

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

14.2 几何视角看 SVD



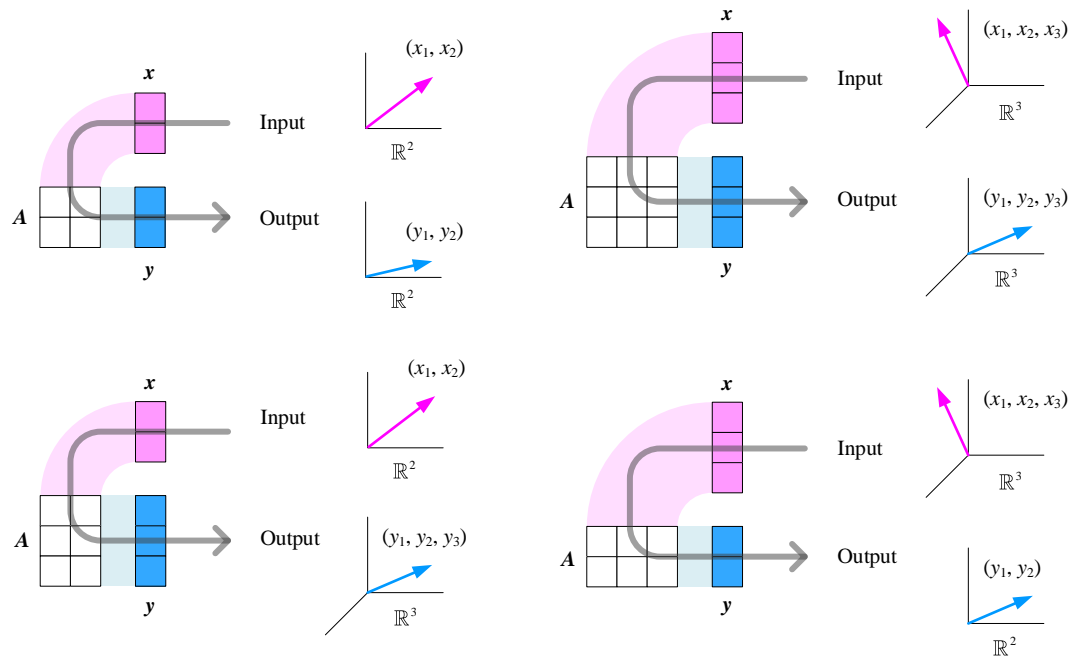
本节你将掌握的核心技能：

- ▶ SVD 将复杂的线性变换分解为“旋转 → 缩放 → 旋转”三个步骤。
- ▶ 旋转可能发生在不同空间。
- ▶ 缩放中可能发生维度变化。
- ▶ 对称矩阵的 SVD 等价于谱分解。
- ▶ 几何变换的面积、体积变换由奇异值控制。
- ▶ 奇异值为 0 意味着降维，向量投影到更低维空间，体现秩的缺失。

为了帮助大家更好理解奇异值分解，本节从几何角度介绍四个不同形状的矩阵的奇异值分解。学习本节之前请大家回顾第 2 章第 5 节，以及本书第 8 章各种几何操作。

四个不同形状矩阵

具体来说，我们将分析图 1 中四种不同形状的矩阵—— 2×2 、 3×3 、 3×2 和 2×3 ——对向量结合操作。

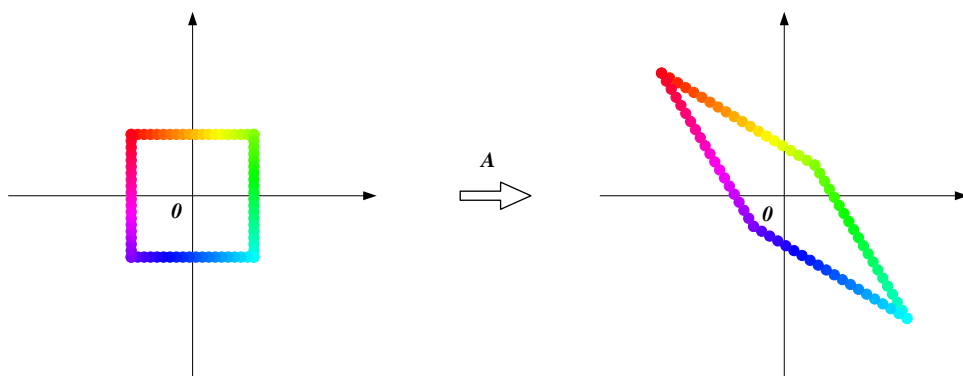
图 1. 四个不同形状的完成的矩阵 A 线性映射

2 × 2 方阵

让我们首先看 2×2 方阵。给定如下矩阵 A

$$A = \begin{bmatrix} 5/4 & -3/4 \\ -3/4 & 5/4 \end{bmatrix} \quad (1)$$

图 2 所示为 (1) 中 2×2 方阵对二维列向量 x 的几何操作。我们发现，该几何变换并不能归结为某一单一类型的操作，而是多个基本几何操作（如旋转、剪切、缩放等等）的有序组合，构成了一种复合性的线性变换。而奇异值分解可以帮助我们理解这个复合线性变换。

图 2. 2×2 对称矩阵 A 的线性变换

矩阵 A 的 SVD 分解结果为

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T = \underbrace{\begin{bmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}}_{\mathbf{U}} \underbrace{\begin{bmatrix} 2 & 0 \\ 0 & 1/2 \end{bmatrix}}_{\mathbf{S}} \underbrace{\begin{bmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}}_{\mathbf{V}^T} = \begin{bmatrix} 5/4 & -3/4 \\ -3/4 & 5/4 \end{bmatrix} \quad (2)$$

$\mathbf{A}\mathbf{x}$ 可以写成 $\mathbf{U}\mathbf{S}\mathbf{V}^T\mathbf{x}$ ，对应的几何操作的顺序 \mathbf{V}^T (旋转) $\rightarrow \mathbf{S}$ (缩放) $\rightarrow \mathbf{U}$ (旋转)，具体如图 3 所示。

由于 (1) 中矩阵 \mathbf{A} 为实对称矩阵，所以 (2) 和谱分解的结果一致。

如所示， \mathbf{V}^T 、 \mathbf{U} 分别对应的旋转操作并不改变面积。而对角矩阵 \mathbf{S} 决定了图形的缩放。

计算的行列式为 1，这意味着图形的面积没有变换，即 $\det(\mathbf{A}) = \det(\mathbf{S})$ 。由于行列式为正，图形也没有发生翻转。

