Государственное образовательное учреждение высшего профессионального образования
Ульяновский государственный университет

Факультет математики и информационных технологий
Кафедра «Информационные технологии»

КУРСОВАЯ РАБОТА

Задача полных наименьших квадратов

(название темы)

010501(65) Прикладная математика и информатика

(код и наименование специальности)

Выполнил студент	РГ–41м		Иванов П. С.
	(группа)	(подпись)	(Ф.И.О.)
Научный руководитель	профессор		Семушин И. В.
	(должность)	(подпись)	(Ф.И.О.)

(оценка)

Ульяновск
2008

Оглавление

Введение

Здесь пишите ...

Задача Введения — обосновать актуальность темы данной работы, т. е. объяснить:

- ✧ Зачем написана эта работа ?
- ✧ Для кого написана эта работа ?
- ✧ Как структурно организован текст работы ?
- ✧ Какой материал помещен в основные структурные единицы работы ?
- ✧ Чем подкреплена данная работа: эксперименты, расчеты... ?

Еще вариант оформления списков:

- ✓ Зачем написана эта работа ?
- ✓ Для кого написана эта работа ?
- ✓ Как структурно организован текст работы ?
- ✓ Какой материал помещен в основные структурные единицы работы ?
- ✓ Чем подкреплена данная работа: эксперименты, расчеты... ?

Еще вариант оформления списков:

- ➠ Зачем написана эта работа ?
- ➠ Для кого написана эта работа ?
- ➠ Как структурно организован текст работы ?
- ➠ Какой материал помещен в основные структурные единицы работы ?
- ➠ Чем подкреплена данная работа: эксперименты, расчеты... ?

Еще вариант оформления списков:

- ➢ Зачем написана эта работа ?
- ➢ Для кого написана эта работа ?
- ➢ Как структурно организован текст работы ?
- ➢ Какой материал помещен в основные структурные единицы работы ?
- ➢ Чем подкреплена данная работа: эксперименты, расчеты... ?

Еще вариант оформления списков:

- ⋗ Зачем написана эта работа ?
- ⋗ Для кого написана эта работа ?
- ⋗ Как структурно организован текст работы ?
- ⋗ Какой материал помещен в основные структурные единицы работы ?
- ⋗ Чем подкреплена данная работа: эксперименты, расчеты... ?

Вариантов оформления списков очень много ... *Еще примеры см. в Заключении.*

Вложенные списки также возможны:

- ⋗ Зачем написана эта работа ?
  - ☞ Для кого написана эта работа ?
  - ☞ Как структурно организован текст работы ?
  - ☞ Какой материал помещен в основные структурные единицы работы ?
  - ☞ Чем подкреплена данная работа: эксперименты, расчеты... ?
- ⋗ Для кого написана эта работа ?
  - ✠ Для кого написана эта работа ?
  - ✠ Как структурно организован текст работы ?
  - ✠ Какой материал помещен в основные структурные единицы работы ?

  - ✠ Чем подкреплена данная работа: эксперименты, расчеты. . . ?

- ⪢ Как структурно организован текст работы ?
  - ➡ Для кого написана эта работа ?
  - ➡ Как структурно организован текст работы ?
  - ➡ Какой материал помещен в основные структурные единицы работы ?
  - ➡ Чем подкреплена данная работа: эксперименты, расчеты. . . ?

- ⪢ Какой материал помещен в основные структурные единицы работы ?
  - ✯ Для кого написана эта работа ?
  - ✯ Как структурно организован текст работы ?
  - ✯ Какой материал помещен в основные структурные единицы работы ?
  - ✯ Чем подкреплена данная работа: эксперименты, расчеты. . . ?

- ⪢ Чем подкреплена данная работа: эксперименты, расчеты. . . ?
  - ❄ Для кого написана эта работа ?
  - ❄ Как структурно организован текст работы ?
  - ❄ Какой материал помещен в основные структурные единицы работы ?
  - ❄ Чем подкреплена данная работа: эксперименты, расчеты. . . ?

Однако слишком «разукрашивать» свою дипломную работу не стоит. Стиль оформления должен быть не рекламный, а достаточно строгий, академический. Желательно остановиться на одном-единственном варианте оформления ненумерованных списков, например, на таком:

- ▹ современный уровень банковских услуг;
- ▹ развитие технической и информационной базы банков и их клиентов;
- ▹ технология работы электронных банковских продуктов;
- ▹ безопасность проведения расчетов;
- ▹ возможности и экономическая целесообразность внедрения электронных розничных услуг в коммерческих банках.

Глава 1

Безусловная оптимизация

Здесь пишите ...
Как показано в гл. 1, подразд. 1.1, пункт 1.1.2 ...

1.1 Функции для минимизации

1.1.1 Название

Здесь пишите ..., что в данном пункте 1.1.1

Название

Здесь пишите ..., что в данном подпункте

Название

Здесь пишите ..., что в данном подпункте

1.1.2 Название

Здесь пишите ..., что в данном пункте 1.1.2

Название

Здесь пишите ..., что в данном подпункте

Название

Здесь пишите ..., что в данном подпункте

1.1.3 Таблица функций для минимизации

Таблица 1.1. Функции для безусловной минимизации

№ п/п	Функция $f(x)$	Начальный вектор x_0	Точка минимума $x^\star$	Значение $f(x^\star)$
1	$(1-x_1)^2+(x_1-x_2)^2+$ $+(x_2-x_3)^2+(x_3-x_4)^2$	$[-2;3;-4;5]$	$[1;1;1;1]$	0
2	$(x_1-1)^2+10(x_2-1)^2+$ $+100(x_3-1)^2+1000(x_4-$ $-1)^2$	$[-1;-2;-3;-4]$	$[1;1;1;1]$	0
3	$100(x_2-x_1^2)^2+(1-x_1)^2$	$[3;4]$	$[1;1]$	0
4	$100(x_2-x_1^3)^2+(1-x_1)^2$	$[-1,2;1]$	$[1;1]$	0
5	$(x_1+10x_2)^2+5(x_3-x_4)^2+$ $+(x_2-2x_3)^4+10(x_1-x_4)^4$	$[3;-1;0;1]$	$[0;0;0;0]$	0
6	$x_1^2+x_2^2+x_3^2+x_4^2+16x_1^2x_2^2+$ $+8x_2^2x_3^2+x_3^2+x_4^2+2$	$[1;2;3;4]$	$[0;0;0;0]$	0
7	$10(x_1-\sin x_2)^2+0,1x_2^2$	$[1,2;3]$	$[0;0]$	0
8	$(1,5-x_1(1-x_2))^2+$ $+(2,25-x_1(1-x_2^2))^2+$ $+(2,625-x_1(1-x_2^3))^2$	$[0;0]$	$[3;0,5]$	0
9	$100(x_2-x_1^2)^2+(1-x_1)^2+$ $+90(x_4-x_3^2)^2+(1-x_3)^3+$ $+10,1((x_2-1)^2+(x_4-$ $-1)^2)+19,8(x_2-1)(x_4-$ $-1)$	$[-3;-1;-3;-1]$	$[1;1;1;1]$	0
10	$-x_1-2x_3-x_2x_3+x_1^2+$ $+x_2^2+x_3^2$	$[0;0;0]$	$\left[\frac{1}{2};\frac{2}{3};\frac{4}{3}\right]$	$\frac{9}{12}$
11	$x_1^3+x_2^3-3x_1x_2$	$[-1;3]$	$[1;1]$	-1
12	$2x_1x_2x_3-4x_1x_3-2x_2x_3+$ $+x_1^2+x_2^2+x_3^2-2x_1-4x_2+$ $+4x_3$	$[-3;-3;-3]$	$[1;2;0]$	-5
13	$x_1^4+x_2^4-(x_1+x_2)^2$	$[0;0]$	$[1;1]\,,[-1;-1]$	-2
14	$x_1^2+x_2^2+x_3^2-x_1x_2+x_1-$ $-2x_3$	$[2;2;2]$	$\left[-\frac{2}{3};\frac{1}{3};1\right]$	$-\frac{4}{3}$

продолжение таблицы на следующей странице

продолжение таблицы с предыдущей страницы

№ п/п	Функция $f(x)$	Начальный вектор x_0	Точка минимума $x^\star$	Значение $f(x^\star)$
15	$3x_1 - x_1 + x_2^3 - 3x_2^2 - 1$	$[-1; -1]$	$\left[\frac{1}{3}; 2\right]$	$-\frac{47}{9}$
16	$6x_1 + 2x_1^2 - 2x_1x_2 + 2x_2^2$	$[-1; -1]$	$[-2; -1]$	-6
17	$x_1 + x_2^2 + ((x_1 + x_2 - 10)/3)^2$	$[-1; -1]$	$[5; 0,5]$	$7,5$
18	$(x_1 - 1)^2 + 100(x_1 - x_2)^2$	$[3; 4]$	$[1; 1]$	0
19	$5(x_1 - 3)^2 + (x_2 - 5)^2$	$[0; 0]$	$[3; 5]$	0
20	$x_1^2 - x_1x_2 + x_2^2$	$[1; 2]$	$[0; 0]$	0
21	$9x_1^2 + 16x_2^2 - 90x_1 - 128x_2$	$[0; 3]$	$[5; 4]$	-481
22	$2x_1^2 + 2x_2^2 + 2x_1x_2 - 4x_1 -$ $- 6x_2$	$[1; 1]$	$\left[\frac{1}{3}; \frac{4}{3}\right]$	$-\frac{14}{3}$
23	$x_1^2 - x_1x_2 + x_2^2 - 2x_1 + x_2$	$[3; 5]$	$[1; 0]$	-1
24	$5x_1 + 4x_1x_2 + x_2^2 - 16x_1 -$ $- 12x_2$	$[1; 1]$	$[-4; 14]$	-152
25	$2x_1^2 + 2x_2^2 + x_1x_2 - 11x_1 -$ $- 8x_2$	$[-3; -5]$	$[2; 3]$	-23
26	$x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$	$[1; 1]$	$[-1; 1,5]$	$-1,25$
27	$x_1^2 + x_2^2 + x_1x_2$	$[1; 1]$	$[0; 0]$	0
28	$x_1^2 + 16x_2^2$	$[2; 2]$	$[0; 0]$	0
29	$(1 - x_1)^2 + (x_1 - x_2)^2$	$[-5; -8]$	$[1; 1]$	0
30	$x_1^2 + 4x_2^2 + 1$	$[3; 5]$	$[0; 0]$	1

1.2 Название

Здесь пишите ...

1.2.1 Название

Здесь пишите ...

Название

Здесь пишите ...

Название

Здесь пишите ...

1.2.2 Название

Здесь пишите ...

Название

Здесь пишите ...

Название

Здесь пишите ...

1.3 Название

Здесь пишите ...

1.3.1 Название

Здесь пишите ...

Название

Здесь пишите ...

Название

Здесь пишите ...

1.3.2 Название

Здесь пишите ...

Название

Здесь пишите ...

Название

Здесь пишите ...

1.4 Название

Здесь пишите ...

1.4.1 Название

Здесь пишите ...

Название

Здесь пишите ...

Название

Здесь пишите ...

1.4.2 Название

Здесь пишите ...

Название

Здесь пишите ...

Название

Здесь пишите ...

Глава 2

Название

Здесь пишите . . .

Как показано в гл. 2, подразд. 2.1, пункт 2.1.2 ...

2.1 Название

Здесь пишите . . .

2.1.1 Название

Здесь пишите . . .

Название

Здесь пишите . . .

Название

Здесь пишите . . .

2.1.2 Название

Здесь пишите . . .

Название

Здесь пишите . . .

Название

Здесь пишите . . .

