

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

14.3 四种奇异值分解



本节你将掌握的核心技能：

- ▶ 完全型 SVD: U 、 V 均为正交矩阵， S 与原矩阵形状一致。
- ▶ 经济型 SVD: 使 S 变为对角方阵， U 变为半正交矩阵。
- ▶ 紧凑型 SVD: 仅当细高矩阵列不满秩时可将经济型压缩为紧凑型，剔除奇异值为 0 部分。
- ▶ 截断型 SVD: 取前 p 个主成分近似还原原始矩阵，是数据压缩常用手段。

本章第一节提到奇异值分解有四种类型：

- ▶ 完全型 (full);
- ▶ 经济型 (economy-size, thin);
- ▶ 紧凑型 (compact);
- ▶ 截断型 (truncated)。

本节将深入介绍这四种奇异值分解。



阅读本节之前，建议大家回顾第 3 章、第 5 节分块矩阵乘法。

完全型： S 和 X 形状一致

给定一个任意实数矩阵 $X_{n \times D}$ ，经过完全型奇异值分解 (SVD) 得到

$$X_{n \times D} = U_{n \times n} S_{n \times D} V_{D \times D}^T \quad (1)$$

如图 1 所示，完全型奇异值分解中 U 、 V 都是正交矩阵， S 的形状和 X 相同。

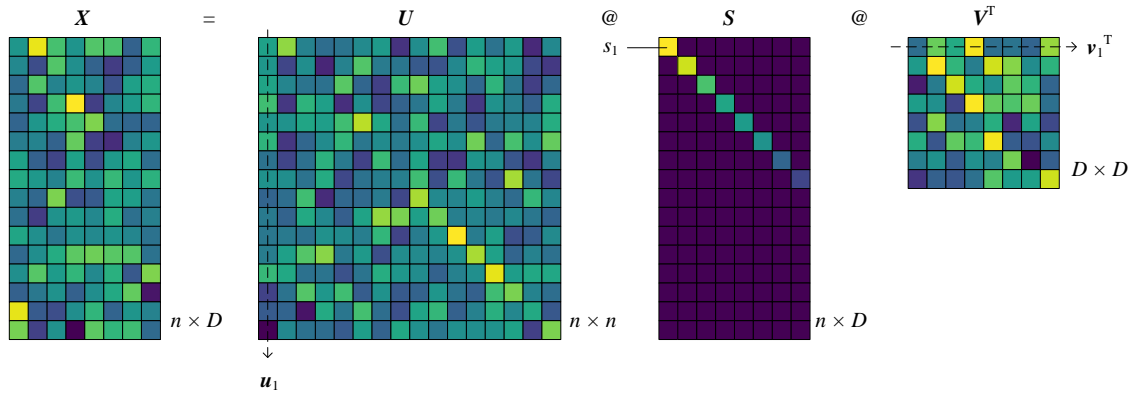


图 1. 完全型 SVD 分解

经济型：S 去掉零矩阵，变方阵

在完全型 SVD 分解基础上，长方对角阵 $S_{n \times D}$ 上下分块为一个对角方阵和一个零矩阵 O ：

$$S_{n \times D} = \begin{bmatrix} s_1 & & & \\ & s_2 & & \\ & & \ddots & \\ & & & s_D \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} = \begin{bmatrix} S_{D \times D} \\ O_{(n-D) \times D} \end{bmatrix} \quad (2)$$

将 $U_{n \times n}$ 写成左右分块矩阵 $[U_{n \times D}, U_{n \times (n-D)}]$ ，其中 $U_{n \times D}$ 和 X 形状相同。

如图 2 所示，利用分块矩阵乘法，完全型 SVD 分解可以简化成经济型 SVD 分解：

$$\begin{aligned} X_{n \times D} &= [U_{n \times D} \quad U_{n \times (n-D)}] \begin{bmatrix} S_{D \times D} \\ O_{(n-D) \times D} \end{bmatrix} V^T \\ &= (U_{n \times D} S_{D \times D} + U_{n \times (n-D)} O_{(n-D) \times D}) V^T \\ &= U_{n \times D} S_{D \times D} V^T \end{aligned} \quad (3)$$

