COURSE 1

Outline of the course:

- Introductive notions. Finite and divided differences.

- Approximation of functions: interpolation of Lagrange, Hermite and Birkhoff type, Least squares approximation.

- Numerical integration. Newton-Cotes quadrature formulas. Repeated quadrature formulas. General quadrature formulas. Romberg's algorithm. Adaptive quadratures formulas. Gauss type quadrature formulas.

- Numerical methods for solving linear systems - direct methods (Gauss, Gauss-Jordan, LU-methods). Perturbations of a linear system.

- Numerical methods for solving linear systems - iterative methods (Jacobi, Gauss-Seidel, SOR).

- Methods for solving nonlinear equations in R: one-step methods (Newton (tangent) method) and multi-step methods (secant, bisection and false position methods). Lagrange, Hermite and Birkhoff inverse interpolation.

- Methods for solving nonlinear systems: successive approximation and Newton methods.

- Numerical methods for solving differential equations: Taylor interpolation, Euler and Runge-Kutta methods.

- Revision of the main types of problems.

Evaluation methods

• Written exam: 70%

• Lab activities (evaluation and continuous observations during the semester): 30%

- Each lab will be evaluated.

- Each lab should be delivered in the same week or the next one.

- Respect delivery dates for each lab assignment. Each delay will be penalized: 1 point/one week of delay.

- A lab that will not be delivered will have grade 1.

- You may deliver 2 lab assignments during one lab.

References:

1. I. Chiorean, T. Cătinaş, R. Trîmbitaş, *Analizā Numericā*, Ed. Presa Univ. Clujeanā, 2010.

2. W. Gander, M. Gander, F. Kwok, *Scientific Computing, An Introduction using Maple and MATLAB*, Springer, 2014.

3. R. L. Burden, J. D. Faires, *Numerical Analysis*, PWS Publishing Company, 2010.

4. R. Trîmbitaş, *Numerical Analysis*, Ed. Presa Univ. Clujeanā, 2007.

Chapter 1. Preliminary notions

We will study numerical methods and algorithms and analyze the error. In the most cases, finding an exact solution is not possible, so we approximate it, we find it numerically.

1.1. Preliminaries

Definition 1 *Let $x^* \in \mathbb{R}$ be an unknown value of interest. An element $\tilde{x} \in \mathbb{R}$ which approximates x^* is called* **the approximation** *or* **the approximant** *of x^*.*

The expression $\Delta x = x^* - \tilde{x}$ or $\Delta x = \tilde{x} - x^*$ is called **the error**.

The value $|\Delta x| = |x^* - \tilde{x}|$ is called **the absolute error** of approximation.

The value $\delta x = \frac{|\Delta x|}{|x^*|} = \frac{|x^* - \tilde{x}|}{|x^*|}$, $x^* \neq 0$ is called **the relative error** of approximation. *The relative error is a proportion, so we can also*

express it as a percentage by multiplying the relative error by 100%. *(The value* $100\% \cdot \delta x$ *is called* **the percent error** of approximation.)

The relative error is used to put error into perspective and it gives an indication of how good a measurement is relative to the size of the thing being measured.

For example, an error of $1cm$ would be a lot if the total length is $15cm$, but insignificant if the length was $5km$.

Example: Consider two approximative measurements of the weight $5.00g$: $5.05g$ and 4.95g. The absolute error is $0.05g$. The relative error is $0.05g/5.00g = 0.01$ or 1%.

The notions are correspondingly extended to normed space.

Definition 2 *If V is a K-linear space then a real functional $p : V \to [0, \infty)$, with the properties:*

$$1)\ p(v_1 + v_2) \leq p(v_1) + p(v_2), \quad \forall v_1, v_2 \in V,$$
$$2)\ p(\alpha v) = |\alpha|\, p(v), \quad \forall \alpha \in K, v \in V,$$
$$3)\ p(v) = 0 \Longrightarrow v = 0$$

is called **a norm** *on V.*

Definition 3 *Let K be a field and V be a given set. We say that V is a K-***linear space** (a linear space over K) *if there exist an internal operation:*

$$" + " : V \times V \to V; \quad (v_1, v_2) \to v_1 + v_2,$$

and an external operation:

$$" \cdot " : K \times V \to V; \quad (\alpha, v) \to \alpha v$$

that satisfy the following conditions:

