

CS 371

Condition of a problem

$$\kappa_A = \frac{\|\Delta z\|}{\|\Delta x\|}$$

$$\kappa_R = \left(\frac{\|\Delta z\|}{\|z\|} \right)$$

$$\left(\frac{\|\Delta x\|}{\|x\|} \right)$$

$$\Delta x = x - \hat{x} \quad \delta x = \frac{|\Delta x|}{x}$$

$$\|x\|_1 = \sum_{i=1}^n |x_i| \quad \|x\|_p = \sqrt[p]{\sum_{i=1}^n x_i^p}$$

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$$

$$|x \cdot y| \leq \|x\| \times \|y\| \quad (\text{Cauchy-Schwartz Inequality.})$$

$$\|x\|_\infty \leq \|x\|_2 \leq \|x\|_1 \leq n \|x\|_\infty$$

$$\|x\|_\infty \leq \|x\|_p \leq n \|x\|_\infty$$

$$e^x = \sum_{n=1}^{\infty} \frac{x^n}{n!}$$

$$\cos(x) = \sum_{n=1}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots$$

$$\sin(x) = \sum_{n=1}^{\infty} \sum_{n=1}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots$$

Numerical Linear Algebra

$$\underbrace{\underbrace{A}_{n \times p}}_{p \times m} \underbrace{B}_{p \times m} \quad \text{flops} = nm(p+p-1)$$

LU Factorization

$$PA = \underbrace{LU}_{\text{Lower and upper triangular matrix}}$$

$$\mathbf{A} \xrightarrow{\text{row operations}} U$$

$$\mathbf{I}_n \xrightarrow{\text{same row operations}} L' \xrightarrow{\text{flip off-diagonal elements}} L$$

$$L\vec{y} = \vec{b} \Rightarrow U\vec{x} = \vec{y}$$

forward sub back sub

Computational cost

$$\text{Decomposing: } \frac{2n^3}{3} + O(n^2)$$

$$\text{Forward and backward sub: } n^2 + O(n)$$

Determinants

1. $\det(BC) = \det(B)\det(C)$
2. **U** upper triangular $\Rightarrow \det U = \prod_{i=1}^n u_{ii}$
3. **L** lower triangular $\Rightarrow \det U = \prod_{i=1}^n u_{ii}$
4. **P** Permutation matrix $\Rightarrow \det P_{\text{even}} = +1, \det P_{\text{odd}} = -1$
5. If $\det \mathbf{A} \neq 0$ then $\mathbf{A}x = b$ has **unique solution**
6. If $\det \mathbf{A} = 0$ then $\mathbf{A}x = b$ has 0 or infinite solutions.

Condition and Stability

$$\|\mathbf{A}\|_p = \max_{\|\vec{x}\|_p \neq 0} \frac{\|\mathbf{A}\vec{x}\|_p}{\|\vec{x}\|_p}$$

$$\|\mathbf{A}\|_1 = \text{maximum abs column sum}$$

$$\|\mathbf{A}\|_\infty = \text{maximum abs row sum}$$

$$\|A\|_p = 0 \iff \|A\| = 0, \quad \|cA\|_p = |c| \|A\|_p,$$

$$\|A+B\|_p \leq \|A\|_p + \|B\|_p$$

$$\kappa_p(A) = \|A\|_p \|A^{-1}\|_p \quad (\text{Condition number})$$

$$\kappa_2(A) = \sqrt{\frac{\lambda_{\max}(A^T A)}{\lambda_{\min}(A^T A)}}$$

$$\kappa(B) = \infty \text{ if } B \text{ is singular}$$

$$\frac{1}{\kappa(B)} \text{ how close } B \text{ is to a singular matrix}$$

$$\vec{r} = \vec{b} - A\vec{u}$$

(residual)

Root finding

Bisection Method make intervals $[a_k, b_k]$, $f(a_0)f(b_0) \leq 0$
(Opposite signs)

$$a_k = \begin{cases} a_{k-1} & f\left(\frac{a_{k-1}+b_{k-1}}{2}\right) \cdot f(a_{k-1}) \leq 0 \\ \frac{a_{k-1}+b_{k-1}}{2} & \text{otherwise} \end{cases}$$

$$b_k = \begin{cases} b_{k-1} & f\left(\frac{a_{k-1}+b_{k-1}}{2}\right) \cdot f(a_{k-1}) > 0 \\ \frac{a_{k-1}+b_{k-1}}{2} & \text{otherwise} \end{cases}$$

$$\# \text{ of steps for tolerance } t \geq \frac{|b_0 - a_0|}{2^N}$$

Newton's Method

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

Secant Method

$$x_{k+1} = x_k - f(x_k) \left[\frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} \right]$$

