МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ

ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ СИСТЕМ
УПРАВЛЕНИЯ И РАДИОЭЛЕКТРОНИКИ (ТУСУР)

Ю.Е. Воскобойников
А.А. Мицель

МЕТОДЫ ПРИКЛАДНОЙ МАТЕМАТИКИ В ЕСТЕСТВЕННЫХ И ГУМАНИТАРНЫХ НАУКАХ

Часть 2. Практикум

Учебное пособие

ТОМСК 2019

УДК 519.2
ББК 22.172
В 650
Воскобойников Ю. Е., Мицель А.А.
Методы прикладной математики в естественных и гуманитарных науках. Часть 2. Практикум: учебное пособие/ Ю. Е. Воскобойников, А.А. Мицель/ Томский гос. ун-т систем управления и радиоэлектроники (ТУСУР). – Томск, 2018. – 32с.

В учебном пособии в первой части приводится системное изложение одного из разделов прикладной математики, связанного с устойчивыми методами и алгоритмами решения систем линейных алгебраических уравнений, возникающих при решении обратных задач. Основное внимание уделяется построению решений с минимальной ошибкой, а также учету имеющейся априорной информации об искомом решении. Во второй части приводится описание практических занятий по созданию алгоритмов построения нормального псевдорешения и регуляризированных решений систем линейных алгебраических уравнений.

Учебное пособие предназначено для магистрантов направления 01.04.02. направления «Прикладная математика и информатика» и аспирантов направления 09.06.01. «Информатика и вычислительная техника» (профиль «Математическое моделирование, численные методы и комплексы программ») Результаты будут полезны также широкому кругу студентов, магистрантов, аспирантов, исследователей, занимающихся решением задач параметрической идентификации и обработки экспериментальных данных.

ОГЛАВЛЕНИЕ

ПРАКТИЧЕСКОЕ ЗАНЯТИЕ №1
Построение нормального псевдорешения СЛАУ 4

§ 1.1. Постановка задачи ……..……..… 5
§ 1.2. SVD-алгоритм построения нормального псевдорешения……..…..……. 6
Задание 1.1…… 8
Задание 1.2……. 12

ПРАКТИЧЕСКОЕ ЗАНЯТИЕ №2
Построение регуляризованного решения СЛАУ . 14

§2.1. Байесовский регуляризирующий алгоритм……... 14
Задание 2.1……. 16
§2.2. Построение регуляризованного решения при неполной информации…… 22
Задание 2.2…… 24

ПРАКТИЧЕСКОЕ ЗАНЯТИЕ №3
Алгоритмы выбора параметра регуляризации . 24

§3.1. Выбор параметра регуляризации на основе критерия оптимальности……. 24
Задание 3.1…… 28
§3.2. Алгоритм выбора параметра регуляризации основе статистического принципа невязки…….. 28
Задание 3.2……... 30
§3.3. Алгоритм поиска α_{CV} методом перекрестной значимости …….. 30
Задание 3.3……. 31

ЛИТЕРАТУРА…….… 36

Практическое занятие №1

Построение нормального псевдорешения СЛАУ

§1.1. Постановка задачи

Дана система линейных алгебраических уравнений матричном виде

$$K\varphi = f\,, \tag{1.1}$$

где K – матрица размером $N \times M$ (N строк и M столбцов), φ – вектор размерности M (содержит M проекций), f – вектор размерности N

$$\tilde{f} = \bar{f} + \eta$$

Здесь $\bar{f}$ - вектор точной правой части, η - вектор ошибок.

Предположим, что матрица K имеет размеры $N \times M$. Вектор $\varphi_{НК}$ размерностью M называют *псевдорешением* (или решением МНК), если он доставляет минимум следующему функционалу

$$\Psi_{НК}(\varphi) = \left\| \; f - K\varphi \; \right\|^2 = (f - K\varphi)^T (f - K\varphi) \tag{1.2}$$

среди всех векторов евклидова пространства E^M.

Решение, обеспечивающее минимум функционалу (2), является решением следующей СЛАУ

$$K^T K\varphi_{НК} = K^T f\,, \tag{1.3}$$

которая называется системой нормальных уравнений.

В отличие от исходной системы $K\varphi = f$ эта система всегда разрешима, т.е. для любой правой части f существует

псевдорешение φ_{HK}. Если матрица K имеет ранг, равный M, то

$$\varphi_{HK} = (K^T K)^{-1} K^T f \ . \tag{1.4}$$

Сингулярным разложением прямоугольной $N \times M$ матрицы K (коротко: SVD-разложением) называется представление:

$$K = U \Lambda V^T \ , \tag{1.5}$$

где U – ортогональная ($N \times N$)-матрица, V – ортогональная ($M \times M$)-матрица, Λ – ($N \times M$)-матрица вида

$$\Lambda = \begin{vmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_M \\ 0 & 0 & 0 & \cdots & 0 \end{vmatrix}, \tag{1.6}$$

в которой последние $N - M$ строки содержат только нулевые элементы. Величины $\lambda_j \geq 0,\ j = 1,...,M$, называются *сингулярными числами* матрицы K, и в дальнейшим полагаем, что λ_j *упорядочены по убыванию*, т.е. $\lambda_j \geq \lambda_{j+1}$. Напомним, что матрица B называется *ортогональной*, если имеет место тождество $B^T B = BB^T = I$

§1.2. SVD-алгоритм построения нормального псевдорешения

Введем векторы

$$y = U^T f, \quad x = V^T \varphi \tag{1.7}$$

размерностью N и M соответственно. Тогда с учетом (1.5) систему $K\varphi = f$ можно преобразовать к эквивалентной системе:

$$\begin{aligned} &\lambda_j x_j = y_j, \quad && j = 1,...,M; \\ &0 = y_j, && j = M+1,..., \ N, \end{aligned} \tag{1.8}$$

которая хорошо характеризует «информативность» правой части: *чем меньше сингулярное число* λ_j, *тем с меньшим весом проекция* x_j *входит в правую часть.* Предельный случай $\lambda_j = 0$, $p+1 \le j \le M$, говорит о вырожденности *K*.

Очевидно, что невыполнение условия $\sum_{j=M+1}^{N} |y_j| = 0$ говорит о несовместности исходной системы.

