Pre-lecture: the goal for today

We wish to write linear systems simply as $A\boldsymbol{x} = \boldsymbol{b}$.

For instance:

$$\begin{array}{rcl} 2x_1 + 3x_2 &=& b_1 \\ 3x_1 + x_2 &=& b_2 \end{array} \iff \begin{bmatrix} 2 & 3 \\ 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

Why?

- It's concise.
- The compactness also sparks associations and ideas!
  - For instance, can we solve by *dividing* by A? $\boldsymbol{x} = A^{-1}\boldsymbol{b}$?
  - If $A\boldsymbol{x} = \boldsymbol{b}$ and $A\boldsymbol{y} = 0$, then $A(\boldsymbol{x} + \boldsymbol{y}) = \boldsymbol{b}$.
- Leads to matrix calculus and deeper understanding.
  - multiplying, inverting, or factoring matrices

Matrix operations

Basic notation

We will use the following notations for an $m \times n$ matrix A (m rows, n columns).

- In terms of the columns of A:

$$A = [\ \boldsymbol{a}_1\ \ \boldsymbol{a}_2\ \ \cdots\ \ \boldsymbol{a}_n\ ] = \begin{bmatrix} | & | & & | \\ \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n \\ | & | & & | \end{bmatrix}$$

- In terms of the entries of A:

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix}, \qquad a_{i,j} = \text{entry in } i\text{-th row, } j\text{-th column}$$

Matrices, just like vectors, are added and scaled componentwise.

Example 1.

(a) $\begin{bmatrix} 1 & 0 \\ 5 & 2 \end{bmatrix} + \begin{bmatrix} 2 & 3 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 8 & 3 \end{bmatrix}$

Armin Straub
astraub@illinois.edu

(b) $7\cdot\begin{bmatrix} 2 & 3 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 14 & 21 \\ 21 & 7 \end{bmatrix}$

Matrix times vector

Recall that $(x_1, x_2, \ldots, x_n)$ solves the linear system with augmented matrix

$$[\ A \quad \boldsymbol{b}\ ] = \begin{bmatrix} | & | & & | & | \\ \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n & \boldsymbol{b} \\ | & | & & | & | \end{bmatrix}$$

if and only if

$$x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \ldots + x_n\boldsymbol{a}_n = \boldsymbol{b}.$$

It is therefore natural to define the **product of matrix times vector** as

$$A\boldsymbol{x} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \ldots + x_n\boldsymbol{a}_n, \qquad \boldsymbol{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

The system of linear equations with augmented matrix $[\ A \quad \boldsymbol{b}\ ]$ can be written in **matrix form** compactly as $A\boldsymbol{x} = \boldsymbol{b}$.

The product of a matrix A with a vector $\boldsymbol{x}$ is a linear combination of the columns of A with weights given by the entries of $\boldsymbol{x}$.

Example 2.

(a) $\begin{bmatrix} 1 & 0 \\ 5 & 2 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 1 \end{bmatrix} = 2\begin{bmatrix} 1 \\ 5 \end{bmatrix} + 1\begin{bmatrix} 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 12 \end{bmatrix}$

(b) $\begin{bmatrix} 2 & 3 \\ 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$

(c) $\begin{bmatrix} 2 & 3 \\ 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1\begin{bmatrix} 2 \\ 3 \end{bmatrix} + x_2\begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 2x_1 + 3x_2 \\ 3x_1 + x_2 \end{bmatrix}$

This illustrates that linear systems can be simply expressed as $A\boldsymbol{x} = \boldsymbol{b}$:

$$\begin{array}{rrcl} 2x_1 & +3x_2 & = & b_1 \\ 3x_1 & +x_2 & = & b_2 \end{array} \iff \begin{bmatrix} 2 & 3 \\ 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

(d) $\begin{bmatrix} 2 & 3 \\ 3 & 1 \\ 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 4 \\ 0 \end{bmatrix}$

Armin Straub
astraub@illinois.edu

Example 3. Suppose A is $m \times n$ and $\boldsymbol{x}$ is in $\mathbb{R}^p$. Under which condition does $A\boldsymbol{x}$ make sense?

We need $n = p$. (Go through the definition of $A\boldsymbol{x}$ to make sure you see why!)

Matrix times matrix

If B has just one column $\boldsymbol{b}$, i.e. $B = [\ \boldsymbol{b}\ ]$, then $AB = [\ A\boldsymbol{b}\ ]$.

In general, the **product of matrix times matrix** is given by

$$AB = [\ A\boldsymbol{b}_1 \quad A\boldsymbol{b}_2 \quad \cdots \quad A\boldsymbol{b}_p\ ], \qquad B = [\ \boldsymbol{b}_1 \quad \boldsymbol{b}_2 \quad \cdots \quad \boldsymbol{b}_p\ ].$$

Example 4.

