

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>

## 5.2 逆矩阵的性质



### 本节你将掌握的核心技能：

- ▶ 建立几何直觉：从旋转、剪切、缩放等角度直观理解各种逆变换的效果。
- ▶ 判断矩阵是否可逆：行列式不为 0 是方阵可逆的充要条件。
- ▶ 不可逆方阵的几何意义：信息丢失，如降维、投影等不可还原操作。
- ▶ 非方阵不可逆原因：输入输出维度不一致，无法唯一反映原始信息。
- ▶ 矩阵连乘的逆：可逆矩阵乘积的逆为各自逆的倒序乘积。
- ▶ 对角方阵可逆时，其逆为对角线元素取倒数构成的新对角矩阵。
- ▶ 对称矩阵的逆仍对称，正交矩阵的逆等于其转置。

有了上一节的几何视角，理解逆矩阵的常用性质就很容易了。

### 不可逆方阵

先说结论，行列式不为 0 是矩阵可逆的充分必要条件。

充分必要条件 (necessary and sufficient condition) 指的是：一个条件的成立既能保证结论成立 (充分性)，且结论成立也能反过来保证该条件成立 (必要性)。

**?** 若方阵的行列式为 0 会发生什么？

给定如下  $2 \times 2$  对称方阵

$$P = \begin{bmatrix} \frac{1}{4} & -\frac{\sqrt{3}}{4} \\ -\frac{\sqrt{3}}{4} & \frac{3}{4} \end{bmatrix} \quad (1)$$

既然一个矩阵是否可逆，取决于它的行列式是否为 0，让我们计算  $\mathbf{P}$  的行列式：

$$\det(\mathbf{P}) = \begin{vmatrix} \frac{1}{4} & -\frac{\sqrt{3}}{4} \\ -\frac{\sqrt{3}}{4} & \frac{3}{4} \end{vmatrix} = 0 \quad (2)$$

显然，这个  $2 \times 2$  对称方阵不可逆。

要理解为什么该矩阵不可逆，我们可以从投影的角度分析它的作用。

如图 1 所示， $2 \times 2$  对称方阵  $\mathbf{P}$  的作用是将平面几何图形投影到过原点的直线上；显然，投影后某些信息丢失，导致无法还原原始图形。

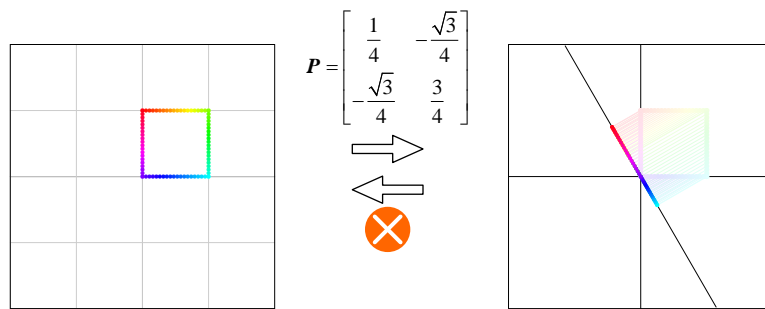


图 1. 平面上的投影不可逆， $2 \times 2$  矩阵，第一例

此外，请大家注意，我们可以把投影矩阵  $\mathbf{P}$  的两个列向量写成

$$\mathbf{p}_1 = \begin{bmatrix} \frac{1}{4} \\ -\frac{\sqrt{3}}{4} \end{bmatrix}, \quad \mathbf{p}_2 = (-\sqrt{3}) \times \begin{bmatrix} \frac{1}{4} \\ -\frac{\sqrt{3}}{4} \end{bmatrix} = -\sqrt{3}\mathbf{p}_1 \quad (3)$$

也就是说， $\mathbf{p}_1$ 、 $\mathbf{p}_2$  共线，线性相关。这就不难解释为什么  $\mathbf{P}$  的行列式为 0 ( $\mathbf{p}_1$ 、 $\mathbf{p}_2$  撑起的平行四边形面积为 0)。

