The Art of Linear Algebra

– Graphic Notes on "Linear Algebra for Everyone" –

Kenji Hiranabe[1]
with the kindest help of Gilbert Strang[2]

September 1, 2021/updated June 30, 2025

Abstract

I try to intuitively visualize some important concepts introduced in "Linear Algebra for Everyone",[3] which include Column-Row ($\boldsymbol{CR}$), Gaussian Elimination ($\boldsymbol{LU}$), Gram-Schmidt Orthogonalization ($\boldsymbol{QR}$), Eigenvalues and Diagonalization ($\boldsymbol{Q\Lambda Q^{\mathrm{T}}}$), and Singular Value Decomposition ($\boldsymbol{U\Sigma V^{\mathrm{T}}}$). This paper aims at promoting the understanding of vector/matrix calculations and algorithms from the perspective of matrix factorization. All the artworks including the article itself are maintained under the GitHub repository `https://github.com/kenjihiranabe/The-Art-of-Linear-Algebra/`.

Foreword

I am happy to see Kenji Hiranabe's pictures of matrix operations in linear algebra ! The pictures are an excellent way to show the algebra. We can think of matrix multiplications by row $\cdot$ column dot products, but that is not all – it is "linear combinations" and "rank 1 matrices" that complete the algebra and the art. I am very grateful to see the books in Japanese translation and the ideas in Kenji's pictures.

– Gilbert Strang
Professor of Mathematics at MIT

Contents

1 Viewing a Matrix – 4 Ways

A matrix ($m \times n$) can be viewed as 1 matrix, mn numbers, n columns and m rows.

[1] twitter: @hiranabe, k-hiranabe@esm.co.jp, `https://anagileway.com`
[2] Massachusetts Institute of Technology, `http://www-math.mit.edu/~gs/`
[3] "Linear Algebra for Everyone": `http://math.mit.edu/everyone/` with Japanese translation from Kindai Kagaku.

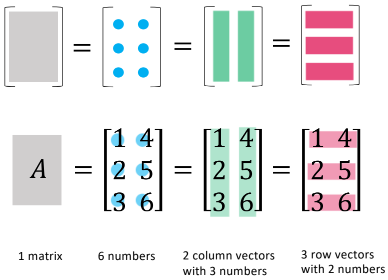


Figure 1: Viewing a Matrix in 4 Ways

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} = \begin{bmatrix} | & | \\ \boldsymbol{a_1} & \boldsymbol{a_2} \\ | & | \end{bmatrix} = \begin{bmatrix} -\boldsymbol{a_1^*}- \\ -\boldsymbol{a_2^*}- \\ -\boldsymbol{a_3^*}- \end{bmatrix}$$

Here, the column vectors are in bold as $\boldsymbol{a_1}$. Row vectors include $*$ as in $\boldsymbol{a_1^*}$. Transposed vectors and matrices are indicated by T as in $\boldsymbol{a}^{\mathrm{T}}$ and A^{T}.

2 Vector times Vector – 2 Ways

Hereafter I point to specific sections of "Linear Algebra for Everyone" and present graphics which illustrate the concepts with short names in gray circles.

- Sec. 1.1 (p.2) Linear combination and dot products
- Sec. 1.3 (p.25) Matrix of Rank One
- Sec. 1.4 (p.29) Row way and column way

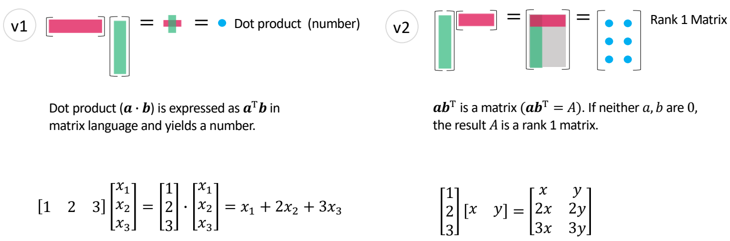


Figure 2: Vector times Vector - (v1), (v2)

(v1) is an elementary operation of two vectors, but (v2) multiplies the column to the row and produces a rank 1 matrix. Knowing this outer product (v2) is the key to the following sections.