特别地，由于 \mathbf{U} 、 \mathbf{V}^T 互为逆变换，所以旋转角度正好相反。

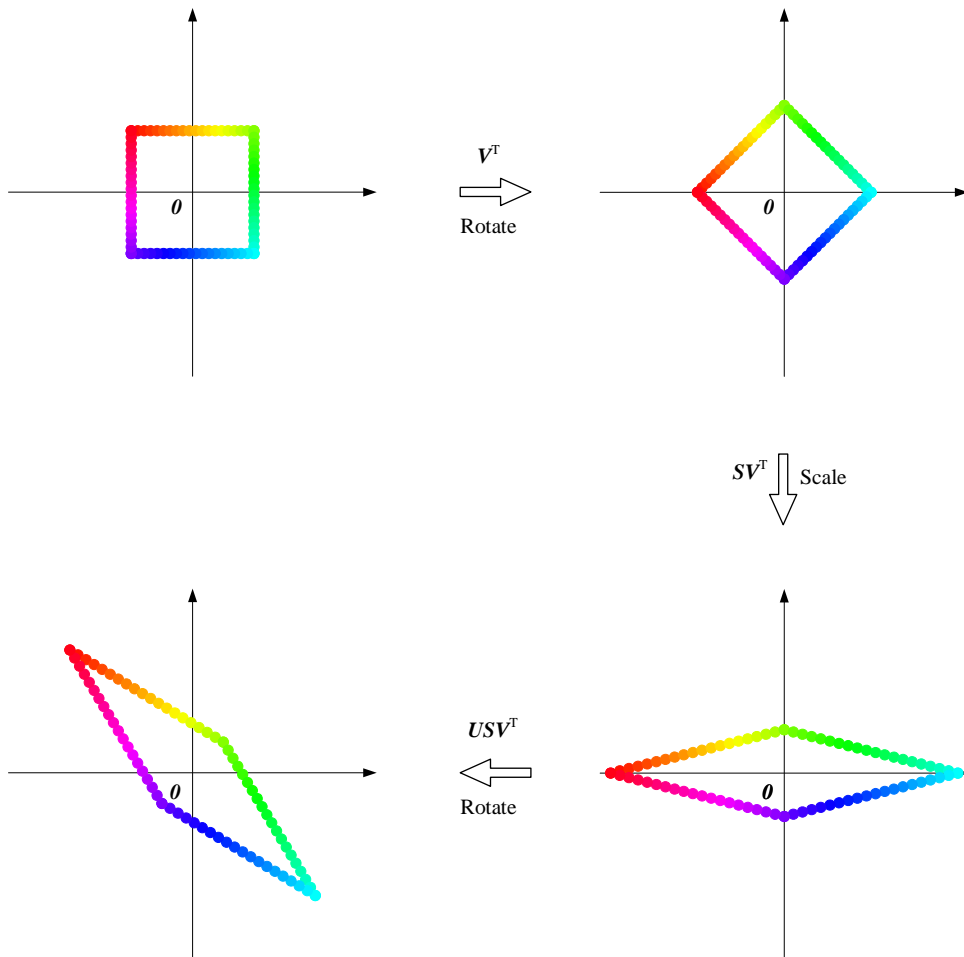


图 3. 2×2 对称矩阵 \mathbf{A} 奇异值分解结果 $\mathbf{U}\mathbf{S}\mathbf{V}^T$ 对应的线性变换

让我们再看一个 2×2 非对称矩阵 \mathbf{A} ，具体如下

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} \quad (3)$$

图 4 所示为 2×2 非对称矩阵 A 产生的线性变换。

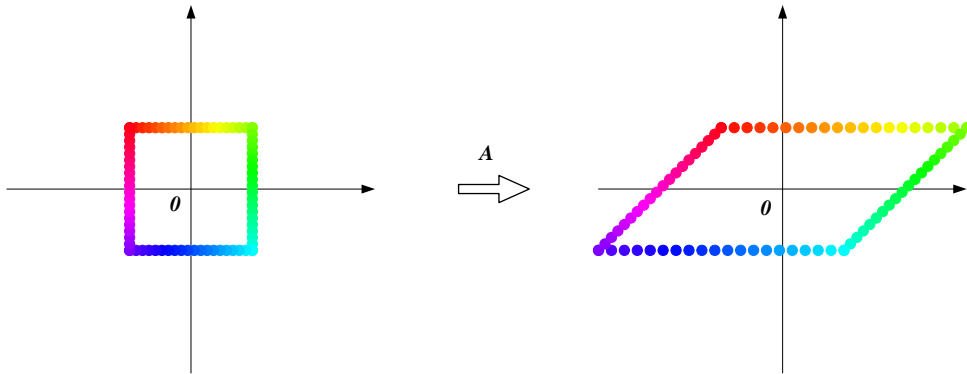


图 4. 2×2 非对称矩阵 A 的线性变换

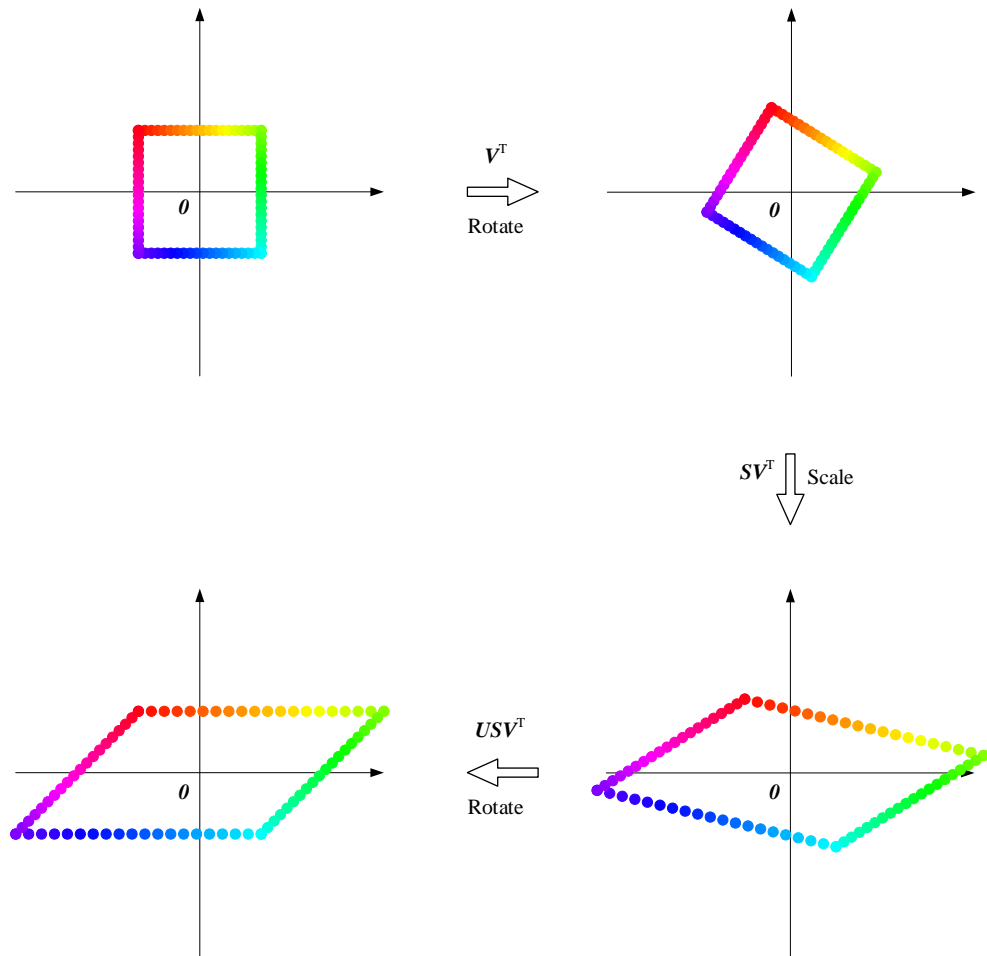
对 (3) 中矩阵 A 的 SVD 分解结果为

$$A = USV^T = \underbrace{\begin{bmatrix} 0.973 & -0.229 \\ 0.229 & 0.973 \end{bmatrix}}_U \underbrace{\begin{bmatrix} 2.288 & 0 \\ 0 & 0.874 \end{bmatrix}}_S \underbrace{\begin{bmatrix} 0.851 & 0.526 \\ -0.526 & 0.851 \end{bmatrix}}_{V^T} = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} \quad (4)$$