Rate of Convergence

Error: $e_k = x_k - x^*$

Order of convergence: x_k converges with order q if and only if $x_i \rightarrow x^*$

$$\lim c_i = N \in [0, \infty)$$

$$|e_{i+1}| = c_i |e_i|^q \quad \begin{cases} \text{Bisection Method} & q = 1 \\ \text{Secant Method} & q = \frac{1+\sqrt{5}}{2} \\ \text{Newton's Method} & q = 2 \end{cases}$$

Fourier Series

$$\omega = 2\pi f$$

$$g_n(x) = \frac{a_0}{2} + \sum_{k=0}^n \left[a_k \cos\left(k \frac{2\pi}{b-a}\right) + b_k \sin\left(k \frac{2\pi}{b-a}\right) \right]$$

$$a_k = \frac{2}{b-a} \int_a^b f(x) \cos\left(k \frac{2\pi x}{b-a}\right)$$

$$b_k = \frac{2}{b-a} \int_a^b f(x) \sin\left(k \frac{2\pi x}{b-a}\right)$$

$$c_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \exp(ikt) dt \quad [a, b] = [-\pi, \pi]$$

$$c_k = \frac{1}{2} (a_k + ib_k)$$

$$f(x) = f(-x) \text{ Even}$$

$$f(-x) = -f(x) \text{ Odd}$$

$$f(t) \text{ even} \Rightarrow b_k = 0 \quad \forall k \quad (\cos(x) = \cos(-x))$$

$$f(t) \text{ odd} \Rightarrow a_k = 0 \quad \forall k \quad (\sin(-x) = -\sin(x))$$

$$V = \left\{ f(x) : \sqrt{\int_a^b |f(x)|^2 dx} < \infty \right\}$$

If $f(x) \in V$, Fourier series of $g_n(x) \rightarrow f(x)$ on $[a, b]$

$$\bar{c}_k = c_{-k}$$

$$a_{-k} = a_k \quad b_{-k} = -b_k$$

$$a_k = 2 \operatorname{Re}(c_k), \quad b_k = -2 \operatorname{Im}(c_k), \quad b_0 = 0, \quad c_0 = \frac{1}{2} a_0$$

Roots of unity

$$W_N = \exp\left(\frac{2\pi i}{N}\right) \quad n\text{-th root of unity}$$

$$W_N^k = \exp\left(\frac{2k\pi i}{N}\right)$$

$$(W_N^k)^N = 1$$

$$(W_N)^{-k} = W_N^{N-k}$$

DFT

$f[i] = f(i)$ for $0 \leq i \leq N-1$ (Sampled for N points)

$$F[k] = DFT\{f[n]\} = \frac{1}{N} \sum_{n=0}^{N-1} f[n] W_N^{-kn}$$

$$f[n] = IDFT\{F[k]\} = \sum_{k=0}^{N-1} F[k] W_N^{kn}$$

$$\vec{F} = W \vec{f}$$

$$W_{ij} = \frac{1}{N} W_N^{-ij} \text{ for } 0 \leq i, j \leq N-1$$

$$W_{ij}^{-1} = W_N^{ij} \text{ for } 0 \leq i, j \leq N-1 \quad W_{ij}^{-1} = N(\overline{W_{ij}})$$

Properties of F

$$F[k] = F[k + sN] \text{ for } s \in \mathbb{Z}$$

$$\overline{F[k]} = F[-k]$$

$\text{Re}\{F[k]\}$ is even in k

$\text{Im}\{F[k]\}$ is odd in k

$$f[n] = f[N-n] = f[-n] \text{ } f \text{ is even in } n$$

$$f[n] = -f[N-n] = -f[-n] \text{ } f \text{ is odd in } n$$

$$f[n] \text{ is even in } n \Rightarrow \text{Im}(F[k]) = 0 \text{ (DFT is real)}$$

$$f[n] \text{ is odd in } n \Rightarrow \text{Re}(F[k]) = 0 \text{ (DFT is purely imaginary)}$$

Aliasing and the Sample Theorem

$$f_s = \frac{N}{T} \text{ Sampling Rate}$$

Sampling Theorem: If a function $f(t)$ is bandwidth limited to frequencies smaller than f_c (max frequency $\leq f_c$) and $f(t)$ is sampled at a rate $\boxed{f_s \geq 2f_c}$ then the function is completely determined by its samples $f[n]$.

