

作者	生姜 DrGinger
脚本	生姜 DrGinger
视频	崔崔 CuiCui
开源学习资源	https://github.com/Visualize-ML
平台	https://www.youtube.com/@DrGinger_Jiang https://space.bilibili.com/3546865719052873 https://space.bilibili.com/513194466

1.8 向量范数



本节你将掌握的核心技能：

- ▶ 向量范数：提供统一标准，将高维向量映射为可比较大小的标量。
- ▶ 三种基本范数： L^1 (绝对值和)、 L^2 (勾股定理)、 L^∞ (分量绝对值最大)。
- ▶ L^p 范数形式： L^1 到 L^∞ 的连续推广。
- ▶ 几何理解范数：平面等高线图形象理解不同范数的形状特征与变化趋势。
- ▶ 向量范数、距离关系：不同范数对应不同的距离度量方式，如欧氏距离、曼哈顿距离等。
- ▶ 成对距离矩阵结构：用矩阵表示成对点之间距离。

在高维空间中，向量的大小并不能直接比较，因此引入了向量范数来度量它们的大小。常见的向量范数包括 L^1 、 L^2 、 L^∞ 范数。 L^1 范数是向量分量的绝对值之和， L^2 范数是欧几里得距离，而 L^∞ 范数是向量分量的最大绝对值。本节在此基础上进而介绍 L^p 范数，并强化这些向量范数的几何视角。

向量谁大？谁小？

简单来说，向量范数实现了向量的大小比较。

数轴上，我们容易知道 5 比 4 大；哪怕符号不同，比如 -5 和 4 比大小，我们可以求**绝对值** (absolute value) 再比大小。绝对值实际上就是数值和原点的距离。

然而，在二维及更高维空间，向量的大小并不能直接比较。比如给定向量 $\mathbf{a} = \begin{bmatrix} 3 \\ -4 \end{bmatrix}$ 和 $\mathbf{b} = \begin{bmatrix} 0 \\ 5 \end{bmatrix}$ ，它们谁大？谁小？

为了使比较变得可行，我们引入**向量范数** (vector norm)；简单来说，向量范数提供了一种度量向量大小的方法；向量范数以特定的规则将向量映射为标量，然后比较大小。

下面让我们先看几种常见的向量范数。

L^1 范数

L^1 范数，也称为**城市街区距离** (city block distance)、**曼哈顿距离** (Manhattan Distance)，计算的是向量分量绝对值之和，即

$$\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i| \quad (1)$$

二维向量 $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ 的 L^1 范数为：

$$\left\| \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right\|_1 = |x_1| + |x_2| \quad (2)$$

例如，向量 $\mathbf{a} = [3, -4]^T$ 的 L^1 范数计算如下：

$$\|\mathbf{a}\|_1 = |3| + |-4| = 3 + 4 = 7 \quad (3)$$

再如，向量 $\mathbf{b} = [0, 5]^T$ 的 L^1 范数计算如下：

$$\|\mathbf{b}\|_1 = |0| + |5| = 0 + 5 = 5 \quad (4)$$

按照 L^1 范数的计算规则，向量 \mathbf{a} 更大。

代码 1 计算 L^1 范数。请大家注意如下两句。

- a** 用 `numpy.linalg.norm()` 计算向量范数，`ord=1` 表示计算 L^1 范数，即对向量所有元素的绝对值求和。
- b** 相当于验证上一个语句的结果。`numpy.abs(a)` 计算向量 \mathbf{a} 中每个元素的绝对值；`numpy.sum([3, 4])` 计算数组中所有元素的总和，结果是 7。

代码 1. L^1 范数 |  LA_Ch08_01_01.ipynb

```
## 初始化
import numpy as np

## 定义向量
a = np.array([3, -4])
b = np.array([0, 5])

## 计算  $L^1$  范数
a L1_a = np.linalg.norm(a, ord=1)
b np.sum(np.abs(a))

L1_b = np.linalg.norm(b, ord=1)
np.sum(np.abs(b))
```

如图 1 所示，2 维向量 $\mathbf{x} = [x_1, x_2]^T$ 的 L^1 范数在平面上的等距线 (iso-distance line) 为旋转正方形 (特殊的菱形)，满足

$$|x_1| + |x_2| = c \quad (5)$$

其中， c 为给定 L^1 范数值， $c \geq 0$ 。

图 1 中黑色箭头代表 $\mathbf{e}_1 = [1, 0]^T$ ， $\mathbf{e}_2 = [0, 1]^T$ 。 \mathbf{e}_1 、 \mathbf{e}_2 的 L^1 范数值为 1。在第一象限 ($x_1 > 0, x_2 > 0$) 上，2 维向量 $\mathbf{x} = [x_1, x_2]^T$ 的 L^1 范数为 1 满足 $x_1 + x_2 = 1$ 。

? 请大家试着写出 (5) 在不同象限的解析式。

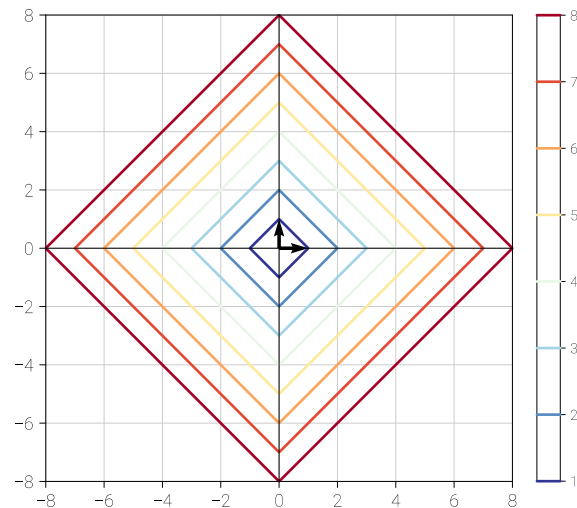



图 1.2 二维向量 L^1 范数平面等距线

代码 2 绘制图 1，下面聊聊其中关键词句。

代码 2. 绘制 L^1 范数平面等距线 |  LA_01_08_02.ipynb

```

## 初始化
import numpy as np
import matplotlib.pyplot as plt

## 生成数据
a [xx, yy = np.meshgrid(np.linspace(-8, 8, 1000),
                        np.linspace(-8, 8, 1000))

## 计算 L1 范数
b [zz = np.abs(xx) + np.abs(yy) # L1 范数计算公式

## 定义标准基向量
e1 = np.array([1, 0])
e2 = np.array([0, 1])

## 可视化
plt.figure(figsize=(6, 6))

# 绘制标准基向量 e1 和 e2
c [plt.quiver(0, 0, e1[0], e1[1],
              angles='xy', scale_units='xy',
              scale=1, color=[0, 0, 0],
              label='e1', zorder=1000)

plt.quiver(0, 0, e2[0], e2[1],
              angles='xy', scale_units='xy',
              scale=1, color=[0, 0, 0],
              label='e2', zorder=1000)

# 绘制 L1 范数的等距线
d [contour = plt.contour(xx, yy, zz,
                        levels=np.arange(1, 9),
                        cmap='RdYlBu_r')

# 添加 colorbar
cbar = plt.colorbar(contour)

# 设置坐标轴范围
plt.xlim(-8, 8); plt.ylim(-8, 8)

# 绘制网格和坐标轴
plt.axhline(0, color='black', linewidth=0.5)
plt.axvline(0, color='black', linewidth=0.5)
plt.xticks(np.arange(-8, 9, 2)); plt.yticks(np.arange(-8, 9, 2))
plt.grid(True, which='both', linestyle='--',
         linewidth=0.5, color='0.8')

# 设置等比例显示
plt.gca().set_aspect('equal', adjustable='box')

