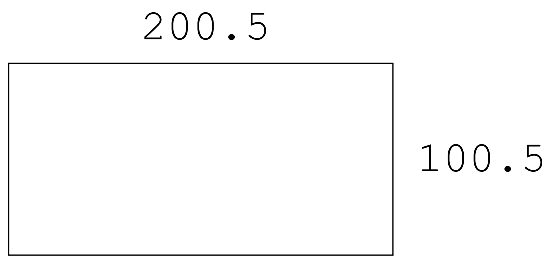
Math 1b TA: Padraic Bartlett

Recitation 3: Row Space and Null Space

Week 3 *Caltech 2011*

1 Random Question

Can you divide the following rectangle up into smaller rectangles, so that each small rectangle has at least one side with integral length?



2 HW comments

- Section average: 85%, or 68/80.
- On the whole, people did rather well! I have just a few comments to make:
- If you use Mathematica when completing your assignment, you **must** include your code to receive full credit! Otherwise we have no way of seeing *how* you got your answers, which is really much more interesting/important than some random $n \times k$ matrix.
- Don't write things like $\sqrt{1/5} = .44$, or $\pi = 3$! They anger your TA. If you want to indicate that you've found a decimal approximation, the appropriate mathematical notation to use is $\sqrt{1/5} \approx .44$, or (sometimes) $\sqrt{1/5} \cong .44$. I won't take off points for this, because I know where you guys are coming from: but seriously, good mathematical notation is important. Also, things like π or $\sqrt{1/5}$ are completely valid answers! – you really don't need to find these approximations at all.
- When you're showing that a collection of vectors V is linearly dependent, you need to do two things:

  1. find a collection of constants $c_i \in \mathbb{R}$ and vectors $v_i \in V$ such that the sum $\sum_{i=1}^{n} a_i v_i = 0$, **and**
  2. show that not all of these constants are 0.

  Almost everyone forgot to check the second condition. In the future, do not forget to check the second condition!

3 A Sea of Definitions

To get to the topic of today's lecture – row space and null space – we first have to go through a number of definitions. We present six concepts here, and provide examples to illustrate what they are trying to describe:

Definition. A **vector space** over $\mathbb{R}$ is a set V and a pair of operations $+ : V \times V \to V$ and $\cdot : \mathbb{R} \times V \to V$, that are in a certain sense "well-behaved:" i.e. the addition operation is associative and commutative, there are additive identites and inverses, the addition and multiplication operations distribute over each other, the scalar multiplication is compatible with multiplication in $\mathbb{R}$, and 1 is the multiplicative identity.[1]

Example.

- $\mathbb{R}^n$ is a vector space, for any $n \in \mathbb{N}$: vector addition is defined by summing each coördinate separately as real numbers, and scalar multiplication multiplies the entire vector by the scalar.
- $\mathbb{R}[x]$, the space of all polynomials with real coefficients, is a vector space: the sum of two polynomials is just their regular sum, and multiplying a polynomial by a scalar also works just as it normally does.
- For a set X, the collection of all functions $X \to \mathbb{R}$ is a vector space. The sum of two functions f, g is the function $f + g$ that yields output $f(\mathbf{x}) + g(\mathbf{x})$ on input $\mathbf{x}$; similarly, multiplying a function f by a scalar a yields the function $a \cdot f$, which outputs $a \cdot f(\mathbf{x})$ on input $\mathbf{x}$.

Definition. A subset S of a vector space V is called a **subspace** iff for any $\mathbf{x}, \mathbf{y} \in S$ and $a, b \in \mathbb{R}$, we have that $a\mathbf{x} + b\mathbf{y}$ is also an element of S.

It bears noting that any subspace is a vector space in its own right.

Example. In $\mathbb{R}^3$, every subspace has one of the following forms:

- $\{(0, 0, 0)\}$, the subspace consisting of one single point, the origin.
- A line in $\mathbb{R}^3$ through the origin.
- A plane in $\mathbb{R}^3$ that contains the origin.
- All of $\mathbb{R}^3$.

Definition. A **basis** for a space V is a collection B of vectors in V, that satisfies the following two properties:

- B is linearly independent: i.e. you cannot write any element in B as a finite linear combination of the other elements in B.
- For any $v \in V$, there are elements $a_1 \ldots a_n \in \mathbb{R}, b_1 \ldots b_n \in B$ such that $v = \sum_{i=1}^{n} a_i b_i$. In other words, we can express any element in V as a finite linear combination of elements in B.

[1] See Wikipedia if you want a precise description of these properties.