Sampling Theorem(ii): For a fixed sample $f[n]$ with a fixed sampling rate f_s . Then if the maximum frequency of a signal $f_c \leq \frac{f_s}{2}$ that can be deconstructed from $f[n]$ such that is free of aliasing errors (DFT if free of aliasing errors).

Fast Fourier Transform

$$\text{flops } F[k] = 2N \Rightarrow \text{flops } \vec{F} = 2N^2 \text{ (Using sums)}$$

$$E[k] \quad \text{(DFT of even indexed part)}$$

$$O[k] \quad \text{(DFT of odd indexed part)}$$

$$F[k] = \frac{1}{2} \left(E[k] + W_N^{-k} O[k] \right) \text{ for } k = 0, \dots, \frac{N}{2} - 1$$

$$F[k + \frac{N}{2}] = \frac{1}{2} \left(E[k] - W_N^{-k} O[k] \right) \text{ for } k = 0, \dots, \frac{N}{2} - 1$$

$$\text{flops } \vec{F} = \frac{5}{2} N \log_2(N)$$

Power Spectrum and Parseval's Theorem

Power spectrum of $f(t)$ or $f[n]$ is $|F[k]|^2$

Parseval's Theorem (Continuous)

$$\frac{1}{b-a} \int_a^b f(t)^2 dt = \sum_{k=-\infty}^{\infty} |c_k|^2$$

Parseval's Theorem (Discrete)

$$\frac{1}{N} \sum_{k=0}^{N-1} |f[n]|^2 = \sum_{k=0}^{N-1} |F[k]|^2$$

Interpolation

For a basis $B = \{\phi_j(x)\}_{j=0}^m$ interpolating points $\{(x_i, f_i)\}_{i=0}^n$

$$y(x) = \sum_{i=0}^m a_i \phi_i(x)$$

$$\Phi = \begin{pmatrix} \phi_0(x_0) & \phi_1(x_0) & \dots & \phi_m(x_0) \\ \phi_0(x_1) & \phi_1(x_1) & \dots & \phi_m(x_1) \\ \vdots & \vdots & \dots & \vdots \\ \phi_0(x_n) & \phi_1(x_n) & \dots & \phi_m(x_n) \end{pmatrix}$$

$$\Phi \vec{a} = \vec{f} \quad \text{(Solves for } a_i \text{)}$$

Vandermonde Polynomial:

$$y(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n \text{ for } (n+1) \text{ points } (x_i, f_i).$$

$$\Phi = V = \begin{pmatrix} 1 & x_0 & x_0^2 & \dots & x_0^n \\ 1 & x_1 & x_1^2 & \dots & x_1^n \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^n \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} f_0 \\ f_1 \\ \vdots \\ f_n \end{pmatrix}$$

$$\det(V) = \prod_{0 \leq i < j \leq n} (x_j - x_i)$$

Computing V takes $\Theta(N^3)$ flops

Computing $y(x)$ takes $3N$ flops

Lagrange Interpolation The Lagrange basis functions are defined as

$$\ell_i(x_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

$$\ell_i(x) = \prod_{j=0, j \neq i}^n \frac{x - x_j}{x_i - x_j} \quad \text{(basis)}$$

$$\Phi = L = I_{n \times n}$$

Computing I takes 0 flops

Computing pre-computed terms takes $2N^2 + 2N$ flops

Computing $y(x)$ takes $5N$ flops

Newton Interpolation

$$\pi_j(x) = \prod_{i=0}^{j-1} (x - x_i) \quad \pi_0(x) = 1 \quad \text{(basis)}$$

$$\pi_j(x_i) = 0 \text{ if } i < j$$

$$\Pi = \begin{pmatrix} \pi_0(x_0) & \dots & \mathbf{0} \\ \vdots & \ddots & \mathbf{0} \\ \pi_n(x_n) & \dots & \pi_n(x_n) \end{pmatrix} \quad (\Pi \text{ is lower triangular})$$

