Holger Broman, August 19, 1999

Matrices, geometry, and mappings

Subspaces, Range and Null spaces

A subspace of a vector space is a part of the vector space that obeys

- if $\mathbf{x}$ and $\mathbf{y}$ belong to the subspace, so does $\mathbf{x} + \mathbf{y}$
- if $\mathbf{x}$ belongs to the subspace, so does $k \cdot \mathbf{x}$, k scalar (real-valued so far)

A subspace is a subset of a vector space, which is also a vector space in its own right.

Examples: The following spaces are all subspaces of $\mathbb{R}^3$:

- the origin
- any single axis
- any plane spanned by two of the axes
- the entire $\mathbb{R}^3$

Note 1: The origin belongs to every subspace. Here is the proof. If $\mathbf{x}$ belongs to the subspace, so does $-\mathbf{x}$, and thus does $\mathbf{x} - \mathbf{x} = 0$.

Two subspaces are said to be orthogonal if all vectors in one are orthogonal to all vectors in the other.

Example: The two subspaces of $\mathbb{R}^2$ spanned by $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ are orthogonal.

If the union of two orthogonal subspaces span the original space, then any one of the subspaces is said to be the orthogonal complement of the other.

Example: Same as above.

——————— o O o ———————

Consider the equation

$$\mathbf{y} = A\,\mathbf{x}, \qquad \mathbf{x} \in \mathbb{R}^N.$$

Re-write it

$$\begin{aligned}\mathbf{y} =& A\,\mathbf{x} = [\mathbf{a}_1, \dots, \mathbf{a}_N]\ \mathbf{x} = \\ =& \sum_{n=1}^{N} x_n\,\mathbf{a}_n\end{aligned}$$

As you can see, $\mathbf{y}$ is some linear combination of the column vectors of A. Thus, $\mathbf{y}$ belongs to the space spanned by these column vectors. If A is $M \times N$, then this space is some subspace of $\mathbb{R}^M$.

> The range-space of A, denoted $\mathcal{R}(A)$, is the space spanned by the column vectors of A.

Note 2: $\dim[\mathcal{R}(A)] \leq \min(M, N)$.

Proof: As $\mathbf{a}_1, \dots, \mathbf{a}_N$ are of length M, then they can span a space of dimension at most $\min(M, N)$.

> The dimension of the column space of A is called the rank of A, denoted r.
>
> $$\dim[\mathcal{R}(A)] = r$$

Note 3: The rank r equals the number of linearly independent column vectors of A. Actually, this number also equals the number of linear independent row vectors of A.

Now, consider the equation

$$A\mathbf{x} = 0.$$

> The vectors x that satisfy $A\mathbf{x} = 0$ define the null-space of A, denoted $\mathcal{N}(A)$.

Note 4: $\dim[\mathcal{N}(A)] \leq N$.

Proof: $\mathbf{x} \in \mathbb{R}^N$

The following theorem is fundamental in linear algebra.

Theorem: Let A be an $M \times N$ matrix. Then

$$\dim[\mathcal{R}(A)] + \dim[\mathcal{N}(A)] = N$$

Proof[1]**:** The null space has a basis, say $\mathbf{e}_1, \dots, \mathbf{e}_d$, where $d = \dim[\mathcal{N}(A)]$. Expand this to a basis for $\mathbb{R}^N$: $\mathbf{e}_{d+1}, \dots, \mathbf{e}_N$. Any vector $\mathbf{x}$ in $\mathbb{R}^N$ can be expressed in this basis

$$\mathbf{x} = \sum_{n=1}^{N} x_n \ \mathbf{e}_n.$$

Now, $A\mathbf{x}$ can be expanded

$$A\mathbf{x} = \sum_{n=1}^{N} x_n \ A\mathbf{e}_n = \sum_{n=d+1}^{N} x_n \ A\mathbf{e}_n,$$

as $A\mathbf{e}_n$ equals zero for $n = 1, \dots, d$. The key idea of the proof is to show that $A\mathbf{e}_n$, $n = d{+}1, \dots, N$ is a basis for the range space of A which thus has dimension $N - d$.

The range space of A is obviously spanned by $A\mathbf{e}_n$, $n = 1, \dots, N$, and thus by $A\mathbf{e}_n$, $n = d+1, \dots, N$ as $A\mathbf{e}_n = 0, n = 1, \dots, d$. Remains to show that $A\mathbf{e}_n$, $n = d+1, \dots, N$ are linearly independent. Assume they are not, i.e. that there exist scalars λ_n, $n = d+1, \dots, N$ so that

$$\sum_{n=d+1}^{N} \lambda_n \ A\mathbf{e}_n = 0.$$

This implies that

$$A \sum_{n=d+1}^{N} \lambda_n \, \mathbf{e}_n = A\mathbf{y} = 0,$$

where $\mathbf{y}$ does not belong to the null space of A as it is not expressible in the basis of the null space $\mathbf{e}_n$, $n = 1, \dots, d$. This contradiction completes the proof.

This theorem describes to what extent the mapping A is invertible; on the range space $\mathcal{R}(A)$ an inverse may be defined. The ambiguity (no inverse exists) is embodied in the null space $\mathcal{N}(A)$.

[1] Adapted from Jöran Bergh