МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ

ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ СИСТЕМ
УПРАВЛЕНИЯ И РАДИОЭЛЕКТРОНИКИ (ТУСУР)

Ю.Е. Воскобойников
А.А. Мицель

СОВРЕМЕННЫЕ ПРОБЛЕМЫ ПРИКЛАДНОЙ МАТЕМАТИКИ

Часть 2. Практикум

Учебное пособие

ТОМСК 2016

УДК 519.2
ББК 22.172
В 650

Воскобойников Ю. Е., Мицель А.А.

Современные проблемы прикладной математики. Часть 2. Практикум: учебное пособие/ Ю. Е. Воскобойников, А.А. Мицель/ Томский гос. ун-т систем управления и радиоэлектроники (ТУСУР). – Томск, 2016. – 52с.

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 параметрической идентификации моделей. Основное внимание уделяется построению решений с минимальной ошибкой или с требуемыми точностными характеристиками, 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 «Прикладная математика и информатика». 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

§1.1. Постановка задачи

Дана система линейных алгебраических уравнений матричном виде

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

где K – матрица размером $N \times M$ (N строк и M столбцов), φ – вектор размерности M (содержит M проекций), $\tilde{f}$ – 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 φ_{HK} размерностью M называют *псевдорешением* (или решением МНК), если он доставляет минимум следующему функционалу

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

среди всех векторов евклидова пространства E^M.

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

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

которая называется системой нормальных уравнений.

В отличие от исходной системы $K\varphi = \tilde{f}$ эта система всегда разрешима, т.е. для любой правой части $\tilde{f}$ 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 $\lambda_j \ge 0,\ j = 1,\dots,M$, называются *сингулярными числами* матрицы K, и в дальнейшим полагаем, что λ_j *упорядочены по убыванию*, т.е. $\lambda_j \ge \lambda_{j+1}$. 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

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

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

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

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 $\sum_{j=M+1}^{N} \left| y_j \right| = 0$ говорит о несовместности исходной системы.

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

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

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

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

а нормальное псевдорешение φ^+ выражается как

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

Напомним, что $\lambda_j > 0,\ j = 1,\dots, p$, 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

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

где практический ранг p_Π матрицы K определяется количеством сингулярных чисел λ_j, удовлетворяющих условию:

$$\frac{\lambda_j}{\lambda_{\max}} \ge \gamma_0 , \tag{1.11}$$

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.

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

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, i*1*, i*2*, j*1*, j*2*).*** 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:

$$Ps\_Solve(K,f,\gamma_0). \quad (1.13)$$

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 $U^{\langle j \rangle}, V^{\langle j \rangle}$. ♦

$$ORIGIN := 1$$

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

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

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

Рис. 1.1. Сингулярное разложение матрицы *K*

$$
Ps\_Solve(K, f, \gamma_0) := \left|
\begin{array}{l}
N \leftarrow rows(K) \\
M \leftarrow cols(K) \\
\lambda \leftarrow svds(K) \\
UV \leftarrow svd(K) \\
U \leftarrow submatrix(UV, 1, N, 1, M) \\
V \leftarrow submatrix(UV, N+1, N+M, 1, M) \\
\lambda_{max} \leftarrow max(\lambda) \\
\phi_{ps} \leftarrow \displaystyle\sum_{j=1}^{M} if\left[\left(\frac{\lambda_j}{\lambda_{max}}\right) \ge \gamma_0, \left(\frac{f \cdot U^{\langle j \rangle}}{\lambda_j}\right) \cdot V^{\langle j \rangle}, 0 \cdot V^{\langle j \rangle}\right]
\end{array}
\right.
$$

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.

$ORIGIN := 1$

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

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

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

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

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

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

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

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

$$\left(K^T V_\eta^{-1} K + V_\varphi^{-1}\right)\varphi_Б = K^T V_\eta^{-1}\tilde{f} + V_\varphi^{-1} m_\varphi . \tag{2.1}$$

Заметим, что вектор $\varphi_Б$ для указанных распределений и квадратичной функции потерь $П\left(\varphi_T, \varphi\right) = \left\|\varphi_T - \varphi\right\|^2$ 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

$$F_Б\left[\varphi\right] = \left\|\tilde{f} - K\varphi\right\|^2_{V_\eta^{-1}} + \left\|\varphi - m_\varphi\right\|^2_{V_\varphi^{-1}}. \tag{2.2}$$

Рассмотрим точность построенного байесовского регуляризированного решения при следующих предположениях:

$$M\left[\eta\right] = 0; \quad M\left[\varphi\eta^T\right] = 0. \tag{2.3}$$

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

$$\varepsilon_Б = \varphi_Б - \overline{\varphi}. \tag{2.4}$$

Ошибка может быть представлена в следующем виде:

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

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

$$M\left[\varepsilon_Б\right] = b_Б = \left(K^T V_\eta^{-1} K + V_\varphi^{-1}\right)^{-1} V_\varphi^{-1}\left(m_\varphi - M_\varphi\left[\varphi\right]\right). \tag{2.6}$$

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

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

с математическим ожиданием $M\left[\xi_Б\right] = 0$ и матрицей ковариации

$$V_{\xi_Б} = \left(K^T V_\eta^{-1} K + V_\varphi^{-1}\right)^{-1} K^T V_\eta^{-1} K \left(K^T V_\eta^{-1} K + V_\varphi^{-1}\right)^{-1}. \tag{2.8}$$

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

$$\Delta_{Б_j} = b_{Б_j} + \sqrt{V_{\xi_{jj}}}, \; j = 1, ..., M. \tag{2.9}$$

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

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. Оптимальный регуляризирующий SVD-алгоритм

Регуляризированное решение запишем в виде

$$\varphi_R = VRU^T C_\eta^{-1/2} \tilde{f} \tag{2.10}$$

или в скалярном виде

$$\varphi_R = \sum_{j=1}^{p} \left[\frac{\lambda_j^2}{\lambda_j^2 + q_j} \overline{x}_j^+ + \frac{\lambda_j}{\lambda_j^2 + q_j} \cdot \langle u_j, n \rangle \right] v_j , \tag{2.11}$$

где матрица R размером $M \times N$ имеет следующую структуру:

$$R = \begin{vmatrix} r_1 & 0 & 0 & \cdot & \cdot & 0 \\ 0 & r_2 & 0 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & r_M & \cdot & \cdot & 0 \end{vmatrix},$$

$$r_j = \frac{\lambda_j}{\lambda_j^2 + q_j}, \quad j = 1, 2, ..., M . \tag{2.12}$$

Величины r_j назовем регуляризирующими множителями.

