专业阅读

主讲 杨勤民

(Email: qmyang@ecust.edu.cn)

华东理工大学数学系

2020年3月～7月

专业阅读课程介绍

我们这门课程是针对数学系各个专业开设的专业英语课程. 一般的英语课程都包括听、说、读、写四个部分, 而我们的课程没这么全面, 重点在于引导同学们阅读国外大学数学教材中的部分章节, 积累数学专业英语阅读和翻译的基本经验, 为将来的工作和学习奠定基础.

本门课程没有合适的教材, 我根据教学大纲, 编写了一本《数学专业英语阅读》讲义, 讲义里面含有我们将要讲授的所有内容, 请同学们自己去复印.

课程重点

我在国外大学数学核心课程，如微积分、复分析、代数学、微分方程、几何学、概率论与数理统计、运筹学和数值分析等课程中挑选了一些基础章节编入了讲义. 同学们对里面涉及的数学知识应当是非常熟悉的, 或者说曾经是非常熟悉的, 这里可以顺带复习一下, 但不是我们课程的重点. 我们课程的重点在于如何理解这些英文, 以及如何将它们翻译成恰当的中文. 同学们可以把这些章节当成精读内容, 仔细体会在阅读和翻译数学专业英语过程中涉及的基本方法和技巧, 举一反三, 切实提高自己专业英语的阅读和理解能力.

课程的预备知识

课程中还包括了一些预备知识, 比如数学英文表达式的读法, 数学专业英语阅读与翻译的要求和基本步骤等, 这些是我们在学习本门课程时经常要用到的.

同时, 正确地读出数学英文表达式在作数学英文报告时也是必不可少的.

数学专业英语的写作与毕业设计

在课程的最后还将讲授英语数学论文的结构和书写要求, 以及如何撰写论文的英文摘要. 希望这些内容为同学们将来写毕业论文或写英文论文有所帮助.

毕业设计包括开题报告和毕业论文两部分, 其中开题报告要求同学们将导师指定的相关英文论文(字符数约4万)翻译成中文, 这一要求将进一步考查和强化同学们专业英语的阅读和翻译水平.

学位英语考试和数学专业英语核心词汇

同学们取得学士学位除了修完所有应该完成的课程之外, 还要通过一个学位英语考试. 学位英语考试分为听力、阅读、词汇翻译和段落翻译四个部分. 其中听力和阅读这2部分由外语学院命题, 词汇翻译和段落翻译由专业英语老师命题. 词汇翻译占10%, 包括10个英译汉, 10个汉译英, 这些词汇来源于讲义后面的数学专业英语词汇. 请同学们务必有计划地记住这些词汇. 我们这门课程的期末考试也要考这些词汇. 段落翻译要将4个英文段落中的划线部分翻译成中文, 占40%. 做段落翻译时可参考一本纸质词典.

华东理工大学专业阅读(QQ群: 515854632)

专业阅读

主讲 杨勤民

(Email: qmyang@ecust.edu.cn)

华东理工大学数学系

2020年3月～7月

Pronunciation of mathematical expressions

1. Logic
2. Sets
3. Real numbers
4. Linear algebra
5. Functions
6. Greek letters
7. Abbreviations

Pronunciation of mathematical expressions

The pronunciations of the most common mathematical expressions are given in the list below. In general, the shortest versions are preferred (unless greater precision is necessary).

1. Logic

$\exists$	there exists	$\forall$	for all
$\neg$	not	$\vee$	or
$\wedge$	and	$\to$	implies

$p \Longrightarrow q$	p implies q; if p, then q
$p \Longleftrightarrow q$	p if and only if q;
	p is equivalent to q;
	p and q are equivalent

2. Sets

$\emptyset$	the empty set
$\mathbb{R}$	the set of real numbers
$\mathbb{N}$	the set of natural numbers
$\mathbb{Z}$	the set of integers
$\mathbb{C}$	the set of complex numbers
$\mathbb{Q}$	the set of rational numbers
$x \in A$	x in A; x belongs to A; x belonging to A;
	x is an element (or a member) of A
$x \notin A$	x does not belong to A;
	x is not an element (or a member) of A

2. Sets(续)

A^c	the complement of A
$A \subset B$	A is contained in B; A is a subset of B
$A \supset B$	A contains B; B is a subset of A
$A \cap B$	A cap B; A meet B;
	A intersection B; A intersected with B
$A \cup B$	A cup B; A join B; A union B
$A \setminus B$	A minus B; the difference between A and B
$A \times B$	A times B; A cross B;
	the Cartesian product of A and B

3. Real numbers

$+$	plus; positive
$-$	minus; negative
$\times$	multiplied by; times
$\div$, $/$	divided by
$x+1$	x plus one
$x-1$	x minus one
$x \pm 1$	x plus or minus one
xy	xy; x multiplied by y
$(x-y)(x+y)$	x minus y, x plus y;
	x minus y into x plus y
$(x+y)$	bracket x plus y bracket closed

3. Real numbers(续)

$\frac{x}{y}$	x over y
$x = 5$	x equals 5; x is equal to 5
$x \approx 5$	x (is) approximately equal to 5
$x \neq 5$	x (is) not equal to 5
$x \equiv y$	x is equivalent to (identical with) y
$x \not\equiv y$	x is not equivalent to (identical with) y
$x \propto y$	x is proportional to y
$x > y$	x (is) greater than y
$x \gg y$	x (is) far greater than y
$x \geq y$	x (is) greater than or equal to y;
	x (is) no less than y

3. Real numbers(续)

$x < y$	x (is) less than y
$x \ll y$	x (is) far less than y
$x \leq y$	x (is) less than or equal to y;
	x (is) no greater than y
$0 < x < 1$	zero (is) less than x (is) less than 1
$\|x\|$	absolute value of x
x^2	x squared; x (raised) to the power 2
x^3	x cubed
x^4	x to the fourth; x to the power four
x^n	x to the nth; x to the power n
x^{-n}	x to the (power) minus n

3. Real numbers(续)

x^{-1}	the reciprocal of x; x inverse
$\sqrt{x}$	(square) root x; the square root of x
$\sqrt[3]{x}$	cube root (of) x
$\sqrt[4]{x}$	fourth root (of) x
$\sqrt[n]{x}$	nth root (of) x
$(x+y)^2$	x plus y all squared
$(\frac{x}{y})^2$	x over y all squared
$n!$	n factorial
$\hat{x}$	x hat
$\bar{x}$	x bar
$\tilde{x}$	x tilde

3. Real numbers(续)

x_i	xi; x subscript i; x suffix i; x sub i
x_i^j	xij [if j is an index, not an exponent!]
$x_1, \cdots, x_n$	$x1$ up to xn
$\sum_{i=1}^{n} a_i$	the sum from i equals one to n ai;
	the sum as i runs from 1 to n of ai;
	the sum for i (running) from 1 to n of ai
$\prod_{n=1}^{\infty} x_n$	the product of xn from n equals 1 to infinity
0.01	zero point zero one; nought point nought one

3. Real numbers(续)

$86.\dot{8}$	eighty six point eight recurring;
	eighty six point eight, eight recurring
$6.6\dot{8}\dot{6}$	six point six eight six, eight six recurring
$\frac{1}{2}, \frac{1}{3}, \frac{1}{4}$	one half, one third, one quarter (one fourth)
$\frac{2}{3}, \frac{9}{100}$	two thirds, nine hundredths
$\frac{7}{1203}$	seven over twelve hundred and three
5%	five percent
30°	thirty degrees
$a : b = c : d$	the ratio of a to b equals that of c to d
$a : b :: c : d$	a is to b as c is to d
$7 \div 3 = 2$余1	3 into 7 goes 2 times and 1 remainder

4. Linear algebra

$\|\boldsymbol{x}\|$	the norm (or modulus) of $\boldsymbol{x}$
$\overrightarrow{OA}$	OA; vector OA
$\overline{OA}$	OA; segment OA
$\boldsymbol{A}^{\mathrm{T}}, \boldsymbol{A}'$	$\boldsymbol{A}$ transpose; the transpose of $\boldsymbol{A}$
$\boldsymbol{A}^{-1}$	$\boldsymbol{A}$ inverse; the inverse of $\boldsymbol{A}$
$\|\boldsymbol{A}\|$	the determinant of $\boldsymbol{A}$

