

Things to Remember

- a_{ij} is the i th row and j th column of A .
- $\pm 0.d_1d_2 \dots d_k \times 10^n$ is the decimal floating point representation of a number.
- Chopping is cheaper than rounding.

Error

- Error: $p - \hat{p}$
- Abs. Err: $|p - \hat{p}|$
- Rel. Err: $\frac{|p - \hat{p}|}{p}$ (for accuracy)

Significant Digits

An aprxmtn \hat{p} has t significant digits if: $\frac{|p - \hat{p}|}{|p|} \leq 5 \times 10^{-t}$

Catastrophic Cancellation (Roundoff)

When subtracting nearly equal numbers, the relative error is large, and you lose a lot of significant digits (and accuracy).

How to Reduce Errors

- Reformat the formula to avoid roundoff
 - Reduce num. of ops (avoid rounding)
 - Nested Arithmetic: Rewrite polynomials to reduce operations
- $$x^3 - 6.1x^2 + 3.2x \rightarrow ((x - 6.1)x + 3.2)x$$

Algorithms and Convergence

- Stable \rightarrow errors grow linearly
- Unstable \rightarrow errors grow exponentially

Rate of Convergence

- For sequences, if $\alpha_n \rightarrow \alpha$ and $|\alpha_n - \alpha| \leq k\beta_n$, $\beta_n \rightarrow 0$ then α_n is $\mathcal{O}(\beta_n)$
- For functions, if $\lim_{h \rightarrow 0} f(h) = L$ and $|f(h)| \leq kh^p$ then $f(h) = L + \mathcal{O}(h^p)$

Taylor Series

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} (x - a)^n$$
$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$
$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} \quad \cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!}$$
$$\ln(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} + \dots$$
$$(1 + x)^{-p} = 1 - px + \frac{p(p+1)x^2}{2} - \frac{p(p+1)(p+2)x^3}{3!} \dots$$

The **Error Term** is the $(n + 1)^{th}$ term.

Root Finding

- Find p such that $f(p) = 0$.

Generic Stopping Criterion

- $\frac{|p_n - p_{n-1}|}{|p_n|} \leq \mathcal{E}; p_n \neq 0$: relative error
- $|f(p_n)| \leq \mathcal{E}$
 - Ensures small $f(p_n)$
 - p_n may differ significantly from p
- Have a fixed number of iterations
- (bisection) $\frac{b_n - a_n}{2} \leq \mathcal{E}$ or $|p_n - p_{n-1}| < \mathcal{E}$
 - Ensures p_n is within \mathcal{E} of p
 - Does not ensure small $f(p_n)$

Bisection Method:

- Conditions:** $f(x) \in C[a, b]$;
 $f(a)$ and $f(b)$ have opposite signs.
- Midpoint:** $x = \frac{a+b}{2}$
- Procedure:** Binary search for the root.
- Error:** Guaranteed quadratic convergence
- Error Formula:** $\frac{b-a}{2^n}$

Newton's Method

- Faster than bisection, quadratic. We follow the tangent line at p_{n-1} to its x -intercept.
 - Requires $f'(p)$ to exist.
 - Requires $f''(p)$ for quadratic convergence.
 - Start with initial guess p_0 and p_1
- $$2. p_n = p_{n-1} - \frac{f(p_{n-1})(p_{n-1} - p_{n-2})}{f'(p_{n-1})f(p_{n-2})}$$

Secant Method

- Does not require $f'(p)$ to exist.
 - Faster than Bisection, order $\phi \approx 1.618$
 - Start with initial guess p_0 and p_1
- $$2. p_n = p_{n-1} - \frac{f(p_{n-1})(p_{n-1} - p_{n-2})}{f(p_{n-1})f(p_{n-2})}$$
- False Position Method**
- Linear or sublinear guaranteed convergence.
 - Convergence can stall if the function has poor behaviour near the root.
 - Start with initial guess p_0 and p_1
- $$2. p_n = p_{n-1}f(p_{n-1}) \cdot \frac{p_{n-1} - p_{n-2}}{f(p_{n-1}) - f(p_{n-2})}$$

Fixed Points

- Start with initial guess p_0
- Generate a sequence $p_n = g(p_{n-1})$
- Stop when $|p_n - p_{n-1}| < \mathcal{E}$
 - A fixed point of f is a point p such that $f(p) = p$.
- Converges if:
 - $g : [a, b] \rightarrow [a, b]$ is continuous
 - $\forall x \in [a, b] : |g'(x)| \leq k < 1$
 - $f(x) = 0$ can be rewritten as $g(x) = x$
- Error:** $\mathcal{O}(q^n)$, for some q , faster when q is small