Здесь $n = C_\eta^{-1/2}\eta$, $\overline{x}_j^+ = \langle v_j, \overline{\varphi}^+ \rangle$, $\overline{\varphi}^+$ - нормальное точное псевдорешение, построенное для точной правой части и определяемое формулой $\overline{\varphi}^+ = \sum_{j=1}^{p_П} \frac{\langle \overline{f}, u_j \rangle}{\lambda_j} \cdot v_j$; v_j, u_j – j-й столбец матрицы V и U соответственно; p – ранг матрицы K.

Матрица ковариации V_η представлена в виде

$$V_\eta = \sigma_\eta^2 \cdot C_\eta \tag{2.13}$$

Очевидно, от величины регуляризирующих множителей зависит ошибка решения φ_R, которую определим функционалом

$$\Delta(R) = M_\eta \left[\left\| \varphi_R - \overline{\varphi}^+ \right\|^2 \right], \tag{2.14}$$

которая может быть записана в виде

$$\Delta(R) = \sum_{j=1}^{p} \left(\frac{q_j}{\lambda_j^2 + q_j} \right)^2 \cdot \left(\overline{x}_j^+ \right)^2 + \sigma_\eta^2 \cdot \sum_{j=1}^{p} \left(\frac{\lambda_j}{\lambda_j^2 + q_j} \right)^2 , \tag{2.15}$$

Определим элементы матрицы R построенного алгоритма путем нахождения величин q_j, $j = 1, 2, ..., p,$ из условия минимума функционала $\Delta(R)$. Дифференцируя (2.15) по q_j и приравнивая производную нулю, получаем оптимальное значение

$$q_{опт_j} = \frac{\sigma_\eta^2}{\left(\overline{x}_j^+ \right)^2}, \quad j = 1, 2, ..., p . \tag{2.16}$$

Оптимальное решение $\varphi_{опт}$, построение при $q_j = q_{опт_j}$ имеет следующее SVD-представление:

$$\varphi_{опт} = \sum_{j=1}^{p}\left[\frac{\lambda_j}{\lambda_j^2 + q_{опт_j}} \cdot \left\langle u_j, C_\eta^{-1/2}\tilde{f}\right\rangle\right] v_j, \qquad (2.17)$$

или в матричном виде:

$$\varphi_{опт} = V_p\, R_{pопт}\, U_p{}^T C_\eta^{-1/2}\tilde{f}, \qquad (2.18)$$

где V_p – матрица размера $M \times p$, составленная из p первых столбцов матрицы V; U_p – матрица размера $N \times p$, составленная из p первых столбцов матрицы U; $R_{pопт}$ – диагональная матрица размера $p \times p$ следующей структуры:

$$R_{pопт} = \begin{vmatrix} \frac{\lambda_1}{\lambda_1^2 + q_{опт_1}} & 0 & \cdots & 0 \\ 0 & \frac{\lambda_1}{\lambda_2^2 + q_{опт_2}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \frac{\lambda_p}{\lambda_p^2 + q_{опт_p}} \end{vmatrix}. \qquad (2.19)$$

Точность построенного решения определим вектором ошибки $\varepsilon_{опт} = \varphi_{опт} - \overline{\varphi}^+$ оптимального решения $\varphi_{опт}$ в виде следующей суммы векторов:

$$\varepsilon_{опт} = \varphi_{опт} - \overline{\varphi}_{опт} + \overline{\varphi}_{опт} - \overline{\varphi}^+ = \xi_{опт} + b_{опт}. \qquad (2.20)$$

Вектор $b_{опт} = \overline{\varphi}_{опт} - \overline{\varphi}^+$ определяет систематическую ошибку решения $\varphi_{опт}$. При этом $b_{опт} = M_\eta\left[\varepsilon_{опт}\right]$ и поэтому вектор $b_{опт}$ можно назвать смещением решения $\varphi_{опт}$. Вектор $\xi_{опт} = \varphi_{опт} - \overline{\varphi}_{опт}$ является случайным вектором с нулевым средним $M\left[\xi_{опт}\right] = 0$ и характеризует случайную ошибку решения $\varphi_{опт}$. Таким образом, для вектора смещения имеем

$$b_{опт} = \overline{\varphi}_{опт} - \overline{\varphi}^+ = VRU^T C^{-1/2}\overline{f} - \overline{\varphi}^+ = V(R\Lambda - I)\overline{x}^+. (2.21)$$

Случайный вектор $\xi_{опт}$

$$\xi_{опт} = \varphi_{опт} - \overline{\varphi}_{опт} = \varphi_{опт} = V_p\, R_{pопт}\, U_p{}^T C_\eta^{-1/2}\eta \qquad (2.22)$$

имеет нулевое среднее и матрицу ковариации

$$V_\xi = \sigma_\eta^2 V_p\, R_{pопт}^2 V_p{}^T. \qquad (2.23)$$

Таким образом, полная ошибка равна

$$\Delta_{опт_j} = b_{опт_j} + \sqrt{V_{\xi_{jj}}},\ j = 1,\dots,p. \qquad (2.24)$$

Задание 2.2.

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

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.

4. Представить матрицу ковариации ошибок V_η в форме $V_\eta = \sigma^2_\eta \cdot C_\eta$, где $C_\eta = \dfrac{1}{\sigma^2_\eta} V_\eta$.

4. Вычислить решение по формуле (2.10) и вектор ошибки (2.24).

