Lecture Notes for Math 104: Fall 2010 (Week 2)

Jacob Bernstein

CHAPTER 1

Fourth Lecture

In this lecture we reviewed Gaussian elimination. We focused on the difference between row operations and column operations and how those could be used to determine different information about (respectively) the null space of a matrix and the column space.

1. Elementary Row and Column Operations

The basic tool we will use are a certain set of operations on the rows and columns of a matrix. These will preserve important features of the matrix while also simplifying the matrix.

Fix a matrix $A \in \mathbb{C}^{m\times n}$. We've already seen how to write A in terms of its columns:

$$A = \begin{pmatrix} \mathbf{a}_1 & | & \dots & | & \mathbf{a}_n \end{pmatrix}$$

here $\mathbf{a}_j$ is vector in $\mathbb{C}^m$ or a $m \times 1$ matrix. it is also helpful to write it in terms of its rows

$$A = \begin{pmatrix} c_1 \\ - \\ \vdots \\ - \\ c_m \end{pmatrix}$$

where here the c_j is a $1 \times n$ matrix (not a vector!).

Starting with A we define an elementary column operation to be one of the following operations: a) Swapping two columns of A, b) scaling the first column by a non-zero scaler and c) adding the second column to the first column. In general we will be iteratively applying a sequence such operations to A.

More precisely, we take and produce a new $m \times n$ matrix A' by doing one of the preceding operations as follows:

$$A = \begin{pmatrix} \mathbf{a}_1 & |\dots| & \mathbf{a}_i & |\dots| & \mathbf{a}_j & |\dots| & \mathbf{a}_n \end{pmatrix}$$

under the swapping operation, a), goes to

$$A' = \begin{pmatrix} \mathbf{a}_1 & |\dots| & \mathbf{a}_j & |\dots| & \mathbf{a}_i & |\dots| & \mathbf{a}_n \end{pmatrix}.$$

Here we are free to choose any $1 \leq i < j \leq n$ we like. Ander the scaling operation, b), one has A going to

$$A' = \begin{pmatrix} \lambda\mathbf{a}_1 & |\dots| & \mathbf{a}_i & |\dots| & \mathbf{a}_j & |\dots| & \mathbf{a}_n \end{pmatrix}$$

where here $\lambda \neq 0$ is a scaler in $\mathbb{C}$. Finally, under the addition operation, c), A goes to

$$A' = \begin{pmatrix} \mathbf{a}_1 + \mathbf{a}_2 & |\mathbf{a}_2| & \dots| & \mathbf{a}_i & |\dots| & \mathbf{a}_j & |\dots| & \mathbf{a}_n \end{pmatrix}.$$

Notice that by combining the swapping operation with the scaling operation, one obtains an operation given by scaling any column by a non-zero scalar. Similarly, by combining the swapping operation with the addition operation one gets an operation wherein any column may be added to any other. This larger set of operations is often refered to as *elementary column operations*.

The nice feature of elementary column operations is that they preserve the range space of a matrix.

THEOREM 1.1. *Let $A \in \mathbb{C}^{m\times n}$. If A' is obtained from A by a (finite) sequence of elementary column operations then $R(A) = R(A')$*

REMARK 1.2. In general $N(A) \neq N(A')$, though it is true that $dim N(A) = dim N(A')$.

PROOF. We verify this only when A' is obtained from A by one elementary column operation. The theorem then follows by induction. Lets first verify that $R(A') = R(A)$ when A' is obtained from A by swapping the ith and jth column. Let $\mathbf{v} \in R(A)$ then there is a $\mathbf{w} \in \mathbb{C}^n$ so that $\mathbf{v} = A\mathbf{w}$. We may write $\mathbf{w} = w_1\mathbf{e}_1 + \cdots + w_i\mathbf{e}_i + \cdots + w_j\mathbf{e}_j + \cdots + w_n\mathbf{e}_n$ where $\mathbf{e}_k$ is the kth standard basis vector. Now let A' be obtained from A by swapping the ith and jth columns. If we set $\mathbf{w}' = \mathbf{w} - w_i\mathbf{e}_i - w_j\mathbf{e}_j + w_j\mathbf{e}_i + w_i\mathbf{e}_j$ then one verifies that $A'\mathbf{w}' = \mathbf{v}$. Hence $R(A) \subset R(A')$. However, reversing the argument works just as well so $R(A) = R(A')$.

