1 Last time: inverses and subspaces

To show that an $n \times n$ matrix A is *invertible*, all we have to do is check that (1) its columns are linearly independent or (2) its columns span $\mathbb{R}^n$. If either (1) or (2) holds, then the other property is also true.

If A is invertible then it has an *inverse* which is an $n \times n$ matrix A^{-1} with

$$AA^{-1} = A^{-1}A = I_n = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix}.$$

A subset H of $\mathbb{R}^n$ is a *subspace* if $0 \in H$ and $u + v \in H$ and $cv \in H$ for all $u, v \in H$ and $c \in \mathbb{R}$.

A subspace is a set which contains all linear combinations of vectors which are already in the set.

Example. Examples of subspaces of $\mathbb{R}^n$:

- The set $\{0\}$ containing just the zero vector.
- The set of all scalar multiples of a single vector.
- $\mathbb{R}^n$ itself.
- The span of any set of vectors in $\mathbb{R}^n$.
- The range of a linear function $T : \mathbb{R}^k \to \mathbb{R}^n$.
- The set of vectors v with $T(v) = 0$ for a linear function $T : \mathbb{R}^n \to \mathbb{R}^k$.

The union of two subspaces is not necessarily a subspace. (Why?)

The intersection of two subspaces is a subspace, however. (Why?)

Definition. To any $m \times n$ matrix A there are two corresponding subspaces of interest:

1. The *column space* of A is the subspace $\operatorname{Col} A \subset \mathbb{R}^m$ given by the span of the columns of A.
2. The *null space* of A is the subspace $\operatorname{Nul} A \subset \mathbb{R}^n$ given by the set of vectors $v \in \mathbb{R}^n$ with $Av = 0$.

It is not obvious from these definitions, but it will turn out that each subspace of $\mathbb{R}^m$ occurs as the column space of some matrix. Likewise, each subspace of $\mathbb{R}^n$ occurs as the null space of some matrix.

Exercise you might want to at least think about: how would you come up with a matrix whose column space or null space is a given subspace H? To answer this, you'll probably need the notion of a *basis*.

Definition. A *basis* of a subspace H of $\mathbb{R}^n$ is a set of linearly independent vectors whose span in H.

The most important basis in linear algebra (and the only one that has a standard notation) is the *standard basis* of $\mathbb{R}^n$: this is the basis consisting of the vectors $e_1, e_2, \dots, e_n$ where e_i is the vector in $\mathbb{R}^n$ with 1 in row i and 0 is all other rows.

The fundamental property of subspaces and bases:

Theorem. Every subspace H of $\mathbb{R}^n$ has a basis of size at most n.

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

How to find a basis of $\mathrm{Nul}\, A$**.**

1. Find all solutions to $Ax = 0$ by row reducing A to echelon form. Recall that x_i is a *basic variable* if column i of $\texttt{RREF}(A)$ contains a pivot position, and that otherwise x_i is a *free variable.*

2. Rewrite each basic variable in terms of the free variables, and then write

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_{i_1} \begin{bmatrix} b_1 \end{bmatrix} + x_{i_2} \begin{bmatrix} b_2 \end{bmatrix} + \cdots + x_{i_k} \begin{bmatrix} b_k \end{bmatrix}$$

where $x_{i_1}, x_{i_2}, \ldots, x_{i_k}$ are the free variables and $b_1, b_2, \ldots, b_k \in \mathbb{R}^n$ are vectors whose entries do not involve any variables.

3. The vectors $b_1, b_2, \ldots, b_k$ then form a basis for $\mathrm{Nul}\, A$.

Example. Suppose $A = \begin{bmatrix} 1 & 2 & 5 & 8 \\ 2 & 3 & 7 & 0 \end{bmatrix}$.