5. Functions

$f(x)$	fx; f of x; the function f of x
$f: S \to T$	a function f from S to T
$x \mapsto y$	x maps to y; x is sent (or mapped) to y
$f = u \circ v$	f is the composite (or composition) of u and v
$\lim\limits_{x \to 0}$	the limit as x approaches zero;
	the limit as x tends/goes to zero
$\lim\limits_{x \to 0^+}$	the limit as x approaches zero from above
$\lim\limits_{x \to 0^-}$	the limit as x approaches zero from below

5. Functions(续)

$f'(x)$	f prime x; f dash x;
	the first derivative of f with respect to x
$f''(x)$	f double-prime x; f double-dash x;
	the second derivative of f with respect to x
$f'''(x)$	f triple-prime x; f triple-dash x;
	the third derivative of f with respect to x
$f^{(4)}(x)$	f four x;
	the fourth derivative of f with respect to x

5. Functions(续)

$\frac{\partial f}{\partial x_1}$	the partial (derivative) of f with respect to $x1$; df by d$x1$
$\frac{\partial^2 f}{\partial x_1^2}$	the second partial (derivative) of f with respect to $x1$
$\int_a^b f(x)\mathrm{d}x$	the integral from a to b (of) fx dx
$\int_0^\infty$	the integral from zero to infinity
$\iint_D$, $\iiint_D$	the double/triple integral over the domain D
$\exp(x)$, e^x	the exponential of x, e to the x
$\log_\mathrm{e} y$, $\ln y$	the logarithm of y to the base e; natural logarithm (of) y

Individual mathematicians often have their own way of pronouncing mathematical expressions and in many cases there is no generally accepted "correct" pronunciation. Distinctions made in writing are often not made explicit in speech; thus the sounds fx may be interpreted as any of: fx, $f(x)$, f_x, FX, $\overline{FX}$, $\overrightarrow{FX}$. The difference is usually made clear by the context. It is only when confusion may occur, or where he/she wishes to emphasize the point, that the mathematician will use the longer forms: f multiplied by x, the function f of x, f subscript x, line FX, the length of the segment FX, vector FX.

Similarly, a mathematician is unlikely to make any distinction in speech (except sometimes a difference in intonation or length of pauses) between pairs such as the following:

$$\begin{array}{lll} x+(y+z) & \text{and} & (x+y)+z, \\ \sqrt{ax}+b & \text{and} & \sqrt{ax+b}, \\ a^n-1 & \text{and} & a^{n-1}. \end{array}$$

The primary reference has been David Hall with Tim Bowyer, Nucleus, English for Science and Technology, Mathematics, Longman 1980. Glen Anderson and Matti Vuorinen have given good comments and supplements.

6. Greek letters

α	alpha [′ælf∂]	β	beta [′beit∂]
γ, Γ	gamma [′gæm∂]	δ, Δ	delta [′delt∂]
ϵ, ε	epsilon [′epsil∂n]	ζ	zeta [′zi:t∂]
η	eta [′it∂]	$\theta, \vartheta, \Theta$	theta [′θi:t∂]
ι	iota [ai′∂ut∂]	κ	kappa [′kæp∂]
λ, Λ	lambda [′læmd∂]	μ	mu [mju:]
ν	nu [nju:]	ξ, Ξ	xi [ksai]
π, ϖ, Π	pi [pai]	ρ, ϱ	rho [r∂u]
$\sigma, \varsigma, \Sigma$	sigma [′sigm∂]	τ	tau [tao]
υ, Υ	upsilon [′ju:psil∂n]	ϕ, φ, Φ	phi [fai]
χ	chi [kai]	ψ, Ψ	psi [psai]
ω, Ω	omega [′∂umig∂]		

7. Abbreviations

etc.	et cetra (= and so on)
e.g.	for example
i.e.	that is
viz.	namely
w.r.t.	with respect to

专业阅读

主讲 杨勤民

(Email: qmyang@ecust.edu.cn)

华东理工大学数学系

2020年3月～7月

Mathematical English reading and translati

阅读与翻译之间的关系

数学专业英语的阅读

数学专业英语的翻译

人名的翻译

数学专业英语的阅读与翻译

阅读是为了获取信息，理解和掌握专业的内容。

翻译是把原文的内容选用适当的中文完整地表达出来。

翻译的前提是通过阅读获得正确的理解，而翻译的过程有助于加深阅读理解。

两者紧密关联，相辅相成。

数学专业英语阅读

除了掌握基本词汇与语法基本知识外，阅读者还应具备运用语言知识和数学知识的能力。这种能力包括：

1. 根据上下文来确定词义和猜测词义；
2. 正确理解句与句、段与段之间的逻辑关系；
3. 对数学内容做必要的、基本的逻辑推理；
4. 归纳段落大意和全文主题。

加强阅读训练

为了提高阅读能力，应该从以下三方面加强训练：

1. 努力练好英语的基本功，掌握基本语法、习惯用法和常用单词的基本用法，逐步扩大常用数学单词和词组的识记和使用范围；

2. 多读多练，在阅读专业英语的过程中，逐步掌握数学专业英语的特点，不断总结提高；

3. 努力掌握正确的阅读方法。

1. 数学专业英语翻译的要求

对于初学者，翻译的基本要求有两个：

第一个要求也是最重要的要求就是准确，即忠实于原文；

第二个要求是译文的表达必须通顺且符合专业规范，包括词句的选用要得当，意义要明确，句子要流畅，让读者易读，并力求简练。

直译与意译

为了达到这两个要求，有的地方可直译，必要时进行意译。

所谓直译，就是在译文语言条件许可时，在译文中既保留原文的内容，又保留原文的形式。

由于每个民族都有自己的词汇、词法结构和表达方法，当原文的思想内容与译文的表达方式有矛盾而不宜采用直译时，就应该采用意译，即按原文的意思（含义）翻译，而不拘泥于原文的形式，有时需要根据译文的表达习惯增加或减少一些词语或作其它必要的变动。

翻译有三忌

一忌错误理解原文或随意添加或遗漏；

二忌逐词死译，勉强凑合，意思含糊不清；

三忌文句不符合译文表达习惯和专业规范，译味太浓。

2. 翻译的步骤

第一步：正确地理解原文。

第二步：恰当地表达成中文。

翻译是一种创造性的工作，译者必须在正确理解原文的基础上选用恰当的中文词句将原文的意思既准确又通顺地表达出来，让它符合汉语的表达习惯和专业的规范。

第三步：认真地进行校对。

认真校对

为了力求译文的表达准确无误，需要译者:

将译文与原文作认真的比较和校对，逐句推敲；

并把公式、图表等作仔细对比，对于无十分把握的句子、词语或有疑惑之处一定要通过各种方式进行查实并作必要的纠正。

同时，对中文的表达也常要进一步整理和完善，力争通顺易读，精益求精。

校对是理解和表达的进一步深化，是使译文符合要求的一个必不可少的步骤。

3. 人名的翻译

外国数学家的名字常以两种形式出现在英文版文献中。

一种是以本国文字出现，通常是拉丁文语系（包括英语）的国家的数学家名字，如Newton（牛顿，英文），Cauchy（柯西，法文），Poisson（泊松，法文），Hilbert（希尔伯特，德文）等；

另一种是按读音翻译（简称“音译”）成英文，如Urysohn（乌雷松，俄文YpbICOH的音译）。

不论上述哪一种形式，译成中文时均采用音译。

翻译人名的三个要点

(一) 著名的数学家的名字按《数学名词》或《中国大百科全书》数学卷的规定统一译法；

(二) 其他数学家的名字一般不予译出，而把姓名原文抄录在译文中（姓的部分全写出，名的部分可缩写），如果要译出，必须在译名后加括号注明原文；

(三) 在同一篇译文中数学家的姓名要么全部不予译出，要么全部译出且对其中不大为人所知的人，特别是不在中国大百科全书数学卷出现的姓名后加括号注上原文。

专业阅读

主讲 杨勤民

(Email: qmyang@ecust.edu.cn)

华东理工大学数学系

2020年3月～7月

Limit

The characteristic feature of calculus is its use of limiting processes. Differentiation and integration involve certain notions of passage to a limit. A fuller discussion of ideas about limits is presented later on in this chapter. Here we wish to touch on only one limit notion, that of the limit of a real function of one real variable. This notion is fundamental in the definition of a derivative.

Suppose f is a function which is defined for all values of x near the fixed value x_0, and possibly, though not necessarily, at x_0 as well. We wish to attach a clear meaning to the statement: $f(x)$ approaches A (or tends to the limit A) as x approaches x_0. The symbolic form of the statement is

$$\lim_{x \to x_0} f(x) = A. \tag{2.1}$$

The symbol A is understood to stand for some particular real number. The arrow is used as a symbol for the word "approaches".

