CHAPTER ONE
MATRICES

SECTION 1: Introduction

Definition: An mxn matrix A is an array (a_{ij}) of real (or complex) numbers, indexed by natural numbers i and j, with $1 \le i \le m$ and $1 \le j \le n$, written like this:

$$\begin{pmatrix} a_{11} & a_{12}\dots & a_{1n} \\ a_{21} & a_{22}\dots & a_{2n} \\ & & \vdots \\ a_{m1} & a_{m2}\dots & a_{mn} \end{pmatrix}$$

mxn is called the order of the matrix. A column vector is simply an mx1 matrix for some m, whereas a row vector is a 1xn matrix for some n. m is called the height of the column vector and n the width of the row vector. We'll write $\mathbb{R}^{mxn}$ for the set of all mxn matrices; we'll often identify $\mathbb{R}^{mx1}$ with $\mathbb{R}^m$ and $\mathbb{R}^{1xn}$ with $\mathbb{R}^n$.

a matrix whose entries are all zero is called a zero or null matrix, denoted 0. Two matrices are equal if they have the same order and corresponding elements are equal.

Matrix addition/subtraction

Only matrices of the same order can be added/subtracted and this is done by adding/subtracting corresponding entries. The result is a matrix of the same order as the original ones.

Matrix Multiplication.

1. If A is an mxn matrix and B is an nxp matrix, then their product AB is athe mxp matrix C with entries

   $C_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj} \quad 1 \le i \le m$ and $1 \le j \le p$.

   If A ismxn and B ispxq where n≠p, we leave AB undefined.
2. If $\lambda \in \mathbb{R}$ and A is a matrix with entries a_{ij}, then λA is the matrix with entries λa_{ij}.

Fact 1. The following hold for all matrices A, B, and C (so long as the sizes make sense), and all $\lambda \in \mathbb{R}$.

1. (AB)C = A(BC). (Associativity.)
2. A(B + C) = AB + AC. (Distributive property.)
3. $\lambda(AB) = (\lambda A)B = A(\lambda B)$. (Commutativvity of scalar multiplication.)
4. $\lambda(A + B) = \lambda A + \lambda B$. (Distributivity of scalar multiplication.)
5. if A is mxn, then $AI_n = A$ and $I_m A = A$. (Identity)

proof:

1. Clearly, (AB)C and A(BC) have the same order, mxq. Now any i≤m and j≤q, we have

$$((AB)C)ij = \sum_{k=1}^{p} (AB)_{ik} C_{kj}$$
$$= \sum_{k=1}^{p} ¿¿$$
$$= \sum_{k} \sum_{l} A_{il} B_{lk} C_{kj}$$
$$= \sum_{l} \sum_{k} A_{il} B_{lk} C_{kj}$$

$$= \sum_{l=1}^{n} A_{il} ¿¿$$
$$= \sum_{l=1}^{n} A_{il} ¿¿$$
$$= (A(BC))_{ij} \qquad \boxdot$$

5. Notice that AI_n has size mxn. Now if $i \le m$ and $j \le n$,

$$(AI)_{ij} = \sum_{k=1}^{n} A_{ik} I_{kj} = A_{ij} \quad \text{for all i and j}$$

since the only term in this sum which is nonzero is the term where k = j, and that term is A_{ij}.

The nxn identity matrix I_n or simply I when there is no possibility of confusion (and sometimes even when there is) is the matrix (a_{ij}) where $I_{kj} = 1$ if k = j and $I_{kj} = 0$ if $k \neq j$. that is matrices of the form: $\begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix}$ e.g

$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, and $I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Fact 2. For all n, there are nxn matrices A and B such that $AB \neq BA$.

Definition: A matrix is in echelon form if the leading term of each row (except the first) is strictly to the right of all the leading terms of the rows above it, and all of the rows without a leading term are below the ones that are. Visually; for

$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & 5 \end{pmatrix} \qquad B = \begin{pmatrix} 2 & 7 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \qquad C = \begin{pmatrix} 13 & 2 & 1 \\ 0 & 3 & 9 \\ 0 & -1 & 0 \end{pmatrix}$

A and B are in echelon form while C is not.

