

**ПРИОРИТЕТНЫЙ НАЦИОНАЛЬНЫЙ ПРОЕКТ «ОБРАЗОВАНИЕ»
РОССИЙСКИЙ УНИВЕРСИТЕТ ДРУЖБЫ НАРОДОВ**

К.П. ЛОВЕЦКИЙ, Л.А. СЕВАСТЬЯНОВ, Е.Б. ЛАНЕЕВ

**РЕГУЛЯРНЫЕ МЕТОДЫ
И АЛГОРИТМЫ РАСЧЕТА
ОБРАТНЫХ ЗАДАЧ
В МОДЕЛЯХ ОПТИЧЕСКИХ СТРУКТУР**

Учебное пособие

Москва

2008

*Инновационная образовательная программа
Российского университета дружбы народов*

**«Создание комплекса инновационных образовательных программ
и формирование инновационной образовательной среды,
позволяющих эффективно реализовывать государственные интересы РФ
через систему экспорта образовательных услуг»**

Экспертное заключение –

доктор физико-математических наук, профессор *И.В. Пузынин*

Ловецкий К.П., Севастьянов Л.А., Ланеев Е.Б.

Регулярные методы и алгоритмы расчета обратных задач в моделях оптических структур: Учеб. пособие. – М.: РУДН, 2008. – 155 с.

Учебное пособие посвящено методам устойчивого решения обратных задач в моделях взаимодействия электромагнитного излучения видимого диапазона с участками среды со сложной геометрией и сложным по составу диэлектрическим наполнением. Приоритетной задачей является использование этих методов для численного решения задач при проектировании оптических устройств. Пособие нацелено на применение теоретических знаний (регулярных методов решения некорректных задач) для разработки оригинальных методов и алгоритмов решения задач математического анализа и синтеза дифракционных оптических элементов и устройств.

Для магистров и аспирантов, обучающихся по направлению «Прикладная математика и информатика».

Учебное пособие выполнено в рамках инновационной образовательной программы Российского университета дружбы народов, направление «Комплекс экспортноориентированных инновационных образовательных программ по приоритетным направлениям науки и технологий», и входит в состав учебно-методического комплекса, включающего описание курса, программу и электронный учебник.

© Ловецкий К.П., Севастьянов Л.А., Ланеев Е.Б., 2008

Содержание

Общее описание курса	5
Иновационность курса	6
Тема 1. Системы уравнений с неточно заданными коэффициентами и правой частью	8
1.1. Влияние неточности исходной информации	8
1.2. Почти вырожденные матрицы	10
1.3. Обусловленность систем уравнений	11
1.4. Псевдорешения и псевдообратные матрицы	13
1.5. Для чего нужны псевдообратные матрицы	19
Тема 2. Решение вырожденных и плохо обусловленных систем линейных алгебраических уравнений	28
2.1. Корректность и некорректность по Адамару	31
2.2. Классические методы решения интегральных уравнений Фредгольма I рода	33
2.3. Метод наименьших квадратов Гаусса	38
2.4. Метод псевдообратной матрицы Мура-Пенроуза	41
Тема 3. Метод регуляризации нахождения нормального решения	44
3.1. Метод регуляризации Тихонова	44
3.2. Метод оптимальной фильтрации Калмана-Бьюси.....	53
3.3. Метод оптимальной линейной фильтрации Винера.....	57
Тема 4. Метод регуляризации нахождения нормального решения системы линейных алгебраических уравнений	60
4.1. Приближенное нахождение нормального решения по неточно известной правой части.....	60
4.2. Приближенное нахождение нормального решения по неточно заданным правой части и матрице	62
Тема 5. Метод регуляризации решения линейных интегральных уравнений первого рода	64
5.1. Существование регуляризирующих операторов для интегральных уравнений первого рода	64
5.2. Интегральное уравнение I рода – некорректная задача:	66
5.3. Редукция задачи построения регуляризирующих операторов к классической вариационной задаче минимизации функционалов с ограничениями.....	70
5.4. Получение семейства регуляризирующих операторов с помощью минимизации сглаживающих функционалов	72

5.5. Алгоритм нахождения приближенных решений, легко реализуемый на ЭВМ.....	73
5.6. О дискретизации задачи нахождения приближенных решений интегральных уравнений первого рода.....	77
Тема 6. Представление решения в виде конечных рядов Фурье.....	80
6.1. Явление Гиббса.....	80
6.2. Метод Фейера:.....	81
6.3. Сигма-множители Ланцоша.....	83
6.4. Сравнение методов сходимости.....	85
6.5. Техника дифференцирования по Ланцошу.....	86
Тема 7. Устойчивые методы суммирования рядов Фурье с приближенными в метрике l_2 коэффициентами.....	87
7.1. Задача устойчивого суммирования ряда Фурье.....	87
7.2. Классы устойчивых методов суммирования рядов Фурье.....	89
Приложение 1. Основные типы уравнений.....	97
Интегральные уравнения.....	97
Системы линейных алгебраических уравнений.....	104
Системы линейно-нелинейных уравнений.....	104
Операторные уравнения.....	105
Некоторые сведения из линейной алгебры.....	106
Система линейных алгебраических уравнений (СЛАУ).....	106
Характеристическое уравнение и типы матриц.....	107
Нормы векторов и матриц.....	109
Число обусловленности.....	109
Умножение матриц и векторов.....	110
Элементы теории вероятностей.....	112
Приложение 2. Обобщенные функции, формула Эйлера и интегральные преобразования.....	116
Элементы теории обобщенных функций.....	116
Формула Эйлера.....	119
Интегральные преобразования.....	122
Литература.....	133
Описание курса и программа.....	134

Общее описание курса

Основы развития устойчивых приближенных методов решения некорректных задач были заложены в фундаментальной работе А.Н. Тихонова «Об устойчивости обратных задач» // ДАН СССР, 1943, т. 39, № 5, где был введен важный класс так называемых обратных задач, связанных с восстановлением количественных характеристик среды по порождаемым ею физическим полям, доступным для измерения. Интенсивное развитие методов решения неустойчивых задач было предопределено широким внедрением компьютеров в математические исследования, что по законам обратной связи вызвало поток разнообразных задач, для решения которых потребовалось развитие новых приближенных методов. Одним из значимых сегментов обратных некорректных задач являются задачи компьютерной оптики, в частности, задач проектирования дифракционных оптических элементов и устройств. Методы устойчивого решения обратных задач в моделях взаимодействия электромагнитного излучения видимого диапазона с участками среды со сложной геометрией и сложным по составу диэлектрическим наполнением занимают важное место в ряду регулярных методов решения некорректных задач.

Целью курса является обучение теоретическим основам регулярных методов решения некорректных задач. Теоретическое изложение при этом опирается на конкретные регулярные методы решения прикладных задач математического анализа и синтеза дифракционных оптических элементов и устройств, оптических наноструктур.

Задача курса - обучение студентов умению и навыкам использования методов устойчивого решения обратных задач в моделях взаимодействия электромагнитного излучения видимого диапазона с участками среды со сложной геометрией и сложным по составу диэлектрическим наполнением.

Более того, приоритетной задачей является применение этих методов для численного решения задач при проектировании оптических устройств. Они обязаны квалифицированно применять теоретические знания (регулярных методов решения некорректных задач) для разработки оригинальных методов и алгоритмов решения задач математического анализа и синтеза дифракционных оптических элементов и устройств.

Инновационность курса

Курс является инновационным по содержанию и литературе, он включает последние научные достижения в области решения задач дифракционной оптики, когда характерные размеры исследуемых объектов не превышают либо сравнимы с длиной волны оптического излучения. Эта область знаний интенсивно развивалась в последнее время, но лишь недавно были созданы устойчивые алгоритмы и разработаны численные методы решения задач для многослойных решеток. Следует отметить, что для оптических однослойных и многослойных решеток с характерными размерами больше длины волны оптического излучения устойчивые методы решения известны с середины прошлого века. Сейчас алгоритмы решения оптических задач в субволновой области распространяются на объекты со сложной геометрией, такие как двухмерные решетки с произвольным профилем, трехмерные решетки (фотонные кристаллы) и на анизотропные материалы. Они востребованы, поскольку позволяют создавать математические модели взаимодействия излучения с веществом в наномасштабах, а затем с их помощью проектировать новые эффективные устройства в высокотехнологичных областях медицины, энергетики, инфокоммуникаций и приборостроения.

В ходе проведения занятий по этому курсу разработчики предполагают использование традиционных методик преподавания, принятых в странах болонской системы образования, то есть с использованием кредитной системы оценки знаний.

Наряду с традиционными элементами преподавания математических методов решения прикладных задач разработчики курса предполагают воспользоваться хорошо зарекомендовавшим себя опытом МФТИ и подобных вузов. А именно, в рамках подпрограммы «Оптика наноструктур» осуществляется закупка уникального аналитического оборудования для измерения разнообразных характеристик оптических наноустройств с целью использования этого оборудования в учебном процессе и для проведения научно-исследовательских работ преподавателями, аспирантами и студентами.

По окончании магистратуры по направлению «Оптика наноструктур» выпускники Российского университета дружбы народов станут конкурентоспособными специалистами в области проектирования современных оптических устройств, которые не будут испытывать затруднений при последующем трудоустройстве.

Курс базируется на публикациях научных статей мировых лидеров исследований в данной области, диссертационных работах их учеников, включающих работы по непосредственному моделированию, дизайну и последующему изготовлению лабораторных образцов оптических элементов и устройств. В список дополнительной и рекомендуемой литературы включены все научно-исследовательские публикации, положенные в основу предлагаемого курса.

В качестве практических заданий, курсовых работ и тем рефератов слушателям магистерской программы будут предложены актуальные проблемы и задачи, решение которых востребовано современным уровнем развития высокотехнологичных отраслей промышленности и научно-исследовательских лабораторий.

Тема 1. Системы уравнений с неточно заданными коэффициентами и правой частью

1.1. Влияние неточности исходной информации

Если какая-либо практическая задача сводится к системе линейных уравнений, то коэффициенты и свободные члены системы, как правило, известны не точно, а с некоторыми погрешностями. Кроме того, как было сказано выше, запись исходных данных и вычисления связаны с некоторыми ошибками округления, влияние которых равносильно некоторому искажению коэффициентов и свободных членов.

Избежать указанных искажений мы не можем, но можем, во-первых, оценить получаемую погрешность, а, во-вторых, постараться выбрать такой метод решения системы, который не увеличивал бы неопределённость результата, уже заложенную в самой системе.

Мы будем стоять на следующей точке зрения: задана система линейных уравнений

$$\mathbf{Ax} = \mathbf{b}, \quad (1.1)$$

которую мы будем называть *исходной*, или *невозмущённой* системой. Рассматривается ещё одна система, называемая *возмущённой* системой, про которую известно, что её коэффициенты и свободные члены лежат в заданных интервалах

$$(a_{ij} - \Delta a_{ij}, a_{ij} + \Delta a_{ij}), (b_i - \Delta b_i, b_i + \Delta b_i), \quad (1.2)$$

где a_{ij} и b_i - коэффициенты и свободные члены исходной системы.

Погрешностью решения мы будем называть разность между решениями исходной и возмущённой систем. (При этом предполагается, что каждая из систем имеет единственное решение.) Согласно этому определению погрешность - столбец, и оценить погрешность можно только по какой-либо норме в арифметическом пространстве. В примерах этого параграфа рассматривается исключительно евклидова норма.

Рассмотрим простейший случай – систему из двух линейных уравнений с двумя неизвестными. Невозмущённая система имеет вид

$$\begin{aligned} ax + by &= c, \\ a_1x + b_1y &= c_1, \end{aligned} \quad (1.3)$$

и пусть известно, что коэффициенты при переменных заданы точно, а возмущены только свободные члены, т.е. $\Delta a = \Delta b = \Delta a_1 = \Delta b_1 = 0$.

Геометрически каждое уравнение невозмущённой системы изображается прямой линией на плоскости. Соответствующее уравнение возмущённой системы – параллельной ей прямой, которая лежит внутри некоторой полосы. Для первого уравнения, например, полоса ограничена прямыми $ax + by = c - \Delta c$ и $ax + by = c + \Delta c$.

Таким образом, решения возмущённой системы лежат внутри параллелограмма, являющегося пересечением полос.

Погрешность решения изображается вектором $(\Delta x, \Delta y)$, длина которого может достигать половины длины наибольшей диагонали параллелограмма. Поэтому искажающее влияние возмущений свободных членов системы при полосах одинаковой ширины тем больше, чем меньше угол между прямыми (1.3).

В общем случае, когда и коэффициенты истинной системы отличаются от коэффициентов исходной, прямые, изображающие уравнения, не только сдвигаются, но и поворачиваются. Параллелограмм заменяется более сложной фигурой, но общий результат остаётся тот же: чем меньше угол между прямыми (1.3), тем, как говорится, *хуже обусловлена* система, т.е. такое же возмущение коэффициентов может вызывать большую погрешность решения.

1.2. Почти вырожденные матрицы

Возмущение коэффициентов системы линейных уравнений может не просто немного исказить её решение, а иметь более серьёзные последствия.

Система линейных уравнений

$$x + 0,99y = 1,01,$$

$$x + 1,01y = 0,99$$

имеет единственное решение, но за счёт изменения её коэффициентов и свободных членов на 1% можно получить как противоречивую систему, так и систему, имеющую бесконечно много решений.

Сейчас нам следует подчеркнуть принципиальную важность этого явления. Если числа не могут быть заданы точно, то стирается чёткая грань между вырожденными и невырожденными матрицами. Появляется класс *почти вырожденных* матриц, границы которого зависят от принятой в данном конкретном исследовании меры точности. Про почти вырожденную матрицу нельзя сказать, вырождена она или нет, так как, изменяя её элементы в пределах заданной точности, можно получить как вырожденную, так и невырожденную матрицу. Поэтому при приближённых вычислениях надо с большой осторожностью относиться к фразам типа: «пусть матрица A невырождена...», которыми пестрит курс линейной алгебры.

Если в системе линейных уравнений матрица оказалась почти вырожденной, то лучше вернуться к постановке той задачи, которая привела к рассматриваемой системе линейных уравнений, для того чтобы получить дополнительную информацию о системе. Может случиться, например, что по смыслу задачи решение должно быть единственным. В действительности, надо чётко представлять себе, что количественная неопределённость исходной информации в этом случае оказывает качественное влияние на решение, и правильнее будет сосредоточить усилия не на решении системы, а на уточнении постановки задачи.

1.3. Обусловленность систем уравнений

Пусть дана исходная система линейных уравнений

$$\mathbf{Ax} = \mathbf{b}, \quad (1.4)$$

где \mathbf{A} – квадратная матрица порядка n и $\det \mathbf{A} \neq 0$. Рассмотрим возмущённую систему

$$(\mathbf{A} + \delta\mathbf{A})\mathbf{y} = \mathbf{b} + \delta\mathbf{b}. \quad (1.5)$$

Прежде всего, не ясно, будет ли система (1.5) иметь единственное решение так же, как и (1.4). Ниже мы наложим на $\delta\mathbf{A}$ достаточное для этого условие. Наша ближайшая задача – оценить при этом условии норму разности решений обеих систем.

Предложение 1. Пусть матричная норма квадратной матрицы \mathbf{B} удовлетворяет условию $\|\mathbf{B}\| \leq \rho < 1$. Тогда существует матрица $(\mathbf{E} + \mathbf{B})^{-1}$ и $\|(\mathbf{E} + \mathbf{B})^{-1}\| \leq (1 - \rho)^{-1}$.

Доказательство. Из оценки спектрального радиуса матрицы вытекает, что все собственные значения матрицы \mathbf{B} лежат в круге $|\lambda| \leq \rho$, т.е. внутри круга сходимости разложения функции $(1 + \lambda)^{-1}$ по степеням λ . Это гарантирует существование матрицы $(\mathbf{E} + \mathbf{B})^{-1}$, равной сумме ряда $\mathbf{E} - \mathbf{B} + \mathbf{B}^2 - \dots$

Для частичной суммы этого ряда имеем оценку

$$\|\mathbf{S}_k\| \leq \sum_0^k \|\mathbf{B}^k\| \leq \sum_0^k \|\mathbf{B}\|^k < \frac{1}{1 - \rho}.$$

Отсюда переходом к пределу при $k \rightarrow \infty$ и получаем требуемое неравенство.

Перейдём теперь к оценке нормы возмущения решения, т.е. $\|\delta\mathbf{x}\| = \|\mathbf{y} - \mathbf{x}\|$, где \mathbf{y} – решение системы (1.5), \mathbf{x} – решение (1.4). Для этого вычтем (1.4) и (1.5). Мы получим

$$(\mathbf{A} + \delta\mathbf{A})(\mathbf{x} + \delta\mathbf{x}) - \mathbf{Ax} = \delta\mathbf{b},$$

или

$$(\mathbf{A} + \delta\mathbf{A})\delta\mathbf{x} + \delta\mathbf{A}\mathbf{x} = \delta\mathbf{b}. \quad (1.6)$$

Запишем $\mathbf{A} + \delta\mathbf{A}$ в виде $\mathbf{A}(\mathbf{E} + \mathbf{A}^{-1}\delta\mathbf{A})$ и предположим, что

$$\|\mathbf{A}^{-1}\| \cdot \|\delta\mathbf{A}\| = \rho < 1. \quad (1.7)$$

Тогда $\|\mathbf{A}^{-1}\delta\mathbf{A}\| \leq \rho$ и, согласно предложению 1, матрица $\mathbf{E} + \mathbf{A}^{-1}\delta\mathbf{A}$ имеет обратную. Следовательно,

$$(\mathbf{A} + \delta\mathbf{A})^{-1} = (\mathbf{E} + \mathbf{A}^{-1}\delta\mathbf{A})^{-1} \mathbf{A}^{-1},$$

и из (1.6)

$$\delta\mathbf{x} = (\mathbf{E} + \mathbf{A}^{-1}\delta\mathbf{A})^{-1} \mathbf{A}^{-1} \delta\mathbf{b} - (\mathbf{E} + \mathbf{A}^{-1}\delta\mathbf{A})^{-1} \mathbf{A}^{-1} \delta\mathbf{A}\mathbf{x}.$$

Отсюда

$$\|\delta\mathbf{x}\| \leq \frac{1}{1-\rho} \|\mathbf{A}^{-1}\| \cdot \|\delta\mathbf{b}\| - \frac{1}{1-\rho} \|\mathbf{A}^{-1}\| \cdot \|\delta\mathbf{A}\| \cdot \|\mathbf{x}\|.$$

Из равенства (1.4) следует, что $\|\mathbf{b}\| \leq \|\mathbf{A}\| \cdot \|\mathbf{x}\|$. Усилим неравенство, умножив первый член правой части на $\|\mathbf{A}\| \cdot \|\mathbf{x}\| / \|\mathbf{b}\|$. Тогда

$$\|\delta\mathbf{x}\| \leq \frac{c(\mathbf{A})\|\mathbf{x}\| \|\delta\mathbf{b}\|}{1-\rho \|\mathbf{b}\|} - \frac{c(\mathbf{A})\|\mathbf{x}\| \|\delta\mathbf{A}\|}{1-\rho \|\mathbf{A}\|},$$

где

$$c(\mathbf{A}) = \|\mathbf{A}^{-1}\| \cdot \|\mathbf{A}\|. \quad (1.8)$$

Разделим на $\|\mathbf{x}\|$ и учтём, что

$$\rho = \|\delta\mathbf{A}\| \cdot \|\mathbf{A}^{-1}\| = c(\mathbf{A}) \frac{\|\delta\mathbf{A}\|}{\|\mathbf{A}\|}.$$

Мы получим окончательную оценку

$$\frac{\|\delta\mathbf{x}\|}{\|\mathbf{x}\|} \leq \frac{c(\mathbf{A})}{1 - c(\mathbf{A}) \frac{\|\delta\mathbf{A}\|}{\|\mathbf{A}\|}} \left(\frac{\|\delta\mathbf{A}\|}{\|\mathbf{A}\|} + \frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|} \right).$$

Число $c(\mathbf{A})$, введённое формулой (1.8), называется *числом обусловленности* матрицы \mathbf{A} в рассматриваемой норме.

Теперь мы можем сформулировать следующий результат.

Предложение 2. Пусть матрица коэффициентов и столбец свободных членов системы (1.4) получили возмущения δA и δb , причём в некоторой матричной норме $\|A^{-1}\| \cdot \|\delta A\| < 1$. Тогда возмущённая система имеет единственное решение, и относительное возмущение $\|\delta x\| / \|x\|$ решения системы (1.4) оценивается через относительные возмущения $\alpha = \|\delta A\| / \|A\|$ и $\beta = \|\delta b\| / \|b\|$ матрицы системы и столбца свободных членов формулой

$$\frac{\|\delta x\|}{\|x\|} \leq \frac{c(\mathbf{A})(\alpha + \beta)}{1 - c(\mathbf{A})\alpha}, \quad (1.9)$$

где $c(\mathbf{A})$ – число обусловленности матрицы \mathbf{A} в рассматриваемой норме.

Примечательно, что в правую часть (1.9) входят только относительные возмущения \mathbf{A} и \mathbf{b} . Матрица \mathbf{A} представлена только своим числом обусловленности, а столбец \mathbf{b} не входит совсем.

1.4. Псевдорешения и псевдообратные матрицы

В практических задачах часто бывает нужно найти решение, удовлетворяющее большому числу возможно противоречивых требований. Если такая задача сводится к системе линейных уравнений, то система оказывается, вообще говоря, несовместной. В этом случае задача может быть решена только путём выбора некоторого компромисса – все требования могут быть удовлетворены не полностью, а лишь до некоторой степени. Поясним это следующим примером.

Пусть из физических соображений можно считать, что в некоторой области их изменения величины y и x связаны линейной зависимостью вида $y = kx + b$, а коэффициенты должны быть установлены

экспериментально. Экспериментальные данные представляют собой m точек на координатной плоскости $(x_1, y_1), \dots, (x_m, y_m)$.

Если эти пары значений действительно связаны искомой зависимостью, то подстановка их в уравнение приводит нас к системе из m линейных уравнений для двух неизвестных k и b :

$$y_i = kx_i + b, \quad i = 1, \dots, m.$$

При любых различных x_i и x_j пара точек (x_i, y_i) и (x_j, y_j) определяет прямую. Но другая пара точек определяет другую прямую, и у нас нет оснований выбрать какую-нибудь одну из всех прямых.

Если экспериментальные данные в достаточной степени заслуживают доверия, то несовместность системы служит основанием для того, чтобы отвергнуть гипотезу о линейной зависимости. Вопрос о совместимости экспериментальных данных с гипотезой линейной зависимости решается статистическим анализом.

Пусть точность исходной информации допускает существование линейной зависимости. В этом случае то, что в действительности нужно, - это найти такую прямую на координатной плоскости, которая, может быть, не проходит ни через одну пару экспериментальных точек, или даже ни через одну из точек, но в каком-то смысле возможно более близко расположена ко всем точкам.

Обычно в этой задаче удалённость точки от прямой измеряют не расстоянием, а разностью ординат $y_i - kx_i - b$, и выбирают прямую так, чтобы сумма квадратов всех таких разностей была минимальна. Коэффициенты k_0 и b_0 уравнения этой прямой дают некоторое решение стоящей перед нами задачи, которое отнюдь не является решением системы линейных уравнений (вообще не имеющей решений). Можно считать числа k_0 и b_0 обобщённым решением системы или, как говорят, псевдорешением. Точное определение этого понятия будет дано ниже.

Мы будем рассматривать систему линейных уравнений

$$\mathbf{Ax} = \mathbf{b}, \quad (1.10)$$

с матрицей A размеров $m \times n$. Буква r будет обозначать ранг этой матрицы. Никаких условий на m, n и r , вообще говоря, не накладывается. Поскольку x - столбец высоты n , а b - столбец высоты m , для геометрической иллюстрации естественно будет использовать арифметические пространства R_n и R_m . Под нормой столбца x с элементами x^1, \dots, x^n мы будем понимать его евклидову норму, т.е. число

$$\|x\| = \sqrt{x^T x} = \sqrt{(x^1)^2 + \dots + (x^n)^2}.$$

Невязкой, которую даёт столбец x при подстановке в систему уравнений (1.10), называется столбец $\mathbf{u} = \mathbf{b} - \mathbf{Ax}$.

Решение системы – это столбец, дающий нулевую невязку.

Если система (1.10) несовместна, естественно постараться найти столбец x , который даёт невязку с минимальной нормой, и если такой столбец найдётся, считать его обобщённым решением системы. Разумеется, если система несовместна, её решение будет также и обобщённым решением.

Для сравнения невязок воспользуемся евклидовой нормой и, следовательно, будем искать столбец x , для которого минимальна величина

$$\|\mathbf{u}\|^2 = (\mathbf{b} - \mathbf{Ax})^T (\mathbf{b} - \mathbf{Ax}). \quad (1.11)$$

Рассматривая элементы столбца x как независимые переменные, найдём полный дифференциал $\|\mathbf{u}\|^2$. Как нетрудно проверить,

$$d\|\mathbf{u}\|^2 = -dx^T A^T (b - Ax) - (b - Ax)^T A dx.$$

Так как второе слагаемое, будучи матрицей первого порядка, не меняется при транспонировании, имеем

$$d\|\mathbf{u}\|^2 = -2dx^T A^T (b - Ax).$$

Поэтому дифференциал равен нулю тогда и только тогда, когда

$$A^T Ax = A^T b . \quad (1.12)$$

Эта система линейных уравнений по отношению к системе (1.10) называется *нормальной* системой. Она является следствием системы (1.10), но независимо от совместности системы (1.10) справедливо

Предложение 1. *Нормальная система уравнений обязательно совместна.*

Доказательство. Матрица $A^T A$ симметрична, и потому транспонированная однородная система для системы (1.12) имеет вид $A^T Ay = 0$. Для любого решения этой системы имеют место равенства, последовательно вытекающие одно из другого:

$$y^T A^T Ay = (Ay)^T (Ay) = 0, \quad Ay = 0, \quad y^T (A^T b) = 0 .$$

Последнее из них означает, для системы (1.12) выполнено условие теоремы Фредгольма. Это заканчивает доказательство.

Предложение 2. *Точная нижняя грань квадрата нормы невязки достигается для всех решений нормальной системы (1.12) и только для них.*

Доказательство. Написав формулу (1.11) для столбца $x_0 + \Delta x$ и раскрыв скобки, мы получаем

$$(b - A(x_0 + \Delta x))^T (b - A(x_0 + \Delta x)) = \\ (b - Ax_0)^T (b - Ax_0) - 2\Delta x^T A^T (b - Ax_0) + \Delta x^T A^T A \Delta x .$$

Последнее слагаемое неотрицательно, так как $\Delta x^T A^T A \Delta x = (A \Delta x)^T (A \Delta x) \geq 0$. Если x_0 удовлетворяет системе (1.12), то второе слагаемое равно нулю, и тогда добавление Δx не уменьшает значения функции, каков бы этот столбец Δx ни был.

Обратно, для функции, определённой для всех x , точная нижняя грань может достигаться только в точке локального экстремума, а в таких точках

дифференциал равен нулю, и выполнено условие (1.12). Предложение доказано.

По известной теореме всё множество решений нормальной системы может быть описано формулой $x = x_0 + z$, в которой x_0 - некоторое фиксированное решение нормальной системы, а z - произвольное решение однородной системы $A^T A z = 0$. Как мы видели при доказательстве предложения 1, последняя система эквивалентна системе $Az = 0$, и мы можем считать, что решение нормальной системы (1.12) определено с точностью до произвольного решения однородной системы $Az = 0$. Таким образом, справедливо

Предложение 3. *Нормальная система имеет единственное решение тогда и только тогда, когда система $Az = 0$ имеет только тривиальное решение, т.е. столбцы матрицы A линейно независимы. В частности, это будет выполнено, если матрица A невырождена.*

Если решение нормальной системы не единственно, то возникает задача выбрать какое-то одно из решений, и выбирается решение с минимальной нормой.

Определение. *Нормальным псевдорешением системы линейных уравнений называется столбец с минимальной нормой среди всех столбцов, дающих минимальную по норме невязку при подстановке в эту систему.*

Поскольку не возникает опасности недоразумения, мы будем называть нормальное псевдорешение просто псевдорешением[1].

Предложение 4. *Пусть x_b и x_c - нормальные псевдорешения двух систем линейных уравнений $Ax = b$ и $Ax = c$. Тогда $\beta x_b + \gamma x_c$ является нормальным псевдорешением системы $Ax = \beta b + \gamma c$.*

Доказательство. Из $A^T A x_b = A^T b$ и $A^T A x_c = A^T c$ следует, что $\beta x_b + \gamma x_c$ удовлетворяет нормальной системе: $A^T A(\beta x_b + \gamma x_c) = A^T(\beta b + \gamma c)$. Далее,

существуют столбцы z_b и z_c такие, что $x_b = A^T z_b$ и $x_c = A^T z_c$. Поэтому $\beta x_b + \gamma x_c = A^T (\beta z_b + \gamma z_c)$, что и заканчивает доказательство.

Естественно, предложение 4 может быть распространено на линейные комбинации произвольного числа столбцов.

Рассмотрим несколько очень простых примеров.

Система из двух уравнений с одной неизвестной:

$$x = 1, \quad x = 2.$$

Нормальная система уравнений для этой системы есть

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} x = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix},$$

или $2x = 3$. Отсюда следует, что псевдорешение равно $3/2$.

Система из одного уравнения с двумя неизвестными:

$$x^1 + x^2 = 2.$$

Нормальной системой уравнений будет система

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} x^1 \\ x^2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} 2,$$

содержащая то же уравнение, повторенное дважды. Её общее решение

$$\begin{pmatrix} x^1 \\ x^2 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} + \alpha \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Псевдорешение будет тем решением, которое получается умножением

$A^T = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ на некоторый столбец z высоты $m=1$. Легко видеть, что таким

решением будет

$$\begin{pmatrix} x^1 \\ x^2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Система из одного уравнения с одним неизвестным $\alpha x = \beta$. Если $\alpha \neq 0$, то псевдорешение совпадает с решением $x = \beta / \alpha$. Если же $\alpha = 0$,

то любое число x при подстановке даёт ту же самую невязку β с нормой $|\beta|$. Из всех чисел мы должны выбрать то, норма которого минимальна, т.е. 0. Итак, псевдорешение уравнения $0x = \beta$ есть 0.

Результат в этом примере такой, какой был бы, если бы мы доопределили функцию $f(\alpha) = \frac{1}{\alpha}$ в нуле числом 0. Здесь видно, что псевдорешение не является непрерывной функцией от элементов матрицы и столбца свободных членов. Этот же недостаток имеет место и в общем случае системы из m уравнений и n неизвестными.

Система линейных уравнений с нулевой матрицей $0x = b$. Псевдорешение находится так же, как и в примере 3: все векторы дают одну и ту же невязку, равную b , и минимальной нормой обладает нулевой вектор, который и будет псевдорешением.

1.5. Для чего нужны псевдообратные матрицы

В процессе работы со статистическими данными часто требуется найти «решение» системы, которая решений (в обычном смысле) не имеет. Выходом из ситуации является нахождение таких значений неизвестных параметров, что все условия системы выполняются «в некоторой степени».

Например, рассмотрим задачу линейной регрессии. Предположим, что переменная есть линейная комбинация предикторных переменных X_1, X_2, \dots, X_m с коэффициентами $\theta_1, \theta_2, \dots, \theta_m$, то есть $Y = \theta_1 X_1 + \theta_2 X_2 + \dots + \theta_m X_m$.

Следовательно, измерения y_i ($i = 1, \dots, n$) переменной Y , отвечающие заданным (не обязательно различным) значениям $x_{i1}, x_{i2}, \dots, x_{im}$, имеют вид

$$y_i = \theta_1 x_{i1} + \theta_2 x_{i2} + \dots + \theta_m x_{im}.$$

Для того чтобы найти значения коэффициентов, необходимо решить систему из n уравнений с m неизвестными. Но эта система может быть несовместной. Вот тут и возникают псевдообратные матрицы.

Определение псевдообратной матрицы

Способ 1

Определение 1. Матрица \mathbf{A}^+ называется псевдообратной матрицей для матрицы \mathbf{A} , если

$$\mathbf{A}\mathbf{A}^+\mathbf{A} = \mathbf{A}.$$

Из этого определения сразу вытекает, что если матрица \mathbf{A} имеет размер $m \times n$, то псевдообратная матрица \mathbf{A}^+ имеет размер $n \times m$.

Способ 2

Опишем и другой, часто встречающийся в литературе, подход к определению этого понятия.

Сначала введём понятие псевдорешения системы уравнений.

Пусть нам дана система уравнений

$$\mathbf{Ax} = \mathbf{b}, \quad (1.13)$$

где \mathbf{A} - матрица размера $m \times n$, \mathbf{b} - вектор из m элементов.

Любое решение этой системы является также и решением системы

$$\mathbf{A}^T \mathbf{Ax} = \mathbf{A}^T \mathbf{b}. \quad (1.14)$$

Определение 2. Псевдорешением системы (1.13) называется решение системы (1.14) с минимальной нормой среди всех столбцов, имеющих минимальную невязку (норма вектора равна квадратному корню из суммы квадратов компонент вектора), а невязкой решения системы (1.13) называется норма вектора $\mathbf{Ax} - \mathbf{b}$.

Поясним определение (1.14) на примере.

Пример 1. Требуется найти псевдорешение системы

$$\begin{cases} x_1 + x_2 = 4, \\ x_1 + x_2 = 6. \end{cases}$$

Очевидно, что эта система «обычных» решений не имеет.

