TMV035 Analysis and Linear Algebra B, 2005

LECTURE 1.1

In this lecture we present the integral. This covers AMBS Ch 27.

1. Introduction

Remember that the goal of ALA-A was to solve systems of algebraic equations:

$$f(x) = 0.$$

We introduced the following functions: the polynomials, the rational functions, and their inverse functions. The polynomial and rational functions can be computed exactly with rational numbers, but the inverse functions, for example, the square root $\sqrt{x}$, or more generally the power functions $x^{p/q}$ with rational exponent, must be computed approximately by some iterative algorithm.

In this course the goal is to solve systems of ordinary differential equations (ODEs):

$$u'(x) = f(x, u(x)).$$

In this way we will introduce the functions $\ln(x)$, $\exp(x)$, $\sin(x)$, $\cos(x)$, and related functions. These are also based on algorithms for approximate computation. We will write MATLAB programs for computing them.

We will also study linear algebra in n dimensions: matrix algebra, linear systems of equations, Gaussian elimination, orhtogonal matrices, least squares method.

We will do two applications projects together with the chemistry course: project 9, reaction kinetics, begins in week 4, examined by a written report; project 12, equilibrium equations, is done and examined in the studio class room in weeks 6–7.

2. The integral

We may consider two tasks in calculus:
1. Given $u(x)$ determine $u'(x)$. This was done in ALA-A.
2. Given $u'(x)$ determine $u(x)$. This is our first topic in ALA-B.

Example 1. We know from ALA-A that

$$Dx^m = mx^{m-1}, \quad m \geq 0,\ m \in \mathbf{Q}.$$

Hence

$$u'(x) = D\frac{x^m}{m} = x^{m-1}, \quad m > 0,$$
$$u'(x) = D\frac{x^{r+1}}{r+1} = x^r, \quad r = m-1 > -1.$$

We may now guess that

$$u(x) = \frac{x^{r+1}}{r+1} + c, \quad r > -1,$$

where c is an arbitrary constant, because $D(\frac{x^{r+1}}{r+1} + c) = x^r$. We determine the constant:

$$u(0) = 0 \implies c = 0 \implies u(x) = \frac{x^{r+1}}{r+1},$$
$$u(0) = u_0 \implies c = u_0 \implies u(x) = \frac{x^{r+1}}{r+1} + u_0.$$

Date: October 23, 2005, Stig Larsson, Mathematical Sciences, Chalmers University of Technology.

Example 2. We also know from ALA-A that

$$Dx^{-m} = -mx^{-m-1}, \quad m \geq 0,\ m \in \mathbf{Q}.$$

Hence

$$u'(x) = D\frac{x^{-m}}{-m} = x^{-m-1}, \quad m > 0,$$
$$u'(x) = D\frac{x^{-r+1}}{-r+1} = x^{-r}, \quad r = m+1 > 1,$$

We conclude that

$$u(x) = \frac{x^{-r+1}}{-r+1} + c, \quad r > 1,$$

where c is an arbitrary constant. We determine the constant:

$$u(1) = 0 \implies c = \frac{-1}{-r+1} \implies u(x) = \frac{x^{-r+1}-1}{-r+1}$$
$$u(1) = u_1 \implies c = \frac{-1}{-r+1} + u_1 \implies u(x) = \frac{x^{-r+1}-1}{-r+1} + u_1.$$

We now make the task more precise: Given an interval $[a, b]$, a function $f : [a, b] \to \mathbf{R}$, and a number u_a, we want to find a unique function u such that

(1) $$\begin{cases} u'(x) = f(x), & x \in [a, b], \\ u(a) = u_a. \end{cases}$$

If we can find such a function u then we write

(2) $$u(x) = u_a + \int_a^x f(y)\,\mathrm{d}y.$$

The term $\int_a^x f(y)\,\mathrm{d}y$ is called the integral of f over the interval $[a, x]$. The problem (1) is called an *initial-value problem* ("begynnelsevärdesproblem"). The first equation in (1) is a differential equation. The second equation is an initial condition and the number u_a is an initial value.

