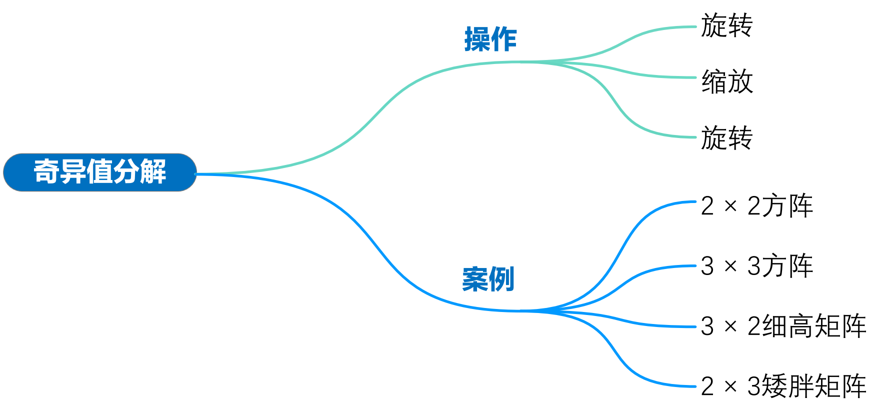
28 Singular Value Decomposition

奇异值分解

用图形视角展示奇异值分解的几何直觉

> 柏拉图我至亲，而真理价更高。
>
> ***Plato is dear to me, but dearer still is truth.***
>
> —— 亚里士多德 (Aristotle) | 古希腊哲学家 | 384 ~ 322 BC

- numpy.ones_like() 用来生成和输入矩阵形状相同的全 1 矩阵
- numpy.zeros_like() 用来生成和输入矩阵形状相同的零矩阵
- numpy.column_stack() 将两个矩阵按列合并
- numpy.vstack() 返回竖直堆叠后的数组
- numpy.linalg.svd() 奇异值分解



28.1 什么是奇异值分解?

这章内容非常重要。一方面，我们要通过这一节回顾二维、三维散点图的各种可视化技巧；另外一方面，我们要利用本书前两章介绍的平面、立体几何变换来展示奇异值分解。

特别希望大家通过本节学习能够建立起对奇异值分解 SVD 几何直觉。

《编程不难》轻描淡写地介绍过奇异值分解，在《编程不难》中我们了解了 NumPy（第 17 章）和 SymPy（第 25 章）中完成奇异值分解的函数。

简单来说，奇异值分解 (Singular Value Decomposition, SVD) 是一种将一个矩阵分解为三个矩阵乘积的线性代数工具。

给定一个矩阵 $\boldsymbol{A}$，它可以表示为 $\boldsymbol{A}=\boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^{\mathrm{T}}$，其中 $\boldsymbol{U}$ 和 $\boldsymbol{V}$ 是正交矩阵，$\boldsymbol{S}$ 是对角矩阵，对角线上的元素称为奇异值。

奇异值分解在线性代数、数据分析、机器学习中是一种无所不包的存在。本章将利用如图 1 所示四个形状各异的矩阵通过几何视角帮助大家理解奇异值分解。图 1 (a)和(b)为方阵；图 1 (c) 叫做细高矩阵，图 1 (d) 叫矮胖矩阵。图 1 (d) 是图 1 (c) 的转置。

注意，图 1 中每个形状矩阵对应一节，这些矩阵都叫矩阵 $\boldsymbol{A}$，SVD 分解的结果也都是 $\boldsymbol{U}$、$\boldsymbol{S}$、$\boldsymbol{V}^{\mathrm{T}}$，请大家注意区分。

此外，为了方便大家对比阅读，本章后续四节将采用几乎一致的编排结构。

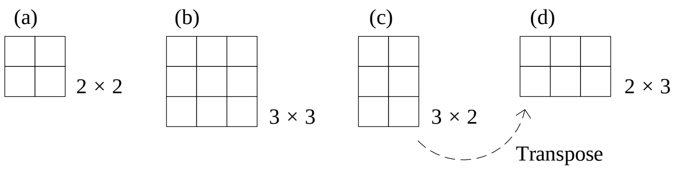


图 1. 四个不同的矩阵 $\boldsymbol{A}$

28.2 2 × 2 方阵

列向量

先从图 2 中展示矩阵乘法 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}$ 说起。

图 2 中矩阵 $\boldsymbol{A}$ 为 2 行 2 列，即形状为 2 × 2；列向量 $\boldsymbol{x}$ 为 2 行 1 列；结果列向量 $\boldsymbol{y}$ 也为 2 行 1 列。从线性代数角度，矩阵 $\boldsymbol{A}$ 完成了 $\boldsymbol{x}$ 到 $\boldsymbol{y}$ 的映射。

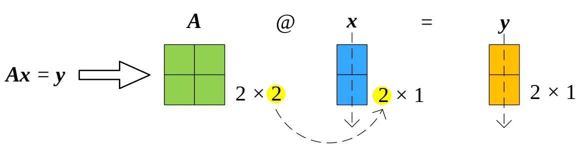


图 2. 列向量 $\boldsymbol{x}$ 在 2 × 2 方阵 $\boldsymbol{A}$ 映射下结果为列向量 $\boldsymbol{y}$

如图 3 所示，列向量 $\boldsymbol{x}=\begin{bmatrix}x_1\\x_2\end{bmatrix}$ 相当于平面直角坐标系中的一个点 (x_1, x_2)；$\boldsymbol{x}=\begin{bmatrix}x_1\\x_2\end{bmatrix}$ 也是平面上一个起点位于原点 (0,0)，终点位于 (x_1, x_2) 的向量。

而列向量 $\boldsymbol{y}=\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ 也是平面直角坐标系中的一个点 (y_1, y_2)；$\boldsymbol{y}=\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ 也是平面直角坐标系中一个起点位于原点 (0,0)，终点位于 $(y_1, y_2,)$ 的向量。大家可以回忆本书前文如何用箭头图可视化向量。

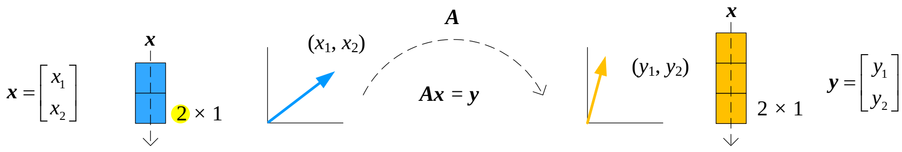


图 3. 2 × 2 方阵 $\boldsymbol{A}$ 映射关系

行向量

当然我们也可以把坐标点 (x_1, x_2) 写成行向量形式 $\boldsymbol{x}^{\mathrm{T}}=\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}^{\mathrm{T}}=\begin{bmatrix} x_1 & x_2 \end{bmatrix}$。

而结果 (y_1, y_2) 则跟着也写成行向量 $\boldsymbol{y}^{\mathrm{T}}=\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}^{\mathrm{T}}=\begin{bmatrix} y_1 & y_2 \end{bmatrix}$。

如图 4 所示，这时矩阵乘法写成 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{y}^{\mathrm{T}}$。这实际上是矩阵乘法等式 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}$ 左右转置的结果。大家需要注意顺序调换。

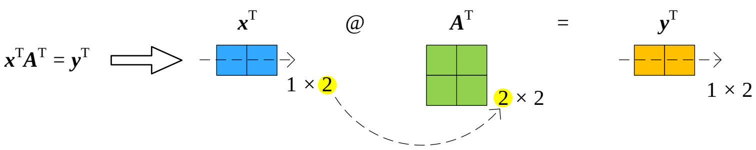


图 4. 行向量 $\boldsymbol{x}^{\mathrm{T}}$ 在 2 × 2 方阵 $\boldsymbol{A}$ 映射下结果为行向量 $\boldsymbol{y}^{\mathrm{T}}$

数据矩阵

为了更好地展示 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}$ 的映射过程，我们用平面上一组散点作为操作对象。

如图 5 所示，我们可以把这个平面上“小彩灯”的坐标排列成矩阵 $\boldsymbol{X}$ 的形式。

矩阵 $\boldsymbol{X}$ 有两列，第一列代表小彩灯横轴坐标 x_1，第二列代表代表小彩灯纵轴坐标 x_2。也就是说，矩阵 $\boldsymbol{X}$ 的每一行代表一个小彩灯。

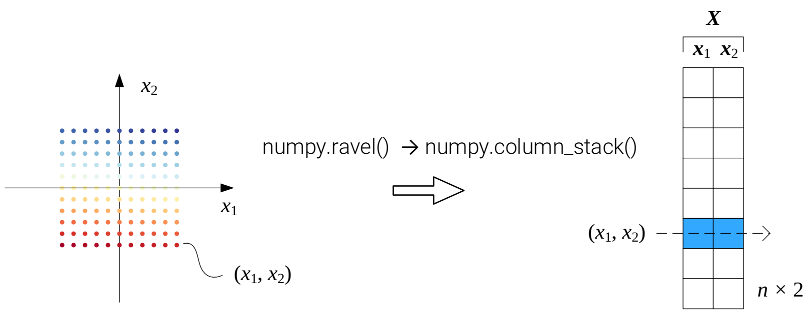


图 5. 用平面散点坐标创建矩阵 $\boldsymbol{X}$

如图 6 所示，在矩阵 $\boldsymbol{A}$ ($\boldsymbol{A}^{\mathrm{T}}$) 映射下，$\boldsymbol{X}$ 转化成了 $\boldsymbol{Y}$，即 $\boldsymbol{X}\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{Y}$。注意，我们还是把 $\boldsymbol{X}$ 的每行看做是坐标点；同时，$\boldsymbol{Y}$ 的每行代表一个坐标点。

