МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ
САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ

А.В. Ушаков, В.В. Хабалов, Н.А. Дударенко

МАТЕМАТИЧЕСКИЕ ОСНОВЫ ТЕОРИИ СИСТЕМ:

элементы теории и практикум

ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Санкт-Петербург

2007

УДК 517/519:62.50:681.3

Ушаков А.В., Хабалов В.В., Дударенко Н.А. Математические основы теории систем: элементы теории и практикум./ Под ред. Ушакова А.В. – СПб: СПбГУИТМО, 2007.

В учебном пособии излагаются теоретические положения, подкрепленные практикумом, основных разделов учебной дисциплины «Математические основы теории систем» естественнонаучного цикла образовательного стандарта направления 651900 – «Автоматизация и управление» подготовки бакалавров и магистров по специальности 220201 – «Управление и информатика в технических системах» подготовки специалистов – инженеров.

Учебное пособие рассчитано на студентов направления 651900 и специальности 220201, тем не менее, оно может быть рекомендовано также аспирантам и молодым специалистам, которым по роду своей деятельности приходится иметь дело с информационными и динамическими системами и математическими проблемами, связанными с построением модельных представлений таких систем, ориентированными на возможности матричного формализма метода пространства состояния.

Утверждено к печати Советом факультета компьютерных технологий и управления, протокол № 4 от 19.12.2006.

В 2007 году СПбГУ ИТМО стал победителем конкурса инновационных образовательных программ вузов России на 2007–2008 годы. Реализация инновационной образовательной программы «Инновационная система подготовки специалистов нового поколения в области информационных и оптических технологий» позволит выйти на качественно новый уровень подготовки выпускников и удовлетворить возрастающий спрос на специалистов в информационной, оптической и других высокотехнологичных отраслях экономики.

ПРИОРИТЕТНЫЕ НАЦИОНАЛЬНЫЕ ПРОЕКТЫ «Образование»

ISBN 5-7577-0321-0

С О Д Е Р Ж А Н И Е

Посвящается шестидесятилетию основания ***кафедры автоматики и телемеханики*** *(ныне* ***кафедры систем управления и информатики****)*

ПРЕДИСЛОВИЕ

Дисциплина «Математические основы теории систем (МОТС)» к настоящему моменту имеет достаточно длинную предысторию. Первоначально в учебных планах подготовки инженеров–электриков по специальности 0606 – «автоматика и телемеханика» в 70-е годы XX-го века появилась дисциплина «Математические основы кибернетики (МОК)». К концу 70-х годов XX-го века название дисциплины претерпевает первое изменение, в результате чего она стала называться «Теоретическими основами кибернетики (ТОК)». Введенная в учебный план специальности 0606 дисциплина как в версии МОК, так и в версии ТОК в основном решала задачи математического обеспечения модельных представлений процессов управления и информационных процессов в канальных средах. В конце 80-х годов XX-го века дисциплина претерпевает очередное изменение названия, в результате чего она начинает называться «Математическими основами исследования процессов управления (МОИПУ)». Из программы дисциплины МОИПУ изымаются положения, связанные с информационными процессами в канальных средах, которые переносятся в программу появившейся в учебном плане специальности дисциплины «Прикладная теория информации (ПТИ)».

Последняя модификация названия дисциплины, в результате которой оно получило действующую в настоящий момент версию «Математические основы теории систем (МОТС)», произошла в начале 90-х годов XX-го века с одновременным изменением номера и названия специальности инженерной подготовки так, что выпускники вузов по данной специальности стали получать квалификацию инженера по специальности 2101 (ныне 220201) – «управление и информатика в технических системах». С середины 90-х годов XX-го века дисциплина МОТС вошла также в структуру учебного плана по разделу естественнонаучных дисциплин образовательного стандарта направления 651900 – «Автоматизация и управление» подготовки бакалавров и магистров.

Таким образом, предлагаемое учебное пособие подготовлено на основе опыта преподавания дисциплин МОК, ТОК, МОИПУ и «Математические основы теории систем», накопленного на кафедре Систем управления и информатики (до 2001-го года – кафедре автоматики и телемеханики) Санкт-Петербургского государственного университета информационных технологий, механики и оптики.

При подготовке учебного пособия использованы материалы ранее опубликованных научных и учебно-методических работ, содержательно коррелирующих с проблематикой дисциплины МОТС: Никифорова Л.Т., Ушаков А.В., Хабалов В.В. Теоретические основы кибернетики. Учебное пособие. Л.: ЛИТМО, 1984; Матричные уравнения в задачах управления и наблюдения непрерывными объектами / Т.А. Акунов, С. Алишеров, Р.О. Оморов, А.В. Ушаков / Под ред. А.В. Ушакова. – Бишкек: Илим, 1991; Модальные оценки качества процессов в линейных многомерных системах / Т.А. Акунов, С. Алишеров, Р.О. Оморов, А.В. Ушаков / Под ред. А.В. Ушакова. – Бишкек: Илим, 1991.; Матричные уравнения в исследовании дискретных процессов над бесконечными и конечными полями / Т.А. Акунов, С. Алишеров, Р.О. Оморов, А.В. Ушаков / Под ред. А.В. Ушакова. – Бишкек: Илим, 1993.; Алгебраические методы в теории устройств дискретной автоматики и телемеханики: Труды лаборатории телемеханики СПбГИТМО(ТУ). / Под ред. А.В.Ушакова. – СПб: СПбГИТМО (ТУ), 2001; Акунова А., Баячорова Б.Ж., Ушаков А.В., Хабалов В.В. Математические основы теории информационных систем в управлении / Под ред. А.В. Ушакова. – Бишкек: Салам, 2003; Дударенко Н.А. Технология контроля вырождения сложных динамических систем с помощью частотных сепаратных чисел обусловленности // Современные технологии: Сборник научных статей / Под. ред. проф. С.А. Козлова. – СПб: СПб ГИТМО (ТУ), 2003; Мельников А.А., Ушаков А.В. Двоичные динамические системы дискретной автоматики / Под ред. А.В. Ушакова. СПб: СПбГУ ИТМО, 2005; Дударенко Н.А., Ушаков А.В. Анализ чувствительности функционала вырождения к параметрической неопределенности функциональных компонентов сложных систем при стохастических экзогенных воздействиях. Мехатроника, автоматизация, управление. №8, 2006.

Искреннюю благодарность авторы выражают своему коллеге Т.А. Акунову за материалы, составившие содержание приложения «Сведения о пакете MATLAB 6.5.»

Конструктивную критику по существу содержания учебного пособия следует направлять автору по: почтовому адресу 195267 до востребования, телефонам 5954128, 2900744 и электронной почте ushakov-AVG@yandex.ru.

Издание настоящего пособия поддержано грантом РФФИ 06–08–01427а.

Используемые обозначения и сокращения

S, X – множество элементов произвольной природы;

$G, G_0, F, GF(p), GF(p^n)$ – алгебраические структуры соответственно группа, подгруппа, поле, простое поле Галуа с характеристикой (модулем) p, расширенное поле Галуа;

$\{X, d\}, X_d$ – метрическое пространство с метрикой $d = d(x, y)$;

$\boldsymbol{A}$, A – соответственно линейный оператор (ЛО) и матрица ЛО;

X^n – n-мерное линейное пространство над полем F;

R^n – линейное вещественное пространство;

I – единичная матрица;

$\boldsymbol{0}$ – нулевой скаляр, вектор, матрица;

A, A^j, A_k – матрица, j-ая строка, k-ый столбец матрицы A;

A^T – матрица, транспонированная к матрице A;

A^* – матрица, сопряженная к матрице A;

A^{-1} – матрица, обратная к матрице A;

A^+ – матрица, псевдообратная к матрице A;

$diag\{\alpha_i, i = \overline{1, n}\}$ – диагональная матрица с элементами α_i на диагонали;

$row\{\alpha_i, i = \overline{1, n}\}$ – строчная матричная структура с элементами α_i в строке;

$col\{\alpha_i, i = \overline{1, n}\}$ – столбцовая матричная структура с элементами α_i в столбце;

$\|(\circ)\|$ – норма элемента $(\circ)$;

$\|(\circ)\|_P$ – норма элемента $(\circ)$ с весом P;

$ang\{x, y\}$ – угол между векторами x и y;

$\stackrel{\Delta}{=}$ – равенство по определению;

$\forall$ – для всех;

$\exists$ – существует;

$\in$ – принадлежит;

$\notin$ – не принадлежит;

$\max\limits_i$ – максимум на множестве элементов с индексом i;

$\bigcup, \bigcap$ – символы объединения и пересечения множеств;

$\gamma = \arg\{\beta(\gamma)\}$ – значение γ, удовлетворяющее *условию* $\beta(\gamma)$;

$det(\circ), tr(\circ), rank(\circ)\{rang(\circ)\}$ – соответственно определитель, след, ранг матрицы $(\circ)$;

$\exp(\circ)$ – матричная экспонента с матричным аргументом $(\circ)$;

$cond\{(\circ)\} = C\{(\circ)\}$ – число обусловленности матрицы $(\circ)$;

$\dim(\circ)$ – размерность элемента $(\circ)$;
$\deg(\circ)$ – степень полинома $(\circ)$;
$Jm(\circ)$ – образ $(\circ)$ ЛО;
$Ker(\circ)$ – ядро $(\circ)$ ЛО;
$\sigma\{\circ\}$, $\sigma_a\{\circ\}$, $\sigma_\alpha\{\circ\}$ – соответственно алгебраические спектры собственных значений, коэффициентов характеристического полинома и сингулярных чисел матрицы $\{\circ\}$;
$\otimes$ – символ кронекеровского произведения;
$contr\{(A,B)\}$ – предикат наличия свойства управляемости пары матриц (A,B);
$observ\{(A,C)\}$ – предикат наличия свойства наблюдаемости пары матриц (A,C);
$\vee$ – логическое "или";
& – логическое "и";
$(\circ){:}\eta$; $(\circ)|\eta$ – предикат наличия характеристического свойства η у элемента $(\circ)$;
$rest(rem)\left\{{*}/{\circ}\right\}$ – остаток от деления ${*}/{\circ}$;
SVD – сингулярное разложение матриц;
МПС – метод пространства состояний;
ВМП – векторно-матричное представление;
ВВ – вход–выход;
ВСВ – вход–состояние–выход;
КС – канал связи;
ОС – обратная связь;
ОУ – объект управления.

Введение. Основные проблемы управления

Основные проблемы управления рассмотрим на примере сложной системы, структурное представление которой приведено на рисунке В.1.

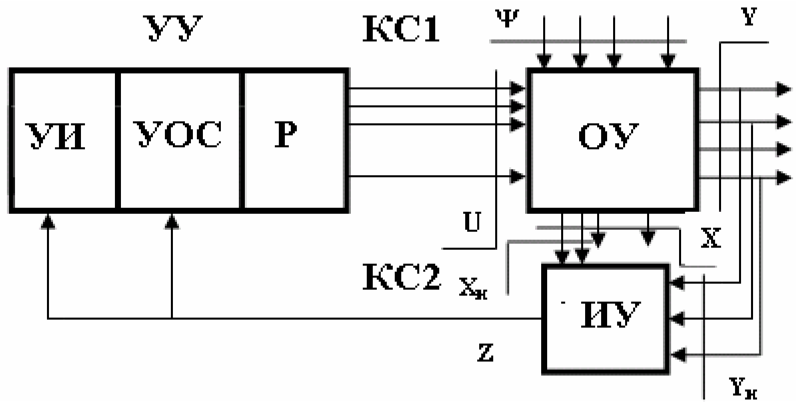

Рисунок В.1. Структурная схема сложной системы управления

На рисунке В.1: ОУ – *объект управления*, представляющий собой некоторый физический объект (технологический процесс), на котором размещены *регулирующие органы* (РО), управляемые сформированными по некоторому закону *сигналами управления*; ИУ – измерительное устройство, преобразующее *доступные непосредственному измерению* компоненты вектора состояния и вектора выхода в электрический сигнал, согласованный с предоставленным каналом связи (КС); УИ – устройство *идентификации* объекта управления; УОС – устройство *оценивания состояния* объекта управления; Р – регулятор, представляющий собой техническую среду, средствами которой создается сигнал управления U, сформированный в соответствии с требуемым *закон управления* (ЗУ) регулирующими органами ОУ; УУ – *устройство управления*, представляющее собой функциональное объединение устройства идентификации объекта, устройства оценивания его состояния и регулятора; КС1, КС2 – соответственно *прямой* (управляющий) и *обратный* (информационный, известительный) *каналы связи*.

Таким образом, сложная система представляет собой функциональное объединение *объекта управления, устройства управления и канальной среды*, образованной прямым и обратным каналами связи.

В современной теории управления объект управления задается с помощью макровектора

ОУ=$\{U,\Psi,X,Y,T,\Omega_U,\Omega_X,\lambda,\delta,\lambda_\Psi,\delta_\Psi,F\}$ (В.1)

В макровекторе (В.1): $U=[U_1,U_2,\ldots,U_r]^T=col\{U_j,j=\overline{1,r}\}$ – r-мерный *вектор управления* объектом;

$\Psi = [\Psi_1, \Psi_2, \ldots, \Psi_v]^T = col\{\Psi_l, l = \overline{1,v}\}$ – v-мерный *вектор внешних возмущающих воздействий*, осуществляющих *нежелательное* управление объектом; $X = [X_1, X_2, \ldots, X_n]^T = col\{X_i, i = \overline{1,n}\}$ – n-мерный *вектор состояния*, содержательно выполняющего функцию памяти объекта; $Y = [Y_1, Y_2, \ldots, Y_m]^T = col\{Y_k, k = \overline{1,m}\}$ – m-мерный *вектор выхода*, содержательно представляющий собой выходную пользовательскую продукцию объекта управления как некоторого технологического процесса; T – временной *интервал управления* объектом, представляющий собой сплошное множество (континуум) моментов управления в случае, если ОУ имеет непрерывную природу, и счетное множество моментов управления в случае, если объект имеет дискретную природу; Ω_U – множество (область в r-мерном пространстве) *допустимых управлений*; Ω_X – множество (область в n-мерном пространстве состояния) *допустимых траекторий*; $\lambda:\ X \times U \Rightarrow X$ – n-мерная векторная *функция перехода*, описывающая процесс перехода из некоторого *исходного* состояния в состояние *перехода* под действием сформированного управления; $\delta:\ X \times U \Rightarrow Y$ – m-мерная векторная *функция выхода*, описывающая процесс *формирования выхода* объекта при переходе из некоторого *исходного* состояния в состояние *перехода* под действием сформированного управления; $\lambda_\Psi:\ X \times U \times \Psi \Rightarrow \Delta X$ – n-мерная векторная функция, описывающая процесс формирования *дополнительного движения* ΔX *по состоянию* под действием внешнего возмущающего воздействия Ψ при переходе из некоторого исходного состояния в состояние перехода под действием сформированного управления; $\delta_\Psi:\ X \times U \times \Psi \Rightarrow \Delta Y$ – m-мерная векторная функция, описывающая процесс формирования *дополнительного движения* ΔY *по выходу* под действием внешнего возмущающего воздействия Ψ при переходе из некоторого исходного состояния в состояние перехода под действием сформированного управления; F – *числовое поле*, которому принадлежат элементы векторов U, Ψ, X, Y, а также системные параметры векторных функций λ, δ, λ_Ψ, δ_Ψ.

В учебной и научной литературе по теории систем управления в основном используется редуцированная версия системного макровектора (В.1), которая имеет представление

$$\text{ОУ} = \{U, X, Y, T, \lambda, \delta\}. \tag{В.2}$$

Компоненты редуцированной версии макровектора (В.2) имеют тот же, что и в (В.1) смысл. Форма (В.2) представления математических моделей объектов управления непрерывной и дискретной природы в учебном пособии будет основной.

Прежде, чем формулировать проблемы управления, необходимо отметить следующее. Любая техническая антропогенная система, то есть система, созданная умом и руками человека, имеет три фазы своего существования. Первой фазой является *фаза разработки*, включающая в себя построение математической модели объекта управления и среды его функционирования, аналитический синтез закона управления, построение алгоритмического обеспечения процедур оценки параметров модели объекта и его состояния, моделирование системы с использованием возможностей современных программных оболочек, разработка технической реализации (программной – SOFT и схемотехнической – HARD) всех компонентов процесса управления, разработка конструкции устройства управления и технологического сопровождения его изготовления и испытания макетного образца устройства управления с использованием стендовых испытательных средств. Второй фазой существования технической системы является *фаза изготовления*, третьей – *фаза эксплуатации*.

Проблемы управления в своей алгоритмической основе решаются в фазе разработки, а реализуются в фазе эксплуатации. Это значит, что математическая постановка задачи (цели) управления должна быть корректно сформулирована, математические модели объекта управления и среды его функционирования должны быть адекватны реальным физическим процессам в них, параметры математических моделей объекта и окружающей среды должны быть оценены с допустимой погрешностью, оценка вектора состояния должна сходиться к вектору состояния, сформированный закон управления должен доставлять процессу управления объектом требуемые динамические качества с одновременным обеспечением стабильности потребительских свойств в условиях возможной параметрической неопределенности, при этом канальная среда в прямом канале должна передавать достоверно сигналы управления к регулирующим органам объекта, а в обратном канале – достоверно передавать информацию о доступных непосредственному измерению компонентах вектора состояния и выхода в устройство управления. Все алгоритмы, задействованные в процессе управления должны быть вычислительно устойчивыми, матричные компоненты используемых математических модельных представлений должны быть хорошо обусловлены.

Приведенная на рисунке В.1 структурная схема сложной системы, представляющей собой функциональное объединение объекта управления, устройства управления и канальной среды, а также сделанный к ней комментарий позволяют сформулировать основные *проблемы управления*.

Первой проблемой является проблема составления математической модели ОУ в форме (В.1) или (В.2), причем ключевыми моментами здесь оказываются назначение *разумной* размерности вектора состояния, а также аналитические представления правил λ и δ. Первая проблема в основном решается *экспертным* образом, который опирается на библиографические источники, опыт специалистов и собственный опыт разработчика.

Второй проблемой является решение задачи идентификации объекта управления, которая сводится при сконструированных аналитических представлениях правил λ и δ к разработке и реализации алгоритма ς формирования оценок $\hat{p}_\lambda$ и $\hat{p}_\delta$ параметров p_λ и p_δ этих правил на основе результатов измерения доступных непосредственному измерению компонентов x_u и y_u векторов x состояния и y выхода ОУ, причем алгоритм должен гарантировать сходимость оценок параметров в форме

$$\varsigma:\{x_u, y_u\} \Rightarrow (\hat{p}_\lambda, \hat{p}_\delta): \lim_{t\to\infty}(\hat{p}_\lambda, \hat{p}_\delta) = (p_\lambda, p_\delta) \tag{В.3}$$

Третьей проблемой является решение задачи оценки состояния объекта, которая сводится к разработке и реализации алгоритма ξ формирования оценки $\hat{x}$ вектора состояния x на основе результатов измерения доступных непосредственному измерению компонентов x_u и y_u векторов x состояния и y выхода ОУ, причем алгоритм должен гарантировать сходимость оценки вектора состояния в форме

$$\xi:\{x_u, y_u\} \Rightarrow \hat{x}: \lim_{t\to\infty}\hat{x}(t) = x(t) \tag{В.4}$$

Четвертой проблемой является решение задачи формирования закона управления, которое является многофазным.

Первая фаза решения состоит в формализации *задачи (цели) управления.* При всем многообразии содержательных постановок задач (целей) управления в формализованном представлении они могут быть сведены к двум версиям. Первая версия, именуемая *задачей перевода (регулирования)*, формулируется следующим образом: *перевести объект управления, находящийся в начальный момент времени $t=t_0$ в состоянии $x(t_0)$, к моменту времени $t=t_k$ в требуемое состояние $x(t_k)$ за минимально возможный на множестве доступных управлений Ω_U промежуток времени $\Delta T = t_k - t_0$*, формализованное представление *задачи перевода (регулирования)* имеет вид

$$x(t=t_0) \Rightarrow x(t=t_k): \Delta T = (t_k - t_0) = \min_{U\in\Omega_U} \tag{В.5}$$

Вторая версия задачи (цели) управления, именуемая *задачей удержания (слежения),* формулируется следующим образом: *удерживать состояние объекта управления $x(t)$ на программной траектории $x_{пр}(t)$ с минимальной на множестве доступных управлений Ω_U нормой вектора ошибки этого удержания,* формализованное представление *задачи удержания (слежения)* принимает вид

$$\left\| x_{np}(t) - x(t) \right\| = \min_{U \in \Omega_U} \tag{B.6}$$

Вторая фаза решения задачи формирования закона управления состоит в формировании *показателя (критерия) качества* протекания управляемого процесса, сформулированного в одной из *постановочных версий.* Показатель качества $J = J(x,u)$ задаётся так, чтобы траекториям управляемого процесса *лучшего качества* соответствовало экстремальное на множествах допустимых управлений Ω_U и допустимых траекторий Ω_X значение $\underset{U \in \Omega_U, X \in \Omega_X}{\text{extrem}} \{J = J(x,u)\}$ этого показателя.

Последняя (финальная) фаза формирования закона управления состоит в формировании сигнала управления как *функции текущего состояния объекта управления,* а в случае непосредственной неизмеримости вектора состояния *его оценки, а также оценки параметров* правил λ и δ его модели так, что *закон управления* принимает аналитическое представление

$$U = U\{\hat{x}, \hat{p}_\lambda, \hat{p}_\delta\} : \begin{Bmatrix} x(t_0) \Rightarrow x(t_\kappa) : \Delta T = (t_\kappa - t_0) = \min\limits_{U \in \Omega_U}; \vee \\ \left\| x_{np}(t) - x(t) \right\| = \min\limits_{U \in \Omega_U}; \& \, \mathrm{J} = \underset{\mathrm{U} \in \Omega_\mathrm{U}, X \in \Omega_X}{\text{extrem}} \{J(x,u)\} \end{Bmatrix} \tag{B.7}$$

Пятой проблемой является проблема канализации информации по прямому каналу связи (КС1) от устройства управления к объекту и по обратному каналу связи (КС2) – от объекта управления к устройству управления. Содержательно проблема канализации информации, как в прямом, так и в обратном каналах сводятся к решению двух задач. *Первая задача* связана с требованием эффективного использования предоставленного канала связи. В вербальной форме задача может быть сформулирована следующим образом: *передачу информации по предоставленному каналу связи следует вести так, чтобы объем сигнала* (V_c) *не превышал емкости* (V_κ) *канала связи, максимально приближаясь к выполнению равенства* $V_c = V_\kappa$, *где*

$$V_c = T_c F_c \log_2 (1 + \frac{P_c}{P_n}); \; V_\kappa = T_\kappa F_\kappa \log_2 (1 + \frac{P_c}{P_n}). \tag{B.8}$$

В формуле (В.8): T_c – временная длительность сигнала, F_c – эффективный спектр сигнала, P_c – мощность сигнала, P_n – мощность помехи, сопровождающей процесс формирования сигнала; T_κ – длительность интервала времени, на который предоставлен канал связи, F_κ – эффективная полоса пропускания канала связи, $\mathrm{P_c}$ – мощность сигнала, фиксируемая в канальной среде, P_n – мощность помехи в канальной среде.

Вторая задача канализации информации связана с удовлетворением требованиям обеспечения достоверности принимаемой информации *(информационной надежности каналообразующих средств).* В вербальной форме задача может быть сформулирована следующим образом: *передачу информации по предоставленному каналу связи в условиях помех следует организовать так, чтобы за счет введения в структуру передаваемых кодов, несущих необходимую получателю информацию, избыточных разрядов, на приемной стороне существовала возможность восстановления искаженного при передаче кода в такой степени, чтобы вероятность $P_{ош}$ исполнения искаженной (ошибочной, ложной) команды не превышала бы вероятности $P_{доп}$, допустимой для данной категории проектируемой системы управления.* В формальной постановке задача обеспечения информационной надежности канальными средствами сводится к обеспечению неравенства

$$P_{ош} = \sum_{i=s+1}^{n} C_n^i p^i (1-p)^{n-i} \le P_{доп}, \qquad (В.9)$$

где n – число разрядов помехозащищенного кода, p – вероятность искажения элементарного сигнала (бита) двоичного кода, s – число исправляемых ошибок средствами помехозащитного декодирования при приеме информации, i – число возможных ошибок, C_n^i – число сочетаний из n по i.

Перечисленные проблемы управления относятся к разряду *«вечных»,* содержательно они инвариантны относительно технологической среды, в которой пребывает конкретное гуманитарное сообщество.

Первая из перечисленных выше проблем является предметной областью дисциплины «Математические основы теории систем». На математическое сопровождение решения перечисленных выше базовых задач управления в их модельном представлении направлено основное содержание учебного пособия. В этой связи в пособии приведены сведения об алгебраических структурах, основных пространствах, матричном формализме, являющемся инструментальной основой метода пространства состояния. Модельные представления динамических объектов (объектов управления) как в классе моделей «вход–состояние–выход (ВСВ)», так и в классе моделей «вход–выход (ВВ)» ограничиваются непрерывными и дискретными по времени объектами. С использованием возможностей ВСВ – модельных представлений решаются задачи анализа структурных свойств динамических объектов – управляемости и наблюдаемости. Проблема конечномерных представлений сигналов, как элементов функционального пространства, решается как в прямой постановке с использованием матрицы Грама, так и в обратной – с использованием теоремы В. Котельникова–К. Шеннона. Освоение основных положений *математических основ теории систем* сопровождено богатым практикумом по базовым проблемам курса.

Вторая, третья и четвертая проблемы являются предметной областью «куста» дисциплин, объединенных названием «Современная теория управления».

Пятая из перечисленных проблем является предметной областью дисциплины «Прикладная теория информации».

1. АЛГЕБРАИЧЕСКИЕ СТРУКТУРЫ

Для изучения основных разделов математических основ теории систем (МОТС) необходимо определенное знакомство с алгебраическими структурами и пространствами. Схема их формирования, взаимной связи и изучения приведена на рисунке 1.1.

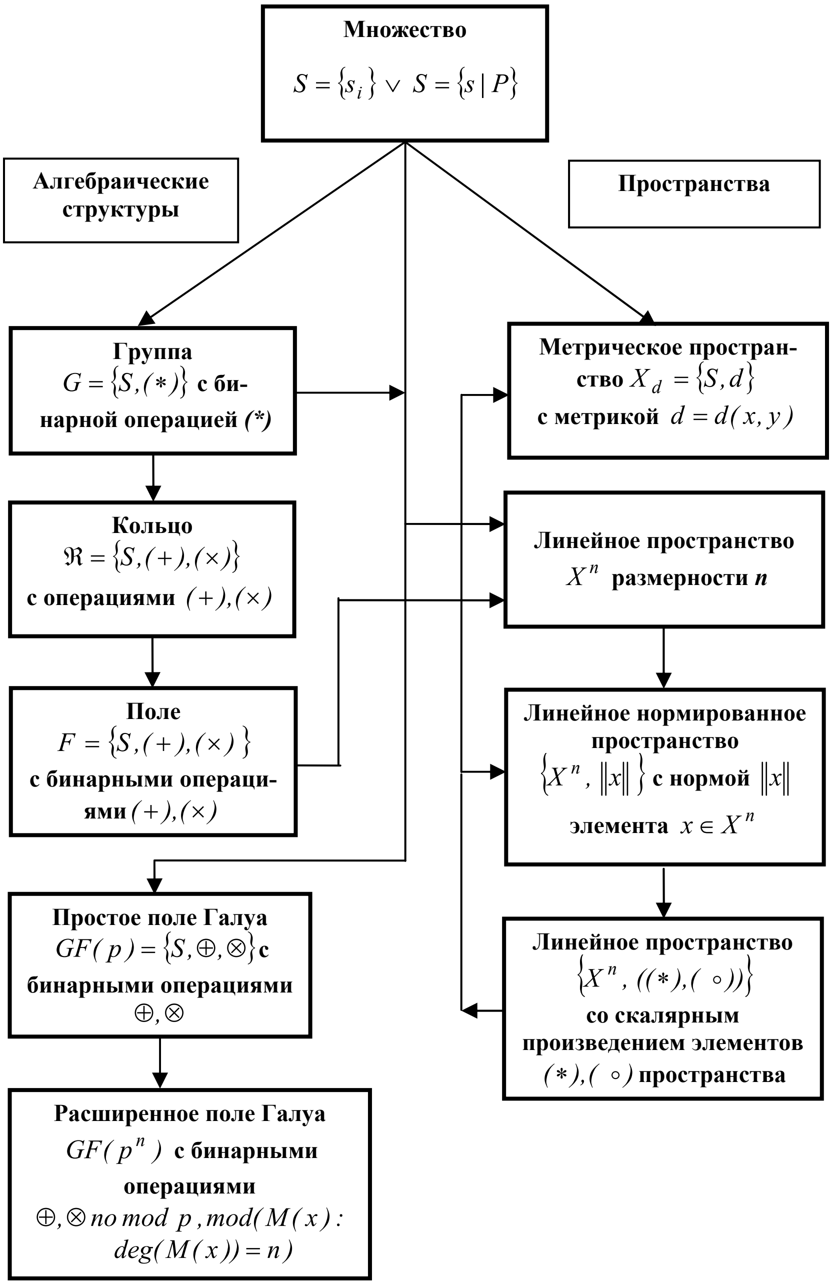

Рисунок 1.1

Определение 1.1 (О1.1). *Множеством* называется совокупность объектов любой природы, задаваемая путем их перечисления или указанием их характеристического свойства P:

$$S = \{s_1, s_2, s_3\} \quad \vee \quad S = \{s \mid P\}.$$

Последняя запись означает, что множество S есть совокупность элементов s, обладающих характеристическим свойством P.