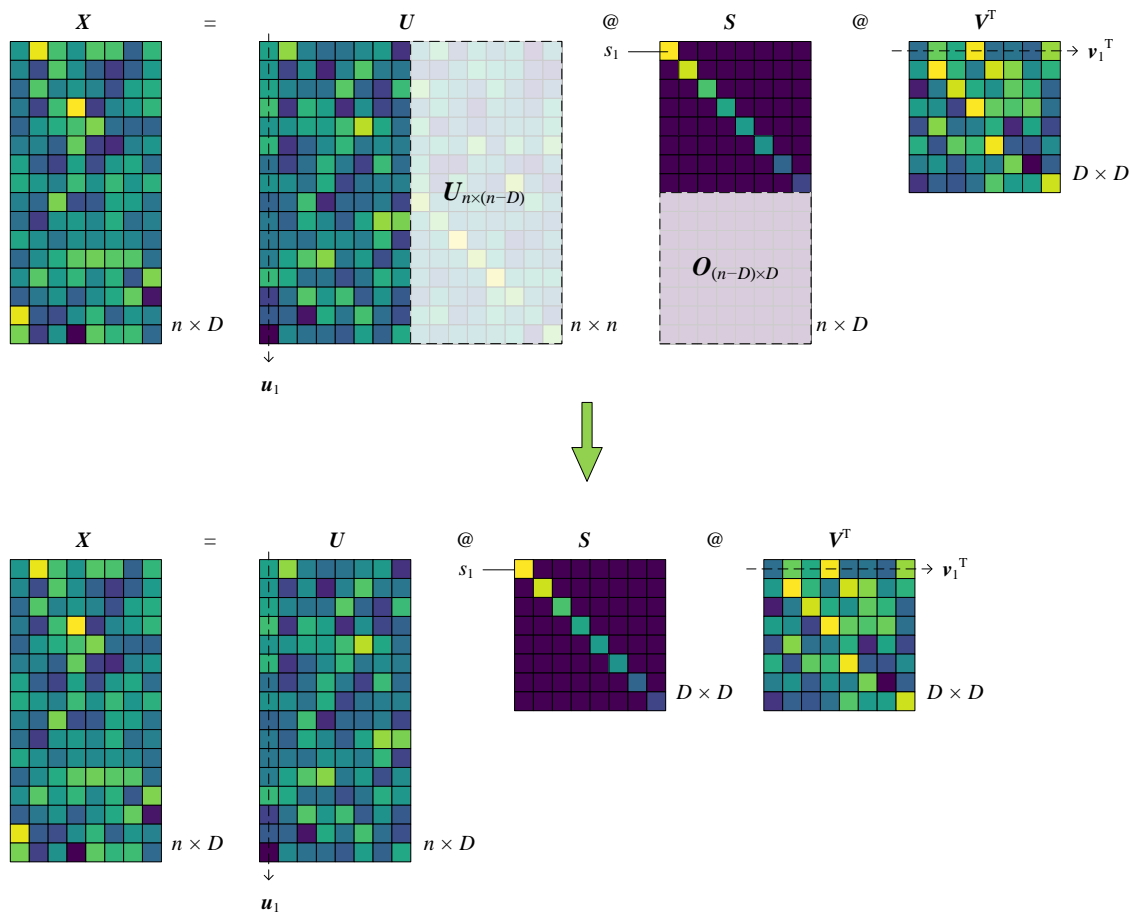
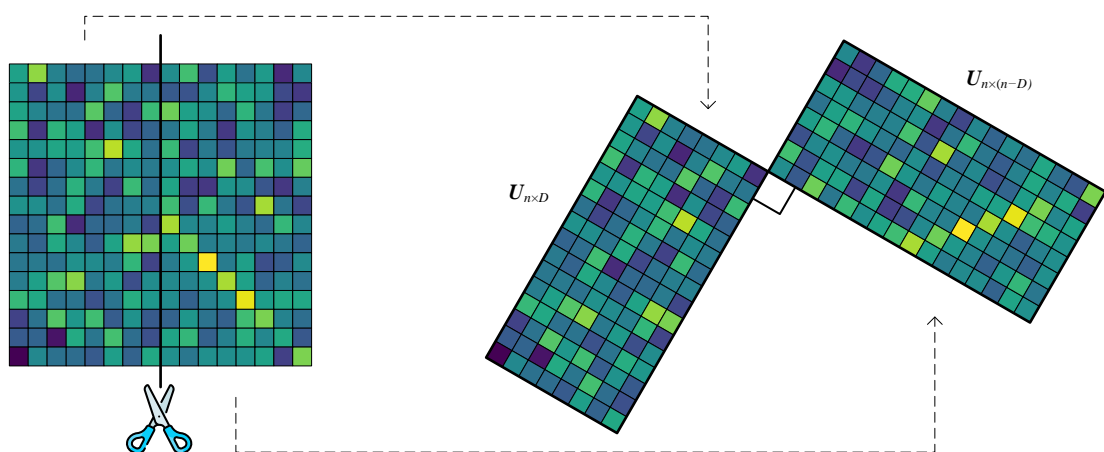