Row operations. There are three different operations that we can use:

(1) Swap: swapping two rows, written $R_i \leftrightarrow R_j$ (where $i \neq j$);

(2) Scaling: multiplying some row by a nonzero number, written $R_i \rightarrow \lambda R_i$, where $\lambda \neq 0$;

(3) Row combination: replacing some row by the sum of itself with a multiple of another row. $R_i \rightarrow R_i + \lambda R_j$(where $i \neq j$)

Row operations can be performed on a matrix. Each row operation can be implemented by multiplication by a certain matrix. For instance, if E is the following matrix, where the zeros along the diagonal occur in rows i and j (and all blank entries are zero), then EA is the result of swapping rows i and j in A.

An elementary matrix can be obtained by performing one row operation on an identity matrix.

We also can perform reverse row operations on elementary matrices to obtain the identity matrix. Below are row operations with their reverses

(1) To reverse a Swap, we apply the same swap again; $R_i \leftrightarrow R_j$ (where $i \neq j$);

(2) To reverse a Scaling operation $R_i \rightarrow \lambda R_i$, we apply $R_i \rightarrow \frac{1}{\lambda} R_i$, where $\lambda \neq 0$;

(3) To reverse a Row combination $R_i \rightarrow R_i + \lambda R_j$(where $i \neq j$), we apply the row combination $R_i \rightarrow R_i - \lambda R_j$.

There is a systematic way of applying these operations to get a matrix in echelon form.

Now, which matrix implements which row operation?

These matrices are called elementary matrices. These are square matrices. In each case, to perform a row operation on an mxn matrix A, we multiply A on the left by its corresponding mxm elementary matrix E, to get EA. Multiplication by E on the right(if possible) would perform a column operation on A.

Fact 3. If E is an mxm elementary matrix, then there is an mxm elementary matrix F such that FE = I.
Proof: We have seen that every row operation is reversible. So let F be the elementary matrix which implements the reverse of the row operation that E implements. Then we have F(EA) = A for all mxn matrices A, for all n. by associativity, this means (FE)A = A for all mxn matrices A, for all n. Thus FE = I

Definition 5. A matrix A is in reduced row echelon form (rref) if

(1) Every zero row in A is below every nonzero row.
(2) The leading term in any nonzero row in A isstrictly to the right of all the leading terms in the rows above it.
(3) The leading term in a nonzero row in A is a 1, and is the only nonzero entry in its column.

Definition 6. Let A be an mxn matrix.

(1) An nxm matrix B such that BA = I_n is called a left inverse for A.
(2) An nxm matrix C such that AC = I_m is called a right inverse for A.
(3) An nxm matrix D which is both a left inverse and a right inverse for A is called an inverse for A.

If A has a left inverse or right inverse, A is called left-invertible or right invertible respectively. A is called invertible if it is both left-invertible and right invertible. Note here that a matrix may be invertible but have no inverse. **Fact 5.**Suppose A is invertible, and B and C are left inverse and right inverse for A, respectively. Then B = C, and hence A has an inverse.
Proof.
B = BI = B(AC) = (BA)C = IC = C

Corollary 1. Suppose A is invertible. Then there is exactly one inverse for A.

Proof. If B and C are inverses for A, then in particular, B is a left inverse and C a right inverse. Then B = C by the above fact.

Definition 7. If A is an invertible matrix, then A^{-1} is the unique inverse of A.

Corollary 2. If A_1, A_2, …,A_k are square invertible matrices of the same size, then $A_1A_2…A_k$ is also invertible and $(A_1A_2…A_k)^{-1} = A_k^{-1}…A_2^{-1}A_1^{-1}$.

Proof. This can be done by induction on k.

Lemma 4. If and B are square, invertible matrices of the same size, then AB is also invertible. Moreover, $(AB)^{-1} = B^{-1}A^{-1}$.

Proof. We have $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$
$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I$

Lemma 5. If A is an invertible matrix, then so is A-1, and $(A^{-1})^{-1} = A$.

Proof. $AA^{-1} = I$ and $A^{-1} A = I$ means A is both left and right inverse for A^{-1}.

Lemma 6. If E is an elementary matrix, then E is invertible and E^{-1} is the elementary matrix which implements the row operation that reverses that which E implements.