Мощность множества $S=\{s_1,s_2,...,s_n\}$ характеризует число элементов множества и обозначается $[S]$.

Определение 1.2 (О1.2). Пусть задано множество $S=\{s_1,s_2,...,s_n\}$, будем говорить, что во множестве S определена бинарная алгебраическая операция, если указано правило, по которому любой паре элементов s_i, s_j из этого множества, взятых в определенном порядке ставится в соответствие единственный элемент s_k того же множества.

Определение 1.3 (О1.3). *Алгебраической структурой* называется множество с заданными в нем алгебраической бинарной операцией (или несколькими операциями) и свойствами элементов относительно этой бинарной операции.

В курсе МОТС изучаются следующие алгебраические структуры: группа (подгруппа), кольцо, идеал, поле, простое и расширенное поля Галуа.

Определение 1.4 (О1.4). Множество G называются *группой*, если для любой пары элементов множества G определена бинарная алгебраическая операция $*$ и выполняются условия:

1. Замкнутости: для $\forall\alpha,\beta\in G$, элемент $\gamma=\alpha*\beta\in G$;
2. Ассоциативности: для $\forall\alpha,\beta,\gamma\in G \quad \alpha*(\beta*\gamma)=(\alpha*\beta)*\gamma$
3. Существования нейтрального элемента (единицы группы): G содержит единственный элемент $e: \forall\alpha\in G \quad e*\alpha=\alpha*e=\alpha;$
4. Существования обратного элемента: для $\forall\alpha\in G \qquad \exists\alpha^{-1}\in G$ (единственный для $\forall\alpha$), называемый элементом, обратным α, такой, что $\alpha*\alpha^{-1}=\alpha^{-1}*\alpha=e$.

Примечание 1.1 (П1.1): Группа G называется *коммутативной* или абелевой группой, если выполняется условие коммутативности: для $\forall\alpha,\beta\in G \quad \alpha*\beta=\beta*\alpha$.

Определение 1.5 (О1.5). Подмножество G_0 группы $G(G_0\subset G)$ называется *подгруппой*, если оно удовлетворяет всем свойствам группы относительно бинарной алгебраической операции *.

Определение 1.6 (О1.6). Пусть G – коммутативная группа, G_0 – подгруппа группы G. Рассмотрим множество:

$$G_0=\{\alpha_1 \quad \alpha_2 \quad \alpha_3 \quad \dots \quad \alpha_k\},$$

$$G_1=\{\alpha_1*\beta_1 \quad \alpha_2*\beta_1 \quad \alpha_3*\beta_1 \quad \dots \quad \alpha_k*\beta_1\}$$

$$\vdots$$

$$G_\nu=\{\alpha_1*\beta_\nu \quad \alpha_2*\beta_\nu \quad \alpha_3*\beta_\nu \quad \dots \quad \alpha_k*\beta_\nu\}$$

где $\beta_i\notin G_0,G_{i-1}, \quad i=\overline{1,\nu}$.

Определенные таким образом множества называются смежными классами группы G по подгруппе G_0 и задают разложение группы G по подгруппе G_0 с образующими элементами β_i, так что $G = \bigcup\limits_{i=0}^{\nu} G_i$, где число ν называется индексом подгруппы G_0 в группе G.

Определение 1.7 (О1.7). Пусть имеются две группы G_1 и G_2 с бинарными операциями «$*$» и «$\circ$» соответственно одной и той же мощности и $\varphi : G_1 \to G_2$ отображение G_1 в G_2 такое, что для всех $x, y \in G_1$ имеет место равенство: $\varphi(x * y) = \varphi(x) \circ \varphi(y)$.

Отображение φ, обладающее таким свойством, называется *изоморфным*. Если между двумя группами G_1 и G_2 можно установить изоморфизм φ, то группы G_1 и G_2 называются *изоморфными*.

Определение 1.8 (О1.8). Множество R называется *кольцом*, если на нем определены бинарные алгебраические операции сложения и умножения и выполняются следующие условия:

1. Множество R является коммутативной группой относительно бинарной операции сложения;
2. Замкнутости относительно бинарной операции умножения: для $\forall \alpha, \beta \in R$ элемент $\alpha\beta \in R$;
3. Ассоциативности относительно бинарной операции умножения: для $\forall \alpha, \beta, \gamma \in R \quad \alpha(\beta\gamma) = (\alpha\beta)\gamma$;
4. Дистрибутивности относительно бинарных операций сложения и умножения: для

$$\forall \alpha, \beta, \gamma \in R \qquad \alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma, \qquad (\beta + \gamma)\alpha = \beta\alpha + \gamma\alpha .$$

Примечание 1.2 (П1.2). Кольцо R называется *коммутативным*, если выполняется условие: $\forall \alpha, \beta \in R \quad \alpha\beta = \beta\alpha$.

Определение 1.9 (О1.9). Подгруппа J аддитивной группы R называется *идеалом*, если для $\forall \alpha \in R$ и $\beta \in J$ элемент $\alpha\beta \in J$.

Примечание 1.3 (П1.3). Идеал, состоящий из всех элементов, кратных некоторому элементу α кольца R, называется *главным* идеалом. Кольцо, в котором каждый идеал главный, называется *кольцом главных идеалов*. Элемент α называется *образующим* (или *порождающим*) элементом идеала.

Поскольку для кольца R справедливы все свойства группы, а для идеала J все свойства подгруппы относительно бинарной операции сложения, то кольцо R можно разложить подобно группе на смежные классы по идеалу J.

Определение 1.10 (О1.10). Коммутативное кольцо F называется *полем*, если выполняются следующие условия:

1. Кольцо F содержит нейтральный элемент 1 относительно бинарной операции умножения такой, что для $\forall \alpha \in F$ $1\alpha = \alpha 1 = \alpha$;
2. Для $\forall \alpha \neq 0$, $\alpha \in F$ существует обратный элемент $\alpha^{-1} \in F : \alpha\alpha^{-1} = \alpha^{-1}\alpha = 1$.
3. Если $\alpha, \beta \in F$, то $\alpha\beta = 0 \Leftrightarrow \alpha = 0$ или $\beta = 0$.

Определение 1.11 (О1.11). Если p – простое число, то кольцо чисел по $\bmod p$ называется *простым полем* Галуа и обозначается $GF(p)$. $GF(p)$ состоит из элементов $0,1,\ldots,p-1$, таким образом $GF(p) = \{0,1,2,\cdots,p-1\}$.

Примечание 1.4 (П1.4). Определим на множестве $GF(p)$ две бинарные операции:

сложения по $\bmod p$, обозначив его $\oplus$, и умножения по $\bmod p$, обозначив его $\otimes$:

1. $c = a \oplus b = rest\dfrac{a+b}{p} \Leftrightarrow a + b = pm + c,$ где $c < p, m$ – целое;
2. $d = a \otimes b = rest\dfrac{ab}{p} \Leftrightarrow ab = pk + d$, где $d < p, k$ – целое.

При этом c и d называются *вычетами*.

Определим понятия сравнимости по $\bmod p$. Два целых числа a и b сравнимы по $\bmod p$:

$a \equiv b(\bmod p) \Leftrightarrow a - b = pm,$ где m – целое.

Определение 1.12 (О1.12). Полином $A(x) = \sum\limits_{i=0}^{n} a_i x^i$ называется *полиномом над полем* $GF(p)$ *или модулярным*, если его коэффициенты $a_i, i = \overline{0,n}$ принадлежат простому полю $GF(p)$. *Степенью полинома* $A(x)$ $\deg\{A(x)\}$ называется наибольшее число $n : a_n \neq 0$.

Примечание 1.5 (П1.5). Сравнение модулярных полиномов $A(x)$ и $B(x)$ по модулю модулярного полинома $F(x)$, производится аналогично сравнению целых чисел по $\bmod p$

$$A(x) \equiv B(x) \bmod [F(x)] \Leftrightarrow A(x) - B(x) = k(x)F(x),$$

где $\deg[k(x)] \leq \deg[F(x)]$.

Аналогично можно ввести операции суммирования (вычитания), умножения по модулю модулярного полинома, при этом приведение подобных членов производится по $\bmod p$. Так

$A(x)+B(x)=C(x)\text{mod}[F(x)] \Leftrightarrow A(x)+B(x)=L(x)F(x)+C(x)$, где $\deg[C(x)] \le \deg[F(x)]$, при этом $C(x)$ называется *вычетом по* $\text{mod}[F(x)]$.

Определение 1.13. Полная система вычетов по двойному модулю $[\text{mod}\, p, \text{mod}[F(x)]]$ образует конечное поле, содержащее p^n элементов, которое обозначается $GF(p^n)$ и называется *расширенным полем Галуа*.

В отличие от простого поля $GF(p)$ элементами расширенного поля $GF(p^n)$ являются уже не числа, а модулярные полиномы степени не выше $(n-1)$ с коэффициентами из простого поля $GF(p)$.

Примеры и задачи

1.1.* Определить, относительно какой бинарной операции: умножения или сложения следующее числовое множество образует группу, или не образует ее вовсе.

а) Множество всех вещественных чисел R.
б) Множество вещественных чисел отличных от нуля.
в) Множество положительных вещественных чисел.
г) Множество всех комплексных чисел.
д) Множество комплексных чисел, отличных от нуля.
е) Множество комплексных чисел с модулем равным единице.
ж) Множество комплексных чисел с модулем большем единицы.
з) Множество чисел, представляющих собой целые положительные степени числа 2 $\{2,4,8,\cdots\}$.
и) Полное множество чисел $\{1,-1,j,-j,\cdots\}$, где $j=\sqrt{-1}$.

1.2. Указать какие из обнаруженных групп в примере 1.1 являются коммутативными группами.

1.3. Дано множество чисел $S=\{0,1,2,3,4,5,6,7\}$, задана бинарная операция $\oplus$ – сложение по $\text{mod}\, p$. Указать, для каких значений p множество S образует группу.

1.4. Дано множество кодовых комбинаций над простым полем $GF(2)=\{0,1\}$ с бинарной операцией сложения комбинаций по $\text{mod}\, 2$ без переносов в старший разряд $S=\{000,001,010,011,100,101,110,111\}$. Определить, является ли множество S группой, указать какой элемент множества S является единицей группы.

1.5. Дано множество квадратных матриц $\{A_i, i=1,2,\cdots\}$ порядка $n\times n$.

а) Доказать, что множество квадратных матриц A_i образует группу относительно бинарной операции сложения.

б) При каком характеристическом свойстве множество матриц A_i образует группу относительно бинарной операции умножения.

1.6. Дано множество n–мерных векторов $X = col\{x_i, i = \overline{1,n}\}, x_i \in R^n$. Выяснить, образует ли множество X группу относительно бинарной операции сложения.

1.7. Дано множество матриц перестановок P_i порядка 3×3:

$$\left\{ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \right\},$$

используемых при перестановках элементов векторов в соответствии с правилами $\overline{x}_i = P_i x$, где $x = (x_1, x_2, x_3)^T$. Указать, относительно какой бинарной операции умножения или сложения множество матриц перестановок образует группу.

1.8. Выяснить, изоморфны ли следующие группы:

а) Вещественных чисел с бинарной операцией сложения и степеней любого положительного целого числа с бинарной операцией умножения.

б) Комплексных чисел с бинарной операцией умножения и натуральных логарифмов комплексных чисел с бинарной операцией сложения.

1.9. Дано множество кодовых комбинаций из элементов $GF(2) = \{0,1\}$ с бинарной операцией сложения по $\bmod 2$ без переноса в старший разряд, образующее группу

$$G = \left\{ \begin{matrix} 0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1000, 1001, 1010, \\ 1011, 1100, 1101, 1110, 1111 \end{matrix} \right\}.$$

а) Указать все возможные подгруппы G_i группы G.

б) Построить разложение группы G по подгруппе $\widetilde{G}$, содержащей минимальное число элементов.

1.10. Дано множество кодовых комбинаций, образующих группу относительно бинарной операции сложения по $\bmod p$ без переноса в старший разряд

$$G = \{x_7, x_6, x_5, x_4, x_3, x_2, x_1\}, \text{ где } x_i \in GF(2) = \{0,1\}, i = \overline{1,7}.$$

а) Доказать, что подмножество кодовых комбинаций, элементы которых удовлетворяют соотношениям:

$x_1 = x_3 \oplus x_5 \oplus x_7$, $x_2 = x_3 \oplus x_6 \oplus x_7$, $x_4 = x_5 \oplus x_6 \oplus x_7$ образует подгруппу G_0.

б) Разложить группу G по подгруппе G_0, взяв в качестве образующих элементов кодовые комбинации

$\{0000001\}, \{0000010\}, \{0000100\}, \{0001000\}, \{0010000\}, \{0100000\}, \{1000000\}$

1.11. * Выяснить, образуют ли кольцо следующие множества:

а) Множество четных чисел;

б) Множество нечетных чисел;

в) Множество рациональных чисел;

г) Множество вещественных чисел;

д) Множество комплексных чисел;

е) Множество целых чисел, кратных данному целому числу большему $q > 0$;

ж) Множество чисел вида $a + b\sqrt{2}$, где a и b целые.

1.12. Выяснить, образуют ли кольцо множество квадратных матриц одной размерности.

1.13. Выяснить, образуют ли кольцо множество:

а) Многочленов $f(x) = \sum\limits_{i=0}^{n} a_i x^{n-i}$ где a_i – целые числа $i = \overline{0,4}$.

б) Линейных комбинаций экспоненциальных функций $g(t) = \sum\limits_{i=0}^{n} g_i e^{it}$.

в) Линейных комбинаций гармонических функций $\varphi(t) = \sum\limits_{i=0}^{n} \varphi_i \cos(it)$, где φ_i – действительные числа.

г) Функций $\xi(t) \in L_T^2$, где $t \in [0, T]$ и $\int\limits_T \xi^2(t) dt < M = const$

1.14. Доказать, что множество парных чисел (a, b) с бинарными операциями, заданными равенствами

$$(a_1, b_1) + (a_2, b_2) = (a_1 + a_2, b_1 + b_2) \quad \text{и} \quad (a_1, b_1)(a_2, b_2) = (a_1 a_2, b_1 b_2)$$

образует кольцо.

1.15. Выяснить, образует ли множество $\{0,1,2,3,4,5\}$ кольцо с бинарными операциями сложения и умножения по $\bmod p$:

а) При $p = 2$;

б) При $p = 3$;

в) При $p = 4$;

г) При $p = 8$.

1.16. Выяснить, образует ли множество многочленов $f(x) = \sum\limits_{i=0}^{n} a_i x^{n-i}$, где $a_i \in GF(p) = \{0,1,2,\cdots, p-1\}$ с бинарными операциями умножения и сложения по двойному модулю $\bmod p$ и $\bmod(x^m + 1)$ кольцо многочленов.

1.17. Дано кольцо многочленов $f(x)=\sum_{i=0}^{n} a_i x^{n-i}$, где $a_i \in GF(2)=\{0,1\}$ с бинарными операциями умножения и сложения по $\mathrm{mod}\,2$ и $\mathrm{mod}\left(x^7+1\right)$. Построить идеалы:

а) J_1 с образующим элементом $g_1(x)=x+1$.

б) J_2 с образующим элементом $g_2(x)=x^3+x+1$.

в) J_3 с образующим элементом $g_3(x)=x^3+x^2+1$.

1.18. В предыдущей задаче, пользуясь многочленами степени меньшей $\deg\{g_i(x)\}$ как элементами, разложить кольцо многочленов на смежные классы по:

а) Идеалу J_1. Если $\alpha,\beta \in F$, то $\alpha\beta=0 \Leftrightarrow \alpha=0$ или $\beta=0$. Если $\alpha,\beta \in F$, то $\alpha\beta=0 \Leftrightarrow \alpha=0$ или $\beta=0$.

б) Идеалу J_2.

в) Идеалу J_3.

1.19. * Выяснить, какие из множеств примера 1.11 образуют поле.

1.20. Выяснить, каким свойством должны обладать квадратные матрицы, чтобы их множество образовало поле.

1.21. Дано множество $\{0,1,2,3,4\}$ с бинарными операциями сложения и умножения по $\mathrm{mod}\,5$. Образует ли это множество поле? Как в этом поле осуществляется обратный элемент для каждого элемента множества?

Примеры решения вариантов задач

Решение задачи 1.1. В соответствие с определением группы G при заданной бинарной операции необходимо выполнение условий аксиом:

- замкнутости;
- ассоциативности;
- существования нейтрального элемента (единицы группы);
- существования обратного элемента.

Тогда имеем в вариантах задачи:

а) Группу относительно сложения и умножения. Действительно:

$\forall \alpha,\beta \in R, \Rightarrow \alpha+\beta \in R, \alpha\beta \in R$

$\forall \alpha,\beta,\gamma \Rightarrow \alpha+(\beta+\gamma)=(\alpha+\beta)+\gamma, \alpha(\beta\gamma)=(\alpha\beta)\gamma$

– нейтральный элемент $e=0$ относительно сложения и $e=1$ относительно умножения;

– обратный элемент α^{-1} $\forall \alpha \in R$ $\alpha^{-1}=-\alpha$ относительно сложения и $\alpha^{-1}=1/\alpha$ относительно умножения, при этом $-\alpha,\ 1/\alpha \in R$.

б) Группу относительно умножения.

в) Группу относительно умножения.

г) Группу относительно сложения и умножения. Действительно: $\forall \alpha, \beta \in C$, где C – множество комплексных чисел, воспользуемся декартовой и полярной формами представления чисел, так, что

$$\alpha = |\alpha| e^{j\varphi_\alpha}, \beta = |\beta| e^{j\varphi_\beta}, \alpha = \mathrm{Re}\,\alpha + jJm\alpha, \beta = \mathrm{Re}\,\beta + jJm\beta;$$

$$\alpha + \beta = (\mathrm{Re}\,\alpha + \mathrm{Re}\,\beta) + j(Jm\alpha + Jm\beta) = (\mathrm{Re}\,\gamma + jJm\gamma) \in C;$$

$$\alpha\beta = |\alpha| e^{j\varphi_\alpha} \cdot |\beta| e^{j\varphi_\beta} = |\alpha| \cdot |\beta| e^{j(\varphi_\alpha + \varphi_\beta)} = |\delta| e^{j\varphi_\delta} \in C;$$

$$\forall \alpha, \beta, \chi \in C \Rightarrow \alpha + (\beta + \chi) = (\alpha + \beta) + \chi, \alpha \cdot (\beta \cdot \chi) = (\alpha \cdot \beta) \cdot \chi.$$

– нейтральный элемент $e = 0 + j0 = 0$ относительно сложения и $e = 1e^{j0} = 1$ относительно умножения:

– обратный элемент $\alpha^{-1} \forall \alpha \in C$ равен: $-\alpha = -\mathrm{Re}\,\alpha - jJm\alpha$ относительно сложения и $\dfrac{1}{\alpha} = \dfrac{1}{|\alpha|} \cdot e^{-j\varphi_\alpha}$ относительно умножения.

д) Группу относительно умножения.

е) Группу относительно умножения.

ж) Не образует группы ни относительно сложения, ни относительно умножения, так как не содержит единиц группы.

з) Не образует группы в обоих случаях, так как относительно сложений не включаются условия замкнутости, относительно умножения не существует α^{-1}.

и) Образует группы относительно умножения.

<u>Решение задачи 1.11</u>. В соответствии с определением 1.8 кольца *R* на элементах кольца должны быть заданы бинарные операции сложения и умножения и выполняться условия (аксиомы):

– коммутативности *R* как группы относительно сложения
– замкнутости относительно умножения
– ассоциативности относительно умножения
– дистрибутивности.

Тогда имеем в вариантах задачи:

а) Образует кольцо. Действительно, множество четных чисел X образует коммутативную группу относительно сложения

$$\forall \alpha, \beta \in X \Rightarrow \alpha\beta \in X$$

$$\forall \alpha, \beta, \gamma \in X \Rightarrow \alpha(\beta\gamma) = (\alpha\beta)\gamma$$

$$\forall \alpha, \beta, \gamma \in X \Rightarrow \alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma, (\beta + \gamma)\alpha = \beta\alpha + \gamma\alpha$$

б) Не образует кольцо, так как не выполняются групповые свойства множества нечетных чисел.

в) Образуют кольцо.

г) Образуют кольцо.

д) Образуют кольцо.

е) Если считать, что ноль кратен целому числу q, то это множество образует кольцо.

ж) Не образует кольцо, так как не выполняются условия замкнутости по умножению. Действительно, пусть $\alpha = a + b\sqrt{2},\ \beta = a - b\sqrt{2}$ где a, b целые, тогда $\gamma = \alpha\beta = a^2 - b^2 = a^2 - \frac{b^2}{\sqrt{2}} \times \sqrt{2}$, где $\frac{b^2}{\sqrt{2}}$ перестает быть целым.

Решение задачи 1.19. Напомним, что для того чтобы кольцо R образовывало поле F необходимо выполнение следующих условий:

– коммутативности кольца,

– существования единиц $1: \forall \alpha \in F$, $\alpha \cdot 1 = 1 \cdot \alpha = \alpha$,

– существования обратного элемента $\alpha^{-1}: \forall \alpha \in F$ $\alpha \cdot \alpha^{-1} = \alpha^{-1} \cdot \alpha = 1$.

Тогда поля будут образовывать множества в задачах в 1.11 в), д), е).

2. ПРОСТРАНСТВА

Определение 2.1 (О2.1). Пространством называется множество объектов математической природы (точка, кривая, вектор, матрица, геометрическая фигура, многообразие и т.д.), именуемых элементами пространства, на которых заданы геометрические характеристики, определяющие расстояние между элементами, их размер, взаимное положение и т.д.

2.1. Метрические пространства. Примеры метрик

Определение 2.2 (О2.2). Пусть произвольные элементы x и y множества X образуют пару $\{x,y\}$, тогда отображение

$$d:\{x,y\}\Rightarrow R \tag{2.1}$$

во множество действительных чисел R, называется *метрикой* и обозначается $d(x,y)$, если оно удовлетворяет:

1. условию неотрицательности:
   $d(x,y)>0$ *для* $\forall\, x,y: x\neq y; d(x,y)=0$ *для* $\forall\, x,y: x\equiv y$;
2. условию симметрии: $d(x,y)=d(y,x)$ *для* $\forall\, x,y$;
3. условию неравенства «треугольника»:
   $d(x,z)\le d(x,y)+d(y,z)$ *для* $x,y,z.$

Содержательно, метрика представляет собой вещественнозначную положительную величину, определяющую *расстояние* между элементами или *степень различия* элементов множества X.

Определение 2.3 (О2.3). Множество X с введенной в нем метрикой $d=d(x,y)$ образует метрическое пространство (МП), обозначаемое в одной из форм $\{X,d\}$ *или* X_d.

Примечание 2.1 (П2.1). Так как на элементах множества X может быть задано бесконечное число метрик $d=d(x,y)$, то на нем может быть построено бесконечное число метрических пространств $\{X,d\}$.

Рассмотрим примеры метрик и метрических пространств.

1. Если $X=R$ – множество действительных чисел, то R образует метрическое пространство $\{R,d\}$ с метрикой

$$d(x,y)=|x-y|=|y-x|; x,y\in R. \tag{2.2}$$

Эта метрика именуется обычной (простой) или абсолютной на R.

2. Если $X=R^n$, то есть оно образовано n–элементными числовыми массивами (именуемыми также n-ками, n–кортежами, n–векторами), представляемыми в виде

$x=(x_1,x_2..x_n);\ y=(y_1,y_2...y_n);x_i,y_j\in R;\ i,j=\overline{1,n};$ то на множестве X может быть задана обобщенная гёльдеровская векторная метрика $d_p(x,y)$, определяемая выражением

$$d_p(x,y)=(\sum_{i=1}^{n}(|x_i-y_i|)^p)^{1/p} \tag{2.3}$$

где p – целое положительное число. Наиболее употребительными векторными метриками являются:

2.1. абсолютная векторная метрика $d_p(x,y)$ при $p=1$

$$d_1(x,y)=\sum_{i=1}^{n}|x_i-y_i|\ ; \tag{2.4}$$

2.2. квадратичная векторная метрика $d_p(x,y)$ при $p=2$

$$d_2(x,y)=(\sum_{i=1}^{n}(|x_i-y_i|)^2)^{1/2}=(\sum_{i=1}^{n}(x_i-y_i)^2)^{1/2}; \tag{2.5}$$

2.3. экстремальная метрика $d_p(x,y)$ при $p\Rightarrow\infty$

$$d_\infty(x,y)=\lim_{p\to\infty}(\sum_{i=1}^{n}(|x_i-y_i|)^p)^{1/p}=\max_i\left\{|x_i-y_i|;i=\overline{1,n}\right\}; \tag{2.6}$$

3. Если множество X образовано n-ками вида $x=(x_1,x_2...x_n);\ y=(y_1,y_2...y_n);$ где элементы $x_i,y_j;\ i,j=\overline{1,n};$ принадлежат простому полю Галуа $GF(p)=\{\,0,1,2...p-1\}$, то n-ки x и y именуются кодами или кодовыми векторами, при этом на множестве X может быть построена метрика Ли

$$d_L(x,y)=\sum_{i=1}^{n}\min\{|x_i-y_i|,p-|x_i-y_i|\}; \tag{2.7}$$

Если $p=2$ так, что $GF(p)=GF(2)=\{0,1\}$, то метрика Ли вырождается в метрику $d_H(x,y)$ Хэмминга

$$d_L(x,y)=d_H(x,y)=\sum_{i=1}^{n}\{x_i\oplus y_i\}, \tag{2.8}$$

где $\oplus$– знак операции суммирования по модулю два ($\bmod 2$). Содержательно метрика Хэмминга определяет число разрядов кодовых векторов x и y, в которых эти векторы отличаются друг от друга. Метрика $d_H(x,y)$ Хэмминга именуется также *кодовым расстоянием.*

4. Если множество X образовано множеством вещественнозначных функций времени $x(t)$ и $y(t)$, заданных на интервале $T=\{\,t:t_0\le t\le t_\kappa\}$, то на множестве X может быть задана p-ичная функциональная метрика $d_p(x,y)$, определяемая интегральным выра-

жением

$$d_p(x,y)=\left\{\int_{t_0}^{t_\kappa}(|x(t)-y(t)|)^p\,dt\right\}^{1/p}. \tag{2.9}$$

Наиболее употребительными функциональными метриками являются:

4.1. абсолютная функциональная метрика при $p=1$

$$d_p(x,y)|_{p=1}=d_1(x,y)=\int_{t_0}^{t_\kappa}(|x(t)-y(t)|)\,dt; \tag{2.10}$$

4.2. квадратичная функциональная метрика при $p=2$

$$d_p(x,y)=d_2(x,y)=\left\{\int_{t_0}^{t_\kappa}(|x(t)-y(t)|)^2\,dt\right\}^{1/2}; \tag{2.11}$$

4.3. экстремальная функциональная метрика при $p\Rightarrow\infty$

$$d_\infty(x,y)=\underbrace{\lim}_{p\to\infty}d_p(x,y)=\underbrace{\sup}_{t\in T}\{|x(t)-y(t)|\} \tag{2.12}$$

Определение 2.4 (О2.4). Метрическое пространство $\{X,d\}$ называется *сепарабельным,* если для любого $\varepsilon>0$ существует счетная последовательность $\{x_1,x_2...\}$ элементов множества X таких, что $d(x,x_i)<\varepsilon$ для некоторого i и любого $x\in$X.

Определение 2.5 (О2.5). Метрическое пространство $\{X,d\}$ называется *компактным*, если можно найти конечную последовательность $\{x_1,x_2,\ldots,x_{n(\varepsilon)}\}$ элементов множества X таких, что $d(x,x_i)<\varepsilon$ для некоторого $i:1\le i\le n(\varepsilon)$ и любого элемента $x\in X$.

Примеры и задачи

2.1.1. Вычислить векторную метрику $d_1(x,y)$ для векторов $x=(-5,7,-2), y=(5,-7,2)$, элементы которых принадлежат множеству действительных чисел R.

2.1.2. Вычислить векторную метрику $d_2(x,y)$ для векторов $x=(-5,7,-2), y=(5,-7,2)$, элементы которых принадлежат множеству действительных чисел R.

2.1.3. Вычислить векторную метрику $d_\infty(x,y)$ для векторов примера 2.1.1.

2.1.4. Построить кривую зависимости значения векторной метрики $d_p(x,y)$ как функцию от $p\in[1,2,\infty)$ для векторов примера 2.1.1.

2.1.5. Вычислить векторную метрику $d_p(x, y)$ для векторов $x=(-5,7,-2)$, $y=(0,0,0)$ для значений $p=1,2,\infty$.