请大家自行分析图 2 不可逆方阵的行列式、列向量是否线性相关。

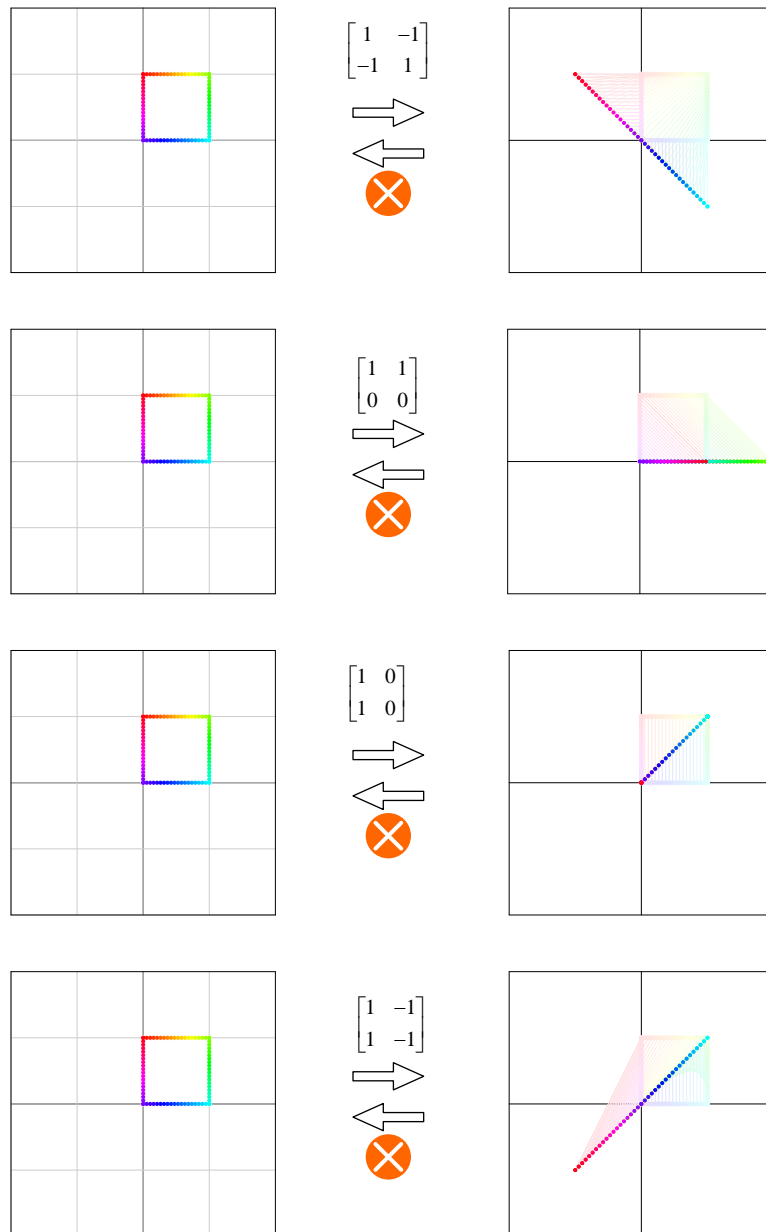


图 2. 平面几何变换不可逆

## 非方阵不可逆

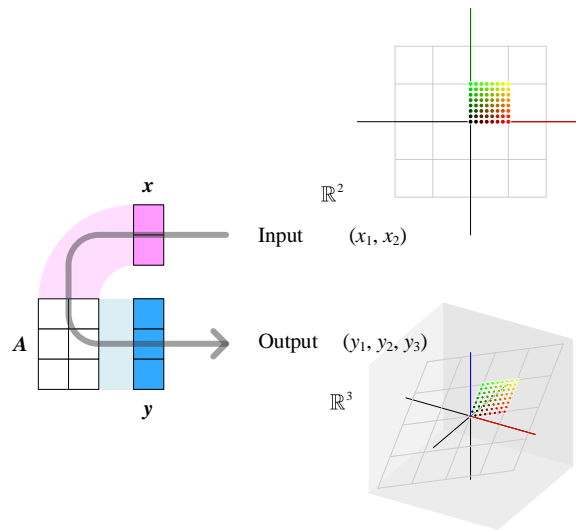
如果一个矩阵不是方阵，即行数  $\neq$  列数，那么它不可逆。非方阵映射的输入维度和输出维度不同，意味着它无法唯一恢复原始输入。

