Numerisk analys för M2 – Formelsammanfattning

Noggrannhetsanalys och datoraritmetik.

Låt $\tilde{x}$ vara en approximation av x.

$\boxed{\text{Absolut fel i } \tilde{x} \text{ är } e_x = \tilde{x} - x}$ $\boxed{\tilde{x} \text{ har } t \text{ korrekta decimaler om } |e_x| \leq \frac{1}{2} \cdot 10^{-t}}$

$\boxed{|e_x| \leq E_x \quad \Leftrightarrow \quad \tilde{x} - E_x \leq x \leq \tilde{x} + E_x \quad \Leftrightarrow \quad x = \tilde{x} \pm E_x}$

$\boxed{\text{Relativt fel i } \tilde{x} \text{ är } r_x = \frac{e_x}{x} = \frac{\tilde{x}-x}{x} \text{ om } x \neq 0}$ Det gäller att $\boxed{\tilde{x} = x(1+r_x)}$

Felfortplantning; $\tilde{x}$ approximation av x $\boxed{e_f = f(\tilde{x}) - f(x) \approx f'(\tilde{x})e_x}$

$\tilde{x} = (\tilde{x}_1, \tilde{x}_2, \cdots, \tilde{x}_n)$ approximation av $x = (x_1, x_2, \cdots, x_n)$

$$\boxed{e_f = f(\tilde{x}) - f(x) \approx \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\tilde{x}) e_{x_i}, \qquad e_{x_i} = \tilde{x}_i - x_i}$$

Låt $fl(x)$ vara flyttalsrepresentationen av x och $\mu = 2^{-(t+1)}$ avrundningsenheten. Vid enkel precision är $t = 23$ och vid dubbel är $t = 52$.

Relativt lagringsfel i $fl(x)$; $\frac{|x - fl(x)|}{|x|} \leq \mu$ om $x \neq 0$ $\boxed{fl(x) = x(1+r),\ |r| \leq \mu}$

Linjär algebra.

LU-faktorisering; A $n \times n$-matris. $PA = LU$, där P är en permutationsmatris, L nedåt triangulär, U uppåt triangulär. Beräkningskostnad $\sim \frac{n^3}{3}$.

$Ax = b$ löses genom (1) Faktorisera $PA = LU$ (2) Lös $Ly = Pb$ (3) Lös $Ux = y$ Beräkningskostnad $\sim \frac{n^2}{2}$ vardera för (2) och (3).

Choleski-faktorisering; $A = C^T C$ om A symmetrisk ($A^T = A$) och positivt definit ($x^T Ax > 0$ för alla $x \neq 0$). Beräkningskostnad $\sim \frac{n^3}{6}$.

A bandmatris, bandbredd $d = p + q + 1$. Faktorisering utan radbyten, beräkningskostnad $\sim nqp$. Med radbyten, beräkningskostnad $\sim nq(p+q)$. Lösningen av triangulära systemen, beräkningskostnad $\sim nd$.

Några vektornormer

$$\boxed{\| x \|_1 = \sum_{i=1}^{n} |x_i|, \quad \| x \|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2}, \quad \| x \|_\infty = \max_{1 \leq i \leq n} |x_i|}$$

och motsvarande matrisnormer

$$\boxed{\| A \|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^{n} |a_{ij}|, \quad \| A \|_2 = \sqrt{\lambda_{\max}(A^T A)}, \quad \| A \|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^{n} |a_{ij}|}$$

som alla är subordinerade, dvs. $\boxed{\| Ax \|_p \leq \| A \|_p \cdot \| x \|_p,\ p = 1, 2, \infty}$

Singulära värden till en matris.

$$s_i(A) = \sqrt{\lambda_i(A^T A)}$$

där $\lambda(A^T A)$ är egenvärdena till matrisen $A^T A$.

Konditionstalet till en matris A.

$$\kappa(A) = \| A \| \cdot \| A^{-1} \|; \quad \kappa(A) = \frac{\max s_i(A)}{\min s_i(A)}; \quad \kappa(A) = \frac{\max \|Ax\|}{\min \|Ax\|},$$

där de sista max, min beräknas över alla normerade vektorer $\|x\| = 1$.

