

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

## 13.2 Cholesky 分解



### 本节你将掌握的核心技能：

- ▶ Cholesky 分解：只能用于正定矩阵。
- ▶ 几何角度看 Cholesky 分解：“缩放 → 剪切 → 剪切 → 缩放”。
- ▶ 换个几何角度看 Cholesky 分解：“剪切 → 缩放 → 缩放 → 剪切”。
- ▶ LDL 分解：适用于对称矩阵。
- ▶ 几何角度看 LDL 分解：“剪切 → 缩放 → 剪切”。
- ▶ 两种方法产生具有特定相关性的随机数。

本节介绍两种特殊的矩阵分解 Cholesky 分解、LDL 分解。Cholesky 分解仅适用于正定矩阵，LDL 分解适用于对称矩阵。

### Cholesky 分解

如图 1 所示，Cholesky 分解 (Cholesky decomposition) 把矩阵  $A$  分解为

$$A = R^T R \quad (1)$$

其中， $R$  为上三角矩阵， $R^T$  为下三角矩阵。

⚠ 注意，只有正定矩阵才能 Cholesky 分解。

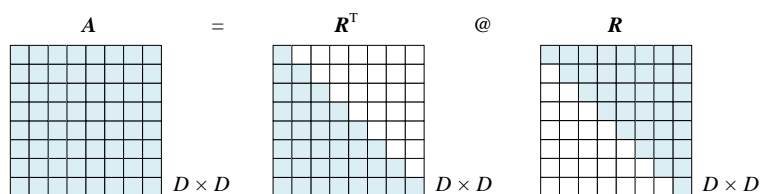


图 1. Cholesky 分解正定矩阵

## 几何视角看 Cholesky 分解

举个例子，给定如下矩阵  $A$

$$A = \begin{bmatrix} 5/4 & -3/4 \\ -3/4 & 5/4 \end{bmatrix} \quad (2)$$

本书之前大家已经对这个矩阵进行特征值分解，大家知道  $A$  的特征值  $\lambda_1 = 2$ 、 $\lambda_2 = 1/2$ 。方阵  $A$  是正定矩阵，可以进行 Cholesky 分解。

对  $A$  进行 Cholesky 分解得到

$$A = \begin{bmatrix} 5/4 & -3/4 \\ -3/4 & 5/4 \end{bmatrix} = \underbrace{\begin{bmatrix} \sqrt{5}/2 & 0 \\ -3\sqrt{5}/10 & 2\sqrt{5}/5 \end{bmatrix}}_{R^T} @ \underbrace{\begin{bmatrix} \sqrt{5}/2 & -3\sqrt{5}/10 \\ 0 & 2\sqrt{5}/5 \end{bmatrix}}_R \quad (3)$$