下面，让我们看两个例子。

如图 3 所示， $3 \times 2$  矩阵  $A$  将二维空间映射到三维空间

$$Ax = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \\ a_{31}x_1 + a_{32}x_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \quad (4)$$

这个矩阵将二维空间的所有点映射到三维空间中的某个二维平面，但不会覆盖整个三维空间。

图 3.  $Ax=y$ , 矩阵  $A$  为  $3 \times 2$ 

信息看上去“增加”了(从 2D 到 3D), 意味着无法**唯一**找回原始的 2D 坐标。额外的 1D 维度可以自由变化, 多个不同的 2D 向量可能映射到同一个 3D 向量。这是  $3 \times 2$  矩阵不可逆的几何解释。

想象你在一张纸 (2D) 上画点, 然后把纸放入三维空间。虽然纸上的点都被映射到 3D 空间中, 但它们仍然在同一个平面上, 无法**唯一**地还原成 2D 坐标。

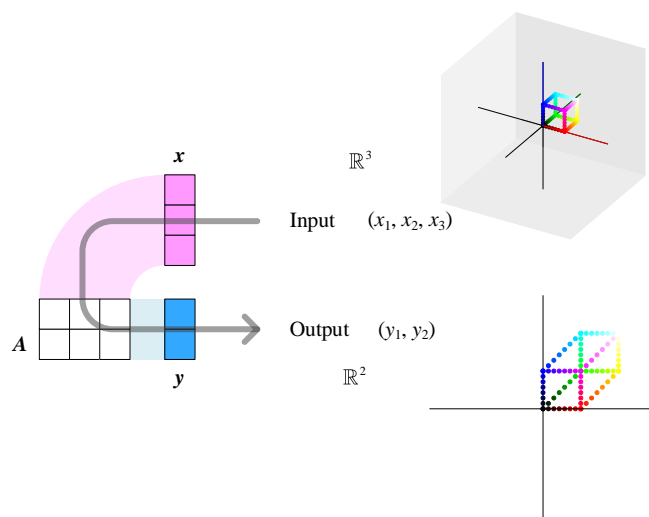
再看一个例子。如图 4 所示,  $2 \times 3$  矩阵  $A$  将三维空间映射到二维空间。

$$Ax = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_{11}x_1 + b_{12}x_2 + b_{13}x_3 \\ b_{21}x_1 + b_{22}x_2 + b_{23}x_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \quad (5)$$

这个不可逆原因是因为映射过程导致信息“丢失”了, 从 3D 到 2D; 不同的 3D 向量可能投影到相同的 2D 向量, 无法**唯一**找回原始的 3D 坐标。

想象你在一个房间 (3D) 里, 看到了物体在墙上的影子 (2D)。虽然你可以看到影子, 但无法仅凭影子**唯一**确定房间里物体的完整三维形状。

非方阵一定不可逆, 因为输入维度  $\neq$  输出维度, 导致信息丢失或增加, 使得逆变换无法**唯一**恢复原始向量。

图 4.  $Ax = y$ , 矩阵  $A$  为  $2 \times 3$ 

⚠ 注意，不可逆方阵、非方阵也有自己的逆——伪逆。伪逆 (pseudoinverse) 是对非方阵或不可逆矩阵的一种广义逆矩阵，能最小化误差并用于求解最小二乘问题。这是本书后续要介绍的内容。

## 两个矩阵乘积的逆

如果形状相同的方阵  $A$ 、 $B$  均可逆，则有

$$(AB)^{-1} = B^{-1}A^{-1} \quad (6)$$

几何角度来看，矩阵乘法  $A @ B @ x$  相当于，对于向量  $x$ ，先用  $B$  完成几何操作，再施加  $A$  对应的几何操作。

用前文的例子， $R$  是旋转矩阵， $S$  是缩放矩阵

$$R = \begin{bmatrix} -1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{bmatrix}, \quad S = \begin{bmatrix} 3/2 & 0 \\ 0 & 2/3 \end{bmatrix} \quad (7)$$