```

a 用 `numpy.meshgrid()` 生成二维网格数组，它们可以组成平面上所有的坐标点。这个过程相当于在平面上铺满网格，准备对每个点进行计算。

b 计算每个网格点的 L^1 范数。`np.abs(xx)` 是对 `xx` 中的每个数字取绝对值，`np.abs(yy)` 同理。然后把它们加起来得到 `zz`，表示每个点的 L^1 范数值。 L^1 范数可以简单理解为“横坐标和纵坐标的绝对值加起来”。

c 画向量 `e1`、`e2`。

`plt.quiver()` 是用来画箭头的函数。

前两个参数 `(0, 0)` 是箭头的起点，也就是从原点开始。

`e1[0]` 和 `e1[1]` 是箭头的方向 `(1, 0)`，表示它指向 x 轴正方向。

`angles='xy'` 和 `scale_units='xy'` 告诉程序：箭头的方向和单位是基于坐标轴的。

`scale=1` 表示不进行缩放。

`color=color_e1` 是黑色，`label='e1'` 给箭头加上标签，`zorder=1000` 表示把这个箭头放在最上面，避免被别的图遮住。

d 画出了 L^1 范数的等高线（等距线）。

`plt.contour()` 是画等高线的函数，相当于在地图上画“等高线”那样，把范数值相同的点连起来。

`xx`、`yy` 是横纵坐标网格，`zz` 是每个点的值，即 L^1 范数。

`levels=np.arange(1, 9)` 表示从 1 到 8 画出每一级的线。

`cmap='RdYlBu_r'` 表示使用“红-黄-蓝，反转”的颜色映射来填充不同等级的线，`_r` 表示颜色反转。

L^2 范数

L^2 范数，也称为**欧几里得范数** (Euclidean norm)、模，计算向量长度、向量大小：

$$\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^n x_i^2} \quad (6)$$

这是物理世界中最常用的度量，例如测量空间中的点之间的直线距离。

二维向量 $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ 的 L^2 范数为：

$$\left\| \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right\|_2 = \sqrt{x_1^2 + x_2^2} \quad (7)$$

例如，向量 $\mathbf{a} = [3, -4]^T$ 的 L^2 范数为：

$$\|\mathbf{a}\|_2 = \sqrt{3^2 + (-4)^2} = \sqrt{9+16} = \sqrt{25} = 5 \quad (8)$$

再如，向量 $\mathbf{b} = [0, 5]^T$ 的 L^2 范数计算如下：

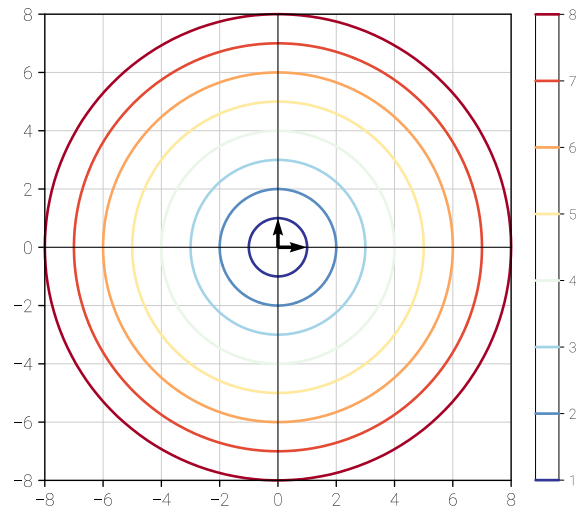
$$\|\mathbf{b}\|_2 = \sqrt{0^2 + 5^2} = \sqrt{0+25} = \sqrt{25} = 5 \quad (9)$$

按照 L^2 向量范数的计算规则，向量 \mathbf{a} 、 \mathbf{b} 相等。

如图 2 所示，2 维向量 $\mathbf{x} = [x_1, x_2]^T$ 的 L^2 范数在平面上的等距线为正圆，满足

$$\sqrt{x_1^2 + x_2^2} = c \quad (10)$$

其中， c 为给定 L^2 范数值， $c \geq 0$ 。

图 2.2 二维向量 L^2 范数平面等距线

LA_01_08_03.ipynb 计算并比较向量 \mathbf{a} 、 \mathbf{b} 的 L^2 范数大小；LA_01_08_04.ipynb 绘制图 2。两个代码文件都很简单，请大家自学。

L^∞ 范数

L^∞ 范数，也称为**切比雪夫范数** (Chebyshev Norm)，计算向量中分量绝对值最大值，即

$$\|\mathbf{x}\|_\infty = \max_i |x_i| \quad (11)$$

二维向量 $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ 的 L^∞ 范数为：

$$\left\| \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right\|_\infty = \max(|x_1|, |x_2|) \quad (12)$$

例如，向量 $\mathbf{a} = [3, -4]^T$ 的 L^∞ 范数计算如下：

$$\|\mathbf{a}\|_\infty = \max(|3|, |-4|) = 4 \quad (13)$$

再如，向量 $\mathbf{b} = [0, 5]^T$ 的 L^∞ 范数计算如下：

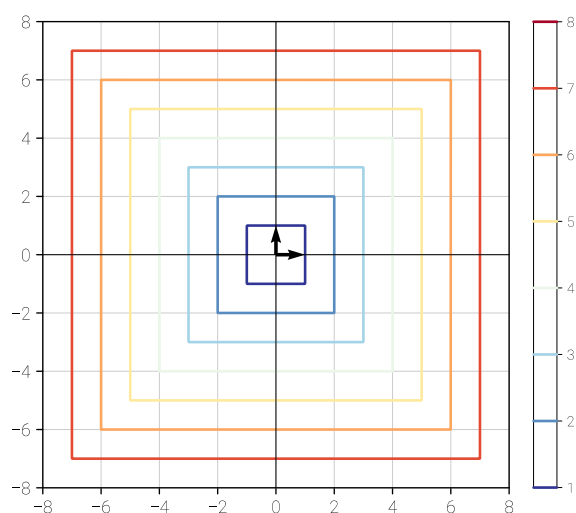
$$\|\mathbf{b}\|_\infty = \max(|0|, |5|) = 5 \quad (14)$$