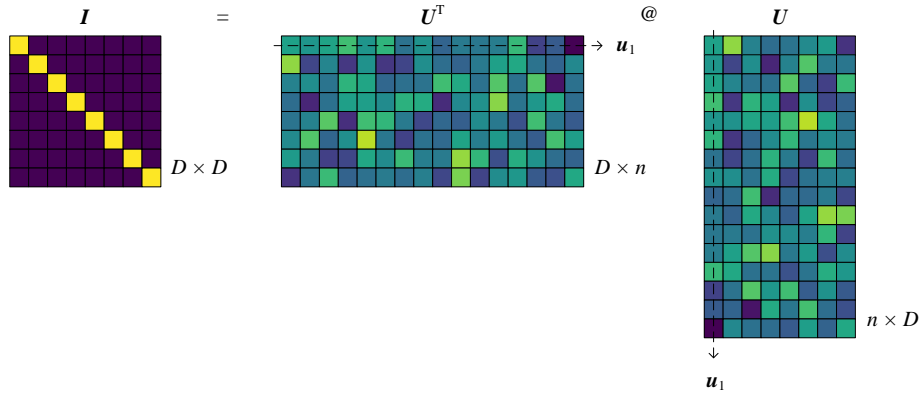
图 2. SVD，从完全型到经济型，X 列满秩

如图 3 所示， $U_{n \times D}$ 和 $U_{n \times (n-D)}$ 列向量张成的空间互为正交补。

图 3. $U_{n \times D}$ 和 $U_{n \times (n-D)}$ 列向量张成的空间互为正交补

$U_{n \times D}$ 不再是正交矩阵，而是半正交矩阵。如图 4 所示， $U_{n \times D}^T$ 和 $U_{n \times D}$ 的矩阵乘法结果仍是单位矩阵，即

$$U_{n \times D}^T U_{n \times D} = \begin{bmatrix} u_1^T \\ u_2^T \\ \vdots \\ u_D^T \end{bmatrix} \begin{bmatrix} u_1 & u_2 & \cdots & u_D \end{bmatrix} = \begin{bmatrix} u_1^T u_1 & u_1^T u_2 & \cdots & u_1^T u_D \\ u_2^T u_1 & u_2^T u_2 & \cdots & u_2^T u_D \\ \vdots & \vdots & \ddots & \vdots \\ u_D^T u_1 & u_D^T u_2 & \cdots & u_D^T u_D \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \quad (4)$$

