

Preface

前言

感谢

首先感谢大家的信任。

作者仅仅是在学习应用数学科学和机器学习算法时，多读了几本数学书，多做了些思考和知识整理而已。知者不言，言者不知。知者不博，博者不知。水平有限，把自己有限所学所思斗胆和大家分享，作者权当无知者无畏。希望大家在 B 站视频下方和 Github 多提意见，让这套书成为作者和读者共同参与创作的优质作品。

特别感谢清华大学出版社的栾大成老师。从选题策划、内容创作、装帧设计，栾老师事无巨细、一路陪伴。每次和栾老师交流，我都能感受到他对优质作品的追求、对知识分享的热情。

出来混总是要还的

曾经，考试是我们学习数学的唯一动力。考试是头悬梁的绳，是锥刺股的锥。我们中的绝大多数人从小到大为各种考试埋头题海，数学味同嚼蜡，甚至让人恨之入骨。

数学给我们带来了无尽的折磨。我们憎恨数学，恐惧数学，恨不得一走出校门就把数学抛之脑后、老死不相往来。

可悲可笑的是，我们其中很多人可能会在毕业的五年或十年以后，因为工作需要，不得不重新学习微积分、线性代数、概率统计，悔恨当初没有学好数学、走了很多弯路、没能学以致用，从而迁怒于教材和老师。

这一切不能都怪数学，值得反思的是我们学习数学的方法、目的。

再给自己一个学数学的理由

为考试而学数学，是被逼无奈的举动。而为数学而数学，则又太过高尚而遥不可及。

相信对于绝大部分的我们来说，数学是工具、是谋生手段，而不是目的。我们主动学数学，是想用数学工具解决具体问题。

现在，这套书给大家一个“学数学、用数学”的全新动力——数据科学、机器学习。

数据科学和机器学习已经深度融合到我们生活的方方面面，而数学正是开启未来大门的钥匙。不是所有人生来都握有一副好牌，但是掌握“数学 + 编程 + 机器学习”绝对是王牌。这次，学习数学不再是为了考试、分数、升学，而是投资时间、自我实现、面向未来。

未来已来，你来不来？

本套丛书如何帮到你

为了让大家学数学、用数学，甚至爱上数学，作者可谓颇费心机。在创作这套书时，作者尽量克服传统数学教材的各种弊端，让大家学习时有兴趣、看得懂、有思考、更自信、用得着。

为此，丛书在内容创作上突出以下几个特点：

- ◀ **数学 + 艺术**——全彩图解，极致可视化，让数学思想跃然纸上、生动有趣、一看就懂，同时提高大家的数据思维、几何想象力、艺术感；
- ◀ **零基础**——从零开始学习 Python 编程，从写第一行代码到搭建数据科学和机器学习应用；
- ◀ **知识网络**——打破数学板块之间的壁垒，让大家看到数学代数、几何、线性代数、微积分、概率统计等板块之间的联系，编织一张绵密的数学知识网络；
- ◀ **动手**——授人以鱼不如授人以渔，和大家一起写代码、用 Streamlit 创作数学动画、交互 App；
- ◀ **学习生态**——构造自主探究式学习生态环境“微课视频 + 纸质图书 + 电子图书 + 代码文件 + 可视化工具 + 思维导图”，提供各种优质学习资源；
- ◀ **理论 + 实践**——从加减乘除到机器学习，丛书内容安排由浅入深、螺旋上升，兼顾理论和实践；在编程中学习数学，学习数学时解决实际问题。

虽然本书标榜“从加减乘除到机器学习”，但是建议读者朋友们至少具备高中数学知识。如果读者正在学习或曾经学过大学数学（微积分、线性代数、概率统计），这套书就更容易读了。

聊聊数学

数学是工具。锤子是工具，剪刀是工具，数学也是工具。

数学是思想。数学是人类思想的高度抽象的结晶体。在其冷酷的外表之下，数学的内核实际上就是人类朴素的思想。学习数学时，知其然，更要知其所以然。不要死记硬背公式定理，理解背后的数学思想才是关键。如果你能画一幅图、用大白话描述清楚一个公式、一则定理，这就说明你真正理解了它。

数学是语言。就好比世界各地不同种族有自己的语言，数学则是人类共同的语言和逻辑。数学这门语言极其精准、高度抽象，放之四海而皆准。虽然我们中绝大多数人没有被数学女神选中，不能为人类的对数学认知开疆扩土；但是，这丝毫不妨碍我们使用数学这门语言。就好比，我们不会成为语言学家，我们完全可以使用母语和外语交流。

数学是体系。代数、几何、线性代数、微积分、概率统计、优化方法等等，看似一个个孤岛，实际上都是数学网络的一条条织线。建议大家学习时，特别关注不同数学板块之间的联系，见树，更要见林。

数学是基石。拿破仑曾说“数学的日臻完善和这个国强民富息息相关。”数学是科学进步的根基，是经济繁荣的支柱，是保家卫国的武器，是探索星辰大海的航船。

数学是艺术。数学和音乐、绘画、建筑一样，都是人类艺术体验。通过可视化工具，我们会在看似枯燥的公式、定理、数据背后，发现数学之美。

数学是历史，是人类共同记忆体。“历史是过去，又属于现在，同时在指引未来。”数学是人类的集体学习思考，她把人的思维符号化、形式化，进而记录、积累、传播、创新、发展。从甲

骨、泥板、石板、竹简、木牍、纸草、羊皮卷、活字印刷、纸质书，到数字媒介，这一过程持续了数千年，至今绵延不息。

数学是无穷无尽的**想象力**，是人类的**好奇心**，是自我挑战的**毅力**，是一个接着一个的**问题**，是看似荒诞不经的**猜想**，是一次次胆大包天的**批判性思考**，是敢于站在前人的肩膀之上的**勇气**，是孜孜不倦地延展人类认知边界的**不懈努力**。

家园、诗、远方

诺瓦利斯曾说：“哲学就是怀着一种乡愁的冲动到处去寻找家园。”

在纷繁复杂的尘世，数学纯粹的就像精神的世外桃源。数学是，一束光，一条巷，一团不灭的希望，一股磅礴的力量，一个值得寄托的避风港。

打破陈腐的锁链，把功利心暂放一边，我们一道怀揣一分乡愁、心存些许诗意、踩着艺术维度，投入数学张开的臂膀，驶入她色彩斑斓、变幻无穷的深港，感受久违的归属，一睹更美、更好的远方。

Acknowledgement

致谢

To my parents.

谨以此书献给我的母亲父亲

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

How to Use the Book

使用本书

丛书资源

本系列丛书提供的配套资源有以下几个：

- ◀ 纸质图书；
- ◀ PDF 文件，方便移动终端学习；请大家注意，纸质图书经过出版社五审五校修改，内容细节上会和 PDF 文件有出入。
- ◀ 每章提供思维导图，纸质书提供全书思维导图海报；
- ◀ Python 代码文件，直接下载运行，或者复制、粘贴到 Jupyter 运行；
- ◀ Python 代码中有专门用 Streamlit 开发数学动画和交互 App 的文件；
- ◀ 微课视频，强调重点、讲解难点、聊聊天。

在纸质书中为了方便大家查找不同配套资源，作者特别设计了如下几个标识。



微课视频

本书配套微课视频均发布在 B 站——生姜 DrGinger：

- ◀ <https://space.bilibili.com/513194466>

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

微课视频是以“聊天”的方式，和大家探讨某个数学话题的重点内容，讲讲代码中可能遇到的难点，甚至侃侃历史、说说时事、聊聊生活。

本书配套的微课视频目的是引导大家自主编程实践、探究式学习，并不是“照本宣科”。

纸质图书上已经写得很清楚的内容，视频课程只会强调重点。需要说明的是，图书内容不是视频的“逐字稿”。

代码文件

本系列丛书的 Python 代码文件下载地址为：

► <https://github.com/Visualize-ML>

Python 代码文件会不定期修改，请大家注意更新。图书配套的 PDF 文件和勘误也会上传到这个 GitHub 账户。因此，建议大家注册 GitHub 账户，给书稿文件夹标星 (star) 或分支克隆 (fork)。

考虑再三，作者还是决定不把代码全文印在纸质书中，以便减少篇幅，节约用纸。

本书编程实践例子中主要使用“鸢尾花数据集”，数据来源是 Scikit-learn 库、Seaborn 库。此外，系列丛书封面设计致敬梵高《鸢尾花》，要是给本系列丛书起个昵称的话，作者乐见“鸢尾花书”。

App 开发

本书几乎每一章都至少有一个用 Streamlit 开发的 App，用来展示数学动画、数据分析、机器学习算法。

Streamlit 是个开源的 Python 库，能够方便快捷搭建、部署交互型网页 App。Streamlit 非常简单易用、很受欢迎。Streamlit 兼容目前主流的 Python 数据分析库，比如 NumPy、Pandas、Scikit-learn、PyTorch、TensorFlow 等等。Streamlit 还支持 Plotly、Bokeh、Altair 等交互可视化库。

本书中很多 App 设计都采用 Streamlit + Plotly 方案。此外，本书专门配套教学视频手把手和大家一起做 App。

大家可以参考如下页面，更多了解 Streamlit：

► <https://streamlit.io/gallery>

► <https://docs.streamlit.io/library/api-reference>

实践平台

本书作者编写代码时采用的 IDE (integrated development environment) 是 Spyder，目的是给大家提供简洁的 Python 代码文件。

但是，建议大家采用 JupyterLab 或 Jupyter notebook 作为本系列丛书配套学习工具。

简单来说，Jupyter 集合“浏览器 + 编程 + 文档 + 绘图 + 多媒体 + 发布”众多功能与一身，非常适合探究式学习。

运行 Jupyter 无需 IDE，只需要浏览器。Jupyter 容易分块执行代码。Jupyter 支持 inline 打印结果，直接将结果图片打印在分块代码下方。Jupyter 还支持很多其他语言，比如 R 和 Julia。

使用 markdown 文档编辑功能，可以编程同时写笔记，不需要额外创建文档。Jupyter 中插入图片和视频链接都很方便。此外，还可以插入 Latex 公式。对于长文档，可以用边栏目录查找特定内容。

Jupyter 发布功能很友好，方便打印成 HTML、PDF 等格式文件。

Jupyter 也并不完美，目前尚待解决的问题有几个。Jupyter 中代码调试不方便，需要安装专门插件（比如 debugger）。Jupyter 没有 variable explorer，要么 inline 打印数据，要么将数据写到 csv 或 Excel 文件中再打开。图像结果不具有交互性，比如不能查看某个点的值，或者旋转 3D 图形，可以考虑安装 (jupyter-matplotlib)。注意，利用 Altair 或 Plotly 绘制的图像支持交互功能。对于自定义函数，目前没有快捷键直接跳转到其定义。但是，很多开发者针对这些问题都开发了插件，请大家留意。

大家可以下载安装 Anaconda，JupyterLab、Spyder、PyCharm 等常用工具都集成在 Anaconda 中。下载 Anaconda 的地址为：

◀ <https://www.anaconda.com/>

学习步骤

大家可以根据自己的偏好制定学习步骤，本书推荐如下步骤。



学完每章后，大家可以在平台上发布自己的 Jupyter 笔记，进一步听取朋友们的意见，共同进步。这样做还可以提高自己学习的动力。

意见建议

欢迎大家对本系列丛书提意见和建议，丛书专属邮箱地址为：

◀ jiang.visualize.ml@gmail.com

也欢迎大家在 B 站视频下方留言互动。

Contents

目录



0.1 本册在全套丛书的定位

鸢尾花书有三大板块——编程、数学、实践。本册《可视之美》是“编程”板块的第二册。上一册《编程不难》着重介绍如何零基础入门学 Python 编程，《可视之美》则在《编程不难》基础之上深入探讨如何用 Python 完成数学、数据可视化。

学习《可视之美》时，希望大家能够掌握各种可视化方案实现手段，但是没有要求深究其背后的数学思想、数学工具。这一点和《编程不难》类似，即“知其然，不需要知其所以然”。

和《编程不难》不同的是，为了缩减篇幅，《可视之美》正文不再提供代码示例，也不再提供练习题目。大家可以在《可视之美》配套的 Jupyter Notebook 中找到相关代码以及注释。

学完《编程不难》、《可视之美》这两册，大家便可以踏上学习“数学三剑客”之路，这时就需要大家知其然，知其所以然。

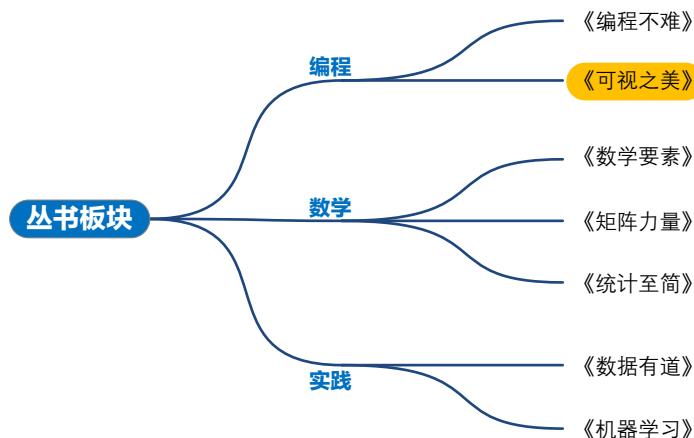


图 1. 鸢尾花书板块布局

0.2 结构：6 大板块

本书一共有 36 章，可以归纳为 6 大板块——图说、美化、色彩、二维、三维、数学之美。

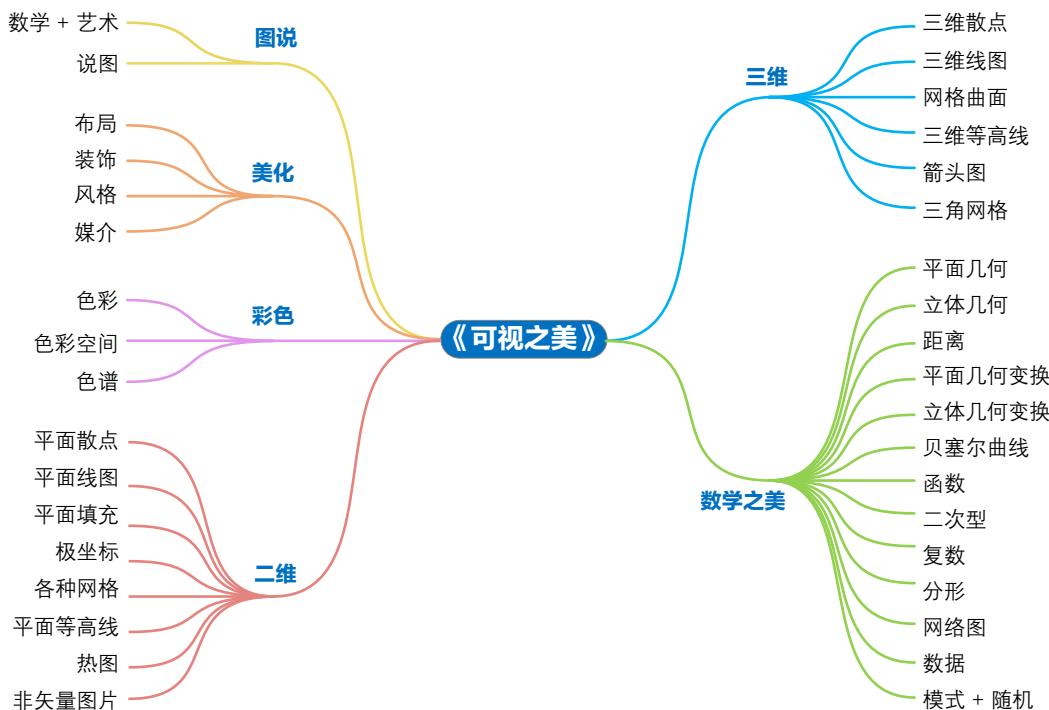


图 2. 《可视之美》板块布局

图说

这部分有两章内容，分别务虚、务实。第 1 章“形而上”地探讨了数学和艺术在解构、重构世界角度的相同之处。第 2 章正式开启了一场“数学 + 艺术”的美学实践之旅，“形而下”地和大家探讨一张图整个生命周期要经过的几个阶段。

美化

“美化”这部分有四章，第 3、4、5 章分别从图片布局、装饰、风格三个角度介绍如何让一幅图尽善尽美。第 6 章则介绍可视化的不同媒介，比如矢量图、非矢量图、Plotly 可交互图片、网页、GIF 动图、Streamlit App 应用等。

想要学好这四章，建议多写代码，多画图。除了书中介绍的案例外，建议大家多多查找 Python 可视化库 (Matplotlib、Seaborn、Plotly、Bokeh、ProPlot 等) 的技术文档，多尝试不同的可视化方案和美化设计。

色彩

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

“彩色”有三章，第 7 章主要介绍了在 Matplotlib 如何定义常用颜色。第 8 章首先介绍了 RGB 色彩模型，鸢尾花书《矩阵力量》还会用到 RGB 色彩模型来讲解向量空间。然后，这一章又介绍了 HSV 色彩模型，HSV 色彩模型的数学本质是圆柱坐标系。第 9 章介绍色谱，即颜色映射。色谱的本质就是函数，将一组数值映射到不同颜色上。

二维

这一板块有八章，主要探讨常用的二维可视化方案：平面散点图（第 10 章）、平面线图（第 11 章）、平面填充（第 12 章）、极坐标（第 13 章）、各种网格（第 14 章）、平面等高线（第 15 章）、热图（第 16 章）、非矢量图片（第 17 章）。

大家学习时要注意，虽然本书讲的是可视化，但是一幅幅图片背后全部都是数学。也就是说，把“艺术”二字拆开来看，数学就是“艺”，而利用 Python 实现可视化方案仅仅是“术”罢了。

三维

“三维”板块是“二维”的升维。这个板块有六章，介绍常用三维可视化方案：三维散点图（第 18 章）、三维线图（第 19 章）、网格曲面（第 20 章）、三维等高线（第 21 章）、箭头图（第 22 章）、三角网格（第 23 章）。注意，第 22、23 章既有二维，也有三维可视化方案。

数学之美

这个板块相当于是全书的实践，利用丰富的可视化方案向大家展示数学之美。“数学之美”这个板块包含 13 个话题，它们分别是：平面几何（第 24 章）、立体几何（第 25 章）、距离（第 26 章）、平面几何变换（第 27 章）、立体几何变换（第 28 章）、贝塞尔曲线（第 29 章）、函数（第 30 章）、二次型（第 31 章）、复数（第 32 章）、分形（第 33 章）、网络图（第 34 章）、数据（第 35 章）、模式 + 随机（第 36 章）。

大家在学习这个板块时多多关注实现这些可视化方案的编程代码，背后的数学原理不在本书核心内容之列。作为《可视之美》的最后一个板块，“数学之美”也引出了鸢尾花书“数学三剑客”——《数学要素》、《矩阵力量》、《统计至简》。在这三本书中，大家将学习“数学之美”采用的大部分数学工具。

0.3 特点：数学之美

万物皆可数学，数学皆可艺术。

《可视之美》可以作为大家的可视化之书、想象力之书、创造实践之书。《可视之美》选图的标准只有一个——尽显数学之美。

《可视之美》用艺术维度打开一扇发现数学之美的窗口。如果在观察某一幅图时，大家特别想要搞清楚其背后数学工具的原理，并且真的付诸行动；那么《可视之美》这本书便目的达成、物有所值！因此，希望大家在阅读本册时要，勤于思考、动手实践。

《可视之美》是鸢尾花书系列中唯一一本真正意义上的“图册”。在这本书中，大家会发现有关科技制图、计算机图形学、创意编程、趣味数学实验、数学科学、机器学习等内容。鸢尾花书其他分册限于篇幅，也限于其核心故事链，不能“肆无忌惮”地给出各种可视化方案，但是《可视之美》没有这样的限制。《可视之美》就是要“毫无节制”地展示数学之美。

可视之美，数学之美；眼见为虚，动手为实。

下面正式邀请大家踏上探索数学之美的旅途。

1

Mathematics + Art

数学 + 艺术

人类存在本质的深层体验和思考



梦里，我梦见我的画作；现实，我手绘我的梦境。

I dream my painting and I paint my dream.

——文森特·梵高 (Vincent van Gogh) | 荷兰后印象派画家 | 1853 ~ 1890

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

1.1 数学 + 艺术

正如本书前言提到的，数学和音乐、绘画、建筑一样，都是人类艺术体验。通过可视化工具，我们会在看似枯燥的公式、定理、数据背后，发现数学之美。

艺术与生俱来。毕加索曾说“每个孩子都是艺术家。问题在于他长大后如何保持艺术家的本质。”几乎所有的孩子在学会读写之前都喜欢涂鸦，这些行为本身都是在无序中创造有序，在无形中创造有形的艺术表达。

没有艺术品可以完全原创。每一件艺术品都是艺术锁链上重要一环，它承前启后。条条艺术锁链织成一张网，每件艺术品都有自己特殊的位置。这一点和数学极为相似。代数、几何、线性代数、微积分、概率统计、优化方法等等，看似一个个孤岛，实际上都是数学网络的一条条织线。

这一章要“形而上”地探讨艺术和数学之间存在着紧密的联系，尽管它们在表现方式上看上去相去甚远。

1.2 物质世界是几何的世界

人类执着于几何，就是因为物质世界本身就是个几何的世界。数学中的几何学原理在艺术中发挥重要作用。艺术家使用几何形状、比例（比如黄金分割比）和对称来构建他们的作品，创造出视觉上的平衡和美感。

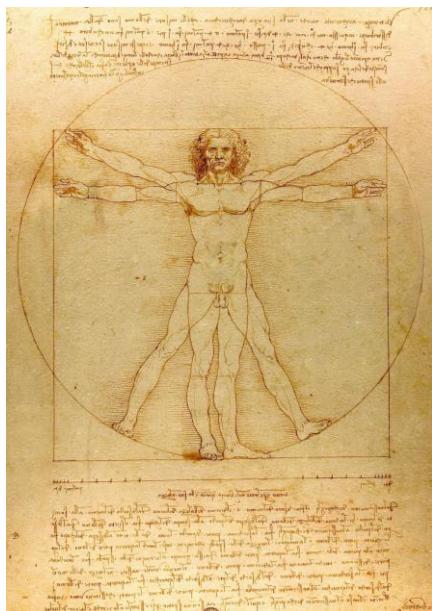


图 1. 达芬奇的《维特鲁威人》，现藏于意大利威尼斯学院美术馆；图片来源：Wikipedia

色彩在艺术中起着重要作用，而色彩的相互关系可以通过数学原理来解释和理解。色彩也是几何，也是空间。RGB 色彩空间就可以看成是一个由红、绿、蓝三色撑起来的正立方体。三色光通过不同比例组合幻化成无数的色彩。而不同色彩放在一起既可以产生和谐，又可以产生视觉的冲突。

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

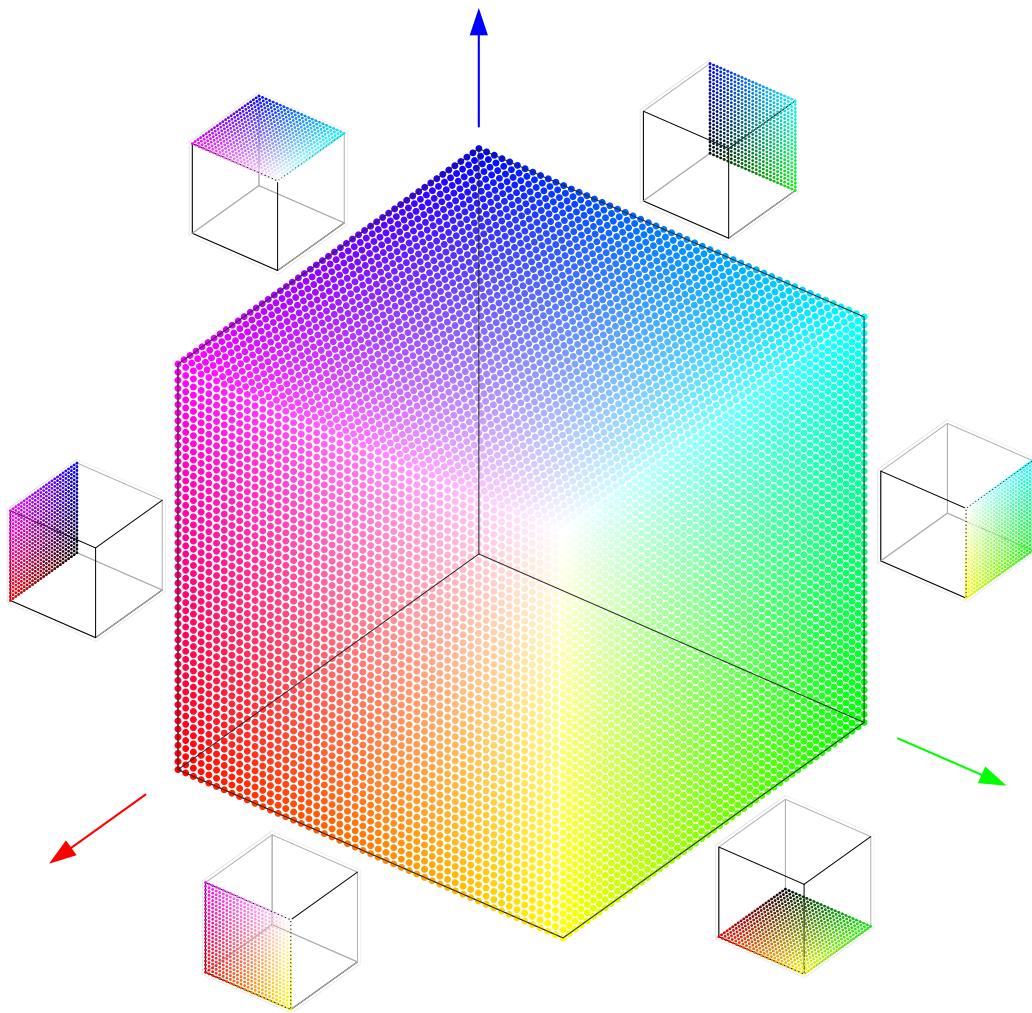
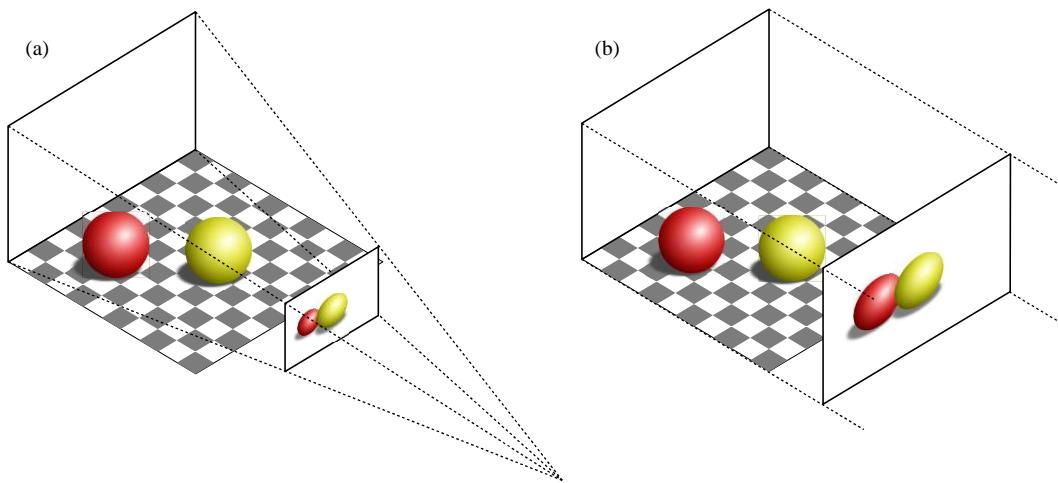


图 2. RGB 色彩空间

有几何的地方，就有空间。人类绘画经历了从二维到三维的维度提升的过程，这是一个关于空间感知和透视的发展历程。

透视是绘画中常用的技巧，用于创造画面中的深度和空间感。透视原理中的视点、消失点和平行线显然基于数学的几何概念。

图 3. 透视投影和正交投影，来源：<https://github.com/rougier/scientific-visualization-book>

在早期的绘画中，艺术家主要关注表面的平面效果，追求形象的符号化和图像的象征性。这些绘画作品通常是二维的，缺乏深度和透视线。然而，随着人类对空间感知的进一步理解，艺术家开始尝试在画布上呈现更加真实的三维效果。

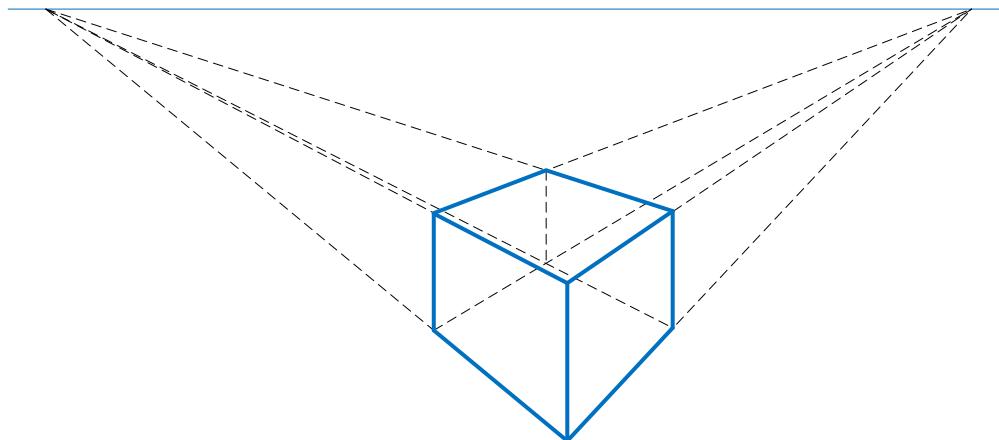


图 4. 三维效果

透视线是绘画中实现维度提升的关键工具之一。在文艺复兴时期，意大利艺术家们如达·芬奇 (Leonardo da Vinci) 和拉斐尔 (Raphael) 等开始系统地研究透视线原理，并将其应用于绘画实践中。此外，雕塑和建筑也是艺术中维度提升的重要领域。

雕塑艺术家通过雕刻实体材料，创造出具有体积、形态和质感的立体作品。立体雕塑的存在使得观众可以从不同角度欣赏作品，营造更强的参与感和沉浸感。建筑艺术则将维度提升应用于空间的创造和设计中。建筑师通过建筑物的结构、布局和比例，创造出具有深度、透视线和立体感的空间。

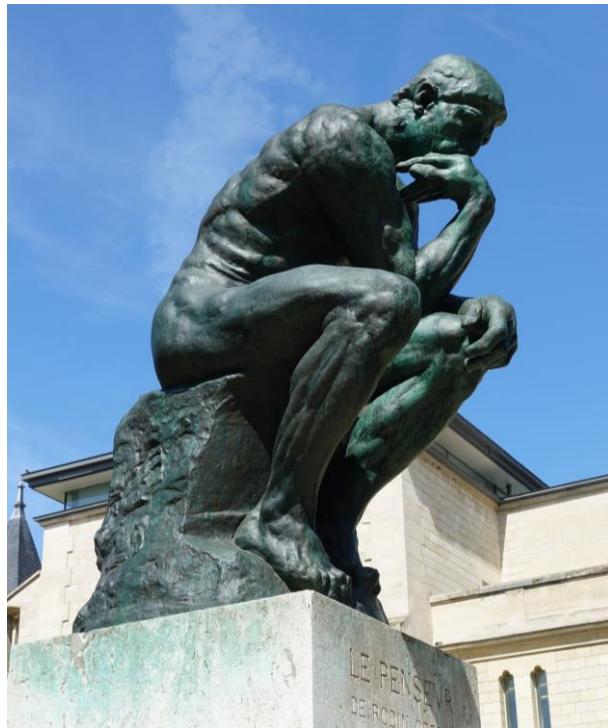


图 5. 罗丹的《沉思者》，现藏于巴黎罗丹美术馆；图片来自 Wikipedia

印象派 (Impressionism) 的画家则试图从时间角度“升维”。“印象”一词即是源自莫奈 (Claude Monet) 的《印象·日出》。创作《印象·日出》时，莫奈 (Claude Monet) 采用多角度绘制同一场景的方法，以捕捉光影的变化和时间的流逝。



图 6. 莫奈的《印象·日出》，现藏于巴黎玛摩丹美术馆；图片来自 Wikipedia

而立体派 (Cubism) 艺术家，比如毕加索 (Pablo Picasso)，则打破传统的透视技法，追求用粗犷的几何形状来解构物体。立体派通过将物体碎裂、解析、重构后在一张画布上呈现让人意想不到的多重视角。



图 7. 毕加索的《亚维农的少女》，现藏于美国纽约现代艺术博物馆 MoMA；图片来自 Google Art Project

艺术不是快照，并不追求事物的外在观感，艺术试图揭示、分享事物更深层、更本质的意义。这一点来看，艺术和数学可谓异曲同工。

1.3 数学 + 艺术 + 人工智能

计算机广泛应用之后，算法艺术应运而生。算法艺术是一种结合数学和艺术的领域，艺术家使用计算机编程和算法来生成艺术作品。这些算法可以基于数学模型、随机性或交互性来创作艺术。

图 9 所示为艺术家 Oliver Brotherhood 创作的开源艺术创意——鸢尾花曲线。这组曲线本质上就是计算机图形学中常用的贝塞尔曲线。

曲线本身尽管和鸢尾花本身并无直接关系，但是在随机数发生器加持的不确定因素下这一组曲线所展现出来的婀娜多姿和鸢尾花在精神层面达到了前所未有的高度统一。这也是本书封面采用这一开源创意的重要原因。

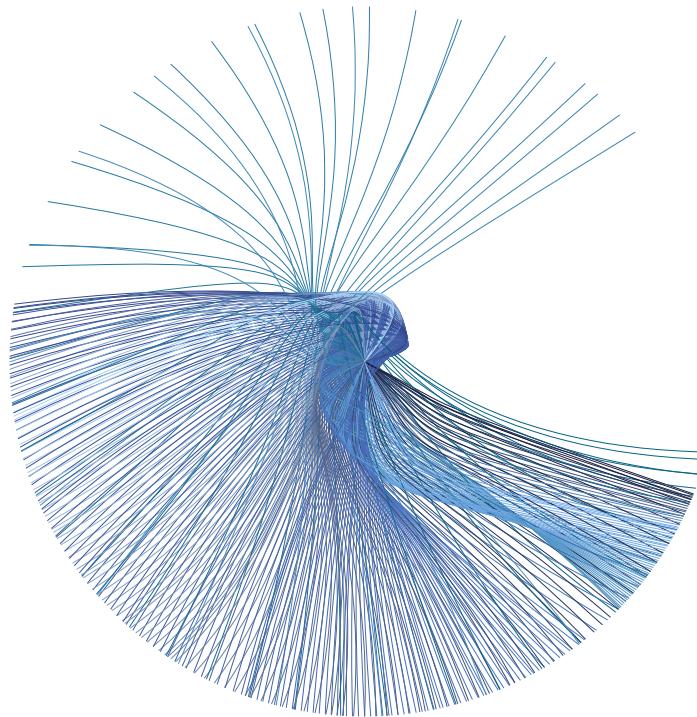


图 8. 鸢尾花曲线，Oliver Brotherhood 开源创意

近些年 AI 的应用发展让艺术王国的版图为之震颤。图 9便是用 Midjourney 训练而成的鸢尾花。本书扉页和半透明硫酸纸上展示的鸢尾花也都是出自 Midjourney 之手。



图 9. Midjourney 训练生成的鸢尾花

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

Midjourney 是一种基于人工智能的创作方式，将数学和艺术巧妙地结合起来。通过数学模型和算法，Midjourney 能够生成独特而引人入胜的艺术作品。

数学提供了创作的理论和框架，如几何学、比例和颜色理论，帮助艺术家创造出具有美感和视觉吸引力的作品。同时，艺术在 Midjourney 中发挥着关键的角色，通过创造力、想象力和表达力，使生成的作品充满了情感和个性。Midjourney 的结合展示了数学与艺术的协同作用，创造出令人惊叹的艺术创作体验。

1.4 解构 + 重构：形而上者谓之道，形而下者谓之器

形而上者谓之道，形而下者谓之器。艺术和数学正是“道”，两者都着力于解构、重构，穿越事物的形，看透事物的神。艺术和数学都有超强能力把事物肢解成基本单位，然后再重新组合构造全新的结构。

北宋王希孟 (1096 - 1119) 创作的《千里江山图》运用了精细的笔触和细致入微的绘画技法，刻画烟波浩渺、崇山峻岭、高崖飞瀑、村舍集市、水榭楼台、渔船客舟、小桥流水、曲径通幽、茂林修竹、柳绿花红 … 特别是，满幅画作巧妙地运用青色、绿色制造立体感和层次感，使得远近山水跃然纸上。

真可谓，只此青绿，千里江山；看山是山，见水是水。看山不是山，见水不是水；看山还是山，见水还是水。



图 10. 王希孟《千里江山图》局部，现藏于北京故宫博物院；图卷纵 51.5 厘米，横 1191.5 厘米；图片来自 www.dpm.org.cn

图 11 所示《鸢尾花》是荷兰后印象派画家梵高 (Vincent van Gogh, 1853 - 1890) 的作品。鲜亮的色调和奔放的纹理完美捕捉了鸢尾花怒放时的妖艳。浓烈的紫色和蓝色跳脱翠绿的草木、棕黄的泥土，营造出一种强烈的视觉冲击力，尽显生命的不朽的张力、微妙的平衡。

梵高的画笔让这丛鸢尾瞬时的绽放成为永恒。



图 11. 梵高《鸢尾花》，现藏于美国加州 J. Paul Getty Museum；图片来自 Google Art Project

蒙德里安的《红、黄、蓝的构成》*Composition with Red, Blue and Yellow* 从几何和色彩角度更是艺术史上独树一帜的作品。《红、黄、蓝的构成》是由荷兰艺术家彼得·蒙德里安 (Piet Mondrian) 于 1930 年创作的一幅抽象绘画作品。这幅画通过几何形状和基本色彩的组合来达到对现实世界的简化和抽象。画面由一系列垂直和水平的黑色线条构成，将画面分割成不同大小和形状的矩形块。这些矩形块填充了红色、蓝色和黄色，形成了一种平衡和谐的色彩组合。

蒙德里安的艺术理念强调对艺术元素的简化和纯粹性的追求。他认为几何形状和基本色彩是最基本的艺术元素，通过将它们组合在一起，可以表达一种超越物质世界的精神和秩序。在《红、黄、蓝的构成》中，蒙德里安通过创造一种平衡的布局和色彩对比，传达了一种对和谐和平静的追求。

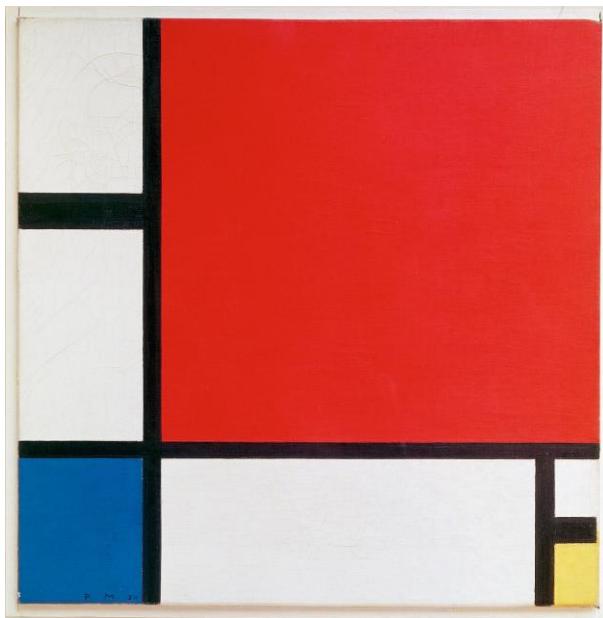


图 12. 蒙德里安《红、黄、蓝的构成》，现藏于苏黎世美术馆；图片来自 Google Art Project

本书作者私以为“解构 + 重构”王者正是汉字。明代画家石涛曾说“一画者，众有之本，万象之根。”而每一个方块字又何尝不是一幅画？

353年，“群贤毕至，少长咸集”，王羲之创作《兰亭集序》。透过这些优雅的线条，跨越千年至今我们依旧感受到历史的线、几何的形、艺术的神、文明的魂。

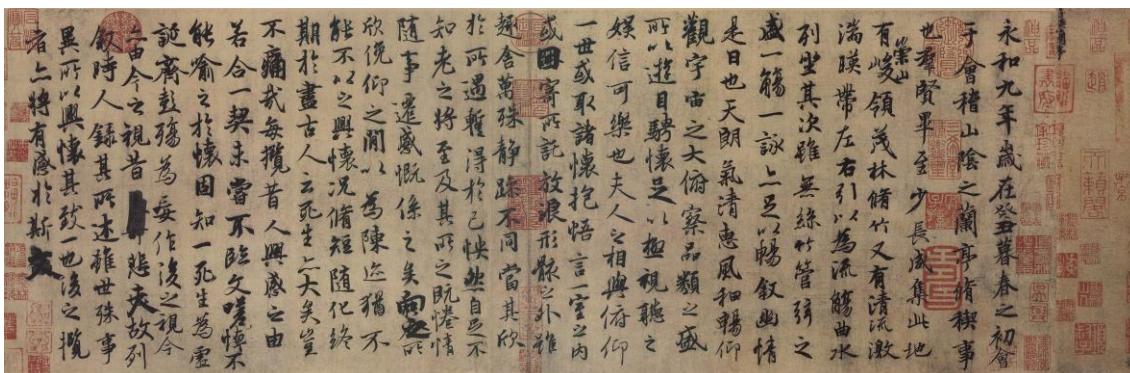


图 13. 王羲之《兰亭集序》，现藏于北京故宫博物院；图片来自 Wikipedia

1.5 师法自然

人法地，地法天，天法道，道法自然。艺术当然离不开人类天马行空的想象力，而给这些想象力持续赋能的沃土正是人类赖以生存的自然界。自然界中，数字、数学可能就是万物之“道”。

道生一，一生二，二生三，三生万物。类似这种二叉树形的增长方式几乎无处不在。

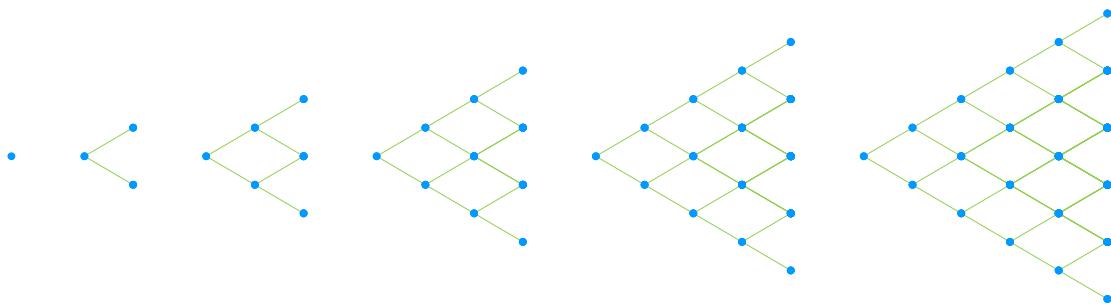


图 14. 二叉树

黄金分割在自然界中广泛存在，如植物的分枝、螺旋壳的结构、人体的比例等，展现了数学和美的奇妙关联。斐波那契数列在自然界中也有许多出现，如植物的叶子排列、花瓣的排列、蜂窝的形状等。这种数列展示了自然界中的规律和对称性。

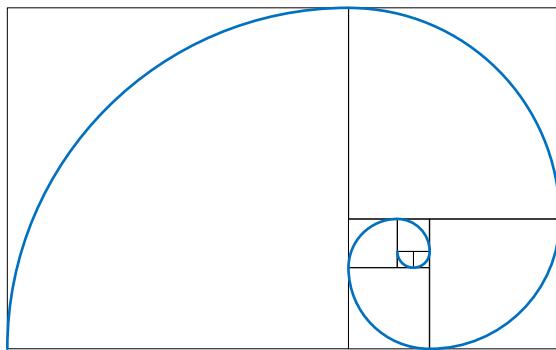


图 15. 斐波那契数列和黄金分割

植物的分枝结构常常呈现出分形的特征。例如，树枝、树叶的排列方式、花朵的形状等都可以看作是分形结构。这种自相似性使得植物在各个尺度上都具有相似的形态，从整体到细节都呈现出美妙的几何模式。有趣的是，观察菌丝我们发现的也是类似树枝分形的结构。从地球的尺度来看，挺拔如云的树木也不过是地表上毫不起眼的菌丝。

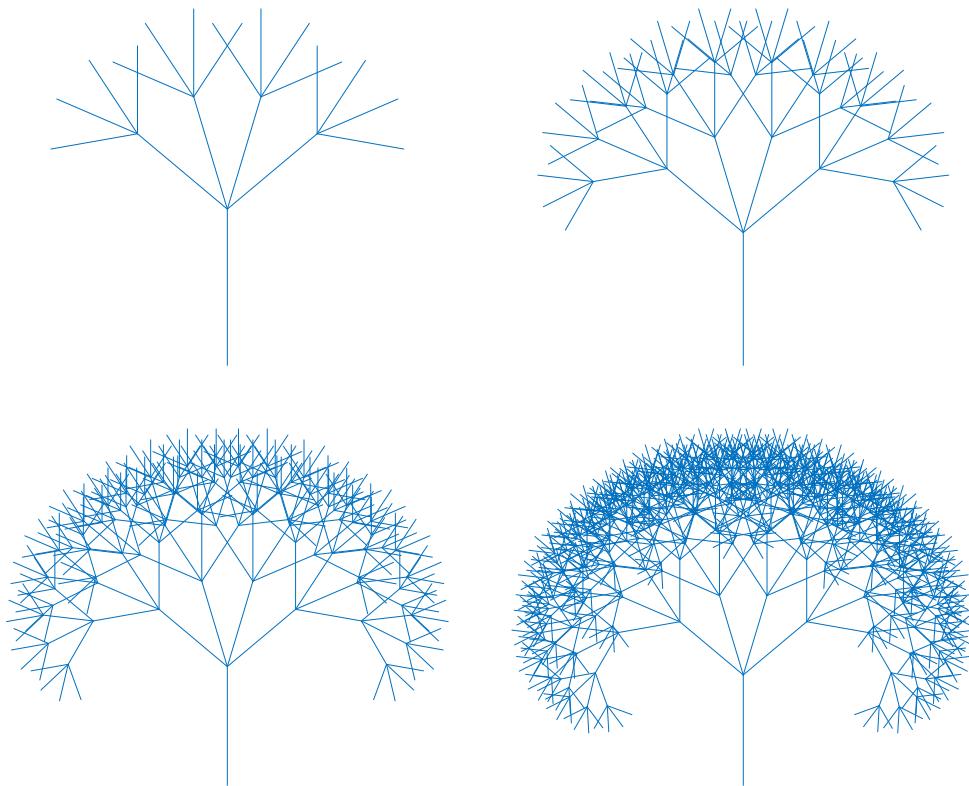


图 16. 分形，树枝

海岸线和山脉的形状也呈现出分形特征。无论是放大还是缩小，它们的形态都是自相似的，具有相似的曲线和起伏。云朵的形状和闪电的分支都可以被视为分形结构。它们在各个尺度上都具有相似的形态和分支模式，展现出自然界中的分形美。

雪花和冰晶是自然界中常见的分形形态。它们的晶体结构在多个尺度上具有相似的形状，形成了复杂而美丽的分形图案。

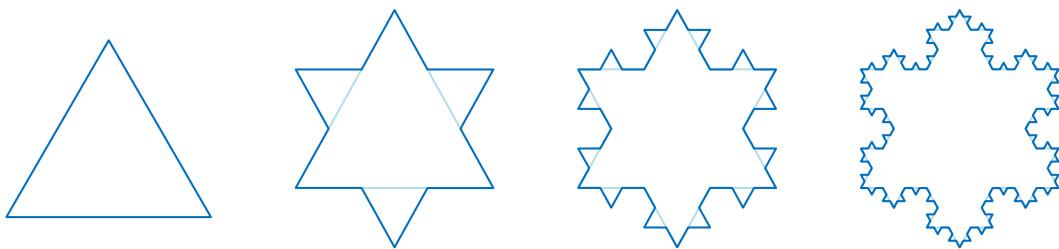


图 17. 分形，科赫雪花 (Koch snowflake)

1.6 模式 + 随机：带着枷锁的翩翩起舞

模式 + 随机，无处不在。模式让物质世界充满秩序，随机让整个寰宇满是精彩。小到一点浮尘、一片雪花、一朵浪花，大到四季变化、动物迁徙、人类社会、满眼繁星。

自然界的草木却没有展现出图 16 这种高度完美的对称。世界上没有两片完全一样的叶子，也没有两片完全一样的雪花。这一点正是“模式 + 随机”的数学体现。

模式代表一种确定，需要站在宏观、大量、长期尺度上观察；随机代表一种不确定，是微观、少量、短期尺度视角。

英国著名植物学家罗伯特·布朗通过显微镜观察悬浮于水中的花粉，发现花粉颗粒迸裂出的微粒呈现出不规则的运动。他说，不断重复地观察这些运动给我极大的满足；它们并非来自水流，也不是源于水的蒸发，这些运动的源头是颗粒自发的行为。

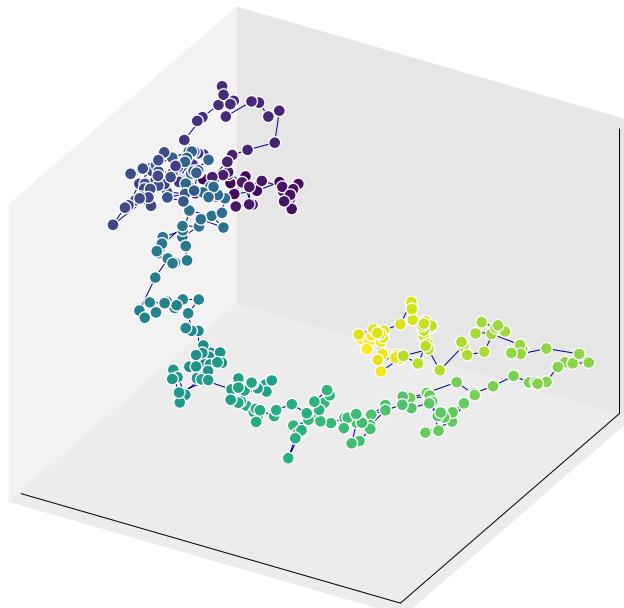


图 18. 模拟某个浮尘在三维空间的随机漫步轨迹

威尔逊·奥尔温·本特利 (Wilson Alwyn Bentley) 是一位美国的自然摄影师和雪花研究家。他被誉为“雪花之父”，因为他是第一个成功地将雪花的照片拍摄下来的人。本特利对雪花的形态和结构产生了浓厚的兴趣，并通过使用特制的显微摄影技术，捕捉到了超过 5000 张雪花的照片。在这 5000 张雪花照片中，他没有发现两片一样的雪花。

但是不管怎么样，大家可以在这些雪花中发现 60 度角、六边形这样的几何模式。这显然不是巧合。究其本质，水分子是由一个氧原子和两个氢原子组成的，呈 V 字型结构。冰的晶体结构称为六方最密堆积结构。在这种结构中，每个水分子与周围六个水分子相邻，并形成六边形的环状结构。这种紧密的排列方式使得冰晶体具有六边形的外观。

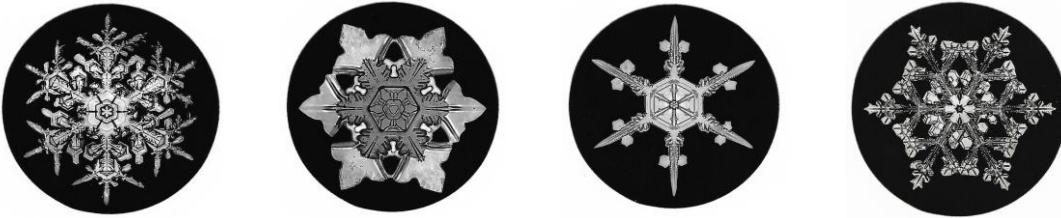


图 19. 威尔逊·奥尔温·本特利拍摄的雪花照片；图片来源：<https://snowflakebentley.com>

微风吹散蒲公英的种子，这些小小降落伞看似做着无规则的随机漫步，但是无时不刻不在气流的支配下运动。宏观尺度上来看，丝丝缕缕的气流、形状各异云朵，极具破坏力的飓风，是在地球的公转和自转影响下运动。



图 20. 飓风，图片来自于 www.nasa.gov

人类之所以能够看到形状和色彩都离不开光。而光具有波粒二象性，表现出既有波动特性又有粒子特性，这又是“模式 + 随机”的一个例子。在波动方面，光可以通过干涉和衍射等现象展示出波的传播特点，并遵循确定性的规律。而在粒子性方面，光表现出随机性，例如光子的发射和探测位置具有一定的随机性。这种波粒二象性的存在使得光在不同实验条件下表现出独特的行为，既有波动的可见光谱特性，也有粒子的能量量子化特点。

某个时间观察特定的一只动物，我们很难发现任何特定规律、模式。长期观察一群动物，我们可以发现四季轮转、草木荣枯支配者动物年复一年地在某个大陆板块的繁衍生息、迁徙移动。

生如夏花之绚烂，逝如秋叶之静美。每个人的一生不也是“模式 + 随机”相互作用的产物。生老病死、起点终点、前浪后浪，这是没人能逃脱的“俗套”的模式。而每个人都能活出自己，走出与众不同的人生轨迹，这便是“精彩”的随机。

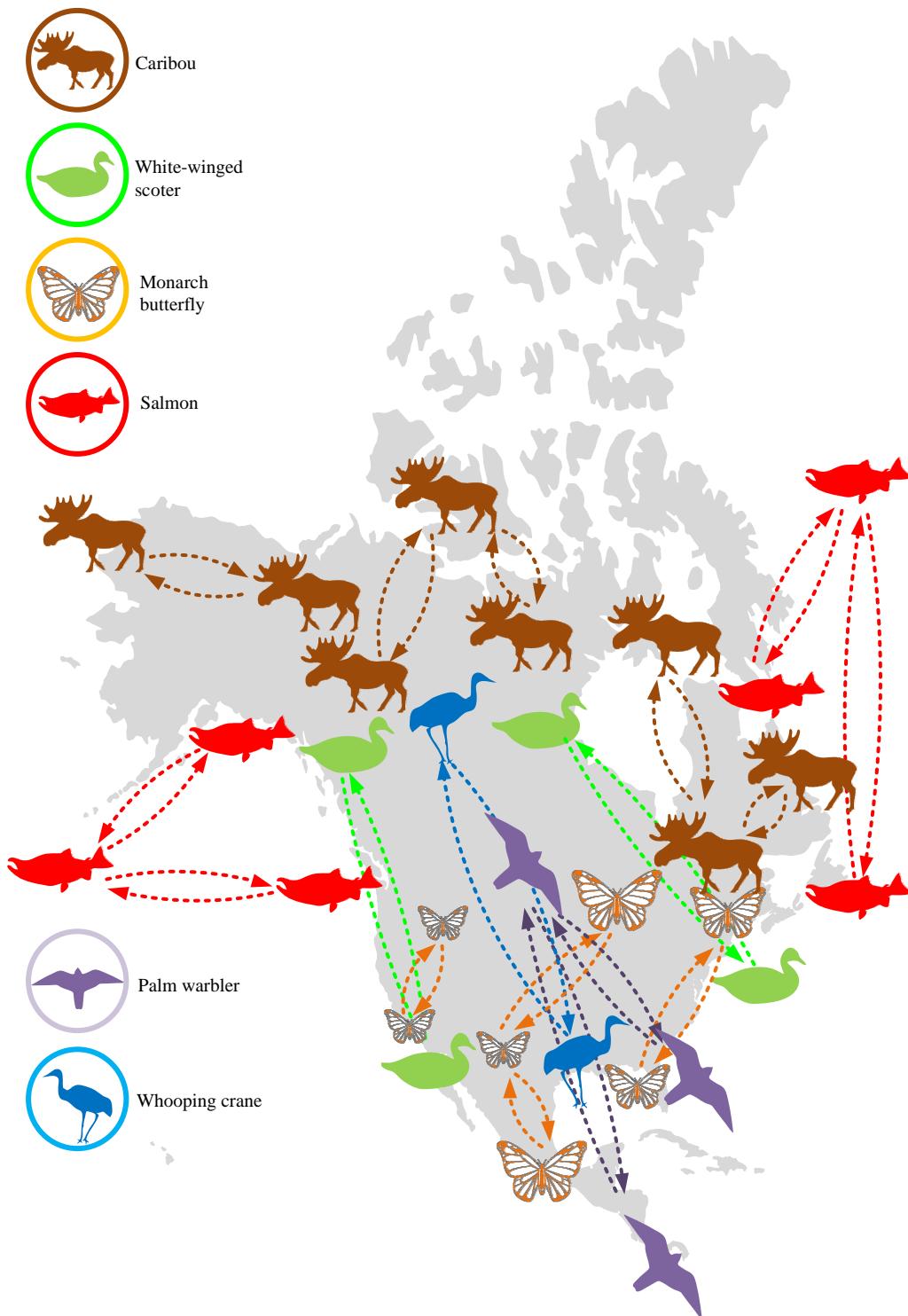


图 21. 北美大陆主要动物迁徙路径

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

1.7 宇宙之道

一沙一世界，一花一天堂。无限掌中置，刹那成永恒。满天的繁星又何尝不在展现这种“模式 + 随机”。

仰观宇宙之大，俯察品类之盛。在科学技术的助力下，我们在微观尺度能够描绘电子轨迹，宏观尺度上能够观察天体运行。图 22 描绘从中心释放的大量电子轨迹。

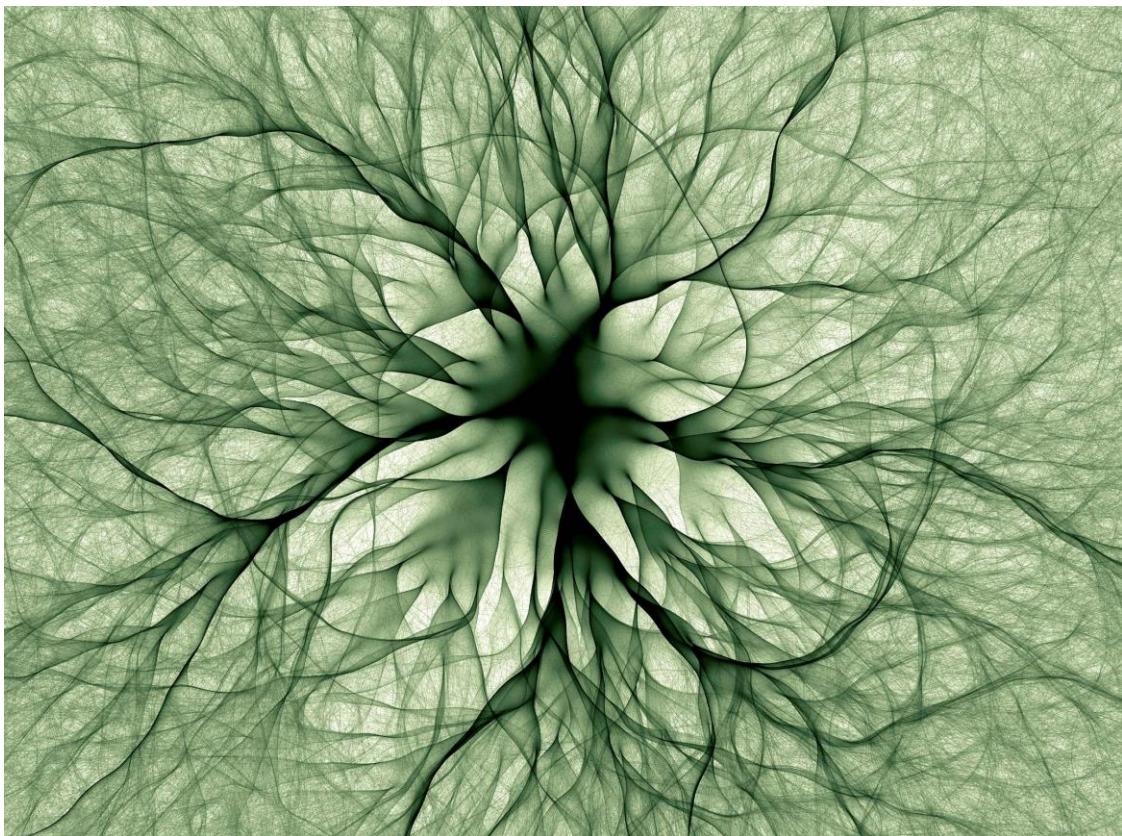


图 22. *Transport II* by Eric J. Heller at Harvard University。图片来自于 www.nsf.gov

微观和宏观达到的模式上的统一，这就是像是一条巨蟒咬住了自己的尾巴，即衔尾蛇 (Ouroboros)。有些人认为，数学中表示无穷的符号 ∞ 也是来自于衔尾蛇 (扭纹形)。

据德国化学家凯库勒 (August Kekulé) 本人的著作称，他梦中梦到一条蛇咬住了自己尾巴，从而得到启发得到苯环结构。

而鸢尾花书选取的编程语言 Python 本意也是蟒蛇，特别地，Anaconda 的标识就是衔尾蛇。这种冥冥之中的巧合耐人寻味。

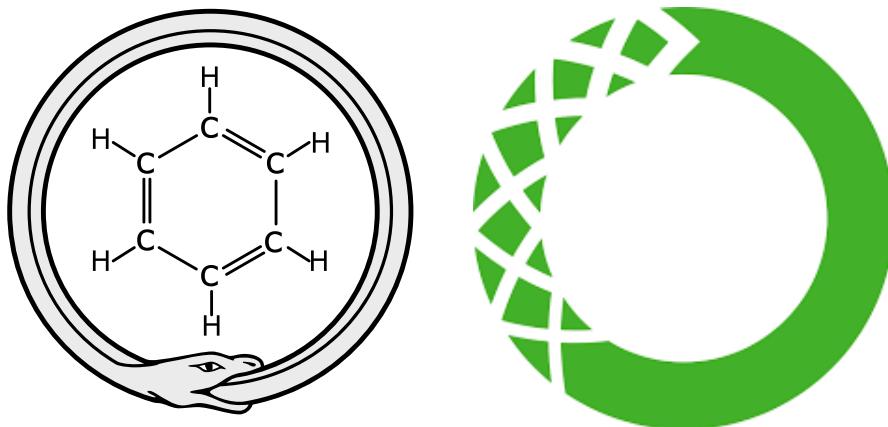


图 23. 衔尾蛇、苯环、Anaconda logo, 图片来自于 Wikipedia

我们身体的某个角落的某个电子在绕着某个原子在做近似椭圆的运动。而我们又乘着地球绕着太阳在椭圆的轨道上运行。而以太阳为中心的太阳系又绕着银河系的中心公转，而银河系…

这繁复的相互缠绕关系，让我们又想到了微观层面的 DNA 结构。宏观、微观，这条蛇再次地咬住了自己的尾巴。

那么，宇宙的图景到底怎样？

在下面这幅银河系的图景中，我们是否既看到了模式的确定，也看到了随机带来的不确定。



图 24. 银河系图景，图片来自于 www.nasa.gov

正如美国物理学家费曼所说，我，一个无数原子组成的宇宙，又是整个宇宙的一粒原子。地球，那个暗淡蓝点，不过是银河系中的一粒沙子，我们又何尝不是蝼蚁。

换个角度来向，数字信息存储在一个个芯片中。这一个个芯片又何尝不是一个个数字“宇宙”！再疯狂一点，我们这个所谓的物质世界是否也是某个矩形芯片的模拟产物？光速仅仅是避免死机的保护机制？

难怪有人说，梵高笔下的星夜更接近真实的星空。

带着这样的几何视角、动态思考，今晚大家不妨去夜观天象，然后再看图 25 梵高的《星夜》中描绘的斗转星移，我们是否会觉得梵高的作画不再“疯狂”？

浮名浮利，虚苦劳神。叹隙中驹，石中火，梦中身。那个时代的梵高是否是举世皆浊我独清，众人皆醉我独醒？



图 25. 梵高的《星夜》，现藏于美国纽约现代艺术博物馆 MoMA；图片来自 Google Art Project



若将贫贱比车马，他得驱驰我得闲。别人笑我太疯癫，我笑他人看不出穿。不为五斗米折腰的梵高是否又在意凡胎俗人看他的眼光？

几时归去，作个闲人。对一张琴，一壶酒，一溪云。数学家和艺术家的“超然物外”本质上是否高度一致？

华罗庚先生说，数缺形时少直观，形少数时难入微；数形结合百般好，隔离分家万事休。不畏浮云遮望眼，忘却浊骨肉眼中的“真实”世界，我们能否通过“数学 + 艺术”凿开的缺口一窥更“本质”的存在？



Let's Talk about Figures

说图

正式开始一场“数学 + 艺术”的美学实践



独处，是创造的秘诀；独处，是好主意诞生的时候。

Be alone, that is the secret of invention; be alone, that is when ideas are born.

——尼古拉·特斯拉 (Nikola Tesla) | 发明家、物理学家 | 1856 ~ 1943



- ◀ XXXXX

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

2.1 一图胜千言

上一章，我们聊了一些有关“数学 + 艺术”形而上务虚的内容，本章开始介绍如何用 Python 完成各种可视化的实操内容。

一图胜千言 (a picture is worth a thousand words)。一说到图片可视化的作用，大家自然而然地会想到这句“陈词滥调”。但是，并不是所有的图片都“胜千言”。

虽然颜值即正义，但是在数学工具、数据科学、机器学习等应用场景，优质可视化方案不仅要读者觉得“眼前一亮”且“言之有物”。

优质可视化方案有助于高效传播信息，在短时间内让读者接收到信息，并促进交流、思考。有效的图片信息传播应注重高颜值、清晰度、专业性、简洁性、准确性和与读者之间的互动。简而言之，运用之妙，存乎一心，让读者主动思考的图片才是优质的可视化方案。

反之，低效可视化方案问题可能涉及：信息密度低、信息过载（比如满纸公式）、图像质量差（非矢量、分辨率低）、设计混乱（中心不明确、分散注意力）、缺乏明确的标签和解释、配色方案失效、缺乏上下文和交互性，以及信息的不准确性（比如手绘高斯函数曲线或曲面）。

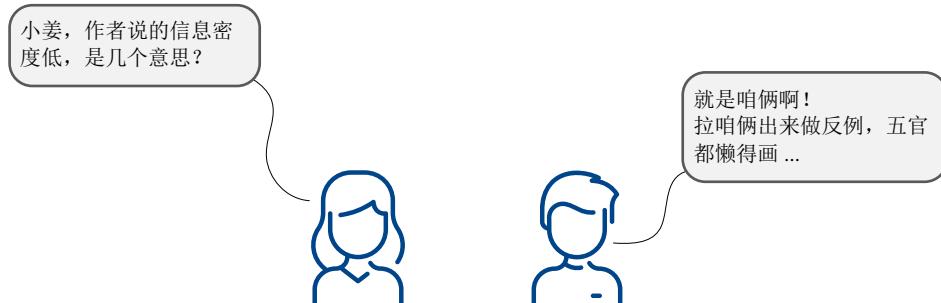


图 1. 低效信息传播

以图 2 为例，为了版面设计，在文本中插入这张鸢尾花照片没有任何意义。如果利用这幅鸢尾花照片讲解一幅彩色照片可以由红、绿、蓝三色组成，这张照片就不仅仅是凑数的“花瓶”。为了保持信息传播的连续、高效，我们还可以用这幅鸢尾花照片讲解矩阵（图 3）、主成分分析（图 4）。



图 2. 鸢尾花照片

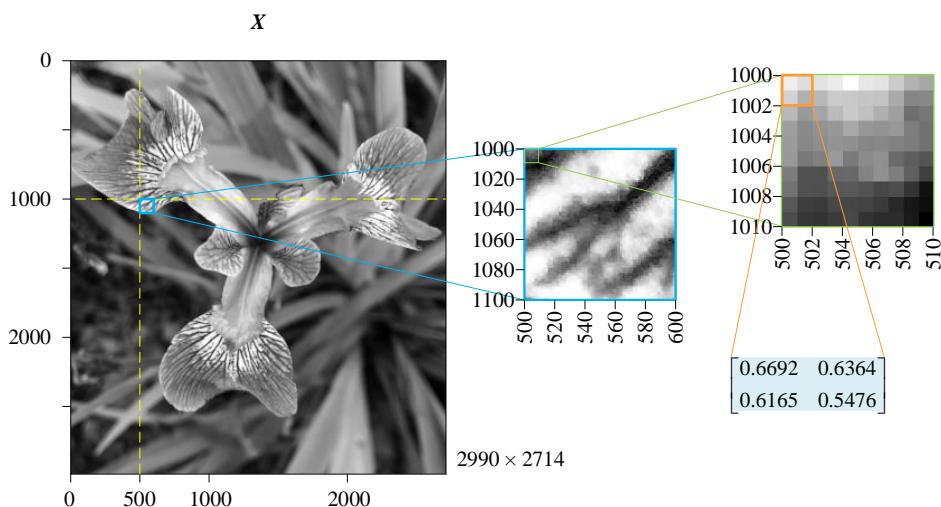


图 3. 照片也是数据矩阵

一张图片的整个生命周期一般要经过如下几个阶段：

- ▶ 了解规则
- ▶ 头脑风暴
- ▶ 编程实现
- ▶ 美化完善（代码层面）
- ▶ 后期制作
- ▶ 发布传播

下面，我们便按这个顺序介绍如何制作一张图片。

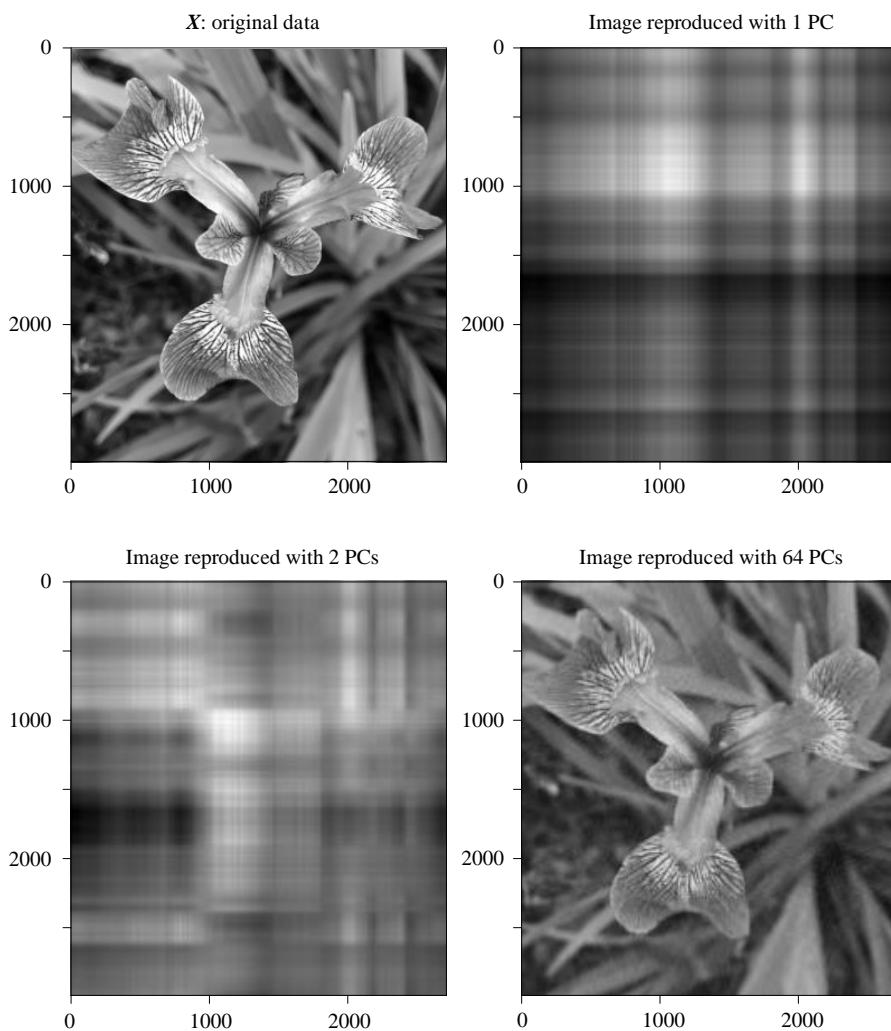


图 4. 对黑白鸢尾花照片的主成分分析

2.2 了解规则：带着枷锁跳舞

在开始可视化之前，明确你的目标是什么以及你的受众是谁。不同的目标和受众可能需要不同类型和风格的可视化呈现。

在这个纸媒和数字媒体共存、共荣的多媒体时代，制作一张图片通常要同时照顾到纸媒、数字媒体的需求。

本书介绍的可视化是在科技制图的范畴之内。因此创作一张图片之前，首先关注制图规则，以便决定图片各种属性。

建议大家在创作图片时考虑以下几个问题：

- ▶ 风格是学术专业？还是轻松活泼？

- ▶ 是否允许手绘？
- ▶ 图片大小尺寸？比例如何？
- ▶ 一幅图是否可以多子图？子图布局有何要求？
- ▶ 图片内文字字体 (Times New Roman, Arial, Roboto, ...)？
- ▶ 文字字号最大、最小几号？文字颜色是否有要求？
- ▶ 图片中的文字是否要求可编辑？
- ▶ 图片中是否可以嵌入公式？
- ▶ 黑白、彩色？配色有何特殊要求？
- ▶ 如何保证彩色图片在黑白灰的呈现效果？
- ▶ 是否需要针对色盲群体调整配色？
- ▶ 颜色采用 RGB，还是 CMYK？
- ▶ 图中线宽、线型是否有要求？
- ▶ 是否有必要删除隐藏图层的元素？
- ▶ 图片是静态，还是交互？
- ▶ 图片的格式？矢量图，还是像素图？
- ▶ 图片如果过大，是否可以光栅化 (rasterize)？
- ▶ 像素图的像素要求如何？最小、最大像素？
- ▶ 图片是否需要单独保存，并提交？
- ▶ 图片文件格式 (JPEG、PNG、GIF、SVG、TIFF、PDF ...)、大小是否有要求？
- ▶ 图片是否要用于演示，比如放在 PPT 中？PPT 中的文字大小？
- ▶ 是否需要制作动画？
- ▶ 是否考虑创作 App 应用、dashboard？

正式出版物 (纸媒、数字媒体、会议) 一般都有专门的制图指南，建议大家在开始创作任何图片之前首先仔细阅读制图指南的细则。

如果找不到相关的制图指南，建议大家参考《自然》杂志的制图指南，链接如下：

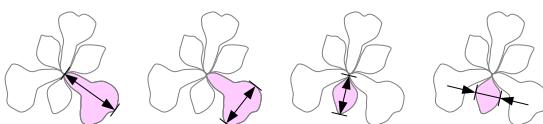
https://www.nature.com/documents/Final_guide_to_authors.pdf

鸢尾花书在创作时，很多制图细节都参考了《自然》的制图指南。

2.3 头脑风暴：知识网络

富有创意的可视化方案可以为数据插上“翅膀”！根据你要传达的信息和数据的性质，选择适合的图表类型，从而提高可视化的效果和可读性。

图 5 所示为鸢尾花数据集，图 16 所示为以鸢尾花数据为起点的知识网。图 6 ~ 图 16 为从鸢尾花书精选出来和鸢尾花数据集有关的可视化方案。



Index	Sepal length X_1	Sepal width X_2	Petal length X_3	Petal width X_4	Species C
1	5.1	3.5	1.4	0.2	Setosa C_1
2	4.9	3	1.4	0.2	
3	4.7	3.2	1.3	0.2	
...	
49	5.3	3.7	1.5	0.2	
50	5	3.3	1.4	0.2	
51	7	3.2	4.7	1.4	
52	6.4	3.2	4.5	1.5	
53	6.9	3.1	4.9	1.5	
...	
99	5.1	2.5	3	1.1	Versicolor C_2
100	5.7	2.8	4.1	1.3	
101	6.3	3.3	6	2.5	
102	5.8	2.7	5.1	1.9	
103	7.1	3	5.9	2.1	
...	
149	6.2	3.4	5.4	2.3	
150	5.9	3	5.1	1.8	



图 5. 鸢尾花数据表格，单位为厘米 (cm)

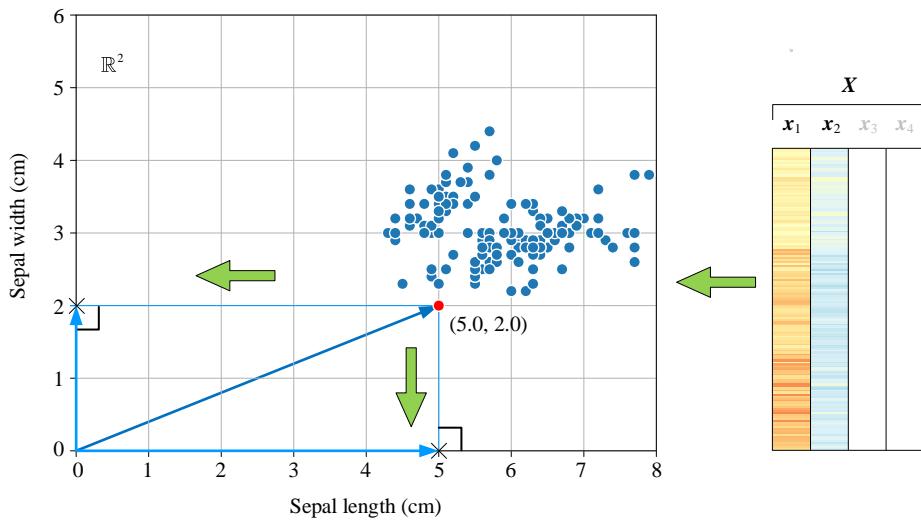


图 6. 鸢尾花前两个特征数据散点图

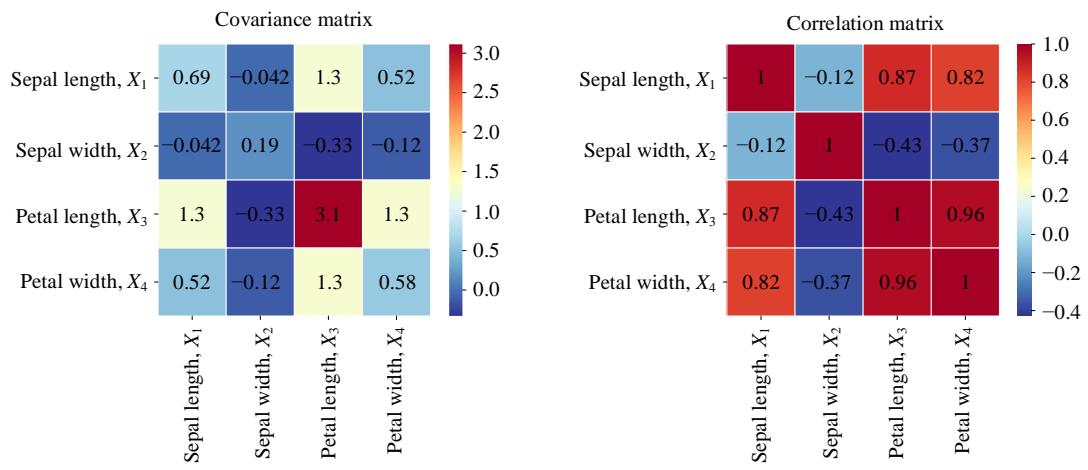


图 14. 协方差矩阵和相关系数矩阵热图

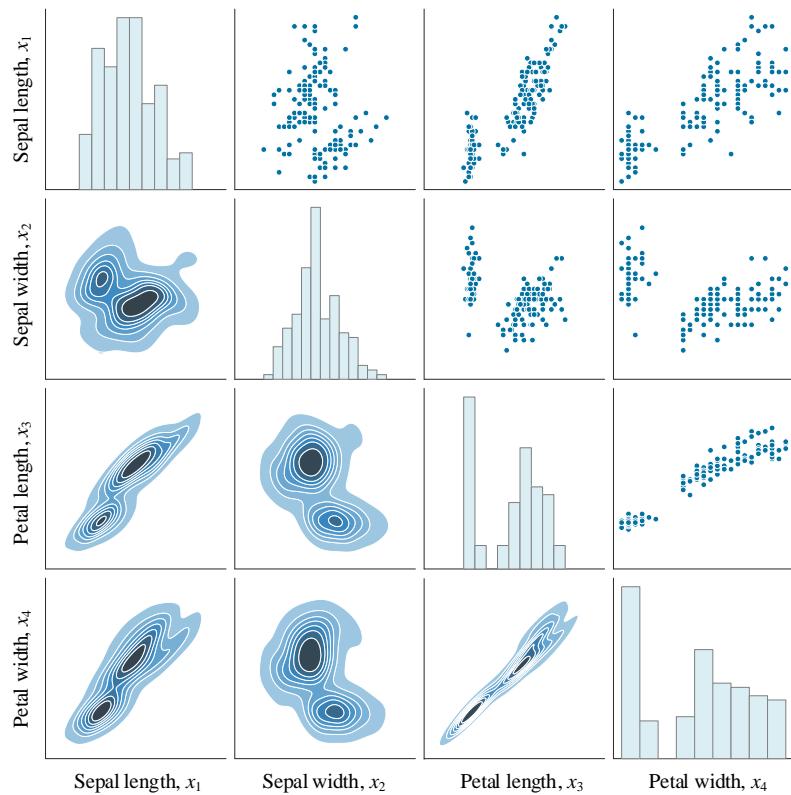


图 7. 鸢尾花数据成对特征分析图，不分类

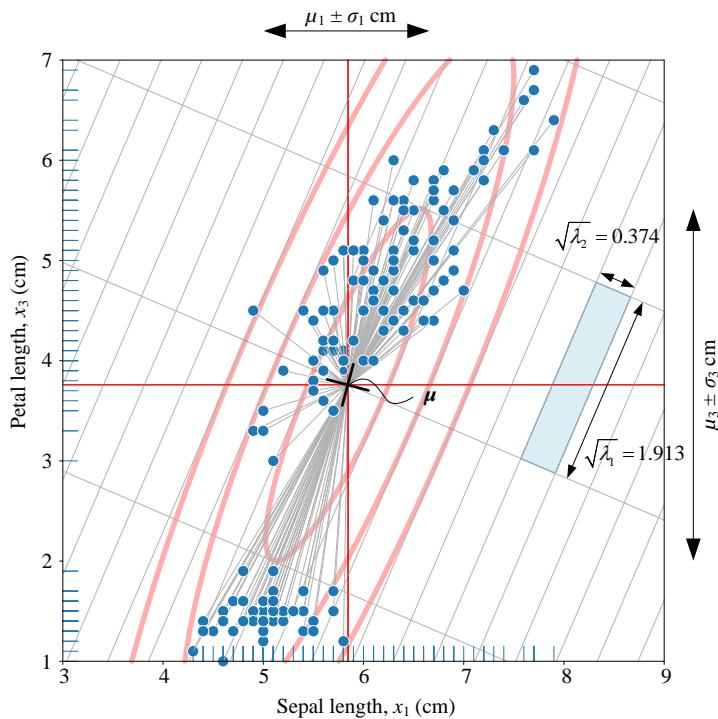


图 8. 花萼长度、花瓣长度平面上的马氏距离等高线和网格

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

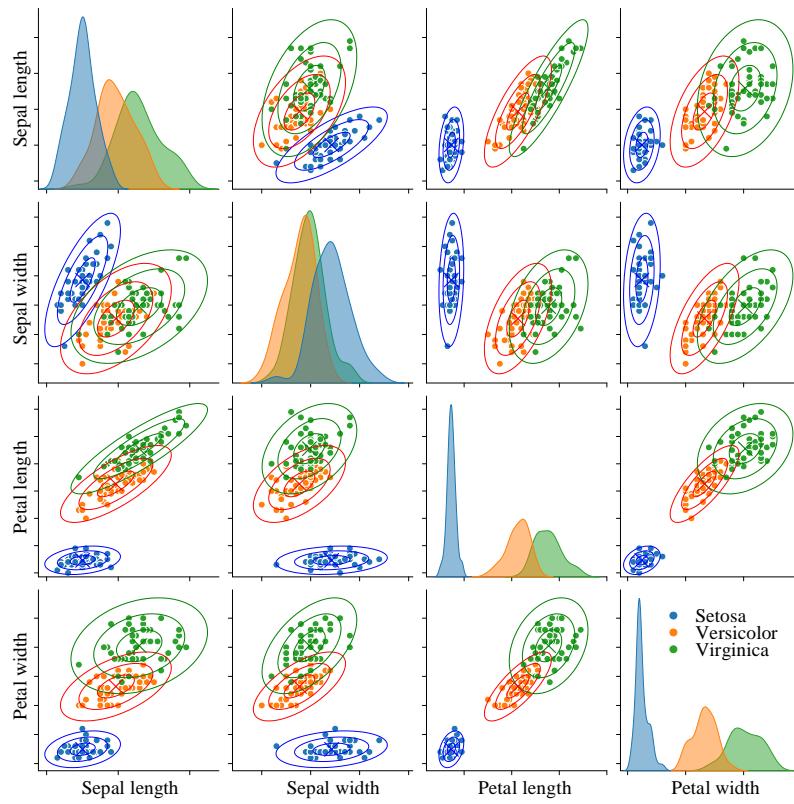
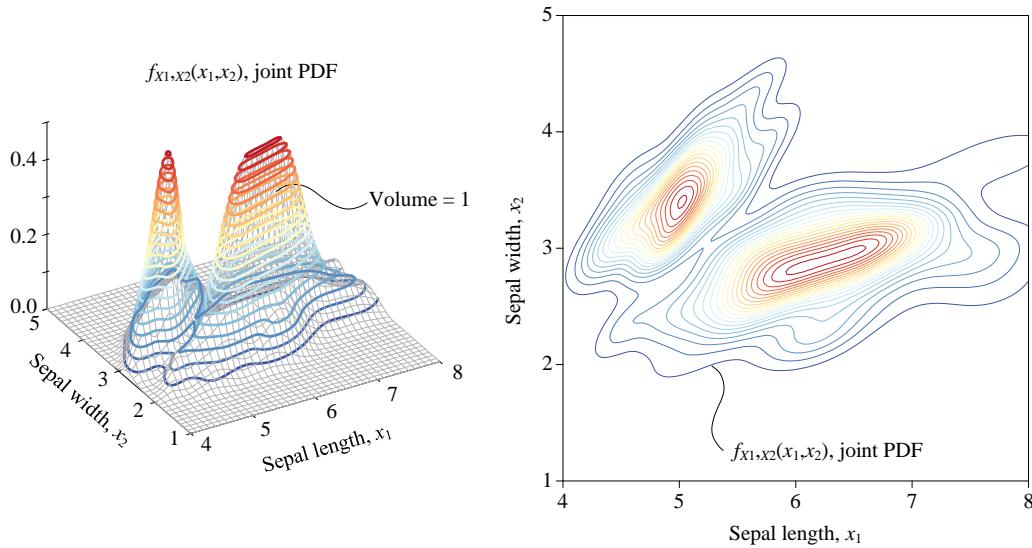


图 18. 协方差矩阵和椭圆的关系，考虑分类

图 9. 联合概率密度函数 $f_{X1,X2}(x_1, x_2)$ 三维等高线和平面等高线，不考虑分类

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

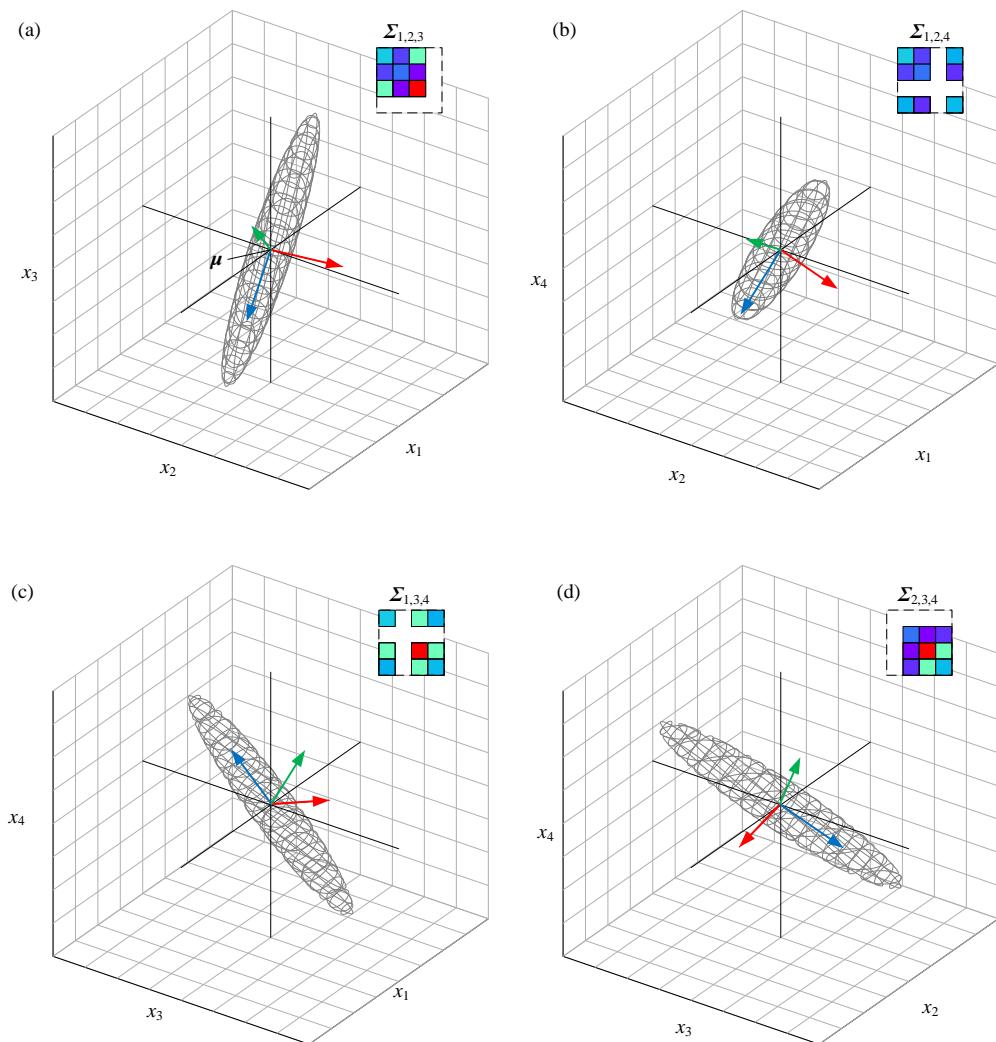


图 10 四维空间的“旋转”超椭球在三维空间中的四个投影

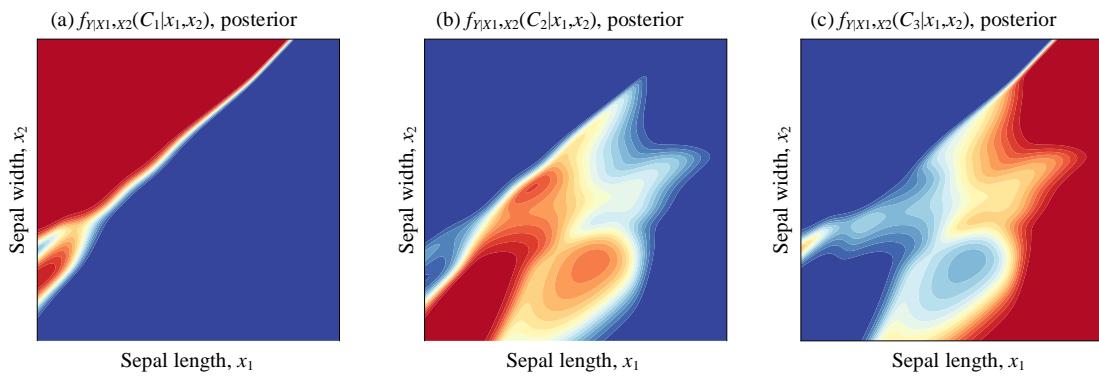


图 11. 比较三个后验概率曲面平面填充等高线

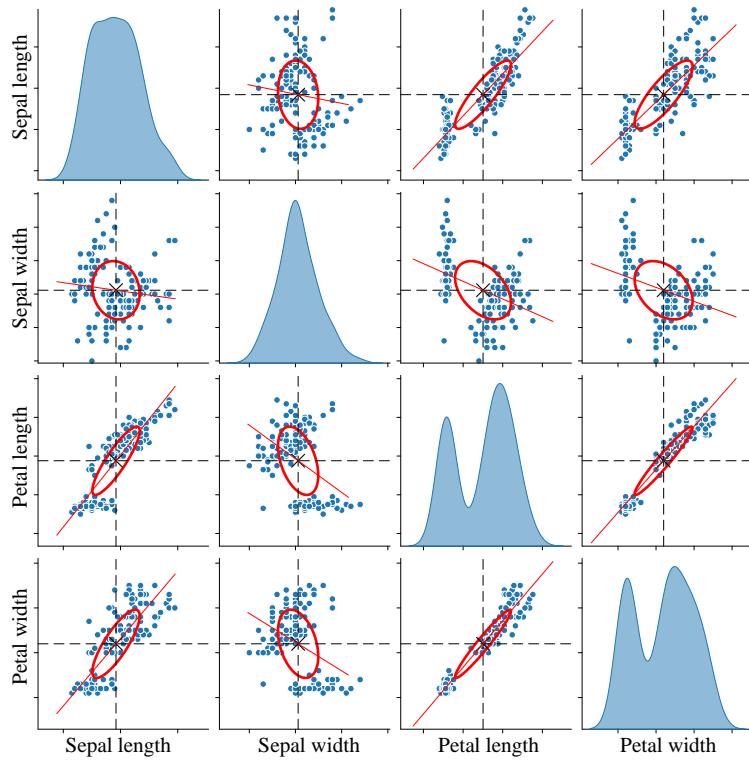


图 12. 鸢尾花数据成对特征图和回归关系

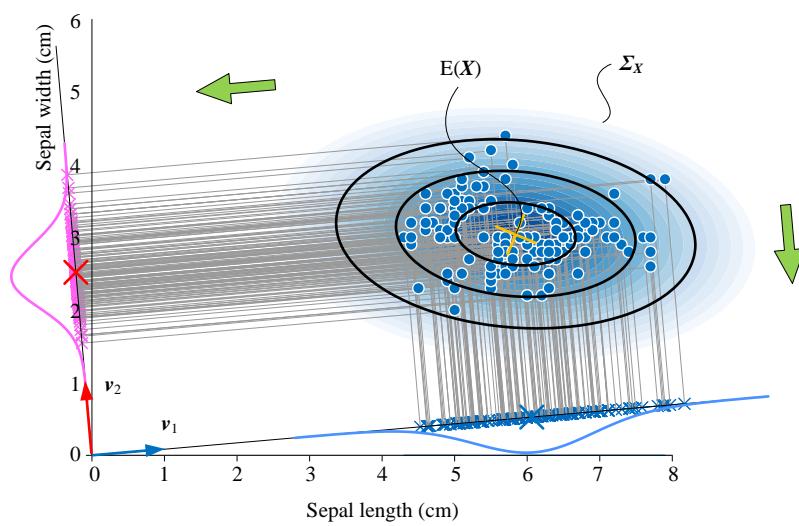


图 13. 鸢尾花数据正交系 V 投影

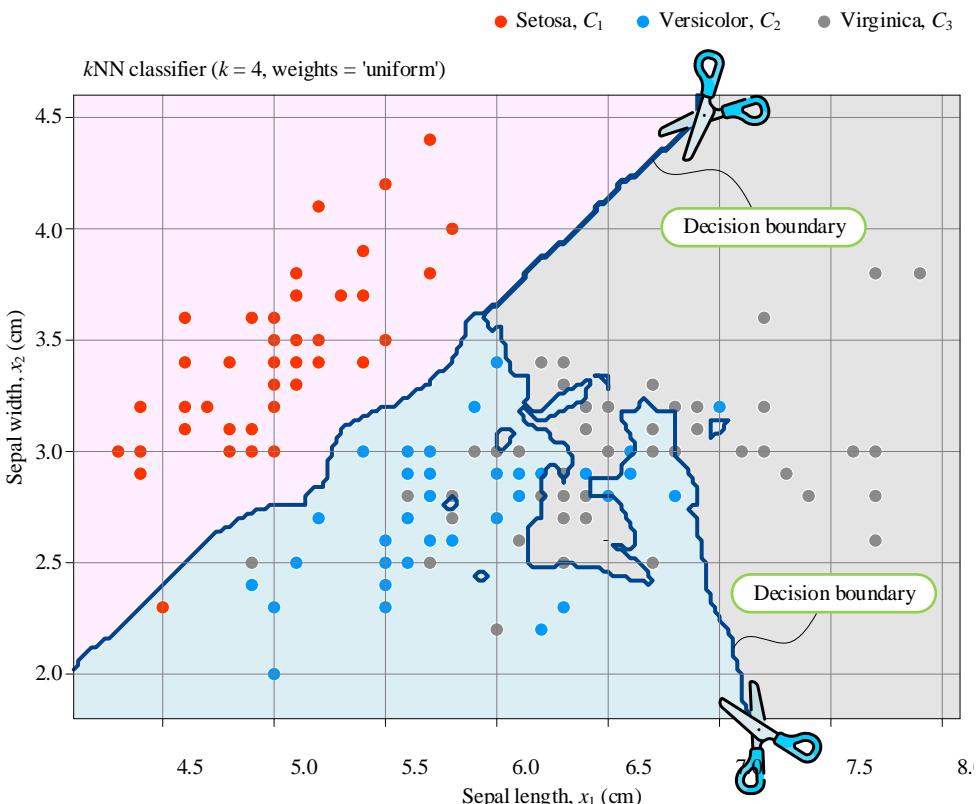


图 14. k 近邻分类, $k=4$, 采用 2 个特征(花萼长度 x_1 , 和花萼宽度 x_2) 分类三种鸢尾花

2.4 编程实现：Python 有大作为

图 6 ~ 图 16 都离不开 Python 编程实现。

Python 拥有多种用于数据可视化的工具。以下是鸢尾花书中一些常用的可视化工具：

Matplotlib: Matplotlib 是 Python 中最流行的绘图库之一，提供了广泛的绘图功能，包括折线图、散点图、柱状图、饼图等。它具有灵活性和广泛的定制选项，可以用于创建静态、交互式和动态的图形。

Seaborn: Seaborn 是建立在 Matplotlib 之上的高级统计数据可视化库。它提供了一组美观且具有统计意义的图表样式和绘图功能，使得数据的可视化变得更加简单和直观。

Plotly: Plotly 是一个交互式的可视化库，支持多种图表类型，包括线图、散点图、柱状图、热力图等。它提供了丰富的交互功能，可以在网页中创建动态和可交互的图形。

图 17 ~ 图 19 提供了一个速查表，用来帮助大家找到合适的可视化方案以及对应的 Python 函数。

除了以上提到的工具，还有其他一些流行的 Python 可视化库，如 Bokeh、Altair、ProPlot、Plotnine 等等，它们都提供了不同的特点和功能，可以根据具体需求选择适合的工具来进行数据可视化。

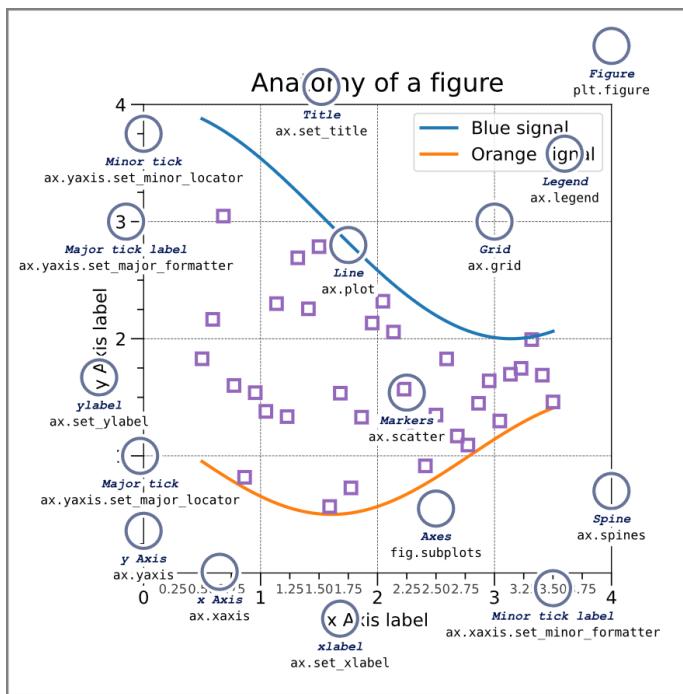
2.5 美化完善：优化默认效果

利用 Python 完成可视化时，利用各种设置美化完善图像是重要的一环。

读过《编程不难》这本书的读者对图 15 都不陌生。如图 15 所示，一幅图的基本构成部分包括以下几个部分：

- ▶ **图像区域** (Figure)：整个绘图区域的边界框，可以包含一个或多个子图。
- ▶ **子图区域** (Axes)：实际绘图区域，包含坐标轴、绘制的图像和文本标签等。
- ▶ **坐标轴** (Axis)：显示子图数据范围并提供刻度标记和标签的对象。
- ▶ **脊柱** (Spine)：连接坐标轴和图像区域的线条，通常包括上下左右四条。
- ▶ **标题** (Title)：描述整个图像内容的文本标签，通常位于图像的中心位置或上方，用于简要概括图像的主题或内容。
- ▶ **刻度** (Tick)：刻度标记，表示坐标轴上的数据值。
- ▶ **标签** (Label)：用于描述坐标轴或图像的文本标签。
- ▶ **图例** (Legend)：标识不同数据系列的图例，通常用于区分不同数据系列或数据类型。
- ▶ **艺术家** (Artist)：在 Matplotlib 中，所有绘图元素都被视为艺术家对象，包括图像区域、子图区域、坐标轴、刻度、标签、图例等等。

美化完善时，以上组成部分都可以调整以便获得更好的可视化效果。本书下一版块专门介绍如何美化完善图像。

图 15. 解剖一幅图，来源 <https://matplotlib.org/stable/gallery/showcase/anatomy.html>

2.6 后期制作：丰富图片细节

后期制作是可视化重要环节之一。以鸢尾花书为例，用 Python 导出的矢量图不会被直接用到图书当中。每一幅都至少经过两个软件后期处理之后才会使用。常用的后期制作软件包括：Adobe Illustrator、Adobe Photoshop、Inkscape（免费）、Microsoft Visio 等。

为什么需要后期制作？下面给出几个理由：

- ▶ 美学设计：尽管 Python 库可以生成基本的图表，但在美学设计方面可能有一些限制。使用其他软件，如 Adobe Illustrator、Photoshop 或 Inkscape 等，可以提供更多的自定义选项，使图表更专业。
- ▶ 复杂效果：某些特殊效果可能在 Python 库中难以实现。其他软件通常提供更高级的图形处理功能。这些效果可以为图表增加视觉吸引力。
- ▶ 排版和注释：在生成图表后，可能需要添加额外的注释、标题、图例或其他文本元素；而编程增加这些元素可能耗时耗力，而且效果差、可编辑性差。其他软件通常提供更灵活的排版选项，可以更好地控制这些元素的位置和外观。
- ▶ 合并图表：如果需要将多个图表组合在一起，或者将图表与其他图像、照片或文本元素进行合并，其他软件通常提供更强大的合并和布局功能。

总之，尽管 Python 库提供了很多绘图功能，但有时候使用其他软件进行后期制作可以提供更大的自由度和更复杂的效果，以满足特定的需求。

注意，美化修饰没有问题，篡改数据必须坚决杜绝。

2.7 发布传播：到什么山上唱什么歌

经过了解规则、头脑风暴、编程实现、编程美化、后期制作等环节，一幅图算是诞生了，现在就差一步——发布传播。发布传播的方式主要分成两大类：

- ▶ 纸媒，比如海报、手册、报刊、杂志、学术论文、书籍等等；
- ▶ 电子媒体、演示 (PPT)、动图、动画、视频、App 应用、dashboard 等等。

特别是学术出版，大家已经要遵循特定的要求。

本章最后给大家提几个建议：

- ▶ 千万不可以抄袭，引用注明出处。
- ▶ 图片除了要美，还要有效，每一幅图都要服务于一条完整故事链。
- ▶ 控制时间成本：和数值相关的可视化部分，建议用编程实现；美化元素建议后期软件处理。
- ▶ 某个可视化方案大量出图，建议写成通用函数。
- ▶ 风格尽量统一，比如颜色、线型、字体、字号、标注等等。
- ▶ 保持可视化的简洁和明了。避免冗余装饰，确保重点和主要信息清晰可见。
- ▶ 使用清晰的标签、标题和图例。
- ▶ 将可视化放入适当的上下文中，以便更好地理解和解释数据和信息。提供相关背景信息、注释和说明，帮助观众理解可视化的含义和重要性；
- ▶ 选择适当的颜色和配色方案，以增强可视化的视觉吸引力和可读性。确保颜色的使用符合信息的含义，避免使用过多的颜色，以免造成混乱。
- ▶ 在电子媒体出版时，考虑为可视化添加交互性和动态效果，以增加用户参与和理解。例如，可以使用交互式工具让用户自由探索数据，或者使用动画效果展示变化和趋势。
- ▶ 倾听反馈，迭代升级。不断提高可视化技能。



艺术可以用来将数学工具、数据转化为生动的可视化方案，帮助人们更好地理解和解释数学原理、挖掘数据背后的故事。

在数学工具、数据科学、机器学习、人工智能等应用场景，优质的“数学 + 艺术”可视化方案可以让人们发现数学之美、数据之美，甚至爱上数学。本书关注的正是这一点。

下面，正式邀请大家踏上本书的“数学 + 艺术”的美学实践之旅！

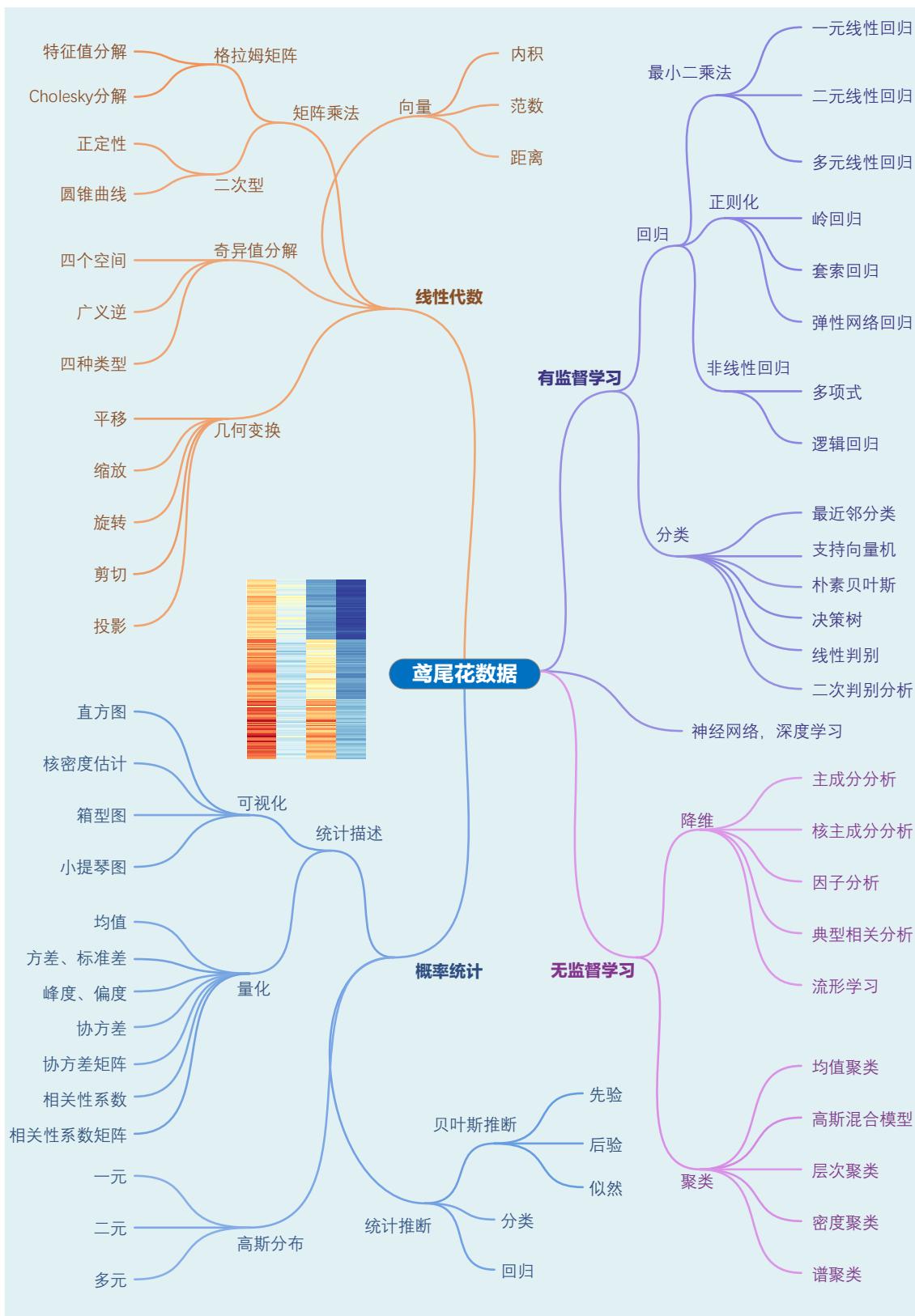


图 16. 有关鸳尾花数据的可视化“头脑风暴”

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

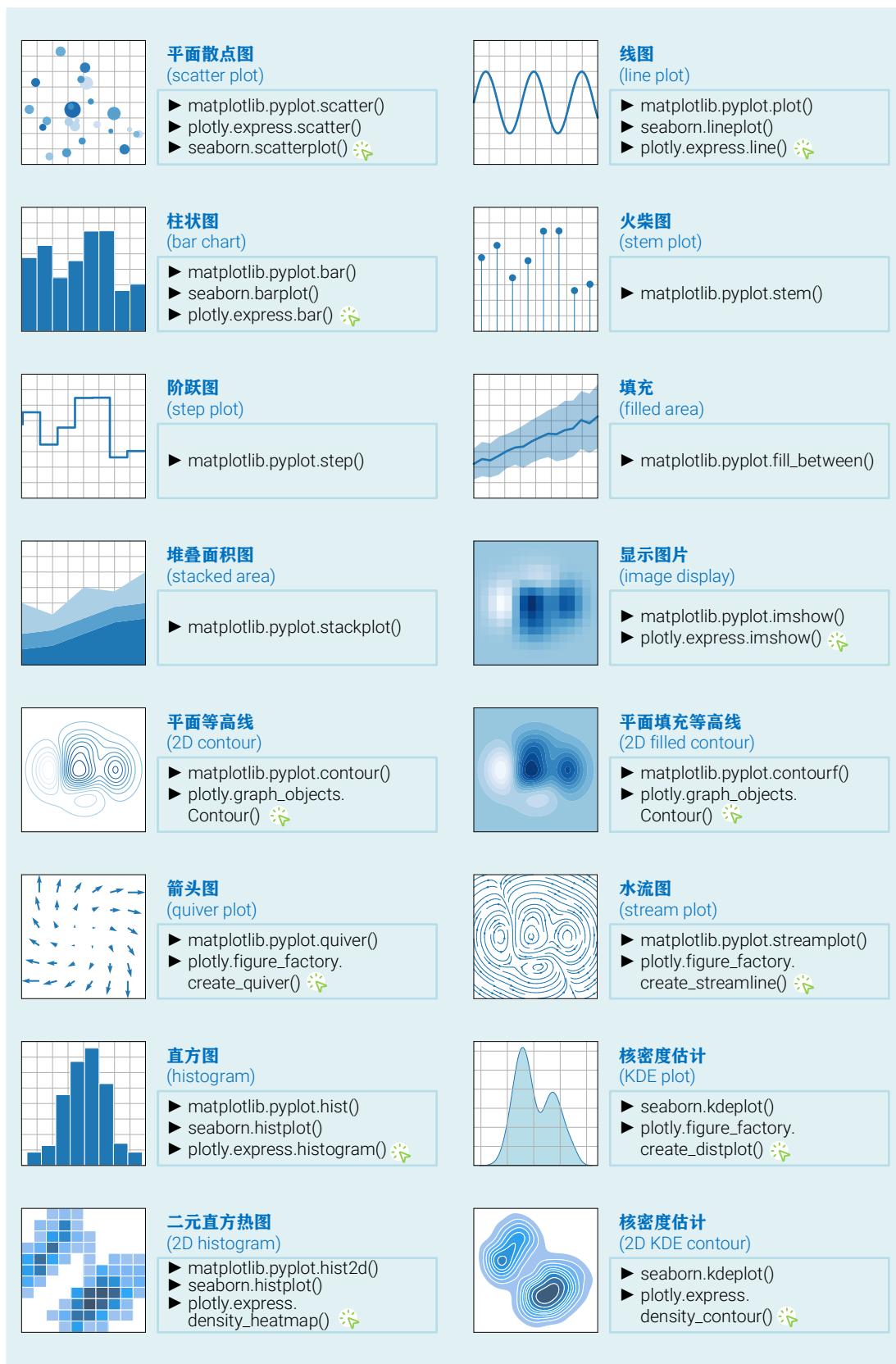


图 17. 鸢尾花书可视化方案目录，第一组

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

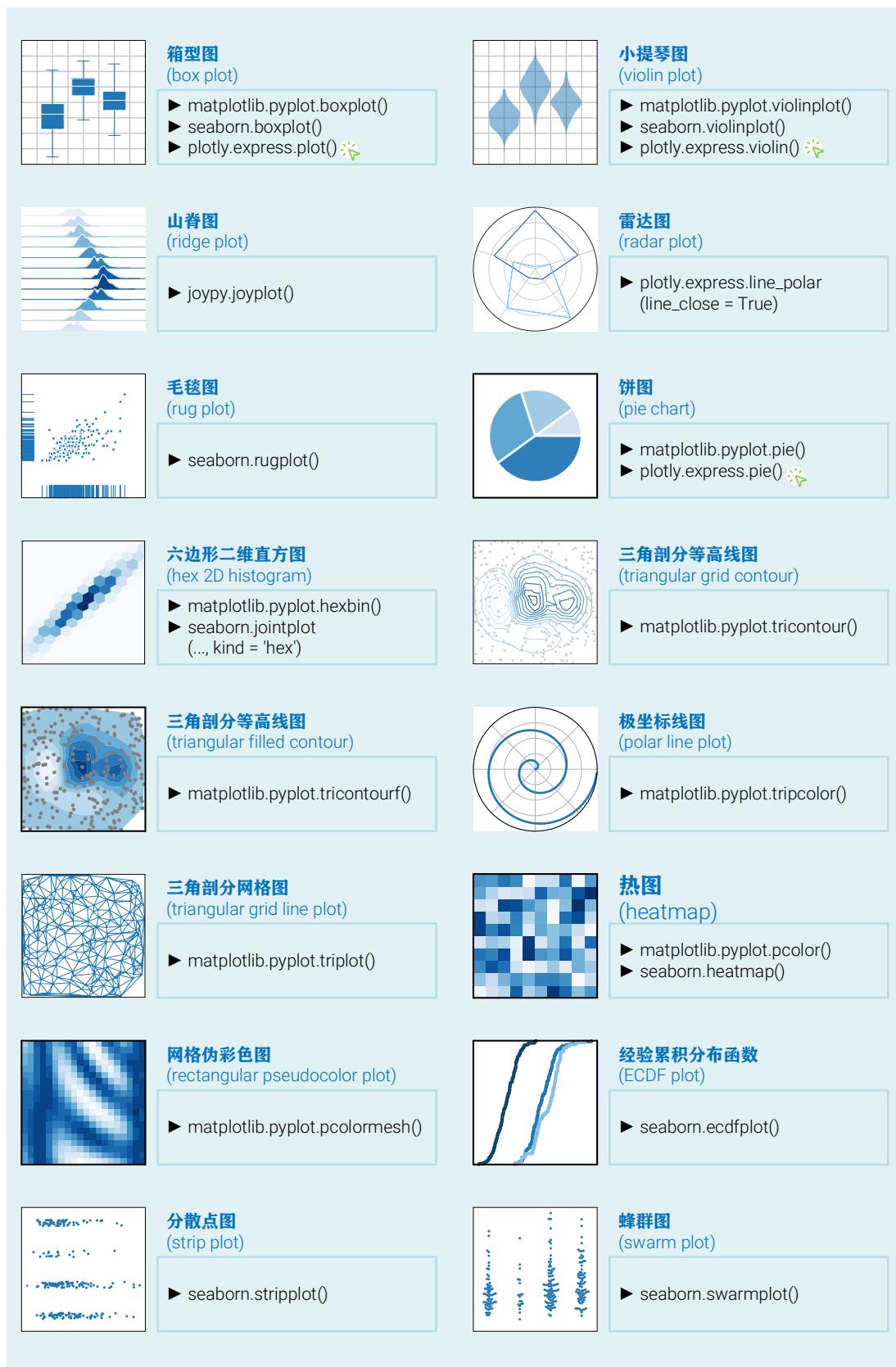


图 18. 鸢尾花书可视化方案目录，第二组

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

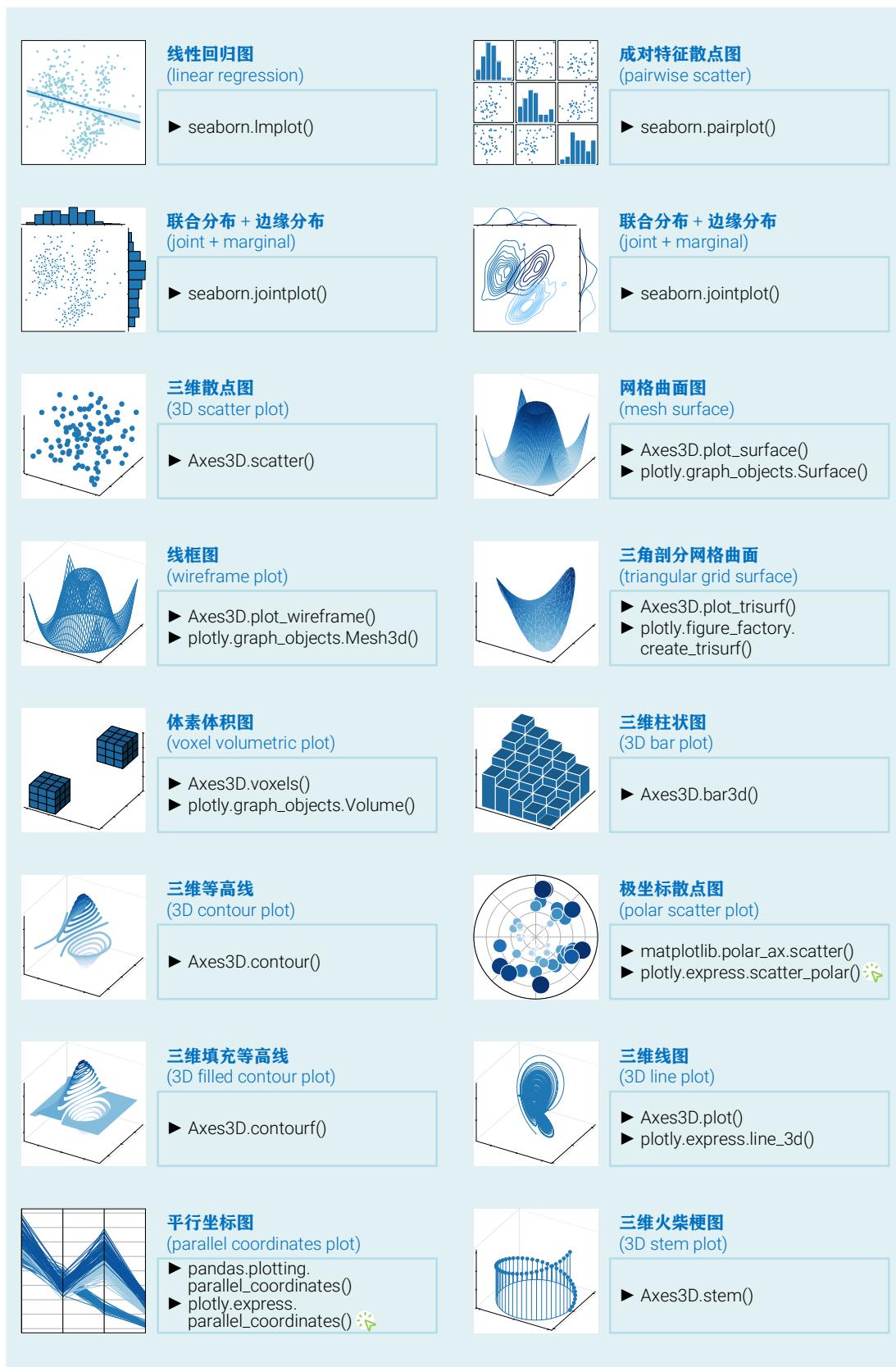


图 19. 鸢尾花书可视化方案目录，第三组

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

3

Layout of A Figure

布局

Matplotlib、Seaborn、Plotly 子图布局方案



艺术洗涤心灵的浮尘。

Art washes away from the soul the dust of everyday life.

—— 毕加索 (Pablo Picasso) | 西班牙艺术家 | 1881 ~ 1973



- ◀ XXXXX

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

3.1 图片规格

大小尺寸

Matplotlib 中默认图片尺寸为：宽 6.4 英寸，高 4.8 英寸。1 英寸 (inch) 约为 2.54 厘米 (cm)。也就是说默认图片尺寸为，宽约 16 厘米，高约 12 厘米。

默认图片以一个 dpi (dots per inch, 每英寸点数) 为 100 的分辨率显示。Matplotlib 图片尺寸可以通过修改参数来自定义图像的尺寸。使用 `figure` 函数，并指定 `figsize` 参数来设置图像的宽度和高度。`figsize` 参数接受一个元组 (宽度, 高度)，单位为英寸。

比如，`import matplotlib.pyplot as plt; plt.figure(figsize=(3, 3))` 将图像尺寸修改为 3 英寸 × 3 英寸。图 4 上图展示为 3 英寸 × 3 英寸图片真实大小。图 4 下图为换算为厘米的图片。

如果想保存图像到文件，可以使用 `savefig` 函数，并通过设置 `dpi` 参数来指定分辨率。例如，如果希望保存图像为 300dpi 的高质量 PNG 文件，可以用 `plt.savefig('plot_name.png', dpi=300)`。

边距

一张图少不了上下左右留白，这个留白就是边距 (margin)。在 matplotlib 默认情况下，图像周围边距为：

- ▶ `figure.subplot.left: 0.125`
- ▶ `figure.subplot.right: 0.9`
- ▶ `figure.subplot.top: 0.88`
- ▶ `figure.subplot.bottom: 0.11`

如图 1 所示，这些参数的值为 0 到 1 之间的浮点数，表示相对于图像的宽度或高度的百分比。

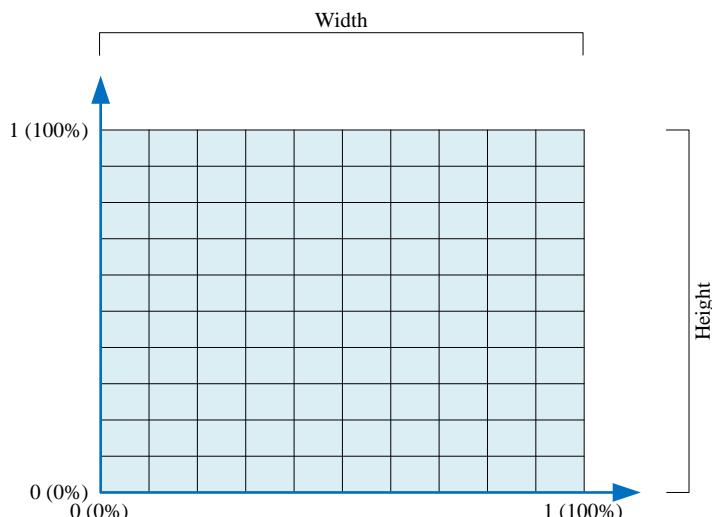


图 1. 图片宽度、高度百分比

如图 5 所示，默认情况下，宽度方向来看，`left = 0.125`，表示左边距相对于图像宽度的 12.5%，`right = 0.9` 表示右边距相对于图像宽度的 90%。`(0, 0)` 表示图形左下角，`(1, 1)` 表示图形右上角。

高度方向来看，`top = 0.88` 表示图片的顶边位于图片高度的 88%。而 `bottom = 0.1` 表示底边距相对于图像高度的 10%。

此外，绘制“图中图”时也需要类似的定位方式，具体如图 8 所示。

从图 5、图 8 中，我们可以看到一个 Figure 中可以有不止一幅子图。子图是指将整个图形区域划分为多个小的绘图区域，每个区域可以用于绘制不同的图形。下面我们就介绍 Matplotlib 常用的子图布局方法。

3.2 使用 subplot

在 Matplotlib 中，`subplot` 是一个函数，用于创建和管理图形中的子图。它的基本语法为 `subplot(nrows, ncols, index)`。其中，`nrows` 为子图的行数，`ncols` 为子图的列数，`index`：当前子图的索引（从 1 开始，按先行后列顺序递增）。

图 6 给出示例介绍如何用 `subplot()` 绘制子图，并分别修饰子图。这个例子来自于 *Scientific Visualization: Python & Matplotlib*。

注意，这幅图中的文本（英文、数字）已经扁平化（flatten）。在矢量图中，“flatten”通常指的是将文本对象转换为矢量线条，以便更好地支持艺术处理。当文本以矢量形式表示时，它由数学定义的几何形状组成。下一章，将介绍如何在 Matplotlib 图像中完成注释。

3.3 使用 add_subplot

`add_subplot()` 函数用于在图形中添加子图。`add_subplot()` 的基本语法为 `add_subplot(nrows, ncols, index)`。其中，`nrows` 为子图的行数，`ncols` 为子图的列数，`index` 为当前子图的索引（从 1 开始，先行后列顺序递增）。

`add_subplot()` 返回一个 `AxesSubplot` 对象，它表示创建的子图。我们可以使用此对象进行进一步的图形操作，例如绘制数据、设置轴标签和标题等。

`add_subplot()` 和 `subplot()` 在功能上是相似的，都可以用于创建和管理图形中的子图。它们的主要区别在于使用方式和语法。

`add_subplot()` 是 `Figure` 对象的方法，用于在特定的 `Figure` 上添加子图。`add_subplot()` 语法为 `fig.add_subplot(nrows, ncols, index)`。因此，使用 `add_subplot()` 方法时，首先需要创建一个 `Figure` 对象，然后调用该方法来添加子图，并将子图对象存储在变量中以进行后续的操作。

`subplot()`是 `pyplot` 模块的函数，用于在当前的图形中添加子图。`subplot()` 语法为 `plt.subplot(nrows, ncols, index)`。使用 `subplot()` 函数时，不需要显式地创建 `Figure` 对象。可以直接调用 `subplot()` 函数，并在同一个代码块中添加多个子图。

图 7 中子图混合平面和三维可视化方案。

此外，我们还可以使用 `insert_axes` 和 `add_axes` 在指定位置插入特定宽高的图像。位置、宽高这四个数值均为 0 和 1 之间的浮点数，代表百分比。`insert_axes` 是 `Figure` 对象的方法，用于在指定位置插入轴。`add_axes` 是 `Figure` 对象或 `Subplot` 对象的方法，用于在指定的图形或子图中添加轴。图 8 给出两个例子。

本章利用了很多对图轴、图脊的操作，这些内容将在下一章系统讲解。

3.4 使用 subplots

在 `Matplotlib` 中，`subplots` 函数用于创建一个包含多个子图的图形布局，并返回一个包含子图对象的元组。以下是使用 `subplots` 函数的基本步骤：

- ▶ 导入绘图模块 `import matplotlib.pyplot as plt`
- ▶ 使用 `subplots` 函数创建子图 `fig, axes = plt.subplots(nrows, ncols)`，其中 `nrows` 和 `ncols` 是整数，分别表示子图行和列的数量。
- ▶ `axes[i, j]` 表示在第 `i` 行和第 `j` 列的位置上的子图对象。

此外，可以对每个子图进行布局调整和美化。

图 9 所示为利用 `subplots` 绘制的一元高斯分布概率密度函数曲线随 μ 、 σ 变化。高斯分布，也被称为正态分布，是概率论和统计学中一种常见的连续概率分布。一元高斯分布以钟形曲线的形式表示，具有对称的特点。

一元高斯分布由两个参数完全描述：均值 μ 和标准差 σ 。均值确定了分布的中心位置，标准差决定了分布的形状和展宽程度。一元高斯分布在许多领域中具有广泛的应用，比如统计描述、统计推断、数据分析建模、机器学习等。

图 10 所示为使用 `subplots` 绘制 Beta 分布概率密度函数曲线随 α 、 β 变化。

Beta 分布是概率论和统计学中常见的连续概率分布，它定义在 0 到 1 之间，并具有灵活的形状。Beta 分布由两个形状参数 α 、 β 控制，用于描述随机变量在 0 到 1 之间的概率分布。Beta 分布在许多领域中有广泛的应用，比如概率建模、贝叶斯推断等。

3.5 使用 GridSpec

在 `Matplotlib` 中，`GridSpec` 是一个用于灵活地布局子图的工具。它允许在绘图区域中创建规则的网格，并指定每个子图的大小、位置和跨过的行列数。

使用 GridSpec，可以以更高级的方式组织和排列多个子图，而不是使用默认的单行单列布局。这对于创建复杂的图形布局非常有用，例如在一个绘图区域中显示多个子图，并使它们具有不同的大小和位置。

图 2 所示为 4×4 网格中两种子图布局。大家可能已经发现这利用了鸢尾花书《编程不难》中介绍的索引和切片。图 2 中蓝色子图为主图，主图分别位于右上、左下。图 11 所示为使用 GridSpec 绘制满足二元高斯分布的随机数散点图和边缘分布直方图。还可以通过 GridSpec 定义子图的宽度比例、高度比例，如图 3 所示。图 12 可视化 Dirichlet 分布、边缘 Beta 分布，子图的宽度比例、高度比例都是 1:3。

本章后文还会介绍如何用 add_gridspec 函数完成类似的可视化方案。

此外，请大家思考通过怎样索引和切片布置能够让主图位于右下、左上。

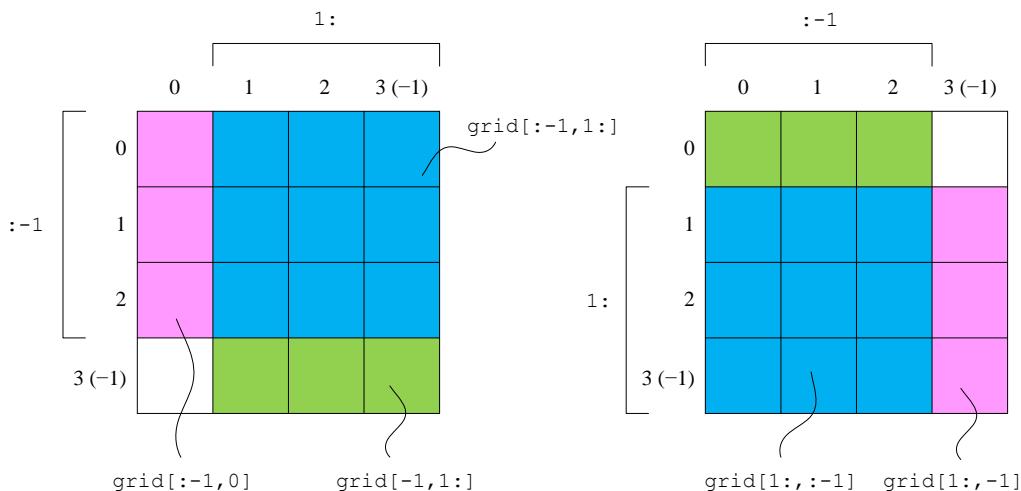


图 2. 4×4 网格中两种子图布局

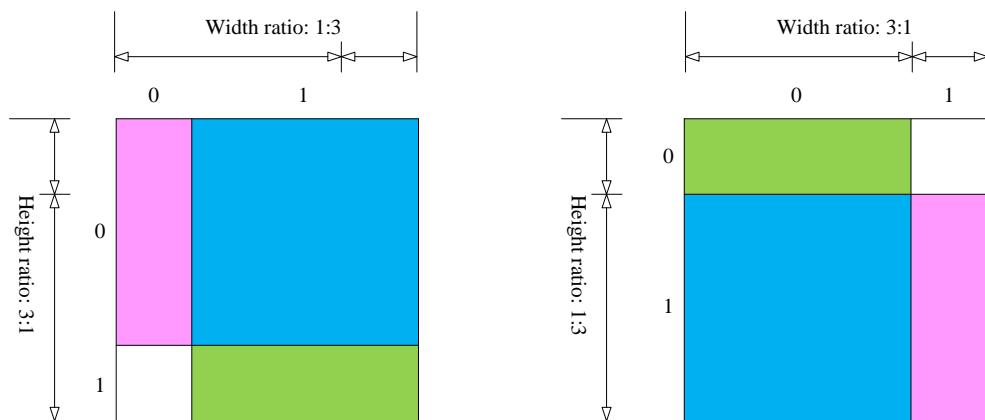


图 3. 2×2 网格中两种子图布局以及宽高比例调整

3.6 使用 add_gridspec

在 Matplotlib 中，`add_gridspec` 函数可以用来创建复杂的图形布局。它允许你在图形中创建多个子图，并指定它们的位置和大小。使用 `add_gridspec` 函数的基本步骤如下：

- ▶ 首先导入 `import matplotlib.pyplot as plt`。
- ▶ 然后创建一个 `Figure` 对象 `fig = plt.figure()`。
- ▶ 再创建一个 `GridSpec` 对象 `gs = fig.add_gridspec(nrows, ncols)`，其中 `nrows` 和 `ncols` 是整数，分别表示行和列的数量。
- ▶ 最后可以使用 `GridSpec` 对象创建子图，`gs[i, j]` 表示在第 `i` 行和第 `j` 列的位置创建一个子图。可以使用切片语法来指定多个位置。

图 13 所示为利用 `add_gridspec` 函数绘制的满足二元高斯分布散点图，这幅图中还绘制了高斯核概率密度估计等高线、边缘概率密度曲线。

图 14、图 15、图 16 所示为利用 `add_gridspec` 函数绘制的利萨茹曲线随 n_x 、 n_y 、 k 变化。

利萨茹曲线 (Lissajous curve) 是一种在平面上生成的特殊曲线，它由两个正弦函数的振幅和频率组合而成。利萨茹曲线通常以参数方程的形式表示。当两个振幅和频率的比例不同或相位差不同时，曲线的形状会发生变化。

利萨茹曲线的形状取决于两个振幅和频率的关系，它可以展示出丰富的几何图案，包括直线、椭圆、环形等。利萨茹曲线具有美观和艺术性。因此，利萨茹曲线常被用作数学教学、科学可视化和艺术创作的工具。

在使用 `add_gridspec` 函数绘制子图时同样可以设定轴类型，比如三维、极坐标等等。

在 Matplotlib 中，`subgridspec` 函数用于创建一个更细粒度的子图网格布局，即嵌套子图。它允许你在一个更大的图形布局中创建具有不同大小和位置的子图。图 18 所示为利用 `subgridspec` 函数创建的嵌套子图，这个例子来自 Matplotlib 官网。

Matplotlib 最近还推出了马赛克 `mosaic` 函数用来完成子图布置，请大家自行学习：

https://matplotlib.org/stable/gallery/subplots_axes_and_figures/mosaic.html

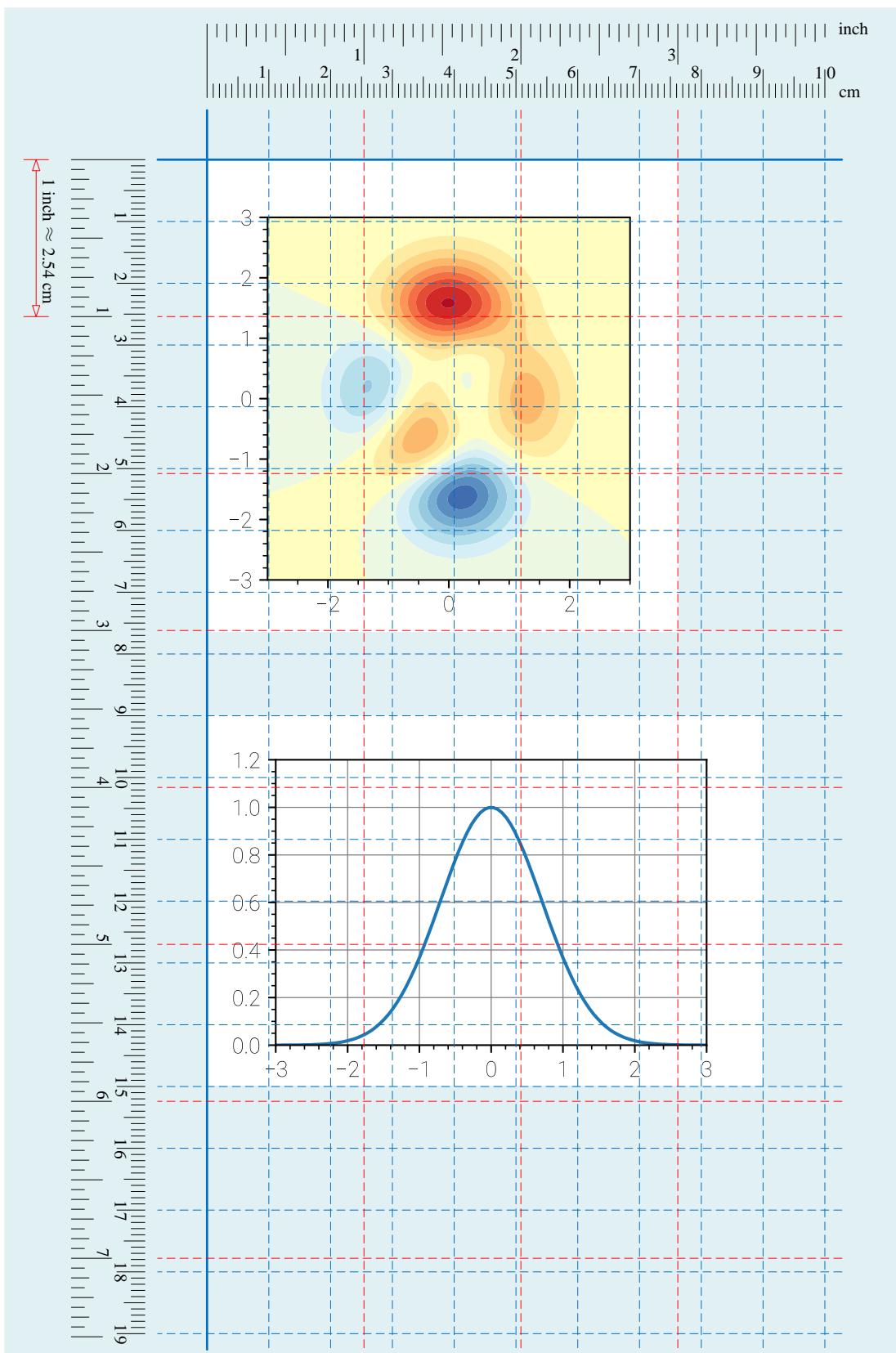


图 4. 自定义图片尺寸

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

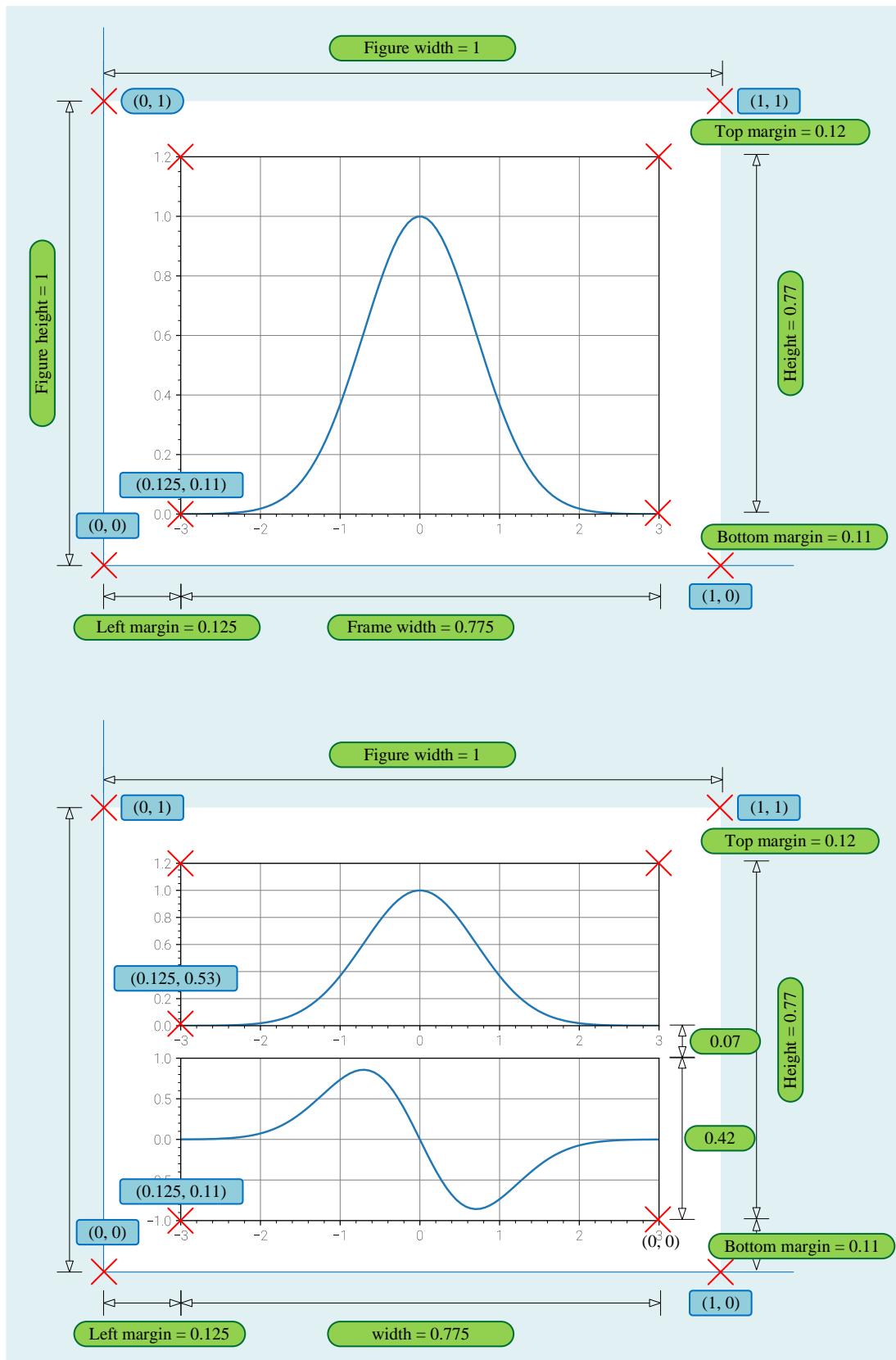


图 5. 图片边距

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

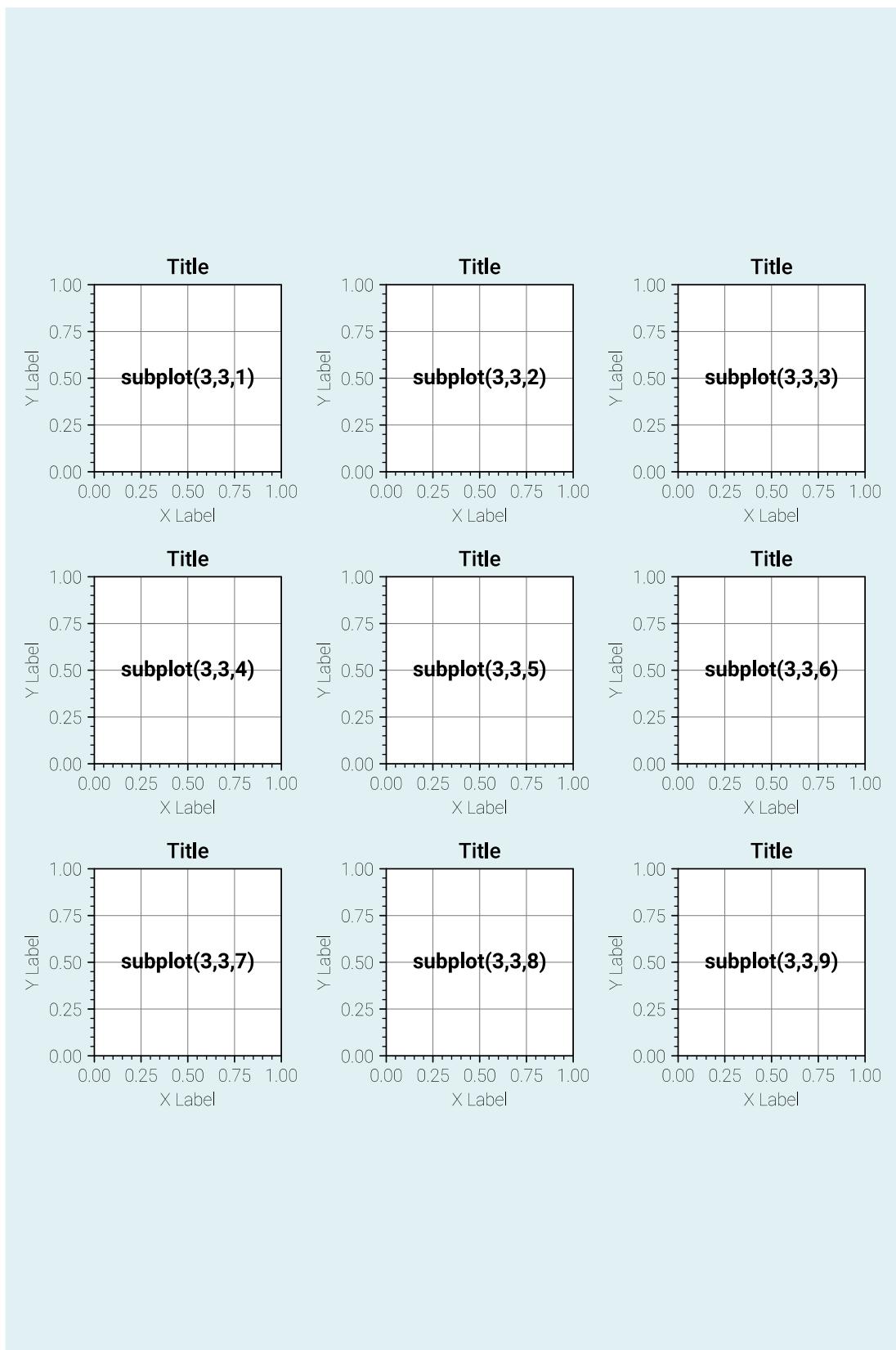


图 6. 使用 subplot 完成子图设置，来源 <https://github.com/rougier/scientific-visualization-book>

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

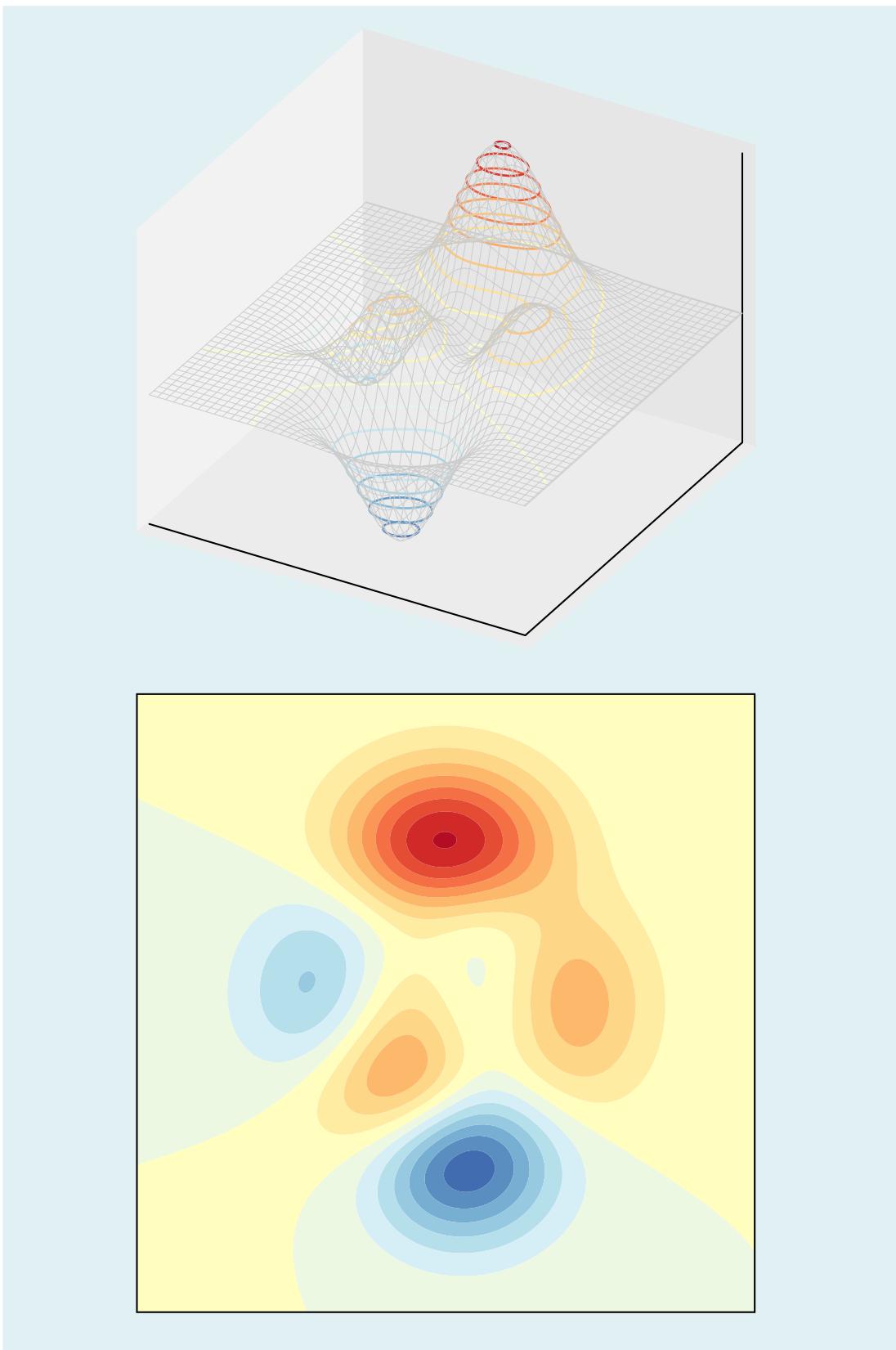


图 7. 使用 add_subplot 混合二维、三维可视化方案

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

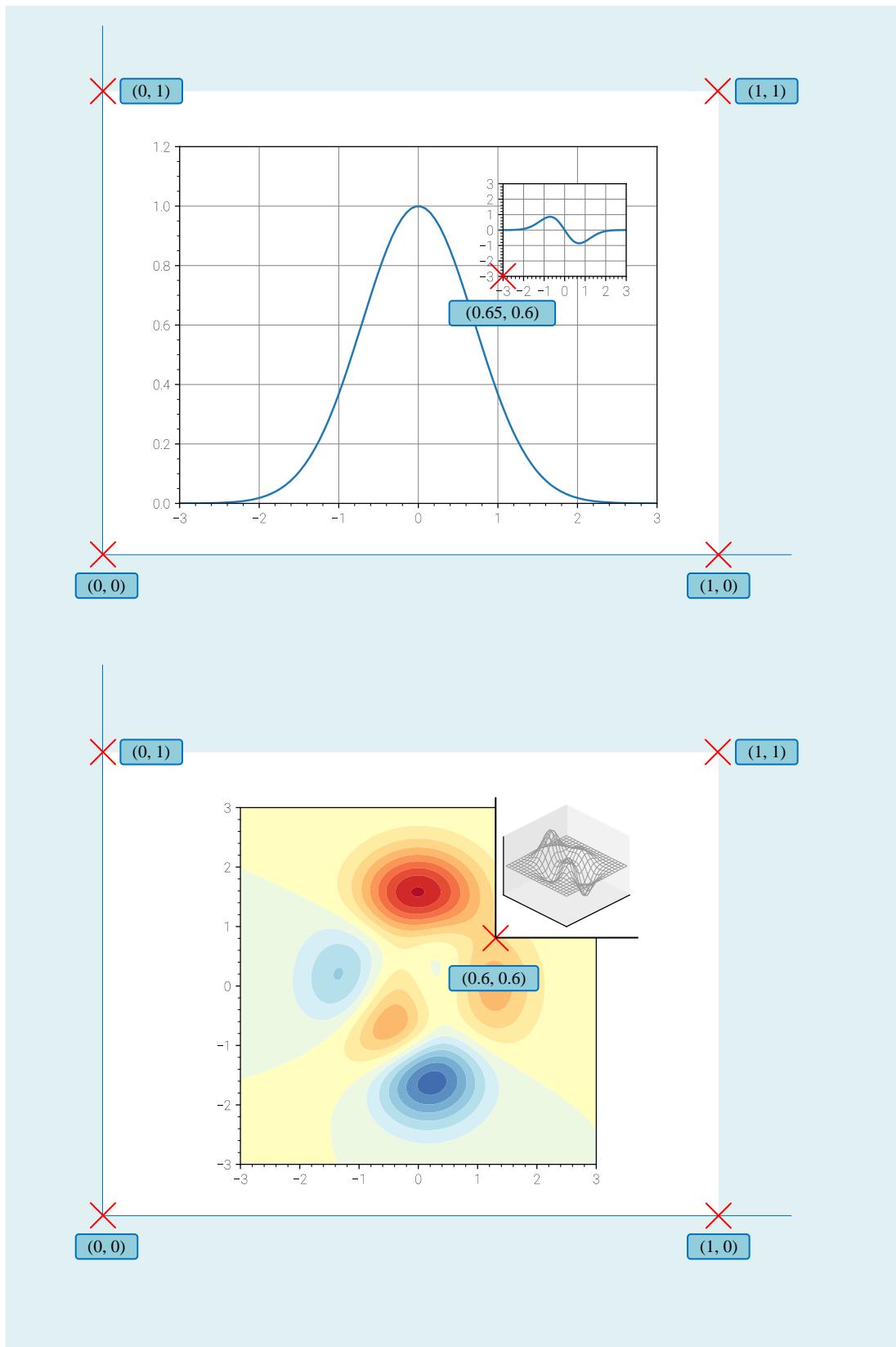


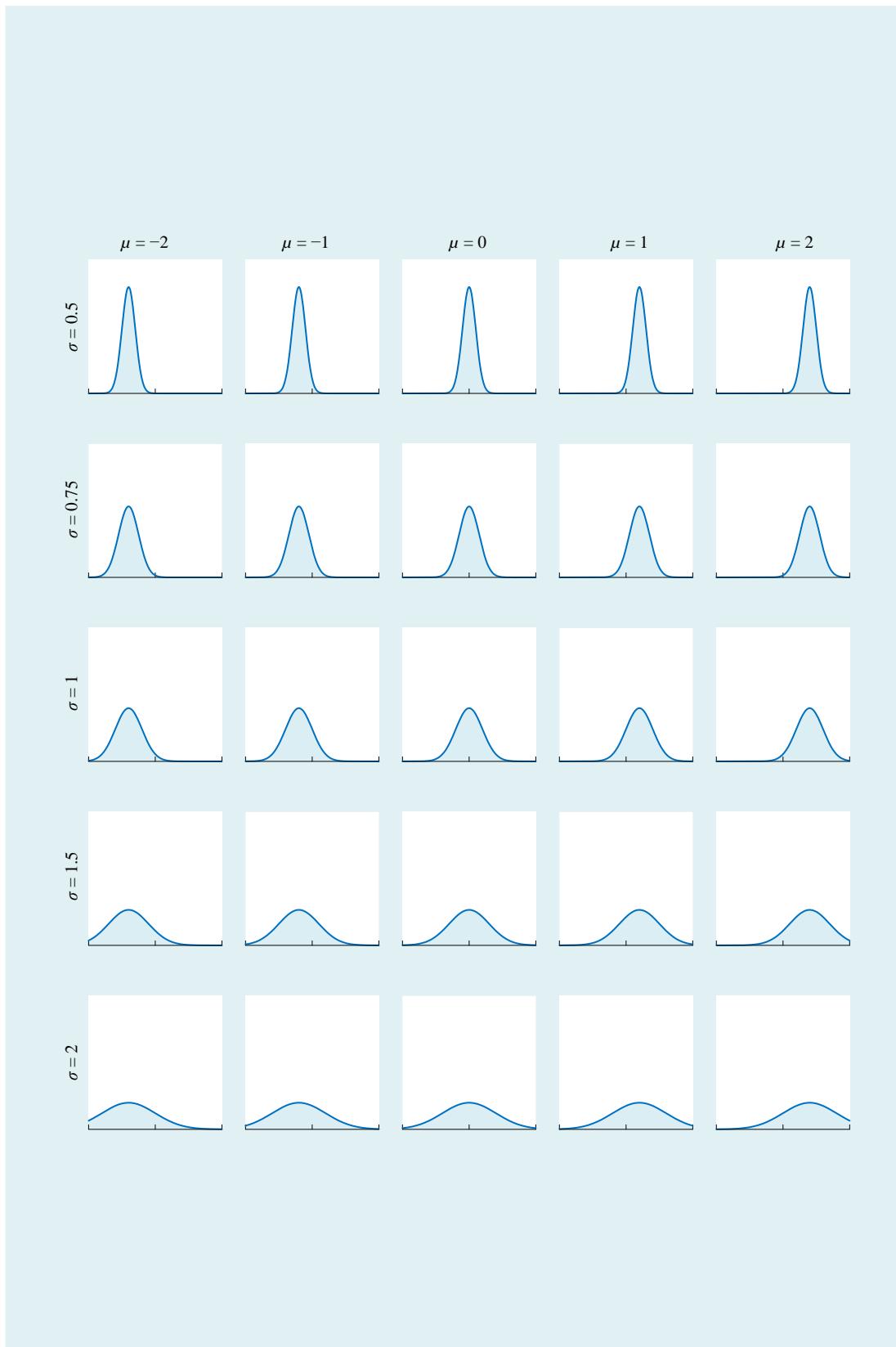
图 8. “图中图”定位

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

图 9. 使用 subplots 绘制一元高斯分布概率密度函数曲线随 μ 、 σ 变化

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

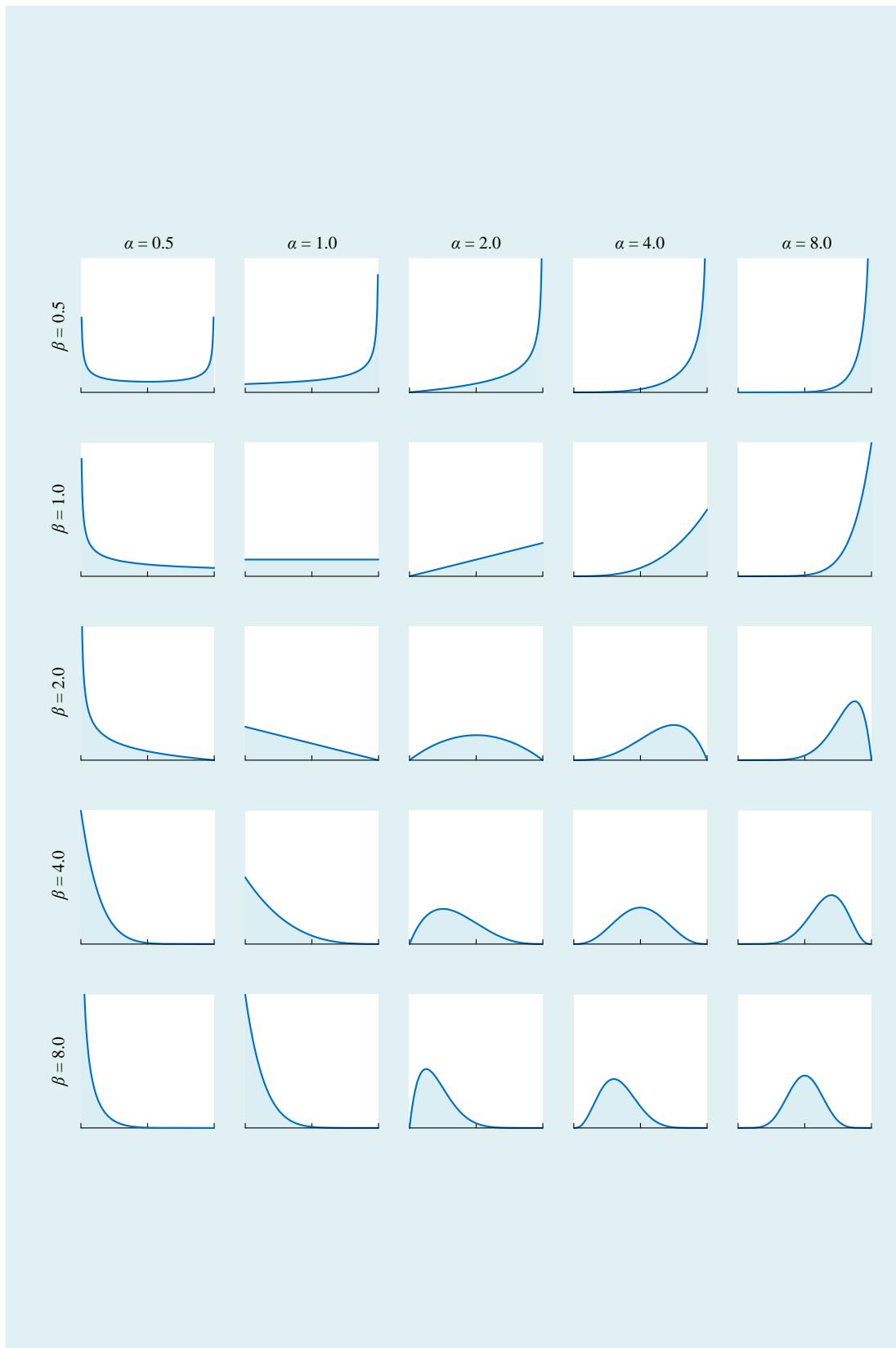


图 10. 使用 subplots 绘制 Beta 分布概率密度函数曲线随 α 、 β 变化

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

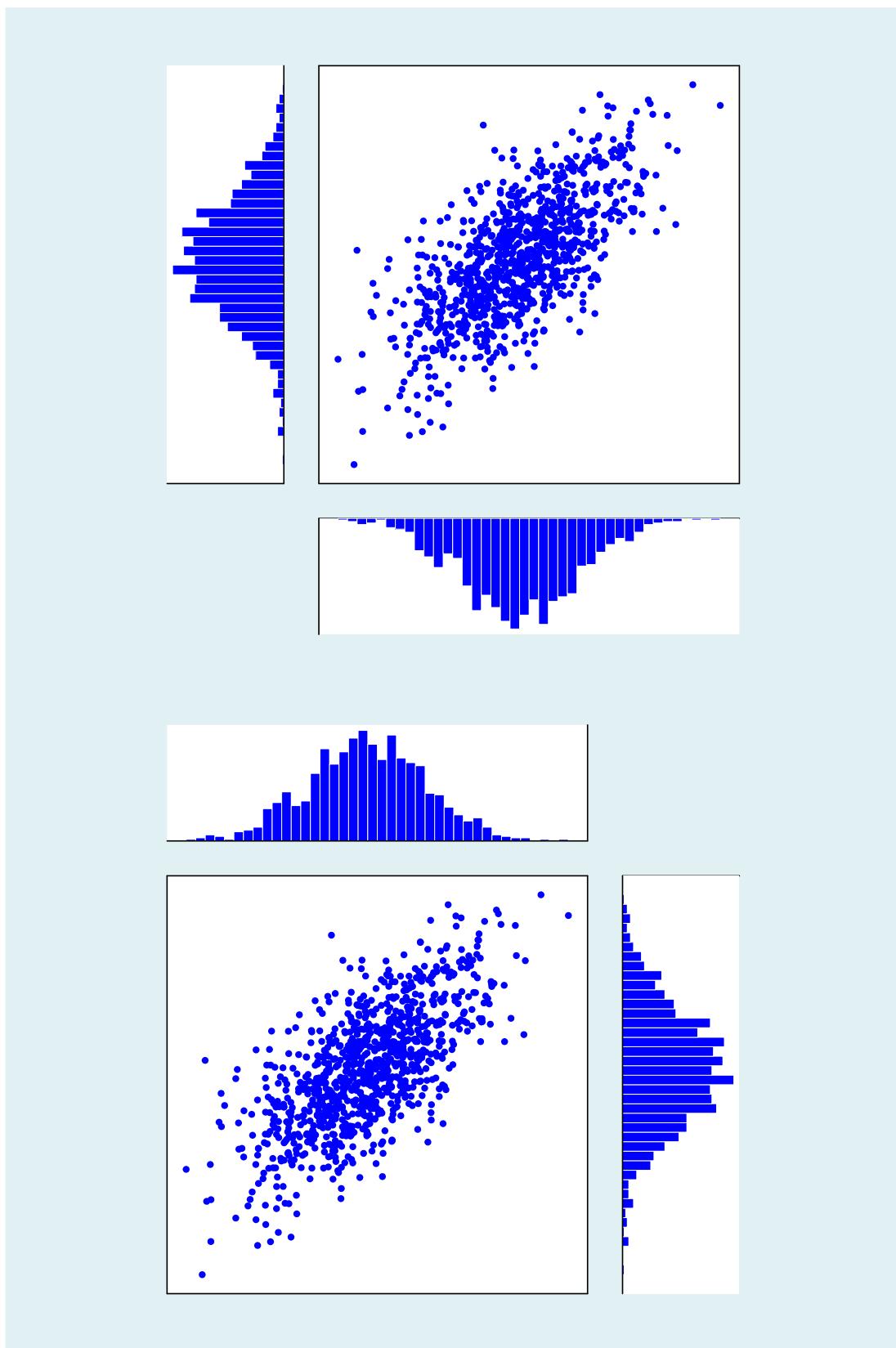


图 11. 使用 GridSpec 绘制满足二元高斯分布的随机数散点图和边缘分布直方图

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

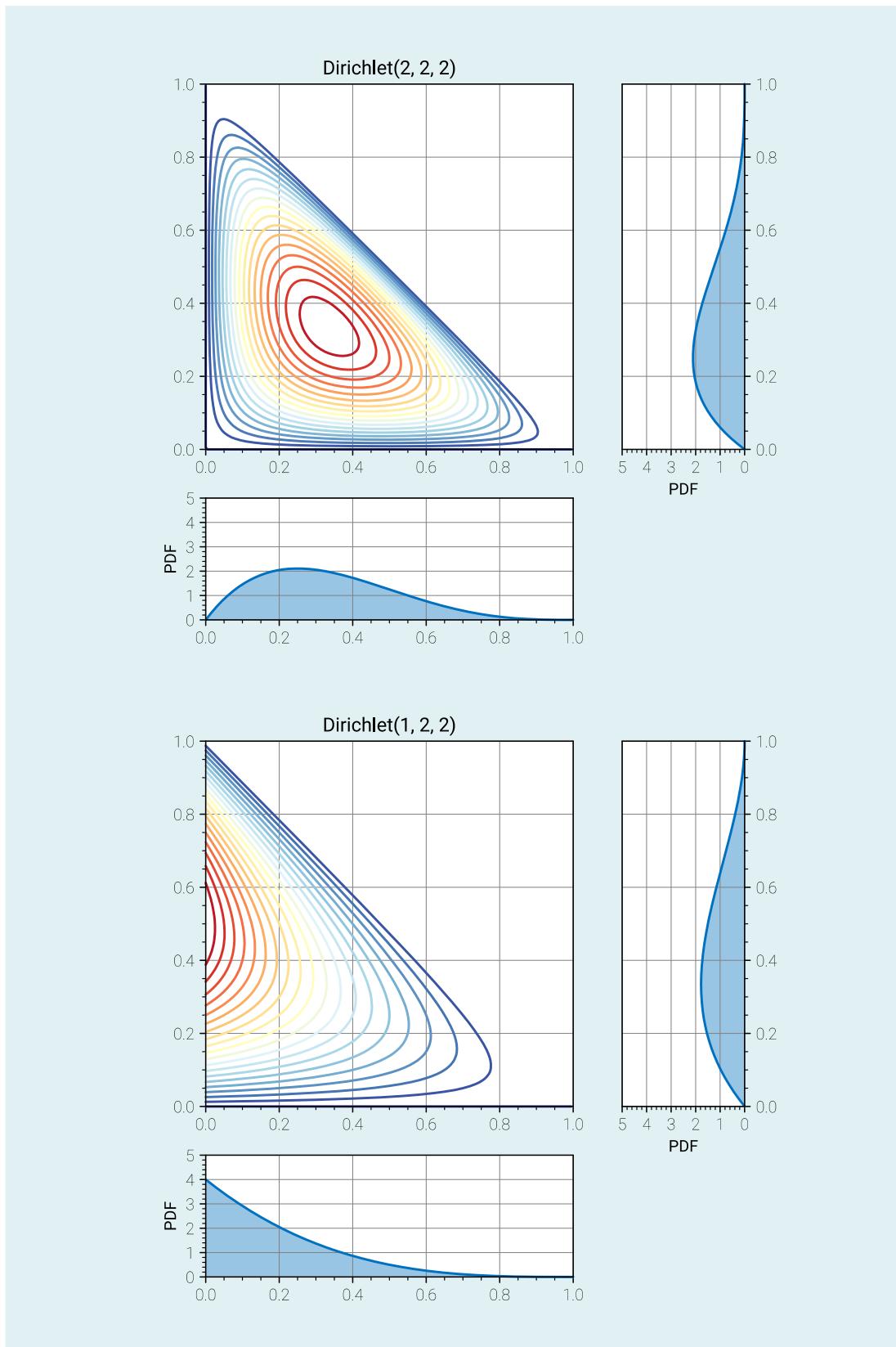


图 12. 使用 GridSpec 绘制的 Dirichlet 分布和边缘 Beta 分布

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

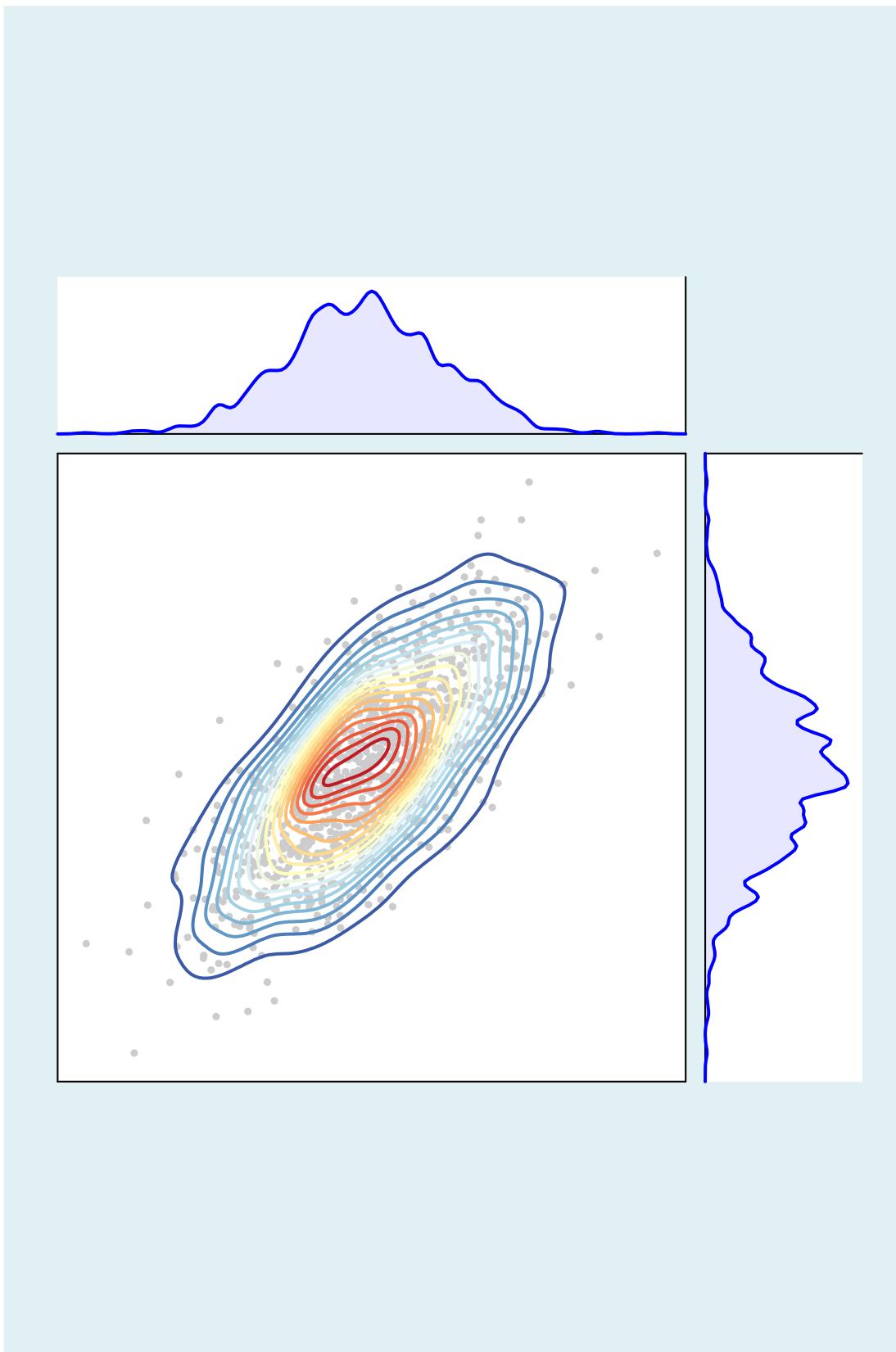


图 13. 使用 add_gridspec 绘制满足二元高斯分布的随机数散点图和边缘分布高斯核概率密度估计

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

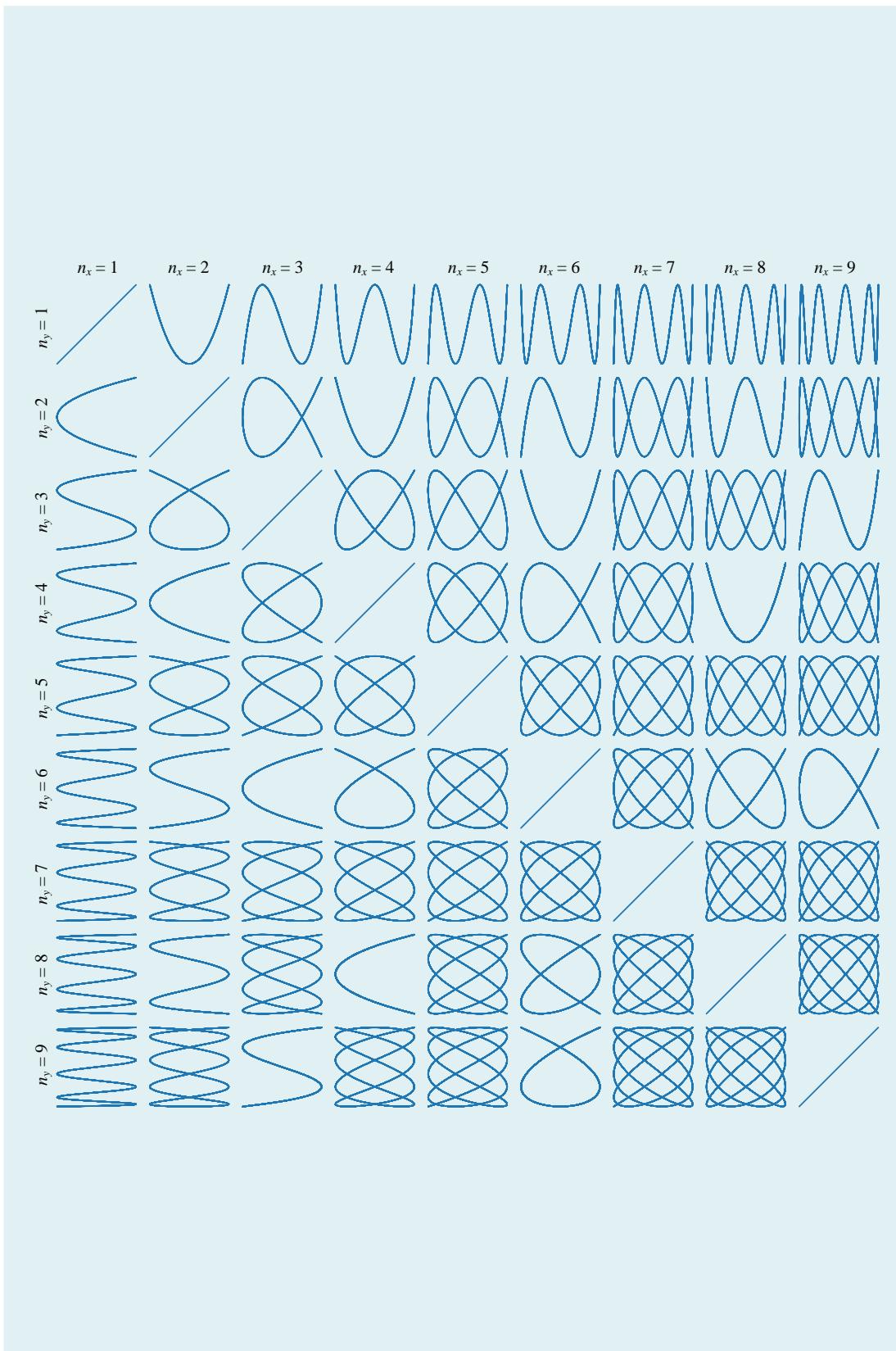


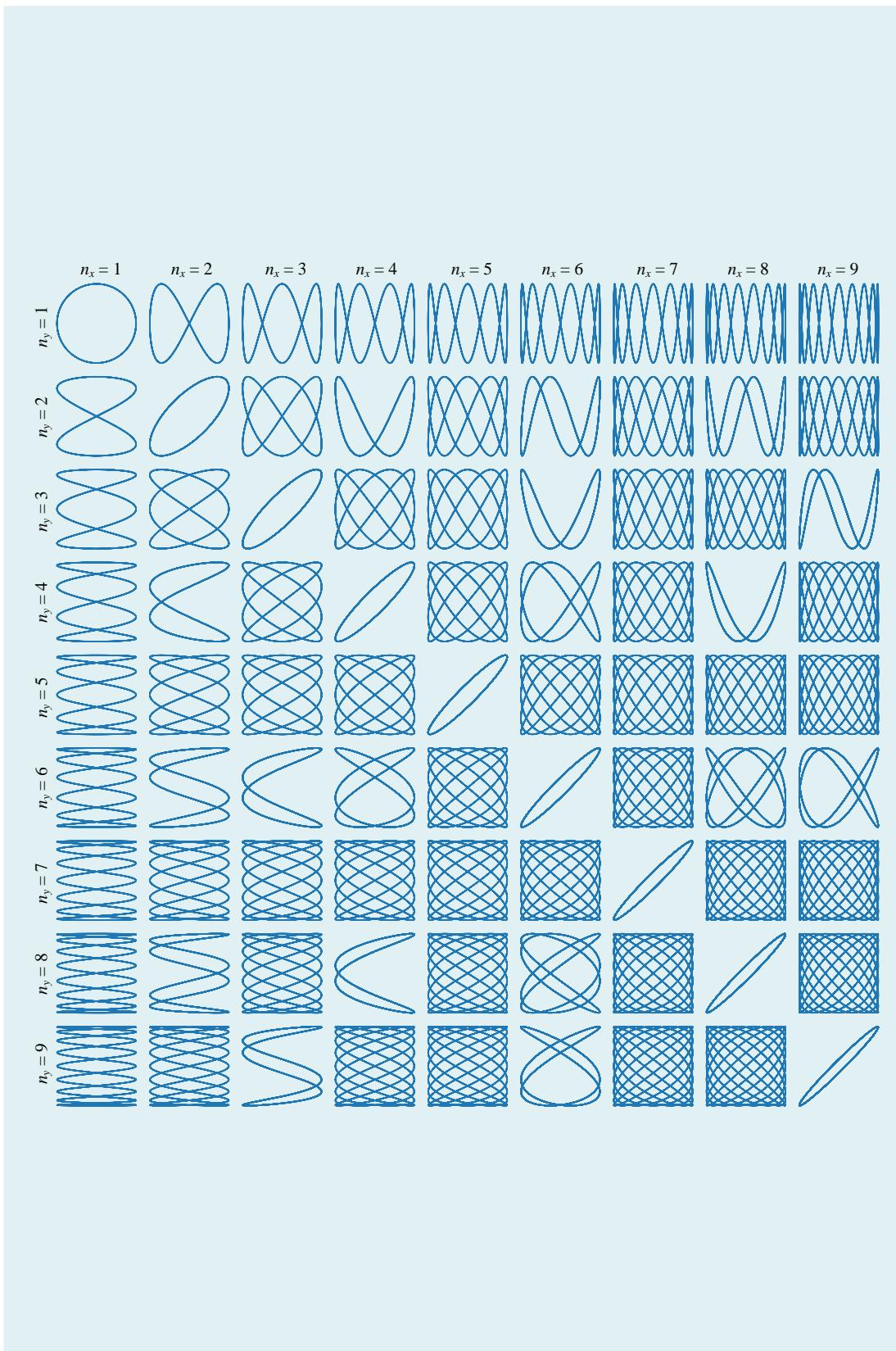
图 14. 利萨茹曲线, $k = 0$

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

图 15. 利萨茹曲线, $k = 2$

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

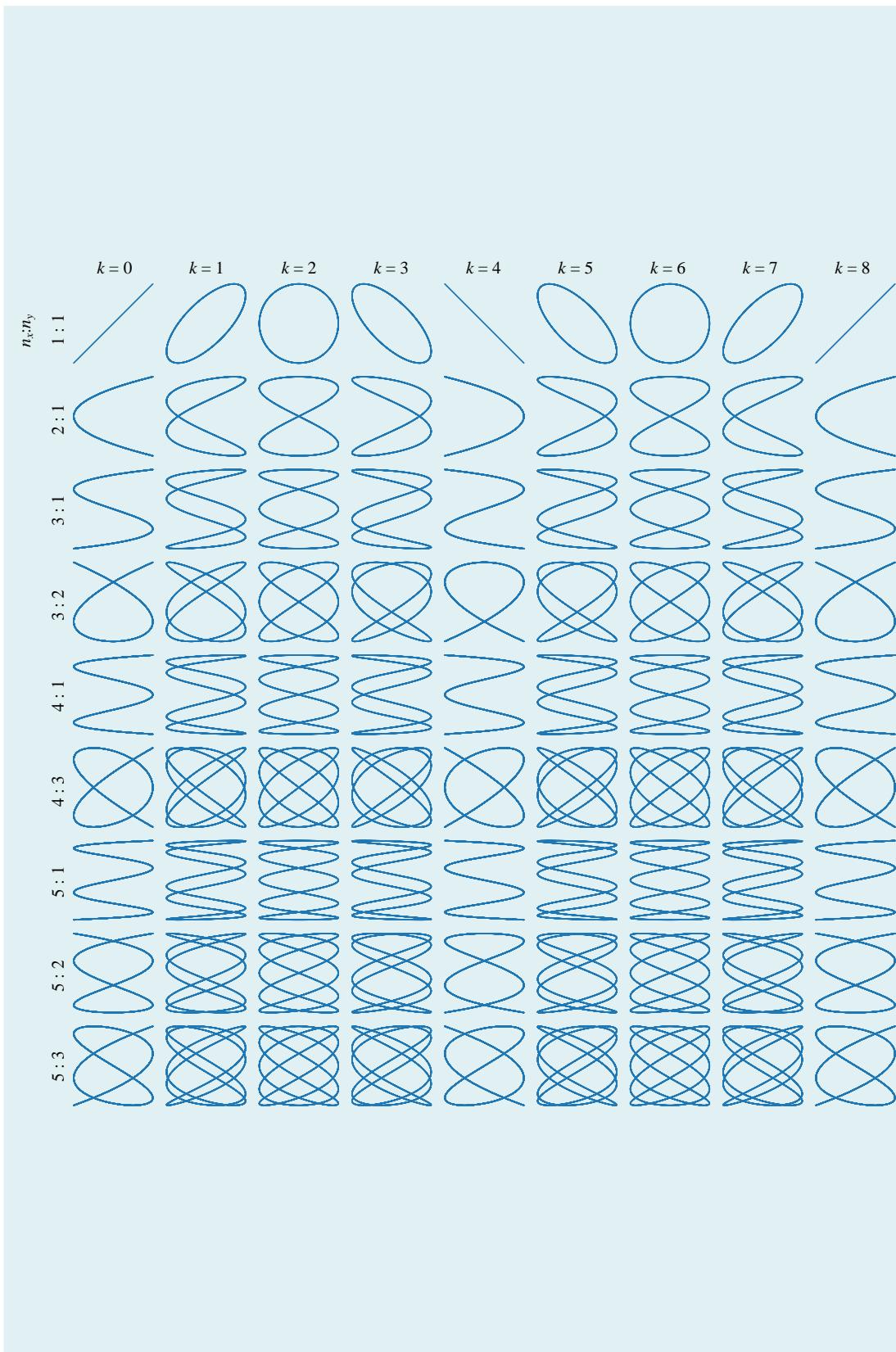


图 16. 利萨茹曲线, k 变化

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

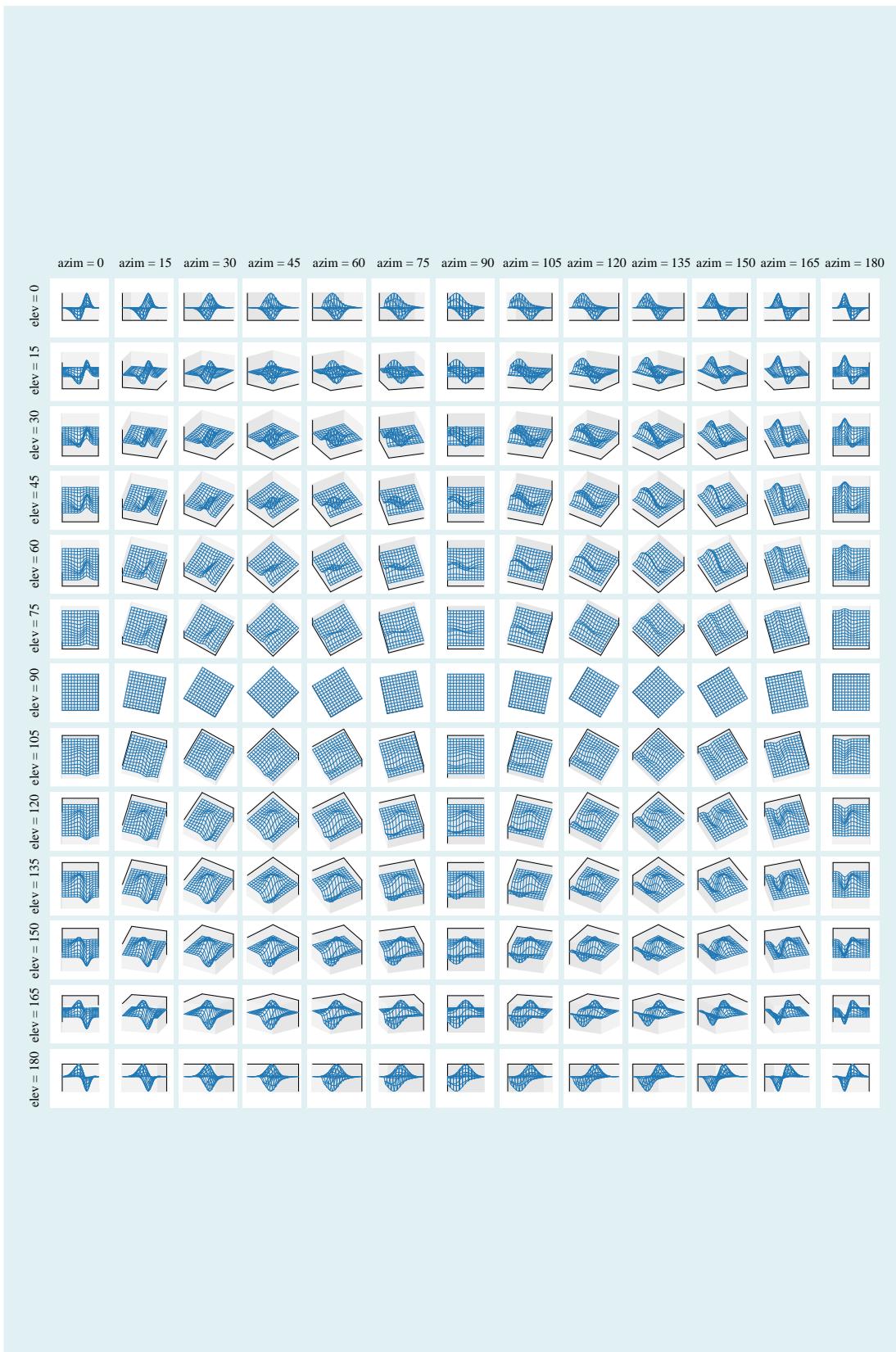


图 17. 利用 add_gridspec 函数绘制子图展示三维网格面随视角变化

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

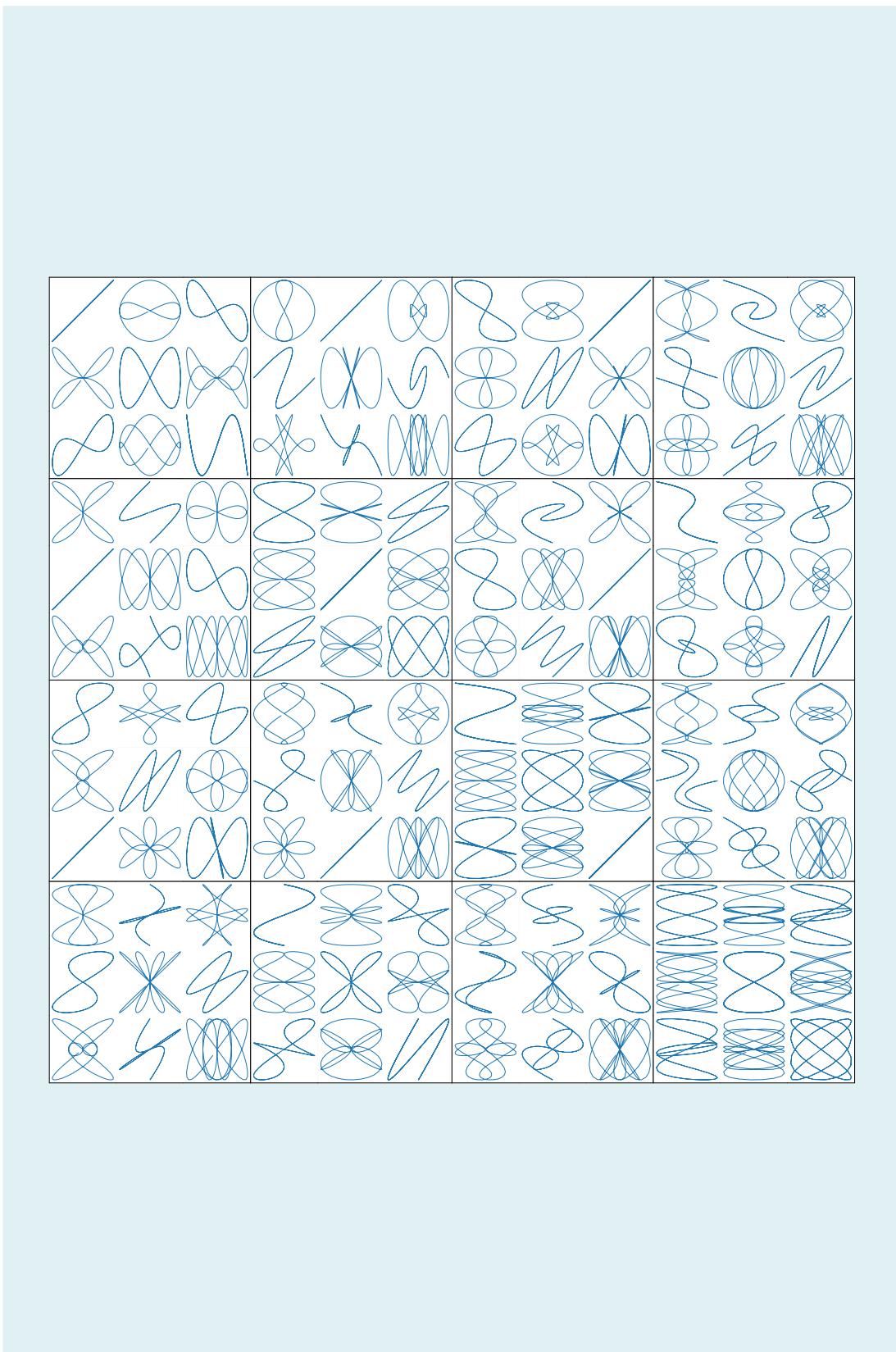


图 18. 利用 subgridspec 函数创建嵌套子图

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

4 Decorations 装饰

图脊、图轴、网格、线条、标注…个个都是艺术家



这个世界就是我们想象力的画布。

This world is but a canvas to our imagination.