按照 L^∞ 向量范数的计算规则，向量 \mathbf{b} 更大。

如图 3 所示，2 维向量 $\mathbf{x} = [x_1, x_2]^T$ 的 L^∞ 范数在平面上的等距线为正方形，满足

$$\max(|x_1|, |x_2|) = c \quad (15)$$

其中， c 为给定 L^∞ 范数值， $c \geq 0$ 。

图 3.2 二维向量 L^∞ 范数平面等距线

请大家试着写代码绘制图 3。



LA_01_08_03.ipynb 计算并比较向量 \mathbf{a} 、 \mathbf{b} 的 L^∞ 范数大小；请大家注意，正无穷用 `numpy.inf`。

L^p 范数

为了统一这些不同的范数，我们可以定义 L^p 范数：

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} \quad (16)$$

注意， $p \geq 1$ 。

二维向量 $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ 的 L^p 范数为：

$$\left\| \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right\|_p = \left(|x_1|^p + |x_2|^p \right)^{\frac{1}{p}} \quad (17)$$

如图 4 所示，在平面上，当 p 取不同值时， $\|\mathbf{x}\|_p = 1$ 对应不同形状。

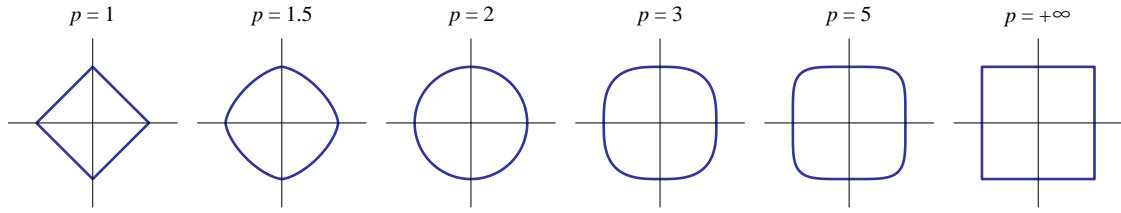
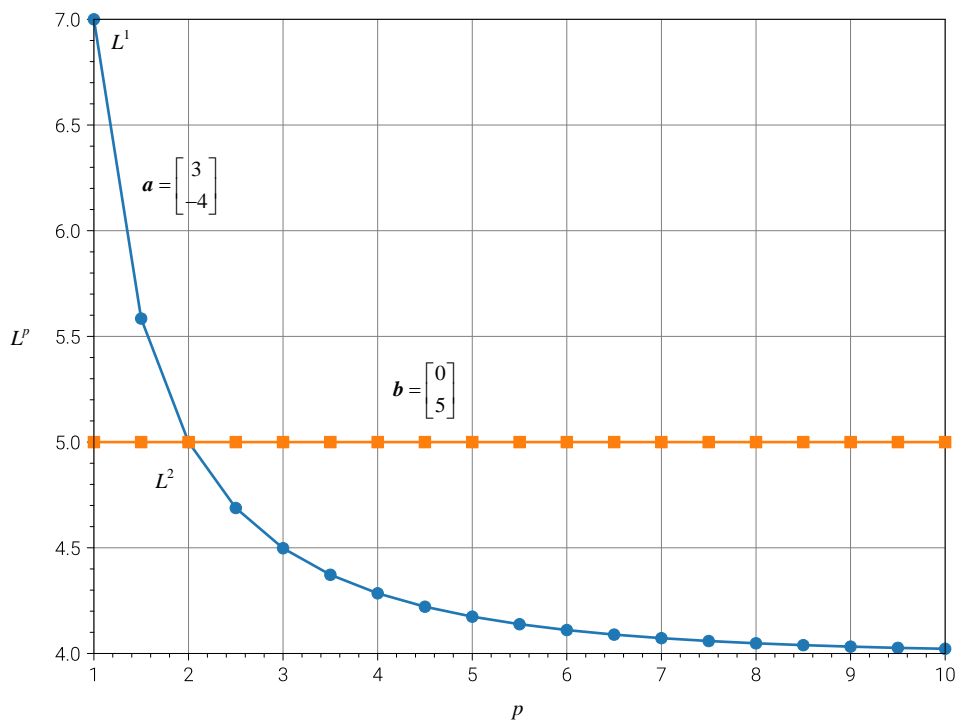
图 4. 在平面上，当 p 取不同值时， L^p 范数对应不同形状

图 5 所示为向量 $\mathbf{a} = [3, -4]^T$ 、 $\mathbf{b} = [0, 5]^T$ 的 L^p 范数随 p 变化。

随着 p 增大，向量 $\mathbf{a} = [3, -4]^T$ 的 L^p 范数不断减小。当 p 靠近 1 时，向量所有分量的贡献都被考虑；当 p 趋向 $+\infty$ 时，向量中最大的绝对值分量作为范数值。

而对于向量 $\mathbf{b} = [0, 5]^T$ ，它的 L^p 范数在 p 变化情况下不变。这是因为向量 \mathbf{b} 只有一个非零分量。

图 5. 向量 \mathbf{a} 、 \mathbf{b} 的 L^p 范数随 p 变化

叠加 L^p 范数等高线

当 $p = 1, 2, \infty$ 时，我们分别得到 L^1 、 L^2 、 L^∞ 范数。我们把这三个范数为 1 的等高线放在同一幅图上，具体如图 6 所示。

如图 6 所示，列向量 $[1/2, 1/2]^T$ 的 L^1 范数为 1，即

$$\left\| \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} \right\|_1 = |1/2| + |1/2| = 1 \quad (18)$$

$\begin{bmatrix} \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}^T$ 的 L^2 范数为 1，即

$$\left\| \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix} \right\|_{\infty} = \sqrt{(\sqrt{2}/2)^2 + (\sqrt{2}/2)^2} = 1 \quad (19)$$

$[1, 1]^T$ 的 L^∞ 范数为 1，即

$$\left\| \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\|_{\infty} = \max(|1|, |1|) = 1 \quad (20)$$

显然，图 6 中三个向量都在同一条直线上，但是向量长度（大小、 L^2 范数、欧几里得范数、模）显然不一致。

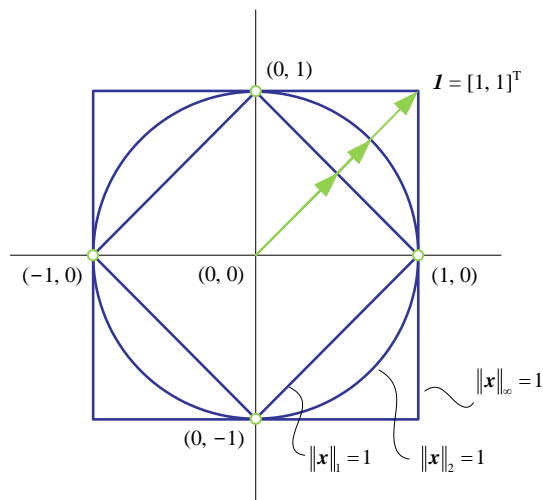


图 6. 在平面上， L^1 、 L^2 、 L^∞ 范数的形状

也请大家注意图 6 中 $(1, 0)$ 、 $(0, 1)$ 、 $(-1, 0)$ 、 $(0, -1)$ 四点处，范数 $\|x\|_p$ 均为 1。

除以上四点以外，在给定点处，随着 p 增大，范数 $\|x\|_p$ 逐渐变小。比如，对于全 1 列向量 $[1, 1]^T$ ， L^1 范数为 2， L^2 范数为 $\sqrt{2}$ ， L^∞ 无穷范数为 1。这是本节后续要展开讲解的部分。