Шаг 1. Перепишем эту систему в матричном виде (то есть в виде (1.13))

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \end{pmatrix}.$$

Запишем систему (1.14). Для этого домножим левую и правую части слева на матрицу $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 4 \\ 6 \end{pmatrix},$$

$$\begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 10 \\ 10 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 5 \end{pmatrix}.$$

Таким образом, система (1.14) – это просто дважды повторённое уравнение $x_1 + x_2 = 5$.

Шаг 2. Любое решение системы (1.14) можно записать в виде

$$x_1 = a, x_2 = 5 - a.$$

В общем случае на этом шаге нужно найти частное решение неоднородной системы и общее решение однородной системы.

В данной ситуации в качестве частного решения можно взять $(0; 5)$, а общим решением однородной «системы» $x_1 + x_2 = 0$ является пара $(a, -a)$.

Шаг 3. Теперь нужно найти решение с минимальной невязкой.

В нашем случае:

$$\mathbf{Ax} - \mathbf{b} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} a \\ 5 - a \end{pmatrix} - \begin{pmatrix} 4 \\ 6 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \mathit{const}$$

Поэтому любое решение есть решение с минимальной невязкой.

Шаг 4. Осталось выбрать из решений с минимальной невязкой (то есть в данном случае из всех решений) решение с минимальной нормой

$$\left\| \begin{pmatrix} a \\ 5 - a \end{pmatrix} \right\|^2 = a^2 + (5 - a)^2 = 2a^2 - 10a + 25.$$

Минимум этого квадратного трёхчлена достигается в точке $a = 5 / 2$.

Таким образом, псевдорешением системы является пара $x_1 = 5 / 2, x_2 = 5 / 2$.

Можно доказать, что:

- 1) система (1.14) всегда имеет хотя бы одно решение;
- 2) решение системы (1.14) с минимальной нормой среди всех столбцов, имеющих минимальную невязку всегда единственно.

Таким образом, определение 2 корректно.

Также необходимо отметить, что если матрица имеет обратную, то псевдообратная матрица совпадёт с обратной.

Теперь определим термин псевдообратная матрица через понятие псевдорешения.

Определение 1*. Псевдообратной матрицей для матрицы \mathbf{A} размера $m \times n$ называется матрица \mathbf{A}^+ , столбцы которой – псевдорешения систем вида $\mathbf{A}\mathbf{x} = \mathbf{e}_i$, где \mathbf{e}_i - i -й столбец единичной матрицы порядка m .

Определения 1 и 1* эквивалентны.

Свойства псевдообратной матрицы

1. Решением системы (1.13) является вектор $\mathbf{x}_0 = \mathbf{A}^+\mathbf{b}$.

Это свойство показывает, что псевдообратная матрица есть некоторое обобщение обратной (кстати, псевдообратную матрицу часто называют обобщенной обратной).

2. Критерий псевдообратной матрицы

Матрица является псевдообратной тогда и только тогда, когда $\mathbf{A}^T \mathbf{A} \mathbf{X} = \mathbf{A}^T$.

3. Другое характеристическое свойство

Псевдообратная матрица \mathbf{A}^+ и только она является решением уравнений (рассмотренных по отдельности)

$$(\mathbf{A}\mathbf{X})^T = \mathbf{A}\mathbf{X}, (\mathbf{X}\mathbf{A})^T = \mathbf{X}\mathbf{A},$$

$$\mathbf{XAX} = \mathbf{X}.$$

4. Для некоторых частных случаев существуют простые формулы для нахождения псевдообратной матрицы. Например, если столбцы матрицы независимы, то $\mathbf{A}^+ = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$.

Если же независимы строки, то $\mathbf{A}^+ = \mathbf{A}^T (\mathbf{A} \mathbf{A}^T)^{-1}$.

5. Для наших целей (практическое нахождение псевдообратной матрицы) наиболее важным является следующее свойство. Пусть \mathbf{A} - $m \times n$ ранга r , $\mathbf{A} = \mathbf{B}\mathbf{C}$ - скелетное разложение матрицы \mathbf{A} (то есть \mathbf{B} - матрица размера $m \times r$, а \mathbf{C} - $r \times n$; значит, \mathbf{B} и \mathbf{C} - матрицы максимального ранга и к ним применимо четвёртое свойство).

Оказывается, что, если известно скелетное разложение, то псевдообратную матрицу можно вычислить следующим образом:

$$\mathbf{A}^+ = \mathbf{C}^+ \mathbf{B}^+ = \mathbf{C}^T (\mathbf{C} \mathbf{C}^T)^{-1} (\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T.$$

Методы вычисления псевдообратной матрицы

1. С помощью скелетного разложения матрицы

Прямое получение скелетного разложения.

Пусть дана матрица \mathbf{A} размера $m \times n$ и ранга r . При помощи элементарных действий со строками эта матрица может быть преобразована в матрицу, r столбцов которой есть первые r столбцов единичной матрицы (пусть это столбцы с номерами j_1, j_2, \dots, j_r).

Введём следующие обозначения. Пусть \mathbf{C} - $r \times n$ - матрица из первых r строк исходной матрицы, а \mathbf{B} - $m \times r$ - матрица, составленная из столбцов исходной матрицы с номерами j_1, j_2, \dots, j_r .

Оказывается, что: $\mathbf{A} = \mathbf{B}\mathbf{C}$ и это разложение – скелетное.

Достоинством этого подхода является относительная простота нахождения матриц \mathbf{B} и \mathbf{C} , а недостатками – трудоёмкость вычислений

псевдообратной матрицы по формуле $\mathbf{A}^+ = \mathbf{C}^+ \mathbf{B}^+ = \mathbf{C}^T (\mathbf{C} \mathbf{C}^T)^{-1} (\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T$ и неопределённость с малостью последних строк матрицы.

Получение скелетного разложения с помощью QR - разложения.

Как известно, любую квадратную матрицу \mathbf{A} можно представить в виде произведения двух матриц $\mathbf{A} = \mathbf{Q} \mathbf{R}$, где \mathbf{Q} - ортогональная матрица, а \mathbf{R} - верхняя треугольная. Для получения этого разложения можно воспользоваться, например, методом отражений.

Если же \mathbf{A} - $m \times n$ - матрица, то для неё существует представление такого же вида, где \mathbf{Q} - ортогональная матрица размера $m \times n$, а \mathbf{R} - матрица размера $m \times n$ с элементами r_{ij} равными 0 при $i > j$. Это разложение можно получить, действуя по алгоритму метода отражений [2].

Нас будет интересовать представление матрицы \mathbf{A} , называемое QR - разложением.

Простейший способ получения этого разложения для $m \times n$ - матриц ($m > n$) максимального ранга состоит в рассмотрении матрицы \mathbf{A} как набора из n столбцов. К этим векторам можно применить процесс ортогонализации Грамма – Шмидта и построить такую верхнюю треугольную матрицу \mathbf{U} , что столбцы матрицы $\mathbf{Q} = \mathbf{A} \mathbf{U}$ будут ортогональны и нормированы. После этого следует вычислить обратную матрицу от матрицы \mathbf{U} (обозначим её $\mathbf{R} : \mathbf{R} = \mathbf{U}^{-1} \mathbf{A}$). Тогда разложение $\mathbf{A} = \mathbf{Q} \mathbf{R}$ и есть QR - разложение.

Существуют и более «продвинутые» методы получения QR - разложения, например, так называемый метод переортогонализации.

Теперь покажем, как, зная QR - разложение, можно построить скелетное разложение матрицы.

Без ограничения общности можно считать, что первые r столбцов матрицы имеют максимальный ранг (если это не так, то можно

переставить столбцы – это лишь вызовет перестановку строк в матрице \mathbf{A}^+), то есть $\mathbf{A} = (\mathbf{B}, \mathbf{A}_1)$, где \mathbf{B} - подматрица размера $m \times r$ имеет ранг r (такая запись означает, что к матрице \mathbf{B} справа приписывается матрица \mathbf{A}_1).

В процессе построения QR - разложения методом ортогонализации, получается равенство $\mathbf{AU} = \mathbf{Q}$, причём первые r столбцов матрицы \mathbf{Q} ортонормированны. Другими словами, матрицу \mathbf{Q} можно «разбить» на две $\mathbf{Q} = (\mathbf{Q}_1, \mathbf{0})$, где подматрица \mathbf{Q}_1 состоит из r ортонормированных столбцов.

Оказывается, что матрицу \mathbf{U} можно выбрать так, чтобы матрица, составленная из последних $n - r$ столбцов матрицы, будет представима в виде

$$\begin{pmatrix} -\mathbf{U} \\ \mathbf{E}_{n-r} \end{pmatrix}$$

(\mathbf{E}_{n-r} - единичная матрица размера $n - r$).

В указанных обозначениях $\mathbf{A} = (\mathbf{B}, \mathbf{BU}) = \mathbf{B}(\mathbf{E}_r, \mathbf{U})$.

Таким образом, найдено скелетное разложение матрицы $\mathbf{A} = \mathbf{BC}$, где \mathbf{B} - подматрица матрицы \mathbf{A} , а матрица $\mathbf{C} = (\mathbf{E}_r, \mathbf{U})$.

При таком выборе скелетного разложения псевдообратные матрицы для матриц \mathbf{B} и \mathbf{C} (и, значит, и для матрицы \mathbf{A}) вычисляются очень просто.

Метод Гревилля

Рассмотрим этот метод на примере матрицы размера 2×4 .

$$\begin{pmatrix} 1 & 9 & 8 & 4 \\ 2 & 0 & 0 & 7 \end{pmatrix}.$$

Шаг 1. Нахождение псевдообратной матрицы для матрицы

$$\mathbf{A}_1 = \begin{pmatrix} 1 & 9 \\ 2 & 0 \end{pmatrix}.$$

В общем случае на этом шаге следует найти псевдообратную матрицу для матрицы, составленной из некоторого количества столбцов исходной матрицы (возможно даже рассмотрение только одного первого столбца).

В нашем случае разумно рассмотреть матрицу из первых двух столбцов – она обратима. Поэтому псевдообратная матрица равна обратной: $\mathbf{A}_1^+ = \begin{pmatrix} 0 & 1/2 \\ 1/9 & -1/18 \end{pmatrix}$.

Шаг 2. Припишем к матрице \mathbf{A}_1 ещё один столбец. Получим новую матрицу \mathbf{A}_2

$$\mathbf{A}_2 = \begin{pmatrix} 1 & 9 & 8 \\ 2 & 0 & 0 \end{pmatrix}.$$

Найдём псевдообратную матрицу для неё. Для этого выполним следующую последовательность действий:

1) Найдём \mathbf{A}_1^* . Так как все элементы матрицы – вещественные числа, то сопряжённая матрица совпадёт с транспонированной:

$$\mathbf{A}_1^* = \begin{pmatrix} 1 & 2 \\ 9 & 0 \end{pmatrix}.$$

2) Перемножим матрицы \mathbf{A}_1 и \mathbf{A}_1^* .

$$\mathbf{A}_1 \mathbf{A}_1^* = \begin{pmatrix} 82 & 2 \\ 2 & 4 \end{pmatrix}.$$

3) Ответим на вопрос: является ли вектор, который мы приписали к таблице \mathbf{A}_1

$$\boldsymbol{\alpha} = \begin{pmatrix} 8 \\ 0 \end{pmatrix},$$

собственным вектором матрицы $\mathbf{A}_1 \mathbf{A}_1^*$. В данном случае ответ отрицательный.

Остальные пункты укажем именно для такой ситуации.

4) Вычисление «направляющего вектора» отклонения от

«собственности» $\beta = \frac{(\mathbf{I} - \mathbf{A}_1 \mathbf{A}_1^*) \alpha}{\|(\mathbf{I} - \mathbf{A}_1 \mathbf{A}_1^*) \alpha\|}$ (\mathbf{I} - единичная матрица).

В нашем случае $\beta = \frac{1}{\sqrt{420004}} \begin{pmatrix} -648 \\ -10 \end{pmatrix} \approx \frac{-1}{648} \begin{pmatrix} 648 \\ 10 \end{pmatrix}$

5) Далее следует вычислить матрицу

$$\begin{aligned} \mathbf{A}_1^+ (\mathbf{I} - \alpha \beta^T) &= \begin{pmatrix} 1 & 2 \\ 9 & 0 \end{pmatrix} \left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{1}{648} \begin{pmatrix} 8 \\ 0 \end{pmatrix} \begin{pmatrix} 648 & 10 \end{pmatrix} \right) = \\ &= \begin{pmatrix} 1 & 2 \\ 9 & 0 \end{pmatrix} \begin{pmatrix} 9 & 10/81 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 9 & 172/81 \\ 81 & 10/9 \end{pmatrix} \end{aligned}$$

6) Для записи матрицы \mathbf{A}_2^+ всё готово. Нужно к матрице, полученной на предыдущем шаге, приписать снизу транспонированный столбец β .

Итак,

$$\mathbf{A}_2^+ = \begin{pmatrix} 9 & 172/81 \\ 81 & 10/9 \\ -1 & -5/342 \end{pmatrix}.$$

Шаг 3. Повторение шага 2. Приписываем следующий столбец матрицы.

В случае, если на третьем этапе ответ будет положительным, то этапы 4-6 будут следующими:

4) Вычисление вектора $\gamma = \frac{(\mathbf{A}^+)^T \mathbf{A}^+ \alpha}{1 + \|\mathbf{A}^+ \alpha\|^2}$.

5) Вычисление матрицы $\mathbf{A}^+ (\mathbf{I} - \alpha \gamma^T)$

6) Теперь следует к матрице, полученной на предыдущем шаге, приписать снизу транспонированный столбец γ .

Тема 2. Решение вырожденных и плохо обусловленных систем линейных алгебраических уравнений

1. Известно, с какими трудностями связано решение так называемых плохо обусловленных систем линейных алгебраических уравнений: малым изменениям правых частей таких систем могут отвечать большие (выходящие за допустимые пределы) изменения решения.

Рассмотрим систему уравнений

$$\mathbf{Az} = \mathbf{u}, \quad (2.1)$$

где \mathbf{A} – матрица с элементами a_{ij} , $\mathbf{A} = \{a_{ij}\}$, \mathbf{z} – искомый вектор с координатами z_j , $\mathbf{z} = \{z_j\}$, \mathbf{u} – известный вектор с координатами u_i , $\mathbf{u} = \{u_i\}$, $i, j = 1, 2, \dots, n$. Система (2.1) называется *вырожденной*, если определитель системы равен нулю, $\det \mathbf{A} = 0$. В этом случае матрица \mathbf{A} имеет равные нулю собственные значения. У плохо обусловленных систем такого вида матрица \mathbf{A} имеет близкие к нулю собственные значения.

Если вычисления производятся с конечной точностью, то в ряде случаев не представляется возможным установить, является ли заданная система уравнений вырожденной или плохо обусловленной. Таким образом, плохо обусловленные и вырожденные системы могут быть неразличимыми в рамках заданной точности. Очевидно, такая ситуация имеет место в случаях, когда матрица \mathbf{A} имеет достаточно близкие к нулю собственные значения.

В практических задачах часто правая часть \mathbf{u} и элементы матрицы \mathbf{A} , т.е. коэффициенты системы (2.1), известны приближенно. В этих случаях вместо системы (2.1) мы имеем дело с некоторой другой системой $\tilde{\mathbf{A}}\mathbf{z} = \tilde{\mathbf{u}}$ такой, что $\|\tilde{\mathbf{A}} - \mathbf{A}\| \leq h$, $\|\tilde{\mathbf{u}} - \mathbf{u}\| \leq \delta$, где смысл норм обычно определяется характером задачи. Имея вместо матрицы \mathbf{A} матрицу $\tilde{\mathbf{A}}$, мы тем более не можем высказать определенного суждения о вырожденности или невырожденности системы (2.1).

В этих случаях о точной системе $\mathbf{Az} = \mathbf{u}$, решение которой надо определить, нам известно лишь то, что для матрицы \mathbf{A} и правой части \mathbf{u} выполняются неравенства $\|\tilde{\mathbf{A}} - \mathbf{A}\| \leq h$ и $\|\tilde{\mathbf{u}} - \mathbf{u}\| \leq \delta$. Но систем с такими исходными данными (\mathbf{A}, \mathbf{u}) бесконечно много, и в рамках известного нам уровня погрешности они неразличимы. Поскольку вместо точной системы (2.1) мы имеем приближенную систему $\tilde{\mathbf{A}}\mathbf{z} = \tilde{\mathbf{u}}$, то речь может идти лишь о нахождении приближенного решения. Но приближенная система $\tilde{\mathbf{A}}\mathbf{z} = \tilde{\mathbf{u}}$ может быть неразрешимой. Возникает вопрос: что надо понимать под приближенным решением системы (2.1) в описанной ситуации?

Среди “возможных точных систем” могут быть и вырожденные. Если они разрешимы, то имеют бесконечно много решений. О приближенном нахождении какого из них должна идти речь?

Таким образом, в большом числе случаев мы должны рассматривать целый класс неразличимых между собой (в рамках заданного уровня погрешности) систем уравнений, среди которых могут быть и вырожденные, и неразрешимые. Методы построения приближенных решений систем этого класса должны быть одними и теми же, общими. Эти решения должны быть устойчивыми к малым изменениям исходных данных (2.1).

В основе построения таких методов лежит идея “отбора”, изложенная в главе при рассмотрении метода регуляризации.

2. Итак, рассмотрим произвольную систему линейных алгебраических уравнений (короче - СЛАУ)

$$\mathbf{Az} = \mathbf{u}, \quad (2.2)$$

в которой \mathbf{z} и \mathbf{u} – векторы, $\mathbf{z} = (z_1, z_2, \dots, z_n) \in \mathbf{R}^n$, $\mathbf{u} = (u_1, u_2, \dots, u_m) \in \mathbf{R}^m$, \mathbf{A} – матрица с элементами a_{ij} , $\mathbf{A} = \{a_{ij}\}$, где $j = 1, 2, \dots, n$; $i = 1, 2, \dots, m$, и число n не обязано быть равным числу m .

Эта система может быть однозначно разрешимой, вырожденной (и иметь бесконечно много решений) и неразрешимой.

Псевдорешением системы (2.2) называют вектор $\tilde{\mathbf{z}}$, минимизирующий невязку $\|\mathbf{Az} - \mathbf{u}\|$ на всем пространстве \mathbf{R}^n . Система (2.2) может иметь не одно псевдорешение. Пусть F_A есть совокупность всех ее псевдорешений и \mathbf{z}^1 - некоторый фиксированный вектор из \mathbf{R}^n , определяемый обычно постановкой задачи.

Нормальным относительно вектора \mathbf{z}^1 решением системы (2.2) будем называть псевдорешение \mathbf{z}^0 с минимальной нормой $\|\mathbf{z} - \mathbf{z}^1\|$, т.е. такое, что $\|\mathbf{z}^0 - \mathbf{z}^1\| = \inf_{\mathbf{z} \in F_A} \|\mathbf{z} - \mathbf{z}^1\|$.

Здесь $\|\mathbf{z}\| = \left\{ \sum_{j=1}^n z_j^2 \right\}^{1/2}$. В дальнейшем для простоты записи будем

считать $\mathbf{z}^1 = \mathbf{0}$ и нормальное относительно вектора $\mathbf{z}^1 = \mathbf{0}$ решение называть просто *нормальным решением*.

Нетрудно показать, что для любой системы вида (2.2) нормальное решение существует и единственно.

3. Нетрудно видеть, что задача нахождения нормального решения системы (2.2) является некорректно поставленной. В самом деле, пусть \mathbf{A} симметричная матрица. Если она невырожденная, то ортогональным преобразованием

$$\mathbf{z} = \mathbf{Vz}^*, \quad \mathbf{u} = \mathbf{Vu}^*$$

её можно привести к диагональному виду и преобразованная система будет иметь вид

$$\lambda_i z_i^* = u_i^*, \quad i = 1, 2, \dots, n,$$

где λ_i – собственные значения матрицы \mathbf{A} .

Если симметричная матрица \mathbf{A} – вырожденная и имеет ранг r , то $n - r$ её собственных значений равны нулю. Пусть

$\lambda_i \neq 0$ для $i = 1, 2, \dots, r$

и

$\lambda_i = 0$ для $i = r + 1, r + 2, \dots, n$.

Полагаем, что система (2.2) разрешима. При этом $u_i^* = 0$ для $i = r + 1, \dots, n$.

Пусть “исходные данные” системы (\mathbf{A} и \mathbf{u}) заданы с погрешностью, т.е. вместо \mathbf{A} и \mathbf{u} заданы их приближения $\tilde{\mathbf{A}}$ и $\tilde{\mathbf{u}}$: $\|\tilde{\mathbf{A}} - \mathbf{A}\| \leq h$, $\|\tilde{\mathbf{u}} - \mathbf{u}\| \leq \delta$.

При этом

$$\|\mathbf{A}\| = \left\{ \sum_{i,j} a_{ij}^2 \right\}^{1/2}, \quad \|\mathbf{u}\| = \left\{ \sum_i u_i^2 \right\}^{1/2}. \quad (2.3)$$

Пусть $\tilde{\lambda}_i$ - собственные значения матрицы $\tilde{\mathbf{A}}$. Известно, что они непрерывно зависят от \mathbf{A} в норме (2.3). Следовательно, собственные значения $\tilde{\lambda}_{r+1}, \tilde{\lambda}_{r+2}, \dots, \tilde{\lambda}_n$ могут быть сколь угодно малыми при достаточно малом h .

Если они не равны нулю, то $\tilde{z}_i^* = \frac{1}{\tilde{\lambda}_i} \tilde{u}_i^*$.

Таким образом, найдутся возмущения системы в пределах любой достаточно малой погрешности $\tilde{\mathbf{A}}$ и $\tilde{\mathbf{u}}$, для которых некоторые \tilde{z}_i^* будут принимать любые наперед заданные значения. Это означает, что задача нахождения нормального решения системы (2.2) является неустойчивой.

2.1. Корректность и некорректность по Адамару

Определения корректности и некорректности. Ж. Адамар в 1902 г. ввел понятия *корректности* и *некорректности* [3-5]. Рассмотрим операторное уравнение:

$$Ay = f, \quad y \in Y, \quad f \in F, \quad (2.4)$$

где y — искомое решение, f — заданная правая часть, Y и F — некоторые гильбертовы пространства (например, W_2^1 и L_2), A — заданный непрерывный оператор (линейный и нелинейный, интегральный, дифференциальный или алгебраический и т.д.).

Определение. Задача решения уравнения (2.4) называется *корректной*, или *корректно поставленной (well-posed)*, если:

1. решение существует,
2. решение единственно,
3. решение устойчиво.

Если хотя бы одно из этих условий не выполняется, то задача называется *некорректной*, или *некорректно поставленной (ill-posed)*.

Более того, Адамар выдвинул утверждение, что некорректные задачи не имеют физического смысла. Другими словами, если уравнение, описывающее некоторую прикладную (физическую, техническую и т.д.) задачу, является некорректным, то или эта задача является искусственной (нереальной), или она описана математически неадекватно. Например, описана интегральным уравнением Фредгольма I рода, задача решения которого некорректна, а нужно бы добавить еще ряд ограничений на решение и тогда задача станет корректной. А поскольку, как выяснилось в последние десятилетия, значительная часть прикладных задач является некорректными, то утверждение Адамара привело к замедлению развития многих разделов чистой и прикладной математики.

Далее будет показано, что если решение не существует, то используют метод типа метода наименьших квадратов Гаусса (получают псевдорешение), если решение неединственно, то используют метод типа метода псевдообратной матрицы Мура-Пенроуза (получают нормальное решение), а если решение неустойчиво, то используют устойчивые (регулярные) методы (регуляризации, фильтрации и др.).

Но прежде чем перейти к этим методам, мы остановимся на классических методах (имея в виду, что они обычно дают решения, некорректные по Адамару).

2.2. Классические методы решения интегральных уравнений Фредгольма I рода

Метод квадратур. Рассмотрим интегральное уравнение Фредгольма I рода:

$$\int_a^b K(x,s)y(s)ds = f(x), c \leq x \leq d, \quad (2.5)$$

где $K(x,s)$ — ядро, $y(s)$ — искомая функция, $f(x)$ — правая часть (имеется в виду зашумленная правая часть $f(x)$), $[a,b]$ — область изменения s , а $[c,d]$ — область изменения x . Метод квадратур заключается в следующем:

Область $[a,b]$ разбиваем через шаг $\Delta s = h_1 = const$, а область $[c,d]$ через шаг $\Delta s = h_2 = const$ (рассмотрим случай постоянства шагов дискретизации h_1 и h_2 , хотя метод можно обобщить и на случай непостоянства h_1 и h_2). Получим число узлов $n = (b-a)/h_1 + 1$ (по s) и $m = (d-c)/h_2 + 1$ (по x).

Интеграл в (2.5) заменяем конечной суммой, расписывая его по некоторой квадратурной формуле, например, по формуле трапеций:

$$\int_a^b K(x,s)y(s)ds \approx \sum_{j=1}^n p_j K(x,s_j)y(s_j), \quad (2.6)$$

где

$$p_j = \begin{cases} 0.5h_1, & j = 1 \text{ или } j=n, \\ h_1, & \text{иначе,} \end{cases} \quad (2.7)$$

$$s_j = a + (j-1)h_1. \quad (2.8)$$

3) Вводя дискретизацию по x :

$$x_i = c + (j-1)h_2, \quad (2.9)$$

окончательно получим:

$$\sum_{j=1}^n A_{ij} y_j = f_i, \quad i = \overline{1, m}, \quad (2.10)$$

где $A_{ij} = p_j K(x_i, s_j)$ — элементы матрицы \mathbf{A} размера $m \times n$, $y_j = y(s_j)$, $f_i = f(x_i)$.

Итак, получили систему m линейных алгебраических уравнений (2.10) относительно n неизвестных y_j . Решая ее, можно получить решение интегрального уравнения (2.5) в дискретном виде.

Матрица A системы уравнений (2.10), вообще говоря, прямоугольна. Если $m = n$, то матрица \mathbf{A} — квадратная и СЛАУ (2.10) можно решать по правилу Крамера, гауссовскими методами и др. Если $m > n$, то СЛАУ (2.10) нужно решать методом наименьших квадратов Гаусса — получим псевдорешение, а если $m < n$, то нужно использовать метод псевдообратной матрицы Мура-Пенроуза — получим нормальное решение. Таким образом, первые два пункта корректности по Адамару будут выполнены.

Однако все эти решения очень неустойчивы, т. е. нарушается 3-й пункт корректности по Адамару. Эта неустойчивость обусловлена тем, что минимальное сингулярное число μ_{\min} интегрального оператора уравнения (2.5) равно нулю и число обусловности $cond = \infty$. Если же n и m конечны, то μ_{\min} может стать несколько отличным от нуля, но решение СЛАУ (2.10) будет по-прежнему очень неустойчивым.

Из изложенного можно сделать следующие выводы:

1) метод квадратур решения интегрального уравнения Фредгольма I рода (а также в меньшей степени уравнения Вольтерры I рода) является крайне неустойчивым (нарушается 3-й пункт корректности по Адамару);

2) классическое определение точного решения \bar{y} как решения, при котором

$$\|A\bar{y} - \mathbf{f}\| = 0, \quad (2.11)$$

вообще говоря, не подходит для случая некорректных задач, так как в случае, когда решение не существует, отсутствует такое \bar{y} , для которого выполняется (2.11), а в случае неединственности решения существует множество \bar{y} , для которых выполняется (2.11), и наконец, в случае неустойчивости критерий (2.11) дает неустойчивое решение.

Метод преобразования Фурье. Рассмотрим интегральное уравнение Фредгольма I рода типа свертки

$$\int_{-\infty}^{\infty} K(x-s)y(s)ds = f(x), \quad -\infty < x < \infty. \quad (2.12)$$

Запишем $y(s)$ в виде обратного преобразования Фурье (приложение 2):

$$y(s) = \frac{1}{2\pi} \int_{-\infty}^{\infty} Y(w')e^{-iw's} dw'. \quad (2.13)$$

Умножим (2.12) на $e^{iw's}$ и проинтегрируем по x от $-\infty$ до ∞ . Получим (заменяя в левой части (2.12) x на x'):

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K(x'-s)Y(w')e^{i(wx'-w's)} dw' ds dx' = F(w), \quad (2.14)$$

где

$$F(w) = \int_{-\infty}^{\infty} f(x)e^{iwx} dx. \quad (2.15)$$

Сделаем замену переменной: $x' - s = x$ в (2.14), получим, изменив порядок интегрирования:

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} K(x)e^{iwx} dx \cdot \int_{-\infty}^{\infty} Y(w') \left[\int_{-\infty}^{\infty} e^{i(w-w')s} ds \right] dw' =$$

$$\begin{aligned}
&= \int_{-\infty}^{\infty} K(x)e^{iwx} dx \cdot \int_{-\infty}^{\infty} Y(w')\delta(w-w')ds = \\
&= \int_{-\infty}^{\infty} K(x)e^{iwx} dx \cdot Y(w) = F(w).
\end{aligned}$$

Обозначив

$$\lambda(w) = \int_{-\infty}^{\infty} K(x)e^{iwx} dx, \quad (2.16)$$

получим

$$\lambda(w)Y(w) = F(w), \quad (2.17)$$

т.е. для уравнения (2.12) справедливо **утверждение**: произведение преобразований Фурье ядра и искомой функции равно преобразованию Фурье правой части. Из (2.17) имеем окончательно:

$$Y(w) = \frac{F(w)}{\lambda(w)}. \quad (2.18)$$

Итак, уравнение (2.12) имеет аналитическое решение и оно записывается в виде обратного преобразования Фурье:

$$y(s) = \frac{1}{2\pi} \int_{-\infty}^{\infty} Y(w)e^{-iwx} dw, \quad (2.19)$$

где преобразования Фурье $Y(w)$, $F(w)$ и $\lambda(w)$ от искомой функции, правой части и ядра записываются в виде (2.18), (2.15) и (2.16).

Формулу (2.19) можно записать в ином виде (изменив порядок интегрирования):

$$\begin{aligned}
y(s) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f(x)e^{-iwx} dx \right] \frac{e^{-iwx}}{\lambda(w)} dw = \\
&= \int_{-\infty}^{\infty} \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-iw(s-x)}}{\lambda(w)} dw \right] f(x) dx
\end{aligned}$$

ИЛИ

$$y(s) = \int_{-\infty}^{\infty} R(s-x)f(x)dx, \quad (2.20)$$

где

$$R(s) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-iws}}{\lambda(w)} dw. \quad (2.21)$$

Формула (2.20) удобна тем, что функция $R(s)$ может быть вычислена заранее и единожды, а затем для ряда функций $f(x)$ могут быть весьма быстро найдены соответствующие решения $y(s)$.

Однако решение (2.19) (или (2.20)), как и решение по методу квадратур, также крайне неустойчиво, и это можно объяснить следующим образом. Ядро $K(x)$ обычно задается в виде гладкой функции, поэтому его спектр $\lambda(w)$ быстро убывает с ростом $|w|$ и $\lim_{w \rightarrow \infty} \lambda(w) = 0$. Функция же $f(x)$ обычно задается в виде таблицы зашумленных значений, т. е. вместо $f(x)$ имеем $\tilde{f}(x) = f(x) + \delta f(x)$, где $\delta f(x)$ — погрешности, обладающие тем свойством, что их спектр при $|w| \rightarrow \infty$ обычно стремится к некоторой (пусть даже очень малой) константе — уровню «белого шума». Поэтому $\lim_{w \rightarrow \infty} F(w) / \lambda(w) = \lim_{w \rightarrow \infty} Y(w) = \infty$ и интеграл (2.19) расходится. Другими словами, неустойчивость метода ПФ обусловлена очень сильной реакцией высоких гармоник в преобразовании Фурье на даже очень малые погрешности измерений $f(x)$. Если же вычисления выполняются по конечным квадратурным формулам, т. е. вместо непрерывного преобразования Фурье $\lambda(w)$, $Y(w)$ и $F(w)$ используются их дискретное преобразование Фурье с конечным числом отсчетов N , то эта неустойчивость несколько уменьшается (становится конечной), но тем не менее остается большой.

Заметим, что как показывает решение большого числа примеров [3, 4], метод преобразования Фурье дает менее неустойчивое решение, чем метод

квадратур, и это обусловлено тем, что, во-первых, метод ПФ дает аналитическое решение (2.19) или (2.20) и, во-вторых, при численной реализации метода преобразования Фурье автоматически срабатывает усечение спектра частот.

Метод преобразования Фурье для двумерного уравнения. Для двумерного интегрального уравнения Фредгольма I рода типа свертки

$$\int \int_{-\infty}^{\infty} K(x_1 - s_1, x_2 - s_2) y(s_1, s_2) ds_1 ds_2 = f(s_1, s_2) \quad (2.22)$$

метод двумерного преобразования Фурье дает следующее решение:

$$y(s_1, s_2) = \frac{1}{4\pi^2} \int \int_{-\infty}^{\infty} Y(w_1, w_2) e^{-i(w_1 s_1 + w_2 s_2)} dw_1 dw_2, \quad (2.23)$$

где

$$Y(w_1, w_2) = \frac{F(w_1, w_2)}{\lambda(w_1, w_2)}, \quad (2.24)$$

$$F(w_1, w_2) = \int \int_{-\infty}^{\infty} f(x_1, x_2) e^{i(w_1 x_1 + w_2 x_2)} dx_1 dx_2, \quad (2.25)$$

$$\lambda(w_1, w_2) = \int \int_{-\infty}^{\infty} K(x_1, x_2) e^{i(w_1 x_1 + w_2 x_2)} dx_1 dx_2. \quad (2.26)$$

В заключение добавим, что помимо метода квадратур и метода преобразования Фурье для решения уравнений Фредгольма I рода развиты: метод разложения по собственным функциям [4, 5], метод итераций (последовательных приближений) [6] и др. Однако и данные методы дают устойчивые решения лишь при использовании регуляризирующих модификаций, специальных стабилизирующих приемов и т. д.