——亨利·戴维·梭罗 (Henry David Thoreau) | 作家、诗人 | 1817 ~ 1862



- ◀ XXXXX

4.1 艺术家

在 Matplotlib 中，艺术家 (artist) 是指图形的每个可见元素，如图脊、图轴、坐标轴、标题、标签、图例、线条、网格、色块等等。每个艺术家对象都有自己的属性和方法，用于控制其外观和行为。

如图 1 所示，图形艺术家构成了一个层级结构。艺术家是图形对象 (Figure)，它包含了所有其他艺术家。图形外框是图脊 (frame)，图形对象之下是图轴 (axis)，用于绘制数据和刻度线。图轴各种其他艺术家，如刻度线、刻度标签和轴标题。图 2 所示为一幅平面等高线中艺术家的层级结构。

绘图时，可以用 zorder 控制艺术家在层级结构中的放置顺序。

本章将专门介绍图脊、图轴、注释这三个艺术家。

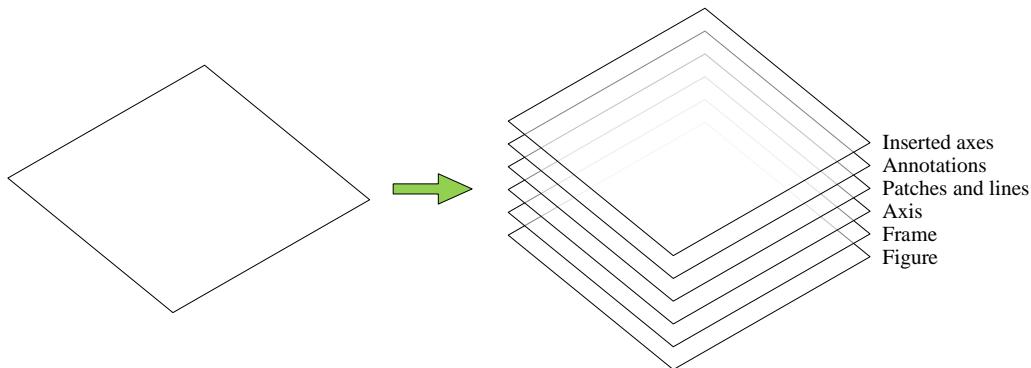


图 1. 一张图片的层级结构

4.2 图脊

在 Matplotlib 中，图脊 (spine) 是指图形中的边框线，用于界定图形的边界。图脊由四条边框线组成：上脊 (top spine)、下脊 (bottom spine)、左脊 (left spine) 和右脊 (right spine)。这些脊线可以通过 Matplotlib 的 Axes.spines 属性进行访问和定制。

图 3、图 4、图 5 所示为一些修改图脊设计方案，图 6 所示为背景网格样式设计方案，请大家参考本章配套 Jupyter Notebook 自行学习。

4.3 图轴

在 Matplotlib 中，图轴 (axis) 是指图形中的坐标轴，用于表示数据的数值范围和刻度。图轴包括轴线 (axis line)、刻度 (tick)、轴标签 (axis label)。

图 7、图 8、图 9 所示为一些修改图轴设计方案，图 10 所示为对数坐标，请大家参考本章配套 Jupyter Notebook 自行学习。

4.4 注释

一幅图的注释可以有很多种类型，比如：

- ▶ 标题 (title)，对应函数为；
- ▶ 横轴标题 (x axis label)，对应函数为 `matplotlib.pyplot.xlabel()` 或 `ax.set_xlabel()`；
- ▶ 纵轴标题 (y axis label)，对应函数为 `matplotlib.pyplot.ylabel()` 或 `ax.set_ylabel()`；
- ▶ 图例 (legend)，对应函数为 `matplotlib.pyplot.legend()`；
- ▶ 添加文字 (text)，对应函数为 `matplotlib.pyplot.text()`。

图 11 给出的几种标题、轴标题、图例布置方案。

本书不特别展开讲解如何用 `text()` 添加文字，请大家自行学习：

https://matplotlib.org/stable/api/_as_gen/matplotlib.pyplot.text.html

https://matplotlib.org/stable/gallery/text_labels_and_annotations/index.html

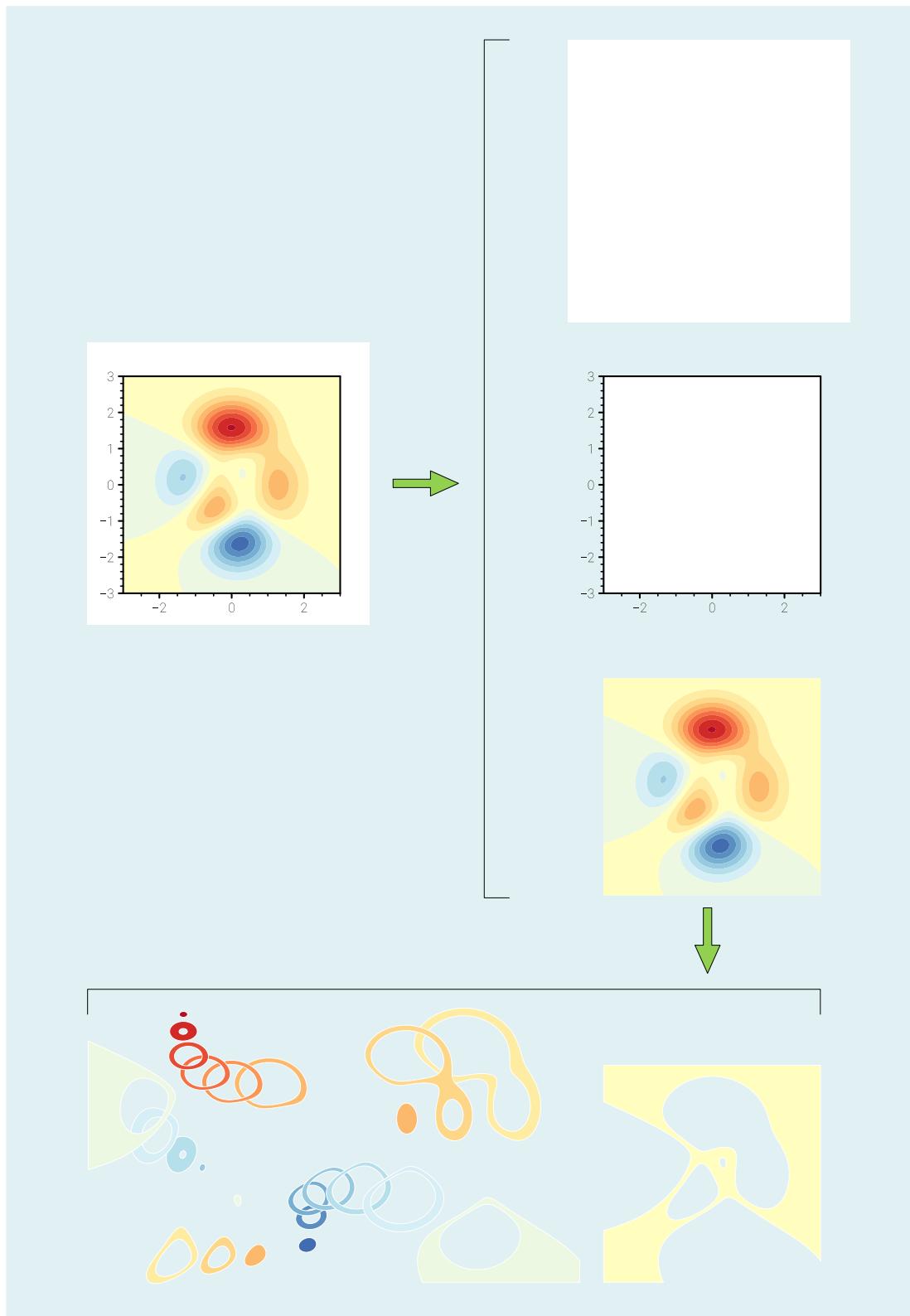


图 2. 一幅等高线的层级结构

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

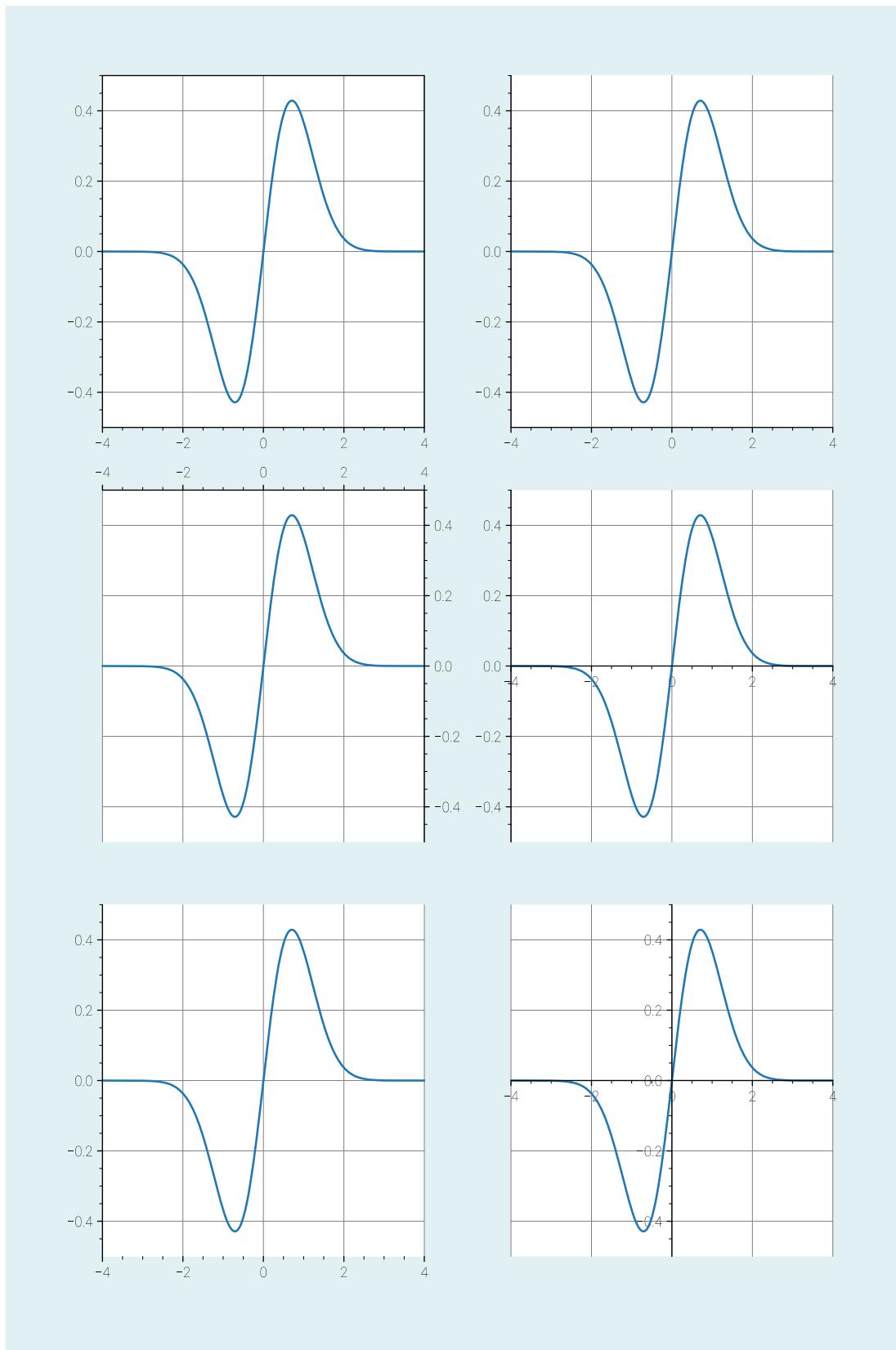


图 3. 修改图脊, 第 1 组

本 PDF 文件为作者草稿, 发布目的为方便读者在移动终端学习, 终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有, 请勿商用, 引用请注明出处。

代码及 PDF 文件下载: <https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger: <https://space.bilibili.com/513194466>

欢迎大家批评指教, 本书专属邮箱: jiang.visualize.ml@gmail.com

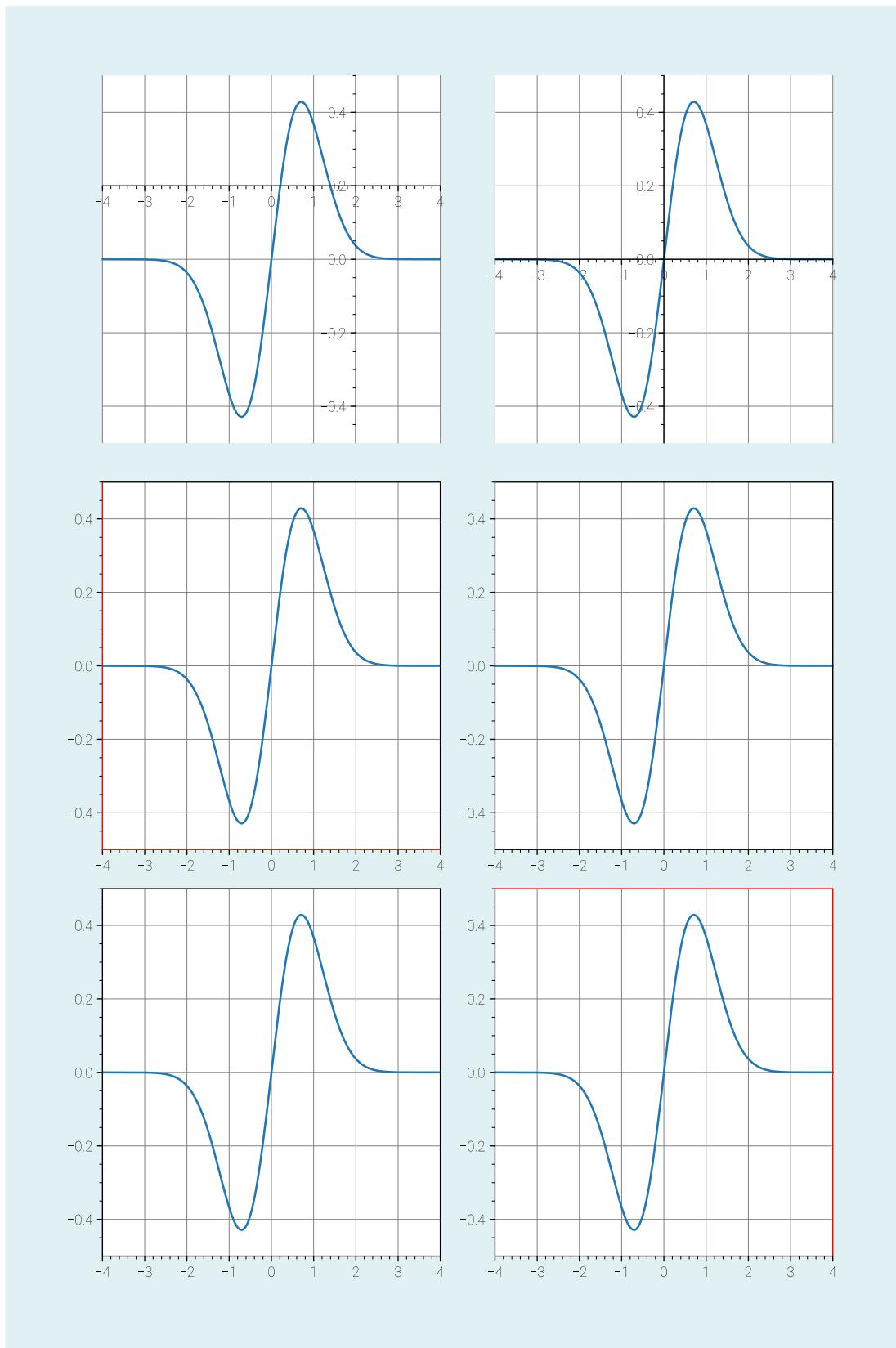


图 4. 修改图脊, 第 2 组

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

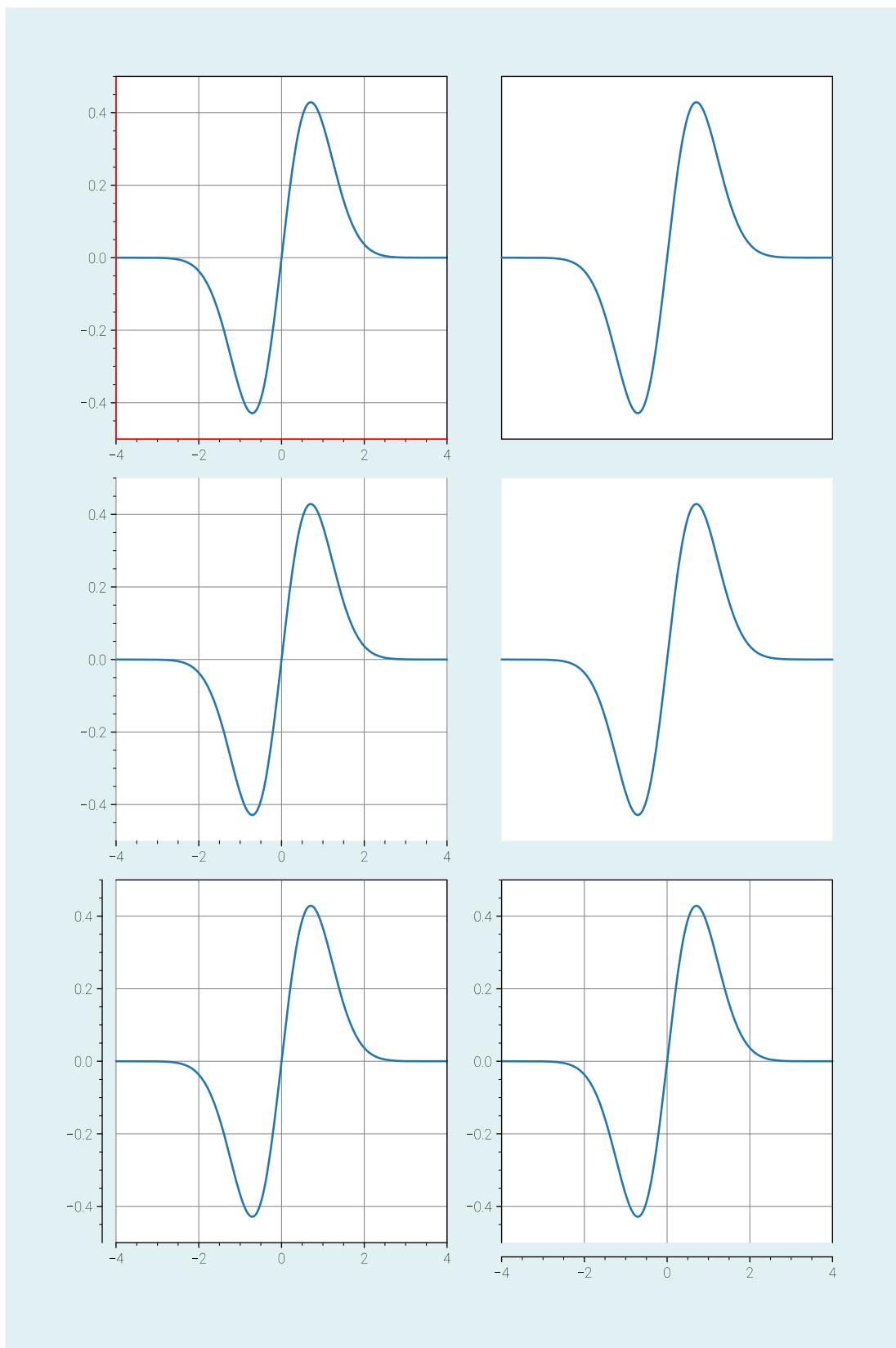


图 5. 修改图脊, 第 3 组

本 PDF 文件为作者草稿, 发布目的为方便读者在移动终端学习, 终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有, 请勿商用, 引用请注明出处。

代码及 PDF 文件下载: <https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger: <https://space.bilibili.com/513194466>

欢迎大家批评指教, 本书专属邮箱: jiang.visualize.ml@gmail.com

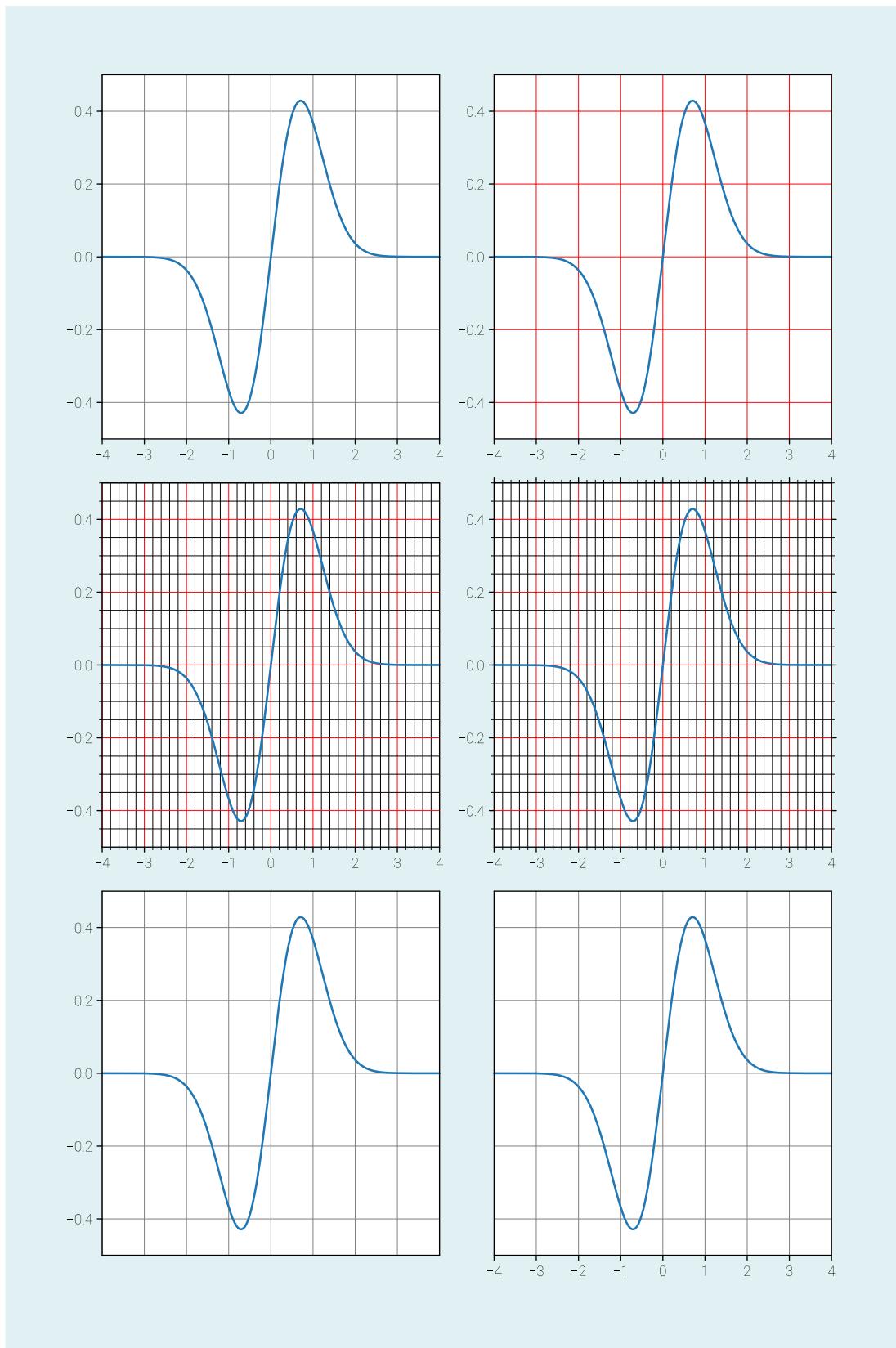


图 6. 背景网格样式

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

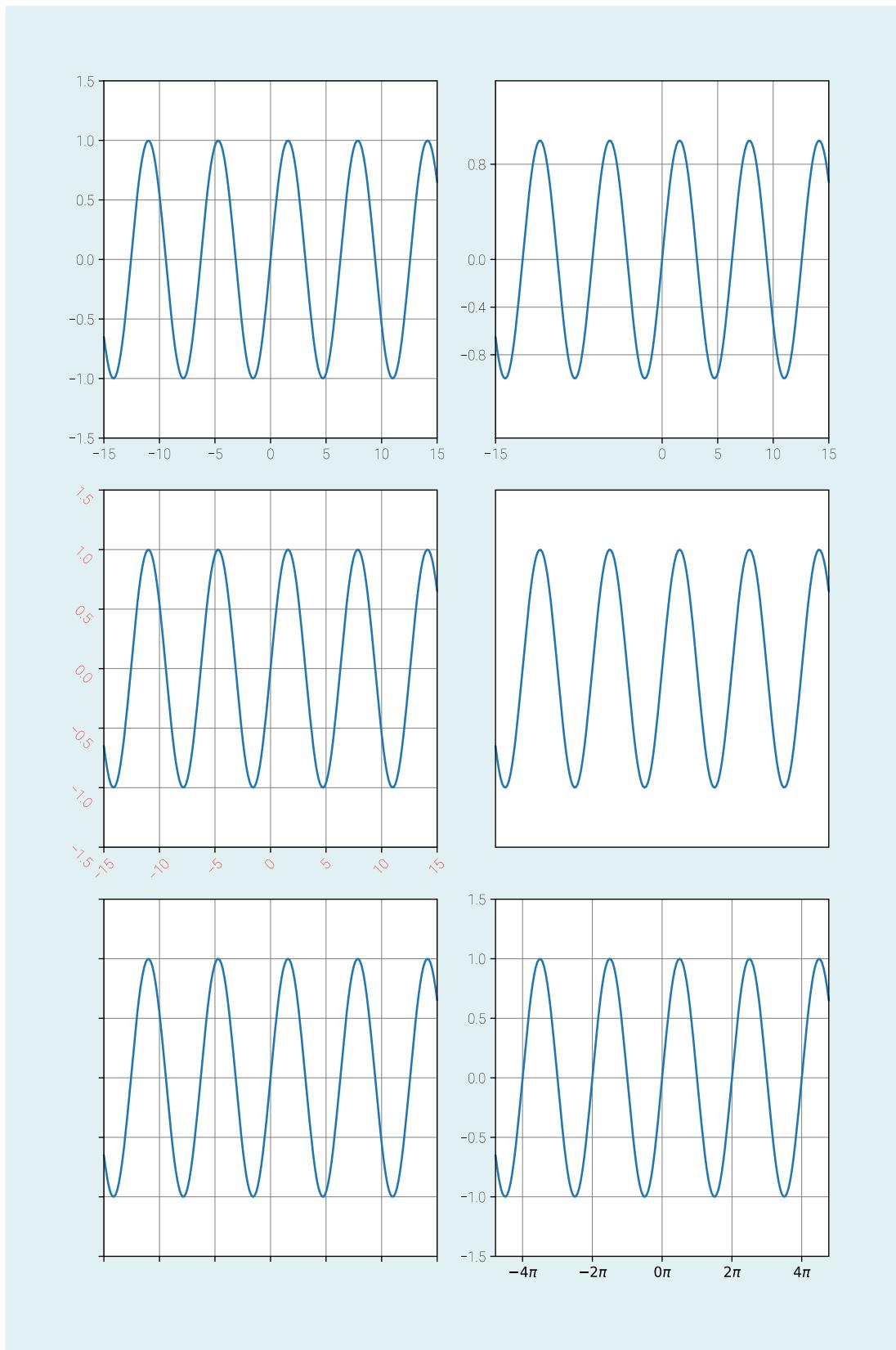


图 7. 修改图轴, 第 1 组

本 PDF 文件为作者草稿, 发布目的为方便读者在移动终端学习, 终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有, 请勿商用, 引用请注明出处。

代码及 PDF 文件下载: <https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger: <https://space.bilibili.com/513194466>

欢迎大家批评指教, 本书专属邮箱: jiang.visualize.ml@gmail.com

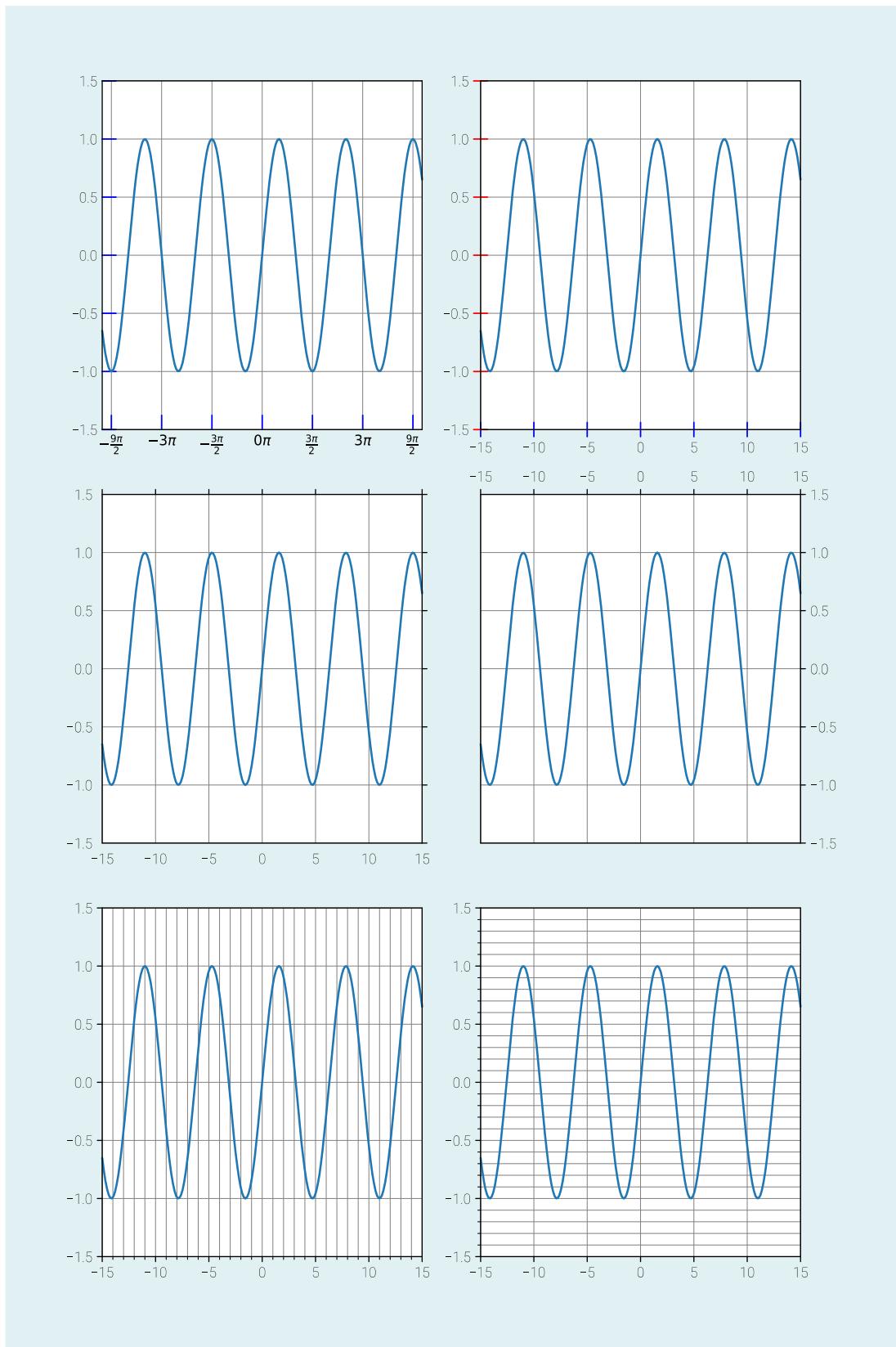


图 8. 修改图轴, 第 2 组

本 PDF 文件为作者草稿, 发布目的为方便读者在移动终端学习, 终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有, 请勿商用, 引用请注明出处。

代码及 PDF 文件下载: <https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger: <https://space.bilibili.com/513194466>
欢迎大家批评指教, 本书专属邮箱: jiang.visualize.ml@gmail.com

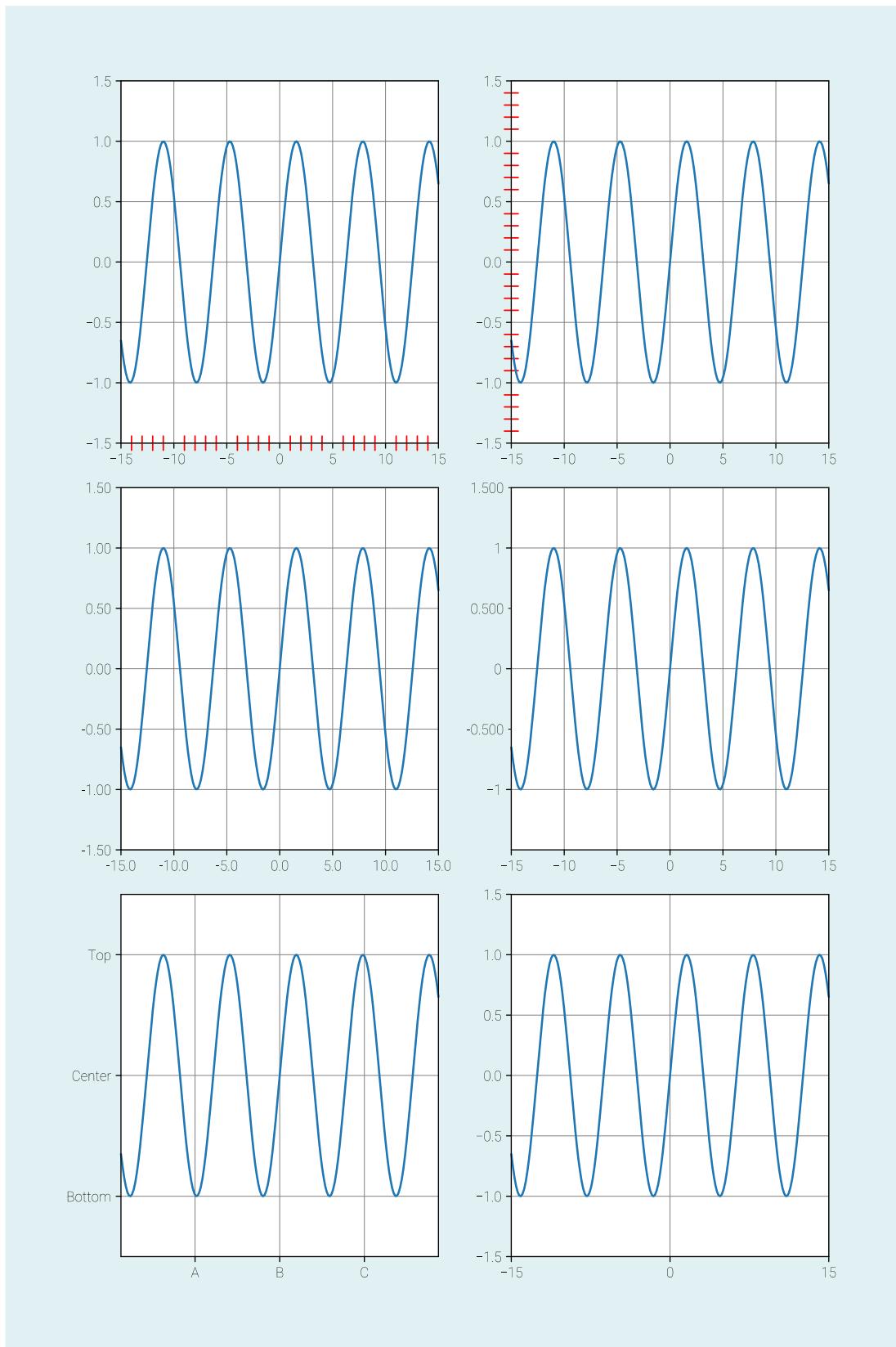


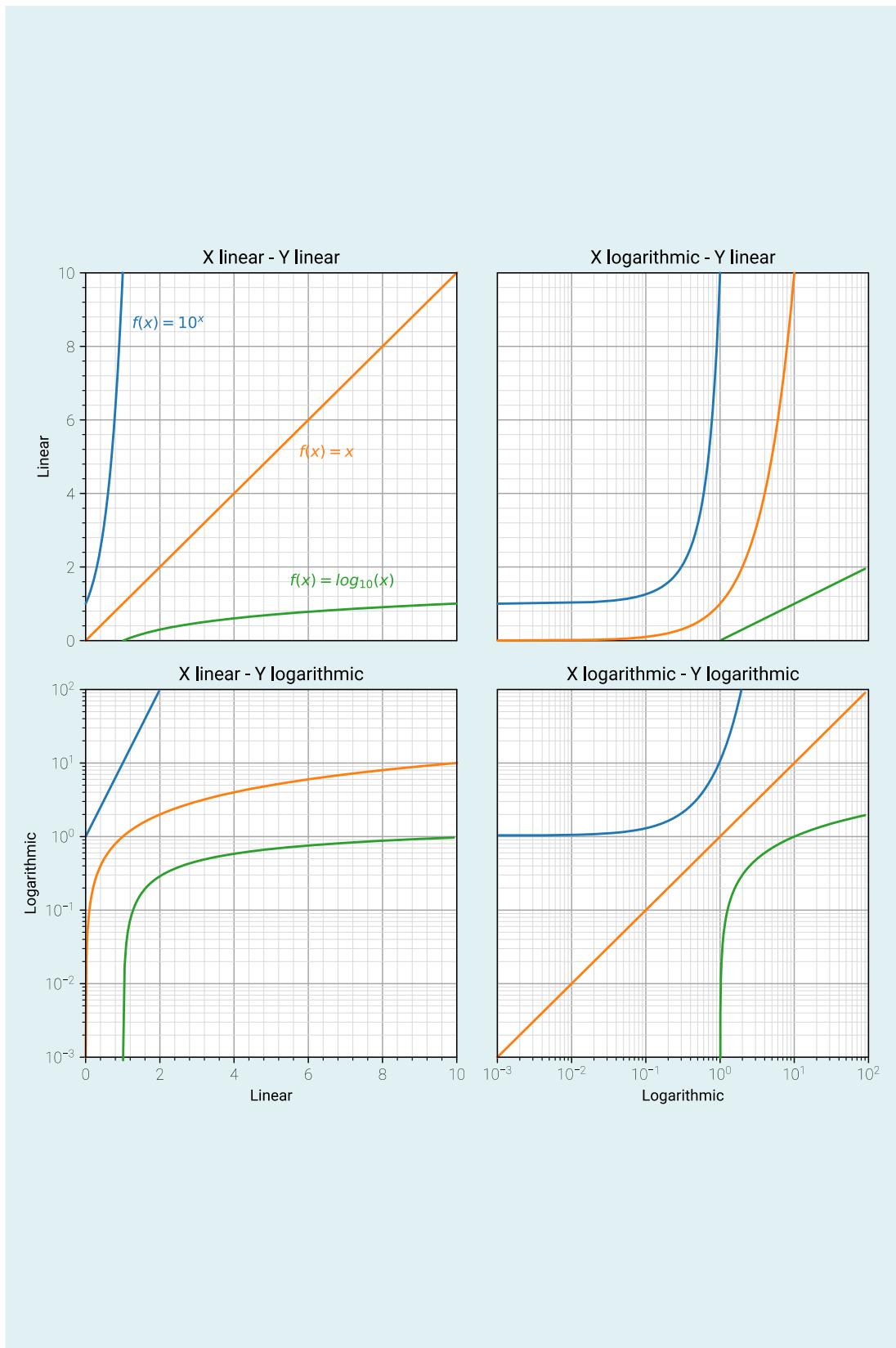
图 9. 修改图轴, 第 3 组

本 PDF 文件为作者草稿, 发布目的为方便读者在移动终端学习, 终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有, 请勿商用, 引用请注明出处。

代码及 PDF 文件下载: <https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger: <https://space.bilibili.com/513194466>

欢迎大家批评指教, 本书专属邮箱: jiang.visualize.ml@gmail.com

图 10. 对数坐标，图片参考 *Scientific Visualization: Python & Matplotlib*

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

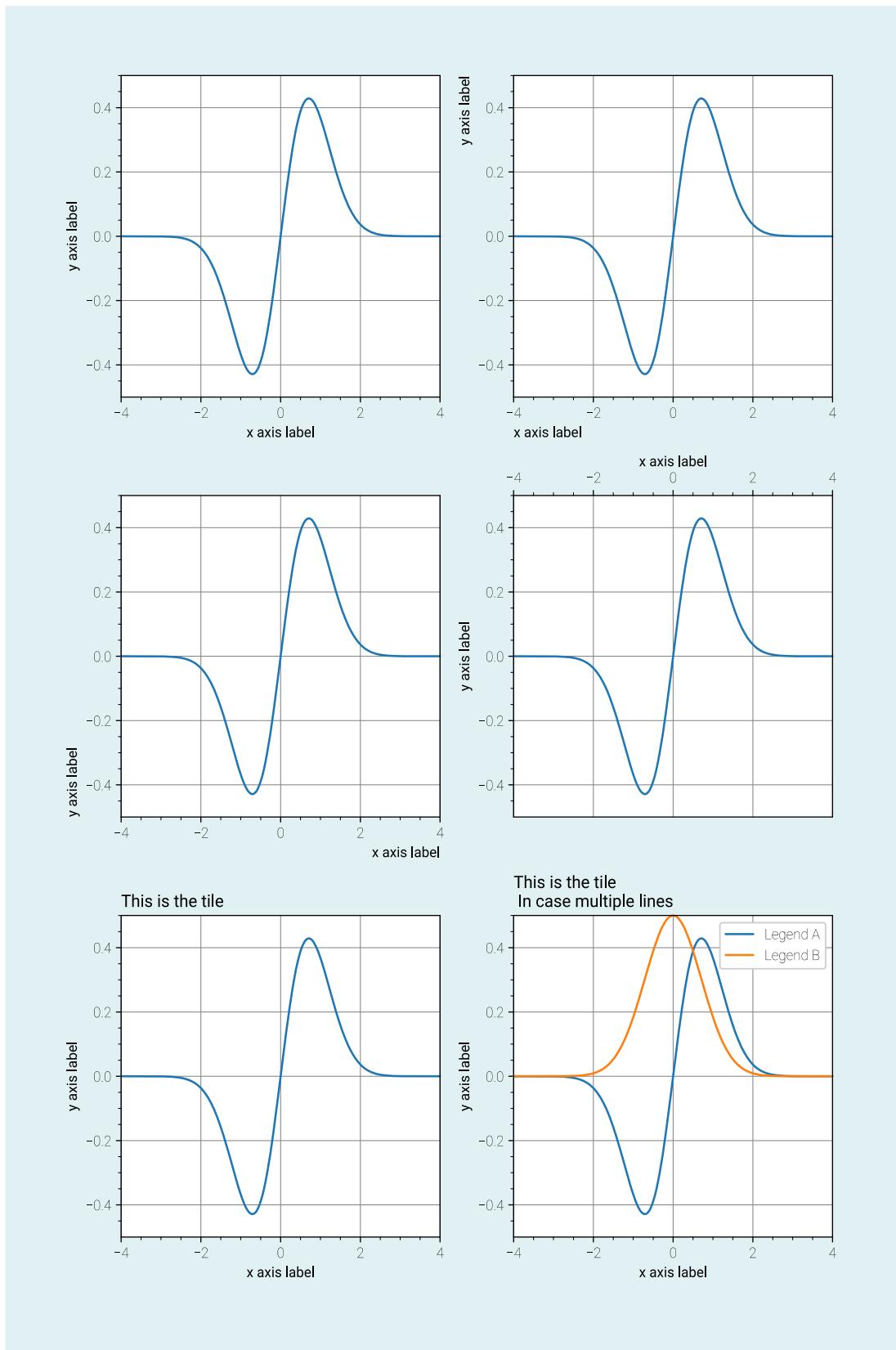


图 11. 标注

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

5 Styles 风格

配色方案、线条、字体、标记符号等等设计元素



像专业人士一样学习规则，这样你就可以像艺术家一样打破它们。

Learn the rules like a pro, so you can break them like an artist.

—— 毕加索 (Pablo Picasso) | 西班牙艺术家 | 1881 ~ 1973



- ◀ XXXXX

5.1 风格

图像的风格指的是图表的整体外观和样式，包括配色方案、线条类型、图脊、图轴、线条宽度、字体、标记符号等等设计元素。

使用过 MATLAB 绘图的读者一定忘不了其严肃，甚至有些呆板的图像样式。而用过 R 语言的 ggplot 的读者，换成 Python 的 Matplotlib 绘图时肯定会有各种视觉上的不适。

Python 中 Plotnine 库的出图风格类似 ggplot，本书不展开介绍。

Matplotlib 提供了一系列的预定义风格，可以通过设置来改变图表的外观。以下是 Matplotlib 中常用的一些图像风格类型：

- ▶ "default"是 Matplotlib 的默认风格，使用蓝色线条和绿色网格。
- ▶ "classic"是一种经典的 Matplotlib 风格，使用黑色线条和白色背景，类似于传统的 Matplotlib 版本。
- ▶ "ggplot"模仿了 R 语言中的 ggplot 库的外观，使用灰色网格和彩色线条。
- ▶ "fivethirtyeight"模仿了流行的数据新闻网站 FiveThirtyEight 的外观，使用红色和蓝色线条，以及灰色网格。
- ▶ "dark_background"使用深色背景和亮色线条，适合用于暗色主题的环境。
- ▶ "seaborn"模仿了 Seaborn 库的外观，使用柔和的颜色和灰色网格。

关于 Matplotlib 绘图风格，请大家参考：

https://matplotlib.org/stable/gallery/style_sheets/style_sheets_reference.html

本章着重展示 ProPlot 提供的可视化方案，因为这个绘图库的出图风格特别“像”科技三大刊——*Cell*、*Nature* 和 *Science*。

5.2 ProPlot：出图风格专业的绘图库

ProPlot 建立在 Matplotlib 之上，提供了更简洁、更便于使用的科学绘图包。ProPlot 支持各种常见的绘图类型，包括线图、散点图、等高线图、柱状图等，并且支持创建多个子图和面板图。

ProPlot 支持高分辨率的输出，可以生成矢量图形（如 PDF、SVG）和栅格图像（如 PNG、JPEG）等多种格式。

图 1~ 图 14 给出的是 ProPlot 官网提供的范例。ProPlot 和 Matplotlib 语法高度一致，因此本书不单独展开讲解。大家需要在科技期刊发表文章的话，可以学习使用 ProPlot。ProPlot 官网地址如下：

<https://proplot.readthedocs.io/en/stable/>

需要注意的是，ProPlot 还在开发中，使用时可能会报错。

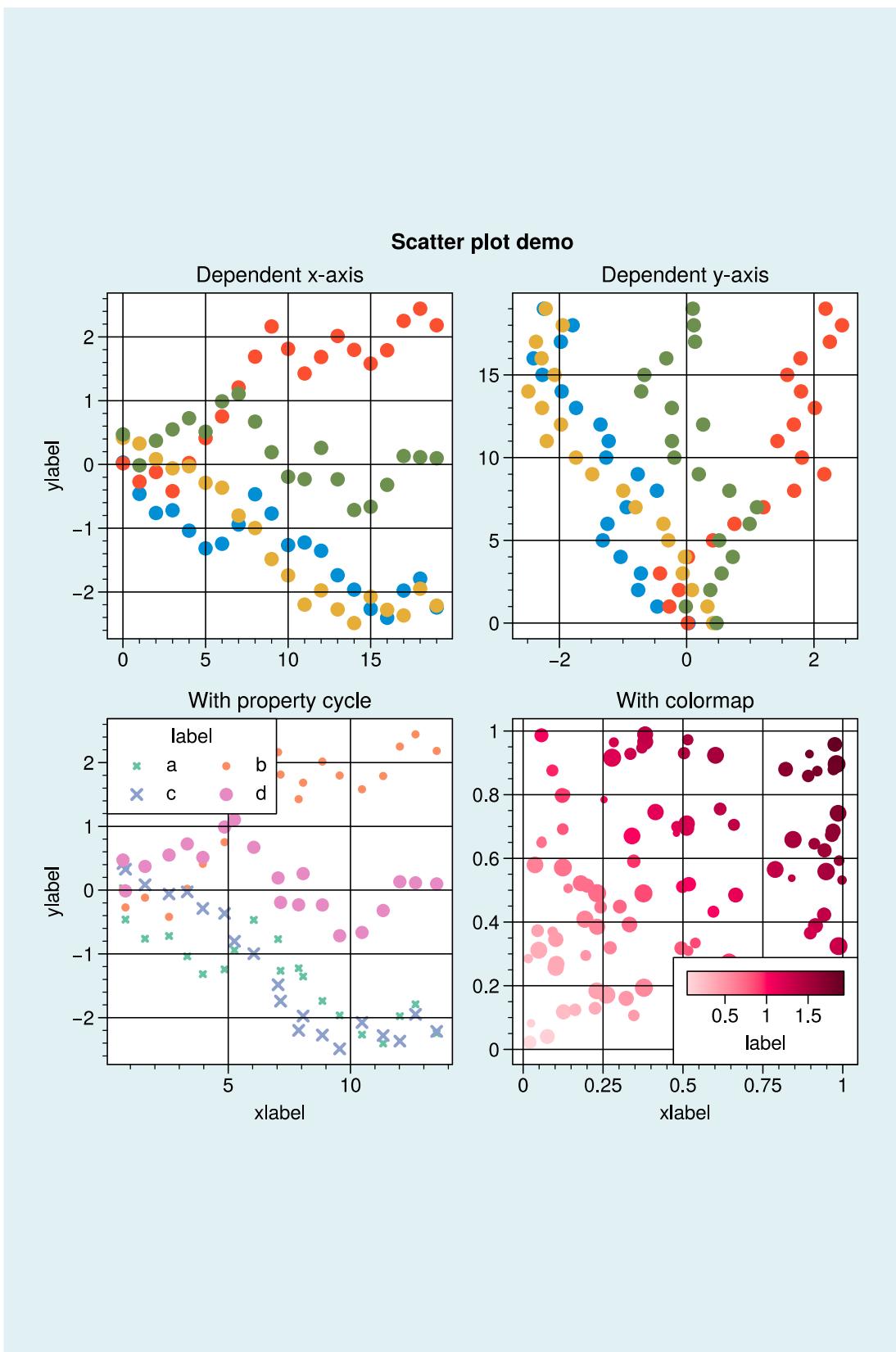


图 1. 散点图

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

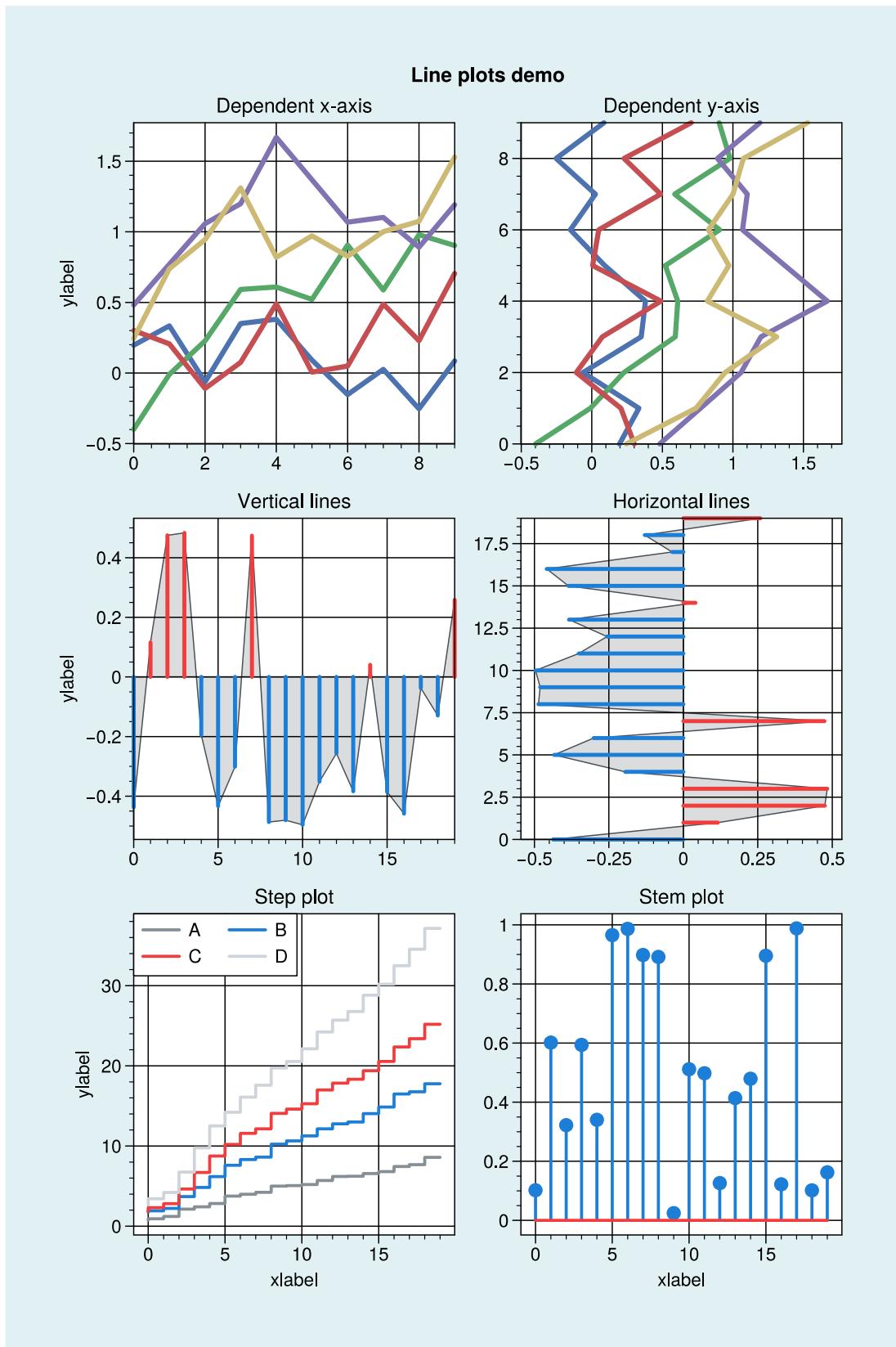


图 2. 线图

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

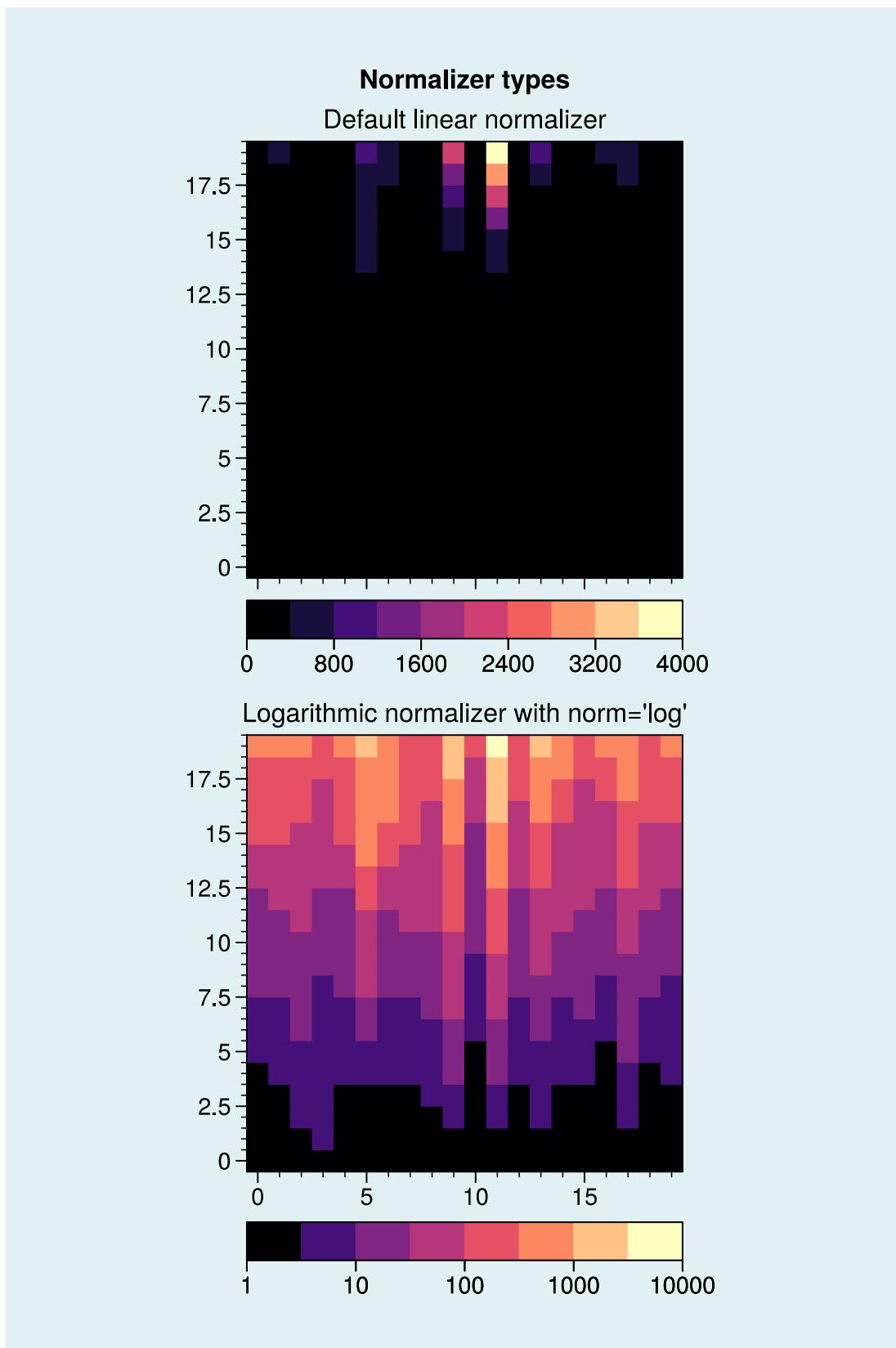


图 3. 热图

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

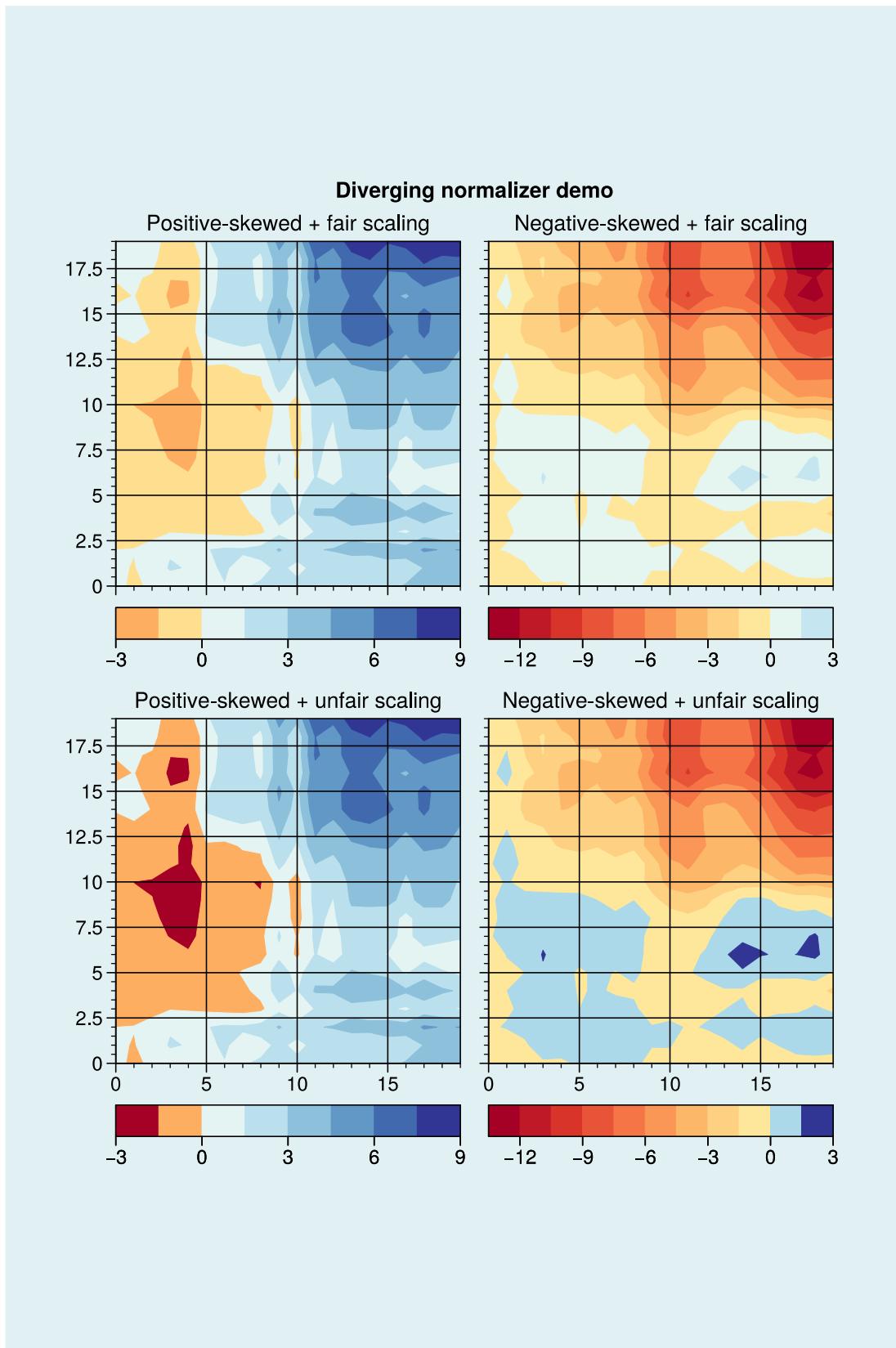


图 4. 等高线图

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

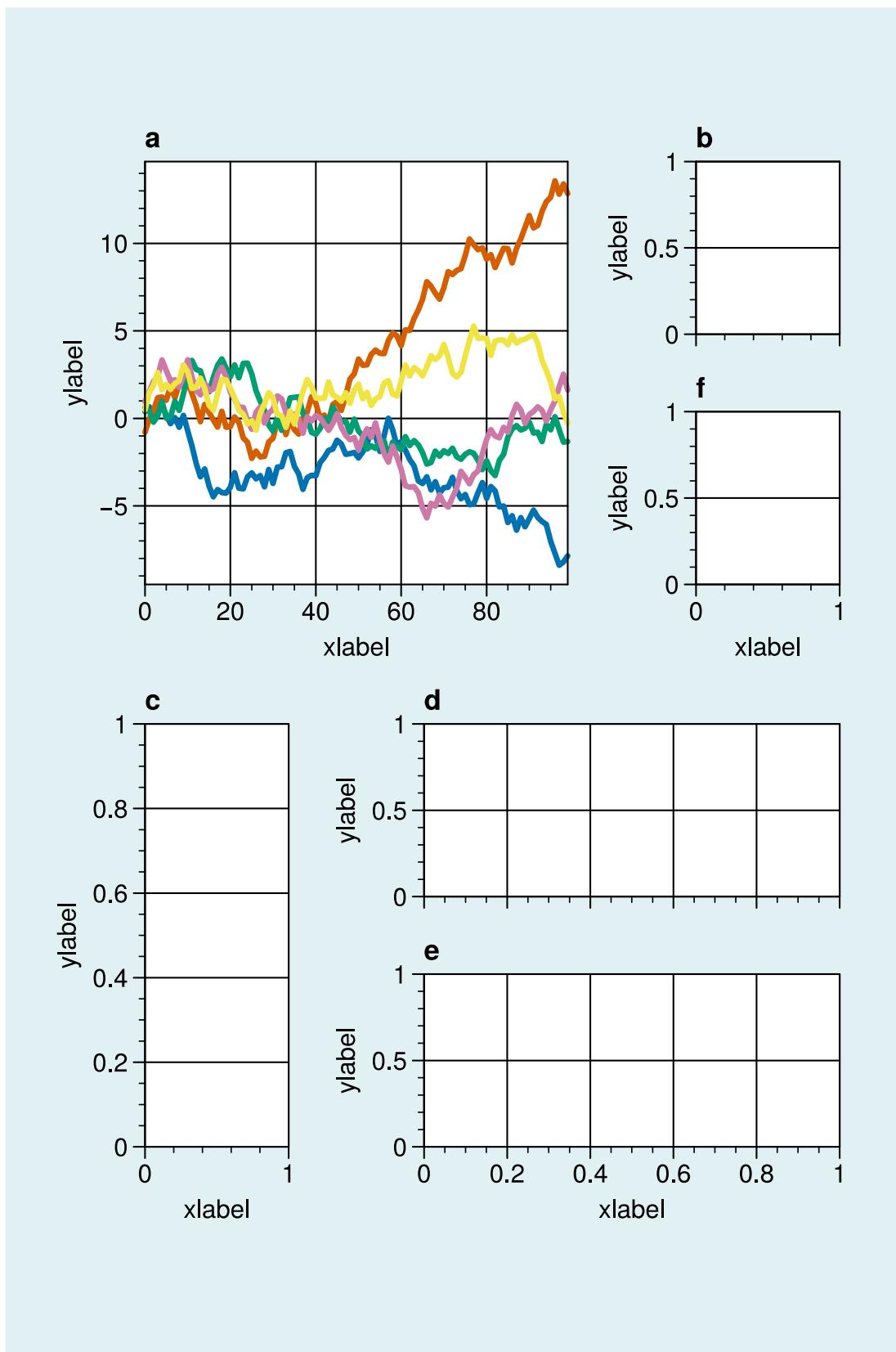


图 5. 复杂子图布置

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

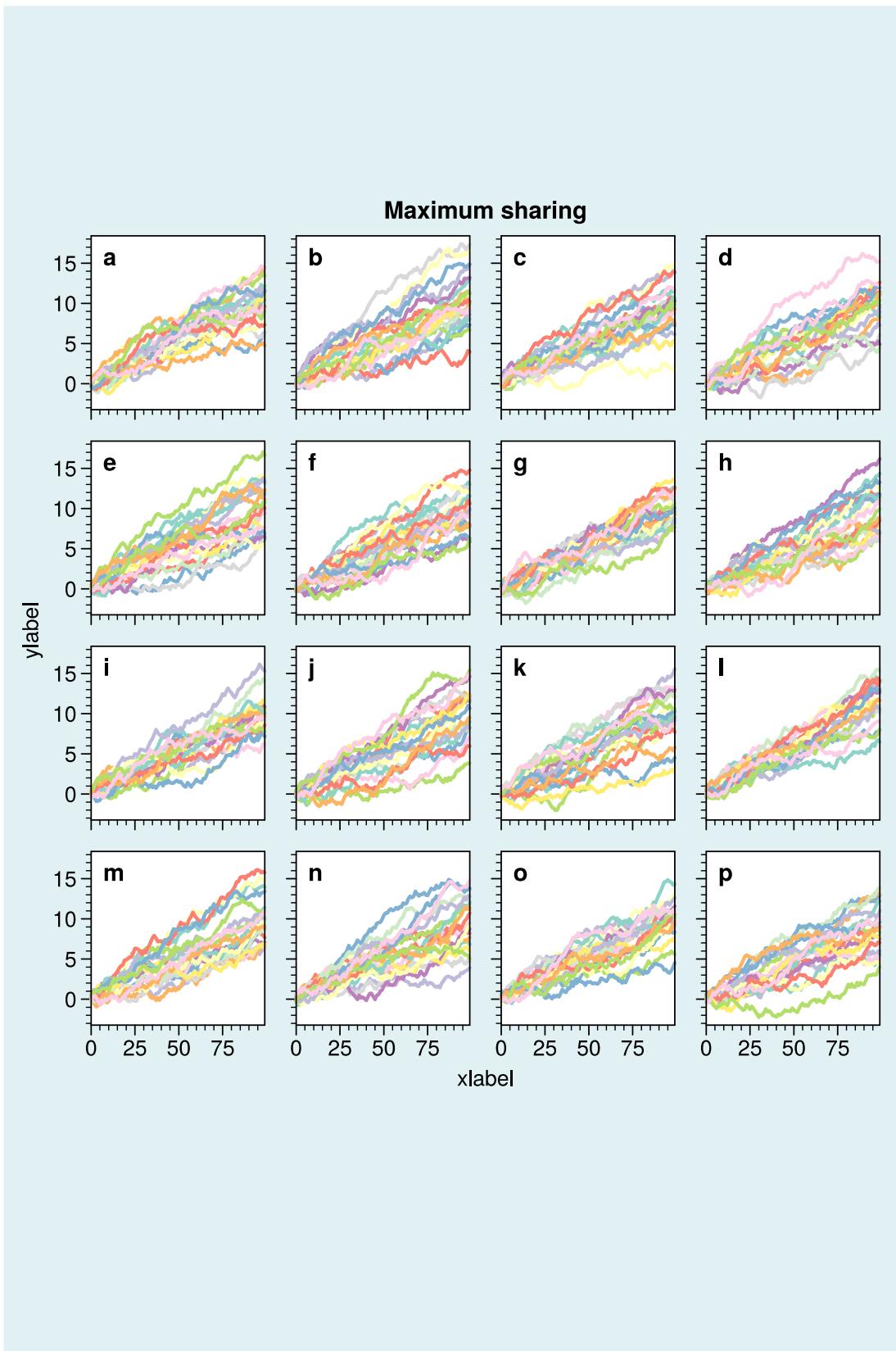


图 6. 共享横纵轴刻度

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

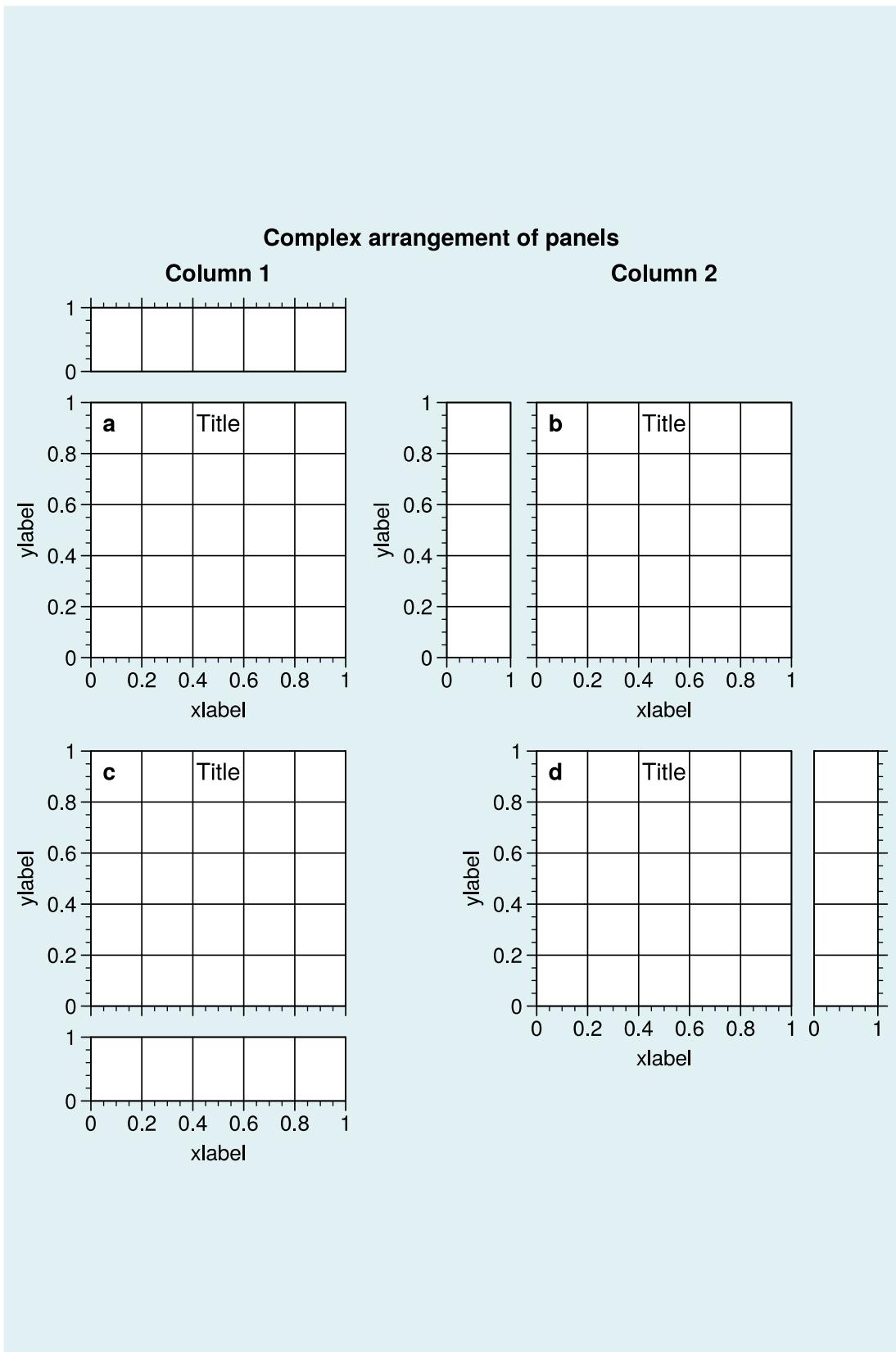


图 7. 嵌套子图布置

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

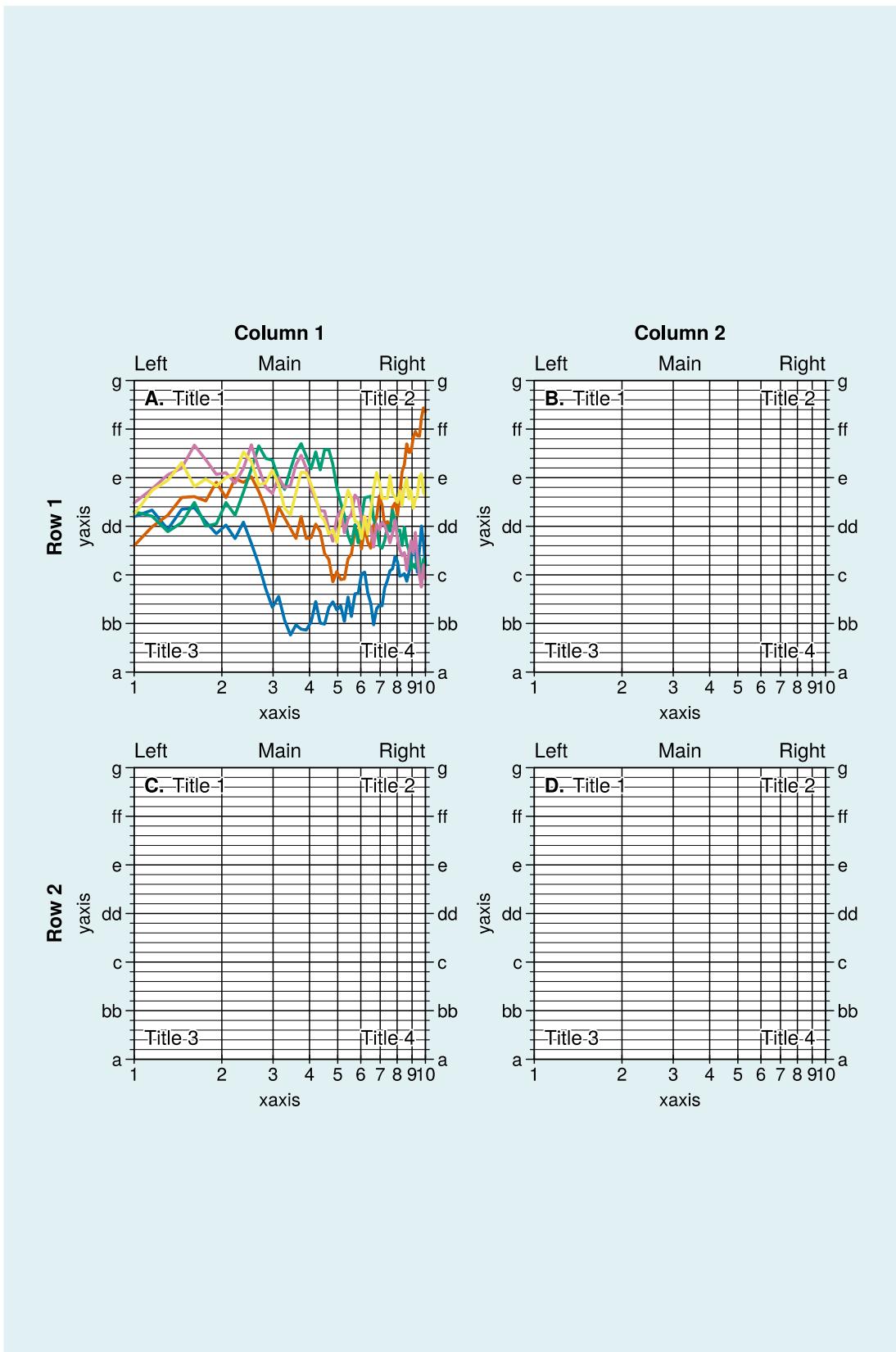


图 8. 对数坐标

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

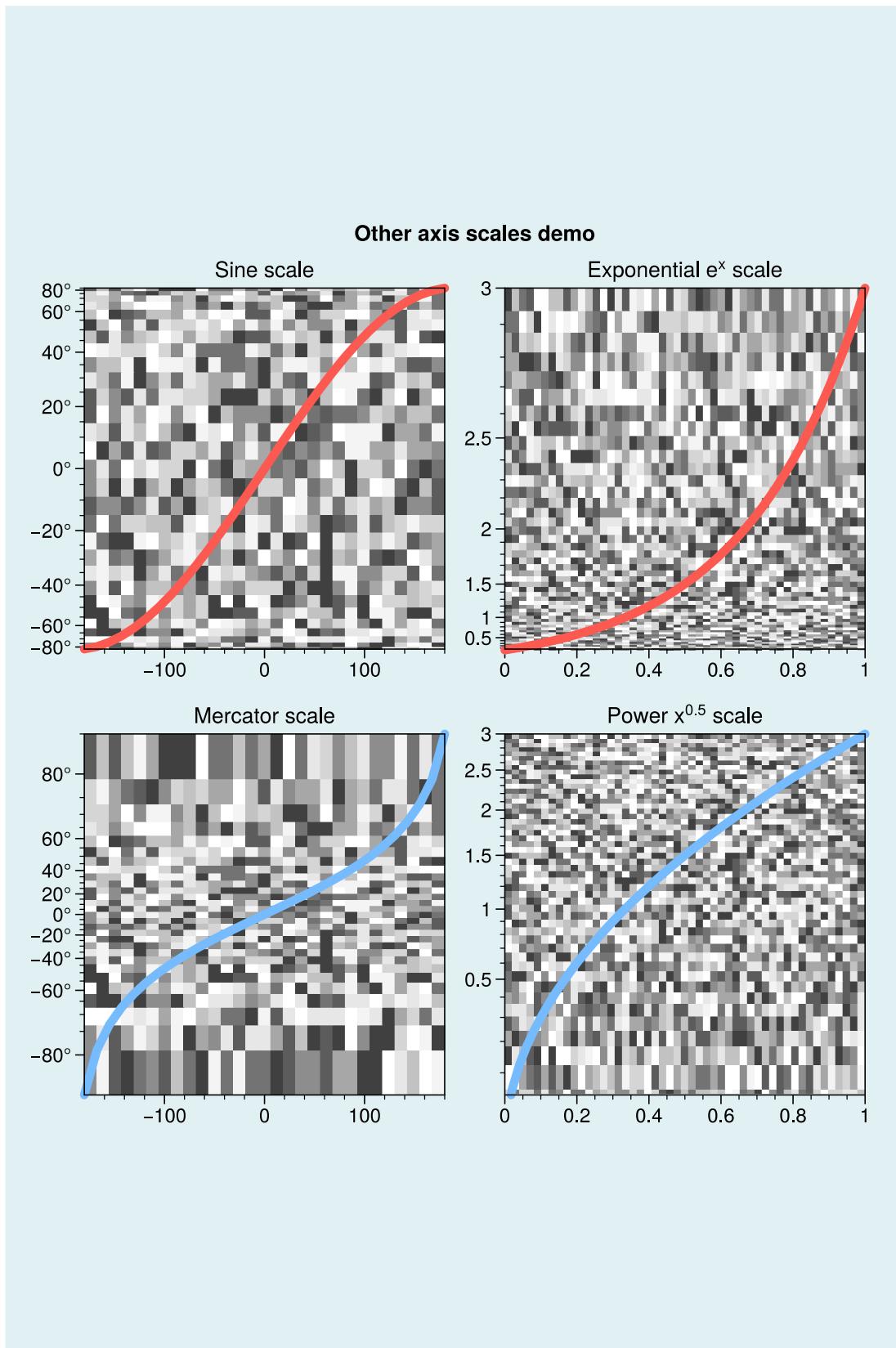


图 9. 其他非线性坐标轴

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

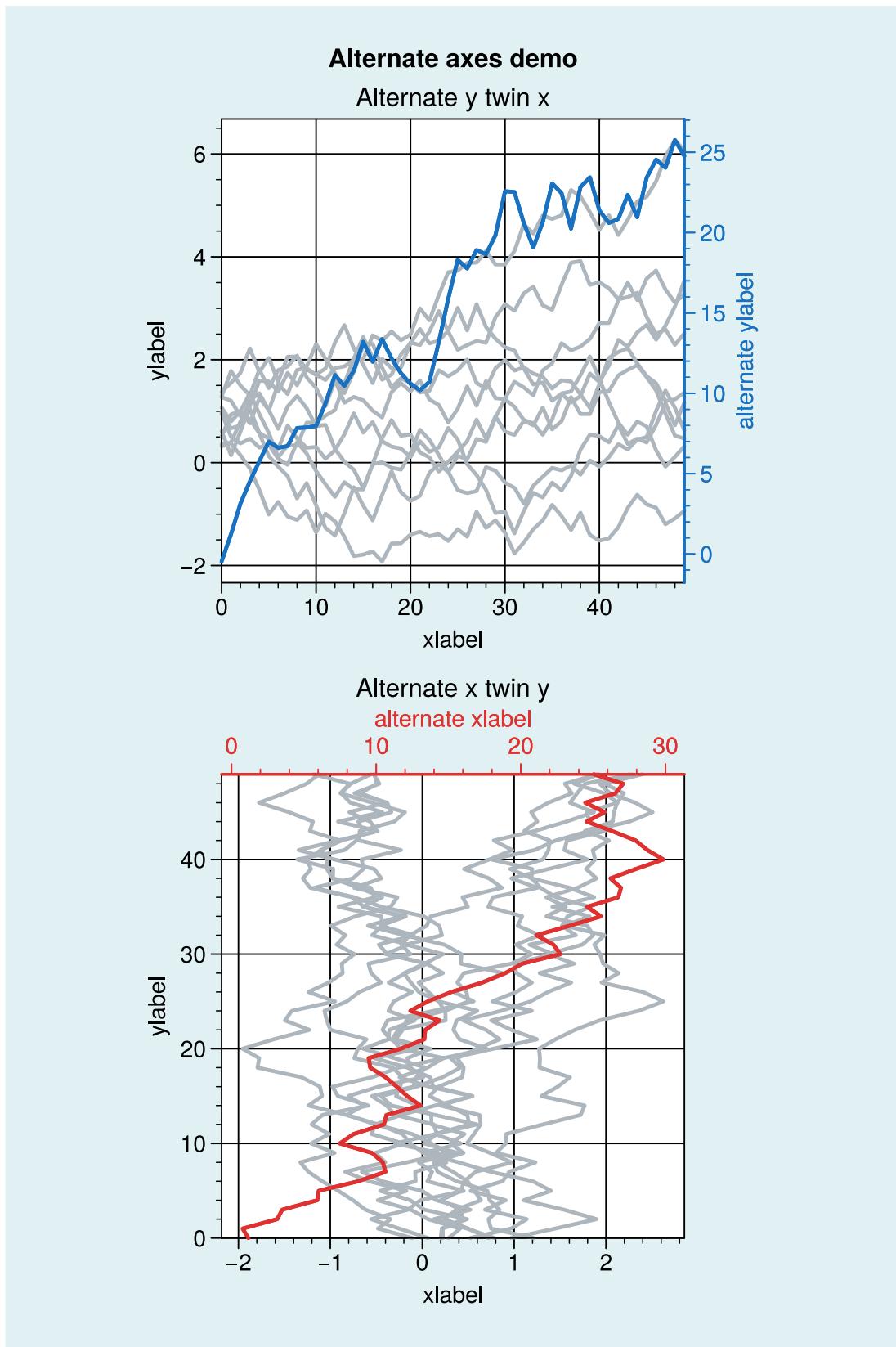


图 10. 双纵轴、双横轴

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

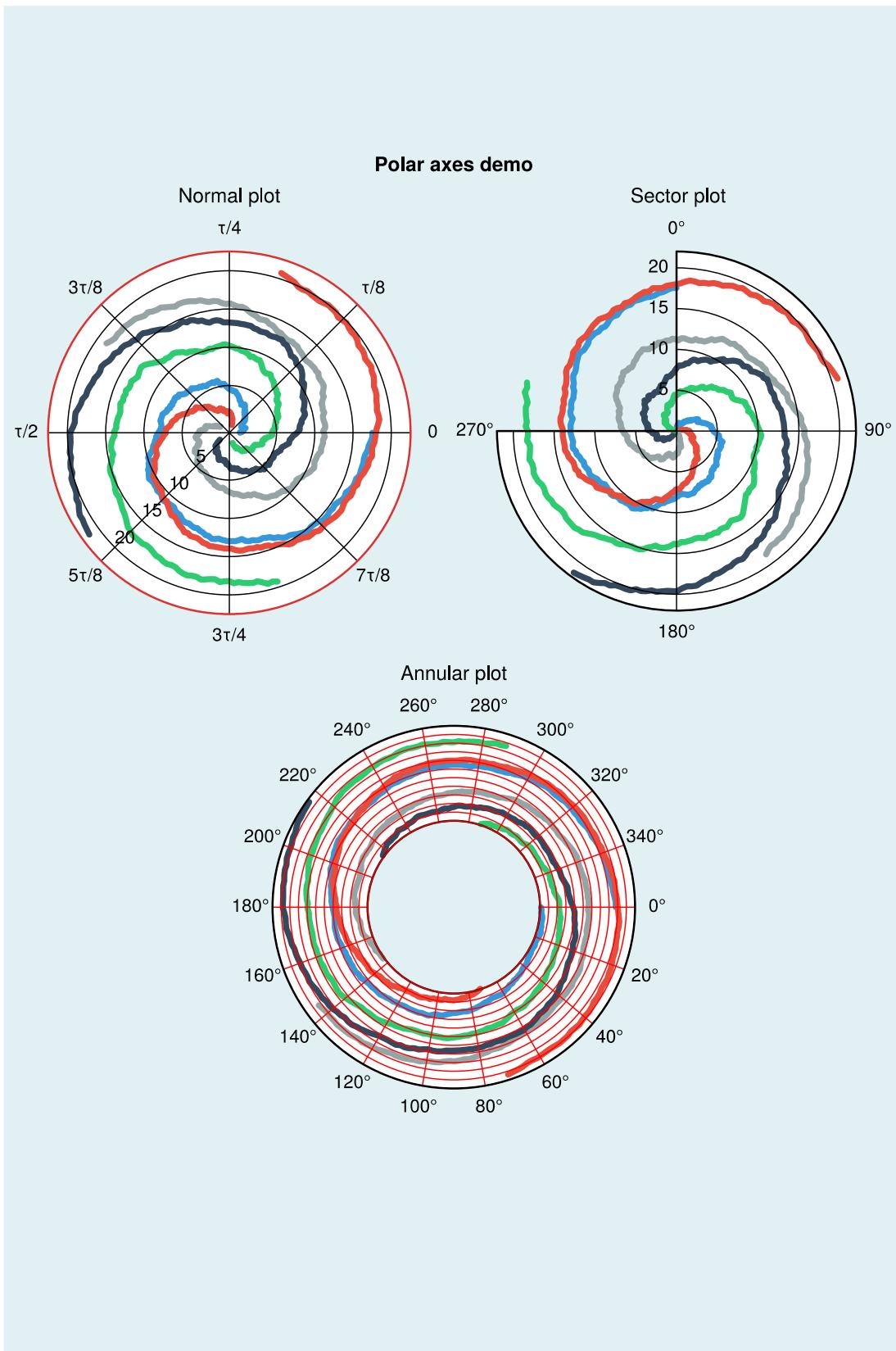


图 11. 极坐标

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

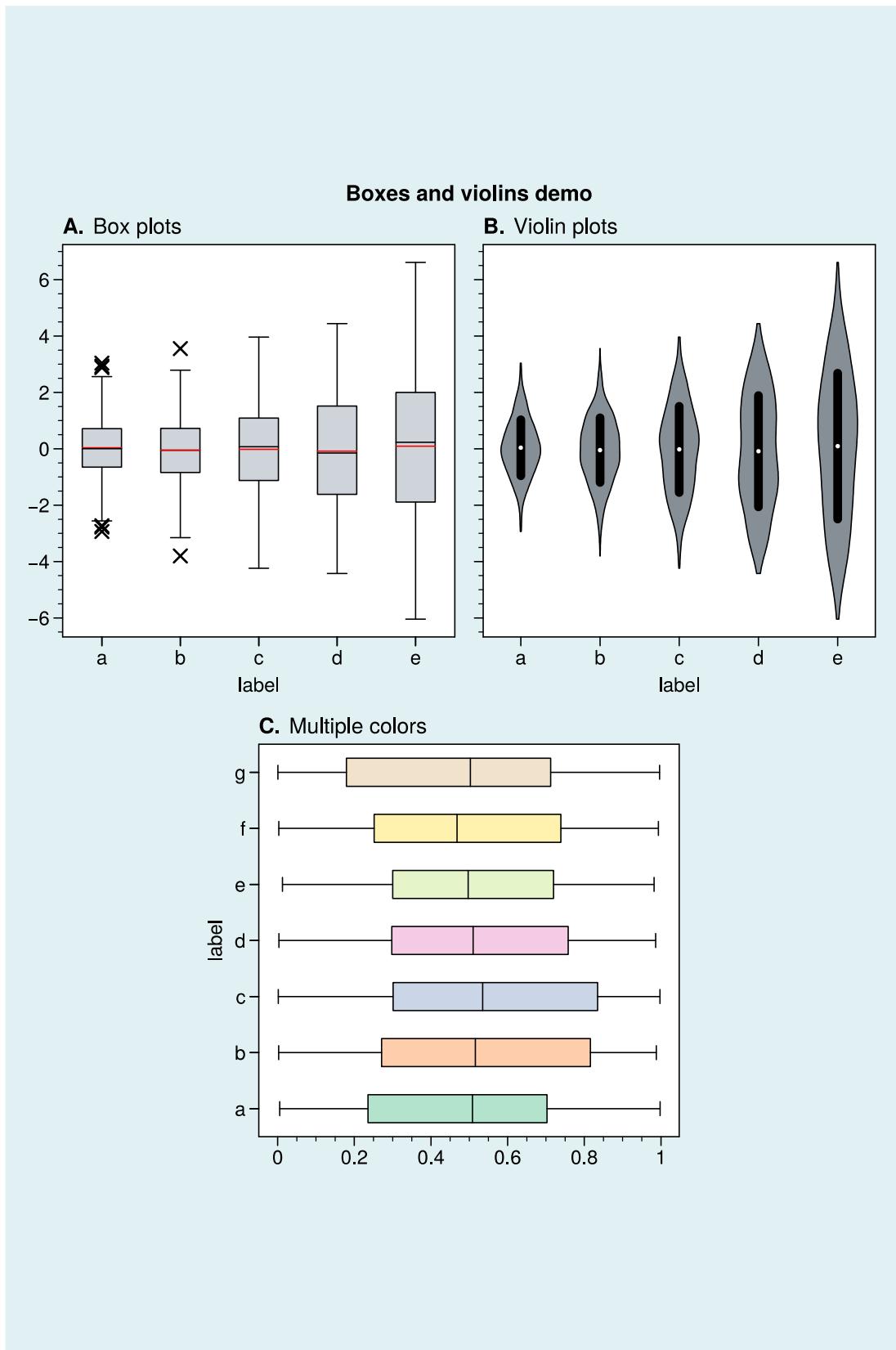


图 12. 统计描述

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

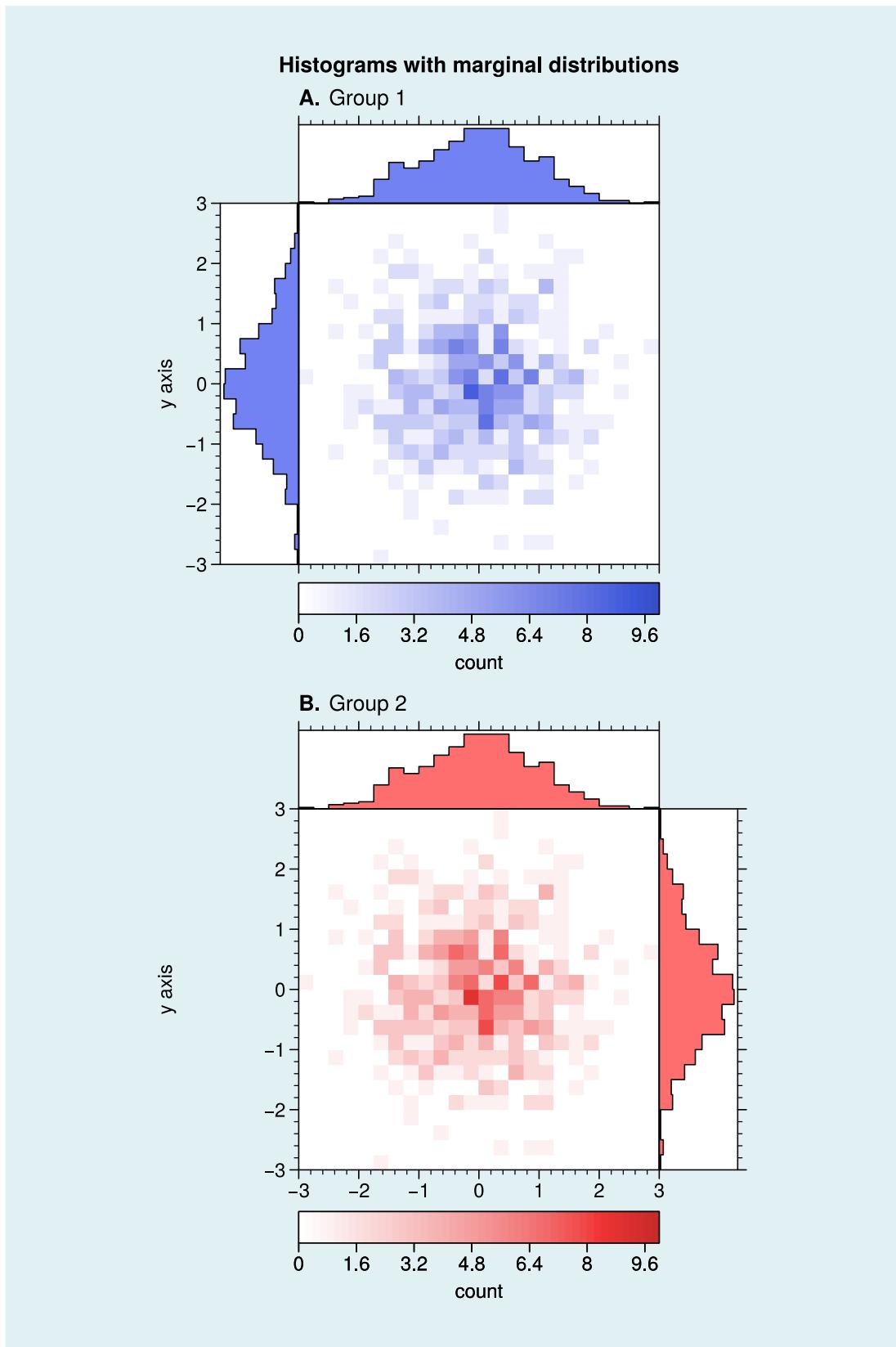


图 13.联合分布、边缘分布

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

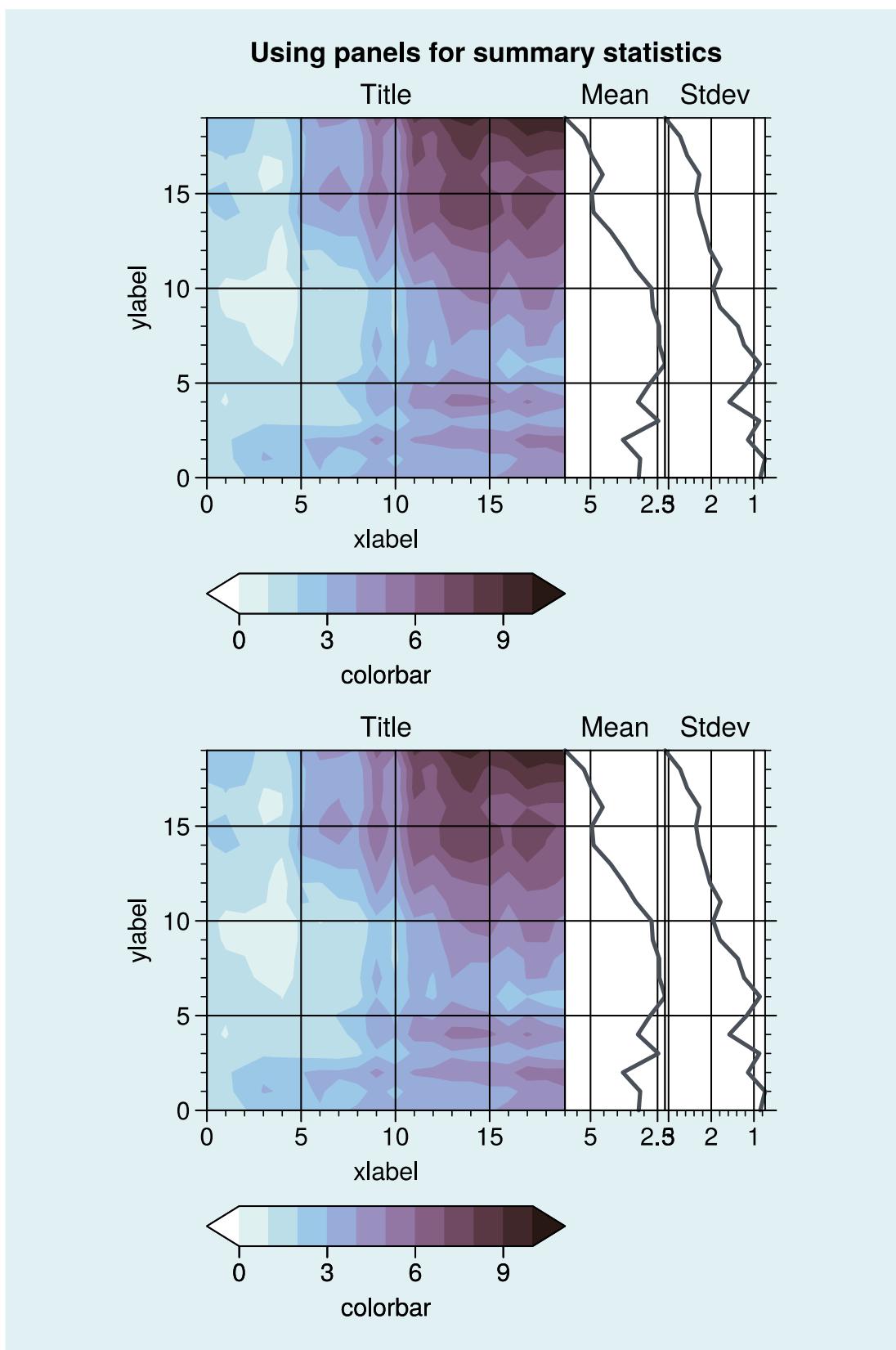


图 14. 边缘统计

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

7 Colors 颜色

Matplotlib 中如何设置颜色



越是垂垂老矣、病痛缠身、捉襟见肘，我越想创造热情洋溢、井然有序、光彩照人的作品还以颜色。

The more ugly, old, nasty, ill, and poor I become the more I want to get my own back by producing vibrant, well-arranged, radiant colour.

—— 文森特·梵高 (Vincent van Gogh) | 荷兰后印象派画家 | 1853 ~ 1890



- ◀ XXXXX

7.1 色彩

色彩是可见光在物体表面反射、折射或透射时产生的感知效果。人眼感知色彩的过程涉及到视觉系统的不同组成部分。

人眼

人眼感知色彩的基本机制是通过视觉感受器官——眼睛的视网膜上的视锥细胞来完成的。视锥细胞包括三种类型：红色感受器、绿色感受器和蓝色感受器，它们分别对应于不同波长的光。当光线进入眼睛并刺激视网膜上的视锥细胞时，它们会产生相应的神经信号，传递到大脑的视觉皮层。大脑对这些信号进行解析和处理，最终形成我们对色彩的感知。

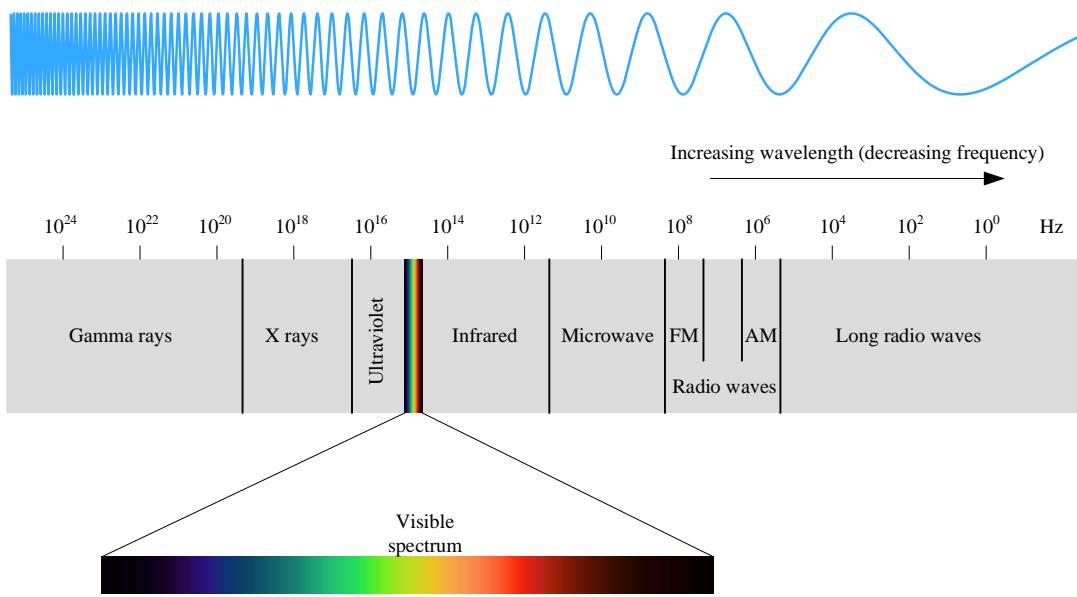


图 1. 光谱

如图 1 所示，可见光只是光谱中一小部分。光谱是指将电磁波按照频率或波长进行分类的方式。从高能量到低能量，光谱的构成包括以下几个部分：

伽马射线是电磁波谱中能量最高的部分，具有极短的波长和高频率。它们常常与核反应、天体物理事件以及放射治疗等相关。

X 射线具有比紫外线更高的能量，波长短，频率高。X 射线在医学成像、材料检测和科学研等应用领域有广泛应用。

紫外线波长比可见光短，能量较高。

可见光是人眼能够感知的电磁波，波长较长。可见光谱从紫色、蓝色、绿色、黄色、橙色到红色。

红外线波长比可见光长，能量较低。红外线在夜视设备、红外热像仪和通信技术等方面有广泛应用。

广播电波具有非常长的波长和低能量，适合用于无线通信和广播传输。

长波指波长非常长的电磁波，如无线电波等，常用于无线通信和远程传输。

色彩心理学是研究色彩对人类情感和行为产生影响的学科。其中，冷暖色调是色彩的一种分类方式。冷色调如蓝色、绿色、紫色等给人以凉爽、安静、宁静的感觉，常用于创造宽松和放松的氛围。暖色调如红色、黄色、橙色等则给人以温暖、充满活力的感觉，常用于刺激和激发人们的情感和能量。

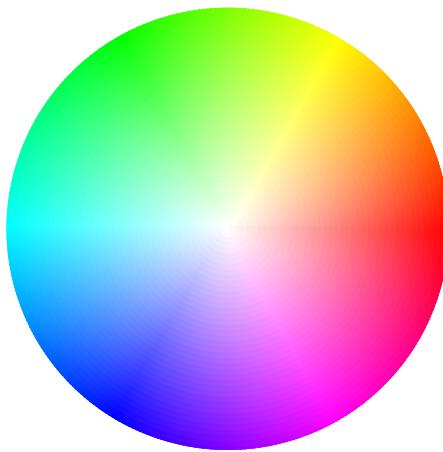


图 2. 冷暖色调色盘

色彩模式

常见的色彩模式有以下几种：

RGB 模式：这是用于数字图像和显示器的最常见的色彩模式。它通过将红、绿、蓝三原色以不同的强度混合来创建各种色彩。本书后续将专门介绍 RGB 色彩模式。

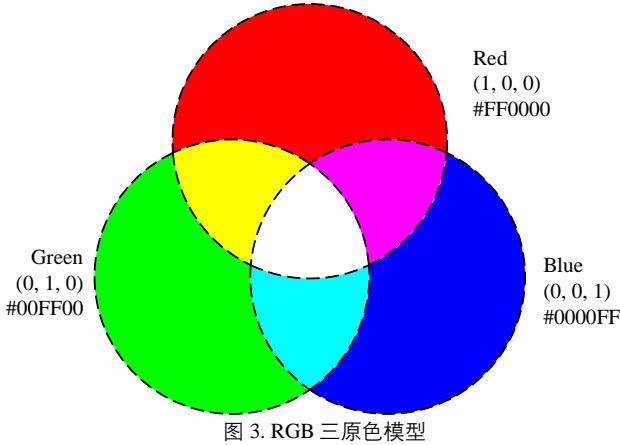
CMYK 模式：这是印刷行业常用的色彩模式。它使用青、洋红、黄和黑四种颜色的墨水混合来创建各种色彩。

HSL/HSV 模式 (色调、饱和度、亮度/明度)：这是一种基于人类感知的色彩模式。色调表示色彩的位置，饱和度表示色彩的纯度，亮度/明度表示色彩的亮暗程度。

7.2 RGB 颜色模型

如图 3 所示，三原色光颜色模型 (RGB color model) 将红 (red)、绿 (green)、蓝 (blue) 色光以不同比例混合得到不同的颜色。Matplotlib 大家也会见到 RGBA，A 代表透明度 alpha。

本书后文大家会看到 RGB 颜色模型的色彩空间。



Matplotlib 中定义颜色

Matplotlib 中，RGB 颜色可以用数组(色号)、十六进制(hexadecimal 简称 hex)、名称等来表示。

比如，绘制一条纯蓝色线，可以设置：

```
color = (0, 0, 1) # 数组
# (red, green, blue)
color = (0, 0, 1, 0.5)
# (red, green, blue, alpha)
# alpha 的含义是透明度
color = [0, 0, 1] # list
color = '#0000FF' # 十六进制 Hex
color = 'blue' # 颜色名称
color = 'b' # 颜色名称简写
```

注意，有些软件中颜色采用三个 0 ~ 255 的数值，比如 MS Word。

常用颜色

表 1 所示为常用颜色的设置方式。

表 1. 常用颜色

	数组	简称	全称	Hex
	[1, 0, 0]	'r'	'red'	'#FF0000'
	[0, 1, 0]	'g'	'green'	'#00FF00'
	[0, 0, 1]	'b'	'blue'	'#0000FF'
	[1, 1, 0]	'y'	'yellow'	'#FFFF00'
	[1, 0, 1]	'm'	'magenta'	'#FF00FF'
	[0, 1, 1]	'c'	'cyan'	'#00FFFF'
	[0, 0, 0]	'k'	'black'	'#000000'
	[1, 1, 1]	'w'	'white'	'#FFFFFF'

注意，Matplotlib 中，'g'和'green'几乎相同，'g'的 RGB 色号为 [0.0, 0.5, 0.0]，'green'的 Hex 色号为 #008000；'y'的 RGB 色号实际上是 [0.75, 0.75, 0.0]，而'yellow' 的色号为 [1.0, 1.0, 0.0]；'m' 的色号为 [0.75, 0.0, 0.75]，而'magenta'的色号为 [1.0, 0.0, 1.0]；'c' 的色号为 [0.0, 0.75, 0.75]，而'cyan'的色号为 [0.0, 1.0, 1.0]。

图 7 所示为 Matplotlib 中定义的颜色名称。图 7 参考如下代码，请大家自行学习：

https://matplotlib.org/stable/gallery/color/named_colors.html

此外，我们还可以用 [0, 1] 之间的数值定义不同深浅的灰色。如图 4 所示，color = '0' 代表纯黑，color = '1' 代表纯白，color = '0.5' 代表 50% 灰。注意，必须使用引号，否则会报错。



图 4. 灰度

图 8 所示为几组渐变色和对应的十六进制值。图 8 参考 *Scientific Visualization: Python + Matplotlib*。下载地址为：

<https://github.com/rougier/scientific-visualization-book>

7.3 CMYK 颜色模型

图 5 所示为 RGB 中任意两个颜色混合得到的三种颜色：青色 (cyan)、品红 (magenta)、黄色 (yellow)。这便是 CMYK 色彩模型的基础。K 代表黑色 (black)。

$$\begin{aligned} r + g &= y \quad \text{Yellow (1, 1, 0) \#FFFF00} \\ r + b &= m \quad \text{Magenta (1, 0, 1) \#FF00FF} \\ g + b &= c \quad \text{Cyan (0, 1, 1) \#00FFFF} \end{aligned}$$

图 5. RGB 中两个颜色混合

如图 6 所示，CMYK 调色盘中，红、绿、蓝三色颜料均匀调色得到黑色。CMYK 一般用在印刷领域，本书不展开讲解。

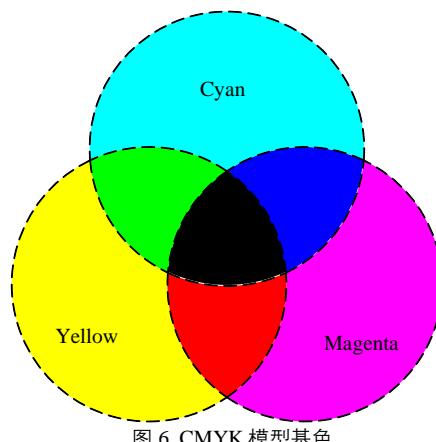


图 6. CMYK 模型基色



图 7. Matplotlib 已定义名称的颜色

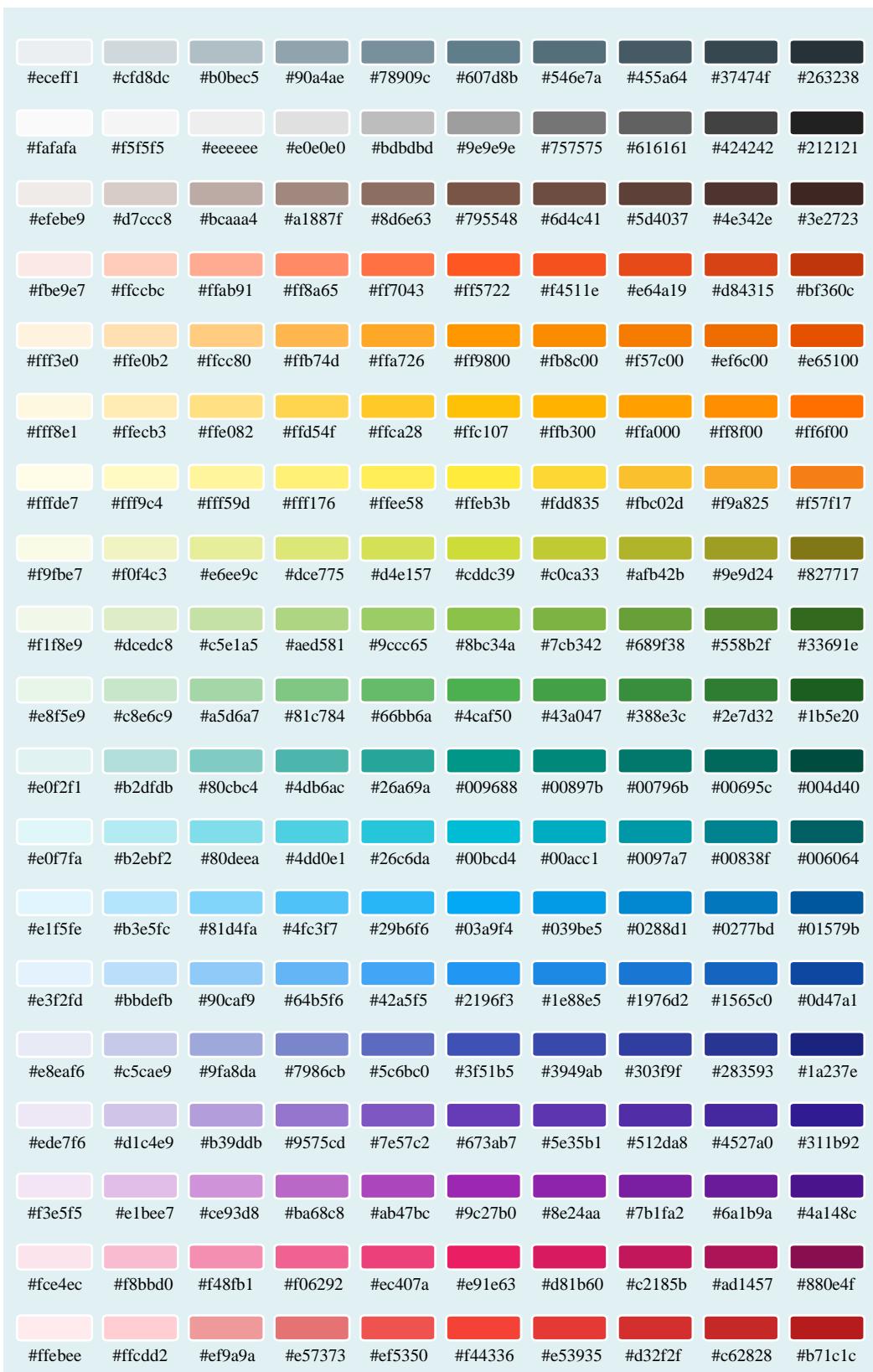


图 8. 几组渐变色和它们的 hex 值

Colorful Spaces

8 色彩空间

红光、绿光、蓝光，色调、饱和度、明暗度



绘画是可以眼见、不可言传的诗歌，而诗歌是可被言传、不可眼见的绘画。

Painting is poetry that is seen rather than felt, and poetry is painting that is felt rather than seen.

——列奥纳多·达·芬奇 (Leonardo da Vinci) | 文艺复兴三杰之一 | 1452 ~ 1519



- ◀ `colorsys.hsv_to_rgb()` 将 HSV (色相、饱和度、亮度) 颜色空间中的颜色值转换为 RGB (红、绿、蓝) 颜色空间中的颜色值
- ◀ `matplotlib.pyplot.scatter()` 绘制散点图
- ◀ `numpy.append()` 将给定的数组或值添加到另一个数组的末尾，返回一个新的数组，用于在 NumPy 中实现数组的扩展和拼接操作
- ◀ `numpy.column_stack()` 将两个矩阵按列合并
- ◀ `numpy.copy()` 创建给定数组的副本，返回一个新的数组，使得修改副本不会影响原始数组，用于实现在 NumPy 中进行数组的深拷贝
- ◀ `numpy.empty()` 创建指定形状 NumPy 空（未初始化）数组
- ◀ `numpy.linspace()` 在指定的间隔内，返回固定步长的数据
- ◀ `numpy.meshgrid()` 创建网格化数据
- ◀ `numpy.ones_like()` 用来生成和输入矩阵形状相同的全 1 矩阵
- ◀ `numpy.vstack()` 返回竖直堆叠后的数组
- ◀ `numpy.zeros_like()` 用来生成和输入矩阵形状相同的零矩阵

8.1 RGB 色彩空间

如图 1 所示，当红、绿、蓝取不同值时，我们便可以得到一个五彩斑斓的 RGB 色彩空间。这个空间可以看成是一个“实心”三维立方体。

这一节，我们将利用平面和三维散点图可视化 RGB 色彩空间。

图 1 的下图，我们仅仅看到 RGB 空间中最鲜亮的三个立面，更多颜色都隐藏在这三个立面之下，比如黑色 $(0, 0, 0)$ 就隐藏在角落里。稍后，我们会用‘切片’可视化空间内部的颜色分布。

《矩阵力量》将用 RGB 色彩空间讲解线性代数中的向量空间。

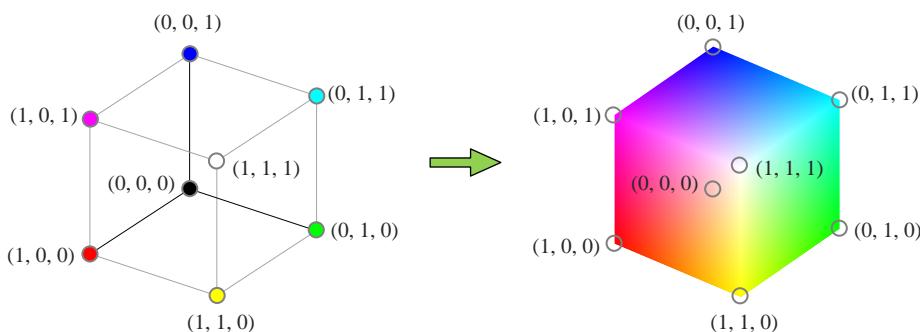


图 1. RGB 色彩空间

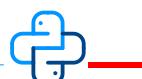
8.2 6 个侧面

我们可以用平面散点图可视化 RGB 色彩空间的 6 个侧面，具体如图 9 所示。

画一个平面散点图时，我们首先用 `numpy.linspace(0, 1, 21)` 生成 $[0, 1]$ 之间的等差数列。然后，用 `numpy.meshgrid()` 将两个等差数列展成二维数据网格。

举个例子，绘制图 9 (a) 这个蓝绿渐变平面时，蓝色、绿色在 $[0, 1]$ 之间渐变，而红色则为用 `numpy.zeros_like()` 生成的全 0。绘制散点图时，我们直接指定每个散点的色号。

图 9 (a) 这个蓝绿色平面，将红色色号加 1，便得到图 9 (b)。几何角度来看，这相当于平面平移。

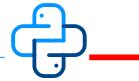


Jupyter 笔记 BK_2_Ch08_1.ipynb 绘制图 9 子图。

有了上面的经验，下一步用三维散点图绘制 RGB 色彩空间的六个立面。

图 10 所示为 RGB 色彩空间外侧三个平面，而图 11 所示为其内侧三个平面。

如图 12 所示，这六个侧面之间存在成对平移关系。绘制这三幅三维散点图时，我们也是指定每个散点的色号。



Jupyter 笔记 BK_2_Ch08_2.ipynb 绘制图 10、图 11、图 12 三幅图子图。

8.3 内部切片

则是用“切片”方法可视化 RGB 空间这个实心立方体的内部。这幅图的子图有三列，每一列子图分别展示红、绿、蓝取特定值 (0.0, 0.2, 0.4, 0.6, 0.8, 1.0)。

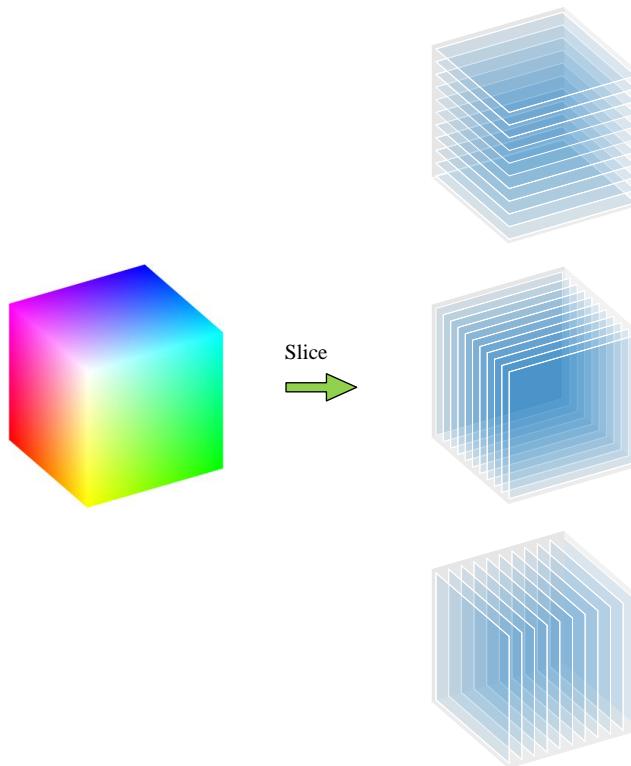
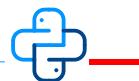


图 2. 三种不同切取 RGB 立方体方式



Jupyter 笔记 BK_2_Ch08_3.ipynb 绘制图 13 子图。

8.4 HSV 色彩空间

RGB 和 CMYK 颜色模型都是面向硬件的，而 HSV 模型更贴合人眼对颜色的感知。

HSV 三个字母分别代表色调 (Hue)、饱和度 (Saturation)、明暗度 (Value)。和 HSV 类似的色彩空间叫 HSL；HSL 中的 L 代表亮度 (lightness)。

`matplotlib.colors.hsv_to_rgb()` 可以将 HSV 色号转换为 RGB 色号。注意，Matplotlib 中 HSV 色号的三个数值也都是在 [0, 1] 之间。

`matplotlib.colors.rgb_to_hsv()` 则将 RGB 色号转换为 HSV 色号。图 14 所示为 RGB 色彩空间到 HSV 和 HSL 色彩空间的转换示意图。特别值得我们关注的是旋转、投影这两步，具体如图 3。本书第 28 章专门讲解三维空间的几何变换。

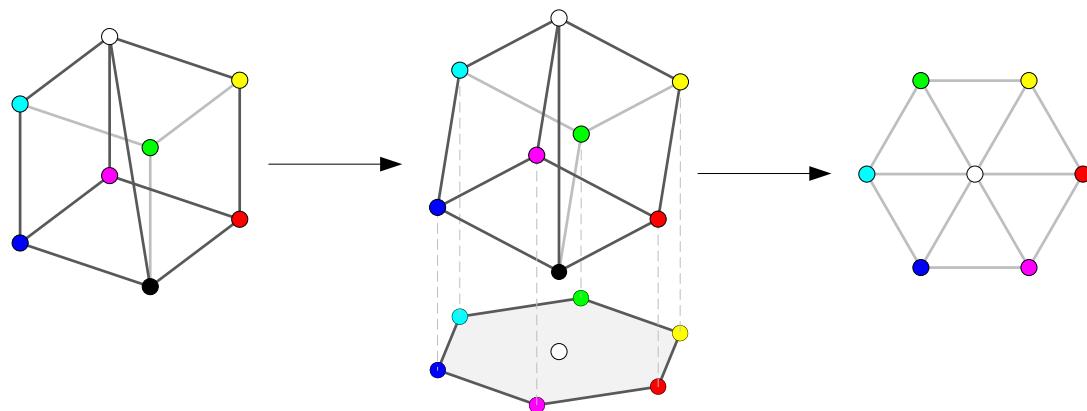


图 3.RGB 色彩空间的旋转和投影，图片改编自 Wikipedia

色调

HSV 中的 H 代表色调 (Hue)。色调一般用角度度量，取值范围为 $0^\circ \sim 360^\circ$ 。

如图 4 所示，从红色开始按逆时针方向计算，红色为 0° ，绿色为 120° ，蓝色为 240° 。红绿蓝的补色分别是黄色 (60°)，青色 (180°)、品红 (300°)。

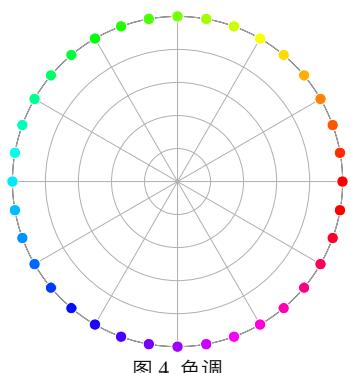


图 4. 色调

图 4 采用的坐标系叫做极坐标系。

在图 5 所示极坐标系中， r 称为**极径** (radial coordinate 或 radial distance)， θ 称为**极角** (angular coordinate 或 polar angle 或 azimuth)。

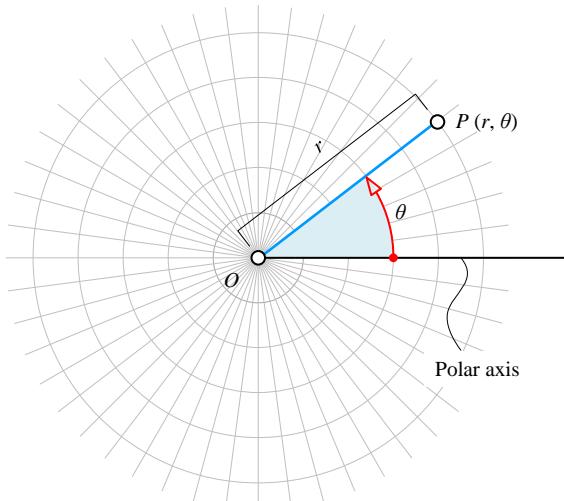


图 5. 极坐标

饱和度

S 代表饱和度 (Saturation)。饱和度的取值范围为 0% ~ 100%，这个值越大，颜色越艳丽。从极坐标角度来看， H 就是极角， S 就是极径。

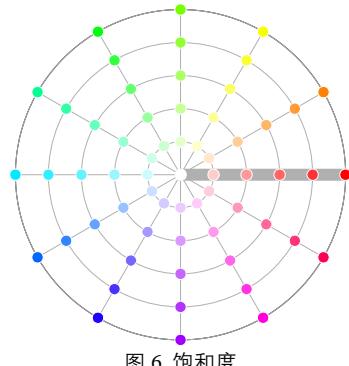


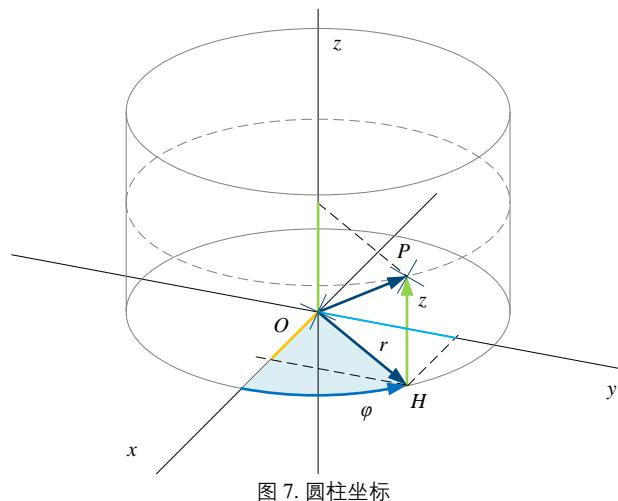
图 6. 饱和度

→ 《数学要素》第 5 章将介绍极坐标系。

明暗度

V 代表明暗度 (Value)。 V 通常取值范围为 0% (黑) 到 100%。如图 7 所示，引入 V ，我们将平面极坐标延展成三维圆柱坐标系。

→ 《数学要素》第 6 章介绍圆柱坐标系。



8.5 两种可视化方案

如果说 RGB 色彩空间是一个实心立方体的话，那么 HSV 色彩空间则是一个实心圆柱体。下面，我们将讨论 HSV 色彩空间的两种可视化方案，具体如图 8 所示。

图 8 (a) 所示为极坐标网格，图 15 所示为 V 取不同值时 HSV 色彩空间“切片”。容易发现这种可视化方案的缺点是，内外圈的散点数量一样多，越往内圈，散点越密。

图 8 (b) 这个可视化方案解决这一问题，每一层圆圈散点数和圆圈半径成正比。这样整幅图的散点看上去类似均匀分布。图 16 便是采用这种方案绘制的可视化方案。

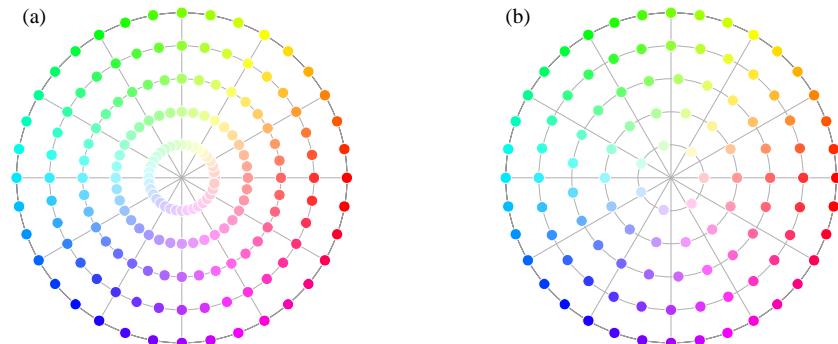


图 8. 两种可视化方案：极坐标网格、均匀分布



Jupyter 笔记 BK_2_Ch08_1.ipynb 和 BK_2_Ch08_2.ipynb 分别绘制图 15、图 16 子图。

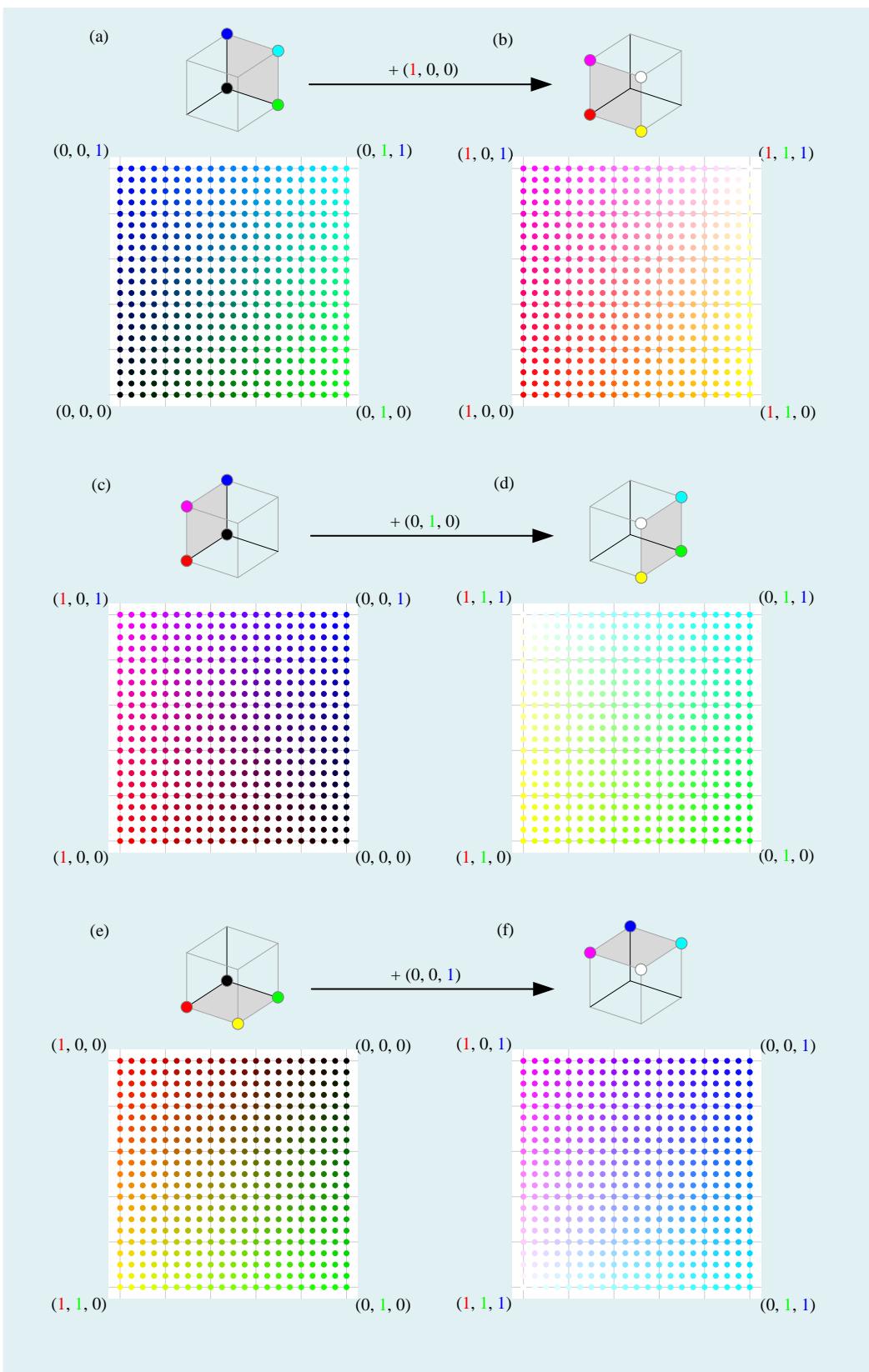


图 9. 用指定色号和平面散点图可视化 RGB 色彩空间的 6 个侧面

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

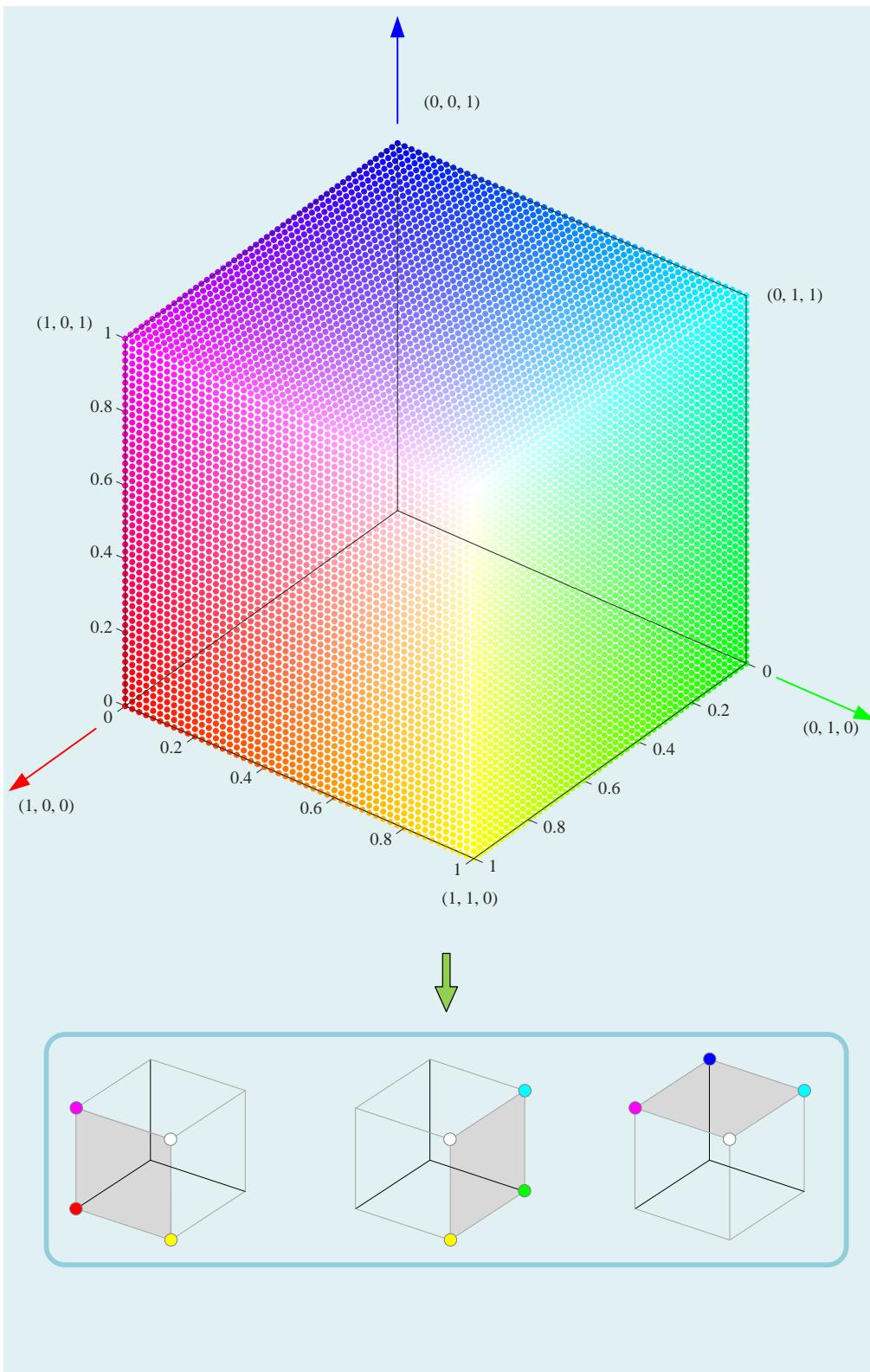


图 10. 三原色张成的彩色空间，外侧三个平面

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

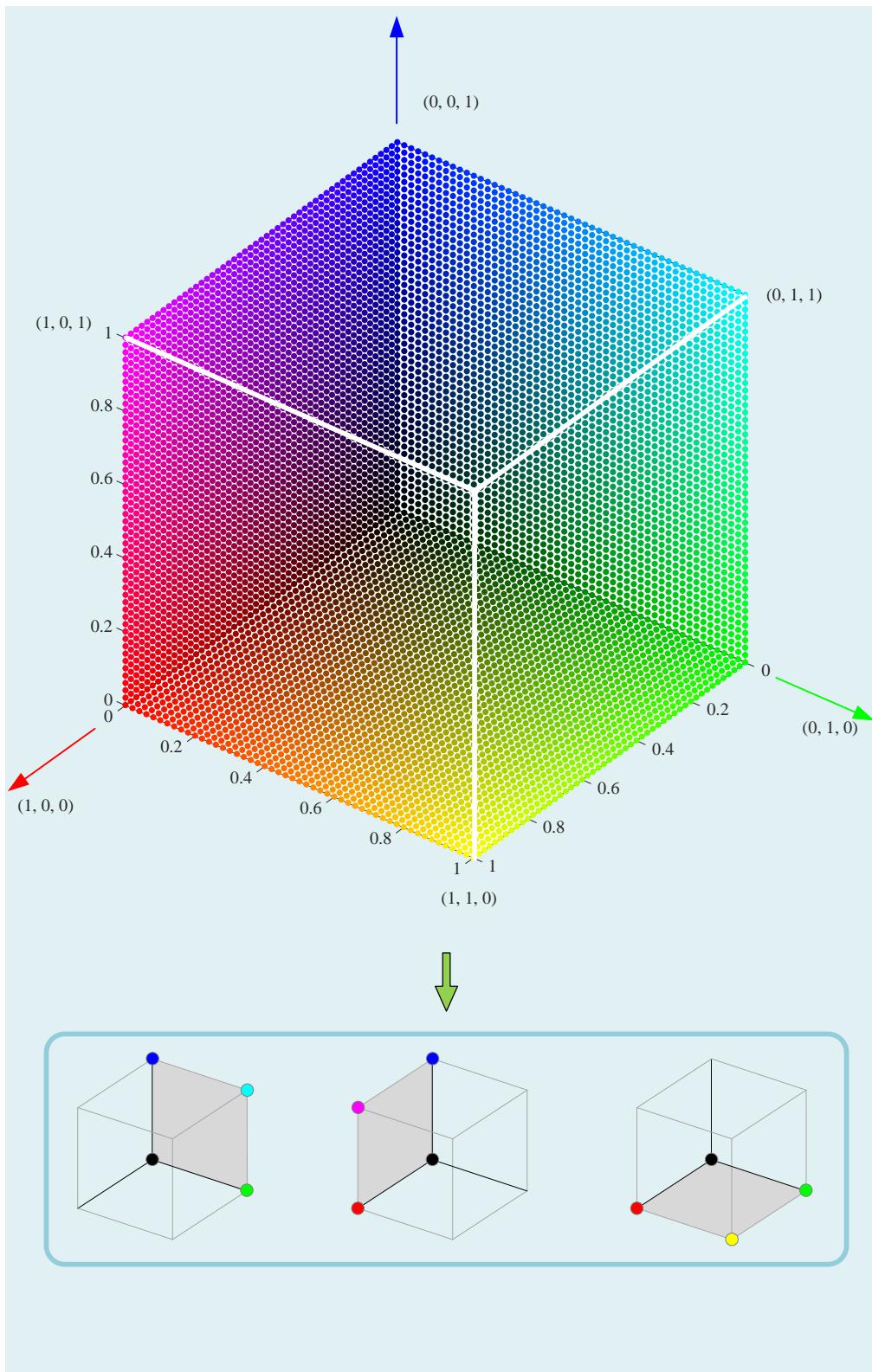


图 11. 三原色张成的彩色空间，内侧三个平面

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

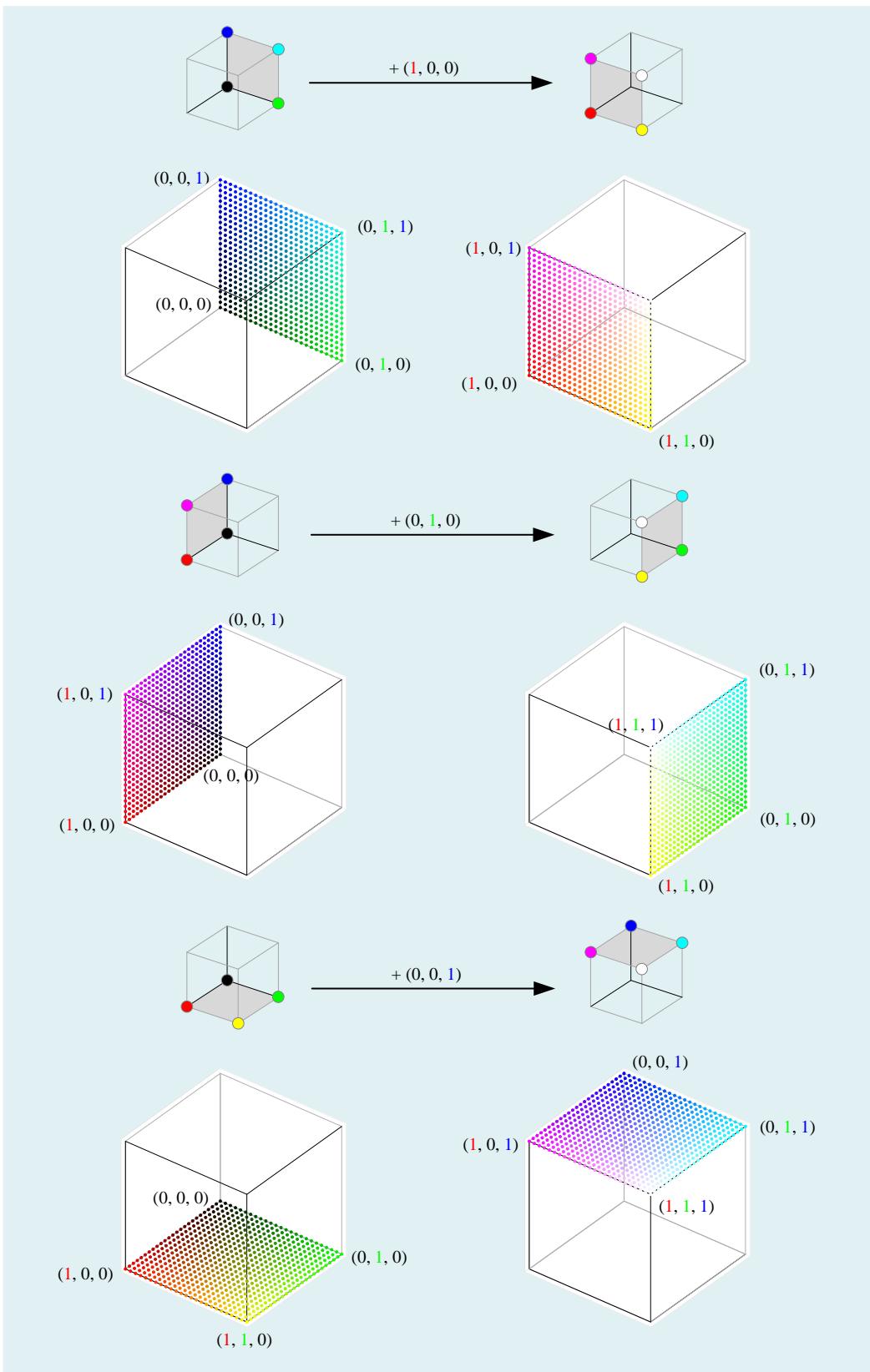


图 12. 三原色张成的彩色空间的六个侧面之间的平移关系

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

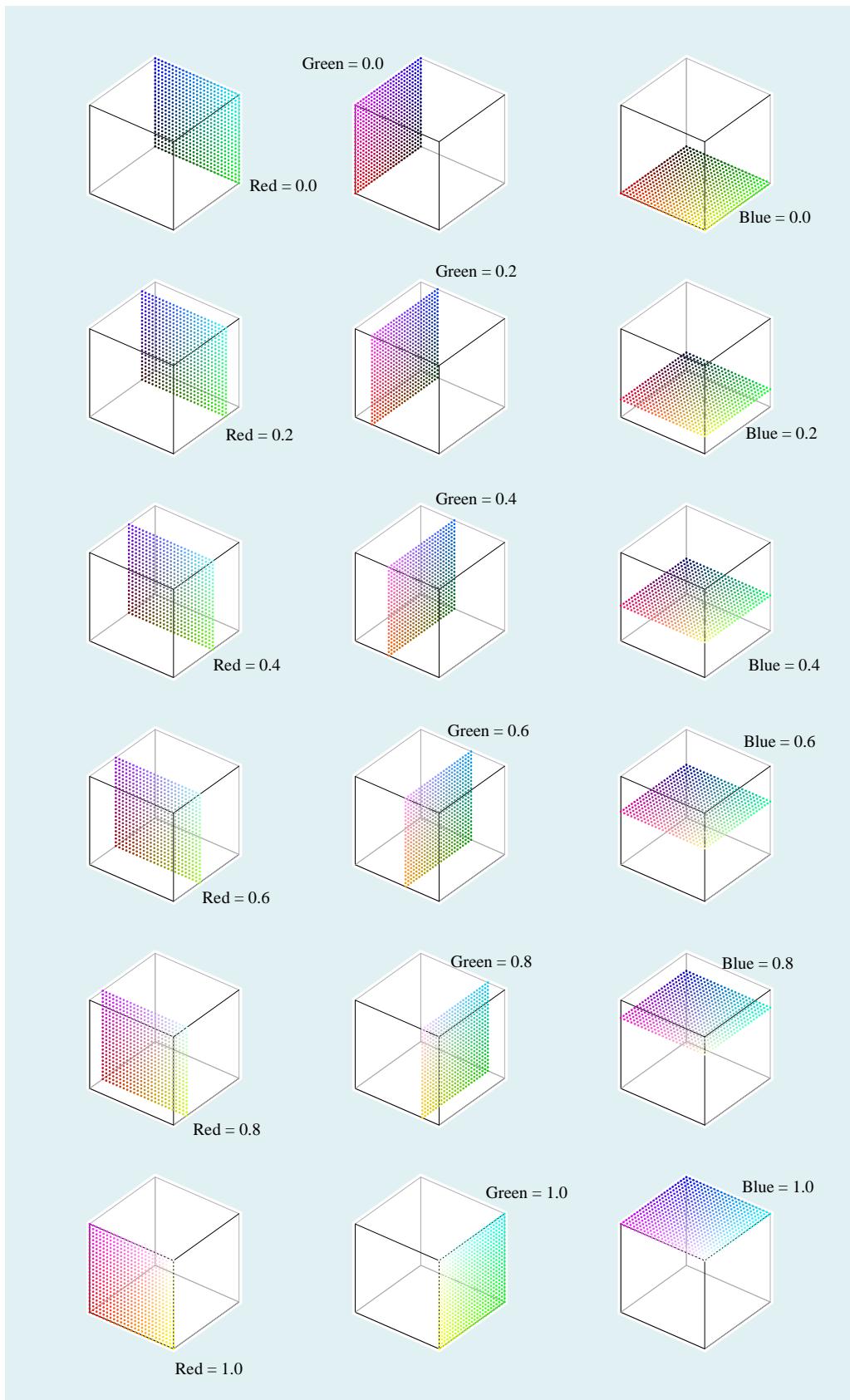


图 13. 用三维切片可视化 RGB 空间内部

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

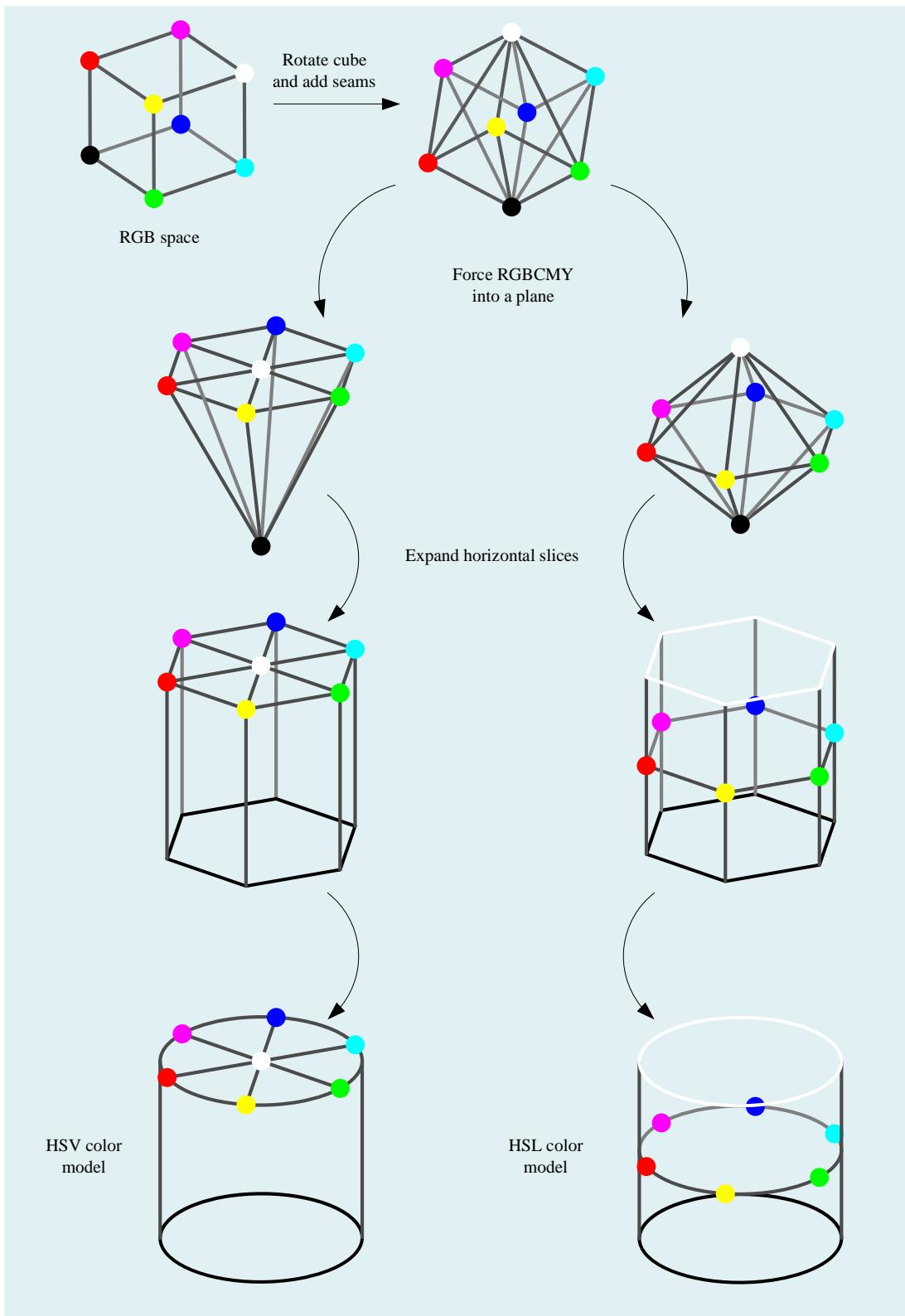


图 14. RGB、HSV、HSL 色彩模型之间的关系，图片改编自 Wikipedia

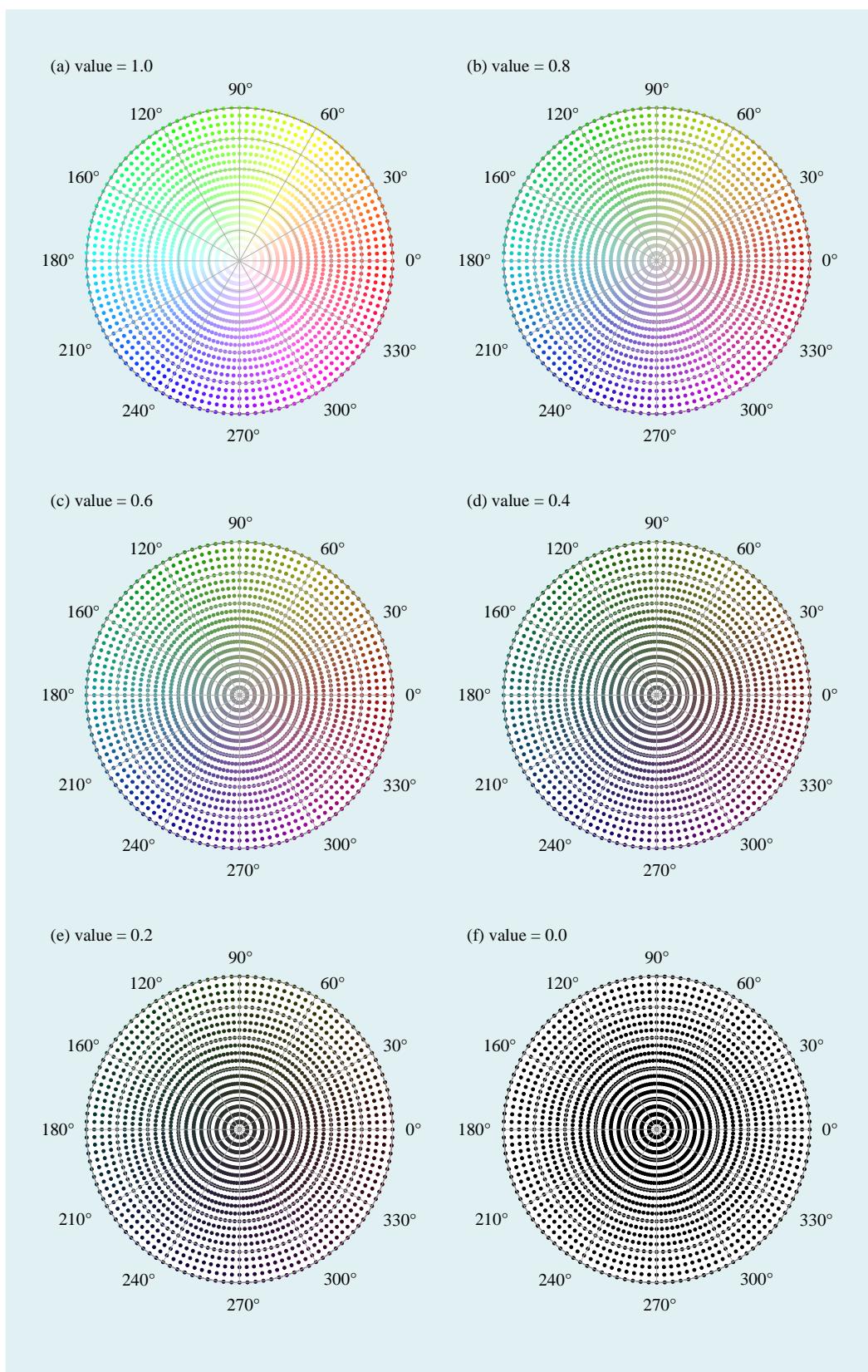


图 15. HSV 色彩空间，极坐标网格数据

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

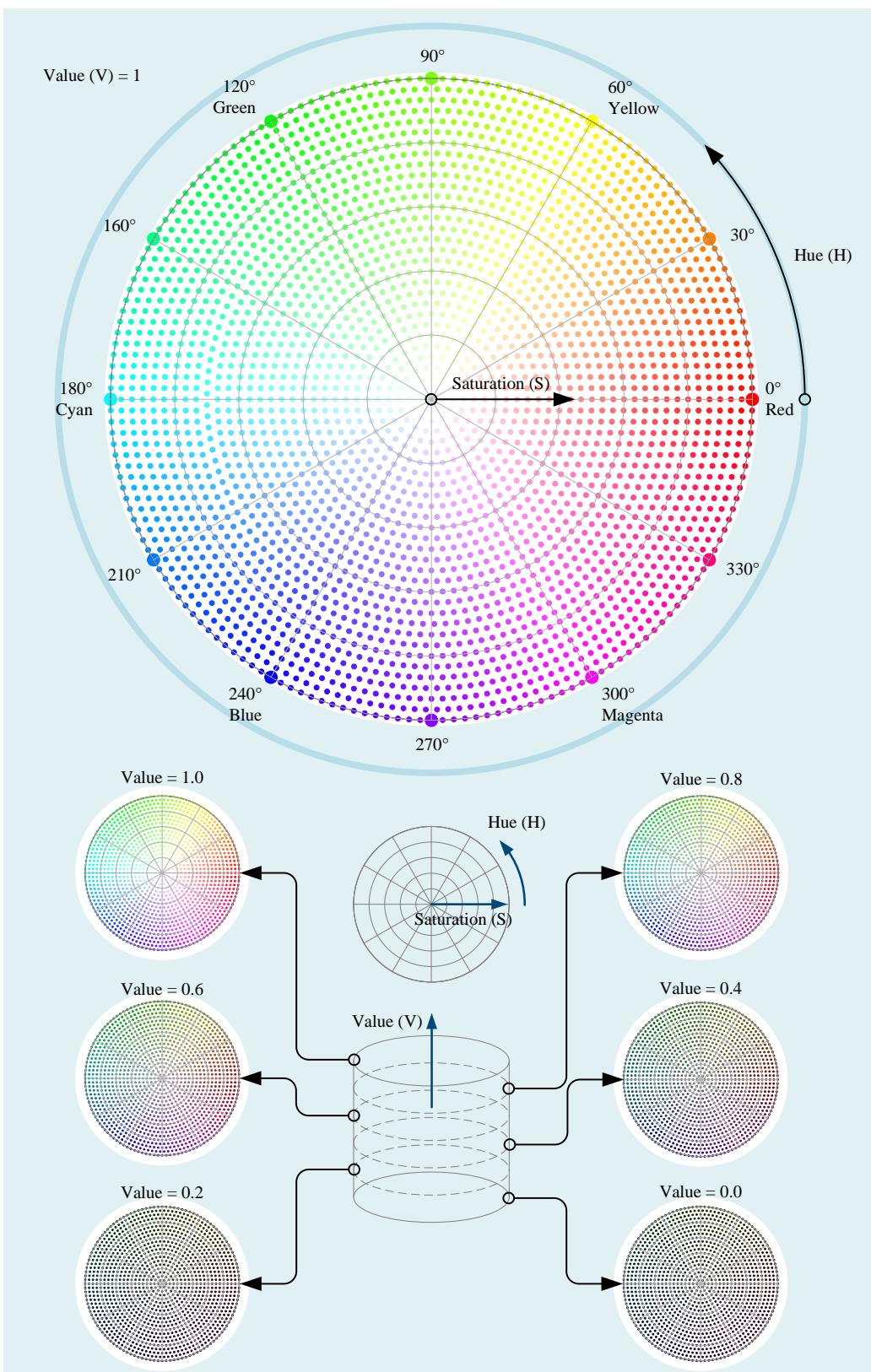


图 16. HSV 色彩空间，散点均匀

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com



Colormaps in Matplotlib

9 色谱

数值序列到连续颜色的映射



一位画家应该在每块画布上都开始一层黑色的底色，因为自然界中的一切事物都是黑暗的，只有在被光线照射的地方才会显现出明亮。

A painter should begin every canvas with a wash of black, because all things in nature are dark except where exposed by the light.

