12

Cholesky Decomposition

# Cholesky 分解

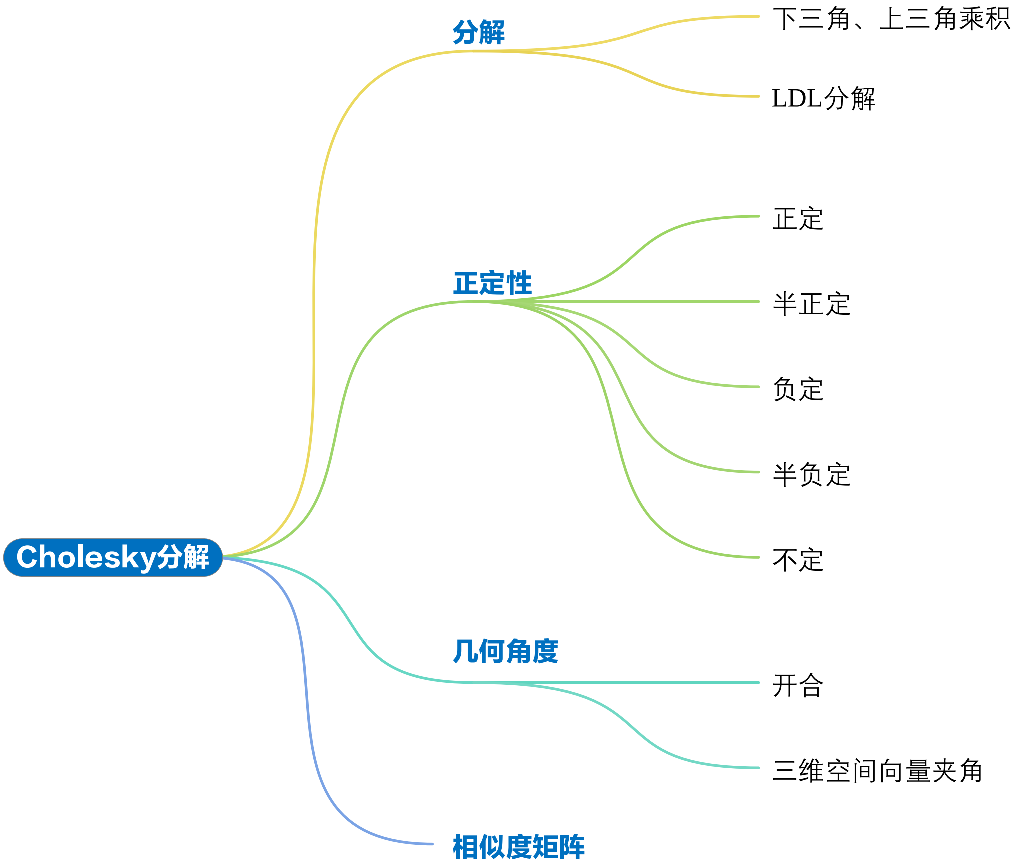
获取列向量空间位置坐标

每个人都是天才。但是，如果你以爬树的能力来判断一条鱼，那么那条鱼终其一生都会认为自己愚蠢无能。

***Everybody is a genius. But if you judge a fish by its ability to climb a tree, it will live its whole life believing that it is stupid.***

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

- ax.contour3D() 绘制三维曲面等高线
- ax.plot_wireframe() 绘制线框图
- math.radians() 将角度转换成弧度
- matplotlib.pyplot.contour()绘制平面等高线
- matplotlib.pyplot.contourf ()绘制平面填充等高线
- matplotlib.pyplot.plot() 绘制线图
- matplotlib.pyplot.quiver() 绘制箭头图
- matplotlib.pyplot.scatter() 绘制散点图
- numpy.arccos() 反余弦
- numpy.cos() 计算余弦值
- numpy.deg2rad() 将角度转化为弧度
- numpy.linalg.cholesky() Cholesky 分解
- numpy.linalg.eig() 特征值分解


Cholesky分解
分解
下三角、上三角乘积
LDL分解
正定性
正定
半正定
负定
半负定
不定
几何角度
开合
三维空间向量夹角
相似度矩阵

# 12.1 Cholesky 分解

实数矩阵的 Cholesky 分解由法国军官、数学家**安德烈·路易·科列斯基** (André-Louis Cholesky) 最先发明。科列斯基本人在一战结束前夕战死沙场，Cholesky 分解是由科列斯基本的同事在他死后发表的，并以科列斯基的名字命名。

通过上一章学习，大家知道 Cholesky 分解将方阵 $\boldsymbol{A}$ 分解为一个下三角矩阵 $\boldsymbol{L}$ 以及它的转置 $\boldsymbol{L}^{\mathrm{T}}$ 的乘积：

$$\boldsymbol{A} = \boldsymbol{L}\boldsymbol{L}^{\mathrm{T}} \tag{1}$$

利用上三角矩阵 $\boldsymbol{R}$ (= $\boldsymbol{L}^{\mathrm{T}}$)，(1) 可以写成：

$$\boldsymbol{A} = \boldsymbol{R}^{\mathrm{T}}\boldsymbol{R} \tag{2}$$

## LDL 分解

在 Cholesky 分解基础上，上一章又介绍了 LDL 分解。LDL 分解将上述矩阵 $\boldsymbol{A}$ 分解成下三角矩阵 $\boldsymbol{L}$、对角阵方阵 $\boldsymbol{D}$、$\boldsymbol{L}^{\mathrm{T}}$ 三者乘积，即，

$$\boldsymbol{A} = \boldsymbol{L}\boldsymbol{D}\boldsymbol{L}^{\mathrm{T}} \tag{3}$$

⚠ 注意，(3) 中下三角矩阵 $\boldsymbol{L}$ 为对角线元素均为 1。

假设对角方阵 $\boldsymbol{D}$ 对角线元素非负，LDL 分解可以进一步写成：

$$\boldsymbol{A} = \boldsymbol{L}\boldsymbol{D}^{1/2}\left(\boldsymbol{D}^{1/2}\right)^{\mathrm{T}}\boldsymbol{L}^{\mathrm{T}} = \boldsymbol{L}\boldsymbol{D}^{1/2}\left(\boldsymbol{L}\boldsymbol{D}^{1/2}\right)^{\mathrm{T}} \tag{4}$$

$\boldsymbol{D}^{1/2}$ 也是个对角方阵，$\boldsymbol{D}^{1/2}$ 对角线上元素是 $\boldsymbol{D}$ 的对角线元素的非负平方根。

令，

$$\boldsymbol{B} = \boldsymbol{D}^{1/2} \tag{5}$$

(4) 可写成：

$$\boldsymbol{A} = \boldsymbol{L}\boldsymbol{B}\left(\boldsymbol{L}\boldsymbol{B}\right)^{\mathrm{T}} \tag{6}$$

用上三角矩阵 $\boldsymbol{R}$ 替换 $\boldsymbol{L}^{\mathrm{T}}$，(6) 可以写成：

$$\boldsymbol{A} = \boldsymbol{R}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{B}\boldsymbol{R} = \left(\boldsymbol{B}\boldsymbol{R}\right)^{\mathrm{T}}\boldsymbol{B}\boldsymbol{R} \tag{7}$$

# 12.2 正定矩阵才可以进行 Cholesky 分解

上一章提到，并非所有矩阵都可以做 Cholesky 分解，只有**正定矩阵** (positive-definite matrix) 才可以完成 Cholesky 分解。

在 $\boldsymbol{x}$ 为非零列向量 ($\boldsymbol{x} \neq \boldsymbol{0}$) 条件下，如果方阵 $\boldsymbol{A}$ 满足：

$$\boldsymbol{x}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{x} > 0 \tag{8}$$

则称方阵 $\boldsymbol{A}$ 为**正定矩阵** (positive definite matrix)。

⚠ 注意，(8) 中列向量 $\boldsymbol{x}$ 的行数和矩阵 $\boldsymbol{A}$ 行数一致。二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 的结果是标量。此外，正定矩阵的特征值均为正。

## 几何视角

从几何角度更容易理解正定矩阵，以如下 $2 \times 2$ 矩阵为例：

$$\boldsymbol{A} = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \tag{9}$$

定义二元函数 $y = f(x_1, x_2)$：

$$\begin{aligned} y &= f\left(x_1, x_2\right) \\ &= \boldsymbol{x}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{x} \\ &= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ &= a x_1^2 + 2b x_1 x_2 + c x_2^2 \end{aligned} \tag{10}$$