如图 2 所示， $Ax$  可以写成  $R^T R x$ ，也就是  $R$  先对  $x$  进行线性变换，然后  $R^T$  再作用。

大家可能已经发现  $R$  对应的几何操作中含有剪切、缩放两个成分，下面让我们展开分析。

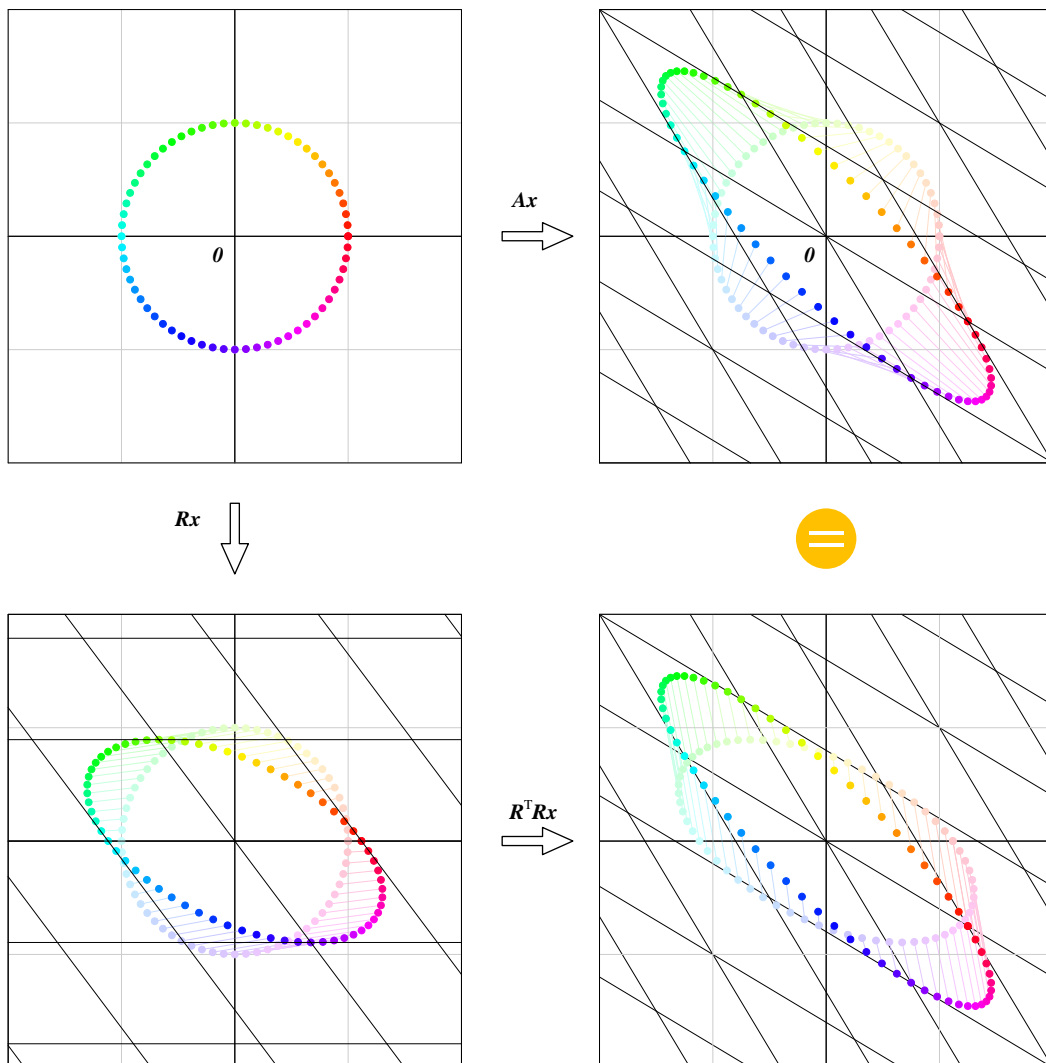


图 2. 矩阵  $A$  的 Cholesky 分解对应的线性变换

缩放  $\rightarrow$  剪切  $\rightarrow$  剪切  $\rightarrow$  缩放

(3) 中  $R^T$  可以写成

$$R^T = \begin{bmatrix} \sqrt{5}/2 & 0 \\ -3\sqrt{5}/10 & 2\sqrt{5}/5 \end{bmatrix} = \underbrace{\begin{bmatrix} \sqrt{5}/2 & 0 \\ 0 & 2\sqrt{5}/5 \end{bmatrix}}_S \underbrace{\begin{bmatrix} 1 & 0 \\ -3/4 & 1 \end{bmatrix}}_K \quad (4)$$

也就是  $R^T$  对应的几何变换可以理解为“剪切  $\rightarrow$  缩放”。 $K$  沿纵轴剪切。

$R$  可以写成

$$R = \begin{bmatrix} \sqrt{5}/2 & -3\sqrt{5}/10 \\ 0 & 2\sqrt{5}/5 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & -3/4 \\ 0 & 1 \end{bmatrix}}_{K^T} \underbrace{\begin{bmatrix} \sqrt{5}/2 & 0 \\ 0 & 2\sqrt{5}/5 \end{bmatrix}}_S \quad (5)$$