Example 1 means that

$$u(x) = u_0 + \int_0^x y^r\,\mathrm{d}y = u_0 + \frac{x^{r+1}}{r+1} \quad (r > 1)$$

is the solution of

$$\begin{cases} u'(x) = x^r, & x \in [0, b], \\ u(0) = u_0, \end{cases}$$

for any $b > 0$.

In a similar way we can integrate a constant function $f(x) = K$:

(3) $$u(x) = u_a + \int_a^x K\,\mathrm{d}y = u_a + K(x-a), \quad x \in [a, b].$$

This function satisfies

$$\begin{cases} u'(x) = K, & x \in [a, b], \\ u(a) = u_a, \end{cases}$$

for any $b > a$.

Example 2 means that

$$u(x) = u_1 + \int_1^x y^{-r}\,\mathrm{d}y = u_1 + \frac{x^{-r+1}-1}{-r+1} \quad (r > 1)$$

is the solution of

$$\begin{cases} u'(x) = x^{-r}, & x \in [1, b], \\ u(1) = u_1, \end{cases}$$

for any $b > 1$.

In this way we can integrate x^r for all rational numbers $r \neq -1$. What about x^{-1}? That is, $\int_1^x y^{-r}\,\mathrm{d}y$? This ought to be $\ln(x)$, but what is this? We will soon answer this question.

And what about an arbitrary function $f(x)$? That is, $\int_1^x f(y)\,\mathrm{d}y$, where f could be any function?

The answer is that we must *construct* the function $u(x)$ if we cannot guess it as we did in the examples. We recall the constructive proof in four steps:

(1) An algoritm which produces a sequence.
(2) The sequence is a Cauchy sequence.
(3) The limit of the sequence solves the problem.
(4) The solution is unique.

Step 1. Algorithm. We begin by creating a mesh by bisecting the interval $[a, b]$ repeatedly. When the number of bisections is $n = 1$ we have

$$x_0^1 = a,\ x_1^1 = \frac{a+b}{2},\ x_2^1 = b.$$

The points x_i^1 are called meshpoints (or nodes) and the distance between them is the steplength $h_1 = (b-a)/2 = (b-a)2^{-1}$. We continue with $n = 2$ and the steplength $h_2 = (b-a)2^{-2}$ and the meshpoints:

$$x_0^2 = a,\ x_1^2 = a + h_2,\ x_2^2 = a + 2h_2,\ x_3^2 = a + 3h_2,\ x_4^2 = b.$$

Note that here the 2 in x_i^2 is not an exponent but an index which indicates that we have bisected $n = 2$ times. With $n = 3$ the steplength is $h_3 = (b-a)2^{-3}$ and

$$\begin{aligned} x_0^3 = a,\ x_1^3 = a + h_3,\ x_2^3 = a + 2h_3,\ x_3^3 = a + 3h_3,\ x_4^3 = a + 4h_3,& \\ x_5^3 = a + 5h_3,\ x_6^3 = a + 6h_3,\ x_7^3 = a + 7h_3,\ x_8^2 = b.& \end{aligned}$$

After n bisections we have $h_n = (b-a)2^{-n}$ and

$$x_0^n = a,\ x_1^n = a + h_n,\ x_2^n = a + 2h_n,\ \ldots,\ x_i^n = a + ih_n,\ \ldots,\ x_N^n = b.$$

The meshpoints are $\{x_i^n\}_{i=0}^N$. Together they form a mesh on the interval $[a, b]$. The number of meshpoints is $N + 1$ and $N = 2^n$.

