



6.4.6 Solutions to systems of linear equations with backward substitution





Introduction

- In this topic, we will
 - Focus on augmented matrices in row-echelon form
 - Determine if such a matrix has no solutions, one solution, or infinitely many solutions
 - Examine the algorithm to find the unique solution if only one solutions exists
 - Introduce the concept of free variables when there are infinitely many solutions
 - Examine the algorithm to find all solutions when there are infinitely many solutions





In row-echelon form...

- Suppose the augmented matrix corresponding to a system of linear equations is in row-echelon form
- We will see how to determine if there
 - Are no solutions
 - One unique solution
 - Infinitely many solutions
- If there is one unique solution or infinitely many solutions, we will describe how to find it or them

$$\left(\begin{array}{cccc|c} -3 & 6 & 0 & 3 & 3.9 \\ 0 & 4 & 9 & 7 & -9.2 \\ 0 & 0 & 0 & 8 & 2.7 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right)$$





No solutions

- If $(A \mid \mathbf{b})$ is in row-echelon form and $\text{rank}(A) \neq \text{rank}(A \mid \mathbf{b})$, then the last non-zero row must be

$$(0 \quad 0 \quad 0 \quad \cdots \quad 0 \mid b_k)$$

where $b_k \neq 0$

- This is the same as saying

$$0x_1 + 0x_2 + 0x_3 + \cdots + 0x_n = b_k \neq 0$$

- This is, of course, impossible, so no solution exists





At least one solution

- If $(A \mid \mathbf{b})$ is in row-echelon form and $\text{rank}(A) = \text{rank}(A \mid \mathbf{b})$, then at least one solution exists
 - We will see how to differentiate there being one unique solution and infinitely many solutions





One unique solution

- If A is an $m \times n$ matrix (meaning, there are n unknowns) and the augmented matrix $(A \parallel \mathbf{b})$ is in row-echelon form,
a unique solution exists if and only if
$$n = \text{rank}(A) = \text{rank}(A \parallel \mathbf{b})$$
 - We can find the solution with backward substitution
- If the rank is n , then there must be n leading non-zero entries
 - If there are n columns of A , then because the matrix is in row-echelon form, each column must have a leading non-zero entry
 - Thus, A must be an upper-triangular matrix with no non-zero entries on the diagonal





One unique solution

- For example, here A is 2×2 , the augmented matrix is in row-echelon form, and the rank of A equals the rank of the augmented matrix

$$\left(\begin{array}{cc|c} 3 & -7 & 4.3 \\ 0 & 5 & -6.5 \end{array} \right)$$

$$5x_2 = -6.5$$

$$x_2 = -1.3$$

$$3x_1 - 7x_2 = 4.3$$

$$3x_1 - 7(-1.3) = 4.3$$

$$3x_1 = 4.3 - 9.1$$

$$= -4.8$$

$$x_1 = -1.6$$

- The solution vector is $\mathbf{x} = \begin{pmatrix} -1.6 \\ -1.3 \end{pmatrix}$





One unique solution

- You can check this:

- Given $\left(\begin{array}{cc|c} 3 & -7 & 4.3 \\ 0 & 5 & -6.5 \end{array}\right)$ with the solution vector $\mathbf{x} = \begin{pmatrix} -1.6 \\ -1.3 \end{pmatrix}$

- You can substitute these values into both equations:

$$3x_1 - 7x_2 = 4.3$$

$$5x_2 = -6.5$$

- Alternatively, you can confirm that

$$-1.6 \begin{pmatrix} 3 \\ 0 \end{pmatrix} - 1.3 \begin{pmatrix} -7 \\ 5 \end{pmatrix} = \begin{pmatrix} 4.3 \\ -6.5 \end{pmatrix}$$