2.3. Метод наименьших квадратов Гаусса

Изложим на примере решения СЛАУ метод наименьших квадратов (МНК) Гаусса.

Переопределенная СЛАУ. Рассмотрим систему m линейных алгебраических уравнений (СЛАУ) относительно n неизвестных, причем $m > n$ и $\text{rang}(\mathbf{A}|\mathbf{f}) > \text{rang}(\mathbf{A})$, т.е. переопределенную СЛАУ:

$$\mathbf{A}\mathbf{y} = \mathbf{f}, \quad (2.27)$$

где \mathbf{A} — матрица $m \times n$, \mathbf{y} — искомый вектор-столбец $n \times 1$, \mathbf{f} — заданная правая часть — вектор-столбец $m \times 1$. Такая СЛАУ не имеет решения, другими словами, нет такого \mathbf{y} , для которого справедливо

$$\|\mathbf{A}\bar{\mathbf{y}} - \mathbf{f}\| = 0, \quad (2.28)$$

т. е. невязка равна нулю. В МНК Гаусса вместо (2.28) вводится условие

$$\|\mathbf{A}\mathbf{y} - \mathbf{f}\| = \min_{\mathbf{y}}. \quad (2.29)$$

Определение. *Псевдорешением* СЛАУ (2.27) называется решение \mathbf{y} , удовлетворяющее условию (2.29), т.е. минимизирующее невязку $\|\mathbf{A}\mathbf{y} - \mathbf{f}\|$.

Таким образом, в МНК условие равенства нулю невязки заменяется на условие ее минимума, а вместо точного решения $\bar{\mathbf{y}}$ рассматривается псевдорешение \mathbf{y} . Заметим, что если $\|\mathbf{A}\bar{\mathbf{y}} - \mathbf{f}\| = 0$, то псевдорешение \mathbf{y} совпадает с точным решением $\bar{\mathbf{y}}$, т. е. псевдорешение обобщает понятие точного решения.

Вывод нормальной СЛАУ. Запишем условие (2.29) в виде:

$$\|\mathbf{A}\mathbf{y} - \mathbf{f}\| = \min_{\mathbf{y}}, \quad (2.30)$$

а нормы будем определять согласно (8.42). Выведем (нестрогое, но наглядно) новую СЛАУ из условия (2.30). Минимизация (2.30) означает равенство нулю вариации (или производной) по \mathbf{y} :

$$2(\mathbf{A}\mathbf{y} - \mathbf{f})\mathbf{A} = 0, \quad (2.31)$$

или, учитывая правила умножения матриц и векторов (8.45), (8.47),

$$\mathbf{A}^*(\mathbf{A}\mathbf{y} - \mathbf{f}) = 0. \quad (2.32)$$

В результате

$$\mathbf{A}^* \mathbf{A} \mathbf{y} = \mathbf{A}^* \mathbf{f}. \quad (2.33)$$

Итак, место переопределенной СЛАУ (2.27) получена новая СЛАУ (2.33), называемая *нормальной* СЛАУ. Запишем (2.33) в виде:

$$\mathbf{B} \mathbf{y} = \mathbf{u}, \quad (2.34)$$

где

$$\mathbf{B} = \mathbf{A}^* \mathbf{A}, \quad (2.35)$$

$$\mathbf{u} = \mathbf{A}^* \mathbf{f}, \quad (2.36)$$

или в случае вещественности \mathbf{A}

$$\mathbf{B} = \mathbf{A}^T \mathbf{A}, \quad (2.37)$$

$$\mathbf{u} = \mathbf{A}^T \mathbf{f}. \quad (2.38)$$

Используя правила (5.46) и (5.48), запишем подробно выражения для элементов новой матрицы \mathbf{B} и новой правой части \mathbf{u} (в случае вещественности \mathbf{A}):

$$B_{ij} = \sum_{k=1}^m A_{ik}^T A_{kj} = \sum_{k=1}^m A_{ki} A_{kj}, \quad (2.39)$$

$$u_i = \sum_{k=1}^m A_{ik}^T f_k = \sum_{k=1}^m A_{ki} f_k. \quad (2.40)$$

СЛАУ (2.33) или (2.34) можно решать по формуле

$$\mathbf{y} = (\mathbf{A}^* \mathbf{A})^{-1} \mathbf{A}^* \mathbf{f} \quad (2.41)$$

или

$$\mathbf{y} = \mathbf{B}^{-1} \mathbf{u}, \quad (2.42)$$

или же методами Гаусса, Холецкого и др.

МНК применительно к интегральному уравнению. Если применить МНК Гаусса к интегральному уравнению (2.5), то получим следующее новое интегральное уравнение (ср. (2.34)):

$$\int_a^b B(t,s) y(s) ds = u(t), \quad a \leq t \leq b, \quad (2.43)$$

где (сравни с (2.39), (2.40))

$$B(t,s) = B(s,t) = \int_c^d K(x,t)K(x,s)dx, \quad (2.44)$$

$$u(t) = \int_c^d K(x,t)f(x)dx. \quad (2.45)$$

Основные особенности МНК Гаусса:

Матрица \mathbf{B} — квадратная $n \times n$, т.е. решается система n уравнений относительно n неизвестных и в случае $\det(\mathbf{B}) \neq 0$ решение СЛАУ (2.34) существует и является единственным.

Матрица \mathbf{B} и новое ядро $\mathbf{B}(t,s)$ являются симметричными и положительно определенными.

Решения уравнений (2.34) и (2.43) неустойчивы.

2.4. Метод псевдообратной матрицы Мура-Пенроуза

Изложим метод псевдообратной матрицы Мура-Пенроуза сначала на примере решения СЛАУ.

Недоопределенная СЛАУ. Рассмотрим СЛАУ

$$\mathbf{A}\mathbf{y} = \mathbf{f}, \quad (2.46)$$

где \mathbf{A} — $m \times n$ -матрица, \mathbf{y} — искомый n -вектор, \mathbf{f} — заданная правая часть — m -вектор, причем $m < n$. Такая недоопределенная СЛАУ имеет множество решений $\bar{\mathbf{y}}$, т.е. нарушается 2-й пункт корректности по Адамару. Для всех них выполняется равенство (2.28).

Нормальное решение и псевдообратная матрица.

Недоопределенная СЛАУ может быть решена *методом псевдообратной матрицы* Мура-Пенроуза (1930г.) [4].

Определение. *Нормальное решение* — это решение с минимальной нормой среди множества решений, т. е. удовлетворяющее условию

$$\|\mathbf{y}\| = \min_{\mathbf{y}} \quad (2.47)$$

или

$$\|y\|^2 = \min_y. \quad (2.48)$$

Нормальное решение — это наиболее гладкое из решений.

Согласно методу псевдообратной матрицы, среди множества решений недоопределенной СЛАУ выбирается нормальное решение. Доказывается [7], что нормальное решение существует и является единственным и находится по формуле:

$$y = A^+ f, \quad (2.49)$$

где A^+ — псевдообратная $n \times m$ -матрица Мура-Пенроуза. Матрица A^+ определяется соотношением:

$$AA^+A = A \quad (2.50)$$

или (теоретическая асимптотическая формула)

$$A^+ = \lim_{\alpha \rightarrow 0} (\alpha E + A^* A)^{-1} A^*. \quad (2.51)$$

Однако соотношения (2.50) и (2.51) неудобны для практического нахождения A^+ . Практически удобный алгоритм отыскания A^+ приведен в [7].

В случае квадратной невырожденной матрицы A имеем: $A^+ = A^{-1}$ (ср. (8.37)), а в случае переопределенной СЛАУ $A^+ = (A^* A)^{-1} A^*$ (ср. (2.41)), т.е. запись (2.49) является общей для недоопределенной, определенной и переопределенной СЛАУ. Кроме того, решение (2.49), которое уместно записать в виде $y^+ = A^+ f$, дает нулевую невязку $\|Ay^+ - f\| = 0$, т.е. оно является псевдорешением (ср. (2.29)) и среди всех псевдорешений (которых в случае недоопределенной СЛАУ множество) имеет, как нормальное решение, минимальную норму. Другими словами, нормальное решение является и псевдорешением. Однако заметим, что пользоваться формулой (2.41) для отыскания любого решения, в том числе и нормального, при $m < n$ нельзя, так как в этом случае $A^* A$ матрица является вырожденной.

Метод псевдообратной матрицы применительно к другим уравнениям. Если под записью (2.46) подразумевать иные уравнения, в частности, интегральное уравнение Фредгольма I рода (2.5) или операторное уравнение (8.34), то изложенный метод также применим, а именно, в качестве решения интегрального или операторного уравнения выбирается нормальное решение, удовлетворяющее условию (2.47) и найденное по формуле (2.49), где A^+ — *псевдообратный оператор*, определяемый посредством (2.51).

Общий вывод. По материалам двух последних пунктов можно сделать следующий вывод: если в качестве решения брать так называемое *нормальное псевдорешение* (нормальное или псевдорешение), то будут выполнены два первых пункта корректности по Адамару. Однако МНК и МПОМ не решают проблему неустойчивости решения, т. е. третий пункт корректности по Адамару для нормального псевдорешения, вообще говоря, не выполняется.

Тема 3. Метод регуляризации нахождения нормального решения

3.1. Метод регуляризации Тихонова

Метод регуляризации Тихонова [5] является дальнейшим развитием метода наименьших квадратов Гаусса (дающего псевдорешение) и метода псевдообратной матрицы Мура-Пенроуза (дающего нормальное решение).

Существо метода. Сначала рассмотрим метод применительно к операторному уравнению:

$$Ay = f, \quad y \in L_2, \quad f \in L_2, \quad (3.1)$$

где A — линейный вполне непрерывный оператор, f — заданная правая часть, а y — искомое решение, причем вместо точных f и A известны их приближения \tilde{f} и \tilde{A} такие, что

$$\|\tilde{f} - f\|_{L_2} \leq \delta, \quad (3.2)$$

$$\|\tilde{A} - A\| \leq \mu, \quad (3.3)$$

где $\delta \geq 0$ и $\mu \geq 0$ — погрешности правой части и оператора (точнее, их верхние оценки, поэтому в (3.2) и (3.3) стоят знаки \leq , а не знаки $=$), т.е. решается уравнение

$$\tilde{A}\tilde{y} = \tilde{f}, \quad \tilde{y} \in L_2, \quad \tilde{f} \in L_2. \quad (3.4)$$

Однако для упрощения записи мы далее будем использовать запись (3.1), подразумевая, что в действительности рассматривается уравнение (3.4).

В методе регуляризации Тихонова ставятся два условия: условие минимизации невязки типа (2.30), как в методе Гаусса, и условие минимизации нормы решения типа (2.48), как в методе Мура-Пенроуза. Это — задача условной минимизации и она решается методом неопределенных множителей Лагранжа, а именно,

$$\|Ay - f\|_{L_2}^2 + \alpha \|y\|_{L_2}^2 = \min_y, \quad (3.5)$$

где $\alpha > 0$ — параметр регуляризации, играющий роль неопределенного множителя Лагранжа. Из условия (3.5) вытекает уравнение Тихонова (ср. (2.33)):

$$(\alpha E + A^* A)y_\alpha = A^* f, \quad (3.6)$$

где E — единичный оператор ($Ey = y$). Итак, вместо уравнения I рода получено уравнение II рода (3.6).

Анализ метода. Проанализируем условие (3.5) и уравнение (3.6).

Если $\alpha = 0$, то метод регуляризации Тихонова переходит в МНК Гаусса с крайне неустойчивым решением, но минимальной невязкой $\|Ay - f\|^2$. С увеличением же α решение становится глаже и устойчивей, т.е. уменьшается норма решения $\|y_\alpha\|^2$, но увеличивается невязка. Истина — посередине, т. е. при некотором умеренном α решение y_α будем иметь и умеренную гладкость, и умеренную невязку. Некоторые способы выбора α изложены ниже.

Если $\delta, \mu \rightarrow 0$, то $\alpha \rightarrow 0$ и

$$y_\alpha = \lim_{\alpha \rightarrow 0} (\alpha E + A^* A)^{-1} A^* f \equiv A^+ f \quad (3.7)$$

(см. (2.51)), т.е. решение y_α переходит в нормальное псевдорешение. Таким образом, метод регуляризации Тихонова является обобщением метода наименьших квадратов Гаусса и метода псевдообратного оператора Мура-Пенроуза.

Метод регуляризации Тихонова устойчив, т. е. выполняется 3-й пункт корректности по Адамару и эта устойчивость обусловлена следующими обстоятельствами. Оператор $A^* A$ в (3.6) является положительно определенным, поэтому все его собственные значения вещественны и неотрицательны: $\lambda_i(A^* A) \geq 0$, причем $\lambda(A^* A)_{\min} = 0$. Наличие же слагаемого αE в (3.6) увеличивает все $\lambda_i(A^* A)$ на α , поэтому $\alpha E + \lambda(A^* A)_{\min} = \alpha$. Вследствие этого, оператор $\alpha E + A^* A$ становится

обратимым, норма обратного оператора $\|(\alpha E + A^* A)^{-1}\| = 1/\alpha \neq \infty$ и задача становится устойчивой.

Решение уравнения (3.6) есть

$$y_\alpha = (\alpha E + A^* A)^{-1} A^* f. \quad (3.8)$$

В методе регуляризации Тихонова рассматривается также более общий подход, когда [19, с. 238]

$$\|Ay - f\|^2 + \alpha \|y - \psi\|^2 = \min_y, \quad (3.9)$$

где ψ — начальное приближение (матожидание, прогноз) решения y . В этом случае решение равно

$$y_\alpha = \psi + (\alpha E + A^* A)^{-1} A^* (f - A\psi). \quad (3.10)$$

Однако соотношения (3.9), (3.10) используются редко из-за неопределенности ψ . Мы их приводим лишь для сопоставления метода регуляризации Тихонова и метода оптимальной фильтрации Калмана (см. п. 3.2).

Регуляризованное интегральное уравнение. Применительно к интегральному уравнению Фредгольма I рода

$$Ay \equiv \int_a^b K(x,s)y(s)ds = f(x), \quad c \leq x \leq d, \quad (3.11)$$

соотношение (3.6) приобретает вид интегрального уравнения Фредгольма II рода с положительно определенным ядром (ср. (2.43)-(2.45)) [4]:

$$\alpha y_\alpha(t) + \int_a^b R(t,s)y_\alpha(s)ds = F(t), \quad a \leq t \leq b, \quad (3.12)$$

где

$$R(t,s) = R(s,t) = \int_c^d K(x,t)K(x,s)dx, \quad (3.13)$$

$$F(t) = \int_c^d K(x,t)f(x)dx. \quad (3.14)$$

Способы выбора параметра регуляризации α . Разработан ряд способов выбора α в методе регуляризации Тихонова [4]. Приведем три из них.

1-й способ — *способ невязки*, согласно которому α выбирается из условия [5] при $\mu = 0$:

$$\|Ay_\alpha - f\|_{L_2} = \delta. \quad (3.15)$$

Если $\|f\|_{L_2} \geq \delta$, то решение уравнения (3.15) относительно α существует и является единственным. При $\mu \neq 0$ способ невязки переходит в *обобщенный принцип невязки* [5].

2-й способ — *способ подбора*. Согласно нему, находятся решения y_α для ряда «разумных» значений α и окончательный выбор α делается на основе дополнительной информации о решении, в основном, визуально. Несмотря на простоту способа, он может быть весьма эффективен в случае, когда имеется немалая информация о решении (степень гладкости, число экстремумов и т.д.), а также выполнена обработка предшествующих «близких» примеров, позволяющая выделить область возможных значений α .

3-й способ — *асимптотический способ*, основанный на следующей зависимости при $\mu = 0$ и $\delta \rightarrow \infty$ [3]:

$$\alpha = C\delta^2, \quad (3.16)$$

где $C > 0$ — некоторая константа. Этот способ можно использовать при малых значениях δ .

Разработаны также другие способы выбора α : способ квазиоптимального (квазинаилучшего) α , способ отношения, способ независимых реализаций, способ перекрестной значимости, способ моделирования и т.д. [5, 6].

Численный алгоритм. Рассмотрим вопрос о численном решении интегрального уравнения (3.12). Остановимся на одном из наиболее эффективных алгоритмов — *методе квадратур*.

Пусть правая часть $f(x)$ задана таблично на следующей, вообще говоря, неравномерной x -сетке узлов:

$$c = x_1 < x_2 < x_3 < \dots < x_l = d, \quad (3.17)$$

а решение $y_\alpha(s)$ ищется на другой неравномерной s -сетке узлов, совпадающей с t -сеткой узлов:

$$a = s_1 = t_1 < s_2 = t_2 < s_3 = t_3 < \dots < s_n = t_n = b, \quad (3.18)$$

причем $l \neq n$. Распишем интеграл в (3.12) по некоторой квадратурной формуле, лучше всего по формуле трапеций. Получим:

$$\alpha y_k + \sum_{j=1}^n r_j R_{kj} y_j = F_k, \quad k = \overline{1, n}, \quad (3.19)$$

где $y_k = y_\alpha(t_k)$, $y_j = y_\alpha(s_j)$, $R_{kj} = R(t_k, s_j)$, $F_k = F(t_k)$. Аналогично интегралы в (3.13) и (3.14) аппроксимируем конечными суммами по квадратурной формуле. Получим:

$$R_{kj} = R_{jk} = \sum_{i=1}^l p_i K_{ik} K_{ij}, \quad k, j = \overline{1, n}, \quad (3.20)$$

$$F_k = \sum_{i=1}^l p_i K_{ik} f_i, \quad k = \overline{1, n}, \quad (3.21)$$

где $K_{ik} = K(x_i, t_k)$, $K_{ij} = K(x_i, s_j)$, $f_i = f(x_i)$, а r_j и p_i — коэффициенты квадратурных формул.

Запись (3.19) есть СЛАУ относительно y_j , $j = \overline{1, n}$.

Метод регуляризации для уравнения типа свертки. Рассмотрим частный случай интегрального уравнения Фредгольма I рода — уравнение типа свертки одномерное (8.3) и двумерное (8.5). Если уравнение общего вида (3.11) при его численном решении методом квадратур требует размещения в компьютерной памяти матрицы СЛАУ (см. (3.19)) и это

ограничивает возможности метода, то для решения одномерного уравнения типа свертки возможно применение метода преобразования Фурье, оперирующего лишь с векторами, что существенно расширяет возможности метода в отношении памяти и времени решения. Сказанное в еще большей степени характерно для двухмерного уравнения.

Рассмотрим *одномерное интегральное уравнение Фредгольма I рода типа свертки*:

$$Ay \equiv \int_{-\infty}^{\infty} K(x-s)y(s)ds = f(x), \quad -\infty < x < \infty. \quad (3.22)$$

Применительно к нему в методе регуляризации Тихонова решение находится из условия *минимума сглаживающего функционала* (ср. (3.5)):

$$\int_{-\infty}^{\infty} [Ay - f(x)]^2 dx + \alpha \int_{-\infty}^{\infty} M(w)|Y(w)|^2 = \min_y, \quad (3.23)$$

где

$$M(w) = |w|^{2q} \quad (3.24)$$

— регуляризатор q -го порядка, причем $q \geq 0$ — задаваемый порядок регуляризации, например, $q = 1$.

Из условия (3.23) получается следующее выражение для регуляризованного решения (ср. (2.19)):

$$y_\alpha(s) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\lambda(-w)F(w)}{L(w) + \alpha M(w)} e^{-iws} dw, \quad (3.25)$$

где

$$L(w) = |\lambda(w)|^2 = \lambda(w)\lambda(-w) = \text{Re}^2 \lambda(w) + \text{Im}^2 \lambda(w) \quad (3.26)$$

(см. также (2.15), (2.16)) или (ср. (2.20))

$$y_\alpha(s) = \int_{-\infty}^{\infty} R_\alpha(s-x)f(x)dx, \quad (3.27)$$

где (ср. (2.21))

$$R_\alpha(s) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\lambda(-w)}{L(w) + \alpha M(w)} e^{-iws} dw. \quad (3.28)$$

Сравним классическое решение (2.19) и регуляризованное решение (3.25). В (3.25) за счет слагаемого $\alpha M(w)$ подынтегральная функция стремится к нулю при $|w| \rightarrow \infty$, т.е. слагаемое $\alpha M(w)$ подавляет реакцию высоких гармоник на погрешность исходных данных, причем подавление тем сильнее, чем больше значения α и q . При этом, чем больше q , тем сильнее подавляются высокие гармоники в решении по сравнению с низкими, параметр же α определяет глобальное подавление: с его увеличением сильнее подавляются все гармоники. Поэтому в отношении q следует руководствоваться следующим правилом: если искомое решение имеет флуктуации, то q следует взять поменьше (например, $q = 1$), а если решение гладкое, то q можно повысить, например, $q = 2$. Что же касается a , то способы его выбора те же, что и для уравнения (3.11) (способы невязки, подбора и др.).

Разработан ряд *численных алгоритмов* получения решения $y_\alpha(s)$ [3]. Все они основаны на замене интегралов в (3.25), (2.15), (2.16), (3.27), (3.28) конечными суммами (по формулам прямоугольников или трапеций), переходе, тем самым, от непрерывного преобразования Фурье к дискретному преобразованию Фурье и использовании алгоритма быстрого преобразования Фурье.

Рассмотрим кратко *двухмерное интегральное уравнение Фредгольма I рода типа свертки*:

$$Ay \equiv \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K(x_1 - s_1, x_2 - s_2) y(s_1, s_2) ds_1 ds_2 = f(x_1, x_2), \quad (3.29)$$

$$-\infty < x < \infty$$

Применим для решения уравнения (3.29) метод регуляризации Тихонова, используя аналогию с одномерным уравнением (3.22).

Введем условие минимума функционала (ср. (3.23))

$$[Ay - f]^2 + \alpha \int \int_{-\infty}^{\infty} M(w_1, w_2) |Y(w_1, w_2)|^2 dw_1 dw_2 = \min_y. \quad (3.30)$$

Из условия (3.30) следует регуляризованное решение (ср. (2.23), (3.25))

$$y_\alpha(s_1, s_2) = \frac{1}{4\pi^2} \int \int_{-\infty}^{\infty} \frac{\lambda(-w_1, -w_2) F(w_1, w_2)}{L(w_1, w_2) + \alpha M(w_1, w_2)} e^{-i(w_1 s_1 + w_2 s_2)} dw_1 dw_2, \quad (3.31)$$

где (ср. (3.26))

$$\begin{aligned} L(w_1, w_2) &= |\lambda(w_1, w_2)|^2 = \lambda(-w_1, -w_2) \lambda(w_1, w_2) = \\ &= \operatorname{Re}^2 \lambda(w_1, w_2) + \operatorname{Im}^2 \lambda(w_1, w_2) \end{aligned} \quad (3.32)$$

(см. также (2.25), (2.26)) или (ср. (3.27))

$$y_\alpha(s_1, s_2) = \int \int_{-\infty}^{\infty} R_\alpha(s_1 - x_1, s_2 - x_2) f(x_1, x_2) dx_1 dx_2, \quad (3.33)$$

где (ср. (3.28))

$$R_\alpha(s_1, s_2) = \frac{1}{4\pi^2} \int \int_{-\infty}^{\infty} \frac{\lambda(-w_1, -w_2)}{L(w_1, w_2) + \alpha M(w_1, w_2)} e^{-i(w_1 s_1 + w_2 s_2)} dw_1 dw_2. \quad (3.34)$$

Регуляризатор $M(w_1, w_2)$ (ср. (3.24)) в работе [8] выбран в виде:

$$M(w_1, w_2) = 1 + (w_1^2 + w_2^2)^2, \quad (3.35)$$

однако сложение безразмерного слагаемого 1 и слагаемого $(w_1^2 + w_2^2)^2$, имеющего размерность, строго говоря, недопустимую, так как в зависимости от выбранной системы единиц соотношение слагаемых будет меняться, и будет меняться степень подавления высоких гармоник в решении. Поэтому более эффективным является, например, выражение

$$M(w_1, w_2) = (w_1^2 + w_2^2)^2, \quad (3.36)$$

хотя и в нем слагаемые w_1^2 и w_2^2 имеют, вообще говоря, разную физическую размерность.

При практической реализации, когда x_1, s_1, x_2, s_2 задаются дискретно, и в конечных пределах двумерные непрерывные преобразования Фурье (2.25), (2.26), (3.31), (3.34) заменяются на двумерные дискретные преобразования Фурье, которые вычисляются как набор одномерных дискретных преобразований Фурье по схеме (9.51). При этом одномерные дискретные преобразования Фурье вычисляются по алгоритму быстрого преобразования Фурье.

Параметр регуляризации α выбирается теми же способами, что и в случае одномерного уравнения (3.22) (способами невязки, обобщенной невязки, подбора и др.).

Отметим также, что для устойчивого решения уравнений I рода (интегральных, дифференциальных, алгебраических и т.д.), помимо метода регуляризации Тихонова, развиты следующие методы: методы регуляризации Лаврентьева, Денисова, «погружения» Бакушинского, максимальной энтропии Берга, итеративной регуляризации Фридмана, Бакушинского, Морозова, локальной регуляризации Арсенина, генератор РА (регуляризирующих алгоритмов) Бакушинского, квазирешений Иванова, метод поиска решения на компакте, дескриптивной регуляризации Морозова и др. (см. [3]). Это всё *методы детерминистской регуляризации*.

Есть еще *методы статистической регуляризации*, использующие статистический (вероятностный) подход. Это (в порядке повышения точности решения и количества дополнительной информации о решении) метод максимального правдоподобия, статистической регуляризации Турчина, Халфина, Лаврентьева, Жуковского, субоптимальной фильтрации, управляемой линейной фильтрации, оптимальной фильтрации Калмана-Бьюси, Винера, максимальной апостериорной вероятности и другие (см. [3]).

3.2. Метод оптимальной фильтрации Калмана-Бьюси

В пунктах. 2.2-2.4, 3.1 были рассмотрены некоторые детерминистские методы решения уравнений. Из них наиболее эффективен (в отношении точности, устойчивости, времени компьютерной реализации, требуемой компьютерной памяти, исходной информации и т.д.) метод регуляризации Тихонова, как показывает решение большого числа прикладных задач [3-5], несмотря на то, что он использует минимум априорной дополнительной информации: лишь значения погрешностей δ и μ (а также иногда прогноз решения ψ).

Еще более точными являются оптимальные методы фильтрации Калмана (Калмана-Бьюси) и Винера — методы, использующие среди устойчивых (регулярных) методов наибольшее количество априорной информации: в методе Калмана — ковариации ошибок и матожидания правой части и решения, а в методе Винера — спектральные плотности мощности шумов правой части и решения. Эти методы относятся к методам статистической регуляризации.

Одношаговый (однократный) фильтр Калмана [4].
Рассматривается СЛАУ

$$Ay + v = f \quad (3.37)$$

или

$$\sum_{j=1}^n a_{ij} y_j + v_i = f_i, \quad i = \overline{1, m}, \quad (3.38)$$

где A — $m \times n$ -матрица, y — искомый n -вектор, f — измеренный m -вектор (замер), v — m -вектор — помеха.

Если рассматривается интегральное уравнение типа (2.5), то вместо него нужно рассматривать его дискретный аналог типа (2.10).

В методе фильтрации Калмана делаются следующие предположения:

1) матожидание случайного вектора v равно нулю:

$$E[v] = 0, \quad (3.39)$$

где запись $E[v]$ означает матожидание (или среднеарифметическое значение) по ансамблю реализаций, т. е.

$$E[v] \equiv \langle v \rangle = \lim_{Q \rightarrow \infty} \frac{1}{Q} \sum_{q=1}^Q v_{i_q} = 0, \quad i = \overline{1, m}, \quad (3.40)$$

где q — номер реализации, или эксперимента, а Q — число реализаций случайного процесса v ;

2) задана симметричная положительно определенная $m \times m$ -матрица - ковариация ошибок правой части:

$$R = E[vv^T] \quad (3.41)$$

или по подробнее:

$$R_{il} = \lim_{Q \rightarrow \infty} \frac{1}{Q} \sum_{q=1}^Q v_{i_q} v_{l_q} = 0, \quad i, l = \overline{1, m}. \quad (3.42)$$

Каждый диагональный элемент матрицы R есть квадрат среднеквадратической погрешности измерения f_i , т.е. $R_{ii} = \delta_i^2 = \sigma_i$, а внедиагональный элемент R_{il} , $i \neq l$, определяет корреляцию погрешностей v_i и v_l ;

3) задан n -вектор

$$\psi = E[y] \quad (3.43)$$

— матожидание (начальное приближение, априорная оценка, прогноз) вектора y или подробнее:

$$\psi_j = \lim_{Q \rightarrow \infty} \frac{1}{Q} \sum_{q=1}^Q y_{j_q}, \quad j = \overline{1, m}; \quad (3.44)$$

4) задана симметричная положительно определенная $n \times n$ -матрица — априорная ковариация ошибок решения:

$$M = E[(y - \psi)(y - \psi)^T]. \quad (3.45)$$

Далее искомое решение y находится из условия минимума квадратичной формы (ср. (3.9)):

$$(Ay - f)^T R^{-1} (Ay - f) + (y - \psi)^T M^{-1} (y - \psi) = \min_y. \quad (3.46)$$

Из условия (3.46) получается (ср. (3.10)) *решение* (апостериорная оценка y , свертка замера с прогнозом):

$$\hat{y} = \psi + (M^{-1} + A^T R^{-1} A)^{-1} A^T R^{-1} (f - A\psi), \quad (3.47)$$

причем апостериорная $n \times n$ -матрица ковариаций ошибок решения \hat{y} равна

$$P \equiv E[(\hat{y} - y)(\hat{y} - y)^T] = (M^{-1} + A^T R^{-1} A)^{-1}. \quad (3.48)$$

Итак, если помимо f и A дополнительно известны R , ψ и M , то уточненное решение уравнения (3.37) согласно методу фильтрации Калмана выразится формулой (3.47), а уточненная матрица ковариаций ошибок решения — формулой (3.48).

Сравнение одношагового фильтра Калмана с методом регуляризации Тихонова. Если A , y и f вещественны, то условие минимизации (3.9) в методе регуляризации Тихонова можно записать в виде:

$$(Ay - f)^T (Ay - f) + \alpha (y - \psi)^T (y - \psi) = \min_y. \quad (3.49)$$

Сравнение (3.46) и (3.49) показывает, что роль α играет (символически) R/M . Наиболее же отчетливо сравнение методов Калмана и Тихонова выполняется в случае, когда

$$M = \varepsilon^2 E, \quad R = \delta^2 E, \quad (3.50)$$

где ε и δ — априорные среднеквадратические ошибки решения и правой части, а E — единичная матрица. В этом случае решение методом Калмана имеет вид (см. (3.47)):

$$\hat{y} = \psi + \left(\frac{\delta^2}{\varepsilon^2} E + A^T A \right)^{-1} A^T (f - A\psi), \quad (3.51)$$

а апостериорная матрица ковариаций ошибок решения равна (см. (3.48)):

$$P = \delta^2 \left(\frac{\delta^2}{\varepsilon^2} E + A^T A \right)^{-1}. \quad (3.52)$$

Решение же методом регуляризации Тихонова имеет вид:

$$y_\alpha = \psi + (\alpha E + A^T A)^{-1} A^T (f - A\psi), \quad (3.53)$$

Сравнение (3.51) и (3.53) показывает, что при

$$\alpha = \frac{\delta^2}{\varepsilon^2} \quad (3.54)$$

и справедливости (3.50) методы Калмана и Тихонова дают одинаковые решения. При этом, как следует из (3.52) и (3.54), апостериорная матрица ковариаций ошибок решения равна

$$P_\alpha = \delta^2 (\alpha E + A^T A)^{-1}, \quad (3.55)$$

где

$$\|\Delta y_\alpha\| = \sqrt{\|P_\alpha\|} \leq \frac{\delta}{\sqrt{\alpha}} \quad (3.56)$$

— оценка ошибки решения по норме.

Многошаговый (многократный) фильтр Калмана. Для одношагового процесса, когда имеется лишь одна реализация вектора f , требование об априорном знании ψ и M , содержащееся в методе Калмана, трудновыполнимо. Поэтому фильтр Калмана обычно применяется для многошаговых процессов, когда в функции времени поступают новые реализации f , а ψ и M итеративно уточняются. Схема многошагового фильтра Калмана выглядит следующим образом [3].

Из априорных соображений выбираются *начальные приближения* для решения $y_0 \equiv \psi$ и матрицы ковариаций ошибок решения $P_0 = M$. Для выбора начальных приближений можно использовать метод регуляризации Тихонова и положить (см. (3.8), (3.55)):

$$y_0 = (\alpha E + A^T A)^{-1} A^T f, \quad (3.57)$$

$$P_0 = \delta^2 (\alpha E + A^T A)^{-1}.$$

Последующие приближения будут найдены согласно следующей итерационной схеме (ср. (3.47), (3.48)):

$$y_k = y_{k-1} + (P_{k-1}^{-1} + A^T R_k^{-1} A)^{-1} A^T R_k^{-1} (f_k - A y_{k-1}), \quad (3.58)$$

$$P_k = (P_{k-1}^{-1} + A^T R_k^{-1} A)^{-1}, \quad k = 1, 2, 3, \dots \quad (3.59)$$

По мере того, как значения $f_k \equiv f_{ik}$ будут флуктуировать вокруг точных f , значения $y_k \equiv y_{ik}$ от итерации к итерации будут приближаться к точным значениям y , демонстрируя сходимость процесса итераций (3.58) к точному решению.