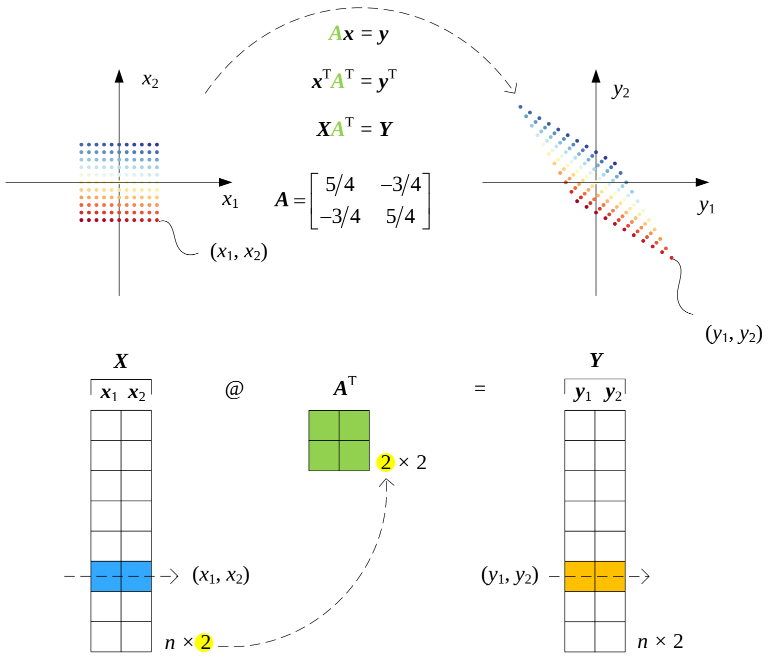
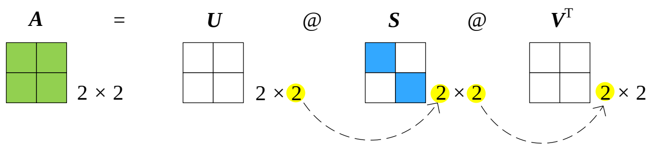


图 6．矩阵 X 在 2 × 2 方阵 A 映射下结果为矩阵 Y

SVD 分解

如图 7 所示，一个 2 × 2 方阵 $A = \begin{bmatrix} 5/4 & -3/4 \\ -3/4 & 5/4 \end{bmatrix}$ 经过完全 SVD 分解得到的结果是三个矩阵的连乘。学过线性代数的读者肯定会发现，图 7 的 SVD 分解也是一个 EVD 特征值分解；更确切地说，方阵 A 对称，图 7 中的矩阵分解也是一个谱分解。

如图 26 所示，U、S、V（V^{T}）这三个矩阵每个都对应一种几何变换！



$$A = \begin{bmatrix} 5/4 & -3/4 \\ -3/4 & 5/4 \end{bmatrix} = \underbrace{\begin{bmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}}_{U} @ \underbrace{\begin{bmatrix} 2 & \\ & 0.5 \end{bmatrix}}_{S} @ \underbrace{\begin{bmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}}_{V^{T}}$$

图 7．对 2 × 2 方阵 A 进行奇异值分解

如图 26 所示，2 × 2 方阵 $A = \begin{bmatrix} 5/4 & -3/4 \\ -3/4 & 5/4 \end{bmatrix}$ 完成的几何操作可以分解成 3 步，即把等式 $Ax = y$ 写成 $USV^{T}x = y$。根据矩阵乘法 $USV^{T}x = y$ 先后顺序，V^{T} 先完成平面旋转。

然后，S 完成缩放，小彩灯排列成“菱形”。

最后，U 还是平面上的旋转。注意，这些小彩灯构成的图形还是个“菱形”！

第二个例子

图 27 给出了另外一个 2 × 2 方阵 A 的几何操作。这个细高矩阵 A 直接将小彩灯排列成平面的一条直线上！

参考图 26，请大家根据 BK_2_Ch28_1.ipynb 将矩阵 A、U、S、V（V^{T}）的具体值自己写到图 27 上。

如果大家之前学过线性代数的话，可以自己算一下图 27 中矩阵 $\boldsymbol{A}$ 的秩，并试着从这个角度解释为什么小彩灯会出现在一条直线上。

28.3 3 × 3 方阵

这一节中，我们再聊一聊 3 × 3 方阵的奇异值分解。

列向量

如图 8 所示，矩阵 $\boldsymbol{A}$ 为 3 行 3 列；列向量 $\boldsymbol{x}$ 为 3 行 1 列；结果列向量 $\boldsymbol{y}$ 为 3 行 1 列。类似前文，矩阵 $\boldsymbol{A}$ 完成了 $\boldsymbol{x}$ 到 $\boldsymbol{y}$ 的映射。

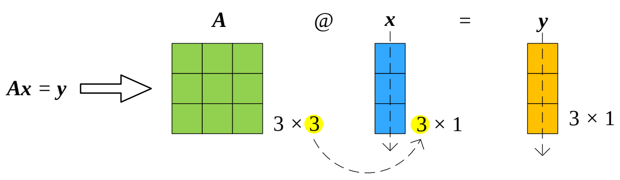


图 8．列向量 $\boldsymbol{x}$ 在 3 × 3 方阵 $\boldsymbol{A}$ 映射下结果为列向量 $\boldsymbol{y}$

如图 16 所示，列向量 $\boldsymbol{x}=\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ 是三维直角坐标系上一个点 (x_1, x_2, x_3)；$\boldsymbol{x}=\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ 也是三维直角坐标系一个起点位于原点 (0,0,0)，终点位于 (x_1, x_2, x_3) 的向量。

列向量 $\boldsymbol{y}=\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$ 则代表三维直角坐标系上一个点 (y_1, y_2, y_3)；$\boldsymbol{y}=\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$ 同样也是三维直角坐标系中一个起点位于原点 (0,0,0)，终点位于 (y_1, y_2, y_3) 的向量。

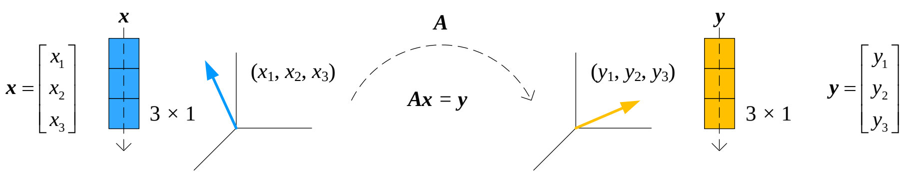


图 9．3 × 3 方阵 $\boldsymbol{A}$ 映射关系

行向量

类似前文，我们也可以把坐标点 (x_1, x_2, x_3) 写成行向量形式 $\boldsymbol{x}^{\mathrm{T}}=\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}^{\mathrm{T}}=\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}$。

坐标点 (y_1, y_2, y_3) 也写成 $\boldsymbol{y}^{\mathrm{T}}=\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}^{\mathrm{T}}=\begin{bmatrix} y_1 & y_2 & y_3 \end{bmatrix}$。

如图 10 所示，矩阵乘法也可以写成 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{y}^{\mathrm{T}}$，即 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}$ 等式左右转置。

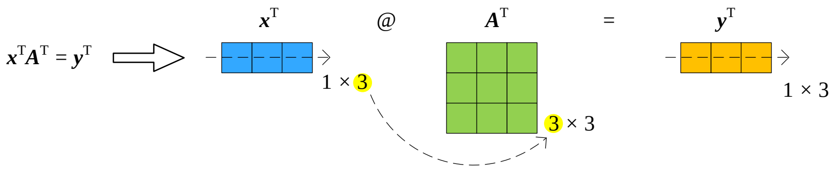


图 10. 行向量 $\boldsymbol{x}^{\mathrm{T}}$ 在 3 × 3 方阵 $\boldsymbol{A}$ 映射下结果为行向量 $\boldsymbol{y}^{\mathrm{T}}$

数据矩阵

为了更好地展示 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}$ 的映射过程，我们采用的操作对象是前文用过的 RGB 三维立方体的外框线。

如图 11 所示，这个三维空间“小彩灯”实际上就是 RGB 的单位立方体的外框线。我们把这些小彩灯的坐标排列成矩阵 $\boldsymbol{X}$ 的形式。本书后续还会用到这个可视化方案。

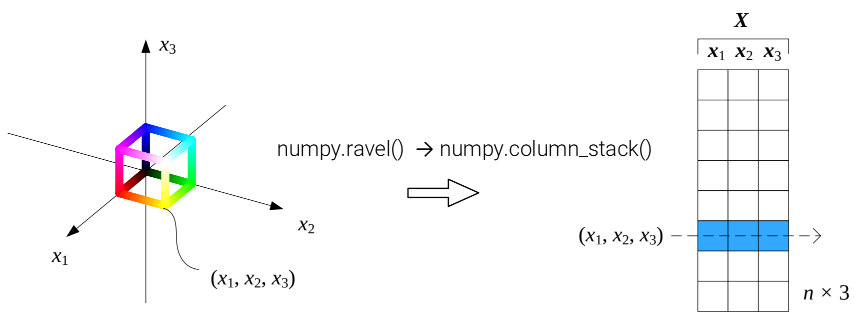


图 11. 用三维空间散点坐标创建矩阵 $\boldsymbol{X}$

如图 12 所示，在矩阵 $\boldsymbol{A}$ （$\boldsymbol{A}^{\mathrm{T}}$） 映射下，$\boldsymbol{X}$ 转化成了 $\boldsymbol{Y}$，即 $\boldsymbol{X}\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{Y}$。图 13 所以为在四个不同视角条件下观察结果结果 $\boldsymbol{Y}$。