Example. The "standard" basis for $\mathbb{R}^4$ consists of the following four vectors:

$$\{(1,0,0,0),(0,1,0,0),(0,0,1,0),(0,0,0,1)\}.$$

However, the following set is also a basis:

$$\{(2,0,0,0),(0,\pi,0,0),(0,0,-1,0),(0,0,0,1)\}.$$

Similarly, the following set is another, third possible choice for a basis:

$$\{(1,1,1,0),(1,1,0,1),(1,0,1,1),(0,1,1,1)\}.$$

Definition. The **dimension** of a space V is the number of elements in any basis of V. (Note that this definition implies that any two bases for a given vector space will always have the same number of elements in them – otherwise, a space could have several different possible dimensions! If you have time, think about why this can never happen.)

Example. $\mathbb{R}^{\ltimes}$ has dimension n. Furthermore, in our example above where we discussed the various subspaces S of $\mathbb{R}^3$,

- when S is a line in $\mathbb{R}^3$, its dimension is 1,
- when S is a plane in $\mathbb{R}^3$, its dimension is 2, and
- when S is all of $\mathbb{R}^3$, its dimension is 3.

As a special case, when the space in question is the trivial subspace $S = \{\mathbf{0}\}$ consisting only of the $\mathbf{0}$-vector, we say that this subspace is 0-dimensional.

Definition. For a set S of vectors inside of some vector space V, the **span** of S is the following subspace:

$$\left\{ v \in V : v = \sum_{i=1}^{n} a_i s_i, \text{ for some } a_1 \dots a_n \in \mathbb{R}, s_1 \dots s_n \in S. \right\}$$

In other words, the span of S is the subspace formed by all of the vectors that can be expressed via finite linear combinations of the elements of S.

Example. The span of the vectors $\{(1,0,0,0),(1,1,0,0),(1,1,1,0)\}$ is the following subspace of $\mathbb{R}^4$:

$$\{(w,x,y,0) : w,x,y \in \mathbb{R}.\}$$

Definition. For a subspace S of a vector space V, we define the **orthogonal complement** $S^{\perp}$ as the following set:

$$S^{\perp} = \{v \in V : \langle v, s\rangle = 0, \forall s \in S\}.$$

In other words, $S^{\perp}$ is the collection of all of the vectors v that are orthogonal[2] to every element of S. As it turns out, this actually forces S to become a subspace. (Why?)

[2] Recall: two vectors $\mathbf{u}, \mathbf{v}$ are said to be orthogonal iff their inner product $\langle \mathbf{u}, \mathbf{v}\rangle$ is 0. Also, recall that the inner product of two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ is their dot product: i.e.$\langle \mathbf{u}, \mathbf{v} = \sum_{i=1}^{n} u_i v_i$.

Example. For the subspace

$$S = \{(w, x, y, 0) : w, x, y \in \mathbb{R}\}$$

we discussed in our previous example, the orthogonal complement $S^\perp$ is the collection of all vectors $(a, b, c, d) \in \mathbb{R}$ such that their inner product with elements in S is always 0. So: because:

- S specifically contains the elements $(1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0)$, and
- $\langle (1, 0, 0, 0), (a, b, c, d) \rangle = a$,
- $\langle (0, 1, 0, 0), (a, b, c, d) \rangle = b$,
- $\langle (0, 0, 1, 0), (a, b, c, d) \rangle = c$,

we know that if (a, b, c, d) is always orthogonal to all of S, we need to have it orthogonal to the three vectors $(1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0)$: i.e. we need a, b, c all equal to 0. This then leaves us with vectors of the form $(0, 0, 0, d)$, which are all orthogonal to all of S: therefore,

$$S^\perp = \{(0, 0, 0, z) : z \in \mathbb{R}\}.$$

4 Row Space and Null Space: Definitions and Observations

With these definitions set, we have the ability to finally define the topics of today's lecture:

Definition. For a $n \times k$ matrix M, the **row space** of M is the subspace of $\mathbb{R}^n$ spanned by the k rows of M.

Example. The row space of the matrix

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

is all of $\mathbb{R}^4$.

Definition. For a $n \times k$ matrix M, the **null space** of M is the following set:

$$\{\mathbf{x} \in \mathbb{R}^n : M \cdot \mathbf{x} = \mathbf{0}\}.$$

In other words, the null space of M is the collection of all of the vectors in $\mathbb{R}^n$ that nullifies M. The null space of M is always a subspace of $\mathbb{R}^n$; prove this, if you are so inclined.

Example. The null space of the matrix

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

is all of $\mathbb{R}^4$.

In lecture, you will hopefully run into the following two remarkable theorems about row spaces and null spaces:

Theorem 1 *For a $n \times k$ matrix M, the dimension of the row space plus the dimension of the null space is always n.*

Theorem 2 *The orthogonal complement to the row space is the null space.*

So: our two examples above were rather "trivial" – you didn't really need to do very much to calculate these specific row and null spaces. How can we find the row and null spaces for general matrices?