При этом может иметь место *стационарный случай*, когда R_k от n не зависит, т. е. $R_k = R$, а индекс k у y и P означает номер уточняющей итерации искомого y и P . Может иметь место и *нестационарный случай*, когда индекс k означает не только номер уточняющей итерации, но и переменность (зависимость от времени $t = t_k$) значений y , P , f , R .

Видим, что многошаговый фильтр Калмана обладает большими возможностями обработки результатов измерений, но требует большого объема данных.

3.3. Метод оптимальной линейной фильтрации Винера

Существо метода. Рассматривается интегральное уравнение Фредгольма I рода типа свертки (ср. (8.3), (3.22)):

$$\int_{-\infty}^{\infty} K(x-s)y(s)ds = f(x), \quad -\infty \leq x \leq \infty. \quad (3.60)$$

В методе фильтрации Винера делаются следующие **предположения** [4]:

1) Искомая функция $y(s)$ и погрешность правой части $v(x)$ являются реализациями стационарных, некоррелированных между собой случайных процессов.

2) Полагаются известными статистические характеристики этих процессов: *спектральная плотность мощности искомого решения*

$$R_y(w) = \lim_{T \rightarrow \infty} \frac{1}{2T} E \left[\left| \int_{-T}^T y(s) e^{iws} ds \right|^2 \right] \quad (3.61)$$

и СПМ помехи

$$R_v(w) = \lim_{T \rightarrow \infty} \frac{1}{2T} E \left[\left| \int_{-T}^T v(x) e^{iwx} dx \right|^2 \right]. \quad (3.62)$$

Заметим, что согласно теореме Винера-Хинчина [4]

$$R_y(w) = \int_{-\infty}^{\infty} r_y(x) e^{iwx} dx, \quad (3.63)$$

$$R_v(w) = \int_{-\infty}^{\infty} r_v(x) e^{iwx} dx, \quad (3.64)$$

где $r_y(x) = E[y(s+x)y(s)]$, $r_v(x) = E[v(x'+x)v(x')]$ *автокорреляционные функции искомого решения и помехи.*

В методе Винера решение $y_R(s)$ ищется исходя из условия минимума величины $E[y_R(s) - y(s)]^2$, где $y(s)$ — точное решение. В результате [4]

$$y_R(s) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\lambda(-w)F(w)}{L(w) + R_v(w) / R_y(w)} e^{-iws} dw, \quad (3.65)$$

$$E[y_R(s) - y(s)]^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{R_v(w)}{L(w) + R_v(w) / R_y(w)} dw. \quad (3.66)$$

Формула (3.65) дает решение уравнения (3.60) методом фильтрации Винера, а формула (3.66) — значение среднеквадратического отклонения этого решения от точного, являющееся минимально возможным — поэтому фильтр называется оптимальным.

Сравнение методов Винера и Тихонова. Сопоставление формул (3.25) и (3.65) показывает, что регуляризованные решения, даваемые методами Тихонова и Винера, переходят одно в другое при

$$\alpha M(\omega) = R_v(\omega) / R_y(\omega). \quad (3.67)$$

Заметим следующее. Если функцию $R_v(\omega)$ (СПМ помехи) можно получить в результате спектральной обработки ряда реализаций чисто шумового процесса (без полезного сигнала), то функцию $R_y(\omega)$ (СПМ решения) аналогичным образом получить невозможно. Поэтому метод фильтрации Винера следует рассматривать, строго говоря, не как рабочий метод, а как теоретический метод, являющийся предельно точным среди всевозможных методов решения уравнения (3.60).

Отметим также, что как следует из (3.65), устойчивость метода Винера тем выше, чем больше отношение $R_v(\omega) / R_y(\omega)$, играющее роль отношения помеха/сигнал. Другими словами, помеха стабилизирует решение. Однако, как следует из (3.66), с ростом помехи $R_v(\omega)$ растет погрешность решения. Вывод: в методе Винера имеет место компромисс между устойчивостью и точностью решения. Впрочем, такого рода компромисс имеет место и в методе регуляризации Тихонова, но он в нем не является оптимальным.

Тема 4. Метод регуляризации нахождения нормального решения системы линейных алгебраических уравнений

4.1. Приближенное нахождение нормального решения по неточно известной правой части

Пусть z^0 есть нормальное решение системы

$$Az = \bar{u}. \quad (4.1)$$

Сначала для простоты будем полагать, что приближенной может быть лишь правая часть, а оператор (матрица) A – точный.

Итак, пусть вместо \bar{u} мы имеем вектор \tilde{u} , $\|\tilde{u} - \bar{u}\| \leq \delta$, т.е. вместо системы (4.1) имеем систему

$$Az = \tilde{u}. \quad (4.2)$$

Требуется найти приближение \tilde{z}_δ к нормальному решению системы (4.1), т.е. к вектору z^0 такое, что $\tilde{z}_\delta \rightarrow z^0$ при $\delta \rightarrow 0$. Отметим, что векторы \bar{u} и \tilde{u} (один из них или оба) могут не удовлетворять классическому условию разрешимости.

Поскольку система (4.1) может быть неразрешимой, то $\inf \|Az - \bar{u}\| = \mu \geq 0$, где \inf берется по всем векторам $z \in R^n$.

Естественно искать приближения z_δ в классе Q_δ векторов z , сопоставимых по точности с исходными данными, т.е. таких, что $\|Az - \tilde{u}\| \leq \mu + \delta$. Но поскольку вместо вектора \bar{u} мы имеем вектор \tilde{u} , то

мы можем найти лишь
$$\tilde{\mu} = \inf_{z \in R^n} \|Az - \tilde{u}\|.$$

Отметим, что из очевидных неравенств

$$\|Az - \tilde{u}\| \leq \|Az - \bar{u}\| + \|\tilde{u} - \bar{u}\|, \quad \|Az - \bar{u}\| \leq \|Az - \tilde{u}\| + \|\tilde{u} - \bar{u}\| \quad \text{следуют}$$

оценки $\tilde{\mu} \leq \mu + \delta$, $\mu \leq \tilde{\mu} + \delta$, приводящие к неравенству $|\tilde{\mu} - \mu| \leq \delta$.

Поэтому будем искать приближение \tilde{z}_δ к нормальному решению z^0 в

классе \tilde{Q}_δ векторов z , для которых $\|Az - \tilde{u}\| \leq \tilde{\mu} + 2\delta$. Отметим, что если имеется информация о разрешимости системы (4.1), то $\mu = 0$ и в качестве класса \tilde{Q}_δ можно брать класс векторов z , для которых $\|Az - \tilde{u}\| \leq \delta$. Класс \tilde{Q}_δ есть класс формально возможных приближенных решений. Но нельзя в качестве \tilde{z}_δ брать произвольный вектор из класса \tilde{Q}_δ , так как такое “приближение” будет неустойчивым к малым изменениям правой части уравнения (4.2). Необходим принцип отбора. Он естественным образом вытекает из постановки задачи. В самом деле, согласно определению нормального решения искомое решение z^0 должно быть псевдорешением с минимальной нормой. Поэтому в качестве приближения к z^0 естественно брать вектор \tilde{z}_δ из \tilde{Q}_δ , минимизирующий функционал $\Omega[z] = \|z\|^2$ на множестве \tilde{Q}_δ .

Таким образом, задача сводится к минимизации функционала $\Omega[z] = \|z\|^2$ на множестве \tilde{Q}_δ векторов z , для которых выполняется условие $\|Az - \tilde{u}\| \leq \tilde{\mu} + 2\delta$.

Пусть z_δ - вектор из \tilde{Q}_δ , на котором функционал $\|z\|^2$ достигает минимума на множестве \tilde{Q}_δ . Его можно рассматривать как результат применения к правой части \tilde{u} уравнения (4.2) некоторого оператора $R_1(\tilde{u}, \delta)$, зависящего от параметра δ . Справедлива [5]

Теорема 1. *Оператор $R_1(\tilde{u}, \delta)$ обладает следующими свойствами:*

- 1) *он определён для всякого $\tilde{u} \in R^m$ и любого $\delta > 0$;*
- 2) *при $\delta \rightarrow 0$ $z_\delta = R_1(\tilde{u}, \delta)$ стремится к нормальному решению z^0 уравнения $Az = \tilde{u}$, т.е. он является регулирующим для уравнения $Az = u$.*

Пусть z_δ - вектор, на котором функционал $\Omega[z] = \|z\|^2$ достигает минимума на множестве \tilde{Q}_δ . Легко видеть из наглядных геометрических

представлений, что вектор z_δ лежит на границе множества \tilde{Q}_δ , т.е. $\|Az_\delta - \tilde{u}\| = \tilde{\mu} + 2\delta = \delta_1$. Это следует непосредственно также из того, что функционал $\Omega[z] = \|z\|^2$ является стабилизирующим и квазимонотонным [5].

Задачу нахождения вектора z_δ можно поставить так: Среди векторов z , удовлетворяющих условию $\|Az - \tilde{u}\| = \tilde{\mu} + 2\delta$, найти вектор z_δ с минимальной нормой, т.е. минимизирующий функционал $\Omega[z] = \|z\|^2$.

Последнюю задачу можно решать методом Лагранжа, т.е. в качестве z_δ брать вектор z^α , минимизирующий функционал

$$M^\alpha[z, \tilde{u}] = \|Az - \tilde{u}\|^2 + \alpha \|z\|^2, \quad \alpha > 0,$$

с параметром α , определяемым по невязке, т.е. из условия $\|Az^\alpha - \tilde{u}\| = \delta_1$. При этом параметр α определяется однозначно [5].

4.2. Приближенное нахождение нормального решения по неточно заданным правой части и матрице

Теперь рассмотрим случай, когда неточно заданы как правая часть u , так и матрица A , т.е. пусть вместо системы уравнений $A\bar{z} = \bar{u}$ с нормальным решением z^0 мы имеем систему

$$A\tilde{z} = \tilde{u}, \quad z \in R^n, \quad \tilde{u} \in R^m, \quad (4.3)$$

где $\|\tilde{u} - \bar{u}\| \leq \delta$ и $\|\tilde{A} - \bar{A}\| = \sup_{\|z\|=1} \|\tilde{A}z - \bar{A}z\| \leq h$. Требуется найти векторы

$z_\gamma, \gamma = (\delta, h)$, такие, что $\lim_{\gamma \rightarrow 0} \|z_\gamma - z^0\| = 0$. Такие векторы z_γ мы будем

называть *приближениями к нормальному решению* z^0 уравнения $A\bar{z} = \bar{u}$, а также *приближенными нормальными решениями*. Их можно находить путем минимизации сглаживающего функционала

$$M^\alpha[z, \tilde{u}, \tilde{A}] = \|\tilde{A}z - \tilde{u}\|^2 + \alpha \|z\|^2, \quad (4.4)$$

подобно изложенному ранее.

Рассмотрим функционал (4.4). Для него справедлива [5]

Теорема 3. Для всяких $\tilde{u} \in R^m$, $\tilde{A}, \alpha > 0$ существует единственный вектор z^α , минимизирующий функционал $M^\alpha[z, \tilde{u}, \tilde{A}]$.

В условиях, когда вместо правой части $u = \bar{u}$ и оператора \bar{A} мы имеем их приближения, \tilde{u} и \tilde{A} , такие, что $\|\tilde{u} - \bar{u}\| \leq \delta$, $\|\tilde{A} - \bar{A}\| \leq h$, классом $\tilde{Q}_\gamma, \gamma \equiv (\delta, h)$, сравнимых по точности с исходными данными (то есть допустимых) приближений к нормальному решению z^0 уравнения $\bar{A}z = \bar{u}$ является множество векторов z , удовлетворяющих условию

$$\|\tilde{A}z - \tilde{u}\| \leq 2(h\|z\| + \delta) + \tilde{\mu}.$$

Поскольку нормальное решение z^0 обладает свойством минимальности нормы, то естественно задачу нахождения приближений к нормальному решению z^0 поставить так: в классе \tilde{Q}_γ векторов z найти вектор z_γ , минимизирующий функционал $\Omega[z] = \|z\|^2$, то есть найти такой вектор

$$z_\gamma \in \tilde{Q}_\gamma \equiv \{z; \|\tilde{A}z - \tilde{u}\| \leq 2(h\|z\| + \delta) + \tilde{\mu}\}, \text{ что } \inf_{z \in \tilde{Q}_\gamma} \|z\|^2 = \|z_\gamma\|^2.$$

Из наглядных геометрических представлений легко видеть, что справедлива

Лемма. Вектор z_γ , минимизирующий функционал $\|z\|^2$ на множестве \tilde{Q}_γ , удовлетворяет условию

$$\|\tilde{A}z_\gamma - \tilde{u}\| \leq 2(h\|z_\gamma\| + \delta) + \tilde{\mu}.$$

Поэтому упомянутая задача сводится к следующей: на множестве векторов z , удовлетворяющих условию

$$\|\tilde{A}z - \tilde{u}\| \leq 2(h\|z\| + \delta) + \tilde{\mu},$$

найти вектор \mathbf{z}_γ минимальной нормы.

Если числа δ и h известны, то эту задачу можно решать методом неопределенных множителей Лагранжа, то есть находить вектор \mathbf{z}^α , минимизирующий (на всем пространстве \mathbf{R}^n) сглаживающий функционал

$$M^\alpha[\mathbf{z}, \tilde{\mathbf{u}}, \tilde{\mathbf{A}}] = \|\tilde{\mathbf{A}}\mathbf{z} - \tilde{\mathbf{u}}\|^2 + \alpha\|\mathbf{z}\|^2,$$

а параметр α определять из условия

$$\|\tilde{\mathbf{A}}\mathbf{z}^\alpha - \tilde{\mathbf{u}}\| = 2(h\|\mathbf{z}^\alpha\| + \delta) + \tilde{\mu}.$$

При этом функции $\varphi(\alpha) = \|\tilde{\mathbf{A}}\mathbf{z}^\alpha - \tilde{\mathbf{u}}\|^2$ и $\psi(\alpha) = (h\|\mathbf{z}^\alpha\| + \delta)^2$ - строго монотонные, $\varphi(\alpha)$ - возрастающая, $\psi(\alpha)$ - убывающая. Поэтому, если числа h и δ известны, то α определяется однозначно.

Тема 5. Метод регуляризации решения линейных интегральных уравнений первого рода

В настоящем разделе подробно излагается метод регуляризации построения приближенных решений линейных интегральных уравнений первого рода. Изложение не опирается на общие результаты, касающиеся метода регуляризации, изложенные ранее, в применении к произвольным уравнениям вида $Az = u$. Из предыдущего материала используется лишь понятие регуляризирующего оператора.

5.1. Существование регуляризирующих операторов для интегральных уравнений первого рода

В настоящем параграфе мы рассмотрим вариационный способ построения регуляризирующих операторов, т.е. вариационный способ построения приближенных решений, для линейных интегральных уравнений первого рода, устойчивых к малым изменениям правой части [5].

1. Будем рассматривать уравнение вида

$$Az \equiv \int_a^b K(x,s)z(s)ds = u(x), \quad x \in [c,d], \quad (5.1)$$

с конечным промежутком интегрирования $[a,b]$, в котором ядро $K(x,s)$ непрерывно по совокупности переменных (x,s) в замкнутой области $\{a \leq s \leq b; c \leq x \leq d\}$. Уклонение правой части $u(x)$ будем оценивать в метрике L_2 , т.е. по формуле

$$\rho_{L_2}(u_1, u_2) = \left\{ \int_c^d [u_1(x) - u_2(x)]^2 dx \right\}^{1/2},$$

а уклонение решения $z(s)$ – в метрике C , т.е. по формуле

$$\rho_C(z_1, z_2) = \max_{s \in [a,b]} |z_1(s) - z_2(s)|.$$

Обозначим через F_1 класс непрерывных на $[a,b]$ функций $z(s)$, имеющих обобщенные производные первого порядка $z'(s)$, интегрируемые с квадратом на $[a,b]$ [9]. Аналогично через F_n будем обозначать класс непрерывных на $[a,b]$ функций $z(s)$, имеющих производные до n -го порядка, интегрируемые с квадратом на $[a,b]$.

Будем полагать, что уравнение (5.1) с точной правой частью $u = u_T(x)$ имеет единственное на множестве F_1 решение $z_T(s)$, $z_T(s) \in F_1$.

Рассмотрим совокупность всех функций $u(x)$, интегрируемых с квадратом на $[c,d]$ с метрикой, определяемой по формуле

$$\rho_{L_2}(u_1, u_2) = \left\{ \int_c^d [u_1(x) - u_2(x)]^2 dx \right\}^{1/2}.$$

Это метрическое пространство $L_2(c,d)$.

Пусть вместо функции $u_T(x)$ мы имеем функцию $u_\delta(x)$ из $L_2(c,d)$ такую, что $\rho_{L_2}(u_\delta, u_T) \leq \delta$, причем

$$\delta^2 < \int_c^d u_\delta^2(x) dx, \left(\|u_\delta\|^2 > \delta^2 \right). \quad (5.2)$$

Таким образом, вместо уравнения

$$\int_a^b K(x, s) z(s) ds = u_T(x)$$

мы имеем уравнение

$$\int_a^b K(x, s) z(s) ds = u_\delta(x). \quad (5.3)$$

В этих условиях речь может идти лишь о нахождении приближенного (к $z_T(s)$) решения уравнения (5.3).

5.2. Интегральное уравнение I рода – некорректная задача:

$$\begin{aligned} \delta z(s) &= \exp(i\omega s), \quad \omega \gg 1: \quad |\delta z| = 1 \\ \Rightarrow \delta u(x) &= \int_a^b K(x, s) \delta u(s) ds = \int_a^b K(x, s) e^{i\omega s} ds. \end{aligned}$$

Интегрируем по частям:

$$\delta u = \frac{1}{i\omega} e^{i\omega s} K(x, s) \Big|_{s=a}^{s=b} - \frac{1}{i\omega} \int_a^b \frac{\partial K(x, s)}{\partial s} e^{i\omega s} ds = O\left(\frac{1}{\omega}\right).$$

При достаточно больших ω $\|\delta u\| \ll \frac{1}{\omega}$ – сколько угодно мала.

Следовательно, малым изменениям правой части соответствуют не малые изменения решения. Таким образом, задача неустойчива.

Кроме того, задача имеет решение не при любых непрерывных правых частях. Непрерывное уравнение с вырожденным ядром:

$$K(x, s) = \sum_{n=1}^N A_n(x) B_n(s) \Rightarrow \sum_{n=1}^N \beta_n A_n(x) = u(x), \quad \beta_n = \int_a^b B_n(s) u(s) ds$$
 – справедливо

для $u(x)$, представимых в виде линейной комбинации функций $A_n(x)$, для других правых частей – не имеет решения. Таким образом, если даже при

некоторой $u = \bar{u}$ существует решение, имеются малые изменения правой части δu , при которых решение не существует.

Очевидно, непосредственно решать задачу при неточно заданной правой части бессмысленно. Если u задача с погрешностью, то соответствующее решение или не существует или отлично от искомого решения z на величину δz , которая может быть большой. Даже если u задано точно, но нахождение решения выполняется численными методами, то неизбежно вносятся погрешности метода и округления. Это снова приводит к большой погрешности решения.

Будем искать приближенное решение среди функций множества F_1 , точнее в классе Q_δ функций из F_1 , для которых

$$\rho_{L_2}(Az, u_\delta) \leq \delta,$$

т.е. в классе функций $F_{1,\delta} = Q_\delta \cap F_1$. Они сопоставимы по точности с исходными данными.

Таким образом, $F_{1,\delta}$ является множеством возможных приближенных решений уравнения (5.3), сопоставимых по точности с исходными данными.

Однако нельзя брать в качестве приближенного решения уравнения (5.3) произвольную функцию из множества $F_{1,\delta}$, так как такое “приближенное решение” не будет устойчивым к малым изменениям правой части $u_\delta(x)$ и может, как угодно сильно отличаться от $z_T(s)$. Множество $F_{1,\delta}$ слишком широкое. Необходим принцип отбора, обеспечивающий получение приближенного решения, устойчивого к малым изменениям правой части $u_\delta(x)$. Отбор можно реализовать следующим образом.

На функциях $z(s)$ из F_1 , очевидно, определены функционалы вида

$$\Omega[z] = \int_a^b \left\{ q(s)z^2(s) + p(s) \left(\frac{dz}{ds} \right)^2 \right\} ds, \quad (5.4)$$

где $q(s)$ и $p(s)$ - заданные неотрицательные непрерывные функции такие, что для всякого $s \in [a, b]$ $q^2(s) + p^2(s) \neq 0$ и $p(s) \geq p_0 > 0$, где p_0 - число. Возьмём один из них, фиксированный.

Поскольку $\Omega[z]$ - неотрицательный функционал, то существует его точная нижняя грань на множестве $F_{1,\delta}$, т.е. число

$$\Omega_0 = \inf_{z \in F_{1,\delta}} \Omega[z].$$

Упомянутый выше принцип отбора состоит в том, что в качестве искомого приближенного решения предлагается брать функцию $z_\delta(s)$, на которой достигается точная нижняя грань функционала $\Omega[z]$ на множестве $F_{1,\delta}$.

Функцию $z_\delta(s)$ можно рассматривать как результат применения к правой части $u_\delta(x)$ уравнения (5.3) некоторого оператора \tilde{R} , зависящего от параметра δ , т.е.

$$z_\delta(s) = \tilde{R}(u_\delta, \delta).$$

Справедливы утверждения:

1) оператор $\tilde{R}(u_\delta, \delta)$ определен на всякой функции u_δ , интегрируемой с квадратом на $[c, d]$ (т.е. $u_\delta \in L_2(c, d)$), и значения его $z_\delta(s) = \tilde{R}(u_\delta, \delta)$ принадлежат множеству F_1 ;

2) оператор $\tilde{R}(u, \delta)$ является регуляризирующим для уравнения (5.1).

Оператор $\tilde{R}(u, \delta)$ называется регуляризирующим для уравнения $Az = u$ в окрестности $u = u_T$, если

1) $\exists \delta_1 > 0$, что $\tilde{R}(u, \delta)$ определен $\forall \delta > 0$ и $\forall u \in U$, для которого $\rho_U(u, u_T) \leq \delta \leq \delta_1$;

2) $\exists \alpha = \alpha(\delta)$, что $\forall \varepsilon > 0 \exists \delta(\varepsilon) \leq \delta_1$, что если $u_\delta \in U$ удовлетворяет $\rho_U(u_\delta, u_T) \leq \delta(\varepsilon)$, то $\rho_F(z_\alpha, z_T) \leq \varepsilon$, где $z_\alpha = \tilde{R}(u_\delta, \alpha(\delta))$.

(Однозначность \tilde{R} в определении не предполагается).

Таким образом, если $\rho_U(u_T, u_\delta) \leq \delta$, то в качестве приближенного решения уравнения $Az = u$ можно взять регуляризованное решение $z_\alpha = \tilde{R}(u_\delta, \alpha)$, где $\alpha = \alpha(\delta)$ – параметр регуляризации.

Если за меру гладкости возможных приближенных решений принять значения функционала $\Omega[z]$, то согласно предлагаемому принципу отбора в качестве приближенного решения уравнения (5.3) берется наиболее гладкое из возможных решений.

Компактное множество: множество $M \in X$, X – линейное пространство, такое что каждая бесконечная последовательность $\{x_i\}$ содержит последовательность, сходящуюся к некоторой точке $x_0 \in X$. Если $x_0 \in M$, то M называется компактным в себе множеством. (Каждое бесконечное подмножество M имеет предельную точку x_0 .)

Компакт – метризуемое компактное пространство.

Примеры: отрезок, окружность, n - мерный куб, шар, сфера. n - мерное Евклидово пространство не является компактом, а его подмножества являются компактами тогда и только тогда, когда оно замкнуто (содержит все свои предельные точки) и ограничено. (Теорема Больцано-Вейерштрасса: из любой ограниченной последовательности можно выделить сходящуюся подпоследовательность. Если она замкнута, то $x_0 \in M$.) Из теоремы Вейерштрасса о достижении непрерывной функции на замкнутом ограниченном множестве своих точных верхних и нижних граней следует

Теорема. Если $z \xrightarrow{A} u$, $z \in F$, $u \in U$, причем F - компакт, а A - непрерывный и взаимно однозначный оператор, то U также компакт и обратный оператор A^{-1} также непрерывен, то есть задача корректна.

5.3. Редукция задачи построения регуляризирующих операторов к классической вариационной задаче минимизации функционалов с ограничениями

1. Таким образом, для нахождения приближенного решения уравнения (5.3) с приближенно известной правой частью $u_\delta(x)$ надо решить следующую вариационную задачу минимизации функционала $\Omega[z]$ с ограничениями: среди функций $z(s)$, принадлежащих множеству F_1 (или F_n) и удовлетворяющих условию $\rho_{L_2}(Az, u_\delta) \leq \delta$, найти такую функцию $z_\delta(s)$, которая минимизирует функционал $\Omega[z]$ на множестве $F_{1,\delta}$ (соответственно на $F_{n,\delta}$).

Такие задачи можно решать прямыми методами минимизации функционалов. Однако для этих задач они трудно реализуемы на ЭВМ. Чтобы естественнее и проще подойти к их решению, воспользуемся следующей леммой.

Лемма 1. Точная нижняя грань функционала $\Omega[z]$ вида (4; 1,4) на множестве $F_{1,\delta}$ достигается на функции $z_\delta(s)$ из F_1 , для которой $\rho_{L_2}(Az_\delta, u_\delta) = \delta$.

2. С учетом этой леммы задача нахождения функции $z_\delta(s)$ из F_1 , удовлетворяющей условию $\rho_{L_2}(Az_\delta, u_\delta) \leq \delta$ и минимизирующей функционал $\Omega[z]$ на множестве $F_{1,\delta}$, сводится к задаче: среди функций $z(s)$ из F_1 , удовлетворяющих условию

$$\rho_{L_2}(Az, u_\delta) \leq \delta,$$

найти функцию $z_\delta(s)$, минимизирующую функционал $\Omega[z]$ на множестве $F_{1,\delta}$.

Это классическая задача вариационного исчисления на условный минимум. Её можно решать методом неопределенных множителей Лагранжа. А именно, находить функцию $z_\alpha(s)$, минимизирующую сглаживающий функционал

$$M^\alpha[z, u_\delta] = \rho_{L_2}^2(Az, u_\delta) + \alpha\Omega[z]$$

(это задача на безусловный минимум), а параметр α определять из условия

$$\rho_{L_2}(Az_\alpha, u_\delta) = \delta,$$

т.е. по невязке. Это уравнение относительно α имеет решение $\alpha = \alpha(\delta)$.

Следовательно, к задаче минимизации сглаживающего функционала $M^\alpha[z, u_\delta]$ с определением параметра α по невязке сводится и исходная задача нахождения функции $z_\delta(s) = z_{\alpha(\delta)}(s)$ из F_1 , удовлетворяющей условию

$$\rho_{L_2}(Az_\delta, u_\delta) \leq \delta$$

и минимизирующей функционал $\Omega[z]$ на множестве $F_{1,\delta}$. Последняя задача и исходная эквивалентны.

Функцию $z_\alpha(s)$, где $\alpha = \alpha(\delta)$, можно рассматривать как результат применения к правой части $u_\delta(x)$ уравнения (4; 1,3) некоторого оператора $R_1(u_\delta, \alpha(\delta))$.

Таким образом, в качестве приближенного решения уравнения (5.3) берется решение другой задачи – задачи минимизации функционала $M^\alpha[z, u_\delta]$ с определением α по невязке, “близкой” при малой погрешности u_δ к исходной задаче нахождения решения уравнения (5.3).

Решение этой последней задачи $z_{\alpha(\delta)}(s)$ устойчиво к малым изменениям в метрике L_2 правой части $u_\delta(x)$.

Аналогичным образом редуцируется исходная вариационная задача к задаче на безусловный минимум при пользовании функционалами $\Omega[z]$ вида (5.4).

Функционалы $\Omega[z]$ играют при этом стабилизирующую роль. Поэтому их называют стабилизирующими функционалами или стабилизаторами, соответственно 1-го и n-го порядков.

5.4. Получение семейства регуляризирующих операторов с помощью минимизации сглаживающих функционалов

Сглаживающий функционал $M^\alpha[z, u]$ можно ввести в рассмотрение формально, не связывая его с редукцией исходной вариационной задачи к классической. В настоящем параграфе будет показано, что путем решения задачи на минимум сглаживающего функционала с фиксированным стабилизатором $\Omega[z]$ и с выбором параметра регуляризации α как некоторой функции $\alpha = \alpha(\delta)$ от уровня погрешности δ правой части $u_\delta(x)$ уравнения (5.3) можно получить бесконечное множество регуляризирующих операторов (а, следовательно, и приближенных решений уравнения (5.3)), различающихся выбором функций $\alpha = \alpha(\delta)$.

1. Итак, рассмотрим сглаживающий функционал

$$M^\alpha[z, u] = \rho_{L_2}^2(Az, u) + \alpha\Omega[z]$$

со стабилизатором $\Omega[z]$ вида (5.4). Тогда справедлива

Теорема 1. Каковы бы ни были параметр $\alpha > 0$ и функция $u(x)$ из $L_2(c, d)$, существует единственная функция $z_\alpha(s)$, имеющая производную $z'_\alpha(s)$, интегрируемую с квадратом на $[a, b]$, на которой функционал $M^\alpha[z, u]$ достигает своей точной нижней грани на множестве F_1 , т.е.

$$M_0^\alpha = \inf_{z \in F_1} M^\alpha[z, u] = M^\alpha[z_\alpha, u].$$

5.5. Алгоритм нахождения приближенных решений, легко реализуемый на ЭВМ

Как было показано в пункте 5.1, для нахождения приближенного (регуляризованного) решения уравнения (5.3) достаточно найти функцию $z_\alpha(s) \in F_1$, минимизирующую сглаживающий функционал $M^\alpha[z, u_\delta]$, и соответствующее значение α . Задачу минимизации функционала $M^\alpha[z, u_\delta]$ можно решать прямыми методами, например, методами наискорейшего спуска. Однако на ЭВМ удобнее находить функцию $z_\alpha(s)$, решая уравнение Эйлера, соответствующее функционалу $M^\alpha[z, u_\delta]$. Более подробному описанию этого способа и посвящен настоящий параграф.

На функциях $z(s)$, имеющих непрерывные производные второго порядка, стабилизирующий функционал $\Omega[z]$ вида

$$\Omega[z] = \int_a^b \left\{ q(s)z^2(s) + p(s) \left(\frac{dz}{ds} \right)^2 \right\} ds$$

путем интегрирования по частям можно записать в виде

$$\Omega[z] = (z, Lz) + p(b)z(b)z'(b) - p(a)z(a)z'(a),$$

где оператор Lz имеет вид

$$Lz \equiv qz - \frac{d}{ds} \left(p \frac{dz}{ds} \right),$$

а (z, Lz) - скалярное произведение вида

$$(z, Lz) = \int_a^b z \cdot Lz \, ds.$$

Следовательно, функционал $M^\alpha[z, u_\delta]$ можно записать в виде

$$M^\alpha[z, u_\delta] = \rho_{L_2}^2 (Az, u_\delta) + \alpha(z, Lz) + \alpha[p(b)z(b)z'(b) - p(a)z(a)z'(a)]$$

или, записывая подробнее $\rho_{L_2}^2(Az, u_\delta)$,

$$M^\alpha[z, u_\delta] = \int_c^d \left\{ \int_a^b K(x, s)z(s)ds - u_\delta(x) \right\}^2 dx + \\ + \alpha(z, Lz) + \alpha[p(b)z(b)z'(b) - p(a)z(a)z'(a)].$$

Необходимым условием минимума функционала $M^\alpha[z, u_\delta]$ является равенство нулю его первой вариации ΔM^α . Известно, что

$$\Delta M^\alpha = \left\{ \frac{d}{d\gamma} M^\alpha[z + \gamma v, u_\delta] \right\}_{\gamma=0}. \quad \text{Выполняя необходимые вычисления,}$$

получим

$$\Delta M^\alpha \equiv A^*Az + \alpha Lz + \alpha[p(b)v(b)z'(b) + p(b)z(b) \times \\ \times v'(b) - p(a)v(a)z'(a) - p(a)z(a)v'(a)] - A^*u_\delta = 0, \quad (5.5)$$

где A^*u - оператор вида

$$A^*u \equiv \int_c^d K(x, s)u(x)dx,$$

сопряженный оператору

$$Az \equiv \int_a^b K(x, s)z(s)ds$$

в пространстве L_2 , а $v(s)$ - вариация функции $z(s)$. Оператор A^*Az можно записать в виде

$$A^*Az \equiv \int_a^b \bar{K}(s, t)z(t)dt, \quad (5.6)$$

где

$$\bar{K}(s, t) = \int_c^d K(x, s)K(x, t)dx.$$

Условие $\Delta M^\alpha = 0$ выполняется, если

$$A^*Az + \alpha Lz = A^*u_\delta \quad (5.7)$$

и

$$p(b)[z(b)v'(b) + z'(b)v(b)] - p(a)[z(a)v'(a) + z'(a)v(a)] = 0. \quad (5.8)$$

Условие (5.8) выполняется, если на концах $s = a$ и $s = b$ промежутка $[a, b]$ решение $z_\alpha(s)$ уравнения (5.3) или его производная $z'_\alpha(s)$ равны нулю.

