Let p^* be an approximation to a number p (not necessarily its floating point approximation). Then:

- **Absolute error:** $|p - p^*|$
- **Relative error:** $\frac{|p-p^*|}{|p|}$ (provided $p \neq 0$)

Significant Digits. The number p^* is said to approximate p to t *significant digits* (or figures) if t is the largest nonnegative integer such that

$$\frac{|p - p^*|}{|p|} \leq 5 \times 10^{-t}.$$

$$\cos(h) = 1 - \frac{h^2}{2!} + \frac{h^4}{4!} - \frac{h^6}{6!} + \cdots$$
$$\sin(h) = h - \frac{h^3}{3!} + \frac{h^5}{5!} - \frac{h^7}{7!} + \cdots$$
$$\tan(h) = h + \frac{h^3}{3} + \frac{2h^5}{15} + \frac{17h^7}{315} + \cdots$$
$$\sinh(x) = \frac{e^x - e^{-x}}{2}$$
$$\cosh(x) = \frac{e^x + e^{-x}}{2}$$
$$e^h = 1 + h + \frac{h^2}{2!} + \frac{h^3}{3!} + \frac{h^4}{4!} + \cdots$$
$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots$$

Suppose $\{\beta_n\}_{n\geq 1}$ is a sequence known to converge to 0 and $\{\alpha_n\}_{n\geq 1}$ a sequence that converges to a value $\alpha \in \mathbb{R}$.
If there exists a constant $K > 0$ such that

$$|\alpha_n - \alpha| \leq K|\beta_n| \quad \text{for large } n,$$

then we say that $\{\alpha_n\}$ converges to α with **rate of convergence** $\mathcal{O}(\beta_n)$. This is denoted as:

$$\alpha_n = \alpha + \mathcal{O}(\beta_n).$$

Order of Convergence. Let $\{p_n\}_{n=0}^{\infty}$ be a sequence with $\lim_{n\to\infty} p_n = p$ and $p_n \neq p$ for all n. If there exist positive constants λ and α such that

$$\lim_{n\to\infty} \frac{|p_{n+1} - p|}{|p_n - p|^\alpha} = \lambda,$$

then we say that $\{p_n\}_{n=0}^{\infty}$ *converges to p with order α and constant λ.*

- In general, the larger α, the more rapid the convergence.
- The constant λ affects this speed, but generally it is not as important as the order α.

The following statements are equivalent for a matrix $A \in \mathbb{R}^{n\times n}$:

- $\det(A) \neq 0$
- The system $Ax = 0$ has a unique solution $x = 0$ (i.e., the column vector with all zero entries)
- The system $Ax = b$ has a unique solution for any n-dimensional column vector b
- The matrix A is *nonsingular*, that is, A^{-1} exists

If GE can be performed on the system $Ax = b$ *without row interchanges*, then the matrix A can be factored as:

$$A = LU, \quad \Rightarrow \quad L(Ux) = b$$

Let $A \in \mathbb{R}^{n\times n}$. A is said to be **strictly diagonally dominant** if

$$|a_{ii}| > \sum_{\substack{j=1\\ j\neq i}}^{n} |a_{ij}| \quad \text{for all } i = 1, \ldots, n,$$

A *strictly diagonally dominant matrix* A is *nonsingular*. Moreover, in this case, GE can be performed on any linear system $Ax = b$ without row or column *interchanges*, and the computations will be robust to *round-off errors*.

If $A \in \mathbb{R}^{n\times n}$ is **positive definite**, then:

- A is *symmetric*, i.e. $A = A^\top$.
- $\mathbf{x}^\top A\mathbf{x} > 0$ for all nonzero $\mathbf{x} \in \mathbb{R}^n$.
- A is *nonsingular*.
- $a_{ii} > 0$ for each $i = 1, \ldots, n$.
- $\max\limits_{1\leq i,j\leq n} |a_{ij}| \leq \max\limits_{1\leq i\leq n} |a_{ii}|$.
- $(a_{ij})^2 < a_{ii}a_{jj}$ for each $i \neq j$.

