Lecture 6 : Solving Linear Systems

Math 3013
Oklahoma State University

January 24, 2022

Lecture 6 : Solving Linear Systems

Agenda:

1. Review: Matrix Operations
2. Linear Systems
3. Interpreting Solution Sets Geometrically
4. The Row Reduction Method

Matrix Multiplication

Suppose **A** and **B** are matrices.
• If $\#Columns\,(\mathbf{A}) \neq \#Rows\,(\mathbf{B})$, **AB** is **not defined**.
• If $\#Columns\,(\mathbf{A}) = \#Rows\,(\mathbf{B})$, **AB** is **is defined** and its entries are given by

$$(\mathbf{AB})_{ij} = Row_i\,(\mathbf{A}) \cdot Col_j\,(\mathbf{B})$$

• the matrix dimensions follow the rule

$$(n \times m) \cdot (m \times p) \longrightarrow (n \times p)$$

The Identity Matrix, Cont'd

Definition

The $n \times n$ **Identity matrix** is the the $n \times n$ matrix $\mathbf{I}$ whose entries are given by

$$I_{ij} = \begin{cases} 1 & \text{if} \quad i = j \\ 0 & \text{if} \quad i \neq j \end{cases} \quad \iff \quad \mathbf{I} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \vdots \\ \vdots & & \ddots & \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

In other words, $\mathbf{I}$ is the $n \times n$ matrix with 1's along the diagonal (running from the upper left to the lower right) and 0's everywhere else.

The Identity Matrix, Cont'd

The identity matrix $\mathbf{I}$ has the property that

$$\begin{aligned} \mathbf{AI} &= \mathbf{A} \quad \text{whenever } \mathbf{AI} \text{ is defined} \\ \mathbf{IA} &= \mathbf{A} \quad \text{whenever } \mathbf{IA} \text{ is defined} \end{aligned}$$

Addition and Scalar Multiplication of Matrices

Definition

Suppose $\mathbf{A}$ and $\mathbf{B}$ are both $n \times m$ matrices. Then the **matrix sum** $\mathbf{A} + \mathbf{B}$ is defined as the $n \times m$ matrix with entries

$$(\mathbf{A} + \mathbf{B})_{ij} = A_{ij} + B_{ij}$$

Definition

Let $\lambda \in \mathbb{R}$, and let $\mathbf{A}$ be an $n \times m$ matrix. Then the **scalar product** of $\mathbf{A}$ by λ is the $n \times m$ matrix $\lambda\mathbf{A}$ with entries

$$(\lambda\mathbf{A})_{ij} = \lambda A_{ij}$$

The Transpose of a Matrix

Definition

Suppose $\mathbf{A}$ is an $n \times m$ matrix. The **transpose** of $\mathbf{A}$ is the $m \times n$ matrix $\mathbf{A}^t$ with entries

$$\left(\mathbf{A}^t\right)_{ij} = A_{ji} \quad ; \quad \text{for } i = 1, \ldots, m \quad j = 1, \ldots, n$$

(note how the row and column indices have been reversed). Equivalently, $\mathbf{A}^t$ is the matrix obtained from $\mathbf{A}$ by converting, in order, the rows of $\mathbf{A}$ into columns.

Example

If

$$\mathbf{A} = \begin{bmatrix} 1 & 3 \\ -2 & 1 \\ 3 & -1 \end{bmatrix}$$

then

$$\mathbf{A}^t = \begin{bmatrix} 1 & -2 & 3 \\ 3 & 1 & -1 \end{bmatrix}$$

Properties of the Transpose Operation

Theorem

Let $\mathbf{A}$ *and* $\mathbf{B}$ *be matrices and let* $\lambda \in \mathbb{R}$*.*

- $\left(\mathbf{A}^t\right)^t = \mathbf{A}$
- $(\lambda\mathbf{A})^t = \lambda\mathbf{A}^t$
- $(\mathbf{A}+\mathbf{B})^t = \mathbf{A}^t + \mathbf{B}^t$
- $(\mathbf{AB})^t = \mathbf{B}^t\mathbf{A}^t$ *(Note how the order of factors changed.)*

Back to Linear Systems

Recall matrices were introduced as a means of writing large systems as linear equations as a single matrix equation.

$$\left.\begin{array}{ccc} a_{11}x_1 + & \cdots & +a_{1m}x_m = b_1 \\ \vdots & & \vdots \\ a_{n1}x_1 + & \cdots & +a_{nm}x_m = b_n \end{array}\right\} \quad \Longleftrightarrow \quad \mathbf{Ax} = \mathbf{b}$$

where

$$\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

Linear Systems

Recall that a linear equation in m unknowns is an equation that can be written in the from

$$a_1 x_1 + a_2 x_2 + \cdots + a_m x_m = b$$

An $n \times m$ linear system is a set of n linear equations in m unknowns:

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m &= b_2 \\
&\vdots \\
a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m &= b_n
\end{aligned}$$

Today, we begin a discussion of how to solve large linear systems.

