LECTURE 7

Solving Linear Systems

The basic method for solving linear systems that taught in Math 3013 is based on the notion of a matrix in *reduced row echelon form.*

Let $\mathbf{A}$ be an $n \times m$ matrix with entries in a field $\mathbb{F}$. Recall that the first (reading left to right) entry in a row of $\mathbf{A}$ that is not equal to $0_{\mathbb{F}}$ is called the *pivot* of the corresponding row, and that a matrix in *row echelon form* has the property that pivots in successive rows always occur off to the right of the pivots in preceding rows. For a matrix to be in *reduced row echelon form*, we require not only that is in row echelon form, but also that each pivot is equal to $1_{\mathbb{F}}$, and that above and below each pivot only $0_{\mathbb{F}}$'s appear.

We note that it is always possible to convert the pivot λ (where of couse $\lambda \neq 0_{\mathbb{F}}$) of a row to $1_{\mathbb{F}}$ by simply scalar multiplying a row by λ^{-1} (a standard elementary operation). We further note that once the matrix is in row echelon form all the matrix entries below a pivot are already $0_{\mathbb{F}}$'s. Once the pivots have all been converted to $1_{\mathbb{F}}$'s, a non-zero entry, say λ, lying in row i above a pivot in row j, can be cleared out by replacing row i with its sum with $-\lambda$ times row j. Note that this operation will not affect any entries in row i that appear to the left of the pivot in row j. Thus, by employing elementary operations we can systematically convert a matrix in row echelon form to a matrix in reduced row echelon form.

Remark 7.1. Suppose $\mathbf{A}$ is an $n \times n$ matrix in reduced row echelon form and moreover has no non-zero rows. Then $\mathbf{A}$ is the $n \times n$ identity matrix. Indeed, each row of $\mathbf{A}$ must have a pivot which must equal $1_{\mathbb{F}}$ since $\mathbf{A}$ is in reduced row echelon form. So $\mathbf{A}$ has exactly n pivots. On the other hand, since $\mathbf{A}$ is also in row echelon form, its pivots must live in different columns. Since there are exactly n pivots and n columns, we must have exactly one pivot in each column and row and that pivot must be equal to $1_{\mathbb{F}}$. Thus, $\mathbf{A}$ must be of the form

$$\mathbf{A} = \begin{pmatrix} 1_{\mathbb{F}} & 0_{\mathbb{F}} & \cdots & \cdots & 0_{\mathbb{F}} \\ 0_{\mathbb{F}} & 1_{\mathbb{F}} & & & 0_{\mathbb{F}} \\ \vdots & & \ddots & & \vdots \\ 0_{\mathbb{F}} & & & 1_{\mathbb{F}} & 0_{\mathbb{F}} \\ 0_{\mathbb{F}} & & & 0_{\mathbb{F}} & 1_{\mathbb{F}} \end{pmatrix}$$

Theorem 7.2. *The reduced row echelon form of an $n \times m$ matrix $\mathbf{A}$ is unique.*

Proof.

We will use a proof by induction, where we induce on the number m of columns of $\mathbf{A}$. For $m = 1$, the result is clear since the only $n \times m$ matrix in reduced row echelon form is the $n \times 1$ matrix with $1_{\mathbb{F}}$ in the first slot and $0_{\mathbb{F}}$'s everywhere else.

Now as an induction hypothesis suppose the statement is true for any matrix with $m-1$ columns. Let $\mathbf{B}$ and $\mathbf{C}$ be two matrices in reduced row echelon form obtained from $\mathbf{A}$ by elementary row operations. Let $\mathbf{A}'$, $\mathbf{B}'$ and $\mathbf{C}'$ be the $n \times (m-1)$ matrices obtained from, respectively, $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ by deleting their last columns. Since the elementary row operations operate column and by column, as does the criteria for being in reduced row echelon form, it is clear that $\mathbf{B}'$ and $\mathbf{C}$' will be matrices in reduced echelon form obtained from $\mathbf{A}'$ by elementary row operations. By the induction hypothesis, then $\mathbf{B}' = \mathbf{C}'$.