Таким образом, искомой функцией $z_\alpha(s)$, минимизирующей функционал $M^\alpha[z, u_\delta]$, т.е. искомым приближенным (регуляризованным) решением уравнения (5.3) будет решение интегро-дифференциального уравнения Эйлера (5.7) или

$$\alpha \left\{ q(s)z - \frac{d}{ds} \left[p(s) \frac{dz}{ds} \right] \right\} + \int_a^b \bar{K}(s, t) z(t) dt = g(s), \quad (5.9)$$

где

$$\bar{K}(s, t) = \int_c^d K(x, s) K(x, t) dx, \quad (5.10)$$

$$g(s) = \int_c^d K(x, s) u_\delta(x) dx, \quad (5.11)$$

удовлетворяющее крайевым условиям одного из следующих типов:

либо

$$z(a) = 0, \quad z(b) = 0, \quad (5.12)$$

либо

$$z(a) = 0, \quad z'(b) = 0, \quad (5.13)$$

либо

$$z'(a) = 0, \quad z(b) = 0, \quad (5.14)$$

либо

$$z'(a) = 0, \quad z'(b) = 0. \quad (5.15)$$

2. Если по смыслу задачи нам известны на концах промежутка $[a, b]$ значения точного решения $z_T(s)$ или его производной $z'_T(s)$, то

естественно искать приближенные решения в классе функций $z(s)$, удовлетворяющих тем же граничным условиям, что и $z_T(s)$, так как в противном случае аппроксимация с помощью приближенных решений значений точного решения $z_T(s)$ на концах промежутка $[a, b]$ может быть неудовлетворительной даже при сколь угодно малом отклонении приближенной правой части $u_\delta(x)$ от точной $u_T(x)$.

3. Если при нахождении приближенных решений уравнения (5.3) пользоваться стабилизаторами n -го порядка, то уравнение Эйлера для сглаживающего функционала $M^\alpha[z, u_\delta]$ будет иметь вид

$$A^*Az + \alpha Lz = A^*u_\delta,$$

где

$$Lz \equiv \sum_{r=0}^n (-1)^r \frac{d^r}{ds^r} \left[q_r(s) \frac{d^r z}{ds^r} \right],$$

так как стабилизатор n -го порядка можно записать в виде скалярного произведения (z, Lz) .

Легко написать и возможные краевые условия для искомых решений уравнения Эйлера. Например,

$$z(a) = z'(a) = \dots = z^{(n-1)}(a) = 0,$$

$$z(b) = z'(b) = \dots = z^{(n-1)}(b) = 0$$

и другие.

Все сказанное в п. 2 настоящего параграфа о согласованности краевых условий точного решения и приближенных решений, получаемых с помощью уравнения Эйлера, в равной мере относится и к этому случаю. Достичь такой согласованности также можно с помощью соответствующих замен неизвестной функции $z(s)$ на новую функцию $\tilde{z}(s)$.

При фактической реализации описанного алгоритма нахождения приближенных решений надо переходить к дискретному аналогу рассмотренной задачи. Такой переход, или, как мы будем говорить, дискретизацию задачи можно произвести по-разному. Вопросам дискретизации посвящен следующий параграф.

5.6. О дискретизации задачи нахождения приближенных решений интегральных уравнений первого рода

Переход к дискретному аналогу задачи нахождения регуляризованных приближенных решений уравнения (5.3) можно произвести по-разному. Представляются очевидными три подхода к дискретизации этой задачи.

Первый подход. Производится дискретизация исходного уравнения (5.3) путем замены интеграла интегральной суммой по некоторой квадратурной формуле. В результате получается система линейных алгебраических уравнений (СЛАУ), вырожденная или плохо обусловленная, приближенное решение которой устойчивое к малым изменениям вектора правой части, и надо находить. Это можно сделать методом регуляризации. Если пользоваться при этом вариационным подходом, то, взяв дискретный аналог стабилизатора $\Omega[z]$, образуем аналог сглаживающего функционала. Затем переходим к его уравнению Эйлера, которое будет представлять собою регуляризованную систему линейных алгебраических уравнений. Решения этой системы (с соответственно подобранными значениями параметра регуляризации α) и будут приближенными решениями уравнения (5.3).

Второй подход. Производится дискретизация сглаживающего функционала, а далее решается задача минимизации функции многих переменных с последующим определением параметра регуляризации α .

Третий подход. Производится дискретизация краевой задачи для уравнения Эйлера (5.7), и далее решается получающаяся при этом система линейных алгебраических уравнений.

Эти способы не эквивалентны друг другу [5]. В некотором смысле более предпочтительным представляется третий подход. Рассмотрим его подробнее.

1. Для простоты изложения дискретизацию будем производить на равномерной сетке.

Будем предполагать, что ядро $K(x,s)$ в уравнении (5.3) есть вещественная, непрерывная в области $\{a \leq s \leq b; c \leq x \leq d\}$ функция. Возьмем в качестве стабилизирующего функционала $\Omega[z]$ функционал вида

$$\Omega[z] = \int_a^b \{z^2 + p(z')^2\} ds, \quad (5.16)$$

где p - положительное число.

Пусть точное решение $z_T(s)$ принадлежит F_1 и удовлетворяет краевым условиям: $z'(a) = 0, z'(b) = 0$. Тогда в качестве регуляризованных решений $z_\alpha(s)$ уравнения (5.3) можно брать функции, являющиеся решениями следующей краевой задачи для уравнения Эйлера:

$$\int_a^b \bar{K}(s,t)z(t)dt + \alpha\{z(s) - pz''(s)\} = g(s), \quad (5.17)$$

$$z'(a) = 0, z'(b) = 0,$$

где

$$\bar{K}(s,t) = \int_c^d K(x,s)K(x,t)dx, \quad g(s) = \int_c^d K(x,s)u_\delta(x)dx.$$

Напишем разностный аналог уравнения (5.17) на равномерной сетке с шагом h . Разобьём промежуток $[a,b]$ на n равных частей и возьмем в качестве узловых точек сетки середины полученных отрезков, т.е. полагаем

$$s_i = a + 0,5 \cdot h + (i-1)h, \quad i = 1, 2, \dots, n; \quad h = \frac{b-a}{n}.$$

Заменяя в левой части уравнения (5.17) интеграл соответствующей ему интегральной суммой, например, по формуле прямоугольников, а $z''(s)$ - соответствующим разностным отношением, получим

$$\sum_{j=1}^n \bar{K}(s_i, t_j) h z_j + \alpha z_i + p \alpha \frac{2z_i - z_{i-1} - z_{i+1}}{h^2} = g_i, \quad (5.18)$$

где

$$i = 1, 2, \dots, n, \quad g_i = \int_c^d K(x, s_i) u_\delta(x) dx.$$

Значения $\bar{K}(s_i, t_j)$ и g_i либо вычисляются аналитически, либо получаются с помощью соответствующих квадратурных формул. Отметим, что при этом число точек сетки по s не связано с числом точек сетки по x .

При $i=1$ и $i=n$ в (5.18) входят не определенные еще значения z_0 и z_{n+1} . Чтобы удовлетворить граничным условиям, полагаем $z_0 = z_1$ и $z_{n+1} = z_n$. Пусть B - матрица с элементами $B_{ij} = \bar{K}(s_i, t_j) h$. Тогда систему уравнений (5.18) относительно вектора z с компонентами (z_1, z_2, \dots, z_n) можно записать в виде

$$B_\alpha z \equiv Bz + \alpha Cz = g, \quad (5.19)$$

где g - вектор с компонентами (g_1, g_2, \dots, g_n) , а αC - симметричная матрица вида

$$\begin{pmatrix} \alpha\left(1+\frac{p}{h^2}\right) & \frac{-\alpha p}{h^2} & 0 & \dots & 0 & 0 \\ \frac{-\alpha p}{h^2} & \alpha\left(1+\frac{2p}{h^2}\right) & \frac{-\alpha p}{h^2} & \dots & 0 & 0 \\ 0 & \frac{-\alpha p}{h^2} & \alpha\left(1+\frac{2p}{h^2}\right) & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & & & \dots & \alpha\left(1+\frac{2p}{h^2}\right) & \frac{-\alpha p}{h^2} \\ 0 & 0 & 0 & \dots & \frac{-\alpha p}{h^2} & \alpha\left(1+\frac{p}{h^2}\right) \end{pmatrix}$$

Таким образом, задача сводится к решению СЛАУ(5.19). Матрица этой системы B_α симметрична. Поэтому для решения её можно использовать весьма экономичные методы.

Тема 6. Представление решения в виде конечных рядов Фурье

6.1. Явление Гиббса

Ряд Фурье:

$$f(x) \approx \sum_{k=0}^{\infty} c_k \varphi_k(x),$$

$\{\varphi_k\}$ - ортонормированная система функций на (a, b) ,

$$c_k = \int_a^b f(x) \varphi_k(x) dx, \quad k = 0, 1, 2, \dots$$

Тригонометрический ряд Фурье:

$$\frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx),$$

$$a_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos kx dx, \quad b_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin kx dx$$

Начнем с частного случая прямоугольной волны $H(t)$ с периодом 2π (рис. 1). Если вычислить сумму первых $2n$ членов, то все члены с косинусами будут равны нулю и получаем

$$H_{2n}(t) = \frac{1}{2} + \frac{2}{\pi} \sum_{k=1}^n \frac{1}{2k-1} \sin(2k-1)t \quad (6.1)$$

Так как

$$a_0 = \frac{1}{\pi} \int_0^{2\pi} H(t) \cos 0 \cdot x \, dx = \frac{1}{\pi} \cdot \pi = 1,$$

$$a_k = \frac{1}{\pi} \int_0^{\pi} \cos kx \, dx = \frac{1}{\pi} \cdot \frac{1}{k} \sin kx \Big|_0^{\pi} = \frac{1}{k\pi} \sin k\pi = 0,$$

$$b_k = \frac{1}{\pi} \int_0^{\pi} \sin kx \, dx = -\frac{1}{\pi} \cdot \frac{1}{k} \cos kx \Big|_0^{\pi} = -\frac{1}{k\pi} [\cos k\pi - 1] = \frac{1}{k\pi} [1 - (-1)^k],$$

$$b_k = \begin{cases} 0, & k\text{- четное: } k=2m, \\ \frac{2}{k\pi}, & k\text{- нечетное: } k=2m-1, \end{cases}$$

то сумма $2n$ первых ненулевых членов для

$$H(t) : H_{2n}(t) = \frac{1}{2} + \frac{2}{\pi} \sum_{k=1}^n \frac{1}{2k-1} \sin(2k-1)t.$$

6.2. Метод Фейера:

$$S_N^{(t)} = \frac{1}{N+1} \sum_{n=0}^N f_n(t),$$

$f_n(t)$ - частичная сумма ряда Фурье для $f(x)$, $f_n(t) = \frac{a_0}{2} + \sum_{k=1}^n (a_k \cos kt + b_k \sin kt)$.

Таким образом,

$$S_N(t) = \frac{a_0}{2} + \sum_{n=1}^N \left(1 - \frac{n}{N+1}\right) (a_n \cos nt + b_n \sin nt).$$

Если $f(t)$, то $S_N(t)$ сходится к $f(t)$ равномерно. Если $f(t)$ принадлежит классу функций, удовлетворяющих условию Липшица

$$|f(x) - f(x')| \leq M \cdot |x - x'|^\alpha, \quad (0 < \alpha \leq 1), x, x' \in [a, b], M = \text{const}, \text{ то } \|f(t) - S_N(t)\|_C = O\left(\frac{1}{n^\alpha}\right).$$

Гиббс отметил, что частичная сумма H_{2n} превосходит функцию $H(t)$ в точках $\frac{\pi}{2n}$ на некоторую величину (рис. 2). Более точно

$$H_{2n}\left(\frac{\pi}{2n}\right) \rightarrow 1.08949\dots, \text{ когда } n \rightarrow \infty. \quad (6.2)$$

Действительно, $H_{2n}(t)$ не только превосходит функцию $H(t)$, но и имеет тенденцию колебаться около $H(t)$, и колебания уменьшаются медленно, когда t удаляется от разрыва.

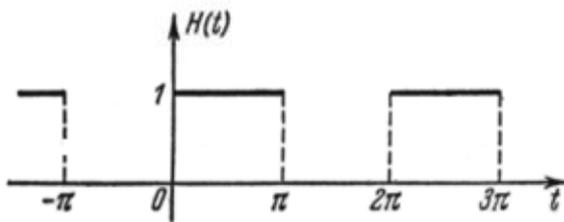


Рис. 1. Прямоугольная волна

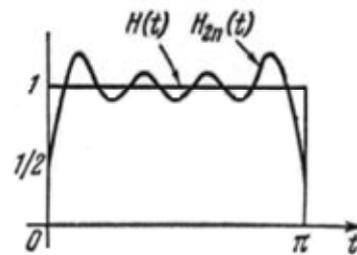


Рис. 2. Явление Гиббса

Чтобы объяснить явление, запишем (см. рис. 1) как

$$\begin{aligned} H_{2n}(t) &= \frac{1}{2} + \frac{2}{\pi} \sum_{k=1}^n \int_0^t \cos(2k-1)x dx = \\ &= \frac{1}{2} + \frac{2}{\pi} \int_0^t \sum_{k=1}^n \cos(2k-1)x dx = \frac{1}{2} + \frac{1}{\pi} \int_0^t \frac{\sin 2nx}{\sin x} dx, \end{aligned} \quad (6.3)$$

где использована формула

$$\sum_{k=1}^n \cos(2k-1)x = \frac{\sin 2nx}{2 \sin x}.$$

Из (6.3) ясно, что максимум и минимум для $0 \leq t \leq \pi$ достигается в точках

$$\frac{dH_{2n}(t)}{dt} = \frac{1}{\pi} \frac{\sin 2nt}{\sin t} = 0,$$

т.е. при

$$t = \frac{m\pi}{2n}, \quad m = \overline{1, 2n-1}, \quad (6.4)$$

и что они чередуются. Их величины были вычислены Карслоу.

То, что верно для этой специальной функции, очевидно, верно и для более общих функций, так как разрыв можно рассматривать как возникающий из прямоугольной волны, прибавленной к главной функции.

6.3. Сигма-множители Ланцоша

Заменяем быстро колеблющуюся функцию $H_{2n}(t)$ сглаженной функцией

$$\overline{H}_{2n}(t) = \frac{n}{\pi} \int_{t-\frac{\pi}{2n}}^{t+\frac{\pi}{2n}} H_{2n}(\tau) d\tau, \quad (6.5)$$

где усредняется одно полное колебание $H_{2n}(t)$, сосредоточенное около t .

Подставив (6.1) в (6.5), получаем

$$\begin{aligned} \overline{H}_{2n}(t) &= \frac{n}{\pi} \int_{t-\frac{\pi}{2n}}^{t+\frac{\pi}{2n}} \left[\frac{1}{2} + \frac{2}{\pi} \sum_{k=1}^n \frac{1}{2k-1} \sin(2k-1)\tau \right] d\tau = \\ &= \frac{n}{\pi} \left[\frac{\pi}{2n} + \frac{2}{\pi} \sum_{k=1}^n \frac{-1}{(2k-1)^2} \cos(2k-1)\tau \right]_{t-\frac{\pi}{2n}}^{t+\frac{\pi}{2n}} = \\ &= \frac{1}{2} + \frac{2}{\pi} \sum_{k=1}^n \frac{1}{2k-1} \frac{\sin\left[(2k-1)\frac{\pi}{2n}\right]}{(2k-1)\frac{\pi}{2n}} \sin(2k-1)t. \end{aligned} \quad (6.6)$$

Если сравнить это выражение с $H_{2n}(t)$, то получим дополнительный множитель

$$\sigma_{2k-1} = \frac{\sin(2k-1)\frac{\pi}{2n}}{(2k-1)\frac{\pi}{2n}} \quad (6.7)$$

для каждого члена суммирования.

Эффект этого множителя σ_k состоит в том, чтобы уменьшить максимум с 0.08949 до 0.01187 и первый минимум с 0.04859 до 0.00473 и

т.д. Таким образом, явление Гиббса сильно уменьшилось от присутствия σ -множителей, которые возникли из сглаживания $H_{2n}(t)$ на коротком интервале длиной π/n .

Мы изучили частный случай прямоугольной волны; покажем теперь, что влияние σ_k -множителей остается тем же для любого ряда Фурье.

Пусть $f(t)$ ($0 \leq t \leq 2\pi$) интегрируема и пусть

$$a_k = \frac{1}{\pi} \int_0^{2\pi} f(t) \cos kt dt, \quad k = 0, 1, 2, \dots,$$

$$b_k = \frac{1}{\pi} \int_0^{2\pi} f(t) \sin kt dt, \quad k = 1, 2, \dots,$$

- коэффициенты Фурье. Тогда

$$f_n(t) = \frac{a_0}{2} + \sum_{k=1}^n (a_k \cos kt + b_k \sin kt).$$

Теперь вычислим

$$\begin{aligned} f_n(t) &= \frac{1}{\pi/n} \int_{t-\frac{\pi}{2n}}^{t+\frac{\pi}{2n}} f_n(\tau) d\tau = \frac{n}{\pi} \left[\frac{a_0}{2} \frac{\pi}{n} + \sum_{k=1}^n \left(a_k \frac{\sin k\tau}{k} - b_k \frac{\cos k\tau}{k} \right) \Big|_{t-\frac{\pi}{2n}}^{t+\frac{\pi}{2n}} \right] = \\ &= \frac{n}{\pi} \left(\frac{\pi}{2n} a_0 + \sum_{k=1}^n \left\{ \frac{a_k}{k} \left[\sin k \left(t + \frac{\pi}{2n} \right) - \sin k \left(t - \frac{\pi}{2n} \right) \right] - \right. \right. \\ &\quad \left. \left. - \frac{b_k}{k} \left[\cos k \left(t + \frac{\pi}{2n} \right) - \cos k \left(t - \frac{\pi}{2n} \right) \right] \right\} \right) = \\ &= \frac{a_0}{2} + \frac{n}{\pi} \sum_{k=1}^n \left(\frac{2a_k}{k} \sin \frac{\pi k}{2n} \cos kt + \frac{2b_k}{k} \sin \frac{\pi k}{2n} \sin kt \right) = \\ &= \frac{a_0}{2} + \sum_{k=1}^n \frac{\sin \frac{\pi k}{2n}}{\frac{\pi k}{2n}} (a_k \cos kt + b_k \sin kt) \end{aligned}$$

и опять получаем σ -множители (равенство (6.7))

$$\sigma_k = \frac{\sin \frac{\pi k}{2n}}{\frac{\pi k}{2n}}, \quad (6.8)$$

вставленные в различные коэффициенты ряда Фурье. Заметим, что члены $\sin kt$ и $\cos kt$ оба имеют множитель σ_k .

6.4. Сравнение методов сходимости

Кроме метода σ -множителей Ланцоша, существует хорошо известный метод Фейера, использующий средние арифметические частичных сумм. Метод Фейера совершенно исключает колебание, тогда как метод Ланцоша только гасит его. Рисунок 3 дает соответствующее сравнение кривых ряда Фурье, суммы Фейера, σ -множителя Ланцоша для двенадцати членной аппроксимации прямоугольной волны. Преимущество метода σ -множителя очевидно.

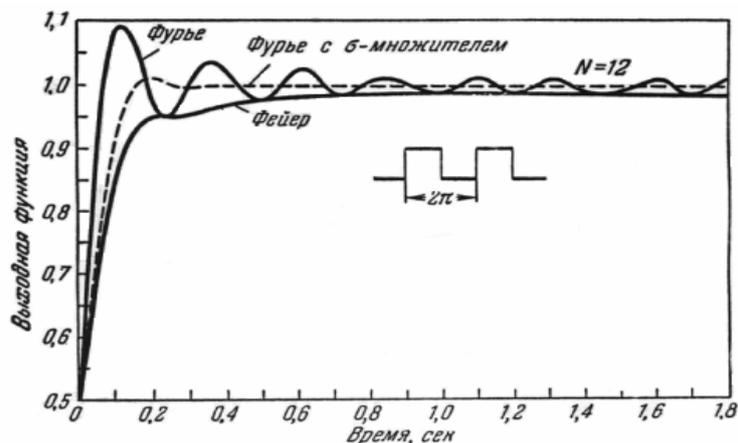


Рис.3. Частичные суммы Фурье $= \frac{1}{2} + \frac{2}{\pi} \sum_{n=0}^N \frac{1}{2n+1} \sin(2n+1)t$;

$$\text{Фейера} = \frac{1}{2} + \frac{1}{\pi(N+1)} \sum_{n=0}^N \frac{2(N-n)+1}{2n+1} \sin(2n+1)t;$$

Фурье с σ -множителем сходимости =

$$= \frac{1}{2} + \frac{2}{\pi} \sum_{n=0}^N \frac{1}{2n+1} \frac{\sin \frac{(2n+1)\pi}{2N+2}}{\frac{(2n+1)\pi}{2N+2}} \sin(2n+1)t.$$

Рисунок 4 показывает скорость, с которой метод σ -множителя приближает прямоугольную волну как функцию n . Таким образом, даже для умеренного n “время роста” кривой очень невелико.

6.5. Техника дифференцирования по Ланцошу

Иногда бывает необходимо продифференцировать ряд Фурье. При этом часто получаются нежелательные в физических задачах большие колебания высокой частоты.

Ланцош предложил практический прием, состоящий в использовании

$$\frac{f\left(t+\frac{\pi}{n}\right)-f\left(t-\frac{\pi}{n}\right)}{\frac{2\pi}{n}} \quad (6.9)$$

в качестве оценки производной $f'(t)$ в точке t , вместо предельного процесса, который имеет сомнительное физическое значение, для дифференцирования приближения усеченным рядом Фурье. Заметим, что то же самое n возникает и в (6.9) и в порядке усеченного ряда. Мы не будем воспроизводить здесь выводы; заметим только, что те же самые σ -множители (6.8), введенные в формально дифференцированный ряд, дают соответствующий результат.

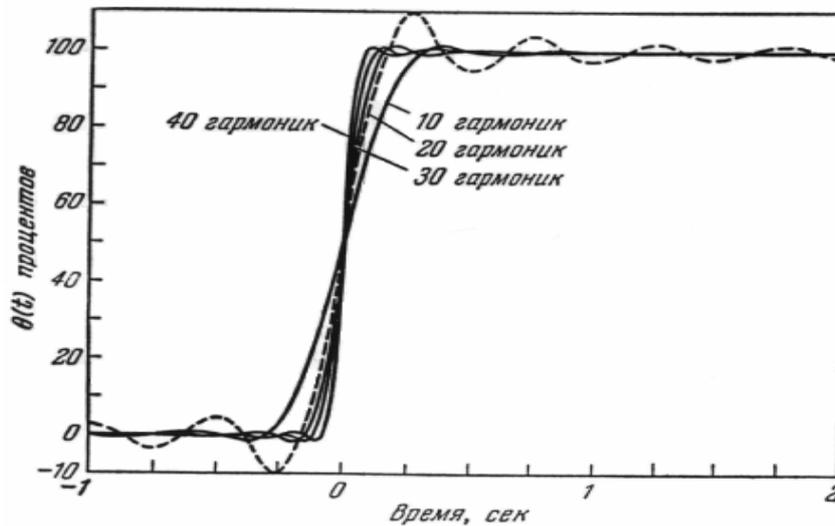


Рис. 4. Прямоугольная функция входа $f = \frac{1}{10} \frac{1}{\text{сек}}$; $T = 10 \text{ сек}$.

Приближение Фурье и приближение Фурье с множителем сходимости.

Тема 7. Устойчивые методы суммирования рядов Фурье с приближенными в метрике l_2 коэффициентами

7.1. Задача устойчивого суммирования ряда Фурье

1. Задача суммирования ряда Фурье по заданной ортонормированной системе функций $\{\varphi_n(t)\}$ состоит в нахождении функции $f(t)$ по ее коэффициентам Фурье.

В ряде случаев в эксперименте для определения интересующей нас функции $f(t)$, характеризующей изучаемый процесс (явление), измеряются ее коэффициенты Фурье $\{a_n\}$ по некоторой ортонормированной системе функций $\{\varphi_n(t)\}$. Измерения всегда носят приближенный характер. Таким образом, вместо a_n получают приближенные значения коэффициентов Фурье c_n . Возникает задача суммирования рядов Фурье с приближенными коэффициентами.

Задача суммирования рядов Фурье не обладает свойством устойчивости к малым изменениям (в метрике l_2) коэффициентов Фурье, если отклонение суммы оценивать в метрике C , и, следовательно, является некорректно поставленной задачей.

Издавна употребляемый метод суммирования рядов Фурье с приближенными коэффициентами состоит в том, что в качестве приближенного значения суммы такого ряда $f(t)$ берется сумма конечного (не слишком большого) числа его первых членов, т.е. полагают

$$f(t) \approx \sum_{n=1}^k c_n \varphi_n(t).$$

Будем называть метод суммирования рядов Фурье функцией $f(t) \in F$ с приближенными в метрике l_2 коэффициентами $\{c_n\}$, устойчивыми в смысле метрики пространства F , если для любого положительного числа $\varepsilon > 0$ можно указать такое число $\delta(\varepsilon)$, что из неравенства

$$\sum_{n=1}^{\infty} (c_n - \tilde{c}_n)^2 \leq \delta^2(\varepsilon)$$

следует неравенство

$$\rho_F(f, \tilde{f}) \leq \varepsilon,$$

где f и \tilde{f} - результаты суммирования данным методом соответственно рядов

$$\sum_{n=1}^{\infty} c_n \varphi_n(t) \text{ и } \sum_{n=1}^{\infty} \tilde{c}_n \varphi_n(t).$$

2. Обозначим через C_D пространство непрерывных в конечной замкнутой области \bar{D} функций с метрикой C . Будем рассматривать ряды Фурье по системе функций $\{\varphi_n(t)\}$ для функций $f(t)$ из C_D .

Построим класс устойчивых методов суммирования рядов Фурье, основанных на идее регуляризации.

Задачу суммирования ряда Фурье функции $f(t)$ можно рассматривать как задачу решения некоторого операторного уравнения относительно функции $f(t)$. В самом деле, если каждой функции из множества F поставить в соответствие элемент u пространства l_2 , а именно, последовательность ее коэффициентов Фурье $\{a_n\}$ по системе $\{\varphi_n(t)\}$ (с весом $\rho(t)$), т.е. элемент $u \equiv \{a_n\}$, то это соответствие можно записать в виде

$$Af = u \quad (7.1)$$

Очевидно, что этот оператор из C_D на l_2 непрерывен на C_D .

Следовательно, задача суммирования ряда Фурье, состоящая в нахождении функции $f(t)$ по заданной последовательности ее коэффициентов Фурье $u \equiv \{a_n\}$, сводится к решению уравнения (7.1)

относительно $f(t)$. Из курса математического анализа известно, что в классе C_D эта задача имеет единственное решение.

7.2. Классы устойчивых методов суммирования рядов Фурье

1. Если нам известны приближенные значения коэффициентов Фурье искомой функции, т.е. элемента u - правой части уравнения (7.1), то речь может идти лишь о нахождении приближенного решения задачи. Поскольку эта задача является некорректно поставленной, то естественно для нахождения ее приближенного решения воспользоваться методом регуляризации.

В качестве стабилизирующего функционала $\Omega[f]$ возьмем некоторый функционал вида

$$\Omega[f] = \sum_{n=1}^{\infty} f_n^2 \cdot \xi_n, \quad (7.2)$$

определяемый заданием последовательности $\{\xi_n\}$. Здесь f_n - коэффициенты Фурье функции $f(t)$ по полной ортонормированной системе функций $\{\varphi_n(t)\}$ (с весом $\rho(t) > 0$), а $\{\xi_n\}$ - последовательность положительных чисел, порядок роста которых при $n \rightarrow \infty$ не ниже, чем $n^{2+\varepsilon}$, где $\varepsilon \geq 0$. Легко видеть, что для любого положительного числа $d > 0$ множество функций $f(t) \in C_D$, для которых $\Omega[f] \leq d$, является компактом в C_D .

Стабилизирующие функционалы такого рода являются естественным обобщением функционалов $\Omega[f]$.

В самом деле, пусть $\{\varphi_n(t)\}$ - полная ортонормированная система собственных функций краевой задачи вида

$$\begin{aligned} \frac{d}{dt} [k(t)\varphi'(t)] - q(t)\varphi(t) + \lambda\varphi(t) &= 0, \quad 0 \leq t \leq l, \\ \varphi(0) &= 0 = \varphi(l), \end{aligned}$$

где $k(t) > 0$ и $q(t) \geq 0$ на отрезке $[0, l]$, а $\{\lambda_n\}$ - совокупность собственных значений этой задачи. Тогда стабилизирующий функционал

$$\Omega[f] = \int_0^l \left\{ k(f')^2 + qf^2 \right\} dt,$$

для функций $f(t)$, удовлетворяющих условиям $f(0) = f(l) = 0$, можно записать в виде ряда $\Omega[f] = \sum_{n=1}^{\infty} f_n^2 \lambda_n$, где f_n - коэффициенты Фурье функции $f(t)$ по системе $\{\varphi_n(t)\}$. Действительно, применяя интегрирование по частям, получаем

$$\int_0^l k(f')^2 dt = \left. f f' k \right|_0^l - \int_0^l f \frac{d}{dt}(k f') dt = - \int_0^l f \frac{d}{dt}(k f') dt.$$

Следовательно, $\Omega[f] = \int_0^l f \left\{ qf - \frac{d}{dt}(k f') \right\} dt$. Подставляя в правую часть

этой формулы вместо функции $f(t)$ ее ряд Фурье $\sum_{n=1}^{\infty} f_n \cdot \varphi_n(t)$, получаем

$$\begin{aligned} \Omega[f] &= \int_0^l \left(\sum_{m=1}^{\infty} f_m \varphi_m(t) \right) \sum_{n=1}^{\infty} f_n \left\{ q \varphi_n(t) - \frac{d}{dt}[k \cdot \varphi_n'] \right\} dt = \\ &= \sum_{n,m=1}^{\infty} f_n f_m \int_0^l \varphi_m(t) \left\{ q \cdot \varphi_n - \frac{d}{dt}(k \cdot \varphi_n') \right\} dt. \end{aligned}$$

Так как $q \varphi_n - \frac{d}{dt}(k \varphi_n') \equiv \lambda_n \varphi_n$, и при $n \neq m$ функции $\varphi_n(t)$ и $\varphi_m(t)$ ортогональны, то

$$\Omega[f] = \sum_{n,m=1}^{\infty} f_n f_m \lambda_n \int_0^l \varphi_m \varphi_n dt = \sum_{n=1}^{\infty} f_n^2 \lambda_n.$$

2. Уточним постановку задачи. Обозначим через \mathfrak{X} совокупность всех последовательностей $\{\xi_n\}$, упомянутых выше, отвечающих всем возможным значениям числа $\varepsilon \geq 0$.

Пусть в конечной замкнутой области \bar{D} N -мерного евклидова пространства \mathbf{R}^N заданы полная ортонормированная (с весом $\rho(t) > 0$)

система функций $\{\varphi_n(t)\}$, где $t = (t_1, t_2, \dots, t_N)$, и непрерывная в \bar{D} функция $f(t)$, представимая своим рядом Фурье по системе $\{\varphi_n(t)\}$,

$$\hat{f}(t) = \sum_{n=1}^{\infty} a_n \varphi_n(t),$$

где

$$a_n = \int_D \rho(t) \hat{f}(t) \varphi_n(t) dt.$$

Пусть вместо коэффициентов Фурье a_n нам известны их приближенные значения $c_n = a_n + \gamma_n$ с малыми (в l_2) погрешностями γ_n ,

$$\sum_{n=1}^{\infty} \gamma_n^2 \leq \delta^2,$$

т.е. вместо последовательности $\hat{u} \equiv \{a_n\}$ нам дана последовательность

$u_\delta \equiv \{c_n\}$, для которой $\rho_{l_2}(\hat{u}, u_\delta) \leq \delta$. Следовательно, вместо ряда Фурье с

точными коэффициентами

$$\sum_{n=1}^{\infty} a_n \varphi_n(t)$$

мы имеем ряд с приближенными коэффициентами

$$\sum_{n=1}^{\infty} c_n \varphi_n(t) \tag{7.3}$$

Нашей целью является отыскание в классе C_D функций $\tilde{f}(t)$, аппроксимирующих функцию $\hat{f}(t)$, по последовательности чисел $\{c_n\}$,

близкой (в метрике l_2) к последовательности $\{a_n\}$ коэффициентов Фурье

функции $\hat{f}(t)$, т.е. такой, что $\sum_{n=1}^{\infty} (c_n - a_n)^2 \leq \delta^2$. При этом аппроксимация

должна быть такой, чтобы при $\delta \rightarrow 0$ $\rho_c(\tilde{f}, \hat{f}) \rightarrow 0$.

В качестве $\tilde{f}(t)$ нельзя брать сумму $S(t)$ ряда (7.3), вычисляемую по правилу

$$S(t) = \lim_{k \rightarrow \infty} \sum_{n=1}^k c_n \varphi_n(t),$$

так как такое суммирование не является устойчивым к малым (в l_2) изменениям коэффициентов $\{c_n\}$.

