

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 8, 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.¹ 它们包括矩阵分解 (Column-Row, \mathbf{CR})、高斯消去法 (Gaussian Elimination, \mathbf{LU})、格拉姆-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	矩阵的五种分解	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/>

‡twitter: @kfchliu, 微博用户: 5717297833

¹“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.

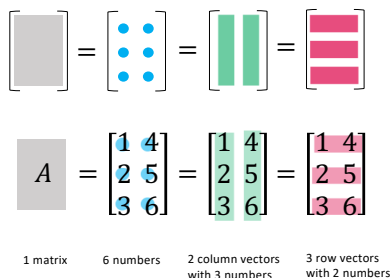


Figure 1: 从四个角度理解矩阵

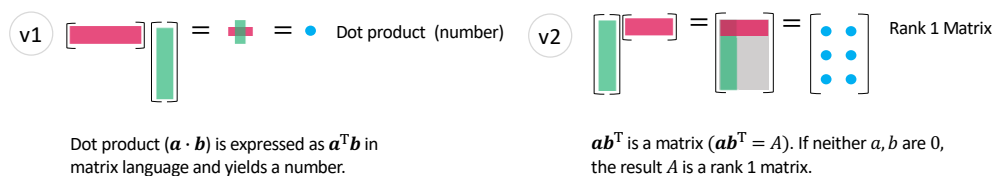
$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} = \begin{bmatrix} | & | \\ \mathbf{a}_1 & \mathbf{a}_2 \\ | & | \end{bmatrix} = \begin{bmatrix} -\mathbf{a}_1^* - \\ -\mathbf{a}_2^* - \\ -\mathbf{a}_3^* - \end{bmatrix}$$

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) Linear combination and dot products
- 1.3 节 (p.25) Matrix of Rank One
- 1.4 节 (p.29) Row way and column way



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

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

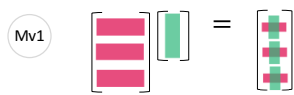
Figure 2: 向量乘以向量 - (v1), (v2)

(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.

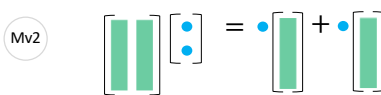
3 矩阵乘以向量——2 个视角

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) Linear combinations
- 1.3 节 (p.21) Matrices and Column Spaces



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 .


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

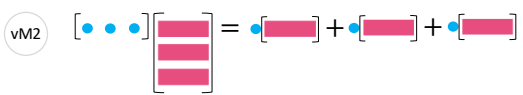
$$A\mathbf{x} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}$$

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) Dimensions of the Four Subspaces

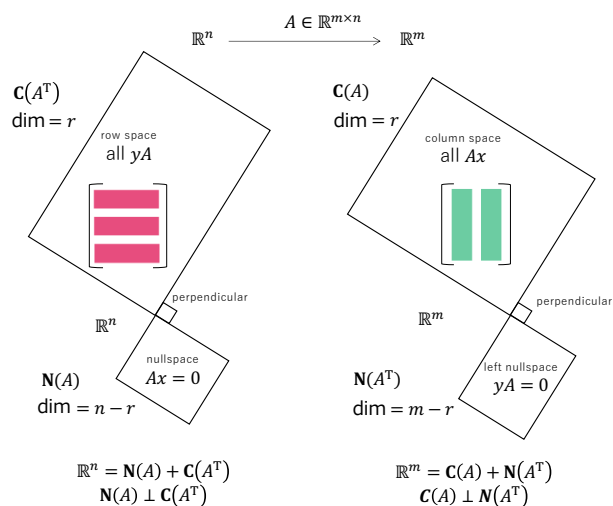


Figure 5: 四个子空间

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) Four ways to multiply $AB = C$
- 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 .