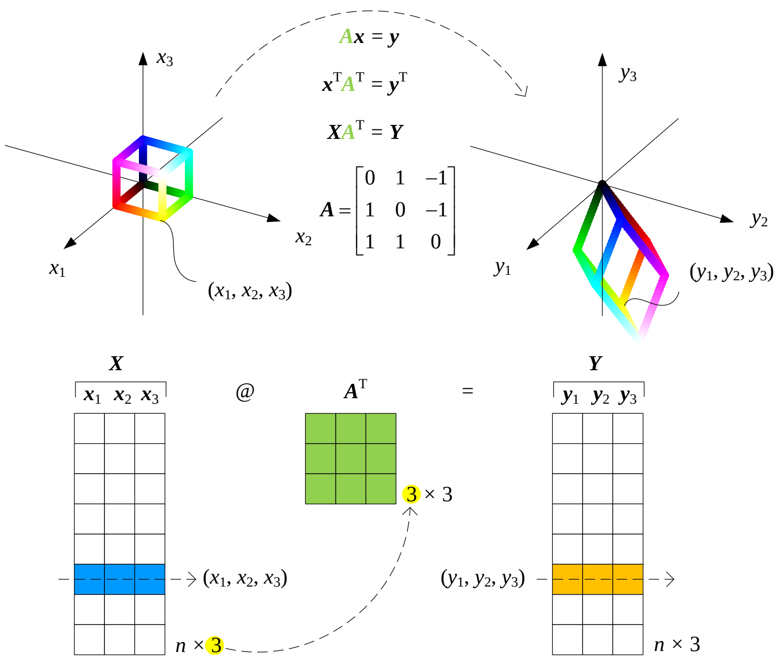


图 12. 矩阵 $\boldsymbol{X}$ 在 3 × 3 方阵 $\boldsymbol{A}$ 映射下结果为矩阵 $\boldsymbol{Y}$

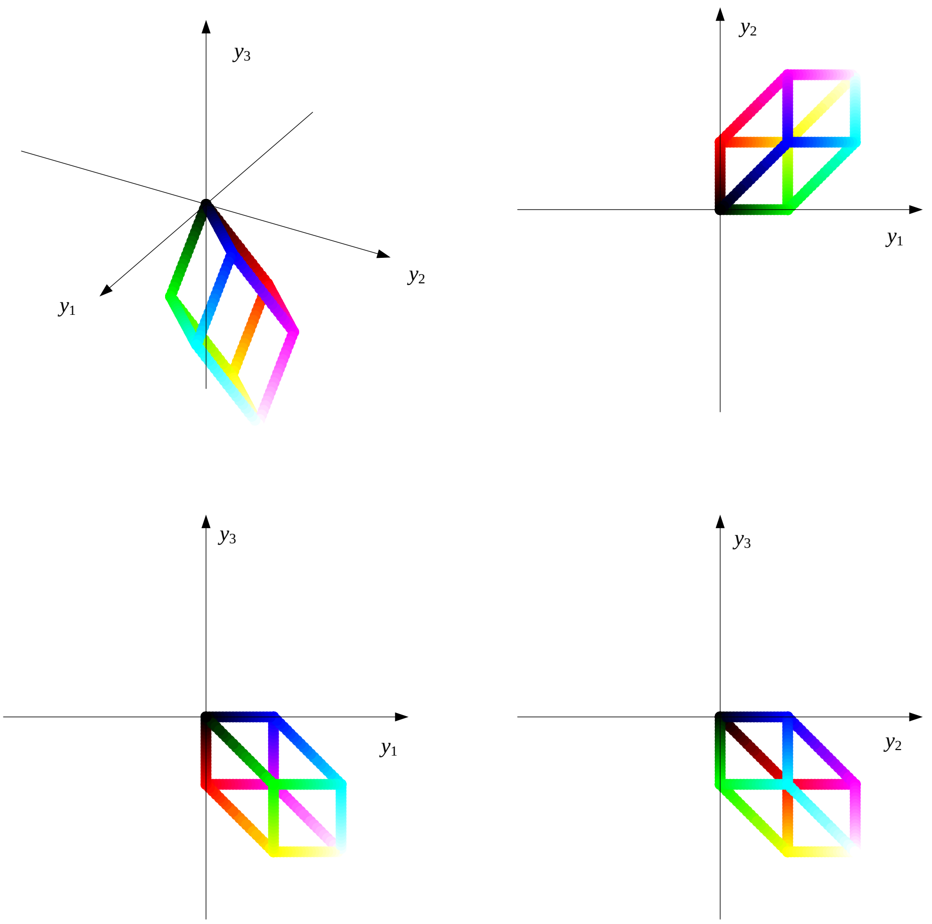
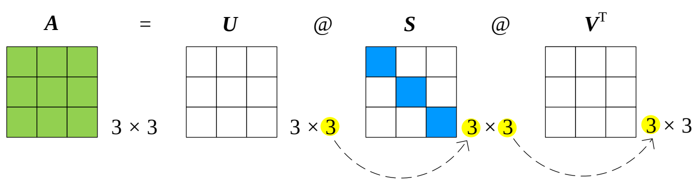


图 13．四个不同视角下的矩阵 $\boldsymbol{Y}$ 形状

SVD 分解

如图 14 所示，一个 3 × 3 方阵 $\boldsymbol{A}$ 经过完全 SVD 分解的结果。



$$\boldsymbol{A}=\begin{bmatrix}0 & 1 & 1\\ 1 & 0 & 1\\ -1 & -1 & 0\end{bmatrix}=\underbrace{\begin{bmatrix}\sqrt{2}/2 & \sqrt{6}/6 & \sqrt{3}/3\\ -\sqrt{2}/2 & \sqrt{6}/6 & \sqrt{3}/3\\ 0 & \sqrt{6}/3 & -\sqrt{3}/3\end{bmatrix}}_{U} @ \underbrace{\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 2\end{bmatrix}}_{S} @ \underbrace{\begin{bmatrix}-\sqrt{2}/2 & \sqrt{2}/2 & 0\\ -\sqrt{6}/6 & -\sqrt{6}/6 & \sqrt{6}/3\\ \sqrt{3}/3 & \sqrt{3}/3 & \sqrt{3}/3\end{bmatrix}}_{V^{\mathrm{T}}}$$

图 14．对 3 × 3 方阵 $\boldsymbol{A}$ 进行奇异值分解

如图 28 所示，3 × 3 方阵 $\boldsymbol{A}=\begin{bmatrix}0 & 1 & 1\\ 1 & 0 & 1\\ -1 & -1 & 0\end{bmatrix}$ 完成的几何操作也可以分解成 3 步。

根据矩阵乘法 $\boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^{\mathrm{T}}\boldsymbol{x}=\boldsymbol{y}$ 先后顺序，$\boldsymbol{V}^{\mathrm{T}}$ 先完成三维空间旋转。

然后，$\boldsymbol{S}$ 完成三维缩放。

最后，$\boldsymbol{U}$ 完成了三维空间的旋转。

第二个例子

图 29 给出了另外一个 3 × 3 方阵 $\boldsymbol{A}$ 对应 SVD 分解的几何操作。请大家根据 BK_2_Ch28_2.ipynb 将矩阵 $\boldsymbol{A}$、$\boldsymbol{U}$、$\boldsymbol{S}$、$\boldsymbol{V}$（$\boldsymbol{V}^{\mathrm{T}}$）的具体值和几何操作写到图 29 上。

28.4 3 × 2 细高矩阵

列向量

类似前文，我们也是先从图 15 中展示的矩阵乘法 $\boldsymbol{Ax} = \boldsymbol{y}$ 说起。

如图 15 所示，矩阵 $\boldsymbol{A}$ 为 3 行 2 列，我们管它叫“细高矩阵”；列向量 $\boldsymbol{x}$ 为 2 行 1 列；结果列向量 $\boldsymbol{y}$ 为 3 行 1 列。从线性代数角度，矩阵 $\boldsymbol{A}$ 完成了 $\boldsymbol{x}$ 到 $\boldsymbol{y}$ 的映射。

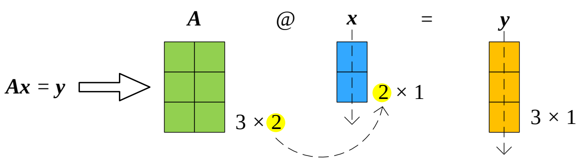


图 15．列向量 $\boldsymbol{x}$ 在细高矩阵 $\boldsymbol{A}$ 映射下结果为列向量 $\boldsymbol{y}$

如图 16 所示，列向量 $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ 相当于平面上一个点 (x_1, x_2)；$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ 也是平面上一个起点位于原点（0,0），终点位于 (x_1, x_2) 的向量。

而列向量 $\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$ 相当于三维直角坐标系上一个点 (y_1, y_2, y_3)；$\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$ 也是三维直角坐标系中一个起点位于原点（0,0,0），终点位于 (y_1, y_2, y_3) 的向量。

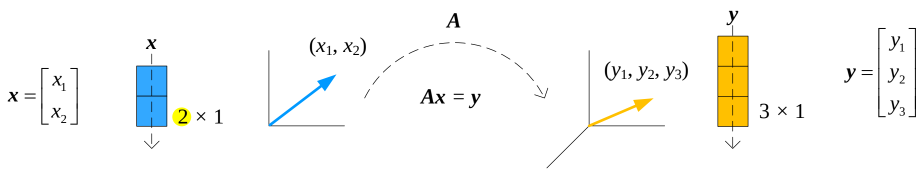


图 16．细高矩阵 $\boldsymbol{A}$ 映射关系

行向量

当然我们也可以把坐标点 (x_1, x_2) 写成行向量形式 $\boldsymbol{x}^{\mathrm{T}} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}$。

而结果 (y_1, y_2, y_3) 则跟着写成 $\boldsymbol{y}^{\mathrm{T}} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} y_1 & y_2 & y_3 \end{bmatrix}$。

如图 17 所示，这时矩阵乘法写成 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{y}^{\mathrm{T}}$。这实际上是矩阵乘法等式 $\boldsymbol{Ax} = \boldsymbol{y}$ 左右转置的结果。大家需要注意顺序调换。