如图 5 所示， $(R @ S) @ x$  代表“一步到位”的复合几何变换。

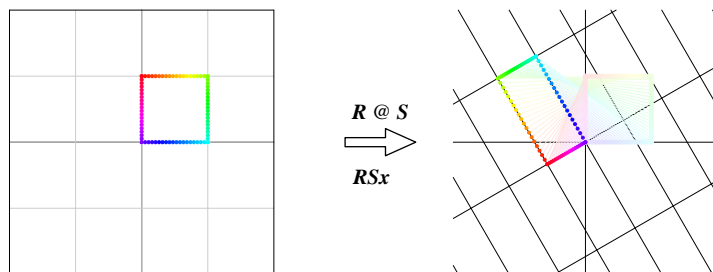


图 5. 复合变换

显然，如图 6 所示， $(R @ S)$  的逆操作为  $(R @ S)^{-1}$ ，即

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

版权归清华大学出版社所有，请勿商用，引用请注明出处。

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)

$$(R @ S)^{-1} @ (R @ S) @ x = I @ x = x \quad (8)$$

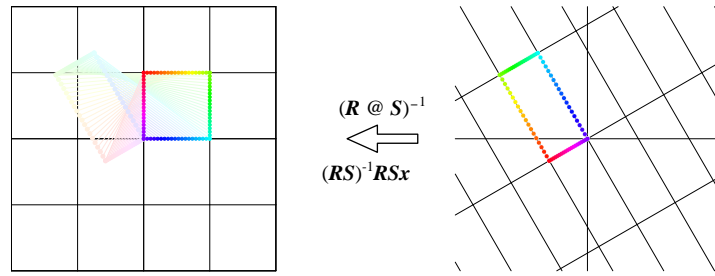


图 6. 复合变换的逆变换

把  $(R @ S)$  复合几何操作展开来看，下式代表“先缩放，再旋转”

$$R @ S @ x = \underbrace{\begin{bmatrix} -1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{bmatrix}}_{\text{Rotate}} @ \underbrace{\begin{bmatrix} 3/2 & 0 \\ 0 & 2/3 \end{bmatrix}}_{\text{Scale}} @ x \quad (9)$$

