

# The Art of Linear Algebra

– Graphic Notes on “Linear Algebra for Everyone” –

Kenji Hiranabe \*

with the kindest help of Gilbert Strang †

translator: Kefang Liu ‡

September 1, 2021/updated November 3, 2022

## Abstract

我尝试为 Gilbert Strang 在书籍 “Linear Algebra for Everyone” 中介绍的矩阵的重要概念进行可视化图释, 以促进从矩阵分解的角度对向量、矩阵计算和算法的理解.<sup>1</sup> 它们包括列-行 (Column-Row,  $\mathbf{CR}$ )、高斯消除 (Gaussian Elimination,  $\mathbf{LU}$ )、施密特正交化 (Gram-Schmidt Orthogonalization,  $\mathbf{QR}$ )、特征值和对角化 (Eigenvalues and Diagonalization,  $\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T$ )、和奇异值分解 (Singular Value Decomposition,  $\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$ ).

## 序言

我很高兴能看到 Kenji Hiranabe 的线性代数中的矩阵运算的图片! 这样的图片是展示代数的绝佳方式. 我们当然可以通过行·列的点乘来想象矩阵乘法, 但那绝非全部——它是“线性组合”与“秩 1 矩阵”组成的代数与艺术. 我很感激能看到日文翻译的书籍和 Kenji 的图片中的想法.

– Gilbert Strang  
麻省理工学院数学教授

## Contents

1 理解矩阵——4 个视角	2
2 向量乘以向量——2 个视角	2
3 矩阵乘以向量——2 个视角	3
4 矩阵乘以矩阵——4 个视角	4
5 实用模式	4
6 The Five Factorizations of a Matrix	7
6.1 $\mathbf{A} = \mathbf{CR}$	7
6.2 $\mathbf{A} = \mathbf{LU}$	8
6.3 $\mathbf{A} = \mathbf{QR}$	8
6.4 $\mathbf{S} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T$	9
6.5 $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$	10

\*twitter: @hiranabe, k-hiranabe@esm.co.jp, <https://anagileway.com>

†Massachusetts Institute of Technology, <http://www-math.mit.edu/~gs/>

‡中文版翻译: <https://github.com/kf-liu/The-Art-of-Linear-Algebra-zh-CN>

<sup>1</sup>“Linear Algebra for Everyone”: <http://math.mit.edu/everyone/>.

## 1 理解矩阵——4 个视角

一个矩阵 ( $m \times n$ ) 可以被视为 1 个矩阵,  $mn$  个数,  $n$  个列和  $m$  个行.

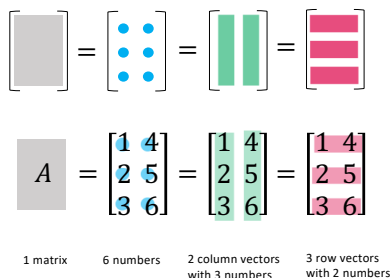


Figure 1: 从四个角度理解矩阵

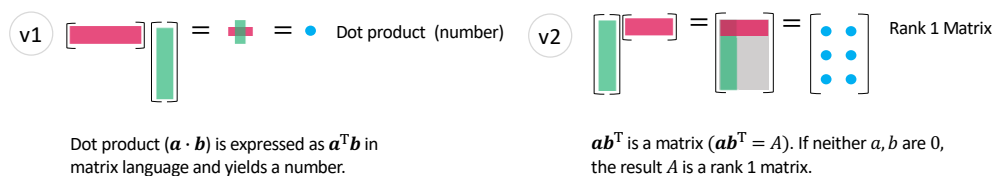
$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} = \begin{bmatrix} | & | \\ \mathbf{a}_1 & \mathbf{a}_2 \\ | & | \end{bmatrix} = \begin{bmatrix} -\mathbf{a}_1^* - \\ -\mathbf{a}_2^* - \\ -\mathbf{a}_3^* - \end{bmatrix}$$

在这里, 列向量被标记为粗体  $\mathbf{a}_1$ . 行向量则有一个 \* 号, 标记为  $\mathbf{a}_1^*$ . 转置向量和矩阵则用 T 标记为  $\mathbf{a}^T$  和  $A^T$ .

## 2 向量乘以向量——2 个视角

后文中, 我将介绍一些概念, 同时列出 “Linear Algebra for Everyone” 一书中的相应部分 (部分编号插入如下). 详细的内容最好看书, 这里我也添加了一个简短的解释, 以便您可以通过这篇文章尽可能多地理解. 此外, 每个图都有一个简短的名称, 例如 v1 (数字 1 表示向量的乘积)、Mv1 (数字 1 表示矩阵和向量的乘积), 以及如下图 (v1) 所示的彩色圆圈. 如你所见, 随着讨论的进行, 该名称将被交叉引用.

