18.700 - LINEAR ALGEBRA, DAY 2
GAUSSIAN ELIMINATION

SAM SCHIAVONE

CONTENTS

I. PRE-CLASS PLANNING

I.1. **Goals for lesson.**

(1) Students will learn the algorithm to row reduce a matrix.
(2) Students will apply row reduction algorithm to examples on worksheet.

I.2. **Methods of assessment.**

(1) Student responses to questions posed during lecture
(2) Student responses to worksheet

I.3. **Materials to bring.** (1) Laptop + adapter (2) Worksheets

(0:00)

II. Lesson Plan

Announcements: • First pset posted on website; due on Thursday on Canvas.

II.1. Review.

- Defined linear systems
- Solved some examples of linear systems
- Defined elementary row operations
- Defined reduced row echelon form

(0:05) Elementary Row Operations:

(1) Add a multiple of a row
(2) Swap two rows
(3) Rescale a row by a nonzero constant

Echelon forms [The word "echelon" comes from a military formation, standing in a diagonal. Also related to French *échelle*, meaning ladder.]

Definition 1.

- A matrix is in *echelon form* (or *row echelon form*) if it satisfies the following three properties:
  (i) All nonzero rows are above any rows of all zeros. (rows of all zeros are at the bottom)
  (ii) Each leading entry of a row is in a column to the right of the leading entry of the row above it.
  (iii) All entries in a column below a leading entry are zero. [Remark how this condition is techinically redundant; implied by previous condition.]
- A matrix is in RREF if it is in echelon form and satisfies two additional conditions:
  (iv) The leading entry in each nonzero row is a 1.
  (v) Each leading 1 is the only nonzero entry in its column.

Theorem (Uniqueness of reduced echelon form)**.** *Every matrix is row equivalent to a* unique *reduced echelon matrix.*

(0:15)

II.2. Gaussian elimination.

Definition 2.

- A *pivot position* in a matrix A is a location in A corresponding to a leading 1 in the reduced echelon form for A.
- A *pivot column* is a column of A that contains a pivot position.
- A *pivot* is a nonzero entry in a pivot position.

(0:20)

II.3. Row reduction algorithm.

The row reduction algorithm consists of 5 steps:

(1) Take the leftmost nonzero column. This is the leftmost pivot column. Make the pivot position at the top of the column.
(2) Select any nonzero entry in the leftmost pivot column as a pivot and move it into the pivot position (top) interchanging rows as necessary.
(3) Use row operations to make all entries under the pivot zero.
(4) Consider the submatrix obtained by removing the row and column containing the pivot. Go to the first step and repeat until there is no submatrix left to consider.

(5) Use row operations to create zeros above each pivot.

[Label (1) - (4) as echelonization, (5) as reduction. Also, DON'T ERASE THE ALGORITHM!]

(0:25)

Example 1. Let's row reduce the following augmented matrix. [Ask students. Explicitly point out the various steps during the example.]

$$\left(\begin{array}{ccc|c} 0 & 0 & 3 & 9 \\ 1 & 5 & -2 & 2 \\ 1/3 & 2 & 0 & 3 \end{array}\right)$$

(Should get

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 3 \end{array}\right)$$

for the RREF.)

Remark 1. There may be many ways to obtain the RREF of a matrix.

(0:30)

II.4. **Worksheet.**

(0:40)

II.5. **Parametric form for solutions.** Consider the augmented matrix

$$\left(\begin{array}{ccc|c} 1 & 0 & 3 & 1 \\ 0 & 1 & -2 & 5 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

This corresponds to the linear system

$$\boxed{\begin{array}{rrl} x_1 & & +3x_3 = 1 \\ & x_2 & -2x_3 = 5 \end{array}} \longrightarrow \boxed{\begin{array}{l} x_1 = 1 - 3x_3 \\ x_2 = 5 + 2x_3 \end{array}}$$

The variables corresponding to pivot columns (i.e x_1 and x_2) are called *pivot* or *leading variables*. The other variables (i.e., x_3) are called *free variables*.