为了更直观地比较范数大小，我们采用两种角度——相同范数值、相同向量长度。

相同范数值

为了更直观地比较不同范数，我们绘制了图 7。图中六张子图保持 L^p 范数值恒定（设定为 1），并观察 p 对满足条件的向量影响。

如图 7 (a) 所示，满足 L^1 范数为 1 的一组向量；这组向量的终点位于旋转正方形上。也就是说，图 7 (a) 中所有向量都满足 $\|x\|_1 = 1$ 。这些向量的长度（大小、 L^2 范数、欧几里得范数、模）显然“参差不齐”。只有 $(1, 0)$ 、 $(0, 1)$ 、 $(-1, 0)$ 、 $(0, -1)$ 四点对应的向量长度（ L^2 范数）为 1；其它向量，比如 $[1/2, 1/2]^T$ ，长度均小于 1。

如图 7 (b) 所示，满足 $L^{1.5}$ 范数的等高线从旋转正方形向圆形过度，边角变得柔和；图 7 (b) 给出的满足 $\|\mathbf{x}\|_{1.5} = 1$ 的一组向量。



请大家自行比较图 7 (b) 这组向量长度 (L^2 范数) 特点。

如图 7 (c) 所示， L^2 范数的等高线变成了圆形，这也是欧几里得距离的等距线；满足 $\|\mathbf{x}\|_2 = 1$ 的一组向量的终点位于单位圆圆周上。显而易见，这组向量的长度 (大小、 L^2 范数、欧几里得范数、模) 完全相同，均为 1。任意向量的 L^2 范数展现出来的是，我们对“距离”与“长度”这些概念的直观认知，它本质上衡量的是一个向量到原点的欧几里得距离。

如图 7 (d)、(e) 所示，当 p 增大到 3、5，范数对应的等高线形状变得更加方正，边缘变陡，逐渐向着更极端的形状发展。如图 7 (f) 所示， L^∞ 范数的等高线构成了一个正方形。

从向量长度 (大小、 L^2 范数、欧几里得范数、模) 来看， $(1, 0)$ 、 $(0, 1)$ 、 $(-1, 0)$ 、 $(0, -1)$ 四点对应的向量长度 (L^2 范数) 还是为 1，其它向量的向量长度 (L^2 范数) 均大于 1。

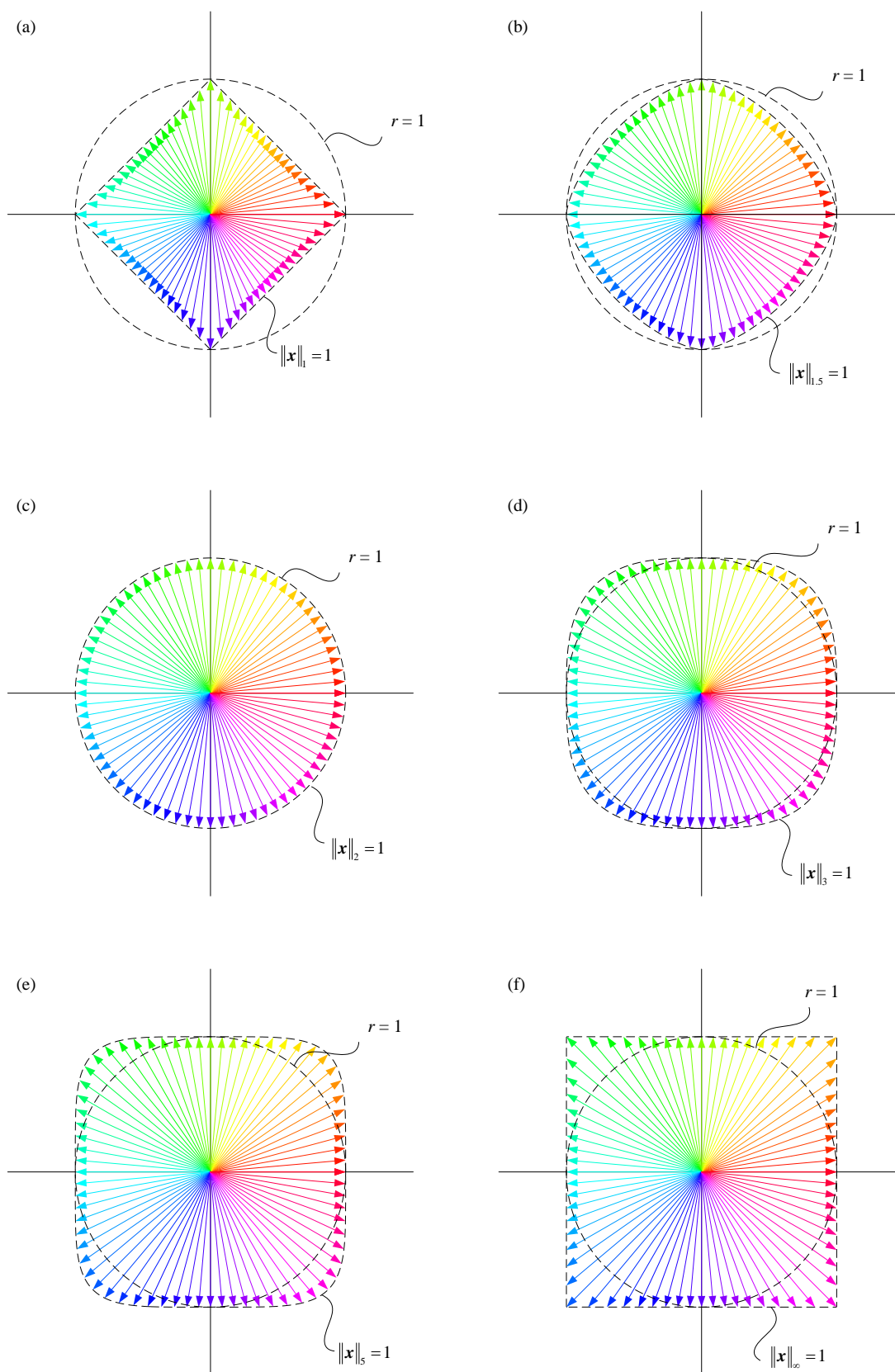


图 7. 满足范数为 1 的一组向量

相同向量长度

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

下面让我们换个视角比较不同向量范数——不同方向单位向量的 L^p 范数大小。

如图 8 (a) 所示，给定一组水平面上的单位向量；显而易见，这些向量的长度（大小、 L^2 范数、欧几里得范数、模）都为 1。这表明它们在欧几里得度量下具有相同的长度。

