Reflections on Linear Algebra: Range and Nullspace

This article assumes readers have already learned basic linear algebra for at least one time.

By Tongyu Lu, written on 13/01/2021

Range and Nullspace

Here we do a brief introduction about the image space (range) and null space (kernel space) of a matrix (or a linear transform) $A_{M\times N}$. The range (or the image space) of A is defined as $R(A) = \mathrm{im}(A) = \{Ax : x \in \mathbb{R}^N\}$, null space (or kernel space) of A is defined as $N(A) = \ker(A) = \{x: Ax = 0, x \in \mathbb{R}^N\}$.

$R(A)$ is subspace of $\mathbb{R}^M$ and $N(A)$ is subspaces of $\mathbb{R}^N$. This can be shown by: $Ax_1 + Ax_2 = A(x_1 + x_2)$.

Now, let us consider the exact meaning of $R(A)$ and $N(A)$. Take A as a linear transformation $A: \mathbb{R}^N \to \mathbb{R}^M$.

- First, consider the meaning of null (kernel) space. Reflect for a while about the naming of kernel space. It means that: any vector in $\ker(A)$ is compressed into the same zero vector when applying A, which is like a kernel.
  If A is injective (which means that "if $Ax_1 = Ax_2$, then $x_1 = x_2$"), then it is safe to say that the null space is $\{0\}$. Why? If $\dim(N(A)) \neq 0,$ then there exist non-zero vectors $v_1, v_2 \in N(A), v_1 \neq v_2$, and then we have (by definition) $Av_1 = 0, Av_2 = 0$, therefore $Av_1 = Av_2$. By the injective assumption, $v_1 = v_2$, which leads to contradiction.
  In the linear equality context, when $N(A) = \{0\}$, then $Ax = 0$ only has zero solution. Therefore, $N(A)$ is also the solution space of $Ax = 0$.
- Then, consider the meaning of the range (or the image space).
  The range of A is straightforward to understand: it consists of the images of linear transformation A.
  If A is surjection (which means that "$\forall b \in \mathbb{R}^M, \exists x \in \mathbb{R}^N$ s.t. $Ax = b$"), then $\dim(R(A)) = M$. How to understand this? It is better to write $A = [a_1, \dots, a_N], a_i \in \mathbb{R}^M$. Ax is actually $a_1x_1 + \dots + a_Nx_N$, which is linear combination of column vectors of A. Therefore, $\mathrm{rank}(a_1, \dots, a_N) = \dim(\mathrm{span}(a_1, \dots, a_N)) = \dim(R(A))$. Then according to the surjection assumption, $a_1x_1 + \dots + a_Nx_N$ could fill up the whole $\mathbb{R}^M$, which means that the N basis vectors $a_1, \dots, a_N$ has rank M.

Now, we move onto a very meaningful theorem: $\dim(R(A)) + \dim(N(A)) = N$

Why?

Before justifying this, we detour a little…

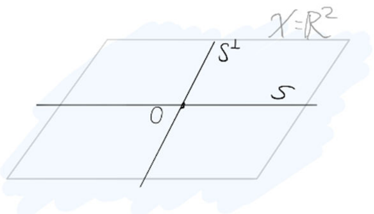
Orthogonal Complement of a Subspace

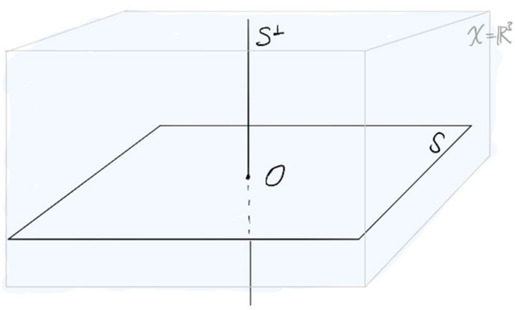
This part assumes readers have known the concepts of inner product and vector orthogonal.

Assume that S is a subspace of an inner-product space $\mathcal{X}$. A vector x is said to be orthogonal to a subspace S iff $\forall s \in S, \langle x, s\rangle = 0$. This relationship is denoted as $x \perp S$.

The orthogonal complement of S is defined as $S^\perp = \{x: x \perp S, x \in \mathcal{X}\}$.