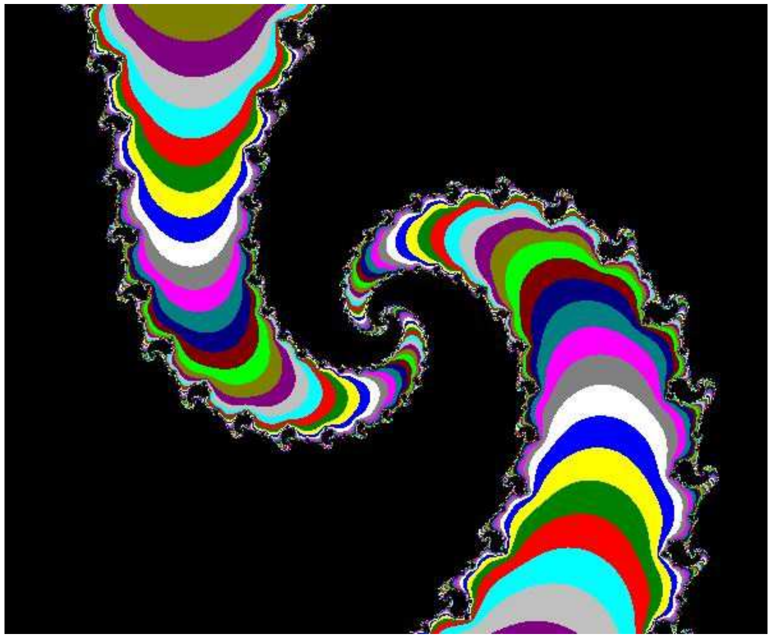


Рис. 2.1. Множество Жюлиа.

2.2 Название

Здесь пишите …

2.2.1 Название

Здесь пишите …

Название

Здесь пишите …

Название

Здесь пишите …

2.2.2 Название

Здесь пишите …

Название

Здесь пишите …

Название

Здесь пишите …

2.3 Название

Здесь пишите …

2.3.1 Название

Здесь пишите …

Название

Здесь пишите …

Название

Здесь пишите …

2.3.2 Название

Здесь пишите …

Название

Здесь пишите …

Название

Здесь пишите …

2.4 Название

Здесь пишите …

2.4.1 Название

Здесь пишите …

Название

Здесь пишите ...

Название

Здесь пишите ...

2.4.2 Название

Здесь пишите ...

Название

Здесь пишите ...

Название

Здесь пишите ...

Глава 3

Название

Здесь пишите ...

3.1 Название

Здесь пишите ...

3.1.1 Название

Здесь пишите ...

Название

Здесь пишите ...

Название

Здесь пишите ...

3.1.2 Название

Здесь пишите ...

Название

Здесь пишите ...

Название

Здесь пишите ...

3.2 Название

Здесь пишите . . .

3.2.1 Название

Здесь пишите . . .

Название

Здесь пишите . . .

Название

Здесь пишите . . .

3.2.2 Название

Здесь пишите . . .

Название

Здесь пишите . . .

Название

Здесь пишите . . .

3.3 Название

Здесь пишите . . .

3.3.1 Название

Здесь пишите . . .

Название

Здесь пишите . . .

Название

Здесь пишите ...

3.3.2 Название

Здесь пишите ...

Название

Здесь пишите ...

Название

Здесь пишите ...

3.4 Название

Здесь пишите ...

3.4.1 Название

Здесь пишите ...

Название

Здесь пишите ...

Название

Здесь пишите ...

3.4.2 Название

Здесь пишите ...

Название

Здесь пишите ...

Название

Здесь пишите ...

Глава 4

Название

Здесь пишите ...

4.1 Название

Здесь пишите ...

4.1.1 Название

Здесь пишите ...

Название

Здесь пишите ...

Название

Здесь пишите ...

4.1.2 Название

Здесь пишите ...

Название

Здесь пишите ...

Название

Здесь пишите ...

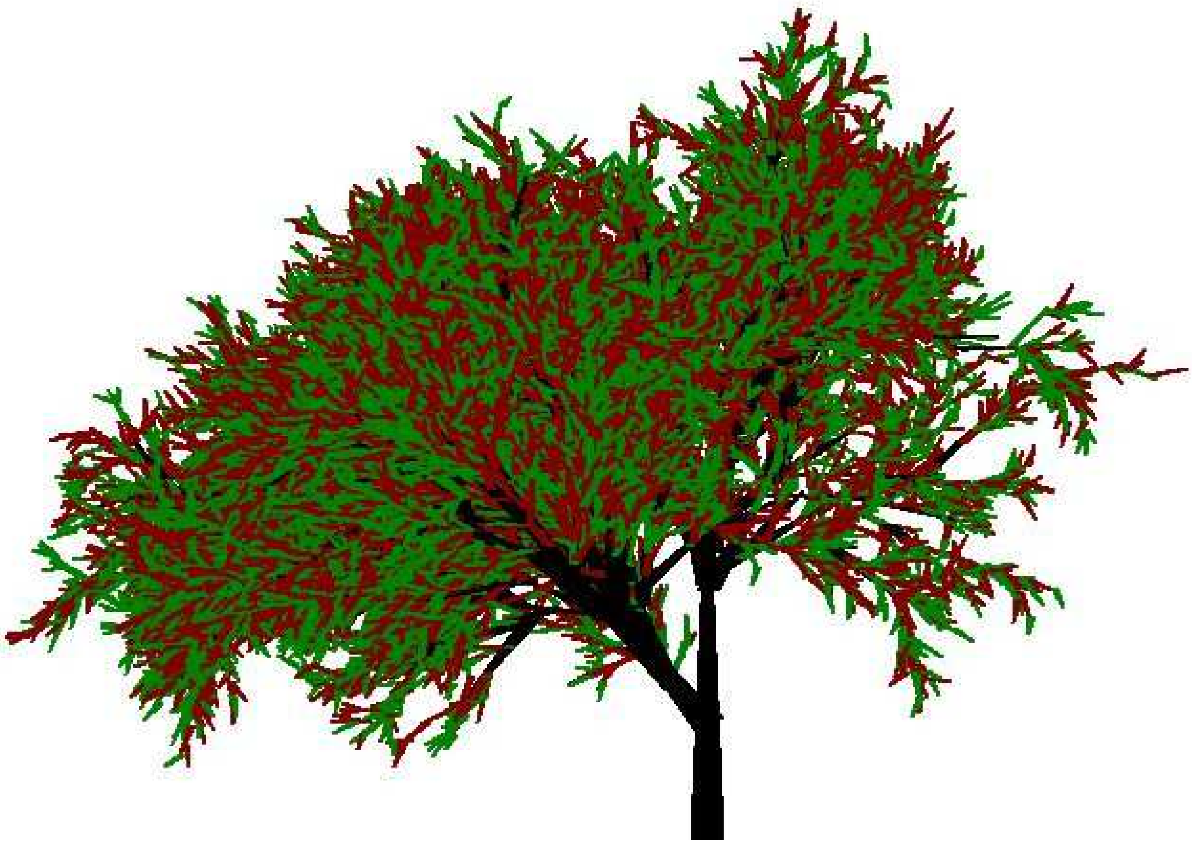


Рис. 4.1. Фрактальное дерево.

4.2 Название

Здесь пишите ...

4.2.1 Название

Здесь пишите ...

Название

Здесь пишите ...

Название

Здесь пишите ...

4.2.2 Название

Здесь пишите ...

Название

Здесь пишите ...

Название

Здесь пишите ...

4.3 Название

Здесь пишите ...

4.3.1 Название

Здесь пишите ...

Название

Здесь пишите ...

Название

Здесь пишите ...

4.3.2 Название

Здесь пишите ...

Название

Здесь пишите ...

Название

Здесь пишите ...

4.4 Название

Здесь пишите ...

4.4.1 Название

Здесь пишите ...

Название

Здесь пишите ...

Название

Здесь пишите ...

4.4.2 Название

Здесь пишите ...

Название

Здесь пишите ...

Название

Здесь пишите ...

Заключение

Здесь пишите …

В данной работе освещены аспекты систем электронных расчетов в современном коммерческом банке со следующих точек зрения:

- ▹ современный уровень банковских услуг;
- ▹ развитие технической и информационной базы банков и их клиентов;
- ▹ технология работы электронных банковских продуктов;
- ▹ безопасность проведения расчетов;
- ▹ возможности и экономическая целесообразность внедрения электронных розничных услуг в коммерческих банках.

Это, конечно, ПРИМЕР. — Заключение не должно пересказывать основной текст, — этого делать не надо.

В Заключении надо:

- перечислить, какие задачи из числа задач этой работы, сформулированных в разделе «Постановка задачи», вами решены,
- сформулировать, какие выводы из этих результатов вами найдены и обоснованы в основном тексте работы.

Еще один пример оформления списков (нумерованный список):

1. Перечислить, какие задачи из числа задач этой работы, сформулированных в разделе «Постановка задачи», вами решены.
2. Сформулировать, какие выводы из этих результатов вами найдены и обоснованы в основном тексте работы.

Здесь пишите …

В данной работе освещены аспекты систем электронных расчетов в современном коммерческом банке со следующих точек зрения:

- ▹ современный уровень банковских услуг;
- ▹ развитие технической и информационной базы банков и их клиентов;
- ▹ технология работы электронных банковских продуктов;
- ▹ безопасность проведения расчетов;
- ▹ возможности и экономическая целесообразность внедрения электронных розничных услуг в коммерческих банках.

Это, конечно, ПРИМЕР. — Заключение не должно пересказывать основной текст, — этого делать не надо.

В Заключении надо:

- перечислить, какие задачи из числа задач этой работы, сформулированных в разделе «Постановка задачи», вами решены,
- сформулировать, какие выводы из этих результатов вами найдены и обоснованы в основном тексте работы.

Еще один пример оформления списков (нумерованный список):

1. Перечислить, какие задачи из числа задач этой работы, сформулированных в разделе «Постановка задачи», вами решены.
2. Сформулировать, какие выводы из этих результатов вами найдены и обоснованы в основном тексте работы.

Список иллюстраций

Список таблиц

Список использованной литературы

1. Андриевский, Б. Р. Избранные главы теории автоматического управления с примерами на языке MATLAB / Б. Р. Андриевский, А. Л. Фрадков. — СПб. : Наука, 2000.
2. Араманович, И. Г. Функции комплексного переменного. Операционное исчисление. Теория устойчивости / И. Г. Араманович, Г. Л. Лунц, Л. Э. Эльсгольц. — М. : Наука, 1965.
3. Евграфов, М. А. Аналитические функции / М. А. Евграфов. — М. : Наука, 1965.
4. Корн, Г. Справочник по математике (для научных работников и инженеров) / Г. Корн, Т. Корн. — М. : Наука, 1974.
5. Лаврентьев, М. А. Методы теории функций комплексного переменного / М. А. Лаврентьев, Б. В. Шабат. — М. : Наука, 1987.
6. Медич, Дж. Статистически оптимальные линейные оценки и управление / Дж. Медич. — М. : Энергия, 1973.
7. Острем, К. Ю. Введение в стохастическую теорию управления / К. Ю. Острем. — М. : Мир, 1973.
8. Пугачев, В. С. Основы стохастической теории автоматических систем / В. С. Пугачев, И. Е. Казаков, Л. Г. Евланов. — М. : Наука, 1980.
9. Свешников, А. Г. Теория функций комплексной переменной / А. Г. Свешников, А. Н. Тихонов. — М. : Наука, 1979.
10. Соломенцев, Е. Д. Функции комплексного переменного и их применения / Е. Д. Соломенцев. — М. : Высшая школа, 1988.
11. Андриевский, Б. Р. Избранные главы теории автоматического управления с примерами на языке MATLAB / Б. Р. Андриевский, А. Л. Фрадков. — СПб. : Наука, 2000.

12. Араманович, И. Г. Функции комплексного переменного. Операционное исчисление. Теория устойчивости / И. Г. Араманович, Г. Л. Лунц, Л. Э. Эльсгольц. — М. : Наука, 1965.

13. Евграфов, М. А. Аналитические функции / М. А. Евграфов. — М. : Наука, 1965.

14. Корн, Г. Справочник по математике (для научных работников и инженеров) / Г. Корн, Т. Корн. — М. : Наука, 1974.

15. Лаврентьев, М. А. Методы теории функций комплексного переменного / М. А. Лаврентьев, Б. В. Шабат. — М. : Наука, 1987.

16. Медич, Дж. Статистически оптимальные линейные оценки и управление / Дж. Медич. — М. : Энергия, 1973.

17. Острем, К. Ю. Введение в стохастическую теорию управления / К. Ю. Острем. — М. : Мир, 1973.

18. Пугачев, В. С. Основы стохастической теории автоматических систем / В. С. Пугачев, И. Е. Казаков, Л. Г. Евланов. — М. : Наука, 1980.

