MAT 308:
Spring 2017

Home
General Information
Syllabus
Homework
Exams

Welcome to MAT 308

Textbook: Multivariable Mathematics, (4th ed.) by Williamson and Trotter.

Lecturer: Sabyasachi Mukherjee

Office: Math Tower 4115

Office Hours: F 10:30am-12:30pm in my office, M 1:00pm-2:00pm in MLC (S235), or by appt.

E-mail: *sabya@math.stonybrook.edu*

Homework

Homework assignments will be posted **here** and on BlackBoard. Please hand them in to your recitation instructor the following week. Please note that your TA will NOT accept late homework.

Quizzes

There will be a short quiz in your recitation session every other week. The first quiz will be taken in the week of Feb 6 - Feb 10.

Exams and Grading

There will be two midterms, and a final exam (dates **here**), whose weights in the overall grade are listed below.

15% Homework

	10% Quizzes 20% Midterm 1 20% Midterm 2 35% Final Exam (cumulative)

MAT 308: Spring 2017

General Information

Stony Brook University expects students to maintain standards of personal integrity that are in harmony with the educational goals of the institution; to observe national, state, and local laws as well as University regulations; and to respect the rights, privileges, and property of other people. Faculty must notify the Office of Judicial Affairs of any disruptive behavior that interferes with their ability to teach, compromises the safety of the learning environment, or inhibits students' ability to learn.

Information for students with disabilities
If you have a physical, psychological, medical, or learning disability that may impact your course work, please contact Disability Support Services at (631) 632-6748 or http://studentaffairs.stonybrook.edu/dss/. They will determine with you what accommodations are necessary and appropriate. All information and documentation is confidential.

Students who might require assistance during emergency evacuation are encouraged to discuss their needs with their professors and Disability Support Services. For procedures and information go to the following website: http://www.sunysb.edu/ehs/fire/disabilities.shtml

MAT 308: Spring 2017

Home
General Information
Syllabus
Homework
Exams

Syllabus and Weekly Plan

Week of	Topics	Comments
Jan 23	10.1: 1st order DE, direction fields 10.2: Separation of variables 10.3: linear equations, integrating factors	
Jan 30	3.1: **Linear Maps/Euclidean spaces** 3.2, 3.3: **Vector Spaces and Linear Maps**	
Feb 6	3.4, 3.5 **Image and Null Space, Coordinates and Dimension** 3.6 **Eigenvalues and Eigenvectors**	
Feb 13	3.6 Eigenvalues and Eigenvectors 3.7 **Inner Products**	
Feb 20	Ch.3/Ch.10/Midterm Review	No HW/Quiz this week,

	Midterm I, Wed. Feb 22	Midterm I in class.
Feb 27	11.1, 11.2 **Differential Operators, Complex Solutions, Higher Order Eqns** 11.3 Non-homogeneous Eqns	
March 6	11.5 **Laplace Transform** 11.6 **Convolution**	
March 13	Spring Break	
March 20	12.1 Vector Fields 12.2 Linear Systems	
March 27	**Sequences and Series in Normed Vector Spaces** **Definition and Basic Properties of Matrix Exponential**	
April 3	**Jordan Canonical Form, Computing Matrix Exponential** 13.1, 13.2 Applications of Diagonalization and Matrix Exponential to Linear Systems	
April 10	Midterm Review Midterm, Wed. April 12	No HW/Quiz this week, Midterm II in class.

April 17	**Nonhomogenous Linear Systems** 14.7 Power Series Solutions	
April 24	13.4 **Equilibrium and Stability**	
May 1	Final Review	
May 9	Final Exam Tuesday, May 9, 8:30pm-11:00pm	

MAT 308: Spring 2017

Homework

Homework Set 1

Homework Set 2

Homework Set 3

Homework Set 4

Homework Set 5

Homework Set 6

Homework Set 7

Homework Set 8

Homework Set 9

MAT 308:
Spring 2017

Exams

Midterm I, Wed. Feb 22

Midterm II, Wed. April 12

Final Exam, Tue. May 9

Differential equations and linearity

Consider the differential equation

$$\frac{dy}{dx} + y = 0 \quad \rightarrow (*)$$

Let $y_1(x)$ and $y_2(x)$ be two distinct solutions of $(*)$.

Then for any two real constants a, b,

we've:

$$\frac{d}{dx}(ay_1 + by_2) + (ay_1 + by_2)$$

$$= a\left(\frac{dy_1}{dx} + y_1\right) + b\left(\frac{dy_2}{dx} + y_2\right)$$

$$= a \cdot 0 + b \cdot 0 \quad \left[\text{Since } y_1 \text{ and } y_2 \text{ are solutions of } (*)\right]$$

$$= 0.$$

This shows that $(ay_1 + by_2)$ is also a solution of $(*)$.

We just observed that the space S of solutions of the differential equation $(*)$ satisfies a certain linearity property:

if $a, b \in \mathbb{R}$, and $y_1, y_2 \in S$, then $(ay_1 + by_2) \in S$.

Hence, any linear combination of elements of S also lies in S.

Loosely, such a space S is called a linear/vector space.

This leads us to the study of linear spaces.

Moreover, the derivative operator also satisfies a linearity property. For any two real constants a, b, and any two differentiable functions y_1, y_2, we've:

$$\frac{d}{dx}(a y_1 + b y_2) = a \frac{dy_1}{dx} + b \frac{dy_2}{dx} \rightarrow (**)$$

~~Loosely~~ An operator with the linearity property $(**)$ is called a

linear operator.

The upshot of the preceding analysis is that the study of derivatives and differential equations naturally lead us to the study of linear spaces

and linear operators.

Before we start a formal discussion of linear spaces/operators, let us restrict our attention ~~to~~ a concrete example; the simplest

linear space $\mathbb{R}^n$.

Note: Understanding an example well enough makes the study of an abstract concept much simpler.

Definition ($\mathbb{R}^n$):

The real Euclidean space of dimension n is defined as the cartesian product:

$$\{(x_1, x_2, \dots, x_n) : x_1, x_2, \dots, x_n \in \mathbb{R}\}.$$

In other words, $\mathbb{R}^n$ consists of all n-tuples of real numbers (we'll sometimes call them vectors).
There's a natural notion of addition on $\mathbb{R}^n$:

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n)$$

One can also define a scalar multiplication. For any $c \in \mathbb{R}$ and $(x_1, \ldots, x_n) \in \mathbb{R}^n$, one defines:

$$c(x_1, \ldots, x_n) = (cx_1, \ldots, cx_n).$$

Definition (Linear maps/transformations on $\mathbb{R}^n$):

A map $T: \mathbb{R}^n \to \mathbb{R}^m$ is called linear if it satisfies

$$T(au + bv) = aT(u) + bT(v),$$

where $a, b \in \mathbb{R}$, and $u, v \in \mathbb{R}^n$.

Example:

1) Let A be an $m \times n$ real matrix; i.e. A has m rows and n columns

In particular:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

Any element of $\mathbb{R}^n$ is of the form:

$$u = (x_1, \ldots, x_n).$$

We define a linear map $T: \mathbb{R}^n \to \mathbb{R}^m$ by:

$$T\big((x_1, \ldots, x_n)\big) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

$$= A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

This is just the usual mult. of a matrix and a column vector

$$= Au$$

$$= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ \text{---} \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix}$$

clearly, $T\big((x_1, \ldots, x_n)\big) \in \mathbb{R}^m$.

Also, matrix multiplication satisfies:

$$A(au + bv) = aAu + bAv,$$

for any $a, b \in \mathbb{R}$, $u, v \in \mathbb{R}^n$

Therefore, T is a linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$.

2) Reflection in $\mathbb{R}^2$ (wrt a line).

3) Rotation in $\mathbb{R}^2$ (wrt the origin),

4) Scaling.

We'll see in class (geometrically) why these define linear maps.

Example (1) had a concrete algebraic description in terms of a matrix. One can ask whether every linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$ has a similar representation.

We'll now proceed to answer this question affirmatively. Let's start with the notion of a basis.

In $\mathbb{R}^n$, the vectors

$$e_1 = (1, 0, \ldots, 0),\ e_2 = (0, 1, 0, \ldots, 0), \ldots, e_n = (0, \ldots, 0, 1)$$

play a special role.

Indeed, any vector $(x_1, x_2, \dots x_n)$ in $\mathbb{R}^n$ can be written as a linear combination of $e_1, \dots, e_n$:

$$(x_1, \dots, x_n) = x_1 (1, 0, \dots, 0) + \dots + x_n (0, \dots, 0, 1)$$

$$= x_1 e_1 + x_2 e_2 + \dots + x_n e_n.$$

we'll see later that $\{e_1, \dots, e_n\}$ is a basis of $\mathbb{R}^n$.

A simple yet crucial observation:

Since the vectors $\{e_1, \dots, e_n\}$ span/generate all of $\mathbb{R}^n$, it is enough to understand the action of a linear map $T: \mathbb{R}^n \to \mathbb{R}^m$ on the vectors $\{e_1, \dots, e_n\}$.

More precisely, any element of $\mathbb{R}^n$ can be written as $x_1 e_1 + \dots + x_n e_n$, for some real numbers $x_1, \dots, x_n$.

Then, $T(x_1 e_1 + \dots + x_n e_n)$

$$= x_1 T(e_1) + \dots + x_n T(e_n).$$

Thus, the action of T on $\{e_1, \dots, e_n\}$ determines the action of T on all of $\mathbb{R}^n$.

Theorem: Let $T: \mathbb{R}^n \to \mathbb{R}^m$ be a linear map. Then there exists an $m \times n$ real matrix A such that for any $u = (x_1, \dots, x_n) \in \mathbb{R}^n$, we've

$$Tu = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Proof: Note that $T(e_1), T(e_2), \dots, T(e_n)$ are elements of $\mathbb{R}^m$ and we can think of them as column vectors.
Let A be ~~a~~ the $m \times n$ matrix whose columns ~~are~~ are:

$$T(e_1), \dots, T(e_n);$$

i.e. the i-th column of A is $T(e_i)$.

Now, a direct computation shows that

$$Ae_i = A \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \leftarrow \text{ith row} = \text{The i-th column of } A.$$

Therefore, we've: $Ae_i = T(e_i)$, for $i = 1, \dots, n$.

By linearity, this yields:

$$T(x_1 e_1 + \cdots + x_n e_n)$$
$$= x_1 T(e_1) + \cdots + x_n T(e_n)$$
$$= x_1 Ae_1 + \cdots + x_n Ae_n$$
$$= A(x_1 e_1 + \cdots + x_n e_n).$$
$$= A\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

$\boxtimes$

Remark: The ~~zero~~ matrix A (as above) is called the matrix of the linear map T with respect to $\{e_1, \ldots, e_n\}$.

Composition of linear maps and multiplication of matrices:

Let $T: \mathbb{R}^n \to \mathbb{R}^m$ and $S: \mathbb{R}^m \to \mathbb{R}^k$ be two linear maps with corresponding matrices A and B respectively.

Then the composition

$S \circ T: \mathbb{R}^n \to \mathbb{R}^k$ is a map satisfying

$$S \circ T(au + bv)$$

$$= S\big(T(au+bv)\big) = S\big(aT(u) + bT(v)\big)$$

$$= aS\big(T(u)\big) + bS\big(T(v)\big)$$

$$= a(S \circ T)(u) + b(S \circ T)(v),$$

for all $a, b \in \mathbb{R}$, $u, v \in \mathbb{R}^n$.

Thus, the composition $S \circ T$ is a linear map from $\mathbb{R}^n$ to $\mathbb{R}^k$.

What is the matrix of $S \circ T$?

Recall that $e_1^n = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, e_n^n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$

span $\mathbb{R}^n$.

Now, by definition of A, we've:

$$T(e_i^n) = A e_i^n = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{pmatrix} = a_{1i} e_1^m + \cdots + a_{mi} e_m^m.$$

Again, $(S \circ T)(e_i) = S(T(e_i))$

$$= S(a_{1i} e_1^m + \cdots + a_{mi} e_m^m)$$

$$= a_{1i} S(e_1^m) + \cdots + a_{mi} S(e_m^m)$$

$$= a_{1i}\, B\begin{pmatrix}1\\0\\ \vdots\\0\end{pmatrix} + \cdots + a_{mi}\, B\begin{pmatrix}0\\0\\ \vdots\\0\\1\end{pmatrix} \quad \left[\text{By definition of } B\right]$$

$$= a_{1i}\begin{pmatrix}b_{11}\\ \vdots\\ b_{K1}\end{pmatrix} + \cdots + a_{mi}\begin{pmatrix}b_{1m}\\ \vdots\\ b_{Km}\end{pmatrix}$$

$$= \begin{pmatrix} a_{1i} b_{11} + \cdots + a_{mi} b_{1m} \\ \cdots \\ a_{1i} b_{K1} + \cdots + a_{mi} b_{Km}\end{pmatrix}$$

$$= i\text{-th column of } \begin{pmatrix} b_{11} & \cdots & b_{1m} \\ \cdots & & \cdots \\ b_{K1} & \cdots & b_{Km}\end{pmatrix}\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & & \cdots \\ a_{m1} & \cdots & a_{mn}\end{pmatrix}$$

$$\Rightarrow (S \circ T)(e_i^n) = i\text{-th column of } BA$$

$$\Rightarrow (S \circ T)(e_i^n) = (BA)(e_i^n) = (BA)\begin{pmatrix}0\\ \vdots\\ 1\\ \vdots\\ 0\end{pmatrix} \to i\text{-th row}$$

Thus, the matrix of $S \circ T : \mathbb{R}^n \to \mathbb{R}^k$ is the $k \times n$ matrix BA.

Hence, the Composition of two linear maps is given by the product of the corresponding matrices.

<u>Note:</u> This is the real reason why matrices are multiplied the way they are.

<u>Definition:</u>

- Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be linear.

Then, $\text{Image}(T) := \{Tu : u \in \mathbb{R}^n\}$.

- $\text{Domain}(T) = \mathbb{R}^n$.
- T is called onto iff

$$\text{Image}(T) = \mathbb{R}^m.$$

- T is called one-to-one iff

$$T(u) = T(v) \Rightarrow u = v.$$

Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be one-to-one and onto. Such a map is called bijective. In this case, there exists an inverse

$T^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ of T. In fact,

$$T \circ T^{-1} = T^{-1} \circ T = Id.$$

Let $T(u) = u'$, $T(v) = v'$.

Then, by definition, $T^{-1}(u') = u$,
$T^{-1}(v') = v$.

Also, by linearity of T,

$$T(au + bv) = aT(u) + bT(v) = au' + bv'.$$

Therefore,

$$T^{-1}(au' + bv')$$
$$= au + bv = aT^{-1}(u') + bT^{-1}(v'),$$
for all $a, b \in \mathbb{R}$, $u', v' \in \mathbb{R}^n$

Hence, $\underline{T^{-1} : \mathbb{R}^n \to \mathbb{R}^n \text{ is also linear.}}$

Let B be the matrix of A.

Since, $T \circ T^{-1} = T^{-1} \circ T = Id$, ~~we have observed~~ and the matrices of $T \circ T^{-1}$, $T^{-1} \circ T$ and Id are AB, BA and Id_n respectively, we've:

$$\underline{AB = BA = Id_n}, \text{ where } Id_n \text{ is the } n \times n \text{ identity matrix.}$$

Thus, the matrix of T^{-1} is $B = A^{-1}$.

☒

Vector spaces and linear maps on them

Definition:

A real/complex vector space V is a set V together with operations

$$+ : V \times V \to V \quad , \quad \cdot : \mathbb{R} \times V \to V$$
$$(\text{or } \mathbb{C} \times V \to V)$$

Satisfying

i) $u+v \in V$, whenever $u, v \in V$

ii) $u+v = v+u$, for all $u, v \in V$

iii) $u+(v+w) = (u+v)+w$, $\forall u, v, w \in V$

iv) $\exists$ a unique element $0 \in V$ such that $0+u = u+0 = u$ $\forall u \in V$.

v) For each $u \in V$, there exists $-u \in V$ with $u+(-u) = (-u)+u = 0$

vi) $c \cdot u \in V$ $\forall c \in \mathbb{R}$ (or $\mathbb{C}$), $\forall u \in V$.

vii) $c \cdot (u+v) = c \cdot u + c \cdot v$ $\forall c \in \mathbb{R}$ (or $\mathbb{C}$), $\forall u, v \in V$.

viii) $(c+d) \cdot u = c \cdot u + d \cdot u$ $\forall c, d \in \mathbb{R}$ (or $\mathbb{C}$), $\forall u \in V$.

ix) $1 \cdot u = u$ $\forall u \in V$,

x) $(cd) \cdot u = c \cdot (d \cdot u)$, $\forall c, d \in \mathbb{R}$ (or $\mathbb{C}$), $\forall u \in V$.

Example: 1) $\mathbb{R}^n$ (or $\mathbb{C}^n$) over $\mathbb{R}$ (or $\mathbb{C}$),

2) All real $m \times n$ matrices over $\mathbb{R}$,

3) All continuous/differentiable functions $f: \mathbb{R} \to \mathbb{R}$ (over $\mathbb{R}$),

4) For any integer $d \geq 1$, the set $P_d = \{a_0 + a_1 x + \cdots + a_d x^d : a_i \in \mathbb{R}\}$ of polynomials of degree at most d. (over $\mathbb{R}$).

Consequences of definition:

1) $0 \cdot u = 0 \quad \forall u \in V$,

2) $c \cdot 0 = 0 \quad \forall c \in \mathbb{R}$ (or $\mathbb{C}$).

3) $c.u = 0 \Rightarrow c = 0$ or $u = 0$

Defn (linear combination):

A vector $u \in V$ is called a linear combination of $u_1, \dots, u_n \in V$ if there exist $c_1, \dots, c_n \in \mathbb{R}$ (or $\mathbb{C}$) such that:

$$u = c_1 u_1 + \dots + c_n u_n.$$
$$= \sum_{i=1}^{n} c_i u_i$$

Defn (subspace):

A subset $W \subseteq V$ is called a (vector) subspace if every linear combination of elements of W lies in W; i.e.

for any $c_1, c_2 \in \mathbb{R}$ (or $\mathbb{C}$) and for any $w_1, w_2 \in W$, we've that $(c_1 w_1 + c_2 w_2) \in W$.

Examples:

1) Any ~~plane~~ hyperplane

$$c_1x_1 + c_2x_2 + \cdots + c_nx_n = 0 \quad \text{in } \mathbb{R}^n.$$

2) The set of all diagonal matrices

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \quad \text{in } \mathrm{Mat}_n(\mathbb{R}).$$

3) All differentiable functions $f : \mathbb{R} \to \mathbb{R}$ with $f'(0) = 0$.

Non-example:

The set $\{(x, y, z) \in \mathbb{R}^3 : z = x^2 + y^2\}$ is not a vector subspace of $\mathbb{R}^3$.

Span:

Let V_1 be a subset of V. The span of V_1 is defined as the set of all possible linear combinations of elements of V_1. In notations,

$$\mathrm{span}(V_1) = \{c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n : c_i \in \mathbb{R}, \alpha_i \in V_1\}$$

Linear independence:

A collection of vectors $\{v_1, \ldots, v_n\}$ in V is called linearly independent if there is no non-trivial ~~ge~~ linear relation between ~~them~~ them; i.e. for some $c_1, \ldots, c_n \in \mathbb{R}$,

if $c_1 v_1 + c_2 v_2 + \cdots + c_n v_n = 0$,

then $c_1 = c_2 = \cdots = c_n = 0$.

Otherwise, the collection of vectors is called linearly dependent.

Remark:

1) A linearly independent set of vectors cannot contain the 0 vector.

2) If $\{v_1, \ldots, v_n\}$ is a linearly dependent set of non-zero vectors, then by definition,

there exist $c_1, \ldots, c_n \in \mathbb{R}$ such that at least one of the c_is is non-zero and

$$c_1 v_1 + \cdots + c_n v_n = 0$$

suppose $c_i \neq 0$, for some $i \in \{1, \ldots, n\}$.

Hence, $c_i v_i = -c_1 v_1 - \cdots - c_{i-1} v_{i-1} - c_{i+1} v_{i+1} - \cdots - c_n v_n$

$$\Rightarrow v_i = -\frac{c_1}{c_i} v_1 - \dots - \frac{c_{i-1}}{c_i} v_{i-1} - \frac{c_{i+1}}{c_i} v_{i+1} - \dots - \frac{c_n}{c_i} v_n.$$

$$\Rightarrow v_i = d_1 v_1 + \dots + d_{i-1} v_{i-1} + d_{i+1} v_{i+1} + \dots + d_n v_n.$$

Also, by our assumption, $v_i \neq 0$.

Therefore, not every d_j is equal to 0.

This shows that in a linearly dependent set of non-zero vectors, there is at least one vector that can be expressed as a linear combination of the others. This justifies the term "dependent"; in the ~~ex~~ worked out example, v_i "depends" on the other vectors non-trivially.

Defn (Basis):

Let β be a subset of a vector space V. β is said to be a basis of V if

i) β is linearly independent,

and ii) $\text{span}(\beta) = V$; if β spans V.

Example: 1) $\{e_1 = (1,0,\ldots,0), e_2 = (0,1,0,\ldots,0), \ldots, e_n = (0,0,\ldots,0,1)\}$

is a basis of $\mathbb{R}^n$.

2) Any two linearly independent vectors, e.g. $\{(1,0),(1,1)\}$ is a basis of $\mathbb{R}^2$. Similarly, any collection of n linearly independent vectors is a basis of $\mathbb{R}^n$.

3) A Basis of $M_2(\mathbb{R}) =$

$$\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

4) A basis for the space of real polynomials of degree at most d:

$\{1, t, t^2, \ldots, t^d\}$.

Coordinates:

Let V be a vector space and $\beta = \{u_1, u_2, \ldots, u_n\}$ be a basis of V. By definition, β is linearly independent and $\mathrm{span}(\beta) = V$.

Choose any $\alpha \in V = \mathrm{span}(\beta)$. Then there exist scalars $c_1, c_2, \ldots, c_n$ (in $\mathbb{R}$ or $\mathbb{C}$) such that

$$\alpha = c_1 u_1 + \cdots + c_n u_n \to (1)$$

We claim that the scalars $\{c_1, \ldots, c_n\}$ are unique. Suppose that $\{d_1, \ldots, d_n\}$ be another set of scalars such that

$$\alpha = d_1 u_1 + \cdots + d_n u_n \to (2)$$

By (1) and (2), we've:

$$c_1 u_1 + \cdots + c_n u_n = d_1 u_1 + \cdots + d_n u_n$$

$$\Rightarrow (c_1 - d_1) u_1 + \cdots + (c_n - d_n) u_n = 0 \to (3)$$

Since $\{u_1, \ldots, u_n\}$ is linearly independent, we must have $c_1 - d_1 = \cdots = c_n - d_n = 0$

$$\Rightarrow c_1 = d_1, \ c_2 = d_2, \ldots, c_n = d_n.$$

Defn: The unique n-tuple $(c_1, \ldots, c_n)$ is called the coordinates of α w.r.t. the basis β.

Dimension:

Let V be a vector space and β be a finite basis of V. In other words, β is a finite set

i) $\text{Span}(\beta) = V$, and
ii) β is linearly independent.

Such a vector space V is called **finite dimensional**.

Theorem: Any two bases of a finite dimensional vector space V have the same cardinality (i.e. the same number of elements).

Defn. Let V be a finite dimensional vector space and β be a basis of V. We define

$$\dim(V) = \text{Cardinality of } \beta$$

The above number is called the dimension of V. In light of the above theorem, $\dim(V)$ doesnot depend on the choice of a basis.

Example: 1) $\dim(\mathbb{R}^n) = n$.

2) $\dim(\mathrm{Mat}_n(\mathbb{R})) = n^2$.

3) $\dim(P_d) = d+1$.

4) The space of all real polynomials (of any degree) is also a vector space. However it has no finite basis. A basis of this vector space is given by:

$$\{1, t, t^2, \ldots, t^n, \ldots\}.$$

Such a space is called infinite dimensional.

5) The space of continuous/differentiable functions $f : \mathbb{R} \to \mathbb{R}$ is another example of an infinite dimensional vector space.

Theorem:

Let V be an n-dimensional vector space. Then

i) any subset of V containing more than n elements is linearly dependent;

ii) no subset of V with fewer than n elements can span V.

Linear maps of vector spaces:

Let V, W be vector spaces. A map $T: V \to W$ is called linear if

$$T(au+bv) = aT(u) + bT(v), \quad \forall a, b \in \mathbb{R}, \ \forall u, v \in V.$$

Examples: 1) $T: \mathbb{R}^2 \to \mathbb{R}^2$

$$T(x,y) = (x+y, x-y).$$

2) $T: P_d \to P_{d-1}$

$$T(f) = f'$$

3) $T : Mat_n(\mathbb{R}) \to \mathbb{R}$

$$T(A) = \text{trace}(A) = \underbrace{a_{11} + a_{22} + \cdots + a_{nn}}_{\text{Sum of the diagonal elements of } A}$$

4) Let $V = \{ f : \mathbb{R} \to \mathbb{R} : f \text{ is continuous} \}$

$T : V \to V$ is defined as:

$(Tf)(x) = \int_0^x f(t)\,dt$, for any $f \in V$.

Consequence of definition:

1) $T(0) = 0$

2) $T(c_1 u_1 + \cdots + c_n u_n) = c_1 T(u_1) + \cdots + c_n T(u_n)$

Theorem (easy consequence of property (2)):

Let V be a finite dimensional vector space and $\{u_1, u_2, \ldots, u_n\}$ be a basis of V. Let W be another vector space and $\{w_1, \ldots, w_n\}$ be any n vectors in W. Then there exists a unique linear map

$$T : V \to W \quad \text{satisfying}$$

$$T(u_k) = w_k, \quad \text{for } k = 1, \ldots, n.$$

Proof: We need to check that T can be extended as a linear map to all of V.
To this end, pick any $\alpha \in V$.
Then (by our discussion of coordinates), there exist unique scalars $c_1, \ldots, c_n$ such that

$$\alpha = c_1 u_1 + \cdots + c_n u_n.$$

We define: $T(\alpha)$

$$\begin{pmatrix}\text{extended to} \\ \text{guarantee} \\ \text{linearity}\end{pmatrix} := c_1 T(u_1) + \cdots + c_n T(u_n)$$

$$= c_1 w_1 + \cdots + c_n w_n.$$

Since the scalars $\{c_1, \dots, c_n\}$ are unique, this defines T in a well-defined way on all of V.

Linearity of T is now an easy exercise. The fact that such an extension is unique follows from the fact that any linear map

$S : V \to W$ satisfying $S(u_k) = w_k$, (for $k = 1, \dots, n$)

must satisfy the property

$$S(\alpha) = S(c_1 u_1 + \dots + c_n u_n) = c_1 S(u_1) + \dots + c_n S(u_n)$$

$$= c_1 w_1 + \dots + c_n w_n.$$

Hence, any such S must be equal to T. ⊗

Remark: Although the proof of the previous theorem is elementary, it turns out to be quite an important result. In fact, this gives us an easy way to define a linear map with very little data.

Image and Null-space:

Let $T: V \to W$ be a linear map.

- $T(V) = \{T(u) : u \in V\} \to$ Image of T

We'll show that $T(V)$ is a subspace of W.

To do so, choose scalars $a, b \in \mathbb{R}$ (or $\mathbb{C}$) and $w_1, w_2 \in T(V)$

Then there exist $u_1, u_2 \in V$ such that

$$w_1 = T(u_1) \quad \text{and} \quad w_2 = T(u_2).$$

Then, $aw_1 + bw_2$

$$= aT(u_1) + bT(u_2)$$

$$= T(au_1 + bu_2) \in T(V).$$

Therefore, $aw_1 + bw_2 \in T(V)$.

This shows that $T(V)$ is a subspace of W.

- T is called onto/surjective if

$$\mathrm{Im}(T) := T(V) = W.$$

- The null space of T, denoted by $\text{Null}(T)$, is defined as:

$$\text{Null}(T) := \{u \in V : T(u) = 0\}.$$

If $u_1, u_2 \in \text{Null}(T)$ and $a, b \in \mathbb{R}$, then

$$\begin{aligned} T(au_1 + bu_2) &= aT(u_1) + bT(u_2) \\ &= a.0 + b.0 \\ &= 0. \end{aligned}$$

$\Rightarrow au_1 + bu_2 \in \text{Null}(T)$.

Hence, $\text{Null}(T)$ is a subspace of V.

- $\text{Image}(T)$ and $\text{Null}(T)$ are the two most important/fundamental subspaces associated with a linear map T. In fact, these subspaces contain substantial information about the map T.

- Since $T(0) = 0$, we always have $0 \in \text{Null}(T)$.

Theorem:

T is one-to-one iff $Null(T)=\{0\}$.

Proof: $\Rightarrow$

Let us assume that T is one-to-one.

If $u \in Null(T)$, then we've

$$T(u)=0$$

But $T(0)=0$ for any linear map.

Hence, $T(u) = T(0)$

Since T is one-to-one, we conclude that $u=0$. So, 0 is the only element of $Null(T)$.

Thus, $Null(T) = \{0\}$.

$\Leftarrow$ Conversely, let us assume that $Null(T)=\{0\}$.

Let $u_1, u_2 \in V$ such that $T(u_1) = T(u_2)$

Then, $T(u_1) - T(u_2) = 0$

$\Rightarrow T(u_1 - u_2) = 0$

$\Rightarrow u_1 - u_2 \in Null(T) = \{0\}$

$\Rightarrow u_1 - u_2 = 0 \Rightarrow u_1 = u_2$. Hence, T is one-to-one.

Matrix of a linear map:

Let V and W be finite dimensional vector spaces with bases

$\beta_V = \{u_1, \dots, u_n\}$ and

$\beta_W = \{w_1, \dots, w_m\}$ respectively.

For any $i \in \{1, \dots, n\}$, we have $T(u_i) \in W$. Hence, there exist unique scalars $\{a_{1i}, a_{2i}, \dots, a_{mi}\}$ such that

$$T(u_i) = a_{1i} w_1 + a_{2i} w_2 + \dots + a_{mi} w_m$$

(Since $\{w_1, \dots, w_m\}$ is a basis of W)

We define the matrix of T with respect to the bases β_V and β_W as:

$$[T]_{\beta_V, \beta_W} := \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1i} & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2i} & a_{2n} \\ - & - & - & & - \\ a_{m1} & a_{m2} & \dots & a_{mi} & \dots a_{mn} \end{pmatrix}$$

- T is formed by the coordinate column vectors of $T(u_1), \dots, T(u_n)$.

(Here and in the next page, V, W, Y are finite dimensional spaces)

• <u>Composition:</u>

Let $T: V \to W$, $S: W \to Y$ be linear maps. Let A be the matrix of T with respect to the bases β_V and β_W, and

B be the matrix of S wrt the bases β_W and β_Y.

Then, the matrix of

$S \circ T: V \to Y$ with respect to the bases β_V and β_Y is given

by BA.

The proof of this fact is similar to the corresponding proof in the setting of Euclidean spaces (see Jan. 30 lecture notes).

Inverse:

Let $T: V \to V$ be a one-to-one and surjective linear map.
Let $\dim(V) = n$, and $\beta = \{u_1, \dots, u_n\}$ be a basis of V.

Further suppose that $[T]_{\beta,\beta} = \underset{\substack{\uparrow \\ \text{The matrix of} \\ T \text{ wrt } \beta}}{A}$

Since T is bijective, there exists a linear map $T^{-1}: V \to V$ such that

$$T \circ T^{-1} = T^{-1} \circ T = Id_V.$$

If B is the matrix of T^{-1} wrt β, then by our discussion ~~of~~ on the previous page, we've:

$$AB = BA = Id_n$$

Hence, $B = A^{-1}$, the inverse of the matrix A.

Therefore, the matrix of a bijective linear map $T: V \to V$ is invertible, and the matrix of the inverse map T^{-1} is given by the inverse of the matrix of T (wrt a fixed basis).

Coordinates:

Let V be a vector space, and $\beta = \{u_1, u_2, \ldots, u_n\}$ be a basis of V. By definition, β is linearly independent and $\text{span}(\beta) = V$.

Choose any $\alpha \in V = \text{span}(\beta)$. Then there exist scalars $c_1, c_2, \ldots, c_n$ (in $\mathbb{R}$ or $\mathbb{C}$) such that

$$\alpha = c_1 u_1 + \cdots + c_n u_n . \rightarrow (1)$$

We claim that the scalars $\{c_1, \ldots, c_n\}$ are unique. Suppose that $\{d_1, \ldots, d_n\}$ be another set of scalars such that

$$\alpha = d_1 u_1 + \cdots + d_n u_n \rightarrow (2)$$

By (1) and (2), we've:

$$c_1 u_1 + \cdots + c_n u_n = d_1 u_1 + \cdots + d_n u_n$$

$$\Rightarrow (c_1 - d_1) u_1 + \cdots + (c_n - d_n) u_n = 0 \rightarrow (3)$$

Since $\{u_1, \ldots, u_n\}$ is linearly independent, we must have $c_1 - d_1 = \cdots = c_n - d_n = 0$

$$\Rightarrow c_1 = d_1,\ c_2 = d_2, \ldots, c_n = d_n .$$

Defn: The unique n-tuple $(c_1, \ldots, c_n)$ is called the coordinates of α w.r.t. the basis β.

Example:

1) P_3 is the vector space of all real polynomials of degree ≤ 3. We know that $\beta = \{1, x, x^2, x^3\}$ is an ordered basis of P_3 (ordered in the sense that we've arranged the elements of the basis in some order).

Then, the coordinates of the vector $2 - 3x + 4x^2 - 5x^3 \in P_3$ with respect to the ordered basis β is:

$$\begin{pmatrix} 2 \\ -3 \\ 4 \\ -5 \end{pmatrix} \begin{array}{l} \rightarrow \text{coefficient of } 1 \\ \rightarrow \text{------- } x \\ \rightarrow \text{------- } x^2 \\ \rightarrow \text{------- } x^3 \end{array}$$

2) $\beta = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$ is an ordered basis of $M_2(\mathbb{R})$.

Then any matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{R})$ can be written as:

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

Hence, the coordinates of $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with respect the basis β is $\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}$.

Dimension:

Let V be a vector space and β be a finite basis of V. In other words, β is a finite set

i) $\text{Span}(\beta) = V$, and
ii) β is linearly independent.

Such a vector space V is called finite dimensional.

Theorem: Any two bases of a finite dimensional vector space V have the same cardinality (i.e. the same number of elements).

Defn. Let V be a finite dimensional vector space and β be a basis of V. We define

$$\dim(V) = \text{Cardinality of } \beta$$

The above number is called the dimension of V. In light of the above theorem, $\dim(V)$ doesnot depend on the choice of a basis.

Example: 1) $\dim(\mathbb{R}^n) = n$.

2) $\dim(\mathrm{Mat}_n(\mathbb{R})) = n^2$.

3) $\dim(P_d) = d+1$.

4) The space of all real polynomials (of any degree) is also a vector space. However it has no finite basis. A basis of this vector space is given by:

$$\{1, t, t^2, \ldots, t^n, \ldots\}.$$

Such a space is called infinite dimensional.

5) The space of continuous/differentiable functions $f: \mathbb{R} \to \mathbb{R}$ is another example of an infinite dimensional vector space.

Theorem:

Let V be an n-dimensional vector space. Then

i) any subset of V containing more than n elements is linearly dependent;

ii) no subset of V with fewer than n elements can span V.

Theorem: Let V be a finite dimensional vector space with $\dim(V) = n$.

Let $\{u_1, \ldots, u_k\}$ be a set of linearly independent vectors in V, with $k < n$.

Then we can choose $(n-k)$ linearly independent vectors $\{u_{k+1}, \ldots, u_n\}$ such that $\{u_1, \ldots, u_k, u_{k+1}, \ldots, u_n\}$ is a basis of V.

Linear maps of vector spaces:

Let V, W be vector spaces. A map $T: V \to W$ is called linear if

$$T(au + bv) = aT(u) + bT(v), \quad \forall a, b \in \mathbb{R}, \ \forall u, v \in V.$$

Examples: 1) $T: \mathbb{R}^2 \to \mathbb{R}^2$

$$T(x,y) = (x+y, x-y).$$

2) $T: P_d \to P_{d-1}$

$$T(f) = f'$$

3) $T: Mat_n(\mathbb{R}) \to \mathbb{R}$

$$T(A) = \text{trace}(A) = \underbrace{a_{11} + a_{22} + \cdots + a_{nn}}$$

Sum of the diagonal elements of A

4) Let $V = \{ f : \mathbb{R} \to \mathbb{R} : f \text{ is continuous} \}$

$T : V \to V$ is defined as:

$$(Tf)(x) = \int_0^x f(t)\,dt \;, \text{ for any } f \in V.$$

Consequence of definition:

1) $T(0) = 0$

2) $T(c_1 u_1 + \cdots + c_n u_n) = c_1 T(u_1) + \cdots + c_n T(u_n)$

Image and Null-space :

Let $T: V \to W$ be a linear map.

- $T(V) = \{T(u) : u \in V\} \to$ Image of T

We'll show that $T(V)$ is a subspace of W.

To do so, choose scalars $a, b \in \mathbb{R}$ (or $\mathbb{C}$) and $w_1, w_2 \in T(V)$.

Then there exist $u_1, u_2 \in V$ such that

$$\underline{w_1 = T(u_1)} \text{ and } \underline{w_2 = T(u_2)}.$$

Then, $aw_1 + bw_2$

$$= aT(u_1) + bT(u_2)$$

$$= T(au_1 + bu_2) \in T(V).$$

Therefore, $aw_1 + bw_2 \in T(V)$.

This shows that $T(V)$ is a subspace of W.

- T is called onto/surjective if

$$\text{Range}(T) = \text{Im}(T) := T(V) = W.$$

Remark: We'll denote the image (or range) of T by $\text{Range}(T)$ (or $\text{Im}(T)$).

Thm. Let V, W be finite dimensional vector spaces and

$\beta = \{u_1, \ldots, u_n\}$ be a basis of V.

Let $T : V \to W$ be a linear map.

Then,

$$\text{range}(T) = \text{span}\{T(u_1), T(u_2), \ldots, T(u_n)\}.$$

Proof: Let $u \in V$.