Sometimes Assertion (2.1) is expressed in the form $f(x) \to A$ as $x \to x_0$. Here are three typical examples of statements of this kind: (a) $x^3 \to 8$ as $x \to 2$, (b) $(x-1)^{1/2} \to 3$ as $x \to 10$, (c) $\log_{10} x \to 2$ as $x \to 100$.

Definition 1. The assertion (2.1) means that we can insure that the absolute value $|f(x)-A|$ is as small as we please merely by requiring that the absolute value $|x-x_0|$ be sufficiently small and different from zero. This verbal statement is expressible in terms of inequalities as follows: Suppose that ε is any positive number. Then there is some positive number δ such that

$$|f(x)-A|<\varepsilon \quad \text{if} \quad 0<|x-x_0|<\delta. \tag{2.2}$$

Note that $0<|x-x_0|$ is the same as $x \neq x_0$. Note also that $|f(x)-A|<\varepsilon$ is the same as $A-\varepsilon<f(x)<A+\varepsilon$, and $|x-x_0|<\delta$ the same as $x_0-\delta<x<x_0+\delta$.

We can give a geometrical portrayal of the inequalities (2.2). Let the points (x, y) with $y = f(x)$ be located on a rectangular coordinate system, also locate the point (x_0, A). For any $\varepsilon > 0$ draw the two horizontal lines $y = A \pm \varepsilon$. Now Assertion (2.1) means that by choosing δ small enough, those points of the graph of $y = f(x)$ which lie between the two vertical lines $x = x_0 \pm \delta$ and not on the line $x = x_0$ will also lie between the horizontal lines $y = A \pm \varepsilon$. Figure 2.1 shows a specimen of this situation. The diagram also shows how δ may have to be made smaller as ε becomes smaller.

It is to be emphasized that Assertion (2.1) places no restrictions whatever on the value of f at x_0, in case it is defined at that point.

Appreciation of the formal definition of the meaning of Assertion (2.1) takes time and experience. The formal definition is the basis for exact reasoning on matters involving the limit concept. But it is also quite important to develop an intuitive understanding of the notion of a limit. This may be done by considering a large number of illustrative examples and by observing the way in which the limit concept is used in the development of calculus. One needs to learn by example how a function $f(x)$ may fail to approach a limit as x approaches x_0.

The variable x may approach x_0 from either of two sides. Let us use $x \to x_0^+$ to indicate that x approaches x_0 from the right, and $x \to x_0^-$ to indicate approach from the left. The conditions for $\lim_{x \to x_0} f(x) = A$ are then that $f(x) \to A$ as $x \to x_0^+$ and also $f(x) \to A$ as $x \to x_0^-$. In terms of inequalities the meaning of $f(x) \to A$ as $x \to x_0^+$ is this: to any $\varepsilon > 0$ corresponds some $\delta > 0$ such that $|f(x) - A| < \varepsilon$ if $x_0 < x < x_0 + \delta$. The meaning of $f(x) \to A$ as $x \to x_0^-$ may be expressed in a similar way.

The limit of $f(x)$ as $x \to x_0$ may fail to exist because:

(a) The limits from right and left exist but are not e-qual. This is the case with $f(x) = 1 + \frac{|x|}{x}$ where $f(x) \to 2$ as $x \to 0^+$ and $f(x) \to 0$ as $x \to 0^-$.

(b) The values of $f(x)$ may get larger and larger (tend to infinity) as $x \to x_0$ from one side or the other, or from both sides. This is the case with $f(x) = 1/x$ as $x \to 0$.

(c) The values of $f(x)$ may oscillate infinitely often, approaching no limit. This is the case with $f(x) = \sin(1/x)$ which oscillates infinitely often between -1 and +1 as $x \to 0$ from either side.

作业

请将下列段落翻译成中文(在作业本上直接写翻译结果, 每两题之间空一行, 不用抄原文)

1. Page3Line1～Page4Line4

2. Page8Line1～Page8Line9

专业阅读

主讲 杨勤民

(Email: qmyang@ecust.edu.cn)

华东理工大学数学系

2020年3月～7月

The calculus of vector-valued functions

Now that we have defined vector-valued functions, we need some tools for examining them. In this section, we begin to explore the calculus of vector-valued functions. As with scalar functions, we begin with the notion of limit and progress to continuity, derivatives and finally, integrals. Take careful note of how our presentation parallels that from Chapters 1,2 and 4. We follow this same kind of progression again when we examine functions of several variables in Chapter 12.

We define everything in this section in terms of vector-valued functions in three dimensions. The definitions can be interpreted for vector-valued functions in two dimensions in the obvious way, by simply dropping the third component everywhere.

For a vector-valued function $\boldsymbol{r}(t) = \langle f(t), g(t), h(t)\rangle$, if we write $\lim_{t\to a} \boldsymbol{r}(t) = \boldsymbol{u}$, we mean that as t gets closer and closer to a, the vector $\boldsymbol{r}(t)$ is getting closer and closer to the vector $\boldsymbol{u}$. If we write $\boldsymbol{u} = \langle u_1, u_2, u_3\rangle$, this means that

$$\lim_{t\to a} \boldsymbol{r}(t) = \lim_{t\to a} \langle f(t), g(t), h(t)\rangle = \boldsymbol{u} = \langle u_1, u_2, u_3\rangle.$$

Notice that for this to occur, we must have that $f(t)$ is approaching u_1, $g(t)$ is approaching u_2 and $h(t)$ is approaching u_3. In view of this, we make the following definition.

Definition 2. For a vector-valued function $\boldsymbol{r}(t) = \langle f(t), g(t), h(t)\rangle$, the limit of $\boldsymbol{r}(t)$ as t approaches a is given by

$$\begin{aligned}\lim_{t\to a}\boldsymbol{r}(t) &= \lim_{t\to a}\langle f(t), g(t), h(t)\rangle \\ &= \langle \lim_{t\to a} f(t), \lim_{t\to a} g(t), \lim_{t\to a} h(t)\rangle, \end{aligned} \tag{2.3}$$

provided all of the indicated limits exist. If any of the limits indicated on the righthand side of (2.3) fail to exist, then $\lim_{t\to a}\boldsymbol{r}(t)$ does not exist.

In the following example, we see that calculating a limit of a vector-valued function simply consists of calculating three separate limits of scalar functions.

Example 1. Find $\lim_{t\to 0}\langle t^2+1, 5\cos t, \sin t\rangle$.

Solution Recall that each of the component functions is continuous (for all t) and so, we can calculate their limits simply by substituting the values for t. We have

$$\lim_{t\to 0}\langle t^2+1, 5\cos t, \sin t\rangle$$

$$= \langle \lim_{t\to 0}(t^2+1), 5\lim_{t\to 0}\cos t, \lim_{t\to 0}\sin t\rangle = \langle 1, 5, 0\rangle.$$

Example 2. Find $\lim_{t\to 0}\langle e^{2t}+5,\, t^2+2t-3,\, \frac{1}{t}\rangle$.

Solution Notice that the limit of the third component is $\lim_{t\to 0}\frac{1}{t}$, which does not exist. So, even though the limits of the first two components exist, the limit of the vector-valued function does not exist.

Recall that for a scalar function f, we say that f is continuous at a if and only if $\lim_{t \to a} f(t) = f(a)$. That is, a scalar function is continuous at a point whenever the limit and the value of the function are the same. We define the continuity of vector-valued functions in the same way.

Definition 3. The vector-valued function

$$\boldsymbol{r}(t) = \langle f(t), g(t), h(t) \rangle$$

is continuous at $t = a$ whenever $\lim_{t \to a} \boldsymbol{r}(t) = \boldsymbol{r}(a)$ (i.e., whenever the limit and the value of the vector-valued function are the same).

Recall that in Chapter 2, we defined the derivative of a scalar function f to be $f'(t) = \lim\limits_{h\to 0} \frac{f(t+h)-f(t)}{h}$. Replacing h by Δt, we can rewrite this as $f'(t) = \lim\limits_{\Delta t\to 0} \frac{f(t+\Delta t)-f(t)}{\Delta t}$. You may be wondering why we want to change from a perfectly nice variable like h to something more unusual like Δt. The only reason is that we want to use the notation to emphasize that Δt is an increment of the variable t. In Chapter 12, we'll define partial derivatives of functions of more than one variable, where we'll use this type of notation to make it clear which variable is being incremented.

We now define the derivative of a vector-valued function in the expected way.