$K^T$  沿横轴剪切。

$R$  可以理解为“缩放  $\rightarrow$  剪切”。

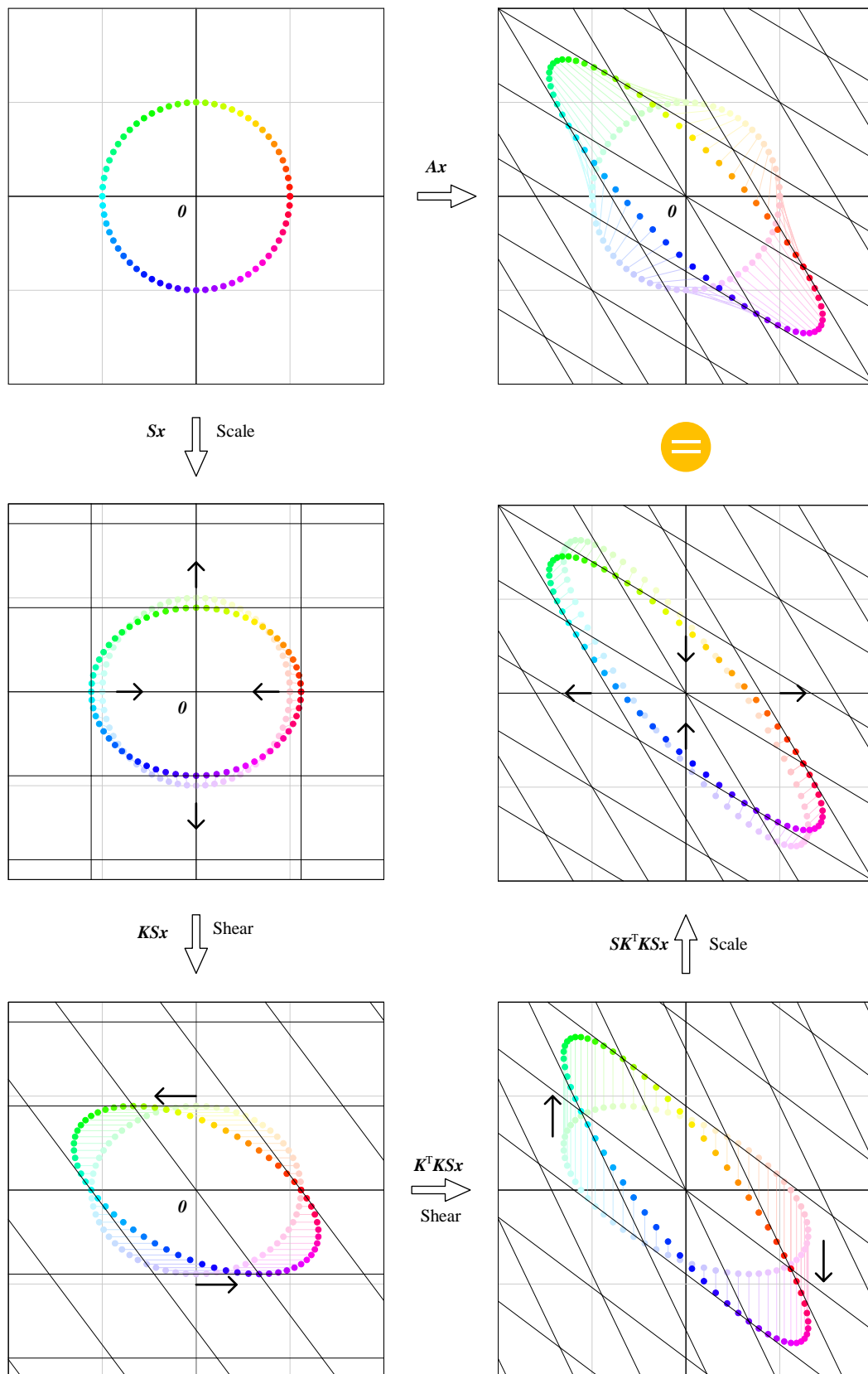
这样  $A$  可以写成

$$A = R^T R = SK(SK)^T = SKK^T S^T = SKK^T S \quad (6)$$

如图 3 所示，这样  $A$  的线性转换可以理解为“缩放 ( $S$ )  $\rightarrow$  剪切 ( $K^T$ )  $\rightarrow$  剪切 ( $K$ )  $\rightarrow$  缩放 ( $S$ )”。



这一点，我们在本书第 8 章第 3 节讲过，请大家回顾。

图 3. 矩阵  $A$  的 Cholesky 分解对应“缩放 → 剪切 → 剪切 → 缩放”

剪切 → 缩放 → 缩放 → 剪切

(3) 中  $\mathbf{R}^T$  还可以写成

$$\mathbf{R}^T = \begin{bmatrix} \sqrt{5}/2 & 0 \\ -3\sqrt{5}/10 & 2\sqrt{5}/5 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 \\ -3/5 & 1 \end{bmatrix}}_{\mathbf{L}} \underbrace{\begin{bmatrix} \sqrt{5}/2 & 0 \\ 0 & 2\sqrt{5}/5 \end{bmatrix}}_{\mathbf{S}} \quad (7)$$

也就是  $\mathbf{R}^T$  对应的几何变换可以理解为“缩放  $\rightarrow$  剪切”。 $\mathbf{L}$  沿纵轴剪切。

$\mathbf{R}$  可以写成

$$\mathbf{R} = \begin{bmatrix} \sqrt{5}/2 & -3\sqrt{5}/10 \\ 0 & 2\sqrt{5}/5 \end{bmatrix} = \underbrace{\begin{bmatrix} \sqrt{5}/2 & 0 \\ 0 & 2\sqrt{5}/5 \end{bmatrix}}_{\mathbf{S}} \underbrace{\begin{bmatrix} 1 & -3/5 \\ 0 & 1 \end{bmatrix}}_{\mathbf{L}^T} \quad (8)$$

$\mathbf{L}^T$  沿横轴剪切。

这样  $\mathbf{A}$  可以写成

$$\mathbf{A} = \mathbf{R}^T \mathbf{R} = \mathbf{L} \mathbf{S} (\mathbf{L} \mathbf{S})^T = \mathbf{L} \mathbf{S} \mathbf{S}^T \mathbf{L}^T = \mathbf{L} \mathbf{S} \mathbf{S} \mathbf{L}^T \quad (9)$$