- 1.1 节 (p.2) 线性组合与点积
- 1.3 节 (p.25) 矩阵与秩 1
- 1.4 节 (p.29) 行方式与列方式



$$\begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 + 2x_2 + 3x_3$$

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \begin{bmatrix} x & y \end{bmatrix} = \begin{bmatrix} x & y \\ 2x & 2y \\ 3x & 3y \end{bmatrix}$$

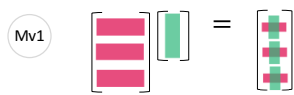
Figure 2: 向量乘以向量 - (v1), (v2)

(v1) 是两个向量之间的基础运算, 而 (v2) 将列乘以行并产生一个秩 1 矩阵. 理解 (v2) 的结果 (秩 1) 是接下来章节的关键.

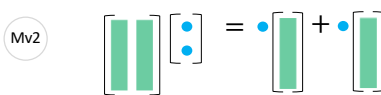
### 3 矩阵乘以向量——2 个视角

一个矩阵乘以一个向量将产生三个点积组成的向量 (Mv1) 和一种  $A$  的列向量的线性组合.

- 1.1 节 (p.3) 线性组合
- 1.3 节 (p.21) 矩阵与列空间



The row vectors of  $A$  are multiplied by a vector  $\mathbf{x}$  and become the three dot-product elements of  $A\mathbf{x}$ .



The product  $A\mathbf{x}$  is a linear combination of the column vectors of  $A$ .

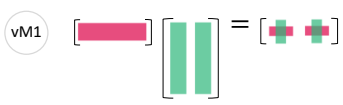
$$A\mathbf{x} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} (x_1 + 2x_2) \\ (3x_1 + 4x_2) \\ (5x_1 + 6x_2) \end{bmatrix}$$

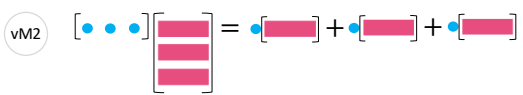
$$A\mathbf{x} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}$$

Figure 3: 矩阵乘以向量- (Mv1), (Mv2)

往往你会先学习 (Mv1). 但当你习惯了从 (Mv2) 的视角看待它, 会理解  $A\mathbf{x}$  是  $A$  的列的线性组合. 矩阵  $A$  的列向量的所有线性组合生成的子空间记为  $\mathbf{C}(A)$ .  $A\mathbf{x} = \mathbf{0}$  的解空间则是零空间, 记为  $\mathbf{N}(A)$ .

同理, 由 (vM1) 和 (vM2) 可见, 行向量乘以矩阵也是同一种理解方式.





$$\mathbf{y}A = [y_1 \quad y_2 \quad y_3] \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} = [(y_1 + 3y_2 + 5y_3) \quad (2y_1 + 4y_2 + 6y_3)]$$

A row vector  $\mathbf{y}$  is multiplied by the two column vectors of  $A$  and become the two dot-product elements of  $\mathbf{y}A$ .

$$\mathbf{y}A = [y_1 \quad y_2 \quad y_3] \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} = y_1[1 \quad 2] + y_2[3 \quad 4] + y_3[5 \quad 6]$$

The product  $\mathbf{y}A$  is a linear combination of the row vectors of  $A$ .

Figure 4: 向量乘以矩阵 - (vM1), (vM2)

上图  $A$  的行向量的所有线性组合生成的子空间记为  $\mathbf{C}(A^T)$ .  $\mathbf{y}A = \mathbf{0}$  的解空间是  $A$  的左零空间, 记为  $\mathbf{N}(A^T)$ .

本书的一大亮点即为四个基本子空间: 在  $\mathbb{R}^n$  上的  $\mathbf{N}(A) + \mathbf{C}(A^T)$  (相互正交) 和在  $\mathbb{R}^m$  上的  $\mathbf{N}(A^T) + \mathbf{C}(A)$  (相互正交).

- 3.5 节 (p.124) 四个子空间的维度

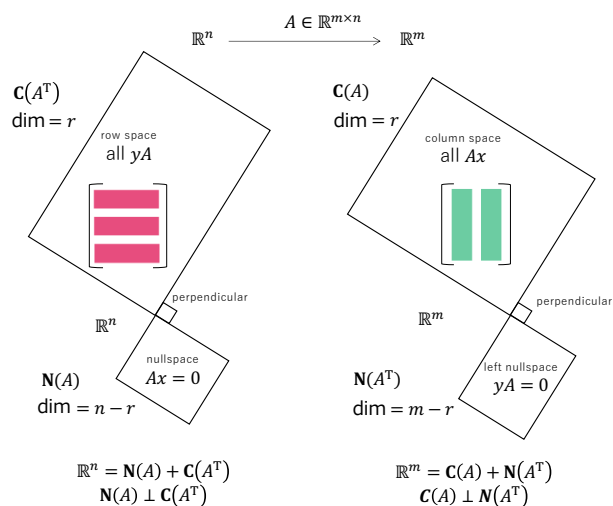


Figure 5: 四个子空间

关于秩  $r$ , 请见  $A = CR$  (6.1 节) .