Definition 4. The derivative $\boldsymbol{r}'(t)$ of the vector-valued function $\boldsymbol{r}(t)$ is defined by

$$\boldsymbol{r}'(t) = \lim_{\Delta t\to 0} \frac{\boldsymbol{r}(t+\Delta t)-\boldsymbol{r}(t)}{\Delta t} \tag{2.4}$$

for any value of t for which the limit exists. When the limit exists for $t = a$, we say that $\boldsymbol{r}(t)$ is differentiable at $t = a$.

Fortunately, you will not need to learn any new differentiation rules, as the derivative of a vector-valued function is found directly from the derivative of the individual components, as we see in the following result.

Theorem 1. Let $\boldsymbol{r}(t) = \langle f(t), g(t), h(t) \rangle$ and suppose that the components f, g and h are all differentiable for some value of t. Then $\boldsymbol{r}(t)$ is also differentiable at that value of t and its derivative is given by $\boldsymbol{r}'(t) = \langle f'(t), g'(t), h'(t) \rangle$.

Proof From the definition of derivative of a vector-valued function (Equation 2.4), we have

$$
\begin{aligned}
\boldsymbol{r}'(t) &= \lim_{\Delta t \to 0} \frac{\boldsymbol{r}(t+\Delta t) - \boldsymbol{r}(t)}{\Delta t} \\
&= \lim_{\Delta t \to 0} \frac{\langle f(t+\Delta t), g(t+\Delta t), h(t+\Delta t) \rangle - \langle f(t), g(t), h(t) \rangle}{\Delta t} \\
&= \lim_{\Delta t \to 0} \frac{\langle f(t+\Delta t) - f(t), g(t+\Delta t) - g(t), h(t+\Delta t) - h(t) \rangle}{\Delta t}
\end{aligned}
$$

where we have used the definition of vector substraction. Distributing the scalar $\frac{1}{\Delta t}$ into each component and using the definition of limit of a vector-valued function (2.3), we have

$$\boldsymbol{r}'(t) = \lim_{\Delta t \to 0} \frac{\langle f(t+\Delta t)-f(t), g(t+\Delta t)-g(t), h(t+\Delta t)-h(t)\rangle}{\Delta t}$$
$$= \lim_{\Delta t \to 0} \langle \frac{f(t+\Delta t)-f(t)}{\Delta t}, \frac{g(t+\Delta t)-g(t)}{\Delta t}, \frac{h(t+\Delta t)-h(t)}{\Delta t} \rangle$$
$$= \langle \lim_{\Delta t \to 0} \frac{f(t+\Delta t)-f(t)}{\Delta t}, \lim_{\Delta t \to 0} \frac{g(t+\Delta t)-g(t)}{\Delta t}, \lim_{\Delta t \to 0} \frac{h(t+\Delta t)-h(t)}{\Delta t} \rangle$$
$$= \langle f'(t), g'(t), h'(t) \rangle$$

where in the last step we recognized the definition of the derivatives of each of the component function f, g and h.

Notice that thanks to Theorem 1, in order to differentiate a vector-valued function, we need only differentiate the individual component functions, using the usual rules of differentiation. We illustrate this in the following example.

Example 3. Find the derivative of

$$\boldsymbol{r}(t) = \langle \sin(t^2),\ \mathrm{e}^{\cos t},\ t\ln t \rangle.$$

Solution Applying the chain rule to the first two components and the product rule to the third, we have (for $t > 0$):

$$\begin{aligned} \boldsymbol{r}'(t) &= \langle \tfrac{\mathrm{d}}{\mathrm{d}t}[\sin(t^2)],\ \tfrac{\mathrm{d}}{\mathrm{d}t}(\mathrm{e}^{\cos t}),\ \tfrac{\mathrm{d}}{\mathrm{d}t}(t\ln t) \rangle \\ &= \langle 2t\cos(t^2),\ -\mathrm{e}^{\cos t}\sin t,\ 1+\ln t \rangle. \end{aligned}$$

For the most part, to compute derivatives of vector-valued functions, we only need to use the already familiar rules for differentiation of scalar functions. There are several special derivative rules, however, which we state in the following theorem.

Theorem 2. Suppose that $\boldsymbol{r}(t)$ and $\boldsymbol{s}(t)$ are differentiable vector-valued functions, $f(t)$ is a differentiable scalar function and c is any scalar constant. Then

1. $\frac{\mathrm{d}}{\mathrm{d}t}[\boldsymbol{r}(t)+\boldsymbol{s}(t)] = \boldsymbol{r}'(t)+\boldsymbol{s}'(t)$
2. $\frac{\mathrm{d}}{\mathrm{d}t}[c\boldsymbol{r}(t)] = c\boldsymbol{r}'(t)$
3. $\frac{\mathrm{d}}{\mathrm{d}t}[f(t)\boldsymbol{r}(t)] = f'(t)\boldsymbol{r}(t)+f(t)\boldsymbol{r}'(t)$
4. $\frac{\mathrm{d}}{\mathrm{d}t}[\boldsymbol{r}(t)\cdot\boldsymbol{s}(t)] = \boldsymbol{r}'(t)\cdot\boldsymbol{s}(t)+\boldsymbol{r}(t)\cdot\boldsymbol{s}'(t)$
5. $\frac{\mathrm{d}}{\mathrm{d}t}[\boldsymbol{r}(t)\times\boldsymbol{s}(t)] = \boldsymbol{r}'(t)\times\boldsymbol{s}(t)+\boldsymbol{r}(t)\times\boldsymbol{s}'(t)$

Notice that parts 3,4 and 5 are the product rules for the various kinds of products we can define. In 3, we have the derivative of a product of a scalar function and a vector-valued function; in 4 we have the derivative of a dot product and in 5, we have the derivative of a cross product. In each of these cases, it's important for you to recognize that these follow the same pattern as the product rule for the derivative of the product of two scalar functions. We leave the proof as exercises.

We conclude this section by making a few straightforward definitions. Recall that when we say that the scalar function $F(t)$ is an antiderivative of the scalar function $f(t)$, we mean that $F(t)$ is any function such that $F'(t) = f(t)$. We extend this notion to vector-valued functions in the following definition.

Definition 5. The vector-valued function $\boldsymbol{R}(t)$ is an antiderivative of the vector-valued function $\boldsymbol{r}(t)$ whenever $\boldsymbol{R}'(t) = \boldsymbol{r}(t)$.

Notice that if $\boldsymbol{r}(t)=\langle f(t), g(t), h(t)\rangle$ and f, g and h have antiderivatives F, G and H, respectively, then $\frac{\mathrm{d}}{\mathrm{d}t}\langle F(t), G(t), H(t)\rangle=\langle f(t), g(t), h(t)\rangle$. That is, $\langle F(t), G(t), H(t)\rangle$ is an antiderivative of $\boldsymbol{r}(t)$. In fact, $\langle F(t)+c_1, G(t)+c_2, H(t)+c_3\rangle$ is also an antiderivative of $\boldsymbol{r}(t)$, for any choice of constants c_1, c_2 and c_3. This leads us to the following definition.

Definition 6. If $\boldsymbol{R}(t)$ is an antiderivative of $\boldsymbol{r}(t)$, the indefinite integral of $\boldsymbol{r}(t)$ is defined to be $\int \boldsymbol{r}(t)\mathrm{d}t=\boldsymbol{R}(t)+\boldsymbol{c}$, where $\boldsymbol{c}$ is an arbitrary constant vector.

As in the scalar case, $\boldsymbol{R}(t)+\boldsymbol{c}$ is the most general antiderivative of $\boldsymbol{r}(t)$. Notice that this says that

$$\begin{aligned}\int \boldsymbol{r}(t)\mathrm{d}t &= \int\langle f(t), g(t), h(t)\rangle\mathrm{d}t\\ &= \left\langle\int f(t)\mathrm{d}t, \int g(t)\mathrm{d}t, \int h(t)\mathrm{d}t\right\rangle. \qquad (2.5)\end{aligned}$$

Example 4. Evaluate the indefinite integral

$$I=\int\langle t^2+2, \sin 2t, 4t\,\mathrm{e}^{t^2}\rangle\mathrm{d}t.$$

Solution From Equation(2.5), we have

$$\begin{aligned}I&=\left\langle\int(t^2+2)\mathrm{d}t, \int\sin 2t\mathrm{d}t, \int 4t\mathrm{e}^{t^2}\mathrm{d}t\right\rangle\\ &=\langle\tfrac{1}{3}t^3+2t+c_1, -\tfrac{1}{2}\cos 2t+c_2, 2\mathrm{e}^{t^2}+c_3\rangle\\ &=\langle\tfrac{1}{3}t^3+2t, -\tfrac{1}{2}\cos 2t, 2\mathrm{e}^{t^2}\rangle+\boldsymbol{c}\end{aligned}$$

where $\boldsymbol{c}=\langle c_1, c_2, c_3\rangle$ is an arbitrary constant vector.