1. Then $A \sim \begin{bmatrix} 1 & 2 & 5 & 8 \\ 0 & -1 & -3 & -16 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -1 & -24 \\ 0 & 1 & 3 & 16 \end{bmatrix}$ so $Ax = 0$ iff $\begin{cases} x_1 - x_3 - 24x_4 = 0 \\ x_2 + 3x_3 + 16x_4 = 0. \end{cases}$

2. This means x_1, x_2 are basic variables and x_3, x_4 are free variables. We have $Ax = 0$ if and only if $x_1 = x_3 + 24x_4$ and $x_2 - 3x_3 - 16x_4$, which means

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} x_3 + 24x_4 \\ -3x_3 - 16x_4 \\ x_3 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} 1 \\ -3 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 24 \\ -16 \\ 0 \\ 1 \end{bmatrix}.$$

3. Thus $\left\{ \begin{bmatrix} 1 \\ -3 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 24 \\ -16 \\ 0 \\ 1 \end{bmatrix} \right\}$ is a basis for $\mathrm{Nul}\, A$.

How to find a basis of $\mathrm{Col}\, A$**.**

1. The pivot columns of A form a basis of $\mathrm{Col}\, A$.

This looks simpler than the previous algorithm, but to find out which columns of A are pivot columns, we have to row reduce A to echelon form, which takes just as much work as finding a basis of $\mathrm{Nul}\, A$.

Example. If $A = \begin{bmatrix} 1 & 2 & 5 & 8 \\ 2 & 3 & 7 & 0 \end{bmatrix}$ then columns 1, 2 have pivots so $\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix} \right\}$ is a basis for $\mathrm{Col}\, A$.

2 Coordinate systems

Suppose H is a subspace of $\mathbb{R}^n$.

Let $b_1, b_2, \ldots, b_k$ be a basis of H.

Theorem. Let $v \in H$. There are unique coefficients $c_1, c_2, \ldots, c_k \in \mathbb{R}$ such that

$$c_1 b_1 + c_2 b_2 + \cdots + c_k b_k = v.$$

In other words, there is a unique vector

$$w = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{bmatrix} \in \mathbb{R}^k$$

such that $v = Bw$ where $B = \begin{bmatrix} b_1 & b_2 & \dots & b_k \end{bmatrix}$.

Proof. Since our basis spans H, there must be some coefficients with $c_1 b_1 + c_2 b_2 + \dots + c_k b_k = v$. If these coefficients were not unique, so that we could write $c_1' b_1 + c_2' b_2 + \dots + c_k' b_k = v$ for some different list of numbers $c_1', c_2', \dots, c_k' \in \mathbb{R}$, then we would have

$$\begin{aligned} 0 = v - v &= (c_1 b_1 + c_2 b_2 + \dots + c_k b_k) - (c_1' b_1 + c_2' b_2 + \dots + c_k' b_k) \\ &= (c_1 - c_1') v_1 + (c_2 - c_2') v_2 + \dots + (c_k - c_k') v_k. \end{aligned}$$

Since our numbers are different, at least one of the differences $c_i - c_i'$ must be nonzero, so what we just wrote is a nontrivial linear dependence among the vectors $b_1, b_2, \dots, b_k$. But this contradicts the condition that elements of a basis be linearly independent. $\square$

Let $\mathcal{B} = (b_1, b_2, \dots, b_k)$ be the list of basis vectors in some fixed order.

Given $v \in H$, define $[v]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{bmatrix} \in \mathbb{R}^k$ as the unique vector with $c_1 b_1 + c_2 b_2 + \dots + c_k v_k = v$.

We call $[v]_{\mathcal{B}}$ the *coordinate vector of* v *in the basis* $\mathcal{B}$ or just v *in the basis* $\mathcal{B}$.

Example. If $H = \mathbb{R}^n$ and $\mathcal{B} = (e_1, e_2, \dots, e_n)$ is the standard basis then $[v]_{\mathcal{B}} = v$.

Example. If $H = \mathbb{R}^n$ and $\mathcal{B} = (e_n, \dots, e_2, e_1)$ then $[v]_{\mathcal{B}} = \begin{bmatrix} v_n \\ \vdots \\ v_2 \\ v_1 \end{bmatrix}$.

Example. Let $b_1 = \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix}$ and $b_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ and $v = \begin{bmatrix} 3 \\ 12 \\ 7 \end{bmatrix}$.

Then $\mathcal{B} = (b_1, b_2)$ is a basis for $H = \mathbb{R}\text{-span}\{b_1, b_2\} \subset \mathbb{R}^3$.