Computing Π takes $\Theta(N^2)$ flops

Computing pre-computed terms takes $\Theta(N^2)$ flops

Computing $y(x)$ takes $\Theta(N)$ flops

Extending newton polynomial with additional point (x_{n+1}, f_{n+1})

$$y_{n+1}(x) = y_n(x) + a_{n+1} \pi_{n+1}(x) \quad a_{n+1} = \frac{f_{n+1} - y_n(x_{n+1})}{\pi_{n+1}(x)}$$

Hermite Interpolation Given $\{(x_i, f_i, f'_i)\}_{i=0}^n$ interpolating polynomial has degree $\boxed{2n+1}$.

Chebyshev Points on interval $[a, b]$ the $n+1$ points are

$$x_j = \frac{a+b}{2} + \frac{b-a}{2} \cos\left(\frac{2j+1}{2n+2}\pi\right) \text{ for } j = 0, \dots, n$$

Cubic Splines Smoothness conditions: $y'_j(x_j) = y'_{j+1}(x_j)$

and $y''_j(x_j) = y''_{j+1}(x_j)$ for cubic splines

free boundary: $y''_1(x_0) = 0, y''_n(x_n) = 0$, clamped boundary:

$y'_1(x_0) = f'_0, y'_n(x_n) = f'_n$, periodic boundary if $f_0 = f_n$:

$y'_1(x_0) = y'_n(x_n)$ and $y''_1(x_0) = y''_n(x_n)$. Cubic splines can be solved in $\Theta(N)$ flops

Regression

Φ with fewer basis function $m+1 < n+1$. the system is over-determined we can find solution \vec{a} such that the residue $\min \|\Phi \vec{a} - \vec{f}\|_2^2$ is minimized.

$$\underbrace{\Phi}_{(n+1) \times (m+1)} \underbrace{\vec{a}}_{(m+1) \times 1} + \vec{r} = \vec{f}$$

$$\boxed{\Phi^T \Phi \vec{a} = \Phi^T \vec{f}} \quad \text{(normal equation)}$$

$$\begin{pmatrix} a \\ b \end{pmatrix} = \frac{1}{(n+1) \sum x_i^2 - (\sum x_i)^2} \begin{pmatrix} (\sum x_i^2)(\sum f_i) - (\sum x_i)(\sum x_i f_i) \\ (\sum x_i f_i) - (\sum x_i)(\sum f_i) \end{pmatrix}$$

Numerical Integration

$$I_M = \sum_{i=1}^n (x_i - x_{i-1}) f\left(\frac{x_i - x_{i-1}}{2}\right) \quad \text{(Midpoint rule)}$$

$$I_T = \sum_{i=1}^n \frac{(x_i - x_{i-1})}{2} (f(x_{i-1}) + f(x_i)) \quad \text{(Trapezoid Rule)}$$

$$I_S = \sum_{i=1}^n \frac{x_i - x_{i-1}}{6} \left[f(x_{i-1}) + 4f(x_{i-\frac{1}{2}}) + f(x_i) \right]$$

$$\text{Where } x_{i-\frac{1}{2}} = \frac{x_{i-1} + x_i}{2} \quad \text{(Simpsons Rule)}$$

Error Bounds

$$|E_M| \leq \frac{(b-a)^3}{24n^2} \max_{x \in [a,b]} |f''(x)| \quad \text{(Midpoint rule error bound)}$$

$$|E_T| \leq \frac{(b-a)^3}{12n^2} \max_{x \in [a,b]} |f''(x)| \quad \text{(Trapezoid rule error bound)}$$

$$|E_S| \leq \frac{(b-a)^5}{2880n^4} \max_{x \in [a,b]} |f^{(4)}(x)| \quad \text{(Simpsons rule error bound)}$$

Degree of precision: \hat{I} has degree of precision m if

$E = I - \hat{I} = 0$ for polynomial $p(x)$ such that $\deg p \leq m$. Or \hat{I} integrates any polynomial p with $\deg p \leq m$ exactly. Midpoint and Trapezium rule: $\deg_p = 1$, Simpsons rule: $\deg_p = 3$.