One unique solution

- For example, here A is 4×4 , the augmented matrix is in row-echelon form, and the rank of A equals the rank of the augmented matrix

$$\left(\begin{array}{cccc|c} -3 & 6 & 0 & 3 & 3.9 \\ 0 & 4 & 9 & 7 & -9.2 \\ 0 & 0 & 5 & 8 & 2.7 \\ 0 & 0 & 0 & -2 & -3.8 \end{array} \right)$$

$$-2x_4 = -3.8$$

$$x_4 = 1.9$$

$$5x_3 + 8x_4 = 2.7$$

$$5x_3 + 8 \cdot 1.9 = 2.7$$

$$5x_3 = 2.7 - 15.2$$

$$= -12.5$$

$$x_3 = -2.5$$





One unique solution

- For example, here A is 4×4 , the augmented matrix is in row-echelon form, and the rank of A equals the rank of the augmented matrix

$$\left(\begin{array}{cccc|c} -3 & 6 & 0 & 3 & 3.9 \\ 0 & 4 & 9 & 7 & -9.2 \\ 0 & 0 & 5 & 8 & 2.7 \\ 0 & 0 & 0 & -2 & -3.8 \end{array} \right) \quad \begin{array}{l} 4x_2 + 9x_3 + 7x_4 = -9.2 \\ 4x_2 + 9(-2.5) + 7 \cdot 1.9 = -9.2 \\ 4x_2 = -9.2 + 22.5 - 13.3 \\ \quad = 0 \\ \quad x_2 = 0 \end{array}$$

$$\text{-- Thus, } \mathbf{x} = \begin{pmatrix} 0.6 \\ 0 \\ -2.5 \\ 1.9 \end{pmatrix} \quad \begin{array}{l} -3x_1 + 6x_2 + 0x_3 + 3x_4 = 3.9 \\ -3x_1 + 6 \cdot 0 + 0(-2.5) + 3 \cdot 1.9 = 3.9 \\ -3x_1 = 3.9 - 5.7 \\ \quad = -1.8 \\ \quad x_1 = 0.6 \end{array}$$





One unique solution

- You can check this:

– Given
$$\begin{pmatrix} -3 & 6 & 0 & 3 & | & 3.9 \\ 0 & 4 & 9 & 7 & | & -9.2 \\ 0 & 0 & 5 & 8 & | & 2.7 \\ 0 & 0 & 0 & -2 & | & -3.8 \end{pmatrix}$$
 with the solution $\mathbf{x} = \begin{pmatrix} 0.6 \\ 0 \\ -2.5 \\ 1.9 \end{pmatrix}$

- You can substitute these values into the linear equations
- Alternatively, you can confirm that $-3x_1 + 6x_2 + 3x_4 = 3.9$

$$0.6 \begin{pmatrix} -3 \\ 0 \\ 0 \\ 0 \end{pmatrix} + 0 \begin{pmatrix} 6 \\ 4 \\ 9 \\ 0 \end{pmatrix} + (-2.5) \begin{pmatrix} 0 \\ 5 \\ 1 \\ 0 \end{pmatrix} + 1.9 \begin{pmatrix} 3 \\ 7 \\ 8 \\ -2 \end{pmatrix} = \begin{pmatrix} 3.9 \\ -9.2 \\ 2.7 \\ -3.8 \end{pmatrix}$$





One unique solution

- For example, here A is 5×3 , the augmented matrix is in row-echelon form, and the rank of A equals the rank of the augmented matrix

$$\left(\begin{array}{ccc|c} 2 & 3 & 6 & -2.3 \\ 0 & 4 & -2 & 4.6 \\ 0 & 0 & 5 & -4.5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

$$\mathbf{x} = \begin{pmatrix} 0.5 \\ 0.7 \\ -0.9 \end{pmatrix}$$

$$5x_3 = -4.5$$

$$x_3 = -0.9$$

$$4x_2 - 2x_3 = 4.6$$

$$4x_2 - 2(-0.9) = 4.6$$

$$4x_2 = 4.6 - 1.8$$

$$= 2.8$$

$$x_2 = 0.7$$

$$2x_1 + 3 \cdot 0.7 + 6(-0.9) = -2.3$$

$$2x_1 = -2.3 + 5.4 - 2.1$$

$$= 1.0$$

$$x_1 = 0.5$$





One unique solution

- Consequently, the vector $\mathbf{x} = \begin{pmatrix} 0.5 \\ 0.7 \\ -0.9 \end{pmatrix}$ is the solution vector to