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

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。

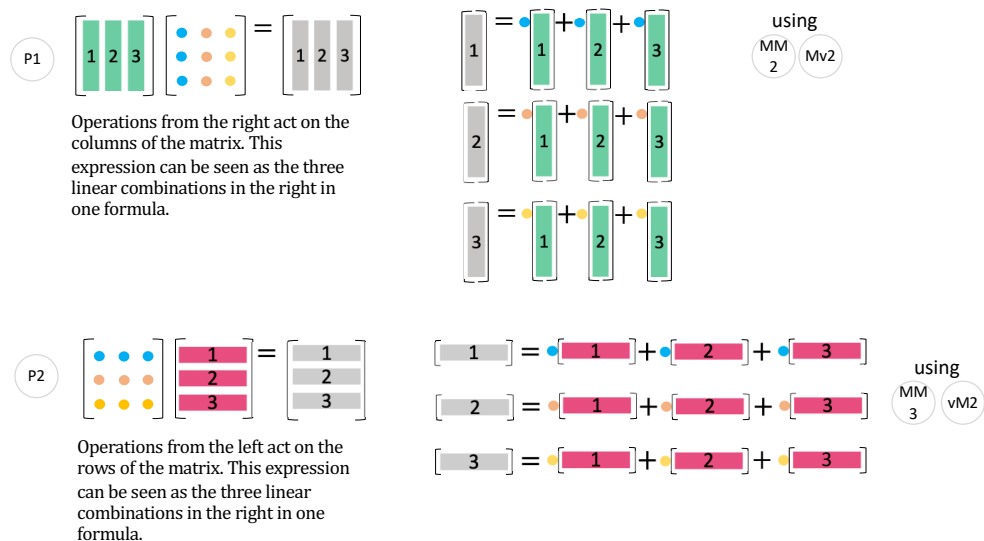


Figure 7: 图 1, 2 - (P1), (P1)

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).

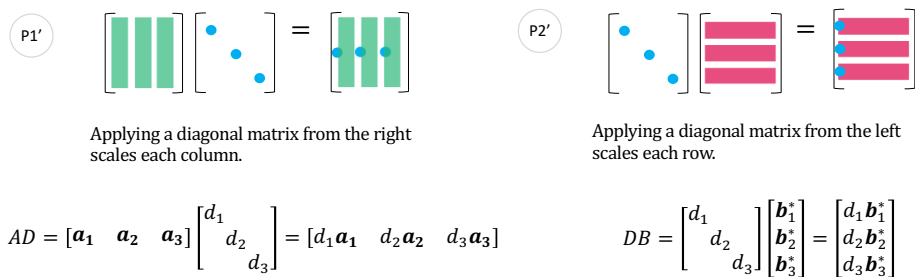


Figure 8: 图 1', 2' - (P1'), (P2')

(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.

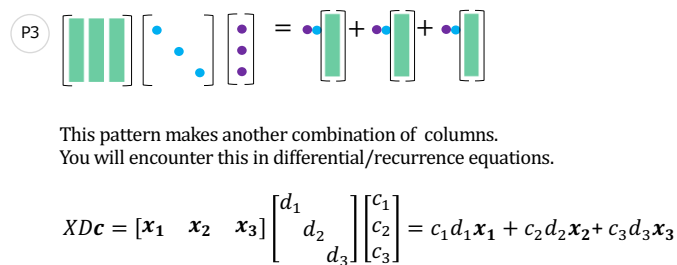


Figure 9: 图 3 - (P3)

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) Eigenvalues and Eigenvectors
- 6.4 节 (p.243) Systems of Differential Equations

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

在两种问题中, 它的解都可以用 A 的特征值 $(\lambda_1, \lambda_2, \lambda_3)$ 、特征向量 $X = [\mathbf{x}_1 \ \mathbf{x}_2 \ \mathbf{x}_3]$ 和系数 $\mathbf{c} = [c_1 \ c_2 \ c_3]^T$ 表示. 其中 \mathbf{c} 是以 X 为基底的初始值 $\mathbf{u}(0) = \mathbf{u}_0$ 的坐标.

