Lecture 4 : Linear Systems and Matrices

Math 3013
Oklahoma State University

January 19, 2022

Lecture 4 : Linear Systems and Matrices

Agenda:

1. WebAssign Demo
2. Hyperplanes
3. Linear Systems
4. Matrices
5. Matrix Multiplication

Hyperplanes

Recall that a hyperplane is an (in general, infinite) set of vectors of the form

$$H = \{\mathbf{p}_0 + t_1\mathbf{d}_1 + t_2\mathbf{d}_2 + \cdots + t_k\mathbf{d}_k \mid t_1, t_{2,} \ldots, t_k \in \mathbb{R}\}$$

Here $\mathbf{p}_0$ is a vector representing an "initial point" in the hyperplane, and the vectors $\mathbf{d}_1, \ldots, \mathbf{d}_k$ is a set of directions used to fill out the rest of the hyperplane starting from $\mathbf{p}_0$.

Today, we'll see how such sets of vectors arise naturally, as the solution sets of systems of linear equations.

Systems of Linear Equations and Matrices

Consider one linear equation in two unknowns

$$x + 2y = 1$$

Solving for the variable y, we see the solution set is

$$\begin{aligned}
\text{Solution Set} &= \left\{ [x, y] \in \mathbb{R}^2 \mid y = \frac{1}{2}(1 - x) \right\} \\
&= \left\{ \left[x, \frac{1}{2}(1 - x) \right] \mid x \in \mathbb{R} \right\} \\
&= \left\{ \left[t, \frac{1}{2}(1 - t) \right] \mid t \in \mathbb{R} \right\} \\
&= \left\{ \left[0, \frac{1}{2} \right] + t \left[1, -\frac{1}{2} \right] \mid t \in \mathbb{R} \right\}
\end{aligned}$$

i.e., the solution set is the line that starting at $\mathbf{p}_0 = \left[0, \frac{1}{2}\right]$ heading in the direction $\mathbf{d} = \left[1, -\frac{1}{2}\right]$.

Two equations in two unknowns

$$\begin{aligned} x - 2y &= 0 \\ x + y &= 3 \end{aligned}$$

First equation $\Rightarrow \quad x = 2y$
Substituting into Second
equation $\Rightarrow \quad 2y + y = 3 \quad \Rightarrow \quad 3y = 3 \quad \Rightarrow \quad y = 1$
Back-substituting into first equation $\Rightarrow \quad x = 2y = (2)(1) = 2$
Answer: $x = 2$, $y = 1$.
And so the solution set is a single vector

$$\text{Solution Set} = \{[2, 1]\}$$

Linear Systems

More generally,

Definition

An **n by m linear system** is a set of n linear equations in m unknowns; that is to say, a set of equations of the form

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

The numbers a_{ij} are referred to as the **coefficients**; more precisely, the number a_{ij} is the coefficient of the variable x_j in the i^{th} equation.

Specification of a linear system is thus corresponds to a list of $n \times m$ **coefficients** a_{ij}, a list of m **variables** x_j, and a list of n numbers b_i which we'll call **inhomogeneous terms**.

Solution Sets for Linear Systems

Theorem

The solution set of a $n \times m$ linear system is either

- *the empty set $\{\}$, or*
- *a hyperplane inside the space of variables. I.e., a set of vectors of the form*

$$S = \{\mathbf{p}_0 + t_1\mathbf{d}_1 + \cdots t_k\mathbf{d}_k \mid t_1, \ldots t_k \in \mathbb{R}\} \subset \mathbb{R}^m$$

(Note: here I am including a single point as a special case of a general hyperplane)

Organizing the Data Needed to Specify a Linear System

Consider an $n \times m$ linear system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{m2}x_2 + \cdots + a_{nm}x_m &= b_n \end{aligned}$$

with mn coefficients $a_{11}, \ldots, a_{nm}$, m variables $x_1, \ldots, x_m$, and n numbers $b_1, \ldots, b_n$

In Linear Algegbra we organize this equation data into "matrices" as follows:

Matrices

Definition

An **n by m matrix** is an arrangement of *nm* numbers a_{ij}, into an ordered rectangular array with *n* rows and *m* columns , such as