2.1.6. Построить кривые постоянных значений $d_p(x, y)=1$ для векторов $x=(x_1, x_2)$, $y=(0,0)$ для значений $p=1,2,\infty$.

2.1.7. Вычислить векторную метрику Ли $d_L(x, y)$ для векторов $x=(1,4,2)$, $y=(2,1,3)$, элементы которых принадлежат простому полю Галуа $GF(5)=\{0,1,2,3,4\}$.

2.1.8. Вычислить векторную метрику Ли $d_L(x, y)$ для векторов примера 2.1.7. при условии, что их элементы принадлежат простому полю Галуа $GF(7)=\{0,1,2,3,4,5,6\}$.

2.1.9. Вычислить векторную метрику $d_H(x, y)$ Хэмминга (кодовое расстояние) для кодовых векторов $x=(1011010)$ *и* $y=(0101101)$.

2.1.10. Вычислить функциональную метрику $d_1(x, y)$ для функций $x(t)=1,\ y(t)=t,\ t\in T=[t:0\le t\le 1]$.

2.1.11. Вычислить функциональную метрику $d_2(x, y)$ для функций $x(t)=1,\ y(t)=t,\ t\in T=[t:0\le t\le 1]$.

2.1.12. Вычислить функциональную метрику $d_\infty(x, y)$ для функций $x(t)=1,\ y(t)=t,\ t\in T=[t:0\le t\le 1]$.

2.1.13. Построить кривую зависимости значения функциональной метрики $d_p(x, y)$ как функцию от $p\in[1,2,\infty)$ для функций $x(t)=1,\ y(t)=t, t\in T=[t:0\le t\le 1]$.

2.1.14. Вычислить функциональную метрику $d_p(x, y)$ для функций $x(t)=t,\ y(t)=t^2,\ t\in T=[t:-1\le t\le 1]$ для значений $p=1,2,\infty$.

Решение вариантов задач

Вычислить векторную метрику Ли $d_L(x,y)$ для векторов $x=(1,4,2)$, $y=(2,1,3)$, элементы которых принадлежат простому полю Галуа $GF(5)=\{0,1,2,3,4\}$.

<u>Решение.</u> В силу определения (2.7) метрики Ли можно записать

$$d_L(x,y)=\sum_{i=1}^{n}\min\{|x_i-y_i|,p-|x_i-y_i|\}=\min\{|1-2|;5-|1-2|\}$$
$$+\min\{|4-1|;5-|4-1|\}+\min\{|2-3|;5-|2-3|\}=\min\{1;4\}+$$
$$+\min\{3;2\}+\min\{1;4\}=1+2+1=4.$$

Ответ: $d_L(x,y)=4$.

2.2. Линейные пространства и операторы. Матрицы линейных операторов

Определение 2.6 (О2.6). *Линейным пространством* (ЛП) $\boldsymbol{X}$ над полем F называется аддитивная абелева группа элементов, именуемых векторами (при этом *сумма двух векторов* совпадает с диагональю параллелограмма, стороны которого совпадают с суммируемыми векторами), дополненная бинарной операцией умножения вектора $\boldsymbol{x}\in\boldsymbol{X}$ на скаляр $\alpha \in F$, удовлетворяющей условиям:

1. ассоциативности: $\forall \alpha, \beta \in F, \forall \boldsymbol{x}\in\boldsymbol{X}$ $\alpha(\beta \boldsymbol{x}) = \alpha\beta \boldsymbol{x}$;

2. дистрибутивности: $\forall \alpha, \beta \in F$, $\forall$ $\boldsymbol{x}_1, \boldsymbol{x}_2\in\boldsymbol{X}$ $\alpha(\boldsymbol{x}_1+\boldsymbol{x}_2)=\alpha\, \boldsymbol{x}_1+\alpha\, \boldsymbol{x}_2$; $(\alpha+\beta)\boldsymbol{x}_1=\alpha \boldsymbol{x}_1+\beta \boldsymbol{x}_1$;

3. умножения вектора на (ноль) $0\in F$ и на (единицу) $1\in F$, осуществляемых по правилам $\forall$ $\boldsymbol{x}\in\boldsymbol{X}$: $1\cdot\boldsymbol{x}=\boldsymbol{x}$, $0\cdot\boldsymbol{x}=\boldsymbol{0}$, где $\boldsymbol{0}\in\boldsymbol{X}$, $\boldsymbol{0}$ – ноль-вектор.

Примечание 2.2 (П2.2). Если $F = R$ (R – множество действительных чисел), то ЛП $\boldsymbol{X}$ называется действительным линейным векторным пространством (ДЛП). Если $F = C$ ($C-$множество комплексных чисел), то ЛП $\boldsymbol{X}$ называется комплексным линейным пространством (КЛП).

Определение 2.7 (О2.7). Пусть $\boldsymbol{x}_1$, $\boldsymbol{x}_2 \ldots$ $\boldsymbol{x}_k$ – векторы ЛП $\boldsymbol{X}$ ($\boldsymbol{x}_i\in\boldsymbol{X}$; $i=\overline{1,k}$) над полем F и $\alpha_1, \alpha_2 ... \alpha_k$ – скаляры из F ($\alpha_i \in F; i=\overline{1,k}$)**,** тогда сумма

$$\alpha_1\boldsymbol{x}_1+\alpha_2\boldsymbol{x}_2+\ldots+\alpha_k\boldsymbol{x}_k=\sum_{i=1}^{k}\alpha_i\,\boldsymbol{x}_i \tag{2.13}$$

именуется *линейной комбинацией* из k векторов $\boldsymbol{x}_i\in\boldsymbol{X}$ ($i=\overline{1,k}$) с коэффициентами $\alpha_i \in F$ ($i=\overline{1,k}$), которая задает некоторый вектор, принадлежащий ЛП $\boldsymbol{X}$.

Определение 2.8 (О2.8). Пусть вектора $\boldsymbol{x}_i\in\boldsymbol{X}$ ($i=\overline{1,k}$), тогда *линейной оболочкой* $\boldsymbol{L}\{\, \boldsymbol{x}_i\, ;\ i=\overline{1,k}\,\}$ над полем F, натянутой на вектора $\boldsymbol{x}_i\in\boldsymbol{X}$ ($i=\overline{1,k}$) называется множество линейных комбинаций из векторов $\boldsymbol{x}_i\in\boldsymbol{X}$ ($i=\overline{1,k}$) с коэффициентами $\alpha_i \in F$ ($i=\overline{1,k}$) (2.9) так, что

$$\boldsymbol{L}\{\, \boldsymbol{x}_i\, ;\ i=\overline{1,k}\,\}=\{\sum_{i=1}^{k}\alpha_i\ \boldsymbol{x}_i\, ;\alpha_i\in F; i=\overline{1,k}\,\}. \tag{2.14}$$

Определение 2.9 (О2.9). Множество векторов ($\boldsymbol{x}_i\in\boldsymbol{X}$ **;** $i=\overline{1,k}$) ЛП $\boldsymbol{X}$ называется *линейно независимым*, если равенство над полем F

$$\sum_{i=1}^{k}\alpha_i\,\boldsymbol{x}_i=0\,. \tag{2.15}$$

возможно лишь при всех $\alpha_i = 0; i = \overline{1,k}.$

Определение 2.10 (О2.10). Множество векторов ($\boldsymbol{x}_i \in \boldsymbol{X}$ **;** $i = \overline{1,k}$) ЛП $\boldsymbol{X}$ называется *линейно зависимым* над полем F, если существуют $\alpha_i \in F; i = \overline{1,k}$, одновременно *не равные нулю*, для которых оказывается справедливым равенство $\sum\limits_{i=1}^{k} \alpha_i \boldsymbol{x}_i = 0$.

Определение 2.11 (О2.11). Если ЛП $\boldsymbol{X}$ представимо линейной оболочкой, натянутой на n линейно независимых векторов ($\boldsymbol{x}_i \in \boldsymbol{X}$ **;** $i = \overline{1,n}$) так, что любое множество из $(n+1)$-го векторов оказывается линейно зависимым, то число n называется *размерностью* пространства $\boldsymbol{X}$, обозначаемой $\dim \boldsymbol{X}$, при этом становится справедливой запись

$$\boldsymbol{X} = \{ \sum_{i=1}^{n} \alpha_i \boldsymbol{x}_i; \alpha_i \in F; i = \overline{1,n}; \boldsymbol{x}_i \in \boldsymbol{X} \},\ \dim \boldsymbol{X} = n. \quad (2.16)$$

Примечание 2.3 (П2.3). В случаях, когда необходимо указывать размерность ЛП $\boldsymbol{X}$ $\dim \boldsymbol{X} = n$, может быть использовано обозначение ЛП в форме $\boldsymbol{X}^n$.Причем, если ЛП $\boldsymbol{X}$ – вещественное, то вместо $\boldsymbol{X}^n$ используется запись R^n; если же ЛП $\boldsymbol{X}$ – комплексное, то- C^n.

Примечание 2.4 (П2.4). Линейное пространство $\boldsymbol{X}^n$ называется конечномерным, если $n < \infty$, и бесконечномерным – в противном случае.

Определение 2.12 (О2.12). Пусть $\dim \boldsymbol{X} = n$, тогда любая система (набор) $\{ e_i \in \boldsymbol{X}\ ; i = \overline{1,n} \}$ *линейно независимых* векторов e_i образует базис e.

Утверждение 2.1 (У2,1). Пусть $e = \{e_1, e_2, ... e_n\}$ базис ЛП $\boldsymbol{X}$ над полем F, тогда любой вектор $\boldsymbol{x} \in \boldsymbol{X}$ представим в форме

$$\boldsymbol{x} = x_1 e_1 + x_2 e_2 + ... + x_n e_n; x_i \in F; i = \overline{1,n} \quad (2.17)$$

При этом представление (2.17) в базисе e над полем F *единственное.*

Доказательство справедливости сформулированного утверждения опирается на факт линейной независимости элементов базиса e. Действительно, допустим, что существует альтернативное представление вектора X в базисе e, задаваемое в форме

$$\boldsymbol{x} = \kappa_1 e_1 + \kappa_2 e_2 + ... + \kappa_n e_n; \kappa_i \in F; i = \overline{1,n}. \quad (2.18)$$

Вычтем равенство (2.18) из (2.17), тогда получим

$$(\kappa_1 - x_1) e_1 + (\kappa_2 - x_2) e_2 + ... + (\kappa_n - x_n) e_n = 0. \quad (2.19)$$

Введем обозначение $\alpha_i = \kappa_i - x_i; i = \overline{1,n}$, тогда (2.19) можно записать в форме (2.14), которое для случая линейно независимых элементов $e_i\, i = \overline{1,n}$ выполняется при всех $\alpha_i = \kappa_i - x_i = 0; i = \overline{1,n}$. Таким образом $\kappa_i = x_i; i = \overline{1,n}.$ ■

Примечание 2.5 (П2.5). Вектор $\boldsymbol{x}$ именуется *бескоординатным* вектором ЛП $\boldsymbol{X}$. Вектор x, составленный из чисел $x_i \; i=\overline{1,n}$, именуемых его *координатами* в базисе e, и сформированный в виде столбца

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}^T = col\{x_i; i=\overline{1,n}\}, \tag{2.20}$$

называется *координатным* вектором, при этом следует помнить о функциональной связи

$$x = x(e).$$

Пространство X^n, составленное из координатных векторов x (2.20), именуется *арифметическим* линейным пространством.

Определение 2.13 (О2.13). Линейная оболочка

$$\boldsymbol{P} \;=\; \boldsymbol{L}\{\boldsymbol{x}_i; i=\overline{1,m};\; m<n=\dim \boldsymbol{X}\}, \tag{2.21}$$

натянутая на систему из m векторов $\boldsymbol{x}_i$, называется *подпространством* пространства $\boldsymbol{X}$, если $\boldsymbol{P}$ обладает свойствами линейного пространства, при этом становится справедливой запись $\boldsymbol{P}\subset\boldsymbol{X}$.

Определение 2.14 (О2.14). Пусть $\boldsymbol{P}$ и $\boldsymbol{R}$ подпространства пространства $\boldsymbol{X}$: $\boldsymbol{P},\; \boldsymbol{R}\subset\boldsymbol{X}$, тогда:

1. *суммой* $\boldsymbol{P}+\boldsymbol{R}$ подпространств $\boldsymbol{P}$ и $\boldsymbol{R}$ называется линейная оболочка $\boldsymbol{P}+\boldsymbol{R} = \boldsymbol{L}\{\boldsymbol{X}= p+r : p\in\boldsymbol{P},\; r\in\boldsymbol{R}\}$, при этом $\boldsymbol{P}+\boldsymbol{R}\subset\boldsymbol{X}$;

2. *пересечением* $\boldsymbol{P}\cap\boldsymbol{R}$ подпространств $\boldsymbol{P}$ и $\boldsymbol{R}$ называется множество элементов $\boldsymbol{x}\in\boldsymbol{X}$, которые одновременно принадлежат подпространствам $\boldsymbol{P}$ и $\boldsymbol{R}$, что записывается в форме $\boldsymbol{P}\cap\boldsymbol{R} =\{\boldsymbol{x} : \boldsymbol{x}\in\boldsymbol{P} \;\;\&\;\; \boldsymbol{x}\in\boldsymbol{R}\}$, при этом $\boldsymbol{P}\cap\boldsymbol{R}\subset\boldsymbol{X}$.

Определение 2.15 (О2.15). Подпространства $\boldsymbol{P},\boldsymbol{R}\subset\boldsymbol{X}$ называются *линейно независимыми*, если $\boldsymbol{P}\cap\boldsymbol{R} = \boldsymbol{0}$, где $\boldsymbol{0}$ - нулевой вектор, являющийся нейтральным элементом аддитивной группы $\boldsymbol{X}$.

Определение 2.16 (О2.16). Пусть $\{\boldsymbol{R}_i\subset\boldsymbol{X}\;\; ; i=\overline{1,k}\}$ – линейно независимые подпространства, тогда сумма этих подпространств $\boldsymbol{R}_1+\boldsymbol{R}_2+\dots+\boldsymbol{R}_k$ является *прямой суммой* подпространств, что обозначается знаком $\oplus$ и записывается в форме

$$\boldsymbol{R}_1+\boldsymbol{R}_2+\dots+\boldsymbol{R}_k = \boldsymbol{R}_1\oplus\boldsymbol{R}_2\oplus\dots\oplus\boldsymbol{R}_k = \oplus\sum_{i=1}^{k}\boldsymbol{R}_i. \tag{2.22}$$

Примечание 2.6 (П2.6). Если подпространства $\boldsymbol{P},\boldsymbol{R}\subset\boldsymbol{X}$:

1. линейно независимы, то

$$\dim\{\;\boldsymbol{P}+\boldsymbol{R}= \;\boldsymbol{P}\oplus\boldsymbol{R}\} = \dim\boldsymbol{P}+\dim\boldsymbol{R}, \tag{2.23}$$

2. произвольны, то

$\dim\{\boldsymbol{P}+\boldsymbol{R}\} = \dim\boldsymbol{P} + \dim\boldsymbol{R} - \dim\{\boldsymbol{P}\cap\boldsymbol{R}\}$. (2.24)

Определение 2.17 (О2.17). Оператор A, отображающий элементы $\boldsymbol{x}\in\boldsymbol{X}$ ЛП $\boldsymbol{X}$ в элементы $\mathbf{y}\in\mathbf{Y}$ ЛП $\mathbf{Y}$, где вектор $\boldsymbol{y}=\boldsymbol{Ax}$ именуется *образом* вектора $\boldsymbol{x}$, а вектор $\boldsymbol{x}$ – *прообразом* вектора $\boldsymbol{y}$, называется *линейным оператором*, если над полем F выполняются линейные соотношения

$\boldsymbol{A}(\alpha_1\boldsymbol{x}_1+\alpha_2\boldsymbol{x}_2)=\alpha_1\boldsymbol{Ax}_1+\alpha_2\boldsymbol{A}\,\boldsymbol{x}_2=\alpha_1\boldsymbol{y}_1+\alpha_2\boldsymbol{y}_2$, (2.25)

где $\alpha_1,\ \alpha_2\in F$; $\boldsymbol{x}_1,\ \boldsymbol{x}_2\in\boldsymbol{X}$; $\boldsymbol{y}_1,\boldsymbol{y}_2\in\boldsymbol{Y}$; $\boldsymbol{y}_1=\boldsymbol{A}\,\boldsymbol{x}_1$; $\boldsymbol{y}_2=\boldsymbol{A}\,\boldsymbol{x}_2$.

Рассмотрим *структуру пространства* линейного оператора (ЛО) $\boldsymbol{A}$, для чего введем следующие определения.

Определение 2.18 (О2.18). Множество всех образов $\boldsymbol{y}=\boldsymbol{Ax}$, где $\boldsymbol{y}\in\boldsymbol{Y}$, $\boldsymbol{x}\in\boldsymbol{X}$ называется *областью значений* ЛО $\boldsymbol{A}$ или его *образом*, обозначается $\mathrm{Im}\{\mathbf{A}\}$ и задается в форме

$\mathrm{Im}\{\boldsymbol{A}\}=\{\boldsymbol{y}\in\boldsymbol{Y}:\ \boldsymbol{y}=\boldsymbol{Ax};\ \boldsymbol{x}\in\boldsymbol{X}\}$. (2.26)

Определение 2.19 (О2.19). Множество всех векторов $\boldsymbol{x}\in\boldsymbol{X}$, для которых выполняется равенство $\boldsymbol{Ax}=\boldsymbol{0}$, образует *ядро* $Ker\{\boldsymbol{A}\}$ или *нуль-пространство* $N\{\boldsymbol{A}\}$ линейного оператора $\boldsymbol{A}$, которое задается в форме

$Ker\{\boldsymbol{A}\}=N\{\boldsymbol{A}\}=\{\boldsymbol{x}\in\boldsymbol{X}\ :\boldsymbol{A}\,\boldsymbol{x}=\boldsymbol{0}\}$. (2.27)

Определение 2.20 (О2.20). Рангом $rang\{\boldsymbol{A}\}=rank\{\boldsymbol{A}\}=r_A$ линейного оператора $\boldsymbol{A}$ называется размерность $\dim\mathrm{Im}\{\boldsymbol{A}\}$ образа $\mathrm{Im}\{\boldsymbol{A}\}$ этого оператора.

Определение 2.21 (О2.21). Дефектом $def\{\boldsymbol{A}\}=\nu_A$ линейного оператора $\boldsymbol{A}$ называется размерность $\dim\{Ker\{\boldsymbol{A}\}=N\{\boldsymbol{A}\}\}$ ядра (нуль–пространства) этого оператора.

Определение 2.22 (О2.22). Подпространство $\varphi\subset\boldsymbol{X}$ линейного пространства $\boldsymbol{X}$ называется *инвариантным* относительно линейного оператора $\boldsymbol{A}$, если выполняется условие

$\boldsymbol{A}\varphi\subset\varphi$, (2.28)

в том смысле, что для $\forall\ \boldsymbol{x}\in\varphi\ \ \boldsymbol{y}=\boldsymbol{A}\,\boldsymbol{x}\in\varphi$.

Определение 2.23 (О2.23). Если $dim\{\varphi\}=1$, то *инвариантное* подпространство φ вырождается в вектор ς, при этом условие (2.28) получает представление

$\boldsymbol{A}\varsigma=\lambda\varsigma$, (2.29)

где «$\boldsymbol{A}\varsigma$» означает «$\boldsymbol{A}$ действует на вектор ς», «$\lambda\varsigma$» означает «λ умножить на ς», λ – скаляр, именуемый *собственным* значением (числом) ЛО $\boldsymbol{A}$, ς – собственный вектор ЛО $\boldsymbol{A}$, линейная оболочка $\boldsymbol{L}\{\varsigma\}$, натянутая на собственный вектор ς, представляет собой *собственное (инвариантное)* подпространство φ.

Примечание 2.7 (П2.7). В силу соотношения (2.29) собственный вектор ς линейного оператора $\boldsymbol{A}$ задается с точностью до мультипликативной константы.

Определение 2.24 (О2.24). Пусть в $n-$мерном ЛП $\boldsymbol{X} = \boldsymbol{X}^n$ выбран базис $e=\{e_i;i=\overline{1,n}\}$, а в $m-$мерном ЛП $\boldsymbol{Y}=\boldsymbol{Y}^m$ выбран базис $f=\{f_j;j=\overline{1,m}\}$, тогда *матрицей* A относительно пары базисов (e,f) *линейного оператора*

$$\boldsymbol{A}:\boldsymbol{X}\Rightarrow\boldsymbol{Y}:\boldsymbol{y}=\boldsymbol{A}\boldsymbol{x}\ ;\ \boldsymbol{x}\in\boldsymbol{X};\ \boldsymbol{y}\in\boldsymbol{Y} \tag{2.30}$$

называется *двумерный* $(m\times n)-$*массив чисел*, столбцовое представление которого

$$A=[A_1\ A_2\ \dots\ A_n]=row\{A_i;i=\overline{1,n}\}, \tag{2.31}$$

таково, что столбцы $A_i\left(i=\overline{1,n}\right)$ его составлены из числовых коэффициентов A_{ji} $\left(j=\overline{1,m};i=\overline{1,n}\right)$ представления вектора $\boldsymbol{A}e_i$ в базисе f линейной комбинацией

$$\boldsymbol{A}e_i=A_{1i}f_1+A_{2i}f_2+\dots+A_{mi}f_m\ . \tag{2.32}$$

Примечание 2.8 (П2.8). Если вектор $\boldsymbol{x}$ записать в координатной форме в базисе $e=\{e_i;i=\overline{1,n}\}$ с помощью представления $\boldsymbol{x}=\sum_{i=1}^{n}x_{ei}e_i$, а вектор $\boldsymbol{y}$ в координатной форме в базисе $f=\{f_j;j=\overline{1,m}\}$ с помощью представления $\boldsymbol{y}=\sum_{j=1}^{m}y_{fj}f_j$, то введение матрицы A линейного оператора $\boldsymbol{A}$ позволяет от *векторно-операторной* формы (2.30) перейти к *векторно-матричной мультипликативной* форме

$$y_f=Ax_e, \tag{2.33}$$

в которой $y_f=col\{y_{fj};j=\overline{1,m}\}$, $x_e=col\{x_{ei};i=\overline{1,n}\}$.

Определение 2.25 (О2.25). Две матрицы A и $\overline{A}$ называются *подобными,* если они задают один и тот же линейный оператор $\boldsymbol{A}$ относительно различных пар базисов (e,f) и $(\overline{e},\overline{f})$.

Примечание 2.9 (П2.9). В силу определения 2.25 линейному оператору $\boldsymbol{A}$ соответствует множество сколь угодной большой *мощности* подобных матриц, каждая из которых порождается своей парой (e,f) базисов.

Введем в рассмотрение *матричное условие подобия* двух матриц одной размерности $(m\times n)$ с помощью следующего утверждения.

Утверждение 2.2 (У2.2). Пусть линейный оператор $\boldsymbol{A}$ реализует отображение $\boldsymbol{A}:\boldsymbol{X}\Rightarrow\boldsymbol{Y}:\boldsymbol{y}=\boldsymbol{A}\boldsymbol{x}$; $\boldsymbol{x}\in\boldsymbol{X}$; $\boldsymbol{y}\in\boldsymbol{Y}$: $\dim\boldsymbol{X}=n$; $\dim\boldsymbol{Y}=m$. Пусть в ЛП $\boldsymbol{X}$ задана пара базисов$(e,\overline{e})$, каждый из которых представляет собой систему линейно независимых векторов соответстенно $e=\{e_i;i=\overline{1,n}\}$ и $\overline{e}=\{\overline{e}_i;i=\overline{1,n}\}$. Базисы связаны *матрицей М преобразования базисов,* задаваемого соотношением

$$e = M\bar{e}, \tag{2.34}$$

где $(n \times n)$ матрица $M = row\{M_i; i = \overline{1,n}\}$ составлена из столбцов M_i, элементы которых M_{ji} представляют собой коэффициенты линейного разложения элементов $\{e_i; i = \overline{1,n}\}$ базиса e по базисным компонентам $\{\bar{e}_i; i = \overline{1,n}\}$ базиса $\bar{e}$, задаваемого в форме

$$e_i = M_{1i}\bar{e}_i + M_{2i}\bar{e}_2 + \ldots + M_{ni}\bar{e}_n. \tag{2.35}$$

Пусть в ЛП ***Y*** задана пара базисов $(f, \bar{f})$, которые составлены из m векторов $f = \{f_l; l = \overline{1,m}\}$ и $\bar{f} = \{\bar{f}_l; l = \overline{1,m}\}$ соответственно и связаны $(m \times m)$ – матрицей T преобразования базисов в форме

$$f = T\bar{f}, \tag{2.36}$$

где матрица T строится по той же схеме, что и матрица M.

Тогда подобные $(m \times n)$ матрицы A и $\bar{A}$, задающие линейный оператор $\boldsymbol{A}$ относительно пар базисов (e, f) и $(\bar{e}, \bar{f})$ соответственно связаны *матричным соотношением (условием) подобия*

$$T\bar{A} = AM. \tag{2.37}$$

Доказательство утверждения строится на представлении *бескоординатного* вектора ***x***∈***X*** относительно базисов e и $\bar{e}$ в форме *координатных* векторов x и $\bar{x}$, связанных в силу (2.34) векторно-матричным соотношением

$$x = M\bar{x}, \tag{2.38}$$

а также *бескоординатного* вектора ***y***∈***Y*** относительно базисов f и $\bar{f}$ в форме *координатных* векторов y и $\bar{y}$, связанных в силу (2.36) векторно-матричным соотношением

$$y = T\bar{y}. \tag{2.39}$$

Если теперь *векторно-операторное* соотношение ***y*= *Ax*** записать в векторно-матричной форме, то для каждой пары базисов (e, f) и $(\bar{e}, \bar{f})$ соответственно получим векторно-матричные представления

$$y = Ax = AM\bar{x}, \tag{2.40}$$

$$\bar{y} = \bar{A}\bar{x}. \tag{2.41}$$

Подстановка (2.41) в (2.39) приводит к векторно-матричному соотношению

$$y = T\bar{A}\bar{x}. \tag{2.42}$$

Сравнение векторно-матричных соотношений (2.40) и (2.42) приводит к соотношению (2.37). ■

Примечание 2.9 (П2.9). Так как матрицы преобразования базисов M и T таковы, что существуют обратные им матрицы M^{-1} и T^{-1}, то (2.37) имеет следующие эквивалентные представления

$$A = T\overline{A}M^{-1}, \quad \overline{A} = T^{-1}AM. \qquad (2.43)$$

Примечание 2.10 (П2.10). Если размерности линейных пространств $\boldsymbol{X}$ и $\boldsymbol{Y}$ таковы, что они совпадают $(n = m)$, базисы $(e, \overline{e})$ и $(f, \overline{f})$ связаны идентичными матрицами их преобразования так, что $T = M$, то матричные условия подобия (2.37) и (2.43) принимают вид

$$M\overline{A} = AM, \quad A = M\overline{A}M^{-1}, \quad \overline{A} = M^{-1}AM, \qquad (2.44)$$

где A и $\overline{A}$ – подобные $(n \times n)$– *квадратные* матрицы.

Введенная в рассмотрение матрица A линейного оператора $\boldsymbol{A}$ позволяет соотношение (2.29) записать в векторно-матричной форме

$$A\xi = \lambda\xi \qquad (2.45)$$

или

$$(A - \lambda I)\xi = 0. \qquad (2.46)$$

В (2.45), (2.46) ξ есть собственный вектор ς ЛО $\boldsymbol{A}$, записанный относительно базиса $e = \{ e_i; i = \overline{1,n} \}$ в координатной форме, именуемый *собственным вектором матрицы* A.

Определение 2.26 (О2.26). Матрица $A - \lambda I$ называется *характеристической матрицей* матрицы A, определитель $\det(A - \lambda I) = \Delta(\lambda)$ не зависит от выбора базиса e и называется *характеристическим многочленом* матрицы A (ЛО $\boldsymbol{A}$), всюду выше и ниже I – единичная матрица, согласованная по размерности с $(n \times n)$-матрицей A.

Система уравнений (2.46) относительно компонентов собственного вектора ξ *совместна,* а потому имеет *ненулевое* решение $\xi \neq 0$, если определитель этой системы равен нулю, т.е. при выполнении равенства

$$\det(A - \lambda I) = \Delta(\lambda) = 0 \qquad (2.47)$$

Уравнение (2.47) называется *характеристическим* уравнением матрицы A (ЛО $\boldsymbol{A}$), а его корни $\{\lambda_i; i = \overline{1,n}\}$ являются собственными значениями этой матрицы (ЛО $\boldsymbol{A}$).

Пусть λ_k – некоторое собственное значение матрицы A. Если характеристический многочлен $\Delta(\lambda)$ можно представить в виде $\Delta(\lambda) = (\lambda - \lambda_k)^{m_k} Q(\lambda_k)$, где $Q(\lambda_k) \neq 0$, то число m_k называется *алгебраической кратностью* собственного значения λ_k.