(a) $\begin{bmatrix} 1 & 0 \\ 5 & 2 \end{bmatrix} \cdot \begin{bmatrix} 2 & -3 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & -3 \\ 12 & -11 \end{bmatrix}$

because $\begin{bmatrix} 1 & 0 \\ 5 & 2 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 1 \end{bmatrix} = 2\begin{bmatrix} 1 \\ 5 \end{bmatrix} + 1\begin{bmatrix} 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 12 \end{bmatrix}$

and $\begin{bmatrix} 1 & 0 \\ 5 & 2 \end{bmatrix} \cdot \begin{bmatrix} -3 \\ 2 \end{bmatrix} = -3\begin{bmatrix} 1 \\ 5 \end{bmatrix} + 2\begin{bmatrix} 0 \\ 2 \end{bmatrix} = \begin{bmatrix} -3 \\ -11 \end{bmatrix}$.

(b) $\begin{bmatrix} 1 & 0 \\ 5 & 2 \end{bmatrix} \cdot \begin{bmatrix} 2 & -3 & 1 \\ 1 & 2 & 0 \end{bmatrix} = \begin{bmatrix} 2 & -3 & 1 \\ 12 & -11 & 5 \end{bmatrix}$

> Each column of AB is a linear combination of the columns of A with weights given by the corresponding column of B.

Remark 5. The definition of the matrix product is inevitable from the multiplication of matrix times vector and the fact that we want AB to be defined such that $(AB)\boldsymbol{x} = A(B\boldsymbol{x})$.

$$\begin{aligned} A(B\boldsymbol{x}) &= A(x_1\boldsymbol{b}_1 + x_2\boldsymbol{b}_2 + \cdots) \\ &= x_1 A\boldsymbol{b}_1 + x_2 A\boldsymbol{b}_2 + \cdots \\ &= (AB)\boldsymbol{x} \quad \text{if the columns of } AB \text{ are } A\boldsymbol{b}_1, A\boldsymbol{b}_2, \ldots \end{aligned}$$

Example 6. Suppose A is $m \times n$ and B is $p \times q$.

(a) Under which condition does AB make sense?

We need $n = p$. (Go through the boxed characterization of AB to make sure you see why!)

(b) What are the dimensions of AB in that case?

AB is a $m \times q$ matrix.

Basic properties

Example 7.

(a) $\begin{bmatrix} 2 & 3 \\ 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 3 & 1 \end{bmatrix}$

(b) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 2 & 3 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 3 & 1 \end{bmatrix}$

This is the 2×2 **identity matrix**.

Theorem 8. Let A, B, C be matrices of appropriate size. Then:

- $A(BC) = (AB)C$ associative
- $A(B + C) = AB + AC$ left-distributive
- $(A + B)C = AC + BC$ right-distributive

Example 9. However, matrix multiplication is not commutative!

(a) $\begin{bmatrix} 2 & 3 \\ 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 5 \\ 3 & 4 \end{bmatrix}$

(b) $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 2 & 3 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 4 \\ 3 & 1 \end{bmatrix}$

Example 10. Also, a product can be zero even though none of the factors is:

$$\begin{bmatrix} 2 & 0 \\ 3 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Transpose of a matrix

Definition 11. The **transpose** A^T of a matrix A is the matrix whose columns are formed from the corresponding rows of A. rows $\leftrightarrow$ columns

Example 12.

(a) $\begin{bmatrix} 2 & 0 \\ 3 & 1 \\ -1 & 4 \end{bmatrix}^T = \begin{bmatrix} 2 & 3 & -1 \\ 0 & 1 & 4 \end{bmatrix}$

(b) $[\, x_1 \;\; x_2 \;\; x_3 \,]^T = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

(c) $\begin{bmatrix} 2 & 3 \\ 3 & 1 \end{bmatrix}^T = \begin{bmatrix} 2 & 3 \\ 3 & 1 \end{bmatrix}$

A matrix A is called **symmetric** if $A = A^T$.

Practice problems

- True or false?
  - AB has as many columns as B.
  - AB has as many rows as B.

The following practice problem illustrates the rule $(AB)^T = B^T A^T$.

Example 13. Consider the matrices

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ -2 & 4 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix}.$$

Compute:

(a) $AB = \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix} =$

(b) $(AB)^T = \begin{bmatrix} & & \\ & & \end{bmatrix}$

(c) $B^T A^T = \begin{bmatrix} 1 & 3 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & -2 \\ 2 & 1 & 4 \end{bmatrix} =$

(d) $A^T B^T = \begin{bmatrix} 1 & 0 & -2 \\ 2 & 1 & 4 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 2 & 0 \end{bmatrix} =$

What's that fishy smell?