Cauchy–Schwarz Inequality. For each $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,

$$\left|\mathbf{x}^\top \mathbf{y}\right| = \left|\sum_{i=1}^n x_i y_i\right| \leq \left(\sum_{i=1}^n x_i^2\right)^{1/2} \left(\sum_{i=1}^n y_i^2\right)^{1/2} = \|\mathbf{x}\|_2 \, \|\mathbf{y}\|_2.$$

Matrix Norms and Distances. A *matrix norm* on $\mathbb{R}^{n\times n}$ is a function $\|\cdot\| : \mathbb{R}^{n\times n} \to \mathbb{R}$ such that for any $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}$ and any $\alpha \in \mathbb{R}$:

- $\|\mathbf{A}\| \geq 0$,
- $\|\mathbf{A}\| = 0$ if and only if $\mathbf{A} = \mathbf{0}$,
- $\|\alpha\mathbf{A}\| = |\alpha|\|\mathbf{A}\|$,
- $\|\mathbf{A} + \mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|$ (triangle inequality),
- $\|\mathbf{AB}\| \leq \|\mathbf{A}\| \cdot \|\mathbf{B}\|$ (submultiplicative property).

Spectral radius $\rho(A)$ of a matrix A is defined by

$$\rho(A) = \max |\lambda|,$$

where λ is an eigenvalue of $\mathbf{A}$.
Let $A \in \mathbb{R}^{n\times n}$. Then,

$$\|A\|_2 = \sqrt{\rho(A^\top A)}.$$

Furthermore, for any natural norm $\|\cdot\|$, we have

$$\rho(A) \leq \|A\|.$$

Condition Number (ℓ_2 norm). Let $A \in \mathbb{R}^{n\times n}$ be nonsingular. The condition number of A (w.r.t. ℓ_2) is

$$\kappa(A) = \|A\|_2 \, \|A^{-1}\|_2 = \sqrt{\frac{\lambda_{\max}(A^\top A)}{\lambda_{\min}(A^\top A)}},$$

where $\lambda_{\max}(A^\top A)$ and $\lambda_{\min}(A^\top A)$ are the largest and smallest eigenvalues (in absolute value) of $A^\top A$.

- Well-conditioned if $\kappa(A) \approx 1$.
- Ill-conditioned if $\kappa(A) \gg 1$.

$$A = D - L - U \quad \text{for iterative}$$

1. **Jacobi:**
$$x^{(k+1)} = D^{-1}(L+U)x^{(k)} + D^{-1}b$$
2. **Gauss-Seidel:**
$$x^{(k+1)} = (D-L)^{-1}Ux^{(k)} + (D-L)^{-1}b$$
3. **Successive Over-Relaxation (SOR):**
$$x^{(k+1)} = (D - \omega L)^{-1}\left[(1-\omega)D + \omega U\right]x^{(k)} + \omega(D - \omega L)^{-1}b$$

For any $\mathbf{x}^{(0)} \in \mathbb{R}^n$, the sequence $\{\mathbf{x}^{(k)}\}_{k=1}^{\infty}$ defined by the iteration

$$\mathbf{x}^{(k+1)} = T\mathbf{x}^{(k)} + \mathbf{c}$$

converges to the unique solution of

$$\mathbf{x} = T\mathbf{x} + \mathbf{c}$$

if and only if $\rho(T) < 1$.

Newton's Method.
Given a function $f(x) \in C^2[a,b]$, Newton's method generates a sequence:

$$p_n = p_{n-1} - \frac{f(p_{n-1})}{f'(p_{n-1})}, \quad n = 1, 2, \ldots$$

Convergence: Quadratic if $f(p) = 0$, $f'(p) \neq 0$, and p_0 is close to the root.

Secant Method (derivative-free):

$$p_n = p_{n-1} - f(p_{n-1}) \cdot \frac{p_{n-1} - p_{n-2}}{f(p_{n-1}) - f(p_{n-2})}$$

Convergence: Superlinear, with order approximately $\alpha = \frac{1+\sqrt{5}}{2} \approx 1.618$

Modified Newton Method (for root of multiplicity m).
If $f(x)$ has a root of multiplicity m at p, standard Newton converges only linearly.
To restore quadratic convergence:

$$p_n = p_{n-1} - \frac{mf(p_{n-1})}{f'(p_{n-1})}, \quad \text{(if multiplicity } m \text{ is known)}$$

Alternative: Apply Newton to $\mu(x) = \frac{f(x)}{f'(x)}$. This leads to:

$$p_n = p_{n-1} - \frac{f(p_{n-1})f'(p_{n-1})}{(f'(p_{n-1}))^2 - f(p_{n-1})f''(p_{n-1})}$$

Note: Requires second derivative $f''(x)$; restores quadratic convergence.