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

§2.3. Построение регуляризованного решения при неполной информации

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

$$F_\alpha[\varphi] = \left\|\tilde{f} - K\varphi\right\|^2_{W_f} + \alpha\left\|\varphi - \omega_\varphi\right\|^2_{W_\varphi} \qquad (2.25)$$

В этом случае вектор φ_α, доставляющий минимум функционалу (2.25), является решением системы

$$\left(K^T W_f K + \alpha W_\varphi\right)\varphi_\alpha = K^T W_f \tilde{f} + \alpha W_\varphi \omega_\varphi, \qquad (2.26)$$

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

$$\varphi_{\min j} \le \bar{\varphi}_j \le \varphi_{\max j},\ j = 1,...,M, \qquad (2.27)$$

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 (2.27). Математическое ожидание $\hat{m}_\varphi$ такого вектора определяется как

$$\hat{m}_\varphi = \left|\frac{\varphi_{\min_1} + \varphi_{\max_1}}{2}, \frac{\varphi_{\min_2} + \varphi_{\max_2}}{2}, ..., \frac{\varphi_{\min_M} + \varphi_{\max_M}}{2}\right|^T, \qquad (2.28)$$

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

$$\hat{V}_\varphi = diag\left\{\frac{\left(\varphi_{\max_1} - \varphi_{\min_1}\right)^2}{4\mu_\beta}, ..., \frac{\left(\varphi_{\max_M} - \varphi_{\min_M}\right)^2}{4\mu_\beta}\right\}. \qquad (2.29)$$

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:

$$\left(K^T V_\eta^{-1} K + \widehat{V}_\varphi^{-1}\right)\varphi_p = K^T V_\eta^{-1}\tilde{f} + \widehat{V}_\varphi^{-1}\widehat{m}_\varphi . \quad (2.30)$$

Если информация о шуме измерения задана в виде системы неравенств

$$\left|\tilde{f}_i - f_i\right| \le \delta_i ,\ 1 \le i \le N , \quad (2.31)$$

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

$$\widehat{V}_\eta = diag\left\{\frac{\delta_1^2}{12}, \frac{\delta_2^2}{12}, \dots, \frac{\delta_N^2}{12}\right\} \quad (2.32)$$

и используем ее в алгоритме (2.30).

Очевидно, что (2.30) является частным случаем системы (2.26) при следующих заменах:

$$W_f = \widehat{V}_\eta^{-1};\ W_\varphi = \widehat{V}_\varphi^{-1};\ \omega_\varphi = \widehat{m}_\varphi;\ \alpha = 1 . \quad (2.33)$$

Стабилизирующий функционал

Введем квадратичную форму

$$\varphi^T W_\varphi \varphi = \|\varphi\|_{W_\varphi}^2 , \quad (2.34)$$

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

$$W_\varphi = D_p^T D_p , \quad (2.35)$$

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:

$$W_\varphi = \begin{vmatrix} 1 & -1 & & 0 \\ -1 & 2 & -1 & \\ & -1 & 2 & -1 \\ 0 & & -1 & 1 \end{vmatrix} \quad (2.36)$$

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

Вернемся к вектору ω_φ, входящему в функционал (2.25) и названному «пробным» решением. При наличии априорной информации вида (2.27) его можно задать как $\omega_\varphi = \widehat{m}_\varphi$, где $\widehat{m}_\varphi$ определяется выражением (2.28). 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.

$$W_f = C_f V_\eta^{-1} , \quad (2.37)$$

где константа $C_f > 0$. При наличии информации вида (31) матрицу V_η можно определить соотношением (2.32). 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.

$$W_f = I_{N\times N};\ W_\varphi = I_{M\times M}, \tag{2.38}$$

Ошибка решения φ_α.

Определим ошибку решения φ_α, определяемым вектором (2.26)

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

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

$$\varepsilon_\alpha = \varphi_\alpha - \bar{\varphi}_\alpha^+ + \bar{\varphi}_\alpha^+ - \bar{\varphi}^+ = \xi_\alpha + b_\alpha. \tag{2.38}$$

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

$$b_\alpha = \alpha\left(K^T W_f K + \alpha W_\varphi\right)^{-1} W_\varphi\left(\omega_\varphi - \bar{\varphi}^+\right). \tag{2.39}$$

Вектор

$$\xi_\alpha = \varepsilon_\alpha - M_\eta\left[\varepsilon_\alpha\right] = \varepsilon_\alpha - b_\alpha \tag{2.40}$$

является случайным вектором с нулевым средним и определяется выражением

$$\xi_\alpha = \left(K^T W_f K + \alpha W_\varphi\right)^{-1} K^T W_f \eta. \tag{2.41}$$

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

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

Полная ошибка решения равна

$$\Delta_{\alpha_j} = b_{\alpha_j} + \sqrt{V_{\xi\alpha_{jj}}},\ j = 1,\dots,M. \tag{2.43}$$

Задание 2.3.

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

2. Вычислить среднее значение $\widehat{m}_\varphi$ и матрицу ковариации $\widehat{V}_\varphi$ по формулам (2.28), (2.29).

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

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

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

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 $\tilde{f}$ вида (2.31) и вычислить матрицу $\widehat{V}_\eta$ 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) $W_f = \widehat{V}_\eta^{-1},\ W_\varphi = \widehat{V}_\varphi^{-1},\ \omega_\varphi = \widehat{m}_\varphi,\ \alpha = 1$.

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

§2.4. Регуляризирующий SVD-алгоритм

Обратимся к системе уравнений (2.26) и вычислим решение этой системы, используя сингулярное разложение.

Предположим, что матрица W_f неотрицательно определена, симметрична и допускает представление

$$W_f = W_f^{1/2} \cdot W_f^{1/2}. \tag{2.44}$$

Запишем сингулярное разложение

$$W_f^{1/2} K = U \Lambda V^T \tag{2.45}$$

и представим матрицу W_φ в форме

$$W_\varphi = V \cdot diag\{m(\lambda_1),...,m(\lambda_M)\} \cdot V^T \tag{2.46}$$

Тогда вектор φ_α регуляризированного решения СЛАУ можно представить как

$$\varphi_\alpha = \sum_{j=1}^{p}\left[\frac{\lambda_j}{\lambda_j^2 + \alpha m(\lambda_j)} \cdot \left\langle u_j, W_f^{1/2}\tilde{f}\right\rangle + \frac{\alpha m(\lambda_j)}{\lambda_j^2 + \alpha m(\lambda_j)}\right] v_j, \tag{2.47}$$

где v_j, u_j – j-е столбцы матриц V, U соответственно: p – ранг (или практический ранг) матрицы K. Из (2.47) непосредственно следует матричное представление решения φ_α:

$$\varphi_\alpha = V_p R_{p\alpha} U_p{}^T W_f^{1/2}\tilde{f} + V_p Z_{p\alpha} V_p^T \omega_\varphi, \tag{2.48}$$