3 Matrix times Vector – 2 Ways

A matrix times a vector creates a vector of three dot products (Mv1) as well as a linear combination (Mv2) of the column vectors of A.

- Sec. 1.1 (p.3) Linear combinations
- Sec. 1.3 (p.21) Matrices and Column Spaces

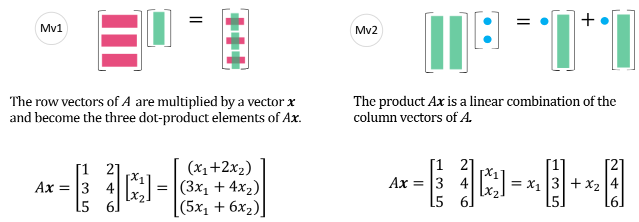


Figure 3: Matrix times Vector - (Mv1), (Mv2)

At first, you learn (Mv1). But when you get used to viewing it as (Mv2), you can understand $A\boldsymbol{x}$ as a linear combination of the columns of A. Those products fill the column space of A denoted as $\mathbf{C}(A)$. The solution space of $A\boldsymbol{x} = \mathbf{0}$ is the nullspace of A denoted as $\mathbf{N}(A)$. To understand the nullspace, let the right-hand side of (Mv1) be $\mathbf{0}$ and see all the dot products are zero.

Also, (vM1) and (vM2) show the same pattern for a row vector times a matrix.

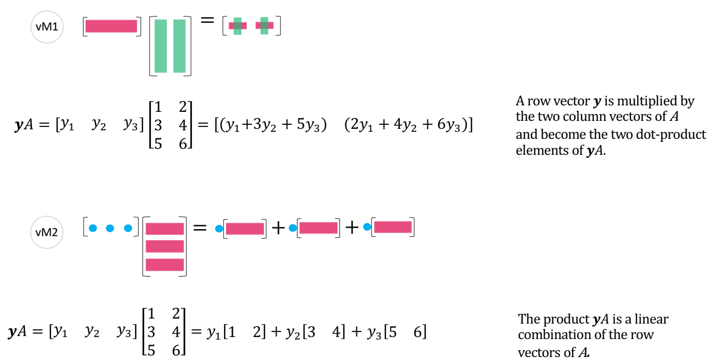


Figure 4: Vector times Matrix - (vM1), (vM2)

The products fill the row space of A denoted as $\mathbf{C}(A^{\mathrm{T}})$. The solution space of $yA = 0$ is the left-nullspace of A, denoted as $\mathbf{N}(A^{\mathrm{T}})$.

The four subspaces consist of $\mathbf{N}(A)$ + $\mathbf{C}(A^{\mathrm{T}})$ (which are perpendicular to each other) in $\mathbb{R}^n$ and $\mathbf{N}(A^{\mathrm{T}})$ + $\mathbf{C}(A)$ in $\mathbb{R}^m$ (which are perpendicular to each other).

- Sec. 3.5 (p.124) Dimensions of the Four Subspaces

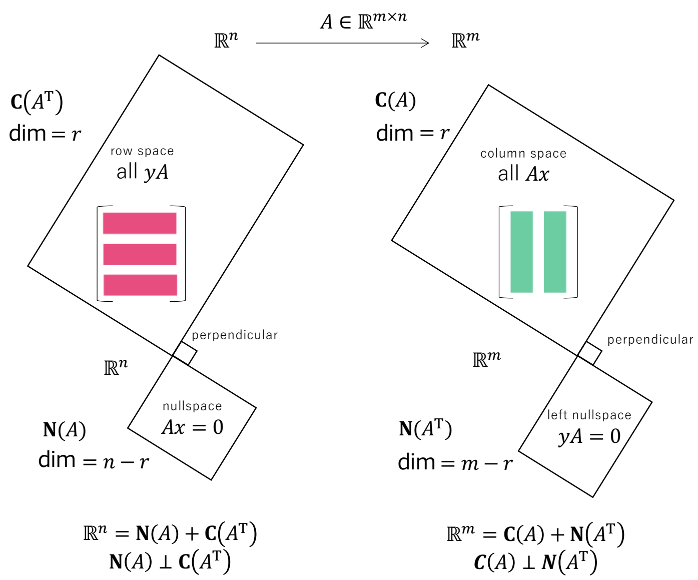


Figure 5: The Four Subspaces

See $A = CR$ (Sec 6.1) for the rank r.