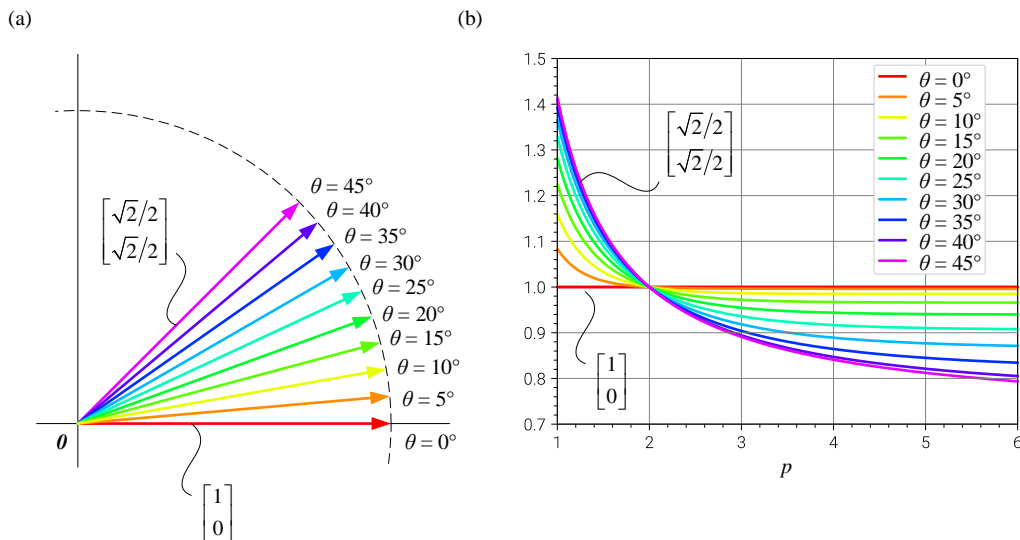


图 8. 单位向量的 L^p 范数随 p 变化

此外，这一组向量位于第一象限，和横轴正方向的夹角从 0 度逐渐变化到 45 度。

然而，当我们切换到 L^p 范数进行度量时，向量的“大小”关系则会随着 p 的变化而变化。注意，这里加引号的“大小”指的是 L^p 范数。

这在图 8 (b) 得到了直观体现；这幅图所示的是这组单位向量范数随 p 变化。

图 8 (b) 的红色线告诉我们，和横轴正方向夹角为 0 度的单位向量 $[1, 0]^T$ 的 L^p 范数大小不随 p 变化。 $[1, 0]^T$ 的 L^p 范数为

$$\left\| \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\|_p = (1^p + 0^p)^{\frac{1}{p}} = 1 \quad (21)$$

$\left[\sqrt{2}/2 \quad \sqrt{2}/2 \right]^T$ 这个单位向量对应图 8 (b) 中“陡峭”的紫色曲线，这意味着 L^p 范数大小受 p 变化影响最为剧烈。单位向量 $\left[\sqrt{2}/2 \quad \sqrt{2}/2 \right]^T$ 和横轴正方向夹角为 45 度。这个向量的 L^p 范数为

$$\left\| \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix} \right\|_p = \left(\left(\sqrt{2}/2 \right)^p + \left(\sqrt{2}/2 \right)^p \right)^{\frac{1}{p}} = 2^{\left(\frac{1}{p} - \frac{1}{2} \right)} \quad (22)$$

当 p 趋向（靠近）1 时， L^p 范数趋向 $\sqrt{2}$ ；当 p 趋向 $+\infty$ 时， L^p 范数趋向 $\sqrt{2}/2$ ；特别地，当 $p = 2$ 时， L^p 范数为 1。

通过这种对比，我们可以更深刻地理解不同方向上的单位向量在不同范数下如何变化。

$0 < p < 1$

如图9所示，对于 L^p 范数的解析式 (17)，当 $0 < p < 1$ 时，我们也能够画出等高线，但此时并非范数。等高线的形状变得更加“尖锐”，形成类似星形的图案。这种几何形状在某些优化问题或数据变换中仍有应用价值，尽管它不符合严格的范数定义。

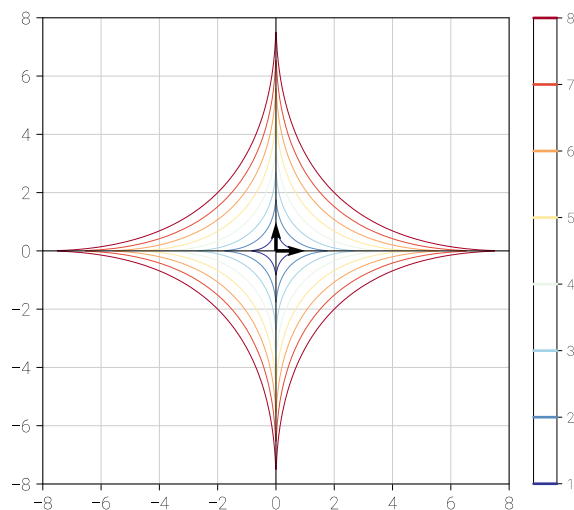


图 9. $\left(\sum_{i=1}^n |x_i|^p\right)^{\frac{1}{p}}$ 等高线, $p = 0.5$

三维向量

三维向量 $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ 的 L^p 范数为：

$$\left\| \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \right\|_p = \left(|x_1|^p + |x_2|^p + |x_3|^p \right)^{\frac{1}{p}} \quad (23)$$

图 10 所示为 3 维向量 L^p 范数等距面随 p 变化。

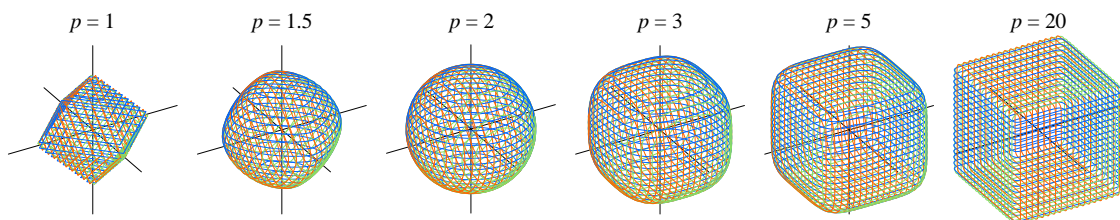


图 10. 3 维向量 L^p 范数等距面随 p 变化

3 维向量 $\mathbf{x} = [x_1, x_2, x_3]^T$ 的 L^1 范数在三维空间中的等距面 (iso-distance surface) 为旋转正方体 (特殊的正八面体), 满足

