LECTURE 8

Homogeneous Linear Systems

We now return to some more theoretical aspects linear systems and their corresponding matrices.

We first note that there is a natural 1:1 correspondence between homogeneous $n \times m$ linear systems and $n \times m$ matrices. For any $n \times m$ matrix

$$\mathbf{A} = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} \tag{1}$$

is interpretable as the coefficient matrix of $n \times m$ homogeneous linear system

$$\begin{aligned} \alpha_{11}x_1 + \cdots + a_{1m}x_m &= 0 \\ &\vdots \\ a_{n1}x_1 + \cdots + a_{nm}x_m &= 0 \end{aligned} \tag{2}$$

In this lecture we'll study the solution spaces of $n \times m$ homogeneous linear systems and relate some of their properties to properties of their coefficient matrices.

Although we have yet to define matrix multiplication or inner products, we can simplify our notation immensely if we simply write

$$\mathbf{Ax} = \mathbf{0}$$

for typical homogeneous linear system, and if we wish to reference a particular equation in this system, say

$$a_{i1}x_1 + \cdots + a_{im}x_m = 0 \quad , \tag{3}$$

we may write

$$\mathbf{r}_i \cdot \mathbf{x} = 0 \tag{4}$$

as right hand side of (3) does in fact coincide with the dot product of the variable vector $\mathbf{x} = [x_1, \ldots, x_m]$ with the i^{th} row of the coefficient matrix. As before, we will also use the notation $\mathbf{c}_j$ for the j^{th} column vector of a matrix.

As an initial remark, let me point out that a homogeneous linear system such as (2) always has a solution. For if we simply set each x_i, $1 \leq i \leq m$, equal to $0_{\mathbb{F}}$, then all the equations in (2) will be satisfied. For this reason, instead of asking if solutions exist, we shall be asking if solutions other than the *trivial solution* $\mathbf{x} = \mathbf{0}_{\mathbb{F}^m}$ exist. The following lemma follows easily from Theorem 7.2.

PROPOSITION 8.1. *A homogeneous linear system with coefficient matrix* $\mathbf{A}$ *has non-trivial solutions if and only if the columns of* $\mathbf{A}$ *are linearly dependent.*

Proof. By theorem 7.2 (i), we know that the linear system $\mathbf{Ax} = \mathbf{0}$ will have a solution if and only if the right hand side $\mathbf{0}$ lies in the span of the columns of $\mathbf{A}$. If $\mathbf{0} \in \mathit{span}(\mathbf{c}_1, \ldots, \mathbf{c}_m)$, then there are field elements $\lambda_1, \ldots, \lambda_m$ such that

$$\mathbf{0} = \lambda_1 \mathbf{c}_1 + \cdots + \lambda_m \mathbf{c}_m$$

If the column vectors $\mathbf{c}_1, \ldots, \mathbf{c}_m$ are linear independent, the only way this equation could be satisfied is by taking all the $\lambda_i = 0$. So if we are to have non-trivial solutions (solutions where some of the variables x_i are non-zero), the columns of $\mathbf{A}$ must be linearly dependent. $\square$

THEOREM 8.2. *The solutions of an $n \times m$ homogeneous linear system form a subspace of $\mathbb{F}^m$.*

Proof. It will suffice to show that any linear combination of two solutions of $\mathbf{Ax} = \mathbf{0}$ is another solution. Let $\mathbf{x}, \mathbf{y} \in \mathbb{F}^m$ be two solutions of $\mathbf{Ax} = \mathbf{0}$. This means that, in the notation of (3), that

$$\mathbf{r}_i \cdot \mathbf{x} = 0 \quad , \quad i = 1, \ldots, n \tag{5}$$
$$\mathbf{r}_i \cdot \mathbf{y} = 0 \quad , \quad i = 1, \ldots, n \tag{6}$$

Multiplying each equation in (5) by some $\alpha \in \mathbb{F}$ and each equation in (6) by β, we get

$$\begin{aligned} 0 &= \alpha (\mathbf{r}_i \cdot \mathbf{x}) = \mathbf{r}_i \cdot (\alpha \mathbf{x}) \quad , \quad i = 1, \ldots, n \\ 0 &= \beta (\mathbf{r}_i \cdot \mathbf{y}) = \mathbf{r}_i \cdot (\beta \mathbf{y}) \quad , \quad i = 1, \ldots, n \end{aligned}$$

and then adding these equations pairwise, we get

$$0 = \mathbf{r}_i \cdot (\alpha \mathbf{x} + \beta \mathbf{y}) \quad , \quad i = 1, \ldots, n$$

which tells us that $\alpha \mathbf{x} + \beta \mathbf{y}$ is also a solution of the original homogeneous linear system. □