4 Matrix times Matrix – 4 Ways

"Matrix times Vector" naturally extends to "Matrix times Matrix".

- Sec. 1.4 (p.35) Four Ways to Multiply $\boldsymbol{AB} = \boldsymbol{C}$
- Also see the back cover of the book

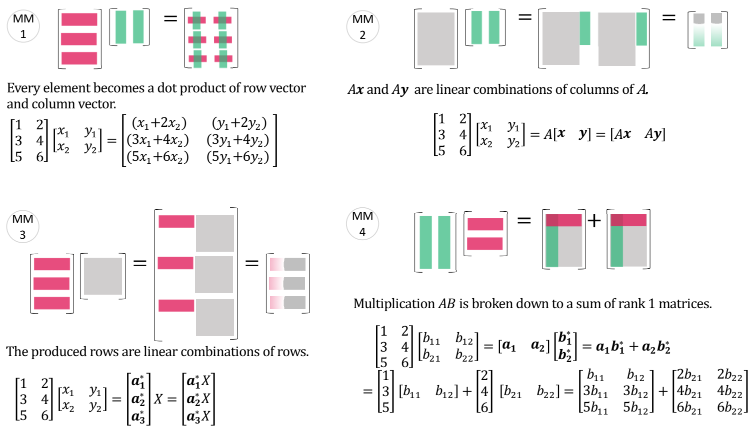


Figure 6: Matrix times Matrix - (MM1), (MM2), (MM3), (MM4)

5 Practical Patterns

Here, I show some practical patterns which allow you to capture the upcoming factorizations in a more intuitive way.

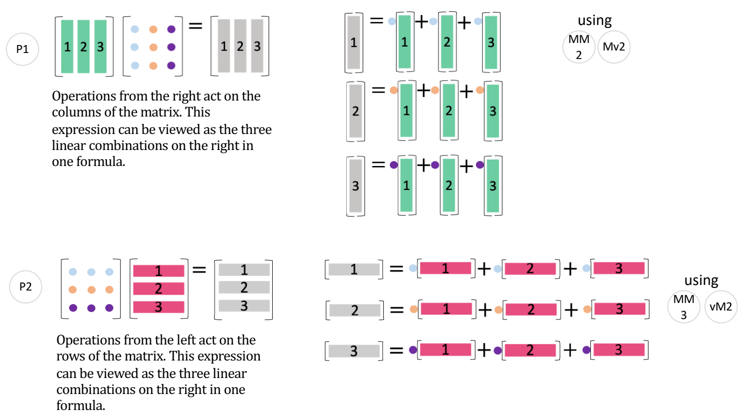


Figure 7: Pattern 1, 2 - (P1), (P1)

Pattern 1 is a combination of (MM2) and (Mv2). Pattern 2 is an extension of (MM3). Note that Pattern 1 is a column operation (multiplying a matrix from right), whereas Pattern 2 is a row operation (multiplying a matrix from left).

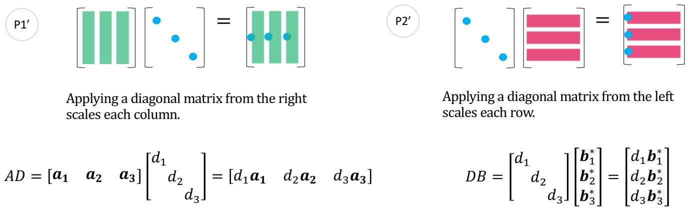


Figure 8: Pattern 1′, 2′ - (P1′), (P2′)

(P1′) multiplies the diagonal numbers to the columns of the matrix, whereas (P2′) multiplies the diagonal numbers to the row of the matrix. Both are variants of (P1) and (P2).