Vill lösa $Ax = b$. Har approximation $\bar{b}$ av b och löser $A\bar{x} = \bar{b}$. Då gäller

$$\frac{\| \bar{x} - x \|}{\| x \|} \leq \kappa(A) \frac{\| \bar{b} - b \|}{\| b \|}, \qquad \kappa(A) - \text{ konditionstalet till } A$$

Minstakvadratlösningen till det överbestämda systemet $Ax = b$, x n-vektor, b m-vektor, är den vektorn $\tilde{x}$ som minimerar normen $\|Ax - b\|_2$. Den uppfyller den s.k. normala ekvationen $A^T A\tilde{x} = A^T b$

Potensmetoden för egenvärdesproblemet $Ax = \lambda x$

Metoden ger det största i absolutbeloppet egenvärdet λ_{max} till A och en egenvektor x som svarar mot λ_{max}.
Låt u_0 vara en slumpvektor, upprepa för $k = 0, 1, \cdots$,

$$(1)\ \ w_k = u_k / \| u_k \|_2 \quad (2)\ \ \mathrm{u}_{k+1} = Aw_k \quad (3)\ \ \sigma_k = w_k^T u_{k+1}$$

Nu kommer $\sigma_k \to \lambda_{max}$, då $k \to \infty$, och $u_k \to x_j$, då $k \to \infty$.

Inversiterationsmetoden för egenvärdesproblemet $Ax = \lambda x$

(Det är faktiskt potensmetoden för matrisen $(A - \mu I)^{-1}$)

Bestäm skift μ nära sökta egenvärdet λ_j, låt u_0 vara en slumpvektor, upprepa för $k = 0, 1, \cdots$,

$$(1)\ \ w_k = u_k / \| u_k \|_2 \quad (2)\ \text{Lös } (A - \mu I)u_{k+1} = w_k \quad (3)\ \ \sigma_k = w_k^T u_{k+1}$$

Nu kommer $\sigma_k \to (\lambda_j - \mu)^{-1}$, då $k \to \infty$, dvs. λ_j approximeras av $\mu + 1/\sigma_k$, och $u_k \to x_j$, då $k \to \infty$.

Ickelinjära ekvationer.

Låt $\{x^{(k)}\}_{k=0}^{\infty}$ vara en talföljd sådan att $x^{(k)} \to x^*$, då $k \to \infty$.

Konvergensordningen är största talet q, så att $\lim_{k\to\infty} \frac{|x^{(k+1)} - x^*|}{|x^{(k)} - x^*|^q} = C < \infty$.
Konvergensen är linjär om $q = 1$, superlinjär om $1 < q < 2$, kvadratisk om $q = 2$. C är asymptotiska felkonstanten.

Newtons metod

$$x^{(k+1)} = x^{(k)} - \frac{f(x^{(k)})}{f'(x^{(k)})}, k = 0, 1, \cdots$$

Sekantmetoden

$$x^{(k+1)} = x^{(k)} - f(x^{(k)}) \frac{x^{(k)} - x^{(k-1)}}{f(x^{(k)}) - f(x^{(k-1)})}, k = 1, 2, \cdots$$

Låt x^* vara roten till $f(x) = 0$ och $\bar{x}$ en approximativ rot.

$$\delta x = \bar{x} - x^* \approx \frac{f(\bar{x})}{f'(x^*)} \approx \frac{f(\bar{x})}{f'(\bar{x})}$$

Parameterkänslighet;

$$\left. \begin{array}{l} f(x, \alpha) = 0 \\ f(x + \delta x, \alpha + \delta\alpha) = 0 \end{array} \right\} \Rightarrow \quad \delta x \approx -\frac{f'_\alpha(x, \alpha)}{f'_x(x, \alpha)} \delta\alpha$$

System $F(x) = 0$ med Jacobianmatris (funktionalmatris) $J(x) = \left[\frac{\partial f_i}{\partial x_j}(x)\right]_{ij}$.