где V_p – матрица размера $M \times p$, составленная из p первых столбцов матрицы V; U_p – матрица размера $N \times p$, составленная из p первых столбцов матрицы U; $R_{p\alpha}$ – диагональная матрица размера $p \times p$ следующей структуры:

$$R_{p\alpha} = \begin{vmatrix} \frac{\lambda_1}{\lambda_1^2 + \alpha m(\lambda_1)} & 0 & \cdots & 0 \\ 0 & \frac{\lambda_2}{\lambda_2^2 + \alpha m(\lambda_2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{\lambda_p}{\lambda_p^2 + \alpha m(\lambda_p)} \end{vmatrix}. \quad (2.49)$$

Матрица $Z_{p\alpha}$ имеет структуру

$$Z_{p\alpha} = \begin{vmatrix} \frac{\alpha m(\lambda_1)}{\lambda_1^2 + \alpha m(\lambda_1)} & 0 & \cdots & 0 \\ 0 & \frac{\alpha m(\lambda_2)}{\lambda_2^2 + \alpha m(\lambda_2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{m(\lambda_p)}{\lambda_p^2 + \alpha m(\lambda_p)} \end{vmatrix}. \quad (2.50)$$

Функция $m(\lambda)$ является неубывающей функцией, например

$$m(\lambda) = \frac{1}{\lambda^\gamma}, \quad (2.51)$$

где $\gamma \geq 0$. Если $\gamma = 0$, то $W_\varphi = I$, что соответствует регуляризации нулевого порядка. Чем больше значение γ, тем в большей степени проекции вектора φ_α взаимосвязаны между собой. Это обусловлено тем, что векторы v_j, соответствующие малым λ_j и имеющие осциллирующие проекции, не войдут в решение φ_α из-за пренебрежимо малого значения множителя $\frac{\lambda_j}{\lambda_j^2 + \alpha m(\lambda_j)}$.

Точность решения оценим как и ранее вектором $\varepsilon_\alpha = \varphi_\alpha - \overline{\varphi}^+$, который можно представить как $\varepsilon_\alpha = \xi_\alpha + b_\alpha$.

Тогда полная погрешность вычисляется по формуле

$$\Delta_{\alpha_j} = b_{\alpha_j} + \sqrt{V_{\xi\alpha_{jj}}}, \; j = 1, \ldots, M, \quad (2.52)$$

где вектор b_α (см. формулу (39)) в сингулярном представлении равен

$$b_\alpha = V_p \cdot Z_{p\alpha} \cdot V_p^T (\omega_\varphi - \overline{\varphi}^+), \quad (2.53)$$

а матрица ковариации решения (42) имеет вид

$$V_{\xi\alpha} = V_p F_{p\alpha} V_p^T. \quad (2.54)$$

Матрица

$$F_{p\alpha} = diag\left\{\left(\frac{\lambda_1}{\lambda_1^2 + \alpha m(\lambda_1)}\right)^2, \left(\frac{\lambda_2}{\lambda_2^2 + \alpha m(\lambda_2)}\right)^2, \ldots, \left(\frac{\lambda_p}{\lambda_p^2 + \alpha m(\lambda_p)}\right)^2\right\},$$

а вектор $\overline{\varphi}^+$ равен $\overline{\varphi}^+ = V\Lambda^{-1}U^T W^{1/2} \overline{f}$.

Задание 2.4.

1. Задать матрицу системы K, вектор решения 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

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

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.

4. Задать ограничения на вектор $\widetilde{f}$ вида (2.31) и вычислить матрицу $\widehat{V}_\eta$ по формуле (2.32).

5. Представить матрицу ковариации ошибок $\widehat{V}_\eta$ в форме $\widehat{V}_\eta = \sigma^2_\eta \cdot C_\eta$, где $C_\eta = \dfrac{1}{\sigma^2_\eta}\widehat{V}_\eta$.

6. Вычислить решение системы по формуле (2.48) и погрешность по формуле (2.52) для следующих данных: $W_f = \widehat{V}_\eta^{-1}$, $W_\varphi = I$, $\omega_\varphi = 0$ для различных значений $\alpha \in [10^{-8}, 1]$. Значения α определять по формуле $\alpha_k = \alpha_{k-1} \cdot 10$. Значения параметра γ выбрать эмпирически.

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

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

Здесь K - матрица системы размерности $N \times M$, $\widetilde{f}$ - вектор правой части размерности N, равный $\widetilde{f} = \overline{f} + \eta$, 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 $V_\eta = \sigma^2_\eta \cdot C_\eta$, где σ^2_η - скалярная величина, C_η - матрица размерности $N \times N$.

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

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

где $\alpha > 0$ - параметр регуляризации.

Так как σ^2_η 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 $e_\alpha = \tilde{f} - K\varphi_\alpha = E(\alpha)\tilde{f}$, где $E(\alpha)$ - оператор невязки, который для решения, определяемого из системы (3), равен

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

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

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

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

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

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

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

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

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

Для проверки гипотезы (3.6) введем статистику

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

Подставляя (3.4) в (3.7), получим

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

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

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

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

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

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

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

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

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

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

$\vartheta_N\left(0.0\right.$	0 .71	1 .14	1 .64	2 .17	2 .73	3 .32	3 .94
$\vartheta_N\left(0.9\right.$	9 .49	1 1.0	1 2.6	1 4.1	1 5.5	1 6.9	1 8.3

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

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

Для вычисления γ_W используем итерационную процедуру ньютоновского типа:

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

с начальным значением $\gamma^{(0)} \sim 10^{-15}$. В качестве γ_W принимается значение $\gamma^{(n)}$, удовлетворяющее (3.12). Заметим, что эта процедура решает нелинейное уравнение $\rho_W\left(\gamma\right) = N$, 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 $\rho_W\left(\gamma\right)$ в форме $\rho_W\left(\gamma\right) = \tilde{f}^T \cdot q_\gamma$, где вектор q_γ является решением системы

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

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

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

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 с использованием SVD-разложения на основе критерия оптимальности

Предположим, что:

- ковариационная матрица V_η допускает представление $V_\eta = \sigma_\eta^2 C_\eta$, и определим сингулярное разложение

$$C_\eta^{-1/2} K = U \Lambda V^T\ , \quad (3.16)$$

допуская при этом, что сингулярные числа λ_j упорядочены по убыванию, т.е. $\lambda_j \ge \lambda_{j+1}$ и $\lambda_j = 0, j = p+1, ..., M$, где p – ранг (или практический ранг) матрицы системы;

- регуляризированное решение φ_α представимо в виде:

$$\varphi_\alpha = \sum_{j=1}^{p}\left[\frac{\lambda_j}{\lambda_j^2 + \alpha m\left(\lambda_j\right)} \cdot \left\langle u_j, C_\eta^{-1/2} \tilde{f}\right\rangle\right] v_j\ , \quad (3.17)$$

где $m(\lambda)$ невозрастающая функция

$m(\lambda)=\frac{1}{\lambda^{\tau}}$, где $\tau \ge 0$.