Then there exist unique scalars $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ such that

$$u = \alpha_1 u_1 + \cdots + \alpha_n u_n$$

$$\Rightarrow T(u) = T(\alpha_1 u_1 + \cdots + \alpha_n u_n)$$

$$\Rightarrow T(u) = \alpha_1 T(u_1) + \cdots + \alpha_n T(u_n)$$

$$\Rightarrow T(u) \in \text{span}\{T(u_1), \ldots, T(u_n)\}$$ for every $u \in V$.

$$\Rightarrow \text{Range}(T) \subseteq \text{span}\{T(u_1), \ldots, T(u_n)\}.$$

The opposite containment is trivial.

Thus, $\text{Range}(T) = \text{span}\{T(u_1), \ldots, T(u_n)\}$.

Remark: T is completely determined by its action on a basis of V.

Theorem (easy consequence of property (2)):

Let V be a finite dimensional vector space and $\{u_1, u_2, \ldots, u_n\}$ be a basis of V. Let W be another vector space and $\{w_1, \ldots, w_n\}$ be any n vectors in W. Then there exists a unique linear map

$$T : V \to W \quad \text{satisfying}$$

$$T(u_k) = w_k, \quad \text{for } k = 1, \ldots, n.$$

Proof: We need to check that T can be extended as a linear map to all of V.
To this end, pick any $\alpha \in V$.
Then (by our discussion of coordinates), there exist unique scalars $c_1, \ldots, c_n$ such that

$$\alpha = c_1 u_1 + \cdots + c_n u_n.$$

We define: $T(\alpha)$

(extended to guarantee linearity)

$$:= c_1 T(u_1) + \cdots + c_n T(u_n)$$

$$= c_1 w_1 + \cdots + c_n w_n.$$

Since the scalars $\{c_1, \ldots, c_n\}$ are unique, this defines T in a well-defined way on all of V.

Linearity of T is now an easy exercise. The fact that such an extension is unique follows from the fact that any linear map

$S : V \to W$ satisfying $S(u_k) = w_k$, (for $k = 1, \ldots, n$)

must satisfy the property

$$S(\alpha) = S(c_1 u_1 + \cdots + c_n u_n) = c_1 S(u_1) + \cdots + c_n S(u_n)$$
$$= c_1 w_1 + \cdots + c_n w_n.$$

Hence, any such S must be equal to T. ⊗

Remark: Although the proof of the previous theorem is elementary, it turns out to be quite an important result. In fact, this gives us an easy way to define a linear map with very little data.

- The null space of T, denoted by Null (T), is defined as:

$$\text{Null}(T) := \{u \in V : T(u) = 0\}.$$

If $u_1, u_2 \in \text{Null}(T)$ and $a, b \in \mathbb{R}$, then

$$\begin{aligned} T(au_1 + bu_2) &= aT(u_1) + bT(u_2) \\ &= a \cdot 0 + b \cdot 0 \\ &= 0. \end{aligned}$$

$\Rightarrow au_1 + bu_2 \in \text{Null}(T).$

Hence, Null (T) is a subspace of V.

- Image (T) and Null (T) are the two most important/fundamental subspaces associated with a linear map T. In fact, these subspaces contain substantial information about the map T.

- Since $T(0) = 0$, we always have $0 \in \text{Null}(T)$.

Theorem:

T is one-to-one iff $\text{Null}(T) = \{0\}$.

Proof: $\Rightarrow$ Let us assume that T is one-to-one.

If $u \in \text{Null}(T)$, then we've

$$T(u) = 0$$

But $T(0) = 0$ for any linear map.

Hence, $T(u) = T(0)$

Since T is one-to-one, we conclude that $u = 0$. So, 0 is the only element of $\text{Null}(T)$.

Thus, $\text{Null}(T) = \{0\}$.

$\Leftarrow$ Conversely, let us assume that $\text{Null}(T) = \{0\}$.

Let $u_1, u_2 \in V$ such that $T(u_1) = T(u_2)$

Then, $T(u_1) - T(u_2) = 0$

$\Rightarrow T(u_1 - u_2) = 0$

$\Rightarrow u_1 - u_2 \in \text{Null}(T) = \{0\}$

$\Rightarrow u_1 - u_2 = 0 \Rightarrow u_1 = u_2$. Hence, T is one-to-one.

Theorem: Let $T: V \to W$ be a linear map. Let $u_0 \in V$ be a solution of the non-homogenous equation

① $\to$ $T(u) = \tilde{w}$, where $\tilde{w}$ is some element of W.

Then the set of all solutions of ① is given by:

$$\{u_0 + u : u \in Null(T)\}.$$

Proof: Let, S be the set of all solutions of ①; i.e.

$$S = \{X \in V : T(X) = \tilde{w}\}.$$

For any $u \in Null(T)$, we have:

$$T(u_0 + u) = T(u_0) + T(u)$$
$$= \tilde{w} + 0 \quad \left[\text{Since } T(u_0) = w, \text{ and } T(u) = 0\right]$$
$$= \tilde{w}$$

$\Rightarrow T(u_0 + u) = \tilde{w}$, for each $u \in Null(T)$

$\Rightarrow u_0 + u \in S$, — — — — — —

$\Rightarrow \{u_0 + u : u \in Null(T)\} \subseteq S. \to$ ②

Converssely, let $x \in S$.

Then, $T(x) = \tilde{w}$

$\Rightarrow T(x) = T(u_0)$ [as $T(u_0) = \tilde{w}$]

$\Rightarrow T(x) - T(u_0) = 0$

$\Rightarrow T(x - u_0) = 0$

$\Rightarrow x - u_0 \in \text{Null}(T)$

$\Rightarrow x - u_0 = u$, for some $u \in \text{Null}(T)$

$\Rightarrow x = u_0 + u$, for some $u \in \text{Null}(T)$

Hence, $x \in \{u_0 + u : u \in \text{Null}(T)\}$

Therefore, $S \subseteq \{u_0 + u : u \in \text{Null}(T)\}$ $\to$ (3)

By (2) and (3), we've that:

$$S = \{u_0 + u : u \in \text{Null}(T)\}.$$

An application:

V = The vector space of all real differentiable functions.

W = The vector space of all real Continuous functions

Then, $T := \frac{d}{dx} : V \to W$ is a linear map.

Fix $g \in W$, and consider the equation

$$T(f) = g \to (*)$$

or $\frac{df}{dx} = g$.

Note that $Null(T)$

$= \{ f \in V : T(f) = 0 \}$

$= \{ f \in V : \frac{df}{dx} \equiv 0 \}$

$= \{ f \in V : f$ is a constant$\}$

$=$ Set of all constant functions

$= \{ c : c \in \mathbb{R} \}$

Now let f_0 be a solution of $(*)$;

i.e, $\frac{df_0}{dx} = g$.

We ~~usut~~ usually call f_0 an anti-derivative/integral of g.

By the Previous theorem, the set of all solutions of (*); i.e. the set of all integrals of g are given by:

$$\{ f_0 + f : f \in Null(T) \}$$

$$= \{ f_0 + c : c \in \mathbb{R} \}.$$

This justifies our old habit of adding constants in indefinite integration.

Matrix of a linear map:

Let V and W be finite dimensional vector spaces with bases

$\beta_V = \{u_1, \dots, u_n\}$ and

$\beta_W = \{w_1, \dots, w_m\}$ respectively.

For any $i \in \{1, \dots, n\}$, we have $T(u_i) \in W$. Hence, there exist unique scalars $\{a_{1i}, a_{2i}, \dots, a_{mi}\}$ such that

$$T(u_i) = a_{1i} w_1 + a_{2i} w_2 + \dots + a_{mi} w_m$$

(Since $\{w_1, \dots, w_m\}$ is a basis of W)

We define the matrix of T with respect to the bases β_V and β_W as:

$$[T]_{\beta_V, \beta_W} := \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1i} & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2i} & a_{2n} \\ - & - & - & - & - \\ a_{m1} & a_{m2} & \dots & a_{mi} & a_{mn} \end{pmatrix}$$

- T is formed by the coordinate column vectors of $T(u_1), \dots, T(u_n)$.

(Here and in the next page, V, W, Y are finite dimensional spaces)

• Composition:

Let $T: V \to W$, $S: W \to Y$ be linear maps. Let A be the matrix of T with respect to the bases β_V and β_W, and

B be the matrix of S wrt the bases β_W and β_Y.

Then, the matrix of

$S \circ T: V \to Y$ with respect to the bases β_V and β_Y is given

by BA.

The proof of this fact is similar to the corresponding proof in the setting of Euclidean spaces (see Jan. 30 lecture notes).

Inverse:

Let $T: V \to V$ be a one-to-one and surjective linear map.
Let $\dim(V) = n$, and $\beta = \{u_1, \dots, u_n\}$ be a basis of V.

Further suppose that $[T]_{\beta,\beta} = A$

The matrix of T wrt β

Since T is ~~a~~ bijective, there exists a linear map $T^{-1}: V \to V$ such that

$$T \circ T^{-1} = T^{-1} \circ T = Id_V.$$

If B is the matrix of T^{-1} wrt β, then by our discussion ~~of~~ on the previous page, we've:

$$AB = BA = Id_n$$

Hence, $B = A^{-1}$, the inverse of the matrix A.

Therefore, the matrix of a bijective linear map $T: V \to V$ is invertible, and the matrix of the inverse map T^{-1} is given by the inverse of the matrix of T (wrt a fixed basis).

Matrix of a linear map:

Let V and W be finite dimensional vector spaces with bases

$\beta_V = \{u_1, \ldots, u_n\}$ and

$\beta_W = \{w_1, \ldots, w_m\}$ respectively.

For any $i \in \{1, \ldots, n\}$, we have $T(u_i) \in W$. Hence, there exist unique scalars $\{a_{1i}, a_{2i}, \ldots, a_{mi}\}$ such that

$$T(u_i) = a_{1i} w_1 + a_{2i} w_2 + \cdots + a_{mi} w_m$$

(Since $\{w_1, \ldots, w_m\}$ is a basis of W)

We define the matrix of T with respect to the bases β_V and β_W as:

$$[T]_{\beta_V, \beta_W} := \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1i} & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2i} & a_{2n} \\ - & - & - & & - \\ a_{m1} & a_{m2} & \cdots & a_{mi} & a_{mn} \end{pmatrix}$$

- T is formed by the coordinate column vectors of $T(u_1), \ldots, T(u_n)$.

(Here and in the next page, V, W, Y are finite dimensional spaces)

• Composition:

Let $T: V \to W$, $S: W \to Y$ be linear maps. Let A be the matrix of T with respect to the bases β_V and β_W, and

B be the matrix of S wrt the bases β_W and β_Y.

Then, the matrix of

$S \circ T : V \to Y$ with respect to the bases β_V and β_Y is given

by BA.

The proof of this fact is similar to the corresponding proof in the setting of Euclidean spaces (see Jan. 30 lecture notes).

Inverse:

Let $T: V \to V$ be a one-to-one and surjective linear map.
Let $\dim(V) = n$, and $\beta = \{u_1, \ldots, u_n\}$ be a basis of V.

Further suppose that $[T]_{\beta,\beta} = \underset{\substack{\uparrow \\ \text{The matrix of} \\ T \text{ wrt } \beta}}{A}$

Since T is ~~a~~ bijective, there exists a linear map $T^{-1}: V \to V$ such that

$$T \circ T^{-1} = T^{-1} \circ T = Id_V.$$

If B is the matrix of T^{-1} wrt β, then by our discussion ~~of~~ on the previous page, we've:

$$AB = BA = Id_n$$

Hence, $B = A^{-1}$, the inverse of the matrix A.

Therefore, the matrix of a bijective linear map $T: V \to V$ is invertible, and the matrix of the inverse map T^{-1} is given by the inverse of the matrix of T (wrt a fixed basis).

We'll now see that if $[T]_\beta$ is the matrix of a linear map $T: V \to V$ w.r.t. the basis β (of V), then the action of T on any vector in V is given by multiplication by $[T]_\beta$.

Let $u \in V$. Since $\beta = \{u_1, \ldots, u_n\}$ is a basis of V, there exist unique scalars $c_1, c_2, \ldots, c_n \in \mathbb{R}$ such that

$$u = c_1 u_1 + \cdots + c_n u_n;$$

i.e. $\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$ are the coordinates of u w.r.t to β.

Suppose that

$$[T]_\beta = \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2j} & \cdots & a_{2n} \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{pmatrix}$$

Then,
$$T(u) = \sum_{j=1}^{n} c_j T(u_j)$$
$$= \sum_{j=1}^{n} c_j \left(\sum_{i=1}^{n} a_{ij} u_i \right)$$
$$= \sum_{i=1}^{n} \sum_{j=1}^{n} c_j a_{ij} u_i$$

$$= \sum_{i=1}^{n} \left(\sum_{j=1}^{n} c_j \, a_{ij} \right) u_i$$

Hence, the coordinates of $T(u)$ w.r.t. β is

$$\begin{pmatrix} \sum_{j=1}^{n} c_j \, a_{1j} \\ \sum_{j=1}^{n} c_j \, a_{2j} \\ \vdots \\ \sum_{j=1}^{n} c_j \, a_{nj} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

$$= [T]_\beta \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

To sum up, we've showed that:

The coordinates of $T(u)$ wrt β $= [T]_\beta \cdot$ (The coordinate column vector of u wrt β)

Hence, the matrix of T (wrt some basis β) contains complete information about T.

However, an arbitrary $n \times n$ matrix may look rather unwieldy and may not give away any geometric information readily. ~~The good news is that~~

The question that we ought to ask at this point is that:

$$\boxed{\text{Is there a better/best choice of a basis } \beta \text{ w.r.t. which the matrix } [T]_\beta \text{ looks simple?}}$$

Example: suppose that $T: \mathbb{R}^2 \to \mathbb{R}^2$ is a linear map whose matrix wrt the standard basis $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ is given by $A = \begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix}$. Evidently, this matrix doesn't say much about the map T, at least you cannot just stare at the matrix to figure what it does to vectors.

Now let's consider the basis

$\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$. [These two vectors are linearly independent. Since $\dim(\mathbb{R}^2) = 2$, they form a basis.

A simple computation shows that:

$$T\begin{pmatrix} 1 \\ 1 \end{pmatrix} = T\left(1\begin{pmatrix} 1 \\ 0 \end{pmatrix} + 1\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right)$$
$$= 1.T\begin{pmatrix} 1 \\ 0 \end{pmatrix} + 1.T\begin{pmatrix} 0 \\ 1 \end{pmatrix}$$
$$= \begin{pmatrix} 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 5 \\ 5 \end{pmatrix} = 5\begin{pmatrix} 1 \\ 1 \end{pmatrix} + 0.\begin{pmatrix} 1 \\ -1 \end{pmatrix},$$

and

$$T\begin{pmatrix} 1 \\ -1 \end{pmatrix} = T\left(1.\begin{pmatrix} 1 \\ 0 \end{pmatrix} - 1.\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right)$$
$$= 1.T\begin{pmatrix} 1 \\ 0 \end{pmatrix} - 1.T\begin{pmatrix} 0 \\ 1 \end{pmatrix}$$
$$= \begin{pmatrix} 2 \\ 3 \end{pmatrix} - \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} = 0.\begin{pmatrix} 1 \\ 1 \end{pmatrix} + (-1).\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Hence, the matrix of T wrt the basis $\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$ is: $B = \begin{pmatrix} 5 & 0 \\ 0 & -1 \end{pmatrix}$.

Wow! A diagonal matrix is much simpler to work with, both algebraically and geometrically. Algebraically, by our discussion on compositions of linear maps, we see that the matrix of

$$T^{\circ n} = \underbrace{T \circ T \circ \cdots \circ T}_{n\text{-times}}$$

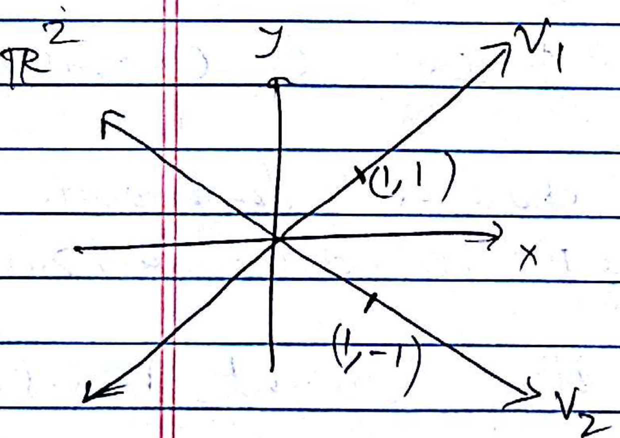
w.r.t $\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$ is given by

$$B^n = \underbrace{B.B \cdots B}_{n\text{-fold product}}$$

$$= \begin{pmatrix} 5^n & 0 \\ 0 & (-1)^n \end{pmatrix}.$$

(multiplying diagonal matrices is much simpler than mult. arbitrary matrices)

Geometrically, T is given by an expansion (by a factor 5) in the direction of V_1 and reflection w.r.t. the origin on the straight line spanned by V_2.



Therefore, having represented the ~~matrix~~ linear map T by a diagonal matrix wrt some basis, we've done ourselves a big geometric and algebraic favor!

Definition (Diagonalizability):

A linear map $T: V \to V$ is called diagonalizable if there exists a basis $\beta = \{u_1, \ldots, u_n\}$ of V w.r.t. which the matrix $[T]_\beta$ of T is a diagonal matrix.

You are perhaps wondering how we arrived at the "good choice of basis" $\left\{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\-1\end{pmatrix}\right\}$ in the previous example. More generally, how do we find out whether a linear map is diagonalizable? Or how do we find the "right" basis wrt which it has a diagonal matrix?

In the rest of these notes, we'll delve into this topic and arrive at somewhat satisfactory answers to these questions.

Definition (Eigenvectors/Eigenvalues):

Let $T: V \to V$ be a linear map. Let $\lambda \in \mathbb{R}$ (or $\mathbb{C}$) and $u \in V$ ($u \neq 0$) be such that:

$$Tu = \lambda u.$$

Then λ is called an eigenvalue of T, and u is an eigenvector of T associated with the eigenvalue λ.

Remark: In the previous example, the eigenvalues of T were 5 and -1. An eigenvector of T associated with 5 (respectively -1) was $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ (respectively $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$).

Theorem: (Diagonalizable $\Leftrightarrow$ basis of eigenvectors)

$T: V \to V$ is diagonalizable if there exists a basis of V consisting of eigenvectors of T.

Proof: $\Rightarrow$ Let's assume that T is diagonalizable. Then there exists a basis $\{u_1, \dots, u_n\} = \beta$ of V s.t. $[T]_\beta$ is a diagonal matrix. Hence,

$$[T]_\beta = \begin{pmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \ddots \\ & & & \lambda_n \end{pmatrix}, \text{ for some } \lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R} \text{ (or } \mathbb{C}).$$

By definition of $[T]_\beta$, this means that

$$\begin{cases} T(u_1) = \lambda_1 u_1 + 0 \cdot u_2 + \cdots + 0 \cdot u_n \\ T(u_2) = 0 \cdot u_1 + \lambda_2 \cdot u_2 + \cdots + 0 \cdot u_n \\ \quad \cdots \\ T(u_j) = 0 \cdot u_1 + \cdots + 0 \cdot u_{j-1} + \lambda_j \cdot u_j + 0 \cdot u_{j+1} + \cdots + 0 \cdot u_n \\ \quad \cdots \\ T(u_n) = 0 \cdot u_1 + \cdots + \lambda_n \cdot u_n \end{cases}$$

$\Rightarrow T(u_1) = \lambda_1 u_1,\ T(u_2) = \lambda_2 u_2, \dots, T(u_j) = \lambda_j u_j, \dots, T(u_n) = \lambda_n u_n.$

clearly, $u_j \neq 0$, for $j = 1, \dots, n$.
Therefore, each u_j (for $j = 1, \dots, n$) is an eigenvector of T associated with the eigenvalue λ_j.

$\Rightarrow \{u_1, \dots, u_n\}$ is a basis of V consisting of eigenvectors of T.

$\Leftarrow$ We now assume that V admits a basis $\beta = \{u_1, \dots, u_n\}$ such that each u_j is an eigenvector of T.

Then, for each $j = 1, \dots, n$, there exists a scalar λ_j (in $\mathbb{R}$ or $\mathbb{C}$) such that

$$T(u_j) = \lambda_j u_j.$$

But then,

$$\begin{cases} T(u_1) = \lambda_1 . u_1 + 0 . u_2 + \cdots + 0 . u_n \\ T(u_2) = 0 . u_1 + \lambda_2 . u_2 + \cdots + 0 . u_n \\ \cdots \\ T(u_n) = 0 . u_1 + 0 . u_2 + \cdots + \lambda_n . u_n \end{cases}$$

By definition, the matrix of T wrt β is given by:

$$[T]_\beta = \begin{pmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ & & \ddots \\ 0 & & \lambda_n \end{pmatrix}.$$

Thus, β is a basis of V w.r.t. which the matrix of T is a diagonal matrix.

$\Rightarrow$ T is diagonalizable. $\boxtimes$

All that sounds good in theory, but how do we find eigenvalues/vectors in practice?

Suppose that $\beta' = \{w_1, \ldots, w_n\}$ is a basis of V, and $T: V \to V$ is a linear map such that

$$[T]_{\beta'} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} =: A.$$

Further suppose that u is an eigenvector of T associated with the eigenvalue λ.

Let $u = c_1 w_1 + c_2 w_2 + \cdots + c_n w_n$, so the coordinates of u w.r.t. β' is

$$\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}.$$

By our assumption, we've:

$T(u) = \lambda u$. Hence, $T(u)$ and λu have the same coordinates wrt β'. The coordinates of $T(u)$ w.r.t. β' are given by:

$$A \cdot \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}, \quad \text{and}$$

the coordinates of λu w.r.t. β' are given by:

$$\begin{pmatrix} \lambda c_1 \\ \lambda c_2 \\ \vdots \\ \lambda c_n \end{pmatrix}.$$

Hence,
$$A \cdot \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} \lambda c_1 \\ \lambda c_2 \\ \vdots \\ \lambda c_n \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} a_{11} & a_{12} & - & a_{1n} \\ a_{21} & a_{22} & - & a_{2n} \\ a_{n1} & a_{n2} & - - - & a_{nn} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} \lambda c_1 \\ \lambda c_2 \\ \vdots \\ \lambda c_n \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} a_{11}c_1 + a_{12}c_2 + \cdots + a_{1n}c_n - \lambda c_1 \\ a_{21}c_1 + a_{22}c_2 + \cdots + a_{2n}c_n - \lambda c_2 \\ \text{---} \\ a_{n1}c_1 + a_{n2}c_2 + \cdots + a_{nn}c_n - \lambda c_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} (a_{11}-\lambda)c_1 + a_{12}c_2 + \cdots + a_{1n}c_n \\ a_{21}c_1 + (a_{22}-\lambda)c_2 + \cdots + a_{2n}c_n \\ \text{---} \\ a_{n1}c_1 + a_{n2}c_2 + \cdots + (a_{nn}-\lambda)c_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \text{---} & \text{---} & & \text{---} \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$\Rightarrow \left(A - \lambda \, Id_n\right) \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \rightarrow (*)$$

(here, Id_n is the identity matrix of size n. Henceforth, we'll denote it simply by I.)

Since u is an eigenvector, u is a non-zero vector. Hence $\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$ is not the zero column vector. Therefore, (*) implies that $\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$ is a non-trivial solution of the equation

(**) $\rightarrow$ $(A - \lambda I)X = 0$, where X is a "column vector variable", and 0 is the zero column vector.

We know that (**) admits a non-trivial solution if and only if

$$\det(A - \lambda I) = 0 \quad (\text{i.e. } (A-\lambda I) \text{ is singular}).$$

Hence, we've proved that:

Theorem: (Finding eigenvalues):
λ is an eigenvalue of T if and only if

$\det(A - \lambda I) = 0$, where A is the matrix of T w.r.t. some basis.

We'll use this theorem to find eigenvalues of a linear map.

Example: Let's go back to our previous example. We consider $T: \mathbb{R}^2 \to \mathbb{R}^2$ whose matrix w.r.t. $\left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\}$ is $\begin{pmatrix}2 & 3\\3 & 2\end{pmatrix} = A$.

To find all the eigenvalues of T, we solve the equation

$$\det(A - \lambda I) = 0$$

$$\Rightarrow \left| \begin{pmatrix}2 & 3\\3 & 2\end{pmatrix} - \lambda \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix} \right| = 0$$

$$\Rightarrow \begin{vmatrix}2-\lambda & 3\\3 & 2-\lambda\end{vmatrix} = 0$$

$$\Rightarrow (2-\lambda)^2 - 9 = 0 \Rightarrow \lambda^2 - 4\lambda + 4 - 9 = 0$$

$$\Rightarrow \lambda^2 - 4\lambda - 5 = 0 \Rightarrow (\lambda - 5)(\lambda + 1) = 0$$

$$\Rightarrow \underline{\lambda = 5, -1}.$$

Hence, the eigenvalues of T are $\{5, -1\}$.

~~To find an eigenvector of T associated with the eigenvalue 5, we solve the system:~~

$$\cancel{(A-\lambda I)\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}}$$

Let $\begin{pmatrix} x \\ y \end{pmatrix}$ be an eigenvector of T associated with the eigenvalue 5. Then,

$$T\begin{pmatrix} x \\ y \end{pmatrix} = 5\begin{pmatrix} x \\ y \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 5x \\ 5y \end{pmatrix}$$

$$\Rightarrow \left.\begin{array}{l} 2x+3y=5x \\ 3x+2y=5y \end{array}\right\}$$

$$\Rightarrow 3y = 3x$$

$$\Rightarrow y = x.$$

Thus, $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ x \end{pmatrix} = x\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

So, any eigenvector of T associated with the eigenvalue 5 is of the form

$$x\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \text{ where } x \in \mathbb{R}.$$

In particular, we can choose $\begin{pmatrix}1\\1\end{pmatrix}$ to be an eigenvector of T associated with the eigenvalue 5.

Now let $\begin{pmatrix}p\\q\end{pmatrix}$ be an eigenvector of T associated with the eigenvalue -1.

Then, $T\begin{pmatrix}p\\q\end{pmatrix} = -1\begin{pmatrix}p\\q\end{pmatrix}$

$$\Rightarrow \begin{pmatrix}2 & 3\\3 & 2\end{pmatrix}\begin{pmatrix}p\\q\end{pmatrix} = \begin{pmatrix}-p\\-q\end{pmatrix}$$

$$\Rightarrow 2p + 3q = -p, \quad 3p + 2q = -q$$

$$\Rightarrow 3q = -3p$$

$$\Rightarrow \underline{q = -p.}$$

Hence, $\begin{pmatrix}p\\q\end{pmatrix} = \begin{pmatrix}p\\-p\end{pmatrix} = p\begin{pmatrix}1\\-1\end{pmatrix}$.

Thus, any eigenvector of T associated with the eigenvalue -1 is of the form: $p\begin{pmatrix}1\\-1\end{pmatrix}$, where $p \in \mathbb{R}$.

In particular, we can choose $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ to be an eigenvector of T associated with the eigenvalue -1.

Therefore, we've found two distinct eigenvectors:

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ and } \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Since they are linearly independent, they form a basis of $\mathbb{R}^2$.

Hence, there exists a basis of $\mathbb{R}^2$, namely $\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$ consisting of eigenvectors of T. Therefore, T is diagonalizable.

Confession: The basis on the top of page 7 didn't pop out of thin air! I had done these computations already (but did not tell you then). But now you know how to arrive at a "good choice of basis".

We just saw how to diagonalize a linear map (when possible). Since this is an extremely important tool, let us record the method as a routine algorithm.

Diagonalizing a linear map:

Let A be the matrix of a linear map $T: V \to V$ wrt some basis β of V.

1) To find eigenvalues of T:

Solve the equation

3* $\rightarrow$ $\det(A - \lambda I) = 0$, which is

a polynomial of degree $n = \dim V$ in λ. List the roots of as $\{\lambda_1, \dots, \lambda_n\}$. These are the eigenvalues of T.

2) Finding the eigenvectors:

For each λ_i, $i = 1, \dots, n$, find the associated eigenvectors by solving the system of equations:

4*: $Ax = \lambda_i x$.

The solutions of 4* are the eigenvectors of T associated with the eigenvalue λ_i.

3) Testing diagonalizability:

Having found all the eigenvalues of eigenvectors T, we need to check whether there is a basis of V consisting of eigenvectors of T. If there exists such a basis $\{u_1, \ldots, u_n\}$, then the matrix of T w.r.t. $\{u_1, \ldots, u_n\}$ will be diagonal.

Step 1 and Step 2 above are rather straightforward. Since a basis of V is simply a linearly independent spanning set of V, step 3 amounts to checking whether there are enough linearly independent eigenvectors (of T) to span all of V.

The next theorem is the first step in analyzing the situation.

Thm. (Eigenvectors of distinct eigenvalues are linearly indep)

Let $u_1, \ldots, u_K$ be eigenvectors of T associated with eigenvalues $\lambda_1, \ldots, \lambda_K$ such that $\lambda_i \neq \lambda_j$ if $i \neq j$ (i.e. all λ_is are distinct).

Then $\{u_1, \ldots, u_K\}$ is a linearly independent set of vectors.

Proof: By assumption, we have that

$$T(u_i) = \lambda_i u_i, \quad i = 1, \ldots, k.$$

Suppose that there exist scalars $c_1, \ldots, c_K$ such that

$$c_1 u_1 + c_2 u_2 + \cdots + c_n u_n = 0 \rightarrow (1)$$

Applying T to both sides of (1), we get:

$$T(c_1 u_1 + c_2 u_2 + \cdots + c_n u_n) = T(0)$$

$$\Rightarrow c_1 \lambda_1 u_1 + c_2 \lambda_2 u_2 + \cdots + c_n \lambda_n u_n = 0 \rightarrow (2)$$

Now, $\lambda_n \times (1) - (2)$ yields:

$$c_1(\lambda_n-\lambda_1)u + c_2(\lambda_n-\lambda_2) + \cdots + c_{n-1}(\lambda_n-\lambda_{n-1})u_{n-1} = 0 \quad \to (3)$$

We now use mathematical induction to complete the proof. Since every eigenvector u_1 is non-zero, $\{u_1\}$ is a linearly independent set. Let us suppose that the set $\{u_1, \ldots, u_{k-1}\}$ is linearly independent.

Then (3) implies that

$$(4) \to \quad c_1(\lambda_n-\lambda_1) = c_2(\lambda_n-\lambda_2) = \cdots = c_{n-1}(\lambda_n-\lambda_{n-1}) = 0.$$

Since all λ_is are distinct, we have that
$\lambda_n-\lambda_1 \neq 0$, $\lambda_n-\lambda_2 \neq 0$, $\ldots$, $\lambda_n-\lambda_{n-1} \neq 0$.

Therefore, $c_1 = c_2 = \cdots = c_{n-1} = 0$ (by (4))

Plugging these in (1), we get:

$$c_n u_n = 0$$
$$\Rightarrow c_n = 0.$$ (As u_n is an eigenvector, we've: $u_n \neq 0$)

Therefore, $c_1 = c_2 = \cdots = c_{n-1} = c_n = 0$.

Thus, we've proved that if $\{u_1, \ldots, u_{k-1}\}$ is a linearly independent set, then $\{u_1, \ldots, u_k\}$ is also a linearly independent set. Hence by math. induction, we conclude that

$\{u_1, \ldots, u_k\}$ is a linearly independent set for any k, whenever they correspond to distinct eigenvalues.

$\boxtimes$

In light of the previous theorem, diagonalizability of T boils down to the existence of sufficiently many eigenvectors of T (so that they can span V).

A relatively mild hypothesis on T can now guarantee diagonalizability of T.

<u>Thm.</u> (n distinct eigenvalues $\Rightarrow$ diagonalizable)

Let V be an n-dim. vector space and $T: V \to V$ be a linear map. If T has n distinct eigenvalues, then T is diagonalizable.

Proof: Let $\{\lambda_1, \dots, \lambda_n\}$ be n distinct eigenvalues of T.

Remark: Recall that the eigenvalues of T are the solutions of the degree n polynomial: $\det(A - \lambda I) = 0$. A polynomial of degree n has at most n roots; in fact, exactly n roots in $\mathbb{C}$. The hypothesis of the theorem is that all these roots are distinct.

Furthermore, let $u_i (\neq 0)$ be an eigenvector of T associated with λ_i. Since all the λ_is are distinct, the previous theorem ascertains that the set of eigenvectors

$\{u_1, \dots, u_n\}$ is linearly independent.

However, as $\dim(V) = n$, any linearly independent set of n vectors is a basis of V. Therefore, $\{u_1, \dots, u_n\}$ is a basis of V consisting of eigenvectors of T.

$\Rightarrow$ T is diagonalizable.

Examples:

1) In the previous example, where $T: \mathbb{R}^2 \to \mathbb{R}^2$ is given by $\begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix}$ w.r.t the standard basis $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$, we found two distinct eigenvalues of T; namely $\{5, -1\}$. Since $\dim(\mathbb{R}^2) = 2$, the previous theorem implies that T is diagonalizable.

2) Now we'll see that $T: V \to V$ can be diagonalizable even if it doesn't have n distinct eigenvalues.

Let $T: \mathbb{R}^3 \to \mathbb{R}^3$ be given by the matrix

$$A = \begin{pmatrix} 2 & -2 & 14 \\ 0 & 3 & -7 \\ 0 & 0 & 2 \end{pmatrix}$$

w.r.t. the standard basis $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$.

To find the eigenvalues of T, we solve the equation:

$$\det(A - \lambda I) = 0$$

$$\Rightarrow \begin{vmatrix} 2-\lambda & -2 & 14 \\ 0 & 3-\lambda & -7 \\ 0 & 0 & 2-\lambda \end{vmatrix} = 0$$

$$\Rightarrow (2-\lambda)(3-\lambda)(2-\lambda) = 0$$

$$\Rightarrow \lambda = 2, 2, 3.$$

Therefore, the eigenvalues of T are $\{2,3\}$. Since T does not have 3 distinct eigenvalues (note that $\dim(\mathbb{R}^3)=3$), we cannot apply the previous theorem to directly conclude that T is diagonalizable. Hence, we now proceed to find all the eigenvectors of T.

Eigenvectors associated with 2:

Let $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$ be an eigenvector of T associated with the eigenvalue 2. Then,

$$T\begin{pmatrix} a \\ b \\ c \end{pmatrix} = 2\begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} 2a-2b+14c \\ 3b-7c \\ 2c \end{pmatrix} = \begin{pmatrix} 2a \\ 2b \\ 2c \end{pmatrix}$$

$$\Rightarrow 2a-2b+14c = 2a, \quad 3b-7c = 2b, \quad 2c = 2c$$

$\Rightarrow 2b = 14c, \quad b = 7c$

$\Rightarrow \underline{b = 7c} \rightarrow$ This is the only constraint

Hence, any eigenvector of T associated with 2 is of the form:

$$\begin{pmatrix} a \\ 7c \\ c \end{pmatrix} = \begin{pmatrix} a \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 7c \\ c \end{pmatrix} = a\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + c\begin{pmatrix} 0 \\ 7 \\ 1 \end{pmatrix},$$

where a, b are any real numbers.

~~So we can choose~~

So we observe that the space of all eigenvectors of T associated with 2 is spanned by two linearly vectors: $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 7 \\ 1 \end{pmatrix}$; the <u>eigenspace</u> of 2 is 2-dimensional. In particular, $\left\{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 7 \\ 1 \end{pmatrix}\right\}$ is a linearly independent set of eigenvectors of T associated with 2.

Eigenvectors of T associated with 3:

Let $\begin{pmatrix} p \\ q \\ r \end{pmatrix}$ be an eigenvector of T associated with the eigenvalue 3.

Hence,
$$T\begin{pmatrix} p \\ q \\ r \end{pmatrix} = 3\begin{pmatrix} p \\ q \\ r \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} 2p-2q+14r \\ 3q-7r \\ 2r \end{pmatrix} = \begin{pmatrix} 3p \\ 3q \\ 3r \end{pmatrix}$$

$$\Rightarrow \left.\begin{array}{r} 2p-2q+14r=3p \\ 3q-7r=3q \\ 2r=3r \end{array}\right\}$$

Thus, $2r=3r \Rightarrow \underline{r=0}$

Putting this in $3q-7r=3q$, we get:
$3q=3q$, which is a trivial condition.

Putting $r=0$ in $2p-2q+14r=3p$,
we get: $2p-2q=3p$
$$\Rightarrow \underline{p=-2q}$$

So, any eigenvector of T associated with 3 is of the form:

$$\begin{pmatrix} -2q \\ q \\ 0 \end{pmatrix} = q\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \text{ where } q \in \mathbb{R}.$$

In this case, the eigenspace of T associated by 3 is 1-dimensional, and it's spanned by $\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}$.

So, there is only one linearly independent eigenvector $\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}$ of T associated with 3.

We now consider the linearly independent set of eigenvectors

$$\beta := \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 7 \\ 1 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} \right\}.$$

(We just listed the linearly independent eigenvectors of 2 and 3.)

Clearly, three linearly independent vectors in $\mathbb{R}^3$ form a basis.

Since,
$$T\begin{pmatrix}1\\0\\0\end{pmatrix}=2\begin{pmatrix}1\\0\\0\end{pmatrix},\quad T\begin{pmatrix}0\\7\\1\end{pmatrix}=2\begin{pmatrix}0\\7\\1\end{pmatrix},$$
→ they're eigenvectors of 2

$$T\begin{pmatrix}-2\\1\\0\end{pmatrix}=3\begin{pmatrix}-2\\1\\0\end{pmatrix},$$
→ this is an eigenvector of 3

the matrix of T w.r.t. β is given by
$$\begin{pmatrix}2&0&0\\0&2&0\\0&0&3\end{pmatrix}.$$

Hence, T is diagonalizable even though it doesn't have 3 distinct eigenvalues.

A moment's reflection (or perhaps a few minutes') will convince you that this was possible because we found 3 linearly independent eigenvectors of T;

i.e. dim (eigenspace of T associated with 2)

+ dim (eigenspace of T associated with 3)

$$= 3 = \dim(\mathbb{R}^3).$$

3) Finally, let's look at a linear map that is not diagonalizable.