С учетом ортогональности матриц *U*, *V* и соотношений (5) функционал (2) можно записать в виде

$$\Psi_{\acute{I}\,\hat{E}}(\varphi) = \Psi_{\acute{I}\,\hat{E}}(x) = \sum_{j=1}^{p}(y_i - \lambda_j x_j)^2 + \\ + \sum_{j=p+1}^{M}(y_j - 0 \cdot x_j)^2 + \sum_{j=M+1}^{N} y_j^2 .$$

Третье слагаемое обусловлено несовместностью исходной системы, и не зависит от x. Второе слагаемое отражает вырожденность системы, и, следуя определению нормального псевдорешения, проекции x_j, входящие во второе слагаемое, следует принять равным 0. Тогда минимум функционала достигается на векторе x^+ размерности p с элементами

$$x_j^+ = \frac{y_j}{\lambda_j}, \quad j = 1,..., p, \tag{1.9}$$

а нормальное псевдорешение φ^+ выражается как

$$\varphi^{+} = \sum_{j=1}^{p} x_j^{+} \cdot v_j \ . \qquad (1.9)$$

Напомним, что $\lambda_j > 0,\ j = 1,..., p$, где p – ранг матрицы K.

Задание 1.1. Построить нормальное псевдорешение с помощью пакета Mathcad

Рассмотрим две функции Mathcad, которые потребуются для построения нормального псевдорешения $\tilde{\varphi}^{+}$, определяемого выражением

$$\tilde{\varphi}^{+} = \sum_{j=1}^{p_{\Pi}} \frac{\left\langle \tilde{f}, u_j \right\rangle}{\lambda_j} \cdot v_j \ , \qquad (1.10)$$

где практический ранг p_{Π} матрицы K определяется количеством сингулярных чисел λ_j, удовлетворяющих условию:

$$\frac{\lambda_j}{\lambda_{\max}} \geq \gamma_0 \ , \qquad (1.11)$$

где γ_0 – достаточно малая величина ($10^{-10} \div 10^{-8}$).

Функция svds. Обращение имеет вид ***svds(K).*** Вычисляет вектор размерности M, состоящий из сингулярных чисел λ_j матрицы K, которые расположены в убывающем порядке.

Функция svd. Обращение имеет вид ***svd(K).*** Вычисляет матрицу UV размером $(N+M)\times M$. Первые N строк этой матрицы соответствуют матрице $\underline{U}$ размером $N \times M$, которая определяет первые M столбцов матрицы U, т.е.

$$U = \left| \underline{U} \vdots u_{M+1} \vdots \cdots \vdots u_N \right| . \qquad (1.12)$$

Последние M строк матрицы UV содержат матрицу V размером $M \times M$.

Заметим, что отсутствие в матрице $\underline{U}$ последних $N - M$ столбцов матрицы U обусловлено тем, что эти столбцы не участвуют в вычислении нормального псевдорешения и поэтому во многих программных реализациях SVD-разложения эти столбцы не вычисляются.

Функция submatrix. Обращение имеет вид ***submatrix (K, i1, i2, j1, j2).*** Формирует новую матрицу из элементов матрицы *K*, стоящих с *i*1 по *i*2 строках и с *j*1 по *j*2 столбцах матрицы *K*.

Пример 1. Дана матрица K размером 6×3. Необходимо вычислить сингулярные числа и матрицы $\underline{U}, V$.

Решение. На рис. 1.1 показан фрагмент документа Mathcad, выполняющий требуемые вычисления. Здесь же приведены вычисление числа обусловленности по формуле

$$cond(K) = \lambda_{\max} / \lambda_{\min}$$

и проверка ортогональности столбцов матриц $\underline{U}, V$. ☻

Перейдем к подпрограмме-функции (П-Ф) *Ps_Solve*, осуществляющей построение нормального псевдорешения СЛАУ по формуле (1.1). Обращение к П-Ф имеет вид:

$$Ps_Solve(K, f, \gamma_0). \qquad (1.13)$$

Формальные параметры: K – матрица системы размером $N \times M$, f – правая часть системы, γ_0 – переменная вещественного типа, входящая в условие (1.11).

На рис. 1.2 приведен фрагмент документа Mathcad с текстом П-Ф *Ps_Solve.*

Замечание 1. При обработке матриц в пакете Mathcad часто используется операция формирования вектора из определенного столбца матрицы. Для этого надо ввести имя матрицы, затем нажать клавиши [Ctrl+6] и в появившихся вверху угловых скобках задать нужный номер столбца. Например, в П-Ф *Ps Solve* стоят операции $U^{\langle j \rangle}, V^{\langle j \rangle}$. ♦

$$ORIGIN := 1$$

$$K := \begin{pmatrix} 1 & 2 & 3 \\ 5 & 7 & 9 \\ 6 & 8 & 9 \\ 4 & 6 & 8 \\ 1 & 3 & 5 \end{pmatrix} \quad N := 5 \quad M := 3 \quad \lambda := svds(K) \quad \frac{max(\lambda)}{min(\lambda)} = 90.55 \quad \lambda = \begin{pmatrix} 22.294 \\ 1.979 \\ 0.246 \end{pmatrix}$$

$$UV := svd(K)$$

$$U := submatrix(UV, 1, N, 1, M) \qquad U = \begin{pmatrix} -0.166 & -0.289 & 0.088 \\ -0.558 & 0.026 & -0.583 \\ -0.602 & 0.528 & 0.599 \\ -0.483 & -0.175 & -0.33 \\ -0.256 & -0.779 & 0.429 \end{pmatrix}$$

$$V := submatrix(UV, N+1, N+M, 1, M)$$

$$V = \begin{pmatrix} -0.393 & 0.772 & -0.5 \\ -0.571 & 0.222 & 0.791 \\ -0.721 & -0.596 & -0.353 \end{pmatrix} \quad U^T \cdot U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad V^T \cdot V = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Рис. 1.1. Сингулярное разложение матрицы *K*

$$
\begin{aligned}
Ps_Solve(K, f, \gamma_0) := \;& N \leftarrow rows(K) \\
& M \leftarrow cols(K) \\
& \lambda \leftarrow svds(K) \\
& UV \leftarrow svd(K) \\
& U \leftarrow submatrix(UV, 1, N, 1, M) \\
& V \leftarrow submatrix(UV, N+1, N+M, 1, M) \\
& \lambda_{max} \leftarrow max(\lambda) \\
& \phi_{ps} \leftarrow \sum_{j=1}^{M} if\left[\left(\frac{\lambda_j}{\lambda_{max}}\right) \geq \gamma_0, \left(\frac{f \cdot U^{\langle j \rangle}}{\lambda_j}\right) \cdot V^{\langle j \rangle}, 0 \cdot V^{\langle j \rangle}\right]
\end{aligned}
$$