19. Свешников, А. Г. Теория функций комплексной переменной / А. Г. Свешников, А. Н. Тихонов. — М. : Наука, 1979.

20. Соломенцев, Е. Д. Функции комплексного переменного и их применения / Е. Д. Соломенцев. — М. : Высшая школа, 1988.

21. Андриевский, Б. Р. Избранные главы теории автоматического управления с примерами на языке MATLAB / Б. Р. Андриевский, А. Л. Фрадков. — СПб. : Наука, 2000.

22. Араманович, И. Г. Функции комплексного переменного. Операционное исчисление. Теория устойчивости / И. Г. Араманович, Г. Л. Лунц, Л. Э. Эльсгольц. — М. : Наука, 1965.

23. Евграфов, М. А. Аналитические функции / М. А. Евграфов. — М. : Наука, 1965.

24. Корн, Г. Справочник по математике (для научных работников и инженеров) / Г. Корн, Т. Корн. — М. : Наука, 1974.

25. Лаврентьев, М. А. Методы теории функций комплексного переменного / М. А. Лаврентьев, Б. В. Шабат. — М. : Наука, 1987.

26. Медич, Дж. Статистически оптимальные линейные оценки и управление / Дж. Медич. — М. : Энергия, 1973.

27. Острем, К. Ю. Введение в стохастическую теорию управления / К. Ю. Острем. — М. : Мир, 1973.

28. Пугачев, В. С. Основы стохастической теории автоматических систем / В. С. Пугачев, И. Е. Казаков, Л. Г. Евланов. — М. : Наука, 1980.

29. Свешников, А. Г. Теория функций комплексной переменной / А. Г. Свешников, А. Н. Тихонов. — М. : Наука, 1979.

30. Соломенцев, Е. Д. Функции комплексного переменного и их применения / Е. Д. Соломенцев. — М. : Высшая школа, 1988.

31. Андриевский, Б. Р. Избранные главы теории автоматического управления с примерами на языке MATLAB / Б. Р. Андриевский, А. Л. Фрадков. — СПб. : Наука, 2000.

32. Араманович, И. Г. Функции комплексного переменного. Операционное исчисление. Теория устойчивости / И. Г. Араманович, Г. Л. Лунц, Л. Э. Эльсгольц. — М. : Наука, 1965.

33. Евграфов, М. А. Аналитические функции / М. А. Евграфов. — М. : Наука, 1965.

34. Корн, Г. Справочник по математике (для научных работников и инженеров) / Г. Корн, Т. Корн. — М. : Наука, 1974.

35. Лаврентьев, М. А. Методы теории функций комплексного переменного / М. А. Лаврентьев, Б. В. Шабат. — М. : Наука, 1987.

36. Медич, Дж. Статистически оптимальные линейные оценки и управление / Дж. Медич. — М. : Энергия, 1973.

37. Острем, К. Ю. Введение в стохастическую теорию управления / К. Ю. Острем. — М. : Мир, 1973.

38. Пугачев, В. С. Основы стохастической теории автоматических систем / В. С. Пугачев, И. Е. Казаков, Л. Г. Евланов. — М. : Наука, 1980.

39. Свешников, А. Г. Теория функций комплексной переменной / А. Г. Свешников, А. Н. Тихонов. — М. : Наука, 1979.

40. Соломенцев, Е. Д. Функции комплексного переменного и их применения / Е. Д. Соломенцев. — М. : Высшая школа, 1988.

Приложение A

Ортогональные преобразования

A.1 Ортогональные матрицы и приложения

В этом разделе напомним определение и некоторые свойства ортогональных матриц, полезные для дальнейшего.

Определение 1.1. Матрица T, имеюшая размер $n \times n$, т. е.$T(n, n)$, есть *ортогональная* матрица, если и только если $TT^T = I$.

Свойство A. Если T_1 и T_1 суть две ортогональные матрицы, то их произведение $T_1 T_2$ есть тоже ортогональная матрица.

Свойство B. $T^{-1} = T^T$ и $T^T T = I$.

Свойство C. Ортогональное преобразование сохраняет скалярное произведение векторов, т. е.$\forall x, y \in \mathbb{R}^n : y^T x \triangleq (x, y) = (Tx, Ty)$, в частности, оно сохраняет (евклидову) норму вектора: $\|Ty\| = \|y\|$.

Свойство D. Если v есть вектор случайных переменных с математическим ожиданием $\mathbf{E}\{v\} = 0$ и ковариацией $\mathbf{E}\left\{vv^T\right\} = I$, то теми же характеристиками обладает вектор $\bar{v} = Tv$, т. е.

$$\mathbf{E}\{\bar{v}\} = 0, \qquad \mathbf{E}\left\{\bar{v}\bar{v}^T\right\} = I.$$

Хотя это свойство легко проверяется, немного удивительно, что компоненты преобразованного вектора остаются взаимно некоррелированными.

Свойства C и D играют существенную роль в квадратно-корневых алгоритмах решения прикладных задач оптимального моделирования и оптимального оценивания методом наименьших квадратов (рис. A.1).

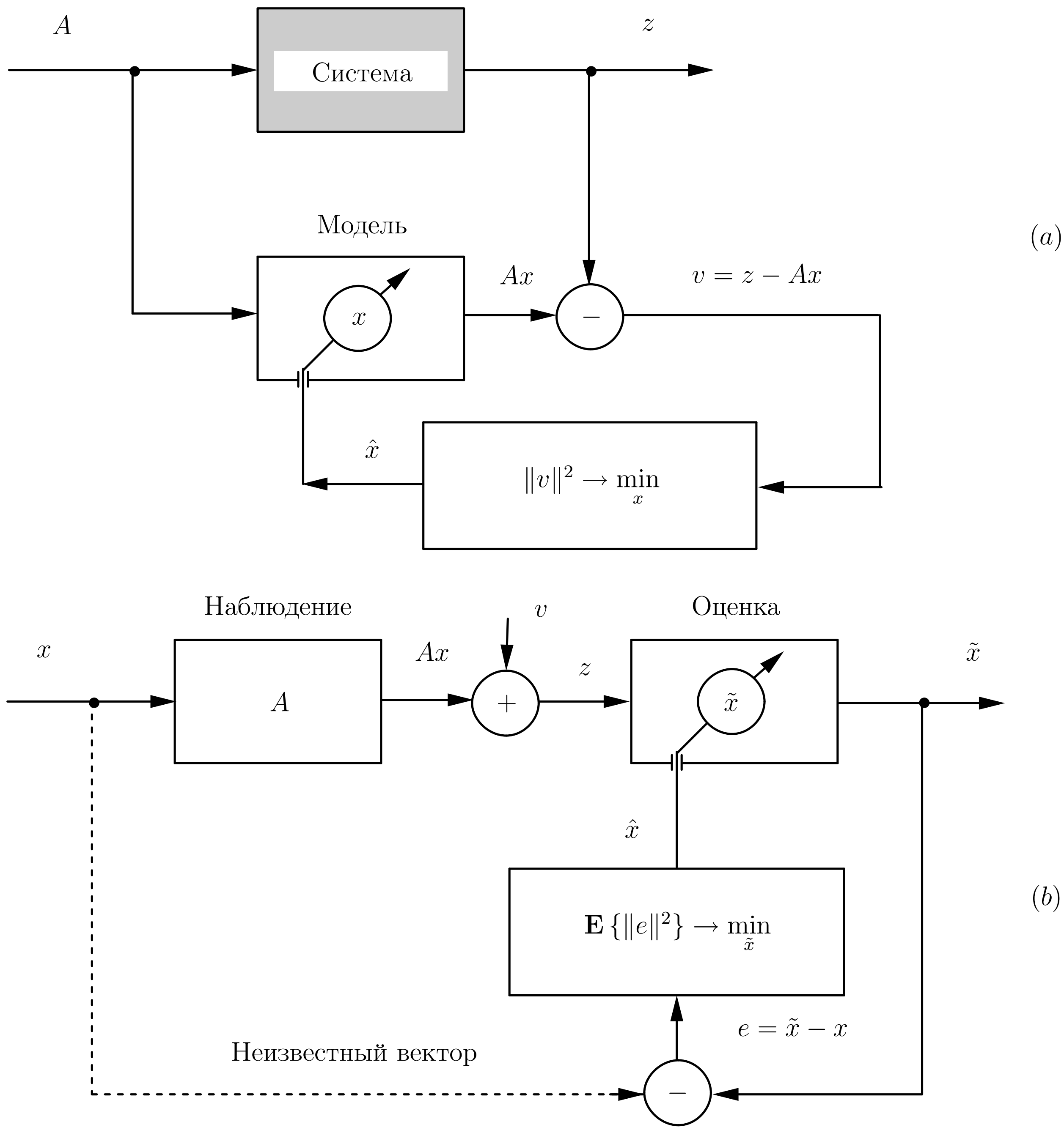


Рис. A.1. Алгебраически эквивалентные задачи, решаемые методом наименьших квадратов значений невязки v или среднего квадрата погрешности e: (a) — оптимальное моделирование неизвестной системы по экспериментальным условиям A и данным z; (b) — оптимальное оценивание неизвестного вектора по наблюдениям Ax в присутствии случайных помех v с характеристиками $\mathbf{E}\{v\} = 0$ и $\mathbf{E}\left\{vv^T\right\} = I$

A.2 Линейная задача наименьших квадратов

Линейная задача наименьших квадратов (см. рис. A.1) ставится следующим образом.

Дано *линейное* уравнение

$$z = Ax + v, \tag{A.1}$$

в котором известны вектор $z \in \mathbb{R}^m$ и $(m \times n)$-матрица $A \in \mathbb{R}^{m\times n}$, т. е.$A = A(m, n)$. Разностный вектор $v \triangleq z - Ax$, называемый *невязкой*, зависит от переменного вектора $x \in \mathbb{R}^n$. Требуется найти значение $\hat{x}$ вектора x, минимизирующее *квадратический* критерий качества

$$J(x) = (z - Ax)^T (z - Ax) = \|v\|^2 \to \min\,. \tag{A.2}$$

Если ни при каком x невязка v не может быть обращена в $\mathbf{0}$ — нулевой вектор, то система $Ax = z$ — несовместная, в противном случае совместная, т. е.критерий (A.2) охватывает оба случая. Однако сам *метод наименьших квадратов* (МНК), выраженный критерием (A.2), создан Лежандром в 1805 году как алгебраическая процедура именно для несовместных систем и подтвержден как статистическая процедура Гауссом в 1809 году [**?**]. МНК как алгебраическая процедура проиллюстрирован выше с помощью рис. A.1(a), а как статистическая процедура — с помощью рис. A.1(b). Замечательно, что обе процедуры имеют одинаковые решения, т. е.алгебраически эти решения эквивалентны и при $\mathbf{E}\{v\} = 0$ и $\mathbf{E}\left\{vv^T\right\} = I$ (см. рис. A.1(b)) совпадают, поэтому можно говорить о едином МНК-решении $\hat{x}$.

МНК-решение $\hat{x}$ всегда существует как решение *нормальных уравнений*

$$A^T A\hat{x} = A^T z, \tag{A.3}$$

выражается формулой

$$\hat{x} = A^+ z + (I - A^+ A)y \tag{A.4}$$

через произвольный вектор $y \in \mathbb{R}^n$, где A^+ — *псевдообратная матрица* для матрицы A, и единственно тогда и только тогда, когда $A^+ A = I$, что равносильно условию, что только нулевой вектор составляет ядро (нуль-пространство) матрицы A, т. е.при $\operatorname{rank} A = n$.

Условие $\operatorname{rank} A = n$, называемое *условием полного столбцового ранга* матрицы A, обусловливает случай $m \geq n$, что при $m > n$ означает переопределенную систему полного ранга в (A.1). Этот типичный для практики случай ниже и рассматривается, при этом из (A.3), (A.4) следует $\hat{x} = A^+ z$ и $A^+ = (A^T A)^{-1} A^T$.

Замечание A.1. Слагаемое $\hat{x}_0 \triangleq A^+ z$ в (A.4) есть единственное МНК-решение с минимальной нормой, называемое *нормальным псевдорешением*. Оно ортогонально второму слагаемому в (A.4), т. е. $A^+ z \perp (I - A^+ A) y$, и лежит в пространстве строк матрицы A, т. е. $\hat{x}_0 \in \mathcal{R}(A^T)$.