Let $T: \mathbb{R}^3 \to \mathbb{R}^3$ be a linear map given by

$$A = \begin{pmatrix} 2 & 2 & 1 \\ 0 & 2 & -1 \\ 0 & 0 & 3 \end{pmatrix}$$ wrt

the standard basis $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$

The eigenvalues of T are solutions of:

$$\det(A - \lambda I) = 0$$

$$\Rightarrow \begin{vmatrix} 2-\lambda & 2 & 1 \\ 0 & 2-\lambda & -1 \\ 0 & 0 & 3-\lambda \end{vmatrix} = 0 \Rightarrow (2-\lambda)(2-\lambda)(3-\lambda) = 0$$

$$\Rightarrow \underline{\lambda = 2, 2, 3}.$$

Hence, the eigenvalues of T are $\{2, 3\}$.

Eigenvectors of T associated with 2:

Let $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$ be an eigenvector of T associated with the eigenvalue 2. Then

$$T\begin{pmatrix} a \\ b \\ c \end{pmatrix} = 2\begin{pmatrix} a \\ b \\ c \end{pmatrix} \Rightarrow \begin{pmatrix} 2a+2b+c \\ 2b-c \\ 3c \end{pmatrix} = \begin{pmatrix} 2a \\ 2b \\ 2c \end{pmatrix}$$

$$\Rightarrow 2a+2b+c = 2a,\ 2b-c = 2b,\ 3c = 2c$$

$$\Rightarrow 2b+c = 0,\ c = 0$$

$$\Rightarrow \underline{b = 0,\ c = 0.}$$

So, any eigenvector of T associated with 2 is of the form:

$$\begin{pmatrix} a \\ 0 \\ 0 \end{pmatrix} = a\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},\quad a \in \mathbb{R}.$$

Thus, the eigenspace of T associated with the eigenvalue 2 is spanned by the single vector $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$. Hence, it's one-dimensional.

In particular, a maximal linearly independent set of eigenvectors of T associated with 2 is

$$\left\{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right\}.$$

Eigenvectors of T associated with 3:

Let $\begin{pmatrix} p \\ q \\ r \end{pmatrix}$ be an eigenvector of T associated with the eigenvalue 3.

Then,

$$T\begin{pmatrix} p \\ q \\ r \end{pmatrix} = 3\begin{pmatrix} p \\ q \\ r \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} 2p+2q+r \\ 2q-r \\ 3r \end{pmatrix} = \begin{pmatrix} 3p \\ 3q \\ 3r \end{pmatrix}$$

$$\Rightarrow 2p+2q+r = 3p, \quad 2q-r = 3q, \quad 3r = 3r$$

$$\Rightarrow p = 2q+r, \quad q = -r$$

$$\Rightarrow p = -2r+r$$

$$\Rightarrow p = -r.$$

Therefore, any ~~et~~ eigenvector of T associated with 3 is of the form:

$$\begin{pmatrix} -r \\ -r \\ r \end{pmatrix} = r\begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}, \text{ where } r \in \mathbb{R}.$$

So, the eigenspace of T associated with 3 is one-dimensional and is spanned by $\begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$. So, a basis of this eigenspace is $\left\{\begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}\right\}$.

As a result, ~~we only~~ there are only two linearly independent eigenvectors of T. So we cannot find a basis of $\mathbb{R}^3$ consisting of eigenvectors of T.

Thus, T is not diagonalizable.

Again, the reason why T is not diagonalizable can be summed up in the following inequality:

$$\dim\left(\text{Eigenspace of } T \text{ associated with the eigenvalue } 2\right)$$

$$+ \dim\left(\text{Eigenspace of } T \text{ associated with the eigenvalue } 3\right)$$

$$= 1 + 1 = 2 < 3 = \dim(\mathbb{R}^3).$$

We conclude our discussion on "whether T is diagonalizable" with the following theorem.

Theorem: Let V be a vector space with $\dim(V) = n$. Let $\{\lambda_1, \ldots, \lambda_K\}$ be all the distinct eigenvalues of a linear map $T: V \to V$.
Let $V_i := \{u \in V : T(u) = \lambda_i u\}$ be the eigenspace of T corresponding to the eigenvalue λ_i, where $i = 1, \ldots, K$.

Then, T is diagonalizable if and only if

$$\dim(V_1) + \dim(V_2) + \cdots + \dim(V_K) = n.$$

Finally, let us investigate the relation between an arbitrary matrix representation of a linear map (w.r.t some basis) and its diagonal matrix representation (w.r.t a basis of eigenvectors).

Let $T: V \to V$ be a linear map, $\beta' = \{w_1, \ldots, w_n\}$ be some basis of V and A be the matrix of T w.r.t. β'.

Furthermore, let $\beta = \{u_1, \ldots, u_n\}$ be a basis of V consisting of eigenvectors of T; i.e.

$$T(u_j) = \lambda_j u_j, \quad j = 1, \ldots, n, \quad \lambda_j \in \mathbb{R}.$$

Since β' is a basis, each u_j has a coordinate vector ~~repr~~ representation w.r.t. β'. Suppose that the coordinates of u_j w.r.t. the basis β' be $\begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{pmatrix}$.

Define $$B := \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1j} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2j} & \cdots & b_{2n} \\ - & - & - & - & - & - \\ b_{n1} & b_{n2} & \cdots & b_{nj} & \cdots & b_{nn} \end{pmatrix}$$

[The j-th column of B is the coordinate column vector of u_j w.r.t. β'.]

Finally, let $$D = \begin{pmatrix} \lambda_1 & & & O \\ & \lambda_2 & & \\ & & \ddots & \\ O & & & \lambda_n \end{pmatrix}$$

Then a straightforward (but tedious) computation shows that

$$AB = BD$$

$$\Longleftrightarrow \quad \underline{B^{-1}AB = D.}$$

<u>Example:</u> We return (or re-return) to $T: \mathbb{R}^2 \to \mathbb{R}^2$ which is given by the matrix $A = \begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix}$ w.r.t. to standard basis $\beta' = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ of $\mathbb{R}^2$.

A basis of $\mathbb{R}^2$ consisting of eigen-vectors of T is:

$$\beta = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$$

Then, $B = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$. [B represent elements of β w.r.t. β']

Finally, $D = \begin{pmatrix} 5 & 0 \\ 0 & -1 \end{pmatrix}$. [Recall that $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is an eigenvector associated with the eigenvalue 5, and $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ is an eigenvector asso. with eigenvalue -1.]

Then we have:

$$B^{-1} A B = D$$

i.e. $$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

$$= \begin{pmatrix} 5 & 0 \\ 0 & -1 \end{pmatrix}.$$

Inner Product Spaces

So far, we have been concerned with general vector spaces and linear maps on them, which generalize linear maps on $\mathbb{R}^n$ (or multiplication by a matrix).

In these notes, we will focus on a special property/structure of $\mathbb{R}^n$ that we're familiar with. Recall that the dot/scalar product of vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$ is defined as:

$$\vec{x} \cdot \vec{y} = x_1 y_1 + x_2 y_2 + x_3 y_3,$$

where
$$\vec{x} = x_1 \vec{i} + x_2 \vec{j} + x_3 \vec{k},$$
$$\vec{y} = y_1 \vec{i} + y_2 \vec{j} + y_3 \vec{k}.$$

We also know that two vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$ are perpendicular if and only if their dot product is zero,

and the angle θ between two vectors $\vec{x}$ and $\vec{y}$ is given by:

$$\theta = \cos^{-1}\left(\frac{\vec{x} \cdot \vec{y}}{\|\vec{x}\| \, \|\vec{y}\|}\right).$$

where $\|\vec{x}\| = \sqrt{\vec{x} \cdot \vec{x}}$, and $\|\vec{y}\| = \sqrt{\vec{y} \cdot \vec{y}}$.

The fundamental importance of dot product in $\mathbb{R}^2$ or $\mathbb{R}^3$ stems from the above facts; indeed, the dot product construction allows us to talk about angles in $\mathbb{R}^2$ or $\mathbb{R}^3$. So we can make sense of angles between vectors and orthogonality of vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$ using dot product.

As we have seen so far in this course, our goal is to develop a general theory of the special properties of $\mathbb{R}^n$. Our definition of vector spaces was in part motivated by the algebraic properties of $\mathbb{R}^n$. Our next definition will follow the same principle, it will generalize the notion of dot products to a much broader class of vector spaces, so that we can talk about angles and orthogonality in a much more general setting (meaning, beyond $\mathbb{R}^n$).

Here is the definition of an inner product on a vector space V (over $\mathbb{R}$).

<u>Defn:</u> An "inner product" is a function $\langle\ ,\ \rangle : V \times V \to \mathbb{R}$ satisfying

i) $\langle u, u\rangle \geq 0$ and equal to 0 iff $u = 0$,

ii) $\langle u, v\rangle = \langle v, u\rangle$,

iii) $\langle u+v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$,

iv) $\langle \alpha u, v\rangle = \alpha\langle u, v\rangle$,

where $u, v, w \in V$, and $\alpha \in \mathbb{R}$

A vector space V with an inner product as defined above is called an "inner product space".

Examples:

a) On $V = \mathbb{R}^n$, the standard "dot" product

$$\left\langle \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \right\rangle = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

defines an inner product (i.e. it satisfies the defining axioms of an inner product).

Just in case you forgot how to compute angles between vectors in $\mathbb{R}^n$, the following example should help.

Let $\vec{x} = (-2, 1, 3)$, $\vec{y} = (0, 1, -1) \in \mathbb{R}^3$

So, $\|\vec{x}\| = \sqrt{4 + 1^2 + 9} = \sqrt{14}$

$\|\vec{y}\| = \sqrt{0 + 1 + 1} = \sqrt{2}$.

$$\langle \vec{x}, \vec{y} \rangle = (-2)(0) + (1)(1) + (3)(-1)$$
$$= -2.$$

So, $\theta = \cos^{-1}\left(\dfrac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{x}\| \, \|\vec{y}\|} \right) = \cos^{-1}\left(\dfrac{-2}{\sqrt{2}\sqrt{14}} \right)$.

$$\Rightarrow \theta = \cos^{-1}\left(-\frac{\sqrt{2}}{14} \right).$$

b) Here's a definition of inner products on 'Some suitable' class of functions (at least integrable) on $[a,b]$:

$$\langle f,g\rangle = \int_a^b f(x)g(x)\,dx$$

i) $$\langle f,f\rangle = \int_a^b (f(x))^2\,dx$$

$$\geq 0 \,, \quad \text{as } (f(x))^2 \geq 0 \ \forall x \in [a,b]$$

Furthermore,

$$\langle f,f\rangle = 0 \Leftrightarrow \int_a^b (f(x))^2\,dx = 0$$

$$\Leftrightarrow (f(x))^2 = 0$$

(If the function is non-zero somewhere, then $(f(x))^2 > 0$, and the integral will be positive.)

$$\Leftrightarrow f(x) = 0 \text{ on } [a,b]$$

$$\Leftrightarrow f \equiv 0$$

So, $\langle f,f\rangle = 0 \Leftrightarrow f \equiv 0$.

ii) for any constant $\alpha \in \mathbb{R}$,

$$\langle \alpha f, g\rangle = \int_a^b (\alpha f(x))\,g(x)\,dx$$

$$\Rightarrow \langle \alpha f, g\rangle = \alpha \int_a^b f(x)g(x)\,dx$$

$$\Rightarrow \langle \alpha f, g\rangle = \alpha \langle f,g\rangle.$$

iii) $\langle f+g, h\rangle = \int_a^b \big(f(x)+g(x)\big)h(x)dx$

$\Rightarrow \langle f+g, h\rangle = \int_a^b f(x)h(x)dx + \int_a^b g(x)h(x)dx$

$\Rightarrow \langle f+g, h\rangle = \langle f, h\rangle + \langle g, h\rangle$.

iv) $\langle f, g\rangle = \int_a^b f(x)g(x)\,dx$

$\Rightarrow \langle f, g\rangle = \int_a^b g(x)f(x)dx$ [Multiplication of real numbers is Commutative.]

$\Rightarrow \langle f, g\rangle = \langle g, f\rangle$.

Therefore, the given product $\langle f, g\rangle = \int_a^b f(x)g(x)\,dx$ satisfies all the defining properties of inner products, and hence it's an inner product.

In particular, the vector space $C[a,b]$ of all continuous functions $f: [a,b] \to \mathbb{R}$ is an inner product space with respect to above inner product:

$$\langle f, g\rangle = \int_a^b f(x)g(x)dx.$$

Angles in inner product spaces

Suppose p and q are two vectors in an inner product space. We denote the inner product between p and q by $\langle p, q \rangle$, and the norms of p and q by $\|p\| = \sqrt{\langle p, p \rangle}$, $\|q\| = \sqrt{\langle q, q \rangle}$.

Then, the angle θ between p and q is defined as:

$$\theta = \cos^{-1}\left(\frac{\langle p, q \rangle}{\|p\| \, \|q\|}\right).$$

8

Let $f(x) = x^4$, $g(x) = x^2$ be elements of $C[-1, 1]$. We'll use the above "integral" inner product to compute the angle between f and g in $C[-1,1]$.

$$f(x) = x^4, \quad g(x) = x^2,$$

$$\langle f, g \rangle = \int_{-1}^{1} f(x) g(x)\, dx$$

$$= \int_{-1}^{1} x^6\, dx = \left[\frac{x^7}{7}\right]_{-1}^{1}$$

$$= \left(\frac{1}{7} + \frac{1}{7}\right) = \frac{2}{7}.$$

$$\|f\| = \sqrt{\int_{-1}^{1} f(x) f(x)\, dx} = \sqrt{\int_{-1}^{1} x^8\, dx}$$

$$= \sqrt{\left(\frac{x^9}{9}\right)_{-1}^{1}} = \sqrt{\frac{1}{9} + \frac{1}{9}} = \sqrt{\frac{2}{9}} = \frac{\sqrt{2}}{3}.$$

$$\|g\| = \sqrt{\int_{-1}^{1} g(x) g(x)\, dx} = \sqrt{\int_{-1}^{1} x^4\, dx} = \sqrt{\left(\frac{x^5}{5}\right)_{-1}^{1}}$$

$$= \sqrt{\frac{1}{5} + \frac{1}{5}} = \frac{\sqrt{2}}{\sqrt{5}}.$$

Hence, $\theta = \cos^{-1}\left(\dfrac{\langle f, g \rangle}{\|f\|\,\|g\|}\right)$

$$\Rightarrow \theta = \cos^{-1}\left(\frac{2/7}{\frac{\sqrt{2}}{3} \cdot \frac{\sqrt{2}}{\sqrt{5}}}\right)$$

$$\Rightarrow \theta = \cos^{-1}\left(\frac{\cancel{2}}{7} \cdot \frac{3\sqrt{5}}{\cancel{2}}\right)$$

$$\Rightarrow \theta = \cos^{-1}\left(\frac{3\sqrt{5}}{7}\right).$$

Defn: Two vectors u_1, u_2 in an inner product space V is called orthogonal if they satisfy $\langle u_1, u_2 \rangle = 0$

Examples: a) The standard basis elements $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$ of $\mathbb{R}^3$ are orthogonal to each other.

b) Let $C[-\pi, \pi]$ be the vector space of real-valued continuous functions on $[-\pi, \pi]$. Then,

$$\int_{-\pi}^{\pi} \sin x \cos x \, dx$$

$$= \frac{1}{2} \int_{-\pi}^{\pi} \sin 2x \, dx = 0 \quad \left[\text{Sin } 2x \text{ is an odd function} \right]$$

Thus, with respect to the inner product

$$\langle f, g \rangle = \int_{-\pi}^{\pi} f(x) \, g(x) \, dx,$$

the functions $\sin x$ and $\cos x$ are orthogonal to each other.

Let us now record a couple of generalizations of well-known results.

Cauchy–Schwarz inequality:

For $u_1, \dots, u_n \in \mathbb{R}$ and $w_1, \dots, w_n \in \mathbb{R}$, the classical C–S inequality states that:

$$|u_1 w_1 + u_2 w_2 + \cdots + u_n w_n|^2$$
$$\leq \left(|u_1|^2 + \cdots + |u_n|^2\right) \cdot \left(|w_1|^2 + \cdots + |w_n|^2\right)$$

i.e. $\langle \vec{u}, \vec{w} \rangle \leq \|\vec{u}\| \, \|\vec{w}\| \rightarrow (*)$

where, $\vec{u} = (u_1, \dots, u_n) \in \mathbb{R}^n$, and
$\vec{w} = (w_1, \dots, w_n) \in \mathbb{R}^n$.

An analogue of $(*)$ holds in arbitrary inner product spaces.

Theorem (Cauchy-Schwarz):

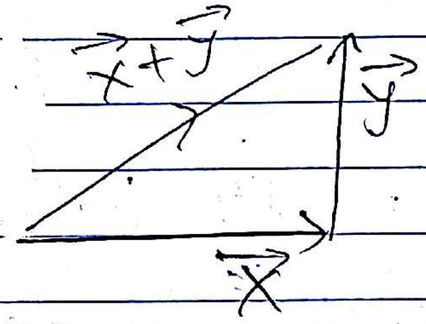
Let V be an inner product space, and $u_1, u_2 \in V$.

Then, $|\langle u_1, u_2 \rangle| \leq \|u_1\| . \|u_2\|$.

The proof can be found in the textbook or in several other places on the "internet".

Pythagorean theorem:

Let $\vec{x} = x_1\vec{i} + x_2\vec{j}$, $\vec{y} = y_1\vec{i} + y_2\vec{j}$ be two perpendicular vectors in $\mathbb{R}^2$.



Then $\vec{x}+\vec{y}$ represents the 'hypotenuse' as shown in the figure. The classical Pythagorean theorem states that

$$\left(l(\vec{x}+\vec{y})\right)^2 = \left(l(\vec{x})\right)^2 + \left(l(\vec{y})\right)^2,$$

where $l()$ denotes the length of a vector. In other words,

$$\left.\begin{array}{l} \|\vec{x}+\vec{y}\|^2 = \|\vec{x}\|^2 + \|\vec{y}\|^2 \\ \text{if } \langle \vec{x}, \vec{y} \rangle = 0 \text{ in } \mathbb{R}^2. \end{array}\right\} \rightarrow (**)$$

The following theorem generalizes (**) to arbitrary inner product spaces.

Theorem (Pythagoras).

Let V be an inner product space and u_1, u_2 be orthogonal in V. Then

$$\|u_1 + u_2\|^2 = \|u_1\|^2 + \|u_2\|^2.$$

Proof: $\|u_1 + u_2\|^2$

$= \langle u_1 + u_2, u_1 + u_2 \rangle$ (by definition)

(using linearity of $\langle , \rangle$) $= \langle u_1, u_1 \rangle + \langle u_2, u_1 \rangle + \langle u_1, u_2 \rangle + \langle u_2, u_2 \rangle$

$= \|u_1\|^2 + 0 + 0 + \|u_2\|^2$ (As u_1 and u_2 are orthogonal)

$= \|u_1\|^2 + \|u_2\|^2.$

Note that the standard basis $\left\{\begin{pmatrix}1\\0\\ \vdots\\0\end{pmatrix},\begin{pmatrix}0\\1\\ \vdots\\0\end{pmatrix},\cdots,\begin{pmatrix}0\\0\\ \vdots\\1\end{pmatrix}\right\}$ of $\mathbb{R}^n$ consists of n orthogonal elements.

Here's another example of a basis of orthogonal vectors: $x_1 = \begin{pmatrix}1\\0\\1\end{pmatrix}$, $x_2 = \begin{pmatrix}1\\0\\-1\end{pmatrix}$, $x_3 = \begin{pmatrix}0\\2\\0\end{pmatrix}$

To show that $\{x_1, x_2, x_3\}$ is an orthogonal set, we need to check that

$$\langle x_1, x_2\rangle = \langle x_1, x_3\rangle = \langle x_2, x_3\rangle = 0$$

Now, $\langle x_1, x_2\rangle = 1\times 1 + 0\times 0 + 1\times(-1)$

$$= 1 - 1 = 0.$$

$$\langle x_1, x_3\rangle = 1\times 0 + 0\times 2 + 0\times 1 = 0.$$

$$\langle x_2, x_3\rangle = 1\times 0 + 0\times 2 + (-1)\times 0$$

$$= 0.$$

Therefore, $\{x_1, x_2, x_3\}$ form an orthogonal set, i.e. they're orthogonal to each other.

Clearly they form a basis of $\mathbb{R}^3$.

The next theorem asserts that orthogonal vectors are always linearly independent.

<u>Theorem</u>:

An orthogonal set of vectors $\{x_1, x_2, \cdots, x_n\} \in V$ is linearly independent.

Proof of theorem—

Let, $c_1, c_2, \ldots, c_n$ be constants ($\in \mathbb{R}$)

Such that

$$c_1 x_1 + c_2 x_2 + \cdots + c_n x_n = 0 \rightarrow (1)$$

In order to prove independence of the vectors $\{x_1, x_2, \ldots, x_n\}$, we need to show that $\boxed{c_i = 0 \text{, for all } i = 1, \ldots, n}$

Taking inner product with the vector x_i, equation (1) yields:

$$\langle c_1 x_1 + c_2 x_2 + \cdots + c_n x_n, x_i \rangle = \langle 0, x_i \rangle$$

$$\Rightarrow \langle c_1 x_1, x_i \rangle + \langle c_2 x_2, x_i \rangle + \cdots + \langle c_i x_i, x_i \rangle + \cdots + \langle c_n x_n, x_i \rangle = 0$$

(inner product of x_i & the 0 vector $= 0$)

$$\Rightarrow c_1 \langle x_1, x_i \rangle + c_2 \langle x_2, x_i \rangle + \cdots + c_i \langle x_i, x_i \rangle + \cdots + c_n \langle x_n, x_i \rangle = 0 \rightarrow (2)$$

Now observe that $\langle x_i, x_i \rangle = \|x_i\|^2 \neq 0$

and

$$\langle x_j, x_i \rangle = 0 \quad \text{for } j \neq i$$

(by definition of orthogonality).

Thus (2) reduces to:

$$c_i \|x_i\|^2 = 0$$

$$\Rightarrow \quad c_i = 0$$ [Since x_i belongs to an orthogonal set of vectors, by definiton $x_i \neq 0$. So $\|x_i\|^2 \neq 0$]

Thus, we have ~~that~~ $c_i = 0$.

But our choice of 'i' was arbitrary.

Therefore, $c_i = 0$ for all $i = 1, 2, \ldots, n$.

$\Rightarrow \{x_1, x_2, \ldots, x_n\}$ form a linearly independent set of vectors.

Theorem: Let $\dim(V) = n$, and $\{x_1, \dots, x_n\}$ be an orthogonal set of n non-zero vectors in V. Then they form a basis of V, and for any $\alpha \in V$, we've:

$$\alpha = \sum_{i=1}^{n} \frac{\langle \alpha, x_i \rangle}{\langle x_i, x_i \rangle} x_i$$

Proof: We've already observed linear independence of $\{x_1, \dots, x_n\}$. The additional hypothesis $\dim(V) = n$ guarantees that it is a basis of V.

Hence, $\text{span}\{x_1, \dots, x_n\} = V$.

For any $\alpha \in V$, there exists unique scalars $c_1, \dots, c_n \in \mathbb{R}$ such that

$$\alpha = c_1 x_1 + \dots + c_i x_i + \dots + c_n x_n$$

$$\Rightarrow \langle \alpha, x_i \rangle = \langle c_1 x_1 + \dots + c_i x_i + \dots + c_n x_n, x_i \rangle$$

$$\Rightarrow \langle \alpha, x_i \rangle = c_i \langle x_i, x_i \rangle \quad \left[\text{Since } \langle x_i, x_j \rangle = 0, \text{ if } i \neq j \right]$$

$$\Rightarrow c_i = \frac{\langle \alpha, x_i \rangle}{\langle x_i, x_i \rangle} \quad \left[x_i \neq 0 \Rightarrow \langle x_i, x_i \rangle \neq 0 \right]$$

So, $$\alpha = \sum_{i=1}^{n} \frac{\langle \alpha, x_i \rangle}{\langle x_i, x_i \rangle} x_i .$$

⊠

Orthonormal set:

A set of vectors $\{x_1, \ldots, x_n\}$ in V is called orthonormal if

i) $\langle x_i, x_j \rangle = 0$ for $i \neq j$,

and ii) $\|x_i\| = 1$, for $i = 1, \ldots, n$.

Thus, an orthonormal set is an orthogonal set of unit vectors.

- For an orthonormal basis $\{x_1, \ldots, x_n\}$ of V (where $\dim(V) = n$) the formula in the previous theorem reduces to:

$$\alpha = \sum_{i=1}^{n} \langle \alpha, x_i \rangle \cdot x_i$$

$$\Leftrightarrow \alpha = \langle \alpha, x_1 \rangle \cdot x_1 + \cdots + \langle \alpha, x_n \rangle x_n,$$

for each $\alpha \in V$.

We'll now show that the functions $\left\{ \frac{\sin(nx)}{\sqrt{\pi}}, \frac{\cos(nx)}{\sqrt{\pi}}, \frac{1}{2\sqrt{\pi}} \right\}_{n=1}^{\infty}$ form an orthonormal set in $C[-\pi, \pi]$ wrt the inner product $\int_{-\pi}^{\pi} f(x) g(x) dx = \langle f, g \rangle$.

This result is important in studying Fourier series.

Let's first assume that
$m, n \in \mathbb{Z}, \quad m \neq n$

$$\int_{-\pi}^{\pi} \sin(nx)\cos(mx)\,dx$$

$$= \int_{-\pi}^{\pi} \frac{1}{2}\left[\sin(n+m)x + \sin(n-m)x\right]dx$$

$$= \frac{1}{2}\left[-\frac{\cos(m+n)x}{(m+n)} - \frac{\cos(n-m)x}{(n-m)}\right]_{-\pi}^{\pi}$$

$$= -\frac{1}{2}\left[\left(\frac{\cos(m+n)\pi}{(m+n)} + \frac{\cos(n-m)\pi}{(n-m)}\right) - \left(\frac{\cos((m+n)(-\pi))}{(m+n)} + \frac{\cos((n-m)(-\pi))}{(n-m)}\right)\right]$$

$= 0.$ (Cosine is an even function,
so, $\cos((m+n)\pi) = \cos((m+n)(-\pi))$
$\cos((n-m)\pi) = \cos((n-m)(-\pi))$)

$$\int_{-\pi}^{\pi} \sin nx \sin mx \, dx$$

$$= \frac{1}{2} \int_{-\pi}^{\pi} \left[\cos(n-m)x - \cos(m+n)x \right] dx$$

$$= \frac{1}{2} \left[\frac{\sin(n-m)x}{n-m} - \frac{\sin(n+m)x}{n+m} \right]_{-\pi}^{\pi}$$

$$= \frac{1}{2} \left[\frac{\sin(n-m)\pi}{n-m} - \frac{\sin(n+m)\pi}{n+m} - \frac{\sin(n-m)(-\pi)}{n-m} + \frac{\sin(n+m)(-\pi)}{n+m} \right]$$

$= 0$ [$\sin(p\pi) = 0$, for any integer p]

$$\int_{-\pi}^{\pi} \cos nx \cos mx \, dx$$

$$= \frac{1}{2} \int_{-\pi}^{\pi} \left[\cos(n+m)x + \cos(n-m)x \right] dx$$

$$= \frac{1}{2} \left[\frac{\sin(n+m)x}{n+m} + \frac{\sin(n-m)x}{n-m} \right]_{-\pi}^{\pi}$$

$$= 0 \quad [\sin(p\pi) = 0, \text{ for any integer } p]$$

Now let $m = n \in \mathbb{Z}$.

Then,
$$\int_{-\pi}^{\pi} \sin(nx)\cos(mx)\,dx$$

$$= \frac{1}{2}\int_{-\pi}^{\pi} \sin(2n\cancel{x})\,dx$$

$$= \frac{1}{2}\left(-\frac{\cos(2nx)}{2n}\right)_{-\pi}^{\pi}$$

$$= -\frac{1}{2}\left(\frac{\cos(2n\pi)}{2n} - \frac{\cos(-2n\pi)}{2n}\right)$$

$$= 0 \quad \left[\text{Cosine is even} \Rightarrow \cos(2n\pi) = \cos(-2n\pi)\right]$$

$$\int_{-\pi}^{\pi} \sin \cancel{n}x \sin mx\,dx$$

$$= \int_{-\pi}^{\pi} \sin^2 nx\,dx = \frac{1}{2}\int_{-\pi}^{\pi}\left(1 - \cos(2nx)\right)dx$$

$$= \frac{1}{2}\left[x - \frac{\sin(2nx)}{2n} \right]_{-\pi}^{\pi}$$

$$= \frac{1}{2}\left[\pi - \frac{\sin(2n\pi)}{2n} + \pi + \frac{\sin(-2n\pi)}{2n} \right]$$

$$= \frac{2\pi}{2} = \pi \quad \left[\sin(p\pi) = 0, \text{ for any integer } p \right]$$

$$\int_{-\pi}^{\pi} \cos nx \cos mx \, dx = \int_{-\pi}^{\pi} \cos^2 nx \, dx$$

$$= \frac{1}{2} \int_{-\pi}^{\pi} (1 + \cos 2nx) \, dx = \frac{1}{2}\left[x + \frac{\sin 2nx}{2n} \right]_{-\pi}^{\pi}$$

$$= \frac{2\pi}{2} \quad \left[\sin(p\pi) = 0, \text{ for any integer } p \right)$$

$$= \pi .$$

Therefore, we've proved that:

$$\int_{-\pi}^{\pi} \sin(nx) \cos(mx) \, dx = 0, \quad \forall m, n \in \mathbb{Z},$$

$$\int_{-\pi}^{\pi} \sin(nx) \sin(mx) \, dx = \begin{cases} 0, & \text{if } m \neq n \\ \pi, & \text{if } m = n \end{cases} ; \quad m, n \in \mathbb{Z},$$

$$\int_{-\pi}^{\pi} \cos(nx) \cos(mx) \, dx = \begin{cases} 0, & \text{if } m \neq n \\ \pi, & \text{if } m = n, \end{cases} \quad m, n \in \mathbb{Z}.$$

Working with the inner product $\langle f, g\rangle = \int_{-\pi}^{\pi} f(x)\, g(x)\, dx$, the results of the previous computations can be phrased as:

$$\langle \sin(nx), \cos(mx)\rangle = 0, \quad m, n \in \mathbb{Z}$$

$\Rightarrow$ $\sin(nx)$ and $\cos(mx)$ are orthogonal for all $m, n \in \mathbb{Z}$.

$$\langle \sin(nx), \sin(mx)\rangle = \begin{cases} 0, & \text{if } n \neq m \\ \pi, & \text{if } n = m \end{cases}$$

and

$$\langle \cos(nx), \cos(mx)\rangle = \begin{cases} 0, & \text{if } n \neq m \\ \pi, & \text{if } n = m \end{cases}$$

we therefore have,

$$\left\langle \frac{\sin(nx)}{\sqrt{\pi}}, \frac{\sin(nx)}{\sqrt{\pi}} \right\rangle = \frac{\pi}{\pi} = 1, \text{ and}$$

$$\left\langle \frac{\cos(nx)}{\sqrt{\pi}}, \frac{\cos(nx)}{\sqrt{\pi}} \right\rangle = \frac{\pi}{\pi} = 1.$$

This proves that $\left\{ \frac{\sin(nx)}{\sqrt{\pi}}, \frac{\cos(nx)}{\sqrt{\pi}} \right\}_{n=1}^{\infty}$ is an orthonormal set over $[-\pi, \pi]$, as they're unit vectors (norm 1), are orthogonal to each other.

Moreover, $\left\langle \frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{2\pi}} \right\rangle$

$$= \int_{-\pi}^{\pi} \frac{dx}{2\pi} = \frac{\pi+\pi}{2\pi} = 1,$$ [so $\frac{1}{\sqrt{2\pi}}$ has norm 1]

$$\left\langle \frac{1}{\sqrt{2\pi}}, \frac{\sin(nx)}{\sqrt{\pi}} \right\rangle$$

$$= \int_{-\pi}^{\pi} \frac{\sin(nx)}{\sqrt{2}\,\pi}\,dx = \frac{-1}{\sqrt{2}\,\pi}\left[\frac{\cos(nx)}{n}\right]_{-\pi}^{\pi}$$

$$= -\frac{1}{\sqrt{2}\,\pi n}\left(\cos(n\pi) - \cos(-n\pi)\right)$$

$$= 0.$$ [Cosine is even $\Rightarrow \cos(n\pi) = \cos(-n\pi)$]

$$\left\langle \frac{1}{\sqrt{2\pi}}, \frac{\cos(nx)}{\sqrt{\pi}} \right\rangle = \frac{1}{\pi\sqrt{2}}\int_{-\pi}^{\pi} \cos(nx)\,dx$$

$$= \frac{1}{\pi\sqrt{2}}\left(\frac{\sin(nx)}{n}\right)_{-\pi}^{\pi}$$

$$= \frac{1}{\pi\sqrt{2}}\left(\frac{\sin(n\pi)}{n} - \frac{\sin(-n\pi)}{n}\right)$$

$$= 0.$$ [$\sin(p\pi) = 0$, for any integer p]

Therefore, we conclude that

$$\left\{ \frac{1}{\sqrt{2\pi}}, \frac{\sin(nx)}{\sqrt{\pi}}, \frac{\cos(nx)}{\sqrt{\pi}} \right\}$$ is an

orthonormal set of functions in $C[-\pi, \pi]$
(as they've norm 1, and are orthogonal to each others).

Gram-Schmidt orthogonalization

We have seen in a previous theorem that a vector can be conveniently represented with respect to an orthonormal basis. In other words, the coordinates of a vector are particularly easy to compute with respect to an orthonormal basis. Fortunately, there is a way to turn any basis (or any linearly independent set of vectors) into an orthonormal basis (or an orthonormal) set of vectors.

Let $\{u_1, \ldots, u_n\}$ be a linearly independent set of vectors.

Define, $\alpha_1 = \dfrac{\beta_1}{\|\beta_1\|}$, where $\beta_1 = u_1$,

$\alpha_2 = \dfrac{\beta_2}{\|\beta_2\|}$, where $\beta_2 = u_2 - \langle u_2, \alpha_1\rangle \alpha_1$

$\alpha_3 = \dfrac{\beta_3}{\|\beta_3\|}$, where $\beta_3 = u_3 - \langle u_3, \alpha_1\rangle \alpha_1 - \langle u_3, \alpha_2\rangle \alpha_2$

$\alpha_n = \dfrac{\beta_n}{\|\beta_n\|}$, where $\beta_n = u_n - \sum_{i=1}^{n-1} \langle u_n, \alpha_i\rangle \alpha_i$

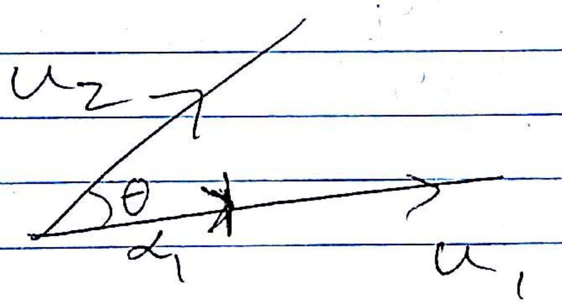
Theorem (Gram-Schmidt):

The set $\{\alpha_1, \ldots, \alpha_n\}$ is orthonormal, and hence linearly independent.

Proof: It's clear from the construction ~~that~~ (using the fact that $u_1, \ldots, u_n$ are linearly independent) that each α_i is a unit vector.

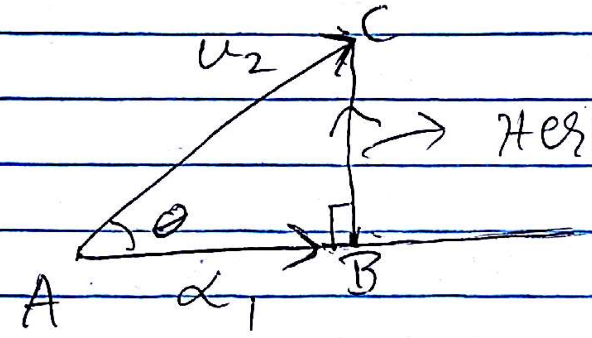
Instead of giving a formal algebraic proof of orthogonality of the vectors, let us convince ourselves geometrically that the vectors α_i are indeed orthogonal to each other.

Take u_1 and u_2.



Since $u_1 \neq 0$ (as it's a part of a linearly independent set), $\alpha_1 = \dfrac{u_1}{\|u_1\|}$ is a unit vector in the direction of u_1.

Now let's drop a perpendicular from u_2 to α_1.



Here's the perpendicular from u_2 to α_1

We know that

$$\langle u_2, \alpha_1 \rangle = \|u_2\| \, \|\alpha_1\| \cos\theta$$

$$\Rightarrow \|u_2\| \cos\theta = \langle u_2, \alpha_1 \rangle \quad (\text{as } \|\alpha_1\| = 1)$$

By our construction, we've

$$|AC| = \|u_2\|,$$

$$|AB| = |AC| \cos\theta = \|u_2\| \cos\theta$$

So, $|AB| = \|u_2\| \cos\theta = \langle u_2, \alpha_1 \rangle$

Observe that the vector $\overrightarrow{AB}$ has magnitude $\langle u_2, \alpha_1 \rangle$ and it's direction is given by the unit vector α_1.

Hence, $\overrightarrow{AB} = \langle u_2, \alpha_1 \rangle \alpha_1$

By usual/Standard vector laws, we know that

$$\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}$$

$$\Rightarrow \langle u_2, \alpha_1 \rangle \alpha_1 + \overrightarrow{BC} = u_2$$

$$\Rightarrow \overrightarrow{BC} = u_2 - \langle u_2, \alpha_1 \rangle \alpha_1 .$$

Evidently, $\overrightarrow{BC} \perp \overrightarrow{AB}$.

Hence, $\overrightarrow{BC} = u_2 - \langle u_2, \alpha_1 \rangle \alpha_1$ is orthogonal to α_1.