## 4 矩阵乘以矩阵——4 个视角

由“矩阵乘以向量”自然延伸到“矩阵乘以矩阵”.

- 1.4 节 (p.35)  $AB = C$  的四种乘法
- 也可以见书的封底

MM 1

Every element becomes a dot product of row vector and column vector.

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \end{bmatrix} = \begin{bmatrix} (x_1+2x_2) & (y_1+2y_2) \\ (3x_1+4x_2) & (3y_1+4y_2) \\ (5x_1+6x_2) & (5y_1+6y_2) \end{bmatrix}$$

MM 2

$Ax$  and  $Ay$  are linear combinations of columns of  $A$ .

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \end{bmatrix} = A \begin{bmatrix} x & y \end{bmatrix} = [Ax \quad Ay]$$

MM 3

The produced rows are linear combinations of rows.

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \end{bmatrix} = \begin{bmatrix} a_1^* \\ a_2^* \\ a_3^* \end{bmatrix} X = \begin{bmatrix} a_1^* X \\ a_2^* X \\ a_3^* X \end{bmatrix}$$

MM 4

Multiplication  $AB$  is broken down to a sum of rank 1 matrices.

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_1 & a_2 \end{bmatrix} \begin{bmatrix} b_1^* \\ b_2^* \end{bmatrix} = a_1 b_1^* + a_2 b_2^* \\ = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \end{bmatrix} + \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix} \begin{bmatrix} b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} \\ 3b_{11} & 3b_{12} \\ 5b_{11} & 5b_{12} \end{bmatrix} + \begin{bmatrix} 2b_{21} & 2b_{22} \\ 4b_{21} & 4b_{22} \\ 6b_{21} & 6b_{22} \end{bmatrix}$$

Figure 6: 矩阵乘以矩阵 - (MM1), (MM2), (MM3), (MM4)

## 5 实用模式

在这里, 我展示了一些实用的模式, 可以让你更直观地理解接下来的内容。

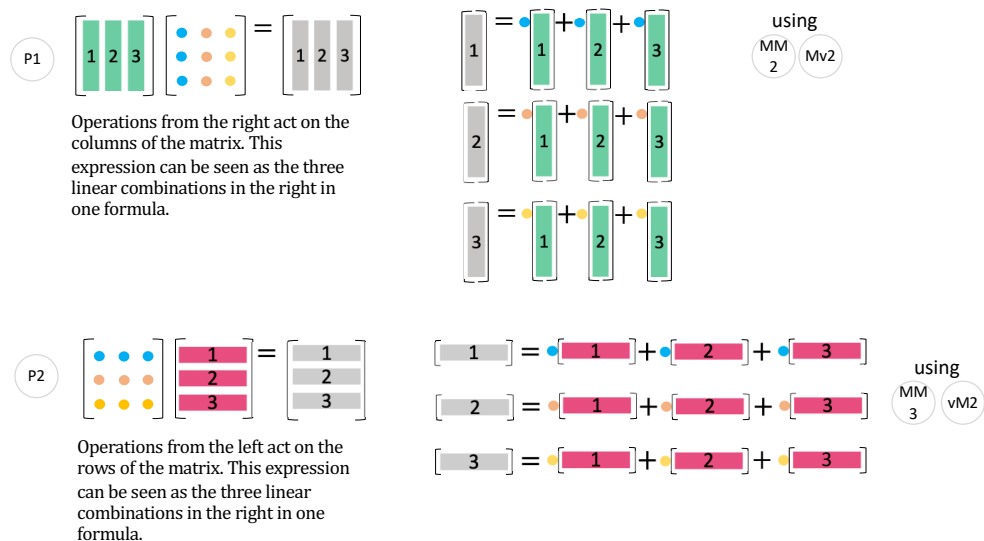


Figure 7: 图 1, 2 - (P1), (P1)

P1 是 (MM2) 和 (Mv2) 的结合. P2 是 (MM3) 和 (vM2) 的扩展. 注意, P1 是列运算 (右乘一个矩阵), 而 P2 是行运算 (左乘一个矩阵).