Understanding Solution Sets Geometrically

Case 0: 0 equations in m unknowns

In this case, there are no conditions placed on the variables

$x_1, \ldots, x_m$

And so each variable x_i is allowed to range over the entire real line.

The "solution vectors" are just

$$\left\{ \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \mid x_1, x_2, \ldots, x_m \in \mathbb{R} \right\} = \mathbb{R}^m$$

Case 1: 1 equation in m unknowns

We now impose a single linear equation

$$a_1x_1 + a_2x_2 + \cdots + a_mx_m = b$$

on m variables $x_1, \ldots, x_m$.
Assume $a_m \neq 0$ (so that the term a_mx_m actually contributes to the equation)
Then this equation is readily solved for x_m :

$$x_m = \frac{b}{a_m} - \frac{a_1}{a_m}x_1 - \cdots - \frac{a_{m-1}}{a_m}x_{m-1}$$

Thus, a solution vector will be a vector of the form

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{m-1} \\ \frac{b}{a_m} - \frac{a_1}{a_m}x_1 - \cdots - \frac{a_{m-1}}{a_m}x_{m-1} \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \frac{b}{a_m} \end{bmatrix} + \begin{bmatrix} x_1 \\ 0 \\ \vdots \\ 0 \\ -\frac{a_1}{a_m}x_1 \end{bmatrix} + \cdots + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ x_{m-1} \\ -\frac{a_{m-1}}{a_m}x_{m-1} \end{bmatrix}$$

(decomposing $\mathbf{X}$ into a sum of vectors where the individual vectors in the sum depend only one of the variables) or, pulling out the common factors x_i of each vector on the R.H.S. (right hand side),

$$\mathbf{x} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \frac{b}{a_m} \end{bmatrix} + x_1 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -\frac{a_1}{a_m} \end{bmatrix} + \cdots + x_{m-1} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ -\frac{a_{m-1}}{a_m} \end{bmatrix}$$

Note how in this last form, the solution vector is expressed as a point on a $(m-1)$ dimensional hyperplane of vectors; i.e., a set of vectors of the form

$$\{\mathbf{p}_0 + t_1\mathbf{v}_1 + \cdots + t_{m-1}\mathbf{v}_{m-1} \mid t_1, \ldots, t_{m-1} \in \mathbb{R}\} \subset \mathbb{R}^m$$

Conclusion: the solution set of a single linear equation in m unknowns is an $(m-1)$-dimensional hyperplane in the vector space $\mathbf{R}^m$.

Example: Systems of Linear Equations in Three Variables.

In order to help visualize the solution set of linear systems, let's now consider what can happen for linear systems in 3 variables. Consider the following set of equations in three variables x, y, and z :

$$x + y + z = 2 \tag{Eq1}$$

$$x - z = 3 \tag{Eq2}$$

$$2x + 2y + 2z = 2 \tag{Eq3}$$

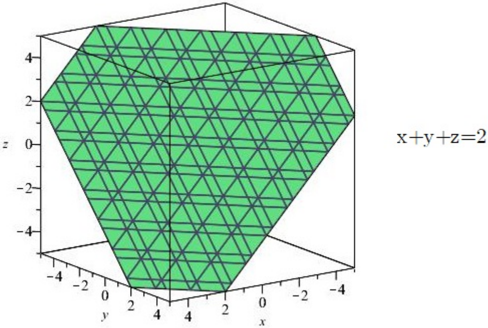
$$x + y + z = -1 \tag{Eq4}$$

$$y - z = -1 \tag{Eq5}$$

$$x + 2y = -2 \tag{Eq6}$$

Each of these equations will correspond to a particular 2-dimensional plane in $\mathbb{R}^3$.

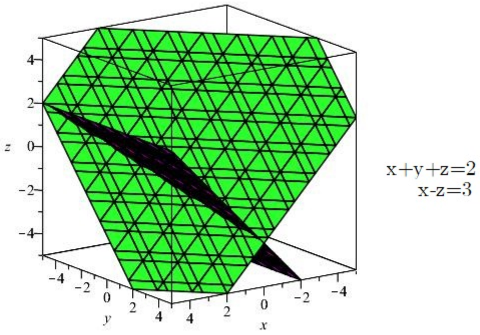
Below is a plot of the solution set of Eq1: $x + y + z = 2$



Note how this appears as a 2-dimensional plane in the vector space $\mathbb{R}^3$

2 Linear Equations in 3 Variables: the usual situation

Below we have plotted the solution set of equations Eq1 and Eq2.

