MATRIX VECTOR SUBSPACES

MATRIX VECTOR SUBSPACES

SYNOPSIS. Operations with vectors have been formally defined by the vector space algebraic structure, and the idea of obtaining new vectors by linear combination has been concisely stated as matrix-vector multiplication. Matrices also describe mappings between vector spaces that preserve linear combinations. Can all vectors of interest be obtained by linear combination of some set of vectors? This is a natural question to ask, and is answered through the concept of vector subspaces associated with a matrix.

1. Vector subspaces

1.1. Vectors reachable by linear combination

A central interest in data science is to seek simple description of complex objects. A typical situation is that many instances of some object of interest are initially given as an m-tuple $\boldsymbol{b} \in \mathbb{R}^m$ with large m. Assuming that addition and scaling of such objects can be cogently defined, a vector space is obtained, say over the field of reals with an Euclidean distance, E_m . Examples include for instance recordings of medical data (electroencephalograms, electrocardiograms), sound recordings, or images, for which m can easily reach in to the millions. A natural question to ask is whether all the m real numbers are actually needed to describe the observed objects. Perhaps instead of specifying all the components of $\boldsymbol{b} \in \mathbb{R}^m$, it might be possible to state that \boldsymbol{v} is a linear combination of n < m vectors

$$\boldsymbol{b} = A\boldsymbol{x} = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \dots & \boldsymbol{a}_n \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \vdots \\ \boldsymbol{x}_n \end{bmatrix} = \boldsymbol{x}_1 \boldsymbol{a}_1 + \dots + \boldsymbol{x}_n \boldsymbol{a}_n.$$

It is then natural to formally define the set of all vectors \boldsymbol{b} that could thus be expressed, meaning that they can be reached by linear combination of the columns of \boldsymbol{A} .

DEFINITION. In vector space $\mathcal{V} = (V, S, +, \cdot)$, the span of vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n \in V$, is the set of vectors reachable by linear combination

$$\mathrm{span}\{\pmb{a}_1,\pmb{a}_2,\ldots,\pmb{a}_n\} = \{\pmb{b} \in V \mid \exists x_1,\ldots,x_n \in S \text{ such that } \pmb{b} = x_1 \, \pmb{a}_1 + \ldots + x_n \, \pmb{a}_n\}.$$

What now is the relationship between this set of reachable vectors and the entire vector space? The mathematical transcription of this question leads to a consideration of another algebraic structure.

DEFINITION. $\mathcal{U} = (U, S, +, \cdot), \ U \neq \emptyset$, is a vector subspace of vector space $\mathcal{V} = (V, S, +, \cdot)$ over the same field of scalars S, denoted by $\mathcal{U} \leq \mathcal{V}$, if $U \subseteq V$ and $\forall a, b \in S, \forall u, v \in U$, the linear combination $au + bv \in U$.

The above states a vector subspace must be closed under linear combination, and have the same vector addition and scaling operations as the enclosing vector space. The simplest vector subspace of a vector space is the null subspace that only contains the null element, $U = \{0\}$. In fact, any subspace must contain the null element $\mathbf{0}$, or otherwise closure would not be verified for the particular linear combination $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$. One can think of $\mathcal{Z} = (\{0\}, S, +, \cdot)$ as the smallest subspace of a vector space. By the above definition, \mathcal{U} is also a subspace of itself, intuitively, the largest subspace. If $U \subset V$, then \mathcal{U} is said to be a *proper subspace* of \mathcal{V} , denoted by $\mathcal{U} < \mathcal{V}$.

Setting n-m components equal to zero in the real space \mathcal{R}_m defines a proper subspace whose elements can be placed into a one-to-one correspondence with the vectors within \mathcal{R}_n . For example, setting component m of $\mathbf{x} \in \mathbb{R}^m$ equal to zero gives $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \dots & x_{m-1} & 0 \end{bmatrix}^T$ that, while not a member of \mathbb{R}^{m-1} , is in a one-to-one relation with $\mathbf{x}' = \begin{bmatrix} x_1 & x_2 & \dots & x_{m-1} \end{bmatrix}^T \in \mathbb{R}^{m-1}$. Dropping the last component of $\mathbf{y} \in \mathbb{R}^m$, $\mathbf{y} = \begin{bmatrix} y_1 & y_2 & \dots & y_{m-1} & y_m \end{bmatrix}^T$ gives vector $\mathbf{y}' = \begin{bmatrix} y_1 & y_2 & \dots & y_{m-1} \end{bmatrix} \in \mathbb{R}^{m-1}$, but this is no longer a one-to-one correspondence since for some given \mathbf{y}' , the last component y_m could take any value.