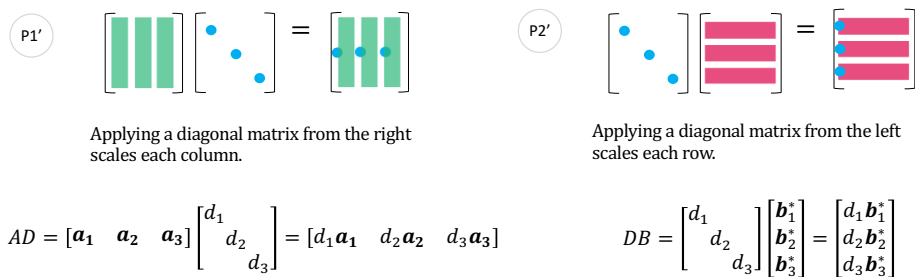


Figure 8: 图 1', 2' - (P1'), (P2')

(P1') 将对角线上的数乘以矩阵的列, 而 (P2') 将对角线上的数乘以矩阵的行. 两个分别为 (P1) 和 (P2) 的变体.

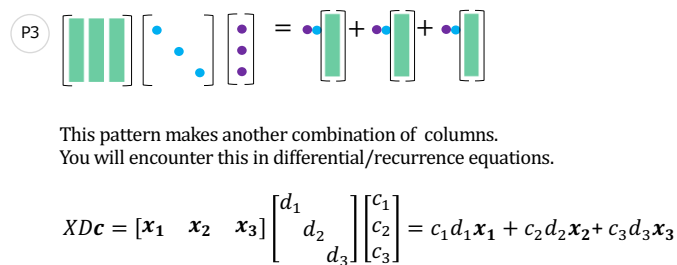


Figure 9: 图 3 - (P3)

当解决微分方程和递归方程时的也会出现这一模式:

- 6 节 (p.201) 特征值和特征向量
- 6.4 节 (p.243) 微分方程组

$$\begin{aligned}\frac{d\mathbf{u}(t)}{dt} &= A\mathbf{u}(t), \quad \mathbf{u}(0) = \mathbf{u}_0 \\ \mathbf{u}_{n+1} &= A\mathbf{u}_n, \quad \mathbf{u}_0 = \mathbf{u}_0\end{aligned}$$

In both cases, the solutions are expressed with eigenvalues  $(\lambda_1, \lambda_2, \lambda_3)$ , eigenvectors  $X = [\mathbf{x}_1 \ \mathbf{x}_2 \ \mathbf{x}_3]$  of  $A$ , and the coefficients  $\mathbf{c} = [c_1 \ c_2 \ c_3]^T$  which are the coordinates of the initial condition  $\mathbf{u}(0) = \mathbf{u}_0$  in terms of the eigenvectors  $X$ .

$$\begin{aligned}\mathbf{u}_0 &= c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + c_3\mathbf{x}_3 \\ \mathbf{c} &= \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = X^{-1}\mathbf{u}_0\end{aligned}$$

and the general solution of the two equations are:

$$\begin{aligned}\mathbf{u}(t) &= e^{At}\mathbf{u}_0 = Xe^{\Lambda t}X^{-1}\mathbf{u}_0 &= Xe^{\Lambda t}\mathbf{c} = c_1e^{\lambda_1 t}\mathbf{x}_1 + c_2e^{\lambda_2 t}\mathbf{x}_2 + c_3e^{\lambda_3 t}\mathbf{x}_3 \\ \mathbf{u}_n &= A^n\mathbf{u}_0 = X\Lambda^nX^{-1}\mathbf{u}_0 &= X\Lambda^n\mathbf{c} = c_1\lambda_1^n\mathbf{x}_1 + c_2\lambda_2^n\mathbf{x}_2 + c_3\lambda_3^n\mathbf{x}_3\end{aligned}$$

See Figure 8: Pattern 3 (P3) above again for  $XDc$ .

P4

A matrix is broken down to a sum of rank 1 matrices,  
as in singular value/eigenvalue decomposition.

$$U\Sigma V^T = [\mathbf{u}_1 \ \mathbf{u}_2 \ \mathbf{u}_3] \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \sigma_3 \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \mathbf{v}_3^T \end{bmatrix} = \sigma_1\mathbf{u}_1\mathbf{v}_1^T + \sigma_2\mathbf{u}_2\mathbf{v}_2^T + \sigma_3\mathbf{u}_3\mathbf{v}_3^T$$

Figure 10: Pattern 4 - (P4)

This pattern (P4) works in both eigenvalue decomposition and singular value decomposition. Both decompositions are expressed as a product of three matrices with a diagonal matrix in the middle, and also a sum of rank 1 matrices with the eigenvalue/singular value coefficients.

More details are discussed in the next section.

## 6 The Five Factorizations of a Matrix

- Preface p.vii, The Plan for the Book.

$A = CR$ ,  $A = LU$ ,  $A = QR$ ,  $A = Q\Lambda Q^T$ ,  $A = U\Sigma V^T$  are illustrated one by one.