——列奥纳多·达·芬奇 (Leonardo da Vinci) | 文艺复兴三杰之一 | 1452 ~ 1519



- ◀ XXXXX

9.1 色谱

色谱，也叫色调搭配、色图，是指一系列连续变化的颜色。色谱的应用很广泛，比如三维网格面、等高线、热图等等。图 1 所示为 Matplotlib 常见的几种色谱。

Matplotlib 默认的色谱为 viridis。本系列丛书中最常见色谱是图 1 (f) 中的 RdBu 和 RdYlBu。

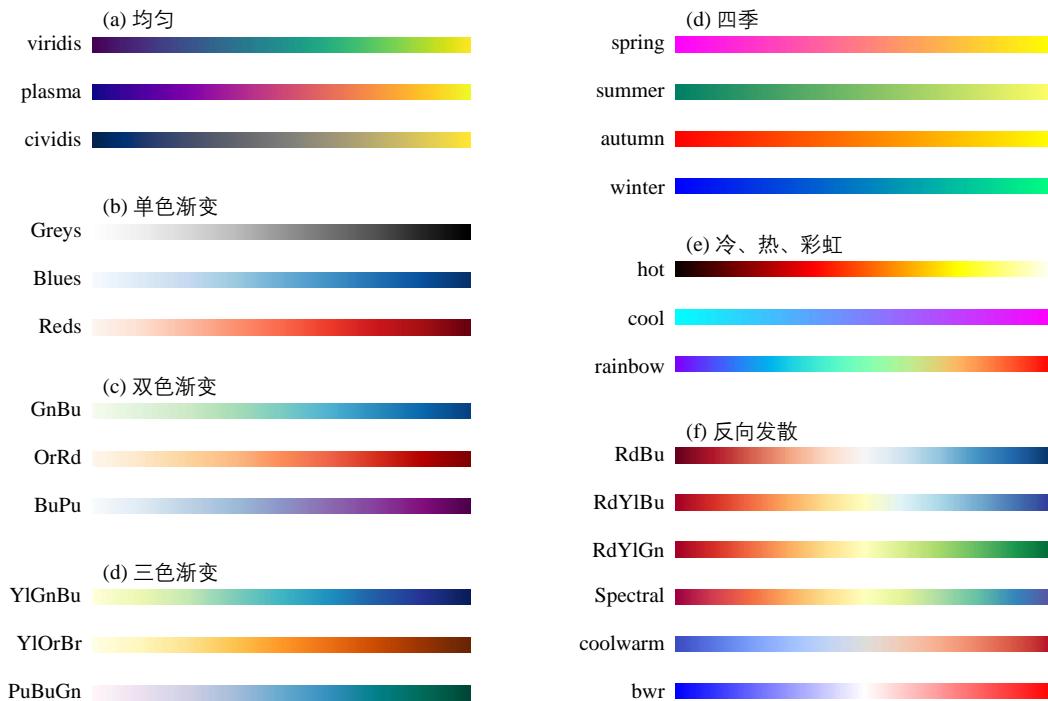


图 1. 几种常用色谱

图 1 (f) 的这几种色谱特别适合可视化穿越 0、且最大值、最小值绝对值近似的函数。此外，相关性系数矩阵热图也常用这类色谱可视化，因为相关性系数取值范围为 $[-1, 1]$ ，也穿越 0 且关于原点对称。

三维网格面

图 16 采用色谱可视化二元函数的取值。图 16 (a) 和 (b) 利用 `plot_surface()` 函数绘制，分别采用 RdYlBu 和 Blues 两种色谱。



《数学要素》第 13 章讲解二元函数。

图 16 (a) 和 (b) 上还加上了色谱条 (color bar) 用来指示不同颜色对应的函数值。

加上“_r”之后，RdYlBu_r 和 Blues_r，色谱的顺序发生调转，如图 16 (c) 和 (d) 所示。

图 16 (e) 和 (f) 则仅仅保留三维网格色谱。这种网格透视效果更好。

等高线

图 17 所示为用几种不同色谱绘制的平面填充等高线、无填充等高线。

本系列丛书中，像 Blues 这种单色渐变色谱，经常用来可视化概率密度函数。因为概率密度函数取值大于等于 0。

本书中，大家还会看到色谱在线图、热图、分类标签等各种可视化应用场景。

更多有关色谱的探讨，请参考：

<https://matplotlib.org/stable/tutorials/colors/colormaps.html>

此外，图 1 也参考上述例子。



Jupyter 笔记 BK_2_Ch09_1.ipynb 和 BK_2_Ch09_2.ipynb 分别绘制图 16、图 17 子图。

下面，我们将用三个视角可视化色谱。这三个视角分别是：RGB 色彩空间、HSV 色彩空间、亮度。

9.2 在 RGB 空间的位置

RdYlBu 在我们眼里是一组渐变的颜色，而每个颜色对应一个 RGB 色号。因此，类似 RdYlBu 这样的色谱实际上就是 RGB 空间的一组坐标。

利用生成的 $[0, 1]$ 之间长度为 100 的等差数列，我们可以从制定的色谱上取得 100 个连续色号。这 100 个色号便对应 RGB 空间 100 个坐标。绘制三维散点时，我们同时给它们赋值对应的色号，图 18、图 19 所示为八个选定的色谱在 RGB 空间的“轨迹”。为了更好地观察，我们设定 4 个观察视角。

特别地，如图 18 (d) 所示，我们发现色谱 cool 实际上就是 cyan 和 magenta 之间的线性插值。

图 19 这四个图谱颜色都很艳丽，但是通过 RGB 这个可视化方案，我们发现四个色谱的差异很明显。

rainbow 的颜色变化较为平滑，而 jet 则多数在 RGB 立方体的三个最鲜亮的面上。turbo 的两个端点的红色和蓝色色号都更靠近原点。也就是所，颜色相对较深。但是 turbo 色号散点轨迹是四个色谱中最平滑的一个，因此颜色过渡均匀。hsv 色谱很特殊，首先它首尾封闭，hsv 的起点和终点都是红色。再者，hsv 所有颜色几乎都在 HSV 色系的饱和度为 1 的边缘上，这一点在 HSV 色彩空间更容易看到。此外，jet 和 hsv 都可以看成由线段构成。



《数据有道》第 5 章将专门介绍插值。

9.3 在 HSV 空间的位置

前文提过，RGB 色彩空间相当于三维直角坐标系，而 HSV 色彩空间相当于圆锥坐标系。既然可以在 RGB 空间可视化色谱，我们也可以在 HSV 色彩空间可视化色谱，具体如图 20、图 21 所示。

用 `matplotlib.colors.rgb_to_hsv()`，我们把 RGB 色号转化为 HSV 色号。H(色调)、S(饱和度)、V(明暗度)三个值都在 [0, 1] 范围之内。而 H 相当于极角，我们需要将其转化成 [0, 2π] 范围之内的值。然后将极坐标转化为直角坐标标。V 值本身就是竖轴值。



Jupyter 笔记 BK_2_Ch09_3.ipynb 绘制图 18、图 19 子图。

Jupyter 笔记 BK_2_Ch09_4.ipynb 绘制图 20、图 21 子图。

亮度

图 2 可可视化八个色谱的亮度。这幅图参考如下 Matplotlib 官方例子：

<https://matplotlib.org/stable/tutorials/colors/colormaps.html>

请大家自行学习，并绘制图 2。

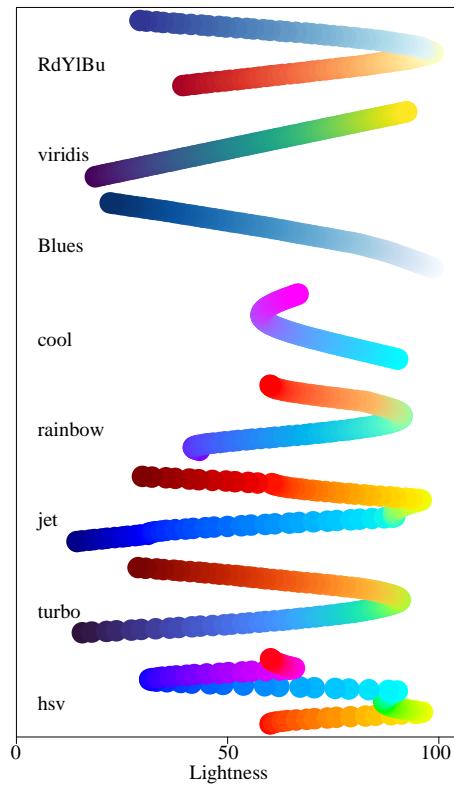


图 2. 八个色谱的亮度

9.4 创建色谱

本节聊聊如何在 Matplotlib 中创建、使用色谱。

两个节点

在 Matplotlib 中创建色谱，至少需要两个颜色作为节点。上一个节，大家已经见过'cool'这个色谱的两端有两个颜色——'cyan' 和 'magenta'，具体如图 3 所示。色谱的左端节点用 0 表示，右侧节点用 1 表示。



图 3. 两个节点，左侧节点 (0.0) 为'cyan'，右侧节点 (1.0) 为'magenta'

下面，我们将左侧节点替换为深蓝色 'darkblue'，得到的色谱如图 4 所示。图 22 (a) 所示为这个色谱在 RGB 空间的具体位置。

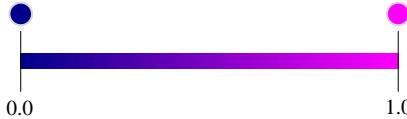


图 4. 两个节点，左侧节点 (0.0) 为'darkblue'，右侧节点 (1.0) 为'magenta'

三个节点

下面，我们用在图 4 色谱的中间 (0.5) 处加一个白色，得到如图 5 所示色谱。这个色谱显然对称。图 22 (b) 所示为这个色谱在 RGB 空间的具体位置。

我们可以用同样的三个颜色构造如图 6 所示的非对称色谱，白色移动到 0.75 处。

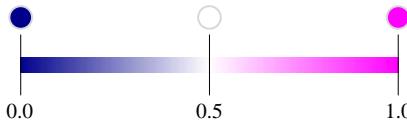


图 5. 三个节点，左侧节点 (0.0) 为'darkblue'，正中间节点 (0.5) 为'white'，右侧节点 (1.0) 为'magenta'



图 6. 三个节点，左侧节点 (0.0) 为'darkblue'，中间节点 (0.75) 为'white'，右侧节点 (1.0) 为'magenta'

五个节点

为了让自定义色谱的颜色渐变更加丰富，我们在图 5 基础上再增加两个节点 ('skyblue' 和 'pink')，得到如图 7 所示色谱。这五个节点均匀布置。这个色谱在 RGB 色彩空间位置如图 22 (c) 所示。将正中间的白色换成黄色，我们便得到图 8 所示色谱。这个色谱在 RGB 色彩空间位置如图 22 (d) 所示。

将图 7 的天蓝色、粉色节点分别向两端靠近，我们便得到如所示图 9 色谱。

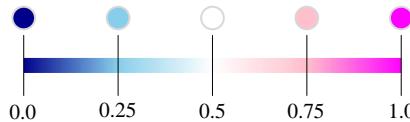


图 7. 五个节点，左侧节点 (0.0) 为'darkblue'，中间有三个节点，右侧节点 (1.0) 为'magenta'，均匀布置

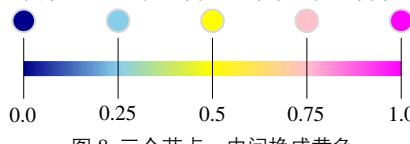


图 8. 三个节点，中间换成黄色

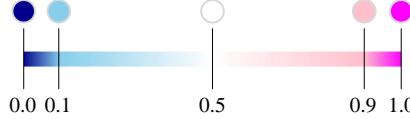


图 9. 五个节点，左侧节点 (0.0) 为'darkblue'，中间有三个节点，右侧节点 (1.0) 为'magenta'，不均匀布置

RGB 色谱

下面，我们用 RGB 三个基色构造一个均匀色谱，具体如图 10 所示。

图 11 则展示了一个首尾连接的循环色谱，这个色谱有四个节点，它们在 RGB 空间的位置如图 23 (a) 所示。这个色谱显得很“暗沉”。我们可以看出，这个色谱线性插值得到的颜色很多都靠近黑色。

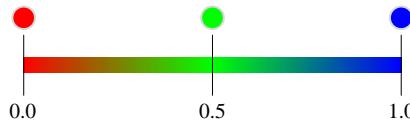


图 10. RGB 色谱，均匀布置

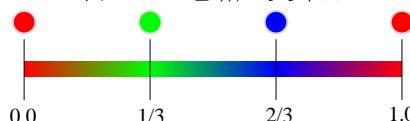


图 11. RGB 色谱，循环

CMY 色谱

作为对比，我们再用 CMY 三个基色构造如图 12 所示色谱。这个色谱的色调显然明亮很多。如图 23 (b) 所示，这个色谱插值得到的颜色都在 RGB 三个最鲜亮的立面上。

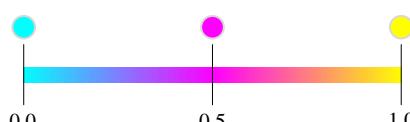


图 12. RGB 色谱，均匀布置

如图 13 所示，我们构造了一个循环色谱。这个色谱有四个节点，它们的位置如图 23 (c) 所示。从 HSV 色彩空间视角来看，这个色谱的所有颜色饱和度 (S) 并非最高。

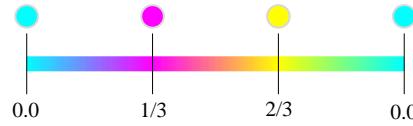


图 13. RGB 色谱，循环

仿制 HSV 色谱

前文大家都见过 hsv 色谱，下面我们自己仿造一个类似色谱。如图 14 所示，这个色谱一共有 7 个节点，首尾循环、均匀布置。色谱在 RGB 色彩空间位置如图 23 (d) 所示。这个色谱所有颜色在 HSV 饱和度最高。

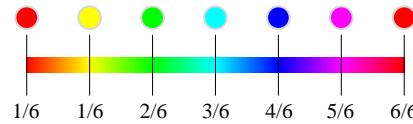
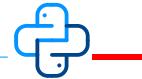


图 14. 仿造 HSV 色谱，循环



Jupyter 笔记 BK_2_Ch09_5.ipynb 绘制图 22、图 23 子图。

热图

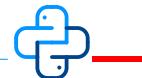
下面，我们用自定义色谱和热图可视化随机数。我们用 seaborn.heat() 绘制随机数，随机数则满足标准正态分布。

图 24 (a) 热图采用图 7 所示热图。为了突出极大、极小的随机数 (可能存在的离群值)，我们可以采用图 9 所示色谱，并得到图 24 (b) 所示热图。

图 24 (c) 则较为特殊，这图热图采用自定义的离散热图。大家可能已经发现，热图变化不再连续。比如， $[-1, 1]$ 之间的随机数都用白色表示。



图 15. 自定义离散色谱



Jupyter 笔记 BK_2_Ch09_6.ipynb 绘制图 24 子图。

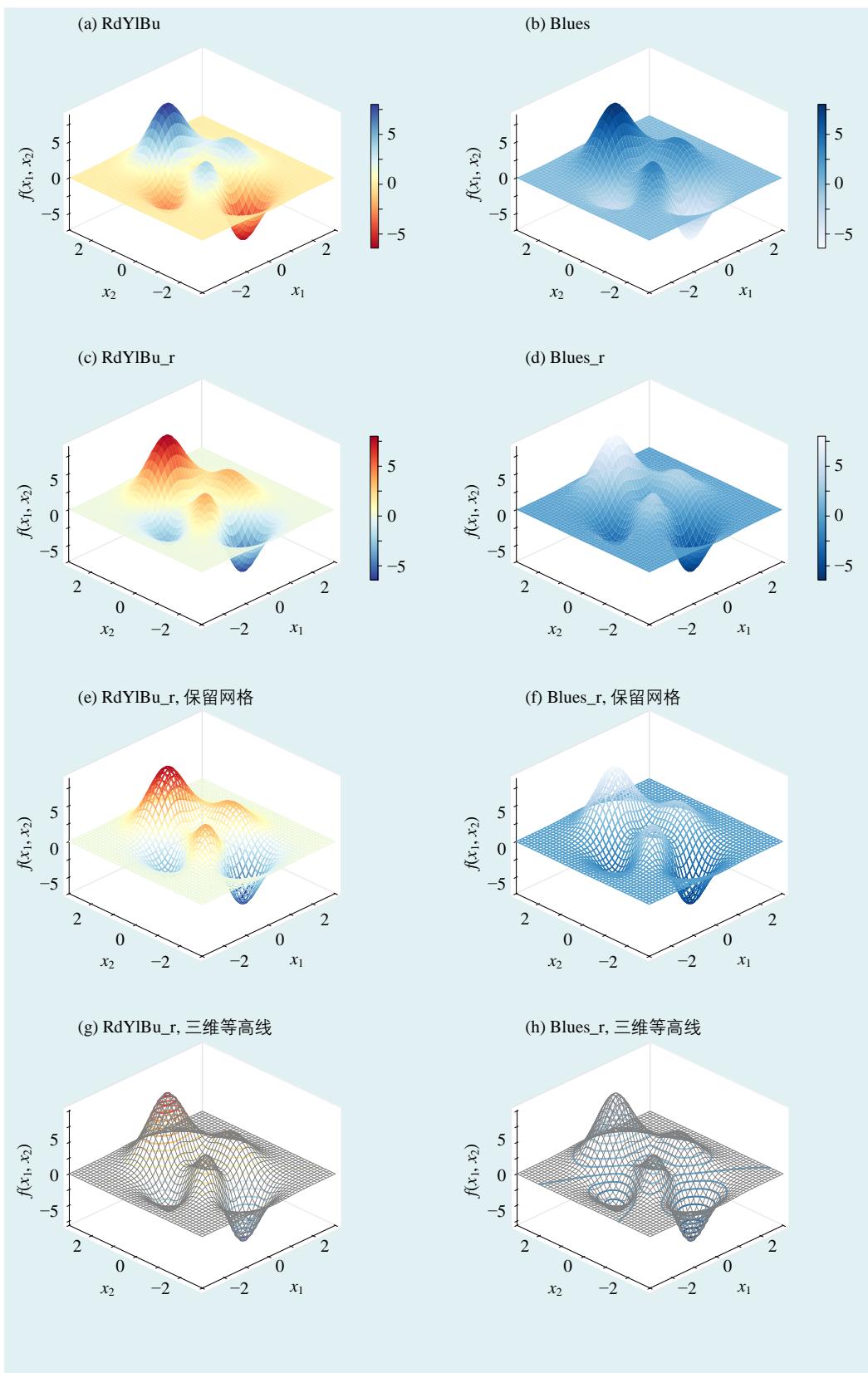


图 16. 三维网格曲面，使用不同色谱

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

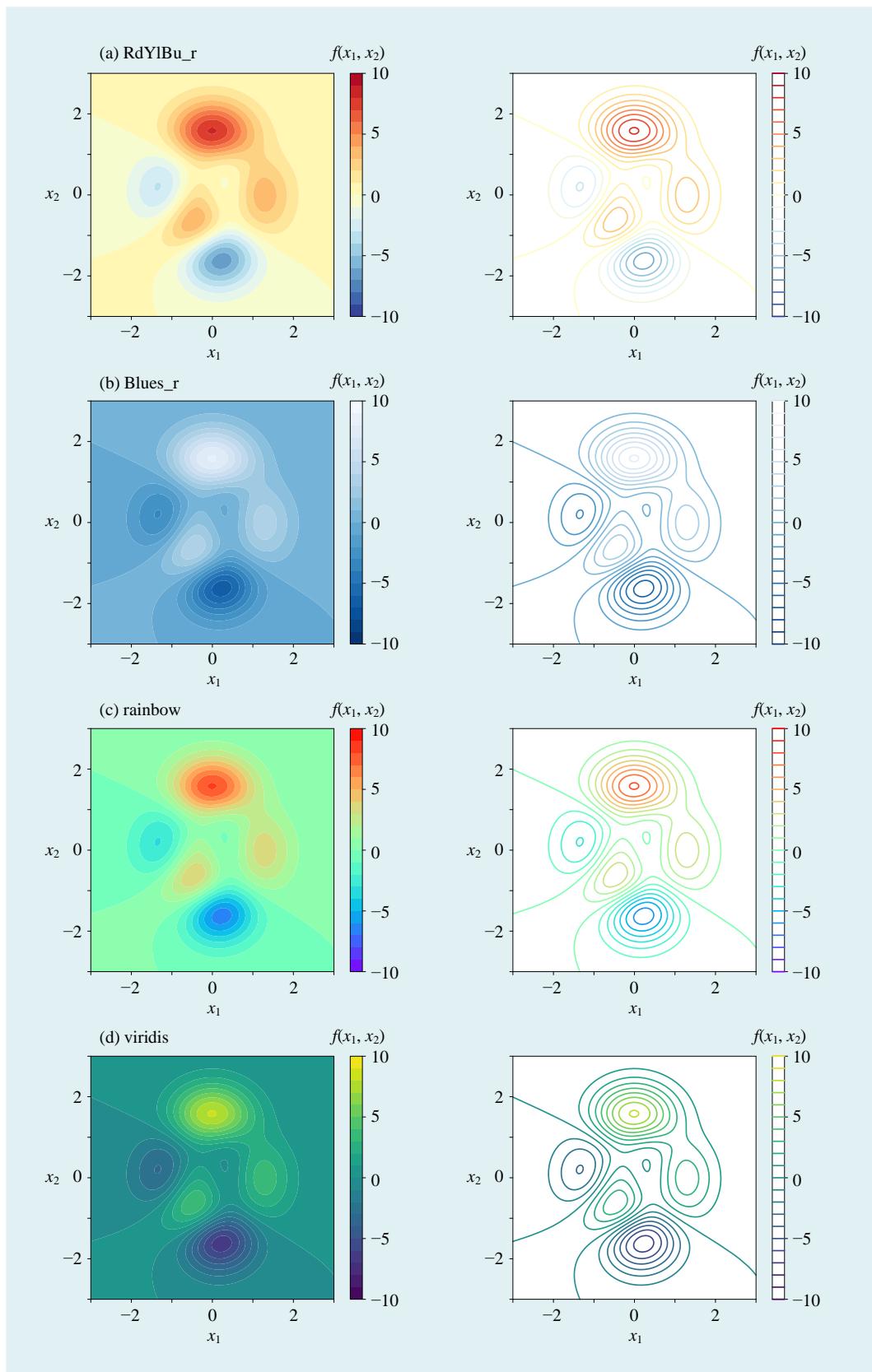


图 17. 平面等高线，使用不同色谱

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

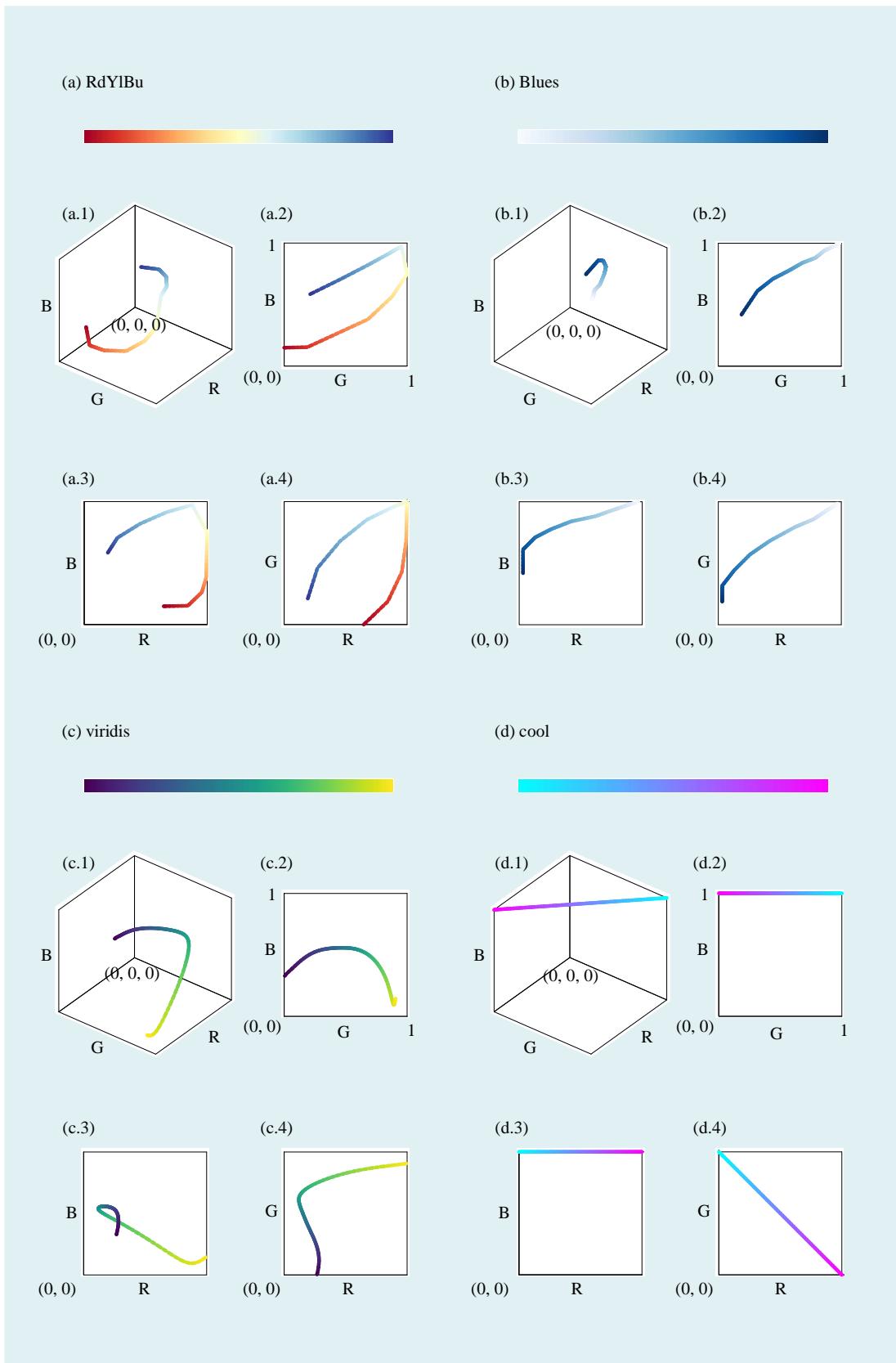


图 18. RGB 空间看 RdYlbu、Blues、viridis、cool 四个色谱

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

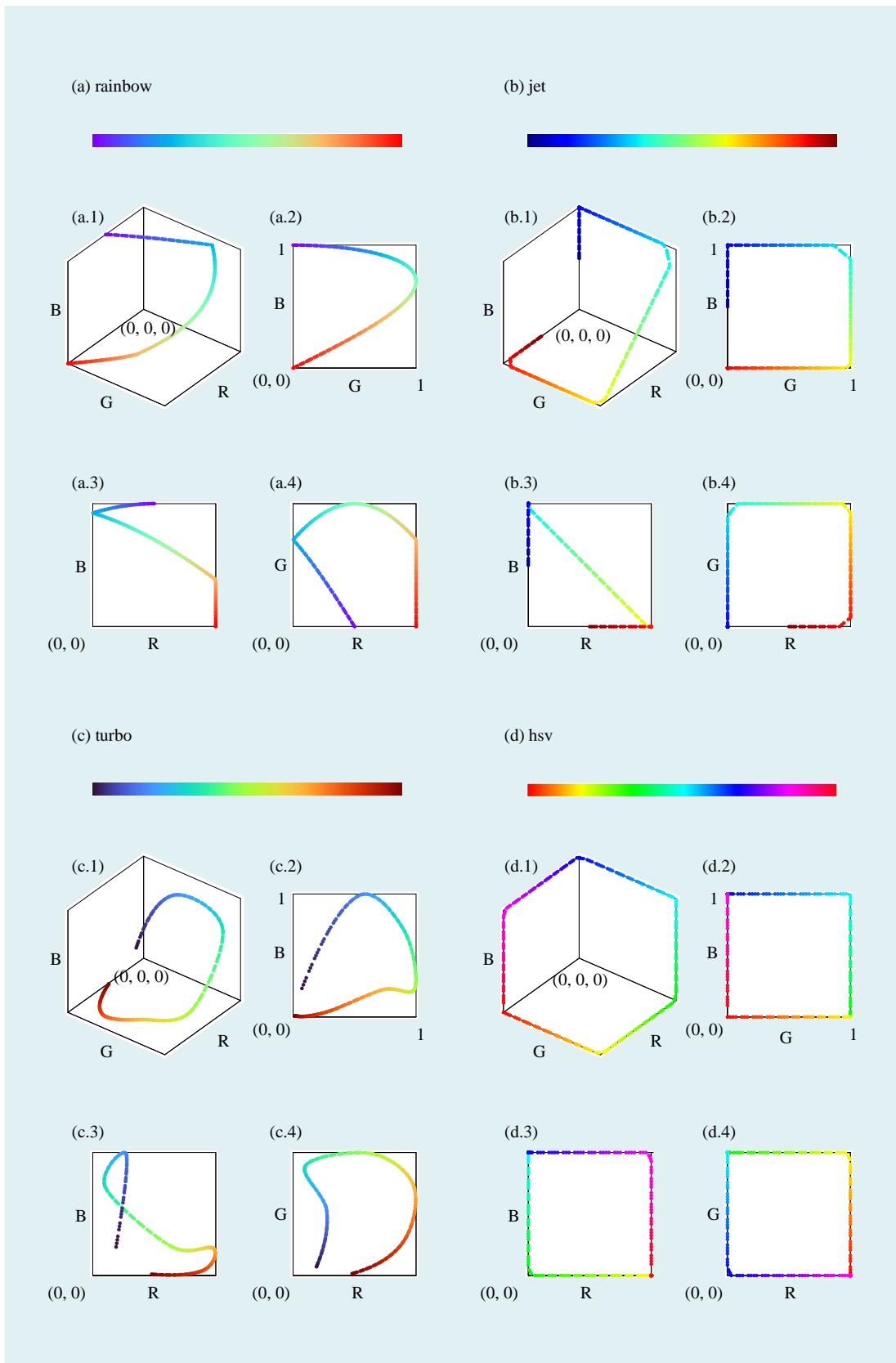


图 19. RGB 空间看 rainbow、jet、turbo、hsv 四个色谱

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

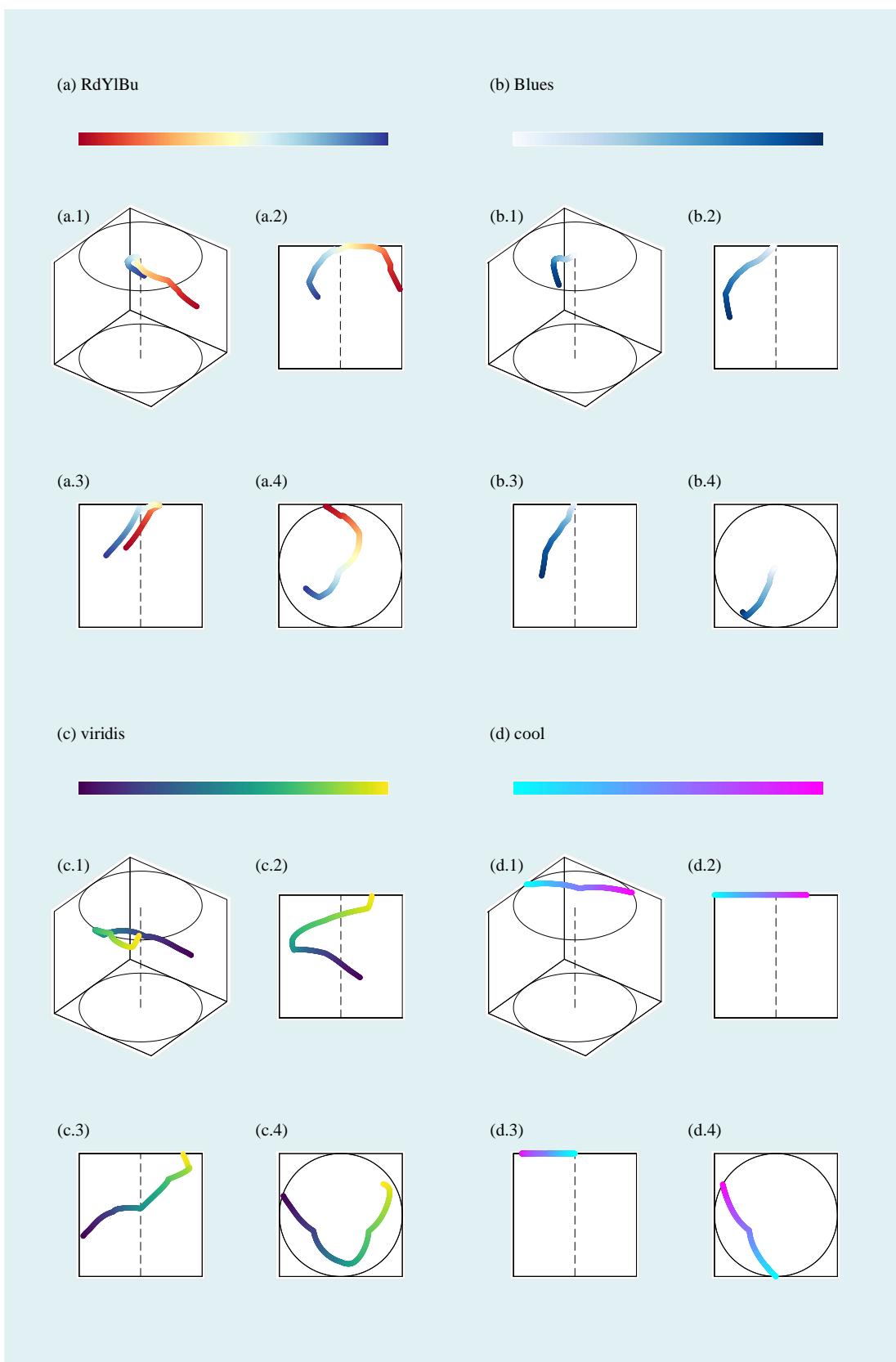


图 20. HSV 空间看 RdYlbu、Blues、viridis、cool 四个色谱

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

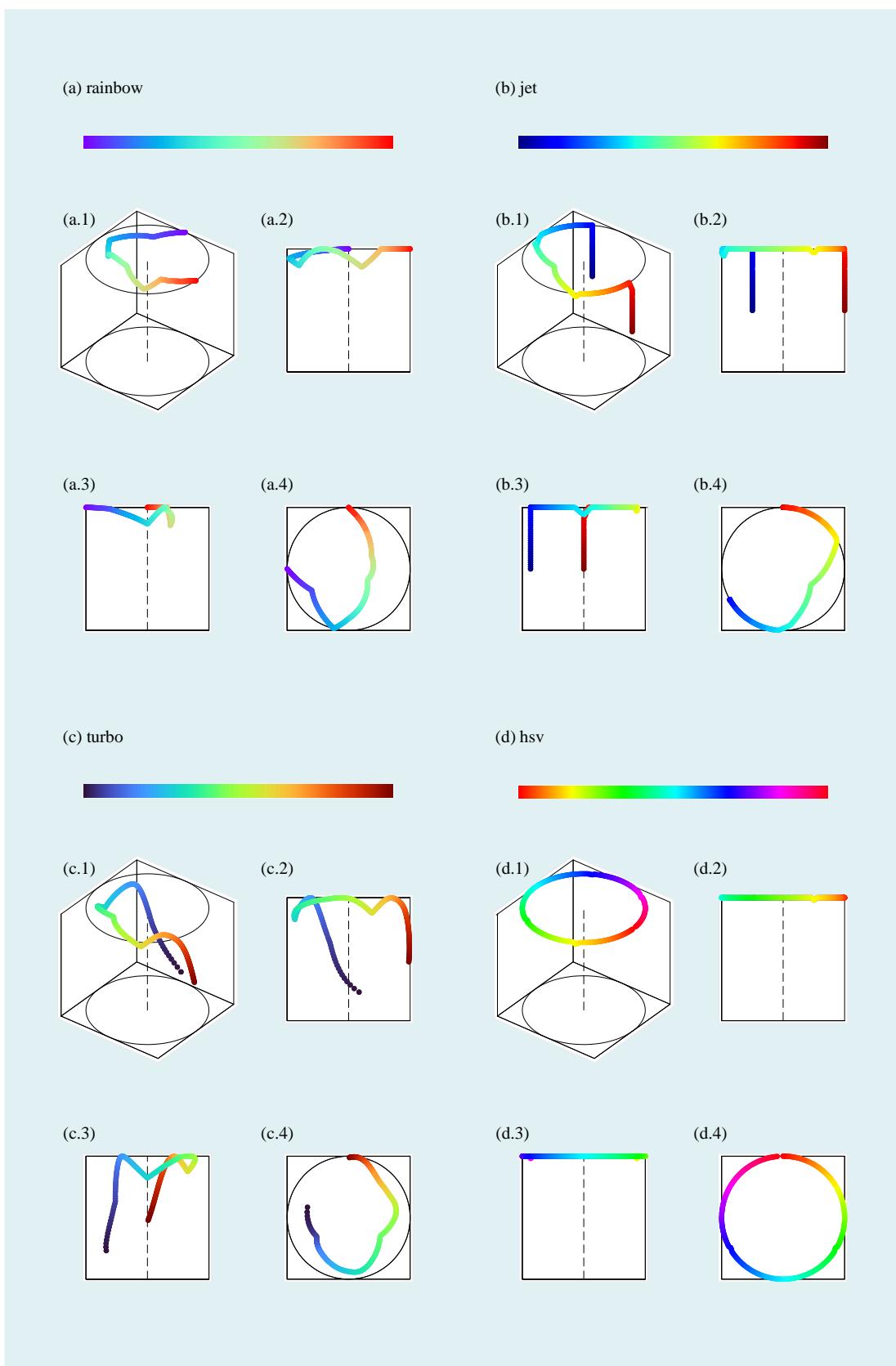


图 21. HSV 空间看 rainbow、jet、turbo、hsv 四个色谱

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

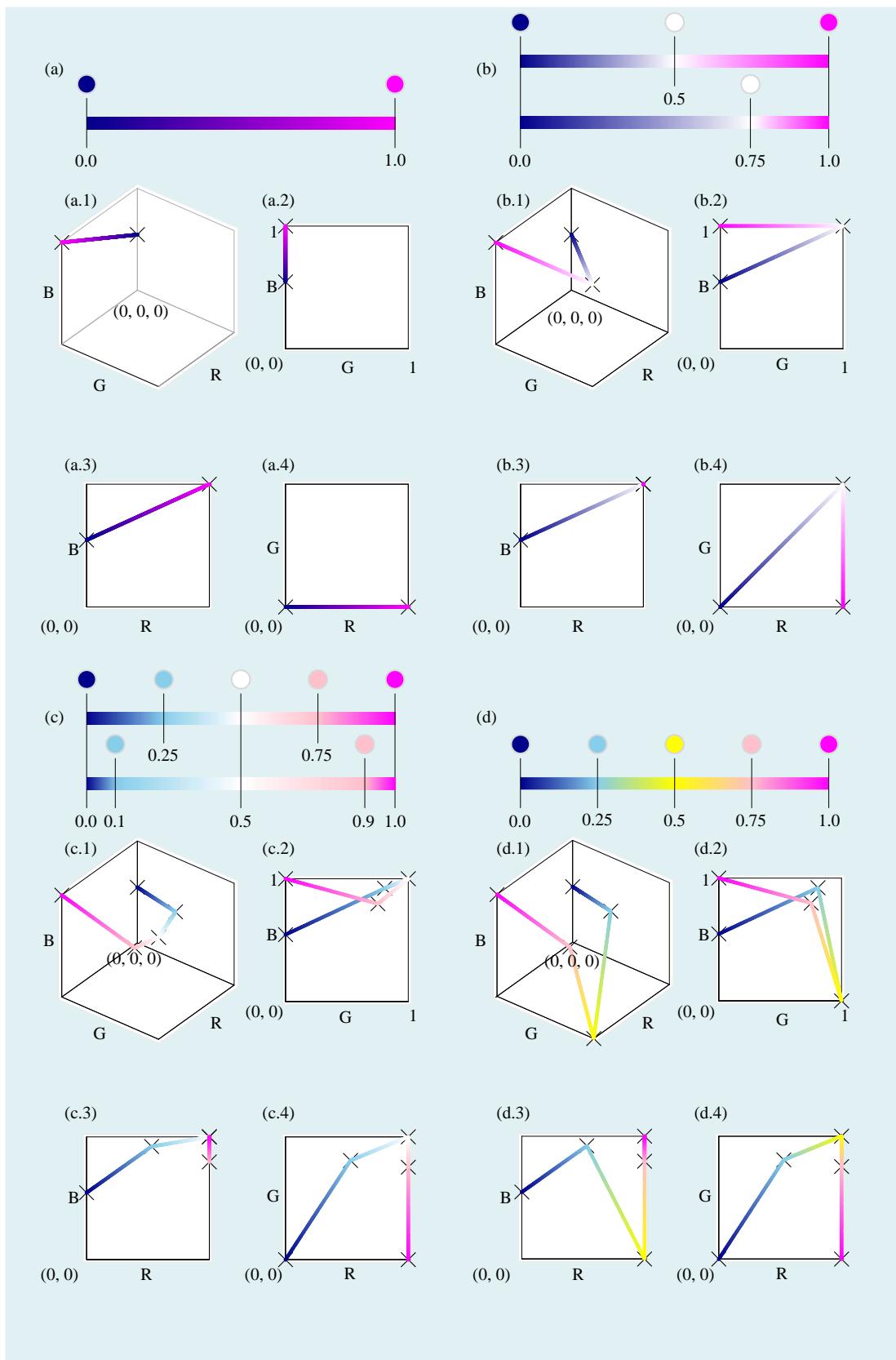


图 22. 自定义色谱在 RGB 色彩空间位置，前四个

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

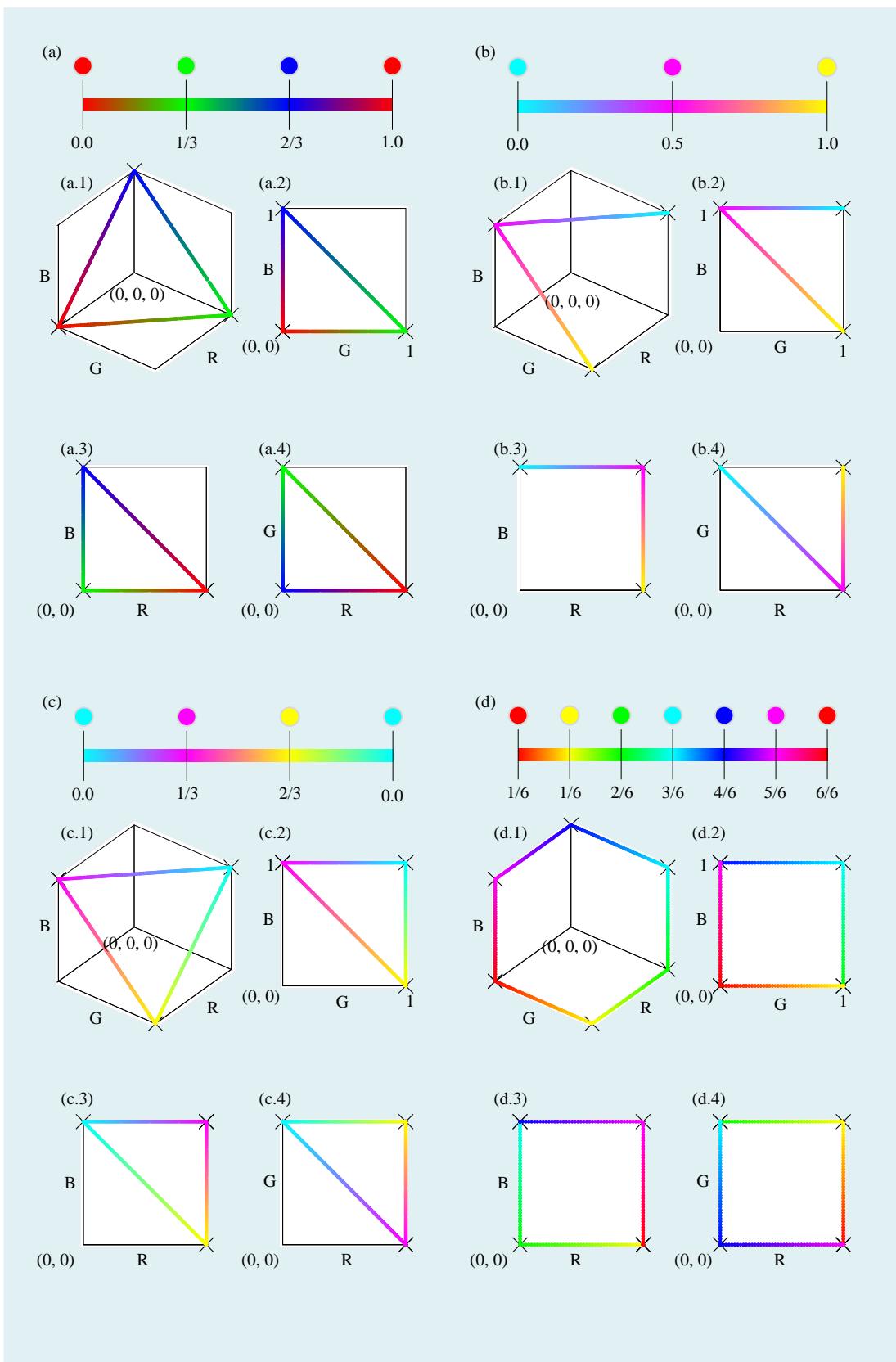


图 23. 自定义色谱在 RGB 色彩空间位置，后四个

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

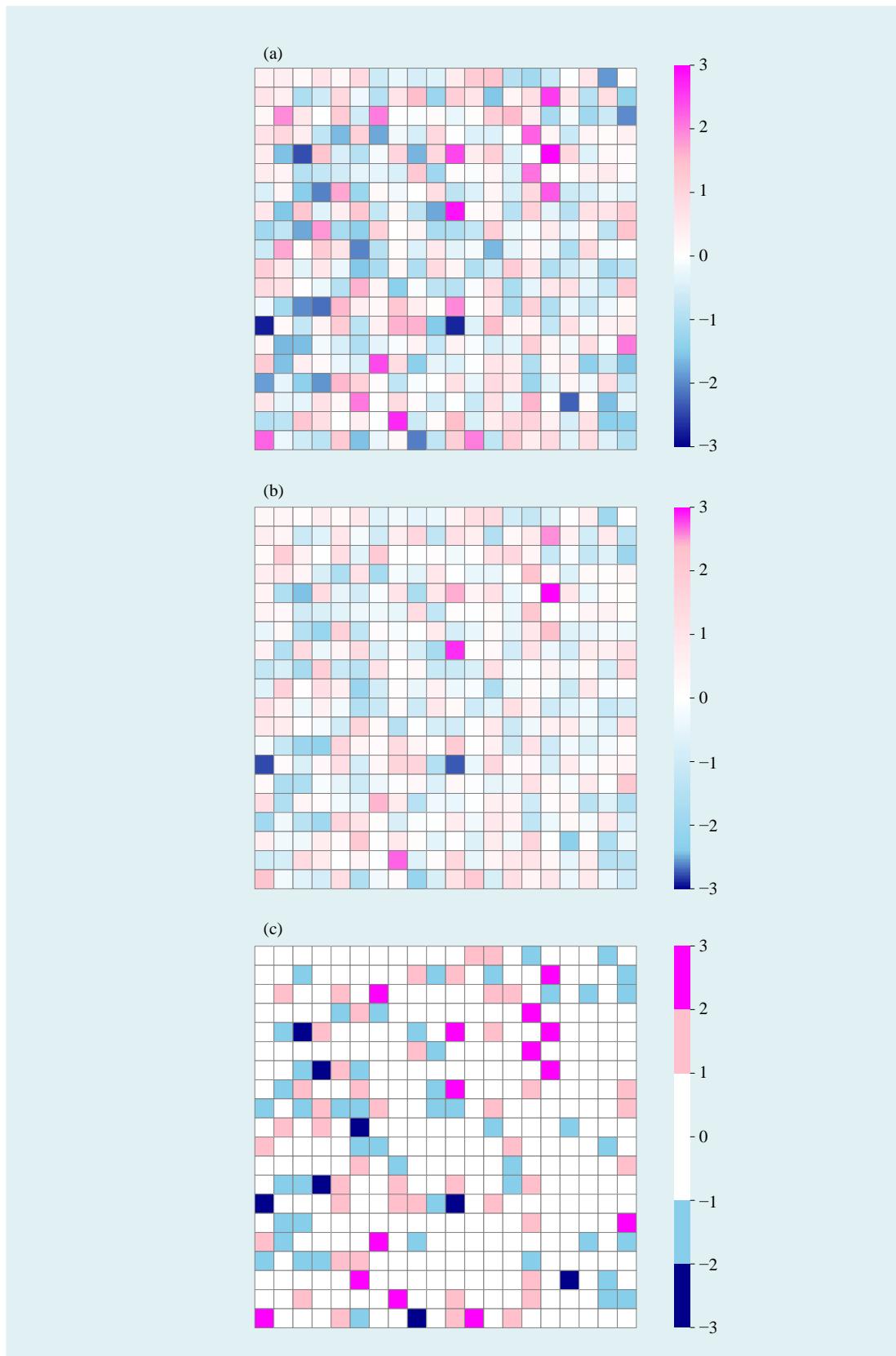


图 24. 热图，使用自定义色谱

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

10

2D Scatter Plots

平面散点

通过散点颜色、大小呈现更多特征



理论上可以用科学描述一切，但这没有意义；这就像把贝多芬的交响乐描述为一组声波，毫无意义。

It would be possible to describe everything scientifically, but it would make no sense; it would be without meaning, as if you described a Beethoven symphony as a variation of wave pressure.

——阿尔伯特·爱因斯坦 (Albert Einstein) | 理论物理学家 | 1879 ~ 1955



- ◀ `matplotlib.patches.Circle()` 创建正圆图形
- ◀ `matplotlib.pyplot.scatter()` 绘制散点图
- ◀ `numpy.exp()` 计算括号中元素的自然指数
- ◀ `numpy.linspace()` 在指定的间隔内，返回固定步长的数据
- ◀ `numpy.meshgrid()` 创建网格化数据
- ◀ `numpy.random.rand()` 生成满足均匀分布的随机数
- ◀ `numpy.random.randn()` 生成满足标准正态分布的随机数
- ◀ `seaborn.scatterplot()` 绘制散点图
- ◀ `sklearn.neighbors.KernelDensity()` 概率密度估计函数

10.1 平面散点图

点动成线，线动成面，面动成体。本章介绍如何在平面上绘制最基本的散点图。本书后中，大家会发现，线图也是散点的连线；等高线、曲面也离不开点。

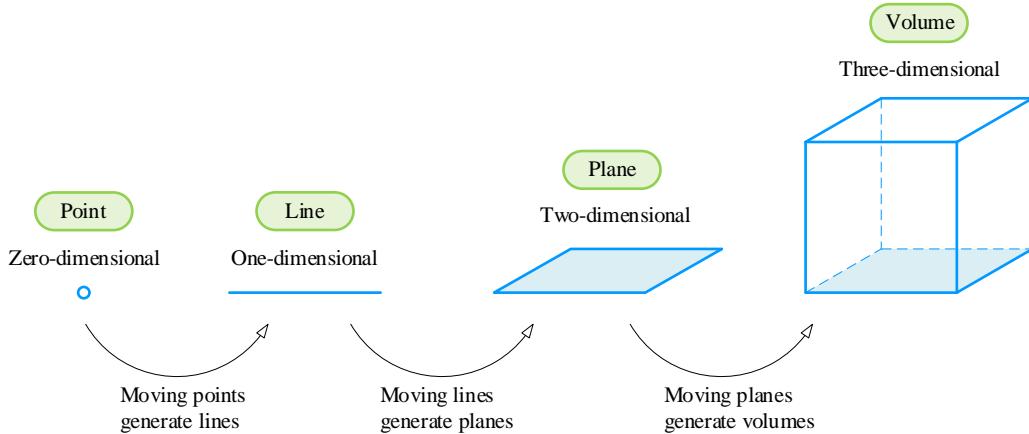


图 1. 点动成线，线动成面，面动成体

规则网格

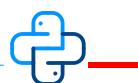
平面散点图是重要的可视化工具。如图 2 (a) 所示，在平面网格散点基础上用颜色渲染可以可视化 3D 数据。进一步提高颗粒度，我们可以得到更加丰满的平面图像，如图 2 (b) 所示。这一点，我们在本书后文三维散点图中还会看到。

散点大小

如图 2 (c) 所示，除了颜色，我们还可以用散点大小展示数据特征。

随机散点

除了规则网格散点，我们更常用平面散点可视化随机散点，比如图 2 (d)。因此，平面散点常用来可视化样本数据。



Jupyter 笔记 BK_2_Ch10_1.ipynb 绘制图 2。

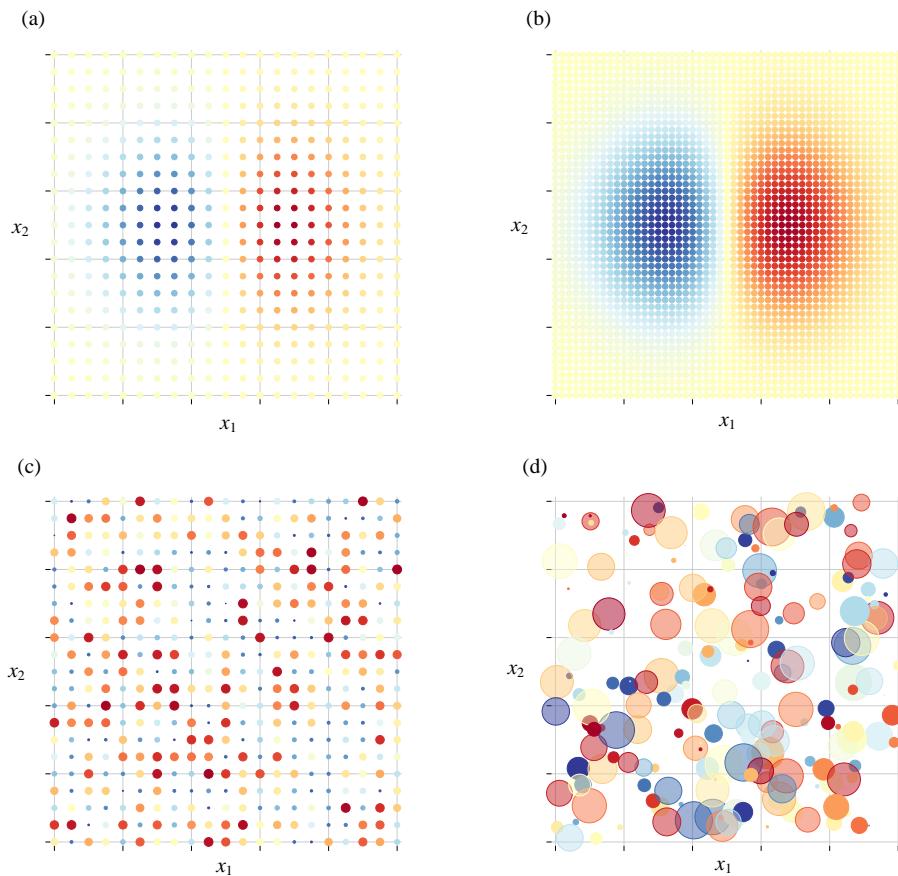


图 2. 使用 matplotlib.pyplot.scatter() 绘制平面散点图

10.2 样本数据

可视化样本数据

图 3 所示为平面散点可视化鸢尾花样本数据。这幅图中，我们可以用颜色、大小、标记符号可视化更多特征。

本书后文将专门介绍用 seaborn.scatterplot() 绘制散点图。



Jupyter 笔记 BK_2_Ch10_2.ipynb 绘制图 3。

10.3 使用面具

图 4、图 5 所示为用面具 (mask)，也叫蒙皮，区分满足不同条件的散点。

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

68-95-99.7 法则

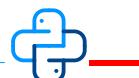
图 4 中，大家会看到一组服从高斯分布的散点。以 0 ± 2 为界， $[-2, 2]$ 区间之内的散点用原点 ● 展示； $[-2, 2]$ 区间之外的散点用叉 × 代表。这体现的实际上是 68-95-99.7 法则。

68-95-99.7 法则是一种统计学中的规则，也被称为“三个标准差法则”或“标准差法则”。该法则用于描述服从高斯分布样本数据分布情况。根据 68-95-99.7 法则，对于一个符合正态分布的数据集，大约：68% 的数据值会落在均值的一个标准差范围内；95% 的数据值会落在均值的两个标准差范围内；99.7% 的数据值会落在均值的三个标准差范围内。

注意，图 4 中样本数据的均值为 0，标准差为 1。 $[-2, 2]$ 区间之内约有 95% 样本数据。

换句话说，大约 68% 的数据会分布在均值左右一个标准差的范围内，约 95% 的数据会分布在均值左右两个标准差的范围内，而约 99.7% 的数据会分布在均值左右三个标准差的范围内。这个法则在统计学和数据分析中被广泛应用，用于估计数据的分布情况和识别异常值。它提供了一种简单而有用的方法来理解和描述正态分布的特性。

鸢尾花书《统计至简》第 9 章将专门讲解一元高斯分布。



Jupyter 笔记 BK_2_Ch10_3.ipynb 绘制图 4。请大家想办法区分 68-95-99.7 对应的不同区间。

蒙特卡罗模拟估算圆周率

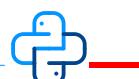
蒙特卡罗模拟是一种使用随机抽样的方法来估算数值的技术，可以用于估算圆周率。下面是使用蒙特卡罗模拟来估算圆周率的一般步骤。

- ▶ 假设有一个边长为 2 的正方形，其中包含一个半径为 1 的圆。
- ▶ 在正方形内部随机生成一组点，可以通过在正方形内均匀抽样得到。每个点都有一个 x 和 y 坐标，均在 $[-1, 1]$ 的范围内。
- ▶ 对于每个生成的点，计算其到原点的距离。
- ▶ 如果距离小于等于 1，表示该点在圆内或圆上，否则在圆外。
- ▶ 统计在圆内的点的数量和在正方形内生成的总点数。
- ▶ 估算圆周率的值可以通过以下公式计算： $\pi \approx (4 \times \text{圆内点的数量}) / (\text{正方形内生成的总点数})$ 。

随着生成的点数增多，根据蒙特卡罗模拟的原理，估算得到的圆周率值会逐渐接近真实值 π 。因此，增加生成的点数可以提高估算的准确性。

需要注意的是，蒙特卡罗模拟是一种概率估算方法，结果的准确性取决于随机性和抽样点的数量。在实际应用中，通常需要生成大量的点才能得到比较准确的估算结果。

鸢尾花书《统计至简》第 15 章将专门讲解蒙特卡罗模拟。



Jupyter 笔记 BK_2_Ch10_3.ipynb 绘制图 5。

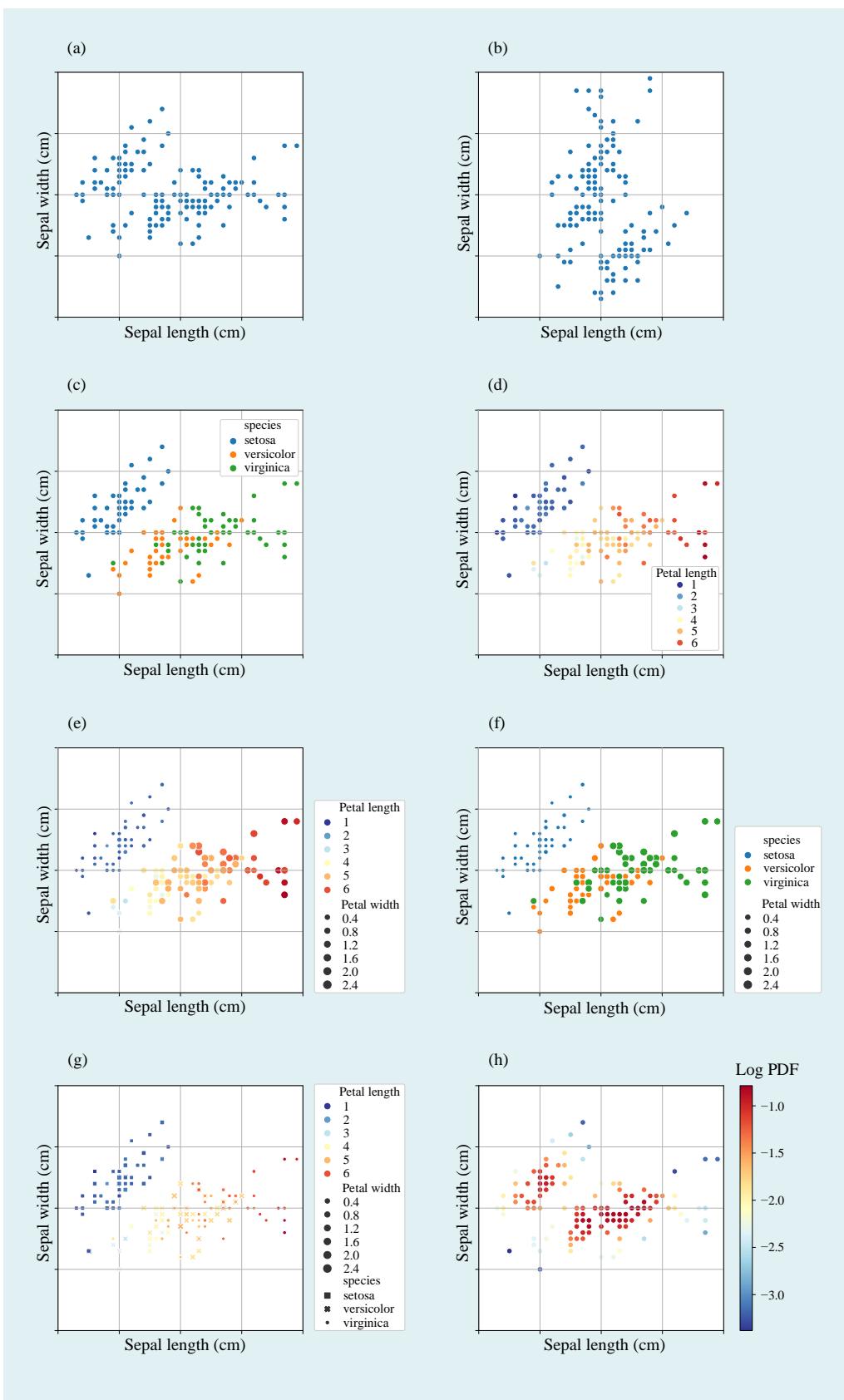


图 3. 用平面散点图可视化鸢尾花数据

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

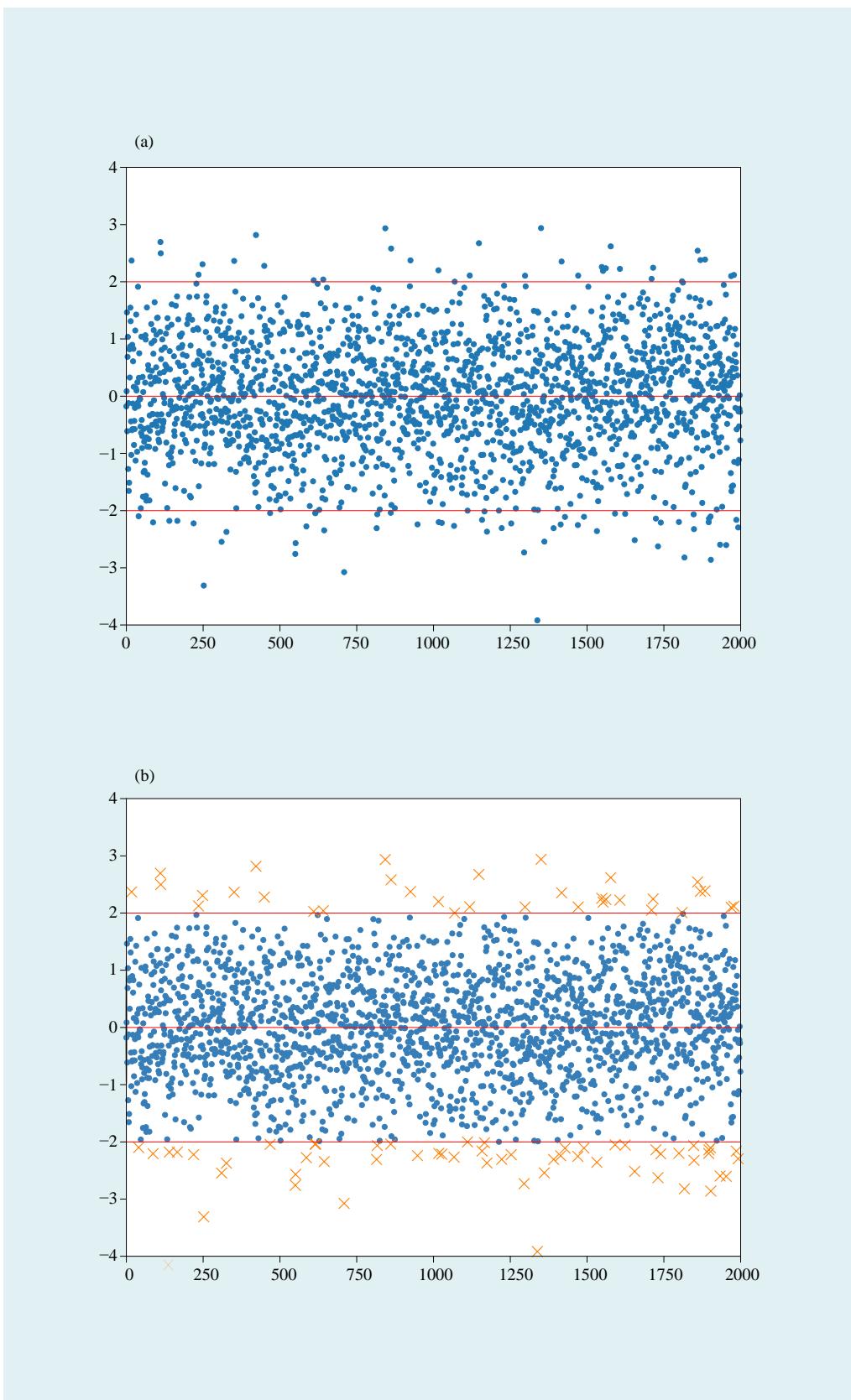


图 4. 使用面具可视化可能的离群值

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

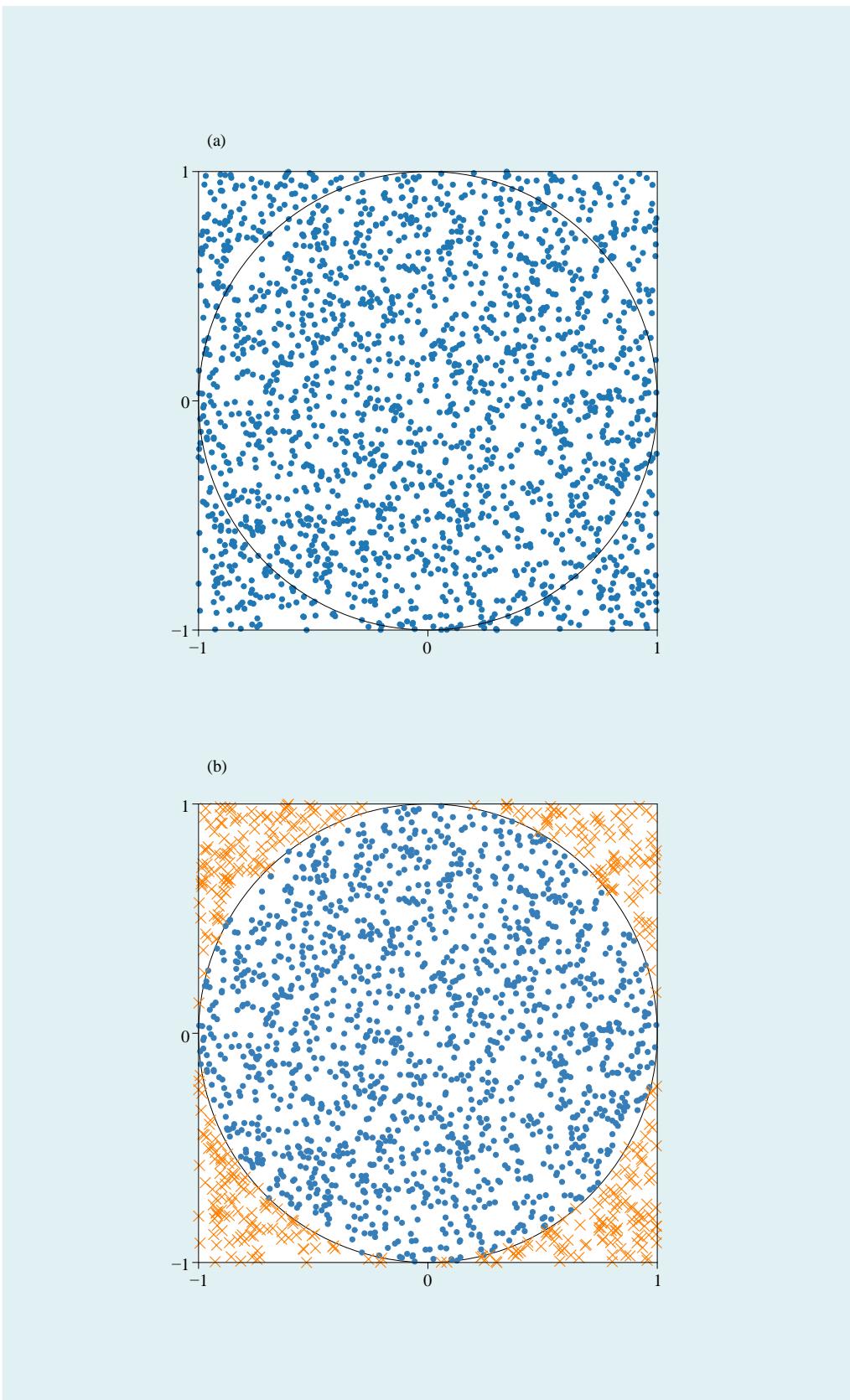


图 5. 蒙特卡洛模拟估算圆周率

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

11

2D Line Plots

平面线图

实际上也是散点顺序相连的折线图



艺术的目的不在于展示事物外在的美，而是其内在的价值。

The aim of art is to represent not the outward appearance of things, but their inward significance.

——亚里士多德 (Aristotle) | 古希腊哲学家 | 384 ~ 322 BC



- ◀ `matplotlib.collections.LineCollection()` 是 Matplotlib 中的一个集合对象，用于绘制多条线段的集合
- ◀ `matplotlib.pyplot.axhline()` 绘制水平线
- ◀ `matplotlib.pyplot.axvline()` 绘制竖直线
- ◀ `matplotlib.pyplot.Normalize()` 函数是用于将数据归一化或标准化到指定的范围内的函数
- ◀ `matplotlib.pyplot.stem()` 绘制火柴梗图
- ◀ `numpy.arange()` 根据指定的范围以及设定的步长，生成一个等差数组
- ◀ `numpy.argwhere()` 返回一个数组中满足指定条件的元素的索引
- ◀ `numpy.concatenate()` 将多个数组进行连接
- ◀ `numpy.cumsum()` 计算累计求和
- ◀ `numpy.linspace()` 在指定的间隔内，返回固定步长的数据
- ◀ `numpy.log()` 底数为 e 自然对数函数
- ◀ `numpy.log10()` 底数为 10 对数函数
- ◀ `numpy.log2()` 底数为 2 对数函数
- ◀ `numpy.random.normal()` 生成满足高斯分布的随机数
- ◀ `numpy.sign()` 函数返回一个数组中每个元素的符号值
- ◀ `numpy.sin()` 计算正弦值
- ◀ `numpy.vstack()` 返回竖直堆叠后的数组
- ◀ `zip(*)` 用于将可迭代的对象作为参数，将对象中对应的元素打包成一个个元组，然后返回由这些元组组成的列表。*代表解包，返回的每一个都是元祖类型，而并非是原来的数据类型

11.1 点动成线

点动成线，线动成面。散点顺序连线的结果就是线图。

颗粒度

绘制线图时，大家首先注意颗粒度 (granularity)，即采样。

多数情况，在绘制一元函数线图时，我们用 `numpy.linspace()` 生成自变量的等差数列。图 1 的两幅图中的散点都来自于正弦函数 $f(x) = \sin(x)$ 。显然，颗粒度粗糙时，用线图可视化一元函数可能会误导读者。

等差数列的公差越小，曲线的颗粒度越高，这样平面线图看上去“光滑”。如图 2 (a) 所示，等差数列有 101 个元素。将这些散点顺序连接便得到图 2 (b)。对于 $f(x) = \sin(x)$ 这个并不复杂的一元函数，图 2 (a) 的颗粒度显然足够用了。

如图 3 所示，为了可视化 $f(x) = 1/\sin(x)$ 在靠近 0 附近的振荡，我们需要极其细腻的颗粒度。

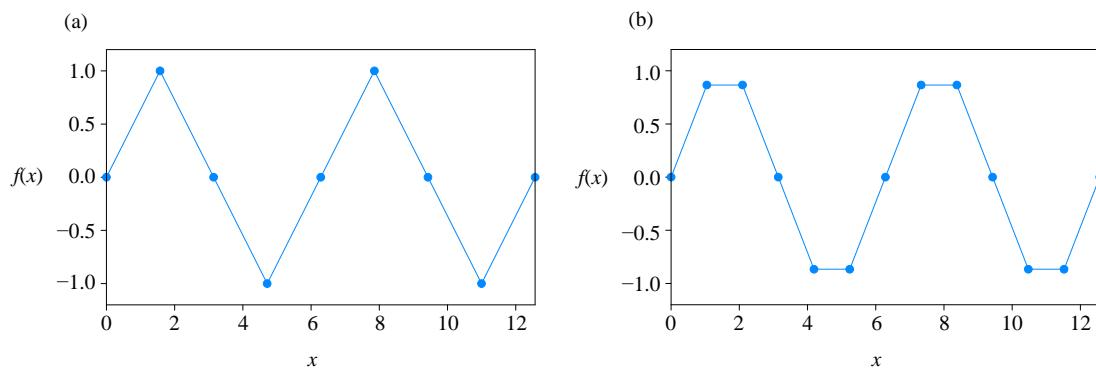


图 1. 颗粒度粗糙

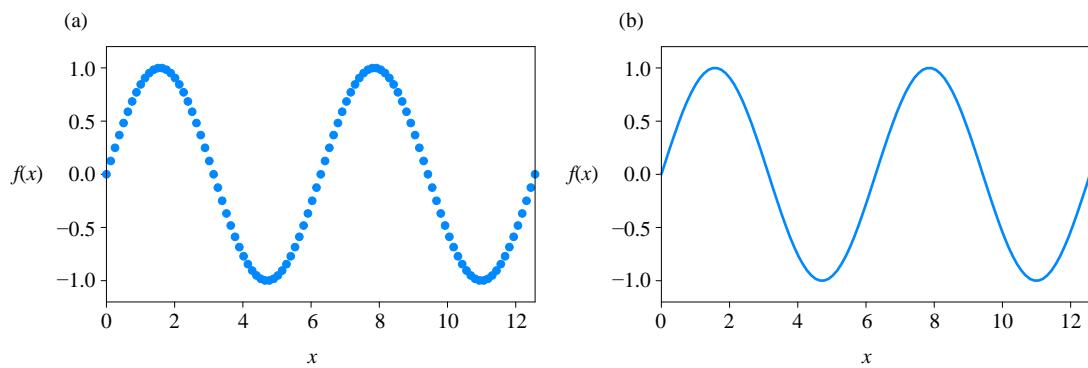


图 2. 颗粒度合适

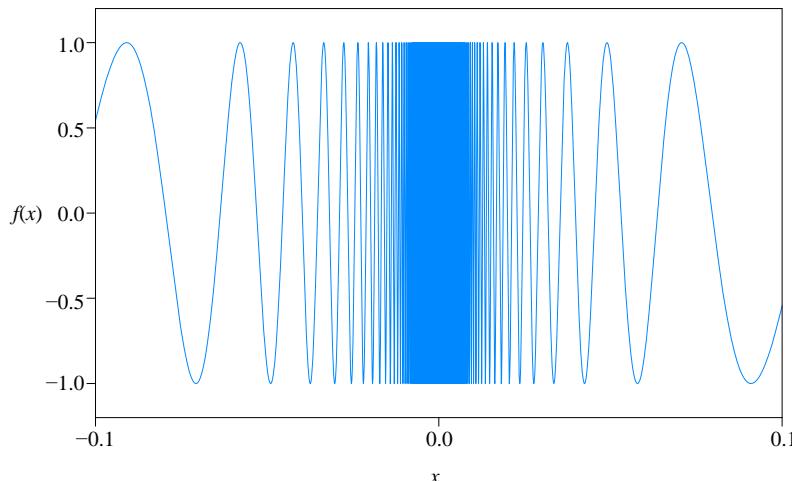


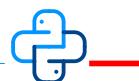
图 3. 特殊函数需要极其细腻的颗粒度

但是，颗粒度过高也不可取，也就是等差数列的公差过小，会增大计算量。这一点在一元函数上并不明显，但是用 `numpy.meshgrid()` 生成网格时，大家就会发现维数灾难 (curse of dimensionality)。

维数灾难是指在高维空间中，数据变得非常稀疏，而且距离变得非常远，使得许多常用的数据分析技术和算法无法有效地处理和分析数据。通俗点讲，假设我们有一个只有两个特征（比如，鸢尾花花萼长度、宽度）的数据集，我们可以很容易地将其可视化成二维平面上的点。但是如果我们有许多特征，比如几百个，那么我们将无法在三维或更高维空间中可视化数据。

此外，保持每个特征的采样数量，当维度增加时会导致数据量急剧增长。比如，单一维度的采样点数为 100，两个特征的网格点数就变成了 10000 (100^2)，三个特征的网格点数就增大到了惊人的 1000000 (100^3)。本书后文还会遇到这个问题。

注意，如果绘图采用对数坐标，建议采用 `numpy.logspace()` 生成数列。



Jupyter 笔记 BK_2_Ch11_1.ipynb 绘制图 1、图 2、图 3。

11.2 阶跃图

再次强调，在绘制线图时，默认散点之间两点顺序连线。这就意味着，任意顺序两点之间的线段是通过线性插值方法得到。

但是，有很多场合，我们需要避免“线性插值”，而采用阶跃方法绘制线图。

`matplotlib.pyplot.plot()` 函数本身可以设定阶跃绘图。此外，`matplotlib.pyplot.step()` 函数是专门绘制阶跃线图的函数。这个函数有三种设置：`'pre'`、`'post'`、`'mid'`。

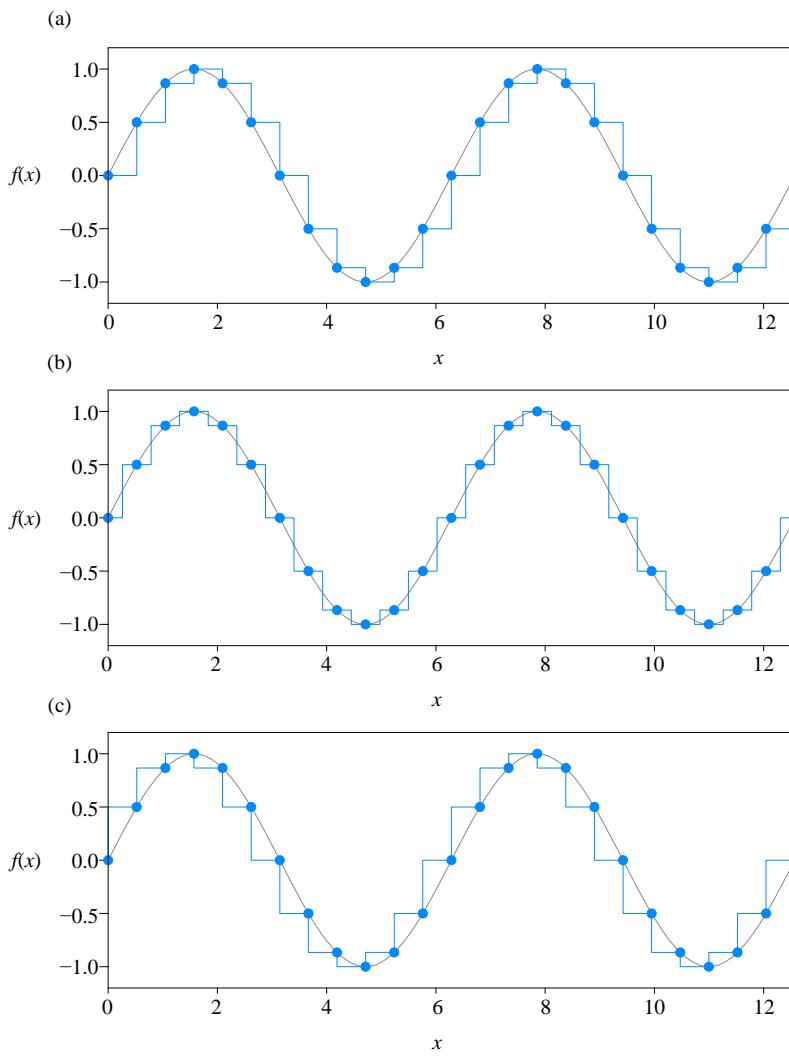


图 4. 三种阶跃

连接两点的插值方法有很多，《数据有道》第 5 章专门介绍。此外，本书后文会介绍贝塞尔曲线 (Bézier curve)。贝塞尔曲线是一种平滑曲线，在计算机图形学、工程和设计领域中广泛应用。

举个例子，贝塞尔二次曲线由三个控制点组成，其中两个控制点定义曲线的端点，第三个控制点定义曲线在端点之间的弯曲。



Jupyter 笔记 BK_2_Ch11_2.ipynb 绘制图 4。

11.3 火柴图

火柴图 (stem plot)，也称火柴梗图、脊柱图，常用来可视化离散数据序列和趋势。火柴图垂直直线所在横轴位置代表样本点的位置，圆点纵轴高度表示样本点的值。

本系列图册中，火柴图常用来可视化数列、离散随机变量概率质量函数 (Probability Mass Function, PMF)，如图 5 所示。

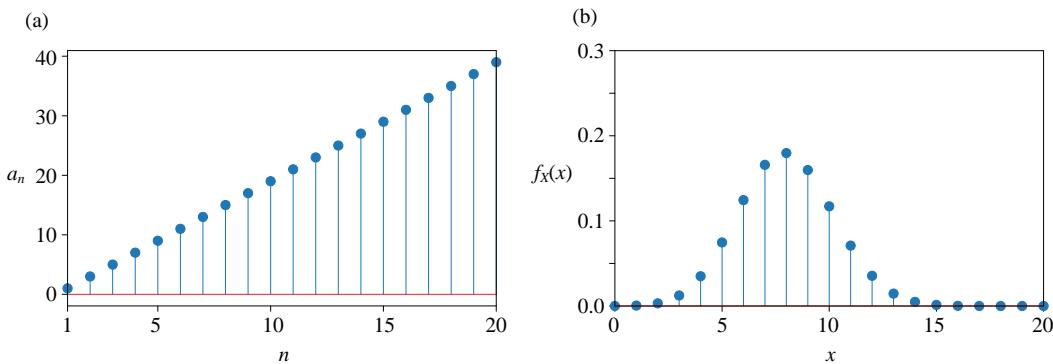


图 5. 火柴图可视化数列、概率质量函数

图 5 的两个子图都可以看成是离散函数。而前文的 $f(x) = \sin(x)$ 则是连续函数。离散函数、连续函数的主要区别在于自变量取值方式不同。离散函数自变量只能取有限或可数无限个值。也就是说，离散函数的函数图像是系列散点。

例如，一个函数 $f(x)$ 表示了投掷一枚骰子后得到点数。因为骰子点数是有限的，所以自变量 x 的取值为 1、2、3、4、5、6 这几个离散值。而连续函数的定义域是一个连续的区间，比如 $(-\infty, \infty)$ 、 $[0, 2]$ 。



Jupyter 笔记 BK_2_Ch11_3.ipynb 绘制图 5。

11.4 参考线

水平线图中，我们经常需要添加水平或竖直参考线。图 6 所示两种不同绘制参考线的方法。

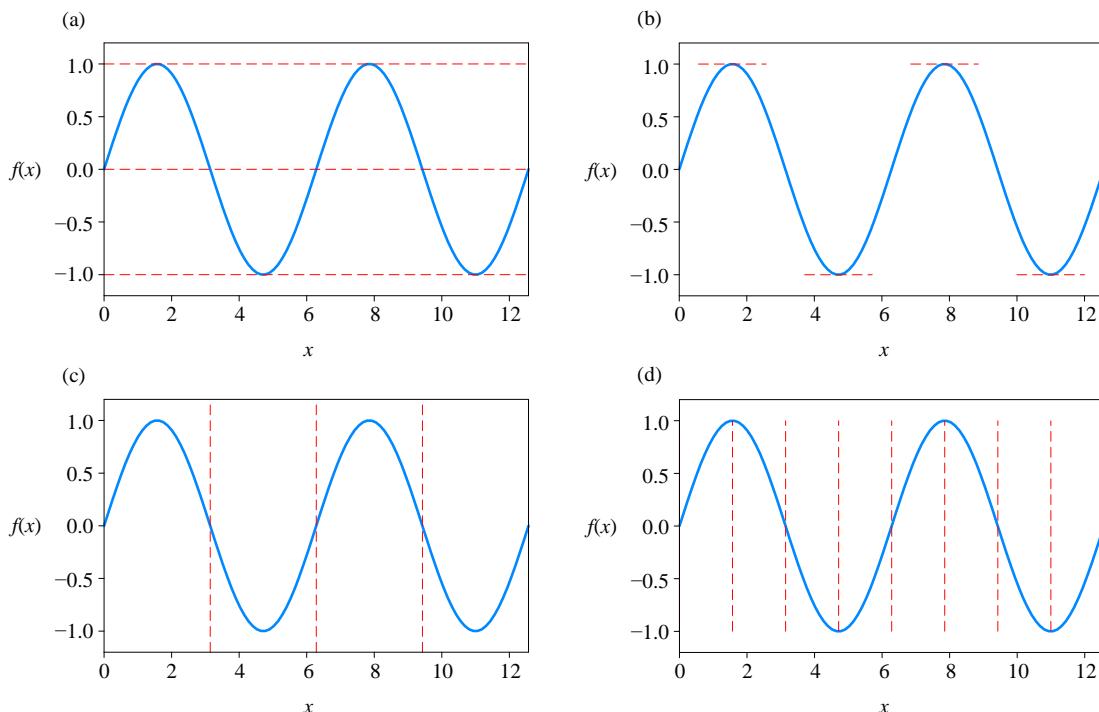


图 6. 两种绘制参考线的方法



Jupyter 笔记 BK_2_Ch11_4.ipynb 绘制图 6。

11.5 使用面具

图 7 所示使用面具 (mask) 分段渲染线图。采用的方法和上一章一致。

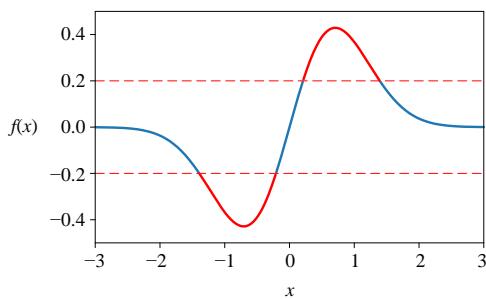


图 7. 分段渲染线图



本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

Jupyter 笔记 BK_2_Ch11_5.ipynb 绘制图 7。

11.6 特殊点线

交点

如图 8 所示，通过寻找 $f_1(x) - f_2(x)$ 的正负号变号的位置，我们可以估计 $f_1(x), f_2(x)$ 的交点。

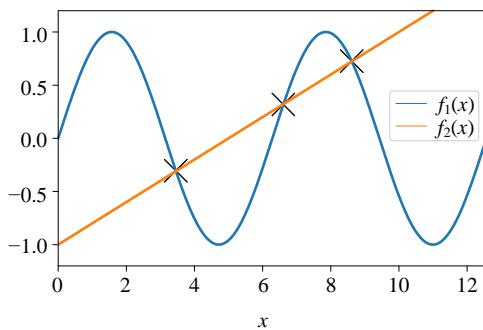
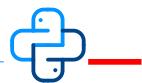


图 8. 可视化交点



Jupyter 笔记 BK_2_Ch11_6.ipynb 绘制图 8。

极大、极小值

`numpy.argmax()`、`numpy.argmin()` 可以寻找数组中的极大、极小值，如图 9 所示。

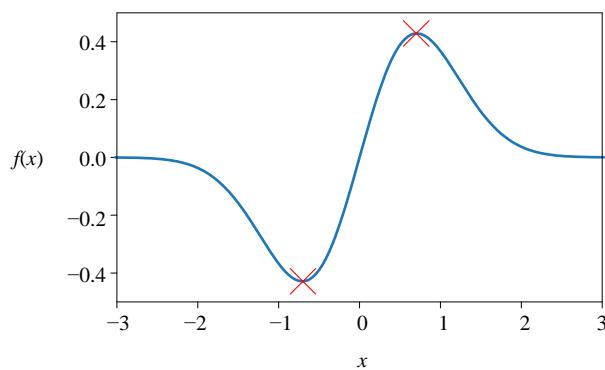
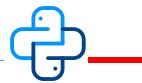


图 9. 可视化极值点



Jupyter 笔记 BK_2_Ch11_7.ipynb 绘制图 9。

11.7 渲染

渲染一组曲线着色

图 12 所示为三种用色谱给一组曲线着色的方法。图 12 (a) 采用 for 循环，分别给每一条曲线着色。

调用 RdYlBu 色谱，用 sigma 数量产生若干连续色号。用 for 循环分别绘制每条曲线，曲线依次调用连续色号。

图 12 (b) 用 LineCollection() 分别渲染每条曲线，并添加色谱条展示 sigma 变化。

图 12 (c) 则用 set_prop_cycle() 修改默认线图颜色。



图 12 中曲线为一元高斯分布的概率密度函数。《统计至简》第 9 章专门讲解一元高斯分布。



Jupyter 笔记 BK_2_Ch11_8.ipynb 绘制图 12 子图。

用三维等高线

图 10 所示为用等高线函数和色谱渲染线图。本书后续将深入讲解这一方法。

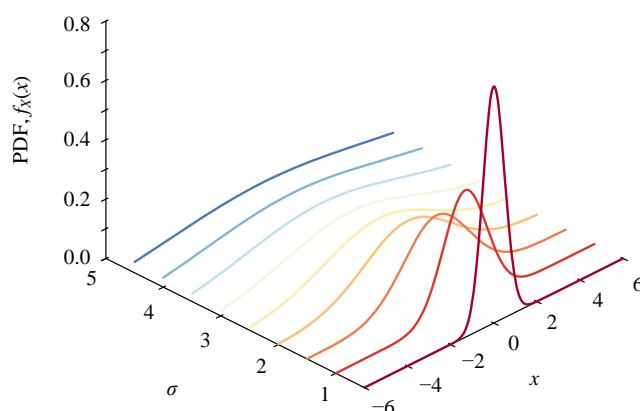


图 10. 用等高线渲染线图

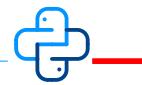
分段渲染曲线

下面，我们用色谱和 LineCollection() 渲染一条曲线的不同分段。

如图 13 (a) 所示，我们先将一条线段打散成一系列线段。然后用 LineCollection()，用 rainbow 色谱分别给每条线段分别着色。图 13 (b) 中有 50 条轨迹。类似地，每条轨迹在水平位置的着色一致。



图 13 两幅图利用随机过程相关数学工具。我们将在《数据有道》第 8 章了解随机过程。



Jupyter 笔记 BK_2_Ch11_9.ipynb 绘制图 13 子图。

绘制网格

图 11 所示为利用 `plot()` 两点连线绘制的正方方格，并采用 rainbow 色谱分段着色渲染。图 14 所示为在此基础上可视化线性、非线性变换。

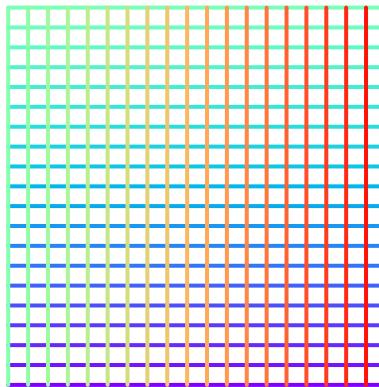


图 11. 用 `plot()` 绘制的网格，利用 rainbow 色谱渲染

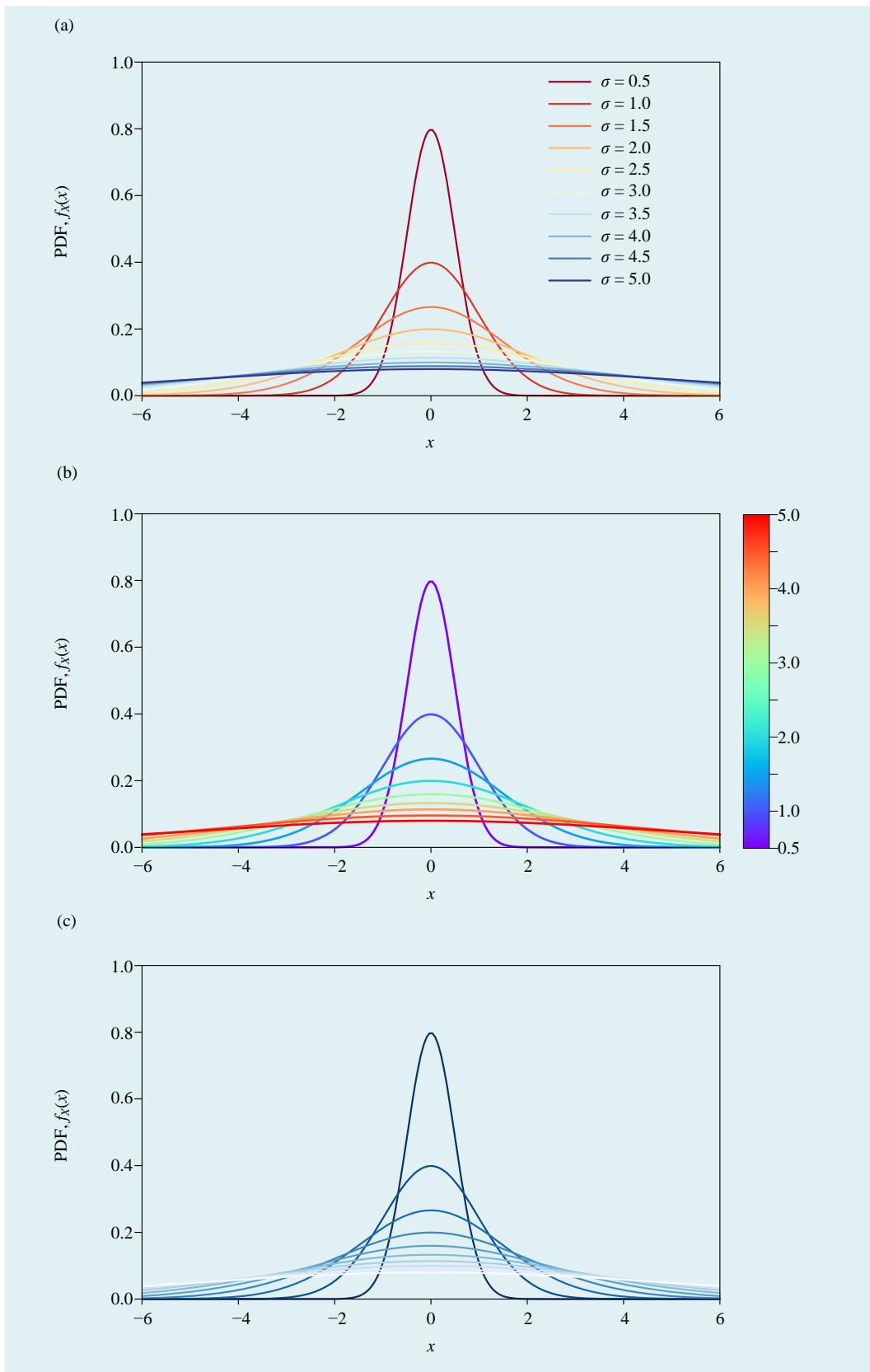


图 12. 用色谱渲染曲线

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

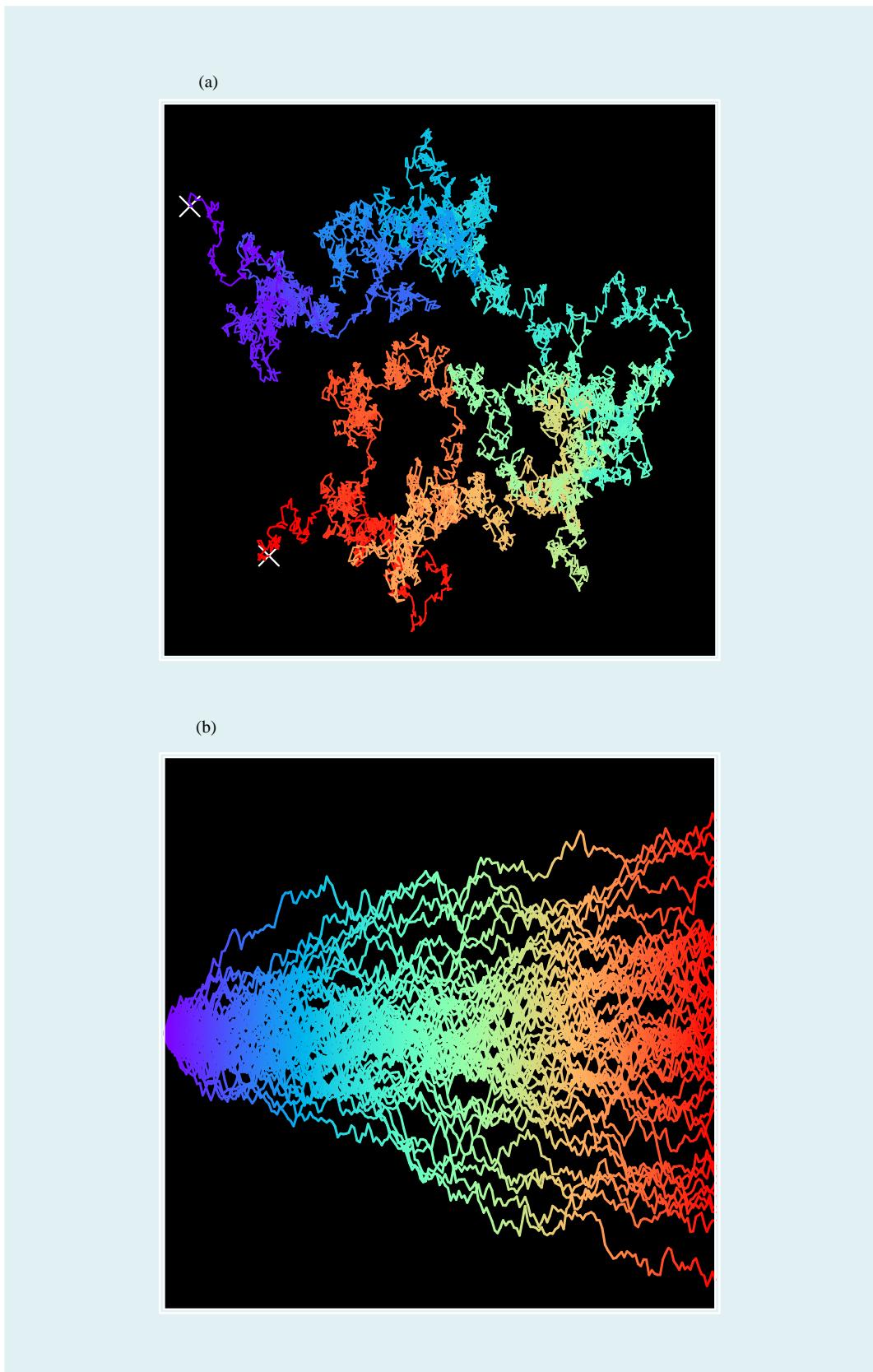


图 13. 用色谱给一条曲线不同线段顺序着色

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

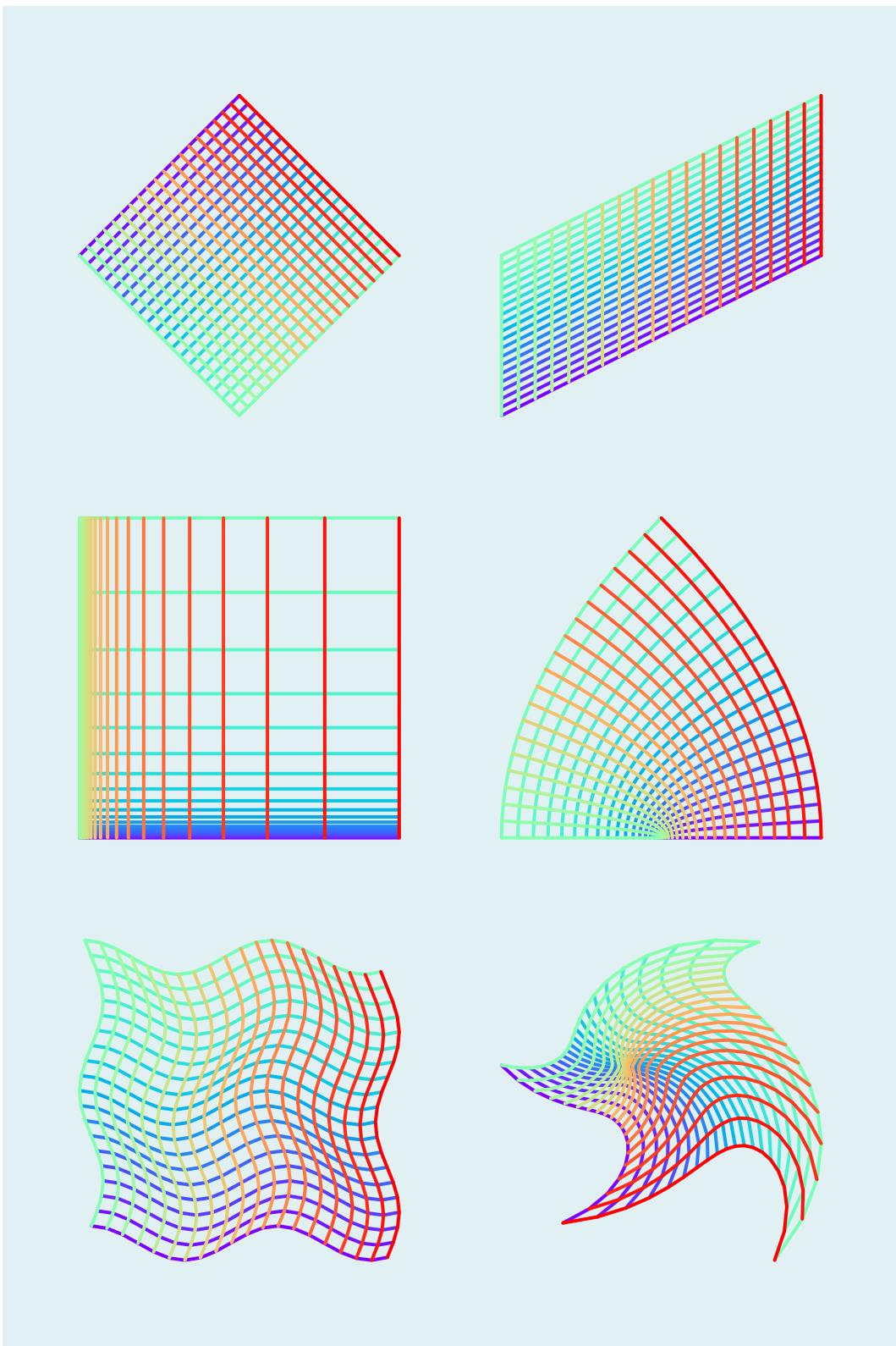


图 14. 可视化线性、非线性变换

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

2D Filled Area

12 平面填充

用面积可视化特定特征



大理石中我看到了天使，我拿起刻刀不停雕刻，直到还它自由。

I saw the angel in the marble and carved until I set him free.

—— 米开朗琪罗 (Michelangelo) | 文艺复兴三杰之一 | 1475 ~ 1564

- ◀ `matplotlib.patches.Rectangle()` 是 Matplotlib 中的一个图形对象，用于绘制矩形形状
 - ◀ `matplotlib.pyplot.axhspan()` 函数用于在水平方向创建一个跨越指定 y 值范围的色块
 - ◀ `matplotlib.pyplot.axvspan()` 函数用于在垂直方向创建一个跨越指定 x 值范围的色块
 - ◀ `matplotlib.pyplot.fill()` 函数用于绘制多边形，并在其中填充颜色，创建一个封闭区域的图形效果
 - ◀ `matplotlib.pyplot.fill_between()` 函数用于在两条曲线之间填充颜色，创建一个区域的图形效果
 - ◀ `matplotlib.pyplot.fill_betweenx()` 函数用于在两条垂直于 x 轴的水平线之间填充颜色，创建一个区域的图形效果
 - ◀ `numpy.linspace()` 在指定的间隔内，返回固定步长的数据
- 

13.1 沿轴填充

沿横轴填充

`fill_between()` 是 `matplotlib.pyplot` 库中的一个函数，用于绘制两个曲线之间的填充区域。

`fill_between(x, y1, y2, ...)` 函数可以接受两个数组 x 和 y_1 ，以及另外一个数组 y_2 ，它们都是相同长度的。这个函数会将 y_1 和 y_2 之间的区域填充，并在 x 上绘制。

图 1 所示三个例子展示曲线和水平线之间沿横轴填充。`fill_between()` 可以使用参数 `where` 来指定 y_1 和 y_2 之间的填充区域，可以使用参数 `facecolor` 来指定填充颜色，使用参数 `alpha` 来指定填充区域的透明度。

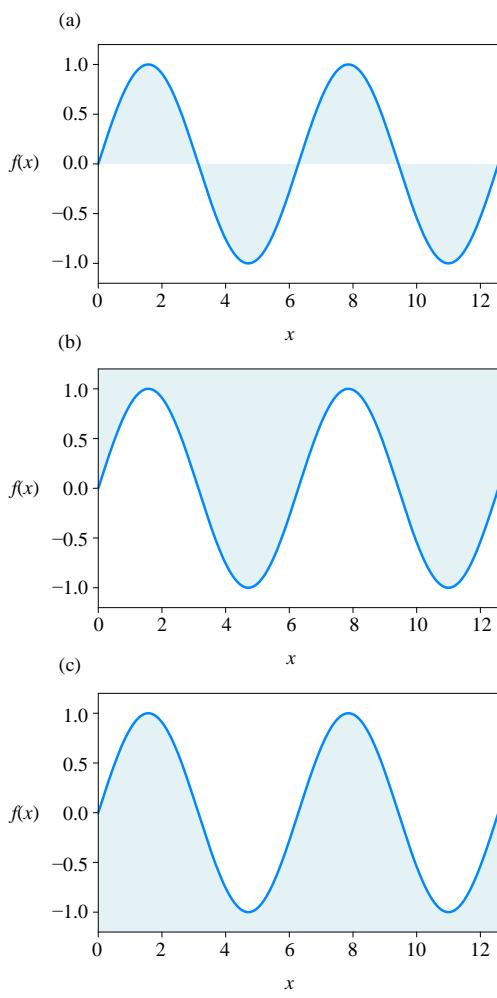


图 1. 曲线和水平线之间沿横轴填充

如图 2 所示，通过设置条件，我们还可以给满足不同条件的区域填充不同颜色。图 3 还给出两个例子，展示两条曲线之间沿横轴填充。

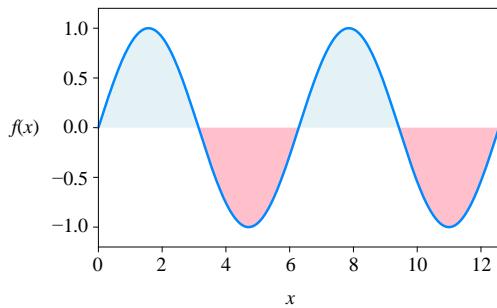


图 2. 曲线和水平线之间沿横轴填充，不同颜色

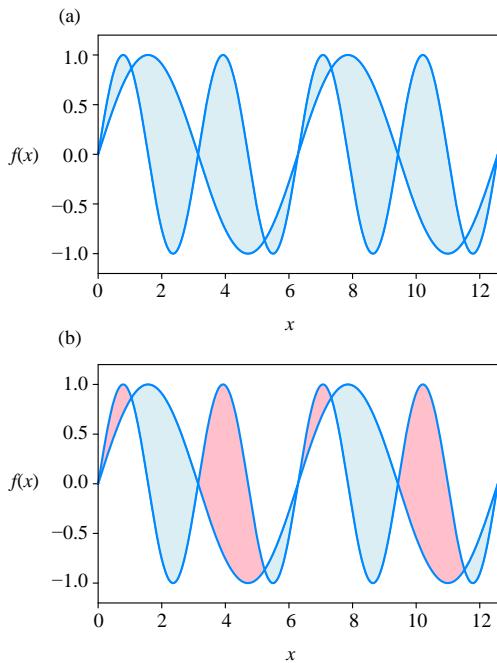


图 3. 两条曲线之间沿横轴填充



Jupyter 笔记 BK_2_Ch13_1.ipynb 绘制图 1、图 2、图 3。

沿纵轴填充

`fill_betweenx` 则可以用来绘制两个曲线在 y 轴方向之间的填充区域。`fill_betweenx(y, x1, x2, ...)` 函数接受两个数组 `y` 和 `x1`, 以及另外一个数组 `x2`, 它们都是相同长度的。这个函数会将 `x1` 和 `x2` 之间的区域填充，并在 `y` 上绘制。**Error! Reference source not found.** 所示为沿纵轴方向填充的 6 个例子。

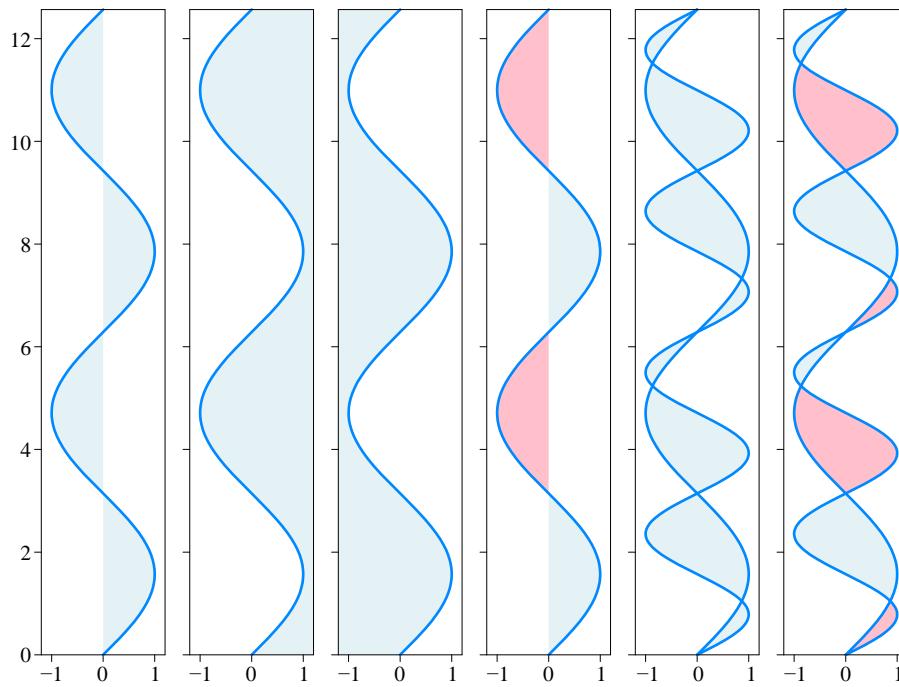
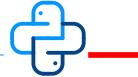


图 4. 沿纵轴方向填充



Jupyter 笔记 BK_2_Ch13_2.ipynb 绘制 Error! Reference source not found.。

填充阴影线

Matplotlib 中我们还可以用 hatch 给各种填充增加阴影线。hatch 是 matplotlib 库中的一个属性，用于给某些图形元素添加填充图案。要使用 hatch 属性，需要将它设置为一个字符串，该字符串描述了所需的填充图案类型。matplotlib 库中提供了多种不同的填充图案类型，下面给出几个例子：

```
'/'：斜杠填充  
\'：反斜杠填充  
'.'：点状填充  
'o'：圆形填充  
'-'：横向线性填充  
'+'：十字线填充  
'x'：斜十字线填充  
'|'：纵向线性填充
```



BK_2_Ch13_3.ipynb 给出几种常见的阴影线，请大家自行学习。也请大家在前两个 Jupyter 笔记使用不同的填充阴影线。

13.2 参考填充色块

本书前文介绍过如何绘制水平、竖直参考线，类似地，我们也可以绘制参考填充色块。`axhspan` 是 `matplotlib` 库中的一个函数，用于在一个子图中绘制一个水平的矩形。

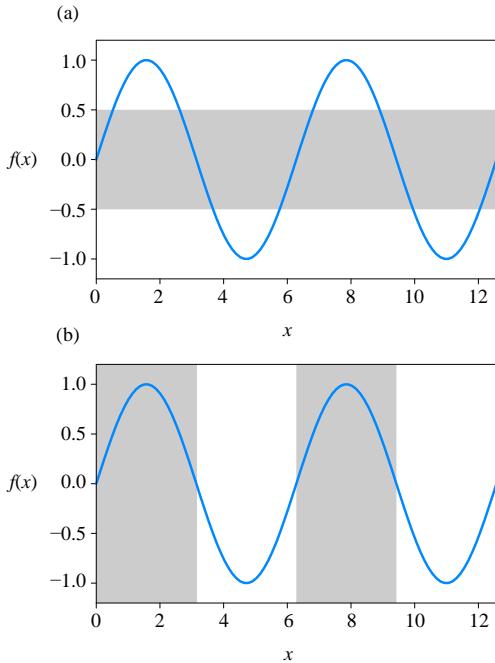


图 5. 参考填充色块

这个函数通常用于强调某个区域的范围或表示一个特定的数据区间，如图 5 (a) 所示。

`axhspan` 函数接受四个参数：`ymin`、`ymax`、`xmin` 和 `xmax`。其中，`ymin` 和 `ymax` 表示矩形的纵向范围，`xmin` 和 `xmax` 表示矩形的横向范围。类似地，`axvspan` 可以绘制竖直方向参考填充色块，如图 5 (b) 所示。

13.3 封闭填充

`fill` 是 `matplotlib` 库中常用的一个函数，用于绘制填充区域。使用 `fill` 函数可以将一个多边形区域填充成指定的颜色。

`fill` 函数接受两个参数：`x` 坐标数组和 `y` 坐标数组，用于指定要填充的多边形区域的顶点坐标。`x` 和 `y` 的长度必须相同，且每个元素都对应一个多边形的顶点。图 6 给出的例子还用到了旋转，《矩阵力量》会介绍如何利用线性代数工具完成旋转操作。

图 7 所示为利用正方形可视化最小二乘回归原理。

`add_patch` 是 `matplotlib` 库中 `Axes` 对象的一个方法，用于向一个子图中添加一个图形元素。这个方法可以添加多种不同类型的图形元素，例如矩形、多边形、圆形、椭圆、箭头等等。在使用 `add_patch` 方法前，需要先创建一个对应的图形元素对象，例如 `Circle`、`Rectangle`、`Polygon`、`Ellipse`、`Arrow` 等。然后，可以使用 `add_patch` 方法将这个对象添加到指定的子图中。在添加完成

后，可以使用 `set_*`（比如 `set_edgecolor`、`set_linewidth`）方法或属性来设置图形元素的属性，例如填充颜色、边框颜色、边框宽度等。本书后续将专门介绍相关用法。

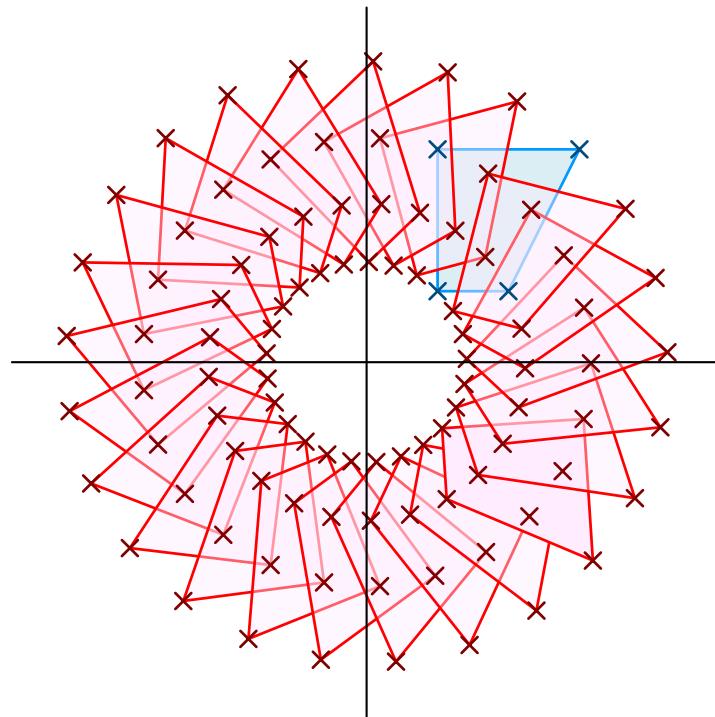


图 6. 用 fill 首尾连接封闭填充

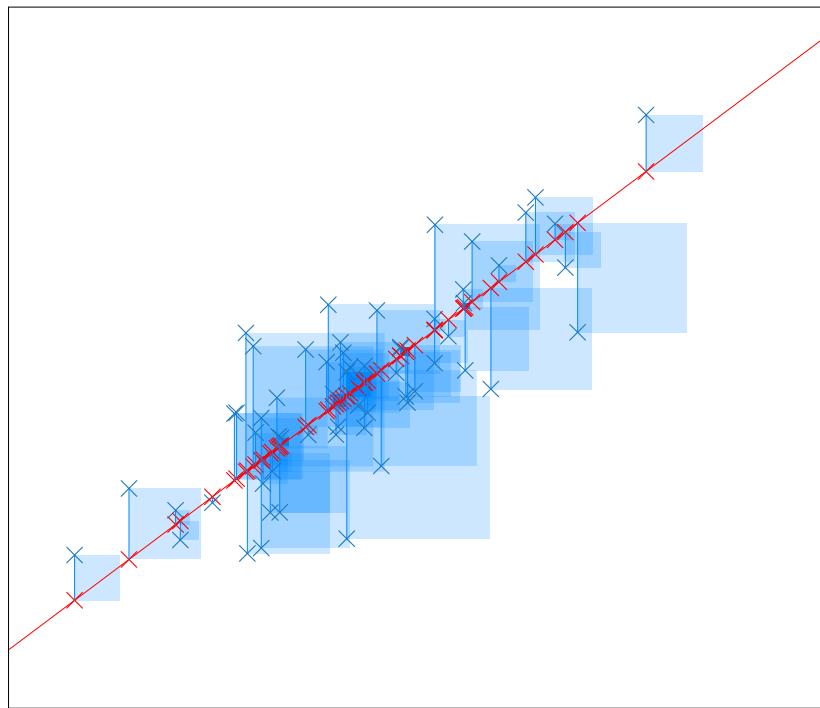
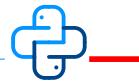


图 7. 添加几何元素



Jupyter 笔记 BK_2_Ch13_5.ipynb 绘制图 6。Jupyter 笔记 BK_2_Ch13_6.ipynb 绘制图 7。

13.4 堆叠面积图

`stackplot` 函数是 `matplotlib` 库中用于绘制堆叠面积图的函数。堆叠面积图可以用于展示多个数据序列之间的比例关系，常用于显示时间序列数据的变化趋势。此外，更常用的是直接用 `dataframe` 绘制堆叠面积图，如图 8 所示。

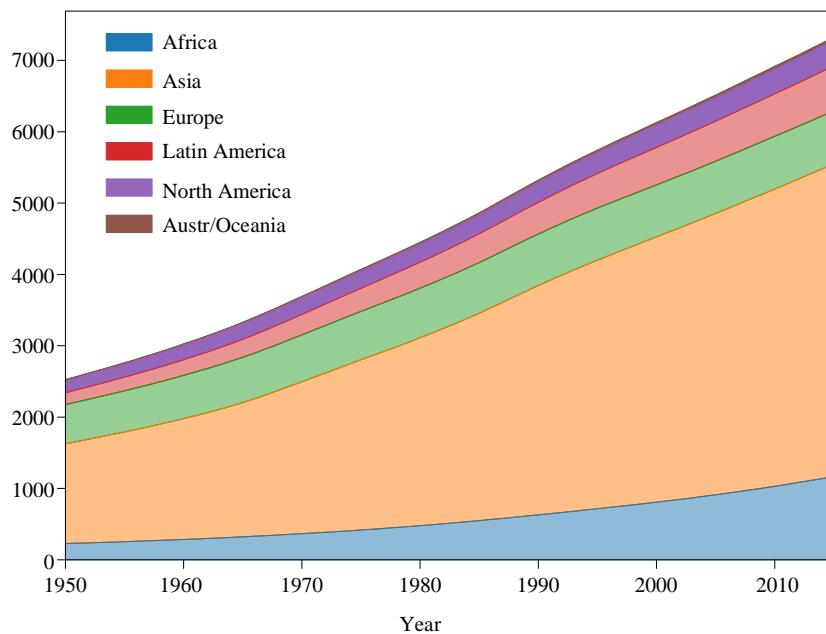
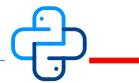


图 8. 用堆叠面积图绘制人口变化



Jupyter 笔记 BK_2_Ch13_7.ipynb 绘制图 8。

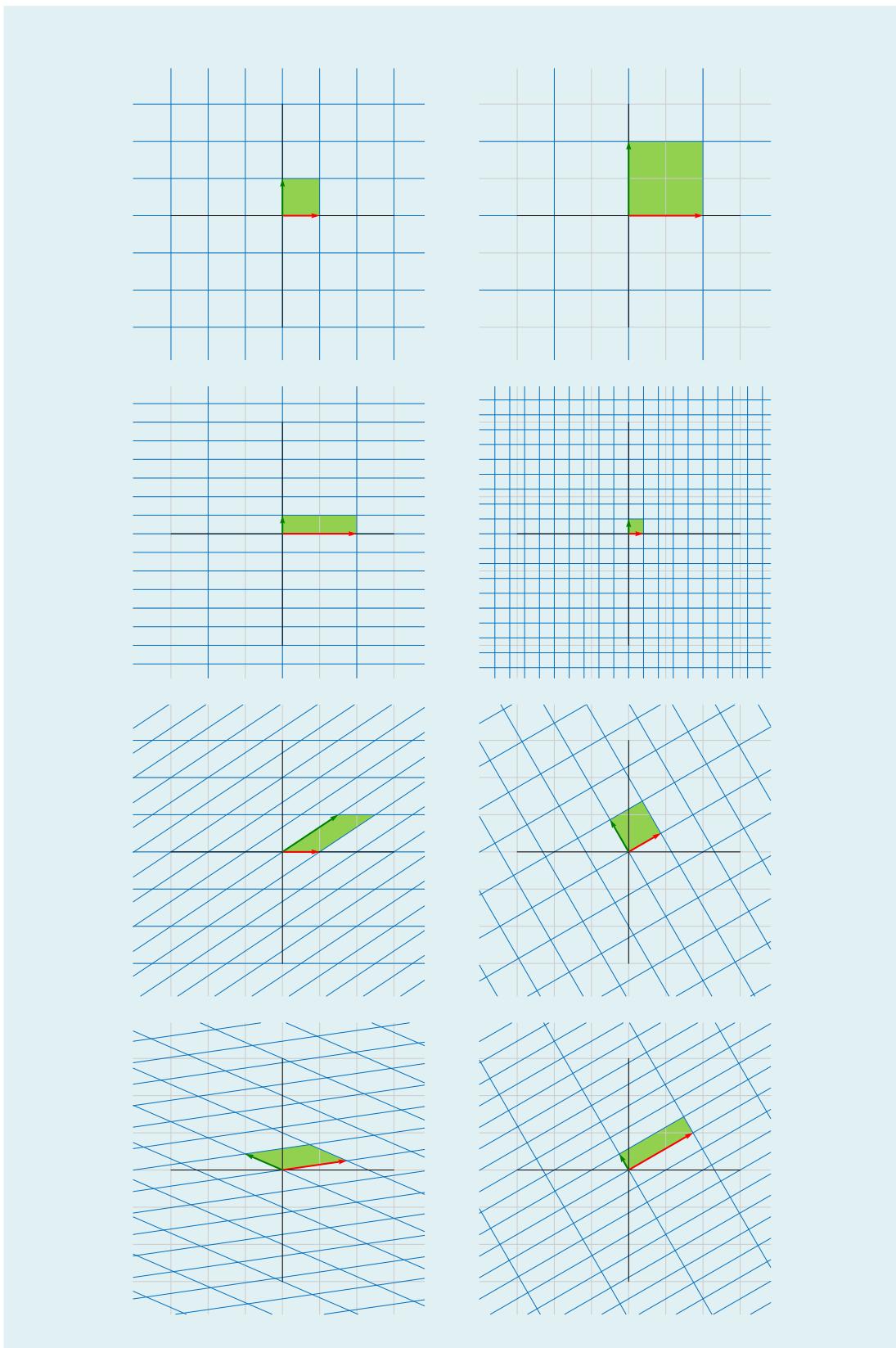


图 9. 几何变换

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

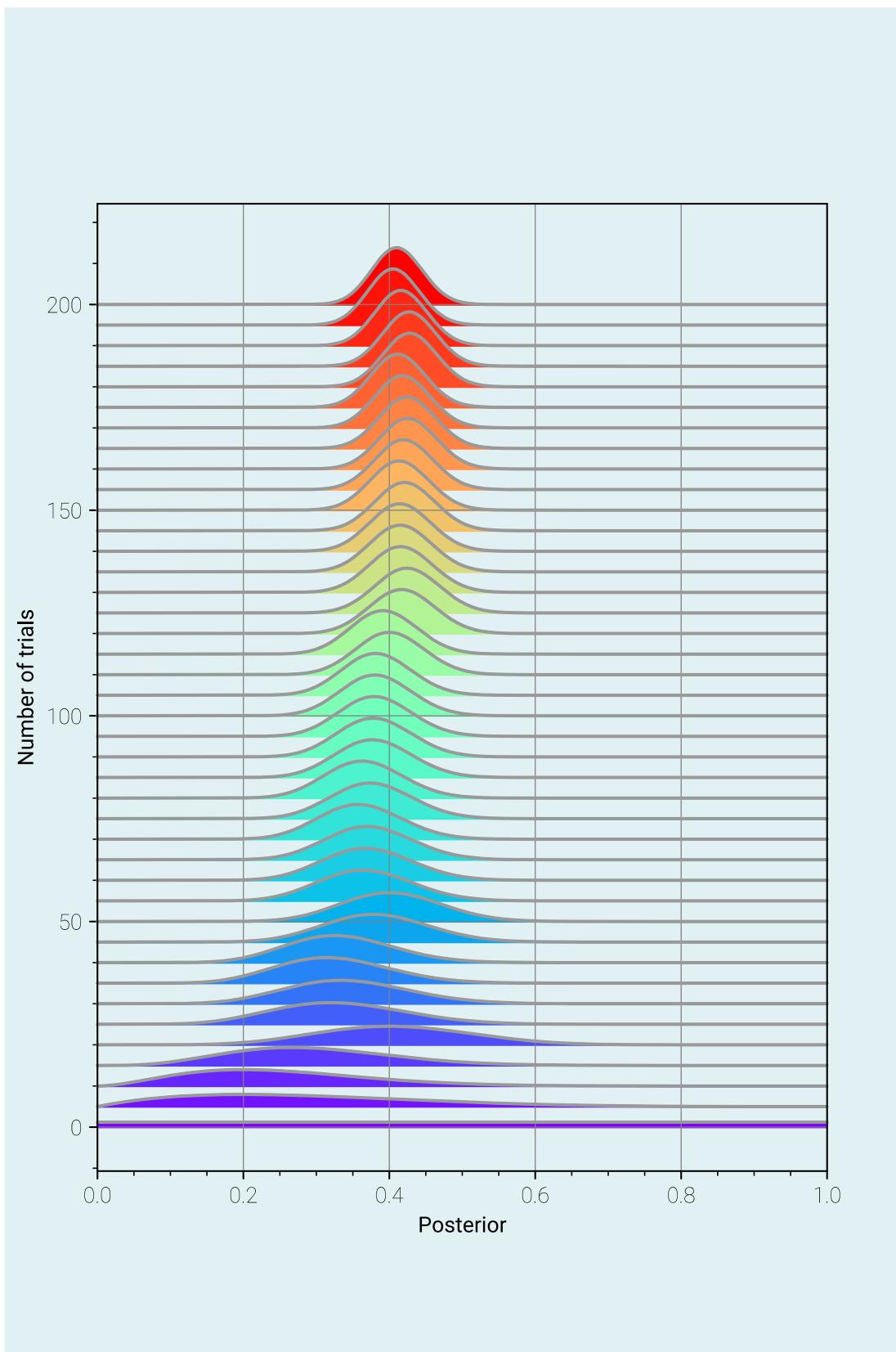


图 10. 复刻山脊图

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

13

Polar Plots

极坐标绘图

以距离和与夹角描述点位置的坐标系统



我们只是一颗普通恒星的小行星上的高级猴子品种。但我们可以理解宇宙，这让我们变得非常特别。

We are just an advanced breed of monkeys on a minor planet of a very average star. But we can understand the Universe. That makes us something very special.

—— 史蒂芬·霍金 (Stephen Hawking) | 英国理论物理学家、宇宙学家 | 1942 ~ 2018



- ◀ `matplotlib.pyplot.bar()` 绘制柱状图
- ◀ `matplotlib.pyplot.cm` 是 Matplotlib 库中的一个模块，用于处理和管理色谱
- ◀ `matplotlib.pyplot.fill()` 绘制封闭填充图形
- ◀ `numpy.linspace()` 在指定的间隔内，返回固定步长的数据
- ◀ `numpy.random.rand()` 返回一个介于 0 和 1 之间的服从均匀分布随机数

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

13.1 线图

相信通过本书前文的学习，大家对极坐标这个概念已经并不陌生。极坐标系统是描述平面上点位置的一种方式，它使用了两个参数：极径和极角。在极坐标系统中，点的位置由它与原点之间的距离和从某个参考方向（通常是 x 轴）逆时针旋转的角度决定。极径表示点到原点的距离。极角表示点到原点的连线与参考方向的夹角，它的单位通常是弧度制。

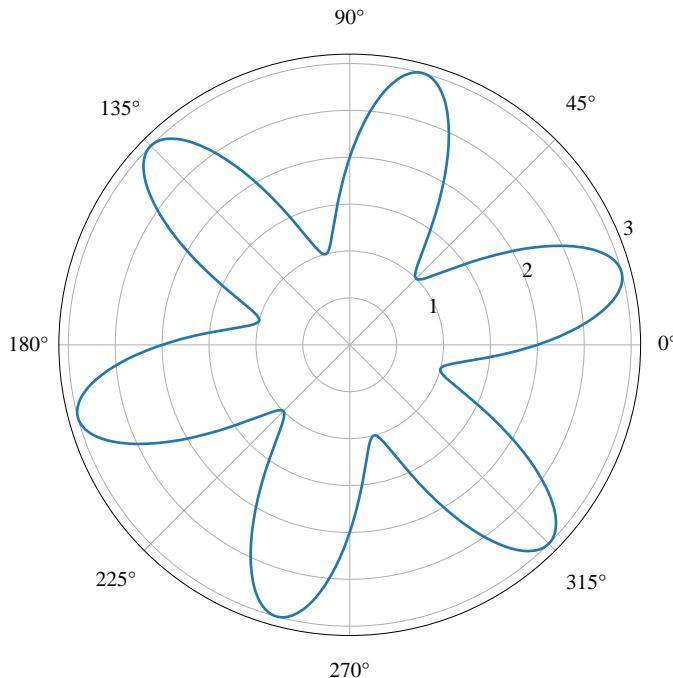
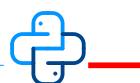


图 1. 极坐标线图

`projection='polar'` 是 Matplotlib 中的一个参数，它可以用来绘制极坐标图。

图 1 所示为极坐标下绘制的线图。我们首先使用 Numpy 生成了极角、极轴两个数组然后，创建极坐标子图，并使用 `ax.plot()` 方法绘制极坐标图。如果想要设置极坐标图中的半径范围，需要使用 `ax.set_rlim()` 方法。**Error! Reference source not found.** 绘制更多的极坐标线图。



Jupyter 笔记 BK_2_Ch13_1.ipynb 绘制图 1。

13.2 散点图

要绘制极坐标下的散点图，需要在创建 Axes 对象时将参数 `projection` 设置为 '`polar`'。然后使用 `scatter()` 函数来添加散点。如图 2 所示，我们也可以指定散点的大小、颜色。

`ax.set_rorigin()` 用来改变极轴原点坐标。`ax.set_rlabel_position()` 用于指定 r-label 的位置相对于轴线的偏移量。

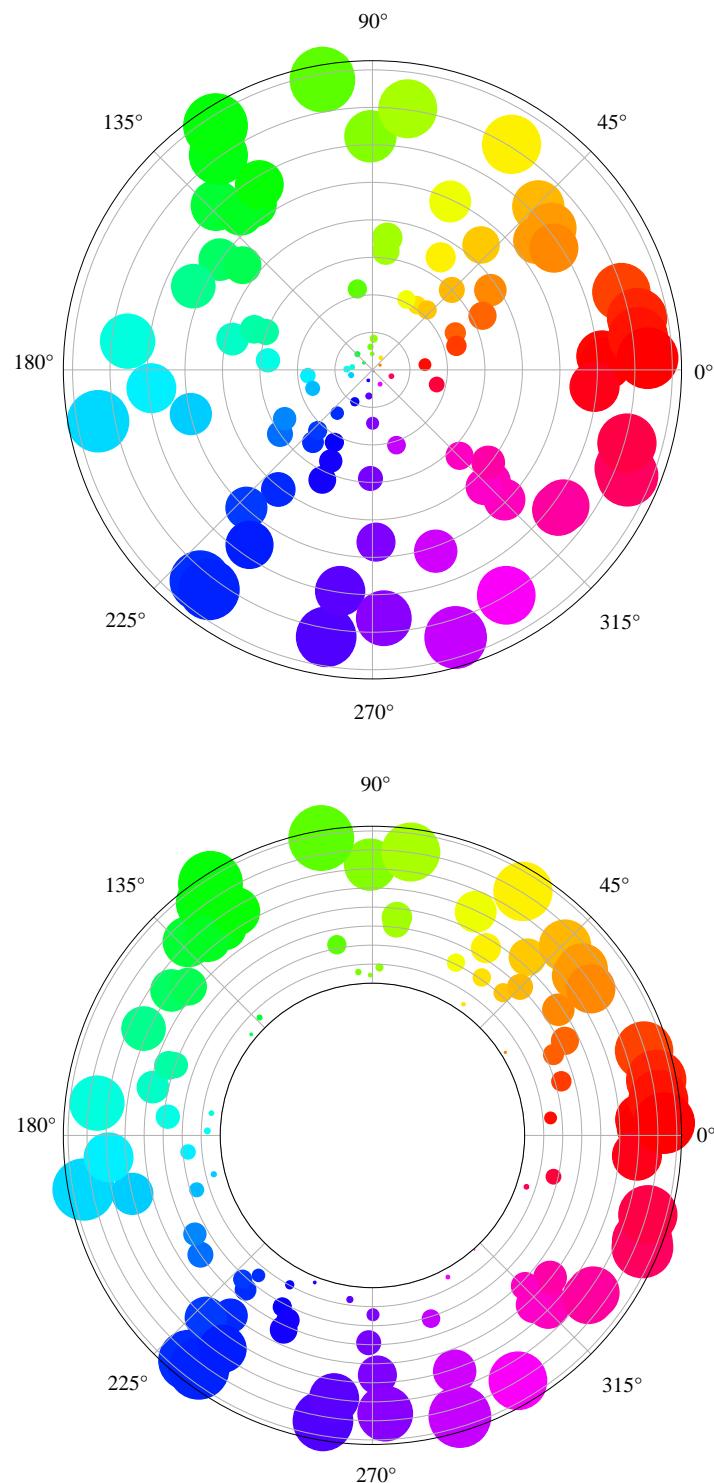
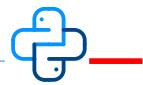


图 2. 极坐标散点图



Jupyter 笔记 BK_2_Ch13_1.ipynb 绘制图 2。

13.3 柱状图

在极坐标中，我们还可以绘制如图 3 所示的柱状图，BK_2_Ch13_4.ipynb 为对应的代码文件。很遗憾，目前 matplotlib 中还没有方便绘制雷达图的工具。

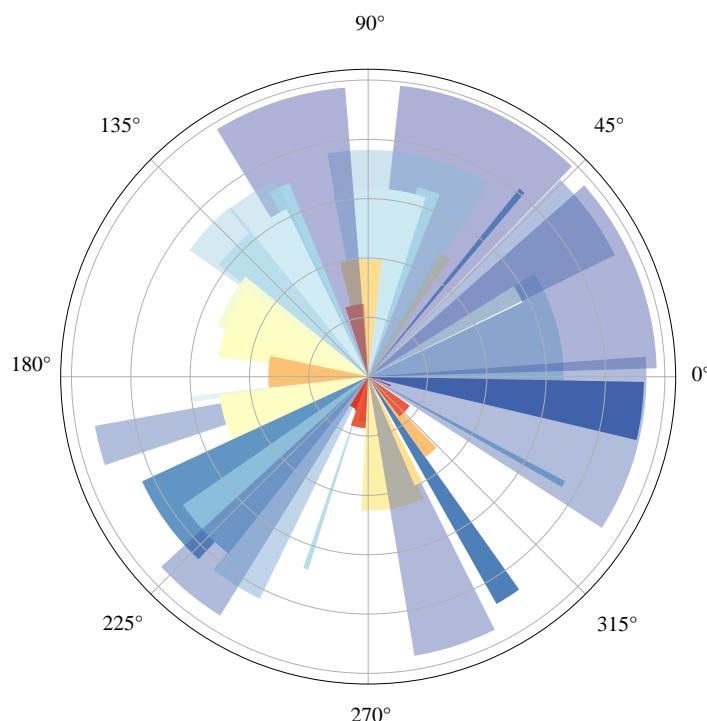


图 3. 极坐标柱状图



想要画雷达图的话，可以参考 BK_2_Ch13_4.ipynb。

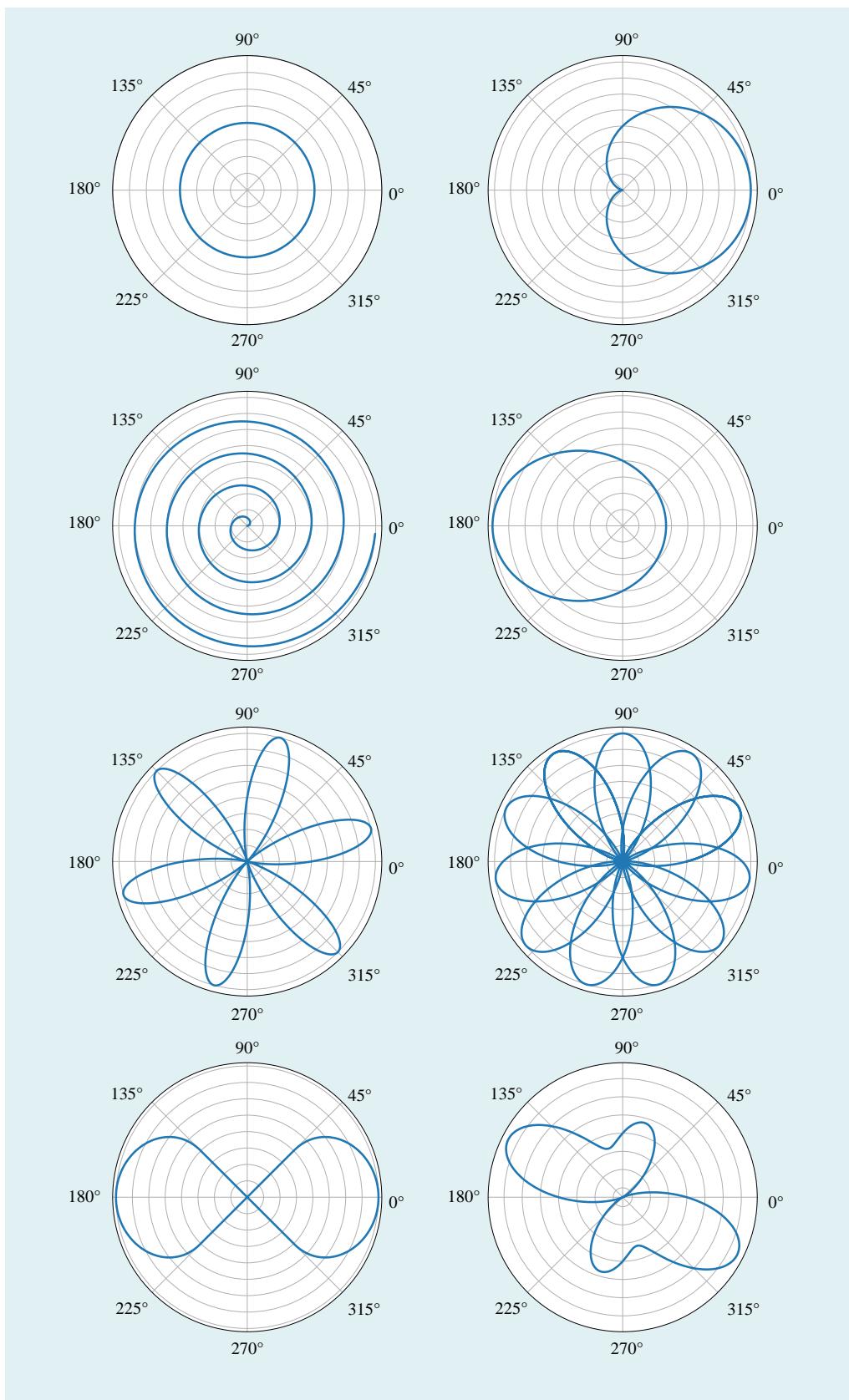


图 4. 更多极坐标线图

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

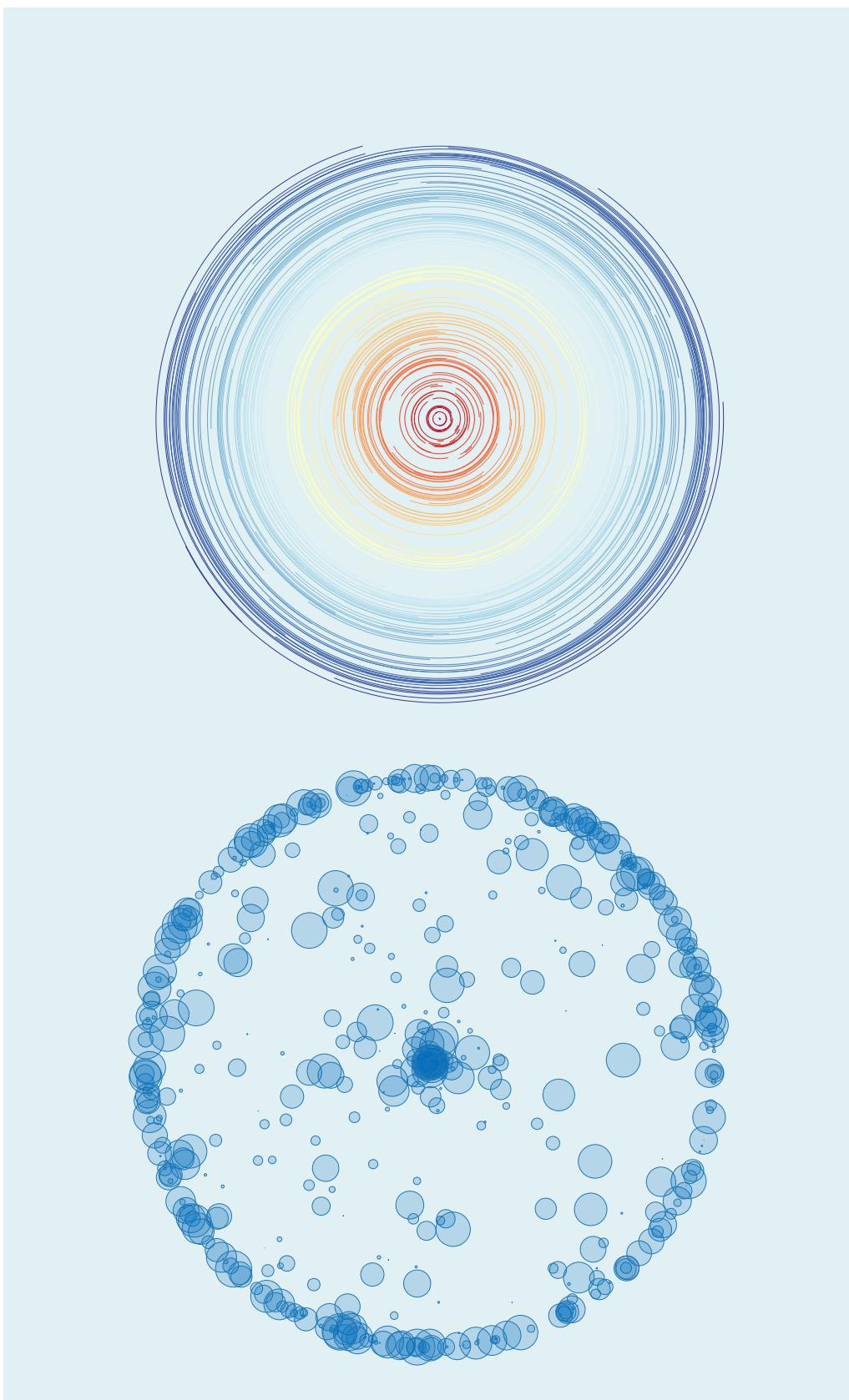


图 5. 基于随机数发生器的极坐标创意编程，第 1 组

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

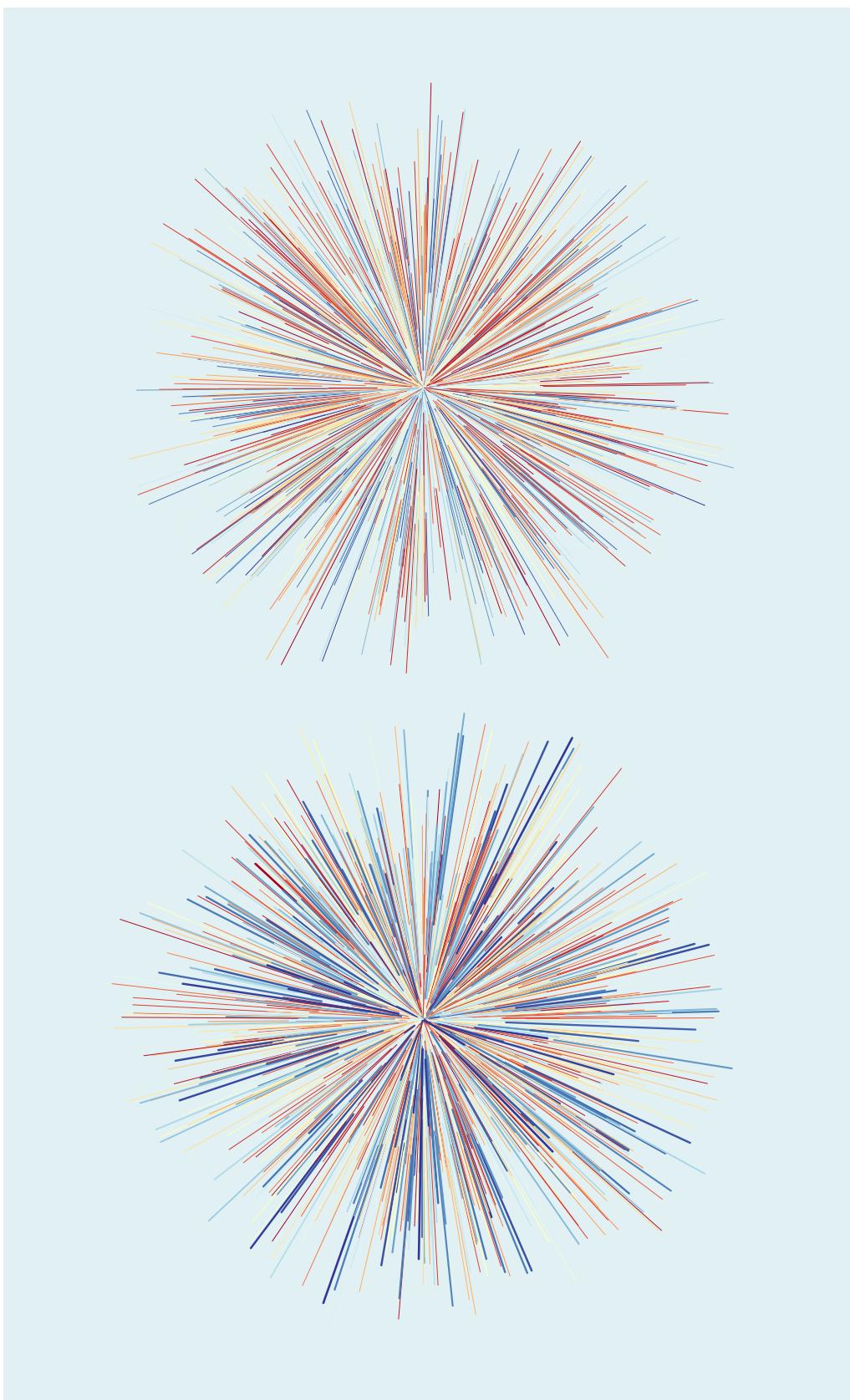


图 6. 基于随机数发生器的极坐标创意编程，第 2 组

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

14

Types of Mesh Grids

各种网格

等高线、三维曲面不可或缺的数据结构



艺术家用头脑绘画，而不是双手。

A man paints with his brains and not with his hands.

—— 米开朗琪罗 (Michelangelo) | 文艺复兴三杰之一 | 1475 ~ 1564



- ◀ `matplotlib.pyplot.plot_wireframe()` 绘制线框图
- ◀ `matplotlib.pyplot.scatter()` 绘制散点图
- ◀ `matplotlib.pyplot.triplot()` 绘制三角剖分图
- ◀ `matplotlib.tri.Triangulation()` 用于创建三角剖分对象
- ◀ `numpy.column_stack()` 将两个矩阵按列合并
- ◀ `numpy.concatenate()` 将多个数组进行连接
- ◀ `numpy.cos()` 计算余弦值
- ◀ `numpy.linspace()` 在指定的间隔内, 返回固定步长的数据
- ◀ `numpy.meshgrid()` 产生网格化数据
- ◀ `numpy.ones_like()` 用来生成和输入矩阵形状相同的全 1 矩阵
- ◀ `numpy.sin()` 计算正弦值

14.1 “方方正正” 网格

相信大家已经对 `numpy.meshgrid()` 函数并不陌生。NumPy 中的 `meshgrid()` 函数用于生成网格状的坐标点矩阵，其作用是将两个或多个一维数组转换为多维数组。具体来说，`meshgrid()` 函数接受两个或多个一维数组作为参数，返回多维坐标矩阵。图 1 所示为生成二维网络状坐标原理。

图 2 所示为从三维空间视角看二维网络状散点。

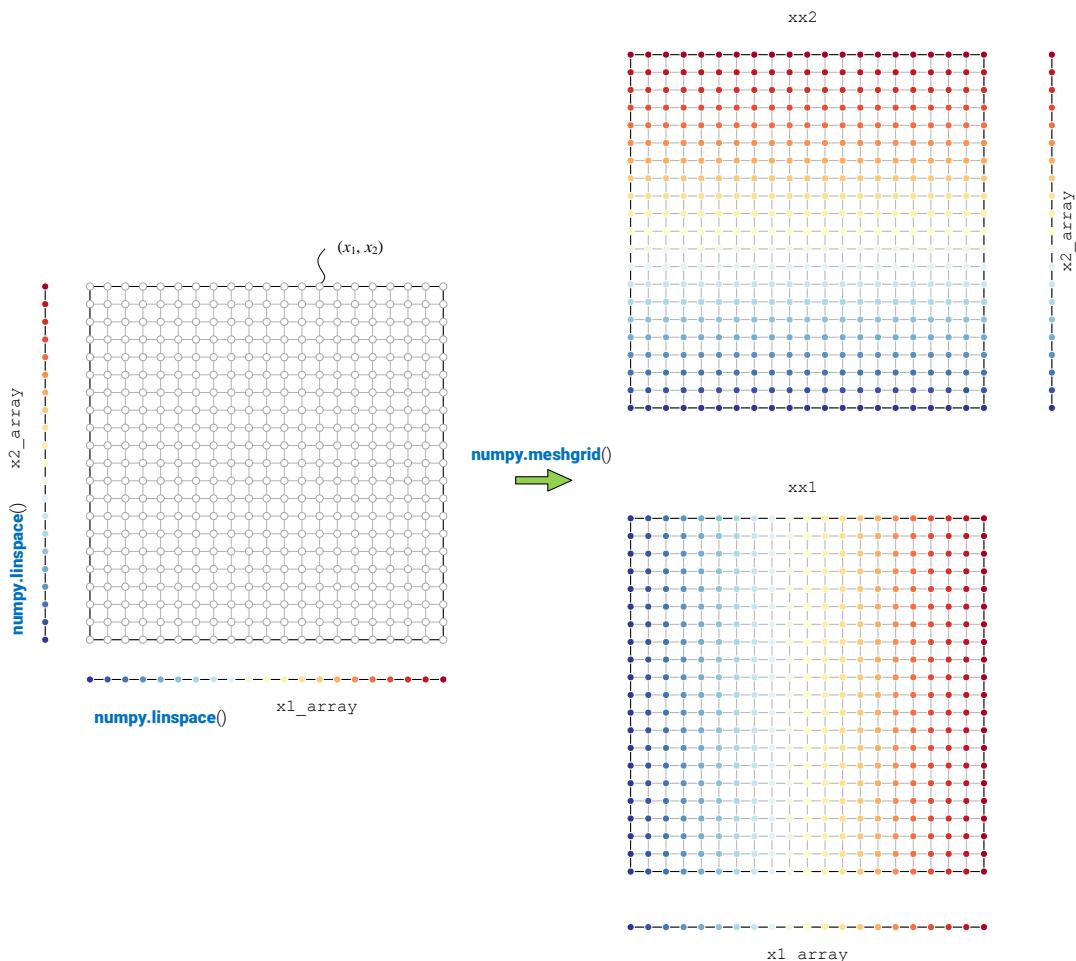


图 1. 用 `numpy.meshgrid()` 生成二维网络数据

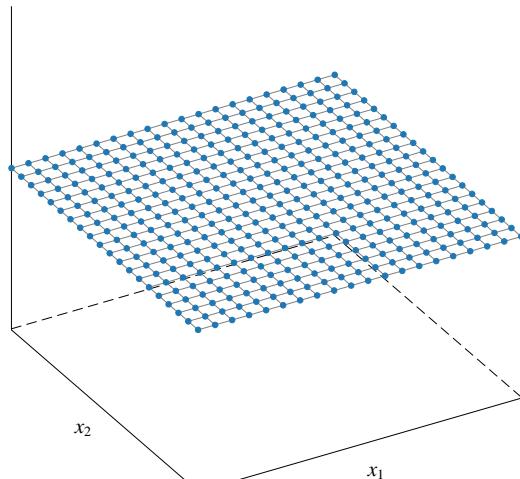


图 2. 三维空间看二维网络状坐标

网格状坐标的用途

`meshgrid()` 产生的二维网络状坐标通常用于绘制网格曲面、等高线等场景。图 3 所示为用二维网络状坐标可视化二元函数 $f(x_1, x_2)$ 。

本书后续，大家会看到我们用网格状坐标绘制等高线。

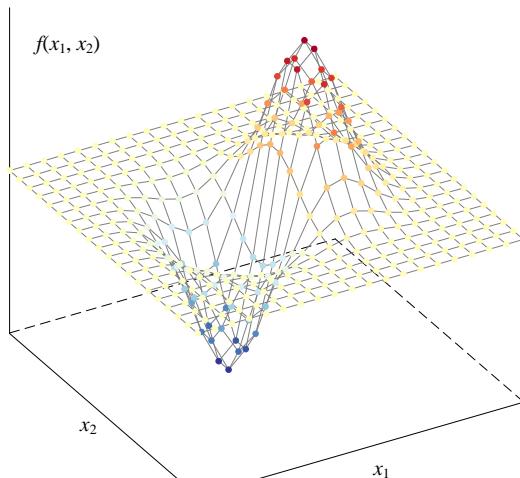


图 3. 二维网络状坐标可视化二元函数

颗粒度

类似平面线图，利用网络状坐标可视化数据时，也会遇到颗粒度的问题。如图 4、图 5 所示，颗粒度过低、过高都会导致可视化效果不理想。本书后文将分别从等高线、网格曲面等几个角度继续颗粒度这个话题。

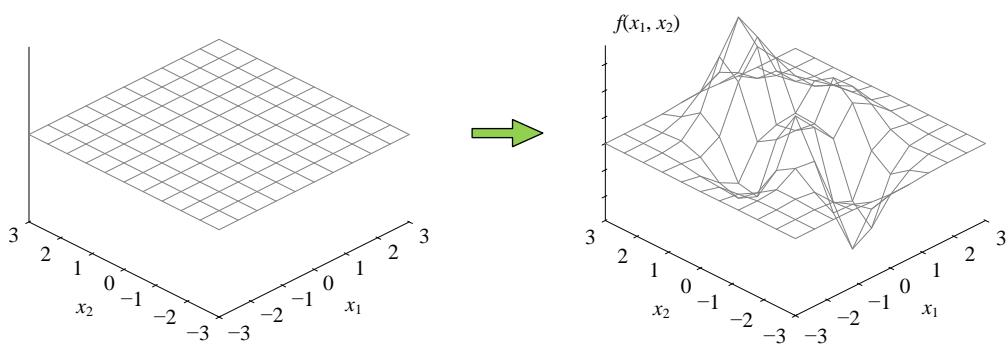


图 4. 颗粒度低

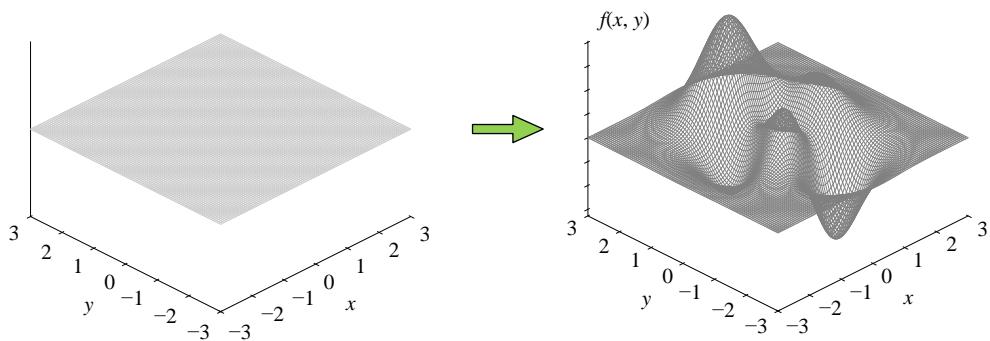


图 5. 颗粒度过高

使用面具

类似前文线图，对于网格我们也可以使用面具（mask）。图 6 所示的两个例子为满足特定条件的部分网格数据。

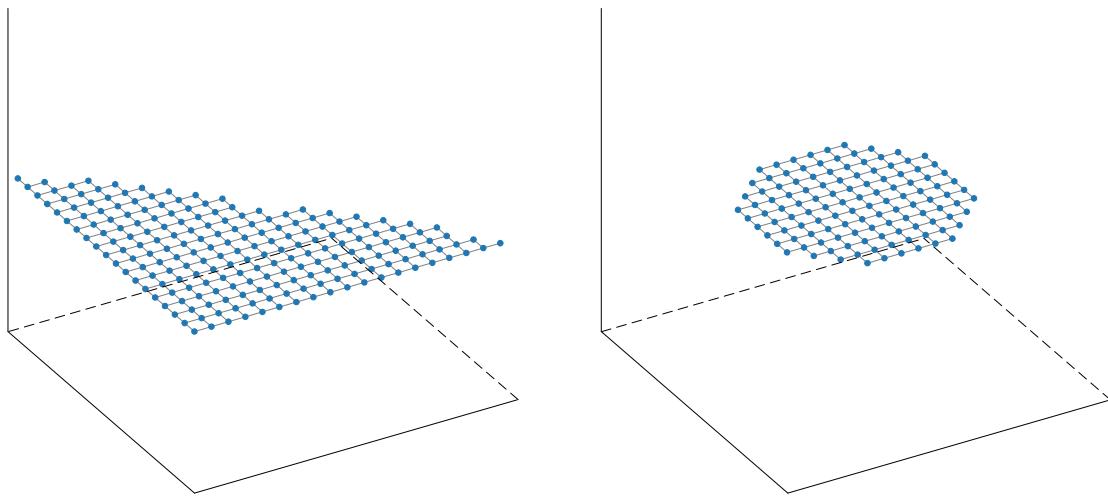


图 6. 使用面具

三维网格

此外，大家对图 7 所示三维网格也应该不陌生。我们在色彩模型中用过它。此外，本书后文还会继续用三维网格散点提供更为丰富的可视化方案。

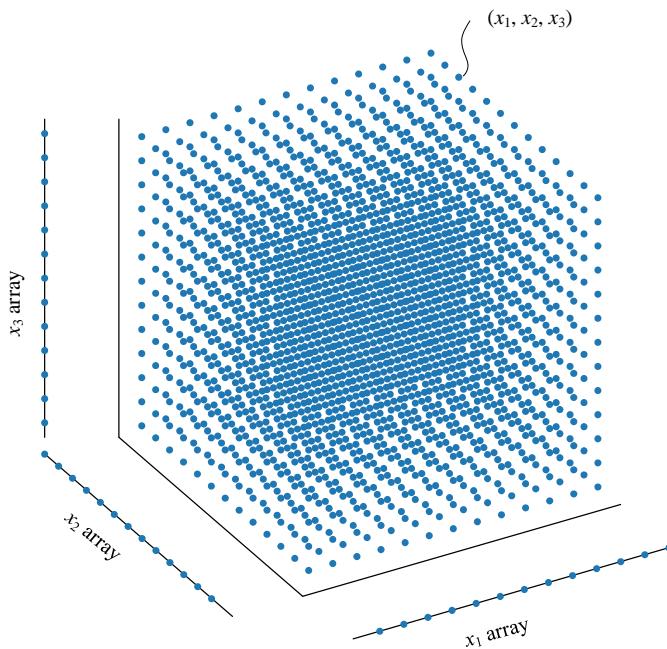


图 7. 三维网格

14.2 极坐标网格

除了方方正正的网格，本系列丛书还会用到极坐标网格。产生如图 8 所示的极坐标网格很容易。首先利用 `numpy.linspace()` 生成极角、极轴的数组，然后用 `numpy.meshgrid()` 生成极坐标网格坐标，最后再将其从极坐标转化为平面直角坐标系坐标。

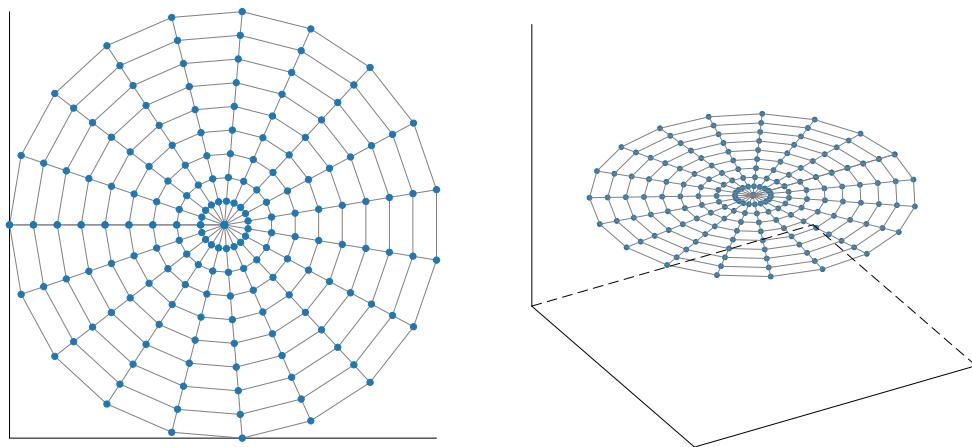


图 8. 极坐标网格

14.3 三角网格

本书后续还会使用三角网格完成特定的可视化方案。三角形网格是由一系列三角形所组成的网格结构。在计算机图形学和计算机模拟等领域，三角形网格常被用于表示复杂的几何体，如曲面、体细胞等，它可以通过三角形边界的拼接来逼近这些复杂的几何形状。三角形网格也常被用于数值计算中，如有限元分析等，因为三角形具有良好的性质，如易于计算、几何尺寸不变等。

三角形网格可以由多种方式生成，其中最常见的是 Delaunay 三角剖分，该方法可以将给定的点集分割成一组不重叠、不交叉的三角形。在 Delaunay 三角剖分中，对于任意三角形，其外接圆不包含其他点，这种性质可以保证三角形的质量较高，从而使得数值计算更加准确和稳定。

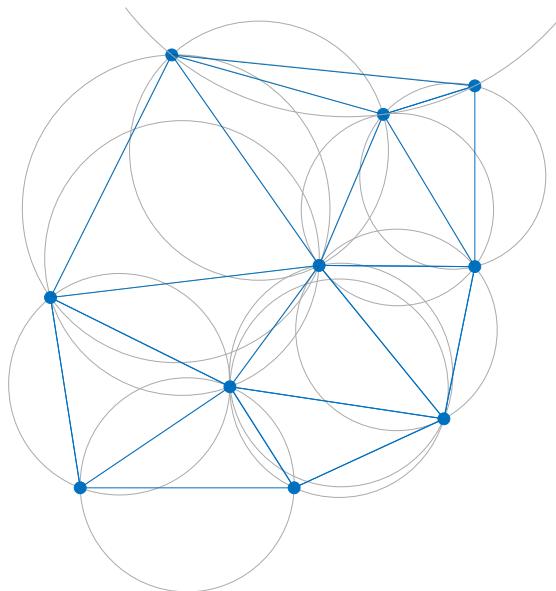


图 9. Delaunay 三角剖分法

`matplotlib.tri` 是一个 Python 库，用于创建和操作三角形网格。它提供了许多用于可视化和分析三角形网格的功能。

`matplotlib.tri` 可以创建三角形网格。可以使用 `Triangulation` 类从给定的点集中创建一个三角形网格，也可以使用其他函数生成各种类型的网格，如 Delaunay 三角剖分等。

`matplotlib.tri` 还可以可视化三角形网格。可以使用 `tripcolor`、`tricontour` 等函数在三角形网格上绘制颜色填充、等高线图等。

`matplotlib.tri` 也可以操作三角形网格。比如，可以使用 `TriAnalyzer`、`TriInterpolator` 等类对三角形网格进行分析、插值等操作。

总的来说，`matplotlib.tri` 为处理三角形网格提供了很多方便的工具和函数，使得用户可以方便地进行可视化和分析。

图 10 所示为常见三种三角网格。本书后续还将深入介绍三角形网格及其应用场景。

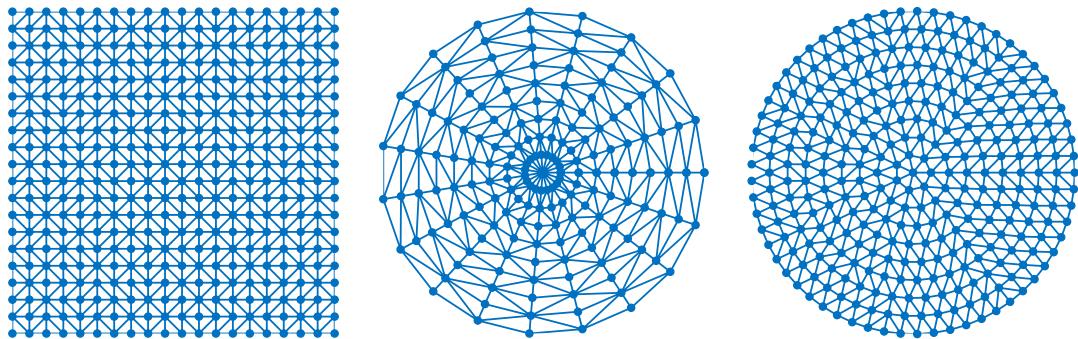


图 10. 三角网格

大家如果对 Delaunay 三角剖分法感兴趣的话，可以参考：

<https://mathworld.wolfram.com/DelaunayTriangulation.html>



Jupyter 笔记 BK_2_Ch14_1.ipynb 绘制本章大部分图片。

15

2D Contours

平面等高线

每条曲线上的点具有相同的数值



征服你自己，而不是世界。

Conquer yourself rather than the world.

——勒内·笛卡尔 (René Descartes) | 法国哲学家、数学家、物理学家 | 1596 ~ 1650



- ◀ `matplotlib.pyplot.contour()` 绘制等高线图
- ◀ `matplotlib.pyplot.contourf()` 绘制平面填充等高线
- ◀ `numpy.diag()` 如果 A 为方阵, `numpy.diag(A)` 函数提取对角线元素, 以向量形式输入结果; 如果 a 为向量, `numpy.diag(a)` 函数将向量展开成方阵, 方阵对角线元素为 a 向量元素
- ◀ `numpy.dot()` 计算向量标量积。值得注意的是, 如果输入为一维数组, `numpy.dot()` 输出结果为标量积; 如果输入为矩阵, `numpy.dot()` 输出结果为矩阵乘积, 相当于矩阵运算符 @
- ◀ `numpy.linalg.det()` 计算行列式值
- ◀ `numpy.linalg.inv()` 矩阵求逆
- ◀ `numpy.meshgrid()` 创建网格化数据
- ◀ `numpy.sqrt()` 计算平方根
- ◀ `numpy.vstack()` 返回竖直堆叠后的数组
- ◀ `numpy.zeros_like()` 用来生成和输入矩阵形状相同的零矩阵
- ◀ `sympy.diff()` 求解符号导数和偏导解析式
- ◀ `sympy.exp()` 符号自然指数
- ◀ `sympy.lambdify()` 将符号表达式转化为函数

15.1 等高线

平面等高线、三维等高线是“鸢尾花书”中非常重要的可视化方案。我们常用等高线可视化二元乃至多元函数、概率密度函数、机器学习中的决策边界等等。

这个话题主要介绍平面等高线，本书后续将专门介绍三维等高线。

等高线原理

等高线图是一种展示三维数据的方式，其中相同数值的数据点被连接成曲线，形成轮廓线。

形象地说，如图 1 所示，二元函数相当于一座山峰。在平行于 x_1x_2 平面在特定高度切一刀，得到的轮廓线就是一条等高线。这是一条三维空间等高线。然后，将等高线投影到 x_1x_2 平面，我们便得到一条平面等高线。

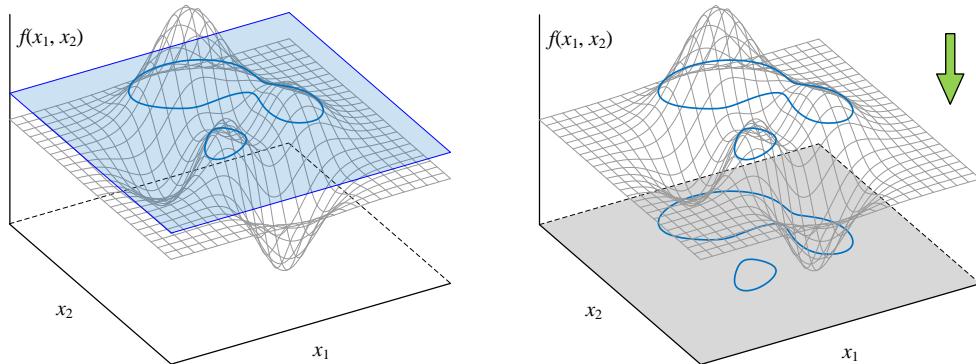


图 1. 平行 x_1x_2 平面切 $f(x_1, x_2)$ ，然后再投影到 x_1x_2 平面

一系列轮廓线的高度一般用不同的颜色或线型表示，使得我们可以通过可视化方式看到数据的分布情况。如图 2 所示，将一组不同高度的等高线投影到平面便得到如图 3 所示的平面等高线。图 3 的右侧还增加了色谱条，用来展示不同等高线对应的具体高度。这一系列高度可以是一组用户输入的数值。大家可能已经发现，等高线图和海拔高度图原理完全相同。类似的图还有，等温线、等降水线、等距线等等。

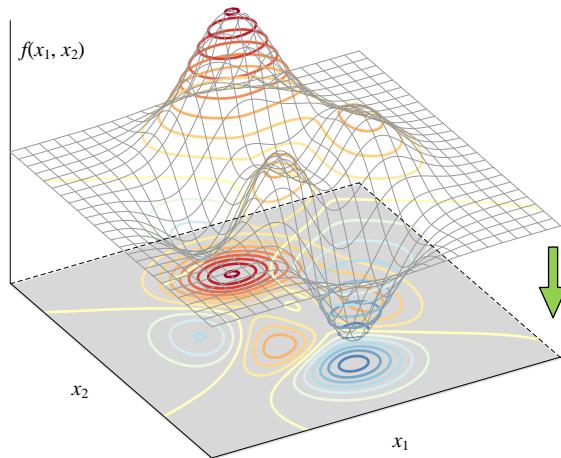


图 2. 将不同高度(值)对应的一组等高线投影到 x_1x_2 平面

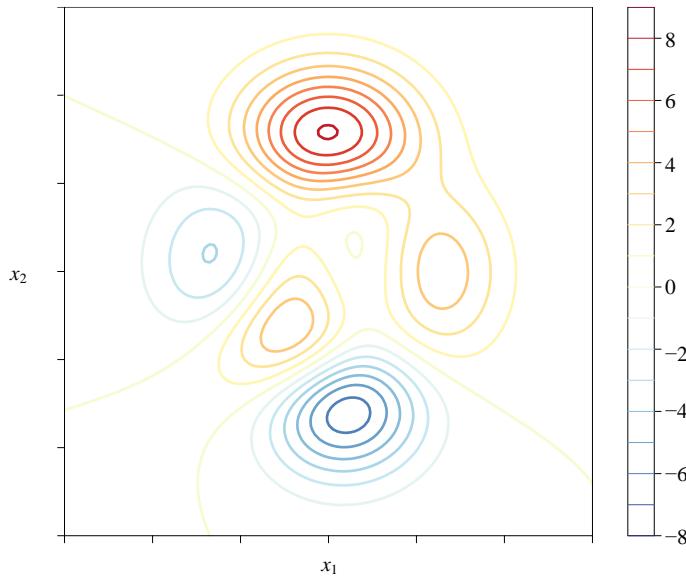


图 3. 平面等高线

步骤

具体来说，Matplotlib 中绘制等高线图需要以下步骤：

准备数据：等高线图需要的数据是一个二维数组，其中每个元素都表示一个点的数值。通常这个二维数组被称为“网格”。

计算轮廓线：Matplotlib 会通过对数据进行插值，计算出一组轮廓线的值，并把它们绘制在二维平面上。轮廓线的数量取决于我们指定的等高线的数量。

绘制轮廓线：Matplotlib 会根据轮廓线的高度在不同的颜色或线型中表示，使得我们可以通过视觉化方式看到数据的分布情况。通常使用 plt.contour() 函数进行绘制。

等高线设置

Matplotlib 还提供了许多与等高线图相关的函数和选项，例如设置轮廓线的样式、标签、标记等。这些选项可以帮助我们更好地展示数据。

图 4 所示为给每条等高线增加高度注释。图 5 所示为单色等高线，matplotlib 会用虚线代表负值。

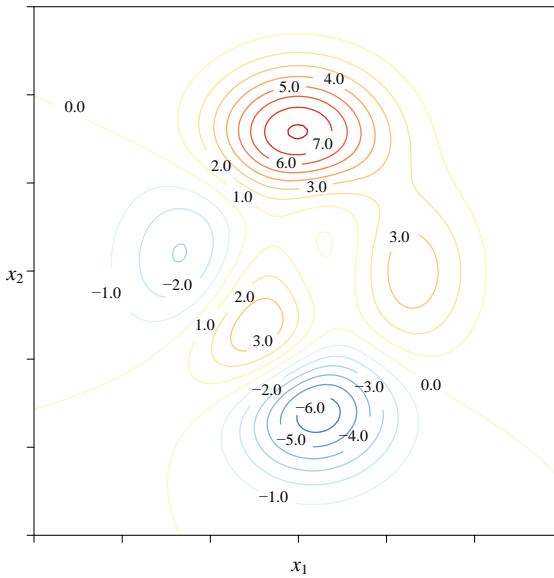


图 4. 给等高线加注释

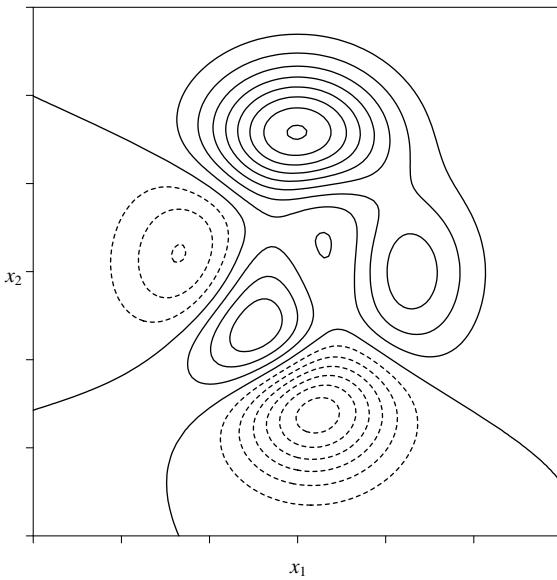


图 5. 单色等高线

填充等高线

`contourf` 是 Matplotlib 库中用于绘制等高线填充图的函数。其原理是通过将数据转换为等高线段的集合，然后通过填充线段之间的空隙来创建颜色填充图。具体地说，`contourf` 函数首先根据数据生成一组等高线，这些等高线可以使用 `contour` 函数绘制。然后 `contourf` 函数会根据这些等高线，将图像中每个等高线所围成的区域填充上颜色。填充的颜色根据指定的 `cmap` 进行选择，可以通过设置 `cmap` 参数来控制。

如图 6 所示，在三维空间中，我们可以把填充等高线想象成是“梯田”。每个颜色代表一定的高度区间。将图 6 填充等高线画在水平面上，我们便得到图 7。图 7 中，我们还绘制一条指定高度 (0.0) 的等高线。

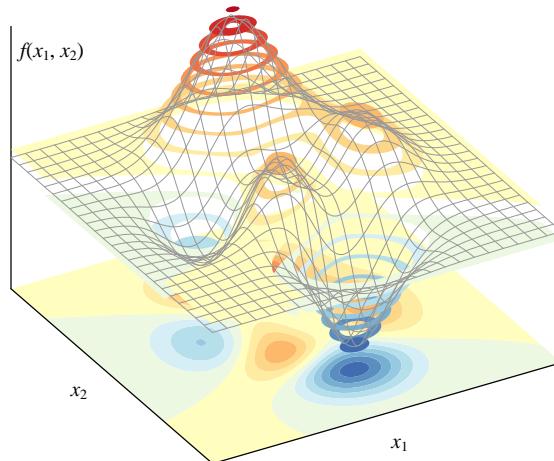


图 6. 填充等高线原理

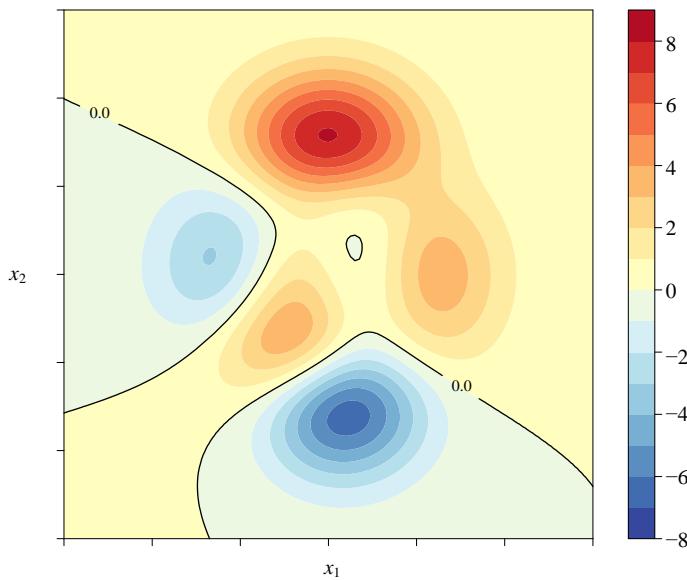


图 7. 平面填充等高线

此外，`contourf` 还可以通过调整填充级别的数量和间隔来控制填充效果，可以通过设置 `levels` 参数来控制填充级别的数量和大小。此外，还可以通过设置 `alpha` 参数来控制填充颜色的透明度，以及通过设置 `extend` 参数来控制在 `levels` 范围之外的值应如何处理。



Jupyter 笔记 BK_2_Ch15_1.ipynb 绘制本话题前几幅图。

在机器学习中，填充等高线常被用来绘制决策边界。图 11 所示为决策树、朴素贝叶斯两种方法分类鸢尾花数据集的决策边界。

15.2 颗粒度

前文提过，绘制连续等高线基于插值，因此等高线是否“平滑”，也取决于网格数据的颗粒度是否足够细腻。

如图 8 所示，绘制等高线采用的网格数据显然过于粗糙，这导致不管是三维等高线，还是平面等高线都非常毛糙。

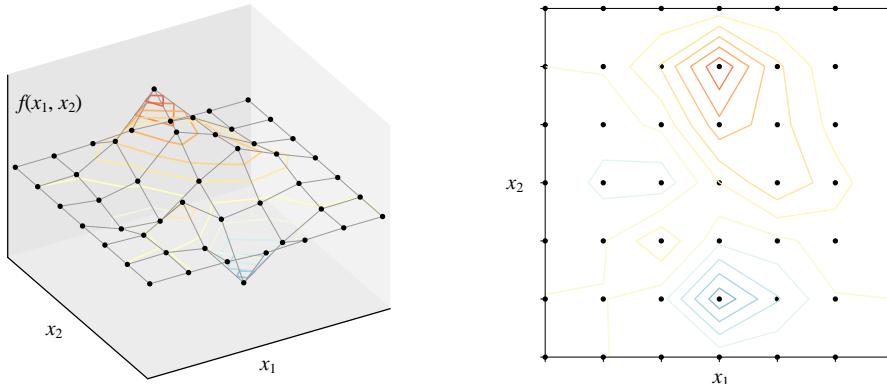


图 8. 颗粒度粗糙

图 9 则采用更为细腻的网格数据。显然，插值得到的三维等高线、平面等高线看上去更加平滑。

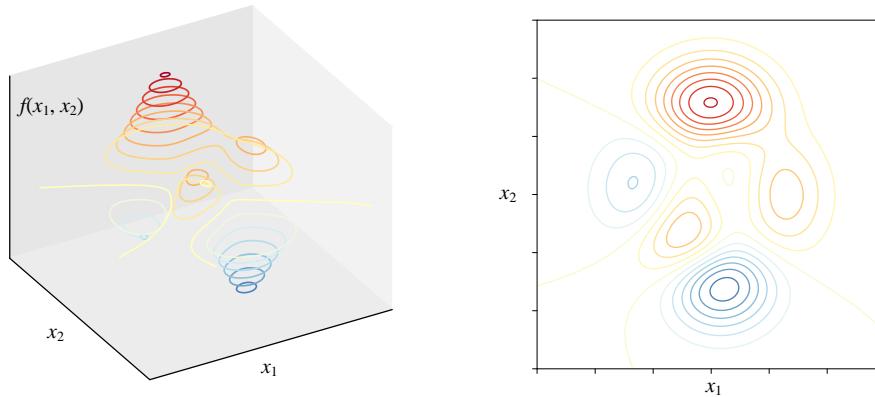


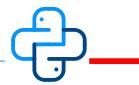
图 9. 颗粒度细腻



Jupyter 笔记 BK_2_Ch15_2.ipynb 绘制图 8、图 9。

网格不是越密越好。网格能够取得多密，还受到算力制约。比如，图 12 所示马氏距离、高斯分布概率密度等高线，进一步提高网格密度会受到内存空间限制。

本书后续还会介绍等高线的更广泛的用途。



Jupyter 笔记 BK_2_Ch15_3.ipynb 绘制图 12。

15.3 可视化线性、非线性变换

如图 10 所示，`contour()` 函数还可以绘制网格。图 13、图 14 展示几种线性、非线性变换，图中网格均用 `contour()` 绘制。其中几幅图用到了复数函数，本书后续还要展示其他可视化方案呈现复数函数。此外，本书后续还会专门讲解线性变换、仿射变换。

复数是数学中的一个概念，用来表示具有实部和虚部的数。复数通常表示为 $a + bi$ 的形式，其中 a 是实部， b 是虚部， i 是虚数单位。实部和虚部都可以是实数。复数的集合用 C 表示。

复数函数是定义在复数域上的函数，即将复数作为输入并产生复数作为输出的函数。复数函数可以包含各种数学操作和运算，例如加法、减法、乘法、除法、指数函数、对数函数等。

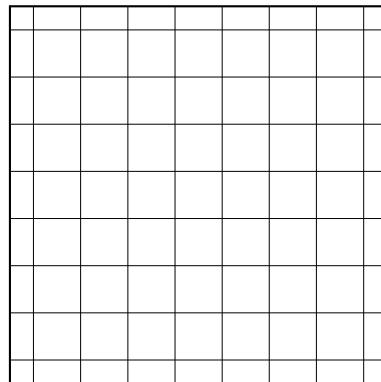


图 10. 用 `contour()` 函数绘制的方格

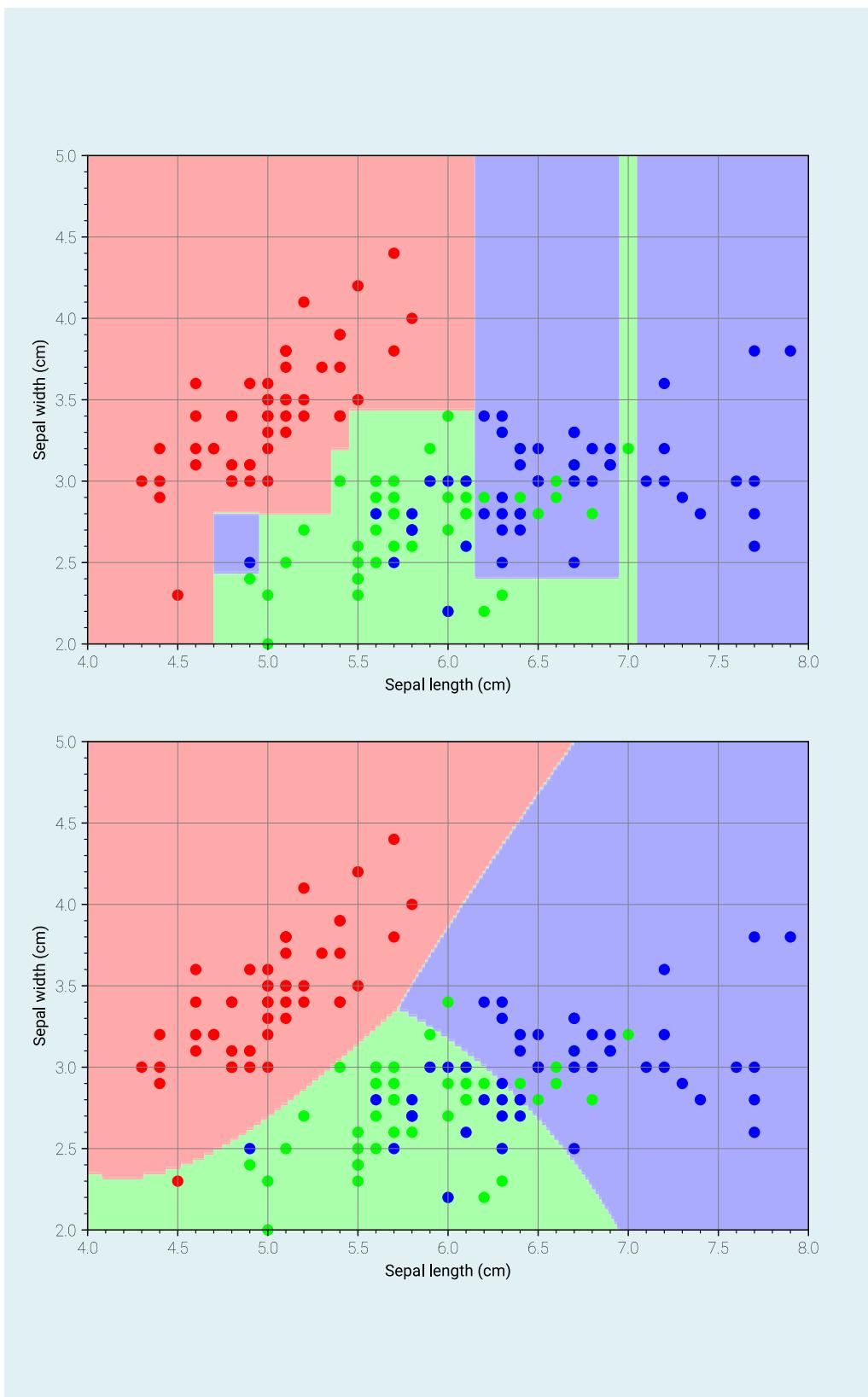


图 11. 利用 contourf() 绘制决策边界等高线

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

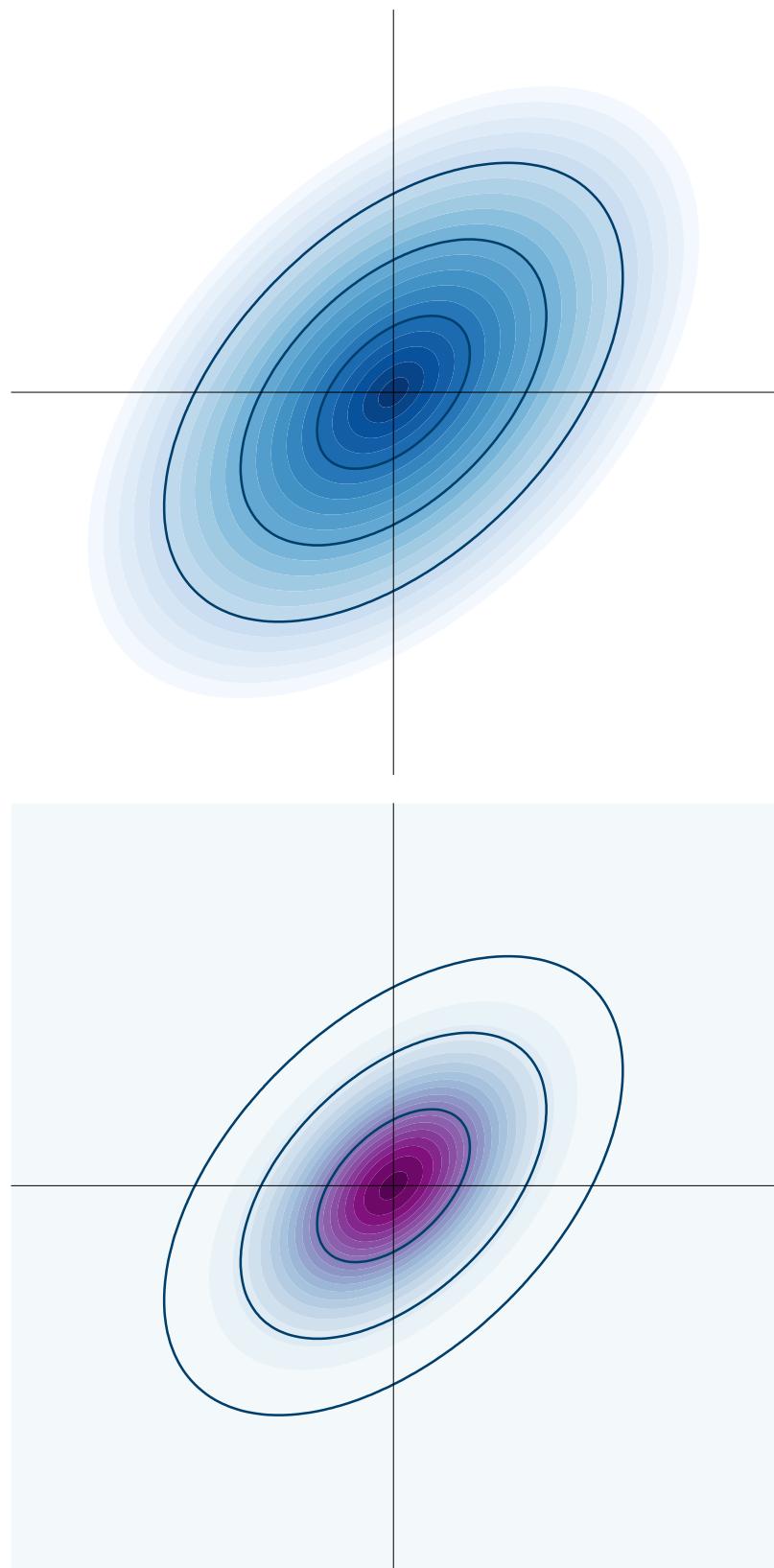


图 12. 马氏距离、高斯概率密度函数等高线

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

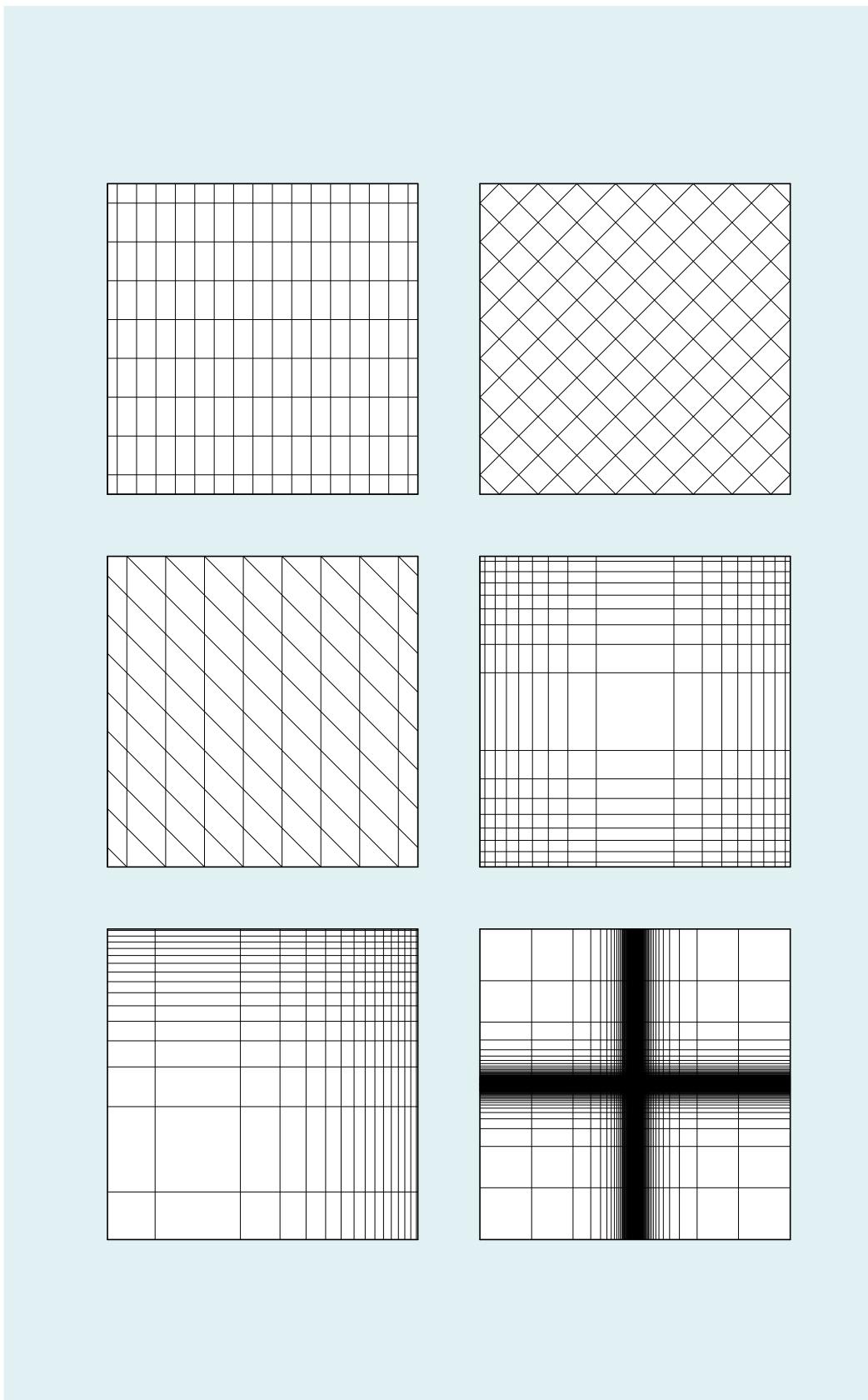


图 13. 线性、非线性变换，利用 contour 绘制，第 1 组

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

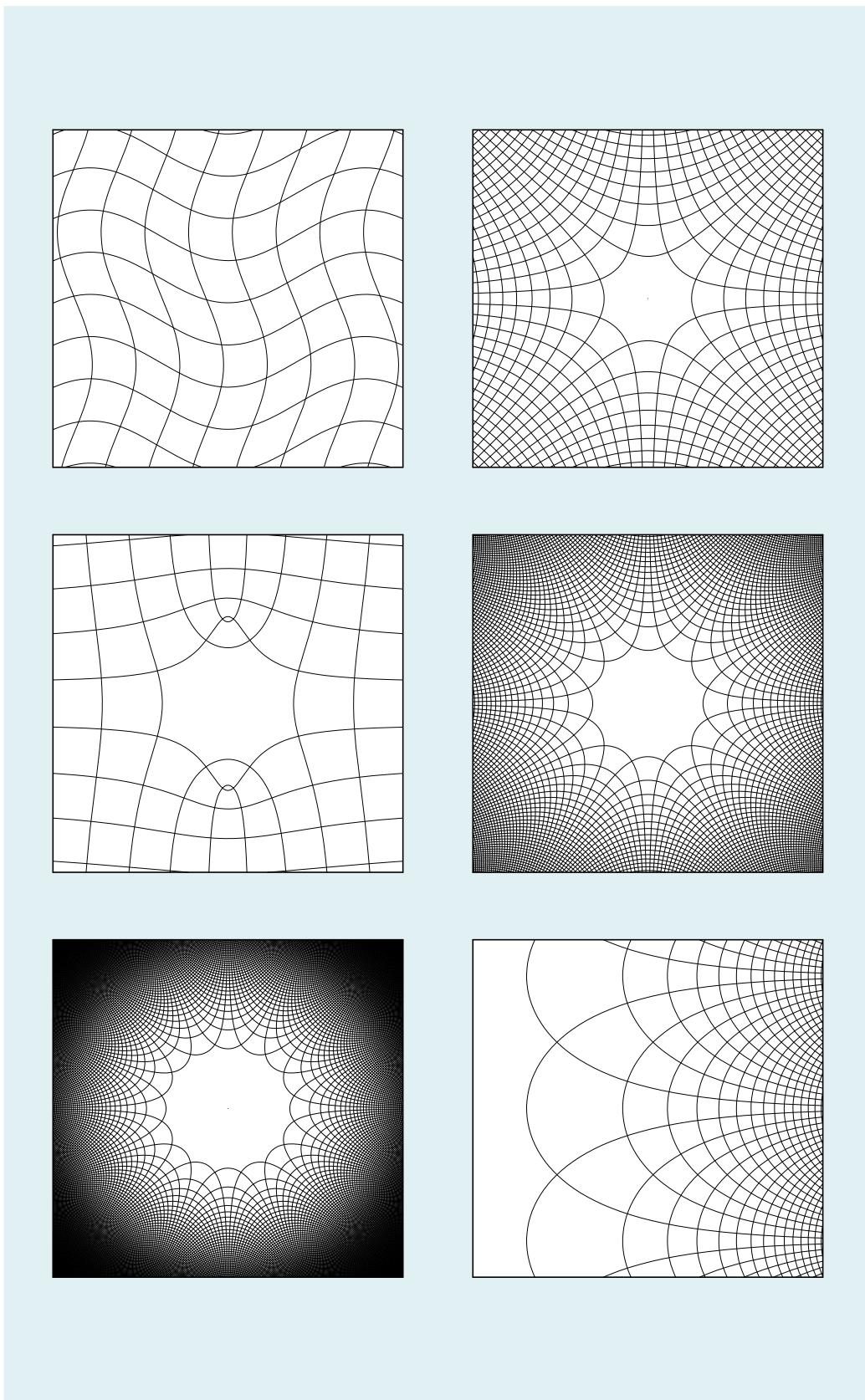


图 14. 线性、非线性变换，利用 contour 绘制，第 2 组

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com



Heatmap

热图

鸢尾花书中常用来可视化矩阵运算



每个孩子都是艺术家。问题在于他长大后如何保持艺术家的本质。

Every child is an artist. The problem is how to remain an artist once he grows up.