$$|x_1| + |x_2| + |x_3| = c \quad (24)$$

$\mathbf{x} = [x_1, x_2, x_3]^T$ 的 L^2 范数在三维空间中的等距面为圆球, 满足

$$x_1^2 + x_2^2 + x_3^2 = c \quad (25)$$

$\mathbf{x} = [x_1, x_2, x_3]^T$ 的 L^∞ 范数在三维空间中的等距面为正方体, 满足

$$\max(|x_1|, |x_2|, |x_3|) = c \quad (26)$$

图 11 向量 $[2, 3, 6]^T$ 、 $[0, 0, 7]^T$ 的 L^p 范数随 p 变化的两条曲线。

? 请大家自行计算 $[2, 3, 6]^T$ 、 $[0, 0, 7]^T$ 两个向量的 L^1 、 L^2 、 L^∞ 范数。

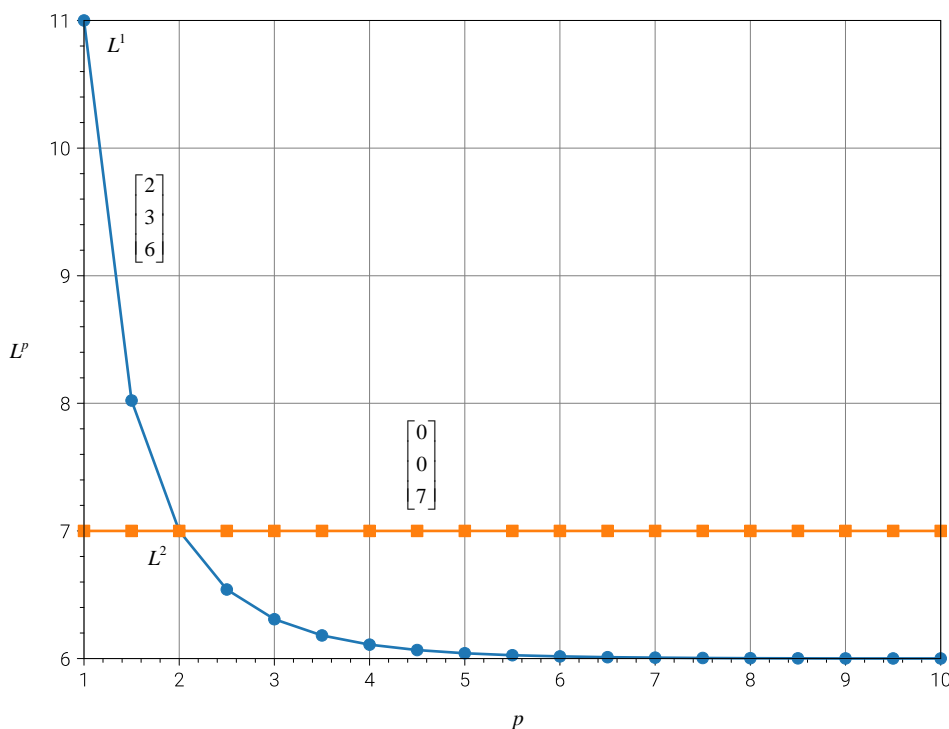


图 11. 向量 $[2, 3, 6]^T$ 、 $[0, 0, 7]^T$ 的 L^p 范数随 p 变化

RGB 颜色空间中的向量范数

让我们再回到 RGB 空间, 看看图 12 (a) 中所示常见的颜色向量的 L^p 范数随 p 变化。

图 12 (b) 所示为红色向量、绿色向量、蓝色向量 L^p 范数为 1, 不受 p 影响。

图 12 (c) 所示为黄色向量、品红向量、蓝色向量 L^p 范数随 p 变化。 p 趋向 1 时, L^p 范数趋向 2; p 趋向 $+\infty$ 时, L^p 范数趋向 1。

图 12 (d) 所示为白色向量 L^p 范数随 p 变化。 p 趋向 1 时, L^p 范数趋向 3; p 趋向 $+\infty$ 时, L^p 范数趋向 1。

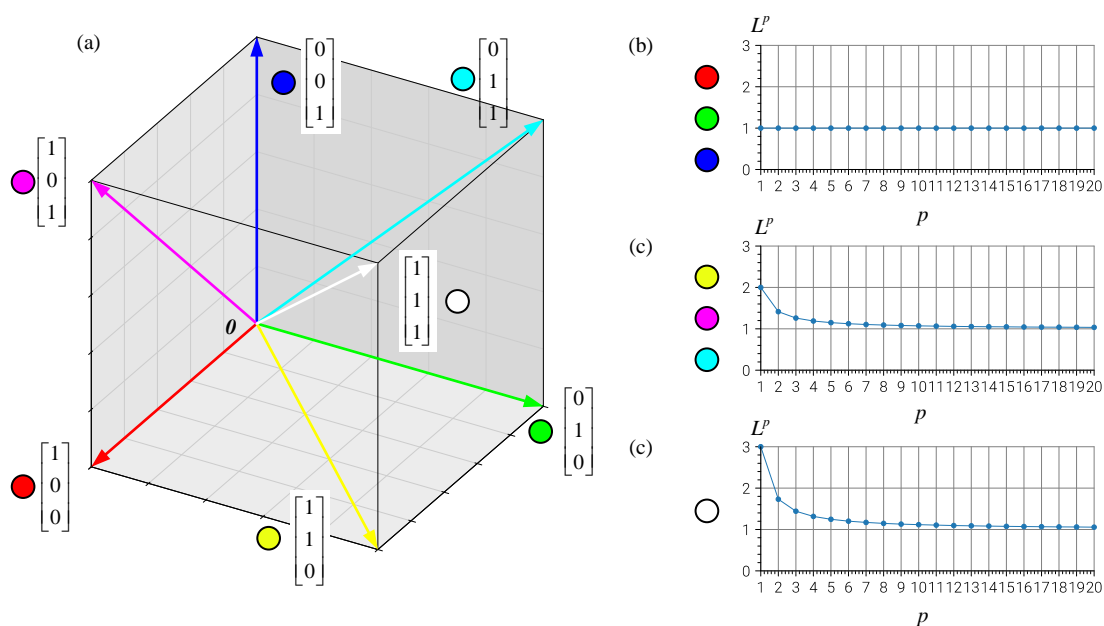
图 12. RGB 中常见颜色向量的 L^p 范数

图 13 所示为 RGB 空间中一组 L^1 范数为 1 的颜色向量。我们可以发现这组向量终点位于一个平面上。不难理解，这些颜色分量满足 $x_1 + x_2 + x_3 = 1$ 。

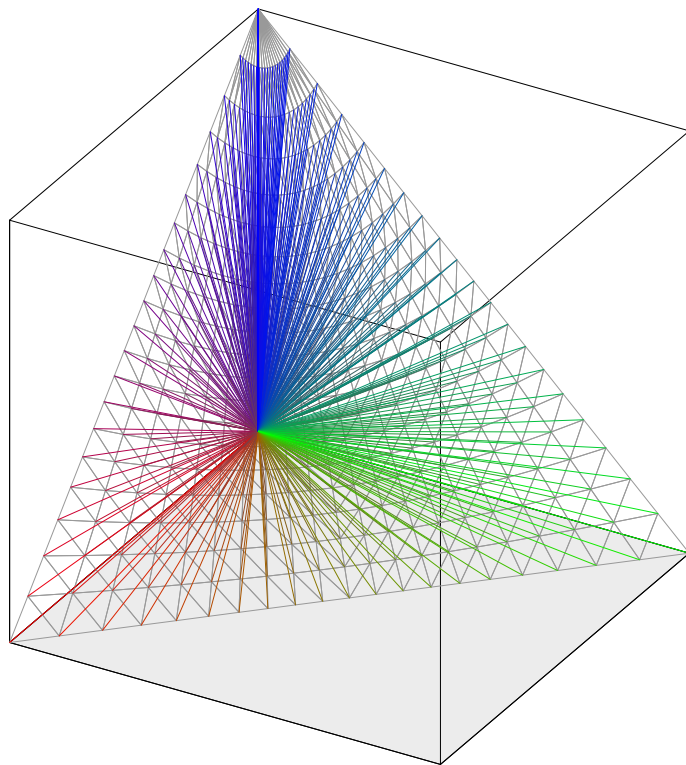
图 13. RGB 中 L^1 范数为 1 的颜色向量

图 14 所示为 RGB 空间中一组 L^2 范数为 1 的颜色向量。相信大家已经很清楚，这些向量终点位于单位圆 $1/8$ 圆面上。图 15 所示为 RGB 空间中一组 L^∞ 范数为 1 的颜色向量 (其中包括白色向量)。这些颜色向量都在 RGB 颜色立方体中最鲜亮三个立面上。