Consider an $n \times m$ linear system $\mathbf{Ax} = \mathbf{0}$. By Theorem 5.3 we can always choose a subset of the columns $\{\mathbf{c}_1, \ldots, \mathbf{c}_m\}$ that provides a basis for the column space. In fact, by relabeling the variables x_i in the original linear system, we can always arrange matters so that the first, say r, column vectors of $\mathbf{A}$ form a basis for the column space. In this situation, all the columns past the r^{th} column can be expressed as linear combinations of the first r columns. That is to say, for each i between $r + 1$ and m, there will be field elements $\lambda_1^{(i)}, \ldots, \lambda_r^{(i)}$ such that

$$\mathbf{c}_i = \lambda_1^{(i)} \mathbf{c}_1 + \lambda_2^{(i)} \mathbf{c}_2 + \cdots + \lambda_r^{(i)} \mathbf{c}_r \quad , \quad r + 1 \leq i \leq m$$

and thus we will have the following $m - r$ dependence relations

$$\lambda_1^{(i)} \mathbf{c}_1 + \lambda_2^{(i)} \mathbf{c}_2 + \cdots + \lambda_r^{(i)} \mathbf{c}_r - \mathbf{c}_i = \mathbf{0} \quad , \quad r + 1 \leq i \leq m \ . \tag{7}$$

THEOREM 8.3. *Let $\mathbf{Ax} = \mathbf{0}$ be a $n \times m$ homogeneous linear system set up in such a way that the first r column vectors of the coefficient matrix $\mathbf{A}$ form a basis for the column space of $\mathbf{A}$ and with the dependence relations (7). Then the vectors*

$$\mathbf{u}_i = \left[\lambda_1^{(i)}, \lambda_2^{(i)}, \ldots, \lambda_r^{(i)}, 0, \ldots, 0, -1, 0, \ldots, 0\right] \quad , \quad r + 1 < i \leq m$$

with the component -1 occurring in the i^{th} slot of the vector on the right, will provide a basis for the solution space of $\mathbf{Ax} = \mathbf{0}$.

Proof. Examining say, the j^{th} component the i^{th} vector equation in (7) component by component, one observes that it is equivalent to

$$\mathbf{r}_j \cdot \mathbf{u}_i = 0$$

Since such a relation will hold for each of the row vectors $\mathbf{A}$, we conclude that each of the vectors $\mathbf{u}_i$ will be solutions of the original homogeneous linear system. Next we note that the $m - r$ vectors $\mathbf{u}_{r+1}, \ldots, \mathbf{u}_m$ are linearly independent. This is because each of the vectors will have exactly one non-zero component in its last $n - r$ entries (the -1 that occurs in the i^{th} slot of $\mathbf{u}_i$). But then the last $n - r$ entries of a linear combination of the these vectors will look like

$$\alpha_{r+1} \mathbf{u}_{r+1} + \cdots + \alpha_m \mathbf{u}_m = [*, \ldots, *, -\alpha_{r+1}, -\alpha_{r+2}, \ldots, -\alpha_m]$$

So such a linear combination can not sum to the zero vector without setting each coefficient $\alpha_{r+1}, \ldots, \alpha_m$ separately equal to 0. We conclude that the vectors $\mathbf{u}_{r+1}, \ldots, \mathbf{u}_m$ will be linearly independent.

Finally, we shall show that these vector generate the entire solution space. Let $\mathbf{x}$ be another solution of $\mathbf{Ax} = \mathbf{0}$ and suppose $x_{r+1}, \ldots, x_m$ are the last $m - r$ components of x. Consider the linear combination

$$\mathbf{x} + \sum_{ii=r+1}^{m} x_i \mathbf{u}_i$$

Noting that the way we have set things up each of last $m-r$ components of $\mathbf{x}$ will cancels with the corresponding component of one (and only one) of the $x_i\mathbf{u}_i$. So we can write

$$\mathbf{x}+\sum_{ii=r+1}^{m} x_i\mathbf{u}_i=[\xi_1,\ldots,\xi_r,0,\ldots,0]$$

On the other hand, xince $\mathbf{x}+\sum_{ii=r+1}^{m} x_i\mathbf{u}_i$ is a sum of solutions of the homogeneous linear system, it too will be a solution. Therefore,

$$\boldsymbol{\xi}_1\mathbf{c}_1+\xi_2\mathbf{c}_2+\cdots+\xi_r\mathbf{c}_r+0\cdot\mathbf{c}_{r+1}+\cdots+0\cdot\mathbf{c}_m=\mathbf{0}$$

or

$$\boldsymbol{\xi}_1\mathbf{c}_1+\xi_2\mathbf{c}_2+\cdots+\xi_r\mathbf{c}_r=\mathbf{0}\quad .$$

