

Applied Linear Algebra

Textbook: Olver and Shakiban, Applied Linear Algebra, 2nd ed.

www-users.cse.umn.edu/~olver/ala.html

Instructor: Jodin Morey moreyjc@umn.edu

Introduction to the Course

- ♦ Recitation Instructor: Ian Ruohoniemi (ruoho021@umn.edu).
- ♦ Office Hours - Jodin: Wed 12:45 - 2:15p, on Zoom. Fri 3:30 - 5p, in my office (Vincent Hall 204)
Ian: TBD.
- ♦ Tentative schedule on Canvas
- ♦ This course assumes knowledge of **introductory linear algebra**

Class Details

Quizzes/Exams/Group-work: Tuesdays.

Grading:

Homework: Weekly, but optional. It is *extremely important* you do the homework. Indeed, almost nothing is more important than the amount of effort you put into the homework. Don't fool yourself into thinking you can afford to skip an assignment.

Use discussion board on Canvas to discuss homework (just don't give solutions).

Quizzes/Exams: I (and your recitation instructor) will be posting scores for quizzes, midterms, and the final exam to Canvas (and Gradescope when applicable). Any Gradescope Quiz/Exam regrade requests must be requested on Gradescope before 11:59 PM, three weeks after it was taken. If you ever have a question about your grade, you can always email me or visit me during office hours.

Learning: You will absorb much more information during lecture and discussion section if you have already oriented yourself to the concepts by **reading the section of the book ahead of time**.

Also, students have been shown to learn best when digesting smaller chunks of information separated by periods of sleep. Subconscious processes occur during sleep and inbetween learning sessions that put the new knowledge into our long-term memory. So, **spread your learning out over the week/semester, and get good sleep**.



Class Notes: My class notes are downloadable as a PDF from the class webpage (even this one).

How Can You be Successful in Calculus?

Obvious Things:

- Read the Book
- Do the Homework
- Attend Lecture
- Attend Discussion
- Study for Tests

Not So Obvious Things:

- Read the Book **BEFORE** Lecture
- Start/Join a Peer Study Group
- Seek Help Early
- Trouble with a Section of the Book?
Find resources/videos on the topic which explain the topic better.
- Do ALL the Homework (and maybe more problems)

1.1 - 1.3 Solutions to Linear Systems, Matrices, Vectors, Gaussian Elimination

Recall that if you have a system of equations, w/the same number of equations as unknowns,
you can use Gaussian elimination to find sols (if they exist).

Say you have the system:

$$\begin{aligned} a_{11}x + a_{12}y + a_{13}z &= b_1 \\ a_{21}x + a_{22}y + a_{23}z &= b_2 \\ a_{31}x + a_{32}y + a_{33}z &= b_3 \\ a_{41}x + a_{42}y + a_{43}z &= b_4 \end{aligned}$$

For example:

$$\begin{aligned} 2x + 5y + 8z &= 0 \\ 2x + 6y + 9z &= 2 \\ 3x + 7y + 11z &= 4 \\ 6x + 14y + 22z &= 8 \end{aligned}$$

Put it into an (augmented) matrix:



Continuing our example:

$$\left[\begin{array}{ccc|c} 2 & 5 & 8 & 0 \\ 2 & 6 & 9 & 2 \\ 3 & 7 & 11 & 4 \\ 6 & 14 & 22 & 8 \end{array} \right]$$

Types of matrices:

$$a \text{ or } [a], \quad \begin{bmatrix} a_1 & a_2 \end{bmatrix}, \quad \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}, \quad \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}.$$

The above matrices are referred to (respectively) as: a scalar (or 1-vector, or 1×1 matrix),
a 2-vector (or row vector, or 1×2 matrix), another 2-vector (or column vector, or 2×1 matrix),
a 2×2 matrix, and a 2×3 matrix.

The a_i are referred to as components, entries, or elements of the matrix.

Matrix Operations

Notating matrices as **A**, **B**, etc., we can treat them as mathematical objects. (?!?)

We can define operations like addition, subtraction, multiplication.

First let's define what it means for **A** to be equal to **B**:

- they're the same size, and
- their elements are equal.

$$\begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix} \neq \begin{bmatrix} 1 & 9 \\ 4 & 5 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \neq \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}.$$

Other operations

Addition: $\mathbf{A} + \mathbf{B} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} a+e & b+f \\ c+g & d+h \end{bmatrix}.$ Requires both matrices be the same size.

