5. Gauss Jordan Elimination

Gauss Jordan elimination is very similar to Gaussian elimination, except that one "keeps going". To apply Gauss Jordan elimination, first apply Gaussian elimination until A is in echelon form. Then pick the pivot furthest to the right (which is the last pivot created). If there is a non-zero entry lying above the pivot (after all, by definition of echelon form there are no non-zero entries below the pivot) then apply elimination to create a zero in that spot, in the usual way. Repeat this until there are no more non-zero entries in that column. Once we are finished with this pivot then move left to the next closest pivot and keep going. An example will hopefully make this clear. Let us start with a matrix in echelon form:

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

The pivot furthest to the right is the $(3,4)$ entry. So we take the third row multiply by 2 and by -3 and add it to the second and first rows, to get

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

So much for the pivot furthest to the right. The next pivot (reading right to left) is at position $(2,3)$. Multiply the second row by -2 and add it to the first row, to get

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

This completes Gauss Jordan elimination.

Definition 5.1. *Let A be an $m \times n$ matrix.*

We say that A is in ***reduced row echelon form*** *if A in echelon form and in addition*

- *every other entry of a column which contains a pivot is zero.*

The end product of Gauss Jordan elimination is a matrix in reduced row echelon form. Note that if one has a matrix in reduced row echelon form, then it is very easy to solve equations. One can read off the solutions with almost no work.

We can exploit this fact to come up with a very pretty way to compute the inverse of a matrix. First a couple of easy (but important):

Lemma 5.2. *Let A be a square matrix.*

If A is invertible then every equation $Ax = b$ has a unique solution.

Proof. Let B be the inverse of A. Suppose that v is a solution of the equation $Ax = b$. Then

$$b = Av.$$

Multiply both sides by B, to get

$$\begin{aligned} Bb &= B(Av) \\ &= (BA)v \\ &= I_n v = v. \end{aligned}$$

Thus $v = Bb$. It follows that there is at most one solution to the equation $Ax = b$. On the other hand let $v = Bb$. Then

$$Av = A(Bb) = (AB)b = I_n b = b,$$

so that $v = Bb$ is a solution to the equation. □

Suppose that we are given a square $n \times n$ invertible matrix A and we want to compute its inverse matrix B. We can view B as being composed of n column vectors, say $v_1, v_2, \dots, v_n \in \mathbb{R}^n$. Suppose that we also view the identity matrix as being composed of columns $e_1, e_2, \dots, e_n$. Then e_i is the column vector whose ith entry is 1 and every other entry is zero. The way that matrix multiplication works is that

$$Av_i = e_i.$$

Since we are trying to determine v_i, we can think of v_i as being the solution to the linear equation

$$Ax = e_i.$$

How can we solve this equation? Well create the augmented matrix

$$(A|e_i).$$

If we apply Gaussian elimination then we get to a matrix U in echelon form

$$(U|f_i).$$

Here f_i is the result of applying the steps of Gaussian elimination to e_i. Now if we continue we get a matrix, call it R, in reduced row echelon form and another column vector g_i,

$$(R|g_i).$$

Now let us make three clever observations. The first is that the matrices U and R do not depend on i. That is to say the steps of both

Gaussian elimination and Gauss Jordan elimination only depend on the coefficient matrix A and not on e_i. The second is that the matrix R must be the identity matrix. Indeed we cannot get a row of zeroes when we apply Gaussian elimination, since we know that every equation has a solution. It follows that every row contains a pivot and so every column contains a pivot (R is a square matrix). Since R is a square matrix in reduced row echelon form it follows that $R = I_n$. But now if we apply back substitution the equations immediately tell us that $v_i = g_i$.