Очевидно, решение надо искать в классе Q_δ функций из C_D , для которых выполняется неравенство

$$\rho_{l_2}(Af, u_\delta) \leq \delta.$$

Но этот класс не является компактным. Он слишком широк. Необходимо сузить его. Для этого возьмем некоторый фиксированный функционал $\Omega_1[f]$ вида (7.2)

$$\Omega_1[f] = \sum_{n=1}^{\infty} f_n^2 \xi_n,$$

где $\{\xi_n\} \in \mathfrak{X}$.

Пусть F_ξ - множество функций из C_D , для которых определен функционал $\Omega_1[f]$, и

$$F_{\delta, \xi} = Q_\delta \cap F_\xi.$$

Приближение функции $\hat{f}(t)$ будем искать на множестве $F_{\delta, \xi} \subset C_D$.

В дальнейшем, для определенности, будем рассматривать одномерную задачу. В этом случае t - координата точки на прямой, а область D - конечный промежуток (a, b) .

Поскольку задача нахождения функций по ее коэффициентам Фурье сводится к решению операторного уравнения (7.1), естественно находить приближение функции $\hat{f}(t)$ методом регуляризации. Для этого рассмотрим функционал

$$M^\alpha [u_\delta, f] = \rho_{t_2}^2 (Af, u_\delta) + \alpha \Omega_1 [f],$$

содержащий числовой параметр α (параметр регуляризации). Его можно написать также в виде

$$M^\alpha [u_\delta, f] = \sum_{n=1}^{\infty} (f_n - c_n)^2 + \alpha \sum_{n=1}^{\infty} f_n^2 \xi_n. \quad (7.4)$$

где f_n - коэффициенты Фурье функции $f(t)$ по системе $\{\varphi_n(t)\}$ с весом $\rho(t) > 0$, т.е.

$$f_n = \int_D f(t) \rho(t) \varphi_n(t) dt. \quad (7.5)$$

Существует функция $\tilde{f}_\alpha(t)$ из $F_{\delta, \xi}$, реализующая минимум функционала $M^\alpha [f, u_\delta]$ на множестве $F_{\delta, \xi}$. Эту функцию и будем принимать в качестве приближения функции $\hat{f}(t)$.

Функционал $\Omega_1 [f]$ будем называть стабилизирующим функционалом для задачи об устойчивом суммировании рядов вида (7.3).

Замечание 1. Если в качестве системы функций $\{\varphi_n(t)\}$ взяты собственные функции краевой задачи

$$\begin{aligned} \operatorname{div}(k \nabla \varphi) - q^2(t) \varphi + \lambda \rho(t) \varphi &= 0, \quad t \in D, \\ \varphi|_S &= 0 \quad \left(\text{или} \quad \frac{\partial \varphi}{\partial n} \Big|_S = 0 \right), \end{aligned}$$

где S - граница области D , в которой ищется решение, то функционал $\Omega_1 [f]$ можно брать в виде

$$\Omega_1 [f] = \int_D \{k (\nabla f)^2 + q^2 f^2\} dt$$

или в эквивалентном виде

$$\Omega_1 [f] = \sum_{n=1}^{\infty} f_n^2 \lambda_n,$$

где λ_n - собственные значения указанной краевой задачи, а f_n - коэффициенты Фурье функции $f(t)$ по системе $\{\varphi_n(t)\}$ с весом $\rho(t)$.

3. Коэффициенты Фурье функции $\tilde{f}_\alpha(t)$ по системе $\{\varphi_n(t)\}$ находим из условия равенства нулю частных производных функционала (7.4) по переменным f_n ($n=1, 2, \dots$). Получим

$$\tilde{f}_{\alpha,n} = \frac{c_n}{1 + \alpha \xi_n}.$$

Таким образом, искомое приближение функции $\hat{f}(t)$ можно записать в виде

$$\tilde{f}_\alpha = \sum_{n=1}^{\infty} r(n, \alpha) c_n \varphi_n(t), \quad (7.6)$$

где

$$r(n, \alpha) = \frac{1}{1 + \alpha \xi_n},$$

и вычислить $\tilde{f}_\alpha(t)$ по правилу

$$\tilde{f}_\alpha = \lim_{k \rightarrow \infty} \sum_{n=1}^k r(n, \alpha) c_n \varphi_n(t). \quad (7.7)$$

4. Формулы (7.6) и (7.7) определяют метод суммирования ряда

$$\sum_{n=1}^k c_n \varphi_n(t),$$

устойчивый в смысле метрики C к малым (в метрике l_2) изменениям его коэффициентов c_n . Действительно, описанная процедура получения функции $\tilde{f}_\alpha(t)$ может быть описана в виде оператора $R(u, \alpha)$, т.е.

$$\tilde{f}_\alpha(t) = R(u_\delta, \alpha).$$

Этот оператор является регуляризирующим для уравнения (7.1) и, следовательно, обладает упомянутым свойством устойчивости.

Заметим, что значение параметра α должно быть взято согласованным с погрешностью исходных данных δ , $\alpha = \alpha(\delta)$. Оно может

быть найдено, например, из условия $\rho_{l_2}(A\tilde{f}_\alpha, u_\delta) = \delta$, которое можно записать также в виде

$$\sum_{n=1}^{\infty} c_n^2 \frac{\alpha^2 \xi_n^2}{(1 + \alpha \xi_n)^2} = \delta^2.$$

При $\delta^2 < \sum_{n=1}^{\infty} c_n^2$ таким способом α определяется однозначно, так как функция

$$\phi(\alpha) = \sum_{n=1}^{\infty} c_n^2 \frac{\alpha^2 \xi_n^2}{(1 + \alpha \xi_n)^2}$$

- непрерывная монотонно возрастающая в области $\alpha > 0$ и $\phi(0) = 0$.

Действительно,

$$\phi'(\alpha) = \sum_{n=1}^{\infty} c_n^2 \left\{ \frac{2\alpha \xi_n^2}{(1 + \alpha \xi_n)^2} - \frac{\alpha^2 \xi_n^2 \cdot 2\xi_n}{(1 + \alpha \xi_n)^3} \right\} = 2\alpha \sum_{n=1}^{\infty} c_n^2 \frac{\xi_n^2}{(1 + \alpha \xi_n)^3} > 0.$$

Так как $\phi(\alpha) \leq \sum_{n=1}^{\infty} c_n^2$, то при $\delta^2 > \sum_{n=1}^{\infty} c_n^2$ уравнение $\phi(\alpha) = \delta^2$ не имеет решение.

Замечание 2. Сумма исходного ряда

$$\sum_{n=1}^{\infty} c_n \varphi_n(t),$$

понимаемая как предел

$$\lim_{k \rightarrow \infty} \sum_{n=1}^k c_n \varphi_n(t),$$

не может служить приближением суммы ряда

$$\sum_{n=1}^{\infty} c_n \varphi_n(t)$$

вследствие неустойчивости ее к малым (в метрике l_2) изменениям коэффициентов c_n . С другой стороны, сумма ряда

$$\sum_{n=1}^{\infty} \frac{1}{1 + \alpha \xi_n} c_n \varphi_n(t),$$

понимаемая так же, как предел

$$\lim_{k \rightarrow \infty} \sum_{n=1}^k \frac{1}{1 + \alpha \xi_n} c_n \varphi_n(t),$$

устойчива к малым (в l_2) изменениям коэффициентов c_n и при значении $\alpha = \alpha(\delta)$, согласованном с погрешностью коэффициентов c_n , равномерно аппроксимирует функцию $\hat{f}(t)$.

Таким образом, множители

$$r(n, \alpha) = \frac{1}{1 + \alpha \xi_n}$$

играют стабилизирующую роль. Мы будем называть их стабилизирующими множителями.

Если положить

$$r(k, \alpha) = \begin{cases} 1, & k \leq n, \\ 0, & k > n, \end{cases}$$

то получим обычно употребляемый метод суммирования. В этом случае надо взять последовательность $\{\xi_n\}$ вида

$$\xi_k < \infty \text{ для } k \leq n, \quad \xi_k = \infty \text{ для } k > n,$$

затем положить $\alpha = 0$.

Замечание 3. Если в классическом способе суммирования ряда Фурье путем перехода к пределу при $n \rightarrow \infty$ в частичной сумме

$$\sum_{k=1}^n c_k \varphi_k(t)$$

число n брать согласованным с погрешностью δ коэффициентов $\{c_n\}$, то такое суммирование будет устойчивым.

Замечание 4. Если функция $\hat{f}(t)$ кусочно-непрерывна, то описанный метод дает устойчивый метод суммирования в каждой точке непрерывности $\hat{f}(t)$.

Приложение 1. Основные типы уравнений

Перечислим основные типы уравнений, которыми описываются прикладные задачи, рассмотренные в основном тексте учебного пособия.

Интегральные уравнения

Одномерное (линейное) интегральное уравнение Фредгольма I рода:

$$\int_a^b K(x,s)y(s)ds = f(x), c \leq x \leq d, \quad (8.1)$$

где $K(x,s)$ - ядро, $f(x)$ - правая часть – известные функции, $y(s)$ - искомая функция ординаты, причем x - наружная переменная, а s - внутренняя переменная, $[c,d]$ - заданная область измерения $f(x)$, $[a,b]$ - задаваемая область поиска $y(s)$. Практически вместо точной $f(x)$ задается экспериментальная зашумленная правая часть $\tilde{f}(x) = f(x) + \delta f(x)$, где $\delta f(x)$ - аддитивная погрешность, вместо точного ядра $K(x,s)$ так же часто известно приближенное $\tilde{K}(x,s)$, а вместо точной $y(s)$ получаем приближенное решение $\tilde{y}(s) = y(s) + \delta y(s)$, причем $\delta y(s)$ обусловлена не только погрешностями $\delta f(x)$ и $\delta K(x,s)$, но и методом решения.

Прикладные задачи, описываемые уравнением (1.1): восстановление непрерывного спектра в обратной задаче спектроскопии (уравнение (2.42)), редукция протяженных сигналов (п. 3.2), восстановление сигналов в системе, не являющейся динамической (уравнение (4.18)), распад клеток и радиоактивных элементов (уравнения (4.26), (4.27)), вентиляция в легких (уравнения (4.26), (4.27)).

Одномерное интегральное уравнение Фредгольма I рода с разностным ядром:

$$\int_a^b K(x-s)y(s)ds = f(x), c \leq x \leq d. \quad (8.2)$$

Прикладная задача[3]: синтез магнитного поля на оси катушки ЯМР - томографа.

Одномерное интегральное уравнение Фредгольма I рода типа свертки:

$$\int_{-\infty}^{\infty} K(x-s)y(s)ds = f(x), -\infty < x < \infty. \quad (8.3)$$

Прикладные задачи[3]: снятие влияния аппаратных искажений, восстановление смазанных фотоизображений, восстановление смазанных фотоизображений биологических микрообъектов, обратная задача речевой акустики.

Одномерное интегральное уравнение Фредгольма I рода типа свертки на полуоси:

$$\int_0^{\infty} K(x-s)y(s)ds = f(x), 0 \leq x < \infty \quad (8.4),$$

Прикладная задача[3]: восстановление непрерывного спектра.

Двухмерное интегральное уравнение Фредгольма I рода типа свертки:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K(x_1-s_1, x_2-s_2)y(s_1, s_2)ds_1 ds_2 = f(x_1, x_2), -\infty < x_1, x_2 < \infty. \quad (8.5)$$

Прикладные задачи: определение плотности вещества $c(x, y)$ в РТ, визуализация результатов $c(x, y)$ на дисплее, восстановление дефокусированных фотоизображений, обратная задача диагностики плазмы, восстановление дефокусированных изображений биологических микрообъектов.

Одномерное интегральное уравнение Фредгольма II рода:

$$y(x) + \int_a^b K(x, s)y(s)ds = f(x), a \leq x \leq b. \quad (8.6)$$

Прикладная задача: восстановление сигнала в системе, не являющейся динамической.

Двухмерное интегральное уравнение Фредгольма II рода:

$$y(x_1, x_2) + \int_{a_1}^{b_1} \int_{a_2}^{b_2} K(x_1, s_1, x_2, s_2) y(s_1, s_2) ds_1 ds_2 = f(x_1, x_2) \quad (8.7)$$
$$a_1 \leq x_1 \leq b_1, a_2 \leq x_2 \leq b_2.$$

Прикладная задача: определение потенциала на поверхности в методе граничных интегральных уравнений [4].

Двухмерное интегральное уравнение Радона (уравнение интегральной геометрии):

$$\int_{\mathcal{L}(l, \theta)} c(x, y) ds = q(l, \theta), \quad (8.8)$$

где $\mathcal{L}(l, \theta)$ - луч зрения (прямая или кривая, уравнение которой задается в функции l и θ), $c(x, y)$ - искомая функция, $q(l, \theta)$ - измеренная правая часть (поглощение). Уравнение (8.8) может быть приведено к уравнению типа (8.5).

Прикладные задачи: определение плотности вещества с помощью рентгеновской томографии, обратная задача диагностики плазмы.

Нелинейное одномерное интегральное уравнение Урысона I рода:

$$\int_a^b K[x, s, y(s)] ds = f(x), c \leq x \leq d, \quad (8.9)$$

где $K[x, s, y(s)]$ - заданная нелинейная функция.

Прикладная задача: синтез магнитного поля на оси катушки ЯМР – томографа.

Линейное одномерное интегральное уравнение Вольтера I рода:

$$\int_a^x K(x, s) y(s) ds = f(x), a \leq x \leq b. \quad (8.10)$$

Прикладная задача: восстановление сигнала в динамической системе без обратной связи.

Линейное одномерное интегральное уравнение Вольтерра I рода с разностным ядром:

$$\int_a^x K(x-s)y(s)ds = f(x), a \leq x \leq b. \quad (8.11)$$

Прикладная задача: восстановление сигнала в стационарной динамической системе без обратной связи.

Линейное одномерное интегральное уравнение Вольтерра I рода типа свертки одностороннее (или на полуоси):

$$\int_0^x K(x-s)y(s)ds = f(x), x \geq 0. \quad (8.12)$$

Прикладная задача: определение импульсной переходной функции в теории управления.

Линейное одномерное интегральное уравнение типа Вольтерры I рода с параметром:

$$\frac{1}{\Delta} \int_x^{x+\Delta} \omega(\xi, y) d\xi = g(x, y), -\infty < x, y < \infty, \quad (8.13)$$

где $\omega(\xi, y)$ - искомая функция, $\Delta > 0$ - некоторое число, $g(x, y)$ - заданная правая часть, y - параметр, т.е. (8.13) можно рассматривать и как совокупность одномерных уравнений (столько уравнений, сколько задано значений). Уравнение (8.13) может быть приведено к уравнению вида (8.3)

Прикладная задача: восстановление смазанных фотоизображений.

Интегральное уравнение Цейпеля:

$$\int_l^R \frac{r}{\sqrt{r^2 - l^2}} \varepsilon(r) dr = \frac{I(l)}{2}, -R \geq l \leq R, \quad (8.14)$$

где $\varepsilon(r)$ - искомая функция, $R > 0$ - некоторое известное число (в частности $R = \infty$, если (8.14) применять к шаровым звездным скоплениям [4]), $I(l)$ -

измеренная функция. Уравнение (8.14) можно классифицировать как *линейное одномерное интегральное уравнение типа Вольтерры I рода сингулярное*.

Прикладная задача: обратная задача диагностики плазмы (случай цилиндрической, круговой или шаровой симметрии).

Рассмотрим линейное одномерное интегральное уравнение Вольтерра:

$$g(x)y(x) + \int_a^x K(x,s)y(s)ds = f(x), a \leq x \leq b, \quad (8.15)$$

где $g(x)$ - заданная функция. Если $g(x) \equiv 0$, то уравнение (8.15) есть уравнение Вольтерры I рода; если $g(x) \neq 0 \forall x \in [a,b]$, то уравнение (8.15) можно записать в виде:

$$y(x) + \int_a^x \frac{K(x,s)}{g(x)} y(s)ds = \frac{f(x)}{g(x)}, a \leq x \leq b, \quad (8.16)$$

т.е. в виде линейного одномерного интегрального уравнения Вольтерры II рода; если же $g(x) = 0$ при некоторых, но не всех значениях $x \in [a,b]$, то (8.15) есть линейное одномерное интегральное уравнение Вольтерра III рода.

Прикладная задача: восстановление сигнала в линейных динамических системах управления с различной «силой» обратной связи (уравнения (4.1), (4.3), (4.5), (4.6), (4.13)-(4.15)).

Нелинейное одномерное интегральное уравнение Вольтерры – Урысона I рода:

$$\int_a^x K[x,s,y(s)]ds = f(x), a \leq x \leq b, \quad (8.17)$$

где $K[x,s,y(s)]$ - заданная нелинейная функция.

Прикладная задача: восстановление сигнала в нелинейной динамической системе управления без обратной связи.

Нелинейное одномерное интегральное уравнение Вольтерры – Урысона II рода:

$$y(x) + \int_a^x K[x, s, y(s)] ds = f(x), a \leq x \leq b, \quad (8.18)$$

где $K[x, s, y(s)]$ - заданная нелинейная функция.

Прикладная задача: восстановление сигнала в нелинейной динамической системе управления с обратной связью.

К линейным интегральным уравнениям Фредгольма I рода также следует отнести *непрерывные интегральные преобразования*. Это – одномерное преобразование Фурье:

$$Y(\omega) = \int_{-\infty}^{\infty} y(t) e^{i\omega t} dt, -\infty < \omega < \infty, \quad (8.19)$$

которое можно записать в виде:

$$\int_{-\infty}^{\infty} e^{i\omega t} y(t) dt = Y(\omega), -\infty < \omega < \infty, \quad (8.20)$$

где $e^{i\omega t}$ можно считать ядром уравнения, $Y(\omega)$ - заданной правой частью, а $y(t)$ - искомой функцией. Тогда соотношение (8.20) можно рассматривать как *линейное одномерное комплексное интегральное уравнение Фредгольма I рода* относительно $y(t)$. Далее, *двухмерное преобразование Фурье*:

$$Y(\omega_1, \omega_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y(t_1, t_2) e^{i(\omega_1 t_1 + \omega_2 t_2)} dt_1 dt_2, -\infty < \omega_1, \omega_2 < \infty, \quad (8.21)$$

или

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i(\omega_1 t_1 + \omega_2 t_2)} y(t_1, t_2) dt_1 dt_2 = Y(\omega_1, \omega_2), -\infty < \omega_1, \omega_2 < \infty. \quad (8.22)$$

Соотношение (8.22) можно рассматривать как *линейное двухмерное комплексное интегральное уравнение Фредгольма I рода*. Имеют место также *косинус-преобразование Фурье*:

$$Y_c(\omega) = \int_{-\infty}^{\infty} y(t) \cos(\omega t) dt, \quad (8.23)$$

синус-преобразование Фурье:

$$Y_s(\omega) = \int_{-\infty}^{\infty} y(t) \sin(\omega t) dt, \quad (8.24)$$

преобразование Хартли:

$$Y_H(\omega) = \int_{-\infty}^{\infty} y(t) \text{cas}(\omega t) dt, \quad (8.25)$$

где $\text{cas}(x) \stackrel{\text{def}}{=} \cos(x) + \sin(x)$, *одностороннее преобразование Лапласа:*

$$\Psi(p) = \int_0^{\infty} y(t) e^{-pt} dt, \quad (8.26)$$

где $p = \lambda + i\sigma$ - комплексная переменная, $y(t)$ - оригинал, $\Psi(p)$ - изображение, а так интегральные преобразования Радона, Меллина, Ганкеля, Бесселя [4] и др., которые так же можно рассматривать как интегральные уравнения Фредгольма I рода относительно $y(t)$.

В уравнениях (8.1) – (8.7), (8.9)-(8.12), (8.15)-(8.18) $y \in Y, f \in F$, где Y и F - некоторые гильбертовы пространства. Обычно $Y = W_2^1$ (пространство Соболева) или $Y = L_2$ (пространство квадратично суммируемых функций), а $F = L_2$.

Определение норм элементов (функций) в этих пространствах:

$$\|y\|_{W_2^1} = \sqrt{\int_a^b y^2(s) ds + \int_a^b y'^2(s) ds}, \quad (8.27)$$

$$\|y\|_{L_2} = \sqrt{\int_a^b y^2(s) ds}, \quad (8.28)$$

$$\|f\|_{L_2} = \sqrt{\int_c^d f^2(x) dx}. \quad (8.29)$$

Следует отметить, что ряд интегральных уравнений решается аналитически. Это – уравнения (8.3)-(8.5), (8.12), (8.14), (8.19)-(8.26). Остальные уравнения решаются, вообще говоря, лишь численно.

Методы решения интегральных уравнений изложены в пункте 3.

Системы линейных алгебраических уравнений

Система линейных алгебраических уравнений (СЛАУ) записывается в виде:

$$Ay = f, \quad (8.30)$$

где A - заданная $m \times n$ матрица элементов a_{ij} , где $i = \overline{1, m}$, $j = \overline{1, n}$ (m - число строк или уравнений, n - число столбцов или неизвестных), y - искомый (неизвестный) вектор-столбец $n \times 1$ элементов y_j , $j = \overline{1, n}$, f - заданная (измеренная) правая часть – вектор-столбец $m \times 1$ элементов f_i , $i = \overline{1, m}$, или

$$\sum_{j=1}^n a_{ij} y_j = f_i, i = \overline{1, m}. \quad (8.31)$$

СЛАУ используется в данном учебном пособии, главным образом, при решении интегральных уравнений методом квадратур и при решении СЛАУ алгоритмом интегральной аппроксимации.

Системы линейно-нелинейных уравнений

Система линейно-нелинейных уравнений (СЛНУ) записывается в виде:

$$\sum_{j=1}^n K(v_i, v'_j) z_j + F = u(v_i), i = \overline{1, m}, \quad (8.32)$$

или более кратко:

$$\sum_{j=1}^n K_{ij} z_j + F = u_i, i = \overline{1, n}, \quad (8.33)$$

где m - число уравнений, n - число линейно входящих неизвестных z_j (а также F) и нелинейно входящих неизвестных v'_j (т.е. всего

$2n+1$ неизвестных), $u_i = u(v_i)$ - заданные (измеренные) правые части, $K_{ij} = K(v_i, v_j)$ - некоторая нелинейная функция.

Прикладные задачи: восстановление дискретного спектра в обратной задаче спектроскопии (уравнение (2.45)), редукция локальных сигналов (уравнение (3.16)).

Операторные уравнения

Все вышеприведенные уравнения (8.1) – (8.26), (8.30) – (8.33) можно записать в виде единого *операторного уравнения*:

$$Ay = F, y \in Y, f \in F, \quad (8.34)$$

где A - заданный оператор (линейный или нелинейный, интегральный, дифференциальный или алгебраический и т. д.), часто называемый *измерительным преобразованием*, f - измеренная правая часть, y - искомое решение, Y и F - некоторые гильбертовы пространства, например, W_2^1, L_2 (см. (1.27) – (5.29)).

Определение. Норма оператора определяется следующим образом:

$$\|A\| = \sup_y \frac{\|Ay\|_F}{\|y\|_Y}. \quad (8.35)$$

Определение. Измерения называются *однократными* (или *одношаговыми*), когда дана одна реализация f , и *многократными* (или *многошаговыми*), когда измерено много реализаций f (см. п. 8.2).

Определения. Задача решения уравнения (8.34) (а значит, и уравнений (8.1) – (8.26), (8.30) – (8.33)) называется *статической*, когда A , u и f не зависят от времени, *динамической*, когда A , u и f зависят от времени, и *квазистатической*, когда A , u и f зависят от номера эксперимента (реализации), но за время эксперимента практически не изменяются.

Определения. Задачу отыскания y в виде: $y = Bf$ будем называть *задачей приведения выхода ко входу*, а оператор B , равный

На практике решение СЛАУ обычно находится не по формуле (8.37), а методами Гаусса, Холецкого и др. [11].

Характеристическое уравнение и типы матриц.

Характеристическое уравнение для квадратной матрицы (при $m = n$) имеет вид:

$$\begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{vmatrix} = 0, \quad (8.40)$$

где $||\cdot||$ есть определитель. Корни уравнения (8.40) называются собственными значениями, или собственными числами λ_i , $i = \overline{1, n}$. Собственные значения произвольной квадратной комплексной матрицы, вообще говоря, комплексны. Величины $1/\lambda_i$, $i = \overline{1, n}$, называются характеристическими числами.

Минор порядка k матрицы A есть определитель k -го порядка, составленный из любой части A с соблюдением расположения элементов a_{ij} . Ранг $r = \text{rang}(A)$ матрицы A - максимальный порядок отличных от нуля миноров. Вводится в рассмотрение также $\rho = \text{rang}(A|f)$ - ранг расширенной матрицы. С использованием понятия ранга можно следующим образом изложить вопрос о решении СЛАУ (8.36). Если $\rho > r$, где $\rho = 3$, а $r = 2$, то СЛАУ не имеет решения и является переопределенной. Если $\rho = r$, то при $\rho = n$ СЛАУ имеет единственное решение, а при $\rho < n$ СЛАУ имеет множество решений и является недоопределенной.

Матрица $A^* = \overline{A}^T = (b_{ij}), i = \overline{1, n}, j = \overline{1, m}$, называется сопряженной с A (или эрмитово сопряженной, или комплексно сопряженной и транспонированной), если $b_{ij} = \overline{a_{ji}}$ (знак $\overline{\cdot}$ означает комплексное

сопряжении, T - транспонирование, а $*$ - эрмитово сопряжение). Квадратная комплексная матрица A называется эрмитовой (самосопряженной), если $A = A^*$ или $a_{ij} = \bar{a}_{ji}$. Если такая матрица вещественна, то она называется симметричной: $A = A^T$ или $a_{ij} = a_{ji}$. У эрмитовой и симметричной матриц все λ_i вещественны (но любого знака). Квадратная матрица A называется положительно определенной, если $\sum_{i,j=1}^n a_{ij} y_i y_j > 0$ при любых вещественных y_i . Примеры положительно определенной матрицы: $A^* A, A A^*, A^T A, A A^T, E$ (где A , вообще говоря, прямоугольна). У положительно определенной матрицы все λ_i вещественны и неотрицательны.

Для прямоугольной $m \times n$ матрицы A вместо собственных значений используются сингулярные числа – это вещественные неотрицательные числа $\mu_i(A) = \sqrt{\lambda_i(A^* A)}, i = \overline{1, n}$, обычно располагаемые в порядке убывания: $\mu_1 \geq \mu_2 \dots \geq \mu_r > \mu_{r+1} = \dots = \mu_n = 0$, где r - ранг матрицы. Если $r < n$, то матрица A есть вырожденная, или особенная матрица; ее определитель $\det(A) \equiv A = 0$ и обратная матрица A^{-1} (при $m = n$) или $(A^* A)^{-1}$ (в общем случае) не существует. Если же $r = n$, то A невырождена, $|A| \sqrt{|A^* A|} = \mu_1 \mu_2 \dots \mu_n$, а обратные матрицы A^{-1} (при $m = n$) и $(A^* A)^{-1}$ (в общем случае) существуют. При этом в случае $m = n$:

$$A^{-1} = \frac{1}{|A|} \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix}, \quad (8.41)$$

где A_{ij} - алгебраические дополнения.

Если A - произвольная (комплексная прямоугольная) матрица, то ее норма $\|A\| = \mu(A)_{\max} = \sqrt{\lambda(A^* A)_{\max}}$, а норма обратной матрицы

$\|A^{-1}\| = 1 / \mu(A)_{\min} = 1 / \sqrt{\lambda(A^* A)_{\min}}$. Если A - эрмитова или симметричная матрица, то $\|A\| = \mu(A)_{\max} = |\lambda(A)|_{\max}$, $\|A^{-1}\| = 1 / \mu(A)_{\min} = 1 / \lambda(A)_{\min}$. Если A - положительно определенная матрица, то $\|A\| = \mu(A)_{\max} = \lambda(A)_{\max}$, $\|A^{-1}\| = 1 / \mu(A)_{\min} = 1 / \lambda(A)_{\min}$.

Нормы векторов и матриц

Нормы комплексных векторов y , f и комплексной квадратной матрицы A определяются так:

$$\|y\| = \sqrt{\sum_{j=1}^n |y_j|^2}, \|f\| = \sqrt{\sum_{i=1}^m |f_i|^2}, \|A\| = \sqrt{\sum_{i,j=1}^n |a_{ij}|^2}. \quad (8.42)$$

Эти нормы называются *эрмитовыми*. Если же y , f и A вещественны, то нормы называются *евклидовыми*.

Число обусловленности

Пусть вместо точных f и A заданы \tilde{f} и \tilde{A} такие, что $\|\tilde{f} - f\| \leq \delta$, $\|\tilde{A} - A\| \leq \varepsilon$, где δ и ε - погрешности задания правой части и матрицы. Тогда относительная погрешность решения СЛАУ (8.36) (по формуле (8.37) или др.) может быть оценена в виде следующего неравенства:

$$\frac{\|\delta y\|}{y} \leq \text{cond}(A) \left(\frac{\delta}{\|f\|} + \frac{\varepsilon}{\|A\|} \right), \quad (8.43)$$

где

$$\text{cond}(A) = \|A\| \cdot \|A\|^{-1} = \mu(A)_{\max} / \mu(A)_{\min} \geq 1 \quad (8.44)$$

- число обусловленности матрицы A .

Если $\text{cond}(A)$ относительно мало (обычно $\leq 10^3$), то матрица A (и СЛАУ) называется *хорошо обусловленной*. Если же $\text{cond}(A)$ относительно велико (обычно $\geq 10^4$), то матрица A (и СЛАУ) называется *плохо*

обусловленной. Заметим, что малость (по сравнению с единицей) определителя $|A|$, вообще говоря, не есть критерий плохой обусловленности.

Умножение матриц и векторов

Умножение двух прямоугольных матриц:

$$Z = X \times Y \quad (8.45)$$

$m \times l$ $m \times n$ $n \times l$

или подробнее

$$z_{ik} = \sum_{j=1}^n x_{ij} y_{jk}, i = \overline{1, m}, k = \overline{1, l} \quad (8.46)$$

Умножение матрицы на вектор:

$$f = A \times y \quad (8.47)$$

$m \times l$ $m \times n$ $n \times 1$

или подробнее:

$$f_i = \sum_{j=1}^n a_{ij} y_j, j = \overline{1, m}. \quad (8.48)$$

Пример 1. Рассмотрим систему двух уравнений с двумя неизвестными:

$$\left. \begin{aligned} y_1 + 2y_2 &= 3, \\ -y_1 + 3y_2 &= 7 \end{aligned} \right\} \quad (8.49)$$

Ее матрица

$$A = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} \quad (8.50)$$

является квадратной, поэтому можно рассматривать следующее характеристическое уравнение:

$$\begin{vmatrix} 1 - \lambda & 2 \\ -1 & 3 - \lambda \end{vmatrix} = 0, \quad (8.51)$$

откуда $\lambda^2 - 4\lambda + 5 = 0$ и корни равны $\lambda_{1,2} = 2 \pm i$, т.е. собственные значения комплексные, что возможно для матрицы (8.50), не являющейся симметричной или положительно определенной.

Определитель

$$|A| = \begin{vmatrix} 1 & 2 \\ -1 & 3 \end{vmatrix} = 5 \quad (8.52)$$

Обратная матрица, в соответствии с (8.41), равна

$$A^{-1} = \frac{1}{5} \begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 3/5 & -2/5 \\ 1/5 & 1/5 \end{pmatrix}. \quad (8.53)$$

Проверка с использованием правила (8.46) умножения матриц:

$$A^{-1}A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (8.54)$$

Решение СЛАУ (8.49) согласно (8.37) с использованием (8.48) равно

$$y = A^{-1}f = \begin{pmatrix} -1 \\ 2 \end{pmatrix} \quad (8.55)$$

или по правилу Крамера

$$y_1 = \frac{\begin{vmatrix} 3 & 2 \\ 7 & 3 \end{vmatrix}}{5} = -1, y_2 = \frac{\begin{vmatrix} 1 & 3 \\ -1 & 7 \end{vmatrix}}{5} = 2 \quad (8.56)$$

Пример 2. Рассмотрим прямоугольную матрицу 3×2 :

$$A = \begin{pmatrix} 2 & -3 \\ -1 & 2 \\ 1 & 4 \end{pmatrix}. \quad (8.57)$$

Эрмитово сопряженная матрица A^* равна

$$A^* = \begin{pmatrix} 2 & -1 & 1 \\ -3 & 2 & 4 \end{pmatrix}. \quad (8.58)$$

Произведение этих матриц дает положительно определенную квадратную матрицу 2×2 :

$$B = A^* A = \begin{pmatrix} 6 & -4 \\ -4 & 29 \end{pmatrix} \quad (8.59)$$

Характеристическое уравнение для матрицы B :

$$\begin{vmatrix} 6 - \lambda & -4 \\ -4 & 29 - \lambda \end{vmatrix} = 0 \quad (8.60)$$

откуда $\lambda^2 - 35\lambda + 158 = 0$ и корни (собственные значения матрицы B) равны $\lambda_1(A^* A) \cong 29.675, \lambda_2(A^* A) \cong 5.325$, т.е. корни вещественны и неотрицательны, как и должно быть для положительно определенной матрицы $A^* A$. Сингулярные числа матрицы A равны $\mu_1(A) = \sqrt{\lambda_1(A^* A)} \cong 5.47 = \mu(A)_{\max} = \|A\|, \mu_2(A) = \sqrt{\lambda_2(A^* A)} \cong 2.32 = \mu(A)_{\min} = \|A^{-1}\|^{-1}$. Число обусловленности матрицы A равно $cond(A) = \mu(A)_{\max} / \mu(A)_{\min} \cong \cong 2.36$, т.е. матрица (8.57) является хорошо обусловленной.