1) $(V, +)$ is a commutative group

2)

$$a)\ (\alpha+\beta)v=\alpha v+\beta v,\quad \forall\alpha,\beta\in K,\quad \forall v\in V,$$
$$b)\ \alpha(v_1+v_2)=\alpha v_1+\alpha v_2,\quad \forall\alpha\in K,\quad \forall v_1,v_2\in V,$$
$$c)\ (\alpha\beta)v=\alpha(\beta v),\quad \forall\alpha,\beta\in K,\quad \forall v\in V,$$
$$d)\ 1\cdot v=v,\quad \forall v\in V.$$

The elements of V are called *vectors* and those of K are called *scalars*.

Definition 4 *Let V and V' be two K-linear spaces. A function $f:V\to V'$ is called* linear transformation *or* linear operator *if:*

1) $f(v_1+v_2)=f(v_1)+f(v_2),\quad \forall v_1,v_2\in V$ *(aditivity)*
2) $f(\alpha v)=\alpha f(v),\quad \forall\alpha\in K,\quad \forall v\in V$ *(homogenity).*

Or, shortly,

$$f(\alpha v_1+\beta v_2)=\alpha f(v_1)+\beta f(v_2),\quad \forall\alpha,\beta\in K,\quad \forall v_1,v_2\in V.$$

Definition 5 *Let V be a linear space on $\mathbb{R}$ or $\mathbb{C}$. A linear operator $P : V \to V$ is called* **projector** *if*

$$P \circ P = P, \quad \textit{(shortly, } P^2 = P\textit{);} \qquad \textit{(idempotence).}$$

Remark 6 *1) The identity operator $I : V \to V,\ I(v) = v$ and the null operator $0 : V \to V,\ 0(v) = 0$ are projectors.*

2) P is projector $\Rightarrow$ $P^C := I - P$, **the complement of** *P, is projector.*

1.2. Finite and divided differences

Finite differences

Let $M = \{a_i \mid a_i = a + ih, \text{ with } i = 0, ..., m;\ a, h \in \mathbb{R}^*,\ m \in \mathbb{N}^*\}$ and $\mathcal{F} = \{f \mid f : M \to \mathbb{R}\}$.

Definition 7 *For $f \in \mathcal{F}$,*

$$(\triangle_h f)(a_i) = f(a_{i+1}) - f(a_i), \qquad i < m$$

is called **the finite difference of the first order** *of the function f, with step h, at point a_i.*

Theorem 8 *The operator $\triangle_h$ is a linear operator with respect to f.*

Proof. If $f, g : M \to \mathbb{R};\ A, B \in \mathbb{R}$ and $i < m$, we have

$$\begin{aligned}(\triangle_h(Af + Bg))(a_i) &= (Af + Bg)(a_{i+1}) - (Af + Bg)(a_i) \\ &= A[f(a_{i+1}) - f(a_i)] + B[g(a_{i+1}) - g(a_i)] \\ &= A(\triangle_h f)(a_i) + B(\triangle_h g)(a_i).\end{aligned} \tag{1}$$

■

Definition 9 *Let* $0 \leq i < m,\ k \in \mathbb{N}$ *and* $1 \leq k \leq m - i$

$$\begin{aligned}(\triangle_h^k f)(a_i) &= (\triangle_h(\triangle_h^{k-1} f))(a_i) \\ &= (\triangle_h^{k-1} f)(a_{i+1}) - (\triangle_h^{k-1} f)(a_i), \quad \textit{with } \triangle_h^0 = I \textit{ şi } \triangle_h^1 = \triangle_h\end{aligned} \tag{2}$$

is called **the k-th order finite difference** *of the function* f*, with step* h*, at point* a_i*.*

Theorem 10 *If* $0 \leq i < m;\ k, p \in \mathbb{N}$ *and* $1 \leq p + k \leq m - i$*, then*

$$(\triangle_h^p(\triangle_h^k f))(a_i) = \triangle_h^k(\triangle_h^p f)(a_i) = (\triangle_h^{p+k} f)(a_i). \tag{3}$$