$A = CR$		Independent columns in $C$ Row echelon form in $R$ Leads to column rank = row rank
$A = LU$		$LU$ decomposition from Gaussian elimination (Lower triangular)(Upper triangular)
$A = QR$		$QR$ decomposition as Gram-Schmidt orthogonalization Orthogonal $Q$ and triangular $R$
$S = Q\Lambda Q^T$		Eigenvalue decomposition of a symmetric matrix $S$ Eigenvectors in $Q$ , eigenvalues in $\Lambda$
$A = U\Sigma V^T$		Singular value decomposition of all matrices $A$ Singular values in $\Sigma$

Table 1: The Five Factorization

### 6.1 $A = CR$

- Sec.1.4 Matrix Multiplication and  $A = CR$  (p.29)

All general rectangular matrices  $A$  have the same row rank as the column rank. This factorization is the most intuitive way to understand this theorem.  $C$  consists of independent columns of  $A$ , and  $R$  is the row reduced echelon form of  $A$ .  $A = CR$  reduces to  $r$  independent columns in  $C$  times  $r$  independent rows in  $R$ .

$$A = CR$$

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

Procedure: Look at the columns of  $A$  from left to right. Keep independent ones, discard dependent ones which can be created by the former columns. The column 1 and the column 2 survive, and the column 3 is discarded because it is expressed as a sum of the former two columns. To rebuild  $A$  by the independent columns 1, 2, you find a row echelon form  $R$  appearing in the right.

$$\begin{matrix} A \\ \begin{bmatrix} \text{grey} & \text{grey} & \text{grey} \end{bmatrix} \end{matrix} = \begin{matrix} C \\ \begin{bmatrix} \text{green} & \text{green} \end{bmatrix} \end{matrix} \begin{matrix} R \\ \begin{bmatrix} \text{blue} & \text{orange} & \text{yellow} \\ \text{blue} & \text{orange} & \text{yellow} \end{bmatrix} \end{matrix} = \begin{bmatrix} \text{blue} & \text{blue} & \text{orange} \\ \text{blue} & \text{orange} & \text{yellow} \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

using  
p1

Figure 11: Column Rank in  $CR$

Now you see the column rank is two because there are only two independent columns in  $C$  and all the columns of  $A$  are linear combinations of the two columns of  $C$ .

Figure 12: Row Rank in  $CR$

And you see the row rank is two because there are only two independent rows in  $R$  and all the rows of  $A$  are linear combinations of the two rows of  $R$ .

## 6.2 $A = LU$

Solving  $A\mathbf{x} = \mathbf{b}$  via Gaussian elimination can be expressed as an  $LU$  factorization. Usually, you apply elementary row operation matrices ( $E$ ) from left of  $A$  to make upper triangular  $U$ .

$$\begin{aligned} EA &= U \\ A &= E^{-1}U \\ \text{let } L &= E^{-1}, \quad A = LU \end{aligned}$$

Now solve  $A\mathbf{x} = \mathbf{b}$  in 2 steps: (1) forward  $L\mathbf{c} = \mathbf{b}$  and (2) back  $U\mathbf{x} = \mathbf{c}$ .

- Sec.2.3 (p.57) Matrix Computations and  $A = LU$

Here, we directly calculate  $L$  and  $U$  from  $A$ .

$$A = \begin{bmatrix} | \\ l_1 \\ | \end{bmatrix} [-\mathbf{u}_1^* -] + \begin{bmatrix} 0 & 0 & 0 \\ 0 & A_2 \\ 0 \end{bmatrix} = \begin{bmatrix} | \\ l_1 \\ | \end{bmatrix} [-\mathbf{u}_1^* -] + \begin{bmatrix} | \\ l_2 \\ | \end{bmatrix} [-\mathbf{u}_2^* -] + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & A_3 \end{bmatrix} = LU$$

Figure 13: Recursive Rank 1 Matrix Peeling from  $A$

To find  $L$  and  $U$ , peel off the rank 1 matrix made of the first row and the first column of  $A$ . This leaves  $A_2$ . Do this recursively and decompose  $A$  into the sum of rank 1 matrices.

Figure 14:  $LU$  rebuilds  $A$

To rebuild  $A$  from  $L$  times  $U$  is easy.

## 6.3 $A = QR$

$A = QR$  changes the columns of  $A$  into perpendicular columns of  $Q$ , keeping  $C(A) = C(Q)$ .