⚠ 注意，(4) 仅保留三位小数。

如图 5 所示，(4) 中 V^T 的几何操作是平面绕原点旋转。 S 沿横轴放大，沿纵轴缩小。 U 的操作也是平面绕原点旋转。

由于(4)中 A 不是对称矩阵， U 的旋转并不 V^T 的逆操作。

图 5. 2×2 非对称矩阵 A 奇异值分解结果 USV^T 对应的线性变换

3 × 3 方阵

下面让我们再聊聊 3×3 方阵的奇异值分解。

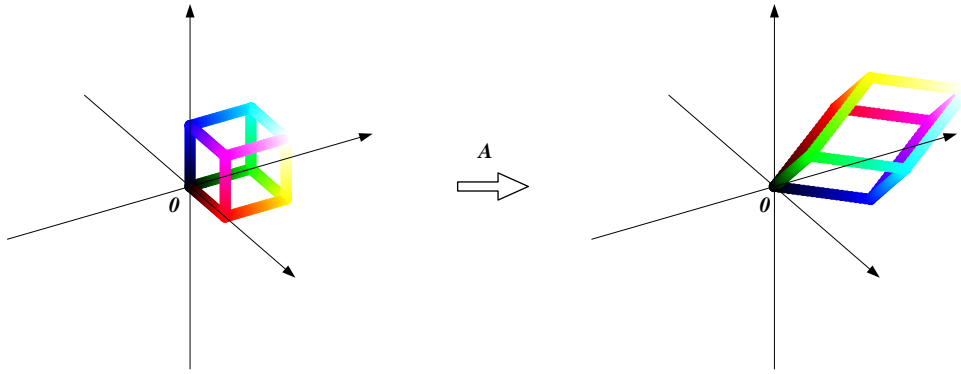
给定如下 3×3 对称方阵 A

$$A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \quad (5)$$

上式中矩阵 A 为满秩，意味着列向量不共线，也意味着行列式不为 0。

图 6 所示为 (5) 中 3×3 对称方阵 A 的几何变换。

? 请大家用拉普拉斯展开计算 (5) 中方阵 A 的行列式。

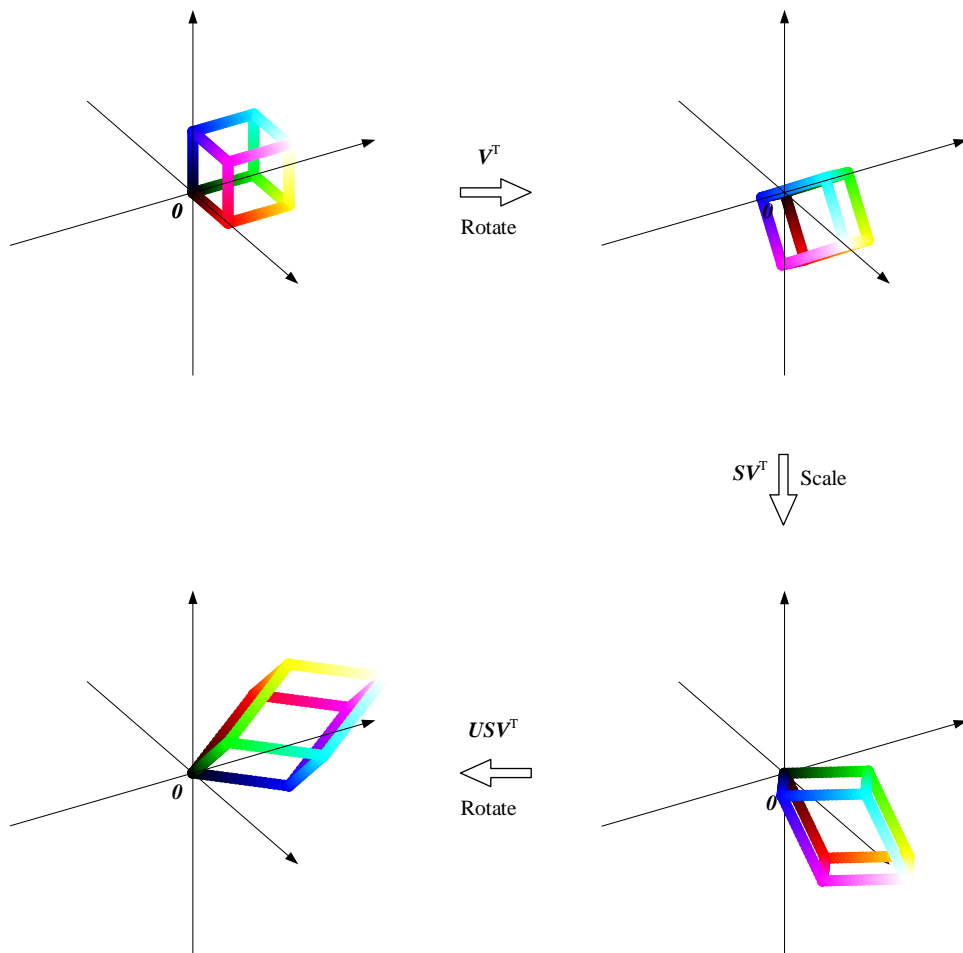
图 6. 3×3 矩阵 A (满秩) 的线性变换

(5) 中 3×3 方阵 A 的奇异值分解为

$$A = USV^T = \underbrace{\begin{bmatrix} \sqrt{3}/3 & 0 & \sqrt{6}/3 \\ \sqrt{3}/3 & -\sqrt{2}/2 & -\sqrt{6}/6 \\ \sqrt{3}/3 & \sqrt{2}/2 & -\sqrt{6}/6 \end{bmatrix}}_U \underbrace{\begin{bmatrix} 2 & & \\ & 1 & \\ & & 1 \end{bmatrix}}_S \underbrace{\begin{bmatrix} \sqrt{3}/3 & \sqrt{3}/3 & \sqrt{3}/3 \\ 0 & \sqrt{2}/2 & -\sqrt{2}/2 \\ -\sqrt{6}/3 & \sqrt{6}/6 & \sqrt{6}/6 \end{bmatrix}}_{V^T} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \quad (6)$$

