Linear Algebra

Oran Danon

July 2017

The following is based on the references list, and lectures taken in Ariel University.

§1. Vectors.

We start by introducing the most basic object in linear algebra *a vector*. We introduce what a vector is using three perspectives:

1. **Physics student.** Vectors are arrows floating in space. They have length and direction.
2. **CS student.** Vectors are order list of numbers.
3. **Mathematics student.** Elements of a vector space. Roughly speaking, mathematicians tries to general all their ideas. It turns out that if a group with two operation satisfies certain properties all results of linear algebra applies to them. From the beginner's prospective it is recommended to think of $\mathbb{R}^n$ with the operations of addition and multiplications. Strictly speaking, a vector space is a nonempty set V of elements called vectors with two operations called addition and multiplication that satisfied the following:
   - Addition is commutative, associative, and has an identity element. Every element has an additive inverse.
   - Scalar multiplication is associative, and it has an identity element called $1 \in \mathbb{F}$ that satisfied $1 \cdot v = v$, for all $v \in V$.
   - Addition and multiplication by scalar are connected by distributive properties. That is, $a(u + v) = au + av$, and $(a + b)v = av + bv$ for all $a, b \in \mathbb{F}$ and all $u, v \in V$.

§2. The Complex Numbers

The complex numbers are numbers of the form $a + bi$ where i is a number that satisfies $i^2 = -1$. Some writes $i = \sqrt{-1}$, but square roots defined only for positive integer and hence this writing is problematic and indeed the rules for squares doesn't hold here.

- $\overline{z \pm w} = \overline{z} \pm \overline{w}$
- $\overline{z \cdot w} = \overline{z} \cdot \overline{w}$
- $\overline{\left(\frac{z_1}{z_2}\right)} = \frac{\overline{z_1}}{\overline{z_2}}$
- $|z_1 \cdot z_2| = |z_1| \cdot |z_2|$
- $\left|\frac{z_1}{z_2}\right| = \frac{|z_1|}{|z_2|}$
- $z_1 \cdot z_2 = r_1 r_2 \left(cis\left(\theta_1 + \theta_2\right)\right)$
- $\frac{z_1}{z_2} = \frac{r_1}{r_2}\left(cis\left(\theta_1 - \theta_2\right)\right)$
- $z^n = r^n \left(cis\left(n\theta\right)\right)$
- $\sqrt[n]{z} = \sqrt[n]{r}\left(cis\left(\frac{\theta + 2\pi k}{n}\right)\right)$
- $e^{i\theta} = \cos\theta + i\sin\theta$

Note: θ needs to be determined by its quarter

§3. Matrices & Linear Transformations

§ 3.1 What is a linear transformation. A linear transformation from V to W is a function $T\colon V \to W$ with the following properties.

- Additivity.
$$T(u+v) = T(u) + T(v).$$
- Homogeneity. For all $v \in V$, it holds that
$$T(\lambda v) = \lambda \cdot T(v), \text{ for all } \lambda \in \mathbb{F}.$$

§ 3.2 Properties of linear map.

- $T(0) = 0$, for all $T\colon V \to W$, and any linear spaces V, W. Put another way, the origin maps to the origin. To see that, note that
$$T(0) = T(0) + T(0) = T(0) + T(0),$$
add the additive inverse of $T(0)$ to both side yields $T(0) = 0$.
- Parallel lines that are evenly spaces remains parallel and evenly spaces. That is, ratios and angles between lines remains the same.

§ 3.3 Matrices. Let $m, n \in \mathbb{N}$. An $m \times n$ matrix A is a rectangular array of elements of $\mathbb{F}$ with m rows and n columns:
$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}.$$
We denote by a_{ij} the entry in the i^{th} row and j^{th} column.

§3.3.1 Matrix of a linear map. Let T be a linear map from U to V. Let $u_1, \ldots, u_n$ be a basis for U and $v_1, \ldots, v_m$ a basis for V. Then the matrix of T, denoted by A is defined as follows:
$$T(u_k) \coloneqq \sum_{i=1}^{m} a_{ik} v_i.$$
Put another way the vector a_{*k} denotes the representation of u_k in the new basis.