—— 毕加索 (Pablo Picasso) | 西班牙艺术家 | 1881 ~ 1973



- ◀ `numpy.linalg.cholesky()` Cholesky 分解
- ◀ `numpy.linalg.eig()` 特征值分解
- ◀ `numpy.linalg.svd()` 奇异值分解
- ◀ `numpy.zeros_like()` 用来生成和输入矩阵形状相同的零矩阵
- ◀ `seaborn.clustermap()` 绘制聚类热图
- ◀ `seaborn.heatmap()` 绘制热图
- ◀ `sklearn.datasets.load_iris()` 加载鸢尾花数据

16.1 热图

Seaborn 中的热图

热图 (heatmap) 是“鸢尾花书”中极为常见的可视化手段。特别是在展示数据、矩阵分解时，我们常用热图可视化矩阵。

虽然，matplotlib 中也有绘制热图的工具；但是，推荐大家使用 seaborn 中的 heatmap 函数。这个函数绘制热图更方便。

Seaborn 是一款基于 matplotlib 的数据可视化库，其中包括了各种绘图函数，其中之一就是 heatmap。使用 Seaborn 的 heatmap 函数可以让大家快速而方便地可视化矩阵数据，使得数据分析更加直观和易于理解。

heatmap 可以用于可视化二维数组。图 1 所示为用热图可视化鸢尾花四个量化特征数据。在 Jupyter notebook 中，大家可以看到我们用 cmap 控制色谱，用 xticklabels、yticklabels 分别控制横轴、纵轴标签，用 cbar_kws 设置色谱条位置，并用 vmin、vmax 控制色谱条起止位置。

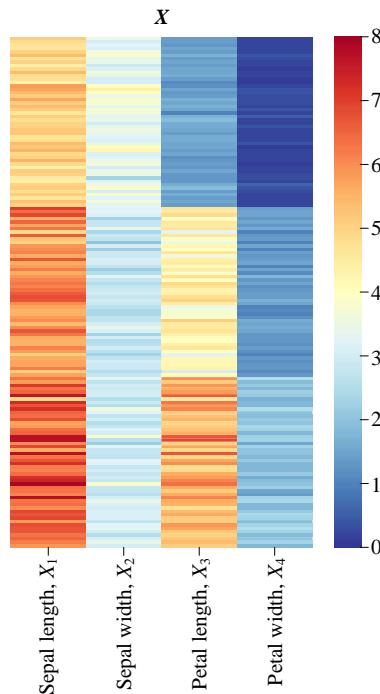


图 1. 热图可视化鸢尾花数据

Seaborn 中的 heatmap 函数还包括许多其他参数，用于自定义热图的外观和行为。例如，大家可以使用 annot 参数在热图中显示数值，使用 fmt 参数指定数字格式，使用 linewidths 参数调整单元格边框宽度等等。

聚类热图

Seaborn 中，`clustermap` 是一个用于绘制聚类热图的函数，其原理是将矩阵中的行和列进行聚类，并以聚类后的顺序重新排列矩阵的行和列。这样可以将具有相似特征的行和列放在一起，从而更容易地发现它们之间的相似性和差异性。图 2 所示为鸢尾花数据的聚类热图。

“鸢尾花书”的《机器学习》将专门讲解各种聚类算法。

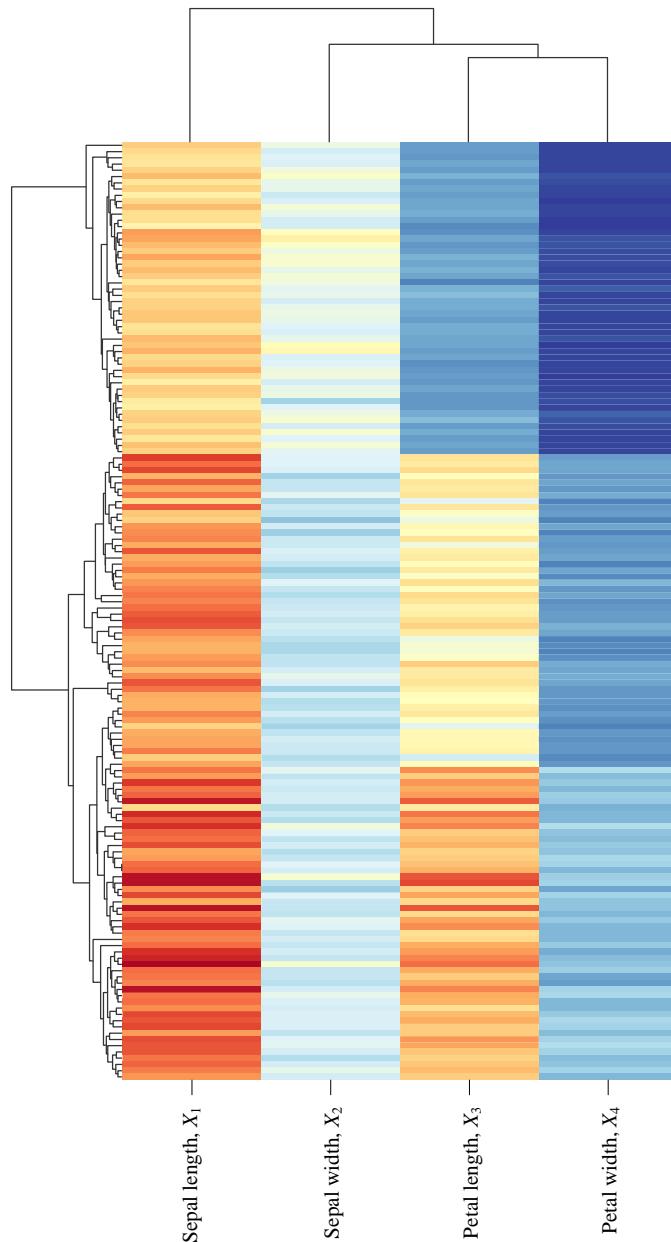


图 2. 热图可视化鸢尾花数据

16.2 矩阵运算

“鸢尾花书”中，大家会经常看到用一组热图可视化矩阵运算，特别是矩阵分解。图 5 所示为常见的几个矩阵运算。注意，后期制作时，热图的形状做了修改。

《矩阵力量》一册将从代数、数据、线性组合、优化、几何、统计等角度和大家讨论这些矩阵运算。此外，大家还会看到我们用热图可视化协方差矩阵、相关性系数矩阵，以及这些矩阵对应的线性代数运算。本节就不再展开讨论了。

16.3 伪彩色网格图

在 Matplotlib 中，`pcolormesh` 函数用于创建一个伪彩色网格图，类似热图。它可用于绘制二维数据的色彩填充图，其中每个数据点的颜色根据其对应的数值进行映射。

在 `pcolormesh` 函数中，可以使用 `rasterized` 参数来控制是否将图形渲染为矢量图形或光栅图像。`rasterized` 参数是一个布尔值，用于指定是否将图形渲染为光栅图像。当设置为 `True` 时，图形将以光栅化的形式保存，这对于包含大量数据点或复杂图形时可以提高渲染性能和文件大小。默认情况下，`rasterized` 参数的值为 `False`，即图形以矢量格式渲染。

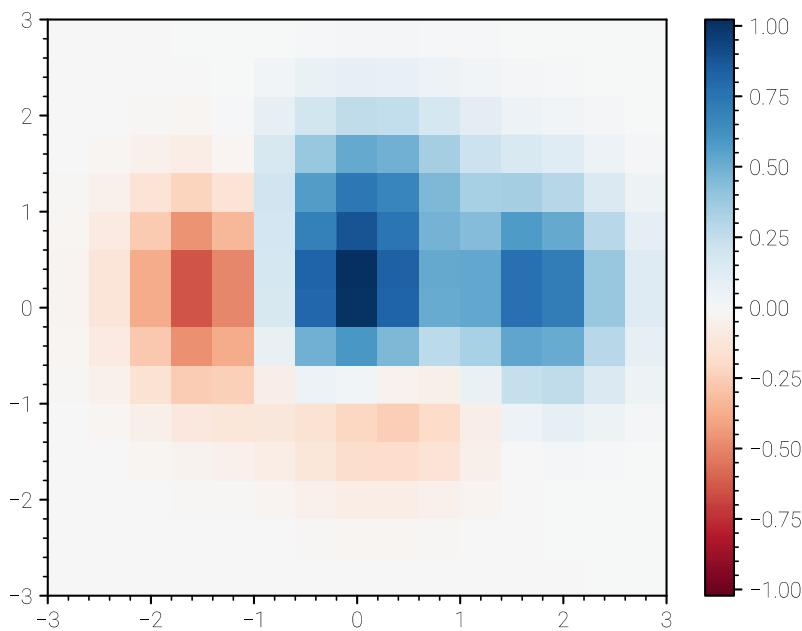
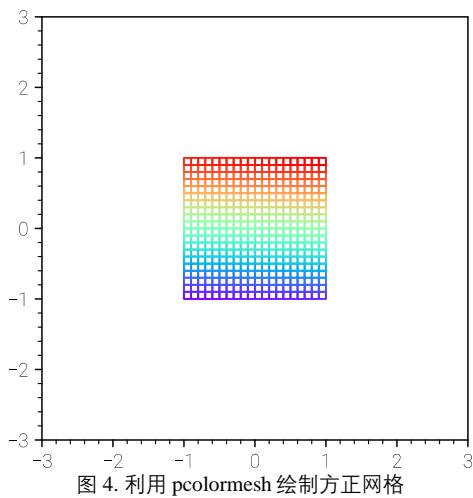


图 3. 利用 `pcolormesh` 绘制的伪彩色图

如图 6 所示，`pcolormesh` 函数还常用来绘制分类算法的决策边界。此外，`pcolormesh` 函数可以绘制网格，并用来可视化线性、非线性变换，具体如图 4、图 7 所示。



pcolor 函数也是 matplotlib 库中的函数，用于绘制伪彩色图，效果和 pcolormesh 类似。与 pcolor 相比，pcolormesh 在效率上更高，特别适用于绘制大型数据集。

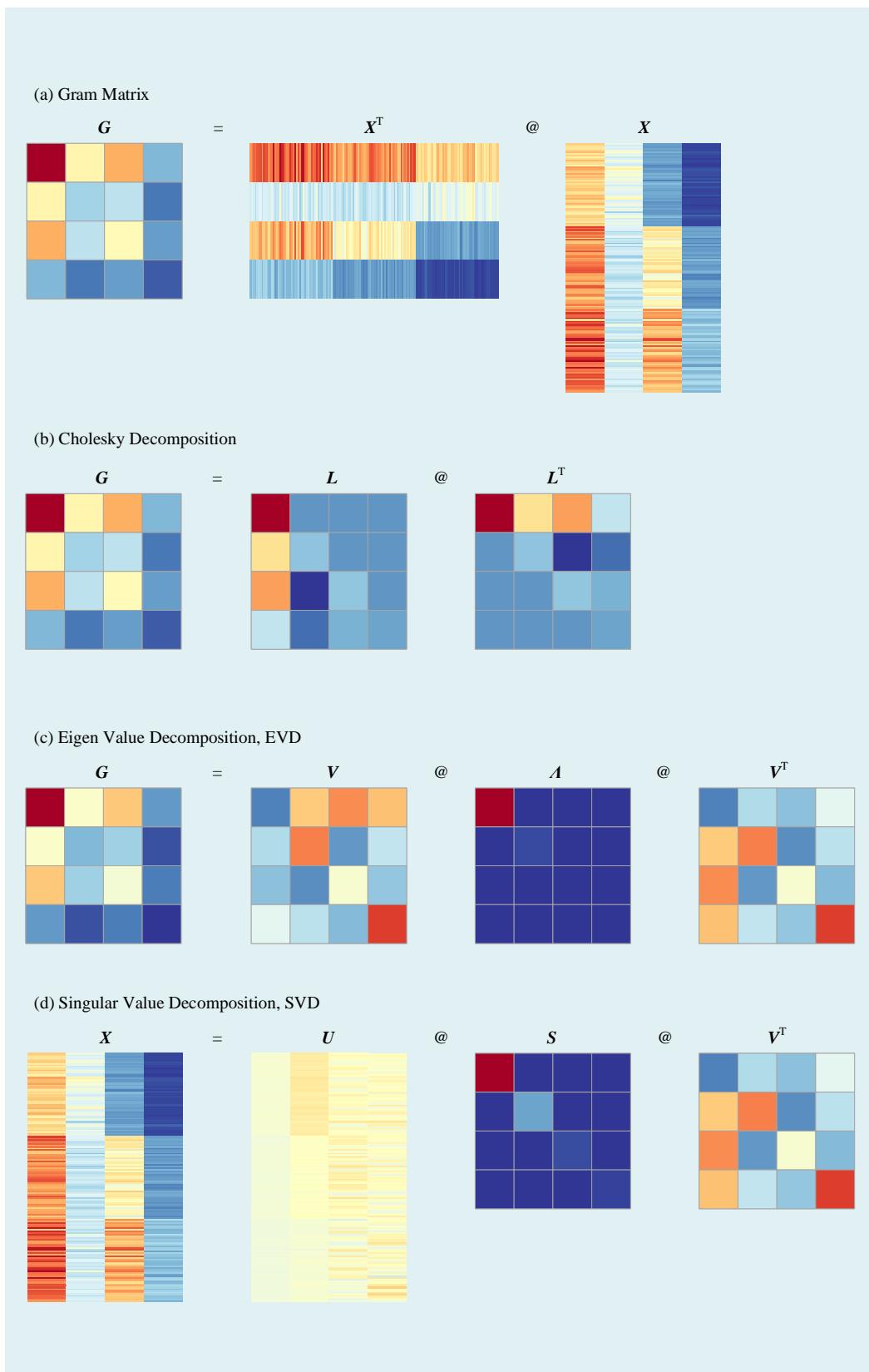


图 5. 用热图可视化矩阵运算

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

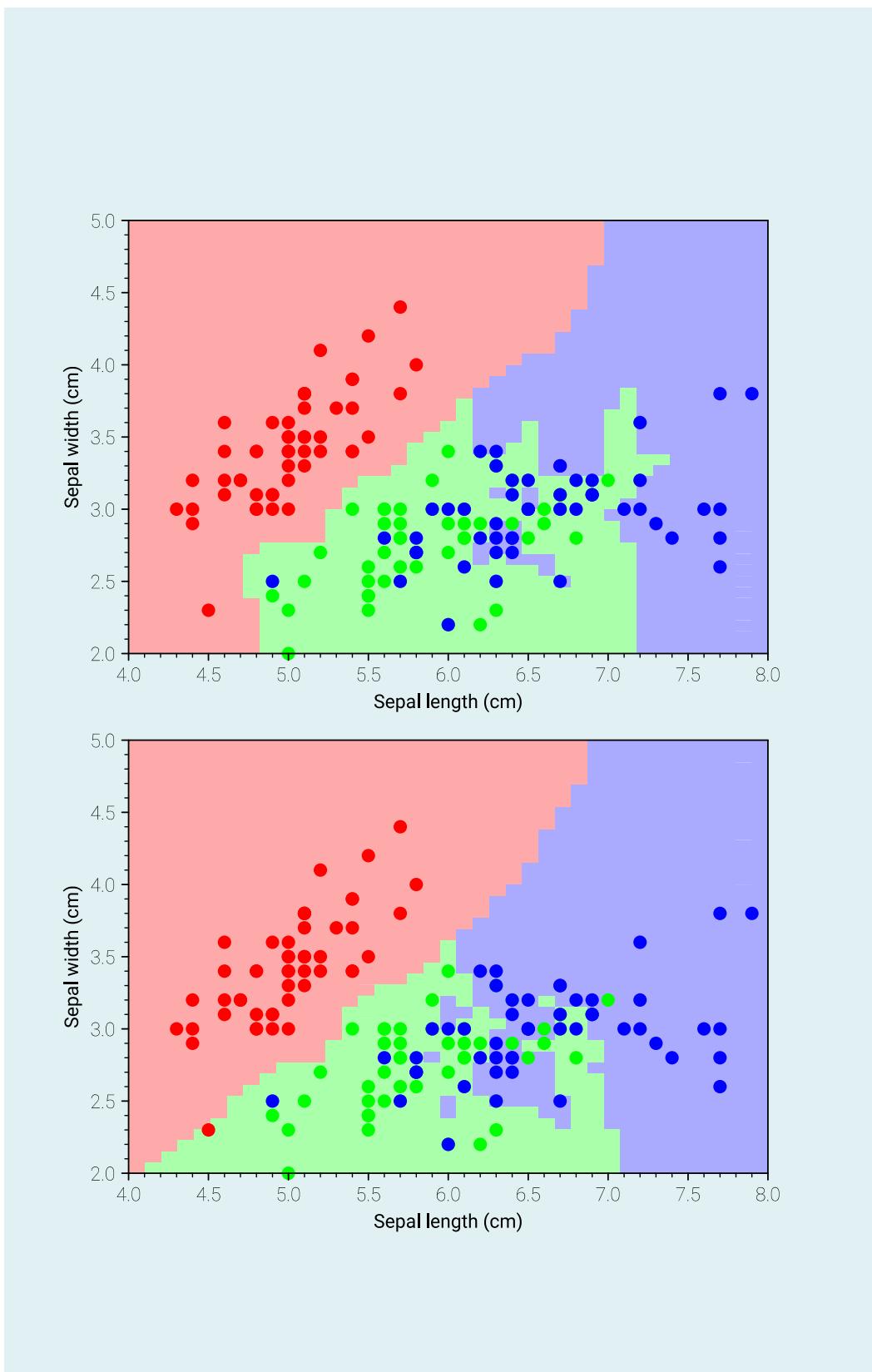


图 6. 用 pcolormesh 函数绘制分类决策边界

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

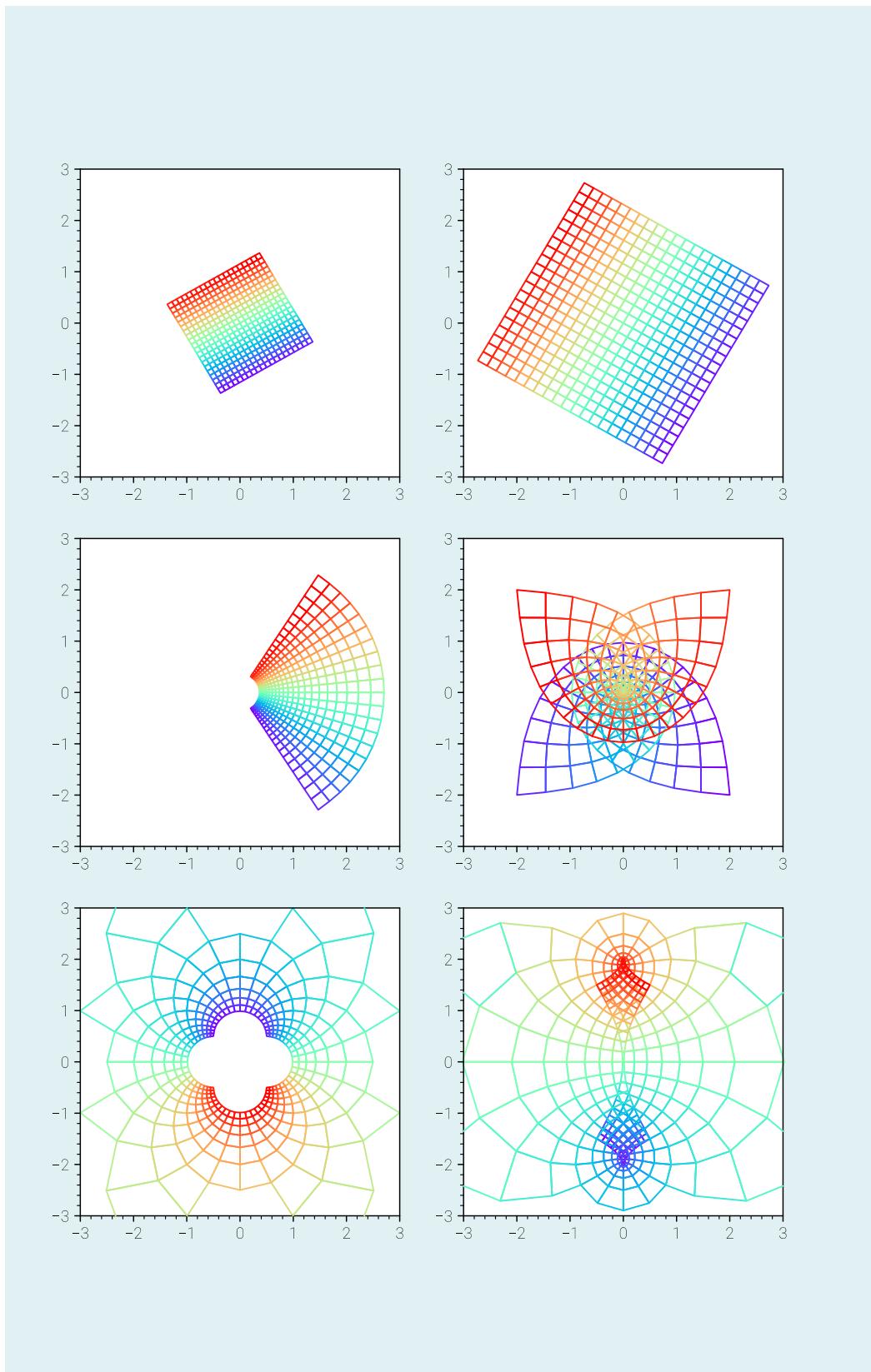


图 7. 用 pcolormesh 函数可视化线性、非线性变换

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

17

Non-Vector Images

非矢量图片

本质上为红绿蓝像素数据矩阵



博爱是好事，因为真正的力量就蕴藏在其中。深爱之人会付出更多，做成更多事情，并且能够取得更大的成就。而在爱的驱使下所做的一切都会变得出色。

It is good to love many things, for therein lies the true strength, and whosoever loves much performs much, and can accomplish much, and what is done in love is well done.

——文森特·梵高 (Vincent van Gogh) | 荷兰后印象派画家 | 1853 ~ 1890



- ◀ `matplotlib.image.imread()` 读取图像文件并返回对应的图像数据
- ◀ `matplotlib.pyplot.hist()` 绘制直方图
- ◀ `matplotlib.pyplot.imshow()` 显示图像数据
- ◀ `numpy.zeros()` 返回给定形状和类型的新数组，用零填充
- ◀ `numpy.zeros_like()` 用来生成和输入矩阵形状相同的零矩阵
- ◀ `skimage.color.rgb2gray()` 将彩色图像转换为灰度图像
- ◀ `skimage.io.imread()` 读取图像文件并返回对应的图像数据

17.1 像素

`matplotlib.image` 模块提供了读取和处理图像的函数，其中最常用的函数是 `imread`。`imread` 函数可以读取图像文件，并将其解码为一个三维的 numpy 数组。

`imshow` 是 `matplotlib` 中用于显示图像的函数。将如图 1 所示鸢尾花照片导入后，容易发现这幅图像实际上是一个 $2990 \times 2714 \times 3$ 的数组。

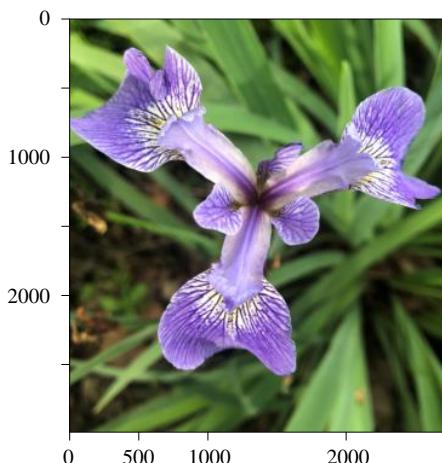


图 1. 鸢尾花照片

图片像素 (pixel) 是图片的基本单位，是构成图片的最小元素。它是一个有限的、离散的、二维的点，有着特定的位置、颜色和亮度值。在数字图像中，每个像素都有一个确定的坐标和值。图片中的像素数量越多，图片的分辨率就越高，图片的清晰度和细节也就越好。

像素的颜色通常使用 RGB 值 (红、绿、蓝三种颜色的强度组合) 表示。每个像素都有一个红、绿、蓝三个通道的值。红、绿、蓝可以分别被编码为一个数字，例如 8 位的数字可以表示 256 种颜色。

也就是说，图 1 这幅图中每个像素首先分解成红绿蓝三个数值。这些数值的取值范围都在 $[0, 255]$ 之间。换个角度，图 1 可以理解成是由三幅图片叠加而成，如图 2 所示。

此外，我们可以获得如图 3 所示的红绿蓝颜色的分布。越靠近 0，颜色越靠近黑，越靠近 255 颜色越靠近纯色。本书前文已经和大家聊过 $[0, 0, 0]$ 代表纯黑， $[255, 255, 255]$ 代表纯白。注意，在 `matplotlib` 中 $[1, 1, 1]$ 代表纯白。

在彩色图像中，每个像素的颜色可以由三个 8 位数字 (红、绿、蓝) 组成，因此彩色图像中的每个像素可以表示 $2^{3 \times 8}$ 种不同的颜色，约为 1600 万种。

在数字图像处理中，对图像进行各种操作，例如缩放、旋转、裁剪、调整亮度和对比度等，都会涉及到像素的处理和修改。

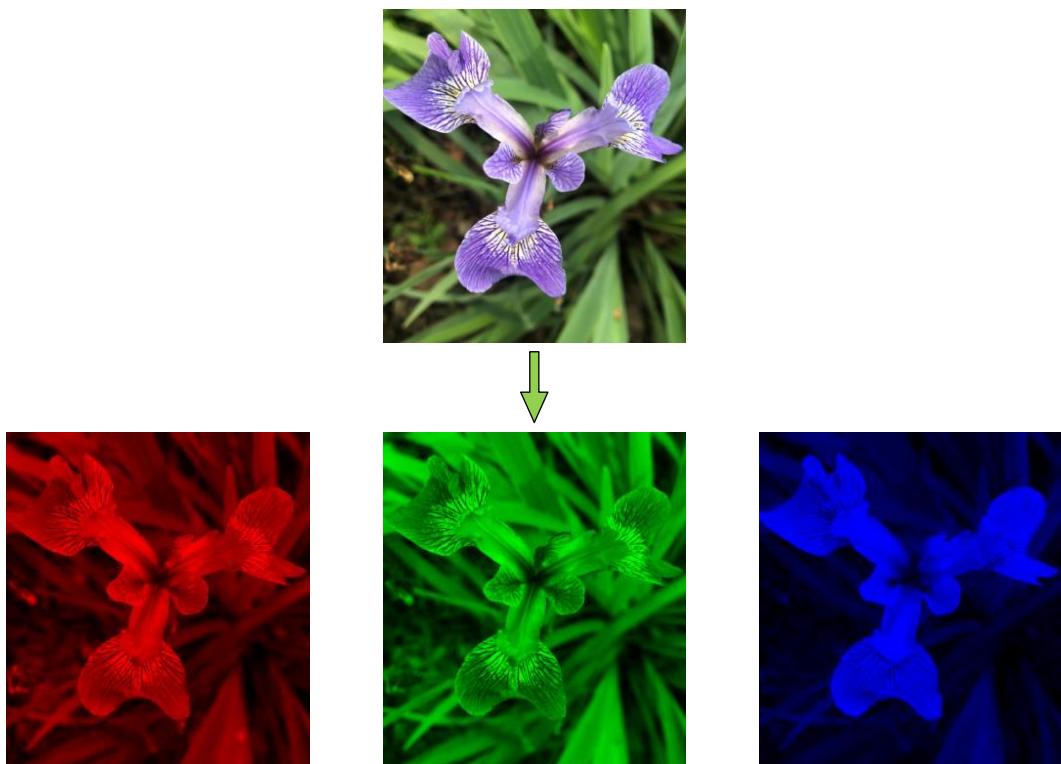


图 2. 鸢尾花照片分解成红绿蓝三个通道

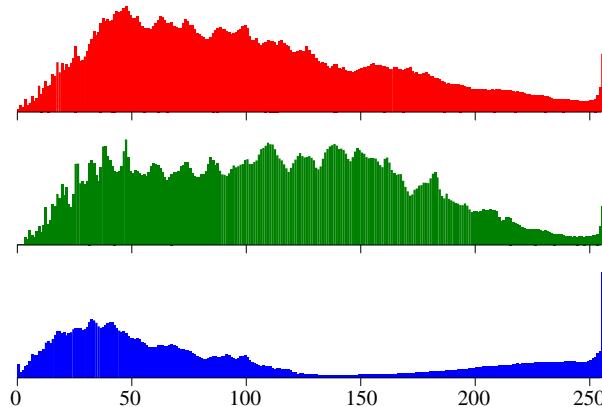


图 3. 鸢尾花照片红绿蓝颜色分布

红绿蓝三个通道

图 4 给出的三幅子图，每幅图仅保留两色通道，另外一个通道数值全部置零。



图 4. 鸢尾花照片，只保留两色通道

色谱

它可以用来显示二维数组或图像文件中的图像。imshow 函数有很多参数可以控制图像的外观。例如，可以使用 cmap 参数指定要使用色谱。图 5 所示为使用色谱展示红色通道。Jupyter notebook 中还给出更多范例。

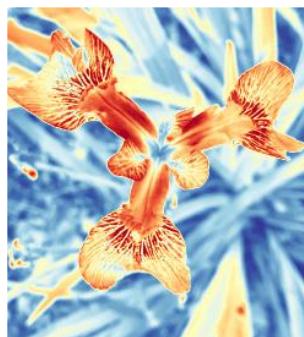


图 5. 使用色谱展示红色通道

灰度

Scikit-image (skimage) 是一个用于图像处理和计算机视觉的 Python 包。它提供了一系列算法，函数和工具，可用于图像处理，包括图像滤波，几何变换，色彩空间转换，图像分割，特征提取等等。具体来说，skimage 可以用于：a) 加载和保存图像；b) 调整图像大小，旋转，裁剪等几何变换；c) 进行图像滤波和增强；d) 在不同颜色空间之间进行转换；e) 检测边缘和角点；f) 进行图像分割和分析；g) 进行特征提取和图像匹配。

图 6 所示为使用 skimage 将彩色图片转化为灰度图片。注意图片的每个像素的取值在 [0, 1] 之间。此外，图像识别一般都使用灰度图像。



图 6. 将彩色图片转化成灰度

修改部分像素

由于图片本身就是一个数组，我们可以通过修改数组的具体值来修改图片。如所示，我们将灰度照片的左上角 500×500 的像素变为白色。



图 7. 修改图片像素

降低像素

图 8 所示为通过采样降低图像像素。图 1 这幅图片的像素大小为 2990×2714 。每 200 个像素采样一个像素，我们便得到图 8。这幅图的像素为 15×14 ，很明显图片的颗粒度很粗糙。



图 8. 采样降低像素

当图像像素较低时，为了让图片看上去更细腻，我们可以采用插值。

17.2 插值

`imshow()` 函数中，我们可以通过设置 `interpolation` 参数来控制如何在图像像素之间进行插值，以生成更平滑的图像。

`imshow()` 函数 `interpolation` 参数的默认值是 '`antialiased`'，它使用反走样技术来平滑图像，使其在缩放时更加清晰。这意味着在缩放图像时，`imshow()` 函数会自动对图像进行插值，以获得更平滑的外观。

除了默认的 '`antialiased`' 插值，`imshow()` 函数还支持其他插值方法，包括 '`nearest`'，'`bilinear`'，'`bicubic`' 等。这些插值方法可以通过 `interpolation` 参数来设置。例如，'`nearest`' 插值只是在最近的

像素值之间进行插值，而 'bicubic' 插值使用更复杂的算法来生成更平滑的图像。图 9 所示为图 8 的两种插值结果。本节的 Jupyter notebook 中给出更多插值方法。

《数据有道》一册将详细讲解常见插值算法。

选择不同的插值方法会影响图像的视觉效果，因此选择合适的插值方法可以使图像更清晰或更平滑，更符合数据的视觉表达。

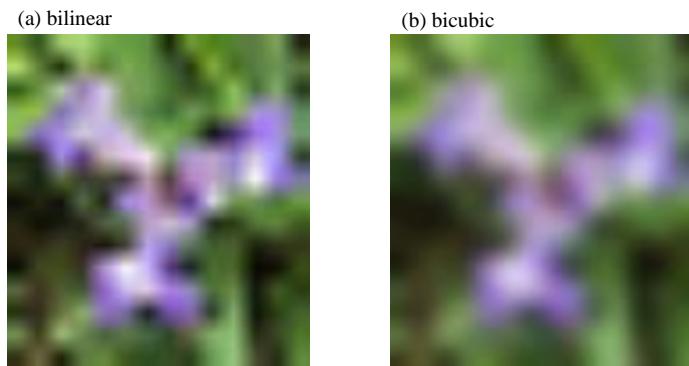


图 9. 插值平滑

18

3D Scatter Plot

三维散点

利用颜色、大小可视化其他特征



当一扇门关闭时，另一扇门打开；但是我们望眼欲穿、死死紧盯那扇关闭的门，看不到为我们打开的门。

When one door closes, another door opens; but we so often look so long and regretfully upon the closed door, that we do not see the ones which open for us.

—— 亚历山大·贝尔 (Alexander Bell) | 发明家、企业家 | 1847 ~ 1822



- ◀ `matplotlib.pyplot.scatter()` 绘制散点图
- ◀ `numpy.dot()` 计算向量标量积。值得注意的是，如果输入为一维数组，`numpy.dot()` 输出结果为标量积；如果输入为矩阵，`numpy.dot()` 输出结果为矩阵乘积，相当于矩阵运算符@
- ◀ `numpy.linalg.det()` 计算行列式值
- ◀ `numpy.linalg.inv()` 矩阵求逆
- ◀ `numpy.meshgrid()` 创建网格化数据
- ◀ `numpy.reshape()` 是一个函数，用于重新重塑一个数组的形状，而不改变其数据内容
- ◀ `numpy.where()` 根据给定的条件返回输入数组中满足条件的元素的索引或值
- ◀ `scipy.stats.dirichlet.pdf()` Dirichlet 分布概率密度函数
- ◀ `scipy.stats.multinomial.pmf()` 多项分布概率质量函数

18.1 三维散点

本书前文，大家已经看过用散点可视化 RGB 色彩空间。本章深入聊一下如何用三维散点可视化各种场景。

各种样式的三维散点

我们首先用三维散点图可视化鸢尾花数据，如图 1 所示。

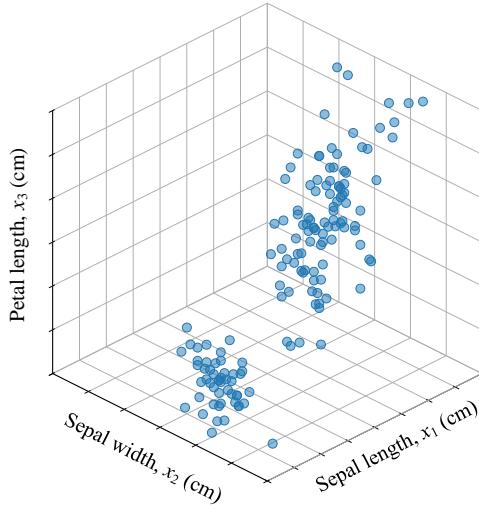


图 1. 空间散点可视化样本数据

在不同平面上的投影

空间散点可以投影到不同平面上。图 2 所示为空间散点投影在 $x_3 = 1$ 平面上，即 $z = 1$ 。图 3 (a) 所示为散点投影在 $x_1 = 8$ 平面上，即 $x = 8$ 。图 3 (b) 所示为散点投影在 $x_2 = 5$ 平面上，即 $y = 5$ 。

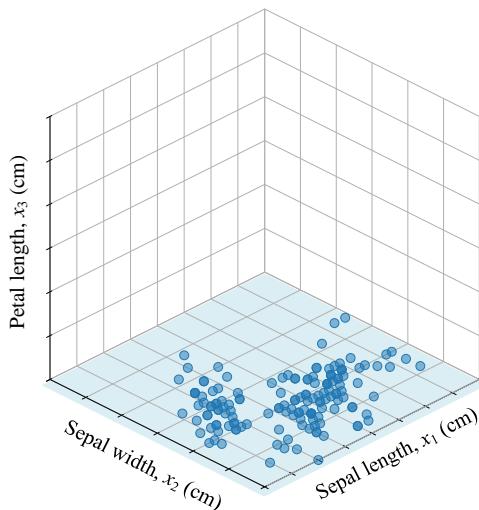
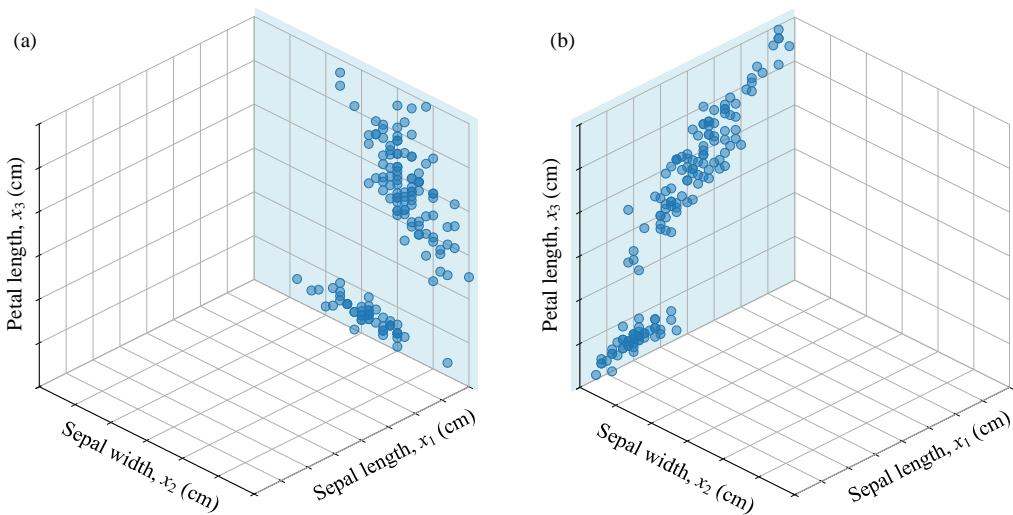


图 2. 空间散点投影在 $x_3 = 1$ ($z = 1$)

图 3. 空间散点投影在 $x_2 = 5$ ($y = 5$)、 $x_2 = 5$ ($y = 5$)

18.2 展示更多特征

类似平面散点图，我们可以用散点大小、颜色可视化更多特征。图 4 所示为用散点大小可视化鸢尾花花瓣宽度。

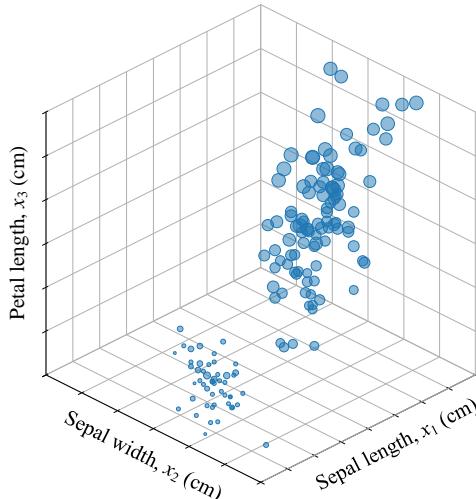


图 4. 用散点大小可视化鸢尾花花瓣宽度

图 5 (a) 所示为用颜色可视化鸢尾花类别。结合图 4、图 5 (a)，我们便得到图 5 (b)。图 5 (b) 可视化鸢尾花四个特征和分类标签。

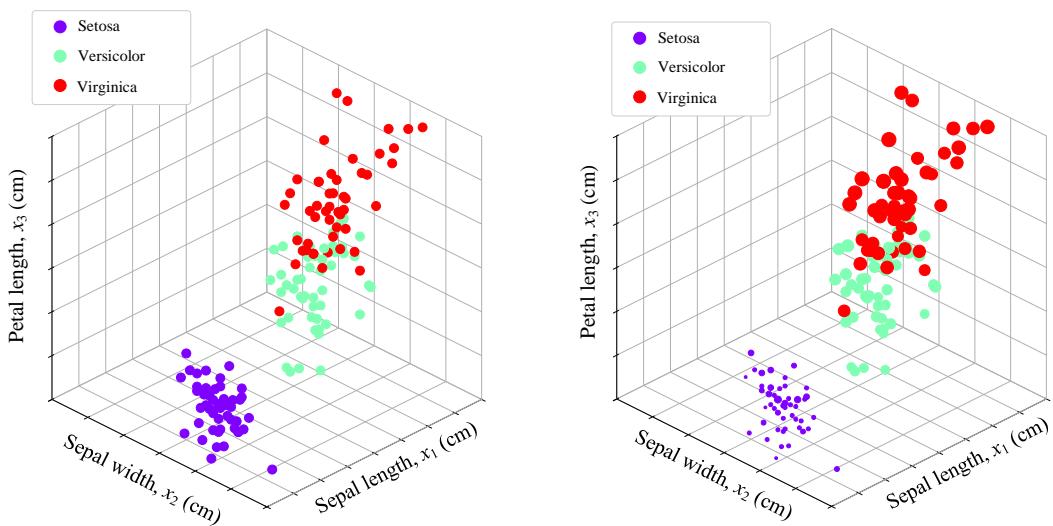


图 5. (a) 用颜色可视化鸢尾花类别；(b) 同时可视化鸢尾花萼宽度、类别

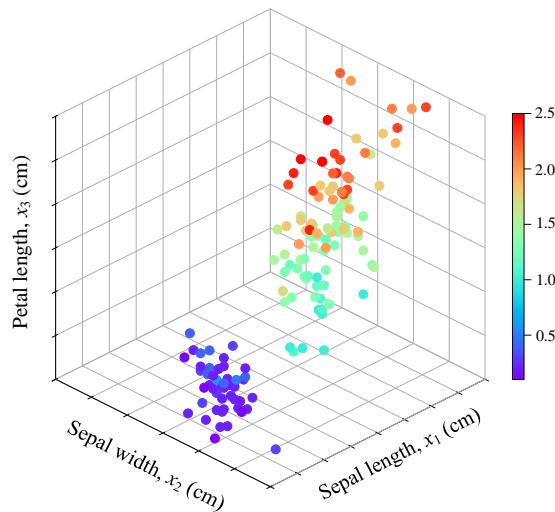


图 6. 用色谱可视化鸢尾花宽度



Jupyter 笔记 BK_2_Ch18_1.ipynb 绘制图 1 ~ 图 6。

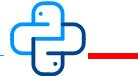
18.3 概率分布

多项分布

图 7 所示为用空间散点可视化多项分布。多项分布 (Multinomial Distribution) 是一种离散型概率分布，用于描述在多项试验中各个可能结果出现次数的概率分布。多项试验是指在一个试验中，每次试验有多个可能的结果，每个结果出现的概率是固定的。

图 7 所示的多项分布中， x_1 、 x_2 、 x_3 的取值范围为 $[0, 20]$ 区间内的整数。我们用散点的颜色代表多项分布的概率质量值。

《统计至简》第 5 章将讲解多项分布。



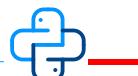
Jupyter 笔记 BK_2_Ch18_2.ipynb 绘制图 7。

高斯分布

三元高斯分布概率密度函数本质上是四维数据。如图 8 所示，分层散点图可以可视化三元高斯分布。

注意 x_1 、 x_2 、 x_3 的取值范围为 $(-\infty, \infty)$ 。

本书后文会用分层等高线可视化高斯分布。

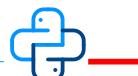


Jupyter 笔记 BK_2_Ch18_3.ipynb 绘制图 8。

Dirichlet 分布

Dirichlet 分布是一种概率分布，用于描述多维随机变量的概率分布。Dirichlet 分布通常用于处理多元分类和多元回归问题，是多项分布的共轭先验分布。

Dirichlet 分布的定义域是 D 维单位超立方体，即所有分量都在 $[0,1]$ 之间且它们之和等于 1。如图 9 所示， θ_1 、 θ_2 、 θ_3 的取值范围为 $[0, 1]$ 实数。



Jupyter 笔记 BK_2_Ch18_4.ipynb 绘制图 9。

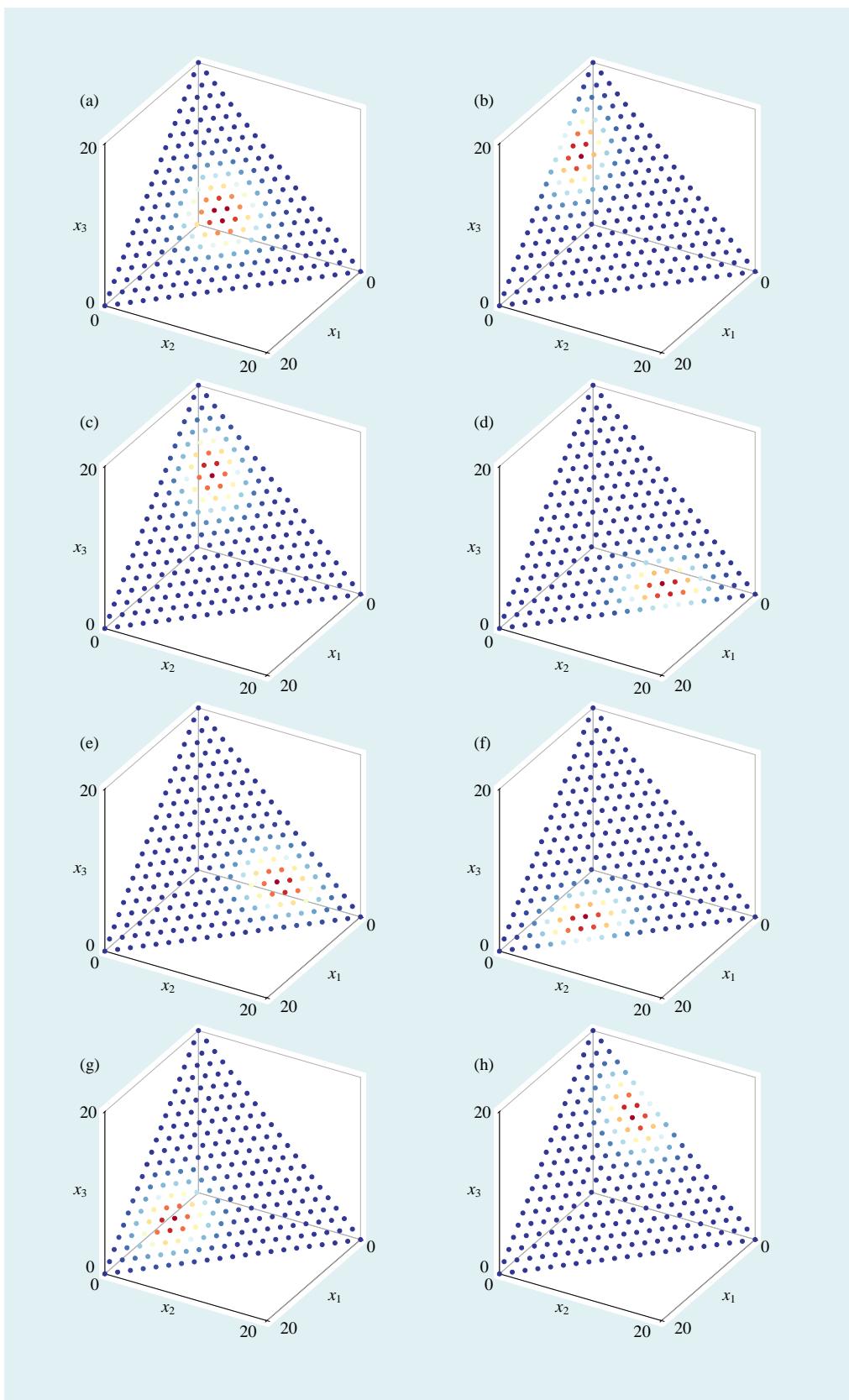


图 7. 用空间散点可视化多项分布

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

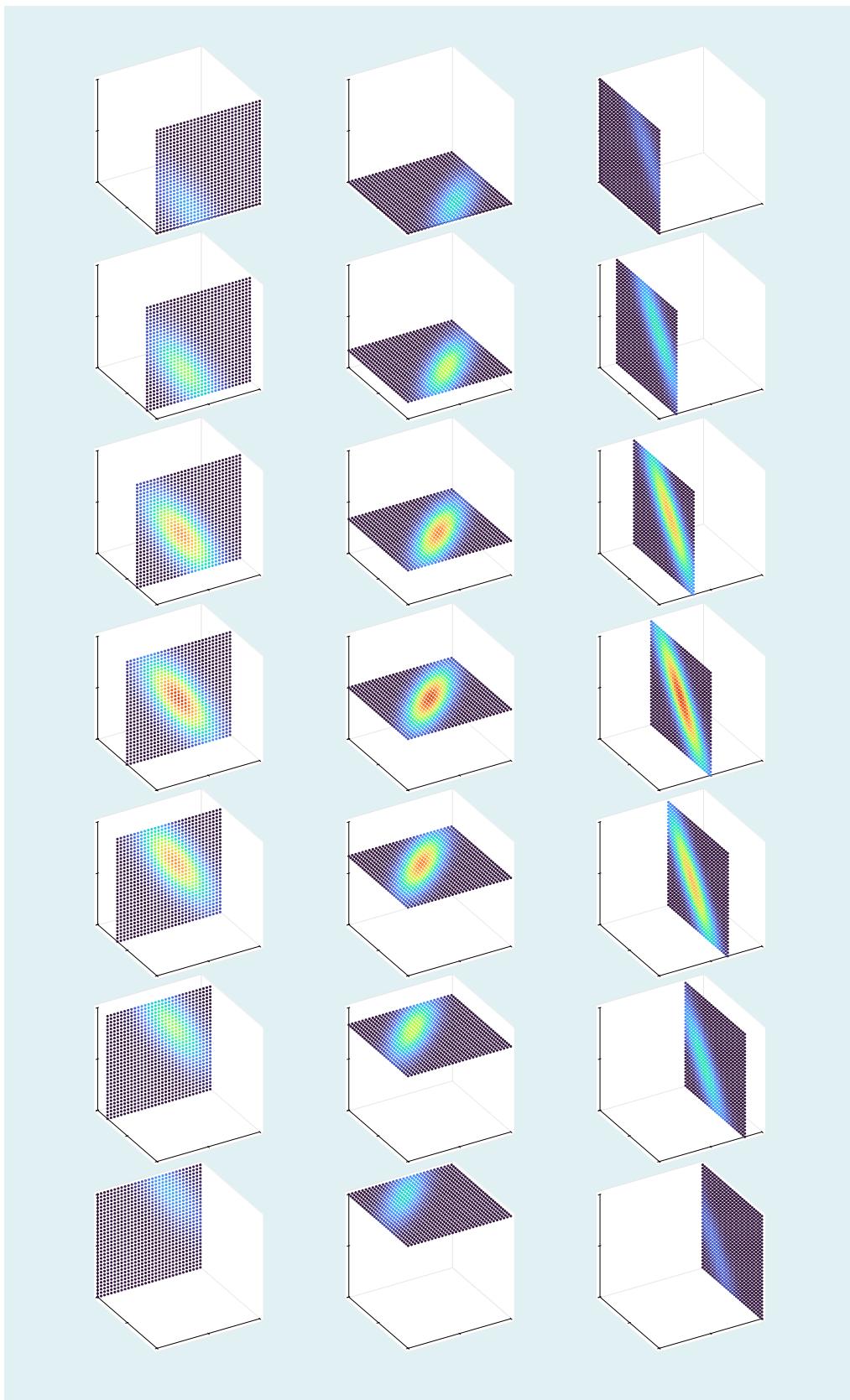


图 8. 用空间散点可视化高斯分布

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

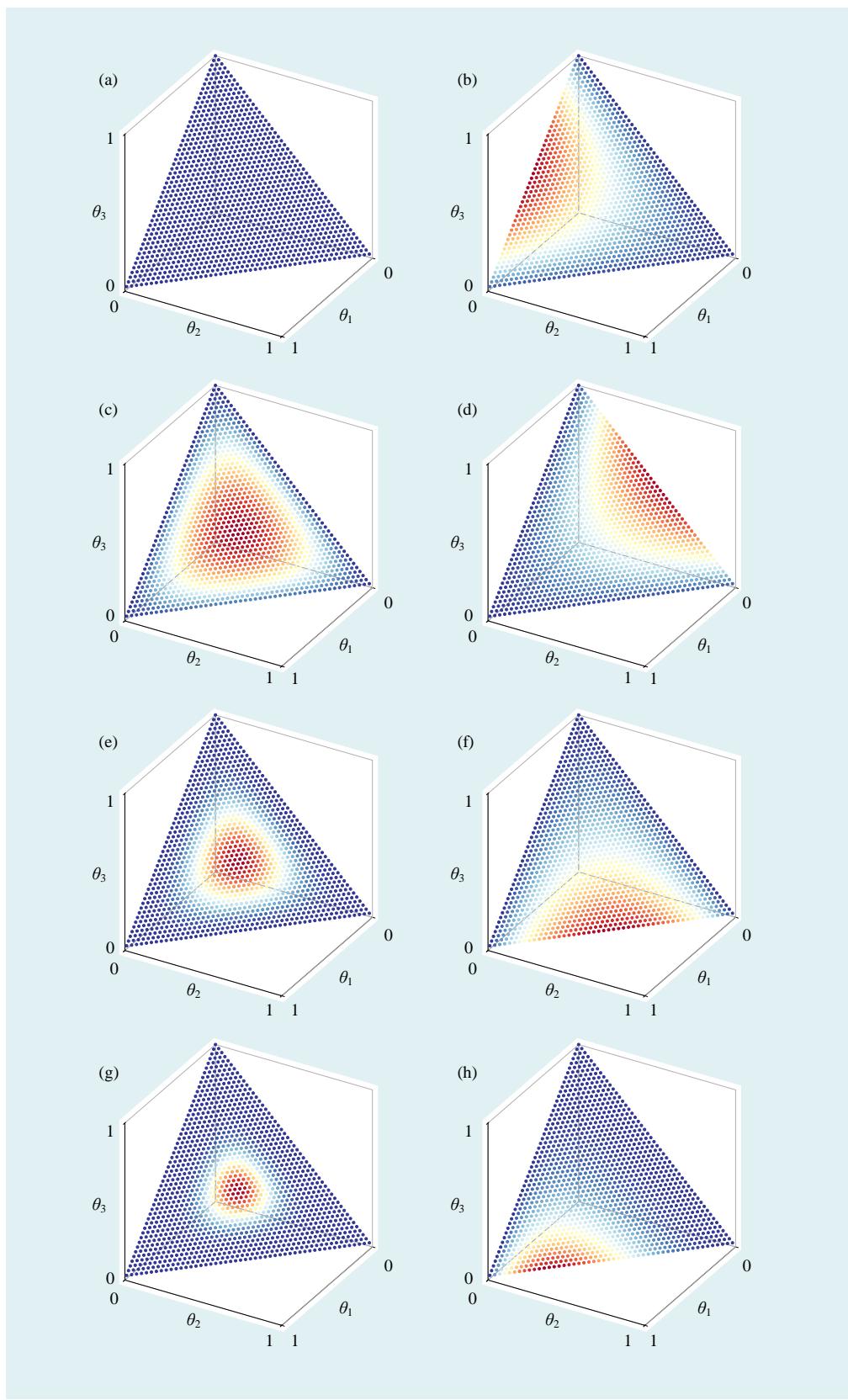


图 9. 用空间散点可视化 Dirichlet 分布

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

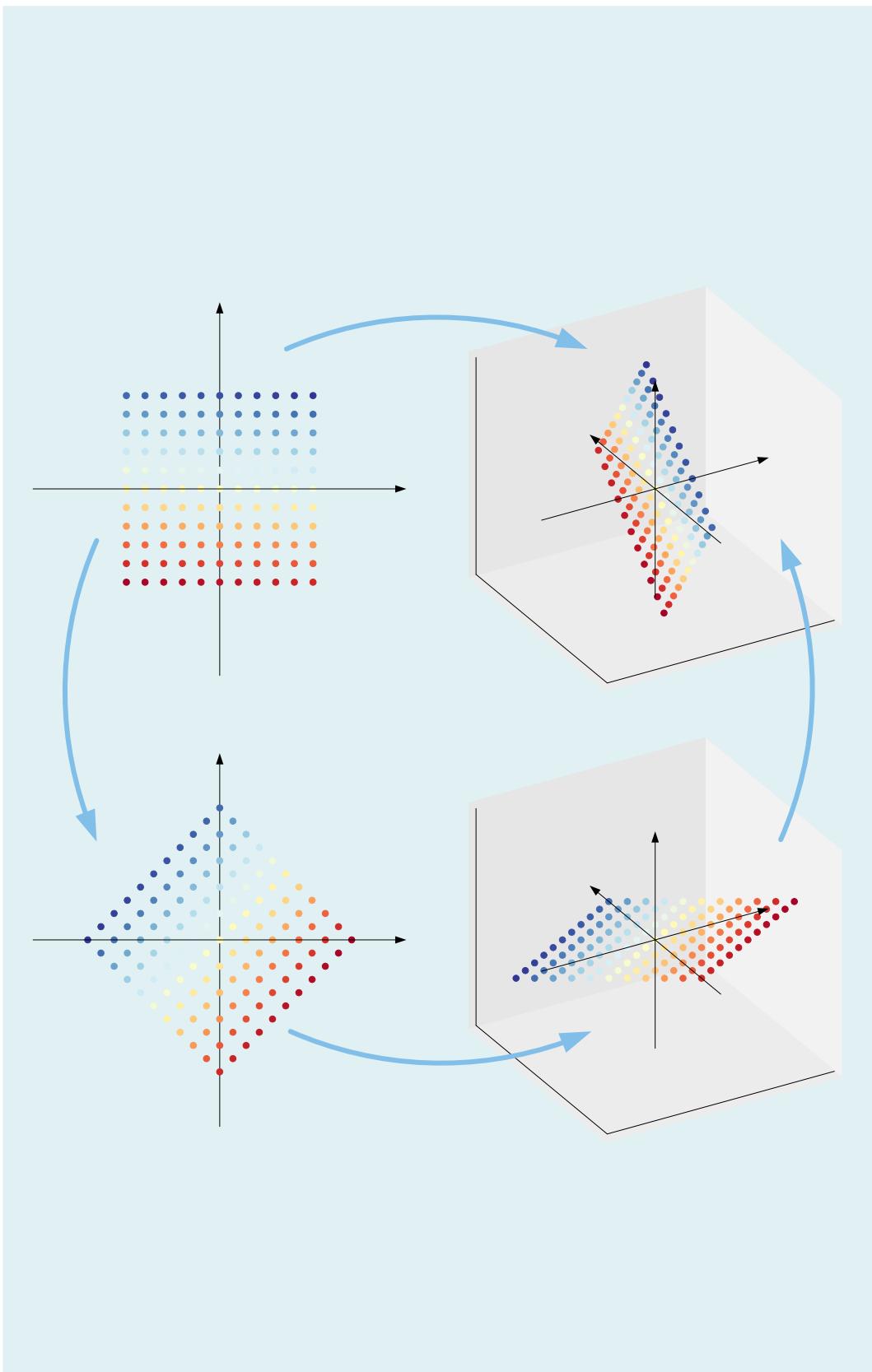


图 10. 可视化奇异值分解，细高矩阵，第 1 组

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

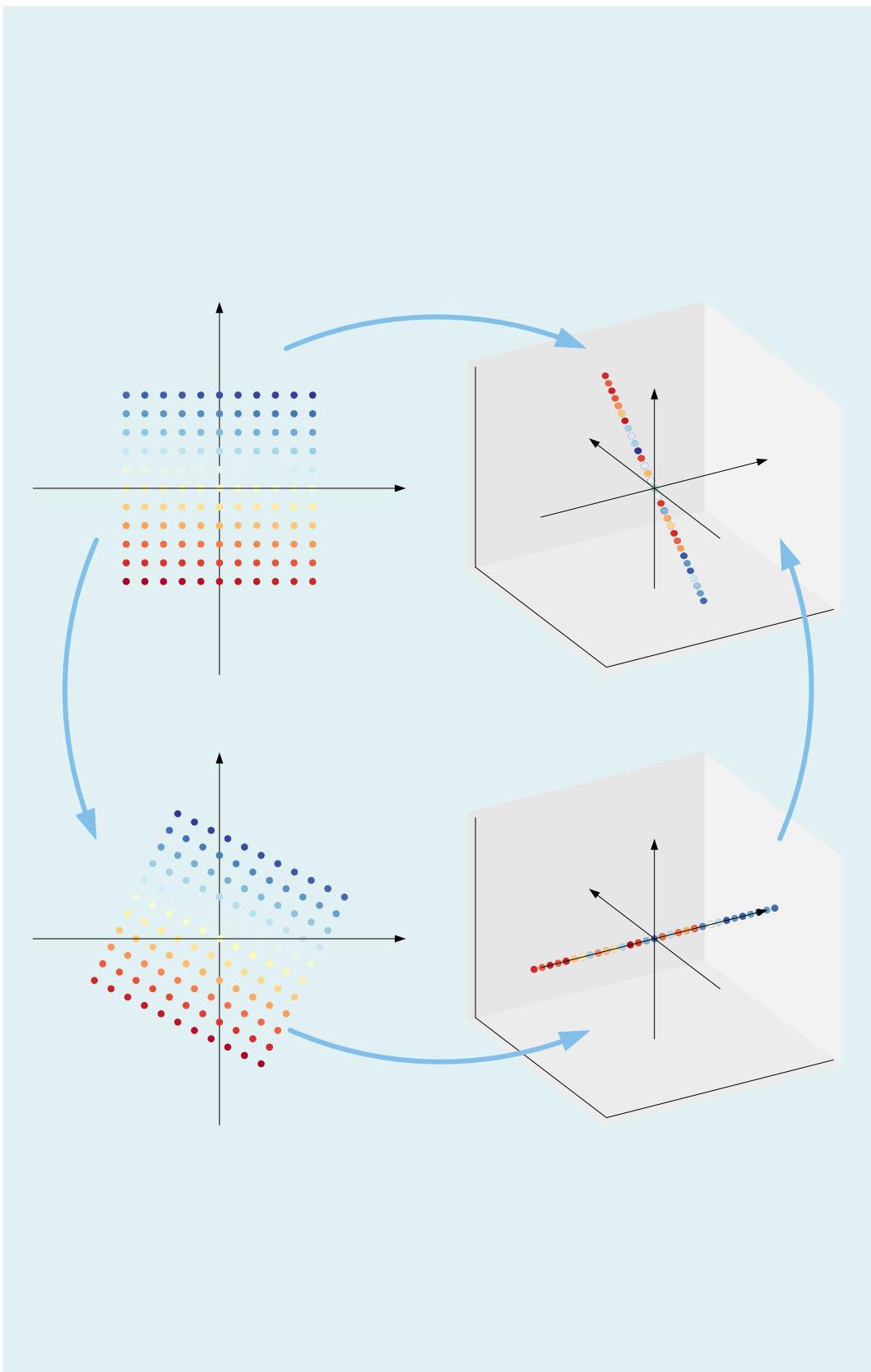


图 11. 可视化奇异值分解，细高矩阵，第 2 组

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

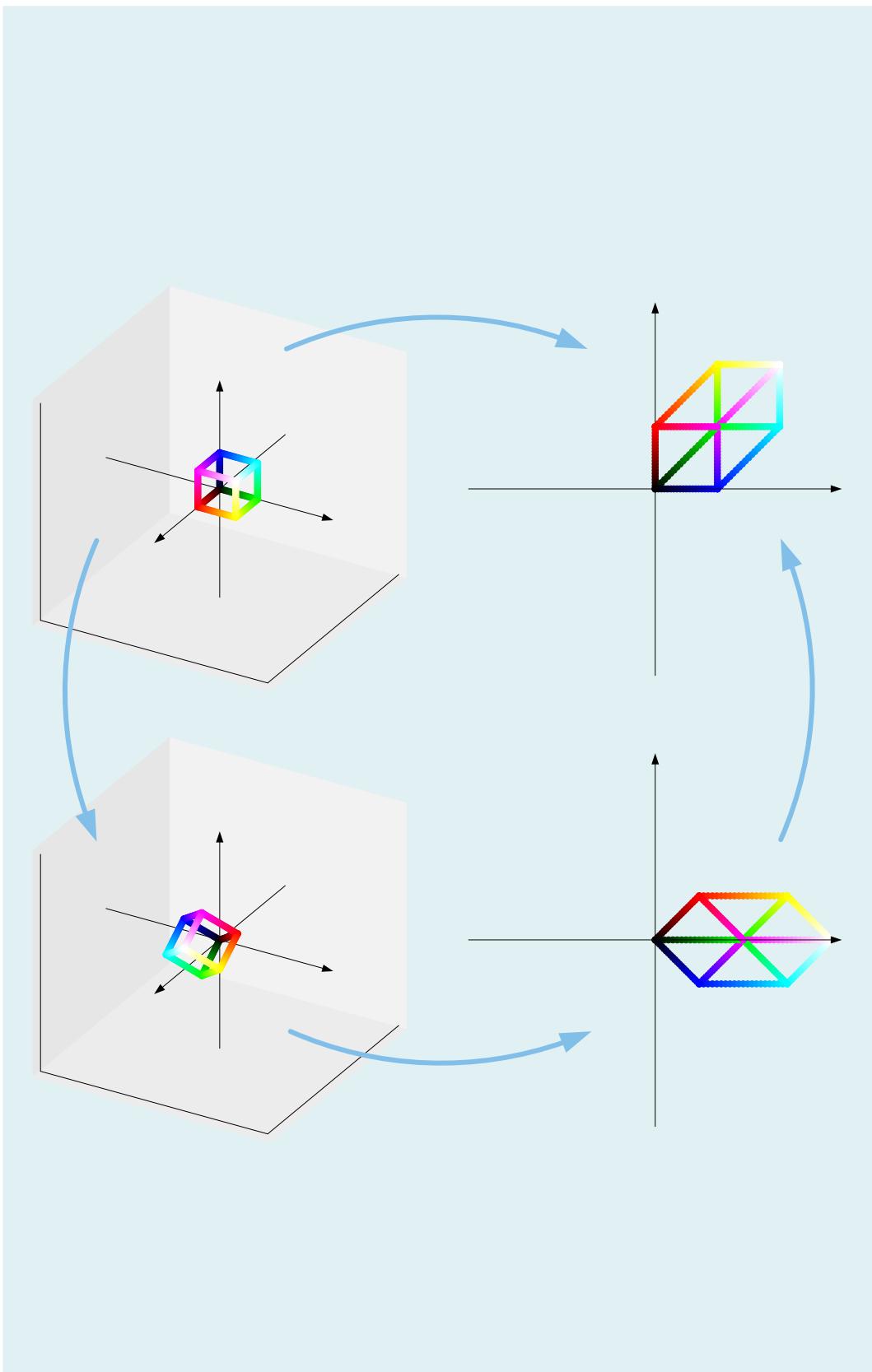


图 12. 可视化奇异值分解，矮胖矩阵，第 1 组

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

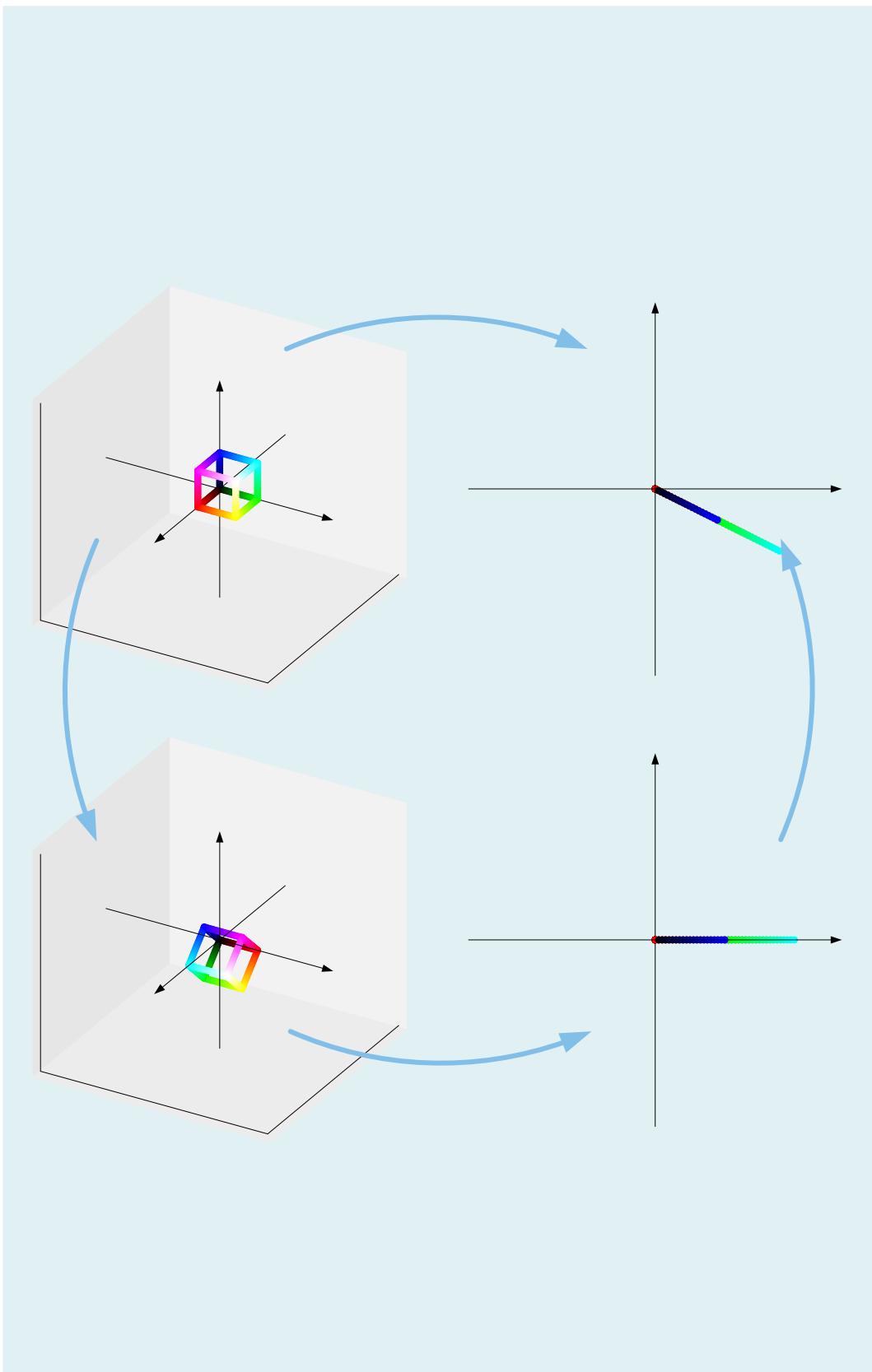


图 13. 可视化奇异值分解，矮胖矩阵，第 2 组

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

19

3D Line Plot

三维线图

将三维散点顺序连线



一，记住要仰望星空，而不是低头看脚。二、永不放弃工作。工作赋予你意义和目的，没有它，生活会变得空虚。三、如果你有幸找到爱情，记住它就在那里，要珍视。

One, remember to look up at the stars and not down at your feet. Two, never give up work. Work gives you meaning and purpose and life is empty without it. Three, if you are lucky enough to find love, remember it is there and don't throw it away.

—— 史蒂芬·霍金 (Stephen Hawking) | 英国理论物理学家、宇宙学家 | 1942 ~ 2018



- ◀ `matplotlib.pyplot.plot_wireframe()` 绘制线框图
- ◀ `matplotlib.pyplot.scatter()` 绘制散点图
- ◀ `matplotlib.pyplot.stem()` 绘制火柴梗图
- ◀ `matplotlib.pyplot.text()` 在图片上打印文字
- ◀ `numpy.arange()` 根据指定的范围以及设定的步长，生成一个等差数组
- ◀ `numpy.exp()` 计算括号中元素的自然指数
- ◀ `numpy.linspace()` 在指定的间隔内，返回固定步长的数据
- ◀ `numpy.meshgrid()` 创建网格化数据

19.1 线图

如图 1 所示，在 Matplotlib 中绘制的三维线图实际上也是三维散点顺序连线得到的“折线”。因此，在绘制三维线图时，大家也需要注意颗粒度的问题。本书前文已经介绍过平面线图中颗粒度这个话题，本章不再展开。

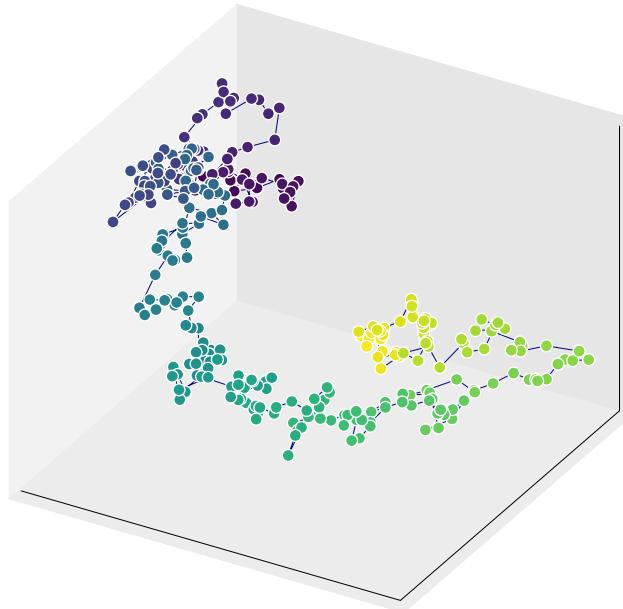


图 1. 用 Matplotlib 绘制微粒随机漫步线图，来自《编程不难》

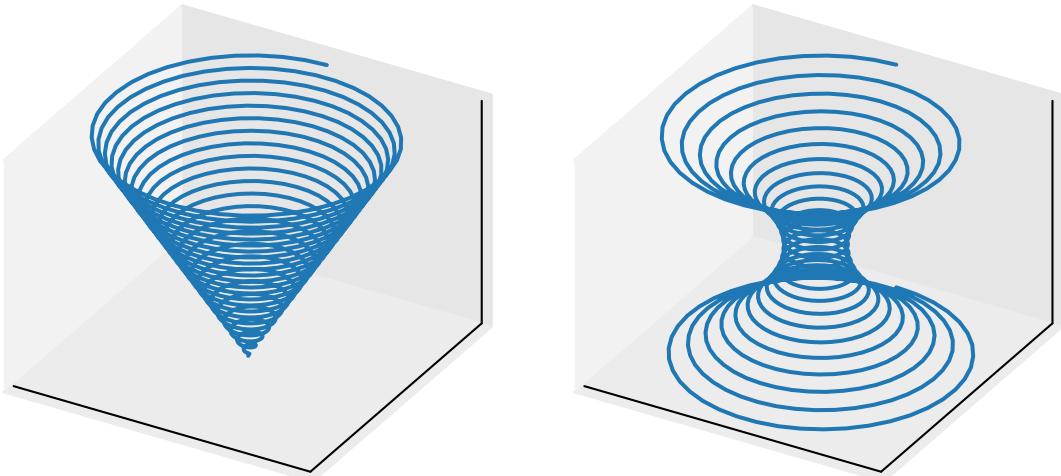
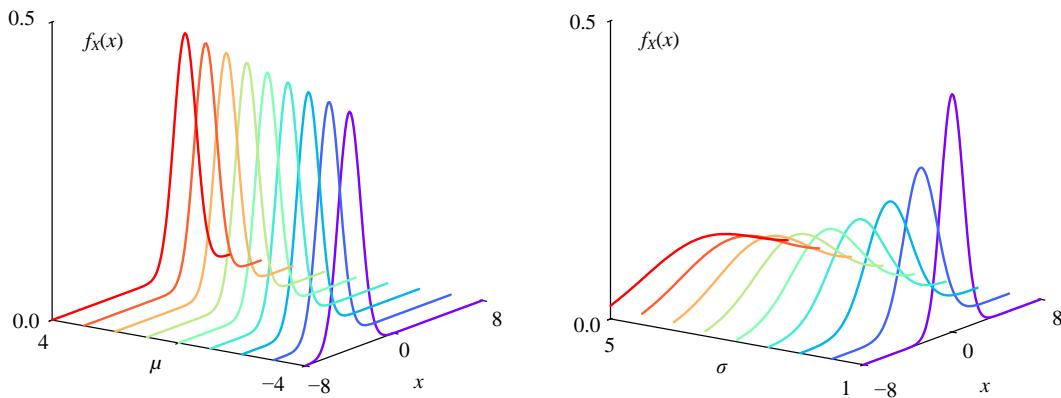


图 2. 颗粒度细腻的三维线图

渲染

`matplotlib.pyplot.plot()` 可以用来绘制平面线图，也可以用来绘制三维线图。

图 3 (a) 所示为一元高斯分布概率密度函数曲线随 μ 变化。图 3 (b) 所示为一元高斯分布概率密度函数曲线随 σ 变化。

图 3. 一元高斯密度函数分别随 μ 、 σ 变化

Jupyter 笔记 BK_2_Topic_5.02_1.ipynb 绘制图 3。

投影

类似上一个话题的散点图，我们也可以在三维空间的特定平面绘制三维线图。图 4 所示为两个例子。图 4 (a) 的蓝色线图绘制在 $x_1 = 3$ 平面上，而橘色线图绘制在 $x_2 = 3$ 平面上。图 4 (b) 的蓝色线图绘制在 $x_1 = -3$ 平面上，而橘色线图绘制在 $x_2 = -3$ 平面上。

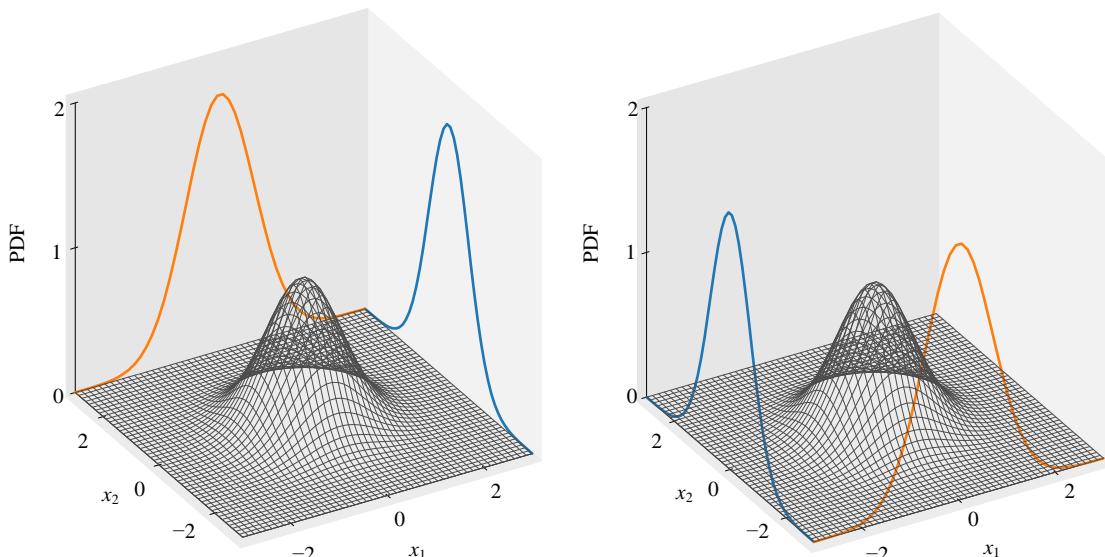


图 4. 投影到背平面、前平面



Jupyter 笔记 BK_2_Topic_5.02_2.ipynb 绘制图 4。

参考线

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

图 5 所示的单位立方体有 8 个顶点。我们可以用三维散点绘制这些顶点，用两点连线绘制这个单位立方体的 12 条边。

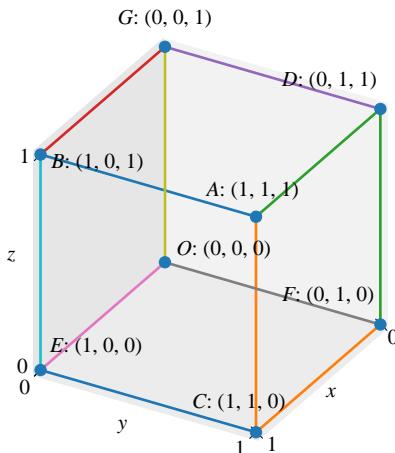


图 5. 单位正方体的 12 条边



Jupyter 笔记 BK_2_Topic_5.02_3.ipynb 绘制图 5。

19.2 火柴梗图

类似平面直角坐标系，在三维直角坐标系中我们也可以用火柴梗图可视化二元离散函数。图 6 (a) 所示为用火柴梗图可视化多项分布。火柴梗图也可以调整投影方便，如图 6 (b) 所示。

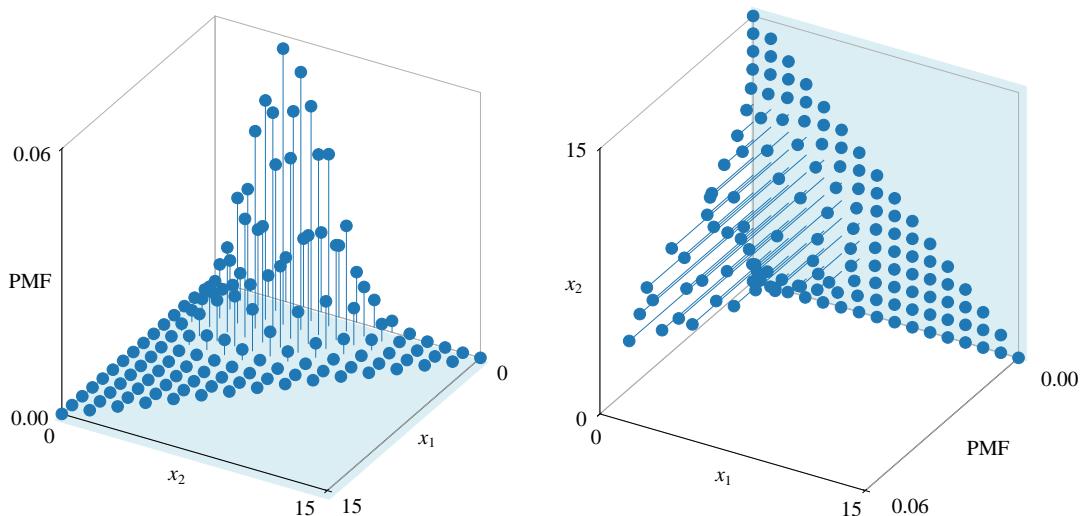


图 6. 沿 z 轴、x 轴方向的火柴梗图



Jupyter 笔记 BK_2_Topic_5.02_3.ipynb 绘制图 6。

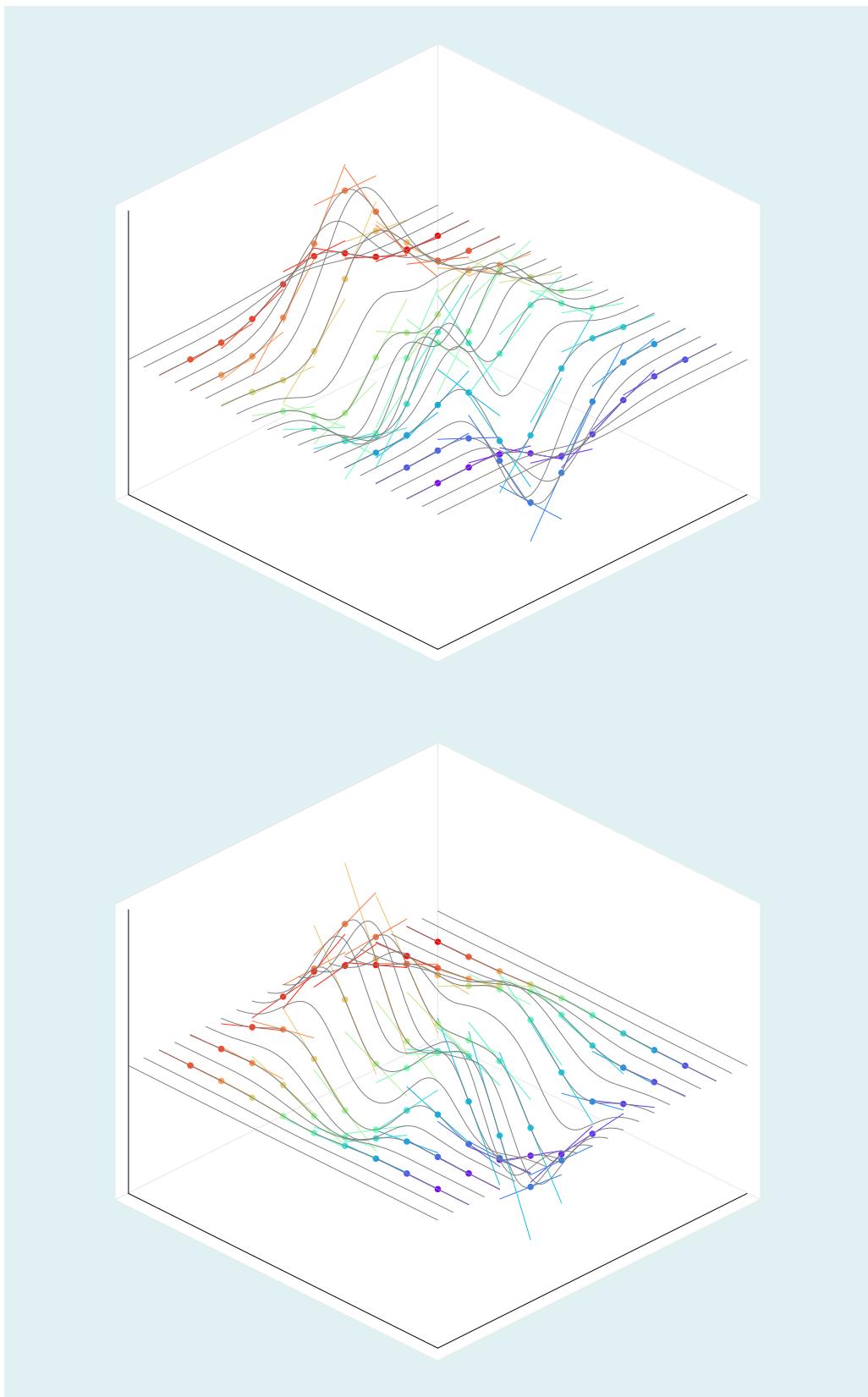


图 7. 可视化偏导数

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

20 Mesh Surface 网格曲面

沿着两个维度的曲线织成的曲面



画家把太阳画成黄点；艺术家把黄点变成太阳。

There are painters who transform the sun to a yellow spot, but there are others who with the help of their art and their intelligence, transform a yellow spot into sun.

—— 毕加索 (Pablo Picasso) | 西班牙艺术家 | 1881 ~ 1973



- ◀ `matplotlib.pyplot.contour()` 绘制等高线图
- ◀ `matplotlib.pyplot.contourf()` 绘制平面填充等高线
- ◀ `matplotlib.pyplot.plot_wireframe()` 绘制线框图
- ◀ `matplotlib.pyplot.scatter()` 绘制散点图
- ◀ `numpy.diag()` 如果 A 为方阵, `numpy.diag(A)` 函数提取对角线元素, 以向量形式输入结果; 如果 a 为向量, `numpy.diag(a)` 函数将向量展开成方阵, 方阵对角线元素为 a 向量元素
- ◀ `numpy.dot()` 计算向量标量积。值得注意的是, 如果输入为一维数组, `numpy.dot()` 输出结果为标量积; 如果输入为矩阵, `numpy.dot()` 输出结果为矩阵乘积, 相当于矩阵运算符`@`
- ◀ `numpy.linalg.det()` 计算行列式值
- ◀ `numpy.linalg.inv()` 矩阵求逆
- ◀ `numpy.meshgrid()` 创建网格化数据
- ◀ `numpy.outer()` 计算两个向量的外积、张量积
- ◀ `numpy.sqrt()` 计算平方根
- ◀ `numpy.vstack()` 返回竖直堆叠后的数组
- ◀ `numpy.zeros_like()` 用来生成和输入矩阵形状相同的零矩阵
- ◀ `scipy.stats.dirichlet.pdf()` 计算 Dirichlet 分布的概率密度函数
- ◀ `scipy.stats.multivariate_normal.pdf()` 多元高斯分布概率密度函数
- ◀ `sympy.diff()` 求解符号导数和偏导解析式
- ◀ `sympy.exp()` 符号自然指数
- ◀ `sympy.lambdify()` 将符号表达式转化为函数

本 PDF 文件为作者草稿, 发布目的为方便读者在移动终端学习, 终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有, 请勿商用, 引用请注明出处。

代码及 PDF 文件下载: <https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger: <https://space.bilibili.com/513194466>

欢迎大家批评指教, 本书专属邮箱: jiang.visualize.ml@gmail.com

20.1 网格曲面

颗粒度

在绘制网格曲面时，我们也会碰到颗粒度问题。如图 1 (a) 所示，当网格稀疏时，生成的网格曲面很粗糙。

另外一个极端，如图 1 (b) 所示，当颗粒度过高时，生成的网格过于绵密，虽然线条变得光滑很多，但是整个曲面变化趋势的辨识度反而降低。

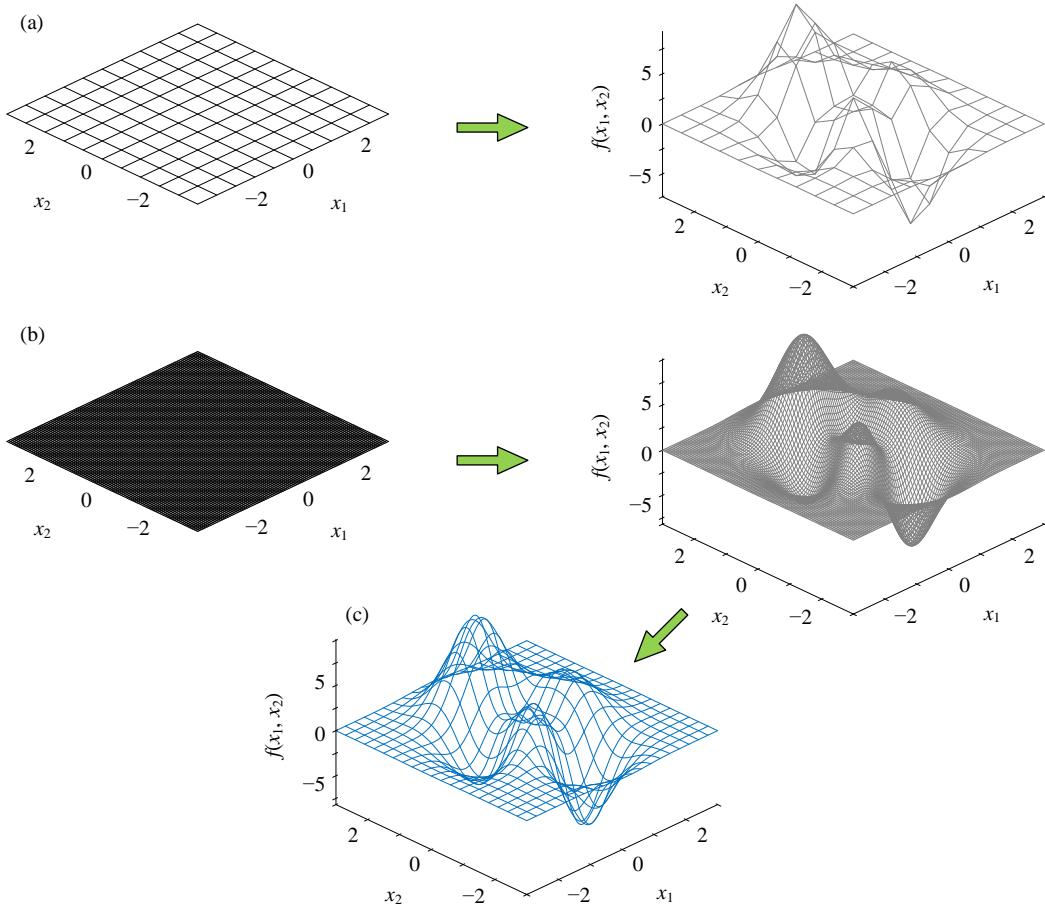


图 1. 网格颗粒度

解决这个问题的办法很简单，在使用 `Axes3D.plot_wireframe()` 绘制网格曲面时，可以使用图 1 (b) 这种颗粒度很高的网格面，同时设置 `rstride`、`cstride` 来调节步幅。如图 1 (c) 所示，增大单一维度上的步幅，可以保证线条的光滑程度，但是网格面看上去更清爽。

注意，使用 `Axes3D.plot_wireframe()` 时，如果不提供 `rstride`、`cstride`，函数会自动设置步幅。但是，为了保证质量可控，建议大家主动设置 `rstride`、`cstride`。

绘制沿特定方向曲线

对于 `Axes3D.plot_wireframe()`，我们可以分别将 `rstride`、`cstride` 设置为 0，从而绘制沿单一方向曲线，如图 2 (a)、(c) 所示。这种可视化方案很适合分析二元函数。这两种曲面可以投影在平面上，如图 2 (b)、(d) 所示。

进一步变形

在网格曲面基础之上，我们还可以绘制并强调特定曲线，如图 2 (e)、(f) 所示。在网格曲面上，我们可以绘制等高线，也可以绘制散点（如图 2 (g)、(h) 所示）。



Jupyter 笔记 BK_2_Ch20_1.ipynb 绘制图 1 所有子图。

20.2 在三维平面展示四维数据

一般情况下，用 `Axes3D.plot_surface()` 函数绘制三维曲面 $f(x, y)$ 时，渲染曲面的颜色也会根据 $f(x, y)$ 取值，如图 3 (a) 所示。

如果，渲染三维曲面 $f(x, y)$ 时采用另外一组数据 $V(x, y)$ ，我们便得到类似图 3 (b) 这幅图。

举个例子， (x, y) 代表经纬度，三维曲面 $f(x, y)$ 代表一座山峰的海拔高度，而 $V(x, y)$ 代表山峰不同位置某个时刻的温度值。

反过来，我们也可以用 $V(x, y)$ 构造曲面，而用 $f(x, y)$ 作为依据渲染曲面，具体如图 3 (c) 所示。



Jupyter 笔记 BK_2_Ch20_2.ipynb 绘制图 3 所有子图。

Dirichlet 分布

图 4 所示为用这种方案可视化 Dirichlet 分布。这个分布 $\theta_1, \theta_2, \theta_3$ 的取值范围都是 $[0, 1]$ ，且满足 $\theta_1 + \theta_2 + \theta_3 = 1$ 。给定不同分布参数 $\alpha_1, \alpha_2, \alpha_3$ ，将不同位置 Dirichlet 分布概率密度值映射到 $\theta_1 + \theta_2 + \theta_3 = 1$ 平面上，我们便得到图 4 图像。

→ 《统计至简》第 7 章将专门讲解 Dirichlet 分布。



Jupyter 笔记 BK_2_Ch20_3.ipynb 绘制图 4 所有子图。

瑞利商

类似地，我们可以用图 5 可视化瑞利商。瑞利商在矩阵特征值和特征向量的计算中有广泛的应用，是线性代数中的一个重要概念。简单来说，在 $x_1^2 + x_2^2 + x_3^2 = 1$ 这个单位球体上，红色代表瑞利商大，而蓝色代表瑞利商小。

→ 《矩阵力量》第 14 章将专门讲解瑞利商。



Jupyter 笔记 BK_2_Ch20_4.ipynb 绘制图 5 所有子图。

20.3 剖面线

绘制剖面

我们可以用 `Axes3D.plot_wireframe()` 绘制剖面。如图 6 所示，剖面可以平行 xy 、 xz 、 yz 平面。结合线图，这些剖面可以用来可视化剖面线。



Jupyter 笔记 BK_2_Ch20_5.ipynb 绘制图 6 子图。

绘制剖面线

如图 7 (a) 所示，这幅图中有几个重要元素：网格曲面、特定高度等高线、剖面。从数学角度来看，浅蓝色剖面切割网格曲面的结果是红色曲线。

图 7 (b)、(c) 两幅子图中的红色剖面线则是用线图绘制。图 7 (b) 的剖面位于 $y = 0$ ，红色剖面线则展示当 $y = 0$ 时，函数 $f(x, y)$ 随 x 变化。图 7 (c) 的剖面位于 $x = 0$ ，红色剖面线则展示当 $x = 0$ 时，函数 $f(x, y)$ 随 y 变化。



Jupyter 笔记 BK_2_Ch20_6.ipynb 绘制图 7 子图。

20.4 平面填充

图 8 两个子图所示为二元高斯分布的概率密度函数曲面。为了可视化 x_1 、 x_2 分别取不同值时函数曲线下方的面积，我们可以采用 `Axes3D.add_collection3d()` 函数在三维空间可视化填充对象。

→ 《统计至简》第 10 章将专门讲解二元高斯分布。



Jupyter 笔记 BK_2_Ch20_7.ipynb 绘制图 8 子图。

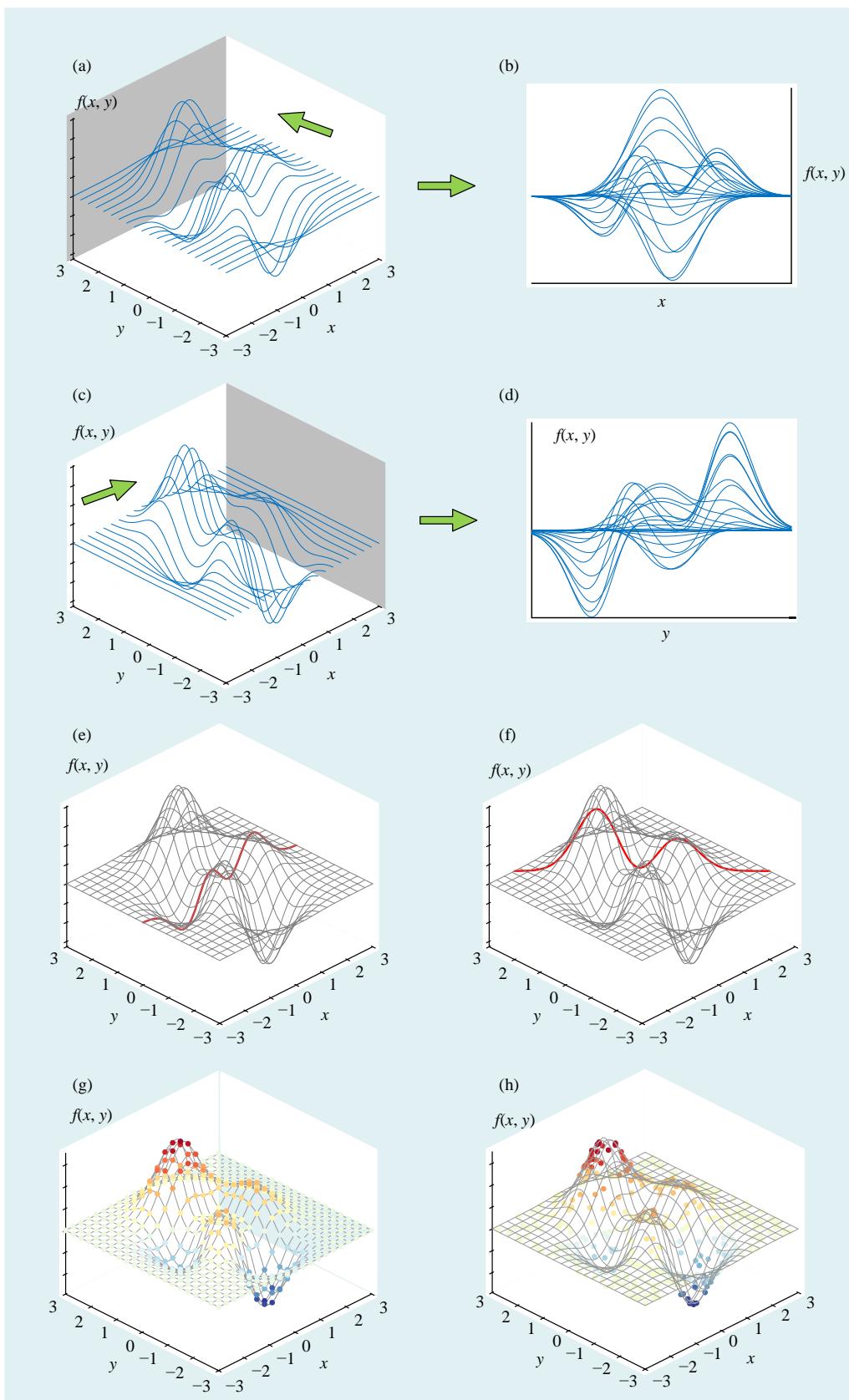


图 2. 网格曲面的进一步变形

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

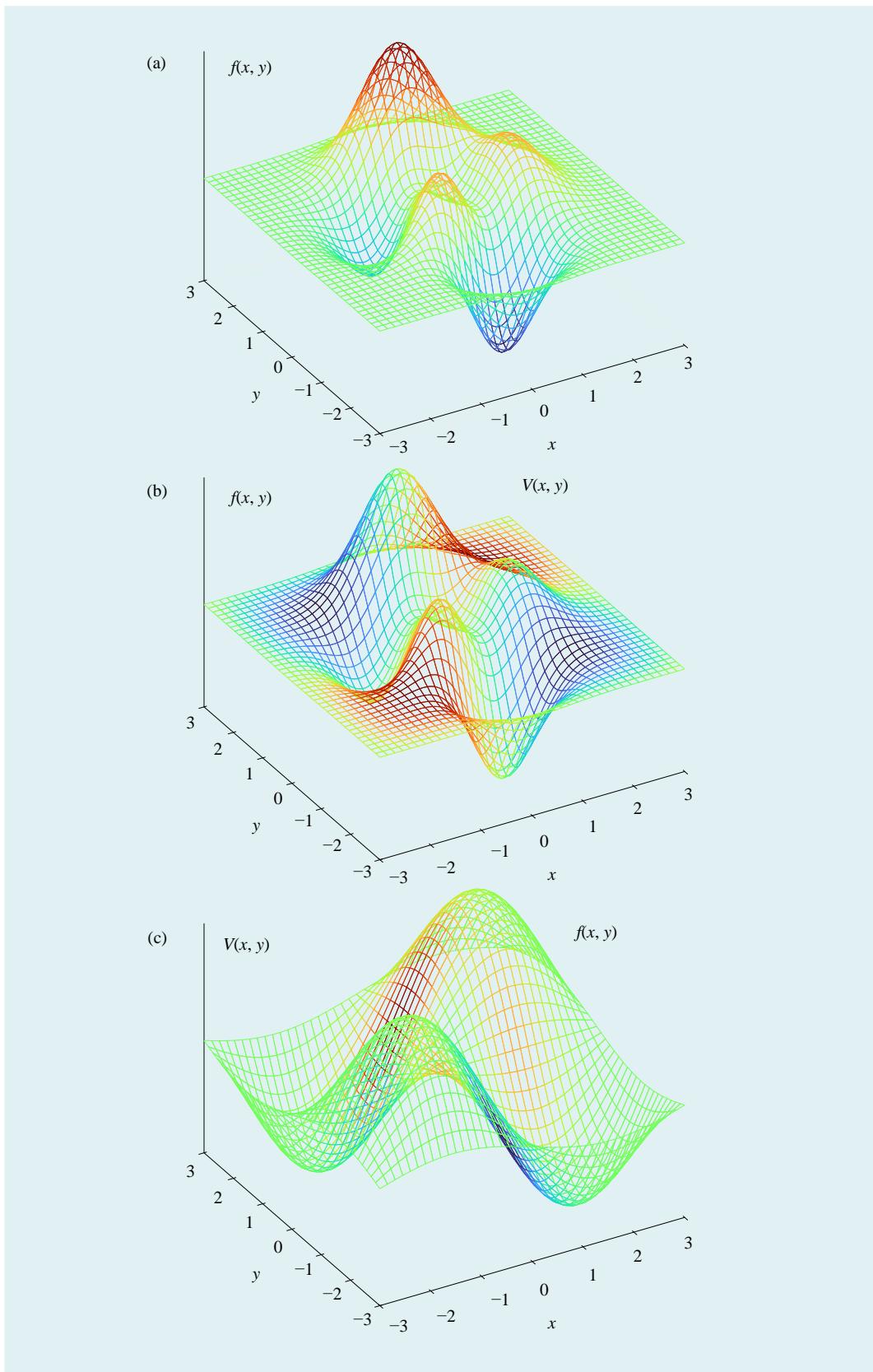


图 3. 渲染三维曲面

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

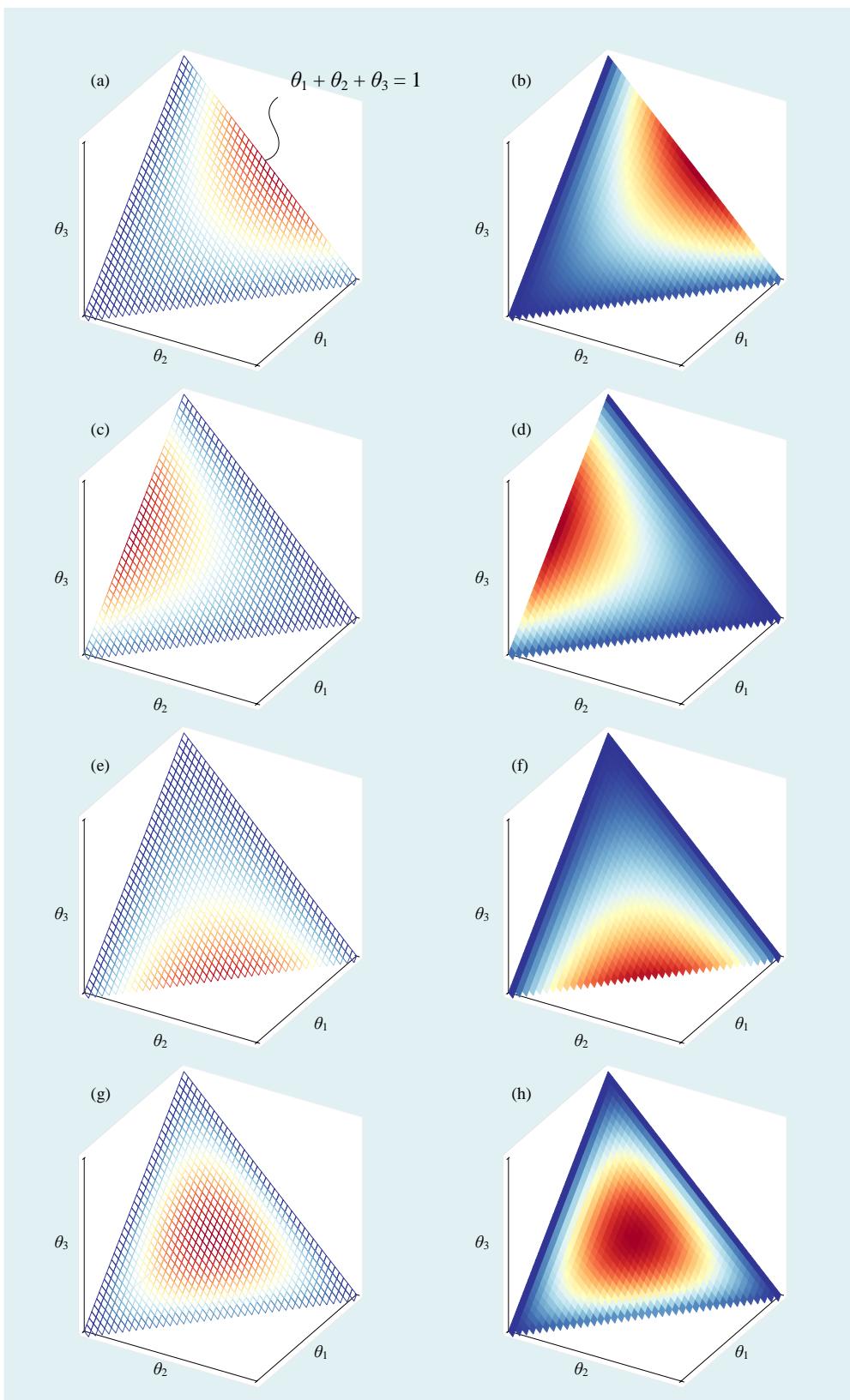


图 4. Dirichlet 分布

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

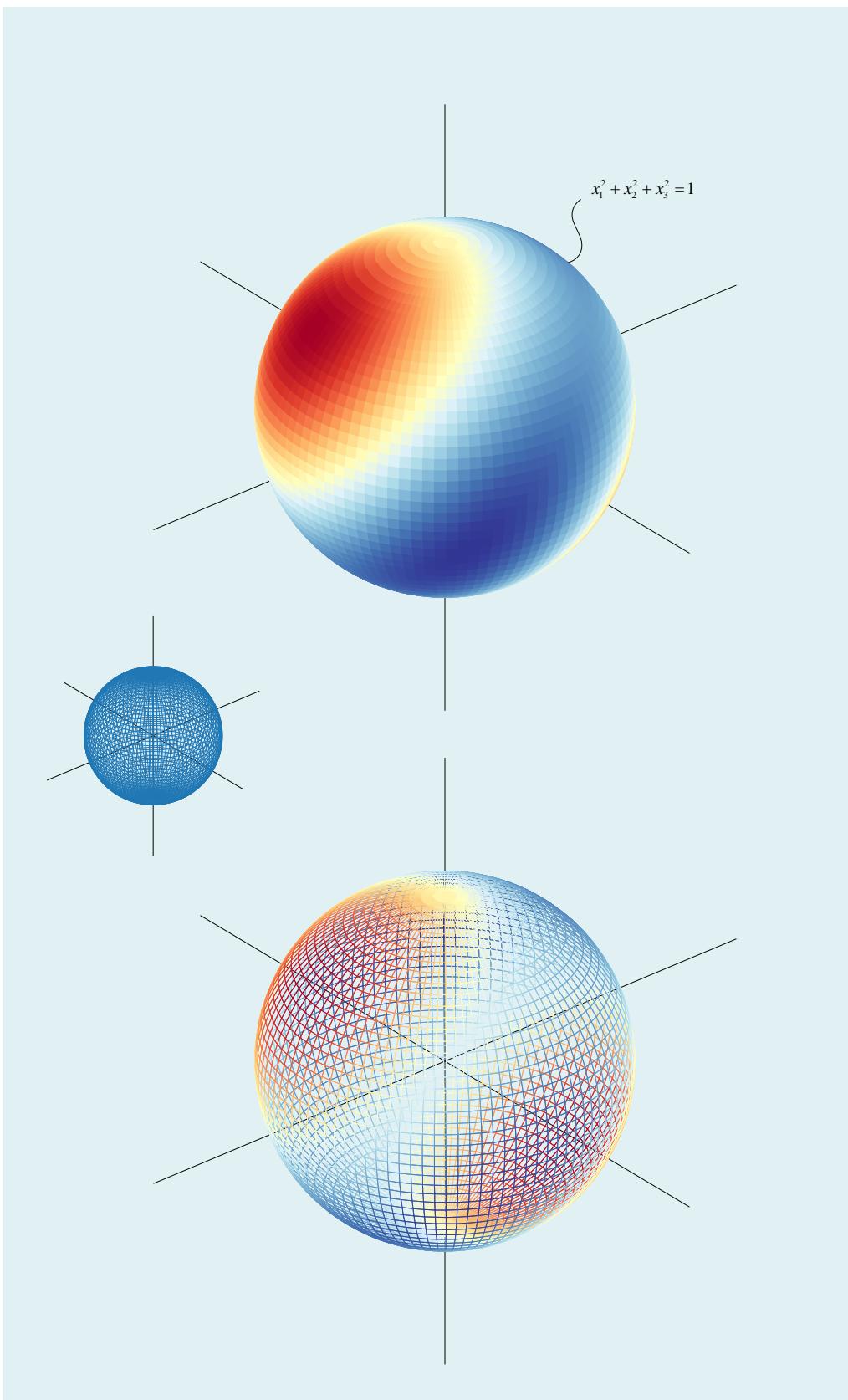


图 5. 瑞利商

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

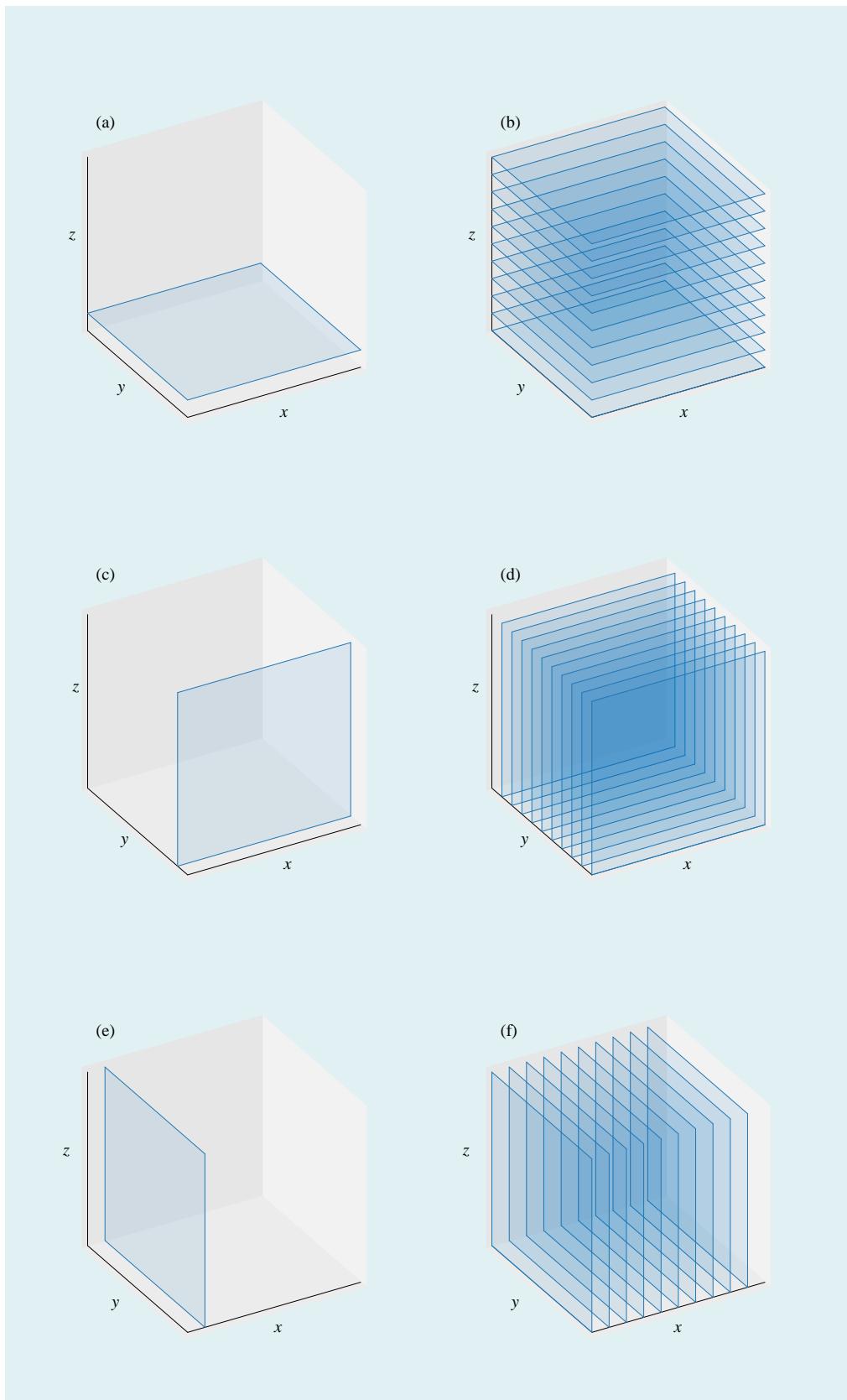


图 6. 平行于不同平面的剖面

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

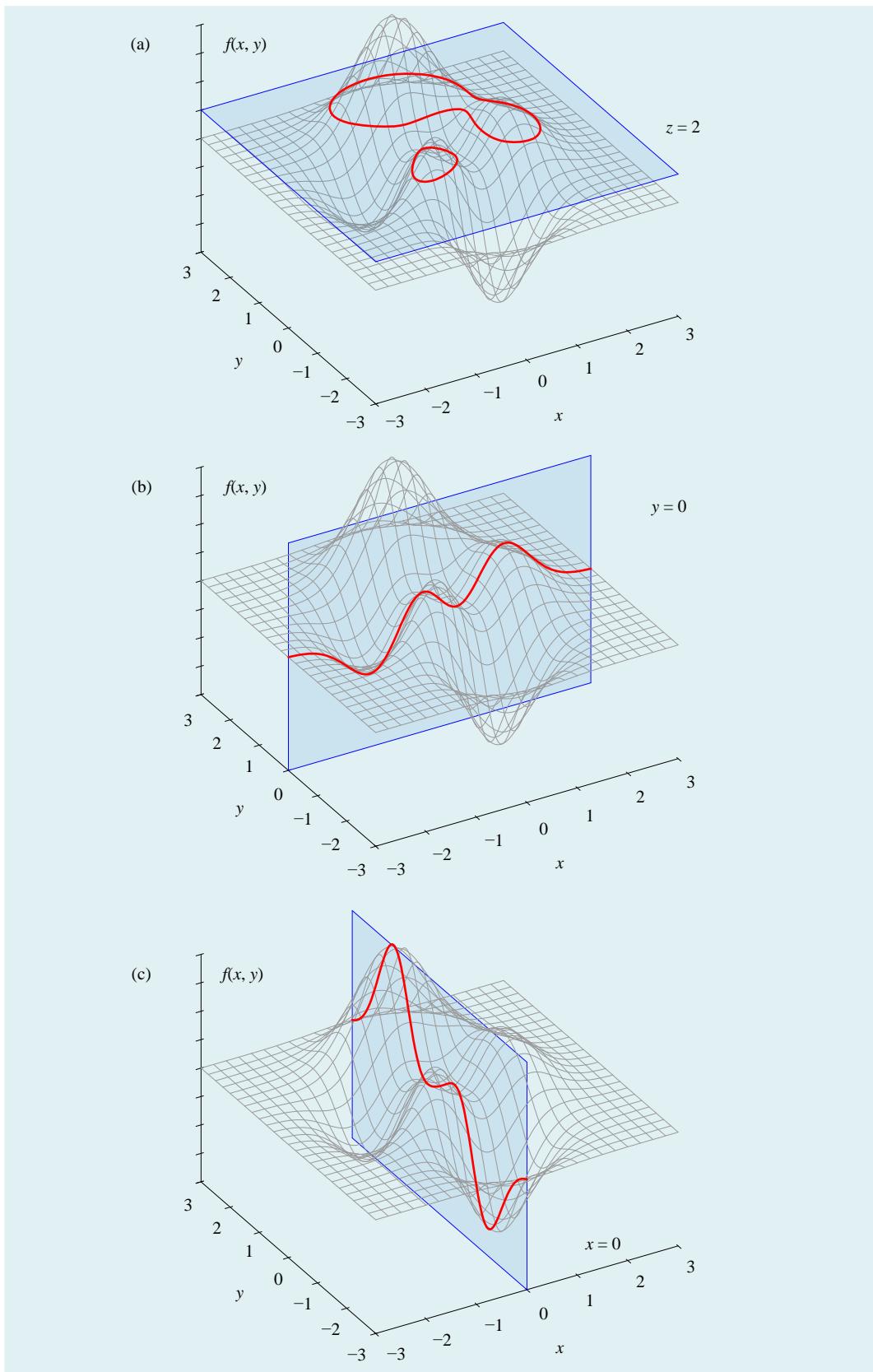


图 7. 二元函数在三个不同剖面上的剖面线

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

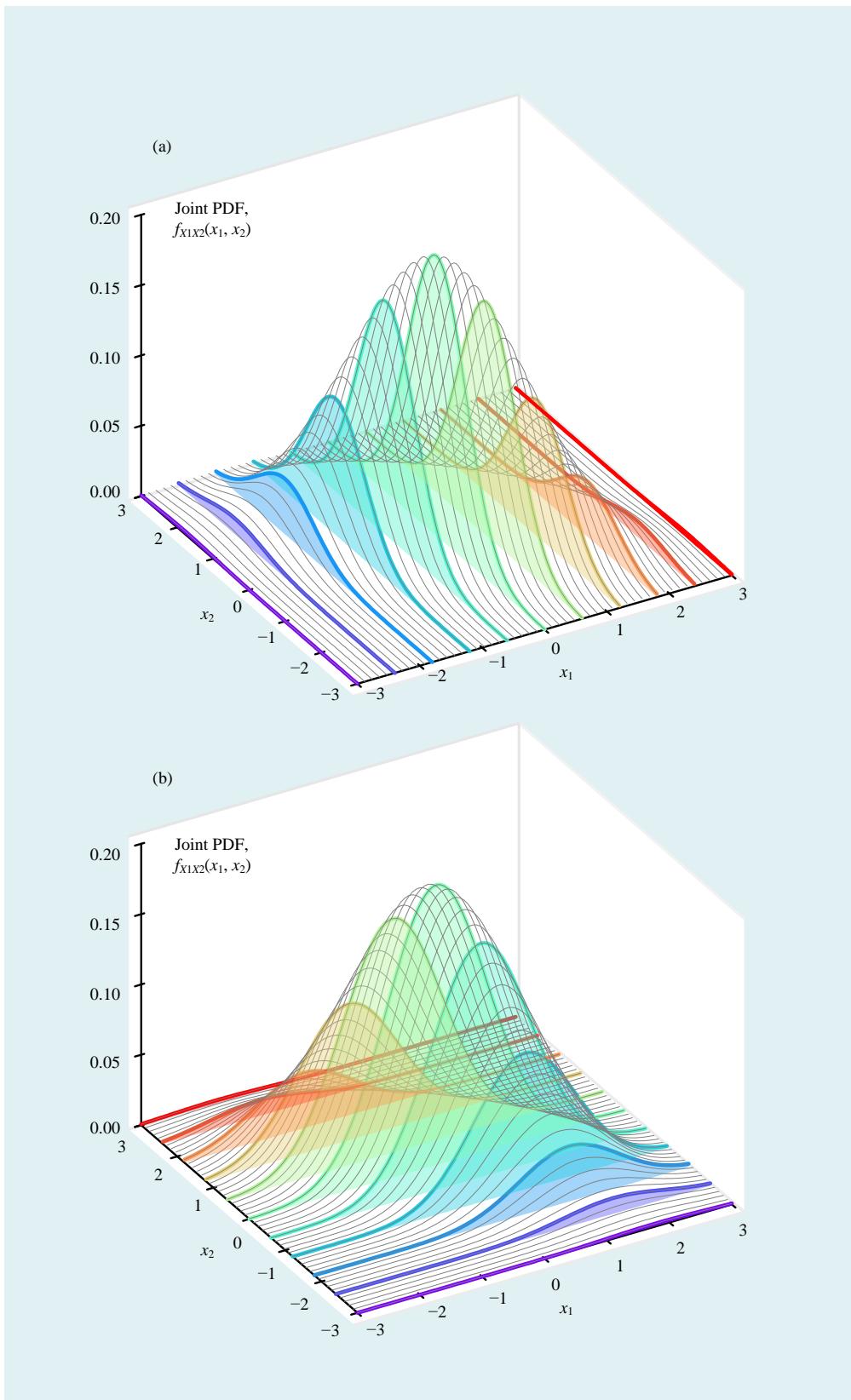


图 8. 三维线图的平面填充

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

21 3D Contours 三维等高线

可视化应用极为灵活



真正的艺术家不是被启发，而是启发别人。

A true artist is not one who is inspired, but one who inspires others.

—— 萨尔瓦多达利 (Salvador Dali) | 西班牙超现实主义画家 | 1904 ~ 1989



- ◀ `matplotlib.pyplot.contour()` 绘制等高线图
- ◀ `matplotlib.pyplot.contourf()` 绘制平面填充等高线
- ◀ `matplotlib.pyplot.plot_wireframe()` 绘制线框图
- ◀ `numpy.dot()` 计算向量标量积。值得注意的是，如果输入为一维数组，`numpy.dot()` 输出结果为标量积；如果输入为矩阵，`numpy.dot()` 输出结果为矩阵乘积，相当于矩阵运算符`@`
- ◀ `numpy.linalg.det()` 计算行列式值
- ◀ `numpy.linalg.inv()` 矩阵求逆
- ◀ `numpy.meshgrid()` 创建网格化数据
- ◀ `numpy.sqrt()` 计算平方根
- ◀ `numpy.vstack()` 返回竖直堆叠后的数组
- ◀ `sympy.diff()` 求解符号导数和偏导解析式
- ◀ `sympy.exp()` 符号自然指数
- ◀ `sympy.lambdify()` 将符号表达式转化为函数

21.1 沿三个方向获取等高线

如图 1 所示，Matplotlib 中三维空间等高线和填充等高线实际上可以指定三个不同方向。下面，我们分别介绍这三种不同获取等高线的方向。

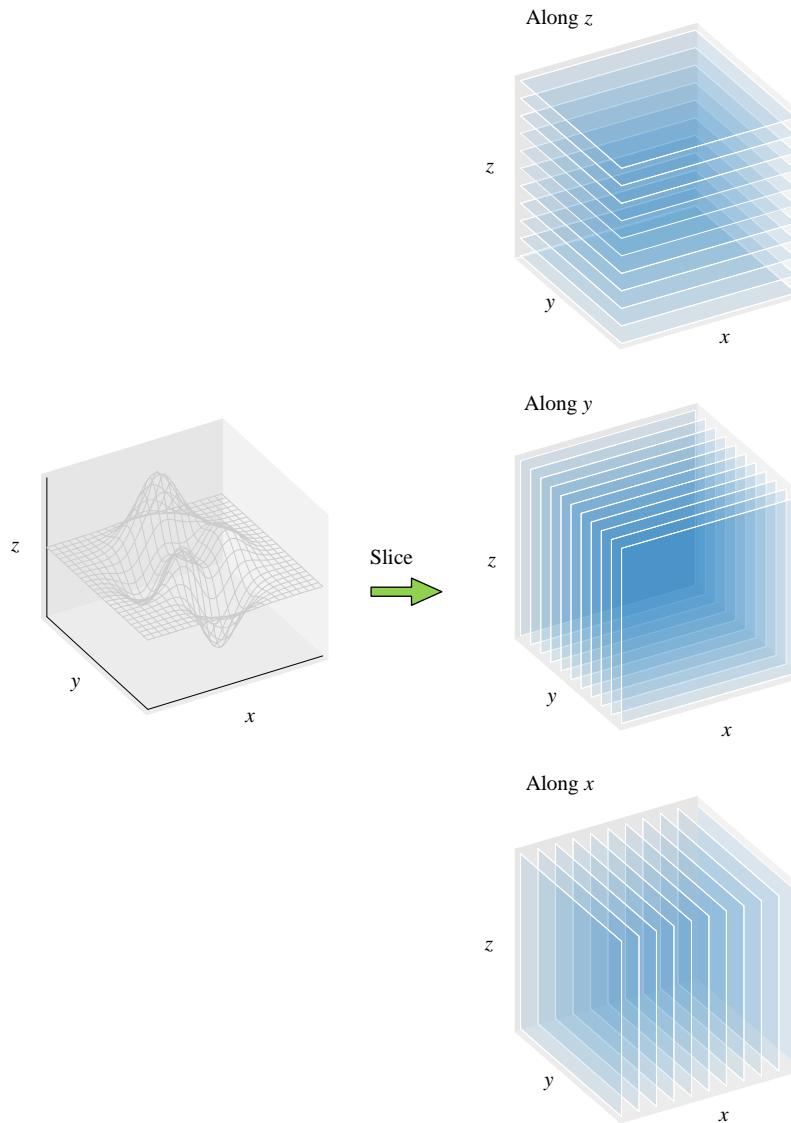
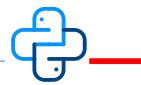


图 1. 三种不同切取等高线的方式

沿 z 方向

大家已经非常熟悉的是其默认竖直方向，即 z 方向，具体如图 6 (a)、(b) 所示。

此外，`matplotlib.pyplot.contour()` 和 `matplotlib.pyplot.contourf()` 还可以通过设置 `offset` 指定绘制所有等高线的具体高度。图 6 剩下几幅子图绘制等高线高度不同。



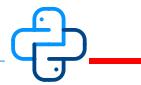
Jupyter 笔记 BK_2_Topic_5.06_1.ipynb 绘制图 6 所有子图。

沿 x 方向

设置 `zdir='x'`, 我们可以绘制沿 x 轴方向的等高线, 如图 7 (a) 所示。注意, 只有在 3D 轴的条件下, 这个设置才会生效。同时设定 `offset`, 我们可以在不同位置绘制这些等高线, 如图 7 (c)、(e)、(g) 所示。

沿 y 方向

类似地, 设置 `zdir='y'` 和不同 `offset` 值, 我们可以绘制沿 y 轴方向的等高线, 如图 7 (b)、(d)、(f)、(h) 所示。通过调整视角我们还可以绘制如图 2 所示平面等高线。



Jupyter 笔记 BK_2_Topic_5.06_2.ipynb 绘制图 7 所有子图。

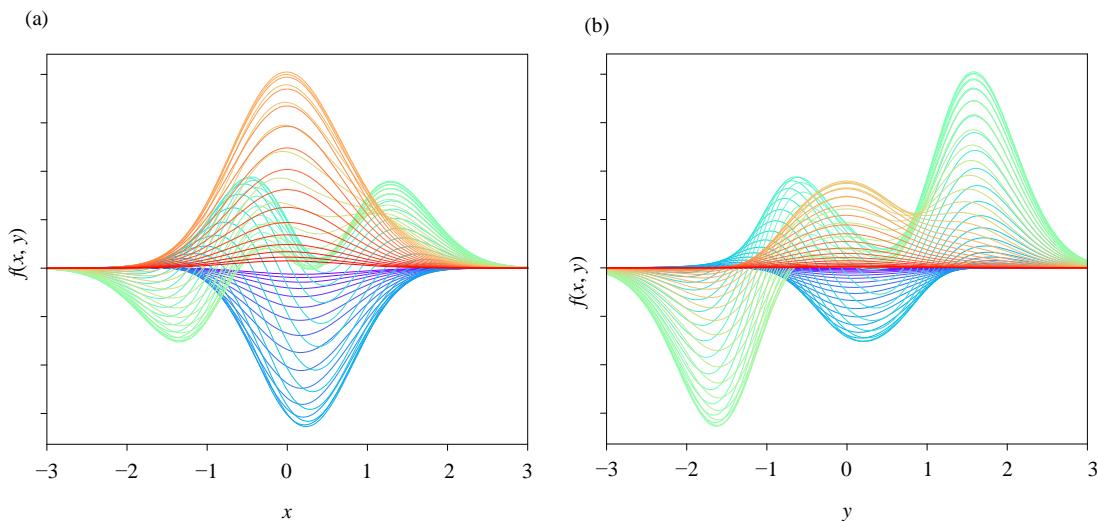


图 2. 通过改变视角绘制两组平面等高线

21.2 特定等高线

将满足单位圆 ($x_1^2 + x_2^2 = 1$) 的坐标映射到不同二次曲面, 我们可以得到如图 8、图 9 所示的几个子图。

对于单位圆, 我们可以用极坐标系很容易获得满足条件的一系列坐标 (x_1, x_2) 。然后再用三维图绘制 $(x_1, x_2, f(x_1, x_2))$ 。

图 8、图 9 这几幅图和正定性、瑞利商有关。《矩阵力量》一册将介绍这两个概念。



Jupyter 笔记 BK_2_Topic_5.07_1.ipynb 绘制图 8、图 9 所有子图。

提取特定等高线数值

下面，我们了解一种相对更为方便的可视化方案。如图 3 所示，我们可以先绘制 $g(x_1, x_2) = x_1^2 + x_2^2$ 。

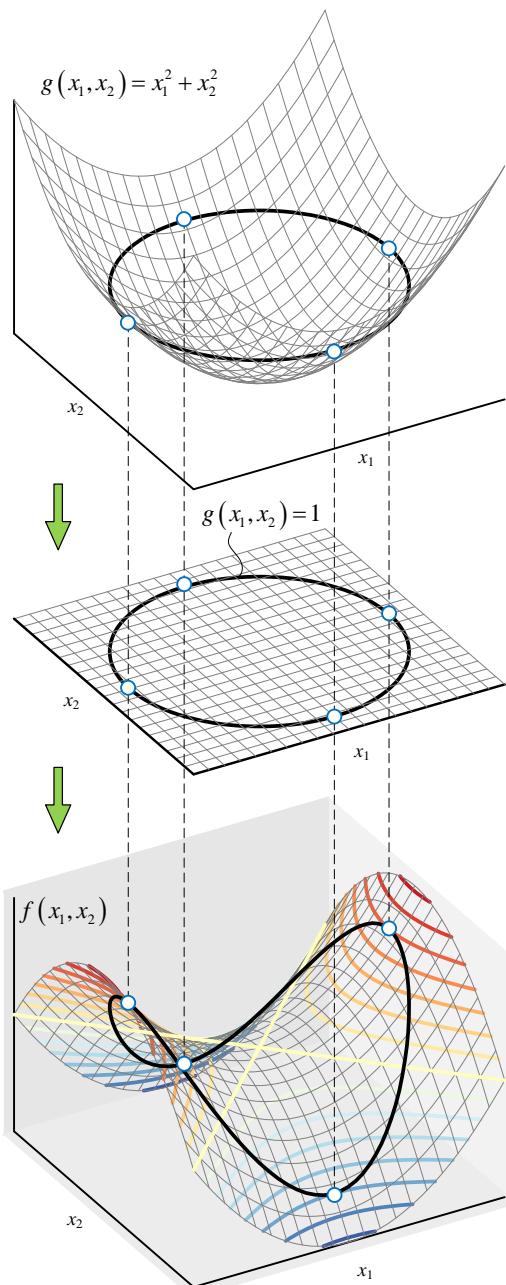
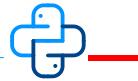


图 3. 提取特定等高线坐标点

然后找到满足 $g(x_1, x_2) = 1$ 的等高线坐标，再将它们映射到 $f(x_1, x_2)$ 曲面上。

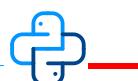
提取特定等高线数值的方法很适合处理较为复杂的等式。如图 10 所示，利用提取等高线数值的方法，我们可以很容易获得满足 $\frac{\partial f}{\partial x_1} = 0$ 或 $\frac{\partial f}{\partial x_2} = 0$ 的坐标点。然后，再将其映射到特定曲面。



Jupyter 笔记 BK_2_Topic_5.07_2.ipynb 绘制图 10 所有子图。

绘制交线

类似地，我们可以用提取等高线的方法绘制如图 11 所示曲面和平面的交线。



Jupyter 笔记 BK_2_Topic_5.07_3.ipynb 绘制图 11 所有子图。

21.3 可视化四维数据

等高线还可以完成很多有趣的可视化方案，这个话题介绍如何用分层等高线可视化四维数据。这个四维数据是三元高斯分布， $f_{X_1, X_2, X_3}(x_1, x_2, x_3)$ 。

《统计至简》第 10 章将介绍二元高斯分布，第 11 章介绍多元高斯分布。

x_1, x_2, x_3 的取值范围都是 $(-\infty, +\infty)$ 。为了方便可视化，我们给 x_1, x_2, x_3 设定的取值范围是 $[-2, 2]$ 。

这样，我们便得到如图 4 左图所示的“豆腐块”。豆腐块表面的“纹理”就是概率密度 $f_{X_1, X_2, X_3}(x_1, x_2, x_3)$ ，第四维数据。

显然，这块豆腐内部每一点都对应一个概率密度。为了可视化这些概率密度值，我们采用“切豆腐”的方法来观察剖面上的概率密度等高线。大家对这种方法应该不陌生，我们在本书前文已经看到好几次。

图 4 右侧三幅子图展示的是三种切豆腐的“手法”。举个例子，如图 5 所以，垂直 x_3 轴切豆腐，意味着绘制等高线时， x_3 固定在某个特定值 c ， $x_3 = c$ 。我们这次用等高线可视化 $f_{X_1, X_2, X_3}(x_1, x_2, x_3 = c)$ 。

为了看到等高线的全貌，我们采用单独子图的可视化方案。图 12 所示为沿着三个不同方向切豆腐的结果。



Jupyter 笔记 BK_2_Topic_5.08_1.ipynb 绘制图 12 所有子图。

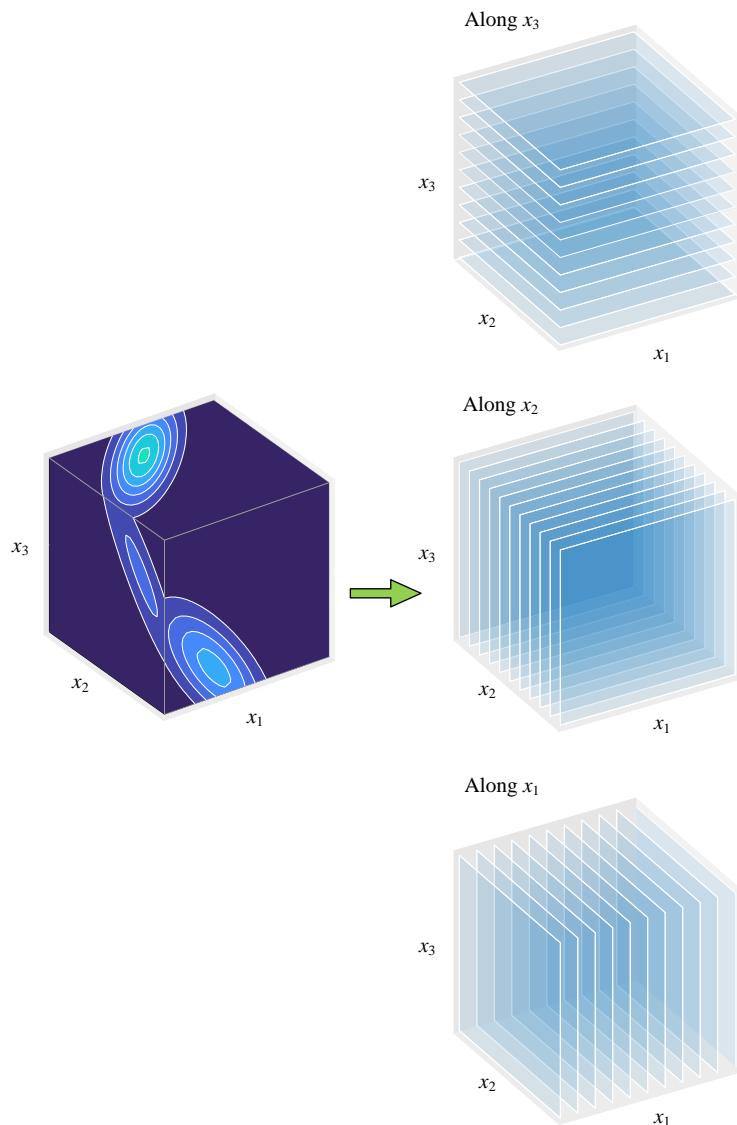


图 4. 三种不同切“豆腐”

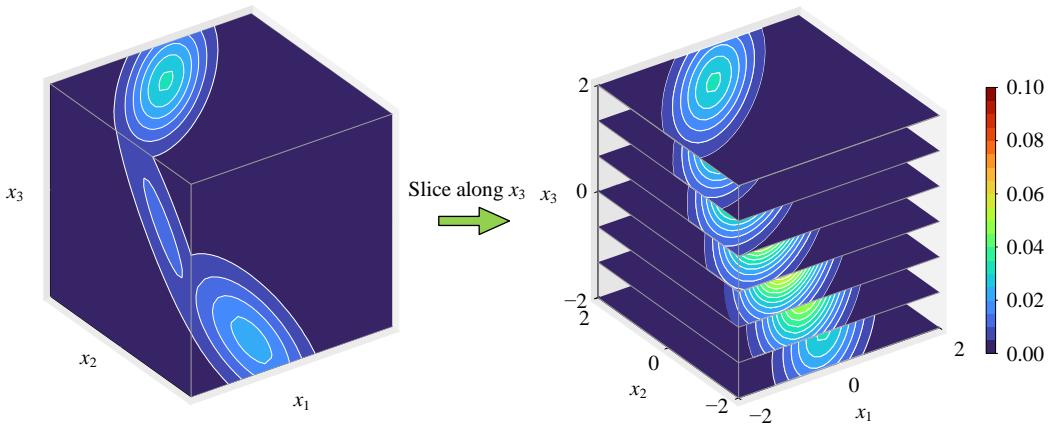


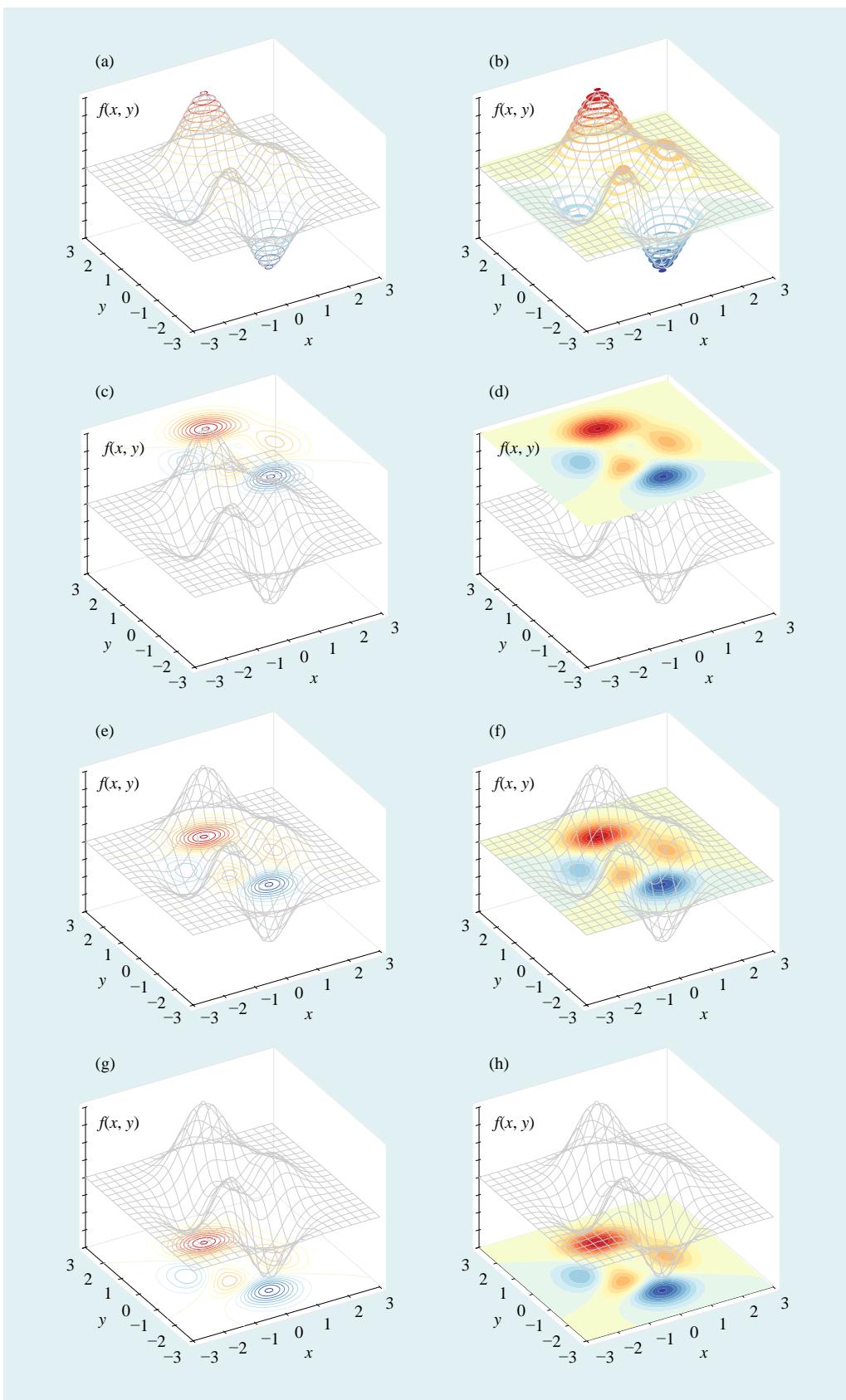
图 5. 垂直 x_3 切豆腐

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

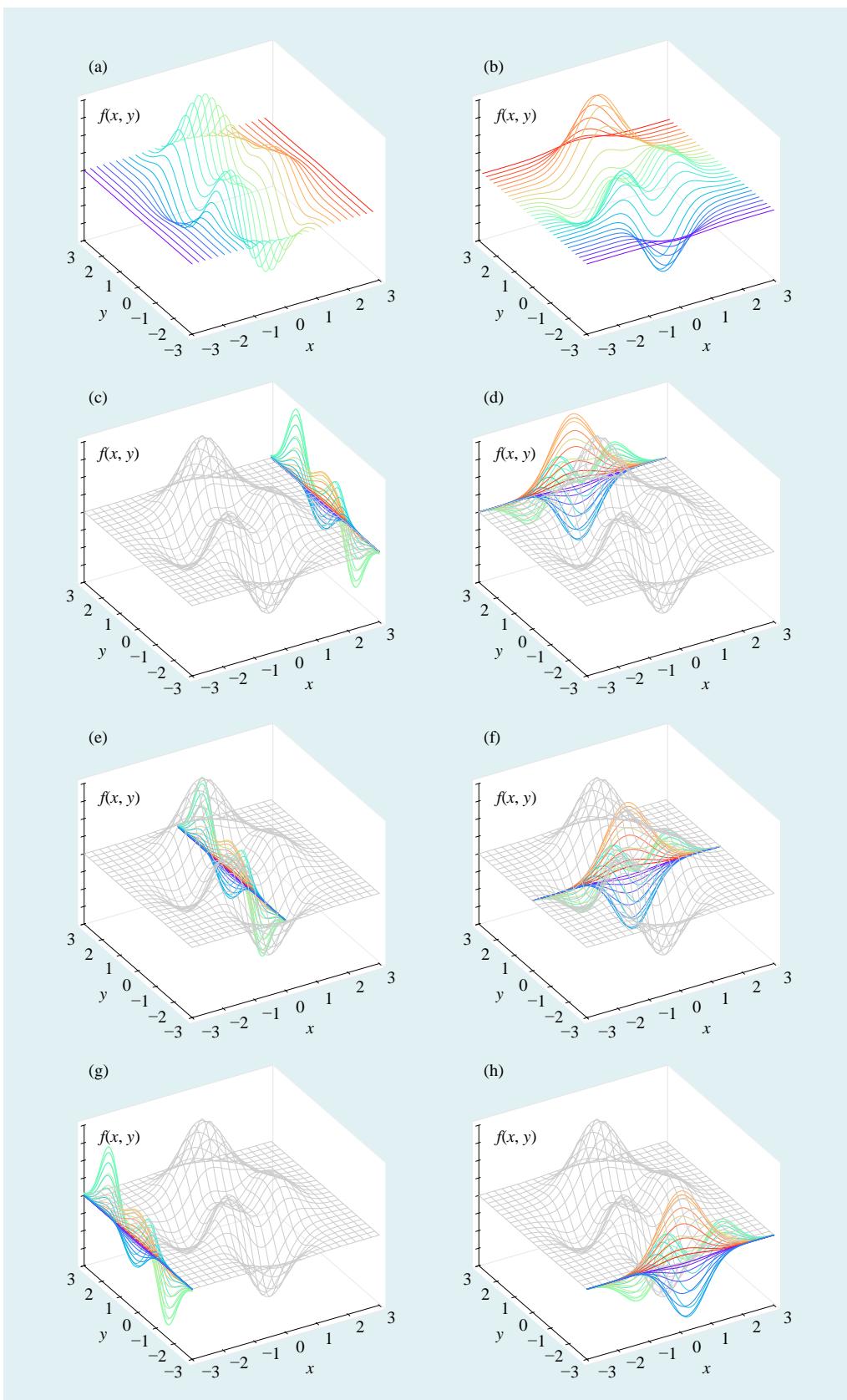
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

图 6. 沿 z 方向获取等高线

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

图 7. 沿 x 、 y 方向获取等高线

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

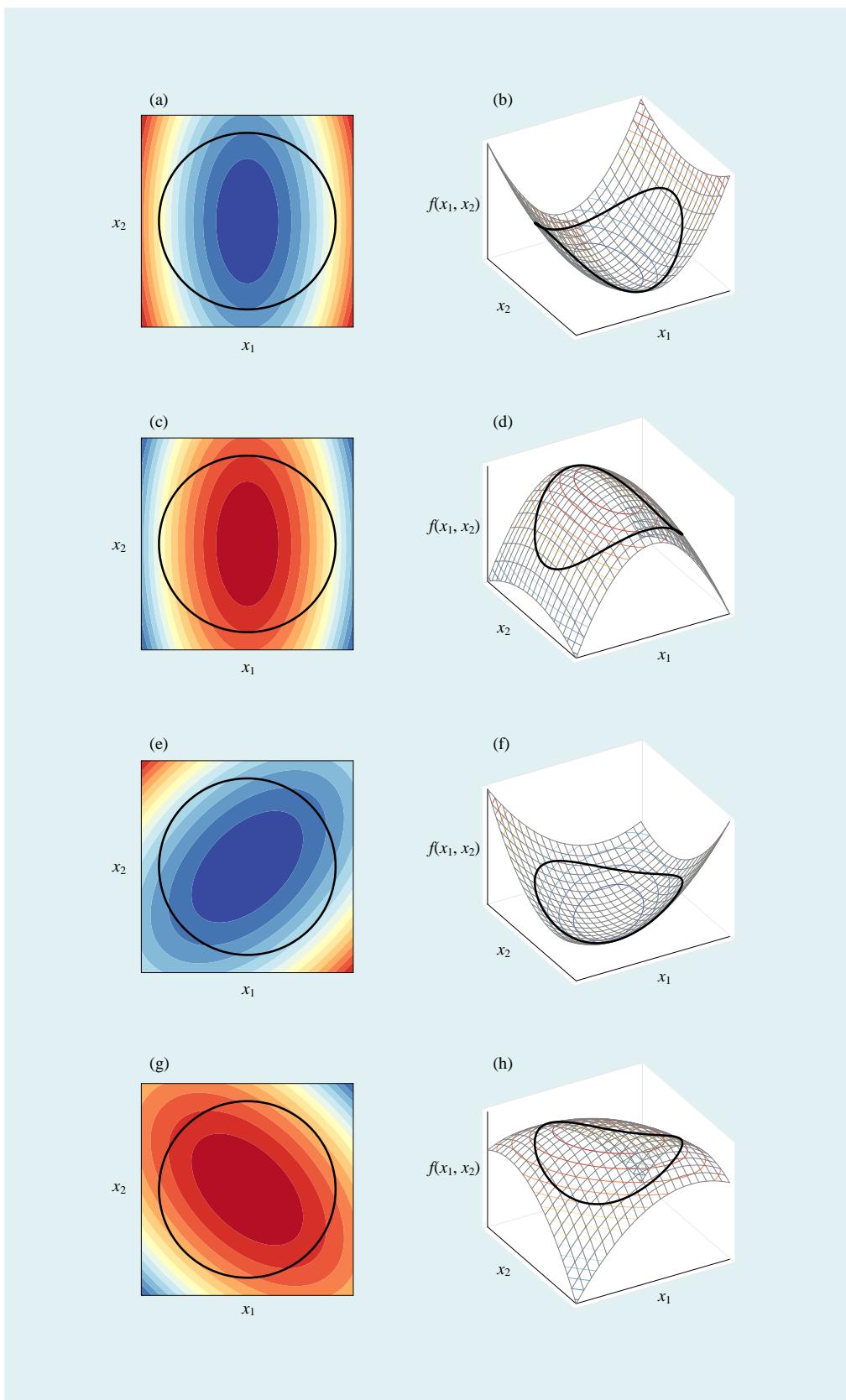


图 8. 将单位圆对应坐标映射到特定曲面，前四个例子

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

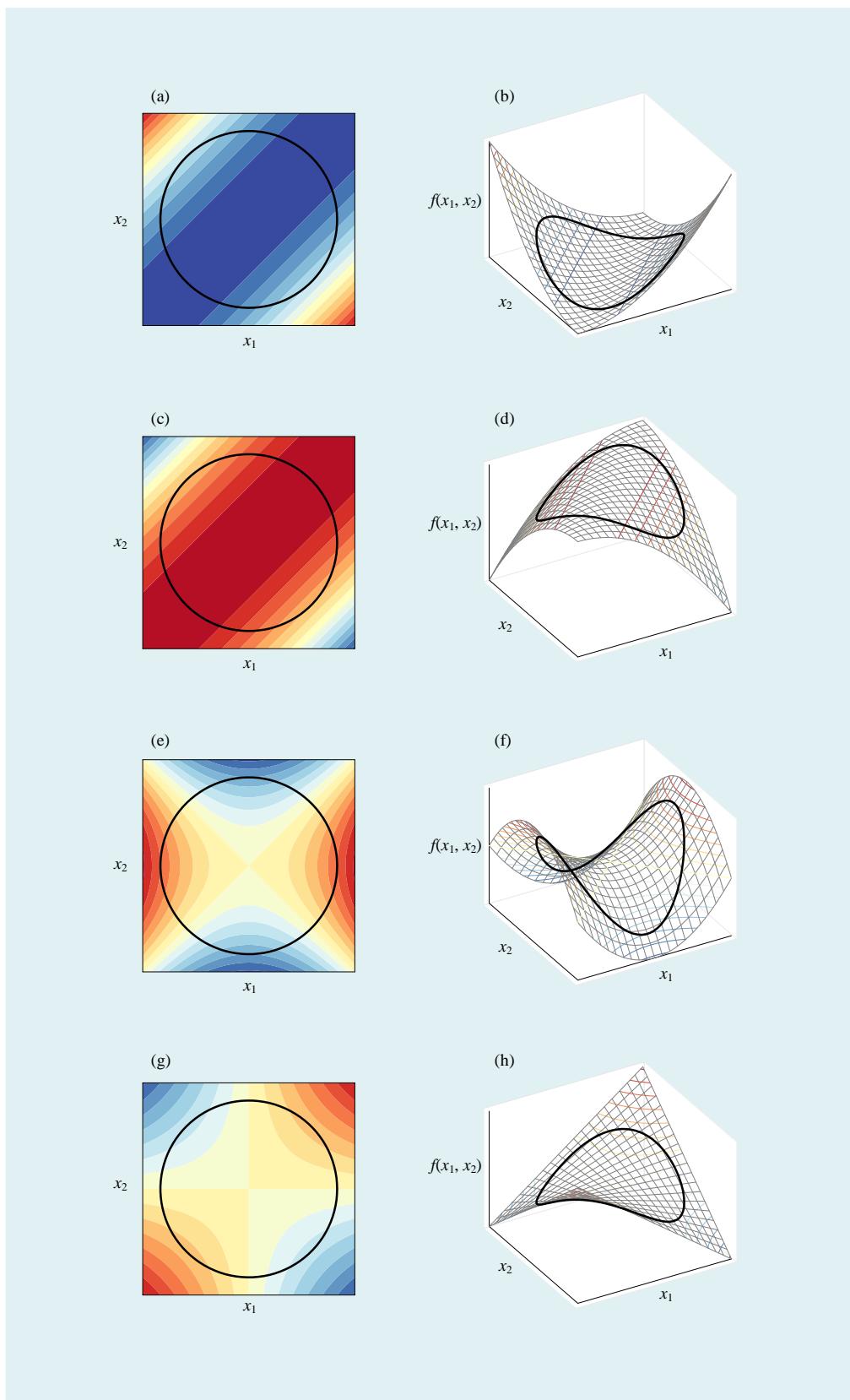


图 9. 将单位圆对应坐标映射到特定曲面，后四个例子

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

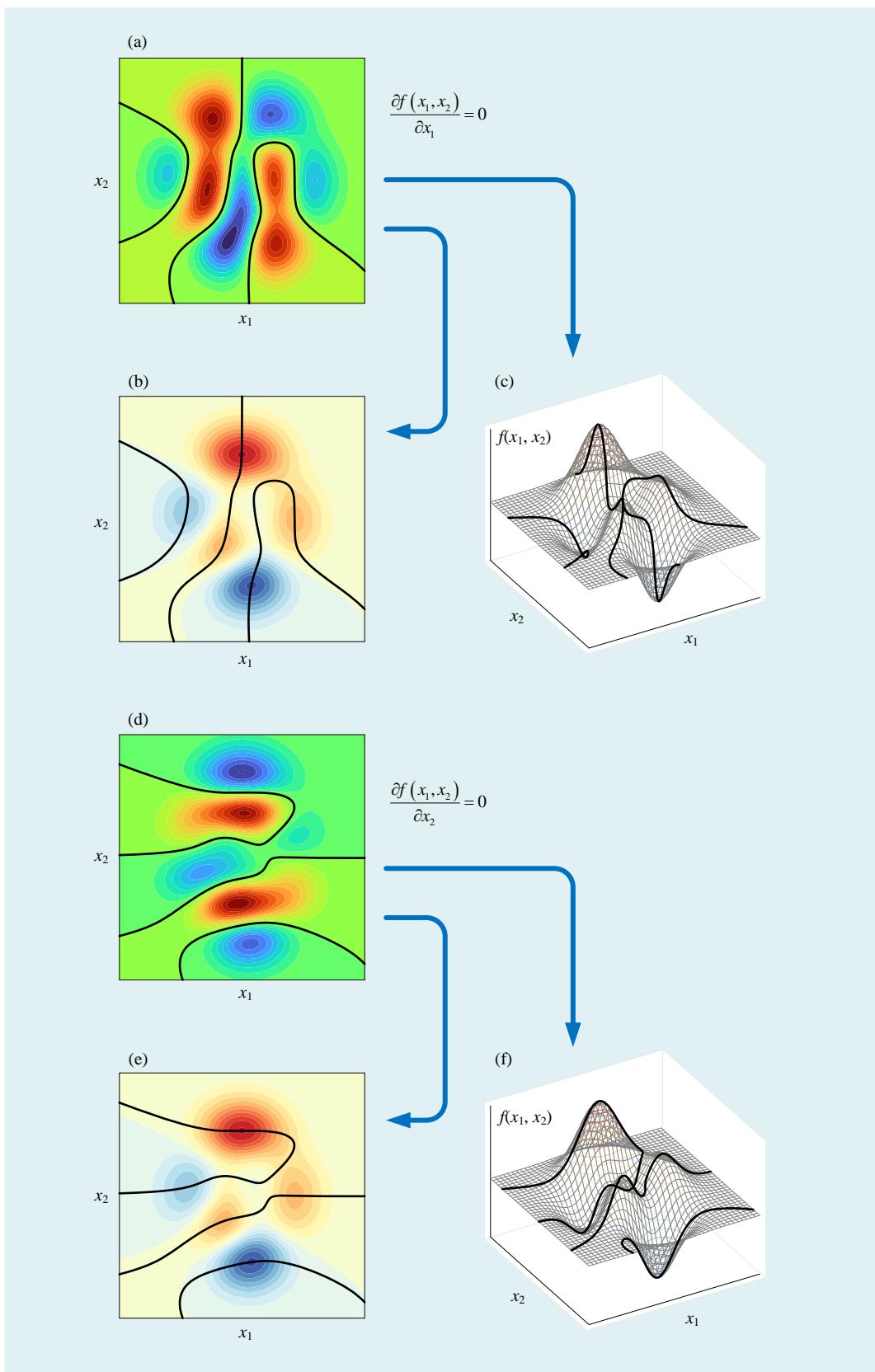


图 10. 提取等高线坐标

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

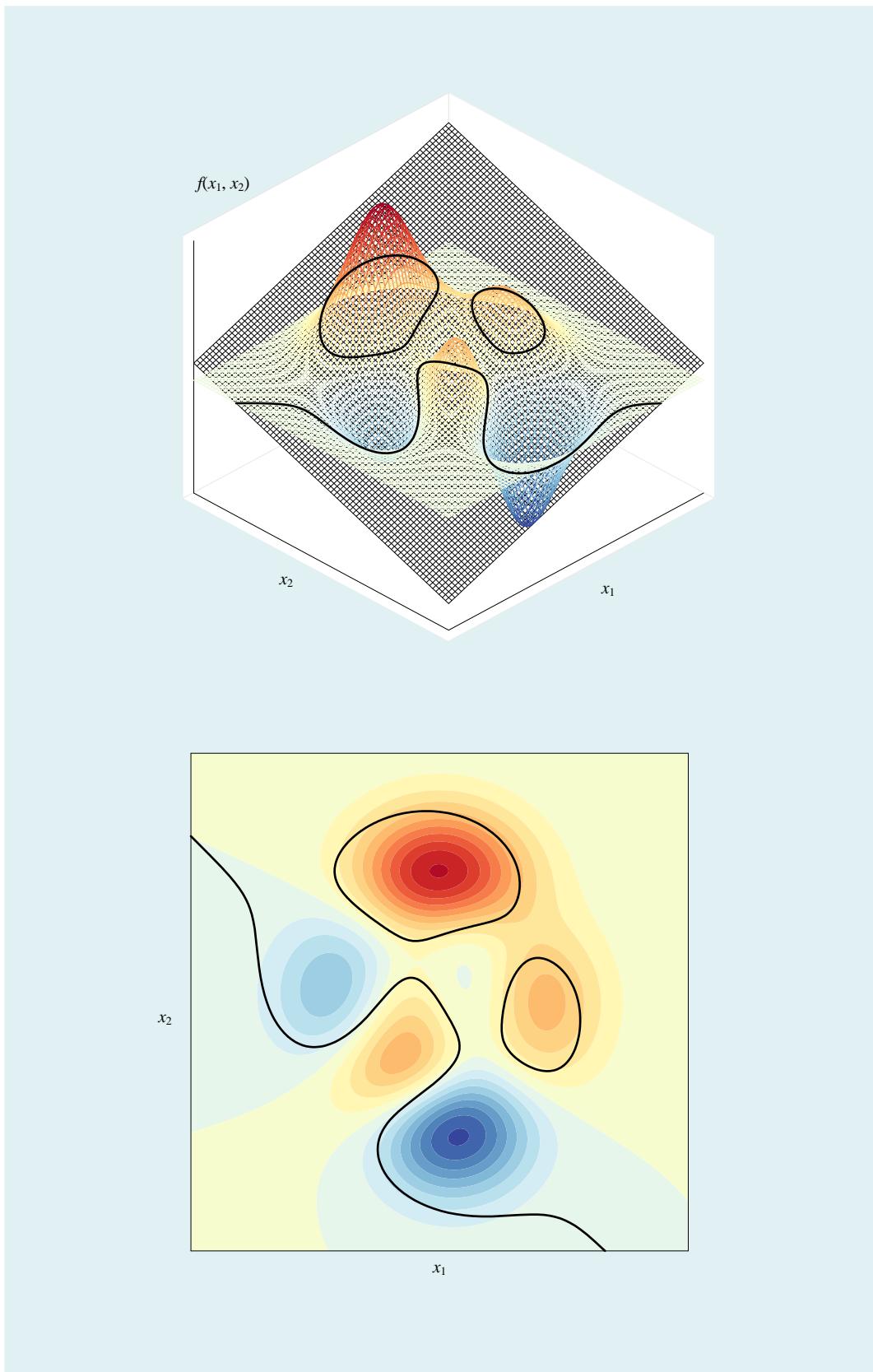


图 11. 用等高线绘制交线

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

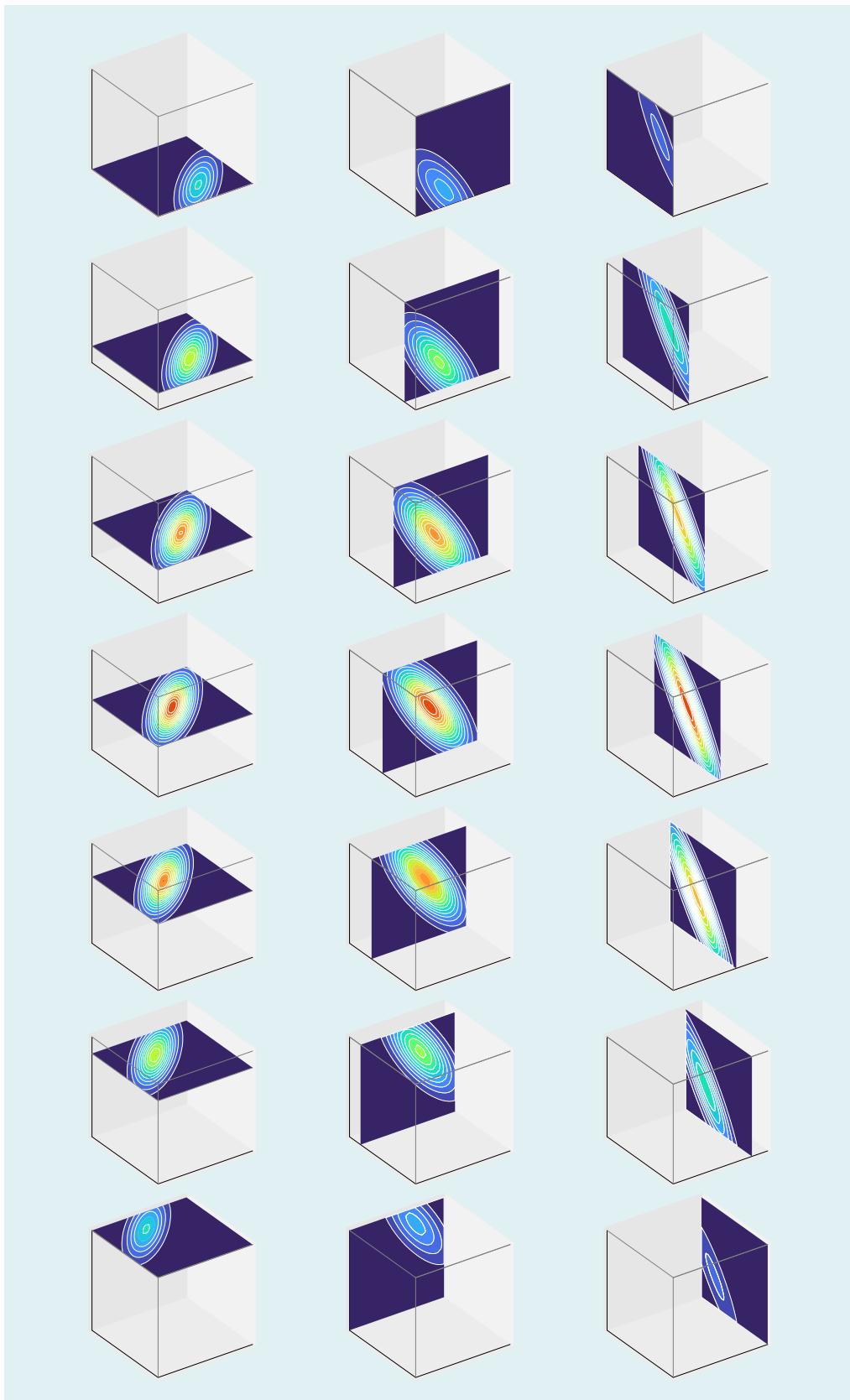


图 12. 沿三个不同方向切豆腐

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

22 Quiver Plots 箭头图 有大小、有方向



生存还是死亡，这是一个问题：

要想活的高贵，到底是该忍气吞声

接受厄运的捶打，

还是该拿起武器痛击无尽的烦恼，

打败一切？

To be, or not to be, that is the question:

Whether 'tis nobler in the mind to suffer

The slings and arrows of outrageous fortune,

Or to take arms against a sea of troubles,

And by opposing end them?