Нетрудно показать, что вектор φ_α является решением системы

$$\left(K^T C_\eta^{-1} K+\alpha W_\varphi\right)\varphi_\alpha = K^T C_\eta^{-1}\tilde{f}\ , \tag{3.18}$$

в которой матрица W_φ выражается соотношением:

$W_\varphi = V_p\, diag\left\{m(\lambda_1), m(\lambda_2),...,m(\lambda_M)\right\}V_p^T$,

где V_p – матрица размера $M\times p$, составленная из p первых столбцов матрицы V, входящей в разложение (3.16).

Теперь статистику $\rho_W(\alpha)$ можно записать в виде

$$\rho_w(\alpha)=\frac{1}{\sigma_\eta^2}\left(C_\eta^{-1/2}\tilde{f}\right)^T\left(C_\eta^{-1/2}e_\alpha\right), \tag{3.19}$$

где векторы $C_\eta^{-1/2}\tilde{f}$, $C_\eta^{-1/2}e_\alpha$ допускают представление:

$$C_\eta^{-1/2}\tilde{f}=\sum_{j=1}^N \langle u_j, C_\eta^{-1/2}\tilde{f}\rangle\cdot u_j=\sum_{j=1}^N \tilde{y}_j u_j\ , \tag{3.20}$$

$$C_\eta^{-1/2}e_\alpha=\sum_{j=1}^p\left[\frac{\alpha m(\lambda_j)}{\lambda_j^2+\alpha m(\lambda_j)}\tilde{y}_j\right]u_j+\sum_{j=p+1}^N \tilde{y}_j u_j \tag{3.21}$$

где $\tilde{y}_j=\left(u_j, C_\eta^{-1/2}\tilde{f}\right)$. Тогда

$$\rho_W(\alpha)=\frac{1}{\sigma_\eta^2}\left[\sum_{j=1}^p\frac{\alpha m(\lambda_j)}{\lambda_j^2+\alpha m(\lambda_j)}\tilde{y}_j^2+\sum_{j=p+1}^N\tilde{y}_j^2\right]. \tag{3.22}$$

Введем $\gamma = 1/\alpha$ и функции

$$R_W(\gamma)=\frac{1}{\sigma_\eta^2}\cdot\sum_{j=1}^p\frac{m(\lambda_j)}{\gamma\lambda_j^2+m(\lambda_j)}\cdot\tilde{y}_j^2\ , \tag{3.23}$$

$$R_W'(\gamma)=\frac{\partial R_W(\gamma)}{\partial\gamma}=-\frac{1}{\sigma_\eta^2}\cdot\sum_{j=1}^p\frac{\lambda_j^2 m(\lambda_j)}{\left[\gamma\lambda_j^2+m(\lambda_j)\right]^2}\cdot\tilde{y}_j^2\ . \tag{3.24}$$

Для вычисления γ_W используем итерационную процедуру:

$$\gamma^{(n)}=\gamma^{(n-1)}-\frac{R_W\left(\gamma^{(n-1)}\right)-p}{R_W'\left(\gamma^{(n-1)}\right)},\ n=1,2,..., \tag{3.25}$$

с начальным значением $\gamma^{(0)}\sim 10^{-15}$. В качестве γ_W принимается значение $\gamma^{(n)}$, удовлетворяющее условию

$$\vartheta_p(\beta/2)\le R_W(\gamma)\le\vartheta_p(1-\beta/2). \tag{3.26}$$

Задание 3.2.

Для построенного регуляризованного решения в лабораторной работе №2 при неполной информации (см. задание 2.4 в работе №2) используя SVD-разложение, 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

§3.3. 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 $e_\alpha = \tilde{f} - K\varphi_\alpha$ - 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:

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

Тогда выражение для статистики примет вид:

$$\rho_V(\alpha) = e_\alpha^T V_\eta^{-1} e_\alpha =$$

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

Введем $\gamma = 1/\alpha$, получим

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

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

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

где

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

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

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

где

$$\begin{aligned} g' &= -\left(K W_\varphi^{-1} K^T\right)\left(V_\eta + \gamma K W_\varphi^{-1} K^T\right)^{-2} \tilde{f} = \\ &= -\left(K W_\varphi^{-1} K^T\right)\left(V_\eta + \gamma K W_\varphi^{-1} K^T\right)^{-1} g \end{aligned} . \qquad (3.33)$$

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

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

Для вычисления γ_V используем итерационную процедуру ньютоновского типа:

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

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

Итерационный процесс завершаем при выполнении условия

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

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 3.3.

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

§3.4. Алгоритм поиска α_V с использованием SVD разложения

Пусть ковариационная матрица V_η допускает представление $V_\eta = \sigma_\eta^2 C_\eta$. Запишем сингулярное разложение (3.16)

$C_\eta^{-1/2} K = U \Lambda V^T$.

Теперь статистику $\rho_V(\alpha)$ можно записать в виде

$$\rho_V(\alpha) = \frac{1}{\sigma_\eta^2}\left(C_\eta^{-1/2} e_\alpha\right)^T \left(C_\eta^{-1/2} e_\alpha\right), \tag{3.37}$$

где вектор $C_\eta^{-1/2} e_\alpha$ допускает представление:

$$C_\eta^{-1/2} e_\alpha = \sum_{j=1}^{p}\left[\frac{\alpha m(\lambda_j)}{\lambda_j^2 + \alpha m(\lambda_j)} \tilde{y}_j\right] u_j + \sum_{j=p+1}^{N} \tilde{y}_j u_j \tag{3.38}$$

Тогда для статистики $R_V(\gamma) = \rho_V(1/\alpha)$ и ее производной получим

$$R_V(\gamma) = \frac{1}{\sigma_\eta^2} \cdot \sum_{j=1}^{p}\left[\frac{m(\lambda_j)}{\gamma \lambda_j^2 + m(\lambda_j)}\right]^2 \cdot \tilde{y}_j^2, \tag{3.39}$$

$$R_V'(\gamma) = \frac{\partial R_V(\gamma)}{\partial \gamma} = -\frac{2}{\sigma_\eta^2} \cdot \sum_{j=1}^{p} \frac{\lambda_j^2 m^2(\lambda_j)}{\left[\gamma \lambda_j^2 + m(\lambda_j)\right]^3} \cdot \tilde{y}_j^2, \tag{3.40}$$

где $\tilde{y}_j = \langle u_j, C_\eta^{-1/2} \tilde{f} \rangle$.