The unique $w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \in \mathbb{R}^2$ such that $\begin{bmatrix} 3 & -1 \\ 6 & 0 \\ 2 & 1 \end{bmatrix} w = \begin{bmatrix} 3 \\ 12 \\ 7 \end{bmatrix}$ is found by row reduction:

$$\begin{bmatrix} 3 & -1 & 3 \\ 6 & 0 & 12 \\ 2 & 1 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 2 \\ 3 & -1 & 3 \\ 2 & 1 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 2 \\ 0 & -1 & -3 \\ 0 & 1 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}.$$

The last matrix implies that $w_1 = 2$ and $w_2 = 3$ so $[v]_{\mathcal{B}} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$.

Example. If $b_1 = e_1 - e_2$, $b_2 = e_2 - e_3$, $b_3 = e_3 - e_4$, $\ldots$, $b_{n-1} = e_{n-1} - e_n$ and

$$v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_{n-1} \\ -v_1 - v_2 - \cdots - v_{n-1} \end{bmatrix}$$

then $v \in H = \mathbb{R}\text{-span}\{b_1, b_2, \ldots, b_{n-1}\}$ and

$$[v]_{\mathcal{B}} = \begin{bmatrix} v_1 \\ v_1 + v_2 \\ v_1 + v_2 + v_3 \\ v_1 + v_2 + v_3 + v_4 \\ \vdots \\ v_1 + v_2 + v_3 + \cdots + v_{n-1} \end{bmatrix} \in \mathbb{R}^{n-1}.$$

The notation $[v]_{\mathcal{B}}$ gives us an easy way to check the following important property:

Theorem. Let H be a subspace of $\mathbb{R}^n$. Then all bases of H have the same number of elements.

Proof. Suppose $\mathcal{B} = (b_1, b_2, \ldots, b_k)$ and $\mathcal{B}' = (b'_1, b'_2, \ldots, b'_l)$ are two ordered bases of H with $k < l$.

Then $[b'_1]_{\mathcal{B}}$, $[b'_2]_{\mathcal{B}}$, $\ldots$, $[b'_l]_{\mathcal{B}}$ are $l > k$ vectors in $\mathbb{R}^k$, so they must be linearly dependent.

This means there exist coefficients $c_1, c_2, \ldots, c_l \in \mathbb{R}$, not all zero, with

$$c_1[b'_1]_{\mathcal{B}} + c_2[b'_2]_{\mathcal{B}} + \cdots + c_l[b'_l]_{\mathcal{B}} = 0.$$

But we have

$$c_1[b'_1]_{\mathcal{B}} + c_2[b'_2]_{\mathcal{B}} + \cdots + c_l[b'_l]_{\mathcal{B}} = [c_1 b'_1 + c_2 b'_2 + \cdots + c_l b'_l]_{\mathcal{B}}.$$

(This is the key step; why is this true? Think about how $[v]_{\mathcal{B}}$ is defined.)

Thus $[c_1 b'_1 + c_2 b'_2 + \cdots + c_l b'_l]_{\mathcal{B}} = 0$, so

$$c_1 b'_1 + c_2 b'_2 + \cdots + c_l b'_l = \begin{bmatrix} b_1 & b_2 & \cdots & b_k \end{bmatrix} [c_1 b'_1 + c_2 b'_2 + \cdots + c_l b'_l]_{\mathcal{B}} = 0.$$

(The first equality holds since by definition $v = B[v]_{\mathcal{B}}$ for $B = \begin{bmatrix} b_1 & b_2 & \cdots & b_k \end{bmatrix}$.)

Since the coefficients c_i are not all zero, this contradicts the fact that $b'_1, b'_2, \ldots, b'_l$ are linearly independent. This mean our original supposition that H has two bases of different sizes can't hold. $\square$

3 Dimension

Points in the subspace $H = \mathbb{R}\text{-span}\{b_1, b_2, \ldots, b_k\} \subset \mathbb{R}^n$, while contained in the set of vectors $\mathbb{R}^n$, are completely determined by their coordinate vectors which belong to $\mathbb{R}^k$.