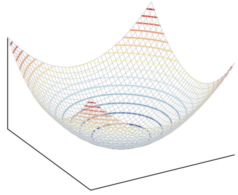
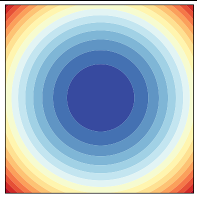
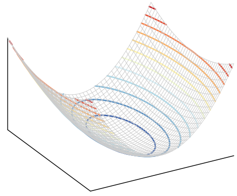
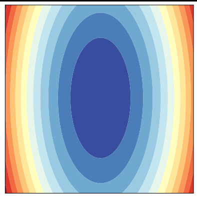
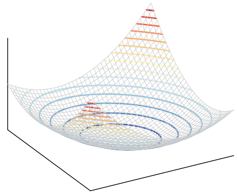
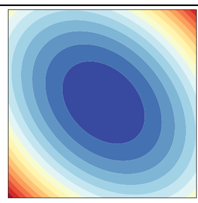
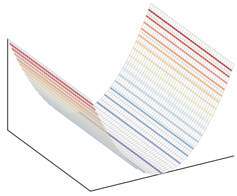
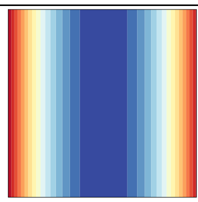
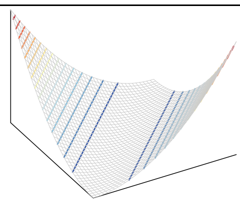
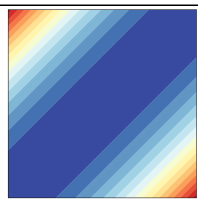
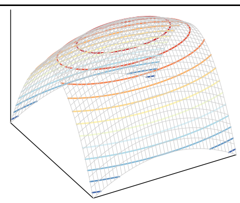
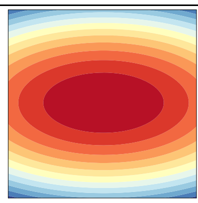
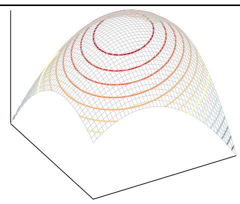
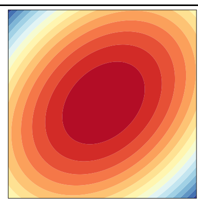
函数 $y = f(x_1, x_2)$ 就是本书第 5 章提到的二次型。更重要的是，上式把正定性和丛书《数学要素》讲过的二次曲面联系起来。

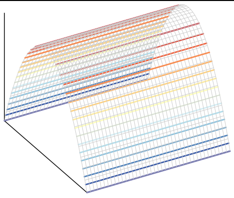
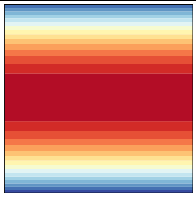
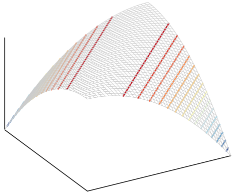
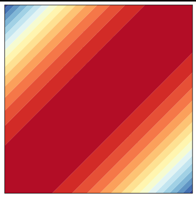
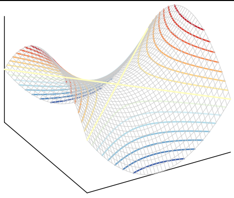
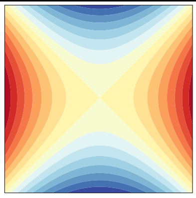
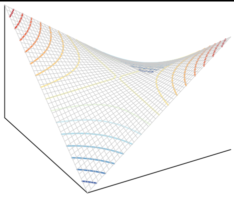
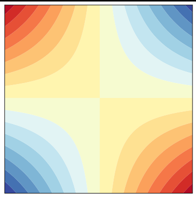
除了正定矩阵，还有半正定、负定、半负定、不定这几种正定性。表 1 总结几种正定性、曲面、等高线特征。希望读者能够通过表中几何图形建立正定性的直观印象。此外，请大家自行分析表中曲面的极值特征。

➜ 本书第 21 章将专门讨论矩阵的正定性。

表 1. 几种正定性

| 正定性 | 例子 | 三维曲面 | 平面等高线 |
| --- | --- | --- | --- |

| | | | |
|---|---|---|---|
| 正定 (positive definite) | 开口向上正圆抛物面 $\boldsymbol{A}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$ |  |  |
| | 开口向上正椭圆抛物面 $\boldsymbol{A}=\begin{bmatrix}2 & 0\\ 0 & 0.5\end{bmatrix}$ |  |  |
| | 开口向上旋转椭圆抛物面 $\boldsymbol{A}=\begin{bmatrix}1.5 & 0.5\\ 0.5 & 1.5\end{bmatrix}$ |  |  |
| 半正定 (positive semi-definite) | 山谷面 $\boldsymbol{A}=\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}$ |  |  |
| | 旋转山谷面 $\boldsymbol{A}=\begin{bmatrix}0.5 & -0.5\\ -0.5 & 0.5\end{bmatrix}$ |  |  |
| 负定 (negative definite) | 开口向下正椭圆抛物面 $\boldsymbol{A}=\begin{bmatrix}-0.5 & 0\\ 0 & -2\end{bmatrix}$ |  |  |
| | 开口向下旋转椭圆抛物面 $\boldsymbol{A}=\begin{bmatrix}-1.5 & 0.5\\ 0.5 & -1.5\end{bmatrix}$ |  |  |

| | | | |
|---|---|---|---|
| 半负定 (negative semi-definite) | 山脊面 $A=\begin{bmatrix}0 & 0\\ 0 & -1\end{bmatrix}$ |  |  |
| | 旋转山脊面 $A=\begin{bmatrix}-0.5 & 0.5\\ 0.5 & -0.5\end{bmatrix}$ |  |  |
| 不定 (indefinite) | 马鞍面 $A=\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}$ |  |  |
| | 旋转马鞍面 $A=\begin{bmatrix}0 & -1\\ -1 & 0\end{bmatrix}$ |  |  |

Bk4_Ch12_01.py 绘制表 1 三维曲面和等高线。请注意改变 a、b、c 三个系数取值。

# 12.3 几何角度：开合

本节，我们从一个有趣的几何视角分析一种特殊矩阵的 Cholesky 分解。

## 以 2 × 2 矩阵为例

给定如 2 × 2 矩阵 $\boldsymbol{P}$，它的主对角元素为 1，非主对角线元素为余弦值 $\cos\theta_{1,2}$：

$$\boldsymbol{P}=\begin{bmatrix}1 & \cos\theta_{1,2}\\ \cos\theta_{1,2} & 1\end{bmatrix} \tag{11}$$

对矩阵 $\boldsymbol{P}$ 进行 Cholesky 分解可以得到：

$$\boldsymbol{P}=\boldsymbol{L}\boldsymbol{L}^{\mathrm{T}}=\underbrace{\begin{bmatrix}1 & 0\\ \cos\theta_{1,2} & \sin\theta_{1,2}\end{bmatrix}}_{\boldsymbol{L}}\underbrace{\begin{bmatrix}1 & \cos\theta_{1,2}\\ 0 & \sin\theta_{1,2}\end{bmatrix}}_{\boldsymbol{L}^{\mathrm{T}}}=\begin{bmatrix}1 & \cos\theta_{1,2}\\ \cos\theta_{1,2} & 1\end{bmatrix} \tag{12}$$

利用上三角矩阵 $\boldsymbol{R}$，矩阵 $\boldsymbol{P}$ 的 Cholesky 分解还可以写成：

$$\boldsymbol{P}=\boldsymbol{R}^{\mathrm{T}}\boldsymbol{R}=\underbrace{\begin{bmatrix}1 & 0\\ \cos\theta_{1,2} & \sin\theta_{1,2}\end{bmatrix}}_{\boldsymbol{R}^{\mathrm{T}}}\underbrace{\begin{bmatrix}1 & \cos\theta_{1,2}\\ 0 & \sin\theta_{1,2}\end{bmatrix}}_{\boldsymbol{R}} \tag{13}$$

将 $\boldsymbol{R}$ 写成：

$$\boldsymbol{R}=\begin{bmatrix}1 & \cos\theta_{1,2}\\ 0 & \sin\theta_{1,2}\end{bmatrix}=\begin{bmatrix}\boldsymbol{r}_1 & \boldsymbol{r}_2\end{bmatrix} \tag{14}$$

在平面直角坐标系中，$\boldsymbol{e}_1$ 和 $\boldsymbol{e}_2$ 分别代表水平和竖直正方向的单位向量，$[\boldsymbol{e}_1, \boldsymbol{e}_2]$ 是 $\mathbb{R}^2$ 空间的标准正交基。$\boldsymbol{R}$ 分别乘 $\boldsymbol{e}_1$ 和 $\boldsymbol{e}_2$，得到 $\boldsymbol{r}_1$ 和 $\boldsymbol{r}_2$：

$$\begin{aligned}\boldsymbol{r}_1&=\boldsymbol{R}\boldsymbol{e}_1=\begin{bmatrix}1 & \cos\theta_{1,2}\\ 0 & \sin\theta_{1,2}\end{bmatrix}\begin{bmatrix}1\\ 0\end{bmatrix}=\begin{bmatrix}1\\ 0\end{bmatrix}\\ \boldsymbol{r}_2&=\boldsymbol{R}\boldsymbol{e}_2=\begin{bmatrix}1 & \cos\theta_{1,2}\\ 0 & \sin\theta_{1,2}\end{bmatrix}\begin{bmatrix}0\\ 1\end{bmatrix}=\begin{bmatrix}\cos\theta_{1,2}\\ \sin\theta_{1,2}\end{bmatrix}\end{aligned} \tag{15}$$

很容易判断 $\boldsymbol{r}_1$ 和 $\boldsymbol{r}_2$ 均为单位向量。

而向量 $\boldsymbol{r}_1$ 和 $\boldsymbol{r}_2$ 夹角余弦值正是 $\cos\theta_{1,2}$：

$$\cos\theta=\frac{\boldsymbol{r}_1\cdot\boldsymbol{r}_2}{\left\|\boldsymbol{r}_1\right\|\left\|\boldsymbol{r}_2\right\|}=\cos\theta_{1,2} \tag{16}$$