Элементы теории вероятностей

Дадим несколько определений.

Определение 1. Действительная переменная, которая в зависимости от исхода опыта, т.е. в зависимости от случая, принимает различные значения, называется *случайной величиной*.

Пусть X - некоторая случайная величина. Тогда имеет место

Определение 2. *Функцией распределения* $F(x)$ случайной величины X называется функция

$$F(x) = P(X < x),$$

где $P(A) \in [0, 1]$ - вероятность события A .

Случайные величины могут быть дискретными и непрерывными. Рассмотрим лишь непрерывные случайные величины.

Определение 3. Случайная величина называется непрерывной, если функция распределения может быть представлена в виде:

$$F(x) = \int_{-\infty}^x f(t) dt,$$

где $f(x)$ - плотность распределения.

При этом

$$f(x) = dF(x) / dx, F(\infty) = \int_{-\infty}^{\infty} f(x) dx = 1,$$

$$P(a \leq X \leq b) = F(b) - F(a) = \int_a^b f(x) dx, P(X = x_0) = 0$$

Определение 4. Равномерное распределение случайной величины x - это распределение с плотностью

$$f(x) = \begin{cases} 1 / (b - a) = const, x \in [a, b], \\ 0, x \notin [a, b] \end{cases} \quad (8.61)$$

В случае равномерного распределения *математическое ожидание* (МО, или первый центральный момент) случайной величины x равно

$$m_0 \equiv MX = \int_{-\infty}^{\infty} xf(x) dx = \int_a^b \frac{x}{b-a} dx = \frac{a+b}{2}, \quad (8.62)$$

а *дисперсия* (второй центральный момент) равна

$$\sigma^2 \equiv DX = \int_{-\infty}^{\infty} (x - m_0)^2 f(x) dx = \int_a^b \left(x - \frac{a+b}{2} \right)^2 \frac{dx}{b-a} = \frac{(b-a)^2}{12}, \quad (8.63)$$

откуда среднеквадратическое отклонение (СКО) равно

$$\sigma = \frac{b-a}{2\sqrt{3}} \approx 0.577 \cdot \frac{b-a}{2} \text{ (рис 5).}$$

Определение 5. Нормальное распределение (распределение Гаусса) случайной величины x - это распределение с плотностью

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}} \quad (8.64)$$

(см. рис 6).

Закону (8.64) соответствует МО

$$m_0 = \int_{-\infty}^{\infty} \frac{x}{\sqrt{2\pi\sigma}} e^{-\frac{(x-a)^2}{2\sigma^2}} dx = a. \quad (8.65)$$

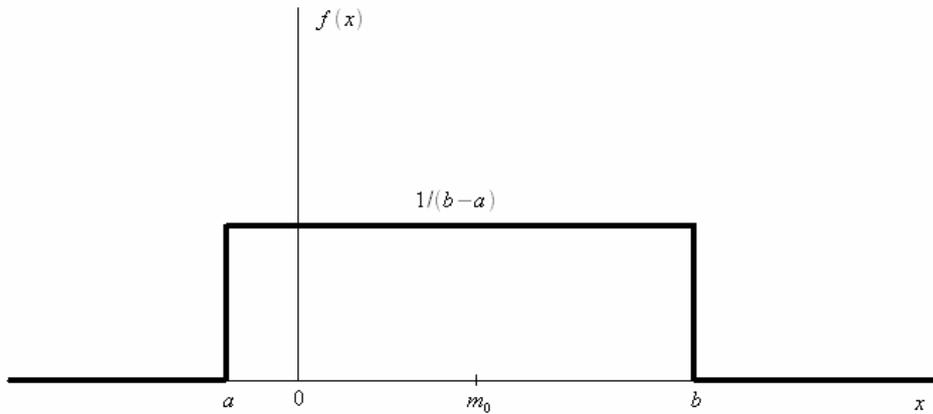


Рис. 5. Равномерное распределение

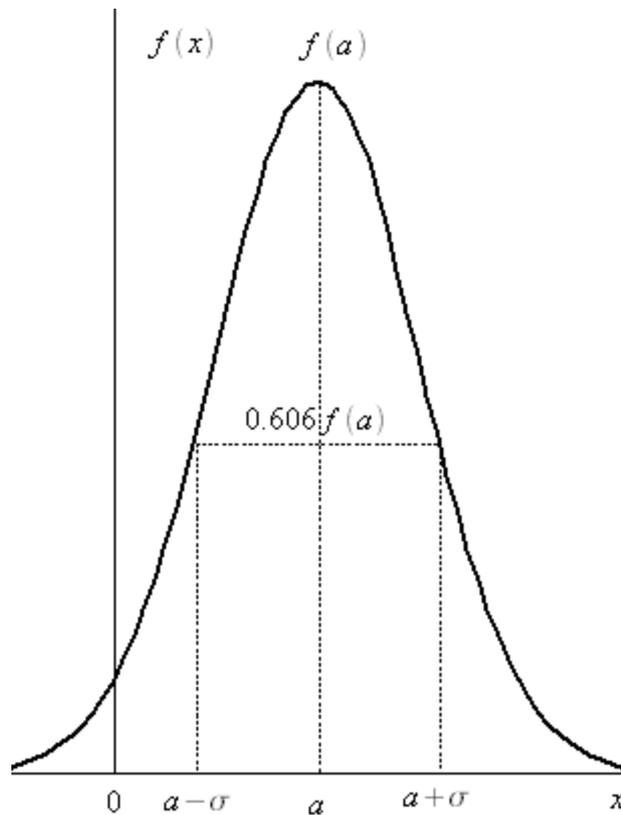


Рис. 6. Нормальное распределение

И дисперсия

$$\sigma^2 = \int_{-\infty}^{\infty} \frac{(x-a)^2}{\sqrt{2\pi\sigma}} e^{-\frac{(x-a)^2}{2\sigma^2}} dx. \quad (8.66)$$

Имеем (см. рис. 6): $f(a + \sigma) \approx 0.606 f(a)$.

Если случайная величина X имеет нормальное распределение с параметрами a и σ , то говорят, что X распределена нормально согласно закону $N(x; a, \sigma)$ или $N(a, \sigma)$ и пишут: $X \in N(x; a, \sigma)$ или $X \in N(a, \sigma)$.

Функция $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, т.е. функция $f(x)$ (см. (8.64)) при $a = 0, \sigma = 1$ называется *плотностью нормированного центрированного нормального распределения*. Плотность $\varphi(x)$ и соответствующая ей функция распределения $\Phi(x) = \int_{-\infty}^x \varphi(t) dt = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$ затабулированы.

Функцию $\Phi(x)$ часто называют *гауссовым интегралом ошибок*. Иногда задается функция $\Phi_0(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$, причем $\Phi(x)$ может быть выражена через $\Phi_0(x)$ посредством соотношения

$$\Phi(x) = \begin{cases} \Phi_0(x) + 0.5, & x \geq 0, \\ -\Phi_0(|x|) + 0.5, & x \leq 0 \end{cases} \quad (8.67)$$

или

$$\Phi(x) = \text{sign}(x) \cdot \Phi_0(|x|) + 0.5 \quad (8.68)$$

Связь между $f(x)$ и $\varphi(x)$, а так же $F(x)$ и $\Phi(x)$ имеет вид:

$$f(x) = \frac{1}{\sigma} \varphi\left(\frac{x-a}{\sigma}\right),$$

$$F(x) = \Phi\left(\frac{x-a}{\sigma}\right) = \text{sign}(x-a) \cdot \Phi_0\left(\frac{|x-a|}{\sigma}\right) + 0.5.$$

Приложение 2. Обобщенные функции, формула Эйлера и интегральные преобразования.

Элементы теории обобщенных функций

Определение обобщенной функции. В современном математическом аппарате широко используются так называемые обобщенные функции [3]. Необходимость их введения была связана с тем, что в классическом математическом анализе постоянно нужно делать оговорки о свойствах функций: при дифференцировании функции нужно требовать, чтобы она была $n \geq 1$ раз дифференцируема, при интегрировании – чтобы она была кусочно-непрерывна и суммируема, и т.д. Вместе с тем очень часто реальные функции (в первую очередь, полученные в результате измерений) такими свойствами не обладают. Поэтому возникла необходимость распространить математический анализ (дифференциальное и интегральное исчисления) на такие функции.

Определение. Обобщенной функцией называется функция, для дифференцирования и интегрирования которой не требуется выполнения классических свойств дифференцируемости, кусочной непрерывности, суммируемости, интегрируемости, и т.д.

Функция Хэвисайда. Исторически первой обобщенной функцией является функция Хэвисайда (или ступенчатая функция) [3]:

$$H(x) = \begin{cases} 0, & x < 0, \\ 1, & x > 0. \end{cases} \quad (9.1)$$

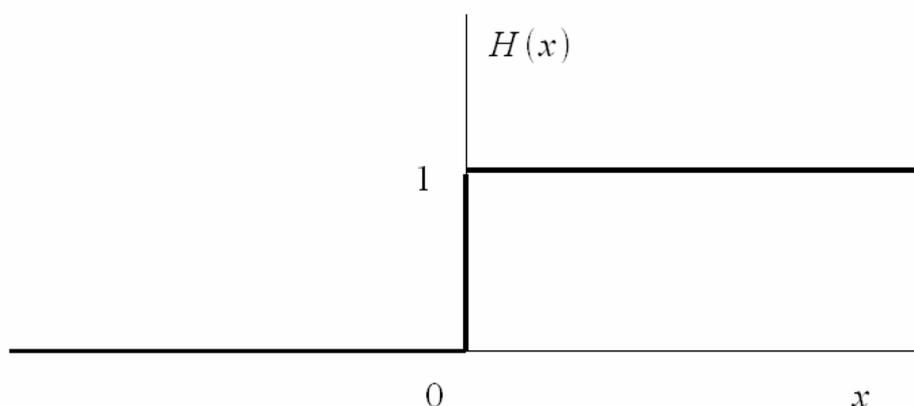


Рис. 7. Функция Хевисайда

δ -функция Дирака. Исторически следующая обобщенная функция – это δ -функция Дирака (или симметричная единичная импульсная функция) [3]. Одномерная δ -функция Дирака определяется следующими тремя пунктами:

1)

$$\delta(x) = \begin{cases} \infty, & x = 0, \\ 0, & x \neq 0, \end{cases} \quad (9.2)$$

2)

$$\int_{-\infty}^{\infty} \delta(x) dx = 1, \quad (9.3)$$

3)

$$\int_{-\infty}^{\infty} f(x) \delta(x - x_0) dx = f(x_0), \quad (9.4)$$

где $f(x)$ - некоторая (классическая) функция, а x_0 - некоторое значение абсциссы x .

Пункту 1) (см. (9.2)) соответствует рис. 8а, а пункту 3) (см. (9.4)) – рис. 8б.

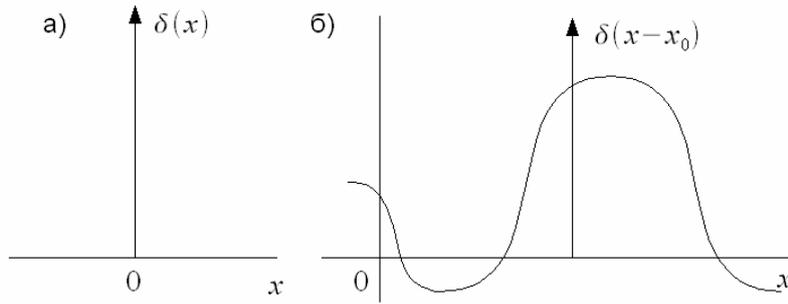


Рис. 8. Иллюстрации к свойствам δ -функции

Из (9.3) и (9.4) следуют, в частности, следующие соотношения:

$$\int_{-\infty}^{\infty} \delta(ax - x_0) dx = \frac{1}{|a|}, \quad (9.5)$$

$$\int_{-\infty}^{\infty} f(x) \delta(ax - x_0) dx = \frac{1}{|a|} f\left(\frac{x_0}{a}\right) \quad (9.6)$$

Функция Хэвисайда и δ -функция Дирака связаны соотношениями:

$$\delta(x) = H'(x) \quad (9.7)$$

$$H(x) = \int_{-\infty}^x \delta(t) dt \quad (9.8)$$

Видим, что эти функции можно дифференцировать и интегрировать (но не в классическом смысле).

Используется также двухмерная δ -функция Дирака:

$$\delta(x, y) = \begin{cases} \infty, & x = y = 0 \\ 0, & \text{иначе} \end{cases} \quad (9.9)$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(x, y) dx dy = 1 \quad (9.10)$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(x - x_0, y - y_0) dx dy = f(x_0, y_0) \quad (9.11)$$

Из (9.10) и (9.11) следуют следующие соотношения:

$$\int \int_{-\infty}^{\infty} \delta(ax - x_0, by - y_0) dx dy = \frac{1}{|ab|} \quad (9.12)$$

$$\int \int_{-\infty}^{\infty} f(x, y) \delta(ax - x_0, by - y_0) dx dy = \frac{1}{|ab|} f\left(\frac{x_0}{a}, \frac{y_0}{b}\right) \quad (9.13)$$

Двухмерную δ -функцию можно представить как произведение одномерных:

$$\delta(x, y) = \delta(x)\delta(y). \quad (9.14)$$

Формула Эйлера

Определение формулы Эйлера. Формула Эйлера имеет вид:

$$Ae^{i\varphi} = A(\cos \varphi + i \sin \varphi). \quad (9.15)$$

Ее можно проиллюстрировать графически. Пусть задана комплексная плоскость $xу$ и на ней окружность радиуса A (рис. 9). Пусть точка (x, y) на этой окружности задается полярным углом (аргументом) φ . Тогда $x = A \cos \varphi, y = A \sin \varphi$, а комплексное число z , соответствующее точке (x, y) , будет равно $z = A(\cos \varphi + i \sin \varphi)$. Итак, формула (9.15) дает связь между показательным и тригонометрическим представлениями комплексного числа (в дополнение к алгебраическому представлению $z = x + iy$).

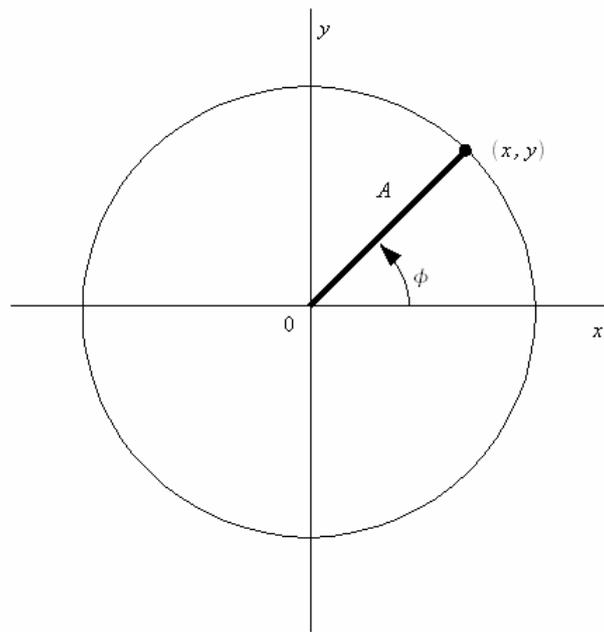


Рис 9. Графическая иллюстрация формулы Эйлера

Формула Эйлера записывается также в виде:

$$e^{i\varphi} = \cos \varphi + i \sin \varphi. \quad (9.16)$$

Она находит широкое применение в различных областях математики. В частности, нам далее потребуется выражение интеграла от формулы Эйлера. Покажем, что он связан с δ -функцией.

Формула Эйлера и δ -функция. Итак, нужно найти интеграл (интеграл Фурье):

$$I(\omega) = \int_{-\infty}^{\infty} e^{i\omega t} dt. \quad (9.17)$$

Используя (9.16), запишем:

$$\int_{-\infty}^{\infty} e^{i\omega t} dt = \int_{-\infty}^{\infty} \cos \omega t dt + i \int_{-\infty}^{\infty} \sin \omega t dt. \quad (9.18)$$

Учтем, что если некоторая функция $f(x)$ является четной, т.е.

$f(-x) = f(x)$, то $\int_{-\infty}^{\infty} f(x) dx = 2 \int_0^{\infty} f(x) dx$, а если $f(x)$ является нечетной

функцией: $f(-x) = -f(x)$, то $\int_{-\infty}^{\infty} f(x) dx = 0$. В правой части (9.18) первый интеграл – это интеграл от четной функции, а второй – от нечетной функции. Поэтому:

$$\int_{-\infty}^{\infty} e^{i\omega t} dt = 2 \int_0^{\infty} \cos \omega t dt. \quad (9.19)$$

При $\omega=0$ интеграл $\int_0^{\infty} \cos \omega t dt = \infty$, а при $\omega \neq 0$

$$\int_0^{\infty} \cos \omega t dt = \frac{1}{\omega} \lim_{t \rightarrow \infty} \sin \omega t. \quad (9.20)$$

Однако, строго говоря, предел $\lim_{t \rightarrow \infty} \sin \omega t$ не существует. Но $\sin \omega t$ - это функция, осциллирующая вокруг нуля, и ее среднее значение (математическое ожидание) равно нулю. Поэтому если отойти от классического математического анализа, то можно считать, что

$$\int_{-\infty}^{\infty} e^{i\omega t} dt = \begin{cases} \infty, & \omega = 0, \\ 0, & \omega \neq 0. \end{cases} \quad (9.21)$$

т.е. интеграл $\int_{-\infty}^{\infty} e^{i\omega t} dt$ пропорционален δ -функции $\delta(\omega)$. Более точное обоснование показывает, что:

$$\int_{-\infty}^{\infty} e^{i\omega t} dt = \int_{-\infty}^{\infty} \cos \omega t dt = 2\pi \delta(\omega). \quad (9.22)$$

Из (9.22) следует интегральное представление δ -функции (ее выражение через интеграл Фурье): $\delta(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega t} dt$.

Аналогичная формула имеет место в двухмерном случае:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i(\omega_1 t_1 + \omega_2 t_2)} dt_1 dt_2 = 4\pi^2 \delta(\omega_1, \omega_2). \quad (9.23)$$

Полученные формулы (9.22) и (9.23) часто используются в математике (см., например, п. 3).

Интегральные преобразования

Рассмотрим интегральные, или непрерывные (а так же дискретные и быстрые) преобразования (8.19)-(8.26), которые находят широкое применение в математике и прикладных задачах.

Непрерывное преобразование Фурье. Пусть задана некоторая (кусочно-непрерывная) функция (исходный процесс) $y(t)$, $-\infty < t < \infty$, где t - время, линейная координата, угловая координата, и.т.д. Если t - время, то $y(t)$ - временной процесс. Тогда интеграл:

$$Y(\omega) = \int_{-\infty}^{\infty} y(t)e^{i\omega t} dt, -\infty < \omega < \infty \quad (9.24)$$

называется (одномерным) *прямым непрерывным преобразованием Фурье (НПФ) или преобразованием Фурье (ПФ), Фурье-образом, спектральной функцией, спектром, изображением по Фурье*, и т.д. А функция $y(t)$ в этом случае называется *обратным преобразованием Фурье (ОПФ) или оригиналом*. Переменная ω называется частотой Фурье. Функция $y(t)$ может быть вещественной или комплексной, а $Y(\omega)$ - вообще говоря, комплексная, поскольку $Y(\omega)$ может быть записана в виде (на основании формулы Эйлера (9.16)):

$$Y(\omega) = \int_{-\infty}^{\infty} y(t) \cos \omega t dt + i \int_{-\infty}^{\infty} y(t) \sin \omega t dt \quad (9.25)$$

или

$$Y(\omega) = \operatorname{Re} Y(\omega) + i \operatorname{Im} Y(\omega), \quad (9.26)$$

где

$$\operatorname{Re} Y(\omega) = \int_{-\infty}^{\infty} y(t) \cos \omega t dt \quad (9.27)$$

$$\operatorname{Im} Y(\omega) = \int_{-\infty}^{\infty} y(t) \sin \omega t dt. \quad (9.28)$$

Соотношение (9.27) называется *косинус - преобразованием* Фурье, а (9.28) – *синус – преобразованием* Фурье. Используют также квадрат модуля преобразования Фурье:

$$|Y(\omega)|^2 = \operatorname{Re}^2 Y(\omega) + \operatorname{Im}^2 Y(\omega). \quad (9.29)$$

Связь между прямым и обратным преобразованиями Фурье.

Запишем (9.24) в виде:

$$Y(\omega) = \int_{-\infty}^{\infty} y(t') e^{i\omega t'} dt'. \quad (9.30)$$

Умножим (9.25) на $e^{-i\omega t}$, проинтегрируем по ω от $-\infty$ до ∞ и изменим порядок интегрирования. Получим:

$$\int_{-\infty}^{\infty} Y(\omega) e^{-i\omega t} d\omega = \int_{-\infty}^{\infty} y(t') \left[\int_{-\infty}^{\infty} e^{i(t'-t)\omega} d\omega \right] dt'. \quad (9.31)$$

Из формулы (9.22) следует, что

$$\int_{-\infty}^{\infty} e^{i(t'-t)\omega} d\omega = 2\pi \delta(t'-t) \quad (9.32)$$

и тогда, используя (9.4), получим:

$$\int_{-\infty}^{\infty} Y(\omega) e^{-i\omega t} d\omega = 2\pi \int_{-\infty}^{\infty} y(t') \delta(t'-t) dt' = 2\pi y(t), \quad (9.33)$$

откуда окончательно

$$y(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} Y(\omega) e^{-i\omega t} d\omega. \quad (9.34)$$

Итак, исходя из прямого преобразования Фурье (9.24), мы получили выражение для *обратного преобразования Фурье* $y(t)$ и это можно рассматривать как результат аналитического решения интегрального уравнения Фредгольма I рода (9.24).

Отметим при этом, что часто в (9.24) ставят знак «-» перед i . В этом случае изменится знак перед i в (9.34). Кроме того, в (9.24) перед интегралом нередко ставят множитель $1/\sqrt{2\pi}$. В этом случае такой же множитель будет перед интегралом в (9.34). Другими словами, (9.24) и (9.34) можно записать в виде:

$$Y(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} y(t) e^{\pm i\omega t} dt, \quad (9.35)$$

$$y(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} Y(\omega) e^{\pm i\omega t} d\omega. \quad (9.36)$$

Однако различие в записях не ведет к различию вычисляемых $Y(\omega)$ и $y(t)$. Поэтому можно пользоваться любой записью. В данном учебном пособии используются соотношения (9.24) и (9.34).

Помимо комплексного преобразования Фурье используются также:

косинус – преобразование Фурье (прямое и обратное):

$$Y_c(\omega) = \int_{-\infty}^{\infty} y(t) \cos(\omega t) dt, \quad (9.37)$$

$$y(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} Y_c(\omega) \cos(\omega t) d\omega \quad (9.38)$$

и *синус – преобразование Фурье* (прямое и обратное):

$$Y_s(\omega) = \int_{-\infty}^{\infty} y(t) \sin(\omega t) dt, \quad (9.39)$$

$$y(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} Y_s(\omega) \sin(\omega t) d\omega. \quad (9.40)$$

Двухмерное преобразование Фурье. Пусть задана двухмерная функция $y(t_1, t_2)$. Тогда двойной интеграл

$$Y(\omega_1, \omega_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y(t_1, t_2) e^{i(\omega_1 t_1 + \omega_2 t_2)} dt_1 dt_2 \quad (9.41)$$

есть двумерное прямое непрерывное преобразование Фурье, а формула обращения, аналогичная (9.34):

$$y(t_1, t_2) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} Y(\omega_1, \omega_2) e^{-i(\omega_1 t_1 + \omega_2 t_2)} d\omega_1 d\omega_2, \quad (9.42)$$

дает двумерное обратное непрерывное преобразование Фурье.

Преобразование Фурье (одномерное и двумерное) широко используется для спектрального анализа временных процессов, для решения интегральных уравнений Вольтерры и Фредгольма I, II и III рода типа свертки одномерных и двумерных, для решения дифференциальных уравнений и т.д.

Дискретное преобразование Фурье. На практике значения функций y и Y задаются и/или вычисляются не непрерывно, а на дискретных сетках узлов, и интегралы в вышеприведенных соотношениях заменяются конечными суммами, т.е. непрерывное преобразование Фурье реализуется в виде дискретного преобразования Фурье. Рассмотрим это на примере вычисления преобразования Фурье согласно (9.24).

Пусть снимаются дискретные отсчеты $y(t)$ на равномерной сетке узлов:

$$t_k = kh, k = \overline{0, N-1}, \quad (9.43)$$

где $h = \Delta t = const$ - интервал (шаг) дискретизации по t , а N - число отсчетов.

Справедлива **теорема Котельникова** [3], в силу которой: линейная частота дискретизации (максимальная линейная частота) в ДПФ равна

$$1) f_g \equiv f_{\max} = \frac{1}{h} \quad (9.44)$$

(заметим, что ω - это круговая частота, связанная с линейной частотой f соотношением: $\omega = 2\pi f$);

интервал (шаг) дискретизации по частоте равен

$$2) \Delta f = \frac{f_{\max}}{N} = \frac{1}{h \cdot N} = \frac{1}{t_{\max}}, \quad (9.45)$$

где $t_g \equiv t_{\max} = t_N = h \cdot N$ - длина выборки;

$$3) \omega_l = 2\pi f_l, l = \overline{0, N-1}, f_l = l \cdot \Delta f = \frac{l}{h \cdot N}. \quad (9.46)$$

Интеграл в (9.24) заменяем конечной суммой по формуле левых прямоугольников с шагом h (рис. 10) и в результате вместо непрерывного соотношения (9.24) получим дискретное:

$$Y_l \equiv Y(\omega_l) = h \sum_{k=0}^{N-1} y_k e^{i2\pi f_l k h}, \quad (9.47)$$

откуда

$$Y_l = h \sum_{k=0}^{N-1} y_k e^{i2\pi l k / N}, l = \overline{0, N-1}, \quad (9.48)$$

где $y_k = y(t_k)$.

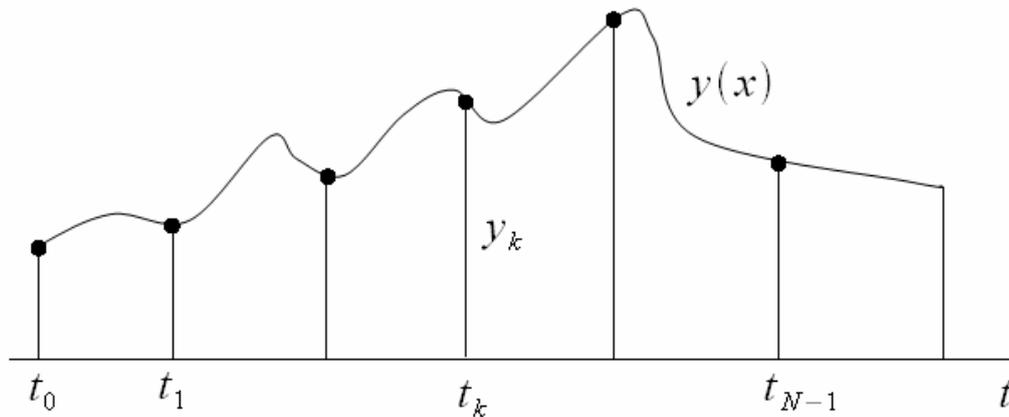


Рис. 10. Дискретизация функции

Получили формулу (9.48) для *дискретного преобразования Фурье* в случае, когда шаги дискретизации по t и ω постоянны и число отсчетов N по t и ω одинаково.

Аналогично обратному непрерывному преобразованию Фурье (9.34) будет соответствовать *обратное дискретное преобразование Фурье*:

$$y_k \equiv y(t_k) = \frac{1}{h \cdot N} \sum_{l=0}^{N-1} Y_l e^{-i2\pi lk/N}, k = \overline{0, N-1}. \quad (9.49)$$

Исследуем формулу (9.48) и сравним ее с формулой (9.24). Результаты будем формулировать в виде выводов.

Вывод 1. В формуле (9.48) не фигурируют значения узлов ω_l и t_k , а фигурируют лишь номера узлов l и k . Это значительно упрощает и убыстряет вычисления, а в конечном итоге позволило создать алгоритм быстрого преобразования Фурье (БПФ).

Далее, в (9.47) к f_l добавим $n \geq 0$ значений $f_g = 1/h$, получим:

$$e^{i2\pi(f_l+n/h)kh} = e^{i2\pi f_l kh} e^{i2\pi nk} = e^{i2\pi lk/N}, n = 0, 1, 2, \dots \quad (9.50)$$

Вывод 2. Y_l - периодическая функция с периодом $f_g = 1/h$.

Вывод 2 порождает следующие следствия. Пусть $Y(\omega)$ или $Y(f)$ - финитная функция, равная нулю при $|\omega| > \omega_B$ или $|f| > f_B$, где f_B - верхняя частота (см. рис. 11, где под $Y(f)$ подразумевается или $\text{Re}Y(f)$, или $|Y(f)|$ и т.д.).

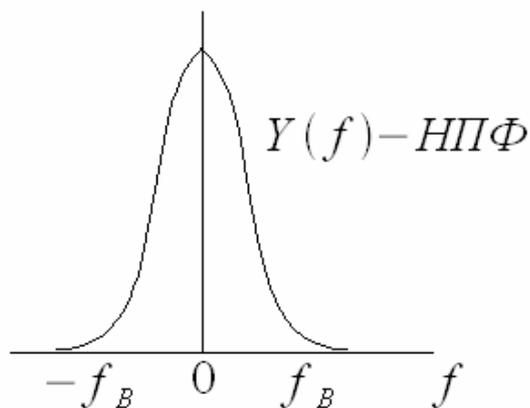


Рис. 11. Непрерывное преобразование Фурье

Тогда при $f_g > 2f_B$ или $h < 1/2f_B$ получим вместо непрерывного преобразования Фурье (см. рис. 11) дискретное преобразование Фурье согласно (9.48) – Рисунок 12, где непрерывная линия соответствует значениям $l \in [0, N-1]$, а пунктир – значениям $l \notin [0, N-1]$. Видим, что дискретное преобразование Фурье заметно отличается от непрерывного преобразования Фурье.

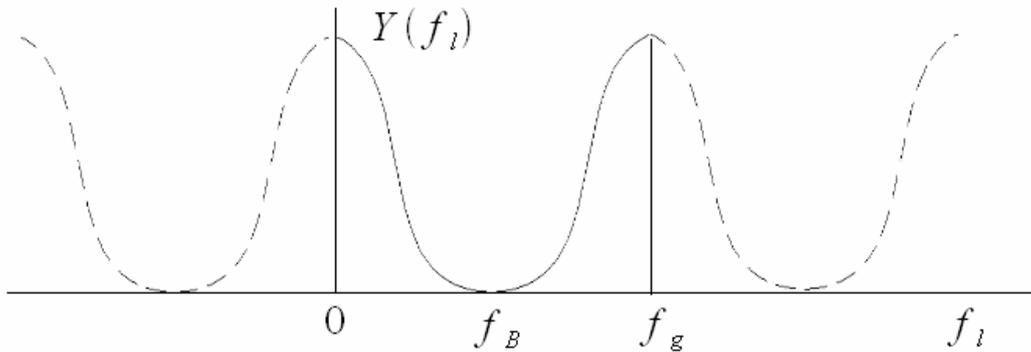


Рис. 12. Дискретное преобразование Фурье ($h < 1/2f_B$)

При $f_g = 2f_B$ или $h = 1/2f_B$ имеем (рис. 13):

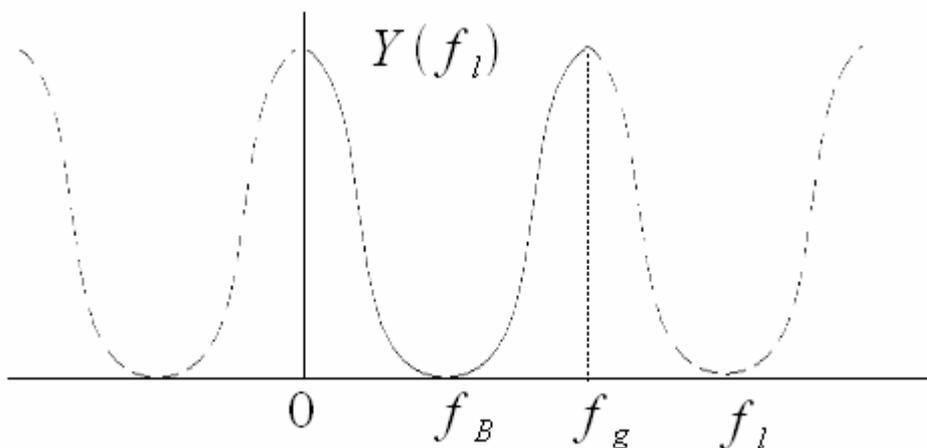


Рис. 13. Дискретное преобразование Фурье ($h = 1/2f_B$)

А при $f_g < 2f_B$ или $h > 1/2f_B$ имеем так называемый *эффект наложения*, искажающий спектр (рис. 14):

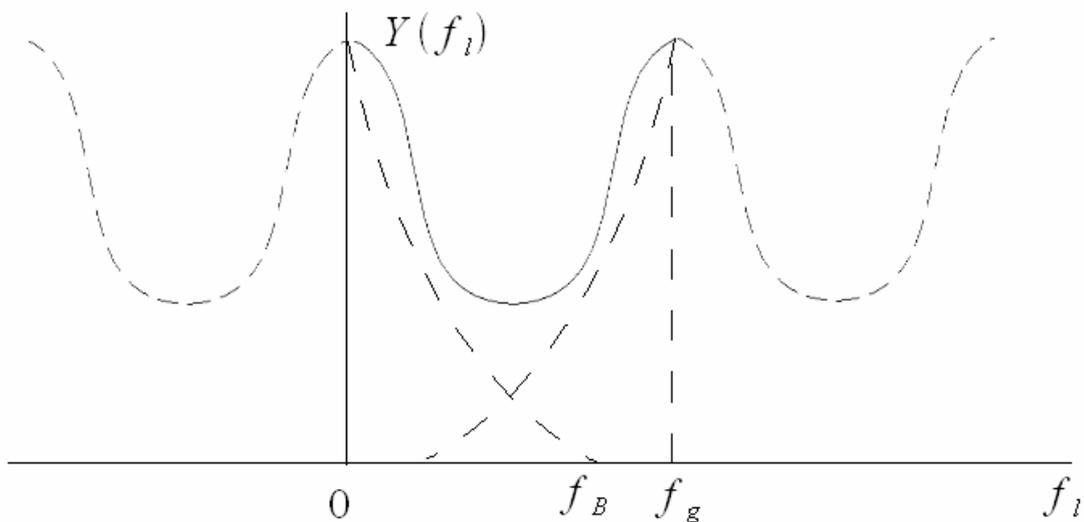


Рис. 14. Дискретное преобразование Фурье ($h > 1/2f_B$)

Поскольку при $h > 1/2f_B$ имеет место эффект наложения, то шаг h дискретизации по t нужно брать в соответствии с неравенством: $h \leq \frac{1}{2f_B}$, оценив при этом каким-то образом f_B .