Рис. 1.2. Текст подпрограммы-функции *Ps_Solve*

Пример 2. Матрица K размером 5×3 формируется с использованием П-Ф *Form_K* (фрагмент документа показан на рис. 1.3). Число обусловленности $1.426 \cdot 10^6$. Для заданного вектора φ вычислены два вектора: вектор «точной» правой части f и вектор «зашумленной» правой части f_η с относительной погрешностью $3.232 \cdot 10^{-3}$ (см. рис. 1.3). По этим двум векторам необходимо построить нормальные псевдорешения с использованием П-Ф *Ps_Solve*.

Здесь вычислены относительные ошибки двух псевдорешений: псевдорешение $\varphi 1_{ps}$, построенное по точной правой части, и псевдорешение $\varphi 2_{ps}$, построенное по искаженной правой части $f\eta$. Несмотря на маленькую погрешность исходных данных, относительная ошибка решения $\varphi 2_{ps}$ достигает большой величины $1.102 \cdot 10^3$, и эта ошибка удовлетворяет неравенству

$$\frac{\|\tilde{\varphi}-\bar{\varphi}\|}{\|\bar{\varphi}\|} \le cond(K)\ \frac{\|\tilde{f}-\bar{f}\|}{\|\bar{f}\|}.$$

Действительно

$1.103\cdot 10^{3} \le cond(K)\cdot 3.232\cdot 10^{-3} = 4.609\cdot 10^{3}$.

Задание 1.2.

Вычислить нормальное решение с помощью обратной матрицы по формуле (1.4) для точной и зашумленной правой части и сравнить с решением, полученным SVD – алгоритмом.

Решение. Два обращения к *Ps_Solve* показаны на рис. 1.3.

$ORIGIN := 1$

$$Form_K(N, M, \sigma) := \left| \begin{array}{l} for\ \ i \in 1..N \\ \quad for\ \ j \in 1..M \\ \qquad K_{i,j} \leftarrow exp\left[-\dfrac{\left(j-\dfrac{M}{N}\cdot i\right)^2}{\sigma^2}\right] \\ K \end{array} \right.$$

$$K := Form_K(5, 3, 30) \qquad Cond(K) := \left| \begin{array}{l} k \leftarrow \dfrac{max(svds(K))}{min(svds(K))} \\ k \end{array} \right.$$

$$Cond(K) = 1.426\times 10^{6}$$

$$\phi := \begin{pmatrix} 1 \\ 3 \\ 6 \end{pmatrix} \qquad f := K\cdot\phi \qquad f = \begin{pmatrix} 9.955 \\ 9.976 \\ 9.99 \\ 9.995 \\ 9.992 \end{pmatrix} \qquad f_\eta := \begin{pmatrix} 10.01 \\ 9.96 \\ 10.03 \\ 9.98 \\ 10.00 \end{pmatrix}$$

$$\frac{|f-f_\eta|}{|f|} = 3.232\times 10^{-3}$$

$$\phi 1_{ps} := Ps_Solve\left(K, f, 10^{-10}\right) \qquad \phi 2_{ps} := Ps_Solve\left(K, f_\eta, 10^{-10}\right)$$

$$\frac{|\phi-\phi 1_{ps}|}{|\phi|} = 7.045\times 10^{-11} \qquad \frac{|\phi-\phi 2_{ps}|}{|\phi|} = 1.102\times 10^{3}$$

Рис 1.3. Построение нормальных псевдорешений

Примечание. При наличии научной задачи магистранта, которую можно свести к системе линейных уравнений, построить нормальное псевдорешение для этой задачи.

Практическое занятие №2

Построение регуляризованного решения СЛАУ

§2.1. Байесовский регуляризирующий алгоритм

Исходной информацией являются:

- априорное распределение $p(\varphi)$ искомого вектора решения φ является нормальным, т.е. $p(\varphi) = const \cdot \exp\left(-\frac{1}{2}\left\|\varphi - m_\varphi\right\|^2_{V_\varphi^{-1}}\right)$, где m_φ и V_φ - математическое ожидание и матрица ковариаций;
- условное распределение $p\left(\tilde{f}\middle|\varphi\right)$, характеризующее распределение вектора измерений $\tilde{f}$ при фиксированном векторе φ является нормальным, $p\left(\tilde{f}\middle|\varphi\right) = const \cdot \exp\left(-\frac{1}{2}\left\|\tilde{f} - K\varphi\right\|^2_{V_\eta^{-1}}\right)$.

Здесь запись $\left\|z - m_z\right\|^2_{V_z^{-1}}$ означает: $\left\|z - m_z\right\|^2_{V_z^{-1}} = \left(z - m_z\right)^T V_z^{-1}\left(z - m_z\right)$.