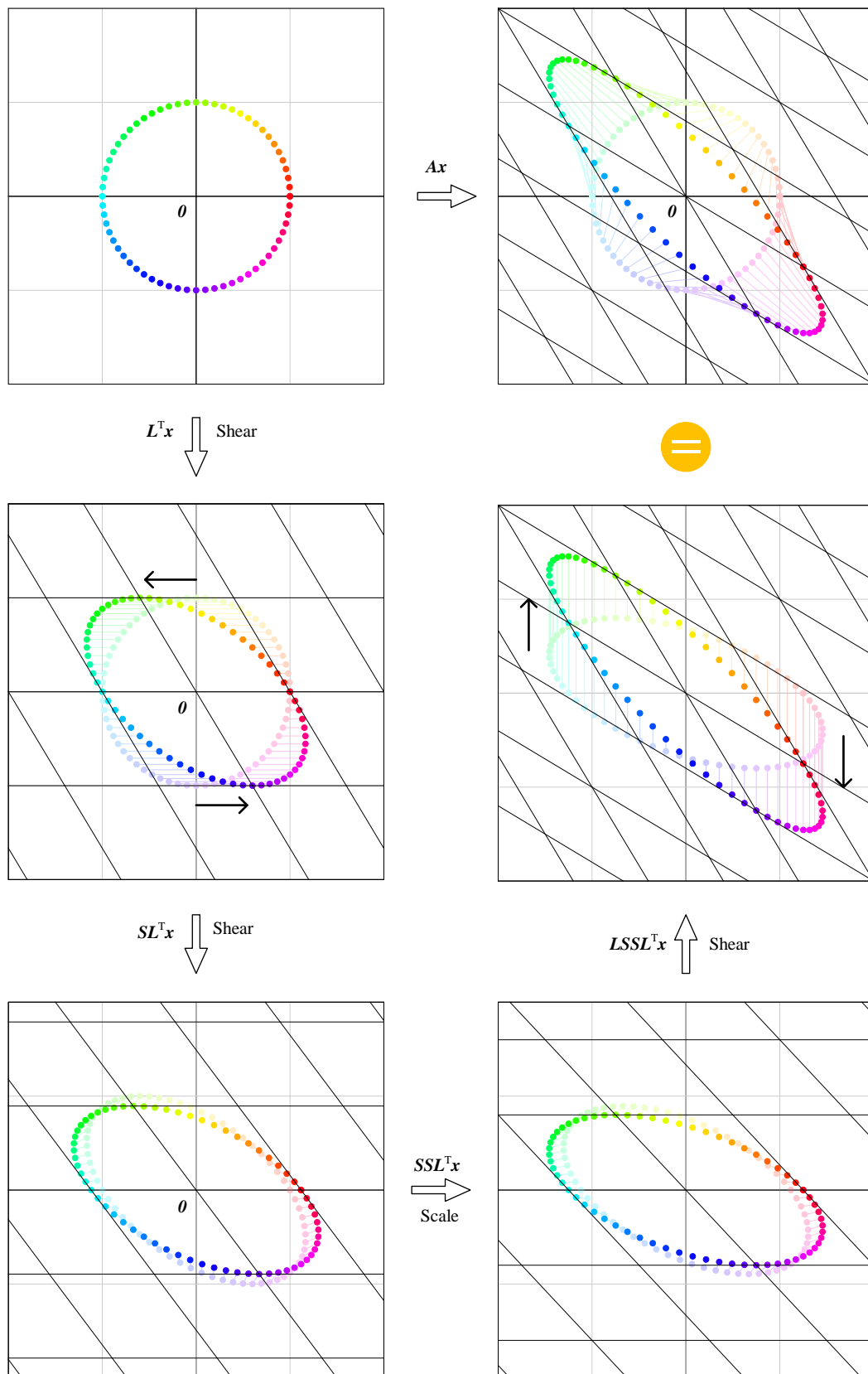
如图 4 所示， $\mathbf{A}$  的线性转换可以理解为“剪切 ( $\mathbf{L}^T$ )  $\rightarrow$  缩放 ( $\mathbf{S}$ )  $\rightarrow$  缩放 ( $\mathbf{S}$ )  $\rightarrow$  剪切 ( $\mathbf{L}$ )”。

把 (9) 中两个  $\mathbf{S}$  合并得到

$$\mathbf{A} = \mathbf{L} \mathbf{S}^2 \mathbf{L}^T \quad (10)$$

$\mathbf{A}$  的线性转换可以理解为“剪切 ( $\mathbf{L}^T$ )  $\rightarrow$  缩放 ( $\mathbf{S}^2$ )  $\rightarrow$  剪切 ( $\mathbf{L}$ )”。

这实际上引出本节第二个矩阵分解——LDL 分解。

图 4. 矩阵  $A$  的 Cholesky 分解对应“剪切 → 缩放 → 缩放 → 剪切”

## LDL 分解

如图 5 所示，Cholesky 分解可以进一步扩展为 **LDL 分解** (LDL decomposition, LDLT decomposition):