But the first r column vectors of $\mathbf{A}$ are linearly independent - hence, each $\xi_i=0$. And this turn means that

$$\mathbf{x}+\sum_{ii=r+1}^{m} x_i\mathbf{u}_i=[0,,\ldots,0]$$

and so we can express $\mathbf{x}$ as a linear combination of the vectors $\mathbf{u}_i$, $r+1\leq i\leq m$. □

COROLLARY 8.4. *The dimension of the solution space of an $n\times m$ homogeneous linear system is $m-r$ where m is the (column) rank of the corresponding coefficient matrix.*

Proof. The preceding theorem produces a basis $\{\mathbf{u}_{r+1},\ldots,\mathbf{u}_m\}$ for the solution space of $n\times m$ homogeneus linear system where the r is the (column) rank of the coefficient matrix. □

Thus far, we have used the terminology *column rank* of a matrix $\mathbf{A}$ to mean the dimension of the subspace of $\mathbb{F}^n$ generated by the column vectors of $\mathbf{A}$. Similarly, we can consider the *row rank* of a matrix to be the dimension of the subspace of $\mathbb{F}^m$ generated by the row vectors of $\mathbf{A}$.

THEOREM 8.5. *Let $\mathbf{A}$ be an $n\times m$ matrix, then the column rank of $\mathbf{A}$ equals its row rank.*

Proof. Recall that elementary row operations on a coefficient matrix do not change the solution set of $\mathbf{Ax}=\mathbf{0}$. By the preceding corollary, then it follows that if $\mathbf{A}'$ is a matrix obtained from $\mathbf{A}$ by elementary row operations, then the dimension of its column space has to coincide with the dimension of the column space of $\mathbf{A}$. For this reason, we can without loss of generality reduce to the situation where $\mathbf{A}$ is a matrix in reduced row echelon form. But when a matrix is in reduced row echelon form it is clear that the linearly independent columns are precisely the columns that contain pivots, and these in turn correspond to non-zero rows of a (reduced) row echelon form. Thus,

$$\begin{aligned}\#\text{ pivots} &= \#\text{ linearly independent row vectors} = \dim\ RowSp\,(\mathbf{A})\equiv RowRank\,(\mathbf{A})\\ &= \#\text{ linearly independent colum vectors} = \dim ColSp\,(\mathbf{A})\equiv ColumnRank\,(\mathbf{A})\end{aligned}$$

□

1. Connection with Lecture 8

In th preceding lecture we described a methodical approach to solving a general $n\times m$ linear system $\mathbf{Ax}=\mathbf{b}$. The crux of this method was to row reduce the augmented matrix $[\mathbf{A}|\mathbf{b}]$ to it reduced row echelon form $[\mathbf{A}'|\mathbf{b}']$, and the regard the components x_i corresponding to the columns of $\mathbf{A}'$ that do not contain pivots as free parameters, and the components of x_i that correspond to columns of $\mathbf{A}'$ that contain pivots as the variables that can be expressed in terms of the free parameters via an equation corresponding to a non-zero row of $[\mathbf{A}'|\mathbf{b}']$. If the matrix $\mathbf{A}'$ has r pivots, then there would be $m-r$ columns without pivots and so

$m-r$ free parameters in the solution. Let $s_1,\dots,s_{m-r}$ be these free parameters. Then each component of a solution could be expressed in terms of the free parameters $s_1,\dots,s_{m-r}$

$$x_i \longleftrightarrow i^{th} \text{ column of } \mathbf{A} \begin{array}{lll} \nearrow & x_i\left(s_1,\dots,s_{m-r}\right) & \text{if } i^{th} \text{ column of } \mathbf{A}' \text{ contains a pivot} \\ & & \\ \searrow & s_j & \text{if } i^{th} \text{ column of } \mathbf{A}' \text{ is the } j^{th} \text{ column without a pivot} \end{array}$$

Via this interpretation of the columns of $\mathbf{A}'$, we could write down the general solution in terms on free parameters $s_1,\dots,s_{m-r}$

$$\mathbf{x} = \mathbf{b}' + s_1\mathbf{b}_1 + \cdots + s_{m-r}\mathbf{b}_{m-r}$$

Here $\mathbf{b}'$ is, on the one hand, the vector in the last column of the augmented matrix in reduced row echeleon form, and on the other hand the solution of $\mathbf{Ax}=\mathbf{b}$ obtained by setting all the free parameters equal to $0_{\mathbb{F}}$. The vectors $\mathbf{b}_1,\dots,\mathbf{b}_{m-r}$ are all solutions of the corresponding homogeneous problem. This can be seen by noting that if we had started with $\mathbf{b}=\mathbf{0}$ the same elementary row opersations that sent $[\mathbf{A}|\mathbf{b}]$ to $[\mathbf{A}'|\mathbf{b}']$ would send $[\mathbf{A}|\mathbf{0}]$ to $[\mathbf{A}'|\mathbf{0}]$; and so, following the method above, we would have the general solution of $\mathbf{Ax}=\mathbf{0}$ presented in the form

$$\mathbf{x} = \mathbf{0} + s_1\mathbf{b}_1 + \cdots + s_{m-r}\mathbf{b}_{m-r}$$

which shows every solution of $\mathbf{Ax}=\mathbf{0}$ as lying in the span of the vectors $\mathbf{b}_1,\dots,\mathbf{b}_{m-r}$.