Finite differences table: (f_i denotes $f(a_i)$)

a	f	$\triangle_h f$	$\triangle_h^2 f$	...	$\triangle_h^{m-1} f$	$\triangle_h^m f$
a_0	f_0	$\triangle_h f_0$	$\triangle_h^2 f_0$	...	$\triangle_h^{m-1} f_0$	$\triangle_h^m f_0$
a_1	f_1	$\triangle_h f_1$	$\triangle_h^2 f_1$	...	$\triangle_h^{m-1} f_1$	
	...					
a_{m-3}	f_{m-3}	$\triangle_h f_{m-3}$	$\triangle_h^2 f_{m-3}$			
a_{m-2}	f_{m-2}	$\triangle_h f_{m-2}$	$\triangle_h^2 f_{m-2}$			
a_{m-1}	f_{m-1}	$\triangle_h f_{m-1}$				
a_m	f_m					

where

$$\triangle_h^k f_i = \triangle_h^{k-1} f_{i+1} - \triangle_h^{k-1} f_i, \;\; k = 1, ..., m; \;\; i = 0, 1, ..., m - k.$$

Examples.

1. Considering $h = 0.25$, $a = 1$, $a_i = a + ih$, $i = \overline{0,4}$, and $f_0 = 0$, $f_1 = 2$, $f_2 = 6$, $f_3 = 14$, $f_4 = 17$ form the finite differences table.

   Sol.: We get:

a	f	$\triangle_h f$	$\triangle_h^2 f$	$\triangle_h^3 f$	$\triangle_h^4 f$
1	0	2	2	2	-11
1.25	2	4	4	-9	
1.50	6	8	-5		
1.75	14	3			
2	17				

2. For $f(x) = e^x$ find $(\triangle_h^k f)(a_i)$, with $a_i = a + ih$, $i \in \mathbb{N}$.

Divided differences

Let $X = \{x_i \mid x_i \in \mathbb{R},\ i = 0, 1, ..., m,\ m \in \mathbb{N}^*\}$ and $f : X \to \mathbb{R}$.

Definition 11 *For $r \in \mathbb{N}$, $r < m$,*

$$(\mathcal{D}f)(x_r) := [x_r, x_{r+1}; f] = \frac{f(x_{r+1}) - f(x_r)}{x_{r+1} - x_r}$$

is called **the first order divided difference** *of the function f, regarding the points x_r and x_{r+1}.*

Theorem 12 *The operator $\mathcal{D}$ is linear with respect to f.*

Proof.

$$\begin{aligned}(\mathcal{D}(\alpha f + \beta g))(x_r) &= \frac{(\alpha f + \beta g)(x_{r+1}) - (\alpha f + \beta g)(x_r)}{x_{r+1} - x_r} \\ &= \alpha(\mathcal{D}f)(x_r) + \beta(\mathcal{D}g)(x_r), \qquad \text{for } \alpha, \beta \in \mathbb{R}.\end{aligned} \tag{4}$$

■

Definition 13 *Let* $r, k \in \mathbb{N}, 0 \leq r < m$ *and* $1 \leq k \leq m - r,\ m \in \mathbb{N}^*$. *The quantity*

$$(\mathcal{D}^k f)(x_r) = \frac{(\mathcal{D}^{k-1} f)(x_{r+1}) - (\mathcal{D}^{k-1} f)(x_r)}{x_{r+k} - x_r}, \quad \text{with } \mathcal{D}^0 = 1,\ \mathcal{D}^1 = \mathcal{D}, \tag{5}$$

is called **the k-th order divided difference** *of the function f, at x_r.*

$(\mathcal{D}^k f)(x_r)$ is also denoted by $\left[x_r, ..., x_{r+k}; f\right]$. Relation (5) can be written as

$$\left[x_r, ..., x_{r+k}; f\right] = \frac{\left[x_{r+1}, ..., x_{r+k}; f\right] - \left[x_r, ..., x_{r+k-1}; f\right]}{x_{r+k} - x_r}. \tag{6}$$

Remark 14 *The operator $\mathcal{D}^k$ is linear with respect to f.*

For $r = 0$ and $k = m$ we have

$$(\mathcal{D}^m f)(x_0) = \sum_{i=0}^{m} \frac{f(x_i)}{(x_i - x_0)...|...(x_i - x_m)}. \tag{7}$$

