tag{8}$$

Отметим, что матрица $\mathbf{B}_k$ имеет положительные элементы на диагонали, т.к. диагонали $\mathbf{F}_{k-1}^{+\mathrm{T}}\mathbf{F}_{k-1}^+$ и $\mathbf{M}_k^{\mathrm{T}}\mathbf{M}_k$ - положительны. Элемент d_k – положителен. Из выражения (7) следует, что матрица $\mathbf{F}_k^{+\mathrm{T}}\mathbf{F}_k^+$ имеет положительную диагональ и рекурсивно изменяющуюся подматрицу $\mathbf{B}_k$. Следовательно, значение стохастической составляющей ошибки возрастает с ростом k.

Рассмотрим зависимость величины детерминированной составляющей ошибки решения дискретной некорректной задачи от числа строк матрицы проектора. Выражение для детерминированной составляющей ошибки имеет следующий вид:

$$e_{2k} = \left\|(\mathbf{F}_k^+\mathbf{F}_k - \mathbf{I})\mathbf{x}\right\|^2 = \mathbf{x}^{\mathrm{T}}\mathbf{x} - \mathbf{x}^{\mathrm{T}}\mathbf{F}_k^+\mathbf{F}_k\mathbf{x}. \tag{9}$$

С целью исследования зависимости детерминированной составляющей ошибки от k, запишем выражение для e_2 в рекурсивном виде:

$$e_{2k} = \mathbf{x}^{\mathrm{T}}\mathbf{x} - (\mathbf{x}^{\mathrm{T}}\mathbf{F}_{k-1}^{+}\mathbf{F}_{k-1}\mathbf{x} + \mathbf{x}^{\mathrm{T}}\mathbf{N}_{k}\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{x} - (\mathbf{x}^{\mathrm{T}}\mathbf{F}_{k-1}^{+}\mathbf{F}_{k-1}\mathbf{x} + \mathbf{x}^{\mathrm{T}}\mathbf{N}_{k}^{1/2\mathrm{T}}\mathbf{N}_{k}^{1/2}\mathbf{x}), \quad (10)$$

где $\mathbf{N}_{\mathrm{k}}$ – квадратная матрица симметричная относительно диагонали:

$$\mathbf{N}_{\mathrm{k}}=\mathbf{F}_{\mathrm{k}}^{+}\mathbf{E}_{\mathrm{k}}\mathbf{Z}_{\mathrm{Fk-1}}^{\perp}. \quad (11)$$

Поскольку $\mathbf{x}^{\mathrm{T}}\mathbf{N}_{k}^{1/2\mathrm{T}}\mathbf{N}_{k}^{1/2}\mathbf{x} > 0$, значение $\mathbf{x}^{\mathrm{T}}\mathbf{F}_{k}^{+}\mathbf{F}_{k}\mathbf{x}$ возрастает с ростом k, а значение детерминированной составляющей ошибки ($\mathbf{x}^{\mathrm{T}}\mathbf{x} - \mathbf{x}^{\mathrm{T}}\mathbf{F}_{k}^{+}\mathbf{F}_{k}\mathbf{x}$) – убывает.

При возрастании k норма детерминированной составляющей уменьшается, и норма стохастической возрастает, так, что суммарная норма ошибки имеет минимум. Такой характер поведения ошибки имел место для всех исследованных нами некорректных обратных задач при уровне шума выше некоторого минимального [3]. Таким образом, для получения решения с минимальной ошибкой необходимо использовать проекционную матрицу с размерностью k, близкой к оптимальной. Однако определение оптимального значения k по графикам зависимости ошибки восстановления истинного сигнала от k представляет только теоретический интерес, т.к. выражение включает вектор истинного решения, который при решении практических задач неизвестен. Чтобы выбрать размерность k проекционной матрицы, при которой ошибка решения близка к минимальной в реальных условиях, т.е. когда точное решение неизвестно, мы предлагаем [3] использовать различные типы критериев выбора k. При этом размерность k выбирается такой, при которой значение критерия минимально.

Выводы

Для рандомизированного метода решения дискретной некорректной задачи приведены результаты аналитического исследования составляющих ошибки восстановления истинного сигнала. Показано, что с ростом числа строк матрицы проектора детерминированная составляющая ошибки убывает, а стохастическая растет.

Литература

1. Hansen P.C. Rank-deficient and discrete ill-posed problems / P.C. Hansen. – Numerical Aspects of Linear Inversion. – SIAM, Philadelphia. – 1998. – 247 p.
2. Tikhonov A.N. Solution of ill-posed problems / A.N. Tikhonov, V.Y. Arsenin. – V.H. Winston. – Washington, DC. – 1977.
3. Рачковский Д.А. Рандомизированный метод решения дискретных некорректных задач / Д.А. Рачковский, Е.Г. Ревунова // Кибернетика и системный анализ – 2012. – № 4 – С. 163-181.

4. Revunova E.G. Using randomized algorithms for solving discrete ill-posed problems / E.G. Revunova, D.A. Rachkovskij // International Journal "Information Theories and Applications" – 2009. – N 2. – P. 176-192.
5. Stewart G.W. On the perturbation of pseudo-inverses, projections and linear least squares problems / G. W. Stewart // SIAM Review. – 1977. – Vol. 19. – N. 4. – P. 634-662.