Now suppose that A' is obtained by A by scaling by $\lambda \neq 0$ the first column of A. If $\mathbf{v} \in R(A)$ then $\mathbf{v} = A\mathbf{w}$ where $\mathbf{w} = \sum_{i=1}^n w_i\mathbf{e}_i$ then if we set $\mathbf{w}' = \frac{w_1}{\lambda}\mathbf{e}_1 + \sum_{i=2}^n w_i\mathbf{e}_i$ then $A'\mathbf{w}' = A\mathbf{w} = \mathbf{v}$. Hence in this case also $R(A) \subset R(A')$. Again the argument is reversible so $R(A) = R(A')$.

Finally, suppose A' is obtained from A by adding the second column to the first. If $\mathbf{v} \in R(A)$ then $\mathbf{v} = A\mathbf{w}$ where $\mathbf{w} = \sum_{i=1}^n w_i\mathbf{e}_i$. Now set $\mathbf{w}' = w_1\mathbf{e}_1 + (w_2 - w_1)\mathbf{e}_2 + \sum_{i=3}^n w_i\mathbf{e}_i$. Then one has $A'\mathbf{w}' = A\mathbf{w} = \mathbf{v}$ so $R(A) \subset R(A')$. Again the argument is reversible so $R(A) = R(A')$. □

In a corresponding way we may define the elementary row operations. Roughly speaking, an *elementary row operation* is one of the following operations: a) swapping two rows, b) scaling the first row by a non-zero scalar, or c) adding the second row to the first row. More precisely, for a matrix A a new matrix A' is obtained by an elementary row operation applied from A if it is given by one of the following formulas:

$$A = \begin{pmatrix} c_1 \\ \vdots \\ c_i \\ \vdots \\ c_j \\ \vdots \\ c_m \end{pmatrix}, A' = \begin{pmatrix} c_1 \\ \vdots \\ c_j \\ \vdots \\ c_i \\ \vdots \\ c_m \end{pmatrix}, A' = \begin{pmatrix} \lambda c_1 \\ \vdots \\ c_i \\ \vdots \\ c_j \\ \vdots \\ c_m \end{pmatrix}, A' = \begin{pmatrix} c_1 + c_2 \\ \vdots \\ c_i \\ \vdots \\ c_j \\ \vdots \\ c_m \end{pmatrix}$$

Unlike the elementary column operations, the elementary row operations preserve the null space, though they change the column space.

THEOREM 1.3. *Let $A \in \mathbb{C}^{m\times n}$ be a matrix. If A' is obtained from A by a (finite) sequence of elementary row operations then $N(A) = N(A')$.*

REMARK 1.4. In general $R(A) \neq R(A')$, though it is true that $dim R(A) = dim R(A')$. We will prove this later.

2. Reduced Echelon Form and Gaussian elimination

Elementary row operations and elementary column operations can be applied to a matrix A to produce new matrices A' and A'' that are "simpler" in a certain sense. In order to make this precise, we need a notion of what a "simple" matrix should be.

We make the following defintions:

DEFINITION 2.1. Let $A \in \mathbb{C}^{m\times n}$ and write

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

we say A is in *row reduced echelon form* (rref) if

(1) A is upper triangular, i.e. if $a_{ij} = 0$ when $i > j$.
(2) The first non-zero entry of each row of A is 1 (note some rows may be all zeros). This entry is called a *pivot*.
(3) The only non-zero entry in a column containing a pivot is the pivot entry.

DEFINITION 2.2. Let $A \in \mathbb{C}^{m\times n}$ and write

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

we say A is in *column reduced echelon form* (cref) if

(1) A is lower triangular, i.e. if $a_{ij} = 0$ when $i < j$.
(2) The first non-zero entry of each column of A is 1 (note some columns may be all zeros). This entry is called a *pivot*.
(3) The only non-zero entry in a row containing a pivot is the pivot entry.

The point about the echelon forms is that (as well discuss below) they are easier to extract information about the range and null space from. The point is that that one can always obtain a matrix A' which is in rref from A by elementary row operations. I won't discuss the details the algorithm to do this – namely Gaussian elimination – as you learned it in Math 51. In a similar manner, and by essentially the same algorithm for any matrix A one can obtain a matrix A' from A by column operations and so that A' is in cref.