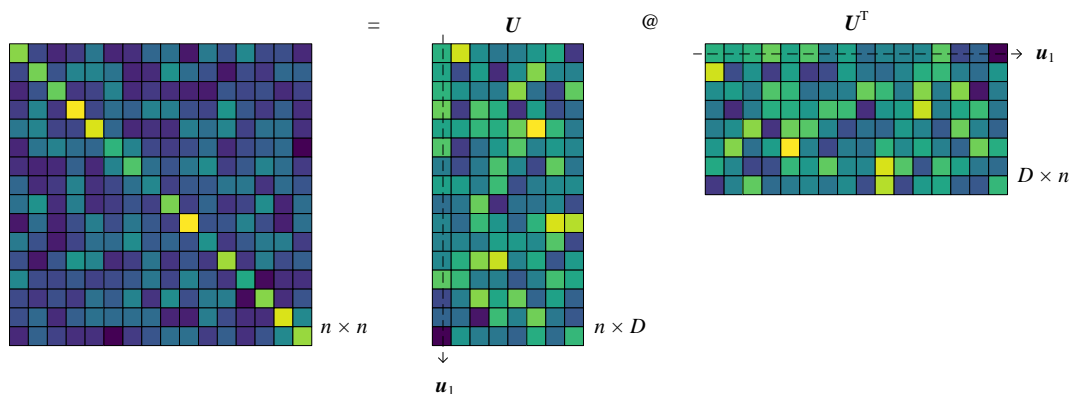
图 4. $U_{n \times D}^T$ 和 $U_{n \times D}$ 的矩阵乘法结果仍是单位矩阵

但是，如图 5 所示， $U_{n \times D}$ 和 $U_{n \times D}^T$ 的乘法，就不再是单位矩阵。将 $U_{n \times D} @ U_{n \times D}^T$ 展开得到

$$U_{n \times D} U_{n \times D}^T = \begin{bmatrix} u_1 & u_2 & \cdots & u_D \end{bmatrix} \begin{bmatrix} u_1^T \\ u_2^T \\ \vdots \\ u_D^T \end{bmatrix} = u_1 u_1^T + u_2 u_2^T + \cdots + u_D u_D^T = I - (u_{D+1} u_{D+1}^T + u_{D+2} u_{D+2}^T + \cdots + u_n u_n^T) \quad (5)$$

