

We will primarily be studying systems of linear equations, and later, systems of ordinary linear differential equations. We begin with some terminology and notation.

An $m \times n$ matrix A is an array with m rows and n columns. We'll write A as

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} = (a_{ij})$$

The general entry a_{ij} is the entry in the i^{th} row and j^{th} column.

So, an $m \times 1$ matrix is a column vector $\begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix} = \bar{a}$,

and a $1 \times n$ matrix is a row vector $\bar{c} = [c_1, \dots, c_n]$.

If $A^{(1)}, \dots, A^{(n)}$ are column m -vectors (i.e., $A^{(i)} \in \mathbb{R}^m$ for each i) we write $[A^{(1)}, \dots, A^{(n)}]$ for the $m \times n$ matrix whose i^{th} column is $A^{(i)}$.

Similarly, if $B_{(1)}, \dots, B_{(m)}$ are row vectors we write $\begin{bmatrix} B_{(1)} \\ \vdots \\ B_{(m)} \end{bmatrix}$ for the matrix whose i^{th} row is $B_{(i)}$.

If $A = (a_{ij})$ is an $m \times n$ matrix the transpose of A is the $n \times m$ matrix $A^T = (a_{ji})$, the matrix obtained from A by switching a_{ij} and a_{ji} . Note that A^T is an $n \times m$ matrix.

For example, if $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ then

$$A^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}.$$

Remark: $(A^T)^T = A$.

If $A = (a_{ij})$ and $B = (b_{ij})$ are each $m \times n$ matrices we define $A+B$ to be the $m \times n$ matrix $(a_{ij} + b_{ij})$.

For example, if $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$ $B = \begin{pmatrix} 7 & 8 \\ 9 & 1 \\ 2 & 3 \end{pmatrix}$

$$\text{then } A+B = \begin{pmatrix} 1+7 & 2+8 \\ 3+9 & 4+1 \\ 5+2 & 6+3 \end{pmatrix} = \begin{pmatrix} 8 & 10 \\ 12 & 5 \\ 7 & 9 \end{pmatrix}$$

Then $(A+B)^T = A^T + B^T$

An $n \times n$ matrix $A = (a_{ij})$ is called a square matrix. The elements $a_{11}, a_{22}, \dots, a_{nn}$ are called the elements of the main diagonal of A . The trace of A is $\text{Tr}(A) = a_{11} + a_{22} + \dots + a_{nn}$.

the sum of the elements on the main diagonal.

- A square matrix is symmetric if $A = A^T$.

For example, $A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 2 & 1 \\ 2 & 1 & 3 \end{pmatrix}$ is symmetric.

- A is anti-symmetric if $A = -A^T$. We will also say that such an A is skew-symmetric. For example, $\begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & -3 \\ -2 & 3 & 0 \end{pmatrix}$ is skew-symmetric.

Notice that the main diagonal elements of a skew-symmetric matrix must be 0 since $a_{ii} = -a_{ii}$.

A square matrix is said to be diagonal if $a_{ij} = 0$ if $i \neq j$. For example, $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ is diagonal.

A square matrix is lower (respectively, upper) triangular if $a_{ij} = 0$ when $i < j$ (respectively, when $i > j$). Thus, a lower triangular matrix has only zero entries above the main diagonal, and an upper triangular matrix has only zero entries below the main diagonal. For example,

$\begin{pmatrix} 1 & 0 & 0 \\ 4 & 2 & 0 \\ 5 & 6 & 3 \end{pmatrix}$ is lower triangular,

and $\begin{pmatrix} 1 & 4 & 5 \\ 0 & 2 & 6 \\ 0 & 0 & 3 \end{pmatrix}$ is upper triangular.

A matrix function is a matrix whose entries are functions of a variable t , for example

$$A(t) = \begin{pmatrix} e^t & e^{2t} \\ \sin t & \cos t \end{pmatrix}.$$

If $A(t) = (a_{ij}(t))$ is a matrix function, and the functions $a_{ij}(t)$ are differentiable (respectively, integrable on $[a, b]$) then $\frac{dA}{dt} = \left(\frac{da_{ij}(t)}{dt} \right)$,

$$\text{and } \int_a^b A(t) dt = \left(\int_a^b a_{ij}(t) dt \right)$$

Matrix Arithmetic

We've already defined $A+B$ when A and B are $m \times n$ matrices. If r is a scalar (i.e., a real or complex number) we define rA to be (ra_{ij}) (where, as usual, $A = (a_{ij})$). We have the following:

(5)

- (1) $A+B = B+A$ (commutativity)
- (2) $A+(B+C) = (A+B)+C$ (associativity) when A, B , and C are each $m \times n$ matrices.
- (3) If r is a scalar $rA = Ar$
- (4) $r(A+B) = rA + rB$
and $(r+s)A = rA + sA$
- (5) $1 \cdot A = A$
- (6) $(rs)A = r(sA)$

Matrix Products: Recall that, if $\bar{a}$ and $\bar{b}$ are either row or column n -vectors, with components $a_1, \dots, a_n$ and $b_1, \dots, b_n$ then the inner (or dot) product $\bar{a} \cdot \bar{b}$ is

$$a_1 b_1 + a_2 b_2 + \dots + a_n b_n = \sum_{i=1}^n a_i b_i$$

For example, if $\bar{a} = (3, 1, 4)$ and $\bar{b} = (x, y, z)$ then $\bar{a} \cdot \bar{b} = 3x + y + 4z$.

If $A = [A^{(1)}, \dots, A^{(n)}]$ is an $m \times n$ matrix with columns $A^{(i)}$, and $\bar{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ then $A\bar{x} = A^{(1)}x_1 + \dots + A^{(n)}x_n$

Notice that, if we write $A = (a_{ij})$, then $A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ is the vector whose i^{th} component

is the inner product of the i^{th} row of A with the vector $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, so if A is an $m \times n$

matrix then $A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ is an m -vector.

Exercise: Compute $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ in two

different ways.

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}x + \begin{bmatrix} 2 \\ 5 \end{bmatrix}y + \begin{bmatrix} 3 \\ 6 \end{bmatrix}z = \begin{bmatrix} x+2y+3z \\ 4x+5y+6z \end{bmatrix}$$

$$\text{Also, } \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} (1 \ 2 \ 3) \cdot (x, y, z) \\ (4 \ 5 \ 6) \cdot (x, y, z) \end{bmatrix} = \begin{bmatrix} x+2y+3z \\ 4x+5y+6z \end{bmatrix}$$

Finally, if A is an $m \times n$ matrix and B is an $n \times s$ matrix with columns $B^{(1)}, \dots, B^{(s)}$ we define the product AB to be the $m \times s$ matrix whose columns are $AB^{(1)}, \dots, AB^{(s)}$.

$$\text{For example, } \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} a & d \\ b & e \\ c & f \end{pmatrix} = \begin{pmatrix} a+2b+3c & d+2e+3f \\ 4a+5b+6c & 4d+5e+6f \end{pmatrix}$$

Remark (1) The ij^{th} entry of AB is the inner product of the i^{th} row of A with the j^{th} column of B , so if $C = (c_{ij}) = AB$ then $c_{ij} = \sum_k a_{ik} b_{kj}$.

(2) If A is an $m \times n$ matrix and $\bar{x}$ is an n -vector then $A\bar{x}$ is an m -vector, so we may view the matrix A as a function $A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ (assuming

all entries are real). From this viewpoint matrix product is simply the composition of functions—

If A is an $m \times n$ matrix, and B is an $n \times s$ matrix then

$$\mathbb{R}^s \xrightarrow{B} \mathbb{R}^n \xrightarrow{A} \mathbb{R}^m$$

and the matrix product AB is the composition $A \circ B$ of functions.

Since composition of functions is associative, so is matrix multiplication, i.e.,

$$A(BC) = (AB)C \quad \text{whenever all products are defined.}$$

Note that, in general, AB and BA are not equal! They are both defined only if A is an $m \times n$ matrix and B is an $n \times m$ matrix. Then we may view A as a function $\mathbb{R}^n \xrightarrow{A} \mathbb{R}^m$ and B as a function $\mathbb{R}^m \rightarrow \mathbb{R}^n$, so AB takes m vectors to m vectors, and BA takes n vectors to n vectors.

Example

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$$

$$B = \begin{pmatrix} 5 & 7 & 3 \\ 6 & 8 & 2 \end{pmatrix}$$

$$AB = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} 5 & 7 & 3 \\ 6 & 8 & 2 \end{pmatrix} = \begin{pmatrix} 1 \cdot 5 + 2 \cdot 6 & 1 \cdot 7 + 2 \cdot 8 & 1 \cdot 3 + 2 \cdot 2 \\ 3 \cdot 5 + 4 \cdot 6 & 3 \cdot 7 + 4 \cdot 8 & 3 \cdot 3 + 4 \cdot 2 \\ 5 \cdot 5 + 6 \cdot 6 & 5 \cdot 7 + 6 \cdot 8 & 5 \cdot 3 + 6 \cdot 2 \end{pmatrix}$$

$$= \begin{pmatrix} 17 & 23 & 7 \\ 39 & 53 & 17 \\ 61 & 83 & 27 \end{pmatrix}$$

$$BA = \begin{pmatrix} 5 & 7 & 3 \\ 6 & 8 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} = \begin{pmatrix} 5 \cdot 1 + 7 \cdot 3 + 3 \cdot 5 & 5 \cdot 2 + 7 \cdot 4 + 3 \cdot 6 \\ 6 \cdot 1 + 8 \cdot 3 + 2 \cdot 5 & 6 \cdot 2 + 8 \cdot 4 + 2 \cdot 6 \end{pmatrix}$$

$$= \begin{pmatrix} 41 & 56 \\ 40 & 56 \end{pmatrix}$$

Lecture 2

Let $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, $B = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}$

Then $AB = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 1 \cdot 5 + 2 \cdot 7 & 1 \cdot 6 + 2 \cdot 8 \\ 3 \cdot 5 + 4 \cdot 7 & 3 \cdot 6 + 4 \cdot 8 \end{pmatrix} = \begin{pmatrix} 19 & 22 \\ 43 & 50 \end{pmatrix}$

and $BA = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 5 \cdot 1 + 6 \cdot 3 & 5 \cdot 2 + 6 \cdot 4 \\ 7 \cdot 1 + 8 \cdot 3 & 7 \cdot 2 + 8 \cdot 4 \end{pmatrix} = \begin{pmatrix} 23 & 34 \\ 31 & 46 \end{pmatrix}$

Remarks: (1) Remember AB and BA are usually different.

(2) $A(B+C) = AB + AC$ when all products are defined

(3) $(A+B)C = AC + BC$ when all products are defined.

(4) If A and B are each upper (respectively, lower) triangular so is AB .

For example, $\begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix} \begin{pmatrix} 2 & 1 & 3 \\ 0 & 7 & 8 \\ 0 & 0 & 9 \end{pmatrix} =$

$$\begin{pmatrix} 1 \cdot 2 & 1 \cdot 1 + 2 \cdot 7 & 1 \cdot 3 + 2 \cdot 8 + 3 \cdot 9 \\ 0 & 4 \cdot 7 & 4 \cdot 8 + 5 \cdot 9 \\ 0 & 0 & 6 \cdot 9 \end{pmatrix} = \begin{pmatrix} 2 & 15 & 46 \\ 0 & 28 & 77 \\ 0 & 0 & 54 \end{pmatrix}$$

is upper triangular.

Linear Systems:

A linear system of m equations in n variables is a system of the form

$$\begin{aligned}
 (*) \quad & a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\
 & a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\
 & \vdots \\
 & a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m
 \end{aligned}$$

where the a_{ij} and b_i are scalars, and $x_1, \dots, x_n$ are variables.

The $m \times n$ matrix $A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$ is called

the coefficient matrix of the system $(*)$. Notice that the left hand side of $(*)$ is just the product $A\bar{x}$ where A is the coefficient matrix and $\bar{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ is a vector of variables.

If we write $\bar{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$ then the system $(*)$

may be written in the form $A\bar{x} = \bar{b}$.

If $b_i = 0$ for all i the system is called homogeneous. If some b_i is non-zero the system is called nonhomogeneous.

A solution to the system $(*)$ is a set $c_1, \dots, c_n$ of scalars which upon substituting $x_1 = c_1, x_2 = c_2, \dots, x_n = c_n$ satisfies the system.

If we let $\bar{c} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$ then $c_1, \dots, c_n$ is a solution to the system if $A\bar{c} = \bar{b}$.

Before we try to solve the system $A\bar{x} = \bar{b}$ we might ask two basic questions:

- (1) When does the system have a solution?
- (2) If the system has a solution, how many solutions does it have.

(1) Not every system has a solution. For example,

$$x + y + z = 0$$

$$x + y = 0$$

$$z = 1$$

has no solution.

A system that has no solutions is said to be inconsistent. A system that has a solution is said to be consistent. Note that a homogeneous system $A\bar{x} = \bar{0}$ is always consistent since it has $\bar{x} = \bar{0}$ as a solution.

Examples: (1) A system of 2 equations in 2 variables

$$ax + by = c$$

$$dx + ey = f$$

The equations represent two lines in the plane which must intersect in a point unless they are parallel, or identical. Thus, such a system may be inconsistent, or it may be consistent with

either one or infinitely many solutions.

(2) A system of 3 equations in 3 unknowns

$$ax + by + cz = d.$$

$$ex + fy + gz = h$$

$$ix + jy + kz = L$$

Each of these equations is the equation of a plane. There are four possibilities - (1) The planes do not intersect

(2) The planes intersect in one point.

(3) The planes intersect in a line

(4) The planes are identical.

Once again, the system may be inconsistent, or it may be consistent with either one or infinitely many solutions.

Remark: The coefficient matrix A of the system (*) is an $m \times n$ matrix. We also associate to (*) an $m \times (n+1)$ matrix $A^\#$ which is obtained from A by adjoining the column vector $\begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$.

$$A^\# = \left(\begin{array}{ccc|c} a_{11} & \dots & a_{1n} & b_1 \\ a_{21} & \dots & a_{2n} & b_2 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{array} \right)$$

The matrix $A^\#$ is called the augmented coefficient matrix of the system (*).

You already know how to solve a system (*). Now we'd like to make the process systematic.

When you try to solve a system you are allowed to perform certain operations on the equations that do not change the solution-set. These are

- (1) Multiply an equation by a nonzero constant
- (2) Add a constant multiple of one equation to another

and the seemingly pointless operation

- (3) permute the equations.

If we write the system as $A\bar{x} = \bar{b}$ then the above operations have the following effects on A and $A^\#$

- (1) Multiply a row by a non-zero constant.
- (2) Add a multiple of one row to another
- (3) Permute the rows.

These three operations are called elementary row operations. We will follow the text, using the following notation.

- (1) $M_i(k)$ means multiply the i^{th} row by the scalar k .
- (2) $A_{ij}(k)$ means add k times the i^{th} row to row j .
- (3) P_{ij} means permute the i^{th} and j^{th} row.

We say that two matrices A and B are row equivalent if B can be obtained from A by a series of elementary row operations. We

will write $A \sim B$ to indicate that A is row equivalent to B .

In our attempt to make systematic the search for the solution to the linear system (*) our immediate goal is to transform the coefficient matrix A (or the augmented coefficient matrix $A^\#$) into a special form.

An $m \times n$ matrix is said to be in row-echelon form if it satisfies the following three conditions

(1) Any row consisting only of zeros lies below all non-zero rows.

(2) The first non-zero entry in a non-zero row is 1.

We call this entry a leading 1.

(3) The leading 1 in any row (below the first row) occurs to the right of the leading 1 in the row above.

The process of reducing $A^\#$ to row-echelon form by a sequence of elementary row operations is called Gaussian elimination.

The elementary row operations don't change the set of solutions to the system (*) so we can read the solutions to (*) from the solutions to the system corresponding to a row-echelon form for $A^\#$.

Remark: If a row-echelon form has more than

one non-zero row then there is more than one row-echelon form for A .

Example: Solve the system

$$x + y + z = 1$$

$$x + y - z = 1$$

$$x - y - z = 0$$

by placing the augmented coefficient matrix in row-echelon form.

$$A^* = \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 0 \end{array} \right)$$

$$\xrightarrow{A_{23}(-1)} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 0 & -2 & 0 & -1 \end{array} \right) \xrightarrow{A_{12}(-1)} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 0 & -2 & 0 \\ 0 & -2 & 0 & -1 \end{array} \right)$$

$$\xrightarrow{M_2(-\frac{1}{2})} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -2 & 0 & -1 \end{array} \right) \xrightarrow{M_3(-\frac{1}{2})} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & \frac{1}{2} \end{array} \right)$$

$$\xrightarrow{P_{2,3}} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 \end{array} \right)$$

This is the augmented matrix for the system

$$x + y + z = 1$$

$$y = \frac{1}{2}$$

$$z = 0$$

Back substituting, we find $x + \frac{1}{2} = 1$, so $x = \frac{1}{2}$, $y = \frac{1}{2}$, $z = 0$.

①

Lecture 3

Consider the system

$$2x + y + z = 1$$

$$x + 3y + z = 0$$

$$4x - y + 2z = 1$$

(1) What is the augmented coefficient matrix $A^\#$?

$$A^\# = \left(\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ 1 & 3 & 1 & 0 \\ 4 & -1 & 2 & 1 \end{array} \right)$$

(2) Place $A^\#$ in row-echelon form.

$$\left(\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ 1 & 3 & 1 & 0 \\ 4 & -1 & 2 & 1 \end{array} \right) \xrightarrow{A_{13}(-2)} \left(\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ 1 & 3 & 1 & 0 \\ 0 & -3 & 0 & -1 \end{array} \right) \xrightarrow{A_{12}(-\frac{1}{2})} \left(\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ 0 & \frac{5}{2} & \frac{1}{2} & -\frac{1}{2} \\ 0 & -3 & 0 & -1 \end{array} \right)$$

$$\xrightarrow{A_{32}(\frac{5}{6})} \left(\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ 0 & 0 & \frac{1}{2} & -\frac{4}{3} \\ 0 & -3 & 0 & -1 \end{array} \right) \xrightarrow{P_{23}} \left(\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ 0 & -3 & 0 & -1 \\ 0 & 0 & \frac{1}{2} & -\frac{4}{3} \end{array} \right)$$

$$\xrightarrow{M_2(-\frac{1}{3})} \left(\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ 0 & 1 & 0 & \frac{1}{3} \\ 0 & 0 & \frac{1}{2} & -\frac{4}{3} \end{array} \right) \xrightarrow{M_3(2)} \left(\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ 0 & 1 & 0 & \frac{1}{3} \\ 0 & 0 & 1 & -\frac{8}{3} \end{array} \right)$$

$$\xrightarrow{M_1(\frac{1}{2})} \left(\begin{array}{ccc|c} 1 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{3} \\ 0 & 0 & 1 & -\frac{8}{3} \end{array} \right).$$

(2)

This is the augmented matrix for the system

$$x + \frac{1}{2}y + \frac{1}{2}z = \frac{1}{2}$$

$$y = \frac{1}{3}$$

$$z = -\frac{8}{3}$$

So, $x = \frac{5}{3}$, $y = \frac{1}{3}$, $z = -\frac{8}{3}$ is the solution to our system.

Another example - Write the augmented matrix for

$$x + y - 2z = 2$$

$$3x - y + z = 1$$

$$5x - 3y + 4z = 0$$

$$A^\# = \left(\begin{array}{ccc|c} 1 & 1 & -2 & 2 \\ 3 & -1 & 1 & 1 \\ 5 & -3 & 4 & 0 \end{array} \right)$$

Place $A^\#$ in row-echelon form

$$\left(\begin{array}{ccc|c} 1 & 1 & -2 & 2 \\ 3 & -1 & 1 & 1 \\ 5 & -3 & 4 & 0 \end{array} \right) \xrightarrow{A_{12}(-3)} \left(\begin{array}{ccc|c} 1 & 1 & -2 & 2 \\ 0 & -4 & 7 & -5 \\ 5 & -3 & 4 & 0 \end{array} \right)$$

$$\xrightarrow{A_{13}(-5)} \left(\begin{array}{ccc|c} 1 & 1 & -2 & 2 \\ 0 & -4 & 7 & -5 \\ 0 & -8 & 14 & -10 \end{array} \right) \xrightarrow{A_{23}(-2)} \left(\begin{array}{ccc|c} 1 & 1 & -2 & 2 \\ 0 & -4 & 7 & -5 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

$$\xrightarrow{M_2[-1/4]} \left(\begin{array}{ccc|c} 1 & 1 & -2 & 2 \\ 0 & 1 & -7/4 & 5/4 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

(3)

This is the augmented matrix for the system

$$x + y - 2z = 2$$

$$y - \frac{7}{4}z = \frac{5}{4}$$

$$\text{So, } y = \frac{5+7z}{4} \quad \text{and}$$

$$x + y - 2z = x + \frac{5+7z}{4} - 2z = x + \frac{5-z}{4} = 2.$$

$$\text{and } x = \frac{3+z}{4}.$$

Every value of z gives us a value of x and of y that yield a solution to the system. Thus, this system has infinitely many solutions.

Now, it's your turn.

Let $A = \begin{pmatrix} 0 & 1 & 2 & 1 \\ 0 & 3 & 1 & 2 \\ 0 & 2 & 0 & 1 \end{pmatrix}$. Find a row-echelon

form for A .

$$\begin{pmatrix} 0 & 1 & 2 & 1 \\ 0 & 3 & 1 & 2 \\ 0 & 2 & 0 & 1 \end{pmatrix} \xrightarrow{A_{12}(-3)} \begin{pmatrix} 0 & 1 & 2 & 1 \\ 0 & 0 & -5 & -1 \\ 0 & 2 & 0 & 1 \end{pmatrix} \xrightarrow{A_{13}(-2)} \begin{pmatrix} 0 & 1 & 2 & 1 \\ 0 & 0 & -5 & -1 \\ 0 & 0 & -4 & -1 \end{pmatrix}$$

$$\xrightarrow{M_2(-1/5)} \begin{pmatrix} 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 1/5 \\ 0 & 0 & -4 & -1 \end{pmatrix} \xrightarrow{A_{23}(4)} \begin{pmatrix} 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 1/5 \\ 0 & 0 & 0 & -1/5 \end{pmatrix} \xrightarrow{M_3(-5)} \begin{pmatrix} 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 1/5 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

As we've mentioned, there may be many row-echelon forms for a matrix, and many sequences of elementary row operations that lead us to one. However, we can give an algorithm for getting there. Suppose A is a non-zero $m \times n$ matrix.

- (1) Find the leftmost non-zero column. We call this a pivot column.
- (2) Using elementary row operations place a 1 in the topmost position in this column (this position is called a pivot).
- (3) Use elementary row operations to put zeros below the pivot.

If there are no other non-zero rows the matrix is in row-echelon form.

Otherwise,

- (4) Apply the above steps to the submatrix consisting of the rows that are below the row containing the pivot.

We will see that the number of non-zero rows in any row-echelon form of A is the same. (i.e., the number of non-zero rows depends only on A , not on any row-echelon form for A). This number is called the rank of A , denoted $\text{rk}(A)$.

Remark: If A is an $m \times n$ matrix then $\text{rk}(A)$ is $\leq m$ and n since the number of non-zero rows is the number of pivots in a row-echelon form, and each pivot lies in exactly one row and one column.

Looking at our earlier examples,

$$\text{rk} \begin{pmatrix} 1 & 1 & -2 & 2 \\ 3 & -1 & 1 & 1 \\ 5 & -3 & 4 & 0 \end{pmatrix} = 2 \quad \text{and} \quad \text{rk} \begin{pmatrix} 0 & 1 & 2 & 1 \\ 0 & 3 & 1 & 2 \\ 0 & 2 & 0 & 1 \end{pmatrix} = 3$$

Definition: A matrix in row-echelon form is said to be in reduced row-echelon form if any column containing a leading 1 has only zeros for its other entries. (The reduced row echelon form turns out to be unique.)

For example,
$$\begin{pmatrix} 1 & 2 & 0 & 0 & 0 & 6 \\ 0 & 0 & 1 & 0 & 0 & 5 \\ 0 & 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$
 is

in reduced row-echelon form.

Let's look at the example we began today's class with.

$$A^{\#} = \left(\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ 1 & 3 & 1 & 0 \\ 4 & -12 & 1 & 1 \end{array} \right).$$

(6)

We saw that

$$\left(\begin{array}{ccc|c} 1 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{3} \\ 0 & 0 & 1 & -\frac{8}{3} \end{array} \right)$$

is a row-echelon form for $A^\#$.

Let's continue to row-reduce it to obtain the reduced row-echelon form for $A^\#$.

$$\left(\begin{array}{ccc|c} 1 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{3} \\ 0 & 0 & 1 & -\frac{8}{3} \end{array} \right) \xrightarrow{A_{21}(-\frac{1}{2})} \left(\begin{array}{ccc|c} 1 & 0 & \frac{1}{2} & \frac{1}{3} \\ 0 & 1 & 0 & \frac{1}{3} \\ 0 & 0 & 1 & -\frac{8}{3} \end{array} \right)$$

$$\xrightarrow{A_{31}(-\frac{1}{2})} \left(\begin{array}{ccc|c} 1 & 0 & 0 & \frac{5}{3} \\ 0 & 1 & 0 & \frac{1}{3} \\ 0 & 0 & 1 & -\frac{8}{3} \end{array} \right)$$

This is in reduced row-echelon form.

It is the augmented matrix for the system

$$x = \frac{5}{3}$$

$$y = \frac{1}{3}$$

$$z = -\frac{8}{3}.$$

In other words, we can read the solutions to the original system directly from the reduced row-echelon form.

The process of using row reduction to find the reduced row-echelon form is sometimes called Gauss-Jordan elimination.

Your turn again - Find the reduced row-echelon form for

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 3 & 2 \\ 4 & 10 & 6 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 1 \\ 1 & 3 & 2 \\ 4 & 10 & 6 \end{pmatrix} \xrightarrow{A_{13}(-4)} \begin{pmatrix} 1 & 2 & 1 \\ 1 & 3 & 2 \\ 0 & 2 & 2 \end{pmatrix} \xrightarrow{A_{12}(-1)} \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 2 & 2 \end{pmatrix}$$

$$\xrightarrow{A_{23}(-2)} \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow{A_{21}(-2)} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Lecture 4

We begin with another example. Let's find the reduced row-echelon form for the augmented coefficient matrix of the system

$$x + 2y + z + w = 1$$

$$x - y + z - w = 0$$

$$2x + w = 1$$

$$A^{\#} = \left(\begin{array}{cccc|c} 1 & 2 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 0 \\ 2 & 0 & 0 & 1 & 1 \end{array} \right) \xrightarrow{A_{13}(-2)} \left(\begin{array}{cccc|c} 1 & 2 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 0 \\ 0 & -4 & -2 & -1 & -1 \end{array} \right)$$

$$\xrightarrow{A_{12}(-1)} \left(\begin{array}{cccc|c} 1 & 2 & 1 & 1 & 1 \\ 0 & -3 & 0 & -2 & -1 \\ 0 & -4 & -2 & -1 & -1 \end{array} \right) \xrightarrow{A_{23}(-\frac{4}{3})} \left(\begin{array}{cccc|c} 1 & 2 & 1 & 1 & 1 \\ 0 & -3 & 0 & -2 & -1 \\ 0 & 0 & -2 & \frac{5}{3} & \frac{1}{3} \end{array} \right)$$

$$\xrightarrow{A_{21}(\frac{2}{3})} \left(\begin{array}{cccc|c} 1 & 0 & 1 & -\frac{1}{3} & \frac{1}{3} \\ 0 & -3 & 0 & -2 & -1 \\ 0 & 0 & -2 & \frac{5}{3} & \frac{1}{3} \end{array} \right) \xrightarrow{M_2(-\frac{1}{3})} \left(\begin{array}{cccc|c} 1 & 0 & 1 & -\frac{1}{3} & \frac{1}{3} \\ 0 & 1 & 0 & \frac{2}{3} & \frac{1}{3} \\ 0 & 0 & -2 & \frac{5}{3} & \frac{1}{3} \end{array} \right)$$

$$\xrightarrow{A_{31}(\frac{1}{2})} \left(\begin{array}{cccc|c} 1 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & \frac{2}{3} & \frac{1}{3} \\ 0 & 0 & -2 & \frac{5}{3} & \frac{1}{3} \end{array} \right) \xrightarrow{M_3(-\frac{1}{2})} \left(\begin{array}{cccc|c} 1 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & \frac{2}{3} & \frac{1}{3} \\ 0 & 0 & 1 & -\frac{5}{6} & -\frac{1}{6} \end{array} \right)$$

This is in reduced row-echelon form. It is the augmented matrix for the system

(2)

$$x + \frac{1}{2}w = \frac{1}{2}$$

$$y + \frac{2}{3}w = \frac{1}{3}$$

$$z - \frac{5}{6}w = -\frac{1}{6}$$

We can choose any value for w , and this value will determine x , y , and z . For this reason we call w a free variable (or free parameter) and we call x , y , and z bound variables.

If we set $w = t$ we find

$$x = -\frac{1}{2}t + \frac{1}{2}$$

$$y = -\frac{2}{3}t + \frac{1}{3}$$

$$z = \frac{5}{6}t - \frac{1}{6}$$

$$w = t$$

So the set $\left\{ \left(-\frac{1}{2}t + \frac{1}{2}, -\frac{2}{3}t + \frac{1}{3}, \frac{5}{6}t - \frac{1}{6}, t \right) : t \in \mathbb{R} \right\}$

is the solution set to our system.

Notice that our free variable is the variable that does not correspond to a leading 1 of the reduced row-echelon form.

Suppose $A\bar{x} = \bar{b}$ is a system of m equations in n variables, and let $A^\#$ be the augmented coefficient matrix of the system.

If $r \neq r^\#$ then any row-echelon form for $A^\#$ will have exactly $r^\#$ non-zero rows, and so the corresponding system has only $r^\#$ equations. Then there are $n - r^\#$ free variables, and the system has infinitely many solutions. This was the case in our last example. We can actually say more.

Theorem: If A is an $m \times n$ matrix, and $A^\#$ is the augmented matrix of the system $A\bar{x} = \bar{b}$ let $r = rK(A)$, and $r^\# = rK(A^\#)$.