All I wanted to point out in this section is the connection between the vectors $\mathbf{b}_1,\dots,\mathbf{b}_{m-r}$ of our concrete solution methodology and the vectors $\mathbf{u}_{r+1},\dots,\mathbf{u}_m$ of Theorem 9.3. Recall that the vectors $\mathbf{u}_i$ were formed by first of all reordering and relabeling the variables so that the first r columns of the coefficient matrix $\mathbf{A}$ provided a basis for the column space. Then the last $m-r$ columns could be expressed as linear combinations of the first r columns, which led to equations like

$$\mathbf{c}_i = \lambda_1^{(i)}\mathbf{c}_1 + \cdots + \lambda_r^{(i)}\mathbf{c}_r \quad , \quad i = r+1,\dots,m$$

We then used the coefficients $\lambda_j^{(i)}$ to form the vectors

$$\mathbf{u}_i = \left[\lambda_1^{(i)}, \lambda_2^{(i)}, \dots, \lambda_r^{(i)}, 0, \dots, -1, \dots, 0\right] \qquad , \qquad i = 1+r,\dots,m$$

that turned out to be a basis for the solution space of $\mathbf{Ax}=\mathbf{0}$.

To see the connection between the vectors $\mathbf{b}_1,\dots,\mathbf{b}_{m-r}$ and the vectors $\mathbf{u}_{r+1},\dots,\mathbf{u}_m$, suppose construct the vectors $\mathbf{u}_i$ starting with a linear system such that $[\mathbf{A}|\mathbf{0}]$ is already in reduced row echelon form. By reordering and relabeling the variables we can arrange so that the pivot columns all proceed the columns without pivots with $[\mathbf{A}|\mathbf{0}]$ still in reduced row echelon form. Then if $\mathbf{A}$ has r pivots, $[\mathbf{A}|\mathbf{0}]$ will have the form

$$\begin{array}{ccccccc|c} 1 & 0 & \cdots & 0 & a_{1,r+1} & \cdots & a_{1,m} & 0 \\ 0 & 1 & 0 & \vdots & a_{2,r+1} & \cdots & a_{2,m} & 0 \\ 0 & 0 & \ddots & \vdots & & \ddots & \vdots & \vdots \\ \vdots & & & 1 & a_{r,r+1} & \cdots & a_{r,m} & 0 \\ \vdots & & & & 0 & \cdots & 0 & 0 \\ \vdots & & & & \vdots & \ddots & \vdots & \vdots \\ 0 & & & & 0 & \cdots & 0 & 0 \end{array}$$

For such an $[\mathbf{A}|\mathbf{0}]$ the vectors $\mathbf{b}_1,\dots,\mathbf{b}_r$ will be

$$\begin{aligned} \mathbf{b}_1 &= [-a_{1,r+1}, -a_{2,r+1}, \dots, -a_{r,r+1}, 1, 0, \dots, 0] \\ \mathbf{b}_2 &= [-a_{1,r+2,}, -a_{2,r+2}, \dots, -a_{r,r+2}, 0, 1, \dots, 0] \\ &\vdots \\ \mathbf{b}_{m-r} &= [-a_{1,m}, -a_{2,m}, \dots, -a_{r,m}, 0, \cdots, 0, 1] \end{aligned}$$

These vectors are just the vectors $\mathbf{u}_{r+1}, \ldots, \mathbf{u}_m$ multiplied by -1. Indeed, each of the last $m-r$ columns of $\mathbf{A}$ (in the above reduced row echelon form) can be express as a linear combintion of the first r columns (indeed the first r columns of $\mathbf{A}$ are effectively a standard basis for the column space). For all $i = r+1, \ldots, m$ we'll have

$$\begin{aligned} \mathbf{c}_i &= a_{1,i}\mathbf{c}_1 + a_{2,i}\mathbf{c}_2 + \cdots + a_{r,i}\mathbf{c}_r \\ &\Rightarrow \quad \mathbf{u}_i = [a_{1,i} \cdots, a_{r,i}, 0, \ldots, 0, -1, 0, \ldots, 0] = -\mathbf{b}_{i-r} \end{aligned}$$

Thus, we fill the gap in the conclustion of Lecture 8 (that the vectors $\mathbf{b}_1, \ldots, \mathbf{b}_{m-r}$ actually provide a basis for the solution space of $\mathbf{A}\mathbf{x} = \mathbf{0}$).