Таким образом, типичный для практики случай имеет формальное решение $\hat{x} = \hat{x}_0 = (A^T A)^{-1} A^T z$, и вычислительная задача наименьших квадратов заключается в его эффективном отыскании.

A.3 Ортогональные матрицы в задаче о наименьших квадратах

В рассматриваемой задаче о наименьших квадратах

$$J(x) = \|z - Ax\|^2, \quad A(m, n), \quad m \geq n, \quad \operatorname{rank} A = n. \tag{A.5}$$

Пусть T, $T(m, m)$, есть матрица некоторого ортогонального преобразования. В силу свойства C (см. разд. A.1) запишем

$$J(x) = \|T(z - Ax)\|^2 = \|(Tz) - (TA)x\|^2. \tag{A.6}$$

При таком представлении видно, что минимум критерия $J(x)$, равный $J(\hat{x})$, не зависит от T. Этом фактом можно воспользоваться, т. е. матрицу T можно выбрать так, что (TA) приобретает привлекательную для вычислений форму. Действительно, в разд. A.4 и A.7 мы покажем, как можно выбрать T, чтобы преобразованная матрица имела вид

$$TA = \begin{bmatrix} R \\ \mathbf{0} \end{bmatrix} \begin{matrix} \}\, n \\ \}\, m - n \end{matrix} \tag{A.7}$$

с верхнетреугольным блоком R, $\operatorname{rank} R = n$.

Если соответственно этому вектор Tz разбить на блоки, т. е. записать

$$Tz = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \begin{matrix} \}\, n \\ \}\, m - n \end{matrix}, \tag{A.8}$$

то $J(x)$ от (A.6) приводится к виду

$$J(x) = \|z_1 - Rx\|^2 + \|z_2\|^2. \tag{A.9}$$

Приведение критерия наименьших квадратов к виду (А.9) позволяет видеть, что искомый вектор $\hat{x}$, минимизирующий этот критерий, должен удовлетворять уравнению

$$R\hat{x} = z_1, \tag{А.10}$$

которое легко решается обратной подстановкой (см. разд. А.6), и кроме того,

$$\min\ J(x) = J(\hat{x}) = \|z_2\|^2. \tag{А.11}$$

В вычислительном отношении эти результаты гораздо более элегантны, чем неразумная трата сил на решение нормальных уравнений (А.3). Но важнее всего то, что решение, использующее ортогональные преобразования (соотношения (А.7), (А.8), (А.10) и (А.11)), менее чувствительны к погрешностям, вызванным ошибками округления в компьютере. Это видно хотя бы из того, что выражение (А.7) влечет равенство

$$R^T R = (TA)^T (TA) = A^T A,$$

которое означает, что R является квадратным корнем из матрицы $(A^T A)$ системы нормальных уравнений (А.3). Следовательно, при решении системы (А.10) вдвое более эффективно используется разрядная сетка компьютера, чем при решении системы (А.3)[1].

А.4 Преобразование Хаусхолдера

Преобразования Хаусхолдера суть матричные представления, которые соответствуют геометрическому понятию отражения. Пусть задан некоторый ненулевой вектор u, который мы называем *направляющим вектором.* Подпространство, ортогональное ему, есть гиперплоскость $U_\perp$. Если взять произвольный вектор y, то можно отразить его от $U_\perp$, в точности соблюдая законы обычного оптического отражения от плоского зеркала (рис. А.2).

Обозначим отраженный вектор y_{r}. Поскольку положение гиперплоскости $U_\perp$ не зависит от длины направляющего вектора, пронормируем его, т. е.образуем орт $\hat{u} = u/\|u\|$. Проекция $(y \mid u)$ вектора y на прямую, задаваемую направлением u, равна $(y^T \hat{u})\hat{u}$. Следовательно,

$$y = (y \mid u) + v, \quad v \perp u, \quad v \in U_\perp. \tag{А.12}$$

[1] Представление в компьютере квадрата a^2 любого действительного числа a требует удвоенной разрядности мантиссы, т. е.счет по уравнению (А.10) равносилен счету с удвоенной разрядностью мантиссы чисел по уравнению (А.3).

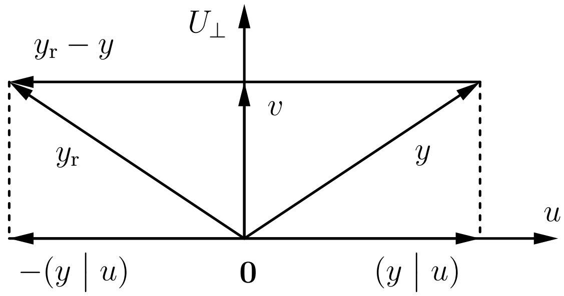


Рис. A.2. Геометрия преобразования Хаусхолдера. Задача 1 (прямая): даны векторы u и y, найти вектор y_r, отраженный от гиперплоскости $U_\perp$

Отраженный вектор y_r, как видно из рис. A.2, имеет разложение

$$y_\mathrm{r} = -(y \mid u) + v, \quad v \perp u, \quad v \in U_\perp \tag{A.13}$$

с той же составляющей v, которая ортогональна вектору u, но с проекцией $-(y \mid u)$, которая (в силу знака $-$) направлена противоположно проекции $(y \mid \hat{u})$ вектора y на направление u. Исключая v из (A.12) и (A.13), находим

$$y_\mathrm{r} = y - 2(y \mid u) = (I - \beta u u^T) y = T_u y, \tag{A.14}$$

где $\beta \triangleq 2/\|u\|^2 = 2/u^T u$. Матрица $T_u \triangleq (I - \beta u u^T)$, в вычислениях явно не участвующая, имеет фундаментальное значение для приложений в силу своих замечательных свойств.

Свойство 1. $T_u = T_u^T$, т. е. T_u — *симметрическая* матрица.

Свойство 2. $T_u^2 = I$, т. е. T_u — *идемпотентная* матрица. Это легко продемострировать алгебраически разложением матрицы T_u^2 или же геометрически по рис. A.2 как двукратное отражение вектора y относительно $U_\perp$.

Свойство 3. Если $u(j) = 0$, то $(T_u y)(j) = y(j)$, т. е. если j-я компонента вектора u — нулевая, то T_u оставляет j-ю компоненту вектора y неизменной.

Свойство 4. Если $u \perp y$, то $T_u y = y$.

Свойство 5.

$$T_u y = y - \gamma u, \qquad \gamma \triangleq 2 y^T u / u^T u = \beta y^T u. \tag{A.15}$$

Свойство 5 важно с практической точки зрения. Формирование матрицы T_u в качестве множителя для y потребовало бы на порядок больше вычислений, чем того требует прямое вычисление $T_u y$ по свойству 5. Это также означает, что не нужно тратить память для хранения T_u, что наиболее существенно проявляется при больших m.

Триангуляризация матрицы преобразованиями Хаусхолдера. Обратимся к основному применению ортогональных преобразований. Для этого решим задачу, обратную к той, что рассмотрена выше, а именно: дан вектор y и дано желаемое расположение отраженного вектора y_{r}, — найти направление u такое, что $T_u y = (s, 0, \ldots, 0)^T$ (рис. А.3). Из свойства С, разд. А.1, норма (евклидова длина) вектора y не изменяется при ортогональном преобразовании, следовательно, определим ее как

$$\sigma \triangleq \|T_u y\| = |s| = (y^T y)^{1/2}. \tag{А.16}$$

Направление u может быть получено из свойства 5 (уравнение (А.15)), т. е.

$$u = \text{const} \cdot (y - s e_1). \tag{А.17}$$

Этот результат приводит к следующему свойству.

Свойство 6. Пусть $s = -\operatorname{sgn}\,[y(1)]\sigma$, где $\operatorname{sgn}\,[\cdot]$ — функция знака,

$$\operatorname{sgn}\,[x] = \begin{cases} 1, & x \geq 0, \\ -1, & x < 0, \end{cases}$$

и элементы вектора u определены выражением (А.17), т. е.$u(1) = y(1) - s$, $u(i) = y(i)$, $i > 1$. Тогда $T_u y = s e_1$ и $\beta \triangleq 2/u^T u = -1/(s u(1))$.

Замечание А.2. Геометрический смысл выражения (А.17) ясен из рис. А.3, где видно, что вектор y_{r} ортогонален гиперплоскости $U_\perp$ и параллелен вектору u.

Непосредственное вычисление $u^T u$ показывает, что $u^T u = -2 s u(1)$, при этом знак для s выбран противоположно знаку первого элемента $y(1)$, т. е.так, чтобы максимизировать $|u(1)|$ и тем уменьшить относительную погрешность вычисления разности $u(1) = y(1) - s$. Если свойство 6 применить к матрице A, взяв в качестве y ее первый столбец, то это даст первый шаг, который послужит основой приведения матрицы к верхнетреугольному виду. Повторение таких действий шаг за шагом позволит осуществлять верхнюю триангуляризацию любой заданной матрицы A.

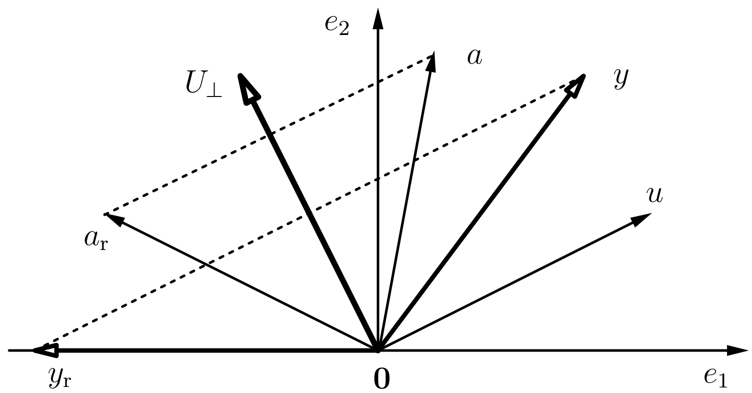


Рис. A.3. Геометрия преобразования Хаусхолдера. Задача 2 (обратная): даны векторы y и y_r, найти вектор u, задающий отражающую гиперплоскость $U_\perp$; здесь $y_\mathrm{r} = se_1 = \left[s \mid 0 \cdots 0\right]^T$

Лемма A.1. Пусть дана матрица $A(m,n)$. Тогда существует ортогональное преобразование Хаусхолдера T_u такое, что

$$T_u A = \begin{array}{r} \\ 1\{ \\ m-1\{ \end{array} \begin{array}{c} \overbrace{}^{1} \quad \overbrace{}^{n-1} \\ \left[\begin{array}{c|c} s & \\ \mathbf{0} & \tilde{A} \end{array}\right] \end{array}. \tag{A.18}$$

Замечание A.3. Скаляр s и матрица $\tilde{A}$ в (A.18) вычисляются непосредственно по данным в матрице A; s — по выражению (A.16) и свойству 6, а $\tilde{A}$ — по свойству 5, (A.15). Первый столбец, который уместно назвать *ведущим* столбцом, используют как вектор y в задаче 2 (см. рис. A.3) для определения вектора u. Второй и далее столбцы, обозначенные на рис. A.3 произвольно как вектор a, отражают от найденной таким образом гиперплоскости $U_\perp$, решая для этого задачу 1 (см. рис. A.2) с $y := a$ и тем самым получая блок $\tilde{A}$.