如图 7 所示, (6) 中 V^T 的几何操作是三维空间旋转。 S 在三维空间缩放。 U 的操作也是三维空间旋转。

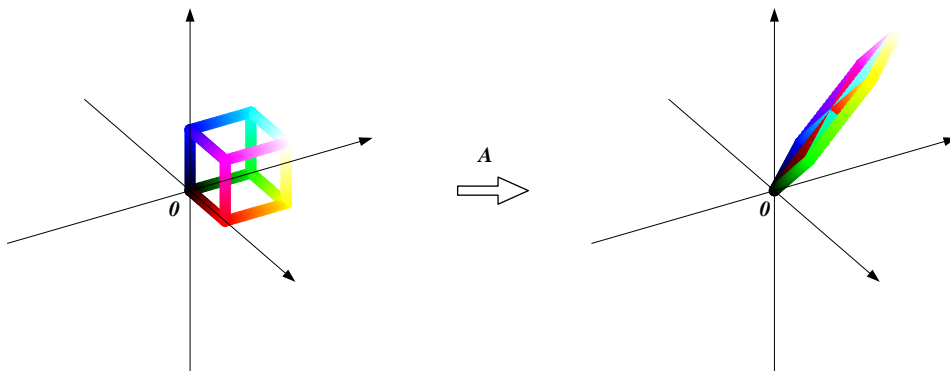
很容易计算得到 $\det(S) = 2$ 。由于 A 为方阵, $\det(A) = \det(S)$, $\det(A)$ 也为 2。

图 7. 3×3 对称矩阵 A (满秩) 奇异值分解结果 USV^T 对应的线性变换

再看一个非满秩 3×3 方阵 A

$$A = \begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \quad (7)$$

很容易发现 $a_1 - a_2 = a_3$ ，这说明方阵 A 列向量相关，行列式显然为 0。如图 8 所示， A 的线性变换导致降维！

图 8. 3×3 矩阵 A (不满秩) 的线性变换

(7) 中 3×3 方阵 A 的奇异值分解

$$A = USV^T = \underbrace{\begin{bmatrix} 0 & \sqrt{6}/3 & -\sqrt{3}/3 \\ \sqrt{2}/2 & -\sqrt{6}/6 & -\sqrt{3}/3 \\ \sqrt{2}/2 & \sqrt{6}/6 & \sqrt{3}/3 \end{bmatrix}}_U \underbrace{\begin{bmatrix} \sqrt{3} & & \\ & \sqrt{3} & \\ & & 0 \end{bmatrix}}_S \underbrace{\begin{bmatrix} \sqrt{6}/3 & \sqrt{6}/6 & \sqrt{6}/6 \\ 0 & \sqrt{2}/2 & -\sqrt{2}/2 \\ -\sqrt{3}/3 & \sqrt{3}/3 & \sqrt{3}/3 \end{bmatrix}}_{V^T} = \begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \quad (8)$$