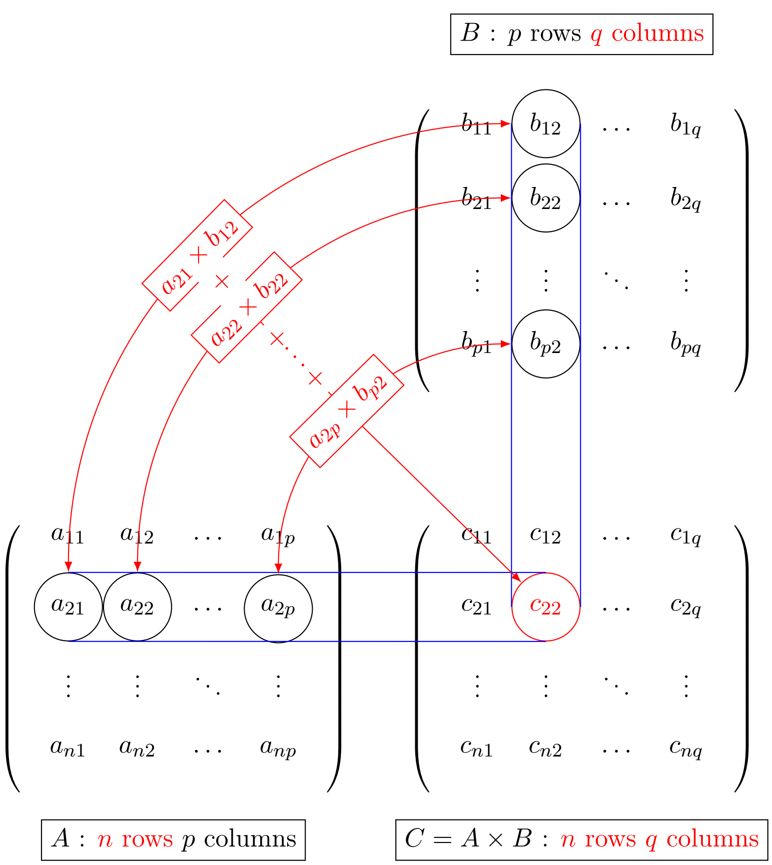
§3.3.2 Why do we like linear transformations, and why do we define the matrix of the linear map as we did. Simply it suffice to track where the basis lends. To see that, suppose T is a matrix denote a linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$. Suppose in addition that $\hat{i}, \hat{j}$ are the basis before the transformation. Then let $v = a\hat{i} + b\hat{j}$, for some $a, b \in \mathbb{F}$. Then we can note that
$$Tv = T(a\hat{i} + b\hat{i}) = aT(\hat{i}) + bT(\hat{j}).$$
If we write the matrix T such that the i^{th} column (vector, denoted by a_{*i}) is the coordinate of where the vector landed after the transformation, and let $v = \lambda_1 a_{*1} + \lambda_2 a_{*2} + \ldots + \lambda_n \ a_{*n}$, then we denote the above calculation by:
$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \lambda_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \lambda_2 \begin{pmatrix} a_{12} \\ \vdots \\ a_{m2} \end{pmatrix} + \ldots + \lambda_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} \lambda_1 a_{11} & \cdots & \lambda_n a_{1n} \\ \vdots & \ddots & \vdots \\ \lambda_1 a_{m1} & \cdots & \lambda_n a_{mn} \end{pmatrix}$$

§3.3.3 Matrix multiplication. Assuming $S\colon U \to V$ and $T\colon V \to U$, then we to look on the composition of $ST \coloneqq S \circ T$ as applying T and then S. That is, $(S \circ T)(v) \coloneqq S\left(T(u)\right)$, for all $u \in U$.

$$
\begin{aligned}
(ST)\, u_k &= S\left(T\left(u_k\right)\right) \\
&= S\left(\sum_{i=1}^{m} M_{ik}^{(T)} \cdot v_k\right) \\
&= \sum_{i=1}^{m} M_{ik}^{(T)} \cdot S\left(u_i\right) \\
&= \sum_{i=1}^{m} M_{ik}^{(T)} \cdot \sum_{j=1}^{l} M_{ji}^{(S)} w_j \\
&= \sum_{i=1}^{m} \left(\sum_{i=1}^{m} M_{ji}^{(S)} M_{ik}^{(T)}\right) w_j,
\end{aligned}
$$

where the third equality holds by the definition of a linear map, and the fifth follows by changing the order of summation which indeed valid as the sum is finite. Hence we derive the following formulas for matrix multiplications. Let $A_{n\times p}, B_{p\times q} \in \mathcal{M}(\mathbf{R})$, then $C = A \cdot B$ can be generating in the following ways:

1. Row by Column. $C_{ij} \coloneqq \sum_{k=1}^{p} a_{ik} b_{kj}$.



2. The columns of C are linear combination of the columns of A. To that end, notice that $c_{11} =$

$\langle a_{1*}, b_{*1}\rangle, c_{21} = \langle a_{2*}, b_{*1}\rangle, \ldots, c_{n1} = \langle a_{n*}, b_{*1}\rangle$, that is $Ab_{*1} = c_{*1}$.

$$A \begin{pmatrix} | & \cdots & | \\ b_{*1} & \cdots & b_{*q} \\ | & \cdots & | \end{pmatrix} = \begin{pmatrix} | & \cdots & | \\ Ab_{*1} & \cdots & Ab_{*q} \\ | & \cdots & | \end{pmatrix}$$

3. The rows of C are linear combination of the rows of B. Consider c_{11}, and note that $c_{11} = \langle a_{1*}, b_{*1}\rangle, c_{12} = \langle a_{1*}, b_{*2}\rangle, \ldots, c_{1q} = \langle a_{1*}, b_{*q}\rangle$. Put another way, $a_{1*}B = c_{1*}$.

$$\begin{pmatrix} — & a_{1*} & — \\ \vdots & \vdots & \vdots \\ — & a_{n*} & — \end{pmatrix} B = \begin{pmatrix} — & a_{1*}B & — \\ \vdots & \vdots & \vdots \\ — & a_{n*}B & — \end{pmatrix}$$

4. Sum of the dot product between the columns of A and the columns of B.

$$\begin{pmatrix} | & \cdots & | \\ a_{*1} & \cdots & a_{*p} \\ | & \cdots & | \end{pmatrix} \begin{pmatrix} — & b_{1*} & — \\ \vdots & \vdots & \vdots \\ — & b_{p*} & — \end{pmatrix} = \sum_{i=1}^{p} a_{i*}b_{*i}.$$

Note that this is sum of p matrices of rank 1.

5. Block matrices.

$$\left(\begin{array}{c|c} A_{1,1} & A_{1,2} \\ \hline A_{2,1} & A_{2,2} \end{array}\right) \left(\begin{array}{c|c} B_{1,1} & B_{1,2} \\ \hline B_{2,1} & B_{2,2} \end{array}\right) = \left(\begin{array}{c|c} A_{1,1}B_{1,1} + A_{1,2}B_{2,1} & A_{1,1}B_{1,2} + A_{1,2}B_{2,2} \\ \hline A_{2,1}B_{1,1} + A_{2,2}B_{2,1} & A_{2,1}B_{1,2} + A_{2,2}B_{2,2} \end{array}\right)$$

§3.3.4 Algebraic properties of product of linear transformations.

1. Identity.
$$TI = T = IT.$$

2. Associativity.
$$T_1(T_2T_3) = (T_1T_2)T_3.$$

3. Distributive properties.
$$(S_1 + S_2)T = S_1T + S_2T, \quad S(T_1 + T_2) = ST_1 + ST_2,$$
whenever $T, T_2, T_3 \colon U \to V,\ S, S_1, S_2 \colon V \to W$.