Similarly, we define the definite integral of a vector-valued function in the obvious way.

Definition 7. For the vector-valued function

$$\boldsymbol{r}(t) = \langle f(t), g(t), h(t) \rangle,$$

we define the definite integral of $\boldsymbol{r}(t)$ by

$$\begin{aligned} \int_a^b \boldsymbol{r}(t)\mathrm{d}t &= \int_a^b \langle f(t), g(t), h(t) \rangle \mathrm{d}t \\ &= \langle \int_a^b f(t)\mathrm{d}t, \int_a^b g(t)\mathrm{d}t, \int_a^b h(t)\mathrm{d}t \rangle. \end{aligned}$$

Notice that this says that the definite integral of a vector-valued function $\boldsymbol{r}(t)$ is simply the vector whose components are the definite integrals of the corresponding components of $\boldsymbol{r}(t)$. With this in mind, we now extend the Fundamental Theorem of Calculus to vector-valued functions.

Theorem 3. Suppose that $\boldsymbol{R}(t)$ is an antiderivative of $\boldsymbol{r}(t)$ on the interval $[a,b]$. Then,

$$\int_a^b \boldsymbol{r}(t)\,\mathrm{d}t = \boldsymbol{R}(b) - \boldsymbol{R}(a).$$

The proof is straightforward and we leave this as an exercise.

Example 5. Evaluate $\int_0^1 \langle \sin \pi t, 6t^2+4t\rangle \mathrm{d}t$.

Solution Notice that an antiderivative for the integrand is

$$\langle -\frac{1}{\pi}\cos \pi t, \frac{6t^3}{3}+4\frac{t^2}{2}\rangle = \langle -\frac{1}{\pi}\cos \pi t, 2t^3+2t^2\rangle.$$

From Theorem 3, we have that

$$\int_0^1 \langle \sin \pi t, 6t^2+4t\rangle \mathrm{d}t$$
$$= \langle -\frac{1}{\pi}\cos \pi t, 2t^3+2t^2\rangle\Big|_0^1$$
$$= \langle -\frac{1}{\pi}\cos \pi, 2+2\rangle - \langle -\frac{1}{\pi}\cos 0, 0\rangle$$
$$= \langle \frac{2}{\pi}, 4\rangle.$$

作业

请将下列段落翻译成中文(在作业本上直接写翻译结果, 每两题之间空一行, 不用抄原文)

3. Page2Line2～Page3Line5

4. Page15

专业阅读

主讲 杨勤民

(Email: qmyang@ecust.edu.cn)

华东理工大学数学系

2020年3月～7月

The Residue Theorem

Let us consider the problem of evaluating the integral $\int_\Gamma f(z)\mathrm{d}z$, where Γ is a simple closed positively oriented contour and $f(z)$ is analytic on and inside Γ except for a single isolated singularity z_0, lying interior to Γ. As we know, the function $f(z)$ has a Laurent series expansion

$$f(z) = \sum_{j=-\infty}^{\infty} a_j(z-z_0)^j, \tag{3.1}$$

converging in some punctured circular neighborhood of z_0; in particular Equation (3.1) is valid for all z on the small positively oriented circle C indicated in Figure 3.1.

By the methods of Section 4.4, integration over Γ can be converted to integration over C without changing the integral $\int_\Gamma f(z)\mathrm{d}z = \int_C f(z)\mathrm{d}z$.

This last integral can be computed by termwise integration of the series (3.1) along C. For all $j \neq -1$ the integral is zero, and for $j = -1$ we obtain the value $2\pi i a_{-1}$. Consequently we have

$$\int_\Gamma f(z)\mathrm{d}z = 2\pi \mathrm{i} a_{-1}. \tag{3.2}$$

(Compare this with the formula for a_{-1} given in Theorem 14 of Chapter 5.)

Thus the constant a_{-1} plays an important role in contour integration. Accordingly, we adopt the following terminology.

Definition 1. If f has an isolated singularity at the point z_0, then the coefficient a_{-1} of $\frac{1}{z-z_0}$ in the Laurent expansion for f around z_0 is called the residue of f at z_0 and is denoted by $\operatorname{Res}(f; z_0)$ or $\operatorname{Res}(z_0)$.

Example 1. Find the residue at $z=0$ of the function $f(z) = z\mathrm{e}^{3/z}$ and compute $\oint_{|z|=4} z\mathrm{e}^{3/z}\mathrm{d}z$.

Solution Since e^{ω} has the Taylor expansion

$$\mathrm{e}^{\omega} = \sum_{j=0}^{\infty} \frac{\omega^j}{j!} \quad (\text{for all } \omega),$$

the Laurent expansion for $z\mathrm{e}^{3/z}$ around $z=0$ is given by

$$z\mathrm{e}^{3/z} = z\sum_{j=0}^{\infty} \frac{1}{j!}\left(\frac{3}{z}\right)^j = z + 3 + \frac{3^2}{2!z} + \frac{3^3}{3!z^2} + \cdots.$$

Hence

$$\operatorname{Res}(0) = \frac{3^2}{2!} = \frac{9}{2},$$

and since $z=0$ is the only singularity inside $|z|=4$, we have, by Formula 3.2,

$$\oint_{|z|=4} z\mathrm{e}^{3/z}\mathrm{d}z = 2\pi\mathrm{i}\cdot\frac{9}{2} = 9\pi\mathrm{i}.$$

Now if f has a removable singularity at z_0, all the coefficients of the negative powers of $(z-z_0)$ in its Laurent expansion are zero, and so, in particular, the residue at z_0 is zero. Furthermore, if f has a pole at z_0, we shall see that its residue there can be computed from a formula. Suppose first that z_0 is a simple pole, that is, a pole of order 1. Then for z near z_0, we have

$$f(z) = \frac{a_{-1}}{z-z_0} + a_1(z-z_0) + a_2(z-z_0)^2 + \cdots,$$

and so

$$\begin{aligned}&(z-z_0)f(z)\\ &= a_{-1} + (z-z_0)\left[a_0 + a_1(z-z_0) + a_2(z-z_0)^2 + \cdots\right].\end{aligned}$$

By taking the limit as $z \to z_0$ we deduce that

$$\lim_{z\to z_0}(z-z_0)f(z) = a_{-1} + 0.$$

Hence at a simple pole

$$\operatorname{Res}(f; z_0) = \lim_{z\to z_0}(z-z_0)f(z). \tag{3.3}$$

For example, the function $f(z) = \frac{e^z}{z(z+1)}$ has simple poles at $z=0$ and $z=-1$, therefore,

$$\operatorname{Res}(f; 0) = \lim_{z\to 0} zf(z) = \lim_{z\to 0}\frac{e^z}{z+1} = 1,$$

and

$$\operatorname{Res}(f; -1) = \lim_{z\to -1}(z+1)f(z) = \lim_{z\to -1}\frac{e^z}{z} = -e^{-1}.$$

Another consequence of Formula (3.3) is illustrated in the next example.

Example 2. Let $f(z)=\frac{P(z)}{Q(z)}$, where the function P and Q are both analytic at z_0, and Q has a simple zero at z_0, while $P(z_0)\neq 0$. Prove that $\mathrm{Res}(f;z_0)=\frac{P(z_0)}{Q'(z_0)}$.

Proof Obviously f has a simple pole at z_0 (see Section 5.6), so we can apply Formula (3.3). Using the fact that $Q(z_0)=0$, we see directly that

$$\mathrm{Res}(f;z_0)=\lim_{z\to z_0}(z-z_0)\frac{P(z)}{Q(z)}=\lim_{z\to z_0}\frac{P(z)}{\frac{Q(z)-Q(z_0)}{z-z_0}}=\frac{P(z_0)}{Q'(z_0)}.$$

Example 3. Compute the residue at each singularity of $f(z)=\cot z$.

Solution Since $\cot z=\cos z/\sin z$, the singularities of this function are simple poles occurring at the points $z=n\pi$, $n=0,\pm1,\pm2,\cdots$. Utilizing Example 2 with $P(z)=\cos z$, $Q(z)=\sin z$, the residues at these points are given by

$$\mathrm{Res}(\cot z;n\pi)=\left.\frac{\cos z}{(\sin z)'}\right|_{z=n\pi}=\frac{\cos n\pi}{\cos n\pi}=1.$$

To obtain the general formula for the residue at a pole of order m we need some method of picking out the coefficient a_{-1} from the Laurent expansion. The reader should encounter no difficulty in following the derivation of the next formula, which was obtained for rational functions in Section 3.1.