我们发现有一个奇异值为 0。

注意，(8) 中 A 的秩为 2， U 、 V 满秩，而 S 的秩也为 2。

如图 9 所示，(8) 中 V^T 的几何操作是三维空间旋转。 S 在三维空间“缩放 + 降维”。 U 的操作也是三维空间旋转。

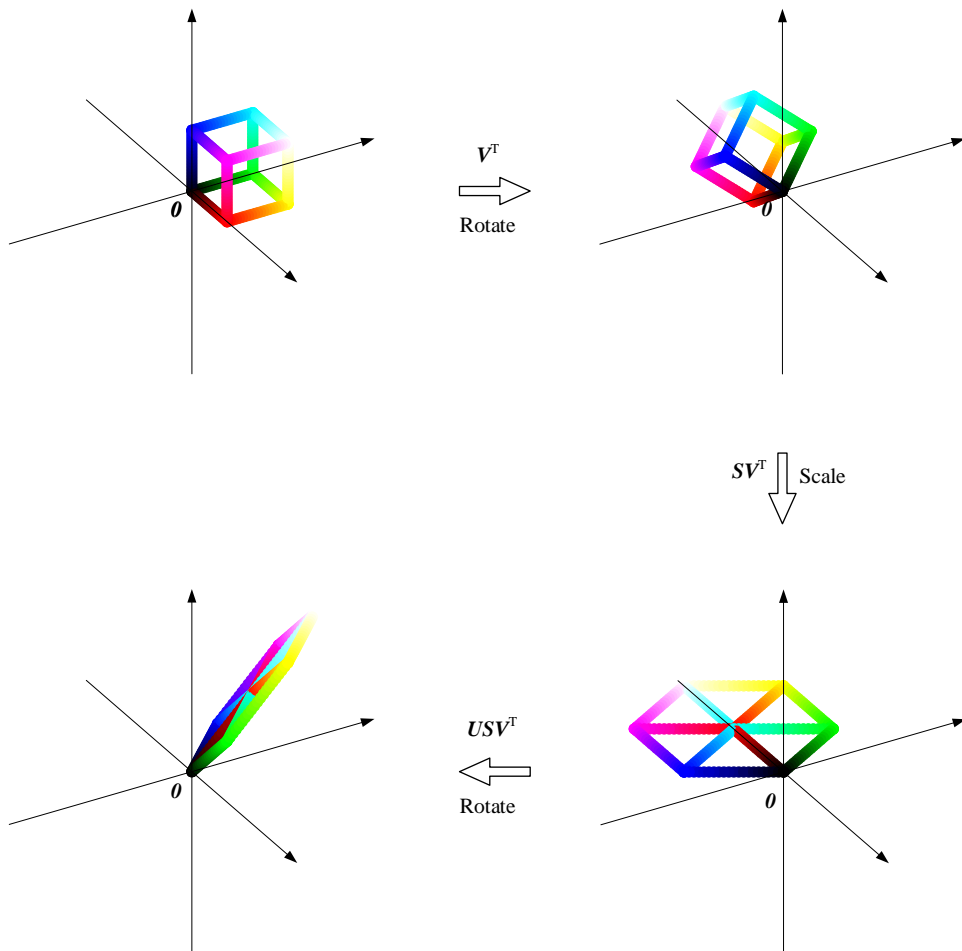


图 9. 3×3 非对称矩阵 A (不满秩) 奇异值分解结果 USV^T 对应的线性变换

3 × 2 细高矩阵

让我们再看 3×2 细高矩阵的奇异值分解。

给定如下 3×2 细高矩阵 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

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix} \quad (9)$$

显然上式中矩阵 A 列满秩。 $Ax=y$ 对应的几何操作如图 10 所示。列向量 x 在平面上，而 y 在三维空间中；所以， $Ax=y$ 完成了“升维”！

观察图 10，容易发现，尽管向量从二维扩展到了三维，但是仍然在一个平面上。



在第 2 章第 5 节讲解矩阵乘法的几何视角时详细讲过这一点，请大家回顾。

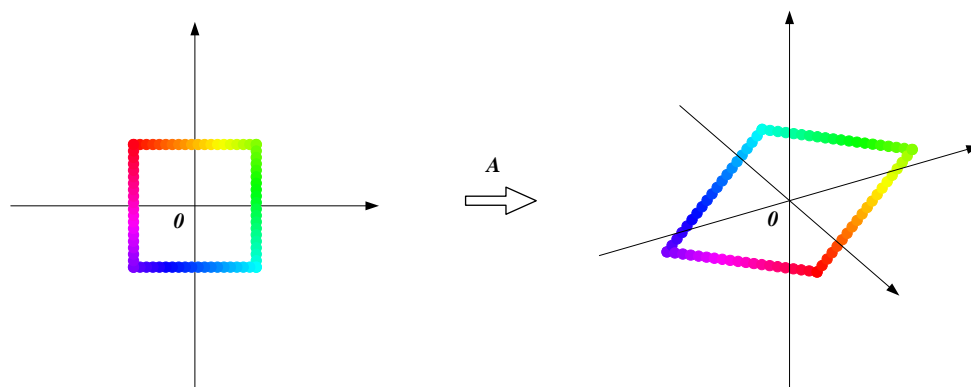


图 10. 3×2 细高矩阵 A （列满秩）的线性变换

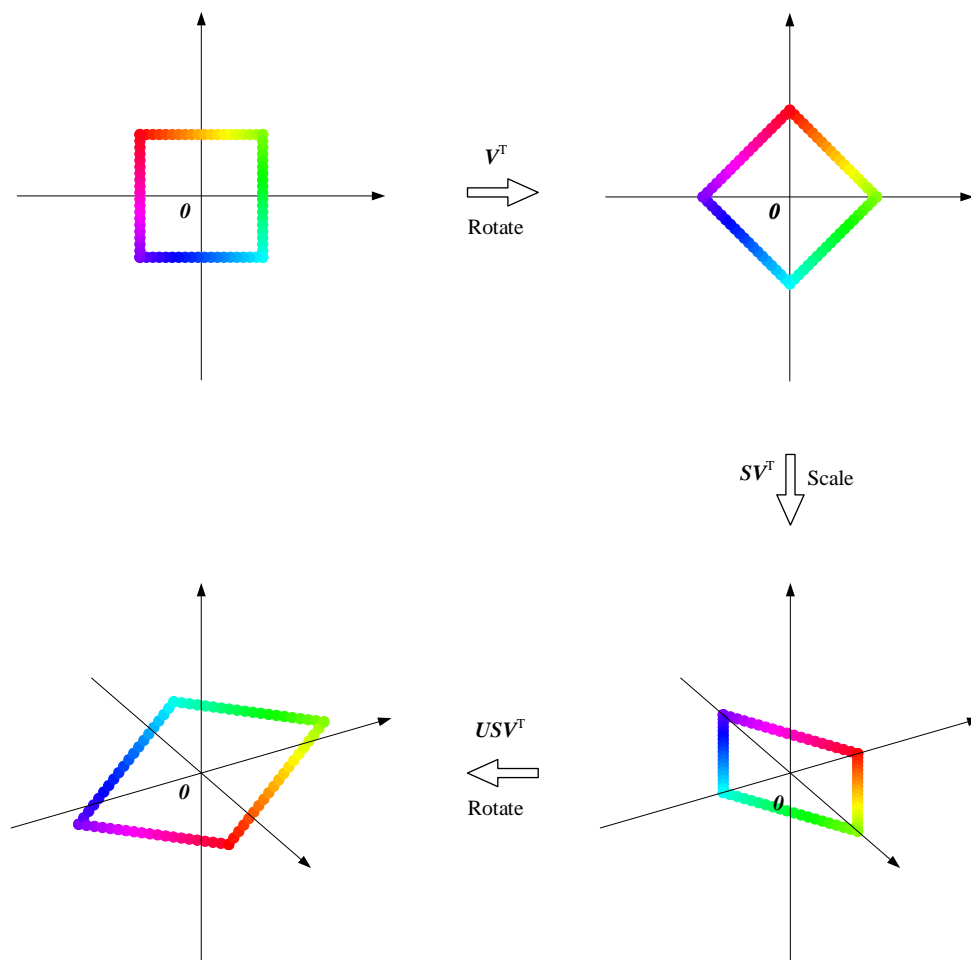
(9) 中这个矩阵 A 的完全型 SVD 分解为

$$A = USV^T = \underbrace{\begin{bmatrix} \sqrt{6}/6 & \sqrt{2}/2 & \sqrt{3}/3 \\ \sqrt{6}/3 & 0 & -\sqrt{3}/3 \\ \sqrt{6}/6 & -\sqrt{2}/2 & \sqrt{3}/3 \end{bmatrix}}_U \underbrace{\begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}}_S \underbrace{\begin{bmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}}_{V^T} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix} \quad (10)$$