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

版权归清华大学出版社所有，请勿商用，引用请注明出处。

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

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

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

$$A = LDL^T = LD^{1/2}(D^{1/2})^T L^T = LD^{1/2}(LD^{1/2})^T \quad (11)$$

其中， $L$  为下三角矩阵，但是对角线元素均为 1； $D$  为对角矩阵，起到缩放作用。

几何角度来看， $L$  的作用就是“剪切”。也就是说，矩阵  $A$  被分解成“剪切 → 缩放 → 剪切”。

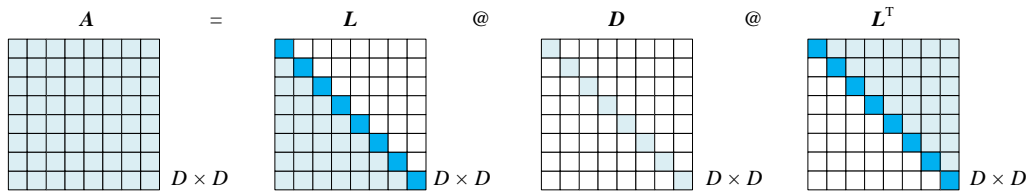


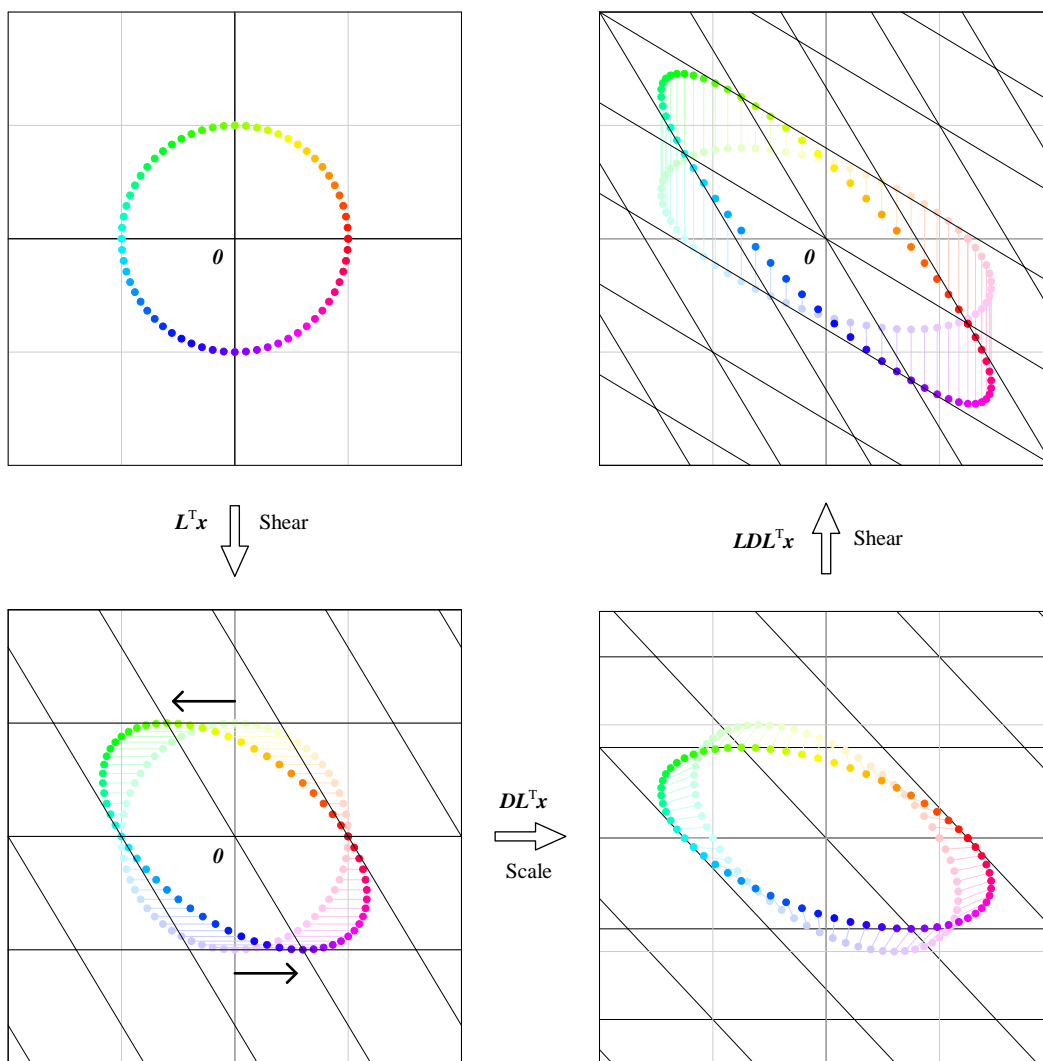
图 5. LDL 分解矩阵运算示意图

对 (2) 中矩阵  $A$  进行 LDL 分解得到，

$$A = \begin{bmatrix} 5/4 & -3/4 \\ -3/4 & 5/4 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 \\ -3/5 & 1 \end{bmatrix}}_L \underbrace{\begin{bmatrix} 5/4 & 0 \\ 0 & 4/5 \end{bmatrix}}_D \underbrace{\begin{bmatrix} 1 & -3/5 \\ 0 & 1 \end{bmatrix}}_{L^T} \quad (12)$$