Теорема A.1 *(Триангуляризация матрицы по методу Хаусхолдера).* Пусть $A_1 := A(m,n)$ и для каждого j выбрано элементарное преобразование Хаусхолдера T_j так, что

$$T_j A_j = \begin{array}{r} \\ 1\{ \\ m-j\{ \end{array} \begin{array}{c} \overbrace{}^{1} \quad \overbrace{\phantom{A_{j+1}}}^{n-j} \\ \left[\begin{array}{c|c} s_j & a_j^T \\ \mathbf{0} & A_{j+1} \end{array}\right] \end{array}, \quad j = 1, \ldots, k; \quad k \le \min(m-1, n). \tag{A.19}$$

Тогда в процессе после k повторных применений свойства 6 и леммы А.1 имеем следующий промежуточный результат триангуляризации матрицы A:

$$T^{(k)}A = \begin{bmatrix} s_1 & & & & a_1^T \\ & s_2 & & & a_2^T \\ & & \ddots & & \vdots \\ & \mathbf{0} & & s_k & a_k^T \\ & & & & A_{k+1} \end{bmatrix} \qquad \text{(A.20)}$$

с отвечающей этому моменту процесса итоговой матрицей преобразований

$$T^{(k)} = \begin{bmatrix} I_{k-1} & 0 \\ 0 & T_k \end{bmatrix} \cdots \begin{bmatrix} I_1 & 0 \\ 0 & T_2 \end{bmatrix} T_1 \,. \qquad \text{(A.21)}$$

Замечание А.4. Важно подчеркнуть, что алгоритм триангуляризации (А.19) не требует вычисления или запоминания ортогональной матрицы $T^{(k)}$, так как правая часть равенства (А.4) вычисляется непосредственно в соответствии с замечанием А.3. Стоит также заметить, как неявно определяется A_{j+1} рекурсией по j в алгоритме (А.19). Кроме A_{j+1}, на шаге j этой рекурсии определяются скаляр s_j и $(n-j)$ компонент вектор-строки a_j^T. Эти неявные соотношения для s_j, a_j^T и A_{j+1} и весь процесс вычислений (рис. А.4) представлены в явном виде в разд. А.5.

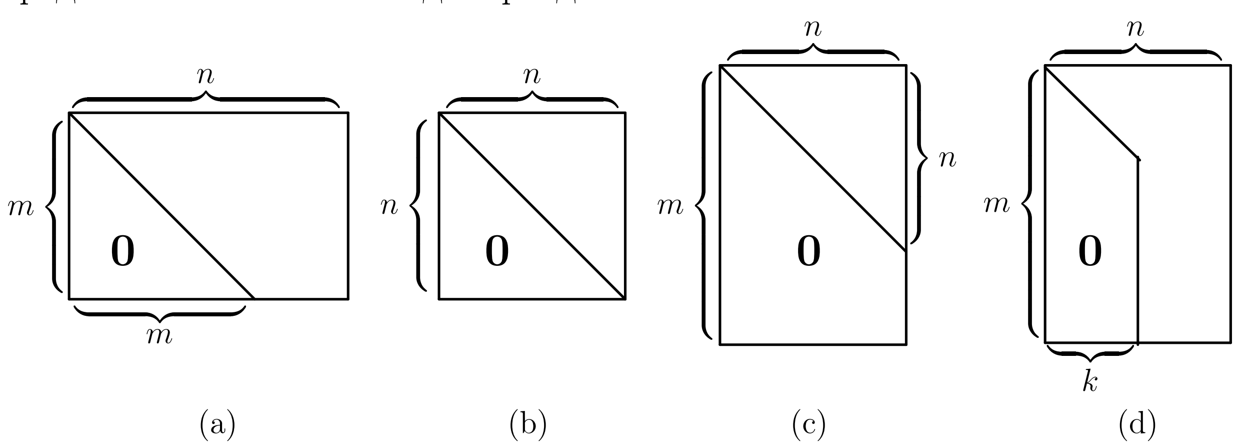


Рис. А.4. Схематическое представление возможных случаев применения теоремы А.1 к матрице $A(m,n)$; (a) недоопределенная система: $k = m - 1 \leq n$; (b) определенная система: $k = n - 1$, $m = n$; (c) переопределенная система: $k = n < m$; (d) $k < n < m$

А.5 Шаг триангуляризации матрицы преобразованием Хаусхолдера

Пусть матрица $A = A(m, n)$ задана. Тогда, согласно лемме А.1, базовая операция процесса триангуляризации заключается в вычислении скаляра s и матрицы $\tilde{A} = \tilde{A}(m, n-1)$ таких, что

$$T_u A = \begin{array}{r} \\ 1\{ \\ m-1\left\{ \right. \end{array} \begin{array}{c} \overbrace{\hphantom{s}}^{1} \quad \overbrace{\hphantom{\tilde{A}}}^{n-1} \\ \left[\begin{array}{c|c} s & \\ \mathbf{0} & \tilde{A} \end{array} \right] \end{array}. \tag{A.22}$$

Алгоритм. Для вычисления s следует применять свойство 6, т.е.выполнять (А.23) для всех $k = 1$ до $\min\,(m-1, n)$. Затем для вычисления $\tilde{A}$ следует применять свойство 5, т.е.последовательно в цикле по $j = 2, \ldots, n$ для всех $i = k, \ldots, m$ выполнять (А.24) с $\lambda \triangleq -\gamma$ (см. (А.15)). Здесь введена величина $\alpha \triangleq -\beta$ (см. (А.14) и свойство 6).

Для $k = 1$ до $\min\,(m-1, n)$

$$\left.\begin{aligned} s_k &= -\operatorname{sgn}\,[A(k,k)] \left(\sum_{i=k}^{m} [A(i,k)]^2 \right)^{1/2}, \\ u_k(1) &= A(k,k) - s_k, \\ u_k(i) &= A(k+i-1, k), \quad i = 2, \ldots, m-k+1, \\ \alpha_k &= 1/(s_k u_k(1)) \qquad (\alpha_k < 0). \end{aligned}\right\} \tag{A.23}$$

Для $j = k+1, \ldots, n$

$$\left.\begin{aligned} &\lambda := \alpha_k \cdot \sum_{i=k}^{m} u_k(i-k+1) A(i,j), \\ &\text{Для } i = k, k+1, \ldots, m \\ &A(i,j) := A(i,j) + \lambda u_k(i-k+1). \end{aligned}\right\} \tag{A.24}$$

Приведенный алгоритм (А.23), (А.24) называют *столбцово ориентированным*, так как операции (А.24) вычисляют целиком каждый j-й столбец матрицы, находящийся справа от ведущего, т. е.k-го столбца.

Альтернативная схема вычислений называется *строчно ориентированным* алгоритмом. Ее можно получить из выражения $T_u = I - \beta u u^T$ для матрицы Хаусхолдера следующим образом.

Введем вспомогательные обозначения: $\mu \triangleq \sqrt{\beta}$, $w \triangleq \mu u$, чтобы записать $T_u = I - ww^T$. Тогда $(T_u A) = A - wz^T$, где $z^T \triangleq w^T A = \mu v^T$, $v^T \triangleq \sum_{i=1}^{m} u(i)A(i,\cdot)$ и $A(i,\cdot)$ есть i-я строка матрицы $A = A(m,n)$. Введем обозначение $\lambda^T = \alpha v^T$, используя ранее введенное (см. (A.23)) $\alpha \triangleq -\beta$. Отсюда получаем формулу для любой i-й строки $(T_u A)(i,\cdot)$ преобразованной матрицы $(T_u A)$ в виде

$$(T_u A)(i,\cdot) = A(i,\cdot) - w(i)z^T = A(i,\cdot) - \mu^2 u(i)v^T = A(i,\cdot) + \lambda^T u(i).$$

Алгоритм (строчно ориентированный), эквивалентный (A.23) и (A.24).

Для $k = 1$ до $\min\,(m-1, n)$

$$\left.\begin{aligned} s_k &= -\operatorname{sgn}\,[A(k,k)] \left(\sum_{i=k}^{m} [A(i,k)]^2\right)^{1/2}, \\ u_k(1) &= A(k,k) - s_k, \\ u_k(i) &= A(k+i-1,k), \quad i = 2,\ldots,m-k+1, \\ \alpha_k &= 1/(s_k u_k(1)) \qquad (\alpha_k < 0). \end{aligned}\right\} \tag{A.25}$$

$$\left.\begin{aligned} &\text{Для } j = k+1,\ldots,n \\ &\qquad \lambda_k(j-k) := \alpha_k \cdot \sum_{i=k}^{m} u_k(i-k+1)A(i,j), \\ &\text{Для } i = k, k+1,\ldots,m \\ &\qquad \text{Для } j = k+1,\ldots,n \\ &\qquad\qquad A(i,j) := A(i,j) + \lambda_k(j-k)u_k(i-k+1). \end{aligned}\right\} \tag{A.26}$$

A.6 Решение треугольной системы $Rx = z$ и обращение матриц R и A

Как отмечено в разд. A.3, мы часто заинтересованы в решении уравнения

$$Rx = z, \tag{A.27}$$

где $R = R(n,n)$ — верхнетреугольная невырожденная матрица. Если нужно иметь только решение x, то R^{-1} (для $x = R^{-1}z$) вычислять не надо. Следующий алгоритм обратной подстановки позволяет вычислить решение x непосредственно.

Алгоритм. Для $j = n, n-1, \dots, 1$ вычислять

$$x(j) = \left(z(j) - \sum_{k=j+1}^{n} R(j,k)x(k) \right) / R(j,j). \tag{A.28}$$

По сложности этот алгоритм почти такой же, как матричное умножение. Он допускает записывать $x(j)$ поверх $z(j)$, что очень удобно в приложениях.

Если все же требуется иметь матрицу $U \triangleq R^{-1}$, то ее можно вычислить по алгоритму окаймления, основанному на следующем легко проверяемом тождестве для треугольных матриц:

$$\begin{bmatrix} R_j & y \\ 0 & \sigma_{j+1} \end{bmatrix}^{-1} = \begin{bmatrix} R_j^{-1} & -R_j^{-1} y \sigma_{j+1}^{-1} \\ 0 & \sigma_{j+1}^{-1} \end{bmatrix} = R_{j+1}^{-1}. \tag{A.29}$$

Это соотношение позволяет вычислять обратную матрицу $U \triangleq R^{-1}$ рекуррентно, т.е., если $R_j^{-1} = U_j$, где R_j обозначает верхнюю левую часть матрицы R, то

$$U_{j+1} = \begin{bmatrix} U_j & -U_j \left[R(1,j+1), \dots, R(j,j+1) \right]^T \sigma_{j+1} \\ 0 & \sigma_{j+1} \end{bmatrix}, \tag{A.30}$$

где $\sigma_{j+1} = 1/R(j+1, j+1)$. Полагая $U = R_{-1}$, этот результат представим в алгоритмической форме.

Алгоритм. Задать начальное значение

$$U(1,1) = 1/R(1,1). \tag{A.31}$$

Для $j = 2, \dots, n$ вычислять по формулам (A.32) и (A.33):

$$U(j,j) = 1/R(j,j), \tag{A.32}$$

$$U(k,j) = -\left(\sum_{i=k}^{j-1} U(k,i)R(i,j) \right) U(j,j), \quad k = 1, \dots, j-1. \tag{A.33}$$

Замечание А.5. R^{-1} вычисляется по столбцам, при этом элементы матрицы R^{-1} могут записываться поверх элементов исходной матрицы R.

В справочных целях приведем примеры реализации данного алгоритма на языке FORTRAN. Эти примеры могут помочь студентам написать свои собственные программы на других языках высокого уровня при выполнении лабораторного проекта № 6, описание которого дано ниже в подразд. **??**.

Обращение верхнетреугольной матрицы: $U := R^{-1}$**.** Реализуются формулы (А.31), (А.32) и (А.33). Если нужно, U может замещать R.

$R(N, N)$, $U(N, N)$, R и U — верхнетреугольные матрицы.
$U(1,1) = 1./R(1,1)$
DO 20 $J = 2, N$
$U(J,J) = 1./R(J,J)$
$JM1 = J - 1$
DO 20 $K = 1, JM1$
$SUM = 0.$
DO 10 $I = K, JM1$
10 $SUM = SUM - U(K,I) * R(I,J)$
20 $U(K,J) = SUM * U(J,J)$

В случаях, когда важно или нужно экономить память компьютера, матрицы в программе объявляют как одномерные массивы (см. подразд. **??**). Хотя в компьютере даже многомерно объявленные массивы всегда хранятся как одномерные, компилятор генерирует индексные выражения с операциями умножения и деления. Операции сложения и вычитания в компьютерах выполняются гораздо быстрее, поэтому индексы для доступа к элементам матриц следует программировать в рекуррентной инкрементной форме, экономя таким образом и время процессора (табл. А.1). В этой программе преобразование в треугольную форму выполняется отождествлением $J(J-1)/1+I$ с (I, J). Рекуррентное инкрементное вычисление KK, JJ и KK экономит вычисления.