Байесовское регуляризированное решение находится из системы линейных алгебраических уравнений

$$\left(K^T V_\eta^{-1} K + V_\varphi^{-1}\right)\varphi_{\acute{A}} = K^T V_\eta^{-1}\tilde{f} + V_\varphi^{-1} m_\varphi . \qquad (2.1)$$

Заметим, что вектор $\varphi_Б$ для указанных распределений и квадратичной функции потерь $\ddot{I}\ \left(\varphi_T, \varphi\right) = \left\|\varphi_T - \varphi\right\|^2$ максимизирует значения апостериорной плотности распределения $p(\varphi \mid \tilde{f})$, которая имеет вид

$$p(\varphi \mid \tilde{f}) = const \cdot \exp\left\{ -\frac{1}{2}\left\| \tilde{f} - K\varphi \right\|^2_{V_\eta^{-1}} - \frac{1}{2}\left\| \varphi - m_{\bar{\varphi}} \right\|^2_{V_{\bar{\varphi}}^{-1}} \right\},$$

а, следовательно, доставляет минимум функционалу

$$F_{Á}\left[\varphi\right] = \left\| \tilde{f} - K\varphi \right\|^2_{V_\eta^{-1}} + \left\| \varphi - m_\varphi \right\|^2_{V_\varphi^{-1}}. \qquad (2.2)$$

Рассмотрим точность построенного байесовского регуляризированного решения при следующих предположениях:

$$M\left[\eta\right] = 0; \;\; M\left[\varphi\eta^T\right] = 0. \qquad (2.3)$$

Последнее условие означает, что проекция φ_j и η_i не коррелированны между собой. Определим вектор ошибки решения как

$$\varepsilon_Б = \varphi_Б - \bar{\varphi}. \qquad (2.4)$$

Ошибка может быть представлена в следующем виде:

$$\begin{aligned} \varepsilon_{Á} &= \left(K^T V_\eta^{-1} K + V_\varphi^{-1}\right)^{-1} \left(K^T V_\eta^{-1}\left(K\bar{\varphi} + \eta\right) + V_\varphi^{-1} m_\varphi\right) - \\ &- \bar{\varphi} = \left(K^T V_\eta^{-1} K + V_\varphi^{-1}\right)^{-1} \cdot \left[K^T V_\eta^{-1}\eta + V_\varphi^{-1}(m_\varphi - \bar{\varphi})\right]. \end{aligned} \qquad (2.5)$$

Математическое ожидание $M\left[\varepsilon_Б\right]$ равно:

$$M\left[\varepsilon_{Á}\right] = b_{Á} = \left(K^T V_\eta^{-1} K + V_\varphi^{-1}\right)^{-1} V_\varphi^{-1}\left(m_\varphi - M_\varphi\left[\varphi\right]\right). \qquad (2.6)$$

Рассмотрим случайную ошибку $\xi_Б = \varepsilon_Б - M\left[\varepsilon_Б\right]$, которая представима в виде:

$$\xi_{Á} = \left(K^T V_\eta^{-1} K + V_\varphi^{-1}\right)^{-1} \cdot K^T V_\eta^{-1}\eta, \qquad (2.7)$$

с математическим ожиданием $M\left[\xi_{Á}\right] = 0$ и матрицей ковариации

$$V_{\xi_{\acute{A}}} = \left(K^T V_\eta^{-1} K + V_\varphi^{-1}\right)^{-1} K^T V_\eta^{-1} K \left(K^T V_\eta^{-1} K + V_\varphi^{-1}\right)^{-1}. \qquad (2.8)$$

Вектор «полной» ошибки байесовского решения можно записать в виде:

$$\Delta_{\acute{A}_j} = b_{\acute{A}_j} + \sqrt{V_{\xi_{jj}}}, \ \ j = 1,...,M \ . \qquad (2.9)$$

Задание 2.1.

1. Задать матрицу системы K, вектор решения φ, среднее значение m_φ и матрицу ковариации V_φ.

2. Для заданной матрицы K и вектора решения φ вычислить правую часть $\overline{f}$.

3. Внести шум в правую часть, т.е. вычислить вектор $f = \overline{f} + \eta$, где η - вектор ошибки, вычисляемый с помощью нормального датчика случайных чисел для двух случаев:

а) ошибки не коррелированны. Значения СКО задать равными $\sigma_{\eta_i} = \delta_i \cdot f_i, \ i = 1,...,N$; $\delta_i = 0.01, 0.05, 0.01$. Матрица ковариации в этом случае будет диагональной $V_\eta = diag\left(\sigma_{\eta_1}^2 ..., \sigma_{\eta_N}^2\right)$.

б) ошибки коррелированны. В этом случае вектор ошибок вычисляется по формуле

$$\eta = V_\eta^{1/2} \xi,$$

где V_η - заданная матрица ковариации; ξ - вектор, проекции которого состоят из нормальных случайных величин с единичной дисперсией и нулевым средним.

4. Вычислить решение по формуле (2.1) и вектор ошибки (2.9).

Примечание. При наличии научной задачи магистранта, которую можно свести к системе линейных уравнений, провести исследования этой задачи.

§2.2. Построение регуляризованного решения при неполной информации

В ситуациях, когда отсутствует априорная информация о числовых характеристиках решения и шума, сглаживающий функционал имеет вид

$$F_\alpha[\varphi] = \left\| \tilde{f} - K\varphi \right\|_{W_f}^2 + \alpha \left\| \varphi - \omega_\varphi \right\|_{W_\varphi}^2 \qquad (2.10)$$

В этом случае вектор φ_α, доставляющий минимум функционалу (2.10), является решением системы

$$\left(K^T W_f K + \alpha W_\varphi \right) \varphi_\alpha = K^T W_f \tilde{f} + \alpha W_\varphi \omega_\varphi , \qquad (2.10)$$

состоящей из M уравнений относительно M неизвестных. Здесь $\alpha > 0$ параметр регуляризации, ω_φ - пробное решение. Параметр регуляризации является неизвестной величиной.

Метод рандомизации. Допустим, что априори известно о принадлежности искомого решения $\bar{\varphi}$ гиперпрямоугольнику, определяемому неравенствами

$$\varphi_{\min_j} \le \bar{\varphi}_j \le \varphi_{\max_j} , \; j = 1, \ldots, M , \qquad (2.11)$$

являющимися детерминированными ограничениями. Метод рандомизации позволяет интерпретировать детерминированные ограничения в терминах числовых характеристик некоторых вероятностных распределений (чаще всего нормального распределения). Первые два момента $\hat{m}_\varphi$, $\hat{V}_\varphi$ нормального распределения $N\left(\hat{m}_\varphi, \hat{V}_\varphi\right)$ определяется таким образом, чтобы случайный вектор, подчиняющийся этому распределению с

вероятностью β попадал в гиперпрямоугольник (2.11). Математическое ожидание $\widehat{m}_{\varphi}$ такого вектора определяется как

$$\widehat{m}_{\varphi} = \left| \frac{\varphi_{\min_1} + \varphi_{\max_1}}{2}, \frac{\varphi_{\min_2} + \varphi_{\max_2}}{2}, ..., \frac{\varphi_{\min_M} + \varphi_{\max_M}}{2} \right|^T, \quad (2.12)$$

а корреляционная матрица $\widehat{V}_{\varphi}$ является диагональной и вычисляется по формуле

$$\widehat{V}_{\varphi} = diag \left\{ \frac{\left(\varphi_{\max_1} - \varphi_{\min_1}\right)^2}{4\mu_{\beta}}, ..., \frac{\left(\varphi_{\max_M} - \varphi_{\min_M}\right)^2}{4\mu_{\beta}} \right\}. \quad (2.13)$$

Значение $\mu_{\beta} = 3$.