- (1) If $r < r^\#$ the system is inconsistent (since there will be at least one row of the reduced row-echelon form of $A^\#$ whose only non-zero entry will be in the last column. Such a row corresponds to the equation $0 = \alpha$ which is impossible).
- (2) If $r = r^\#$ the system is consistent. Moreover, if $r^\# < n$ there are infinitely many solutions. If $r^\# = n$ there is a unique solution (since there are no free variables).

We've already noticed that a homogeneous system is always consistent since it has $\bar{x} = \bar{0}$ for a solution. Part (2) of the above theorem tells us that if $A\bar{x} = \bar{0}$ is a system of m equations in n variables, and $m \leq n$, then the system

has infinitely many solutions.

Your turn - (1) Find the reduced row echelon form for $A^\#$ for the system

$$x + 2y + 3z = 1$$

$$2x + 3y + z = 0$$

$$7x + 12y + 11z = 3$$

Then find the solution set for this system.

Answer:

$$\left(\begin{array}{ccc|c} 1 & 0 & -7 & -3 \\ 0 & 1 & 5 & 2 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

$$x - 7z = -3$$

So,

$$y + 5z = 2$$

The variable z is a free variable, and x and y are bound variables. If we let $z = t$ we find $\{(-3 + 7t, 2 - 5t, t) : t \in \mathbb{R}\}$ is the solution set.

(2) Same questions for the system

$$x + y + z = 1$$

$$2x + y - z = 0$$

$$5x + 3y - z = -1$$

(5)

Answer: The reduced row-echelon form is

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

The system is inconsistent. ($\text{rk}(A) = 2 < \text{rk}(A^*) = 3$)

Inverse of a square matrix

Let A be an $n \times n$ matrix, and let I_n be the $n \times n$ identity matrix (i.e., the matrix whose main diagonal entries are 1, and whose off-diagonal entries are 0)

$$I_n = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$$

An $n \times n$ matrix B is called the inverse of the matrix A if $AB = BA = I_n$.

Such a matrix B may not exist but when it does we say that A is invertible, and we say that B is the inverse of A . We will denote B by A^{-1} .

If A is ^{not} invertible we say that A is singular (so invertible matrices are sometimes called non-singular).

Remarks: (1) If B exists it is unique, i.e., if C is another $n \times n$ matrix satisfying $AC = CA = I_n$ then $B = C$.

To see this suppose $AB=BA=I$
and $AC=CA=I$

Then $C(AB) = C \cdot I = C$

$$\stackrel{||}{(CA)B} = I \cdot B = B$$

(2) If B is the inverse of A then A is the inverse of B

(3) If $AB=I$ and $CA=I$ then $B=C$.

To see this we write

$$C(AB) = C \cdot I = C$$

$$\stackrel{||}{(CA)B} = I \cdot B = B$$

(4) If A^{-1} exists then A is one-to-one, and the system $A\bar{x} = \bar{b}$ has a unique solution for every vector $\bar{b}$. The solution is $\bar{x} = A^{-1}\bar{b}$ (since $A\bar{x} = A(A^{-1}\bar{b}) = (AA^{-1})\bar{b} = I \cdot \bar{b} = \bar{b}$)

Next time we'll prove the following.

Theorem: The $n \times n$ matrix A is invertible if and only if $\text{rk}(A) = n$.

Lecture 5

Theorem: The $n \times n$ matrix A is invertible if and only if $\text{rk}(A) = n$.

proof:

If A is invertible then, for each vector $\bar{b}$, the system $A\bar{x} = \bar{b}$ has the unique solution $\bar{x} = A^{-1}\bar{b}$. We know that if such a system has a unique solution then $\text{rk}(A) = \text{rk}(A^*) = n$.

Conversely, let's assume $\text{rk}(A) = n$.

We write $\bar{e}_i$ for the standard unit vector that has 1 for its i^{th} entry, and 0 for its other entries. Since $\text{rk}(A) = n$ we can solve the system $A\bar{x} = \bar{e}_i$, and our solution will be unique. Let $\bar{y}_i$ be the solution (so that $A\bar{y}_i = \bar{e}_i$), and let $Y = [\bar{y}_1, \dots, \bar{y}_n]$ be the $n \times n$ matrix whose i^{th} column is $\bar{y}_i$.

$$\begin{aligned} \text{Then } AY &= A[\bar{y}_1, \dots, \bar{y}_n] = \\ &= [A\bar{y}_1, \dots, A\bar{y}_n] = [\bar{e}_1, \dots, \bar{e}_n] = I. \end{aligned}$$

Remark (4) of Lecture (4) tells us that $YA = I$ as well.

As a corollary, we note that if $A\bar{x} = \bar{b}$ has a unique solution for any vector $\bar{b}$ then A is invertible (since if $A\bar{x} = \bar{b}$ has a unique solution we know $\text{rk}(A) = n$).

Remarks: (1) If A and B are invertible so is AB , and
 $(AB)^{-1} = B^{-1}A^{-1}$ (since $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$).

(2) If A and B are $n \times n$ matrices and AB is invertible then A and B are each invertible

proof: We know that AB is invertible if and only if $\text{rk}(AB) = n$

If B is not invertible (so $\text{rk}(B) < n$) then $B\bar{x} = \bar{0}$ has a solution $\bar{y} \neq \bar{0}$. Then $(AB)\bar{y} = A(B\bar{y}) = A\bar{0} = \bar{0}$, so $\text{rk}(AB) < n$, and AB is not invertible.

If B is invertible, but A is not then $\text{rk}(A) < n$ so $A\bar{x} = \bar{0}$ has a non-trivial solution $\bar{z} \neq \bar{0}$. Since B is invertible $B\bar{x} = \bar{z}$ has a unique solution $\bar{u}$ (so $B\bar{u} = \bar{z}$). Then $\bar{0} = A\bar{z} = A(B\bar{u}) = (AB)\bar{u}$ so

$\text{rk}(AB) < n$ (since $\bar{u} \neq \bar{0}$). Once again, we see that AB is not invertible.

So AB is only invertible if A and B are each invertible.

(3) $(A^T)^{-1} = (A^{-1})^T$, i.e., A^T is invertible if A is invertible. (since $A^T(A^{-1})^T = (A^{-1}A)^T = I_n^T = I_n$)

How do we find A^{-1} ? The answer is contained in the proof of the theorem with which we began today's lecture. We solve $A\bar{x} = \bar{e}_i$ for $i = 1, \dots, n$. If the solutions are $\bar{x}_1, \dots, \bar{x}_n$ (so $A\bar{x}_i = \bar{e}_i$) then $A^{-1} = [\bar{x}_1, \dots, \bar{x}_n]$.

Notice that if we solve $A\bar{x} = \bar{e}_i$ using Gauss-Jordan elimination then the steps we take to find the reduced row-echelon form for $A^\#$ are the same for each i . Thus, we can do all of these calculations simultaneously - we write

$$(A | I_n)$$

and reduce to reduced row-echelon form. When the left hand side is reduced to I_n the resulting right-hand side is A^{-1} .

Examples: (i) $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}$

$$\left(\begin{array}{ccc|ccc} 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{array} \right) \xrightarrow{P_{13}} \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \end{array} \right)$$

$$\xrightarrow{P_{23}} \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{array} \right) \xrightarrow{A_{31}(-1)} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & -1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{array} \right)$$

$$\text{So, } A^{-1} = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Let's check - $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

The case of 2×2 invertible matrices will show up in a lot of our examples, so we're going to compute the inverse in general.

Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and assume $ad - bc \neq 0$.

$$\left(\begin{array}{cc|cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array} \right) \xrightarrow{A_{12} \left(-\frac{c}{a} \right)} \left(\begin{array}{cc|cc} a & b & 1 & 0 \\ 0 & d - \frac{bc}{a} & -\frac{c}{a} & 1 \end{array} \right)$$

$$\xrightarrow{M_2 \left(\frac{a}{ad - bc} \right)} \left(\begin{array}{cc|cc} a & b & 1 & 0 \\ 0 & 1 & -\frac{c}{ad - bc} & \frac{a}{ad - bc} \end{array} \right) \xrightarrow{A_{21}(-b)}$$

$$\left(\begin{array}{cc|cc} a & 0 & 1 + \frac{bc}{ad - bc} & \frac{-ab}{ad - bc} \\ 0 & 1 & \frac{-c}{ad - bc} & \frac{a}{ad - bc} \end{array} \right) \xrightarrow{M_1 \left(\frac{1}{a} \right)}$$

$$\left(\begin{array}{cc|cc} 1 & 0 & \frac{d}{ad - bc} & \frac{-b}{ad - bc} \\ 0 & 1 & \frac{-c}{ad - bc} & \frac{a}{ad - bc} \end{array} \right)$$

$$\text{So, } A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

$$\begin{aligned} \text{Let's check - } & \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \cdot \frac{1}{ad - bc} \\ &= \begin{pmatrix} ad - bc & 0 \\ 0 & ad - bc \end{pmatrix} \cdot \frac{1}{ad - bc} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

Now, it's your turn. (1) Find A^{-1} if $A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix}$.

Then solve $A\bar{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$.

Answer: $A^{-1} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1/2 & -1/6 \\ 0 & 0 & 1/3 \end{pmatrix}$

If $A\bar{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ then $\bar{x} = A^{-1} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1/2 & -1/6 \\ 0 & 0 & 1/3 \end{pmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$
 $= \begin{bmatrix} -1 \\ 1/2 \\ 1 \end{bmatrix}$

Notice that A and A^{-1} are each upper triangular. This is true in general. The inverse of an upper (respectively, lower) triangular matrix is upper (respectively, lower) triangular.

Questions: Let $A = \begin{pmatrix} 1 & 2 & 1 & 2 \\ 3 & 1 & 3 & 1 \\ 1 & 4 & 1 & 4 \\ 5 & 7 & 5 & 7 \end{pmatrix}$

Notice that the last row is the sum of the first three rows. Is A invertible? Why?

One more exercise—find A^{-1} if $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix}$.

(6)

Answer: $A^{-1} = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$

Let's check - $AA^{-1} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

Notice that A and A^{-1} are each symmetric. If A is symmetric and invertible then A^{-1} is also symmetric (A is symmetric means $A = A^T$, Then $(A^{-1})^T = (A^T)^{-1} = A^{-1}$).

Lecture 6

Determinants - For us these will be a computational tool given by a formula. However, I would like you to have some idea how determinants arise, so I'll explain this in the simplest case.

First we need a few ideas -

A function $f(x)$ is said to be linear if

$$f(ax+by) = af(x)+bf(y)$$

for all scalars a, b . We assume f takes values in $\mathbb{R}$ or $\mathbb{C}$.

A function f of two variables is said to be bilinear if it is linear in each variable

$$\text{i.e. } f(ax+by, z) = af(x, z) + bf(y, z)$$

$$\text{and } f(x, cy+dz) = cf(x, y) + df(x, z)$$

We say that f is alternating if $f(x, x) = 0$.

The reason for that name is the following.

$$\begin{aligned} 0 &= f(x+y, x+y) = f(x, x+y) + f(y, x+y) \\ &= f(x, x) + f(x, y) + f(y, x) + f(y, y) \end{aligned}$$

The first and last terms $f(x, x)$ and $f(y, y)$ are each 0, so $0 = f(x, y) + f(y, x)$,
i.e., $f(x, y) = -f(y, x)$

We're interested in functions on the columns of 2×2 matrices that are both bilinear and alternating.

(2)

We start with a 2×2 matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and compute $f\left(\begin{pmatrix} a \\ c \end{pmatrix}, \begin{pmatrix} b \\ d \end{pmatrix}\right)$.

$$\text{Now, } \begin{pmatrix} a \\ c \end{pmatrix} = a \begin{pmatrix} 1 \\ 0 \end{pmatrix} + c \begin{pmatrix} 0 \\ 1 \end{pmatrix} = a \bar{e}_1 + c \bar{e}_2$$

$$\text{and } \begin{pmatrix} b \\ d \end{pmatrix} = b \begin{pmatrix} 1 \\ 0 \end{pmatrix} + d \begin{pmatrix} 0 \\ 1 \end{pmatrix} = b \bar{e}_1 + d \bar{e}_2$$

$$\text{So, } f\left(\begin{pmatrix} a \\ c \end{pmatrix}, \begin{pmatrix} b \\ d \end{pmatrix}\right) = f(a \bar{e}_1 + c \bar{e}_2, b \bar{e}_1 + d \bar{e}_2)$$

$$= a f(\bar{e}_1, b \bar{e}_1 + d \bar{e}_2) + c f(\bar{e}_2, b \bar{e}_1 + d \bar{e}_2)$$

$$= ab f(\bar{e}_1, \bar{e}_1) + ad f(\bar{e}_1, \bar{e}_2) + cb f(\bar{e}_2, \bar{e}_1) + cd f(\bar{e}_2, \bar{e}_2)$$

from the bilinearity of f .

Since f is alternating we know $f(\bar{e}_1, \bar{e}_1) = 0$
and $f(\bar{e}_2, \bar{e}_2) = 0$. Moreover, $f(\bar{e}_2, \bar{e}_1) = -f(\bar{e}_1, \bar{e}_2)$,
so $f\left(\begin{pmatrix} a \\ c \end{pmatrix}, \begin{pmatrix} b \\ d \end{pmatrix}\right) = (ad - bc) \cdot f(\bar{e}_1, \bar{e}_2)$.

Thus, the value of f on the columns of $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ will depend on $f(\bar{e}_1, \bar{e}_2)$.

The determinant of a 2×2 matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is the unique bilinear alternating function on the columns of $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ that takes the value 1 on the identity matrix $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.
(Note that the columns of $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ are $\bar{e}_1$ and $\bar{e}_2$.)

We define the determinant of an $n \times n$ matrix in an analogous way. The determinant of an $n \times n$ matrix is a real (or complex) valued function on the columns of the $n \times n$ matrix A that is linear in each variable (we say f is multilinear), alternating (meaning that if two columns of A are identical the f takes the value 0), and that takes the value 1 on the identity matrix. A calculation then shows that we can compute the determinant of A in the following way.

The minor M_{ij} of A is the determinant of the $(n-1) \times (n-1)$ submatrix obtained from A by deleting the i^{th} row and j^{th} column of A . Then the determinant, $\det(A)$, of A is

$$\det(A) = \sum_{k=1}^n (-1)^{i+k} a_{ik} M_{ik}$$

The term $(-1)^{i+k} M_{ik}$ is sometimes called the ik^{th} cofactor.

The above expression for $\det(A)$ is called the expansion along the i^{th} row.

We can also expand along the j^{th} column.

$$\det(A) = \sum_{k=1}^n (-1)^{k+j} a_{kj} M_{kj}$$

The proof that the above expressions are the determinant relies on the notion of permutations to keep track of interchanging columns

Examples:

(1) $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Let's expand along the first column.

$$\det(A) = a \cdot (-1)^{1+1} \cdot d + c \cdot (-1)^{1+2} \cdot b = ad - bc$$

(2) $\det \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$. Let's expand along the first row.

$$\det(A) = 1 \cdot (-1)^{1+1} \cdot \det \begin{pmatrix} 5 & 6 \\ 8 & 9 \end{pmatrix} + 2 \cdot (-1)^{1+2} \cdot \det \begin{pmatrix} 4 & 6 \\ 7 & 9 \end{pmatrix}$$

$$+ 3 \cdot (-1)^{1+3} \cdot \det \begin{pmatrix} 4 & 5 \\ 7 & 8 \end{pmatrix} = (5 \cdot 9 - 6 \cdot 8) + 2(-1) \cdot (4 \cdot 9 - 7 \cdot 6)$$

$$+ 3(4 \cdot 8 - 5 \cdot 7) = -3 + 2 \cdot 6 - 3 \cdot 3 = 0.$$

(3) $\det \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix}$. Let's expand along the bottom row.

$$\det(A) = 6 \cdot (-1)^{3+3} \cdot \det \begin{pmatrix} 1 & 2 \\ 0 & 4 \end{pmatrix} = 6 \cdot 4 \cdot 1 = 24$$

Notice that, for this upper triangular matrix, $\det(A)$ is the product of the diagonal elements. This is true in general - If A is upper (respectively, lower) triangular then $\det(A)$ is the product of the diagonal entries

We'll write $C_{ij} = (-1)^{i+j} M_{ij}$ for the ij^{th} -cofactor.

Then $\sum_{k=1}^n a_{ik} C_{jk} = 0$ if $i \neq j$ (the elements a_{ik} are

the elements of the i^{th} row of A , while the C_{jk} are the cofactors of row j). To see this let's fix values of i and j with $i \neq j$, and

let B be the matrix we get by adding the i^{th} row to the j^{th} row of A (i.e., we perform $A_{ij} \leftarrow A_{ij} + 1$).

Since determinants are multilinear this operation doesn't change the determinant, so

$$\det(A) = \det(B)$$

We then compute $\det(B)$ using the expansion along row j . We get

$$\det(B) = \sum_{k=1}^n (a_{jk} + a_{ik}) C_{jk} = \sum_{k=1}^n a_{jk} C_{jk} + \sum_{k=1}^n a_{ik} C_{jk}$$

The first sum, $\sum a_{jk} C_{jk}$, is $\det(A)$, so we have

$$\det(A) = \det(B) = \det(A) + \sum a_{ik} C_{jk}$$

$$\text{Thus } \sum a_{ik} C_{jk} = 0$$

Similarly, if $i \neq j$,

$$\sum a_{ki} C_{kj} = 0$$

(a_{ki} are the elements of column i , and the C_{kj} are the cofactors of column j).

The adjoint of A is the matrix $\text{adj}(A) = (C_{ij})^T = (C_{ji})$

Theorem: Let A be an $n \times n$ matrix with $\det(A) \neq 0$.

$$\text{Then } \det(A) \cdot A^{-1} = \text{adj}(A).$$

(6)

In other words, we have an explicit expression for A^{-1} :

$$A^{-1} = \frac{1}{\det(A)} \cdot \text{adj}(A).$$

proof: We compute $A \cdot \text{adj}(A) = B = (b_{ij})$.

If $i \neq j$ then the remarks preceding the theorem tell us that $b_{ij} = 0$. If $i = j$ we have

$$b_{ii} = \det A, \text{ so } A \cdot \text{adj}(A) = \begin{pmatrix} \det A & 0 & \dots & 0 \\ 0 & \det A & 0 & \dots & 0 \\ 0 & & & & \vdots \\ \vdots & & & & 0 \\ 0 & \dots & 0 & \det A \end{pmatrix} = (\det A) \cdot I.$$

As an example, we again compute A^{-1} when $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ when $\det(A) = ad - bc \neq 0$.

$$\text{adj}(A) = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}^T$$

$$c_{11} = (-1)^{(1+1)} d = d, \quad c_{12} = (-1)^{(1+2)} c = -c$$

$$c_{21} = (-1)^{(2+1)} b = -b, \quad c_{22} = (-1)^{(2+2)} a = a.$$

$$\text{So, } \text{adj}(A) = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}^T = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

$$\text{and } A^{-1} = \frac{1}{\det(A)} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

(7)

Here is an exercise for you - Compute A^{-1} by computing $\text{adj}(A)$ if $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix}$

Answer: $\text{adj}(A) = \begin{pmatrix} 24 & -12 & -2 \\ 0 & 6 & -5 \\ 0 & 0 & 4 \end{pmatrix}$

So $A^{-1} = \frac{1}{24} \begin{pmatrix} 24 & -12 & -2 \\ 0 & 6 & -5 \\ 0 & 0 & 4 \end{pmatrix}$

Lecture 7

Here's one more application of determinants

Cramer's Rule: Let A be an $n \times n$ matrix with $\det(A) \neq 0$, and let $\bar{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$.

The solution to the system $A\bar{x} = \bar{b}$ is given by

$$x_i = \frac{\det(B_i)}{\det(A)} \quad (i=1, \dots, n), \text{ where } B_i \text{ is the}$$

matrix obtained from A by replacing the i^{th} column of A with the vector $\bar{b}$.

proof:

$$\det(B_i) = \det \begin{pmatrix} a_{11} & \dots & a_{1,i-1} & b_1 & a_{1,i+1} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & & a_{n,i-1} & b_n & a_{n,i+1} & \dots & a_{nn} \end{pmatrix}$$

$$\begin{aligned} \text{But } b_1 &= a_{11}x_1 + \dots + a_{1n}x_n \\ b_2 &= a_{21}x_1 + \dots + a_{2n}x_n \\ &\vdots \\ b_n &= a_{n1}x_1 + \dots + a_{nn}x_n \end{aligned}$$

We substitute this into the i^{th} column, and use the fact that the determinant is multilinear and alternating. This yields $\det(B_i) = x_i \cdot \det(A)$.

To make this explicit we write out the case of a 2×2 matrix.

Let $A = \begin{pmatrix} a & e \\ c & d \end{pmatrix}$, and let $\bar{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$. We want to solve the system $A\bar{x} = \bar{b}$.

The matrix B_1 is $\begin{pmatrix} b_1 & e \\ b_2 & d \end{pmatrix}$, so

$$\det(B_1) = \det \begin{pmatrix} b_1 & e \\ b_2 & d \end{pmatrix} = \det \begin{pmatrix} ax_1 + ex_2 & e \\ cx_1 + dx_2 & d \end{pmatrix}$$

$$= \det \begin{pmatrix} ax_1 & e \\ cx_1 & d \end{pmatrix} + \det \begin{pmatrix} ex_2 & e \\ dx_2 & d \end{pmatrix} \quad (\text{since det is multilinear})$$

$$= x_1 \det \begin{pmatrix} a & e \\ c & d \end{pmatrix} + x_2 \det \begin{pmatrix} e & e \\ d & d \end{pmatrix}.$$

The last term is 0 since det is alternating.

$$\text{So } \det(B_1) = x_1 \det(A) \Leftrightarrow x_1 = \frac{\det(B_1)}{\det(A)}.$$

$$\text{Similarly, } x_2 = \frac{\det(B_2)}{\det(A)}.$$

Remarks: (i) $\det(A) = \det(A^T)$ (since ^{when we} can expand

using the multilinearity of det, every term in the sum is a product of n elements, one from each row, and each column of A . The sign of the term is 1 or -1 , depending upon how many ^{times} we must switch adjacent columns to end up with I_n .

(3)

The answer is the same for A and A^T .

Again, the simplest case is $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. We've already computed $\det(A)$. Let's compute $\det(A^T) = \det \begin{pmatrix} a & c \\ b & d \end{pmatrix}$

$$\det \begin{pmatrix} a & c \\ b & d \end{pmatrix} = \det(ae_1 + be_2, ce_1 + de_2)$$

$$= a \det(e_1, ce_1 + de_2) + b \det(e_2, ce_1 + de_2)$$

$$= ac \det(e_1, e_1) + ad \det(e_1, e_2) + bc \det(e_2, e_1) + bd \det(e_2, e_2)$$

$$= ad \det(e_1, e_2) - bc \det(e_1, e_2)$$

$$= ad - bc = \det(A).$$

(2) If A and B are $n \times n$ matrices then
 $\det(AB) = \det(A) \cdot \det(B)$

To see this suppose Δ is another multilinear, alternating functional on the columns of $n \times n$ matrices (a functional is a scalar-valued function). Fix an $n \times n$ matrix A . If $\bar{v}_1, \dots, \bar{v}_n$ are n -vectors set $A(\bar{v}_1, \dots, \bar{v}_n) = (\bar{w}_1, \dots, \bar{w}_n)$. Then $\Delta(\bar{w}_1, \dots, \bar{w}_n) = \det(A) \cdot \Delta(\bar{v}_1, \dots, \bar{v}_n)$ by the same sort of calculation used to compute $\det(A)$ (i.e., using multilinearity, and the alternating property).

(4)

Now, let B be an $n \times n$ matrix, and let $B(\bar{e}_1, \dots, \bar{e}_n) = (\bar{u}_1, \dots, \bar{u}_n)$.

We take $D = \det$ so

$$\begin{aligned} \det(AB) &= \det(AB(\bar{e}_1, \dots, \bar{e}_n)) = \det(A(u_1, \dots, u_n)) = \\ &= \det(A) \cdot \det(\bar{u}_1, \dots, \bar{u}_n) = \det(A) \cdot \det(B(\bar{e}_1, \dots, \bar{e}_n)) \\ &= \det(A) \cdot \det(B). \end{aligned}$$

(3) If A is invertible then $\det(A^{-1}) = \frac{1}{\det A}$

$$(\text{since } \det(A \cdot A^{-1}) = \det(A) \cdot \det(A^{-1}))$$

$$\det(I) = 1$$

$$\text{So } \det(A^{-1}) = \frac{1}{\det(A)}.$$

(4) Two $n \times n$ matrices A and B are said to be similar if there is an invertible matrix S such that $A = S^{-1}BS$. If A and B are similar then $\det A = \det B$ (since $\det(A) = \det(S^{-1}BS) = \det(S^{-1}) \cdot \det(B) \cdot \det(S) = \det(B)$).

Vector Spaces

We recall the basic properties of vectors in $\mathbb{R}^n$

- (1) If $\bar{x} = (x_1, \dots, x_n)$, $\bar{y} = (y_1, \dots, y_n) \in \mathbb{R}^n$ then $\bar{x} + \bar{y} = (x_1 + y_1, \dots, x_n + y_n) \in \mathbb{R}^n$.
- (2) If $\bar{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$, and $a \in \mathbb{R}$ then $a\bar{x} = (ax_1, \dots, ax_n) \in \mathbb{R}^n$.

We describe (1) by saying that $\mathbb{R}^n$ is closed under vector addition, and we describe (2) by saying $\mathbb{R}^n$ is closed under scalar multiplication.

(3) The vector $\bar{0} = (0, \dots, 0)$ satisfies $\bar{x} + \bar{0} = \bar{0} + \bar{x} = \bar{x}$ for all $\bar{x} \in \mathbb{R}^n$. We say that $\bar{0}$ is the additive identity of $\mathbb{R}^n$.

(4) (Commutativity) $\bar{x} + \bar{y} = \bar{y} + \bar{x}$ for all $\bar{x}, \bar{y} \in \mathbb{R}^n$

(5) (Associativity) $\bar{x} + (\bar{y} + \bar{z}) = (\bar{x} + \bar{y}) + \bar{z}$ for all $\bar{x}, \bar{y}, \bar{z}$

(6) (Distributive laws) $r(\bar{x} + \bar{y}) = r\bar{x} + r\bar{y}$ for all $r \in \mathbb{R}, \bar{x}, \bar{y} \in \mathbb{R}^n$.

$(r+s)\bar{x} = r\bar{x} + s\bar{x}$ for all $r, s \in \mathbb{R}, \bar{x} \in \mathbb{R}^n$.

(7) $1 \cdot \bar{x} = \bar{x}$ for all $\bar{x} \in \mathbb{R}^n$, and $(rs)\bar{x} = r(s\bar{x})$ for all $r, s \in \mathbb{R}, \bar{x} \in \mathbb{R}^n$.

(8) (Existence of an inverse) For each $\bar{x} \in \mathbb{R}^n$ the vector $-\bar{x} (= -1 \cdot \bar{x})$ satisfies $\bar{x} + (-\bar{x}) = \bar{0}$.

We want to study other sets that satisfy the above properties. For example, the set $\mathcal{C}[0,1]$ of all continuous functions on the interval $[0,1]$ with addition defined by

$$(f+g)(x) = f(x) + g(x)$$

satisfies the above properties (with $\bar{0}$ being the function that is identically zero).

Definition: Let V be a non-empty set that is equipped with an addition, and a scalar multiplication by elements of $F = \mathbb{R}$ (or $\mathbb{C}$) satisfying

(1) (Closure) if $\bar{x}, \bar{y} \in V$ then $\bar{x} + \bar{y} \in V$

(2) (Closure under scalar multiplication) if $\bar{x} \in V, r \in F$ then $r\bar{x} \in V$

- (3) There is an (identity) element $\bar{0} \in V$ satisfying $\bar{x} + \bar{0} = \bar{0} + \bar{x} = \bar{x}$ for all $\bar{x} \in V$.
- (4) $\bar{x} + \bar{y} = \bar{y} + \bar{x}$ for all $\bar{x}, \bar{y} \in V$.
- (5) $\bar{x} + (\bar{y} + \bar{z}) = (\bar{x} + \bar{y}) + \bar{z}$ for all $\bar{x}, \bar{y}, \bar{z} \in V$.
- (6) $r(\bar{x} + \bar{y}) = r\bar{x} + r\bar{y}$ for all $r \in F, \bar{x}, \bar{y} \in V$.
 $(r+s)\bar{x} = r\bar{x} + s\bar{x}$ for all $r, s \in F, \bar{x} \in V$.
- (7) $1 \cdot \bar{x} = \bar{x}$ for all $\bar{x} \in V$.
- (8) For each $\bar{x} \in V$ there is an inverse $(-\bar{x})$ satisfying $\bar{x} + (-\bar{x}) = \bar{0}$.

Remark: The vector space V depends upon the choice of scalars. For example, $\mathbb{C}$ is a vector space with scalar field $\mathbb{C}$, but it is also a vector space with scalar field $\mathbb{R}$. We will see that these two vector spaces are quite different.

Basic Properties: (1) The identity $\bar{0}$ is unique, i.e., if $\bar{\alpha}$ is a vector that also satisfies $\bar{\alpha} + \bar{x} = \bar{x} + \bar{\alpha} = \bar{x}$ for all $\bar{x} \in V$ then $\bar{\alpha} = \bar{0}$. To see this take $\bar{x} = \bar{0}$ so we have $\bar{\alpha} + \bar{0} = \bar{0} + \bar{\alpha} = \bar{0}$. On the other hand, by the defining property (3) of $\bar{0}$ we have $\bar{x} + \bar{0} = \bar{0} + \bar{x} = \bar{x}$ for all $\bar{x}$. So, if we now take $\bar{x} = \bar{\alpha}$ we get $\bar{\alpha} + \bar{0} = \bar{0} + \bar{\alpha} = \bar{\alpha}$. Thus, $\bar{0} = \bar{\alpha}$.

- (2) $0 \cdot \bar{x} = \bar{0}$ for all $\bar{x} \in V$. (since $0+0=0$, so $0 \cdot \bar{x} = (0+0) \cdot \bar{x} = 0\bar{x} + 0\bar{x}$. Subtracting $0\bar{x}$ from each side gives us $\bar{0} = 0 \cdot \bar{x}$.)