- Sec.4.4 Orthogonal matrices and Gram-Schmidt (p.165)

In Gram-Schmidt, the normalized  $\mathbf{a}_1$  is picked up as  $\mathbf{q}_1$  first and then  $\mathbf{a}_2$  is adjusted to be perpendicular to  $\mathbf{q}_1$  to create  $\mathbf{q}_2$ , and this procedure goes on.



$$\begin{aligned}
\mathbf{q}_1 &= \mathbf{a}_1 / \|\mathbf{a}_1\| \\
\mathbf{q}_2 &= \mathbf{a}_2 - (\mathbf{q}_1^T \mathbf{a}_2) \mathbf{q}_1, \quad \mathbf{q}_2 = \mathbf{q}_2 / \|\mathbf{q}_2\| \\
\mathbf{q}_3 &= \mathbf{a}_3 - (\mathbf{q}_1^T \mathbf{a}_3) \mathbf{q}_1 - (\mathbf{q}_2^T \mathbf{a}_3) \mathbf{q}_2, \quad \mathbf{q}_3 = \mathbf{q}_3 / \|\mathbf{q}_3\|
\end{aligned}$$

or you can write with  $r_{ij} = \mathbf{q}_i^T \mathbf{a}_j$ :

$$\begin{aligned}
\mathbf{a}_1 &= r_{11} \mathbf{q}_1 \\
\mathbf{a}_2 &= r_{12} \mathbf{q}_1 + r_{22} \mathbf{q}_2 \\
\mathbf{a}_3 &= r_{13} \mathbf{q}_1 + r_{23} \mathbf{q}_2 + r_{33} \mathbf{q}_3
\end{aligned}$$

The original  $A$  becomes  $QR$ : orthogonal times triangular.

$$A = \begin{bmatrix} | & | & | \\ \mathbf{q}_1 & \mathbf{q}_2 & \mathbf{q}_3 \\ | & | & | \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ & r_{22} & r_{23} \\ & & r_{33} \end{bmatrix} = QR$$

$$QQ^T = Q^T Q = I$$

Figure 15:  $A = QR$

The column vectors of  $A$  can be adjusted into an orthonormal set: the column vectors of  $Q$ . Each column vector of  $A$  can be rebuilt from  $Q$  and the upper triangular matrix  $R$ .

See Pattern 1 (P1) again for the graphic interpretation.

## 6.4 $S = Q\Lambda Q^T$

All symmetric matrices  $S$  must have real eigenvalues and orthogonal eigenvectors. The eigenvalues are the diagonal elements of  $\Lambda$  and the eigenvectors are in  $Q$ .

- Sec.6.3 (p.227) Symmetric Positive Definite Matrices

$$\begin{aligned}
S &= Q\Lambda Q^T = \begin{bmatrix} | & | & | \\ \mathbf{q}_1 & \mathbf{q}_2 & \mathbf{q}_3 \\ | & | & | \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3 \end{bmatrix} \begin{bmatrix} -\mathbf{q}_1^T - \\ -\mathbf{q}_2^T - \\ -\mathbf{q}_3^T - \end{bmatrix} \\
&= \lambda_1 \begin{bmatrix} | \\ \mathbf{q}_1 \\ | \end{bmatrix} \begin{bmatrix} -\mathbf{q}_1^T - \end{bmatrix} + \lambda_2 \begin{bmatrix} | \\ \mathbf{q}_2 \\ | \end{bmatrix} \begin{bmatrix} -\mathbf{q}_2^T - \end{bmatrix} + \lambda_3 \begin{bmatrix} | \\ \mathbf{q}_3 \\ | \end{bmatrix} \begin{bmatrix} -\mathbf{q}_3^T - \end{bmatrix} \\
&= \lambda_1 P_1 + \lambda_2 P_2 + \lambda_3 P_3
\end{aligned}$$

$$P_1 = \mathbf{q}_1 \mathbf{q}_1^T, \quad P_2 = \mathbf{q}_2 \mathbf{q}_2^T, \quad P_3 = \mathbf{q}_3 \mathbf{q}_3^T$$

$$S = Q \Lambda Q^T = \begin{bmatrix} \text{col 1} & \text{col 2} & \text{col 3} \end{bmatrix} \begin{bmatrix} 1 & & \\ & 2 & \\ & & 3 \end{bmatrix} \begin{bmatrix} \text{row 1} \\ \text{row 2} \\ \text{row 3} \end{bmatrix} = \lambda_1 q_1 q_1^T + \lambda_2 q_2 q_2^T + \lambda_3 q_3 q_3^T \quad \text{using P4}$$

Figure 16:  $S = Q\Lambda Q^T$

A symmetric matrix  $S$  is diagonalized into  $\Lambda$  by an orthogonal matrix  $Q$  and its transpose. And it is broken down into a combination of rank 1 projection matrices  $P = qq^T$ . This is the spectral theorem.

