Lectures - Week 5
Four Basic Spaces

1. The **column space** (or equivalently the **range**) of A, where A is $m \times n$ matrix is all linear combinations of the columns of A. We denote this by $\mathcal{R}(A)$.

- By definition (because it contains all linear combinations and is thus closed under addition and scalar multiplication) the column space is a subspace of R^m.
- An equivalent statement to $A\vec{x} = \vec{b}$ being solvable is that $\vec{b}$ is in the range or column space of A.

2. The **null space** of A, denoted $\mathcal{N}(A)$, where A is $m \times n$ matrix is the set of all vectors $\vec{z} \in \mathrm{R}^n$ such that $A\vec{z} = \vec{0}$.

- The null space is a subspace of R^n because it consists of vectors in R^n and is closed under addition and scalar multiplication:

$$A\vec{y} = 0, A\vec{z} = 0 \Rightarrow A(\vec{y} + \vec{z}) = A\vec{y} + A\vec{z} = 0$$

$$A\vec{z} = 0, k \in \mathrm{R}^1 \Rightarrow A(k\vec{z}) = k(A\vec{z}) = k(0) = 0$$

3. The **row space** of A is the span of the rows of A and is thus a subspace of R^n. It can be found by row reducing A because the resulting upper triangular matrix U has the same row space because we are always taking linear combinations of the rows of A to get U.

4. The **null space of** A^T, $\mathcal{N}(A^T)$, is a subspace of R^n and consists of all $\vec{z} \in \mathrm{R}^n$ such that $A^T\vec{z} = \vec{0}$. This space is often call the *left null space of* A.

Note that two of these spaces are subspaces of R^n and two of R^m. We want to be able to find a basis and its dimension for each of these spaces and see any relationships they have with each other.

Example Find the range of each matrix

$$A_1 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad A_2 = \begin{pmatrix} 3 & 2 \\ 1 & 0 \\ 1 & 6 \end{pmatrix}$$

The $\mathcal{R}(A_1) = \mathrm{R}^2$ because the columns are two linearly independent vectors in R^2 and thus form a basis for R^2. The $\mathcal{R}(A_2)$ is the span$\{(3,1,1)^T, (2,0,6)^T\}$ which is a subspace of R^3.

Example Find the null space of each matrix

$$A_1 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad A_2 = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \quad A_3 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

For A_1 we have that $\mathcal{N}(A_1) = \vec{0}$ because the matrix is invertible. To see this we could take the determinant or perform GE and get the result.

For A_2 we see that $\mathcal{N}(A_2)$ is $\alpha(-2x_2, x_2)^T$, i.e., the all points on the line through the origin $y = -.5x$. To see this consider GE for the system

$$\begin{pmatrix} 1 & 2 & \\ 2 & 4 & \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & \\ 0 & 0 & \end{pmatrix} \Rightarrow 0 \cdot x_2 = 0, \quad x_1 + 2x_2 = 0$$

This says that x_2 is arbitrary and $x_1 = -2x_2$.

For A_3 we see that $\mathcal{N}(A_2)$ is all of $\mathbf{R}^2$.

Example What are the possible null spaces of a 3×3 matrix?

For an invertible matrix it is the (i) zero vector in $\mathbf{R}^3$, we could have (ii) a line through the origin, (iii) a plane through the origin or all of (iv) $\mathbf{R}^3$.

Theorem If A is an $n \times n$ invertible matrix then

(i) $\mathcal{N}(A) = \vec{0}$

(ii) $\mathcal{R}(A) = \mathbf{R}^n$

The first part of the theorem says that the only solution to $A\vec{x} = \vec{0}$ is $\vec{x} = \vec{0}$ for an invertible matrix. The second part says that for A invertible, $A\vec{x} = \vec{b}$ is solvable for any right hand side because $\vec{b} \in \mathbf{R}^n$ and the range of A is all of $\mathbf{R}^n$. So the columns of A form a basis for $\mathbf{R}^n$.
We now want to look at the general case where A is not invertible and A is rectangular.

Example Consider the system $A\vec{x} = \vec{b}$ where

$$A = \begin{pmatrix} 2 & 2 & 1 \\ 0 & 6 & 3 \\ 5 & 8 & 4 \end{pmatrix}$$

Are the rows of A linearly independent? Give the row space of A and its dimension. What is the null space of A and its dimension? Are the columns of A linearly independent? What is the range of A of its dimension? Is this system solvable for any right hand side vector $\vec{b}$? What is the range of A^T and its dimension? What is the $\mathcal{N}(A^T)$?