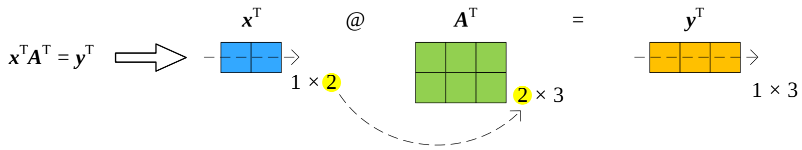


图 17．行向量 x^T 在细高矩阵 A 映射下结果为行向量 y^T

数据矩阵

如图 18 所示，在矩阵 A（A^T）映射下，X 转化成了 Y，即 $XA^T = Y$。注意，我们还是把 X 的每行看做是坐标点；同时，Y 的每行代表一个坐标点。

? 请大家思考图 18 中矩阵 A 对应的几何操作是否可逆？

读到这，有读者可能会问，这和奇异值有什么关系？

答案就在图 18 中！

观察图 18，我们会发现本来在 x_1-x_2 平面中的排列整齐的小彩灯一下子跳到了三维空间中，完成了华丽的“升维”！

但是，仔细观察，我们发现三维空间中的小彩灯似乎还是在一个特殊平面中！

而且原本方方正正的排列，似乎变成了一个菱形？

矩阵 A 到底施了怎样的“魔法”完成了图 18 的几何变换？

奇异值分解就能帮助我们回答这些问题！

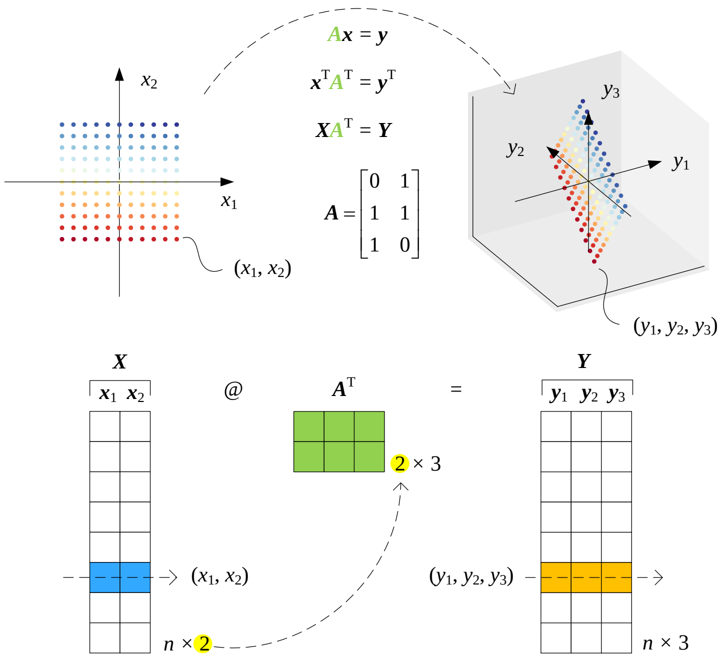
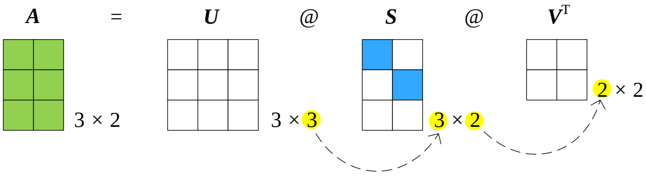


图 18．矩阵 X 在细高矩阵 A 映射下结果为矩阵 Y

SVD 分解

如图 19 所示，一个 3 × 2 的细高矩阵 A 经过完全 SVD 分解得到的结果是三个矩阵的连乘。

图 19 中 U、S、V（V^T）这三个矩阵每个也是都对应一种几何变换！



$$A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix} = \underbrace{\begin{bmatrix} 1/\sqrt{6} & \sqrt{2}/2 & \sqrt{3}/3 \\ 2/\sqrt{6} & 0 & -\sqrt{3}/3 \\ 1/\sqrt{6} & -\sqrt{2}/2 & \sqrt{3}/3 \end{bmatrix}}_{U} @ \underbrace{\begin{bmatrix} \sqrt{3} & \\ & 1 \\ & \end{bmatrix}}_{S} @ \underbrace{\begin{bmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}}_{V^{\mathrm{T}}}$$

图 19. 对细高矩阵 A 进行奇异值分解

如图 30 所示，细高矩阵 $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix}$ 完成的几何操作可以分解成 3 步，即把等式 $Ax = y$ 写成 $USV^{\mathrm{T}}x = y$。

根据矩阵乘法 $USV^{\mathrm{T}}x = y$ 先后顺序，V^{T} 先完成平面旋转。

然后，S 完成缩放 + “升维”。

注意，“升维”打引号是因为数据维度并没有提高。大家查看配套 Jupyter 笔记 BK_2_Ch28_3.ipynb 就会发现经过 S 变换后两列数据变成了三列数据，增加的第三列是一列 0。在图 20 中，我们更容易看到 S 是如何完成缩放这个几何操作的，我们同时看到了小彩灯构成的几何形状确实是“菱形”。

最后，U 完成了三维空间的旋转。这个三维旋转不会改变图形大小，也就是说在三维空间中，这些小彩灯构成的图形还是个“菱形”！

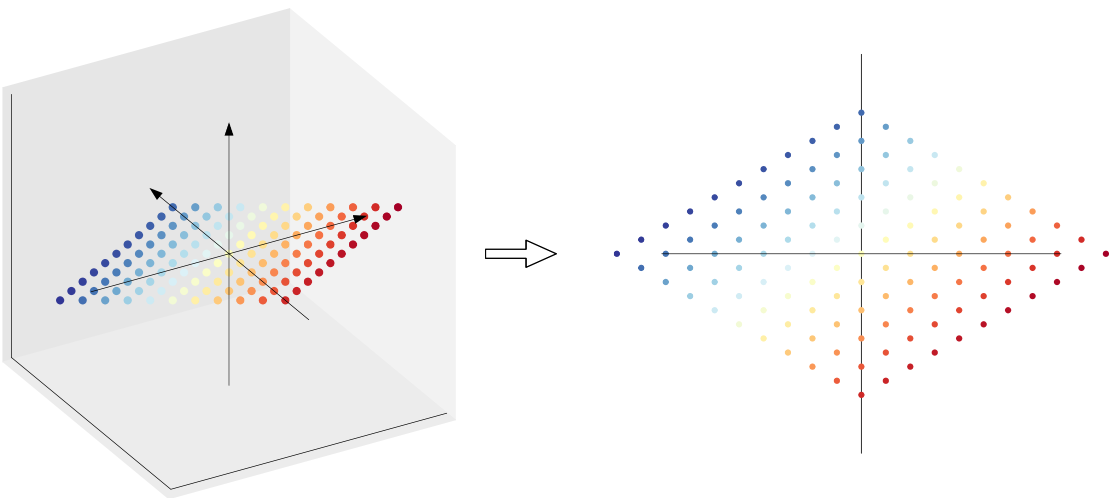


图 20. 增加了一列全 0 的“升维”

第二个例子

图 31 给出了另外一个细高矩阵 A 的几何操作。这个细高矩阵 A 直接将小彩灯排列成三维空间中的一条直线上！

请大家根据 BK_2_Ch28_3.ipynb 将矩阵 A、U、S、V（V^{T}）的具体值自己写到图 31 上。

如果大家之前学过线性代数的话，可以自己算一下这个细高矩阵 A 的秩，并试着从这个角度解释为什么小彩灯会出现在一条直线上。

⚠ 注意，本章仅仅要求大家建立对 U、S、V（V^{T}）的几何直觉，不需要大家掌握这三个矩阵的性质。相关展开讲解内容都在《矩阵力量》。

28.5 2 × 3 矮胖矩阵

有了前文内容，理解图 32 和图 33 就不难了。但是本着“重复 + 精进”这个原则，为了让大家更深刻地理解奇异值分解的几何直觉，我们还是耐着性子按部就班地照着上文思路再分析一遍。

需要强调的是，这两组例子中矩阵的名称完全不变，但是它们完全不同。请大家试着把前文所有有关矩阵形状的记忆全部擦去，重新开始！

列向量

如图 21 所示，“矮胖”矩阵 $\boldsymbol{A}$ 为 2 行 3 列；列向量 $\boldsymbol{x}$ 为 3 行 1 列；结果列向量 $\boldsymbol{y}$ 为 2 行 1 列。从线性代数角度，矮胖矩阵 $\boldsymbol{A}$ 完成了 $\boldsymbol{x}$ 到 $\boldsymbol{y}$ 的映射。

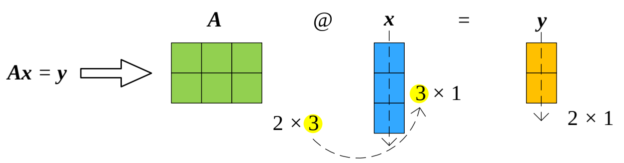


图 21．列向量 $\boldsymbol{x}$ 在矮胖矩阵 $\boldsymbol{A}$ 映射下结果为列向量 $\boldsymbol{y}$

如图 22 所示，列向量 $\boldsymbol{x}=\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$ 相当于三维直角坐标系上一个点 (x_1, x_2, x_3)；$\boldsymbol{x}=\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$ 也是三维直角坐标系一个起点位于原点 $(0,0,0)$，终点位于 (x_1, x_2, x_3) 的向量。

而列向量 $\boldsymbol{y}=\begin{bmatrix}y_1\\y_2\end{bmatrix}$ 相当于平面直角坐标系上一个点 (y_1, y_2)；$\boldsymbol{y}=\begin{bmatrix}y_1\\y_2\end{bmatrix}$ 也是平面直角坐标系中一个起点位于原点 $(0,0)$，终点位于 (y_1, y_2) 的向量。