Как отмечалось на с. 44, иногда требуется вычислять R^{-1}. Такая ситуация появляется, если требуется найти A^{-1}, для которой уже выполнено преобразование $TA = R$, где $T = T^{(n-1)}$ по формуле (А.21), так как в теореме А.1 для этого случая $m = n$ и $A^{-1} = R^{-1}T$. Последнее означает, что то же самое ортогональное преобразование T теперь надо применить к строкам матрицы R^{-1}, но уже в обратном порядке следования элементарных преобразований, составляющих полное преобразование $T = T^{(n-1)}$ по формуле (А.21). Таким образом, возникает проблема запоминания элементарных преобразований, составляющих полное преобразование $T = T^{(n-1)}$, чтобы позже можно было его применить в задаче отыскания A^{-1} или же для решения уравнения $Ax = z$ с невырожденной матрицей A после пребразования $TA = R$.

Как видно из (А.24), для отражения любого вектора $y = T^{(k)}z$ от гиперплоскости $U_{\perp}$, заданной вектором u, требуется иметь сохраненными две

Таблица А.1. Эффективное обращение верхнетреугольной матрицы: $U := R^{-1}$. Реализуются формулы (А.31), (А.32) и (А.33). Верхнетреугольные матрицы R и U хранятся как векторы размерности $N(N+1)/2$. Если нужно, U может замещать R.

	$R(N,N)$, $U(N,N)$, R и U — верхнетреугольные матрицы.	
	$U(1) = 1./R(1)$	© $R(1) \equiv R(1,1)$
	$JJ = 1$	
	DO 20 $J = 2, N$	
	$\overline{JJ} = JJ$	© $\overline{JJ} = J(J-1)/2 = (J-1, J-1)$
	$JJ = JJ + J$	© $JJ = J(J+1)/2 = (J,J)$
	$U(JJ) = 1./R(JJ)$	
	$JM1 = J - 1$	
	$KK = 0$	
	DO 20 $K = 1, JM1$	
	$KK = KK + K$	© $KK = K(K+1)/2$
	$KI = KK$	
	$SUM = 0.$	
	DO 10 $I = K, JM1$	
	$SUM = SUM - U(KI) * R(\overline{JJ} + I)$	© $KI = (K,I)$, $\overline{JJ} + 1 = = (I,J)$
10	$KI = KI + I$	© $KI = (K, I+1)$
20	$U(\overline{JJ} + K) = SUM * U(JJ)$	© $\overline{JJ} + K = (K,J)$

величины: вектор u и скаляр α. Поскольку нули ниже диагонали, получающиеся в результате триангуляризации, хранить не нужно, это место можно отвести для сохранения вектора u (вместе с диагональю, поэтому диагональные элементы s_k приходится хранить отдельно). Данное предложение схематически показано на рис. А.5. Каким образом можно выполнить вычисление матрицы A^{-1}, показано на рис. А.6 на конкретном примере размерностей, а именно: $m = 4$, $n = 4$.

А.7 Преобразование Гивенса

Преобразование Гивенса осуществляет вращение вектора в плоскости двух координат. Поскольку поворот вектора $y = (y_1 \mid y_2)^T$ на угол θ по часовой стрелке эквивалентен повороту системы координат против часовой стрелки на тот же угол, легко найти (рис. А.7), что координаты y'_1, y'_2 повернутого вектора $y_\mathrm{r} = (y'_1 \mid y'_2)^T$ определяются в виде $y'_1 = y_1 \cos\theta + y_2 \sin\theta$, $y'_2 = -y_1 \sin\theta + y_2 \cos\theta$.

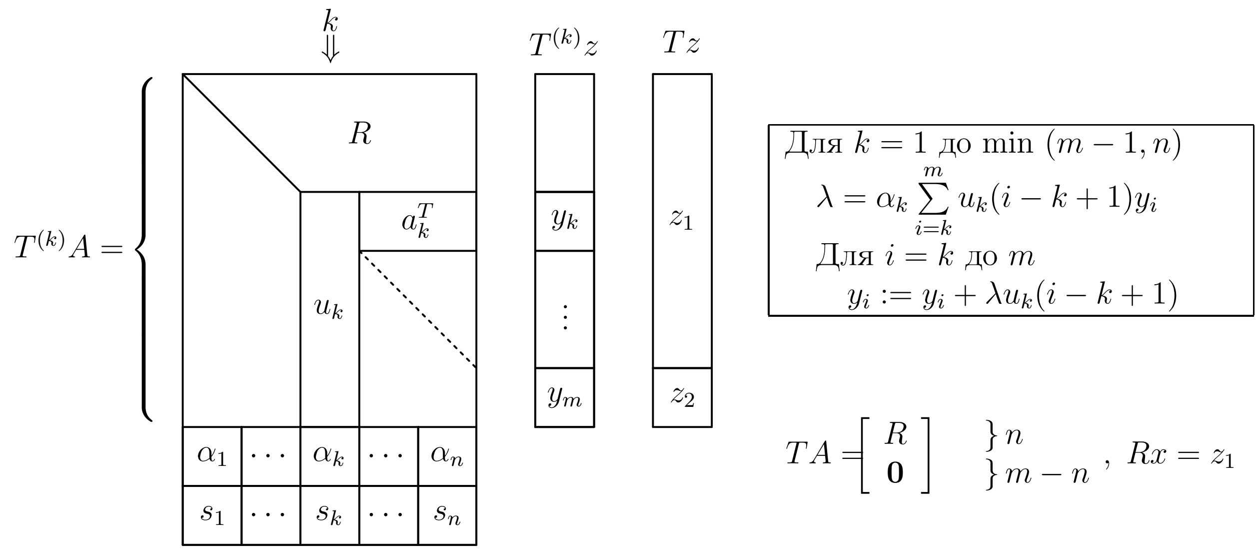


Рис. А.5. Сохранение преобразования T и вычисление вектора $y = Tz$, $\forall y \in \mathbb{R}^m$

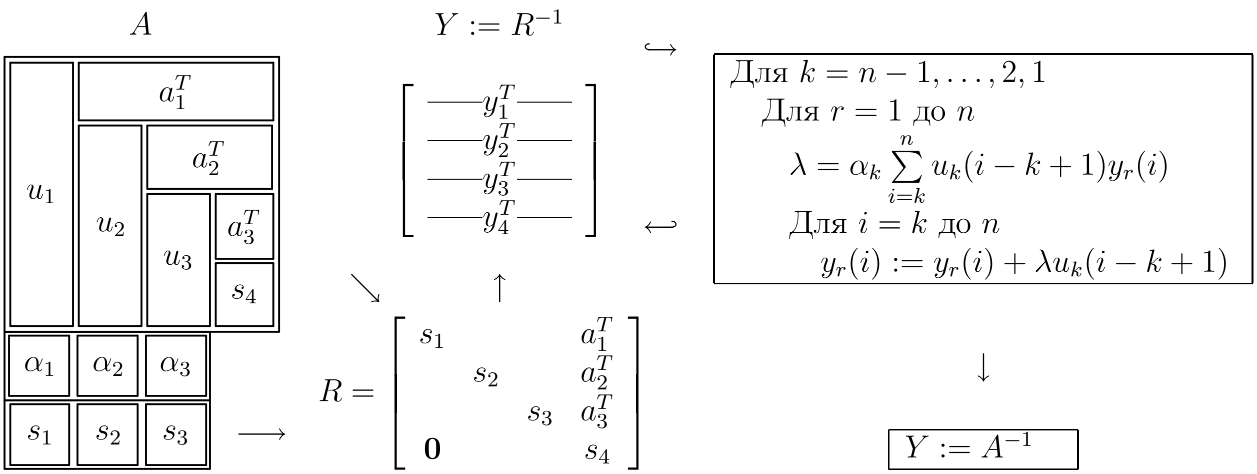


Рис. А.6. Вычисление матрицы A^{-1} после сохранения преобразования T

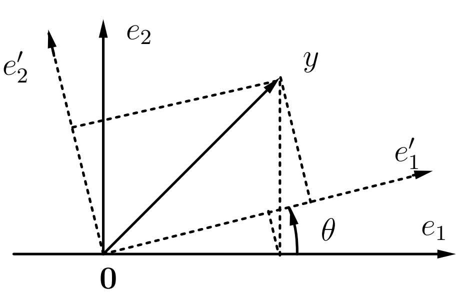


Рис. А.7. Геометрия вращений

Записывая это в матричной форме и требуя, чтобы поворот $P_{1,2}$ в плоскости (e_1, e_2) происходил до совмещения с первой координатной осью, получим

$$y_{\mathrm{r}} = \begin{bmatrix} c & s \\ -s & c \end{bmatrix} y = P_{1,2} y = \begin{bmatrix} r \\ 0 \end{bmatrix}, \quad \left\{ \begin{array}{c} c \triangleq \cos\theta = y_1/r \\ s \triangleq \sin\theta = y_2/r \end{array} \right\}, \quad r \triangleq \sqrt{y_1^2 + y_2^2},$$

где, очевидно, матрица $P_{1,2}$ плоского вращения ортогональна при любом θ.

Триангуляризация матрицы преобразованиями Гивенса. Выбор θ такой, что вторая координата вектора y_{r} становится равной нулю, используем для триангуляризации матрицы $A(m,n)$. На первом шаге нужны преобразования, делающие равными нулю все элементы ниже первого диагонального элемента. Для этого, очевидно, нужно выполнить последовательно элементарные вращения $P_{1,2}$, $P_{1,3}$, $\cdots$, $P_{1,m}$. Так определенные преобразования воздействуют на все столбцы матрицы, но только первый столбец, который уместно назвать *ведущим столбцом*, приобретает желаемый вид.

Лемма А.2. Пусть дана матрица $A(m,n)$ и y — ее ведущий (левый) столбец. Тогда существует ортогональное преобразование Гивенса, задаваемое матрицей $P_1 = P_1(m,m)$, такое, что

$$P_1 A = \begin{array}{r} 1\{ \\ m-1\left\{ \vphantom{\begin{array}{c} 0 \\ 0 \end{array}} \right. \end{array} \! \left[\begin{array}{c|c} \overbrace{\; r \;}^{1} & \overbrace{\qquad}^{n-1} \\ \mathbf{0} & \tilde{A} \end{array} \right], \tag{A.34}$$

$$P_1 = P_{1,m} \cdots P_{1,3} P_{1,2}\ ,$$

$$\begin{array}{ccccc}
P_{1,i}\ (i = 2, 3, \dots, m) & & & & y \\
\begin{array}{c|ccccccc|}
 & 1 & & & i & & & m \\ \hline
1 & c_{1,i} & & & s_{1,i} & & & \\
 & & \ddots & & & & & \\
 & & & 1 & & \mathbf{0} & & \\
 & & & & \ddots & & & \\
i & -s_{1,i} & & & c_{1,i} & & & \\
 & & & & & 1 & & \\
 & & & \mathbf{0} & & & \ddots & \\
m & & & & & & & 1 \\ \hline
\end{array}
& \longleftarrow &
\boxed{\begin{array}{l}
r_{1,1} := y_1 \\
\text{Для } i = 2, 3, \dots, m \\
\quad r_{1,i} := \sqrt{r_{1,i-1}^2 + y_i^2} \\
\quad c_{1,i} := \dfrac{r_{1,i-1}}{r_{1,i}} \\
\quad s_{1,i} := \dfrac{y_i}{r_{1,i}} \\
r \triangleq r_{1,m} = \sum\limits_{i=1}^{m} y_i^2 = \|y\|^2
\end{array}}
& \longleftarrow &
\left[\begin{array}{c} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_m \end{array} \right]
\end{array}$$

Теорема А.2 *(Триангуляризация матрицы по методу Гивенса).* Пусть $A_1 := A(m,n)$ и для каждого $j = 1.2.\ldots,k$, $k \le \min(m-1,n)$ серия элементарных преобразований Гивенса, задаваемая матрицей P_j размера $(m+1-j)\times(m+1-j)$, выбрана для ведущего (левого) столбца y_j матрицы A_j так, что

$$P_j A_j = \begin{array}{r} \\ 1\{ \\ m-j\{ \end{array}\!\!\begin{array}{c} \overbrace{\quad}^{1}\;\;\overbrace{\qquad\quad}^{n-j} \\ \left[\begin{array}{c|c} r_j & a_j^T \\ \mathbf{0} & A_{j+1} \end{array}\right] \end{array}, \quad j = 1,\ldots,k; \quad k \le \min(m-1,n). \quad (\text{A.35})$$