To see if the rows are linearly independent and to get the row space of A we simply row reduce A to get

$$\begin{pmatrix} 2 & 2 & 1 \\ 0 & 6 & 3 \\ 5 & 8 & 4 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 2 & 1 \\ 0 & 6 & 3 \\ 0 & 3 & 1.5 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 2 & 1 \\ 0 & 6 & 3 \\ 0 & 0 & 0 \end{pmatrix}$$

so the row space is a subspace of $\mathbf{R}^3$ given by $\text{span}\{(2,2,1)^T, (0,6,3)^T\}$ so it has dimension 2. Clearly the rows of A are linearly dependent.

This also tells us about the null space of A; clearly there is a vector in the null space of A and is $\alpha(0, -3, 6)^T$.

The three columns are in R^3; we want to know if one of the columns can be formed by a linear combination of the other two; if so they are linearly dependent. From inspection we see that column two is twice column three so they are linearly dependent because

$$0\begin{pmatrix}2\\0\\5\end{pmatrix}+1\begin{pmatrix}2\\6\\8\end{pmatrix}-2\begin{pmatrix}1\\3\\4\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}$$

If we didn't see this relationship what could we do? If we perform GE as above this doesn't help us because the column space of the original matrix and the resulting upper triangular matrix is NOT the same. Here we know that the column space of A is the span $\{(2,0,5)^T,(2,6,8)^T\}$ whereas the column space of U is the span $\{(2,0,0)^T,(2,6,0)^T\}$ which is clearly not the same space. What we could do is form A^T and row reduce it because each time we are taking a linear combinations of the rows of A^T (i.e., columns of A) so we are not changing the space. We have

$$\begin{pmatrix}2&0&5\\2&6&8\\1&3&4\end{pmatrix}\rightarrow\begin{pmatrix}2&0&5\\0&6&3\\0&3&1.5\end{pmatrix}\rightarrow\begin{pmatrix}2&0&5\\0&6&3\\0&0&0\end{pmatrix}$$

So we get the same result as by inspection, $\mathcal{R}(A)$ =span$\{(2,0,5)^T,(0,6,3)^T\}$ because $(2,6,8)^T=(2,0,5)^T+(0,6,3)^T$. Clearly the dimension is 2. This also says that we can find a $\vec{b}$ which is not in the span of the columns of A; e.g., $\vec{b}=(4,10,1)^T\neq c_1(2,2,1)^T+c_2(0,6,3)^T$.

Now the $\mathcal{R}(A^T)$ is a subspace of R^3 and we should be able to find it in an analogous way to how we found $\mathcal{R}(A)$. We simply take the transpose and row reduce it; but we have already row reduced A so we know that $\mathcal{R}(A^T)$ =span$\{(2,2,1)^T,(0,6,3)^T\}$ so it has dimension 2.

Lastly the $\mathcal{N}(A^T)$ is found by row reducing A^T which we have already done. We get that it has dimension 1 and is $\alpha(0,1,-2)$.

Two things to note are for this matrix (i) there are two rows and two columns that are linearly dependent and (ii) $\mathcal{R}(A^T)$ is the row space of A.

Example Consider the under-determined system $A\vec{x}=\vec{b}$ where

$$A=\begin{pmatrix}2&3&-1\\0&4&6\end{pmatrix}$$

Are the rows of A linearly independent? Give the row space of A and its dimension. What is the null space of A and its dimension? Are the columns of A linearly independent? What is the range of A of its dimension? Is this system solvable for any right hand side vector $\vec{b}$? What is the range of A^T and its dimension? What is the $\mathcal{N}(A^T)$?

The rows are vectors in R^3 and because there are only two of them, we can see by inspection that they are linearly independent. So the row space has dimension 2 and is given by span$\{(2,3,-1)^T,(0,4,6)^T\}$.

For the columns we have three vectors in $\mathbf{R}^2$ so we know that they are linearly dependent. Any two of the columns are linearly independent so the dimension is 2 and is all of $\mathbf{R}^2$. The system is solvable for any right hand side in $\mathbf{R}^2$ because although the columns are linearly dependent, they span all of R^2 and so the $\mathcal{R}(A) = \mathrm{R}^2$.

Does A map anything to $\vec{0}$ other than the zero vector? Clearly yes because the system is already reduced by GE and

$$4x_2 + 6x_3 = 0 \Rightarrow x_3 \text{ is arbitrary.}$$

So if we choose, e.g., $\alpha(11/2, -3, 2)^T$ it is in the null space of A.

The $\mathcal{R}(A^T)$ is found by row reducing A so clearly it is just $\text{span}\{(2,3,-1)^T, (0,4,6)^T\}$.