So in principle we could compute the inverse matrix B, column by column, by solving n sets of linear equations. But this sounds like a lot of hard work. The final observation is that for each one of these n sets of linear equations the coefficient matrix is the same, namely A, and the steps of Gauss Jordan are the same. So why not put all n sets of equations into one super augmented matrix? If one puts the n column vectors $e_1, e_2, \dots, e_n$ together side by side, then one will automatically get the identity matrix I_n,

$$(A|I_n).$$

Let us see an example. Suppose that

$$A = \begin{pmatrix} 1 & 2 & -3 \\ 2 & 5 & -8 \\ -3 & -5 & 8 \end{pmatrix}.$$

First we make the super augmented matrix

$$\begin{pmatrix} 1 & 2 & -3 & | & 1 & 0 & 0 \\ 2 & 5 & -8 & | & 0 & 1 & 0 \\ -3 & -5 & 8 & | & 0 & 0 & 1 \end{pmatrix}.$$

Now apply Gaussian elimination. Multiply the first row by -2 and 3 and add them to the second and third row.

$$\begin{pmatrix} 1 & 2 & -3 & | & 1 & 0 & 0 \\ 0 & 1 & -2 & | & -2 & 1 & 0 \\ 0 & 1 & -1 & | & 3 & 0 & 1 \end{pmatrix}.$$

Now multiply the second row by -1 and add it to the third row to get

$$\begin{pmatrix} 1 & 2 & -3 & | & 1 & 0 & 0 \\ 0 & 1 & -2 & | & -2 & 1 & 0 \\ 0 & 0 & 1 & | & 5 & -1 & 1 \end{pmatrix}.$$

This completes the Gaussian elimination. Now we continue Gauss Jordan elimination Multiply the third row by 2 and 3 and add it to the second and first row, to get

$$\begin{pmatrix} 1 & 2 & 0 & | & 16 & -3 & 3 \\ 0 & 1 & 0 & | & 8 & -1 & 2 \\ 0 & 0 & 1 & | & 5 & -1 & 1 \end{pmatrix}.$$

The last step is to multiply the second row by -2 and add it to the first row,

$$\begin{pmatrix} 1 & 0 & 0 & | & 0 & -1 & -1 \\ 0 & 1 & 0 & | & 8 & -1 & 2 \\ 0 & 0 & 1 & | & 5 & -1 & -1 \end{pmatrix}.$$

The inverse matrix is

$$B = \begin{pmatrix} 0 & -1 & -1 \\ 8 & -1 & 2 \\ 5 & -1 & -1 \end{pmatrix}.$$

One can check that indeed $AB = BA = I_3$.

Let us summarise how to solve a system of m equations in n unknowns. First replace the linear equations by a matrix equation

$$Av = b.$$

Then form the augmented matrix $C = (A|b)$. Apply Gaussian elimination to get a matrix $C' = (A'|b')$. If there are any rows of the coefficient matrix A which are a row of zeroes and yet the corresponding row of C' is not zero then there are no solutions. Otherwise the columns which don't contain pivots give the free variables. If we use back substitution then we are free to assign values to those variables and use the equations to determine the other variables.

Definition 5.3. *Let A be a $m \times n$ matrix. The system of equations $Ax = 0$ is called **homogeneous**.*

One particular nice feature about linear equations is that to solve a system of linear equations it suffices to solve the homogeneous equation and add it to any particular solution.

Proposition 5.4. *Let A be a $m \times n$ matrix, let b be a $m \times 1$ vector and let v_0 be a solution to the equation $Ax = b$.*

Then the solutions to the equation $Ax = b$ all have the form $v = v_0 + v'$ where v' is any solution of the homogeneous equation $Ax = 0$.

Proof. Suppose that v' is a solution of the homogeneous. Then

$$A(v_0 + v') = Av_0 + Av' = b + 0 = b.$$

So $v = v_0 + v'$ is a solution to $Ax = b$. Conversely suppose that v is a solution to $Ax = b$. Let $v' = v - v_0$. Then

$$Av' = A(v - v_0) = Av - Av_0 = b - b = 0.$$

Therefore v' is a solution to the equation $Ax = 0$. □

We can put together all that we have learned so far about what happens when we have the same number of equations and variables:

Theorem 5.5. *Let n be a positive integer and let A be a square $n \times n$ matrix. Let $\phi\colon \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be the function $\phi(v) = Av$.*