Proof. Suppose E is an mxm. Let f be the mxm elementary matrix which implements the reverse of the row operation E implements. . Then (EF)A = A and (FE)A = A for all mxn matrices A, for all n. in particular, with $A = I_m$, we get $EF = (EF)I_m = I_m$ $FE = (FE)I_m = I_m$

Theorem 4. Let A be a square matrix. Then A is invertible if and only if A is row-equivalent to I.

Proof.Let R be the reduced row echelon form of A. Since we can obtain R from A by row operations, there are elementary matrices $E_1, \ldots, E_k$ such that
$R = E_kE_{k-1}\ldots E_1A$
As we've seen before, elementary matrices are invertible. Hence by the above, R is invertible and square. Then R is actually I.
If A is row equivalent to I, then there are elementary matrices $E_1, \ldots, E_k$ such that $E_kE_{k-1}\ldots E_1A = I$
Then $A = IA = (E_kE_{k-1}\ldots E_1)^{-1}(E_kE_{k-1}\ldots E_1)A = (E_kE_{k-1}\ldots E_1)^{-1}I = (E_kE_{k-1}\ldots E_1)^{-1}$ ⊡
This theorem actually gives us a way of computing the inverse of a square matrix as follows:
Fact 6. Let A be an invertible matrix, and let $E_1, \ldots, E_k$ be elementary matrices reducing A to I, ie, such that $E_kE_{k-1}\ldots E_1A = I$
Then $A^{-1} = E_kE_{k-1}\ldots E_1I$. That is to find the inverse of A, we start with I and apply the same row operations we used to reduce A to I.
Typically, we perform the row operations on A and I in parallel to form the augmented matrix and then reduce it keeping in mind the side for A
Theorem 18. Let A be a square matrix.

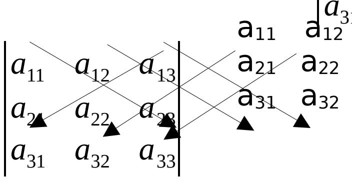
- **a.** Then the following are equivalent.
  (1) A is invertible.
  (2) $Det(A) \neq 0$.A
- **b.** If A is an invertible matrix, then A-1 is invertible and $(A^{-1})^{-1} = A$.
- **c.** If B is an nxn invertible matrix, then so is AB and the inverse of the product is the product of their inverses. That is $(AB)^{-1} = B^{-1}A^{-1}$.
- **d.** If A is invertible, then so is AT, and the inverse of AT is the transpose of A-1. That is $(A^T)^{-1} = (A^{-1})^T$.

Definition. The determinant is a special number associated with any square matrix.

For example: $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ then det A = ad – bc

For a 3x3 matrix, we write the 1st and 2nd columns again beside the determinant, We then compose products as shown by the arrows in the figure below.

That is, to find find $A = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$



To have $|A| = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$
The determinant of A is denoted as |A| or detA and is given by

$|A| = \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det A_{1j}$ be an nxn matrix. …………*

Eg compute the determinant of $A = \begin{pmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{pmatrix}$

Definition: if A is a square matrix, then the **Minor** of entry a_{ij} (called the ijth minor of A) is denoted by Mij and is defined to be the determinant of the submatrix that remains when the ith row and jth column of A are deleted.

Cofactor of an element a_{ij} is $c_{ij} = (-1)^{i+j}M_{ij}$.
When the + or – sign is attached to the Minor, then Minor becomes a cofactor. The formula * above is called the cofactor expansion across row 1 of A.
Theorem: The determinant of an nxn matrix A can be computed by a cofactor expansion across any row or down any column.
The cofactor matrix is the matrix of all the cofactors of the elements of a matrix
That is if $A = (a_{ij})$ nxn , then
cof A = cof $(a_{ij}) = (c_{ij})$ where $c_{ij} = (-1)^{i+j}M_{ij}$ $i = 1,2, \ldots n$ and $j = 1, 2,\ldots,n$.

A triangular matrix is a special kind of m x n matrix where the entries either bellow or above the main diagonal are zero. Eg

$A = \begin{pmatrix} 1 & 4 & 2 \\ 0 & 3 & 4 \\ 0 & 0 & 1 \end{pmatrix}$ is upper triangular and $B = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 8 & 0 \\ 4 & 9 & 7 \end{pmatrix}$ is lower triangular.