The existence of such the Gaussian elimination algorithm is then a proof of the following theorems which we will use in this class.

THEOREM 2.3. *For any matrix $A \in \mathbb{C}^{m\times n}$ there is a unique matrix A' obtained by a finite number of elementary row operations from A so that A' is in rref. We write $A' = rref(A)$.*

Similarly, there is a unique matrix A'' obtained by a finite number of elementary column operations from A and so that A'' is in cref. We write $A'' = cref(A)$.

3. Pivots

The pivots of $cref(A)$ and $rref(A)$ allow one to determine (respectively) the range or the null space of A more easily. For instance one can check that the set of columns of $cref(A)$ containing a pivot forms a basis of $R(cref(A)) = R(A)$. Thus, $dimR(A)$ is the number of pivots in $cref(A)$. A more involved argument shows that, in $rref(A)$ looking at the null space, non-pivot columns correspond to "free variables" while pivot columns correspond to"pivot variables" that are determined by the free variables. In particular, one expects the dimension of the $N(rref(A)) = N(A)$ to be the number of "free variables".

4. Examples

I'll illustrate some of the preceding with an example. Let

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

column operation steps

$$\begin{pmatrix} 2 & 3 & 1 \\ 0 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 1/2 & 1/2 & 1/2 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1/2 & -1/2 & 0 \end{pmatrix}$$

so A has rank 2 and the range is spanned by the vectors

$$\begin{bmatrix} 1 \\ 0 \\ 1/2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1/2 \end{bmatrix}$$

Notice that $A' = cref(A)$ has nullity 1 and null space spanned by $\mathbf{e}_2$ which is NOT in the null space of A. Row operations give:

$$\begin{pmatrix} 2 & 3 & 1 \\ 0 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 3/2 & 1/2 \\ 0 & 1 & 1 \\ 0 & 1/2 & 1/2 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

It is straightforward to see that for $A' = rref(A)$, $Null(A')$ is spanned by $\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$ and hence so is $Null(A)$. In particular as expected the nullity is 1. I leave it to you to see that the$R(A)$ is not the same as $R(A')$.

5. Column and row operations as matrix multiplications

We claim (we will come back to this later in the course that there are $n \times n$ (non-singular) matrices S_{ij}, M_λ and C (corresponding to elementary row column operations a), b) and c) respectively) so that multiplication of A on the right by the matrix produces A', i.e. $A' = AC$. Where A' is obtained from A by one of the elementary column operations.

Similarly, there are $m \times m$ (non-singular) matrices so that multiplication of A on the left by these matrices yields the elementary row operations.

CHAPTER 2

Fifth Lecture

We present some (selected) proofs of important linear algebra facts.

THEOREM 0.1. *Every vector space $E \subset \mathbb{C}^n$ admits a basis.*

PROOF. This is such a fundamental result that it can be a bit difficult to prove and so we don't do it here. $\square$

LEMMA 0.2. *If A is a $m \times n$ matrix and $m < n$ (i.e. A is short and wide) then there is a non-zero vector $\mathbf{v} \in \mathbb{C}^n$ so $A\mathbf{v} = 0$. (i.e. $N(A)$ is non-trivial).*

PROOF. Let $A' = rref(A)$. We verified last time that $N(A') = N(A)$, so we just need to find a non-zero vector in $N(A')$. Write

$$A' = \begin{bmatrix} \mathbf{a}'_1 | & \cdots & | \mathbf{a}'_n \end{bmatrix}$$

so $\mathbf{a}'_i$ are the columns of A'. Since there is only one pivot (at most) in each column and row) and $m < n$ there must be a column $\mathbf{a}'_{j_0}$ without a pivot. Write

$$\mathbf{a}_{i_0} = \begin{bmatrix} a_{1j_0} \\ \vdots \\ a_{mj_0} \end{bmatrix}$$

Let

$$\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

where

$$v_i = \begin{cases} -a_{ij_0} & \text{if } \mathbf{a}'_i \text{ has a pivot and } i < j_0 \\ 0 & \text{if } \mathbf{a}'_i \text{ has no pivot and } i < j_0 \\ 1 & \text{if } i = j_0 \\ 0 & \text{if } i > j_0 \end{cases}$$