Lecture 8

We begin with more examples of vector spaces over $\mathbb{R}$.

- (1) V is the set of all 2×2 symmetric matrices with real entries. This is a subspace since

$$\begin{pmatrix} a & c \\ c & b \end{pmatrix} + \begin{pmatrix} d & f \\ f & e \end{pmatrix} = \begin{pmatrix} a+d & c+f \\ c+f & b+e \end{pmatrix}$$

$$\text{and } r \begin{pmatrix} a & c \\ c & b \end{pmatrix} = \begin{pmatrix} ra & rc \\ rc & rb \end{pmatrix} \text{ and } 1 \begin{pmatrix} a & c \\ c & b \end{pmatrix} = \begin{pmatrix} a & c \\ c & b \end{pmatrix}$$

are each symmetric. The various associative and distributive laws hold for all matrices, and, in particular, they hold for symmetric matrices. Note that $1 \cdot \begin{pmatrix} a & c \\ c & b \end{pmatrix} = \begin{pmatrix} a & c \\ c & b \end{pmatrix}$.

- (2) V is the set of continuous even functions on the interval $[-1, 1]$ (the function f is even if $f(x) = f(-x)$).

$$\text{If } f, g \in V \text{ then } (f+g)(x) = f(x) + g(x) = f(-x) + g(-x) = (f+g)(-x) \text{ and } rf(-x) = rf(x) \text{ if } r \in \mathbb{R}.$$

Once again, the various associative and distributive laws hold for all functions, and so they hold for even functions. Note that $1 \cdot f = f$.

- (3) Let A be an $m \times n$ matrix. The set of solutions to a homogeneous system $A\bar{x} = \bar{0}$ is a vector space since if $\bar{u}$ and $\bar{v}$ are

solutions (i.e., $A\bar{u} = \bar{0}$ and $A\bar{v} = \bar{0}$) then

$$A(\bar{u} + \bar{v}) = \bar{0}$$

$$\text{and } A(r\bar{v}) = rA\bar{v} = \bar{0}$$

so $\bar{u} + \bar{v}$ and $r\bar{v}$ are solutions. The various associative and distributive laws hold for all vectors in $\mathbb{R}^n$, and so hold for the solutions to $A\bar{x} = \bar{0}$. Note that $1 \cdot \bar{u} = \bar{u}$.

The set of solutions to $A\bar{x} = \bar{0}$ is called the kernel or nullspace of A , and is denoted by $\text{Ker } A$.

(4) We again suppose A is an $m \times n$ matrix. The set $\{\bar{b} \in \mathbb{R}^m : \text{there exists } \bar{u} \in \mathbb{R}^n \text{ with } A\bar{u} = \bar{b}\}$ $= \{A\bar{u} : \bar{u} \in \mathbb{R}^n\}$ is a vector space called the image of A (and is denoted by $\text{Im}(A)$).

To see $\text{Im}(A)$ is a vector space suppose $\bar{b}, \bar{c} \in \text{Im}(A)$. Then $\bar{b} = A\bar{u}$ and $\bar{c} = A\bar{v}$ for some vector $\bar{u}, \bar{v} \in \mathbb{R}^n$, so

$$\bar{b} + \bar{c} = A\bar{u} + A\bar{v} = A(\bar{u} + \bar{v}) \in \text{Im}(A)$$

$$\text{and } r\bar{b} = rA\bar{u} = A(r\bar{u}) \in \text{Im}(A).$$

Note that $1 \cdot \bar{u} = \bar{u}$.

(5) $V = \{a \cos \theta + b \sin \theta : a, b \in \mathbb{R}\}$ is a vector space with addition $(a_1 \cos \theta + b_1 \sin \theta) + (a_2 \cos \theta + b_2 \sin \theta)$ $= (a_1 + a_2) \cos \theta + (b_1 + b_2) \sin \theta$ and scalar multiplication $r(a \cos \theta + b \sin \theta)$ $= ra \cos \theta + rb \sin \theta$

(6) This example looks a little different from the previous ones.

The space V is $\mathbb{R}_+ = \{r \in \mathbb{R} : r > 0\}$, the space of positive real numbers. We write $\bar{r}$ for the element of V corresponding to the real number r .

The addition law is $\bar{u} + \bar{v} = \overline{u \cdot v}$

Scalar multiplication is $r\bar{u} = \overline{u^r}$

The identity element is $\bar{1}$

The inverse $-\bar{u}$ is $\overline{\left(\frac{1}{u}\right)}$.

We check, for example, the distributive laws,

$$r(\bar{u} + \bar{v}) = \overline{(u \cdot v)^r} = \overline{u^r \cdot v^r} = r\bar{u} + s\bar{v}$$

$$\text{and } (r+s)\bar{u} = \overline{u^{(r+s)}} = \overline{u^r \cdot u^s} = r\bar{u} + s\bar{u}$$

Now, you try some. Which of the following are vector spaces?

(1) $V = \{ax^2 + b : a, b \in \mathbb{R}\}$.

(2) $V = \{f(x) \in C[0,1] : f(0) = 2\}$

(3) $V = \{(x, 1, y) : x, y \in \mathbb{R}\}$.

(4) $V = \{(x, y, z) : 3x + 2y + 5z = 0\}$

(5) $V = \{(x, y, z) : 3x + 2y + 5z^2 = 0\}$

(6) $V = \{(x, y, z) : 3x + 2y + 5z = 1\}$.

(7) $V = \{f \in C[0,1] : \int_0^1 f(x) dx = 0\}$

The course is about to become more abstract so this is a good time to stop, and review what we've done so far.

The first part of the course is concerned with matrix arithmetic and the solutions to systems of linear equations.

We defined the sum of 2 $m \times n$ matrices $A = (a_{ij})$ and $B = (b_{ij})$ by $A+B = (a_{ij}+b_{ij})$, and we defined

scalar multiplication of a matrix A by $rA = (ra_{ij})$. In our language of the last lecture this turns the set of $m \times n$ matrices with real entries into a real vector space $M_{m \times n}(\mathbb{R})$.

We viewed an $m \times n$ matrix A as a function

$$A: \mathbb{R}^n \rightarrow \mathbb{R}^m$$

$$\text{since } (*) \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ a_{21}x_1 + \dots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{bmatrix} \in \mathbb{R}^m.$$

Using (*) we defined the product of an $m \times n$ matrix A and an $n \times s$ matrix B as follows.

We write $B^{(i)}$ for the i^{th} column of B , so $B = [B^{(1)}, \dots, B^{(n)}]$. Then $AB = [AB^{(1)}, \dots, AB^{(n)}]$ is the matrix whose i^{th} column is $AB^{(i)}$ (and the

(5)

product $AB^{(i)}$ is defined as in $(*)$.

Viewing A as a function $\mathbb{R}^n \rightarrow \mathbb{R}^m$ and B as a function $\mathbb{R}^s \rightarrow \mathbb{R}^n$ then the product AB is the $m \times s$ matrix that, when viewed as a function, is the function $A \circ B$ (the composition of the function B with the function A)

$$\begin{array}{c} \mathbb{R}^s \xrightarrow{B} \mathbb{R}^n \xrightarrow{A} \mathbb{R}^m \\ \quad \quad \quad \searrow \quad \nearrow \\ \quad \quad \quad A \circ B \end{array}$$

This shows that matrix multiplication is associative (i.e., $A(BC) = (AB)C$ when all products are defined.)

It is important to remember that, even when AB and BA are defined, AB is not usually equal to BA .

We noticed that matrix multiplication is distributive over matrix addition $A(B+C) = AB+AC$

$$\text{and } (B+C)A = BA+CA$$

(when the products are defined.)

We turned our attention to solving systems of linear equations. With this in mind we introduced Gaussian elimination and Gauss-Jordan elimination to place matrices in a simpler form.

Lecture 9

We studied systems $A\bar{x} = \bar{b}$. The matrix A is the coefficient matrix of the system, and the matrix $A^\# = (A|\bar{b})$ is the augmented coefficient matrix.

We were allowed to perform any of the following elementary row operations.

(1) $M_i(k)$: Multiply the i^{th} row by the scalar k .

(2) $A_{ij}(k)$: Add k times the i^{th} row to row j .

(3) P_{ij} : Permute rows i and j .

Using these operations we placed our matrix in row-echelon form B , i.e., our matrix B satisfies the following.

(1) Any row consisting only of zeros lies below all non-zero rows.

(2) The first non-zero entry of a non-zero row is a 1 (called a leading 1).

(3) The leading 1 in any row below the first row occurs to the right of the leading 1 in the above rows.

The process of placing A in row-echelon form is called Gaussian elimination. The number of non-zero rows in any row-echelon form of A is the same. This number is called the

rank of A , denoted $\text{rk}(A)$.

A matrix in row-echelon form is in reduced row-echelon form if any column containing a leading 1 has only zeros for its other entries. The reduced row echelon form of a matrix is unique.

The process of applying elementary row operations to a matrix to place it in reduced row-echelon form is called Gauss-Jordan elimination.

If we let r be the rank of A , and $r^\#$ the rank of the augmented coefficient matrix $A^\#$ then, if A is an $m \times n$ matrix,

- (1) If $r < r^\#$ the system is inconsistent (i.e., it has no solutions)
- (2) If $r = r^\#$ the system is consistent. Moreover, if $r < n$ there are infinitely many solutions. If $r = n$ there is a unique solution.

One of our most important uses of Gauss-Jordan elimination was to find the inverse of an invertible $n \times n$ matrix. Recall that an $n \times n$ matrix A is invertible if there is an $n \times n$ matrix B with $AB = BA = I$, the $n \times n$ identity matrix. We write A^{-1} for B , and call it the inverse of A . It is unique when it exists, and in that case we say

(3)

A is invertible.

If A is invertible then $A\bar{x} = \bar{b}$ has a unique solution for every $\bar{b} \in \mathbb{R}^n$, namely $\bar{x} = A^{-1}\bar{b}$. Consequently, the rank of A must be n .

If B is a left inverse for A , i.e., if $BA = I$, then B is also a right inverse for A (so $B = A^{-1}$). To see this note that $BA = I$ means $A\bar{x} = \bar{b}$ has a unique solution for every $\bar{b} \in \mathbb{R}^n$ (So A is both one-to-one and onto when viewed as a function $\mathbb{R}^n \xrightarrow{A} \mathbb{R}^n$).

If $\text{rk}(B) < n$ then $B\bar{x} = \bar{0}$ has a non-trivial solution $\bar{z}$ (so $B\bar{z} = \bar{0}$ but $\bar{z} \neq \bar{0}$). Let $\bar{u} \in \mathbb{R}^n$ be the vector with $A\bar{u} = \bar{z}$. Then $\bar{0} = B\bar{z} = BA\bar{u} = \bar{u}$.

Since $\bar{u} = \bar{0}$ we know $A\bar{u} = \bar{0}$. But $A\bar{u} = \bar{z}$ which was taken to be a non-zero solution to $B\bar{x} = \bar{0}$.

This contradiction shows $\text{rk}(B)$ must be n , so B is one-to-one and onto. Then B has a left inverse C . We have $CB = I$ (and $BA = I$)

so $(CB)A = I \cdot A = A$ Thus, $A = C$, and

$$\overset{''}{C}(BA) = C \cdot I = C$$

A is both a left and a right inverse for B , and so B is a left and right inverse for A .

We checked that if A and B are $n \times n$ matrices and (AB) is invertible so are A and B . In this case $(AB)^{-1} = B^{-1}A^{-1}$.

We find the inverse of a non-singular (i.e., invertible) matrix by solving the systems $Ax = \bar{e}_i$ ($i=1, \dots, n$). Explicitly, we construct a "big" augmented matrix

$$(A | I_n).$$

We apply elementary row operations to reduce the left hand side to I_n (which is the reduced row-echelon form of an invertible matrix). The right hand side is A^{-1} .

We observed that if A is upper (respectively, lower) triangular, and invertible, then A^{-1} is also upper (respectively, lower) triangular.

Also, $(A^T)^{-1} = (A^{-1})^T$, so A^T is invertible if A is, and A^{-1} is symmetric if A is.

We computed the inverse of the 2×2 matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ using the above method and found it to be $\frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

The expression $ad-bc$ is the determinant of the matrix A . We defined determinants for an $n \times n$ matrix in the following way. The determinant, $\det A$, of the $n \times n$ matrix A is the unique multilinear, alternating scalar valued function on the columns of A that

(5)

satisfies $\det(I_n) = 1$ (where I_n is the $n \times n$ identity matrix).

Recall that multilinear means that $\det$ is linear in each variable.

Alternating means that if two columns of A are identical then $\det(A) = 0$. We use the word alternating to describe this because if $f(\bar{x}_1, \dots, \bar{x}_n)$ is a multilinear alternating $\mathbb{R}$ (or $\mathbb{Q}$) - valued function then the value of f changes sign each time we switch two adjacent columns (so the sign alternates). Using only the definition (and the idea of permutations) one checks that we can compute $\det(A)$ in the following ways.

The ij -minor M_{ij} of A is the determinant of the submatrix obtained from A by deleting the i^{th} row and j^{th} column of A . The ij -cofactor C_{ij} is $(-1)^{(i+j)} M_{ij}$. Then

$$\det A = \sum_{k=1}^n a_{ik} C_{ik} = \sum_{k=1}^n a_{kj} C_{kj}$$

For triangular matrices A $\det(A)$ is the product of the diagonal entries, $\det A = \prod_{i=1}^n a_{ii}$.

The adjoint of A is the transpose of the matrix of cofactors: $\text{adj}(A) = (C_{ij})^T = (C_{ji})$. We checked that, if $\det(A) \neq 0$, the matrix A is non-singular, and $A^{-1} = \frac{1}{\det(A)} \cdot \text{adj}(A)$.

(6)

We saw that $\det(A) = \det(A^T)$, and if A and B are $n \times n$ matrices, $\det(AB) = \det(A) \cdot \det(B)$. As a consequence we have $\det(A^{-1}) = \frac{1}{\det(A)}$.

Finally, we proved Cramer's Rule: Let A be an $n \times n$ invertible matrix, and let $\bar{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$.

The solution to the system $A\bar{x} = \bar{b}$ is given by

$$x_i = \frac{\det(B_i)}{\det(A)} \quad (i=1, \dots, n), \text{ where } B_i \text{ is the}$$

$n \times n$ matrix obtained from A by replacing the i^{th} column of A with the vector $\bar{b}$.

Lecture 10

There are some subsets of an F -vector space V that are particularly important. These are the subsets that are themselves F -vector spaces with the same addition and scalar multiplication as in V .

Definition: Let V be an F -vector space. A subset W of V is called a subspace of V if W is an F -vector space under the same operations (vector addition and scalar multiplication) used in V .

The last example in Lecture 7

$$V = \{(x, y) \in \mathbb{R}^2 : 2x + 3y = 0\}$$

is a subspace of $\mathbb{R}^2$, while the two non-examples

$$V = \{(x, y, z) \in \mathbb{R}^3 : 2x + y - z = 1\} \text{ and}$$

$$V = \{(x, y, z) \in \mathbb{R}^3 : 2x^2 + y + z = 0\}$$

are not subspaces of $\mathbb{R}^3$.

In each of these cases the deciding factor is whether or not

V is closed under both vector addition and scalar multiplication.

In fact, this is always the case as the following result shows.

Theorem: Let W be a subset of the F -vector space V . Then W is a subspace of V if and only if

(2)

W is closed under vector addition and scalar multiplication.

proof:

If W is a subspace then it is, by definition, closed under these two operations.

On the other hand, if W is closed under both operations then $0 \cdot \bar{x} = \bar{0}$ and $(-1) \cdot \bar{x} = -\bar{x}$ are in W if $\bar{x} \in W$. Thus, W contains the identity element, and the inverse of each vector in W . The various associative, distributive, and commutative laws are inherited from V (i.e., they hold in V so they hold in every subset of V).

If W is a subspace of $\mathbb{R}^n$ defined by algebraic equations then $\bar{0} \in W$, i.e., the hyperplanes (or planes or lines) defined by the equations must pass through the origin, so the equations cannot have a non-zero constant term. Moreover, for vector addition and scalar multiplication to hold each of the equations must be linear.

Let's look at more examples

(i) The subset $\{\bar{0}\}$ is always a subspace since it satisfies the conditions in the above theorem.

(2) $V = M_2(\mathbb{R}) = 2 \times 2$ matrices with real entries.

$W = 2 \times 2$ symmetric matrices contained in V .

Then W is a subspace since

$$\begin{pmatrix} a & c \\ c & b \end{pmatrix} + \begin{pmatrix} d & f \\ f & e \end{pmatrix} = \begin{pmatrix} a+d & c+f \\ c+f & b+e \end{pmatrix} \text{ is symmetric}$$

$$\text{and } r \begin{pmatrix} a & c \\ c & b \end{pmatrix} = \begin{pmatrix} ra & rc \\ rc & rb \end{pmatrix} \text{ is symmetric.}$$

(3) $V = \mathcal{C}([-1, 1]) =$ the real vector space of continuous functions on the interval $[-1, 1]$, W is the subset of functions $f(x) \in V$ satisfying $f(x) = f(-x)$ (i.e., f is an even function).

W is a subspace since, if $f(x) = f(-x)$ and $g(x) = g(-x)$, then $(f+g)(x) = f(x) + g(x) = f(-x) + g(-x) = (f+g)(-x)$ is even, and $rf(-x) = rf(x)$.

(4) $V = M_2(\mathbb{R})$, $W =$ all invertible matrices in V . Then W is not a subspace since the zero matrix $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, which is the identity in V , is not invertible.

(5) Let A be an $m \times n$ matrix, and let $\bar{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} \neq \bar{0}$. The set of solutions to $A\bar{x} = \bar{b}$ is not a subspace of $\mathbb{R}^n$ since it doesn't contain $\bar{0}$.

(6) The set of solutions to $A\bar{x} = \bar{0}$ is a subspace of $\mathbb{R}^n$ since if $\bar{u}, \bar{v}$ are each solutions then $A(\bar{u} + \bar{v}) = A\bar{u} + A\bar{v} = \bar{0} + \bar{0} = \bar{0}$ and $A(r\bar{u}) = rA\bar{u} = \bar{0}$, so $\bar{u} + \bar{v}$ and $r\bar{u}$ are each solutions.

This subspace is called the null space or Kernel of A , denoted by $\text{Ker}(A)$.

In other words, $\text{Ker}(A) = \{\bar{x} \in \mathbb{R}^n : A\bar{x} = \bar{0}\}$.

(7) If A is an $m \times n$ matrix the set $\text{Image of } A = \{\bar{b} \in \mathbb{R}^m : A\bar{x} = \bar{b} \text{ for some vector } \bar{x} \in \mathbb{R}^n\}$ is a subspace of $\mathbb{R}^m$ (denoted by $\text{Im}(A)$).

To see this suppose $A\bar{x} = \bar{b}$ and $A\bar{y} = \bar{c}$.

Then $\bar{b} + \bar{c} = A\bar{x} + A\bar{y} = A(\bar{x} + \bar{y}) \in \text{Im}(A)$

and $r\bar{b} = rA\bar{x} = A(r\bar{x}) \in \text{Im}(A)$.

Let V be a vector space over $F = \mathbb{R}$ or $\mathbb{C}$.

A set $\{\bar{v}_1, \dots, \bar{v}_n\}$ of vectors in V is said to span V if every vector $\bar{x} \in V$ can be written as $\bar{x} = c_1\bar{v}_1 + \dots + c_n\bar{v}_n$ for some scalars $c_1, \dots, c_n \in F$.

We'll also say that $\{\bar{v}_1, \dots, \bar{v}_n\}$ generate V , or that V is the span of $\{\bar{v}_1, \dots, \bar{v}_n\}$.

The expression $c_1\bar{v}_1 + \dots + c_n\bar{v}_n$ is called a linear combination of the vectors $\bar{v}_1, \dots, \bar{v}_n$.

Examples:

(1) The vectors $\bar{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\bar{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, and $\bar{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ span $\mathbb{R}^3$ since $\begin{bmatrix} a \\ b \\ c \end{bmatrix} = a\bar{e}_1 + b\bar{e}_2 + c\bar{e}_3$.

(2) The vectors $\bar{e}_1, \dots, \bar{e}_n$ span $\mathbb{R}^n$.

(3) The vectors $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ span $\mathbb{R}^2$ since $\bar{e}_1 = \frac{1}{2}(\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \end{bmatrix})$ and $\bar{e}_2 = \frac{1}{2}(\begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ -1 \end{bmatrix})$, so

$$\begin{aligned} \begin{bmatrix} a \\ b \end{bmatrix} &= a\bar{e}_1 + b\bar{e}_2 = a \cdot \frac{1}{2}(\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \end{bmatrix}) + b \cdot \frac{1}{2}(\begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ -1 \end{bmatrix}) \\ &= \frac{a+b}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{a-b}{2} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \end{aligned}$$

(4) The polynomials $1, x$, and x^2 span the space $\mathcal{P}_2$ of polynomials of degree ≤ 2 .

So do the polynomials $1+x$, $1-x$, and $x+x^2$ since $1 = \frac{1}{2}((1+x) + (1-x))$, $x = \frac{1}{2}((1+x) - (1-x))$

$$\text{and } x^2 = x+x^2 - x = x+x^2 - \frac{1}{2}((1+x) - (1-x))$$

(5) If $\bar{v}_1, \dots, \bar{v}_n$ span V so does any set of vectors that contains $\{\bar{v}_1, \dots, \bar{v}_n\}$.

Remark: Let V be an F -vector space, and let $\{\bar{v}_1, \dots, \bar{v}_n\}$ be vectors in V . The span of $\{\bar{v}_1, \dots, \bar{v}_n\}$ is a subspace of V (since if $a_1\bar{v}_1 + \dots + a_n\bar{v}_n$ and $b_1\bar{v}_1 + \dots + b_n\bar{v}_n \in \text{span}(V)$)

(6)

so is $(a_1 \bar{v}_1 + \dots + a_n \bar{v}_n) + (b_1 \bar{v}_1 + \dots + b_n \bar{v}_n) = (a_1 + b_1) \bar{v}_1 + \dots + (a_n + b_n) \bar{v}_n$

and so is $r(a_1 \bar{v}_1 + \dots + a_n \bar{v}_n) = r a_1 \bar{v}_1 + \dots + r a_n \bar{v}_n$ for any scalar r .

Question: How can we tell if 2 vectors in $\mathbb{R}^2$ span $\mathbb{R}^2$?

If $\begin{bmatrix} a \\ b \end{bmatrix}$ and $\begin{bmatrix} c \\ d \end{bmatrix}$ span $\mathbb{R}^2$ then we can write an arbitrary vector $\begin{bmatrix} e \\ f \end{bmatrix}$ as $d_1 \begin{bmatrix} a \\ b \end{bmatrix} + d_2 \begin{bmatrix} c \\ d \end{bmatrix}$,

i.e. we can find values of x_1 and x_2 satisfying

$$a x_1 + c x_2 = e$$

$$b x_1 + d x_2 = f$$

for each choice of e and f . We know we can do this if and only if the rank of the coefficient matrix $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$ is two, and we know this happens if and only if $\det \begin{pmatrix} a & c \\ b & d \end{pmatrix} \neq 0$
 " $ad - bc$

Here's one for you - Does the vector $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ lie in the span of $\begin{bmatrix} 3 \\ 5 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 6 \end{bmatrix}$?

Lecture 11

Recall that $\{\bar{v}_1, \dots, \bar{v}_n\}$ span the F -vector space W if every vector $\bar{x}$ in W can be written as an F -linear combination $\bar{x} = c_1 \bar{v}_1 + \dots + c_n \bar{v}_n$.

More examples:

(1) The matrices $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, and $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ span $M_2(\mathbb{R})$, the space of 2×2 matrices with real entries since

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = a \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + c \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + d \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

(2) Let V be the vector space $C^{(\infty)}(0,1)$, the space of infinitely differentiable functions on the interval $(0,1)$, and let W be the subspace of V spanned by e^x and e^{-x} . Then W is also spanned by $\cosh x = \frac{e^x + e^{-x}}{2}$ and $\sinh x = \frac{e^x - e^{-x}}{2}$

since $e^x = \cosh x + \sinh x$ and $e^{-x} = \cosh x - \sinh x$.

(3) Let V be the $\mathbb{R}$ -vector space of 2×2 symmetric matrices. $V = \left\{ \begin{pmatrix} a & c \\ c & b \end{pmatrix} : a, b, c \in \mathbb{R} \right\}$.

Then V is spanned by $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, and $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

since $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = a \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + c \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

(2)

Problem: Find a set of vectors that spans the vector space of 2×2 skew symmetric matrices.

Notation: If $\bar{v}_1, \dots, \bar{v}_n$ lie in the F -vector space V we write $\langle \bar{v}_1, \dots, \bar{v}_n \rangle$ for the subspace of V spanned by $\bar{v}_1, \dots, \bar{v}_n$. Thus, $\langle \bar{v}_1, \dots, \bar{v}_n \rangle$ is the set of all linear combinations $c_1 \bar{v}_1 + \dots + c_n \bar{v}_n$.

We know that if a set S of vectors spans a vector space V then any set of vectors of V that contains S also spans V , so if we enlarge a spanning set the bigger set still spans V . In the other direction we may ask for a minimal spanning set for V , i.e., a set T of spanning vectors with the property that, if we eliminate any vector from T , the resulting set no longer spans V .

For example, the vectors $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$, and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ span $V = \mathbb{R}^2$. If we eliminate any of these vectors the remaining two vectors still span $\mathbb{R}^2$. However, if we eliminate one more vector then the final remaining vector no longer spans $\mathbb{R}^2$. So, any of $\{\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}\}$, $\{\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}\}$, or $\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}\}$ is a minimal spanning set.

Remark (4) If the vector space V is spanned by

a finite collection S of vectors then some subset of S is a minimal spanning set for V .

More precisely,

(2) If $\{\bar{v}_1, \dots, \bar{v}_n\} = S \text{ span } W$, and any one of the $\bar{v}_i$, say $\bar{v}_j$, is a linear combination of the others then we may eliminate $\bar{v}_j$ from S and the remaining vectors still span W .

Important Definition: A finite, non-empty set $\{\bar{v}_1, \dots, \bar{v}_n\}$ of vectors in V is said to be linearly independent if whenever $c_1\bar{v}_1 + \dots + c_n\bar{v}_n = \bar{0}$ then $c_1 = c_2 = \dots = c_n = 0$.

For example, the vectors $\bar{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\bar{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, and $\bar{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ are linearly independent since

$$c_1\bar{e}_1 + c_2\bar{e}_2 + c_3\bar{e}_3 = \bar{0} \quad \text{means}$$

$$\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \bar{0} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

$$\text{so } c_1 = c_2 = c_3 = 0.$$

Notice that two vectors $\bar{v}$ and $\bar{w}$ are linearly independent if and only if $\bar{v}$ is not a scalar multiple of $\bar{w}$ since, if $a\bar{v} + b\bar{w} = \bar{0}$ and $b \neq 0$, then $\bar{w} = -\frac{a}{b}\bar{v}$.

If $\bar{v}_1, \dots, \bar{v}_n$ are not linearly independent we say they are linearly dependent. In that case, there are scalars $c_1, \dots, c_n$, not all zero,

with $c_1 \bar{v}_1 + \dots + c_n \bar{v}_n = 0$. In this case we can write one of the $\bar{v}_i$ as a linear combination of the others. For example, if $c_1 \neq 0$ then $\bar{v}_1 = -\frac{1}{c_1}(c_2 \bar{v}_2 + \dots + c_n \bar{v}_n)$.

Here are some problems for you. Decide if the following vectors are dependent or independent.

(1) $V = \mathbb{R}^3$, $S = \{ \bar{v}_1 = \bar{e}_1, \bar{v}_2 = \bar{e}_1 + \bar{e}_2, \bar{v}_3 = \bar{e}_2 + \bar{e}_3 \}$

(2) $V = \mathcal{P}_2$, the space of polynomials of degree ≤ 2 .
 $S = \{ 1, x^2 + x + 1, x - 1 \}$

$T = \{ 1, x^2 + 1, x + 1, x^2 + x \}$.

(3) $V = M_2(\mathbb{R})$, $S = \left\{ \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}, \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \right\}$.

(4) $V = \mathbb{R}^3$, $S = \left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \right\}$.

One more problem - Let $\bar{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$, $\bar{v}_2 = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$, $\bar{v}_3 = \begin{bmatrix} 6 \\ 5 \\ 3 \end{bmatrix}$, $\bar{v}_4 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. Find an independent subset of $\bar{v}_1, \bar{v}_2, \bar{v}_3, \bar{v}_4$ that span the same space as $\bar{v}_1, \bar{v}_2, \bar{v}_3, \bar{v}_4$.

Remarks: (i) If $\bar{v}_1, \dots, \bar{v}_k \in \mathbb{R}^n$ then $\bar{v}_1, \dots, \bar{v}_k$ are linearly dependent if and only if the system

$$[\bar{v}_1, \dots, \bar{v}_k] \bar{x} = \bar{0}$$

has a nontrivial solution. If we write $\bar{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix}$ then $[\bar{v}_1, \dots, \bar{v}_k] \bar{x} = \bar{0}$ is the same as $x_1 \bar{v}_1 + \dots + x_k \bar{v}_k = \bar{0}$.

A non-trivial solution to $[\bar{v}_1, \dots, \bar{v}_k] \bar{x} = \bar{0}$ gives us a relation of linear dependence among the $\bar{v}_i$.

(2) A consequence of (1) is that, if $k > n$, the vectors $\bar{v}_1, \dots, \bar{v}_k$ are linearly dependent since $[\bar{v}_1, \dots, \bar{v}_k] \bar{x} = \bar{0}$ has infinitely many solutions.

(3) If $k = n$ then $\bar{v}_1, \dots, \bar{v}_k$ are linearly independent if and only if $\det[\bar{v}_1, \dots, \bar{v}_k] \neq 0$. In this case, $[\bar{v}_1, \dots, \bar{v}_k] \bar{x}$ has the unique solution $\bar{x} = \bar{0}$.