—— 威廉·莎士比亚 (William Shakespeare) | 英国剧作家 | 1564 ~ 1616



- ◀ `matplotlib.pyplot.axhline()` 绘制水平线
- ◀ `matplotlib.pyplot.axvline()` 绘制竖直线
- ◀ `matplotlib.pyplot.fill_between()` 区域填充颜色
- ◀ `matplotlib.pyplot.plot()` 绘制线图
- ◀ `matplotlib.pyplot.quiver()` 绘制箭头图
- ◀ `matplotlib.pyplot.scatter()` 绘制散点图
- ◀ `matplotlib.pyplot.text()` 在图片上打印文字
- ◀ `numpy.flip()` 指定轴翻转数组
- ◀ `numpy.fliplr()` 左右翻转数组
- ◀ `numpy.flipud()` 上下翻转数组
- ◀ `numpy.meshgrid()` 创建网格化数据
- ◀ `numpy.prod()` 指定轴的元素乘积
- ◀ `sympy.diff()` 求解符号导数和偏导解析式
- ◀ `sympy.lambdify()` 将符号表达式转化为函数
- ◀ `sympy.symbols()` 定义符号变量

22.1 向量

向量 (vector) 是数学中的一种概念，常用于表示空间中的物理量，如速度、力、位移等。向量可以用有向线段表示，具有方向和大小两个属性。与之相对的标量 (scaler) 只有大小这一个属性。

在二维空间中，一个向量可以表示为一个有序的数对 (x, y) 、 $[x, y]$ 、 $[x, y]^T$ 。向量也可以用一个有向线段来表示，线段的起点为原点 $(0, 0)$ ，终点为 (x, y) 。其中， x 表示向量在 x 轴上的投影， y 表示向量在 y 轴上的投影，也称为向量的横纵坐标。

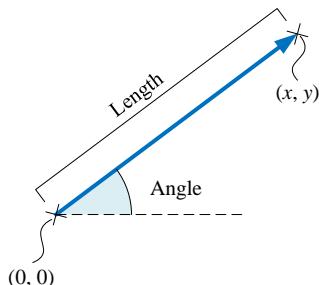


图 1. 向量起点、终点、大小和方向

类似地，在三维空间中，一个向量可以表示为 (x, y, z) 、 $[x, y, z]$ 、 $[x, y, z]^T$ 。三维向量也可以用一个有向线段来表示，线段的起点为原点 $(0, 0, 0)$ ，终点为 (x, y, z) 。其中， x 、 y 和 z 分别表示向量在 x 轴、 y 轴和 z 轴上的投影长度，也称为向量的三个坐标。

向量的大小，也称为向量的长度或模，可以用勾股定理求出。

《矩阵力量》一册将专门讲解向量。

22.2 箭头

quiver 是 matplotlib 中的一个函数，用于绘制二维、三维箭头图。

举个例子，二维箭头图的函数和基本参数为 `matplotlib.pyplot.quiver(x, y, u, v, scale=1)`。其中， x 和 y 是箭头起始点的坐标， u 和 v 是箭头在两个方向的投影量。默认情况下，箭头的长度是按照输入数据的比例来绘制的，可以通过 `scale` 参数进行调整。图 2 所示为 quiver 箭头的常用参数。

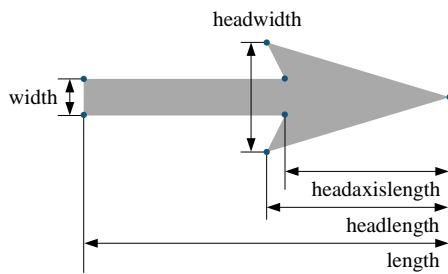


图 2. 箭头的参数

二维箭头

图 3 所示为利用箭头图可视化向量加法。图 4 所示为利用箭头图展示向量长度。图 5 可可视化向量减法。

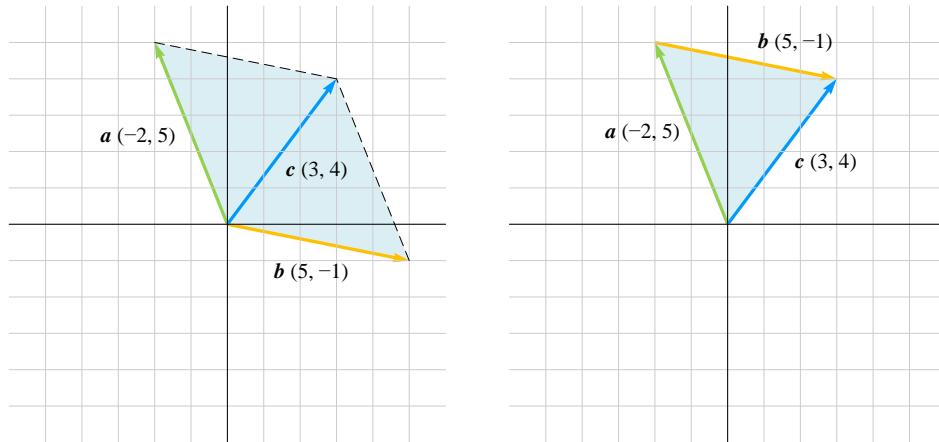


图 3. 可视化二维向量加法

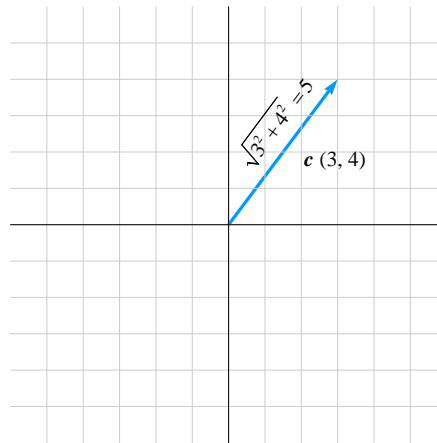


图 4. 可视化二维向量长度

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

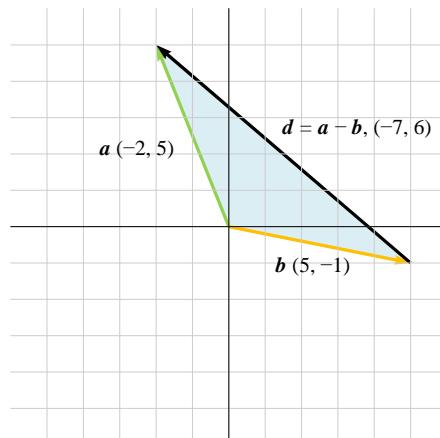


图 5. 可视化二维向量减法

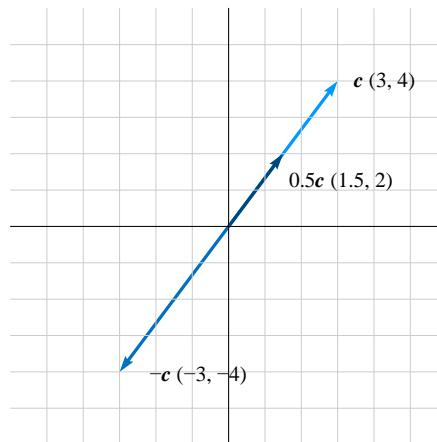


图 6. 可视化二维向量乘法

三维箭头

图 7 所示为利用三维箭头图可视化三维向量加法。图 8 所示为向量投影到 xy 平面、 xz 平面、 yz 平面。图 9 所示为向量投影到 x 轴、 y 轴、 z 轴。

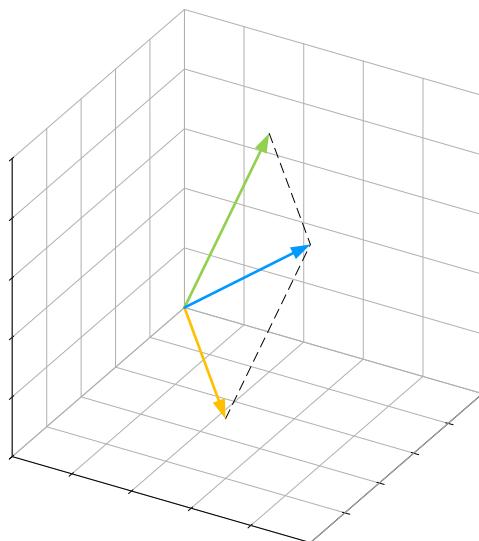


图 7. 可视化三维向量加法

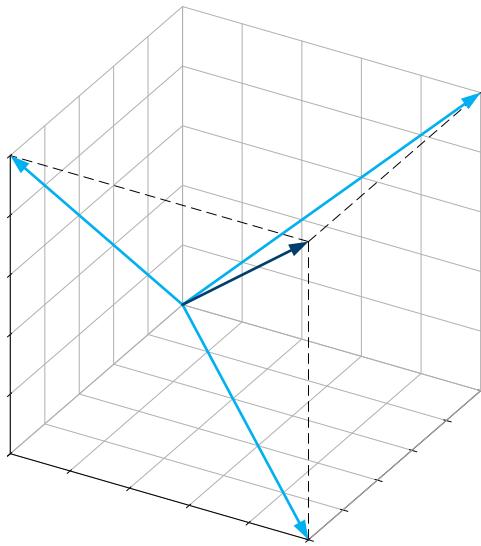


图 8. 三维向量投影到平面

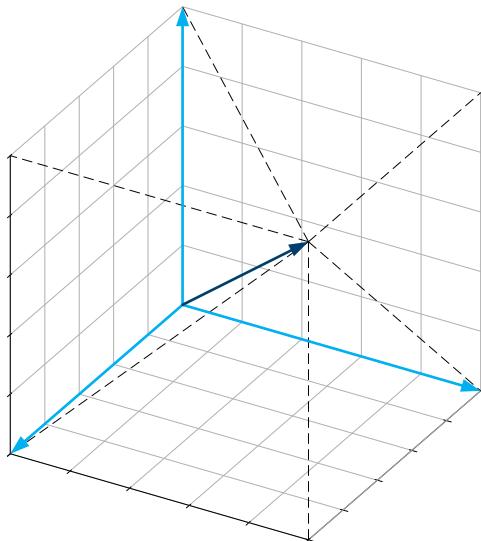


图 9. 三维向量投影到轴

22.3 向量场

除了单独绘制箭头图，我们还可以绘制向量场。向量场是指在空间中的每一个点都存在一个向量的集合。在数学中，向量场通常用函数来描述，这个函数将每个点映射到该点处的向量。这个函数被称为向量场的场函数或者向量场的定义式。向量场可以用来描述许多物理现象，例如流体力学中的速度场，电场、磁场、水流、风向等等。

图 10 所示为利用向量场可视化特征向量。图 13 所示为利用向量场可视化曲面梯度向量、法向量。

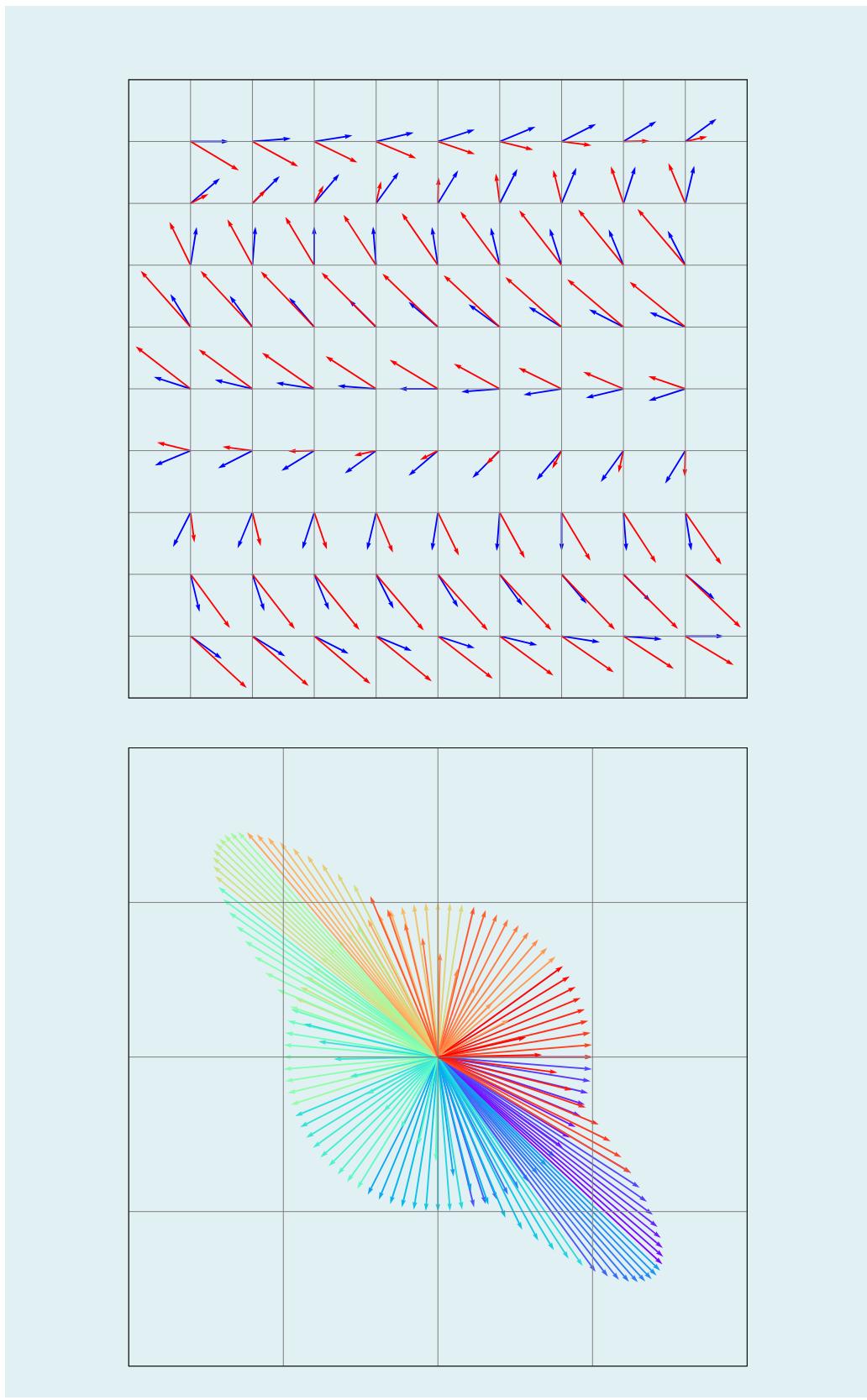


图 10. 可视化特征向量

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

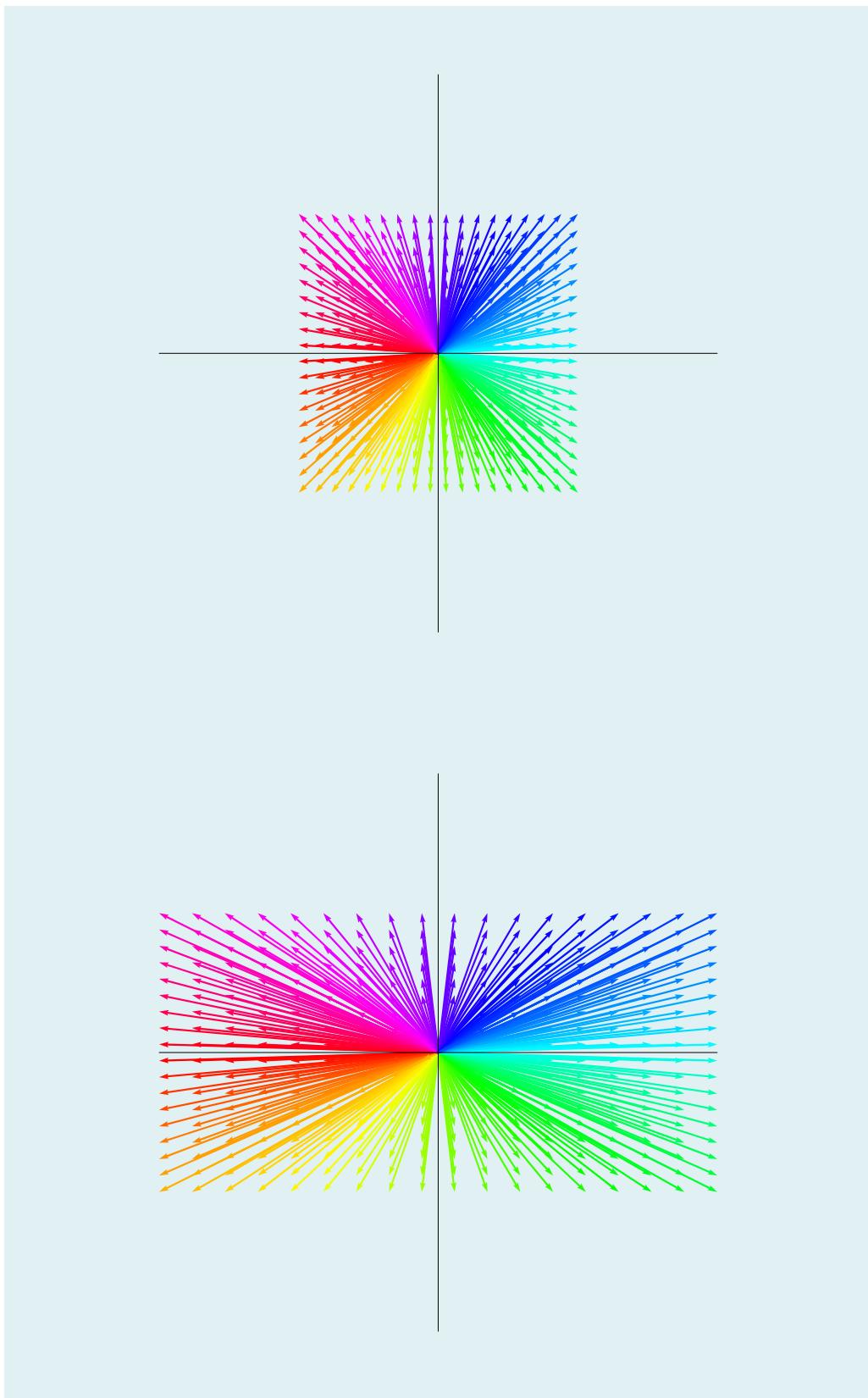


图 11. 可视化几何变换，第 1 组

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

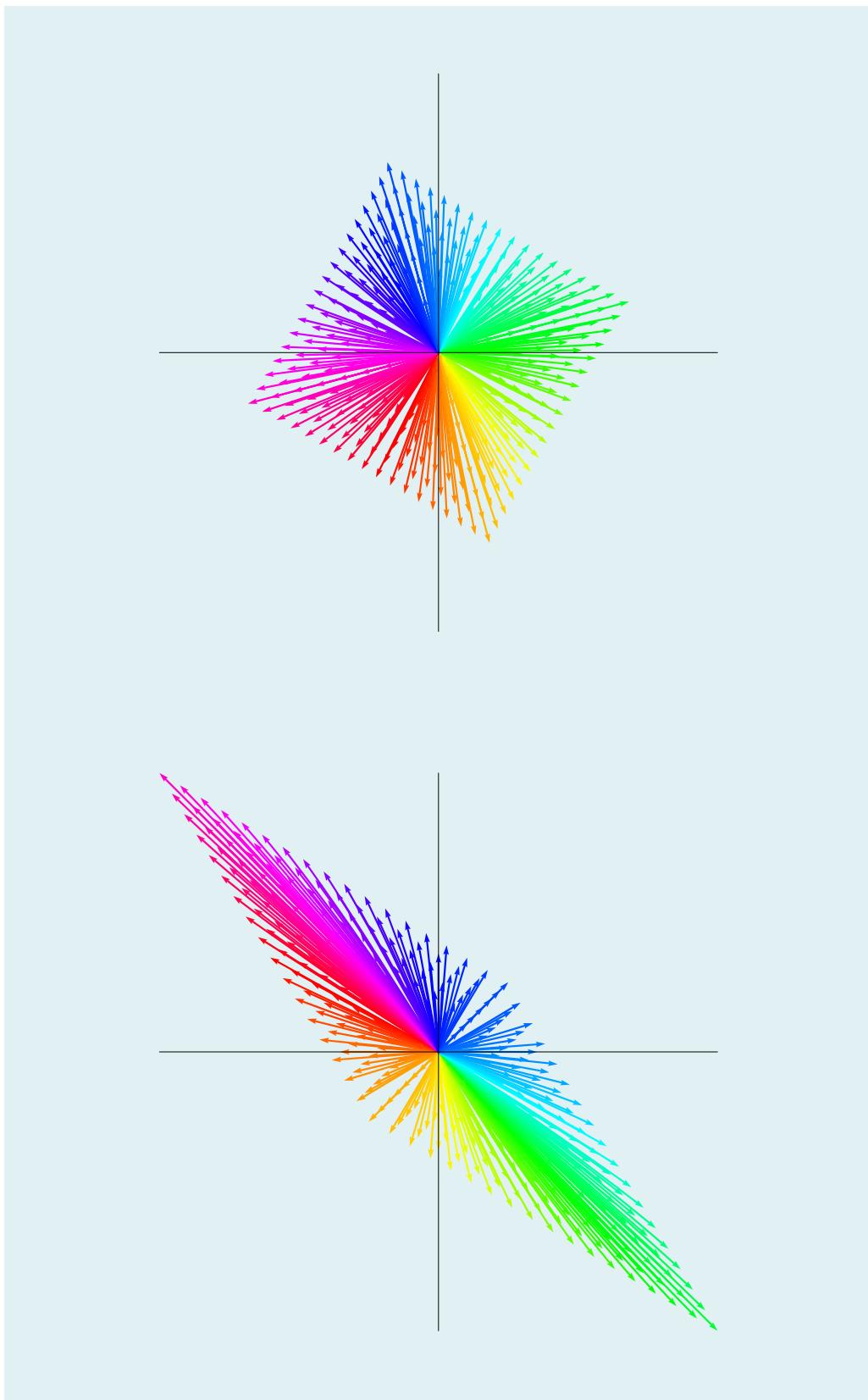


图 12. 可视化几何变换，第 2 组

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

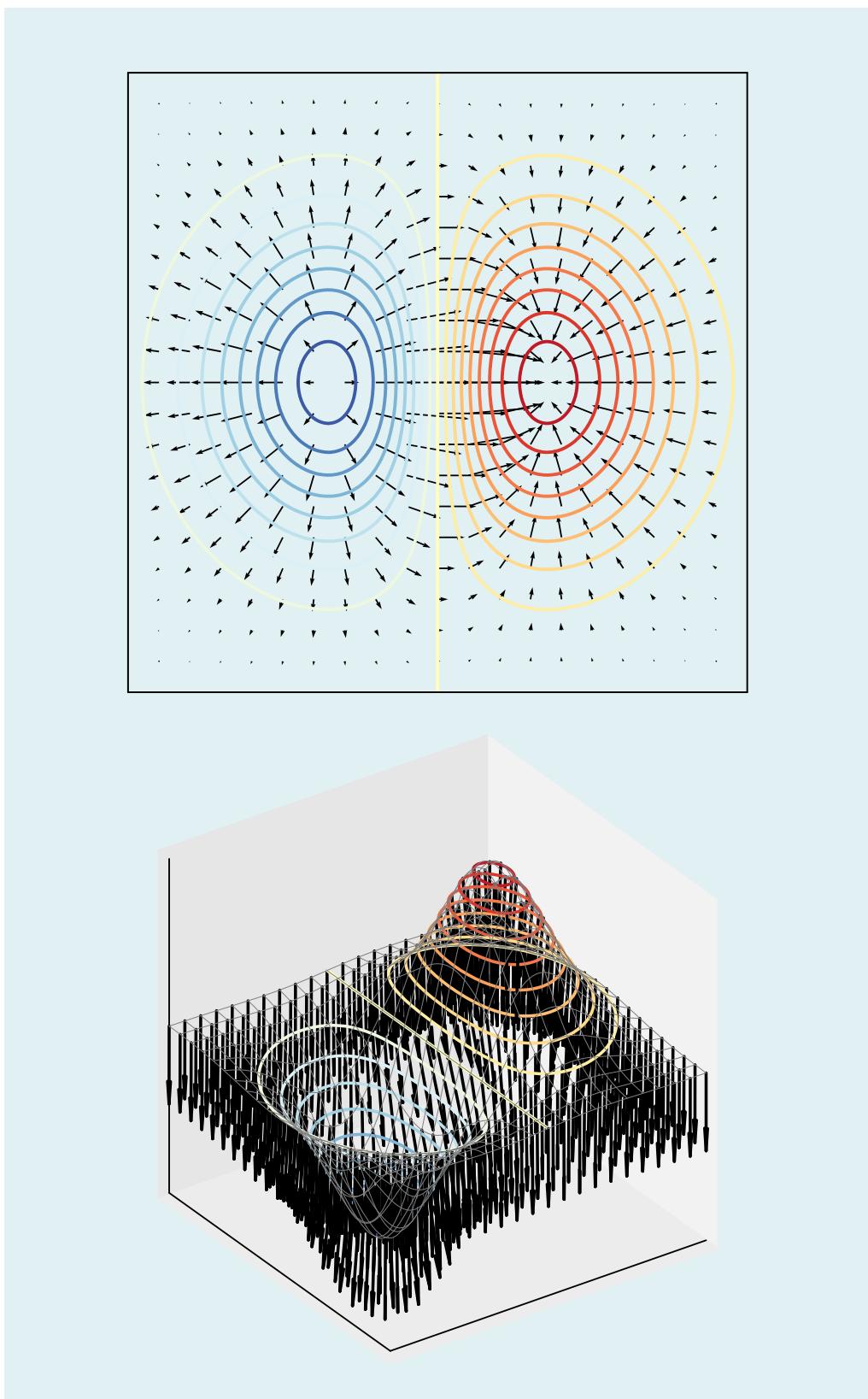


图 13. 可视化曲面梯度向量、法向量

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

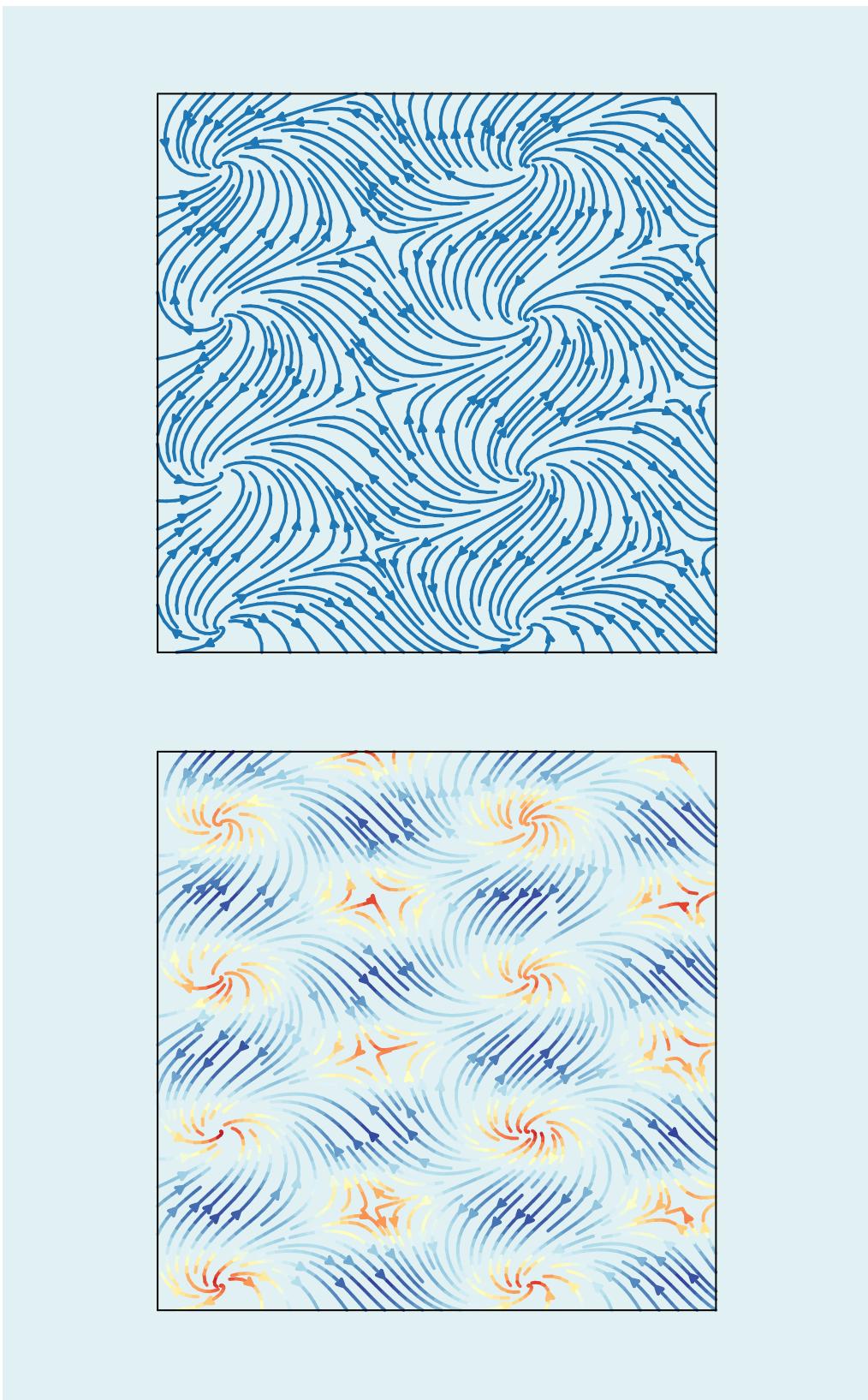


图 14. 水流图

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

23

Triangular Grid

三角网格

将给定的几何形状划分为一组不重叠的三角形



天地间唯一的英雄主义——参破尘世真相，依旧热爱生活。

There is only one heroism in the world: to see the world as it is, and to love it.

—— 罗曼·罗兰 (Romain Rolland) | 法国作家、诺贝尔文学奖得主 | 1866 ~ 1944



- ◀ `matplotlib.pyplot.plot_trisurf()` 在三角形网格上绘制平滑的三维曲面图
- ◀ `matplotlib.pyplot.tricontourf()` 在三角形网格上绘制填充的等高线图
- ◀ `matplotlib.pyplot.triplot()` 在三角形网格上绘制线条
- ◀ `matplotlib.tri.Triangulation()` 生成三角剖分对象
- ◀ `matplotlib.tri.UniformTriRefiner()` 对三角形网格进行均匀细化，生成更密集的三角形网格，以提高绘制的精细度和准确性
- ◀ `numpy.asarray()` 将输入数据，比如列表、元组等，转换为 NumPy 数组
- ◀ `numpy.clip()` 用于将数组中的元素限制在指定的范围内，将小于最小值的元素设置为最小值，将大于最大值的元素设置为最大值，返回经过剪切后的数组
- ◀ `numpy.column_stack()` 将两个矩阵按列合并
- ◀ `numpy.meshgrid()` 产生网格化数据
- ◀ `scipy.spatial.Delaunay()` 生成一个点集的 Delaunay 三角剖分
- ◀ `scipy.stats.dirichlet.pdf()` 计算 Dirichlet 分布的概率密度函数
- ◀ `sympy.diff()` 求解符号导数和偏导解析式
- ◀ `sympy.exp()` 符号自然指数
- ◀ `sympy.lambdify()` 将符号表达式转化为函数

23.1 三角剖分

三角剖分是计算几何学中的一个重要概念，用于将给定的几何形状划分为一组不重叠的三角形。详细来说，三角剖分的目标是将一个多边形或多边形的集合分解成一组互不相交的三角形，使得这些三角形的顶点恰好是原始几何形状的顶点，并且任意两个三角形之间的交集只能是共享一个边或顶点。

`matplotlib.tri` 是 Matplotlib 中的一个模块，提供了三角剖分的绘图功能。`scipy.spatial` 中的 `Delaunay` 类可以帮助我们生成一个点集的 Delaunay 三角剖分，它可以用于构建三角形网格、寻找最近邻等等。

图 1 (a) 所示的网格可以手动设定，也可以自动生成。自动生成的三角网格采用 Delaunay 三角剖分。三角网格可以帮助我们绘制各种类型的三角形网格，例如等高线图、三角形色块图和三角形曲面图等。图 1 (b) 所示为三角网格等高线图。

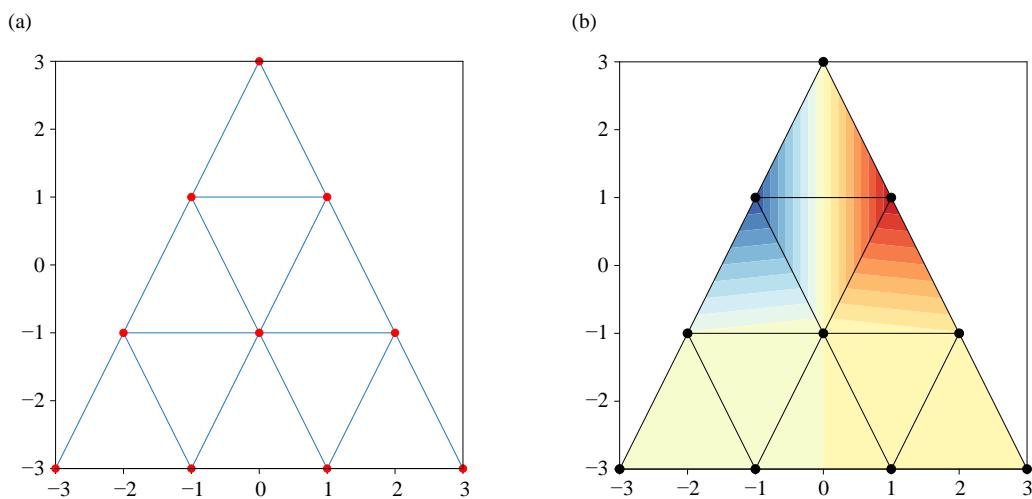


图 1. 三角网格和等高线



Bk_2_Ch23_1.ipynb 笔记中给出更多三角剖分的范例，请大家自行学习。

Delaunay 三角剖分

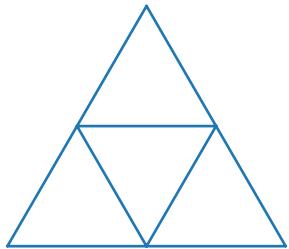
Delaunay 三角剖分是一个用于将点集分割为三角形的算法，其中任何三角形的外接圆不包含其他任何点。它的名字来源于它的发明者 Boris Delaunay。

Delaunay 三角剖分是计算机图形学、计算几何和计算机视觉中常用的技术之一，它在三维重建、图像处理、自然语言处理、机器学习等领域都有广泛的应用。

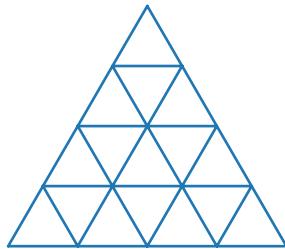
颗粒度

三角网格也存在颗粒度的问题。图 2 所示为给定等边三角形不同的颗粒度的三角网格剖分。颗粒度越高，三角网格越细腻，但是计算量也急剧增大。

(a) small triangles = 4



(b) small triangles = 16



(c) small triangles = 64

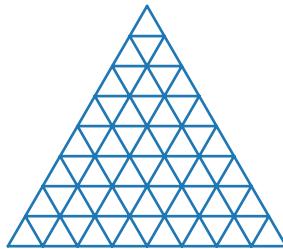
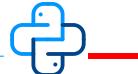


图 2. 三角网格的颗粒度

Bk_2_Ch23_2.ipynb 绘制图 2。



23.2 重心坐标系

“鸢尾花书”中三维网格常常用来可视化重心坐标系。从物理角度来看，重心坐标系 (barycentric coordinate system) 是一种描述一个几何形状内部任意点位置的方法。它是以该形状的重心作为原点建立的坐标系。

在平面上的一个三角形中，任何一点都可以表示为三个定点的加权平均值，其中每个定点的权重由它到该点的距离与该三角形的周长之比确定。这些权重称为该点在三角形的重心坐标。

实际上，用三维直角坐标系解释重心坐标系更方便。图 3 左图所示为三维直角坐标系，为了区分坐标系的横轴、纵轴、竖轴分别记做 θ_1 、 θ_2 、 θ_3 。这个三维直角坐标系中坐标可以记做 $(\theta_1, \theta_2, \theta_3)$ 。

图 3 左图浅蓝色平面上的每个坐标 $(\theta_1, \theta_2, \theta_3)$ 都满足 $\theta_1 + \theta_2 + \theta_3 = 1$ 。 $\theta_1 + \theta_2 + \theta_3 = 1$ 这个限制条件，让原本三维的空间降维成二维。即便如此，如图 3 右图所示，三角网格的每一点仍旧对应 $(\theta_1, \theta_2, \theta_3)$ 。

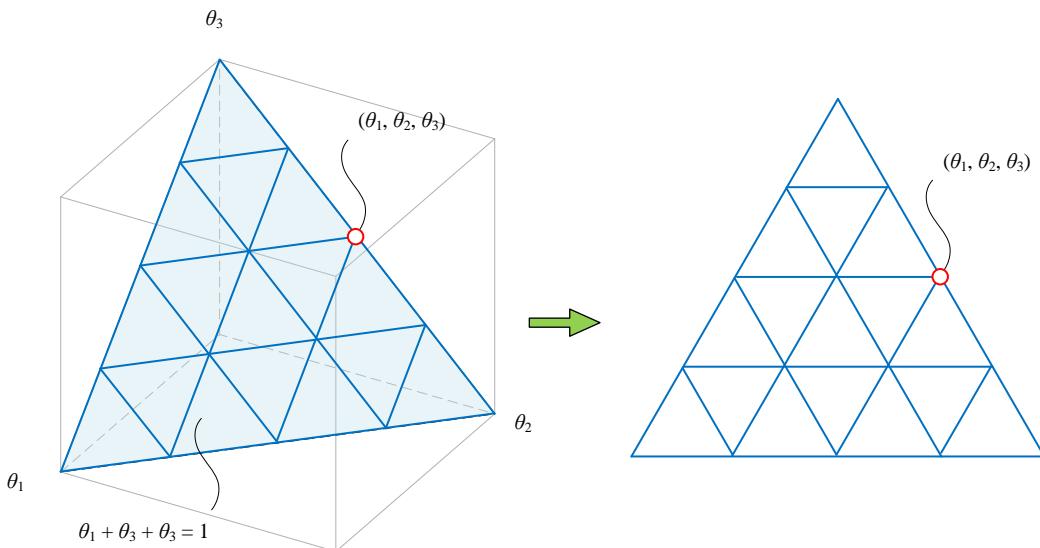


图 3. 从三维直角坐标系到重心坐标系

图 4 从三个不同的视角分别看 θ_1 、 θ_2 、 θ_3 。图 5 上下两幅子图比较两个坐标系。上图给出的是利用三角网格表达平面直角坐标系。下图则是利用相同网格表达重心坐标系坐标。请大家注意区分。



Bk_2_Ch23_3.ipynb 绘制图 5。

23.3 Dirichlet 分布

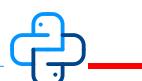
“鸢尾花书”中，重心坐标系常用来可视化 Dirichlet 分布概率密度函数。

《统计至简》一册将专门讲解 Dirichlet 分布及其在贝叶斯推断的应用。

Dirichlet 分布是一种连续的概率分布，通常用于描述一个多元随机变量的概率分布。它是以德国数学家 Peter Gustav Lejeune Dirichlet 的名字命名的。

Dirichlet 分布是一种在开区间 (0,1) 上的多维概率分布，参数是一个向量 α 。

本书前文利用其它可视化方案展示过 Dirichlet 分布。图 6、图 7 向大家展示如何利用重心坐标系可视化三元 Dirichlet 分布随参数 $[\alpha_1, \alpha_2, \alpha_3]$ 变化。



Bk_2_Ch23_4.ipynb 绘制图 6、图 7。

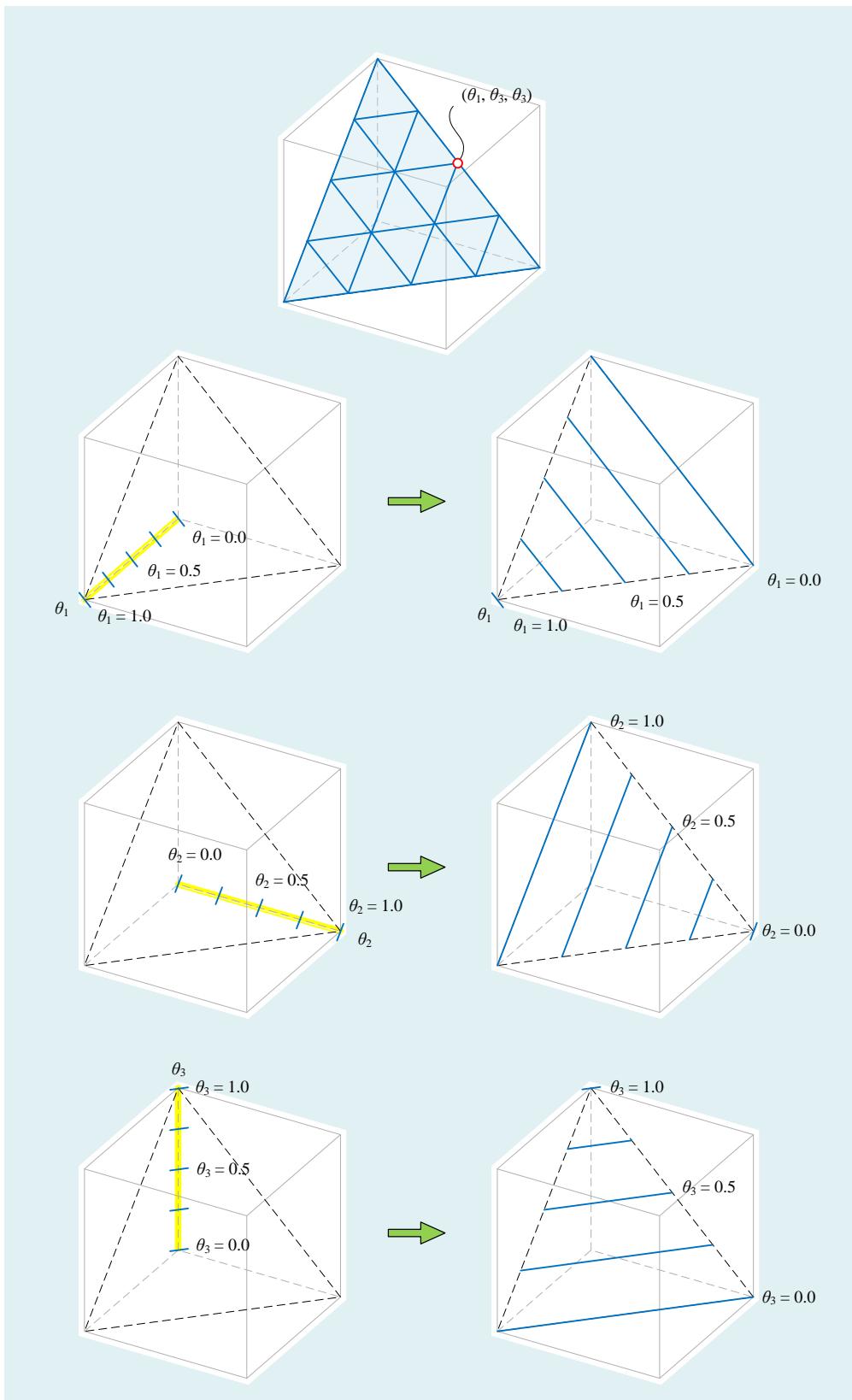


图 4. 三维直角坐标系角度看重心坐标系

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

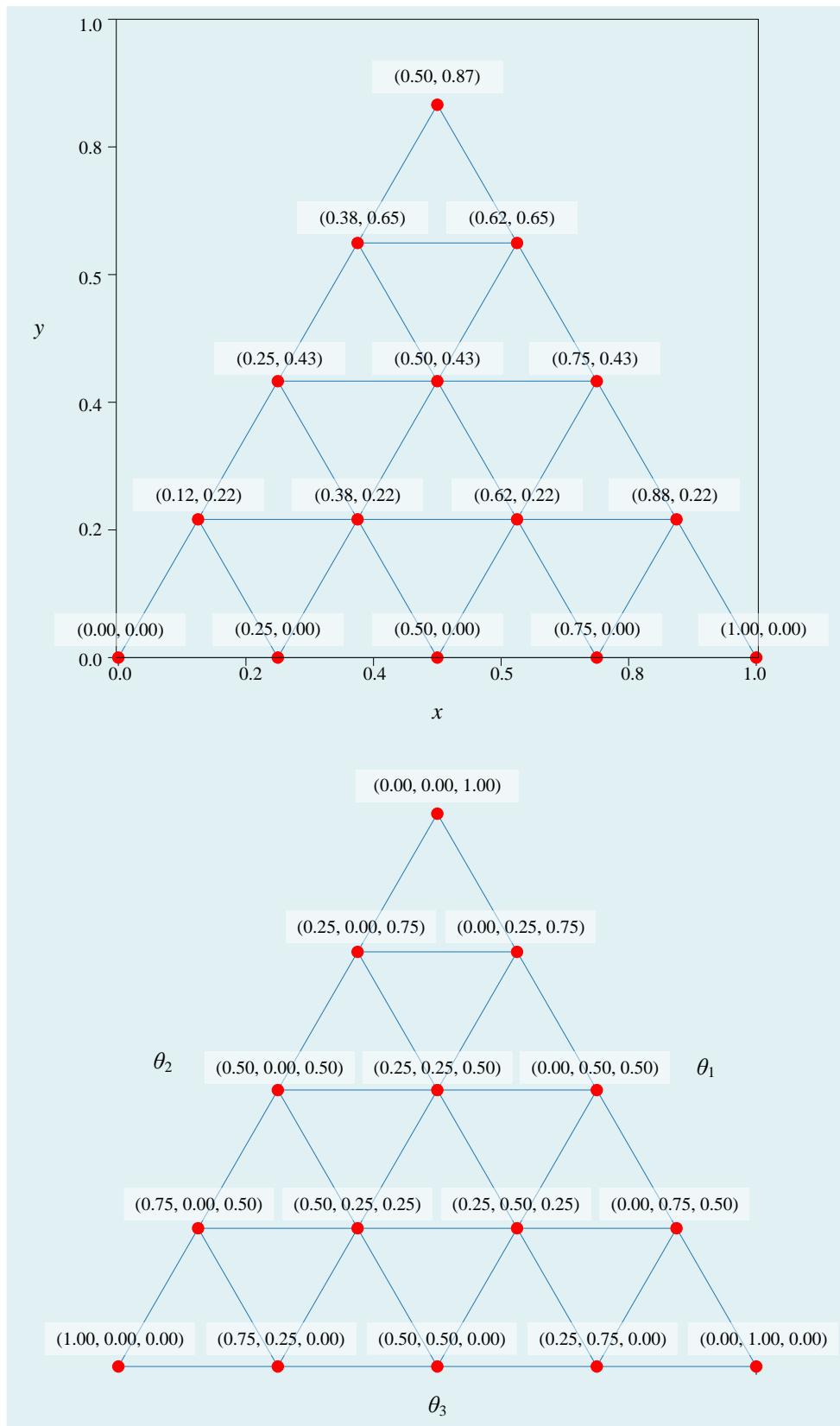


图 5. 比较平面直角坐标系坐标、重心坐标系坐标

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

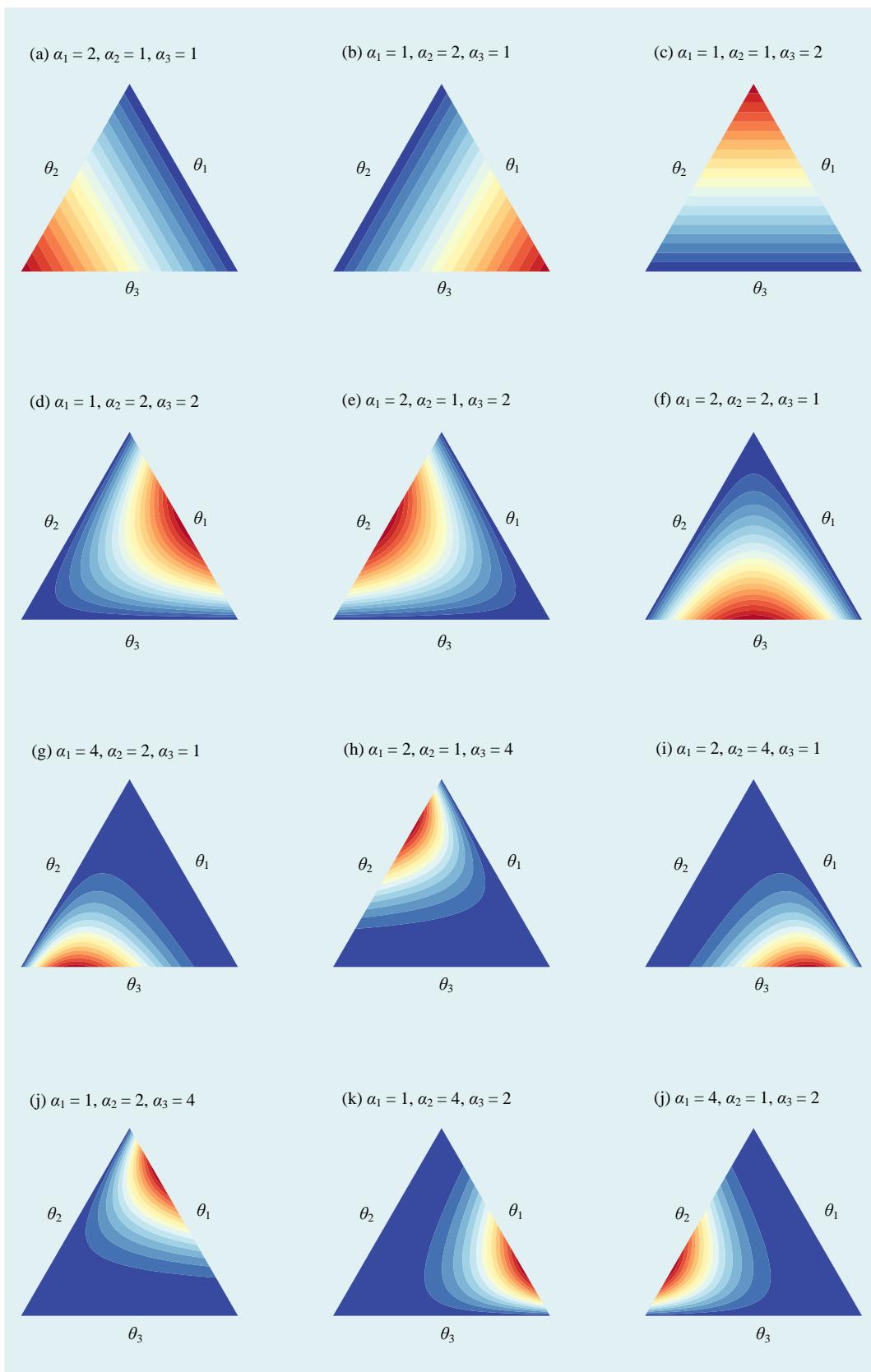


图 6. Dirichlet 分布, 第 1 组

本 PDF 文件为作者草稿, 发布目的为方便读者在移动终端学习, 终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有, 请勿商用, 引用请注明出处。

代码及 PDF 文件下载: <https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger: <https://space.bilibili.com/513194466>

欢迎大家批评指教, 本书专属邮箱: jiang.visualize.ml@gmail.com

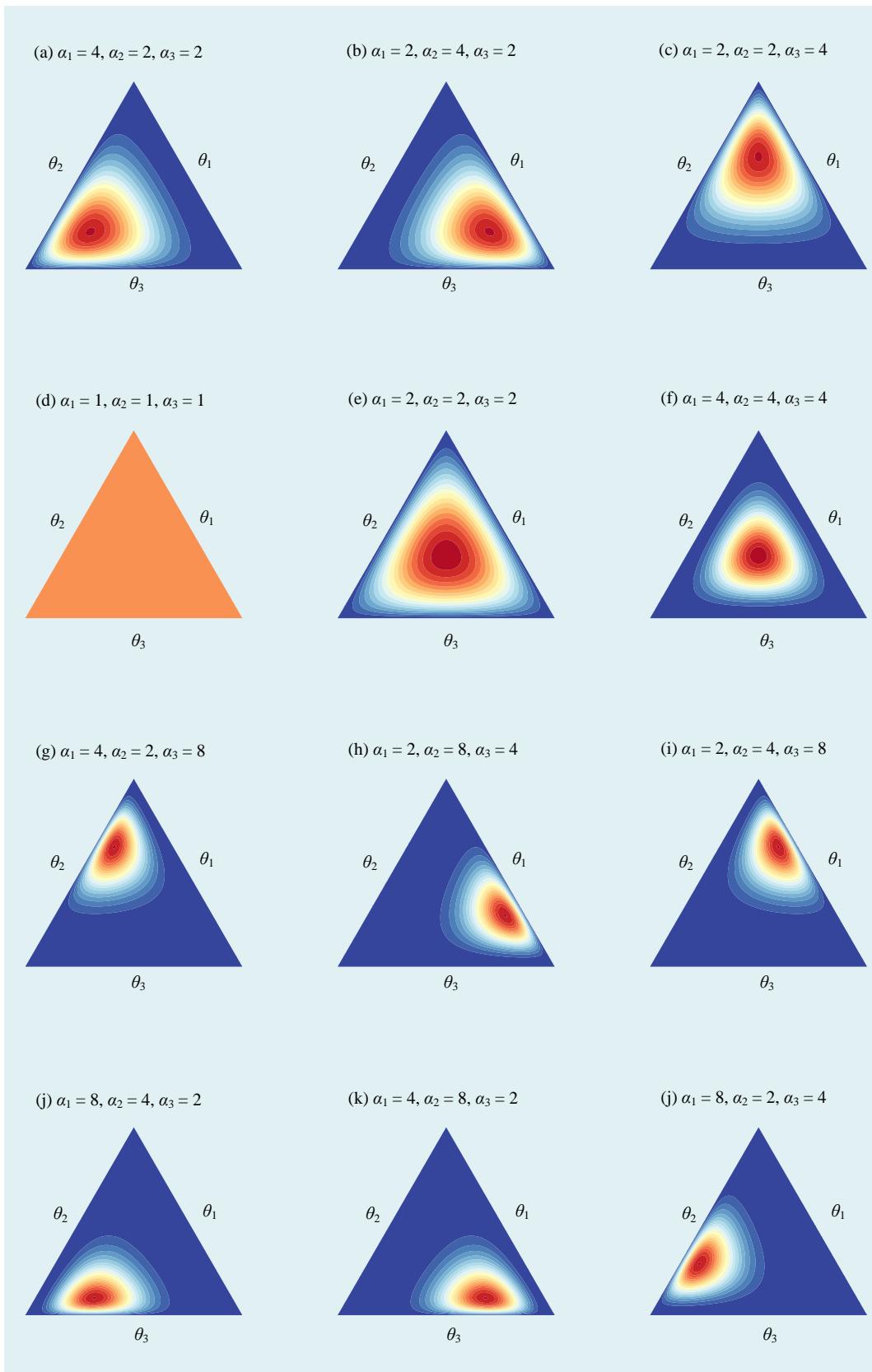


图 7. Dirichlet 分布, 第 2 组

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com



Plane Geometry

平面几何

用 Matplotlib 绘制各种平面几何形状



艺术就是谎言，但艺术让我们看清真相。

Art is the lie that enables us to realize the truth.

——毕加索 (Pablo Picasso) | 西班牙艺术家 | 1881 ~ 1973



- ◀ `matplotlib.patches.Arc()` 绘制弧线
- ◀ `matplotlib.patches.Arrow()` 绘制箭头
- ◀ `matplotlib.patches.Circle()` 绘制正圆
- ◀ `matplotlib.patches.Ellipse()` 绘制椭圆
- ◀ `matplotlib.patches.FancyBboxPatch()` 绘制 Fancy 矩形框
- ◀ `matplotlib.patches.Polygon()` 绘制多边形
- ◀ `matplotlib.patches.Rectangle()` 绘制长方形
- ◀ `matplotlib.patches.RegularPolygon()` 绘制正多边形
- ◀ `matplotlib.pyplot.cm` 提供各种预定义色谱方案，比如 `matplotlib.pyplot.cm.rainbow`
- ◀ `matplotlib.pyplot.contour()` 绘制平面等高线
- ◀ `matplotlib.pyplot.contourf()` 绘制填充等高线图
- ◀ `numpy.cos()` 计算余弦值
- ◀ `numpy.diag()` 如果 A 为方阵，`numpy.diag(A)` 函数提取对角线元素，以向量形式输入结果；如果 a 为向量，`numpy.diag(a)` 函数将向量展开成方阵，方阵对角线元素为 a 向量元素
- ◀ `numpy.dot()` 计算向量标量积。值得注意的是，如果输入为一维数组，`numpy.dot()` 输出结果为标量积；如果输入为矩阵，`numpy.dot()` 输出结果为矩阵乘积，相当于矩阵运算符 @
- ◀ `numpy.linalg.inv()` 矩阵求逆
- ◀ `numpy.linalg.norm()` 计算范数
- ◀ `numpy.meshgrid()` 创建网格化数据
- ◀ `numpy.sin()` 计算正弦值
- ◀ `numpy.sqrt()` 计算平方根

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

24.1 使用 patches 绘制平面几何形状

相信大家对 matplotlib.patches 已经不陌生了。matplotlib.patches 是 Matplotlib 库中的一个模块，可以使用它来绘制如图 1 圆形、矩形、多边形、箭头等等。

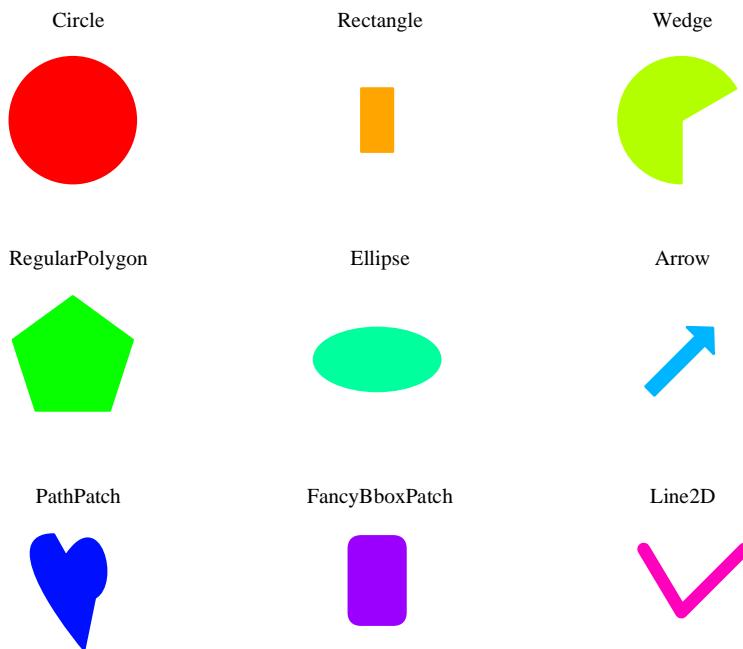


图 1. matplotlib.patches 中常见的几何图形

图 1 参考如下示例：

https://matplotlib.org/stable/gallery/shapes_and_collections/artist_reference.html

图 2 所示为利用 matplotlib.patches 绘制一组单位圆内接、外接正多边形。举个例子，patches.Circle 可以创建一个圆形对象。这个对象可以具有不同的参数，如位置、大小、边框颜色、填充颜色、阴影线等。

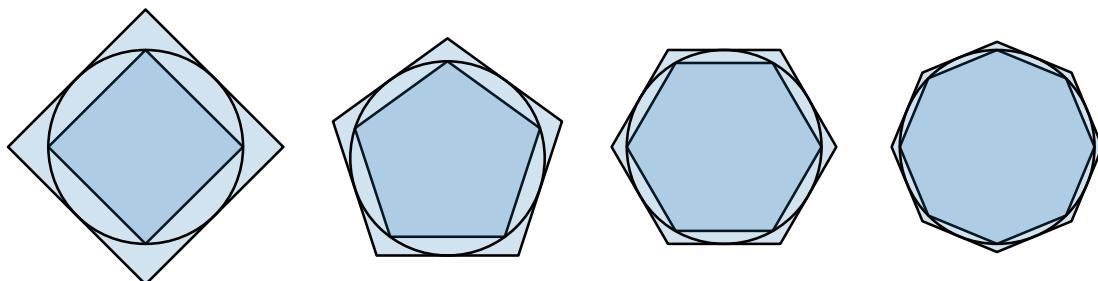


图 2. 利用 patches 绘制正圆，以及外切、内接正多边形

《数学要素》第 3 章中，大家会看到我们利用图 2 介绍如何估算圆周率。



Jupyter 笔记 BK_2_Ch24_1.ipynb 绘制图 2。

斐波那契数列可以在自然界中找到很多例子，如植物的分枝结构、蜂窝的排列方式等。这是因为斐波那契数列具有一种黄金分割比例的特性，在自然界中被广泛运用。斐波那契数列在计算机科学和编程中有着广泛的应用。例如，它可以用来设计递归算法和动态规划算法，解决一些复杂的问题。斐波那契数列与黄金分割比例、数学规律等有着密切的联系。它在代数、数论、几何等数学领域中都有重要的应用。

24.2 使用等高线绘制平面几何形状

“鸢尾花书”中，我们更常见的是利用平面等高线可视化平面几何形状。

图 16 上图所示为利用等高线绘制的一组圆锥曲线。通过在 $[0, 3]$ 范围之内改变离心率，圆锥曲线从正圆、椭圆，最终变成双曲线。绘制每条曲线时，我们先设置离心率，然后利用网格数据生成特定圆锥曲线的数据。绘制等高线时，仅仅绘制等高线值为 1 的那一条曲线。并且，利用色谱我们生成一组连续变化的颜色，分别渲染每一条圆锥曲线。



《数学要素》第 9 章将介绍如何通过设定离心率改变圆锥曲线类型。

图 16 下图绘制的是在给定椭圆上不同点处的切线。绘制这幅图时需要用到椭圆切线的解析式，《矩阵力量》第 20 章将专门讲解这一话题。

下面我们看两个更复杂的例子。如图 17 上图所示，给定矩形，绘制一组和矩形相切的椭圆。图中的矩形用 `matplotlib.patches` 绘制。而椭圆采用等高线绘制。



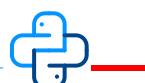
《数学要素》第 9 章会专门讲解这组椭圆的性质。

如图 17 下图所示，给定旋转椭圆，绘制一组和椭圆相切的矩形。椭圆采用参数方程绘制，而矩形采用 `matplotlib.patches`。

绘制矩形还用到了仿射变换 (affine transformation)。本书后续将专门讲解仿射变换。



《统计至简》第 14 章将讲解图 17 下图用到的数学工具。



Jupyter 笔记 BK_2_Ch24_2.ipynb 绘制图 16 上图。Jupyter 笔记 BK_2_Ch24_3.ipynb 绘制图 16 下图。
Jupyter 笔记 BK_2_Ch24_4.ipynb 绘制图 17 上图。Jupyter 笔记 BK_2_Ch24_5.ipynb 绘制图 17 下图。



想要理解如何用 `patches` 绘制各种几何图形，大家可以参考如下链接：

https://matplotlib.org/stable/api/patches_api.html

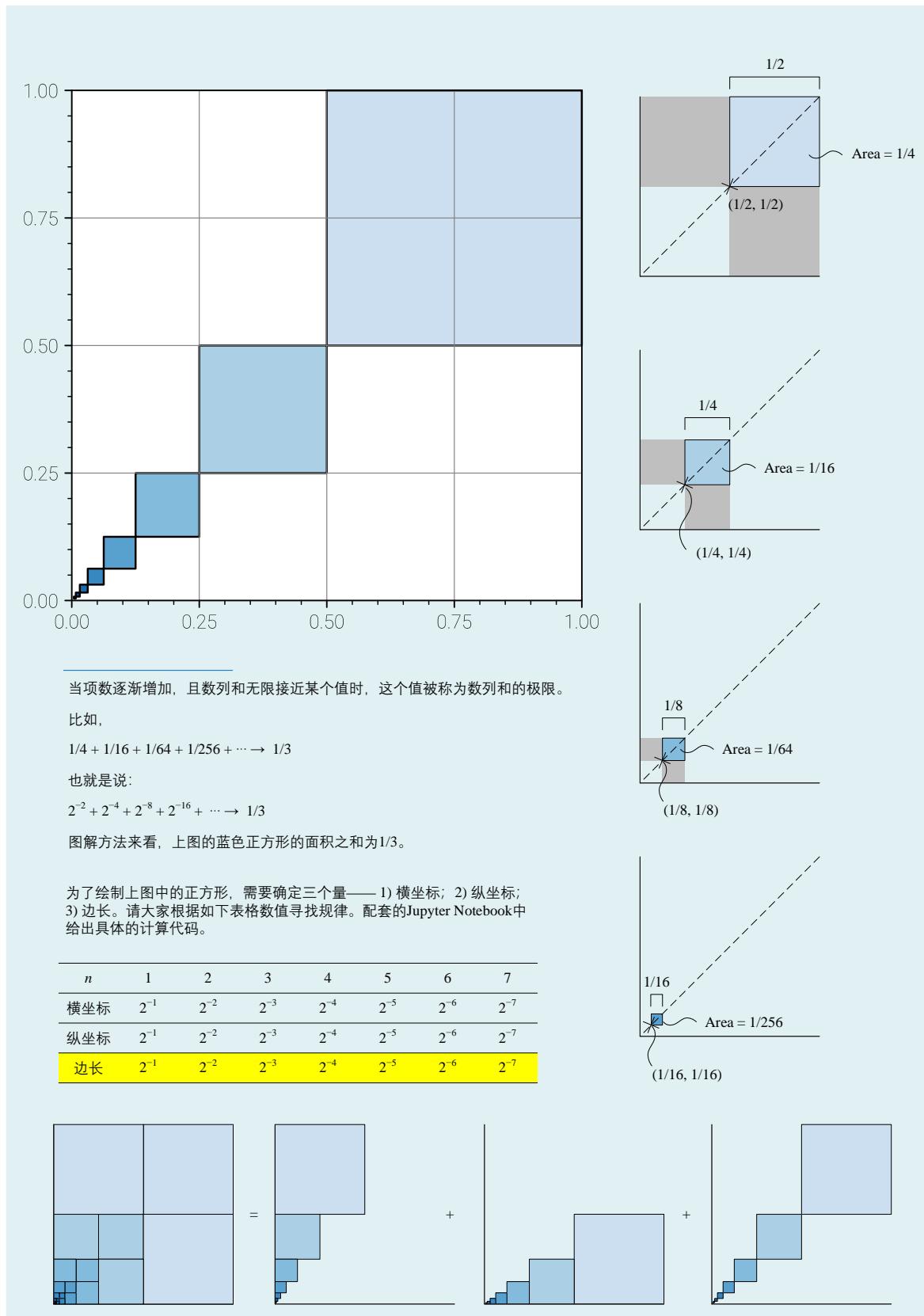


图 3. 可视化极限，第 1 组

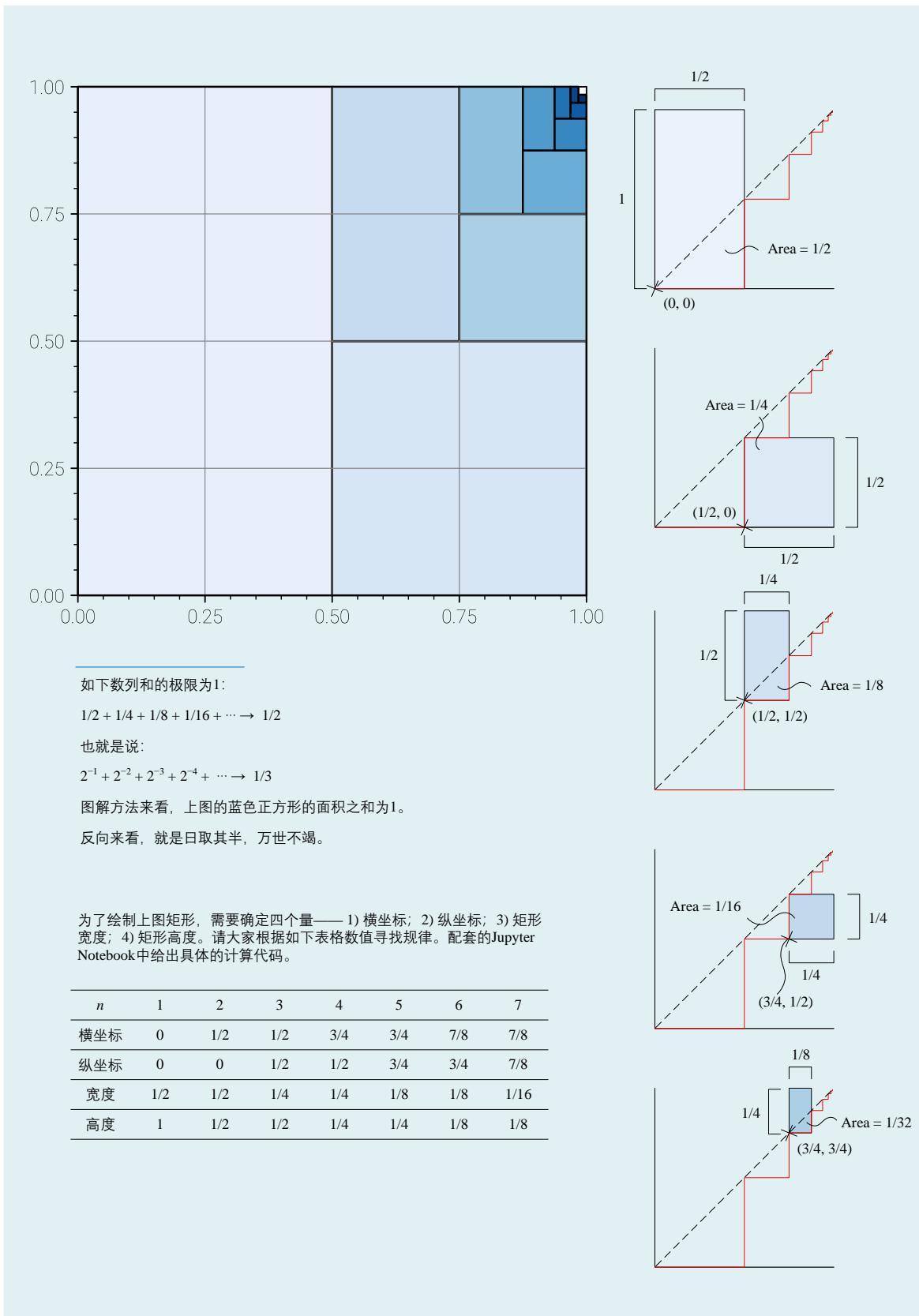


图 4. 可视化极限，第 2 组

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

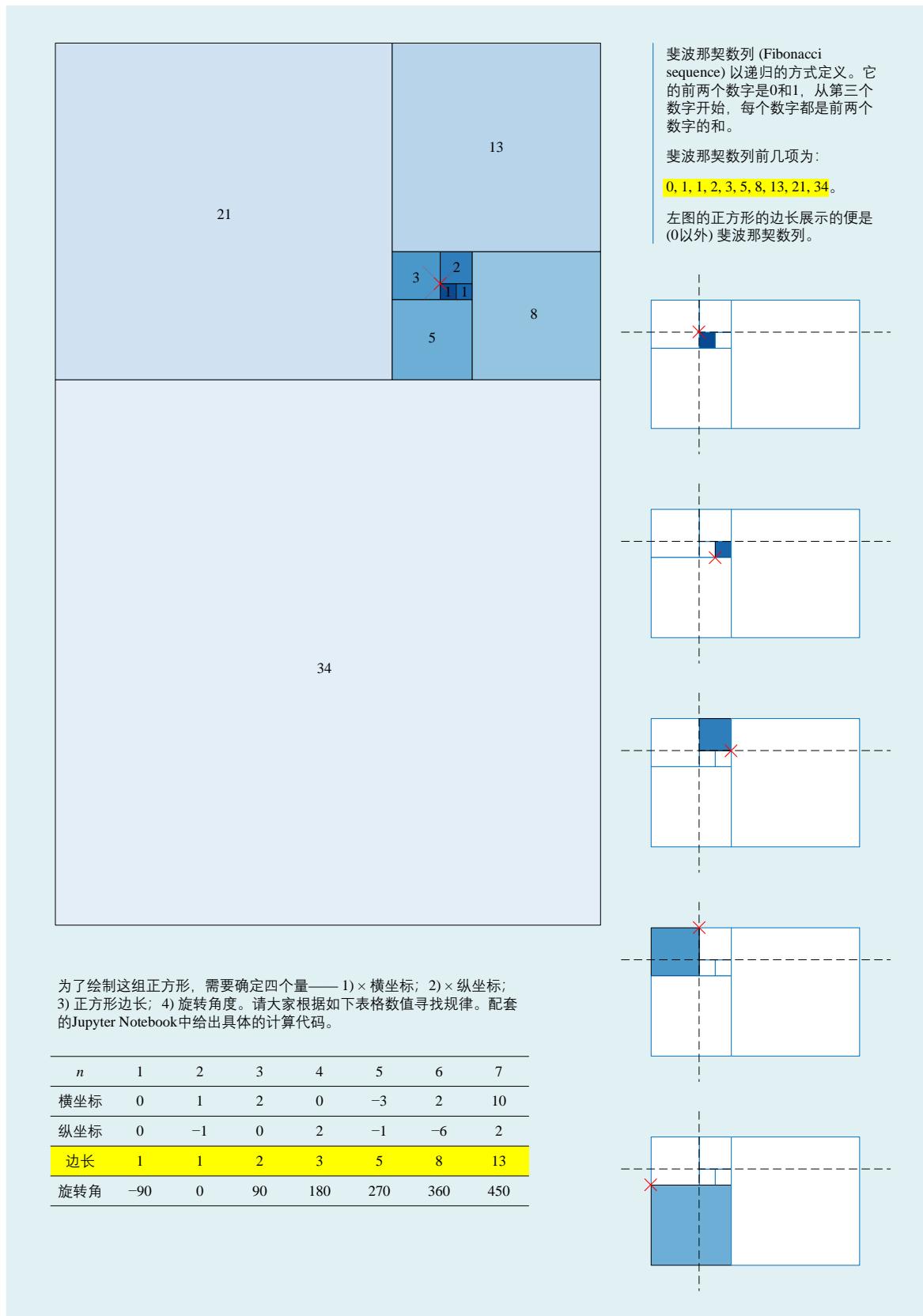


图 5. 可视化斐波那契数列

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

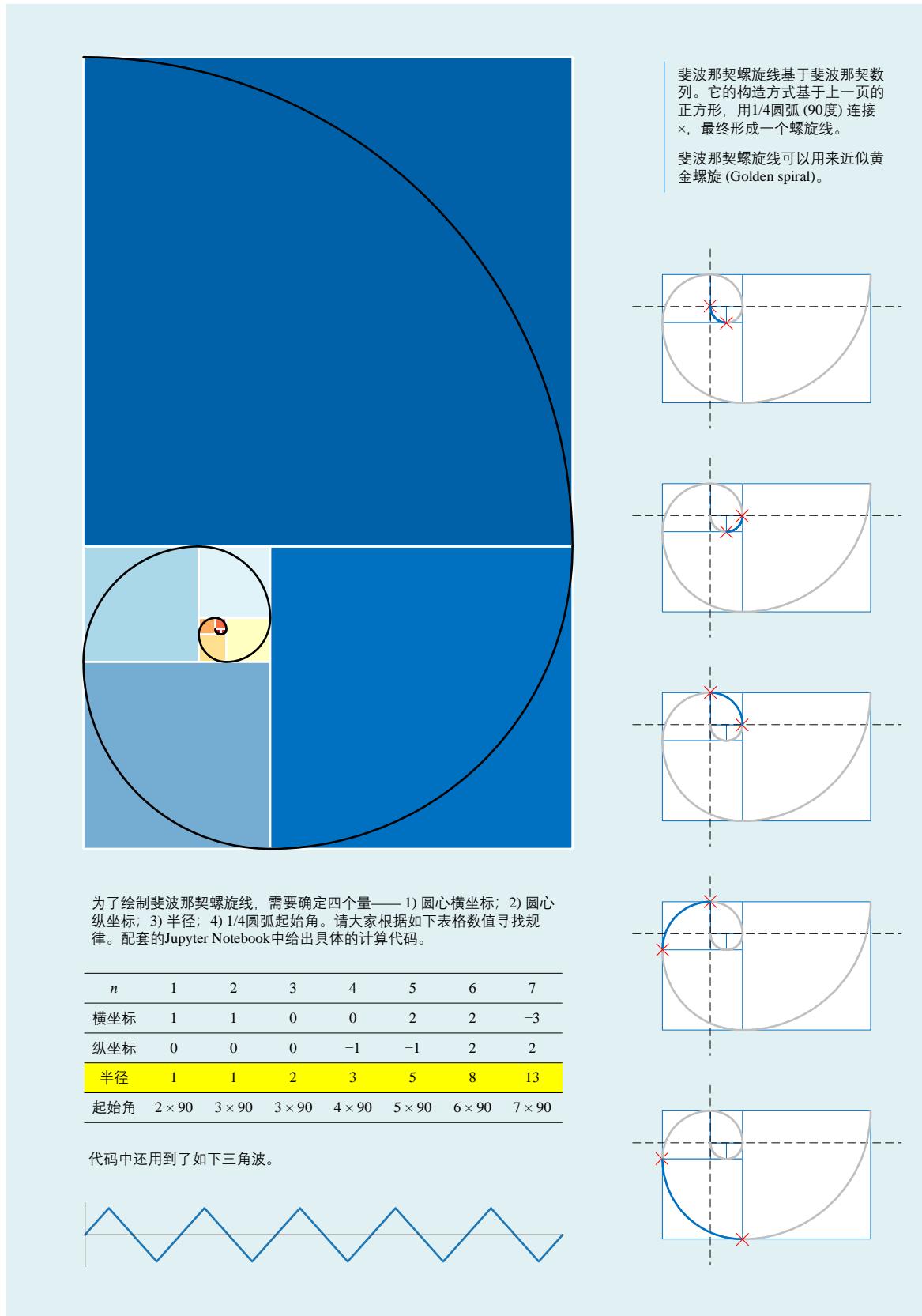


图 6. 可视化斐波那契螺旋线

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

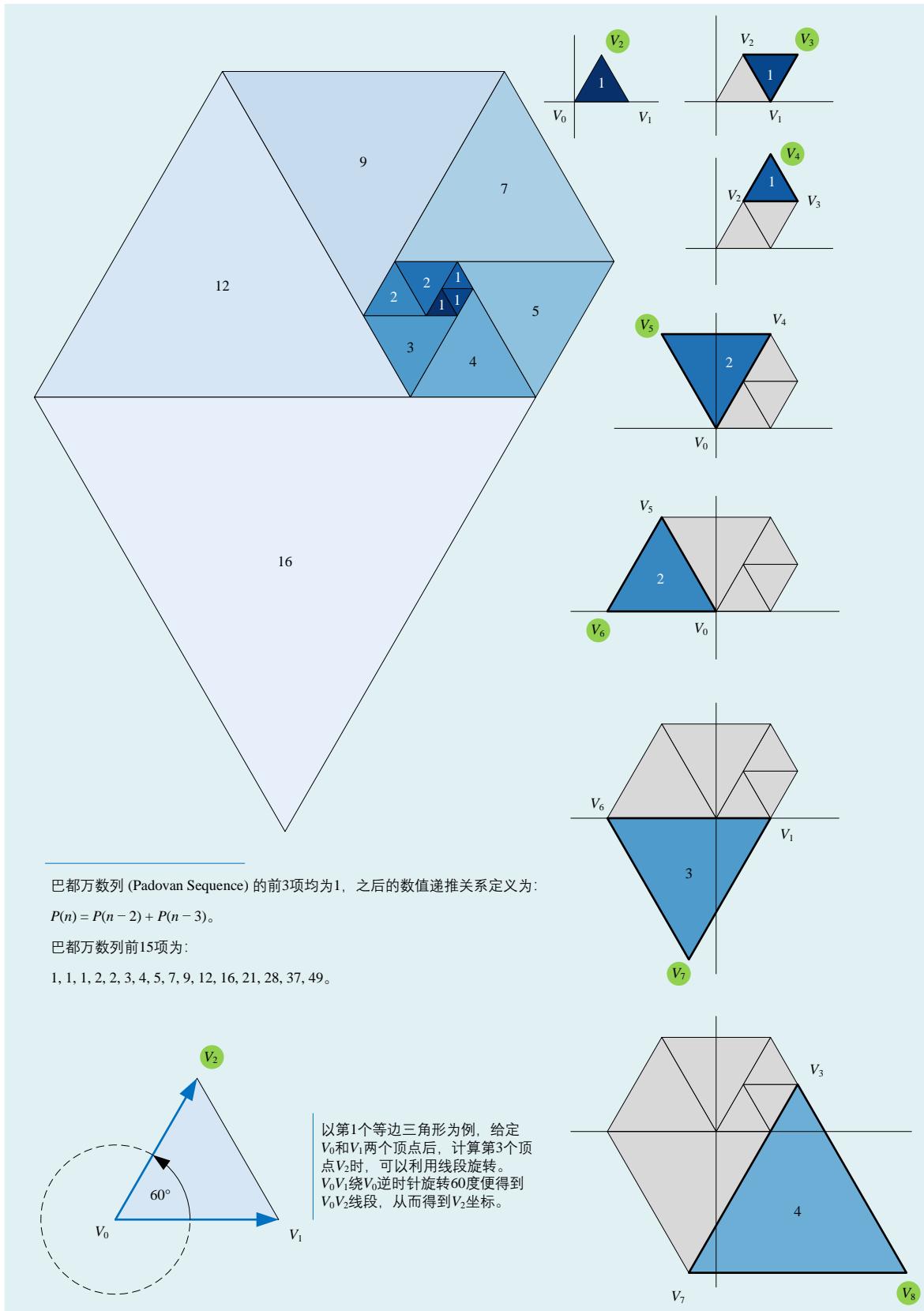


图 7. 可视化雷卡曼数列

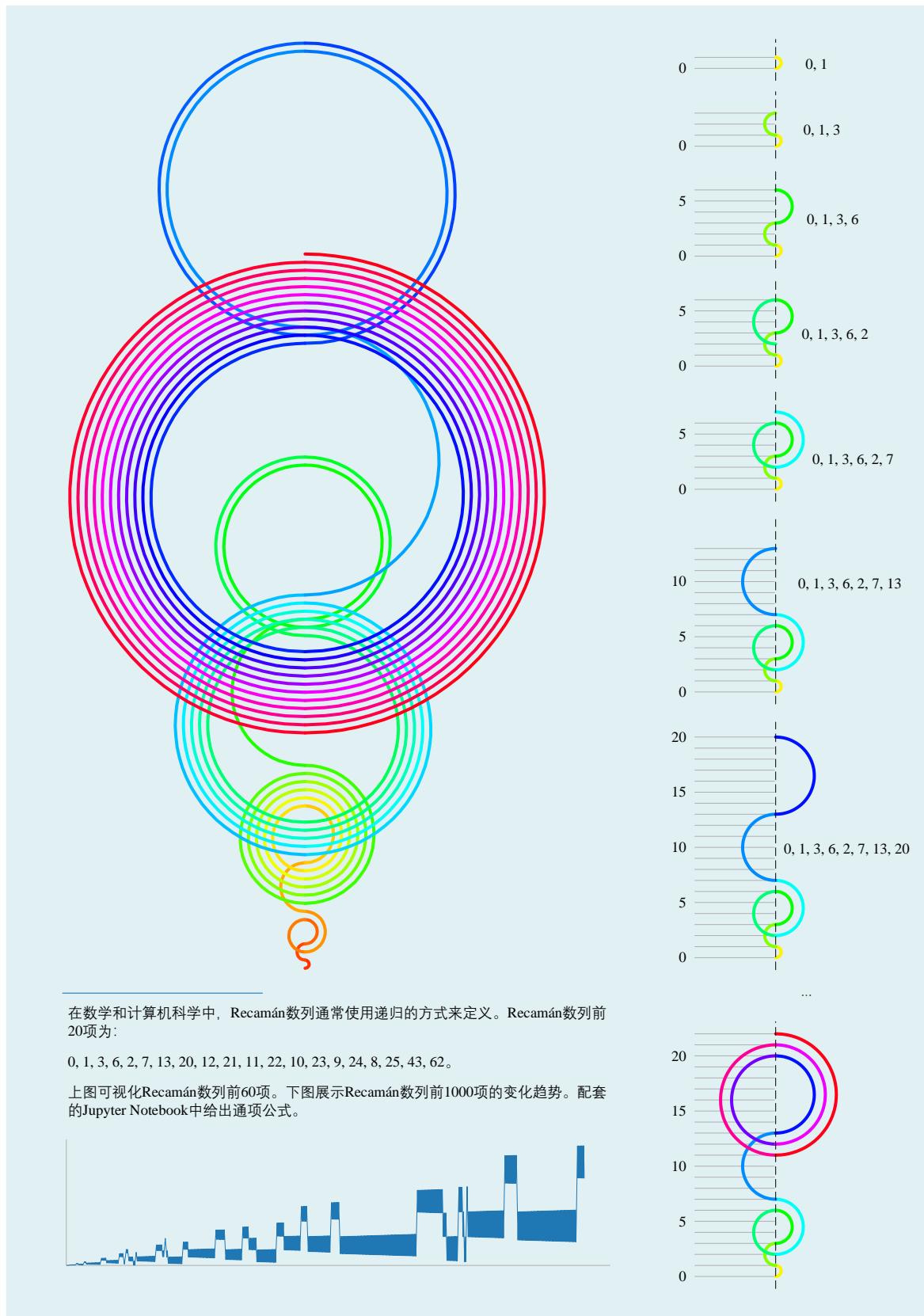
本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com



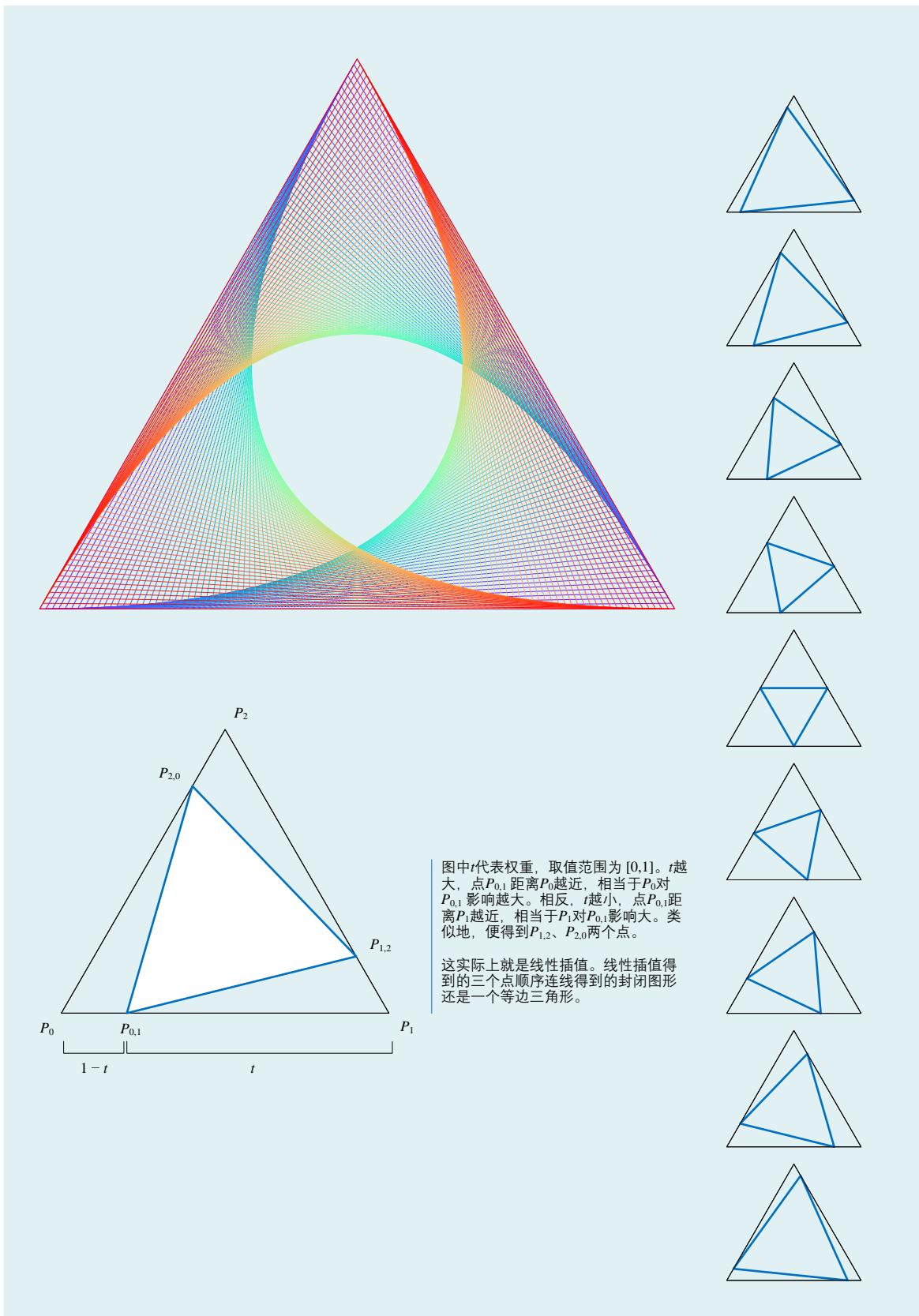


图 9. 线性插值，等边三角形

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

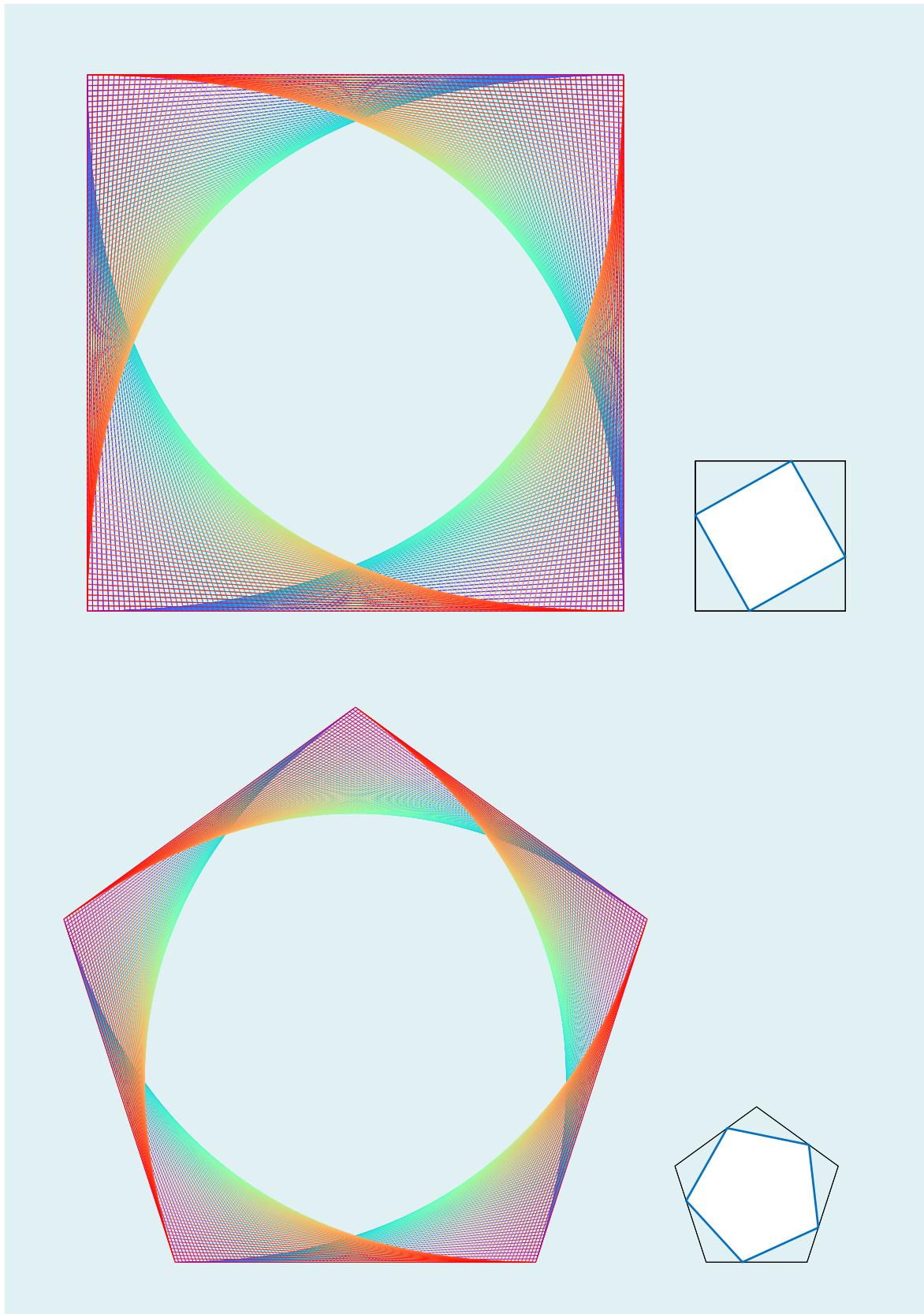


图 10. 线性插值，正方形、正五边形

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

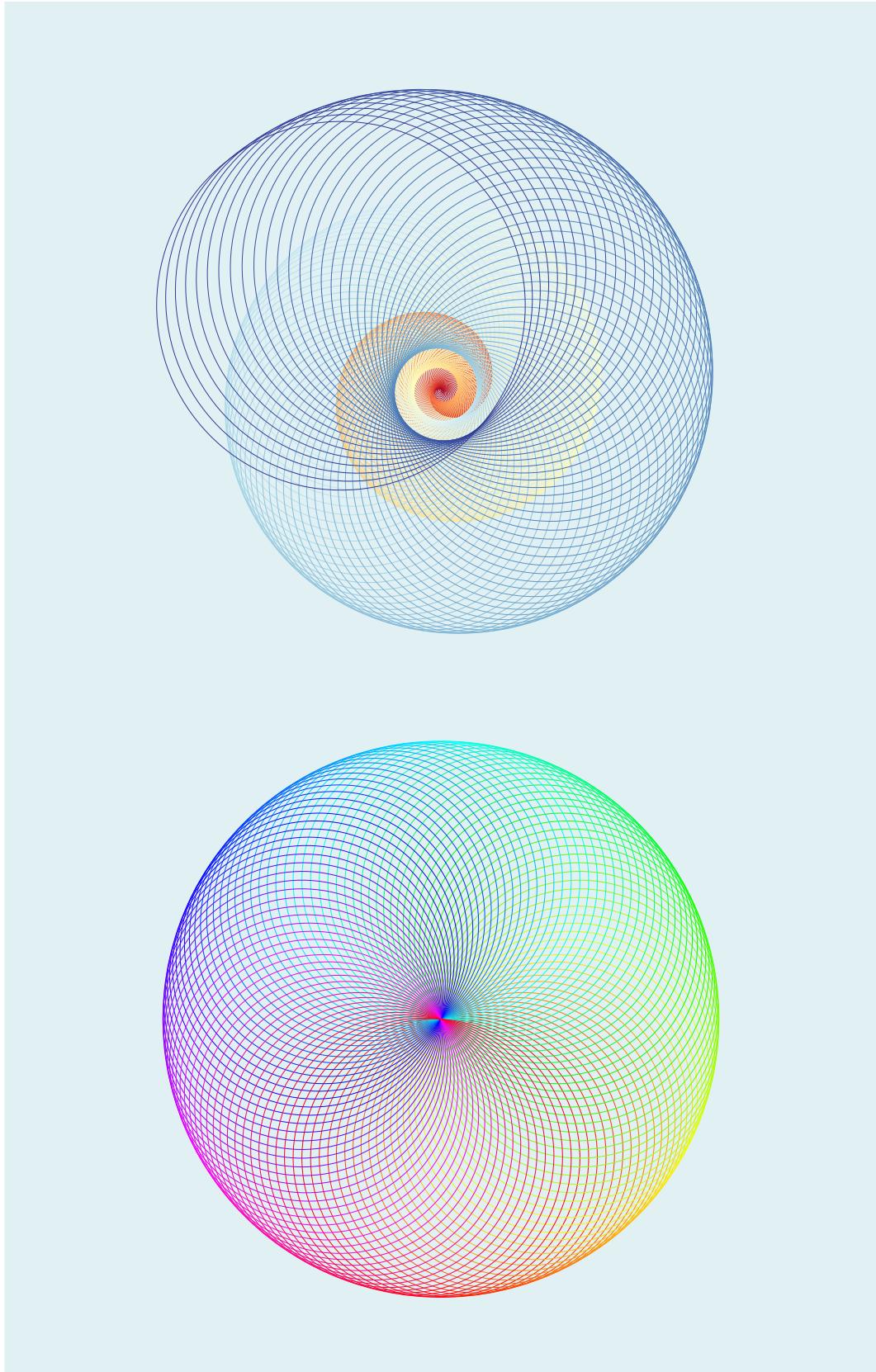


图 11. 两组旋转正圆

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

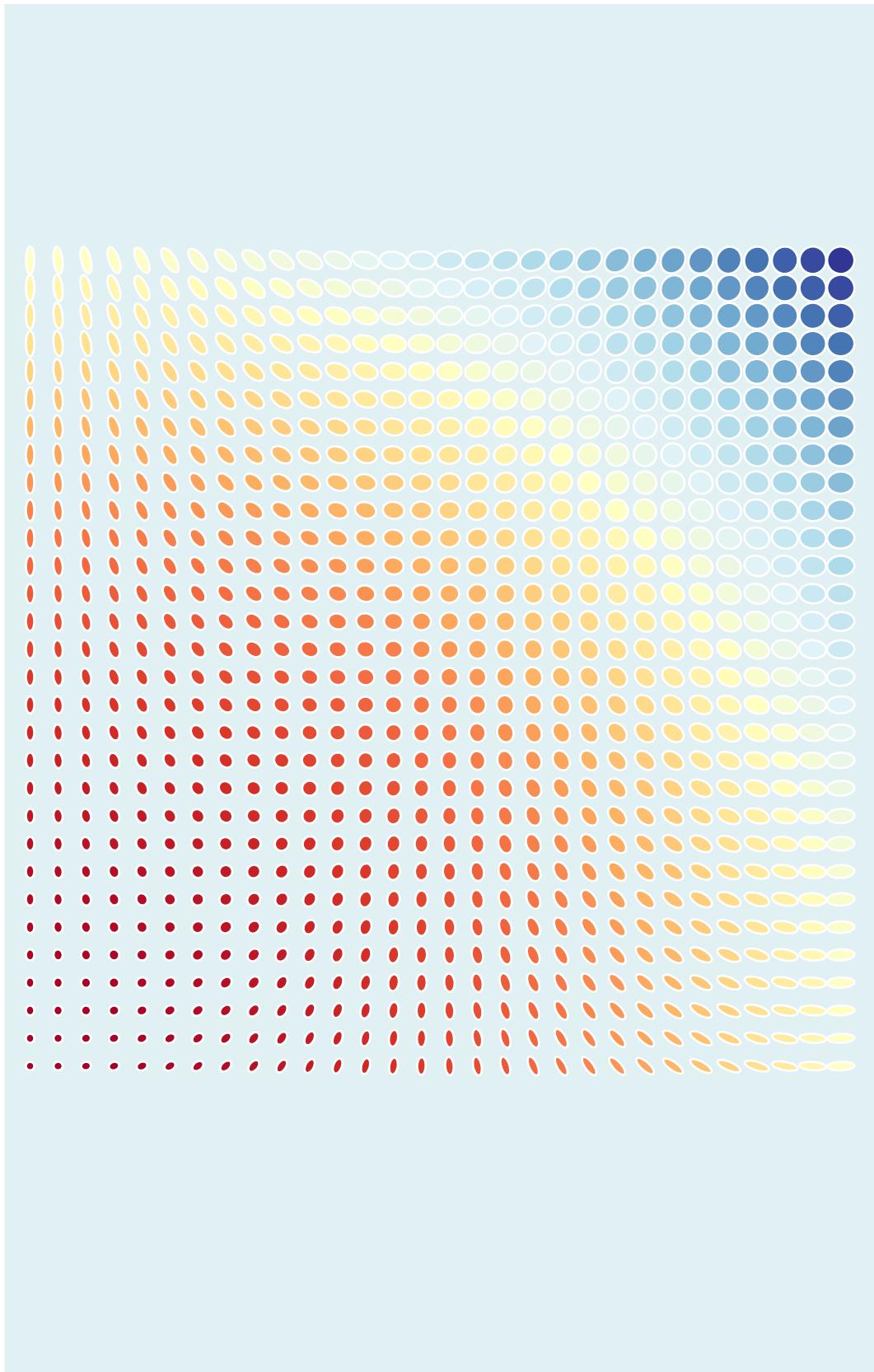


图 12. 一组旋转椭圆

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

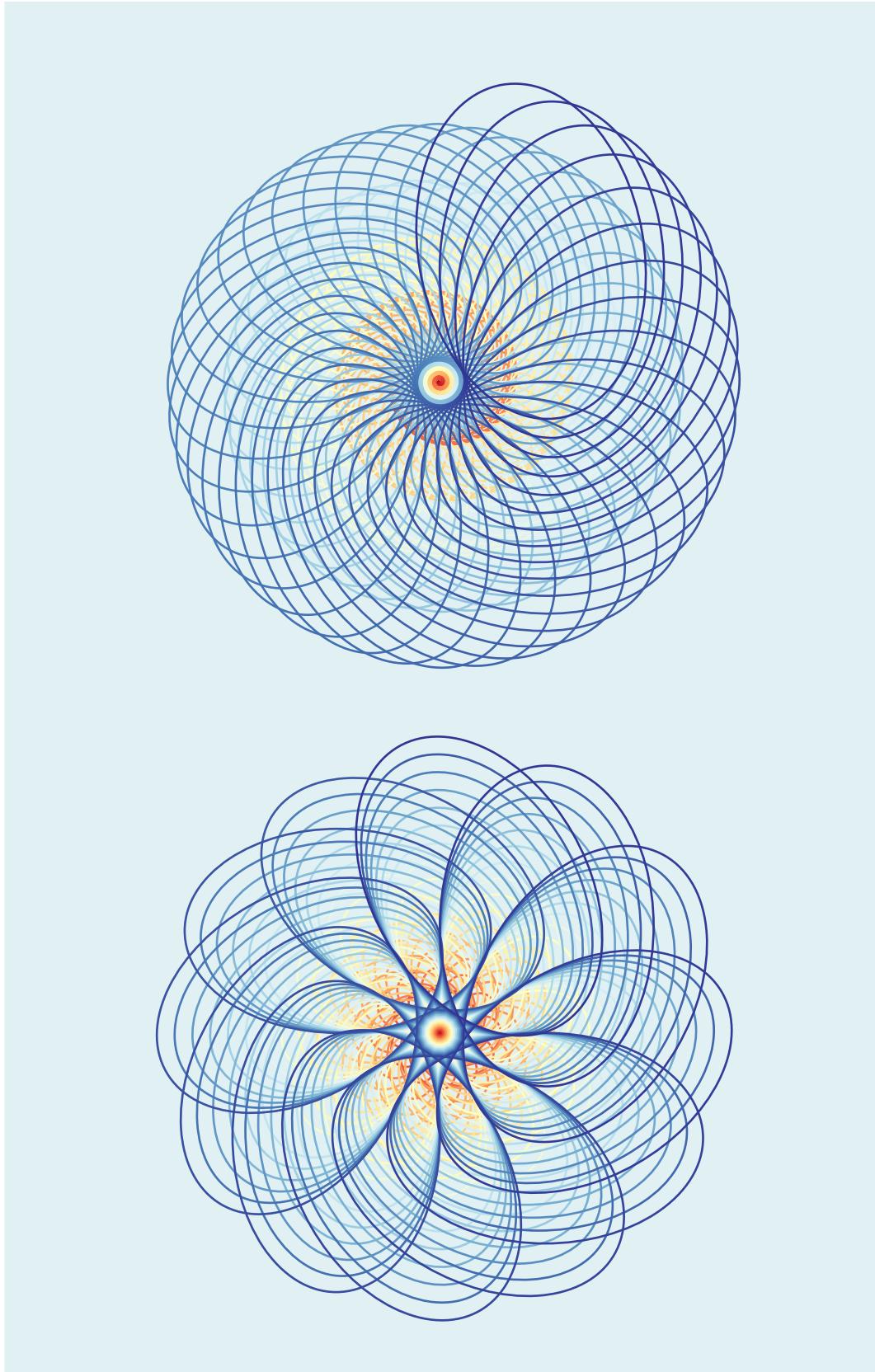


图 13. 两组旋转椭圆

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

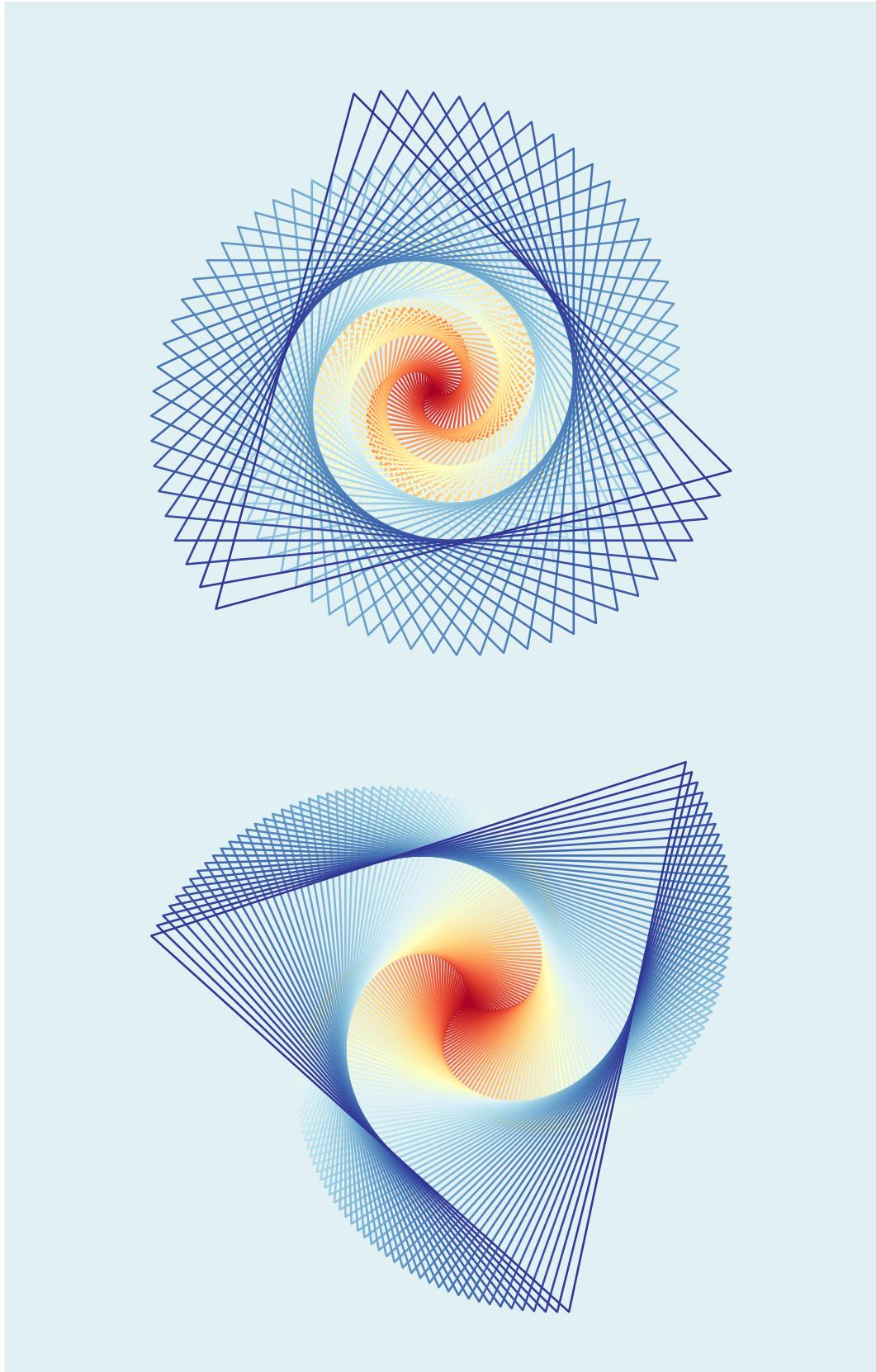


图 14. 两组旋转三角形

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

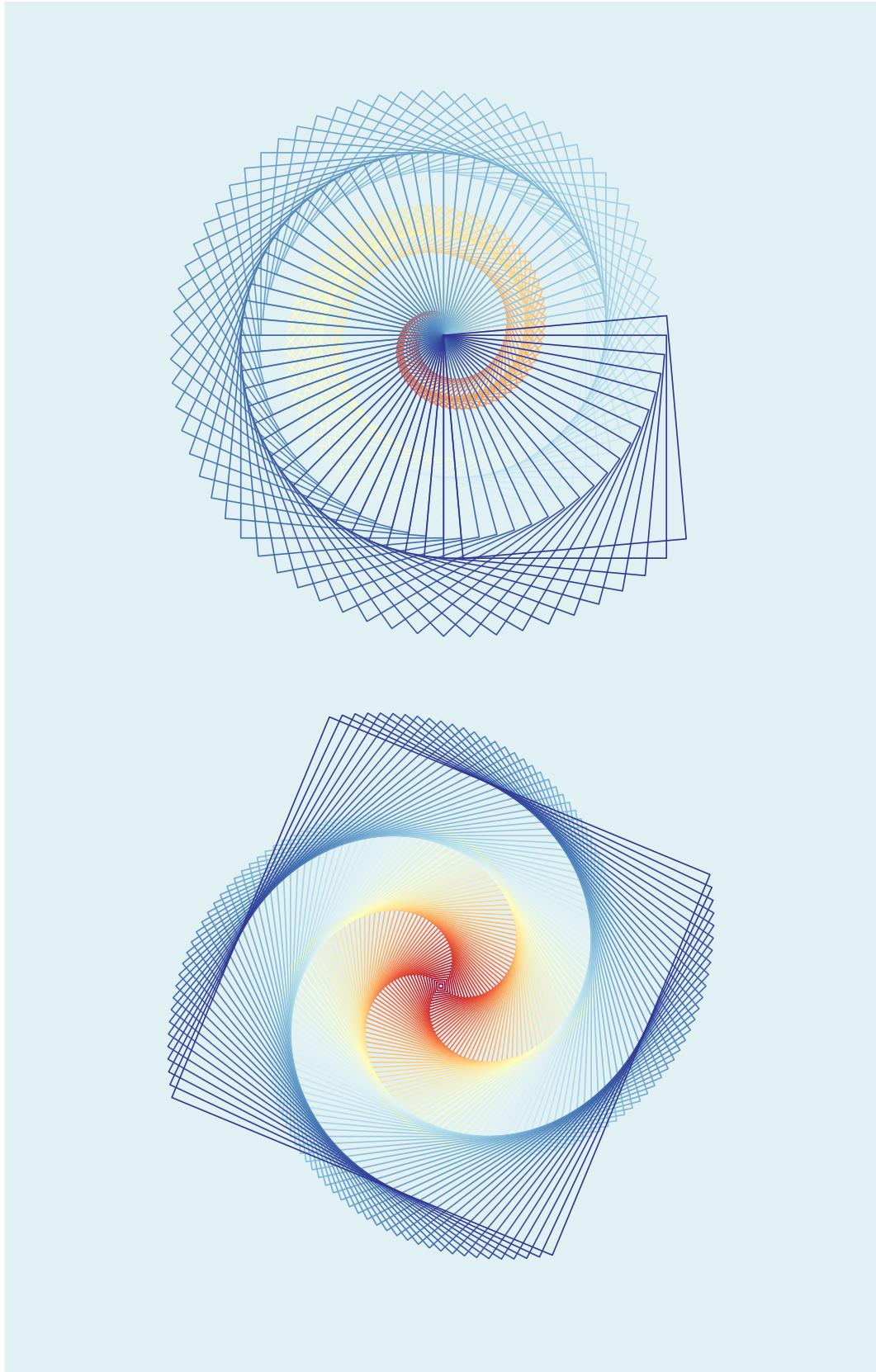


图 15. 两组旋转正方形

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

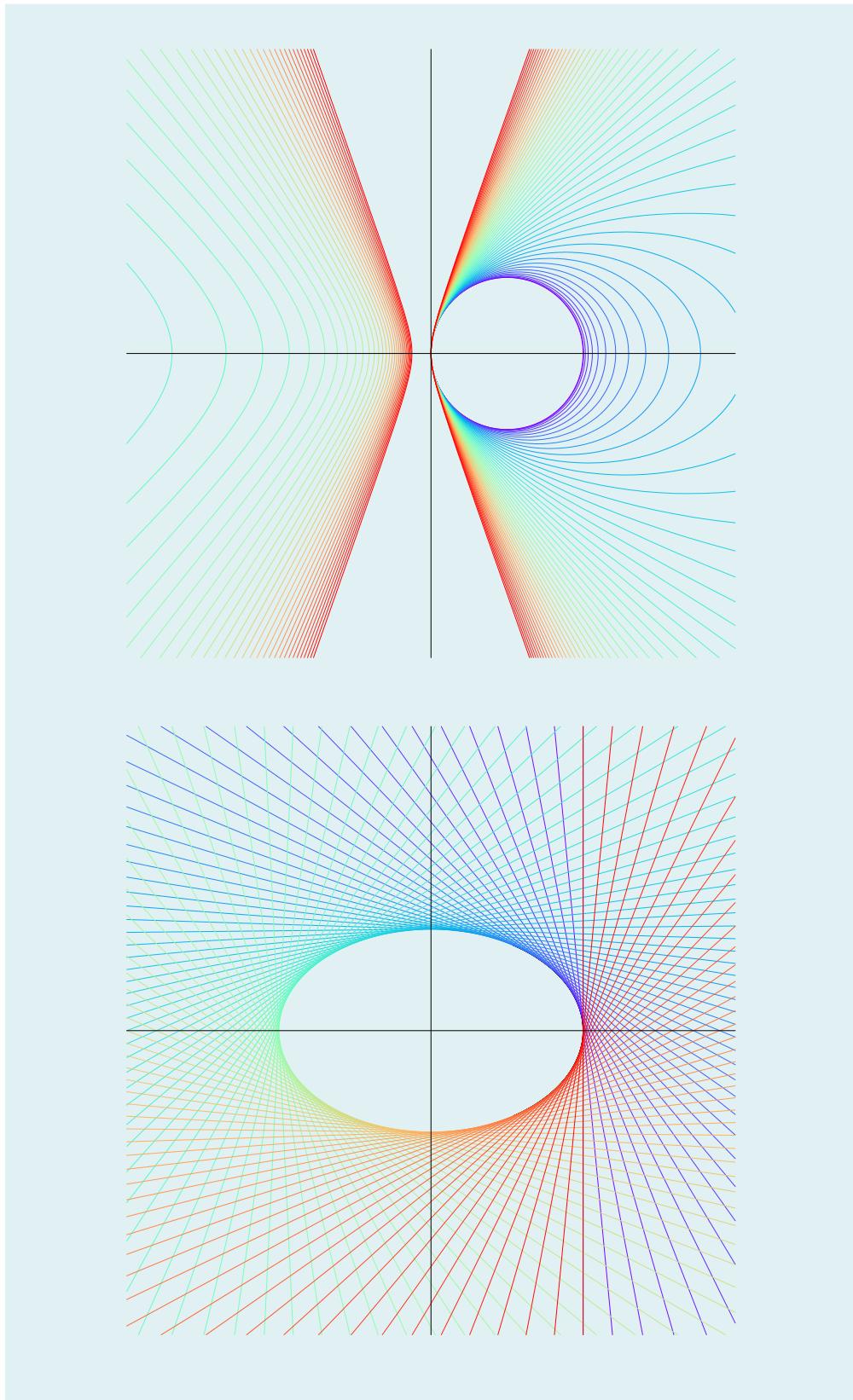


图 16. 利用等高线绘制圆锥曲线、椭圆切线

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

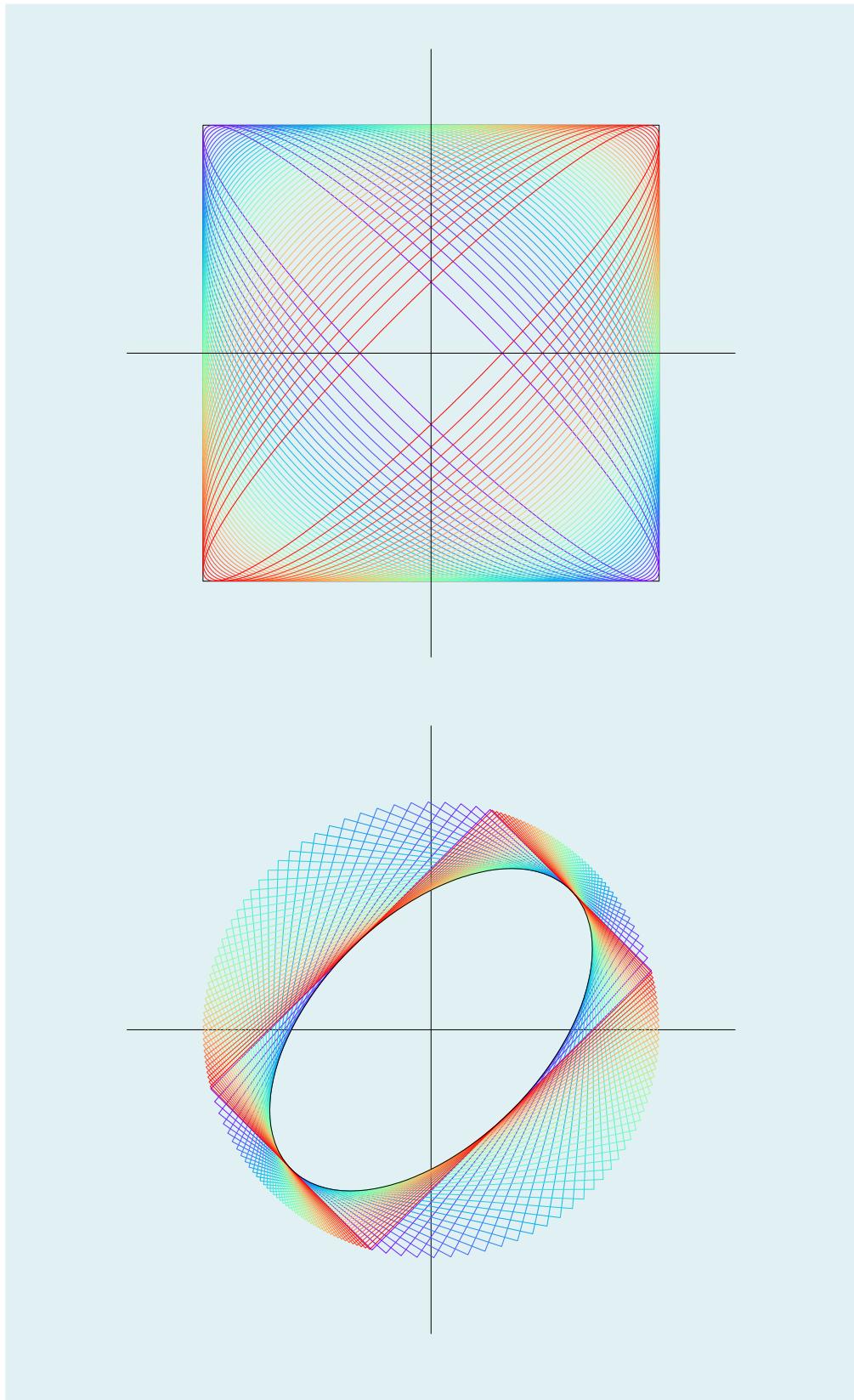


图 17. 给定矩形相切的一组椭圆、给定椭圆相切的一组矩形

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

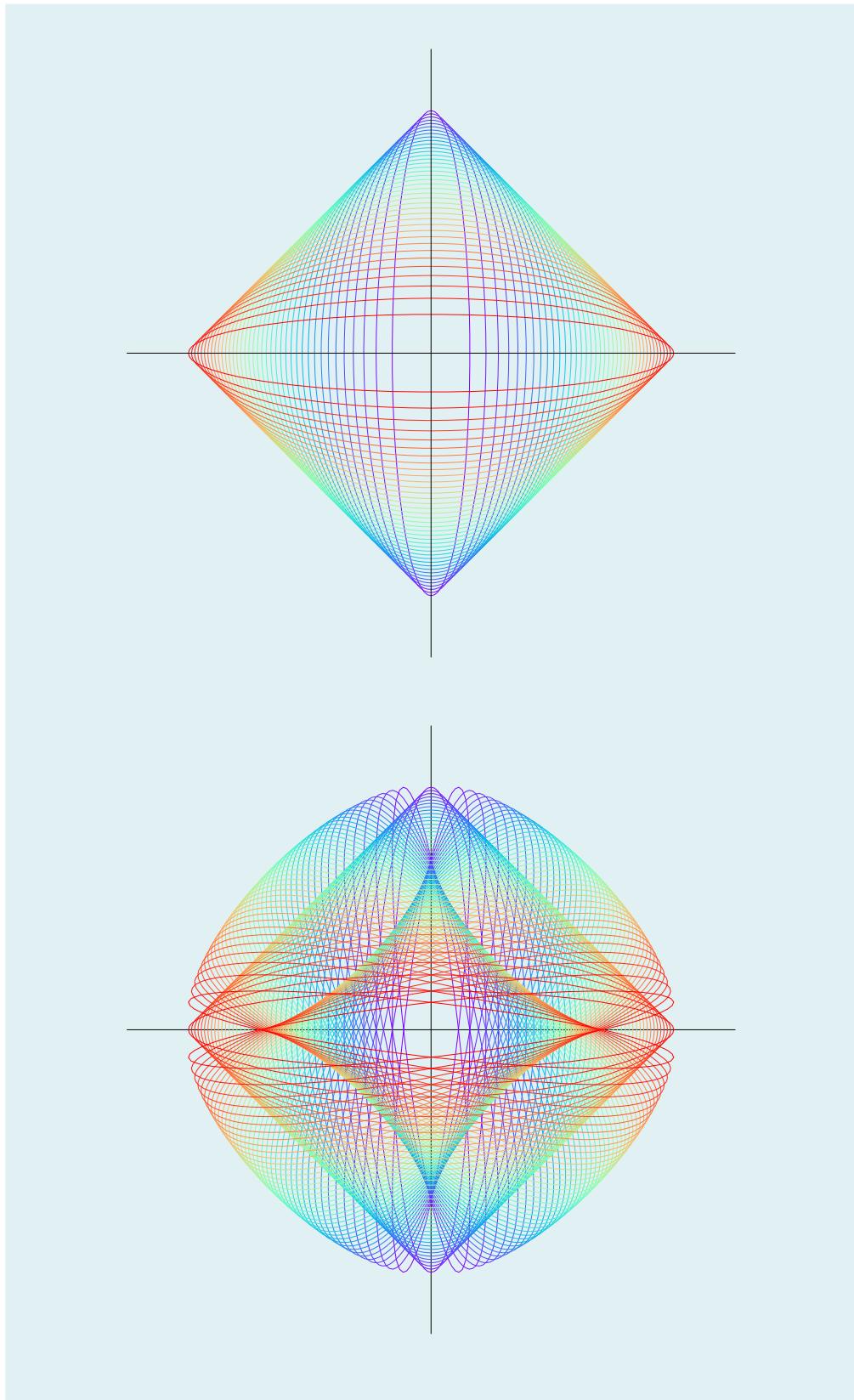


图 18. 一组椭圆，长半轴平方、短半轴平方之和为定值

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

25

Solid Geometry

立体几何

用 Matplotlib 绘制立体几何形状



艺术家，胸怀宇宙，手握繁星。

The painter has the Universe in his mind and hands.

——列奥纳多·达·芬奇 (Leonardo da Vinci) | 文艺复兴三杰之一 | 1452 ~ 1519



- ◀ Axes3D.bar3d() 绘制三维柱状图
- ◀ Axes3D.plot_surface() 绘制三维曲面
- ◀ Axes3D.voxels() 绘制三维 voxels 图
- ◀ matplotlib.pyplot.contour() 绘制等高线图
- ◀ matplotlib.pyplot.plot_trisurf() 在三角形网格上绘制平滑的三维曲面图
- ◀ sympy.abc import x 定义符号变量 x
- ◀ sympy.integrate() 符号积分
- ◀ sympy.symbols() 定义符号变量

25.1 绘制几何体的几种方法

本书前文介绍过利用散点、线图、等高线、网格面等三维空间可视化方案。

用网格面

图 1 和图 4 所示为利用参数方程绘制的四个几何体。

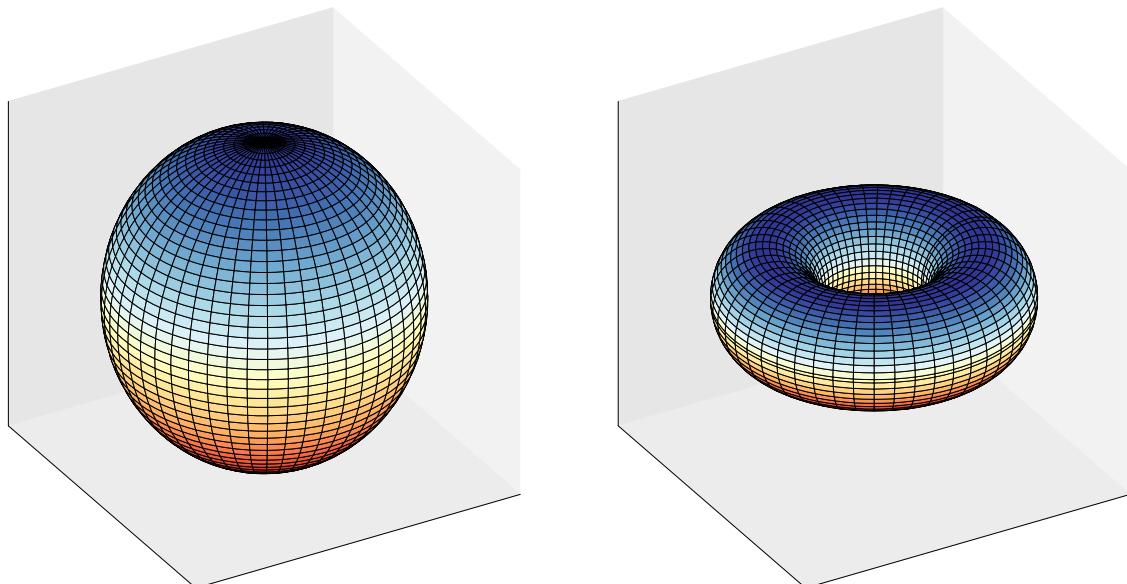


图 1. 用网格面绘制几何体，参数方程，第 1 组

用三维柱状图

图 2 所示为利用三维柱状图绘制的几何体。

图 2 实际上展示的二重积分估算，《数学要素》第 18 章将深入介绍这个数学工具。

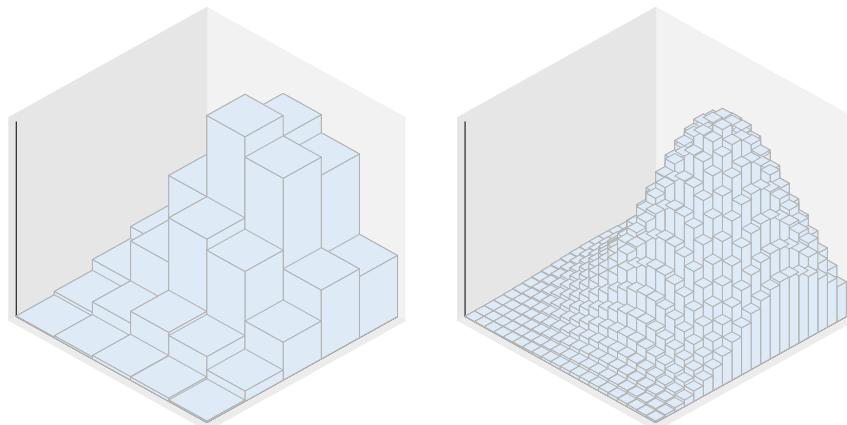


图 2. 用柱状图绘制几何体

三角网格

图 5 所示为三角网格绘制的几何体。本书后续将专门介绍三角网格。

用 voxels 绘制立体几何体

Matplotlib 中的 voxels 是一种用于三维数据可视化的功能。它将三维数据集表示为一系列的小方块，其中每个方块的位置、大小和颜色都可以自定义。通过使用 voxels，可以直观地展示复杂的三维数据结构，如体积数据、分子模型等。voxels 支持不同的绘制样式，包括实心方块和透明方块，可以根据需要进行调整。此外，还可以添加轴标签、标题和图例等来增强可视化效果。

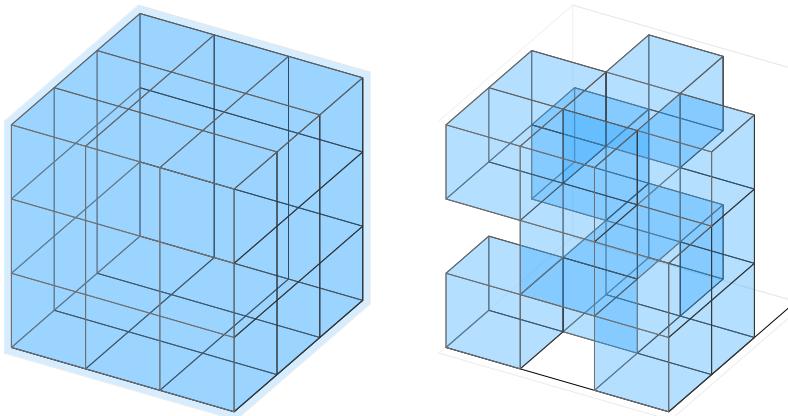


图 3. 用 voxels 绘制几何体

25.2 用等高线绘制三维几何体

本章前文提到的这几种绘制立体几何形状的可视化方案外，本章要重点展示的是用沿 x 、 y 、 z 方向等高线可视化三维几何体。如图 7 ~ 图 11 所示，前三列图形叠加得到第四列几何体。

这些几何体都是用隐函数写成，《数学要素》第 10 章将会介绍隐函数。

图 13 所示为一个旋转椭球 (ellipsoid) 在三个不同平面的投影。注意，这个椭球不是一般的椭球，它代表了一个 3×3 的协方差矩阵的马氏距离为 1 的“等距线”。图 14 所示为图 13 旋转椭球“摆正”后的椭圆，及其在三个平面的投影。也就是说，图 13 和图 14 中两个椭球大小完全一致，空间旋转角度不同而已。

本书后续将介绍更多立体几何变换。

鸢尾花书《统计至简》将会详细介绍相关数学工具。

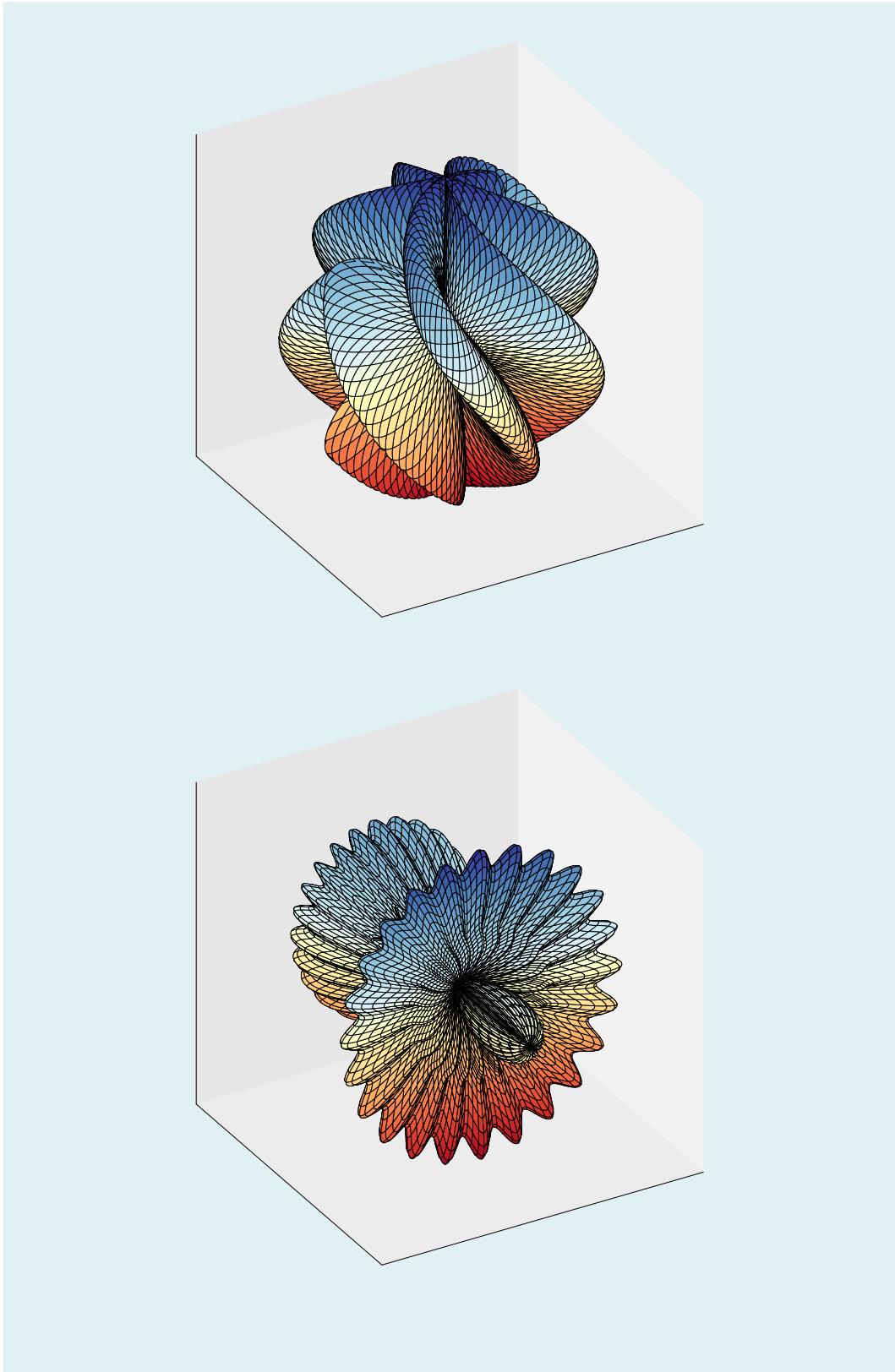


图 4. 用网格面绘制几何体，参数方程，第 2 组

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

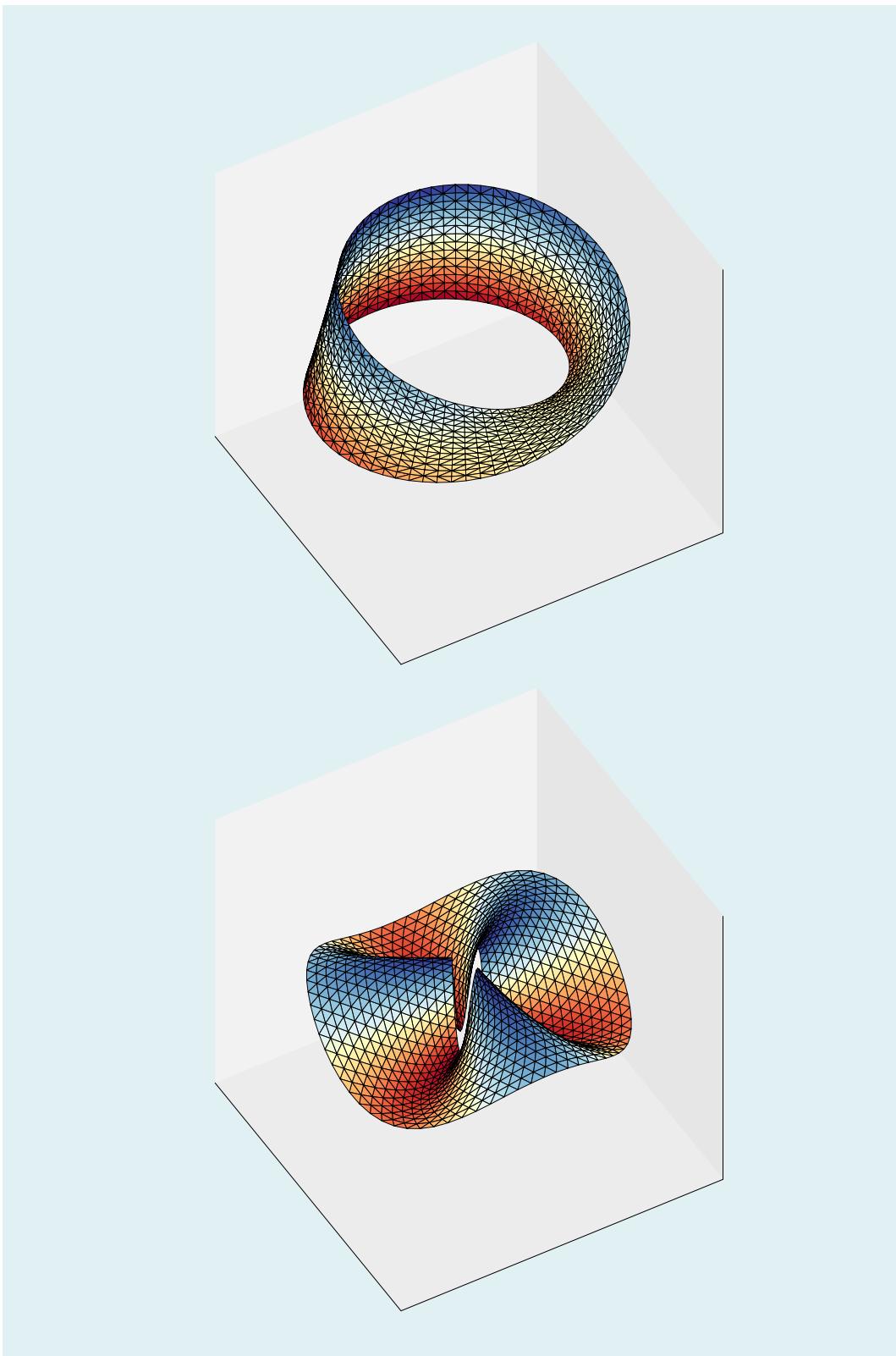


图 5. 用三角网格绘制几何体，来自 Matplotlib 示例

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

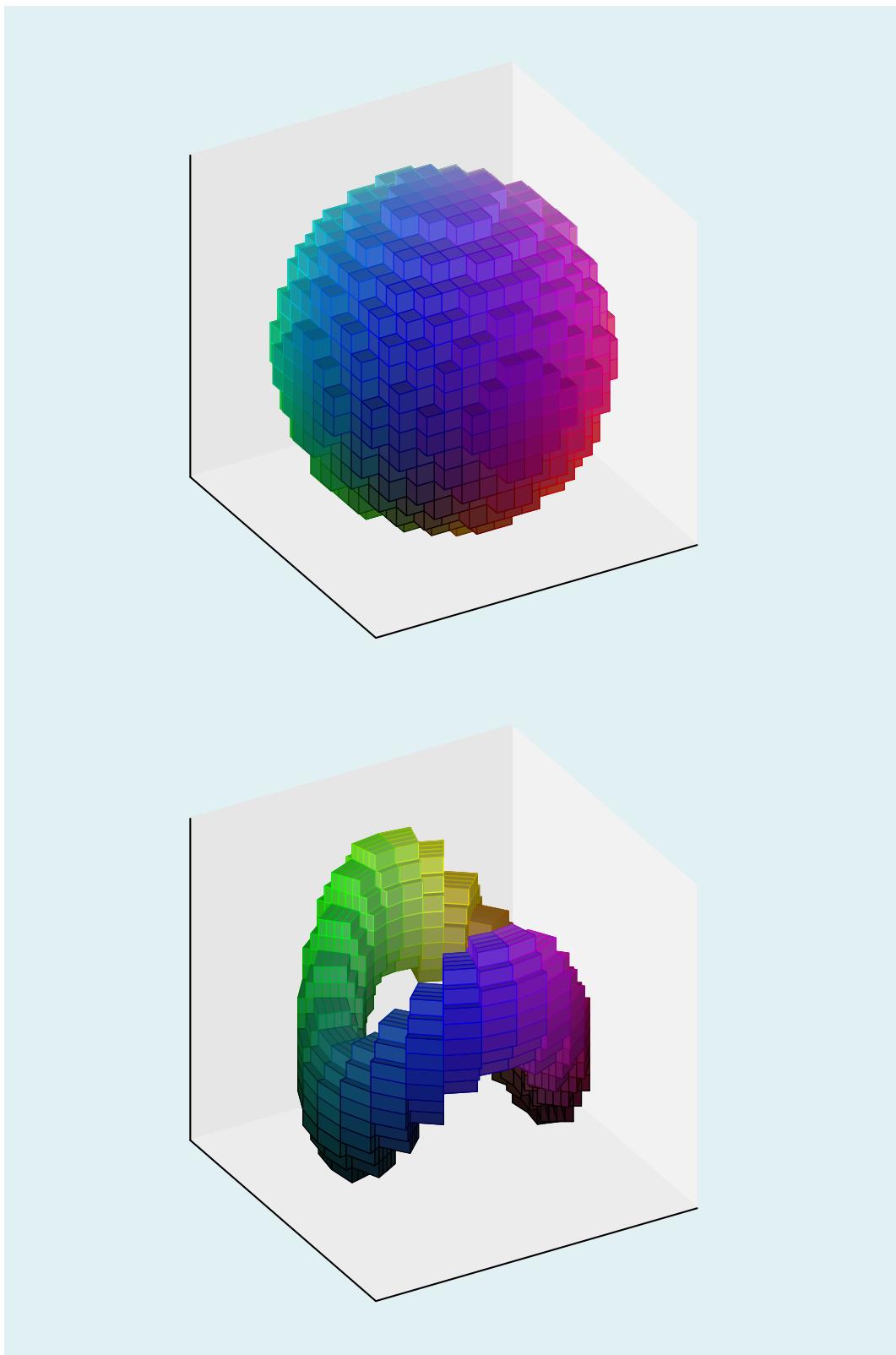


图 6. 用 voxels 绘制几何体，RGB、CMYK 色彩空间

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

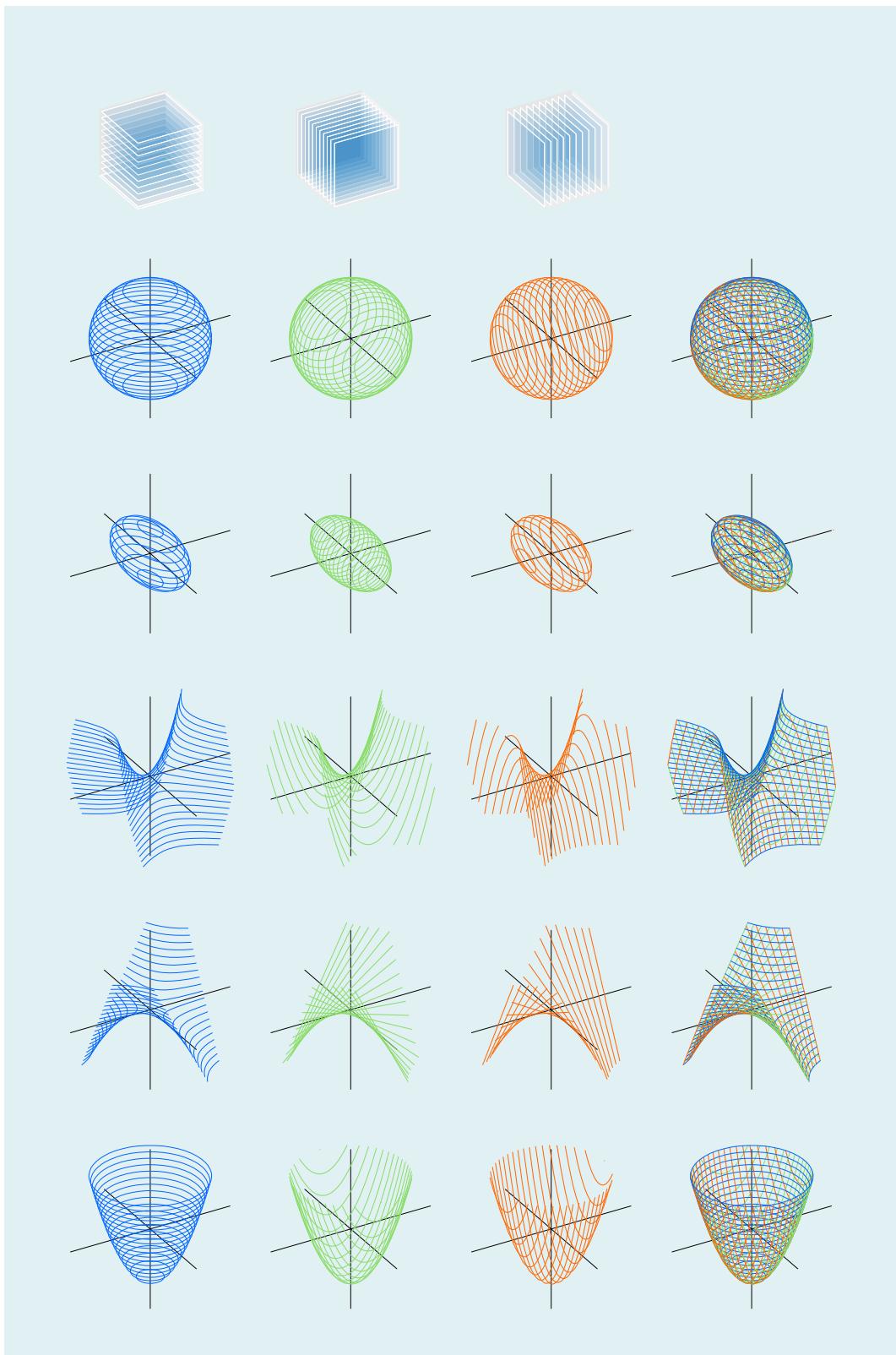


图 7. 用等高线可视化隐函数曲面，第 1 组图形

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

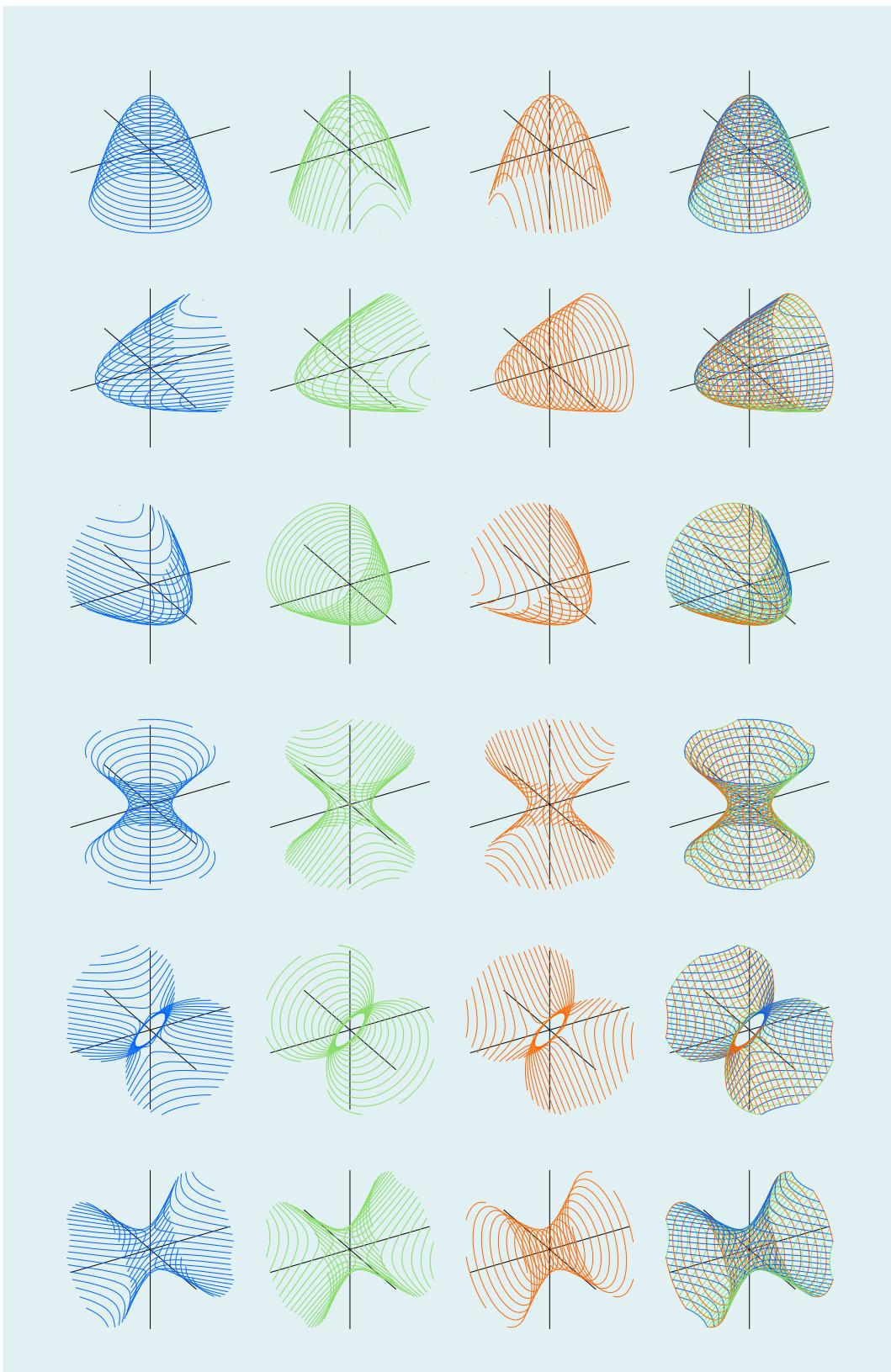


图 8. 用等高线可视化隐函数曲面，第 2 组图形

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

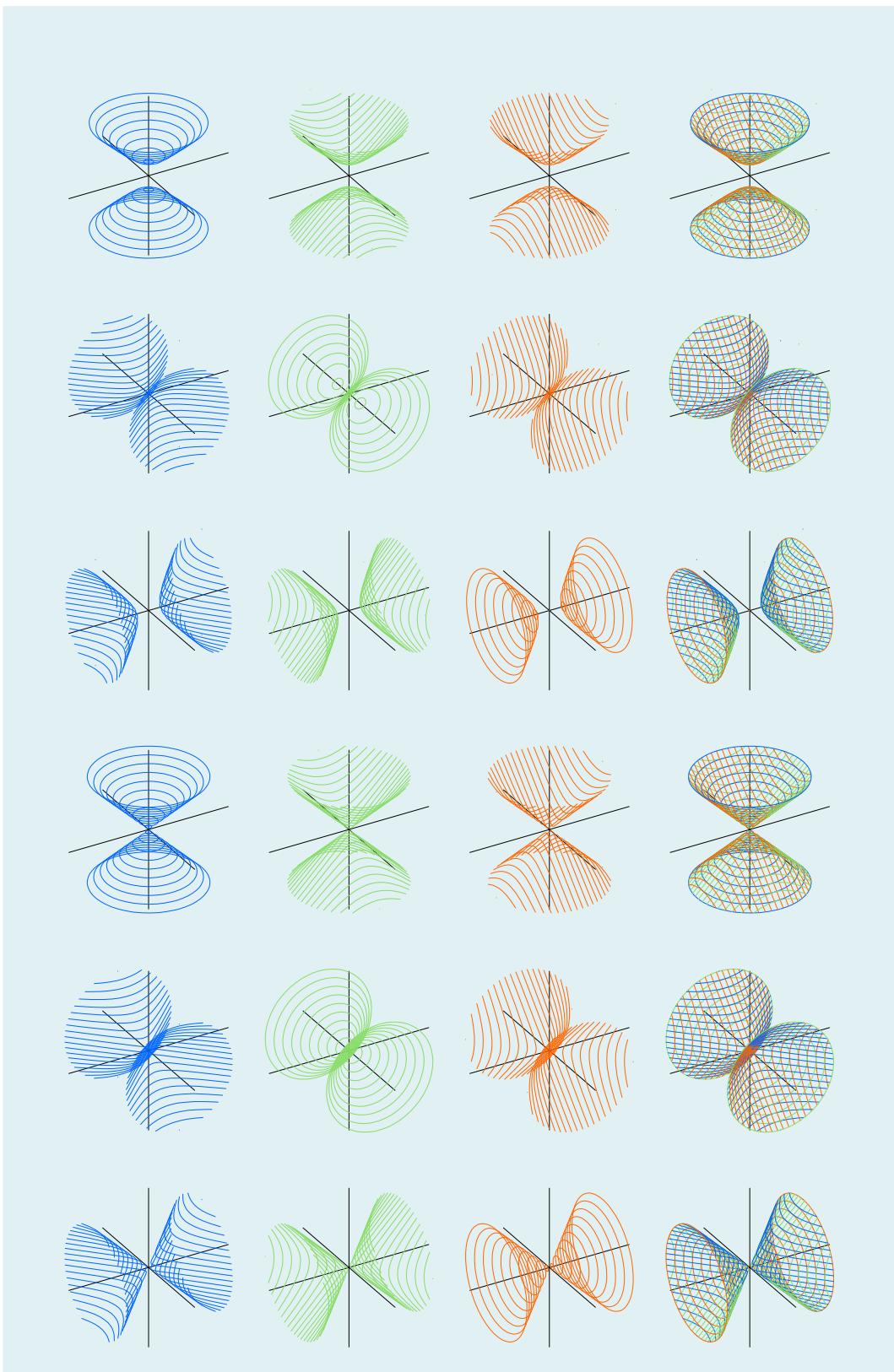


图 9. 用等高线可视化隐函数曲面，第 3 组图形

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

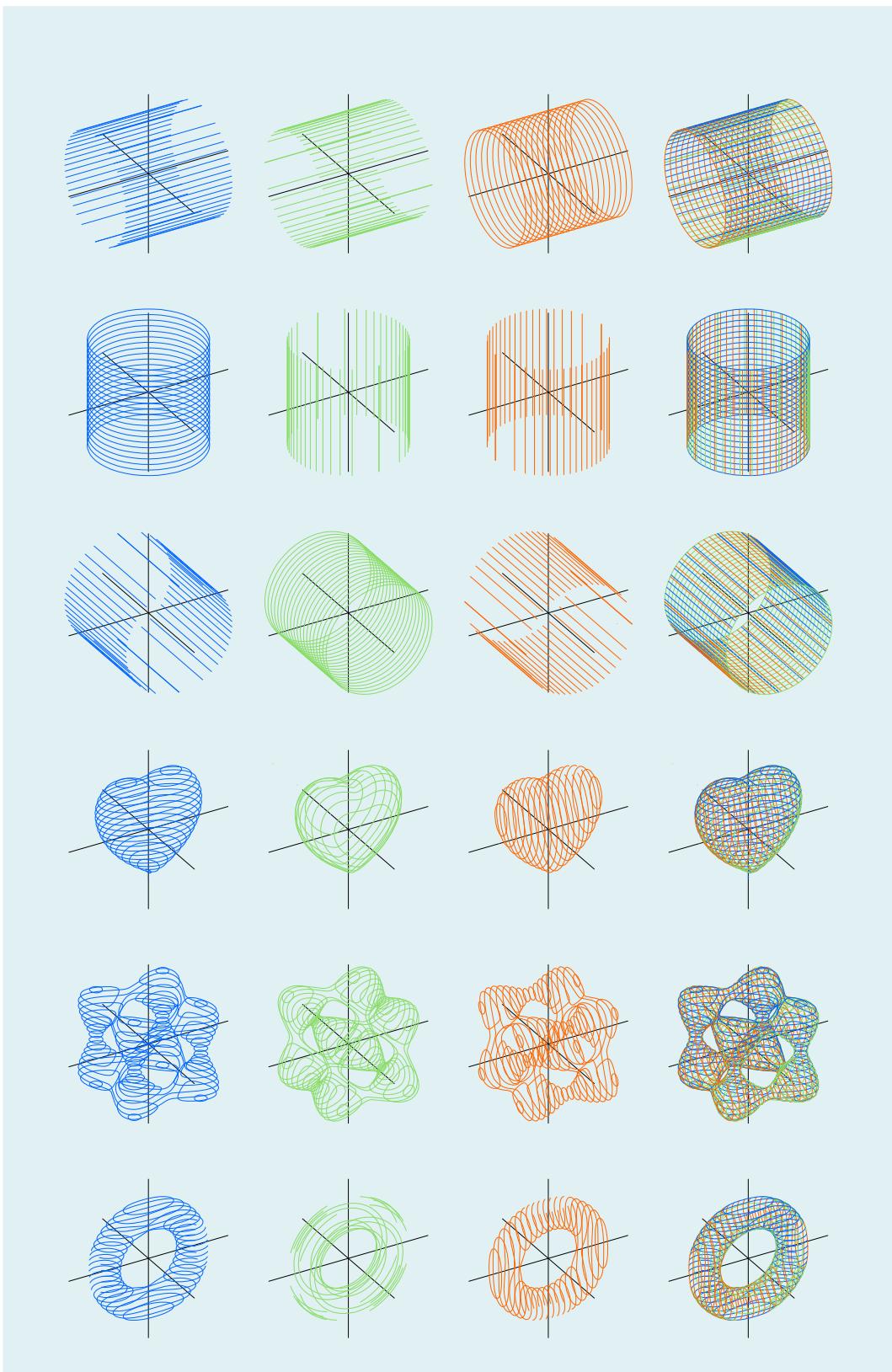


图 10. 用等高线可视化隐函数曲面，第 4 组图形

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

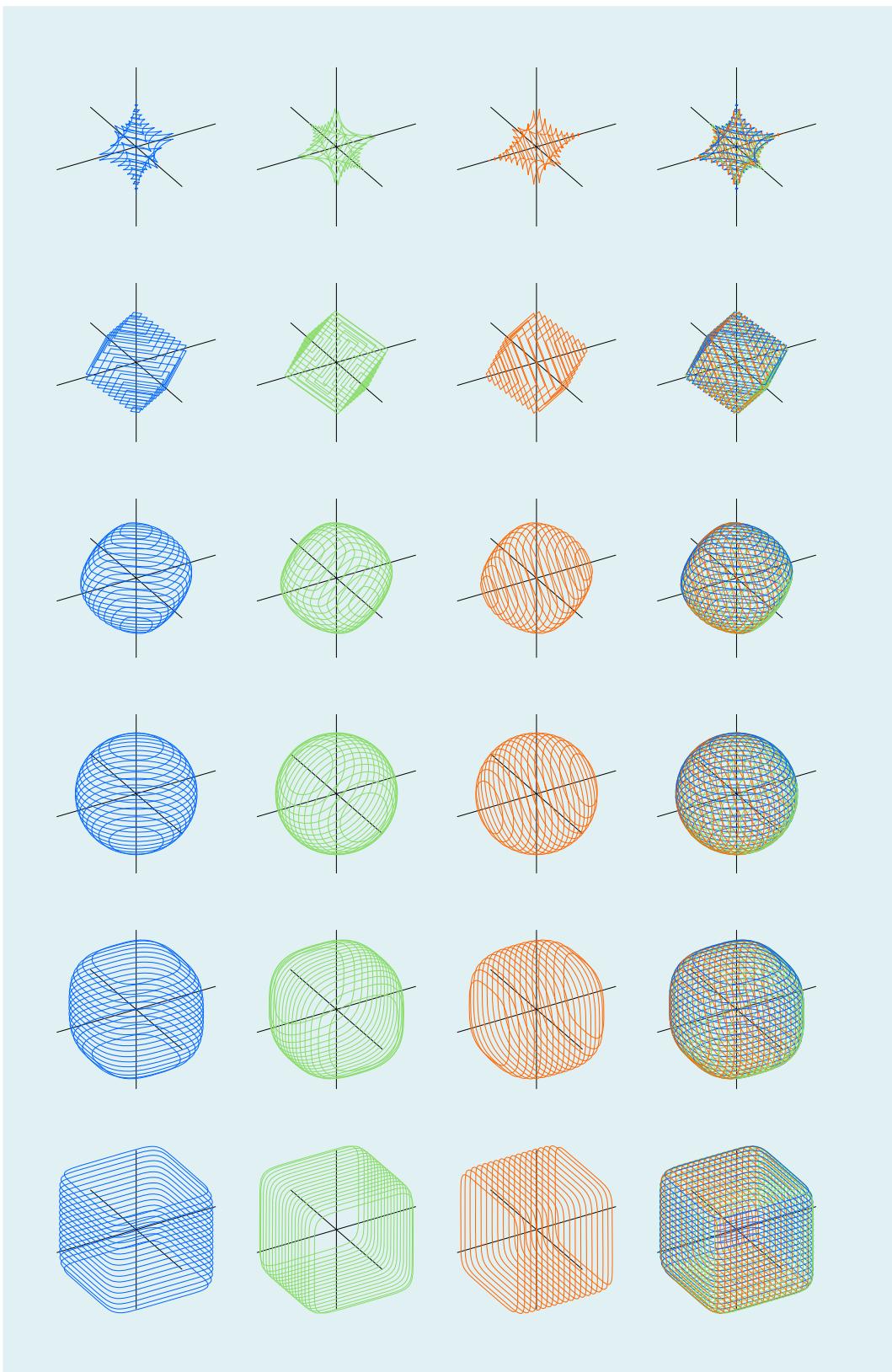


图 11. 用等高线可视化隐函数曲面，第 5 组图形

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

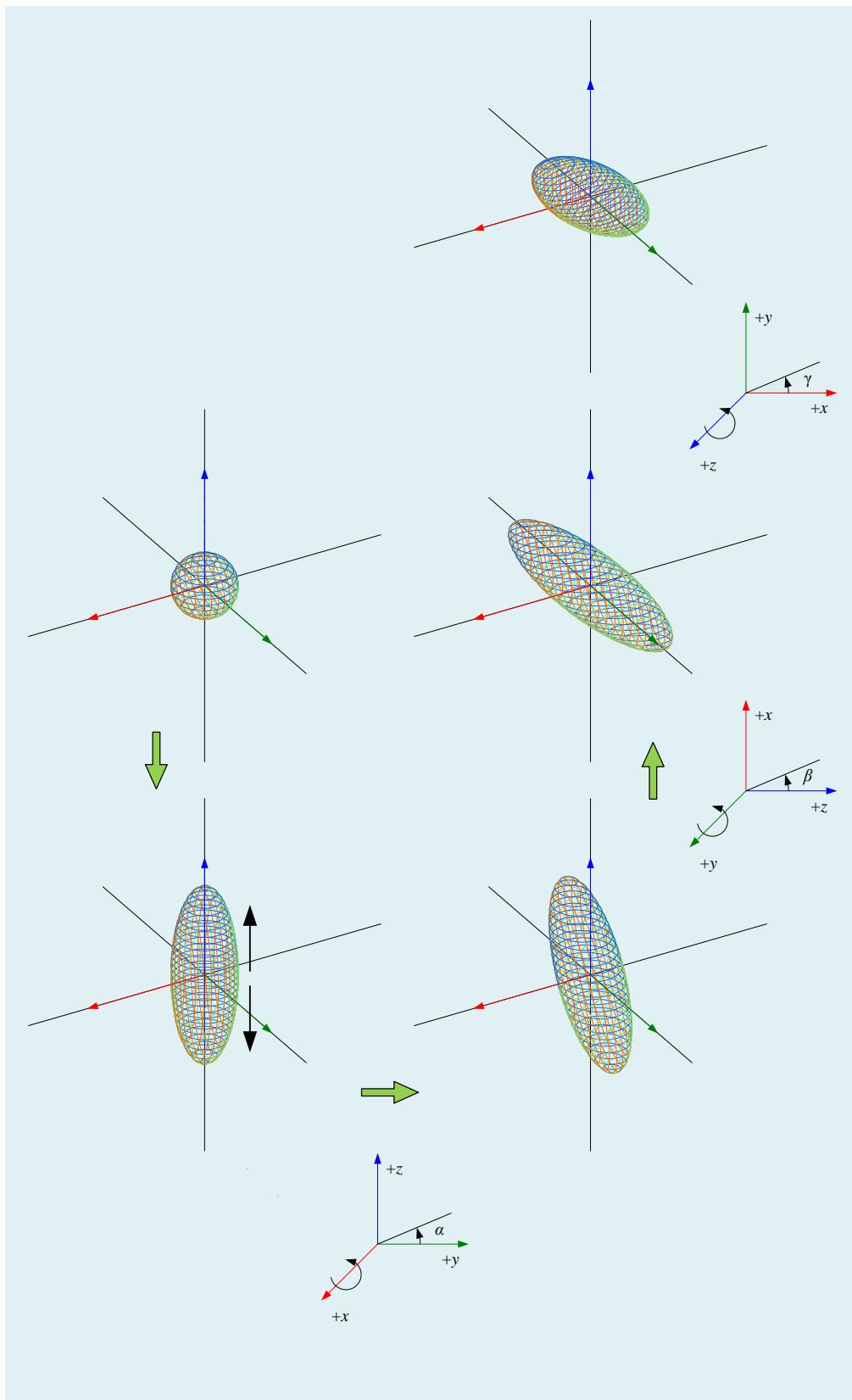


图 12. 从单位球到旋转椭球

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

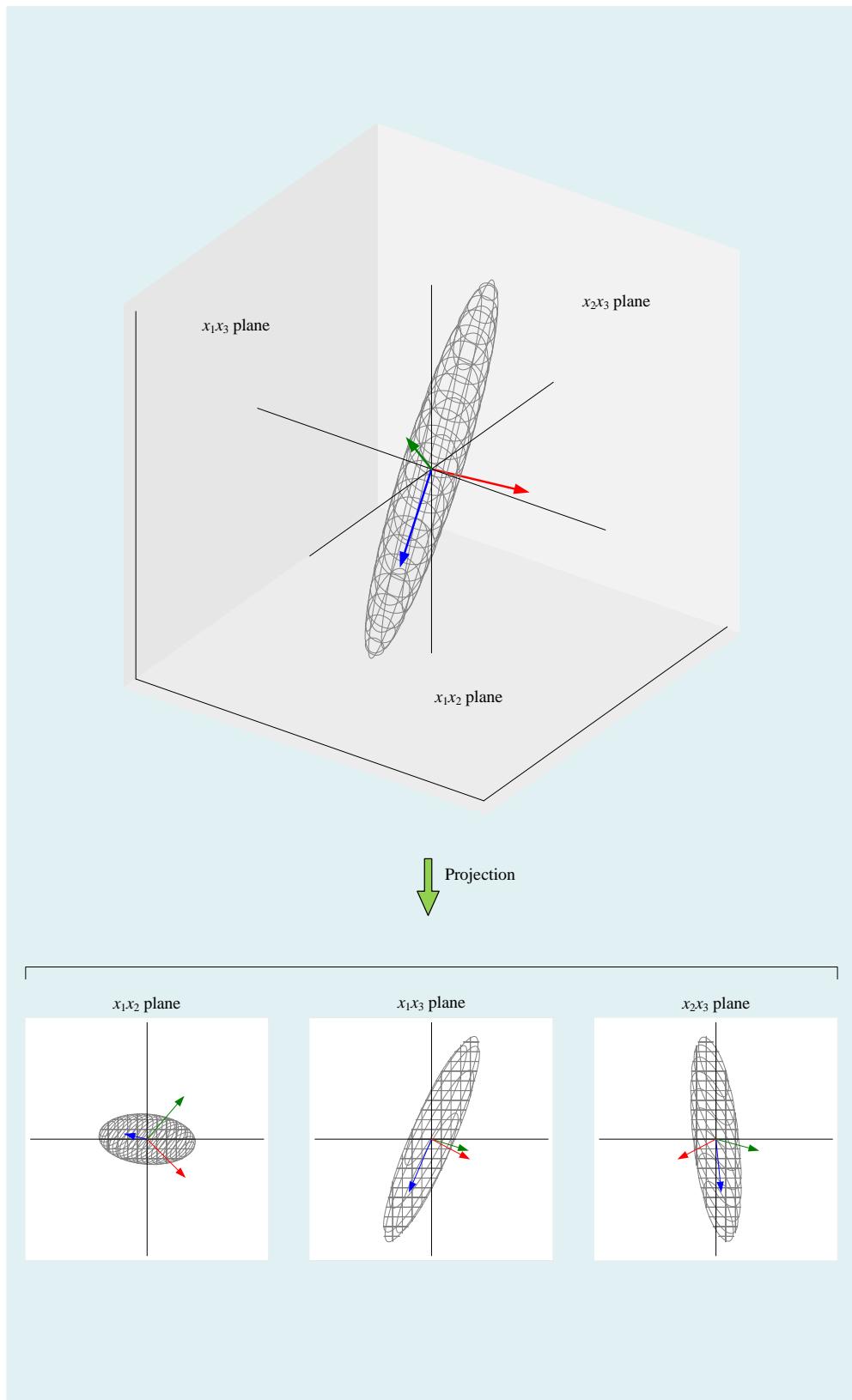


图 13. 旋转椭球在三个平面的投影

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

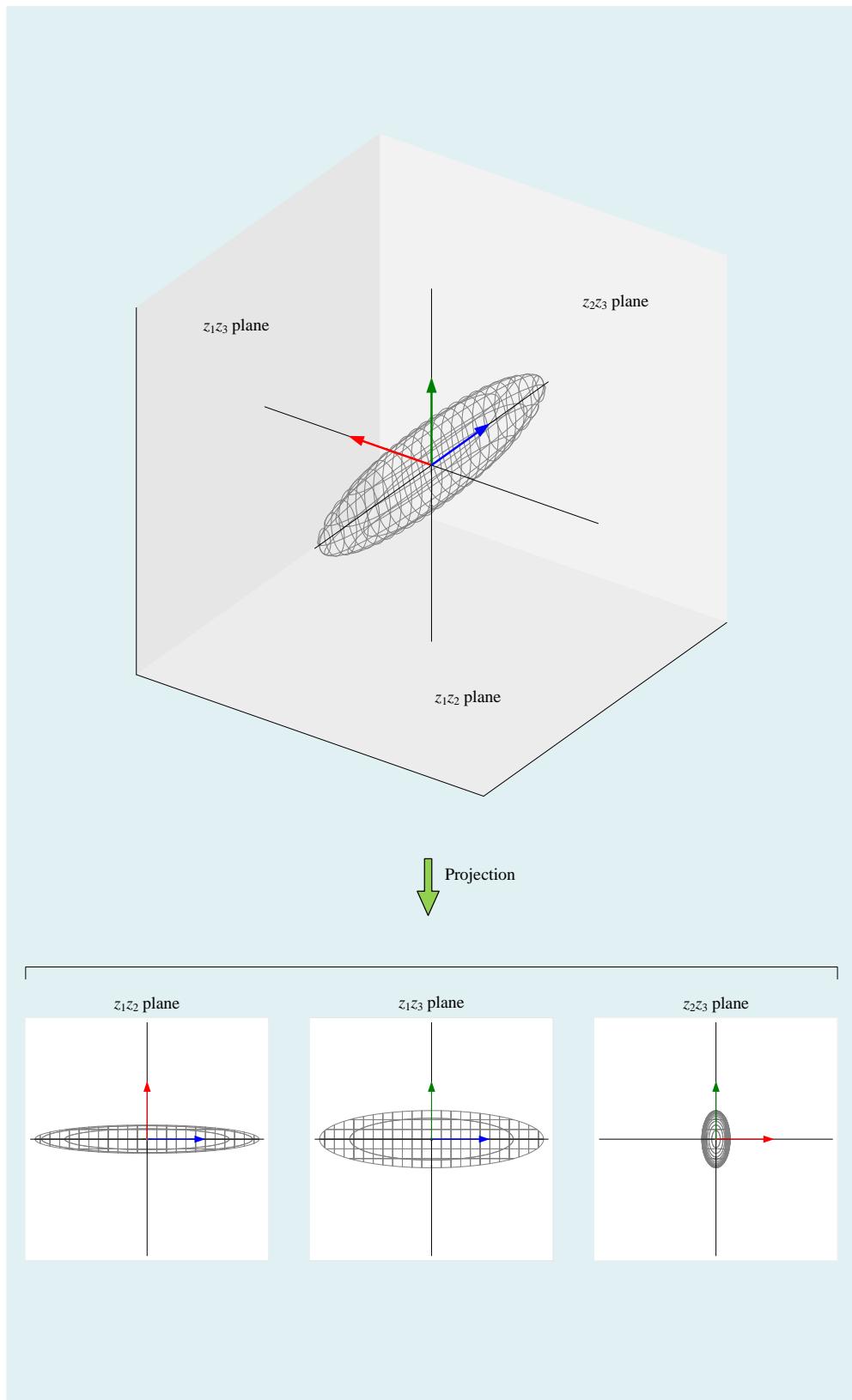


图 14. 正椭球在三个平面的投影

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com



Types of Distances

距离

两点连线、欧氏距离、闵氏距离、马氏距离



没有人会两次踏入同一条河流；江河川流不息，红尘物是人非。

No man ever steps in the same river twice. For it's not the same river and he's not the same man.

——赫拉克利特 (Heraclitus) | 古希腊哲学家 | 535 ~ 475 BC



- ◀ `numpy.diag()` 如果 A 为方阵, `numpy.diag(A)` 函数提取对角线元素, 以向量形式输入结果; 如果 a 为向量, `numpy.diag(a)` 函数将向量展开成方阵, 方阵对角线元素为 a 向量元素
- ◀ `numpy.linalg.inv()` 计算逆矩阵
- ◀ `numpy.linalg.norm()` 计算范数
- ◀ `scipy.spatial.distance.chebyshev()` 计算切比雪夫距离
- ◀ `scipy.spatial.distance.cityblock()` 计算城市街区距离
- ◀ `scipy.spatial.distance.euclidean()` 计算欧氏距离
- ◀ `scipy.spatial.distance.mahalanobis()` 计算马氏距离
- ◀ `scipy.spatial.distance.minkowski()` 计算闵氏距离
- ◀ `scipy.spatial.distance.seuclidean()` 计算标准化欧氏距离
- ◀ `sklearn.metrics.pairwise.euclidean_distances()` 计算成对欧氏距离矩阵
- ◀ `sklearn.metrics.pairwise_distances()` 计算成对距离矩阵

26.1 欧氏距离

欧氏距离，也叫欧几里得距离 (Euclidean Distance)，是最常见的距离度量方法，它计算两个点之间的直线距离。如图 1 所示，在一维数轴上，任意一点 x 到原点的距离就是 x 的绝对值 $|x|$ 。

对于平面直角坐标系，欧氏距离可以通过使用勾股定理来计算。平面上任意一点 (x_1, x_2) 和原点 $(0, 0)$ 的距离为 $\sqrt{x_1^2 + x_2^2}$ 。如图 1 所示，欧氏距离的等距线为一系列同心圆。如图 5 所示为平面上常见的两点连线的可视化方案。

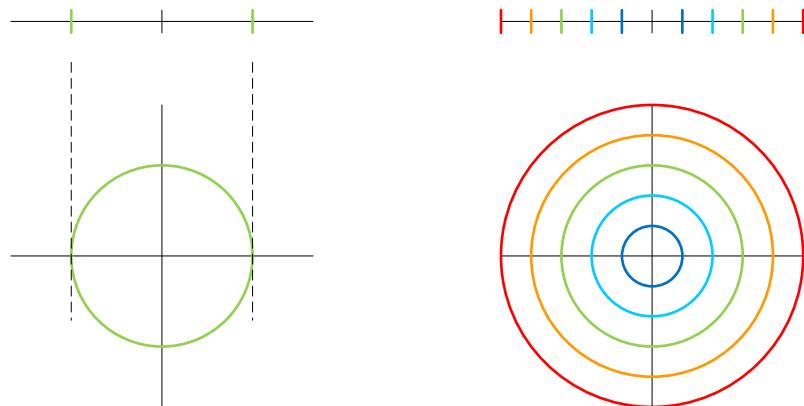


图 1. 数轴、平面直角坐标系上的欧氏距离等距线

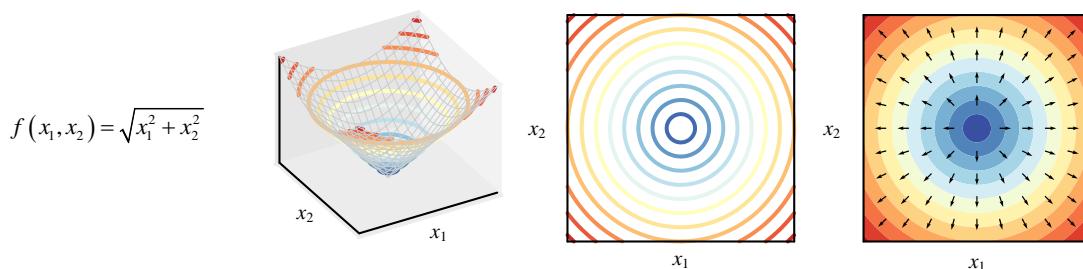


图 2. 平面上任意一点 (x_1, x_2) 和 $(0, 0)$ 之间的欧氏距离的几种可视化方案

如图 3 所示，在三维直角坐标系中，任意一点 (x_1, x_2, x_3) 和 $(0, 0, 0)$ 之间的欧氏距离为 $\sqrt{x_1^2 + x_2^2 + x_3^2}$ ，其等距线为正球体。

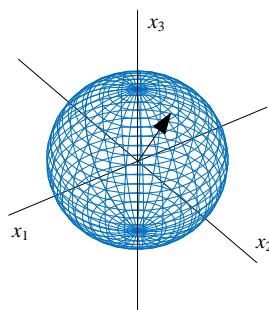


图 3. 三维直角坐标系上的欧氏距离等距线

26.2 其他距离

在机器学习中，距离不再仅仅是两点之间最短的线段。距离变成用于衡量两个对象或数据之间的相似性或差异性的概念。不同的距离度量方法可以基于不同的度量标准和算法来计算。以下是几种常见的距离度量方法。

如图 4 所示，在平面上，城市街区距离，也叫曼哈顿距离 (Manhattan Distance) 或 L^1 距离，通过沿着坐标轴的垂直和水平线段的长度之和来测量两个点之间。

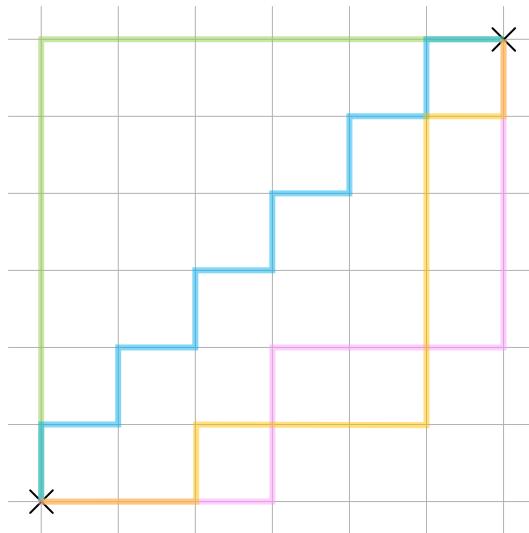


图 4. 城市街区距离

L^p 范数是向量的一种度量方式，用来衡量向量中各个元素的大小。在 L^p 范数中， p 是一个实数，并且 p 大于等于 1。

《矩阵力量》第 3 章将专门讲解向量 L^p 范数。

L^p 距离，也叫闵氏距离 (Minkowski Distance)，是使用 L^p 范数来度量两个向量之间的距离。当 $p = 1$ 时，得到的是城市街区距离；当 $p = 2$ 时，得到的是欧氏距离；当 p 趋近于无穷大时，得到的是切比雪夫距离 (Chebyshev distance)。

标准化欧氏距离 (Standardized Euclidean Distance) 是对欧氏距离进行标准化的一种方法。在计算标准化欧氏距离时，对每个维度的值进行标准化处理，然后再计算欧氏距离。

马氏距离 (Mahalanobis Distance) 是一种考虑特征之间相关性的距离度量方法。它使用协方差矩阵来衡量特征之间的相关性，从而在计算距离时考虑到了特征之间的相关性。

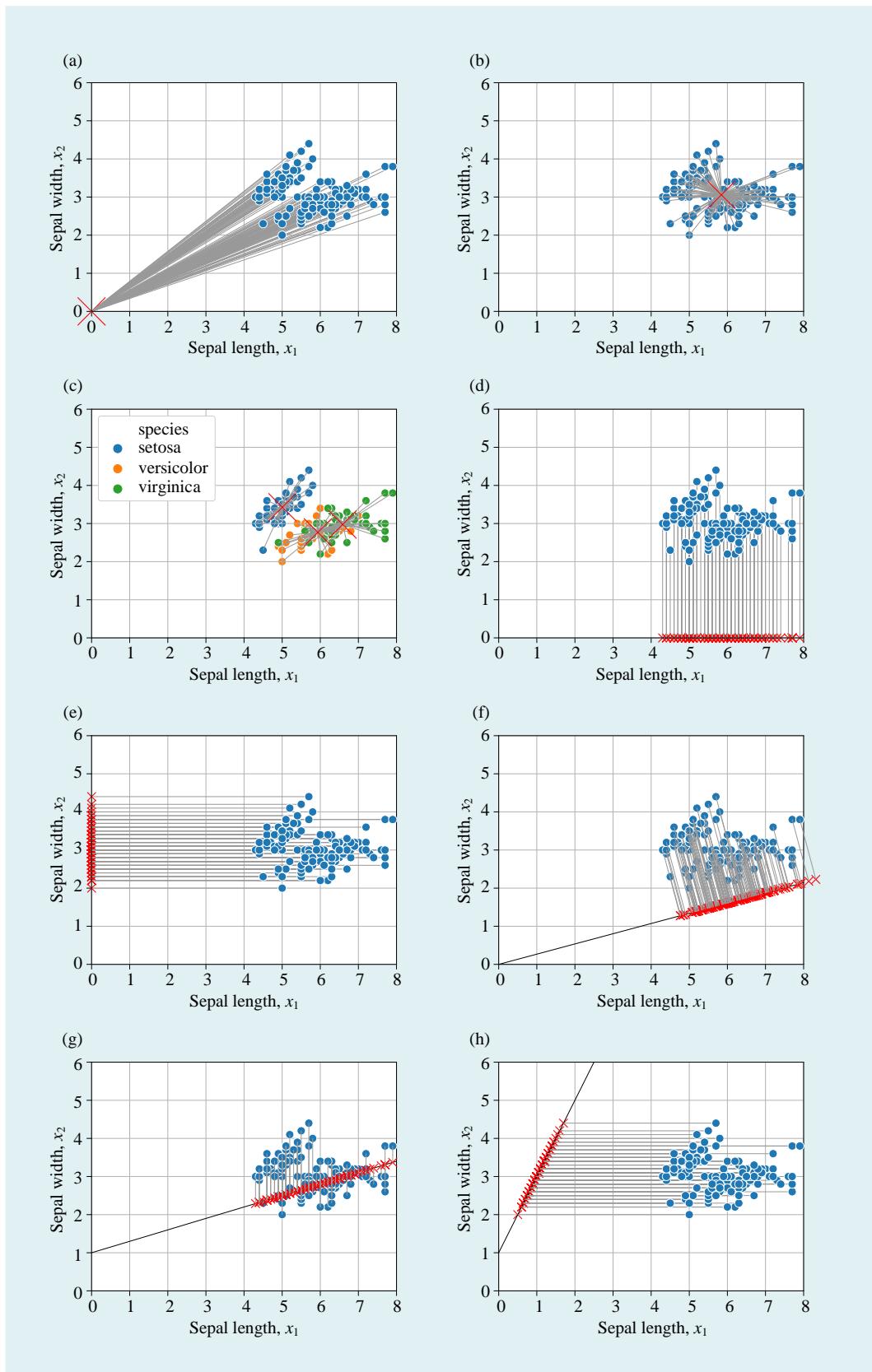


图 5. 平面两点距离的几种情况

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

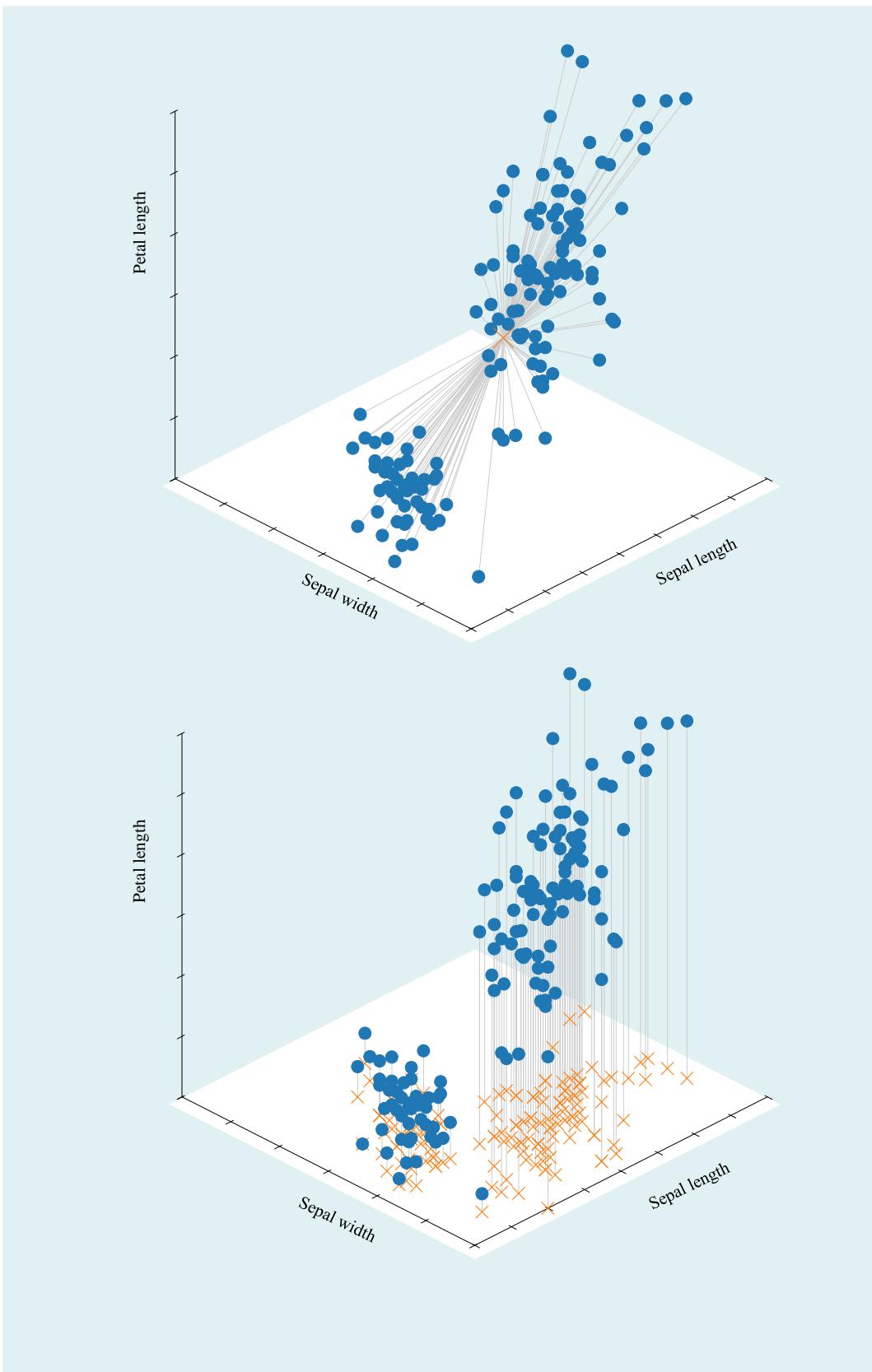
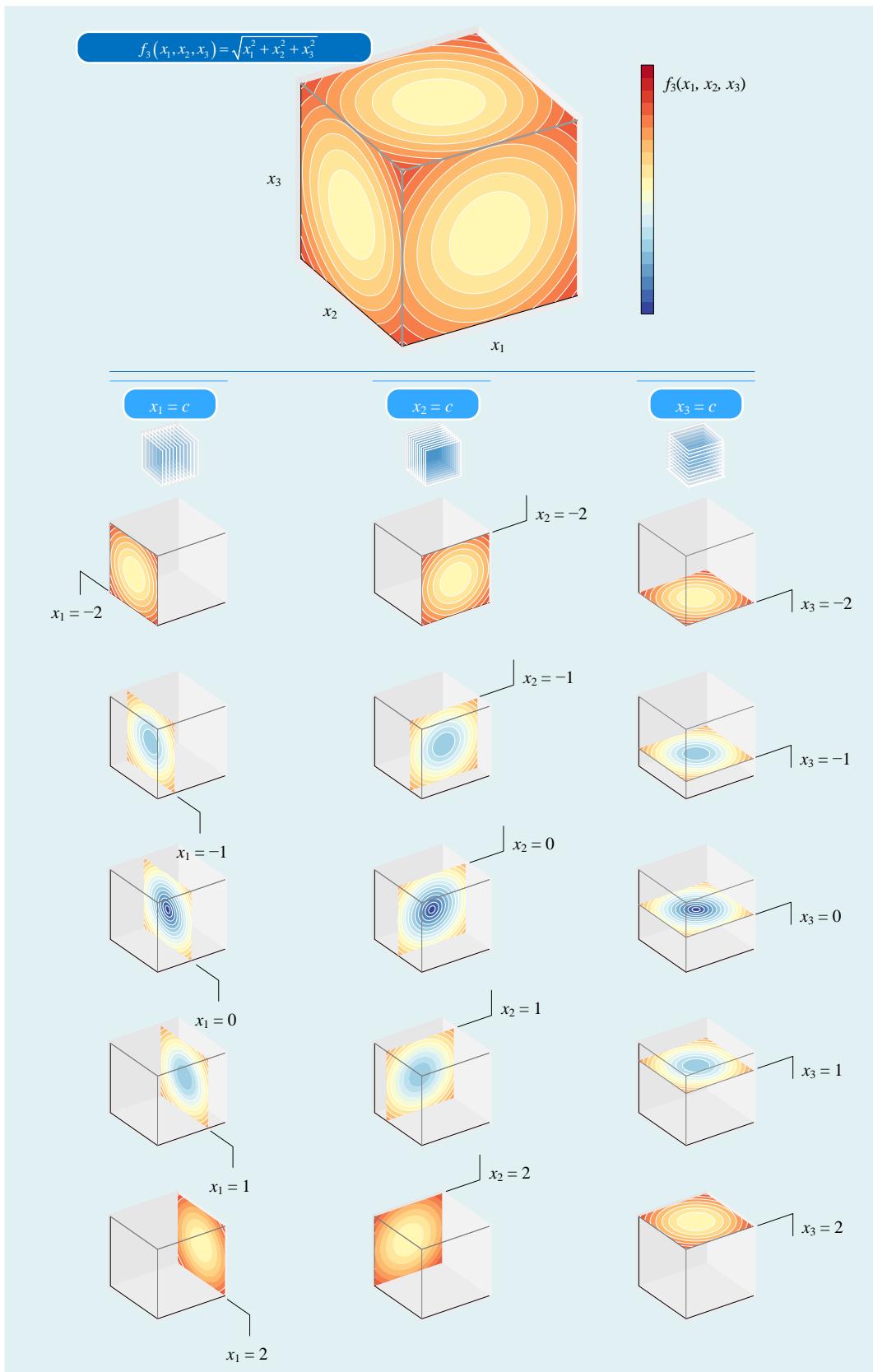


图 6. 三维空间距离

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

图 7. 三维直角坐标系中，任意一点 (x_1, x_2, x_3) 和 $(0, 0, 0)$ 之间的欧氏距离

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

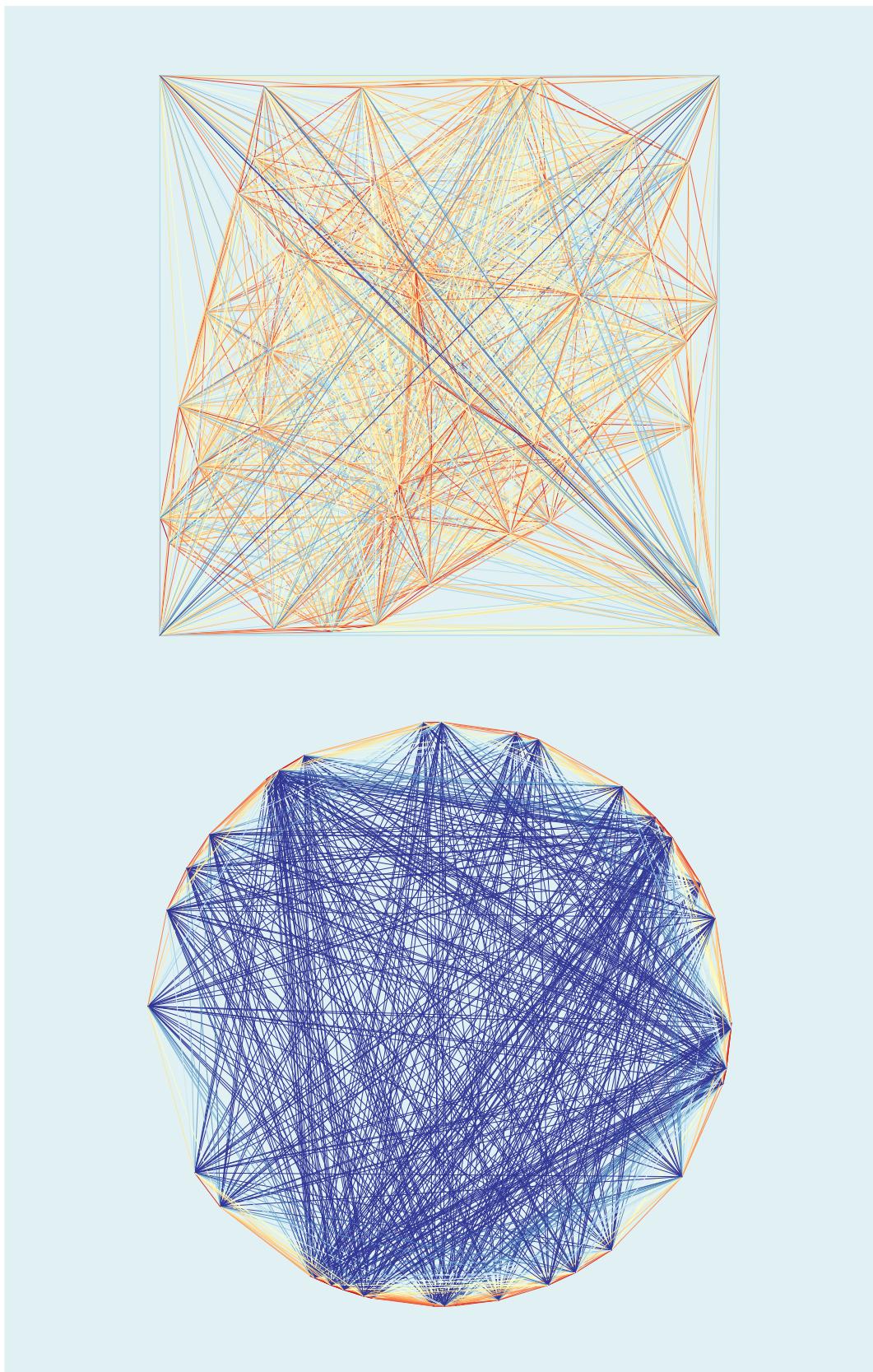


图 8. 根据欧氏距离远近渲染两点连线

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

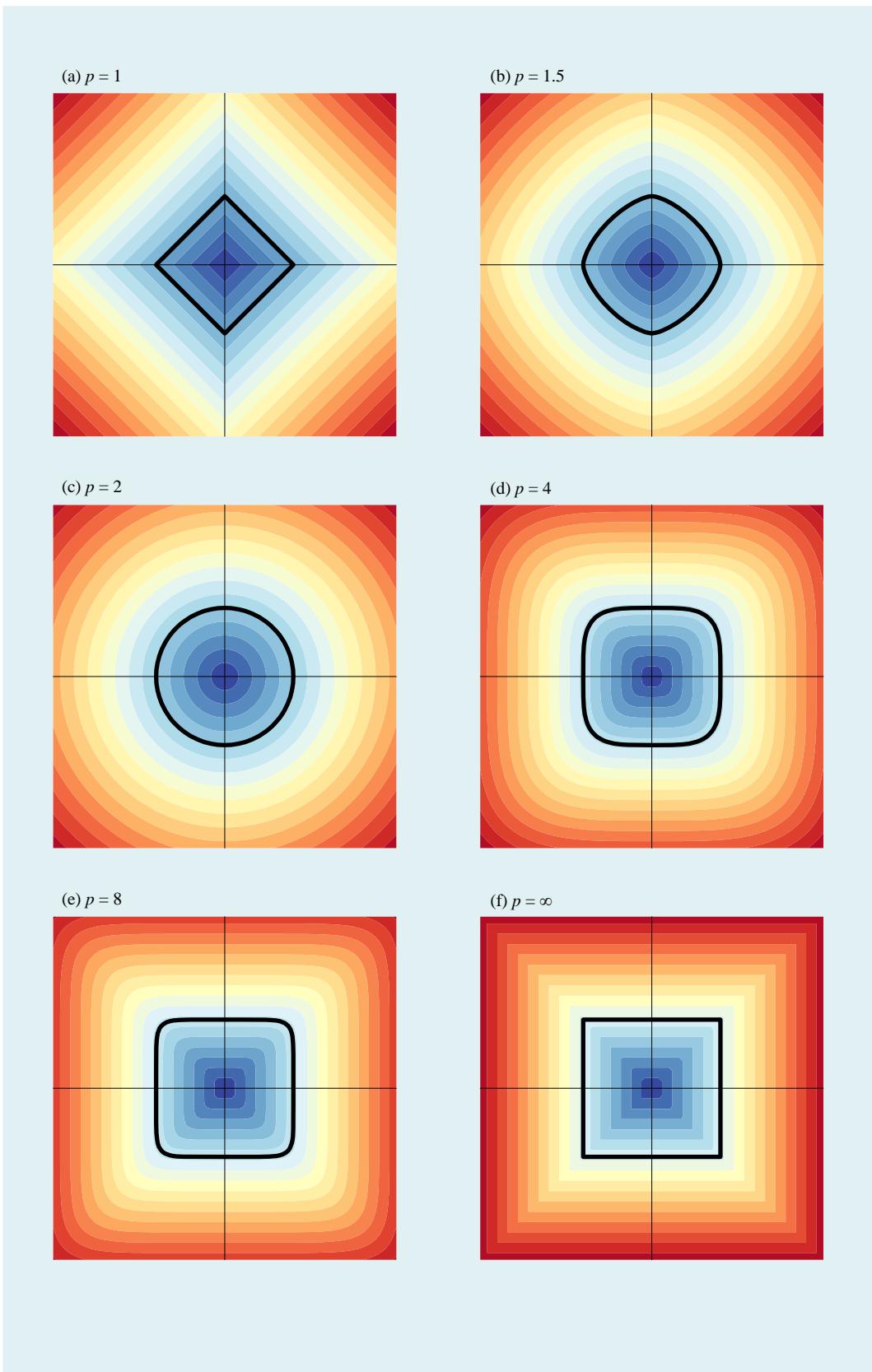


图 9. 向量范数，二维等高线

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

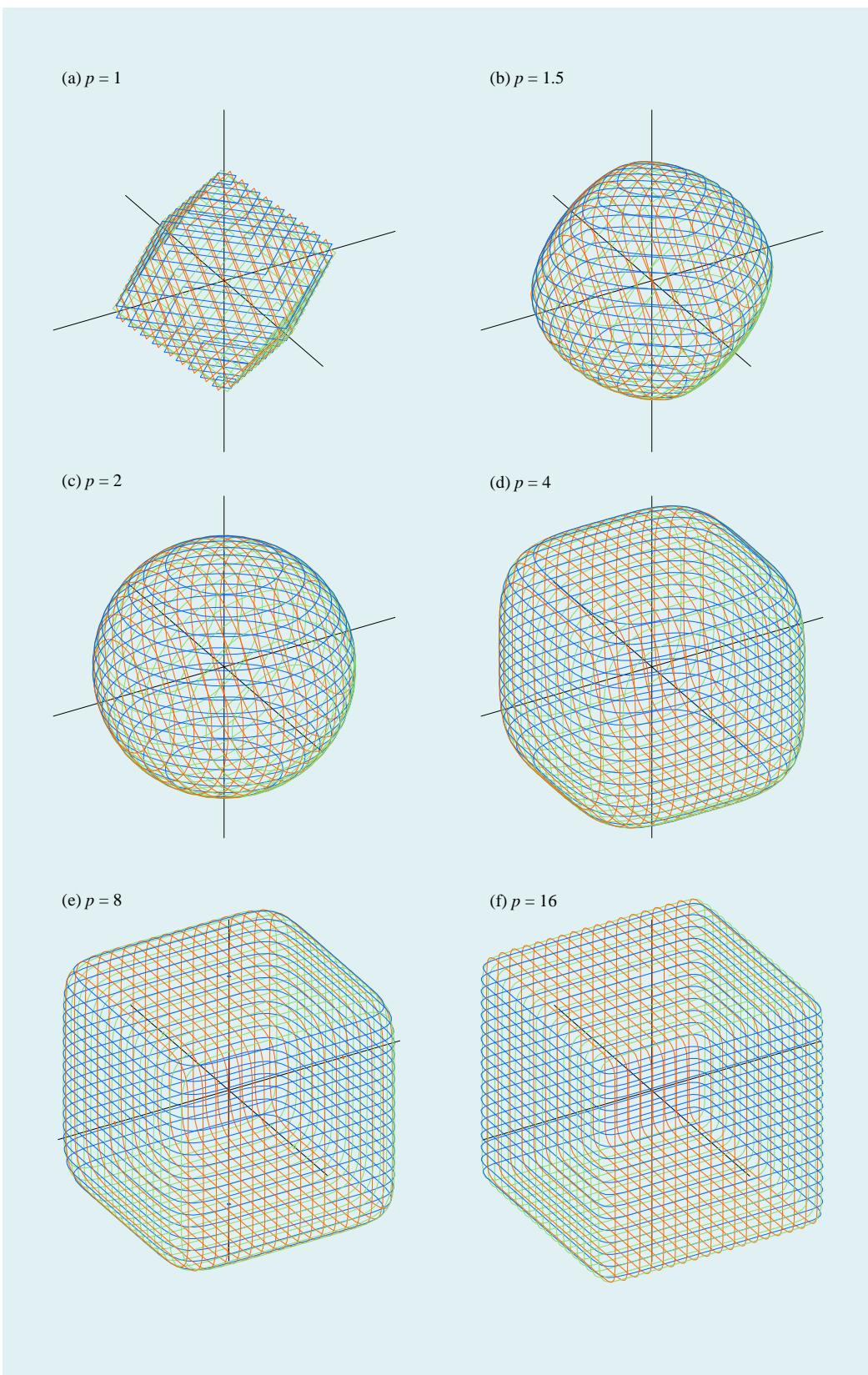


图 10. 向量范数，三维几何体

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

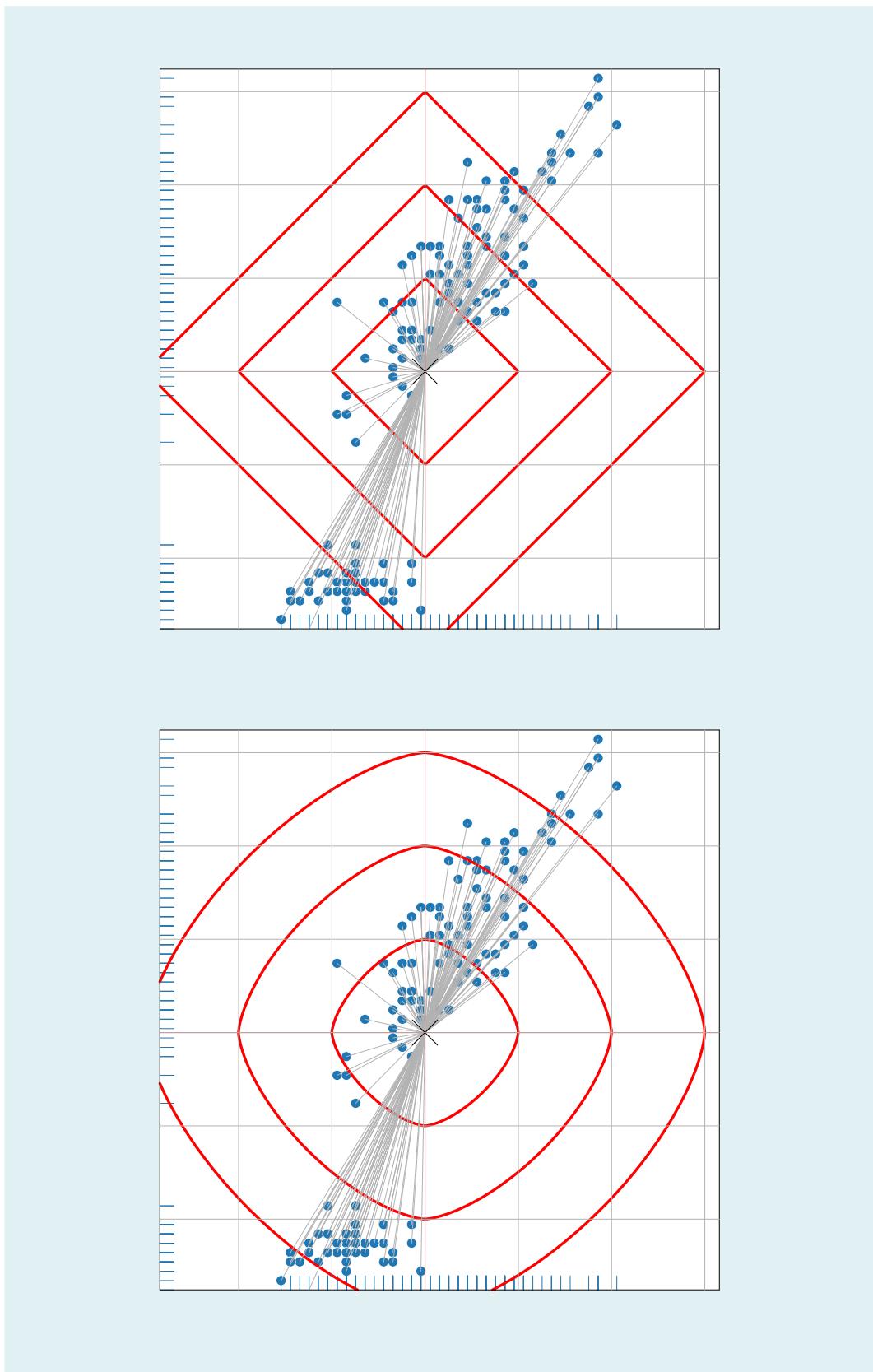


图 11. 鸢尾花数据，距离质心， L^1 、 $L^{1.5}$ 距离

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

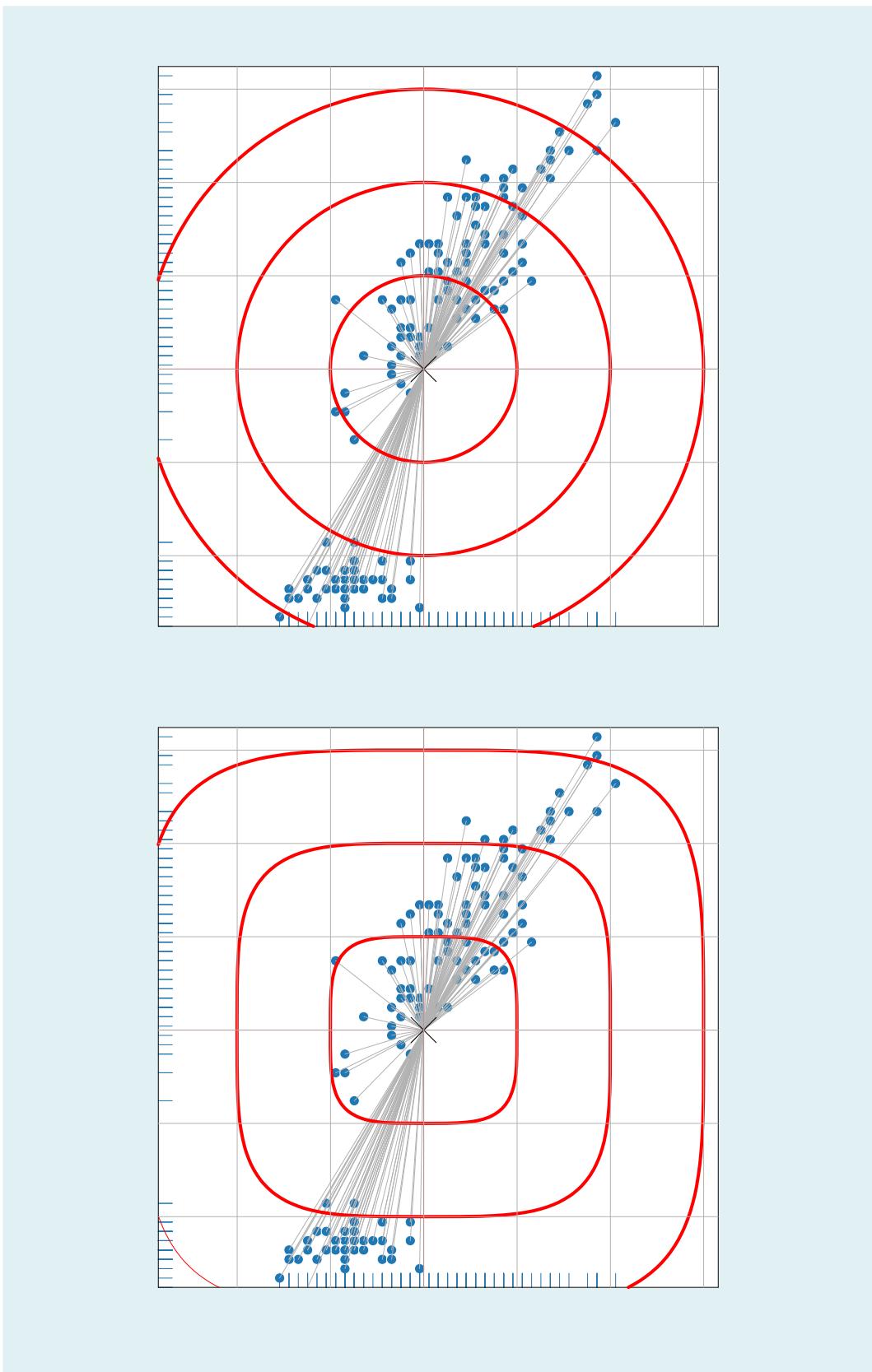


图 12. 鸢尾花数据，距离质心，L2、L4 距离

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

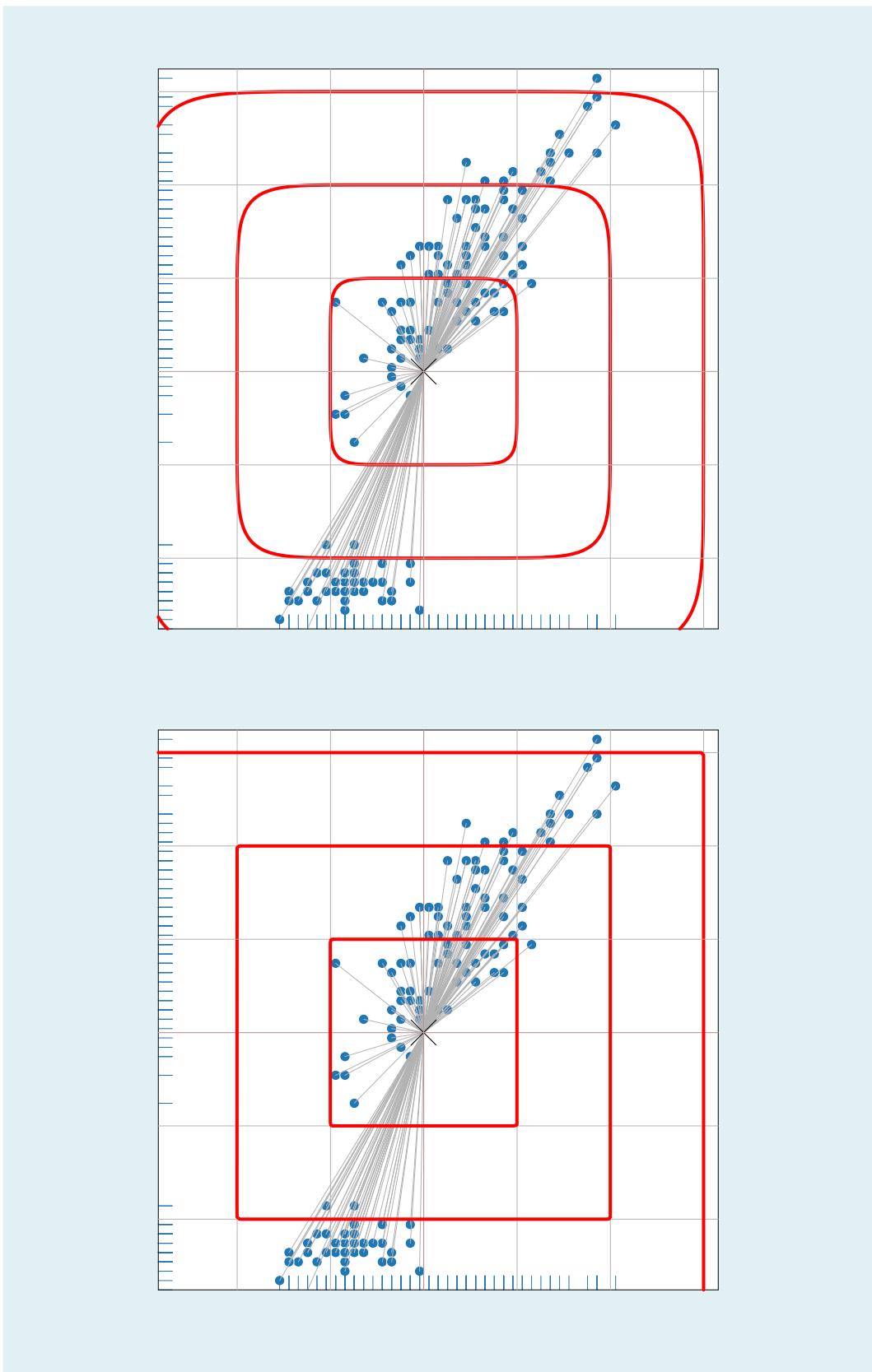


图 13. 鸢尾花数据，距离质心， L^2 、 L^∞ 距离

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

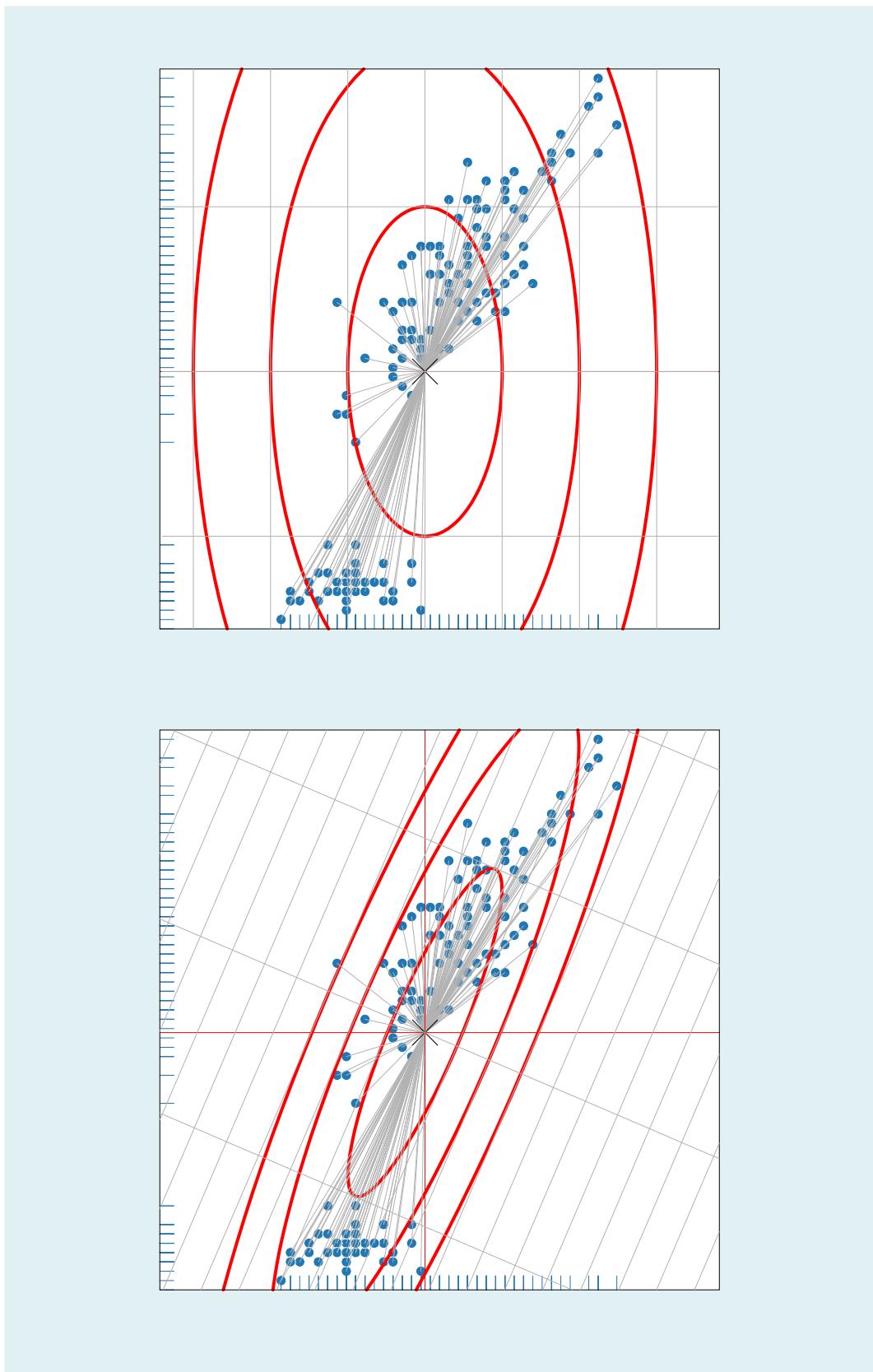


图 14. 鸢尾花数据，距离质心，标准化欧氏距离、马氏距离

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

27

Geometric Transformations on a Plane

平面几何变换

平移、缩放、旋转、镜像、投影、剪切



哲学就是怀着一种乡愁的冲动到处去寻找家园。

Philosophy is properly home-sickness; the wish to be everywhere at home.