$$\left(\begin{array}{ccc|c} 2 & 3 & 6 & -2.3 \\ 0 & 4 & -2 & 4.6 \\ 0 & 0 & 5 & -4.5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

$$2x_1 + 3x_2 + 6x_3 = -2.3$$

$$4x_2 - 2x_3 = 4.6$$

$$5x_3 = -4.5$$

$$0x_1 + 0x_2 + 0x_3 = 0$$

$$0x_1 + 0x_2 + 0x_3 = 0$$

$$0.5 \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + 0.7 \begin{pmatrix} 3 \\ 4 \\ 0 \\ 0 \\ 0 \end{pmatrix} - 0.9 \begin{pmatrix} 6 \\ -2 \\ 5 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -2.3 \\ 4.6 \\ -4.5 \\ 0 \\ 0 \end{pmatrix}$$





The backward-substitution algorithm

- The algorithm for finding a unique solution is as follows:
 - If $(A \parallel \mathbf{b})$ is an augmented matrix corresponding to a system of m equations in n unknowns where $m \geq n$,
and the augmented matrix is in row-echelon form
and where $\text{rank}(A) = \text{rank}(A \parallel \mathbf{b}) = n$
 - Then start with $k \leftarrow n$
 1. Substitute all values into the equation defined by Row k that have been determined, which must leave a linear equation in x_k
 2. Then solve for x_k
 3. If $k = 1$, we are done,
otherwise set $k \leftarrow k - 1$ and return to Step 1.





The last possibility

- The story so far, for an augmented matrix in row-echelon form:
 - If $\text{rank}(A) \neq \text{rank}(A \parallel \mathbf{b})$, no solution exists
 - If A is $m \times n$, meaning there are n unknowns,
and $\text{rank}(A) = \text{rank}(A \parallel \mathbf{b}) = n$, then a unique solution exists
 - In this case, $m \geq n$, meaning, there must be at least as many equations than there are unknowns
 - Can $\text{rank}(A) = \text{rank}(A \parallel \mathbf{b}) > n$?
 - This means there are more leading non-zero terms than columns
 - This is impossible
 - The only other possibility is that $\text{rank}(A) = \text{rank}(A \parallel \mathbf{b}) < n$





Infinitely many solutions

- If A is $m \times n$, meaning there are n unknowns,
and $\text{rank}(A) = \text{rank}(A \parallel \mathbf{b}) < n$,
then there are fewer non-zero equations than unknowns
- For example, $x_1 + x_2 = 0$ is a single equation with two unknowns
 - The corresponding augmented matrix is $\begin{pmatrix} 1 & 1 & \vdots & 0 \end{pmatrix}$
 - There are infinitely many solutions: whatever value x_2 takes on,

if $x_1 = -x_2$, then $\mathbf{x} = \begin{pmatrix} -x_2 \\ x_2 \end{pmatrix}$ is a solution vector





Infinitely many solutions

- Looking at the augmented matrix in row-echelon form, we note that if a variable is associated with a leading non-zero coefficient, this means it can be written as a linear combination of all subsequent variables
- For example, given

$$\left(\begin{array}{cccc|c} 7.6 & -0.3 & 1.5 & -6.0 & 8.9 \\ 0 & 0 & 9.2 & 3.8 & -5.4 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right) \quad \begin{aligned} x_1 &= \frac{8.9 + 0.3x_2 - 1.5x_3 + 6.0x_4}{7.6} \\ x_3 &= \frac{-5.4 - 3.8x_4}{9.2} \end{aligned}$$

- There are no such constraints for x_2 or x_4
 - Thus, these two unknowns are *free variables*