That's precisely why we defined:

$$\beta_2 = u_2 - \langle u_2, \alpha_1 \rangle \alpha_1 .$$

This takes care of orthogonality and we make it a unit vector by dividing it by $\|\beta_2\|$

So, $\alpha_2 = \dfrac{\beta_2}{\|\beta_2\|}$ is a unit vector

that is orthogonal to the unit vector α_1.

And the rest is a straight-forward generalization of the previous argument.

This completes the geometric proof of the theorem. $\boxtimes$

Problem (application of Gram-Schmidt)

The vectors $(1,1,1)$, $(1,2,1)$ span a plane P in $\mathbb{R}^3$.
Use the Gram-Schmidt process to find an orthonormal basis of $\mathbb{R}^3$ in which the first two vectors form an orthonormal basis for P.

Solution: clearly, P is 2-dimensional.

Let us find an orthonormal basis for P first (by orthonormalizining the given vectors).
Set $u_1 = (1,1,1)$, $u_2 = (1,2,1)$.

Following Gram-Schmidt we've:

$\beta_1 = u_1 = (1,1,1)$, and $\alpha_1 = \dfrac{\beta_1}{\|\beta_1\|}$

$$\Rightarrow \alpha_1 = \frac{(1,1,1)}{\sqrt{1^2+1^2+1^2}} = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right).$$

So, $\beta_2 = u_2 - \langle u_2, \alpha_1 \rangle \alpha_1$

$$= (1,2,1) - \left\langle (1,2,1), \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right) \right\rangle \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$$

$$= (1,2,1) - \left(\frac{1}{\sqrt{3}} + \frac{2}{\sqrt{3}} + \frac{1}{\sqrt{3}}\right)\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$$

$$= (1,2,1) - \frac{4}{\sqrt{3}}\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$$

$$= (1,2,1) - \left(\frac{4}{3}, \frac{4}{3}, \frac{4}{3}\right)$$

$$= \left(-\frac{1}{3}, \frac{2}{3}, -\frac{1}{3}\right), \text{ So, } \|\beta_2\| = \sqrt{\left(\frac{-1}{3}\right)^2 + \left(\frac{2}{3}\right)^2 + \left(\frac{-1}{3}\right)^2}$$

$$= \sqrt{6}/3$$

Finally, $\alpha_2 = \beta_2 / \|\beta_2\|$

$$= \left(-\frac{1}{3}, \frac{2}{3}, -\frac{1}{3}\right) \Big/ \frac{\sqrt{6}}{3}$$

$$= \frac{3}{\sqrt{6}}\left(-\frac{1}{3}, \frac{2}{3}, -\frac{1}{3}\right)$$

$$= \left(-\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, -\frac{1}{\sqrt{6}}\right)$$

Therefore, an orthonormal basis for P is given by:

$$\left\{ \alpha_1 = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \alpha_2 = \left(\frac{-1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, \frac{-1}{\sqrt{6}}\right) \right\}$$

Remark: Our Construction of α_1 and α_2 guarantees that both α_1 and α_2 are linear Combinations of u_1 and u_2. Hence, $\alpha_1, \alpha_2 \in P$. Since $\{\alpha_1, \alpha_2\}$ is a linearly independent set and $\dim(P) = 2$, we conclude that $\{\alpha_1, \alpha_2\}$ is an orthonormal basis for P.

Now we need to extend $\{\alpha_1, \alpha_2\}$ to a basis of $\mathbb{R}^3$ (in an orthonormal manner)

To this end, we first need to choose some vector $u_3 \in \mathbb{R}^3$ such that

$\{\alpha_1, \alpha_2, u_3\}$ is a basis of $\mathbb{R}^3$ (not necessarily orthonormal). But this is equivalent to choosing u_3 outside

$\operatorname{span}\{\alpha_1, \alpha_2\}$.

We claim that $(1,0,0) \notin \operatorname{span}\{\alpha_1, \alpha_2\} = P$

Indeed if $(1,0,0) \in \operatorname{span}\{\alpha_1, \alpha_2\}$,

Then (by a previous theorem) we would have:

$$(1,0,0) = \langle (1,0,0), \alpha_1 \rangle \alpha_1 + \langle (1,0,0), \alpha_2 \rangle \alpha_2$$

or $$(1,0,0) = \frac{1}{\sqrt{3}}\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right) + \left(-\frac{1}{\sqrt{6}}\right)\left(-\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, \frac{-1}{\sqrt{6}}\right)$$

or $$(1,0,0) = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) + \left(\frac{1}{6}, -\frac{2}{6}, \frac{1}{6}\right)$$

or $$(1,0,0) = \left(\frac{1}{2}, 0, \frac{1}{2}\right) \quad !!$$

This contradiction shows that
$(1,0,0) \notin \operatorname{span}\{\alpha_1, \alpha_2\}$.

Hence, $\{\alpha_1, \alpha_2, \underset{\substack{\parallel \\ u_3}}{(1,0,0)}\}$ is a basis of $\mathbb{R}^3$.

Since we already have that $\langle \alpha_1, \alpha_2 \rangle = 0$, in order to turn this basis into an orthonormal basis of $\mathbb{R}^3$, we only need to consider:

$$\beta_3 := u_3 - \langle u_3, \alpha_1 \rangle \alpha_1 - \langle u_3, \alpha_2 \rangle \alpha_2$$

$$= (1,0,0) - \left(\frac{1}{2}, 0, \frac{1}{2}\right)$$

$$= \left(\frac{1}{2}, 0, -\frac{1}{2}\right).$$

So, $\|\beta_3\| = \sqrt{\left(\frac{1}{2}\right)^2 + 0^2 + \left(\frac{-1}{2}\right)^2} = \frac{1}{\sqrt{2}}$.

So, $\alpha_3 = \dfrac{\beta_3}{\|\beta_3\|} = \sqrt{2}\left(\frac{1}{2}, 0, \frac{-1}{2}\right)$

$$\Rightarrow \alpha_3 = \left(\frac{1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}}\right).$$

Therefore, the ~~required~~ basis

$$\left\{ \alpha_1 = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \alpha_2 = \left(-\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, -\frac{1}{\sqrt{6}}\right), \alpha_3 = \left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right) \right\}$$ is

orthonormal and its first two vectors span P. $\boxtimes$

We consider differential equations of the form

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = 0 \, . \rightarrow (*)$$

An ad hoc method of solving (*) is to try an exponential solution

$$y = e^{rx}.$$

With this choice of y, (*) reduces to:

$$ar^2e^{rx} + br\,e^{rx} + c\,e^{rx} = 0$$

$$\Rightarrow (ar^2 + br + c)\,e^{rx} = 0$$

$$\Rightarrow ar^2 + br + c = 0 \quad (\text{as } e^{rx} \neq 0, \text{ for all } x)$$

$\hookrightarrow$ (1)

Thus, if $r = \dfrac{-b \pm \sqrt{b^2 - 4ac}}{2a}$, then

$y = e^{rx}$ is a solution of (*).

However, this ad hoc doesnot tell us whether we managed to find all solutions of (*). These lecture notes will be devoted to finding the most general solution of (*).

We'll use our knowledge of linear algebra to attack this problem. To connect "Solving the diff. eqn. (*)" to "linear algebra", we need to first choose the right vector space. Since we are going to differentiate functions, a safe choice is:

$$V := C^\infty(\mathbb{R}) = \{ f : \mathbb{R} \to \mathbb{R} : f \text{ is infinitely differentiable} \}$$

We denote the linear map $\frac{d}{dx} : V \to V$ by D.

Then, $\frac{d^2}{dx^2} : V \to V$ is the two-fold composition of the operator D (i.e. $D \circ D = \frac{d}{dx}\left(\frac{d}{dx}\right) = \frac{d^2}{dx^2}$).

We denote $\frac{d^2}{dx^2} : V \to V$ by $D^2 (= D \circ D)$.

In general, D^n stands for $\underbrace{D \circ D \circ \cdots \circ D}_{n\text{-fold}} = \frac{d^n}{dx^n}$.

With these terminology solving ⊛ is equivalent to solving the equation

$$(aD^2 + bD + cI)(y) = 0 \rightarrow (**)$$

where $I : V \to V$ is the identity linear map; i.e. $I(f) = f$.

Let S be the space of all solutions of ⊛ or $(**)$ (they're equivalent). Observe that

$aD^2 + bD + cI : V \to V$ is a linear map.

setting $T := (aD^2 + bD + cI)$, it follows that

$$S = \text{null space of } T = \text{null}(T).$$

It is immediate (since null(T) is a subspace of V) that S is closed under linear combinations. We are interested in finding a basis for S, which will give us a general solution to $(*)$.

Abuse of notation: In the sequel, we will denote the identity map $I: V \to V$ by 1. So, e.g., (D^2+I) will be written as (D^2+1), and (D^2-3I) will written as (D^2-3) and so on.

"Factoring" differential operator (not literally):

Consider the linear map $(D^2-1): V \to V$. Note that

$$(D+1)\circ(D-1)(f)$$

We aren't "multiplying" $(D+1)$ and $(D-1)$, we're composing them.

$$= (D+1)(Df-f)$$

$$= D(Df-f) + I(Df-f)$$

$$= D^2f - Df + Df - f = D^2f - f$$
$$= (D^2-1)(f)$$

Thus, $(D+1)\cdot(D-1)(f) = (D^2-1)(f)$,

for all $f \in V$.

Thus justifies the equality of linear map:

$$(D^2-1) = (D+1)\circ(D-1).$$

We'll use this technique to find general solutions of $(*)$.

Recall that $(*)$ was equivalent to $(**)$:

$$(aD^2+bD+c)(y) = 0.$$

Dividing $(**)$ by a, we can assume that $(**)$ is of the form:

$$(D^2+PD+q)(y) = 0.$$

Our plan is to "factorize" [I'll stop putting quotation marks from now on, this was the last reminder that factoring indicates composition, and not usual multiplication in this context] the linear map

$$(D^2+PD+q).$$

But this is the same as factoring the polynomial (x^2+Px+q). We know that if r_1, r_2 are solutions of the quadratic equation $x^2+Px+q=0$, then,

$$x^2 + px + q \equiv (x - r_1)(x - r_2).$$

[This is an identity, it holds for all x]

One possible dilemma is that not all real quadratic equations have real roots! But we are brave enough to venture into the world of complex numbers. In fact, the use of complex numbers is unavoidable although we are dealing with differential equations (or quadratic equations) with real coefficients.

Theorem (Fundamental theorem of algebra - special case)

Let $P(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$ be a degree n real polynomial; i.e, $a_1, \ldots, a_n \in \mathbb{R}$. Then there exist complex numbers $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ such that

$$P(x) = a_n (x - \lambda_1) \cdots (x - \lambda_n).$$

In other words, $P(x)$ can be completely factorized over $\mathbb{C}$. However, the λ_is are not necessarily distinct.

Example: The Polynomial (x^2+1) factors as $(x+i)(x-i)$ over $\mathbb{C}$.
It's not possible to factorize (x^2+1) over $\mathbb{R}$. So the passage from $\mathbb{R}$ to $\mathbb{C}$ is completely natural.

Now we return to the topic of solving
$$(D^2+PD+q)(y)=0.$$

Step-I: Solve the quadratic equation
$$x^2+Px+q=0.$$

Suppose that the roots are r_1, r_2.

Then, $x^2+Px+q=(x-r_1)(x-r_2)$.
Therefore, we've a corresponding factorization of differential operators:
$$(D^2+PD+q)=(D-r_1)\circ(D-r_2).$$

Step-II: Write $(D^2+PD+r)(y)=0$ as
$$(D-r_1)\circ(D-r_2)(y)=0. \rightarrow (2).$$

Let us set $(D-r_2)(y)=u. \rightarrow (3)$

Then 2 reduces to:

$$(D - r_1)(u) = 0$$

$$\Rightarrow \frac{du}{dx} = r_1 u \quad \Rightarrow \int \frac{du}{u} = \int r_1\, dx$$

$$\Rightarrow \ln u = r_1 x + c_1,$$

note that adding the constant c_1 ensures that $(r_1 x + c_1)$ is the most general integral = anti-derivative of r_1.

$$\Rightarrow u = e^{c_1} e^{r_1 x}$$

$$\Rightarrow u = A e^{r_1 x}$$

(setting $A = e^{c_1}$)

Now, plugging $u = A e^{r_1 x}$ in 3, we get:

$$(D - r_2)(y) = A e^{r_1 x}$$

$$\Rightarrow \frac{dy}{dx} - r_2 y = A e^{r_1 x}$$

$$\Rightarrow e^{-r_2 x} \frac{dy}{dx} - r_2 e^{-r_2 x} y = A e^{r_1 x} \cdot e^{-r_2 x}$$

$e^{-r_2 x}$ is the integrating factor

$$\Rightarrow \frac{d}{dx}\left(y e^{-r_2 x}\right) = A e^{(r_1 - r_2) x}$$

$$\Rightarrow \int d\left(y e^{-r_2 x}\right) = \int A e^{(r_1 - r_2) x} dx$$

$$\Rightarrow ye^{-r_2 x} = \begin{cases} \dfrac{A e^{(r_1 - r_2)x}}{r_1 - r_2} + C_2, & \text{when } r_1 \neq r_2 \\ Ax + C_2, & \text{when } r_1 = r_2 \end{cases}$$

(once again, adding the constant C_2 ensures that we considered the most general integral of $Ae^{(r_1-r_2)}$.)

$$\Rightarrow y = \begin{cases} \dfrac{A}{r_1 - r_2} e^{r_1 x} + C_2 e^{r_2 x}, & \text{if } r_1 \neq r_2 \\ Axe^{r_2 x} + C_2 e^{r_2 x}, & \text{if } r_1 = r_2 \end{cases}$$

$$\Rightarrow y = \begin{cases} \tilde{A} e^{r_1 x} + \tilde{B} e^{r_2 x}, & \text{if } r_1 \neq r_2 \\ \tilde{A} x e^{r_2 x} + \tilde{B} e^{r_2 x}, & \text{if } r_1 = r_2. \end{cases}$$

(here, $\tilde{A}, \tilde{B}$ are arbitrary constants.)

We've now found the most general solution of the equation

$$\frac{d^2y}{dx^2} + P\frac{dy}{dx} + qy = 0$$

$$\Leftrightarrow \left(D^2 + PD + q\right)(y) = 0.$$

It turns out that the space of solutions S of $(D^2+pD+q)(y)=0$ is:

$$S=\begin{cases}\text{span}\{e^{r_1x}, e^{r_2x}\}, & \text{if } r_1\neq r_2\\ \text{span}\{e^{rx}, xe^{rx}\}, & r:=r_1=r_2\end{cases}$$

What about $\dim(S)$?

Well, it seems that S is 2-dimensional, a fact that will follow if we could prove that the sets $\{e^{r_1x}, e^{r_2x}\}$ (where $r_1\neq r_2$) and $\{e^{rx}, xe^{rx}\}$ are linearly independent.

It's rather simple to prove these facts directly, but we want to device a method that can be generalized to higher order diff. eqns, meaning we are seeking a method of proving linear independence of any finite collection of smooth functions $\{f_1, f_2, \dots, f_n\}$.

This leads us to the discussion of Wronskians.

Let $\{f_1, f_2, \ldots, f_n\}$ be a linearly dependent set of smooth functions.
(smooth means infinitely differentiable)

Consider the matrix:

$$\begin{pmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ f_1'(x) & f_2'(x) & \cdots & f_n'(x) \\ \vdots & & & \vdots \\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \cdots & f_n^{(n-1)}(x) \end{pmatrix}$$

[$f_i^{(k)}$ denotes the k-th derivative of f_i]

Since $\{f_1, \ldots, f_n\}$ is linearly dependent, there exist constants $\{a_1, \ldots, a_n\}$ with

$(a_1, \ldots, a_n) \neq (0, \ldots, 0)$ [meaning not all a_is are 0]

Such that $\underbrace{a_1 f_1(x) + \cdots + a_n f_n(x) = 0}_{\text{for all } x \in \mathbb{R}}$ → (4)

Differentiating 4 $(n-1)$-times, we get:

$$\left.\begin{aligned} a_1 f_1(x) + \cdots + a_n f_n(x) &= 0 \\ a_1 f_1'(x) + \cdots + a_n f_n'(x) &= 0 \\ \cdots\cdots \\ a_1 f_1^{(n-1)}(x) + \cdots + a_n f_n^{(n-1)}(x) &= 0 \end{aligned}\right\} \text{for all } x \in \mathbb{R}$$

$$\Rightarrow \begin{pmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ f_1'(x) & f_2'(x) & \cdots & f_n'(x) \\ \cdots & & & \\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \cdots & f_n^{(n-1)}(x) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = 0,$$

for all $x \in \mathbb{R}$.

Now, for any $x \in \mathbb{R}$, the above system of linear equations admit a non-trivial solution $(a_1, a_2, \ldots, a_n) \neq (0, 0, \ldots, 0)$

Hence, $$\boxed{\det \begin{pmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ \cdots & & & \\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \cdots & f_n^{(n-1)}(x) \end{pmatrix} = 0},$$

for every $x \in \mathbb{R}$.

We define:

$$W(f_1, f_2, \ldots, f_n)(x)$$

$$= \det \begin{pmatrix} f_1(x) & \cdots & f_n(x) \\ \vdots & & \vdots \\ f_1^{(n-1)}(x) & \cdots & f_n^{(n-1)}(x) \end{pmatrix}$$

The above argument proves that:

Theorem: If $\{f_1, \ldots, f_n\}$ is a linearly dependent set of smooth functions, then $W(f_1, \ldots, f_n)(x) = 0$, for all $x \in \mathbb{R}$.

Equivalently, if the function $W(f_1, \ldots, f_n)(x)$ is not identically zero, then the functions $\{f_1, \ldots, f_n\}$ is linearly independent.

We now use this theorem to show that the functions $\{e^{r_1 x}, e^{r_2 x}\}$ (where $r_1 \neq r_2$) and $\{e^{rx}, xe^{rx}\}$ are linearly independent over $\mathbb{R}$.

A simple computation shows that

$$W(e^{r_1 x}, e^{r_2 x})(x) = \begin{vmatrix} e^{r_1 x} & e^{r_2 x} \\ r_1 e^{r_1 x} & r_2 e^{r_2 x} \end{vmatrix}$$

$$= (r_2 - r_1)\, e^{(r_1 + r_2)x}$$

$$\neq 0, \text{ for all } x \in \mathbb{R}$$

(note that $r_2 - r_1 \neq 0$, $e^{(r_2 - r_1)x} \neq 0$)

Thus, $W(e^{r_1 x}, e^{r_2 x})(x)$ is not identically zero (in fact, never zero) and hence $\{e^{r_1 x}, e^{r_2 x}\}$ is a linearly independent set.

I leave it to you show (by a similar computation) that $W(e^{rx}, x e^{rx})(x)$ is not identically zero, and hence $\{e^{rx}, x e^{rx}\}$ is a linearly independent set.

This proves that for any linear 2nd order diff eqn. (with constant coeff.)

$$(D^2 + pD + q)(y) = 0,$$

the solution space S admits a

basis $\{e^{r_1 x}, e^{r_2 x}\}$, [when r_1, r_2 are two distinct roots of $x^2+px+q=0$]

or $\{e^{rx}, xe^{rx}\}$. [when r is the unique root of $x^2+px+q=0$]

In either case, $\dim(S)=2$, = order of the diff. eqn.

We have now completed our analysis of solutions of 2nd order linear diff eqns. with constant coefficients.

Let us summarize our observation.

For the diff equation $(D^2+pD+q)(y)=0$, let the roots of the quadratic eqn.

$$x^2+px+q=0 \quad \text{be}$$

$\{r_1, r_2\}$.

Case-I: $(r_1, r_2 \in \mathbb{R},\ r_1 \neq r_2)$.

The general solution in this case is:

$$y = Ae^{r_1 x} + Be^{r_2 x}.$$

Case-II: $(r_1, r_2 \in \mathbb{R},\ r_1 = r_2)$

The general solution is:

$$y = Ae^{r_1 x} + Bxe^{r_1 x}.$$

Case-III: (r_1, r_2 are complex conjugate)

Let $r_1 = \alpha + i\beta$, $r_2 = \alpha - i\beta$, where $\alpha, \beta \in \mathbb{R}$.

The general solution in this case is:

$$\begin{aligned}
y &= Ae^{r_1 x} + Be^{r_2 x} \\
&= Ae^{(\alpha+i\beta)x} + Be^{(\alpha-i\beta)x} \\
&= Ae^{\alpha x}(\cos\beta x + i\sin\beta x) + Be^{\alpha x}(\cos\beta x - i\sin\beta x) \\
&= e^{\alpha x}\cos\beta x\,(A+B) + e^{\alpha x}\sin\beta x\,(iA - iB) \\
&= \tilde{A}e^{\alpha x}\cos\beta x + \tilde{B}e^{\alpha x}\sin\beta x \\
&= e^{\alpha x}\left(\tilde{A}\cos(\beta x) + \tilde{B}\sin(\beta x)\right).
\end{aligned}$$

Finally, if we have an initial-value problem:

$$\begin{cases} (D^2 + PD + q)(y) = 0, \\ y'(\alpha_1) = \beta_1, \ y(\alpha_1) = \beta_2 \end{cases}$$

or

$$\begin{cases} (D^2 + PD + q)(y) = 0, \\ y(\alpha_1) = \beta_1, \quad y(\alpha_2) = \beta_2 \end{cases}$$

(or some other condition), then we can use these conditions to find the constants and get a particular solution.

Of course, all that we did so far, can be generalized to solve an n-th order linear diff. eqn. with constant coefficients. Since the generalization is straightforward, we'll skip the details.

A bit more on Wronskians.

Recall that a set of smooth (or sufficiently diff.) functions is linearly independent (over $\mathbb{R}$) if their Wronskian

$W(f_1, \ldots, f_n)(x)$ is not identically zero on $\mathbb{R}$.

A natural question arises: if $W(f_1, \ldots, f_n)(x)$ is identically zero on $\mathbb{R}$, does it follow that $\{f_1, \ldots, f_n\}$ is a linearly dependent set?

The answer is, unfortunately, no!

In a homework problem, you'll prove that the functions

$\{x^2, x|x|\}$ have vanishing Wronskian on $\mathbb{R}$, but they are linearly independent.

A usual approach in mathematics is to hope for the best, but not to give up if the best is too good to hold!

Now, we have that:

$$\left\{\begin{array}{l}\text{Wronkian}\\ \text{vanishing}\\ \text{identically}\end{array}\right\} \underset{\substack{\text{doesnot}\\ \text{necessarily}\\ \text{imply}}}{\overset{\text{implies}}{\rightleftarrows}} \left\{\begin{array}{l}\text{Linear}\\ \text{dependence}\end{array}\right\}.$$

Can we come up with a condition (Ω) such that:

$$\left\{\begin{array}{l}\text{Wronskian}\\ \text{Vanishing}\\ \text{identically}\end{array} + (\Omega)\right\} \overset{\text{implies}}{\Longrightarrow} \left\{\begin{array}{l}\text{Linear}\\ \text{dependence}\end{array}\right\}?$$

The answer is **yes!**

For two functions f_1, f_2, the condition (Ω) is particularly simple.

Theorem:

Let f_1, f_2 be real differentiable functions such that $f_1(x) \neq 0$, for all $x \in \mathbb{R}$.
(i.e. f_1 doesnot vanish at any point).

If $W(f_1, f_2)(x) = 0$, for all $x \in \mathbb{R}$, then $\{f_1, f_2\}$ are linearly dependent.

Proof: The condition
$$W(f_1, f_2)(x) = 0 \quad \text{means}$$

$$\det \begin{vmatrix} f_1(x) & f_2(x) \\ f_1'(x) & f_2'(x) \end{vmatrix} = 0, \text{ for all } x \in \mathbb{R}$$

$$\Rightarrow f_1(x) f_2'(x) - f_2(x) f_1'(x) = 0, \text{ for all } x \in \mathbb{R}$$

$$\Rightarrow \frac{f_1(x) f_2'(x) - f_2(x) f_1'(x)}{(f_1(x))^2} = 0, \text{ for all } x \in \mathbb{R}$$

(Dividing by $(f_1(x))^2$ makes sense because $f_1(x) \neq 0$, for all $x \in \mathbb{R}$)

$$\Rightarrow \frac{d}{dx}\left(\frac{f_2(x)}{f_1(x)}\right) = 0, \text{ for all } x \in \mathbb{R} \quad \left[\text{Quotient rule of differentiation}\right]$$

$$\Rightarrow \frac{f_2(x)}{f_1(x)} = c, \text{ for some } c \in \mathbb{R}$$

$$\Rightarrow \underline{f_2(x) = c f_1(x).}$$

This shows that $\{f_1(x), f_2(x)\}$ is a linearly dependent set of functions. ⊠

<u>Remark</u>: In the case of two functions $\{f_1, f_2\}$, the extra condition (Ω) is that $f_1(x) \neq 0$, for all $x \in \mathbb{R}$.

The general answer to the question is given in the following theorem.

<u>Theorem</u>: Let $\{f_1(x), f_2(x), \ldots, f_n(x)\}$ be $(n-1)$ times differentiable functions. Suppose that:

i) $W(f_1, f_2, \cdots, f_n)(x) = 0$, for all $x \in \mathbb{R}$,

and

ii) $W(f_1, f_2, \cdots, f_{n-1})(x) \neq 0$, for all $x \in \mathbb{R}$.

(This is Condition (Ω))

Then, $\{f_1, \cdots, f_n\}$ is a linearly dependent set of functions, and there exist $c_1, \cdots, c_{n-1} \in \mathbb{R}$ such that

$$f_n \equiv c_1 f_1 + c_2 f_2 + \cdots + c_{n-1} f_{n-1}.$$

As an application of the previous two theorems, a homework problem will outline an alternative proof of the fact that:

$$\dim\left(\mathrm{Null}\left(D^2 + PD + q : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})\right)\right)$$

$$= \dim\left(\text{space of solutions of } \frac{d^2y}{dx^2} + P\frac{dy}{dx} + qy = 0\right)$$

$$= 2.$$

Laplace transforms

For a function $f:[0,+\infty) \to \mathbb{R}$ satisfying some suitable growth condition, the Laplace transform of f, denoted by $\mathcal{L}[f]$ is another function

$$\mathcal{L}[f](s) = \int_0^\infty e^{-st} f(t)\,dt$$

Since this is an improper integral, we must impose some 'slow growth' condition on f to guarantee convergence of the integral.

The most basic example:

Let $f:[0,+\infty) \to \mathbb{R}$
$f(t) = e^{at}$

Then
$$\mathcal{L}[f](s) = \int_0^\infty e^{-st} f(t)\,dt$$
$$= \int_0^\infty e^{(a-s)t}\,dt$$
$$= \lim_{T\to+\infty} \int_0^T e^{(a-s)t}\,dt$$
$$= \begin{cases} \lim_{T\to+\infty} \left[\dfrac{e^{(a-s)t}}{(a-s)} \right]_0^T, & \text{if } s \neq a \end{cases}$$

②

$$\begin{cases} \lim_{T\to+\infty} \left[t \right]_0^T & , \text{if } a=s \end{cases}$$

$$= \begin{cases} \lim_{T\to+\infty} \left[\dfrac{e^{(a-s)T} - 1}{a-s} \right] & , \text{if } s \neq a \\ \lim_{T\to+\infty} (T) & , \text{if } s=a \end{cases}$$

Now, the limits in the above improper exist only if $\underline{s > a}$.

Therefore, $\mathcal{L}[f](s)$ is defined when $s > a$, and

$$\mathcal{L}[f](s) = \frac{1}{s-a}, \text{ for } s > a.$$

We'll now see that under suitable conditions computing the Laplace transform of $f^{(n)}$

is rather simple; in fact, we'll give an explicit formula for $\mathcal{L}[f^{(n)}]$.

First let, $n=1$.

Now, $\mathcal{L}[f'](s) = \int_0^\infty e^{-st} f'(t)\,dt$

Using integration by parts, we've :

$$\int e^{-st} f'(t)\,dt$$

$$= e^{-st}\int f'(t)\,dt - \int\left(\frac{d}{dt}(e^{-st})\right)\left(\int f'(t)\,dt\right)dt$$

$$= e^{-st} f(t) + s\int e^{-st} f(t)\,dt.$$

Therefore, $\mathcal{L}[f'](s) = \lim\limits_{T\to+\infty} \int_0^T e^{-st} f'(t)\,dt$

$$= \lim_{T\to+\infty}\left(\left(e^{-st}f(t)\right)_0^T + s\int_0^T e^{-st} f(t)\,dt\right)$$

$$= \lim_{T\to+\infty}\left(\frac{f(T)}{e^{sT}} - \frac{f(0)}{e^0} + s\int_0^T e^{-st} f(t)\,dt\right)$$

If we assume that $\mathcal{L}[f](s)$ exists;

i.e. $\lim\limits_{T\to+\infty}\int_0^T e^{-st} f(t)\,dt$ is convergent and f satisfying ~~the~~ "tame growth" Condition:

$$\lim_{T\to+\infty} \frac{f(T)}{e^{ST}} = 0$$

(meaning that f grows sub-exponentially i.e. f grows slower than e^{ST})

then $\boxed{\mathcal{L}[f'](S) = S\,\mathcal{L}[f](S) - f(0)} \to (*)$

Using $(*)$ on $f'' = (f')'$ we get:

$$\mathcal{L}[f''](S) = S\,\mathcal{L}[f'](S) - f'(0)$$

$$\Rightarrow \mathcal{L}[f''](S) = S\left(S\,\mathcal{L}[f](S) - f(0)\right) - f'(0)$$

$$\Rightarrow \mathcal{L}[f''](S) = S^2\mathcal{L}[f](S) - Sf(0) - f'(0)$$

Of course, we need to assume here that

i) $\mathcal{L}[f](S)$ exists,

(i.e. f and f' grow sub-exponentially)

ii) $\lim_{T\to+\infty} \frac{f(T)}{e^{ST}} = 0$

iii) $\lim_{T\to+\infty} \frac{f'(T)}{e^{ST}} = 0.$

An inductive argument now shows that

$$\boxed{\mathcal{L}[f^{(n)}](s) = s^n \mathcal{L}[f](s) - s^{n-1} f(0) - \cdots - f^{(n-1)}(0).}$$

This gives a formula for finding $\mathcal{L}[f^{(n)}]$ directly from $\mathcal{L}[f]$.

It'll also be useful to find a formula for the integral of a function f in terms of $\mathcal{L}[f]$.
To do so, set

$$g := \int_0^t f(u)\, du$$

Applying (*) on g, we've:

$$\mathcal{L}[g'](s) = s\mathcal{L}[g](s) - g(0)$$

$$\Rightarrow \mathcal{L}[f](s) = s\mathcal{L}\left[\int_0^t f(u)\,du\right](s) - \int_0^0 f(u)\,du$$

[Fundamental thm. of Calculus: $g' = f$]

$$\Rightarrow \boxed{\mathcal{L}\left[\int_0^t f(u)\,du\right] = \frac{1}{s}\mathcal{L}[f](s)} \rightarrow (**)$$

Formulas (*) and (**) can be used to compute various other Laplace transforms.

Let's compute ~~$\mathcal{L}[t](s)$~~ $\mathcal{L}[f]$, where

$f(t) = t$.

$$\mathcal{L}[f](s) = \int_0^{+\infty} e^{-st} f(t)\,dt \quad , \quad s > 0$$

$$= \int_0^{\infty} t e^{-st}\,dt$$

$$= \lim_{T \to +\infty} \int_0^T \left(t e^{-st}\right) dt$$

$$= \lim_{T \to +\infty} \left(\frac{t e^{-st}}{(-s)} + \frac{1}{s} \int e^{-st}\,dt \right)\Bigg|_0^T$$

$$= \lim_{T \to +\infty} \left(-\frac{t}{s} e^{-st} - \frac{1}{s^2} e^{-st} \right)\Bigg|_0^T$$

$$= \lim_{T \to +\infty} \left(\left(\frac{-T}{s e^{sT}} - \frac{1}{s^2 e^{sT}} \right) - \left(-0 - \frac{1}{s^2} e^0 \right) \right)$$

$$= \frac{1}{s^2}.$$

Thus, $\boxed{\mathcal{L}[f](s) = \frac{1}{s^2}, \text{ where } f(t) = t,}$

(on $s > 0$)

Applying (**) on this, we get:

$$\mathcal{L}\left[\int_0^t u\,du\right] = \frac{1}{s}\mathcal{L}[t](s)$$

$$\Rightarrow \mathcal{L}\left[\frac{t^2}{2}\right] = \frac{1}{s^3}$$

$$\Rightarrow \boxed{\mathcal{L}[t^2](s) = \frac{2}{s^3}}$$

Inductively we've:

$$\boxed{\mathcal{L}[t^n](s) = \frac{n!}{s^{n+1}}}, \quad s>0$$

A fundamental property of Laplace transforms that follows directly from the definition is that:

$$\underline{\mathcal{L}[af+bg] \equiv a\mathcal{L}[f] + b\mathcal{L}[g]}$$

where $a, b \in \mathbb{R}$.

So, $\mathcal{L}$ is a linear operator.

Now let: $f(t) = te^{at}$

so, $f'(t) = ate^{at} + e^{at}$

$\Rightarrow f'(t) = af(t) + e^{at}$

$\Rightarrow \mathcal{L}[f'](s) = a\mathcal{L}[f](s) + \mathcal{L}[e^{at}](s)$ (linearity)

(by ⊛) $\Rightarrow s\mathcal{L}[f](s) - f(0) = a\mathcal{L}[f](s) + \dfrac{1}{s-a}$, (when $s>a$)

$\Rightarrow (s-a)\,\mathcal{L}[f](s) = \dfrac{1}{s-a}$ $\quad (f(0)=0)$

$\Rightarrow \boxed{\mathcal{L}[f](s) = \dfrac{1}{(s-a)^2}}$, when $s>a$

Using the above trick repeatedly, we get:

$$\boxed{\mathcal{L}[t^n e^{at}](s) = \frac{n!}{(s-a)^{n+1}}}, \quad n=0,1,2,\ldots$$

for $s>a$.

In order to find $\mathcal{L}[\sin at]$, we'll use a "complex" trick.

For $a \in \mathbb{R}$, we've:

$$e^{iat} = \cos at + i \sin at$$

$$e^{-iat} = \cos at - i \sin at.$$

So, $\cos(at) = \dfrac{e^{iat} + e^{-iat}}{2}$,

$$\sin(at) = \frac{e^{iat} - e^{-iat}}{2i}.$$

By linearity of $\mathcal{L}$, we've:

$$\mathcal{L}[\cos(at)](s) = \frac{1}{2}\mathcal{L}[e^{iat}](s) + \frac{1}{2}\mathcal{L}[e^{-iat}](s)$$

$$= \frac{1}{2}\frac{1}{(s-ia)} + \frac{1}{2}\frac{1}{(s+ia)}$$

$$= \frac{1}{2}\left(\frac{s+i\cancel{a}+s-i\cancel{a}}{(s-ia)(s+ia)}\right)$$

$$\boxed{\mathcal{L}[\cos(at)](s) = \frac{s}{s^2+a^2}}$$

Similarly, $\mathcal{L}[\sin at](s)$

$$= \frac{1}{2i}\left(\mathcal{L}[e^{iat}](s) - \mathcal{L}[e^{-iat}](s)\right)$$

$$= \frac{1}{2i}\left(\frac{1}{s-ia} - \frac{1}{s+ia}\right)$$

$$= \frac{1}{2i}\frac{2ia}{(s^2+a^2)} = \frac{a}{s^2+a^2}.$$

So, $$\boxed{\mathcal{L}[\sin at](s) = \frac{a}{s^2+a^2}.}$$

Finally, if

$$f(t) = t\sin at,$$

then,

$$f'(t) = at\cos at + \sin at$$

$\Rightarrow$ ~~$\mathcal{L}[f'](s) =$~~

$$\Rightarrow f''(t) = a\cos at - a^2 t\sin at + a\cos at$$

$$\Rightarrow f''(t) = 2a\cos at - a^2 f(t)$$

$$\Rightarrow \mathcal{L}[f''](s) = 2a\,\mathcal{L}[\cos at](s) - a^2\,\mathcal{L}[f](s)$$

(linearity of $\mathcal{L}$)

$$\Rightarrow s^2 \mathcal{L}[f](s) - s f(0) - f'(0) = \frac{2as}{s^2+a^2} - a^2 \mathcal{L}[f](s)$$

$$\Rightarrow (s^2+a^2)\mathcal{L}[f](s) - 0 - 0 = \frac{2as}{s^2+a^2}$$

$$\Rightarrow \boxed{\mathcal{L}[f](s) = \frac{2as}{(s^2+a^2)^2}}$$

Similarly, we've:

$$\boxed{\mathcal{L}[t\cos at](s) = \frac{s^2-a^2}{(s^2+a^2)^2}}.$$

<u>Inverse Laplace transform:</u>

If $\mathcal{L}[f] = g$, then we say that

$$\underline{\mathcal{L}^{-1}[g] = f},$$ and f is

called the inverse Laplace transform of g. For instance,

$$\underline{\mathcal{L}^{-1}\left[\frac{1}{s^2}\right] = t}.$$

Now we're in a position to apply Laplace transforms to solve linear differential equation.

<u>Example:</u> We want to solve

$$y'' + y = e^{-t} + 1, \quad y(0) = -1, \quad y'(0) = 1.$$

Assuming that the solution $y(t)$ satisfies all the required growth conditions, we've:

$$\mathcal{L}[y''+y](s) = \mathcal{L}[e^{-t}+1](s)$$

$$\Rightarrow \mathcal{L}[y''](s) + \mathcal{L}[y](s) = \mathcal{L}[e^{-t}](s) + \mathcal{L}[1](s)$$

$$\Rightarrow s^2\mathcal{L}[y](s) - s\,y(0) - y'(0) + \mathcal{L}[y](s) = \frac{1}{s-(-1)} + \frac{1}{s}$$

$$\Rightarrow (s^2+1)\mathcal{L}[y](s) + s - 1 = \frac{1}{s+1} + \frac{1}{s}$$

$$\Rightarrow (s^2+1)\mathcal{L}[y](s) = \frac{1}{s+1} + \frac{1}{s} + 1 - s$$

$$\cancel{\Rightarrow (s^2+1)\mathcal{L}[y](s) = \frac{s^4(s+1) + s(s+1) - s^4(s+1)}{s(s+1)}}$$

$$\Rightarrow \mathcal{L}[y](s) = \frac{1}{(s+1)(s^2+1)} + \frac{1}{s(s^2+1)} + \frac{1}{s^2+1} - \frac{s}{s^2+1}$$

$$\Rightarrow y(t) = \mathcal{L}^{-1}\left[\frac{1}{(s+1)(s^2+1)}\right](t) + \mathcal{L}^{-1}\left[\frac{1}{s(s^2+1)}\right](t)$$

$$+ \mathcal{L}^{-1}\left[\frac{1}{s^2+1}\right](t) - \mathcal{L}^{-1}\left[\frac{s}{s^2+1}\right](t)$$

Now, $\mathcal{L}^{-1}\left[\frac{1}{s^2+1}\right](t) = \sin t$,

$$\mathcal{L}^{-1}\left[\frac{s}{s^2+1}\right](t) = \cos t$$

In order to find the other two inverse Laplace transforms, we want to use the method of partial fractions.