Theorem 15 *If* $f, g : X \to \mathbb{R}$ *then*

$$[x_0, ..., x_m; fg] = \sum_{k=0}^{m} [x_0, ..., x_k; f][x_k, ..., x_m; g].$$

Proof. The proof follows by complete induction with respect to m. ■

Table of divided differences:

x	f	$\mathcal{D}f$	$\mathcal{D}^2 f$	...	$\mathcal{D}^{\mathbf{m-1}} f$	$\mathcal{D}^m f$
x_0	f_0	$\mathcal{D}f_0$	$\mathcal{D}^2 f_0$	...	$\mathcal{D}^{m-1} f_0$	$\mathcal{D}^m f_0$
x_1	f_1	$\mathcal{D}f_1$	$\mathcal{D}^2 f_1$		$\mathcal{D}^{m-1} f_1$	
x_2	f_2	$\mathcal{D}f_2$	$\mathcal{D}^2 f_2$			
...	...	...				
x_{m-2}	f_{m-2}	$\mathcal{D}f_{m-2}$	$\mathcal{D}^2 f_{m-2}$			
x_{m-1}	f_{m-1}	$\mathcal{D}f_{m-1}$				
x_m	f_m					

with $f_i = f(x_i), \quad i = 0, 1, ..., m.$

Example 16 *For* $x_0 = 0$, $x_1 = 1$, $x_2 = 2$, $x_3 = 4$ *and* $f_0 = 3$, $f_1 = 4$, $f_2 = 7$, $f_3 = 19$ *form the divided differences table.*

x	f	Df	D^2f	D^3f
0	3	1	1	0
1	4	3	1	
2	7	6		
4	19			

Example 17 *Form the divided differences table for* $x_0 = 2$, $x_1 = 4$, $x_2 = 6$, $x_3 = 8$ *and* $f_0 = 4$, $f_1 = 8$, $f_2 = 20$, $f_3 = 48$.

Example 18 *Form the divided differences table for* $x_0 = 1$, $x_1 = 2$, $x_2 = 3$, $x_3 = 5$, $x_4 = 7$ *and* $f_0 = 3$, $f_1 = 5$, $f_2 = 9$, $f_3 = 11$, $f_4 = 15$.

Chapter 2. Polynomial interpolation

Interpolation is the science of *"reading between the lines of a mathematical table"* (E. Whittaker, G. Robinson)

Assume we know only some values $f(x_i)$, $i = 0, ..., m$ of a function f.

x	$x_0,$	$x_1,$	...	z	...	x_m
$y = f(x)$	$y_0,$	$y_1,$	...	?	...	y_m

Is there a way to compute or approximate the function value $f(z)$ for some given z without evaluating f?

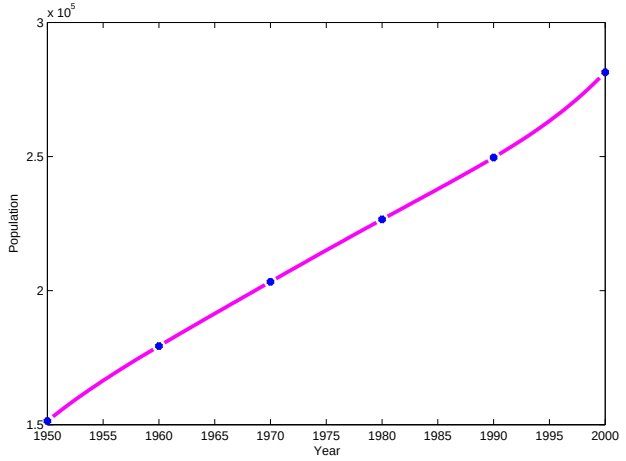
Applications:

1. approximating data at points where measurements are not available

2. sketching a function f which is expensive to evaluate if it is evaluated at some given points.

Example 19 *A census of the population of the United States is taken every 10 years. The following table lists the population, in thousands of people, from 1950 to 2000.*

1950	1960	1970	1980	1990	2000
151326	179323	203302	226542	249633	281422



Question: these data could be used to provide a reasonable estimate of the population in 1975? Answer: population in 1975 is 215042.