—— 诺瓦利斯 (Novalis) | 德国作家 | 1772 ~ 1801



- ◀ matplotlib.pyplot.axhline() 绘制水平线
- ◀ matplotlib.pyplot.axvline() 绘制竖直线
- ◀ matplotlib.pyplot.scatter() 绘制散点图
- ◀ matplotlib.pyplot.scatter() 绘制散点图
- ◀ numpy.arange() 根据指定的范围以及设定的步长，生成一个等差数组
- ◀ numpy.c_() 用于按列连接数组的函数，用于快速组合多个数组成为一个新的二维数组
- ◀ numpy.column_stack() 将两个矩阵按列合并
- ◀ numpy.exp() 计算括号中元素的自然指数
- ◀ numpy.linalg.cholesky() 矩阵 Cholesky 分解
- ◀ numpy.meshgrid() 创建网格化数据
- ◀ numpy.ravel() 用于将多维数组展平为一维数组
- ◀ numpy.reshape() 用于重新调整数组的形状
- ◀ numpy.sqrt() 计算平方根
- ◀ seaborn.load_dataset() 加载 Seaborn 中数据集
- ◀ sklearn.covariance.EmpiricalCovariance() 估算协方差矩阵
- ◀ sklearn.datasets.load_iris 加载 Scikit-Learn 中鸢尾花数据

27.1 常见几何变换

图 1 所示为常见的几何变换。图 1 (a) 所示为原始棋盘黑白格子。

图 1 (b) 为仿射变换 (affine transformation) 的结果。仿射变换是本节的重点。

图 1 (c) 投射变换 (projective transformation) 是一种将三维空间中的点映射到二维平面上的变换，也被称为透视变换 (perspective transformation)。

图 1 (d) 所示为非线性变换的一种形式。

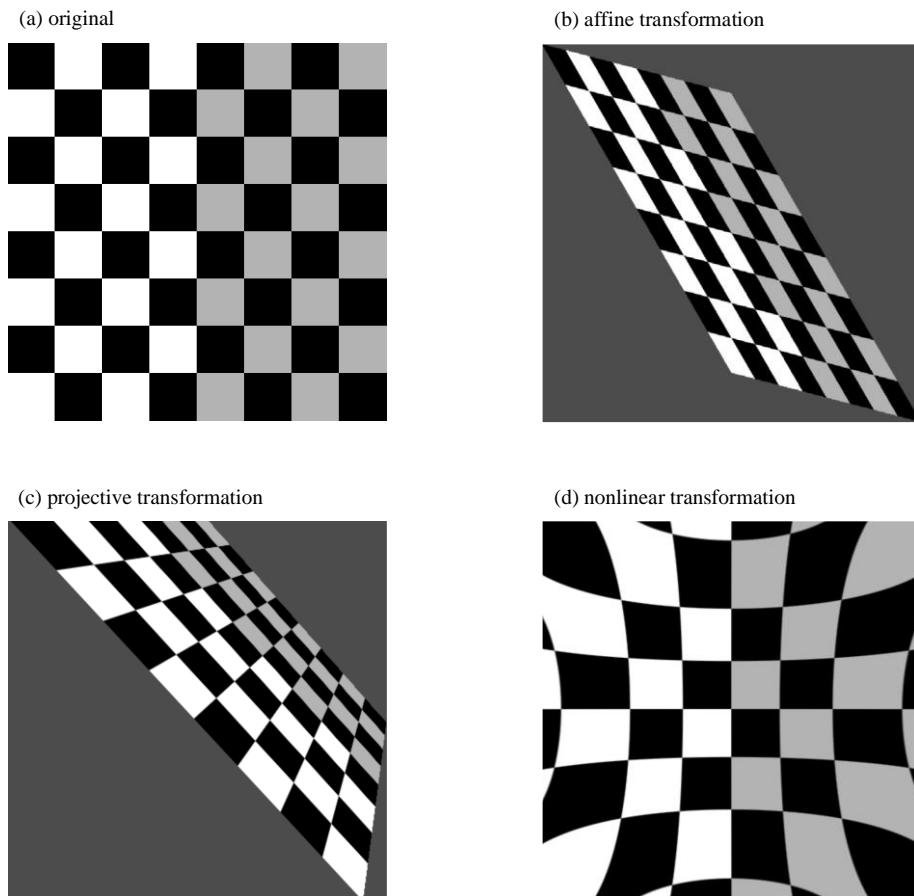


图 1. 常见几何变换

27.2 仿射变换

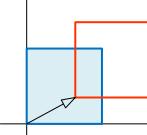
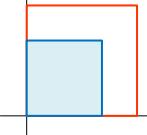
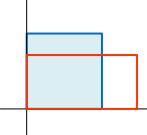
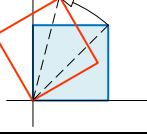
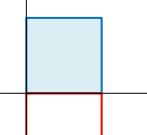
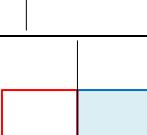
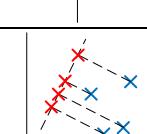
仿射变换是一种线性变换，它可以将一个二维或三维空间中的点映射到另一个二维或三维空间中的点。

在仿射变换中，原始对象的形状、大小、角度和比例等属性可能会发生变化，但它们之间的相对位置和平行关系将保持不变。

仿射变换可以应用于许多不同的领域，如计算机视觉、计算机图形学、机器学习等。在计算机视觉中，仿射变换可以用于图像的平移、缩放、旋转、对称、投影、剪切等操作。[表 1](#) 总结常见几种仿射变换。

在计算机图形学中，仿射变换可以用于三维图形的变换和投影等操作。在机器学习中，仿射变换可以用于特征提取、数据增强和数据对齐等操作。

[表 1. 常见仿射变换](#)

几何变换	
平移 (translation)	
等比例缩放 s 倍 (scaling)	
非等比例缩放 (unequal scaling)	
逆时针旋转 (counterclockwise rotation)	
关于横轴镜像对称 (reflection along x-axis)	
关于纵轴镜像对称 (reflection along y-axis)	
向通过原点直线投影	

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

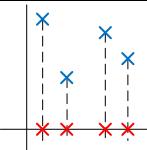
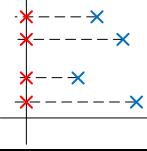
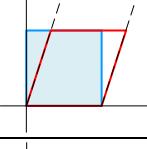
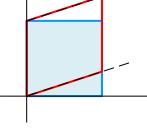
向横轴投影	
向纵轴投影	
沿水平方向剪切 (shear along x-axis)	
沿竖直方向剪切 (shear along y-axis)	

图 2 所示为原始网格散点。图 9 所示为各种仿射变换及其组合的结果。请大家注意，多数情况仿射变换先后影响结果。

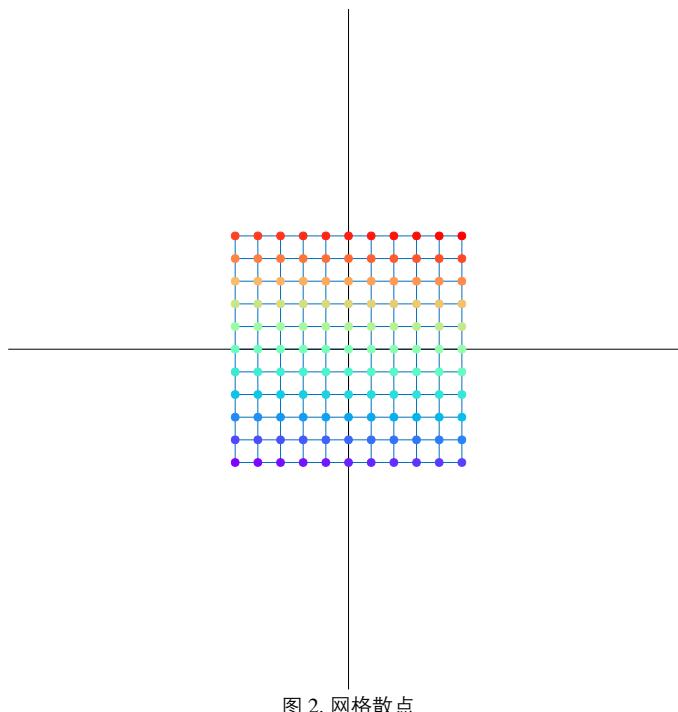
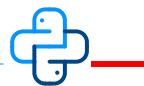


图 2. 网格散点



Jupyter 笔记 BK_2_Ch27_1.ipynb 绘制图 9 所有子图。

图 3 给出第二个例子。我们在单位圆上不同位置用不同颜色标记位置，管它们叫“小彩灯”。然后，对这些“小彩灯”先旋转，再剪切。同时，相同颜色的“小彩灯”之间再绘制一条线段，用来标识运动轨迹。

图 10 所示当旋转角度不同时，经过“旋转 → 剪切”变化的小彩灯都在同一个椭圆上。但是显然，每幅子图的小彩灯位置不同。此外，我们也在这些图上利用大小两个正方形来可视化旋转。

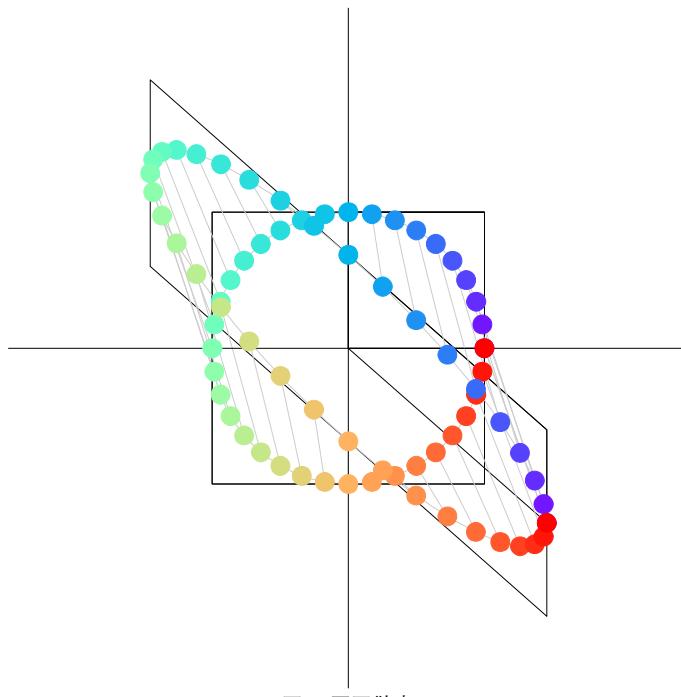
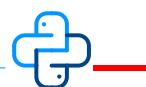


图 3. 正圆散点



Jupyter 笔记 BK_2_Ch27_2.ipynb 绘制图 10 所有子图。

《矩阵力量》将专门讲解仿射变换中用到的一系列数学工具。《统计至简》将会利用图 10 解析蒙特卡罗模拟产生的具有一定相关性的随机数。

27.3 投影

如一般情况，投影是指将一个三维物体，比如图 4 所示的马克杯，的映射到一个平面上的过程。在投影中，平面通常称为投影面，被投影的物体通常称为投影体。图 5 所示为马克杯在不同平面上的投影。

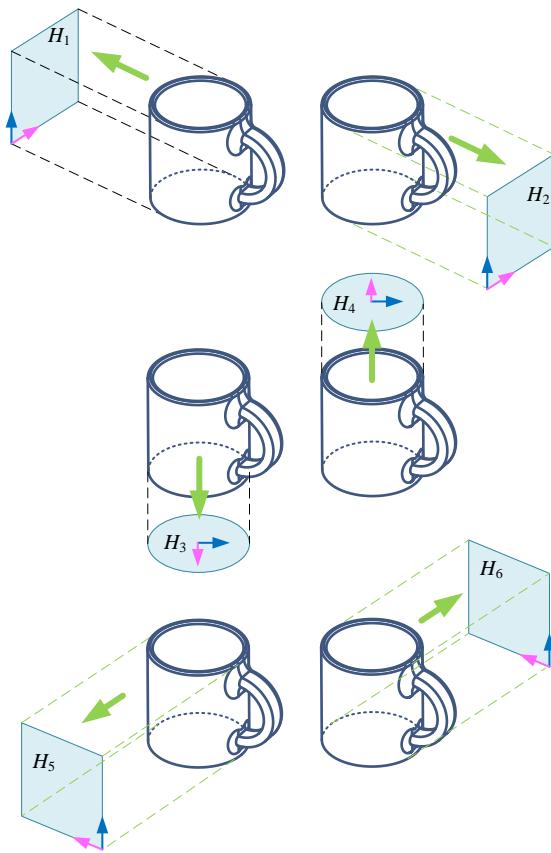
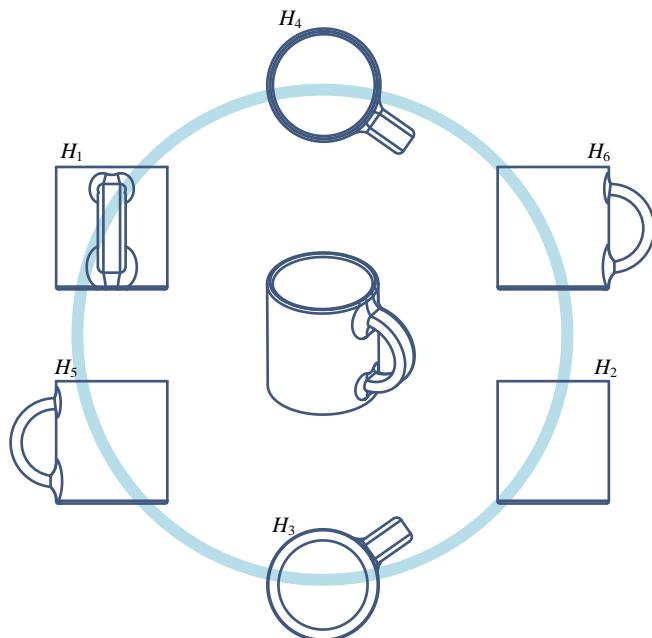


图 4. 马克杯六个投影方向



本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

图 5. 马克杯在六个方向投影图像

从数据角度来看，投影相当于一种数据降维。投影得到的像虽然可以“管窥一豹”，但是信息毕竟发生了“降维”压缩。仅此仅仅透过某一个角度的投影的像不能完全获得投影体的全部原始细节。

投影有不同的类型，其中最常见的包括正交投影 (orthogonal projection) 和透视投影 (perspective projection)。

在正交投影中，物体被投影到一个平行于投影面的平面上，从而保留了物体的真实形状和大小。而在透视投影中，物体被投影到一个与投影面不平行的平面上，从而产生了一种远近透视的效果，使得远离投影面的物体部分看起来比靠近投影面的物体部分更小。

投影在许多领域中都有应用，如建筑设计、计算机图形学、视觉艺术等。在建筑设计中，投影通常用于绘制建筑平面图、立面图和剖面图。在计算机图形学中，投影用于创建三维场景的二维表示，以及在计算机游戏和虚拟现实中实现视觉效果。在视觉艺术中，投影可以用于创造一种深度感或透视效果，以增强画面的艺术效果。

本书前文提到，matplotlib 在绘制三维图形时，默认透视投影。本书建议科学技术作品中静态图形最好使用正交投影。“鸢尾花书”系列作品中三维图形大多采用正交投影。

平面投影

图 6 所示为平面上的散点、曲线投影到横轴的结果。“鸢尾花书”中用这幅图正交投影、马氏距离、数据分布、数据投影等数学概念。大家将会在《矩阵力量》、《统计至简》两册看到相关的数学工具。图 11 展示更多不同角度的平面点线投影。

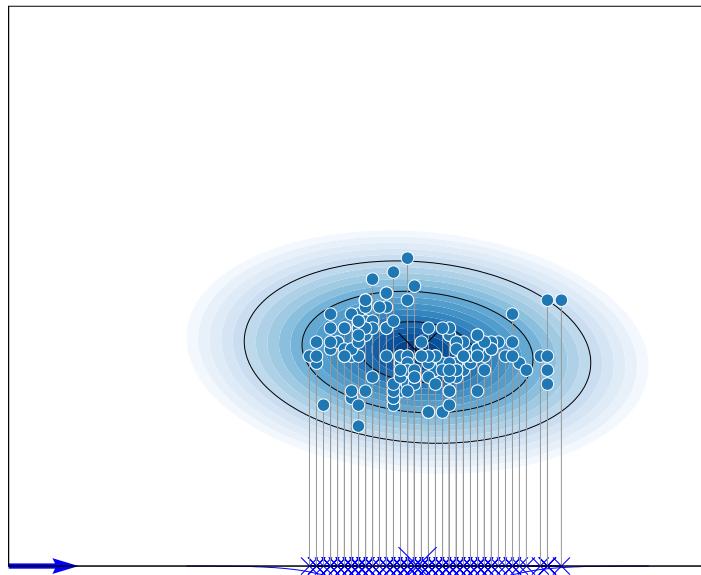
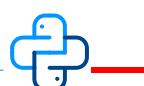


图 6. 平面投影



Jupyter 笔记 BK_2_Ch27_3.ipynb 绘制图 11 所有子图。

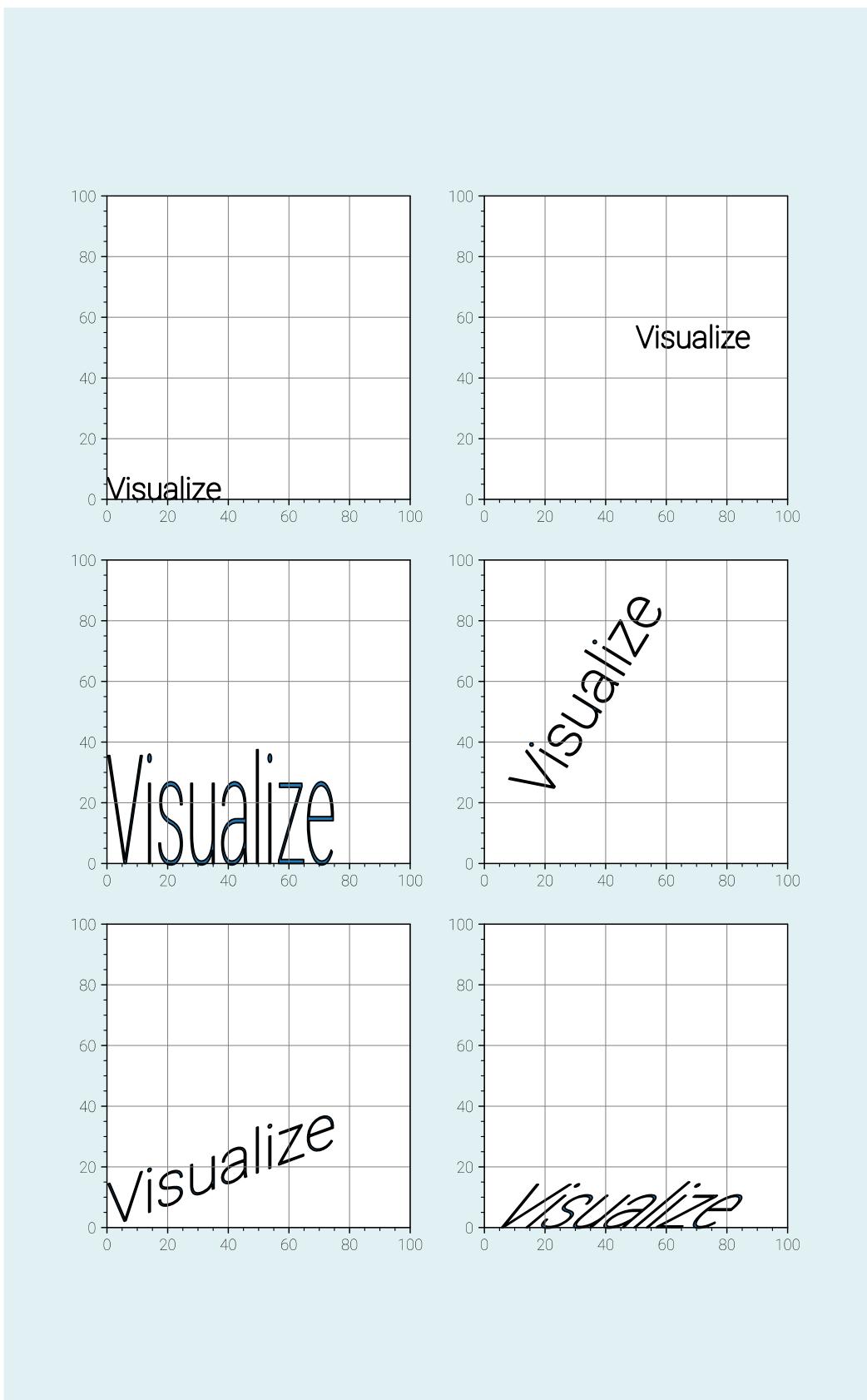


图 7. 扁平化文字对象的仿射变换

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

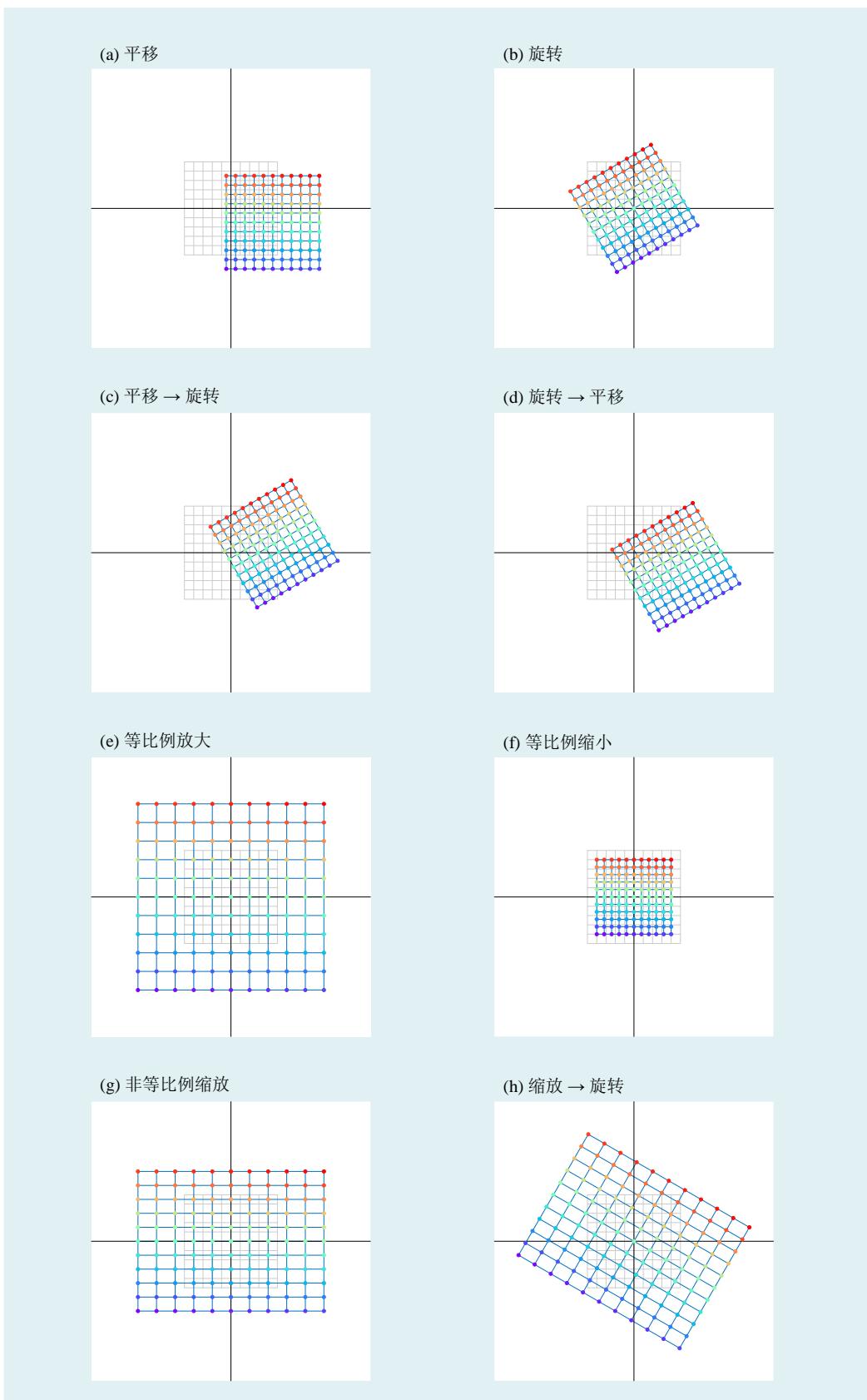


图 8. 仿射变换，第 1 组

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

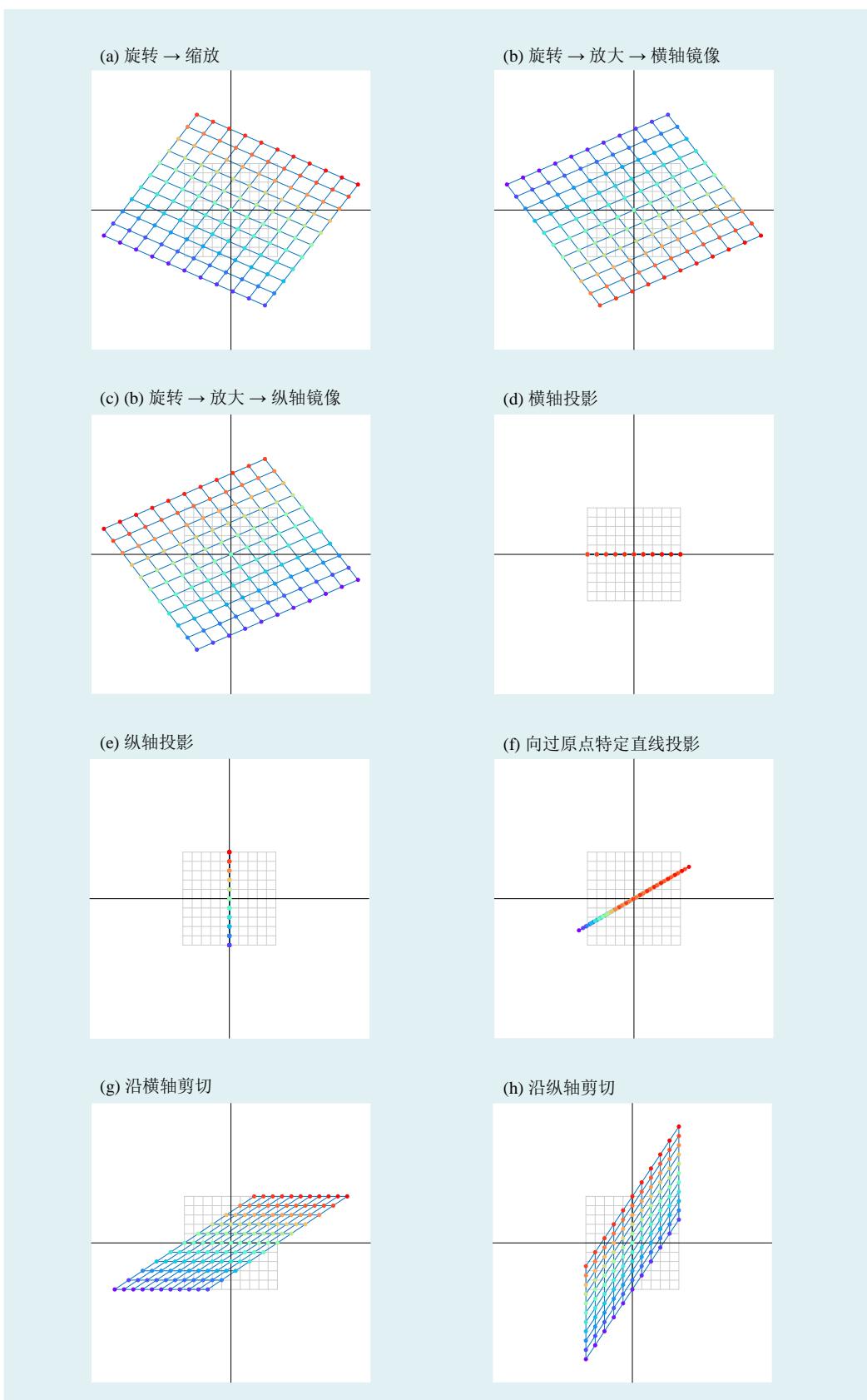


图 9. 仿射变换，第 2 组

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

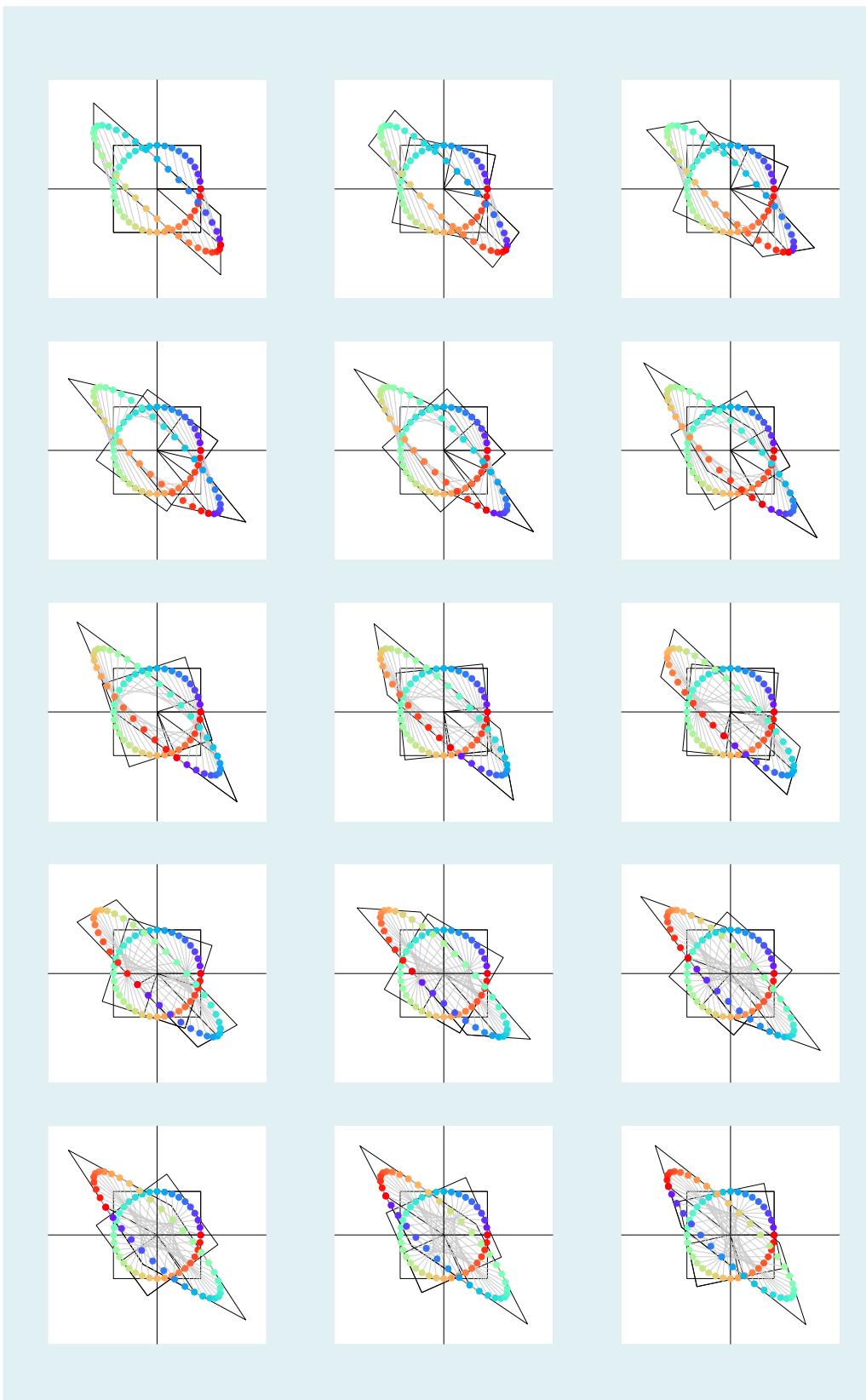


图 10. 正方形、单位圆剪切、旋转变换

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

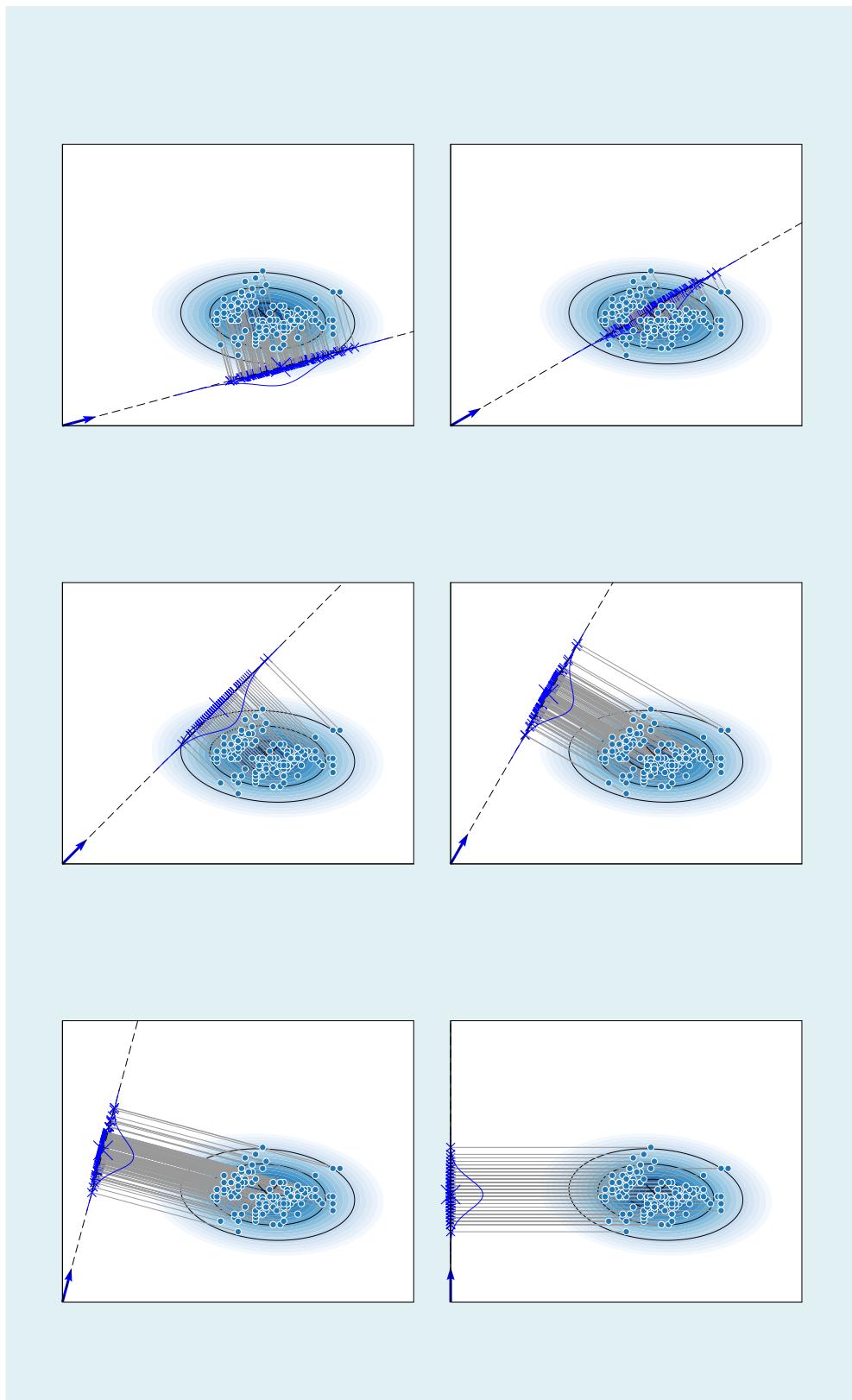


图 11. 平面上点、线投影

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

28

Geometric Transformations in 3D Space

立体几何变换

平移、缩放、旋转、镜像、投影、剪切



超现实主义极具破坏力，它打破视觉的枷锁。

Surrealism is destructive, but it destroys only what it considers to be shackles limiting our vision.

—— 萨尔瓦多·达利 (Salvador Dali) | 西班牙超现实主义画家 | 1904 ~ 1989



- ◀ `matplotlib.pyplot.quiver()` 绘制箭头图
- ◀ `matplotlib.pyplot.scatter()` 绘制散点图
- ◀ `numpy.column_stack()` 将两个矩阵按列合并
- ◀ `numpy.concatenate()` 将多个数组进行连接
- ◀ `numpy.cos()` 计算余弦
- ◀ `numpy.deg2rad()` 将角度转化为弧度
- ◀ `numpy.linspace()` 在指定的间隔内, 返回固定步长的数据
- ◀ `numpy.ones_like()` 用来生成和输入矩阵形状相同的全 1 矩阵
- ◀ `numpy.roll()` 将数组中的元素按照指定的偏移量进行循环移动，并返回一个新的数组。
- ◀ `numpy.sin()` 计算正弦
- ◀ `numpy.vstack()` 返回竖直堆叠后的数组

28.1 立体几何变换

上一章介绍的在平面上的几何变换（平移、缩放、旋转、镜像、投影、剪切）也可以用在三维空间中。表 1 总结常见立体几何仿射变换。在鸢尾花书中，最常用的 4 种几何变换为平移、缩放、旋转、投影，下面我们逐个可视化这三种几何变换。

本章介绍的这些立体几何变换涉及的数学工具将在《矩阵力量》中展开讲解。

表 1. 常见仿射变换，立体几何

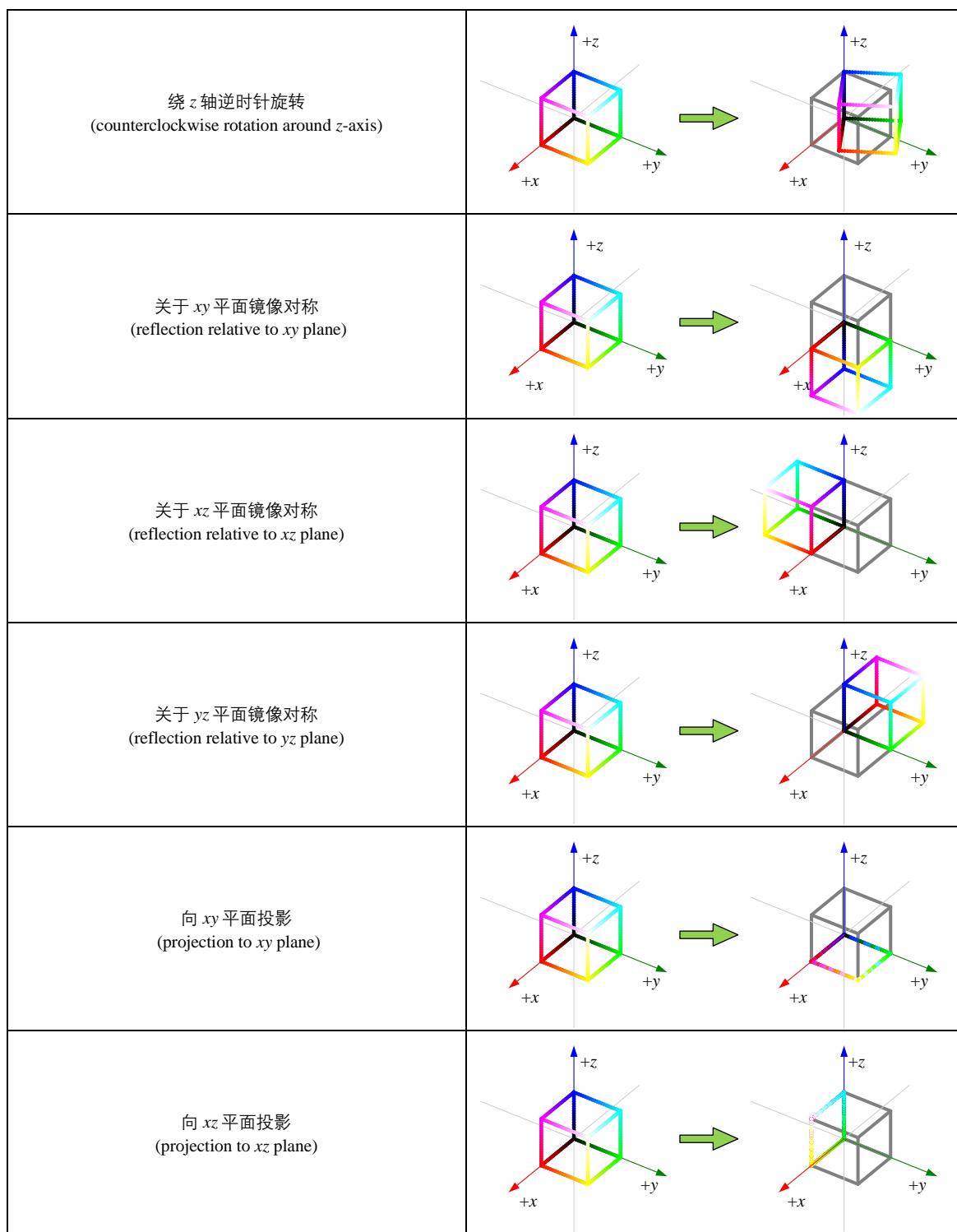
几何变换	示例
平移 (translation)	
等比例缩放 s 倍 (equal scaling)	
非等比例缩放 (unequal scaling)	
绕 x 轴逆时针旋转 (counterclockwise rotation around x -axis)	
绕 y 轴逆时针旋转 (counterclockwise rotation around y -axis)	

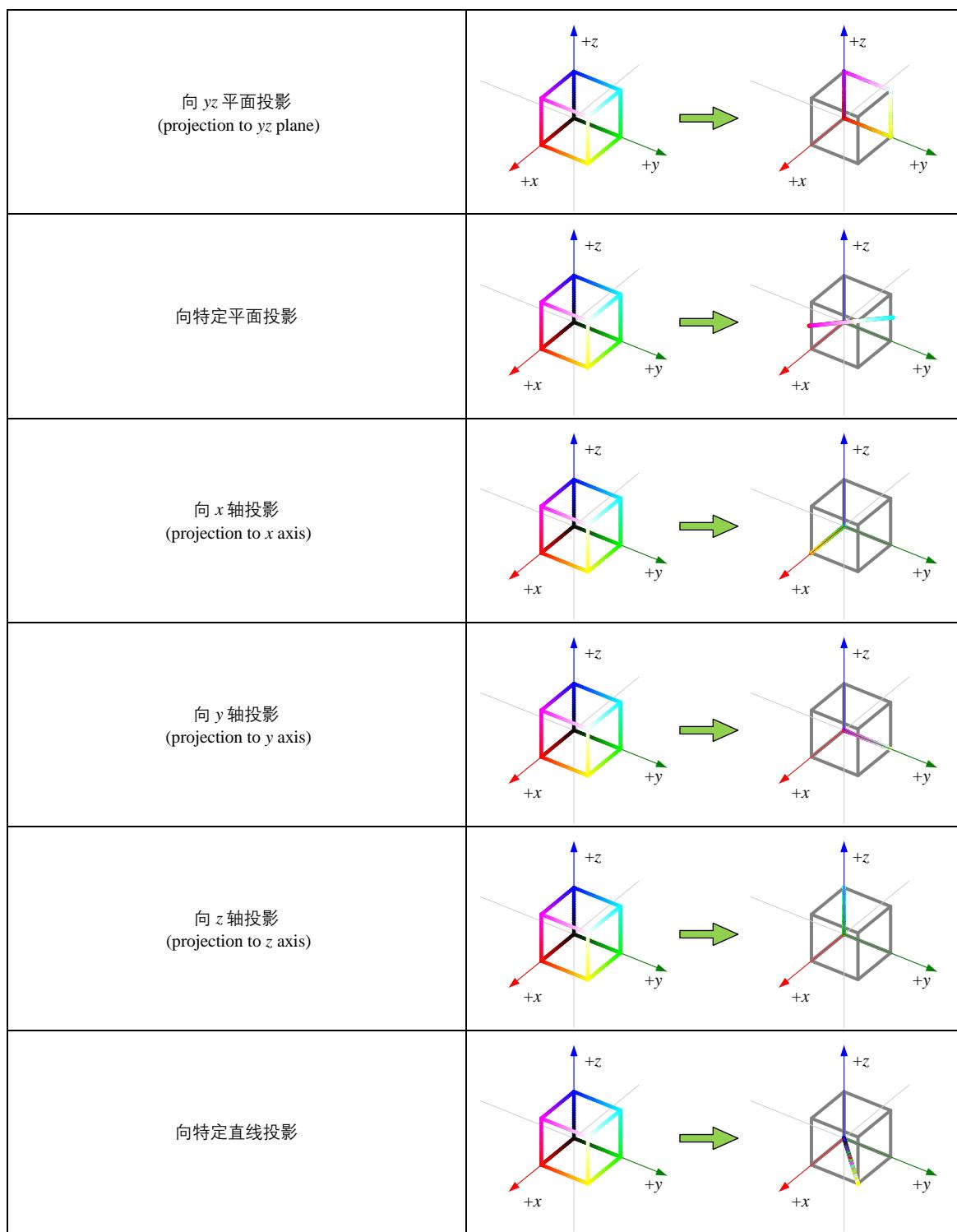
本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

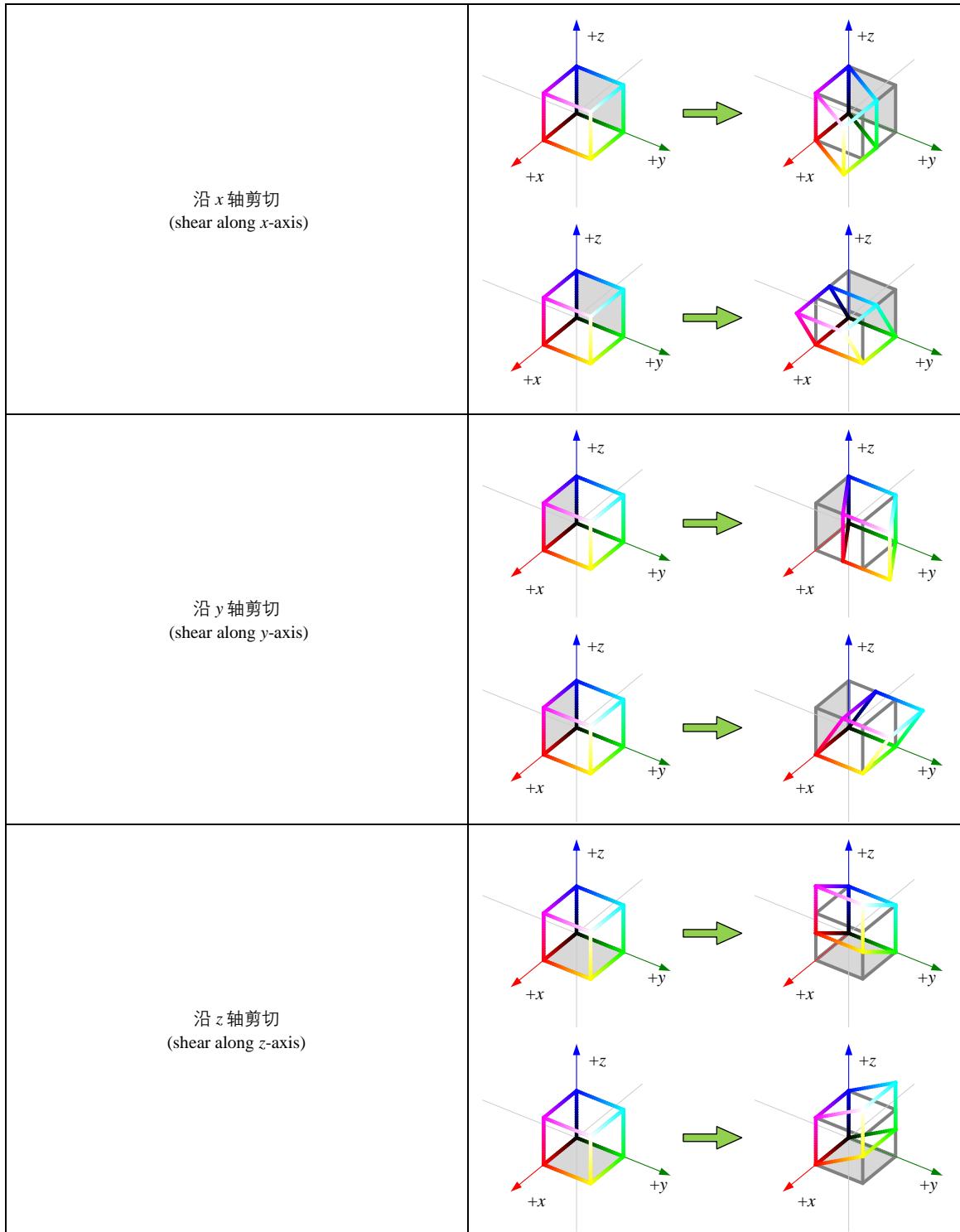
代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com







28.2 4 种常用几何变换

为了方便可视化下文 4 种常用的立体几何变换，我们给出如图 1 所示的可视化方案。

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

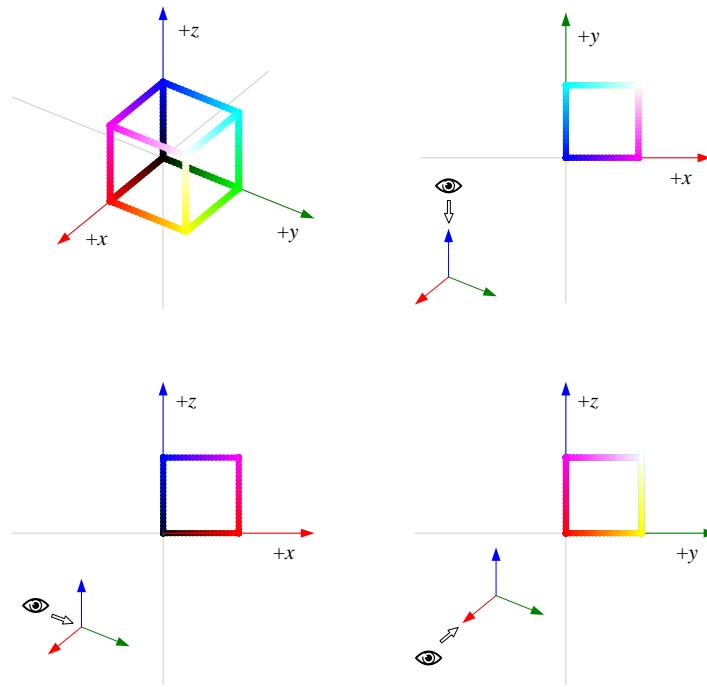


图 1. 原始数据，四个投影视角

平移

图 2 所示为从四个投影视角展示平移。

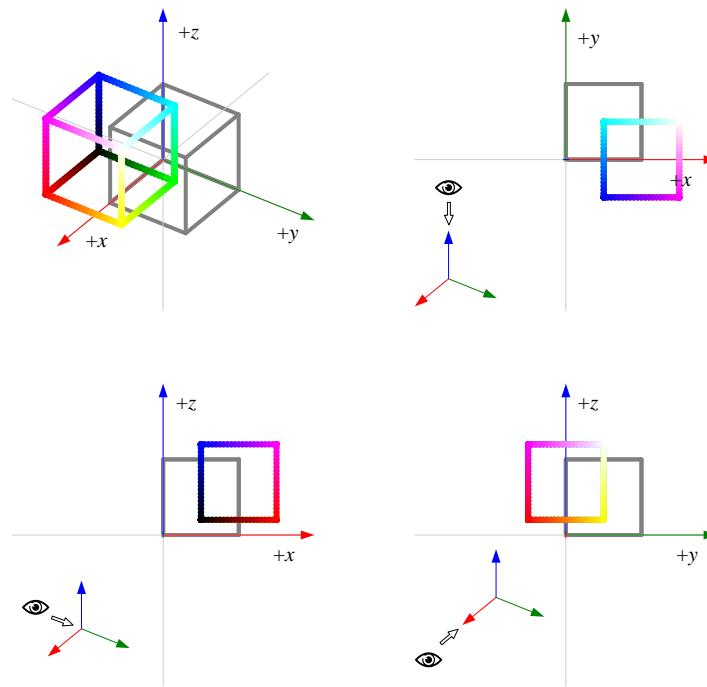


图 2. 平移，四个投影视角

缩放

图 5 所示为从四个投影视角展示等比例缩放、非等比例缩放。

旋转

三维空间中，三个旋转角度和飞机姿态的三个角度密切相关。如图 3 所示，翻滚角 (Roll Angle) 是飞机绕其纵轴旋转的角度，用于描述飞机的侧倾程度。当飞机向右侧倾斜时，翻滚角为正值；向左倾斜时，翻滚角为负值。

俯仰角 (Pitch Angle) 是飞机绕其横轴旋转的角度，用于描述飞机的仰角或俯角。当飞机向上抬头时，俯仰角为正值；向下俯冲时，俯仰角为负值。

偏航角 (Yaw Angle) 是飞机绕其垂直轴旋转的角度，用于描述飞机的航向偏转。当飞机顺时针旋转时，偏航角为正值；逆时针旋转时，偏航角为负值。

如图 4 所示，这些角度通常使用欧拉角 (Euler Angles) 系统来表示，其中翻滚角、俯仰角和偏航角分别绕飞机的纵轴、横轴和垂直轴旋转。图 6、图 7、图 8 所示为从四个投影视角展示旋转操作。特别地，图 8 中正方体分别经过三个方向旋转。

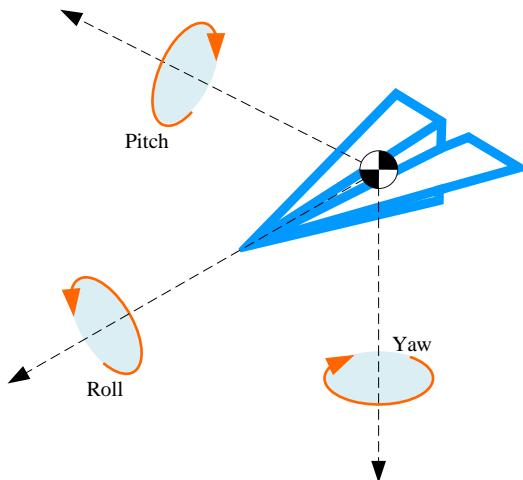


图 3. 飞机姿态的三个角度

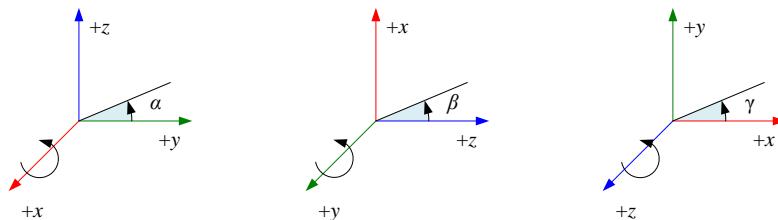


图 4. 三个旋转角度

投影

图 9、图 10 所示为从四个投影视角展示投影操作。

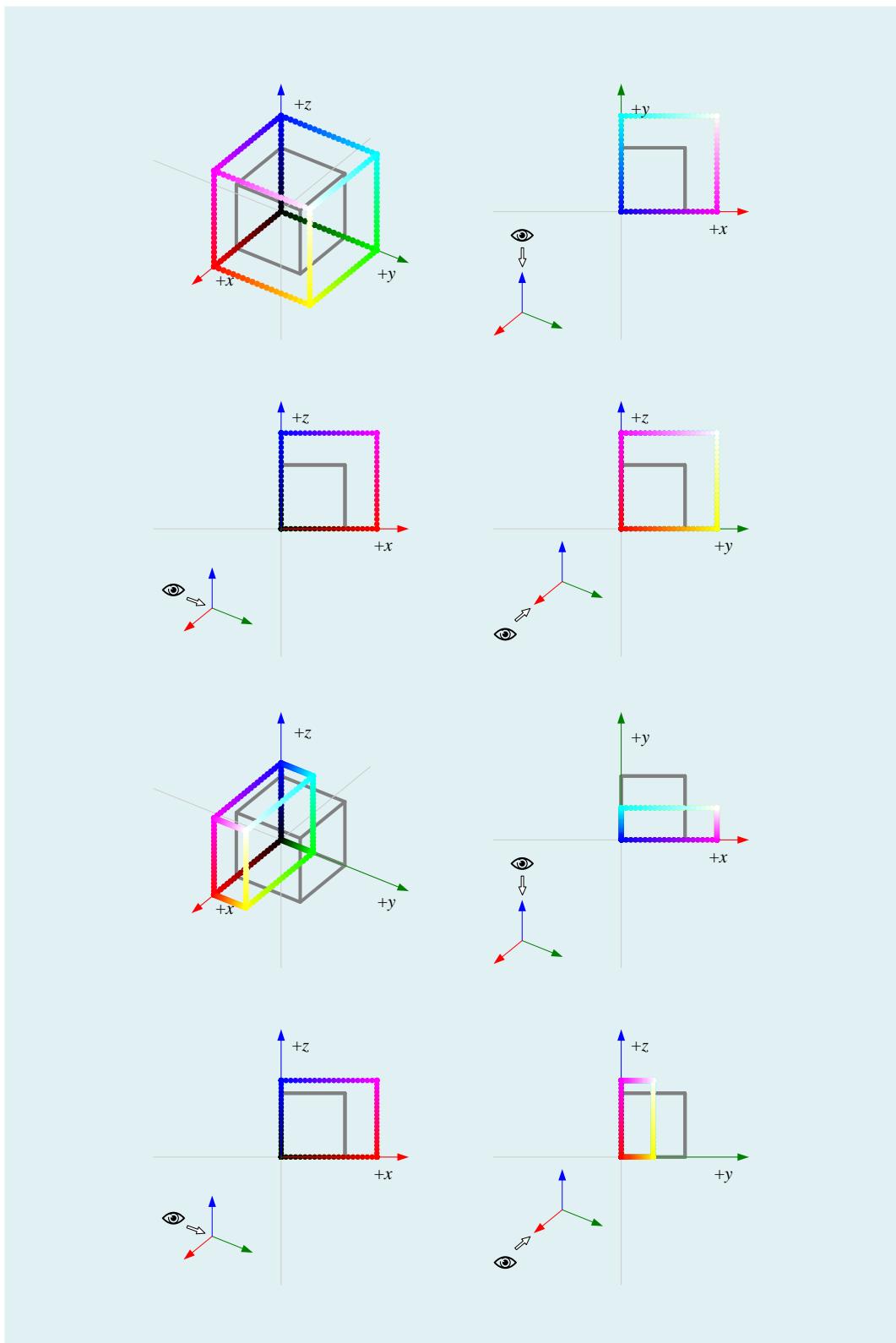


图 5. RGB 立方体框线，等比例缩放、非等比例缩放，四个投影视角

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

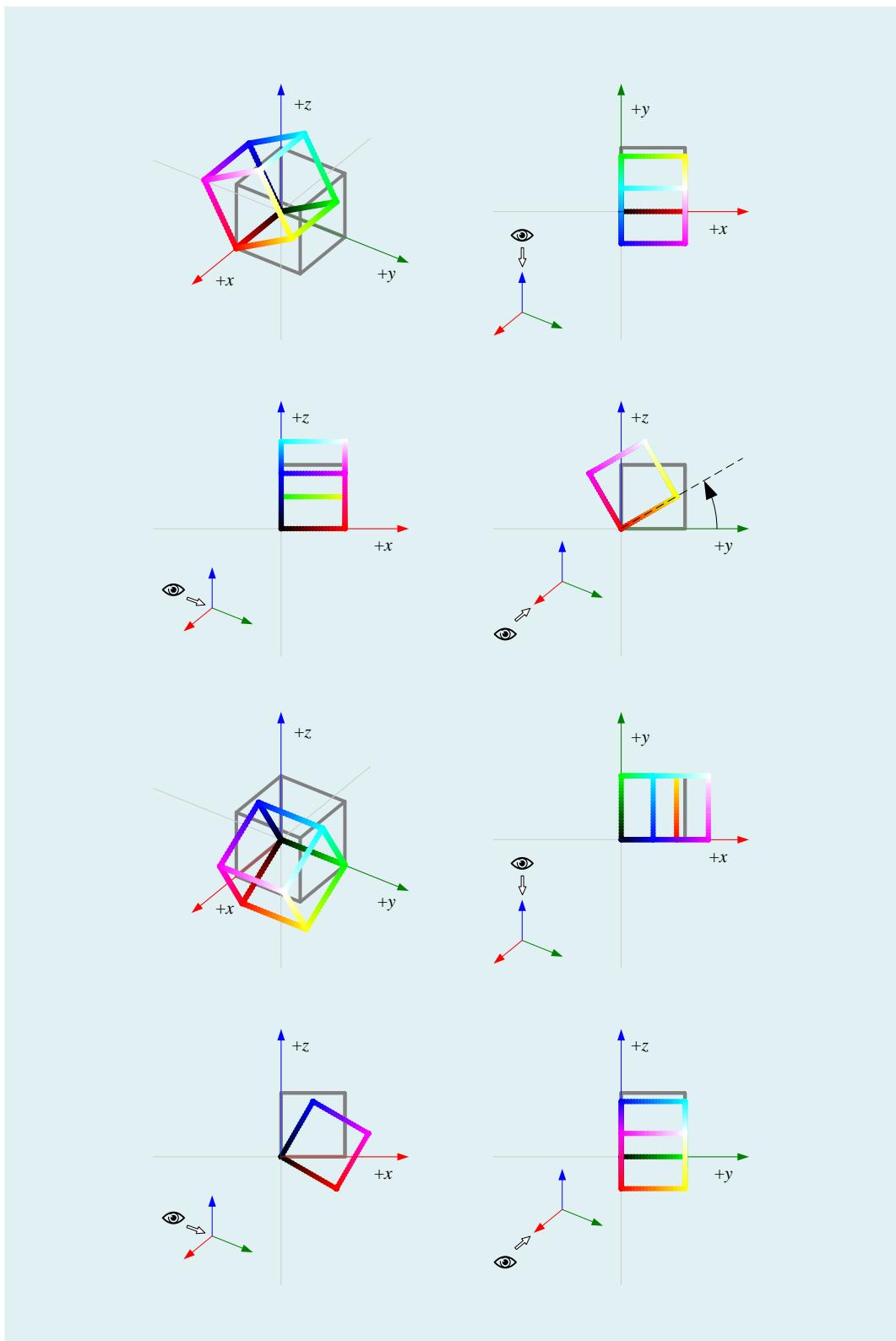


图 6. RGB 立方体框线，旋转，第 1 组，四个投影视角

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

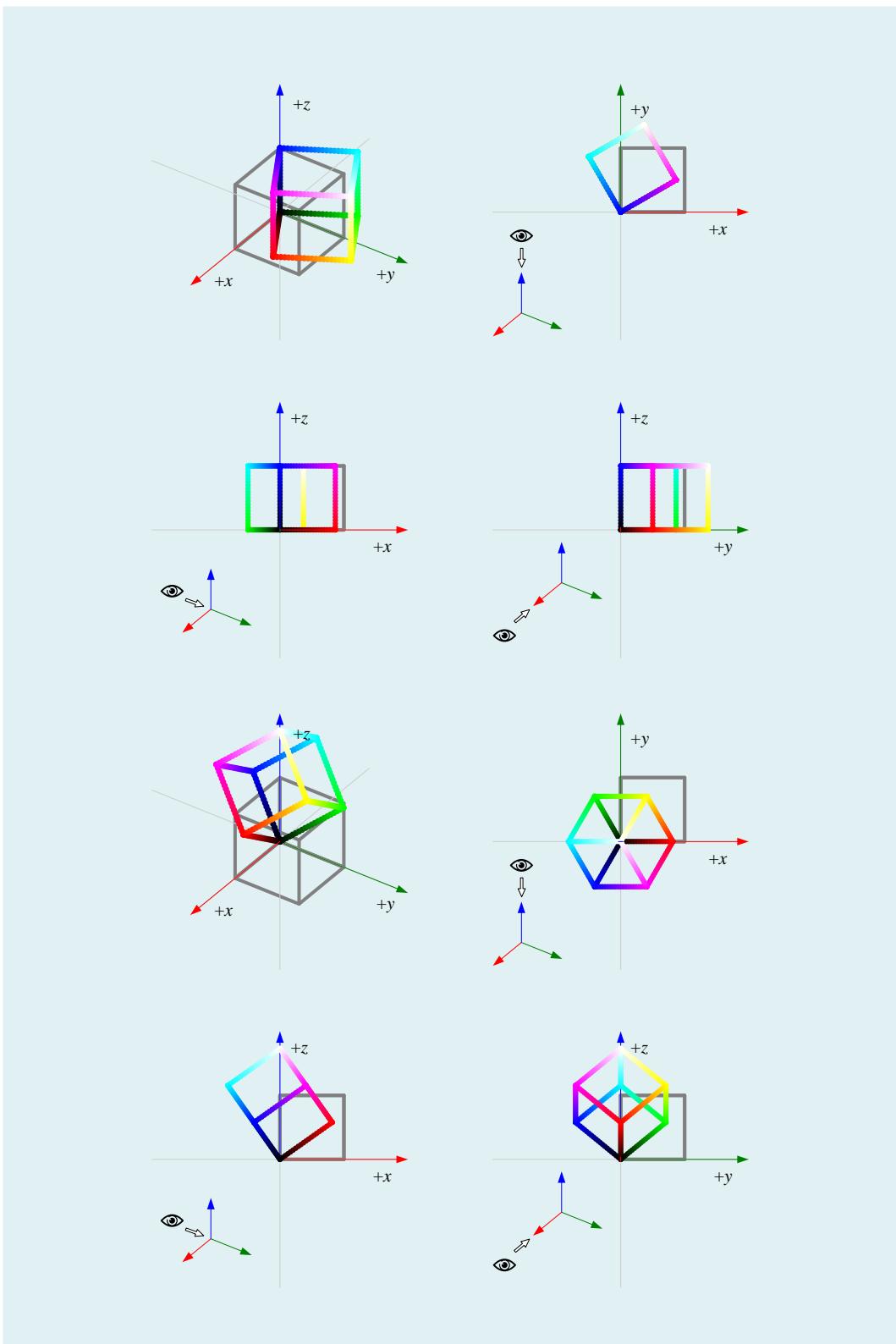


图 7. RGB 立方体框线，旋转，第 2 组，四个投影视角

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

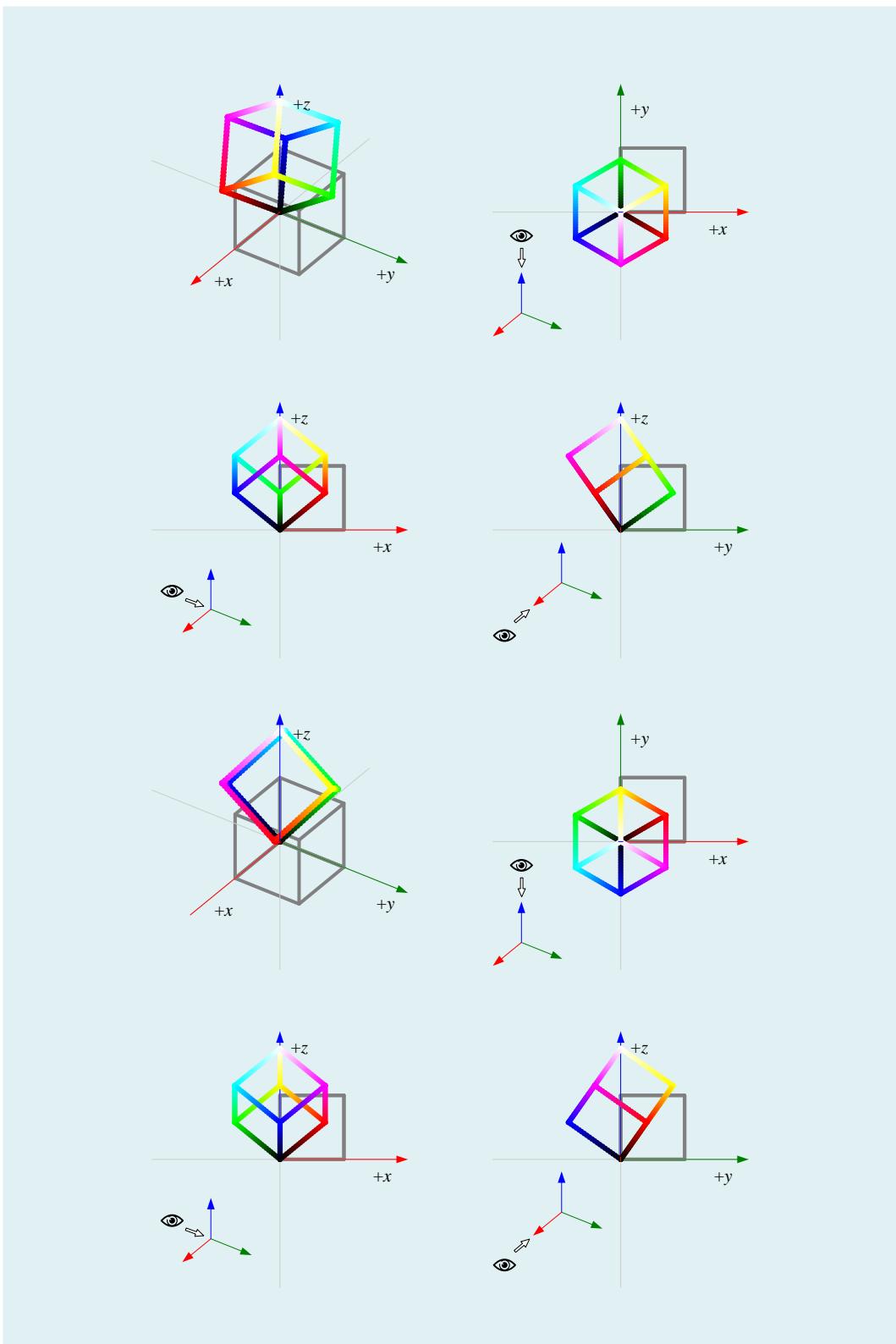


图 8. RGB 立方体框线, 旋转, 第 3 组, 四个投影视角

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

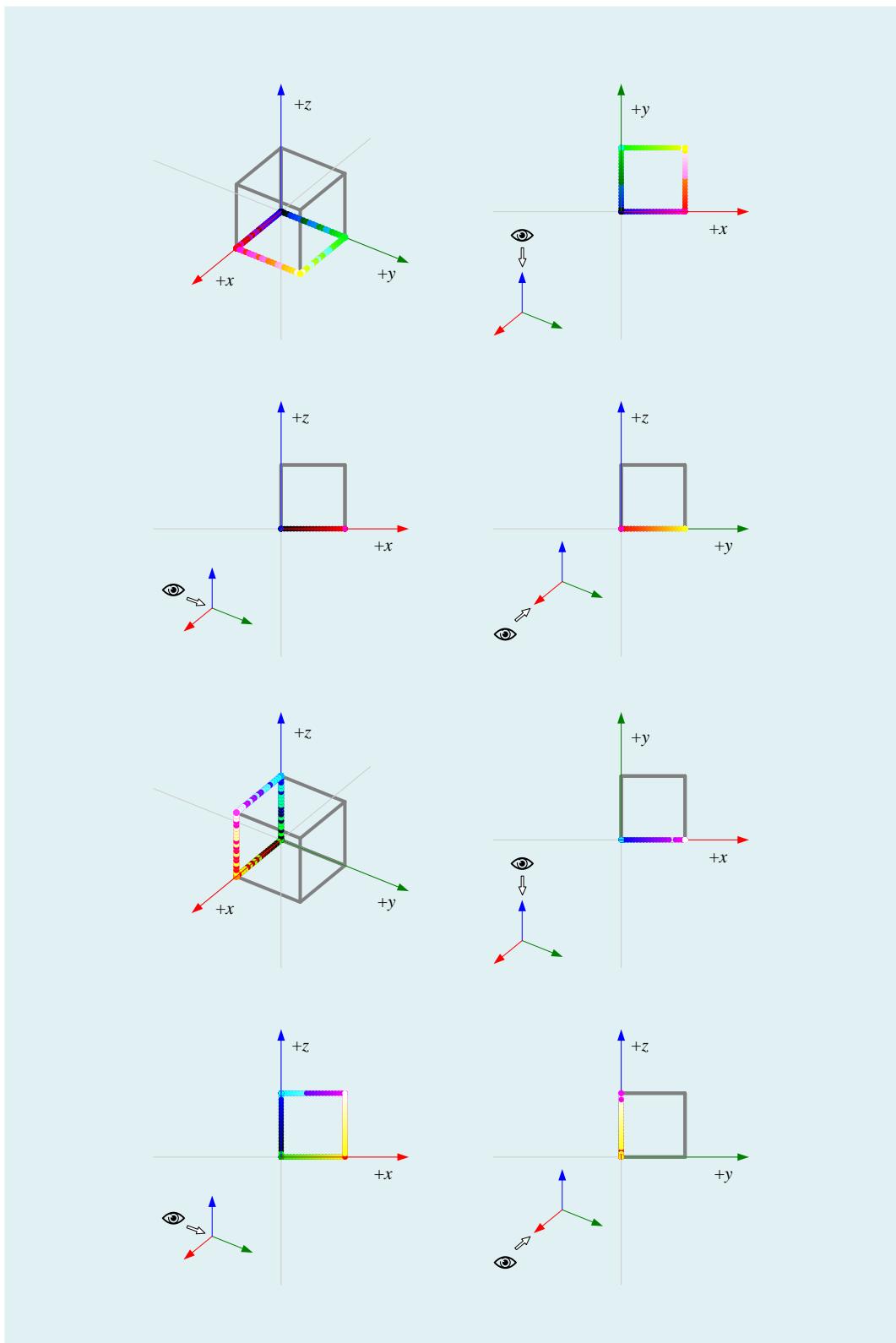


图 9. RGB 立方体框线，投影，第 1 组，四个投影视角

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

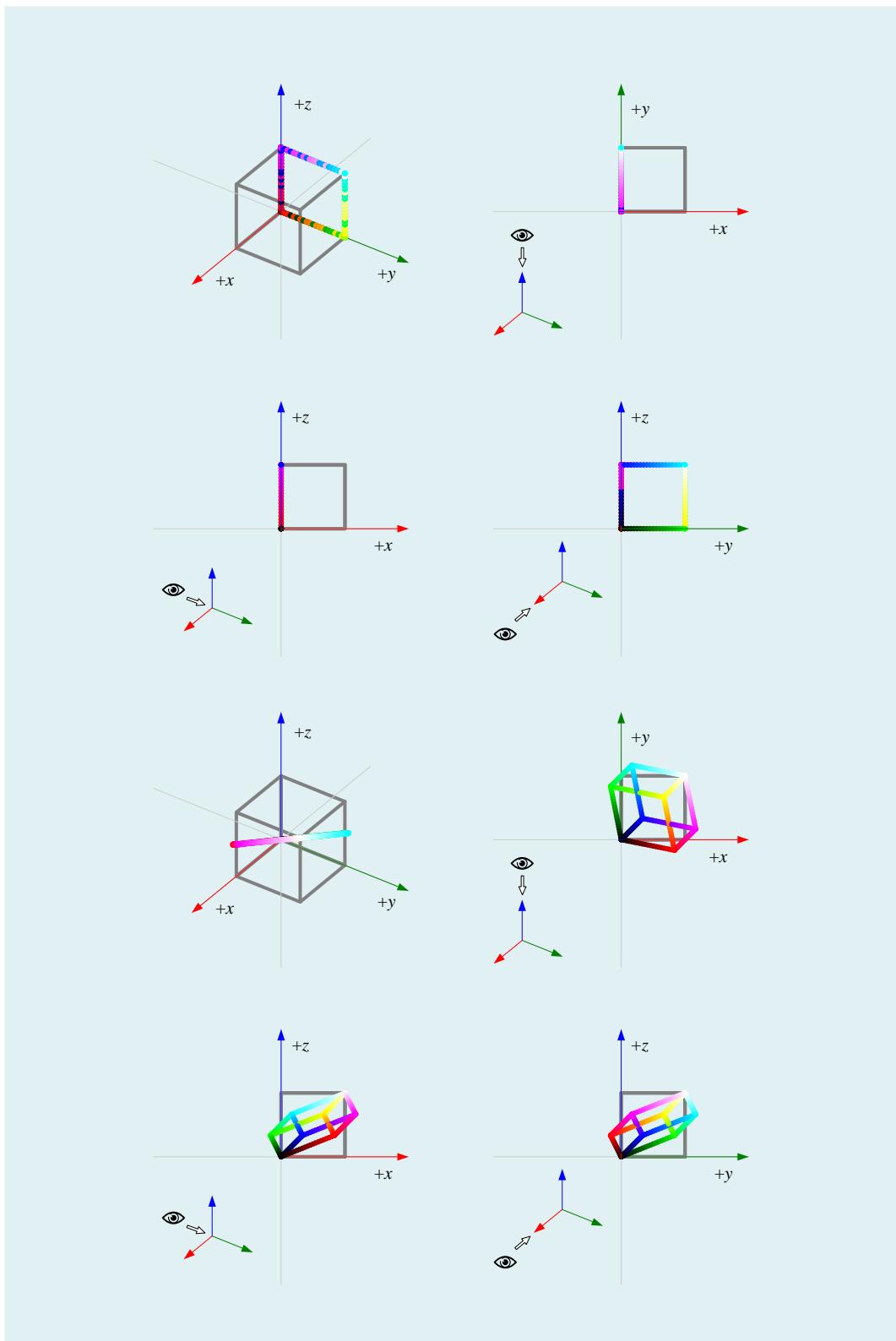


图 10. RGB 立方体框线，投影，第 2 组，四个投影视角

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

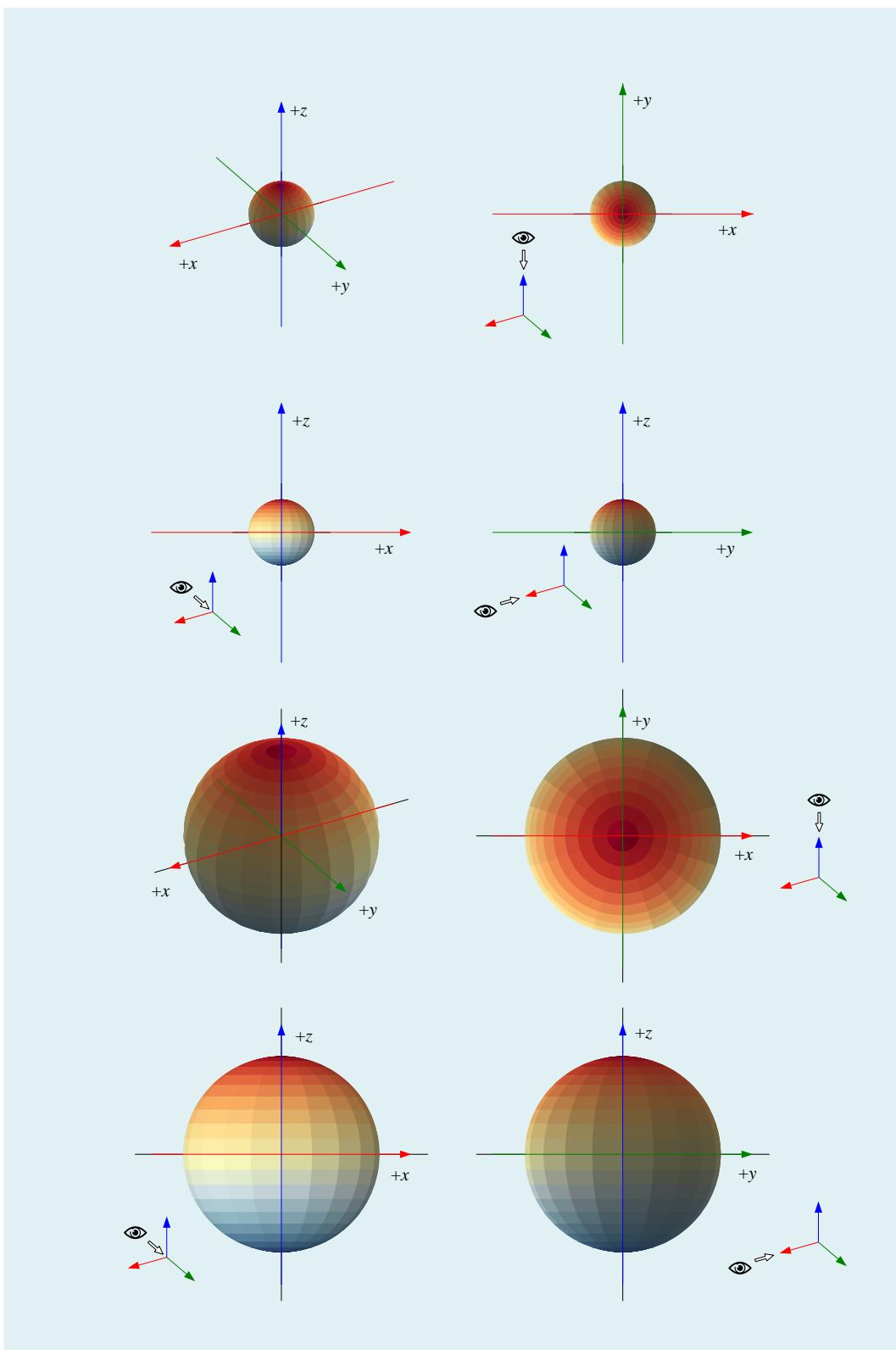


图 11. 单位球，单位球等比例放大

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

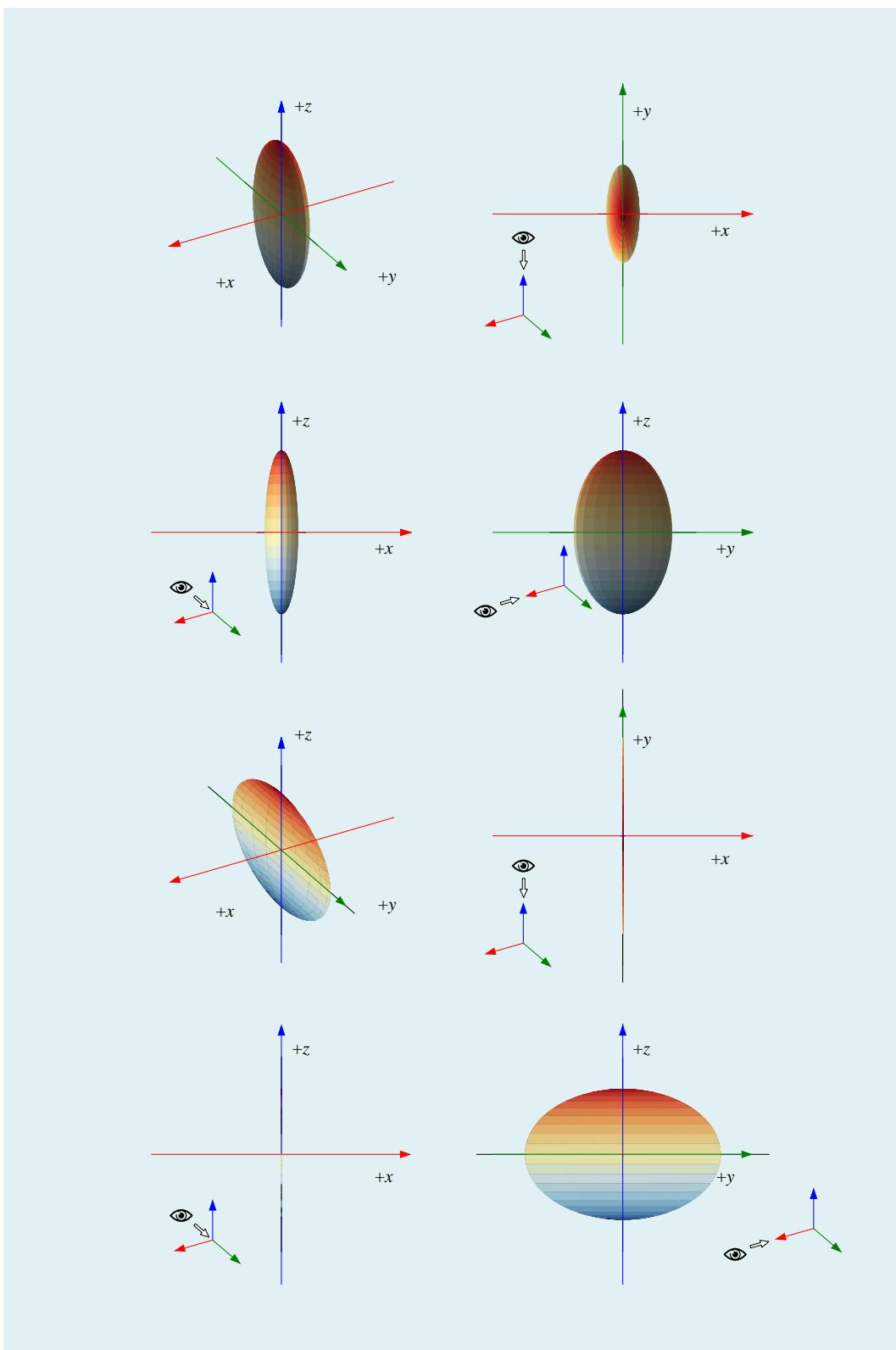


图 12. 单位球非等比例缩放，单位球 x 、 z 方向放大且 y 方向压扁

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

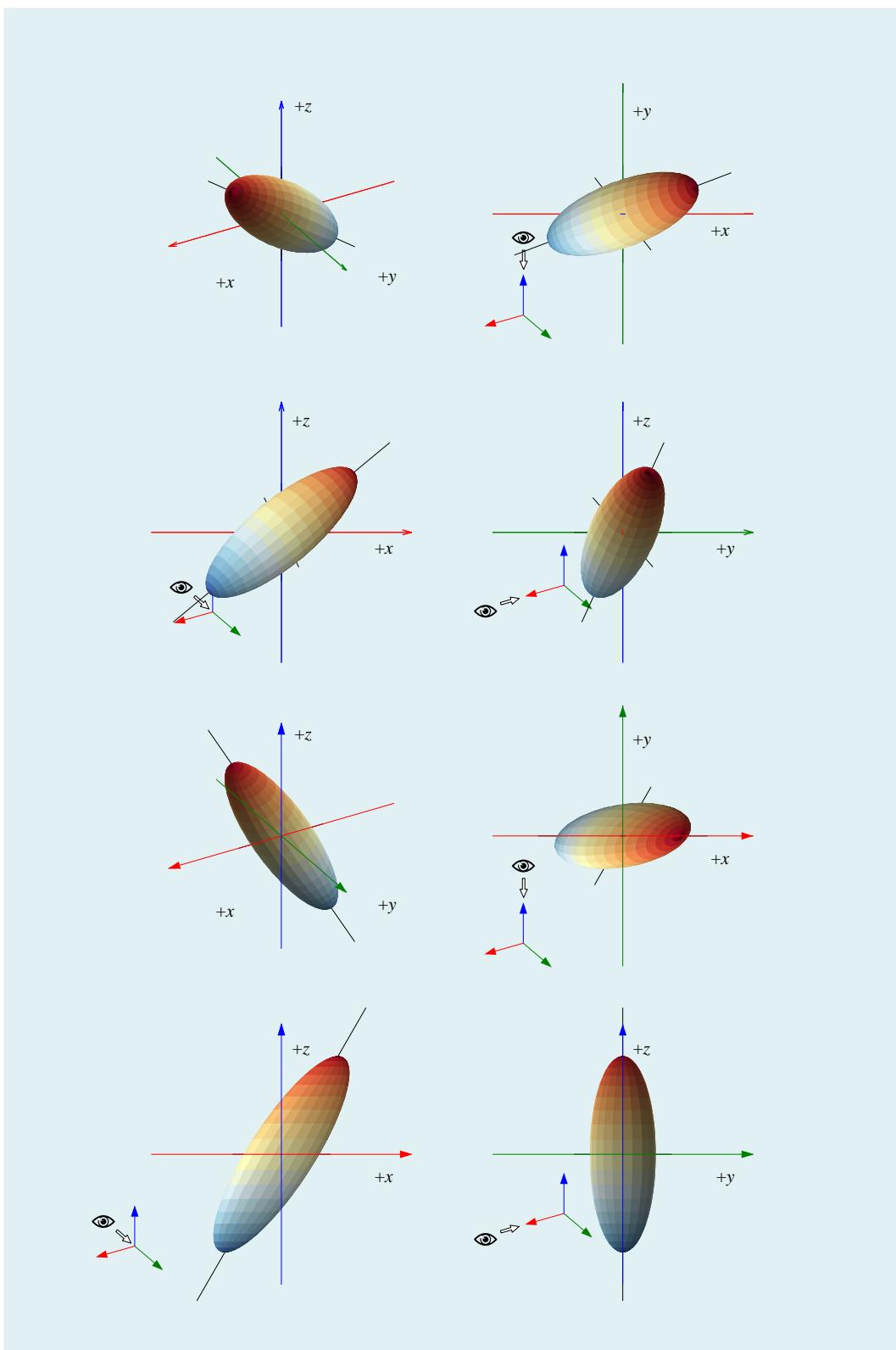


图 13. 单位球先缩放、再三轴旋转，单位球先缩放、再沿 x 剪切

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com



Bézier curve

贝塞尔曲线

计算机图形学中特别重要的参数曲线



我要扼住命运的喉咙。

I shall seize fate by the throat.

——路德维希·范·贝多芬 (Ludwig van Beethoven) | 德意志作曲家、钢琴演奏家 | 1770 ~ 1827



- ◀ `math.factorial()` 计算给定整数的阶乘
- ◀ `numpy.column_stack()` 将两个矩阵按列合并
- ◀ `numpy.interp()` 给定的一维数组上进行线性插值
- ◀ `numpy.random.rand()` 用于生成指定形状的随机数组，随机数服从 0 到 1 之间的均匀分布
- ◀ `random.random()` 生成一个介于 0 到 1 之间的随机浮点数，包括 0 但不包括 1
- ◀ `scipy.interpolate.interp1d()` 一维插值
- ◀ `scipy.interpolate.interp2d()` 二维插值

29.1 贝塞尔曲线

贝塞尔曲线 (Bézier curve) 是一种常用于计算机图形学中的数学曲线。它由法国工程师皮埃尔·贝塞尔 (Pierre Bézier) 在 19 世纪中叶发明。

贝塞尔曲线最初是为了描述船只的水线曲线。后来，贝塞尔曲线被广泛应用于计算机图形学中，用于绘制平滑曲线，如字体、二维图形和三维模型等。多数矢量图形都离不开贝塞尔曲线。

贝塞尔曲线是由一组控制点和一个阶数确定的曲线。控制点是定义曲线形状的关键点，阶数是定义贝塞尔曲线逼近实际曲线的程度的参数。通常情况下，阶数等于控制点的数量减 1。

贝塞尔曲线的特点是它们具有局部控制性，这意味着通过调整单个控制点的位置，可以轻松地改变曲线的形状。此外，它们也具有平滑的曲线形状和良好的数学性质。Adobe Photoshop、Illustrator 中的钢笔曲线绘图工具实际上使用的便是贝塞尔曲线。

本质上来说，贝塞尔曲线就是一种插值方法。贝塞尔曲线可以是一阶曲线、二阶曲线、三阶曲线等，其阶数决定了曲线的平滑程度。下面首先介绍一阶贝塞尔曲线原理。

29.2 一阶

一阶曲线由两个控制点组成，形成一条直线。如图 1 所示，简单来说一阶贝塞尔曲线就是两点之间连线。图中 t 代表权重，取值范围为 $[0, 1]$ 。 t 越大，点 $B(t)$ 距离 P_0 越近，如图中暖色 \times ，相当于 P_0 对 $B(t)$ 影响越大。相反， t 越小，点 $B(t)$ 距离 P_1 越近，如图中冷色 \times ，相当于 P_1 对 $B(t)$ 影响大。

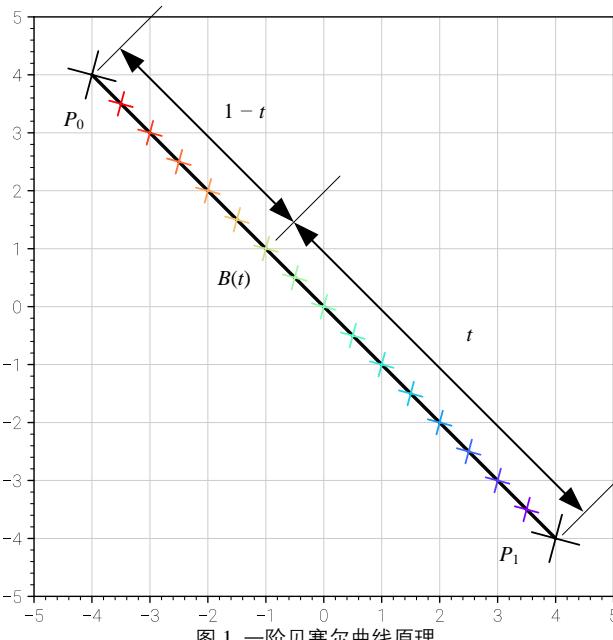


图 1. 一阶贝塞尔曲线原理

29.3 二阶

二阶贝塞尔曲线由三个控制点组成，形成一条弯曲的曲线。如图2所示， P_0 和 P_2 点控制了曲线（黑色线）的两个端点，而 P_1 则决定的曲线的弯曲行为。实际上图2中黑色二阶贝塞尔曲线上每一个点都经历了两组线性插值得到。如图3所示，设定 $t = 13/16$ ，通过第一组线性插值，我们分别得到了 P_0P_1 线段上的 A_0 ，以及 P_1P_2 线段上的 A_1 。然后通过第二组线性插值，我们便得到 A_0A_1 线段上的 $B(13/16)$ 。当 t 在 $[0, 1]$ 之间连续取值时，我们便得到了二阶贝塞尔曲线上的一系列点。

图5给出几个不同的贝塞尔曲线， P_1 点坐标为随机生成。大家可能已经发现，贝塞尔曲线一般不会经过 P_1 点，除非 P_0 、 P_1 、 P_2 三点在同一条直线上。

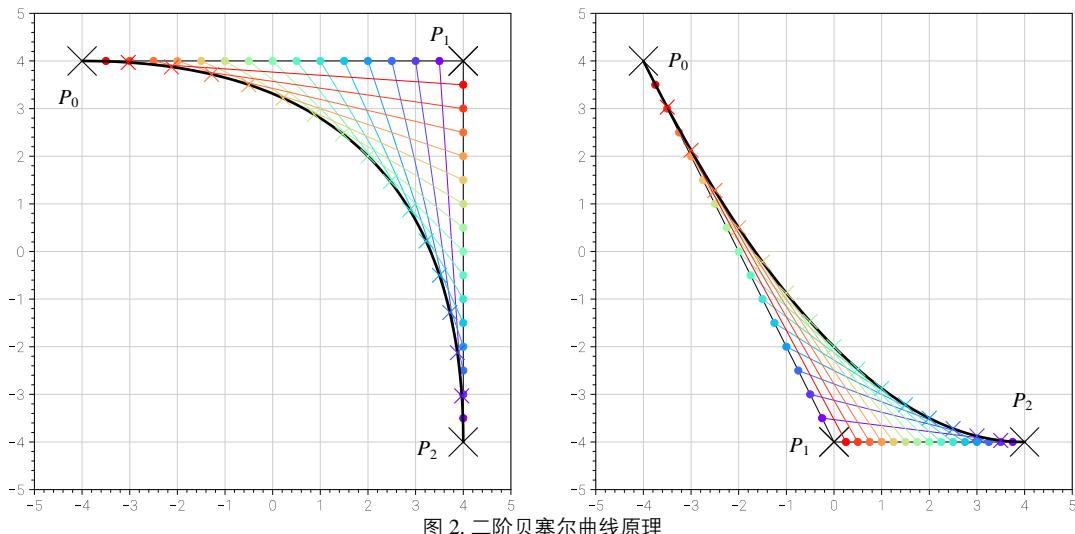
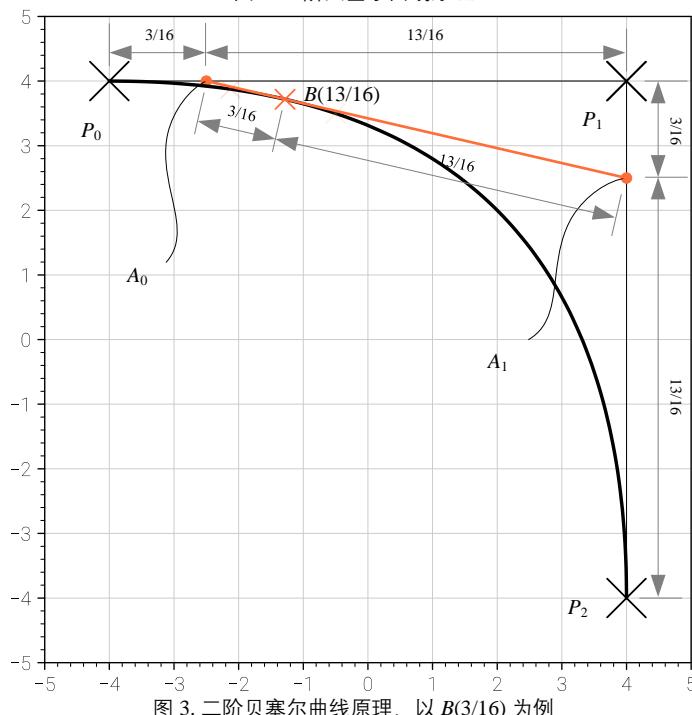


图 2. 二阶贝塞尔曲线原理

图 3. 二阶贝塞尔曲线原理，以 $B(13/16)$ 为例

29.4 三阶

图 4 比较一阶、二阶、三阶贝塞尔曲线。三阶贝塞尔曲线由四个控制点组成，形成更加复杂的曲线。如图 4 (c) 所示， P_0 和 P_3 点同样控制了曲线的两个端点，而 P_1 和 P_2 两点决定的曲线的弯曲行为。图 6 所示为一系列三阶贝塞尔曲线， P_1 和 P_2 为随机数。

图 7 所示为一组四阶贝塞尔曲线，曲线的弯曲行为更加复杂。

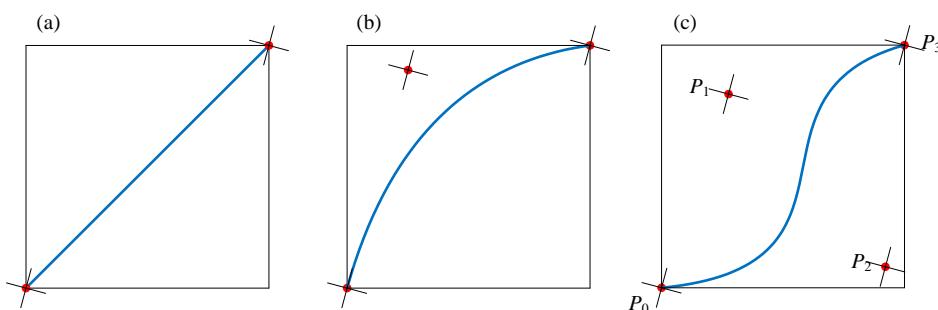


图 4. 贝塞尔曲线原理，比较一阶、二阶、三阶

图 8 则是采用 Python 复刻的用贝塞尔曲线创作的“鸢尾花曲线”。“鸢尾花曲线”来自于 Oliver Brotherhood 的开源设计创意。

29.5 三维空间

上述贝塞尔曲线还都仅限于平面，贝塞尔曲线也可以很容易扩展到三维空间。为了可视化贝塞尔曲线，我们把它们放在 RGB 色彩空间中。

图 9 所示为 RGB 色彩空间中 1 ~ 6 阶贝塞尔曲线。图中控制点的表示为 \times ，控制点之间的顺序连线为划线。图 9 中这些贝塞尔曲线有一个共同特点，它们的首尾两个控制点分别是黑色 (0, 0, 0)、白色 (1, 1, 1)。其他控制点则均由随机数发生器生成。

图 10 所示 8 阶贝塞尔曲线的 9 个控制点都是随机数发生器生成。

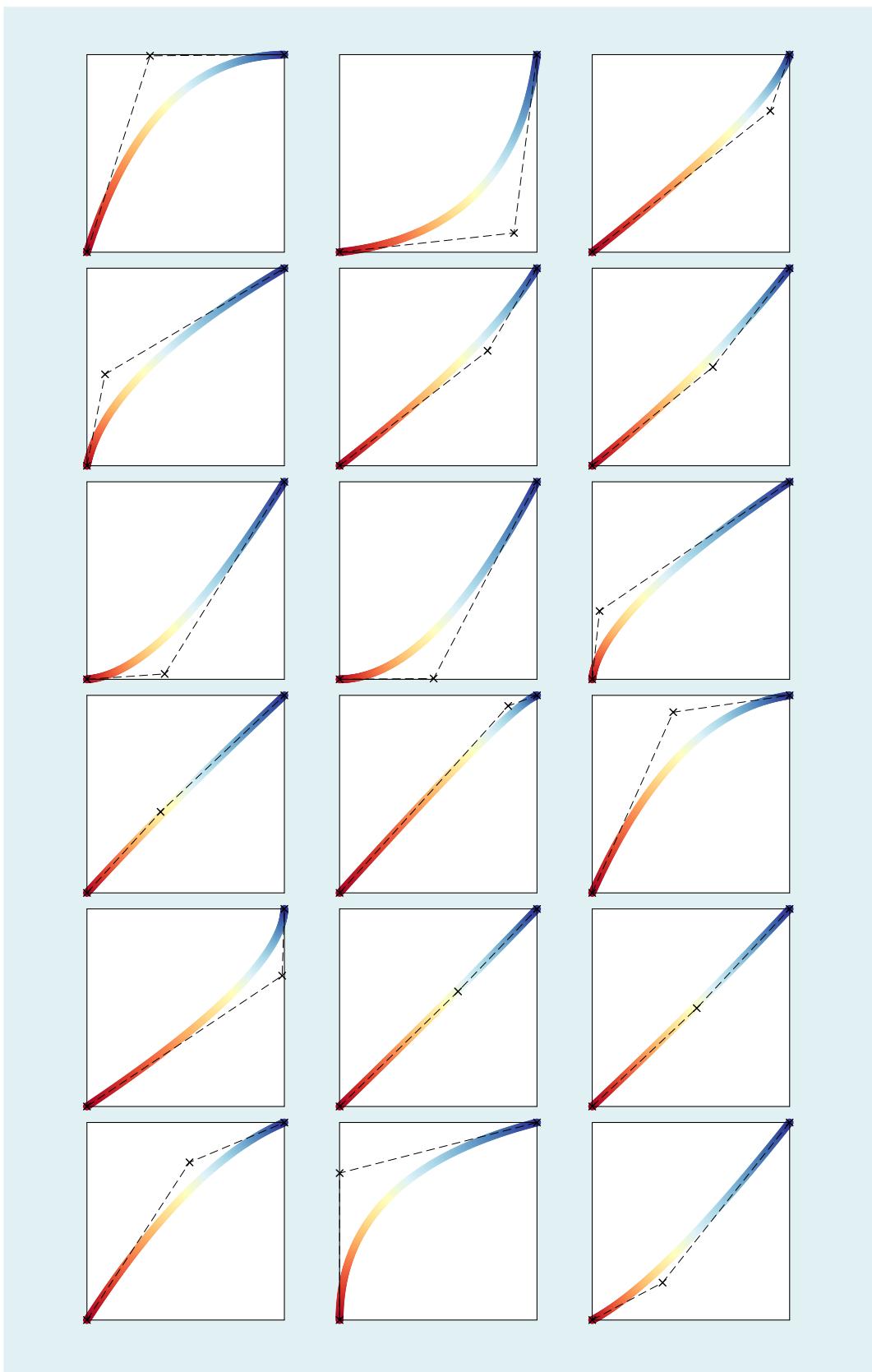


图 5. 二阶贝塞尔曲线

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

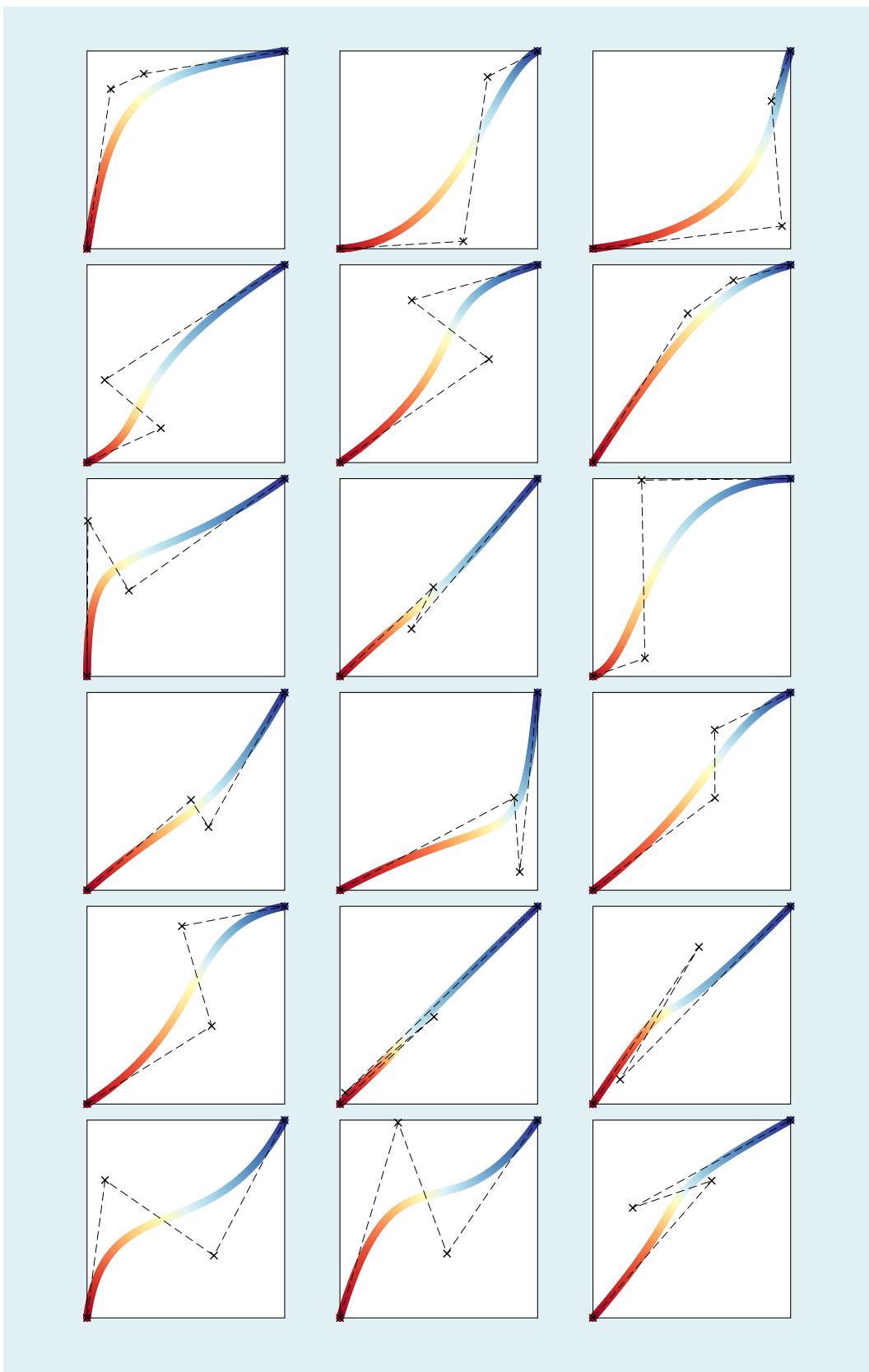


图 6. 三阶贝塞尔曲线

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

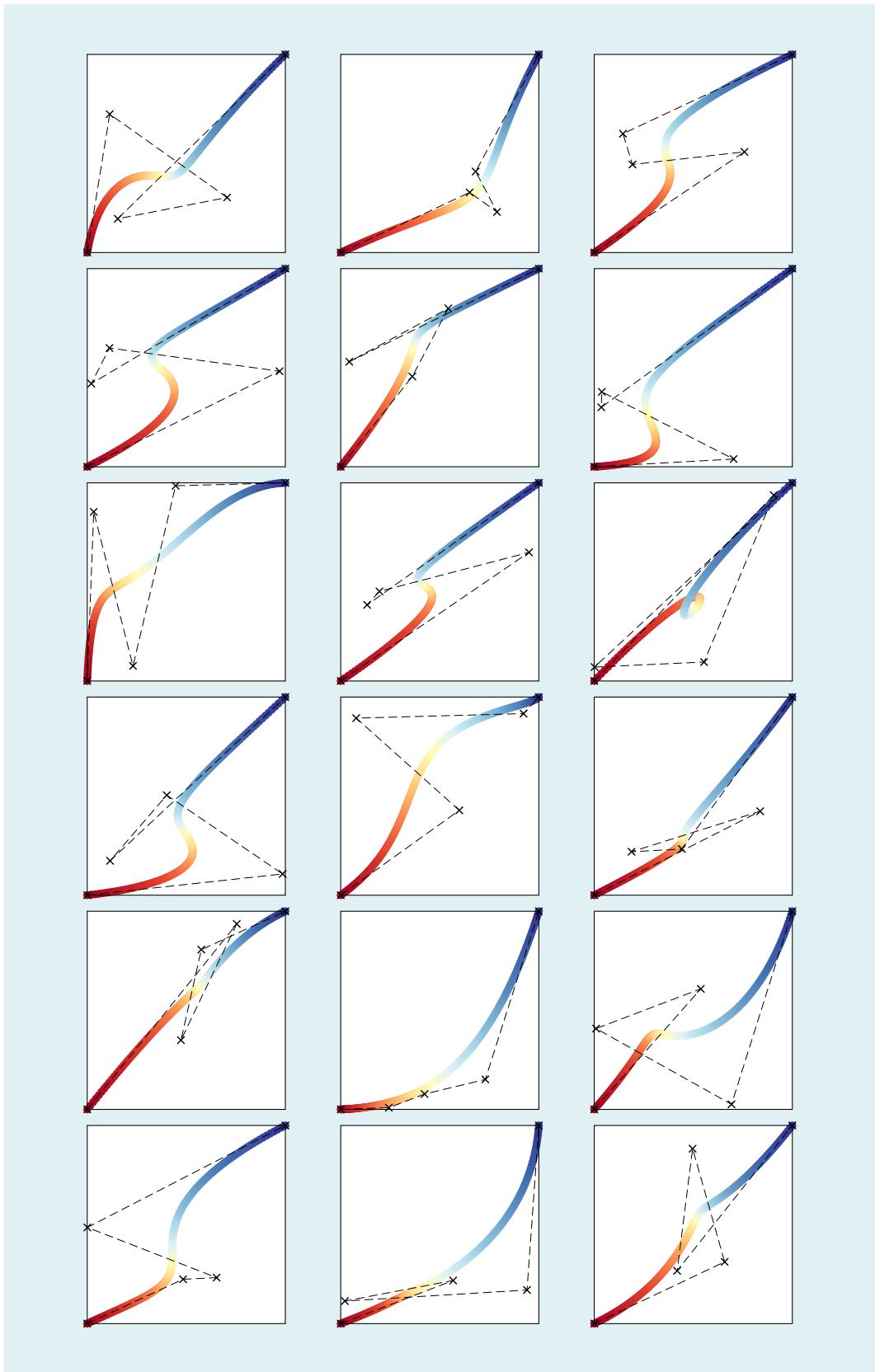


图 7. 四阶贝塞尔曲线

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

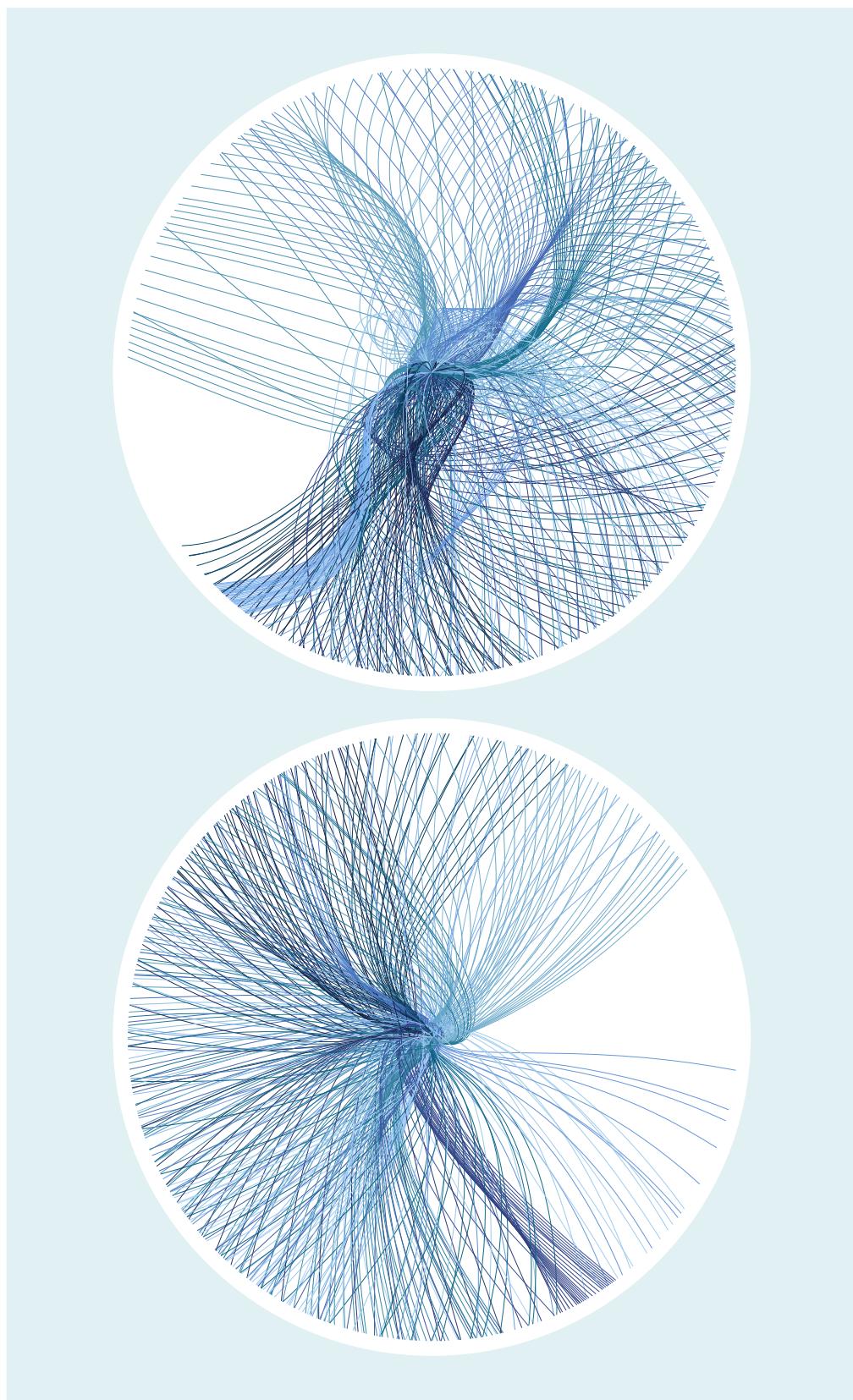


图 8. 用贝塞尔曲线绘制的“鸢尾花曲线”

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

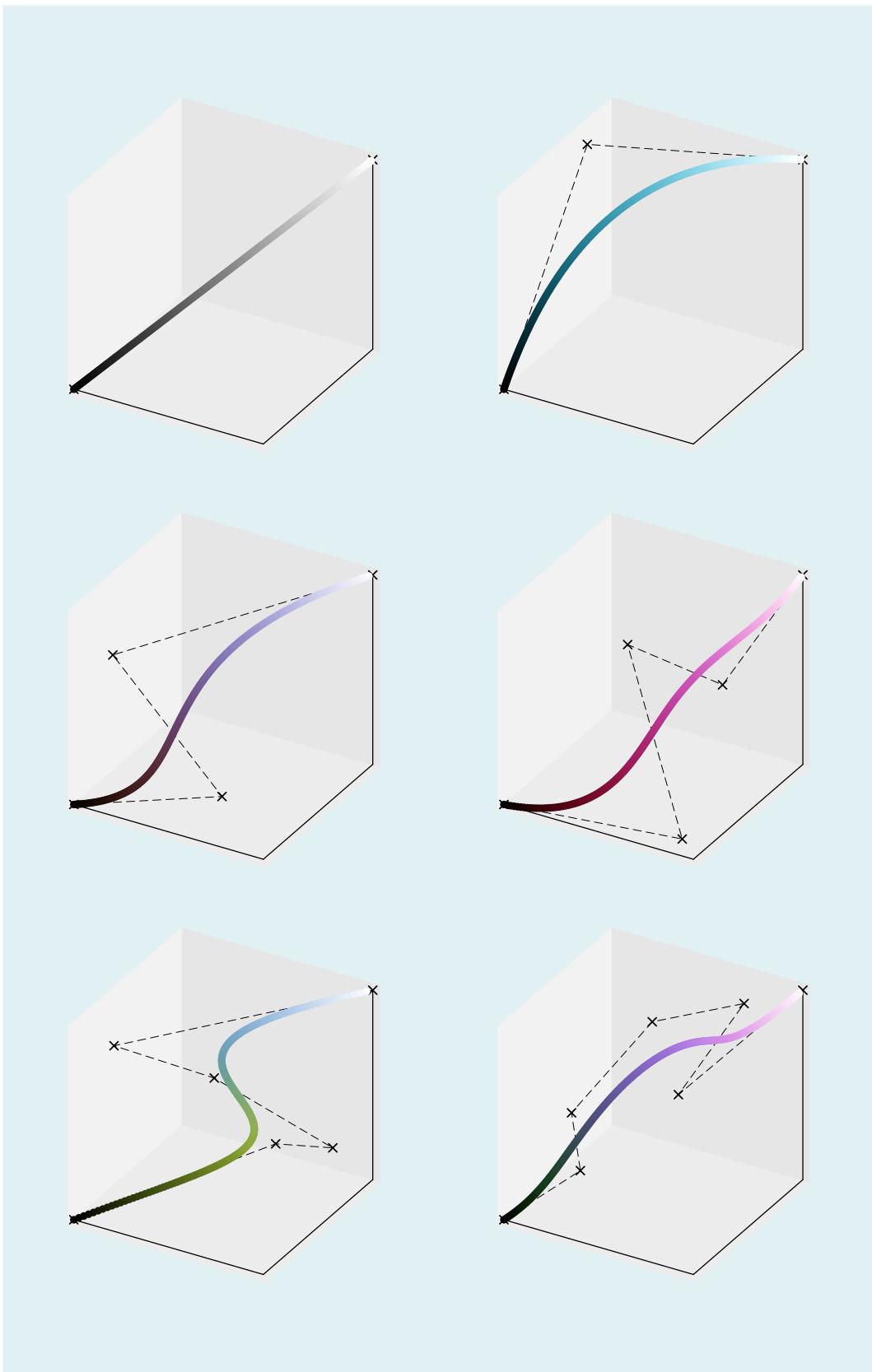


图 9. RGB 色彩空间中的 1 ~ 6 阶贝塞尔曲线

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

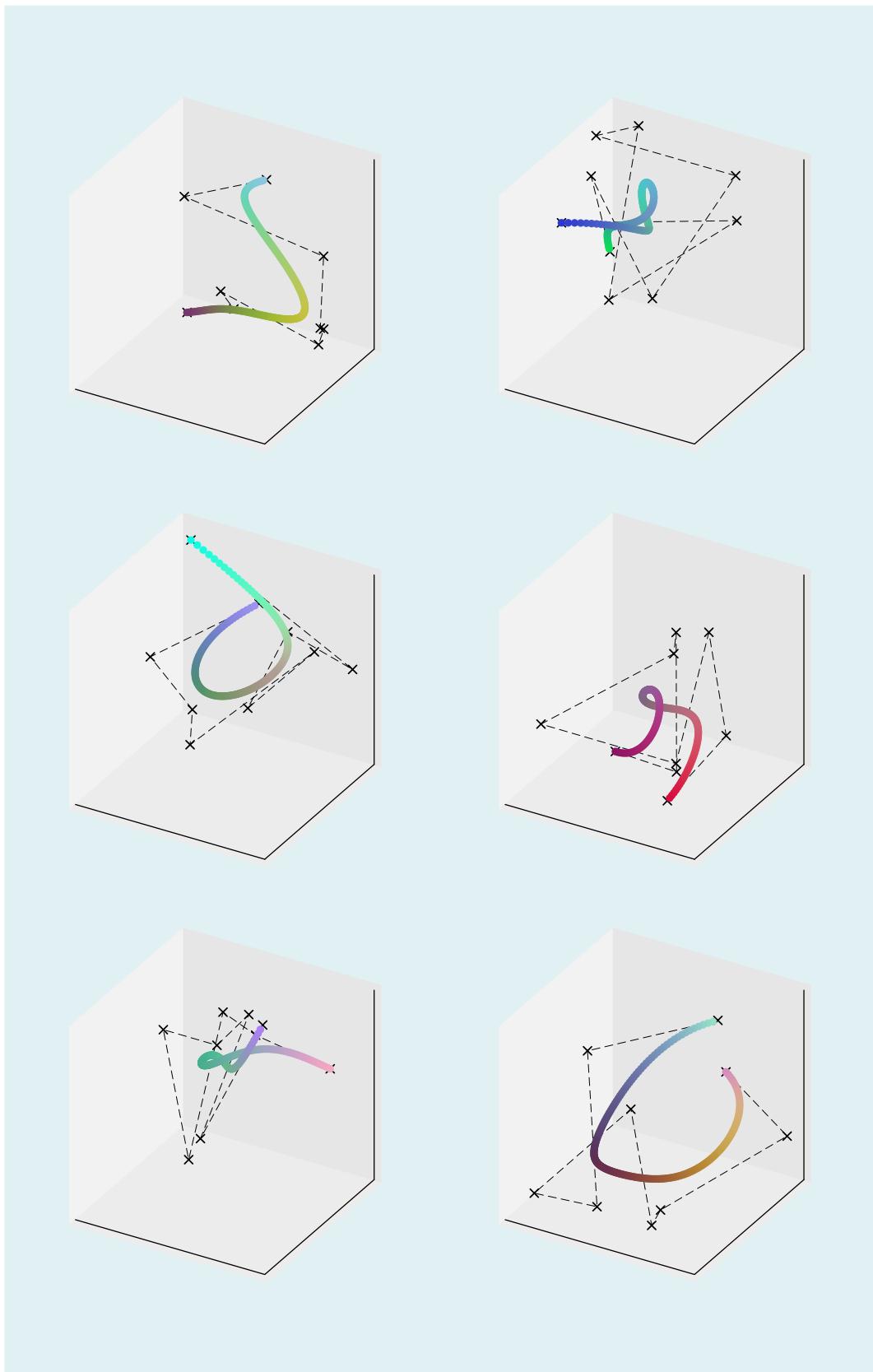


图 10. RGB 色彩空间中的几个 8 阶贝塞尔曲线，9 个控制点均由随机数发生器生成

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

30 Function 函数

可视化一元、二元、三元函数



至暗深夜也必然结束，太阳终将升起。

Even the darkest night will end and the sun will rise.

——雨果 (Victor Hugo) | 法国文学家 | 1802 ~ 1885



- ◀ `matplotlib.pyplot.contour()` 绘制等高线图
- ◀ `matplotlib.pyplot.contourf()` 绘制填充等高线图
- ◀ `matplotlib.pyplot.plot_wireframe()` 绘制线框图
- ◀ `numpy.linspace()` 在指定的间隔内, 返回固定步长的数据
- ◀ `numpy.meshgrid()` 创建网格化数据

30.1 函数

鸢尾花书的读者对于数学中函数这个概念当然不陌生。本章聊一聊函数的各种常见可视化方案。

什么是函数？

在代数中，函数是一种数学关系，它将一个或多个输入值映射 (mapping) 到唯一的输出值。函数可以用一个规则或方程式来表示，其中输入值称为自变量，输出值称为因变量。函数描述了不同变量之间的依赖关系，并且可以用来表示数学问题的模型。函数可以通过数学符号、图表或文字描述来表示，它们在代数中广泛应用于方程求解、图形绘制和数值计算等领域。

一句话概括来说，函数就是映射，输入值映射到唯一的输出值。如图 1 所示，我们设计了两个函数：左侧函数 Shape() 输入为彩色几何图形，函数输出为图形形状；右侧函数 Color() 输入还是彩色几何形状，函数输出为图形颜色。

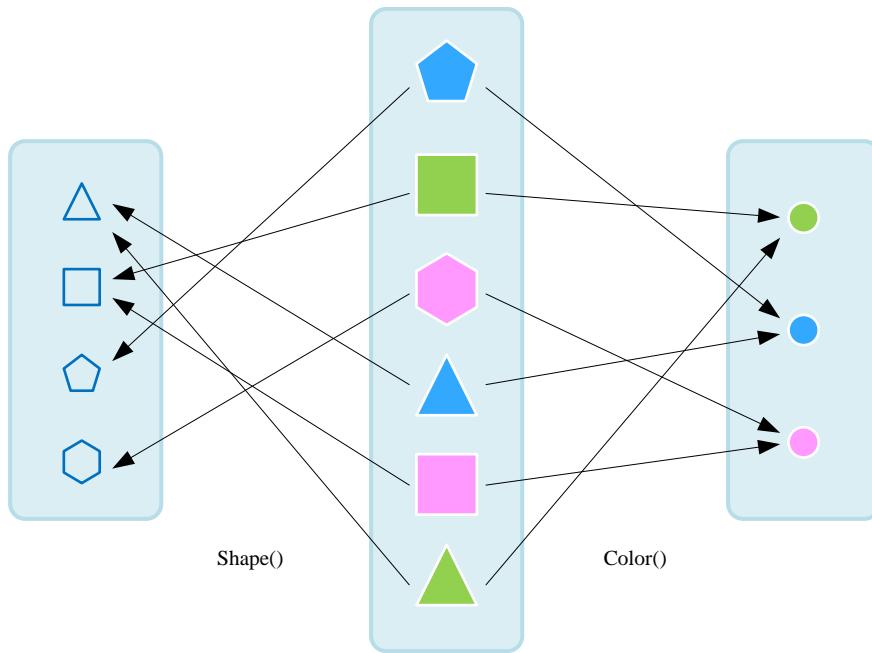


图 1. 识别颜色、形状的函数

单射、满射

单射、非单射、满射和非满射是函数映射中的性质，描述了输入值和输出值之间的关系。

单射 (injective) 是指函数中不同的输入值对应着不同的输出值，即每个输出值只有一个对应的输入值。

非单射 (non-injective) 是指函数中存在多个不同的输入值对应着相同的输出值，即至少有一个输出值有多个对应的输入值。

满射 (surjective) 是指函数的所有可能的输出值都能够被映射到，即每个输出值都有至少一个对应的输入值。

非满射 (non-surjective) 是指函数中存在至少一个输出值无法被映射到，即存在某些输出值没有对应的输入值。

图 2 所示为单射、非单射、满射、非满射构成的“四象限”。单射、非单射更关注输入值，而满射、非满射则关注输出值。同时满足单射与满射叫双射 (bijective)，也叫一一映射。

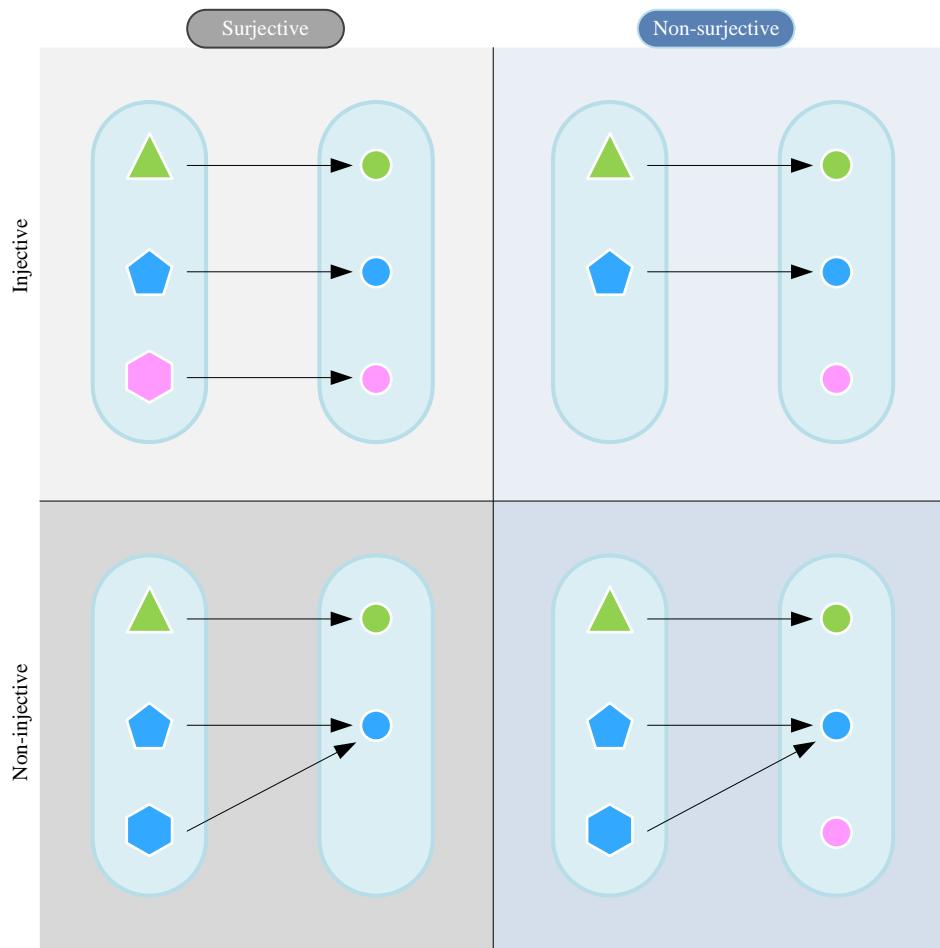


图 2. 单射、非单射、满射、非满射构成的四象限

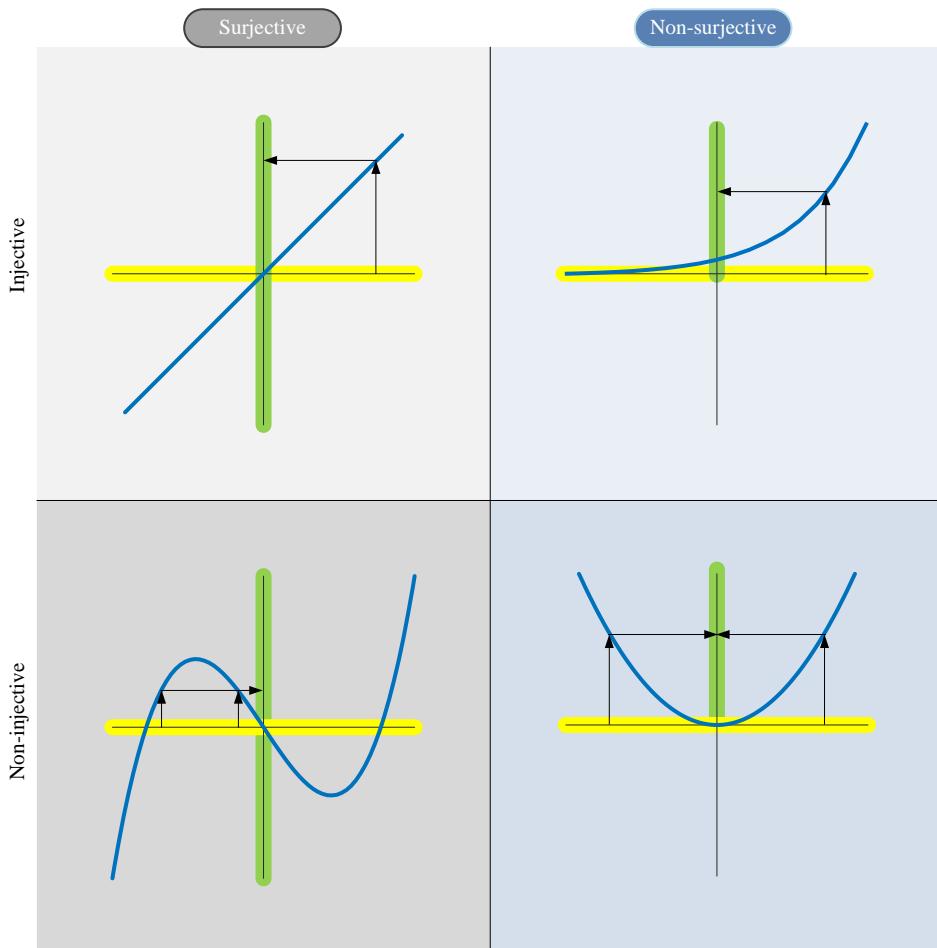


图 3. 单射、非单射、满射、非满射构成的四象限，具体实例

一元、二元、三元、多元

在数学中，函数的元 (arity) 指的是函数接受的参数个数。

常见的函数元数包括：

一元函数 (unary function) 接受一个参数。例如， $f_1(x) = x$ 是一个一元函数，它接受一个参数 x 。

二元函数 (binary function) 接受两个参数。例如， $f_2(x_1, x_2) = x_1 + x_2$ 是一个二元函数，它接受两个参数 x_1 和 x_2 。

三元函数 (ternary function) 接受三个参数。例如， $f_3(x_1, x_2, x_3) = x_1 + x_2 + x_3$ 是一个三元函数，它接受三个参数 x_1 、 x_2 和 x_3 。

多元函数 (n -ary function) 接受 n 个参数。多元函数的参数个数可以是任意多个，例如 $f_n(x_1, x_2, \dots, x_n) = x_1 + x_2 + \dots + x_n$ 是一个多元函数，它接受任意 n 个参数 x_1 、 x_2 、 \dots 、 x_n 。

下面我们就来聊聊如何利用本书前文介绍的各种可视化方案展示一元、二元、三元函数。

30.2 几个函数例子

一次函数

一元一次函数是一种形式为 $f(x) = ax + b$ 的函数， a 和 b 是实数常数。它是最简单的线性函数，描述了一个直线的斜率和截距。图 4、图 5 所示为一元、二元、三元一次函数的几种可视化方案。

一次函数在数学和实际应用中非常有用。在统计学中，它被广泛用于线性回归分析，其中通过拟合一条直线来描述自变量与因变量之间的关系。这可以帮助我们预测和解释数据的变化趋势。

此外，一次函数还在分类决策面中扮演重要角色。在机器学习和模式识别中，一次函数可以用来划分特征空间，将数据点分为不同的类别。

鸢尾花书《数学要素》将专门介绍一次函数。

二次函数

一元二次函数是一种形式为 $f(x) = ax^2 + bx + c$ 的函数，其中 a 、 b 、 c 是实数常数且 a 不为零。图 6、图 7 所示为一元、二元、三元二次函数的几种可视化方案。

二次函数在数学和实际应用中有广泛的应用。在数学领域，二次函数是研究代数和解析几何的重要对象，它们具有丰富的性质和特征。

在实际应用中，二次函数可以用于建模和预测各种现象。例如，在物理学中，二次函数可以描述自由落体运动的高度随时间变化的关系。此外，二次函数还用于优化问题和曲线拟合。例如，在最优化领域，通过分析二次函数的性质，可以找到最小值或最大值的位置；在曲线拟合中，可以利用二次函数来逼近实际数据，并进行预测和插值。

鸢尾花书《数学要素》将专门介绍二次函数。

绝对值函数

绝对值函数是一种形式为 $f(x) = a|x - b|$ 的函数， a 和 b 是实数常数。图 8、图 9 所示为一元、二元、三元绝对值函数的几种可视化方案。

绝对值函数在数学和实际应用中都具有重要的作用。在数学中，绝对值函数是一种特殊的分段函数，具有简单的图像和性质。它在不等式和绝对值相关的问题时经常出现。

绝对值函数可以用于测量误差或距离的绝对值，使得结果不受正负号影响。在优化问题中，绝对值函数常用于表示目标函数或约束条件，帮助找到最优解。此外，绝对值函数还在统计学和数据分析中发挥作用。它可以用于计算绝对误差或残差，评估模型的拟合程度。

鸢尾花书《数学要素》将专门介绍不同绝对值函数。

高斯函数

图 10、图 11 所示为一元、二元、三元高斯函数的几种可视化方案。高斯函数常用于高斯分布，因此高斯函数在统计学和概率论中广泛应用。高斯函数在数据分析和模型拟合中非常有用。它可以用来描述数据集的分布情况，并进行概率推断和统计推断。在机器学习中，高斯函数常用于聚类算法、异常检测和生成模型。此外，高斯函数还在信号处理中广泛应用，如图像处理和滤波。它可以用来平滑信号、降噪和边缘检测。

鸢尾花书《统计至简》将专门介绍高斯分布。

拉普拉斯核函数

拉普拉斯核函数是一种常用的非线性核函数，用于支持向量机 (Support Vector Machine, SVM) 和其他机器学习算法中的非线性分类和回归任务。图 12、图 13 所示为一元、二元、三元拉普拉斯核函数的几种可视化方案。

鸢尾花书《机器学习》将专门介绍支持向量机。

拉普拉斯核函数可以将数据映射到高维特征空间，从而在低维空间中实现非线性分类。与其他核函数相比，拉普拉斯核函数在决策边界附近具有更陡峭的变化，因此对异常点更敏感，使其能够更好地捕捉数据中的局部结构。

拉普拉斯核函数在图像识别、文本分类、生物信息学等领域广泛应用。它具有良好的分类性能，并且在处理高维数据和处理小样本问题时表现出色。

逻辑函数

逻辑函数是一种常用的非线性函数，也称为 Sigmoid 函数或逻辑曲线。逻辑函数的取值范围在 0 到 1 之间，具有平滑的 S 形曲线。图 14、图 15 所示为一元、二元、三元逻辑函数的几种可视化方案。

逻辑函数在机器学习、神经网络和统计学中广泛应用。它主要用于将输入值映射到概率值，将连续变量转换为概率分布。在二元分类任务中，逻辑函数可以将输入的线性组合转化为 0 和 1 之间的概率值，用于判断样本属于某一类别的概率。

逻辑函数在逻辑回归模型中扮演重要角色，用于建立分类模型和进行概率预测。它能够对数据进行非线性建模，拟合复杂的分类决策边界。

此外，逻辑函数还用于激活函数，如在神经网络中，逻辑函数可用于将神经元的输出转化为非线性响应，提高模型的表达能力。

鸢尾花书《数据有道》将专门介绍逻辑回归。

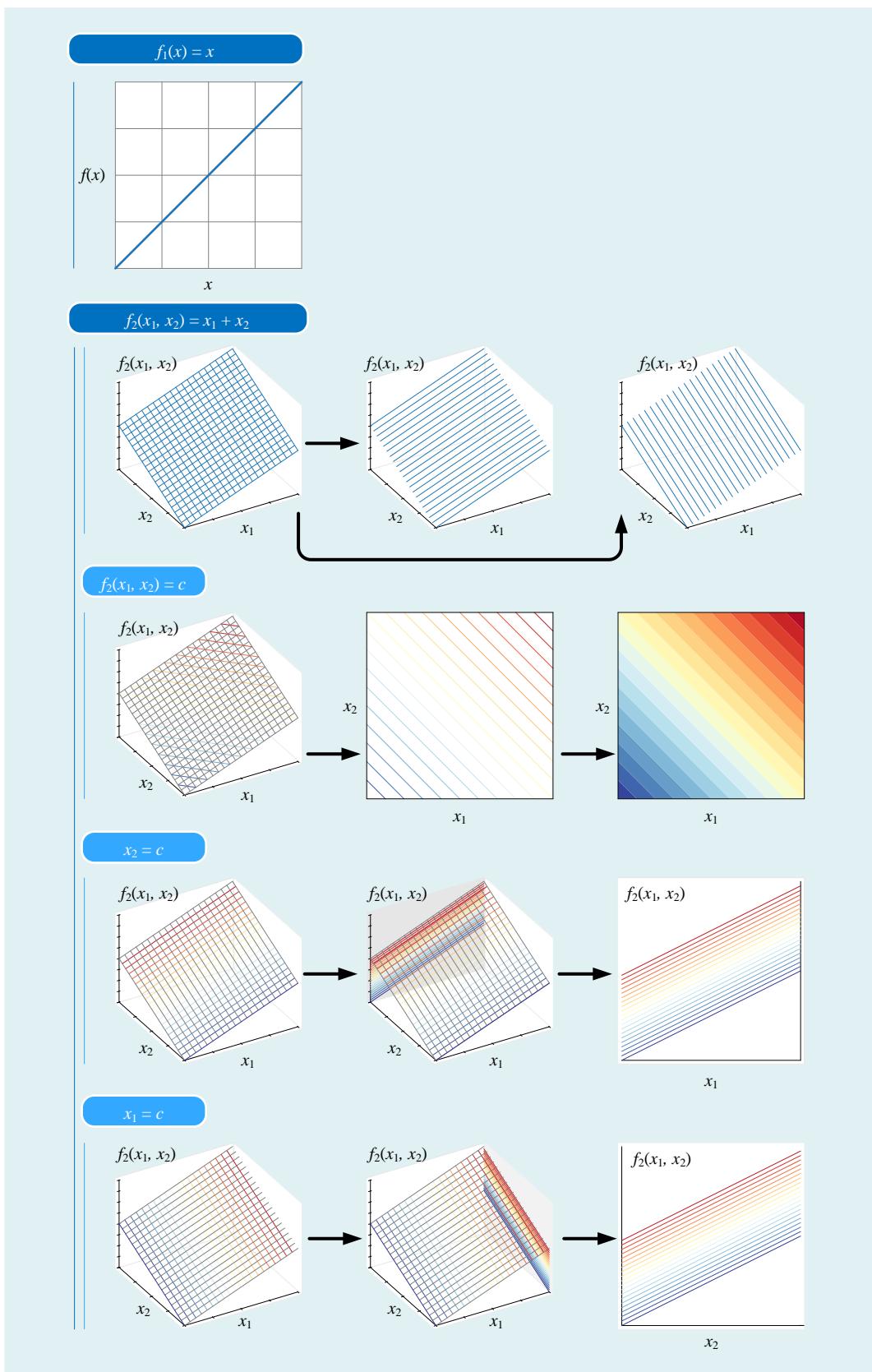


图 4. 一次函数，一元、二元

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

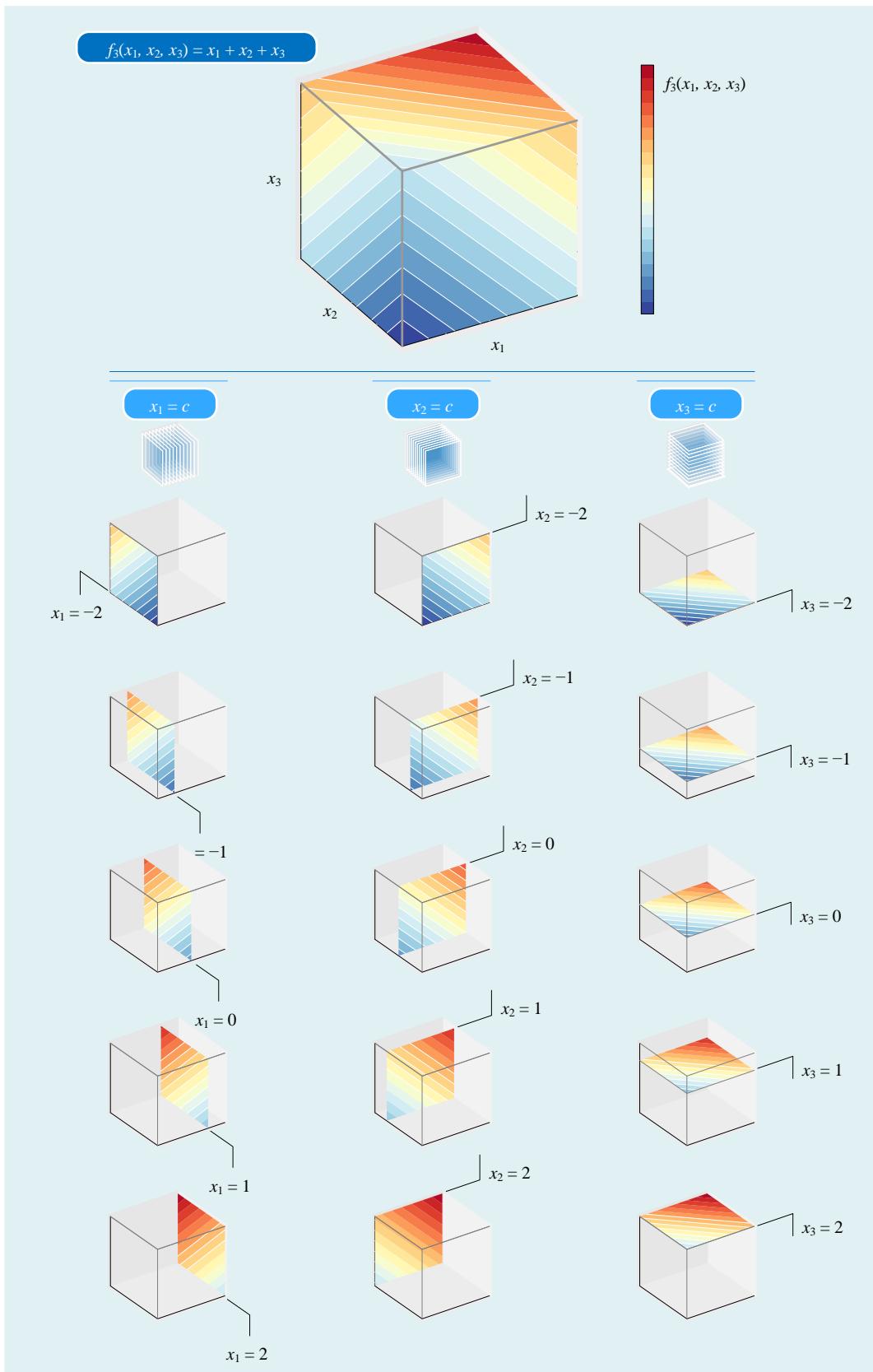


图 5. 一次函数，三元

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

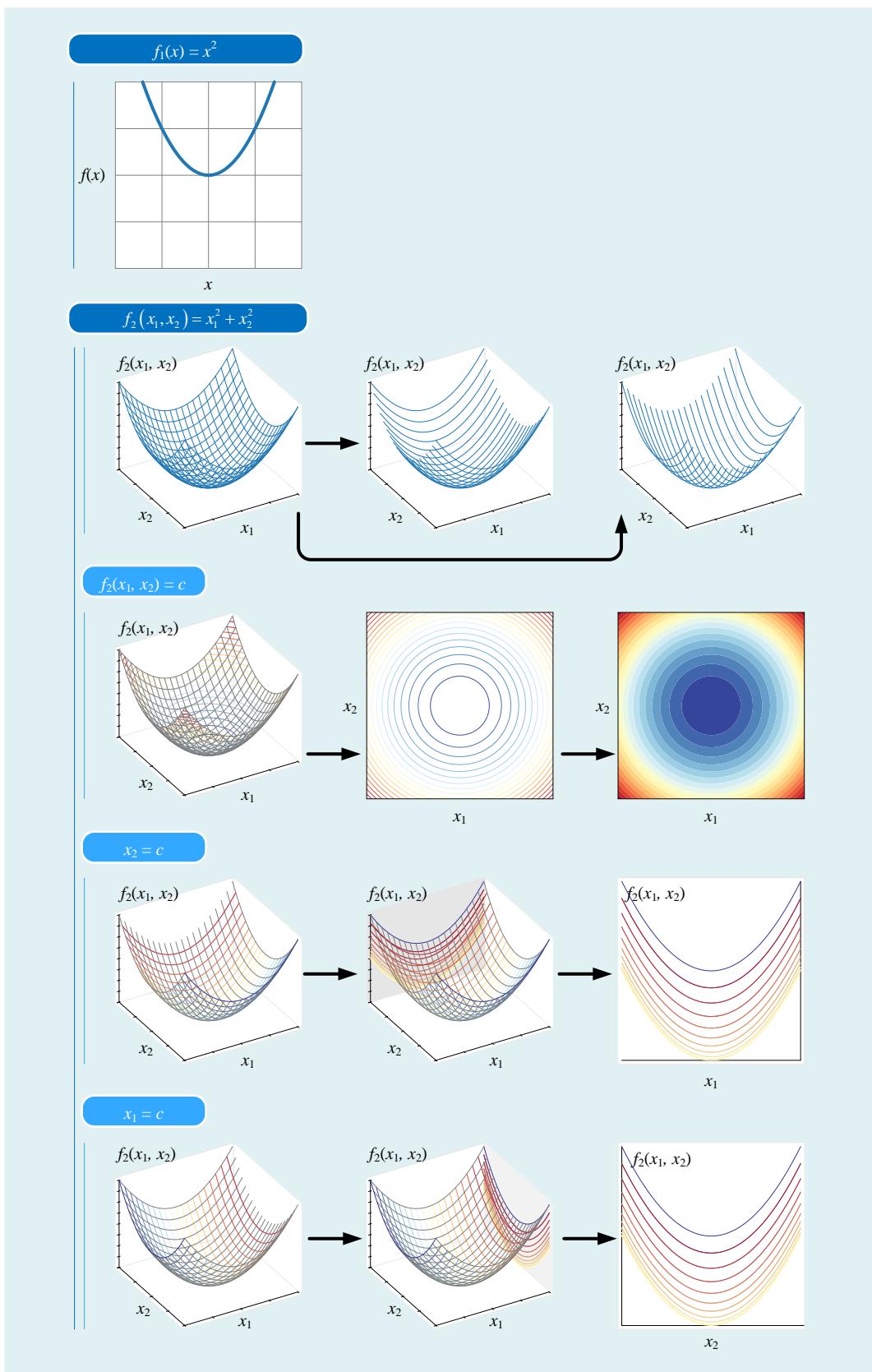


图 6. 二次函数, 一元、二元

本 PDF 文件为作者草稿, 发布目的为方便读者在移动终端学习, 终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有, 请勿商用, 引用请注明出处。

代码及 PDF 文件下载: <https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger: <https://space.bilibili.com/513194466>

欢迎大家批评指教, 本书专属邮箱: jiang.visualize.ml@gmail.com

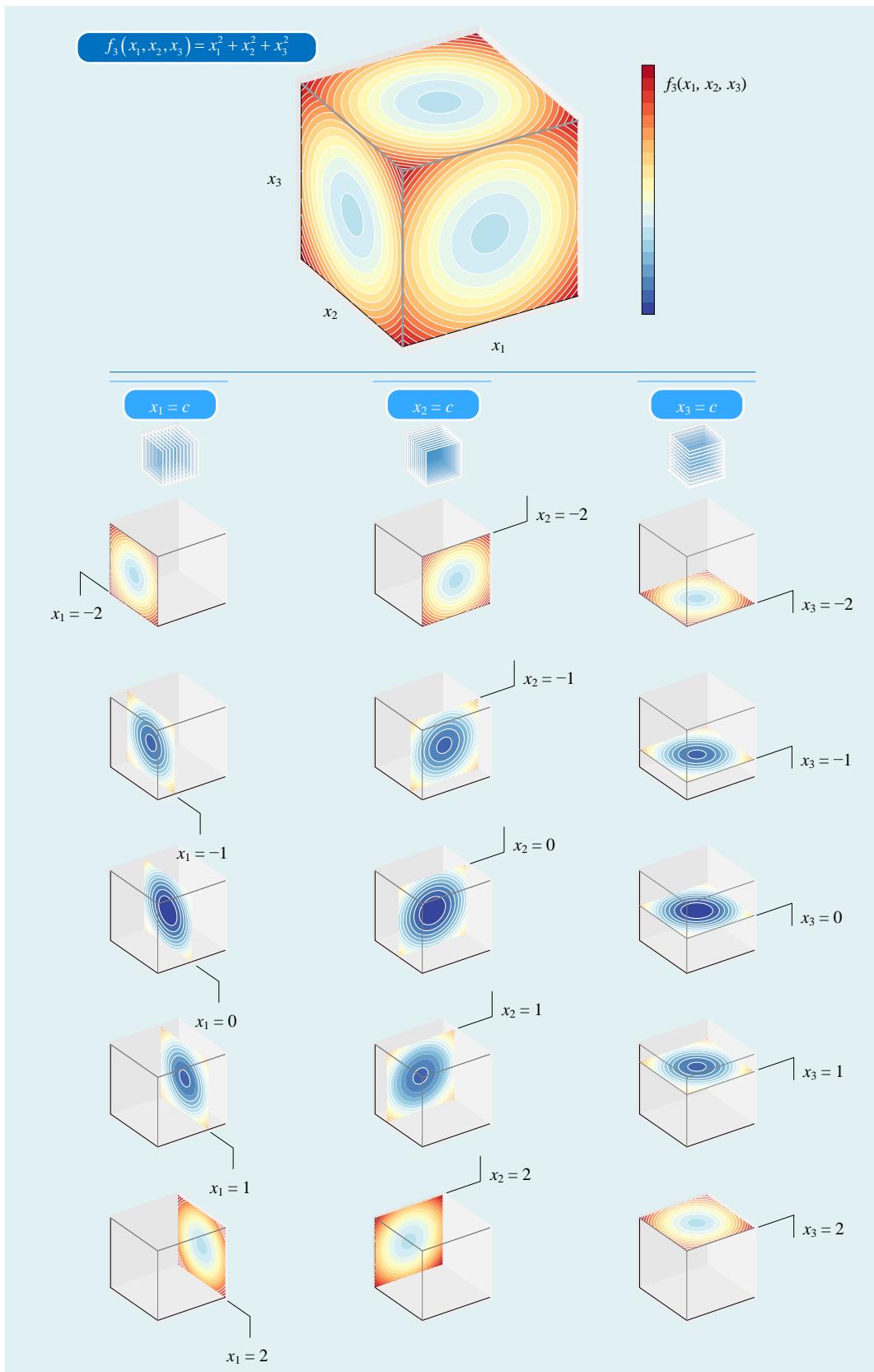


图 7. 二次函数，三元

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

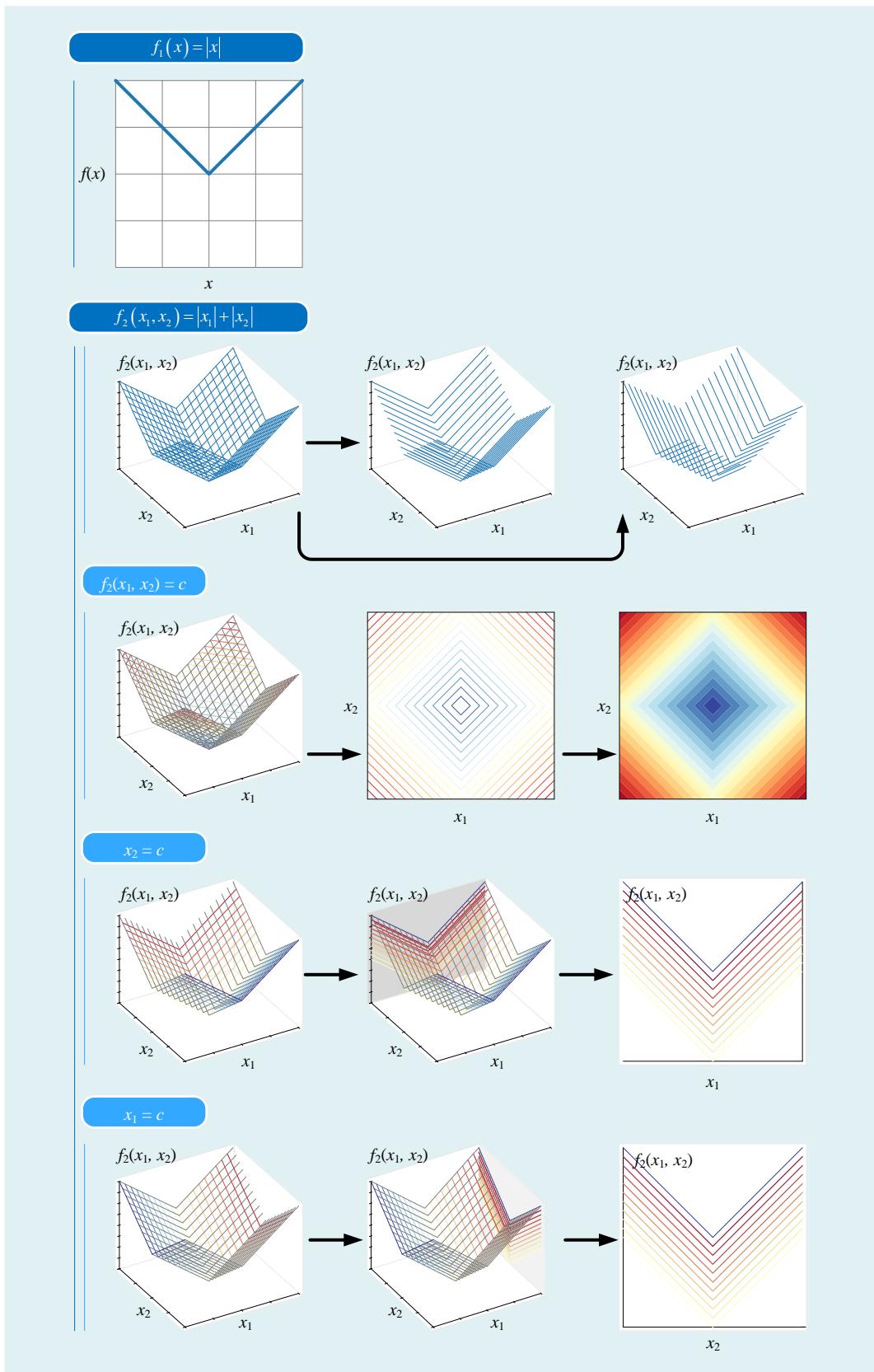


图 8. 绝对值函数，一元、二元

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

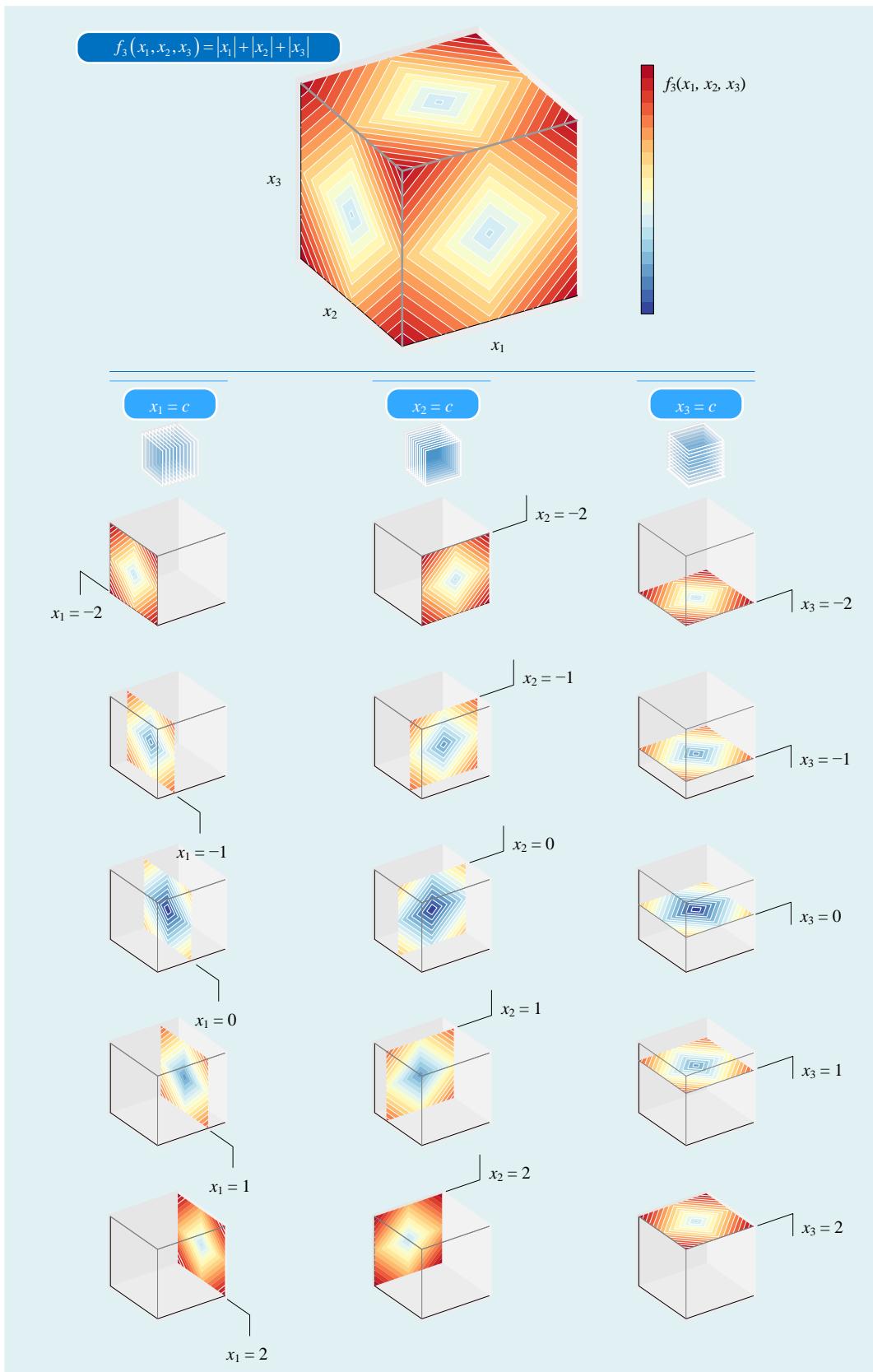


图 9. 绝对值函数，三元

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

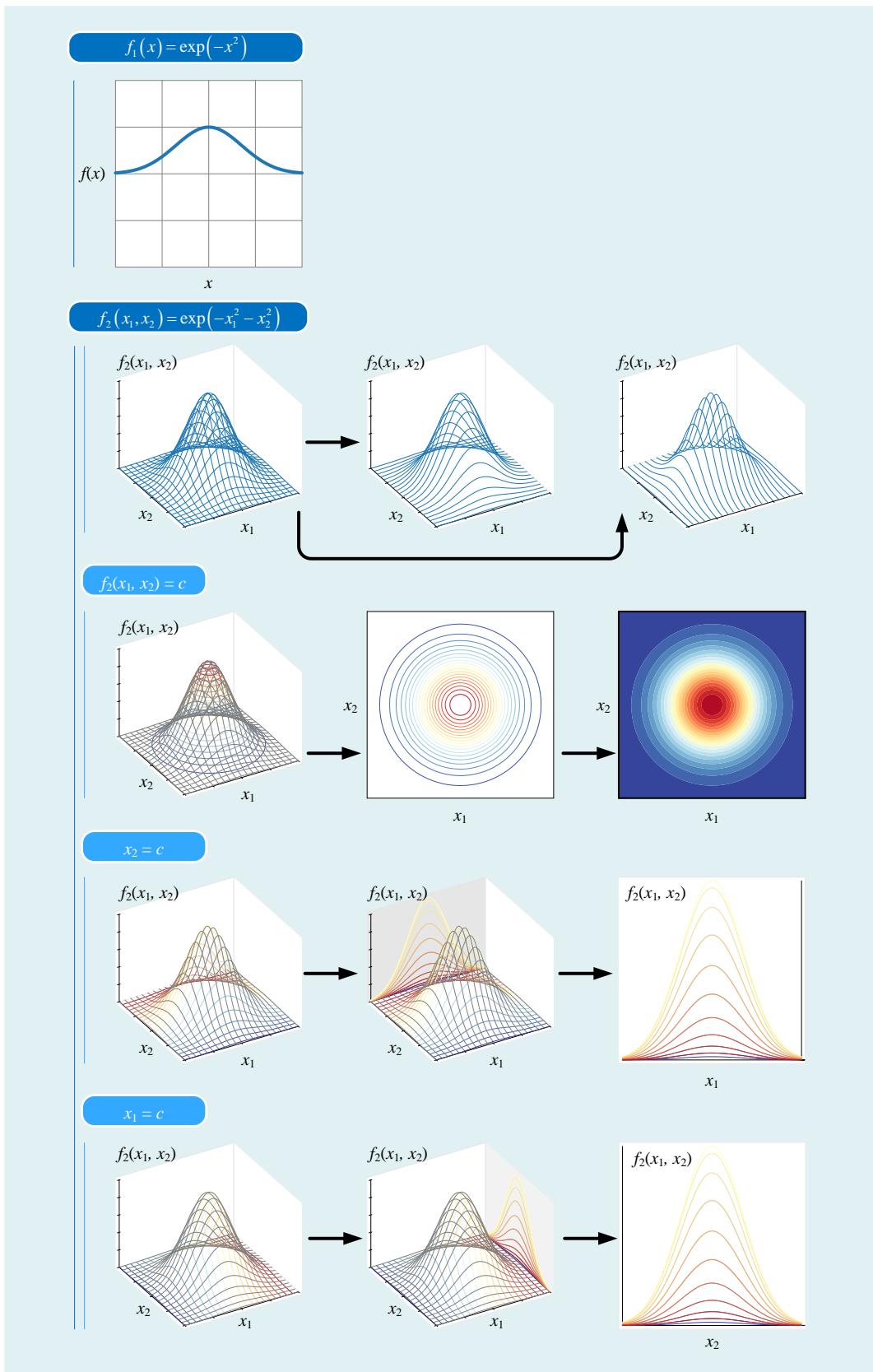


图 10. 高斯函数，一元、二元

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

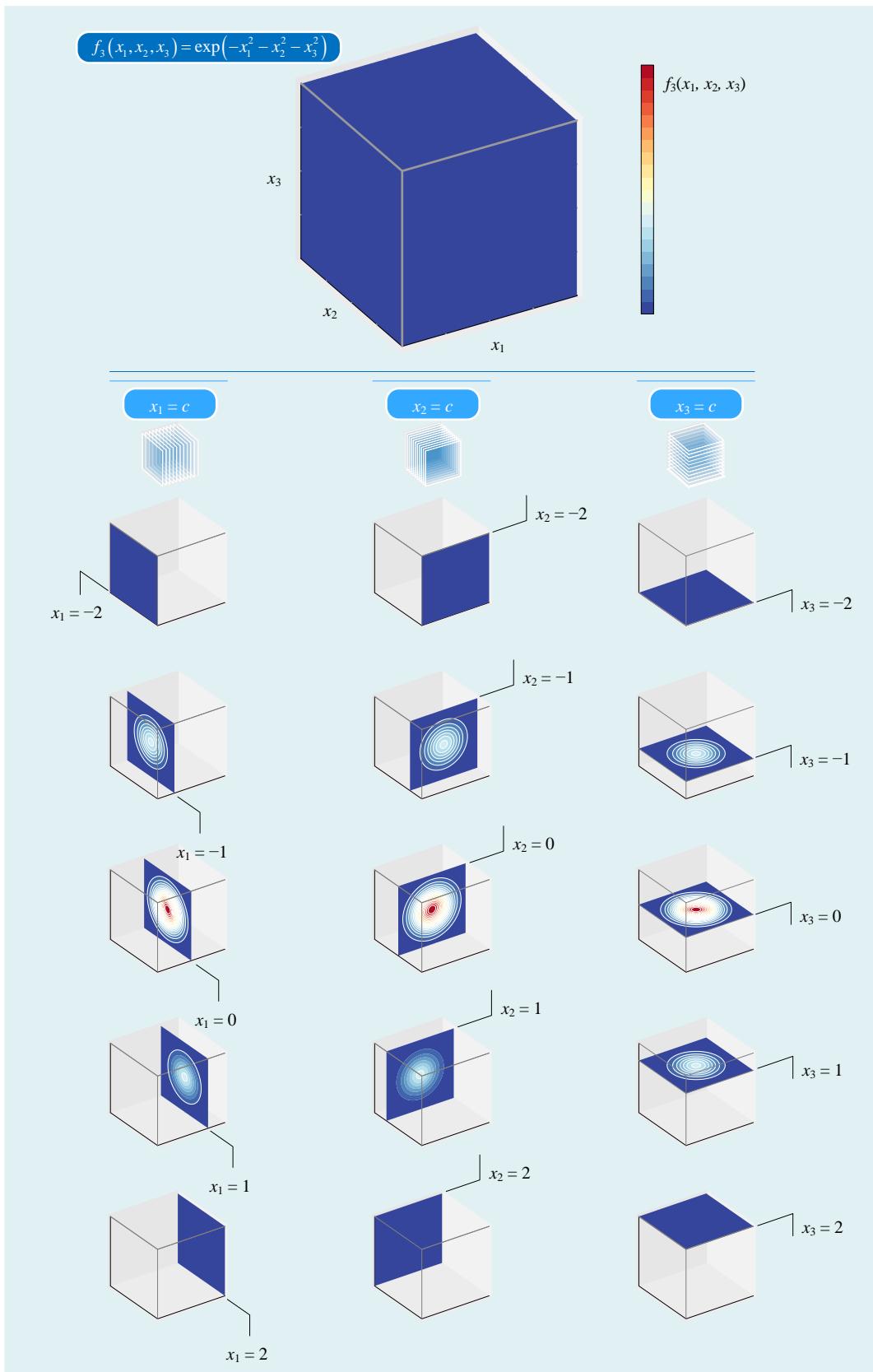


图 11. 高斯函数，三元

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

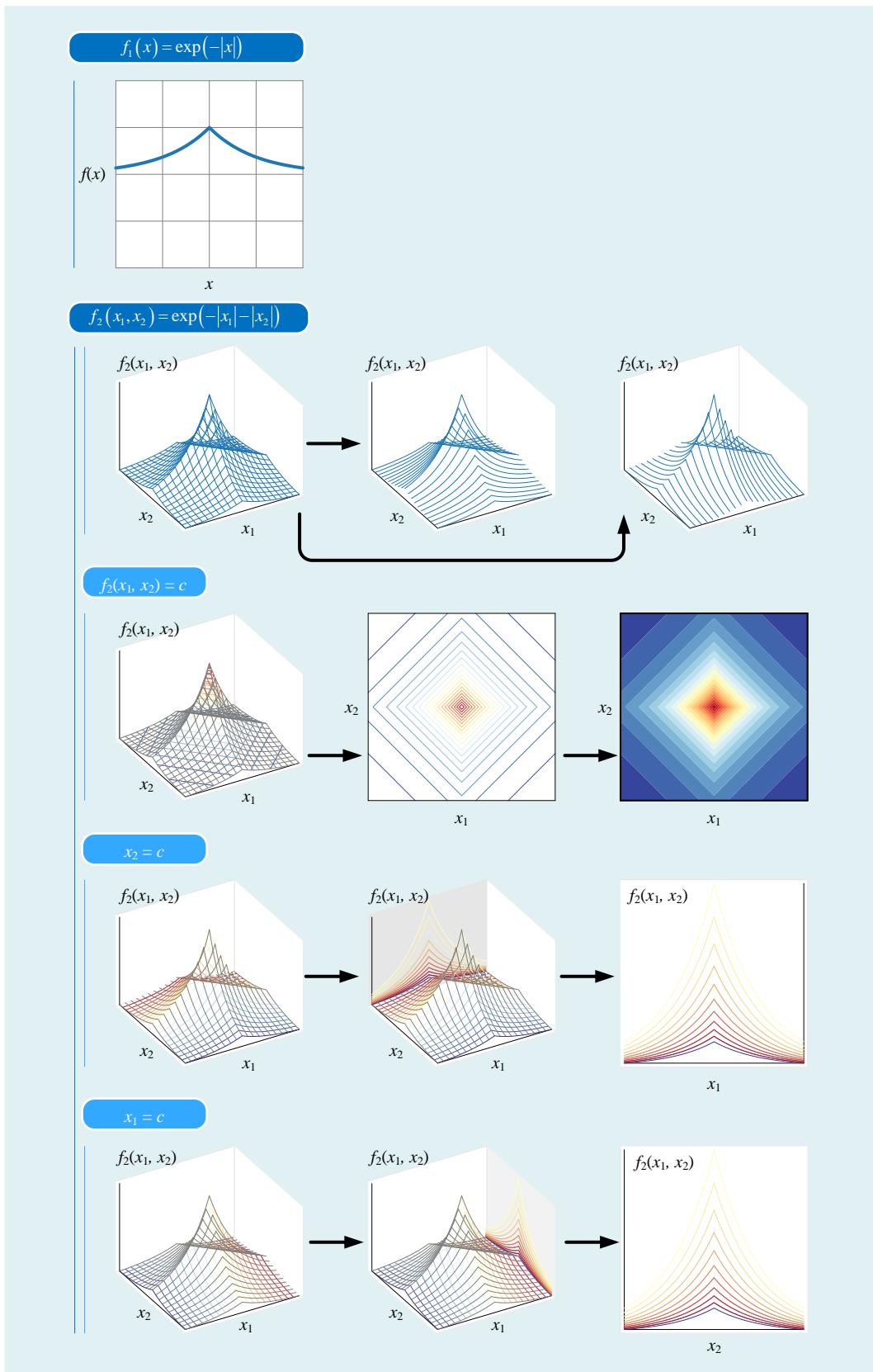


图 12. 拉普拉斯核函数，一元、二元

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

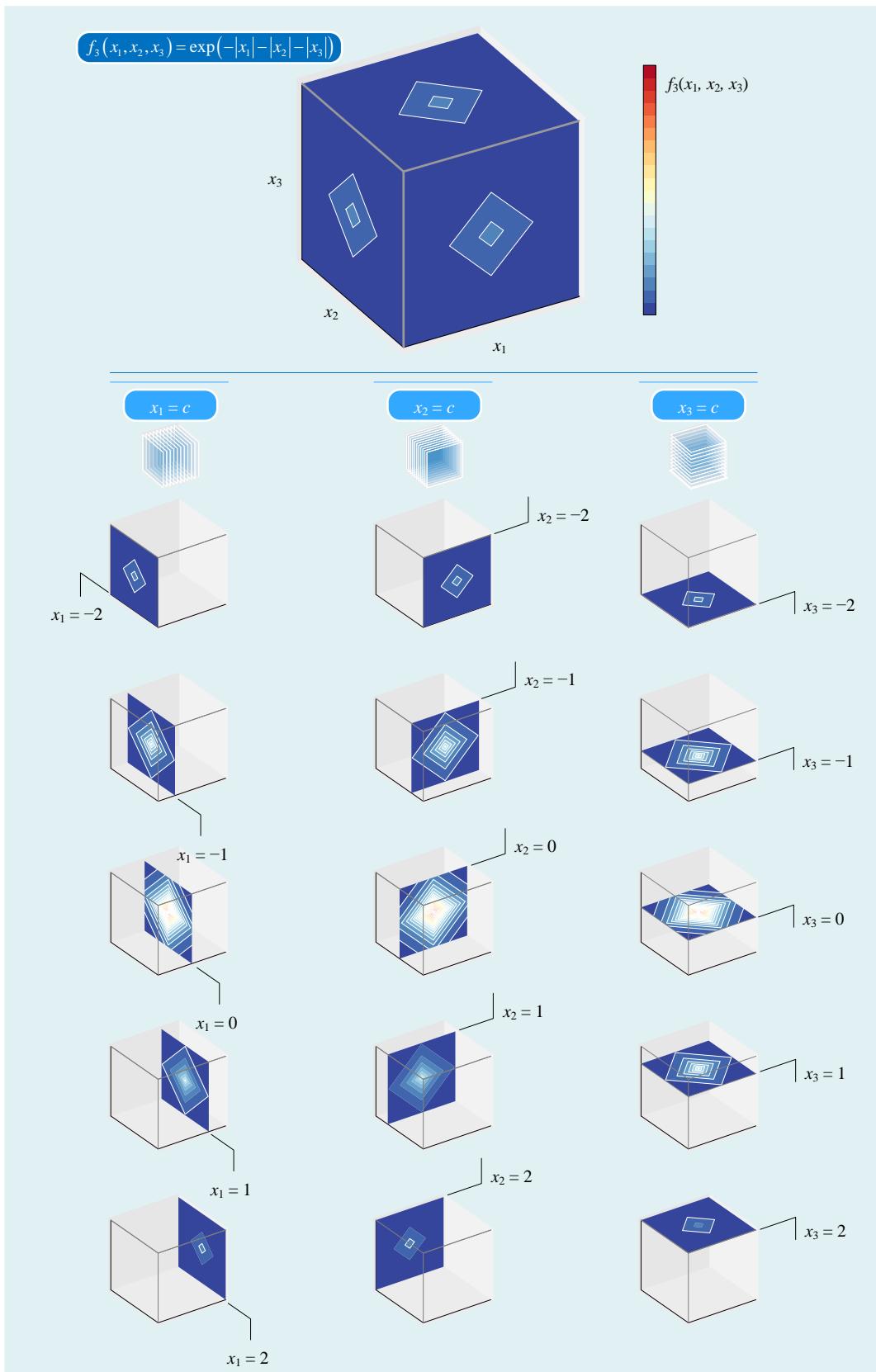


图 13. 拉普拉斯核函数，三元

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

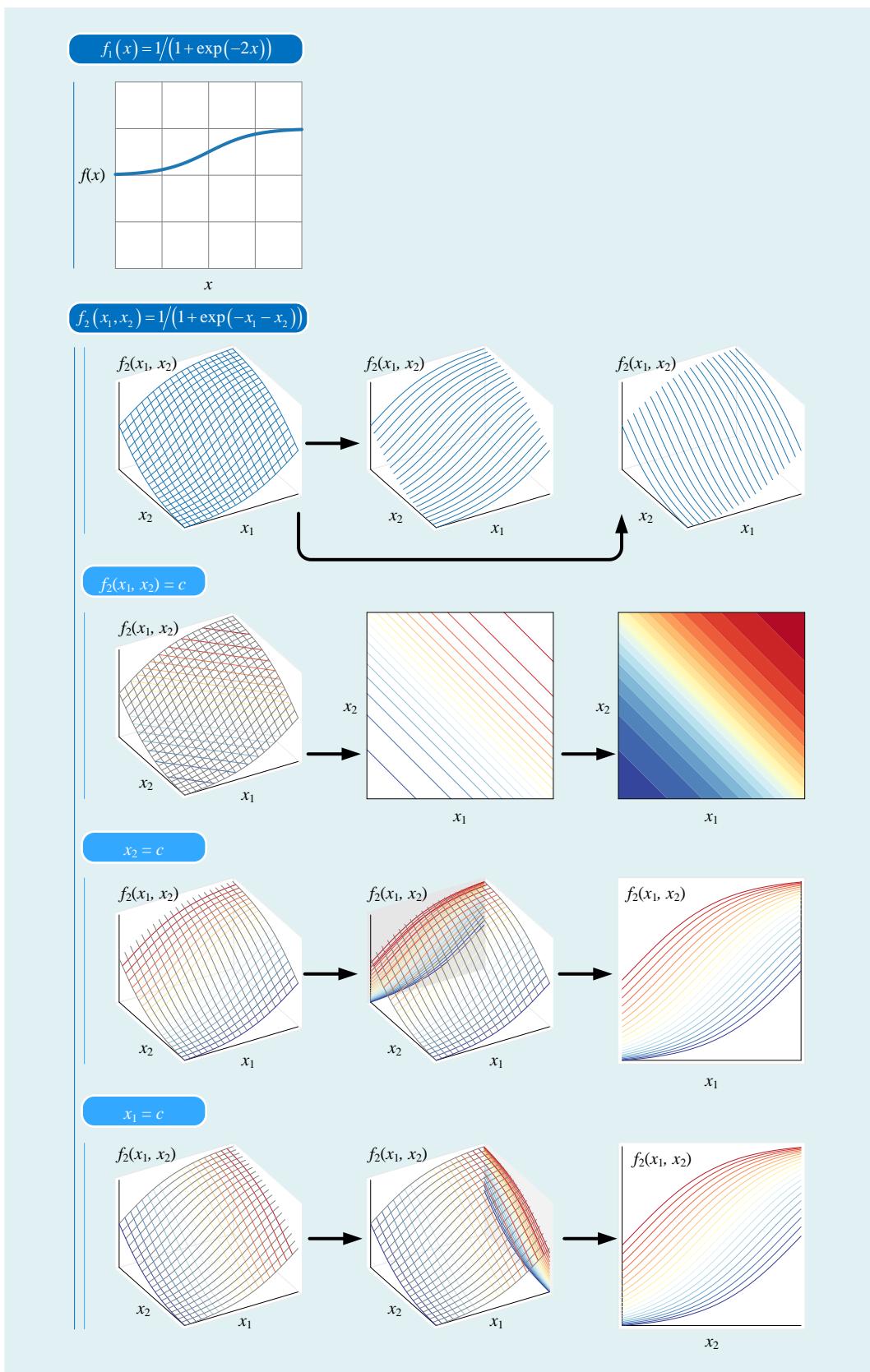


图 14. 逻辑函数，一元、二元

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

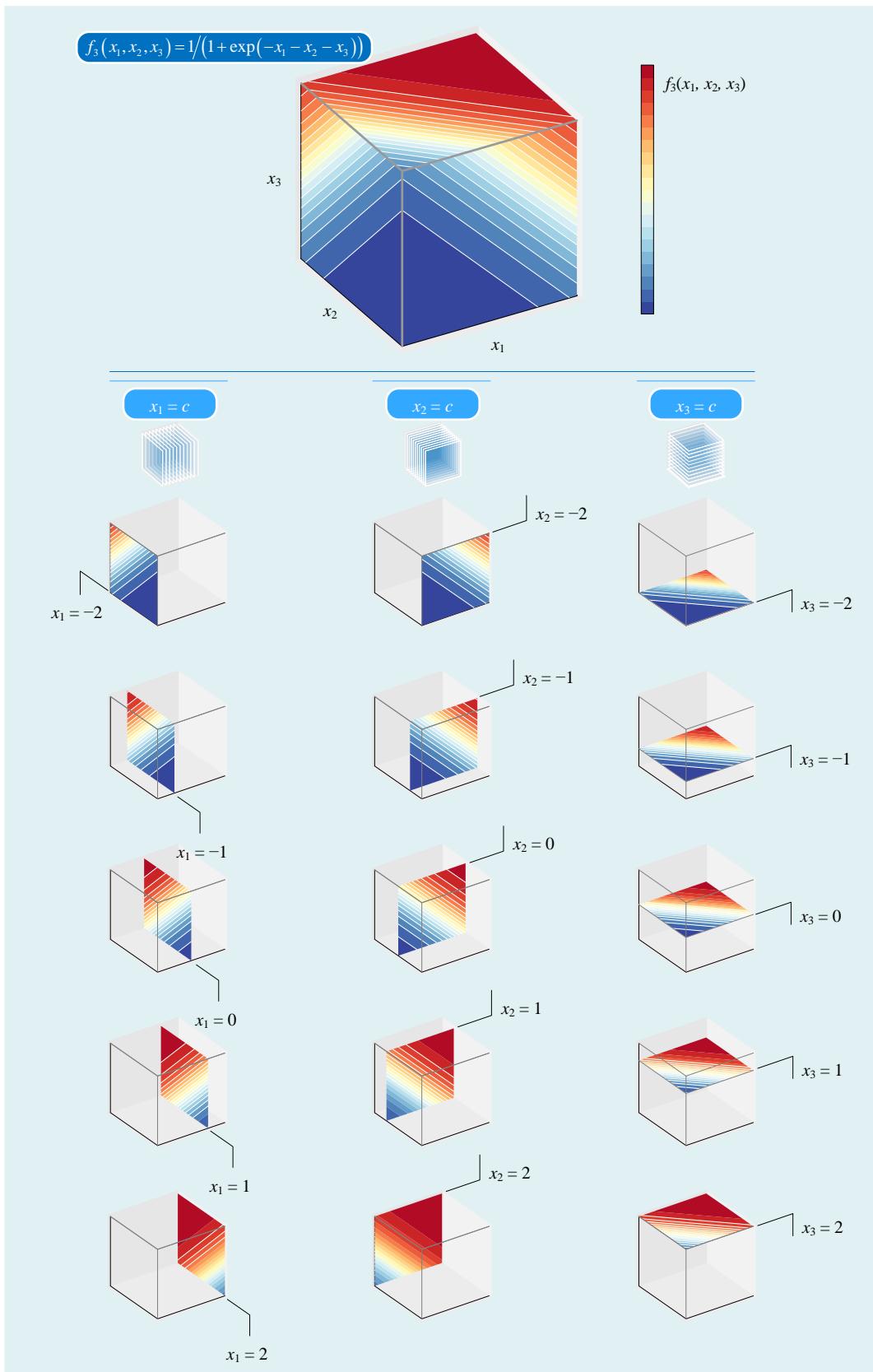


图 15. 逻辑函数，三元

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com



Quadratic Form

二次型

可视化三元二次型正定性



别怕完美，因为我们永远也达不到。

Have no fear of perfection - you'll never reach it.