Значение γ_V, при котором принимается гипотеза (3.27), удовлетворяет условию

$$\vartheta_p(\beta/2) \le R_V(\gamma_V) \le \vartheta_p(1-\beta/2), \tag{3.41}$$

а $\vartheta_p(\beta/2)$, $\vartheta_p(1-\beta/2)$ – квантили χ^2-распределения с p степенями свободы уровней $\beta/2$, $1-\beta/2$ соответственно.

Для вычисления γ_V вновь используем итерационную процедуру ньютоновского типа:

$$\gamma^{(n)} = \gamma^{(n-1)} - \frac{R_V\left(\gamma^{(n-1)}\right) - p}{R_V'\left(\gamma^{(n-1)}\right)}, \; n = 1, 2, \ldots, \tag{3.42}$$

с начальным значением $\gamma^{(0)} \sim 10^{-15}$. В качестве γ_V принимается значение $\gamma^{(n)}$, удовлетворяющее (3.41).

Задание 3.4.

Для построенного регуляризованного решения в лабораторной работе №2 при неполной информации (см. задание 2.4 в работе №2) используя SVD-разложение, вычислить:

1) параметр регуляризации α_V;

2) регуляризованное решение φ_{α_V};

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

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

Практическое занятие №4

Локальная регуляризация

§4.1. Векторный параметр регуляризации

Рассмотрим сглаживающий функционал вида

$$F_\alpha[\varphi] = \| \tilde{f} - K\varphi \|^2_{W_f} + \alpha \cdot \sum_{j=2}^{M} (\varphi_j - \varphi_{j-1})^2 , \qquad (4.1)$$

в котором в качестве матрицы W_φ используется трехдиагональная матрица (размером $M \times M$) вида

$$W_\varphi = \begin{vmatrix} 1 & -1 & & & 0 \\ -1 & 2 & -1 & & \\ & & \ddots & & \\ & & & \ddots & \\ 0 & & & -1 & 1 \end{vmatrix} \qquad (4.2)$$

Вместо скалярного параметра регуляризации введем во второе слагаемое этого функционала *векторный параметр регуляризации*

$$\mu = \{ \mu_2, \mu_3, \dots, \mu_M \} . \qquad (4.3)$$

Тогда имеем новый сглаживающий функционал:

$$F[\varphi, \mu] = \| \tilde{f} - K\varphi \|^2_{W_f} + \sum_{j=2}^{M} \mu_j^2 (\varphi_j - \varphi_{j-1})^2 . \qquad (4.4)$$

Определив матрицу:

$$M(\mu) = \begin{vmatrix} \mu_2^2 & -\mu_2^2 & & & 0 \\ -\mu_2^2 & (\mu_2^2 + \mu_3^2) & -\mu_3^2 & & \\ 0 & -\mu_3^2 & (\mu_3^2 + \mu_4^2) & -\mu_4^2 & \\ & & \ddots & & \ddots \\ & 0 & & -\mu_M^2 & \mu_M^2 \end{vmatrix} , \qquad (4.5)$$

функционал (4.4) можно переписать в виде:

$$F[\varphi, \mu] = \| \tilde{f} - K\varphi \|^2_{W_f} + \| \varphi \|^2_{M(\mu)} . \qquad (4.6)$$

Введем функционал

$$\Gamma[\mu] = \gamma_1^2 \cdot \sum_{j=2}^{M} (\mu_j - \gamma_0)^2 + \gamma_2^2 \cdot \sum_{j=3}^{M} (\mu_j - \mu_{j-1})^2 \qquad (4.7)$$

и ограничения $0 \le \mu_j \le \gamma_0, \ j = 2, \dots, M$.

Введем новый сглаживающий функционал (назовем его *локальным сглаживающим функционалом*)

$$\Phi[\varphi, \mu] = F[\varphi, \mu] + \Gamma[\mu] \qquad (4.8)$$

и определим точку его минимума (φ_μ^*, μ^*) из условий:

$$\frac{\partial \Phi[\varphi, \mu]}{\partial \varphi_j} = 0, \quad j = 1, \dots, M; \qquad (4.9)$$

$$\frac{\partial \Phi[\varphi, \mu]}{\partial \mu_j} = 0, \quad j = 2, \dots, M. \qquad (4.10)$$

Точку минимума (φ^*, μ^*) функционала $\Phi[\varphi, \mu]$ ищем итерационным путем из решения совместных систем вида

$$\sum_{j=1}^{M} \left(K^T W_f K + M(\mu^{(l)}) \right)_{i,j} \varphi_j^{(l)} = \left(K^T W_f \tilde{f} \right)_i , \ i = 1, \dots, M , \qquad (4.11)$$

$$\left[\left(\varphi_2^{(l)}-\varphi_1^{(l)}\right)^2+\gamma_1^2+2\gamma_2^2\right]\cdot\mu_2^{(l+1)}-\gamma_2^2\cdot\mu_3^{(l+1)}=\gamma_1^2\gamma_0;$$

$$-\gamma_2^2\cdot\mu_{i-1}^{(l+1)}+\left[\left(\varphi_i^{(l)}-\varphi_{i-1}^{(l)}\right)^2+\gamma_1^2+2\gamma_2^2\right]\cdot\mu_i^{(l+1)}-\gamma_2^2\cdot\mu_{i+1}^{(l+1)}=\gamma_1^2\gamma_0, i=3,...,M-1;$$

$$-\gamma_2^2\cdot\mu_{M-1}^{(l+1)}+\left[\left(\varphi_M^{(l)}-\varphi_{M-1}^{(l)}\right)^2+\gamma_1^2+2\gamma_2^2\right]\cdot\mu_M^{(l+1)}=\gamma_1^2\gamma_0. \quad (4.12)$$

$$\mu_i^{(l+1)}=\begin{cases}\mu_i^{(l+1)},\ \text{если}\ 0\le\ \mu_i^{(l+1)}\le\gamma_0;\\ 0,\ \text{если}\ \mu_i^{(l+1)}<0;\\ \gamma_0,\ \text{если}\ \mu_i^{(l+1)}>\gamma_0.\end{cases}, \ l=1,2,.... \quad (4.13)$$

$$\gamma_0=\sqrt{\alpha_W}\ . \quad (4.14)$$

Условием прекращения итераций является одновременное выполнение условий:

$$\frac{\left\|\ \mu^{(l)}-\mu^{(l-1)}\right\|}{\left\|\ \mu^{(l)}\right\|}\le\varepsilon\ ;\ \frac{\left\|\ \varphi^{(l)}-\varphi^{(l-1)}\right\|}{\left\|\ \varphi^{(l)}\right\|}\le\varepsilon\ , \quad (4.15)$$