Predictions of this type can be obtained by using a function that fits the given data. This process is called **interpolation.** (If the desired

value z is within the range of the interpolation points x_i, then we have *interpolation;* if z is outside the range, the process is called *extrapolation.)*

Example 20 *a) Some values obtained by physical measurements: The temperature of the air outside a house during a day:*

t	8*am*	9*am*	11*am*	1*pm*	5*pm*
T *in* $^\circ C$	12.1	13.6	15.9	18.5	16.1

What temperature was at 10am?

b) Estimate $\sin 1$ *if we know* $\sin\frac{\pi}{6} = \frac{1}{2}, \sin\frac{\pi}{4} = \frac{\sqrt{2}}{2}$ *and* $\sin\frac{\pi}{3} = \frac{\sqrt{3}}{2}$.

c) Estimate $\log_{10}(z)$ *for values of the argument* z *that are intermediate between the tabulated values.*

One of the most useful classes of functions mapping the set of real numbers into itself is the polynomials.

Polynomials are used as the basic means of approximation in nearly all areas of numerical analysis: the solutions of equations, the approximation of functions, of integrals and derivatives, solutions of integral and differential equations, etc.

Polynomials owe this popularity to their simple structure, which makes it easy to construct effective approximations and then make use of them.

Advantages:

- Easy to handle calculus with polynomials: the derivative and the primitive of a polynomial are easy to determine and are also polynomials. The evaluation of a polynomial can be made efficient (the Horner scheme).

- Properly chosen, they can approximate arbitrary well any continuous function:

(Weierstrass Approximation Theorem) Given any $f \in C[a,b]$, $\forall \varepsilon > 0$ (arbitr. small) $\exists P(x)$ polynomial that is as "close" to f as desired:

$$|f(x) - P(x)| < \varepsilon, \quad \forall x \in [a,b].$$

Of course, the smaller ε, the greater the degree of P may become.

2.1. Taylor interpolation

Theorem 21 *(Taylor theorem) Let $f \in C^n[a,b]$, such that there exists $f^{(n+1)}$ on $[a,b]$ and consider $x_0 \in [a,b]$. The Taylor polynomial is*

$$T_n(x) = \sum_{k=0}^{n} \frac{(x-x_0)^k}{k!} f^{(k)}(x_0) \tag{8}$$

and we have the approximation formula

$$f(x) = T_n(x) + R_n(x),$$

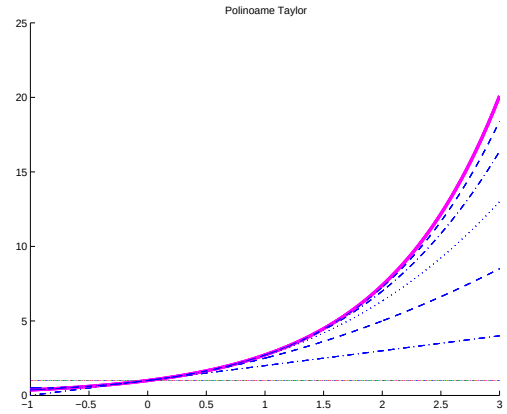
R_n denoting the remainder (the error).

For $\forall x \in [a,b]$ there exists a number ξ between x_0 and x such that

$$R_n(x) = \frac{(x-x_0)^{n+1}}{(n+1)!} f^{(n+1)}(\xi).$$

Remark 22 *Taylor polynomials agree as closely as possible with a given function around the specific point x_0, but not on the entire interval.*

Example 23 *We calculate the first six Taylor polynomials about* $x_0 = 0$ *for* $f(x) = e^x$.



Notice that even for the higher-degree polynomials, the error becomes progressively worse as we move away from $x_0 = 0$.

Example 24 *Consider* $f(x) = \frac{1}{x}$ *and* $x_0 = 1$. *Approximate the value of* $f(3)$ *by the first and the second degree Taylor polynomials.*

n	0	1	2
$T_n(3)$	1	-1	3

Taylor polynomial approximation is used when approximations are needed only at numbers close to x_0. It is more efficient to use methods that include information at various points.

2.2. Lagrange interpolation

Let $[a,b] \subset \mathbb{R},\ x_i \in [a,b],\ i = 0,1,...,m$ such that $x_i \neq x_j$ for $i \neq j$ and consider $f : [a,b] \to \mathbb{R}$.