Then for each level of bisection n we create an approximate solution $U^n(x)$ as follows. We first set

$$U^n(x_0^n) = U^n(a) = u_a.$$

Then we approximate the function $f(x)$ on the interval $[x_0^n, x_1^n]$ by its value $f(x_0^n)$ in the left endpoint and integrate as in (3):

$$\begin{aligned} U^n(x) &= U^n(x_0^n) + \int_{x_0^n}^x f(x_0^n)\,\mathrm{d}y \\ &= U^n(x_0^n) + f(x_0^n)(x - x_0^n), \quad x \in [x_0^n, x_1^n]. \end{aligned}$$

At the next meshpoint we then have

$$\begin{aligned} U^n(x_1^n) &= U^n(x_0^n) + f(x_0^n)(x_1^n - x_0^n) \\ &= U^n(x_0^n) + f(x_0^n)h_n. \end{aligned}$$

We continue in the same way:

$$\begin{aligned} U^n(x) &= U^n(x_1^n) + \int_{x_1^n}^x f(x_1^n)\,\mathrm{d}y \\ &= U^n(x_1^n) + f(x_1^n)(x - x_1^n), \quad x \in [x_1^n, x_2^n], \end{aligned}$$

and

$$U^n(x_2^n) = U^n(x_1^n) + f(x_1^n)h_n.$$

In this way we construct $U^n(x)$ in each mesh interval $I_i^n = [x_{i-1}^n, x_i^n]$. It is sufficient to compute the node values $U^n(x_{i-1}^n)$ and $U^n(x_i^n)$; then the intermediate values are obtained by

$$\begin{aligned} U^n(x) &= U^n(x_{i-1}^n) + f(x_{i-1}^n)(x - x_{i-1}^n) \\ &= U^n(x_{i-1}^n) + \frac{U^n(x_i^n) - U^n(x_{i-1}^n)}{h_n}(x - x_{i-1}^n), \quad x \in [x_{i-1}^n, x_i^n]. \end{aligned}$$

In other words: the graph of U^n is obtained by drawing a straight line between the node values.

We can now formulate the algorithm in a compact form:

Algorithm. (Rectangle rule) First set the initial values:

$$\begin{cases} x_0^n = a, \\ U(x_0^n) = u_a. \end{cases} \tag{4}$$

Then set $h_n = 2^{-n}(b-a)$, $N = 2^n$, and compute for $n = 1, \dots, N$:

$$\begin{cases} x_i^n = x_{i-1}^n + h_n, \\ U^n(x_i^n) = U^n(x_{i-1}^n) + f(x_{i-1}^n)h_n. \end{cases} \tag{5}$$

This is very easy to program in MATLAB. You will soon do this in the studio class.

Another motivation for the algorithm is that if a solution u exists, then it is differentiable with $u' = f$ so that

$$\begin{aligned} u(x_i^n) &= u(x_{i-1}^n) + u'(x_{i-1}^n)(x_i^n - x_{i-1}^n) + E_u(x_i^n, x_{i-1}^n) \\ &= u(x_{i-1}^n) + f(x_{i-1}^n)h_n + E_u(x_i^n, x_{i-1}^n). \end{aligned}$$

If we skip the remainder then we obtain (5).

It is important to note that $U^n(x_i^n)$ can also be expressed as a sum:

$$\begin{aligned} U^n(x_i^n) &= U^n(x_{i-1}^n) + f(x_{i-1}^n)h_n \\ &= U^n(x_{i-2}^n) + f(x_{i-2}^n)h_n + f(x_{i-1}^n)h_n \\ &\vdots \\ &= U^n(x_0^n) + f(x_0^n)h_n + \dots + f(x_{i-1}^n)h_n \\ &= u_a + \sum_{j=1}^{i} f(x_{j-1}^n)h_n, \end{aligned} \tag{6}$$

that is,

$$U^n(x_i^n) = u_a + \sum_{j=1}^{i} f(x_{j-1}^n)h_n. \tag{7}$$

Step 2. Cauchy sequence. We must show that $\{U^n(x)\}_{n=1}^\infty$ is a Cauchy sequence (for each fixed x), i.e.,

$$|U^n(x) - U^m(x)| \to 0, \quad \text{as } m, n \to \infty.$$

The proof of this is rather long and complicated. It is written in detail in the book. In the proof it is important that f is Lipschitz continuous on $[a, b]$. You may skip this part of the proof if you like.