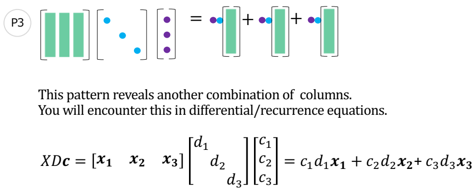


Figure 9: Pattern 3 - (P3)

This pattern emerges when you solve differential equations and recurrence equations:

- Sec. 6 (p.201) Eigenvalues and Eigenvectors
- Sec. 6.4 (p.243) Systems of Differential Equations

$$\frac{d\boldsymbol{u}(t)}{dt} = A\boldsymbol{u}(t), \quad \boldsymbol{u}(0) = \boldsymbol{u}_0$$
$$\boldsymbol{u}_{n+1} = A\boldsymbol{u}_n, \quad \boldsymbol{u}_{\boldsymbol{0}} = \boldsymbol{u}_0$$

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

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

and the general solution of the two equations are:

$$\begin{aligned} \boldsymbol{u}(t) &= e^{At}\boldsymbol{u}_0 = Xe^{\Lambda t}X^{-1}\boldsymbol{u}_{\boldsymbol{0}} && = Xe^{\Lambda t}\boldsymbol{c} = c_1e^{\lambda_1 t}\boldsymbol{x}_1 + c_2e^{\lambda_2 t}\boldsymbol{x}_2 + c_3e^{\lambda_3 t}\boldsymbol{x}_3 \\ \boldsymbol{u}_n &= A^n\boldsymbol{u}_0 = X\Lambda^nX^{-1}\boldsymbol{u}_{\boldsymbol{0}} && = X\Lambda^n\boldsymbol{c} = c_1\lambda_1^n\boldsymbol{x}_1 + c_2\lambda_2^n\boldsymbol{x}_2 + c_3\lambda_3^n\boldsymbol{x}_3 \end{aligned}$$

See Figure 9: Pattern 3 (P3) above again to get XDc.

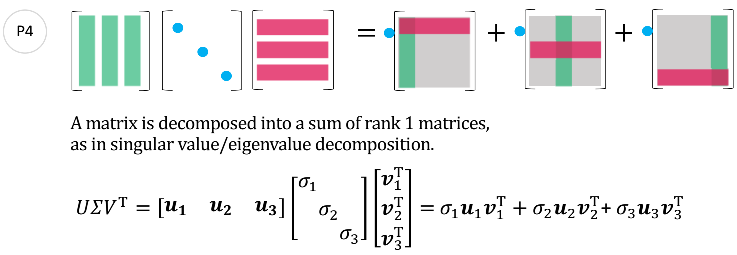


Figure 10: Pattern 4 - (P4)

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.

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

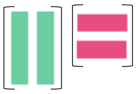
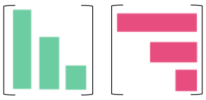
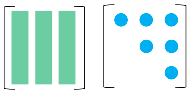
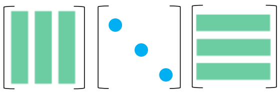
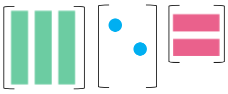
$\boldsymbol{A = CR}$		Independent columns in C Row echelon form in R Leads to column rank = row rank
$\boldsymbol{A = LU}$		LU decomposition from Gaussian elimination (Lower triangular)(Upper triangular)
$\boldsymbol{A = QR}$		QR decomposition as Gram-Schmidt orthogonalization Orthogonal Q and triangular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

Table 1: The Five Factorization

6.1 $\boldsymbol{A = CR}$

- Sec.1.4 Matrix Multiplication and $\boldsymbol{A = CR}$ (p.29)

The row rank and the column rank of a general rectangular matrix A are equal. 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 and 2, you find a row echelon form R appearing on the right.

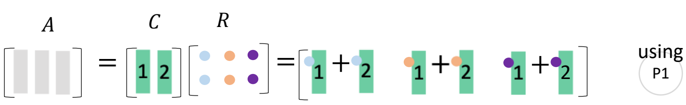


Figure 11: Column Rank in CR

Now 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.

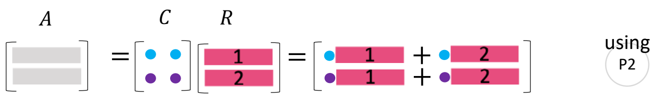


Figure 12: Row Rank in CR

And the row rank is also 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 $\boldsymbol{A = LU}$

Solving $A\boldsymbol{x} = \boldsymbol{b}$ via Gaussian elimination can be represented as an LU factorization. Usually, you apply elementary row operation matrices (E) to 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 $A\boldsymbol{x} = \boldsymbol{b}$ in 2 steps: (1) forward $L\boldsymbol{c} = \boldsymbol{b}$ and (2) back $U\boldsymbol{x} = \boldsymbol{c}$.