The $\mathcal{N}(A^T)$ is found by row reducing A^T

$$\begin{pmatrix} 2 & 0 \\ 3 & 4 \\ -1 & 6 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 0 \\ 0 & 4 \\ 0 & 6 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 0 \\ 0 & 4 \\ 0 & 0 \end{pmatrix}$$

and thus $\mathcal{N}(A^T)$ is $\vec{0}$.

Here note that for the 2×3 matrix A there are two linearly independent rows and columns and $\mathcal{R}(A^T)$ is the row space of A.

In these two examples the number of linearly independent rows and columns is the same. This is *always* the case and we give it a name.

Definition The *rank* of a matrix is the number of linearly independent rows or columns (it is the same).

If we know that there are e.g., three linearly independent rows then there are three linearly independent columns and the rank is three.

Example For the two matrices in the previous examples, find their rank.

Clearly both have rank two.

The next result gives the dimensions of these spaces and is often called the Fundamental Theorem of Linear Algebra, Part I.

Theorem Fundamental Theorem of Linear Algebra, Part I. Let A be an $m \times n$ matrix. Then the following conditions hold.

- The $\mathcal{R}(A)$ is the column space of A and is a subspace of R^m. The dimension $\dim(\mathcal{R}(A))$ is the rank r and $r \leq m$.
- The null space of A, $\mathcal{N}(A)$ is a subspace of R^n and has dimension $n - r$ where r is the rank of A
- The row space of A is a subspace of $\mathbf{R}^n$ and is the column space of A^T, $\mathcal{R}(A^T)$ and has dimension r.
- The $\mathcal{N}(A^T)$ is the left null space of A and is a subspace of $\mathbf{R}^m$ whose dimension is $m - r$.

Lets return to our two examples and look at them in light of this theorem.

Example For

$$A = \begin{pmatrix} 2 & 2 & 1 \\ 0 & 6 & 3 \\ 5 & 8 & 4 \end{pmatrix}$$

we have $m = n = 3$. We row-reduced the matrix to get 2 linearly independent rows so the rank is 2. We immediately know that there are also 2 linearly independent columns and we know that the $\mathcal{R}(A^T)$ is the same as the row space. We know that the dimension of the null space of A has to be 3-2=1 and the dimension of $\mathcal{N}(A^T)$ is 3-2=1. For the second matrix

$$A = \begin{pmatrix} 2 & 3 & -1 \\ 0 & 4 & 6 \end{pmatrix}$$

$m = 2$ and $n = 3$. We found that the rank was 2 so the null space of A has dimension $n - r = 3 - 2 = 1$. Also the left null space $\mathcal{N}(A^T)$ has dimension $m - r = 2 - 2 = 0$ so it has dimension zero.

Example Analyze the four fundamental spaces for the matrix

$$A = \begin{pmatrix} 2 & 1 & 4 & 4 \\ -4 & 2 & -2 & 0 \\ 0 & 4 & 6 & 2 \end{pmatrix}$$

In this case $m = 3$ and $n = 4$ so the maximum the rank can be is 3. To see the row space we row reduce A to get

$$\begin{pmatrix} 2 & 1 & 4 & 4 \\ -4 & 2 & -2 & 0 \\ 0 & 4 & 6 & 2 \end{pmatrix} \to \begin{pmatrix} 2 & 1 & 4 & 4 \\ 0 & 4 & 6 & 8 \\ 0 & 4 & 6 & 2 \end{pmatrix} \to \begin{pmatrix} 2 & 1 & 4 & 4 \\ 0 & 4 & 6 & 8 \\ 0 & 0 & 0 & -6 \end{pmatrix}$$

so the rank is 3 and a basis for the row space is $\{(2, 1, 4, 4), (0, 4, 6, 8), (0, 0, 0, -6)\}$ or we could have taken the three rows of A. The column space or range of A is a subspace of $\mathbf{R}^3$ but its dimension is three (because the rank is three) so the column space is all of $\mathbf{R}^3$ and we can take the standard basis. The dimension of the null space $\mathcal{N}(A)$ is 4-3=1 and from our row reduction we have $x_4 = 0$, $x_3 = \alpha$, $x_2 = -1.5x_3$ and $x_1 = (-x_2 - 4x_3)/2 = -5/4\alpha$ so a basis for $\mathcal{N}(A)$ is $(-5, -6, 4, 0)^T$. Finally the left null space of A, $\mathcal{N}(A^T)$ has dimension 3-3=0.

What do we know about these spaces for an invertible matrix? The following are equivalent statements.