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

以上两个问题的通解为:

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

见 Figure9: P3 中的 $XD\mathbf{c}$.

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)

P4 在特征值分解和特异值分解中都会用到. 两种分解都可以表示为三个矩阵之积, 其中中间的矩阵均为对角矩阵. 且都可以表示为带特征值/特异值系数的秩 1 矩阵之积. 更多细节将在下一节中讨论.

6 矩阵的五种分解

- 前言 p.vii, The Plan for the Book.

$A = CR, A = LU, A = QR, A = Q\Lambda Q^T, A = U\Sigma V^T$ 将一一说明.

$A = CR$		C 为 A 的线性无关列 R 为 A 的行阶梯形矩阵 可推知列秩 = 行秩
$A = LU$		LU 分解通过 高斯消去法 (下三角)(上三角)
$A = QR$		QR 分解为 格拉姆-施密特正交化中的 正交矩阵 Q 和三角矩阵 R
$S = Q\Lambda Q^T$		对称矩阵 S 可以进行 特征值分解 特征向量组成 Q , 特征值组成 Λ
$A = U\Sigma V^T$		所有矩阵 A 的 奇异值分解 奇异值组成 Σ

Table 1: 五种分解

6.1 $A = CR$

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

所有一般的长矩阵 A 都有相同的行秩和列秩. 这个分解是理解这一定理最直观的方法. C 由 A 的线性无关列组成, R 为 A 的行阶梯形矩阵 (消除了零行). $A = CR$ 将 A 化简为 r 的线性无关列 C 和线性无关行 R 的乘积.

$$A = CR$$

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

推导过程: 从左往右看 A 的列. 保留其中线性无关的列, 去掉可以由前者线性表出的列. 则第 1、2 列被保留, 而第三列因为可以由前两列之和表示而被去掉. 而要通过线性无关的 1、2 两列重新构造出 A , 需要右乘一个行阶梯矩阵 R .

$$\begin{matrix} A & & C & & R \\ \left[\begin{array}{|c|c|c|} \hline \text{grey} & \text{grey} & \text{grey} \\ \hline \end{array} \right] & = & \left[\begin{array}{|c|c|} \hline \text{green} & \text{green} \\ \hline \end{array} \right] & \left[\begin{array}{|c|c|c|} \hline \text{blue} & \text{orange} & \text{yellow} \\ \hline \end{array} \right] & = & \left[\begin{array}{|c|c|c|} \hline \text{blue} + \text{orange} & \text{orange} & \text{yellow} \\ \hline \end{array} \right] & \left[\begin{array}{|c|c|c|} \hline \text{orange} & \text{orange} & \text{yellow} \\ \hline \end{array} \right] & \left[\begin{array}{|c|c|c|} \hline \text{yellow} & \text{orange} & \text{yellow} \\ \hline \end{array} \right] \end{matrix}$$

using
P1

Figure 11: CR 中列的秩

现在你会发现行的秩为 2, 因为 C 中只有 2 个线性无关列. 而 A 中所有的列都可以由 C 中的 2 列线性表出.

$$\begin{matrix} A \\ \left[\begin{array}{c} \text{---} \\ \text{---} \end{array} \right] \end{matrix} = \begin{matrix} C \\ \left[\begin{array}{cc} \bullet & \bullet \\ \bullet & \bullet \end{array} \right] \end{matrix} \begin{matrix} R \\ \left[\begin{array}{c} \text{1} \\ \text{2} \end{array} \right] \end{matrix} = \begin{matrix} \left[\begin{array}{c} \bullet \text{1} \\ \bullet \text{2} \end{array} \right] + \left[\begin{array}{c} \bullet \text{2} \\ \bullet \text{2} \end{array} \right] \end{matrix}$$

using
p2

Figure 12: CR 中行的秩

同样, 列秩也为 2, 因为 R 中只有 2 个线性无关行, 且 A 中所有的行都可以由 R 中的 2 行线性表出.