Если матрица V_{η} задана, то, подставляя вычислительные описанным образом $\widehat{m}_{\varphi}$, $\widehat{V}_{\varphi}$ в систему уравнений (2.29), получаем матричную запись алгоритма нахождения «рандомизированного» регуляризированного решения φ_p:

$$\left(K^T V_{\eta}^{-1} K + \widehat{V}_{\varphi}^{-1}\right)\varphi_p = K^T V_{\eta}^{-1} \tilde{f} + \widehat{V}_{\varphi}^{-1} \widehat{m}_{\varphi}. \quad (2.14)$$

Если информация о шуме измерения задана в виде системы неравенств

$$\left| \tilde{f}_i - f_i \right| \le \delta_i, \; 1 \le i \le N, \quad (2.15)$$

то вновь обращаемся к методу рандомизации и вычисляем корреляционную матрицу $\widehat{V}_{\eta}$ по формуле

$$\widehat{V}_{\eta} = diag \left\{ \frac{\delta_1^2}{12}, \frac{\delta_2^2}{12}, ..., \frac{\delta_N^2}{12} \right\} \quad (2.16)$$

и используем ее в алгоритме (2.14).

Очевидно, что (2.14) является частным случаем системы (2.10) при следующих заменах:

$$W_f = V_\eta^{-1};\ \ W_\varphi = \widehat{V}_\varphi^{-1};\ \ \omega_\varphi = \widehat{m}_\varphi;\ \ \alpha = 1. \qquad (2.17)$$

Стабилизирующий функционал

Введем квадратичную форму

$$\varphi^T W_\varphi \varphi = \|\varphi\|^2_{W_\varphi}, \qquad (2.18)$$

которую назовем стабилизирующим функционалом. Неотрицательно определенная матрица W_φ находится из условия: чем «глаже» вектор φ, тем меньшее значение принимает функционал (34). Исходя из этого условия, часто матрицу V_φ формируют как

$$W_\varphi = D_p^T D_p, \qquad (2.19)$$

где D_p – матрица, являющаяся дискретным аналогом оператора дифференцирования p-го порядка (и тогда говорят о регуляризации p-го порядка). Так, при $p = 0$ матрица W_φ является единичной размером $M \times M$.

Для $p = 1$ матрица W_φ имеет вид:

$$W_\varphi = \begin{vmatrix} 1 & -1 & & 0 \\ -1 & 2 & -1 & \\ & -1 & 2 & -1 \\ 0 & & -1 & 1 \end{vmatrix} \qquad (2.20)$$

Если решение ищется на множестве векторов с ограниченной нормой, для которых отсутствует взаимосвязь между «соседними» проекциями, то целесообразно использовать единичную матрицу либо диагональную матрицу вида (2.13), в

этом случае $W_\varphi = \widehat{V}_\varphi^{-1}$ и это будет соответствовать регуляризации нулевого порядка.

Вернемся к вектору ω_φ, входящему в функционал (2.9) и названному «пробным» решением. При наличии априорной информации вида (2.11) его можно задать как $\omega_\varphi = \widehat{m}_\varphi$, где $\widehat{m}_\varphi$ определяется выражением (2.12). При отсутствии такой информации традиционным заданием является $\omega_\varphi = 0_M$.

Матрицу W_f рекомендуется задавать с точностью до константы равной обратной матрицы V_η^{-1}, т.е.

$$W_f = C_f V_\eta^{-1}, \tag{2.21}$$

где константа $C_f > 0$. При наличии информации вида (31) матрицу V_η можно определить соотношением (2.16). При отсутствии информации о числовых характеристик погрешностей η_i матрицу W_f можно задать диагональной. Ненулевые элементы такой матрицы интерпретируются как *весовые множители, определяющие значимость (или информативность) соответствующих проекций вектора правой части* $\tilde{f}$.

В предельном случае (соответствующем отсутствию информации об искомом решении и шуме измерения) матрицы W_f и W_φ задаются единичными, т.е.

$$W_f = I_{N\times N}; \;\; W_\varphi = I_{M\times M}, \tag{2.22}$$

Ошибка решения φ_α.

Определим ошибку решения φ_α, определяемым вектором (2.10)

$$\varepsilon_\alpha = \varphi_\alpha - \overline{\varphi}^+ ,$$

где $\overline{\varphi}^+$ – нормальное псевдорешение системы $K\varphi = \overline{f}$ при точной правой части $\overline{f}$, т.е. $\overline{\varphi}^+ = (K^T W_f K)^{-1} K^T W_f \overline{f}$. Как и прежде, вектор ε_α представим суммой векторов случайной ξ_α и систематической b_α ошибок:

$$\varepsilon_\alpha = \varphi_\alpha - \overline{\varphi}_\alpha^+ + \overline{\varphi}_\alpha^+ - \overline{\varphi}^+ = \xi_\alpha + b_\alpha . \qquad (2.23)$$

Вектор $b_\alpha = M_\eta \left[\varepsilon_\alpha \right]$ можно назвать смещением решения φ_α. Систематическая ошибка b_α имеет вид

$$b_\alpha = \alpha \left(K^T W_f K + \alpha W_\varphi \right)^{-1} W_\varphi \left(\omega_\varphi - \overline{\varphi}^+ \right) . \qquad (2.24)$$

Вектор

$$\xi_\alpha = \varepsilon_\alpha - M_\eta \left[\varepsilon_\alpha \right] = \varepsilon_\alpha - b_\alpha \qquad (2.25)$$