Scalar Multiplication: $k\mathbf{A} = k \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} ka & kb \\ kc & kd \end{bmatrix}.$

Negative of a matrix: $-\mathbf{A} = -1 \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} -a & -b \\ -c & -d \end{bmatrix}.$

And we can define subtraction as a kind of addition: $\mathbf{A} - \mathbf{B} = \mathbf{A} + (-\mathbf{B}).$

Matrix Multiplication

Matrix multiplication is less intuitive. We wish to define it such that systems like:

$$\begin{aligned} 2x + 5y + 8z &= 0 \\ 2x + 6y + 9z &= 2 \quad , \\ 3x + 7y + 11z &= 4 \end{aligned}$$

with the coefficient matrix $\mathbf{A} = \begin{bmatrix} 2 & 5 & 8 \\ 2 & 6 & 9 \\ 3 & 7 & 11 \end{bmatrix}$, variable vector $\vec{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$,

and constant vector $\vec{b} = \begin{bmatrix} 0 \\ 2 \\ 4 \end{bmatrix}$, can be written as $\mathbf{A}\vec{x} = \vec{b}.$

Matrix multiplication is defined for $\mathbf{A}^{m \times n}$ and $\mathbf{B}^{n \times p}$, where the number of columns (n) in **A** is equal to the number of rows in **B**.

The result is $\mathbf{A}^{m \times n} \mathbf{B}^{n \times p} = \mathbf{C}^{m \times p}$, so the inner indices "cancel."

Example: $\mathbf{A}^{3 \times 3} \mathbf{x}^{3 \times 1} = \begin{bmatrix} 2 & 5 & 8 \\ 2 & 6 & 9 \\ 4 & 7 & 11 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2x + 5y + 8z \\ 2x + 6y + 9z \\ 4x + 7y + 11z \end{bmatrix} = \mathbf{C}^{3 \times 1}.$

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} =$$

See animated in class

And similarly:

$$\mathbf{a}^{1 \times 3} \mathbf{b}^{3 \times 1} = \begin{bmatrix} 1 & 5 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix} = 1 \cdot 1 + 5(-2) + 8 \cdot 3 = 15 =: \mathbf{c}^{1 \times 1}$$

$$= \sum_{k=1}^n a_k b_k \quad (\text{dot product of vectors}), \text{ and}$$

$$\mathbf{AB} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \cdot \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix} = [c_{ij}],$$

where $c_{ij} = \sum_{k=1}^n a_{ik} b_{kj}$, where n is the "inner" index.

When the size of the matrices are not compatible (as defined above), matrix multiplication is not defined.

Identity and Zero Matrices; and Algebraic Properties:

Let $\mathbf{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\mathbf{0} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. Which results in: $\mathbf{AI} = \mathbf{A}$, $\mathbf{A} + \mathbf{0} = \mathbf{A}$.

$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$ **Additive Commutativity**

$\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}$ **Additive Associativity**

$\mathbf{A(BC)} = (\mathbf{AB})\mathbf{C}$ **Multiplicative Associativity**

$\mathbf{A(B + C)} = \mathbf{AB} + \mathbf{AC}$, and **Distributive**
 $\mathbf{(A + B)C} = \mathbf{AC} + \mathbf{BC}$

However, AB is usually NOT equal to BA.
(i.e., matrices usually don't have "multiplicative commutivity")

If $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$, then notice that ...

$\mathbf{AB} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, however $\mathbf{BA} = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$. So, $\mathbf{AB} \neq \mathbf{BA}$.

Column $\begin{bmatrix} a \\ b \end{bmatrix}$ and row $\begin{bmatrix} e & f \end{bmatrix}$ vector notation.

$\begin{bmatrix} a \\ b \end{bmatrix} = (a,b) \neq \begin{bmatrix} a & b \end{bmatrix}$. The parenthetical notation for $\begin{bmatrix} a & b \end{bmatrix}$ is $(a,b)^T$.

Video Tutorial (visually rich/intuitive, although you may have to pause & rewatch parts till it sinks in):
<https://youtu.be/XkY2DOUCWMU>

Definition: A matrix is called *upper triangular* if all of the entries below the main diagonal are zero (or *lower triangular* if above are zero).

Example: $\begin{bmatrix} 1 & 5 & 8 \\ 0 & 0 & 9 \\ 0 & 0 & 11 \end{bmatrix}$ is upper triangular.