TFAE

(1) *The equation $Ax = b$ has a unique solution for every $b \in \mathbb{R}^n$.*
(2) *The equation $Ax = b$ has at least one solution for every $b \in \mathbb{R}^n$.*
(3) *A is invertible.*
(4) *ϕ is bijective.*
(5) *ϕ is injective.*
(6) *The equation $Ax = 0$ has at most one solution.*
(7) *There is a vector $b \in \mathbb{R}^n$ such that the equation $Ax = b$ has a unique solution.*
(8) *The rank of A is n.*

Proof. Suppose that (1) holds. Then it is clear that (2) holds. Thus (1) implies (2).

Now suppose that (2) holds. Let e_i be the $m \times 1$ vector whose ith row contains 1 and whose other entries are all zero. Let b_i be any solution to the equation $Ax = e_i$. Let B be the matrix whose ith column is b_i. Then $AB = I_n$. Now, since matrix multiplication is not commutative, we don't actually know right off that $BA = I_n$. We need to appeal to Gaussian elimination in one form or another. The key point is that if we apply Gaussian Jordan elimination then we get the identity matrix.

For example we know that the super augmented matrices $(A|I_n)$ and $(I_n|B)$ are row equivalent. Permuting the columns, it follows that the super augmented matrices $(B|I_n)$ and $(I_n|A)$ are row equivalent. But this says that $BA = I_n$. Hence B is the inverse of A and so A is invertible.

Another way to argue is that we know that there are elemenary matrices $E_1, E_2, \dots, E_k$ such that $A = E_1E_2 \cdots E_kI_n = E_1E_2 \cdots E_k$. But the product of elementary matrices is invertible (and the inverse is the product of the inverse elementary matrices in the opposite order).

Either way, (2) implies (3).

Now suppose that A is invertible. Let B be the inverse of A and let ψ be the function $\psi\colon \mathbb{R}^n \longrightarrow \mathbb{R}^n$ given by $\psi(v) = Bv$. Then

$$(\psi \circ \phi)(v) = \psi(\phi(v)) = B(A(v)) = (BA)v = I_n v = v.$$

Thus $\psi \circ \phi$ is the identity function. Similarly $\phi \circ \psi$ is the identity. It follows that ϕ is bijective. Therefore (3) implies (4).

Clearly (4) implies (5). Now suppose that ϕ is injective. The solutions of the equation $Ax = 0$ is the inverse image of zero. As ϕ is injective there is only at most one element in the inverse image and so the equation $Ax = 0$ has at most one solution. Therefore (5) implies (6).

Now suppose that the equation $Ax = 0$ has at most one solution. As $x = 0$ is a solution then $Ax = 0$ has a unique solution. Thus (6) implies (7).

Now suppose that the equation $Ax = b$ has a unique solution, for some vector $b \in \mathbb{R}^n$. We can solve this equation using Gaussian elimination applied to the augmented matrix $(A|b)$. Suppose we end up with $(U|c)$. But then the equation $Ux = c$ has a unique solution, since row equivalent augmented matrices have the same solution set. U is in echelon form and the rank of A is the number of pivots. Suppose that there are less than n pivots. Then U must contain a row of zeroes. There are two possibilities. The corresponding entry in c is non-zero. But then the equation $Ux = c$ is inconsistent, a contradiction. Otherwise, the corresponding entry is zero and this happens for every such row. But then we could solve the system $Ux = c$ using backwards substitution. Since at least one of the variables is a free variable, there would be infinitely many solutions, a contradiction. So the rank of A is n. Thus (7) implies (8).

Now suppose that (8) holds. Suppose we start with an equation $Ax = b$. Form the augmented matrix $(A|b)$ and apply Gaussian elimination to get $(U|c)$. By assumption the number of pivots is n and so there are no rows of zeroes. But then if we solve the system $Ux = c$ using back substitution then the solution is unique and this is the same as the solution to the linear equation $Ax = b$. $\square$