## 几何视角

如图 1 所示，从几何角度来讲，$\boldsymbol{R}$ ($\boldsymbol{L}^{\mathrm{T}}$) 的相当于把原本正交的 $[\boldsymbol{e}_1, \boldsymbol{e}_2]$ 标准正交基转化成具有一定夹角的 $[\boldsymbol{r}_1, \boldsymbol{r}_2]$ 非正交基，且 $\boldsymbol{e}_1 = \boldsymbol{r}_1$，相当于“锚定”。这类似 QR 分解中，$\boldsymbol{q}_1$ 平行于 $\boldsymbol{x}_1$。

⚠ 再次强调，虽然 $[\boldsymbol{r}_1, \boldsymbol{r}_2]$ 中每个列向量为单位向量，但是并不正交，因此 $[\boldsymbol{r}_1, \boldsymbol{r}_2]$ 为非正交基。

图 1 所示这种几何变换像是“门合页”的开合。我们给这种几何变换取个名字，就叫做“开合”。

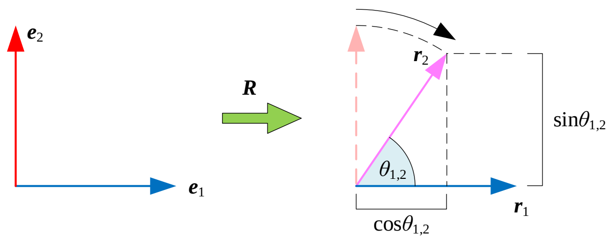



图 1. 开合

图 2 所示为四种不同开合角度。$0 < \cos\theta_{1,2} < 1$ 时，即 $0° < \theta_{1,2} < 90°$，“门合页”从直角 90°关小角度至 $\theta_{1,2}$，具体如图 2 (a) (b) 所示两例。

$-1 < \cos\theta_{1,2} < 0$ 时，即 $90° < \theta_{1,2} < 180°$，“合页”从直角 90°打开至 $\theta_{1,2}$，具体如图 2 (c) (d) 所示两例。

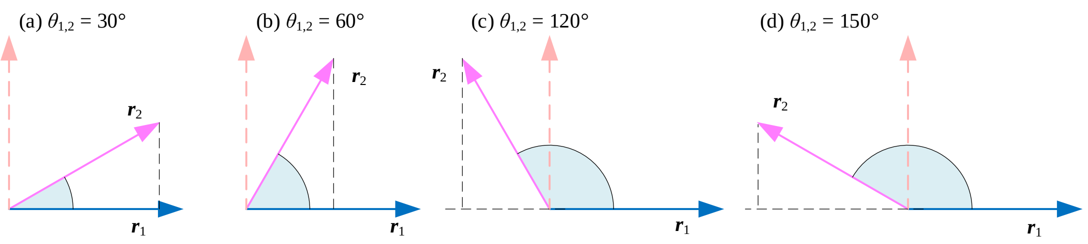



图 2. 不同的开合角度 $\cos\theta_{1,2}$

## 行列式值

计算 (14) 中 $\boldsymbol{R}$ 的行列式值：

$$|\boldsymbol{R}| = \begin{vmatrix} 1 & \cos\theta_{1,2} \\ 0 & \sin\theta_{1,2} \end{vmatrix} = \sin\theta_{1,2} \tag{17}$$

这个行列式值结果表明“开合”前后，图形的面积缩比为 $\sin\theta_{1,2}$。这和我们在图 3 中看到一致。$[\boldsymbol{e}_1, \boldsymbol{e}_2]$ 构造正方形面积为 1，而 $[\boldsymbol{r}_1, \boldsymbol{r}_2]$ 构造的平行四边形面积为 $\sin\theta_{1,2}$。

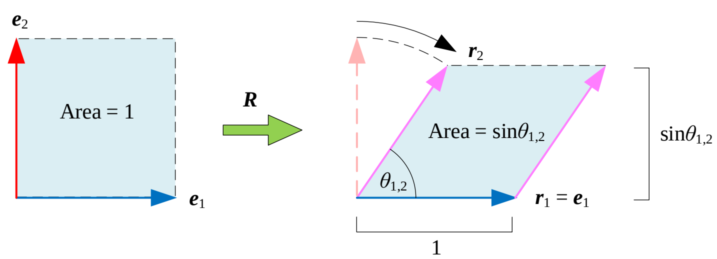


图 3. 开合对应的面积变化

### 举个例子

给定 $\boldsymbol{P}$ 为:

$$\boldsymbol{P}=\begin{bmatrix} 1 & \cos 60° \\ \cos 60° & 1 \end{bmatrix}=\begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix} \tag{18}$$

对 $\boldsymbol{P}$ 进行 Cholesky 分解得到:

$$\boldsymbol{P}=\boldsymbol{R}^{\mathrm{T}}\boldsymbol{R}=\underbrace{\begin{bmatrix} 1 & 0 \\ 0.5 & \sqrt{3}/2 \end{bmatrix}}_{\boldsymbol{R}^{\mathrm{T}}}\underbrace{\begin{bmatrix} 1 & 0.5 \\ 0 & \sqrt{3}/2 \end{bmatrix}}_{\boldsymbol{R}} \tag{19}$$

图 4 所示为 $\boldsymbol{e}_1$ 和 $\boldsymbol{e}_2$ 经过 (19) 中 $\boldsymbol{R}$ 转换得到向量 $\boldsymbol{r}_1$ 和 $\boldsymbol{r}_2$，而正圆经过 $\boldsymbol{R}$ 转换变成椭圆。

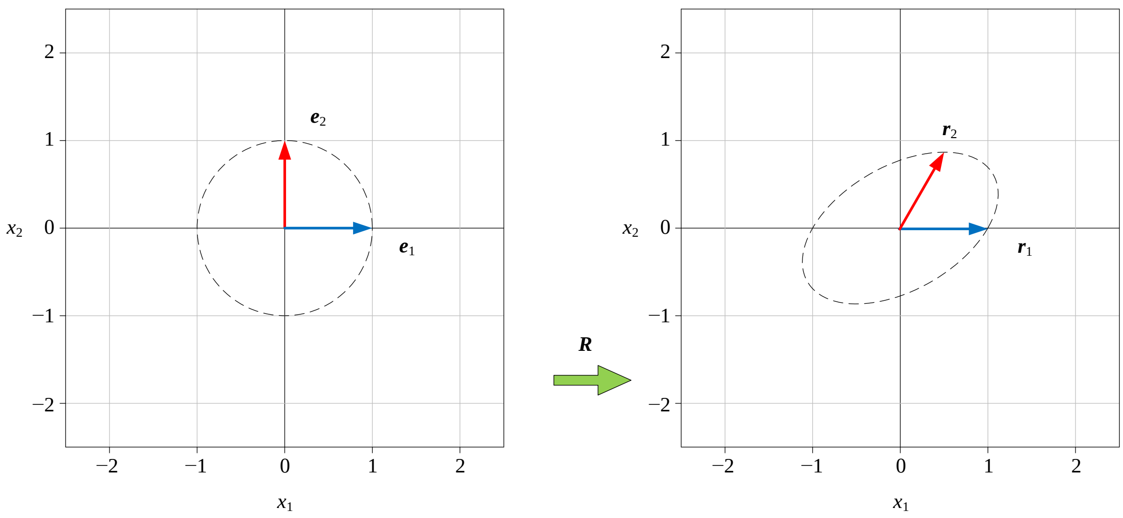



图 4. $\boldsymbol{e}_1$ 和 $\boldsymbol{e}_2$ 经过 $\boldsymbol{R}$ 转换得到向量 $\boldsymbol{r}_1$ 和 $\boldsymbol{r}_2$

## 12.4 几何变换：缩放 → 开合

给定 $\boldsymbol{\Sigma}$ 具体形式如下:

$$\boldsymbol{\Sigma}=\begin{bmatrix} a^2 & a\cdot b\cdot\cos\theta_{1,2} \\ a\cdot b\cdot\cos\theta_{1,2} & b^2 \end{bmatrix} \tag{20}$$

其中，$a$ 和 $b$ 都是正数。

先把 $\boldsymbol{\Sigma}$ 写成:

$$\Sigma = \underbrace{\begin{bmatrix} a & \\ & b \end{bmatrix}}_{S} \underbrace{\begin{bmatrix} 1 & \cos\theta_{1,2} \\ \cos\theta_{1,2} & 1 \end{bmatrix}}_{P} \underbrace{\begin{bmatrix} a & \\ & b \end{bmatrix}}_{S} \tag{21}$$

将 (11) 代入 (21)，得到：

$$\Sigma = (RS)^{\mathrm{T}}(RS) = \underbrace{\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}}_{S} \underbrace{\begin{bmatrix} 1 & 0 \\ \cos\theta_{1,2} & \sin\theta_{1,2} \end{bmatrix}}_{R^{\mathrm{T}}} \underbrace{\begin{bmatrix} 1 & \cos\theta_{1,2} \\ 0 & \sin\theta_{1,2} \end{bmatrix}}_{R} \underbrace{\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}}_{S} \tag{22}$$

上式相当于对 $\Sigma$ 直接进行 Cholesky 分解的结果。

将 $RS$ ($S$ 先、$R$ 后) 作用在在 $e_1$ 和 $e_2$ 上，得到 $x_1$ 和 $x_2$：

$$\begin{aligned} x_1 = RSe_1 &= \begin{bmatrix} 1 & \cos\theta_{1,2} \\ 0 & \sin\theta_{1,2} \end{bmatrix} \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = a \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ x_2 = RSe_2 &= \begin{bmatrix} 1 & \cos\theta_{1,2} \\ 0 & \sin\theta_{1,2} \end{bmatrix} \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = b \begin{bmatrix} \cos\theta_{1,2} \\ \sin\theta_{1,2} \end{bmatrix} \end{aligned} \tag{23}$$