(12) 对应的几何操作如图 6 所示。

因此，Cholesky 分解可以看作是 LDL 分解的特例，可以对 LDL 分解得到的  $D$  进一步开平方。

图 6. 矩阵  $A$  的 LDL 分解对应“剪切  $\rightarrow$  缩放  $\rightarrow$  剪切”

### 产生具有一定相关性的随机数

图 7 所示为质心位于原点、协方差矩阵为单位矩阵的随机数，把图中散点数据写成数据矩阵  $Z$ 。



注意，数据矩阵  $Z$  的行向量代表图 7 中散点。

数据矩阵  $Z$  质心位于原点意味着

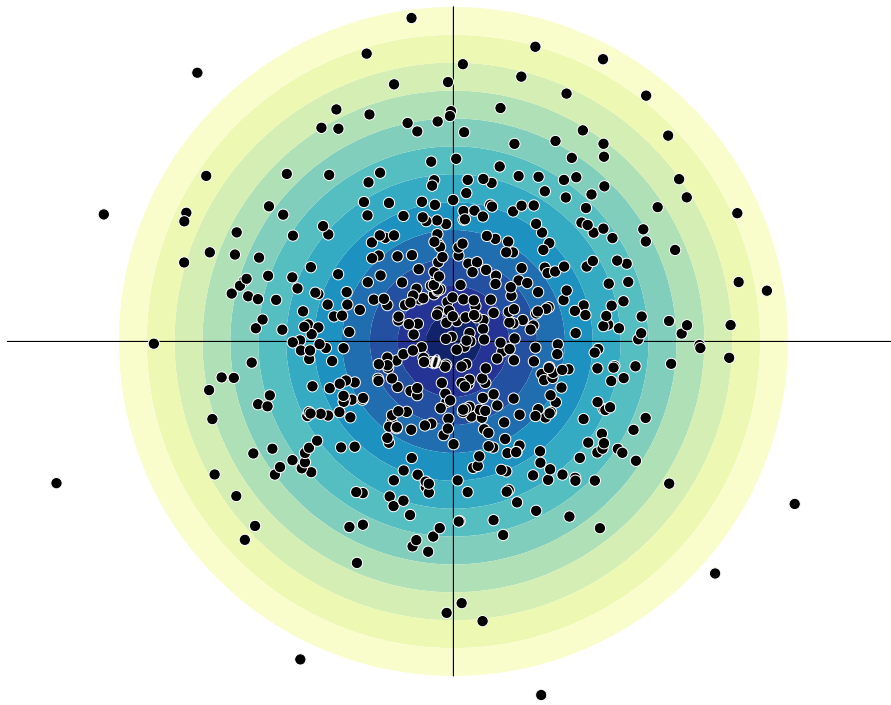
$$\mu_z = \frac{Z^T \mathbf{1}}{n} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (13)$$

其中， $n$  为样本数据数量。

数据矩阵  $Z$  协方差矩阵则为

$$\Sigma_z = \frac{Z^T Z}{n-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I \quad (14)$$

由于数据矩阵  $Z$  的质心已经是零向量，所以上式没有中心化（去均值）。

图 7. 质心位于原点、协方差矩阵为单位矩阵的数据矩阵  $Z$ 

而图 8 所示为质心位于  $[4, 3]^T$ 、协方差矩阵为  $\begin{bmatrix} 1 & 1/2 \\ 1/2 & 1 \end{bmatrix}$  的随机数，把图中散点数据写成数据矩阵  $X$ 。

数据矩阵  $X$  质心位于  $[4, 3]^T$  意味着

$$\mu_X = \frac{X^T \mathbf{1}}{n} = \begin{bmatrix} 4 \\ 3 \end{bmatrix} \quad (15)$$

数据矩阵  $X$  协方差矩阵则为

$$\Sigma_X = \frac{X_c^T X_c}{n-1} = \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1 \end{bmatrix} \quad (16)$$