Let, $$\frac{1}{(s+1)(s^2+1)} = \frac{A}{s+1} + \frac{Bs+C}{s^2+1}$$

$$= \frac{As^2+A+Bs^2+Bs+Cs+C}{(s+1)(s^2+1)}$$

$$\Rightarrow \frac{1}{(s+1)(s^2+1)} = \frac{s^2(A+B) + (B+C)s + (A+C)}{(s+1)(s^2+1)}$$

Hence, $A+B=0 \Rightarrow \underline{B=-A}$.

$B+C=0 \Rightarrow \underline{C=-B=A}$,

$A+C=1 \Rightarrow 2A=1 \Rightarrow A=\frac{1}{2}, C=\frac{1}{2}, B=-\frac{1}{2}$

Hence, $\mathcal{L}^{-1}\left[\frac{1}{(s+1)(s^2+1)}\right](t)$

$$=\frac{1}{2}\mathcal{L}^{-1}\left[\frac{1}{(s+1)}\right](t)+\frac{1}{2}\mathcal{L}^{-1}\left[\frac{1}{s^2+1}\right](t)$$

$$-\frac{1}{2}\mathcal{L}^{-1}\left[\frac{s}{s^2+1}\right](t)$$

$$=\left(\frac{1}{2}e^{-t}+\frac{1}{2}\sin t-\frac{1}{2}\cos t\right).$$

Finally, let

$$\frac{1}{s(s^2+1)}=\frac{A'}{s}+\frac{B's+C'}{s^2+1}$$

$$\Rightarrow \frac{1}{s(s^2+1)}=\frac{A's^2+A'+B's^2+C's}{s(s^2+1)}$$

So, $A'+B'=0 \Rightarrow \underline{B'=-A'}$

$\underline{C'=0}$, $\underline{A'=1}$, so $B'=-1$.

Therefore,

$$\mathcal{L}^{-1}\left[\frac{1}{s(s^2+1)}\right](t) = \mathcal{L}^{-1}\left[\frac{1}{s}\right](t) - \mathcal{L}^{-1}\left[\frac{s}{s^2+1}\right](t)$$

$$= (1 - \cos t)$$

Hence, the solution of the initial value problem is:

$$y(t) = \left(\frac{1}{2}e^{-t} + \frac{1}{2}\sin t - \frac{1}{2}\cos t\right) + (1 - \cos t)$$
$$+ \sin t - \cos t$$

$$\Rightarrow \boxed{y(t) = \frac{1}{2}e^{-t} + \frac{3}{2}\sin t - \frac{5}{2}\cos t + 1}$$

Remark: Note that all the Laplace transforms that we've computed so far are rational functions (i.e. ratio of two polynomials).

Moreover, each of them are of the form

$$\frac{p(s)}{q(s)}, \text{ where } \deg(p) < \deg(q).$$

Therefore, the functions that arose as Laplace transforms tend to zero as $s \to +\infty$. In other words, functions that arose as Laplace transforms <u>decay</u> to 0 as $s \to +\infty$.

Meaning of Laplace transform:

In order to explain the meaning of the Laplace transform of a function, it will be useful to analyze the inverse Laplace transform.

If the Laplace transform of f exists and

$$\mathcal{L}[f] = F, \quad \text{then}$$

$$(*) \rightarrow \boxed{f(t) = \mathcal{L}^{-1}[F](t) = \frac{1}{2\pi i} \lim_{T \to +\infty} \int_{\gamma - iT}^{\gamma + iT} e^{st} F(s)\, ds}$$

where $[\gamma - iT, \gamma + iT]$ lies in the domain of definition of F. The above formula, which gives an explicit way of computing

f from $\mathcal{L}[f]$ [i.e. a function from its Laplace transform) is called the Bromwich integral.

Now recall the Fourier Series of a 2π-periodic function f:

$$\frac{a_0}{2} + \sum_{n=1}^{\infty}\left(a_n \cos(nx) + b_n \sin(nx)\right),$$

where $a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos(nx)\,dx$, $n \geq 0$

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(nx)\,dx.$$

Under suitable conditions, the Fourier series of f converges to f allowing us to view $f(x)$ as a "Sum" of simpler 2π-periodic functions $\sin(nx)$ and $\cos(nx)$.

The coefficients a_n, b_n can be thought of as weights corresponding to these "simpler building block functions" ($\sin(nx)$ and $\cos(nx)$) telling the importance of $\sin(nx)$ / $\cos(nx)$ in the decomposition of f. In ~~order~~ other words, a_n's and b_n's measure to what extent the function $f(x)$ "looks like" $\cos(nx)$ and $\sin(nx)$.

The philosophy behind the definition of Laplace transforms is quite similar. Instead of comparing the function with trigonometric functions, Laplace transform compares f with exponential functions e^{st}.

The other difference is that we want to express f as an integral of various e^{st}, not as an infinite sum (in this sense, the Laplace transform is a continuous decomposition as opposed to the Fourier transform, which is a discrete decomposition).

The Browwich integral formula (*) asserts that a function f can be written as a weighted continuous sum of the functions e^{sT} with associated weight $F(s) = \mathcal{L}[f](s) = \int_0^\infty e^{-st} f(t)\,dt.$

Hence, $F(s) = \mathcal{L}[f](S)$ measures the extent to which $f(t)$ resembles the function e^{st}, for any given S.

By the Bromwich integral formula, f is equal to:

$$\left(\text{Constant}\right) \int \begin{pmatrix} \text{importance} \\ \text{of } e^{st} \text{ in the} \\ \text{function } f(t) \end{pmatrix} \times \begin{pmatrix} \text{building block} \\ \text{function} \\ e^{st} \end{pmatrix}.$$

Therefore, f can be decomposed in terms of exponential functions e^{st} such that $\mathcal{L}[f](S)$ is the weight associated to e^{st}.

To conclude, let us mention a sufficient condition for existence of the Laplace transform of f.

Theorem: If $f : [0, +\infty) \to \mathbb{R}$ is continuous (or piecewise continuous) and if there exist $M, c, T \geq 0$ such that

$$|f(t)| \leq M e^{ct}, \text{ for all } t \geq T,$$

then $\mathcal{L}[f]$ exists. Moreover, we have an estimate:

$$|\mathcal{L}[f](s)| \leq \frac{M}{s-c}, \quad \text{for } s > c.$$

The proof is straightforward, the given "sub-exponential growth" condition assures convergence of the improper integral integral involved in the defition of Laplace transform.

However, there's a bonus; ~~we~~ it also follows from the above estimate that

$$\lim_{s \to +\infty} \mathcal{L}[f](s) = 0.$$

This means that the higher the value of s, the lesser the weight of e^{st} in the decomposition of $f(t)$. This is ~~comp~~ intuitively clear; if f has sub-exponential growth, then f barely "looks like" e^{st} when s is very large.

Convolutions and

More techniques to find inverse Laplace transforms

In the previous set of lecture notes, we saw how to compute Laplace transforms of polynomials, exponential functions and simple trigonometric functions. We also used the technique of partial fractions to compute inverse Laplace transforms of complicated rational functions. A combination of these methods along with a ~~co~~ formula for $\mathcal{L}[f^{(n)}]$ (Laplace transform of derivatives) allowed us to solve initial-value problems.

In these notes, we'll discuss a couple of more powerful techniques of computing inverse Laplace transforms.

Theorem:

$$\frac{d}{ds}\left(\mathcal{L}[f](s)\right) = -\mathcal{L}[tf(t)](s).$$

Proof: Recall that

$$\mathcal{L}[f](s) = \int_0^\infty e^{-st} f(t)\,dt$$

(differentiating under the integral sign) $\Rightarrow$

$$\frac{d}{ds}\mathcal{L}[f](s) = \int_0^\infty \frac{d}{ds}\left(e^{-st}\right) f(t)\,dt$$

$$\Rightarrow \frac{d}{ds}\left(\mathcal{L}[f](s)\right) = -\int_0^\infty e^{-st}\left(t f(t)\right) dt$$

$$\Rightarrow \frac{d}{ds}\left(\mathcal{L}[f](s)\right) = -\mathcal{L}[t f(t)](s).$$

As an application recall that

$$\mathcal{L}[e^{at}](s) = \frac{1}{s-a}$$

$$\Rightarrow \frac{d}{ds}\left(\frac{1}{s-a}\right) = -\mathcal{L}[t e^{at}](s)$$

$$\Rightarrow \mathcal{L}[t e^{at}](s) = \frac{1}{(s-a)^2}.$$

Repeating this argument, we get:

$$\mathcal{L}[t^n e^{at}](s) = \frac{n!}{(s-a)^{n+1}}, \quad n = 0, 1, \ldots, \quad s > a$$

Similarly, if $f(t) = \sin at$, then

$$\mathcal{L}[f](s) = \frac{a}{s^2+a^2}$$

$$\Rightarrow \frac{d}{ds}\left(\frac{a}{s^2+a^2}\right) = -\mathcal{L}[t\sin at](s)$$

$$\Rightarrow \mathcal{L}[t\sin at](s) = -\left(\frac{-a.(2s)}{(s^2+a^2)^2}\right)$$

$$\Rightarrow \mathcal{L}[t\sin at](s) = \frac{2as}{(s^2+a^2)^2}.$$

Putting $f(t) = \cos(at)$, we get:

$$\mathcal{L}[t\cos at](t) = \frac{s^2-a^2}{(s^2+a^2)^2}$$

Our next goal is to answer the following question:

If $\mathcal{L}[f](s) = F(s)$ and $\mathcal{L}[g](s) = G(s)$, what is $\mathcal{L}^{-1}[FG]$?

To answer this question, we'll have to define another strange object, whose meaning will be discussed later.

If f and g are integrable functions, we define their convolution $f*g$ as a function:

$$(f*g)(t) = \int_0^t f(u)\, g(t-u)\, du$$

We have the following theorem:

Theorem: If f, g are integrable and if $\mathcal{L}[f], \mathcal{L}[g]$ exist, then

i) $f*g = g*f$,

ii) $\mathcal{L}[f*g](s) = \mathcal{L}[f](s) \cdot \mathcal{L}[g](s)$.

Therefore, if $\mathcal{L}[f] = F$ and $\mathcal{L}[g] = G$,
then

$$\mathcal{L}^{-1}[FG] = f * g.$$

Computations:

1. Let, $f(t) = t$, $g(t) = \cos t$, $t \geq 0$.

Then, for $t \geq 0$,

$$(f*g)(t) = \int_0^t f(u)\, g(t-u)\, du$$

$$= \int_0^t u\cos(t-u)\, du$$

$$= \Big[-u\sin(t-u) + \int \sin(t-u)\, du\Big]_0^t$$

$$= \Big[-u\sin(t-u) + \cos(t-u)\Big]_0^t$$

$$= \big(-t\sin(t-t) + \cos(t-t)\big) - \big(-0\sin(t-0) + \cos(t-0)\big)$$

$$= 1 - (\cos t) = 1 - \cos t.$$

~~Hence,~~ $\cancel{\mathcal{L}^{-1}[t\cos}$

Now, $\mathcal{L}[t](s) = \dfrac{1}{s^2}$, $\mathcal{L}[\cos t](s) = \dfrac{s}{s^2+1}$

So, $\mathcal{L}^{-1}\left[\frac{1}{s^2}\cdot\frac{s}{s^2+1}\right](t) = (f * g)(t)$

$\Rightarrow \mathcal{L}^{-1}\left[\frac{1}{s(s^2+1)}\right](t) = (1-\cos t)$

[We could prove the above formula using partial fractions too]

(2) Now we want to find $\mathcal{L}^{-1}\left[\frac{e^{-2s}}{s(s^2+4)}\right]$

Recall that $\mathcal{L}[\sin 2t](s) = \frac{2}{s^2+4}$.

Using the formula

$$\mathcal{L}\left[\int_0^t f(u)\,du\right](s) = \frac{1}{s}\mathcal{L}[f](s)$$

with $f(t) = \frac{\sin 2t}{2}$, we get:

$$\mathcal{L}\left[\int_0^t \frac{\sin 2u}{2}\,du\right](s) = \frac{1}{s}\cdot\mathcal{L}\left[\frac{\sin 2t}{2}\right](s)$$

$$\Rightarrow \mathcal{L}\left[\left(-\frac{\cos 2u}{4}\right)_0^t\right](s) = \frac{1}{s}\cdot\frac{1}{s^2+4}$$

$$\Rightarrow \mathcal{L}\left[-\frac{\cos 2t}{4} + \frac{1}{4}\right](s) = \frac{1}{s(s^2+4)}$$

$$\Rightarrow \mathcal{L}^{-1}\left[\frac{1}{s(s^2+4)}\right](t) = \frac{1-\cos 2t}{4}$$

$$\Rightarrow \mathcal{L}^{-1}\left[\frac{1}{s(s^2+4)}\right](t) = \frac{1}{2}\sin^2(t) \rightarrow A$$

We now use another formula that's easy to prove :

B $\Rightarrow$ $\mathcal{L}[f(t-a)](s) = \mathcal{L}[f](s) \cdot e^{-as}$.

$(a>0)$

Therefore, by A and B, we've:

$$\mathcal{L}\left[\frac{1}{2}\sin^2(t-2)\right](s)$$

$$= \frac{e^{-2s}}{s(s^2+4)}$$

$$\Rightarrow \mathcal{L}^{-1}\left[\frac{e^{-2s}}{s(s^2+4)}\right](t) = \frac{1}{2}\sin^2(t-2).$$

3. We want to solve the initial value problem:

$$y' - y = \int_0^t y(u)\,du, \quad y(0) = 1$$

Let $y(t)$ be the solution; i.e.

$$y'(t) - y(t) = \int_0^t y(u)\,du$$

$$\Rightarrow \mathcal{L}[y'](s) - \mathcal{L}[y](s) = \mathcal{L}\left[\int_0^t y(u)\,du\right]$$

$$\Rightarrow s\mathcal{L}[y](s) - y(0) - \mathcal{L}[y](s) = \frac{1}{s}\mathcal{L}[y](s)$$

$$\Rightarrow \left(s - 1 - \frac{1}{s}\right)\mathcal{L}[y](s) = y(0)$$

$$\Rightarrow \frac{s^2 - s - 1}{s}\mathcal{L}[y](s) = 1$$

$$\Rightarrow \mathcal{L}[y](s) = \frac{s}{s^2 - s - 1} = \frac{s}{\left(s - \left(\frac{1+\sqrt{5}}{2}\right)\right)\left(s - \left(\frac{1-\sqrt{5}}{2}\right)\right)}$$

Let $\dfrac{s}{\left(s - \left(\frac{1+\sqrt{5}}{2}\right)\right)\left(s - \left(\frac{1-\sqrt{5}}{2}\right)\right)} = \dfrac{A}{\left(s - \left(\frac{1+\sqrt{5}}{2}\right)\right)} + \dfrac{B}{\left(s - \left(\frac{1-\sqrt{5}}{2}\right)\right)}$

$$\Rightarrow s = As + Bs - A\left(\frac{1-\sqrt{5}}{2}\right) - B\left(\frac{1+\sqrt{5}}{2}\right)$$

$$\Rightarrow \begin{cases} A + B = 1 \\ A\left(\frac{1-\sqrt{5}}{2}\right) + B\left(\frac{1+\sqrt{5}}{2}\right) = 0 \end{cases}$$

$$\Rightarrow B(\sqrt{5}+1) = A(\sqrt{5}-1), \quad B = 1 - A$$

$$\Rightarrow B = A\frac{(\sqrt{5}-1)}{\sqrt{5}+1}, \quad B = 1 - A$$

$$\Rightarrow B = A\frac{(\sqrt{5}-1)^2}{4}, \quad B = 1 - A$$

$$\Rightarrow 1 - A = A\left(\frac{6-2\sqrt{5}}{4}\right), \quad B = 1 - A$$

$$\Rightarrow 1 - A = A\left(\frac{3-\sqrt{5}}{2}\right), \quad B = 1 - A$$

$$\Rightarrow 2 - 2A = 3A - \sqrt{5}A, \quad B = 1 - A$$

$$\Rightarrow 2 = 5A - \sqrt{5}A, \quad B = 1 - A$$

$$\Rightarrow 2 = A\sqrt{5}(\sqrt{5}-1), \quad B = 1 - A$$

$$\Rightarrow A = \frac{2}{\sqrt{5}(\sqrt{5}-1)}, \quad B = 1 - A$$

$$\Rightarrow A = \frac{2(\sqrt{5}+1)\sqrt{5}}{4 \cdot 5}, \quad B = 1 - A$$

$$\Rightarrow A = \frac{\sqrt{5}(\sqrt{5}+1)}{10}, \quad B = 1 - A$$

$$\Rightarrow A = \frac{1}{2} + \frac{\sqrt{5}}{10}, \quad B = 1 - \frac{1}{2} - \frac{\sqrt{5}}{10}$$

$$\Rightarrow A = \left(\frac{1}{2} + \frac{\sqrt{5}}{10}\right), \quad B = \left(\frac{1}{2} - \frac{\sqrt{5}}{10}\right)$$

Thus, $\mathcal{L}[y(t)](s) = \left(\frac{1}{2} + \frac{\sqrt{5}}{10}\right) \cdot \frac{1}{\left(s - \left(\frac{1+\sqrt{5}}{2}\right)\right)}$

$$+ \left(\frac{1}{2} - \frac{\sqrt{5}}{10}\right) \cdot \left(\frac{1}{s - \left(\frac{1-\sqrt{5}}{2}\right)}\right)$$

$$\Rightarrow \mathcal{L}[y(t)](s) = \left(\frac{1}{2} + \frac{\sqrt{5}}{10}\right) \mathcal{L}\left[e^{\left(\frac{1+\sqrt{5}}{2}\right)t}\right](s) + \left(\frac{1}{2} - \frac{\sqrt{5}}{10}\right) \mathcal{L}\left[e^{\left(\frac{1-\sqrt{5}}{2}\right)t}\right](s)$$

$$\Rightarrow y(t) = \left(\frac{1}{2} + \frac{\sqrt{5}}{10}\right) e^{\left(\frac{1+\sqrt{5}}{2}\right)t} + \left(\frac{1}{2} - \frac{\sqrt{5}}{10}\right) e^{\left(\frac{1-\sqrt{5}}{2}\right)t}.$$

$\boxtimes$

What does convolution (or convolving two functions) mean?

Let us first define convolution for two integrable real functions $f, g: \mathbb{R} \to \mathbb{R}$ (note that f, g are defined on all of $\mathbb{R}$, not just on $[0, +\infty)$.

Then, $$\boxed{(f * g)(t) = \int_{-\infty}^{\infty} f(t-u)\, g(u)\, du}$$

In what follows, we will see that convolving a function f with a suitably chosen "peaking" sequence of functions φ_δ (as $\delta \to 0$) yields a smooth approximation of f.

Suppose that f is a fairly irregular (e.g. non-differentiable) function and we want to do "calculus" with f. Clearly, non-differentiability of f is a major obstruction to performing differential calculus on f. So our goal is to replace f by a differentiable function f_δ such that:

i) f_δ and f are very close (so we don't change f too much), and

ii) f_δ is differentiable.

One way of constructing such a function f_δ is to apply a moving average formula. For $\delta > 0$ small enough, we define:

$$f_\delta(t) = \frac{1}{2\delta}\int_{t-\delta}^{t+\delta} f(u)\,du \,.$$

Intuitively we are considering the average value of f locally near t. Since f_δ is defined as an integral,

the fundamental theorem of calculus proves that f_δ is differentiable and

$$(f_\delta)'(t) = \frac{f(t+\delta) - f(t-\delta)}{2\delta}.$$

Moreover, if we assume that f is continuous at t, then for any pre-assigned $\varepsilon > 0$, there exists a $\tilde{\delta} > 0$ such that

$$|f(t) - f(u)| < \varepsilon \,, \text{ whenever } |t-u| < \tilde{\delta}.$$

Thus, $|f_\delta(t) - f(t)|$

$$= \left| \frac{1}{2\delta}\int_{t-\delta}^{t+\delta} f(u)\,du - \frac{1}{2\delta}\int_{t-\delta}^{t+\delta} f(t)\,du \right|$$

$$= \frac{1}{2\delta} \left| \int_{t-\delta}^{t+\delta} (f(u) - f(t)) \, du \right|$$

$$\leq \frac{1}{2\delta} \int_{t-\delta}^{t+\delta} |f(u) - f(t)| \, du$$

$$< \frac{1}{2\delta} \int_{t-\delta}^{t+\delta} \varepsilon \, du \quad \left[\text{where } 0 < \delta < \tilde{\delta} \right]$$

$$= \varepsilon.$$

Thus, whenever $0 < \delta < \tilde{\delta}$, we have that:

$$|f_\delta(t) - f(t)| < \varepsilon.$$

This means that $\lim_{\delta \to 0} f_\delta(t) = f(t)$, $\forall t \in \mathbb{R}$

In other words, we've showed that our moving average formula produces a smooth approximation f_δ of f. [Here, smooth means differentiable.]

Let us define the function:

$$\varphi_\delta : \mathbb{R} \to \mathbb{R}$$

$$\varphi_\delta(u) = \begin{cases} \frac{1}{2\delta} & , \text{if } -\delta \leq u \leq \delta \\ 0 & , \text{otherwise} \end{cases}$$

for any $\delta > 0$.

The crucial observation is that f_δ can be written as the convolution of f with φ_δ:

$$(f * \varphi_\delta)(t) = \int_{-\infty}^{\infty} f(t-u)\,\varphi_\delta(u)\,du$$

$$= \int_{-\delta}^{\delta} \frac{1}{2\delta} f(t-u)\,du \quad \left[\begin{array}{l}\text{Since } \varphi_\delta \equiv 0 \\ \text{outside of} \\ [-\delta, \delta]\end{array}\right]$$

$$= \frac{1}{2\delta} \int_{-\delta}^{\delta} f(t-u)\,du$$

Substituting $t - u = x$, we get:

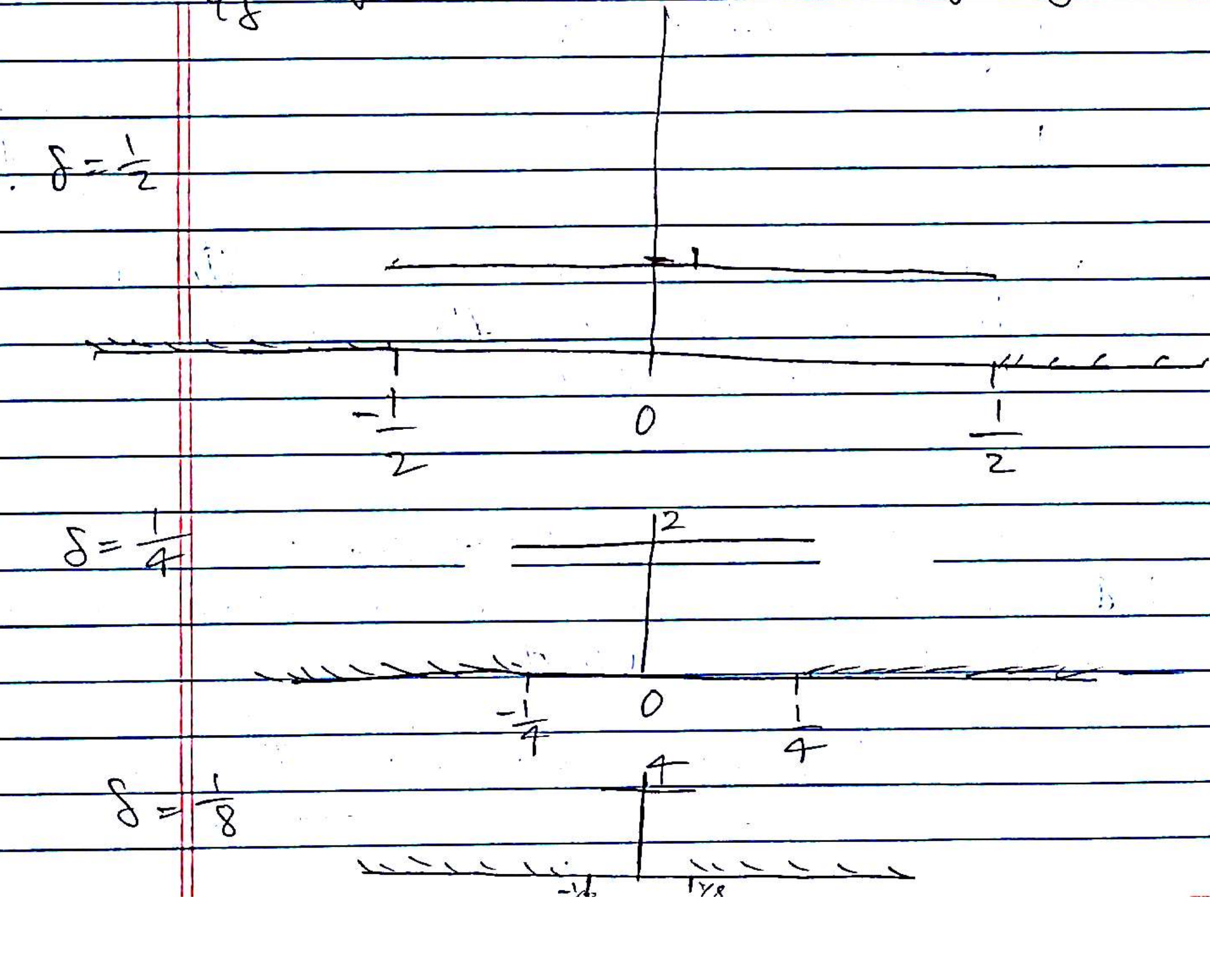
$$(f * \varphi_\delta)(t) = \frac{1}{2\delta} (-) \int_{t+\delta}^{t-\delta} f(x)\,dx$$

$$= \frac{1}{2\delta} \int_{t-\delta}^{t+\delta} f(x)\,dx = f_\delta(t)$$

Thus, $\boxed{f * \varphi_\delta = f_\delta}$; i.e.

Convolving f with the functions φ_δ yields a smooth approximation of f (essentially by blurring or taking a local average of f at each point).

The functions $\{\varphi_\delta\}_{\delta > 0}$ are also interesting in their own right. Here are the graphs of the functions φ_δ for some values of δ:



Evidently, as $\delta \to 0$, the function φ_δ is non-zero only on a very tiny interval $[-\delta, \delta]$ and the value of φ_δ there is very large $\left(= \frac{1}{2\delta}\right)$.

Thus, $\{\varphi_\delta\}_{\delta > 0}$ can be thought of as a <u>peaking</u> sequence of functions supported on increasingly smaller intervals.

As $\delta \to 0$, we've

$$\boxed{\lim_{\delta \to 0} \varphi_\delta(t) = \begin{cases} +\infty, & \text{if } t = 0 \\ 0, & \text{otherwise} \end{cases}}$$

This "function" is called the Dirac delta function, except it's not a function mathematically.

If such a "Dirac delta" function existed in a precise mathematical sense, it would have satisfied:

$$f * (\text{Dirac delta}) = f \; ; \text{ i.e.}$$

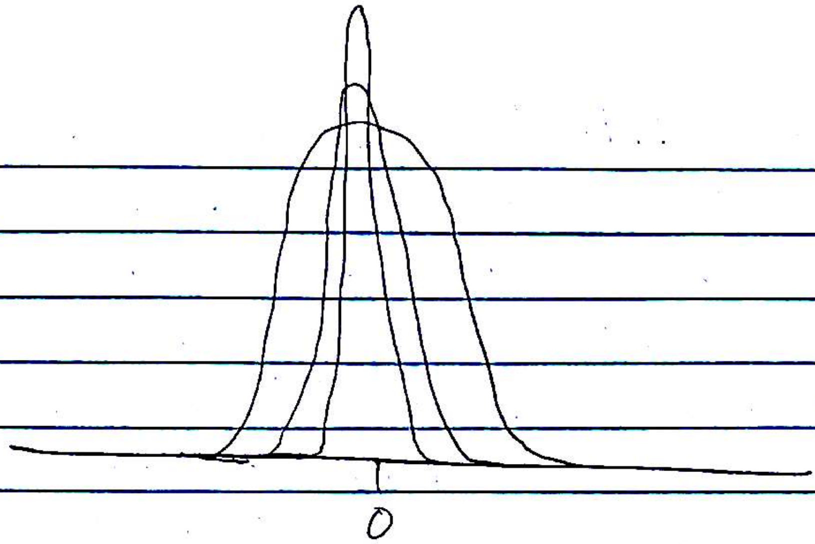
the "Dirac delta" function is a hypothetical identity element with respect to the convolution operation. Since it doesnot exist mathematically, we have to be happy with the "approximate identity" $\{\varphi_\delta\}$ which satisfy:

$$\boxed{f * \varphi_\delta \xrightarrow{\delta \to 0} f}$$

To conclude, let us mention that in practice, there is a better choice for the functions $\{\varphi_\delta\}$. We can choose $\{\varphi_\delta\}$ to be an infinitely differentiable sequence of functions with the same peaking property:

$$\varphi_\delta(x) = \begin{cases} \frac{1}{\delta^n} e^{-1/\left(1 - \frac{|x|^2}{\delta^2}\right)}, & |x| < \delta \\ 0, & |x| \geq \delta \end{cases}$$

One can check that each φ_δ is infinitely differentiable. Their graphs look like:



(as $\delta \to 0$)

(Support of φ_δ shrinks to 0 as $\delta \to 0$
and $\varphi_\delta(0) \to +\infty$ as $\delta \to 0$)

These functions are called bump functions.

They also satisfy:

$$\boxed{\lim_{\delta \to 0} (f * \varphi_\delta)(t) = f(t), \quad \forall t \in \mathbb{R}.}$$

Moreover, as φ_δ is infinitely diff., it follows that each $(f * \varphi_\delta)$ is also infinitely differentiable.

Hence with these choices of $\{\varphi_\delta\}$, we obtain an infinitely differentiable approximation of f via convolution.

Metric Spaces and Convergence of Sequences/Series

Defn: (metric)
Let X be a non-empty set. A function

$$d : X \times X \to \mathbb{R}^+ \cup \{0\}$$ is called

a metric on X if d satisfies the following three conditions:

i) $d(a,b) \geq 0$, and $d(a,b) = 0$ if and only if $a = b$;

ii) $d(a,b) = d(b,a)$ (symmetry);

iii) $d(a,b) \leq d(a,c) + d(c,b)$ (triangle inequality).

$\forall\, a, b, c \in X$.

Intuitively, we think of d as a "distance" function between points of X.

Defn (metric space)
A non-empty set X endowed with a metric d is called a metric space.

Examples: i) Let $X = \mathbb{R}$ or $\mathbb{C}$. We define $d(a,b) = |a-b|$, where $|\cdot|$ is the usual absolute value. Then (X,d) is a metric space.

ii) Let $X = \mathbb{R}^n$ or $\mathbb{C}^n$. We define

$$d((a_1, \ldots, a_n), (b_1, \ldots, b_n)) = \left(\sum_{i=1}^{n} |a_i - b_i|^2 \right)^{1/2}.$$

Then (X,d) is a metric space (here the triangle equality follows from the classical Cauchy-Schwarz inequality).

In what follows, we will see an important connection between inner products and metrics.

Let V be a vector space with the inner product $\langle \cdot \rangle$.

Recall that the norm of an element $u \in V$ is defined as:

$$\|u\| = \sqrt{\langle u, u \rangle}.$$

Now, V naturally becomes a metric space with respect to function $d : V \times V \to \mathbb{R}^+ \cup \{0\}$ defined as:

$$d(u_1, u_2) = \|u_1 - u_2\|$$
$$= \sqrt{\langle u_1 - u_2, u_1 - u_2 \rangle}.$$

(It's easy to show that d satisfies the defining properties of a metric)

Thus, every inner product space is a metric space.

For us, the most important example would be:

$$V = M_n(\mathbb{R})$$

We define an inner product on V:

$$\langle A, B \rangle = \text{trace}(A^* B), \; A, B \in V.$$

(Here, trace(A) of a square matrix A denotes the sum of its diagonal entries)

Again, it is straight-forward to check that this is an inner product.

Now, with respect to this inner product, we have

$$\|A\|^2 = tr(A^* A)$$

$$= \sum_{i,j=1}^{n} |a_{ij}|^2, \text{ where } A = (a_{ij})_{i,j=1}^{n}$$

Therefore, V becomes a metric space with the metric:

$$d(A, B) = \|A - B\|$$

$$= \sqrt{\sum_{i,j=1}^{n} |a_{ij} - b_{ij}|^2},$$

where $A = (a_{ij}), B = (b_{ij}) \in V$.

The fact that d is a metric follows from the definition of $\|\cdot\|$ and the fact that

$\underline{\|A+B\| \leq \|A\| + \|B\|}$, for all $A, B \in M_n(\mathbb{R})$

A metric space is the ideal setting to talk about convergence of sequences and series.

In a general metric space (X, d), we've the following definition:

Defn. (Convergence of sequences)

Let $\{x_K\}_K$ be a sequence in a metric space (X, d). We say that $\{x_n\}_n$ is convergent if there exists an element $x \in X$ such that (*)

$$d(x_K, x) \to 0 \text{ as } K \to \infty$$

(Think about it as the distance of x_K from x going to 0.)

In the metric space $(V = M_n(\mathbb{R}), d)$, with the metric defined above, convergence of a sequence of matrices $\{A_K\}_K$ to a matrix A means:

$$d(A_K, A) \to 0$$

(*) The precise definition is as follows:
for any $\varepsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that

$$d(x_K, x) < \varepsilon, \ \forall K \geq n_0.$$

i.e. $\|A_K - A\| \to 0$ as $K \to \infty$

i.e. $\sum_{i,j=1}^{n} |a_K^{i,j} - a^{i,j}|^2 \to 0$ as $K \to \infty$

i.e. $|a_K^{i,j} - a^{i,j}| \to 0$ as $K \to \infty$, for $i,j = 1, \ldots, n$

i.e. $\lim_{K \to \infty} a_K^{i,j} = a^{i,j}$, for all $i,j = 1, \ldots, n$.

Therefore, in $(M_n(\mathbb{R}), d)$, a sequence of matrices $\{A_K\}_K$ converges to A iff each entry of A_K converges to the corresponding entry of A.

Since $M_n(\mathbb{R})$ is a vector space, we can talk about sums of matrices. Thus, it makes sense to talk about convergence of infinite sums too.

Defn (Convergence of Series):

i) In $(M_n(\mathbb{R}), d)$, a series $\sum_{K=1}^{\infty} A_K$ is convergent iff the sequence of partial sums $\{S_K\}$, where $S_K = A_1 + \cdots + A_K$, is convergent.

ii) A series $\sum_{K=1}^{\infty} A_K$ is called **absolutely convergent** if the series of real numbers $\sum_{K=1}^{\infty} \|A_K\|$ converges

---

We need another important definition.

Defn (Cauchy sequences):

In a metric space (X, d), a sequence $\{x_K\}$ is called **Cauchy** if

$$d(x_i, x_j) \to 0 \text{ as } i, j \to \infty;$$

i.e. for every $\varepsilon > 0$, there exists an $n_0 \in \mathbb{N}$ such that

$$d(x_i, x_j) < \varepsilon \quad \forall\ i, j \geq n_0.$$

A Fundamental theorem:

Every Cauchy sequence of real numbers is convergent. Conversely every convergent sequence is Cauchy.

Remark: Here, it's important to consider the set of real numbers. For instance, $(\mathbb{Q}, d)$ (where $d(x,y) = |x-y|$) does not satisfy the above theorem. ~~above~~ In particular, there are Cauchy sequences of rational numbers that converge to irrationals; i.e. outside of $\mathbb{Q}$. Similarly, the open interval $(0,1)$ with the same absolute value metric does not satisfy the theorem either: the sequence $\left\{\frac{1}{n}\right\}$ is a Cauchy sequence in $(0,1)$, but the limit point 0 lies outside of $(0,1)$.

Intuitively, the terms of a cauchy sequence eventually get very close to each other. But they may or may not get arbitrary close to a limit point; i.e. the sequence may not converge.

Lemma: Let $\sum_{k=1}^{\infty} A_k$ be an absolutely convergent series in $M_n(\mathbb{R})$. Then the sequence of partial sums $\{S_k\}$, where $S_k = A_1 + \cdots + A_k$, is a Cauchy sequence in $M_n(\mathbb{R})$.

Proof: By assumption, the series $\sum_{k=1}^{\infty} \|A_k\|$ is convergent. Fix an $\varepsilon > 0$. Since $\sum_{k=1}^{\infty} \|A_k\|$ is convergent, the sequence of partial sums of this series is convergent. More precisely, the sequence $\{\tilde{S}_k\}$, where $\tilde{S}_k = \|A_1\| + \cdots + \|A_k\|$, is convergent.