- Sec.2.3 (p.57) Matrix Computations and $\boldsymbol{A = LU}$

Here, we directly calculate L and U from A.

$$A = \begin{bmatrix} | \\ \boldsymbol{l}_1 \\ | \end{bmatrix} \begin{bmatrix} -\boldsymbol{u}_1^* - \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & & \\ 0 & & A_2 \end{bmatrix} = \begin{bmatrix} | \\ \boldsymbol{l}_1 \\ | \end{bmatrix} \begin{bmatrix} -\boldsymbol{u}_1^* - \end{bmatrix} + \begin{bmatrix} | \\ \boldsymbol{l}_2 \\ | \end{bmatrix} \begin{bmatrix} -\boldsymbol{u}_2^* - \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & A_3 \end{bmatrix} = LU$$

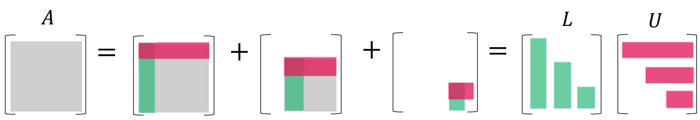


Figure 13: Recursive Rank 1 Matrix Peeling from A

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.

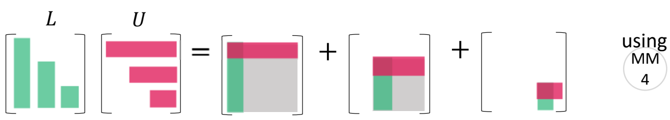


Figure 14: LU rebuilds A

To rebuild A from L times U, use column-row multiplication.

6.3 $\boldsymbol{A = QR}$

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

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

In Gram-Schmidt, the normalized $\boldsymbol{a}_1$ is $\boldsymbol{q}_1$. Then $\boldsymbol{a}_2$ is adjusted to be perpendicular to $\boldsymbol{q}_1$ to create $\boldsymbol{q}_2$. This procedure gives:

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

In the reverse direction, let $r_{ij} = \boldsymbol{q}_i^{\mathrm{T}}\boldsymbol{a}_j$ and you will get:

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

The original A becomes QR: orthogonal Q times upper triangular R.

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

$$
QQ^{\mathrm{T}} = Q^{\mathrm{T}}Q = I
$$

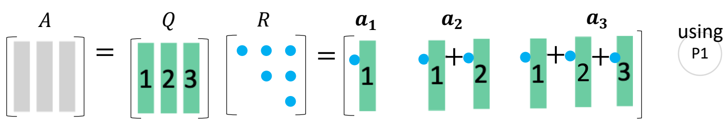


Figure 15: $A = QR$

Each column vector of A can be rebuilt from Q and R.
See Pattern 1 (P1) again for the graphic interpretation.

6.4 $S = Q\Lambda Q^{\mathrm{T}}$

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

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

$$
S = Q\Lambda Q^{\mathrm{T}} = \begin{bmatrix} | & | & | \\ \boldsymbol{q}_1 & \boldsymbol{q}_2 & \boldsymbol{q}_3 \\ | & | & | \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3 \end{bmatrix} \begin{bmatrix} -\boldsymbol{q}_1^{\mathrm{T}}- \\ -\boldsymbol{q}_2^{\mathrm{T}}- \\ -\boldsymbol{q}_3^{\mathrm{T}}- \end{bmatrix}
$$

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

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

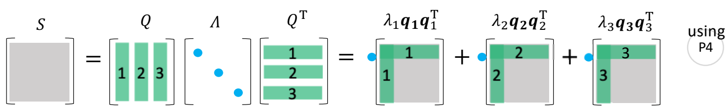


Figure 16: $S = Q\Lambda Q^{\mathrm{T}}$

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

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

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

6.5 $\boldsymbol{A = U\Sigma V^{\mathrm{T}}}$

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

Every matrix (including rectangular one) has a singular value decomposition (SVD). $A = U\Sigma V^{\mathrm{T}}$ has the singular vectors of A in U and V. The following figure illustrates the 'reduced' SVD.

