8.1 Subspace

In previous lecture notes, we introduced the concept of a vector space and the notion of basis and dimension. In this note, we introduce the idea of subspaces, as it is often useful to look at part of the entire set of vectors in a vector space.

Definition 8.1 (Subspace): A subspace $\mathbb{U}$ consists of a subset of the set $\mathbb{V}$ in the vector space $(\mathbb{V},\mathbb{F})$ that satisfies the following three properties:

- Contains the zero vector: $\vec{0} \in \mathbb{U}$.
- Closed under vector addition: For any two vectors $\vec{v}_1, \vec{v}_2 \in \mathbb{U}$, their sum $\vec{v}_1 + \vec{v}_2$ must also be in $\mathbb{U}$.
- Closed under scalar multiplication: For any vector $\vec{v} \in \mathbb{U}$ and scalar $\alpha \in \mathbb{F}$, the product $\alpha\vec{v}$ must also be in $\mathbb{U}$.

Equivalently, a subspace is a subset of the vectors in a vector space where any linear combination of the vectors in the set lies within the set.

Recall that the range of a matrix $A \in \mathbb{R}^{n\times m}$ is the space of all outputs that the operator A can map to. We know that $range(A)$ is a subset of $\mathbb{R}^n$. However, is $range(A)$ a subspace? Let's see if it satisfies each condition described above.

- We know that the zero vector is in $range(A)$ because A operating on the zero vector gives the zero vector: $A\vec{0} = \vec{0}$.
- If $\vec{v}_1, \vec{v}_2$ are in $range(A)$, then there exist $\vec{u}_1, \vec{u}_2 \in \mathbb{R}^m$ such that $A\vec{u}_1 = \vec{v}_1$ and $A\vec{u}_2 = \vec{v}_2$. Adding the two equations together, we have (due to the distributivity of matrix-vector multiplication):

  $$\vec{v}_1 + \vec{v}_2 = A\vec{u}_1 + A\vec{u}_2 = A\left(\vec{u}_1 + \vec{u}_2\right). \tag{1}$$

  This tells us the $\vec{v}_1 + \vec{v}_2$ is in $range(A)$ as well.
- If $\vec{v}$ is in $range(A)$, then there exists $\vec{u} \in \mathbb{R}^m$ such that $A\vec{u} = \vec{v}$. For any scalar α, we can write

  $$\alpha\vec{v} = \alpha A\vec{u} = A\left(\alpha\vec{u}\right). \tag{2}$$

  This says that $\alpha\vec{v}$ is also in $range(A)$.

As a result, we can see that $range(A)$ is a subspace.

Now let's introduce two new terms: **column space** and **row space**. **Column space** is is defined as the span of the columns of a matrix, which is just the range of the matrix. Similarly, **row space** is defined as the span of the rows of a matrix. Directly following from the way column and row spaces are defined, we can see that both of these spaces are subspaces.

Just as basis and dimension are defined for vector spaces, they have equivalent definitions for subspaces. A basis of a subspace is a set of linearly independent vectors that span the subspace, and the dimension of a subspace is the number of vectors in its bases.

Let's look at an example.

Example 8.1 (Subspace): Consider the linear operator $A = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix}$. Then the range of A would be a subspace consisting of vectors in $\mathbb{R}^3$

$$range(A) = \left\{ \alpha \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + \gamma \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\} \tag{3}$$

where α, β, γ are any real scalars. We know that the input space $\mathbb{R}^3$ has dimension of 3. How about the dimension of the output space? We can see that a basis for the output space $range(A)$ is just the first two columns of A since the first two columns are linearly independent of each other and the third column does not contribute to the range since it is just a zero vector. So the dimension of the range of A is equal to 2. Notice that the dimension of the output space is smaller than the dimension of input space. We will investigate this further in the next section.

8.2 Loss of Dimensionality and Nullspace

It might seem a little strange that the dimension of the output space can be smaller than the output of the input space. After all, for any matrix $A \in \mathbb{R}^{n \times m}$, we have m parameters to fiddle with when we construct our input vector, so how can it be that what we are seeing on the output can be described by fewer parameters? Where is the "remaining dimensionality" going? The answer can be found in something called the **nullspace**. The nullspace of A consists of all vectors $\vec{x}$ in $\mathbb{R}^m$ such that $A\vec{x} = \vec{0}$:

$$N(A) = \{\vec{x} \mid A\vec{x} = \vec{0},\ \vec{x} \in \mathbb{R}^m\}. \tag{4}$$