Notice that $\mathbf{v} \neq 0$ and $A'\mathbf{v} = 0$ hence $\mathbf{v} \in N(A)$. $\square$

THEOREM 0.3. *If $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is a basis of $E \subset \mathbb{C}^n$ and $\mathbf{w}_1, \ldots \mathbf{w}_l$ is also a basis for E then $k = l$.*

PROOF. We argue by contradiction. Up to a relabelling we may assume that $k > l$. The fact that the $\mathbf{w}_i$ are a basis means there are c_{ij} so that $\mathbf{v}_j = \sum_{i=1}^k c_{ij}\mathbf{w}_i$. Now let V be the $n \times k$ matrix whose columns are $\mathbf{v}_j$ and W be the $n \times l$ matrix whose columns are the $\mathbf{w}_i$. Let C be the $l \times k$ matrix with entries c_{ij}. Then we have

$$V = WC$$

Now C is a $l \times k$ matrix and $l < k$ so by Lemma 0.2 there is a non-trival $\mathbf{x}$ so that $C\mathbf{x} = 0$ but then $V\mathbf{x} = 0$ but this implies the columns of V (i.e. the $\mathbf{v}_j$s) are linearly dependent. A contradiction. □

Lemma 0.4. *Let $\mathbf{v}_1, \ldots, \mathbf{v}_l$ be linearly independent vectors in E. Then $dim(E) \geq l$.*

Proof. This is very similar to the previous argument. We argue by contradiction. Suppose that $dim(E) = k < l$. By Theorem 0.1, there is a basis of E and by Theorem 0.3 we may write it as $\mathbf{w}_1, \ldots, \mathbf{w}_k$. As before writing the $\mathbf{v}_j$ in terms of the $\mathbf{w}_i$ yields a contradiction to Lemma 0.2 Hence $dim(E) = k \geq l$. □

Theorem 0.5. *Any set of linearly independent vectors $\mathbf{v}_1, \ldots, \mathbf{v}_l$ in a vector space E can be extended to a basis $\mathbf{v}_1, \ldots, \mathbf{v}_k$ where $k = dim(E) \geq l$.*

Proof. Set $m = k - l$ i.e. the difference between number of linear inependent vectors $\mathbf{v}$ we start with and the number we want to end up with. By Lemma 0.4 $m \geq 0$. By Theorems 0.1 and 0.3 and Lemma 0.4 when $m = 0$ then the vectors we started with form a basis. To really see this just need to verify that $\mathbf{v}_1, \ldots, \mathbf{v}_l$ generate E. To that end pick any $\mathbf{w} \in E$. If $\mathbf{w} \notin Span\{\mathbf{v}_1, \ldots, \mathbf{v}_l\}$ then $\{\mathbf{w}, \ldots, \mathbf{v}_1, \ldots, \mathbf{v}_l\}$ is linearly independent, then Lemma 4 implies $\dim E \geq l + 1$ contradiction $m = 0$.

We now prove the theorem by induction on m. That is fix $m \geq 0$: suppose that we know that for any vector space E' and lin indep vectors $\mathbf{a}_1, \ldots, \mathbf{a}_{l'}$ in F with $dim(E') = k'$ satisfies $k' - l' = m$ then $\mathbf{a}_1, \ldots, \mathbf{a}_{l'}$ can be extended to a basis of F $\mathbf{a}_1, \ldots, \mathbf{a}_{k'}$. (i.e. we added m new vectors) We want to conclude that for any other vector space E'' with lin indep vectors $\mathbf{b}_1, \ldots, \mathbf{b}_{l''}$ so that $dim(E'') = k''$ satisfies $k'' - l'' = m + 1$ then $\mathbf{b}_1 \ldots, \mathbf{b}_{l''}$ extends to a basis $\mathbf{b}_1, \ldots, \mathbf{b}_{k''}$ (i.e. we added $m+1$ new vectors). We see this as follows: By Theorem **??** the $\mathbf{b}_1, \ldots, \mathbf{b}_{l'}$ cannot span E'' as otherwise they would form a basis but $m + 1 > 0$. Hence pick any vector $\mathbf{w}$ in E'' not in the span of the $\mathbf{b}_1, \ldots, \mathbf{b}_{l'}$. Now if set $\mathbf{b}_{l''+1} = \mathbf{w}$ then the new set of vectors $\mathbf{b}_1, \ldots, \mathbf{b}_{l''}, \mathbf{b}_{l''+1}$ is a) Linearly independent in F'' and hence together with F'' satisfies the induction hypotheses.