1.2. Vector space composition

Vector subspaces arise from the decomposition of a vector space, the idea of breaking up a complex object into component parts. The converse, composition of vector spaces $\mathcal{U} = (U, S, +, \cdot)$ $\mathcal{V} = (V, S, +, \cdot)$ is also defined in terms of linear combination. A vector $\mathbf{x} \in \mathbb{R}^3$ can be obtained as the linear combination

$$\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ x_2 \\ x_3 \end{bmatrix},$$

but also as

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 - a \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ a \\ x_3 \end{bmatrix},$$

for some arbitrary $a \in \mathbb{R}$. In the first case, x is obtained as a unique linear combination of a vector from the set $U = \{ [x_1 \ 0 \ 0]^T | x_1 \in \mathbb{R} \}$ with a vector from $V = \{ [0 \ x_2 \ x_3]^T | x_2, x_3 \in \mathbb{R} \}$. In the second case, there is an infinity of linear combinations of a vector from V with another from $W = \{ [x_1 \ x_2 \ 0]^T | x_1, x_2 \in \mathbb{R} \}$ to the vector x. This is captured by a pair of definitions to describe the two types of vector space composition.

MATRIX VECTOR SUBSPACES

DEFINITION. Given two vector subspaces $\mathcal{U} = (U, S, +, \cdot)$, $\mathcal{V} = (V, S, +, \cdot)$ of the space $\mathcal{W} = (W, S, +, \cdot)$, the direct sum is the vector space $\mathcal{U} \oplus \mathcal{V} = (U \oplus V, S, +, \cdot)$, where the direct sum of the two sets of vectors U, V is $U \oplus V = \{u + v : \exists! u \in U, \exists! v \in V\}$. (unique decomposition)

DEFINITION. Given two vector subspaces $\mathcal{U} = (U, S, +, \cdot)$, $\mathcal{V} = (V, S, +, \cdot)$ of the space $\mathcal{W} = (W, S, +, \cdot)$, the sum is the vector space $\mathcal{U} + \mathcal{V} = (U + V, S, +, \cdot)$, where the sum of the two sets of vectors U, V is $U + V = \{u + v : u \in U, v \in V\}$.

Since the same scalar field, vector addition, and scaling is used, it is convenient to refer to vector space sums simply by the sum of the vector sets U + V, or $U \oplus V$, instead of specifying the full 4-tuplet for each space. This shall be adopted henceforth to simplify the notation.

∴ u=[1; 0; 0]; v=[0; 2; 3]; vp=[0; 1; 3]; w=[1; 1; 0]; [u+v vp+w]

$$\begin{bmatrix} 1 & 1 \\ 2 & 2 \\ 3 & 3 \end{bmatrix}$$
∴ (4)

In the above computational example, the essential difference between the two ways to express $\mathbf{x} \in \mathbb{R}^3$ is that $U \cap V = \{\mathbf{0}\}$, but $V \cap W = \{[0 \ a \ 0]^T : a \in \mathbb{R}\} \neq \{\mathbf{0}\}$, and in general if the zero vector is the only common element of two vector spaces then the sum of the vector spaces becomes a direct sum. In general, the common elements of two vector subspaces can also be defined.

DEFINITION. Given two vector subspaces $(U, S, +, \cdot)$, $(V, S, +, \cdot)$ of the space $(W, S, +, \cdot)$, the intersection is the set $U \cap V = \{x | x \in U, x \in V\}$.

In practice, the most important procedure to construct direct sums or to check when an intersection of two vector subspaces reduces to the zero vector is through an inner product.

DEFINITION. Two vector subspaces U, V of the real vector space \mathbb{R}^m are orthogonal, denoted as $U \perp V$ if $\mathbf{u}^T \mathbf{v} = 0$ for any $\mathbf{u} \in U, \mathbf{v} \in V$.