这相当于，对 $e_1$ 和 $e_2$ 先缩放 ($S$)，再开合 ($R$)。

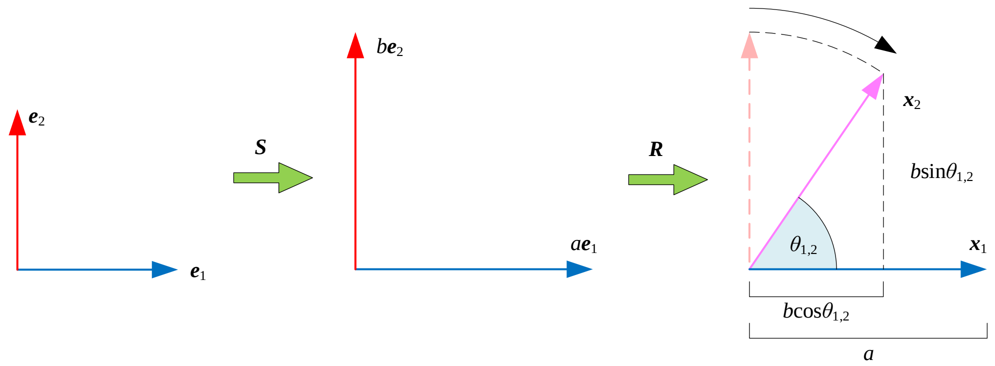



图 5. 先缩放再开合

计算 (23) 中，向量 $x_1$ 和 $x_2$ 夹角余弦值为：

$$\cos\theta = \frac{x_1 \cdot x_2}{\|x_1\|\|x_2\|} = \frac{a \cdot b \cdot \cos\theta_{1,2}}{a \cdot b} = \cos\theta_{1,2} \tag{24}$$

发现向量 $x_1$ 和 $x_2$ 夹角等同于向量 $r_1$ 和 $r_2$ 夹角夹角。

## 举个例子

给定 $\Sigma$ 具体值为：

$$\boldsymbol{\Sigma}=\begin{bmatrix}1.5^2 & 1.5\times 2\times\cos 60° \\ 1.5\times 2\times\cos 60° & 2^2\end{bmatrix}=\begin{bmatrix}2.25 & 1.5 \\ 1.5 & 4\end{bmatrix} \tag{25}$$

对 $\boldsymbol{\Sigma}$ 进行 Cholesky 分解得到：

$$\boldsymbol{\Sigma}=\left(\boldsymbol{R}_{\Sigma}\right)^{\mathrm{T}}\left(\boldsymbol{R}_{\Sigma}\right)=\begin{bmatrix}1.5 & 0 \\ 1 & 1.732\end{bmatrix}\begin{bmatrix}1.5 & 1 \\ 0 & 1.732\end{bmatrix} \tag{26}$$

图 6 所示为 $\boldsymbol{e}_1$ 和 $\boldsymbol{e}_2$ 经过 $\boldsymbol{R}_{\Sigma}$ 转换得到向量 $\boldsymbol{x}_1$ 和 $\boldsymbol{x}_2$。

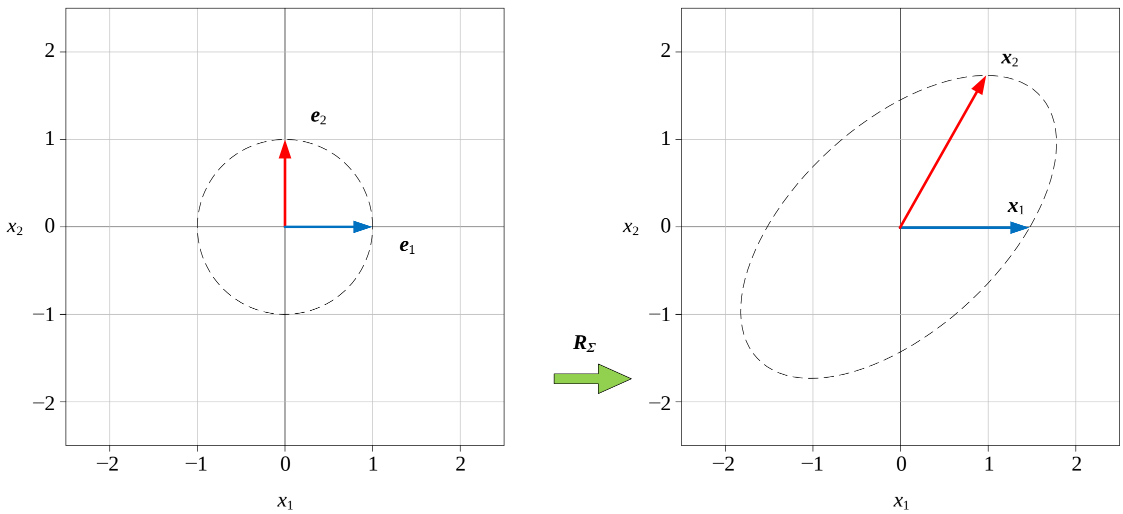



图 6. $\boldsymbol{e}_1$ 和 $\boldsymbol{e}_2$ 经过 $\boldsymbol{R}_{\Sigma}$ 转换得到向量 $\boldsymbol{x}_1$ 和 $\boldsymbol{x}_2$

按照 (22)，$\boldsymbol{\Sigma}$ 可以分解成：

$$\boldsymbol{\Sigma}=\underbrace{\begin{bmatrix}1.5 & 0 \\ 0 & 2\end{bmatrix}}_{\boldsymbol{S}}\underbrace{\begin{bmatrix}1 & 0 \\ 0.5 & \sqrt{3}/2\end{bmatrix}}_{\boldsymbol{R}^{\mathrm{T}}}\underbrace{\begin{bmatrix}1 & 0.5 \\ 0 & \sqrt{3}/2\end{bmatrix}}_{\boldsymbol{R}}\underbrace{\begin{bmatrix}1.5 & 0 \\ 0 & 2\end{bmatrix}}_{\boldsymbol{S}} \tag{27}$$

图 7 所示为 $\boldsymbol{e}_1$ 和 $\boldsymbol{e}_2$ 分别经过 $\boldsymbol{S}$ 和 $\boldsymbol{R}$ 转换，得到向量 $\boldsymbol{x}_1$ 和 $\boldsymbol{x}_2$。

本系列丛书一般用 $\boldsymbol{\Sigma}$ 来代表协方差矩阵。本节之所以用矩阵 $\boldsymbol{\Sigma}$，这是因为大家很快会发现 Cholesky 分解和协方差矩阵之间的紧密联系。而本章前文中提到的矩阵 $\boldsymbol{P}$，就是本书之后要讲的相关性系数矩阵。类比的话，矩阵 $\boldsymbol{P}$ 中的余弦值就是相关性系数。

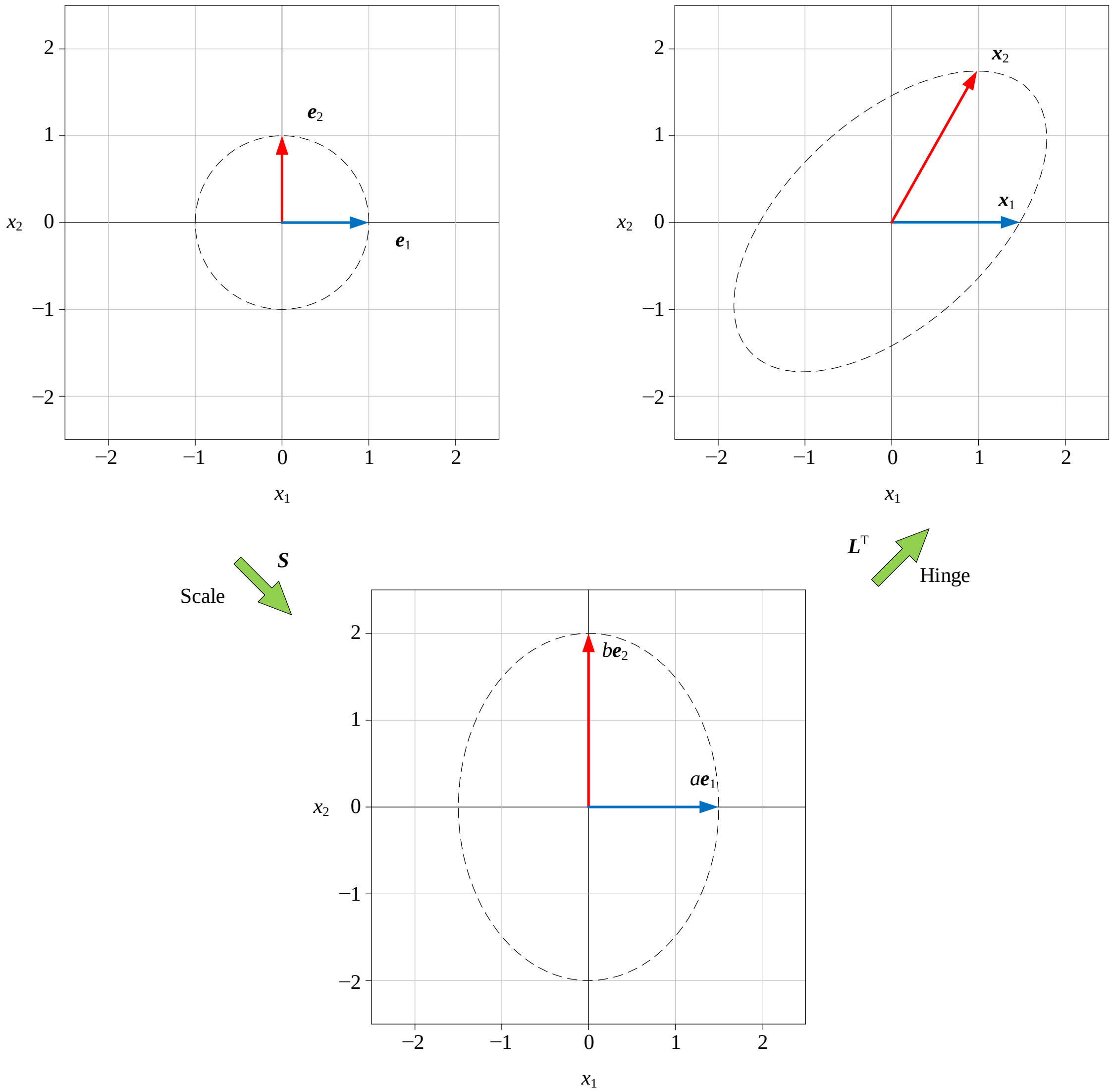



图 7. $\boldsymbol{e}_1$ 和 $\boldsymbol{e}_2$ 分别经过 $\boldsymbol{S}$ 和 $\boldsymbol{R}$ 转换