Examples: (1) $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 4 \\ 1 \end{bmatrix}, \begin{bmatrix} 6 \\ 4 \\ 2 \end{bmatrix}, \begin{bmatrix} 19 \\ 12 \\ 3 \end{bmatrix}$ are dependent

since $k = 4 > n = 3$.

(2) The vectors $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 4 \\ 1 \end{bmatrix}, \begin{bmatrix} 6 \\ 4 \\ 2 \end{bmatrix}$ are independent

since $\det \begin{pmatrix} 1 & 3 & 6 \\ 2 & 4 & 4 \\ 1 & 1 & 2 \end{pmatrix} \neq 0$.

Let A be an $n \times n$ matrix. We know that the columns of A are independent if and only if $\det(A) \neq 0$. Since $\det(A) = \det(A^T)$ we also know that the rows of A are independent if and only if the columns of A are independent.

(6)

Let A be an $m \times n$ matrix, and write $A^{(i)}$ for the i^{th} column of A . (so $A = [A^{(1)}, \dots, A^{(n)}]$). If $\vec{c} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \in \mathbb{R}^n$ then $A\vec{c} = c_1 A^{(1)} + \dots + c_n A^{(n)}$, so the image of A , $\text{Im}(A) = \{A\vec{c} : \vec{c} \in \mathbb{R}^n\}$ is the space spanned by the columns of A . It is, as we have seen, a subspace of $\mathbb{R}^m$. The number of independent columns is the rank of A .

As an example, consider the 3×8 matrix

$$A = \begin{pmatrix} 1 & 2 & 3 & 5 & 7 & 8 \\ 0 & 1 & 1 & 1 & 9 & 7 \\ 0 & 0 & 1 & 1 & 2 & 19 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The matrix A is in row-echelon form so we can see that its rank is 3. Without doing any additional work we know that there are only 3 independent columns. Clearly, the first three columns are independent, so these span the image of A .

Lecture 12

The set of functions $\{f_1(x), \dots, f_n(x)\}$ is said to be linearly independent on the interval I if

$$c_1 f_1(x) + \dots + c_n f_n(x) = 0 \text{ for all } x \in I$$

implies $c_1 = c_2 = \dots = c_n = 0$.

A useful tool for deciding the linear independence of functions is the following.

Suppose $f_1, \dots, f_n \in \mathcal{C}^{(n-1)}(I)$, the space of $(n-1)$ -times continuously differentiable functions on I . The Wronskian of $f_1, \dots, f_n$ is

$$W(f_1, \dots, f_n)(x) = \det \begin{pmatrix} f_1(x) & \dots & f_n(x) \\ f_1'(x) & \dots & f_n'(x) \\ \vdots & & \vdots \\ f_1^{(n-1)}(x) & \dots & f_n^{(n-1)}(x) \end{pmatrix}.$$

Of course, if we switch two adjacent columns the Wronskian changes sign since the determinant is alternating. However, we are primarily interested in whether or not there is a point in I at which the Wronskian is non-zero. This is explained in the following.

Theorem: Let $f_1, \dots, f_n \in \mathcal{C}^{(n-1)}(I)$. If $W(f_1, \dots, f_n)$ is non-zero at some point x_0 in I then the f_i are linearly independent on I .

proof: If $c_1 f_1(x) + c_2 f_2(x) + \dots + c_n f_n(x) = 0$ for all $x \in I$ we differentiate $(n-1)$ times to get the system

$$c_1 f_1(x) + \dots + c_n f_n(x) = 0$$

$$c_1 f_1'(x) + \dots + c_n f_n'(x) = 0$$

$$\vdots$$

$$c_1 f_1^{(n-1)}(x) + \dots + c_n f_n^{(n-1)}(x) = 0$$

This is a system of n equations in the n unknowns $c_1, \dots, c_n$. The Wronskian is the determinant of the coefficient matrix, so if it is non-zero at the point $x_0 \in I$ the system has only the trivial solution $c_1 = c_2 = \dots = c_n = 0$, i.e., the functions $f_1(x), \dots, f_n(x)$ are linearly independent on I .

Remark: If $W(f_1, \dots, f_n) = 0$ for all $x \in I$ then the theorem doesn't tell us anything.

Examples: (1) The functions e^x and e^{2x} are linearly independent on $(-\infty, \infty)$ since

$$W(e^x, e^{2x}) = \det \begin{pmatrix} e^x & e^{2x} \\ e^x & 2e^{2x} \end{pmatrix} = 2e^{3x} - e^{3x} = e^{3x},$$

and so $W(e^x, e^{2x})$ is not zero at any point.

(2) The functions $\cos x$ and $\sin x$ are independent on $(-\infty, \infty)$ since $W(\cos x, \sin x) = \det \begin{pmatrix} \cos x & \sin x \\ -\sin x & \cos x \end{pmatrix} =$

$$\cos^2 x + \sin^2 x = 1.$$

$$(3) \text{ Let } f(x) = x^3 \text{ and } g(x) = \begin{cases} 3x^3 & x \geq 0 \\ -x^3 & x < 0 \end{cases}.$$

$$\text{Then, for } x \geq 0, \quad W(f, g) = \det \begin{pmatrix} x^3 & 3x^3 \\ 3x^2 & 9x^2 \end{pmatrix} = 9x^5 - 9x^5 = 0,$$

$$\text{and for } x < 0, \quad W(f, g) = \det \begin{pmatrix} x^3 & -x^3 \\ 3x^2 & -3x^2 \end{pmatrix} = -3x^5 + 3x^5 = 0.$$

So, the Wronskian is identically zero on $(-\infty, \infty)$.
The theorem fails to tell us anything. Let's examine $f(x)$ and $g(x)$ more carefully.

On $[0, \infty)$ we have $-3f(x) + g(x) = 0$; so $f(x)$ and $g(x)$ are linearly dependent on $[0, \infty)$.

On $(-\infty, 0)$ we have $f(x) + g(x) = 0$, so $f(x)$ and $g(x)$ are dependent on $(-\infty, 0)$.

However, we can not find two non-zero constants c_1 and c_2 satisfying $c_1 f(x) + c_2 g(x) = 0$ on $(-\infty, \infty)$, so $f(x)$ and $g(x)$ actually are independent on $(-\infty, \infty)$.

Problems for you —

(1) For which values of a and b are the functions e^{ax} and e^{bx} independent on $(-\infty, \infty)$.

(2) Find the Wronskian of the functions

$$f(x) = 2x^4 \quad g(x) = \begin{cases} 8x^4 & \text{if } x \geq 0 \\ -x^4 & \text{if } x < 0 \end{cases}.$$

Determine if $f(x)$ and $g(x)$ are independent on $(-\infty, \infty)$.

(3) Show that the vectors $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$, and $\begin{bmatrix} 7 \\ 8 \\ 5 \end{bmatrix}$ are linearly dependent, and that any two of these vectors are linearly independent.

Find an equation for the plane spanned by these vectors.

We have seen that a minimal spanning set of vectors for a vector space V must consist of linearly independent vectors.

On the other hand, a maximal linearly independent set of vectors must span V .

To see this let's suppose there is a vector $\bar{x} \in V$ that does not lie in $\langle \bar{v}_1, \dots, \bar{v}_k \rangle$.

Since $\{\bar{v}_1, \dots, \bar{v}_k\}$ is maximal the vectors

$\bar{x}, \bar{v}_1, \dots, \bar{v}_k$ must be linearly dependent, i.e.

there are scalars $c_0, c_1, \dots, c_k$, not all 0, with

$$c_0 \bar{x} + c_1 \bar{v}_1 + \dots + c_k \bar{v}_k = \bar{0}$$

Moreover, $c_0 \neq 0$ since $\bar{v}_1, \dots, \bar{v}_k$ are independent.

Then $\bar{x} = -\frac{1}{c_0} (c_1 \bar{v}_1 + \dots + c_k \bar{v}_k)$, i.e. $\bar{x}$ does lie

in $\langle \bar{v}_1, \dots, \bar{v}_k \rangle$. Every vector in V lies in the span of $\bar{v}_1, \dots, \bar{v}_k$.

Remark: In the above we assumed V can be spanned by finitely many vectors. This is not always the case

Important Definition: A set $\{\bar{v}_1, \dots, \bar{v}_k\}$ of vectors in a vector space V is called a basis of V if $\bar{v}_1, \dots, \bar{v}_k$ span V and are linearly independent.

Examples: (1) $\bar{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\bar{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $\bar{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ are a basis for $\mathbb{R}^3$.

(2) $V = P_2$, the space of real valued polynomials of degree ≤ 2 . The set $\{1, x, x^2\}$ is a basis for P_2 .

(3) Any 3 independent vectors in $\mathbb{R}^3$ are a basis for $\mathbb{R}^3$. Why?

The last example suggests the following questions

(1) Suppose the vector space V has a finite basis. Are all bases of V necessarily finite?

(2) If B and C are two bases for the vector space V do B and C have the same number of vectors in them, i.e. if $B = \{\bar{v}_1, \dots, \bar{v}_k\}$, $C = \{\bar{w}_1, \dots, \bar{w}_t\}$ are each bases of V is $k=t$.

We will see that the answers to (1) and (2) are each yes, and we define the dimension of V (denoted $\dim(V)$) to be the cardinality of any basis of V .

Examples: (1) $\dim \mathbb{R}^n = n$ since $\{\bar{e}_1, \dots, \bar{e}_n\}$ is a basis.

(2) $\dim P_n = n+1$ since $\{1, x, x^2, \dots, x^n\}$ is a basis.

Question: What is $\dim(M_2(\mathbb{R}))$?

Lectures 13 (and 14, part 1)

Suppose that V is an F -vector space with a finite basis consisting of the m elements $\bar{v}_1, \dots, \bar{v}_m$.

Theorem: Every basis of V has exactly m elements.

proof:

Among all finite bases for V there is at least one with a minimal number of elements. We may take m to be this minimal number.

We prove that if $\{\bar{w}_1, \dots, \bar{w}_n\}$ is a set of vectors in V with $n > m$ then the vectors $\bar{w}_1, \dots, \bar{w}_n$ are linearly dependent.

We write $\bar{w}_1 = a_{11}\bar{v}_1 + \dots + a_{m1}\bar{v}_m$

$$\bar{w}_2 = a_{12}\bar{v}_1 + \dots + a_{m2}\bar{v}_m$$

$\vdots$

$$\bar{w}_n = a_{1n}\bar{v}_1 + \dots + a_{mn}\bar{v}_m$$

Now, suppose $c_1\bar{w}_1 + \dots + c_n\bar{w}_n = \bar{0}$ for some scalars $c_1, \dots, c_n$.

$$\text{Then } c_1(a_{11}\bar{v}_1 + \dots + a_{m1}\bar{v}_m) + \dots + c_n(a_{1n}\bar{v}_1 + \dots + a_{mn}\bar{v}_m) = \bar{0}$$

Now, we collect like terms:

$$(a_{11}c_1 + a_{12}c_2 + \dots + a_{1n}c_n)\bar{v}_1 + (a_{21}c_1 + a_{22}c_2 + \dots + a_{2n}c_n)\bar{v}_2 \\ + \dots + (a_{m1}c_1 + a_{m2}c_2 + \dots + a_{mn}c_n)\bar{v}_m = \bar{0}$$

Since the $\bar{v}_i$ are independent each of the parenthetical coefficients must be 0, i.e.

$$a_{11}c_1 + a_{12}c_2 + \dots + a_{1n}c_n = 0$$

$$a_{21}c_1 + a_{22}c_2 + \dots + a_{2n}c_n = 0$$

$$\vdots$$

$$a_{m1}c_1 + a_{m2}c_2 + \dots + a_{mn}c_n = 0$$

This is a homogeneous system of m equations in the n unknowns $c_1, \dots, c_n$. Since $n > m$ we know that such a system has non-trivial solutions, i.e., there are constants $c_1, \dots, c_n$, not all zero with $c_1\bar{w}_1 + \dots + c_n\bar{w}_n = \bar{0}$. Thus, the $\bar{w}_i$ are linearly dependent.

So, if $\bar{w}_1, \dots, \bar{w}_n$ is a basis of V we see that n must equal m , i.e., all bases have the same cardinality. (called the dimension of V).

Remarks: (1) The dimension of a vector space depends on the scalar field. For example, if $V = \mathbb{C}$ and $F = \mathbb{C}$ then $\{1\}$ is a basis for $\mathbb{C}$, so $\mathbb{C}$ is one-dimensional as a $\mathbb{C}$ -vector space. However, if $V = \mathbb{C}$ and $F = \mathbb{R}$ then $\{1, i\}$ is a basis for $\mathbb{C}$, so $\mathbb{C}$ is two dimensional as an $\mathbb{R}$ -vector space.

(2) There are vector spaces that are not finite dimensional. For example, the vector space $\mathcal{C}(-\infty, \infty)$ of continuous functions on $(-\infty, \infty)$ contains the linearly independent functions

$1, x, x^2, \dots, x^n, \dots$ The proof of the theorem shows $C(-\infty, \infty)$ is infinite dimensional.

We make a few observations

- (1) If $\dim V = n$ then any set S that spans V must contain at least n vectors. Otherwise, there is a subset of V that spans V and consists of linearly independent vectors. This subset is then a basis of V but contains fewer than n elements, an impossibility.
- (2) If $\dim V = n$ then any set S of n linearly independent vectors is a basis of V . Otherwise there is a vector $\bar{x} \in V$ that is not a linear combination of the vectors in S . The set $S \cup \{\bar{x}\}$ is then a set of $(n+1)$ linearly independent vectors in V , an impossibility.
- (3) If $\dim V = n$ then any set S of n vectors that span V is a basis. To see this notice that if S is not a basis then the vectors in S are not linearly independent. In that case we may eliminate some of the vectors in S to obtain a linearly independent spanning set, i.e. a basis of V . This is impossible since $\dim V = n$.

(4) Suppose $\dim V = n$, and that W is a subspace of V . Then $\dim W \leq n$. To see this note that a basis for W is a set of linearly independent vectors in V that span W . Any set of linearly independent vectors in V has cardinality $\leq n$.

(5) Any set S of linearly independent vectors in V is part of a basis of V . To see this suppose the cardinality of $S = \dim V$. Then S spans V and so is a basis. If the cardinality of S is less than $\dim V$ then there is a vector $\bar{x} \in V$ that is not in the span of S , and so is independent of the vectors in S . Then $\{\bar{x}\} \cup S$ is a set of independent vectors in V . If the cardinality of $\{\bar{x}\} \cup S$ is $\dim V$ then $\{\bar{x}\} \cup S$ is a basis for V . Otherwise, we repeat this process until we have obtained a basis.

As an example, we have seen that the skew-symmetric matrices are a subspace of $M_2(\mathbb{R})$, the space of 2×2 matrices with real entries. We have also seen that $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ is a basis for the skew symmetric matrices; and that $\dim M_2(\mathbb{R}) = 4$. The set $\left\{ \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \right\}$ is a basis for $M_2(\mathbb{R})$. Why?

(5)

Problems for you -

(1) Find a basis for the null space of $\begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{pmatrix}$.

(2) Let W be the subspace of $\mathbb{R}^3$ defined by the equation $x + 2y + 3z = 0$.

(3) Show that the matrices in $M_2(\mathbb{R})$ that have trace zero form a subspace of $M_2(\mathbb{R})$, and find a basis for it.

(4) Find a basis for the space of 2×2 symmetric matrices with real entries.

(5) The vectors $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ are linearly independent. Find a vector $\bar{x} \in \mathbb{R}^3$ such that $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and $\bar{x}$ are a basis for $\mathbb{R}^3$.

(1)

More problems (lecture 14, part 2)

- (1) a) Find a basis for the space of 4×4 symmetric matrices.
 b) Find a basis for the space of 4×4 skew-symmetric matrices.
 c) Show that every matrix is the sum of a symmetric matrix and a skew symmetric matrix.

- (2) For which values of a is the set
 $\{2+ax^2, 1+x+x^2, 1-x\}$
 a basis for P_2

- (3) Find a basis for the column space of the matrix

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 & 6 \\ 0 & 3 & 1 & 5 & 8 \\ 0 & 0 & 2 & 9 & 9 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

- (4) For which values of a is the set
 $\left\{ \begin{bmatrix} 1 \\ a \\ 1 \end{bmatrix}, \begin{bmatrix} a \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$
 a basis of $\mathbb{R}^3$.

- (5) Find a basis for $\mathbb{R}^3$ that contains the vectors $\begin{bmatrix} 8 \\ 5 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 6 \\ 7 \end{bmatrix}$

Lecture 15

Let V be a vector space with basis $\{\bar{v}_1, \dots, \bar{v}_n\}$. Every vector $\bar{x} \in V$ can be expressed as

$$\bar{x} = c_1 \bar{v}_1 + \dots + c_n \bar{v}_n$$

Since $\bar{v}_1, \dots, \bar{v}_n$ span V . Moreover, this expression is unique, i.e., $c_1, \dots, c_n$ are uniquely determined by $\bar{x}$. To see this suppose we can write $\bar{x}$ in two ways as a linear combination of $\bar{v}_1, \dots, \bar{v}_n$, say

$$\bar{x} = c_1 \bar{v}_1 + \dots + c_n \bar{v}_n = d_1 \bar{v}_1 + \dots + d_n \bar{v}_n.$$

$$\text{Then } \underbrace{c_1 \bar{v}_1 + \dots + c_n \bar{v}_n - (d_1 \bar{v}_1 + \dots + d_n \bar{v}_n)}_{(c_1 - d_1) \bar{v}_1 + \dots + (c_n - d_n) \bar{v}_n} = \bar{x} - \bar{x} = \bar{0}$$

Since $\bar{v}_1, \dots, \bar{v}_n$ are linearly independent we must have $(c_1 - d_1) = (c_2 - d_2) = \dots = (c_n - d_n) = 0$, i.e., $c_i = d_i$ for $i = 1, \dots, n$.

Now we suppose our basis is ordered (i.e., the order in which we write the $\bar{v}_i$ matter so while $\{\bar{v}_1, \bar{v}_2, \bar{v}_3, \dots, \bar{v}_n\}$ and $\{\bar{v}_2, \bar{v}_1, \bar{v}_3, \dots, \bar{v}_n\}$ are the same basis of V they are different ordered bases of V).

If $B = \{\bar{v}_1, \dots, \bar{v}_n\}$ is an ordered basis of V , and $\bar{x} \in V$ with $\bar{x} = c_1 \bar{v}_1 + \dots + c_n \bar{v}_n$ we call the n -tuple $\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$ the component vector of $\bar{x}$ with respect to B , and c_i is the i^{th} component of $\bar{x}$

(2)

with respect to B we write $[\bar{x}]_B$ for $\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$.

For example, if $V = P_2$, $B = \{1, x, x^2\}$,
 $C = \{x, 1, x^2\}$ and $D = \{x+1, x-1, x^2\}$, then

if $\bar{f} = 2x^2 + 2x - 2$ we have

$$[\bar{f}]_B = \begin{bmatrix} -2 \\ 2 \\ 2 \end{bmatrix}, [\bar{f}]_C = \begin{bmatrix} 2 \\ -2 \\ 2 \end{bmatrix}, [\bar{f}]_D = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix}$$

If $V = \mathbb{R}^3$, $B = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$ let's

find $\begin{bmatrix} 4 \\ 6 \\ 2 \end{bmatrix}_B$. We need to write $\begin{bmatrix} 4 \\ 6 \\ 2 \end{bmatrix}$ as

$$\begin{bmatrix} 4 \\ 6 \\ 2 \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + c_3 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}. \text{ Then}$$

$$4 = c_1 + c_2$$

$$6 = c_1 + c_3$$

and $2 = c_3$ so $c_1 = 4$ and $c_2 = 0$.

$$\text{Then } \begin{bmatrix} 4 \\ 6 \\ 2 \end{bmatrix}_B = \begin{bmatrix} 4 \\ 0 \\ 2 \end{bmatrix}$$

Here's one for you - Let $V = \mathbb{R}^3$, $B = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\}$

Find $\begin{bmatrix} 8 \\ 4 \\ 2 \end{bmatrix}_B$.

Remark: $[\bar{x} + \bar{y}]_B = [\bar{x}]_B + [\bar{y}]_B$ and $[r\bar{x}]_B = r[\bar{x}]_B$.

You might wonder if there is a relationship between the component vectors associated to a vector $\bar{x}$ with respect to different bases of V . The answer is yes, and is explained in the following.

Suppose $B = \{\bar{v}_1, \dots, \bar{v}_n\}$ and $C = \{\bar{w}_1, \dots, \bar{w}_n\}$ are two bases of V . We may express each $\bar{v}_i$ as a linear combination of $\{\bar{w}_1, \dots, \bar{w}_n\}$, and so associate $[\bar{v}_i]_C$ to $\bar{v}_i$. We build an $n \times n$ matrix by placing $[\bar{v}_i]_C$ as the i^{th} column. This matrix is called

the Change of Basis matrix from B to C . We follow the

book in denoting this matrix by $P_{C \leftarrow B}$.

If $\bar{x} \in V$ write $\bar{x} = a_1 \bar{v}_1 + \dots + a_n \bar{v}_n$, so $[\bar{x}]_B = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$

Then $P_{C \leftarrow B} [\bar{x}]_B = a_1 [\bar{v}_1]_C + \dots + a_n [\bar{v}_n]_C =$

$$[a_1 \bar{v}_1 + \dots + a_n \bar{v}_n]_C = [\bar{x}]_C.$$

i.e., the change of basis matrix from B to C takes the component vector of $\bar{x}$ with respect to B to the component vector of $\bar{x}$ with respect to C .

As a consequence we see $P_{B \leftarrow C}$ takes $[\bar{x}]_C$ to $[\bar{x}]_B$.

Thus, the product $P_{B \leftarrow C} P_{C \leftarrow B}$ takes $[\bar{x}]_B$ to itself for every $\bar{x} \in V$, and so $P_{B \leftarrow C} \cdot P_{C \leftarrow B}$ must be the identity matrix I_n , i.e., $P_{B \leftarrow C}$ and $P_{C \leftarrow B}$ are inverses.

Example: $V = P_2$, the space of polynomials of degree ≤ 2 .
 $B = \{1+x, 1-x, x^2\}$, $C = \{1, 1+x, 1+x^2\}$.
 $\bar{y} = 4x^2 + 6x + 3$.

a) Find $P_{C \leftarrow B}$ and $P_{B \leftarrow C}$

b) Find $[\bar{y}]_B$ and $[\bar{y}]_C$

(a) To find $P_{C \leftarrow B}$ we write the vectors in B as

linear combinations of the vectors in C , and place the component vectors into the columns of $P_{C \leftarrow B}$.

$$1+x = 0 \cdot 1 + 1 \cdot (1+x) + 0 \cdot (1+x^2), \quad [1+x]_C = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$$1-x = 2 \cdot 1 - 1 \cdot (1+x) + 0 \cdot (1+x^2), \quad [1-x]_C = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}$$

$$x^2 = -1 \cdot 1 + 0 \cdot (1+x) + 1 \cdot (1+x^2), \quad [x^2]_C = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

$$P_{C \leftarrow B} = \begin{pmatrix} 0 & 2 & -1 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

To find $P_{B \leftarrow C}$ we write the vectors in C as linear combinations of the vectors in B .

$$1 = \frac{1}{2}(1+x) + \frac{1}{2}(1-x) + 0 \cdot x^2, \quad [1]_B = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \end{bmatrix}$$

$$1+x = 1 \cdot (1+x) + 0 \cdot (1-x) + 0 \cdot x^2, \quad [1+x]_B = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$1+x^2 = \frac{1}{2}(1+x) + \frac{1}{2}(1-x) + 1 \cdot x^2, \quad [1+x^2]_B = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 1 \end{bmatrix}$$

$$P_{B \leftarrow C} = \begin{pmatrix} \frac{1}{2} & 1 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix}$$

As a check we compute the product $P_{B \leftarrow C} P_{C \leftarrow B}$.

$$\begin{pmatrix} \frac{1}{2} & 1 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 2 & -1 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \checkmark$$

(b) To find $[\bar{y}]_B$ we write $\bar{y} = 4x^2 + 6x + 3$ as a linear combination of the vectors in B .

$$\bar{y} = 4x^2 + 6x + 3 = \frac{9}{2}(1+x) - \frac{3}{2}(1-x) + 4 \cdot x^2$$

$$\text{So, } [\bar{y}]_B = \begin{bmatrix} \frac{9}{2} \\ -\frac{3}{2} \\ 4 \end{bmatrix}.$$

(6)

to find $[\bar{y}]_C$ we write $\bar{y}$ as a linear combination of the vectors in C

$$4x^2 + 6x + 3 = \bar{y} = -7 \cdot 1 + 6(1+x) + 4(1+x^2)$$

$$\text{So } [\bar{y}]_C = \begin{bmatrix} -7 \\ 6 \\ 4 \end{bmatrix}.$$

As a check we notice that

$$[\bar{y}]_C = \begin{bmatrix} -7 \\ 6 \\ 4 \end{bmatrix} = \begin{pmatrix} 0 & 2 & -1 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{bmatrix} 9/2 \\ -3/2 \\ 4 \end{bmatrix} = P_{C \leftarrow B} [\bar{y}]_B.$$

Problem: Let $V = \mathbb{R}^3$, $B = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$,

and let $\bar{x} = \begin{bmatrix} 5 \\ 4 \\ 3 \end{bmatrix}$.

(a) Find $[\bar{x}]_B$.

(b) Let $C = \left\{ \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$

Find $P_{C \leftarrow B}$

Lecture 16

Let's begin with a little more practice.

Let $V = \mathbb{R}^3$ $B = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$, $C = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$.

Find $P_{B \leftarrow C}$.

We write each vector in C as a linear combination of the vectors in B .

$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = -1 \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + 1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 0 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \text{ so } \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}_B = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}.$$

$$\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = -1 \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + 0 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \text{ so } \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}_B = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

$$\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = 1 \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} - 1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \text{ so } \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}_B = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}.$$

$$\text{Then } P_{B \leftarrow C} = \begin{pmatrix} -1 & -1 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix}$$

We note that change of basis matrices have a nice type of compatibility. If A, B, C are ordered bases of the vector space V then

$$P_{C \leftarrow A} = P_{C \leftarrow B} \cdot P_{B \leftarrow A}$$

since, if $\bar{x} \in V$,

$$P_{C \in B} P_{B \leftarrow A} [\bar{x}]_A = P_{C \in B} [\bar{x}]_B = [\bar{x}]_C \Rightarrow P_{C \in A} [\bar{x}]_A$$

Row and Column Spaces

If A is an $m \times n$ matrix the row space of A is defined to be the subspace of $\mathbb{R}^n$ spanned by the rows of A .

The column space of A is the subspace of $\mathbb{R}^m$ spanned by the columns of A . We have seen that this is the same as the image of A .

Remark: If A and B are row equivalent matrices they have the same row space, since none of the elementary row operations change the space spanned by the rows. As a result, we note that a basis for the row space of A is the set of non-zero rows in any row-echelon form of A .

Example: Suppose we want to find a basis for the subspace W of $\mathbb{R}^5$ spanned by the vectors

$$\begin{bmatrix} 1 \\ 2 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \text{ and } \begin{bmatrix} 4 \\ 7 \\ 3 \\ 3 \\ 2 \end{bmatrix}$$

We transpose the vectors and place them in the rows of a 3×5 matrix

$$A = \begin{pmatrix} 1 & 2 & 1 & 1 & 1 \\ 2 & 3 & 1 & 1 & 0 \\ 4 & 7 & 3 & 3 & 2 \end{pmatrix}$$

We row-reduce A to find a row-echelon form.

$$\begin{pmatrix} 1 & 2 & 1 & 1 & 1 \\ 2 & 3 & 1 & 1 & 0 \\ 4 & 7 & 3 & 3 & 2 \end{pmatrix} \xrightarrow{A_{12}(-2)} \begin{pmatrix} 1 & 2 & 1 & 1 & 1 \\ 0 & -1 & -1 & -1 & -2 \\ 4 & 7 & 3 & 3 & 2 \end{pmatrix}$$

$$\xrightarrow{A_{13}(-4)} \begin{pmatrix} 1 & 2 & 1 & 1 & 1 \\ 0 & -1 & -1 & -1 & -2 \\ 0 & -1 & -1 & -1 & -2 \end{pmatrix} \xrightarrow{A_{32}(-1)} \begin{pmatrix} 1 & 2 & 1 & 1 & 1 \\ 0 & -1 & -1 & -1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\xrightarrow{M_2(-1)} \begin{pmatrix} 1 & 2 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The first two rows are a basis for the row space of A , so

$$\begin{bmatrix} 1 \\ 2 \\ 1 \\ 1 \\ 1 \end{bmatrix} \text{ and } \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 2 \end{bmatrix}$$

are a basis of W .

Problem: Find a basis for the row space of

$$A = \begin{pmatrix} 1 & 2 & 2 & 2 & 3 \\ 3 & 2 & 2 & 2 & 3 \\ 8 & 8 & 8 & 8 & 12 \end{pmatrix}$$

Now we'd like to find a basis for the column space of an $m \times n$ matrix A . One way would be to transpose A , find a basis for the row space of A^T , and then transpose again. But, if we want bases for both the row and column spaces this is a lot of work. There is an easy way to find a basis for the column space if we have already found a basis for the row space by placing A in row-echelon form.

Recall that elementary row operations don't change the solution set for $A\bar{x} = \bar{0}$. Let's write $A = [A^{(1)} \dots A^{(n)}]$, and suppose $E = [E^{(1)} \dots E^{(n)}]$ is a row-echelon form for A . Then $A\bar{x} = \bar{0}$ if and only if $E\bar{x} = \bar{0}$, so, if $\bar{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, $x_1 A^{(1)} + \dots + x_n A^{(n)} = x_1 E^{(1)} + \dots + x_n E^{(n)} = \bar{0}$.

In other words a linear relation among the columns of A holds if and only if the same relation holds among the columns of E . Thus, the linearly independent columns of E correspond to the linearly independent columns of A . We know that the columns of E that contain leading 1's are a basis for the space

(5)

spanned by the columns of E , so the corresponding columns of A are a basis for the column space of A .

Warning! The columns of E need not be in the column space of A since elementary row operations don't preserve column spaces.