This leads to an important theorem: any inner-product space $\mathcal{X}$ could be decomposed into a subspace $S \subseteq \mathcal{X}$ and its orthogonal complement $S^\perp$, written as $\mathcal{X} = S \oplus S^\perp = \{s + s^\perp : \forall s \in S, s^\perp \in S^\perp\}$





Proof is left out, but here shows several figures to illustrate this idea. The proof follows this logic line: suppose that $S \oplus S^\perp \neq \mathcal{X}$, then there exists a vector $z \in \mathcal{X}$ and $z \perp S$ but $z \notin S \oplus S^\perp$, then $z \in S^\perp$, which says $z \in S \oplus S^\perp$.

Here is a property: $S \cap S^\perp = \{0\}$. This is because if $v \in S \cap S^\perp$, then $\langle v, v\rangle = 0$

Orthogonal Complement of Range/Nullspace

Equipped with the concept of orthogonal complement, we are now able to discover the relationship between range and nullspace of A.

We could observe that $\forall x \in R(A^T), \forall z \in N(A)$, there must be $\langle x, z \rangle = 0$. This could be proved by:

$$x^T z = (A^T y)^T z = y^T (Az) = 0$$

Therefore, we have $N(A) \perp R(A^T)$

Or equivalently, $R(A^T) = N(A)^\perp$

Applying the orthogonal complement direct sum theorem in the previous section, we have:

$$\mathbb{R}^N = N(A) \oplus N(A)^\perp = N(A) \oplus R(A^T)$$

Symmetrically, we study

$$R(A)^\perp = \{y \in \mathbb{R}^M : y^T z = 0, \forall z \in R(A)\} = \{y \in \mathbb{R}^M : y^T A x = 0, \forall x \in \mathbb{R}^N\} = \{y \in \mathbb{R}^M : A^T y = 0\} = N(A^T)$$

Therefore, $R(A) \perp N(A^T)$

Applying the orthogonal complement direct sum theorem in the previous section, we have:

$$\mathbb{R}^M = R(A) \oplus R(A)^\perp = R(A) \oplus N(A^T)$$

Summarizing the discussions above, we have **the fundamental theorem of linear algebra**:

$$R(A^T) = N(A)^\perp, \qquad R(A)^\perp = N(A^T)$$
$$\mathbb{R}^N = N(A) \oplus R(A^T), \qquad \mathbb{R}^M = R(A) \oplus N(A^T)$$
$$\dim\big(N(A)\big) + \operatorname{rank}(A) = N, \qquad \dim\big(N(A^T)\big) + \operatorname{rank}(A) = M$$
$$\dim\big(N(A)\big) + \dim\big(N(A)^\perp\big) = N, \qquad \dim\big(R(A)^\perp\big) + \dim\big(R(A)\big) = M$$
$$\dim\big(N(A)\big) + \dim\big(R(A)\big) = N, \qquad \dim\big(R(A)^\perp\big) + \dim\big(N(A)^\perp\big) = M$$

The last three lines are because: the row rank is equal to the column rank:

$$\operatorname{rank}(A) = \dim\big(R(A)\big) = \dim\big(R(A^T)\big) \leq \max\{M, N\}$$

but the relationship between $\dim\big(N(A)\big), \dim\big(N(A^T)\big)$ are not so straightforward.

This theorem is beautiful for its high symmetry! In addition, it could explain a lot of phenomenon in the world of linear equality.

Applications of Range/Nullspace

- If $M < N$, then A cannot be injective.
  How comes? $\dim\big(N(A)\big) = N - \operatorname{rank}(A) \geq N - M > 0$, which says $N(A)$ has non-zero vectors.
- Given linear equation $Ax = 0$ (N variables and M equations). If the number of equations is smaller than the number of variables ($M < N$), then there exist non-zero solutions.
- If $M > N$, then A cannot be surjective.
  How comes? $\dim\big(R(A)\big) = N - \dim\big(N(A)\big) \leq N < M$, which says $R(A)$ cannot fill up the whole image space.
- Given linear equation $Ax = b$ (N variables and M equations). If the number of equations is larger than the number of variables ($M > N$), then the equation may be infeasible. $Ax = b$ is infeasible iff $b \in \mathbb{R}^M \backslash R(A)$.

Reflections: all the discussions above are happening within the real number field $\mathbb{F} = \mathbb{R}$ and the vector spaces are Euclidian. We did not discuss the situation when matrix (transformation) A is abstract. I may step onto that further.

References

[1]Optimization Models, Book, G.C. Calafiore and L. El Ghaoui, Cambridge University Press, October 2014

[2] A Chinese Article

https://github.com/frank1ma/LinearAlgebraQuickReview/blob/master/pdf/Introduction_to_Linear_Algebra-ch03.pdf