§4. The following are equivalent

- $A\vec{x} = \vec{b}$, has one solution $\vec{x} = A^{-1}\vec{b}$
- The sole solution to $A\vec{x} = \vec{0}$ is the trivial solution $\vec{0}$.
- All eigenvalues are nonzero
- $|A| \neq 0$
- The rows are independent
- The column are independent
- The row space is all of $\mathbb{R}^n$
- The column space is all of $\mathbb{R}^n$

§5. The Four Subspaces

The Four Subspaces				
Name	**Symbols**	**Definition**	**Dimension**	**Subspace of**
Column space Image of A	$\mathbf{C}(A), \mathbf{Im}(A)$	$\mathbf{C}(A) := \{Ax : x \in \mathbb{R}^n\}$	$\dim(\mathbf{C}(A)) = r$	$\mathbf{C}(A) \subseteq \mathbb{R}^m$
Kernel Null space	$\mathbf{ker}(A)$ $\mathbf{Null}(A), \mathbf{N}(A)$	$\mathbf{N}(A) := \left\{ \begin{array}{c} x \in \mathbb{R}^n : \\ Ax = \mathbf{0} \end{array} \right\}$	$\dim(\mathbf{N}(A)) = n - r$	$\mathbf{ker}(A) \subseteq \mathbb{R}^n$
Row space	$\mathbf{R}(A) = \mathbf{C}(A^T)$	$\mathbf{C}(A^T)$	$\dim\left(\mathbf{C}\left(A^T\right)\right) = r$	$\mathbf{R}(A) \subseteq \mathbb{R}^n$
Left null space	$\mathbf{N}(A^T)$	$N(A^T) = \left\{ \begin{array}{c} x \in \mathbb{R}^n : \\ x^T A = \mathbf{0} \end{array} \right\}$	$\dim\left(\mathbf{N}\left(A^T\right)\right) = m - r$	$\mathbf{N}(A^T) \subseteq \mathbb{R}^m$

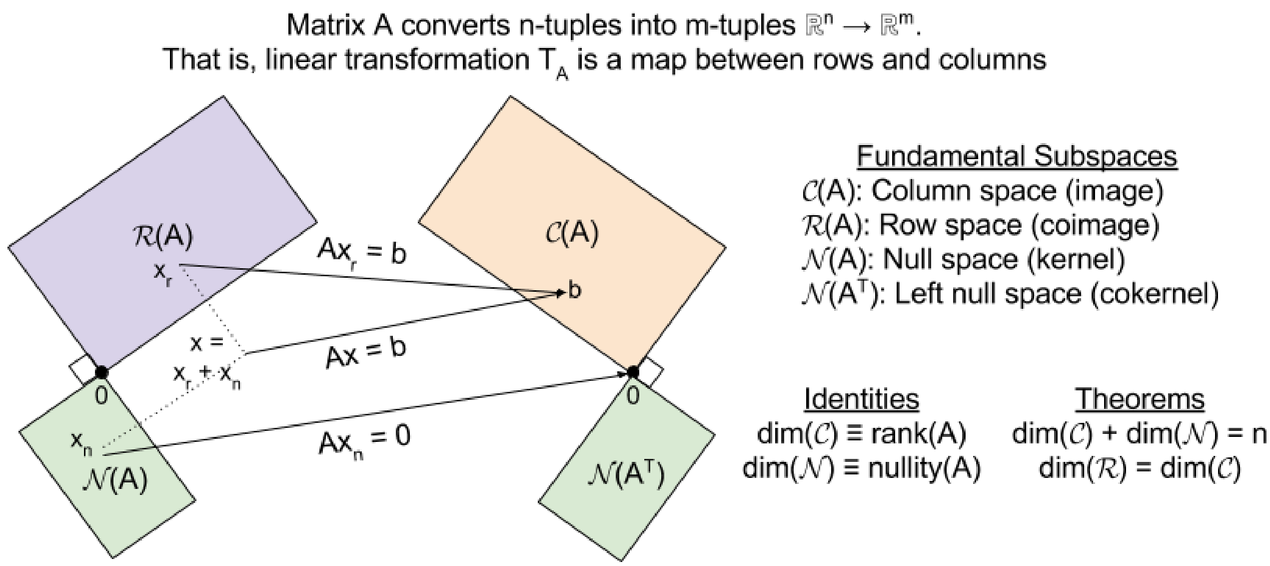


Figure 5.1: The Four Subspaces

Determinate of $\mathcal{M}_n$ Matrix		
Elementary row operation	**Elementary column operation**	**Effect on determinant**
$R_i \leftrightarrow R_j$	$C_i \leftrightarrow C_j$	change the sign of the determinant
$R_i \to kR_i,\ k \in \mathbb{R} \setminus \{0\}$	$C_i \to kC_i,\ k \in \mathbb{R} \setminus \{0\}$	multiplies the determinant by k
$R_i \to R_i + kR_j,\ j \neq i$	$C_i \to R_i + kR_j,\ j \neq i$	no effect on the determinant

Let $T\colon \underset{\Cup}{\mathbb{R}^n} \to \underset{\Cup}{\mathbb{R}^m}$, and let $\mathcal{M}_{m\times n}$ be the corresponding matrix, Then:

$\mathcal{M}_{m\times n} \cdot \vec{x}_{n\times 1} = \vec{x}_{m\times 1}$

§6. CR-Decomposition.