Infinitely many solutions

- Thus, if A is $m \times n$ and $(A \mid \mathbf{b})$ is in row-echelon form, and $\text{rank}(A) = \text{rank}(A \mid \mathbf{b}) < n$, then each unknown associated with a column that does not contain a leading non-zero entry is a free variable

$$\begin{array}{cccc|c} & x_2 & & x_4 & \\ \hline 7.6 & -0.3 & 1.5 & -6.0 & 8.9 \\ 0 & 0 & 9.2 & 3.8 & -5.4 \\ 0 & 0 & 0 & 0 & 0 \end{array}$$





Infinitely many solutions

- If $(A \mid \mathbf{b})$ is in row-echelon form where A is $m \times n$, and at least one solution exists, then the number of free variables is $n - \text{rank}(A)$

- For example:

$$\begin{array}{cccc|c} & & x_3 & x_4 & \\ \left(\begin{array}{ccccc} 8.5 & 2.1 & 0 & 2.0 & 0 \\ 0 & 8.3 & 1.3 & 8.9 & 1.3 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right) \end{array}$$

$$\begin{array}{cc|c} & x_2 & \\ \left(\begin{array}{ccc} 7.6 & -0.3 & 1.5 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array} \right) \end{array}$$

$$\begin{array}{ccccc|c} & x_2 & & x_4 & & \\ \left(\begin{array}{cccccc} 1.7 & -0.8 & 9.5 & -4.8 & 9.1 & 2.3 \\ 0 & 0 & 2.9 & 5.4 & -2.0 & 9.2 \\ 0 & 0 & 0 & 0 & 3.7 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right) \end{array}$$

$$\begin{array}{cccc|c} x_1 & & x_3 & x_4 & \\ (0 & 2.0 & 0.2 & -1.5 & \vdots & 0) \end{array}$$





Example

- For example,

$$\begin{pmatrix} \textcolor{red}{2} & \textcolor{red}{x_2} & \vdots & 4 \end{pmatrix}$$

$$2x_1 + 3x_2 = 4$$

$$2x_1 = 4 - 3x_2$$

$$x_1 = 2 - 1.5x_2$$

- Thus, the solution vector is $\mathbf{x} = \begin{pmatrix} 2 - 1.5x_2 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} + \begin{pmatrix} -1.5x_2 \\ x_2 \end{pmatrix}$
$$= \begin{pmatrix} 2 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} -1.5 \\ 1 \end{pmatrix}$$





IMPORTANT

- Top students will always ask, so I will address it now:
There are alternative ways of coming up with a solution, but all solutions are equivalent
- For example, given $\begin{pmatrix} 2 & 1 & \vdots & 3 \end{pmatrix}$, you could either
 - Let x_2 be a free variable, in which case the solution is
$$\mathbf{x} = \begin{pmatrix} 1.5 - 0.5x_2 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1.5 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} -0.5 \\ 1 \end{pmatrix}$$
 - Let x_1 be a free variable, in which case the solution is
$$\mathbf{x} = \begin{pmatrix} x_1 \\ 3 - 2x_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \end{pmatrix} + x_1 \begin{pmatrix} 1 \\ -2 \end{pmatrix}$$
- These describe the same solutions in different ways
- We are describing an *algorithm* that gives a solution
 - An algorithm means you can implement it in a computer





Example

- For example,
$$\begin{pmatrix} 2 & -1 & 4 & | & -6 \\ 0 & 0 & 5 & | & 7 \end{pmatrix}$$

$$5x_3 = 7 \text{ or } x_3 = 1.4$$

$$2x_1 - x_2 + 4x_3 = -6$$

$$2x_1 = -6 + x_2 - 4 \cdot 1.4$$

$$x_1 = -5.8 + 0.5x_2$$

- Thus, the solution is
$$\mathbf{x} = \begin{pmatrix} -5.8 + 0.5x_2 \\ x_2 \\ 1.4 \end{pmatrix} = \begin{pmatrix} -5.8 \\ 0 \\ 1.4 \end{pmatrix} + \begin{pmatrix} 0.5x_2 \\ x_2 \\ 0 \end{pmatrix}$$
$$= \begin{pmatrix} -5.8 \\ 0 \\ 1.4 \end{pmatrix} + x_2 \begin{pmatrix} 0.5 \\ 1 \\ 0 \end{pmatrix}$$