The determinant of triangular matrices are also easy to evaluate regardless of size.
Theorem: If A is a triangular matrix, then det(A) is the product of the entries on the main diagonal.
Theorem (Row Operations): Let A be a square matrix.

a. If a multiple of one row of A is added to another row, the resulting determinant will remain same.
b. If two rows of A are interchanged to produce B, then det B = -det A.
c. If one row of A is multiplied by k to produce B, then det B = k.det A.

e.g without expansion, show that $\begin{vmatrix} x & a+x & b+c \\ x & b+x & c+a \\ x & c+x & a+b \end{vmatrix} = 0$

Theorems:

1. If A is an n x n matrix, then det A^T = det A.
2. If A and B are n x n matrices, then det(AB) = (det A)(det B).

SUMMARY

1. The inverse of a matrix can be computed by
   a. Forming an augmented matrix with the identity matrix of same order and then by row operations, reduce the matrix to and identity matrix
   b. Computing the cofactor matrix and getting its transpose
2. The determinant of a matrix can be computed by
   a. Looking for an equivalent triangular matrix using row operations and taking note of the sign changes and scalars and applying them later
   b. Using expansion about a row or column of your choice (The one with many zeros)

SECTION 2: SYSTEMS OF LINEAR EQUATIONS

Consider the following Chemistry problem: We've got x molecules of toluene, C_7H_8 and y molecules of nitric oxide, HNO_3. Putting them together, we can produce trinitrotoluene (TNT), which has the chemistry form $C_7H_5O_6N_3$, with some water (H2O) by product. Say we produce z molecules of the first and w of the second. We'd like to balance this equation:
$xC_7H_8 + yHNO_3 \rightarrow zC_7H_5O_6N_3 + wH_2O$

That is,we will like to find values of x, y, z, and w so that the number of atoms of each type is the same before and after the reaction. Counting them each individually, this leads to the following system of equations:

(1) $7x = 7z$
(2) $8x + 1y = 5z + 2w$
(3) $1y = 3z$
(4) $3y = 6z + 1w$

Of course, we are interested in nonnegative integer values of x, y, z, w, so we will have to watch out for that when we are searching for solutions. Equation (1) clearly implies $x = z$, and so subtracting equation (3) from equation (2), we get

$$8z = 2z + 2w \Longrightarrow 6z = 2w \Longrightarrow w = 3z$$

So if (x, y, z, w) is any solution, then $x = z$, $y = 3z$ and $w = 3z$. Moreover, it is easy to check that no matter what choice we make for the value of z, letting $x = z$ and $y, w = 3z$, we get a solution to the above system, and if z is a nonnegative integer then so are x, y, and w. So we have a solution and this kind of solution is called an infinitely many solutions.

We are going to see some methods in this course that will tell us about the kind of solution that we have.

Definition . Let $x_1, \ldots, x_n$ be variables. A linear combination of $x_1, \ldots, x_n$ is an expression of the form $a_1x_1 + a_2x_2 + \ldots + a_nx_n$ where $a_1, \ldots, a_n \in \mathbb{R}$. $a_1, \ldots, a_n$ are called the coefficients of the combination. A linear equation in variables $x_1, \ldots, x_n$ is an equation of the form $a_1x_1 + a_2x_2 + \ldots + a_nx_n = b$, where $b \in \mathbb{R}$. b is called the constant of the equation.

A system of linear equations is simply a finite set of linear equations in the same variables e.g

$$\begin{array}{c} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{array}$$

A tuple $(s_1, \ldots, s_n) \in \mathbb{R}^n$ is a solution to this system of equations if the equations obtained by replacing xi by s_i for each i, in each equation, results in unanimously true statements. If S is a system of linear equations then we write sol(S) for its set of solutions. So sol(S) $\subseteq \mathbb{R}^n$.

When a system of equations is written like the above, we often refer to the ith equation down as row i,or ρ_i. Generally, the order of the equations does not matter when considering solutions; it only matters to our written system. The first term $a_{ij}x_j$ in row I such that $a_{ij} \neq 0$ is called the leading term of that row, and a_{ij} the leading coefficient. Note that a row may not have a leading term, eg, if all of the coefficients are zero; $0x_1 + 0x_2 + \ldots + 0x_n = b$

A system is in echelon form if the leading term of each row (except the first) is strictly to the right of all the leading terms of the rows above it, and all of the rows without a leading term are below the ones that are. For instance, the following is in echelon form:

$$\begin{aligned} x - 2y + z &= 0 \\ z &= 2 \end{aligned}$$

while

$$y + 3z = 1$$
$$-x - y - z = 0$$
$$5y - 2z = -1$$

Is not in echelon form.