Note that Pattern 4 (P4) is working for the decomposition.

$$\begin{aligned} S &= S^T = \lambda_1 P_1 + \lambda_2 P_2 + \lambda_3 P_3 \\ QQ^T &= P_1 + P_2 + P_3 = I \\ P_1 P_2 &= P_2 P_3 = P_3 P_1 = O \\ P_1^2 &= P_1 = P_1^T, \quad P_2^2 = P_2 = P_2^T, \quad P_3^2 = P_3 = P_3^T \end{aligned}$$

## 6.5 $A = U\Sigma V^T$

- Sec.7.1 (p.259) Singular Values and Singular Vectors

All matrices including rectangular ones have a singular value decomposition (SVD).  $A = U\Sigma V^T$  has the singular vectors of  $A$  in  $U$  and  $V$ . And its singular values line up in  $\Sigma$ 's diagonal elements. The following illustrates the 'reduced' SVD.

$$A = U \Sigma V^T = \begin{bmatrix} \text{col 1} & \text{col 2} & \text{col 3} \end{bmatrix} \begin{bmatrix} 1 & \\ & 2 \end{bmatrix} \begin{bmatrix} \text{row 1} \\ \text{row 2} \end{bmatrix} = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T \quad \text{using P4}$$

Figure 17:  $A = U\Sigma V^T$

You can find  $V$  as an orthonormal basis of  $\mathbb{R}^n$  (eigenvectors of  $A^T A$ ), and  $U$  as an orthonormal basis of  $\mathbb{R}^m$  (eigenvectors of  $AA^T$ ). Together they diagonalize  $A$  into  $\Sigma$ . This is also expressed as a combination of rank 1 matrices.

$$\begin{aligned} A = U\Sigma V^T &= \begin{bmatrix} | & | & | \\ \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \\ | & | & | \end{bmatrix} \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \end{bmatrix} \begin{bmatrix} -\mathbf{v}_1^T - \\ -\mathbf{v}_2^T - \end{bmatrix} = \sigma_1 \begin{bmatrix} | \\ \mathbf{u}_1 \\ | \end{bmatrix} \begin{bmatrix} -\mathbf{v}_1^T - \end{bmatrix} + \sigma_2 \begin{bmatrix} | \\ \mathbf{u}_2 \\ | \end{bmatrix} \begin{bmatrix} -\mathbf{v}_2^T - \end{bmatrix} \\ &= \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T \end{aligned}$$

Note that:

$$\begin{aligned} UU^T &= I_m \\ VV^T &= I_n \end{aligned}$$

See Pattern 4 (P4) for the graphic notation.

## Conclusion and Acknowledgements

I presented systematic visualizations of matrix/vector multiplication and their application to the Five Matrix Factorizations. I hope you enjoyed them and will use them in your understanding of Linear Algebra.

Ashley Fernandes helped me with beautifying this paper in typesetting and made it much more consistent and professional.

To conclude this paper, I'd like to thank Prof. Gilbert Strang for publishing “Linear Algebra for Everyone”. It guides us through a new vision to these beautiful landscapes in Linear Algebra. Everyone can reach a fundamental understanding of its underlying ideas in a practical manner that introduces us to contemporary and also traditional data science and machine learning. An important part of the matrix world.

## References and Related Works

1. Gilbert Strang(2020), *Linear Algebra for Everyone*, Wellesley Cambridge Press.,  
<http://math.mit.edu/everyone>
2. Gilbert Strang(2016), *Introduction to Linear Algebra*, Wellesley Cambridge Press, 5th ed.,  
<http://math.mit.edu/linearalgebra>
3. Kenji Hiranabe(2021), *Map of Eigenvalues*, An Agile Way(blog),  
<https://anagileway.com/2021/10/01/map-of-eigenvalues/>