The Lagrange interpolation problem (LIP) consists in determining the polynomial P of the smallest degree for which

$$P(x_i) = f(x_i),\ i = 0,1,...,m \tag{9}$$

i.e., the polynomial of the smallest degree which passes through the distinct points $(x_i, f(x_i))$, $i = 0,1,...,m$.

Since in (9) there are $m+1$ conditions to be satisfied, we need $m+1$ degrees of freedom. Consider the m-th degree polynomial

$$P(x) = a_0 + a_1 x + ... + a_{m-1} x^{m-1} + a_m x^m. \tag{10}$$

The $m+1$ coefficients $\{a_i\}$ have to be determined in such way that (9) are satisfied. This leads to the linear system of equations:

$$\begin{cases} a_0 + a_1 x_0 + ... + a_{m-1} x_0^{m-1} + a_m x_0^m = f(x_0) \\ a_0 + a_1 x_1 + ... + a_{m-1} x_1^{m-1} + a_m x_1^m = f(x_1) \\ \\ a_0 + a_1 x_m + ... + a_{m-1} x_m^{m-1} + a_m x_m^m = f(x_m). \end{cases}$$

Written in the matrix form, the system is

$$\underbrace{\begin{pmatrix} 1 & x_0 & ... & x_0^{m-1} & x_0^m \\ 1 & x_1 & ... & x_1^{m-1} & x_1^m \\ \vdots & & & & \\ 1 & x_m & ... & x_m^{m-1} & x_m^m \end{pmatrix}}_{V} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{pmatrix} = \begin{pmatrix} f(x_0) \\ f(x_1) \\ \vdots \\ f(x_m). \end{pmatrix}.$$

The matrix V with the special structure containing the powers of the nodes is called a Vandermonde Matrix.

Remark 25 *For $m+1$ distinct nodes the Vandermonde matrix is non-singular and there exists a unique interpolating polynomial P of degree less or equal to m with $P(x_i) = f(x_i),\ i = 0, 1, ..., m.$*

Definition 26 *A solution of (LIP) is called* **Lagrange interpolation polynomial**, *denoted by* $L_m f$.

Remark 27 *We have* $(L_m f)(x_i) = f(x_i),\ i = 0, 1, ..., m.$

$L_m f \in \mathbb{P}_m$ ($\mathbb{P}_m$ *is the space of polynomials of at most* m*-th degree*).

The Lagrange interpolation polynomial is given by

$$(L_m f)(x) = \sum_{i=0}^{m} \ell_i(x) f(x_i), \tag{11}$$

where by $\ell_i(x)$ denote **the Lagrange fundamental interpolation polynomials.**

We have

$$u(x) = \prod_{j=0}^{m} (x - x_j),$$

$$u_i(x) = \frac{u(x)}{x - x_i} = (x - x_0)...(x - x_{i-1})(x - x_{i+1})...(x - x_m) = \prod_{\substack{j=0 \\ j \neq i}}^{m} (x - x_j)$$

and

$$\ell_i(x) = \frac{u_i(x)}{u_i(x_i)} = \frac{(x-x_0)...(x-x_{i-1})(x-x_{i+1})...(x-x_m)}{(x_i-x_0)...(x_i-x_{i-1})(x_i-x_{i+1})...(x_i-x_m)} = \prod_{\substack{j=0 \\ j\neq i}}^{m} \frac{x-x_j}{x_i-x_j}, \tag{12}$$

for $i = 0, 1, ..., m$.

How do we know that the interpolation polynomial expanded in powers of x as in (10) and the polynomial constructed as in (11) represent the same polynomial?

Assume we have computed two interpolating polynomials $Q(x)$ and $P(x)$ each of degree m such that

$$Q(x_j) = f(x_j) = P(x_j), \quad j = 0, ..., m.$$

Then we can form the difference

$$d(x) = Q(x) - P(x),$$

that is a polynomial of degree less or equal to m.

Because of the interpolation property of P and Q, we have

$$d(x_j) = Q(x_j) - P(x_j) = 0, \quad j = 0, ..., m.$$

A non-zero polynomial of degree less than or equal to m cannot have more than m zeros. But d has $m + 1$ distint zeros, hence it must be indentically zero, so $Q(x) = P(x)$.

Remark 28 *Because the Vandermonde Matrix is ill conditioned the method (11) is recomended for computing the Lagrange polynomial.*