Theorem 1. If f has a pole of order m at z_0, then

$$\operatorname{Res}(f; z_0) = \lim_{z\to z_0} \frac{1}{(m-1)!}\frac{\mathrm{d}^{m-1}}{\mathrm{d}z^{m-1}}\left[(z-z_0)^m f(z)\right]. \quad (3.4)$$

Proof

Starting with the Laurent expansion for f around z_0,

$$f(z) = \frac{a_{-m}}{(z-z_0)^m} + \cdots + \frac{a_{-2}}{(z-z_0)^2} + \frac{a_{-1}}{z-z_0} + a_0 + a_1(z-z_0) + \cdots,$$

we multiply by $(z-z_0)^m$,

$$(z-z_0)^m f(z) = a_{-m} + \cdots + a_{-2}(z-z_0)^{m-2}$$
$$+ a_{-1}(z-z_0)^{m-1} + a_0(z-z_0)^m + a_1(z-z_0)^{m+1} + \cdots.$$

and differentiate $m-1$ times to derive

$$\frac{\mathrm{d}^{m-1}}{\mathrm{d}z^{m-1}}[(z-z_0)^m f(z)] = (m-1)!a_{-1}$$
$$+ m!a_0(z-z_0) + \frac{(m+1)!}{2}a_1(z-z_0)^2 + \cdots.$$

Hence

$$\lim_{z\to z_0} \frac{\mathrm{d}^{m-1}}{\mathrm{d}z^{m-1}}[(z-z_0)^m f(z)] = (m-1)!a_{-1} + 0,$$

which is equivalent to Equation (3.4).

Example 4. Compute the residues at the singularities of $f(z)=\frac{\cos z}{z^2(z-\pi)^3}$.

Solution This function has a pole of order 2 at $z=0$ and a pole of order 3 at $z=\pi$. Applying Formula (3.4) we find

$$\operatorname{Res}(0)=\lim_{z\to 0}\frac{1}{1!}\frac{\mathrm{d}}{\mathrm{d}z}\left[z^2 f(z)\right]=\lim_{z\to 0}\frac{\mathrm{d}}{\mathrm{d}z}\left[\frac{\cos z}{(z-\pi)^3}\right]$$

$$=\lim_{z\to 0}\left[\frac{-(z-\pi)\sin z-3\cos z}{(z-\pi)^4}\right]=-\frac{3}{\pi^4},$$

$$\operatorname{Res}(\pi)=\lim_{z\to\pi}\frac{1}{2!}\frac{\mathrm{d}^2}{\mathrm{d}z^2}\left[(z-\pi)^3 f(z)\right]$$

$$=\lim_{z\to\pi}\frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}z^2}\left[\frac{\cos z}{z^2}\right]=\lim_{z\to\pi}\frac{1}{2}\left[\frac{(6-z^2)\cos z+4z\sin z}{z^4}\right]=\frac{\pi^2-6}{2\pi^4}.$$

We have already seen how to compute the integral $\int_\Gamma f(z)\mathrm{d}z$ when $f(z)$ has only one singularity inside Γ. Let's now turn to the more general case where Γ is a simple closed positively oriented contour and $f(z)$ is analytic inside and on Γ except for a finite number of isolated singularities at the points $z_1, z_2, \ldots, z_n$ inside Γ (see Figure 3.2). Notice that by the methods of Section 4.4 we can express the integral along Γ in terms of the integrals around the circles C_j in Figure 3.3:

$$\int_\Gamma f(z)\mathrm{d}z=\sum_{j=1}^{n}\int_{C_j} f(z)\mathrm{d}z.$$

However, because z_j is the only singularity of f inside C_j, we know that

$$\int_{C_j} f(z)\mathrm{d}z = 2\pi\mathrm{i}\,\mathrm{Res}(z_j).$$

Hence we have established the following important result.

Theorem 2. (Cauchy's Residue Theorem.) If Γ is a simple closed positively oriented contour and f is analytic inside and on Γ except at the points $z_1, z_2, \cdots, z_n$ inside Γ, then

$$\int_{\Gamma} f(z)\mathrm{d}z = 2\pi\mathrm{i}\sum_{j=1}^{n} \mathrm{Res}(z_j).$$

Example 5. Evaluate $\oint_{|z|=2} \frac{1-2z}{z(z-1)(z-3)}\mathrm{d}z$.

Solution The integrand $f(z) = (1-2z)/[z(z-1)(z-3)]$ has simple poles at $z=0$, $z=1$, and $z=3$. However, only the first two of these points lie inside Γ: $|z|=2$. Thus by the residue theorem

$$\oint_{|z|=2} f(z)\mathrm{d}z = 2\pi\mathrm{i}[\mathrm{Res}(0)+\mathrm{Res}(1)],$$

and since

$$\mathrm{Res}(0) = \lim_{z\to 0} zf(z) = \lim_{z\to 0}\frac{1-2z}{(z-1)(z-3)} = \frac{1}{3},$$
$$\mathrm{Res}(1) = \lim_{z\to 1}(z-1)f(z) = \lim_{z\to 1}\frac{1-2z}{z(z-3)} = \frac{1}{2},$$

we obtain

$$\oint_{|z|=2} f(z)\mathrm{d}z = 2\pi\mathrm{i}\left(\frac{1}{3}+\frac{1}{2}\right) = \frac{5\pi\mathrm{i}}{3}.$$

作业

请将下列段落翻译成中文(在作业本上直接写翻译结果, 每两题之间空一行, 不用抄原文)

5. Page2Line2～Page3Line3

6. Page9Line1～Page10Line6

专业阅读

主讲 杨勤民

(Email: qmyang@ecust.edu.cn)

华东理工大学数学系

2020年3月～7月

Basic Matrix Operations

Matrices play a central role in this book. They form an important part of the theory, and many concrete examples are based on them. Therefore it is essential to develop facility in matrix manipulation. Since matrices pervade much of mathematics, the techniques needed here are sure to be useful elsewhere.

Let m, n be positive integers. An $m \times n$ matrix is a collection of mn numbers arranged in a rectangular array:

$$\begin{array}{cc} & n \text{ columns} \\ m \text{ rows} & \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \end{array}$$

For example,

$$\begin{bmatrix} 2 & 1 & 0 \\ 1 & 3 & 5 \end{bmatrix}$$

is a 2×3 matrix.

The numbers in a matrix are called the matrix entries and are denoted by a_{ij}, where i,j are indices (integers) with $1 \leq i \leq m$ and $1 \leq j \leq n$. The index i is called the row index, and j is the column index. So a_{ij} is the entry which appears in the ith row and jth column of the matrix:

$$\begin{array}{cc} & \begin{array}{cccc} & j & & \end{array} \\ i & \begin{bmatrix} & \vdots & & \\ \cdots & a_{ij} & \cdots & \cdots \\ & \vdots & & \end{bmatrix} \end{array}$$

In the example above, $a_{11} = 2$, $a_{13} = 0$, and $a_{23} = 5$.

We usually introduce a symbol such as A to denote a matrix, or we may write it as (a_{ij}).

A $1 \times n$ matrix is called an n-dimensional row vector. We will drop the index i when $m = 1$ and write a row vector as

$$A = [a_1 \cdots a_n], \quad \text{or as} \quad A = (a_1, \ldots, a_n). \tag{4.1}$$

The commas in this row vector are optional. Similarly, an $m \times 1$ matrix is an m-dimensional column vector:

$$B = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}. \tag{4.2}$$

A 1×1 matrix $[a]$ contains a single number, and we do not distinguish such a matrix from its entry.

Addition of matrices is vector addition:

$$(a_{ij}) + (b_{ij}) = (s_{ij}),$$

where $s_{ij} = a_{ij} + b_{ij}$ for all i, j. Thus

$$\begin{bmatrix} 2 & 1 & 0 \\ 1 & 3 & 5 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 3 \\ 4 & -3 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 1 & 3 \\ 5 & 0 & 6 \end{bmatrix}.$$

The sum of two matrices A, B is defined only when they are both of the same shape, that is, when they are $m \times n$ matrices with the same m and n.