上述几何操作对应图 7。

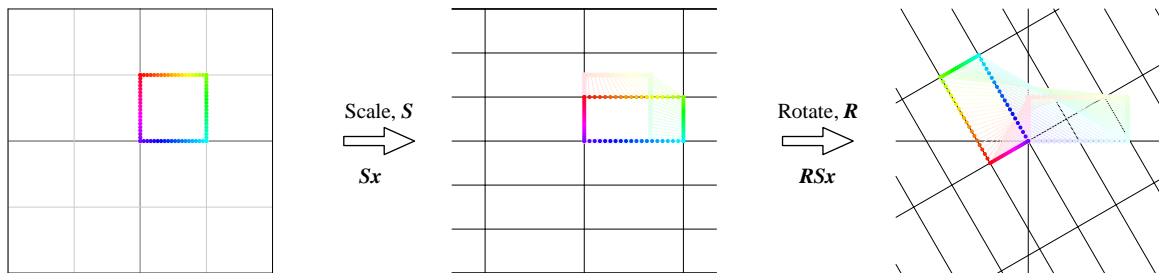


图 7. 先缩放、再旋转

逆向来看，我们需要先消去  $R$ ，再消去  $S$

$$S^{-1} @ R^{-1} @ R @ S @ x = I @ x = x \quad (10)$$

逆向先后操作对应图 8。

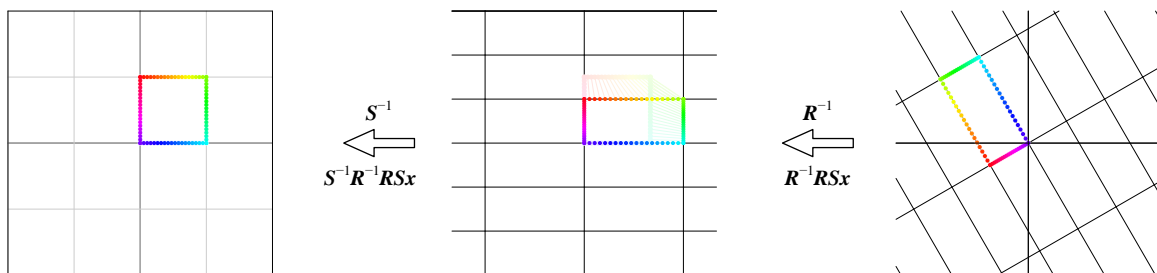


图 8. 逆变换：先反向旋转、再反向缩放

如果先反向缩放，再反向旋转，会发生什么？

$$R @ S @ R^{-1} @ S^{-1} @ x \quad (11)$$

图 9 所示为错误顺序的逆向变换；显然，我们没有恢复原始图形（单位正方形）。

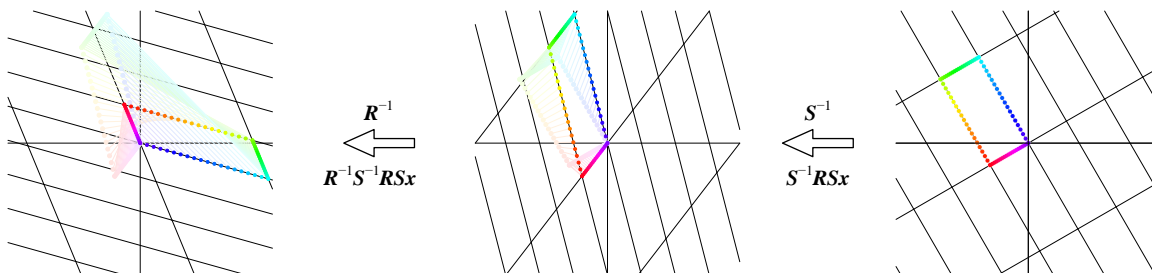


图 9. 错误的逆变换：先反向缩放、再反向旋转

## 对角方阵的逆

上一节，我们已经看到对角矩阵的作用：每个对角元素都是一个独立的缩放因子。

比如，给定缩放矩阵  $S$

$$S = \begin{bmatrix} 2/3 & 0 \\ 0 & 3/2 \end{bmatrix} \quad (12)$$

如图 10 所示， $S$  对二维空间中的向量分别沿  $x_1$  轴、 $x_2$  轴进行缩放。

$S$  的逆矩阵

$$S^{-1} = \begin{bmatrix} 3/2 & 0 \\ 0 & 2/3 \end{bmatrix} \quad (13)$$

实际上就是对角元素的倒数。

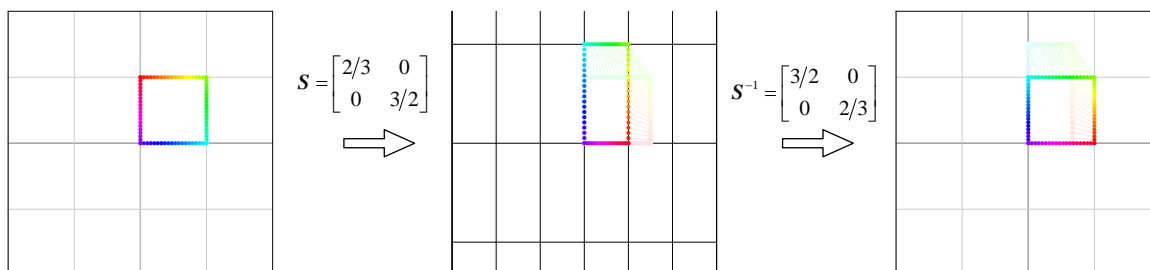


图 10. 平面上的缩放、缩放的逆操作

## 对称矩阵的逆

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

版权归清华大学出版社所有，请勿商用，引用请注明出处。

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)

大家已经清楚，如果  $A$  是一个对称矩阵，即满足  $A = A^T$ ，那么它的逆矩阵  $A^{-1}$  也是对称矩阵

$$(A^{-1})^T = A^{-1} \quad (14)$$

简单证明一下

$$(A^{-1})^T \underset{I}{AA^{-1}} = (A^{-1})^T \underset{I}{A^T A^{-1}} = (AA^{-1})^T \underset{I}{A^{-1}} = A^{-1} \quad (15)$$