Рассмотрим случай, когда все корни характеристического уравнения (2.43) являются различными, т.е. $\Delta(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)...(\lambda - \lambda_n)$. Тогда собственный вектор ξ_k матрицы A, соответствующий значению λ_k, является решением уравнения

$$\left(A-\lambda_k I\right)\xi_k=0 \tag{2.48}$$

Это означает, что вектор ξ_k принадлежит ядру матрицы $\left(A-\lambda_k I\right)$ т.е. $\xi_k \in N\left(A-\lambda_k I\right)$, таким образом, ядро $N\left(A-\lambda_k I\right)$ является *собственным* подпространством матрицы A, порожденным собственным вектором ξ_k. Это подпространство является одномерным, если дефект матрицы $\left(A-\lambda_k I\right)$ равен единице.

Собственные векторы $\xi_1, \xi_2, \ldots \xi_n$, соответствующие различным собственным значениям $\lambda_1, \lambda_2, \ldots \lambda_n$, являются линейно независимыми, образуют базис в пространстве X^n, само пространство X^n расщепляется оператором ***A*** с матрицей A в прямую сумму одномерных собственных подпространств в форме

$$X^n=\oplus\sum_{k=1}^{n} H_k\ ;\ H_k=N\left(A-\lambda_k I\right); k=\overline{1,n}. \tag{2.49}$$

Матрица A оператора ***A*** в базисе из собственных векторов является диагональной с собственными значениями $\lambda_1, \lambda_2, \ldots, \lambda_n$ на главной диагонали, при этом справедливо равенство

$$\Lambda=diag\left\{\lambda_k; k=\overline{1,n}\right\}=H^{-1}AH; H=row\left\{H_k; k=\overline{1,n}\right\}, \tag{2.50}$$

где H – матрица линейно независимых *собственных* векторов ξ_k матрицы A.

Определение 2.27 (О2.2) Оператор ***A*** и его матрица A, имеющие ровно *n* линейно-независимых собственных векторов, называются оператором и матрицей *простой структуры*. Таким образом, любая матрица, имеющая различные собственные значения, подобна диагональной матрице Λ и является матрицей простой структуры.

Рассмотрим случай кратных корней характеристического уравнения. Тогда $\Delta(\lambda)=\left(\lambda-\lambda_1\right)^{m_1}\left(\lambda-\lambda_1\right)^{m_2}\ldots\left(\lambda-\lambda_1\right)^{m_k}$ и все собственные векторы ξ_k, соответствующие значению λ_k, также являются решениями уравнения (2.46).

Определение 2.28 (О2.28). Число μ_k линейно-независимых собственных векторов равно размерности собственного подпространства $H_k=N(A-\lambda_k I)$ (или дефекту матрицы $A-\lambda_k I$), которое называется *геометрической кратностью* или *степенью вырожденности* собственного значения λ_k.

Ядро матрицы $A-\lambda_k I$ называется *корневым подпространством* K_k соответствующим значению λ_k. Геометрическая кратность корня μ_k больше его алгебраической кратности m_k, а собственное подпространство H_k содержится в корневом подпространстве K_k.

Для любого оператора с матрицей A существует аналогичное (2.49) разложение пространства X^n на прямую сумму его корневых подпространств:

$X^n = K_1 \oplus K_2 \oplus \ldots \oplus K_r.$ (2.51)

В том случае и только тогда, когда алгебраические и геометрические кратности всех собственных значений матрицы A совпадают, совпадают и их собственные и корневые подпространства и, следовательно, матрица имеет простую структуру.

В общем случае произвольная матрица A линейного оператора преобразованием подобия может быть приведена к *нормальной жордановой форме:*

$$J = diag\{J_1, J_2, ..., J_r\} = T^{-1}AT, \qquad (2.52)$$

где блоки J_i имеют размерность $m_i \times m_i$ и могут быть представлены следующим образом: $J_i = diag\{J_{i_1}, J_{i_2}, ..., J_{i_{\mu i}}\}$, причем каждый подблок имеет вид

$$J_{ij} = \begin{bmatrix} \lambda_i & 1 & 0 & \vdots & 0 \\ 0 & \lambda_i & 1 & \vdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0, & .... & .... & 0, & \lambda_i \end{bmatrix}; \qquad (2.53)$$

каждый блок T_i матрицы $T = [T_1\ T_2, ..., T_r]$ состоит из столбцов координат одного из линейно-независимых векторов и других корневых векторов, соответствующих собственному значению λ_i. Если матрица имеет простую структуру разложения, то (2.52) совпадает с (2.50).

Введем в рассмотрение численную (скалярную) характеристику элементов (векторов) $x = [x_1, x_2, ..., x_n]^T$ арифметического пространства X^n, именуемую *нормой* вектора, полагая ниже, что пространство действительное, что позволяет записать $X^n = R^n$.

Определение 2.29 (O2.29). Пусть функция $\varphi(\cdot)$ сопоставляет каждому вектору $x \in R^n$ – линейного вещественного пространства вещественное число $\|x\|$, называемое *нормой* (размером) этого вектора, если выполняются условия:

1) $\varphi(x) = \|x\| > 0$ для $\forall x \neq 0$ и $\varphi(x) = \|x\| = 0$ при $x = 0$;

2) $\varphi(\alpha x) = \|\alpha x\| = |\alpha| \cdot \|x\|$;

3) $\varphi(x + y) = \|x + y\| \leq \|x\| + \|y\|$.

Универсальной векторной нормой является векторная норма Гельдера, задаваемая выражением:

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}; \quad p \text{ – целое положительное.}$$

Наиболее употребительными векторными нормами являются нормы при $p = 1, 2$ и ∞:

1. $\|x\|_1 = \sum_{i=1}^{n} |x_i|$ – абсолютная норма вектора;

2. $\|x\|_2 = \left(\sum_{i=1}^{n} x_i^2 \right)^{1/2}$ – квадратичная или *евклидова* норма вектора;

3. $\|x\|_\infty = \lim_{P \to \infty} \|x\|_P = \max_{i=\overline{1,n}} |x_i|$ – бесконечная или экстремальная норма вектора.

Приведенные *векторные нормы эквивалентны* в том смысле, что для норм $\|x\|_\mu$ и $\|x\|_\nu$ существуют положительные числа β_1 и β_2 такие, что выполняются неравенства: $\beta_1 \|x\|_\mu \le \|x\|_\nu \le \beta_2 \|x\|_\mu$.

Так для норм $\|x\|_1$, $\|x\|_2$ и $\|x\|_\infty$ выполняются оценочные неравенства:

$\|x\|_2 \le \|x\|_1 \le \sqrt{n}\|x\|_2$,

$\|x\|_\infty \le \|x\|_2 \le \sqrt{n}\|x\|_\infty$,

$\|x\|_\infty \le \|x\|_1 \le n\|x\|_\infty$.

Линейное арифметическое пространство X^n с введенной векторной нормой $\|x\|_p$ образует линейное *нормированное* пространство, обозначаемое в одной из форм $\left\{ X^n, \|x\|_p \right\}$ или X_p^n, в дальнейшем в основном рассматривается случай $p = 2$, при этом индекс нормы опускается.

Примечание 2.11 (П2.11). Линейное нормированное пространство является *метрическим* так, как выполняется цепочка равенств $\|x\| = \|x - 0\| = d(x,0)$, где *0*-нулевой вектор.

Введем в рассмотрение численную характеристику, являющуюся оценкой *взаимного положения* элементов (векторов) линейного пространства. Такой характеристикой является *скалярное произведение* двух координатных векторов.

Определение 2.30 (О2.30). Пусть функция $\varphi\{\cdot\}$ сопоставляет каждой паре $\{\cdot\} = \{x, y\}$ векторов $x, y \in X^n = R^n$ линейного вещественного арифметического пространства вещественное *число*, обозначаемое в одной из форм (x, y) *или* $\langle x, y \rangle$, называемое *скалярным произведением* (СП) векторов x и y, если выполняются условия:

1. коммутативности: $(x, y) = (y, x)$;
2. дистрибутивности: $(x + y, z) = (x, z) + (y, z)$;
3. линейности: $(\alpha \cdot x, y) = \alpha \cdot (x, y), \alpha \in R$;

4. $(x, x) = \|x\|^2$, где $\|x\| = \|x\|_E$ – евклидова норма вектора x;

5. неравенства Коши-Буняковского: $|(x, y)| \le \|x\| \cdot \|y\|$;

6. оценки взаимного положения векторов x и y:

$(x, y) = \|x\| \cdot \|y\| \cdot \cos\{ang\{x, y\}\}$,

где $ang\{x, y\}$ – угол между векторами x и y.

Примечание 2.12 (П2.12). Вычисление скалярного произведения (x, y) векторов $x = col\{x_i; i = \overline{1,n}\}$, $y = col\{y_i; i = \overline{1,n}\}$ производится в силу соотношений

$$(x, y) = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n = \sum_{i=1}^{n} x_i y_i = \sum_{i=1}^{n} y_i x_i = (y, x). \tag{2.54}$$

Соотношения (2.54) делают справедливыми следующие представления скалярного произведения

$$(x, y) = x^T y = y^T x = x^T I\, y = y^T I\, x = (x, y)_I = (y, x)_I. \tag{2.55}$$

Соотношения (2.55), содержащие представление скалярного произведения в виде мультипликативной структуры «вектор–строка – единичная матрица – вектор–столбец», позволяют ввести в рассмотрение понятие *скалярного произведения с весом* (в (2.55) с *единичным* весом, задаваемым *единичной весовой* матрицей I). По аналогии может быть введено скалярное произведение $(x, y)_P$ с *неединичным весом,* порождаемое *неединичной весовой* матрицей P в форме

$$(x, y)_P = (Px, y) = (x, Py) = x^T P^T y = x^T P y. \tag{2.56}$$

Весовая матрица P должна быть: *симметричной* $P = P^T$ и *положительно-определенной* так, что собственные значения этой матрицы $\{\lambda_{Pi} : det(\lambda_P I - P) = 0 : \lambda_{Pi} > 0; i = \overline{1,n}\}$ *положительны*.

Линейное арифметическое пространство со скалярным произведением в силу удовлетворения СП условию 4. определения 2.30 является *нормированным* ЛП и *метрическим.* ЛП со скалярным произведением именуется *гильбертовым* линейным пространством.

Частным случаем *гильбертова* ЛП является *евклидово* линейное пространство.

Определение 2.31 (О2.31). Векторы x и y называются *ортогональными* ($x \perp y$), если их скалярное произведение равно нулю, т.е. выполняется равенство $(x, y) = 0$.

Определение 2.32 (О2.32). Система $\{e_i; i = \overline{1,n}\}$ векторов ЛП называется *ортогональной*, если все векторы этой системы попарно ортогональны: $(e_i, e_j) = 0; i, j = \overline{1,n}; i \ne j$.

Определение 2.33 (О2.33). Система $\{e_i; i = \overline{1,n}\}$ векторов ЛП называется *ортонормированной*, если выполняется равенство:

$\left(e_i, e_j\right) = \delta_{ij}; i, j = \overline{1,n};$ где символ Кронекера δ_{ij} удовлетворяет условиям $\delta_{ij} = 1$ при $i = j$ и $\delta_{ij} = 0$ при $i \neq j$.

Определение 2.34 (О2.34). *Евклидовым* линейным пространством E^n называется $n-$мерное ЛП**,** в качестве базиса в котором используется система $\left\{e_i; i = \overline{1,n}\right\}$ ортонормированных векторов таких, что матрица E, имеющая своими столбцами вектора $\left\{e_i; i = \overline{1,n}\right\}$, образует единичную матрицу, т.е. выполняется матричное равенство

$$E = row\left\{e_i; i = \overline{1,n}\right\} = I. \tag{2.57}$$

Определение 2.35 (О2.35). Базис, построенный на системе $\left\{e_i; i = \overline{1,n}\right\}$ ортонормированных векторов, удовлетворяющих соотношению (2.57) образует *евклидов* или *естественный базис*.

Ортонормированные базисы обладают рядом существенных достоинств перед другими базисными системами в ЛП X^n. Переход от *произвольной исходной* базисной системы $\left\{e_i; i = \overline{1,n}\right\}$ к *ортонормированной* базисной системе $\left\{u_i : \left(u_i, u_j\right) = \delta_{ij}; i, j = \overline{1,n};\right\}$ может быть осуществлен с помощью алгоритма Грама-Шмидта. Алгоритм Грамма–Шмидта является двухшаговой *рекуррентной процедурой*, при этом на *первом шаге* осуществляется переход от *произвольной исходной* базисной системы $\left\{e_i; i = \overline{1,n}\right\}$ к *ортогональной* промежуточной базисной системе $\left\{v_i : \left(v_i, v_j\right) = 0; i, j = \overline{1,n}; i \neq j\right\}$, на *втором шаге* осуществляется *нормировка* элементов построенного базиса, приводящая к искомой *ортонормированной* базисной системе. В результате, **алгоритм Грамма-Шмидта** принимает вид:

$$\begin{aligned}
&v_1 = e_1;\ u_1 = \frac{v_1}{\|v_1\|};\\
&v_2 = e_2 - \left(e_2, u_1\right) u_1;\ u_2 = \frac{v_2}{\|v_2\|};\\
&\vdots\\
&v_k = e_k - \sum_{i=1}^{k-1} (e_k, u_i) u_i;,\ u_k = \frac{v_k}{\|v_k\|};\ k = \overline{2,n}.
\end{aligned} \tag{2.58}$$

В заключение раздела рассмотрим решение следующей задачи. В ЛП R^n задан некоторый базис $\left\{e_i; i = \overline{1,n}\right\}$, относительно этого базиса задан координатный вектор $x = col\left\{x_i; i = \overline{1,n}\right\}$, задан в ЛП и базис $\left\{f_i; i = \overline{1,n}\right\}$, компоненты которого f_i заданы в координатной форме относительно того же базиса $\{e_i; i=\overline{1,n}\}$. Ставится задача конструирования представления вектора x относительно базиса $\{f_i; i=\overline{1,n}\}$ в форме

$$x = x_{f1}f_1 + x_{f2}f_2 + ... + x_{fn}f_n = \sum_{i=1}^{n} x_{fi}f_i, \tag{2.59}$$

так что относительно базиса $\{f_i; i=\overline{1,n}\}$ вектор x получает представление

$$x_f = col\{x_{fi}; i=\overline{1,n}\}. \tag{2.60}$$

Если умножить выражение (2.59) *скалярно* последовательно n раз на элементы $\{f_i; i=\overline{1,n}\}$, то будет получена система уравнений относительно искомых *неизвестных* $\{x_{fi}; i=\overline{1,n}\}$, решение которой в векторно-матричной форме имеет вид

$$x_f = col\{x_{fi}; i=\overline{1,n}\} = \Gamma^{-1}\beta, \tag{2.61}$$

где $\Gamma = col\{row\{(f_i, f_j); j=\overline{1,n}\}; i=\overline{1,n}\}$ – матрица Грама, построенная на скалярных произведениях всех пар элементов базиса $\{f_i; i=\overline{1,n}\}$ друг на друга; $\beta = col\{(x, f_i); i=\overline{1,n}\}$ – вектор–столбец из скалярных произведений вектора x на элементы базиса представления $\{f_i; i=\overline{1,n}\}$. Следует заметить, что, если базис f состоит из ортонормированных элементов, то матрица Грама становится единичной ($\Gamma = I$), при этом (2.61) принимает вид $x_f = \beta$.

Следует заметить, что невырожденность ($\exists\Gamma^{-1}$) матрицы Грама является критерием *линейной независимости* системы векторов, на скалярных произведениях которой матрица Грама построена. Более того, решение задачи в форме (2.61) *геометрически* представляет собой решение задачи проектирования исходного вектора x на пространство, натянутое на систему векторов $\{f_i; i=\overline{1,m}; m \le n\}$..

Примеры и задачи

2.2.1*. Доказать, что для любого оператора A, действующего из $\boldsymbol{X}^n$ в $\boldsymbol{Y}^n$ сумма ранга r_A и дефекта n_A равна размерности n пространства $\boldsymbol{X}^n$.

2.2.2*. Привести матрицу $A = \begin{bmatrix} -1 & 1 & -1 \\ 0 & 1 & -4 \\ 0 & 1 & -3 \end{bmatrix}$ к нормальной жордановой форме и определить матрицу преобразования T.

2.2.3*. Определить ортогональную проекцию вектора $x=\begin{bmatrix}1 & 1\end{bmatrix}^T$ на линейную оболочку $\boldsymbol{L}(y)$ натянутую на вектор $y=\begin{bmatrix}\sqrt{3} & 1\end{bmatrix}^T$.

2.2.4. Выяснить, являются ли системы векторов арифметических пространств линейно зависимыми.

а) $x_1=\begin{bmatrix}-3 & 1 & 5\end{bmatrix}^T$, $x_2=\begin{bmatrix}6 & -2 & 15\end{bmatrix}^T$;

б) $x_1=\begin{bmatrix}1 & 2 & 3\end{bmatrix}^T$, $x_2=\begin{bmatrix}2 & 5 & 7\end{bmatrix}^T$, $x_3=\begin{bmatrix}3 & 7 & 10\end{bmatrix}^T$.

2.2.5. Проверить, что векторы

$e_1=\begin{bmatrix}2 & 2 & -1\end{bmatrix}^T$, $e_2=\begin{bmatrix}2 & -1 & 2\end{bmatrix}^T$, $e_3=\begin{bmatrix}-1 & 2 & 2\end{bmatrix}^T$ образуют базис пространства R^3 и найти координаты вектора $x=[1\ \ 1\ \ 1]^T$ в этом базисе.

2.2.6. Вектор x в естественном базисе пространства R^2 имеет координаты $x=\begin{bmatrix}1 & 1\end{bmatrix}^T$, а оператору $\boldsymbol{A}$ в этом базисе соответствует матрица $A=\begin{bmatrix}2 & -1\\ 1 & 1\end{bmatrix}$. Определить координаты векторов x и Ax в базисе, элементами которого являются столбцы матрицы A.

2.2.7. Определить область значений и ядро оператора с матрицей $A=\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}$ в естественном базисе R^2.

2.2.8. Применить алгоритм ортогонализации Грамма–Шмидта к следующим системам векторов:

а) $e_1=\begin{bmatrix}1 & -2 & 2\end{bmatrix}^T$, $e_2=\begin{bmatrix}-1 & 0 & -1\end{bmatrix}^T$, $e_3=\begin{bmatrix}5 & -3 & -7\end{bmatrix}^T$;

б) $q_1=\begin{bmatrix}2 & 3 & -4 & -6\end{bmatrix}^T$, $q_2=\begin{bmatrix}1 & 8 & -2 & -16\end{bmatrix}^T$,

$q_3=\begin{bmatrix}12 & 5 & -14 & 5\end{bmatrix}^T$, $q_4=\begin{bmatrix}3 & 11 & 4 & -7\end{bmatrix}^T$.

2.2.9. Найти матрицу A оператора $\boldsymbol{A}$ в R^3, отображающего векторы:

$x_1=\begin{bmatrix}0 & 0 & 1\end{bmatrix}^T$ в $y_1=\begin{bmatrix}2 & 3 & 5\end{bmatrix}^T$;

$x_2=\begin{bmatrix}0 & 1 & 1\end{bmatrix}^T$ в $y_2=\begin{bmatrix}1 & 0 & 0\end{bmatrix}^T$;

$x_3=\begin{bmatrix}1 & 1 & 1\end{bmatrix}^T$ в $y_3=\begin{bmatrix}0 & 1 & -1\end{bmatrix}^T$:

а) в естественном базисе;

б) в базисе x_1, x_2, x_3.

2.2.10. Найти базис ортогонального дополнения $L^{\perp}$ линейной оболочки L системы векторов пространства R^4:

$e_1=[1, 3, 0, 2]^T$, $e_2=[3, 7, -1, 2]^T$, $e_3=[2, 4, -1, 0]^T$.

2.2.11. Вычислить собственные значения и собственные векторы следующих матриц:

а) $A=\begin{bmatrix}2 & 0\\ 0 & 2\end{bmatrix}$; б) $A=\begin{bmatrix}4 & -1 & -2\\ 2 & 1 & -2\\ 1 & -1 & 1\end{bmatrix}$; в) $A=\begin{bmatrix}1 & 0 & -4\\ 0 & 3 & 0\\ -2 & 0 & -1\end{bmatrix}$.

2.2.12. Привести следующие матрицы к нормальной жордановой форме и определить матрицы преобразования координат:

а) $A=\begin{bmatrix}8 & -8 & -2\\ 4 & -3 & -2\\ 3 & -4 & 1\end{bmatrix}$; б) $A=\begin{bmatrix}1 & 0.6 & -0.8\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$.

2.2.13*. Показать, что матрица ортогонального проектирования на линейную оболочку линейно независимых векторов $L(u_1,u_2,...,u_k)$ равна $P=U(U^TU)^{-1}$, где U – матрица, составленная из столбцов координат векторов $u_1,u_2,...,u_k$. Как выглядит матрица P, когда векторы $u_1,u_2,...,u_k$ образуют ортонормированную систему?

Решение вариантов задач

Решение задачи 2.2.1. Разложим пространство X^n в прямую сумму пространства $X^n=N(A)+L_A$, где $N(A)$ – ядро (нуль–пространство) матрицы A, L_A – любое дополнительное к $N(A)$ подпространство. Тогда для любого $x\in X^n$ имеет место единственное представление $x=x_N+x_L$, где $x_N\in N(A)$, $x_L\in L_A$.

Поскольку $y=Ax=A(x_N+x_L)=Ax_L\in Jm(A)$, так как $Ax_N=0$, то любой вектор $y\in Jm(A)$ имеет прообраз из подпространства L_A, причем единственный, откуда следует, что размерности подпространств $Jm(A)$ и L_A совпадают. Но $\dim X^n=\dim N(A)+\dim L_A$ и $\dim L_A=r_A$, откуда $n=n_A+r_A$.

Решение задачи 2.2.2. Решение характеристического уравнения $\Delta(\lambda)=0$ исходной матрицы A дает для нее три одинаковых собственных значения $\lambda_1=\lambda_2=\lambda_3=-1$, следовательно, алгебраическая кратность корня равна трем. Дефект матрицы

$$A-\lambda I=A\text{-}(\text{-}1)I=\begin{bmatrix}0 & 1 & -1\\ 0 & 2 & 24\\ 0 & 1 & -2\end{bmatrix}$$

равен единице, следовательно, собственное подпространство $N\{A-(-1)I\}$ является одномерным.

Поскольку собственное пространство, соответствующее собственному значению $\lambda=-1$, является одномерным, форма Жордана J состоит из единственного блока, отвечающего этому значению $\lambda = -1$ и принимает вид

$$J = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} -1 & 1 & -1 \\ 0 & 1 & -4 \\ 0 & 1 & -3 \end{bmatrix}.$$

Уравнение (2.52) $TJ = AT$, записанное в столбцовой форме $\begin{bmatrix} T_1 & T_2 & T_3 \end{bmatrix} \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix} = A\begin{bmatrix} T_1 & T_2 & T_3 \end{bmatrix}$ порождает систему соотношений

$$AT_1 = \lambda T_1;\ AT_2 = \lambda T_2 + T_1;\ AT_3 = \lambda T_3 + T_2.$$

Последовательное решение этих уравнений *без использования процедуры обращения* позволяет сконструировать матрицу T в форме

$T = \begin{bmatrix} T_1 & T_2 & T_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & 1 & -1 \end{bmatrix}$, в результате чего уравнение подобия (2.52) приводит к искомому результату

$$J = T^{-1}AT = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & 1 & -1 \end{bmatrix}^{-1} \begin{bmatrix} -1 & 1 & -1 \\ 0 & 1 & -4 \\ 0 & 1 & -3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & 1 & -1 \end{bmatrix} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{bmatrix}.$$

<u>Решение задачи 2.2.3.</u> Если $\hat{x}$ – ортогональная проекция вектора x на оболочку $L(y)$, то $\hat{x} = \alpha y$, где $\alpha = const$. Поскольку $(x - \hat{x}) \perp L$, то имеем $(x - \hat{x}, y) = 0$, или $x^T y = \hat{x}^T y$, откуда получим $x = \dfrac{x^T y}{(y,y)} y = \dfrac{x^T y}{\|y\|^2} y = \left[\dfrac{3+\sqrt{3}}{4}, \dfrac{1+\sqrt{3}}{4} \right]^T$. Задача может быть также решена с использованием матрицы Грама.

3. МАТРИЧНЫЕ ИНВАРИАНТЫ И НЕИНВАРИАНТЫ. СИНГУЛЯРНОЕ РАЗЛОЖЕНИЕ МАТРИЦ

Рассматриваются подобные матрицы A и $\overline{A}$ размерности $(n \times n)$, задающие один и тот же линейный оператор ***A*** относительно различных пар базисов. Для этих матриц существует невырожденная $(n \times n)$ – матрица M, связывающая матрицы A и $\overline{A}$ матричным соотношением подобия

$$M\overline{A} = AM\,. \tag{3.1}$$

Матричное соотношение подобия (3.1) имеет два матричных аналога

$$\overline{A} = M^{-1}AM \text{ и } A = M\overline{A}M^{-1}. \tag{3.2}$$

Возникает вопрос: *какие характеристик* подобных матриц A и $\overline{A}$ при преобразованиях подобия вида (3.2) *сохраняются*, а *какие* – нет?

Определение 3.1 (О3.1). Характеристики матриц размерности $(n \times n)$, принадлежащих классу подобных, то есть задающих один и тот же линейный оператор ***A***, которые *сохраняются* неизменными на всем классе подобных представлений, называются *матричными инвариантами.*

Определение 3.2 (О3.2). Характеристики матриц размерности $(n \times n)$, принадлежащих классу подобных, то есть задающих один и тот же линейный оператор ***A***, которые для каждой реализации подобной матрицы оказываются *своими*, называются *матричными неинвариантами.*

Рассмотрим ***матричные инварианты*** на примере двух подобных матриц A и $\overline{A}$ вида (3.2).

1. Первым матричным *инвариантом* являются характеристические полиномы $\det(\lambda I - A)$ и $\det(\lambda I - \overline{A})$ подобных матриц A и $\overline{A}$. В этой связи сформулируем и докажем следующее утверждение.

Утверждение 3.1 (У3.1). Характеристические полиномы подобных матриц A и $\overline{A}$ совпадают так, что выполняется соотношение

$$\det(\lambda I - A) = \lambda^n + a_1\lambda^{n-1} + a_2\lambda^{n-2} + \ldots + a_{n-1}\lambda + a_n = \det(\lambda I - \overline{A})\,. \quad \square(3.3)$$

Доказательство. Подставим в (3.3) матрицу A с использованием представления (3.2), а также представление единичной матрицы в форме $I = MM^{-1} = MIM^{-1}$, тогда получим скалярно-матричное соотношение

$$\det(\lambda MIM^{-1} - M\overline{A}M^{-1}) = \det\{M(\lambda I - \overline{A})M^{-1}\} \tag{3.4}$$

Воспользуемся положением о том, что детерминант произведения матриц равен произведению детерминантов, с учетом которого (3.4) принимает вид

$$\det\left\{M(\lambda I-\overline{A})M^{-1}\right\}=\det(M)\det\left(\lambda I-\overline{A}\right)\det(M^{-1}).$$

Если учесть, что детерминант это скаляр, и воспользоваться свойством детерминанта произведения матриц в обратном порядке так, что произведение детерминантов равняется детерминанту произведения матриц, согласованных по размерности, то получим

$$\det(M)\det(M^{-1})\det\left(\lambda I-\overline{A}\right)=\det(MM^{-1}=I)\det\left(\lambda I-\overline{A}\right)=\det\left(\lambda I-\overline{A}\right).$$

Таким образом, установлено равенство $\det(\lambda I-A)=\det(\lambda I-\overline{A})$. ■

2. Вторым матричным *инвариантом* являются алгебраические спектры $\sigma\{A\}$ и $\sigma\{\overline{A}\}$ собственных значений подобных матриц A и $\overline{A}$, что записывается в форме

$$\sigma\{A\}=\sigma\{\overline{A}\}=\left\{\lambda_i:\lambda_i{}^n+a_1\lambda_i{}^{n-1}+\ldots+a_{n-1}\lambda_i+a_n=0;i=\overline{1,n}\right\}. \tag{3.5}$$

3. Третьим матричным инвариантом являются детерминанты $\det(A)$ и $\det(\overline{A})$ подобных матриц A и $\overline{A}$, определяемые соотношением

$$\det(A)=\det(\overline{A})=\prod_{i=1}^{n}\lambda_i=\lambda_1\lambda_2\ldots\lambda_{n-1}\lambda_n \tag{3.6}$$

4. Четвертым матричным *инвариантом* являются следы $tr(A)$ и $tr(\overline{A})$ подобных матриц A и $\overline{A}$, определяемые выражением

$$tr(A)=tr(\overline{A})=\sum_{i=l}^{n}\lambda_i=\lambda_1+\lambda_2+\ldots+\lambda_{n-1}+\lambda_n, \tag{3.7}$$

и вычисляемые с помощью соотношений

$$tr(A)=\sum_{i=1}^{n}A_{ii}=A_{11}+A_{22}+\ldots+A_{n-1,n-1}+A_{nn}, \tag{3.8}$$

$$tr(\overline{A})=\sum_{i=1}^{n}\overline{A}_{ii}=\overline{A}_{11}+\overline{A}_{22}+\ldots+\overline{A}_{n-1,n-1}+\overline{A}_{nn}. \tag{3.9}$$

Примечание 3.1 (П3.1). Матричным *инвариантом* являются также ранги $rang(A)$ и $rang(\overline{A})$ подобных матриц A и $\overline{A}$, как число линейно независимых столбцов, как размерности $\dim(\mathrm{Im}(A))$ и $\dim(\mathrm{Im}(\overline{A}))$ образов этих матриц, определяемые числом *ненулевых* собственных значений подобных матриц A и $\overline{A}$.