DEFINITION. Two vector subspaces U,V of U+V are orthogonal complements, denoted $U=V^{\perp}, V=U^{\perp}$ if they are orthogonal subspaces, $U \perp V$, and $U \cap V = \{\mathbf{0}\}$, i.e., the null vector is the only common element of both subspaces.

Continuing the above computational example where the same vector was obtained through two different linear combinations, z = u + v = v' + w, the essential difference between the two is u is orthogonal to v, whereas v' is not orthogonal to w.

2. Vector subspaces of a linear mapping

2.1. Matrix column and left null spaces

The wide-ranging utility of linear algebra essentially results a complete characterization of the behavior of a linear mapping between vector spaces $f: U \to V$, $f(a\mathbf{u} + b\mathbf{v}) = af(\mathbf{u}) + bf(\mathbf{v})$. For some given linear mapping the questions that arise are:

- 1. Can any vector within *V* be obtained by evaluation of *f*?
- 2. Is there a single way that a vector within V can be obtained by evaluation of f?

Linear mappings between real vector spaces $f: \mathbb{R}^n \to \mathbb{R}^m$, have been seen to be completely specified by a matrix $A \in \mathbb{R}^{m \times n}$. It is common to frame the above questions about the behavior of the linear mapping f(x) = Ax through sets associated with the matrix A. To frame an answer to the first question, a set of reachable vectors is first defined.

DEFINITION. The column space (or range) of matrix $A \in \mathbb{R}^{m \times n}$ is the set of vectors reachable by linear combination of the matrix column vectors

$$C(A) = \text{range}(A) = \{ \boldsymbol{b} \in \mathbb{R}^m : \exists \boldsymbol{x} \in \mathbb{R}^n \text{ such that } \boldsymbol{b} = A\boldsymbol{x} \}.$$

By definition, the column space is included in the co-domain of the function f(x) = Ax, $C(A) \subseteq \mathbb{R}^m$, and is readily seen to be a vector subspace of \mathbb{R}^m . Having defined the set of vectors reachable by linear combination, two questions arise:

- 1. Is the column space the entire co-domain, $C(A) = \mathbb{R}^m$? This would signify that any vector can be reached by linear combination of columns of A.
- 2. What co-domain vectors are not reachable by linear combination of columns of A?

Consider the orthogonal complement of C(A) defined as the set vectors y orthogonal to all of the column vectors of $A = [a_1 \ a_2 \ \dots \ a_n]$, expressed through inner products as

$$a_1^T y = 0, a_2^T y = 0, \dots, a_n^T y = 0.$$

This can be expressed more concisely through the transpose operation

$$A = \begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix}, A^T y = \begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_n^T \end{bmatrix} y = \begin{bmatrix} a_1^T y \\ a_2^T y \\ \vdots \\ a_n^T y \end{bmatrix},$$

and leads to the definition of a set of vectors for which $A^Ty = 0$

DEFINITION. The left null space (or cokernel) of a matrix $A \in \mathbb{R}^{m \times n}$ is the set

$$N(\mathbf{A}^T) = \text{null}(\mathbf{A}^T) = \{ \mathbf{y} \in \mathbb{R}^m : \mathbf{A}^T \mathbf{y} = \mathbf{0} \}.$$

Note that the left null space is also a vector subspace of the co-domain of f(x) = Ax, $N(A^T) \subseteq \mathbb{R}^m$. The above definitions suggest that both the matrix and its transpose play a role in characterizing the behavior of the linear mapping f = Ax, so analogous sets are define for the transpose A^T .

Matrix Vector Subspaces 5

DEFINITION. The row space (or corange) of a matrix $A \in \mathbb{R}^{m \times n}$ is the set

$$R(A) = C(A^T) = \operatorname{range}(A^T) = \{c \in \mathbb{R}^n : \exists y \in \mathbb{R}^m c = A^T y\} \subseteq \mathbb{R}^n$$

DEFINITION. The null space of a matrix $A \in \mathbb{R}^{m \times n}$ is the set

$$N(A) = \text{null}(A) = \{x \in \mathbb{R}^n : Ax = 0\} \subseteq \mathbb{R}^n$$