$$\mathbf{B} = \left(\begin{array}{c|c} & b_{1m} \\ \mathbf{B}' & \vdots \\ & b_{nm} \end{array} \right) \quad , \quad \mathbf{C} = \left(\begin{array}{c|c} & c_{1m} \\ \mathbf{C}' & \vdots \\ & c_{nm} \end{array} \right)$$

Let's dispense with an easy case write away. Suppose the last columns of $\mathbf{B}$ and $\mathbf{C}$ contain a pivot (if one does so does the other – for, otherwise, $\dim RowSp(\mathbf{B})$ would not equal $\dim RowSp(\mathbf{C})$, which would contract the hypothesis that both $\mathbf{B}$ and $\mathbf{C}$ are row equivalent to $\mathbf{A}$). Since the submatrices $\mathbf{B}' = \mathbf{C}'$ are already in (reduced) row echelon form, there is only one place to put a pivot in the last column of these matrices – at the end of the row that follows the last non-zero of $\mathbf{B}' = \mathbf{C}'$. Since $\mathbf{B}$ and $\mathbf{C}$ are to be in reduced row echelon form, that pivot has to be equal to $1_{\mathbb{F}}$. Hence, in this situation, there is is only one way to add a column to $\mathbf{B}' = \mathbf{C}'$ to get a matrix in reduced row echelon form, and so we must have $\mathbf{C} = \mathbf{B}$.

Now suppose $\mathbf{B} \neq \mathbf{C}$ and the last colunms of $\mathbf{B}$ and $\mathbf{C}$ do not contain a pivot. Since we still have $\mathbf{B}' = \mathbf{C}'$, $\mathbf{B}$ and $\mathbf{C}$ can differ at most by entries in the m^{th} column. Suppose the first time an entry b_{jm} in the m^{th} column of $\mathbf{B}$ does not match the corresponding entry c_{jm} of $\mathbf{C}$ happens when $j = i$. Let $\mathbf{u}$ be any solution of $\mathbf{Bu} = \mathbf{0}_{\mathbb{F}^n}$. Since $\mathbf{C}$ is row equivalent to $\mathbf{B}$ (since both $\mathbf{B}$ and $\mathbf{C}$ are row equivalent to $\mathbf{A}$), $\mathbf{u}$ will also satisfy $\mathbf{Cu} = \mathbf{0}_{\mathbb{F}^n}$. Thus,

$$(\mathbf{B} - \mathbf{C})\,\mathbf{u} = \mathbf{0}_{\mathbb{F}^n}$$

On the other hand, the i^{th} coordinate of $(\mathbf{B} - \mathbf{C})\,\mathbf{u}$ is $(b_{im} - c_{im})\,u_m$. Since $b_{im} \neq c_{im}$ we must have $u_m = 0$. Thus, any solution $\mathbf{u}$ of $\mathbf{Bu} = \mathbf{0}_{\mathbb{F}^n} = \mathbf{Cu}$ must have $u_m = 0$. It follows that the last columns of $\mathbf{B}$ and $\mathbf{C}$ must contain some pivot otherwise, the last column would be a free column that puts no restriction on u_m. But this contradicts the premises of the situation of this paragraph, that $\mathbf{B} \neq \mathbf{C}$ and the last columns of $\mathbf{B}$ and $\mathbf{C}$ do not contain a pivot.

We can thus conclude that if $\mathbf{B}$ and $\mathbf{C}$ are each row equivalent to a given matrix $\mathbf{A}$, and both $\mathbf{B}$ and $\mathbf{C}$ are in row echelon form, then $\mathbf{B} = \mathbf{C}$. $\square$

Here is why we are interested in matrices in reduced row echeleon form. Suppose you are given an $n \times m$ linear system $S(\mathbf{A}, \mathbf{b})$

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1m}x_m &= b_1 \\ &\vdots \\ a_{n1}x_1 + \cdots + a_{nm}x_m &= b_n \end{aligned}$$

with coefficient matrix $\mathbf{A} = (a_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}}$ and inhomogeneous part $\mathbf{b} = (b_i)_{1 \leq i \leq n}$. Form the *augmented matrix* $[\mathbf{A} \mid \mathbf{b}]$: this will be the $n \times (m+1)$ matrix whose first m columns coincide with the corresponding columns of $\mathbf{A}$ and whose $(m+1)^{th}$ column coincides with the vector $\mathbf{b}$.