Since every convergent sequence is Cauchy, we have that $\{\tilde{S}_k\}$ is a Cauchy sequence. Thus, there exists an $n_0 \in \mathbb{N}$ such that

$$d_{\mathbb{R}}(\tilde{S}_i, \tilde{S}_j) < \varepsilon \quad \forall\ i > j \geq n_0.$$

This means that:

$$\left|\widetilde{S}_i - \widetilde{S}_j\right| < \varepsilon, \quad \forall\, i > j \geq n_0$$

$$\Rightarrow \left| \left(\|A_1\| + \cdots + \|A_j\| + \|A_{j+1}\| + \cdots + \|A_i\|\right) - \left(\|A_1\| + \cdots + \|A_j\|\right) \right| < \varepsilon,$$

$$\forall\, i > j \geq n_0$$

$$\Rightarrow \underline{\|A_{j+1}\| + \cdots + \|A_i\| < \varepsilon, \quad \forall\, i > j \geq n_0} \quad \hookrightarrow (*)$$

But then, $d(S_i, S_j)$

$$= \|(A_1 + \cdots + A_j + A_{j+1} + \cdots + A_i) - (A_1 + \cdots + A_j)\|$$

$$= \|A_{j+1} + \cdots + A_i\|$$

$$\leq \|A_{j+1}\| + \cdots + \|A_i\| < \varepsilon, \quad \forall\, i > j \geq n_0$$

(by $(*)$)

Thus, $\{S_K\}$ is a Cauchy sequence in $M_n(\mathbb{R})$. ☒

We continue with the terminology of the previous lemma. (11)

The definition of the metric d on $M_n(\mathbb{R})$ implies that for a Cauchy sequence $\{S_k\}$ in $M_n(\mathbb{R})$, the (i,j)-the entries of $\{S_k\}$ form a Cauchy sequence in $\mathbb{R}$. Since every Cauchy sequence of real numbers has a limit, we've proved that:

<u>Lemma:</u> If the series $\sum_{k=1}^{\infty} A_k$ in $M_n(\mathbb{R})$ is absolutely convergent, then $\sum_{k=1}^{\infty} A_k$ is convergent.

In Section 2.1 of the book "Lie groups, Lie algebras, and representations" by Brian C. Hall, these analytic results are used to prove the existence of matrix exponentials.

---

14. *The connectedness of* $\mathsf{SL}(n;\mathbb{R})$. Using the polar decomposition of $\mathsf{SL}(n;\mathbb{R})$ (Proposition 1.16) and the connectedness of $\mathsf{SO}(n)$ (Exercise 13), show that $\mathsf{SL}(n;\mathbb{R})$ is connected.
    Hint: Recall that if P is a real, symmetric matrix, then there exists a real, orthogonal matrix R_1 such that $P = R_1 D R_1^{-1}$, where D is diagonal.
15. *The connectedness of* $\mathsf{GL}(n;\mathbb{R})^+$. Using the connectedness of $\mathsf{SL}(n;\mathbb{R})$ (Exercise 14) show that $\mathsf{GL}(n;\mathbb{R})^+$ is connected.
16. If R is an element of $\mathsf{SO}(3)$, show that R must have an eigenvector with eigenvalue 1.
    Hint: Since $\mathsf{SO}(3) \subset \mathsf{SU}(3)$, every (real or complex) eigenvalue of R must have absolute value 1.
17. Show that the set of translations is a normal subgroup of the Euclidean group $\mathsf{E}(n)$. Show that the quotient group $\mathsf{E}(n)/(\text{translations})$ is isomorphic to $\mathsf{O}(n)$. (Assume Proposition 1.5.)
18. Let a be an irrational real number. Show that the set of numbers of the form $e^{2\pi i n a}$, $n \in \mathbb{Z}$, is dense in S^1. (See Problem 1.)
19. Show that every continuous homomorphism Φ from $\mathbb{R}$ to S^1 is of the form $\Phi(x) = e^{iax}$ for some $a \in \mathbb{R}$. (This shows in particular that every such homomorphism is smooth.)
20. Suppose $G \subset \mathsf{GL}(n_1;\mathbb{C})$ and $H \subset \mathsf{GL}(n_2;\mathbb{C})$ are matrix Lie groups and that $\Phi : G \to H$ is a Lie group homomorphism. Then, the image of G under Φ is a subgroup of H and thus of $\mathsf{GL}(n_2;\mathbb{C})$. Is the image of G under Φ necessarily a matrix Lie group? Prove or give a counter-example.
21. Suppose P is a real, positive, symmetric matrix with determinant one. Show that there is a unique real, positive, symmetric matrix Q whose square is P.
    Hint: The existence of Q was discussed in Section 1.7. To prove uniqueness, consider two real, positive, symmetric square roots Q_1 and Q_2 of P and show that the eigenspaces of both Q_1 and Q_2 coincide with the eigenspaces of P.

2

Lie Algebras and the Exponential Mapping

2.1 The Matrix Exponential

The exponential of a matrix plays a crucial role in the theory of Lie groups. The exponential enters into the definition of the Lie algebra of a matrix Lie group (Section 2.5) and is the mechanism for passing information from the Lie algebra to the Lie group. Since many computations are done much more easily at the level of the Lie algebra, the exponential is indispensable in studying (matrix) Lie groups.

Let X be an $n \times n$ real or complex matrix. We wish to define the exponential of X, denoted e^X or $\exp X$, by the usual power series

$$e^X = \sum_{m=0}^{\infty} \frac{X^m}{m!}. \tag{2.1}$$

We will follow the convention of using letters such as X and Y for the variable in the matrix exponential.

Proposition 2.1. *For any $n \times n$ real or complex matrix X, the series (2.1) converges. The matrix exponential e^X is a continuous function of X.*

Before proving this, let us review some elementary analysis. Recall that the norm of a vector $x = (x_1, \ldots, x_n)$ in $\mathbb{C}^n$ is defined to be

$$\|x\| = \sqrt{\langle x, x\rangle} = \left(\sum_{k=1}^{n} |x_k|^2\right)^{1/2}.$$

We now define the norm of a matrix by thinking of the space $M_n(\mathbb{C})$ of all $n \times n$ matrices as $\mathbb{C}^{n^2}$. This means that we define

$$\|X\| = \left(\sum_{k,l=1}^{n} |X_{kl}|^2\right)^{1/2}. \tag{2.2}$$

This norm satisfies the inequalities

$$\|X+Y\| \leq \|X\| + \|Y\|, \tag{2.3}$$
$$\|XY\| \leq \|X\|\,\|Y\| \tag{2.4}$$

for all $X, Y \in M_n(\mathbb{C})$. The first of these inequalities is the triangle inequality and is a standard result from elementary analysis. The second of these inequalities follows from the Schwarz inequality (Exercise 1). If X_m is a sequence of matrices, then it is easy to see that X_m converges to a matrix X in the sense of Definition 1.3 if and only if $\|X_m - X\| \to 0$ as $m \to \infty$.

The norm (2.2) is called the **Hilbert–Schmidt** norm. There is another commonly used norm on the space of matrices, called the **operator norm**, whose definition is not relevant to us. It is easily shown that convergence in the Hilbert–Schmidt norm is equivalent to convergence in the operator norm. (This is true because we work with linear operators on the *finite-dimensional* space $\mathbb{C}^n$.) Furthermore, the operator norm also satisfies (2.3) and (2.4). Thus, it matters little whether we use the operator norm or the Hilbert–Schmidt norm.

A sequence X_m of matrices is said to be a **Cauchy sequence** if

$$\|X_m - X_l\| \to 0$$

as $m, l \to \infty$. Thinking of the space $M_n(\mathbb{C})$ of matrices as $\mathbb{C}^{n^2}$ and using a standard result from analysis, we have the following.

Proposition 2.2. *If X_m is a Cauchy sequence in $M_n(\mathbb{C})$, then there exists a unique matrix X such that X_m converges to X.*

That is, every Cauchy sequence in $M_n(\mathbb{C})$ converges.

Now, consider an infinite series whose terms are matrices:

$$X_0 + X_1 + X_2 + \cdots. \tag{2.5}$$

If

$$\sum_{m=0}^{\infty} \|X_m\| < \infty,$$

then the series (2.5) is said to **converge absolutely**. If a series converges absolutely, then it is not hard to show that the partial sums of the series form a Cauchy sequence, and, hence, by Proposition 2.2, the series converges. That is, any series which converges absolutely also converges. (The converse is not true; a series of matrices can converge without converging absolutely.)

We now turn to the proof of Proposition 2.1.

Proof. In light of (2.4), we see that

$$\|X^m\| \leq \|X\|^m,$$

and, hence,

$$\sum_{m=0}^{\infty} \left\| \frac{X^m}{m!} \right\| \leq \sum_{m=0}^{\infty} \frac{\|X\|^m}{m!} = e^{\|X\|} < \infty.$$

Thus, the series (2.1) converges absolutely, and so it converges.

To show continuity, note that since X^m is a continuous function of X, the partial sums of (2.1) are continuous. However, it is easy to see that (2.1) converges uniformly on each set of the form $\{\|X\| \leq R\}$, and so the sum is, again, continuous. $\square$

We now list some elementary properties of the matrix exponential.

Proposition 2.3. *Let X and Y be arbitrary $n \times n$ matrices. Then, we have the following:*

1. *$e^0 = I$.*
2. *$\left(e^X\right)^* = e^{X^*}$.*
3. *e^X is invertible and $\left(e^X\right)^{-1} = e^{-X}$.*
4. *$e^{(\alpha+\beta)X} = e^{\alpha X} e^{\beta X}$ for all α and β in $\mathbb{C}$.*
5. *If $XY = YX$, then $e^{X+Y} = e^X e^Y = e^Y e^X$.*
6. *If C is invertible, then $e^{CXC^{-1}} = Ce^X C^{-1}$.*
7. *$\left\|e^X\right\| \leq e^{\|X\|}$.*

It is *not* true in general that $e^{X+Y} = e^X e^Y$, although, by Point 4, it is true if X and Y commute. This is a crucial point, which we will consider in detail later. (See the Lie product formula in Section 2.4 and the Baker–Campbell–Hausdorff formula in Chapter 3.)

Proof. Point 1 is obvious and Point 2 follows from taking term-by-term adjoints of the series for e^X. Points 3 and 4 are special cases of Point 5. To verify Point 5, we simply multiply the power series term by term. (It is left to the reader to verify that this is legal.) Thus,

$$e^X e^Y = \left(I + X + \frac{X^2}{2!} + \cdots\right)\left(I + Y + \frac{Y^2}{2!} + \cdots\right).$$

Multiplying this out and collecting terms where the power of X plus the power of Y equals m, we get

$$e^X e^Y = \sum_{m=0}^{\infty} \sum_{k=0}^{m} \frac{X^k}{k!} \frac{Y^{m-k}}{(m-k)!} = \sum_{m=0}^{\infty} \frac{1}{m!} \sum_{k=0}^{m} \frac{m!}{k!(m-k)!} X^k Y^{m-k}. \tag{2.6}$$

Now, because (and *only* because) X and Y commute,

$$(X+Y)^m = \sum_{k=0}^{m} \frac{m!}{k!(m-k)!} X^k Y^{m-k},$$

and, thus, (2.6) becomes

$$e^X e^Y = \sum_{m=0}^{\infty} \frac{1}{m!}(X+Y)^m = e^{X+Y}.$$

To prove Point 6, simply note that

$$\left(CXC^{-1}\right)^m = CX^mC^{-1}$$

and, thus, the two sides of Point 6 are equal term by term.

Point 7 is evident from the proof of Proposition 2.1. □

Proposition 2.4. *Let X be a $n \times n$ complex matrix. Then, e^{tX} is a smooth curve in $M_n(\mathbb{C})$ and*

$$\frac{d}{dt}e^{tX} = Xe^{tX} = e^{tX}X.$$

In particular,

$$\left.\frac{d}{dt}e^{tX}\right|_{t=0} = X.$$

Proof. Differentiate the power series for e^{tX} term by term. (This is permitted because, for each i and j, $\left(e^{tX}\right)_{ij}$ is given by a convergent power series in t, and it is a standard theorem that one can differentiate power series term by term.) □

2.2 Computing the Exponential of a Matrix

We consider here methods for exponentiating general matrices. A special method for exponentiating 2×2 matrices is described in Exercises 6 and 7.

2.2.1 Case 1: X is diagonalizable

Suppose that X is an $n\times n$ real or complex matrix and that X is diagonalizable over $\mathbb{C}$; that is, there exists an invertible complex matrix C such that $X = CDC^{-1}$, with

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

It is easily verified that e^D is the diagonal matrix with eigenvalues $e^{\lambda_1}, \ldots, e^{\lambda_n}$, and so in light of Proposition 2.3, we have

$$e^X = C\begin{pmatrix} e^{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n} \end{pmatrix}C^{-1}.$$

Thus, if we can explicitly diagonalize X, we can explicitly compute e^X. Note that if X is real, then although C may be complex and the λ_k's may be complex, e^X must come out to be real, since each term in the series (2.1) is real.

For example, take

$$X = \begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix}.$$

Then, the eigenvectors of X are $\begin{pmatrix} 1 \\ i \end{pmatrix}$ and $\begin{pmatrix} i \\ 1 \end{pmatrix}$, with eigenvalues $-ia$ and ia, respectively. Thus, the invertible matrix

$$C = \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}$$

maps the basis vectors $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ to the eigenvectors of X, and so (check) $C^{-1}XC$ is a diagonal matrix D. Thus, $X = CDC^{-1}$ and

$$\begin{aligned} e^X &= \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}\begin{pmatrix} e^{-ia} & 0 \\ 0 & e^{ia} \end{pmatrix}\begin{pmatrix} 1/2 & -i/2 \\ -i/2 & 1/2 \end{pmatrix} \\ &= \begin{pmatrix} \cos a & -\sin a \\ \sin a & \cos a \end{pmatrix}. \end{aligned} \tag{2.7}$$

Note that explicitly if X (and hence a) is real, then e^X is real. See Exercise 6 for an alternative method of calculation.

2.2.2 Case 2: X is nilpotent

An $n \times n$ matrix X is said to be **nilpotent** if $X^m = 0$ for some positive integer m. Of course, if $X^m = 0$, then $X^l = 0$ for all $l > m$. In this case, the series (2.1), which defines e^X, terminates after the first m terms, and so can be computed explicitly.

For example, let us compute e^X, where

$$X = \begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix}.$$

Note that

$$X^2 = \begin{pmatrix} 0 & 0 & ac \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

and that $X^3 = 0$. Thus,

$$e^X = \begin{pmatrix} 1 & a & b+\frac{1}{2}ac \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}.$$

2.2.3 Case 3: X arbitrary

A general matrix X may be neither nilpotent nor diagonalizable. However, by Theorem B.6, every matrix X can be written (uniquely) in the form $X = S + N$, with S diagonalizable, N nilpotent, and $SN = NS$. Then, since N and S commute,

$$e^X = e^{S+N} = e^S e^N$$

and e^S and e^N can be computed as in the two previous subsections.

For example, take

$$X = \begin{pmatrix} a & b \\ 0 & a \end{pmatrix}.$$

Then,

$$X = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} + \begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix}.$$

The two terms clearly commute (since the first one is a multiple of the identity), and, so,

$$e^X = \begin{pmatrix} e^a & 0 \\ 0 & e^a \end{pmatrix}\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} e^a & e^a b \\ 0 & e^a \end{pmatrix}.$$

2.3 The Matrix Logarithm

We wish to define a matrix logarithm, which should be an inverse function (to the extent possible) to the matrix exponential. Let us recall the situation for the logarithm of complex numbers, in order to see what is reasonable to expect in the matrix case. Since e^z is never zero, only nonzero numbers can have a logarithm. Every nonzero complex number can be written as e^z for some z, but the z is not unique. There is no continuous way to define the logarithm on the set of all nonzero complex numbers. The situation for matrices is similar. For any $X \in M_n(\mathbb{C})$, e^X is invertible; therefore, only invertible matrices can possibly have a logarithm. We will see (Theorem 2.9) that every invertible matrix can be written as e^X, for some $X \in M_n(\mathbb{C})$. However, the X is not unique and there is no continuous way to define the matrix logarithm on the set of all invertible matrices.

The simplest way to define the matrix logarithm is by a power series. We recall how this works in the complex case.

Lemma 2.5. *The function*

$$\log z = \sum_{m=1}^{\infty} (-1)^{m+1} \frac{(z-1)^m}{m} \tag{2.8}$$

is defined and analytic in a circle of radius 1 about $z = 1$.

For all z with $|z - 1| < 1$,

$$e^{\log z} = z.$$

For all u with $|u| < \log 2$, $|e^u - 1| < 1$ and

$$\log e^u = u.$$

Proof. The usual logarithm for real, positive numbers satisfies

$$\frac{d}{dx}\log(1-x) = \frac{-1}{1-x} = -\left(1 + x + x^2 + \cdots\right)$$

for $|x| < 1$. Integrating term by term and noting that $\log 1 = 0$ gives

$$\log(1-x) = -\left(x + \tfrac{x^2}{2} + \tfrac{x^3}{3} + \cdots\right).$$

Taking $z = 1 - x$ (so that $x = 1 - z$), we have

$$\begin{aligned}\log z &= -\left((1-z) + \frac{(1-z)^2}{2} + \frac{(1-z)^3}{3} + \cdots\right) \\ &= \sum_{m=1}^{\infty} (-1)^{m+1} \frac{(z-1)^m}{m}.\end{aligned}$$

This series has radius of convergence 1 and defines a complex analytic function on the set $\{|z-1| < 1\}$, which coincides with the usual logarithm for real z in the interval $(0, 2)$. Now, $\exp(\log z) = z$ for $z \in (0,2)$, and by analyticity, this identity continues to hold on the whole set $\{|z-1| < 1\}$. (That is to say, the functions $z \to \exp(\log z)$ and $z \to z$ are both complex analytic functions and they agree on the interval $(0,2)$; therefore they must agree on the whole disk $\{|z-1| < 1\}$.)

On the other hand, if $|u| < \log 2$, then

$$|e^u - 1| = \left|u + \tfrac{u^2}{2!} + \cdots\right| \le |u| + \frac{|u|^2}{2!} + \cdots = e^{|u|} - 1 < 1.$$

Thus, $\log(\exp u)$ makes sense for all such u. Since $\log(\exp u) = u$ for real u with $|u| < \log 2$, it follows by analyticity that $\log(\exp u) = u$ for all complex numbers with $|u| < \log 2$. □

Definition 2.6. *For any $n \times n$ matrix A, define $\log A$ by*

$$\log A = \sum_{m=1}^{\infty} (-1)^{m+1} \frac{(A-I)^m}{m} \tag{2.9}$$

whenever the series converges.

Since the complex-valued series (2.8) has radius of convergence 1 and since $\|(A-I)^m\| \le \|A-I\|^m$, the matrix-valued series (2.9) will converge

Jordan Canonical form, nilpotent-diagonalizable decomposition, and Computation of matrix exponentials

In last week's lecture notes, we have seen the definition of matrix exponential:

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!}, \quad A \in M_n(\mathbb{C}).$$

We will now discuss effective methods of computing e^A.

Case-I: A is diagonalizable:

Let A be a diagonalizable matrix. Then there exists an invertible matrix $P \in M_n(\mathbb{C})$ such that

$$P^{-1}AP = D, \text{ where } D \text{ is a diagonal matrix.}$$

Then, $e^A = e^{PDP^{-1}} = Pe^D P^{-1}$.

clearly, if the eigenvalues of A are $\{\lambda_1, \dots, \lambda_n\}$ (not necessarily distinct), then we have:

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Then, it is easy to compute that

$$D^k = \begin{pmatrix} \lambda_1^k & & 0 \\ & \ddots & \\ 0 & & \lambda_n^k \end{pmatrix}, \quad \text{for all } k \geq 0.$$

Hence, e^D

$$= \sum_{k=0}^{\infty} \frac{D^k}{k!} = \begin{pmatrix} \sum_{k=0}^{\infty} \frac{\lambda_1^k}{k!} & & 0 \\ & \ddots & \\ 0 & & \sum_{k=0}^{\infty} \frac{\lambda_n^k}{k!} \end{pmatrix}$$

$$= \begin{pmatrix} e^{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n} \end{pmatrix}.$$

Finally, $$e^A = P\begin{pmatrix} e^{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n} \end{pmatrix} P^{-1} \rightarrow (1)$$ (3)

Since P is the matrix of linearly independent ~~e~~eigenvectors of A (arranged in the correct order), P is explicitly computable.

Hence, (1) gives a formula for computing e^A, when A is diagonalizable.

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

The eigenvalues of A are 5 and -1,

and $$A\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 5\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad A\begin{pmatrix} 1 \\ -1 \end{pmatrix} = -1\begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

[Compare pages 17-20 of the lecture notes on eigenvalues/eigenvectors.]

Setting $P = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, we see that

$$PAP^{-1} = \begin{pmatrix} 5 & 0 \\ 0 & -1 \end{pmatrix}$$

i.e. $A = P^{-1} \begin{pmatrix} 5 & 0 \\ 0 & -1 \end{pmatrix} P$.

$$\text{So, } e^A = P^{-1} \begin{pmatrix} e^5 & 0 \\ 0 & e^{-1} \end{pmatrix} P$$

$$= \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^{-1} \begin{pmatrix} e^5 & 0 \\ 0 & e^{-1} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} e^5 & e^5 \\ e^{-1} & -e^{-1} \end{pmatrix}$$

$$= \begin{pmatrix} \frac{e^5 + e^{-1}}{2} & \frac{e^5 - e^{-1}}{2} \\ \frac{e^5 - e^{-1}}{2} & \frac{e^5 + e^{-1}}{2} \end{pmatrix}$$

$\boxtimes$

Case - II : Nilpotent matrices

A matrix A is called nilpotent if there exists some $K \in \mathbb{N}$ such that

$$A^K = O.$$

For a nilpotent matrix A, we have:

$$e^A = \sum_{j=0}^{\infty} \frac{A^j}{j!} = \sum_{j=0}^{K-1} \frac{A^j}{j}$$

as the subsequence terms are all 0.

So, e^A is simply a finite sum.

Example:

$$A = \begin{pmatrix} 5 & -3 & 2 \\ 15 & -9 & 6 \\ 10 & -6 & 4 \end{pmatrix}.$$

A simple computation shows that

$$A^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad \text{So, } e^A = I + A = \begin{pmatrix} 6 & -3 & 2 \\ 15 & -8 & 6 \\ 10 & -6 & 5 \end{pmatrix}.$$

We now come to the general case. An arbitrary matrix is neither diagonalizable, nor nilpotent. However, one can write any matrix A as:

$$A = D + N,$$ where D is diagonalizable, N is nilpotent, and

$$DN = ND.$$

This is called the N-D decomposition of A.

Then, e^A

$$= e^{D+N}$$

$$= e^D \cdot e^N \quad (\text{Since } DN = ND)$$

As we already know how to compute e^D and e^N, we can easily compute e^A.

In the rest of the lecture notes, we'll see how to compute the N-D decomposition of a matrix A. The general theorem that guarantees the existence of such a decomposition follows from the existence of the Jordan canonical form of a matrix A. Discussing the general theory of Jordan canonical forms is beyond the scope of this course, so we only focus on computations. To make life easier, we will only consider 2×2 and 3×3 matrices.

Example: e^A of a non-diagonalizable 2×2 matrix A: (8)

Let $A = \begin{pmatrix} \frac{5}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{3}{2} \end{pmatrix}$

The characteristic equation of A is:

$$\det(\lambda I - A) = 0$$

$$\Rightarrow \det\begin{pmatrix} \lambda - \frac{5}{2} & \frac{1}{2} \\ -\frac{1}{2} & \lambda - \frac{3}{2} \end{pmatrix} = 0$$

$$\Rightarrow \lambda^2 - 4\lambda + 4 = 0 \Rightarrow \underline{(\lambda-2)^2 = 0}$$

So, 2 is the only eigenvalue of A.

Let $\begin{pmatrix} a \\ b \end{pmatrix}$ be an eigenvector corresponding to 2.

Then $A\begin{pmatrix} a \\ b \end{pmatrix} = 2\begin{pmatrix} a \\ b \end{pmatrix}$

$$\Rightarrow \begin{pmatrix} \frac{5a}{2} - \frac{b}{2} \\ \frac{a}{2} + \frac{3b}{2} \end{pmatrix} = \begin{pmatrix} 2a \\ 2b \end{pmatrix}$$

$\Rightarrow \underline{a = b}$.

Hence, the eigenspace of 2 is:

$$\left\{\begin{pmatrix} a \\ a \end{pmatrix} : a \in \mathbb{C}\right\} = \operatorname{span}\left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}.$$

Therefore, A has a unique 1-dimensional eigenspace ; i.e. only one linearly independent eigenvector. It follows that A is not diagonalizable.

we set $V_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, and seek a vector V_2 such that

$$(A - 2I) V_2 = V_1$$

$\boxed{\text{For such a } V_2 \text{, we'll have } (A-2I)^2 V_2 = 0. \text{ So, } V_2 \text{ is called a generalized eigenvector.}}$

If $V_2 = \begin{pmatrix} c \\ d \end{pmatrix}$, then,

$$\begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$\Rightarrow$ $\underline{c - d = 2}$. $\rightarrow$ Unique condition on c and d

Let's choose $c = 2,\ d = 0$.

Then, $V_2 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$.

~~Define~~ By construction, we have

$$\left.\begin{aligned} AV_1 &= 2V_1, \\ AV_2 &= V_1 + 2V_2. \end{aligned}\right\}$$

Setting $P = (V_1 \ \ V_2) = \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix}$,

we get:

$$AP = P\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$$

$$\Rightarrow A = P\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}P^{-1}$$

$$\Rightarrow A = \underbrace{P\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}P^{-1}}_{\text{diagonalizable}} + \underbrace{P\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}P^{-1}}_{\text{nilPotent.}}$$

clearly, $e^{\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}} = \begin{pmatrix} e^2 & 0 \\ 0 & e^2 \end{pmatrix}$,

$$e^{\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Hence, $e^A = \left(P\begin{pmatrix} e^2 & 0 \\ 0 & e^2 \end{pmatrix}P^{-1}\right)\cdot\left(P\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}P^{-1}\right)$

$$= P\begin{pmatrix} e^2 & 0 \\ 0 & e^2 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}P^{-1}$$

$$= P\begin{pmatrix} e^2 & e^2 \\ 0 & e^2 \end{pmatrix}P^{-1},$$

where, $P = \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix}$.

$\boxtimes$

Example: exponential of a 3×3 non-diagonalizable matrix :

Let $A = \begin{pmatrix} 2 & 2 & 3 \\ 1 & 3 & 3 \\ -1 & -2 & -2 \end{pmatrix}$

The eigenvalues of A are roots of:

$\det(\lambda I - A) = 0$

$\Leftrightarrow \quad x^3 - 3x^2 + 3x - 1 = 0$

$\Leftrightarrow \quad (x-1)^3 = 0.$

Hence, 1 is the unique eigenvalue of A.

Let $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$ be an eigenvector of A corresponding to the eigenvalue 1.

Then $A \begin{pmatrix} a \\ b \\ c \end{pmatrix} = 1 \begin{pmatrix} a \\ b \\ c \end{pmatrix}$

$$\Rightarrow \begin{pmatrix} 2a+2b+3c \\ a+3b+3c \\ -a-2b-2c \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

$\Rightarrow \quad a+2b+3c=0 \quad \rightarrow$ unique condition

So, eigenspace of A corresponding to 1

$$= \left\{ \begin{pmatrix} -2b-3c \\ b \\ c \end{pmatrix} : \; b, c \in \mathbb{R} \right\}$$

Clearly, this eigenspace has dimension 2 and hence, A is not diagonalizable (we only get 2 linearly independent eigenvectors, but diagonalizability would require 3.)

In order to express A in its Jordan canonical form $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$,

we need vectors V_1, V_2, V_3 such that

$$(A-I)V_1 = 0, \; (A-I)V_2 = 0,$$
$$(A-I)V_3 = V_2.$$

In particular V_2 must belong to $\ker(A-I) \cap \text{Image}(A-I)$.

Now, Image $(A-I)$

$$= \text{Image}\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ -1 & -2 & -3 \end{pmatrix}$$

$$= \left\{ \begin{pmatrix} p+2q+3r \\ p+2q+3r \\ -(p+2q+3r) \end{pmatrix} : p, q, r \in \mathbb{R} \right\}$$

$$= \text{Span}\left\{ \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \right\}$$

Also, $(A-I)\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ -1 & -2 & -3 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$

So, $\ker(A-I) \cap \text{Image}(A-I) = \text{Span}\left\{ \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \right\}$

We choose $V_2 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$.

Then, V_3 has to be chosen such that

$(A-I)V_3 = V_2$.

If $V_3 = \begin{pmatrix} p \\ q \\ r \end{pmatrix}$, then we have:

$$\Rightarrow (p+2q+3r)\begin{pmatrix}1\\1\\-1\end{pmatrix} = \begin{pmatrix}1\\1\\-1\end{pmatrix}$$

$$\Rightarrow p+2q+3r = 1$$

Let $q = r = 0$, $p = 1$.

so we set $V_3 = \begin{pmatrix}1\\0\\0\end{pmatrix}$.

Finally, $V_1 \in \ker(A-I)$

$$\Rightarrow V_1 = \begin{pmatrix}-2b-3c\\b\\c\end{pmatrix}, \text{ for some } b, c \in \mathbb{R}$$

Let, $c = 0$, $b = 1$, (this choice is guided by our goal: V_1, V_2, V_3 have to be linearly indep.)

So, $V_1 = \begin{pmatrix}-2\\1\\0\end{pmatrix}$.

Thus, with respect to the basis $\left\{ V_1 = \begin{pmatrix}-2\\1\\0\end{pmatrix}, V_2 = \begin{pmatrix}1\\1\\-1\end{pmatrix}, V_3 = \begin{pmatrix}1\\0\\0\end{pmatrix} \right\}$ we're:

$$AV_1 = V_1, \quad AV_2 = V_2, \quad AV_3 = V_2 + V_3.$$

So,

Setting $P = \begin{pmatrix} v_1 & v_2 & v_3 \end{pmatrix} = \begin{pmatrix} -2 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & -1 & 0 \end{pmatrix}$,

we have that $\boxed{\text{the matrix } A \text{ is represented by } \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \text{ wrt } \{v_1, v_2, v_3\}}$.

$$AP = P \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\Rightarrow A = P \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} P^{-1}.$$

So, $e^A = P \exp \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} P^{-1}$

$$= P \exp\left(\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \right) P^{-1}$$

$$= P \exp \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \exp \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} P^{-1}$$

$$= P\begin{pmatrix} e & 0 & 0 \\ 0 & e & 0 \\ 0 & 0 & e \end{pmatrix}\left(I + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}\right)P^{-1}$$

$$= P\begin{pmatrix} e & 0 & 0 \\ 0 & e & 0 \\ 0 & 0 & e \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}P^{-1}$$

$$= P\begin{pmatrix} e & 0 & 0 \\ 0 & e & e \\ 0 & 0 & e \end{pmatrix}P^{-1}$$

where $P = \begin{pmatrix} -2 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & -1 & 0 \end{pmatrix}.$

5.7 Nonhomogeneous Linear Systems

In Section 3.5 we exhibited two techniques for finding a single particular solution of a single nonhomogeneous nth-order linear differential equation—the method of undetermined coefficients and the method of variation of parameters. Each of these may be generalized to nonhomogeneous linear systems. In a linear system modeling a physical situation, nonhomogeneous terms typically correspond to external influences, such as the inflow of liquid to a cascade of brine tanks or an external force acting on a mass-and-spring system.

Given the nonhomogeneous first-order linear system

$$\mathbf{x}' = \mathbf{A}\mathbf{x} + \mathbf{f}(t) \tag{1}$$

where $\mathbf{A}$ is an $n \times n$ constant matrix and the "nonhomogeneous term" $\mathbf{f}(t)$ is a given continuous vector-valued function, we know from Theorem 4 of Section 5.1 that a general solution of Eq. (1) has the form

$$\mathbf{x}(t) = \mathbf{x}_c(t) + \mathbf{x}_p(t), \tag{2}$$

where

- $\mathbf{x}_c(t) = c_1\mathbf{x}_1(t) + c_2\mathbf{x}_2(t) + \cdots + c_n\mathbf{x}_n(t)$ is a general solution of the associated *homogeneous* system $\mathbf{x}' = \mathbf{A}\mathbf{x}$, and
- $\mathbf{x}_p(t)$ is a single particular solution of the original nonhomogeneous system in (1).

Preceding sections have dealt with $\mathbf{x}_c(t)$, so our task now is to find $\mathbf{x}_p(t)$.

Undetermined Coefficients

First we suppose that the nonhomogeneous term $\mathbf{f}(t)$ in (1) is a linear combination (with constant vector coefficients) of products of polynomials, exponential functions, and sines and cosines. Then the method of undetermined coefficients for systems is essentially the same as for a single linear differential equation. We make an intelligent guess as to the *general form* of a particular solution $\mathbf{x}_p$, then attempt to determine the coefficients in $\mathbf{x}_p$ by substitution in Eq. (1). Moreover, the choice of this general form is essentially the same as in the case of a single equation (discussed in Section 3.5); we modify it only by using undetermined *vector* coefficients rather than undetermined scalars. We will therefore confine the present discussion to illustrative examples.

Example 1 Find a particular solution of the nonhomogeneous system

$$\mathbf{x}' = \begin{bmatrix} 3 & 2 \\ 7 & 5 \end{bmatrix}\mathbf{x} + \begin{bmatrix} 3 \\ 2t \end{bmatrix}. \tag{3}$$

Solution The nonhomogeneous term $\mathbf{f} = \begin{bmatrix} 3 & 2t \end{bmatrix}^T$ is linear, so it is reasonable to select a linear trial particular solution of the form

$$\mathbf{x}_p(t) = \mathbf{a}t + \mathbf{b} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} t + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}. \tag{4}$$

Upon substitution of $\mathbf{x} = \mathbf{x}_p$ in Eq. (3), we get

$$\begin{aligned}
\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} &= \begin{bmatrix} 3 & 2 \\ 7 & 5 \end{bmatrix}\begin{bmatrix} a_1t + b_1 \\ a_2t + b_2 \end{bmatrix} + \begin{bmatrix} 3 \\ 2t \end{bmatrix} \\
&= \begin{bmatrix} 3a_1 + 2a_2 \\ 7a_1 + 5a_2 + 2 \end{bmatrix} t + \begin{bmatrix} 3b_1 + 2b_2 + 3 \\ 7b_1 + 5b_2 \end{bmatrix}.
\end{aligned}$$

We equate the coefficients of t and the constant terms (in both x_1- and x_2-compon-ents) and thereby obtain the equations

$$\begin{aligned} 3a_1 + 2a_2 &= 0, \\ 7a_1 + 5a_2 + 2 &= 0, \\ 3b_1 + 2b_2 + 3 &= a_1, \\ 7b_1 + 5b_2 &= a_2. \end{aligned} \tag{5}$$

We solve the first two equations in (5) for $a_1 = 4$ and $a_2 = -6$. With these values we can then solve the last two equations in (5) for $b_1 = 17$ and $b_2 = -25$. Substitution of these coefficients in Eq. (4) gives the particular solution $\mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^T$ of (3) described in scalar form by

$$\begin{aligned} x_1(t) &= 4t + 17, \\ x_2(t) &= -6t - 25. \end{aligned}$$

Example 2

Figure 5.7.1 shows the system of three brine tanks investigated in Example 2 of Section 5.2. The volumes of the three tanks are $V_1 = 20$, $V_2 = 40$, and $V_3 = 50$ (gal), and the common flow rate is $r = 10$ (gal/min). Suppose that all three tanks contain fresh water initially, but that the inflow to tank 1 is brine containing 2 pounds of salt per gallon, so that 20 pounds of salt flow into tank 1 per minute. Referring to Eq. (18) in Section 5.2, we see that the vector $\mathbf{x}(t) = \begin{bmatrix} x_1(t) & x_2(t) & x_3(t) \end{bmatrix}^T$ of amounts of salt (in pounds) in the three tanks at time t satisfies the nonhomogeneous initial value problem

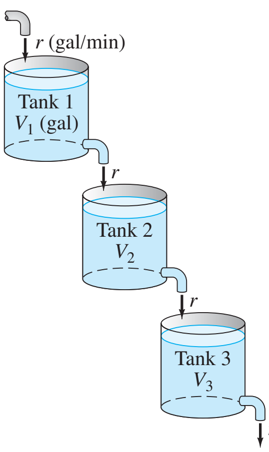


FIGURE 5.7.1. The three brine tanks of Example 2.

$$\frac{d\mathbf{x}}{dt} = \begin{bmatrix} -0.5 & 0 & 0 \\ 0.5 & -0.25 & 0 \\ 0 & 0.25 & -0.2 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 20 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{x}(0) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \tag{6}$$

The nonhomogeneous term $\mathbf{f} = \begin{bmatrix} 20 & 0 & 0 \end{bmatrix}^T$ here corresponds to the 20 lb/min inflow of salt to tank 1, with no (external) inflow of salt into tanks 2 and 3.

Because the nonhomogeneous term is constant, we naturally select a constant trial function $\mathbf{x}_p = \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix}^T$, for which $\mathbf{x}'_p \equiv \mathbf{0}$. Then substitution of $\mathbf{x} = \mathbf{x}_p$ in (6) yields the system

$$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -0.5 & 0 & 0 \\ 0.5 & -0.25 & 0 \\ 0 & 0.25 & -0.2 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} + \begin{bmatrix} 20 \\ 0 \\ 0 \end{bmatrix}$$

that we readily solve for $a_1 = 40$, $a_2 = 80$, and $a_3 = 100$ in turn. Thus our particular solution is $\mathbf{x}_p(t) = \begin{bmatrix} 40 & 80 & 100 \end{bmatrix}^T$.