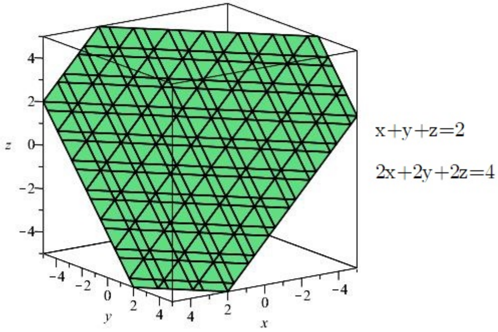


Note how the points common to both solution planes is a line in $\mathbb{R}^3$

Moral: Generally speaking, **the imposition of an additional equation reduces the dimension of the hyperplane solution space by 1**

2 Linear Equations in 3 Variables: redundant equations

Below we have plotted the solution set of equations Eq1 and Eq3.

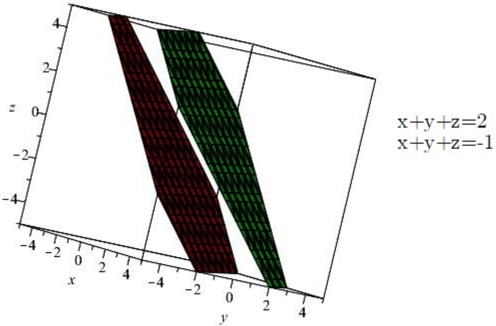


Note how we have the same planar solution set as that of Eq1. This is because the condition implied by Eq3 is already implied Eq1.

Moral: Additional equations do not always reduce the dimension of the solution sets

2 Linear Equations in 3 Variables: contradictory equations

Below we have plotted the solution set of equations Eq1 and Eq4.



Note the two solution planes do not intersect at all.
This is because if both equations were true, $2 = -1 \longrightarrow$ a mathematical contradiction.

Moral: Additional equations can lead to no solutions

Summary: The Geometry of Planar Intersections in $\mathbb{R}^3$

When 2 solution planes intersect, either

- the intersection is a line (general situation)
- the intersection is a plane (redundant equations)
- the intersection is empty (contradictory equations)

More generally, for large systems of equations

An additional equation causes either

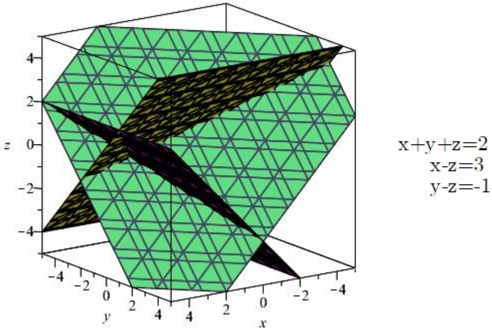
- the dimension of the hyperplane solution set to be reduced by 1 (the general situation)
- the dimension of the hyperplane solution is unchanged (additional equation is redundant to conditions already imposed)
- the solution set is made empty (addition equation contradicts previous equations)

3 linear equations in 3 unknowns: Case 1

The general situation:

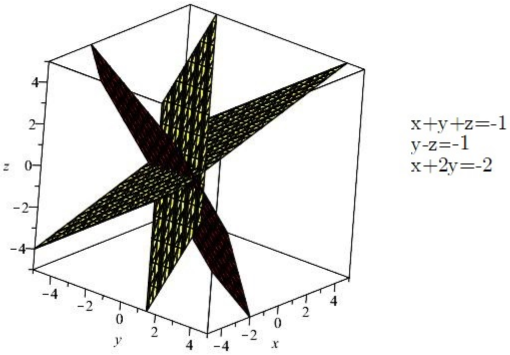
$$\#\text{ unknowns} - \#\text{ equations} = 0 \text{ free variables}$$

and so we should expect a 0-dimensional solution hyperplane (i.e. a unique vector solution)
To illustrate this case, I have plotted the solution set of equations Eq1, Eq2, and Eq5 below



Case 2: Redundant Equations

If an additional equation is already implied by previous equations, then the solution set is unchanged.
Below we have plotted the solution set of equations Eq4, Eq5 and Eq6.

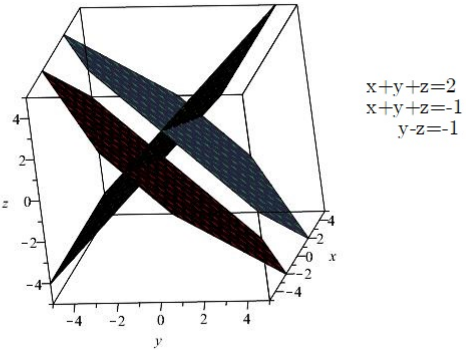


Note how the last equation is actually the sum of the first two equations. So the last equation does not put any new condition on

Case 3: Contradictory Equations

If an additional equation contradicts previous equations, then the solution set is empty.