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}$$

Special Cases

An **n-dimensional column vector** is a n by 1 matrix, such as

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

An **m-dimensional row vector** is a 1 by m matrix, such as

$$\mathbf{c} = \begin{bmatrix} c_1 & c_2 & \cdots & c_m \end{bmatrix}$$

Row Vectors and Column Vectors, Cont'd

Note: Each of these two special types of matrices corresponds to an ordered list of numbers. (for column vectors, the numbers are arranged and ordered vertically; while, for row vectors, the numbers are arranged and ordered horizontally).
Thus, for a given vector $[a_1, \ldots, a_n]$ (i.e., a given ordered list of numbers), we can associate two different matrices:

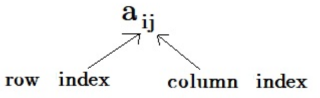
$$\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \qquad \text{a column vector (i.e., a } n \times 1 \text{ matrix}$$

and

$$\begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} \qquad \text{, a row vector (i.e., a } 1 \times n \text{ matrix)}$$

An Important Notational Convention

When specifying the individual entries of a matrix we'll use a index notation whereby the position of the element within its matrix will be specified by two subscripts:
Thus, for a given matrix **A**, we'll tend to denote the entry in the i^{th} row and j^{th} column of the matrix by something like a_{ij} ,
We'll refer to the first subscript i as the element's **row index** and the second subscript j will be the element's **column index**
By convention, the first row index (i) **always** precedes the column index (j), when one specifies a matrix entry a_{ij}.



Using Matrices to Specify an $n \times m$ Linear System

- Let $\mathbf{x}$ be an m-dimensional column vector of unknowns

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$$

- Let $\mathbf{A}$ be an n by m matrix of coefficients

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ldots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}$$

- Let $\mathbf{b}$ be an n-dimensional column vector of values

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

This notation allows us to write (formally at this point), the linear system

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_m & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_m & = & b_2 \\ \vdots & & \vdots \\ a_{n1}x_1 + a_{m2}x_2 + \cdots + a_{nm}x_n & = & b_n \end{array}$$

as

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

or, more succinctly, as

$$\mathbf{Ax} = \mathbf{b}$$

Example

Find the matrix $\mathbf{A}$ and the column vector $\mathbf{b}$ corresponding to the following linear system

$$\begin{aligned} x_1 + 2x_3 + x_4 &= 1 \\ 2x_1 - x_2 &= 3 \\ x_2 + x_3 - x_4 &= -2 \end{aligned}$$

The coefficient matrix is

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 2 & 1 \\ 2 & -1 & 0 & 0 \\ 0 & 1 & 1 & -1 \end{bmatrix}$$

and the (vector) inhomogeneous term is

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

So the "matrix equation" equivalent to the 4×3 linear system

$$\begin{aligned} x_1 + 2x_3 + x_4 &= 1 \\ 2x_1 - x_2 &= 3 \\ x_2 + x_3 - x_4 &= -2 \end{aligned}$$

is

$$\begin{bmatrix} 1 & 0 & 2 & 1 \\ 2 & -1 & 0 & 0 \\ 0 & 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix}$$

Matrix Multiplication

We have seen that an $n \times m$ linear system

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

corresponds to a "matrix equation" of the form

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

or, symbolically,

$$\mathbf{Ax} = \mathbf{b}$$

What I'll show next is how to interpret the left hand side of (1) as the "product" of the coefficient matrix $\mathbf{A}$ with the column vector $\mathbf{x}$ of variables.

Case 1: a row vector times a column vector

Definition

Let **a** be an n-dimensional row vector and let **b** be an n-dimensional column vector. Then the **matrix product ab** is the dot product of the vector **a** and the vector **b**

$$\begin{aligned}\mathbf{ab} &= \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \\ &= a_1 b_1 + a_2 b_2 + \cdots + a_n b_n\end{aligned}$$

Note that the right hand side is just the dot product of **a** and **b** thought of as ordered lists of n numbers.