注意，(10) 中 A 的秩为 2， U 、 V 满秩，而 S 的秩也为 2。

如图 11 所示， V^T 在平面（绕原点）旋转。 S 完成“缩放 + 升维”。 U 的操作是三维空间旋转。

打个比方，上式中 S 相当于把一张照片（二维）放在桌子上（三维）。

图 11. 3×2 细高矩阵 A (列满秩) 奇异值分解结果 USV^T 对应的线性变换

给定如下列不满秩的 3×2 细高矩阵 A

$$A = \begin{bmatrix} 0 & 0 \\ 1 & -1 \\ 1 & -1 \end{bmatrix} \quad (11)$$

图 12 所示为 (11) 的线性转换。我们发现所有的向量都落在了一条直线上。

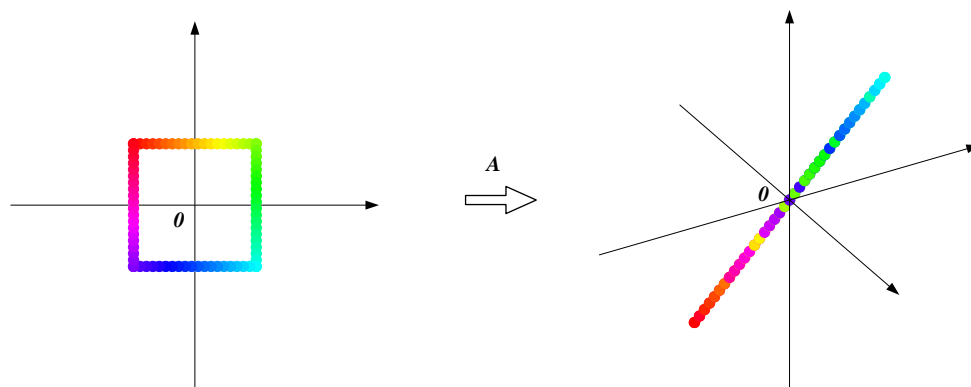


图 12. 3×2 细高矩阵 A (列不满秩) 的线性变换

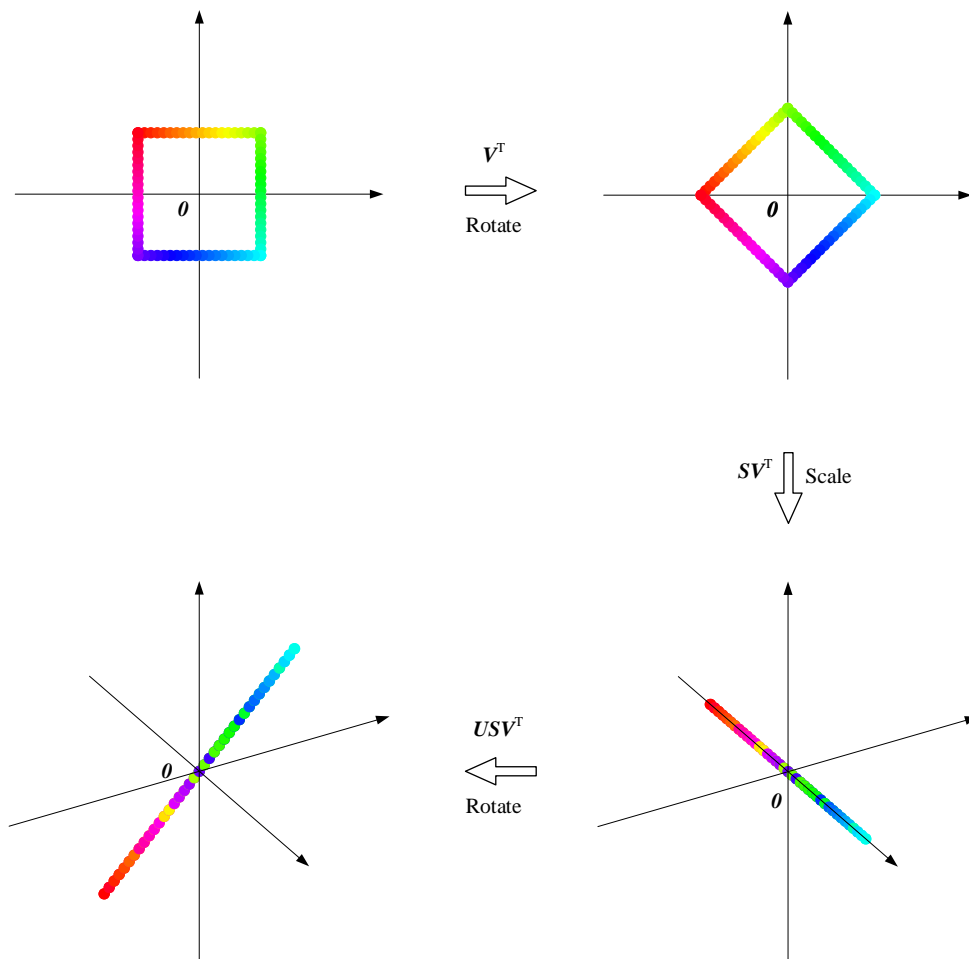
(11) 中矩阵 A 的完全型 SVD 分解为

$$A = USV^T = \underbrace{\begin{bmatrix} 0 & 1 & 0 \\ \sqrt{2}/2 & 0 & -\sqrt{2}/2 \\ \sqrt{2}/2 & 0 & \sqrt{2}/2 \end{bmatrix}}_U \underbrace{\begin{bmatrix} 2 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}}_S \underbrace{\begin{bmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}}_{V^T} = \begin{bmatrix} 0 & 0 \\ 1 & -1 \\ 1 & -1 \end{bmatrix} \quad (12)$$