In Example 2 of Section 5.2 we found the general solution

$$\mathbf{x}_c(t) = c_1 \begin{bmatrix} 3 \\ -6 \\ 5 \end{bmatrix} e^{-t/2} + c_2 \begin{bmatrix} 0 \\ 1 \\ -5 \end{bmatrix} e^{-t/4} + c_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} e^{-t/5}$$

of the associated homogeneous system, so a general solution $\mathbf{x} = \mathbf{x}_c + \mathbf{x}_p$ of the nonhomogeneous system in (6) is given by

$$\mathbf{x}(t) = c_1 \begin{bmatrix} 3 \\ -6 \\ 5 \end{bmatrix} e^{-t/2} + c_2 \begin{bmatrix} 0 \\ 1 \\ -5 \end{bmatrix} e^{-t/4} + c_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} e^{-t/5} + \begin{bmatrix} 40 \\ 80 \\ 100 \end{bmatrix}. \tag{7}$$

When we apply the zero initial conditions in (6), we get the scalar equations

$$\begin{aligned} 3c_1 \qquad\qquad\quad + 40 &= 0, \\ -6c_1 + c_2 \qquad + 80 &= 0, \\ 5c_1 - 5c_2 + c_3 + 100 &= 0 \end{aligned}$$

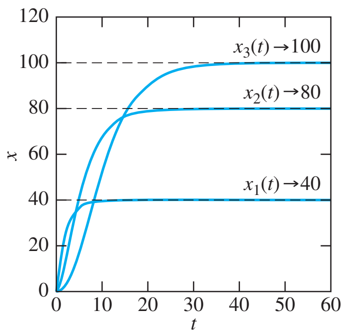


FIGURE 5.7.2. Solution curves for the amount of salt defined in (8).

that are readily solved for $c_1 = -\frac{40}{3}$, $c_2 = -160$, and $c_3 = -\frac{2500}{3}$. Substituting these coefficients in Eq. (7), we find that the amounts of salt in the three tanks at time t are given by

$$\begin{aligned} x_1(t) &= 40 - 40e^{-t/2}, \\ x_2(t) &= 80 + 80e^{-t/2} - 160e^{-t/4}, \\ x_3(t) &= 100 + \tfrac{100}{3}\left(-2e^{-t/2} + 24e^{-t/4} - 25e^{-t/5}\right). \end{aligned} \tag{8}$$

As illustrated in Fig. 5.7.2, we see the salt in each of the three tanks approaching, as $t \to +\infty$, a uniform density of 2 lb/gal—the same as the salt density in the inflow to tank 1. ■

In the case of duplicate expressions in the complementary function and the nonhomogeneous terms, there is one difference between the method of undetermined coefficients for systems and for single equations (Rule 2 in Section 3.5). For a system, the usual first choice for a trial solution must be multiplied not only by the smallest integral power of t that will eliminate duplication, but also by all lower (nonnegative integral) powers of t as well, and all the resulting terms must be included in the trial solution.

Example 3 Consider the nonhomogeneous system

$$\mathbf{x}' = \begin{bmatrix} 4 & 2 \\ 3 & -1 \end{bmatrix}\mathbf{x} - \begin{bmatrix} 15 \\ 4 \end{bmatrix} te^{-2t}. \tag{9}$$

In Example 1 of Section 5.2 we found the solution

$$\mathbf{x}_c(t) = c_1\begin{bmatrix} 1 \\ -3 \end{bmatrix}e^{-2t} + c_2\begin{bmatrix} 2 \\ 1 \end{bmatrix}e^{5t} \tag{10}$$

of the associated homogeneous system. A preliminary trial solution $\mathbf{x}_p(t) = \mathbf{a}te^{-2t} + \mathbf{b}e^{-2t}$ exhibits duplication with the complementary function in (10). We would therefore select

$$\mathbf{x}_p(t) = \mathbf{a}t^2e^{-2t} + \mathbf{b}te^{-2t} + \mathbf{c}e^{-2t}$$

as our trial solution, and we would then have six scalar coefficients to determine. It is simpler to use the method of variation of parameters, our next topic. ■

Variation of Parameters

Recall from Section 3.5 that the method of variation of parameters may be applied to a linear differential equation with variable coefficients and is not restricted to nonhomogeneous terms involving only polynomials, exponentials, and sinusoidal functions. The method of variation of parameters for systems enjoys the same flexibility and has a concise matrix formulation that is convenient for both practical and theoretical purposes.

We want to find a particular solution $\mathbf{x}_p$ of the nonhomogeneous linear system

➤ $$\mathbf{x}' = \mathbf{P}(t)\mathbf{x} + \mathbf{f}(t), \tag{11}$$

given that we have already found a general solution

➤ $$\mathbf{x}_c(t) = c_1\mathbf{x}_1(t) + c_2\mathbf{x}_2(t) + \cdots + c_n\mathbf{x}_n(t) \tag{12}$$

of the associated homogeneous system

➤ $$\mathbf{x}' = \mathbf{P}(t)\mathbf{x}. \tag{13}$$

We first use the fundamental matrix $\mathbf{\Phi}(t)$ with column vectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ to rewrite the complementary function in (12) as

$$\mathbf{x}_c(t) = \mathbf{\Phi}(t)\mathbf{c}, \tag{14}$$

where $\mathbf{c}$ denotes the column vector whose entries are the coefficients $c_1, c_2, \dots, c_n$. Our idea is to replace the vector "parameter" $\mathbf{c}$ with a variable vector $\mathbf{u}(t)$. Thus we seek a particular solution of the form

$$\mathbf{x}_p(t) = \mathbf{\Phi}(t)\mathbf{u}(t). \tag{15}$$

We must determine $\mathbf{u}(t)$ so that $\mathbf{x}_p$ does, indeed, satisfy Eq. (11).

The derivative of $\mathbf{x}_p(t)$ is (by the product rule)

$$\mathbf{x}_p'(t) = \mathbf{\Phi}'(t)\mathbf{u}(t) + \mathbf{\Phi}(t)\mathbf{u}'(t). \tag{16}$$

Hence substitution of Eqs. (15) and (16) in (11) yields

$$\mathbf{\Phi}'(t)\mathbf{u}(t) + \mathbf{\Phi}(t)\mathbf{u}'(t) = \mathbf{P}(t)\mathbf{\Phi}(t)\mathbf{u}(t) + \mathbf{f}(t). \tag{17}$$

But

$$\mathbf{\Phi}'(t) = \mathbf{P}(t)\mathbf{\Phi}(t) \tag{18}$$

because each column vector of $\mathbf{\Phi}(t)$ satisfies Eq. (13). Therefore, Eq. (17) reduces to

$$\mathbf{\Phi}(t)\mathbf{u}'(t) = \mathbf{f}(t). \tag{19}$$

Thus it suffices to choose $\mathbf{u}(t)$ so that

$$\mathbf{u}'(t) = \mathbf{\Phi}(t)^{-1}\mathbf{f}(t); \tag{20}$$

that is, so that

$$\mathbf{u}(t) = \int \mathbf{\Phi}(t)^{-1}\mathbf{f}(t)\,dt. \tag{21}$$

Upon substitution of (21) in (15), we finally obtain the desired particular solution, as stated in the following theorem.

THEOREM 1 Variation of Parameters

If $\mathbf{\Phi}(t)$ is a fundamental matrix for the homogeneous system $\mathbf{x}' = \mathbf{P}(t)\mathbf{x}$ on some interval where $\mathbf{P}(t)$ and $\mathbf{f}(t)$ are continuous, then a particular solution of the nonhomogeneous system

$$\mathbf{x}' = \mathbf{P}(t)\mathbf{x} + \mathbf{f}(t)$$

is given by

$$\mathbf{x}_p(t) = \mathbf{\Phi}(t)\int \mathbf{\Phi}(t)^{-1}\mathbf{f}(t)\,dt. \tag{22}$$

This is the **variation of parameters formula** for first-order linear systems. If we add this particular solution and the complementary function in (14), we get the general solution

$$\mathbf{x}(t) = \mathbf{\Phi}(t)\mathbf{c} + \mathbf{\Phi}(t)\int \mathbf{\Phi}(t)^{-1}\mathbf{f}(t)\,dt \tag{23}$$

of the nonhomogeneous system in (11).

The choice of the constant of integration in Eq. (22) is immaterial, for we need only a single particular solution. In solving initial value problems it often is convenient to choose the constant of integration so that $\mathbf{x}_p(a) = \mathbf{0}$, and thus integrate from a to t:

$$\mathbf{x}_p(t) = \mathbf{\Phi}(t)\int_a^t \mathbf{\Phi}(s)^{-1}\mathbf{f}(s)\,ds. \tag{24}$$

If we add the particular solution of the nonhomogeneous problem

$$\mathbf{x}' = \mathbf{P}(t)\mathbf{x} + \mathbf{f}(t), \quad \mathbf{x}(a) = \mathbf{0}$$

in (24) to the solution $\mathbf{x}_c(t) = \mathbf{\Phi}(t)\mathbf{\Phi}(a)^{-1}\mathbf{x}_a$ of the associated homogeneous problem $\mathbf{x}' = \mathbf{P}(t)\mathbf{x}$, $\mathbf{x}(a) = \mathbf{x}_a$, we get the solution

$$\mathbf{x}(t) = \mathbf{\Phi}(t)\mathbf{\Phi}(a)^{-1}\mathbf{x}_a + \mathbf{\Phi}(t)\int_a^t \mathbf{\Phi}(s)^{-1}\mathbf{f}(s)\,ds \tag{25}$$

of the nonhomogeneous initial value problem

$$\mathbf{x}' = \mathbf{P}(t)\mathbf{x} + \mathbf{f}(t), \quad \mathbf{x}(a) = \mathbf{x}_a. \tag{26}$$

Equations (22) and (25) hold for any fundamental matrix $\mathbf{\Phi}(t)$ of the homogeneous system $\mathbf{x}' = \mathbf{P}(t)\mathbf{x}$. In the constant-coefficient case $\mathbf{P}(t) \equiv \mathbf{A}$ we can use for $\mathbf{\Phi}(t)$ the exponential matrix $e^{\mathbf{A}t}$—that is, the particular fundamental matrix such that $\mathbf{\Phi}(0) = \mathbf{I}$. Then, because $(e^{\mathbf{A}t})^{-1} = e^{-\mathbf{A}t}$, substitution of $\mathbf{\Phi}(t) = e^{\mathbf{A}t}$ in (22) yields the particular solution

$$\mathbf{x}_p(t) = e^{\mathbf{A}t}\int e^{-\mathbf{A}t}\mathbf{f}(t)\,dt \tag{27}$$

of the nonhomogeneous system $\mathbf{x}' = \mathbf{P}(t)\mathbf{x} + \mathbf{f}(t)$. Similarly, substitution of $\mathbf{\Phi}(t) = e^{\mathbf{A}t}$ in Eq. (25) with $a = 0$ yields the solution

$$\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}_0 + e^{\mathbf{A}t}\int_0^t e^{-\mathbf{A}t}\mathbf{f}(t)\,dt \tag{28}$$

of the initial value problem

$$\mathbf{x}' = \mathbf{P}(t)\mathbf{x} + \mathbf{f}(t), \quad \mathbf{x}(0) = \mathbf{x}_0. \tag{29}$$

Remark If we retain t as the independent variable but use s for the variable of integration, then the solutions in (27) and (28) can be rewritten in the forms

$$\mathbf{x}_p(t) = \int e^{-\mathbf{A}(s-t)}\mathbf{f}(s)\,ds \quad \text{and} \quad \mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}_0 + \int_0^t e^{-\mathbf{A}(s-t)}\mathbf{f}(s)\,ds.$$ ■

Example 4 Solve the initial value problem

$$\mathbf{x}' = \begin{bmatrix} 4 & 2 \\ 3 & -1 \end{bmatrix}\mathbf{x} - \begin{bmatrix} 15 \\ 4 \end{bmatrix} te^{-2t}, \quad \mathbf{x}(0) = \begin{bmatrix} 7 \\ 3 \end{bmatrix}. \tag{30}$$

Solution The solution of the associated homogeneous system is displayed in Eq. (10). It gives the fundamental matrix

$$\mathbf{\Phi}(t) = \begin{bmatrix} e^{-2t} & 2e^{5t} \\ -3e^{-2t} & e^{5t} \end{bmatrix} \quad \text{with} \quad \mathbf{\Phi}(0)^{-1} = \frac{1}{7}\begin{bmatrix} 1 & -2 \\ 3 & 1 \end{bmatrix}.$$

It follows by Eq. (28) in Section 5.6 that the matrix exponential for the coefficient matrix **A** in (30) is

$$\begin{aligned} e^{\mathbf{A}t} = \mathbf{\Phi}(t)\mathbf{\Phi}(0)^{-1} &= \begin{bmatrix} e^{-2t} & 2e^{5t} \\ -3e^{-2t} & e^{5t} \end{bmatrix} \cdot \frac{1}{7}\begin{bmatrix} 1 & -2 \\ 3 & 1 \end{bmatrix} \\ &= \frac{1}{7}\begin{bmatrix} e^{-2t} + 6e^{5t} & -2e^{-2t} + 2e^{5t} \\ -3e^{-2t} + 3e^{5t} & 6e^{-2t} + e^{5t} \end{bmatrix}. \end{aligned}$$

Then the variation of parameters formula in Eq. (28) gives

$$\begin{aligned} e^{-\mathbf{A}t}\mathbf{x}(t) &= \mathbf{x}_0 + \int_0^t e^{-\mathbf{A}s}\mathbf{f}(s)\,ds \\ &= \begin{bmatrix} 7 \\ 3 \end{bmatrix} + \int_0^t \frac{1}{7}\begin{bmatrix} e^{2s} + 6e^{-5s} & -2e^{2s} + 2e^{-5s} \\ -3e^{2s} + 3e^{-5s} & 6e^{2s} + e^{-5s} \end{bmatrix}\begin{bmatrix} -15se^{-2s} \\ -4se^{-2s} \end{bmatrix} ds \\ &= \begin{bmatrix} 7 \\ 3 \end{bmatrix} + \int_0^t \begin{bmatrix} -s - 14se^{-7s} \\ 3s - 7se^{-7s} \end{bmatrix} ds \\ &= \begin{bmatrix} 7 \\ 3 \end{bmatrix} + \frac{1}{14}\begin{bmatrix} -4 - 7t^2 + 4e^{-7t} + 28te^{-7t} \\ -2 + 21t^2 + 2e^{-7t} + 14te^{-7t} \end{bmatrix}. \end{aligned}$$

Therefore,

$$e^{-\mathbf{A}t}\mathbf{x}(t) = \frac{1}{14}\begin{bmatrix} 94 - 7t^2 + 4e^{-7t} + 28te^{-7t} \\ 40 + 21t^2 + 2e^{-7t} + 14te^{-7t} \end{bmatrix}.$$

Upon multiplication of the right-hand side here by $e^{\mathbf{A}t}$, we find that the solution of the initial value problem in (30) is given by

$$\begin{aligned} \mathbf{x}(t) &= \frac{1}{7}\begin{bmatrix} e^{-2t} + 6e^{5t} & -2e^{-2t} + 2e^{5t} \\ -3e^{-2t} + 3e^{5t} & 6e^{-2t} + e^{5t} \end{bmatrix} \cdot \frac{1}{14}\begin{bmatrix} 94 - 7t^2 + 4e^{-7t} + 28te^{-7t} \\ 40 + 21t^2 + 2e^{-7t} + 14te^{-7t} \end{bmatrix} \\ &= \frac{1}{14}\begin{bmatrix} (6 + 28t - 7t^2)e^{-2t} + 92e^{5t} \\ (-4 + 14t + 21t^2)e^{-2t} + 46e^{5t} \end{bmatrix}. \end{aligned}$$

■

In conclusion, let us investigate how the variation of parameters formula in (22) "reconciles" with the variation of parameters formula in Theorem 1 of Section 3.5 for the second-order linear differential equation

$$y'' + Py' + Qy = f(t). \tag{31}$$

If we write $y = x_1$, $y' = x_1' = x_2$, $y'' = x_1'' = x_2'$, then the single equation in (31) is equivalent to the linear system $x_1' = x_2$, $x_2' = -Qx_1 - Px_2 + f(t)$, that is,

$$\mathbf{x}' = \mathbf{P}(t)\mathbf{x} + \mathbf{f}(t), \tag{32}$$

where

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} y \\ y' \end{bmatrix}, \quad \mathbf{P}(t) = \begin{bmatrix} 0 & 1 \\ -Q & -P \end{bmatrix}, \quad \text{and} \quad \mathbf{f}(t) = \begin{bmatrix} 0 \\ f(t) \end{bmatrix}.$$

Now two linearly independent solutions y_1 and y_2 of the homogeneous system $y'' + Py' + Qy = 0$ associated with (31) provide two linearly independent solutions

$$\mathbf{x}_1 = \begin{bmatrix} y_1 \\ y_1' \end{bmatrix} \quad \text{and} \quad \mathbf{x}_2 = \begin{bmatrix} y_2 \\ y_2' \end{bmatrix}$$

of the homogeneous system $\mathbf{x}' = \mathbf{P}(t)\mathbf{x}$ associated with (32). Observe that the determinant of the fundamental matrix $\mathbf{\Phi} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 \end{bmatrix}$ is simply the Wronskian

$$W = \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix}$$

of the solutions y_1 and y_2, so the inverse fundamental matrix is

$$\mathbf{\Phi}^{-1} = \frac{1}{W}\begin{vmatrix} y_2' & -y_2 \\ -y_1' & y_1 \end{vmatrix}.$$

Therefore the variation of parameters formula $\mathbf{x}_p = \mathbf{\Phi} \int \mathbf{\Phi}^{-1}\mathbf{f}\, dt$ in (22) yields

$$\begin{aligned} \begin{bmatrix} y_p \\ y_p' \end{bmatrix} &= \begin{bmatrix} y_1 & y_2 \\ y_1' & y_2' \end{bmatrix} \int \frac{1}{W} \begin{bmatrix} y_2' & -y_2 \\ -y_1' & y_1 \end{bmatrix} \begin{bmatrix} 0 \\ f \end{bmatrix} dt \\ &= \begin{bmatrix} y_1 & y_2 \\ y_1' & y_2' \end{bmatrix} \int \frac{1}{W} \begin{bmatrix} -y_2 f \\ y_1 f \end{bmatrix} dt. \end{aligned}$$

The first component of this column vector is

$$y_p = \begin{bmatrix} y_1 & y_2 \end{bmatrix} \int \frac{1}{W} \begin{bmatrix} -y_2 f \\ y_1 f \end{bmatrix} dt = -y_1 \int \frac{y_2 f}{W}\, dt + y_2 \int \frac{y_1 f}{W}\, dt.$$

If, finally, we supply the independent variable t throughout, the final result on the right-hand side here is simply the variation of parameters formula in Eq. (33) of Section 3.5 (where, however, the independent variable is denoted by x).

5.7 Problems

Apply the method of undetermined coefficients to find a particular solution of each of the systems in Problems 1 through 14. If initial conditions are given, find the particular solution that satisfies these conditions. Primes denote derivatives with respect to t.

1. $x' = x + 2y + 3$, $y' = 2x + y - 2$
2. $x' = 2x + 3y + 5$, $y' = 2x + y - 2t$
3. $x' = 3x + 4y$, $y' = 3x + 2y + t^2$; $x(0) = y(0) = 0$
4. $x' = 4x + y + e^t$, $y' = 6x - y - e^t$; $x(0) = y(0) = 1$
5. $x' = 6x - 7y + 10$, $y' = x - 2y - 2e^{-t}$
6. $x' = 9x + y + 2e^t$, $y' = -8x - 2y + te^t$
7. $x' = -3x + 4y + \sin t$, $y' = 6x - 5y$; $x(0) = 1$, $y(0) = 0$
8. $x' = x - 5y + 2\sin t$, $y' = x - y - 3\cos t$
9. $x' = x - 5y + \cos 2t$, $y' = x - y$
10. $x' = x - 2y$, $y' = 2x - y + e^t \sin t$
11. $x' = 2x + 4y + 2$, $y' = x + 2y + 3$; $x(0) = 1$, $y(0) = -1$
12. $x' = x + y + 2t$, $y' = x + y - 2t$
13. $x' = 2x + y + 2e^t$, $y' = x + 2y - 3e^t$
14. $x' = 2x + y + 1$, $y' = 4x + 2y + e^{4t}$

Problems 15 and 16 are similar to Example 2, but with two brine tanks (having volumes V_1 and V_2 gallons as in Fig. 5.7.1) instead of three tanks. Each tank initially contains fresh water, and the inflow to tank 1 at the rate of r gallons per minute has a salt concentration of c_0 pounds per gallon. ***(a)*** *Find the amounts $x_1(t)$ and $x_2(t)$ of salt in the two tanks after t minutes.* ***(b)*** *Find the limiting (long-term) amount of salt in each tank.* ***(c)*** *Find how long it takes for each tank to reach a salt concentration of 1 lb/gal.*

Plot similarly some solution curves for the following differential equations.

1. $\dfrac{dy}{dx} = \dfrac{4x - 5y}{2x + 3y}$

2. $\dfrac{dy}{dx} = \dfrac{4x - 5y}{2x - 3y}$

3. $\dfrac{dy}{dx} = \dfrac{4x - 3y}{2x - 5y}$

4. $\dfrac{dy}{dx} = \dfrac{2xy}{x^2 - y^2}$

5. $\dfrac{dy}{dx} = \dfrac{x^2 + 2xy}{y^2 + 2xy}$

Now construct some examples of your own. Homogeneous functions like those in Problems 1 through 5—rational functions with numerator and denominator of the same degree in x and y—work well. The differential equation

$$\frac{dy}{dx} = \frac{25x + y(1 - x^2 - y^2)(4 - x^2 - y^2)}{-25y + x(1 - x^2 - y^2)(4 - x^2 - y^2)} \tag{5}$$

of this form generalizes Example 5 in this section but would be inconvenient to solve explicitly. Its phase portrait (Fig. 6.1.22) shows two periodic closed trajectories—the circles $r = 1$ and $r = 2$. Anyone want to try for three circles?

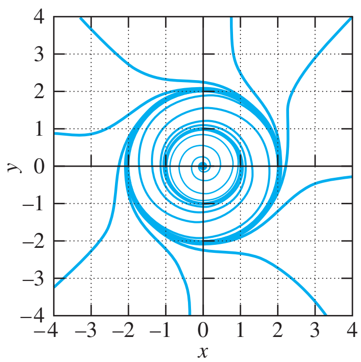


FIGURE 6.1.22. Phase portrait for the system corresponding to Eq. (5).

6.2 Linear and Almost Linear Systems

We now discuss the behavior of solutions of the autonomous system

$$\frac{dx}{dt} = f(x, y), \qquad \frac{dy}{dt} = g(x, y) \tag{1}$$

near an isolated critical point (x_0, y_0) where $f(x_0, y_0) = g(x_0, y_0) = 0$. A critical point is called **isolated** if some neighborhood of it contains no other critical point. We assume throughout that the functions f and g are continuously differentiable in a neighborhood of (x_0, y_0).

We can assume without loss of generality that $x_0 = y_0 = 0$. Otherwise, we make the substitutions $u = x - x_0$, $v = y - y_0$. Then $dx/dt = du/dt$ and $dy/dt = dv/dt$, so (1) is equivalent to the system

$$\begin{aligned} \frac{du}{dt} &= f(u + x_0, v + y_0) = f_1(u, v), \\ \frac{dv}{dt} &= g(u + x_0, v + y_0) = g_1(u, v) \end{aligned} \tag{2}$$

that has $(0, 0)$ as an isolated critical point.

Example 1 The system

$$\begin{aligned} \frac{dx}{dt} &= 3x - x^2 - xy = x(3 - x - y), \\ \frac{dy}{dt} &= y + y^2 - 3xy = y(1 - 3x + y) \end{aligned} \tag{3}$$

has $(1, 2)$ as one of its critical points. We substitute $u = x - 1$, $v = y - 2$; that is, $x = u + 1$, $y = v + 2$. Then

$$3 - x - y = 3 - (u + 1) - (v + 2) = -u - v$$

and

$$1 - 3x + y = 1 - 3(u + 1) + (v + 2) = -3u + v,$$

so the system in (3) takes the form

$$\begin{aligned} \frac{du}{dt} &= (u + 1)(-u - v) = -u - v - u^2 - uv, \\ \frac{dv}{dt} &= (v + 2)(-3u + v) = -6u + 2v + v^2 - 3uv \end{aligned} \tag{4}$$

and has $(0, 0)$ as a critical point. If we can determine the trajectories of the system in (4) near $(0, 0)$, then their translations under the rigid motion that carries $(0, 0)$ to $(1, 2)$ will be the trajectories near $(1, 2)$ of the original system in (3). This equivalence is illustrated by Fig. 6.2.1 (which shows computer-plotted trajectories of the system in (3) near the critical point $(1, 2)$ in the xy-plane) and Fig. 6.2.2 (which shows computer-plotted trajectories of the system in (4) near the critical point $(0, 0)$ in the uv-plane). ■

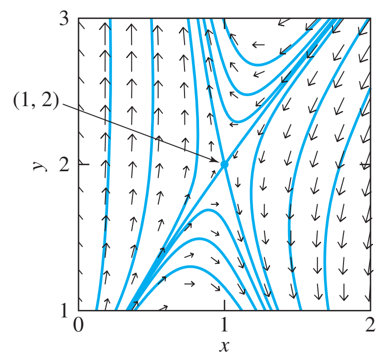


FIGURE 6.2.1. The saddle point $(1, 2)$ for the system $x' = 3x - x^2 - xy$, $y' = y + y^2 - 3xy$ of Example 1.

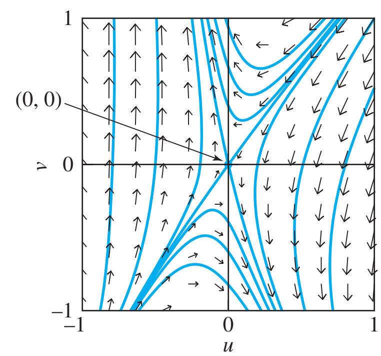


FIGURE 6.2.2. The saddle point $(0, 0)$ for the equivalent system $u' = -u - v - u^2 - uv$, $v' = -6u + 2v + v^2 - 3uv$.

Figures 6.2.1 and 6.2.2 illustrate the fact that the solution curves of the xy-system in (1) are simply the images under the translation $(u, v) \to (u + x_0, v + y_0)$ of the solution curves of the uv-system in (2). Near the two corresponding critical points—(x_0, y_0) in the xy-plane and $(0, 0)$ in the uv-plane—the two phase portraits therefore look precisely the same.

Linearization Near a Critical Point

Taylor's formula for functions of two variables implies that—if the function $f(x, y)$ is continuously differentiable near the fixed point (x_0, y_0)—then

$$f(x_0 + u, y_0 + v) = f(x_0, y_0) + f_x(x_0, y_0)u + f_y(x_0, y_0)v + r(u, v)$$

where the "remainder term" $r(u, v)$ satisfies the condition

$$\lim_{(u,v)\to(0,0)} \frac{r(u, v)}{\sqrt{u^2 + v^2}} = 0.$$

(Note that this condition would not be satisfied if $r(u, v)$ were a sum containing either constants or terms linear in u or v. In this sense, $r(u, v)$ consists of the "nonlinear part" of the function $f(x_0 + u, y_0 + v)$ of u and v.)

If we apply Taylor's formula to both f and g in (2) and assume that (x_0, y_0) is an isolated critical point so $f(x_0, y_0) = g(x_0, y_0) = 0$, the result is

$$\begin{aligned}\frac{du}{dt} &= f_x(x_0, y_0)u + f_y(x_0, y_0)v + r(u, v),\\ \frac{dv}{dt} &= g_x(x_0, y_0)u + g_y(x_0, y_0)v + s(u, v)\end{aligned} \tag{5}$$

where $r(u, v)$ and the analogous remainder term $s(u, v)$ for g satisfy the condition

$$\lim_{(u,v)\to(0,0)} \frac{r(u, v)}{\sqrt{u^2 + v^2}} = \lim_{(u,v)\to(0,0)} \frac{s(u, v)}{\sqrt{u^2 + v^2}} = 0. \tag{6}$$

Then, when the values u and v are small, the remainder terms $r(u, v)$ and $s(u, v)$ are *very* small (being small even in comparison with u and v).

If we drop the presumably small nonlinear terms $r(u, v)$ and $s(u, v)$ in (5), the result is the *linear* system

➤ $$\begin{aligned}\frac{du}{dt} &= f_x(x_0, y_0)u + f_y(x_0, y_0)v,\\ \frac{dv}{dt} &= g_x(x_0, y_0)u + g_y(x_0, y_0)v\end{aligned} \tag{7}$$

whose constant coefficients (of the variables u and v) are the values $f_x(x_0, y_0)$, $f_y(x_0, y_0)$ and $g_x(x_0, y_0)$, $g_y(x_0, y_0)$ of the functions f and g at the critical point (x_0, y_0). Because (5) is equivalent to the original (and generally) nonlinear system $u' = f(x_0 + u, y_0 + v)$, $v' = g(x_0 + u, y_0 + v)$ in (2), the conditions in (6) suggest that the **linearized system** in (7) closely approximates the given nonlinear system when (u, v) is close to $(0, 0)$.

Assuming that $(0, 0)$ is also an isolated critical point of the linear system, and that the remainder terms in (5) satisfy the condition in (6), the original system $x' = f(x, y)$, $y' = g(x, y)$ is said to be **almost linear** at the isolated critical point (x_0, y_0). In this case, its **linearization** at (x_0, y_0) is the linear system in (7). In short, this linearization is the linear system $\mathbf{u}' = \mathbf{J}\mathbf{u}$ (where $\mathbf{u} = \begin{bmatrix} u & v \end{bmatrix}^T$) whose coefficient matrix is the so-called **Jacobian matrix**

➤ $$\mathbf{J}(x_0, y_0) = \begin{bmatrix} f_x(x_0, y_0) & f_y(x_0, y_0) \\ g_x(x_0, y_0) & g_y(x_0, y_0) \end{bmatrix} \tag{8}$$

of the functions f and g, evaluated at the point (x_0, y_0).

Example 1 *Continued*

In (3) we have $f(x, y) = 3x - x^2 - xy$ and $g(x, y) = y + y^2 - 3xy$. Then

$$\mathbf{J}(x, y) = \begin{bmatrix} 3 - 2x - y & -x \\ -3y & 1 + 2y - 3x \end{bmatrix}, \quad \text{so} \quad \mathbf{J}(1, 2) = \begin{bmatrix} -1 & -1 \\ -6 & 2 \end{bmatrix}.$$

Hence the linearization of the system $x' = 3x - x^2 - xy$, $y' = y + y^2 - 3xy$ at its critical point $(1, 2)$ is the linear system

$$\begin{aligned}u' &= -u - v,\\ v' &= -6u + 2v\end{aligned}$$

that we get when we drop the nonlinear (quadratic) terms in (4). ■

It turns out that in most (though not all) cases, the phase portrait of an almost linear system near an isolated critical point (x_0, y_0) strongly resembles—qualitatively—the phase portrait of its linearization near the origin. Consequently, *the first step toward understanding general autonomous systems is to characterize the isolated critical points of linear systems.*

Isolated Critical Points of Linear Systems

In Section 5.3 we used the eigenvalue-eigenvector method to study the 2×2 linear system

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{9}$$

with constant-coefficient matrix **A**. The origin $(0, 0)$ is a critical point of the system regardless of the matrix **A**, but if we further require the origin to be an *isolated* critical point, then (by a standard theorem of linear algebra) the determinant $ad - bc$ of **A** must be nonzero. From this we can conclude that *the eigenvalues* λ_1 *and* λ_2 *of* **A** *must be nonzero*. Indeed, λ_1 and λ_2 are the solutions of the characteristic equation

$$\begin{aligned} \det(\mathbf{A} - \lambda\mathbf{I}) &= \begin{vmatrix} a - \lambda & b \\ c & d - \lambda \end{vmatrix} \\ &= (a - \lambda)(d - \lambda) - bc \\ &= \lambda^2 - (a + d)\lambda + (ad - bc) \\ &= 0, \end{aligned} \tag{10}$$

and the fact that $ad - bc \neq 0$ implies that $\lambda = 0$ cannot satisfy Eq. (10); hence λ_1 and λ_2 are nonzero. The converse also holds: If the characteristic equation (10) has no zero solution—that is, if all eigenvalues of the matrix **A** are nonzero—then the determinant $ad - bc$ is nonzero. Altogether, we conclude that the origin $(0, 0)$ is an isolated critical point of the system in Eq. (9) if and only if the eigenvalues of **A** are all nonzero. Our study of this critical point can be divided, therefore, into the five cases listed in the table in Fig. 6.2.3. This table also gives the type of each critical point as found in Section 5.3 and shown in our gallery Fig. 5.3.16 of typical phase plane portraits:

Eigenvalues of A	Type of Critical Point
Real, unequal, same sign	Improper node
Real, unequal, opposite sign	Saddle point
Real and equal	Proper or improper node
Complex conjugate	Spiral point
Pure imaginary	Center

FIGURE 6.2.3. Classification of the isolated critical point $(0, 0)$ of the two-dimensional system $\mathbf{x}' = \mathbf{A}\mathbf{x}$.

Closer inspection of that gallery, however, reveals a striking connection between the *stability* properties of the critical point and the eigenvalues λ_1 and λ_2 of **A**. For example, if λ_1 and λ_2 are real, unequal, and negative, then the origin represents an improper nodal sink; because all trajectories approach the origin as $t \to +\infty$, the critical point is asymptotically stable. Likewise, if λ_1 and λ_2 are real, equal, and negative, then the origin is a proper nodal sink, and is again asymptotically stable. Further, if λ_1 and λ_2 are complex conjugate with negative real part, then the origin is a spiral sink, and is once more asymptotically stable. All three of these cases can be captured as follows: *If the real parts of* λ_1 *and* λ_2 *are negative, then the origin is an asymptotically stable critical point*. (Note that if λ_1 and λ_2 are real, then they are themselves their real parts.)

Similar generalizations can be made for other combinations of signs of the real parts of λ_1 and λ_2. Indeed, as the table in Fig. 6.2.4 shows, the stability properties of the isolated critical point $(0, 0)$ of the system in Eq. (9) are always determined by the signs of the real parts of λ_1 and λ_2. (We invite you to use the gallery in Fig. 5.3.16 to verify the conclusions in the table.)

Real Parts of λ_1 and λ_2	Type of Critical Point	Stability
Both negative	• Proper or improper nodal sink, *or* • Spiral sink	Asymptotically stable
Both zero (*i.e.*, λ_1 and λ_2 are given by $\pm iq$ with $q \neq 0$)	• Center	Stable but not asymptotically stable
At least one positive	• Proper or improper nodal source, *or* • Spiral source, *or* • Saddle point	Unstable

FIGURE 6.2.4. Stability properties of the isolated critical point $(0, 0)$ of the system in Eq. (9) with nonzero eigenvalues λ_1 and λ_2.

These findings are summarized in Theorem 1:

THEOREM 1 Stability of Linear Systems

Let λ_1 and λ_2 be the eigenvalues of the coefficient matrix **A** of the two-dimensional linear system

$$\begin{aligned} \frac{dx}{dt} &= ax + by, \\ \frac{dy}{dt} &= cx + dy \end{aligned} \tag{11}$$

with $ad - bc \neq 0$. Then the critical point $(0, 0)$ is

1. Asymptotically stable if the real parts of λ_1 and λ_2 are both negative;
2. Stable but not asymptotically stable if the real parts of λ_1 and λ_2 are both zero (so that $\lambda_1, \lambda_2 = \pm qi$);
3. Unstable if either λ_1 or λ_2 has a positive real part.

It is worthwhile to consider the effect of small perturbations in the coefficients a, b, c, and d of the linear system in (11), which result in small perturbations of the eigenvalues λ_1 and λ_2. If these perturbations are sufficiently small, then positive real parts (of λ_1 and λ_2) remain positive and negative real parts remain negative. Hence an asymptotically stable critical point remains asymptotically stable and an unstable critical point remains unstable. Part 2 of Theorem 1 is therefore the only case in which arbitrarily small perturbations can affect the stability of the critical point $(0, 0)$. In this case pure imaginary roots $\lambda_1, \lambda_2 = \pm qi$ of the characteristic equation can be changed to nearby complex roots $\mu_1, \mu_2 = r \pm si$, with r either positive or negative (see Fig. 6.2.5). Consequently, a small perturbation of the coefficients of the linear system in (11) can change a stable center to a spiral point that is either unstable or asymptotically stable.

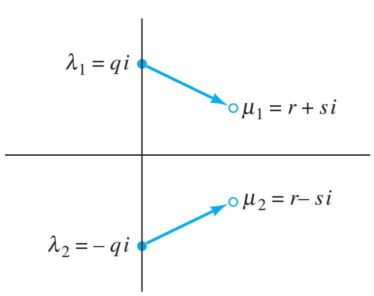


FIGURE 6.2.5. The effects of perturbation of pure imaginary roots.

There is one other exceptional case in which the type, though not the stability, of the critical point $(0, 0)$ can be altered by a small perturbation of its coefficients. This is the case with $\lambda_1 = \lambda_2$, equal roots that (under a small perturbation of the coefficients) can split into two roots μ_1 and μ_2, which are either complex conjugates or unequal real roots (see Fig. 6.2.6). In either case, the sign of the real parts of the roots is preserved, so the stability of the critical point is unaltered. Its nature may change, however; the table in Fig. 6.2.3 shows that a node with $\lambda_1 = \lambda_2$ can either remain a node (if μ_1 and μ_2 are real) or change to a spiral point (if μ_1 and μ_2 are complex conjugates).

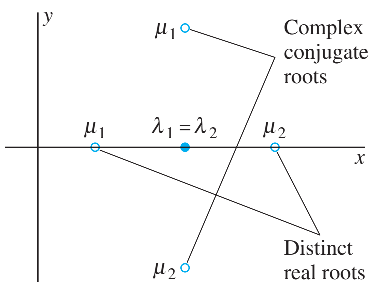


FIGURE 6.2.6. The effects of perturbation of real equal roots.

Suppose that the linear system in (11) is used to model a physical situation. It is unlikely that the coefficients in (11) can be measured with total accuracy, so let the unknown precise linear model be

$$\begin{aligned} \frac{dx}{dt} &= a^\star x + b^\star y, \\ \frac{dy}{dt} &= c^\star x + d^\star y. \end{aligned} \tag{11*}$$

If the coefficients in (11) are sufficiently close to those in (11*), it then follows from the discussion in the preceding paragraph that the origin $(0, 0)$ is an asymptotically stable critical point for (11) if it is an asymptotically stable critical point for (11*), and is an unstable critical point for (11) if it is an unstable critical point for (11*). Thus in this case the approximate model in (11) and the precise model in (11*) predict the same qualitative behavior (with respect to asymptotic stability versus instability).