Example: Find a basis for the row space, and a basis for the column space of

$$A = \begin{pmatrix} 2 & 1 & 3 & 4 & 1 \\ 1 & 2 & 1 & 2 & 1 \\ 2 & 1 & 4 & 3 & 1 \\ 5 & 4 & 8 & 9 & 3 \end{pmatrix}$$

We row-reduce to find a row-echelon form for A .

$$\begin{pmatrix} 2 & 1 & 3 & 4 & 1 \\ 1 & 2 & 1 & 2 & 1 \\ 2 & 1 & 4 & 3 & 1 \\ 5 & 4 & 8 & 9 & 3 \end{pmatrix} \xrightarrow{R_{12}} \begin{pmatrix} 1 & 2 & 1 & 2 & 1 \\ 2 & 1 & 3 & 4 & 1 \\ 2 & 1 & 4 & 3 & 1 \\ 5 & 4 & 8 & 9 & 3 \end{pmatrix} \xrightarrow{A_{12}(-2)} \begin{pmatrix} 1 & 2 & 1 & 2 & 1 \\ 0 & -3 & 1 & 0 & -1 \\ 2 & 1 & 4 & 3 & 1 \\ 5 & 4 & 8 & 9 & 3 \end{pmatrix}$$

$$\xrightarrow{A_{13}(-2)} \begin{pmatrix} 1 & 2 & 1 & 2 & 1 \\ 0 & -3 & 1 & 0 & -1 \\ 0 & -3 & 2 & -1 & -1 \\ 5 & 4 & 8 & 9 & 3 \end{pmatrix} \xrightarrow{A_{14}(-5)} \begin{pmatrix} 1 & 2 & 1 & 2 & 1 \\ 0 & -3 & 1 & 0 & -1 \\ 0 & -3 & 2 & -1 & -1 \\ 0 & -6 & 3 & -1 & -2 \end{pmatrix} \xrightarrow{A_{24}(-2)}$$

(6)

$$\begin{pmatrix} 1 & 2 & 1 & 2 & 1 \\ 0 & -3 & 1 & 0 & -1 \\ 0 & -3 & 2 & -1 & -1 \\ 0 & 0 & 1 & -1 & 0 \end{pmatrix} \xrightarrow{A_{23}(-1)} \begin{pmatrix} 1 & 2 & 1 & 2 & 1 \\ 0 & -3 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 & 0 \end{pmatrix} \xrightarrow{A_{34}(-1)} \begin{pmatrix} 1 & 2 & 1 & 2 & 1 \\ 0 & -3 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\xrightarrow{M_2(-1/3)} \begin{pmatrix} 1 & 2 & 1 & 2 & 1 \\ 0 & 1 & -1/3 & 0 & -1/3 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = E$$

The first three columns of the last matrix E are a basis for the column space of E , so the first three columns of A are a basis for the column space of A .

The first three rows of E are a basis for the row space of A .

Problem: Find a basis for the row space and a basis for the column space of the matrix

$$A = \begin{pmatrix} 1 & 3 & 1 & 2 \\ 2 & 3 & 0 & 1 \\ 4 & 9 & 2 & 5 \end{pmatrix}.$$

Lecture 17

Recall that the Kernel or nullspace of a matrix A is $\{\bar{x} : A\bar{x} = \bar{0}\}$. The dimension of the kernel of A is called the nullity of A .

Rank-Nullity Theorem: Let A be an $m \times n$ matrix.
Then $\text{rk}(A) + \text{nullity}(A) = n$.

proof: If A is invertible then $\text{nullity}(A) = 0$
since $\text{Ker}(A) = \{\bar{0}\}$ (the only solution to $A\bar{x} = \bar{0}$ is $\bar{x} = \bar{0}$ in this case), and $\text{rk}(A) = n$.

If A is not invertible, say A has rank r , then, when we solve $A\bar{x} = \bar{0}$, we find $n-r$ free variables. If $x_1, \dots, x_{n-r}$ are the free variables. Let $\bar{v}_i$ be the vector obtained by setting $x_i = 1$, and $x_j = 0$ if $i \neq j$. Then $\bar{v}_1, \dots, \bar{v}_{n-r}$ are linearly independent, and span $\text{Ker}(A)$, i.e., $\{\bar{v}_1, \dots, \bar{v}_{n-r}\}$ is a basis for $\text{Ker}(A)$.

So, $\dim[\text{Ker}(A)] = \text{nullity}(A) = n-r$ and we have
 $\text{rk}(A) + \text{nullity of } A = r + n-r = n$.

Remark: Recall that $\text{rk}(A)$ is the number of independent rows of A , and that this is also the number of independent columns of A . Since the column space of A is the image of A we may rephrase the rank-nullity

(2)

theorem as follows. If A is an $m \times n$ matrix then
 $\dim(\text{Im}(A)) + \dim(\text{Ker}(A)) = n$.

Example! Let $A = \begin{pmatrix} 1 & 3 & 4 & 5 & 9 & 6 \\ 0 & 2 & 7 & 9 & 6 & 8 \\ 0 & 0 & 0 & 3 & 5 & 4 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$

This is a 4×6 matrix in which the rows are clearly independent. So the dimension of the column space is 4, and the dimension of the null space is $6 - 4 = 2$. Without doing any work we can see that the image of A is $\mathbb{R}^4$, and $\text{Ker}(A)$ is a two-dimensional subspace of $\mathbb{R}^6$. (i.e., $\text{Ker}(A)$ is a plane through the origin in $\mathbb{R}^6$.)

Remark! We may rephrase our earlier work on the solutions to a system of linear equations as follows. Let A be an $m \times n$ matrix. The system $A\bar{x} = \bar{b}$ is inconsistent if $\bar{b}$ is not in the column space of A .

If $\bar{b}$ is in the column space of A the system $A\bar{x} = \bar{b}$ is consistent. If the dimension of the column space is n (so the nullity is 0) the system has a unique solution. If the dimension of the column space is less than n the system has infinitely

many solutions.

In the latter case we would like to find all solutions. To do so let's find the relationship between different solutions - Suppose $\bar{u}$ and $\bar{v}$ are each solutions to $A\bar{x} = \bar{b}$, so $A\bar{u} = \bar{b}$ and $A\bar{v} = \bar{b}$.

Then $A\bar{u} - A\bar{v} = A(\bar{u} - \bar{v}) = \bar{b} - \bar{b} = \bar{0}$, i.e., $(\bar{u} - \bar{v}) \in \text{Ker}(A)$.

So, any two solutions to $A\bar{x} = \bar{b}$ differ by an element of $\text{Ker}(A)$. In other words, the general solution to $A\bar{x} = \bar{b}$ is

$$\bar{x} = c_1 \bar{v}_1 + \dots + c_{n-r} \bar{v}_{n-r} + \bar{v}_p$$

where $\{\bar{v}_1, \dots, \bar{v}_{n-r}\}$ is a basis for $\text{Ker}(A)$, and $\bar{v}_p$ is any solution to $A\bar{x} = \bar{b}$.

Example: Let $A = \begin{pmatrix} 1 & 2 & 11 \\ 2 & 1 & 12 \\ 3 & 1 & 13 \end{pmatrix}$.

(1) Find $\text{rk}(A)$ and nullity (A) .

(2) Find the general solution to $A\bar{x} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$.

We will read the results off of a row-echelon form for the augmented matrix.

$$\left(\begin{array}{ccc|c} 1 & 2 & 11 & 1 \\ 2 & 1 & 12 & 2 \\ 3 & 1 & 13 & 1 \end{array} \right) \xrightarrow{A_{12}(-2)} \left(\begin{array}{ccc|c} 1 & 2 & 11 & 1 \\ 0 & -3 & -10 & 0 \\ 3 & 1 & 13 & 1 \end{array} \right) \xrightarrow{A_{13}(-3)} \dots$$

$$\left(\begin{array}{cccc|c} 1 & 2 & 1 & 1 & 1 \\ 0 & -3 & -1 & 0 & 0 \\ 0 & -5 & -2 & 0 & -2 \end{array} \right) \xrightarrow{M_2(-1/3)} \left(\begin{array}{cccc|c} 1 & 2 & 1 & 1 & 1 \\ 0 & 1 & 1/3 & 0 & 0 \\ 0 & -5 & -2 & 0 & -2 \end{array} \right) \xrightarrow{A_{23}(5)} \left(\begin{array}{cccc|c} 1 & 2 & 1 & 1 & 1 \\ 0 & 1 & 1/3 & 0 & 0 \\ 0 & 0 & -11/3 & 0 & -2 \end{array} \right)$$

$$\xrightarrow{M_3(-3)} \left(\begin{array}{cccc|c} 1 & 2 & 1 & 1 & 1 \\ 0 & 1 & 1/3 & 0 & 0 \\ 0 & 0 & 10 & 0 & 6 \end{array} \right)$$

(1) $\text{rk}(A)=3$ so $\text{nullity}(A)=1$.

(2) The augmented matrix corresponds to the system

$$\begin{aligned} x+2y+z+w &= 1 \\ y+\frac{1}{3}z &= 0 \\ z &= 6 \end{aligned}$$

There is one free variable which we take as w .
Then $z=6, y=-2, w=0, x=-1$ is a particular solution to $A\bar{x}=\bar{b}$; i.e. $\bar{x} = \begin{bmatrix} -1 \\ -2 \\ 6 \\ 0 \end{bmatrix}$ is a solution

The homogeneous system is

$$\begin{aligned} x+2y+z+w &= 0 \\ y+\frac{1}{3}z &= 0 \\ z &= 0 \end{aligned}$$

So $x=-w$. The null space has $\begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$ as a basis.

The general solution to $A\bar{x}=\bar{b}$ is $\bar{x} = c \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix} + \begin{bmatrix} -1 \\ -2 \\ 6 \\ 0 \end{bmatrix}$.

(5)

Problem: Let $A = \begin{pmatrix} 2 & 1 & 2 \\ 4 & 1 & 4 \\ 8 & 3 & 8 \end{pmatrix}$, and $\bar{b} = \begin{bmatrix} 4 \\ 6 \\ 8 \end{bmatrix}$.

- (1) Find the rank and nullity of A .
- (2) Find a basis for $\text{Ker}(A)$, and the general solution to $A\bar{x} = \bar{b}$.

Finally, we summarize much of our past work in the following theorems.

Invertible matrix theorem: Let A be an $n \times n$ matrix.

The following are equivalent.

- (1) A is invertible
- (2) $\text{rk}(A) = n$, i.e., $\text{Im}(A) = \text{column space of } A = \mathbb{R}^n$
- (3) The rows of A are independent
- (4) The columns of A are independent
- (5) $\text{Ker}(A) = \{\vec{0}\}$, i.e., $\text{nullity}(A) = 0$.
- (6) $\det(A) \neq 0$.
- (7) A^T is invertible
- (8) The columns of A are a basis for $\mathbb{R}^n$
- (9) The rows of A (when transposed) are a basis for $\mathbb{R}^n$

Lecture 18

Let V and W be F -vector spaces. A mapping T from V to W is just a function $T: V \rightarrow W$, i.e., a rule that assigns one vector in W to each vector in V .

For example, an $m \times n$ matrix A is a mapping $A: \mathbb{R}^n \rightarrow \mathbb{R}^m$. The determinant is a mapping from $M_n(\mathbb{R})$ to $\mathbb{R}$.

A mapping $T: V \rightarrow W$ is called a linear transformation if $T(\bar{x} + \bar{y}) = T(\bar{x}) + T(\bar{y})$ for all $\bar{x}, \bar{y} \in V$ and $T(r\bar{x}) = rT(\bar{x})$ for all $r \in F$ and all $\bar{x} \in V$.

We know that, if A is an $m \times n$ matrix, then $A(\bar{x} + \bar{y}) = A\bar{x} + A\bar{y}$ and $A(r\bar{x}) = rA(\bar{x})$ for all $\bar{x}, \bar{y} \in \mathbb{R}^n, r \in \mathbb{R}$.

so A is a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$.

However, $\det(rA) = r^n \det(A)$ so $\det$ is not a linear transformation for $n \geq 2$.

Note that, if T is a linear transformation, $V \rightarrow W$, $T(c_1\bar{x}_1 + \dots + c_n\bar{x}_n) = c_1T(\bar{x}_1) + \dots + c_nT(\bar{x}_n)$, so linear transformations preserve linear combinations. (here, $c_1, \dots, c_n \in F$ and $\bar{x}_1, \dots, \bar{x}_n \in V$).

V is called the domain of T , and W is called the codomain of T .

Later in the semester we will be studying linear transformations that are defined using derivatives. For example, if $f \in C^{(3)}(I)$, the vector space of 3-times continuously differentiable functions on the interval I , the mapping $T: C^{(3)}(I) \rightarrow C^{(0)}(I)$ defined by $T(f) = 3 \frac{d^3 f}{dx^3} + 12 \frac{d^2 f}{dx^2} + f(x)$ is a linear

transformation because $\frac{d}{dx}$ is. In fact, if $g, f \in C^{(n)}(I)$ then $\frac{d(f+g)}{dx} = \frac{df}{dx} + \frac{dg}{dx}$ and $\frac{d(rf)}{dx} = r \frac{df}{dx}$ so

$\frac{d}{dx}$ is a linear transformation from $C^{(n)}(I)$ to $C^{(n-1)}(I)$.

Remark: Let $T: V \rightarrow W$ be a linear transformation, and let $\{\bar{v}_1, \dots, \bar{v}_n\}$ be a basis for V . If we write $\bar{x} \in V$ as a linear combination of $\bar{v}_1, \dots, \bar{v}_n$

$$\bar{x} = c_1 \bar{v}_1 + \dots + c_n \bar{v}_n$$

then $T(\bar{x}) = T(c_1 \bar{v}_1 + \dots + c_n \bar{v}_n) = c_1 T(\bar{v}_1) + \dots + c_n T(\bar{v}_n)$, so the effect of T on a basis of V completely determines the effect of T on every vector in V .

Example: Let $T: \mathbb{R}^3 \rightarrow \mathbb{R}^4$ be a linear transformation satisfying $T(\bar{e}_1) = (1, 4, 7, 0)$, $T(\bar{e}_2) = (3, 2, 5, 0)$, and $T(\bar{e}_3) = (0, 1, 4, 0)$. Find $T(7, 8, 9)$. (We've written $\mathbb{R}^n$ as row vectors, rather than column vectors).

$$\begin{aligned}
 (7, 8, 9) &= 7\bar{e}_1 + 8\bar{e}_2 + 9\bar{e}_3, \text{ so } T(7, 8, 9) = 7T(\bar{e}_1) + 8T(\bar{e}_2) + 9T(\bar{e}_3) \\
 &= 7 \cdot (1, 4, 7, 0) + 8 \cdot (3, 2, 5, 0) + 9 \cdot (0, 1, 4, 0) \\
 &= (31, 44, 89, 0).
 \end{aligned}$$

Problem: Let $T: \mathbb{P}_2 \rightarrow \mathbb{P}_2$ be the linear transformation satisfying $T(1) = 5 + 2x$, $T(x) = 1 + x^2$, $T(x^2) = 7 - 9x + 2x^2$. What is $T(7x^2 + 3x - 5)$?

Remarks: (1) Let $T: V \rightarrow W$ be a linear transformation. Then $T(\bar{0}_V) = \bar{0}_W$. (since for any $\bar{x} \in V$, $T(\bar{x} + \bar{0}_V) = T(\bar{x})$
 $\parallel$
 $T(\bar{x}) + T(\bar{0}_V)$.)

Subtracting $T(\bar{x})$ from each side we get $T(\bar{0}_V) = \bar{0}_W$.

(2) For any $\bar{x} \in V$ $T(-\bar{x}) = -T(\bar{x})$ (since $-\bar{x} = -1 \cdot \bar{x}$, so $T(-\bar{x}) = T(-1 \cdot \bar{x}) = -1T(\bar{x}) = -T(\bar{x})$).

Our first example of a linear transformation was an $m \times n$ matrix, which we've seen is a linear transformation $\mathbb{R}^n \xrightarrow{A} \mathbb{R}^m$. In fact, these are the only linear transformations $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$. We build a matrix A whose i^{th} column is $T(\bar{e}_i)$ (so, $A = [T(\bar{e}_1), \dots, T(\bar{e}_n)]$). If we write $\bar{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ then $\bar{x} = x_1\bar{e}_1 + \dots + x_n\bar{e}_n$, and $T(\bar{x}) = T(x_1\bar{e}_1 + \dots + x_n\bar{e}_n) = x_1T(\bar{e}_1) + \dots + x_nT(\bar{e}_n) = A\bar{x}$.

So T and A have the same effect on every $\bar{x} \in V$.

Definition: The matrix $A = [T(\bar{e}_1), \dots, T(\bar{e}_n)]$ is called the matrix of T .

Example: Let $T: \mathbb{R}^5 \rightarrow \mathbb{R}^3$ be the linear transformation defined by $T(x_1, x_2, x_3, x_4, x_5) = (2x_1 - 3x_2 + x_5, x_3 + x_4, x_1 - x_3 + x_5)$. Find the matrix of T .

$$\begin{aligned} T(\bar{e}_1) &= (2, 0, 1)^T, \quad T(\bar{e}_2) = (-3, 0, 0)^T, \quad T(\bar{e}_3) = (0, 1, -1)^T \\ T(\bar{e}_4) &= (0, 1, 0)^T, \quad T(\bar{e}_5) = (1, 0, 1)^T. \end{aligned}$$

So the matrix of T is

$$A = \begin{pmatrix} 2 & -3 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & -1 & 0 & 1 \end{pmatrix}$$

Problem: Let T be the linear transformation $\mathbb{R}^4 \rightarrow \mathbb{R}^3$ defined by $T(x_1, x_2, x_3, x_4) = (7x_1 + x_3 - x_4, 2x_2 + x_3, x_1 - x_4)$.

- (1) Find the matrix A of T .
- (2) Find the rank of A and the nullity of A .
- (3) Find a basis for $\text{Ker}(T)$.

Let $T: V \rightarrow W$ be a linear transformation. The set $\{\bar{x} \in V: T(\bar{x}) = \bar{0}_W\}$ is called the kernel of T , denoted $\text{Ker}(T)$.

This generalizes the notion of the null space of a matrix.

Note that $\text{Ker}(T)$ is a subspace of V since, if $\bar{x}, \bar{y} \in \text{Ker}(T)$,

$$T(r\bar{x} + s\bar{y}) = rT(\bar{x}) + sT(\bar{y}) = \bar{0}_W + \bar{0}_W = \bar{0}_W$$

for all $r, s \in F$.

The range of T is the image of the function T , i.e., $\{T(\bar{x}) : \bar{x} \in V\}$.

Notice that the range of T is a subspace of W since if $\bar{u} = T(\bar{x})$ and $\bar{v} = T(\bar{y})$ then $r\bar{u} + s\bar{v} = rT(\bar{x}) + sT(\bar{y}) = T(r\bar{x} + s\bar{y})$ for all $r, s \in F$.

Example: Let $T: M_2(\mathbb{R}) \rightarrow M_2(\mathbb{R})$ be the linear transformation $T(A) = A + A^T$.

The Kernel of T is $\{A \in M_2(\mathbb{R}) : T(A) = A + A^T = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\}$.

So $\text{Ker } T = \{A \in M_2(\mathbb{R}) : A = -A^T\}$, i.e., $\text{Ker } T$ consists entirely of skew-symmetric matrices.

The range of T is $\{T(A) : A \in M_2(\mathbb{R})\}$
 $= \{A + A^T : A \in M_2(\mathbb{R})\}$

The matrices $A + A^T$ are the symmetric matrices.

Problem (i) Let $T: P_2 \rightarrow P_2$ be the linear transformation satisfying $T(1) = x^2 + 1$, $T(x) = x^2 - 1$, and $T(x^2) = x + 1$

(a) Find $T(ax^2 + bx + c)$

(b) Find the Kernel and image of T .

(2) Same questions if $T(1) = x^2 + 1$, $T(x) = x^2 - 1$ and $T(x^2) = x^2 + 2$

Lecture 19

Let $T: U \rightarrow V$, $S: V \rightarrow W$ be linear transformations of F -vector spaces. The composition $S \circ T: U \rightarrow W$ is also a linear transformation since

$$\begin{aligned} S \circ T(a\bar{x} + b\bar{y}) &= S(T(a\bar{x} + b\bar{y})) = S(aT(\bar{x}) + bT(\bar{y})) \\ &= aS \circ T(\bar{x}) + bS \circ T(\bar{y}). \end{aligned}$$

We already knew this when $U = \mathbb{R}^n$, $V = \mathbb{R}^m$, $W = \mathbb{R}^k$ and T is an $m \times n$ matrix, S a $k \times m$ matrix.

The composition $S \circ T$ is just the $k \times n$ matrix product ST .

Definitions: (1) A linear transformation $T: V \rightarrow W$ is injective (or one-to-one) if distinct elements of V are mapped to distinct elements of W , i.e., if $\bar{x} \neq \bar{y}$ then $T(\bar{x}) \neq T(\bar{y})$.

(2) A linear transformation T is onto if the range of T ($= \text{Im}(T)$) is all of W . (We'll also say T is surjective).

(3) A linear transformation that is both injective and surjective is called an isomorphism.

If $T: V \rightarrow W$ is an isomorphism we say V and W are isomorphic.

Examples: (1) The mapping $T: P_2 \rightarrow P_2$ defined by $T(ax^2 + bx + c) = ax^2 + bx$ is not an injective linear transformation. It is not surjective either.

The mapping $T: P_2 \rightarrow P_3$ given by $T(ax^2+bx+c) = ax^3+bx^2+cx$ is injective, but not surjective. (2)

(2) The mapping $T: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ defined by $T(x,y,z) = (x,y)$ is a surjective linear transformation. It is not injective.

(3) The mapping $T: \mathbb{R}^3 \rightarrow P_2$ defined by $T(a,b,c) = ax^2+bx+c$ is an isomorphism.

(4) Let A be an $n \times n$ matrix which we view as a linear transformation $\mathbb{R}^n \rightarrow \mathbb{R}^n$. Then A is an isomorphism if and only if $\det(A) \neq 0$. Why?

Remark: If $T: U \rightarrow V$ is an injective linear transformation then $\text{Im}(T)$ is isomorphic to U . Since the mapping $T: U \rightarrow \text{Im}(T)$ is both injective and surjective.

It is often useful to know if a linear transformation is injective. The next theorem provides a criterion for injectivity.

Theorem: Let $T: V \rightarrow W$ be a linear transformation. T is injective if and only if $\text{Ker}(T) = \{\bar{0}_V\}$.

proof: We already know that $T(\bar{0}_V) = \bar{0}_W$. If T is injective then no other vector in V can map to $\bar{0}_W$ so $\text{Ker}(T) = \{\bar{0}_V\}$.

Conversely, suppose $\text{Ker}(T) = \{\bar{0}_V\}$. If $\bar{x} \neq \bar{y}$ are distinct vectors in V then $\bar{x} - \bar{y} \neq \bar{0}_V$.

Since $\text{Ker}(T) = \{\bar{0}_V\}$ we see $T(\bar{x} - \bar{y}) \neq 0$, or equivalently, $T(\bar{x}) \neq T(\bar{y})$, i.e., T is injective.

Examples: (1) The linear transformation $T: P_3 \rightarrow P_2$ defined by $T(f(x)) = \frac{df}{dx}$ is not one-to-one since $T(1) = 0$,

so $\text{Ker}(T) \neq \{0\}$. It is onto. Why?

(2) The linear transformation $T: P_1 \rightarrow P_2$ defined by $T(ax+b) = bx+a$

is injective since $T(ax+b) = bx+a = 0$ if and only if both a and b are 0. Thus, $\text{Ker}(T) = \{0\}$.

However, T is not surjective. Why?

In the following we assume V and W are finite dimensional.

Remarks: (1) If $T: V \rightarrow W$ is injective then $\dim(V) \leq \dim(W)$.

To see this suppose $\{\bar{v}_1, \dots, \bar{v}_n\}$ is a basis for V , and suppose there is a linear relation

$$c_1 T(\bar{v}_1) + \dots + c_n T(\bar{v}_n) = \bar{0}_W$$

$$\text{Then } T(c_1 \bar{v}_1 + \dots + c_n \bar{v}_n) = \bar{0}_W$$

Since T is one-to-one we know $\text{Ker}(T) = \{\bar{0}_V\}$, so $c_1 \bar{v}_1 + \dots + c_n \bar{v}_n = \bar{0}$. Since $\bar{v}_1, \dots, \bar{v}_n$ are independent we must have

$c_1 = c_2 = \dots = c_n = 0$, i.e., $T(\bar{v}_1), \dots, T(\bar{v}_n)$ are independent. Then the dimension of W must be $\geq n$.

(2) If T is surjective then $\dim(V) \geq \dim(W)$.

To see this note that if T is onto then the range

of T is W . We choose a basis $\{\bar{v}_1, \dots, \bar{v}_k\}$ of $\text{Ker } T$, and complete it to a basis $\{\bar{v}_1, \dots, \bar{v}_k, \bar{v}_{k+1}, \dots, \bar{v}_n\}$ of V . Since T is onto $T(\bar{v}_1), \dots, T(\bar{v}_n)$ span W . But $T(\bar{v}_1), \dots, T(\bar{v}_k) = \bar{0}_W$, so $T(\bar{v}_{k+1}), \dots, T(\bar{v}_n)$ span W . So $\dim W \leq n - k \leq \dim V$.

(4) If $\dim(V) = \dim(W)$ then a linear transformation $T: V \rightarrow W$ is one-to-one if and only if it is onto.

If T is injective the proof of remark 1 shows the image of a basis of V is a basis of W , i.e., if $\bar{v}_1, \dots, \bar{v}_n$ are a basis for V then

$T(\bar{v}_1), \dots, T(\bar{v}_n)$ are a basis for W , so every vector in W can be written as $c_1 T(\bar{v}_1) + \dots + c_n T(\bar{v}_n)$ for some scalars $c_1, \dots, c_n$. But

$c_1 T(\bar{v}_1) + \dots + c_n T(\bar{v}_n) = T(c_1 \bar{v}_1 + \dots + c_n \bar{v}_n) \in \text{Im}(T)$, so T is onto.

If T is onto we want to show $\text{Ker}(T) = \{\bar{0}\}$.

Suppose $\bar{x} \in V$ and $T(\bar{x}) = \bar{0}_W$. Write $\bar{x}$ as a linear combination of basis elements

$$\bar{x} = c_1 \bar{v}_1 + \dots + c_n \bar{v}_n$$

Then $T(\bar{x}) = T(c_1 \bar{v}_1 + \dots + c_n \bar{v}_n) = c_1 T(\bar{v}_1) + \dots + c_n T(\bar{v}_n) = \bar{0}_W$

We claim $T(\bar{v}_1), \dots, T(\bar{v}_n)$ are independent which forces $c_1 = c_2 = \dots = c_n = 0$, so $\bar{x} = \bar{0}_V$, i.e. $\text{Ker}(T) = \{\bar{0}_V\}$, so T is one-to-one.

To see $T(\bar{v}_1), \dots, T(\bar{v}_n)$ are a basis of W notice

(5)

that we can write any vector in W as a linear combination of $T(\bar{v}_1), \dots, T(\bar{v}_n)$ because T is onto (to make this explicit if $\bar{z} \in W$, say $\bar{z} = T(\bar{u})$).

Write $\bar{u} = a_1 \bar{v}_1 + \dots + a_n \bar{v}_n$, so $\bar{z} = T(\bar{u}) = T(a_1 \bar{v}_1 + \dots + a_n \bar{v}_n) = a_1 T(\bar{v}_1) + \dots + a_n T(\bar{v}_n)$.

Thus, $T(\bar{v}_1), \dots, T(\bar{v}_n)$ span an n -dimensional subspace of W , and since $n = \dim W$ we see that this subspace must be W .

If $T: V \rightarrow W$ is an isomorphism of F -vector spaces then the inverse map $T^{-1}: W \rightarrow V$ (defined by $T^{-1}(\bar{x}) = \bar{y}$ if and only if $T(\bar{y}) = \bar{x}$) is an isomorphism $T^{-1}: W \rightarrow V$. It is one-to-one and onto because T is. To see that T^{-1} is a linear transformation suppose $T(\bar{y}) = \bar{x}$ and $T(\bar{u}) = \bar{v}$. Then $T(r\bar{y} + s\bar{u}) = rT(\bar{y}) + sT(\bar{u}) = r\bar{x} + s\bar{v}$, so $T^{-1}(r\bar{x} + s\bar{v}) = r\bar{y} + s\bar{u} = rT^{-1}(\bar{x}) + sT^{-1}(\bar{v})$.

If $V = W = \mathbb{R}^n$ and T is the linear transformation with matrix A then T is an isomorphism if and only if A is. This is the case if and only if $\det(A) \neq 0$ (since $\det(A) \neq 0$ if and only if A is invertible, or equivalently A has rank n , so $\text{Ker}(A) = \{0\}$ and $\text{Im}(A) = \mathbb{R}^n$).

Example: Let $T: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ be the linear transformation given by $T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} 2x+y-z \\ 3x-y+z \\ 7x+y-z \end{bmatrix}$. Is T an isomorphism?

We compute the matrix $A = [T(\bar{e}_1), T(\bar{e}_2), T(\bar{e}_3)]$ of T .

$$T(\bar{e}_1) = \begin{bmatrix} 2 \\ 3 \\ 7 \end{bmatrix}, \quad T(\bar{e}_2) = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \quad T(\bar{e}_3) = \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}.$$

The matrix A of T is $\begin{bmatrix} 2 & 1 & -1 \\ 3 & -1 & 1 \\ 7 & 1 & -1 \end{bmatrix}$ which is

only of rank 2, so A is not an isomorphism, nor is T .

Problem: Let $T: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ be the linear transformation given by $T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x+y+z \\ x-y-z \\ x+y-z \end{bmatrix}$. Is T an isomorphism?

Lecture 20

Let V and W be F -vector spaces with ordered bases $B = [\bar{v}_1, \dots, \bar{v}_n]$, $C = [\bar{w}_1, \dots, \bar{w}_m]$, respectively. If $T: V \rightarrow W$ is a linear transformation we associate an $m \times n$ matrix to T as follows. The i^{th} column of this matrix is $[T(\bar{v}_i)]_C$, the component vector of $T(\bar{v}_i)$ with respect to the basis C of W . We denote this matrix by $[T]_B^C = ([T(\bar{v}_1)]_C, \dots, [T(\bar{v}_n)]_C)$

and call it the matrix of T relative to the bases B and C .