Lemma 7.3. *If $[\mathbf{B} \mid \mathbf{c}]$ is an (augmented) matrix obtained from $[\mathbf{A} \mid \mathbf{b}]$ by applying an elementary row operation, then the solutions of the linear system with coefficient matrix $\mathbf{B}$ and inhomogeneous part $\mathbf{c}$ coincide with the solutions of the linear system with coefficient matrix $\mathbf{A}$ and inhomogeneous part $\mathbf{b}$.*

Proof. Each row of an augmented matrix corresponds to a particular equation of the corresponding linear system. For example the i^{th} row of $[\mathbf{A} \mid \mathbf{b}]$ will be $[a_{i1}, a_{i2}, \ldots, a_{im}, b_i]$ and this corresponds to the equation

$$a_{i1}x_1 + \cdots + a_{im}x_m = b_i \quad .$$

There are three basic elementary operations to consider.

- If $[\mathbf{B} \mid \mathbf{c}]$ is obtained from $[\mathbf{A} \mid \mathbf{b}]$ by interchanging two rows, then clearly the solution sets are not going to change since the set of equations for $[\mathbf{B} \mid \mathbf{c}]$ will be the same as that of $[\mathbf{A} \mid \mathbf{b}]$, it's just that they'll be written in a slightly different order.

- If $[\mathbf{B} \mid \mathbf{c}]$ is obtained from $[\mathbf{A} \mid \mathbf{b}]$ by replacing a row by its scalar multiple by an non-zero element $\lambda \in \mathbb{F}$. Say this is done to the i^{th} row of $[\mathbf{A} \mid \mathbf{b}]$. The equation
$$a_{i1}x_1 + \cdots + a_{im}x_m = b_i$$
has exactly the same solutions as
$$\lambda a_{i1}x_1 + \cdots + \lambda a_{im}x_m = \lambda b_i \quad ,$$
it is clear that in modifying one row in this way is not going to affect the solutions of the corresponding linear systems.
- Suppose we replace the i^{th} row of $[\mathbf{A} \mid \mathbf{b}]$ with its sum with λ times its j^{th} row. In this case, we observe that any solution of
$$\begin{aligned} a_{ji}x_1 + \cdots + a_{jm}x_m &= b_j \\ a_{i1}x_1 + \cdots + a_{im}x_m &= b_m \end{aligned}$$
also satisfies
$$\begin{aligned} a_{j1}x_1 + \cdots + a_{jm}x_m &= b_j \\ (a_{i1} + \lambda a_{j1})\, x_1 + \cdots + (a_{im} + \lambda a_{jm})\, x_m &= b_i + \lambda b_j \end{aligned}$$
and vice versa.[1]

□

From this lemma, it follows that

Corollary 7.4. *If $[\mathbf{B} \mid \mathbf{c}]$ is a matrix in reduced row echelon form obtained from $[\mathbf{A} \mid \mathbf{b}]$ by a sequence of elementary row operations, then the solutions to the linear system corresponding to $[\mathbf{A} \mid \mathbf{b}]$ will be the same as the solutions to the linear system corresponding to $[\mathbf{B} \mid \mathbf{c}]$.*

Now I'll try to indicate what's especially nice about the set of equations corresponding to an augmented matrix in reduced row echelong form. This I'll do via a sequence of examples.

Example 7.5. Consider the system
$$\begin{aligned} x_1 + 2x_2 + x_3 &= 8 \\ x_1 - x_2 + x_3 &= 2 \\ 2x_1 + x_2 - x_3 &= 1 \end{aligned}$$
This system has augmented matrix
$$[\mathbf{A} \mid \mathbf{b}] = \left[\begin{array}{ccc|c} 1 & 2 & 1 & 8 \\ 1 & -1 & 1 & 2 \\ 2 & 1 & -1 & 1 \end{array}\right]$$

[1] Suppose $(x_1, \ldots, x_m)$ satisfies both
$$\begin{aligned} a_{j1}x_1 + \cdots + a_{jm}x_m &= b_j \qquad (1a) \\ a_{i1}x_1 + \cdots + a_{im}x_m &= b_i \qquad (1b) \end{aligned}$$
Then, for any $\lambda \in \mathbb{F}$ we have an identity
$$\lambda a_{j1}x_1 + \cdots + \lambda a_{jm}x_m = \lambda b_j \qquad (2)$$
And so if we add the left hand side of (2) to the left hand side of (1b) and the right hand side of (2) to the right hand side of (1b) we get the identity
$$(a_{i1} + \lambda a_{j1})\, x_1 + \cdots + (a_{im} + \lambda a_{jm})\, x_m = b_i + \lambda b_j \qquad (3)$$
and so any solution of (1a) and (1b) will also be a solution of (1a) and (3). The opposite inclusion is proved similarly.