? 请大家从正交投影矩阵的角度分析 (5)。

➔ 请大家复习本书第 9 章前三节有关正交矩阵的内容。

图 5. $U_{n \times D}$ 和 $U_{n \times D}^T$ 的乘法，就不再是单位矩阵

比较图 2 中完全型 SVD、经济型 SVD，分解结果中唯一不变的就是矩阵 V ，它一直保持方阵形态。

扁平矩阵

请大家注意，扁平矩阵的完全型 SVD 分解 (图 6)、经济型 SVD 分解 (图 7) 不同于细高矩阵。

? 请大家自行分析图 6、图 7。

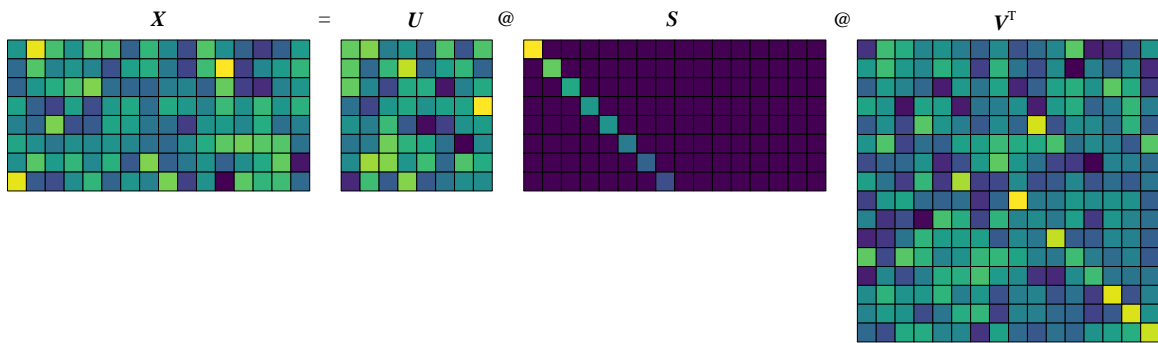


图 6. 完全型 SVD 分解，扁平矩阵

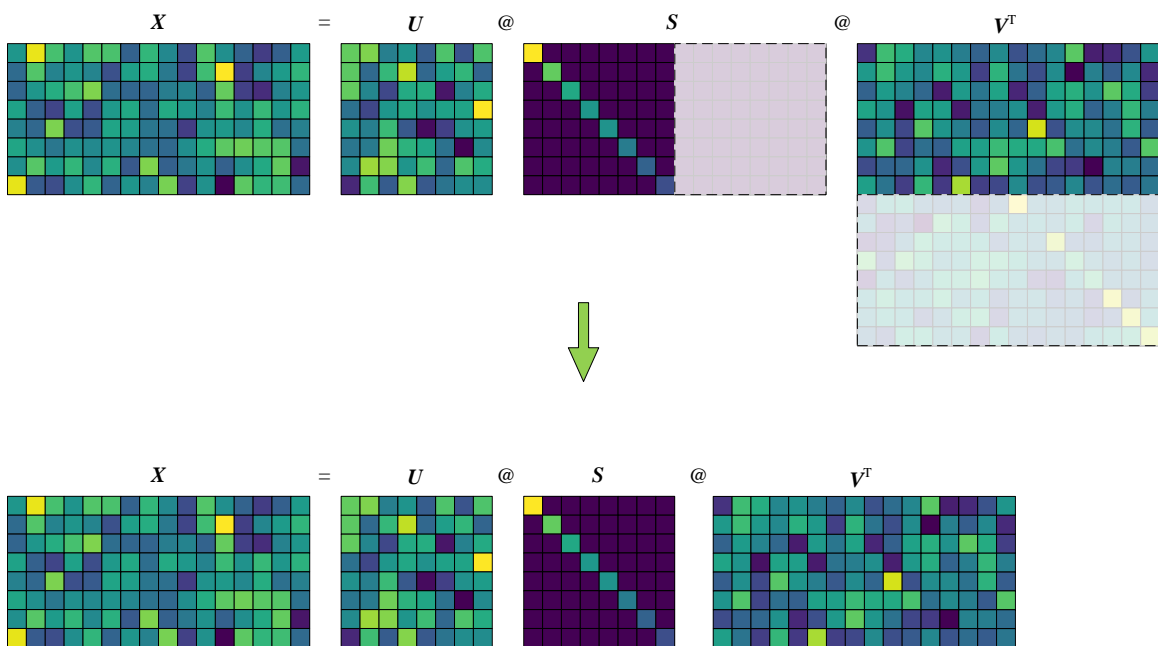


图 7. SVD，从完全型到经济型，扁平矩阵

紧凑型：细高矩阵列非满秩

下面介绍在经济型 SVD 分解基础上获得紧凑型 SVD 分解。

⚠ 注意，仅当细高矩阵列不满秩时才存在紧凑型奇异值分解。

特别地，如果 $\text{rank}(X) = r < D$ ，奇异值 s_j 满足：

$$s_1 \geq s_2 \geq \cdots s_r > 0, \quad s_{r+1} = s_{r+2} = \cdots s_D = 0 \quad (6)$$

这种条件下，经济型 SVD 分解得到的奇异值方阵 S 可以分成四个子块：

$$S = \begin{bmatrix} S_{r \times r} & \mathbf{O}_{r \times (D-r)} \\ \mathbf{O}_{(D-r) \times r} & \mathbf{O}_{(D-r) \times (D-r)} \end{bmatrix} \quad (7)$$

上式中，矩阵 $S_{r \times r}$ 对角线元素奇异值均大于 0。

将 (7) 代入经济型 SVD 分解 (3)，整理得到：

$$\begin{aligned} X_{n \times D} &= \begin{bmatrix} U_{n \times r} & U_{n \times (D-r)} \end{bmatrix} \begin{bmatrix} S_{r \times r} & \mathbf{O}_{r \times (D-r)} \\ \mathbf{O}_{(D-r) \times r} & \mathbf{O}_{(D-r) \times (D-r)} \end{bmatrix} \begin{bmatrix} V_{D \times r} & V_{D \times (D-r)} \end{bmatrix}^T \\ &= \begin{bmatrix} U_{n \times r} S_{r \times r} & \mathbf{O}_{n \times (D-r)} \end{bmatrix} \begin{bmatrix} (V_{D \times r})^T \\ (V_{D \times (D-r)})^T \end{bmatrix} \\ &= U_{n \times r} S_{r \times r} (V_{D \times r})^T \end{aligned} \quad (8)$$

