Applied Linear Algebra

Instructor: Jodin Morey moreyjc@umn.edu

Announcements

- TA is Ratul (biswa087@umn.edu).
- Office Hours: Jodin: Mon, 1:00-2:15p. Fri 3:35-4:50p. Vincent 8 (No office hours no Mon. 1/2)
  Ratul: Thurs, 1:25-3:20p (umn.zoom.us/j/2024003004).
- Use discussion board on Canvas to discuss homework (just don't give solutions).
- Tentative schedule on Canvas
- These lecture notes (PDFs) will be available on Canvas
- HW sets are available on Canvas and will be turned in on Gradescope every Wednesday (no late HW accepted).
  Any regrade requests must be requested on Gradescope before 11:30 PM, on the Wednesday, two weeks after it was due.

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

Recall that if you have a system of equations, with the same number of equations as unknowns, you can use Gaussian elimination to find solutions (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:

$$\begin{aligned} 1x + 5y + 8z &= 0 \\ 2x + 6y + 9z &= 2 \\ 3x + 7y + 11z &= 4 \\ 6x + 14y + 22z &= 8 \end{aligned}$$

Put it into an (augmented) matrix:

Continuing our example: $\begin{bmatrix} 1 & 5 & 8 & | & 0 \\ 2 & 6 & 9 & | & 2 \\ 3 & 7 & 11 & | & 4 \\ 6 & 14 & 22 & | & 8 \end{bmatrix}$

Types of matrices: a or $[a]$, $\begin{bmatrix} a_1 & a_2 \end{bmatrix}$, $\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$, $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, $\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$.

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 $\mathbf{A}, \mathbf{B},$ etc., we can treat them as mathematical objects. (?!?)

We define operations like addition, subtraction, multiplication.

First let's define what it means for a matrix $\mathbf{A}$ to be equal to another matrix $\mathbf{B}$: they are the same size, and their elements are equal.

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

Here are some other operations:

Addition: $\mathbf{A} + \mathbf{B} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} a+e & b+f \\ c+g & d+h \end{bmatrix}.$

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

Addition requires that both matrices be the same size.

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:

$$\begin{aligned} 1x + 5y + 8z &= 0 \\ 2x + 6y + 9z &= 2 \\ 3x + 7y + 11z &= 4 \end{aligned},$$

with the coefficient matrix $\mathbf{A} = \begin{bmatrix} 1 & 5 & 8 \\ 2 & 6 & 9 \\ 3 & 7 & 11 \end{bmatrix}$, 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 $\mathbf{A}$ is equal to the number of rows in $\mathbf{B}$.

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

Example: $\mathbf{A}^{3\times3}\mathbf{x}^{3\times1} = \begin{bmatrix} 1 & 5 & 8 \\ 2 & 6 & 9 \\ 3 & 7 & 11 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1x+5y+8z \\ 2x+6y+9z \\ 3x+7y+11z \end{bmatrix} = \mathbf{C}^{3\times1}$.

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

See animated in class

And similarly:

$$\mathbf{a}^{1\times3}\mathbf{b}^{3\times1} = \begin{bmatrix} 1 & 5 & 8 \end{bmatrix}\begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix} = 1\cdot1 + 5(-2) + 8\cdot3 = 15 = \mathbf{c}^{1\times1}$$

$= \sum_{k=1}^{n} a_k b_k$ (dot product of vectors), 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 $c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$.

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

Algebraic Properties: Identity and Zero Matrices

$\mathbf{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\quad \mathbf{0} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, $\quad \mathbf{AI} = \mathbf{A}$, $\quad \mathbf{A} + \mathbf{0} = \mathbf{A}$.

$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$ **Additive Commutative Property**

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

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

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

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 all sinks in):
https://youtu.be/XkY2DOUCWMU

Definition: A matrix is called *upper* (*lower*) *triangular* if all of the entries below (above) the main diagonal 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 (lower) *unitriangular* if it is upper (lower) triangular **and** all of the main diagonal entries 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 one row by a nonzero constant.