# 12.5 推广到三维空间

本节利用立体几何视角探讨 Cholesky 分解。

给定如下 3 × 3 矩阵 $\boldsymbol{P}$,

$$\boldsymbol{P} = \begin{bmatrix} 1 & \cos\theta_{1,2} & \cos\theta_{1,3} \\ \cos\theta_{1,2} & 1 & \cos\theta_{2,3} \\ \cos\theta_{1,3} & \cos\theta_{2,3} & 1 \end{bmatrix} \tag{28}$$

其中，$\theta_{1,2}$、$\theta_{1,3}$、$\theta_{2,3}$ 三个角度均大于等于 0°。

对 $\boldsymbol{P}$ 进行 Cholesky 分解:

$$\boldsymbol{P}=\boldsymbol{R}^{\mathrm{T}}\boldsymbol{R} \tag{29}$$

其中,

$$\boldsymbol{R}=\begin{bmatrix} 1 & \cos\theta_{1,2} & \cos\theta_{1,3} \\ 0 & \sqrt{1-\cos\theta_{1,2}^2} & \dfrac{\cos\theta_{2,3}-\cos\theta_{1,3}\cos\theta_{1,2}}{\sqrt{1-\cos\theta_{1,2}^2}} \\ 0 & 0 & \sqrt{1-\cos\theta_{1,3}^2-\dfrac{\left(\cos\theta_{2,3}-\cos\theta_{1,3}\cos\theta_{1,2}\right)^2}{1-\cos\theta_{1,2}^2}} \end{bmatrix} \tag{30}$$

相当于:

$$\boldsymbol{r}_1=\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},\quad \boldsymbol{r}_2=\begin{bmatrix} \cos\theta_{1,2} \\ \sqrt{1-\cos\theta_{1,2}^2} \\ 0 \end{bmatrix},\quad \boldsymbol{r}_2=\begin{bmatrix} \cos\theta_{1,3} \\ \dfrac{\cos\theta_{2,3}-\cos\theta_{1,3}\cos\theta_{1,2}}{\sqrt{1-\cos\theta_{1,2}^2}} \\ \sqrt{1-\cos\theta_{1,3}^2-\dfrac{\left(\cos\theta_{2,3}-\cos\theta_{1,3}\cos\theta_{1,2}\right)^2}{1-\cos\theta_{1,2}^2}} \end{bmatrix} \tag{31}$$

将 $\boldsymbol{R}=[\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3]$ 代入 (29) 得到:

$$\boldsymbol{P}=\begin{bmatrix} 1 & \cos\theta_{1,2} & \cos\theta_{1,3} \\ \cos\theta_{1,2} & 1 & \cos\theta_{2,3} \\ \cos\theta_{1,3} & \cos\theta_{2,3} & 1 \end{bmatrix}=\begin{bmatrix} \boldsymbol{r}_1^{\mathrm{T}} \\ \boldsymbol{r}_2^{\mathrm{T}} \\ \boldsymbol{r}_3^{\mathrm{T}} \end{bmatrix}\begin{bmatrix} \boldsymbol{r}_1 & \boldsymbol{r}_2 & \boldsymbol{r}_3 \end{bmatrix}=\begin{bmatrix} \boldsymbol{r}_1^{\mathrm{T}}\boldsymbol{r}_1 & \boldsymbol{r}_1^{\mathrm{T}}\boldsymbol{r}_2 & \boldsymbol{r}_1^{\mathrm{T}}\boldsymbol{r}_3 \\ \boldsymbol{r}_2^{\mathrm{T}}\boldsymbol{r}_1 & \boldsymbol{r}_2^{\mathrm{T}}\boldsymbol{r}_2 & \boldsymbol{r}_2^{\mathrm{T}}\boldsymbol{r}_3 \\ \boldsymbol{r}_3^{\mathrm{T}}\boldsymbol{r}_1 & \boldsymbol{r}_3^{\mathrm{T}}\boldsymbol{r}_2 & \boldsymbol{r}_3^{\mathrm{T}}\boldsymbol{r}_3 \end{bmatrix}=\begin{bmatrix} \boldsymbol{r}_1\cdot\boldsymbol{r}_1 & \boldsymbol{r}_1\cdot\boldsymbol{r}_2 & \boldsymbol{r}_1\cdot\boldsymbol{r}_3 \\ \boldsymbol{r}_2\cdot\boldsymbol{r}_1 & \boldsymbol{r}_2\cdot\boldsymbol{r}_2 & \boldsymbol{r}_2\cdot\boldsymbol{r}_3 \\ \boldsymbol{r}_3\cdot\boldsymbol{r}_1 & \boldsymbol{r}_3\cdot\boldsymbol{r}_2 & \boldsymbol{r}_3\cdot\boldsymbol{r}_3 \end{bmatrix} \tag{32}$$

观察 (32) 对角线，可以容易判断 $\boldsymbol{r}_1$、$\boldsymbol{r}_2$、$\boldsymbol{r}_3$ 均为单位向量，但是 $[\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3]$ 为非正交基。

而 $\boldsymbol{P}$ 中非对角线元素 $\cos\theta_{i,j}$ 就是 $\boldsymbol{r}_i$ 和 $\boldsymbol{r}_j$ 向量夹角的余弦值。下面验证一下。

计算向量 $\boldsymbol{r}_1$ 和 $\boldsymbol{r}_2$ 角度的余弦值:

$$\frac{\boldsymbol{r}_1\cdot\boldsymbol{r}_2}{\left\|\boldsymbol{r}_1\right\|\left\|\boldsymbol{r}_2\right\|}=\cos\theta_{1,2} \tag{33}$$

$\boldsymbol{r}_1$ 和 $\boldsymbol{r}_3$ 角度的余弦值为:

$$\frac{\boldsymbol{r}_1\cdot\boldsymbol{r}_3}{\left\|\boldsymbol{r}_1\right\|\left\|\boldsymbol{r}_3\right\|}=\cos\theta_{1,3} \tag{34}$$

$\boldsymbol{r}_2$ 和 $\boldsymbol{r}_3$ 角度的余弦值为:

$$\frac{\boldsymbol{r}_1\cdot\boldsymbol{r}_3}{\left\|\boldsymbol{r}_1\right\|\left\|\boldsymbol{r}_3\right\|}=\cos\theta_{2,3} \tag{35}$$

## 几何视角

如图 8 所示，利用 $\boldsymbol{R}$，我们完成了标准正交基 $[\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3]$ 向非正交基 $[\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3]$ 的转换。

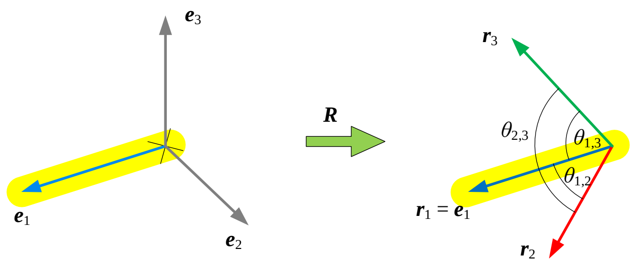



图 8. 三维系转化成满足指定两两夹角的坐标系

换个角度，(28) 中矩阵 $\boldsymbol{P}$ 指定了目标向量两两"相对夹角" $\cos\theta_{1,2}$、$\cos\theta_{1,3}$、$\cos\theta_{2,3}$。即 $\boldsymbol{r}_1$ 和 $\boldsymbol{r}_2$ 的相对夹角为 $\cos\theta_{1,2}$，$\boldsymbol{r}_1$ 和 $\boldsymbol{r}_3$ 的相对夹角为 $\cos\theta_{1,3}$，$\boldsymbol{r}_2$ 和 $\boldsymbol{r}_3$ 的相对夹角为 $\cos\theta_{2,3}$。我们想要找到空间中满足这个条件的三个单位向量。

对 $\boldsymbol{P}$ 进行 Cholesky 分解得到矩阵 $\boldsymbol{R}$，它的列向量 $\boldsymbol{r}_1$、$\boldsymbol{r}_2$、$\boldsymbol{r}_3$ 就是我们想要找的三个向量的空间坐标点。特别地，$\boldsymbol{r}_1$ 和 $\boldsymbol{e}_1$ 相同。好就好比，在构造 $[\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3]$ 这个非正交基时，$\boldsymbol{r}_1$ 锚定在 $\boldsymbol{e}_1$。