In the special case when $V=W$ and $B=C$ we call $[T]_B^B$ the matrix of T relative to B . We will often be interested in this special case.

Notice that if $V = \mathbb{R}^n$, $B = \{\bar{e}_1, \dots, \bar{e}_n\}$ and $W = \mathbb{R}^m$, $C = \{\bar{e}_1, \dots, \bar{e}_m\}$ then $[T]_B^C$ is just

the matrix of T

Remark: If $V=W$, and T is the identity map, $T(\bar{v}) = \bar{v}$ for all $\bar{v} \in V$, then $[T]_B^B$ is $I_{n \times n}$, the change of basis matrix.

Example (1) Let $V = P_2$, $W = \mathbb{R}^3$, $B = \{1, x, x^2\}$, $C = \{\bar{e}_1, \bar{e}_2, \bar{e}_3\}$ and let $T: V \rightarrow W$ be defined by $T(ax^2 + bx + c) = (a, b, c)$. Find $[T]_B^C$.
We compute $[T(1)]_C$, $[T(x)]_C$, $[T(x^2)]_C$.

$$[T(1)]_C = (0, 0, 1)^T = \bar{e}_3$$

$$[T(x)]_C = (0, 1, 0)^T = \bar{e}_2$$

$$[T(x^2)]_C = (1, 0, 0)^T = \bar{e}_1$$

$$\text{So, } [T]_B^C = (\bar{e}_3, \bar{e}_2, \bar{e}_1) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

(2) Let $V = P_1 = W$ $B = \{2, 2+x\}$, $C = \{1+x, 1-x\}$,
and let $T: P_1 \rightarrow P_1$ be given by $T(ax+b) = bx+a$.
Find $[T]_B^C$. We compute $[T(2)]_C$ and $[T(2+x)]_C$

$$[T(2)]_C = [2x]_C = [1 \cdot (1+x) - 1 \cdot (1-x)]_C = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$[T(2+x)]_C = [(2x+1)]_C = \left[\frac{3}{2}(1+x) - \frac{1}{2}(1-x) \right]_C = \begin{bmatrix} 3/2 \\ -1/2 \end{bmatrix}.$$

$$\text{So, } [T]_B^C = \begin{pmatrix} 1 & 3/2 \\ -1 & -1/2 \end{pmatrix}.$$

Problem: Let $T: \mathbb{R}^3 \rightarrow P_1$ be the linear transformation
 $T\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) = ax+c.$

Let $B = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$, $C = \{2x, 1\}$.

Find $[T]_B^C$.

Let $T: V \rightarrow W$ be a linear transformation, and let B, C be ordered bases of V and W , respectively. Then, if $\bar{x} \in V$ we have $[T(\bar{x})]_C = [T]_B^C [\bar{x}]_B$ (since, if we

write $\bar{x} = d_1 \bar{v}_1 + \dots + d_n \bar{v}_n$ as a linear combination of the vectors in B ,

$$\begin{aligned} [T]_B^C [\bar{x}]_B &= d_1 [T(\bar{v}_1)]_C + d_2 [T(\bar{v}_2)]_C + \dots + d_n [T(\bar{v}_n)]_C \\ &= [T(d_1 \bar{v}_1 + \dots + d_n \bar{v}_n)]_C = [T(\bar{x})]_C. \end{aligned}$$

We will often be interested in the case where $V=W$. Suppose $T: V \rightarrow V$ is a linear transformation, and that B and C are ordered bases of V . Then $[T]_C^C = P_{C \leftarrow B} [T]_B^B P_{B \leftarrow C}$.

We just saw that $[T]_C^C [\bar{x}]_C = [T(\bar{x})]_C$. We compute

$$P_{C \leftarrow B} [T]_B^B P_{B \leftarrow C} ([\bar{x}]_C) = P_{C \leftarrow B} [T]_B^B [\bar{x}]_B =$$

$$P_{C \leftarrow B} [T(\bar{x})]_B = [T(\bar{x})]_C \text{ for any } \bar{x} \in V.$$

Recall that $P_{C \leftarrow B}$ and $P_{B \leftarrow C}$ are inverses of each other. If we write S for $P_{B \leftarrow C}$ then $S^{-1} = P_{C \leftarrow B}$. What we just proved says

$$S^{-1} [T]_B^B S = [T]_C^C.$$

Definition: Let A and B be $n \times n$ matrices. We say that A is similar to B if there is an

(4)

invertible matrix S with $A = S^{-1}BS$.

In this case $B = SAS^{-1} = (S^{-1})^{-1}AS^{-1}$ so B is similar to A .

We have shown that if $T: V \rightarrow V$ is a linear transformation, and B and C are ordered bases of V , then the matrices $[T]_B^B$ and $[T]_C^C$ are similar.

If U, V, W are vector spaces with ordered bases A, B , and C , respectively, and $S: U \rightarrow V, T: V \rightarrow W$ are linear transformations then $TS: U \rightarrow W$ is a linear transformation with

$$[TS]_A^C = [T]_B^C [S]_A^B.$$

To see this suppose $\bar{x} \in U$. Then

$$[T]_B^C [S]_A^B [\bar{x}]_A = [T]_B^C [S(\bar{x})]_B = [TS(\bar{x})]_C = [TS]_A^C [\bar{x}]_A$$

In other words, the matrix associated to the composition of linear transformations (relative to ordered bases) is the product (i.e., composition) of the matrices associated to each transformation.

Remarks: Let V, W be F -vector spaces with ordered bases B and C , and let $T: V \rightarrow W$ be a linear transformation. Then $\bar{x} \in V$ is in $\text{Ker}(T)$ if and only if $[\bar{x}]_B \in \text{Ker}([T]_B^C)$. (since $T(\bar{x}) = \bar{0}$ if and only

(5)

$$\text{if } [T(x)]_C = \bar{0}$$

$$[T]_B^C [\bar{x}]_B, \text{ i.e., } [\bar{x}]_B \in \text{Ker } [T]_B^C$$

Similarly, $\bar{z}$ lies in $\text{Im}(T)$ (= range space of T) if and only if $[\bar{z}]_C$ lies in $\text{Im } [T]_B^C$ (= column space of $[T]_B^C$). To see this note that $\bar{z} \in \text{Im}(T)$ means there exists $\bar{x} \in V$ with $T(\bar{x}) = \bar{z}$. Then $[T(\bar{x})]_C = [\bar{z}]_C$.

$$[T]_B^C [\bar{x}]_B$$

From our earlier work we see that T is one-to-one if and only if $\text{Ker } [T]_B^C = \{\bar{0}\}$, and T is surjective if and only if the column space of $[T]_B^C$ (= $\text{Im } [T]_B^C$) is $\mathbb{R}^n$ (where $n = \dim W$).

Combining these observations we see that T is invertible if and only if $[T]_B^C$ is invertible for some ordered bases B and C .

$$\text{Finally, if } T \text{ is invertible then } [T^{-1}]_C^B \cdot [T]_B^C = [I]_B^B$$

$$\text{so } [T^{-1}]_C^B = ([T]_B^C)^{-1}.$$

Example: Let $T: P_2 \rightarrow \mathbb{R}^2$ be the linear transformation given by $T(ax^2 + bx + c) = (a+c, a-c)^T$.

Let $B = \{x^2 + x, 1 - x\}$, $C = \{e_1, e_2\}$.

Find the matrix of T relative to the bases B and C .

Determine if T is injective and/or surjective

We compute $[T(x^2)]_C$, $[T(1+x)]_C$, and $[T(1-x)]_C$.

$$[T(x^2)]_C = [(1, 1)^T]_C = 0 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 1 \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$[T(1+x)]_C = [(1, -1)^T]_C = 2 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} - 1 \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

$$[T(1-x)]_C = [(1, -1)^T]_C = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

$$\text{So } [T]_B^C = \begin{pmatrix} 0 & 2 & 2 \\ 1 & -1 & -1 \end{pmatrix}$$

This is a rank 2 matrix so T is surjective, but $[T]_B^C$ has a one-dimensional Kernel, so T is not injective.

Problem: Let $T: \text{Sym}_2(\mathbb{R}) \rightarrow \mathbb{R}^3$ be the linear transformation $T\begin{pmatrix} a & c \\ c & b \end{pmatrix} = (a, b, c)^T$

$$\text{Let } B = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}, \quad C = \{ \bar{e}_1, e_3, e_1 + e_2 \} \\ = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\}.$$

Find $[T]_B^C$ and determine if T is injective and/or surjective.

Lecture 21

Let A be an $n \times n$ matrix. We will be interested in scalars λ and vectors $\bar{x} \neq \bar{0}$ that satisfy $A\bar{x} = \lambda\bar{x}$, i.e., matrix multiplication acts like scalar multiplication on $\bar{x}$. The scalar λ is called an eigenvalue (or characteristic value) of A , and the vector $\bar{x}$ is called an eigenvector corresponding to the eigenvalue λ .

Given the matrix A how do we find its eigenvalues? If $A\bar{x} = \lambda\bar{x}$ then $(A - \lambda I)\bar{x} = \bar{0}$. We know that this system has a non-zero solution if and only if $(A - \lambda I)$ is singular, i.e. $\det(A - \lambda I) = 0$. So, to find the eigenvalues of A we solve $\det(A - \lambda I) = 0$.

Remarks: (1) The polynomial $p(\lambda) = \det(A - \lambda I)$ is called the characteristic polynomial of A . The equation $p(\lambda) = 0$ is called the characteristic equation. It is also common to see $\det(\lambda I - A)$ called the characteristic polynomial of A . Since we are setting the characteristic polynomial equal to 0 either choice leads to the same solutions. The advantage of using $\det(\lambda I - A)$ is that its leading coefficient is always 1, and this is often convenient.

(2) If A is an $n \times n$ matrix $\det(A - \lambda I)$ is a degree n polynomial in λ .

(3) A is non-singular if and only if $\det(A) \neq 0$ which happens if and only if 0 is not an eigenvalue of A (If 0 is an eigenvalue then A has a non-trivial null space).

(4) Even if the entries in A are all real numbers the eigenvalues may not be. For example, if

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \text{ then}$$

$$\det(A - \lambda I) = \det \left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \right) = \det \begin{pmatrix} -\lambda & 1 \\ -1 & -\lambda \end{pmatrix} = \lambda^2 + 1$$

So $\det(A - \lambda I) = 0$ if and only if $\lambda = \pm i$. ($= \sqrt{-1}$).

Examples: Find the eigenvalues and corresponding eigenvectors if $A =$

(1) $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. We've just seen that the eigenvalues are $\lambda = \pm i$. For each eigenvalue we'll look for a corresponding eigenvector.

$\lambda = i$: We want to find $\bar{x}$ with $A\bar{x} = i\bar{x}$

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = i \begin{pmatrix} u \\ v \end{pmatrix}$$

So $v = iu$ and $-u = iv$. Notice that these two equations say the same thing. We pick any non-zero value for u , say $u = 1$. Then $v = iu = i$, so $\begin{bmatrix} 1 \\ i \end{bmatrix}$ is an eigenvector corresponding to $\lambda = i$.

We check this: $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{bmatrix} 1 \\ i \end{bmatrix} = \begin{bmatrix} i \\ -1 \end{bmatrix} = i \begin{bmatrix} 1 \\ i \end{bmatrix} \quad \checkmark$

$$\lambda = -i \quad \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = -i \begin{bmatrix} u \\ v \end{bmatrix}$$

$$\text{So, } \begin{aligned} v &= -iu \\ -u &= -iv \end{aligned}$$

Again, these equations say the same thing. We take $u=1$ so $v=-i$, and $\begin{bmatrix} 1 \\ -i \end{bmatrix}$ is an eigenvector corresponding to $\lambda = -i$.

We check this $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{bmatrix} 1 \\ -i \end{bmatrix} = \begin{bmatrix} -i \\ -1 \end{bmatrix} = -i \begin{bmatrix} 1 \\ -i \end{bmatrix} \quad \checkmark$.

(2)

$$A = \begin{pmatrix} 1 & 5 & 6 \\ 0 & 2 & 7 \\ 0 & 0 & 4 \end{pmatrix}$$

$$\det(\lambda I - A) = \det \begin{pmatrix} \lambda - 1 & -5 & -6 \\ 0 & \lambda - 2 & -7 \\ 0 & 0 & \lambda - 4 \end{pmatrix} = (\lambda - 1)(\lambda - 2)(\lambda - 4) = 0$$

if and only if $\lambda = 1, 2$, or 4 . Notice that the eigenvalues of a triangular matrix are the diagonal entries.

$\lambda = 1$

$$\begin{pmatrix} 1 & 5 & 6 \\ 0 & 2 & 7 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$x + 5y + 6z = x \quad x = x$$

$$2y + 7z = y \quad y = -7z = 0$$

$$4z = z \quad z = 0$$

$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ is an eigenvector corresponding to $\lambda = 1$.

$$\lambda = 2$$

$$\begin{pmatrix} 1 & 5 & 6 \\ 0 & 2 & 7 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 2 \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$x + 5y + 6z = 2x \leftarrow x = 5y$$

$$2y + 7z = 2y \leftarrow y = y$$

$$4z = 2z \leftarrow z = 0$$

We take $y=1$ and see that $\begin{bmatrix} 5 \\ 1 \\ 0 \end{bmatrix}$ is an eigenvector corresponding to $\lambda=2$.

$$\lambda = 4$$

$$\begin{pmatrix} 1 & 5 & 6 \\ 0 & 2 & 7 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 4 \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

$$x + 5y + 6z = 4x$$

$$2y + 7z = 4y \leftarrow y = \frac{7}{2}z, x = \frac{47}{6}z$$

$$4z = 4z$$

Taking $z=6$ we get $\begin{bmatrix} 47 \\ 21 \\ 6 \end{bmatrix}$ is an eigenvector

corresponding to $\lambda=4$.

Remarks: (1) Notice that we were able to specify one of the values of x, y, z . If we chose a different value we would multiply our eigenvector by a scalar. The result is still an eigenvector since, if $A\bar{x} = \lambda\bar{x}$ then $A(c\bar{x}) = cA\bar{x} = c\lambda\bar{x} = \lambda(c\bar{x})$.

(2) In our first example the eigenvalues $\lambda = \pm i$ are complex conjugates (Recall that the complex conjugate of $a+bi$ is $a-bi$). If A is a matrix with real entries and $A\bar{x} = \lambda x$ then A is fixed by complex conjugation (i.e., $\bar{A} = A$), so

$$\overline{A x} = A \bar{x} = \bar{\lambda} \bar{x}$$

i.e., $\bar{\lambda}$ is also an eigenvalue of A and $\bar{x}$ is a corresponding eigenvector.

Here we write a vector as $\underline{x}$ instead of $\bar{x}$ so that we don't confuse vectors with complex conjugates.

(3) If $A\bar{x} = \lambda \bar{x}$ then $A^2 \bar{x} = A \cdot A\bar{x} = A \cdot \lambda \bar{x} = \lambda A\bar{x} = \lambda^2 \bar{x}$, and $A^n \bar{x} = \lambda^n \bar{x}$, i.e., λ^n is an eigenvalue of A^n , and $\bar{x}$ is an eigenvector.

(4) If $\bar{u}, \bar{v}$ are eigenvectors for A corresponding to the same eigenvalue λ then $A(c\bar{u} + d\bar{v})$

$$= cA\bar{u} + dA\bar{v} = c\lambda\bar{u} + d\lambda\bar{v} = \lambda(c\bar{u} + d\bar{v}),$$

so a linear combination of eigenvectors corresponding to the same eigenvalue is again an eigenvector corresponding to λ . Is this true if $\bar{u}$ and $\bar{v}$ correspond to different eigenvalues?

Problems: Find the eigenvalues, and corresponding eigenvectors.

$$(1) A = \begin{pmatrix} 3 & 0 & 0 \\ 2 & 7 & 0 \\ 1 & 1 & 5 \end{pmatrix}.$$

$$(2) A = \begin{pmatrix} 1 & -1 \\ 2 & 2 \end{pmatrix}.$$

$$(3) A = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 2 & 2 \end{pmatrix}$$

Lecture 22

We'll start with a few more examples. Let's find the eigenvalues and corresponding eigenvectors for A if.

$$A = \begin{pmatrix} 0 & 2 & 2 \\ 2 & 0 & 2 \\ 2 & 2 & 0 \end{pmatrix} \quad \det(A - \lambda I) = \det \begin{pmatrix} -\lambda & 2 & 2 \\ 2 & -\lambda & 2 \\ 2 & 2 & -\lambda \end{pmatrix}$$

$$= -\lambda^3 + 12\lambda + 16 = 0?$$

$\lambda = -2$ is a root. To find the others we divide

$$\begin{array}{r} -\lambda^2 + 2\lambda + 8 \\ \lambda + 2 \overline{) -\lambda^3 + 12\lambda + 16} \\ \underline{-\lambda^3 - 2\lambda^2} \\ +2\lambda^2 + 12\lambda + 16 \\ \underline{+2\lambda^2 + 4\lambda} \\ 8\lambda + 16 \\ \underline{8\lambda + 16} \\ 0 \end{array}$$

So $\lambda = -2$ and $\lambda = 4$ are the other roots, and $\lambda = -2$ is a root of multiplicity 2.

We now find eigenvectors

$$\lambda = 4 \quad \begin{pmatrix} 0 & 2 & 2 \\ 2 & 0 & 2 \\ 2 & 2 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 4 \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad \begin{array}{l} 2y + 2z = 4x \\ 2x + 2z = 4y \\ 2x + 2y = 4z \end{array}$$

There is one free variable, and we take $z = 1$, so $x = y = 1$

and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector corresponding to $\lambda=4$.

$\lambda=-2$: We write $A\bar{x} = -2\bar{x}$ as $(A+2I)\bar{x} = \bar{0}$.

$$A+2I = \begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix}.$$

The reduced row echelon form is $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

So there are 2 free variables and $x+y+z=0$.

We first take $y=1, z=0$ and then $y=0, z=1$ to obtain 2 independent solutions

$$\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

Problem: Find eigenvalues and corresponding eigenvectors

for $A = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 4 \end{pmatrix}$

In the 2 preceding examples the matrix A had an eigenvalue λ that occurred with multiplicity greater than one as a root of the characteristic polynomial. The multiplicity of λ as a root of the characteristic polynomial is called the algebraic multiplicity of λ .

If A is an $n \times n$ matrix with eigenvalue λ then the set of all eigenvectors (together with $\vec{0}$) form a subspace E_λ of $\mathbb{R}^n$. To see this suppose $\vec{u}, \vec{v} \in E_\lambda$,

i.e., $A\vec{u} = \lambda\vec{u}$ and $A\vec{v} = \lambda\vec{v}$. Then $A(r\vec{u} + s\vec{v}) = rA\vec{u} + sA\vec{v} = r\lambda\vec{u} + s\lambda\vec{v} = \lambda(r\vec{u} + s\vec{v})$. Of course, we had already observed this in the last lecture.

The dimension of E_λ is called the geometric multiplicity of λ . In general, the geometric multiplicity is less than or equal to the algebraic multiplicity.

Examples: (1) Let $A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix}$.

The eigenvalues are 1 (algebraic multiplicity 2), 2 and 3. We want to find eigenvectors corresponding to $\lambda = 1$. We solve

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

$$\begin{aligned} \text{Then } y &= 0 \\ z + w &= 0 \\ 2w &= 0 \end{aligned}$$

so $y = z = w = 0$, but there is no condition on x . We take $x = 1$ to obtain the eigenvector $\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$.

Any other eigenvector corresponding to $\lambda = 1$ is a scalar multiple of $\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$, so the $\lambda = 1$ -eigenspace is one dimensional; i.e. the geometric multiplicity is one.

$$(2) A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

The eigenvalues are again 1 (with algebraic multiplicity 2), 2, and 3. We look for eigenvectors corresponding to 1.

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

So $z+w=0$, $2w=0$. Thus, $w=z=0$, but there are

no conditions on x or y . If we take $x=1, y=0$, we see $\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$ is an eigenvector corresponding to 1.

If we take $x=0, y=1$ we see $\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$ is also

an eigenvector corresponding to 1.

In this case the geometric multiplicity is 2, and E_1 has $\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$ for a basis.

Suppose A has more than one eigenvalue. What can we say about their associated eigenvectors? The answer is given in the following.

Theorem: Eigenvectors corresponding to distinct eigenvalues are linearly independent.

proof

Let $\lambda_1, \dots, \lambda_m$ be distinct eigenvalues of A with corresponding eigenvectors $\bar{v}_1, \dots, \bar{v}_m$, respectively. If there is a linear relation among the $\bar{v}_i$, then there is one involving the fewest number of the $\bar{v}_i$, say $c_1 \bar{v}_1 + \dots + c_k \bar{v}_k = \bar{0} \dots (*)$ is a relation with k as small as possible. We apply the matrix A to the relation $(*)$

$$c_1 A \bar{v}_1 + \dots + c_k A \bar{v}_k = \bar{0}$$

||

$$c_1 \lambda_1 \bar{v}_1 + \dots + c_k \lambda_k \bar{v}_k = \bar{0} \dots (**)$$

Next, we multiply $(*)$ by a non-zero eigenvalue (which we take to be λ_1) to obtain $c_1 \lambda_1 \bar{v}_1 + c_2 \lambda_1 \bar{v}_1 + \dots + c_k \lambda_1 \bar{v}_k = \bar{0} \dots (***)$

We subtract $(**)$ from $(***)$ to obtain $c_2 (\lambda_1 - \lambda_2) \bar{v}_1 + \dots + c_k (\lambda_1 - \lambda_k) \bar{v}_k = \bar{0}$.

This is a linear relation involving fewer $\bar{v}_i$, a contradiction to the minimality of the relation $(*)$. Thus, there does not exist any linear relation among the $\bar{v}_i$, i.e., they are independent.

An $n \times n$ matrix that has n linearly independent eigenvectors is called non-defective. If A does not have n linearly independent eigenvectors we say A is defective.

Examples: $\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix}$ is defective.

$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix}$ is non-defective.

Remark: The theorem shows that if A has n distinct eigenvalues then A is non-defective.

For example, $A = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & 3 & 0 & 0 & 0 \\ 4 & 7 & 6 & 0 & 0 \\ 8 & 1 & 7 & 5 & 0 \\ 9 & 1 & 6 & 3 & 2 \end{pmatrix}$ is non-defective.

If A is non-defective then any set of n independent eigenvectors is a basis of $\mathbb{R}^n$. We call such a basis an eigenbasis associated to A .

Remark: A is non-defective if and only if the geometric multiplicity of each eigenvalue λ_i is equal to the algebraic multiplicity of λ_i . To see this

suppose n_i is the geometric multiplicity of λ_i , and m_i is the algebraic multiplicity.

If A is non-defective then $\sum n_i = n$, and $n_i \leq m_i$ for each i . So $n = \sum n_i \leq \sum m_i = n$, i.e.,

$$\sum n_i = \sum m_i \quad \text{and} \quad n_i = m_i \quad \text{for each } i.$$

Conversely, if $n_i = m_i$ for each i then $\sum n_i = \sum m_i = n$. The union of the bases for each eigenspace is an eigenbasis associated to A .

Lecture 23

Recall that the $n \times n$ matrices A and B are similar if there is an invertible matrix S with $A = S^{-1}BS$. We have seen that if T is a linear transformation $V \rightarrow V$, and B and C are ordered bases of V , then the matrices $[T]_B^B$ and $[T]_C^C$ are similar with the matrix S being the change of basis matrix. In other words, similar matrices represent the same linear transformation, but with respect to different bases.

Suppose A is an $n \times n$ matrix, and that there is an eigenbasis B for $\mathbb{R}^n$ associated to A ; say $B = (\bar{v}_1, \dots, \bar{v}_n)$ with $A\bar{v}_i = \lambda_i \bar{v}_i$. We build an $n \times n$ matrix S whose i th column is $\bar{v}_i$. Then

$$\begin{aligned} AS &= A[\bar{v}_1, \dots, \bar{v}_n] = [A\bar{v}_1, \dots, A\bar{v}_n] \\ &= [\lambda_1 \bar{v}_1, \dots, \lambda_n \bar{v}_n] \end{aligned}$$

$$[\bar{v}_1, \dots, \bar{v}_n] \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = S \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

The matrix $D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ is diagonal, and the matrix S is invertible since its columns are independent. Thus, $AS = SD$ or $S^{-1}AS = D$, i.e., A is similar to a diagonal matrix. In this case we say A is diagonalizable. The advantage of working with D instead of A is

the simplicity of its arithmetic. If we wish to do calculations involving the linear transformation A we choose a basis B such that the matrix of A with respect to B is diagonal, and then compute using the diagonal matrix.

Now, not every $n \times n$ matrix is diagonalizable. We've seen that the matrix $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ has

only one independent eigenvector, so $\mathbb{R}^2$ does not have an eigenbasis associated to A .

The $n \times n$ matrix A is diagonalizable if and only if it is non-defective. We've already seen that if A is non-defective then it is diagonalizable. Conversely, if $S^{-1}AS = D$ is a diagonal matrix then $AS = SD$. If we write $S = [S^{(1)}, \dots, S^{(n)}]$ with $S^{(i)}$ the i th column of S , and $D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$. Then $A[S^{(1)}, \dots, S^{(n)}] = [S^{(1)}, \dots, S^{(n)}] \begin{pmatrix} \lambda_1 & 0 \\ & \ddots \\ 0 & \lambda_n \end{pmatrix}$.

So, $AS^{(i)} = \lambda_i S^{(i)}$, i.e., the columns $S^{(i)}$ are eigenvectors for A , and they are independent since S is invertible.

Examples: (1) The matrix $A = \begin{pmatrix} 1 & 0 & 0 \\ 4 & 2 & 0 \\ 5 & 6 & 3 \end{pmatrix}$ is diagonalizable

since it has 3 distinct eigenvalues, and so it has 3 independent eigenvectors. Let's find the matrix

(3)

S that diagonalizes A .

We need to find eigenvectors corresponding to the eigenvalues $\lambda = 1, 2, 3$.

$$\lambda = 1 \quad \begin{pmatrix} 0 & 0 & 0 \\ 4 & 1 & 0 \\ 5 & 6 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

We get $4x + y = 0$ and $5x + 6y + 2z = 0$.

So $y = -4x$ and $x = \frac{2}{19}z$. We take $z = 19$, so $y = -8$ and $x = 2$ and obtain $\begin{bmatrix} 2 \\ -8 \\ 19 \end{bmatrix}$ is an eigenvector.

$$\lambda = 2 \quad \begin{pmatrix} -1 & 0 & 0 \\ 4 & 0 & 0 \\ 5 & 6 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Then $x = 0$ and $5x + 6y + z = 0$ so $z = -6y$. We take $y = 1$, so $z = -6$. Then $\begin{bmatrix} 0 \\ 1 \\ -6 \end{bmatrix}$ is an eigenvector.

$$\lambda = 3 \quad \begin{pmatrix} -2 & 0 & 0 \\ 4 & -1 & 0 \\ 5 & 6 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$-2x = 0$, $4x - y = 0$, $5x + 6y = 0$, so $x = y = 0$. We take $z = 1$ and obtain the eigenvector $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.

We place the eigenvectors into the columns of S .

$$S = \begin{pmatrix} 2 & 0 & 0 \\ -8 & 1 & 0 \\ 19 & -6 & 1 \end{pmatrix}. \quad \text{Then } S^{-1}AS = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} = D$$

If we permute the columns of S we permute the columns of D as well.

Problem: Find the matrix S that diagonalizes

$$A = \begin{pmatrix} 2 & 3 & 5 \\ 0 & 4 & 5 \\ 0 & 0 & 5 \end{pmatrix}.$$

Remarks: (1) If A and C are similar they have the same traces. This follows from the fact $\text{Tr}(AB) = \text{Tr}(BA)$. To see this suppose $C = S^{-1}AS$.

$$\begin{aligned} \text{Then } \text{Tr}(C) &= \text{Tr}(S^{-1}AS) = \text{Tr}((S^{-1}A)S) = \text{Tr}(S(S^{-1}A)) \\ &= \text{Tr}((SS^{-1})A) = \text{Tr}(I \cdot A) = \text{Tr}(A), \end{aligned}$$

(2) If A and C are similar they have the same determinant since, if $C = S^{-1}AS$, $\det(C) = \det(S^{-1}AS) = \det(S^{-1}) \det(A) \cdot \det(S) = \det(A)$.

(3) If A and B are $n \times n$ matrices with the same eigenvalues they need not be similar. For example, $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ are not similar since A is not diagonalizable.

(5)

Question: Which of the following matrices are similar to one another?

$$A = \begin{pmatrix} 1 & 2 & 5 \\ 0 & 2 & 6 \\ 0 & 0 & 3 \end{pmatrix}$$

$$B = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

$$C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 3 \end{pmatrix}$$

$$D = \begin{pmatrix} 2 & 4 & 7 \\ 0 & 1 & 9 \\ 0 & 0 & 3 \end{pmatrix}$$

$$E = \begin{pmatrix} 3 & 0 & 0 \\ 1 & 1 & 0 \\ 9 & 2 & 2 \end{pmatrix}$$

$$F = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 4 & 7 & 1 \end{pmatrix}.$$

Lecture 24

An n^{th} order linear differential equation is an equation of the form

$$(*) \quad a_n(x) \frac{d^n y}{dx^n} + a_{n-1}(x) \frac{d^{n-1} y}{dx^{n-1}} + \dots + a_1(x) \frac{dy}{dx} + a_0(x)y = F(x).$$

The integer n is the order of the highest occurring derivative. Linear refers to the fact that y and its derivatives occur to the first order.

Sometimes one call this an ordinary differential equation which means that it doesn't contain any partial derivative.

We write D^k for the k^{th} derivative $\frac{d^k}{dx^k}$, and $L(y)$ for the left hand side of (*),

$$L(y) = (a_n(x)D^n + \dots + a_1(x)D + a_0(x))(y).$$

The operator L is a linear transformation on the space $C^{(n)}$ of n -times continuously differentiable functions.