Case 2: a matrix times a column vector

Let

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix} \qquad , \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}$$

Then the **matrix product Ax** is the n-dimensional column vector with components

$$\mathbf{Ax} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m \end{bmatrix}$$

Note how the i^{th} entry of the result $\mathbf{Ax}$ is the dot product of the vector (i.e., ordered list) corresponding to the i^{th} row of $\mathbf{A}$ and the vector corresponding to $\mathbf{x}$

$$\begin{aligned}
(\mathbf{Ax})_i &= Row_i(\mathbf{A}) \cdot \mathbf{x} \\
&= a_{i1}x_1 + \cdots + a_{im}x_m \\
&= \sum_{j=1}^{m} a_{ij}x_j
\end{aligned}$$

Example

Compute the matrix product $\mathbf{Ax}$ where

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 2 & 1 \\ 2 & -1 & 0 & 0 \\ 0 & 1 & 1 & -1 \end{bmatrix} \quad , \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

We have

$$\mathbf{Ax} = \begin{bmatrix} 1 & 0 & 2 & 1 \\ 2 & -1 & 0 & 0 \\ 0 & 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} (1,0,2,1)\cdot(x_1,x_2,x_3,x_4) \\ (2,-1,0,0)\cdot(x_1,x_2,x_3,x_4) \\ (0,1,1,-1)\cdot(x_1,x_2,x_3,x_4) \end{bmatrix}$$

$$= \begin{bmatrix} x_1 + 2x_3 + x_4 \\ 2x_1 - x_2 \\ x_2 + x_3 - x_4 \end{bmatrix}$$

Note that the matrix equation

$$\begin{bmatrix} 1 & 0 & 2 & 1 \\ 2 & -1 & 0 & 0 \\ 0 & 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix} \qquad (*)$$

is thus equivalent to the following linear system

$$\begin{aligned} x_1 + 2x_3 + x_4 &= 1 \\ 2x_2 - x_2 &= 3 \\ x_2 + x_3 - x_4 &= -2 \end{aligned}$$

once the product on the left side of (*) is computed and we compare both sides component-by-component.

Matrix Multiplication: the general definition

Definition

Let $\mathbf{A}$ be an n by m matrix and let $\mathbf{B}$ be an s by t matrix.

(i) If $m \neq s$ the matrix product $\mathbf{AB}$ is **undefined** (i.e. if the number of columns of $\mathbf{A}$ does not equal the number of rows of $\mathbf{B}$, the matrix product does not exist).

(ii) If $m = s$, then the matrix product $\mathbf{AB}$ **is defined** and it is the n by t matrix whose entries $(\mathbf{AB})_{ij}$ are prescribed by

$$\begin{aligned}(\mathbf{AB})_{ij} &= Row_i(\mathbf{A}) \cdot Col_j(\mathbf{B}) \\ &= a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{im}b_{mj} \\ &= \sum_{k=1}^{m} a_{ik}b_{kj}\end{aligned}$$

The Basic Rules of Matrix Multiplication

Let $\mathbf{A}$ be an $n \times m$ matrix and let $\mathbf{B}$ be a $m \times p$ matrix.

- Then the matrix product $\mathbf{AB}$ is well-defined and it will be the $n \times p$ matrix with entries

$$
\begin{aligned}
(\mathbf{AB})_{ij} &= Row_i(\mathbf{A}) \cdot Col_j(\mathbf{B}) \\
&= \sum_{k=1}^{m} a_{ik} b_{kj} \\
&= a_{i1} b_{1j} + \cdots + a_{im} b_{mj}
\end{aligned}
$$

- The matrix dimensions follow the rule

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

A simple symbolic example

Let $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and let $\mathbf{B} = \begin{bmatrix} e & f \\ g & h \end{bmatrix}$. The matrix product $\mathbf{AB}$ is defined; since the number of rows in $\mathbf{A}$ is the same as the number of columns of $\mathbf{B}$.

The product matrix $\mathbf{AB}$ will be a 2×2 matrix with entries

$$\begin{aligned}
(\mathbf{AB})_{11} &= Row_1(\mathbf{A}) \cdot Col_1(\mathbf{B}) = ae + bg \\
(\mathbf{AB})_{12} &= Row_1(\mathbf{A}) \cdot Col_2(\mathbf{B}) = af + bh \\
(\mathbf{AB})_{21} &= Row_2(\mathbf{A}) \cdot Col_1(\mathbf{B}) = ce + dg \\
(\mathbf{AB})_{22} &= Row_2(\mathbf{A}) \cdot Col_2(\mathbf{B}) = cf + dh
\end{aligned}$$

Thus,

$$\mathbf{AB} = \begin{bmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{bmatrix}$$