Almost Linear Systems

Recall that we first encountered an almost linear system at the beginning of this section, when we used Taylor's formula to write the nonlinear system (2) in the almost linear form (5) which led to the linearization (7) of the original nonlinear system. In case the nonlinear system $x' = f(x, y)$, $y' = g(x, y)$ has $(0, 0)$ as an isolated critical point, the corresponding almost linear system is

$$\begin{aligned} \frac{dx}{dt} &= ax + by + r(x, y), \\ \frac{dy}{dt} &= cx + dy + s(x, y), \end{aligned} \tag{12}$$

where $a = f_x(0, 0)$, $b = f_y(0, 0)$ and $c = g_x(0, 0)$, $d = g_y(0, 0)$; we assume also that $ad - bc \neq 0$. Theorem 2, which we state without proof, essentially implies that—with regard to the type and stability of the critical point $(0, 0)$—the effect of the small nonlinear terms $r(x, y)$ and $s(x, y)$ is equivalent to the effect of a small perturbation in the coefficients of the associated *linear* system in (11).

THEOREM 2 Stability of Almost Linear Systems

Let λ_1 and λ_2 be the eigenvalues of the coefficient matrix of the linear system in (11) associated with the almost linear system in (12). Then

1. If $\lambda_1 = \lambda_2$ are equal real eigenvalues, then the critical point $(0, 0)$ of (12) is either a node or a spiral point, and is asymptotically stable if $\lambda_1 = \lambda_2 < 0$, unstable if $\lambda_1 = \lambda_2 > 0$.
2. If λ_1 and λ_2 are pure imaginary, then $(0, 0)$ is either a center or a spiral point, and may be either asymptotically stable, stable, or unstable.

3. Otherwise—that is, unless λ_1 and λ_2 are either real equal or pure imaginary—the critical point $(0, 0)$ of the almost linear system in (12) is of the same type and stability as the critical point $(0, 0)$ of the associated linear system in (11).

Thus, if $\lambda_1 \neq \lambda_2$ and $\text{Re}(\lambda_1) \neq 0$, then the type and stability of the critical point of the almost linear system in (12) can be determined by analysis of its associated linear system in (11), and only in the case of pure imaginary eigenvalues is the stability of $(0, 0)$ not determined by the linear system. Except in the sensitive cases $\lambda_1 = \lambda_2$ and $\text{Re}(\lambda_i) = 0$, the trajectories near $(0, 0)$ will resemble qualitatively those of the associated linear system—they enter or leave the critical point in the same way, but may be "deformed" in a nonlinear manner. The table in Fig. 6.2.7 summarizes the situation.

Eigenvalues λ_1, λ_2 for the Linearized System	**Type of Critical Point of the Almost Linear System**
$\lambda_1 < \lambda_2 < 0$	Stable improper node
$\lambda_1 = \lambda_2 < 0$	Stable node or spiral point
$\lambda_1 < 0 < \lambda_2$	Unstable saddle point
$\lambda_1 = \lambda_2 > 0$	Unstable node or spiral point
$\lambda_1 > \lambda_2 > 0$	Unstable improper node
$\lambda_1, \lambda_2 = a \pm bi \quad (a < 0)$	Stable spiral point
$\lambda_1, \lambda_2 = a \pm bi \quad (a > 0)$	Unstable spiral point
$\lambda_1, \lambda_2 = \pm bi$	Stable or unstable, center or spiral point

FIGURE 6.2.7. Classification of critical points of an almost linear system.

An important consequence of the classification of cases in Theorem 2 is that a *critical point of an almost linear system is asymptotically stable if it is an asymptotically stable critical point of the linearization of the system.* Moreover, a critical point of the almost linear system is unstable if it is an unstable critical point of the linearized system. If an almost linear system is used to model a physical situation, then—apart from the sensitive cases mentioned earlier—it follows that the qualitative behavior of the system near a critical point can be determined by examining its linearization.

Example 2 Determine the type and stability of the critical point $(0, 0)$ of the almost linear system

$$\begin{aligned} \frac{dx}{dt} &= 4x + 2y + 2x^2 - 3y^2, \\ \frac{dy}{dt} &= 4x - 3y + 7xy. \end{aligned} \tag{13}$$

Solution The characteristic equation for the associated linear system (obtained simply by deleting the quadratic terms in (13)) is

$$(4 - \lambda)(-3 - \lambda) - 8 = (\lambda - 5)(\lambda + 4) = 0,$$

so the eigenvalues $\lambda_1 = 5$ and $\lambda_2 = -4$ are real, unequal, and have opposite signs. By our discussion of this case we know that $(0, 0)$ is an unstable saddle point of the linear system, and hence by Part 3 of Theorem 2, it is also an unstable saddle point of the almost linear system in (13). The trajectories of the linear system near $(0, 0)$ are shown in Fig. 6.2.8, and those of the nonlinear system in (13) are shown in Fig. 6.2.9. Figure 6.2.10 shows a phase

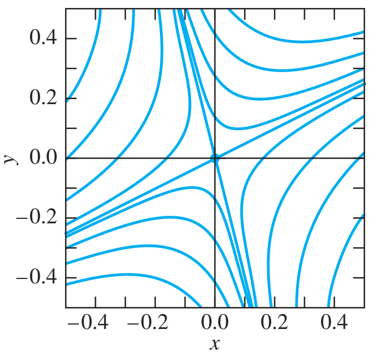


FIGURE 6.2.8. Trajectories of the linearized system of Example 2.

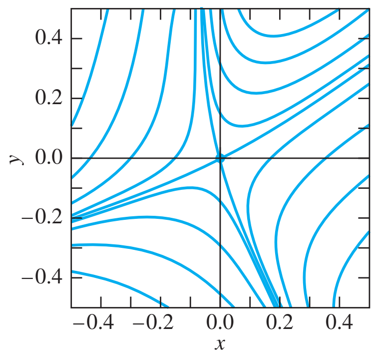


FIGURE 6.2.9. Trajectories of the original almost linear system of Example 2.

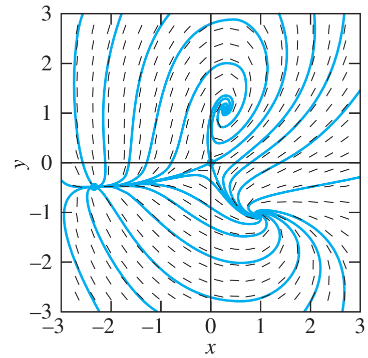


FIGURE 6.2.10. Phase portrait for the almost linear system in Eq. (13).

portrait of the nonlinear system in (13) from a "wider view." In addition to the saddle point at $(0, 0)$, there are spiral points near the points $(0.279, 1.065)$ and $(0.933, -1.057)$, and a node near $(-2.354, -0.483)$. ■

We have seen that the system $x' = f(x, y)$, $y' = g(x, y)$ with isolated critical point (x_0, y_0) transforms via the substitution $x = u + x_0$, $y = v + y_0$ to an equivalent uv-system with corresponding critical point $(0, 0)$ and linearization $\mathbf{u}' = \mathbf{J}\mathbf{u}$, whose coefficient matrix $\mathbf{J}$ is the Jacobian matrix in (8) of the functions f and g at (x_0, y_0). Consequently we need not carry out the substitution explicitly; instead, we can proceed directly to calculate the eigenvalues of $\mathbf{J}$ preparatory to application of Theorem 2.

Example 3 Determine the type and stability of the critical point $(4, 3)$ of the almost linear system

$$\begin{aligned} \frac{dx}{dt} &= 33 - 10x - 3y + x^2, \\ \frac{dy}{dt} &= -18 + 6x + 2y - xy. \end{aligned} \tag{14}$$

Solution With $f(x, y) = 33 - 10x - 3y + x^2$, $g(x, y) = -18 + 6x + 2y - xy$ and $x_0 = 4$, $y_0 = 3$ we have

$$\mathbf{J}(x, y) = \begin{bmatrix} -10 + 2x & -3 \\ 6 - y & 2 - x \end{bmatrix}, \quad \text{so} \quad \mathbf{J}(4, 3) = \begin{bmatrix} -2 & -3 \\ 3 & -2 \end{bmatrix}.$$

The associated linear system

$$\begin{aligned} \frac{du}{dt} &= -2u - 3v, \\ \frac{dv}{dt} &= 3u - 2v \end{aligned} \tag{15}$$

has characteristic equation $(\lambda + 2)^2 + 9 = 0$, with complex conjugate roots $\lambda = -2 \pm 3i$. Hence $(0, 0)$ is an asymptotically stable spiral point of the linear system in (15), so Theorem 2 implies that $(4, 3)$ is an asymptotically stable spiral point of the original almost linear system in (14). Figure 6.2.11 shows some typical trajectories of the linear system in (15), and Fig. 6.2.12 shows how this spiral point fits into the phase portrait for the original almost linear system in (14). ■

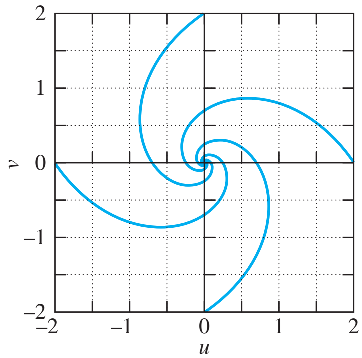


FIGURE 6.2.11. Spiral trajectories of the linear system in Eq. (15).

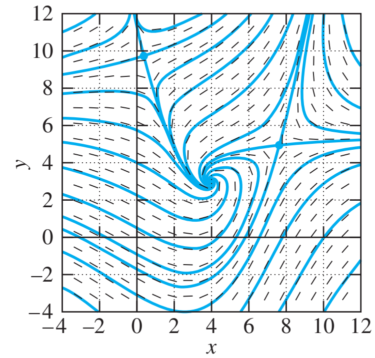


FIGURE 6.2.12. Phase portrait for the almost linear system in Eq. (14).

6.2 Problems

In Problems 1 through 10, apply Theorem 1 to determine the type of the critical point $(0, 0)$ *and whether it is asymptotically stable, stable, or unstable. Verify your conclusion by using a computer system or graphing calculator to construct a phase portrait for the given linear system.*

1. $\dfrac{dx}{dt} = -2x + y, \quad \dfrac{dy}{dt} = x - 2y$

2. $\dfrac{dx}{dt} = 4x - y, \quad \dfrac{dy}{dt} = 2x + y$

3. $\dfrac{dx}{dt} = x + 2y, \quad \dfrac{dy}{dt} = 2x + y$

4. $\dfrac{dx}{dt} = 3x + y, \quad \dfrac{dy}{dt} = 5x - y$

5. $\dfrac{dx}{dt} = x - 2y, \quad \dfrac{dy}{dt} = 2x - 3y$

6. $\dfrac{dx}{dt} = 5x - 3y, \quad \dfrac{dy}{dt} = 3x - y$

7. $\dfrac{dx}{dt} = 3x - 2y, \quad \dfrac{dy}{dt} = 4x - y$

8. $\dfrac{dx}{dt} = x - 3y, \quad \dfrac{dy}{dt} = 6x - 5y$

9. $\dfrac{dx}{dt} = 2x - 2y, \quad \dfrac{dy}{dt} = 4x - 2y$

10. $\dfrac{dx}{dt} = x - 2y, \quad \dfrac{dy}{dt} = 5x - y$

Each of the systems in Problems 11 through 18 has a single critical point (x_0, y_0)*. Apply Theorem 2 to classify this critical point as to type and stability. Verify your conclusion by using a computer system or graphing calculator to construct a phase portrait for the given system.*

11. $\dfrac{dx}{dt} = x - 2y, \quad \dfrac{dy}{dt} = 3x - 4y - 2$

12. $\dfrac{dx}{dt} = x - 2y - 8, \quad \dfrac{dy}{dt} = x + 4y + 10$

13. $\dfrac{dx}{dt} = 2x - y - 2, \quad \dfrac{dy}{dt} = 3x - 2y - 2$

14. $\dfrac{dx}{dt} = x + y - 7, \quad \dfrac{dy}{dt} = 3x - y - 5$

15. $\dfrac{dx}{dt} = x - y, \quad \dfrac{dy}{dt} = 5x - 3y - 2$

16. $\dfrac{dx}{dt} = x - 2y + 1, \quad \dfrac{dy}{dt} = x + 3y - 9$

17. $\dfrac{dx}{dt} = x - 5y - 5, \quad \dfrac{dy}{dt} = x - y - 3$

18. $\dfrac{dx}{dt} = 4x - 5y + 3, \quad \dfrac{dy}{dt} = 5x - 4y + 6$

In Problems 19 through 28, investigate the type of the critical point $(0, 0)$ *of the given almost linear system. Verify your conclusion by using a computer system or graphing calculator to construct a phase portrait. Also, describe the approximate locations and apparent types of any other critical points that are visible in your figure. Feel free to investigate these additional critical points; you can use the computational methods discussed in the application material for this section.*

19. $\dfrac{dx}{dt} = x - 3y + 2xy, \quad \dfrac{dy}{dt} = 4x - 6y - xy$

20. $\dfrac{dx}{dt} = 6x - 5y + x^2, \quad \dfrac{dy}{dt} = 2x - y + y^2$

21. $\dfrac{dx}{dt} = x + 2y + x^2 + y^2, \quad \dfrac{dy}{dt} = 2x - 2y - 3xy$

22. $\dfrac{dx}{dt} = x + 4y - xy^2, \quad \dfrac{dy}{dt} = 2x - y + x^2y$

23. $\dfrac{dx}{dt} = 2x - 5y + x^3, \quad \dfrac{dy}{dt} = 4x - 6y + y^4$

24. $\dfrac{dx}{dt} = 5x - 3y + y(x^2 + y^2), \quad \dfrac{dy}{dt} = 5x + y(x^2 + y^2)$

25. $\dfrac{dx}{dt} = x - 2y + 3xy, \quad \dfrac{dy}{dt} = 2x - 3y - x^2 - y^2$

26. $\dfrac{dx}{dt} = 3x - 2y - x^2 - y^2, \quad \dfrac{dy}{dt} = 2x - y - 3xy$

27. $\dfrac{dx}{dt} = x - y + x^4 - y^2, \quad \dfrac{dy}{dt} = 2x - y + y^4 - x^2$

28. $\dfrac{dx}{dt} = 3x - y + x^3 + y^3, \quad \dfrac{dy}{dt} = 13x - 3y + 3xy$

Spring 2017

Stony Brook University
Sabyasachi Mukherjee, Matthew Dannenberg

Problem Set 1

Differential Equations with Linear Algebra

Homework Problems

1.1. *Explicit Solutions.* (25 points)

Solve the following equations using the methods (separation of variables, integrating factors) presented in class.

- $x' + x = \frac{1}{e^t}$
- $3e^t \tan x \frac{dt}{dx} + (1 - e^t)\sec^2 x = 0$
- $x' + \frac{1-2t}{t^2} x = 1$
- $\frac{dy}{dx} = x + y$
- $xy' + (2x - 3)y = x^4$

1.2. *Initial Value Problem.* (10 points)

Solve the following initial value problem.

$y'' = \sin(x), \ y(0) = 1, \ y'(0) = 1.$

1.3. *Linear Combination.* (5 points)

Verify that if $y_1(x)$ and $y_2(x)$ are solutions of the respective equations

$$y' + gy = f_1 \quad \text{and} \quad y' + gy = f_2$$

then $c_1 y_1 + c_2 y_2$ is, for every pair of constants c_1, c_2, a solutions of the equation

$$y' + gy = c_1 f_1 + c_2 f_2$$

1.4. *Characterizing Isoclines.* (10 points)

Consider the linear differential equation $y' + ay = c$, with a, c constant, $a \neq 0$. Prove that the isoclines of the direction field of this equation are horizontal lines and that every horizontal line is an isocline.

1.5. *(Bonus problem) A Bernoulli Equation.*

Solve the following differential equation.

$3\frac{dx}{dt} = 2x + \frac{t+1}{x^2}.$

Due Date: Thursday, February 2, at the beginning of recitation.

Stony Brook University
Sabyasachi Mukherjee, Matthew Dannenberg

Spring 2017
Problem Set 2

Differential Equations with Linear Algebra

Homework Problems

2.1. *A Non-standard Vector Space Structure on* $\mathbb{R}^2$. (20 points)

Show that $(\mathbb{R}^2, \mathbb{R}, \oplus, \odot)$ with the operations defined as follows is a vector space.

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \oplus \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 - 1 \\ y_1 + y_2 + 2 \end{bmatrix}$$

$$c \odot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} cx - c + 1 \\ cy + 2c - 2 \end{bmatrix}$$

Here, $+, -$ denote the usual addition and subtraction of real numbers.

2.2. *A subspace of* $M_n(\mathbb{R})$. (10 points)

Show that the set of all real $n \times n$ upper triangular matrices is a subspace of $M_n(\mathbb{R})$.

2.3. *Finding a Basis.* (10 points)

Let $\mathcal{P}_3$ be the vector space of real polynomials of degree at most 3 (with respect to usual addition of polynomials and multiplication of scalars with polynomials). Let V be the subspace of $\mathcal{P}_3$ defined as:

$$V = \{f(x) \in \mathcal{P}_3 : f(0) + f(1) = 0,\ f'(0) = f'(1)\}$$

Find a basis for V.

2.4. *Containment of subspaces.* (10 points)

Let W_1, W_2 and W_3 be subspaces of a vector space V such that W_1 is contained in $W_2 \cup W_3$. Show that W_1 is either contained in W_2, or contained in W_3.

2.5. *Describing Linear Maps.* (10 points)

Describe explicitly a linear map from $\mathbb{R}^3$ into $\mathbb{R}^3$ which has as its range the subspace spanned by $(1, 0, -1)$ and $(1, 2, 2)$.

2.6. *(Bonus problem) Range and Null Space.* (10 points)

Let V be a vector space and $T : V \to V$ be a linear map. Show that the following two statements about T are equivalent.

(a) $\text{Range}(T) \cap \text{Null}(T) = \{0\}$.

(b) $\text{Null}(T \circ T) \subseteq \text{Null}(T)$.

Due Date: Thursday, February 9, at the beginning of recitation.

Stony Brook University
Sabyasachi Mukherjee, Matthew Dannenberg

Spring 2017

Problem Set 3

Differential Equations with Linear Algebra

Homework Problems

3.1. *Vector Spaces and Dimension.* (10 points)

For each of the following spaces, show whether or not it is a vector space over the scalar field $\mathbb{R}$. If it is a vector space, give its dimension.

(a) Symmetric 2×2 real matrices, i.e. matrices A such that the transpose A^T is equal to A (with respect to usual matrix addition and multiplication of scalars with matrices).

(b) $\{(x, y) \in \mathbb{R}^2 : y > 0\}$ (with respect to the standard operations on $\mathbb{R}^2$).

3.2. *Range and Null Space.* (10 points)

Find the null space and range of the map $f : \mathbb{R}^3 \to \mathbb{R}^3$ defined by $f(x, y, z) = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 2 & 0 \\ 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$.

What is the sum of the dimensions of these two subspaces?

3.3. *Coordinates of Vectors.* (10 points)

Show that $\mathcal{B} = \{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \}$ is a basis of $\mathbb{R}^3$. What are the coordinates of the vector $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ with respect to the ordered basis $\mathcal{B}$?

3.4. *Linear Independence.* (10 points)

Suppose that the vectors u_1, u_2 and u_3 in a vector space V are linearly independent. Show that the vectors $u_1 + u_2$, $u_2 + u_3$ and $u_3 + u_1$ are also linearly independent.

3.5. *Diagonalizing Linear Maps.* (20 points)

The following matrices A represent linear maps $T : \mathbb{R}^3 \to \mathbb{R}^3$ with respect to the standard (ordered) basis $\mathcal{B} = \{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \}$. For each of them, determine whether or not T is diagonalizable. If T is diagonalizable, find a basis of $\mathbb{R}^3$ consisting of eigenvectors of T and find an invertible matrix P such that $P^{-1}AP$ is a diagonal matrix.

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

(b) $A = \begin{bmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{bmatrix}$.

3.6. *(Bonus problem) A Basis of* $\mathcal{P}_3$. (10 points)

Let $\mathcal{P}_3$ be the vector space of all real polynomials of degree at most 3, and $f(x)$ be a real polynomial of degree 3. Show that $\{f(x), f'(x), f''(x), 1\}$ is a basis of $\mathcal{P}_3$.

Due Date: Thursday, February 16, at the beginning of recitation.

Spring 2017

Stony Brook University
Sabyasachi Mukherjee, Matthew Dannenberg

Problem Set 4

Differential Equations with Linear Algebra

Homework Problems

4.1. *Eigenvalues of Linear Maps.* (5+10 points)

(a) Let $T : V \to V$ be a linear map with an eigenvalue λ. Show that λ^2 is an eigenvalue of $T \circ T$.

(b) Let $\mathcal{C}^\infty(\mathbb{R})$ be the vector space of all infinitely differentiable real functions (with respect to addition and scalar multiplication of functions). Consider the linear map $T := \frac{d^2}{dx^2} : \mathcal{C}^\infty(\mathbb{R}) \to \mathcal{C}^\infty(\mathbb{R})$. For $\lambda > 0$, prove that any linear combination of $e^{x\sqrt{\lambda}}$ and $e^{-x\sqrt{\lambda}}$ is an eigenvector for λ.

4.2. *Computing Powers of Matrices.* (10 points)

Show that if $A = \begin{bmatrix} 0 & 2 \\ -1 & 3 \end{bmatrix}$, then $A^{10} = \begin{bmatrix} -1022 & 2046 \\ -1023 & 2047 \end{bmatrix}$.

(Hint: Write A as PDP^{-1}, where D is diagonal.)

4.3. *Inner Product or Not?.* (5 points)

Consider the vector space $\mathbb{R}^2$ with respect to usual addition and scalar multiplication of vectors. Does the formula $\langle \mathbf{x}, \mathbf{y} \rangle = x_1y_1 - x_2y_2$ (where $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = (y_1, y_2)$) define an inner product on $\mathbb{R}^2$?

4.4. *Recovering Angle from Length.* (10 points)

Prove that if $||\mathbf{x}|| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$ is the norm defined by an inner product $\langle \mathbf{x}, \mathbf{y} \rangle$, then $\langle \mathbf{x}, \mathbf{y} \rangle = \frac{1}{4}\left(||\mathbf{x} + \mathbf{y}||^2 - ||\mathbf{x} - \mathbf{y}||^2\right)$.

4.5. *Symmetric Matrices.* (10 points)

Let A be a real symmetric 2×2 matrix. Show that $\langle A(\mathbf{x}), \mathbf{y} \rangle = \langle \mathbf{x}, A(\mathbf{y}) \rangle$ for every $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$ (here $\langle\ ,\ \rangle$ denotes the usual dot product in $\mathbb{R}^2$).

4.6. *Finding an Orthonormal Basis.* (10 points)

Let $V = \mathcal{P}_2\,[0, 1]$ be the vector space of all real polynomials of degree at most 2 restricted to $[0, 1]$. If V is given the inner product

$$\langle f, g \rangle = \int_0^1 f(x)g(x)dx,$$

find an orthonormal basis for V.

(Hint: Apply Gram-Schmidt on the basis $\{1, x, x^2\}$.)

Due Date: Thursday, February 23, at the beginning of recitation.

Stony Brook University
Sabyasachi Mukherjee, Matthew Dannenberg

Differential Equations with Linear Algebra

Homework Problems

5.1. *Linear Differential Equation with Constant Coefficients.* (10+10 points)

Solve the following initial value problems (do not use trial solutions, use the method of repeated integration):

(a) $y'' + y' - 6y = 0,\ y(0) = 2,\ y'(0) = 2.$

(b) $y'' - 2y' + 2y = 0,\ y(\pi) = 2,\ y'(\pi) = 0.$

5.2. *Differential Equations with A Prescribed Solution.* (5+5 points)

Find linear differential equations of minimal order (with constant coefficients) that are satisfied by the following functions:

(a) $f(x) = 2xe^{-x} + e^{-x}.$

(b) $g(x) = 3\cos(4x) - 5e^{2x}\sin 3x.$

5.3. *Undetermined Coefficients.* (10 points)

Find the general solution of the following linear non-homogenous differential equation using the method of undetermined coefficients:

$y'' - 4y = 2e^{2x}.$

5.4. *Variation of Parameters.* (10 points)

For the following differential equation, find or guess a solution y_1 of the associated homogenous equation. Then determine $u(x)$ so that $y(x) = u(x)y_1(x)$ is a solution of the differential equation containing two arbitrary constants.

$x^2y'' - 3xy' + 3y = x^4,\ x > 0.$

(Hint: try $y_1 = x^n$, for some positive integer n)

5.5. *Linear Independence and Wronskian.* (10 points)

Compute the Wronskian of the functions $y_1(x) = e^{-3x}$, $y_2(x) = \cos 2x$, and $y_3(x) = \sin 2x$, and conclude that they are linearly independent (on $\mathbb{R}$).

5.6. (Bonus Problem) *Exploiting The Power of Wronskians.* (10 points)

Let $r_1,\ r_2$ be two distinct real roots of the quadratic equation $x^2 + px + q = 0$, where $p,\ q \in \mathbb{R}$. We have seen, using exponential functions as trial solutions, that $y_1(x) = e^{r_1 x}$ and $y_2(x) = e^{r_2 x}$ are solutions of the linear differential equation

$$\frac{d^2y}{dx^2} + p\frac{dy}{dx} + qy = 0. \tag{1}$$

(a) Prove that the Wronskian of $y_1(x)$ and $y_2(x)$ never vanishes on $\mathbb{R}$.

(b) Let $f(x)$ be an arbitrary solution of (1). Prove that the Wronskian of the functions $\{y_1(x), y_2(x), f(x)\}$ is identically zero. Now use a theorem from the lecture notes to conclude that $\{y_1(x), y_2(x)\}$ is a basis for the vector space of all solutions of (1). In particular, the space has dimension 2.

Remark. The assumption that r_1 and r_2 are distinct real numbers is unnecessary, with minor modifications the above proof goes through in the other cases (i.e. if $r_1 = r_2 \in \mathbb{R}$ or if r_1 and r_2 are complex conjugate) as well.

Due Date: Thursday, March 9, at the beginning of recitation.

Stony Brook University
Sabyasachi Mukherjee, Matthew Dannenberg

Differential Equations with Linear Algebra

Homework Problems

6.1. *Computing Inverse Laplace Transforms.* (7+7+7+9 points)

(a) Find the inverse Laplace transforms of the following functions (you may use any standard formula involving Laplace transforms listed in the book):

(a) $\frac{s^2-2s}{s^4+5s^2+4}$,

(b) $\frac{1}{s^2(s^2-1)}$,

(c) $\frac{s}{(s-3)(s^2+1)}$.

(b) Find $\mathcal{L}^{-1}(\frac{1}{(s^2+a^2)^2})$.

6.2. *Initial Value Problems via Laplace Transform.* (10+10 points)

(a) Solve the following initial value problem using Laplace transform:
$x'' + 4x' + 13x = te^{-t};\ x(0) = 0,\ x'(0) = 2.$

(b) Use the convolution theorem to derive the indicated solution $x(t)$ of the given initial value problem:
$x'' + 4x' + 13x = f(t);\ x(0) = 0,\ x'(0) = 0.$
$$x(t) = \frac{1}{3}\int_0^t f(t-u)e^{-2u}\sin(3u)du.$$

6.3. *Laplace Transform of Discontinuous Function.* (10 points)

(a) Define the Heaviside function

$$H(t) = \begin{cases} 0 & \text{for } t < 0 \\ 1 & \text{for } 0 \le t. \end{cases}$$

Show that $\mathcal{L}[H(t-a)](s) = \frac{1}{s}e^{-as}$.

(b) Solve the differential equation $y'' = H(t-a)$ $(0 < a)$, with initial conditions $y(0) = 1$, $y'(0) = 0$.

6.4. (Bonus Problem) *Bump Function.* (10 points)

Prove that the function

$$\Psi(x) = \begin{cases} \exp\left(-\frac{1}{1-x^2}\right) & \text{for } |x| < 1 \\ 0 & \text{otherwise} \end{cases}$$

is everywhere differentiable and that its derivative is continuous. (In fact, the function has continuous derivatives of all orders. This is an example of a bump function.)

Due Date: Thursday, March 23, at the beginning of recitation.

Stony Brook University
Sabyasachi Mukherjee, Matthew Dannenberg

Differential Equations with Linear Algebra

Homework Problems

7.1. *Existence and Uniqueness of Solutions.* (10 points)

(a) Explain which part(s) of the existence and uniqueness theorem (of solutions of differential equations) fail(s) to apply to the initial value problem

$$\dot{x} = \begin{cases} \sqrt{x}, & \text{if } x \geq 0, \\ 0, & \text{if } x < 0, \end{cases}$$

and $x(0) = 0$.

(b) Find two distinct solutions of this equation.

7.2. *Inverse of a Matrix Using Cayley-Hamilton Theorem.* (10 points)

Find the inverse of the matrix $\begin{bmatrix} 2 & 4 & 8 \\ 1 & 0 & 0 \\ 1 & -3 & -7 \end{bmatrix}$ using the Cayley-Hamilton theorem.

7.3. *Computing Matrix Exponential: Brute Force Method.* (10 points)

Find e^{tA} by computing the successive terms $I,\ tA,\ t^2A^2/2!, \cdots$ in the series definition, where $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

7.4. *Solving Systems of Linear Differential Equations by Two Different Methods.* (20 points)

Consider the system of linear differential equations

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} y \\ -6x + 5y \end{bmatrix}, \ (x(0), y(0)) = (1, 2). \tag{1}$$

(a) Solve (1) using the method of elimination.

(b) Solve (1) using the matrix exponential method.

(c) Did we know a priori that the two methods would produce the same solution?

7.5. *Interplay between The Elimination Method and The Matrix Exponential Method.* (20 points)

Consider the system of linear differential equations

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} 9x - 4y \\ 4x + y \end{bmatrix}. \tag{2}$$

(a) Find the general solution of (2) using the method of elimination.

(b) Now use the general solution obtained in part (a) to find particular solutions satisfying $(x(0), y(0)) = (1, 0)$ and $(x(0), y(0)) = (0, 1)$.

(c) Use the results of part (b) to find e^{tA}, where $A = \begin{bmatrix} 9 & -4 \\ 4 & 1 \end{bmatrix}$

7.6. *Matrix Exponential of a 2×2 Matrix with Trace Zero.* (10 points)

Let $A \in M_2(\mathbb{R})$ with $\operatorname{trace}(A) = 0$. Show that $exp(A) = \cos(\sqrt{\det(A)})I + \frac{\sin(\sqrt{\det(A)})}{\sqrt{\det(A)}}A$, where $\frac{\sin(\sqrt{\det(A)})}{\sqrt{\det(A)}}$ is interpreted as 1 when $\det(A) = 0$, in accordance with the limit $\lim_{t\to 0} \frac{\sin t}{t} = 1$.

Hint: Recall the power series expansions of sin and cos.

7.7. *Matrix Exponential Using N-D Decomposition.* (10 points)

Find e^A using the N-D decomposition of A, where $A = \begin{bmatrix} 2 & 0 & 1 \\ -1 & 3 & 1 \\ -1 & 0 & 4 \end{bmatrix}$.

7.8. (Bonus Problem) *An Application of Jordan Canonical Forms.* (20 points)

The goal of this problem is to prove that

$$\det\left(e^A\right) = e^{trace(A)},$$

for every $A \in M_n(\mathbb{C})$. Let us start with a couple of important definitions.

Definition 1 *Two matrices A and B are called* similar *over $\mathbb{C}$ if there exists an invertible matrix $P \in M_n(\mathbb{C})$ such that $A = P^{-1}AP$.*

Definition 2 *Let $A = (a_{ij}) \in M_n(\mathbb{C})$. The characteristic polynomial of A is defined as $\det(\lambda I - A)$. Clearly, the characteristic polynomial of A is a degree n polynomial in λ.*

(a) Show that similar matrices have the same determinant and the same characteristic polynomial. Conclude that similar matrices have the same eigenvalues with the same multiplicities.

(b) Let $A = (a_{ij}) \in M_n(\mathbb{R})$, $\lambda^n + p_1\lambda^{n-1} + \cdots + p_n$ be the characteristic polynomial of A, and $\{\lambda_1, \cdots, \lambda_n\}$ be the set of all eigenvalues (not necessarily distinct) of A. Show that $-\sum_{i=1}^{n} \lambda_i = p_1 = -\sum_{i=1}^{n} a_{ii}$. Conclude that similar matrices have the same trace, which is equal to the sum of all the (common) eigenvalues. In particular, we have $e^{trace(A)} = \exp\left(\sum_{i=1}^{n} \lambda_i\right)$.

(c) Recall that every complex $n \times n$ matrix is similar to its Jordan canonical form. Using the Jordan canonical form of A, show that e^A is similar to an upper triangular matrix with entries $e^{\lambda_1}, e^{\lambda_2}, \cdots, e^{\lambda_n}$ on its principal diagonal. Conclude that $\det\left(e^A\right) = \exp\left(\sum_{i=1}^{n} \lambda_i\right)$.

(d) Quod erat demonstrandum.

Due Date: Monday, April 10, at the beginning of class.

Differential Equations with Linear Algebra

Homework Problems

8.1. *Method of Undetermined Coefficients.* (15 points)

Apply the method of undetermined coefficients to find the general solution of the following system.

$x' = 6x - 7y + 10,\ y' = x - 2y - 2e^{-t}.$

8.2. *Method of Variation of Parameters.* (15 points)

Apply the method of variation of parameters to solve the initial value problem

$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 2 & -4 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 36t^2 \\ 6t \end{bmatrix},\ \begin{bmatrix} x(0) \\ y(0) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$

8.3. *Limitations of Power Series Method.* (10 points)

Show that the power series method fails to yield a power series solution of the form $y = \sum_{n=0}^{\infty} c_n x^n$ for the differential equation $x^2 y' + y = 0$.

8.4. *Truncated Power Series Solution.* (20 points)

Find the first six non-zero terms of the power series solution of the following differential equation around $x = 0$.

$(x^2 - 4)y'' + 3xy' + y = 0,\ y(0) = 4,\ y'(0) = 1$

8.5. (Bonus Problem) *Power Series Solution Via Recurrence Relation.* (15 points)

Find the general solution in powers of x of the following differential equation. State the recurrence relation and the radius of convergence of the power series.

$(x^2 - 1)y'' + 4xy' + 2y = 0.$

Due Date: Thursday, April 27, at the beginning of class.

Stony Brook University
Sabyasachi Mukherjee, Matthew Dannenberg

Spring 2017

Problem Set 9

Differential Equations with Linear Algebra

Homework Problems

9.1. *Stability Analysis-I.* (10 points)

Carry out a stability analysis for the equilibrium point $(0, 0)$ of the following system.

$$\frac{dx}{dt} = 5x - 3y + y(x^2 + y^2), \quad \frac{dy}{dt} = 5x + y(x^2 + y^2)$$

9.2. *Stability Analysis-II.* (15 points)

Find all equilibrium points of the given system, and carry out a stability analysis for each equilibrium point.

$$\frac{dx}{dt} = xy - 2, \quad \frac{dy}{dt} = x - 2y$$

9.3. *Stability Analysis and Bifurcation.* (15 points)

We discussed the following theorem in class.

THEOREM 1 Stability of Linear Systems

Let λ_1 and λ_2 be the eigenvalues of the coefficient matrix $\mathbf{A}$ of the two-dimensional linear system

$$\begin{aligned} \frac{dx}{dt} &= ax + by, \\ \frac{dy}{dt} &= cx + dy \end{aligned} \tag{11}$$

with $ad - bc \neq 0$. Then the critical point $(0, 0)$ is

1. Asymptotically stable if the real parts of λ_1 and λ_2 are both negative;
2. Stable but not asymptotically stable if the real parts of λ_1 and λ_2 are both zero (so that $\lambda_1, \lambda_2 = \pm qi$);
3. Unstable if either λ_1 or λ_2 has a positive real part.

The following problem, which discusses the behavior of an equilibrium point for various parameters, is an interesting application of the above theorem.

First note that the characteristic equation of the 2×2 matrix $\mathbf{A}$ can be written in the form $\lambda^2 - T\lambda + D = 0$, where D is the determinant of $\mathbf{A}$ and the trace T of the matrix $\mathbf{A}$ is the sum of its two diagonal elements. Then apply Theorem 1 to show that the type of the critical point $(0, 0)$ of the system $\mathbf{x}' = \mathbf{A}\mathbf{x}$ is determined—as indicated in Fig. 1 —by the location of the point (T, D) in the *trace-determinant plane* with horizontal T-axis and vertical D-axis.

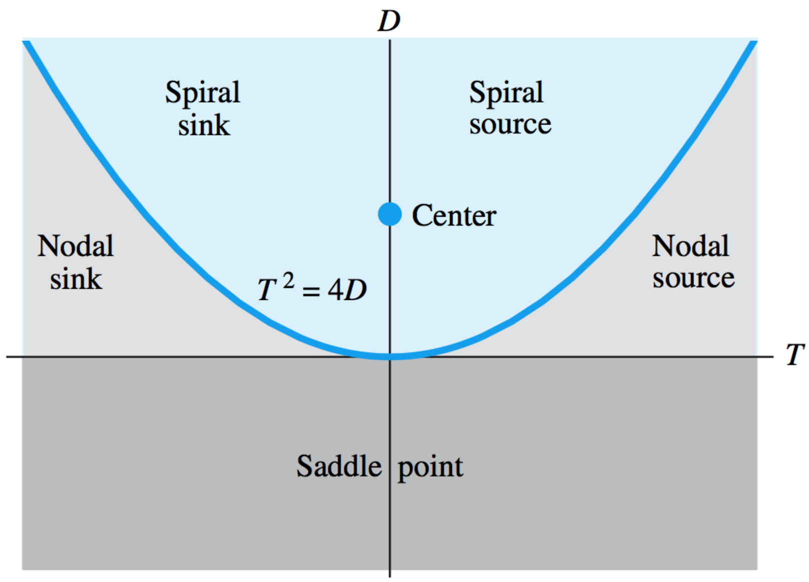


FIGURE 1. The critical point $(0, 0)$ of the system $\mathbf{x}' = \mathbf{A}\mathbf{x}$ is a

- spiral sink or source if the point (T, D) lies above the parabola $T^2 = 4D$ but off the D-axis;
- stable center if (T, D) lies on the positive D-axis;
- nodal sink or source if (T, D) lies between the parabola and the T-axis;
- saddle point if (T, D) lies beneath the T-axis.

Due Date: Thursday, May 4, at the beginning of recitation.