Примечание 3.2 (П3.2). Необходимым, но недостаточным условием подобия двух матриц A и $\overline{A}$ являются равенства (3.6) детерминантов матриц и (3.7) следов матриц. Достаточным условием подобия матриц A и $\overline{A}$ являются равенства (3.3) равенства характеристических полиномов матриц и (3.5) совпадения алгебраических спектров собственных значений этих матриц.

Рассмотрим ***матричные неинварианты*** на примере двух подобных матриц A и $\overline{A}$ вида (3.2).

1. Первым матричным неинвариантом являются геометрические спектры $\{\xi_i; i=\overline{1,n}\}$ и $\{\overline{\xi}_i; i=\overline{1,n}\}$ собственных векторов ξ_i и $\overline{\xi}_i$ подобных матриц A и $\overline{A}$, задаваемых в форме

$$\xi_i = \arg\left\{A\xi_i = \lambda_i\xi_i; \lambda_i \in \sigma\left\{A\right\}; i=\overline{1,n}\right\}, \tag{3.10}$$

$$\overline{\xi_i} = \arg\left\{A\overline{\xi}_i = \lambda_i\overline{\xi}_i; \lambda_i \in \sigma\left\{\overline{A}\right\}; i=\overline{1,n}\right\}. \tag{3.11}$$

При этом (как правило) выполняется отношение неравенства

$$\xi_i \neq \overline{\xi}_i. \tag{3.12}$$

2. Вторым матричным *неинвариантом* являются нормы $\|A\|_{(*)}$ и $\|\overline{A}\|_{(*)}$ подобных матриц A и $\overline{A}$, которые в зависимости от индекса (*) нормы задаются в формах.

2.1. Евклидова (Фробениусова) норма

$$\|A\|_E = \|A\|_F = \left\{\sum_{i,j=1}^{n}\left|A_{ij}\right|^2\right\}^{1/2}; \|\overline{A}\|_E = \|\overline{A}\|_F = \left\{\sum_{i,j=1}^{n}\left|\overline{A}_{ij}\right|^2\right\}^{1/2} \tag{3.13}$$

При этом (как правило) выполняется отношение неравенства

$$\|A\|_E\left(\|A\|_F\right) \neq \|\overline{A}\|_E\left(\|\overline{A}\|_F\right) \tag{3.14}$$

2.2. Операторные (индуцированные) нормы

$$\|A\|_p = \max_x \frac{\|Ax\|_p}{\|x\|_p}; \ \|\overline{A}\|_p = \max_x \frac{\|\overline{A}x\|_p}{\|x\|_p} \tag{3.15}$$

2.2.1. При $p=1$ $\|A\|_{p=1}; \|\overline{A}\|_{p=1}$ – столбцовые нормы A и $\overline{A}$;

2.2.2. При $p \Rightarrow \infty$ $\|A\|_{p=\infty}; \|\overline{A}\|_{p=\infty}$ – строчные нормы A и $\overline{A}$;

2.2.3. При $p=2$ $\|A\|_{p=2}; \|\overline{A}\|_{p=2}$ – спектральные нормы A и $\overline{A}$, вычисляемые в силу соотношений

$$\|A\|_2 = \max_x \frac{\|Ax\|_2}{\|x\|_2} = \alpha_M(A) : \alpha_M(A) = \left|\mu_M^{1/2}\right| : \det(\mu I - A^T A) = 0;$$

$$\|\overline{A}\|_2 = \max_x \frac{\|\overline{A}x\|_2}{\|x\|_2} = \alpha_M(\overline{A}) : \alpha_M(\overline{A}) = \left|\mu_M^{1/2}\right| : \det(\mu I - \overline{A}^T\overline{A}) = 0,$$

где $\alpha_M(A)$ и $\alpha_M(\overline{A})$ – соответственно максимальные сингулярные числа матриц A и $\overline{A}$. Приведенные матричные нормы удовлетворяют оценочным неравенствам, конструируемым на примере матрицы A:

$$\|A\|_2 \leq \|A\|_1 \leq \sqrt{n}\|A\|_2;$$

$$\max_{i,j}\left|A_{ij}\right| \le \|A\|_2 \sqrt{nm}\max_{i,j}\left|A_{ij}\right|;$$

$$\frac{1}{\sqrt{n}}\|A\|_\infty \le \|A\|_2 \le \sqrt{m}\|A\|_\infty;$$

$$\frac{1}{\sqrt{m}}\|A\|_\infty \le \|A\|_2 \le \sqrt{n}\|A\|_\infty;$$

$$\|A\|_2 \le \left(\|A\|_1\|A\|_\infty\right)^{1/2}.$$

При этом (как правило) выполняется отношение неравенства

$$\|A\|_p \ne \|\bar{A}\|_p \text{ при } p = 1,2,\infty. \tag{3.16}$$

3. Третьим матричным *неинвариантом* являются алгебраические спектры $\sigma_\alpha\{A\}$ и $\sigma_\alpha\{\bar{A}\}$ сингулярных чисел подобных матриц A и $\bar{A}$. С понятием сингулярное число связана процедура сингулярного разложения матриц (процедура *SVD–разложения*), которую рассмотрим для общего случая.

Определение 3.3 (О3.3). Сингулярным разложением вещественнозначной матрицы N размерности $(m \times n)$ называется ее *факторизация*, задаваемая в виде

$$N = U\Sigma V^T, \tag{3.17}$$

где U – ортогональная $(m \times m)$ матрица, V – ортогональная $(n \times n)$ матрица, образующие соответственно левый и правый сингулярные базисы и обладающие свойствами:

$$UU^T = U^T U = I, \quad VV^T = V^T V = I. \tag{3.18}$$

Σ – матрица сингулярных чисел α_i, которая принимает вид

$$\Sigma = diag\{\alpha_i; i = \overline{1,m}\} \text{ при } m = n, \tag{3.19}$$

$$\Sigma = [diag\{\alpha_i; i = \overline{1,m}\} \mid 0_{m,n-m}] \text{ при } m < n, \tag{3.20}$$

$$\Sigma = \begin{bmatrix} diag\{\alpha_i; i = \overline{1,n}\} \\ \hline 0_{m-n,n} \end{bmatrix} \text{ при } m > n. \tag{3.21}$$

Положим пока $m = n$ и протранспонируем матричное выражение (3.17), тогда получим

$$N^T = V\Sigma^T U^T\Big|_{n=m} = V\Sigma U^T \tag{3.22}$$

Умножим (3.17) на (3.22), тогда с использованием свойства (3.18) получим цепочку равенств

$$NN^T = U\Sigma V^T V\Sigma U^T = V\Sigma^2 U^T. \tag{3.23}$$

Теперь умножим (3.22) слева на (3.17), получим:

$$N^T N = V\Sigma U^T U\Sigma V^T = V\Sigma^2 V^T. \tag{3.24}$$

Умножим матричное уравнение (3.23) на матрицу U справа, тогда с учетом (3.18) получим матричное соотношение:

$$NN^T U = U\Sigma^2 . \qquad (3.25)$$

Перейдем в (3.25) к столбцовой форме записи правых матричных компонентов:

$$NN^T \left\{col(U_i; i = \overline{1,m}\right\} = U\left\{col(\Sigma^2)_i; i = \overline{1,m})\right\},$$

что эквивалентно матрично-векторному представлению

$$NN^T U_i = U(\Sigma^2)_i; i = \overline{1,m} . \qquad (3.26)$$

Если учесть, что столбец $(\Sigma^2)_i$ имеет вид

$$(\Sigma^2)_i = \left[0_{1\times(i-1)} \;\vdots\; \alpha_i^2 \;\vdots\; 0_{1\times(m-i)}\right]^T, \qquad (3.27)$$

то с учетом (3.27) соотношение (3.26) записывается

$$NN^T U_i = \alpha_i^2 U_i; i = \overline{1,m} . \qquad (3.28)$$

Векторно-матричное соотношение (3.28) представляет собой полное решение проблемы собственных значений α_i^2 и собственных векторов U_i матрицы NN^T. В результате чего получаем, что $\alpha_i^2 (i = \overline{1,m})$ ищутся как решения характеристического уравнения

$$det(\alpha^2 I - NN^T) = 0, \qquad (3.29)$$

а матрица U оказывается составленной из собственных векторов U_i матрицы NN^T единичной нормы в форме

$$U = row\left\{U_i : \|U_i\| = 1; i = \overline{1,m}\right\}. \qquad (3.30)$$

Умножим теперь матричное уравнение (3.24) на матрицу V справа, тогда с учетом (3.18) получим:

$$N^T NV = V\Sigma^2 . \qquad (3.31)$$

По аналогии с (3.26) ÷ (3.29) соотношение (3.31) запишем в форме m матрично-векторных выражений:

$$N^T NV_i = \alpha_i^2 V_i; i = \overline{1,m} , \qquad (3.32)$$

которые представляют собой задачу на собственные значения α_i^2 и собственные векторы V_i матрицы NN^T. Последнее позволяет составить характеристическое уравнение

$$\det(\alpha^2 I - N^T N) = 0 , \qquad (3.33)$$

позволяющее вычислить все $\alpha_i^2 (i = \overline{1,m})$, знание которых в силу (3.32) позволяет найти собственные векторы V_i единичной нормы матрицы $N^T N$. Матрица V правого сингулярного базиса в итоге по аналогии с (3.30) записывается в форме:

$$V = row\left\{V_i : \|V_i\| = 1; i = \overline{1,m}\right\}. \qquad (3.34)$$

Следует заметить, что в случае $m = n$ матрицы NN^T и $N^T N$ обладают одним и тем же спектром собственных значений так, что

$\sigma\left\{NN^T\right\}=\sigma\left\{N^T N\right\}=\left\{\alpha_i^2; i=\overline{1,m}\right\}$. Если $m \neq n$, то спектр $\sigma\{N^T N\}$ содержит n собственных значений, а спектр $\sigma\{NN^T\}$ содержит m собственных значений, причем количество *ненулевых* элементов этих спектров оказываются равными.

Дадим теперь *геометрическую* интерпретацию сингулярного разложения матрицы N (3.17). Для этой цели умножим (3.17) на матрицу V справа и воспользуемся свойствами (3.18), тогда получим:

$$NV = U\Sigma. \tag{3.35}$$

Запишем (3.35) по аналогии с (3.28) и (3.32) в столбцовой форме

$$NV_i = \alpha_i U_i; i=\overline{1,m}. \tag{3.36}$$

Сконструируем теперь на векторно-матричном соотношении (3.36) согласованные тройки $\left\{V_i, \alpha_i, U_i; i=\overline{1,m}\right\}$, которые несут информацию о том, что в силу (3.36) эффект действия оператора с матрицей N на *i*-й элемент V_i правого сингулярного базиса V состоит в умножении на *i*-ое сингулярное число α_i *i*-го элемента U_i левого сингулярного базиса U.

Если теперь с помощью матрицы N в силу линейного векторно-матричного соотношения:

$$\eta = N\chi \tag{3.37}$$

отобразить *сферу* $\|\chi\|=1$, то она отобразится в *эллипсоид*, *положение* полуосей которого *определяется* элементами U_i левого сингулярного базиса U, а *длины* этих полуосей в силу (3.36) будут равны $\alpha_i\|\chi\|$.

В заключение заметим, что в англоязычной литературе сингулярное разложение матриц именуется SVD–разложением (SVD–процедурой). Во всех версиях пакета MATLAB существует функция SVD(*N*), которая выводит матричные компоненты факторизации (3.17).

Возвращаясь к проблеме матричных *неинвариантов*, следует констатировать, что если в качестве $(n \times n)$ – матрицы N взять подобные матрицы A и $\overline{A}$, то, как правило, на спектрах их сингулярных чисел выполняется отношение неравенства $\sigma_\alpha\{A\} \neq \sigma_\alpha\left\{\overline{A}\right\}$.

4. Четвертым матричным *неинвариантом* являются числа обусловленности $cond\{A\}=C\{A\}$, $cond\left\{\overline{A}\right\}=C\left\{\overline{A}\right\}$ подобных матриц A и $\overline{A}$.

Дадим определение числу обусловленности произвольной $(n \times n)$–матрице N, первое из них будет *геометрическим*, а второе – *алгебраическим*.

Определение 3.4 (О3.4). Числом обусловленности $cond\{A\}=C\{A\}$ произвольной $(n \times n)$–матрицы N называется положительнозначная скалярная характеристика этой матрицы, задаваемая в форме

$$C\{N\} = \|N\|_p \left\|N^{-1}\right\|_p \tag{3.38}$$

Примечание 3.3 (П3.3). Численно значение числа обусловленности матрицы зависит от типа используемой в (3.38) матричной нормы. Если в (3.38) используется спектральная норма матриц ($p = 2$), то выполняются соотношения

$$\|N\|_{p=2} = \alpha_M(N), \left\|N^{-1}\right\|_{p=2} = \alpha_m^{-1}(N). \tag{3.39}$$

Число обусловленности (3.38), построенное на спектральных нормах (3.39), принимает вид

$$C\{N\} = \|N\|_{p=2} \left\|N^{-1}\right\|_{p=2} = \alpha_M(N)\alpha_m^{-1}(N). \tag{3.40}$$

Выражение (3.40) показывает, что число обусловленности матрицы N линейной алгебраической задачи (3.37) *геометрически* характеризует степень *сплющивания* эллипсоида, получаемого при отображении сферы $\|\chi\| = 1$ единичного радиуса.

Алгебраическое определение числа обусловленности $C\{N\}$ матрицы N введем на базе следующего утверждения.

Утверждение 3.2 (У3.2). Число обусловленности матрицы N, заданное в форме (3.38), содержательно представляет собой *коэффициент усиления относительных погрешностей* задания (знания) компонентов правой части ЛАЗ (3.37) в относительную погрешность ее левой части. □

Доказательство. Введем в рассмотрение помимо номинальной версии ЛАЗ (3.37) ее возмущенную версию

$$\eta + \Delta\eta = (N + \Delta N)(\chi + \Delta\chi). \tag{3.41}$$

Перейдем от задачи (3.41) к задаче в абсолютных приращениях, которая с использованием (3.37) и (3.41) запишется в форме

$$\Delta\eta = \Delta N\chi + N\Delta\chi + \Delta N\Delta\chi \tag{3.42}$$

Переход в (3.42) к согласованным матричным и векторным нормам позволяет записать

$$\|\Delta\eta\| \le \|\Delta N\|\|\chi\| + \|N\|\|\Delta\chi\| + \|\Delta N\|\|\Delta\chi\|. \tag{3.43}$$

Введем в рассмотрение относительные погрешности представления компонентов ЛАЗ и ее решения, определив их следующими соотношениями:

$$\delta_\eta \stackrel{\Delta}{=} \frac{\|\Delta\eta\|}{\|\eta\|};\ \delta_N \stackrel{\Delta}{=} \frac{\|\Delta N\|}{\|N\|};\ \delta_\chi \stackrel{\Delta}{=} \frac{\|\Delta\chi\|}{\|\chi\|}. \tag{3.44}$$

Свяжем относительные погрешности (3.44) аналитической зависимостью, опираясь на соотношение (3.43). Для этих целей представим номинальную версию ЛАЗ (3.37) в форме

$$\chi = N^{-1}\eta,$$

которая в согласованных матричных и векторных нормах позволяет записать

$$\|\eta\| \ge \frac{1}{\left\|N^{-1}\right\|}\|\chi\|. \tag{3.45}$$

Деление левой части неравенства (3.43) на левую часть второго неравенства (3.45) и соответственно правой части неравенства (3.43) на правую часть второго неравенства (3.45) усиливает выполнение условий исходного неравенства (3.43) и принимает вид

$$\frac{\|\Delta\eta\|}{\|\eta\|} \le \|N\|\left\|N^{-1}\right\|\left\{\frac{\|\Delta N\|}{\|N\|} + \frac{\|\Delta\chi\|}{\|\chi\|} + \frac{\|\Delta N\|}{\|N\|}\frac{\|\Delta\chi\|}{\|\chi\|}\right\}. \tag{3.46}$$

Если в (3.46) учесть (3.38), а также выражение для относительных погрешностей (3.44), то неравенство (1.8) примет вид

$$\delta_\eta \le C_N(\delta_N + \delta_\chi + \delta_N\delta_\chi). \qquad \blacksquare(3.47)$$

Таким образом, *алгебраическое* определение числа обусловленности матрицы совпадает с выдвинутым положением утверждения 3.2 и имеет следующую формулировку.

Определение 3.5 (О3.5). Число обусловленности, заданное в форме (3.38), произвольной квадратной матрицы N, порождающей линейную алгебраическую задачу вида (3.37), содержательно представляет собой *коэффициент усиления относительных погрешностей* задания (знания) компонентов правой части ЛАЗ (3.37) в относительную погрешность ее левой части. □

В заключение заметим, что числа обусловленности $C\{A\}$ *и* $C\{\bar{A}\}$ подобных матриц A и $\bar{A}$, являясь матричными *неинвариантами*, как правило, связаны отношением неравенства $C\{A\} \ne C\{\bar{A}\}$.

Примеры и задачи

Выбрать из приводимых ниже матриц пару подобных путем вычисление матричных *инвариантов*, в случае положительного исхода выбора вычислить все матричные *инварианты* и *неинварианты* этих матриц.

3.1. $\begin{bmatrix} -3 & 0 \\ 0 & -5 \end{bmatrix}$ 3.2. $\begin{bmatrix} 7 & 0 \\ 0 & -3 \end{bmatrix}$ 3.3. $\begin{bmatrix} 1 & 0 \\ 0 & -5 \end{bmatrix}$

3.4. $\begin{bmatrix} -2 & 0 \\ 0 & 6 \end{bmatrix}$ 3.5. $\begin{bmatrix} -3 & 0 \\ 7 & -5 \end{bmatrix}$ 3.6. $\begin{bmatrix} 0 & 1 \\ 21 & 4 \end{bmatrix}$

3.7. $\begin{bmatrix} 0 & 1 \\ 5 & -4 \end{bmatrix}$ 3.8. $\begin{bmatrix} 0 & 1 \\ 12 & 4 \end{bmatrix}$ 3.9. $\begin{bmatrix} -3 & 9 \\ 0 & -5 \end{bmatrix}$

3.10. $\begin{bmatrix} 0 & 21 \\ 1 & 4 \end{bmatrix}$ 3.11. $\begin{bmatrix} 0 & 5 \\ 1 & -4 \end{bmatrix}$ 3.12. $\begin{bmatrix} 0 & 12 \\ 1 & 4 \end{bmatrix}$

3.13. $\begin{bmatrix} 0 & 1 \\ -15 & -8 \end{bmatrix}$ 3.14. $\begin{bmatrix} -1 & 2 \\ 8 & 5 \end{bmatrix}$ 3.15. $\begin{bmatrix} -1 & 4 \\ 2 & -3 \end{bmatrix}$

3.16. $\begin{bmatrix} -2 & 0 \\ 7 & 6 \end{bmatrix}$ 3.17. $\begin{bmatrix} 0 & -15 \\ 1 & -8 \end{bmatrix}$ 3.18. $\begin{bmatrix} 7 & 0 \\ 8 & -3 \end{bmatrix}$

3.19. $\begin{bmatrix} -3 & 8 \\ 1 & -1 \end{bmatrix}$ 3.20. $\begin{bmatrix} 6 & 10 \\ 0 & -2 \end{bmatrix}$ 3.21. $\begin{bmatrix} -1 & 1 \\ -8 & -7 \end{bmatrix}$

3.22. $\begin{bmatrix} 5 & 16 \\ 1 & -1 \end{bmatrix}$ 3.23. $\begin{bmatrix} 5 & -8 \\ 5 & -9 \end{bmatrix}$ 3.24. $\begin{bmatrix} 12 & -7 \\ 12 & -8 \end{bmatrix}$

3.25. $\begin{bmatrix} 1 & -1 \\ 24 & -9 \end{bmatrix}$ 3.26. $\begin{bmatrix} 5 & 8 \\ 2 & -1 \end{bmatrix}$ 3.27. $\begin{bmatrix} 5 & 8 \\ -5 & -9 \end{bmatrix}$

3.28. $\begin{bmatrix} 12 & 3.5 \\ 24 & -8 \end{bmatrix}$ 3.29. $\begin{bmatrix} -15 & 4 \\ -30 & 7 \end{bmatrix}$ 3.30. $\begin{bmatrix} 21 & -16 \\ 21 & -17 \end{bmatrix}$

3.31. $\begin{bmatrix} 5 & -4 \\ 10 & -9 \end{bmatrix}$ 3.32. $\begin{bmatrix} 6 & 0 \\ 6 & -2 \end{bmatrix}$ 3.33. $\begin{bmatrix} -15 & 8 \\ -15 & 7 \end{bmatrix}$

3.34. $\begin{bmatrix} 21 & 16 \\ -21 & -17 \end{bmatrix}$ 3.35. $\begin{bmatrix} -3 & 2.67 \\ 3 & -1 \end{bmatrix}$ 3.36. $\begin{bmatrix} -3.33 & -5.33 \\ 2.33 & 7.33 \end{bmatrix}$

3.37. $\begin{bmatrix} -0.5 & 0.5 \\ -22.5 & -7.5 \end{bmatrix}$ 3.38. $\begin{bmatrix} 21 & -8 \\ 42 & -17 \end{bmatrix}$ 3.39. $\begin{bmatrix} 2.5 & -4.5 \\ 2.5 & -6.5 \end{bmatrix}$

3.40. $\begin{bmatrix} 7.33 & -4.67 \\ -2.67 & -3.33 \end{bmatrix}$

Решение вариантов задач

Решение задачи на примере пары матриц 3.15 и 3.7. Выдвинем *гипотезу*, что матрицы

$A = \begin{bmatrix} -1 & 4 \\ 2 & -3 \end{bmatrix}$ и $\overline{A} = \begin{bmatrix} 0 & 1 \\ 5 & -4 \end{bmatrix}$ подобны.

Вычислим матричные *инварианты* этих матриц:

1. Характеристические полиномы, которые принимают вид

$$\det(\lambda I - A) = \det\begin{bmatrix} \lambda+1 & -4 \\ -2 & \lambda+3 \end{bmatrix} = (\lambda+1)(\lambda+3) = \lambda^2 + 4\lambda - 5;$$

$$\det(\lambda I - \overline{A}) = \det\begin{bmatrix} \lambda & -1 \\ -5 & \lambda+4 \end{bmatrix} = \lambda(\lambda+4) - 5 = \lambda^2 + 4\lambda - 5.$$

Гипотеза верна, так как $\det(\lambda I - A) = \det(\lambda I - \overline{A})$, поэтому выбранные матрицы A и $\overline{A}$ подобны.

2. Алгебраические спектры собственных значений матриц

$\sigma\{A\} = \sigma\{\overline{A}\} = \{\lambda_1 = 1; \lambda_2 = -5 : \lambda^2 + 4\lambda - 5 = 0;\}$;

3. Определители (детерминанты) матриц

$$\det(A) = \det\begin{bmatrix} -1 & 4 \\ 2 & -3 \end{bmatrix} = (-1)(-3) - (2)(4) = 4 - 8 = -5,$$

$$\det(\overline{A}) = \det\begin{bmatrix} 0 & 1 \\ 5 & -4 \end{bmatrix} = (0)(-4) - (1)(5) = -5,$$

$$\det(A) = \det(\overline{A}) = \lambda_1\lambda_2 = (1)(-5) = -5,$$

$$\det(A) = \det(\overline{A}).$$

4. Следы матриц

$$tr(A) = tr\begin{bmatrix} -1 & 4 \\ 2 & -3 \end{bmatrix} = \sum_{i=1}^{n} A_{ii} = A_{11} + A_{22} = (-1) + (-3) = -4,$$

$$tr(\overline{A}) = tr\begin{bmatrix} 0 & 1 \\ 5 & -4 \end{bmatrix} = \sum_{i=1}^{n} \overline{A}_{ii} = \overline{A}_{11} + \overline{A}_{22} = (-1) + (-3) = -4,$$

$$tr(A) = tr(\overline{A}) = \sum_{i=1}^{n} \lambda_i = \lambda_1 + \lambda_2 = (1) + (-5) = -4.$$

Вычислим матричные *неинварианты* этих матриц:

1. Спектры собственных векторов $\{\xi_i; i = \overline{1,n}\}$ и $\{\overline{\xi}_i; i = \overline{1,n}\}$

$$A\xi_i = \lambda_i\xi_i \Rightarrow \begin{bmatrix} -1 & 4 \\ 2 & -3 \end{bmatrix}\xi_i = \lambda_i\xi\,|_{\lambda_1 = 1, \lambda_2 = -5} \Rightarrow \xi_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \xi_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix};$$

$$\overline{A}\,\overline{\xi}_i = \lambda_i\overline{\xi}_i \Rightarrow \begin{bmatrix} 0 & 1 \\ 5 & -4 \end{bmatrix}\overline{\xi}_i = \lambda_i\overline{\xi}\,|_{\lambda_1 = 1, \lambda_2 = -5} \Rightarrow \overline{\xi}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \overline{\xi}_2 = \begin{bmatrix} 1 \\ -5 \end{bmatrix}.$$

$\{\xi_i; i = \overline{1,n}\} \neq \{\overline{\xi}_i; i = \overline{1,n}\}$.

2. Нормы $\|A\|_{(*)}$ и $\|\overline{A}\|_{(*)}$.

2.1. Евклидовы (Фробениусовы) матричные нормы

$$\|A\|_E = \|A\|_F = \left\{\sum_{i,j=1}^{n} |A_{ij}|^2\right\}^{1/2} = \{1^2 + 2^2 + 4^2 + 3^2\}^{1/2} = 30^{1/2}$$

$$\|\overline{A}\|_E = \|\overline{A}\|_F = \left\{\sum_{i,j=1}^{n} |\overline{A}_{ij}|^2\right\}^{1/2} = \{0^2 + 5^2 + 1^2 + 4^2\}^{1/2} = 42^{1/2}$$

2.2. Операторные (индуцированные) нормы

$$\|A\|_p = \max_x \frac{\|Ax\|_p}{\|x\|_p}; \ \|\overline{A}\|_p = \max_x \frac{\|\overline{A}x\|_p}{\|x\|_p}$$

2.2.1. При $p = 1$ $\|A\|_{p=1}; \|\overline{A}\|_{p=1}$ – столбцовые нормы.

$$\|A\|_{p=1} = \max_i \{\|A_i\|_1\} = \max\{(|-1| + |2|), (|4| + |-3|)\} = 7,$$

$$\|\overline{A}\|_{p=1} = \max_i \{\|\overline{A}_i\|_1\} = \max\{(|0| + |5|), (|1| + |-4|)\} = 5.$$

2.2.2. При $p = \infty$ $\|A\|_{p=\infty}; \|\overline{A}\|_{p=\infty}$ – строчные нормы.

$$\|A\|_{p=\infty} = \max_j \{\|A^j\|_1\} = \max\{(|-1| + |4|), (|2| + |-3|)\} = 5,$$

$$\|\overline{A}\|_{p=\infty} = \max_j \left\{\|\overline{A}^j\|_1\right\} = \max\{(|0| + |1|), (|5| + |-4|)\} = 9.$$

2.2.3. При $p = 2$ $\|A\|_{p=2}; \|\overline{A}\|_{p=2}$ – спектральные нормы A и $\overline{A}$, вычисляемые в силу соотношений

$$\|A\|_2 = \max_x \frac{\|Ax\|_2}{\|x\|_2} = \alpha_M(A) : \alpha_M(A) = \left|\mu_M^{1/2}\right| : \det(\mu I - A^T A) = 0;$$

$$A^T A = \begin{bmatrix} -1 & 2 \\ 4 & -3 \end{bmatrix} \begin{bmatrix} -1 & 4 \\ 2 & -3 \end{bmatrix} = \begin{bmatrix} 5 & -10 \\ -10 & 25 \end{bmatrix};$$

$$\det(\mu I - A^T A) = \mu^2 - 30\mu + 25 = 0;$$

$$\mu_1 = 29.42; \ \mu_2 = 0.858; \ \alpha_M(A) = \alpha_1 = 5.424, \alpha_m(A) = \alpha_2 = 0.9263,$$

$$\|A\|_2 = \max_x \frac{\|Ax\|_2}{\|x\|_2} = \alpha_M(A) = 5.424.$$

$$\|\overline{A}\|_2 = \max_x \frac{\|\overline{A}x\|_2}{\|x\|_2} = \alpha_M(\overline{A}) : \alpha_M(\overline{A}) = \left|\overline{\mu}_M^{1/2}\right| : \det(\overline{\mu} I - \overline{A}^T \overline{A}) = 0.$$

$$\overline{A}^T \overline{A} = \begin{bmatrix} 0 & 5 \\ 1 & -4 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 5 & -4 \end{bmatrix} = \begin{bmatrix} 25 & -20 \\ -20 & 17 \end{bmatrix};$$

$$\det(\overline{\mu} I - \overline{A}^T \overline{A}) = \overline{\mu}^2 - 42\overline{\mu} + 25 = 0;$$

$$\overline{\mu}_1 = 41.396; \ \overline{\mu}_2 = 0.604; \ \alpha_M(\overline{A}) = \overline{\alpha}_1 = 6.434, \alpha_m(\overline{A}) = \overline{\alpha}_2 = 0.7772,$$

$$\|\overline{A}\|_2 = \max_x \frac{\|\overline{A}x\|_2}{\|x\|_2} = \alpha_M(\overline{A}) = 6.434.$$

3. Алгебраические спектры $\sigma_\alpha\{A\}$ и $\sigma_\alpha\{\overline{A}\}$ сингулярных чисел подобных матриц A и $\overline{A}$ вычислены в предыдущем пункте и имеют представления

$$\sigma_{\alpha}\{A\}=\{\alpha_1=5.424;\alpha_2=0.9263\},$$

$$\sigma_{\alpha}\{\overline{A}\}=\{\overline{\alpha}_1=6.434;\overline{\alpha}_2=0.7772\}.$$

4. Спектральные числа обусловленности подобных матриц A и $\overline{A}$, вычисляемые в силу соотношений

$$C_{\lambda}\{A\}=C_2\{A\}=\alpha_M(A)\alpha_m^{-1}(A)=5.424(0.9263)^{-1}=5.8557,$$

$$C_{\lambda}\{\overline{A}\}=C_2\{\overline{A}\}=\alpha_M(\overline{A})\alpha_m^{-1}(\overline{A})=6.434(0.7772)^{-1}=8.2787.$$

■

4. КАНОНИЧЕСКИЕ ФОРМЫ МАТРИЦ. МАТРИЦЫ ПРИВЕДЕНИЯ ПОДОБИЯ

Рассматриваются подобные матрицы A и $\overline{A}$, связанные матричным условием подобия (3.1) с матрицей M приведения подобия. Будем полагать, что $(n \times n)-$матрица A задана в *произвольном базисе* (имеет произвольную форму), а $(n \times n)-$матрица $\overline{A}$ задана в *каноническом базисе* (имеет каноническую форму). В связи со сказанным встают два вопроса. *Первый вопрос*: как формируются матрицы в *канонической* форме? *Второй вопрос*: как формируется матрица M приведения подобия, позволяющая с помощью матричного соотношения

$$\overline{A} = M^{-1} A M \tag{4.1}$$

осуществить переход от матрицы A, заданной в произвольном базисе к матрице $\overline{A}$, задаваемой в некотором каноническом базисе?