2.2. Geometric description of subspaces

The concepts of Euclidean geometry are widely used to characterize subspaces of a vector space. Consider the familiar example of $\mathscr{E}_2 = (\mathbb{R}^2, \mathbb{R}, +, \cdot)$ the Euclidean 2-space or plane,

$$\mathbb{R}^2 = \left\{ \left[\begin{array}{c} x_1 \\ x_2 \end{array} \right] : x_1, x_2 \in \mathbb{R} \right\}.$$

As is common the vector space representing the plane is referred to either by its full name \mathscr{E}_2 or in shorthand form as \mathbb{R}^2 . The trivial subspaces of \mathbb{R}^2 are the zero vector space $\mathscr{Z} = (\{0\}, \mathbb{R}, +, \cdot)$, and \mathbb{R}^2 itself. Any line passing through the origin is a non-trivial subspace $\mathscr{Z} = (L, \mathbb{R}, +, \cdot)$ with

$$L_{p,q} = \left\{ \left[\begin{array}{c} ap \\ aq \end{array} \right] : a \in \mathbb{R} \right\}.$$

In particular the x_1, x_2 axes are

$$L_{1,0} = \left\{ \left[\begin{array}{c} a \\ 0 \end{array} \right] : a \in \mathbb{R} \right\}, L_{0,1} = \left\{ \left[\begin{array}{c} 0 \\ a \end{array} \right] : a \in \mathbb{R} \right\},$$

respectively. In \mathbb{R}^3 , the non-trivial subspaces are lines passing through the origin

$$L_{p,q,r} = \left\{ \left[\begin{array}{c} ap \\ aq \\ ar \end{array} \right] : a \in \mathbb{R} \right\},\,$$

and planes passing through the origin

$$P_{\mathbf{n}} = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} : \mathbf{n}^T \mathbf{x} = n_1 x_1 + n_2 x_2 + n_3 x_3 = 0, x_1, x_2, x_3 \in \mathbb{R} \right\},\,$$

where n is the normal vector of the plane. An intuitive understanding of subspace geometry is essential and built up from instructive computational examples.

Examples. Consider a linear mapping between real spaces $f: \mathbb{R}^n \to \mathbb{R}^m$, defined by $y = f(x) = Ax = [y_1 \dots y_n]^T$, with $A \in \mathbb{R}^{m \times n}$. Julia provides the nullspace function to return a set of vectors that span a null space. A function colspace to provide a set of vectors to span the column space is not yet in the general libraries, but can be readily defined, together with a function to display numerical results to a default precision of p = 6 digits.

```
in function colspace(A,p=6)
    return round.(Matrix(qr(A).Q)[:,1:rank(A)],digits=p)
    end;
```

```
∴ short(x) = round(x,digits=6);
∴ short(pi)
```

3.141593

.

With these functions defined, the following examples provide great insight into the significance of column and null spaces and their associated spanning sets. For these small-dimensional, simple examples geometric insight is sufficient to understand what column and null spaces represent. Computational procedures can be devised for much higher number of components, and the geometric insights gained here carry over.

1. For m = 3, n = 1,

$$\boldsymbol{A} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \boldsymbol{A}^T = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix},$$

the column space C(A) is the y_1 -axis, and the left null space $N(A^T)$ is the y_2y_3 -plane since the condition $A^Ty = 0$ reduces to $y_1 = 0$. Spanning vector sets for C(A) and $N(A^T)$ can be computed as follows, confirming the previous geometric descriptions. Note that combining the two leads to the identity matrix, an observation whose significance will soon become apparent.

∴ nullspace(A')

$$\begin{bmatrix}
0.0 & 0.0 \\
1.0 & 0.0 \\
0.0 & 1.0
\end{bmatrix}$$
(7)

∴ [colspace(A) nullspace(A')]

$$\begin{bmatrix}
1.0 & 0.0 & 0.0 \\
0.0 & 1.0 & 0.0 \\
0.0 & 0.0 & 1.0
\end{bmatrix}$$
(8)

:

2. For m = 3, n = 2,

$$A = \begin{bmatrix} 1 & -1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} a_1 & a_2 \end{bmatrix}, A^T = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix},$$

the columns of A are colinear, $a_2 = -a_1$, and the column space C(A) is the y_1 -axis, and the left null space $N(A^T)$ is the y_2y_3 -plane, as before.