Если добавляются отсчеты t_k изнутри области $[0, t_g)$, т.е. уменьшается h при неизменном t_g , то f_g увеличивается (см. (9.44)) и эффект наложения снижается, однако шаг дискретизации по частоте $\Delta f = 1/t_g$ не изменяется и, следовательно, не изменяется разрешение по частоте f .

Итак, основные свойства дискретного преобразования Фурье:

- периодичность с периодом f_g ;
- эффект наложения (при $f_g < 2f_B$);
- изменение эффекта наложения и разрешения по f при добавлении новых отсчетов по t .

Об алгоритмах быстрого преобразования Фурье. На практике дискретное преобразование Фурье обычно реализуется в виде алгоритма быстрого преобразования Фурье (Fast Fourier Transform – FFT). Впервые

такой алгоритм предложили Кули и Тьюки; в этом алгоритме число отсчетов по t и f равно $N = 2^m$, где $m \in \mathbb{Z}$, например, $N = 1024$; если обычное дискретное преобразование Фурье требует N^2 операций, то быстрое преобразование Фурье требует $N \log_2 N$ операций (если $N = 1024$, то соответственно $\approx 10^6$ и $\approx 10^4$ операций – разница существенная).

К настоящему времени разработано много *стандартных программ* для быстрого преобразования Фурье.

Что касается *двухмерного непрерывного преобразования Фурье*, то оно на практике реализуется в виде двухмерного дискретного преобразования Фурье, которое может быть записано как набор одномерных дискретных преобразований Фурье. Например, двухмерное непрерывное преобразование Фурье (9.41) может быть записано в виде:

$$Y(\omega_1, \omega_2) = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} y(t_1, t_2) e^{i\omega_1 t_1} dt_1 \right] e^{i\omega_2 t_2} dt_2,$$

т.е. в виде одномерного непрерывного преобразования Фурье от одномерного непрерывного преобразования Фурье. А двухмерное дискретное преобразование Фурье можно записать в виде (см. (9.48)):

$$Y_{lj} = h_1 h_2 \sum_{k=0}^{N-1} \sum_{m=0}^{M-1} y_{km} e^{i2\pi(1k/N + jm/M)} = h_2 \sum_{m=0}^{M-1} \left(h_1 \sum_{k=0}^{N-1} e^{i2\pi k/N} \right) e^{i2\pi jm/M}, \quad (9.51)$$

т.е. в виде одномерного дискретного преобразования Фурье от одномерного дискретного преобразования Фурье.

Использование регуляризации. Задача вычисления преобразования Фурье является, строго говоря, некорректной (неустойчивой), так как связана с решением интегрального уравнения Фредгольма I рода, хотя «степень некорректности» в значительной степени снижается благодаря тому, что это уравнение решается аналитически.

Для повышения устойчивости можно использовать метод регуляризации Тихонова, в результате чего вместо формулы (9.24) нужно будет использовать формулу

$$Y_{\alpha}(\omega) = \int_{-\infty}^{\infty} \frac{y(t)}{1 + \alpha t^{2n}} e^{i\omega t} dt, \quad (9.52)$$

где $\alpha > 0$ - параметр регуляризации.

Преобразование Лапласа. Лежащее в основе оперативного метода *одностороннее прямое преобразование Лапласа* есть интегральное преобразование вида [4]:

$$\Psi(p) = \int_0^{\infty} \varphi(x) e^{-px} dx, \quad (9.53)$$

где $p = \lambda + i\sigma$ – комплексная переменная, $\varphi(x)$ – функция действительной переменной x (обычно времени), называемая *оригиналом*; $\psi(p)$ – *изображение* функции $\varphi(x)$, что часто записывается как $\varphi(x) \rightarrow \psi(p)$ или $\psi(p) \rightarrow \varphi(x)$, или $\psi(p) = L[\varphi(x)]$.

Двустороннее преобразование Лапласа отличается от (9.53) нижним пределом интегрирования, равным $-\infty$. Обычно под термином «преобразование Лапласа» подразумевается одностороннее преобразование (9.53).

Оригинал $\varphi(x)$ должен удовлетворять следующим условиям:

- $\varphi(x)$ - кусочно-непрерывная функция,
- $\varphi(x) = 0$ при $x < 0$,
- $|\varphi(x)| < M e^{cx}$ при $x > 0$, где $M > 0$ и $c \geq 0$ – некоторые константы, причем если $|\varphi(x)| \leq |\varphi(0)|$, то $c = 0$.

Тогда справедливо *обратное преобразование Лапласа*, дающее выражение для оригинала:

$$\varphi(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \psi(p) e^{px} dp. \quad (9.54)$$

Пример. Пусть задано изображение $L[\varphi(x) = \psi(p) = 1/p]$, а в отношении оригинала $\varphi(x)$ известно, что $c = 0$. Тогда $p = i\sigma$, $\lambda = 0$ и, в соответствии с (9.54),

$$\begin{aligned}\varphi(x) &= \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} \frac{1}{p} e^{px} dp = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{1}{\sigma} e^{i\sigma x} d\sigma = \\ &= \frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{\cos \sigma x}{\sigma} d\sigma + \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\sin \sigma x}{\sigma} d\sigma.\end{aligned}\tag{9.55}$$

Первый интеграл в (9.55) равен нулю, так как подынтегральная функция нечетна, а второй (табличный) интеграл равен 2π .

В результате

$$\varphi(x) = 1.\tag{9.56}$$

Метод, основанный на применении преобразования Лапласа (метод преобразования Лапласа, операционный метод), широко используется (наряду с методом преобразования Фурье) для решения интегральных уравнений типа свертки, линейных дифференциальных уравнений с постоянными коэффициентами и интегро-дифференциальных уравнений [4].

Литература

1. *Беклемишев Д.В.* Дополнительные главы линейной алгебры. - М.: Наука, 1983.
2. *Богачев К.Ю.* Практикум на ЭВМ. Методы решения линейных систем и нахождения собственных значений. – М., 1998. -137 с.
3. *Сизиков В.С.* Устойчивые методы обработки результатов измерений. - СПб.: Специальная литература, 1999. – 239 с.
4. *Верлань А.Ф. Сизиков В.С.* Интегральные уравнения: методы, алгоритмы, программы. - Киев: Наук. Думка, 1986.
5. *Тихонов А.Н., Арсенин В.Я.* Методы решения некорректных задач. - М.: Наука, 1986.
6. *Бакушинский А.Б., Гончарский А.В.* Некорректные задачи. Численные методы и приложения. - М.: Изд-во МГУ, 1989.
7. *Гантмахер Ф.Р.* Теория матриц. - М.: Наука, 1966.
8. *Тихонов А.Н., Гончарский А.В., Степанов В.В., Ягола А.Г.* Численные методы решения некорректных задач. – М.: Наука, 1990.
9. *Соболев С.Л.* Некоторые приложения функционального анализа в математической физике. - Л.: Изд. ЛГУ, 1950.
10. *Воеводин В.В.* Линейная алгебра. - (2-е изд.). – М.: Наука, 1980. – 400 с.
11. *Воеводин В.В.* Численные методы алгебры (теория и алгоритмы). М.: Наука, 1966. - 248 с.

ОПИСАНИЕ КУРСА И ПРОГРАММА

Цели и задачи магистерской программы «Регулярные методы и алгоритмы расчета обратных задач в моделях оптических структур»

Целью учебно-методического комплекса магистерской программы «Оптика наноструктур» является формирование у студентов четкого представления об основных принципах функционирования современных дифракционных оптических элементов и устройств, тонкопленочных многослойных покрытий; о законах взаимодействия электромагнитного излучения видимого диапазона с материалом. Целью является также изучение способов и возможностей математического синтеза и компьютерного проектирования дифракционных оптических покрытий. Полученные знания закрепляются в оптической лаборатории и дисплейном классе на примерах изучения конкретных дифракционных оптических элементов и многослойных покрытий со сложной геометрией.

Задачей учебно-методического комплекса магистерской программы «Оптика наноструктур» является обучение студентов навыкам самостоятельного анализа технических заданий на проектирование дифракционных оптических элементов и устройств. Они должны научиться выбирать из имеющихся в наличии алгоритмов и программ математического синтеза или разрабатывать их самостоятельно. В результате обучения обретут навыки ориентации в научной и бизнес информации с целью выбора нужной функции или нужного инструмента для реализации известной функции в области проектирования и создания дифракционных оптических наноструктур.

Цели и задачи учебно-методического курса «Регулярные методы и алгоритмы расчета обратных задач в моделях оптических структур»

Курс «Регулярные методы и алгоритмы расчета обратных задач в моделях оптических структур» является составной частью магистерской программы «Оптика наноструктур». Магистерская программа «Оптика наноструктур» реализуется в рамках направления «Прикладная математика и информатика» и направления «Прикладная математика и физика», а возможно и других направлений. В составе магистерской программы «Оптика наноструктур» курс «Регулярные методы и алгоритмы расчета обратных задач в моделях оптических структур» является обязательным привязанным к семестру, для других магистерских программ этот курс может быть курсом по выбору без привязки к семестру или факультативным на усмотрение методической комиссии программы. Курс носит как теоретический, так и прикладной характер.

Основы развития устойчивых приближенных методов решения некорректных задач были заложены в фундаментальной работе А.Н.Тихонова «Об устойчивости обратных задач»// ДАН СССР, 1943, т. 39, № 5., где был введен важный класс так называемых обратных задач, связанных с восстановлением количественных характеристик среды по порождаемым ею физическим полям, доступным для измерения. Интенсивное развитие методов решения неустойчивых задач было предопределено широким внедрением компьютеров в математические исследования, что по законам обратной связи вызвало поток разнообразных задач, для решения которых потребовалось развитие новых приближенных методов. Одним из значимых сегментов обратных некорректных задач являются задачи компьютерной оптики, в частности, задач проектирования дифракционных оптических элементов и устройств.

Методы устойчивого решения обратных задач в моделях взаимодействия электромагнитного излучения видимого диапазона с участками среды со сложной геометрией и сложным по составу диэлектрическим наполнением занимают важное место в ряду регулярных методов решения некорректных задач.

В связи с вышеизложенным целью курса «Регулярные методы и алгоритмы расчета обратных задач в моделях оптических структур» является обучение учащихся теоретическим основам регулярных методов решения некорректных задач, а также конкретным регулярным методам решения задач математического анализа и синтеза дифракционных оптических элементов и устройств, оптических наноструктур.

Задачей курса «Регулярные методы и алгоритмы расчета обратных задач в моделях оптических структур» является обучение студентов умению и навыкам использования методов устойчивого решения обратных задач в моделях взаимодействия электромагнитного излучения видимого диапазона с участками среды со сложной геометрией и сложным по составу диэлектрическим наполнением. Более того, приоритетной задачей является применение этих методов для численного решения задач при проектировании оптических устройств. Они обязаны квалифицированно применять теоретические знания (регулярных методов решения некорректных задач) для разработки оригинальных методов и алгоритмов решения задач математического анализа и синтеза дифракционных оптических элементов и устройств.

Инновационность курса.

Курс является инновационным по содержанию и по литературе, он включает в себя последние научные достижения в области решения задач дифракционной оптики, когда характерные размеры исследуемых объектов не превышают либо сравнимы с длиной волны оптического

излучения. Эта область знаний интенсивно развивалась в последнее время, но лишь недавно были созданы устойчивые алгоритмы и разработаны численные методы решения задач для многослойных решеток. Следует отметить, что для оптических однослойных и многослойных решеток с характерными размерами больше длины волны оптического излучения устойчивые методы решения известны с середины прошлого века. Сейчас алгоритмы решения оптических задач в субволновой области распространяются на объекты со сложной геометрией, такие как двумерные решетки с произвольным профилем, трехмерные решетки (фотонные кристаллы) и на анизотропные материалы. Они востребованы, поскольку позволяют создавать математические модели взаимодействия излучения с веществом в наномасштабах, а затем с их помощью проектировать новые эффективные устройства в высокотехнологичных областях медицины, энергетики, инфокоммуникаций и приборостроения.

В ходе проведения занятий по этому курсу разработчики предполагают использование традиционных методик преподавания, принятой в странах болонской системы образования, то есть с использованием кредитной системы оценки знаний.

Наряду с традиционными элементами преподавания математических методов решения прикладных задач разработчики курса предполагают воспользоваться хорошо зарекомендовавшим себя опытом МФТИ и подобных вузов. А именно, в рамках подпрограммы «Оптика наноструктур» осуществляется закупка уникального аналитического оборудования для измерения разнообразных характеристик оптических наноустройств с целью использования этого оборудования в учебном процессе и для проведения научно-исследовательских работ преподавателями, аспирантами и студентами.

По окончании магистратуры по направлению «Оптика наноструктур» выпускники Российского университета дружбы народов станут конкурентно-способными специалистами в области проектирования современных оптических устройств, которые не будут испытывать затруднений при последующем трудоустройстве.

Курс базируется на публикациях научных статей мировых лидеров исследований в данной области в научной периодике, диссертационных работах их учеников, включающих работы по непосредственному моделированию, дизайну и последующему изготовлению лабораторных образцов оптических элементов и устройств. В список дополнительной и рекомендуемой литературы включены все научно-исследовательские публикации, положенные в основу предлагаемого курса.

В качестве практических заданий, курсовых работ и тем рефератов слушателям магистерской программы будут предложены актуальные проблемы и задачи, решение которых востребовано современным уровнем развития высокотехнологичных отраслей промышленности и научно-исследовательских лабораторий.

Структура курса (с указанием количества часов аудиторных/самостоятельной работы на темы)

Темы лекций

Тема 1. Введение – общее представление об обратных задачах, возникающих при моделировании оптических структур – дифракционных оптических элементов, фотоэмиссионных диодов, солнечные батареи и т.п. Матрица Мюллера – определение ее коэффициентов по методу наименьших квадратов. Восстановление показателя преломления для

изотропных/анизотропных материалов по методам Берремана и Шуберта для многослойных покрытий. Модель просветляющего покрытия - некорректность решения прямой задачи на собственные векторы при кратных собственных значениях (1 пара).

Тема 2. Векторы, операторы, базисы. Матрицы, вектор-столбцы координат, базис в координатной записи. Арифметические пространства R^n и R^m

Матрицы-операторы $A: R^n \rightarrow R^m$

$m = n$ – квадратная матрица A системы линейных алгебраических уравнений $Ax = b$;

$m < n$ – недоопределенная системы линейных алгебраических уравнений $Ax = b$;

$m > n$ – переопределенная системы линейных алгебраических уравнений $Ax = b$;

(1 пара)

Тема 3. Нормы векторов и матриц (операторов)

Зависимость относительной погрешности решения от относительных погрешностей матрицы и правой части

$$\frac{\|\delta x\|}{\|x\|} = F \left(\frac{\|\delta A\|}{\|A\|}, \frac{\|\delta b\|}{\|b\|} \right)$$

Число обусловленности обратимой матрицы (оператора).

Если A необратима, то $x = A^{-1}b$ не определено

Альтернатива Фредгольма:

Существует – не существует

Единственное – не единственное

$$x = (A^T A)^{-1} A^T b, \quad x = A^+ b$$

(2 пары)

Тема 4. Гильбертовы пространства, норма, скалярное произведение, ортонормированный базис. Норма оператора, неограниченные операторы, примеры.

Конечномерные аппроксимации, сходимость, теоремы Бесселя и Планшереля. Дискретизация и обрезание рядов – примеры возмущений. Возмущенные конечномерные СЛАУ

(2 пары)

Тема 5. Задачи, приводящие к необходимости решать некорректные задачи с возмущенными СЛАУ, с интегральными уравнениями 1 рода, с уравнениями типа свертки, родственные им задачи о восстановлении коэффициентов Фурье возмущенных в C метрике функций

Примеры: восстановление аппаратной функции; восстановление коэффициентов матрицы Мюллера по измеренным данным.

(1 пара)

Тема 6. Подробное изложение нескольких конкретных задач:

Обратные задачи спектроскопии.

Редукция измерений к идеальному измерительному прибору.

Восстановление аппаратной функции экранирующей маски при напылении тонких диэлектрических пленок. (2 пары)

Тема 7. Формирование идей построения регулярных алгоритмов и методов решения неустойчивых некорректных обратных задач.

(1 пара)

Тема 8. Регуляризованные методы решения систем линейных алгебраических уравнений. Метод псевдообратных матриц, метод Лаврентьева для систем линейных алгебраических уравнений с симметричными матрицами, метод Тихонова для систем линейных

алгебраических уравнений с произвольными матрицами. Метод сингулярного разложения прямоугольной матрицы. Сравнение детерминированных регуляризованных методов с методами фильтрации Калмана-Бьюси и Винера.

(2 пары)

Тема 9. Аппроксимация сеточных функций с помощью сплайнов. Вейвлет аппроксимация. Методы приближения функций в классах монотонных, выпуклых функций, аналитических функций. Согласованная аппроксимация действительной и мнимой части аналитической функции на примере дисперсионного соотношения Крамерса-Кронига.

(2 пары)

Тема 10. Методы минимизации: метод деформируемого многогранника, градиентный метод, метод Левенберга-Марквардта. Пример: проектирование дифракционного оптического элемента градиентными методами.

(2 пары)

Тема 11. Обзор программного обеспечения по методам решения систем линейных и нелинейных уравнений, минимизации функционалов, аппроксимации функций – библиотеки численных алгоритмов и программ.

(1 пара)

Темы практических занятий (лабораторные работы)

в дисплейном классе

Лабораторные занятия посвящены устойчивым регуляризованным методам решения линейных и нелинейных обратных задач по восстановлению коэффициентов систем уравнений.

1. Вычисление коэффициентов матрицы Мюллера (по входным и выходным значениям интенсивностей и/или различных поляризаций) по методу наименьших квадратов (SVD) для переопределенных матриц.
2. Вычисление коэффициентов матрицы Мюллера (по входным и выходным значениям интенсивностей и/или различных поляризаций). Метод тихоновской регуляризации.
3. Вычисление коэффициентов матрицы Мюллера (по входным и выходным значениям интенсивностей и/или различных поляризаций). Псевдообратные матрицы.
4. Вычисление коэффициентов матрицы Мюллера (по входным и выходным значениям интенсивностей и/или различных поляризаций). Метод деформируемого многогранника.
5. Восстановление показателя преломления для изотропных/анизотропных материалов по методу Берремана (нелинейная задача) по методу наименьших квадратов для переопределенных систем (метод покоординатного спуска).
6. Восстановление показателя преломления для изотропных/анизотропных материалов по методу Берремана (нелинейная задача) по методу тихоновской регуляризации.
7. Расчет толщины просветляющего (или затемняющего) покрытия на определенной длине волны для очков с заданным коэффициентом преломления стекол (задача Тихонравова).

Примечание. Все задачи лабораторных работ решаются при заданной длине волны.

Темы коллоквиумов и контрольных работ

1. Эквивалентность норм в конечномерном векторном пространстве. Зависимость относительной ошибки решения возмущенной системы линейных алгебраических уравнений от относительных ошибок матрицы и вектора правой части..
2. Слабая и сильная сходимость конечномерных аппроксимаций обобщенных рядов Фурье.
3. Эффект Гиббса, регуляризирующие множители Ланцоша и Фейера.
4. Тихоновская регуляризация устойчивого суммирования рядов Фурье.
5. Метод Тихонова регуляризации решения возмущенной системы линейных алгебраических уравнений.
6. Метод тихоновской регуляризации решения интегральных уравнений первого рода.

Описание системы контроля знаний:

Общие правила выполнения контрольных заданий;

Требования к оформлению работы

Постановка задачи.

1. Краткая формулировка задачи.
2. Развернутая постановка задачи с указанием основных режимов работы и их сценариев.

Алгоритм решения.

1. Математическое описание алгоритма.
2. Структура алгоритма ядра программы (укрупненная блок-схема).

Тестирование.

1. Описание основных режимов тестирования алгоритма и программы и результатов работы программы.
2. Список возможных ошибок и аномалий, описание реакции программы на них.

Заключение.

Содержит общие комментарии и замечания исполнителя о выполненной работе.

Приложение.

Приложение должно содержать текст программы (полная распечатка или распечатка алгоритма ядра программы).

Работа должна быть представлена в виде распечатанного текста и на дискете (Word + Delphi и/или C++).

Рекомендации к составлению отчета

Оформление.

отчет по работе должен быть оформлен в форме Word-файла.

Содержание отчета.

Каждый пункт задания вычислительного эксперимента должен найти свое отражение в отчете.

Каждый раздел отчета должен содержать:

формулировку цели эксперимента

описание исходных данных - приближаемая функция, интервал и порядок приближения, метод приближения и т.п.

результаты эксперимента, представленные в форме таблиц, гистограмм и графиков

иллюстрационный материал в виде копий экрана с графиками зависимостей погрешности приближения, вида приближаемой функции и т.п.

выводы, следующие из результатов эксперимента в контексте его цели.

Примерные типы письменных работ и форм устного контроля

Примерный образец оформления курсовой работы на тему «Интегральные уравнения первого рода. Регуляризованный метод Тихонова» приведен в приложении 1.

Шкала оценок, итоговые оценки (методика выставления)

Бально-рейтинговая методика оценки уровня знаний по обязательной дисциплине «Регулярные методы и алгоритмы расчета обратных задач в моделях оптических структур», привязанной к семестру

Порядок начисления баллов за семестр.

Контрольная работа № 1: 0 – 40 баллов

Теоретические вопросы: 0 – 10 баллов

Практические задания: 0 – 30 баллов

Контрольная работа № 2: 0 – 40 баллов

Теоретические вопросы: 0 – 10 баллов

Практические задания: 0 – 30 баллов

Контрольная работа № 3: 0 – 20 баллов

Теоретические вопросы: 0 – 20 баллов

Шкала бально-рейтинговой системы.

Баллы за семестр	Автоматическая оценка		Баллы за итоговый контроль знаний	Общая сумма баллов	Итоговая оценка
	Итоговая оценка	Дополнительные баллы			
78 – 80	зачет	по 5 баллов за каждый свыше 76**	0 – 20*	86 – 100	зачет
41 – 77	Нет	Нет	0 – 20	51 – 97	зачет
			0 – 20	41 – 50	незачет
< 41	незачет	Нет	Нет	Нет	незачет

* студент имеет право не проходить итоговый контроль знаний.

** дополнительные баллы начисляются автоматически:

за 86 баллов, набранных в семестре, начисляется дополнительно 6 баллов

(общая сумма баллов – 92);

за 87 баллов – 12 баллов (99);

за 88 баллов – 18 баллов (106);

за 89 баллов – 24 балла (113);

за 90 баллов – 30 баллов (120).

Академическая этика, соблюдение авторских прав.

Все имеющиеся в тексте сноски тщательно выверены и снабжены «адресами». Авторы не включали в свою работу выдержки из работ других авторов без указания на это, не пересказывали чужих работ близко к тексту без отсылки к ним. Авторы также не использовали чужих идей без указания первоисточников. Это касается и источников, найденных в интернете. В необходимых случаях указан полный адрес сайта.

Программа курса УМК:

Аннотированное содержание курса.

Первый модуль трудоемкостью в 1 кредит составляют:

- теоретический материал, излагаемый в четырех первых темах,
- практическими занятиями в дисплейном классе в течение 12 академических часов.

В конце этого модуля проводится промежуточный контроль знаний.

Второй модуль трудоемкостью в 1 кредит составляют:

- теоретический материал, излагаемый в четырех темах с пятой по восьмую,
- практическими занятиями в дисплейном классе в течение 12 академических часов.

В конце этого модуля проводится промежуточный контроль знаний.

Третий модуль трудоемкостью в 1 кредит составляют:

- теоретический материал, излагаемый в девятой – одиннадцатой темах,
- практическими занятиями в дисплейном классе в течение 10 академических часов.

В конце этого модуля проводится итоговый контроль знаний.

Список обязательной и дополнительной литературы с указанием соответствия разделов источника (постранично) разделам читаемого курса

Список обязательной литературы.

1. Регулярные методы и алгоритмы расчета обратных задач в моделях оптических структур/ Под ред. Л.А. Севастьянова: Учебное пособие. - М.: Изд-во РУДН (готовится к печати).

2. Тихонов А.Н., Арсенин В.Я. Методы решения некорректных задач. М: Наука, 1979. 288 с.
3. А.Ф. Верлань, В.С. Сизиков. Интегральные уравнения: методы, алгоритмы, программы. Справочное пособие. – Киев: Наукова думка, 1986, 544с.
4. В.С. Сизиков. Устойчивые методы обработки результатов измерений. Учебное пособие.-СПб.: «СпецЛит», 1999, 240 с.
5. А.Н.Тихонов, А.С.Леонов, А.Г.Ягола. Нелинейные некорректные задачи.- М.:Наука, 1995
6. А.Н.Тихонов, А.В.Гончарский, В.В.Степанов, А.Г.Ягола. Численные методы решения некорректных задач. -М.:Наука, 1990
7. Севастьянов Л.А. Математическая модель экранируемого напыления: вычислительный эксперимент, использующий результаты натуральных экспериментов. Дисс. ... Докт. Физ.-мат. Наук, 1999, ОИЯИ, Дубна.
8. Пытьев Ю.П. Методы математического моделирования измерительно-вычислительных систем. 2-е изд., перераб. М.: Физматлит, 2004, 400с.
9. <http://alglib.sources.ru/optimization/levenbergmarquardt.php>
- 10.

Список дополнительной литературы и источников в интернете.

- 11.Гантмахер Ф. Р., Теория матриц. М.: Наука, 1966, 576 стр.
- 12.В.В. Воеводин, Ю.А. Кузнецов. Матрицы и вычисления.
- 13.Ярив А., Юх П. Оптические волны в кристаллах. М. Мир.1983.
- 14.Уилкинсон Дж., Алгебраическая проблема собственных значений, М.: Наука, 1970, 564 с.
15. www.netlib.org
16. www.exponenta.ru
17. <http://alglib.sources.ru/>

Темы рефератов, курсовых работ, эссе

Темы рефератов

1. Редукция к идеальному измерительному прибору (обзор работ Пытьева Ю.П.)
2. Обзор методов линейной фильтрации (Сизиков и др.)

Темы курсовых работ

1. Решение интегрального уравнения Фредгольма первого рода (дискретизация по методу конечных разностей) на примере восстановления гравитационного потенциала.
2. Решение интегрального уравнения Фредгольма первого рода с помощью разложения решения в ряд Фурье на примере восстановления гравитационного потенциала.
3. Градиентные методы проектирования дифракционных оптических элементов (с ограничениями и без).

Темы курсовых работ с последующим продолжением в качестве магистерской диссертации.

1. Интегральное преобразование Крамерса-Кронига. По работе: A. B. Kuzmenko, Kramers-Kronig constrained variational analysis of optical spectra DPMC, University of Geneva, 24 Quai Ernest-Ansermet, 1211 Geneva 4, Switzerland, arXiv:cond-mat/0503565 v1 23 Mar 2005
2. Методы проектирования многослойных зеркал с заданным коэффициентом отражения в фиксированном диапазоне длин волн.
3. Методы проектирования зеркал на основе дифракционных решеток с заданным коэффициентом отражения в фиксированном диапазоне длин волн. По диссертационной

работе **Tunable Optoelectronic Devices** by Carlos Fernando Rondina Mateus, UNIVERSITY of CALIFORNIA at BERKELEY, 2004, 217 p.

Учебный тематический план курса УМК (календарный план, структурированный по видам учебных занятий)

Календарный план (20 недель) учебных занятий по обязательной учебной дисциплине «Методы и алгоритмы решения задач в моделях оптических покрытий», привязанный ко второму семестру магистратуры.

Виды и содержание учебных занятий				
Неделя	Лекции	Число часов	Лабораторные занятия	Число часов
1	Введение – общее представление об обратных задачах, возникающих при моделировании оптических структур Матрица Мюллера. Восстановление показателя преломления для изотропных/анизотропных материалов.	2	Подготовка графической оболочки программы для прикладных расчетов	2

2	Векторы, операторы, матрицы. Системы линейных алгебраических уравнений.	2	Вычисление коэффициентов матрицы Мюллера (по входным и выходным значениям интенсивностей) по методу наименьших квадратов (SVD) для переопределенных матриц.	2
3	Нормы векторов и матриц. Зависимость относительной погрешности решения от относительных погрешностей матрицы и правой части.	2	Вычисление коэффициентов матрицы Мюллера (по входным и выходным значениям различных поляризаций) по методу наименьших квадратов (SVD) для переопределенных матриц.	2
4	Число обусловленности обратной матрицы. Альтернатива Фредгольма	2	Вычисление коэффициентов матрицы Мюллера (по входным и выходным значениям различных поляризаций). Метод тихоновской регуляризации.	2
5	Гильбертовы	2	Вычисление	2

	пространства, норма, скалярное произведение, ортонормированный базис. Норма оператора.		коэффициентов матрицы Мюллера (по входным и выходным значениям интенсивностей). Метод тихоновской регуляризации.	
6	Конечномерные аппроксимации, сходимость. Возмущенные конечномерные СЛАУ	2	Вычисление коэффициентов матрицы Мюллера (по входным и выходным значениям различных поляризации). Псевдообратные матрицы.	2
7	Промежуточный контроль (Контрольная работа №1)			2
8	Практические примеры некорректных задач.	2	Вычисление коэффициентов матрицы Мюллера (по входным и выходным значениям интенсивностей). Псевдообратные матрицы.	2
9	Редукция измерений к идеальному измерительному прибору.	2	Вычисление коэффициентов матрицы Мюллера (по входным и выходным значениям интенсивностей). Метод	2

			деформируемого многогранника.	
10	Восстановление аппаратной функции экранирующей маски при напылении тонких диэлектрических пленок.	2	Вычисление коэффициентов матрицы Мюллера (по входным и выходным значениям различных поляризацій). Метод деформируемого многогранника.	2
11	Формирование идей построения регулярных алгоритмов и методов решения неустойчивых некорректных обратных задач.	2	Восстановление показателя преломления для изотропных материалов по методу Берремана (нелинейная задача) по методу наименьших квадратов для переопределенных систем.	2
12	Регуляризованные методы решения систем линейных алгебраических уравнений	2	Восстановление показателя преломления для анизотропных материалов по методу Берремана (нелинейная задача) по методу наименьших квадратов для переопределенных систем.	2

13	Сравнение детерминированных регуляризованных методов с методами фильтрации Калмана-Бьюси и Винера.	2	Восстановление показателя преломления для изотропных материалов по методу Берремана (нелинейная задача) по методу покоординатного спуска.	2
14	Промежуточный контроль (Контрольная работа №2)			2
15	Аппроксимация сеточных функций с помощью сплайнов. Вейвлет аппроксимация.	2	Восстановление показателя преломления для изотропных материалов по методу Берремана (нелинейная задача) с помощью тихоновской регуляризации.	2
16	Согласованная аппроксимация действительной и мнимой части аналитической функции на примере дисперсионного соотношения Крамерса-Кронига.	2	Восстановление показателя преломления для анизотропных материалов по методу Берремана (нелинейная задача) с помощью тихоновской регуляризации.	2
17	Метод деформируемого	2	Расчет толщины просветляющего	2

	многогранника, градиентный метод, метод Левенберга- Маркгвардта		покрытия на определенной длине волны для очков с заданным коэффициентом преломления стекол (задача Тихонравова).	
18	Проектирование дифракционного оптического элемента градиентными методами.	2	Расчет толщины затемняющего покрытия на определенной длине волны для очков с заданным коэффициентом преломления стекол (задача Тихонравова).	2
19	Обзор программного обеспечения по методам решения систем линейных и нелинейных уравнений, минимизации функционалов, аппроксимации функций.	2	Заключительный обзор курса. Консультации.	2
20	Итоговый контроль знаний			2