其中， $X_c$  为  $X$  的中心化，即

$$X_c = X - \mu_X^T \quad (17)$$

几何上来看，上式相当于平移。



数据中心化在本书第 12 章第 1 节讲过，请大家回顾。

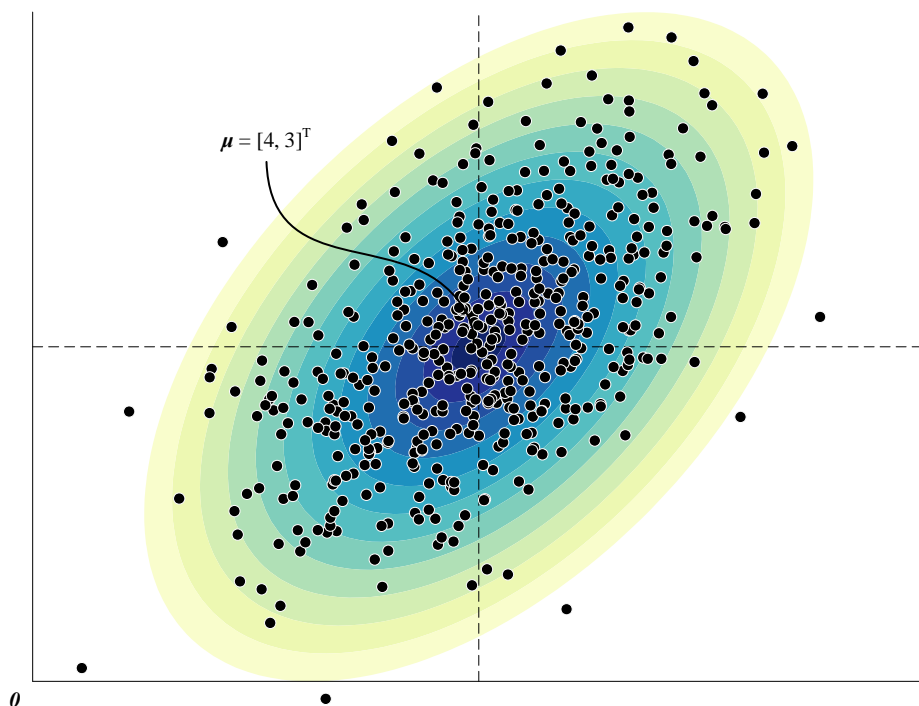


图 8. 质心位于  $[4, 3]^T$ 、协方差矩阵为  $\begin{bmatrix} 1 & 1/2 \\ 1/2 & 1 \end{bmatrix}$  的数据矩阵  $X$

下面介绍如何利用 Cholesky 分解、平移完成  $Z$ 、 $X$  相互转换。

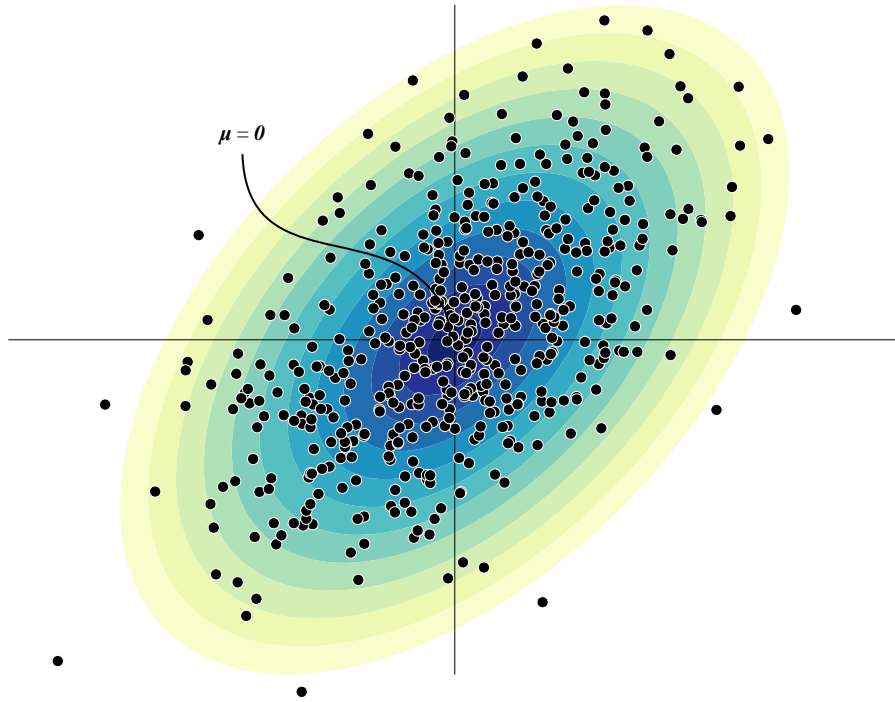
首先对 (16) 中  $\Sigma_X$  进行 Cholesky 分解

$$\Sigma_X = R^T R = \begin{bmatrix} 1 & 0 \\ 1/2 & \sqrt{3}/2 \end{bmatrix} \begin{bmatrix} 1 & 1/2 \\ 0 & \sqrt{3}/2 \end{bmatrix} = \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1 \end{bmatrix} \quad (18)$$

$X$  可以写成

$$X = ZR + \mu_X^T \quad (19)$$

$Z$  (图 7) 经过  $R$  线性变换后得到图 9，数据中心还在原点；然后再用广播原则对数据平移便得到图 8。

图 9. 数据矩阵  $Z @ R$ 

验证  $X$  的协方差矩阵

$$\Sigma_X = \frac{(X - \mu_X^T)^T (X - \mu_X^T)}{n-1} = \frac{R^T Z^T Z R}{n-1} = R^T \frac{Z^T Z}{n-1} R = R^T R \quad (20)$$