Using the elementary row operations we can transform $[\mathbf{A} \mid \mathbf{b}]$ to the following reduced row echelon form:

$$[\mathbf{A}' \mid \mathbf{b}'] = \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

The solutions to the system of equations corresponding to $[\mathbf{A}' \mid \mathbf{b}]$ must be same as the solution to the original system. But once we write them down

$$\left.\begin{array}{l} 1 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 = 1 \\ 0 \cdot x_1 + 1 \cdot x_2 + 0 \cdot x_3 = 2 \\ 0 \cdot x_1 + 0 \cdot x_2 + 1 \cdot x_3 = 3 \end{array}\right\} \quad \Longleftrightarrow \quad \left\{\begin{array}{l} x_1 = 1 \\ x_2 = 2 \\ x_3 = 3 \end{array}\right.$$

We see that the equations corresponding to the reduced row echelon form *effectively state the solution.*

EXAMPLE 7.6. Consider the linear system

$$\begin{array}{rcl} x_1 - x_2 + x_3 & = & 2 \\ x_1 + 2x_2 + x_3 & = & 5 \\ x_1 + x_2 - x_3 & = & 0 \\ -x_1 + x_2 & = & 0 \end{array}$$

It's augmented matrix

$$\left[\begin{array}{ccc|c} 1 & -1 & 1 & 2 \\ 1 & 2 & 1 & 5 \\ 1 & 1 & -1 & 0 \\ -1 & 1 & 0 & 0 \end{array}\right]$$

row reduces to the following matrix in reduced row echelon form

$$: \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

From the augmented matrix in reduced row echelon form we can again read off the solution

$$\begin{array}{rcl} x_1 & = & 1 \\ x_2 & = & 1 \\ x_3 & = & 2 \\ 0 & = & 0 \end{array}$$

of the original linear system.

In the two preceding examples, we had systems for which there is just a single solution (just one set of values of x_1, x_2 and x_3 that would provide a solution). The next example will serve to illustrate the way things turn out when there are multiple solutions.

EXAMPLE 7.7. Consider the linear system

$$\begin{array}{rcl} x_1 + x_2 + x_3 - x_4 & = & 2 \\ x_1 - x_2 - x_3 + x_4 & = & 0 \\ 3x_1 + x_2 + x_3 - x_4 & = & 4 \end{array}$$

(Note that we have fewer equations than unknowns - so we should expect multiple solutions.) As before we write down the augmented matrix and then transform it to a matrix in reduced row echelon form

$$\left[\begin{array}{cccc|c} 1 & 1 & 1 & -1 & 2 \\ 1 & -1 & -1 & 1 & 0 \\ 3 & 1 & 1 & -1 & 4 \end{array}\right] \quad \rightarrow \quad \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

The equations corresponding to the augmented matrix in reduced row echelon form are

$$\begin{aligned} x_1 &= 1 \\ x_2 + x_3 - x_4 &= 1 \\ 0 &= 0 \end{aligned}$$

This presentation of the solution space is not nearly as clean as in the preceding two examples, but it's about the best we can do, because we simply don't have enough independent equations to determine the solution uniquely. All we can say is that x_1 has to equal 1 and x_2, x_3 and x_4 are related by $x_2+x_3-x_4 = 1$.

However, we can provide a slightly better interpretation of the equations. Let us adopt the convention that when we write down the equations corresponding to a augmented matrix in reduced row echelon form, we keep the variables corresponding to columns with pivots on the left hand side and move the variable corresponding to columns without pivots to the right hand side (after multiplying by -1).