These are the same operations you did in prior math on systems of equations, but notated with a matrix.

Solve a system by putting it into a matrix, then transforming it into an **Echelon Form** matrix by creating pivot columns (with "pivot points"):

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

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

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

where **U** is upper triangular.

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

Definition: If **A** can be transformed into upper triangular form **U** with nonzero pivot points, using only regular Gaussian Elimination, **A** is called a *regular matrix*.

Continuing an example from above: $\left[\begin{array}{ccc|c} 1 & 5 & 8 & 0 \\ 2 & 6 & 9 & 2 \\ 3 & 7 & 11 & 4 \end{array}\right]$

$$\overset{R_2+(-2R_1)}{\Rightarrow} \left[\begin{array}{ccc|c} 1 & 5 & 8 & 0 \\ 0 & -4 & -7 & 2 \\ 3 & 7 & 11 & 4 \end{array}\right] \overset{R_3+(-3R_1)}{\Rightarrow} \left[\begin{array}{ccc|c} 1 & 5 & 8 & 0 \\ 0 & -4 & -7 & 2 \\ 0 & -8 & -13 & 4 \end{array}\right]$$

$$\overset{R_3+(-2R_2)}{\Rightarrow} \left[\begin{array}{ccc|c} 1 & 5 & 8 & 0 \\ 0 & -4 & -7 & 2 \\ 0 & 0 & 1 & 0 \end{array}\right] \text{ or } \mathbf{U}\vec{x} = \vec{c}.$$

Only used operation #1. So **A** is regular.

LU Factorization

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

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

Definition: An *elementary matrix* $\mathbf{E}$ associated with an elementary row operation for an m-rowed matrix is the $m \times m$ matrix obtained by applying the row operation to $\mathbf{I}_m$.

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

Instead of the normal way, do this 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 accomplished 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 elementary row operations take the form:

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

To undo a type 1 elementary row operation, one merely changes the sign of a, so the inverse operations $\mathbf{L}$ take the form:

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

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

And using the inverses above, we define $\mathbf{L} := \mathbf{L}_n, \ldots, \mathbf{L}_1$.

And using matrix multiplication associativity, we find:

$\mathbf{LU} = (\mathbf{L}_n, \ldots, \mathbf{L}_1)(\mathbf{E}_1, \ldots, \mathbf{E}_n\mathbf{A}) = (\mathbf{L}_n, \ldots, \mathbf{L}_2)(\mathbf{L}_1\mathbf{E}_1)(\mathbf{E}_2, \ldots, \mathbf{E}_n\mathbf{A}) = (\mathbf{L}_n, \ldots, \mathbf{L}_2)(\mathbf{E}_2, \ldots, \mathbf{E}_n\mathbf{A}) = \ldots = \mathbf{A}$.

Theorem: $\mathbf{A}$ is regular *iff* it can be factored $\mathbf{A} = \mathbf{LU}$, where $\mathbf{L}$ is a lower unitriangular matrix, having all 1's on the diagonal, and **U** is upper triangular with nonzero diagonal entries, which are called the pivots of **A**. The nonzero off diagonal entries l_{ij} for $i > j$ appearing in **L** prescribe the **elementary row operations** that bring **A** into upper triangular form.

But why would we want to do this? LU factorization provides a way to solve the system of equations!

But why not use Gaussian elimination? Turns out that with 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

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

$\mathbf{A}\vec{x} = \mathbf{LU}\vec{x} = \mathbf{L}\vec{c} = \vec{b}$, for some unkown $\vec{c}$.

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

then solve $\mathbf{U}\vec{x} = \vec{c}$ for $\vec{x}$.

But why do this? 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 solution to our system is $\vec{x} = \begin{bmatrix} 1 & 0 & -2 \end{bmatrix}^T$.