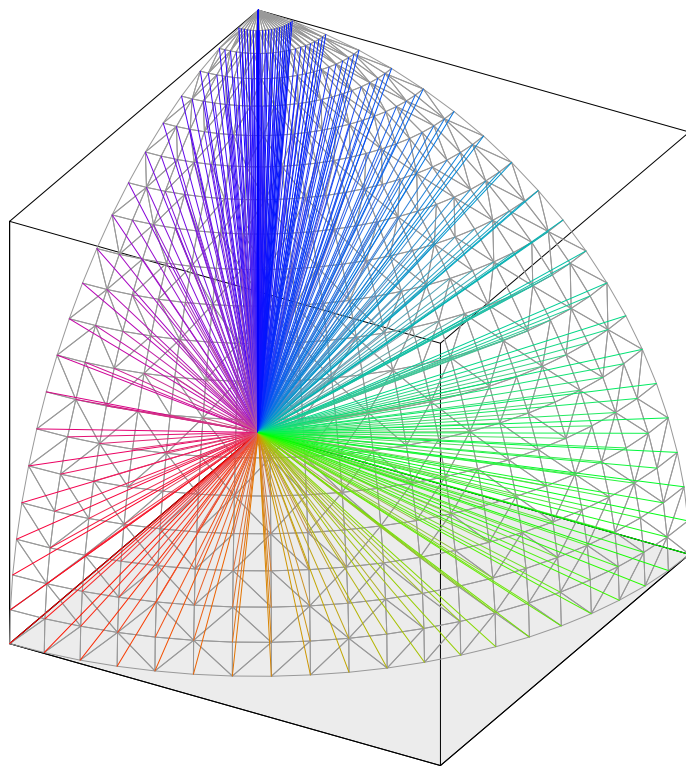


图 14. RGB 中 L^2 范数为 1 的颜色向量

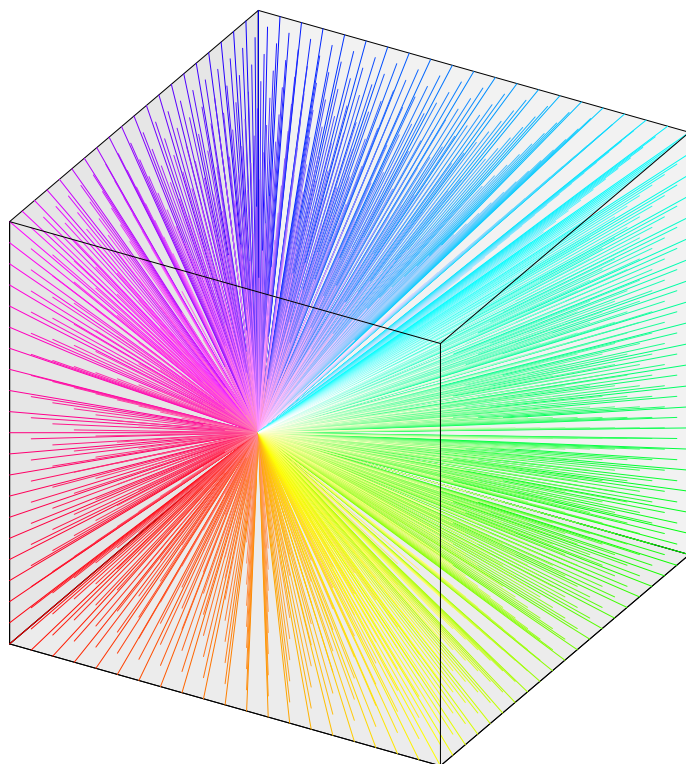


图 15. RGB 中 L^∞ 范数为 1 的颜色向量

距离度量

在数学中，我们经常关心两个向量之间有多“远”。

给定平面上两个向量 a 、 b ，

$$a = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \quad (27)$$

我们可以通过不同的“距离定义”来衡量它们之间的差异。不特别强调的话， a 、 b 的起点位于原点，这里的距离指的是 a 、 b 的终点之间的距离。

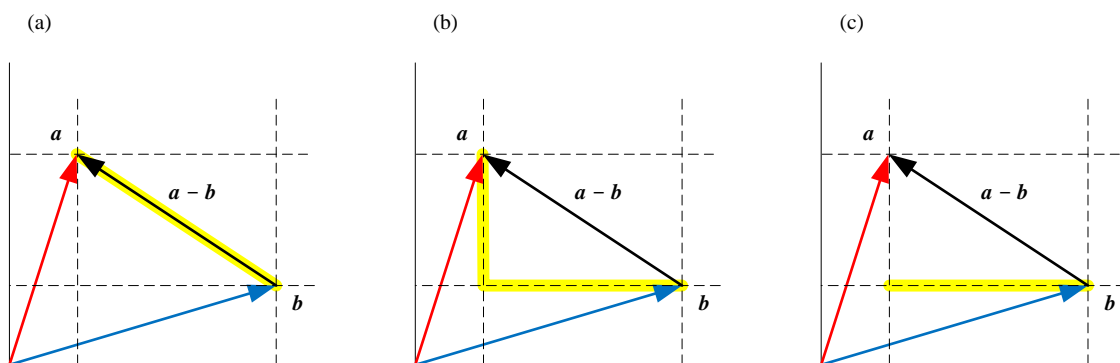


图 16. 三种不同的距离度量

用到本节前文的不同距离度量，用 L^2 范数，即欧几里得范数，度量 a 、 b 的距离

$$\|a - b\|_2 = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2} \quad (28)$$

如图 16 (a) 所示，这是我们日常最直觉的距离概念，相当于一根尺子量出的直线长度。

如果用 L^1 范数度量 a 、 b 的距离

$$\|a - b\|_1 = \sqrt{|a_1 - b_1| + |a_2 - b_2|} \quad (29)$$

如图 16 (b) 所示，就像在方方正正的城市街道中步行，只能沿着水平和垂直方向走。

如果用 L^∞ 范数度量 a 、 b 的距离

$$\|a - b\|_\infty = \max(|a_1 - b_1|, |a_2 - b_2|) \quad (30)$$

如图 16 (c) 所示， L^∞ 范数取两个方向中偏移最大的那个方向作为距离。

机器学习中还有很多距离度量，请大家自行学习。

成对距离矩阵

如图 17 所示，我们在二维平面上布置了 26 个点，分别命名为 A, B, C, \dots, Z ；其中 B 和 S 重叠， D 和 O 重叠。显然，每两个点之间都可以计算出一个距离。