In the present case, the pivots of

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

occur in columns 1 and 2; and so when we write down the equations for this augmented matrix we'll keep the variables x_1 and x_2 on the left hand side and move the variables x_3 and x_4 over to the right hand side (as these variables correspond to columns without pivots). Thus, we write

$$\begin{aligned} x_1 &= 1 \\ x_2 &= 1 - x_3 + x_4 \end{aligned}$$

We now interprete this last set of equations as saying $x_1 = 1$, x_2 is a certain function of x_3 and x_4, but there is no restriction placed on the variables x_3 and x_4.

Let us now write down a typical solution vector. Since no restriction is placed on the variables x_3 and x_4, we can allow these variables to be any real number. Say, $x_3 = r \in \mathbb{R}$ and $x_4 = s \in \mathbb{R}$. Then a typical solution vector will have the form

$$(*) \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 1-r+s \\ r \\ s \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + r \begin{bmatrix} 0 \\ -1 \\ 1 \\ 0 \end{bmatrix} + s \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \quad ; \quad r, s \in \mathbb{R}$$

or

$$\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \text{ something in the span of } \begin{bmatrix} 0 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

We note that this last presentation of the solution is somewhat reminiscent of Theorem 7.7 In fact, as we'll now see, that this last presentation corresponds precisely to the way Theorem 7.7 tells us to write down solutions to an inhomogeneous linear system.

Consider now the homogeneous linear system corresponding to the present example:

$$\begin{aligned} x_1 + x_2 + x_3 - x_4 &= 0 \\ x_1 - x_2 - x_3 + x_4 &= 0 \\ 3x_1 + x_2 + x_3 - x_4 &= 0 \end{aligned}$$

Row reduceing its augmented matrix we get

$$\left[\begin{array}{cccc|c} 1 & 1 & 1 & -1 & 0 \\ 1 & -1 & -1 & 1 & 0 \\ 3 & 1 & 1 & -1 & 0 \end{array}\right] \quad \rightarrow \quad \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

From which we conclude, using the convention above,

$$\begin{aligned} x_1 &= 0 \\ x_2 &= -x_3 + x_4 \end{aligned}$$

Any vector $\mathbf{x}_0$ satisfying these two equations will have to have the form

(**) $$\mathbf{x}_0 = \begin{bmatrix} 0 \\ -r+s \\ r \\ s \end{bmatrix} = r \begin{bmatrix} 0 \\ -1 \\ 1 \\ 0 \end{bmatrix} + s \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

Comparing (**) with (*), we see that the last two terms on the right in (*) correspond precisely to $\mathbf{x}_0$ which is the general solution of the corresponding homogeneous system. The first term on the right hand side of (*)

$$\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

is, by itself, a solution of the original inhomogeneous system. It is the solution obtained by choosing r and s to be 0.

In summary, the reduced echelon form of the augmented matrix gives us a means of writing the general solution to an inhomogeneous linear system as the sum of a particular solution plus the general solution of the corresponding homogeneous system.

Example 7.8. Express the solution of the following linear system in the form expressed by Theorem 7.7.

$$\begin{aligned} x_1 - x_2 + 2x_3 &= 1 \\ x_1 + x_2 - x_3 &= 2 \end{aligned}$$

Following our by now standard procedure we calculate the corresponding augmented matrix in reduced row echelon form:

$$\left[\begin{array}{ccc|c} 1 & -1 & 2 & 1 \\ 1 & 1 & -1 & 2 \end{array}\right] \quad \rightarrow \quad \left[\begin{array}{ccc|c} 1 & 0 & \frac{1}{2} & \frac{3}{2} \\ 0 & 1 & -\frac{3}{2} & \frac{1}{2} \end{array}\right]$$

So we have

$$\begin{aligned} x_1 &= \frac{3}{2} - \frac{1}{2}x_3 \\ x_2 &= \frac{1}{2} + \frac{3}{2}x_3 \end{aligned}$$

Regarding x_3 as a free variable, we can write the general solution as

(***) $$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \frac{3}{2} - \frac{1}{2}s \\ \frac{1}{2} + \frac{3}{2}s \\ s \end{bmatrix} = \begin{bmatrix} \frac{3}{2} \\ \frac{1}{2} \\ 0 \end{bmatrix} + s \begin{bmatrix} -\frac{1}{2} \\ \frac{3}{2} \\ 1 \end{bmatrix}$$

Regarding the vector $\begin{bmatrix} \frac{3}{2} \\ \frac{1}{2} \\ 0 \end{bmatrix}$ as a particular solution of the original linear system and $s \begin{bmatrix} -\frac{1}{2} \\ \frac{3}{2} \\ 1 \end{bmatrix}$ as the general solution of the corresponding homogeneous system, the right hand side of (***) expresses the general solution of the original linear system in the form given by Theorem 7.7.