如图 13 所示, V^T 在平面 (绕原点) 旋转。 S 完成“缩放 + 升维”。 U 的操作是三维空间旋转。

由于 S 中一个缩放系数为 0, 所以该方向本质上是降维, 这也就是为什么图 13 中所有散点落在同一条直线上。

注意, (12) 中 A 的秩为 1, U 、 V 满秩, 而 S 的秩也为 1。

图 13. 3×2 细高矩阵 A (列满秩) 奇异值分解结果 USV^T 对应的线性变换

2 × 3 扁平矩阵

本节最后让我们看 2×3 扁平矩阵。

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

给定如下行满秩 3×2 扁平矩阵 A

$$A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} \quad (13)$$

$Ax = y$ 对应的几何操作如图 14 所示。列向量 x 在三维空间中，而 y 在平面上；所以， $Ax = y$ 完成了“降维”。我们在图 14 中看到，单位立方体确实被压缩到平面上。

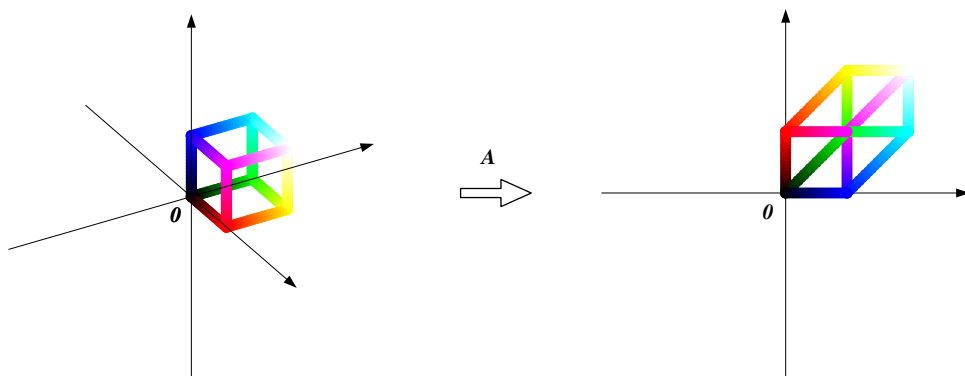


图 14. 2×3 扁平矩阵 A （行满秩）的线性变换

(13) 中这个矩阵 A 的完全型 SVD 分解为

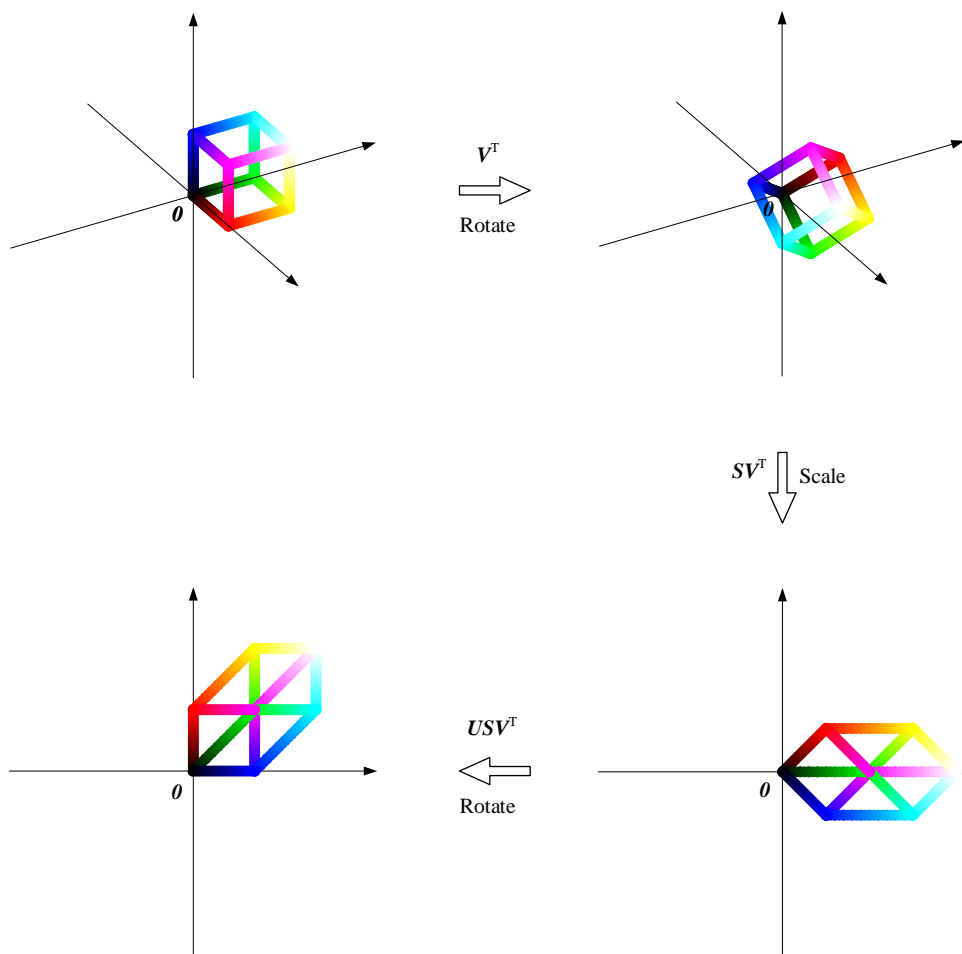
$$A = USV^T = \underbrace{\begin{bmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}}_U \underbrace{\begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}}_S \underbrace{\begin{bmatrix} \sqrt{6}/6 & \sqrt{6}/3 & \sqrt{6}/6 \\ \sqrt{2}/2 & 0 & -\sqrt{2}/2 \\ \sqrt{3}/3 & -\sqrt{3}/3 & \sqrt{3}/3 \end{bmatrix}}_{V^T} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} \quad (14)$$