用排列组合方法来算的话，一共有 325 个距离值！

这么多距离值该如何记录呢？

我们可以用矩阵！

图 18 所示为成对距离矩阵。这个矩阵的第 i 行、第 j 列元素代表第 i 个点和第 j 个点之间的距离。

在下一章中，我们将正式介绍“矩阵”这一概念。矩阵不仅能存储这样的距离信息，也能线性变换中扮演关键角色！

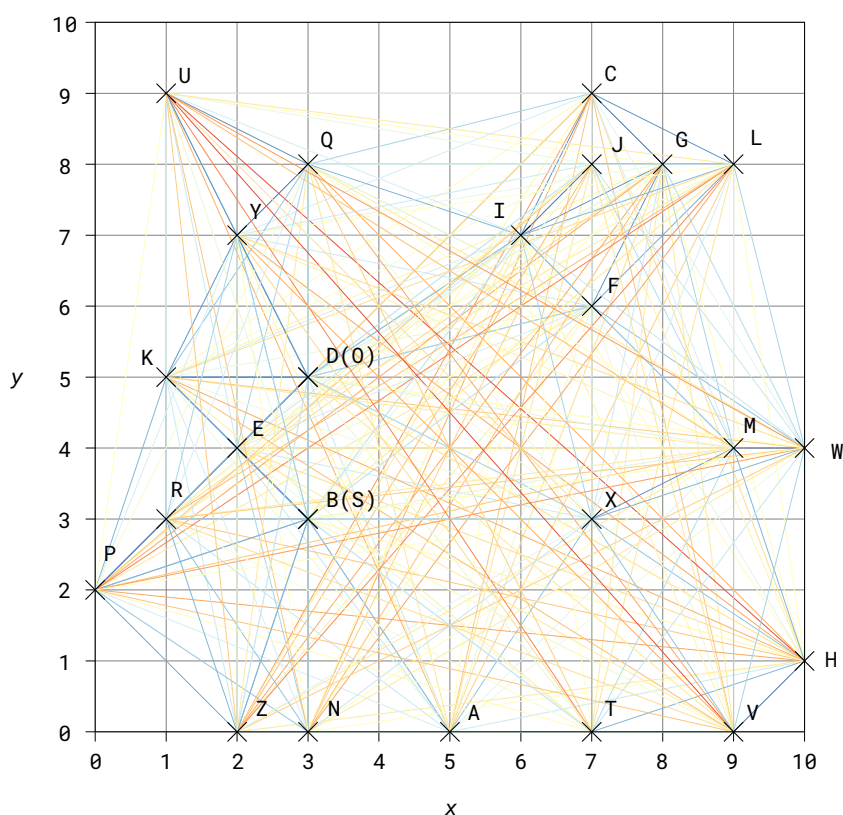


图 17. 平面上 26 个点之间的两两欧氏距离，图片来自《编程不难》

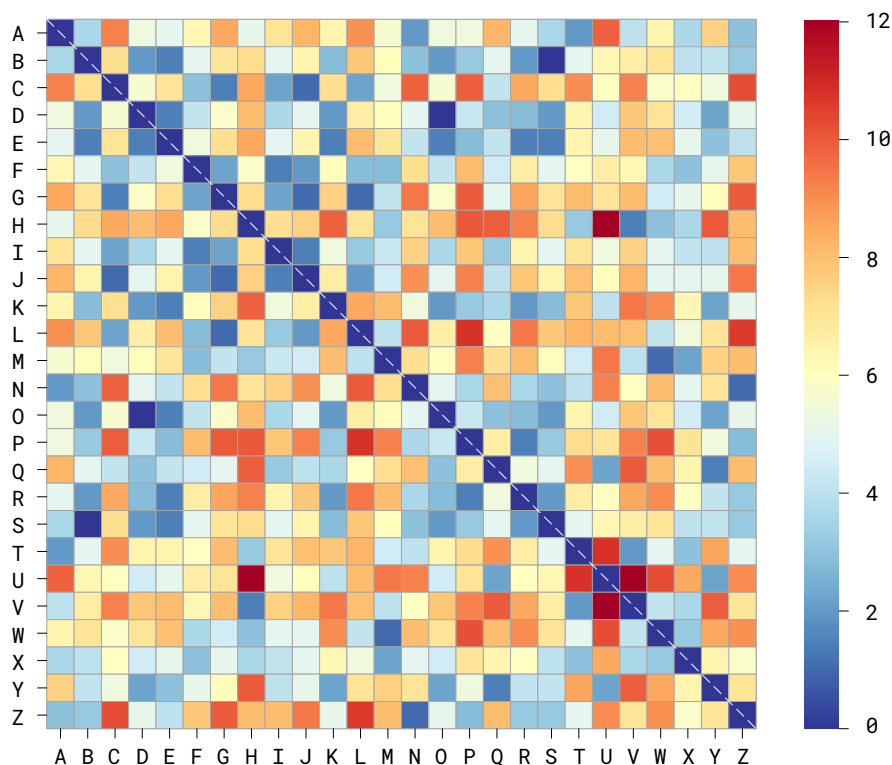


图 18. 成对欧氏距离矩阵, 图片来自《编程不难》



请大家用 DeepSeek/ChatGPT 等工具完成本节如下习题。

Q1. 学习使用 Matplotlib 绘制 $f(x_1, x_2) = x_1^2 + x_2^2$ 的等高线。

Q2. 给定向量 $[1, 2, 3]^T$, 手算其 L^1 、 L^2 、 L^∞ 范数。

Q3. 给定向量 $[1, 2, 3]^T$, 绘制向量 L^p 范数随 p 变化的曲线。

Q4. 比较以下成对向量在 p 取不同值时 L^p 范数的大小关系, 并解释为什么。

▶ $[1, 2, 2]^T, [0, 0, 3]^T$

▶ $[2, 4, 4]^T, [0, 0, 6]^T$

▶ $[2, 3, 6]^T, [0, 0, 7]^T$

▶ $[1, 4, 8]^T, [0, 0, 9]^T$

Q5. 请写 Python 代码绘制图 5 和图 11。

Q6. 什么是次可加性, 和向量范数有什么关系。

Q7. 为什么 $0 < p < 1$, (16) 不是范数?

Q8. 请计算以下颜色向量的 L^1 、 L^2 、 L^∞ 范数。

▶ 红色向量 $[1, 0, 0]^T$

▶ 黄色向量 $[1, 1, 0]^T$

► 白色向量 $[1, 1, 1]^T$

Q9. 请用 `matplotlib.pyplot.contourf()` 试着绘制图 1、图 2、图 3 的填充等高线图。