Тогда в процессе после k повторных применений леммы А.2 имеем следующий промежуточный результат триангуляризации матрицы A:

$$P^{(k)}A = \left[\begin{array}{ccccc} r_1 & & & & a_1^T \\ & r_2 & & & a_2^T \\ & & \ddots & & \vdots \\ & & & r_k & a_k^T \\ & \mathbf{0} & & & A_{k+1} \end{array}\right] \qquad (\text{A.36})$$

с отвечающей этому моменту процесса итоговой матрицей преобразований

$$P^{(k)} = \begin{bmatrix} I_{k-1} & 0 \\ 0 & P_k \end{bmatrix} \cdots \begin{bmatrix} I_1 & 0 \\ 0 & P_2 \end{bmatrix} P_1\,, \quad P_j = P_{j,m-j+1}\cdots P_{j,3}P_{j,2}\,, \quad (\text{A.37})$$

$$P_{j,t}\ (t = 2,3,\ldots,l),\ l \triangleq m-j+1,\ j = 1,2,\ldots,k,\ i = j+t-1.$$

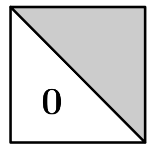
$$\begin{array}{c} \\ 1 \\ \\ \\ t \\ \\ \\ l \end{array}\begin{array}{c} \begin{array}{ccc} 1 & t & l \end{array} \\ \boxed{\begin{array}{ccccccc} c_{j,i} & & & s_{j,i} & & & \\ & \ddots & & & & & \\ & & 1 & & \mathbf{0} & & \\ & & & \ddots & & & \\ -s_{j,i} & & & c_{j,i} & & & \\ & & & & 1 & & \\ & \mathbf{0} & & & & \ddots & \\ & & & & & & 1 \end{array}} \end{array} \longleftarrow \boxed{\begin{array}{l} r_{j,1} := y_{1,j} \\ \text{Для } t = 2,\ldots,l \\ \quad r_{j,t} := \sqrt{r_{j,t-1}^2 + y_{t,j}^2} \\ \quad c_{j,j+t-1} := \dfrac{r_{j,t-1}}{r_{j,t}} \\ \quad s_{j,j+t-1} := \dfrac{y_{t,j}}{r_{j,t}} \\ r_j \triangleq r_{j,l} = \sum\limits_{t=1}^{l} y_{t,j}^2 = \|y_j\|^2 \end{array}} \longleftarrow \begin{array}{c} y_j \\ \Downarrow \\ \boxed{\begin{array}{c} y_{1,j} \\ \\ \\ y_{t,j} \\ \\ \\ y_{l,j} \end{array}} \end{array}$$

Формула (А.37) имеет рекуррентный вид произведения

$$\left.\begin{aligned} P^{(j)} &= \begin{bmatrix} I_{j-1} & 0 \\ 0 & P_j \end{bmatrix} P^{(j-1)}, \quad P^{(1)} = P_1 \\ P_j &= P_{j,m-j+1} \cdots P_{j,3} P_{j,2}\,, \quad j = 2, \ldots, N \end{aligned}\right\}, \; N = \min(m-1, n). \tag{A.38}$$

Все участвующие здесь матрицы являются ортогональными, поэтому финальная матрица $P \triangleq P^{(N)}$ также ортогональна. Общее число используемых при этом элементарных матриц вращения равно $(m-1) + (m-2) + \ldots + (m-N) = (2m-N-1)N/2$. В результате (в случае $m > n$) получим

$$PA = \begin{bmatrix} R \\ \cdots \\ \mathbf{0} \end{bmatrix}, \quad R =$$

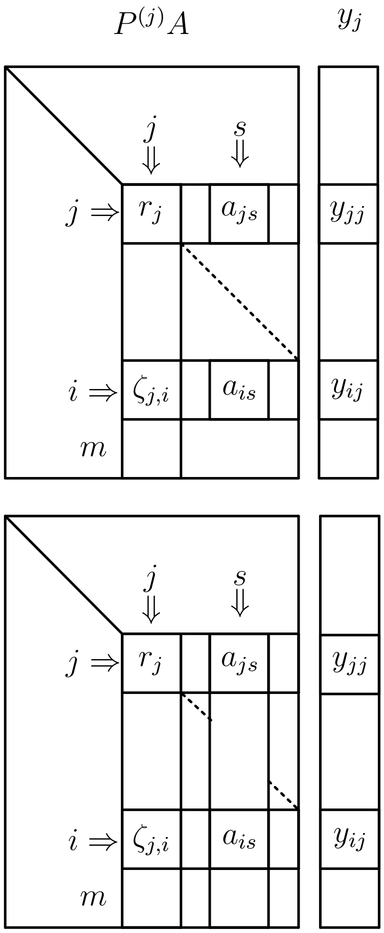


$$\triangleq R_{\mathrm{ne}}, \tag{A.39}$$

где индекс $_{\mathrm{ne}}$ подчеркивает, что в треугольной матрице R заполненной частью может быть только «северо-восточная» (*north-by-east*) часть. Полагая $Q = P^T$, при $m = n$ имеем QR-разложение матрицы $A = A(n,n)$, т. е.$A = QR$. Матрицы в (А.37) и (А.38) непосредственно не вычисляются.

Для алгоритма Гивенса — так же, как и для других матричных алгоритмов, — имеются две схемы вычислений: (1) строчно ориентированная схема и (2) столбцово ориентированная схема (рис. А.8). Как и в алгоритме преобразований Хаусхолдера (см. рис. А.5), здесь обычно требуется сохранять информацию о произведенных по ходу алгоритма (А.38) элементарных преобразованиях, чтобы впоследствии иметь возможность решать системы уравнений $Ax = z$ (совместные или несовместные, в последнем случае — по методу наименьших kdадратов, см. подразд. А.3) или же находить обратную матрицу A^{-1} (когда $m = n$).

Необходимая информация означает значения косинуса и синуса, однако их сохранение было бы неэффективным решением. Gentleman (1973) предложил эффективный способ, включенный в рис. А.8(б) и (г) с геометрической иллюстрацией его действия на рис. А.8(е). Введенный им рабочий признак ζ — это одно число, которое можно хранить в позиции (i,j) как $\zeta_{j,i}$ вместо нулевого элемента, появляющегося в позиции (i,j) матрицы (А.39) в момент преобразования $P_{j,t}$ $(t = i+1-j)$ в (А.37). Как и с преобразованиями Хаусхолдера, нахождение A^{-1} после преобразований Гивенса требует такой же последовательности процедур: сначала находят R^{-1} (см. А.6), затем к R^{-1} применяют с правой стороны финальное преобразование $P \triangleq P^{(N)}$ (А.38), так как $A^{-1} = R^{-1}P$. Для этого на рис. А.6 надо взять алгоритм из рис. А.8(д), который также отыскивает Pz при решении уравнения $Ax = z$.



(а)

Для $j = 1$ до $\min(m-1, n)$
$r_j := a_{jj}$
Для $i = j+1$ до m
$a := r_j;\ b := a_{ij}$
получить (r, c, s, ζ) из (a, b)
$(a_{jj} \triangleq r_j) := r$
$(a_{ij} \triangleq \zeta_{j,i}) := \zeta$
$c_{j,i} := c;\ s_{j,i} := s$
Для $s = j+1$ до n
Для $i = j+1$ до m
$\alpha = c_{j,i}a_{js} + s_{j,i}a_{is}$
$a_{is} = -s_{j,i}a_{js} + c_{j,i}a_{is}$
$a_{js} := \alpha$

(б)

........Процедура........
получить (r, c, s, ζ) из (a, b)

$$\sigma = \begin{cases} \operatorname{sgn}[a], & |a| > |b| \\ \operatorname{sgn}[b], & |a| \le |b| \end{cases}$$

$$r = \sigma(a^2 + b^2)^{1/2}$$

$$c = \begin{cases} a/r, & r \ne 0 \\ 1, & r = 0 \end{cases}$$

$$s = \begin{cases} b/r, & r \ne 0 \\ 0, & r = 0 \end{cases}$$

$$\zeta = \begin{cases} s, & |a| > |b| \\ 1/c, & \begin{cases} |a| \le |b| \ \& \\ \&\ c \ne 0 \end{cases} \\ 1, & c = 0 \end{cases}$$

(в)

Для $j = 1$ до $\min(m-1, n)$
Для $i = j+1$ до m
$a := a_{jj};\ b := a_{ij}$
получить (r, c, s, ζ) из (a, b)
$(a_{jj} \triangleq r_j) := r$
$(a_{ij} \triangleq \zeta_{j,i}) := \zeta$
Для $s = j+1$ до n
$\alpha = ca_{js} + sa_{is}$
$a_{is} = -sa_{js} + ca_{is}$
$a_{js} := \alpha$

(г)

........Процедура........
получить (c, s) из (ζ)

$$\zeta = 1 \ \Rightarrow \begin{cases} c := 0 \\ s := 1 \end{cases}$$

$$d := (1 - \zeta^2)^{1/2}$$

$$\zeta < 1 \ \Rightarrow \begin{cases} c := d \\ s := \zeta \end{cases}$$

$$\zeta > 1 \ \Rightarrow \begin{cases} c := \zeta \\ s := d \end{cases}$$

(д)

Для $j = 1$ до $\min(m-1, n)$
Для $i = j+1$ до m
получить (c, s) из $\zeta_{j,i}$
$\alpha = cy_{jj} + sy_{ij}$
$y_{ij} = -sy_{jj} + cy_{ij}$
$y_{jj} := \alpha$

(е)

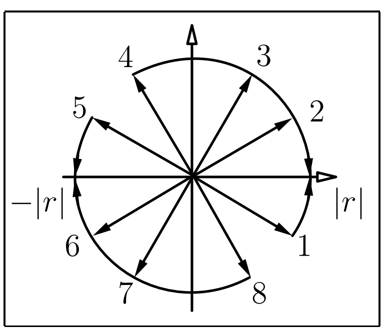
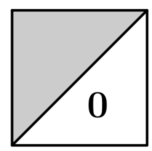


Рис. A.8. Преобразование Гивенса: (а) столбцово ориентированная схема вычисления матрицы PA, где $P = P^{(j)}$ при $j = \min(m-1, n)$ (нижняя матрица слева); (б) вычисление координаты r вектора $(a,\ b)^T$, повернутого до совмещения с первой осью, а также косинуса и синуса угла поворота и рабочего признака ζ; (в) строчно ориентированная схема вычисления матрицы PA (верхняя матрица слева); (г) восстановление косинуса и синуса угла поворота из признака ζ; (д) получение вектора y теми преобразованиями $P_{j,i}$ произвольного вектора $z \in \mathbb{R}^m$, которые сохранены в рабочих признаках $\zeta_{j,i}$ и восстанавливаются из них; (е) вследствие (б), векторы 1, 2, 3 и 4 поворачиваются к положительному направлению первой координатной оси, а векторы 5, 6, 7 и 8 поворачиваются к отрицательному направлению этой оси.

A.8 Варианты заполнения матрицы R

Традиционно, любые ортогональные преобразования (выше рассмотрены T — преобразование Хаусхолдера и P — преобразование Гивенса, ниже будет рассмотрено Q — преобразование Грама-Шмидта) приводят матрицу к виду, показанному на рис. А.4 или в выражении (А.39). Однако выбор того угла матрицы, который должен остаться треугольно заполненым, естественно, произволен. Предпочтения диктуются целями использования, т. е.предназначением преобразования. Преследуя здесь учебно-тренировочные цели, включим в проект (см. подразд. **??**) все четыре возможных варианта заполнения матрицы R, а именно: первый вариант показан в (А.39), следующие три имеют вид

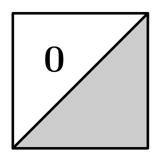
$$PA = \begin{bmatrix} R \\ \cdots \\ \mathbf{0} \end{bmatrix}, \quad R =$$



$$\triangleq R_{\mathrm{nw}}, \qquad \text{(A.40)}$$

где индекс $_{\mathrm{nw}}$ подчеркивает, что в треугольной матрице R заполненной частью может быть только «северо-западная» (*north-by-west*) часть,

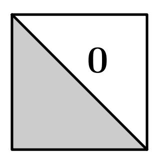
$$PA = \begin{bmatrix} \mathbf{0} \\ \cdots \\ R \end{bmatrix}, \quad R =$$



$$\triangleq R_{\mathrm{se}}, \qquad \text{(A.41)}$$

где индекс $_{\mathrm{se}}$ подчеркивает, что в треугольной матрице R заполненной частью может быть только «юго-восточная» (*south-by-east*) часть, и

$$PA = \begin{bmatrix} \mathbf{0} \\ \cdots \\ R \end{bmatrix}, \quad R =$$



$$\triangleq R_{\mathrm{sw}}, \qquad \text{(A.42)}$$

где индекс $_{\mathrm{sw}}$ подчеркивает, что в треугольной матрице R заполненной частью может быть только «юго-западная» (*south-by-west*) часть. Вполне очевидно, что эти вариенты получаются простым изменение порядка действий в алгоритмах преобразований.