6.2 $A = LU$

用高斯消除法求解 $Ax = b$ 也被称为 LU 分解. 通常, 是 A 左乘一个初等行变换矩阵 (E) 来得到一个上三角矩阵 U .

$$EA = U$$

$$A = E^{-1}U$$

$$\text{let } L = E^{-1}, \quad A = LU$$

现在, 求解 $Ax = b$ 有 2 步: (1) 求解 $Lc = b$, (2) 代回 $Ux = c$.

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

在这里, 我们直接通过 A 计算 L 和 U .

$$A = \begin{bmatrix} | \\ l_1 \\ | \end{bmatrix} [-u_1^* -] + \begin{bmatrix} 0 & 0 & 0 \\ 0 & A_2 \\ 0 \end{bmatrix} = \begin{bmatrix} | \\ l_1 \\ | \end{bmatrix} [-u_1^* -] + \begin{bmatrix} | \\ l_2 \\ | \end{bmatrix} [-u_2^* -] + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & A_3 \end{bmatrix} = LU$$

$$\begin{matrix} A \\ \left[\begin{array}{c} \text{---} \\ \text{---} \end{array} \right] \end{matrix} = \begin{matrix} \left[\begin{array}{c} \text{---} \\ \text{---} \end{array} \right] \end{matrix} + \begin{matrix} \left[\begin{array}{c} \text{---} \\ \text{---} \end{array} \right] \end{matrix} + \begin{matrix} \left[\begin{array}{c} \text{---} \\ \text{---} \end{array} \right] \end{matrix} = \begin{matrix} L \\ \left[\begin{array}{c} \text{---} \\ \text{---} \end{array} \right] \end{matrix} \begin{matrix} U \\ \left[\begin{array}{c} \text{---} \\ \text{---} \end{array} \right] \end{matrix}$$

Figure 13: $A = LU$

要计算 L 和 U , 首先分离出由 A 的第一行和第一列组成的外积. 余下的部分为 A_2 . 递归执行此操作, 将 A 分解为秩 1 矩阵之和.

$$\begin{matrix} L \\ \left[\begin{array}{c} \text{---} \\ \text{---} \end{array} \right] \end{matrix} \begin{matrix} U \\ \left[\begin{array}{c} \text{---} \\ \text{---} \end{array} \right] \end{matrix} = \begin{matrix} \left[\begin{array}{c} \text{---} \\ \text{---} \end{array} \right] \end{matrix} + \begin{matrix} \left[\begin{array}{c} \text{---} \\ \text{---} \end{array} \right] \end{matrix} + \begin{matrix} \left[\begin{array}{c} \text{---} \\ \text{---} \end{array} \right] \end{matrix}$$

using
MM
4

Figure 14: 由 LU 重新构造 A

由 L 乘以 U 来重新构造 A 则相对简单.

6.3 $A = QR$

$A = QR$ 是在保持 $C(A) = C(Q)$ 的条件下, 将 A 转化为正交矩阵 Q .

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

在格拉姆-施密特正交化中, 首先, 单位化的 a_1 被用作 q_1 , 然后求出 a_2 与 q_1 正交所得到的 q_2 , 以此类推.

$$\begin{aligned} q_1 &= a_1 / \|a_1\| \\ q_2 &= a_2 - (q_1^T a_2) q_1, \quad q_2 = q_2 / \|q_2\| \\ q_3 &= a_3 - (q_1^T a_3) q_1 - (q_2^T a_3) q_2, \quad q_3 = q_3 / \|q_3\| \end{aligned}$$

或者你也可以写作 $r_{ij} = \mathbf{q}_i^T \mathbf{a}_j$:

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

原本的 A 就可以表示为 QR : 正交矩阵乘以上三角矩阵.