As it turns out, we have a number of methods for finding the row space or null space of a given matrix. We detail them in the next two sections:

5 Finding the Row Space: Two Methods

Here, we describe two methods for finding the row space of a matrix, give examples of how each method works, and talk a bit about why these methods "work." (If you would like to see a formal proof of any of the following algorithms, ask me in office hours or by email! Hopefully, though, the ideas here should make the path to actually proving these theorems fairly clear.)

Theorem 3 *If M is a matrix, and M' is the matrix M after reducing it into a basic form, then the nonzero rows of M' form a basis for the row space of M.*

Example. To find the row space of the matrix

$$M = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 10 \\ 11 & 12 & 13 & 14 & 15 \\ 16 & 17 & 18 & 19 & 20 \\ 21 & 22 & 23 & 24 & 25 \end{pmatrix}$$

via this method, we first reduce it into basic form, by pivoting at (1,1):

$$M = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & -5 & -10 & -15 & -20 \\ 0 & -10 & -20 & -30 & -40 \\ 0 & -15 & -30 & -45 & -50 \\ 0 & -20 & -40 & -60 & -80 \end{pmatrix}$$

and then at (2,2):

$$M = \begin{pmatrix} 1 & 0 & -1 & -2 & -3 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

By our theorem, the two nonzero rows $(1, 0, -1, -2, -3)$ and $(0, 1, 2, 3, 4)$ form a basis for the row space of M.

Why does this work? Because all of the operations that we perform on a matrix to put it in basic form are reversible, we know that the rows of M' have the same span as the rows of M, because we can simply *create* the rows of M from M' rows. So the only interesting thing to check is whether M's nonzero rows form a linearly independent set. But this is immediate, as any nonzero row in M' has to have a 1 in some slot from a unit column vector in it, and all of the other rows in M' have to have a 0 in that same slot. Therefore, no nontrivial combination of these rows can kill off these unit-column parts – so the set must be linearly independent, and therefore a basis for M's row space.

Our second theorem is similar:

Theorem 4 *If M is a matrix, and A is the matrix M^T after reducing it into a basic form, then the rows of M that correspond to the unit column vectors in A form a basis for M's row space.*

Example. To find the row space of the matrix

$$M = \begin{pmatrix} 2 & 7 & 1 & 8 & 2 & 8 \\ 1 & 4 & 1 & 4 & 2 & 1 \\ 0 & 1 & 1 & 0 & 2 & -6 \end{pmatrix}$$

via this method, we first take its transpose

$$M^T = \begin{pmatrix} 2 & 1 & 0 \\ 7 & 4 & 1 \\ 1 & 1 & 1 \\ 8 & 4 & 0 \\ 2 & 2 & 2 \\ 8 & 1 & -6 \end{pmatrix}$$

and then reduce this matrix into a basic form, by first pivoting at (2,1):

$$\begin{pmatrix} 2 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \\ 2 & 0 & 2 \\ -6 & 0 & -6 \end{pmatrix},$$

and then at (3,2):

$$\begin{pmatrix} 2 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Then, our theorem tells us that the rows of M corresponding to the second and third columns of this matrix – i.e. the rows $(1, 4, 1, 4, 2, 1)$ and $(0, 1, 1, 0, 2, -6)$ – form a basis for M's row space.

Why does this work? So: in the example above, pivoting at $(2,1)$ caused us to subtract 2 times the second column from all of the entries in the first column below the first row. Similarly, pivoting at $(3,2)$ caused us to subtract 1 copy of the third column from the first column in all of the entries below the second row. Doing these two operations then zeroed out the entire first column beneath the first two rows! What does this mean? Well, if we stop looking at M^T and return to just M, we've just shown that the first row can be written as two copies of the second row minus the third row!

In general, this process works in this fashion: pivoting in M^T amounts to row operations in M, and the columns in M^T that don't turn into unit column vectors always correspond to rows in M that are linearly dependent on the others (and thus can be discarded.)

Now, we turn our attention to finding the nullspace of a given matrix:

6 Finding the Null Space: Two Methods

Theorem 5 *Suppose that we start with a $n \times k$ matrix A. Suppose that the matrix*

$$\left.\left(\underbrace{A^T}_{k}\,\middle|\,\underbrace{I}_{n}\right)\right\}n$$

formed by taking A^T and attaching a $k \times k$ identity matrix to it, has the following reduced row-echelon form:

$$\left(\begin{array}{c|c} E & F \\ \hline 0 & G \end{array}\right)$$
$$\underbrace{\quad}_{k}\;\underbrace{\quad}_{n}$$

Then the nonzero rows of G form a basis for the nullspace of A.