где ε – достаточно малая величина – порядка $10^{-4}\div10^{-3}$. В качестве начального значения $\mu^{(0)}$ примем вектор с проекциями

$$\mu_i^{(0)}=\frac{\sqrt{\alpha_W}}{2},\quad i=2,...,\ M, \quad (4.16)$$

где α_W – оценка α_{opt} по критерию оптимальности (см. п. 3.1).

Значения γ_1 и γ_2 в системе (4.12) равны: $\gamma_1=5$, $\gamma_2=0.5$

Задание 4.1.
По формулам (4.11) – (4.16) построить локальное регуляризованное решение, используя найденное в лабораторной работе №3 (см. задание 3.1) глобальное регуляризованное решение.

Практическое занятие №5

Дескриптивные алгоритмы решения СЛАУ

§ 5.1. Глобальный дескриптивный регуляризирующий алгоритм

Пусть вектор решения φ принадлежит выпуклому множеству Ф, задаваемому системой линейных неравенств:

$$G\varphi\le g\ , \quad (5.1)$$

где G – матрица ограничений размером $L\times M$; g – вектор размерности L. Ограничения (5.1) определяют допустимое множество векторов φ, из которых должен находиться вектор регуляризированного решения. Например, для условия неотрицательности проекций вектора φ ($\varphi_i\ge0,\ i=1,2,...,M$) система неравенств (5.1) преобразуется к виду

$$-I_{M\times M}\cdot\varphi\le0_M\ , \quad (5.3)$$

здесь $I_{M\times M}$ – единичная матрица; 0_M – нулевой вектор.

В этом случае искомый вектор φ_α^*, будет являться решением следующей задачи квадратичного программирования :

$$\inf\left[\left\|\tilde f-K\varphi\right\|^2_{C_\eta^{-1}}+\alpha\left\|\varphi\right\|^2_{W_\varphi}\right] \quad (5.3)$$

при ограничениях $G\varphi\le g$.

Для приведения исходной задачи к двойственной вновь запишем диагональную матрицу $R_{p\alpha}$ (см. (2.49)):

$$R_{p\alpha}=diag\{r_1(\alpha),r_2(\alpha),...,r_p(\alpha)\} \quad (5.4)$$

размером $p\times p$ с элементами

$$r_j(\alpha)=\frac{\lambda_j}{\lambda_j^2+\alpha m(\lambda_j)},\quad j=1,...,p\ , \quad (5.5)$$

где p – ранг матрицы, и введем вектор $z_p = \left| y_1, y_2, ..., y_p \right|^T$, составленный из первых p проекций вектора $y = U^T C_\eta^{-1/2} \tilde{f}$.

Тогда регуляризированное решение φ_α, доставляющее минимум функционалу (5.3) с матрицей

$$W_\varphi = V_p diag\{m(\lambda_1), ..., m(\lambda_p)\} V_p^T,$$

определяется как $\varphi_\alpha = V_p x_\alpha$, где V_p – матрица размером $M \times p$, составленная из p первых столбцов матрицы V. Вектор

$$x_\alpha = R_{p\alpha} z_p, \qquad (5.5)$$

состоящий из p проекций, доставляет минимум функционалу:

$$F_\alpha(x) = x^T R_{p\alpha}^{-1} x - 2x^T z_p + const. \qquad (5.6)$$

Введя обозначения $D_{p\alpha} = 2R_{p\alpha}^{-1}$, $d_p = -2z_p$, с учетом $\varphi_\alpha = V_p x_\alpha$ и ограничений (5.1), приходим к задаче квадратичного программирования, а именно:

Задача А. *Найти вектор* x_α^* *размерности* M, *доставляющий минимум функционалу*

$$F_\alpha(x) = \frac{1}{2} x^T D_{p\alpha} x + x^T d_p + const \qquad (5.7)$$

при ограничении

$$G V_p x \le g. \qquad (5.8)$$

Двойственная по Лагранжу задача формулируется следующим образом:

Задача В. *Найти вектор* μ^* *размерности* L, *доставляющий минимум функционалу*

$$\tilde{\psi}_\alpha(\mu) = \frac{1}{4} \mu^T G V_p R_\alpha V_p^T G^T \mu - \mu^T(-g + G V_p x_\alpha) + const \quad (4.1.15)$$

при ограничении

$$\mu \ge 0_L \qquad (5.9)$$

Решение *задачи В* можно осуществить используя известные алгоритмы квадратичного программирования.

После вычисления μ^* решение x_α^* *задачи А* находится из выражения

$$x_\alpha^* = x_\alpha - \frac{1}{2} R_{p\alpha} V_p^T G^T \mu^* \qquad (5.10)$$

и состоит из двух слагаемых: регуляризированного решения x_α (5.5), полученного безусловной минимизацией функционала (5.6), и вектора, зависящего от решения μ^* двойственной *задачи В*. Очевидно, что если $\mu^* = 0$, то $x_\alpha^* = x_\alpha$.

Вектор дескриптивного решения находится как

$$\varphi_\alpha^* = V_p x_\alpha^*. \qquad (5.11)$$

Таким образом, построение дескриптивного решения φ_α^* можно представить следующими шагами:

- выполнение сингулярного разложения $C_\eta^{-1/2} K = U \Lambda V^T$, где $V_\eta = \sigma_\eta^2 C_\eta$, λ_j – сингулярные числа, U, V – ортогональные матрицы;
- вычисление вектора x_α из условия минимума функционала $F_\alpha(x)$ (см. (5.5)), т.е.

$$x_\alpha = diag\left\{ \frac{\lambda_1}{\lambda_1^2 + \alpha m(\lambda_1)}, ..., \frac{\lambda_p}{\lambda_M^2 + \alpha m(\lambda_p)} \right\} \cdot z_p, \qquad (5.12)$$

где p – ранг матрицы системы. При этом вектор α осуществляется одним из способов, описанных в лабораторной работе №3;
- проверка ограничений задачи (5.8). Если ограничения (5.8) выполняются, то $x_\alpha^* = x_\alpha$;
- если ограничения (5.8) не выполняются, то решение вариационной *задачи В;*

- формирование вектора x^*_α, определяемого выражением (5.10);
- вычисление вектора дескриптивного решения (5.11).