Scalar multiplication of a matrix by a number is defined as with vectors. The result of multiplying a number c and a matrix (a_{ij}) is another matrix:

$$c(a_{ij}) = (b_{ij}),$$

where $b_{ij} = ca_{ij}$ for all i,j. Thus

$$2\begin{bmatrix} 0 & 1 \\ 2 & 3 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ 4 & 6 \\ 4 & 2 \end{bmatrix}.$$

Numbers will also be referred to as scalars.

The complicated notion is that of matrix multiplication. The first case to learn is the product AB of a row vector A (4.1) and a column vector B (4.2) which is defined when both are the same size, that is, $m = n$. Then the product AB is the 1×1 matrix or scalar

$$a_1b_1 + a_2b_2 + \cdots + a_mb_m. \tag{4.3}$$

(This product is often called the "dot product" of the two vectors.) Thus

$$[3 \;\; 1 \;\; 2]\begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix} = 3 \cdot 1 + 1 \cdot (-1) + 2 \cdot 4 = 10.$$

The usefulness of this definition becomes apparent when we regard A and B as vectors which represent indexed quantities. For example, consider a candy bar containing m ingredients. Let a_i denote the number of grams of $(\text{ingredient})_i$ per candy bar, and let b_i denote the cost of $(\text{ingredient})_i$ per gram. Then the matrix product $AB = c$ computes the cost per candy bar:

$$(\text{grams}/\text{bar}) \cdot (\text{cost}/\text{gram}) = (\text{cost}/\text{bar}).$$

作业

请将下列段落翻译成中文(在作业本上直接写翻译结果, 每两题之间空一行, 不用抄原文)

作业第7题 Page9

In general, the product of two matrices A and B is defined if the number of columns of A is equal to the number of rows of B, say if A is an $l \times m$ matrix and B is an $m \times n$ matrix. In this case, the product is an $l \times n$ matrix. Symbolically, $(l \times m) \cdot (m \times n) = (l \times n)$. The entries of the product matrix are computed by multiplying all rows of A by all columns of B, using rule Equation (4.3) above. Thus if we denote the product AB by P, then

$$p_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{im}b_{mj}. \tag{4.4}$$

This is the product of the ith row of A and the jth column of B.

$$\begin{bmatrix} \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i_2} & \cdots & a_{im} \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} \cdots & b_{1j} & \cdots \\ \cdots & b_{2j} & \cdots \\ \vdots & \vdots & \vdots \\ \cdots & b_{mj} & \cdots \end{bmatrix} = \begin{bmatrix} \ddots & \vdots & \ddots \\ \cdots & p_{ij} & \cdots \\ \ddots & \vdots & \ddots \end{bmatrix}$$

For example,

$$\begin{bmatrix} 0 & -1 & -2 \\ 3 & 4 & -6 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}. \tag{4.5}$$

This definition of matrix multiplication has turned out to provide a very convenient computational tool.

Going back to our candy bar example, suppose that there are l candy bars. Then we may form a matrix A whose ith row measures the ingredients of $(\text{bar})_i$. If the cost is to be computed each year for n years, we may form a matrix B whose jth column measures the cost of the ingredients in $(\text{year})_j$. The matrix product $AB = P$ computes the cost per bar:

$$p_{ij} = \text{cost of}(\text{bar})_i \text{ in } (\text{year})_j.$$

Matrix notation was introduced in the nineteenth century to provide a shorthand way of writing linear equations. The system of equations

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

can be written in matrix notation as

$$AX = B,$$

where A denotes the coefficient matrix (a_{ij}), X and B are column vectors, and AX is the matrix product

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

Thus the matrix equation

$$\begin{bmatrix} 0 & -1 & 2 \\ 3 & 4 & -6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

represents the following system of two equations in three unknowns:

$$\begin{aligned} -x_2 + 2x_3 &= 2, \\ 3x_1 + 4x_2 - 6x_3 &= 1. \end{aligned}$$

Equation (4.5) exhibits one solution: $x_1 = 1$, $x_2 = 4$, $x_3 = 3$.

Formula (4.4) defining the product can also be written in "sigma" notation as

$$p_{ij} = \sum_{k=1}^{m} a_{ik} b_{kj} = \sum_{k} a_{ik} b_{kj}.$$

Each of these expressions is a shorthand notation for the sum (4.4) which defines the product matrix.

Our two most important notations for handling sets of numbers are the $\sum$ or sum notation as used above and matrix notation. The $\sum$ notation is actually the more versatile of the two, but because matrices are much more compact we will use them whenever possible. One of our tasks in later chapters will be to translate complicated mathematical structures into matrix notation in order to be able to work with them conveniently.

Various identities are satisfied by the matrix operation, such as the distributive laws

$A(B+B') = AB+AB' \quad \text{and} \quad (A+A')B = AB+A'B$

and the associative law

$$(AB)C = A(BC). \tag{4.6}$$

These laws hold whenever the matrices involved have suitable sizes, so that the products are defined. For the associative law, for example, the sizes should be $A = l \times m$, $B = m \times n$ and, $C = n \times p$, for some l, m, n, p. Since the two products (4.6) are equal, the parentheses are not required, and we will denote them by ABC . The triple product ABC is then an $l \times p$ matrix.

For example, the two ways of computing the product

$$ABC = \begin{bmatrix} 1 \\ 2 \end{bmatrix} [1 \;\; 0 \;\; 1] \begin{bmatrix} 2 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$$

are

$$(AB)C = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 0 & 2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 6 & 2 \end{bmatrix}$$

and

$$A(BC) = \begin{bmatrix} 1 \\ 2 \end{bmatrix} [2 \;\; 1] = \begin{bmatrix} 3 & 1 \\ 6 & 2 \end{bmatrix}.$$

作业

请将下列段落翻译成中文(在作业本上直接写翻译结果, 每两题之间空一行, 不用抄原文)

作业第8题 Page18

Scalar multiplication is compatible with matrix multiplication in the obvious sense:

$$c(AB) = (cA)B = A(cB).$$

The proofs of these identities are straightforward and not very interesting.

In contrast, the commutative law does not hold for matrix multiplication; that is, $AB \neq BA$, usually.

In fact, if A is an $l \times m$ matrix and B is an $m \times l$ matrix, so that AB and BA are both defined, then AB is $l \times l$ while BA is $m \times m$. Even if both matrices are square, say $m \times m$, the two products tend to be different. For instance,

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},$$

while

$$\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Since matrix multiplication is not commutative, care must be taken when working with matrix equations. We can multiply both sides of an equation $B = C$ on the left by a matrix A, to conclude that $AB = AC$, provided that the products are defined. Similarly, if the products are defined, then we can conclude that $BA = CA$. We can not derive $AB = CA$ from $B = C$!

Any matrix all of whose entries are 0 is called a zero matrix and is denoted by 0, though its size is arbitrary. Maybe $0_{m\times n}$ would be better.

The entries a_{ii} of a matrix A are called its diagonal entries, and a matrix A is called a diagonal matrix if its only nonzero entries are diagonal entries.

The square $n \times n$ matrix whose only nonzero entries are 1 in each diagonal position,

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

is called the $n \times n$ identity matrix. It behaves like 1 in multiplication: If A is an $m \times n$ matrix, then

$$I_m A = A \quad \text{and} \quad A I_n = A.$$

Here are some shorthand ways of drawing the matrix I_n:

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

We often indicate that a whole region in a matrix consists of zeros by leaving it blank or by putting in a single 0.

We will use "$*$" to indicate an arbitrary undetermined entry of a matrix. Thus

$$I_n = \begin{bmatrix} * & & * \\ & \ddots & \\ 0 & & * \end{bmatrix}.$$

may denote a square matrix whose entries below the diagonal are 0, the other entries being undetermined. Such a matrix is called an upper triangular matrix.

Let A be a (square) $n \times n$ matrix. If there is a matrix B such that

$$AB = I_n \quad \text{and} \quad BA = I_n, \tag{4.7}$$

then B is called an inverse of A and is denoted by A^{-1}:

$$A^{-1}A = I_n = AA^{-1}.$$

When A has an inverse, it is said to be an invertible matrix. For example, the matrix $A = \begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix}$ is invertible. Its inverse is $A^{-1} = \begin{bmatrix} 3 & -1 \\ -5 & 2 \end{bmatrix}$, as is seen by computing the products AA^{-1} and $A^{-1}A$.

Two more examples are:

$$\begin{bmatrix} 1 & \\ & 2 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & \\ & \frac{1}{2} \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 1 \\ & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -1 \\ & 1 \end{bmatrix}.$$

We will see later that A is invertible if there is a matrix B such that either one of the two relations $AB = I_n$ or $BA = I_n$ holds, and that B is then the inverse. But since multiplication of matrices is not commutative, this fact is not obvious. It fails for matrices which aren't square. For example, let $A = [1 \;\; 2]$ and let $B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Then $AB = [1] = I_1$, but $BA = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} \neq I_1$.