Дадим *ответ* на *первый* из поставленных вопросов, то есть построим матрицы ЛО ***A*** в *канонических* формах.

Определение 4.1 (О4.1). Канонической формой $(n \times n)$ матрицы линейного оператора ***A*** будем называть форму $(n \times n)$ матрицы линейного оператора (ЛО), которая построена в соответствии с некоторым правилом (законом, *каноном*) с тем, чтобы решить одну из возможных задач: сокращение объема матричных вычислений путем минимизации числа *ненулевых* элементов матрицы; облегчение анализа *структуры пространства* ЛО ***A***, обеспечение вычислительной устойчивости всех матричных процедур путем уменьшения числа обусловленности матрицы ЛО и т.д.

К настоящему моменту сконструировано большое число канонических форм задания $(n \times n)$ матрицы линейного оператора ***A***, ниже рассматриваются только базовые канонические формы.

Базовые канонические формы $(n \times n)$ матрицы линейного оператора ***A*** строятся на двух алгебраических спектрах исходной матрицы A, заданной в *произвольном базисе*.

Первый алгебраический спектр

$$\sigma\{A\} = \left\{\lambda_i : A\xi_i = \lambda_i \xi_i : \det\left(\lambda_i I - A\right)\right) = 0 : i = \overline{1,n}\}$$

представляет собой спектр собственных значений $\left\{\lambda_i : i = \overline{1,n}\right\}$ матрицы A.

Второй алгебраический спектр

$$\sigma_a\{A\} = \{a_i : \det(\lambda I - A) = \lambda^n + a_1\lambda^{n-1} + \ldots + a_{n-1}\lambda + a_n; i = \overline{1,n}\}$$

представляет спектр коэффициентов $\left\{a_i : i = \overline{1,n}\right\}$ характеристического полинома $D(\lambda) = \det\left(\lambda I - A\right)$ матрицы A.

Рассмотрим канонические представления $\overline{A}$ исходной матрицы A, которые конструируются на алгебраическом спектре собственных значений матриц, для различных случаев его реализации.

1. *Диагональная каноническая форма* матрицы может быть построена, когда алгебраический спектр собственных значений имеет реализацию

$$\sigma\{A\} = \left\{\lambda_i : \mathrm{Jm}(\lambda_i) = 0; \lambda_i \neq \lambda_l; i \neq l; i,l = \overline{1,n}\right\}. \tag{4.2}$$

Алгебраический спектр вида (4.2) порождает множество подобных матриц линейного оператора ***A***, именуемых матрицами *простой структуры.*

В случае реализации алгебраического спектра $\sigma\{A\}$ в форме (4.2), когда все собственные значения *вещественные и простые (различные, не кратные)*, может быть построена диагональная матрица Λ с элементами λ_i на главной диагонали и нулями на остальных позициях этой матрицы так, что она принимает вид

$$\overline{A} = \Lambda = \begin{bmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_n \end{bmatrix} = diag\{\lambda_i, i = \overline{1,n}\}. \tag{4.3}$$

2. *Блочно-диагональная каноническая форма* матрицы может быть построена, когда алгебраический спектр собственных значений имеет реализацию

$$\sigma\{A\} = \left\{\mathrm{Jm}\,(\lambda_{2i-1,2i}) \neq 0 : \lambda_{2i-1} = \alpha_i + j\beta_i; \lambda_{2i} = \alpha_i - j\beta_i : \lambda_i \neq \lambda_l; i \neq l; i,l = \overline{1,n/2}\right\} \tag{4.4}$$

В случае реализации алгебраического спектра $\sigma\{A\}$ в форме (4.4), когда все собственные значения *комплексно-сопряженные и простые (не кратные)*, может быть построена блочно-диагональная матрица $\tilde{\Lambda}$ с вещественнозначными матричными блоками $\tilde{\Lambda}_{ii} = \begin{bmatrix} \alpha_i & \beta_i \\ -\beta_i & \alpha_i \end{bmatrix}$ на главной диагонали и нулями на остальных позициях этой матрицы так, что она принимает вид

$$\overline{A} = \tilde{\Lambda} = diag\left\{\tilde{\Lambda}_{ii} = \begin{bmatrix} \alpha_i & \beta_i \\ -\beta_i & \alpha_i \end{bmatrix}; i = \overline{1,n/2};\right\}. \tag{4.5}$$

3. *Комбинированная блочно-диагональная каноническая форма* матрицы может быть построена, когда алгебраический спектр собственных значений содержит только *простые* собственные значения, часть которых числом n_R являются *вещественными,* а другая часть числом

$n_c = 2m_c$ – *комплексно-сопряженными,* при этом выполняется соотношение $n = n_c + n_R$.

Комбинированная блочно-диагональная матрица имеет на своей главной диагонали диагональную матрицу вида (4.3) размерности $(n_R \times n_R)$ и блочно-диагональную матрицу вида (4.5) размерности $(n_c \times n_c)$ так, что она примет вид

$$\overline{A} = \tilde{\tilde{\Lambda}} = diag\ \left\{\Lambda_{(n_R \times n_R)}; \tilde{\Lambda}_{(n_c \times n_c)}\right\}. \qquad (4.6)$$

Матричные блоки на диагонали комбинированной блочно-диагональной матрицы можно менять местами, так что наряду с формой (4.6) матрица $\overline{A}$ может иметь представление

$$\overline{A} = \tilde{\Lambda} = diag\{\tilde{\Lambda}_{(n_c \times n_c)}; \Lambda_{(n_R \times n_R)}\}. \qquad (4.7)$$

Так, например, если алгебраический спектр собственных значений матриц ЛО ***A*** имеет реализацию

$$\sigma\{A\} = \left\{\lambda_{1,2} = \alpha \pm j\beta; \lambda_i : Jm(\lambda_i) = 0; \lambda_i \neq \lambda_l; i \neq l; i, l = \overline{3, n}\right\}, \qquad (4.8)$$

то комбинированное блочно-диагональное представление канонической матрицы $\overline{A}$ принимает вид

$$\overline{A} = \tilde{\tilde{\Lambda}} = \begin{bmatrix} \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix} & \begin{bmatrix} 0_{2\times(n-2)} \end{bmatrix} \\ \begin{bmatrix} 0_{(n-2)\times 2} \end{bmatrix} & \begin{bmatrix} \Lambda_{(n-2)\times(n-2)} \end{bmatrix} \end{bmatrix}, \qquad (4.9)$$

где $0_{2\times(n-2)}, 0_{(n-2)\times 2}, \Lambda_{(n-2)\times(n-2)}$ – соответственно нулевые матрицы размерности $2\times(n-2)$ и $(n-2)\times 2$ и диагональная матрица размерности $(n-2)\times(n-2)$.

4. *Жорданова каноническая форма* матрицы может быть построена, когда алгебраический спектр собственных значений имеет реализацию

$$\sigma\{A\} = \left\{\lambda_k - кратности\ \mu_k, k = \overline{1, p}; \sum_{k=1}^{p} \mu_k = n; \mathrm{Jm}(\lambda_k) = 0;\right\}. \qquad (4.10)$$

Тогда *жорданова каноническая форма* матрицы $\overline{A}$, максимально близкая к диагональной форме для случая вещественных кратных собственных значений матриц ЛО ***A***, в соответствии со структурой алгебраического спектра (4.10) примет вид

$$\overline{A} = J = diag\left\{ J_{kk} = \begin{bmatrix} \lambda_k & 1 & 0 & . & 0 & 0 \\ & \lambda_k & 1 & . & 0 & 0 \\ . & . & . & . & . & . \\ 0 & 0 & 0 & . & \lambda_k & 1 \\ 0 & 0 & 0 & . & 0 & \lambda_k \end{bmatrix}_{(\mu_k \times \mu_k)} ; k = \overline{1, p} \right\}. \qquad (4.11)$$

Жорданова каноническая форма (4.11) представляет собой блочно-диагональную матрицу, составленную из *жордановых блоков* J_{kk} размерности $(\mu_k \times \mu_k)$, имеющих на своей главной диагонали собственное значение λ_k кратности μ_k, единицы на первой *наддиагонали* и нули на остальных позициях блока. *Жорданова каноническая форма* (4.11) является верхней жордановой формой, наряду с которой может быть построена *нижняя жорданова каноническая форма*, которая характеризуется тем, что единицы в жордановых блоках размещаются на первой *поддиагонали.* Следует заметить, что *жорданова каноническая форма* может быть построена и для случая матриц ЛО, алгебраический спектр собственных значений которых содержит кратные комплексно-сопряженные элементы, причем возможны как комплексно-значная так и вещественно-значная формы.

5. Рассмотрим теперь *канонические представления* $\overline{A}$ исходной матрицы A, которые конструируются на алгебраическом спектре $\sigma_a\{A\}$ коэффициентов характеристического полинома

$$D(\lambda) = \det(\lambda I - A) = \lambda^{n} + \sum_{i=1}^{n} a_i \lambda^{n-i} \tag{4.12}$$

матриц линейного оператора $\boldsymbol{A}$. Этих представлений два, они совпадают с точностью до транспонирования. Канонические представления имеют вид

$$\overline{A} = \begin{bmatrix} 0 & 1 & 0 & . & 0 & 0 \\ 0 & 0 & 1 & . & 0 & 0 \\ . & . & . & . & . & . \\ 0 & 0 & 0 & . & 0 & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & . & -a_2 & -a_1 \end{bmatrix} = \begin{bmatrix} 0_{(n-1)\times 1} & I_{(n-1)\times(n-1)} \\ & \\ & -a \end{bmatrix} \tag{4.13}$$

и

$$\overline{A}^{T} = \begin{bmatrix} 0 & 0 & . & ..0 & -a_n \\ 1 & 0 & . & ..0 & -a_{n-1} \\ 0 & 1 & . & ..0 & -a_{n-2} \\ . & . & . & . & . \\ 0 & 0 & . & ..0 & -a_2 \\ 0 & 0 & . & ..1 & -a_1 \end{bmatrix} = \begin{bmatrix} 0_{1\times(n-1)} & \\ & -a^{T} \\ I_{(n-1)\times(n-1)} & \end{bmatrix}. \tag{4.14}$$

В *канонических формах* (4.13) и (4.14) $a-$ n-мерный вектор–строка коэффициентов, записанных в обратном по отношению их размещения в характеристическом полиноме порядке так, что он принимает вид

$$a = [a_n, a_{n-1}, a_{n-2}, \ldots a_2, a_1] = row\{a_{n+1-i} : i = \overline{1,n}\}, \tag{4.15}$$

$0_{(n-1)\times 1}; 0_{1\times(n-1)}; I_{(n-1)\times(n-1)}$ – соответственно $(n-1)$-мерные матрица-столбец и матрица–строка, а также $(n-1)\times(n-1)$ – единичная матрица.

Обе *канонические формы* (4.13) и (4.14) именуются *нормальной, сопровождающей* (свой характеристический полином) и *фробениусовой канонической* формой. С тем, чтобы их различать текстуально форма (4.13) названа *строчной* нормальной, сопровождающей или фробениусовой, а (4.14) – *столбцовой*. Для канонической формы (4.13) используется обозначение $\overline{A} = A_F$.

Строчная сопровождающая каноническая форма матрицы ЛО $\boldsymbol{A}$ имеет в последней строке коэффициенты характеристического полинома, записанные с обратными знаками и в обратном порядке, первую *наддиагональ*, заполненную единицами, остальные позиции матрицы заполнены нулями. При использовании этой формы матрицы для модельных представлений динамических объектов она именуется *канонической управляемой фробениусовой (сопровождающей) формой.*

Столбцовая сопровождающая каноническая форма матрицы ЛО $\boldsymbol{A}$ имеет в последнем столбце коэффициенты характеристического полинома, записанные с обратными знаками и в обратном порядке, первую *поддиагональ*, заполненную единицами, остальные позиции матрицы, заполненные нулями. При использовании этой формы матрицы для модельных представлений динамических объектов она именуется *канонической наблюдаемой фробениусовой (сопровождающей) формой.*

Теперь дадим ответ на *второй* вопрос, поставленный в начале раздела, то есть построим матрицы приведения подобия произвольной матрицы ЛО $\boldsymbol{A}$ к каноническим формам.

Приведение матрицы A простой структуры ЛО $\boldsymbol{A}$ к диагональной форме (4.3) строится на положениях следующих утверждений.

Утверждение 4.1 (У4.1). Матрица M, приводящая произвольную $(n\times n)$ квадратную матрицу A простой структуры ЛО $\boldsymbol{A}$ к диагональной форме Λ в силу соотношения (4.1), принимающего для $\overline{A} = \Lambda$ представление

$$\Lambda = M^{-1}AM, \tag{4.16}$$

имеет своими столбцами собственные векторы матрицы A.

Доказательство. Запишем базовое уравнение матричного подобия для рассматриваемого случая

$$M\Lambda = AM \tag{4.17}$$

в столбцовой форме

$$M[\Lambda_1 \quad \Lambda_2 \quad \ldots \quad \Lambda_i \quad \ldots \quad \Lambda_n] = A[M_1 \quad M_2 \quad \ldots \quad M_i \quad \ldots \quad M_n] \tag{4.18}$$

где Λ_i, M_i – $i-$*ые* столбцы соответственно матриц Λ и $M, (i=\overline{1,n})$. Перейдем теперь от матричного уравнения (4.18) к $n-$векторно-матричным уравнениям вида

$$M\Lambda_i = AM_i\,;\,\left(i=\overline{1,n}\right),\tag{4.19}$$

где $i-ый$ столбец Λ_i диагональной матрицы Λ имеет вид

$$\Lambda_i = \begin{bmatrix}0 & \dots & 0 & \lambda_i & 0 & \dots & 0\end{bmatrix}^T.\tag{4.20}$$

Нетрудно видеть, что с учетом (4.20) векторно-матричное уравнение (4.19) принимает вид

$$\lambda_i M_i = AM_i\;;i=\overline{1,n}\;.\tag{4.21}$$

Векторно-матричное соотношение (4.21) является определением собственного вектора матрицы A, откуда следует, что M_i – собственный вектор матрицы A. ■

Утверждение 4.2 (У4.2). Пусть матрица A ЛО $\boldsymbol{A}$ является матрицей простой структуры, тогда ее каноническая строчная фробениусова форма A_F, имеющая представление (4.13), обладает собственными векторами

$$\xi_i = \arg\left\{A_F\xi_i = \lambda_i\xi_i\,;i=\overline{1,n}\right\},\tag{4.22}$$

которые строятся по *схеме* Вандермонда так, что они принимают вид

$$\xi_i = \begin{bmatrix}1 & \lambda_i & \lambda_i^2 & \dots & \lambda_i^{n-1}\end{bmatrix}^T;\,i=\overline{1,n}.\tag{4.23}$$

Доказательство сформулированного утверждения строится на непосредственной подстановке в (4.22) представлений (4.13) и (4.23), в результате получается следующая цепочка векторно-матричных равенств

$$A_F\xi_i = \begin{bmatrix}0 & 1 & 0 & \dots & 0\\ 0 & 0 & 1 & \dots & 0\\ \dots & \dots & \dots & \dots & \dots\\ 0 & 0 & 0 & \dots & 1\\ -a_n & -a_{n-1} & -a_{n-2} & \dots & -a_1\end{bmatrix}\begin{bmatrix}1\\ \lambda_i\\ \lambda_i^2\\ ..\\ \lambda_i^{n-1}\end{bmatrix} = \begin{bmatrix}\lambda_i\\ \lambda_i^2\\ \lambda_i^3\\ ..\\ -\sum\limits_{l=1}^{n}a_l\lambda_i^{n-l}\end{bmatrix}.\tag{4.24}$$

Но в силу характеристического уравнения матриц ЛО $\boldsymbol{A}$ оказывается справедливой запись

$$\lambda_i^{\,n} + \sum_{l=1}^{n}a_l\lambda_i^{n-l} = 0,\tag{4.25}$$

из которой следует справедливость представления

$$-\sum_{l=1}^{n}a_l\lambda_i^{n-l} = \lambda_i^{\,n},\tag{4.26}$$

подстановка которого в (4.24) приводит последнее к виду

$$A_F\xi_i = \lambda_i\begin{bmatrix}1 & \lambda_i & \lambda_i^2 & \dots & \lambda_i^{n-1}\end{bmatrix}^T.\tag{4.27}$$

Соотношение (4.27) делает справедливым утверждение 4.2. ■

Доказательство утверждения 4.2. и положения утверждения 4.1 содержат доказательство утверждения 4.3.

Утверждение 4.3 (У4.3). Пусть матрица A_F является канонической строчной фробениусовой формой матриц ЛО ***A*** простой структуры, тогда матрица A_F может быть приведена к канонической диагональной форме (4.3) с помощью матрицы Вандермонда M_B, столбцы которой M_{Bi} $\left(i=\overline{1,n}\right)$ суть собственные вектора вида (4.23) так, что она принимает вид

$$M_B = row\left\{M_{Bi} = \xi_i = \begin{bmatrix}1 & \lambda_i & \lambda_i^2 & \ldots & \lambda_i^{n-1}\end{bmatrix}^T;\ i=\overline{1,n.}\right\} = \\ = \begin{bmatrix} 1 & 1 & \ldots & 1 \\ \lambda_1 & \lambda_2 & \ldots & \lambda_n \\ \lambda_1^2 & \lambda_2^2 & \ldots & \lambda_n^2 \\ \ldots & \ldots & \ldots & \ldots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \ldots & \lambda_n^{n-1} \end{bmatrix} \qquad \blacksquare(4.28)$$

Рассмотрим теперь задачу конструирования матрицы $\widetilde{M}$ приведения исходной матрицы A, обладающей комплексно-значным спектром собственных значений (4.4), к канонической блочно-диагональной форме (4.5) в силу матричного соотношения $\widetilde{\Lambda} = \widetilde{M}^{-1} A \widetilde{M}$.

Утверждение 4.4 (У4.4). Пусть $(n\times n)-$матрица A такова, что алгебраический спектр ее собственных значений составлен из комплексно-сопряженных чисел так, что он имеет вид (4.4). Геометрический спектр собственных векторов этой матрицы составлен из комплексно-сопряженных векторов так, что он имеет вид

$$\{\xi_{2i-1};\xi_{2i};i=\overline{1,n/2}\} = \left\{\begin{matrix}\xi_{2i-1}=\xi_{R,2i-1}+j\xi_{J,2i-1};\ \xi_{2i}=\xi_{R,2i}-j\xi_{J,2i}; \\ A\xi_{2i-1}=\lambda_{2i-1}\xi_{2i-1};\ A\xi_{2i}=\lambda_{2i}\xi_{2i};i=\overline{1,n/2.}\end{matrix}\right\}.$$

Тогда матрица

$$\widetilde{M} = \arg\left\{\widetilde{\Lambda} = \widetilde{M}^{-1} A \widetilde{M} = diag\left\{\widetilde{\Lambda}_{ii} = \begin{bmatrix}\alpha_i & \beta_i \\ -\beta_i & \alpha_i\end{bmatrix}; i=\overline{1,n/2;}\right\}\right\} = \\ = row\left\{\left[\widetilde{M}_{2i-1} \vdots \widetilde{M}_{2i}\right]; i=\overline{1,n/2;}\right\}$$

имеет своими столбцами $\widetilde{M}_{2i-1}, \widetilde{M}_{2i}$ соответственно вещественный и мнимый компоненты собственных векторов, что записывается в форме

$$\widetilde{M}_{2i-1} = \xi_{R,2i-1}, \widetilde{M}_{2i} = \xi_{J,2i-1}. \qquad \square$$

Доказательство утверждения строится на непосредственной подстановке в векторно-матричные соотношения для собственных векторов

$$A\xi_{2i-1}=\lambda_{2i-1}\xi_{2i-1}; A\xi_{2i}=\lambda_{2i}\xi_{2i}; i=\overline{1,n/2}$$

представлений собственных значений и векторов в форме

$$\left\{\lambda_{2i-1}=\alpha_i+j\beta_i;\lambda_{2i}=\alpha_i-j\beta_i:i=\overline{1,n/2}\right\}$$
$$\xi_{2i-1}=\xi_{R,2i-1}+j\xi_{J,2i-1};\xi_{2i}=\xi_{R,2i}-j\xi_{J,2i};i=\overline{1,n/2}.$$

и последующем разделении полученных выражений на вещественный и мнимый компоненты, что в итоге приводит к двум векторно-матричным равенствам

$$(A-\alpha_i I)\xi_{Ri}=-\beta_i\xi_{Ji}\ ;\ (A-\alpha_i I)\xi_{Ji}=\beta_i\xi_{Ri}. \qquad (4.29)$$

В свою очередь, если записать уравнение подобия $\widetilde{M}\widetilde{\Lambda}=A\widetilde{M}$, в котором выделить блоки соответствующие собственным значениям λ_{2i-1} и λ_{2i} то получим

$$\left[\widetilde{M}_{2i-1}\vdots\widetilde{M}_{2i}\right]\begin{bmatrix}\alpha_i & \beta_i\\ -\beta_i & \alpha_i\end{bmatrix}=A\left[\widetilde{M}_{2i-1}\vdots\widetilde{M}_{2i}\right]$$

Решение последнего матричного уравнения относительно матриц-столбцов $\widetilde{M}_{2i-1},\widetilde{M}_{2i}$ приводит к соотношениям

$$(A-\alpha_i I)\widetilde{M}_{2i-1}=-\beta_i\widetilde{M}_{2i}\ ;\ (A-\alpha_i I)\widetilde{M}_{2i}=\beta_i\widetilde{M}_{2i-1}.$$

Сравнение последних соотношений с соотношениями (4.29) делает справедливыми положения утверждения. ■

Примечание 4.1 (ПР4.1). Если спектр собственных значений матриц ЛО ***A*** является комбинированным так, что он содержит как вещественные некратные собственные значения, так и комплексно-сопряженные некратные собственные значения, при этом примера ради имеет реализацию вида (4.8) $\sigma\{A\}=\left\{\lambda_{1,2}=\alpha\pm j\beta;\lambda_i:Jm(\lambda_i)=0;\lambda_i\neq\lambda_l;i\neq l;i,l=\overline{3,n}\right\}$, то исходная матрица A приводима к блочно диагональной канонической форме вида (4.9) с помощью обобщенной матрицы $\widetilde{\widetilde{M}}$, имеющей своими столбцами собственные вектора, а также их вещественные и мнимые компоненты, согласованные с вещественными и комплексно-значными собственными значениями. Так в случае, когда исходная матрица A ЛО ***A*** имеет каноническую строчную фробениусову форму A_F, то $\widetilde{\widetilde{M}}$ является обобщенной матрицей Вандермонда $\widetilde{M}_B$, имеющей представление

$$\widetilde{M}_B=\begin{bmatrix}\mathrm{Re}(\lambda_1^0) & \mathrm{Jm}(\lambda_1^0) & \lambda_3^0 & \ldots & \lambda_n^0\\ \mathrm{Re}(\lambda_1^1) & \mathrm{Jm}(\lambda_1^1) & \lambda_3^1 & \ldots & \lambda_n^1\\ \mathrm{Re}(\lambda_1^2) & \mathrm{Jm}(\lambda_1^2) & \lambda_3^2 & \ldots & \lambda_n^2\\ \ldots & \ldots & \ldots & \ldots & \ldots\\ \mathrm{Re}(\lambda_1^{n-1}) & \mathrm{Jm}(\lambda_1^{n-1}) & \lambda_3^{n-1} & \ldots & \lambda_n^{n-1}\end{bmatrix}. \qquad (4.30)$$

В заключение покажем, что матрица Вандермонда M_B и матрица M собственных векторов произвольной матрицы A простой структуры

ЛО $\boldsymbol{A}$ позволяют приводить матрицу A к канонической строчной фробениусовой форме A_F. Действительно, обе матрицы M и M_B решают задачу диагонализации матриц A и A_F в формах

$$\Lambda = M^{-1}AM \quad \Lambda = {M_B}^{-1}A_F M_B . \tag{4.31}$$

Приведенные матричные соотношения позволяют составить матричное равенство

$${M_B}^{-1}A_F M_B = M^{-1}AM ,$$

которое в разрешенном относительно матрицы A_F виде записывается

$$A_F = M_B M^{-1}AMM_B^{-1} = M_F^{-1}AM_F , \tag{4.32}$$

$M_F = MM_B^{-1}$ – матрица приведения произвольной матрицы A простой структуры к матрице A_F, являющейся канонической строчной фробениусовой формой матриц ЛО $\boldsymbol{A}$.

Примеры и задачи

Приводимые ниже матрицы *простой структуры* привести к канонической *диагональной и строчной сопровождающей (фробениусовой)* формам, построить матрицы приведения подобия к указанным каноническим базисам.