**?** 本节前文介绍过， $R$  可以进一步拆解为“剪切 → 缩放”或“缩放 → 剪切”，请大家写成对应矩阵分解。

把 (19) 反过来看，通过如下运算  $X$  还可以变为  $Z$

$$Z = (X - \mu_X^T) R^{-1} \quad (21)$$

注意，上式  $R$  可逆；这意味着， $R$  满秩。也就是说， $\Sigma_X$  正定。

显然，协方差矩阵  $\Sigma_X$  为对称矩阵，我们可以对其进行谱分解。既然 Cholesky 可以作为  $X$ 、 $Z$  的桥梁，谱分解也可以。

对协方差矩阵  $\Sigma_X$  进行谱分解

$$\Sigma_X = V \Lambda V^T = \begin{bmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix} \begin{bmatrix} 3/2 & 0 \\ 0 & 1/2 \end{bmatrix} \begin{bmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix} = \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1 \end{bmatrix} \quad (22)$$

由于特征值均非负，我们可以把上式写成

$$\Sigma_X = \left( V \Lambda^{\frac{1}{2}} \right) \left( V \Lambda^{\frac{1}{2}} \right)^T = \begin{bmatrix} \sqrt{3}/2 & -1/2 \\ \sqrt{3}/2 & 1/2 \end{bmatrix} \begin{bmatrix} \sqrt{3}/2 & \sqrt{3}/2 \\ -1/2 & 1/2 \end{bmatrix} = \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1 \end{bmatrix} \quad (23)$$

也就是说， $\mathbf{X}$  和  $\mathbf{Z}$  的关系也可以写成

$$\mathbf{X} = \mathbf{Z} \left( \mathbf{V} \mathbf{A}^{\frac{1}{2}} \right)^T + \boldsymbol{\mu}_X^T = \mathbf{Z} \mathbf{A}^{\frac{1}{2}} \mathbf{V}^T + \boldsymbol{\mu}_X^T \quad (24)$$

上式的分步几何操作为“缩放 → 旋转 → 平移”。

单独看缩放的话，数据矩阵  $\mathbf{Z}$  (图 7) 经过  $\mathbf{A}^{\frac{1}{2}}$  缩放后得到图 10，然后再旋转，最后平移也可以得到图 8。

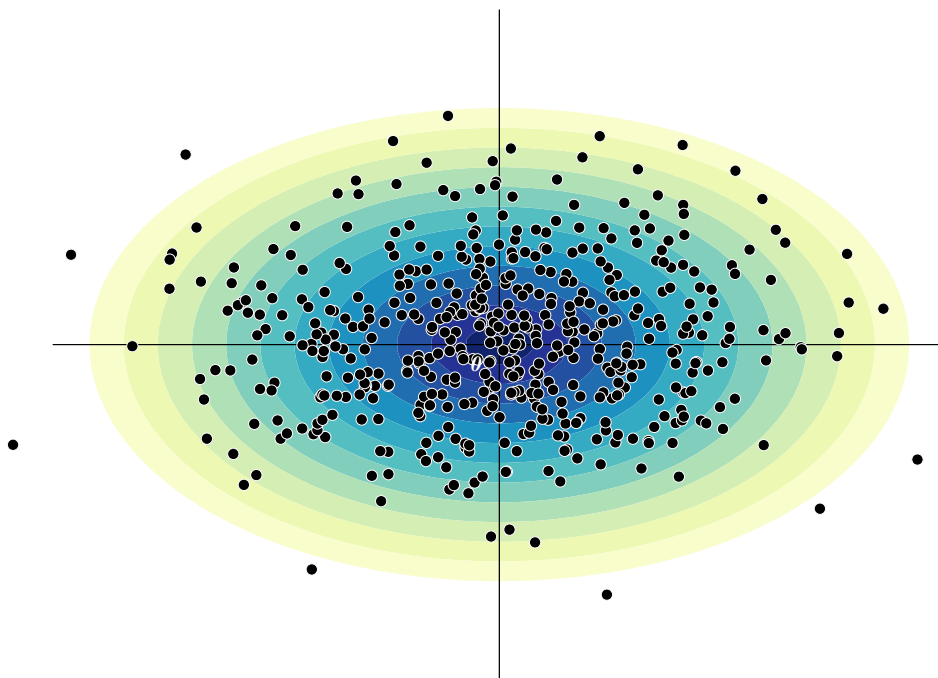


图 10. 数据矩阵  $\mathbf{Z} @ \mathbf{A}^{\frac{1}{2}}$

比较 (18)、(23)，我们发现两者有相同的形式！这两个式子都相当于对  $\Sigma \mathbf{x}$  开平方。



请大家用 DeepSeek/ChatGPT 等工具完成本节如下习题。

**Q1.** 请大家学习使用 `numpy.linalg.cholesky()` 函数，并对 (3) Cholesky 分解

<https://numpy.org/doc/2.2/reference/generated/numpy.linalg.cholesky.html>

**Q2.** 请大家学习使用 `scipy.linalg.ldl()` 函数对 (3) LDL 分解

<https://docs.scipy.org/doc/scipy/reference/generated/scipy.linalg.ldl.html>