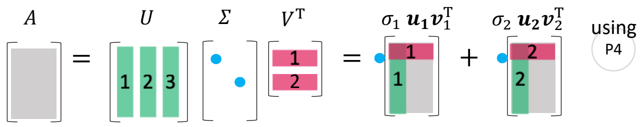


Figure 17: $A = U\Sigma V^{\mathrm{T}}$

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

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

Note that:

$$
\begin{gathered}
UU^{\mathrm{T}} = I_m \\
VV^{\mathrm{T}} = I_n
\end{gathered}
$$

See Pattern 4 (P4) for the graphic notation.

Conclusion and Acknowledgements

I have presented a systematic visualization of matrix/vector multiplication and its applications to the Five Matrix Factorizations. I hope you enjoy them and find them useful in understanding Linear Algebra.

Ashley Fernandes helped me with typesetting, which makes this paper much more appealing and professional.

To conclude this paper, I'd like to thank Prof. Gilbert Strang for publishing "Linear Algebra for Everyone". It presents a new pathway 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.

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, 6th ed., `http://math.mit.edu/linearalgebra`

3. Kenji Hiranabe(2021), *Map of Eigenvalues*, Slidedeck, `https://github.com/kenjihiranabe/The-Art-of-Linear-Algebra/blob/main/MapofEigenvalues.pdf`

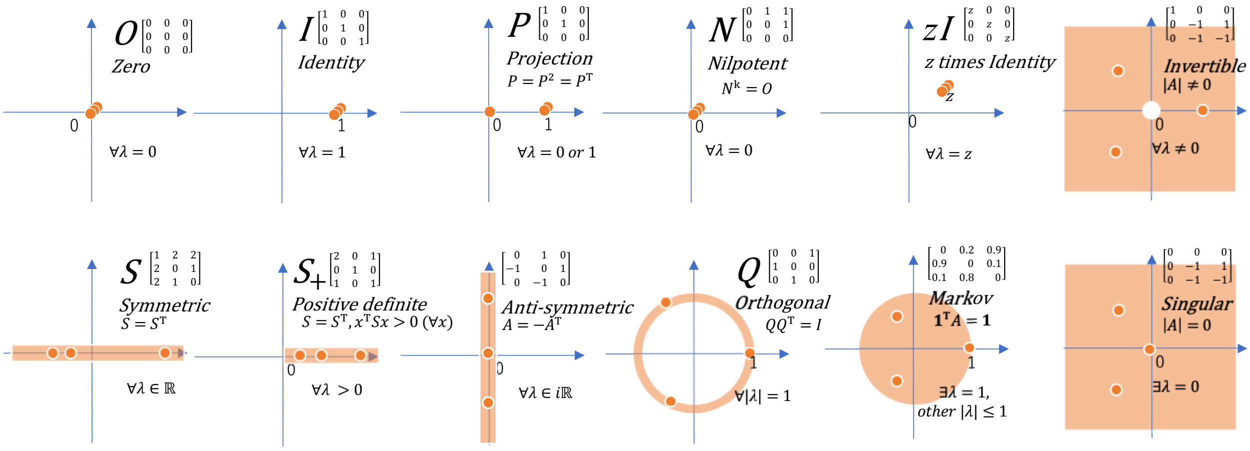


Figure 18: Map of Eigenvalues

4. Kenji Hiranabe(2020), *Matrix World*, Slidedeck, `https://github.com/kenjihiranabe/The-Art-of-Linear-Algebra/blob/main/MatrixWorld.pdf`