⚠ 再次强调一下，$\cos\theta_{1,2}$、$\cos\theta_{1,3}$、$\cos\theta_{2,3}$ 确定的角度是向量之间的"相对夹角"。而 $[\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3]$ 两两列向量确定的角度则是参考标准正交基的"绝对夹角"，这是因为 $\boldsymbol{r}_1 = \boldsymbol{e}_1$。

## 两个例子

图 9 给出两个例子，在给定 $\cos\theta_{1,2}$、$\cos\theta_{1,3}$、$\cos\theta_{2,3}$ 三个角度条件下，我们可以利用 Cholesky 分解矩阵 $\boldsymbol{P}$ 计算得到满足夹角条件的三个单位向量 $\boldsymbol{r}_1$、$\boldsymbol{r}_2$、$\boldsymbol{r}_3$。

(a) $\theta_{1,2} = 60°, \theta_{1,3} = 90°, \theta_{2,3} = 120°$　　(b) $\theta_{1,2} = 135°, \theta_{1,3} = 60°, \theta_{2,3} = 120°$

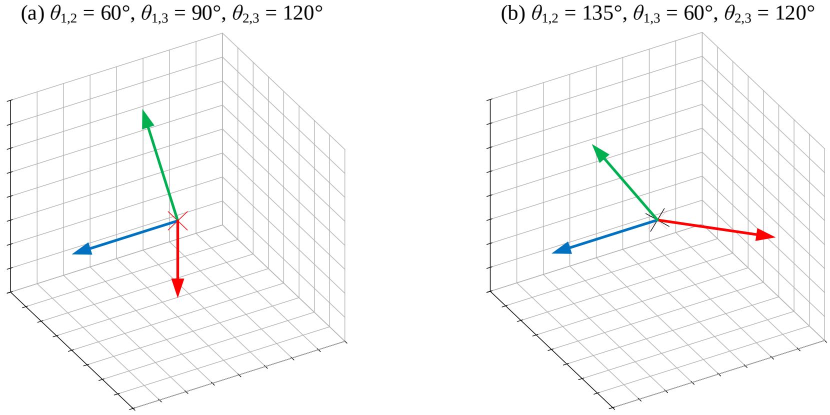


图 9. 给定三个夹角，确定向量三维空间位置

## 前提条件

在图 8 中，任意两个夹角之和必须大于等于第三个夹角，且任意角度不能为 0°，也就是必须满足如下三个不等式：

$$\begin{aligned}\theta_{1,2}+\theta_{1,3}\geq\theta_{2,3}>0^\circ\\ \theta_{1,2}+\theta_{2,3}\geq\theta_{1,3}>0^\circ\\ \theta_{1,3}+\theta_{2,3}\geq\theta_{1,2}>0^\circ\end{aligned} \tag{36}$$

另外，三个角度夹角必须小于等于 360°：

$$\theta_{1,2}+\theta_{1,3}+\theta_{2,3}\leq 360^\circ \tag{37}$$

试想一个有趣的现象，在图 8 中，如果 $\theta_{1,2} = \theta_{1,3} + \theta_{2,3}$，这意味着 $\boldsymbol{r}_1$、$\boldsymbol{r}_2$、$\boldsymbol{r}_3$ 三个向量在一个平面上，$\boldsymbol{r}_1$、$\boldsymbol{r}_2$、$\boldsymbol{r}_3$ 线性相关。这种情况，矩阵 $\boldsymbol{R}$ 不满秩，也就是说 $\boldsymbol{P}$ 也不满秩。正定矩阵满秩，这种情形 $\boldsymbol{P}$ 不可以 Cholesky 分解。

而三个夹角之和等于 360°时，即 $\theta_{1,2} + \theta_{1,3} + \theta_{2,3} = 360^\circ$，$\boldsymbol{r}_1$、$\boldsymbol{r}_2$、$\boldsymbol{r}_3$ 三个向量也在一个平面上，$\boldsymbol{P}$ 也不可以 Cholesky 分解。

最后，如果 $\theta_{1,2}$、$\theta_{1,3}$、$\theta_{2,3}$ 任一角度为 0°，这意味着存在两个向量共线，这种情况 $\boldsymbol{P}$ 也不可以 Cholesky 分解。

也就是为了保证 (28) 中 $\boldsymbol{P}$ 可以 Cholesky 分解，即正定，需要满足以下条件：

$$\begin{aligned}&\theta_{1,2}>0^\circ,\quad \theta_{1,3}>0^\circ,\quad \theta_{2,3}>0^\circ\\ &\theta_{1,2}+\theta_{1,3}>\theta_{2,3},\quad \theta_{1,2}+\theta_{2,3}>\theta_{1,3},\quad \theta_{1,3}+\theta_{2,3}>\theta_{1,2}\\ &\theta_{1,2}+\theta_{1,3}+\theta_{2,3}<360^\circ\end{aligned} \tag{38}$$

## 夹角相同

再看一组特殊情况，(28) 中 $\boldsymbol{P}$ 两两夹角相同，即，

$$\theta_{1,2}=\theta_{1,3}=\theta_{2,3}=\theta \tag{39}$$

此时，$\boldsymbol{P}$ 可以写成：

$$\boldsymbol{P}=\begin{bmatrix}1 & \cos\theta & \cos\theta\\ \cos\theta & 1 & \cos\theta\\ \cos\theta & \cos\theta & 1\end{bmatrix} \tag{40}$$

打个比方，这个例子像是一把雨伞的开合。假设雨伞只有三个伞骨，雨伞开合时，伞骨之间的两两夹角相等。

雨伞合起来时，三个伞骨并拢，相当于三个向量之间夹角为 0°，即共线。

如果雨伞最大开度可以让伞面为平面，这时三个伞骨之间夹角为 120°，也就是三个向量在一个平面上。

有了这两个极限情况，我们知道向量之间夹角 $\theta$ 取值范围为 [0°, 120°]，而 $\cos\theta$ 的取值范围为 [−0.5, 1] (cos(120°) = −0.5, cos(0°) = 1)。这也就是说，这种情况下，$\boldsymbol{P}$ 的两个极端取值为：

$$\boldsymbol{P}=\begin{bmatrix}1 & -0.5 & -0.5\\ -0.5 & 1 & -0.5\\ -0.5 & -0.5 & 1\end{bmatrix},\quad \boldsymbol{P}=\begin{bmatrix}1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\end{bmatrix} \tag{41}$$

上式中两个 $\boldsymbol{P}$ 都不能进行 Cholesky 分解，因为 $\boldsymbol{P}$ 都不满秩。

图 10 给出四个不同开合角度。图 10 (d) 对应的 (41) 第一个矩阵 $\boldsymbol{P}$，$\theta_{1,2}$、$\theta_{1,3}$、$\theta_{2,3}$ 三个角度都是 120°，因此 $\boldsymbol{r}_1$、$\boldsymbol{r}_2$、$\boldsymbol{r}_3$ 在一个平面上，线性相关。

从统计角度来看，$\boldsymbol{P}$ 代表相关性系数矩阵。如果其中任意两个随机变量的相关性系数相等，满足 (40) 相关性系数的取值范围为 [−0.5, 1]。