Gaussian Elimination: the object of Gaussian elimination is to produce, through a series of operations on a given system of linear equations, a final system which is in echelon form. There row operations given above are used.

Theorem 1. If S is a system of liner equations and T is the result of applying one of the above row operations to S, then S and T have exactly the same set of solutions, ie, sol(S) = sol(T).

Proof: Let S be the following system of linear equations;

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \end{aligned}$$

we want to show that

(i) if $r \in \mathbb{R}^n$ is a solution to S then r is also a solution to T and

(ii) there is a row-operation which when applied to T, produces S.

(i) proves that sol(S) $\subseteq$ sol(T); then, (ii) and (i) prove together that sol(T) $\subseteq$ sol(S).

Let's prove (i) first. Let $(s_1, \dots, s_n) \in \mathbb{R}^n$ be a solution to S. This means we have

$$\begin{aligned} a_{11}s_1 + a_{12}s_2 + \dots + a_{1n}s_n &= b_1 \\ a_{21}s_1 + a_{22}s_2 + \dots + a_{2n}s_n &= b_2 \\ &\vdots \\ a_{m1}s_1 + a_{m2}s_2 + \dots + a_{mn}s_n &= b_m \end{aligned}$$

to fully prove (i), we need to handle three cases according to which type of row operation we applied to S to get T; we'll do one for the row combination case. Now for $i \neq j$ and $\lambda \in \mathbb{R}$, T is the system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ &\vdots \\ (a_{i1}x_1 + \cdots + a_{1n}x_n) + \lambda(a_{j1}x_1 + \cdots + a_{jn}x_n) &= b_i + \lambda b_j \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \end{aligned}$$

since $(s_1, \dots, s_n)$ is a solution to S, we have

$a_{i1}s_1 + \cdots + a_{1n}s_n = b_i$ and

$a_{j1}s_1 + \cdots + a_{jn}s_n = b_j$ so

$(a_{i1}s_1 + \cdots + a_{1n}s_n) + \lambda(a_{j1}s_1 + \cdots + a_{jn}s_n) = b_i + \lambda b_j$

The other equations in T are the same as in S. Hence $(s_1, \dots, s_n)$ is a solution to T.

To prove (ii), we again case by case depending on which operation we applied to get T from S keeping in mind the reverse row operations seen above.

.Now, let S be the following system.

$$a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1$$
$$a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2$$
$$\vdots$$
$$a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m$$

perform the following operations for each i from 1 to m in that order, or until you are told to stop.

- Find the row, below row, with the leftmost leading term among those rows. If there is no such thing (ie, if ajk = 0 for all $j \geq i$ and $1 \leq k \leq m$), stop. Otherwise swap this row with row i.
- Say the leading coefficient in row I is in column j. For each k > i, perform the row combination $\rho_k \rightarrow \rho_k - (a_{kj}/a_{ij})\rho_i$.
- Repeat.

After the first i steps of this algorithm, we've ensured that the leading terms of the first I rows go from left to right (this is easily proven by induction on i). If we go through every row, then we've ensured that the leading terms are ordered this way throughout the whole matrix. If we stop at row I, then none of the rows below have leading terms. Either way, after finishing, the system is in echelon form.

Now, we look at the equations without leading terms viz: $0x_1 + 0x_2 + \ldots + 0x_n = b$

This equation is true when b = 0 and false when $b \neq 0$. If $b \neq 0$, we say the system is inconsistent hence, it has no solution. That is sol (S) = $\emptyset$. If b = 0, we have infinitely many solutions.

Now if a system is not inconsistent, we back-substitute. To perform back substitution, we do the following for each i, starting with the largest I such that row i has a leading variable, and going down to 1.

- Solve equation i for the leading variable, in terms of all the other variables in equation i.
- In your formula for the leading variable, for each j> i, replace the leading variable of row j with formula you found for it previously.