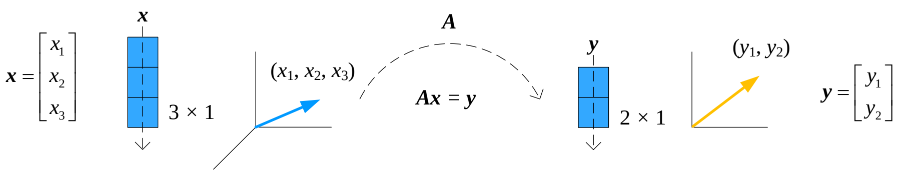


图 22．矮胖矩阵 $\boldsymbol{A}$ 映射关系

行向量

当然我们也可以把坐标点 (x_1, x_2, x_3) 写成行向量形式 $\boldsymbol{x}^{\mathrm{T}}=\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}^{\mathrm{T}}=\begin{bmatrix}x_1 & x_2 & x_3\end{bmatrix}$。

而结果 (y_1, y_2) 则跟着写成 $\boldsymbol{y}^{\mathrm{T}}=\begin{bmatrix}y_1\\y_2\end{bmatrix}^{\mathrm{T}}=\begin{bmatrix}y_1 & y_2\end{bmatrix}$。

如图 23 所示，这时矩阵乘法写成 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{y}^{\mathrm{T}}$。这实际上是矩阵乘法等式 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}$ 左右转置的结果。大家需要注意顺序调换。

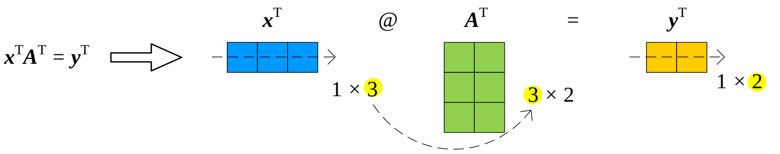


图 23．行向量 $\boldsymbol{x}^{\mathrm{T}}$ 在矮胖矩阵 $\boldsymbol{A}$ 映射下结果为行向量 $\boldsymbol{y}^{\mathrm{T}}$

数据矩阵

如图 24 所示，在矮胖矩阵 $\boldsymbol{A}$ 映射下，$\boldsymbol{X}$ 转化成了 $\boldsymbol{Y}$，即 $\boldsymbol{X}\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{Y}$。注意，我们还是把 $\boldsymbol{X}$ 的每行看做是坐标点；同时，$\boldsymbol{Y}$ 的每行代表一个坐标点。

观察图 24，我们发现本来在三维空间中的 RGB 框线一下子匍匐卧倒在二维平面中，发生了坍塌式的“降维”！

下面，我们就利用奇异值分解来看看到底哪个几何变换环节导致了这次整列垮掉！

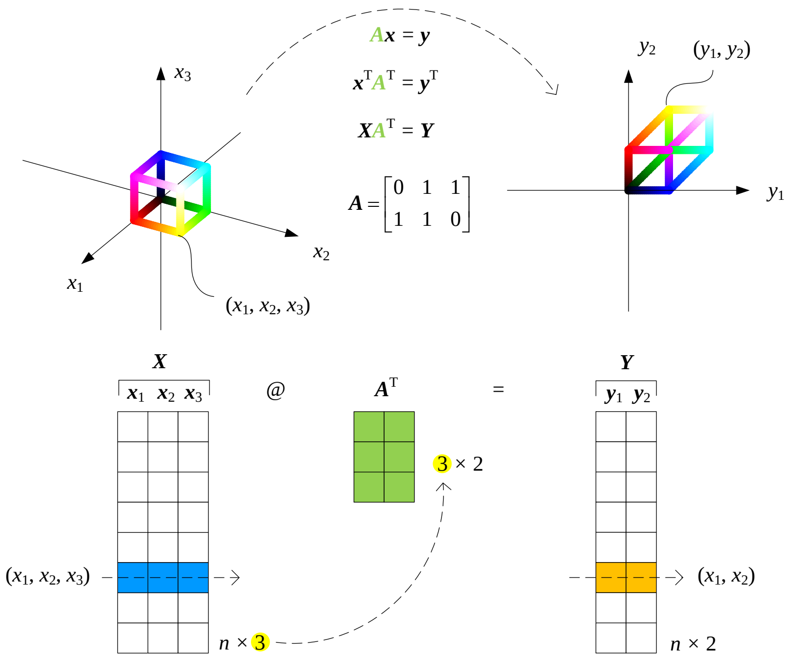
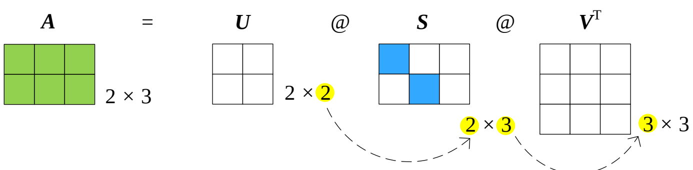


图 24．矩阵 $\boldsymbol{X}$ 在矮胖矩阵 $\boldsymbol{A}$ 映射下结果为矩阵 $\boldsymbol{Y}$

SVD 分解

如图 25 所示，一个 2 × 3 的矮胖矩阵 $\boldsymbol{A}$ 经过完全 SVD 分解得到的结果是三个矩阵的连乘。图 25 中 $\boldsymbol{U}$、$\boldsymbol{S}$、$\boldsymbol{V}$（$\boldsymbol{V}^{\mathrm{T}}$）这三个矩阵每个都对应一种几何变换！再次惊叹！



$$
\boldsymbol{A} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} = \underbrace{\begin{bmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}}_{U} @ \underbrace{\begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}}_{S} @ \underbrace{\begin{bmatrix} 1/\sqrt{6} & 2/\sqrt{6} & 1/\sqrt{6} \\ \sqrt{2}/2 & 0 & -\sqrt{2}/2 \\ \sqrt{3}/3 & -\sqrt{3}/3 & \sqrt{3}/3 \end{bmatrix}}_{V^{\mathrm{T}}}
$$

图 25. 对细高矩阵 $\boldsymbol{A}$ 进行奇异值分解

如图 32 所示，矮胖矩阵 $\boldsymbol{A}$ 完成的几何操作可以分解成 3 步，即把等式 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{y}$ 写成 $\boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^{\mathrm{T}}\boldsymbol{x}=\boldsymbol{y}$。根据矩阵乘法 $\boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^{\mathrm{T}}\boldsymbol{x}=\boldsymbol{y}$ 先后顺序，$\boldsymbol{V}^{\mathrm{T}}$ 先完成三维空间旋转。

然后，$\boldsymbol{S}$ 完成缩放 + "降维"。

最后，$\boldsymbol{U}$ 完成了二维平面的旋转。这个二维旋转不会改变图形大小！

第二个例子

图 33 给出了另外一个矮胖矩阵 $\boldsymbol{A}$ 的几何操作。这个矮胖矩阵 $\boldsymbol{A}$ 直接将 RGB 线框降维到平面的一条直线上！

请大家根据 BK_2_Ch28_4.ipynb 将矩阵 $\boldsymbol{A}$、$\boldsymbol{U}$、$\boldsymbol{S}$、$\boldsymbol{V}$（$\boldsymbol{V}^{\mathrm{T}}$）的具体值自己写到图 33 上。

⚠ 注意，本节有一个矩阵乘法细节希望大家一定要记住——$(\boldsymbol{A}\boldsymbol{B})^{\mathrm{T}}=\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$。