Lagrange Interpolation Formula.
Given $n+1$ distinct data points $(x_0, f(x_0))$, $(x_1, f(x_1))$, …, $(x_n, f(x_n))$, the unique polynomial $P(x) \in \mathbb{P}_n$ that interpolates the data is:

$$P(x) = \sum_{j=0}^{n} f(x_j) L_j(x)$$

where the Lagrange basis polynomials $L_j(x)$ are defined by:

$$L_j(x) = \prod_{\substack{0 \le k \le n \\ k \ne j}} \frac{x - x_k}{x_j - x_k}$$

Numerical Differentiation (Equally Spaced Nodes, $n = 2$). If $x_1 = x_0 + h$ and $x_2 = x_0 + 2h$, then:

$$f'(x_0) = \frac{1}{h}\left(-\frac{3}{2}f(x_0) + 2f(x_0+h) - \frac{1}{2}f(x_0+2h)\right) + \frac{h^2}{3}f^{(3)}(\xi(x_0)),$$
$$f'(x_1) = \frac{1}{h}\left(-\frac{1}{2}f(x_1-h) + \frac{1}{2}f(x_1+h)\right) - \frac{h^2}{6}f^{(3)}(\xi(x_1)),$$
$$f'(x_2) = \frac{1}{h}\left(\frac{1}{2}f(x_2-2h) - 2f(x_2-h) + \frac{3}{2}f(x_2)\right) + \frac{h^2}{3}f^{(3)}(\xi(x_2)).$$

Divided Differences & Newton Form (compact). Let distinct nodes $x_0, \dots, x_n$. Define divided differences recursively

$$f[x_i] = f(x_i),\ f[x_i, \dots, x_{i+k}] = \frac{f[x_{i+1}, \dots, x_{i+k}] - f[x_i, \dots, x_{i+k-1}]}{x_{i+k} - x_i}$$

The Newton interpolant is

$$P_n(x) = \sum_{k=0}^{n} f[x_0, \dots, x_k] \prod_{j=0}^{k-1}(x - x_j),$$

Cubic Spline Interpolation Types. Given data points $(x_0, y_0), \dots, (x_n, y_n)$, a cubic spline consists of piecewise cubics $S_i(x)$ with continuous first and second derivatives.

- **Natural (Free) Spline:**
  $S''(x_0) = 0$ and $S''(x_n) = 0$
- **Clamped (Complete) Spline:**
  $S'(x_0) = f'(x_0)$ and $S'(x_n) = f'(x_n)$ (derivatives specified)
- **Not-a-Knot Spline:**
  S''' is continuous at x_1 and x_{n-1} (i.e., $S_0'''(x_1) = S_1'''(x_1)$, $S_{n-2}'''(x_{n-1}) = S_{n-1}'''(x_{n-1})$)

Numerical Differentiation via Lagrange Interpolation.
Let $x_0, x_1, \dots, x_n \in [a, b]$ and suppose $f \in C^{n+1}[a, b]$. Then the Lagrange interpolation formula gives:

$$f(x) = \sum_{k=0}^{n} f(x_k)\, L_k(x) + \frac{f^{(n+1)}(\xi(x))}{(n+1)!}(x - x_0)(x - x_1)\cdots(x - x_n)$$

for some $\xi(x) \in [a, b]$, where $L_k(x)$ are the Lagrange basis polynomials. Differentiating and evaluating at $x = x_j$, we obtain:

$$f'(x_j) = \sum_{k=0}^{n} f(x_k)\, L_k'(x_j) + \frac{f^{(n+1)}(\xi)}{(n+1)!} \prod_{\substack{0 \le k \le n \\ k \ne j}} (x_j - x_k)$$

This is an $(n+1)$-point formula for estimating $f'(x_j)$ from values $f(x_0), \dots, f(x_n)$.

Gaussian Quadrature (5-point Rule).
Gaussian quadrature approximates

$$\int_{-1}^{1} f(x)\, dx \approx \sum_{i=1}^{n} w_i f(x_i)$$

where x_i are the roots of the nth Legendre polynomial $P_n(x)$, and w_i are weights.