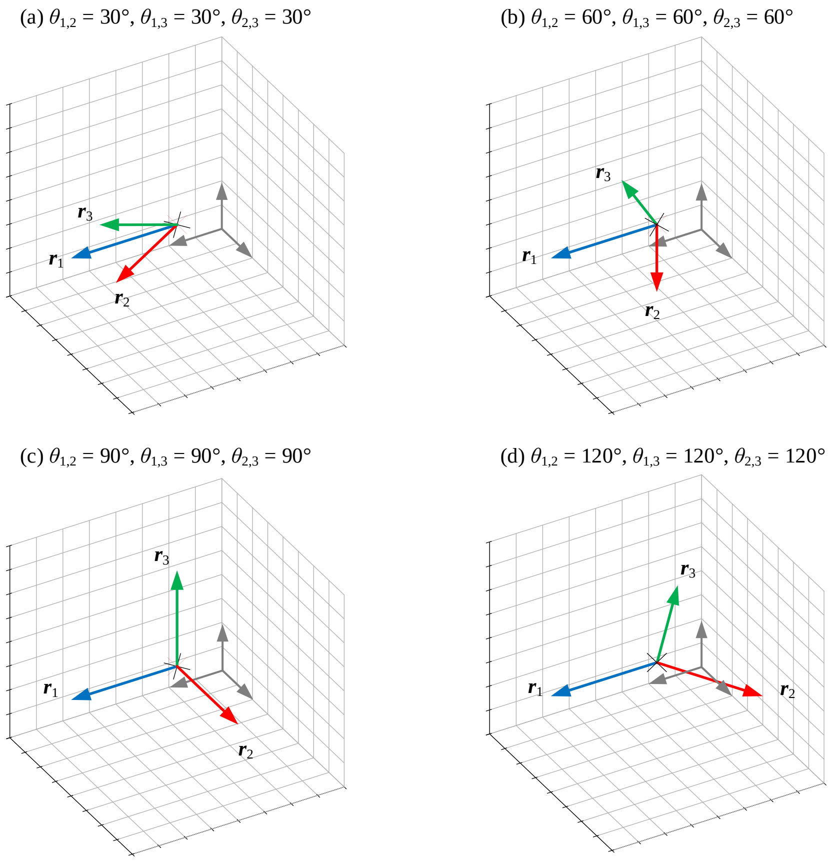



图 10. 相等角度条件下，确定向量三维空间位置

至此，我们利用空间几何视角，探讨了 Cholesky 分解以及满足 Cholesky 分解条件。

Bk4_Ch12_02.py 绘制图 9 和图 10。请读者自行设定夹角条件，看看哪些角度组合能够进行 Cholesky 分解，哪些不能。

# 12.6 从格拉姆矩阵到相似度矩阵

有了本章前文内容铺垫，下面我们回头来看一下格拉姆矩阵。

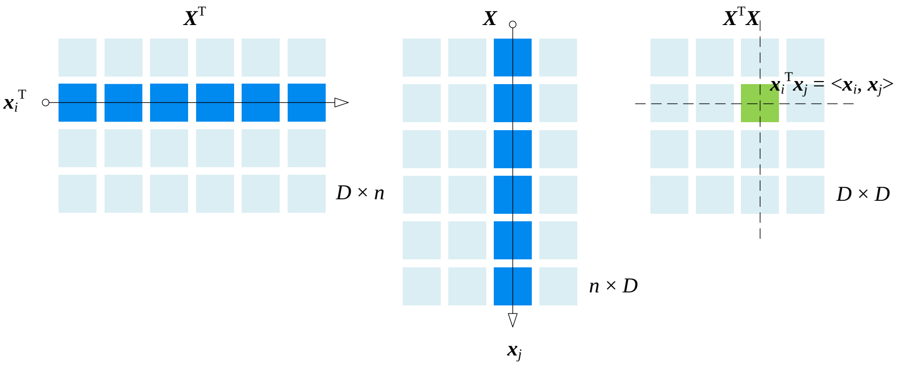



图 11. 格拉姆矩阵

如图 11 所示，数据矩阵 $X$ 的格拉姆矩阵 $G$ 可以写成标量积形式：

$$G = \begin{bmatrix} x_1 \cdot x_1 & x_1 \cdot x_2 & \cdots & x_1 \cdot x_D \\ x_2 \cdot x_1 & x_2 \cdot x_2 & \cdots & x_2 \cdot x_D \\ \vdots & \vdots & \ddots & \vdots \\ x_D \cdot x_1 & x_D \cdot x_2 & \cdots & x_D \cdot x_D \end{bmatrix} = \begin{bmatrix} \langle x_1, x_1 \rangle & \langle x_1, x_2 \rangle & \cdots & \langle x_1, x_D \rangle \\ \langle x_2, x_1 \rangle & \langle x_2, x_2 \rangle & \cdots & \langle x_2, x_D \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle x_D, x_1 \rangle & \langle x_D, x_2 \rangle & \cdots & \langle x_D, x_D \rangle \end{bmatrix} \tag{42}$$

## 确定列向量坐标

对 $G$ 进行 Cholesky 分解得到：

$$G = R_G^T R_G \tag{43}$$

将 $R_G$ 写成一排列向量：

$$R_G = \begin{bmatrix} r_{G,1} & r_{G,2} & \cdots & r_{G,D} \end{bmatrix} \tag{44}$$

将 (44) 代入 (43) 得到：

$$
\boldsymbol{G}=\begin{bmatrix}\boldsymbol{r}_{G,1}{}^{\mathrm{T}}\\ \boldsymbol{r}_{G,2}{}^{\mathrm{T}}\\ \vdots\\ \boldsymbol{r}_{G,D}{}^{\mathrm{T}}\end{bmatrix}\begin{bmatrix}\boldsymbol{r}_{G,1} & \boldsymbol{r}_{G,2} & \cdots & \boldsymbol{r}_{G,D}\end{bmatrix}=\begin{bmatrix}\langle\boldsymbol{r}_{G,1},\boldsymbol{r}_{G,1}\rangle & \langle\boldsymbol{r}_{G,1},\boldsymbol{r}_{G,2}\rangle & \cdots & \langle\boldsymbol{r}_{G,1},\boldsymbol{r}_{G,D}\rangle\\ \langle\boldsymbol{r}_{G,2},\boldsymbol{r}_{G,1}\rangle & \langle\boldsymbol{r}_{G,2},\boldsymbol{r}_{G,2}\rangle & \cdots & \langle\boldsymbol{r}_{G,1},\boldsymbol{r}_{G,D}\rangle\\ \vdots & \vdots & \ddots & \vdots\\ \langle\boldsymbol{r}_{G,D},\boldsymbol{r}_{G,1}\rangle & \langle\boldsymbol{r}_{G,D},\boldsymbol{r}_{G,2}\rangle & \cdots & \langle\boldsymbol{r}_{G,D},\boldsymbol{r}_{G,D}\rangle\end{bmatrix} \tag{45}
$$

(42) 等价于 (45)，向量模和向量夹角之间完全等价。这相当于我们在 $\mathbb{R}^D$ 中找到了 $\boldsymbol{X}$ 每个列向量的具体坐标！

以鸢尾花数据矩阵 $\boldsymbol{X}$ 为例，$\boldsymbol{X}$ 可以写成四个列向量左右排列，即 $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4]$。这些列向量都有 150 个元素，显然不能直接在 $\mathbb{R}^4$ 空间中展示。

图 12 所示为计算 $\boldsymbol{X}$ 的 Gram 矩阵 $\boldsymbol{G}$ 过程热图。如前文所述，矩阵 $\boldsymbol{G}$ 中包含了 $[\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4]$ 各个列向量的模，以及它们之间两两夹角余弦值。

一个向量就两个元素——大小和方向，$\boldsymbol{G}$ 这相当于集成了 $[\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4]$ 每个向量关键信息。

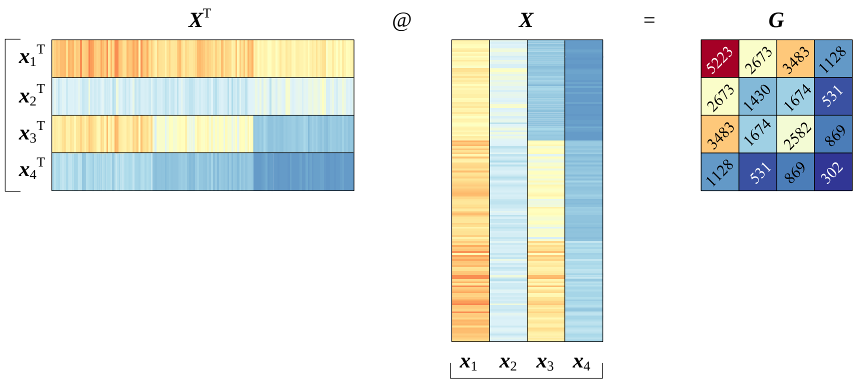



图 12. 鸢尾花数据矩阵 $\boldsymbol{X}$ 格拉姆矩阵，图片来自本书第 10 章

如图 13 所示，对 Gram 矩阵 $\boldsymbol{G}$ 进行 Cholesky 分解得到上三角矩阵 $\boldsymbol{R}_G$，$\boldsymbol{R}_G$ 的列向量长度为 4，它们在在 $\mathbb{R}^4$ 空间中，等价于 $[\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4]$。

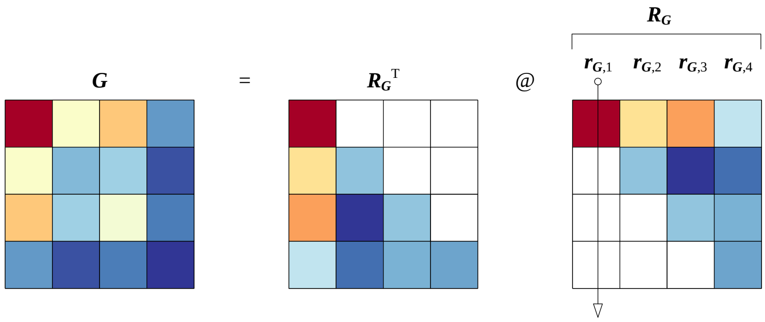



图 13. 对格拉姆矩阵 $\boldsymbol{G}$ 进行 Cholesky 分解

### 向量夹角

以向量夹角余弦形式展开 $\boldsymbol{G}$ 中向量积：

$$\boldsymbol{G}=\begin{bmatrix} \|\boldsymbol{x}_1\|\|\boldsymbol{x}_1\|\cos\theta_{1,1} & \|\boldsymbol{x}_1\|\|\boldsymbol{x}_2\|\cos\theta_{2,1} & \cdots & \|\boldsymbol{x}_1\|\|\boldsymbol{x}_D\|\cos\theta_{1,D} \\ \|\boldsymbol{x}_2\|\|\boldsymbol{x}_1\|\cos\theta_{1,2} & \|\boldsymbol{x}_2\|\|\boldsymbol{x}_2\|\cos\theta_{2,2} & \cdots & \|\boldsymbol{x}_2\|\|\boldsymbol{x}_D\|\cos\theta_{2,D} \\ \vdots & \vdots & \ddots & \vdots \\ \|\boldsymbol{x}_D\|\|\boldsymbol{x}_1\|\cos\theta_{1,D} & \|\boldsymbol{x}_D\|\|\boldsymbol{x}_2\|\cos\theta_{2,D} & \cdots & \|\boldsymbol{x}_D\|\|\boldsymbol{x}_D\|\cos\theta_{D,D} \end{bmatrix} \tag{46}$$

观察矩阵 $\boldsymbol{G}$，它包含了数据矩阵 $\boldsymbol{X}$ 中列向量的两个重要信息——模 $\|\boldsymbol{x}_i\|$、方向 (向量两两相对夹角余弦值 $\cos\theta_{i,j}$)。