Norms

Vector Norms

- $l_1 : \|x\|_1 = \sum x_i$
- $l_2 : \|x\|_2 = \sqrt{x_1^2 + \dots + x_n^2}$ (Euclidean)
- $l_\infty : \|x\|_\infty = \max\{|x_1|, \dots, |x_n|\}$ (∞)

- Properties**
- Scalability: $\|\alpha x\| = |\alpha| \|x\|$
- Triangle Inequality: $\|x + y\| \leq \|x\| + \|y\|$

Vector Distances

- l_α distance: $\|x - y\|_\alpha$

Matrix Norms

- The Natural Norm $\|\cdot\|_*$ for $A, B \in \mathbb{R}^{n \times n}; \alpha \in \mathbb{R}$ is defined as a function that satisfies:
 - $\|A\| \geq 0$
 - $\|A\| = 0 \iff A = 0$
 - $\|\alpha A\| = |\alpha| \|A\|$
 - $\|A + B\| \leq \|A\| + \|B\|$
- Def.** $\|A\|_* = \max_{\|x\|=1} \|Ax\|_*$ where $\|Ax\|$ is any vector norm.
- $l_\infty : \|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|$ (row sum)

Special Properties

- For any natural norm $\|\cdot\|_\alpha : \rho(A) \leq \|A\|_\alpha$
- For $l_2 : \|A\|_2 = \sqrt{\rho(A^T A)}$

Vector Sequence Convergence

- $\{x^{(k)}\}$ converges to x for any small $\mathcal{E} > 0$ eventually every $x^{(k)}$ is within \mathcal{E} of x

Eigenvalues and Eigenvectors

E.value (λ): Scalar s.t. $A\vec{x} = \lambda\vec{x}$
E.vector (\vec{x}): Nonzero vector only scaled by A
Spectral Radius: $\rho(A) = \max\{|\lambda_i|\}$

Properties

- $\det(A - \lambda I) = 0 \iff \lambda$ is an eigenvalue. Solve the characteristic polynomial for λ .
- $\forall \lambda [(A - \lambda I)\vec{x} = 0 \iff \vec{x}$ is an eigenvector]
- If $\rho A < 1$, A is convergent $\implies \lim_{k \rightarrow \infty} A^k = 0$

Linear Systems - Pivoting Strategies

If the pivot is small, large errors can occur. Pivoting helps maintain numerical stability.

Partial Pivoting

Choose the largest element in the **current column** (below or at the pivot) to avoid dividing by a small number.

- For $k = 1 \dots n - 1$:
 - Find $r = \arg \max\{|a_{ik}|\}$
 $k \leq i \leq n$

- If $r \neq k$, swap rows: $E_k \leftrightarrow E_r$
- Continue Gaussian Elimination as usual

Scaled Partial Pivoting

Handles rows with vastly different magnitudes by normalizing.

- For each row $i = 1 \dots n$, compute the scale factor: $s_i = \max_j |a_{ij}|$
- For pivot column k , choose the row r such that $\frac{|a_{rk}|}{s_r}$ is maximal for $r \geq k$
- If $r \neq k$, swap rows: $E_k \leftrightarrow E_r$
- Proceed with Gaussian Elimination

Full Pivoting

Most stable but most expensive. Swap both rows and columns.

- At step k , find the largest element $|a_{ij}|$ in the submatrix $A_{k:n, k:n}$
- Swap row k with row i , and column k with column j
- Update row and column permutations
- Continue Gaussian Elimination

Linear Algebra

- To multiply $A \cdot B$, dot-product the rows of A by the columns of B .
- $AA^{-1} = A^{-1}A = I$
- To find A^{-1} , row reduce the aug. matrix $[A|I]$.
- A^T is A flipped over the main diagonal.

Determinant

- $\det(A) \neq 0 \implies \begin{cases} A^{-1} & \text{exists} \\ Ax = b & \text{has a unique solution} \end{cases}$

- Cofactor Expansion (Laplace Expansion):**
For an $n \times n$ matrix $A = [a_{ij}]$, the determinant is given by:

$$\det(A) = \sum_{j=1}^n a_{ij} C_{ij} \quad (\text{row expansion on row } i)$$

or

$$\det(A) = \sum_{i=1}^n a_{ij} C_{ij} \quad (\text{column expansion on column } j)$$

where $C_{ij} = (-1)^{i+j} \det(M_{ij})$ is the cofactor, and M_{ij} is the $(n - 1) \times (n - 1)$ minor formed by deleting the i -th row and j -th column of A .

Matrix Factorization

LU Decomposition

Row Swaps

Special Matrices

Permutation Matrices

Singular

Banded Matrices

Tridiagonal Matrices

Diagonally Dominant Matrices

Positive Definite Matrices

Iterative Methods for Linear Systems

Jacobi Method

Gauss-Seidel Method