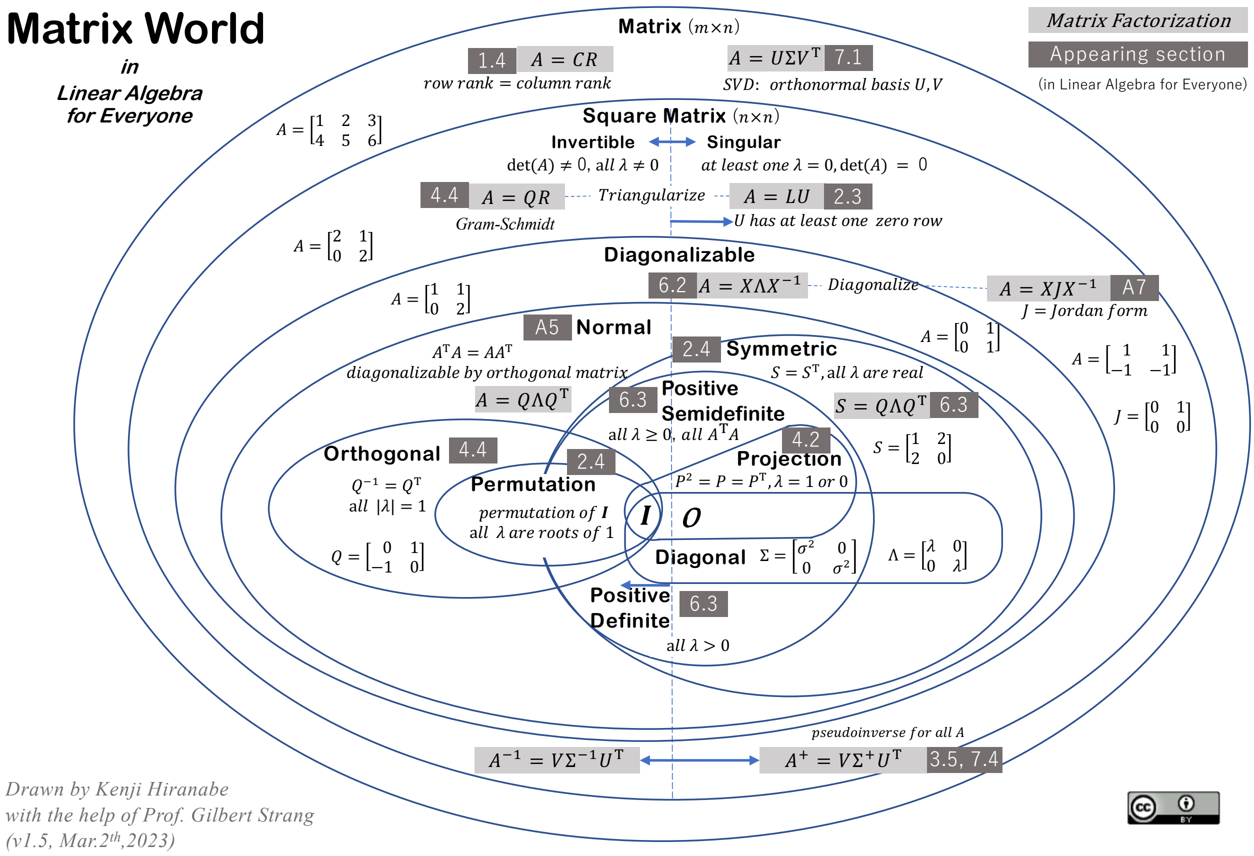


Figure 19: Matrix World

5. Gilbert Strang, artwork by Kenji Hiranabe, *The Four Subspaces and the solutions to* $A\boldsymbol{x} = \boldsymbol{b}$

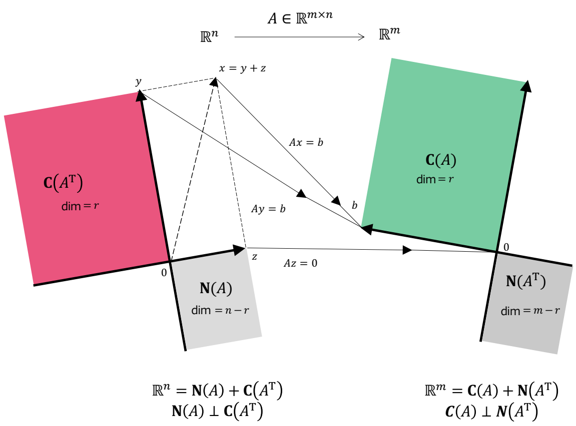


Figure 20: The Four Subspaces and the solutions to $A\boldsymbol{x} = \boldsymbol{b}$