4.1. $\begin{bmatrix} -3 & 0 \\ 7 & -5 \end{bmatrix}$ 4.2. $\begin{bmatrix} 0 & 1 \\ 21 & 4 \end{bmatrix}$ 4.3. $\begin{bmatrix} 0 & 1 \\ 5 & -4 \end{bmatrix}$ 4.4. $\begin{bmatrix} 0 & 12 \\ 1 & 4 \end{bmatrix}$

4.5. $\begin{bmatrix} -3 & 9 \\ 0 & -5 \end{bmatrix}$ 4.6. $\begin{bmatrix} 0 & 21 \\ 1 & 4 \end{bmatrix}$ 4.7. $\begin{bmatrix} 0 & 5 \\ 1 & -4 \end{bmatrix}$ 4.8. $\begin{bmatrix} -2 & 0 \\ 7 & 6 \end{bmatrix}$

4.9. $\begin{bmatrix} -1 & 2 \\ 8 & 5 \end{bmatrix}$ 4.10. $\begin{bmatrix} -1 & 4 \\ 2 & -3 \end{bmatrix}$ 4.11. $\begin{bmatrix} 0 & -15 \\ 1` & -8 \end{bmatrix}$ 4.12. $\begin{bmatrix} 7 & 0 \\ 8 & -3 \end{bmatrix}$

4.13. $\begin{bmatrix} -3 & 8 \\ 1 & -1 \end{bmatrix}$ 4.14. $\begin{bmatrix} 6 & 10 \\ 0 & -2 \end{bmatrix}$ 4.15. $\begin{bmatrix} -1 & 1 \\ -8 & -7 \end{bmatrix}$ 4.16. $\begin{bmatrix} 5 & 16 \\ 1 & -1 \end{bmatrix}$

4.17. $\begin{bmatrix} 5 & -8 \\ 5 & -9 \end{bmatrix}$ 4.18. $\begin{bmatrix} 12 & -7 \\ 12 & -8 \end{bmatrix}$ 4.19. $\begin{bmatrix} 1 & -1 \\ 24 & -9 \end{bmatrix}$ 4.20. $\begin{bmatrix} 5 & 8 \\ 2 & -1 \end{bmatrix}$

Решение вариантов задач

В качестве примера произвольной матрицы A ЛО $\boldsymbol{A}$ возьмем матрицу 4.10, воспользовавшись при этом результатами изучения ее в предыдущем разделе в виде спектров собственных значений и векторов так, что можно записать

$$A=\begin{bmatrix}-1 & 4\\ 2 & -3\end{bmatrix}:\sigma\{A\}=\{\lambda_1=1,\lambda_2=-5\};\xi_1=\begin{bmatrix}2\\1\end{bmatrix},\xi_2=\begin{bmatrix}1\\-1\end{bmatrix}$$

<u>Задача 1</u>: Привести матрицу A к диагональной форме $\Lambda=\begin{bmatrix}1 & 0\\ 0 & -5\end{bmatrix}$ с помощью матричного соотношения $\Lambda=M^{-1}AM$, где

$$M=\begin{bmatrix}M_1=\xi_1 & M_2=\xi_2\end{bmatrix}=\begin{bmatrix}2 & 1\\ 1 & -1\end{bmatrix};M^{-1}=\begin{bmatrix}2 & 1\\ 1 & -1\end{bmatrix}^{-1}=\begin{bmatrix}1/3 & 1/3\\ 1/3 & -2/3\end{bmatrix},$$

тогда

$$\Lambda=M^{-1}AM=\begin{bmatrix}1/3 & 1/3\\ 1/3 & -2/3\end{bmatrix}\begin{bmatrix}-1 & 4\\ 2 & -3\end{bmatrix}\begin{bmatrix}2 & 1\\ 1 & -1\end{bmatrix}=\begin{bmatrix}1 & 0\\ 0 & -5\end{bmatrix}.$$

<u>Задача 2</u>: Привести матрицу A к строчной сопровождающей (фробениусовой) форме, для построения которой составим характеристический полином матрицы A

$$D(\lambda)=\det(\lambda I-A)=(\lambda-\lambda_1)(\lambda-\lambda_2)=(\lambda-2)(\lambda+5)=\lambda^2+4\lambda-5.$$

Сопровождающая форма A_F исходной матрицы A принимает вид $A_F=\begin{bmatrix}0 & 1\\ 5 & -4\end{bmatrix}$, решим задачу ее диагонализации с помощью матрицы Вандермонда.

$$\Lambda=M_B^{-1}A_FM_B=\begin{bmatrix}1 & 1\\ \lambda_1 & \lambda_2\end{bmatrix}^{-1}A_F\begin{bmatrix}1 & 1\\ \lambda_1 & \lambda_2\end{bmatrix}=\begin{bmatrix}1 & 1\\ 1 & -5\end{bmatrix}^{-1}\begin{bmatrix}0 & 1\\ 5 & -4\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & -5\end{bmatrix}=$$

$$=\begin{bmatrix}5/6 & 1/6\\ 1/6 & -1/6\end{bmatrix}\begin{bmatrix}1 & -5\\ 1 & 25\end{bmatrix}=\begin{bmatrix}1 & 0\\ 0 & -5\end{bmatrix}.$$

Теперь приведем исходную матрицу A к сопровождающей (фробениусовой) форме A_F с помощью матричных соотношений

$$A_F=M_BM^{-1}AMM_B^{-1}=\begin{bmatrix}1 & 1\\ 1 & -5\end{bmatrix}\begin{bmatrix}2 & 1\\ 1 & -1\end{bmatrix}^{-1}\begin{bmatrix}-1 & 4\\ 2 & -3\end{bmatrix}\begin{bmatrix}2 & 1\\ 1 & -1\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & -5\end{bmatrix}^{-1}=\begin{bmatrix}0 & 1\\ 5 & -4\end{bmatrix}$$

Поставленная задача решена.

5. ЛИНЕЙНЫЕ И КВАДРАТИЧНЫЕ ФОРМЫ. ДИФФЕРЕНЦИРОВАНИЕ ФУНКЦИЙ ОТ ВЕКТОРОВ И МАТРИЦ ПО СКАЛЯРНЫМ, ВЕКТОРНЫМ И МАТРИЧНЫМ ПЕРЕМЕННЫМ

Определение 5.1 (О5.1). Пусть каждому вектору x линейного действительного пространства R^n ставится в соответствие вполне определенное число из R. Тогда говорят, что в линейном пространстве R^n определена *скалярная* функция от *вектора* $F(x): R^n \to R$.

Определение 5.2 (О5.2). Функция $F_1(x)$, областью определения которой является линейное пространство R^n, а областью значений – совокупность действительных чисел R, называется действительной *линейной формой* (линейным функционалом), если выполняется соотношение

$$F_1(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 F_1(x_1) + \alpha_2 F_2(x_2) \tag{5.1}$$

для любых векторов x_1 и x_2 и любых действительных чисел α_1 и α_2.

Пусть $\{e_1, e_2, ..., e_n\}$ – естественный базис в пространстве R^n, $x = [x_1, x_2, ..., x_n]^T$ – вектор–столбец координат вектора x в этом базисе, тогда любая линейная форма $F_1(x)$ может быть представлена в следующем виде:

$$F_1(x) = \alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_n x_n, \tag{5.2}$$

где $\alpha_k = F_1(e_k), k = \overline{1,n}$. Наоборот, при любых действительных числах $\alpha_1, \alpha_2, ..., \alpha_n$ выражение (5.2) определяет некоторую линейную форму в R^n.

Определение 5.3 (О5.3). Множество всех векторов $x \in R^n$, для которых $F_1(x) = 0$, называется *ядром* линейной формы (функционала) и обозначается $N(F_1)$:

$$N(F_1) = \{x \in R^n : F_1(x) = 0\} \tag{5.3}$$

Линейную форму (5.2) можно записать в E^n как скалярное произведение

$$F_1(x) = (x, \alpha), \tag{5.4}$$

где $\alpha = [\alpha_1, \alpha_2, ..., \alpha_n]^T$.

Определение 5.4 (О5.4). Пусть L – некоторое подпространство пространства R^n. Выберем в R^n произвольный вектор u, тогда множество векторов $z=u+v$, где $v \in L$ называют *плоскостью* в пространстве

R^n. Вектор u называется вектором сдвига, а подпространство L – *направляющим подпространством* этой плоскости.

Определение 5.5 (О.5). *Гиперплоскостью* H в пространстве R^n называется плоскость размерностью *n–1*. Если $L^{\perp}$ – ортогональное дополнение направляющего подпространства L гиперплоскости H и N – любой его базисный вектор, то уравнение гиперплоскости можно записать в следующем виде:

$$(x,N)=(N,x)=b, \tag{5.5}$$

где вектор $N\in L^{\perp}$ есть нормаль к гиперплоскости H, b – действительное число.

Определение 5.6 (О5.6). *Квадратичной формой* от n действительных переменных $x_1, x_2, ..., x_n$ называется функция вида

$$F_2(x)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j, \tag{5.6}$$

где a_{ij} – действительные числа.

Если составить симметричную матрицу A из коэффициентов a_{ij}, называемую матрицей квадратичной формы, и рассматривать величины $x_1, x_2,..., x_n$ как координаты вектора $x\in E^n$ в некотором ортонормированном базисе (например, естественном), то квадратичная форма может быть записана как *скалярное произведение* или квадрат *евклидовой векторной нормы* с весом A:

$$F_2(x)=(Ax,x)=x^TAx=(\|x\|)_A^2. \tag{5.7}$$

Рангом квадратичной формы $F_2(x)$ называется ранг ее матрицы A. При замене переменных $x=Ty$ форма $F_2(x)$ становиться квадратичной формой $F_2(y)$ новых переменных $y_1, y_2,..., y_n$, причем матрица B в этой форме связана с матрицей A соотношением

$$B=T^T\mathrm{A}T, \tag{5.8}$$

при этом если матрица T неособенная, то ранг квадратичной формы не меняется.

Любую квадратичную форму $F_2(x)$ ранга r можно неособенным линейным преобразованием привести к *каноническому* виду.

$$F_2(y)=\alpha_1y_1^2+\alpha_2y_2^2+...+\alpha_ry_r^2, \tag{5.9}$$

где $\alpha_k, k=\overline{1,r}$ – все отличные от нуля числа. Канонический вид называется *нормальным* видом, если все коэффициенты α_k в (5.9) равны 1 или –1. Число положительных коэффициентов в выражении (5.9) называется положительным индексом инерции, число отрицательных коэффициентов – отрицательным индексом инерции, а разность между ними – сигнатурой квадратичной формы.

Симметричная матрица A квадратичной формы имеет ортонормированную систему собственных векторов в евклидовом пространстве E^n, соответствующих собственным значениям $\lambda_1, \lambda_2, ..., \lambda_n$ матрицы A, которые все являются действительными числами. Поэтому матрица A квадратичной формы ортогонально подобна матрице с действительными собственными значениями матрицы A:

$$\Lambda = \mathrm{diag}\{\lambda_1, \lambda_2, ..., \lambda_n\} = T^T A T, (T^{-1} = T^T), \tag{5.10}$$

где $T = [T_1, T_2, ..., T_n]$ – ортогональная матрица, составленная из столбцов координат ортонормированных собственных векторов матрицы A в том же базисе, в котором задана A.

Определение 5.7 (О5.7). Квадратичная форма $F_2(x) = (Ax, x)$ называется *положительно определенной*, если $(Ax, x) > 0$ при $x \neq 0$, и *неотрицательно определенной* если $(Ax, x) \geq 0$ при любых $x \in E^n$. Аналогично, определяются *отрицательно определенная* и *неположительно определенная* квадратичные формы. Если форма $F_2(x)$ принимает разные знаки при некоторых $x \in E^n$, то она называется *неопределенной*. Для того, чтобы квадратичная форма была положительно определенной, необходимо и достаточно, чтобы все главные миноры матрицы этой формы были положительны (критерий Сильвестра).

Пусть $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n$ – собственные значения положительно определенной матрицы A, тогда для всех векторов $x \in E^n$ справедливы неравенства

$$\lambda_n \leq \frac{x^T\ Ax}{x^T x} \leq \lambda_1. \tag{5.11}$$

Определение 5.8 (О5.8). Если все собственные значения матрицы квадратичной формы имеют одинаковый знак, то форма называется *эллиптической*, а уравнение $x^T Ax = c$, где $c - const$, определяет в пространстве E^n гиперэллипсоид постоянного значения (уровня) c.

Рассмотрим основные правила дифференцирования функций от векторов и матриц по *скалярным, векторным и матричным переменным.*

1. Пусть $A = A(q)$ – матрица, элементы которой суть функции $A_{ij} = A_{ij}(q)$ скалярной переменной q. Тогда производной $A_q = \dfrac{\partial A(q)}{\partial q}$ от матрицы $A(q)$ по q является матрица, составленная из производных $A_{ij_q} = \dfrac{\partial A_{ij}(q)}{\partial q}$ ее элементов по переменной q, что может быть записано в форме

$A_q = row\{col(A_{ij}; i=\overline{1,m}); j=\overline{1,n}\}$.

Для производной от суммы и произведения матриц, зависящих от скалярной переменной q по этой переменной справедливы представления

$$C_q = \frac{\partial}{\partial q}\{C(q) = A(q) + B(q)\} = A_q + B_q;$$

$$D_q = \frac{\partial}{\partial q}\{D(q) = A(q)\cdot B(q)\} = A_q \cdot B(q) + A(q)\cdot B_q.$$

Для степенной матричной функции $f(A(q)) = (A(q))^p$ от квадратной $(n \times n)$ – матрицы $A(q)$, где p – целое положительное число, производная по скалярной переменной q вычисляется в силу соотношения

$$\frac{\partial}{\partial q}\left\{(A(q))^p\right\} = A_q(A(q))^{p-1} + (A(q))A_q(A(q))^{p-2} + \ldots + (A(q))^{p-1}A_q.$$

Для вычисления производной от обратной матрицы $(A(q))^{-1}$ сформулируем и докажем следующее утверждение.

Утверждение 5.1 (У5.1). Производная $\frac{\partial}{\partial q}\left(A(q)\right)^{-1}$ от обратной матрицы $\left(A(q)\right)^{-1}$ по скалярной переменной q вычисляется по формуле

$$\frac{\partial}{\partial q}\left(A(q)\right)^{-1} = -\left(A(q)\right)^{-1} A_q \left(A(q)\right)^{-1}. \quad \square \; (5.12)$$

Доказательство утверждения строится на дифференцировании по скалярному параметру q матричного уравнения $\left(A(q)\right)^{-1}\cdot\left(A(q)\right) = I$, где I – единичная матрица, в результате которого получим

$$\frac{\partial}{\partial q}\left\{\left(A(q)\right)^{-1}\cdot\left(A(q)\right) = I\right\} = \frac{\partial}{\partial q}\left(A(q)\right)^{-1}\cdot A(q) + \left(A(q)\right)^{-1}\cdot A_q = 0.$$

Разрешение полученного матричного уравнения относительно производной $\frac{\partial}{\partial q}\left(A(q)\right)^{-1}$ приводит к (5.12). ■

2. Пусть $J = J(x)$ – *скалярная функция* векторного аргумента $x = [x_1,\ldots,x_n]^T$. Тогда, обозначив символом ∇ оператор градиента, для *производной* $\frac{\partial J}{\partial x}$ *от этой функции по векторному аргументу* и градиента можно записать следующие представления:

$$\nabla_x J = \left(\frac{\partial J}{\partial x}\right)^T = \left[\frac{\partial J}{\partial x_1}, \frac{\partial J}{\partial x_2}, \ldots, \frac{\partial J}{\partial x_n}\right]^T \text{ – вектор–столбец;}$$

$$(\nabla_x J)^T = \frac{\partial J}{\partial x} = \left[\frac{\partial J}{\partial x_1}, \frac{\partial J}{\partial x_2}, ..., \frac{\partial J}{\partial x_n}\right] \text{ – вектор–строка;}$$

$$\nabla_{xx} J = \frac{\partial}{\partial x} \cdot \frac{\partial J}{\partial x} = \begin{bmatrix} \frac{\partial^2 J}{\partial x_1^2} & \cdots & \frac{\partial^2 J}{\partial x_n \partial x_1} \\ \cdots & \cdots & \cdots \\ \frac{\partial^2 J}{\partial x_1 \partial x_n} & \cdots & \frac{\partial^2 J}{\partial x_n^2} \end{bmatrix} \text{ – } (n \times n) \text{ – матрица.}$$

3. Пусть $y = [y_1(x), y_2(x), ..., y_m(x)]^T$ – вектор–столбец скалярных функций от *вектора* x (*векторная функция от векторного аргумента*), тогда

$$\nabla_x y = [\nabla_x y_1, \nabla_x y_2, ..., \nabla_x y_m] = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_1} \\ \cdots & \cdots & \cdots \\ \frac{\partial y_1}{\partial x_n} & \cdots & \frac{\partial y_m}{\partial x_n} \end{bmatrix}. \tag{5.13}$$

– матрица размерами $(n \times m)$. Аналогично определяется производная

$$(\nabla_x y)^T = [\nabla_x y_1, \nabla_x y_2, ..., \nabla_x y_m]^T = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial y_m}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_n} \end{bmatrix}. \tag{5.14}$$

4. Пусть $z = z(x)$ и $y = y(x)$ – векторы–столбцы размерности m, и x – вектор–столбец размерности n. Тогда производная по x от скалярного произведения векторов z и y (градиент скалярного произведения) определяется следующим образом:

$$\nabla_x (y, z) = (\nabla_x y)^T z + (\nabla_x z)^T y. \tag{5.15}$$

Примеры и задачи

5.1.* Записать квадратичную форму $F_2(x) = 2x_1^2 + 5x_2^2 - 4x_1x_2$ с симметричной матрицей этой формы.

5.2.* Привести матрицу $A = \begin{bmatrix} 3/2 & 0 & -1/2 \\ 0 & 1 & 0 \\ -1/2 & 0 & 3/2 \end{bmatrix}$ квадратичной формы ортогональным преобразованием к каноническому виду.

$$TT^T = I. \tag{5.16}$$

5.3.* Найти значение x, при котором положительно определенная форма $F_2(x) = (R_x, x) + 2(x, Sb) + C$, где $R > 0, C = const$ принимает минимальное значение. Вычислить это значение.

5.4.* Вычислить производные от следующих скалярных функций от вектора x:

а) $J = \lambda^T Ax$; б) $J = x^T x$; в) $J = x^T Ax$.

5.5.* Пусть X – квадратная матрица размером $(n \times n)$ и $J(X) = trX$ – след этой матрицы, равный сумме ее диагональных элементов. Показать, что

$$\text{а) } \nabla_x trX = \frac{\partial trX}{\partial X} = I; \quad \text{б) } \frac{\partial tr(AX)}{\partial X} = A^{\mathrm{T}}. \tag{5.17}$$

5.6. Определить дефект линейной формы $F_1(x)$ на пространстве R^n.

5.7. Показать, что любой вектор z пространства R^n может быть представлен единственным образом в виде $z = \alpha x + y$, где $y \in N(F_1)$, x – фиксированный вектор R^n, α – действительное число.

5.8. Определить расстояние $\mu(x, H)$ от произвольного вектора $z \in E^n$ до гиперплоскости $(n, x) = b$.

5.9. Для каждой из квадратичных форм найти ортогональное преобразование T неизвестных, приводящее эту форму к каноническому виду, и записать полученный канонический вид:

а) $2x_1^2 + 5x_2^2 + 2x_3^2 - 4x_1x_2 - 2x_1x_3 + 4x_2x_3$;

б) $-3x_2^2 + 4x_1x_2 + 10x_1x_3 - 4x_2x_3$;

в) $-x_1^2 + x_2^2 - 5x_3^2 + 6x_1x_3 + 4x_2x_3$;

г) $2x_1^2 + x_2^2 - 4x_1x_2 - 4x_2x_3$.

5.10. Выяснить, являются ли положительно определенными следующие квадратичные формы:

а) $x_1^2 + 2x_1x_2 + 4x_2^2 + 4x_2x_3 + 2x_3^2$;

б) $5x_1^2 + x_2^2 + 5x_3^2 + 4x_1x_2 - 8x_1x_3 - 4x_2x_3$;

в) $3x_1^2 + x_2^2 + 5x_3^2 + 4x_1x_2 - 8x_1x_3 - 4x_2x_3$.

5.11. Доказать, что положительно определенная матрица является неособенной.

5.12. Доказать, что если A – положительно определенная симметричная матрица, то A^{-1} – также положительно определенная матрица.

5.13. Доказать, что если $|A| \neq 0$, то $A^T A$ и AA^T – положительно определенные матрицы; если $|A| = 0$, $A^T A$ и AA^T – неотрицательно определенные матрицы.

5.14. Привести матрицу $A=\begin{bmatrix}3 & 1 & 1\\ 1 & 0 & 2\\ 1 & 2 & 0\end{bmatrix}$ квадратичной формы к каноническому виду и записать полученный канонический вид.

5.15. Доказать справедливость для любой симметричной матрицы A спектрального разложения:

$A=\lambda_1\xi_1\xi_1^T+\lambda_2\xi_2\xi_2^T+...+\lambda_n\xi_n\xi_n^T$, где $\lambda_1,\lambda_2,...,\lambda_n$ – собственные значения матрицы $A:\xi_1,\xi_2,...,\xi_n$ – ортонормированная система собственных векторов этой матрицы.

5.16. Доказать, что $x^T Ax=tr(Axx^T)$, где tr обозначает след матрицы.

5.17. Вычислить производные от следующих функций:

а) $J(x)=(y-Ax)^T Q(y-AX)$, где $Q^T=Q>0$;

б) $J(X)=tr(AX^T)$;

в) $J(X)=tr(X^T AX)$;

г) $J(X)=tr(X^{-1})$;

д) $J(X)=|X|$.

5.18. Определить минимальное значение квадратичной формы $F_2(x)=(Rx,x)+2(x,Su)+(u,Tu)$, где $R>0$; S,T – матрицы, u – не зависящий от x вектор.

Решение вариантов задач

Решение задачи 5.1. Матрица A исходной формы: $A=\begin{bmatrix}2 & -4\\ 0 & 5\end{bmatrix}$.

Любая действительная матрица может быть представлена в виде: $A=A^c+A^k$, где $A^{\text{c}}=\frac{1}{2}(A+A^T)$ – симметричная матрица, $A^k=\frac{1}{2}(A-A^T)$ – кососимметричная матрица. Поскольку для любого вектора x $A^{\text{k}}x\perp x$, то $x^T A^k x=0$, поэтому имеем $F_2(x)=x^T A^{\text{c}}x$, где A^c в данном случае равна $A^{\text{c}}=\begin{bmatrix}2 & -2\\ -2 & 5\end{bmatrix}$.

Решение задачи 5.2. Вычисление корней характеристического уравнения $\det[A-\lambda I]=0$ дает следующие собственные значения матрицы A квадратичной формы: $\lambda_1=\lambda_2=1,\quad \lambda_3=2$. Решив систему уравнений $(A-I)t=0$, получим два ортонормированных собственных вектора, соответствующих собственному значению $\lambda = 1$: $\xi_1 = \left[\sqrt{2/2} \ \vdots 0 \ \vdots \sqrt{2/2}\right]^T$, $\xi_2 = [0 \ \vdots 1 \ \vdots 0]^T$. Решив систему уравнений $(A - 2I)\xi = 0$, получим третий вектор ортонормированной системы, соответствующий собственному значению $\lambda = 2$: $\xi_3 = \left[\sqrt{2/2} \ \vdots 0 \ \vdots -\sqrt{2/2}\right]^T$. В итоге матрица ортогонального преобразования равна $T = [\xi_1, \xi_2, \xi_3] = \begin{bmatrix} \sqrt{2/2} \vdots & 0 \vdots & \sqrt{2/2} \\ 0 \ \ \vdots & 1 \vdots & 0 \\ \sqrt{2/2} \vdots & 0 \vdots & -\sqrt{2/2} \end{bmatrix}$.

Матрица канонического вида формы равна $\Lambda = diag\{1, 1, 2\}$ так что $\Lambda = T^T A T$, при этом $T^T T = I$.

Решение задачи 5.3. Рассмотрим решение данной задачи методом, не связанным с вычислением производных. Дополним форму $F_2(x)$ до полного квадрата:

$$\begin{aligned} &x^T Rx + 2x^T Sb + c = x^T Rx + 2x^T RR^{-1}Sb + b^T S^T R^{-1}Sb + c - \\ &- b^T S^T R^{-1}Sb = (x + R^{-1}Sb)^T R(x + R^{-1}Sb) + c - b^T S^T R^{-1}Sb \end{aligned} \tag{5.18}$$

Из выражения (5.18) и положительной определенности матрицы R следует, что данная квадратичная форма принимает минимальное значение при $x + R^{-1}Sb = 0$, откуда получим $x_{\min} = -R^{-1}Sb$ и $F_{2\min} = c - b^T S^T R^{-1}Sb$.

Решение задачи 5.4.

а) Положим $A^T \lambda = C$, тогда исходную функцию можно записать в виде $J = C^T x$. Согласно выражению (5.13) имеем

$$\nabla_x J = \nabla_x (C^T x) = \nabla_x (C_1 x_1 + ... + C_n x_n) = C,$$

откуда получаем

$$\nabla_x J = A^T \lambda \tag{5.19}$$

б) Поскольку $x^T x = (x, x)$ то

$\nabla_x (x, x) = (\nabla_x x^T)x + (\nabla_x x^T)x = 2(\nabla_x x^T)x$. Но $\nabla_x x^T = I$, поэтому окончательно получим

$$\nabla_x J = 2x \tag{5.20}$$

в) Положим $y = Ax$, тогда можем записать $J = x^T Ax = (Ax, x) = (y, x)$. Согласно выражению (5.15)

$$\nabla_x J = \nabla_x (y^T x) = (\nabla_x y^T)x + (\nabla_x x^T)y.$$

Но $\nabla_x y^T = \nabla_x (Ax)^T = [\nabla_x y_1, ..., \nabla_x y_n] = A^{\mathrm{T}}$ по формуле (5.13), а $\nabla_x x^T = I$, поэтому в итоге получим $\nabla_x J = A^T x + Ax$. Если матрица A симметричная, то будем иметь

$$\nabla_x J = 2Ax. \tag{5.21}$$

Решение задачи 5.5.

а) Поскольку, $trX = x_{11} + ... + x_{nn}$, то из формулы (5.13) имеем

$$\nabla_x trX = \nabla_x x_{11} + ... + \nabla_x x_{nn} = I.$$

б) Поскольку $tr(AX) = a_{11}x_{11} + a_{12}x_{21} + ... + a_{nn}x_{nn}$, то

$$\nabla_x tr(AX) = a_{11}\nabla_x x_{11} + ... + a_{nn}\nabla_x x_{nn}. \tag{5.22}$$

Но $\nabla_x x_{ij}$ есть матрица размером $(n \times n)$, имеющая единственный отличный от нуля и равный единице элемент, стоящий в i-й строке и j-м столбце. Сложив все слагаемые правой части выражения (5.22), получим требуемый результат.

6. ФУНКЦИИ ОТ МАТРИЦ. МАТРИЧНАЯ ЭКСПОНЕНТА И ЕЕ СВОЙСТВА

Рассматривается $(n \times n)$ – квадратная матрица A, на которой конструируются функции от матрицы $f(A)$ трех типов: *скалярная* функция от матрицы, *векторная* функция от матрицы и *матричная* функция от матрицы.

Определение 6.1 (О6.1). *Скалярной функцией* (СФМ) от квадратной матрицы A называется функция $f(A)$, которая реализует отображение $f(A): R^{n\times n} \Rightarrow R$, где R – множество действительных чисел.

Примерами скалярных функций от матрицы являются: $f(A) = \det(A), f(A) = tr(A), f(A) = C\{A\}, f(A) = \|A\|$ – детерминант, след, число обусловленности и норма матрицы соответственно, СФМ является квадратичная форма $f(A) = x^T Ax$.

Определение 6.2 (О6.2). *Векторной функцией* от квадратной матрицы A называется функция $f(A)$, которая реализует отображение $f(A): R^{n\times n} \Rightarrow R^n$, где R^n – n-мерное действительное пространство.

Примерами векторных функций от матрицы (ВФМ) являются такие, как $f(A) = col\{\lambda_i; i = \overline{1,n}\}$, $f(A) = col\{\alpha_i; i = \overline{1,n}\}$ – векторы, построенные на элементах алгебраических спектров соответственно собственных значений $\{\lambda_i; i = \overline{1,n}\}$ и сингулярных чисел $\{\alpha_i; i = \overline{1,n}\}$ матрицы A.

Матричная функция от матрицы (МФМ) реализует отображение $f(A): R^{n\times n} \Rightarrow R^{n\times n}$. *Исходное определение* матричной функции от матрицы задается следующим образом.

Определение 6.3 (О6.3). Пусть $f(\alpha)$ – скалярный степенной ряд (многочлен) относительно скалярной переменной α.

$$f(\alpha) = a_0 + a_1\alpha + a_2\alpha^2 + \ldots + a_p\alpha^p. \tag{6.1}$$

Тогда скалярный ряд $f(\alpha)$ порождает матричную функцию $f(A)$ от матрицы A в виде матричного ряда, если в представлении (6.1) для $f(\alpha)$ скалярную переменную α заменить на матрицу A так, что $f(A)$ запишется в форме

$$f(A) = a_0 I + a_1 A + a_2 A^2 + \ldots + a_p A^p. \tag{6.2}$$

Поставим задачу построения перехода от исходного представления МФМ $f(A)$ в форме (6.2) к ее *минимальному представлению*, то есть к представлению матричным многочленом минимальной степени. Начнем решение этой задачи с *теоремы Гамильтона–Кэли*.

Утверждение 6.1 (У6.1) (Теорема Гамильтона–Кэли).

Квадратная $(n \times n)$- матрица A с характеристическим полиномом

$D(\lambda) = \det(\lambda I - A) = \lambda^n + a_1\lambda^{n-1} + a_2\lambda^{n-2} + \ldots + a_{n-1}\lambda + a_n$, обнуляет свой характеристический полином так, что выполняется матричное соотношение

$$D(A) = A^n + a_1 A^{n-1} + a_2 A^{n-2} + \ldots + a_{n-1}A + a_n I = 0, \tag{6.3}$$

где *0* – $(n \times n)$ нулевая матрица. □

Доказательство справедливости сформулированного утверждения осуществим для случая матрицы A простой структуры, характеризующейся алгебраическим спектром $\sigma\{A\} = \{\lambda_i : \lambda_i \neq \lambda_j; i \neq j; \mathrm{Jm}\lambda_i = 0; i = \overline{1,n}\}$ вещественных и некратных собственных значений так, что на нем может быть сконструирована диагональная матрица $\Lambda = diag\{\lambda_i \,; i = \overline{1,n}\}$. Если теперь воспользоваться матричным соотношением подобия (2.30), то матрицу A можно представить в форме $A = M\Lambda M^{-1}$, что в свою очередь для (6.3) позволяет записать

$$D(A) = M\left\{\Lambda^n + a_1\Lambda^{n-1} + a_2\Lambda^{n-2} + \ldots + a_{n-1}\Lambda + a_n I\right\}M^{-1} =$$
$$= M\,diag\left\{\lambda_i^n + a_1\lambda_i^{n-1} + a_2\lambda_i^{n-2} + \ldots + a_{n-1}\lambda_i + a_n I; i = \overline{1,n}\right\}M^{-1} = 0. \quad ■$$

Теорема Гамильтона-Кэли позволяет ввести следующие определения.