- A is an invertible $n \times n$ matrix.
- A is an $n \times n$ matrix with rank n.
- A is an $n \times n$ matrix and the range of A, $\mathcal{R}(A)$ is all of $\mathbf{R}^n$ and has dimension n
- A is an $n \times n$ matrix and the null space of A, $\mathcal{N}(A)$ is the zero vector and has dimension 0.

- A is an $n \times n$ matrix and the left null space of A, $\mathcal{N}(A^T)$ is the zero vector and has dimension 0.
- A is an $n \times n$ matrix and the range of A^T, $\mathcal{R}(A^T)$ is all of $\mathbf{R}^n$ and has dimension n.
- A is an $n \times n$ matrix and the row space of A is all of $\mathbf{R}^n$ and has dimension n.

Orthogonal Spaces

In the last section, we concentrated on determining the dimensions of the four fundamental spaces. In this section we want to look at their orientation with respect to each other.

Recall that two vectors are orthogonal if their dot product is zero. Suppose we had two spaces V, W and they had the property that for any $\vec{v} \in V$ and any $\vec{w} \in W$, $\vec{v}^T \vec{w} = 0$. Then this would be an analogous definition of orthogonality for spaces.

Definition Two spaces V, W are *orthogonal* provided $\vec{v}^T \vec{w} = 0$ for any $\vec{v} \in V$ and $\vec{w} \in W$.

If two spaces are orthogonal then the only vector they have in common is the zero vector. If every $\vec{v} \in V$ is orthogonal to each basis vector of W then it is orthogonal to all of W because every other vector in W can be written as a linear combination of the basis vectors. Specifically if $\vec{w}_i \in W$, $i = 1, \ldots, n$ form a basis for W and $\vec{v}^T \vec{w}_i = 0$ for all i then

$$\vec{p} = \sum_{i=1}^{n} c_i \vec{w}_i \Rightarrow \vec{v}^T \vec{p} = \vec{v}^T \Big(\sum_{i=1}^{n} c_i \vec{w}_i \Big) = \sum_{i=1}^{n} c_i (\vec{v}^T \vec{w}_i) = \sum_{i=1}^{n} c_i (0) = 0$$

Example Let V be the plane spanned by the vectors $\vec{v}_1 = (1,0,0,0)$, $\vec{v}_2 = (1,1,0,0)$ and W the line spanned by $\vec{w} = (0,0,4,5)$. Then $\vec{w}^T \vec{v}_1 = 0$ and $\vec{w}^T \vec{v}_2 = 0$ so for any other element of V, say $\alpha \vec{v}_1 + \beta \vec{v}_2$, we have $\vec{w}^T (\alpha \vec{v}_1 + \beta \vec{v}_2) = \alpha \vec{w}^T \vec{v}_1 + \beta \vec{w}^T \vec{v}_2 = 0$ and the two spaces are orthogonal.

We might ask ourselves if any of our four fundamental spaces are orthogonal. The following theorem answers this question.

Theorem Let A be an $m \times n$ matrix.

(i) The null space of A, $\mathcal{N}(A)$ and the row space of A are orthogonal spaces.
(ii) The left null space of A and the column space of A, $\mathcal{R}(A)$ are orthogonal spaces.

Proof To see why (i) holds we first take a vector in $\mathcal{N}(A)$, say $\vec{w} \in \mathcal{N}(A)$; thus $A\vec{w} = \vec{0}$. This means that we take the first row of A and dot it into $\vec{w}$ to get 0; then the second row of A and dot it into $\vec{w}$ and get zero and so forth. Clearly this implies that $\vec{w}$ if orthogonal to each row of A. To see why (ii) holds we have to take $\vec{w} \in \mathcal{N}(A^T)$ and $\vec{b} \in \mathcal{R}(A)$ and show that $\vec{w}^T \vec{b} = 0$ where $\vec{w}, \vec{b}$ were arbitrary. Now $A^T w = \vec{0}$ and as in the previous argument this implies $\vec{w}$ is orthogonal to each row of A^T. But the rows of A^T are the columns of A so $\vec{w}$ is orthogonal to each column of A, i.e., orthogonal to $\mathcal{R}(A)$.

We could prove these results in a more rigorous way. For example, another way to prove (i) is a bit more abstract, but insightful. As before let $\vec{w} \in \mathcal{N}(A)$; another way to say that

$\vec{v}$ is in the row space of A is to say $\vec{v} \in \mathcal{R}(A^T)$. This says there is a vector $\vec{x}$ such that $A^T\vec{x} = \vec{v}$. Using this we have

$$\vec{w}^T\vec{v} = \vec{w}^T A^T \vec{x} = (A\vec{w})^T\vec{x} = 0$$

because $A\vec{w} = \vec{0}$.