The nullspace of A is the set of vectors that get mapped to zero by A. What is the dimension of the nullspace? We know that it can be at most m, since all of the input vectors have m components. However, unless A is the zero matrix, not every input gets mapped to zero, so in general the dimension should be less than m. The question we need to ask is how many independent ways can we create the zero vector from taking linear combinations of the columns of A. Recall that

$$A\vec{x} = \sum_{k=1}^{m} x_i \vec{a}_k, \tag{5}$$

where again x_i are the free parameters. So our task is to find vectors $\vec{x}$ such that

$$\sum_{i=1}^{m} x_i \vec{a}_i = \vec{0} \tag{6}$$

First note that the only way $\sum_{i=1}^{m} x_i \vec{a}_i = \vec{0}$ (non-trivially) is if the columns of A are not all linearly independent. This holds by definition of linear independence. We can represent A in terms of its linearly independent columns and dependent columns, $\vec{a}^i$ and $\vec{a}^d$ respectively. Assuming there are $j < m$ linearly independent columns[1],

$$A = \begin{bmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \dots & \vec{a}_m \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & & | & | & & | \\ \vec{a}^i_1 & \dots & \vec{a}^i_j & \vec{a}^d_1 & \dots & \vec{a}^d_{m-j} \\ | & & | & | & & | \end{bmatrix}. \tag{7}$$

We can then break up the summation in equation (6) into two summations one for the linearly independent columns and another for the linearly dependent columns,

$$\sum_{k=1}^{j} x^i_k \vec{a}^i_k + \sum_{k=1}^{m-j} x^d_k \vec{a}^d_k = \vec{0}, \tag{8}$$

where x^i and x^d are the parameters of $\vec{x}$ that multiply the linearly independent and dependent columns of A respectively. Rearranging a bit we get

$$\sum_{k=1}^{j} x^i_k \vec{a}^i_k = -\sum_{k=1}^{m-j} x^d_k \vec{a}^d_k \tag{9}$$

Remember we get to choose the parameters in our vector $\vec{x}$ that will satisfy (6). We know that in the total summation at least one linearly dependent vector must be multiplied by a nonzero parameter, since the linearly independent vectors alone cannot be linearly combined (non-trivially) to get $\vec{0}$. With this constraint let us then simplify our problem. We will impose that x^d_1 be nonzero, and set the other parameters multiplying linearly dependent vectors equal to zero. That is $x^d_2 = \ldots = x^d_{m-j} = 0$. Since $\vec{a}^d_1$ is linearly dependent on the $\vec{a}^i_k$'s we know that there exist a unique set of numbers $\beta^1_1, \beta^1_2, \ldots, \beta^1_j$ such that

$$\sum_{k=1}^{j} \beta^1_k \vec{a}_k = \vec{a}^d_1. \tag{10}$$

In other words there is a unique linear combination of our linearly independent vectors that equals $\vec{a}^d_1$. **As an aside, if a vector can be represented as a linear combination of linearly independent vectors then this representation is unique. You can try to prove this. Hint: Assume that two representations exists, set the two representations equal to one another, and see if the linear independence still holds.** Rearranging we get

[1] Please note that the linearly independent columns do not all have to be next to another, we just write it this way to ease the presentation. The results we show will still hold even if the linearly independent columns are not side by side.

$$\sum_{k=1}^{j} -\beta_k^1 \vec{a}_k^i + \vec{a}_1^d = \vec{0}, \tag{11}$$

which is also equal to

$$\sum_{k=1}^{j} -\beta_k^1 \vec{a}_k^i + \vec{a}_1^d + \sum_{k=2}^{m-j} 0\vec{a}_k^d = \vec{0}. \tag{12}$$

Notice that the last summation on the left hand side is equal to zero, and we only include it to more clearly show one of the vectors in the nullspace, namely

$$\vec{x} = \begin{bmatrix} -\beta_1^1 \\ -\beta_2^1 \\ \vdots \\ -\beta_j^1 \\ 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{13}$$

Can we find others? Well the first thing we can do is multiply equation (10) by our free parameter x_1^d,

$$x_1^d (\sum_{k=1}^{j} \beta_k^1 \vec{a}_k) = x_1^d \vec{a}_1^d. \tag{14}$$

Similarly we can conclude

$$\sum_{k=1}^{j} -x_1^d \beta_k^1 \vec{a}_k^i + x_1^d \vec{a}_1^d + \sum_{k=2}^{m-j} 0\vec{a}_k^d = \vec{0}. \tag{15}$$