After this process, you have a formula for each leading variable in the system, in terms of only free variables. Any choice of values for the free variables then results in a solution to the given system of linear equations.
A system of linear equations can be written as a single equation involving matrices and vectors; namely, the system

$$a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1$$
$$a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2$$
$$\vdots$$
$$a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m$$

can bewritten as Ax = b, where A is an mxn matrix with entries a_{ij}, x is the column vector with entries x_i, and b is the column vector with entries b_i. We also sometimes represent the entire system as a matrix;

$$\begin{pmatrix} a_{11}a_{12}\ldots a_{1n}b_1 \\ a_{21}a_{22}\ldots a_{2n}b_2 \\ .. \\ .. \\ .. \\ a_{m1}a_{m2}\ldots a_{mn}b_m \end{pmatrix}$$

We write (A|b) for this matrix, and call it the augmented matrix of the system.
A system of linear equations Ax = b is in reduced row echelon form if and only if A is.

Definition . A linear equation with a constant term zero, ie of the form
$a_1x_1 + a_2x_2 + \ldots + a_nx_n = 0$
is called homogeneous. A system of linear equations is called homogeneous if each of its equations is homogeneous.
Note that a homogeneous system always has at least one solution, namely the zero vector.

Lemma 1. Let A be an mxn matrix. If A is left-invertible then the only solution to the homogeneous system Ax = 0 is the zero vector.

Proof. Let B be a left-inverse for A. (Why is it incorrect to talk about A^{-1} here?) Then if s is any solution to the homogeneous system Ax = 0, we have

$$S = Is = (BA)s = B(As) = B0 = 0$$

So in fact, 0 is the only solution. ⊡

Lemma 2. Suppose a system Ax = b, where A is m x n, has at least one solution, $p \in \mathbb{R}^n$. let k be the number of free variables in some echelon form of the system Ax = b. then there exist solutions $h_1, \ldots, h_k$ to the homogeneous equation Ax = 0 such that
$\{s\in \mathbb{R}^n \mid As = b\} = \{p + c_1h_1 + \ldots + c_kh_k \mid c_1, \ldots, c_k\in \mathbb{R}\}$
proof. The main idea here is that if $h \in \mathbb{R}^n$ I any solution to the homogeneous system Ax = 0, then

$$A(p + h) = Ap + Ah = Ap + 0 = Ap = b$$

and then so is p + h. Moreover if any $s\in \mathbb{R}^n$ is a solution to Ax = b, then

$$A(s - p) = As - Ap = b - b = 0$$

andso s – p is a solution to the homogeneous system Ax = 0. This tells us that a vector $s\in \mathbb{R}^n$ is a solution to Ax = b if and only if it's of the form p + h, where h i a solution to the homogeneous system Ax = 0. To see that there are fixed vectors $h_1, \ldots, h_k\in \mathbb{R}^n$ which generate the solutions to Ax = 0 will require more work; it essentially falls out of back-substitution. ⊡

Example. The system

$$x+z+w=2$$

$2x-y+w=1$
$x+y+3z+2w=5$

Has a solution $\begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$. what about the homogeneous system? It reduces to $x+z+w=0$

$y+2z+w=0$

$$0 = 0$$

Then the solution set to the homogeneous system is

$$\left\{ z\begin{pmatrix}-1\\-2\\1\\0\end{pmatrix}+w\begin{pmatrix}-1\\-1\\0\\1\end{pmatrix} \middle| z,w\in R \right\}$$

And so the solution set to the original system of equations is

$$\left\{ z\begin{pmatrix}-1\\-2\\1\\0\end{pmatrix}+w\begin{pmatrix}-1\\-1\\0\\1\end{pmatrix} \middle| z,w\in R \right\}$$

$$\left\{ \begin{pmatrix}0\\0\\1\\1\end{pmatrix}+z\begin{pmatrix}-1\\-2\\1\\0\end{pmatrix}+w\begin{pmatrix}-1\\-1\\0\\1\end{pmatrix} \middle| z,w\in R \right\}$$

Lemma 3. Suppose R is an invertible nxn matrix in reduced row echelon form. Then R = I.

Proof. Consider the homogeneous system Rx = 0. If there were any free variables in this system, then there would be a nonzero vector h $\mathbb{R}^n$ such that Rh = 0. But as R is invertible, by the lemma above, this can't happen. So there can't be any free variables in this system. Since R is in reduced row echelon form, this means exactly that R = I_n.