A.9 Правосторонние ортогональные преобразования и их применение

С правосторонними ортогональными преобразованиями мы уже сталкивались (см. подразд. А.6); тогда для квадратной матрицы A после $TA = R$

вычисляли $A^{-1} = R^{-1}T$. Однако, можно начинать с правостороннего преобразования матрицы A; тогда отыскание A^{-1} потребует, соответственно, левостороннего преобразования.

Пусть $A = A(n, n)$ — квадратная невырожденная матрица. Будем рассматривать ее строки как векторы в $\mathbb{R}^n$. Преобразования вектора как матрицы-строки в n-мерном линейном пространстве задаются умножением ее на преобразующую матрицу справа. Поэтому правосторонним ортогональным преобразованием Q можно привести матрицу A к виду $AQ = R$, где применена ортогональная матрица Q одного из типов, а R — треугольная матрица, имеющая форму одного из возможных вариантов заполнения (см. подразд. А.9). При этом преобразованию Q подвергаются не столбцы, а строки матрицы A, и преобразование Q запоминается по принципу, показанному ранее на рис. А.5 и рис. А.8, на месте элементов, обращаемых в нуль.

После такого преобразования матрицы A решение системы $Ax = z$ сводится к решению эквивалентной системы с треугольной матрицей $Ry = z$. Затем искомый вектор определяется через сохраненное преобразование Q как $x = Qy$. Обратная матрица A^{-1}, соответственно, находится как решение системы $RY = I$ с последующим преобразованием Q матрицы Y, т. е.$X = A^{-1} = QY$. Матрица Q не формируется, из чего видна необходимость запоминания преобразований, обеспечивших $AQ = R$.

А.10 Двусторонние ортогональные преобразования и их применение

Ортогональные преобразования, будучи применены одновременно слева и справа к данной матрице A, позволяют приводить ее к формам с нулями как ниже, так и выше диагонали. Это, в свою очередь, облегчает решение других сложных задач. С помощью ортогональных преобразований для квадратной матрицы широко распространены: приведение симметрической матрицы к трехдиагональному виду и приведение квадратной матрицы к двухдиагональному виду. При этом в качестве ортогональных преобразований одинаково успешно могут быть использованы преобразования Хаусхолдера или преобразования Гивенса.

Приведение симметрической матрицы к трехдиагональному виду. Применим к симметрической матрице слева и справа преобразование Хаусхолдера (или Гивенса), выбирая его из задачи желаемого преобразования ведущего столбца и ведущей строки, а именно: сохранение первого диагонального элемента, получение ненулевых элементов в двух смежных с ним позициях и получение нулевых элементов в остальных позициях.

Лемма А.3. Пусть дана матрица $A = A(n,n) = A^T$. Тогда существует ортогональное преобразование Q_2 (Хаусхолдера T_2 или Гивенса P_2) такое, что

$$\left[\begin{array}{c|c} I_1 & \mathbf{0} \\ \mathbf{0} & Q_2 \end{array}\right] A \left[\begin{array}{c|c} I_1 & \mathbf{0} \\ \mathbf{0} & Q_2^T \end{array}\right] = \begin{array}{r} \\ 1\{ \\ 1\{ \\ n-2\{ \end{array}\!\!\begin{array}{c} \overbrace{}^{1}\;\overbrace{}^{1}\;\overbrace{\phantom{\mathbf{0}}}^{n-2} \\ \left[\begin{array}{c|cc} a_1 & s_1 & \mathbf{0} \\ \hline s_1 & & \\ \mathbf{0} & & \tilde{A} \end{array}\right] \end{array}. \tag{А.43}$$

Замечание А.6. В (А.43) транспонирование Q_2^T не требуется, если в качестве Q_2 взято преобразование Хаусхолдера (в силу его симметричности). При этом индекс «$_2$» указывает на позицию того элемента в ведущем столбце (для левостороннего преобразования) или в ведущей строке (для правостороннего преобразования), который остается ненулевым в этом столбце (в результате применения Q_2) или в этой строке (в результате применения Q_2^T). В данном случае, т. е.в (А.43), эти элементы суть s_1 и s_1. Элемент a_1 не изменяется, так как I_1 — единичная матрица размера 1×1.

Теорема А.3 *(Тридиагонализация симметрической матрицы).* Пусть дана симметрическая матрица $A = A(n,n) = A^T$, $A_1 := A(n,n)$ и для каждого $j = 1, \ldots, k$, где $k \le N = n-2$, выбрано элементарное преобразование Q_{j+1} (Хаусхолдера T_{j+1} или Гивенса P_{j+1}) так, что

$$\left[\begin{array}{c|c} I_1 & \mathbf{0} \\ \mathbf{0} & Q_{j+1} \end{array}\right] A_j \left[\begin{array}{c|c} I_1 & \mathbf{0} \\ \mathbf{0} & Q_{j+1}^T \end{array}\right] = \begin{array}{r} \\ 1\{ \\ 1\{ \\ n-j-1\{ \end{array}\!\!\begin{array}{c} \overbrace{}^{1}\;\overbrace{}^{1}\;\overbrace{\phantom{\mathbf{0}\,\mathbf{0}}}^{n-j-1} \\ \left[\begin{array}{c||cc} a_j & s_j & \mathbf{0} \\ \hline s_j & & \\ \mathbf{0} & & A_{j+1} \end{array}\right] \end{array}. \tag{А.44}$$

Тогда в процессе после k повторных применений леммы А.3 имеем отвечающую этому моменту процесса итоговую матрицу преобразований

$$Q^{(k)} = \begin{bmatrix} I_k & 0 \\ 0 & Q_{k+1} \end{bmatrix} Q^{(k-1)}, \quad 1 \le k \le N = n-2, \quad Q^{(0)} = I_n \tag{А.45}$$

и промежуточный результат тридиагонализации данной матрицы A в виде

$$
Q^{(k)} A \left(Q^{(k)}\right)^T = \left[\begin{array}{cccccc|c}
a_1 & s_1 & & & & & \\
s_1 & a_2 & s_2 & & & \mathbf{0} & \\
 & s_2 & \ddots & \ddots & & & \\
 & & \ddots & a_k & s_k & & \\
\cline{6-7}
 & & & s_k & & & \\
 & \mathbf{0} & & & & A_{k+1} & \\
 & & & & & &
\end{array}\right].
$$

Приведение квадратной матрицы к двухдиагональному виду. Применим к произвольной квадратной матрице слева преобразование Q_1 и справа преобразование S_2 (беря любое из них как преобразование Хаусхолдера или как преобразование Гивенса), при этом Q_1 выберем из задачи желаемого преобразования ведущего столбца и S_2 — из задачи желаемого преобразования ведущей строки, а именно: при действии Q_1 — получение ненулевого диагонального элемента и нулевых элементов ниже него в первом (ведущем) столбце; при действии S_2 — сохранение диагонального элемента, получение в смежной с ним позиции ненулевого элемента и нулевых элементов правее него в первой (ведущей) строке.

Лемма А.4. Пусть дана матрица $A = A(n, n)$. Тогда существуют ортогональное преобразование Q_1 (Хаусхолдера или Гивенса) и ортогональное преобразование S_2 (Хаусхолдера или Гивенса) такие, что

$$
Q^{(1)} A S^{(1)} = \begin{array}{r}
\\
1\{ \\
n-2\{
\end{array}
\begin{array}{c}
\overbrace{\qquad}^{1}\;\overbrace{\qquad}^{1}\;\overbrace{\qquad}^{n-2} \\
\left[\begin{array}{c|cc}
s_1 & a_1 & \mathbf{0} \\
\cline{2-3}
\mathbf{0} & \multicolumn{2}{c}{\tilde{A}}
\end{array}\right]
\end{array},
\qquad
\begin{cases}
Q^{(1)} = Q_1, \\
S^{(1)} = \left[\begin{array}{c|c} I_1 & \mathbf{0} \\ \mathbf{0} & S_2 \end{array}\right].
\end{cases}
\tag{А.46}
$$

Теорема А.4 *(Бидиагонализация квадратной матрицы).* Пусть дана квадратная матрица $A = A(n, n)$, $A_1 := A$ и для каждого $j = 1, \ldots, k$, где $k \leq n - 2$, выбраны элементарное преобразование Q_j (Хаусхолдера типа T_j или Гивенса типа P_j) и элементарное преобразование S_{j+1} (Хаусхолдера типа T_{j+1} или Гивенса типа P_{j+1}) таким образом, что в результате получаем

$$
Q_j A_j \left[\begin{array}{c|c} I_1 & \mathbf{0} \\ \mathbf{0} & S_{j+1} \end{array}\right] =
\begin{array}{r}
\\
1\{ \\
n-j\Big\{
\end{array}
\begin{array}{c}
\begin{array}{ccc} \overbrace{\qquad}^{1} & \overbrace{\qquad}^{1} & \overbrace{\qquad\quad}^{n-j-1} \end{array} \\
\left[\begin{array}{c|cc} s_j & a_j & \mathbf{0} \\ \hline \mathbf{0} & \multicolumn{2}{c}{A_{j+1}} \end{array}\right]
\end{array}. \tag{A.47}
$$

Тогда в процессе после k повторных применений леммы A.4 имеем отвечающие этому моменту процесса итоговые матрицы преобразований

$$
\left.
\begin{array}{l}
Q^{(k)} = \begin{bmatrix} I_{k-1} & 0 \\ 0 & Q_k \end{bmatrix} Q^{(k-1)}, \quad k \le n-2, \quad Q^{(0)} = I_n, \quad Q^{(1)} = Q_1, \\
S^{(k)} = S^{(k-1)} \begin{bmatrix} I_k & 0 \\ 0 & S_{k+1} \end{bmatrix}, \quad k \le n-2, \quad S^{(1)} = \begin{bmatrix} I_1 & 0 \\ 0 & S_2 \end{bmatrix}
\end{array}
\right\} \tag{A.48}
$$

и промежуточный результат бидиагонализации данной матрицы A в виде

$$
Q^{(k)} A S^{(k)} = \left[\begin{array}{ccccc|c}
s_1 & a_1 & & & & \\
& s_2 & a_2 & & & \mathbf{0} \\
& & \ddots & \ddots & & \\
& & & s_k & a_k & \\ \cline{6-6}
& & \mathbf{0} & & & A_{k+1}
\end{array}\right].
$$

Выполнив после $k = n-2$ еще одно левостороннее преобразование Q_{n-1} (что отвечает применению верхней формулы (A.48) для $k = n-1$), получаем окончательно

$$
Q^{(n-1)} A S^{(n-2)} = \begin{bmatrix}
s_1 & a_1 & & & \\
& s_2 & a_2 & & \mathbf{0} \\
& & \ddots & \ddots & \\
& & & s_{n-1} & a_{n-1} \\
& \mathbf{0} & & & s_n
\end{bmatrix}.
$$

Приложение В

Название приложения В

В.1 Название подраздела В.1 в приложении В

Как показано в приложении В, подразд. В.2, пункт В.2.1 ...

В.1.1 Название пункта В.1.1 в приложении В, подразд. В.1

Название подпункта пункта В.1.1 в приложении В, подразд. В.1

В.1.2 Название

Название

В.2 Название подраздела в приложении 2

В.2.1 Название

Название

В.2.2 Название

Название