奇异值分解 SVD，可谓是宇宙第一矩阵分解，在线性代数中有着举足轻重的作用。

但凡大家对奇异值分解 SVD 有一点点好奇，推荐大家首先回顾本章前文介绍的平面、三维空间几何变换。如果还觉得不过瘾的话，请大家移步《矩阵力量》！

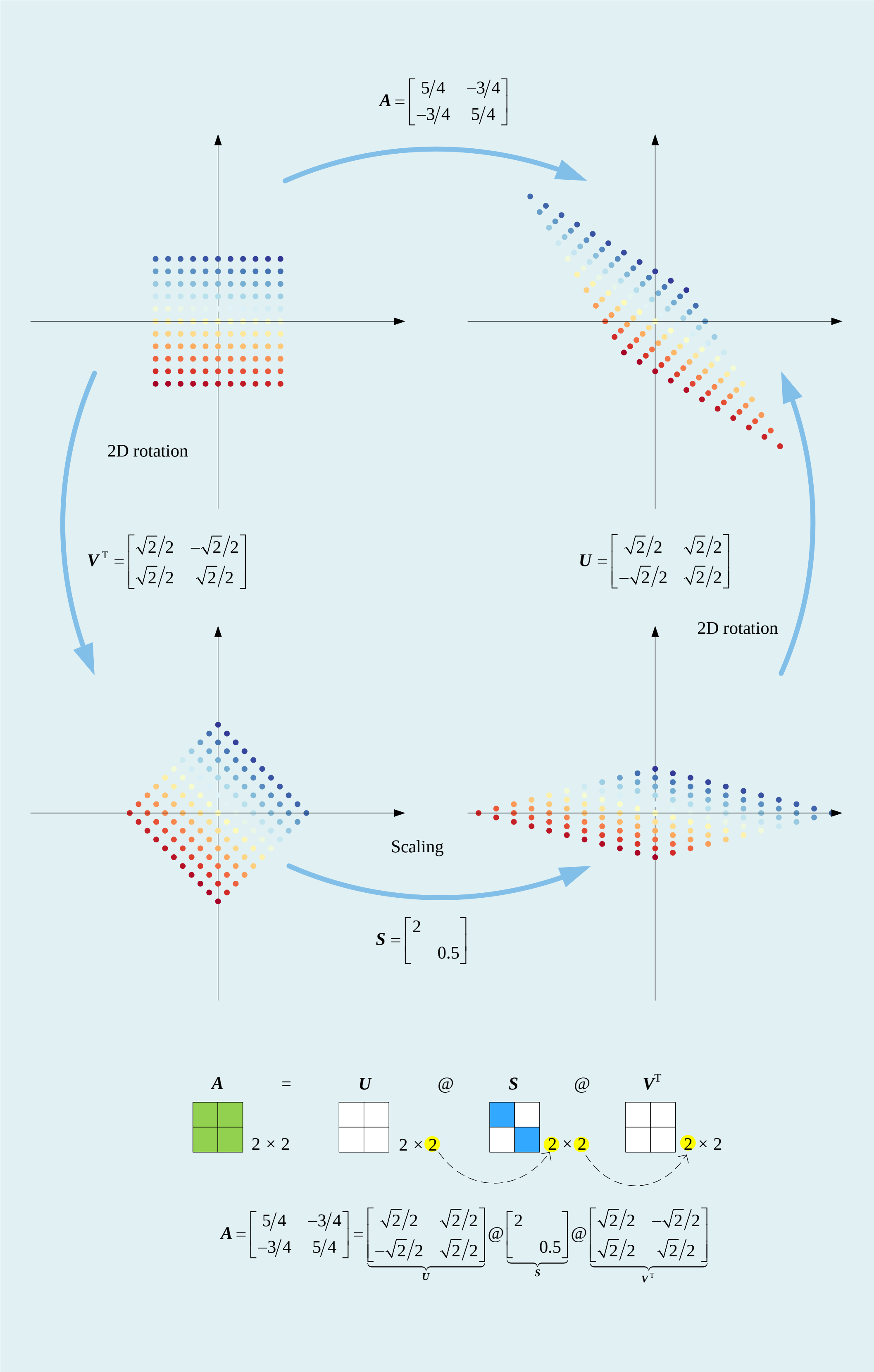


图 26. 可视化奇异值分解，2 × 2 方阵，第 1 组 | BK_2_Ch28_1.ipynb

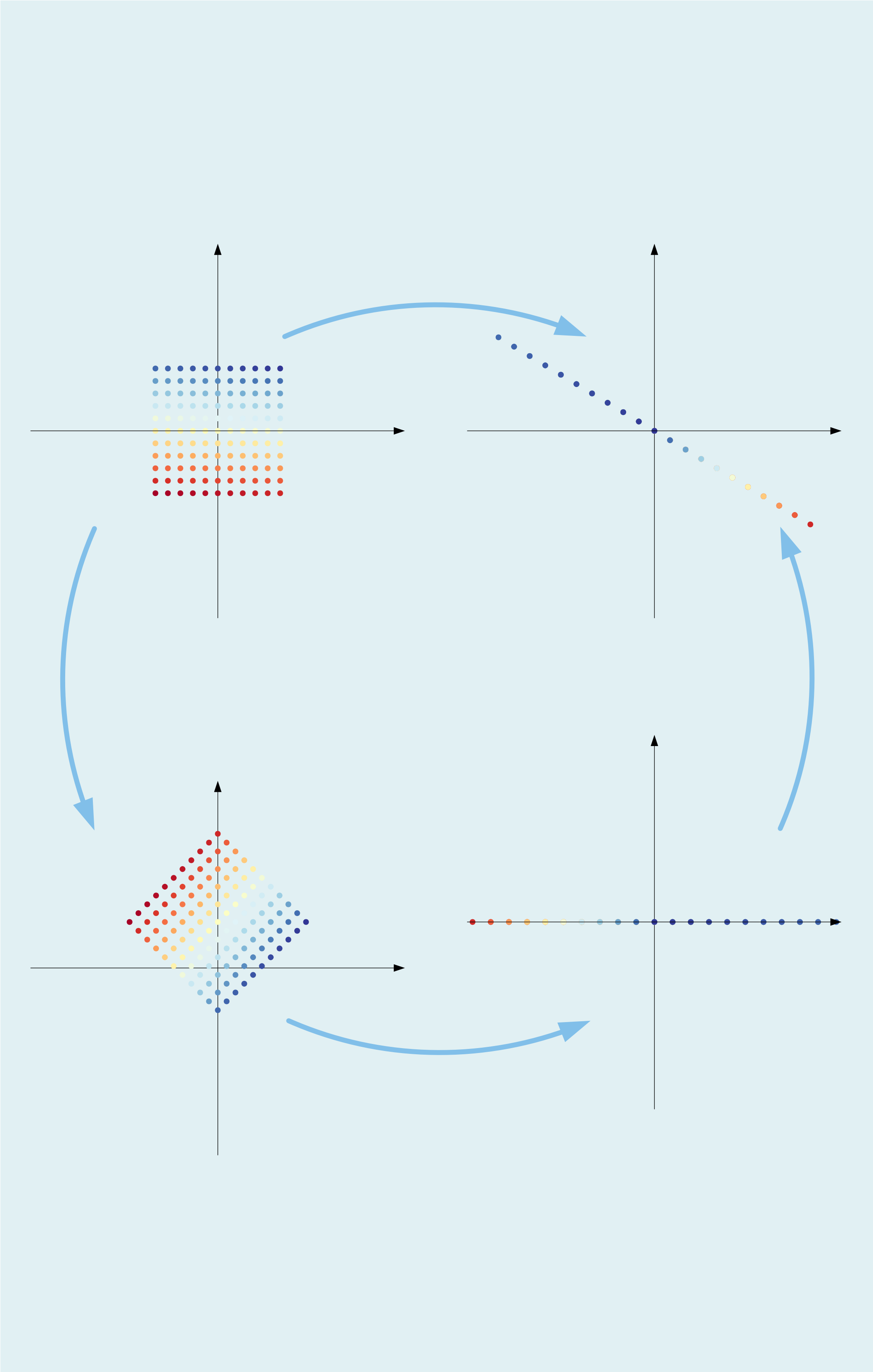


图 27. 可视化奇异值分解，2 × 2 方阵，第 2 组 | BK_2_Ch28_1.ipynb