This theorem says that every vector in the null space of A is perpendicular to every vector in the row space of A; however something stronger is actually true. The fact is that every vector in R^n which is perpendicular to the row space of A is in the null space of A; that is, the null space contains *every* vector in R^n which is orthogonal to row space of A. The analogous condition holds for (ii). To this end, we make the following definition.

Definition Let V be a given subspace of R^n. Then the set of all vectors in R^n which are orthogonal to V is called the *orthogonal complement* of V and is denoted $V^\perp$.

We want to say that the null space is the orthogonal complement of the row space and the row space is the orthogonal complement of the null space. Our theorem didn't prove this; it only proved they were orthogonal spaces. Why can't there be another vector orthogonal to the null space of A other than a linear combination of the rows of A? We will demonstrate this by "proof by contradiction"; i.e., we will assume such a vector exists and then get a contradiction. Let $\vec{z}$ be another vector orthogonal to any $\vec{w} \in \mathcal{N}(A)$; i.e., $\vec{z}^T\vec{w} = 0$ and assume that $\vec{z}$ is NOT a linear combination of the rows of A. Now lets form another matrix B which is $(m+1)\times n$ and is the same as A except in the $(m+1)$ row we add our vector $\vec{z}$. Now because $\vec{z}$ is not in the row space of A, our new matrix B has one more linearly independent row than A and so its rank must be increased by one compared to the rank of A. On the other hand the null space of B is the same as the null space of A. However, we know that n is the sum of the rank and the dimension of the null space so we have our contradiction. The second part of the Fundamental Theorem of Linear Algebra summarizes these results.

Theorem Fundamental Theorem of Linear Algebra, Part II.

$$\mathcal{N}(A) = \Big(\mathcal{R}(A^T)\Big)^\perp \qquad \mathcal{R}(A^T) = \Big(\mathcal{N}(A)\Big)^\perp$$

$$\mathcal{N}(A^T) = \Big(\mathcal{R}(A)\Big)^\perp \qquad \mathcal{R}(A) = \Big(\mathcal{N}(A^T)\Big)^\perp$$

We already know that $A\vec{x} = \vec{b}$ has a solution if $\vec{b} \in \mathcal{R}(A)$. This second equality also says that $A\vec{x} = \vec{b}$ has a solution if $\vec{b}$ is orthogonal to every vector in $\mathcal{N}(A^T)$; i.e., if $A^T\vec{w} = \vec{0}$, then $\vec{b}^T\vec{w} = 0$.

Two spaces can be orthogonal and not be orthogonal complements of each other. However, in R^n if V, W are orthogonal and if the dimension of V plus the dimension of W is n, then they are orthogonal complements. A very useful fact is that if V, W are orthogonal complements in R^n then any vector $\vec{x} \in \mathrm{R}^n$ can be written as the sum of a vector in V

and one in W, i.e., there exists $\vec{v} \in V$, $\vec{w} \in W$ such that $\vec{x} = \vec{v} + \vec{w}$. However, this result only says it is possible to do this, it doesn't say how to construct such a decomposition. In the next section we will see how these results can be used to understand how to solve the linear least squares problem of fitting a polynomial to a set of data where there are more data points than degrees of freedom in the polynomial; i.e., we have an over-determined system which probably doesn't have a solution.

Before we move on to linear least squares, lets first look at the matrix $A^T A$ where A is $m \times n$. Clearly it is a square symmetric matrix because

$$(A^T A)^T = A^T (A^T)^T = A^T A$$

but is it positive definite? Recall that a square matrix B is positive definite if $\vec{x}^T B \vec{x} > 0$ for all $\vec{x} \neq 0$. We have

$$x^T (A^T A) x = (x^T A^T)(Ax) = (Ax)^T (Ax) = y^T y \qquad \text{where } y = Ax$$

Now $y^T y$ is just the scalar (or inner or dot) product of a vector with itself or also it is the square of the Euclidean length of $\vec{y}$ which is always non-negative. It is only zero if $\vec{y} = \vec{0}$. Can $\vec{y}$ ever be zero? Remember that $y = A\vec{x}$ so if $\vec{x} \in \mathcal{N}(A)$ then $\vec{y} = \vec{0}$. When can the rectangular matrix A have something in the null space other than the zero vector? If we can take a linear combination of the columns of A (with coefficients nonzero) and get zero, i.e., if the columns of A are linearly dependent. In other words, if the rank of A is not n. We say that $A^T A$ is symmetric positive definite if the columns of A are linearly independent; otherwise it is positive semi-definite (meaning that $x^T A^T A x \geq 0$).