This proves the theorem. □

Remark 0.6. If you are having problems with this think about the case $m = 1$ and $m = 2$.

Some applications to matrices.

Theorem 0.7. *Let A be a $m \times m$ matrix . Then A is full rank if and only if the columns of A form a basis of $\mathbb{C}^m$*

Proof. $\Rightarrow$. We need to check that the columns are a basis that is the generate and are linearly independent. We know that A has full rank implies the dimension of $R(A)$ is n which implies by Theorems 0.1 and 0.3 that $R(A)$ has basis $\mathbf{w}_1, \ldots, \mathbf{w}_m \in \mathbb{C}^m$. As the $\mathbf{w}_i$s are linearly indpendet, by Theorem 5 we can extend $\mathbf{w}_1, \ldots, \mathbf{w}_m$ to a basis of $\mathbb{C}^m$ but by Theorem 0.3 this extension must be trival hence $R(A) = \mathbb{C}^m$ and since the columns spance $R(A)$ they span $\mathbb{C}^n$. We check that the columns are linearly indendent as follows: if they failed to be linearly independent then one could remove one of the columns and have $m-1$ vectors spanning $\mathbb{C}^n$. In particular, there would be a $m \times (m-1)$ matrix A' with $R(A') = \mathbb{C}^m$. In particular, if $\mathbf{e}_i$ is the standard basis of $\mathbb{C}^m$ we can find $\mathbf{c}_i \in \mathbb{C}^{m-1}$ so that $\mathbf{e}_i = A'\mathbf{c}_i$. Let C be the

$m-1 \times m$ matrix with columns $\mathbf{c}_i$ then $Id = A'C$, but C is short and wide and so by Lemma 2 there is a non-trival $\mathbf{x} \in \mathbb{C}^m$ so that $C\mathbf{x} = 0$. Since $Id\mathbf{x} = \mathbf{x} \neq 0$ this is a contradiction.

$\Leftarrow$. Since the columns form a basis their span has dimension m. Hence $dim R(A) = n$ and hence A has full rank. □

COROLLARY 0.8. *If A is an $m \times m$ matrix with linearly independent columns then A has full rank.*

PROOF. As the columns are linearly independent and there are m of them, they must span $\mathbb{C}^m$ hence they are a basis. □

CHAPTER 3

Sixth Lecture

In this lecture we discussed how to interpret matrices as systems of linear equations. We also discussed non-singular matrices.

1. Systems of Linear Equations

We disucss one of the classic applications of linear algebra. Namely solving systems of linear equations. We've already seen these sorts of questions in other guises.

The basic set up is let $A \in \mathbb{C}^{m\times n}$ be a $m \times n$ matrix. We can interpret this matrix as a system of m linear equations in n unknowns by letting $\mathbf{x} \in \mathbb{C}^n$ be a vector of variables and $\mathbf{b} \in \mathbb{C}^m$ be fixed and try and look for solutions to:

$$A\mathbf{x} = \mathbf{b}$$

Two important and natural questions immediately arise. Is there always a solution? And if so is it unique? In general both anwers are false so it is a good idea to be able to quantify them. Of course one would also like to find a solution if it exists, but that is a more computational question.

It follows from the definitions pretty much directly that our system has a solution when and only when $\mathbf{b} \in R(A)$. On the other hand if $\mathbf{x} = \mathbf{v}$ is a solution and $\mathbf{v}' \in N(A)$ then it is also clear $\mathbf{x} = \mathbf{v} + \mathbf{v}'$ is also a solution. Similarly, if $\mathbf{x} = \mathbf{v}$ and $\mathbf{x}\mathbf{w}$ are two solutions then $\mathbf{v} - \mathbf{w} \in N(A)$. In other words the null space precisely characterizes how solutions fail to be unique while the range characterizes which inputs lead to solutions.