That $\{U^n(x)\}_{n=1}^\infty$ is a Cauchy sequence means that it generates a decimal expansion, and a decimal expansion is the same as a real number. Therefore we get a real number $u(x)$ such that

$$u(x) = \lim_{n \to \infty} U^n(x).$$

Using (7) this can also be expressed as

$$u(x) = u_a + \lim_{n\to\infty} \sum_{j=1}^{i} f(x_{j-1}^n) h_n, \tag{8}$$
$$\text{if } i \text{ is related to } n \text{ so that } x = x_i^n, \text{ that is, } i = \frac{x-a}{h_n} = \frac{x-a}{b-a} 2^n.$$

We remark that this is possible (i.e., i is an integer) only if x is a meshpoint. But the set of all meshpoints x_i^n for all n and i is very large and densely spread out in $[a, b]$, in fact all real numbers x can be approximated by a sequence of meshpoints. We don't prove this.

We have now *constructed a new function* $u(x)$.

Step 3. The limit solves the equation. We have to show that the new new function $u(x)$ satisfies the initial-value problem (1). First of all, we have $U^n(a) = u_a$ so that $u(a) = \lim_{n\to\infty} U^n(a) = u_a$, so that the initial condition is satisfied. Then we must show that u is differentiable and that $u'(x) = f(x)$. We will do this later, AMBS 28.8.

Step 4. Uniqueness. We must show that u is the only solution of (1). Assume then that v is another solution. This means that

$$\begin{cases} v'(x) = f(x), & x \in [a, b], \\ v(a) = u_a. \end{cases}$$

Since v is differentiable we have as in (6) and (7)

$$\begin{aligned} v(x) = v(x_i^n) &= v(x_{i-1}^n) + v'(x_{i-1}^n)(x_i^n - x_{i-1}^n) + E_v(x_i^n, x_{i-1}^n) \\ &= v(x_{i-1}^n) + f(x_{i-1}^n) h_n + E_v(x_i^n, x_{i-1}^n) \\ &= u_a + \sum_{j=1}^{i} f(x_{j-1}^n) h_n + \sum_{j=1}^{i} E_v(x_j^n, x_{j-1}^n), \end{aligned}$$

if i is related to n so that $x = x_i^n$, that is, $i = \dfrac{x-a}{b-a} 2^n = \dfrac{x-a}{h_n}$. We know from (8) that

$$u_a + \sum_{j=1}^{i} f(x_{j-1}^n) h_n \to u(x) \quad \text{as } n \to \infty.$$

For the remainder we have

$$\begin{aligned} \Big| \sum_{j=1}^{i} E_v(x_j^n, x_{j-1}^n) \Big| &\le \sum_{j=1}^{i} |E_v(x_j^n, x_{j-1}^n)| \le \sum_{j=1}^{i} K_v |x_j^n - x_{j-1}^n|^2 = \sum_{j=1}^{i} K_v h_n^2 \\ &= K_v h_n^2 i = K_v (i h_n) h_n = K_v (x-a) h_n \quad \Big\{\text{because } i = \frac{x-a}{h_n}\Big\} \\ &\le K_v h_n (b-a) \to 0. \end{aligned}$$

Therefore

$$v(x) = v(x_i^n) = u_a + \sum_{j=1}^{i} f(x_{j-1}^n) h_n + \sum_{j=1}^{i} E_v(x_j^n, x_{j-1}^n) \to u(x) \quad \text{as } n \to \infty.$$

We conclude that $v(x) = u(x)$ so there is only one solution.

The fundamental theorem of calculus. We can now formulate what we have done as a theorem.