We can use this description to write the set of solutions parametrically:

$$\begin{cases} x_1 = 1 - 3x_3 \\ x_2 = 5 + 2x_3 \\ x_3 \text{ free} \end{cases}$$

We call this the *parametric form* of the solutions. We can express this in vector notation as

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 - 3t \\ 5 + 2t \\ t \end{bmatrix} = \begin{bmatrix} -3t \\ 2t \\ t \end{bmatrix} + \begin{bmatrix} 1 \\ 5 \\ 0 \end{bmatrix} = \begin{bmatrix} -3 \\ 2 \\ 1 \end{bmatrix} t + \begin{bmatrix} 1 \\ 5 \\ 0 \end{bmatrix}.$$

[Ask students how many solutions this system has.]

(0:45)

II.6. **Worksheet part 2.**

(0:50)

II.7. **Existence and uniqueness of solutions.**

Theorem.

(1) *A linear system is consistent iff the right most column of the augmented matrix is* not *a pivot column, i.e., iff an echelon form for the matrix does* not *contain a row of the form*

$$\begin{bmatrix} 0 & \cdots & 0 & | \; b \end{bmatrix}$$

with b nonzero.

(2) *If the linear system is consistent, then the solution set contains either*
   (i) *a unique solution, if there are no free variables*
   (ii) *infinitely many solutions, if there is one or more free variables.*

[Remark on proofs briefly. (1) $\Rightarrow$ by contrapositive ($0 = b \neq 0$). If system has n variables, and has a unique solution, how many pivots does it have?]

(0:55)

II.8. **Summary.** Solving a linear system via row reduction:

(1) Write the augmented matrix of the system.
(2) Use the row reduction algorithm to put the matrix in echelon form. If the matrix is inconsistent, stop; otherwise continue to the next step.
(3) Keep row reducing to put the matrix in reduced echelon form.
(4) Write the linear system corresponding to the reduced echelon form.
(5) Express the solution to the system in parametric form.

(1:00) II.9. **1A: $\mathbb{R}^n$ and $\mathbb{C}^n$.** [Now following Axler, Chapter 1.]

The notion of a vector space was created to abstract many objects that share similar properties. The most classical example of vector spaces are $\mathbb{R}^2$ and $\mathbb{R}^3$, thought of as the real plane and real 3-space. We begin by generalizing these to higher dimensions.

II.9.1. *Definitions.*

Definition 3.

- Let $n \in \mathbb{Z}_{\geq 0}$ be a nonnegative integer. An *n-tuple* or *list* of *length n* is an ordered collection of n elements.
- Two lists are equal if they have the same length and the same entries in the same order.

Remark 2.

- Lists are usually written in the form $(z_1, z_2, \ldots, z_n)$. Note that lists must have **finite length**! The object $(z_1, z_2, \ldots)$ is not a list, but rather an infinite sequence.
- Unlike sets, the order and multiplicities of elements in tuples matters! So $\{2, 3\} = \{3, 2\}$ and $\{4, 4, 4\} = \{4\}$, but $(2, 3) \neq (3, 2)$ and $(4, 4, 4) \neq (4)$.

For the rest of the lecture, fix $n \in \mathbb{Z}_{>0}$.

Definition 4. Let $\mathbb{F}^n$ be the set of all n-tuples with entries in $\mathbb{F}$:

$$\mathbb{F}^n := \{(x_1, \ldots, x_n) : x_k \in \mathbb{F} \text{ for all } k = 1, \ldots, n\}.$$

Given $(x_1, \ldots, x_n) \in \mathbb{F}^n$, its k^{th} *coordinate* or *entry* is x_k.

Remark 3. It is sometimes convenient to write elements of $\mathbb{F}^n$ vertically, as column vectors.

Example 2. $\mathbb{C}^{17} = \{(z_1, \ldots, z_{17}) : z_1, \ldots, z_{17} \in \mathbb{C}\}$. [Can't visualize it, but still makes sense algebraically.]

(1:10)

II.9.2. *Algebraic properties of* $\mathbb{F}^n$.

(1) Addition operation
(2) 0-vector
(3) Additive inverses
(4) Scalar multiplication
(5) Commutativity of addition

[Draw examples in $\mathbb{R}^2$ to illustrate geometric intution behind these definitions. Prove commutativity of addition.]