定义缩放矩阵 $\boldsymbol{S}$，具体形式如下：

$$\boldsymbol{S}=\begin{bmatrix} \|\boldsymbol{x}_1\| & & & \\ & \|\boldsymbol{x}_2\| & & \\ & & \ddots & \\ & & & \|\boldsymbol{x}_D\| \end{bmatrix} \tag{47}$$

对 $\boldsymbol{G}$ 左右分别乘上 $\boldsymbol{S}$ 的逆，得到 $\boldsymbol{C}$：

$$\boldsymbol{C}=\boldsymbol{S}^{-1}\boldsymbol{G}\boldsymbol{S}^{-1}=\begin{bmatrix} \frac{\boldsymbol{x}_1\cdot\boldsymbol{x}_1}{\|\boldsymbol{x}_1\|\|\boldsymbol{x}_1\|} & \frac{\boldsymbol{x}_1\cdot\boldsymbol{x}_2}{\|\boldsymbol{x}_1\|\|\boldsymbol{x}_2\|} & \cdots & \frac{\boldsymbol{x}_1\cdot\boldsymbol{x}_D}{\|\boldsymbol{x}_1\|\|\boldsymbol{x}_D\|} \\ \frac{\boldsymbol{x}_2\cdot\boldsymbol{x}_1}{\|\boldsymbol{x}_2\|\|\boldsymbol{x}_1\|} & \frac{\boldsymbol{x}_2\cdot\boldsymbol{x}_2}{\|\boldsymbol{x}_2\|\|\boldsymbol{x}_2\|} & \cdots & \frac{\boldsymbol{x}_2\cdot\boldsymbol{x}_D}{\|\boldsymbol{x}_2\|\|\boldsymbol{x}_D\|} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\boldsymbol{x}_D\cdot\boldsymbol{x}_1}{\|\boldsymbol{x}_D\|\|\boldsymbol{x}_1\|} & \frac{\boldsymbol{x}_D\cdot\boldsymbol{x}_2}{\|\boldsymbol{x}_D\|\|\boldsymbol{x}_2\|} & \cdots & \frac{\boldsymbol{x}_D\cdot\boldsymbol{x}_D}{\|\boldsymbol{x}_D\|\|\boldsymbol{x}_D\|} \end{bmatrix} \tag{48}$$

矩阵 $\boldsymbol{C}$ 中元素就是向量两两夹角余弦值。

### 余弦相似度矩阵

矩阵 $\boldsymbol{C}$ 有自己的名字——**余弦相似度矩阵** (cosine similarity matrix)。这是因为 $\boldsymbol{C}$ 的每个元素实际上计算的是 $\boldsymbol{x}_i$ 和 $\boldsymbol{x}_j$ 向量的相对夹角 $\theta_{i,j}$ 余弦值 $\cos\theta_{i,j}$，即，

$$\boldsymbol{C}=\begin{bmatrix} 1 & \cos\theta_{2,1} & \cdots & \cos\theta_{1,D} \\ \cos\theta_{1,2} & 1 & \cdots & \cos\theta_{2,D} \\ \vdots & \vdots & \ddots & \vdots \\ \cos\theta_{1,D} & \cos\theta_{2,D} & \cdots & 1 \end{bmatrix} \tag{49}$$

相比格拉姆矩阵 $\boldsymbol{G}$，余弦相似度矩阵 $\boldsymbol{C}$ 中只包含了 $\boldsymbol{X}$ 列向量两两夹角 $\cos\theta_{i,j}$ 这个单一信息。

对 $\boldsymbol{C}$ 进行 Cholesky 分解得到：

$$\boldsymbol{C}=\boldsymbol{L}\boldsymbol{L}^{\mathrm{T}}=\boldsymbol{R}^{\mathrm{T}}\boldsymbol{R} \tag{50}$$

将 $\boldsymbol{R}$ 写成 $[\boldsymbol{r}_1, \boldsymbol{r}_1, \ldots, \boldsymbol{r}_D]$，$\boldsymbol{C}$ 可以写成：

$$\boldsymbol{C} = \boldsymbol{R}^{\mathrm{T}}\boldsymbol{R} = \begin{bmatrix} \boldsymbol{r}_1^{\mathrm{T}} \\ \boldsymbol{r}_2^{\mathrm{T}} \\ \vdots \\ \boldsymbol{r}_D^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} \boldsymbol{r}_1 & \boldsymbol{r}_2 & \cdots & \boldsymbol{r}_D \end{bmatrix} = \begin{bmatrix} \boldsymbol{r}_1^{\mathrm{T}}\boldsymbol{r}_1 & \boldsymbol{r}_1^{\mathrm{T}}\boldsymbol{r}_2 & \cdots & \boldsymbol{r}_1^{\mathrm{T}}\boldsymbol{r}_D \\ \boldsymbol{r}_2^{\mathrm{T}}\boldsymbol{r}_1 & \boldsymbol{r}_2^{\mathrm{T}}\boldsymbol{r}_2 & \cdots & \boldsymbol{r}_2^{\mathrm{T}}\boldsymbol{r}_D \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{r}_D^{\mathrm{T}}\boldsymbol{r}_1 & \boldsymbol{r}_D^{\mathrm{T}}\boldsymbol{r}_2 & \cdots & \boldsymbol{r}_D^{\mathrm{T}}\boldsymbol{r}_D \end{bmatrix} = \begin{bmatrix} 1 & \cos\theta_{2,1} & \cdots & \cos\theta_{1,D} \\ \cos\theta_{1,2} & 1 & \cdots & \cos\theta_{2,D} \\ \vdots & \vdots & \ddots & \vdots \\ \cos\theta_{1,D} & \cos\theta_{2,D} & \cdots & 1 \end{bmatrix} \tag{51}$$

根据本章前文分析，我们知道 $\boldsymbol{r}_1, \boldsymbol{r}_1, \ldots, \boldsymbol{r}_D$ 都是单位向量。$\boldsymbol{r}_1, \boldsymbol{r}_1, \ldots, \boldsymbol{r}_D$ 两两夹角余弦值满足相似度矩阵 $\boldsymbol{C}$ 的要求。

图 14 所示为鸢尾花数据矩阵的格拉姆矩阵 $\boldsymbol{G}$，先转化成相似度矩阵 $\boldsymbol{C}$，再转化成角度矩阵。角度越小说明特征越相似。

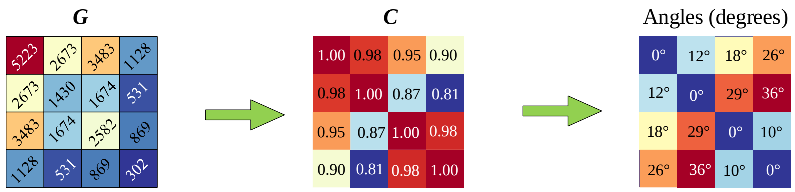



图 14. 格拉姆矩阵 $\boldsymbol{G}$ 转化成相似度矩阵 $\boldsymbol{C}$，再转化成角度

本节介绍的内容在**蒙特卡洛模拟** (Monte Carlo simulation) 中有重要应用。如图 15，本章介绍的 Cholesky 分解结果可以用来产生满足指定相关性系数的随机数。

本系列丛书《概率统计》和《数据科学》两本会从理论、应用两个角度讲解蒙特卡洛模拟。

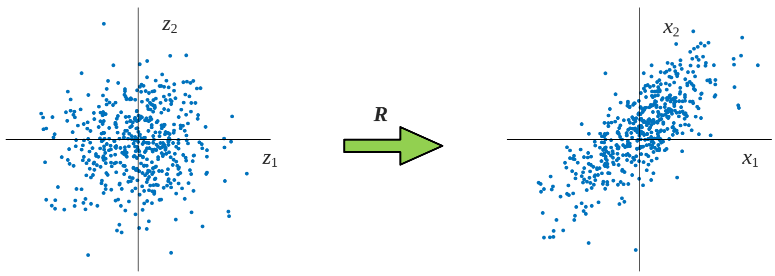



图 15. 产生满足指定相关性矩阵要求的随机数

本章从几何视角讲解了 Cholesky 分解。只有正定矩阵才可以进行 Cholesky 分解，这一点可以用来判断矩阵是否为正定。我们创造了“开合”这个词用来描述 Cholesky 分解得到的上三角矩阵对应的几何变换。

对 Gram 矩阵进行 Cholesky 分解可以帮我们确定原数据矩阵的列向量空间等价坐标。此外，我们将在本系列丛书《概率统计》中有关协方差矩阵和蒙特卡罗模拟中再聊到 Cholesky 分解。

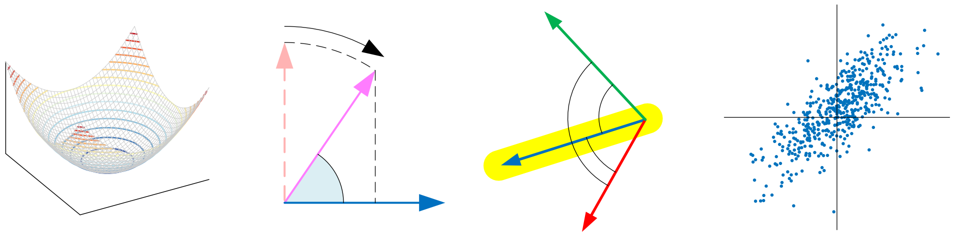


图 16. 总结本章重要内容的四副图