1. Summary: Solving Linear Systems

Below are the steps we've employed to solve a linear system an $n \times m$ linear system $\mathbf{Ax} = \mathbf{b}$ with coefficient matrix $\mathbf{A} \in Mat_{n,m}(\mathbb{F})$ and inhomogeneous part $\mathbf{b} \in \mathbb{F}^m$.

(1) Form the augmented matrix $[\mathbf{A}|\mathbf{b}] \in Mat_{n,m+1}(\mathbb{F})$.
(2) Apply elementary row operations to reduce $[\mathbf{A}|\mathbf{b}]$ to a matrix $[\mathbf{A}'|\mathbf{b}']$ in *reduced row echelon form.*
(3) Identify the columns of $\mathbf{A}'$ that contain pivots (first non-zero entries of a row). The columns without pivots will correspond to the free parameters of the solution, the columns with pivots will correspond to the variables that can be expressed in terms of the free parameters of the solution. Thus, if the matrix $\mathbf{A}'$ has r pivots, there will be $m-r$ free parameters in the solution; let's denote these free parameters by $s_1, s_2, \ldots, s_{m-r}$.
(4) Write down the linear system corresponding to the augmented matrix $[\mathbf{A}'|\mathbf{b}']$ and then move the variables corresponding to free parameters (corresponding in turn to columns without pivots) to the right hand side. With these equations, you can now express each component x_i of a solution vector $\mathbf{x}$ as a linear function of the free parameters

$$x_i = \begin{cases} x_i\left(s_1, \ldots, s_{m-r}\right) & ;\text{ if } x_i \text{ corresponds to a column with a pivot} \\ s_j & ;\text{ if } x_i \text{ corresponds to the } j^{th} \text{ column without a pivot} \end{cases}$$

(5) Write down the general solution vector as

$$\mathbf{x} = \begin{bmatrix} x_1\left(s_1, \ldots, s_{m-r}\right) \\ x_2\left(s_1, \ldots, s_{m-r}\right) \\ \vdots \\ x_m\left(s_1, \ldots, s_{m-r}\right) \end{bmatrix}$$

and then $\mathbf{x}$ with respect the free parameters. This will lead to a vector equation in the form

$$\mathbf{x} = \mathbf{b}' + s_1\mathbf{b}_1 + s_2\mathbf{b}_2 + \cdots + s_{m-r}\mathbf{b}_{m-r} \tag{*}$$

Here $\mathbf{b}'$ corresponds precisely to the last column of the augmented matrix $[\mathbf{A}'|\mathbf{b}']$ in reduced row echelon form. The vectors $\mathbf{b}_i$ in this expansion are formed via

$$j^{th} \text{ component of } \mathbf{b}_i := \text{coefficient of } s_i \text{ in the } j^{th} \text{ component of } \mathbf{x}$$

We shall see latter that the vectors $\mathbf{b}_i, \ldots, \mathbf{b}_{m-r}$ actually form a basis for the solution space of the corresponding homogeneous linear system:

$$\text{solution set of } \mathbf{Ax} = \mathbf{0} \quad = \quad span_{\mathbb{F}}\left(\mathbf{b}_1, \ldots, \mathbf{b}_{m-r}\right)$$

Also, since the $s_1, \ldots, s_{m-r}$ are free parameters, we can obtain a particularly simple solution by setting each $s_i = 0_{\mathbb{F}}$. Thus, equation (*) amounts to writing the solution of $\mathbf{Ax} = \mathbf{0}$ in the form expressed by Theorem 7.7:

$$\mathbf{x} = (\text{a particular solution of } \mathbf{Ax} = \mathbf{b}) + (\text{a solution of } \mathbf{Ax} = \mathbf{0})$$