Definition: A matrix is called upper *unitriangular* if it is upper triangular **and** all of the main diagonal entries are 1 (or *lower unitriangular* if lower triangular and diagonals are 1)

Example: $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 8 & 9 & 1 \end{bmatrix}$ is lower unitriangular.

Gaussian Elimination

- Know how to do *elementary row operations* on matrices!
- #1. Add to a lower row, a constant multiple of a higher row.
 - #2. Interchange two rows (called pivoting).
 - #3. Multiply a row by a nonzero constant.



These are the same operations you did in prior math on systems of equations, but notated w/matrices.

Solve a system by putting its coefficients into a matrix, then transforming it into an **Echelon Form** matrix.

Given system: $a_{11}x + a_{12}y + a_{13}z = b_1,$
 $a_{21}x + a_{22}y + a_{23}z = b_2,$
 $a_{31}x + a_{32}y + a_{33}z = b_3.$

Rewrite as $[A|\vec{b}] = \left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & b_1 \\ a_{21} & a_{22} & a_{23} & b_2 \\ a_{31} & a_{32} & a_{33} & b_3 \end{array} \right]$. Then transform into:

$$\left[\begin{array}{ccc|c} a'_{11} & a'_{12} & a'_{13} & b'_1 \\ 0 & a'_{22} & a'_{23} & b'_2 \\ 0 & 0 & a'_{33} & b'_3 \end{array} \right] = [U|\vec{b}'],$$



where **U** is upper triangular.

The leading nonzero entries in each row ($a'_{11}, a'_{22}, a'_{33}$) are *pivots*, and the columns they're in are called *pivot columns*.

Definition: Transforming a matrix using only row operation #1 is called *regular Gaussian Elimination*.

Definition: If one restricts to only using regular Gaussian Elimination on matrix **A**, and they can transform it into upper triangular form **U** w/non-zero diagonals, **A** is a *regular matrix*.

Continuing an example from above: $\left[\begin{array}{ccc|c} 2 & 5 & 8 & 0 \\ 2 & 6 & 9 & 2 \\ 4 & 7 & 11 & 4 \end{array} \right]$

$$\begin{array}{c} R_2+(-R_1) \\ \Rightarrow \end{array} \left[\begin{array}{ccc|c} 2 & 5 & 8 & 0 \\ 0 & 1 & 1 & 2 \\ 4 & 7 & 11 & 4 \end{array} \right] \quad \begin{array}{c} R_3+(-2R_1) \\ \Rightarrow \end{array} \left[\begin{array}{ccc|c} 2 & 5 & 8 & 0 \\ 0 & 1 & 1 & 2 \\ 0 & -3 & -5 & 4 \end{array} \right]$$

$$\begin{array}{c} R_3+3R_2 \\ \Rightarrow \end{array} \left[\begin{array}{ccc|c} 2 & 5 & 8 & 0 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & -2 & 10 \end{array} \right] \text{ or } U\vec{x} = \vec{c}.$$

Non-zero diagonals, and we only used operation #1 (Adding to a lower row, a constant multiple of a higher row). So **A** is regular.



LU Factorization



Gaussian elimination is not the most efficient way to solve a system.

Toward this efficiency goal, let's add some tools to our toolbox.

Definition: Say you wanted to apply a particular row operation to a matrix $\mathbf{A}^{m \times n}$.

The *elementary matrix* $\mathbf{E}^{m \times m}$ is the matrix you could multiply $\mathbf{A}$ by to accomplish this row operation.

You can discover what $\mathbf{E}^{m \times m}$ is by applying the intended row operation on identity matrix $\mathbf{I}_m$.

Example: Given $\mathbf{A} := \begin{bmatrix} 1 & 2 & 1 \\ 2 & 6 & 1 \\ 1 & 1 & 4 \end{bmatrix}$. Say you want to add (-2) times the 1st row to the 2nd row?

Instead of the usual way, do the intended operation to the identity matrix: $\mathbf{E}_1 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

Now you see that the same operation can be done to $\mathbf{A}$ by matrix multiplication:

$$\mathbf{E}_1 \mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 6 & 1 \\ 1 & 1 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & -1 \\ 1 & 1 & 4 \end{bmatrix}.$$

Elementary matrices $\mathbf{E}$ corresponding to **type 1** operations take the form:

$$\mathbf{E}_1 = \begin{bmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{E}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ a & 0 & 1 \end{bmatrix}, \quad \text{or} \quad \mathbf{E}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & a & 1 \end{bmatrix}.$$

So, if $\mathbf{A}$ is regular then, there exists $\mathbf{E}_1, \dots, \mathbf{E}_n$ such that $\mathbf{E}\mathbf{A} := \mathbf{E}_1 \mathbf{E}_2 \dots \mathbf{E}_n \mathbf{A} = \mathbf{U}$.