Every matrix A can be decompose into two matrix C, R such that $A = CR$ and:

1. The column of C are independent.
2. The rows of R are independent.
3. It turns out that R is the reduce echelon form.

§6.0.1 Example.. Note that

$$A = \begin{bmatrix} 1 & 4 & 5 \\ 3 & 2 & 5 \\ 2 & 1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 3 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

We take the first two independent columns of A and then express all the columns of A using R. Indeed, that R tells what are the coefficients for obtaining each column of A:

$$\begin{bmatrix} 1 & 4 \\ 3 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = a \cdot \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} + b \cdot \begin{bmatrix} 4 \\ 2 \\ 1 \end{bmatrix}.$$

§7. LU-factorization.

Algorithm Solve $Ax = b$ by elimination Factor $A = LU$

procedure LU-factorization(A)
  $A_0 \leftarrow A$
  for $i = 1, \dots, n$ **do**
    $A_i, \ell_{i*} \leftarrow$ Reduce the i^{th} row & column(A_{i-1}, i)
  return

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \ell_{21} & 1 & 0 & 0 \\ \vdots & & \ddots & 0 \\ \ell_{n1} & \ell_{n2} & \dots & 1 \end{bmatrix} \begin{bmatrix} \text{row 1 of } A_0 & & & \\ 0 & \text{row 1 of } A_1 & & \\ 0 & 0 & \ddots & \\ 0 & 0 & \dots & \text{row 1 of } A_n \end{bmatrix}$$

procedure Reduce the i^{th} row & column(A, i)
  for $j = 1, \dots, n$ **do**
    Subtract ℓ_{ij} times row j from row i to produce zeros in column 1.
  Comment. After the first step we have:

$$A = \begin{bmatrix} 1 \\ \ell_{21} \\ \vdots \\ \ell_{n1} \end{bmatrix} \begin{bmatrix} \text{row 1 of } A \end{bmatrix} + \begin{bmatrix} 0 & \dots & \dots & 0 \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{bmatrix}$$

return A_i and the ℓ_{*i} vector.

§8. QR-factorization

§9. Eigen-Decomposition

§10. SVD - Singular Value Decomposition

We consider the solutions to

$$Av_i = \sigma_i u_i.$$

Put another way, $AV = U\Sigma$ with $U^T U = I$ and $V^T V = I$. This means that

$$A \begin{bmatrix} v_1 & \dots & v_r \end{bmatrix} = \begin{bmatrix} u_1 & \dots & u_r \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_r \end{bmatrix}.$$

Note that r is the rank of A.

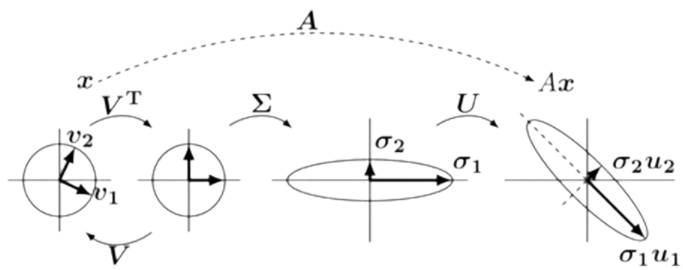


Figure 10.1: SVD

References

[1] S. Axler. *Linear algebra done right*. Undergraduate Texts in Mathematics. Springer, Cham, third edition, 2015.

[2] G. Sanderson. Essence of linear algebra.

[3] G. Stang. Mit a 2020 vision of linear algebra, spring 2020.

[4] G. Strang. *Introduction to linear algebra*. Wellesley Cambridge Press, fifth edition edition, 2016.