We point out that thinking about the A as a linear transformation leads to some important (and equivalent) notions. Namely, that of *surjective* and *injective* maps. Consider the linear map $\hat{A}$ associated to A from $\mathbb{C}^n$ to $\mathbb{C}^m$ given by

$$\begin{aligned}\hat{A} : \mathbb{C}^n &\to \mathbb{C}^m \\ \mathbf{x} &\mapsto A\mathbf{x}\end{aligned}$$

We say that $\hat{A}$ is surjective (or onto) precisely when for all $\mathbf{b} \in \mathbb{C}^m$ there is an $\mathbf{a} \in \mathbb{C}^n$ so $\hat{A}(\mathbf{a}) = \mathbf{b}$ that is one can always solve $A\mathbf{x} = \mathbf{b}$ for any $\mathbf{b}$. We say $\hat{A}$ is injective (or 1-1) when $\hat{A}(\mathbf{x}) = \hat{A}(\mathbf{y})$ implies that $\mathbf{x} = \mathbf{y}$ that is there is at *most* one solution to $A\mathbf{x} = \mathbf{b}$ (there may be none).

We note the following equivalent properties for $A \in \mathbb{C}^{m\times n}$: A is surjective as a linear map, $A\mathbf{x} = \mathbf{b}$ always has a solution, $R(A) = \mathbb{C}^m$, the columns of A span $\mathbb{C}^m$. Similiarly: A is injective: $A\mathbf{x} = \mathbf{b}$ has at most one solution, $N(A) = \{0\}$, the columns of A are linearly independent.

2. Non-Singular Matrices

We won't at present discuss further the mechanics of solving systems. The standard approach to this is Gaussian elimination a topic that should have been covered in great depth in Math 51. Rather we will specialize to a very special case. Namely when A is surjective *and* injective, i.e. $R(A) = \mathbb{C}^n$ and $N(A) = \mathbb{C}^m$.

One easy to see fact about A in this case is that A is $m \times m$ i.e. square. This follows as each column is an m vector and there are n of them in $m \times n$ matrix. Injective implies the columns of A are linearly-independent while surjective implies they span $\mathbb{C}^n$. That is they form a basis of $\mathbb{C}^m$ and so $m = n$. We call such a matrix A *Non-Singular.* A $m \times m$ matrix that is not non-singular is called singular. Notice that by Theorem 0.7 and Corollary 0.8 any $m \times m$ matrix is non-singular provided that either columns linearly-independent *or* span

Let A be a $m \times m$ non-singular matrix. One of the most imporant facts about such A is that there exists another matrix which we denote by A^{-1} so that $AA^{-1} = I$. We see this as follows: Let $\mathbf{a}_i$ be the unique solution to $A\mathbf{x} = \mathbf{e}_i$. Then one has:

$$A^{-1} = \begin{bmatrix}\mathbf{a}_1 & | \dots | & \mathbf{a}_m\end{bmatrix}$$

Then we check that $AA^{-1} = I$. A^{-1} is the inverse and so also say that A is invertible

We list some useful facts and indicate an rough idea of the proofs.

- A^{-1} is non-singular. To see this note that $N(A^{-1}) = \{0\}$. Indeed, if $\mathbf{x} \in N(A^{-1})$ then $\mathbf{x} = AA^{-1}\mathbf{x} = A0 = 0$. Hence, A^{-1} is non-singular.
- $A^{-1}A = I$. To see this, we note that $A^{-1} = A^{-1}I = A^{-1}AA^{-1}$. Hence, $I = A^{-1}(A^{-1})^{-1} = A^{-1}AA^{-1}(A^{-1})^{-1} = A^{-1}A$. Here $(A^{-1})^{-1}$ exists as A^{-1} is non-singular.
- If $BA = Id$ or $AB = Id$ then A is non-singular and $B = A^{-1}$. To see this multiply (on the right or left) by A^{-1}.
- $(A^{-1})^{-1} = A$.
- If A, B non-singular then so is AB and $(AB)^{-1} = B^{-1}A^{-1}$. Conversely, if AB is non-singular then both A and B are.