Since x_1^d is a free parameter any vector of the form

$$\vec{x} = \begin{bmatrix} -\beta_1^1 x_1^d \\ -\beta_2^1 x_1^d \\ \vdots \\ -\beta_j^1 x_1^d \\ x_1^d \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} -\beta_1^1 \\ -\beta_2^1 \\ \vdots \\ -\beta_j^1 \\ 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} x_1^d \tag{16}$$

will also be in the nullspace! Are there others? Yes. There is nothing special about choosing the parameter of the first linearly dependent vector to be the nonzero parameter. We can repeat the same procedure for each of the linearly dependent columns, to obtain new vectors in the nullspace. For example say that we set $x_1^d = x_3^d = \ldots = x_{m-j}^d = 0$, and leave x_2^d as our nonzero parameter we will find that

$$\vec{x} = \begin{bmatrix} -\beta_1^2 \\ -\beta_2^2 \\ \vdots \\ -\beta_j^2 \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} x_2^d, \tag{17}$$

is also in the nullspace, where $\sum_{k=1}^{j} \beta_k^2 \vec{a}_k = \vec{a}_2^d$. This procedure can be done for each linearly dependent vector, for example if x_3^d is the nonzero parameter we will get

$$\vec{x} = \begin{bmatrix} -\beta_1^3 \\ -\beta_2^3 \\ \vdots \\ -\beta_j^3 \\ 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} x_3^d, \tag{18}$$

Furthermore, we can add vectors together from our nullspace together to get other vectors in the nullspace. **Aside: Try to prove this. Hint: if $\vec{x}_1$ and $\vec{x}_2$ are in the nullspace of A, what can be said about $A(\vec{x}_1 + \vec{x}_2)$?** This means that

$$\vec{x} = \begin{bmatrix} -\beta_1^1 \\ -\beta_2^1 \\ \vdots \\ -\beta_j^1 \\ 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} x_1^d + \begin{bmatrix} -\beta_1^2 \\ -\beta_2^2 \\ \vdots \\ -\beta_j^2 \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} x_2^d + \begin{bmatrix} -\beta_1^3 \\ -\beta_2^3 \\ \vdots \\ -\beta_j^3 \\ 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} x_3^d + \ldots + \begin{bmatrix} -\beta_1^{m-j} \\ -\beta_2^{m-j} \\ \vdots \\ -\beta_j^{m-j} \\ 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} x_{m-j}^d \tag{19}$$

is also in the nullspace. Notice that once we choose the x^d parameters then the x^i parameters are fixed. This is because the x^d parameters control how much of the linearly dependent columns of A are being put into the summation, and the x^i parameters must ensure that the exact amount of linearly independent columns are included to cancel out the linearly dependent columns so that the output be zero. So we finally conclude that the dimension of the nullspace is equal to the number of our x^d parameters, which is equal to the number of linearly dependent columns of A. We will work out some examples in the next section.

The last point we would like to highlight here is that the dimension of the range of A is equal to the number of linearly independent columns, and the dimension of the nullspace of A is equal to the number of linearly dependent columns. Thus

$$m - \dim(range(A)) = \dim(N(A)), \tag{20}$$

so the loss of dimensionality from the input space to the output space shows up in the nullspace! This result is called the **rank-nullity theorem**.

8.3 Computing the Nullspace

So now we will show you how to compute the nullspace of a matrix systematically. Hopefully our analysis in the previous section will prove to be fruitful. For a vector to be in the nullspace it must weight the linearly independent columns appropriately to cancel out the weighted linearly dependent columns, and what we will show next will find all such vectors that do so.

In solving for the nullspace, we are fundamentally trying to solve the system of equations $A\vec{x} = \vec{0}$. We know that row reducing does not effect the solution of the system of equations, so we will assume that every matrix we have is already in reduced row echelon form (RREF). If the matrix is not in RREF then it can first be row reduced and the techniques will apply. RREF will make our lives much easier for two reasons: first, it is really easy to find the linearly independent columns (the columns with pivots), and it is easy to figure out how the linearly independent columns should be combined to cancel out the linearly dependent columns. We will work with the matrix

$$A = \begin{bmatrix} 1 & 2 & 0 & 3 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{21}$$

First we identify the linearly dependent columns, which in this case could be columns 2 and 4. To be clear, the set of linearly dependent columns we chose is

$$\{ \begin{bmatrix} 2 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 0 \\ 0 \\ 0 \end{bmatrix} \} \tag{22}$$

and the set of linearly independent columns includes columns 1, 3, and 5,

$$\left\{ \begin{bmatrix}1\\0\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\\0\\0\end{bmatrix} \right\}. \tag{23}$$

Note that the choice of linearly dependent columns need not be unique. All that is needed is that any vector in the linearly dependent columns we choose can be written as a linear combination of the vectors in the set of linearly independent columns and, of course, the columns in the linearly independent set should be linearly independent.