如图 15 所示， V^T 先让单位立方体在三维空间中旋转。

S 让向量发生“缩放 + 降维”。这好比给空间几何形状拍了张照片，图像变成了二维。

最后， U 让平面几何图形绕原点旋转。

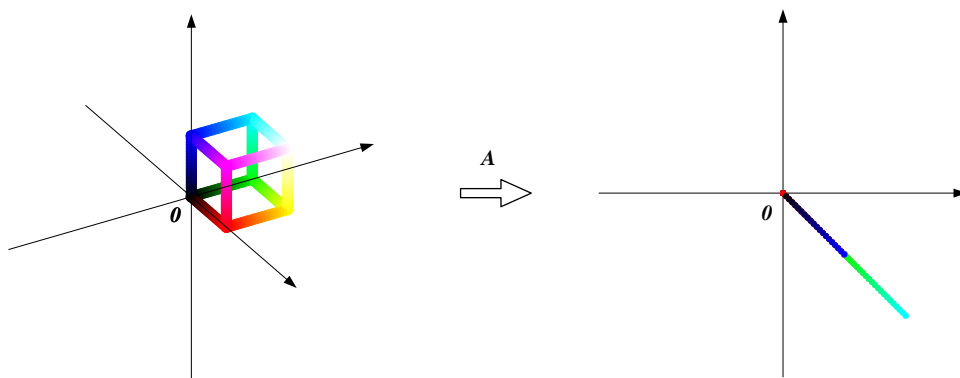
注意，(14) 中 A 的秩为 2， U 、 V 满秩，而 S 的秩也为 2。

图 15. 2×3 扁平矩阵 A (行满秩) 奇异值分解结果 USV^T 对应的线性变换

让我们看本节最后一个例子。给定如下 3×2 扁平矩阵 A

$$A = \begin{bmatrix} 0 & 1 & 1 \\ 0 & -1 & -1 \end{bmatrix} \quad (15)$$

很容易发现上式这个矩阵行不满秩，它对应的线性变换如图 17 所示。

图 16. 2×3 扁平矩阵 A (行不满秩) 的线性变换

(15) 中矩阵 A 的完全型 SVD 分解为

$$A = USV^T = \underbrace{\begin{bmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}}_U \underbrace{\begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_S \underbrace{\begin{bmatrix} 0 & \sqrt{2}/2 & \sqrt{2}/2 \\ 1 & 0 & 0 \\ 0 & -\sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}}_{V^T} = \begin{bmatrix} 0 & 1 & 1 \\ 0 & -1 & -1 \end{bmatrix} \quad (16)$$

(16) 中分解得到的矩阵对应的顺序几何操作如图 17 所示。

? 请大家自行分析图 17。

注意，(16) 中 A 的秩为 1， U 、 V 满秩，而 S 的秩也为 1。

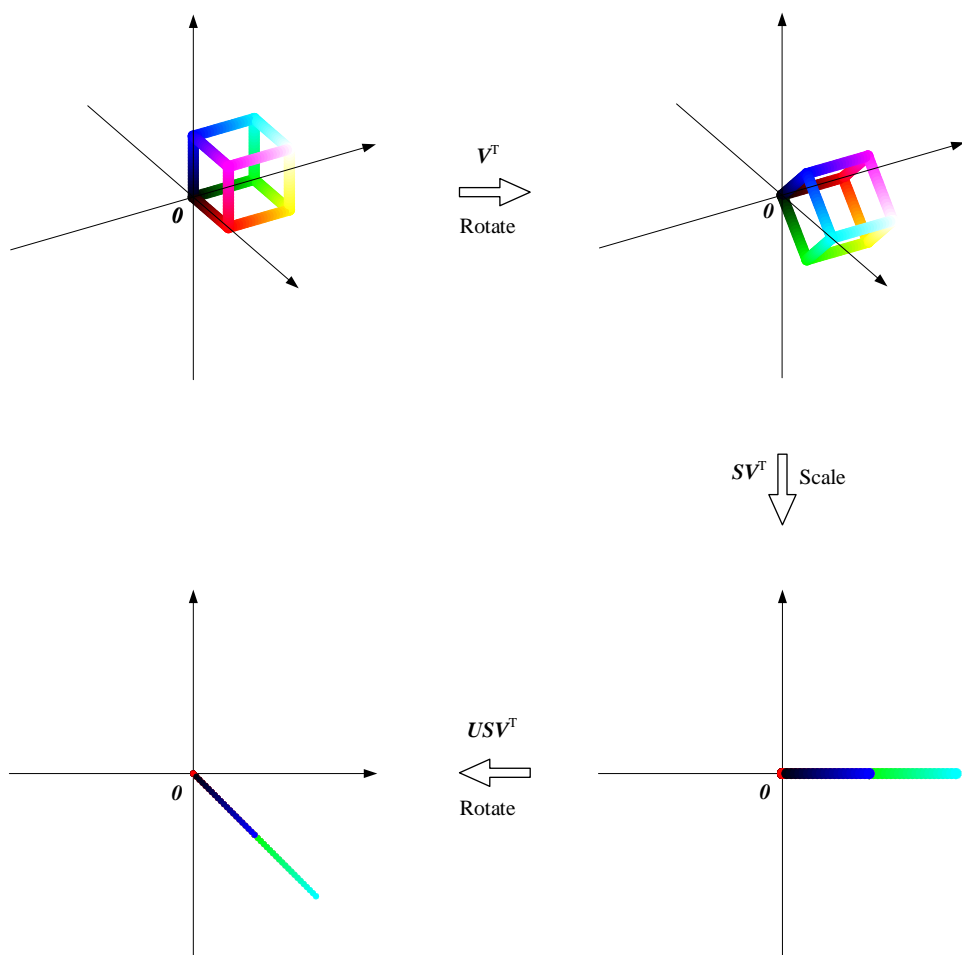


图 17. 2×3 扁平矩阵 A (行不满秩) 奇异值分解结果 USV^T 对应的线性变换



请大家用 DeepSeek/ChatGPT 等工具完成本节如下习题。

Q1. 请大家手算本节所有例子的奇异值分解，并用 Python 编程检验结果。