An equation $Ly = F(x)$ is called inhomogeneous if $F(x)$ is not identically zero. An equation $Ly = 0$ is said to be homogeneous.

The set of solutions to a homogeneous n^{th} order differential equation is a vector space of dimension n . If $y_1, \dots, y_n$ are n independent

solutions they form a basis for the solution space, so every solution to $Ly=0$ is a linear combination

$$y = c_1 y_1 + \dots + c_n y_n.$$

The textbook sketches a proof in the case $n=2$, but it relies upon an unproven result. Rather than repeating this I'll sketch the idea. A first order homogeneous differential equation $\frac{dy}{dx} + a_0(x)y = 0$ is

solved with a single integration: $\int \frac{dy}{y} = \int -a_0(x) dx$

So $y = Ce^{-\int a_0(x) dx}$ for a constant C . Thus, there is one independent solution. A 2nd order equation involves two constants of integration. So, for a first order equation we can specify an initial condition $y(x_0) = A$ and obtain a unique solution to the initial value problem

$$\frac{dy}{dx} + a_0(x)y = 0 \quad y(x_0) = A$$

For a second order equation we may specify two initial conditions because there are 2 constants of integration.

$$\frac{d^2 y}{dx^2} + a_1 \frac{dy}{dx} + a_0 y = 0 \quad y(x_0) = A, \quad y'(x_0) = B,$$

and this will determine a unique solution to the initial value problem.

For an n^{th} order equation there will be n constants and so we may specify n initial conditions $y(x_0) = A_1, \dots, y^{(n-1)}(x_0) = A_n$

We write D^k for $\frac{d^k}{dx^k}$ so our equation (*)

becomes $(a_n(x)D^n + \dots + a_1(x)D + a_0(x))y = 0$ ($F(x)=0$ in our case). The parentetical expression is a linear operator L on the space of n times differentiable functions, so we rewrite the equation as $Ly = 0$.

We look at the initial value problems

$$Ly = 0 \quad y(x_0) = 1 \quad y'(x_0) = \dots = y^{(n-1)}(x_0) = 0$$

$$Ly = 0 \quad y(x_0) = 0 \quad y'(x_0) = 1 \quad y^{(2)}(x_0) = \dots = y^{(n-1)}(x_0) = 0$$

$$Ly = 0 \quad y(x_0) = 0 = \dots = y^{(n-2)}(x_0), \quad y^{(n-1)}(x_0) = 1$$

Each one has a unique solution, and the Wronskian of the solutions is 1, so the solutions are independent.

Remark! We've seen that, if the Wronskian of n functions doesn't provide any information in the case where the Wronskian is identically zero. However, if the n functions are solutions to an n^{th} order homogeneous differential equation then the vanishing of the Wronskian at any point in the interval I means that the n functions are linearly dependent on I .

To see this let x_0 be a point at which $W(y_1, \dots, y_n) = 0$.

We study the system

$$c_1 y_1(x_0) + \dots + c_n y_n(x_0) = 0$$

⋮

$$c_1 y_1^{(n-1)}(x_0) + \dots + c_n y_n^{(n-1)}(x_0) = 0$$

The determinant of the system is the Wronskian $W(y_1, \dots, y_n)(x_0)$, and since this is zero our system has a non-trivial solution

$(a_1, \dots, a_n)$, and let $u(x) = a_1 y_1(x) + \dots + a_n y_n(x)$.

Then $y = u(x)$ satisfies the initial value problem

$$L(y) = 0, \quad y(x_0) = 0 = \dots = y^{(n-1)}(x_0)$$

The trivial solution $y(x) = 0$ also satisfies this initial value problem, so the uniqueness of the solution to an initial value problem forces $u(x)$ to be 0, i.e., $a_1 y_1(x) + \dots + a_n y_n(x) = 0$, but the a_i are not all zero. Thus, $y_1, \dots, y_n$ are linearly dependent.

Finally, we return to non-homogeneous equations $L(y) = F(x) \neq 0$. The general solution uses the same ideas as those used in finding the general solution to a non-homogeneous system of linear equations.

Suppose $u(x)$ and $v(x)$ are each solutions to $L(y) = F(x)$, so $L(u) = L(v) = F$. Then $L(u-v) = L(u) - L(v) = F - F = 0$,

so $u-v \in \text{Ker}(L)$, i.e., the general solution

to $L(y) = F(x)$ is $y = y_p + c_1 y_1 + \dots + c_n y_n$ where y_p is any particular solution to $L(y) = F$ and $y_1, \dots, y_n$ are a basis for the solution space to the homogeneous system $L(y) = 0$.

Let's practice re-writing differential equations in terms of the linear operator $D = \frac{d}{dx}$.

$$e^x \frac{d^3 y}{dx^3} + \sin x \cdot \frac{d^2 y}{dx^2} + 3 \frac{dy}{dx} + (\cos x) y = \tan x$$

$$\frac{d^3 y}{dx^3} = D^3 y, \quad \frac{d^2 y}{dx^2} = D^2 y \quad \text{so the equation becomes}$$

$$(e^x D^3 + \sin x D^2 + 3D + \cos x) y = \tan x.$$

We will mostly be concerned with differential equations with constant coefficients. So let's practice re-writing those. Write the following in terms of D .

$$(1) \quad 4 \frac{d^4 y}{dx^4} + 2 \frac{d^2 y}{dx^2} + 3 \frac{dy}{dx} - 7y = e^x$$

$$(2) \quad 9 \frac{d^8 y}{dx^8} + 12 \frac{d^6 y}{dx^6} + \frac{d^2 y}{dx^2} - y = \cos x.$$

Lecture 25

We begin with a study of homogeneous ordinary differential equations with constant coefficients.

$$y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_1y' + a_0y = 0.$$

We rewrite this as $P(D)y = 0$ where $P(D) = D^n + a_{n-1}D^{n-1} + \dots + a_1D + a_0$.

We say that $P(D)$ is a polynomial differential operator.

If r is a real variable the polynomial $P(r) = r^n + a_{n-1}r^{n-1} + \dots + a_0$ is called the auxiliary polynomial.

Remark: (1) If $P(D)$ and $Q(D)$ are polynomial differential operators then $P(D)Q(D) = Q(D)P(D)$. This follows from the fact that $(D-\alpha)(D-\beta) = D^2 - (\alpha+\beta)D + \alpha\beta = (D-\beta)(D-\alpha)$.

(2) If $P(D) = P_1(D) \dots P_r(D)$ then any solution to $P_i(D)y = 0$ is a solution to $P(D)y = 0$ (since

$$P(D)y = P_1(D) \dots P_{i-1}(D) P_{i+1}(D) \dots P_r(D) P_i(D)y = 0.)$$

(3) If r is a real or complex number then

$$\frac{d}{dx}(e^{rx}) = r e^{rx}$$

(4) If $u(x)$ is a sufficiently differentiable function

$$\text{then } \frac{d}{dx}(e^{rx}u(x)) = e^{rx}u' + u(x) \cdot r e^{rx}$$

$$D(e^{rx}u(x))$$

$$\text{So } (D-r)(e^{rx}u(x)) = e^{rx}u' + u(x) \cdot r e^{rx} - r u(x) e^{rx} = e^{rx}u'.$$

$$\text{Thus } (D-r)^m(e^{rx}u(x)) = e^{rx}u^{(m)}(x).$$

In the special case when $u(x) = x^k$ with $0 \leq k \leq (m-1)$

(2)

we get $(D-r)^m (x^k e^{rx}) = 0$.

Thus, the differential equation $(D-r)^m y = 0$ has

$e^{rx}, x e^{rx}, x^2 e^{rx}, \dots, x^{(m-1)} e^{rx}$ as solutions.

Moreover, these solutions are independent since $1, x, \dots, x^{(m-1)}$ are independent, so $e^{rx}, x e^{rx}, \dots, x^{(m-1)} e^{rx}$ a basis for the solution space for $(D-r)^m y = 0$.

Example: Find the general solution to
 $(D-1)^2 (D+2)^3 (D+7) y = 0$

$$y = C_1 e^x + C_2 x e^x + C_3 e^{-2x} + C_4 x e^{-2x} + C_5 x^2 e^{-2x} + C_6 e^{-7x}$$

Remark: A polynomial with real coefficients can always be factored into a product of linear polynomials, and irreducible quadratic polynomials with real coefficients

An irreducible quadratic polynomial $x^2 + bx + c$ can be factored over the complex numbers into linear factors $(x - (\alpha + \beta i))(x - (\alpha - \beta i))$, with α, β real numbers. ($b = -2\alpha, c = \alpha^2 + \beta^2$)

For example, $x^2 - 2x + 5 = (x - (1 + 2i))(x - (1 - 2i))$

Euler's formula says that $e^{(a+bi)x} = e^{ax} (\cos(bx) + i \sin(bx))$

Combining all of the above we get that the differential equation $P(D)y = (D^2 - 2\alpha D + (\alpha^2 + \beta^2))y = 0$ [with $P(D)$ an

irreducible quadratic differential operator) has

$y_1 = e^{ax}(\cos bx + i \sin bx)$ and $y_2 = e^{ax}(\cos bx - i \sin bx)$ as solutions.

Then $\frac{1}{2}(y_1 + y_2) = e^{ax} \cos(bx)$ and $\frac{1}{2i}(y_1 - y_2) = e^{ax} \sin(bx)$

are solutions, and these are both real. Moreover, they are independent since $\cos(bx)$ and $\sin(bx)$ are independent.

Example: Find the general solution to $P(D)y = (D^2 - 2D + 5)y = 0$.

We factor the auxiliary polynomial

$$P(r) = (r - (1 + 2i))(r - (1 - 2i)),$$

so the roots are $r = 1 \pm 2i$. Then $e^x \cos(2x)$ and $e^x \sin(2x)$ are a basis for the solution space, and

$$y = c_1 e^x \cos(2x) + c_2 e^x \sin(2x)$$

is the general solution to $P(D)y = 0$.

Problem: Find the general solution to

$$(D^2 - 4D + 13)y = 0.$$

We now put all of our work together. If

$P(D) = (D^2 - 2\alpha D + \alpha^2 + \beta^2)^m$ then the equation $P(D)y = 0$ has $e^{\alpha x} \cos(\beta x)$, $e^{\alpha x} \sin(\beta x)$, $x e^{\alpha x} \cos(\beta x)$, $x e^{\alpha x} \sin(\beta x)$,

$\dots x^{(m-1)} e^{\alpha x} \cos(\beta x), x^{(m-1)} e^{\alpha x} \sin(\beta x)$
as a basis for the solution space.

Example: Find the general solution to

$$P(D)y = (D^2+1)^2 (D-3)^2 (D+1)y = 0.$$

The roots of the auxiliary polynomial $P(r)$ are
 $r = (\pm i)$ with multiplicity 2
 $r = 3$ with multiplicity 2
 $r = -1$ with multiplicity 1.

Then, for $r = i$ (we have $a=0, b=1$) we see that
 $\cos x, \sin x, x \cos x, x \sin x$ are solutions.

For $r = 3$ $e^{3x}, x e^{3x}$ are solutions

For $r = -1$ e^{-x} is a solution

So, the general solution is

$$y = C_1 \cos x + C_2 \sin x + C_3 x \cos x + C_4 x \sin x + C_5 e^{3x} + C_6 x e^{3x} + C_7 e^{-x}$$

In summary, $P(r) = (r-r_1)^{n_1} \dots (r-r_k)^{n_k}$ with the r_i distinct ^{real or complex} numbers. The solutions to the differential equation $P(D)y = 0$ are as follows.

(1) If r_i is real then $e^{r_i x}, x e^{r_i x}, \dots, x^{(n_i-1)} e^{r_i x}$ are linearly independent solutions

(2) If $r_j = a \pm bi$ is complex (but not real, i.e., $b \neq 0$) then
 $e^{ax} \cos(bx), xe^{ax} \cos(bx), \dots, x^{(n_j-1)} e^{ax} \cos(bx)$

and $e^{ax} \sin(bx), xe^{ax} \sin(bx), \dots, x^{(n_j-1)} \sin(bx)$
 are linearly independent solutions.

(3) The general solution to $P(D)y = 0$ is a linear combination of the solutions in (1) and (2).

Problems: Find the general solution to $P(D)y = 0$ if

(1) $P(D) = (D^2 + 9)^3 (D^2 - 1)$

(2) $P(D) = (D^2 - 2D + 2)^2 (D^2 + 1)$

(3) $P(D) = (D^3 - 2D^2 - 4D + 8)^2$

Lecture 26

We turn our attention to non-homogeneous differential equations, i.e., equations of the form $P(D)y = F(x) \neq 0$. We already know that the general solution has the form $y = y_c + y_p$ where y_c is the general solution

to $P(D)y = 0$, and y_p is any particular solution to $P(D)y = F(x)$. In the last lecture we learned how to find y_c . Now we discuss some methods for finding a y_p . The first idea is simple, but it increases the amount of computation we must do.

We want to solve $P(D)y = F(x)$. Suppose $L(D)$ is a linear differential operator that annihilates $F(x)$, i.e., $L(D)F(x) = 0$. Then $L(D)P(D)y = L(D)F(x) = 0$, and this is a homogeneous equation that we already know how to solve.

Example: Solve $(D^2 - 4)y = e^x$.

We first notice that $De^x = e^x$, so $(D-1)e^x = 0$. We take $L(D) = (D-1)$. Then

$$(D-1)(D^2-4)y = (D-1)e^x = 0 \dots (*)$$

The auxiliary polynomial is $(r-1)(r^2-4)$, so the roots are $1, \pm 4$. The general solution to the homogeneous equation (*) is $y = c_1 e^x + c_2 e^{4x} + c_3 e^{-4x}$. The general solution to $(D^2-4)y = 0$ is $c_2 e^{4x} + c_3 e^{-4x}$, and we know that the general solution to

the non-homogeneous equation $(D^2-4)y=e^x$ is the sum of the general solution to the homogeneous equation $(D^2-4)y=0$ and any particular solution to $(D^2-4)y=e^x$. So, for some value of c , $y=c_1e^x+c_2e^{4x}+c_3e^{-4x}$ is the general solution to $(D^2-4)y=e^x$. To find the value of c , that works we substitute $y=c_1e^x$ into the non-homogeneous equation $(D^2-4)y=e^x$. We find

$$(D^2-4)(c_1e^x) = c_1e^x - 4c_1e^x = e^x,$$

so $-3c_1e^x = e^x$, or $-3c_1 = 1$. Thus, $c_1 = -\frac{1}{3}$, and the general solution to $(D^2-4)y=e^x$ is $y = -\frac{1}{3}e^x + c_2e^{4x} + c_3e^{-4x}$.

The part of the solution $y_p = c_1e^x$ is called a trial

solution. It contains the undetermined coefficient c_1 . For that reason this method is called the method of undetermined coefficients.

Example! Find the general solution to $(D^2+1)(D-2)y = e^{2x}$.

Since $De^{2x} = 2e^{2x}$ we see that $(D-2)$ annihilates e^{2x} , so $(D^2+1)(D-2)^2 = (D-2)e^{2x} = 0$.

The general solution to this homogeneous equation is $y = c_1\cos x + c_2\sin x + c_3e^{2x} + c_4xe^{2x}$.

(since the auxiliary polynomial $P(r) = (r^2+1)(r-2)^2$ has roots $r = \pm i, 2, 2$).

Now, the general solution to the homogeneous equation $(D^2+1)(D-2)y=0$ is $y=c_1\cos x+c_2\sin x+c_3e^{2x}$, so for some value of the undetermined coefficient c_4 we have $y_p=c_4xe^{2x}$ is a particular solution

to $(D^2+1)(D-2)y=e^{2x}$. To find c_4 we substitute $y_p=c_4xe^{2x}$ into the non-homogeneous equation. This requires a little calculation:-

$$D(xe^{2x}) = e^{2x} + 2xe^{2x} = e^{2x}(1+2x)$$

$$\text{So } (D-2)xe^{2x} = e^{2x}(1+2x) - 2xe^{2x} = e^{2x}$$

$$\text{Then } (D^2+1)(D-2)xe^{2x} = (D^2+1)e^{2x} = 4e^{2x} + e^{2x} = 5e^{2x},$$

$$\text{and } (D^2+1)(D-2)c_4xe^{2x} = 5c_4e^{2x}.$$

Substituting this into the non-homogeneous equation we find $(D^2+1)(D-2)c_4xe^{2x} = 5c_4e^{2x} = e^{2x}$,

$$\text{So } c_4 = \frac{1}{5}, \text{ and } y_p = \frac{1}{5}xe^{2x}.$$

The general solution to the non-homogeneous equation is $y=c_1\cos x+c_2\sin x+c_3e^{2x}+\frac{1}{5}xe^{2x}$.

Problem: Find the general solution to $(D-1)^2(D^2-9)y=e^{3x}$.

The Key to this method is finding a polynomial differential operator $L(D)$ that annihilates $F(x)$.

If $L_1(D)$ annihilates $F(x)$ and $L_2(D)$ annihilates $G(x)$ then $L_1(D)L_2(D)(aF(x)+bG(x)) = aL_1(D)L_2(D)(F(x)) + bL_1(D)L_2(D)(G(x)) = aL_2(D)L_1(D)(F(x)) + bL_1(D)L_2(D)(G(x))$

$= 0 + 0 = 0$, so $L_1(D)L_2(D)$ annihilates every linear combination of $F(x)$ and $G(x)$.

We note the following.

(1) $L(D) = (D-a)^{(k+1)}$ annihilates $e^{ax}, xe^{ax}, \dots, x^k e^{ax}$; and $(a_k x^k + \dots + a_1 x + a_0) e^{ax}$.

If $a=0$ we get $D^{(k+1)}$ annihilates $a_k x^k + \dots + a_0$.

(2) $L(D) = D^2 - 2aD + a^2 + b^2$ annihilates $e^{ax} \cos(bx)$ and $e^{ax} \sin(bx)$

and $L(D) = (D^2 - 2aD + a^2 + b^2)^{(k+1)}$ annihilates $e^{ax} \cos(bx), xe^{ax} \cos(bx), \dots, x^k e^{ax} \cos(bx)$
 $e^{ax} \sin(bx), xe^{ax} \sin(bx), \dots, x^k e^{ax} \sin(bx)$

Example: solve the initial value problem.

$$y'' + 16y = 4 \cdot \cos(x).$$

We notice that $D^2 + 16$ annihilates $4 \cos(x)$.

The differential equation can be re-written as

$$(D^2 + 16)y = 4 \cdot \cos(x). \quad (*)$$

The homogeneous equation $(D^2 + 16)y = 0$ has the auxiliary equation $P(r) = r^2 + 16$, and so has $\pm 4i$ as roots. The general solution to $(D^2 + 16)y = 0$ is $y = c_1 \cos(4x) + c_2 \sin(4x)$.

We multiply the non-homogeneous equation (*) by $(D^2 + 1)$ to obtain $(D^2 + 1)(D^2 + 16)y = (D^2 + 1)(4 \cos x) = 0$. The auxiliary equation for this equation is

$(r^2+1)(r^2+16)=0$, so its roots are $\pm i$ and $\pm 4i$. The general solution to $(D^2+1)(D^2+16)y=0$ is $y = c_1 \cos(4x) + c_2 \sin(4x) + c_3 \cos x + c_4 \sin x$.

Then, for a particular choice of c_3 and c_4 , $y_p = c_3 \cos x + c_4 \sin x$ is a particular solution to (*). To find the values of c_3 and c_4 we substitute y_p into (*).

$$(D^2+16)(c_3 \cos x + c_4 \sin x) = 4 \cos x$$

$$\underbrace{\quad}_{''} = c_3 \cos x - c_4 \sin x + 16c_3 \cos x + 16c_4 \sin x = 15c_3 \cos x + 15c_4 \sin x.$$

Thus, $15c_3 = 4$ and $c_4 = 0$, so $y_p = \frac{4}{15} \cos x$, and the

general solution to (*) is $y = c_1 \cos(4x) + c_2 \sin(4x) + \frac{4}{15} \cos x$

Problems: Find the general solution to

(1) $(D+3)(D^2+1)y = 2e^{-3x}$

(2) $y^{(3)} - y^{(2)} + y' - y = 9e^{-3}$

(3) $(D+3)(D-1)y = \sin^2 x$ (Hint: $\sin^2 x = \frac{1 - \cos 2x}{2}$).

Variation of Parameters Problem.

We were computing a particular solution to

$$y'' - 6y' + 9y = 4e^{3x} \ln x, \quad x > 0,$$

and I apparently miscopied the term for u_1' . The answer is worked out below.

We first solve the homogeneous equation $(D^2 - 6D + 9)y = 0$.

The general solution is $y = c_1 e^{3x} + c_2 x e^{3x}$. We look for a particular solution to the non-homogeneous equation of the form $y = u_1(x) e^{3x} + u_2(x) \cdot x e^{3x}$. (so $v_1 = e^{3x}$ $v_2 = x e^{3x}$).

After differentiating y we imposed the condition $u_1' v_1 + u_2' v_2 = 0$, so $y' = u_1 v_1' + u_2 v_2'$. Then $y'' = u_1' v_1' + u_1 v_1'' + u_2' v_2' + u_2 v_2''$. When we substitute y'' , y' , and y into the non-homogeneous equation we see that $u_1' v_1' + u_2' v_2' = F(x)$. Together with our imposed condition $u_1' v_1 + u_2' v_2 = 0$ we have a system of two equations in the variables u_1' and u_2' . The solution is

$$u_1' = \frac{-v_2 \cdot F(x)}{W(v_1, v_2)} \quad u_2' = \frac{v_1 \cdot F(x)}{W(v_1, v_2)} \quad \text{where } W \text{ is the Wronskian.}$$

$$\text{In our problem } W = \det \begin{pmatrix} e^{3x} & x e^{3x} \\ 3e^{3x} & e^{3x} + 3x e^{3x} \end{pmatrix} = e^{6x},$$

$$\text{So } u_1' = \frac{-x e^{3x} \cdot 4e^{3x} \ln x}{e^{6x}} \quad \text{and} \quad u_2' = \frac{e^{3x} \cdot 4e^{3x} \ln x}{e^{6x}}$$
$$= -4x \ln x \quad \quad \quad = 4 \ln x.$$

$$\text{Then } u_1 = \int -4x \ln x \, dx \quad \text{and } u_2 = \int 4 \ln x \, dx$$

We do each integral by parts

$$u_1 = \int -4x \ln x \, dx$$

$$\text{Let } u = \ln x \quad dv = -4x \, dx \\ du = \frac{1}{x} \, dx \quad v = -2x^2$$

$$u_1 = -2x^2 \ln x - \int -2x^2 \cdot \frac{1}{x} \, dx \\ = -2x^2 \ln x + x^2$$

$$u_2 = \int 4 \ln x \, dx$$

$$\text{Let } u = \ln x \quad dv = 4 \, dx \\ du = \frac{1}{x} \, dx \quad v = 4x$$

$$u_2 = 4x \ln x - \int 4 \, dx \\ = 4x \ln x - 4x$$

$$y = u_1 v_1 + u_2 v_2 = (-2x^2 \ln x + x^2) e^{3x} + (4x \ln x - 4x) x e^{3x}$$

Lecture 27

We have seen that the general solution to the homogeneous equation $L(D)y=0$ occurs as part of the general solution to the non-homogeneous equation $L(D)y=F(x)$. So, when we look for a particular solution to $L(D)y=F(x)$ we may exclude terms that correspond to the solution of $L(D)y=0$. In other words, we can choose a trial solution that doesn't contain the terms in the solution of $L(D)y=0$. The following table summarizes the explicit trial solutions for certain $F(x)$.

$F(x)$	Trial solution if $p(a) \neq 0$	Trial solution if a is a root of $P(r)=0$ of multiplicity m .
$Cx^k e^{ax}$	$y_p = e^{ax} (A_0 + \dots + A_k x^k)$	$y_p = x^m e^{ax} (A_0 + \dots + A_k x^k)$
$Cx^k e^{ax} \cos(bx)$ or $Cx^k e^{ax} \sin(bx)$	If $P(a+bi) \neq 0$ $y_p = e^{ax} [A_0 \cos bx + B_0 \sin bx + \dots + x^k (A_k \cos bx + B_k \sin bx)]$	If $a+bi$ is a root of $P(r)=0$ of multiplicity m $y_p = x^m e^{ax} [(A_0 \cos bx + B_0 \sin bx) + \dots + x^k (A_k \cos bx + B_k \sin bx)]$

Example 1. Find the general solution to

$$\frac{d^3 y}{dx^3} - 3 \frac{d^2 y}{dx^2} - 16 \frac{dy}{dx} + 48y = 6e^{3x}.$$

We first re-write this as an equation in terms of the differential operator D .

$$L(D)y = (D^3 - 3D^2 - 16D + 48)y = 6e^{3x}$$

The auxiliary equation is $P(r) = r^3 - 3r^2 - 16r + 48$. Its roots are $3, \pm 4$.

We consult the table and see that we should use the trial solution $y_p(x) = xe^{3x}(A)$. We substitute this trial solution into $L(D)y = 6e^{3x}$. We obtain:

$$\begin{aligned} L(D)y_p &= (D^2 - 4)(D - 3)(Axe^{3x}) = (D^2 - 4)Ae^{3x} \\ &= 9Ae^{3x} - 4Ae^{3x} = 5Ae^{3x} = 6e^{3x} \end{aligned}$$

So, $A = \frac{6}{5}$, and $y_p = \frac{6}{5}xe^{3x}$.

The solution to the homogeneous equation $L(D)y = 0$ is $y = c_1e^{3x} + c_2e^{4x} + c_3e^{-4x}$. Thus, the general solution to $L(D)y = 6e^{3x}$ is $y = \frac{6}{5}xe^{3x} + c_1e^{3x} + c_2e^{4x} + c_3e^{-4x}$.

Problem: Find the general solution to $y^{(3)} + 3y^{(2)} - 4y' - 12y = 4\cos x$.

There is a very slightly different approach to finding solutions to $L(D)y = F(x)$. Recall that when $r = a + bi$ is a complex root of the auxiliary polynomial then $e^{ax}\cos bx$ and $e^{ax}\sin x$ are independent real solutions to $L(D)y = 0$. The expressions $e^{ax}\cos bx$ and $e^{ax}\sin bx$ are the real and imaginary parts of $e^{(a+bi)x} = e^{ax}\cos bx + ie^{ax}\sin bx$. This observation leads to

the following.

Theorem: If $y(x) = u(x) + iv(x)$ is a complex solution to

$$y'' + a_1 y' + a_0 y = F(x) + iG(x)$$

then $u'' + a_1 u' + a_0 u = F(x)$ and $v'' + a_1 v' + a_0 v = G(x)$

The proof is just to equate the real parts of each side, and to equate the imaginary parts of each side.

Example: Find a particular solution to

$$y'' - y = e^{2x} \cos x.$$

We note that $e^{2x} \cos x$ is the real part of $e^{(2+i)x} = e^{2x} e^{ix} = e^{2x} (\cos x + i \sin x)$, so we look at the complex differential equation (*) $z'' - z = e^{(2+i)x}$. The auxiliary polynomial $P(r) = r^2 - 1$ has ± 1 as roots, so a trial solution for this equation is $Ae^{(2+i)x}$. In order to find A we substitute the trial solution into the complex differential equation

$$\begin{aligned} z'' - z &= A(2+i)^2 e^{(2+i)x} - A e^{(2+i)x} \\ &= A(2+i)^2 e^{(2+i)x} - A e^{(2+i)x} \\ &= A(3+4i) e^{(2+i)x} - A e^{(2+i)x} = A(2+4i) e^{(2+i)x} \\ &= e^{(2+i)x} \quad (\text{from } (*)) \end{aligned}$$

$$\text{So, } A(2+4i) = 1, \text{ and } A = \frac{1}{2+4i} = \frac{1}{2+4i} \cdot \frac{2-4i}{2-4i} = \frac{1-2i}{10}$$

$$\begin{aligned} \text{Then } z &= A e^{(2+i)x} = \frac{1-2i}{10} e^{2x} (\cos x + i \sin x) \\ &= \frac{e^{2x}}{10} (\cos x + i \sin x - 2i \cos x + 2 \sin x) \end{aligned}$$

is a complex solution to (*).

So, $y = \operatorname{Re}(z)$ (= real part of z) $= \frac{e^{2x}}{10} (\cos x + 2 \sin x)$.

is a real solution to $y'' - y = e^{2x} \cos x$.

Let's check this: $y' = \frac{1}{10} e^{2x} (4 \cos x + 3 \sin x)$

$$y'' = \frac{e^{2x}}{10} (11 \cos x + 2 \sin x)$$

and $y'' - y = e^{2x} \cos x$.

Problem: Use a complex-valued trial solution to solve $y'' - 4y' + 4y = e^x \sin x$.

Variation of Parameters

There is a useful method for finding the solution to a non-homogeneous equation that doesn't rely on the form of the non-homogeneous term, or the constancy of the coefficients of the equation, if we know the general solution to the homogeneous equation (so, for us the coefficients will still be constant).

We'd like to solve $y'' + a_1 y' + a_0 y = F(x)$ where a_1, a_0 , and F are continuous. Let v_1 and v_2 be a basis for the solution space of the homogeneous equation, so the general solution to the homogeneous equation is $c_1 v_1 + c_2 v_2$.

The idea is to replace the constants c_1 and c_2 with differentiable functions $u_1(x)$ and $u_2(x)$, and to look for a particular solution to the non-homogeneous equation of the form $y = u_1 v_1 + u_2 v_2$.

Differentiating y we get $y' = u_1' v_1 + u_1 v_1' + u_2' v_2 + u_2 v_2'$. If we try to compute y'' , and then substitute into the non-homogeneous equation we obtain a more complicated equation than our original one.

However, the non-homogeneous equation imposes only one condition on our functions v_1 and v_2 . This allows us to impose a second condition which we take to be $u_1' v_1 + u_2' v_2 = 0$.