Example

- For example,
$$\begin{pmatrix} 20 & -2 & 7 & \vdots & 13 \\ 0 & -10 & 11 & \vdots & -9 \end{pmatrix}$$

$$-10x_2 + 11x_3 = -9$$

$$-10x_2 = -9 - 11x_3$$

$$x_2 = 0.9 + 1.1x_3$$

$$20x_1 - 2x_2 + 7x_3 = 13$$

$$20x_1 = 13 + 2(0.9 + 1.1x_3) - 7x_3$$

$$20x_1 = 14.8 - 4.8x_3$$

$$x_1 = 0.74 - 0.24x_3$$

- Thus, the solution is
$$\mathbf{x} = \begin{pmatrix} 0.74 - 0.24x_3 \\ 0.9 + 1.1x_3 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0.74 \\ 0.9 \\ 0 \end{pmatrix} + \begin{pmatrix} -0.24x_3 \\ 1.1x_3 \\ x_3 \end{pmatrix}$$
$$= \begin{pmatrix} 0.74 \\ 0.9 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} -0.24 \\ 1.1 \\ 1 \end{pmatrix}$$





Example

- Let's examine this solution to this system of linear equations:

$$\left(\begin{array}{ccc|c} 20 & -2 & 7 & 13 \\ 0 & -10 & 11 & -9 \end{array} \right) \quad \begin{array}{l} 20x_1 - 2x_2 + 7x_3 = 13 \\ -10x_2 + 11x_3 = -9 \end{array} \quad \mathbf{x} = \begin{pmatrix} 0.74 \\ 0.9 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} -0.24 \\ 1.1 \\ 1 \end{pmatrix}$$

- If $x_3 = 1$, then $\mathbf{x} = \begin{pmatrix} 0.5 \\ 2 \\ 1 \end{pmatrix}$ $\begin{array}{l} 20 \cdot 0.5 - 2 \cdot 2 + 7 \cdot 1 = 13 \\ -10 \cdot 2 + 11 \cdot 1 = -9 \end{array}$

- If $x_3 = -19$, then $\mathbf{x} = \begin{pmatrix} 5.3 \\ -20 \\ -19 \end{pmatrix}$ $\begin{array}{l} 20 \cdot 5.3 - 2 \cdot (-20) + 7 \cdot (-19) = 13 \\ -10 \cdot (-20) + 11 \cdot (-19) = -9 \end{array}$





Example

• For example,
$$\begin{array}{cccccc|c} & x_2 & & & x_5 & x_6 & \\ \left(\begin{array}{cccccc|c} 2 & 3 & 4 & -6 & 1 & -5 & 1 \\ 0 & 0 & 5 & 2 & 9 & 3 & 2 \\ 0 & 0 & 0 & -2 & -3 & 2 & 3 \end{array} \right) \end{array}$$