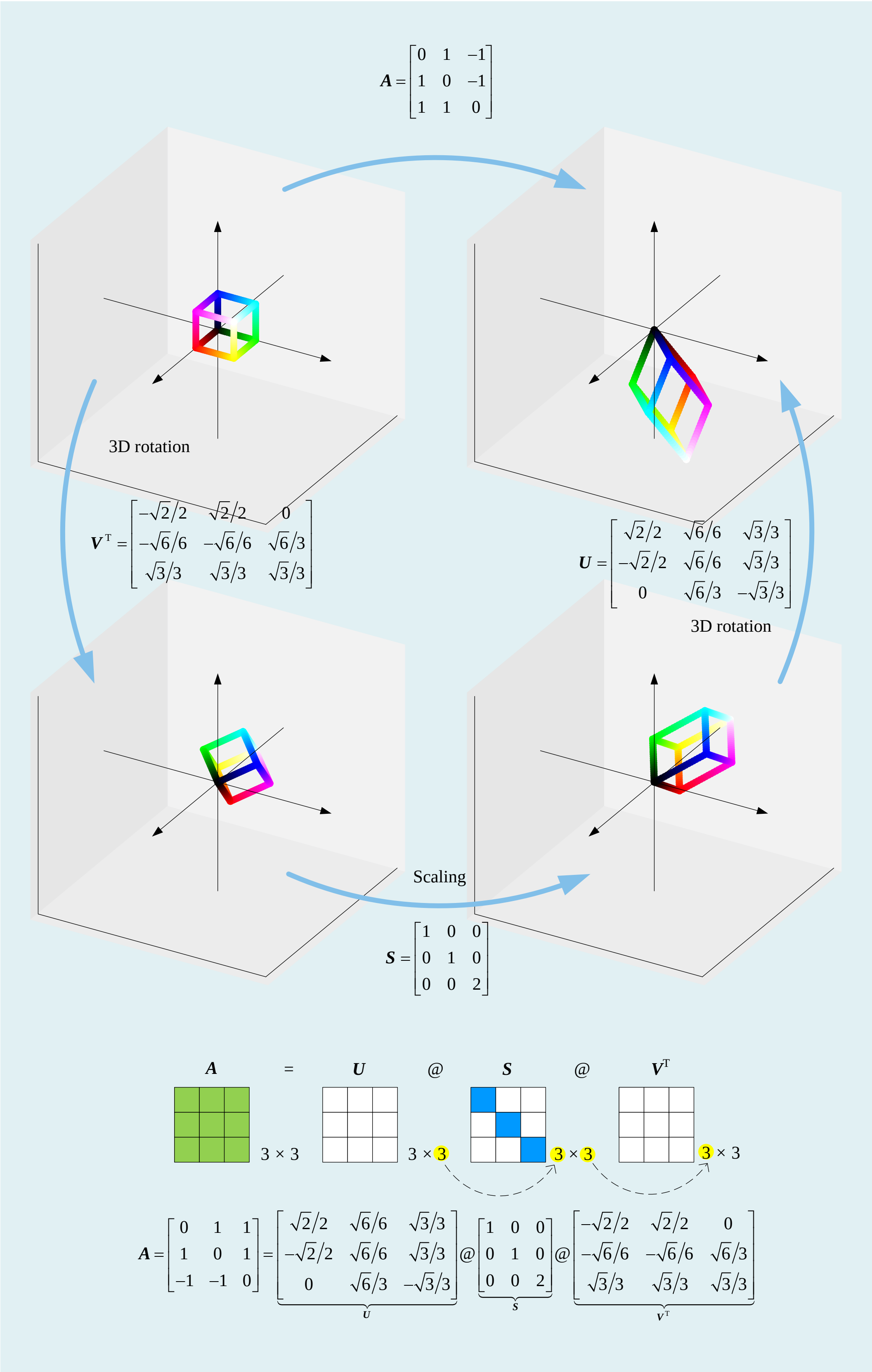


图 28. 可视化奇异值分解，3 × 3 方阵，第 1 组 | BK_2_Ch28_2.ipynb

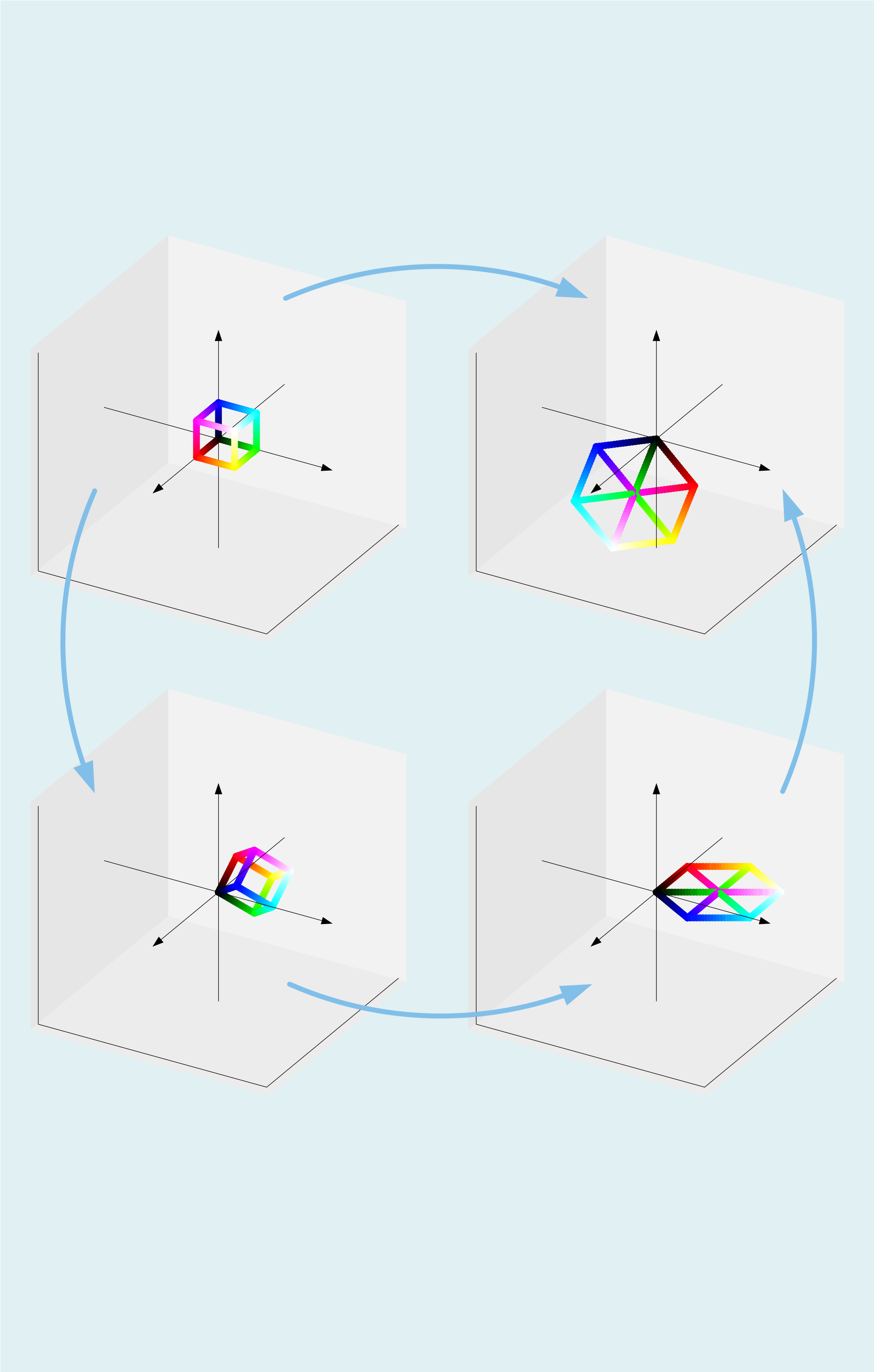


图 29. 可视化奇异值分解，3 × 3 方阵，第 2 组 | BK_2_Ch28_2.ipynb

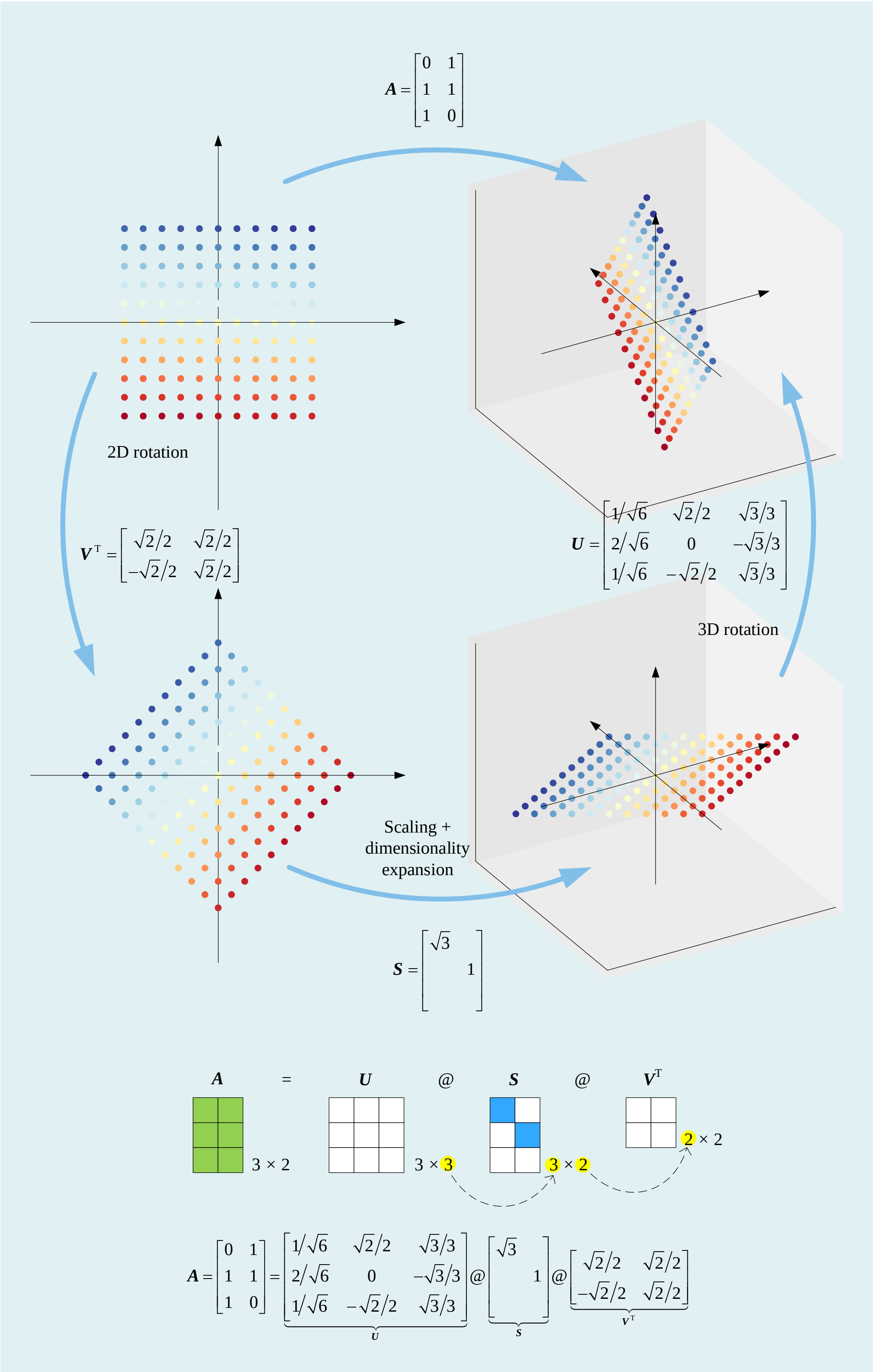


图 30. 可视化奇异值分解，细高矩阵，第 1 组 | BK_2_Ch28_3.ipynb