With this condition y' becomes $y' = u_1 v_1' + u_2 v_2'$,

so $y'' = u_1' v_1' + u_1 v_1'' + u_2' v_2' + u_2 v_2''$. We substitute the expressions for y'' and y' into the non-homogeneous equation and collect terms to obtain

$$u_1 (v_1'' + a_1 v_1' + a_0 v_1) + u_2 (v_2'' + a_1 v_2' + a_0 v_2) + (u_1' v_1' + u_2' v_2') = F(x)$$

The coefficients of u_1 and u_2 are both zero since v_1 and v_2 are solutions to the homogeneous equation, so $u_1' v_1' + u_2' v_2' = F(x)$. Combining this with our imposed condition $u_1' v_1 + u_2' v_2 = 0$ we obtain the system

$$\begin{aligned} u_1' v_1 + u_2' v_2 &= 0 \\ u_1' v_1' + u_2' v_2' &= F(x) \end{aligned}$$

This is a linear system of two equations in the

unknowns u_1' and u_2' . Solving this system we obtain

$$u_1' = \frac{-v_2 \cdot F(x)}{W(v_1, v_2)} \quad \text{and} \quad u_2' = \frac{v_1 \cdot F(x)}{W(v_1, v_2)}$$

from Cramer's rule (or high school algebra).

Integrating we find the functions $u_1(x)$ and $u_2(x)$, and the particular solution $y = u_1 v_1 + u_2 v_2$ to the nonhomogeneous equation.

Lecture 28

We begin with an example of the variation of parameters method. We'll solve $y'' - 4y = \frac{8}{e^{2x} + 1}$.

The homogeneous equation $y'' - 4y = 0$ has auxiliary equation $r^2 - 4 = 0$, so $r = \pm 2$. Then $y = c_1 e^{2x} + c_2 e^{-2x}$ is the general solution to the homogeneous equation. We look for a solution to the non-homogeneous equation of the form $y = u_1 e^{2x} + u_2 e^{-2x}$ (so $v_1 = e^{2x}$ and $v_2 = e^{-2x}$).

$$\text{Then } u_1' e^{2x} + u_2' e^{-2x} = 0$$

$$\text{and } 2u_1' e^{2x} - 2u_2' e^{-2x} = \frac{8}{e^{2x} + 1}$$

$$u_1' = \underbrace{-e^{-2x} \cdot \frac{8}{e^{2x} + 1}}_W$$

$$u_2' = e^{2x} \cdot \frac{8}{\underbrace{e^{2x} + 1}_W}$$

$$W = \begin{vmatrix} e^{2x} & e^{-2x} \\ 2e^{2x} & -2e^{-2x} \end{vmatrix} = -4$$

$$u_1 = \int \frac{2e^{-2x}}{e^{2x} + 1} dx$$

$$u_2 = \int \frac{2e^{2x}}{e^{2x} + 1} dx$$

We'll find u_2 first - We let $w = e^{2x} + 1$, $dw = 2e^{2x} dx$
so $u_2 = \int \frac{dw}{w} = \ln|w| = \ln(e^{2x} + 1)$

Now we compute $u_1 = \int \frac{2e^{-2x}}{e^{2x}+1} dx$.

We let $w = e^{2x} + 1$,

so $\frac{1}{e^{2x}+1} = \frac{1}{w}$;

and $e^{-2x} = \frac{1}{w-1}$. Then $-2e^{-2x} dx = \frac{-1}{(w-1)^2} dw$

So, $u_1 = \int \frac{dw}{(w-1)^2 \cdot w}$

We compute this integral using partial fractions

$$\begin{aligned} \text{We write } \frac{1}{(w-1)^2 w} &= \frac{A}{w} + \frac{B}{w-1} + \frac{C}{(w-1)^2} \\ &= \frac{A(w-1)^2 + B(w-1)w + Cw}{w(w-1)^2} \end{aligned}$$

$$\text{Then } 1 = A(w-1)^2 + B(w-1)w + Cw =$$

$$A(w^2 - 2w + 1) + B(w^2 - w) + Cw =$$

$$w^2(A+B) + w(-2A-B+C) + A$$

$$\text{So, } A=1, B=-1, C=1$$

$$\text{Then } u_1 = \int \frac{dw}{w} + \int \frac{-dw}{w-1} + \int \frac{dw}{(w-1)^2}$$

$$= \ln|w| - \ln|w-1| - \frac{1}{w-1} = \ln(e^{2x}+1) - 2x - \frac{1}{e^{2x}}$$

And $y = c_1 e^{2x} + c_2 e^{-2x} + u_1 e^{2x} + u_2 e^{-2x}$ is the general solution

Problem: Find the general solution to

$$y'' - 6y' + 9y = 4e^{3x} \ln x, \quad x > 0.$$

Variation of parameters works for n^{th} order differential equations as well. If $L(D)y = F(x)$ is an n^{th} order differential equation we suppose $y = c_1 v_1 + \dots + c_n v_n$ is the solution to the homogeneous equation $L(D)y = 0$ we look for a solution to the non-homogeneous equation $L(D)y = F(x)$ of the form (*) $y = u_1(x)v_1(x) + \dots + u_n(x)v_n(x)$. We differentiate this and impose the condition that

$$u_1'v_1 + \dots + u_n'v_n = 0 \text{ so that}$$

$$y' = u_1v_1' + \dots + u_nv_n'$$

We differentiate Equation (*) k times, and for $1 \leq k \leq n-1$ we impose the conditions "the sum of the terms involving $u_1', \dots, u_n'$ is zero". We then

differentiate one more time, and solve the resulting system to find $u_1', \dots, u_n'$. Finally, we integrate to find $u_1(x), \dots, u_n(x)$.

Systems of Differential Equations

A first order linear system of differential equations is a system of the following form.

$$\begin{aligned}
 (*) \quad \frac{dx_1}{dt} &= a_{11}(t)x_1(t) + a_{12}(t)x_2(t) + \dots + a_{1n}(t)x_n(t) + b_1(t) \\
 &\vdots \\
 \frac{dx_n}{dt} &= a_{n1}(t)x_1(t) + a_{n2}(t)x_2(t) + \dots + a_{nn}(t)x_n(t) + b_n(t)
 \end{aligned}$$

or, if $\bar{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, $\frac{d\bar{x}}{dt} = A(t)\bar{x}(t) + \bar{b}(t)$, where

$$\bar{b}(t) = \begin{bmatrix} b_1(t) \\ \vdots \\ b_n(t) \end{bmatrix} \quad \text{and} \quad A(t) = (a_{ij}(t)).$$

If $\bar{b}(t) = \bar{0}$ the system is called homogeneous, and

if $\bar{b}(t) \neq \bar{0}$ the system is called non-homogeneous.

We may rewrite the system (*) as

$$D\bar{x} = A\bar{x} + \bar{b}(t)$$

$$\text{or } (D\bar{I} - A)\bar{x} = \bar{b}(t)$$

$$\begin{aligned}
 \text{i.e. } (D - a_{11})x_1 - a_{12}x_2 - \dots - a_{1n}x_n &= b_1(t) \\
 -a_{21}x_1 + (D - a_{22})x_2 - \dots - a_{2n}x_n &= b_2(t) \\
 &\vdots
 \end{aligned}$$

$$-a_{n1}x_1 - a_{n2}x_2 - \dots - (D - a_{nn})x_n = b_n(t).$$

With the appropriate manipulation we can eliminate $(n-1)$ variables leaving us with a higher order

equation in one variable which we can solve by our previous techniques. For example, if $n=2$ we consider the system

$$\frac{dx_1}{dt} = a_{11}x_1 + a_{12}x_2 + b_1(t)$$

$$\frac{dx_2}{dt} = a_{21}x_1 + a_{22}x_2 + b_2(t)$$

which we re-write as

$$(D - a_{11})x_1 - a_{12}x_2 = b_1 \quad \dots (1)$$

$$-a_{21}x_1 + (D - a_{22})x_2 = b_2 \quad \dots (2)$$

Multiplying (2) by a_{12} and applying $(D - a_{22})$ to (1) we get

$$(D - a_{22})(D - a_{11})x_1 - (D - a_{22})a_{12}x_2 = (D - a_{22})b_1(t) \quad \dots (3)$$

$$-a_{12}a_{21}x_1 - a_{12}(D - a_{22})a_{12}x_2 = a_{12}b_2(t) \quad \dots (4)$$

Subtracting (4) from (3) gives us a second order equation in $x_1(t)$.

Conversely, an n^{th} order equation

$$y^{(n)} + \dots + a_1 y' + a_0 y = f(x)$$

may be re-written as a system by introducing new variables.

For example, if $n=2$, the equation

$$y'' + a_1 y' + a_2 y = f(x)$$

can be re-written by letting $x_1 = y$, $x_2 = y'$. Then we have

$$\frac{dx_1}{dt} = x_2 \quad \text{and} \quad \frac{dx_2}{dt} = -a_1 x_2 - a_2 x_1 + f(x)$$

Even a higher order linear system may be re-written as a linear system. For example,

$$\frac{d^2 x}{dt^2} + a(t)x = b_1(t) \quad \dots$$

$$\frac{d^2 y}{dt^2} + c(t)x = b_2(t)$$

can be re-written by letting $x_1 = x(t)$, $x_2 = \frac{dx}{dt}$, $x_3 = y(t)$ and $x_4 = \frac{dy}{dt}$. Then the above system becomes

$$\frac{dx_2}{dt} + a(t)x_3 = b_1(t)$$

$$\frac{dx_3}{dt} + c(t)x_1 = b_2(t)$$

$$\frac{dx_1}{dt} = x_2$$

$$\frac{dx_3}{dt} = x_4$$

The point of the above is that there is nothing lost if we simply consider first order linear systems of differential equations.

Finally, we note that for the system $x'(t) = Ax + b(t)$

with $\bar{x} = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix}$ we may specify an initial condition $\bar{x}(t_0) = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$. This is called an initial value problem.

Lecture 29

Following the textbook we write $V_n(I)$ for the vector space of column n -vectors whose components are functions defined on the interval I .

Let $\bar{x}_1(t), \dots, \bar{x}_n(t) \in V_n(I)$. The Wronskian of these vectors is $W(\bar{x}_1, \dots, \bar{x}_n) = \det(\bar{x}_1, \dots, \bar{x}_n)$.

We observe that if $W(\bar{x}_1, \dots, \bar{x}_n)(t_0)$ is non-zero at any point $t_0 \in I$ then $\bar{x}_1(t), \dots, \bar{x}_n(t)$ are linearly independent (since if $c_1 \bar{x}_1 + \dots + c_n \bar{x}_n = \bar{0}$, or equivalently $[\bar{x}_1(t), \dots, \bar{x}_n(t)] \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \bar{0}$ then $\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \bar{0}$ since

$$\det[\bar{x}_1(t_0), \dots, \bar{x}_n(t_0)] = W(\bar{x}_1, \dots, \bar{x}_n)(t_0) \neq 0$$

If $A = (a_{ij}(t))$ is an $n \times n$ matrix of functions $a_{ij}(t)$, then the function $T: V_n(I) \rightarrow V_n(I)$ defined by $T(\bar{x}(t)) = A(t)\bar{x}(t)$ is a linear transformation, and the space of solutions to a homogeneous system $\bar{x}'(t) = A(t)\bar{x}(t)$ of differential equations is an n -dimensional subspace of $V_n(I)$. A basis $\bar{x}_1(t), \dots, \bar{x}_n(t)$ for this solution space is called a fundamental set of solutions for $\bar{x}' = A\bar{x}(t)$. The matrix $X = [\bar{x}_1, \dots, \bar{x}_n]$ is called a fundamental matrix for the differential equation $\bar{x}' = A\bar{x}$. Finally, we have a now familiar fact about non-homogeneous equations. Let $\bar{x}_1, \dots, \bar{x}_n$ be a fundamental set of solutions for the homogeneous

equation $\bar{x}'(t) = A\bar{x}(t)$. Then every solution to the non-homogeneous equation $\bar{x}'(t) = A(t)\bar{x}(t) + \bar{b}(t)$ is of the form $\bar{x}(t) = c_1\bar{x}_1(t) + \dots + c_n\bar{x}_n(t) + \bar{x}_p(t)$, where $\bar{x}_p(t)$ is any solution to $\bar{x}' = A\bar{x}(t) + \bar{b}(t)$.

We've seen the proof of an analogous fact in other contexts - Let $\bar{u}(t), \bar{v}(t)$ be solutions to $\bar{x}' = A\bar{x} + \bar{b}(t)$, so $\bar{u}' = A\bar{u} + \bar{b}(t)$ and $\bar{v}' = A\bar{v} + \bar{b}(t)$. Then

$$\bar{u}' - \bar{v}' = A\bar{u} + \bar{b}(t) - (A\bar{v} + \bar{b}(t)) = A\bar{u} - A\bar{v} = A(\bar{u} - \bar{v}),$$

so any two solutions to the non-homogeneous system differ by a solution to the homogeneous system.

We now try to solve a homogeneous system $\bar{x}' = A\bar{x}$. We begin with the case when A is non-defective.

Theorem: Let A be an $n \times n$ matrix whose entries are real numbers (not functions). Let λ be an eigenvalue of A with $\bar{v}$ a corresponding eigenvector. (so $A\bar{v} = \lambda\bar{v}$). Then $\bar{x}(t) = e^{\lambda t}\bar{v}$ is a solution to $\bar{x}' = A\bar{x}$.

proof: If $\bar{x}(t) = e^{\lambda t}\bar{v}$ then $\bar{x}'(t) = \lambda e^{\lambda t}\bar{v} = e^{\lambda t}\lambda\bar{v} = e^{\lambda t}A\bar{v} = A e^{\lambda t}\bar{v}$, so $\bar{x}' = A\bar{x}$.

Example: Find a fundamental set of solutions for $\bar{x}' = A\bar{x}(t)$ if $A = \begin{pmatrix} 1 & 2 & 4 \\ 0 & 3 & 1 \\ 0 & 0 & 2 \end{pmatrix}$.

We can read off the eigenvalues since A is upper triangular - $\lambda = 1, 3, 2$.

We compute an eigenvector corresponding to each eigenvalue.

$$\lambda=1 \quad \begin{pmatrix} 1 & 2 & 4 \\ 0 & 3 & 1 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} u \\ v \\ w \end{pmatrix}$$

$$\text{So } u+2v+4w=u$$

$$3v+w=v \leftarrow v=0$$

$$2w=w \leftarrow w=0$$

, so $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ is an eigenvector corresponding to $\lambda=1$.

$$\lambda=3 \quad \begin{pmatrix} 1 & 2 & 4 \\ 0 & 3 & 1 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix} = 3 \begin{pmatrix} u \\ v \\ w \end{pmatrix}$$

$$u+2v+4w=3u$$

$$3v+w=3v$$

$$2w=3w \leftarrow w=0$$

Then $u+2v=3u$ (since $w=0$), so $2v=2u$ and $v=u$

$\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ is an eigenvector corresponding to $\lambda=3$.

$$\lambda=2 \quad \begin{pmatrix} 1 & 2 & 4 \\ 0 & 3 & 1 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix} = 2 \begin{pmatrix} u \\ v \\ w \end{pmatrix}$$

$$u+2v+4w=2u$$

$$3v+w=2v \leftarrow v=-w$$

$$2w=2w$$

Then $u=2w$, so $\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$ is an eigenvector corresponding

to 2.

Then $e^t \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, $e^{3t} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $e^{2t} \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$ is a fundamental set of solutions.

The key point in the last problem is that the solution space is 3 dimensional, and we have 3 independent solutions corresponding to our three independent eigenvectors since the Wronskian

$$W = \det \left[e^t \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, e^{3t} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, e^{2t} \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} \right] = e^{(t+3t+2t)} \det \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \neq 0.$$

This is a general fact - The solution space to $\bar{x}' = A\bar{x}(t)$ is n -dimensional if A is an $n \times n$ matrix. If A is non-defective with eigenvalues $\lambda_1, \dots, \lambda_n$ and corresponding eigenvectors $\bar{v}_1, \dots, \bar{v}_n$ then $\{e^{\lambda_1 t} \bar{v}_1, \dots, e^{\lambda_n t} \bar{v}_n\}$ is a fundamental set of solutions since $W(e^{\lambda_1 t} \bar{v}_1, \dots, e^{\lambda_n t} \bar{v}_n) \neq 0$.

$$\det(e^{\lambda_1 t} \bar{v}_1, \dots, e^{\lambda_n t} \bar{v}_n) = e^{(\lambda_1 + \dots + \lambda_n)t} \det(\bar{v}_1, \dots, \bar{v}_n) \neq 0.$$

Problem: Find the general solution to $\bar{x}'(t) = A\bar{x}(t)$ if $A = \begin{pmatrix} -2 & -7 \\ -1 & 4 \end{pmatrix}$.

(5)

Recall that A is diagonalizable if and only if A is non-defective. In that case, if S is the matrix whose columns are n independent eigenvectors for A then $S^{-1}AS = D$ is the diagonal matrix where the i^{th} diagonal entry is the eigenvalue corresponding to the i^{th} column of S .

We examine the system $\bar{x}'(t) = A\bar{x}$.

We make a change of variables $\bar{y} = S^{-1}\bar{x}$ (so $\bar{x} = S\bar{y}$).

Then $\bar{x}' = S\bar{y}'$ and $\bar{x}' = A\bar{x} = SDS^{-1}\bar{x} = SD\bar{y}$
 $\parallel$
 $S\bar{y}'$

Finally, since S is invertible we get $\bar{y}' = D\bar{y}$.

A fundamental matrix for this system is $\begin{pmatrix} e^{\lambda_1 t} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n t} \end{pmatrix}$

Since $\bar{x} = S\bar{y}$ we see a fundamental matrix for $\bar{x}' = A\bar{x} = (e^{\lambda_1 t}\bar{v}_1, \dots, e^{\lambda_n t}\bar{v}_n)$.

We have seen that a real matrix may have complex eigenvalues and eigenvectors. What happens in this case?

Suppose $\bar{w}_1(t) = \bar{u}(t) + i\bar{v}(t)$ and $\bar{w}_2(t) = \bar{u}(t) - i\bar{v}(t)$ are complex conjugate solutions to $\bar{x}' = A\bar{x}$. Then $\frac{1}{2}(\bar{w}_1 + \bar{w}_2) = \bar{u}(t) (= \operatorname{Re}(\bar{w}_1))$ and $\frac{1}{2i}(\bar{w}_1 - \bar{w}_2) = \bar{v}(t) (= \operatorname{Im}(\bar{w}_1))$ are also solutions to $\bar{x}' = A\bar{x}$, but $\bar{u}(t)$ and $\bar{v}(t)$ are real.

(6)

Example: Find the general solution to $\bar{x}'(t) = A\bar{x}(t)$ if
 $A = \begin{pmatrix} -2 & -6 \\ 3 & 4 \end{pmatrix}$.

The characteristic polynomial is $\det(A - \lambda I) =$

$$\det \begin{pmatrix} -2-\lambda & -6 \\ 3 & 4-\lambda \end{pmatrix} = (-2-\lambda)(4-\lambda) + 18 = \lambda^2 - 2\lambda + 10$$

So, the eigenvalues are $\lambda = 1 \pm 3i$.

An eigenvector corresponding to $1+3i$ is $\begin{bmatrix} i-1 \\ 1 \end{bmatrix}$

$$\text{Then } \bar{x}(t) = e^{(1+3i)t} \begin{bmatrix} i-1 \\ 1 \end{bmatrix} = e^t (\cos(3t) + i\sin(3t)) \begin{bmatrix} i-1 \\ 1 \end{bmatrix}$$

$$= e^t \begin{bmatrix} i\cos(3t) - \sin(3t) - \cos(3t) - i\sin(3t) \\ \cos(3t) + i\sin(3t) \end{bmatrix} =$$

$$e^t \begin{bmatrix} -\sin(3t) - \cos(3t) \\ \cos(3t) \end{bmatrix} + ie^t \begin{bmatrix} \cos(3t) - \sin(3t) \\ \sin(3t) \end{bmatrix}$$

The real and imaginary parts are each solutions,
 so $\bar{x}_1 = e^t \begin{bmatrix} -\sin(3t) - \cos(3t) \\ \cos(3t) \end{bmatrix}$ and $\bar{x}_2 = e^t \begin{bmatrix} \cos(3t) - \sin(3t) \\ \sin(3t) \end{bmatrix}$

are a fundamental set of solutions, and

$$\bar{x}(t) = c_1 \bar{x}_1 + c_2 \bar{x}_2$$

is the general solution.

Problem: Find the general solution to $\bar{x}' = A\bar{x}$
 if $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

Lecture 30

Let A be an $n \times n$ matrix. Recall that an eigenvector $\bar{v}$ (corresponding to the eigenvalue λ) is a vector satisfying $A\bar{v} = \lambda\bar{v}$. Equivalently, $\bar{v}$ satisfies $(A - \lambda I)\bar{v} = \bar{0}$. We are interested in vectors $\bar{x}$ satisfying $(A - \lambda I)^k \bar{x} = \bar{0}$ for some positive integer k . These vectors are called generalized eigenvectors.

If $\bar{v}$ is an eigenvector, and $\bar{w}$ is a vector satisfying $(A - \lambda I)^{(k-1)} \bar{w} = \bar{v}$ then $(A - \lambda I)^k \bar{w} = (A - \lambda I)\bar{v} = \bar{0}$, so $\bar{w}$ is a generalized eigenvector.

Example: let $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. The only eigenvalue is 1, and

$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is an eigenvector. If $\begin{bmatrix} x \\ y \end{bmatrix}$ satisfies $(A - I)\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ then $y = 1$, so $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ is a generalized eigenvector.

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Now, suppose we want to find the general solution to $\bar{x}'(t) = A\bar{x}(t)$ with $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Since A is

defective our earlier work won't give us a fundamental set of solutions, so we go back to basic algebra. The system $\bar{x}' = A\bar{x}$ is the matrix form of the system

$$\begin{aligned} \bar{x}_1' &= x_1 + x_2 \\ \bar{x}_2' &= x_2 \end{aligned} \quad , \quad \text{where } \bar{x}(t) = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

We re-write this in terms of the differential operator D :

$$\begin{aligned} (D-1)x_1 - x_2 &= 0 & \dots (1) \\ (D-1)x_2 &= 0 & \dots (2) \end{aligned}$$

Multiplying the first equation by $(D-1)$ we get

$$(D-1)^2 x_1 - (D-1)x_2 = 0.$$

So, $(D-1)^2 x_1 = 0$. The auxiliary polynomial is $P(r) = (r-1)^2$, and so 1 is a root of multiplicity 2. Then $x_1(t) = c_1 e^t + c_2 t e^t$. Substituting this into equation (1) we get

$$x_2 = (D-1)x_1 = c_2 e^t.$$

$$\text{Then } \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} c_1 e^t + c_2 t e^t \\ c_2 e^t \end{bmatrix} = c_1 e^t \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2 e^t \left(\begin{bmatrix} 0 \\ 1 \end{bmatrix} + t \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right)$$

Recall that $\bar{v} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is an eigenvector, and $\bar{w} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ is a generalized eigenvector, so our solution is

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = c_1 e^t \bar{v} + c_2 e^t (\bar{w} + t \bar{v})$$

This example suggests that, in the case of a defective 2×2 matrix A , the general solution to $\bar{x}'(t) = A\bar{x}(t)$ is $\bar{x}(t) = c_1 e^{\lambda t} \bar{v} + c_2 e^{\lambda t} (\bar{w} + t \bar{v})$. In fact, that is exactly what happens.

To see this suppose there is a solution to $\bar{x}' = A\bar{x}$ of the form $e^{\lambda t}(\bar{v}_1 + t\bar{v}_0)$ for some vectors $\bar{v}_1$ and $\bar{v}_0$. Then $\bar{x}' = e^{\lambda t}(\lambda\bar{v}_1 + \bar{v}_0) + \lambda t\bar{v}_0$, and $A\bar{x} = e^{\lambda t}(A\bar{v}_1 + tA\bar{v}_0)$. If $\bar{x}' = A\bar{x}$ then these two expressions are equal, so $(\lambda\bar{v}_1 + \bar{v}_0 + \lambda t\bar{v}_0) = A\bar{v}_1 + tA\bar{v}_0 = A\bar{v}_1 + tA\bar{v}_0$.

$$\text{So, } A\bar{v}_1 = \lambda\bar{v}_1 + \bar{v}_0 \text{ or } (A - \lambda I)\bar{v}_1 = \bar{v}_0 \\ \text{and } \lambda\bar{v}_0 = A\bar{v}_0.$$

Thus, $\bar{v}_0$ is an eigenvector corresponding to the eigenvalue λ , and $\bar{v}_1$ is a generalized eigenvector.

Suppose $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$. This matrix is defective with the algebraic multiplicity of 1 equal to 3, while the geometric multiplicity is only 2. The vectors $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ are independent eigenvectors.

If we look for a generalized eigenvector, we can try to find $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ with $(A - I)\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

The top row corresponds to $0x + 0y + 0z = 1$, which is impossible. However, if we look for $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$

satisfying $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ which tells us $z = 1$, so $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ is a generalized eigenvector.

The general solution to $\bar{x}' = A\bar{x}$ when $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$

$$\text{is } y = c_1 e^t \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + c_2 e^t \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + c_3 \left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + t \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right).$$

The two independent eigenvectors each give rise to a solution, and the generalized eigenvector gives us a 3rd independent solution.

In general, if λ has algebraic multiplicity m and geometric multiplicity $r < m$ then there are m linearly independent solutions to $\bar{x}' = A\bar{x}$ which are built in cycles of the form

$$\bar{x}_0(t) = e^{\lambda t} \bar{v}_0$$

$$\bar{x}_1(t) = e^{\lambda t} (\bar{v}_1 + t \bar{v}_0)$$

$$\vdots$$

$$\bar{x}_k(t) = e^{\lambda t} \left(\bar{v}_k + t \bar{v}_{k-1} + \dots + \frac{1}{k!} t^k \bar{v}_0 \right)$$

$$\text{where } (A - \lambda I)^{(k+1)} \bar{v}_k = \bar{0} \text{ and } (A - \lambda I)^k \bar{v}_k \neq \bar{0}$$

There are r cycles of the above form (with varying k), and the sum of the varying k is $m - r$.

Applying this to each eigenvalue yields a fundamental set of solutions to $\bar{x}'(t) = A\bar{x}$.

Problem: Find the general solution to $\bar{x}' = A\bar{x}$ if (i) $A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 2 \end{pmatrix}$.

(5)

$$(2) \cdot A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$

$$(3) \quad A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$

Non-homogeneous Systems

We're interested in solving the system $\bar{x}'(t) = A\bar{x}(t) + \bar{b}(t)$, when $\bar{b}(t) \neq \bar{0}$.

Suppose $Y = [\bar{x}_1, \dots, \bar{x}_n]$ is a fundamental matrix for $\bar{x}'(t) = A\bar{x}(t)$, so the general solution to $\bar{x}' = A\bar{x}$ is $\bar{x}(t) = c_1\bar{x}_1 + \dots + c_n\bar{x}_n = Y \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$.

We use the variation of parameters trick, and look for a solution of the form

$$\bar{x}_p(t) = Y \begin{bmatrix} \bar{u}_1(t) \\ \vdots \\ \bar{u}_n(t) \end{bmatrix} = Y \bar{u}(t)$$

$$\text{Then } \begin{matrix} \bar{x}_p' \\ \text{"} \end{matrix} = A\bar{x}_p + \bar{b}(t) = AY\bar{u}(t) + \bar{b}(t)$$

$$Y'\bar{u} + Y\bar{u}'$$

$$\text{So, } Y' = AY \text{ and } Y\bar{u}' = \bar{b}(t)$$

$$\text{Then } \bar{u}' = Y^{-1}\bar{b} \text{ and } \bar{u} = \int Y^{-1}\bar{b}(t)dt$$

$$\text{Finally, } \bar{x}_p = Y\bar{u}(t)$$

(6)

Example: Let $A = \begin{pmatrix} 4 & -3 \\ 2 & -1 \end{pmatrix}$, $\bar{b}(t) = \begin{pmatrix} e^{2t} \\ e^t \end{pmatrix}$.

Solve $\bar{X}' = A\bar{X} + \bar{b}$.

We first solve the homogeneous equation $\bar{X}' = A\bar{X}$.

$$\det(A - \lambda I) = \det \begin{pmatrix} 4-\lambda & -3 \\ 2 & -1-\lambda \end{pmatrix} = (4-\lambda)(-1-\lambda) + 6 = \lambda^2 - 3\lambda + 2 = 0$$

So, $\lambda = 1, 2$.

$$\lambda = 1 \quad \begin{pmatrix} 4-3 \\ 2-1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \quad 4x - 3y = x \quad \text{so } x = y$$

$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector

$$\lambda = 2 \quad \begin{pmatrix} 4-3 \\ 2-1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 2 \begin{pmatrix} x \\ y \end{pmatrix} \quad 4x - 3y = 2x \quad \text{so } y = \frac{2}{3}x$$

$\begin{bmatrix} 3 \\ 2 \end{bmatrix}$ is an eigenvector

Then $\begin{pmatrix} 3e^{2t} & e^t \\ 2e^{2t} & e^t \end{pmatrix}$ is a fundamental matrix.

$$Y^{-1} = e^{-3t} \begin{pmatrix} e^t & -e^t \\ -2e^{2t} & 3e^{2t} \end{pmatrix} = \begin{pmatrix} e^{-2t} & -e^{-2t} \\ -2e^{-t} & 3e^{-t} \end{pmatrix}$$

$$Y^{-1}\bar{b}(t) = \begin{pmatrix} e^{-2t} & -e^{-2t} \\ -2e^{-t} & 3e^{-t} \end{pmatrix} \begin{pmatrix} e^{2t} \\ e^t \end{pmatrix} = \begin{pmatrix} 1 - e^{-t} \\ -2e^t + 3 \end{pmatrix}$$

$$\bar{u} = \int \begin{pmatrix} 1 - e^{-t} \\ -2e^t + 3 \end{pmatrix} dt = \begin{pmatrix} t + e^{-t} \\ -2e^t + 3t \end{pmatrix}$$

$$\bar{X}_p = Y\bar{u} = \begin{pmatrix} 3e^{2t} & e^t \\ 2e^{2t} & e^t \end{pmatrix} \begin{pmatrix} t+e^{-t} \\ -2e^t+3t \end{pmatrix}$$

and $\bar{X} = Y \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \bar{X}_p$ is the general solution to the non-homogeneous system.