We are going to see non-singular matrices again and again. This is because and this is really important *the columns of a non-singular matrix form a basis* and *a basis gives a non-singular matrix by taking them as columns*! More precisely, given a basis $\mathbf{a}_i$ of $\mathbb{C}^m$ we can write some vector

$$\mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} = \sum_{i=1}^m b_i \mathbf{e}_i = I\mathbf{b}$$

as

$$\mathbf{b} = \sum_{i=1}^m c_i \mathbf{a}_i = A\mathbf{c}$$

Here A is $m \times m$ matrix with colums $\mathbf{a}_i$. The c_i are the coefficients of $\mathbf{b}$ in the basis $\mathbf{a}_i$. We say that the c_is (equvialently the $\mathbf{c}$) are the *coefficients of* $\mathbf{b}$ *with respect to the basis* $\{a_i\}$

The point is the matrix A tells us how to determine the coefficients of $\mathbf{b}$ with respect to the *standard basis* in terms of the coefficients $\mathbf{c}$. On the other hand, multiplying by A^{-1} we have $A^{-1}\mathbf{b} = \mathbf{c}$ in other words A^{-1} tells us how to write the c_i in terms of the b_i. In other words how to write the coefficients of $\mathbf{b}$ in terms of the basis $\mathbf{a}$ in terms of the coefficients of the standard basis. For this reason non-singular matrices are sometimes said to give a *change of basis*.

Why is changing basis good? Well some problems are easier to understand in a given basis. More importantly a matrix X often has a natural basis (usually different form the standard basis) on which it behaves particularly simple. For example, let us write

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

We claim that the three vectors form a basis of $\mathbb{C}^3$. Now let us try and write

$$\mathbf{w} = \begin{bmatrix} 2, 3, 1 \end{bmatrix}$$

in terms of this basis. That is find c_1, c_2, c_3 so that $\mathbf{w} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$ This amounts to looking at

$$\begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$$

Then we have

$$\begin{bmatrix} 1 & 0 & -1 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$$

computing out

$$\begin{bmatrix} 1 & 0 & 1 \\ -1/2 & 1/2 & -1/2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$$

3. How to tell if A is non-singular

There are a number of ways to recognize that an $m \times m$ matrix A is non-singular we won't discuss all of them. When $A \in \mathbb{C}^{m \times m}$ then A is nonsingular is equivalent to...

- $N(A) = \{0\}$ equivalently the columns of A are linearly independent equivalently for all $\mathbf{b}$, $A\mathbf{x} = \mathbf{b}$ has at most one solution.
- $R(A) = \mathbb{C}^m$ equivalently the columns span $\mathbb{C}^m$ equivalently for all $\mathbf{b}$ $A\mathbf{x} = \mathbf{b}$ has a solution.
- There is a $m \times m$ matrix B so $AB = I$ or $BA = I$
- $\det(A) \neq 0$

The last condition is that the determinant of the matrix. We won't discuss this much further as it is a concept that while useful theoretically, almost never gets used in the algorithms we will study. As I consequence, I'll defer defining the determinant till we use it (if ever).

Lets see another example. Lets say I tell you that $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ is a basis of $\mathbb{C}^3$. Then I define $\mathbf{w}_1 = \mathbf{v}_1 - 2\mathbf{v}_2 + \mathbf{v}_3$, $\mathbf{w}_2 = \mathbf{v}_2 - \mathbf{v}_3$ and $\mathbf{w}_3 = \mathbf{v}_1 + \mathbf{v}_3$. How can we determine if $\mathbf{w}_i$s are a basis of $\mathbb{C}^3$? To answer this lets rewrite this as a matrix problem.

$$\begin{bmatrix}\mathbf{w}_1 & |\mathbf{w}_2| & \mathbf{w}_3\end{bmatrix} = \begin{bmatrix}\mathbf{v}_1 & |\mathbf{v}_2| & \mathbf{v}_3\end{bmatrix}\begin{bmatrix}1 & 0 & 1\\ -2 & 1 & 0\\ 1 & -1 & 1\end{bmatrix}$$

I.e. for each column multiplying by V replaces standard basis by basis $\mathbf{v}_i$. For $\mathbf{w}_i$ to be a basis W needs to be non-singular. V is non-singular so it is enougth that the right hand side is non-singular. Row reducing shows that it is.

Now how do we write $\mathbf{v}_i$'s in terms of the $\mathbf{w}_i$'s?

$$\begin{bmatrix}\mathbf{w}_1 & |\mathbf{w}_2| & \mathbf{w}_3\end{bmatrix}\begin{bmatrix}1 & 0 & 1\\ -2 & 1 & 0\\ 1 & -1 & 1\end{bmatrix}^{-1} = \begin{bmatrix}\mathbf{v}_1 & |\mathbf{v}_2| & \mathbf{v}_3\end{bmatrix}$$