To *undo* a type 1 operation, change the sign of a . So inverse operations $\mathbf{L}$ take the form:

$$\mathbf{L}_1 = \mathbf{E}_1^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -a & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{L}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -a & 0 & 1 \end{bmatrix}, \quad \mathbf{L}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -a & 1 \end{bmatrix}.$$

And using the inverses above, we can define $\mathbf{L} := \mathbf{L}_n \mathbf{L}_{n-1} \dots \mathbf{L}_1$.

And using multiplicative associativity, we find:

$$\mathbf{L}\mathbf{U} = (\mathbf{L}_n \dots \mathbf{L}_1)(\mathbf{E}_1 \dots \mathbf{E}_n \mathbf{A}) = (\mathbf{L}_n \dots \mathbf{L}_2)(\mathbf{L}_1 \mathbf{E}_1)(\mathbf{E}_2 \dots \mathbf{E}_n \mathbf{A}) = (\mathbf{L}_n \dots \mathbf{L}_2)(\mathbf{E}_2 \dots \mathbf{E}_n \mathbf{A}) = \dots = \mathbf{A}.$$

Theorem: $\mathbf{A}$ is regular *iff* it can be factored $\mathbf{A} = \mathbf{LU}$, where $\mathbf{U}$ is upper triangular with nonzero diagonal entries, and $\mathbf{L}$ is a lower unitriangular matrix, having all 1's on the diagonal.

The nonzero off diagonal entries l_{ij} for $i > j$ appearing in $\mathbf{L}$ are the elimination multipliers.

To reduce $\mathbf{A}$ to $\mathbf{U}$, perform the **elementary row operations** of subtracting $l_{ij} \times (\text{Row } j)$ from Row i .

But why would we want to do this? Why not use Gaussian elimination?

Turns out that w/large matrices (the types you might run into when doing mathematics in the real world), using LU factorization is computationally more efficient (takes less time on a computer).



Forward and Back Substitution (the efficiency)

Observe that. If $\mathbf{A}\vec{x} = \vec{b}$, and $\mathbf{A}$ is regular, then

$$\mathbf{A}\vec{x} = \mathbf{LU}\vec{x} = \mathbf{L}\vec{c} = \vec{b}, \text{ for some unknown } \vec{c}. \quad (\text{we define } \vec{c} := \mathbf{U}\vec{x})$$

Therefore, instead of solving the original equation, we can instead solve $\mathbf{L}\vec{c} = \vec{b}$ for $\vec{c}$,

then solve $\mathbf{LU}\vec{x} = \mathbf{L}\vec{c}$, or rather $\mathbf{U}\vec{x} = \vec{c}$ for $\vec{x}$.

But why break one problem into two problems? Because solving systems in upper or lower triangular form is much easier!

Example: $\mathbf{A} = \begin{bmatrix} 2 & 1 & 1 \\ 4 & 5 & 2 \\ 2 & -2 & 0 \end{bmatrix}$, and $\mathbf{A} = \mathbf{LU} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{bmatrix}$.

Solve $\mathbf{A}\vec{x} = \vec{b}$, where $\vec{b} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$.

First solve $\mathbf{L}\vec{c} = \vec{b}$ or $\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$

Observe we can do this with *forward substitution*. That is, starting with the first row: $c_1 = 1$.

Therefore from the second row: $2(1) + c_2 = 2$ or $c_2 = 0$. And from the third row: $(1) - (0) + c_3 = 2$ or $c_3 = 1$.

So we have $\vec{c} = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix}^T$.

Now we must solve $\mathbf{U}\vec{x} = \vec{c}$ or
$$\begin{bmatrix} 2 & 1 & 1 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

Observe we can do this with *backward substitution*:

From the third row we get: $x_3 = -1$. From the second row we get: $x_2 = 0$.

And from the first row we get: $2x_1 + 1(0) + 1(-1) = 1$ or $x_1 = 1$.

Therefore, the soln to our system is $\vec{x} = [1 \ 0 \ -2]^T$.