Change of Interval: To apply over interval $[a, b]$, transform:

$$\int_a^b f(x)\, dx = \frac{b-a}{2} \int_{-1}^{1} f\left(\frac{b-a}{2}t + \frac{a+b}{2}\right) dt$$

where the *nodes* x_i are the n distinct zeros of the Legendre polynomial $P_n(x)$ (orthogonal on $[-1, 1]$ with weight 1), and the *weights* are

$$w_i \;=\; \frac{2}{\left(1 - x_i^2\right)\left[P_n'(x_i)\right]^2}.$$

Legendre polynomials (Bonnet's recursion).

$$P_0(x) = 1, \qquad P_1(x) = x,$$
$$(n+1)P_{n+1}(x) = (2n+1)\, x\, P_n(x) - n\, P_{n-1}(x) \quad (n \ge 1).$$

Useful derivative identity (for weights):

$$P_n'(x) = \frac{n}{1 - x^2}\left(P_{n-1}(x) - x\, P_n(x)\right).$$

Degree of Precision (Exactness). A quadrature rule

$$\int_a^b f(x)\, dx \approx \sum_{i=1}^{n} w_i f(x_i)$$

is said to have *degree of precision* m if it integrates exactly every polynomial $p(x)$ with $\deg p \le m$, but fails for some polynomial of degree $m + 1$.

For Gaussian quadrature with n nodes, the degree of precision is $2n - 1$.

Lipschitz Condition. A function $f(t, y)$ satisfies a Lipschitz condition in y if there exists $L > 0$ such that

$$|f(t, y_1) - f(t, y_2)| \le L|y_1 - y_2| \quad \text{for all } y_1, y_2 \in \mathbb{R},\ a \le t \le b.$$

Then we say f is *Lipschitz in* y. Or, sufficient condition $\left|\frac{\partial f}{\partial y}(t, y)\right| \le L$

Existence and Uniqueness Theorem. Let $D = \{(t, y) \in \mathbb{R}^2 : a \le t \le b, -\infty < y < \infty\}$. Suppose $f(t, y)$ is continuous on D, and Lipschitz in y. Then the initial value problem:

$$y'(t) = f(t, y), \quad y(a) = \alpha, \quad a \le t \le b,$$

has a unique solution $y(t)$ on $[a, b]$.

Well-posedness: The same conditions (continuity and Lipschitz) imply the IVP is *well posed*: small changes in inputs (e.g. initial value or $f(t, y)$) lead to small changes in solution $y(t)$.

Euler's Method (for IVPs).
Consider the initial value problem (IVP):

$$y'(t) = f(t, y), \quad y(a) = \alpha, \quad a \le t \le b.$$

Let $h > 0$ be the *step size*, assumed constant for simplicity:

$$h = \frac{b-a}{N}, \quad t_j = a + jh, \quad j = 0, 1, \dots, N.$$

We seek approximations $y_j \approx y(t_j)$ via:

$$y_{j+1} = y_j + h\, f(t_j, y_j), \quad j = 0, 1, \dots, N-1.$$

Remarks:

- Local truncation error: $\mathcal{O}(h^2)$
- Global truncation error: $\mathcal{O}(h)$ (first-order method)
- Variable step sizes h_j may also be used when needed.

Runge–Kutta Methods (for IVPs).
Consider the IVP

$$y'(t) = f(t, y), \quad y(a) = \alpha, \quad a \le t \le b$$

with step size $h = \frac{b-a}{N}$, $t_i = a + ih$.

General explicit two-stage RK method:

$$k_1 = h\, f(t_i, y_i),$$
$$k_2 = h\, f(t_i + \alpha h,\, y_i + \beta k_1),$$
$$y_{i+1} = y_i + a\, k_1 + b\, k_2.$$

Choose a, b, α, β to maximize local order.

Two common $\mathcal{O}(h^2)$ choices:

- **Midpoint method:** $a = 0,\ b = 1,\ \alpha = \beta = \frac{1}{2}$

$$k_1 = h\, f(t_i, y_i), \quad k_2 = h\, f\left(t_i + \frac{h}{2},\, y_i + \frac{k_1}{2}\right), \quad y_{i+1} = y_i + k_2.$$

- **Modified Euler:** $a = b = \frac{1}{2},\ \alpha = \beta = 1$

$$k_1 = h\, f(t_i, y_i), \quad k_2 = h\, f(t_i + h,\, y_i + k_1), \quad y_{i+1} = y_i + \frac{k_1 + k_2}{2}.$$