We would want to find all possible scalars x_1, x_2, x_3, x_4, x_5 such that

$$x_2 \times \begin{bmatrix}2\\0\\0\\0\\0\end{bmatrix} + x_4 \times \begin{bmatrix}3\\2\\0\\0\\0\end{bmatrix} + x_1 \times \begin{bmatrix}1\\0\\0\\0\\0\end{bmatrix} + x_3 \times \begin{bmatrix}0\\1\\0\\0\\0\end{bmatrix} + x_5 \times \begin{bmatrix}0\\0\\1\\0\\0\end{bmatrix} = \vec{0}. \tag{24}$$

Rather than considering all the linearly dependent vectors at once, we can consider them individually and sum up the contribution from each linearly dependent columns at the end. Let us first impose that the first linearly dependent column must have a weight of one in the summation and the other linearly dependent columns have weights of zero. After this, we will repeat this and only allow the second linearly dependent column to show up in the summation. Let's start with the first linearly dependent column, we want to find the unique weighting of the linearly independent columns so that the resulting sum cancels out the first linearly dependent column. It is easy to see that

$$1 \times \begin{bmatrix}2\\0\\0\\0\\0\end{bmatrix} + 0 \times \begin{bmatrix}3\\2\\0\\0\\0\end{bmatrix} = 2 \times \begin{bmatrix}1\\0\\0\\0\\0\end{bmatrix} + 0 \times \begin{bmatrix}0\\1\\0\\0\\0\end{bmatrix} + 0 \times \begin{bmatrix}0\\0\\1\\0\\0\end{bmatrix}. \tag{25}$$

Now, as before, the equality would hold if we multiply both sides of the equation by any scalar $\alpha \in \mathbb{R}$

$$\alpha \times \begin{bmatrix}2\\0\\0\\0\\0\end{bmatrix} + 0 \times \begin{bmatrix}3\\2\\0\\0\\0\end{bmatrix} = 2\alpha \times \begin{bmatrix}1\\0\\0\\0\\0\end{bmatrix} + 0 \times \begin{bmatrix}0\\1\\0\\0\\0\end{bmatrix} + 0 \times \begin{bmatrix}0\\0\\1\\0\\0\end{bmatrix}. \tag{26}$$

Note that essentially we can treat α as a free variable that can vary its value however we want while satisfying the above equation. Moving everything to the left hand side, we have

$$\alpha \times \begin{bmatrix}2\\0\\0\\0\\0\end{bmatrix} + 0 \times \begin{bmatrix}3\\2\\0\\0\\0\end{bmatrix} + (-2\alpha) \times \begin{bmatrix}1\\0\\0\\0\\0\end{bmatrix} + 0 \times \begin{bmatrix}0\\1\\0\\0\\0\end{bmatrix} + 0 \times \begin{bmatrix}0\\0\\1\\0\\0\end{bmatrix} = 0. \tag{27}$$

Now we see that any vector of the form $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -2\alpha \\ \alpha \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \alpha$ is in the nullspace of A. Now let's set the weight of the second linearly dependent column to one and set the weight of the first linearly dependent to zero. Similarly, we can find the unique weightings of the linearly independent columns that sum to the second linearly dependent column.

$$0 \times \begin{bmatrix} 2 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + 1 \times \begin{bmatrix} 3 \\ 2 \\ 0 \\ 0 \\ 0 \end{bmatrix} = 3 \times \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + 2 \times \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + 0 \times \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}. \tag{28}$$

The equality would still hold if we multiply both sides of the equation by any scalar $\beta \in \mathbb{R}$,

$$0 \times \begin{bmatrix} 2 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \beta \times \begin{bmatrix} 3 \\ 2 \\ 0 \\ 0 \\ 0 \end{bmatrix} = 3\beta \times \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + 2\beta \times \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + 0 \times \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}. \tag{29}$$

Again, moving everything to the left hand side of the equation, we have

$$0 \times \begin{bmatrix} 2 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \beta \times \begin{bmatrix} 3 \\ 2 \\ 0 \\ 0 \\ 0 \end{bmatrix} + (-3\beta) \times \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + (-2\beta) \times \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + 0 \times \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \vec{0}. \tag{30}$$

We have that any vector of the form $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -3\beta \\ 0 \\ -2\beta \\ \beta \\ 0 \end{bmatrix} = \begin{bmatrix} -3 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix} \beta$ is in the nullspace of A. We know that if any two vectors are in the nullspace, then their sum is also in the nullspace. Thus we can conclude that the nullspace of A is

$$N(A) = \{ \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \alpha + \begin{bmatrix} -3 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix} \beta \mid \alpha, \beta \in \mathbb{R} \}. \tag{31}$$

Notice that the dimension of the nullspace is 2 and the dimension of the range is 3.