—— 萨尔瓦多达利 (Salvador Dali) | 西班牙超现实主义画家 | 1904 ~ 1989



- ◀ `matplotlib.pyplot.contour()` 绘制等高线图
- ◀ `matplotlib.pyplot.contourf()` 绘制填充等高线图
- ◀ `matplotlib.pyplot.plot_wireframe()` 绘制线框图
- ◀ `matplotlib.pyplot.quiver()` 绘制箭头图
- ◀ `numpy.linspace()` 在指定的间隔内, 返回固定步长的数据
- ◀ `numpy.meshgrid()` 创建网格化数据
- ◀ `sympy.diff()` 求解符号导数和偏导解析式
- ◀ `sympy.expand()` 展开代数式
- ◀ `sympy.lambdify()` 将符号表达式转化为函数
- ◀ `sympy.simplify()` 简化代数式
- ◀ `sympy.symbols()` 定义符号变量

311 二元二次型

读过《编程不难》一册的读者对二次型 (quadratic form) 这个概念应该不陌生。

二次型是一个二次多项式函数，通常表示为 $Q(\mathbf{x}) = \mathbf{x}^T @ \mathbf{A} @ \mathbf{x}$ ，其中 \mathbf{x} 是一个向量 $\mathbf{x} = [x_1, x_2, \dots, x_n]^T$ ， $\mathbf{A}_{n \times n}$ 是一个实对称矩阵。举例来说，对于一元函数， $f(x) = x^2$ 就是一个二次型；对于二元函数， $f(x_1, x_2) = x_1 x_2$ 也是个二次型。

二次型在数学和应用领域中广泛应用，用于描述曲线拟合、优化问题等。它通过特征值和特征向量的研究揭示矩阵的性质和行为，如正定性和负定性。二次型在矩阵分析、优化方法、机器学习中具有重要意义。通过研究二次型，我们可以深入理解多元函数的形式和行为，为解决实际问题提供数学基础。

说到二次型，就必须要聊聊正定性。《编程不难》用图解法介绍过 $A_{2 \times 2}$ 正定性。下面我们逐一回顾。

正定

如图 1 所示，一个矩阵 $A_{2 \times 2}$ 是正定，意味着 $f(\mathbf{x}) = \mathbf{x}^T @ A_{2 \times 2} @ \mathbf{x}$ 是个开口朝上的抛物面，形状像是碗。除了 $(0, 0)$ ， $f(\mathbf{x}) = \mathbf{x}^T @ A_{2 \times 2} @ \mathbf{x}$ 均大于 0。 $(0, 0)$ 为最小值，图中箭头都背离 $(0, 0)$ 。

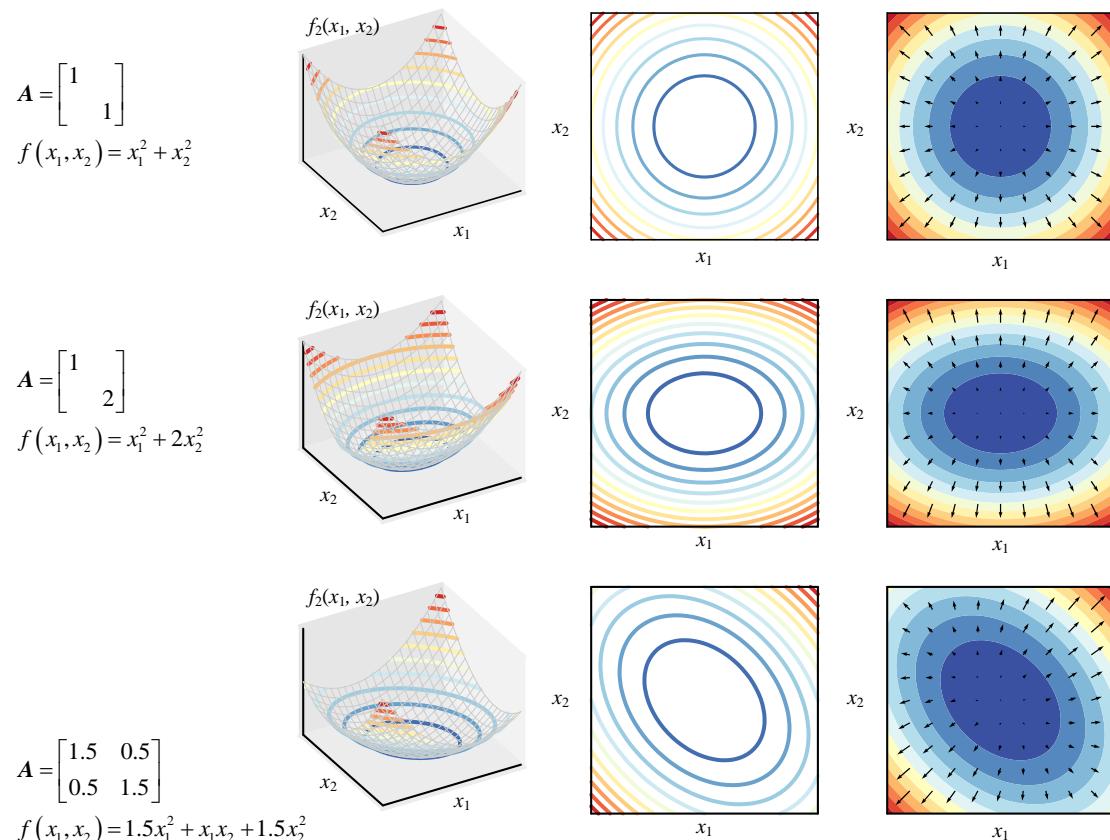


图 1. 正定， $A_{2 \times 2}$

半正定

如图2所示，一个矩阵 $A_{2 \times 2}$ 是半正定，意味着 $f(\mathbf{x}) = \mathbf{x}^T @ A_{2 \times 2} @ \mathbf{x}$ 是个开口朝上的山谷面。除了 $(0, 0)$ ， $f(\mathbf{x}) = \mathbf{x}^T @ A_{2 \times 2} @ \mathbf{x}$ 均大于等于 0。山谷的谷底都是极小值，图中箭头都背离谷底所在直线。

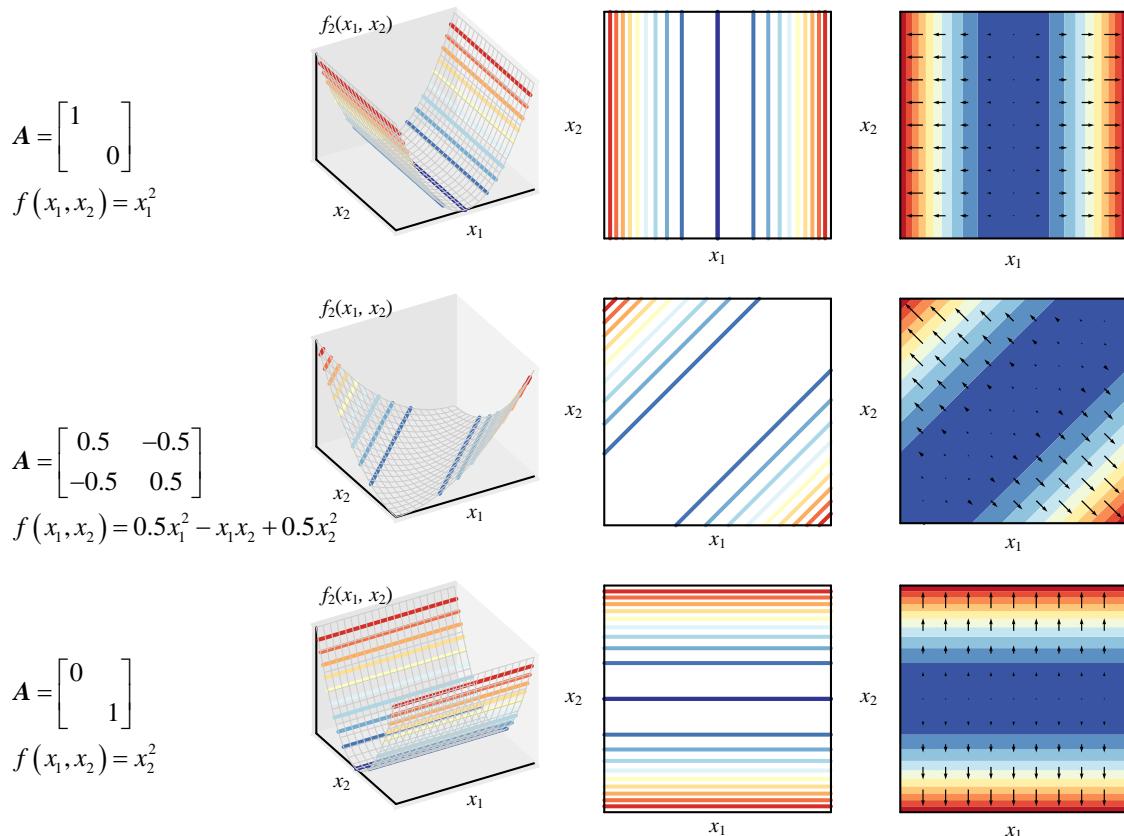
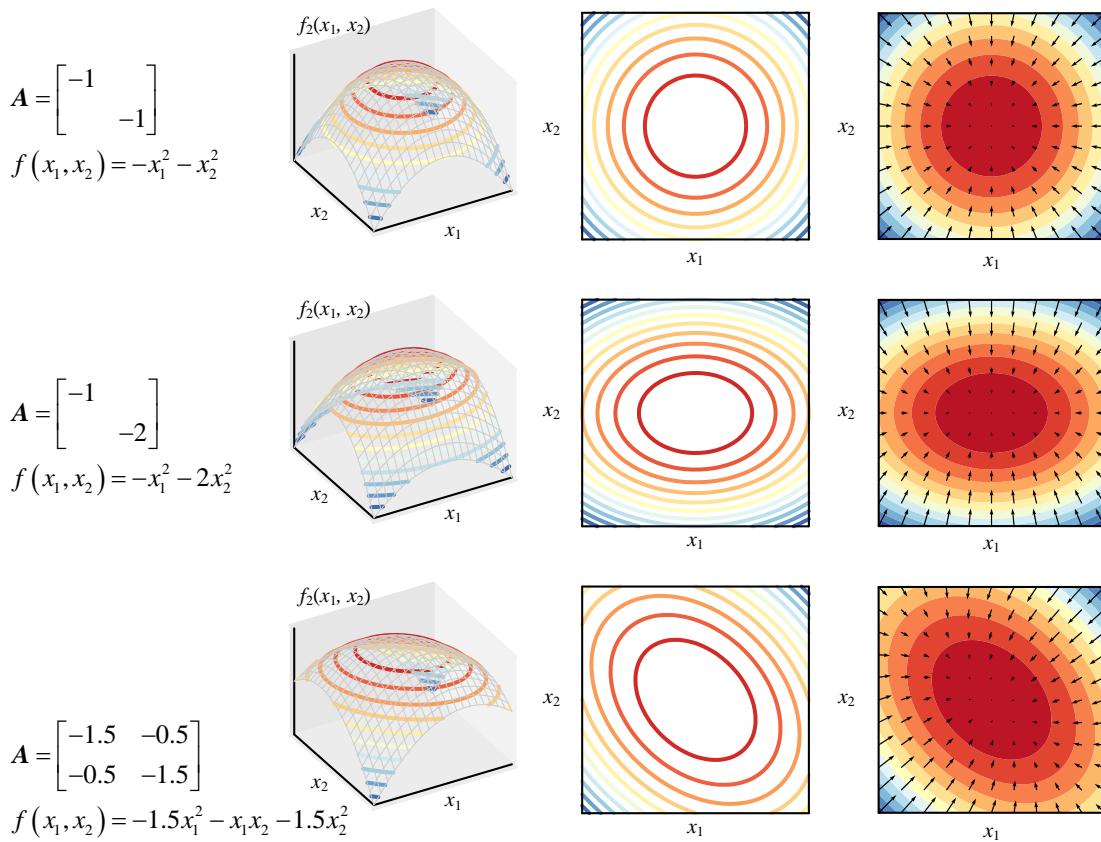


图 2. 半正定， $A_{2 \times 2}$

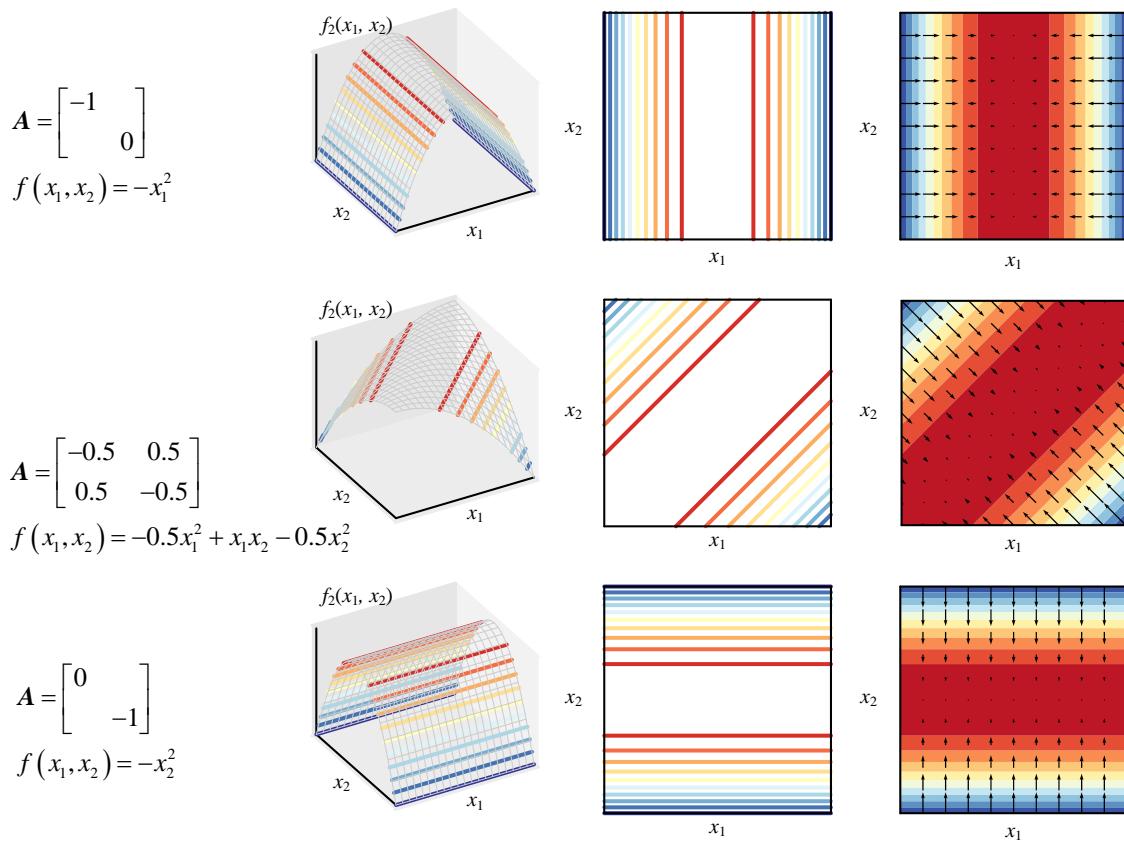
负定

如图3所示，一个矩阵 $A_{2 \times 2}$ 是负定，意味着 $f(\mathbf{x}) = \mathbf{x}^T @ A_{2 \times 2} @ \mathbf{x}$ 是个开口朝下的抛物面。除了 $(0, 0)$ ， $f(\mathbf{x}) = \mathbf{x}^T @ A_{2 \times 2} @ \mathbf{x}$ 均小于 0。 $(0, 0)$ 为最大值，图中箭头都指向 $(0, 0)$ 。

图 3. 负定, $A_{2 \times 2}$

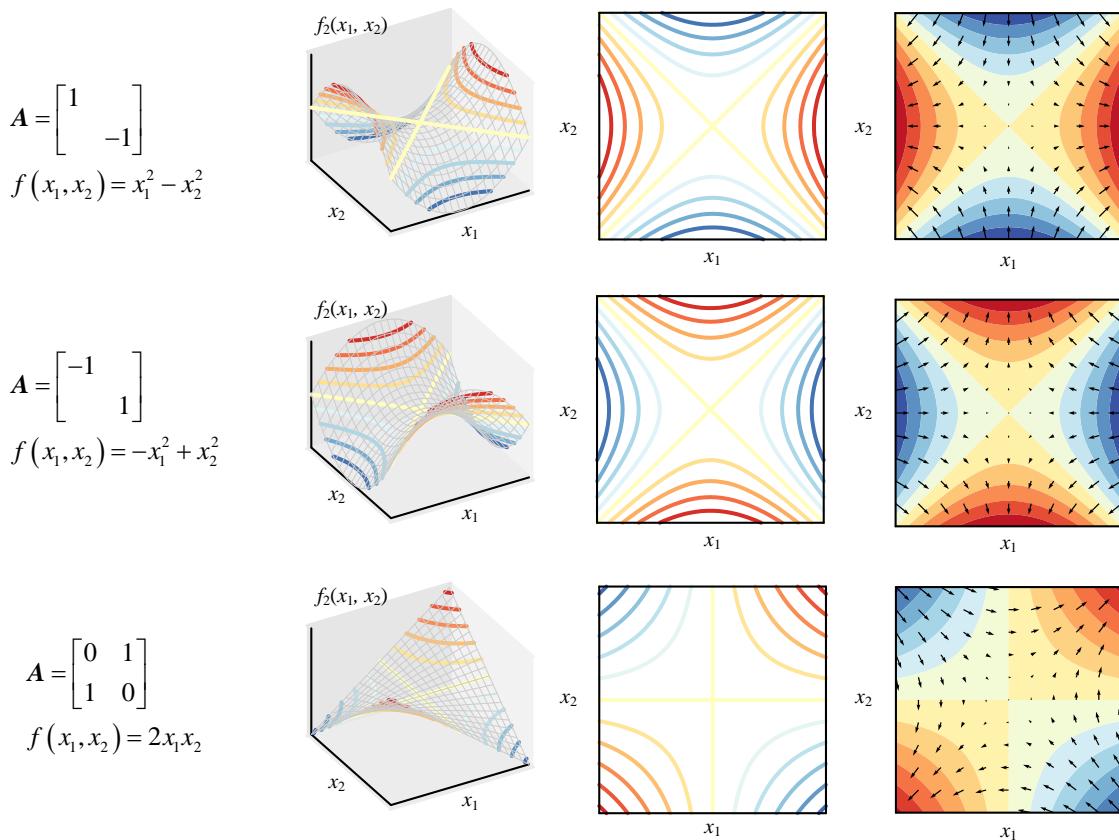
半负定

如图4所示，一个矩阵 $A_{2 \times 2}$ 是半负定，意味着 $f(\mathbf{x}) = \mathbf{x}^T @ A_{2 \times 2} @ \mathbf{x}$ 是个开口朝下的山脊面。除了 $(0, 0)$, $f(\mathbf{x}) = \mathbf{x}^T @ A_{2 \times 2} @ \mathbf{x}$ 均小于等于 0。山脊的顶端都是极大值，图中箭头指向山脊顶端所在直线。

图 4. 半负定, $A_{2\times 2}$

不定

如图4所示, 一个矩阵 $A_{2\times 2}$ 不定, 意味着 $f(\mathbf{x}) = \mathbf{x}^T @ A_{2\times 2} @ \mathbf{x}$ 是个马鞍面, $(0, 0)$ 为鞍点。 $f(\mathbf{x}) = \mathbf{x}^T @ A_{2\times 2} @ \mathbf{x}$ 符号不定。图中有些箭头背离 $(0, 0)$, 有些指向 $(0, 0)$ 。

图 5. 不定, $A_{2 \times 2}$

31.2 三元二次型

二元二次型不是本章可视化的核心，本章想要聊聊三元二次型有哪些有趣的特征。图 6 ~ 图 21 给出若干种三元二次型，请大家逐个分析。特别注意等高线变化趋势、极值位置、剖面等高线的几何形状（正圆、椭圆、旋转椭圆、平行线、抛物线、双曲线），并试着解释为什么会出现这些几何形状。

《矩阵力量》第 21 章将专门讲解正定性。

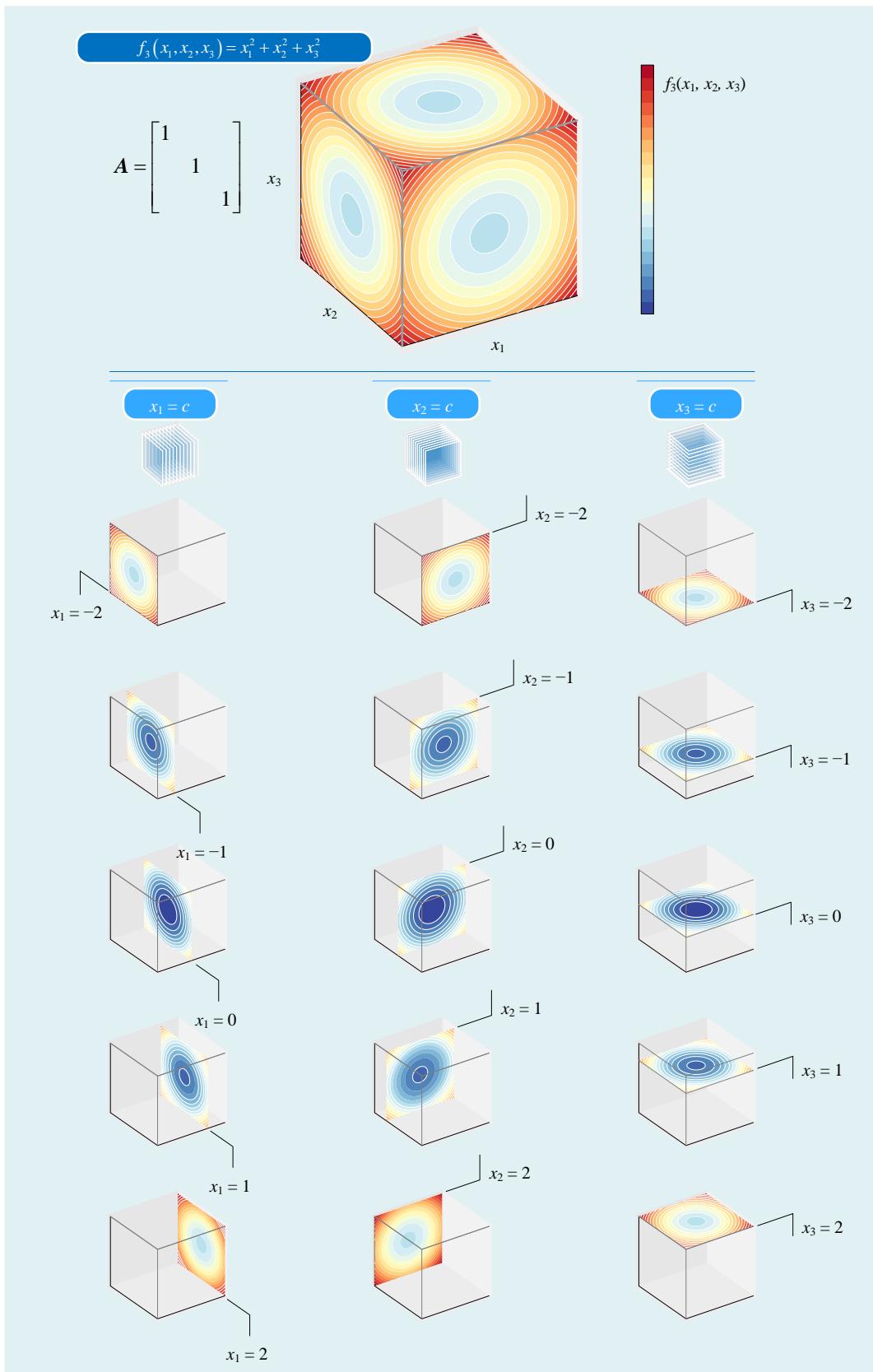


图 6. 三元二次型, 正定, 情况 A

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

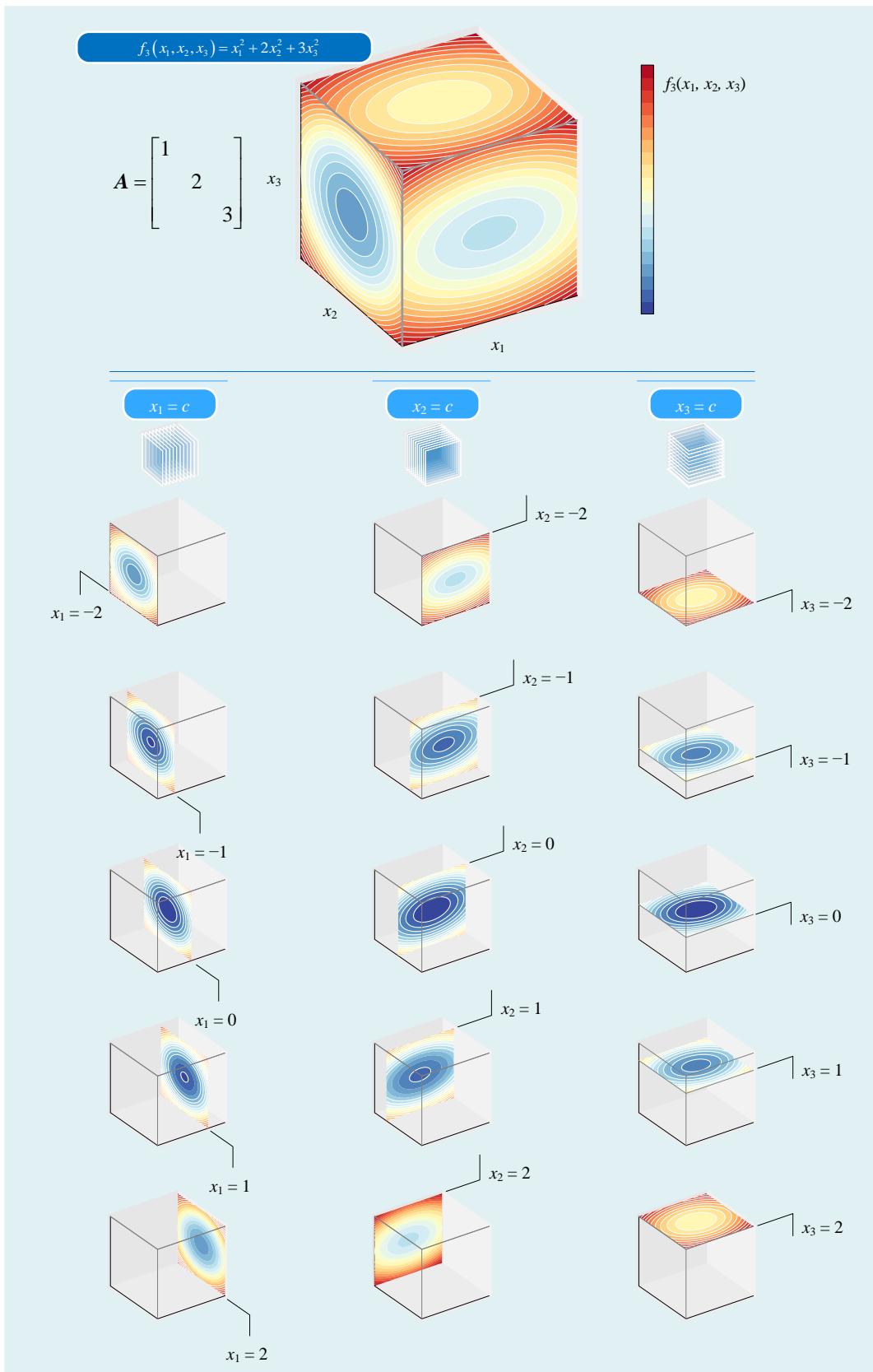


图 7. 三元二次型，正定，情况 B

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

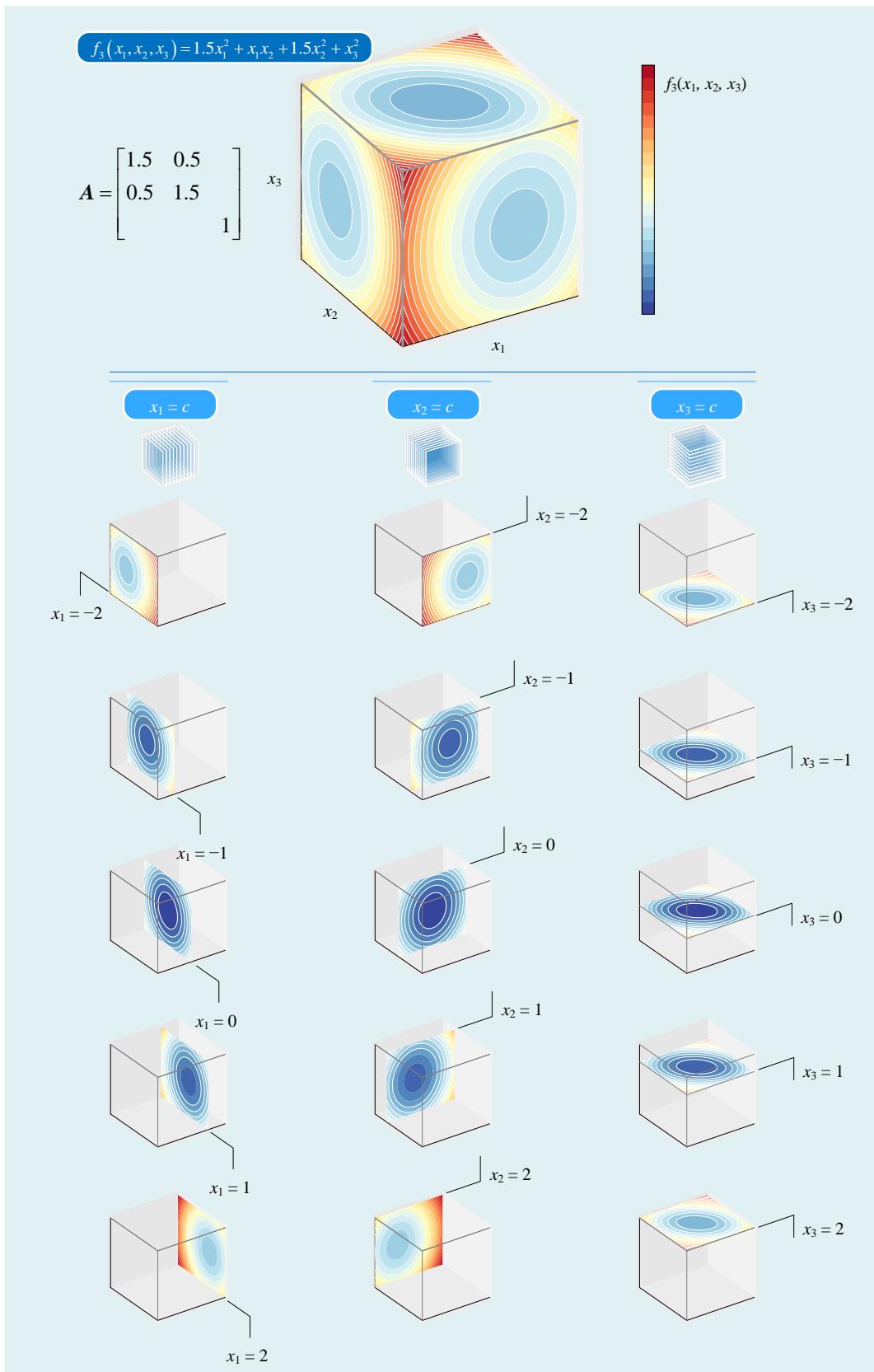


图 8. 三元二次型，正定，情况 C

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

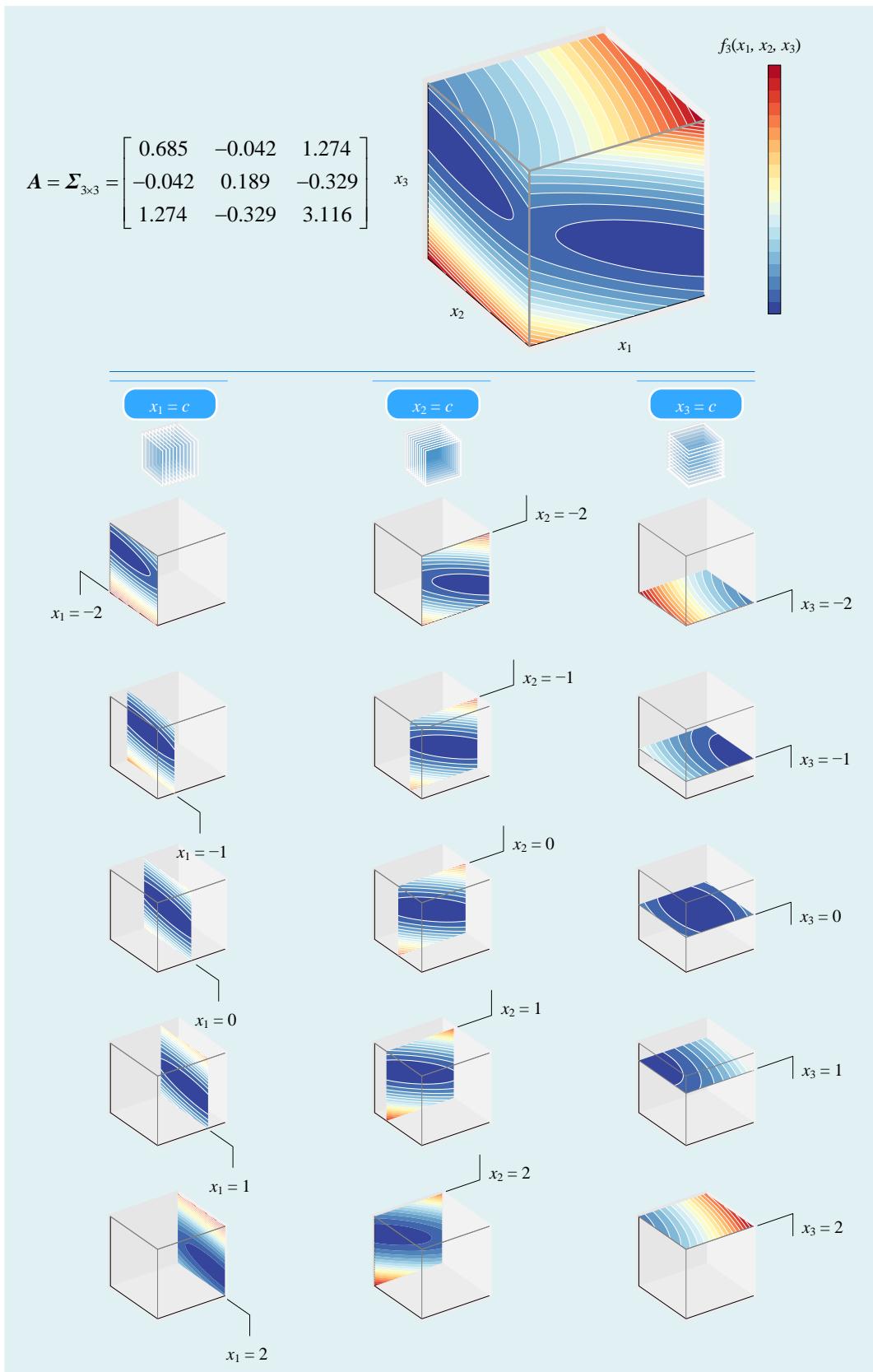


图 9. 三元二次型, 正定, 情况 D, (鸢尾花数据前三个特征的协方差矩阵)

本 PDF 文件为作者草稿, 发布目的为方便读者在移动终端学习, 终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有, 请勿商用, 引用请注明出处。

代码及 PDF 文件下载: <https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger: <https://space.bilibili.com/513194466>

欢迎大家批评指教, 本书专属邮箱: jiang.visualize.ml@gmail.com

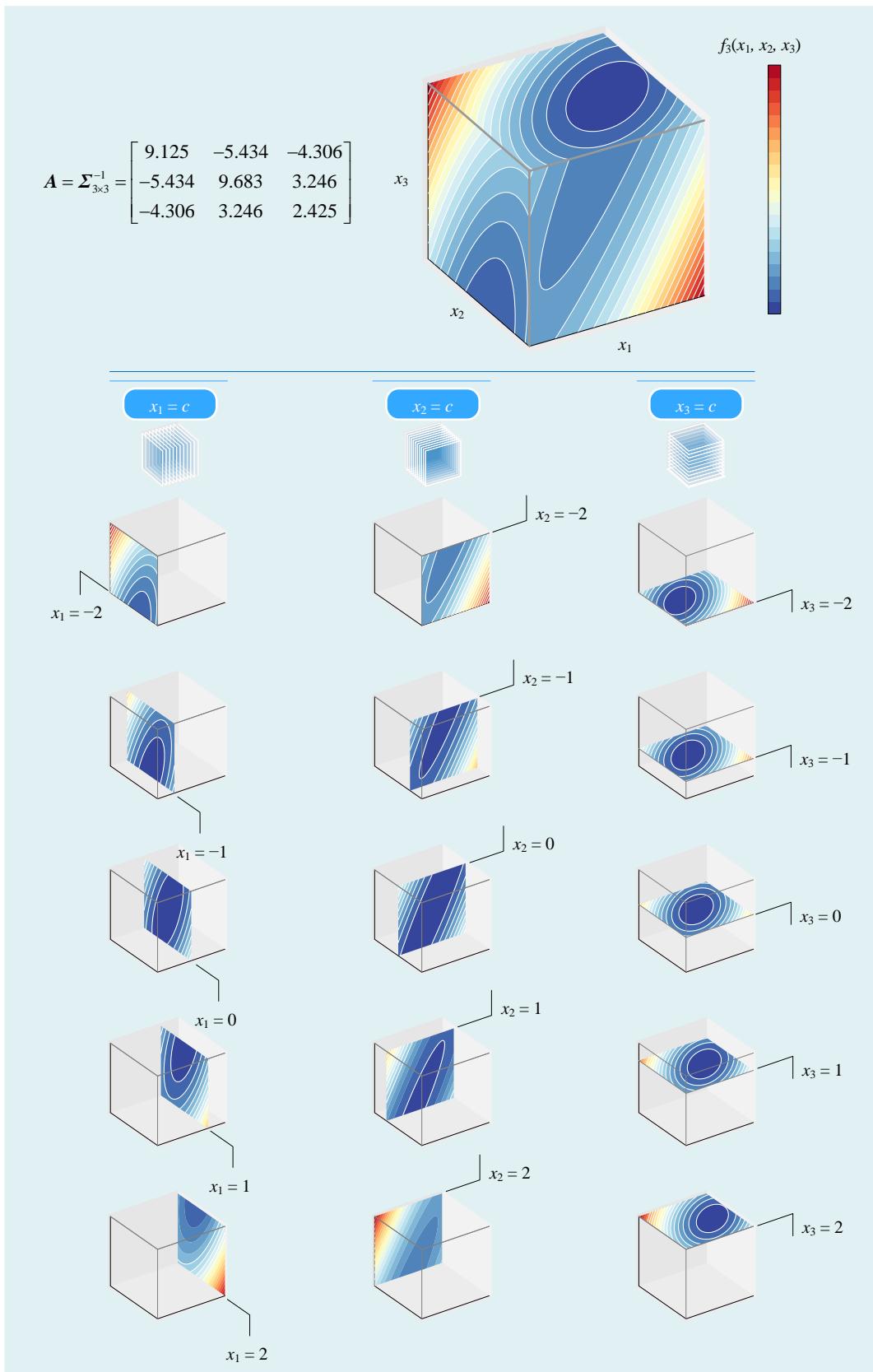


图 10. 三元二次型, 正定, 情况 E, (鸢尾花数据前三个特征的协方差矩阵逆矩阵)

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

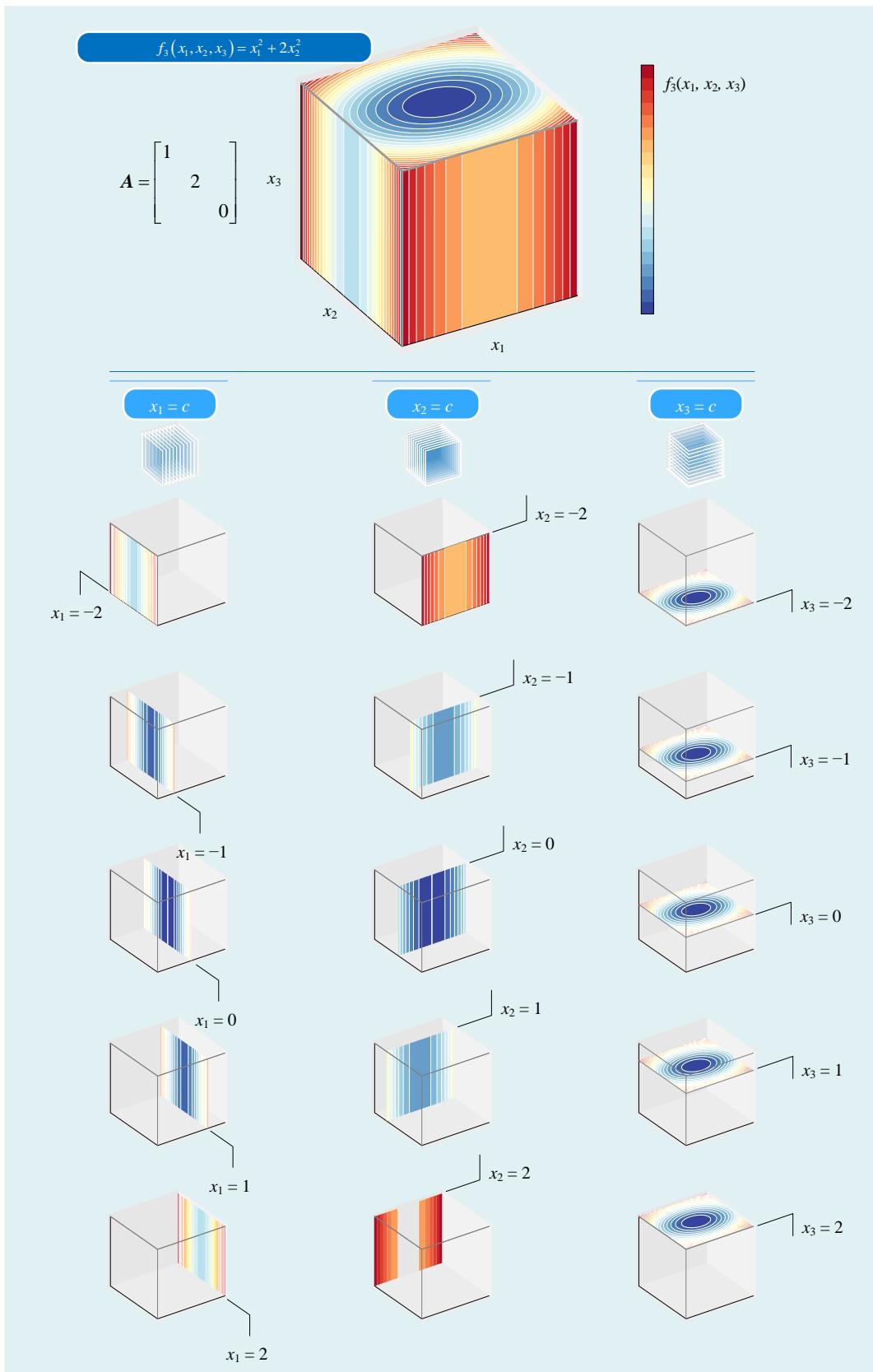


图 11. 三元二次型，半正定，情况 F

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

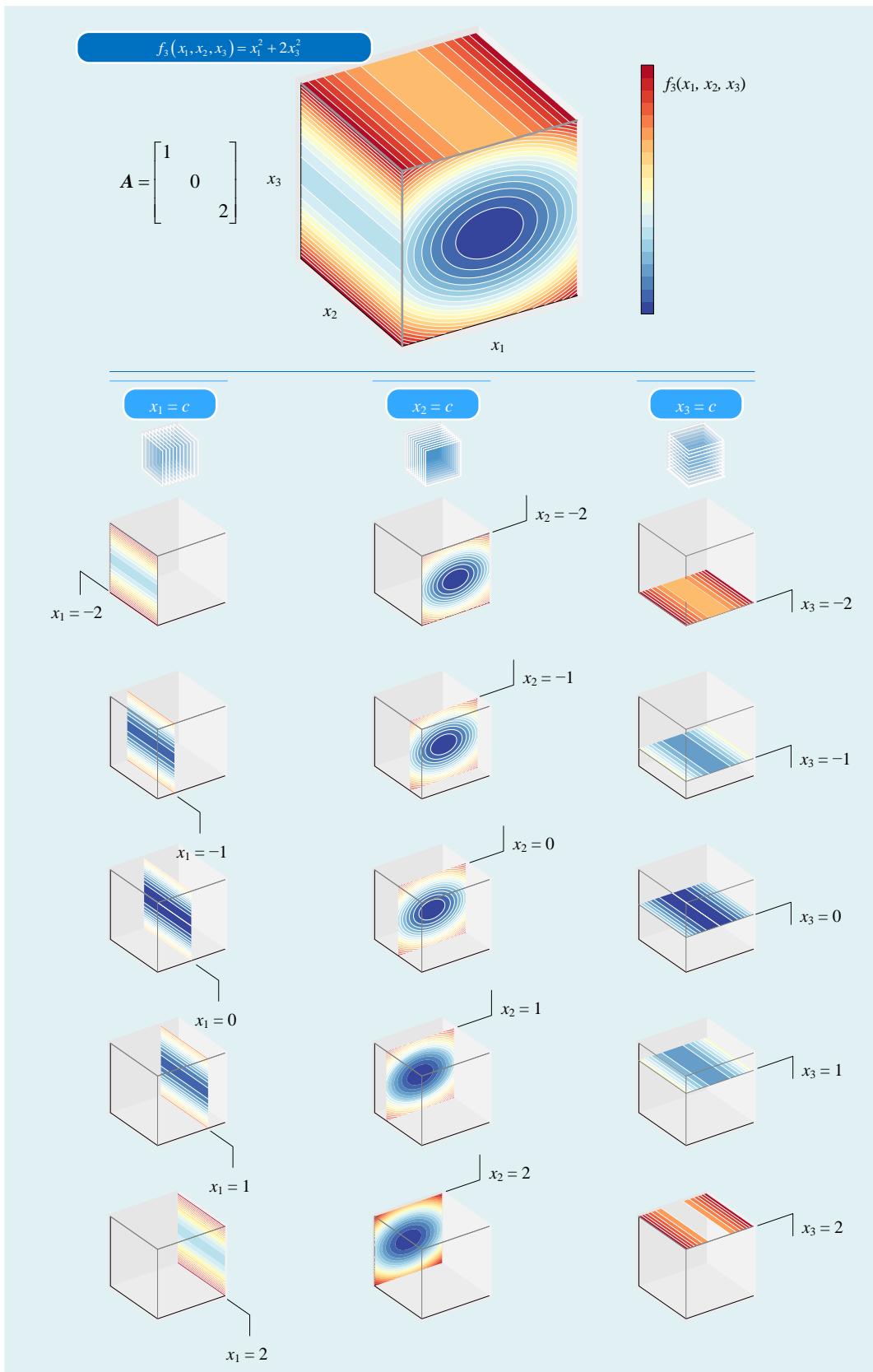


图 12. 三元二次型，半正定，情况 G

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

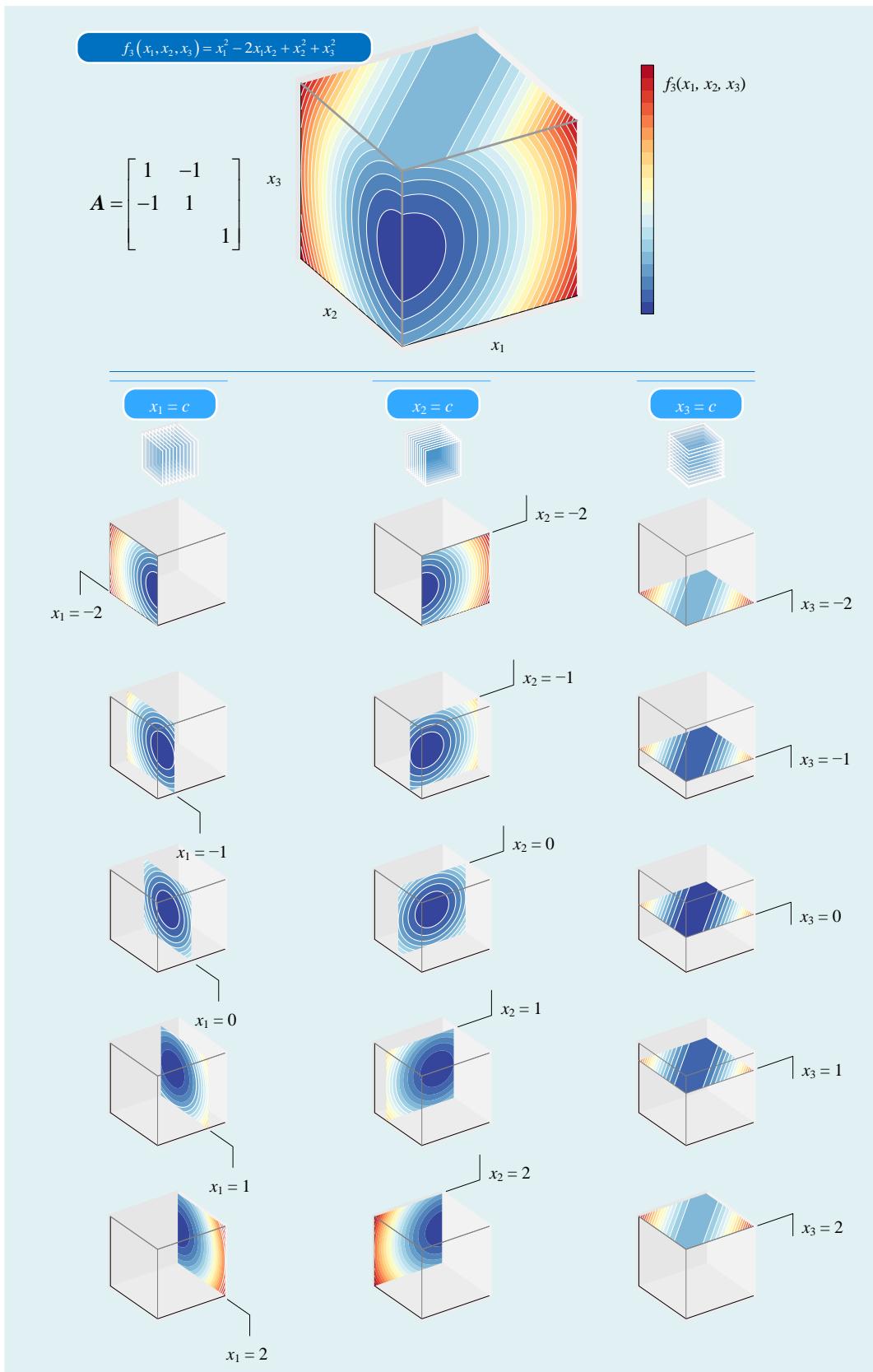


图 13. 三元二次型，半正定，情况 H

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

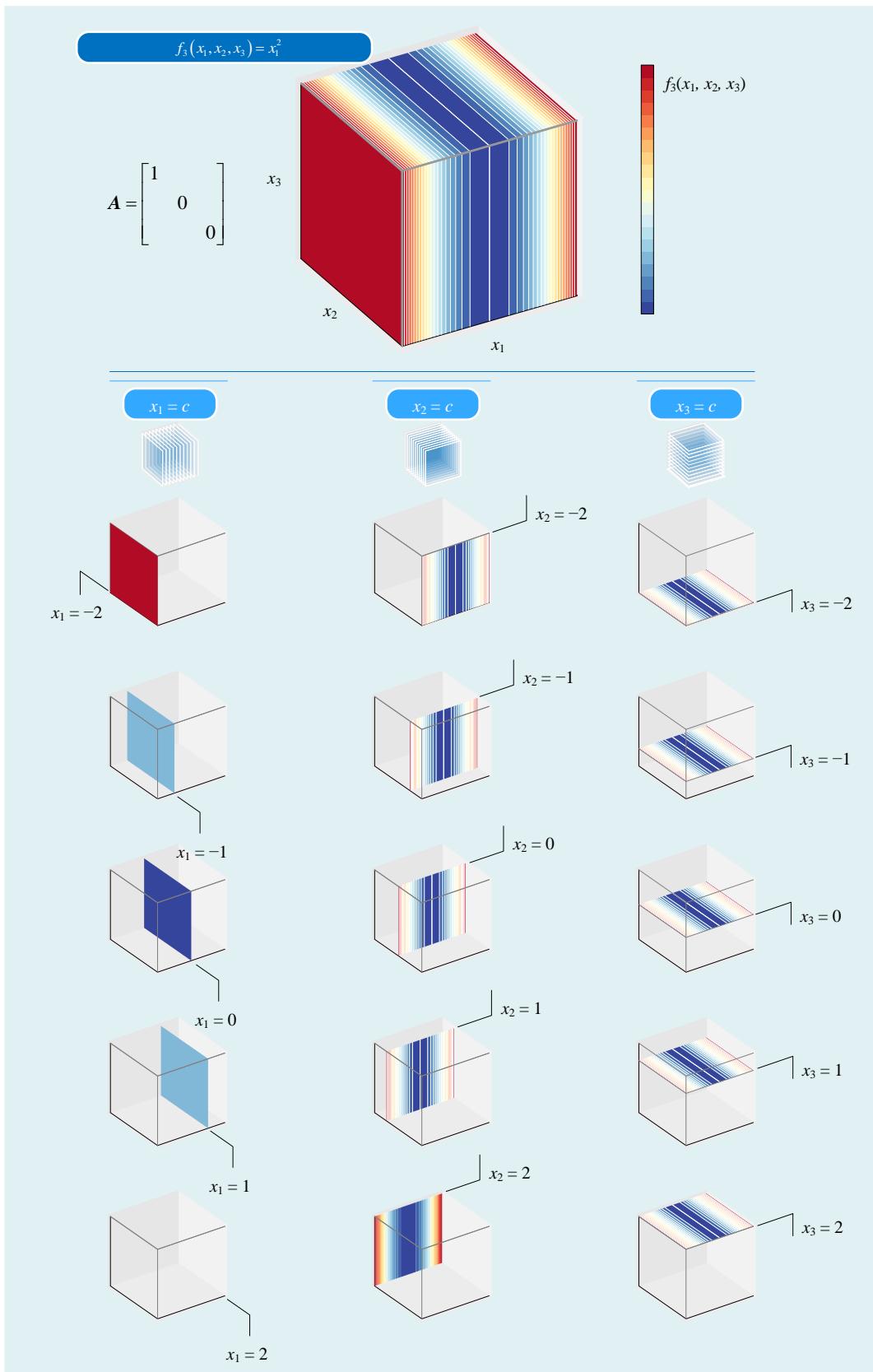


图 14. 三元二次型，半正定，情况 I

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

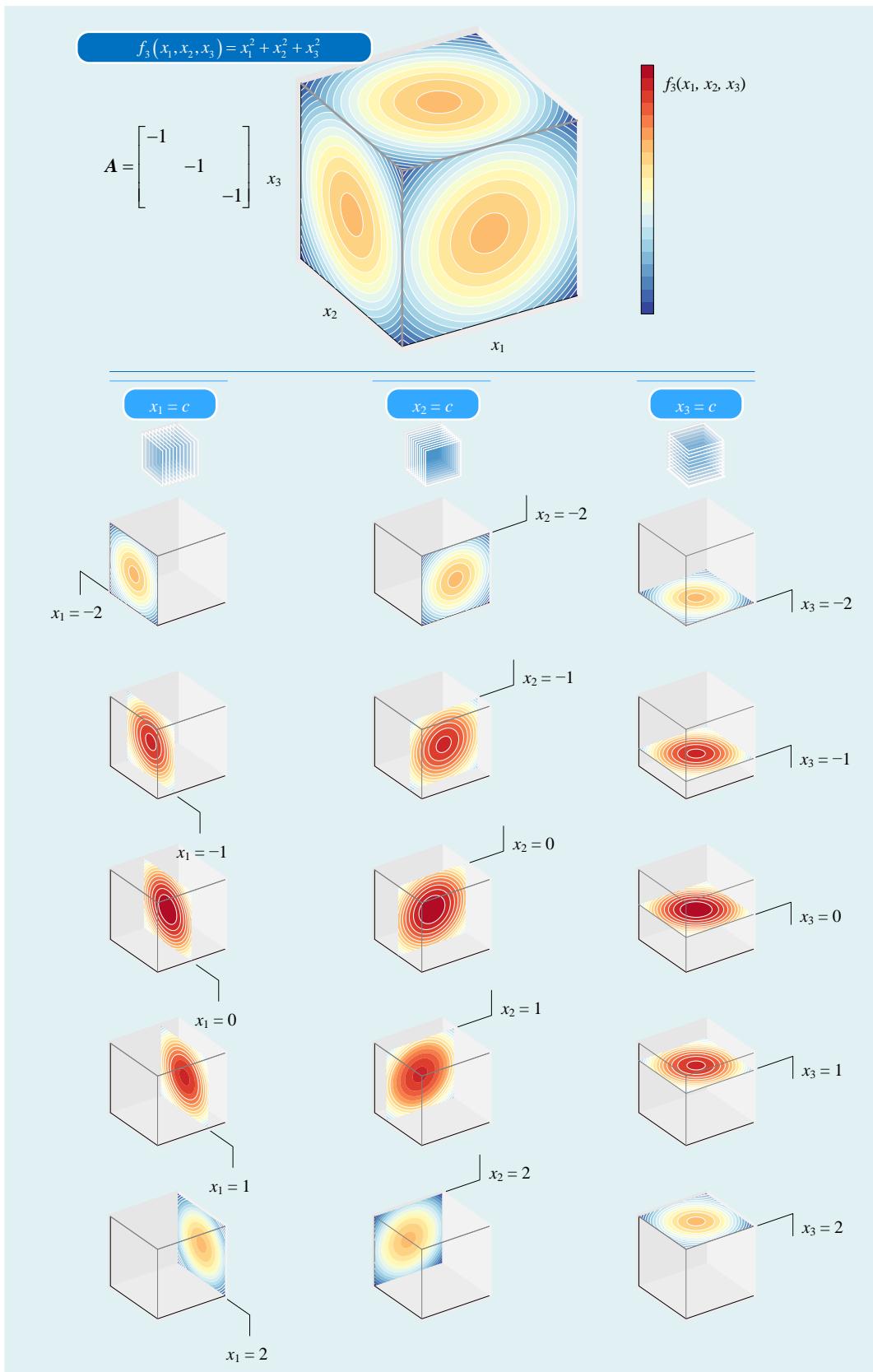


图 15. 三元二次型, 负定, 情况 J

本 PDF 文件为作者草稿, 发布目的为方便读者在移动终端学习, 终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有, 请勿商用, 引用请注明出处。

代码及 PDF 文件下载: <https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger: <https://space.bilibili.com/513194466>

欢迎大家批评指教, 本书专属邮箱: jiang.visualize.ml@gmail.com

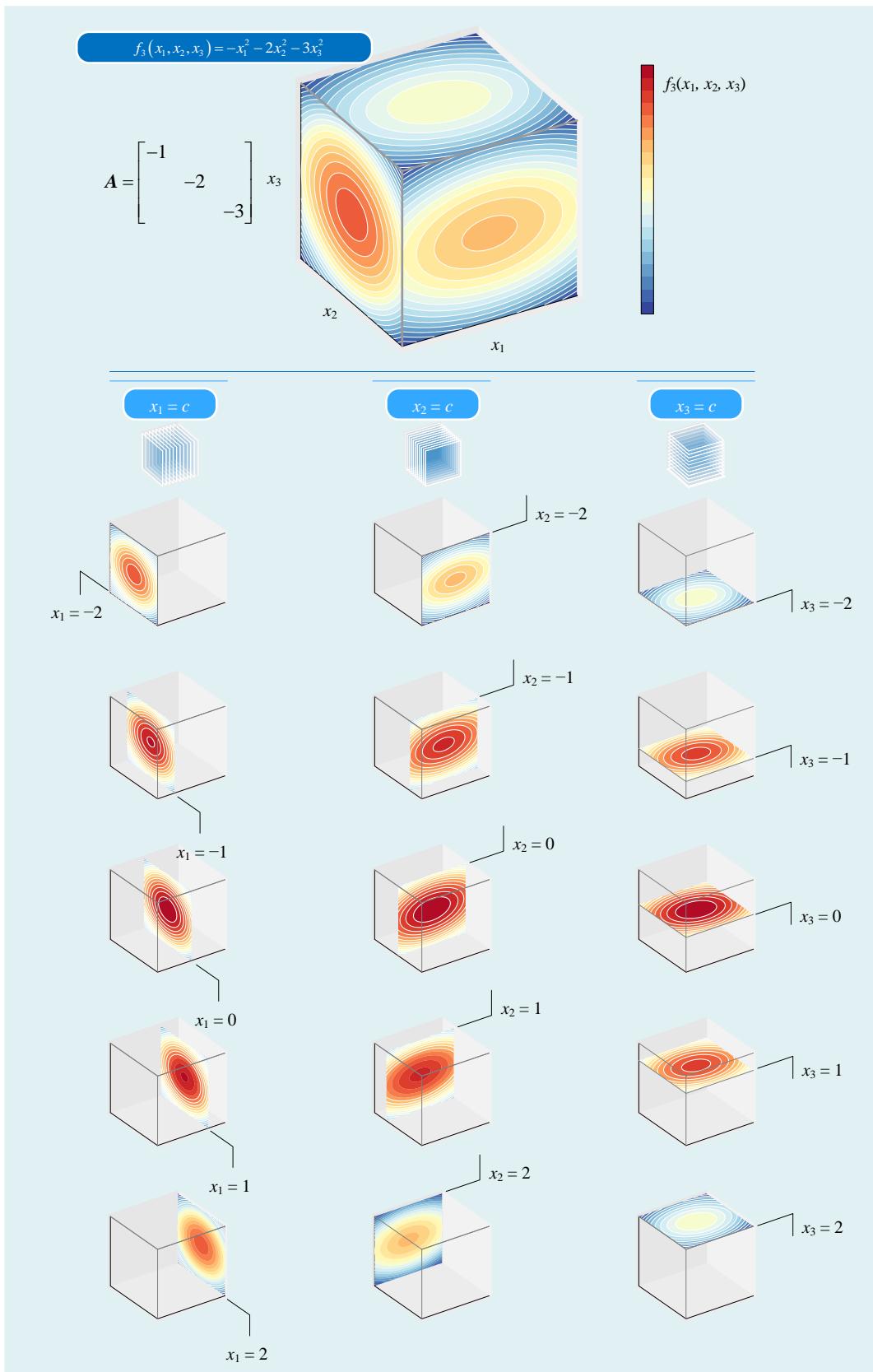


图 16. 三元二次型, 负定, 情况 K

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

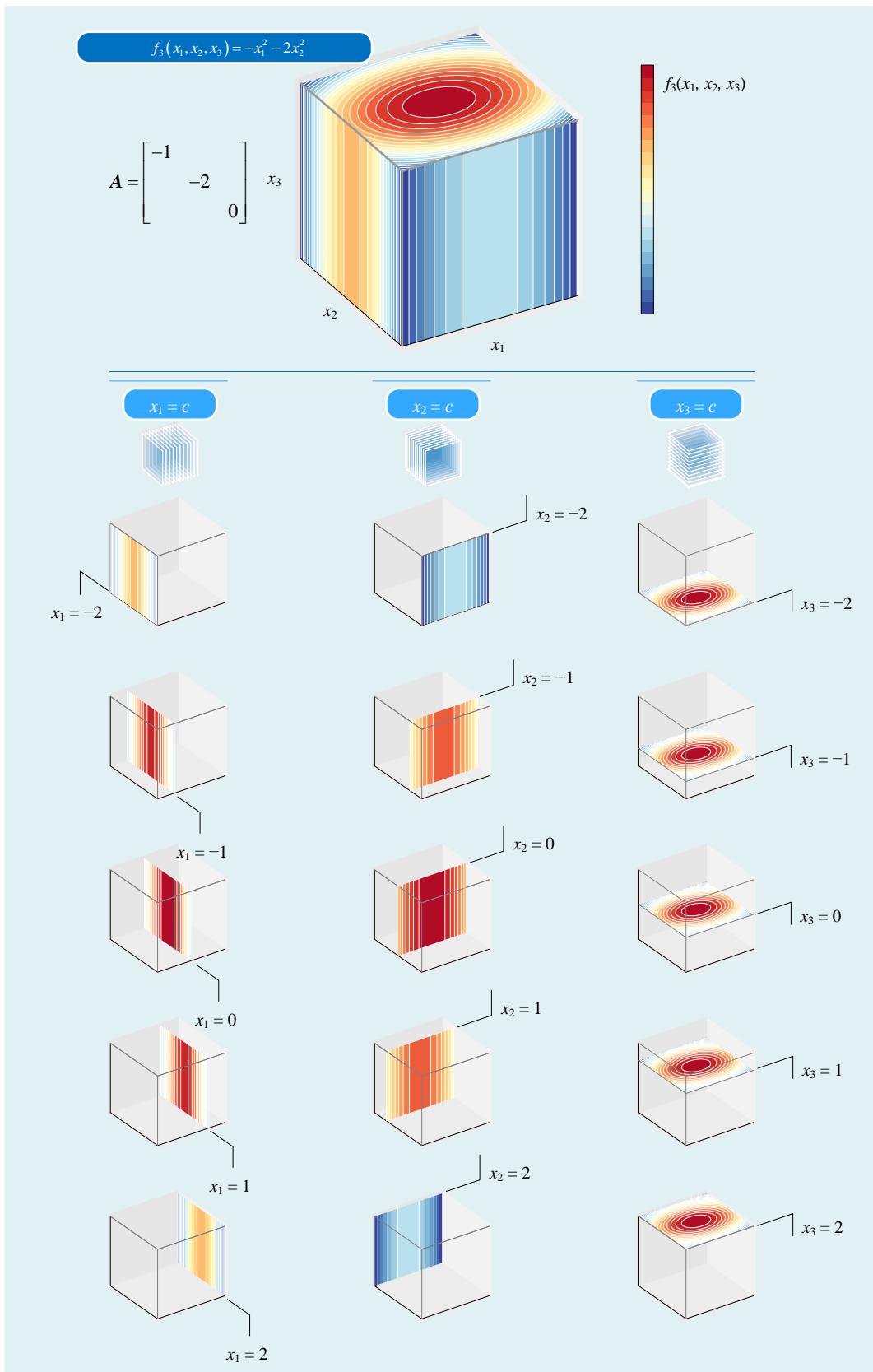


图 17. 三元二次型，半负定，情况 L

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

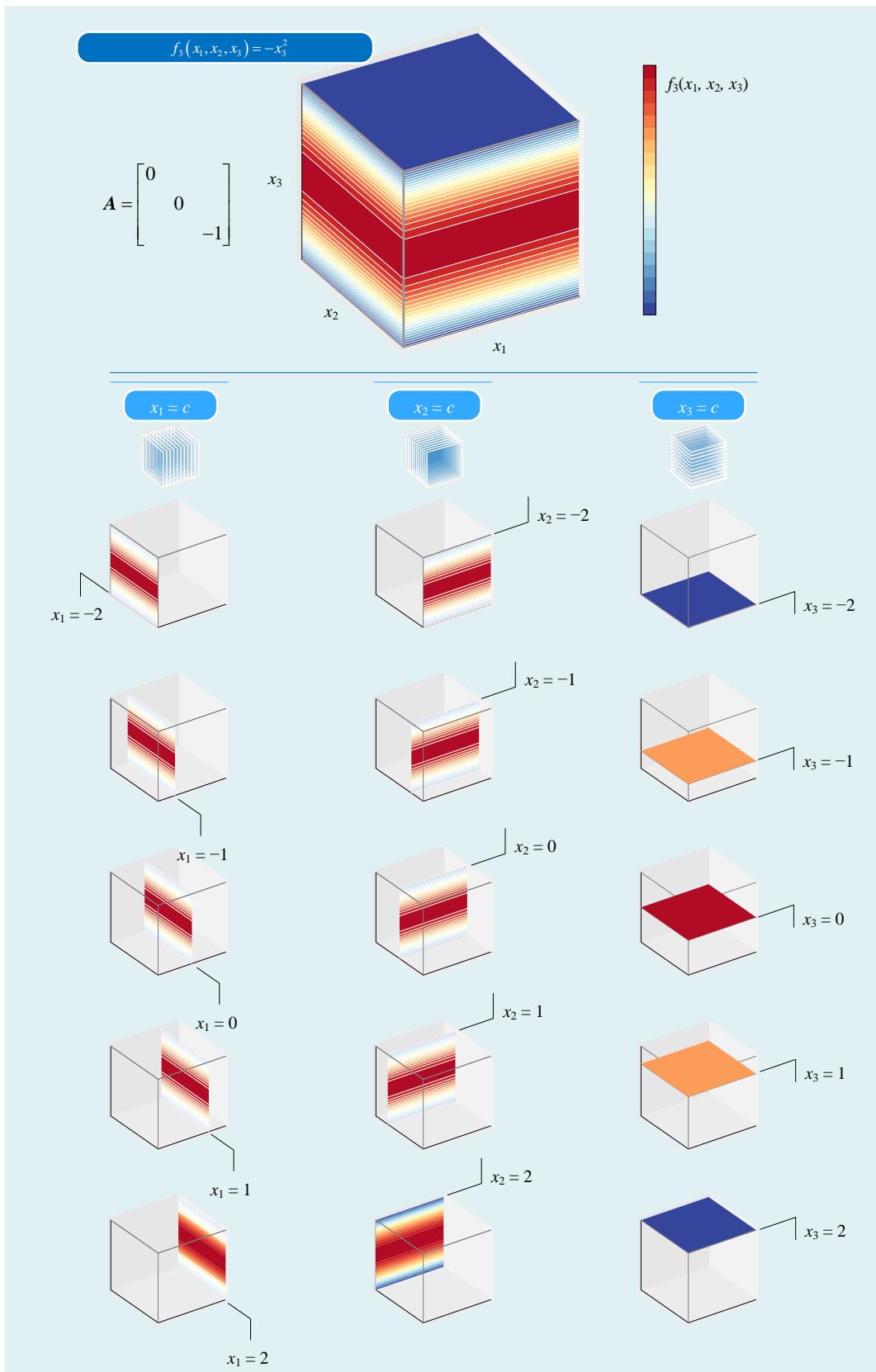


图 18. 三元二次型，半负定，情况 M

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

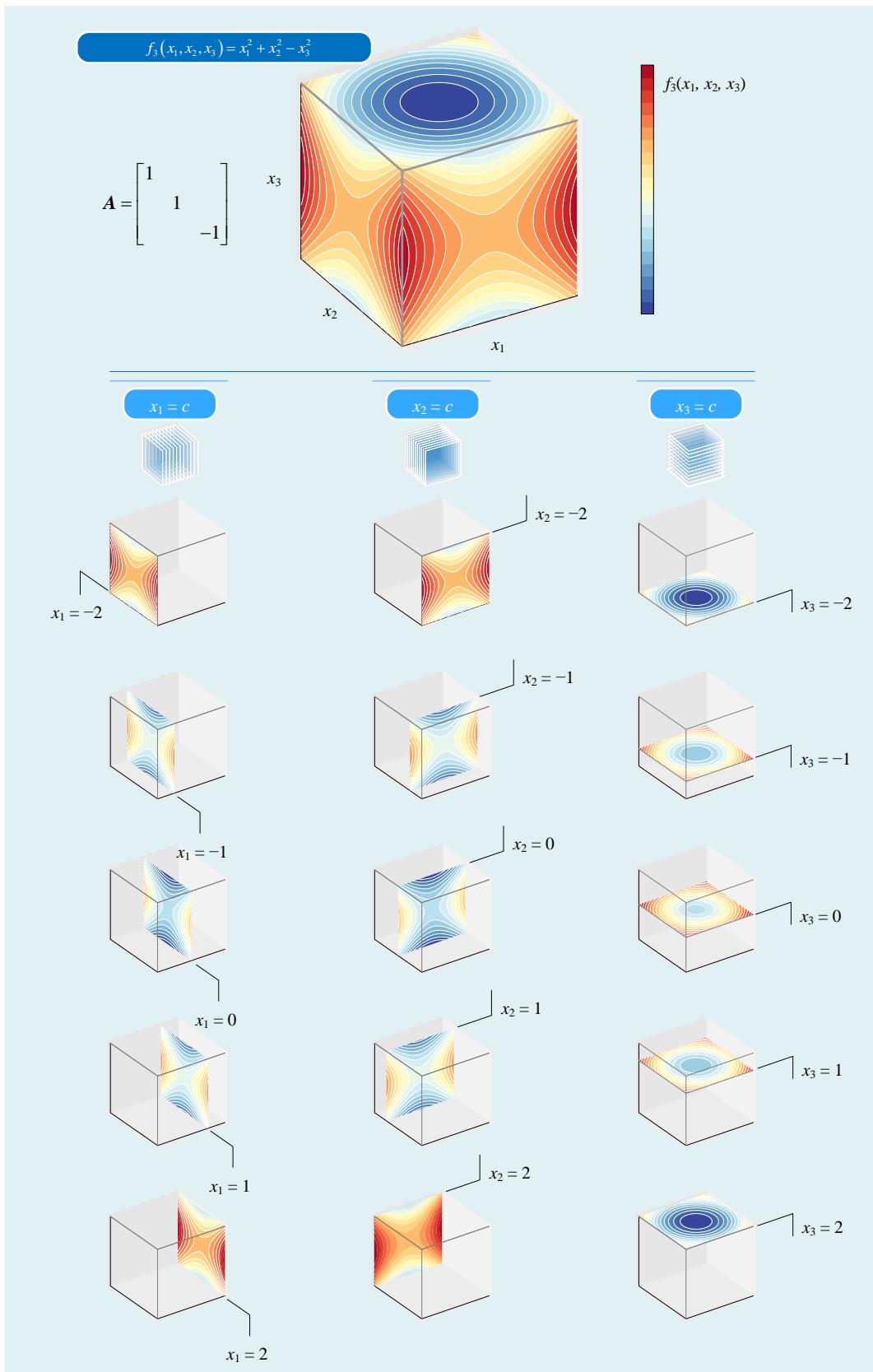


图 19. 三元二次型, 不定, 情况 N

本 PDF 文件为作者草稿, 发布目的为方便读者在移动终端学习, 终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有, 请勿商用, 引用请注明出处。

代码及 PDF 文件下载: <https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger: <https://space.bilibili.com/513194466>

欢迎大家批评指教, 本书专属邮箱: jiang.visualize.ml@gmail.com

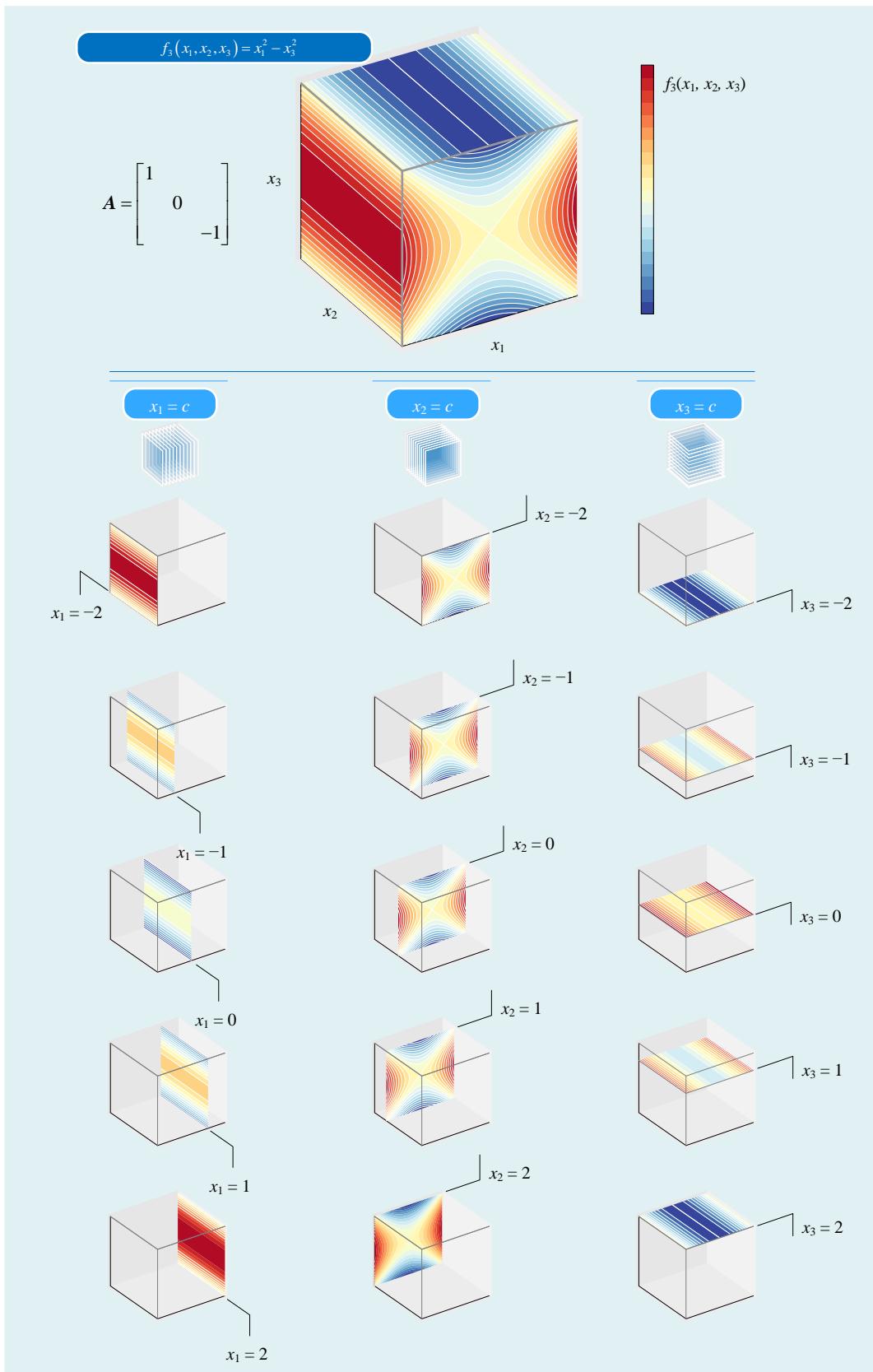


图 20. 三元二次型，不定，情况 O

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

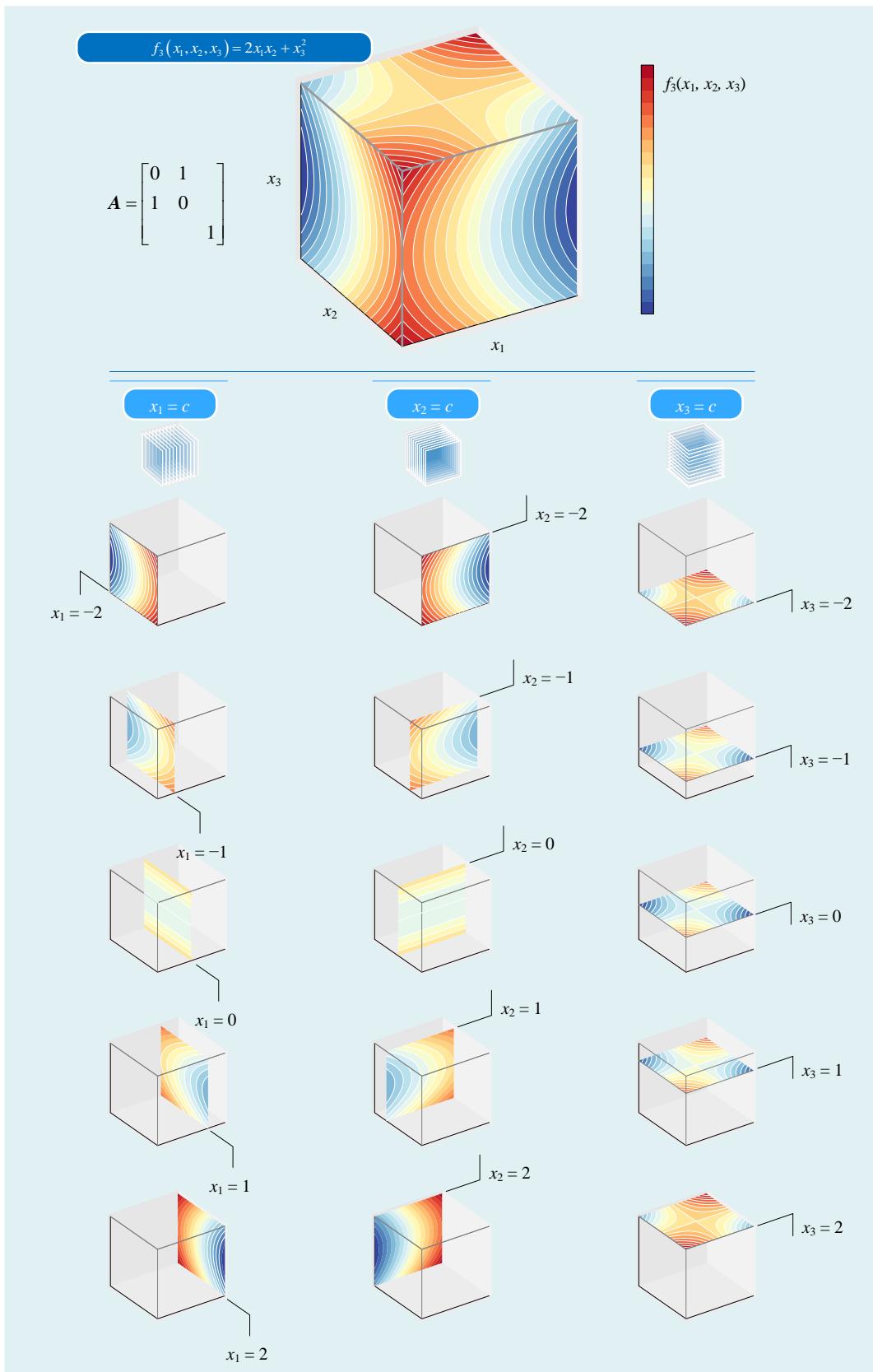


图 21. 三元二次型，不定，情况 P

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

32

Visualize Complex Functions

复数

丰富且美丽的线性、非线性变换



我内心燃烧着一团巨大的火焰，但没有人停下来在它身上取暖，路人只看到一缕烟雾。

A great fire burns within me, but no one stops to warm themselves at it, and passers-by only see a wisp of smoke.

——文森特·梵高 (Vincent van Gogh) | 荷兰后印象派画家 | 1853 ~ 1890



- ◀ `matplotlib.pyplot.axhline()` 绘制水平线
- ◀ `matplotlib.pyplot.axvline()` 绘制竖直线
- ◀ `matplotlib.pyplot.contour()` 绘制等高线图
- ◀ `matplotlib.pyplot.contourf()` 绘制填充等高线图
- ◀ `matplotlib.pyplot.pcolormesh()` 绘制二维网格数据的伪彩色图
- ◀ `matplotlib.pyplot.quiver()` 绘制箭头图
- ◀ `matplotlib.pyplot.scatter()` 绘制散点图
- ◀ `numpy.abs()` 计算复数模
- ◀ `numpy.angle()` 计算复数辐角
- ◀ `numpy.meshgrid()` 创建网格化数据
- ◀ `numpy.zeros_like()` 用来生成和输入矩阵形状相同的零矩阵

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

32.1 复数

什么是复数？

复数是由实部和虚部组成的数。实部表示复数在实轴上的投影，虚部表示复数在虚轴上的投影。复数可以用形如 $a + bi$ 的形式表示，其中 a 是实部， b 是虚部， i 是虚数单位，满足 $i^2 = -1$ 。

复平面

复平面是一个用于表示复数的平面，其中实轴和虚轴分别对应平面上的横轴和纵轴。

在复平面上，每个复数可以表示为一个点 (a, b) ，其中实部 a 对应于点在实轴上的投影，虚部 b 对应于点在虚轴上的投影。复数的实部和虚部可以用来确定平面上的点的位置。

复平面可以被看作是一个笛卡尔坐标系，其中实轴表示实数轴，从左向右增加，虚轴表示纯虚数轴，从下向上增加。原点 $(0, 0)$ 对应于零值复数。

通过复平面，可以直观地理解复数的性质和运算。例如，复数的加法和减法对应于在复平面上的向量相加和相减，复数的乘法对应于向量的伸缩和旋转，复数的除法对应于向量的缩放和旋转等。

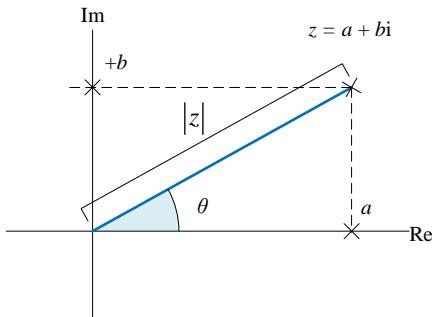


图 1. 复数

复数模、辐角

复平面的极坐标形式也很常见，其中复数可以表示为模和辐角的组合。在极坐标形式下，复数的模对应于从原点到点的距离，辐角对应于从正实轴逆时针旋转到点的角度。如图 2 所示，复平面上一系列点可以看成是一组“箭头”。

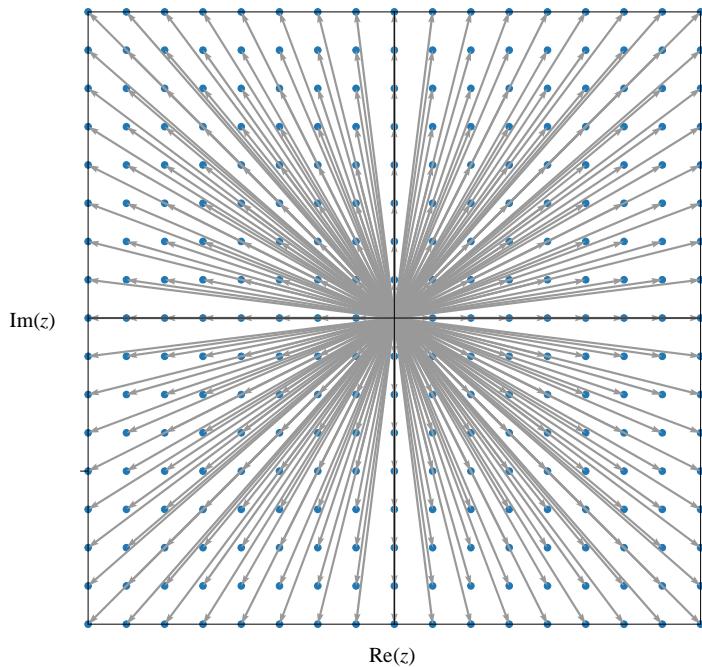


图 2. 复平面上，一系列“箭头”

复数模是指复数的绝对值或者长度，表示从原点到复数在复平面上的距离。对于一个复数 $z = a + bi$ ，其模可以用 $|z| = \sqrt{a^2 + b^2}$ 计算。图 3 所示为两种可视化复数模的方案，本书采用平面等高线。

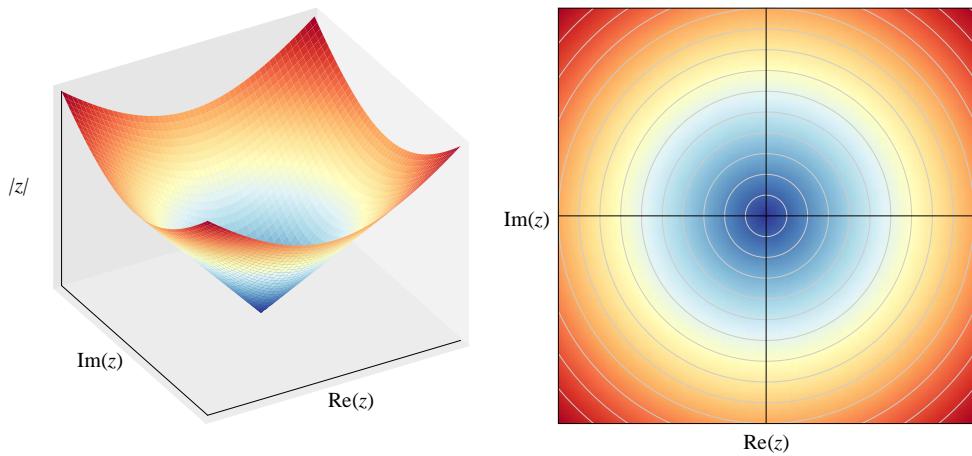


图 3. 可视化复数的模

辐角，或幅角，是指从正实轴逆时针旋转到复数所在位置的角度。可以使用三角函数来计算角度。对于一个非零复数 $z = a + bi$ ，其角度可以用 $\theta = \arctan(b, a)$ 计算。`arctan()` 是一个可以返回

带符号角度的反正切函数。角度的单位通常是弧度。图 4 所示为在复平面上用 hsv 色谱可视化辐角，这是因为 hsv 色谱为循环色谱，首尾颜色咬合。

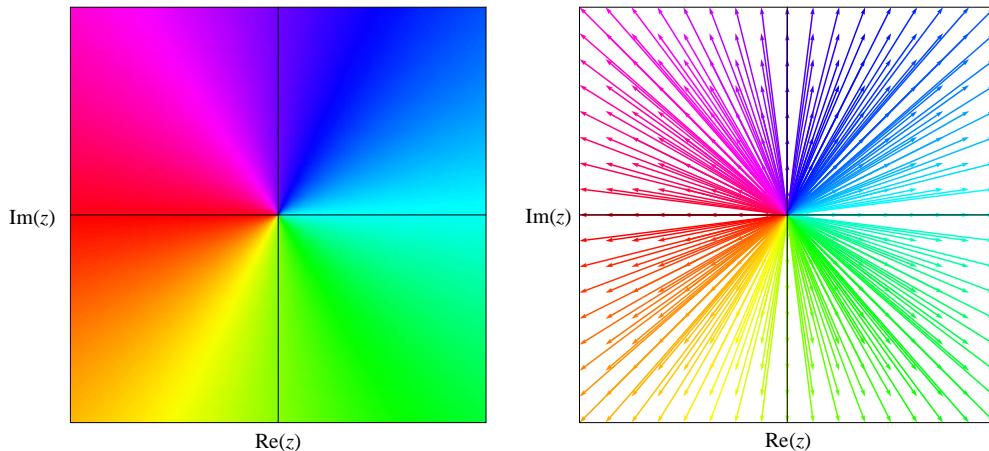


图 4. 复平面上，用 hsv 色谱可视化辐角

在 NumPy 中，可以使用 `numpy.abs()` 函数来计算复数的模，使用 `numpy.angle()` 函数来计算复数的辐角。

复数的模和角度提供了对复数在复平面上的位置和特征的描述。它们在复数的运算、变换和表示中都起着重要的作用。可以使用模和角度来进行复数的乘法、除法、幂运算等操作，并且它们还可以方便地表示复数的极坐标形式。

复数在许多领域有广泛的应用。在物理学中，复数用于描述交流电路中的电压和电流。在工程学中，复数在信号处理和控制系统中起着重要作用。在数学领域，复数广泛用于解析几何和复变函数等领域。

在机器学习领域，复数也有一些应用。例如，复数神经网络是一种利用复数权重和激活函数的神经网络模型。这种模型可以更好地处理复杂的信号和数据，特别是在音频和图像处理等领域。复数神经网络在语音识别、图像处理和模式识别等任务中取得了一定的成功。

32.1 复数函数

复数函数是指接受复数作为输入并返回复数作为输出的函数。复数函数可以将一个或多个复数作为参数，然后根据特定的规则进行计算，并返回一个或多个复数作为结果。

复数函数可以涉及各种数学运算和操作，例如加法、减法、乘法、除法、幂运算、三角函数、指数函数、对数函数等。这些函数在复数领域中扮演着重要的角色，因为复数具有实部和虚部，可以进行各种运算和转换。

一些常见的复数函数包括：

复数加法和减法函数：将两个或多个复数相加或相减。

复数乘法和除法函数：将两个或多个复数相乘或相除。

复数幂函数：计算一个复数的幂，其中指数可以是实数或复数。

复数三角函数：包括正弦、余弦、正切等，可以用来计算复数的三角函数值。

复数指数函数和对数函数：计算复数的指数和对数值。

本章重点内容就是用前文的可视化方案展示复数函数。本章最后选定图 5 所示两个可视化方案展示复数函数。图 5 (a) 背景是辐角，等高线为复数模；图 5 (b) 的正方形网格用来可视化复数函数导致的线性或非线性变换。本书前文介绍过如何用等高线绘制这种网格。

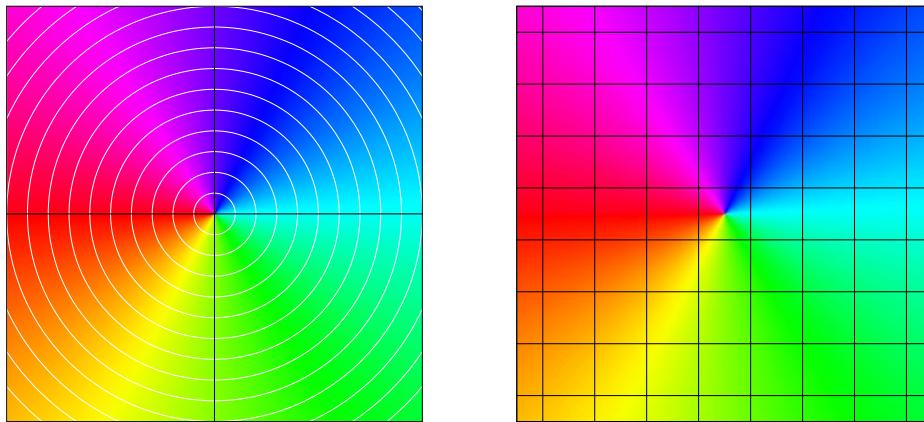


图 5. 复数函数可视化方案

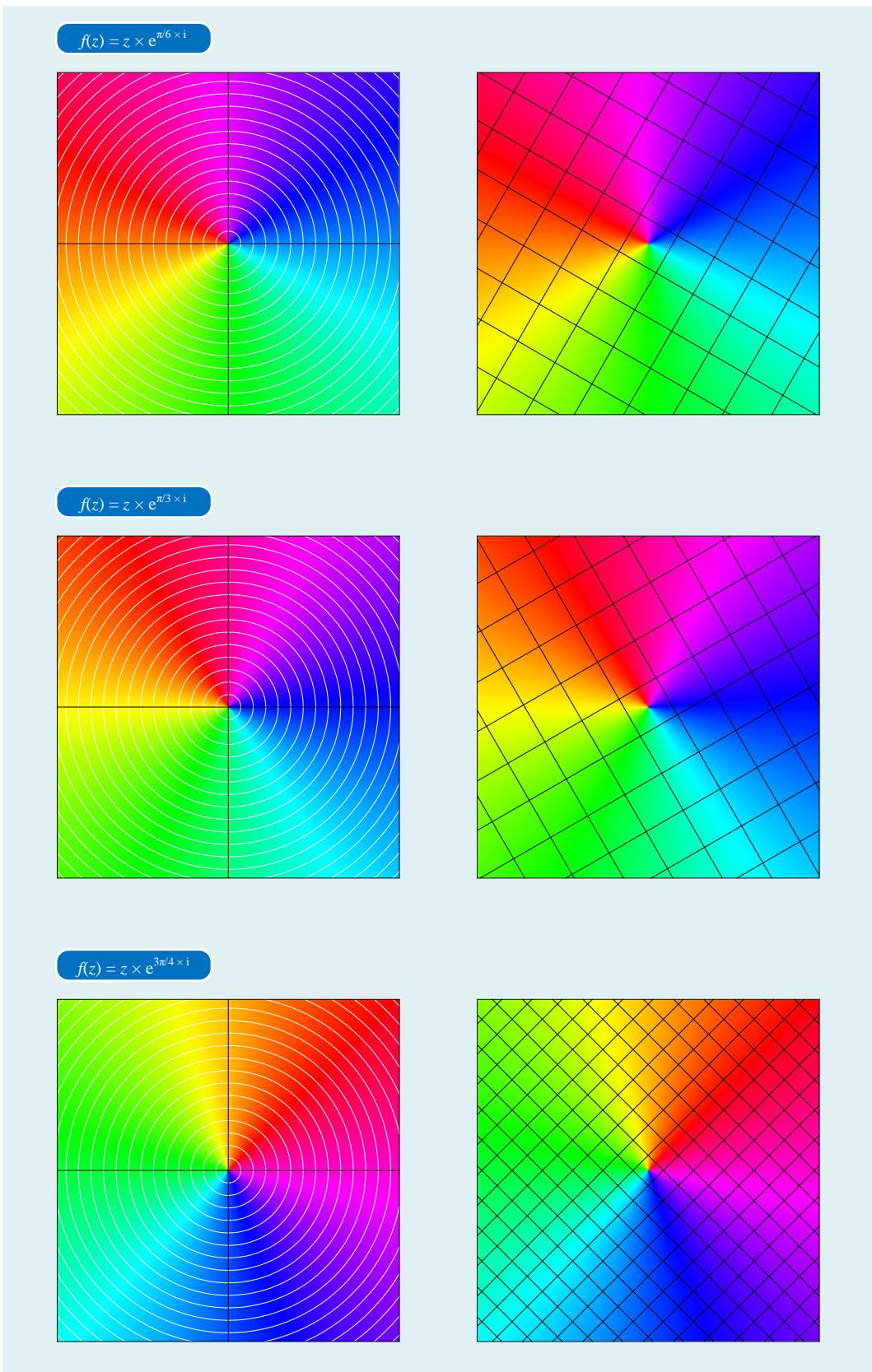


图 6. 旋转、缩放

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

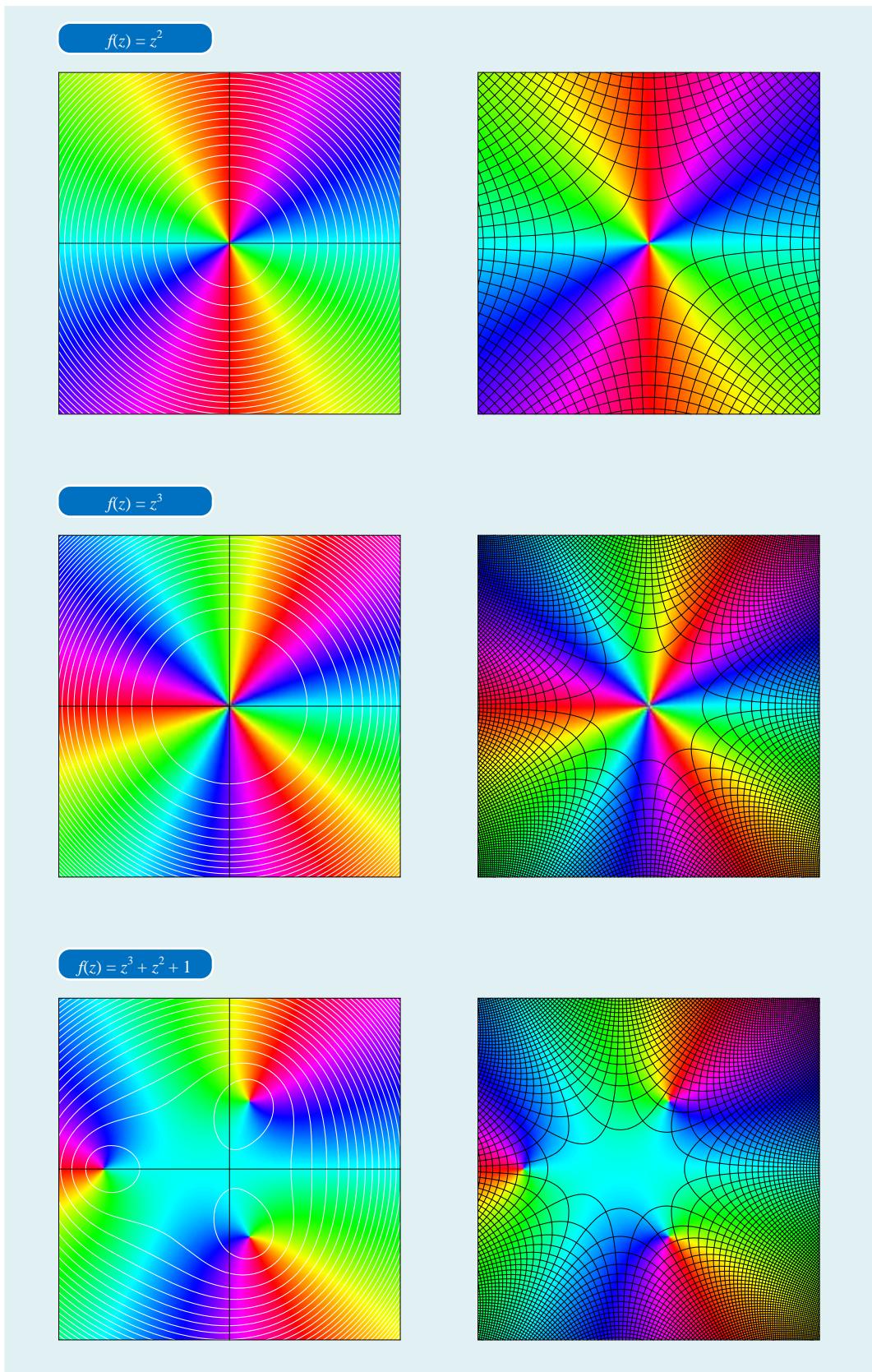


图 7. 多项式

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

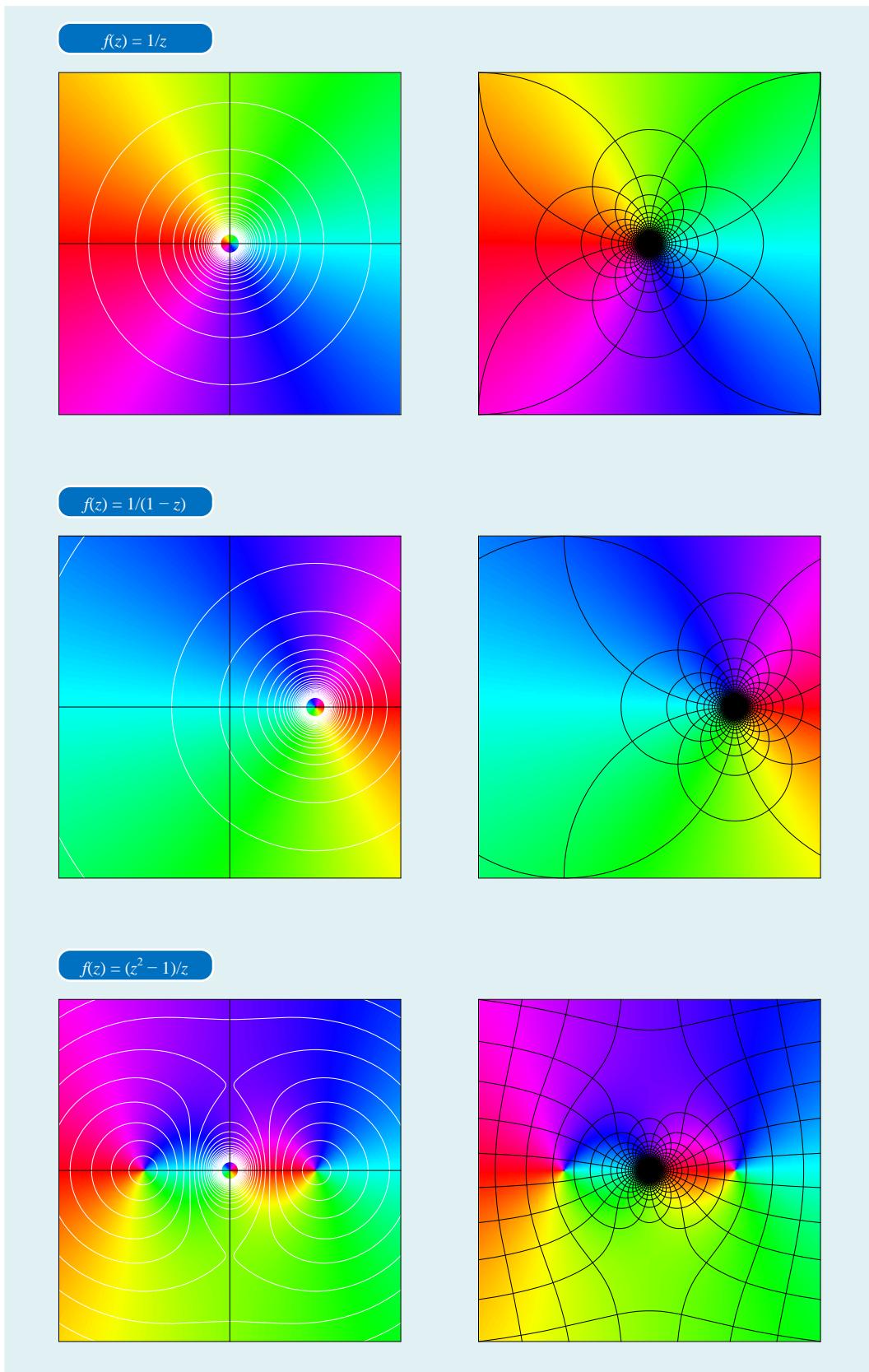


图 8. 分式, 第 1 组

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

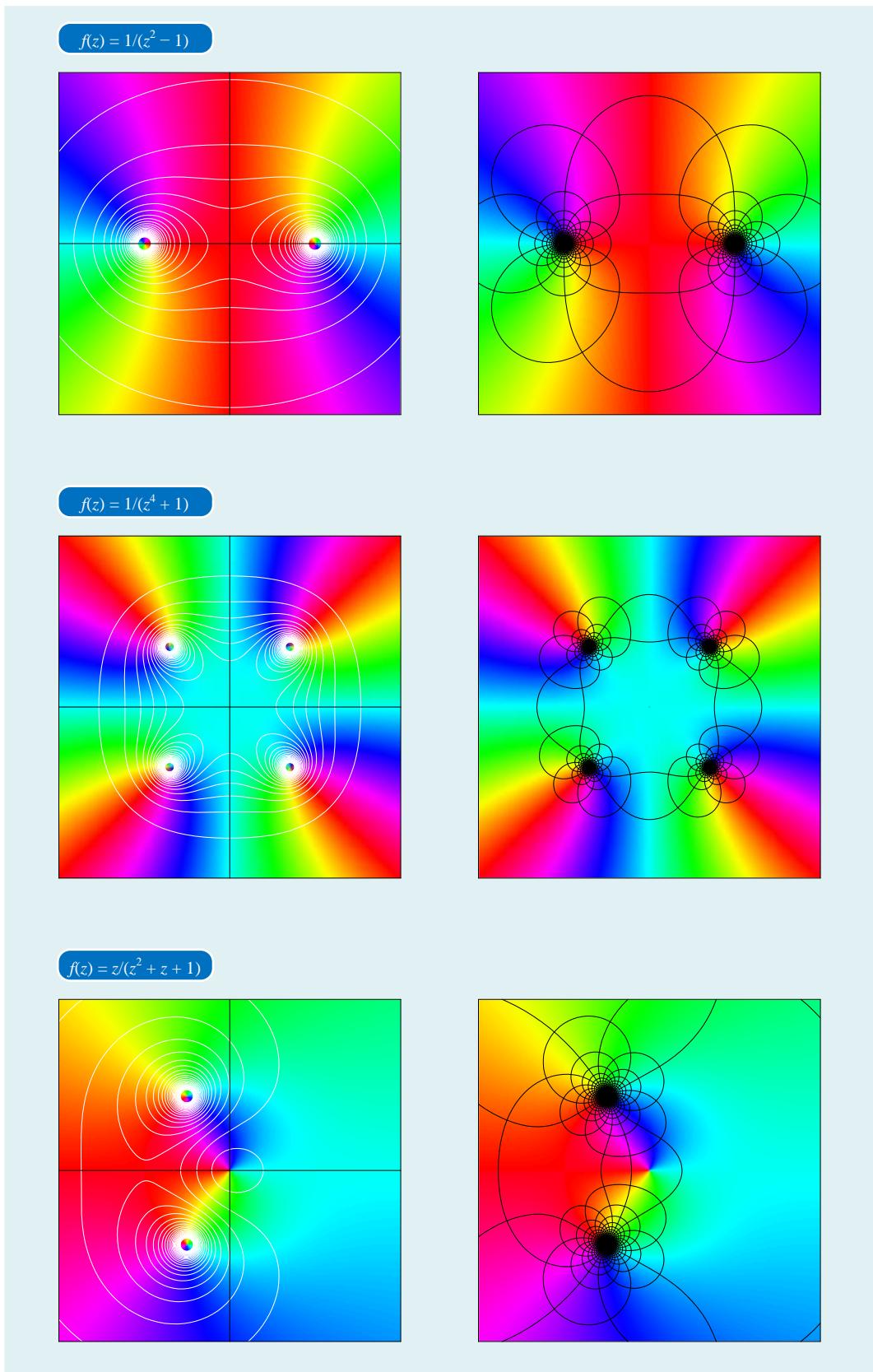


图 9. 分式, 第 2 组

本 PDF 文件为作者草稿, 发布目的为方便读者在移动终端学习, 终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有, 请勿商用, 引用请注明出处。

代码及 PDF 文件下载: <https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger: <https://space.bilibili.com/513194466>
欢迎大家批评指教, 本书专属邮箱: jiang.visualize.ml@gmail.com

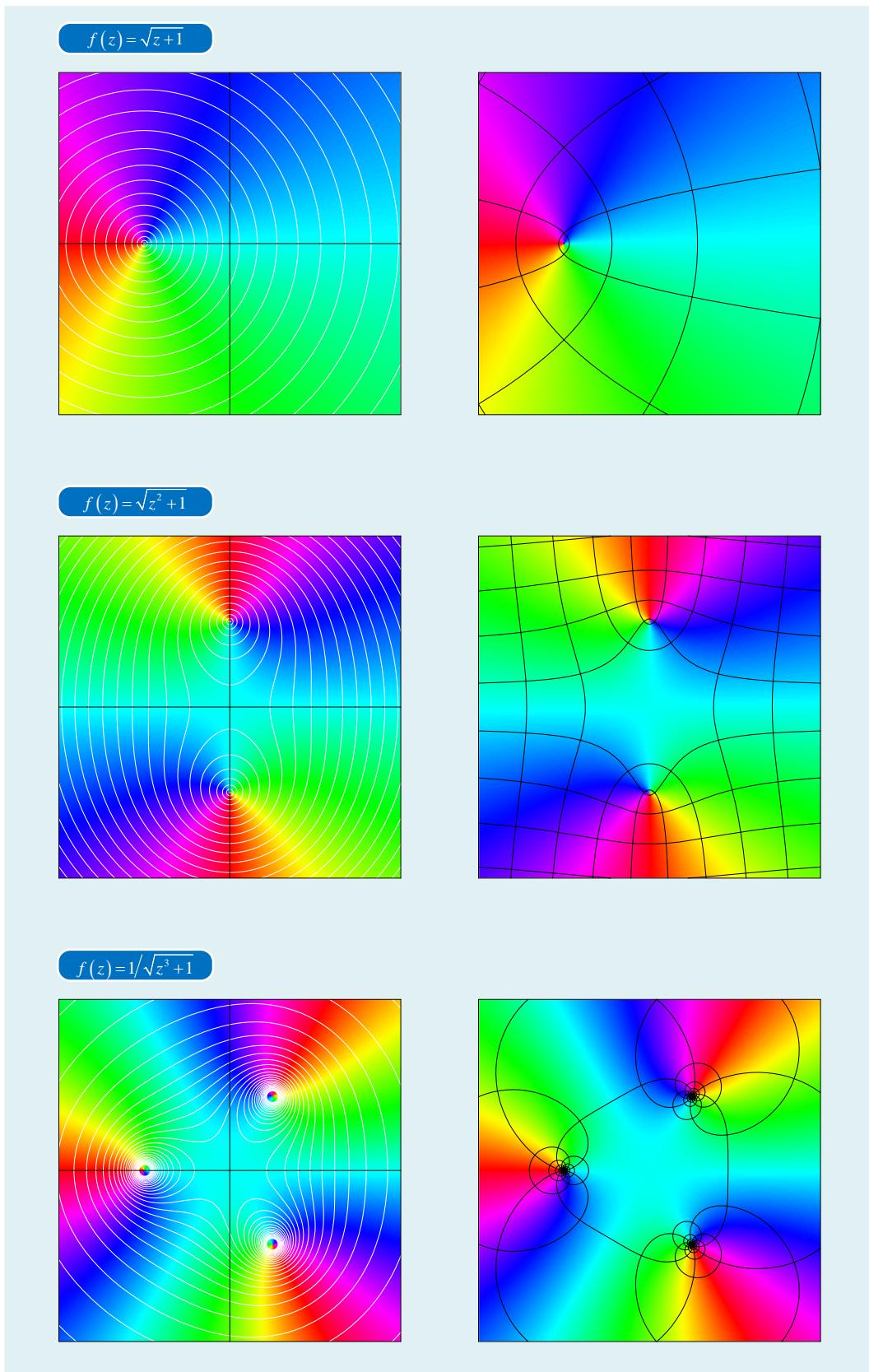


图 10. 根式

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

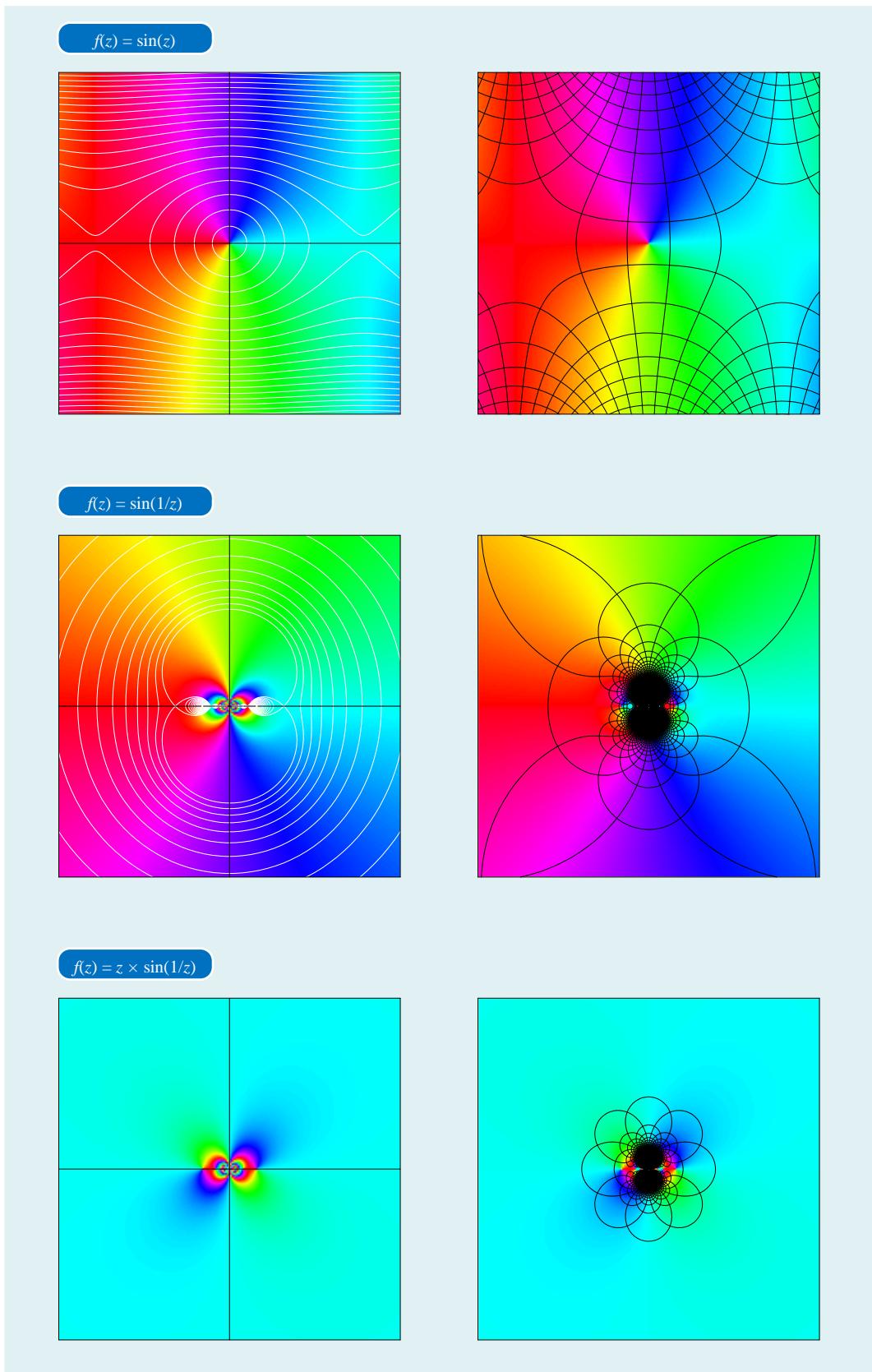


图 11. 三角正弦

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

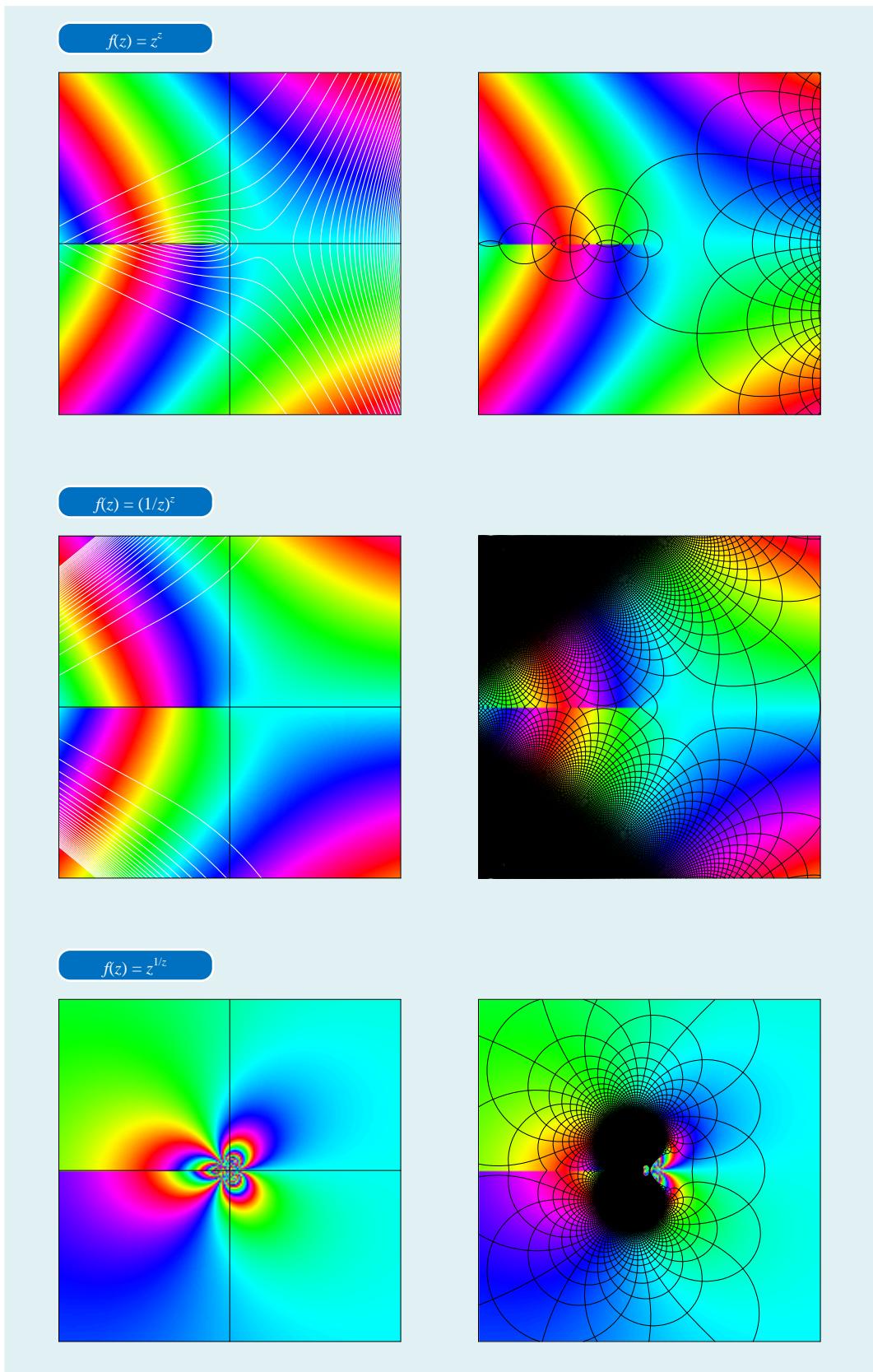


图 12. 乘幂

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

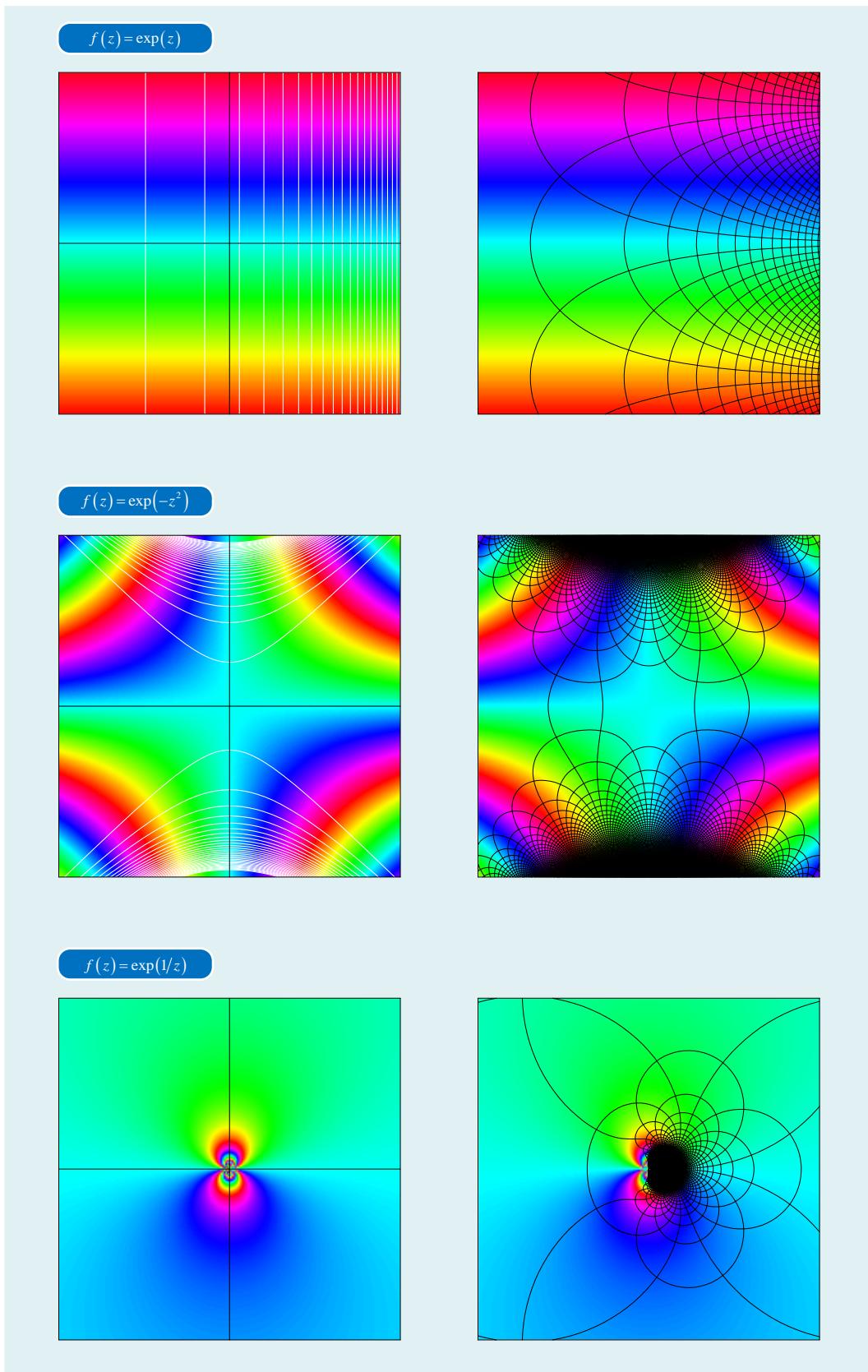


图 13. 指数

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

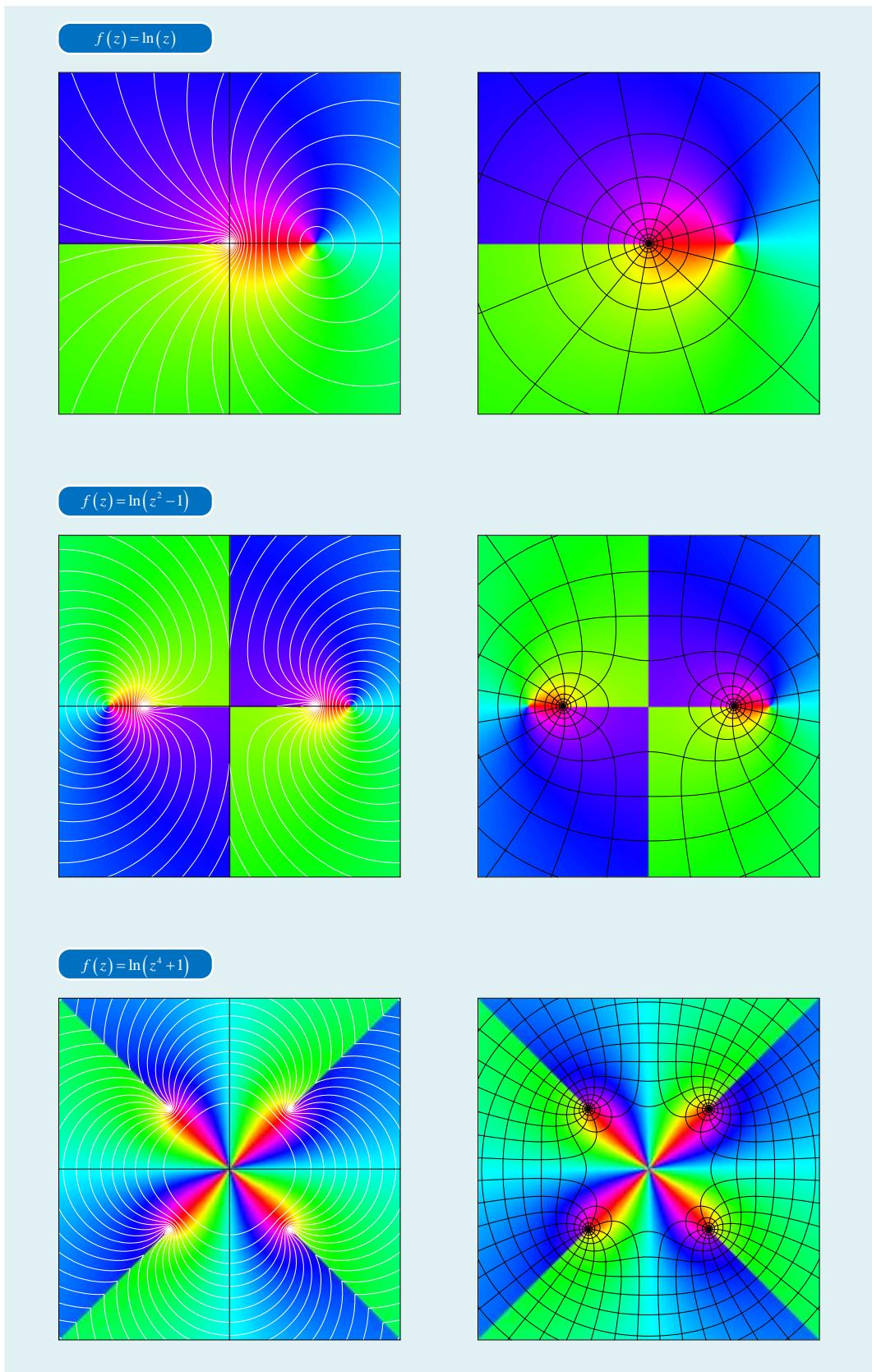


图 14. 对数

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com



Fractal

分形

自然界无处不在的自相似



今天，是别人的；我苦苦耕耘的未来，是我的。

The present is theirs; the future, for which I really worked, is mine.

——尼古拉·特斯拉 (Nikola Tesla) | 发明家、物理学家 | 1856 ~ 1943



- ◀ XXXXX

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

33.1 分形

分形 (fractal) 是指具有自相似性质的几何形状或图案。它们在不同的尺度上都呈现出类似的结构，即无论是观察整体还是细节部分，都可以看到相似的形态。分形通常通过迭代的方式生成，即通过重复一定的规则或操作来构造出越来越复杂的形状。反过来看，再复杂的图形通过解构可以得到简单的几何形状或图案。

自然界中存在许多分形结构。科赫曲线是由无限个自相似的三角形组成的曲线。科赫曲线在自然界随处可见。

树木的分枝结构呈现出分形特征。树干分出树枝，树枝再分出更小的枝干，如此类推。这种分形结构使得树木能够有效地获取光线和水分。

闪电的形状也可以被视为一种分形。它的形态在不同的尺度上都存在类似的形状特征。

山脉的轮廓也展现出分形的特征。无论是在整个山脉的缩放上还是在山脉的岩层中，都可以看到类似的形状。

本章介绍如何利用 Python 可视化几种常见的分形。

33.2 Koch 雪花

Koch 雪花 (Koch Snowflake)由瑞典数学家 Helge von Koch 在 1904 年引入，是分形几何中的经典案例之一。

Koch 雪花是通过一系列简单的规则和迭代生成的。如图 1 所示，Koch 雪花起始于一个等边三角形，通过以下步骤来构造 Koch 雪花：

- ▶ 将每条边分成三等分。
- ▶ 在中间的一段边上，建立一个等边三角形，向外伸出。
- ▶ 移除初始三角形的底边。

在每次迭代中，对于每条线段，都会重复上述步骤。通过不断重复这些规则，Koch 雪花的形状逐渐复杂起来，边缘颗粒度不断提高，类似于一个由无数小三角形构成的雪花。无论观察 Koch 雪花的哪一部分，都可以发现与整体相似的几何形状。

Koch 雪花的具体算法，请参考：

<https://mathworld.wolfram.com/KochSnowflake.html>

33.3 谢尔宾斯基三角形

谢尔宾斯基三角形 (Sierpinski triangle) 是一种数学图形，是以波兰数学家 Waclaw Sierpiński 的名字命名的。如图 2 所示，谢尔宾斯基三角形是一个由等边三角形构成的图形，每一步都通过以下规则生成新的图形：

- ▶ 将初始的等边三角形分成四个等边子三角形，中央的三角形被去除。
- ▶ 对每个剩余的子三角形，重复步骤 1。

通过不断重复这个过程，谢尔宾斯基三角形会越来越复杂，每一步都生成更多的小三角形。最终的图形是一个无限细分的结构，看起来像一个由三角形构成的海绵或细分的地形。

谢尔宾斯基三角形具体算法，请参考：

<https://mathworld.wolfram.com/SierpinskiSieve.html>

谢尔宾斯基地毡 (Sierpinski carpet) 的构造与谢尔宾斯基三角形相似，区别仅在于谢尔宾斯基地毡是以正方形而非等边三角形为基础的。如图 3 所示，将一个实心正方形划分为 3×3 的 9 个小正方形，去掉中间的小正方形，再对余下的小正方形重复这一操作便能得到谢尔宾斯基地毡。谢尔宾斯基地毡具体算法，请参考：

<https://mathworld.wolfram.com/SierpinskiCarpet.html>

33.4 Vicsek 正方形分形

Vicsek 正方形分形 (Vicsek box fractal) 是一种分形结构，以匈牙利物理学家 Tamás Vicsek 的名字命名。如图 4、图 5，Vicsek 正方形分形生成方法如下：

- ▶ 起始于一个正方形。
- ▶ 将正方形分成 9 个相等的小正方形，中间的正方形保留。
- ▶ 对于每个保留的正方形，将其分成 9 个相等的小正方形，再将中间的正方形保留。

重复上述步骤，对每个保留的正方形进行相同的分割。

Vicsek 正方形分形具体算法，请参考：

<https://mathworld.wolfram.com/BoxFractal.html>

33.5 龙曲线

龙曲线 (dragon curve) 是一种分形曲线，以其蜿蜒曲折、复杂而美丽的形状而闻名。龙曲线由两个简单的规则构建而成：

- ▶ 起始于一条直线，可以将其视为一条 "1" 的序列。
- ▶ 对于每一次迭代，将当前序列复制一份，并在复制的序列中，将每个元素都反转（将 "1" 变为 "0"，将 "0" 变为 "1"）。
- ▶ 将复制的序列连接到原始序列的末尾。
- ▶ 重复上述步骤，不断迭代生成更长的序列。

通过迭代应用这些规则，龙曲线的形状逐渐复杂起来，呈现出蜿蜒曲折、自相似的特征。如图 6 所示，龙曲线可以在二维平面上绘制出来，形状类似于一条盘旋蜿蜒的龙。

33.6 巴恩斯利蕨

巴恩斯利蕨 (Barnsley fern) 是一种分形植物形状，以 Michael Barnsley 命名，他在 1988 年的书籍 *Fractals Everywhere* 中首次介绍了它。

巴恩斯利蕨的形状类似于蕨类植物的叶子。如图 7 所示，巴恩斯利蕨由一系列的线段构成，这些线段按照一定的规则进行迭代生成。巴恩斯利蕨分形的具体算法，请参考：

<https://mathworld.wolfram.com/BarnsleysFern.html>

33.7 毕达哥拉斯树

毕达哥拉斯树 (Pythagoras Tree)，也称勾股树，以古希腊数学家毕达哥拉斯的名字命名。

如图 8 所示，在毕达哥拉斯树分形中，每个矩形都成为下一级矩形的主干，并且每个新添加的矩形都相对于前一级的矩形进行几何变换。最终，毕达哥拉斯树呈现出一个由许多嵌套的矩形组成的树状结构。

毕达哥拉斯树的具体算法，请参考：

<https://mathworld.wolfram.com/PythagorasTree.html>

33.8 曼德博集合

曼德博集合 (Mandelbrot Set) 是一种在复平面上的分形图形，以法国数学家 Benoit Mandelbrot 的名字命名。曼德博集合是由复数运算和迭代生成的。曼德博集合的具体算法，请大家参考：

<https://mathworld.wolfram.com/MandelbrotSet.html>

如图 9 所示，曼德博集合的最显著特点是其自相似性。即使放大曼德博集合的一小部分，也可以发现与整体相似的结构。这种自相似性是通过迭代函数和复数运算的属性得到的。

33.9 朱利亚集合

朱利亚集合 (Julia Set) 也是一种在复平面上的分形图形，以法国数学家 Gaston Julia 的名字命名。与曼德博集合类似，朱利亚集合也是通过复数运算和迭代生成的。图 10 ~ 图 13 展示三组朱利亚几何分形曲线。特别是图 12、图 13，可以看成是朱利亚集合分形曲线连续变化的 12 个快照。

有关朱利亚集合的算法，请大家参考：

<https://mathworld.wolfram.com/JuliaSet.html>

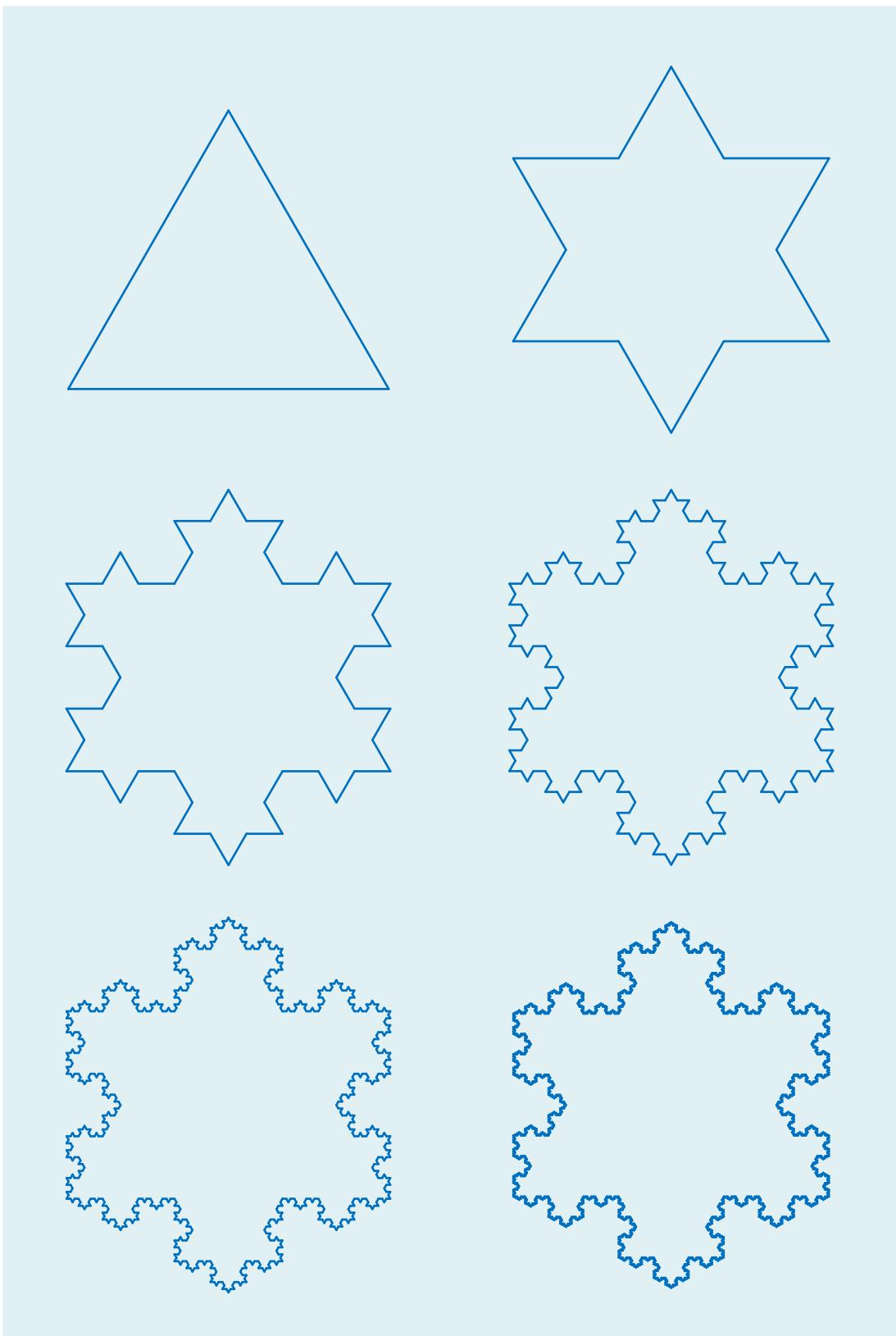


图 1. Koch 雪花

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

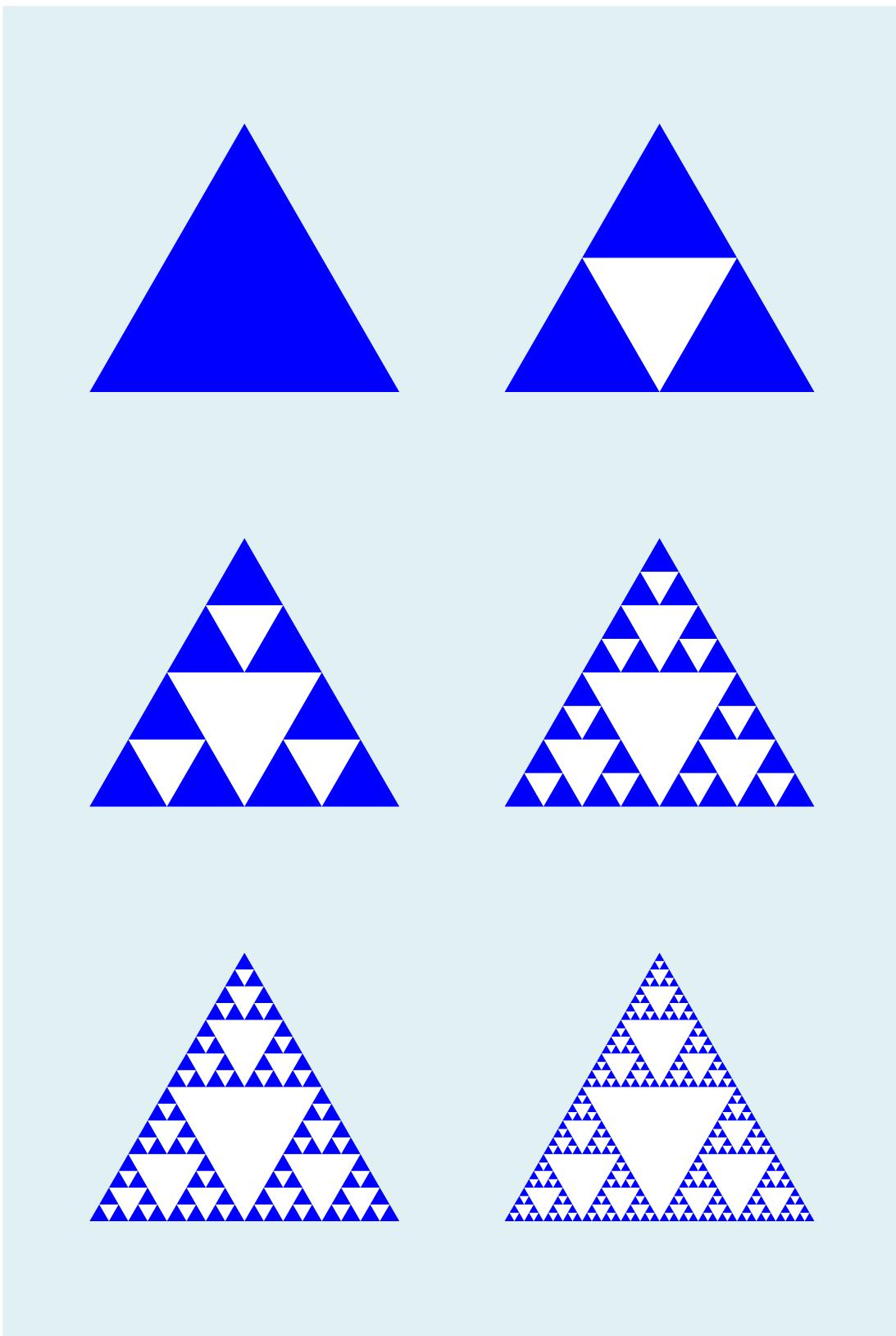


图 2. 谢尔宾斯基三角形

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

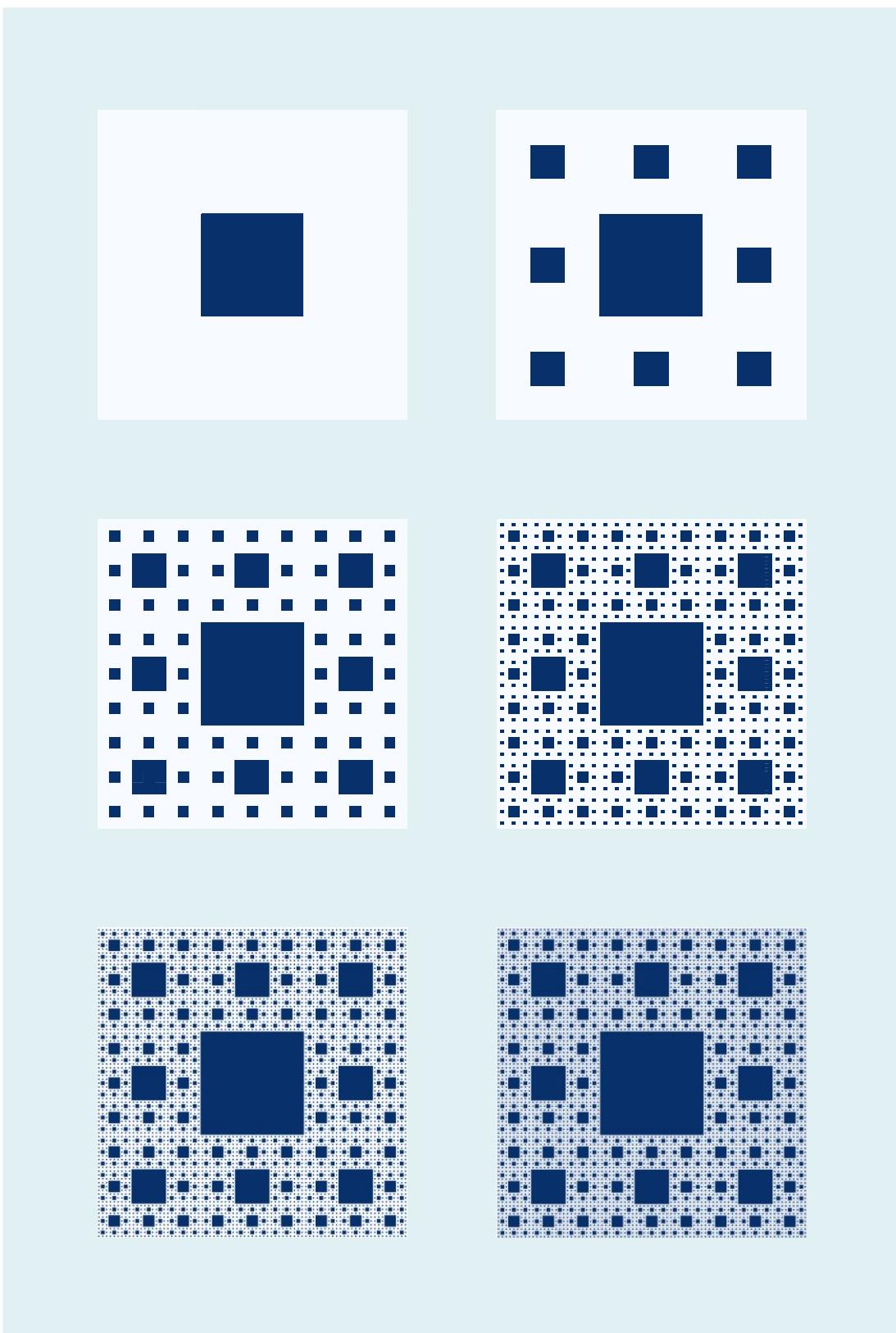


图 3. 谢尔宾斯基地毯

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

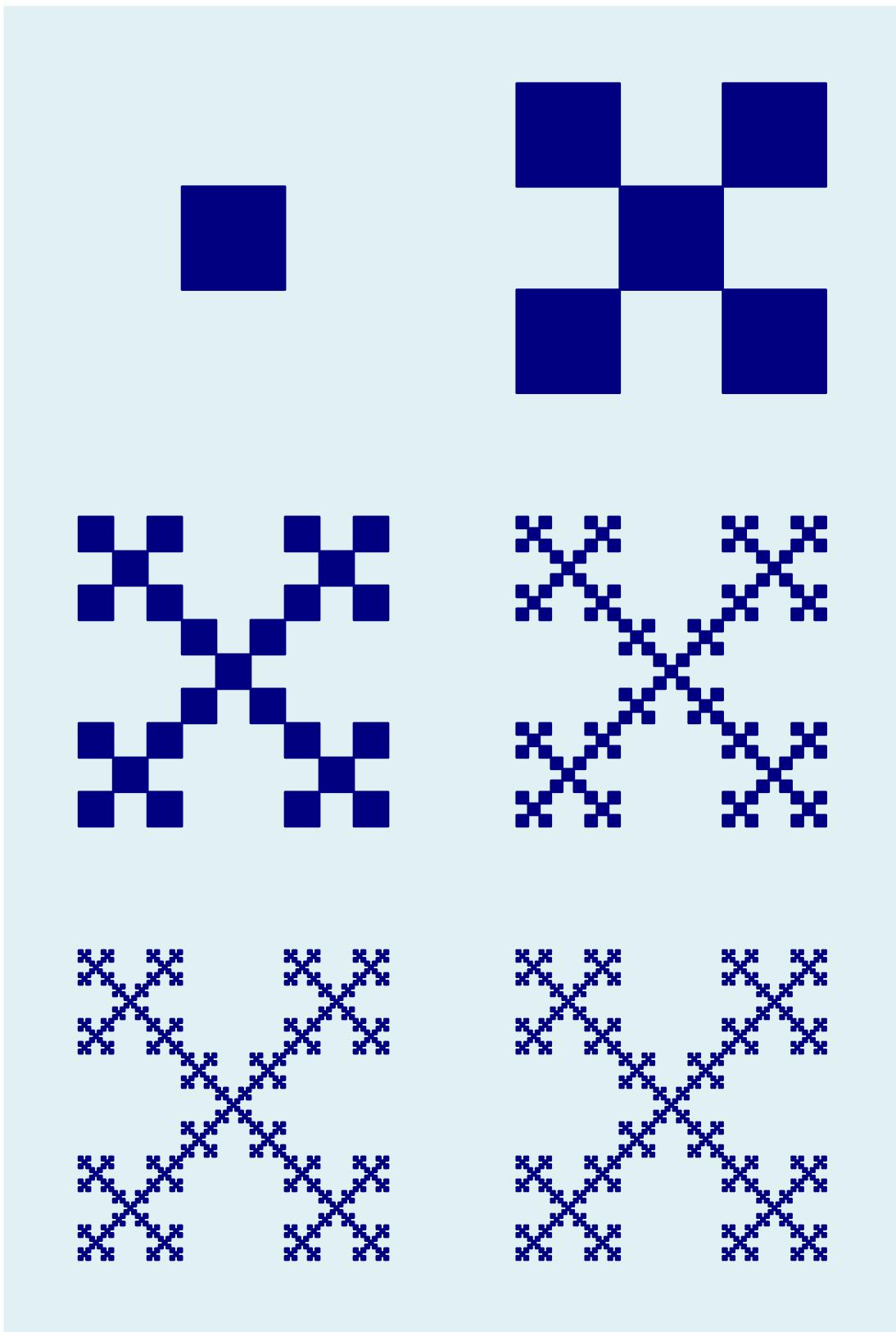


图 4. Vicsek 正方形分形，第 1 组

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

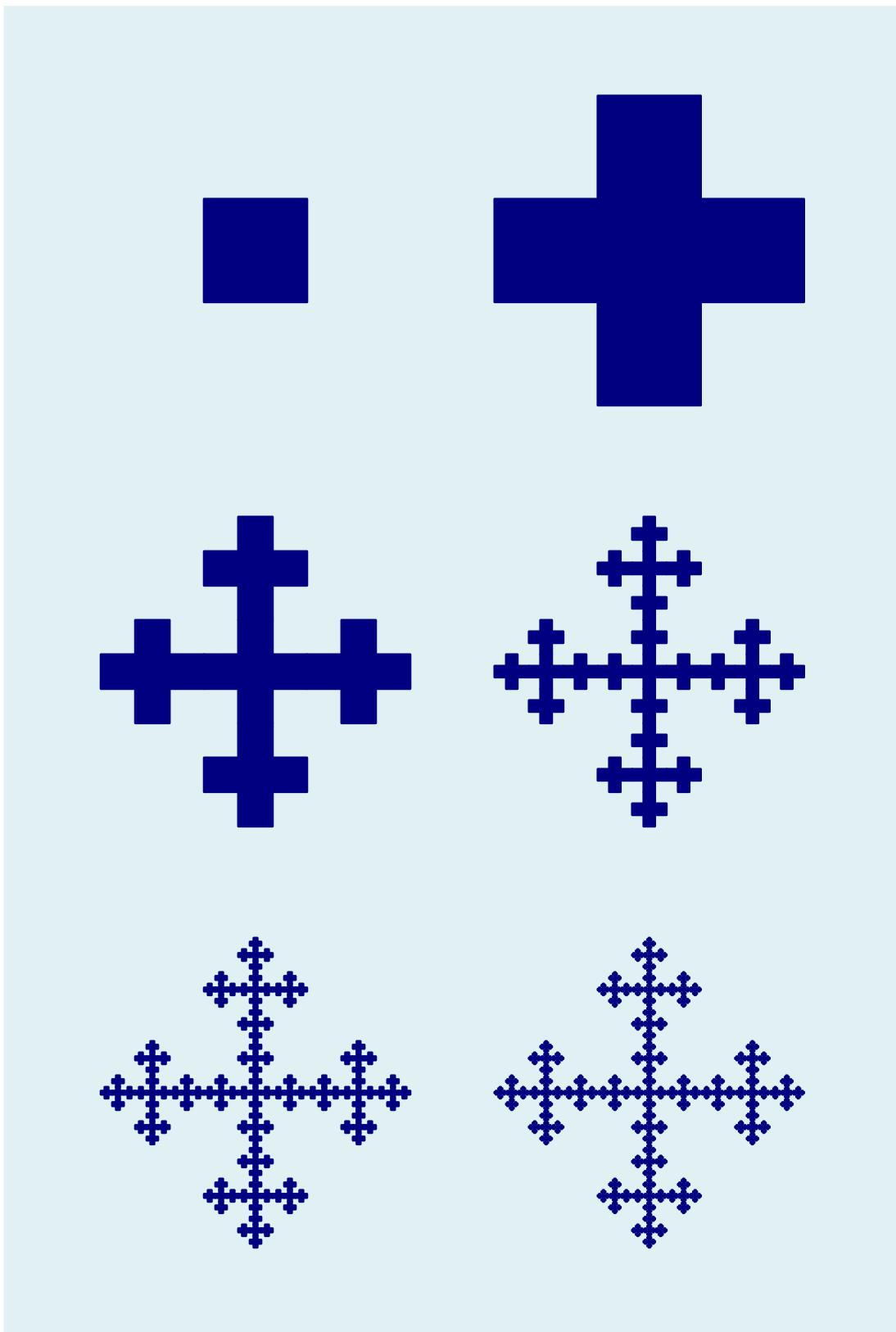


图 5. Vicsek 正方形分形，第 2 组

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

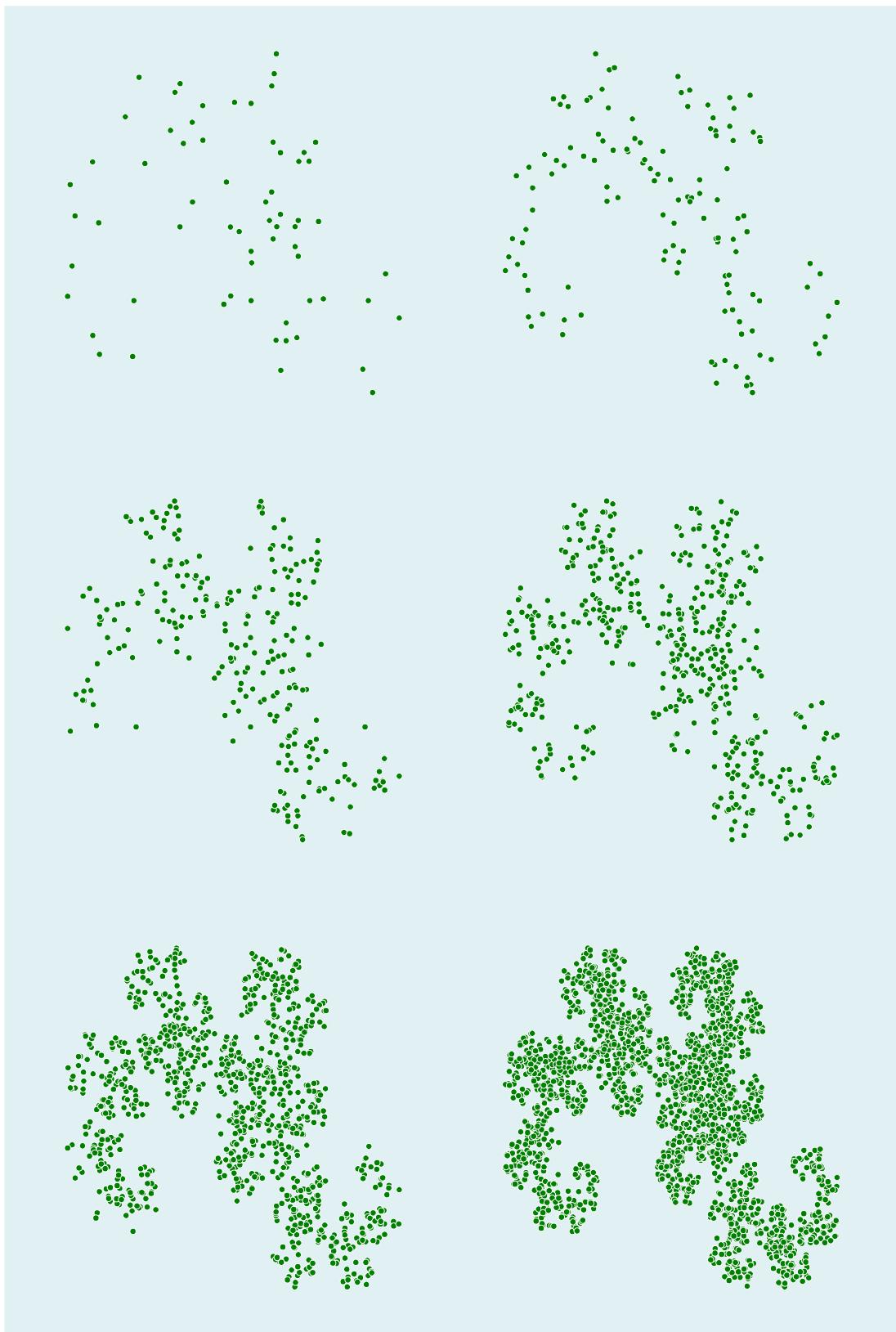


图 6. 龙曲线，基于随机数

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

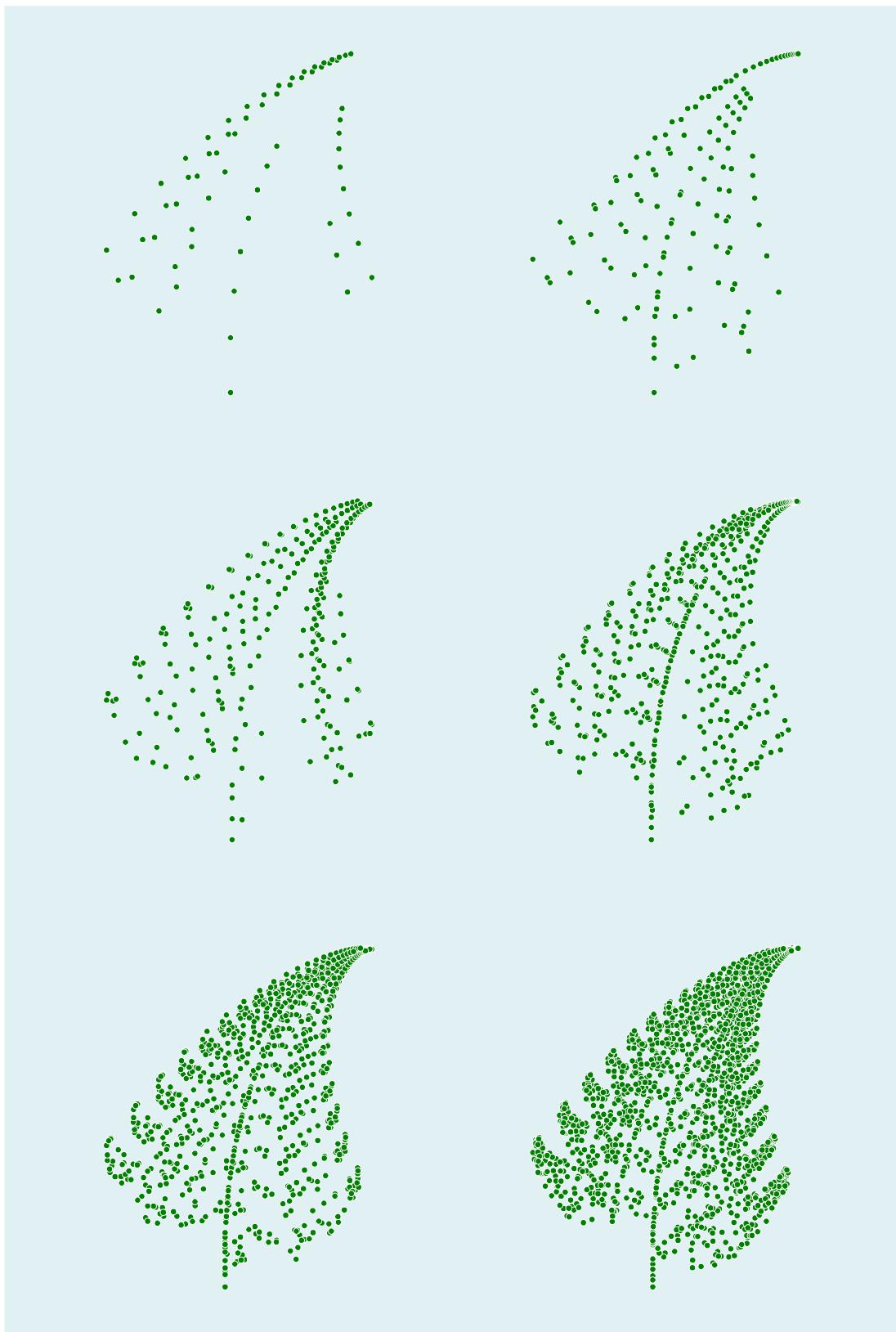


图 7. 巴恩斯利蕨，基于随机数

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

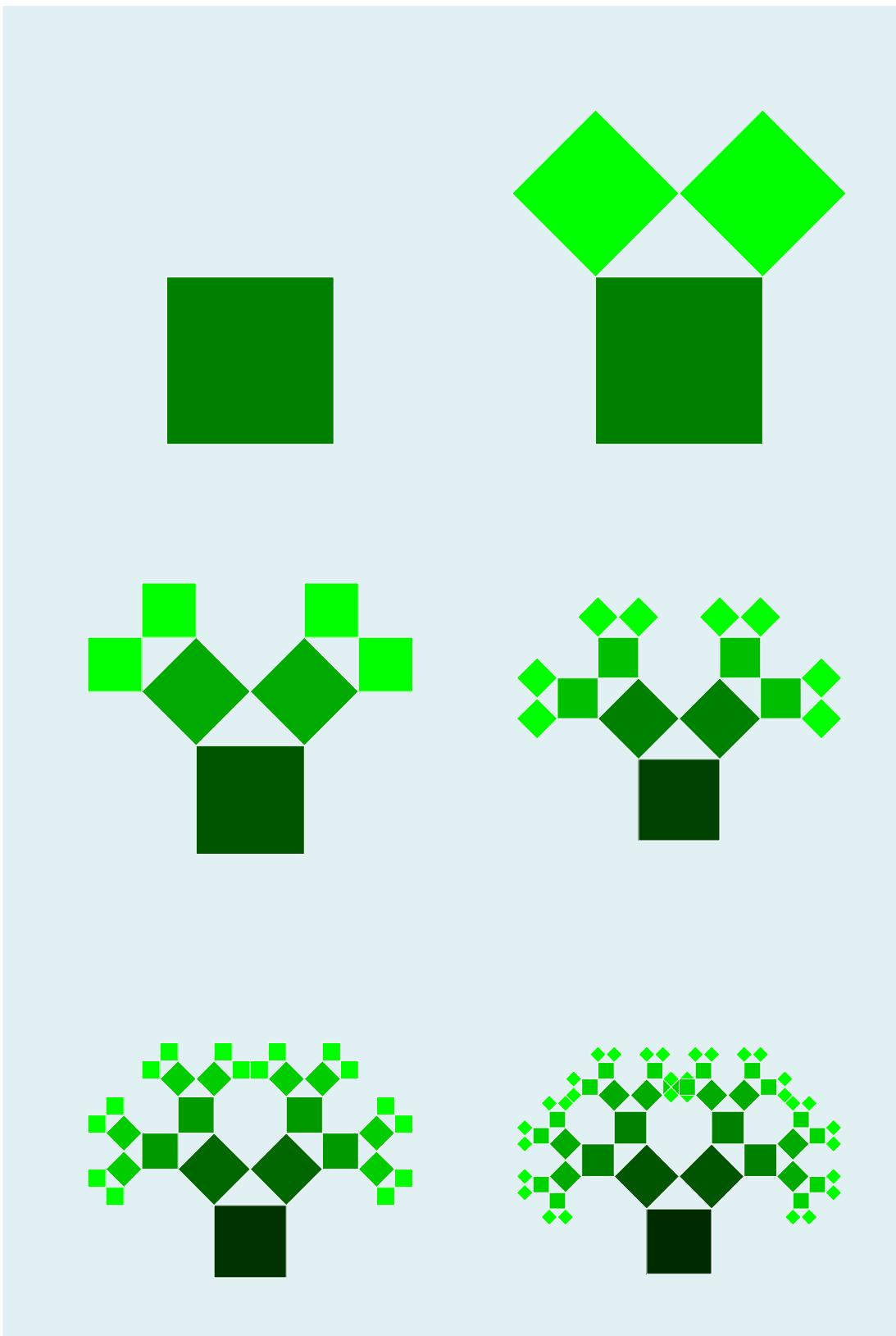


图 8. 勾股树

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

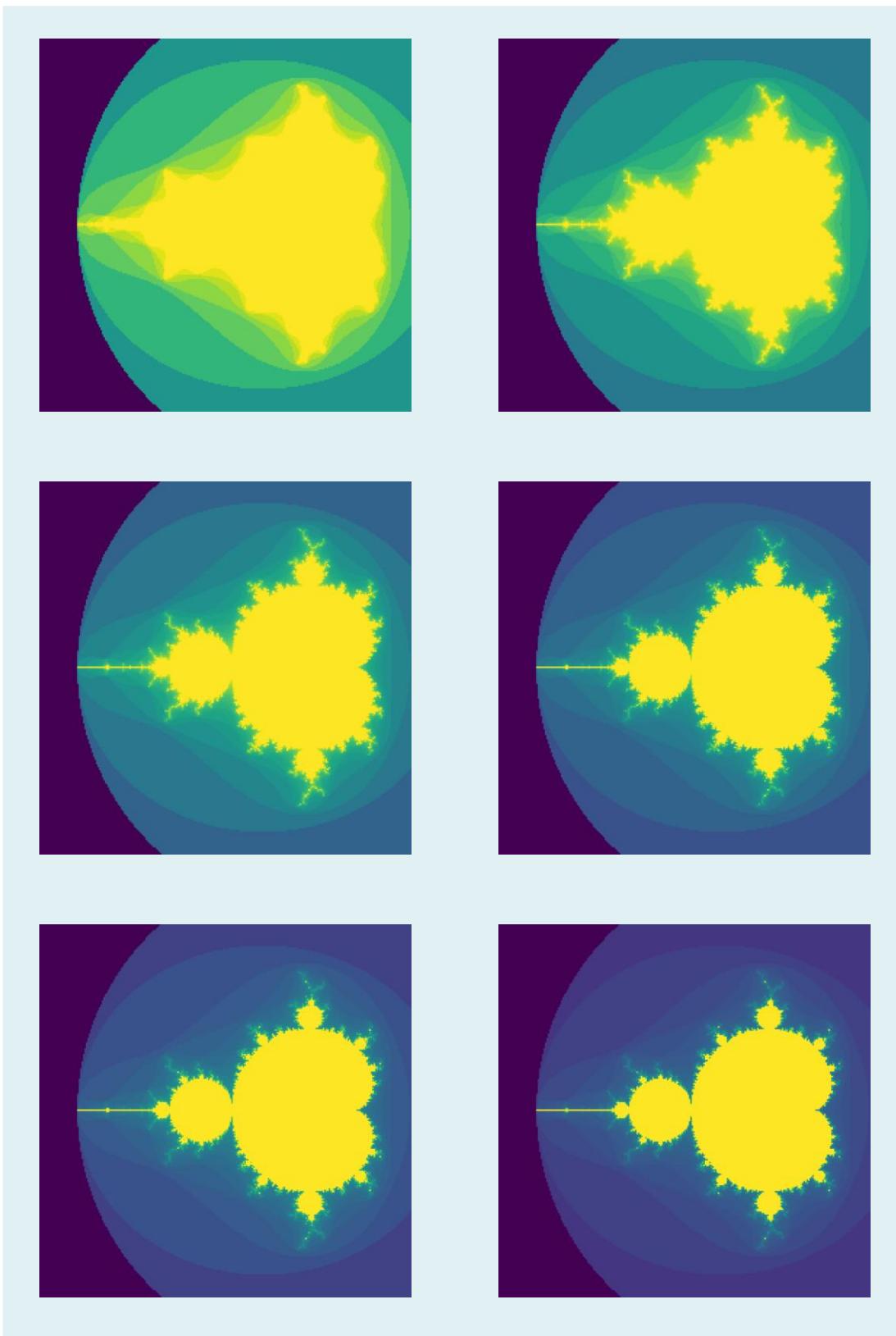


图 9. 曼德博集合

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

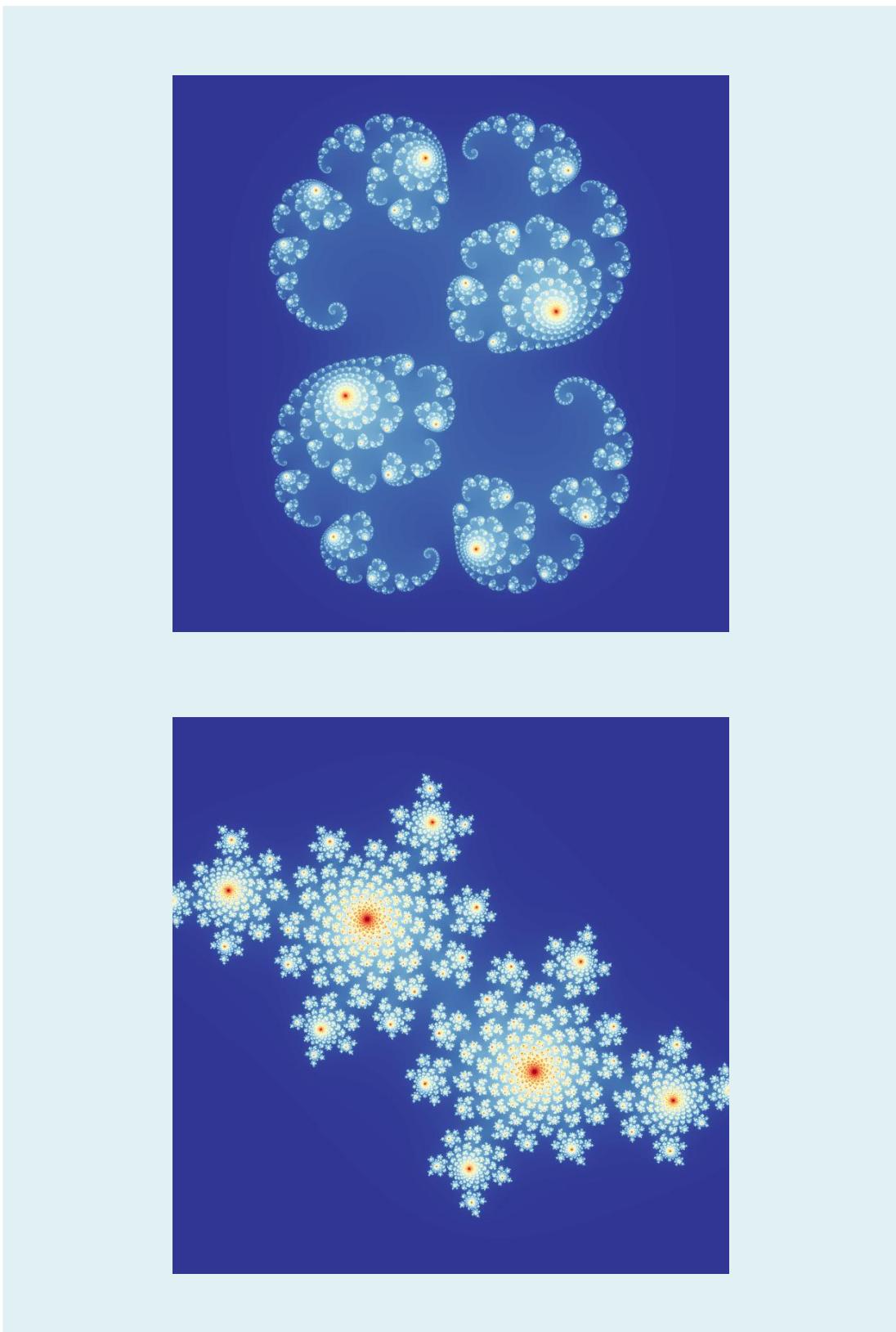


图 10. 朱利亚集合，第 1 组

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

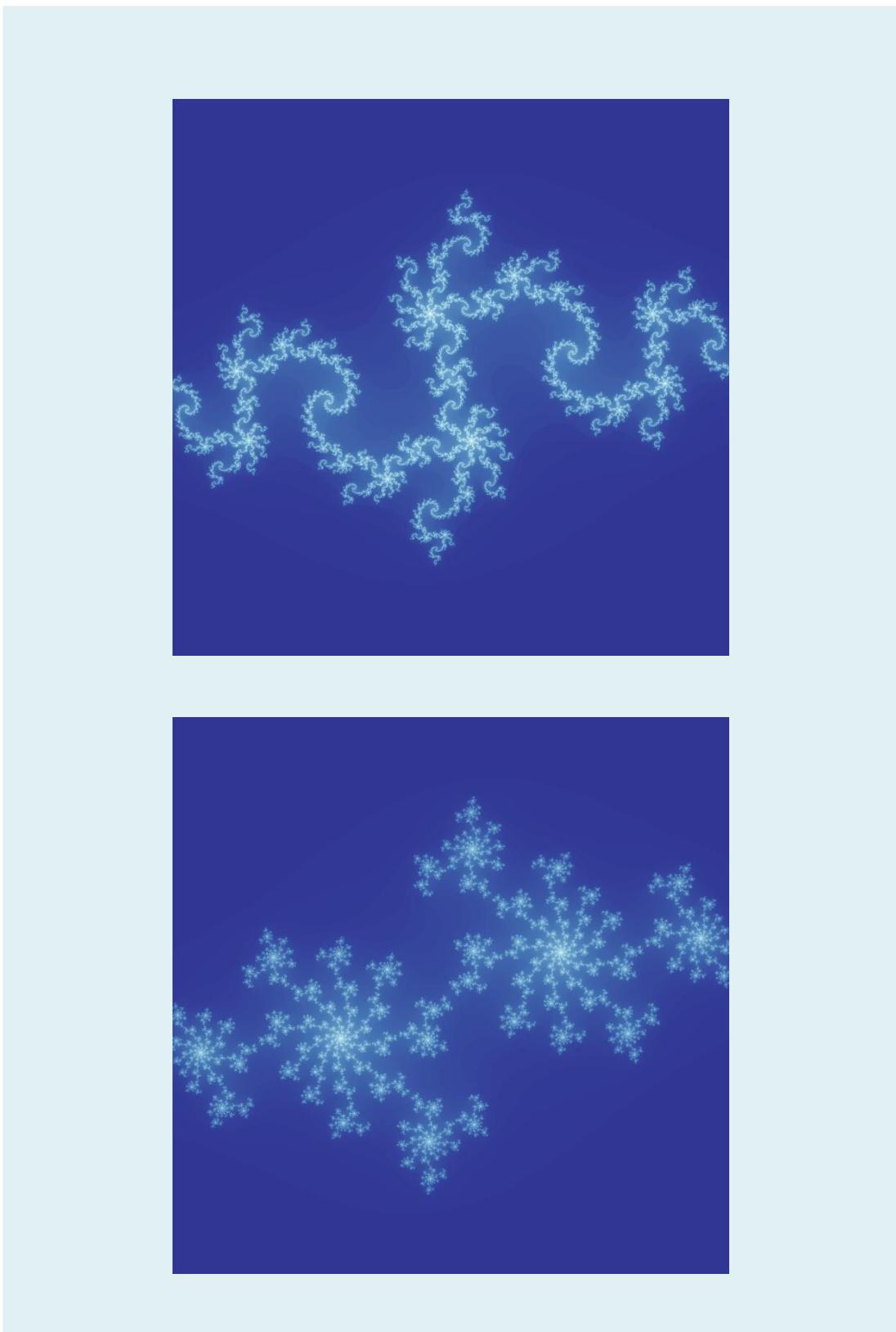


图 11. 朱利亚集合，第 2 组

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

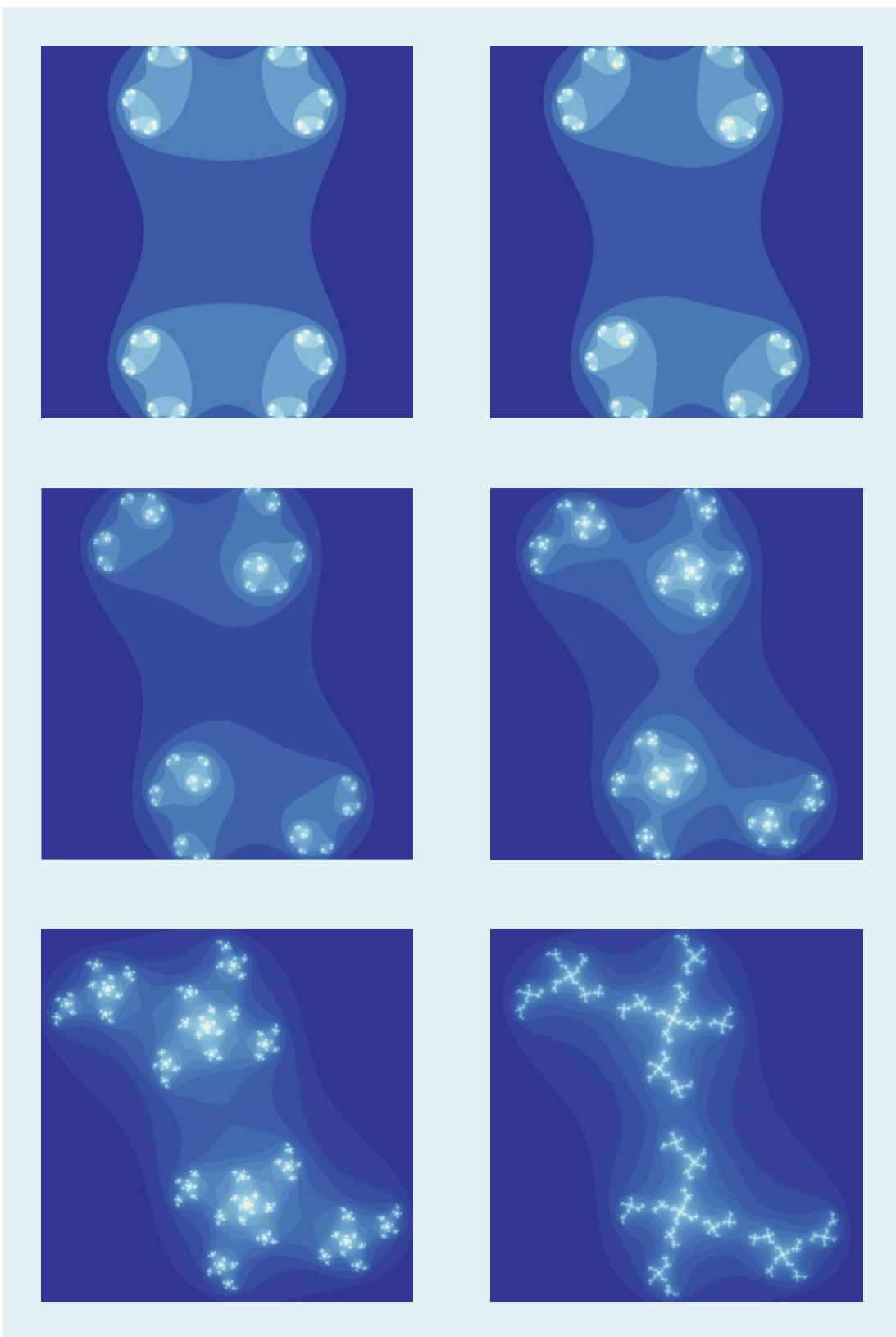


图 12. 朱利亚集合，第 3 组

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

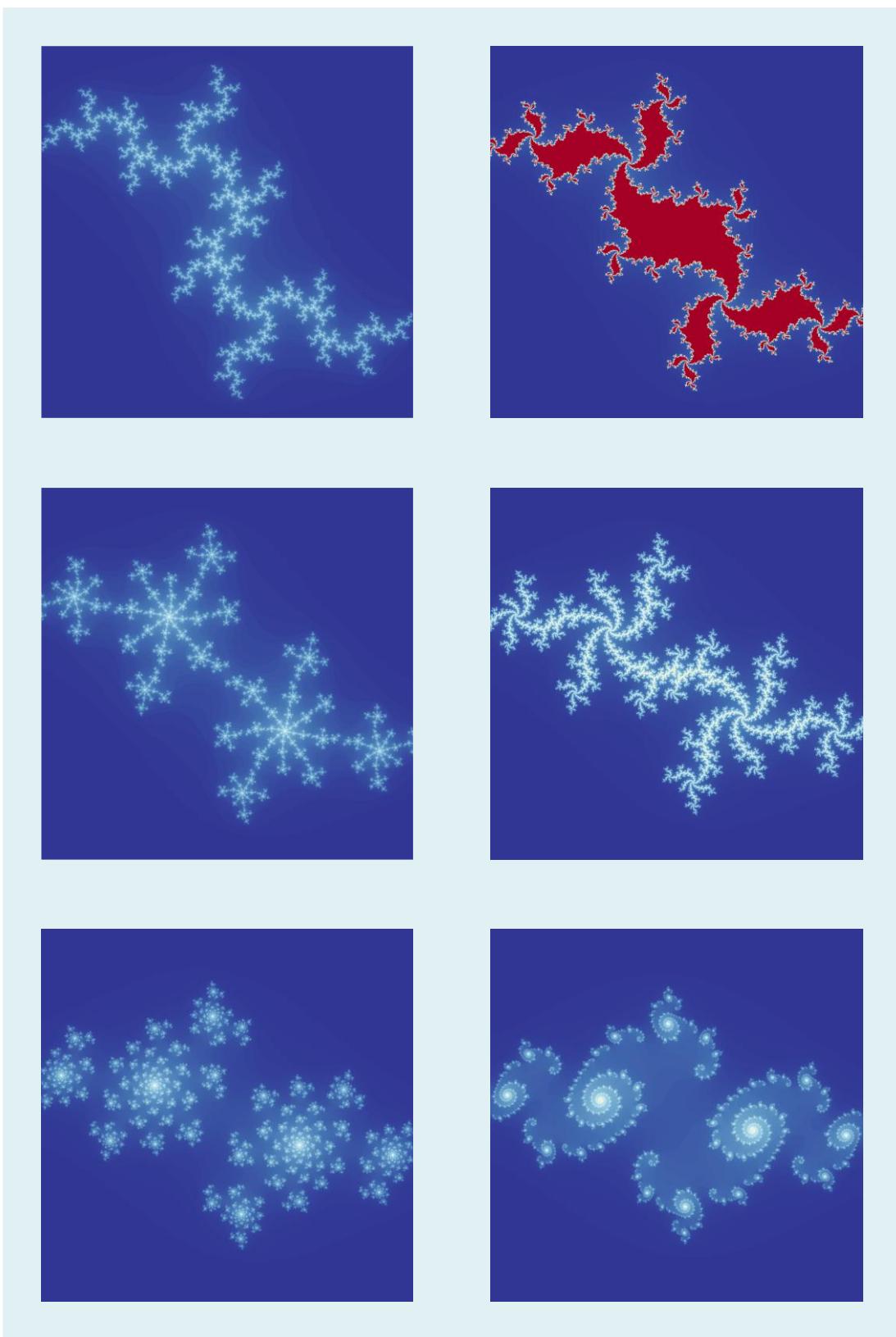


图 13. 朱利亚集合，第 4 组

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

34

Network Diagrams

网络图

可视化网络的层级结构和信息路径



艺术不是工艺品，艺术传递艺术家的体验。

Art is not a handicraft, it is the transmission of feeling the artist has experienced.