## 矩阵转置的逆

如果  $A$  可逆，则有

$$(A^{-1})^T = (A^T)^{-1} \quad (16)$$

简单证明一下

$$(A^{-1})^T = (A^{-1})^T \underset{I}{A^T (A^T)^{-1}} = \left( (A^{-1})^T A^T \right) (A^T)^{-1} = \underset{I}{(AA^{-1})^T} (A^T)^{-1} = (A^T)^{-1} \quad (17)$$

要想直观理解 (16) 这个等式背后的几何原理，我们需要借助奇异值分解 (Singular Value Decomposition, SVD)。



奇异值分解是本书后续要探讨的重要话题之一。

## 正交矩阵的逆

上一节告诉我们，如果矩阵  $A$  为正交矩阵， $A$  的逆为其转置，即

$$A^T = A^{-1} \quad (18)$$

旋转矩阵是一种正交矩阵，比如矩阵  $V$  让平面几何体绕原点逆时针旋转  $\theta$

$$V = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \quad (19)$$

$V$  的逆矩阵为：

$$V^{-1} = V^T = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \quad (20)$$

$V$  和  $V^T$  的乘积为单位矩阵

$$V @ V^T = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (21)$$

$V^T$  和  $V$  的乘积也是单位矩阵 (形状和上式相同)

$$V @ V^T = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (22)$$

需要大家注意的是，单位矩阵、镜像矩阵、置换矩阵也都是正交矩阵。



## $kA$ 的逆

如果方阵  $A$  可逆，则  $kA$  的逆为

$$(kA)^{-1} = \frac{1}{k} A^{-1} \quad (23)$$

其中， $k$  不为 0。简单来说，逆变换时，除了  $A^{-1}$ ，还需要反向缩放，即乘以  $1/k$ ，来抵消  $k$  的影响。

换个角度看， $kA$  相当于

$$kA = k \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix} A = \begin{bmatrix} k & & & \\ & k & & \\ & & \ddots & \\ & & & k \end{bmatrix} A \quad (24)$$

把  $n \times n$  方阵  $A$  写成一组行向量，上式可以写成

$$\begin{bmatrix} k & & & \\ & k & & \\ & & \ddots & \\ & & & k \end{bmatrix} A = \begin{bmatrix} k & & & \\ & k & & \\ & & \ddots & \\ & & & k \end{bmatrix} \begin{bmatrix} \mathbf{a}^{(1)} \\ \mathbf{a}^{(2)} \\ \vdots \\ \mathbf{a}^{(n)} \end{bmatrix} = \begin{bmatrix} k\mathbf{a}^{(1)} \\ k\mathbf{a}^{(2)} \\ \vdots \\ k\mathbf{a}^{(n)} \end{bmatrix} \quad (25)$$

我们发现，对角方阵和  $A$  相乘，对角方阵的对角线元素对  $A$  的行向量进行了缩放；如图 11 所示，如果对角方阵对角线元素不同，则对每一行的缩放系数不同。

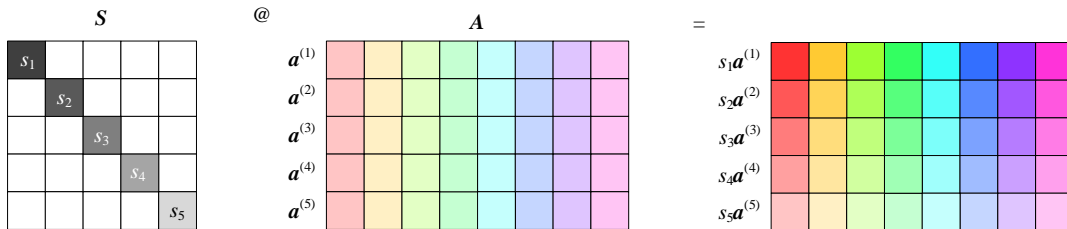


图 11. 对扁平矩阵  $A$  的每行缩放

反过来，如果把  $kA$  写成

$$kA = A \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix} k = A \begin{bmatrix} k & & & \\ & k & & \\ & & \ddots & \\ & & & k \end{bmatrix} \quad (26)$$