Задание 5.1.
Построить дескриптивный алгоритм решения СЛАУ.
Обратиться к выполненному заданию 3.1. Задать ограничения на искомое решение и найти глобальное регуляризованное решение, удовлетворяющее заданным ограничениям.

§ 5.2. Локальный дескриптивный регуляризирующий алгоритм

Запишем оптимальное решение (см. (2.18)) для l-й итерации

$$x_s^{(l)} = R_{pS}^{(l)} z_p, \quad l = 0,1,2,... \tag{5.13}$$

где диагональная матрица $R_{pS}^{(l)}$ (см. (2.19))

$$R_{pS}^{(l)} = diag\{r_{S_1}^{(l)}, r_{S_2}^{(l)}, ..., r_{S_p}^{(l)}\} \tag{5.14}$$

размером $p \times p$ имеет элементы

$$r_{S_j}^{(l)} = \frac{1}{\lambda_j (1 + \widehat{S}_j^{(l)})}, \; j = 1,...,p, \tag{5.15}$$

а вектор $z_p = \left| \tilde{y}_1, \tilde{y}_2, ..., \tilde{y}_p \right|^T$ составлен из первых p проекций вектора $\tilde{y} = U^T C_\eta^{-1/2} \tilde{f}$.

Величины $\widehat{S}_j^{(l)}$ определяются следующим выражением:

$$\widehat{S}_j^{(l+1)} = \begin{cases} S_{j,1}^{*(l)}, \; если \quad \tilde{S}_j \le 1/4 \;\; и \quad 0 \le S_j^{(l)} < S_{j,2}^{*(l)}; \\ S_{j,2}^{*(l)}, \; если \quad \tilde{S}_j \le 1/4 \;\; и \quad S_j^{(l)} = S_{j,2}^{*(l)}; \\ 0, \quad если \quad \tilde{S}_j \le 1/4 \;\; и \quad S_j^{(l)} > S_{j,2}^{*(l)}; \\ 0, \quad если \quad \tilde{S}_j > 1/4. \end{cases} \tag{5.16}$$

Здесь $S_{j,1}^{*(l)}$ и $S_{j,2}^{*(l)}$ определяются как корни квадратного уравнения

$$\left(S_j^{*(l)} \right)^2 + \left(2 - \frac{1}{\tilde{S}_j} \right) S_j^{*(l)} + 1 = 0, \tag{5.17}$$

при этом $S_{j,1}^{*(l)} \le S_{j,2}^{*(l)}$. В формулах (5.19) и (5.20) введены величины:

$$\tilde{S}_j = \sigma_\eta^2 / \tilde{y}_j^2, \; S_j^{(l)} = \sigma_\eta^2 / \left(\lambda_j x_j^{(l)} \right)^2, \; j = 1,...,p, \tag{5.18}$$

где $x_j^{(l)}$ – проекции вектора $x^{(0)} = V_p^T \varphi^{(0)}$. В качестве «стартового» решения $\varphi^{(0)}$ будем использовать регуляризированное решение φ_α, построенное при $\alpha = \alpha_W$.

Итерационный процесс прекращаем при одновременном выполнении условий:

$$\frac{\left\| \widehat{S}^{(l)} - \widehat{S}^{(l-1)} \right\|}{\left\| \widehat{S}^{(l)} \right\|} \le \varepsilon \, ; \; \frac{\left\| x_S^{(l)} - x_S^{(l-1)} \right\|}{\left\| x_S^{(l)} \right\|} \le \varepsilon \, , \tag{5.19}$$

где ε – достаточно малая величина – порядка $10^{-4} \div 10^{-3}$.

Введем обозначения: $d_p = -2 z_p$, $D_{pS} = 2 R_p^{-1}$. С учетом $\varphi_S = V_p x_S$, приходим к следующей задаче квадратичного программирования:

Задача А. *Найти вектор* x_S^* *размерности* p, *доставляющий минимум функционалу*

$$F_S(x) = \frac{1}{2} x^T D_{pS} x + x^T d_p + const \tag{5.20}$$

при ограничении

$$G V_p^T x_S \le g \, . \tag{5.21}$$

Для данной прямой задачи сформулируем двойственную по Лагранжу задачу

Задача В. *Найти вектор* μ^* *размерности* L, *доставляющий минимум*

$$\tilde{\psi}_S(\mu) = \frac{1}{4}\mu^T G V_p R_{pS} V_p^T G^T \mu - \mu^T(-g + G V_p x_S) + const \quad (5.22)$$

при ограничении

$$\mu \geq 0_L . \quad (5.23)$$

Решая данную задачу методами математического программирования, находим вектор μ^* и вектор

$$x_S^* = x_S - \frac{1}{2} R_{pS} V_p^T G^T \mu , \quad (5.24)$$

где x_S определяется из (5.16). Очевидно, что если $\mu^* = 0$, то $x_S^* = x_S$.

Таким образом, построение дескриптивного локального регуляризированного решения φ_S^*, удовлетворяющего ограничениям $G\varphi \leq g$, можно представить следующими этапами:

- вычисляется вектор x_S локального регуляризированного решения (см. (5.13)), удовлетворяющий ограничениям (5.19);
- проверяются ограничения $G V_p x_S \leq g$;
- если эти ограничения выполняются, то $x_S^* = x_S$;
- если ограничения нарушаются, то находится решение μ^* двойственной задачи (5.22), (5.23) и вычисляется вектор x_S^* по формуле (5.24);
- строится вектор решения

$$\varphi_S^* = V_p x_S^* . \quad (5.25)$$

Задание 5.2.

Построить локальный дескриптивный алгоритм решения СЛАУ. Обратиться к заданию 5.1.

Литература

1. Воскобойников Ю.Е., Мицель А.А. Современные проблемы прикладной математики. Часть 1. Лекционный курс: учебное пособие/ Томский гос. ун-т систем управления и радиоэлектроники (ТУСУР). – Томск: 2016. – 136 с.