является случайным вектором с нулевым средним и определяется выражением

$$\xi_\alpha = \left(K^T W_f K + \alpha W_\varphi \right)^{-1} K^T W_f \eta . \qquad (2.26)$$

Ковариационная матрица $V_{\xi_\alpha} = M\left[\xi_\alpha \xi_\alpha^T \right]$ этого вектора определяется выражением:

$$V_{\xi\alpha} = (K^T W_f K + \alpha W_\varphi)^{-1} K^T W_f V_\eta W_f K (K^T W_f K + \alpha W_\varphi)^{-1},$$

Полагая $W_f = V_\eta^{-1}$, получим

$$V_{\xi\alpha} = (K^T W_f K + \alpha W_\varphi)^{-1} K^T W_f K (K^T W_f K + \alpha W_\varphi)^{-1}, \qquad (2.42)$$

Полная ошибка решения равна

$$\Delta_{\alpha_j} = b_{\alpha_j} + \sqrt{V_{\xi\alpha_{jj}}}, \ \ j = 1,...,M \ . \qquad (2.27)$$

Задание 2.2.

1. Задать матрицу системы K, вектор решения φ, ограничения на вектор φ вида (2.27).

2. Вычислить среднее значение m_φ и матрицу ковариации V_φ по формулам (2.28), (2.29).

3. Для заданной матрицы K и вектора решения φ вычислить правую часть $\overline{f}$.

4. Внести шум в правую часть, т.е. вычислить вектор

$$f = \overline{f} + \eta,$$

где η - вектор ошибки, вычисляемый с помощью нормального датчика случайных чисел для двух случаев:

а) ошибки не коррелированны. Значения СКО задать равными $\sigma_{\eta_i} = \delta_i \cdot f_i, \ i = 1,...,N$; $\delta_i = 0.01, 0.05, 0.01$. Матрица ковариации в этом случае будет диагональной $V_\eta = diag\left(\sigma^2_{\eta_1}...,\sigma^2_{\eta_N}\right)$.

б) ошибки коррелированны. В этом случае вектор ошибок вычисляется по формуле $\eta = V_\eta^{1/2}\xi$, где V_η - заданная матрица ковариации; ξ - вектор, проекции которого состоят из нормальных случайных величин с единичной дисперсией и нулевым средним.

5. Задать ограничения на вектор f вида (2.31) и вычислить матрицу V_η по формуле (2.32).

6. Вычислить решение системы (2.26) и погрешность по формуле (2.43) для следующих вариантов:

а) $W_f = I,\ W_\varphi = I,\ \omega_\varphi = 0$, для различных значений $\alpha \in [10^{-8}, 1]$. Значения α определять по формуле $\alpha_k = \alpha_{k-1} \cdot 10$;

б) $W_f = V_\eta^{-1},\ W_\varphi = V_\varphi^{-1},\ \omega_\varphi = m_\varphi,\ \alpha = 1$.

Примечание. При наличии научной задачи магистранта, которую можно свести к системе линейных уравнений, провести исследования этой задачи.

Практическое занятие №3

Алгоритмы выбора параметра регуляризации

§3.1. Выбор параметра регуляризации на основе критерия оптимальности

Рассмотрим систему уравнений

$$K\varphi = f \ . \tag{3.1}$$

Здесь K - матрица системы размерности $N \times M$, f - вектор правой части размерности N, равный $f = \overline{f} + \eta$, где $\overline{f}$ - точная правая часть, η - погрешность правой части, которая описывается моментами 1-го и 2-го порядка $M[\eta] = 0,\ M[\eta\eta^T] = V_\eta$. Матрица V_η допускает представление $V_\eta = \sigma_\eta^2 \cdot C_\eta$, где σ_η^2 - скалярная величина, C_η - матрица размерности $N \times N$.

Регуляризованное решение системы (3.1) φ_α является решением следующей системы уравнений

$$\left(K^T V_\eta^{-1} K + \alpha\sigma_\eta^2 W_\varphi\right)\varphi_\alpha = K^T V_\eta^{-1} \tilde{f} \ , \tag{3.2}$$

где $\alpha > 0$ - параметр регуляризации.

Так как σ_η^2 является константой, то в дальнейшем будем рассматривать систему вида

$$\left(K^T V_\eta^{-1} K + \alpha W_\varphi\right)\varphi_\alpha = K^T V_\eta^{-1} \tilde{f} \ . \tag{3.3}$$

В качестве матрицы W_φ используется либо единичная матрица (регуляризация нулевого порядка), либо матрица вида

$$W_\varphi = \begin{vmatrix} 1 & -1 & & 0 \\ -1 & 2 & -1 & \\ & -1 & 2 & -1 \\ 0 & & -1 & 1 \end{vmatrix}$$

(регуляризация первого порядка).

Рассмотрим вектор невязки $e_\alpha = f - K\varphi_\alpha = E(\alpha)\tilde{f}$, где $E(\alpha)$ - оператор невязки, который для решения, определяемого из системы (3), равен

$$E(\alpha) = I - K\left(K^T V_\eta^{-1} K + \alpha W_\varphi\right)^{-1} K^T V_\eta^{-1} \text{ или}$$

$$E(\alpha) = V_\eta \left(V_\eta + \alpha^{-1} K W_\varphi^{-1} K^T\right)^{-1}. \qquad (3.4)$$

Тогда для матрицы ковариации вектора невязки $V_e(\alpha) = M[e_\alpha e_\alpha^T]$ мы можем записать:

$$V_e(\alpha) = E(\alpha) V_{\tilde{f}} E^T(\alpha). \qquad (3.5)$$

где $V_{\tilde{f}}$ - матрица ковариации вектора правой части

$$V_f = M[f \cdot f^T].$$

В качестве α_{opt} возьмем такое значение α_W, при котором принимается основная статистическая гипотеза:

$$H_0: \; V_e(\alpha) = V_\eta E^T(\alpha). \qquad (3.6)$$

Таким образом, значение α_W можно рассматривать как оценку оптимального параметра регуляризации α_{opt}.