——列夫·托尔斯泰 (Leo Tolstoy) | 俄国作家 | 1828 ~ 1910

34.1 网络图

在深度学习中，网络图（network diagram）是指描述神经网络架构和连接关系的图形表示。它是一种可视化工具，用于展示神经网络中的各个层以及它们之间的连接。

网络图通常由节点（node）和边（edge）组成。节点代表神经网络的层或单个神经元，而边表示它们之间的连接或信息传递。节点之间的连接可以是有向的或无向的，这取决于神经网络的结构。

通过网络图，可以清晰地展示神经网络的层级结构和信息流动的路径。它可以帮助人们理解和分析神经网络的架构，以及如何在不同层之间传递和转换数据。

网络图还可以用于可视化神经网络的参数和权重，以及它们在不同层之间的传递。这有助于深入了解神经网络的工作原理，以及每个层对输入数据的处理方式。

NetworkX 是一个用于创建、操作和研究复杂网络的 Python 软件包。它提供了一种灵活且高效的方式来表示、分析和可视化各种类型的网络结构，包括社交网络、生物网络、物流网络等等。

NetworkX 具有丰富的功能，可用于执行各种网络分析任务。它支持创建网络的节点和边，并提供了各种算法和方法来研究网络的特性和行为。通过 NetworkX，大家可以进行节点和边的属性设置、图形布局、路径查找、子图提取、连通性分析、社区检测以及中心性和影响力度量等。

NetworkX 的官方网站上提供了丰富的示例和图库，展示了该软件包在各种应用领域的应用。图 1 ~ 图 10 所示为精选出来的几个例子，下面逐一简介。

图 1 上图展示的是多层网络。图 1 下图可视化 13 个节点的循环布局图。边根据节点之间的距离着色。节点之间的距离是指在圆上的任意两个节点之间沿弧线穿越的最小节点数。

图 2 上图为随机几何位置的无向图，下图为 128 个美国城市人口和距离组成的无向图。无向图是一种图论中的基本概念，表示节点之间的双向连接关系。它由一组节点和连接这些节点的边组成，边没有方向性，即两个节点之间的连接是相互的。

图 3 可视化互联网上的 186 个站点到洛斯阿拉莫斯国家实验室的路由 LANL Routes 信息。

旅行推销员问题（Traveling Salesman Problem, TSP）是一种著名的组合优化问题，属于图论领域。问题的目标是找到一条最短路径，使得旅行推销员能够访问一组给定的城市，并返回起始城市，同时经过每个城市一次且仅一次。

如图 4 所示，每个节点表示一个城市，边的权重表示城市之间的距离或成本。旅行推销员问题要求找到一条路径，使得旅行者从起始城市出发，经过所有其他城市，最后回到起始城市，并使得路径的总长度最小。图 4 中红色线代表最优路径。

图 5 可视化 1886 年至 1985 年的所有 685 场世界国际象棋锦标赛比赛参赛者、赛事、成绩。边宽度代表对弈的数量，点的大小代表获胜棋局数量。

图 6 所示为 5757 个 5 个字母的单词上生成一个无向图。如果两个单词在一个字母上不同，它们之间就会有一条边。

图 7 上图为随机几何位置的无向图，图 7 下图为 128 个美国城市人口和距离组成的无向图。

图 8 所示为 16 种可能的三元组类型。三元组类型是指在社交网络或其他网络中，根据节点之间的连接关系，将节点组合成不同类型的三元组。三元组由三个节点组成，它们之间存在特定的连接模式。图中，前三位数字表示相互、非对称和空值二元组（双向、单向和非连接边）的数量，字母表示方向，分别是向上 (U)、向下 (D)、循环 (C) 或传递 (T)。

图 9 可视化具有最多 6 个节点的所有连通图的图谱。

图 10 展示基因之间的介数中心性 (Betweenness Centrality)，它使用了 WormNet v.3-GS 数据来测量基因之间的正向功能关联。WormNet 是一个用于研究秀丽隐杆线虫 (*Caenorhabditis elegans*) 基因功能网络的资源。

以上这些例子对应的代码、数据可以在如下链接找到：

https://networkx.org/documentation/stable/auto_examples/index.html

<https://github.com/networkx/networkx/tree/main/examples>

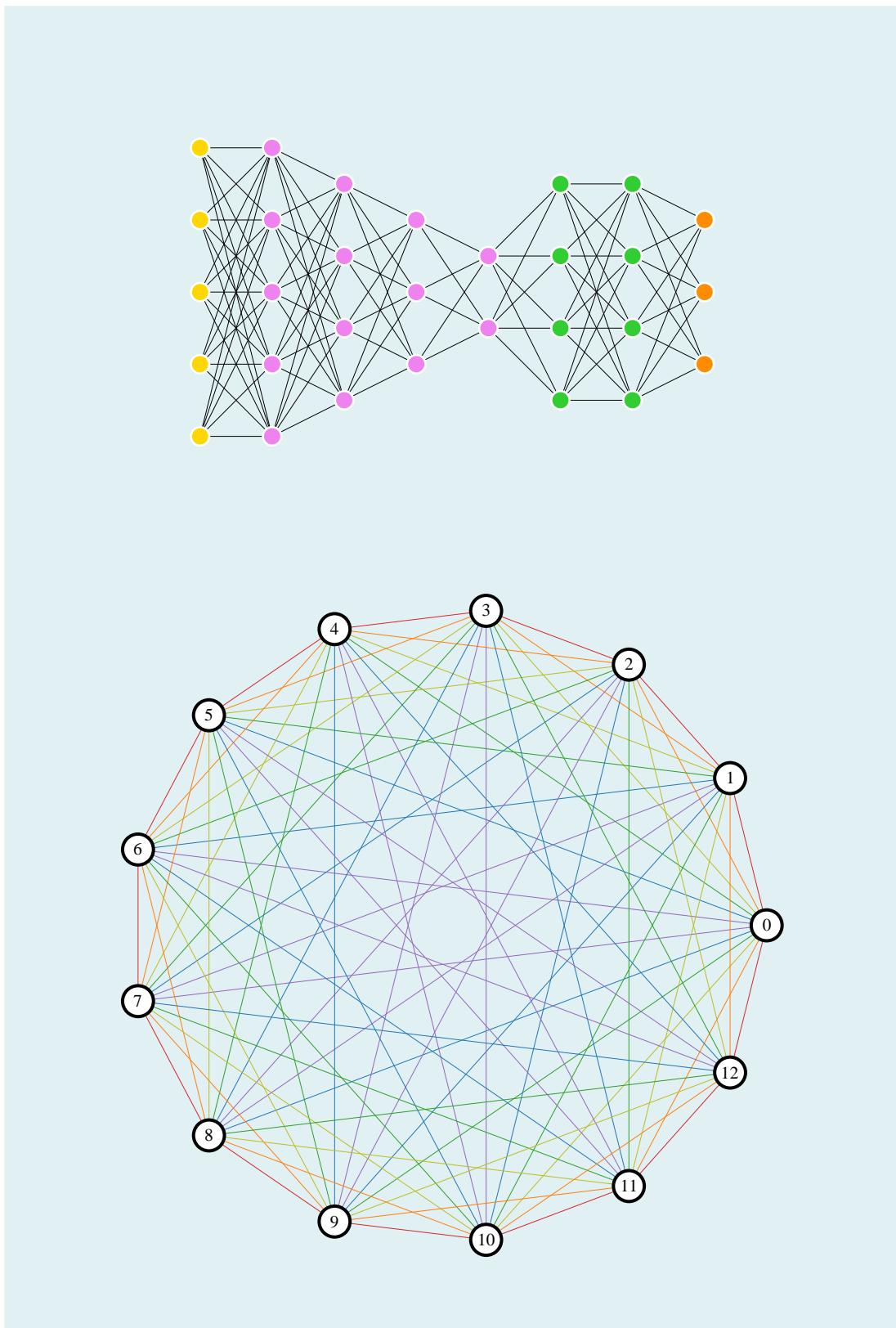


图 1. 上图展示的是多层网络，下图可视化 13 个节点的循环布局图

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

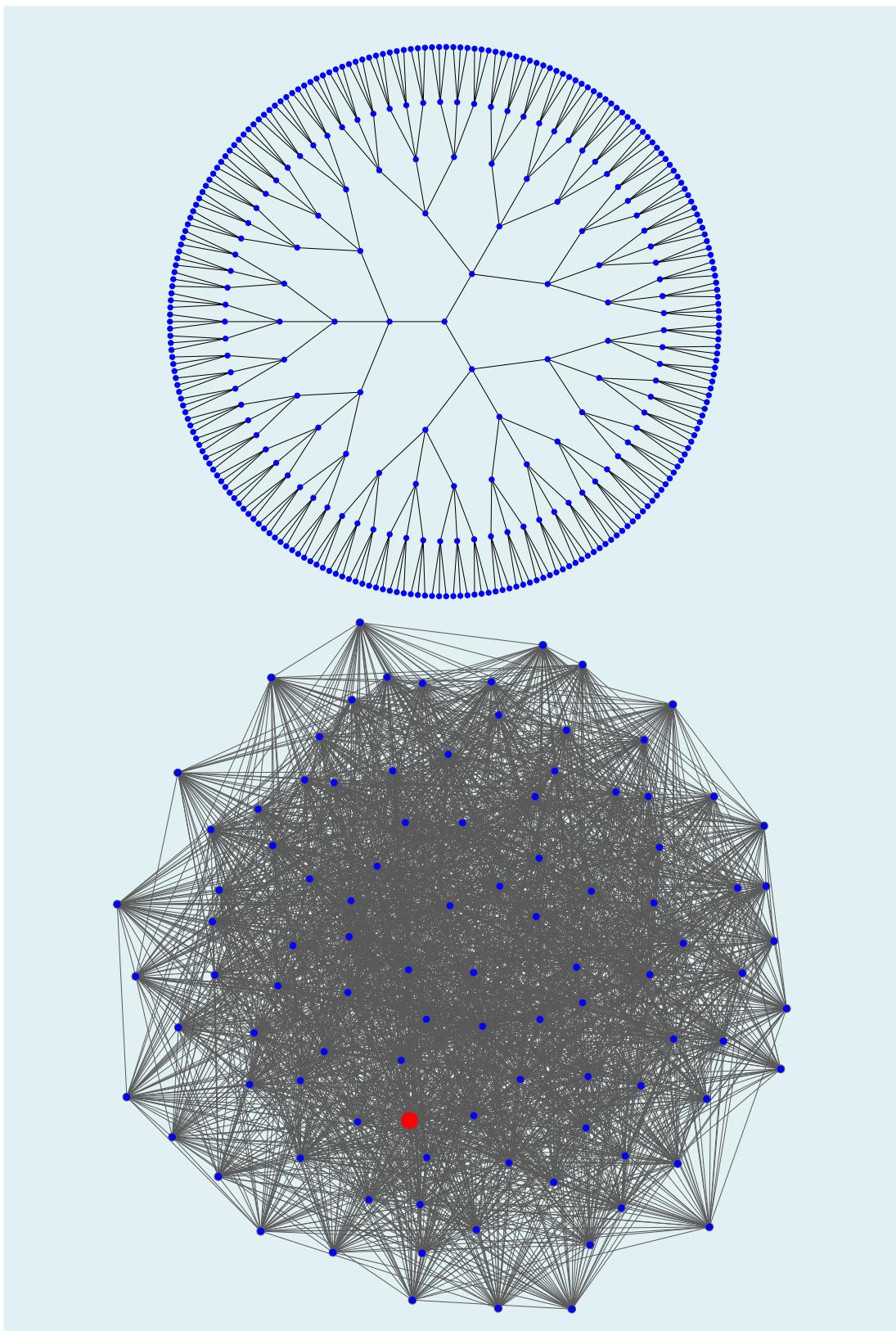


图 2. 下图可视化了一种搜索节点算法

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com



图 3. 互联网上的 186 个站点到洛斯阿拉莫斯国家实验室的路由 LANL Routes 信息

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

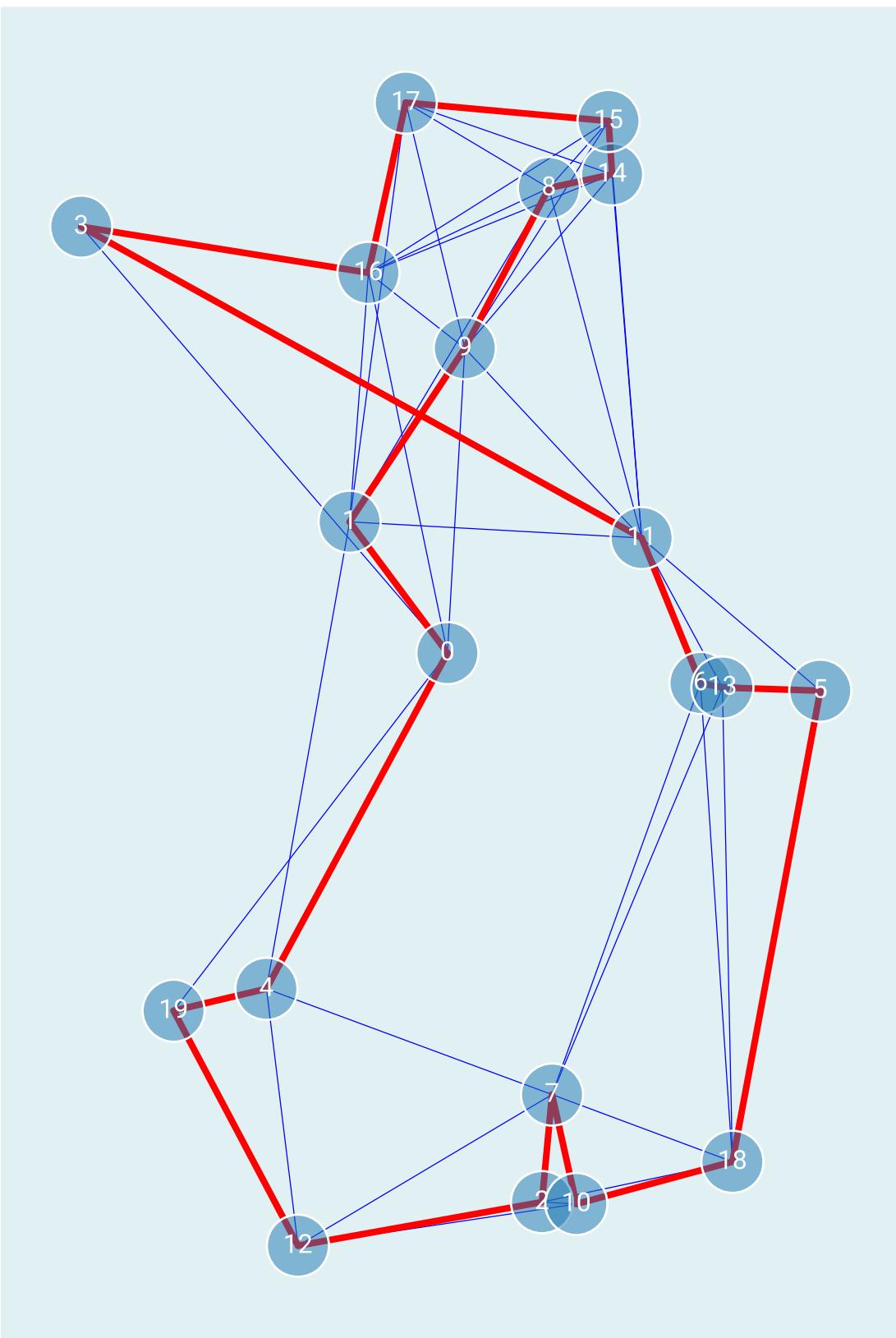


图 4. 销售员问题

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

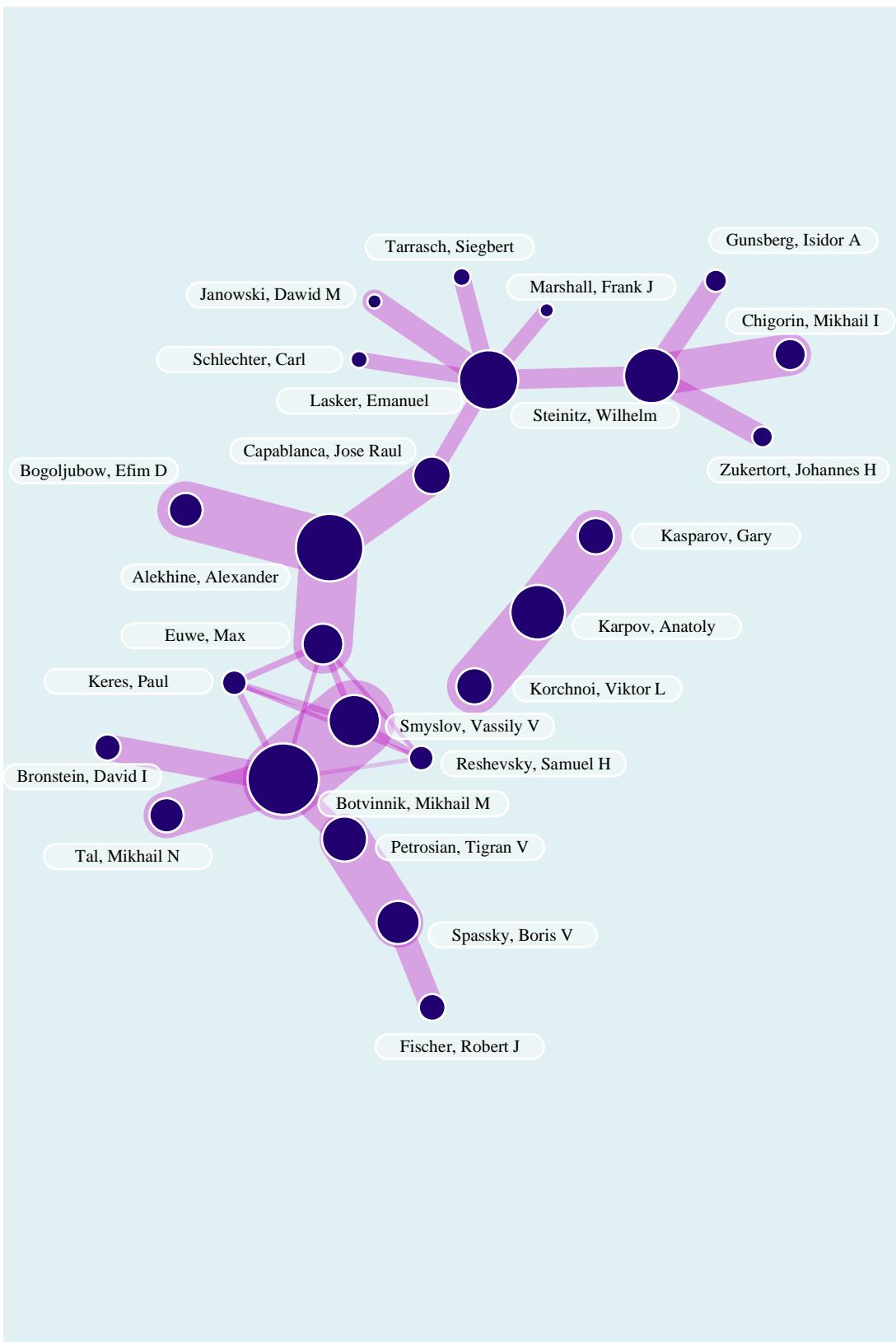


图 5. 可视化 1886 年至 1985 年的所有 685 场世界国际象棋锦标赛比赛参赛者、赛事、成绩

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

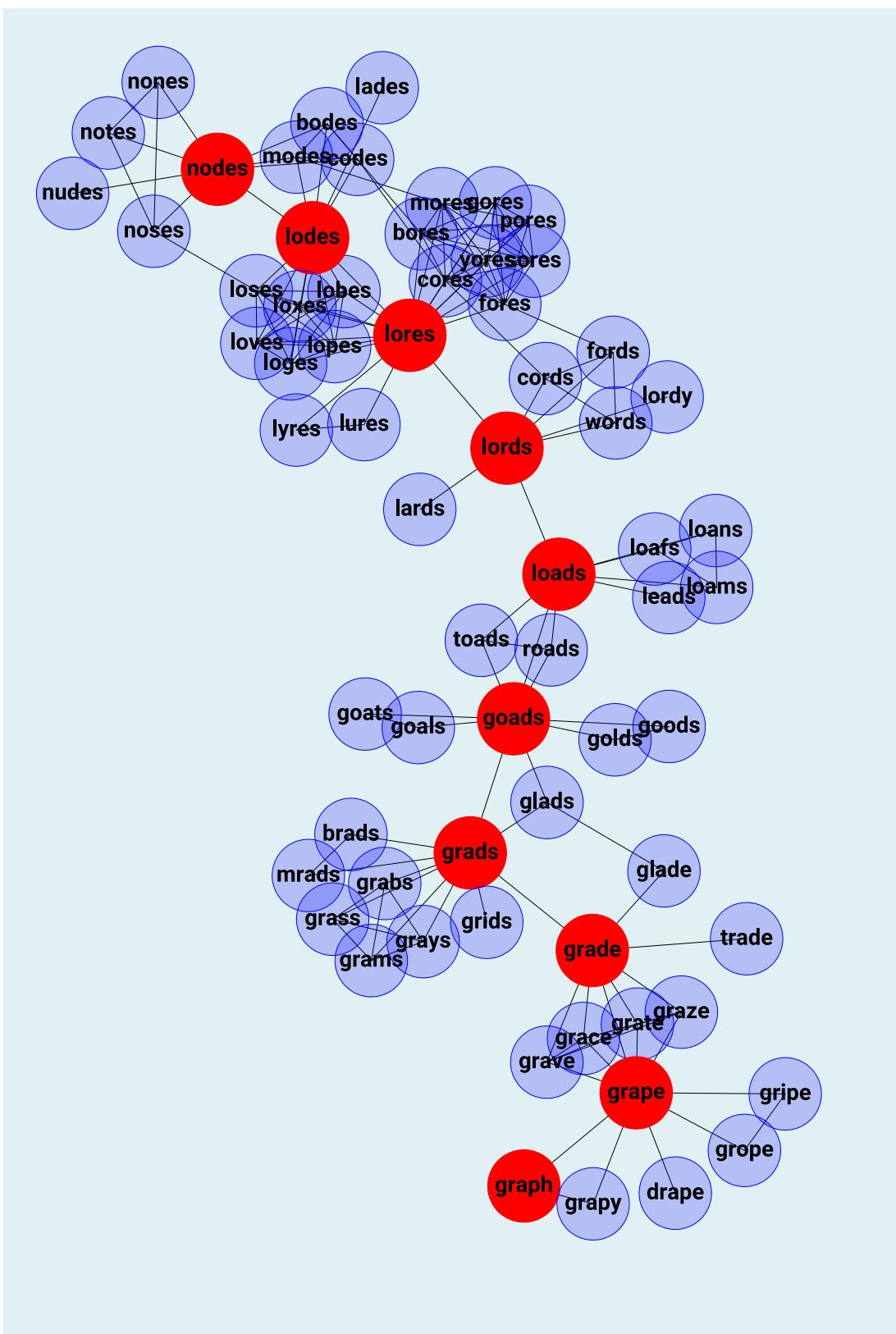


图 6. 5757 个 5 个字母的单词上生成一个无向图。如果两个单词在一个字母上不同，它们之间就会有一条边

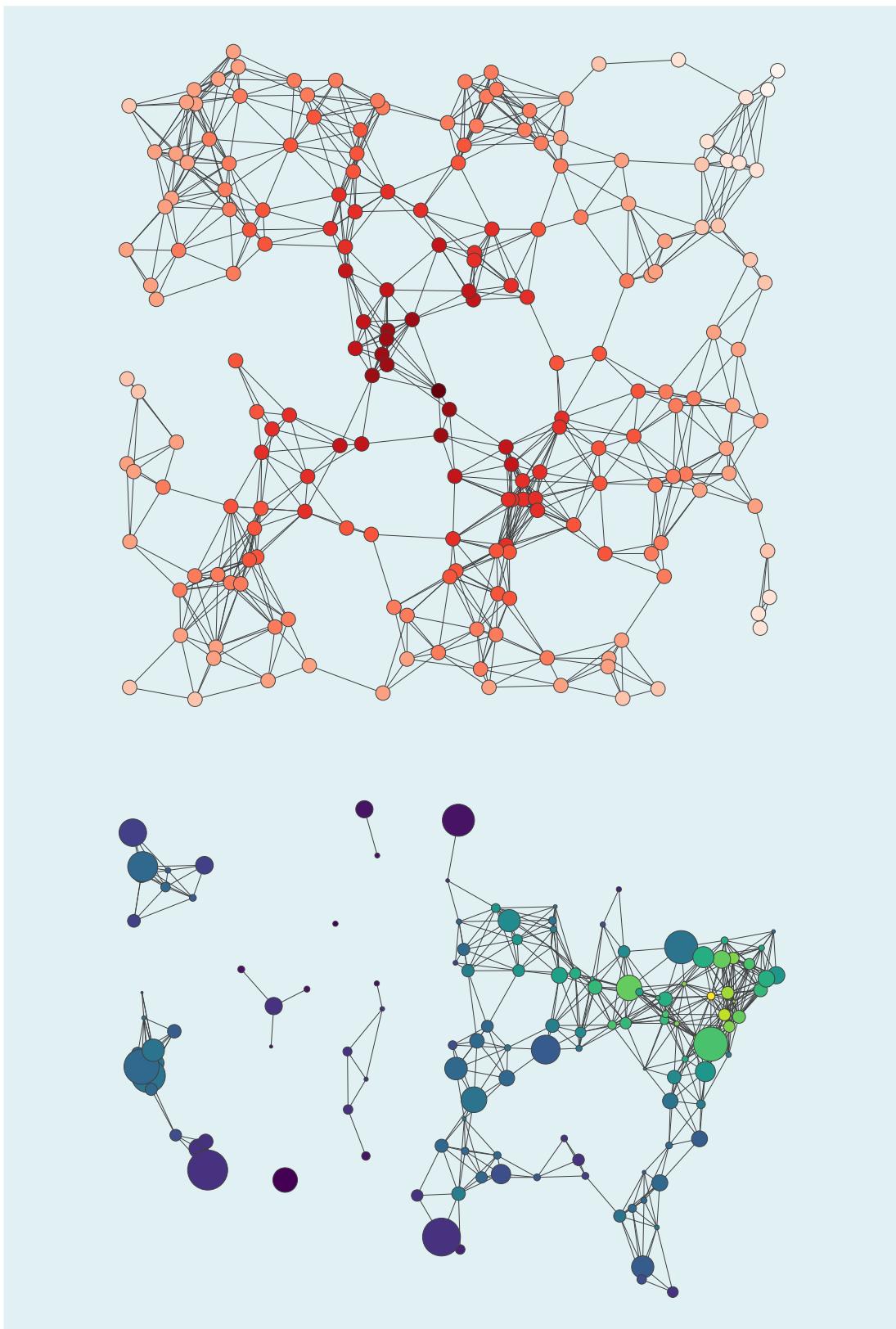


图 7. 上图为随机几何位置的无向图，下图为 128 个美国城市人口和距离组成的无向图

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

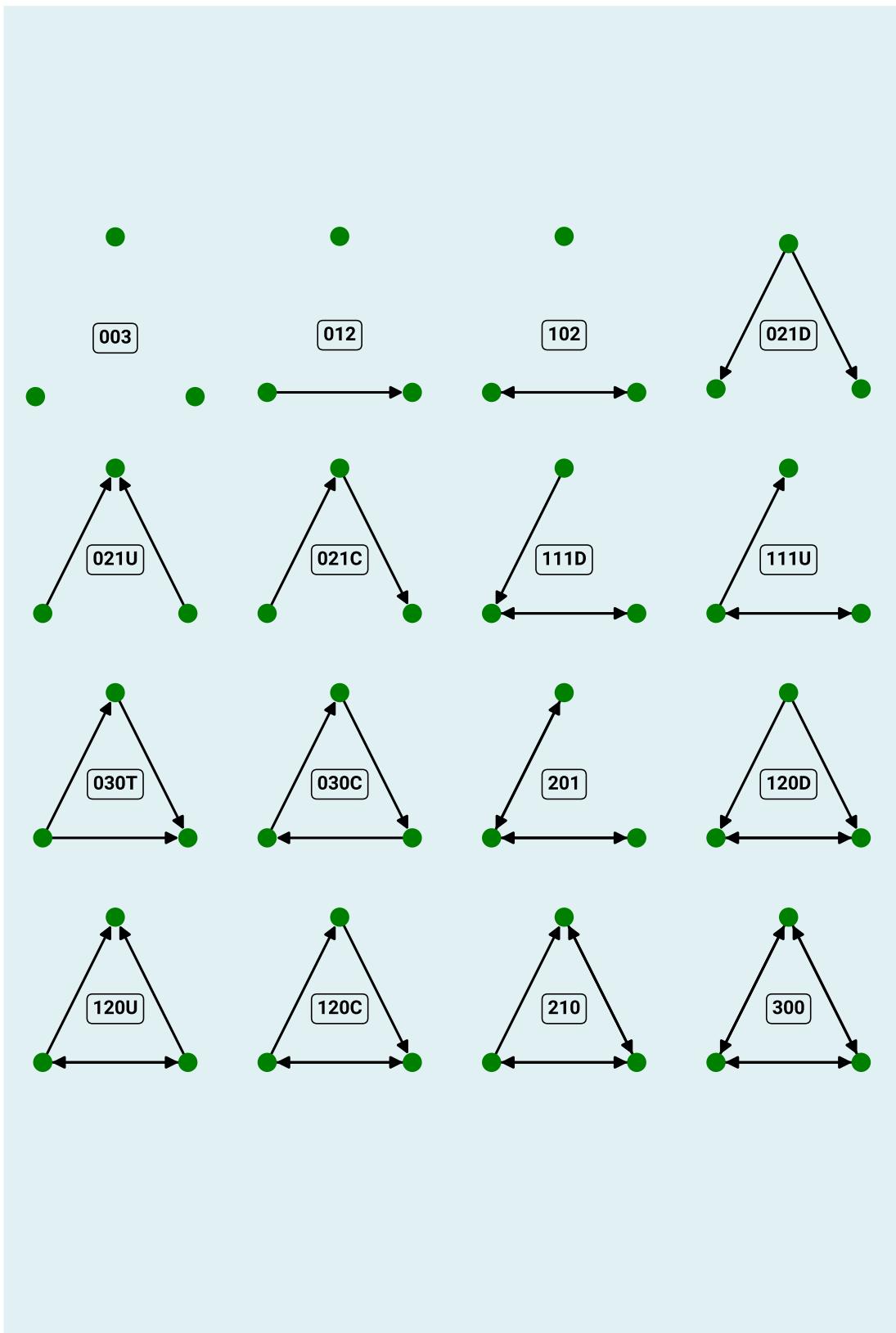


图 8.16 种可能的三元组类型

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

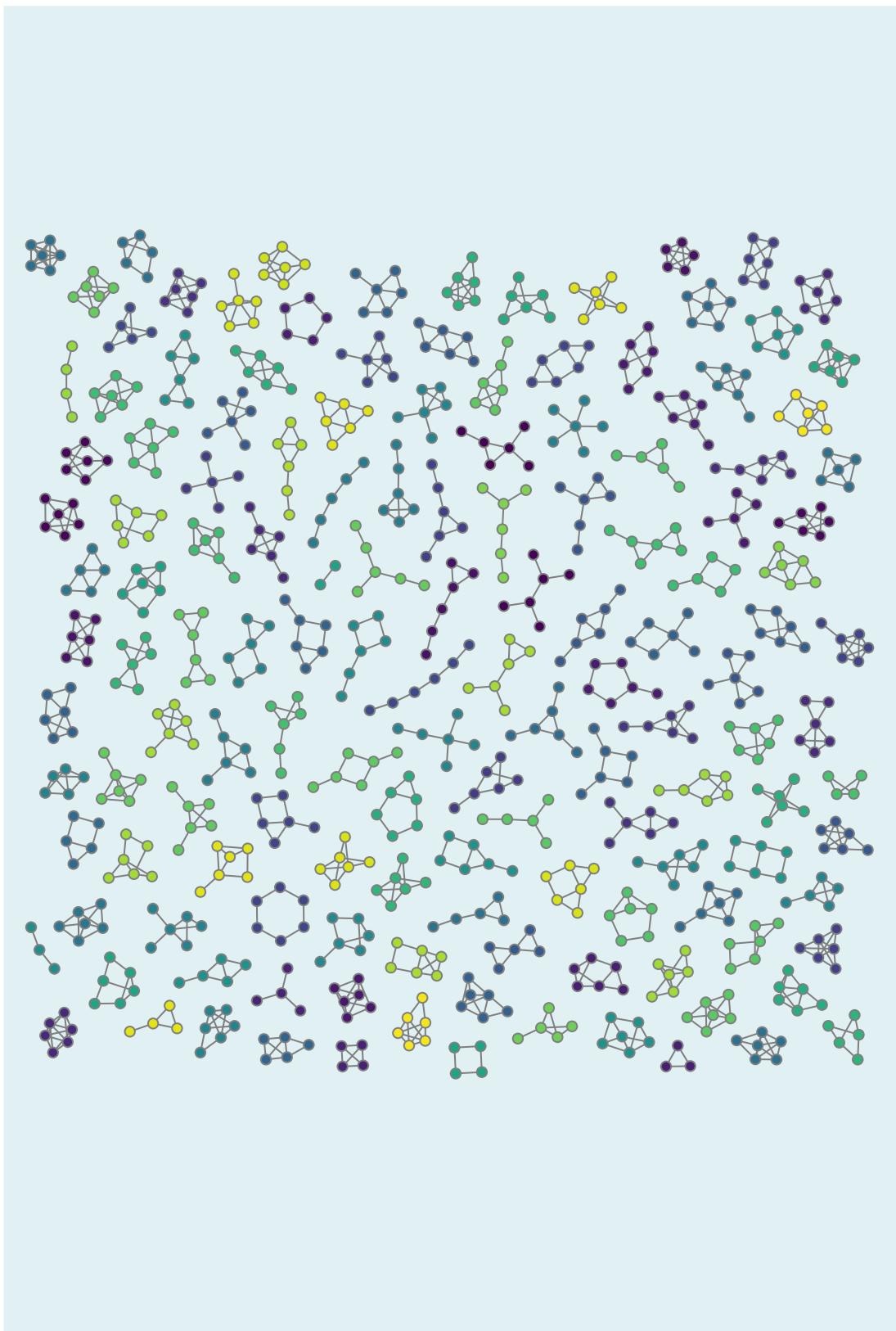


图 9. 最多 6 个节点的所有连通图的图谱

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

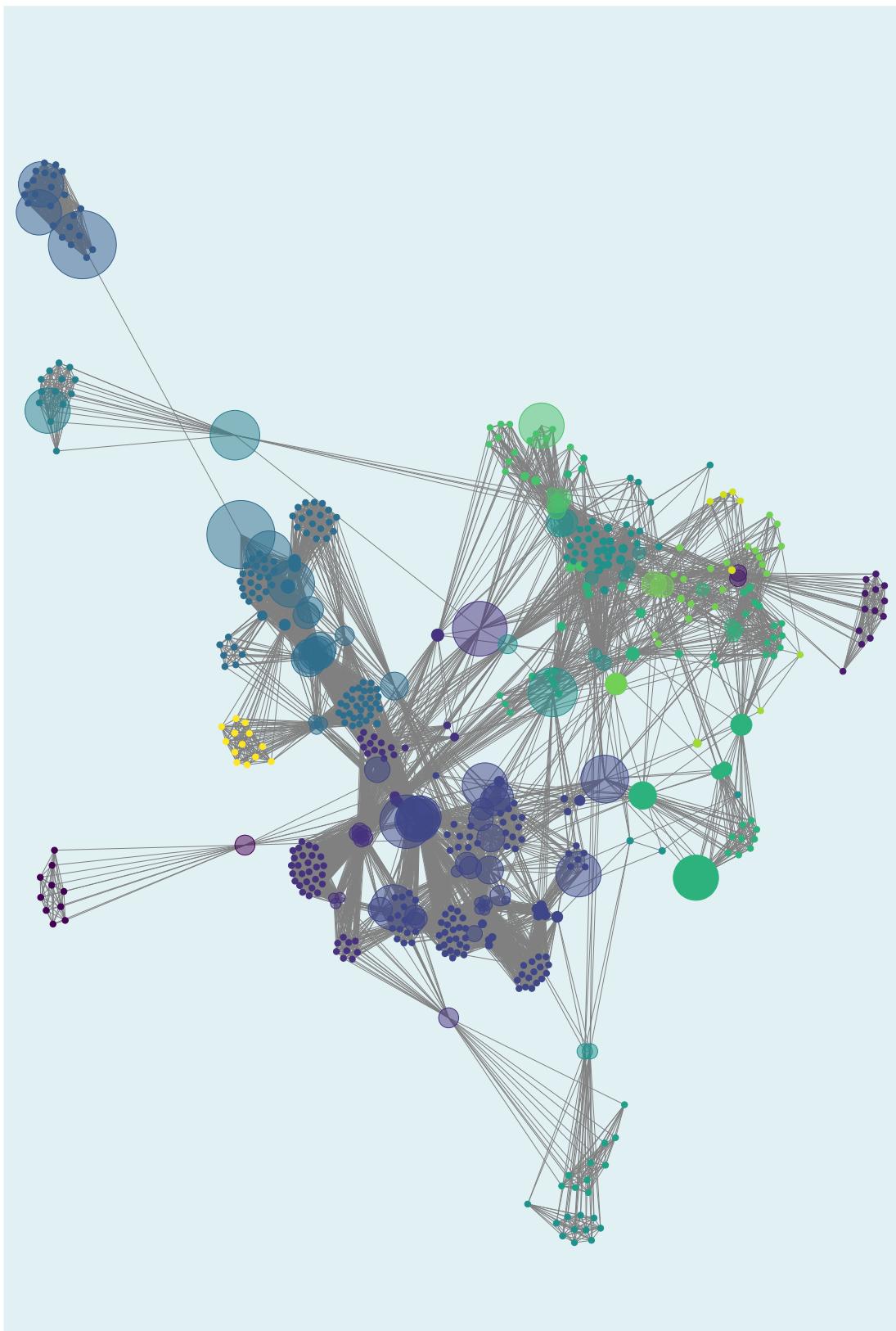


图 10. WormNet v.3-GS 数据来测量基因之间的正向功能关联

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

35

Distributions of Sample Data

数据分布

先验分布、证据因子、后验分布



青春不是芳华年少，而是一种心境；青春不是桃花红颜、盈盈朱唇、轻柔体态，而是积极的心志，丰富的想象，炙热的感情；青春是充满生机、清新盎然的生命源泉。

Youth is not a time of life; it is a state of mind; it is not a matter of rosy cheeks, red lips and supple knees; it is a matter of the will, a quality of the imagination, a vigor of the emotions; it is the freshness of the deep springs of life.

—— 塞缪尔·厄尔曼 (Samuel Ullman) | 美国诗人 | 1840 ~ 1924



- ◀ `matplotlib.pyplot.contour3D()` 绘制三维等高线图
- ◀ `matplotlib.pyplot.contourf()` 绘制平面填充等高线
- ◀ `matplotlib.pyplot.fill_between()` 区域填充颜色
- ◀ `matplotlib.pyplot.plot_wireframe()` 绘制线框图
- ◀ `matplotlib.pyplot.scatter()` 绘制散点图
- ◀ `numpy.ones_like()` 用来生成和输入矩阵形状相同的全 1 矩阵
- ◀ `numpy.outer()` 计算外积，张量积
- ◀ `numpy.vstack()` 返回竖直堆叠后的数组
- ◀ `scipy.stats.gaussian_kde()` 高斯核密度估计
- ◀ `seaborn.kdeplot()` 绘制 KDE 概率密度估计曲线
- ◀ `statsmodels.api.nonparametric.KDEUnivariate()` 构造一元 KDE
- ◀ `statsmodels.nonparametric.kde.kernel_switch()` 更换核函数
- ◀ `statsmodels.nonparametric.kernel_density.KDEMultivariate()` 构造多元 KDE

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

35.1 引入新视角：贝叶斯推断

《编程不难》和本书前文介绍过常见数据分布的可视化方案，特别是利用 Seaborn 各种函数展示鸢尾花数据分布。本章引入一个全新视角，贝叶斯推断 (Bayesian statistical inference)，来观察、分析数据。

贝叶斯统计推断是一种基于贝叶斯定理的统计推断方法。它利用先验知识和观测数据来更新对未知量的信念或概率分布，并计算后验概率。

以下是与贝叶斯统计推断相关的一些关键概念。

贝叶斯定理描述了在已知先验概率和条件概率的情况下，如何计算后验概率。条件概率指在某个条件下某事件发生的概率。

似然概率 (likelihood probability) 是在特定条件下，观测数据出现的概率。如图 1、图 2 所示，给定鸢尾花的分类条件下，估计样本数据分布得到的结果便是似然概率。图 1 采用二元高斯分布估算似然概率，等高线为椭圆。图 2 采用高斯核函数估计似然概率。

本章中，证据因子 (evidence) 描述鸢尾花数据的分布，证据因子根据似然概率计算得到。具体计算方法请大家参考《统计至简》第 18、19 章。

后验概率 (posterior probability) 是在考虑了观测数据后，对未知量的概率分布进行更新得到的概率分布。

35.2 高斯分布

图 3、图 4、图 5 所示为利用二元高斯分布估算得到的三个似然概率结果。

图 6 为利用图 3、图 4、图 5 计算得到的证据因子结果。

图 7、图 8、图 9 所示为计算得到的后验概率结果。后验概率常用来分类决策。比如，给定某朵鸢尾花花萼长度为 5 cm、花萼宽度 3 cm 条件下，它最可能是哪一类鸢尾花？回答这个问题就可以用后验概率的具体值。

35.3 高斯核密度估计

图 10、图 11、图 12 为利用二元高斯核密度估计估算得到的三个似然概率结果。

图 13 为利用图 10、图 11、图 12 计算得到的证据因子结果。

图 14、图 15、图 16 所示为计算得到的后验概率结果。

本章不会展开讨论这些统计学概念。只给定性描述，不会定量计算。大家可以在《统计至简》一书中找到相关的数学工具具体介绍。

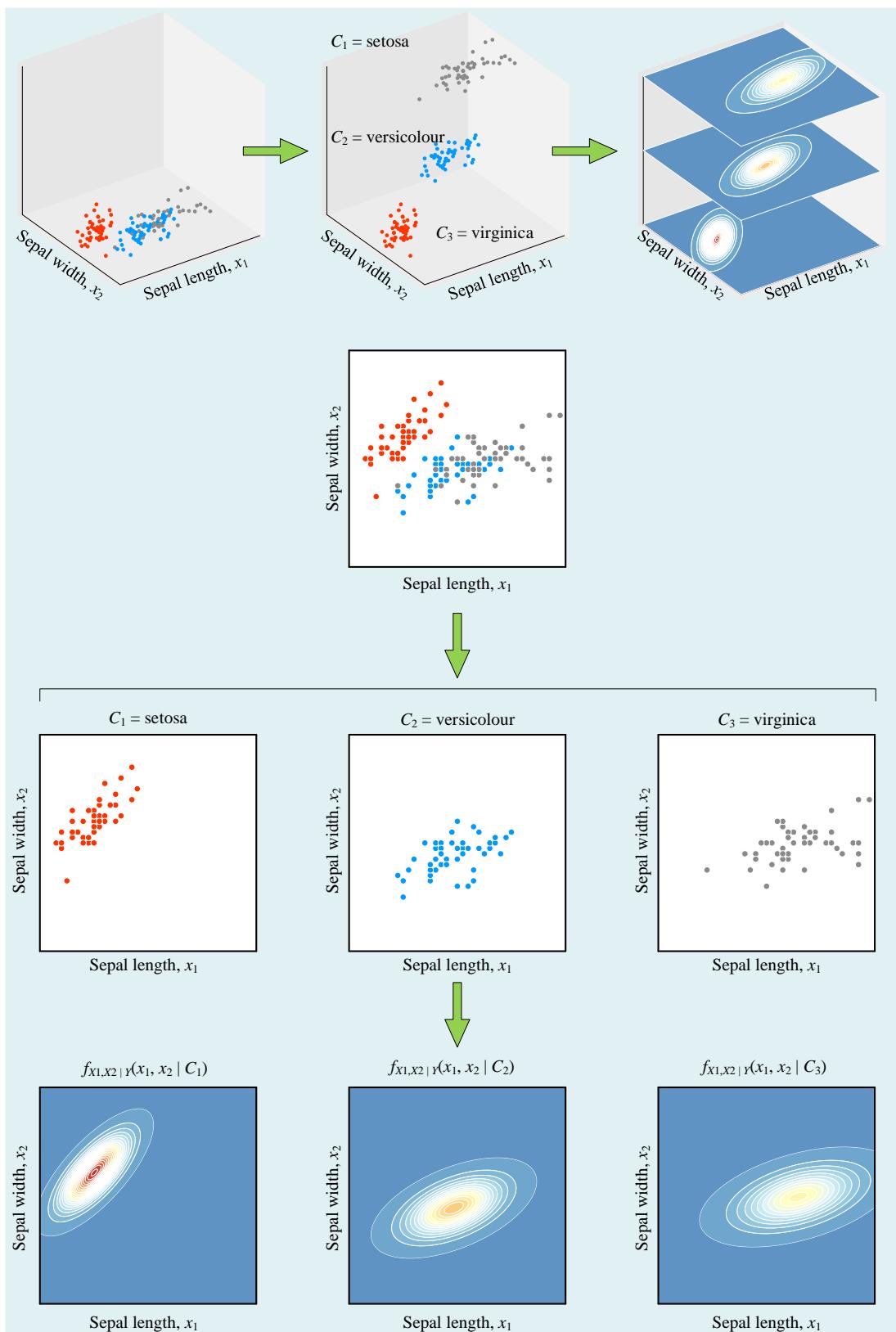


图 1. 似然概率可视化方案，似然概率基于高斯分布

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

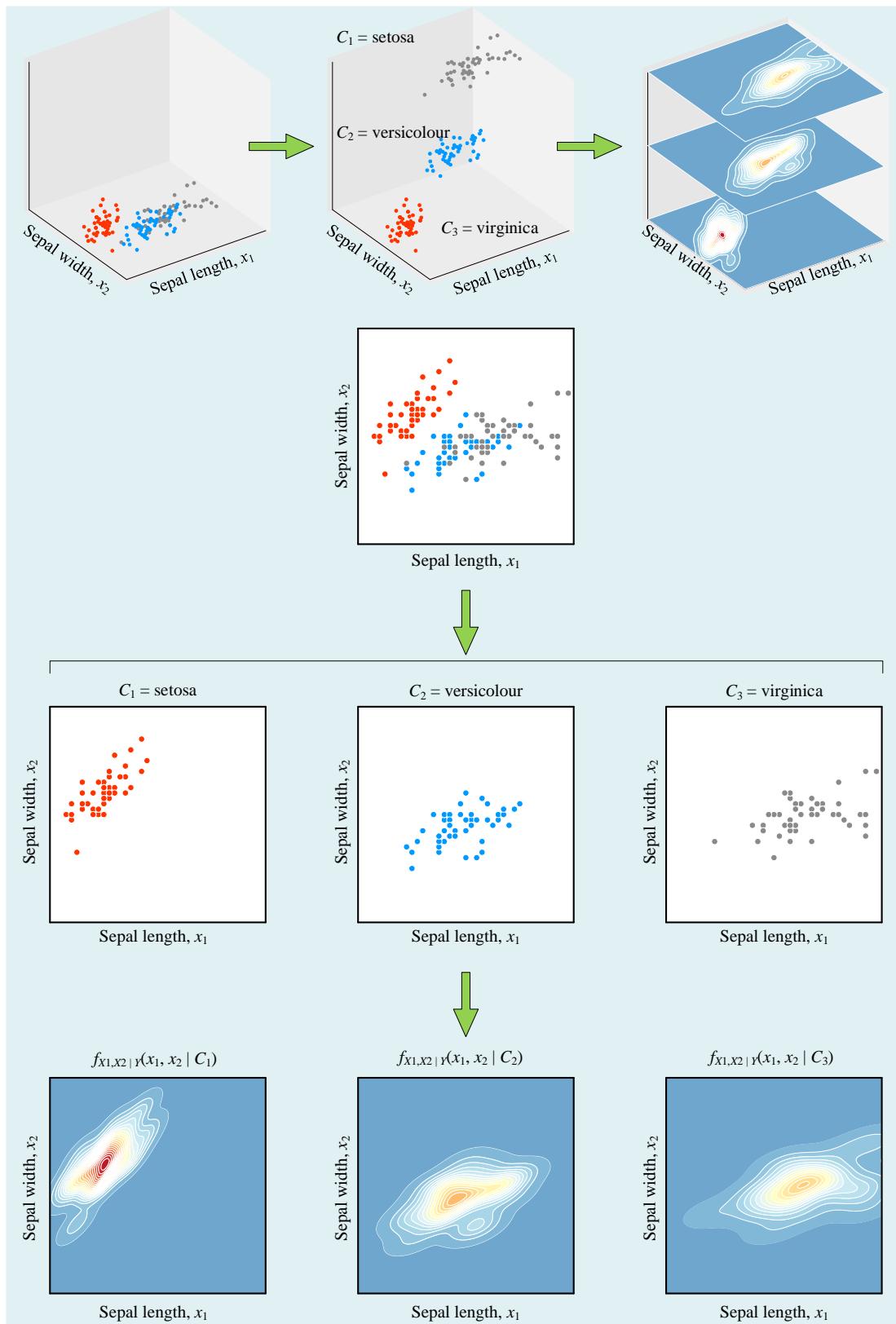


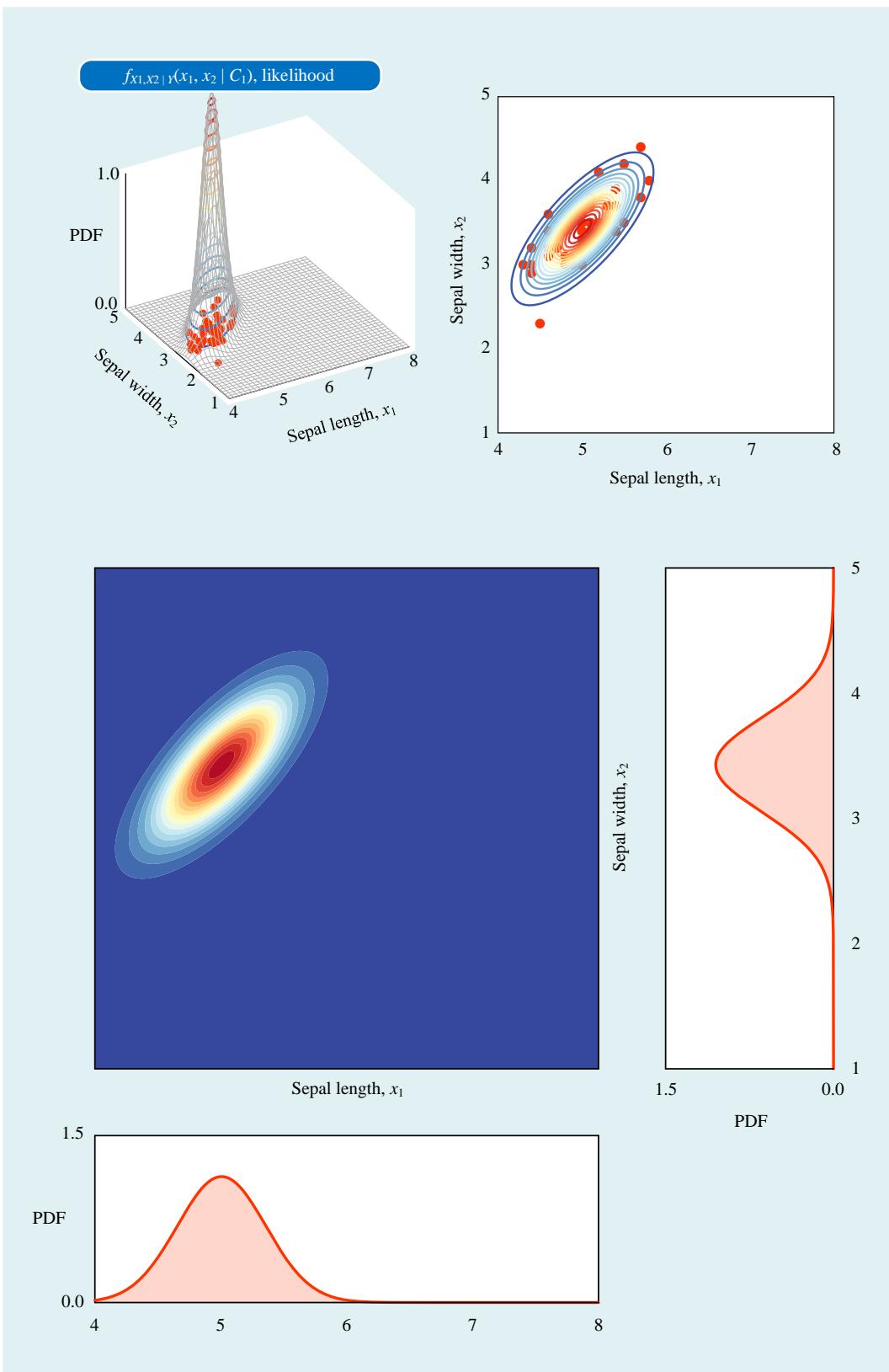
图 2. 似然概率可视化方案，似然概率基于高斯核密度估计

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

图 3. 似然概率, $f_{x1,x2|y}(x_1, x_2 | C_1)$, 似然概率基于高斯分布

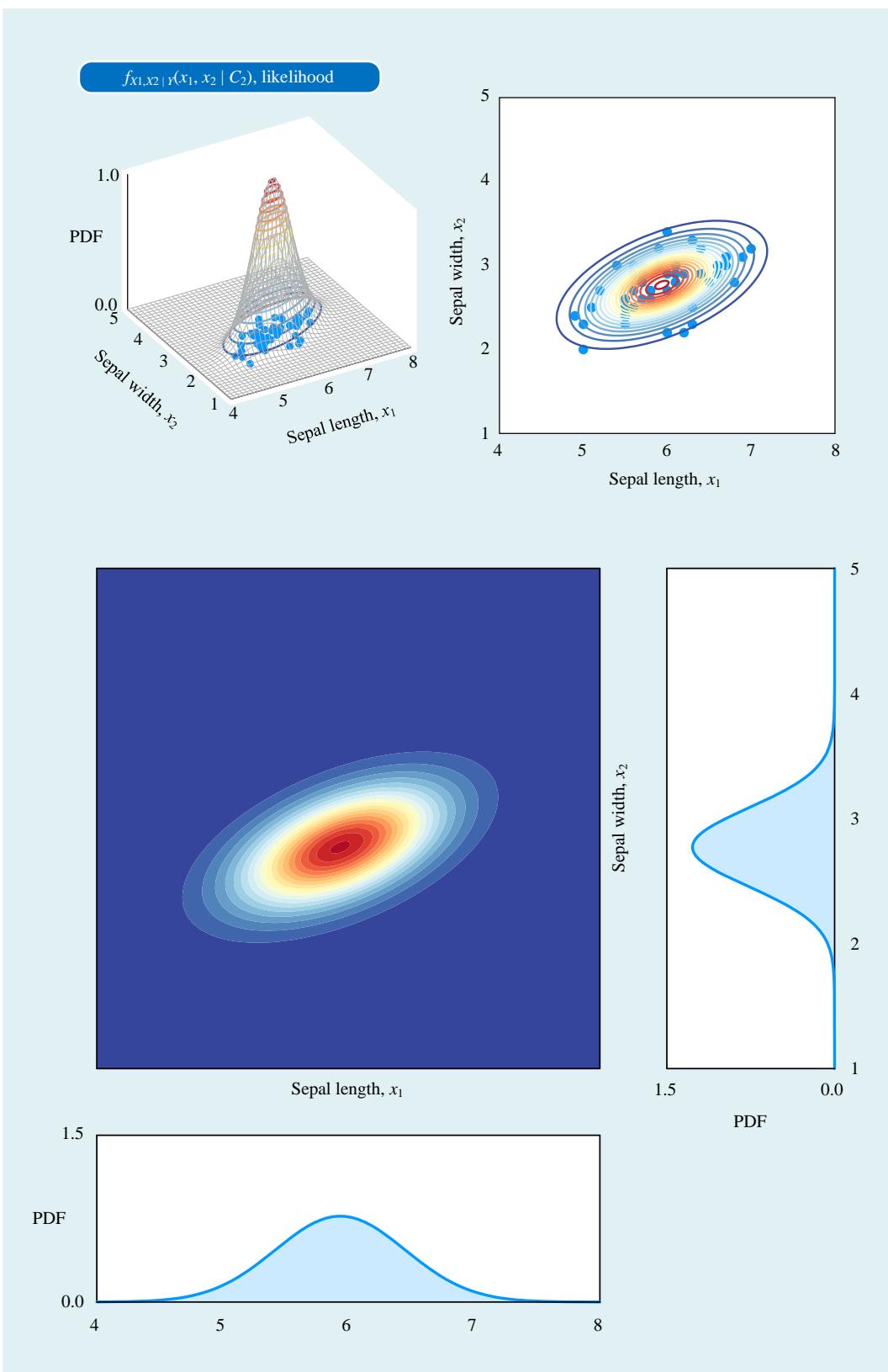
本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

图 4. 似然概率, $f_{x1,x2|y}(x_1, x_2 | C_2)$, 似然概率基于高斯分布

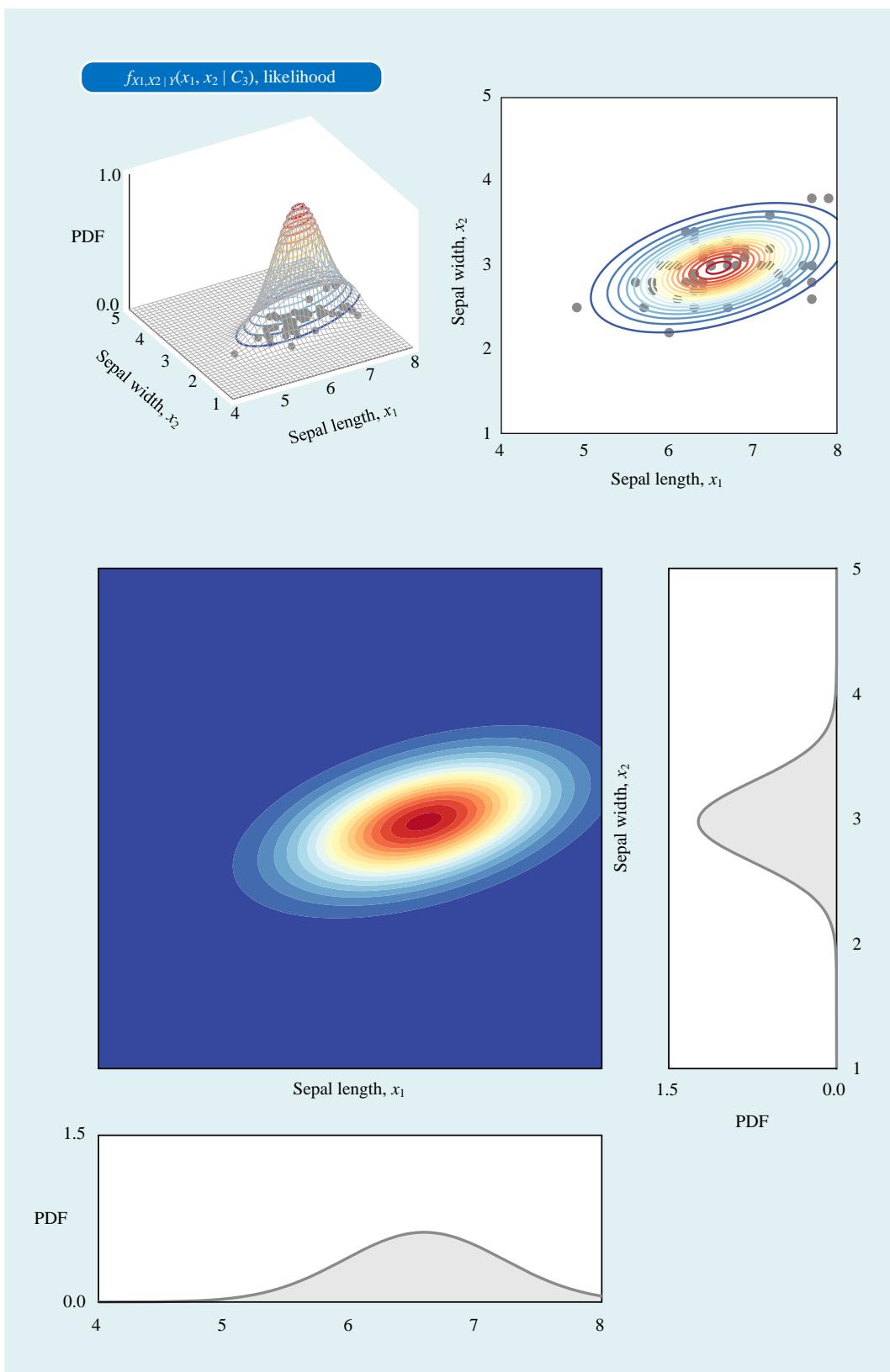
本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

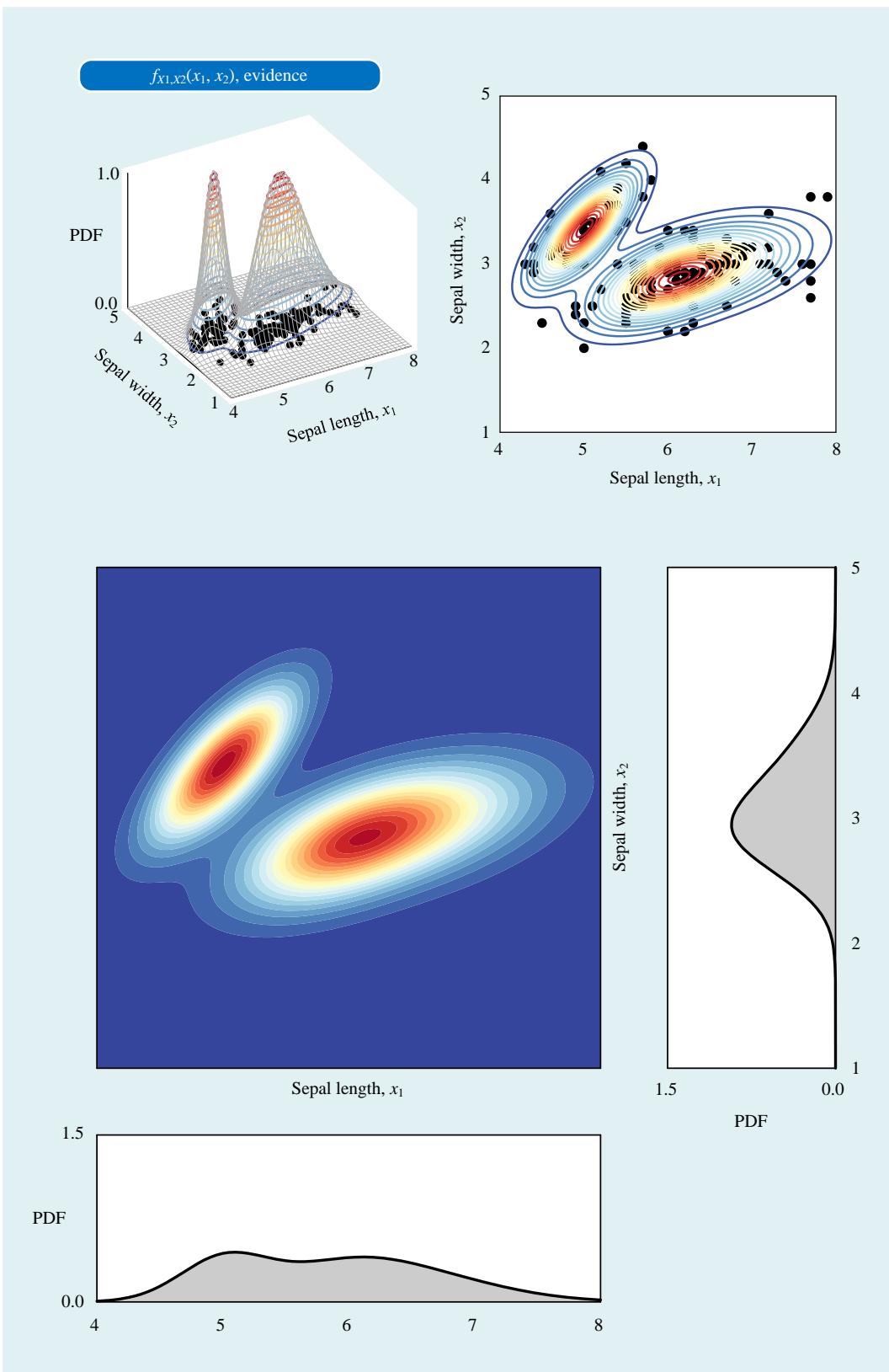
图 5. 似然概率, $f_{x1,x2|y}(x_1, x_2 | C_3)$, 似然概率基于高斯分布

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

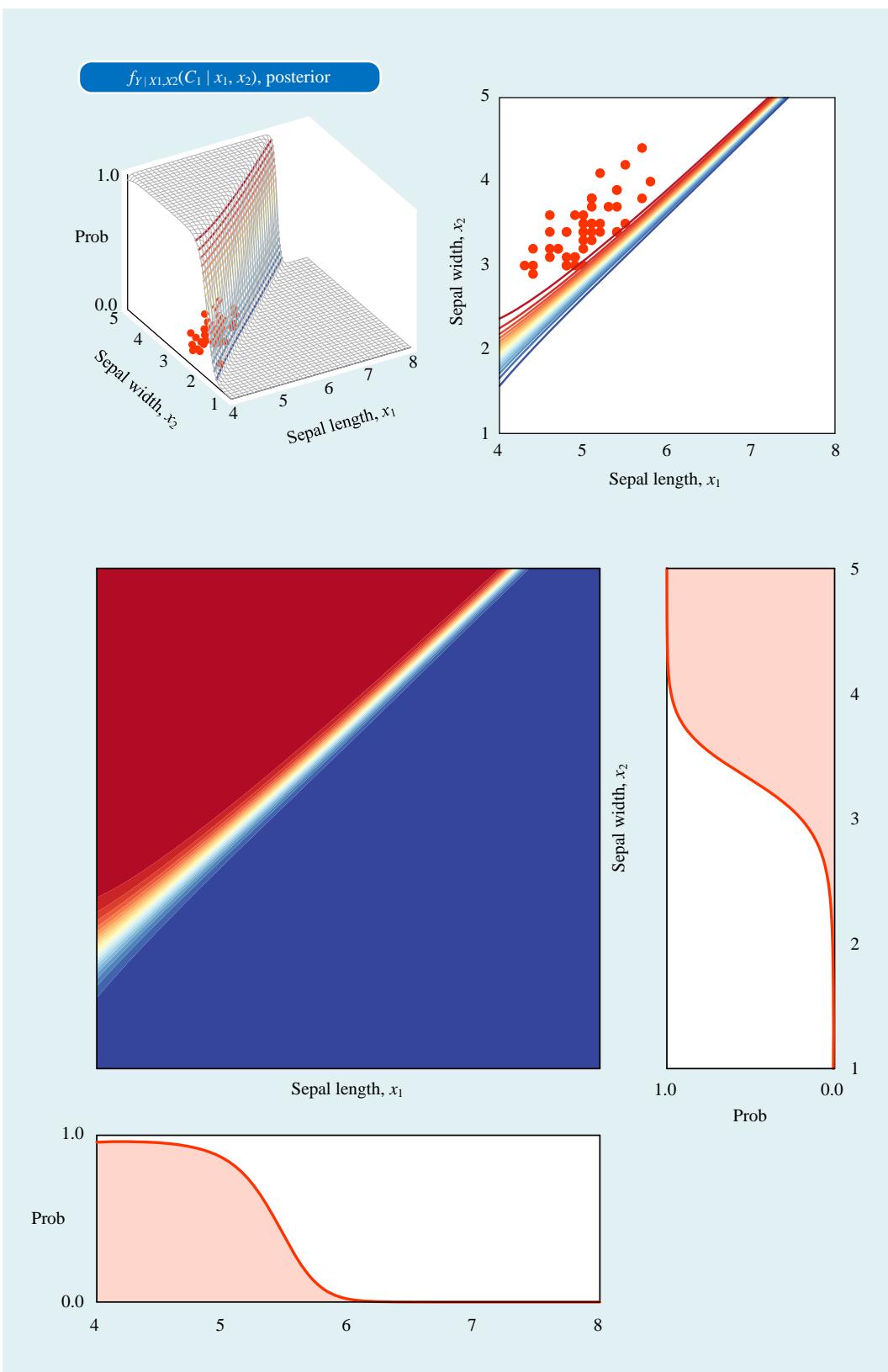
图 6. 证据因子, $f_{x1,x2}(x_1, x_2)$, 似然概率基于高斯分布

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

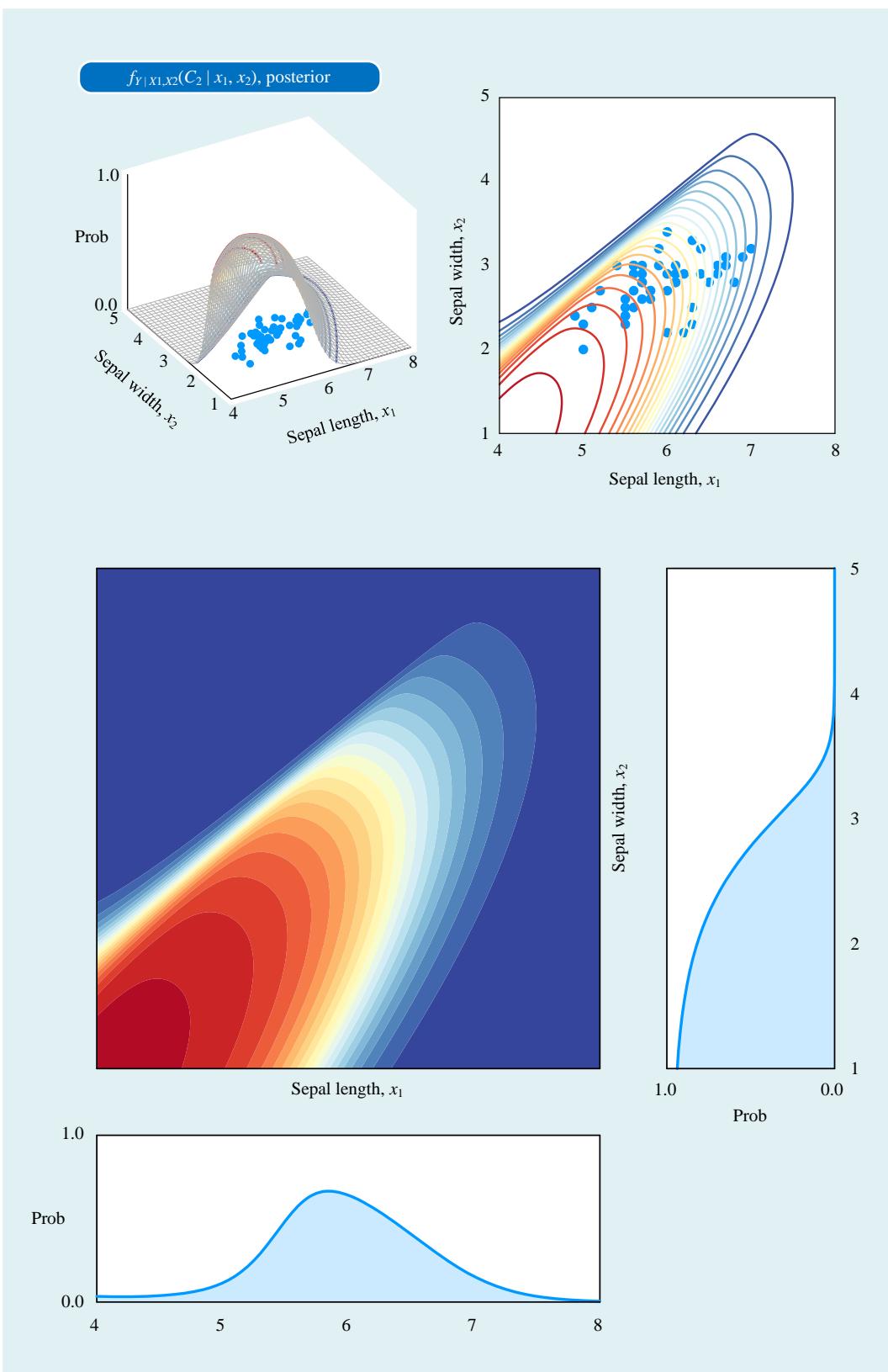
图 7. 后验概率, $f_{Y|X_1,X_2}(C_1 | x_1, x_2)$, 似然概率基于高斯分布

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

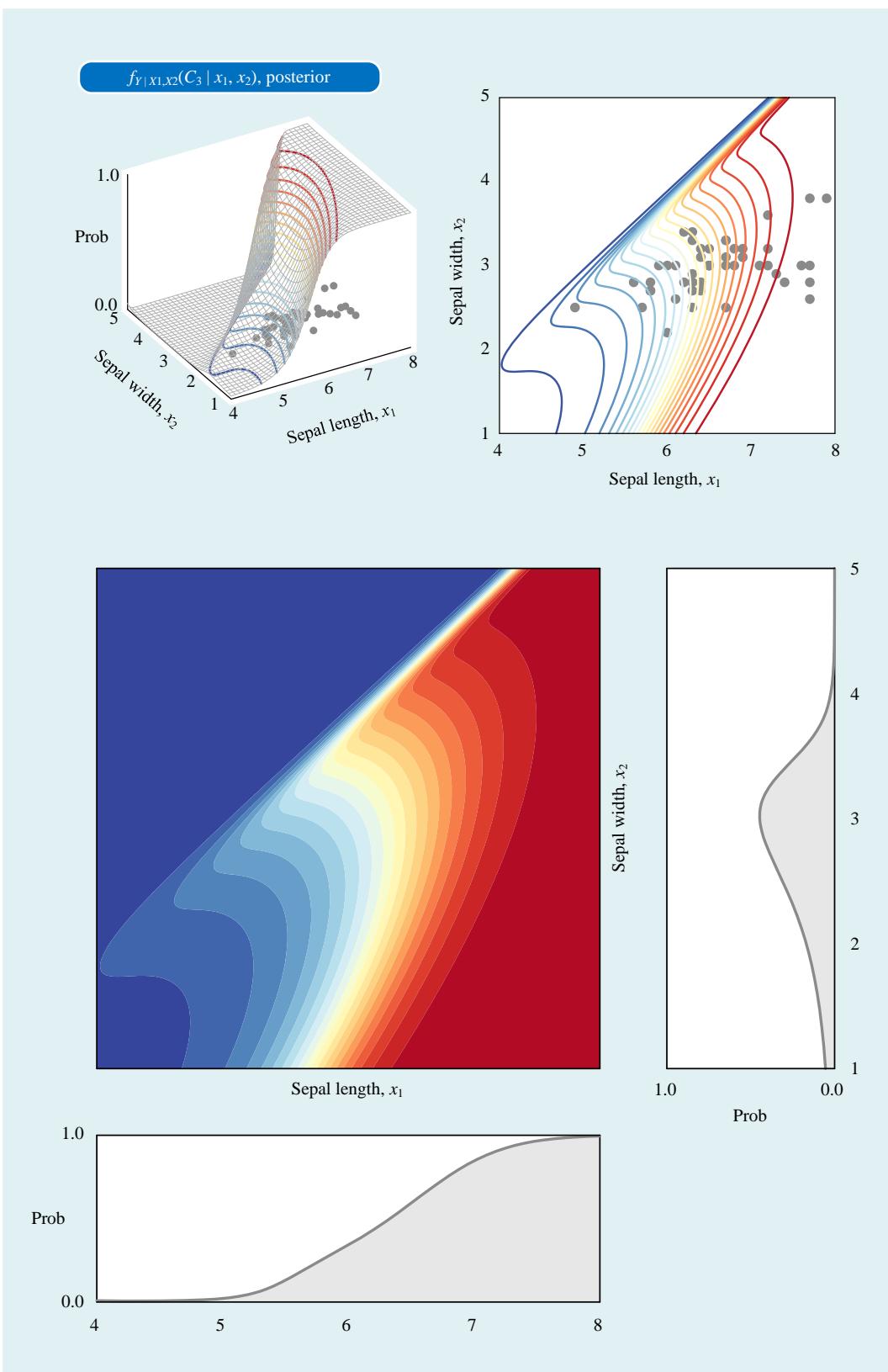
图 8. 后验概率, $f_{Y|X_1,X_2}(C_2 | x_1, x_2)$, 似然概率基于高斯分布

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

图 9. 后验概率, $f_{Y|X_1, X_2}(C_3 | x_1, x_2)$, 似然概率基于高斯分布

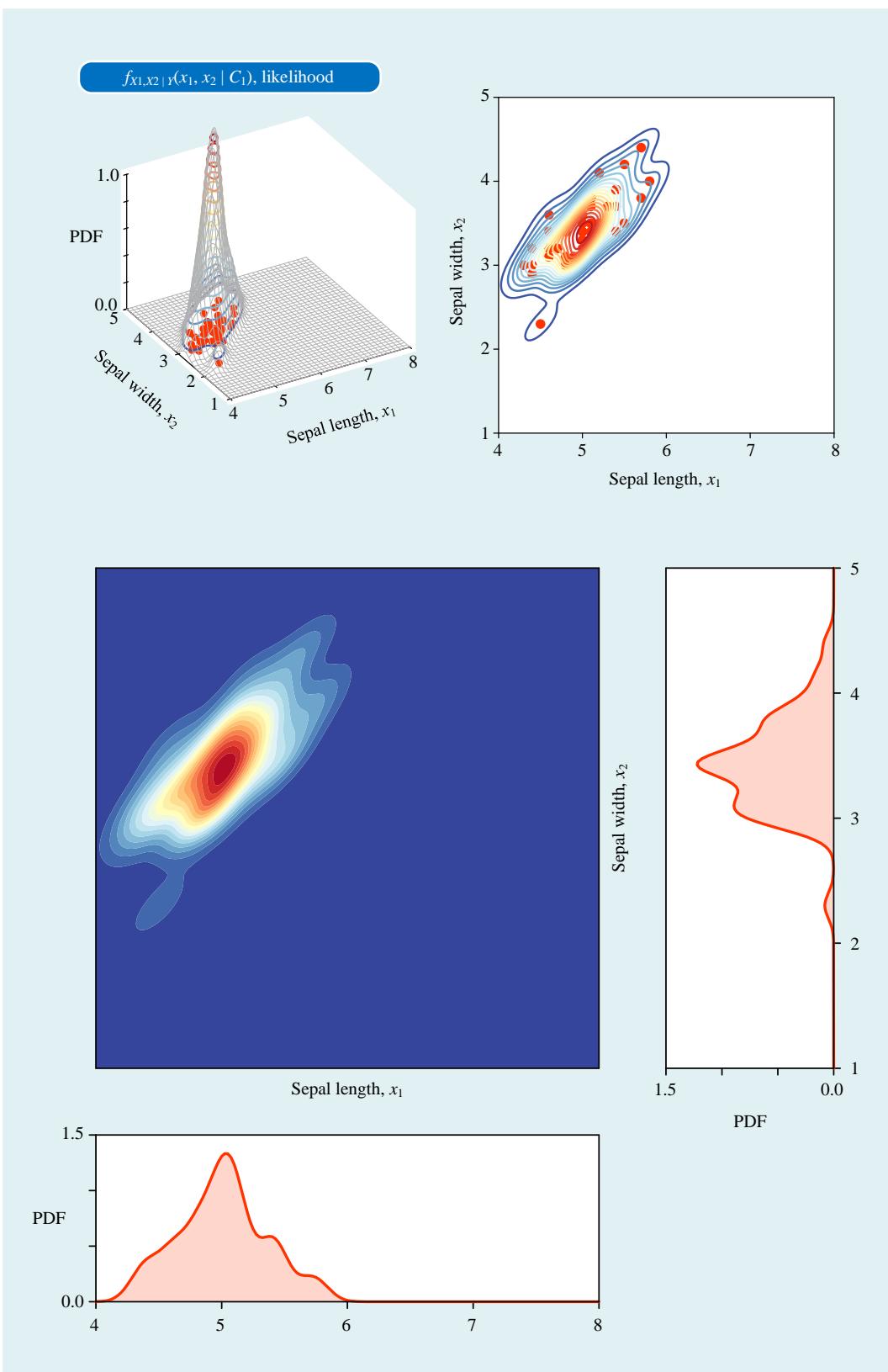
本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

图 10. 似然概率, $f_{x1,x2|y}(x_1, x_2 | C_1)$, 似然概率基于高斯核密度估计

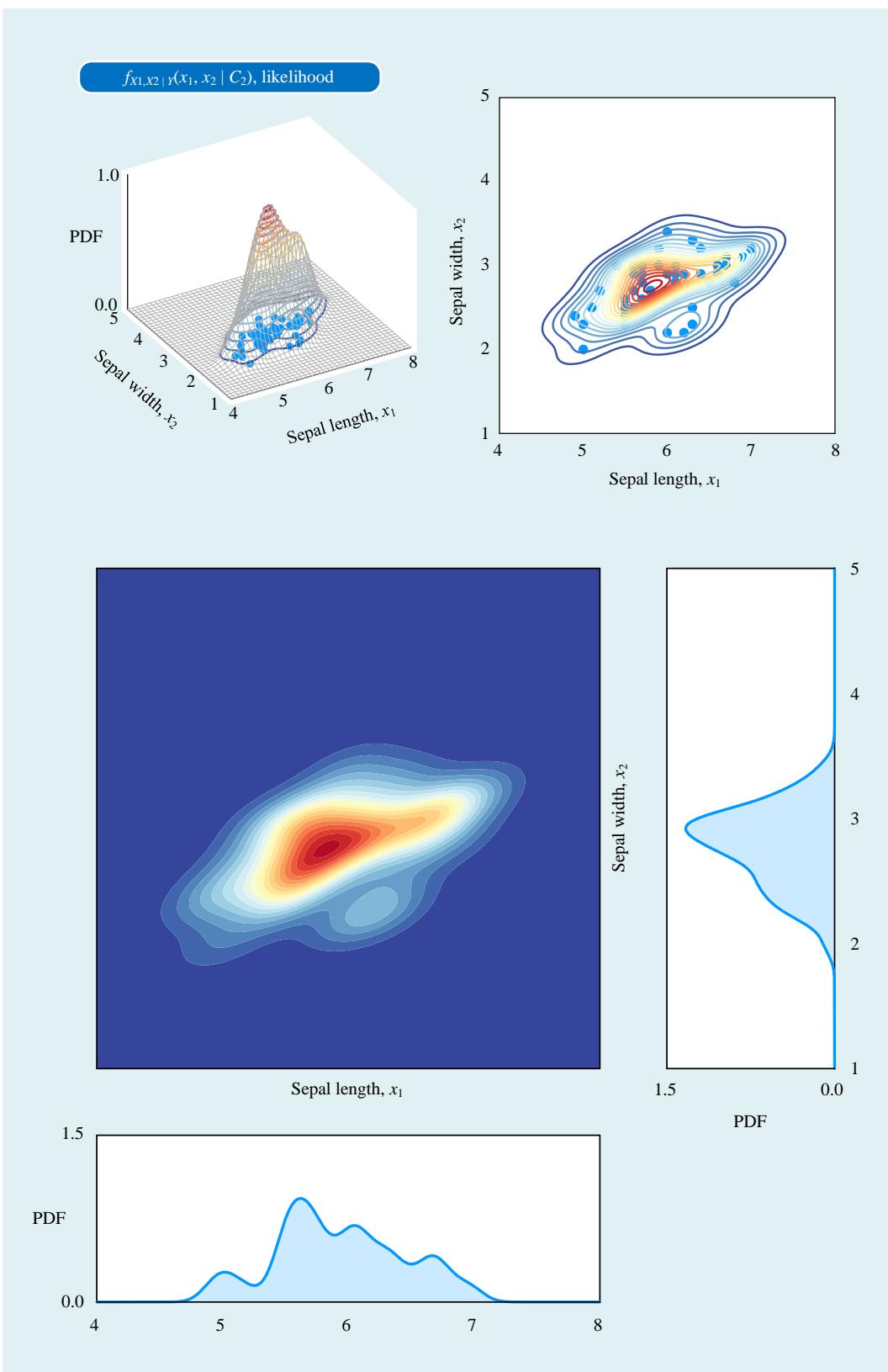
本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

图 11. 似然概率, $f_{x1,x2|y}(x_1, x_2 | C_2)$, 似然概率基于高斯核密度估计

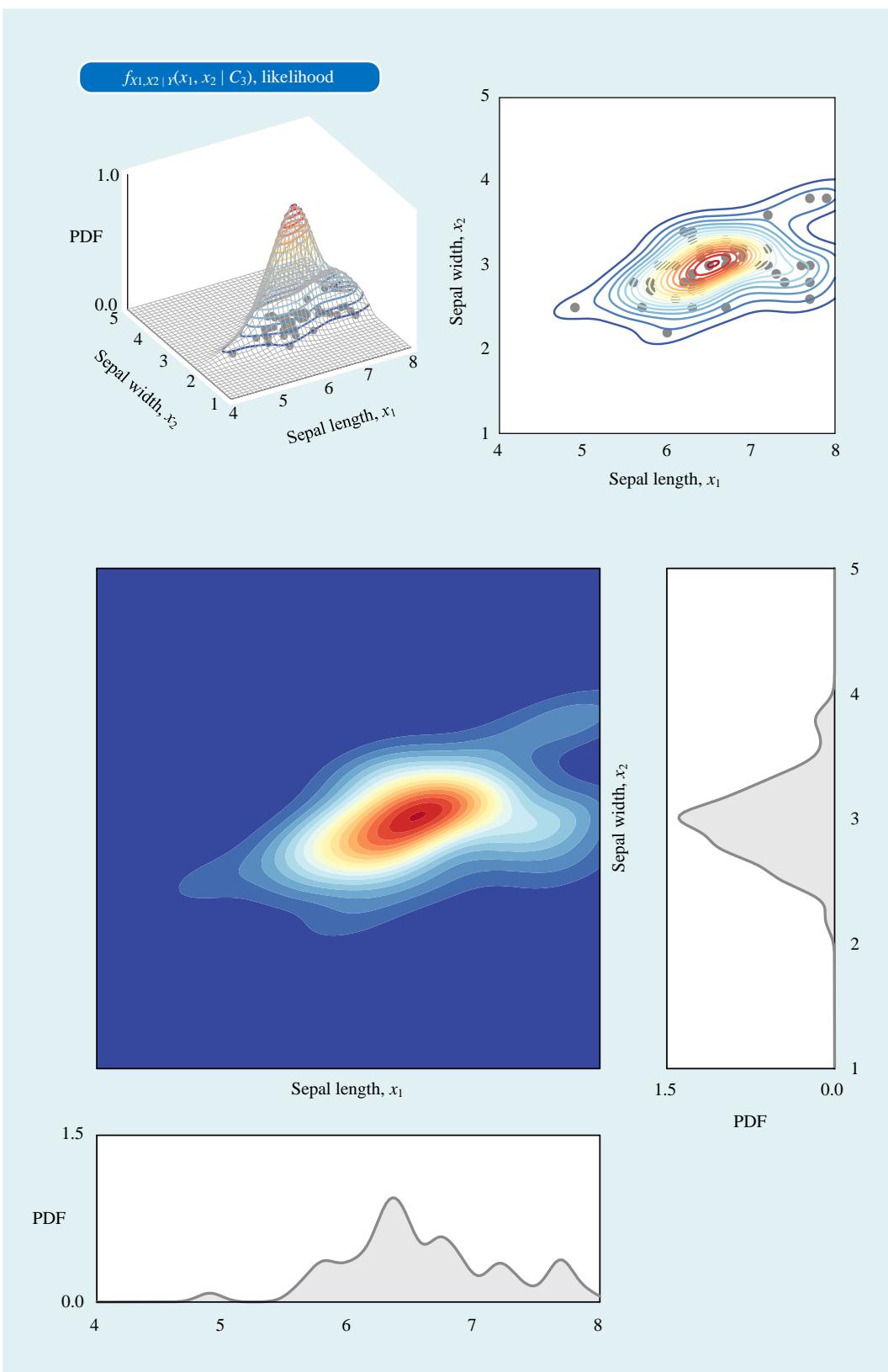
本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

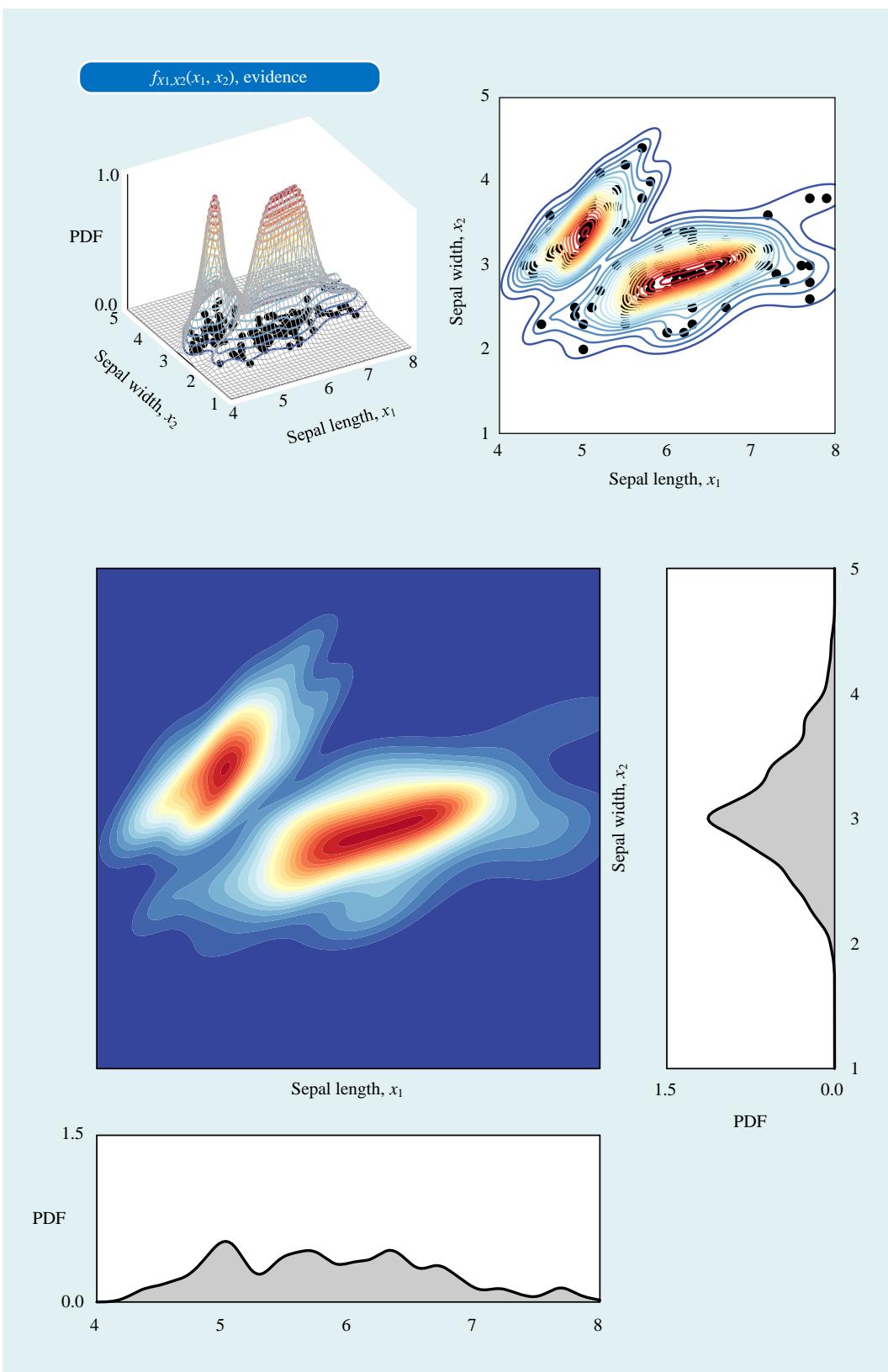
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

图 12. 似然概率, $f_{x1,x2|y}(x_1, x_2 | C_3)$, 似然概率基于高斯核密度估计

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

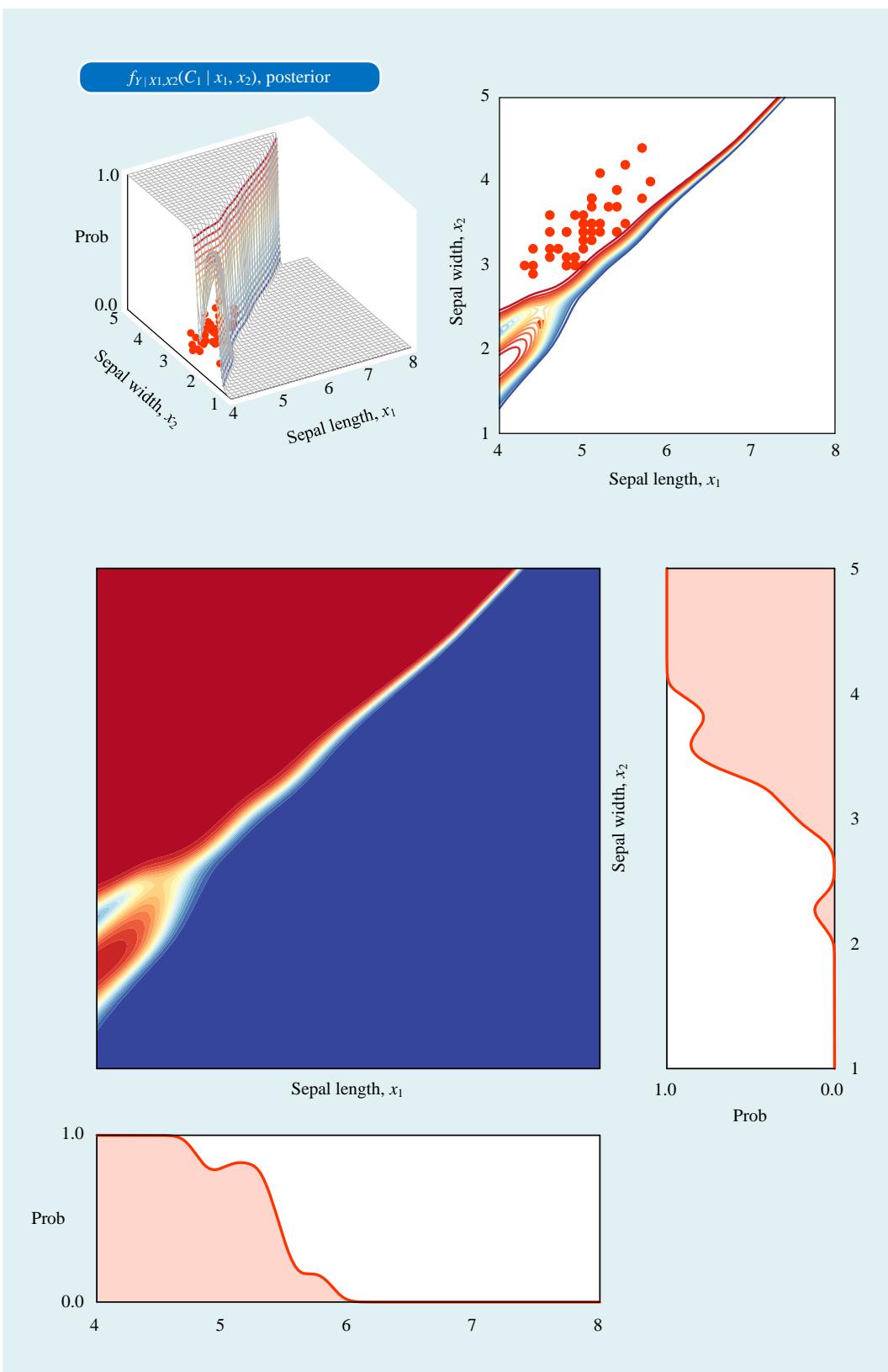
图 13. 证据因子, $f_{x1,x2}(x_1, x_2)$, 似然概率基于高斯核密度估计

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
 版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

图 14. 后验概率, $f_{Y|X1,X2}(C_1 | x_1, x_2)$, 似然概率基于高斯核密度估计

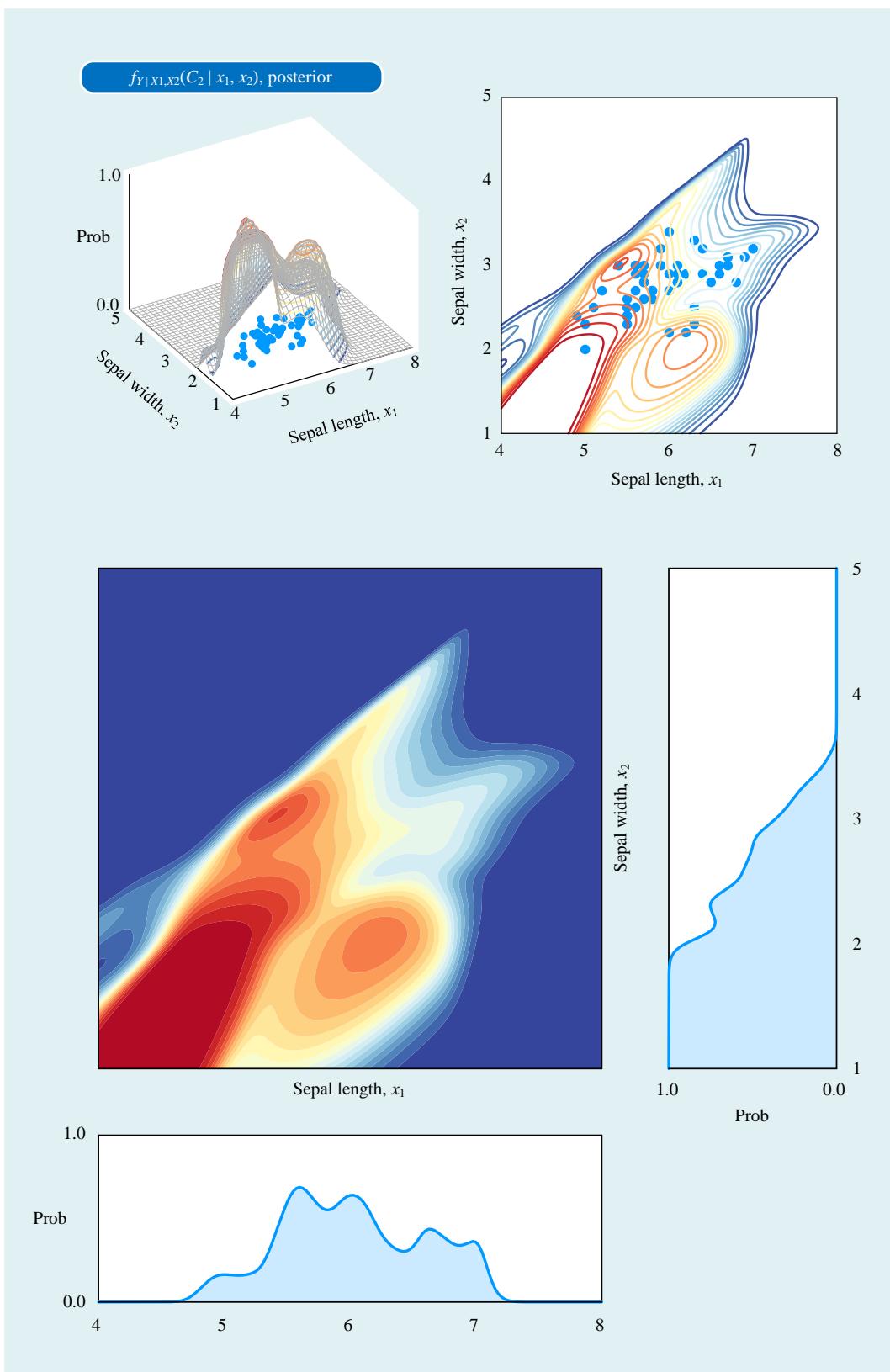
本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

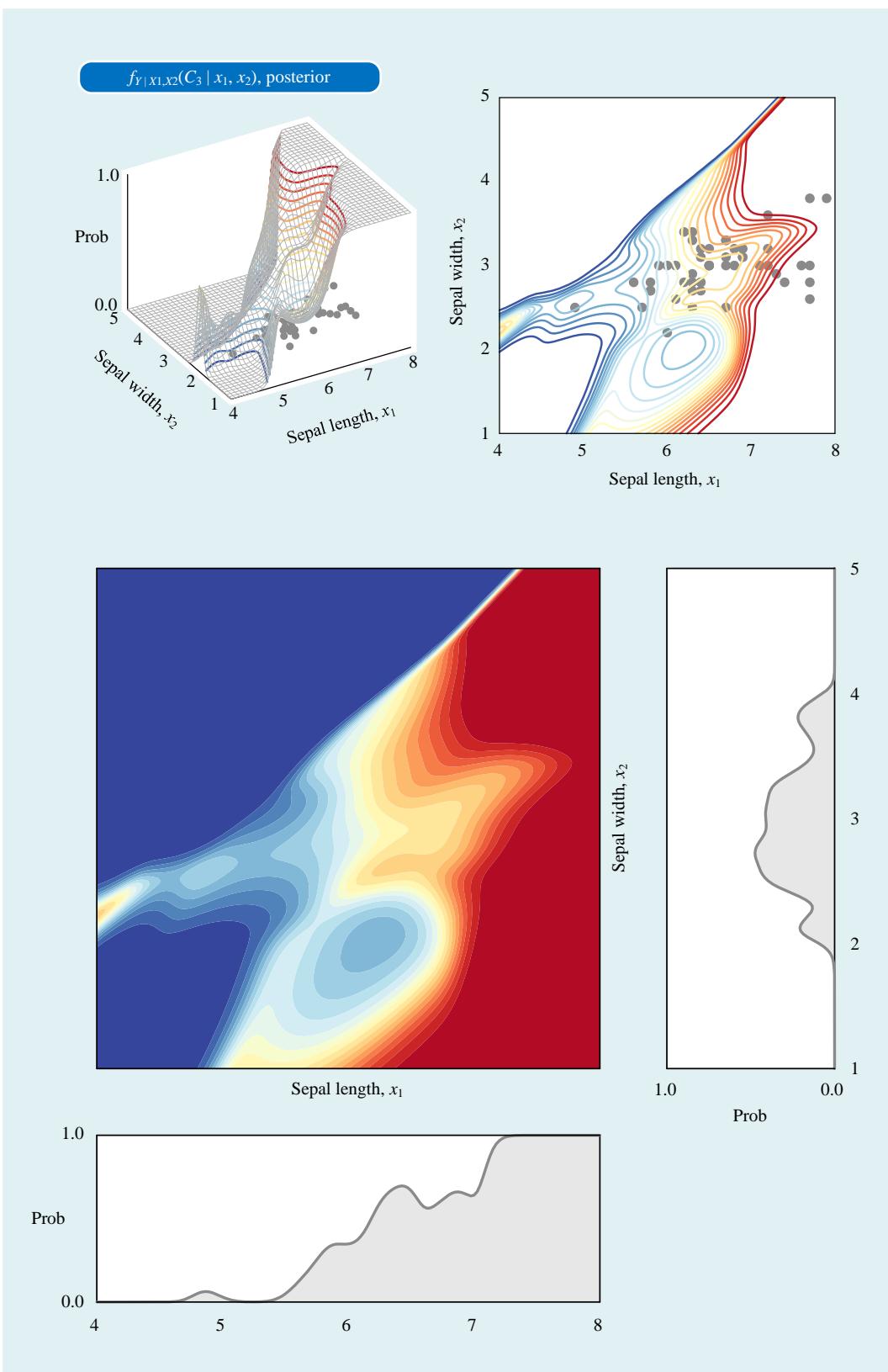
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

图 15. 后验概率, $f_{Y|X1,X2}(C_2 | x_1, x_2)$, 似然概率基于高斯核密度估计

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>
本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

图 16. 后验概率, $f_{Y|X1,X2}(C_3 | x_1, x_2)$, 似然概率基于高斯核密度估计

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

36

Pattern + Randomness

模式 + 随机

模式让世界充满秩序，随机让寰宇满是精彩



你能想象的所有东西都是真的。

Everything you can imagine is real.

——毕加索 (Pablo Picasso) | 西班牙艺术家 | 1881 ~ 1973



- ◀ matplotlib.patches.Circle() 绘制正圆
- ◀ matplotlib.pyplot.axhline() 绘制水平线
- ◀ matplotlib.pyplot.axvline() 绘制竖直线
- ◀ matplotlib.pyplot.contour() 绘制等高线图
- ◀ matplotlib.pyplot.contourf() 绘制填充等高线图
- ◀ mpmath.e mpmath 库中的欧拉数
- ◀ mpmath.pi mpmath 库中的圆周率
- ◀ mpmath.sqrt(2) mpmath 库计算 2 的平方根
- ◀ np.random.normal() 产生服从正态分布随机数
- ◀ numpy.cumsum() 累加
- ◀ numpy.fliplr() 上下翻转矩阵
- ◀ scipy.stats.multivariate_normal() 多元高斯分布
- ◀ scipy.stats.multivariate_normal.pdf() 多元高斯分布 PDF 函数
- ◀ seaborn.distplot() 绘制频率直方图和 KDE 曲线

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

36.1 二叉树：道生一，一生二，二生三，三生万物

在雪花、树叶中，我们看到了分形和基本几何形状，这就是模式；世上没有两片完全一样的雪花，没有两片一样的叶子，这便是随机。

人生要经历生老病死，后浪把前浪拍在沙滩上，这也都是模式；即便如此，每个人都可以活出自己的精彩。

《道德经》中“道生一，一生二，二生三，三生万物”一句完美地描述了“模式 + 随机”。“道”就是“模式”，而万物生的繁复则来自于“随机”。“道生一，一生二，二生三”像极了二叉树。图 1 所示为利用 Python 绘制的二叉树，这个二叉树顺时针旋转 90 度便得到杨辉三角结构。

图 2 所示为在二叉树中随机行走轨迹。

数学和艺术中，“模式 + 随机”更是无处不在。图 3、图 4 给出的圆周率、自然对数底数、根号 2、黄金分割比小数点后超过 1000 位 0 ~ 9 数字热图，我们似乎看不到任何规律；但是，统计之后，我们会发现 0 ~ 9 似乎分布均匀。虽然，目前数学上很难证明这四个数是正规数。热爱数学、艺术的话，把这四幅热图又何尝不是完美的艺术品。

36.2 蒙特卡罗模拟

随机是指在一系列事件或结果中没有可预测模式或规律性，无法通过已知信息准确预测的性质。随机数发生器是一种用来生成随机数的工具或算法。它可以产生看似无规律、不可预测的数字序列。图 5 所示为满足不同相关性系数的二元高斯分布随机数，这些子图中我们可以看到随机数和椭圆的关系。

蒙特卡洛模拟是一种基于概率统计的计算方法，通过随机抽样的方式模拟和分析现实问题。它以概率和统计的原理为基础，通过大量的随机抽样实验来模拟系统的行为和结果，以获得对系统行为的估计或预测。图 6 所示为一维随机行走，三种不同的模式——向上、居中、向下。图 7 像是随风飘动的丝线，每增加一处约束就限制了丝线局部的运动。



读完《可视之美》后，大家可能会问，数学研究、艺术创作，到底有什么用？

想想我们今天的物质世界有多少和牛顿力学直接间接相关。再想想，牛顿同时代的人有多少人根本不理解牛顿力学三定义，甚至可能都没听过牛顿这个人。

不管是宏观，还是微观，只有不断“升维”，提升认知的维度，才能让我们跳脱刘慈欣《三体》描述的“火鸡视角”。而数学和艺术就是射穿“火鸡视角”的那束光。

套用刘慈欣的话，我们都是阴沟里的虫子，抑或是农场里的火鸡，但总有人要投身数学、热爱艺术、仰望星空。

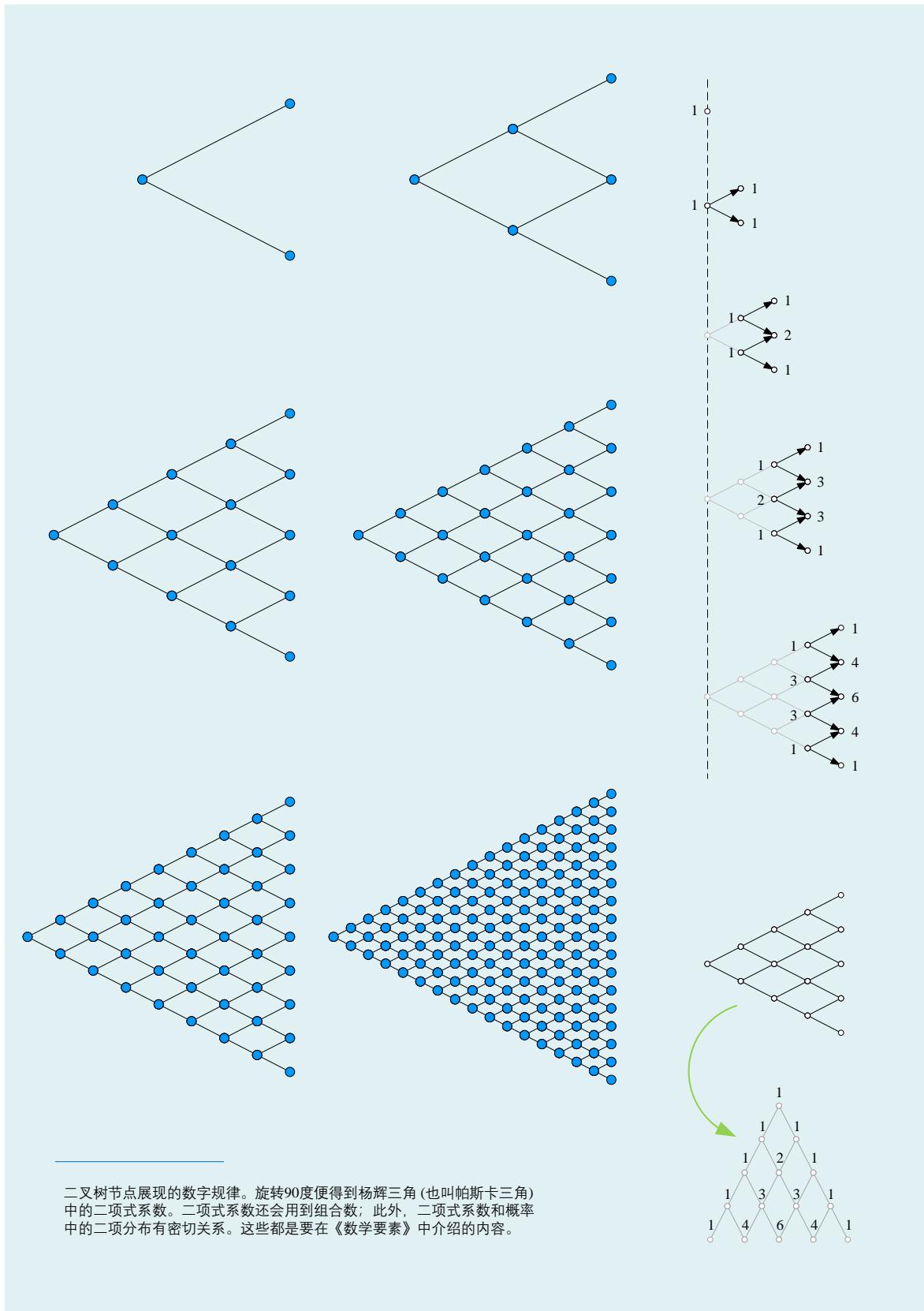


图 1. 二叉树

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

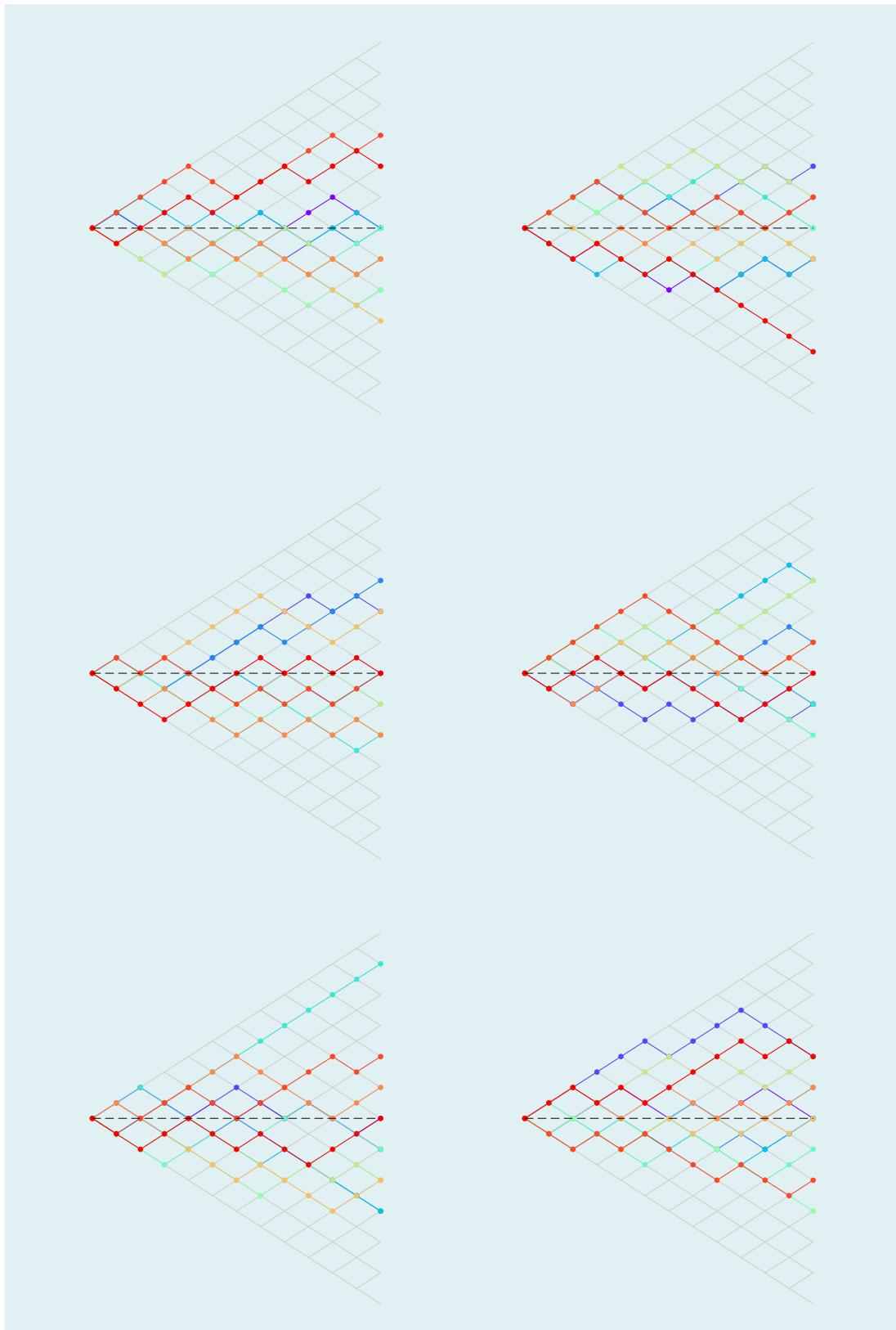


图 2. 二叉树中的随机行走

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

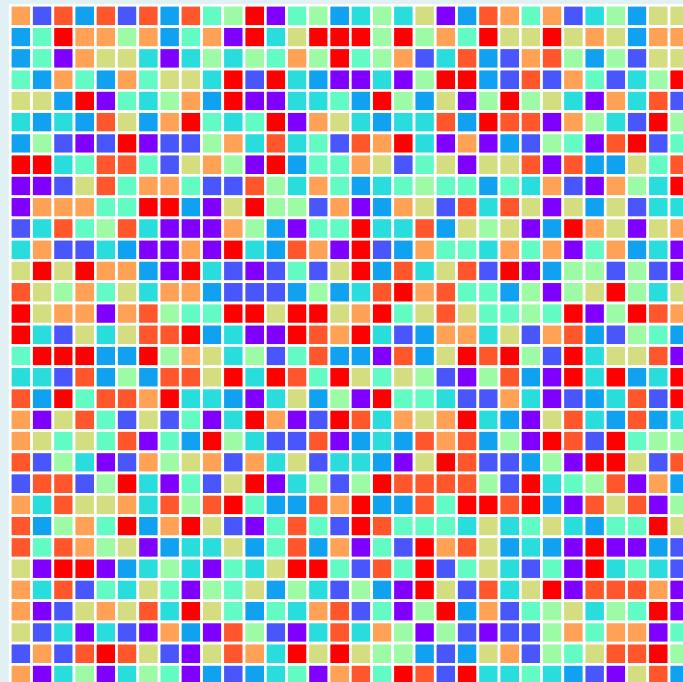
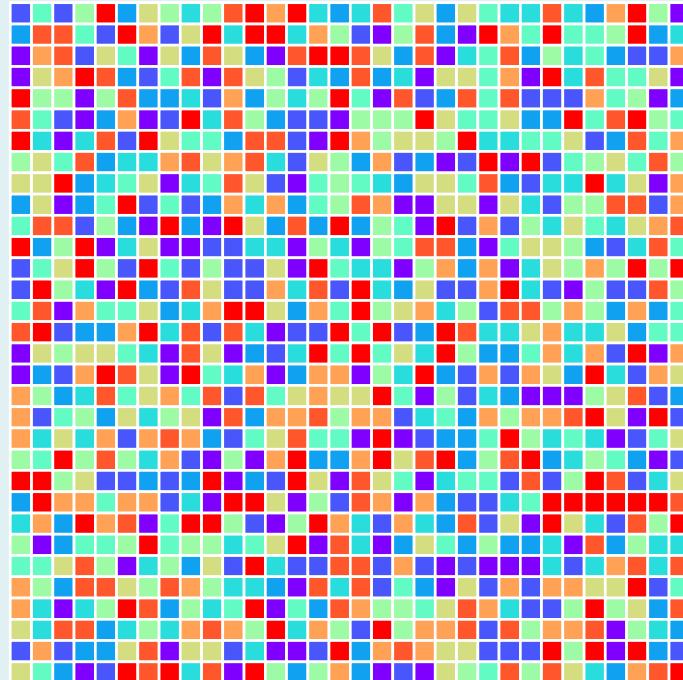


图 3. 圆周率 π 、自然对数底数 e ，小数点后 1024 位热图

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

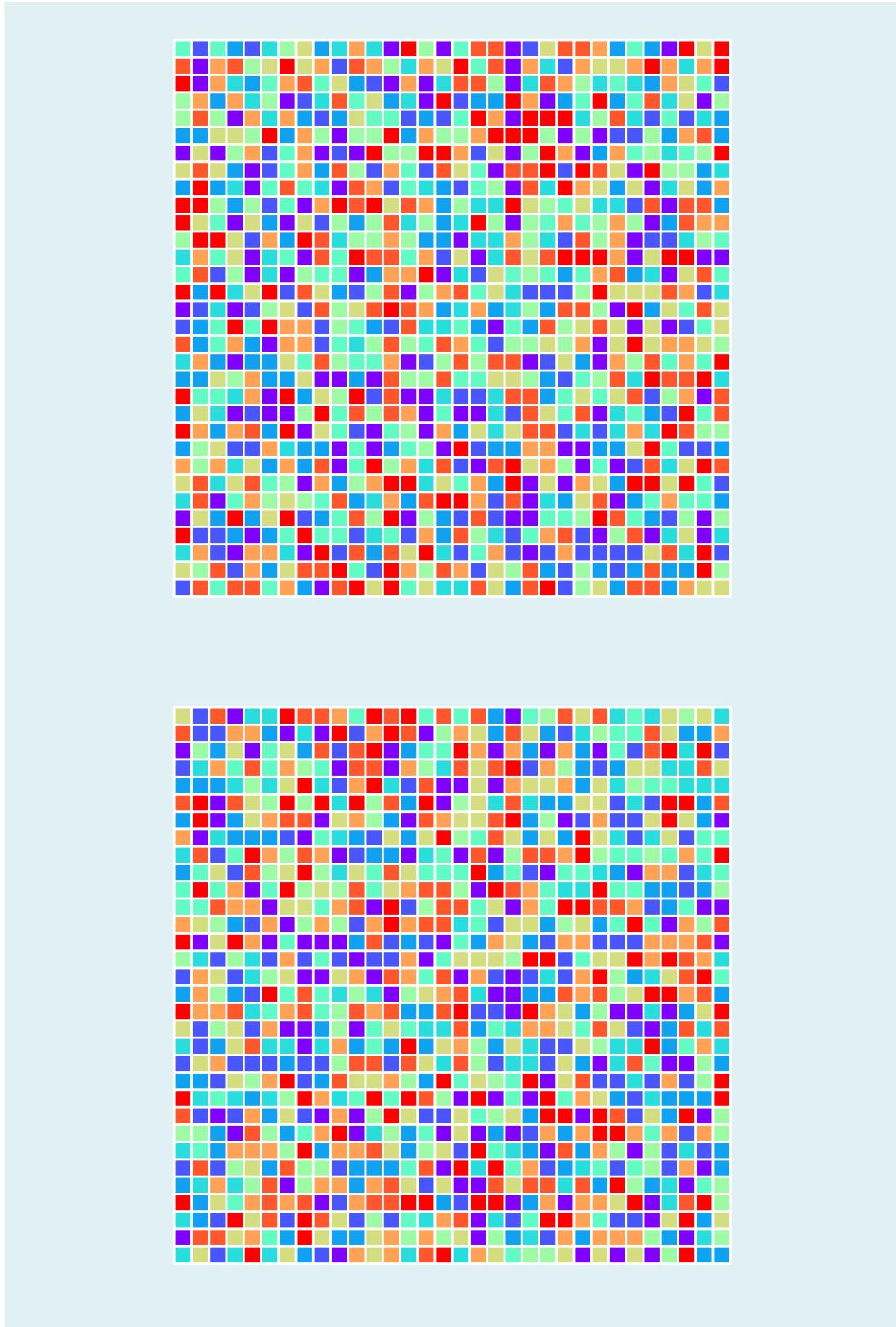


图 4. 根号 2、黄金分割比，小数点后 1024 位热图

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

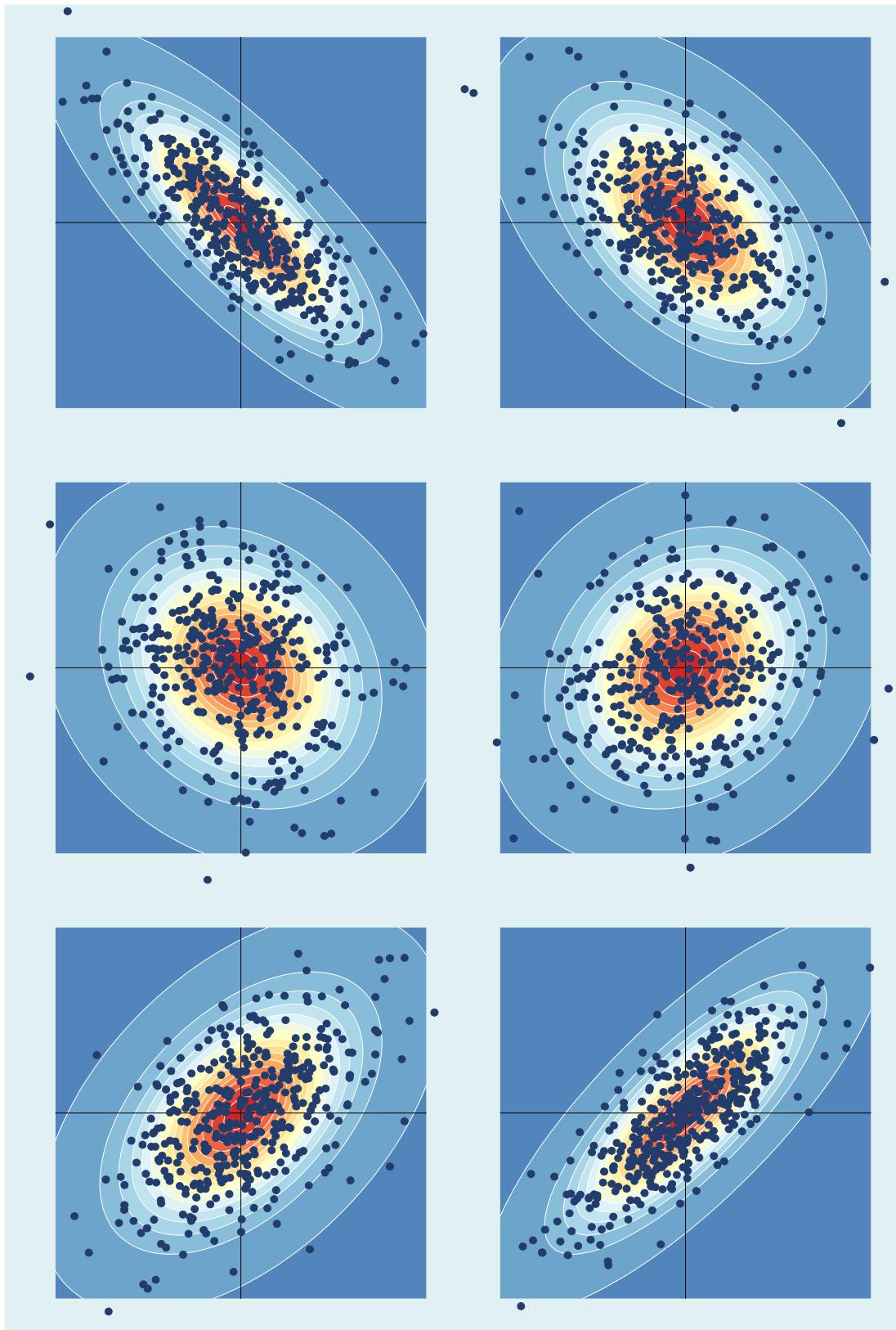


图 5. 满足二元高斯分布的随机数

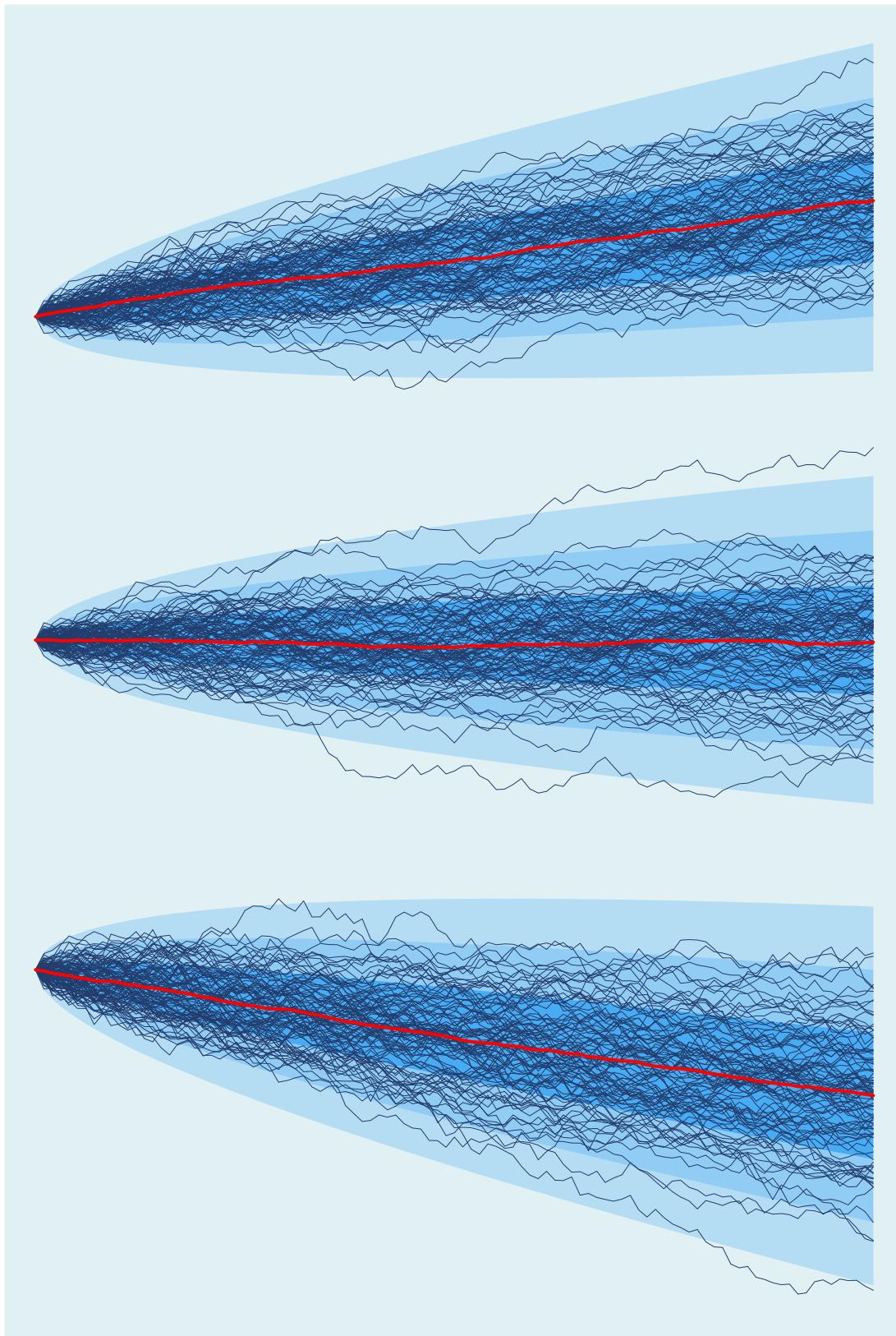


图 6. 一维随机行走，100 条轨迹

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

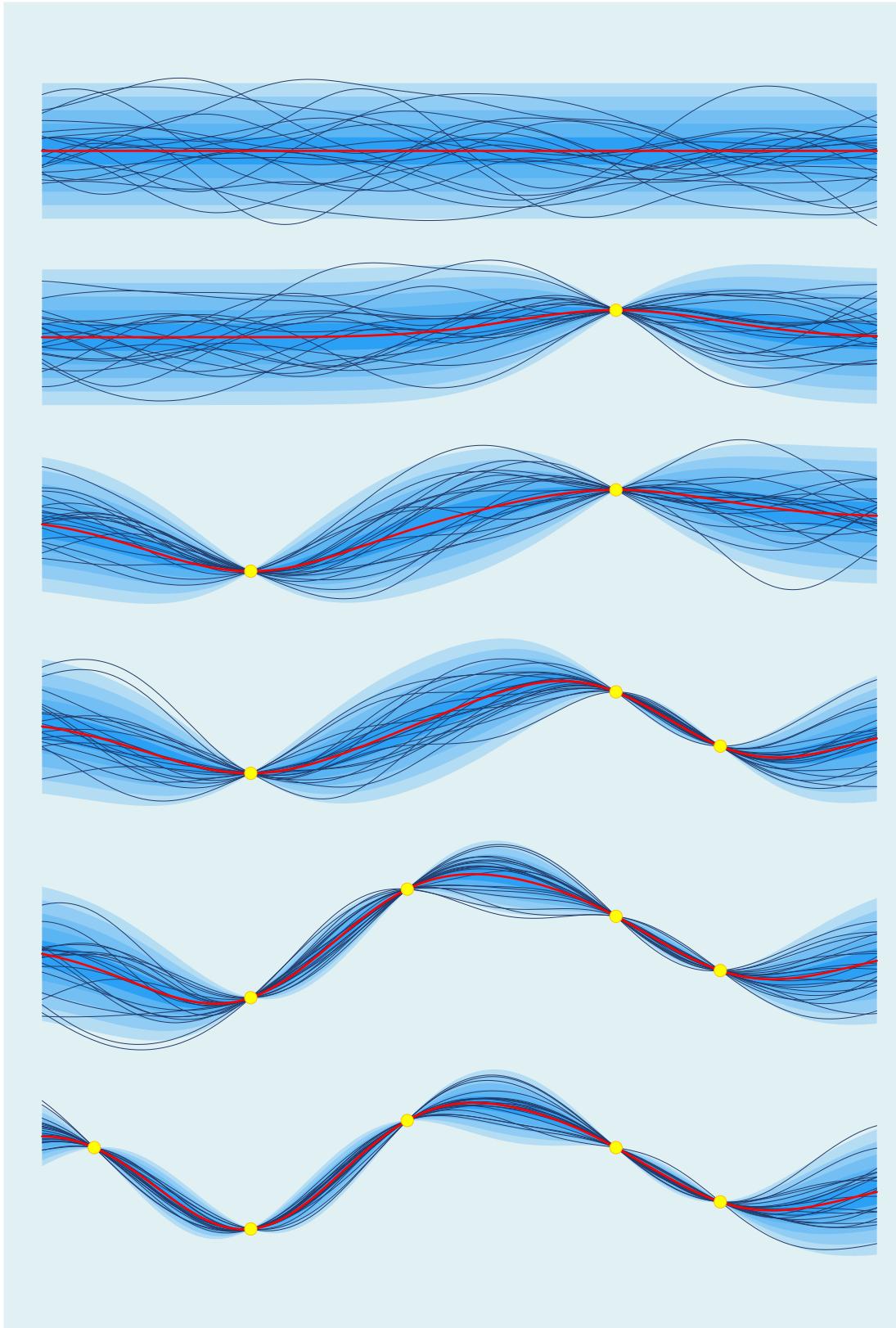


图 7. 高斯过程，20 条轨迹

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微课视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

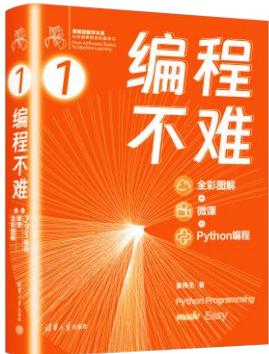
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com



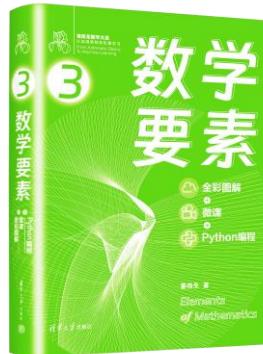
“鸢尾花书”的整体布局

数学

数学基础



线性代数



概率统计



Python编程

数据可视化

回归、降维

分类、聚类

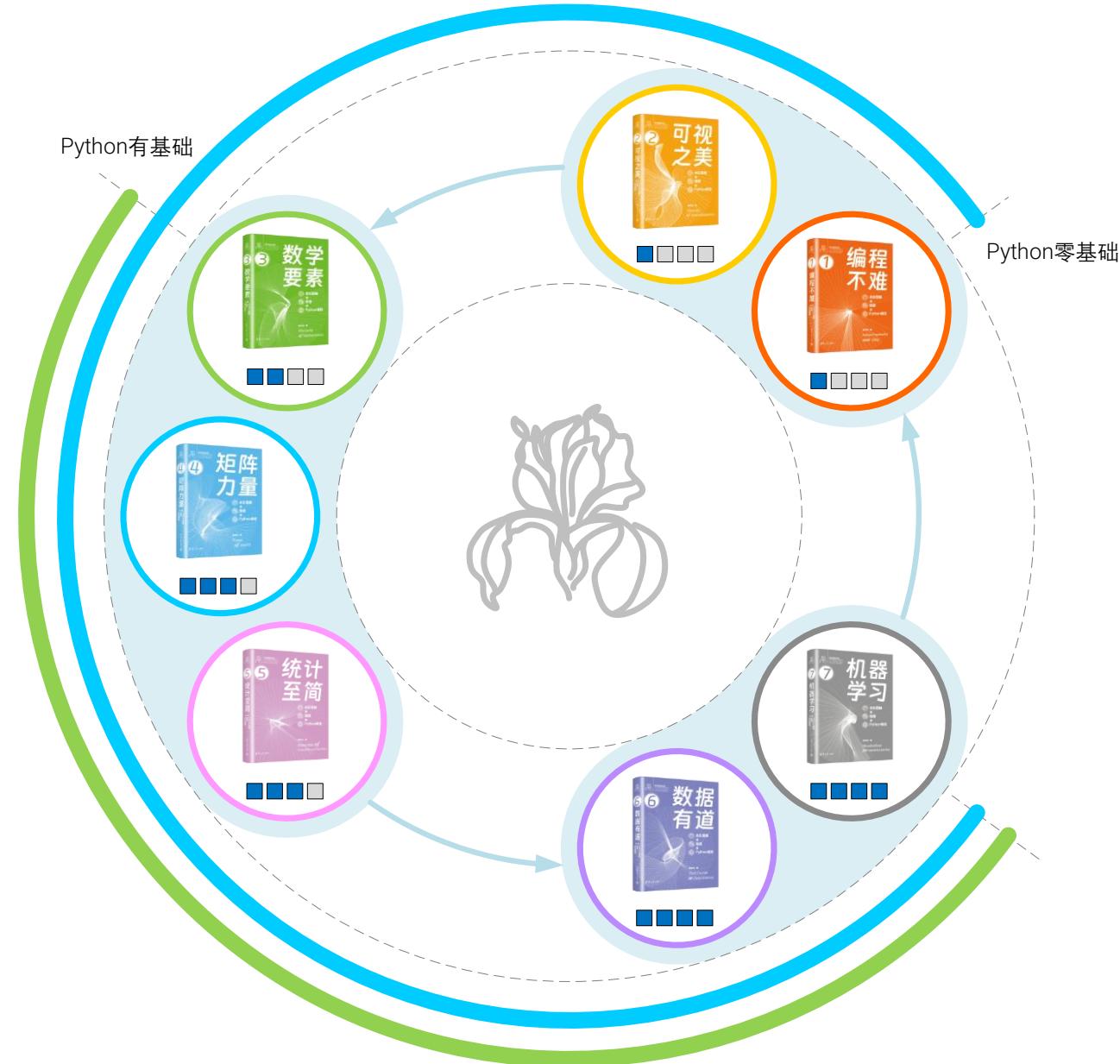
工具

实践

数学 + Python编程 + 可视化 + 机器学习实践



“鸢尾花书”的学习顺序

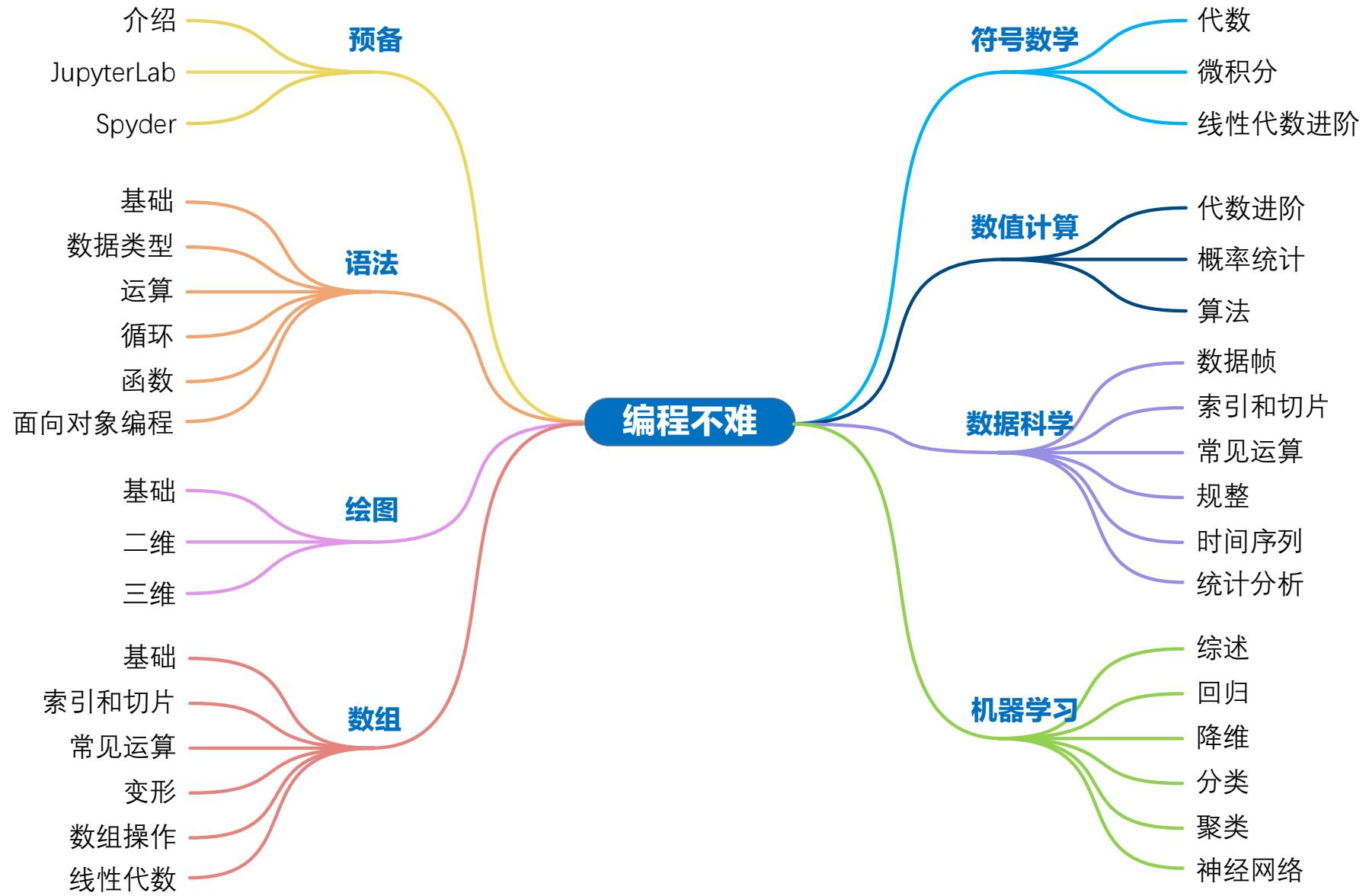


分册进度状态

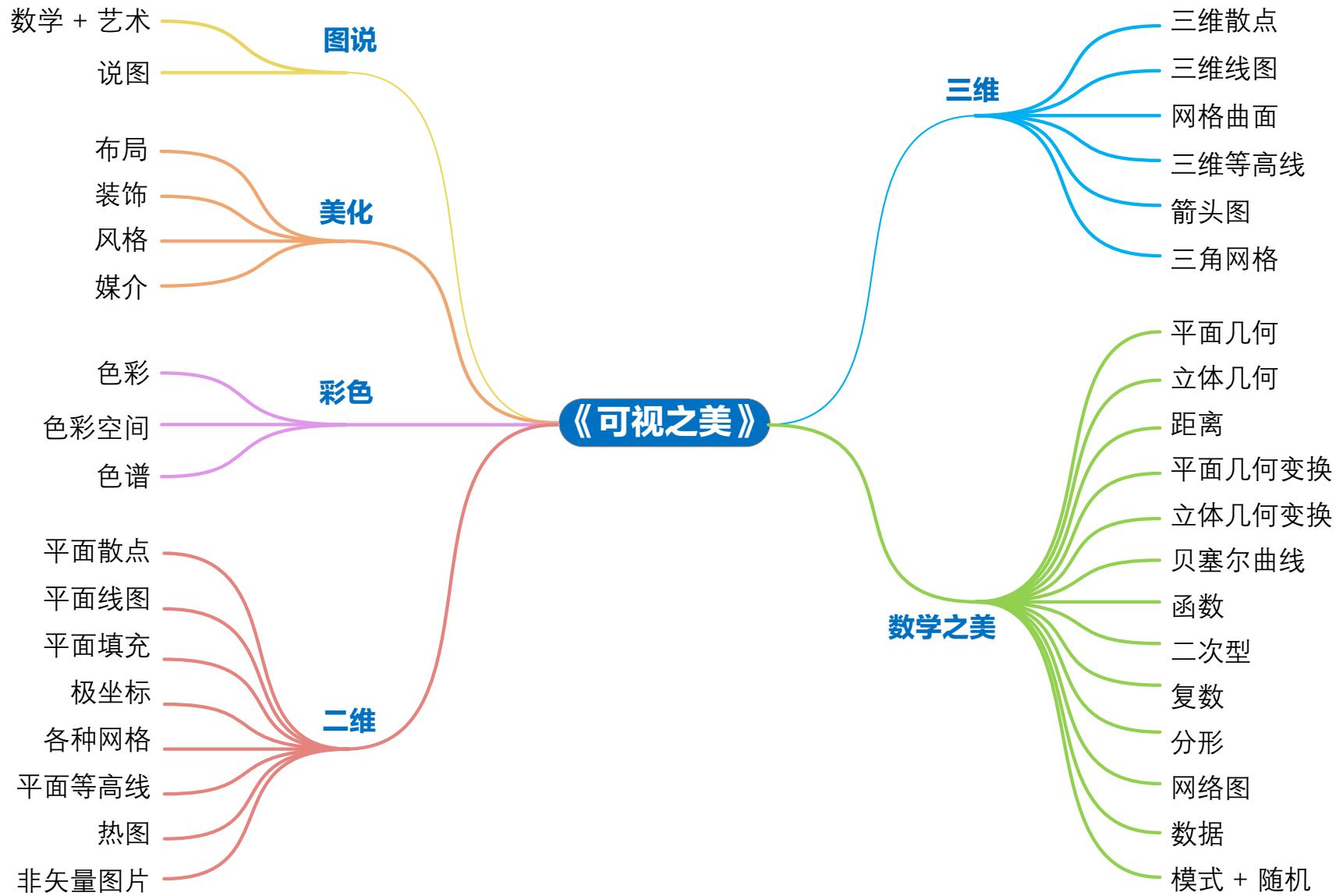
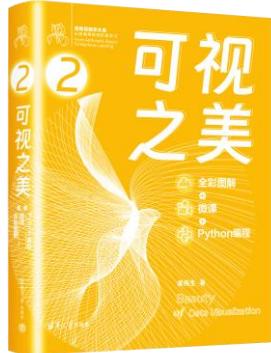
	草稿、Python	打磨、视频	清华社五审五校	上架
1 编程不难		~50%		
2 可视之美		~100%	2023, 10	
3 数学要素		100%	完成	完成
4 矩阵力量		100%	完成	完成
5 统计至简		100%	2023, 07	2023, 08
6 数据有道		80%	2024年初	
7 机器学习		80%	2024年初	



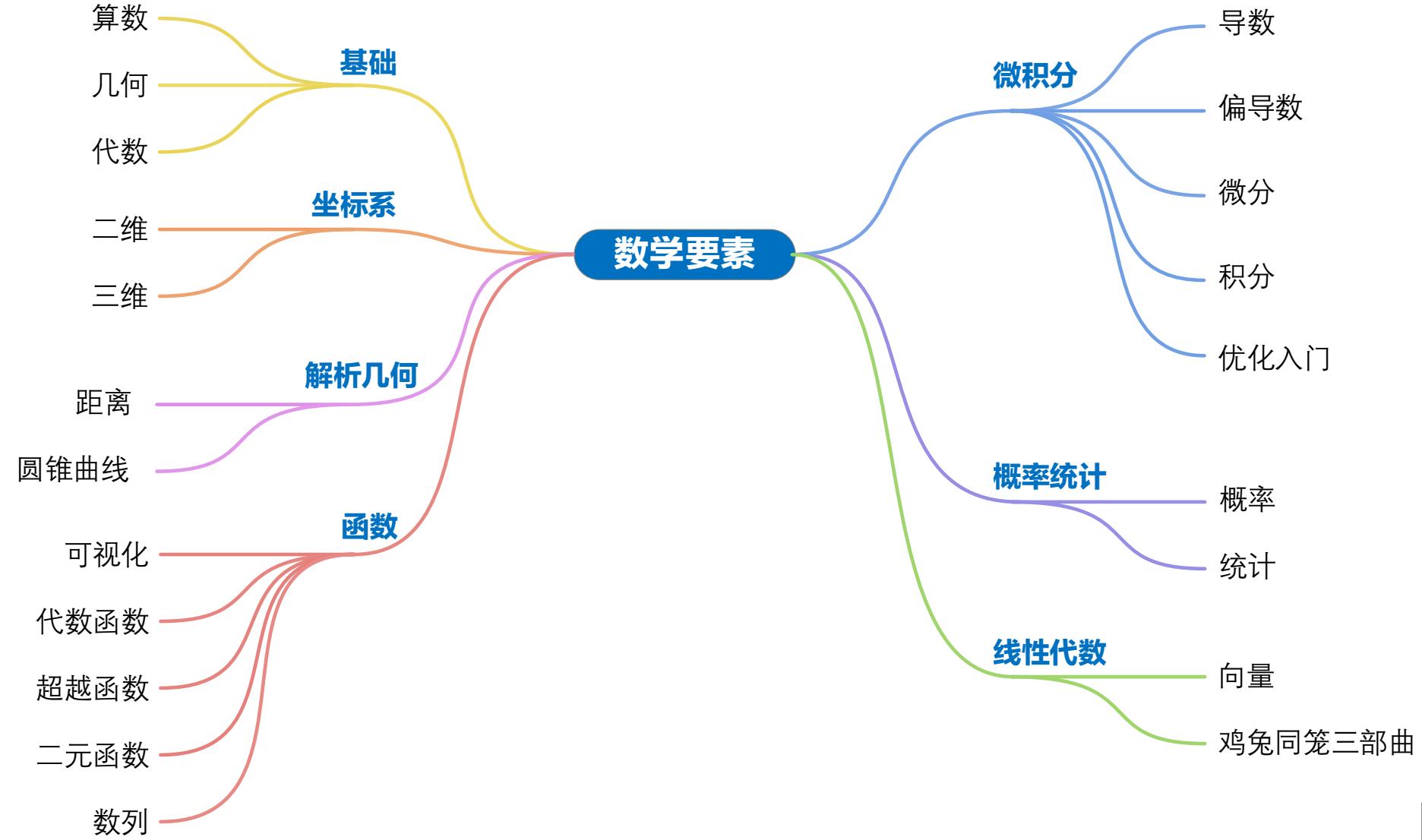
Book 1 《编程不难》



Book 2 《可视之美》



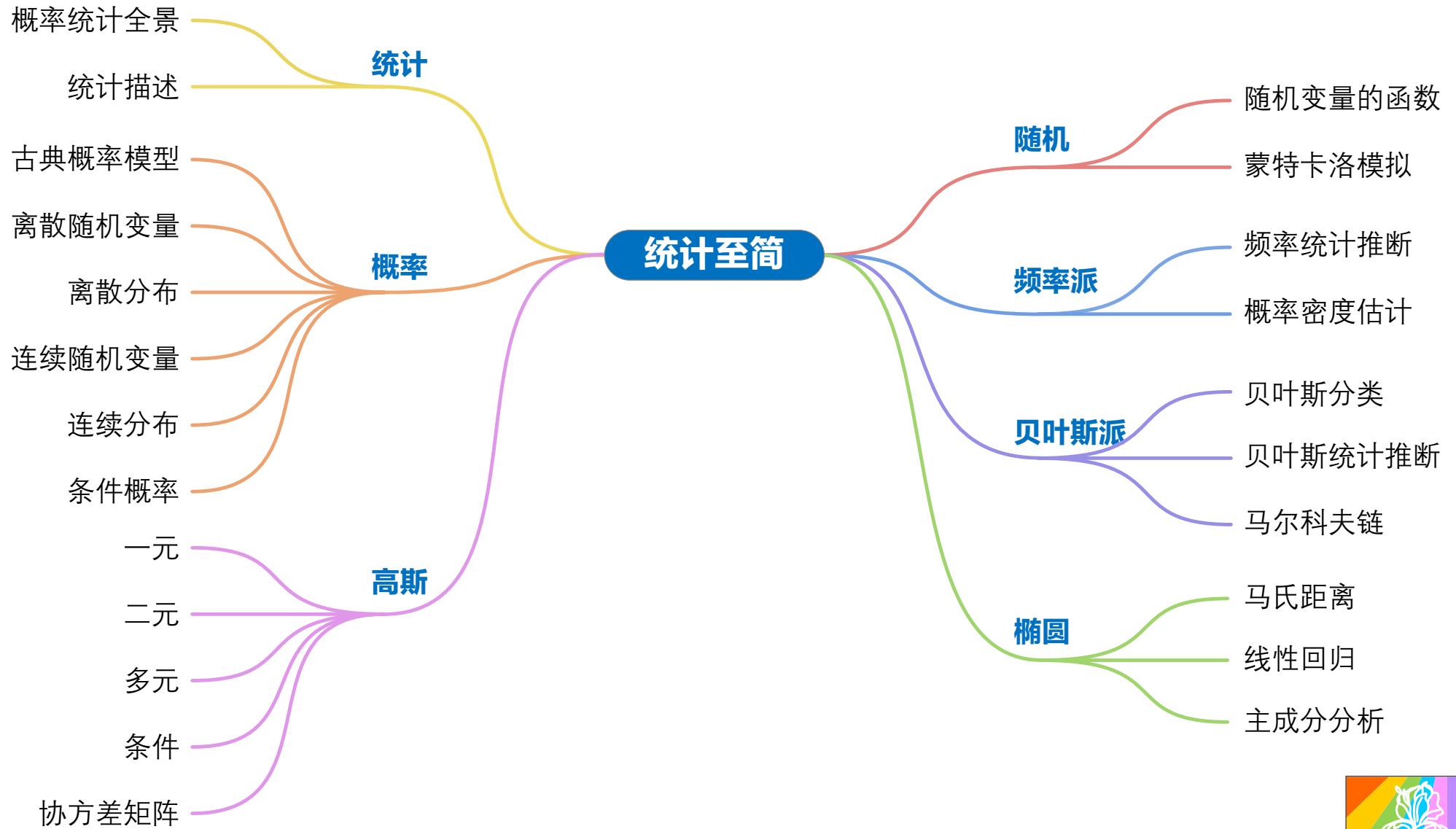
Book 3 《数学要素》



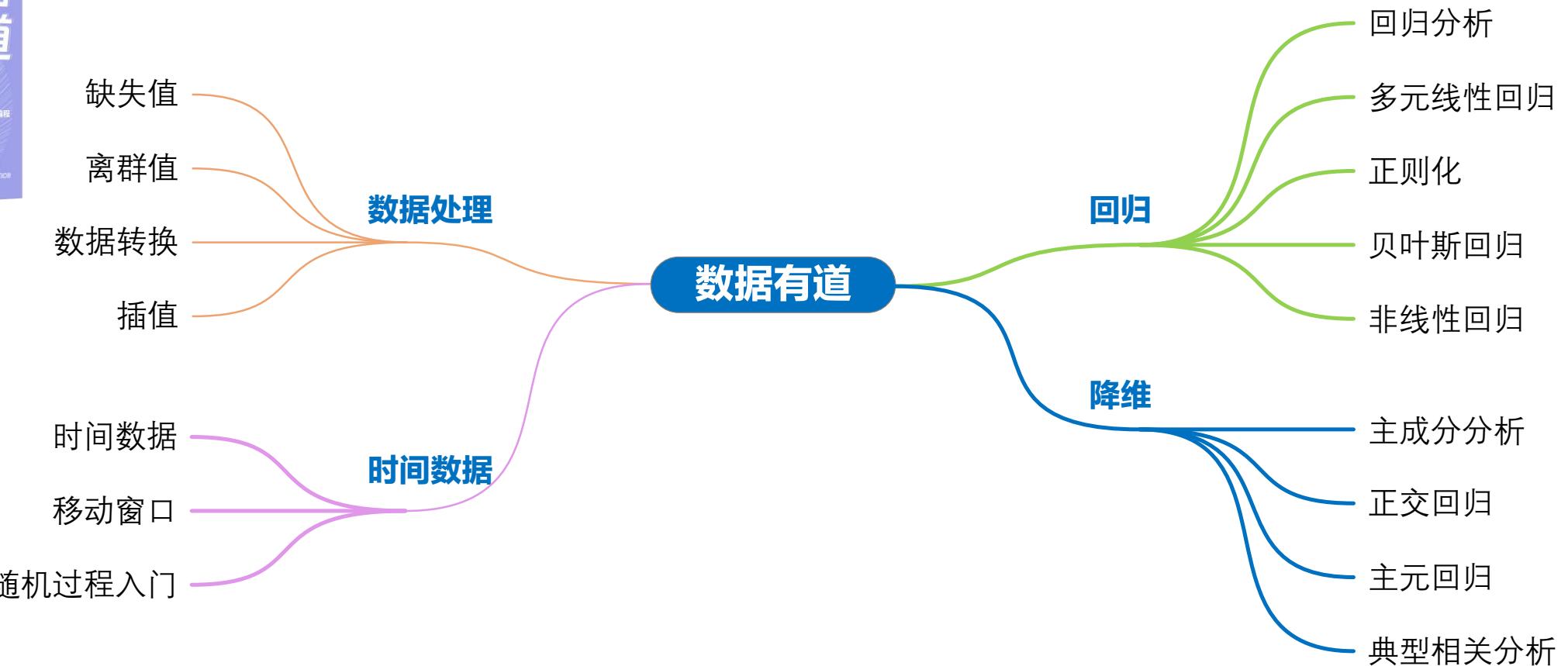
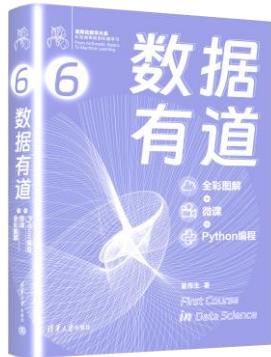
Book 4 《矩阵力量》



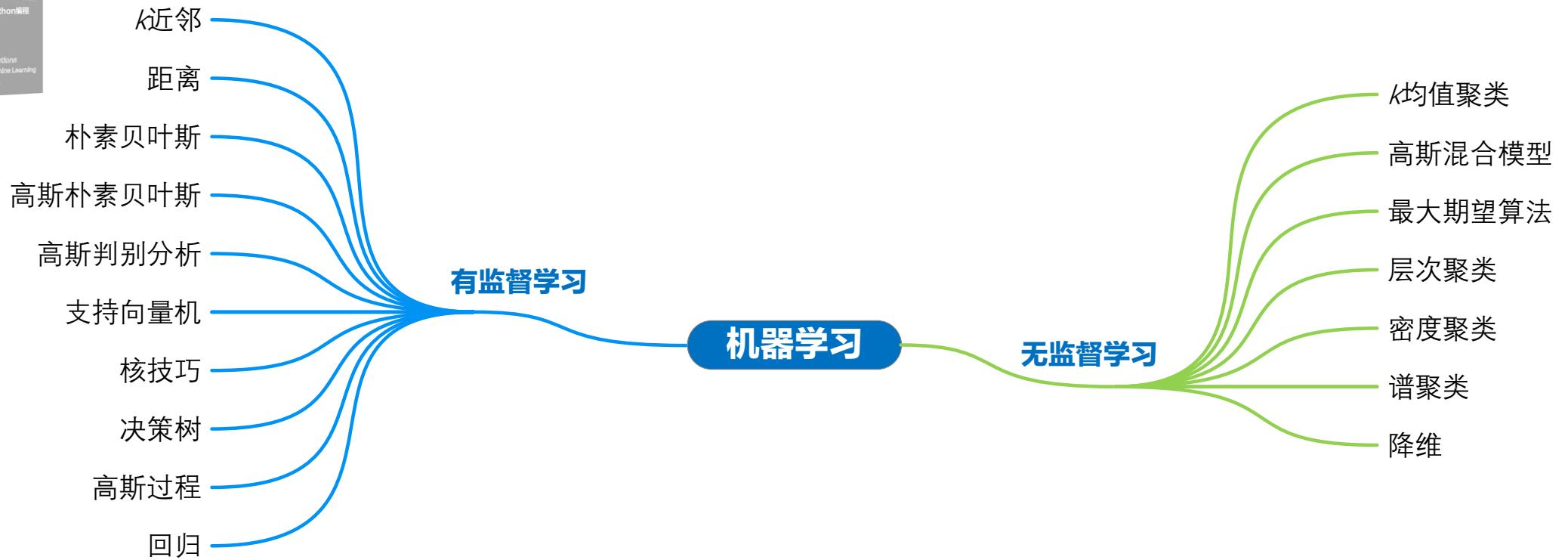
Book 5 《统计至简》



Book 6 《数据有道》



Book 7 《机器学习》



开源资源

PDF书稿、代码: <https://github.com/Visualize-ML>

微课视频: <https://space.bilibili.com/513194466>

信息发布: <https://www.zhihu.com/people/jamestong-xue>

专属邮箱: jiang.visualize.ml@gmail.com