On the other hand, an inverse is unique if it exists at all. In other words, there can be only one inverse. Let B, B' be two matrices satisfying Equations (4.7), for the same matrix A. We need only know that $AB = I_n$ (B is a right inverse) and that $B'A = I_n$ (B' is a left inverse). By the associative law, $B'(AB) = (B'A)B$. Thus

$$B' = B'I = B'(AB) = (B'A)B = IB = B,$$

and so $B' = B$.

作业

请将下列段落翻译成中文(在作业本上直接写翻译结果, 每两题之间空一行, 不用抄原文)

作业第9题 Page23

Proposition 1. Let A, B be $n \times n$ matrices. If both are invertible, so is their product AB, and

$(AB)^{-1} = B^{-1}A^{-1}$.

More generally, if $A_1, \ldots, A_m$ are invertible, then so is the product $A_1 \cdots A_m$, and its inverse is $A_m^{-1} \cdots A_1^{-1}$.

Thus the inverse of

$$\begin{bmatrix} 1 & \\ & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ & 2 \end{bmatrix}$$

is

$$\begin{bmatrix} 1 & -1 \\ & 1 \end{bmatrix} \begin{bmatrix} 1 & \\ & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 1 & -\frac{1}{2} \\ & \frac{1}{2} \end{bmatrix}.$$

Proof Assume that A, B are invertible. Then we check that $B^{-1}A^{-1}$ is the inverse of AB:

$$ABB^{-1}A^{-1} = AIA^{-1} = AA^{-1} = I,$$

and similarly

$$B^{-1}A^{-1}AB = \cdots = I.$$

The last assertion is proved by induction on m. When $m = 1$, the assertion is that if A_1 is invertible then A_1^{-1} is the inverse of A_1, which is trivial.

Next we assume that the assertion is true for $m = k$, and we proceed to check it for $m = k+1$. We suppose that $A_1, \ldots, A_{k+1}$ are invertible $n \times n$ matrices, and we denote by P the product $A_1 \cdots A_k$ of the first k matrices. By the induction hypothesis, P is invertible, and its inverse is $A_k^{-1} \cdots A_1^{-1}$. Also, A_{k+1} is invertible. So, by what has been shown for two invertible matrices, the product $PA_{k+1} = A_1 \cdots A_k A_{k+1}$ is invertible, and its inverse is $A_{k+1}^{-1} P^{-1} = A_{k+1}^{-1} A_k^{-1} \cdots A_1^{-1}$. This shows that the assertion is true for $m = k+1$, which completes the induction proof.

Though this isn't clear from the definition of matrix multiplication, we will see that most square matrices are invertible. But finding the inverse explicitly is not a simple problem when the matrix is large.

The set of all invertible $n \times n$ matrices is called the n-dimensional general linear group and is denoted by GL_n. The general linear groups will be among our most important examples when we study the basic concept of a group in the next chapter.

Various tricks simplify matrix multiplication in favorable cases. Block multiplication is one of them. Let M, M' be $m \times n$ and $n \times p$ matrices, and let r be an integer less than n. We may decompose the two matrices into blocks as follows:

$$M = [A \mid B] \quad \text{and} \quad M' = \left[\frac{A'}{B'}\right],$$

where A has r columns and A' has r rows. Then the matrix product can be computed as follows:

$$MM' = AA' + BB'. \tag{4.8}$$

This decomposition of the product follows directly from the definition of multiplication, and it may facilitate computation. For example,

$$\left[\begin{array}{cc|c} 1 & 0 & 5 \\ 0 & 1 & 7 \end{array}\right] \left[\begin{array}{cc} 2 & 3 \\ 4 & 8 \\ \hline 0 & 0 \end{array}\right]$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 4 & 8 \end{bmatrix} + \begin{bmatrix} 5 \\ 7 \end{bmatrix} [0 \quad 0] = \begin{bmatrix} 2 & 3 \\ 4 & 8 \end{bmatrix}.$$

Note that Formula (4.8) looks the same as Formula (4.3) for multiplying a row vector and a column vector.

We may also multiply matrices divided into more blocks. For our purposes, a decomposition into four blocks will be the most useful. In this case the rule for block multiplication is the same as for multiplication of 2×2 matrices. Let $r+s=n$ and let $k+l=m$. Suppose we decompose an $m \times n$ matrix M and an $n \times p$ matrix M' into submatrices

$$M = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right], \qquad M' = \left[\begin{array}{c|c} A' & B' \\ \hline C' & D' \end{array}\right]$$

where the number of columns of A is equal to the number of rows of A'. Then the rule for block multiplication is

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]\left[\begin{array}{c|c} A' & B' \\ \hline C' & D' \end{array}\right]=\left[\begin{array}{c|c} AA'+BC' & AB'+BD' \\ \hline CA'+DC' & CB'+DD' \end{array}\right].$$

For example,

$$\left[\begin{array}{cc|c} 1 & 0 & 5 \\ \hline 0 & 1 & 7 \end{array}\right]\left[\begin{array}{cc|cc} 2 & 3 & 1 & 1 \\ 4 & 1 & 0 & 0 \\ \hline 0 & 1 & 1 & 0 \end{array}\right]=\left[\begin{array}{cc|cc} 2 & 8 & 6 & 1 \\ \hline 4 & 8 & 7 & 0 \end{array}\right].$$

In this product, the upper left block is

$$[1 \ \ 0]\begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix}+[5][0 \ \ 1]=[2 \ \ 8],$$

etc.

Again, this rule can be verified directly from the definition of matrix multiplication. In general, block multiplication can be used whenever two matrices are decomposed into submatrices in such a way that the necessary products are defined.

Besides facilitating computations, block multiplication is a useful tool for proving facts about matrices by induction.

作业

请将下列段落翻译成中文(在作业本上直接写翻译结果, 每两题之间空一行, 不用抄原文)

作业第10题　Page33Line7～Page34Line11

专业阅读

主讲 杨勤民

(Email: qmyang@ecust.edu.cn)

华东理工大学数学系

2020年3月～7月

Integrals as General and Particular Solutions

The first-order equation $\mathrm{d}y/\mathrm{d}x = f(x,y)$ takes an especially simple form if the function f is independent of the dependent variable y:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f(x). \tag{5.1}$$

In this special case we need only integrate both sides of Equation (5.1) to obtain

$$y(x) = \int f(x)\mathrm{d}x + C. \tag{5.2}$$

This is a general solution of Equation (5.1), meaning that it involves an arbitrary constant C, and for every choice of C it is a solution of the differential equation in Equation (5.1).

If $G(x)$ is a particular antiderivative of f — that is, if $G'(x) \equiv f(x)$ — then

$$y(x) = G(x) + C. \tag{5.3}$$

The graphs of any two such solutions $y_1(x) = G(x) + C_1$ and $y_2(x) = G(x) + C_2$ on the same interval I are "parallel" in the sense illustrated by Figures 5.1 and 5.2. There we see that the constant C is geometrically the vertical distance between the two curves $y(x) = G(x)$ and $y(x) = G(x) + C$.

To satisfy an initial condition $y(x_0) = y_0$, we need only substitute $x = x_0$ and $y = y_0$ into Equation (5.3) to obtain $y_0 = G(x_0) + C$, so that $C = y_0 - G(x_0)$.

With this choice of C, we obtain the particular solution of Equation (5.1) satisfying the initial value problem

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f(x), \qquad y(x_0) = y_0.$$

We will see that this is the typical pattern for solution of first-order differential equations. Ordinarily, we will first find a general solution involving an arbitrary constant C. We can then attempt to obtain, by appropriate choice of C, a particular solution satisfying a given initial condition $y(x_0) = y_0$.

Remark As the term used in the previous paragraph, a general solution of a first-order differential equation is simply a one-parameter family of solutions. A natural question is whether a given general solution contains every particular solution of the differential equation. When this is known to be true, we call it **the** general solution of the differential equation. For example, because any two antiderivatives of the same function $f(x)$ can differ only by a constant, it follows that every solution of Equation (5.1) is of the form in Equation (5.2). Thus Equation (5.2) serves to define **the** general solution of Equation (5.1).

作业

请将下列段落翻译成中文(在作业本上直接写翻译结果, 每两题之间空一行, 不用抄原文)

作业第11题　Page6