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

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

$$\begin{bmatrix} | & | & | \\ A \\ | & | & | \end{bmatrix} = \begin{bmatrix} | & | & | \\ Q \\ | & | & | \end{bmatrix} \begin{bmatrix} \bullet & \bullet & \bullet \\ R \\ \bullet & \bullet & \bullet \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \\ \bullet & \bullet & \bullet \\ | & | & | \end{bmatrix} = \begin{bmatrix} | & | & | \\ \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \\ | & | & | \end{bmatrix} \quad \text{using P1}$$

Figure 15: $A = QR$

A 的列向量就可以转化为一个正交集: Q 的列向量. A 的每一个列向量都可以用 Q 和上三角矩阵 R 重新构造出.

图释可以回头看 P1.

6.4 $S = Q\Lambda Q^T$

所有对称矩阵 S 都必须有实特征值和正交特征向量. 特征值是 Λ 的对角元素, 特征向量在 Q 中.

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

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

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

$$\begin{bmatrix} S \\ | & | & | \end{bmatrix} = \begin{bmatrix} | & | & | \\ Q \\ | & | & | \end{bmatrix} \begin{bmatrix} \bullet & \bullet & \bullet \\ \Lambda \\ \bullet & \bullet & \bullet \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} \lambda_1 \mathbf{q}_1 \mathbf{q}_1^T \\ \bullet & \bullet & \bullet \end{bmatrix} + \begin{bmatrix} \lambda_2 \mathbf{q}_2 \mathbf{q}_2^T \\ \bullet & \bullet & \bullet \end{bmatrix} + \begin{bmatrix} \lambda_3 \mathbf{q}_3 \mathbf{q}_3^T \\ \bullet & \bullet & \bullet \end{bmatrix} \quad \text{using P4}$$

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

一个对称矩阵 S 通过一个正交矩阵 Q 和它的转置矩阵, 对角化为 Λ . 然后被分解为一阶投影矩阵 $P = \mathbf{q}\mathbf{q}^T$ 的组合. 这就是谱定理.

注意, 这里的分解用到了 P4.

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

6.5 $A = U\Sigma V^T$

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

包括长方阵在内的所有矩阵都具有奇异值分解 (SVD). $A = U\Sigma V^T$ 中, 有 A 的奇异向量 U 和 V . 奇异值则排列在 Σ 的对角线上. 下图就是“简化版”的 SVD.

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

你可以发现, V 是 \mathbb{R}^n ($A^T A$ 的特征向量) 的标准正交基, 而 U 是 \mathbb{R}^m (AA^T 的特征向量) 的标准正交基. 它们共同将 A 对角化为 Σ . 这也可以表示为秩 1 矩阵的线性组合.

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

注意:

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

图释见 P4.

总结和致谢

我展示了矩阵/向量乘法的系统可视化与它们在五种矩阵分解中的应用. 我希望你能够喜欢它们、通过它们加深对线性代数的理解.

Ashley Fernandes 在排版时帮我美化了这篇论文, 使它更加一致和专业.