大家特别注意 (8) 中，矩阵 V 先分块后再转置。

⚠ 注意，紧缩型 SVD 分解中， U 和 V 都不是方阵。

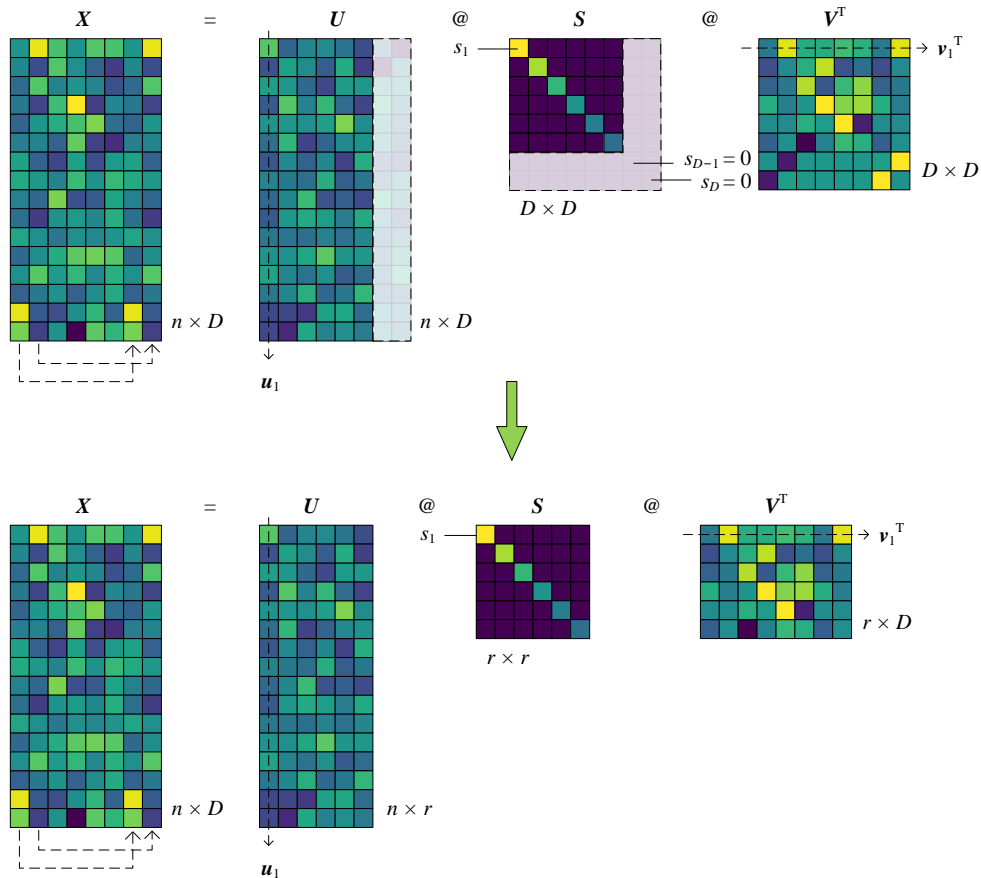


图 8. SVD，从经济型到紧凑型， X 列不满秩

特别值得大家注意的是，一个矩阵奇异值分解后，非零奇异值的数量对应矩阵的列秩 (row rank)、行秩 (column rank)。

矩阵的列秩是矩阵所有列向量最大线性无关个数；而行秩是矩阵所有行向量最大线性无关个数。

图 8 告诉我们， U 的列向量张成的空间等同于 X 列向量张成的空间； V^T 的行向量张成的空间等同于 X 行向量张成的空间。而非零奇异值的个数就决定了空间的维度，即矩阵秩、列秩、行秩。