(1:20)

II.10. **Abstract vector spaces.** Now that we've observed some of the algebraic properties of $\mathbb{F}^n$, we're going to give an abstract definition of a vector space as something with these properties, i.e.,a set equipped with addition and scalar multiplication operations satisfying some conditions.

Definition 5. A *vector space* is a set V equipped with an addition operation

$$\begin{aligned} + : V \times V &\to V \\ (u, v) &\mapsto u + v \end{aligned}$$

and a scalar multiplication operation

$$\begin{aligned} \mathbb{F} \times V &\to V \\ (\lambda, v) &\mapsto \lambda v \end{aligned}$$

satisfying the following properties.

(1) (Associativity of addition):
$$(u + v) + w = u + (v + w)$$
for all $u, v, w \in V$.
(2) (Additive identity): There exists an element $0 \in V$ such that $v + 0 = 0 + v = v$ for all $v \in V$.
(3) (Additive inverses): For each $v \in V$, there exists $w \in V$ such that $v + w = w + v = 0$.
(4) (Commutative of addition): $u + v = v + u$ for all $u, v \in V$.
(5) (Scalar multiplicative identity): $1\,v = v$ for all $v \in V$.
(6) (Associativity of scalar multiplication): $a(bv) = (ab)v$ for all $a, b \in \mathbb{F}$ and $v \in V$.
(7) (Distributive laws): [ask students]
    (a) $a(u + v) = au + av$ for all $a \in \mathbb{F}$ and all $u, v, \in V$.
    (b) $(a + b)v = av + bv$ for all $a, b \in F$ and all $v \in V$.

Elements of a vector space are called *vectors*. [Tell engineer, physicist, mathematician joke.]

Remark 4. Structures satisfying (1), (2), (3) are called *groups*; structures satisfying (1), (2), (3), (4) are called *abelian groups*.

Remark 5. Elements of $\mathbb{F}$ are sometimes called *scalars*. When we need to specify the underlying field, we will say that V is a *vector space over* $\mathbb{F}$ or an $\mathbb{F}$*-vector space*.

(1:30) **Example 3.**

- $\mathbb{R}^2$ is a vector space over $\mathbb{R}$.
- The *zero vector space* or *trivial vector space* is the set $\{0\}$ consisting of just the zero vector. Exercise: Check that it satisfies the axioms of a vector space.
- Let
$$\mathbb{F}^\infty := \{(x_1, x_2, \ldots) : x_k \in \mathbb{F} \text{ for all } k = 1, 2, \ldots\},$$
with addition and scalar multiplication defined componentwise:
$$\begin{aligned}(x_1, x_2, \ldots) + (y_1, y_2, \ldots) &= \cdots \\ \lambda(x_1, x_2, \ldots) &= \cdots .\end{aligned}$$
Exercise: Check that $\mathbb{F}^\infty$ satisfies the axioms of a vector space. What is the zero vector?
- Let S be any set. Define
$$\mathbb{F}^S := \{f : S \to \mathbb{F}\}$$
with addition and scalar multiplication defined pointwise, i.e., for any $f, g \in \mathbb{F}^S$ and any $\lambda \in \mathbb{F}$
$$(f+g)(x) := f(x) + g(x)(\lambda f)(x) := \lambda f(x)$$
for all $x \in S$.

Exercise: $\mathbb{F}^S$ is a vector space.

Remark 6. One can view $\mathbb{F}^n$ as $\mathbb{F}^S$ where $S = \{1, 2, \ldots, n\}$.
$$\begin{aligned}\mathbb{F}^S &\leftrightarrow \mathbb{F}^n \\ (f : S \to \mathbb{F}) &\mapsto (f(1), f(2), \ldots, f(n)) \\ (g : S \to \mathbb{F}) &\leftarrow\!\shortmid (x_1, x_2, \ldots, x_n) \\ k \mapsto x_k\end{aligned}$$
One can similarly view $\mathbb{F}^\infty$ as $\mathbb{F}^S$ where $S = \mathbb{Z}_{>0} = \{1, 2, \ldots\}$.