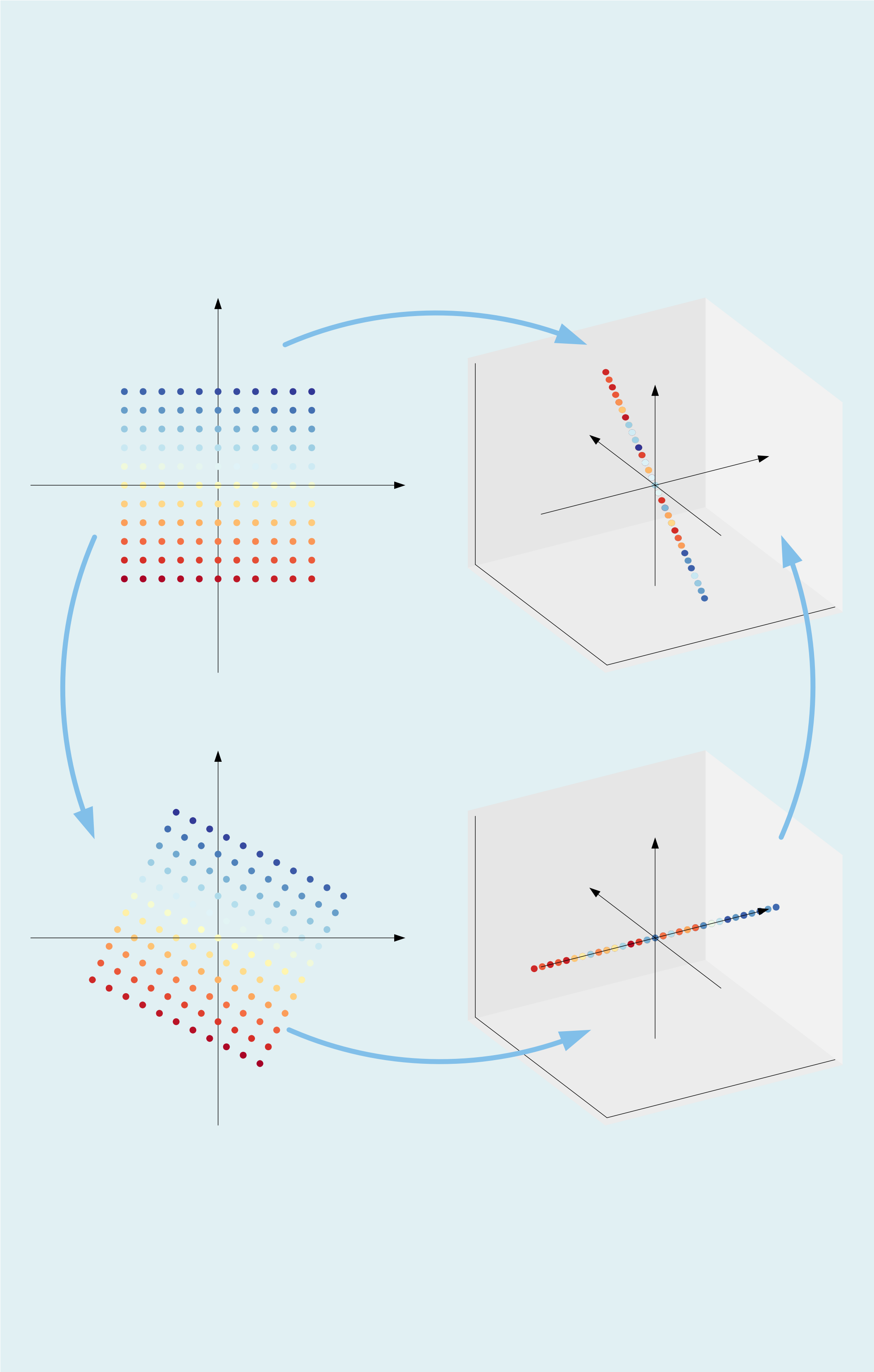


图 31. 可视化奇异值分解，细高矩阵，第 2 组 | BK_2_Ch28_3.ipynb

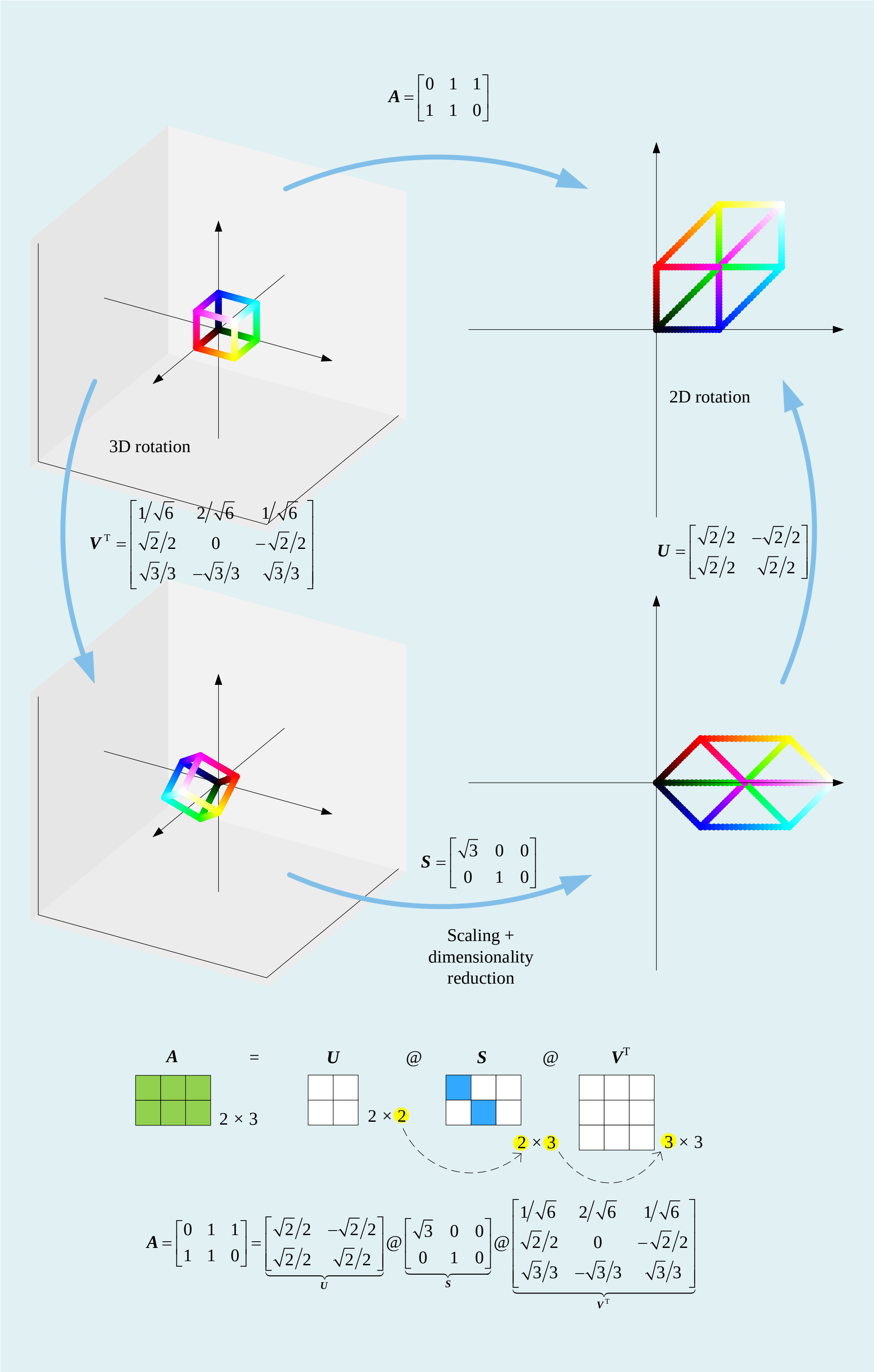


图 32. 可视化奇异值分解，矮胖矩阵，第 1 组 | BK_2_Ch28_4.ipynb

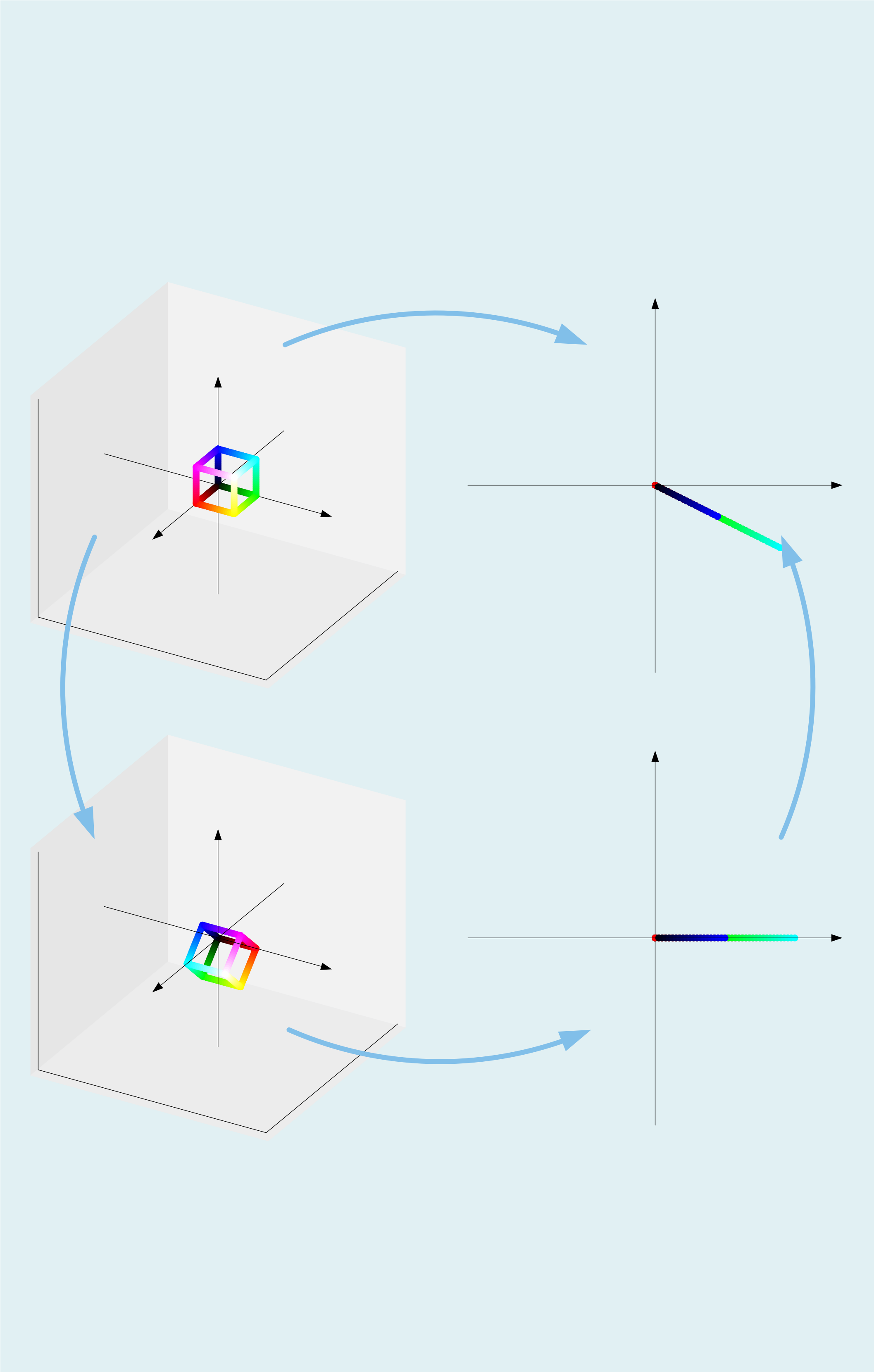


图 33. 可视化奇异值分解，矮胖矩阵，第 2 组 | BK_2_Ch28_4.ipynb