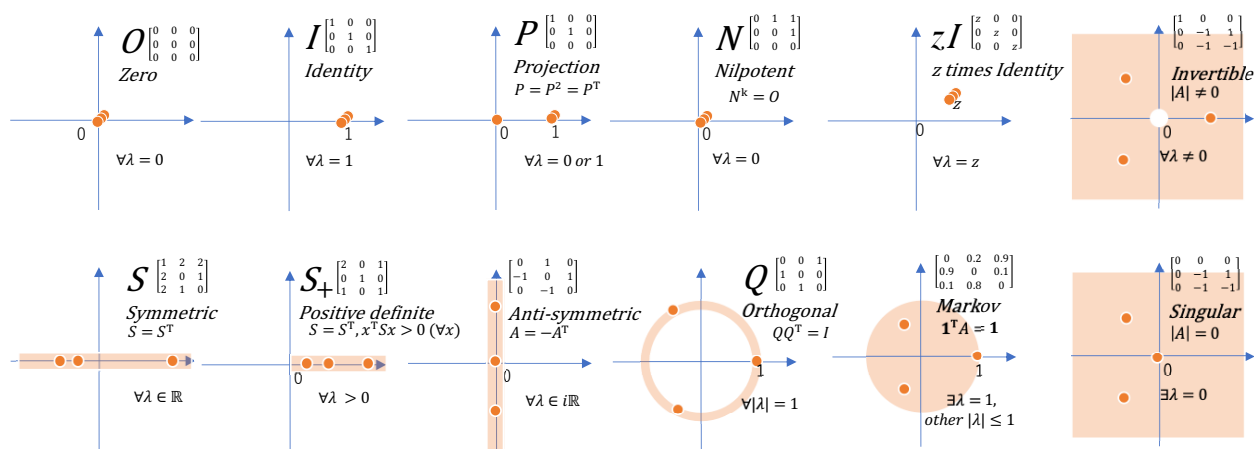


Figure 18: Map of Eigenvalues

4. Kenji Hiranabe(2020), *Matrix World*, An Agile Way(blog),  
<https://anagileway.com/2020/09/29/matrix-world-in-linear-algebra-for-everyone/>

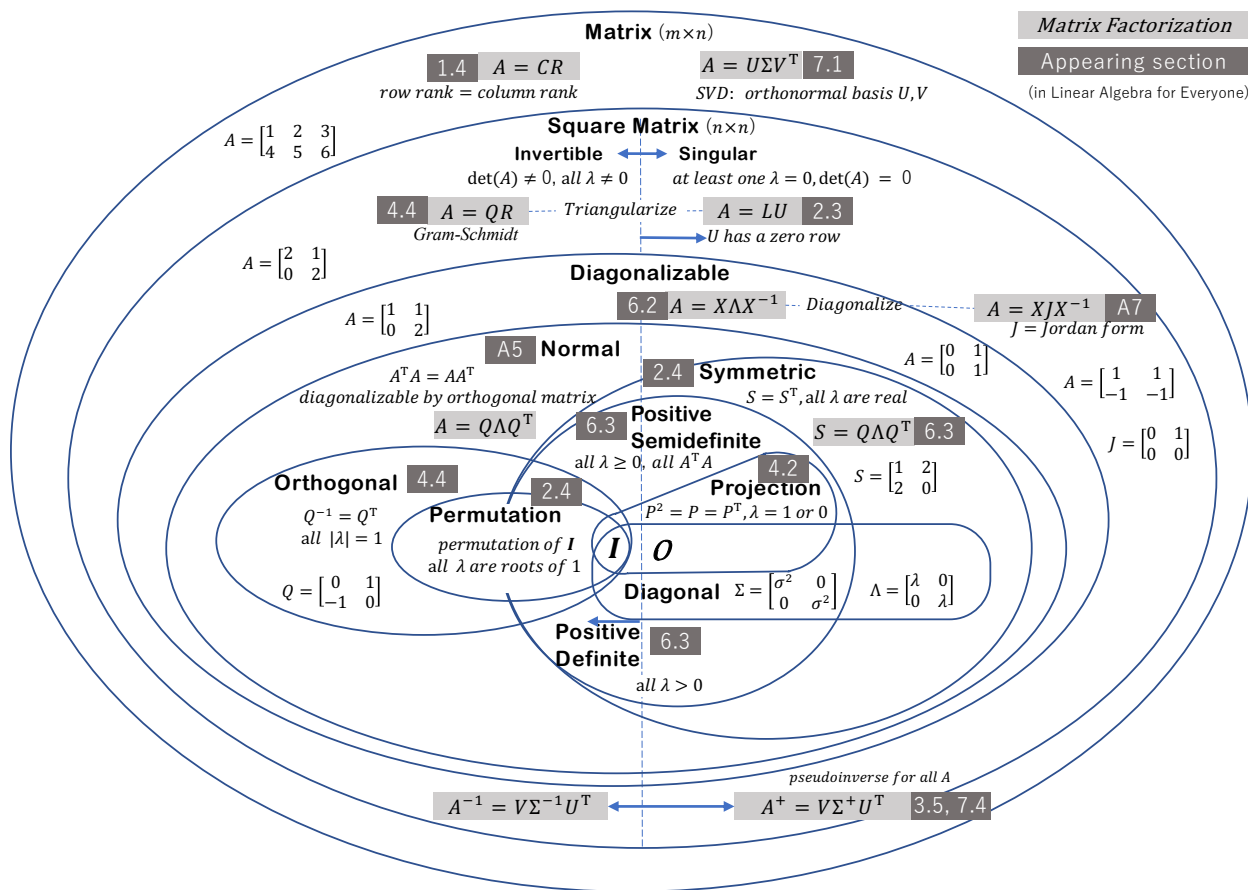


Figure 19: Matrix World