$$-2x_4 = 3 + 3x_5 - 2x_6$$

$$x_4 = -1.5 - 1.5x_5 + x_6$$

$$5x_3 = 2 - 2x_4 - 9x_5 - 3x_6$$

$$= 2 - 2(-1.5 - 1.5x_5 + x_6) - 9x_5 - 3x_6 \quad x_3 = 1 - 1.2x_5 - x_6$$

$$= 2 + (3 + 3x_5 - 2x_6) - 9x_5 - 3x_6$$

$$= 5 - 6x_5 - 5x_6$$

$$2x_1 = 1 - 3x_2 - 4x_3 + 6x_4 - x_5 + 5x_6$$

$$= 1 - 3x_2 - 4(1 - 1.2x_5 - x_6) + 6(-1.5 - 1.5x_5 + x_6) - x_5 + 5x_6$$

$$= 1 - 3x_2 - 4 + 4.8x_5 + 4x_6 - 9 - 9x_5 + 6x_6 - x_5 + 5x_6$$

$$= -12 - 3x_2 - 5.2x_5 + 15x_6$$

$$x_1 = -6 - 1.5x_2 - 2.6x_5 + 7.5x_6$$





Example

- Thus, given

$$x_1 = -6 - 1.5x_2 - 2.6x_5 + 7.5x_6$$

$$x_3 = 1 - 1.2x_5 - x_6$$

$$x_4 = -1.5 - 1.5x_5 + x_6$$

we may deduce the solution vector is

$$\mathbf{x} = \begin{pmatrix} -6 - 1.5x_2 - 2.6x_5 + 7.5x_6 \\ x_2 \\ 1 - 1.2x_5 - x_6 \\ -1.5 - 1.5x_5 + x_6 \\ x_5 \\ x_6 \end{pmatrix} = \begin{pmatrix} -6 \\ 0 \\ 1 \\ -1.5 \\ 0 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} -1.5 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_5 \begin{pmatrix} -2.6 \\ 0 \\ -1.2 \\ -1.5 \\ 1 \\ 0 \end{pmatrix} + x_6 \begin{pmatrix} 7.5 \\ 0 \\ -1 \\ 1 \\ 0 \\ 1 \end{pmatrix}$$





The backward-substitution algorithm

- The algorithm for finding all solution is as follows:
 - If $(A \parallel \mathbf{b})$ is an augmented matrix corresponding to a system of m equations in n unknowns,
and the augmented matrix is in row-echelon form
and where $\text{rank}(A) = \text{rank}(A \parallel \mathbf{b}) < n$
then there are $n - \text{rank}(A)$ free variables
 - Declare any unknown associated with a column not containing a leading non-zero entry as a free variable





The backward-substitution algorithm

- The algorithm for finding all solution is as follows:
 - Then start with $k \leftarrow \text{rank}(A)$
 1. Determine which variable x_ℓ is associated with the leading non-zero entry of Row k
 2. Substitute any values into the equation defined by Row k that have been determined, which must leave a linear equation in x_ℓ and any free variables greater than ℓ
 3. Then solve for x_ℓ
 4. If $k = 1$, we are done,
otherwise set $k \leftarrow k - 1$ and return to Step 1.





The backward-substitution algorithm

- Interestingly enough, this second algorithm also works if there is a single unique solution
 - The only difference is that there will be no free variables
- Thus, in all cases, the solution vector will be of the form:
 - A fixed vector plus all linear combinations of a set of vectors, each of which is associated with one of the free variables
 - If there is only one solution, it is that fixed vector





Summary

- Following this topic, you now
 - Understand that if an augmented matrix is in row-echelon form, we can determine if there are no solutions, one unique solution, or infinitely many solutions
 - Have seen the algorithm for finding the unique solution if there is only one solution
 - Have seen the algorithm for finding all solutions if there are infinitely many solutions





References

- [1] https://en.wikipedia.org/wiki/Row_echelon_form
- [2] [https://en.wikipedia.org/wiki/Rank_\(linear_algebra\)](https://en.wikipedia.org/wiki/Rank_(linear_algebra))
- [3] https://en.wikipedia.org/wiki/Triangular_matrix





Acknowledgments

None so far.





Colophon

These slides were prepared using the Cambria typeface. Mathematical equations use Times New Roman, and source code is presented using Consolas. Mathematical equations are prepared in MathType by Design Science, Inc. Examples may be formulated and checked using Maple by Maplesoft, Inc.

The photographs of flowers and a monarch butter appearing on the title slide and accenting the top of each other slide were taken at the Royal Botanical Gardens in October of 2017 by Douglas Wilhelm Harder. Please see

<https://www.rbg.ca/>

for more information.





Disclaimer

These slides are provided for the NE 112 *Linear algebra for nanotechnology engineering* course taught at the University of Waterloo. The material in it reflects the authors' best judgment in light of the information available to them at the time of preparation. Any reliance on these course slides by any party for any other purpose are the responsibility of such parties. The authors accept no responsibility for damages, if any, suffered by any party as a result of decisions made or actions based on these course slides for any other purpose than that for which it was intended.