Определение 6.4 (О6.4). Многочлен (степенной ряд) $\varphi(\alpha)$ относительно скалярной переменной α называется *аннулирующим многочленом* квадратной матрицы A, если выполняется условие

$$\varphi(A) = 0. \tag{6.4}$$

Очевидно, *аннулирующим многочленом* матрицы A в силу теоремы Гамильтона-Кэли является в первую очередь ее *характеристический полином.* Ясно, что существует множество аннулирующих многочленов матрицы A степени большей, чем n. Но могут существовать аннулирующие многочлены степени $m < n$.

Определение 6.5 (О6.5). Аннулирующий многочлен $\psi(\alpha)$ наименьшей степени m со старшим коэффициентом при α^m, равным единице, называется *минимальным многочленом* матрицы A.

Построим разложение многочлена $f(\alpha)$ (6.1), задающего матричную функцию $f(A)$ от матрицы в форме (6.2), по модулю минимального многочлена $\psi(\alpha)$ матрицы A, представив его выражением

$$f(\alpha) = \varphi(\alpha)\psi(\alpha) + r(\alpha), \tag{6.5}$$

где многочлен $r(\alpha)$ имеет степень $\deg(r(\alpha))$ меньше степени $\deg(\psi(\alpha)) = m$ минимального многочлена $\psi(\alpha)$ матрицы A. Выражение (6.5) позволяет дать следующее определение матричной функции от матрицы.

Определение 6.6 (О6.6). Пусть многочлен $f(\alpha)$ относительно скалярной переменной α допускает представление в форме (6.5), тогда матричная функция $f(A)$ может быть задана в *минимальной* форме

$$f(A) = r(A). \tag{6.6}$$

Заметим, что основной проблемой при задании матричной функции от матрицы в форме (6.6) является вычисление многочлена $r(\alpha)$.

Основной способ вычисления многочлена $r(\alpha)$ в силу (6.5) опирается на то, что $r(\alpha)$ является остатком от деления $f(\alpha)$ на минимальный многочлен

$$r(\alpha) = rest\frac{f(\alpha)}{\psi(\alpha)}. \tag{6.7}$$

Если $f(\alpha)$ *не является рядом или многочленом вида* (6.1), а является произвольной аналитической функцией со значениями на алгебраическом спектре собственных значений матрицы A, то формирование матричной функции $f(A)$ от матрицы A, опирается на представление $f(\alpha)$ в соответствии с интерполяционной схемой Лагранжа в виде мультипликативной структуры из двучленов $(\alpha - \lambda_i)$ или в соответствии с интерполяционной схемой Ньютона в виде ряда по степеням двучленов $(\alpha - \lambda_i)$, число членов которых определяется минимальным многочленом $\psi(\lambda)$. Для реализации интерполяционной схемы Лагранжа, которая в случае размещения интерполяционных узлов на собственных значениях λ_i матрицы A приобретает название интерполяционной схемы Лагранжа–Сильвестра, требуется знание значений $f(\lambda_i)$. Для реализации интерполяционной схемы Ньютона требуется знание значений $f(\lambda_i), f'(\lambda_i) \ldots f^{(m_i-1)}(\lambda_i)$.

Если минимальный многочлен $\psi(\alpha)$ степени m в силу его определения записать в форме

$$\Psi(\alpha) = (\alpha - \lambda_1)^{m_1}(\alpha - \lambda_2)^{m_2} \ldots (\alpha - \lambda_r)^{m_r}, \tag{6.8}$$

где $m_1 + m_2 + \ldots + m_r = m$, $\left\{\lambda_i; i = \overline{1, r}\right\} \in \sigma\{A\}$,

то можно построить представление для функции $f(\alpha)$ в форме

$$f(\alpha) = r(\alpha), \tag{6.9}$$

где $r(\alpha)$ – *интерполяционный многочлен* Лагранжа–Сильвестра или Ньютона, сформированный на алгебраическом спектре $\sigma\{A\}$ собственных значений $\left\{\lambda_i; i = \overline{1, r}\right\}$ матрицы A, характеризующийся степенью меньшей степени m минимального многочлена $\psi(\alpha)$, а потому удовлетворяющий условиям (6.5), (6.7).

Рассмотрим случай, когда нули минимального многочлена (6.8) являются простыми, т.е. при $m_1 = m_2 = ... = m_r = 1$, минимальный многочлен и характеристический совпадают так, что выполняются равенства $\psi(\alpha) = D(\alpha)$ и $r = n$, тогда представление $f(\alpha) = r(\alpha)$ в форме *интерполяционного многочлена* Лагранжа–Сильвестра принимает вид

$$r(\alpha) = \sum_{i=1}^{n} \frac{(\alpha - \lambda_1)...(\alpha - \lambda_{i-1})(\alpha - \lambda_{i+1})...(\alpha - \lambda_n)}{(\lambda_i - \lambda_1)...(\lambda_i - \lambda_{i-1})(\lambda_i - \lambda_{i+1})...(\lambda_i - \lambda_n)} f(\lambda_i). \tag{6.10}$$

Матричная функция от матрицы для случая некратных собственных значений матрицы A принимает с использованием (6.10) вид

$$f(A) = r(A) = \sum_{i=1}^{n} \frac{(A - \lambda_1 I)...(A - \lambda_{i-1} I)(A - \lambda_{i+1} I)...(A - \lambda_n I)}{(\lambda_i - \lambda_1)...(\lambda_i - \lambda_{i-1})(\lambda_i - \lambda_{i+1})...(\lambda_i - \lambda_n)} f(\lambda_i). \tag{6.11}$$

Теперь допустим, что характеристический многочлен $D(\alpha)$ имеет кратные корни, но *минимальный* многочлен $\psi(\alpha)$, являясь делителем $D(\alpha)$, имеет только простые корни

$$\psi(\alpha) = (\alpha - \lambda_1)(\alpha - \lambda_2)...(\alpha - \lambda_m).$$

В этом случае интерполяционный многочлен $r(\alpha)$ совпадает с точностью до замены числа членов n на m с представлением (6.10). Как следствие матричная функция $f(A)$ от матрицы A принимает вид

$$f(A) = r(A) = \sum_{i=1}^{m} \frac{(A - \lambda_1 I)...(A - \lambda_{i-1} I)(A - \lambda_{i+1} I)...(A - \lambda_m I)}{(\lambda_i - \lambda_1)...(\lambda_i - \lambda_{i-1})(\lambda_i - \lambda_{i+1})...(\lambda_i - \lambda_m)} f(\lambda_i). \tag{6.12}$$

В заключение рассмотрим общий случай, когда минимальный многочлен матрицы A имеет вид (6.8). Для случая кратных нулей минимального многочлена, то есть когда он имеет вид (6.8), представление $r(\alpha)$ в форме *интерполяционного многочлена* Лагранжа-Сильвестра, содержащего элементы интерполяционной схемы Ньютона, принимает вид

$$r(\alpha) = \sum_{i=1}^{r} \sum_{j=1}^{m_i} \frac{1}{(j-1)!} \left[\frac{\partial^{(j-1)}}{\partial \alpha^{(j-1)}} \frac{f(\alpha)}{\psi_i(\alpha)} \right] \Big|_{\alpha = \lambda_i} \frac{\psi(\alpha)}{(\alpha - \lambda_i)^{m_i}}, \tag{6.13}$$

где для компактности записи использовано обозначение

$$\psi_i(\lambda_i) = \frac{\psi(\alpha)}{(\alpha - \lambda_i)^{m_i}} \Big|_{\alpha = \lambda_i}.$$

Если ввести обозначение

$$\alpha_{i,j} = \frac{1}{(j-1)!} \left[\frac{\partial^{(j-1)}}{\partial \alpha^{j-1}} \frac{f(\alpha)}{\psi_k(\alpha)} \right]_{\alpha = \lambda_i}, \tag{6.14}$$

то выражение (6.13) для $f(\alpha)=r(\alpha)$ принимает вид

$$r(\alpha)=\sum_{i=1}^{r}\left\{\alpha_{i,1}+\alpha_{i,2}(\alpha-\lambda_i)+...+\alpha_{i,m_i}(\alpha-\lambda_i)^{m_i-1}\right\}\psi_i(\alpha). \qquad (6.15)$$

Если воспользоваться представлением (6.15), то для МФМ можно записать

$$f(A)=r(A)=\sum_{i=1}^{r}\left\{\alpha_{i,1}I+\alpha_{i,2}(A-\lambda_i I)+...+\alpha_{i,m_i}(A-\lambda_i I)^{m_i-1}\right\}\psi_i(A). \qquad (6.16)$$

Если матрица A представляет собой $(n\times n)-$жорданову клетку, порождаемую собственным значением λ_0 кратности n, так что матрица A принимает вид

$$A=J=\begin{bmatrix}\lambda_0 & 1 & 0 & ... & 0\\ 0 & \lambda_0 & 1 & ... & 0\\ . & . & . & ... & .\\ . & . & . & ... & .\\ 0 & 0 & 0 & ... & 1\\ 0 & 0 & 0 & ... & \lambda_0\end{bmatrix}, \qquad (6.17)$$

то интерполяционный многочлен $r(\alpha)$, так как минимальный многочлен матрицы A (6.17) имеет вид $\psi(\alpha)=(\alpha-\lambda_0)^n$, для функции $f(\alpha)$ полностью строится по интерполяционной схеме Ньютона и определяется выражением

$$r(\alpha)=f(\lambda_0)+\frac{f'(\lambda_0)}{1!}(\alpha-\lambda_0)+...+\frac{f^{(n-1)}(\lambda_0)}{(n-1)!}(\alpha-\lambda_0)^{n-1}. \qquad (6.18)$$

В силу (6.18), (6.5), (6.7) матричная функция $f(A)$ от матрицы $A=J$ принимает вид

$$r(J)=f(\lambda_0)I+\frac{f'(\lambda_0)}{1!}(J-\lambda_0 I)+...+\frac{f^{(n-1)}(\lambda_0)}{(n-1)!}(J-\lambda_0 I)^{n-1}=$$

$$=\begin{bmatrix} f(\lambda_0) & \frac{f'(\lambda_0)}{1!} & \frac{f''(\lambda_0)}{2!} & ... & \frac{f^{(n-1)}(\lambda_0)}{(n-1)!}\\ 0 & f(\lambda_0) & \frac{f'(\lambda_0)}{1!} & ... & \frac{f^{(n-2)}(\lambda_0)}{(n-2)!}\\ . & . & . & ... & .\\ 0 & 0 & 0 & ... & \frac{f'(\lambda_0)}{1!}\\ 0 & 0 & 0 & ... & f(\lambda_0)\end{bmatrix} \qquad (6.19)$$

Рассмотрим теперь случай, когда матрица A имеет вид $A=diag\left\{J_i;\sum_{i=1}^{r}m_i=n\right\}$, где $J_i-(m_i\times m_i)$ – жорданова клетка, порождаемая собственным значением λ_i кратности m_i, так что матрица A принимает вид

$$A = diag\left\{ J_i = \begin{bmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 \\ \vdots & \cdots & \cdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & \lambda_i \end{bmatrix}; \sum_{i=1}^{r} m_i = n \right\}, \qquad (6.20)$$

тогда в силу (6.18) и (6.19) матричная функция $f(A)$ от матрицы A (6.20) принимает вид

$$f(A) = diag\left\{ f(J_i) = \begin{bmatrix} f(\lambda_i) & \frac{f'(\lambda_i)}{1!} & \frac{f''(\lambda_i)}{2!} & \ldots & \frac{f^{(n-1)}(\lambda_i)}{(n-1)!} \\ 0 & f(\lambda_i) & \frac{f'(\lambda_i)}{1!} & \ldots & \frac{f^{(n-2)}(\lambda_i)}{(n-2)!} \\ . & . & . & \ldots & . \\ 0 & 0 & 0 & \ldots & \frac{f'(\lambda_i)}{1!} \\ 0 & 0 & 0 & \ldots & f(\lambda_i) \end{bmatrix}; \sum_{i=1}^{r} m_i = n \right\}. \qquad (6.21)$$

Из определения матричной функции от матрицы во всех формах следуют ее основные свойства:

Свойство 6.1 (СВ6.1). Матричная функция от матрицы $f(A)$ сохраняет геометрический спектр $\left\{\xi_i; i = \overline{1,n}\right\}$ собственных векторов ξ_i матрицы A: $A\xi_i = \lambda_i \xi_i$, так что выполняется соотношение

$$f(A)\xi_i = f(\lambda_i)\xi_i, \qquad (6.22)$$

где $f(\lambda_i) = \lambda_{fi}$ – собственные значения матрицы $f(A)$, удовлетворяющие ее характеристическому уравнению $\det\{\lambda_f I - f(A)\} = 0$ и вычисляемые как функция $f(\lambda)$ на спектре $\sigma\{A\} = \left\{\lambda_i; i = \overline{1,n}\right\}$ собственных значений матрицы $f(A)$.

Свойство 6.2 (СВ6.2). Матричная функция от матрицы $f(A)$ сохраняет матричное отношение подобие в том смысле, что если матрицы A и B подобны, т.е. $B = T^{-1}AT$, то

$$f(B) = T^{-1} f(\mathrm{A}) T. \qquad (6.23)$$

Свойство 6.3 (СВ6.3). Матричная функция от матрицы $f(A)$ сохраняет блочно-диагональную форму матрицы A в том смысле, что, если $A = diag\{A_1\ A_2\ \ldots\ A_\mu\}$, то

$$f(A) = diag\{f(A_1)\ f(A_2) \ldots f(A_\mu)\}. \qquad (6.24)$$

Теперь распространим полученные результаты на задачи формирования способов аналитического представления и вычисления *матричной экспоненты* e^{At}, параметризованной непрерывным временем t, исходное задание которой в форме (6.1) порождено скалярной экспонентой $e^{\alpha t}$ или $\exp(\alpha t)$, записанной в форме бесконечного скалярного ряда

$$e^{\alpha t}=1+\alpha t+\frac{(\alpha t)^2}{2!}+\frac{(\alpha t)^3}{3!}+\ldots+\frac{(\alpha t)^p}{p!}+\ldots=\sum_{k=0}^{\infty}\frac{(\alpha t)^k}{k!}\quad ,$$

и принимает вид

$$e^{At}=I+At+\frac{(At)^2}{2!}+\frac{(At)^3}{3!}+\ldots+\frac{(At)^p}{p!}+\ldots=\sum_{k=0}^{\infty}\frac{(At)^k}{k!}. \qquad (6.25)$$

Следует заметить, что аналогичным образом может быть задана любая матричная функция от матрицы, для скалярного прототипа которой известен ряд ее представляющий.

В связи со сказанным и проведенными выше исследованиями, а также упомянутыми свойствами матричных функций от матриц, перечислим основные способы вычисления и построения аналитических представлений матричной экспоненты.

1. *Численный способ*, основанный на переходе от непрерывного времени t к дискретному k, выраженному в числе интервалов дискретности длительности Δt так, что $t=(\Delta t)k$, в результате чего матричная экспонента e^{At} получает представление

$$e^{At}=e^{A\Delta tk}=(e^{A\Delta t})^k=\left(\overline{A}\right)^k, \qquad (6.26)$$

где матрица $\overline{A}=e^{A\Delta t}$ при правильном выборе интервала дискретности Δt задается конечным числом $(p+1)$ членов степенного матричного представления

$$\overline{A}=e^{A\Delta t}=I+A\Delta t+\frac{1}{2!}(A\Delta t)^2+\frac{1}{3!}(A\Delta t)^3+\ldots+\frac{1}{p!}(A\Delta t)^p. \qquad (6.27)$$

При чем, если $p\geq n$, то с помощью (6.7) ряд (6.27) может быть приведен к *минимальной форме,* т.е. матричному ряду степени $m-1$, а в случае $\psi(\alpha)=D(\alpha)$ – к матричному ряду степени $n-1$. Для вычисления интервала дискретности Δt можно воспользоваться соотношением

$$\Delta t=0.05n\left(\left|tr(A)\right|\right)^{-1}. \qquad (6.28)$$

2. *Способ диагонализации* матрицы A, именуемый иначе способом собственных значений. Способ применим к матрицам A простой структуры так, что ее спектр собственных значений имеет вид $\sigma\{A\}=\left\{\lambda_i:\lambda_i\neq\lambda_j;i\neq j;Jm\lambda_i=0;i=\overline{1,n}\right\}$, а потому оказывается спра-

ведливым матричное соотношение приведения подобия $M\Lambda = AM$, где $\Lambda = diag\left\{\lambda_i; i = \overline{1,n}\right\}$.

Тогда матричная экспонента принимает вид

$$e^{At} = M\, e^{\Lambda t} M^{-1} = M\, diag\left\{e^{\lambda_i t}; i = \overline{1,n}\right\} M^{-1}, \qquad (6.29)$$

где

$$M = row\left\{M_i = \xi_i = arg\left(A\xi_i = \lambda_i \xi_i\right); i = \overline{1,n}\right\}, \qquad (6.30)$$

то есть M – матрица собственных векторов матрицы A.

3. *Способ, основанный на приведении к нормальной форме Жордана* матрицы A. Способ применим к матрицам A, спектр собственных значений которых $\sigma\{A\} = \left\{\lambda_i; i = \overline{1,r}; \sum_{i=1}^{r} m_i = n\right\}$ содержит r кратных собственных значений λ_i кратности m_i каждый. Для этого случая оказывается справедливым матричное соотношение приведения подобия $TJ = AT$, где

$$T = row\{T_i = \left\lfloor T_{i1} = \xi_i; T_{i2} = \left(A - \lambda_i I\right)^{+} T_{i1} .. T_{im_i} = \left(A - \lambda_i I\right)^{+} T_{im_i - 1} \right\rfloor; i = \overline{1,r}\},$$

здесь ξ_i – собственный вектор матрицы A, соответствующий собственному значению λ_i: $\xi_i = arg\left\{A\xi_i = \lambda_i \xi_i; i = \overline{1,r}\right\}$; $(*)^{+}$ – операция псевдообращения матрицы (*).

$$J = diag\left\{ J_i = \begin{bmatrix} \lambda_i & 1 & 0 & \dots & 0 \\ 0 & \lambda_i & 1 & \dots & 0 \\ . & . & . & \dots & . \\ . & . & . & \dots & . \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & \lambda_i \end{bmatrix}; \sum_{i=1}^{r} m_i = n \right\}.$$

В результате для матричной экспоненты e^{At} можно записать $e^{At} = Te^{Jt}T^{-1}$, где матричная экспонента e^{Jt} имеет вид

$$e^{Jt} = diag\left\{ e^{J_i t} = \begin{bmatrix} e^{\lambda_i t} & \frac{te^{\lambda_i t}}{1!} & \frac{t^2 e^{\lambda_i t}}{2!} & \dots & \frac{t^{n-1} e^{\lambda_i t}}{(n-1)!} \\ 0 & e^{\lambda_i t} & \frac{te^{\lambda_i t}}{1!} & \dots & \frac{t^{n-2} e^{\lambda_i t}}{(n-2)!} \\ . & . & . & \dots & . \\ 0 & 0 & 0 & \dots & \frac{te^{\lambda_i t}}{1!} \\ 0 & 0 & 0 & \dots & e^{\lambda_i t} \end{bmatrix}; \sum_{i=1}^{r} m_i = n \right\}. \qquad (6.31)$$

4. *Способ преобразования Лапласа* заключается в вычислении обратного преобразования Лапласа от резолвенты $(sI - A)^{-1}$ в форме

$$e^{At} = L^{-1}\{(sI - A)^{-1}\} \qquad (6.32)$$

Способ поддерживается алгоритмом Фаддеева–Леверье разложения резолвенты без ее обращения на основе представления

$$(sI-A)^{-1} = \frac{1}{\det(sI-A)}\left[\Delta(sI-A)\right]^T = \frac{s^{n-1}H_0 + s^{n-2}H_1 + \ldots + H_{n-1}}{s^n + a_1 s^{n-1} + \ldots + a_{n-1}s + a_n},$$

где $(n \times n)$ – матрицы $H_i \ (i = \overline{0, n-1})$ и коэффициенты характеристического уравнения вычисляются с помощью рекуррентной процедуры алгоритма Фаддеева–Леверье:

$$\begin{array}{ll} H_0 = I, & a_1 = -tr(AH_0) \\ H_1 = AH_0 + a_1 I, & a_2 = -tr(AH_1)/2 \\ \ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots \\ H_k = AH_{k-1} + a_k I, & a_{k+1} = -tr(AH_k)/k \end{array} . \qquad (6.33)$$

С использованием матриц $H_k \ (k = \overline{0, n-1})$ для резолвенты $(sI - A)^{-1}$ можно записать в форме

$$(sI - A)^{-1} = \frac{s^{n-1}}{D(s)}H_0 + \frac{s^{n-2}}{D(s)}H_1 + \ldots + \frac{s}{D(s)}H_{n-2} + \frac{1}{D(s)}H_{n-1}. \qquad (6.34)$$

Матричная экспонента (6.32) с использованием (6.34) получает представление

$$e^{At} = L^{-1}\left\{\frac{s^{n-1}}{D(s)}\right\}H_0 + L^{-1}\left\{\frac{s^{n-2}}{D(s)}\right\}H_1 + \ldots + L^{-1}\left\{\frac{1}{D(s)}\right\}H_{n-1}. \qquad (6.35)$$

Запишем характеристический многочлен $D(s)$ в форме

$$D(s) = (s - \lambda_1)^{m_1}(s - \lambda_1)^{m_2} \ldots (s - \lambda_1)^{m_r},$$

тогда становится справедливым представление

$$\frac{s^k}{D(s)} = \sum_{i=1}^{r}\left\{\frac{\beta_{k1}}{s - \lambda_i} + \frac{\beta_{k2}}{(s-\lambda_i)^2} + \ldots + \frac{\beta_{km_i}}{(s-\lambda_i)^{m_i}}\right\}; k = \overline{0, n-1}. \qquad (6.36)$$

Тогда

$$L^{-1}\left\{\frac{s^k}{D(s)}\right\} = \sum_{i=1}^{r}\left\{\beta_{k1} + \beta_{k2}t + \frac{\beta_{k3}}{2!}t^2 + \ldots + \frac{\beta_{km_i}}{(m_i - 1)!}t^{m_i - 1}\right\}e^{\lambda_i t}; k = \overline{0, n-1}. \qquad (6.37)$$

Подставляя (6.37) в (6.35) окончательно получим

$$e^{At} = \sum_{k=n-1}^{0} H_{n-1-k}\left\{\sum_{i=1}^{r}\left\{\beta_{k1} + \beta_{k2}t + \ldots + \frac{\beta_{km_i}}{(m_i - 1)!}t^{m_i - 1}\right\}e^{\lambda_i t}; k = \overline{0, n-1}\right\}.$$

5. *Способ Лагранжа–Сильвестра.* Интерполяционный многочлен Лагранжа–Сильвестра в зависимости от свойств минимального многочлена $\psi(\alpha)$ определяется выражениями (6.11),(6.12),(6.16), которые

после замены функции $f(\lambda_i)$ на $e^{\lambda_i t}$, A на At, λ_i на $\lambda_i t$ дают представлдение матричной экспоненты e^{At}.

Примеры и задачи

6.1. Найти матричную экспоненту e^{At} способом, основанным на приведении к нормальной форме Жордана для матрицы

$$A=\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -8 & -12 & -6 \end{bmatrix}.$$

6.2. Найти e^{At} методами собственных значений (диагонализации матрицы A) и с помощью преобразования Лапласа для матриц:

а) $A=\begin{bmatrix} 8 & -8 & -2 \\ 4 & -3 & -2 \\ 3 & -4 & 1 \end{bmatrix}$; б) $A=\begin{bmatrix} -3 & 1 & 0 \\ 1 & -3 & 0 \\ 0 & 0 & -3 \end{bmatrix}$.

6.3. Найти e^{At} методом собственных значений (диагонализации матрицы A) и с помощью преобразования Лапласа для матриц:

а) $A=\begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$; б) $A=\begin{bmatrix} -1 & -2 & 3 \\ 0 & 1 & -1 \\ -1 & -1 & 2 \end{bmatrix}$;

в) $A=\begin{bmatrix} 2 & -2 \\ 0 & 0 \end{bmatrix}$; г) $A=\begin{bmatrix} \sigma & \omega \\ -\omega & \sigma \end{bmatrix}$.

6.4. Вычислить e^{At} любым методом для матриц примера 6.3.

6.5. Вычислить e^{At} численным методом для матриц:

а) $A=\begin{bmatrix} 8 & -8 & -2 \\ 4 & -3 & -2 \\ 3 & -4 & 1 \end{bmatrix}$; б) $A=\begin{bmatrix} -3 & 1 & 0 \\ 1 & -3 & 0 \\ 0 & 0 & -3 \end{bmatrix}$;

в) $A=\begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$; г) $A=\begin{bmatrix} -1 & -2 & 3 \\ 0 & 1 & -1 \\ -1 & -1 & 2 \end{bmatrix}$;

д) $A=\begin{bmatrix} 2 & -2 \\ 0 & 0 \end{bmatrix}$,

подготовив схему вычислений в соответствии с соотношениями (6.26), (6.27) и (6.28), положив в разложении матричной экспоненты $e^{A\Delta t}$

(6.28) $p=4$ и построив его минимальное представление с использованием (6.7), полагая $\psi(\alpha)=D(\alpha)$.

6.6. При каких свойствах матриц A и B справедливо $e^A \cdot e^B = e^B \cdot e^A$?

6.7. Доказать справедливость следующих равенств:

а) $e^{A(t+\tau)} = e^{At} \cdot e^{A\tau}$

б) $e^{-At} = (e^{At})^{-1}$

в) $\dfrac{d}{dt}e^{At} = Ae^{At} = e^{At}A.$

г) $e^{(A+B)t} = e^{At} \cdot e^{Bt}$, если $AB=BA$.

д) $\int\limits_0^t e^{At}dt = A^{-1}(e^{At}-I) = (e^{At}-I)A^{-1}$ при $\det(A)\neq 0$.

Решение вариантов задач

Решение задачи 6.1. Характеристический многочлен $D(\alpha)$ матрицы A имеет вид $D(\alpha)=(\alpha+2)^3$ так, что собственное значение $\lambda=-2$ характеризуется кратностью $m=n=3.$ В свою очередь характеристическая матрица $A-\lambda I = A+2I$ обладает нуль–пространством $N\{A+2I\}$ размерности $r_N=1$, которому принадлежит один собственный вектор $\xi=(1,-2,\ 4)^T$. В связи со сказанным нормальная форма Жордана матрицы A принимает канонический вид (6.17) и записывается в форме $J=\begin{bmatrix} -2 & 1 & 0 \\ 0 & -2 & 1 \\ 0 & 0 & -2 \end{bmatrix}$. Матрица T отношения подобия $TJ=AT$, так что $A=TJT^{-1}$, имеет представление

$$T=\left[T_1=\xi;T_2=\left(A+2I\right)^{+}T_1;T_3=\left(A+2I\right)^{+}T_2\right]=\begin{bmatrix} 1 & 0.8573 & 0.5442 \\ -2 & -0.7142 & -0.2314 \\ 4 & -0.5713 & -0.2517 \end{bmatrix}.$$

Тогда в силу свойства 6.2, а также представления (6.31) e^{Jt} искомая матричная экспонента e^{At} принимает вид

$$e^{At}=Te^{Jt}T^{-1}=T\begin{bmatrix} e^{-2t} & te^{-2t} & \frac{1}{2}t^2e^{-2t} \\ 0 & e^{-2t} & te^{-2t} \\ 0 & 0 & e^{-2t} \end{bmatrix}T^{-1}=$$

$$=T\left\{\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}e^{-2t}+\begin{bmatrix}0&1&0\\0&0&1\\0&0&0\end{bmatrix}te^{-2t}+\begin{bmatrix}0&0&1\\0&0&0\\0&0&0\end{bmatrix}\frac{1}{2}t^2e^{-2t}\right\}T^{-1}=$$

$$=\begin{bmatrix}1&0.857&0.544\\-2&-0.714&-0.231\\4&-0.571&-0.252\end{bmatrix}\left\{\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}e^{-2t}+\begin{bmatrix}0&1&0\\0&0&1\\0&0&0\end{bmatrix}te^{-2t}+\begin{bmatrix}0&0&1\\0&0&0\\0&0&0\end{bmatrix}\frac{1}{2}t^2e^{-2t}\right\}$$

$$\begin{bmatrix}0.048&-0.095&0.190\\-1.430&-2.431&-0.838\\4.004&4.005&1.001\end{bmatrix}=\left\{\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}e^{-2t}+\begin{bmatrix}2&1&0\\0&2&1\\-8&-12&-4\end{bmatrix}te^{-2t}+\begin{bmatrix}4&4&1\\-8&-8&-2\\16&16&4\end{bmatrix}\frac{1}{2}t^2e^{-2t}\right\}.$$

Поставленная задача решена.