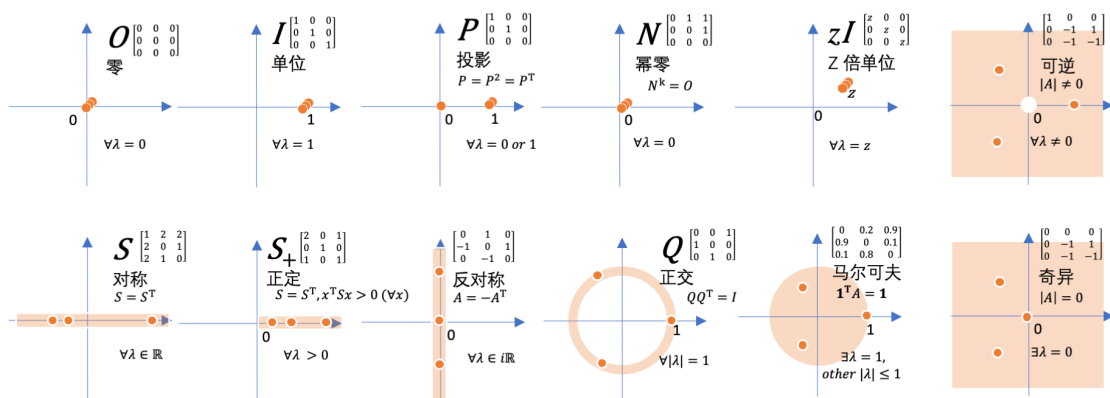
在结束这篇论文之前, 我要感谢 Gilbert Strang 教授出版了《Linear Algebra for Everyone》一书. 它引导我们通过新的视角去了解线性代数中这些美丽的风景. 其中介绍了当代和传统的数据科学和机器学习, 每个人都可以通过实用的方式对它的基本思想进行基本理解. 矩阵世界的重要组成部分.

参考文献与相关工作

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/>

实 $n \times n$ 方阵的特征值映射



By Kenji Hiranabe with the kindest help of Prof. Gilbert Strang.
 Translator: Kefang Liu



Figure 18: 特征值图

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

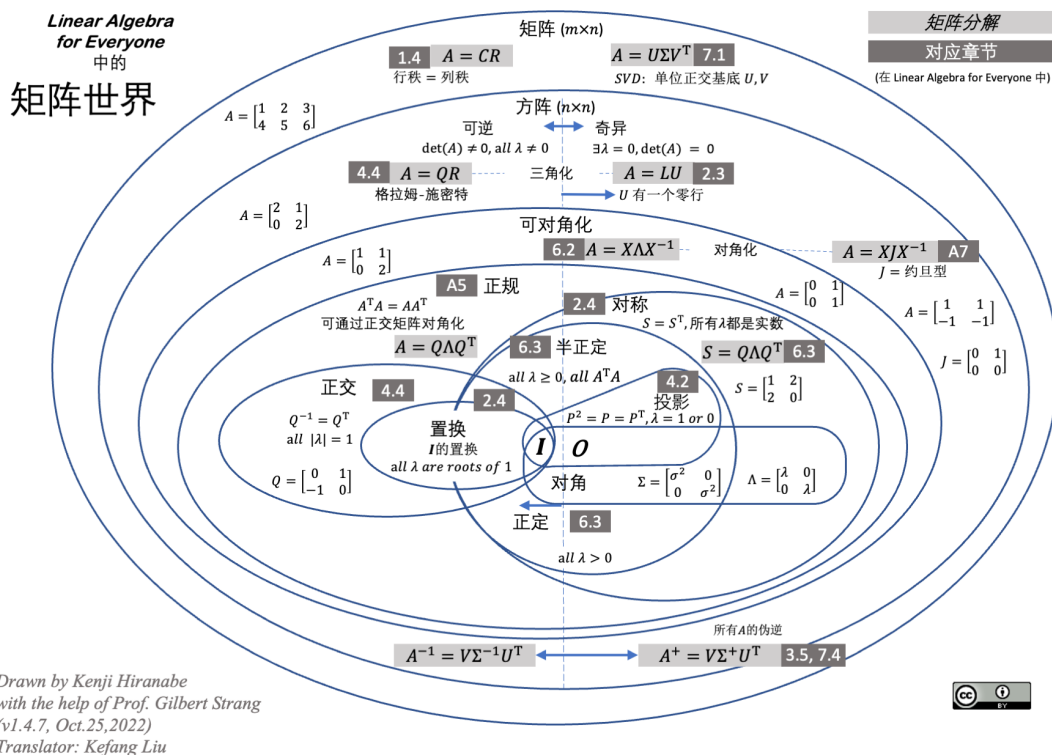


Figure 19: 矩阵世界