再把  $n \times n$  方阵  $A$  写成一组列向量，上式可以写成

$$A \begin{bmatrix} k & & & \\ & k & & \\ & & \ddots & \\ & & & k \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} k & & & \\ & k & & \\ & & \ddots & \\ & & & k \end{bmatrix} = \begin{bmatrix} k\mathbf{a}_1 & k\mathbf{a}_2 & \cdots & k\mathbf{a}_n \end{bmatrix} \quad (27)$$

也就是说，如图 12 所示， $A$  和对角方阵相乘，对角方阵的对角线元素对  $A$  的列向量进行缩放。

注意，图 11 和图 12 中矩阵  $A$  形状不一致。

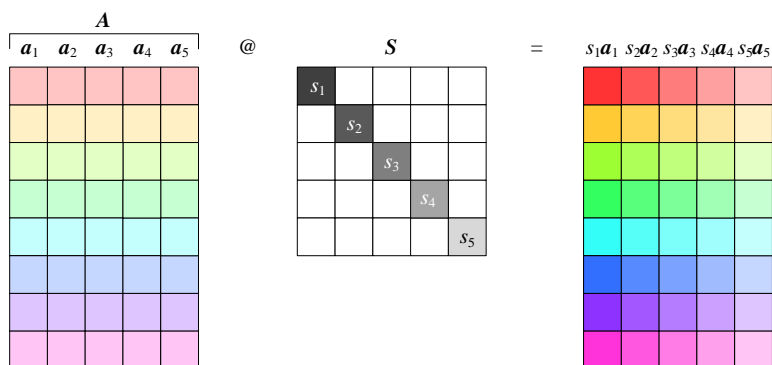


图 12. 对细高矩阵  $A$  的每列缩放

## 逆矩阵的行列式

大家已经很清楚，矩阵  $A$  的行列式  $\det(A)$  表示变换对面积、体积的缩放因子。如果矩阵  $A$  可逆，其逆矩阵的行列式为

$$\det(A^{-1}) = \frac{1}{\det(A)} \quad (28)$$

如图 13 所示，若  $A$  将单位正方形的面积放大 2 倍，则  $\det(A) = 2$ 。

其逆矩阵  $A^{-1}$  会把面积缩小为 1/2，所以  $\det(A^{-1}) = 1/2$ 。

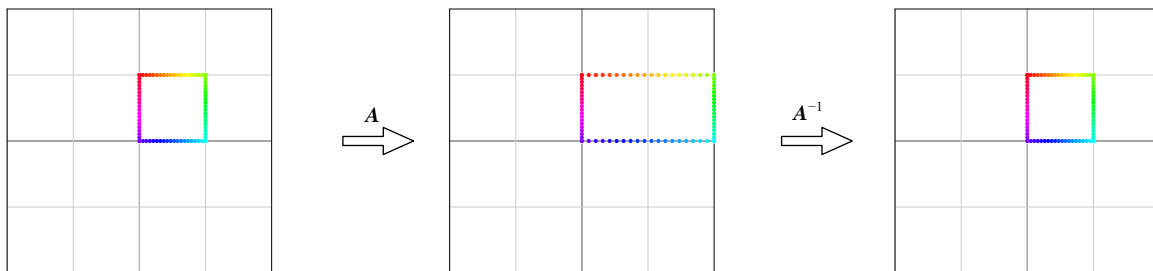


图 13. 面积的缩放



请大家用 DeepSeek/ChatGPT 等工具完成本节如下习题。

**Q1.** 以下哪些矩阵是正交矩阵

►  $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

▶  $\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$

▶  $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

▶  $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

▶  $\begin{bmatrix} 3/5 & -4/5 \\ 4/5 & 3/5 \end{bmatrix}$

▶  $\begin{bmatrix} 3 & -4 \\ 4 & 3 \end{bmatrix}$

**Q2.** 请计算如下矩阵的逆，并解释几何意义。

▶  $\begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}$

▶  $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

▶  $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

▶  $\begin{bmatrix} 3/5 & -4/5 \\ 4/5 & 3/5 \end{bmatrix}$

**Q3.** 请指出哪些方阵不可逆，并说出你的根据。

▶  $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$

▶  $\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$

▶  $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$

▶  $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$