The correspondence $v \mapsto [v]_{\mathcal{B}}$ is an invertible function $H \to \mathbb{R}^k$. We call this function an *isomorphism* between H and $\mathbb{R}^k$: its existence means that H "looks" the same as $\mathbb{R}^k$.

For this reason we say that a subspace H with a basis of size k is *k-dimensional*. More generally:

Definition. The *dimension* of a subspace H is the number of vectors in any basis of H.

We denote the dimension of H by $\dim H$. This number belongs to $\{0, 1, 2, 3, \dots\}$.

If $H = \{0\}$ then we define $\dim H = 0$.

Example. We have $\dim \mathbb{R}^n = n$.

If H is the set of all vectors of the form $\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_k \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \in \mathbb{R}^n$, then H is a subspace and $\dim H = k$.

Note that $e_1, e_2, \dots, e_k$ is a basis for H.

A line in $\mathbb{R}^2$ through the origin is a 1-dimensional subspace.

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

The processes we gave to construct bases of $\operatorname{Nul} A$ and $\operatorname{Col} A$ imply that:

Corollary. The dimension of $\operatorname{Nul} A$ is the number of free variables in the linear system $Ax = 0$.

Corollary. The dimension of $\operatorname{Col} A$ is the number of pivot columns in A.

There is a special name for the dimensional of the column space of a matrix:

Definition. The *rank* of a matrix A is $\operatorname{rank} A = \dim \operatorname{Col} A$.

Putting everything together gives the following pair of important results.

Theorem (Rank-Nullity theorem)**.** If A is a matrix with n columns then $\operatorname{rank} A + \dim \operatorname{Nul} A = n$.

Proof. Every column of A which is not a pivot column indexes a free variable in the system $Ax = 0$, so

$$\begin{aligned} n &= \#\{ \text{ pivot columns of } A \ \} + \#\{ \text{ non-pivot columns } \} \\ &= \#\{ \text{ pivot columns of } A \ \} + \#\{ \text{ free variables in } Ax = 0 \ \} \\ &= \dim \operatorname{Col} A + \dim \operatorname{Nul} A \\ &= \operatorname{rank} A = \dim \operatorname{Nul} A. \end{aligned}$$

□

Theorem (Basis theorem)**.** If H is a subspace of $\mathbb{R}^n$ with $\dim H = p$ then

1. Any set of p linearly independent vectors in H is a basis for H.
2. Any set of p vectors in H which span H is a basis for H.

Proof. Suppose we have p linearly independent vectors in H. If these vecttors do not span H, then adding a vector which is in H but not in their span would produce a set of $p+1$ linearly independent vectors in H. If this larger set still does not span H, then adding a vector from H that is not in the span gives an even larger linearly independent set of $p+2$ vectors. Continuing in this way must eventually produce a basis for H, but this basis will have more than p elements, contradicting $\dim H = p$.

Suppose we instead have p vectors which span H. If these vectors are not linearly independent, then one of the vectors is a linear combination of the others. Remove this vector; we then have $p-1$ vectors which span H. If these vectors are still not linearly independent, then one is a linear combination of the others and removing this vector gives a set of $p-2$ vectors which span H. Continuing in this way must eventually produce a basis for H, but this basis will have fewer than p elements, contradicting $\dim H = p$. □

Corollary. If $H \subset \mathbb{R}^n$ is an n-dimensional subspace then $H = \mathbb{R}^n$.

Proof. If H has a basis with n elements then these elements are linearly independent, so form a basis for $\mathbb{R}^n$. Then every vector in $\mathbb{R}^n$ is a linear combination of the basis vectors, so belongs to H. □

Corollary. Let A be an $n \times n$ matrix. The following are equivalent:

(a) A is invertible.

(b) The columns of A form a basis for $\mathbb{R}^n$.

(c) $\operatorname{rank} A = \dim \operatorname{Col} A = n$.

(d) $\dim \operatorname{Nul} A = 0$.

Proof. We have already seen that (a) and (b) are equivalent.

(c) holds if and only if the columns of A span $\mathbb{R}^n$ which is equivalent to (a).

(d) holds if and only if the columns of A are linearly independent which is equivalent to (a). □