截断型：近似

如果 $\text{rank}(X) = r \leq D$ ，取经济型奇异值分解中前 p 个奇异值 ($p < r$) 对应的 U 、 S 、 V 矩阵成分，用它们还原原始数据就是截断型奇异值分解：

$$X_{n \times D} \approx \hat{X}_{n \times D} = U_{n \times p} S_{p \times p} (V_{D \times p})^T \quad (9)$$

请大家自行补足上式中矩阵分块和对应的乘法运算。

(9) 不是等号，也就是截断型奇异值分解不能完全还原原始数据。换言之，截断型奇异值分解是对原矩阵 X 的一种近似。

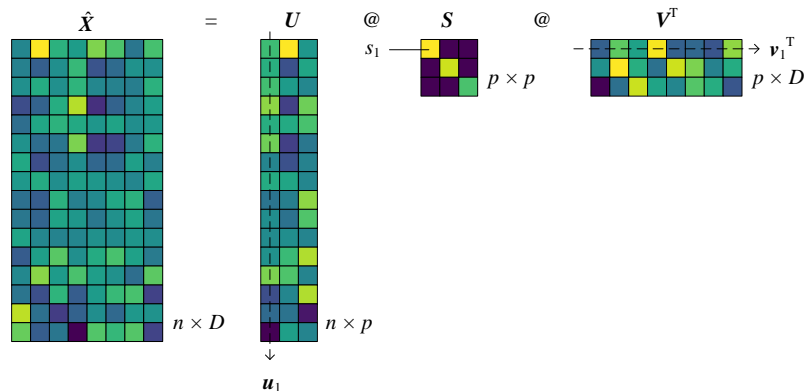


图 9. 采用截断型 SVD 分解近似原始数据

图 9 所示为 SVD 截断型分解热图，可以发现 $X_{n \times D}$ 和 $\hat{X}_{n \times D}$ 两幅热图存在一定“色差”，即如下两个 (形状相同) 矩阵的差

$$X_{\text{error}} = X_{n \times D} - \hat{X}_{n \times D} \quad (10)$$

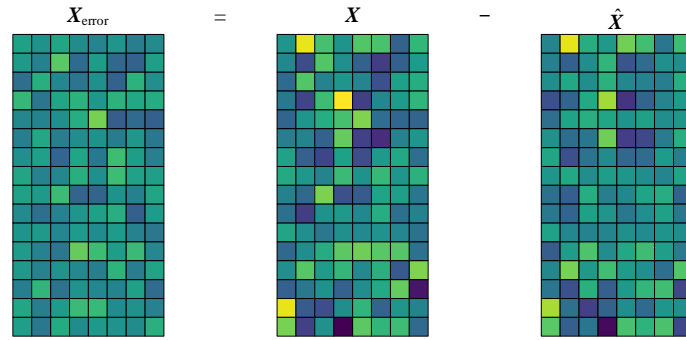


图 10. 误差

误差是哪里来的呢？这是下一节要讨论的话题！



请大家用 DeepSeek/ChatGPT 等工具完成本节如下习题。

Q1. 请对如下矩阵完成完全型奇异值分解，然后再将结果写成经济型奇异值分解。

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix}$$

Q2. 请写成如下矩阵的完全型 SVD、经济型 SVD、紧凑型 SVD。

$$\mathbf{A} = \begin{bmatrix} 0 & 0 \\ 1 & -1 \\ 1 & -1 \end{bmatrix}$$

Q3. 分别计算 **Q1**、**Q2** 矩阵 \mathbf{A} 的行秩、列秩。

Q4. 分别计算 **Q1**、**Q2** 矩阵 \mathbf{A} 格拉姆矩阵 $\mathbf{A}^T\mathbf{A}$ ，以及 $\mathbf{A}^T\mathbf{A}$ 的行秩、列秩。

Q5. 分别计算 **Q1**、**Q2** 矩阵 \mathbf{A}^T 格拉姆矩阵 $\mathbf{A}\mathbf{A}^T$ ，以及 $\mathbf{A}\mathbf{A}^T$ 的行秩、列秩。