Below we have plotted the solution set of equations Eq4, Eq5 and Eq6.



Note how the last equation is actually the sum of the first two equations. So the last equation does not put any new condition on

Generalization and Summary: the solution sets of $n \times m$ linear systems

The solution set of n linear equations in m unknowns is either

- A hyperplane in $\mathbb{R}^m$ of dimension $m - n$ (if all equations are independent)
- A hyperplane in $\mathbb{R}^m$ of dimension $> m - n$ (if there are some redundant equations)
- The empty set (if there are contradictory equations)

Elementary Operations

Elementary Operations are things we can do to equations that do not change their solution:
Consider the following 2 linear equations

$$x + y = 2 \tag{Eq1}$$

$$x - y = 1 \tag{Eq2}$$

Then

(i) we can change the order of equations

$$Eq1 \leftrightarrow Eq2 : \qquad \left\{ \begin{array}{ccc} x - y & = & 1 \\ x + y & = & 2 \end{array} \right\} \text{ has the same solution set}$$

Elementary Operations, Cont'd

(ii) we can replace an equation by a non-zero scalar multiple of itself

$$Eq2 \to 2{*}Eq2: \qquad \left\{ \begin{array}{ccc} x+y & = & 2 \\ 2x-2y & = & 2 \end{array} \right\} \text{ has the same solution set}$$

(iii) we can replace an equation by its sum with another equation

$$Eq2 \to Eq2{+}Eq1: \qquad \left\{ \begin{array}{ccc} x+y & = & 2 \\ 2x+0 & = & 3 \end{array} \right\} \text{ has the same solution set}$$

Solving Linear Systems by Using Elementary Operations

Below is an example of how to find a solution to a linear system by manipulating the equations (instead of solving for variables and then substituting for variables).

$$\left\{ \begin{array}{c} x + y = 1 \\ x - y = 3 \end{array} \right\} \xrightarrow{Eq2 \rightarrow Eq2 + Eq1} \left\{ \begin{array}{c} x + y = 1 \\ 2x = 4 \end{array} \right\}$$

$$\xrightarrow{Eq2 \rightarrow \frac{1}{2} Eq2} \left\{ \begin{array}{c} x + y = 1 \\ x = 2 \end{array} \right\}$$

$$\xrightarrow{Eq1 \rightarrow Eq1 - Eq2} \left\{ \begin{array}{c} y = -1 \\ x = 2 \end{array} \right\}$$

$$\xrightarrow{Eq1 \longleftrightarrow Eq2} \left\{ \begin{array}{c} x = 2 \\ y = -1 \end{array} \right\} \quad \text{(the solution)}$$

The Matrix Method for Solving Linear Systems

Recall that the data that goes into specifying an $n \times m$ linear system

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

is an $n \times m$ matrix $\mathbf{A}$, an $m \times 1$ matrix of variables $\mathbf{x}$, and a $n \times 1$ matrix of numbers $\mathbf{b}$.

$$\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \ldots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix} \quad , \quad \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} \quad , \quad \mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

In fact, the matrix equation $\mathbf{Ax} = \mathbf{b}$ is equivalent to the original linear system.

Augmented Matrices

Definition

The **augmented matrix** of an $n \times m$ linear system $\mathbf{Ax} = \mathbf{b}$ is the $n \times (m+1)$ matrix $[\mathbf{A} \mid \mathbf{b}]$ formed by adjoining the column vector $\mathbf{b}$ to the $n \times m$ matrix $\mathbf{A}$

$$[\mathbf{A} \mid \mathbf{b}] = \left[\begin{array}{ccc|c} a_{11} & \cdots & a_{1m} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nm} & b_n \end{array}\right]$$

Augmented Matrices and Equations: Example

The augmented matrix of the 3×3 linear system

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

is

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

The Idea To Be Pursued

Schematically,

$$\begin{aligned} \text{linear system} \quad &\rightarrow \quad \text{augmented matrix } [\mathbf{A}|\mathbf{b}] \\ [\mathbf{A} \mid \mathbf{b}] \quad &\rightarrow \quad \left[\mathbf{A}' \mid \mathbf{b}'\right] \text{ , the augmented matrix of the solution} \\ \left[\mathbf{A}'|\mathbf{b}'\right] \quad &\rightarrow \quad \text{equations of solution} \end{aligned}$$

In the second step we will be using operations on augmented matrices that correspond to operations on equations that don't change their solution.