Theorem. *Assume that* $f : [a, b] \to \mathbf{R}$ *is Lipschitz continuous and* u_a *is a number. Then there is a unique function* $u : [a, b] \to \mathbf{R}$ *such that*

$$\begin{cases} u'(x) = f(x), & x \in [a, b], \\ u(a) = u_a. \end{cases}$$

This function can be written

$$u(x) = u_a + \int_a^x f(y)\, dy, \tag{9}$$

where the integral is constructed as the limit

$$\int_a^x f(y)\, dy = \lim_{n\to\infty} \sum_{j=1}^{i} f(x_{j-1}^n) h_n,$$

where $h_n = (b-a)2^{-n}$, $x_j^n = a + jh_n$, *if* i *is related to* n *so that* $x = x_i^n$, *that is,* $i = \dfrac{x-a}{h_n} = \dfrac{x-a}{b-a}2^n$.

The function u *is uniformly differentiable on* $[a,b]$ *with constant* $K_u = \frac{1}{2}L_f$.

The function u is a *primitive function of* f. A primitive function of f is a function F such that $F' = f$. It is only determined up to a constant: $F(x) + C$ is also a primitive function. It can be made unique by adding an initial condition.

Example. $f(x) = x^2$ and $F(x) = \frac{1}{3}x^3$ is a primitive function. We set $u(x) = F(x) + C = \frac{1}{3}x^3 + C$ and $u(1) = 1$. Then $C = \frac{2}{3}$ and $u(x) = \frac{1}{3}x^3 + \frac{2}{3}$.

Sometimes we write (9) (with $x = b$) as

$$u(b) = u(a) + \int_a^b f(x)\,\mathrm{d}x = u(a) + \int_a^b u'(x)\,\mathrm{d}x,$$

which means that the integral of u' can be expressed as

$$\int_a^b u'(x)\,\mathrm{d}x = u(b) - u(a) = \Big[u(x)\Big]_{x=a}^{b}.$$

Here we introduced the convenient notation

$$\Big[u(x)\Big]_{x=a}^{b} = u(b) - u(a).$$

Example.

$$\int_1^2 x^2\,\mathrm{d}x = \left[\frac{x^3}{3}\right]_{x=1}^{2} = \frac{8}{3} - \frac{1}{3} = \frac{7}{3}.$$

It is important to note that the integration variable y in (9) is a “dummy variable” which can be replaced by anything but x and a:

$$\int_a^x f(y)\,\mathrm{d}y = \int_a^x f(s)\,\mathrm{d}s = \int_a^x f(\ddot{\mathrm{o}})\,\mathrm{d}\ddot{\mathrm{o}}.$$

Example.

$$\int_1^x y^2\,\mathrm{d}y = \left[\frac{y^3}{3}\right]_{y=1}^{x} = \frac{x^3}{3} - \frac{1}{3},$$
$$\int_1^x \ddot{\mathrm{o}}^2\,\mathrm{d}\ddot{\mathrm{o}} = \left[\frac{\ddot{\mathrm{o}}^3}{3}\right]_{\ddot{\mathrm{o}}=1}^{x} = \frac{x^3}{3} - \frac{1}{3}.$$

The same thing holds for sums:

$$\sum_{i=1}^{N} a_i = \sum_{n=1}^{N} a_n = a_1 + \cdots + a_N.$$

Interpretation as area. Assume now that $f : [a, b] \to \mathbf{R}_+$ is a nonnegative function. Then the sum in (9) (with $x = b = x_N^n$)

$$\sum_{j=1}^{N} f(x_{j-1}^n) h_n$$

is a sum of small rectangle areas $f(x_{j-1}^n) h_n$, each rectangle with base h_n and height $f(x_{j-1}^n)$. When $n \to \infty$ and $h_n \to 0$ this sum converges to the area under the graph $y = f(x)$, $x \in [a, b]$. On the other hand we know that the limit is

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{n \to \infty} \sum_{j=1}^{N} f(x_{j-1}^n) h_n.$$

We conclude that

$$A = \int_a^b f(x)\,\mathrm{d}x$$

is the area under the graph.

Example. The area under $y = x^2$, $x \in [1, 2]$, is

$$A = \int_1^2 x^2\,\mathrm{d}x = \Big[\frac{x^3}{3}\Big]_{x=1}^{2} = \frac{8}{3} - \frac{1}{3} = \frac{7}{3}.$$

`/stig`