Для проверки гипотезы (3.6) введем статистику

$$\rho_W(\alpha) = e_\alpha^T \left[V_\eta E^T(\alpha)\right]^{-1} e_\alpha, \qquad (3.7)$$

Подставляя (3.4) в (3.7), получим

$\rho_W(\alpha) = \tilde{f}^T V_\eta^{-1} e_\alpha$ или

$$\rho_W(\alpha) = \tilde{f}^T \left(V_\eta + \alpha^{-1} K W_\varphi^{-1} K^T\right)^{-1} \tilde{f} \ . \qquad (3.8)$$

Введем параметр $\gamma = 1/\alpha$. Тогда выражение (3.9) примет вид

$$\rho_W(\gamma) = \tilde{f}^T \left(V_\eta + \gamma K W_\varphi^{-1} K^T\right)^{-1} \tilde{f} \ . \qquad (3.9)$$

$$M\left[\rho_W(\gamma)\right] = Sp\left[\left(V_\eta E^T(\gamma)\right)^{-1}\left(V_\eta E^T(\gamma)\right)\right] = Sp\left[I_N\right] = N \ ,$$

где I_N – единичная матрица размера $N \times N$.

Статистика $\rho_W(\gamma)$ пари $\gamma = \gamma_W$ подчиняется χ^2-распределению с N степенями свободы. Тогда проверка гипотезы (3.6) сводится к проверке предположения: *подчиняется ли величина* $\rho_W(\gamma)$ χ^2*-распределению с* N *степенями свободы.* Для этого построим интервал

$$\Theta_N(\beta) = \left[\vartheta_N(\beta/2), \vartheta_N(1-\beta/2)\right], \qquad (3.10)$$

где $\vartheta_N(\beta/2)$ – квантиль χ^2-распределения уровня $\beta/2$. Если $\rho_W(\alpha)$ попадает в интервал (3.10), т.е. выполняется неравенство

$$\vartheta_N(\beta/2) \le \rho_W(\gamma) \le \vartheta_N(1-\beta/2), \qquad (3.11)$$

то гипотеза (3.6) может быть принята с вероятностью ошибки первого рода, равной β. Следовательно, значение α_W, при котором выполняется (3.11), является оценкой для α_{opt}.

Для $N \le 10$ граничные точки интервала $\Theta_N(\beta)$ – квантили $\vartheta_N(\beta/2)$, $\vartheta_N(1-\beta/2)$ при $\beta = 0.1$ приведены в таблице 3.1.

Таблица 3.1

N	4	5	6	7	8	9	10

$\vartheta_N(0.05)$	0.71	1.14	1.64	2.17	2.73	3.32	3.94
$\vartheta_N(0.95)$	9.49	1.0	2.6	4.1	5.5	6.9	8.3

Если $N > 10$, то квантили достаточно точно могут аппроксимироваться следующими выражениями:

$$\vartheta_N(0.05) \approx N - 1.64\sqrt{2N}\,;\ \vartheta_N(0.95) \approx N + 1.64\sqrt{2N}\,. \quad (3.12)$$

Для вычисления γ_W используем итерационную процедуру ньютоновского типа:

$$\gamma^{(n)} = \gamma^{(n-1)} - \frac{\rho_W\left(\gamma^{(n-1)}\right) - N}{\rho'_W\left(\gamma^{(n-1)}\right)},\ n = 1, 2, \ldots, \quad (3.13)$$

с начальным значением $\gamma^{(0)} \square 10^{-15}$. В качестве γ_W принимается значение $\gamma^{(n)}$, удовлетворяющее (3.12). Заметим, что эта процедура решает нелинейное уравнение $\rho_W(\gamma) = N$, но момент останова определяется условием (3.12).

Замечание. Для вычисления статистики (3.9) необходимо вычислять обратную матрицу $\left(V_\eta + \gamma K W_\varphi^{-1} K^T\right)^{-1}$, что приводит к увеличению трудоемкости вычислений. Чтобы этого избежать, запишем статистику $\rho_W(\gamma)$ в форме $\rho_W(\gamma) = f^T \cdot q_\gamma$, где вектор q_γ является решением системы

$$\left(V_\eta + \gamma K W_\varphi^{-1} K^T\right) \cdot q = f\,. \quad (3.14)$$

В формуле (3.14) используется производная $\rho'_W(\gamma)$, равная $\rho'_W(\gamma) = f^T q'_\gamma$, где вектор q'_γ является решением системы уравнений

$$\left(V_\eta + \gamma K W_\varphi^{-1} K^T\right) \cdot \left(K W_\varphi^{-1} K^T\right)^{-1} q' = -q_\gamma \,. \qquad (3.15).$$

Задание 3.1.

Для построенного регуляризованного решения в лабораторной работе №2 при неполной информации (см. задание 2.3 в работе №2) вычислить:

1) параметр регуляризации α_W;

2) регуляризованное решение φ_{α_W};

3) ошибку решения $\Delta_{\alpha_j} = b_{\alpha_j} + \sqrt{V_{\xi\alpha_{jj}}}$, $j = 1,...,M$

Примечание. При наличии научной задачи магистранта, которую можно свести к системе линейных уравнений, провести исследования этой задачи

§3.2. Алгоритм выбора параметра регуляризации основе статистического принципа невязки

Для вычисления значения α_V проверяется статическая гипотеза

$$H_0 : V_e(\alpha) = V_\eta \qquad (3.27)$$

Для проверки (3.27) введем статистику $\rho_V(\alpha) = e_\alpha^T V_\eta^{-1} e_\alpha$.