MATRIX VECTOR SUBSPACES

∴ A=[1 -1; 0 0; 0 0]; CA=colspace(A)

$$\begin{bmatrix} 1.0 \\ 0.0 \\ 0.0 \end{bmatrix}$$
 (9)

∴ NAt=short.(nullspace(A'))

$$\begin{bmatrix}
0.0 & 0.0 \\
1.0 & 0.0 \\
0.0 & 1.0
\end{bmatrix}$$
(10)

∴ [CA NAt]

$$\begin{bmatrix}
1.0 & 0.0 & 0.0 \\
0.0 & 1.0 & 0.0 \\
0.0 & 0.0 & 1.0
\end{bmatrix}$$
(11)

:.

3. For m = 3, n = 2,

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \mathbf{A}^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix},$$

the column space C(A) is the y_1y_2 -plane, and the left null space $N(A^T)$ is the y_3 -axis.

∴ A=[1 0; 0 1; 0 0]; CA=colspace(A)

$$\begin{bmatrix}
1.0 & 0.0 \\
0.0 & 1.0 \\
0.0 & 0.0
\end{bmatrix}$$
(12)

∴ NAt=short.(nullspace(A'))

$$\begin{bmatrix}
0.0 \\
0.0 \\
1.0
\end{bmatrix}$$
(13)

∴ [CA NAt]

$$\begin{bmatrix}
1.0 & 0.0 & 0.0 \\
0.0 & 1.0 & 0.0 \\
0.0 & 0.0 & 1.0
\end{bmatrix}$$
(14)

:.

4. For m = 3, n = 2,

$$A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 0 & 0 \end{bmatrix}, A^{T} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \end{bmatrix},$$

the same C(A), $N(A^T)$ are obtained, albeit with a different set of spanning vectors returned by colspace.

∴ A=[1 1; 1 -1; 0 0]; CA=colspace(A)

$$\begin{bmatrix}
-0.707107 & -0.707107 \\
-0.707107 & 0.707107 \\
0.0 & 0.0
\end{bmatrix}$$
(15)

∴ NAt=short.(nullspace(A'))

$$\begin{bmatrix} 0.0 \\ 0.0 \\ 1.0 \end{bmatrix}$$
 (16)

∴ [CA NAt]

$$\begin{bmatrix}
-0.707107 & -0.707107 & 0.0 \\
-0.707107 & 0.707107 & 0.0 \\
0.0 & 0.0 & 1.0
\end{bmatrix}$$
(17)

:.

5. For m = 3, n = 3,

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 3 \\ 1 & -1 & -1 \\ 1 & 1 & 3 \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{bmatrix},$$

$$A^{T} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 3 & -1 & 3 \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_{1}^{T} \\ \boldsymbol{a}_{2}^{T} \\ \boldsymbol{a}_{3}^{T} \end{bmatrix}, A^{T}\boldsymbol{y} = \begin{bmatrix} \boldsymbol{a}_{1}^{T} \boldsymbol{y} \\ \boldsymbol{a}_{2}^{T} \boldsymbol{y} \\ \boldsymbol{a}_{3}^{T} \boldsymbol{y} \end{bmatrix}.$$

Since $\mathbf{a}_3 = \mathbf{a}_1 + 2\mathbf{a}_2$, the orthogonality condition $\mathbf{A}^T \mathbf{y} = \mathbf{0}$ is satisfied by vectors of form $\mathbf{y} = [a \ 0 \ -a], \forall a \in \mathbb{R}$.

$$\therefore$$
 A=[1 1 3; 1 -1 -1; 1 1 3]; CA=colspace(A)

∴ NAt=short.(nullspace(A'))

$$\begin{bmatrix}
-0.707107 \\
0.0 \\
0.707107
\end{bmatrix} (19)$$

∴ [CA NAt]

$$\begin{bmatrix} -0.5773502691896257 & 0.40824829046386313 & -0.707107 \\ -0.5773502691896257 & -0.816496580927726 & 0.0 \\ -0.5773502691896257 & 0.40824829046386313 & 0.707107 \end{bmatrix}$$
(20)

:.