Example. If we let

$$A = \begin{pmatrix} 0 & 8 & 6 & 7 \\ 5 & 3 & 0 & 9, \end{pmatrix},$$

then the matrix $(A^T|I)$ is simply

$$\begin{pmatrix} 0 & 5 & 1 & 0 & 0 & 0 \\ 8 & 3 & 0 & 1 & 0 & 0 \\ 6 & 0 & 0 & 0 & 1 & 0 \\ 7 & 9 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Switching rows 1 and 3 and then pivoting at (1,1) yields

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 1/6 & 0 \\ 0 & 3 & 0 & 1 & -4/3 & 0 \\ 0 & 5 & 1 & 0 & 0 & 0 \\ 0 & 9 & 0 & 0 & -7/6 & 1 \end{pmatrix};$$

finally, switching rows 2 and 3 and pivoting at (2,2) yields

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 1/6 & 0 \\ 0 & 1 & 1/5 & 0 & 0 & 0 \\ 0 & 0 & -3/5 & 1 & -4/3 & 0 \\ 0 & 0 & -9/5 & 0 & -7/6 & 1 \end{pmatrix}.$$

If we partition up the matrix into four blocks, the bottom-left one of which is all-zeroes, according to the picture in our theorem, this matrix becomes

$$\left(\begin{array}{cc|cccc} 1 & 0 & 0 & 0 & 1/6 & 0 \\ 0 & 1 & 1/5 & 0 & 0 & 0 \\ \hline 0 & 0 & -3/5 & 1 & -4/3 & 0 \\ 0 & 0 & -9/5 & 0 & -7/6 & 1 \end{array}\right).$$

Therefore, according to our theorem, the two rows of the bottom-right block

$$\{(-3/5, 1, -4/3, 0), (-9/5, 0, -7/6, 1)\}$$

form a basis for the nullspace of A.

Why does this work? First, notice that the nullspace of A is the same as the nullspace of the reduced row-echelon form of A, as being orthogonal to all of A's rows is equivalent to being orthogonal to all of the rows in any matrix acquired via elementary row operations from A.

So: when we convert $(A^T|I)$ into its $\left(\begin{array}{c|c} E & F \\ 0 & G \end{array}\right)$ form, what do the zero rows in the A^T-part of the matrix represent? Well, they represent combinations of the rows of A^T that combine to 0 – combinations that are recorded by the I-part of the matrix, as the rows of G! So, if we multiply a row of G on the right by A, this matrix tells us that the product will always be a zero vector – therefore, these rows are giving us precisely a basis for the null space of A.

Theorem 6 *Suppose that the reduced row-echelon form of some matrix A, with zero rows deleted, is of the form $(I|B)$ (where I is the identity matrix of some appropriate dimension.) Then a basis for the nullspace of A is given by the rows of the matrix $(-B^T|I)$, where I is again an identity matrix of the right dimension to be appended to $-B^T$.*

Example. Let

$$A = \begin{pmatrix} 6 & 6 & 2 & 6 & 10 & -34 \\ 6 & 0 & 2 & 2 & 10 & 23 \\ 0 & 6 & 0 & 4 & 0 & -57 \end{pmatrix}.$$

Pivot at (1,1):

$$\begin{pmatrix} 1 & 1 & 1/3 & 1 & 5/3 & -17/3 \\ 0 & -6 & 0 & -4 & 0 & 57 \\ 0 & 6 & 0 & 4 & 0 & -57 \end{pmatrix}.$$

Now, pivot at (2,2):

$$\begin{pmatrix} 1 & 0 & 1/3 & 1/3 & 5/3 & 23/6 \\ 0 & 1 & 0 & 2/3 & 0 & -57/6 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Discard the zero row of this matrix, and divide it up into its I and B parts:

$$\left(\begin{array}{cc|cccc} 1 & 0 & 1/3 & 1/3 & 5/3 & 23/6 \\ 0 & 1 & 0 & 2/3 & 0 & -57/6 \end{array}\right).$$

Take the B-part of this matrix, take its transpose, make it negative, and append an identity matrix of the right size (in this case, 4×4) to it:

$$\begin{pmatrix} -1/3 & 0 & 1 & 0 & 0 & 0 \\ -1/3 & -2/3 & 0 & 1 & 0 & 0 \\ -5/3 & 0 & 0 & 0 & 1 & 0 \\ -23/6 & 57/6 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

According to our theorem, the rows of the above matrix form a basis for the null space of A.

Why does this work? Just like before, we first notice that finding a basis for A's nullspace is equivalent to finding a basis for the reduced row-echelon form of A. The zero rows of this reduced row-echelon form will never cause us any trouble – so all we have to care about is zeroing out the rows of the $(I|B)$ matrix.

So: to do this, simply notice (by working it out by hand) that the product

$$(I|B) \cdot \left(-B^T|I\right)$$

is always an all-zeroes matrix. Consequently, the rows of $(-B^T|I)$ form a basis for the nullspace of A.