Здесь $e_\alpha = f - K\varphi_\alpha$ - вектор невязки. Учитывая, что $\varphi_\alpha = \left(K^T V_\eta^{-1} K + \alpha W_\varphi\right)^{-1} K^T V_\eta^{-1} \tilde{f}$, получим:

$$e_\alpha = f - K\varphi_\alpha = \tilde{f} - K\left(K^T V_\eta^{-1} K + \alpha W_\varphi\right)^{-1} K^T V_\eta^{-1} \tilde{f} = \\ = V_\eta \left(V_\eta + \alpha^{-1} K W_\varphi^{-1} K^T\right)^{-1} f \quad (3.28)$$

Тогда выражение для статистики примет вид:

$$\rho_V(\alpha) = e_\alpha^T V_\eta^{-1} e_\alpha = \\ = f^T \left(V_\eta + \alpha^{-1} K W_\varphi^{-1} K^T\right)^{-1} V_\eta \left(V_\eta + \alpha^{-1} K W_\varphi^{-1} K^T\right)^{-1} f$$

Введем $\gamma = 1/\alpha$, получим

$$\rho_V(\gamma) = f^T \left(V_\eta + \gamma K W_\varphi^{-1} K^T\right)^{-1} V_\eta \left(V_\eta + \gamma K W_\varphi^{-1} K^T\right)^{-1} f \,. \quad (3.29)$$

Вычислим производную по γ. Для этого представим $\rho_V(\gamma)$ в виде

$$\rho_V(\gamma) = g^T V_\eta g \,, \quad (3.30)$$

где

$$g = \left(V_\eta + \gamma K W_\varphi^{-1} K^T\right)^{-1} f \,. \quad (3.31)$$

Тогда производная от $\rho_V(\gamma)$ равна

$$(\rho_V(\gamma))' = 2 g^T V_\eta g' \,, \quad (3.32)$$

где

$$g' = -\left(K W_\varphi^{-1} K^T\right)\left(V_\eta + \gamma K W_\varphi^{-1} K^T\right)^{-2} f = \\ = -\left(K W_\varphi^{-1} K^T\right)\left(V_\eta + \gamma K W_\varphi^{-1} K^T\right)^{-1} g \,. \quad (3.33)$$

Вычисление g' по формуле (3.33) эквивалентно решению системы уравнений

$$\left(V_\eta + \gamma K W_\varphi^{-1} K^T\right)\left(K W_\varphi^{-1} K^T\right)^{-1} g' = -g \,. \quad (3.34)$$

Для вычисления γ_V используем итерационную процедуру ньютоновского типа:

$$\gamma^{(n)} = \gamma^{(n-1)} - \frac{\rho_V\left(\gamma^{(n-1)}\right) - N}{\rho_V'\left(\gamma^{(n-1)}\right)}, \; n = 1, 2, \ldots, \tag{3.35}$$

с начальным условием $\gamma^{(0)} \square 10^{-15}$.

Итерационный процесс завершаем при выполнении условия

$$\vartheta_N\left(\beta/2\right) \le \rho_V\left(\gamma\right) \le \vartheta_N\left(1-\beta/2\right). \tag{3.36}$$

где $\vartheta_N\left(\beta/2\right)$, $\vartheta_N\left(1-\beta/2\right)$ – квантили χ^2-распределения с N степенями свободы уровней $\beta/2$, $1-\beta/2$ соответственно.

Задание 3.2.

Для построенного регуляризованного решения в лабораторной работе №2 при неполной информации (см. задание 4 в работе №2) вычислить:

1) параметр регуляризации α_V;

2) регуляризованное решение φ_{α_V};

3) ошибку решения $\Delta_{\alpha_j} = b_{\alpha_j} + \sqrt{V_{\xi\alpha_{jj}}}$, $j = 1, \ldots, M$.

§3.3. Алгоритм поиска α_{CV} методом перекрестной значимости

Предположим, что корреляционная матрица шума измерений V_η неизвестна. В этом случае регуляризированное решение φ_α определяется из системы уравнений

$$\left(K^T K + \alpha W_\varphi\right)\varphi_\alpha = K^T \tilde{f}. \tag{3.36}$$

Выбор параметра методом перекрестной значимости (cross-validation method CV) осуществляется из условия минимума функционала

$$U(\alpha)=\frac{\frac{1}{N}\sum_{i\in I_u}\left(\tilde{f}_i-\left\{K\varphi_\alpha^{(i)}\right\}_i\right)^2}{\left[\frac{1}{N}Sp\left(I-A(\alpha)\right)\right]_2}, \tag{3.39}$$

где $\varphi_\alpha^{(i)}$ – решение, построенное по вектору $\tilde{f}^{(i)}$, полученному из вектора $\tilde{f}$ путем удаления проекции $\tilde{f}_i$; I_U – множество индексов, состоящее из значений $\{1,2,...,N\}$; $A(\alpha)=K\left(K^TK+\alpha I\right)^{-1}K^T$. Величина α_U, доставляющая минимум $U(\alpha)$, применяется в качестве параметра регуляризации, выбранного по методу перекрестной значимости.

Задание 3.3. Для построенного регуляризованного решения (3.38) вычислить провести следующие вычисления.

Шаг 1.

a) Задать начальное значение параметра регуляризации $\alpha_{CV}=10^{-10}$.

b) «Проредить компоненты вектора правой части f, удаляя все четные компоненты, начиная со второй, т.е. $f_2, f_4,..., f_{n-1}$ (здесь n – нечетное). По полученному вектору $\tilde{f}^{(1)}$ с нечетными проекциями $f_1, f_3,..., f_n$ при заданном значении параметра α_{CV} построить решение (3.38) $\varphi_{\alpha_{CV}}^{(1)}$ и затем вычислить невязку $\frac{1}{N}\sum_{i\in I_u}\left(\tilde{f}_i-\left\{K\varphi_{\alpha_{CV}}^{(1)}\right\}_i\right)^2$ в точках, которые были удалены

при построении регуляризованного решения, т.е. использовать проекции $f_2, f_4, \ldots, f_{n-1}$.

c) Изменить значение параметра α_{CV} на порядок и выполнить шаг 2.

Вычисления закончить при достижении минимума (3.39).

Шаг 2. С найденным значением параметра α_{CV} построить регуляризованное решение (3.38) $\varphi_{\alpha_{CV}}$.

Шаг 3. Вычислить ошибку решения $\Delta_{\alpha_j} = b_{\alpha_j} + \sqrt{V_{\xi\alpha_{jj}}}$, $j = 1, \ldots, M$.

Примечание. При наличии научной задачи магистранта, которую можно свести к системе линейных уравнений, провести исследования этой задачи

Литература

1. Воскобойников Ю.Е., Мицель А.А. Методы прикладной математики в естественных и гуманитарных науках. Часть 1. Лекционный курс: учебное пособие/ Томский гос. ун-т систем управления и радиоэлектроники (ТУСУР). – Томск: 2018. – 100 с.