Newtons metod

$$x^{(k+1)} = x^{(k)} + d^{(k)}, k = 0, 1, \cdots, \quad \text{där} \quad J(x^{(k)}) d^{(k)} = -F(x^{(k)})$$

Inbäddning. $H(x, \lambda) = 0, 0 \leq \lambda \leq 1$, så att $H(x, 0) = E(x)$ enkel att lösa och $H(x, 1) = F(x)$.

Fixpunktsiteration;

$$x^{(k+1)} = G(x^{(k)}), k = 0, 1, \cdots$$

x^* är en fixpunkt till funktionen $G(x)$ om $x^* = G(x^*)$.
Fixpunktsiterationerna $x^{(k)} \to x^*$ om $\| G(u) - G(v) \| \leq \mu \| u - v \|$ med $\mu < 1$.

Differensapproximation av derivator.

Framåtdifferens och bakåtdifferens:

$$D_+ y(x) = \frac{y(x + h) - y(x)}{h} = y'(x) + c_1 h + c_2 h^2 + \cdots$$

$$D_- y(x) = \frac{y(x) - y(x - h)}{h} = y'(x) + c_1 h + c_2 h^2 + \cdots$$

Centraldifferens:

$$\boxed{D_0 y(x) = \frac{y(x+h) - y(x-h)}{2h} = y'(x) + c_2 h^2 + c_4 h^4 \cdots}$$

Ensidiga differenser:

$$\boxed{\begin{aligned} \frac{-y(x+2h) + 4y(x+h) - 3y(x)}{2h} &= y'(x) + \mathcal{O}(h^2) \\ \frac{3y(x) - 4y(x-h) + y(x-2h)}{2h} &= y'(x) + \mathcal{O}(h^2) \end{aligned}}$$

Differensapproximationer för andra och fjärde derivator:

$$\boxed{D_+ D_- y(x) = \frac{y(x+h) - 2y(x) + y(x-h)}{h^2} = y''(x) + c_2 h^2 + c_4 h^4 \cdots}$$

$$\boxed{(D_+ D_-)^2 y(x) = \frac{y(x+2h) - 4y(x+h) + 6y(x) - 4y(x-h) + y(x-2h)}{h^4} = y^{(4)}(x) + \mathcal{O}(h^2)}$$

Richardsonextrapolation.

Om en h - beroende approximation $F(h)$ för en storhet a har felet på formen $a_1 h^p + O(h^r)$ med $r > p$:

$$F(h) = a + a_1 h^p + O(h^r).$$

då kan a approximeras med större noggranhet med hjälp av Richardsonextrapolation:

$\boxed{a_{extr} = F(h) + \frac{F(h) - F(qh)}{q^p - 1}}$; I fallet när $p = 2$, $q = 2$ $\boxed{a_{extr} = F(h) + \frac{F(h) - F(2h)}{3}}$

Begynnelsevärdesproblem.

$$\begin{cases} y' = f(t, y), a \leq t \leq b \\ y(a) = c. \end{cases}$$

Euler framåtmetoden $\boxed{y_{n+1} = y_n + hf(t_n, y_n)}$

Euler bakåtmetoden $\boxed{y_{n+1} = y_n + hf(t_{n+1}, y_{n+1})}$

Mittpunktsmetoden $\boxed{y_{n+1} = y_{n-1} + 2hf(t_n, y_n)}$

Trapetsmetoden $\boxed{y_{n+1} = y_n + \frac{h}{2}\{f(t_n, y_n) + f(t_{n+1}, y_{n+1})\}}$

Störningsresultat;

Om $\| f(t,u) - f(t,v) \| \leq L \| u - v \|$ och

$$\begin{cases} y' = f(t,y), a \leq t \leq b \\ y(a) = c_1 \end{cases} \qquad \begin{cases} z' = f(t,z), a \leq t \leq b \\ z(a) = c_2 \end{cases}$$

så gäller $\| y(t) - z(t) \| \leq e^{L(t-\hat{a})} \| y(\hat{a}) - z(\hat{a}) \|, \quad a \leq \hat{a} \leq t \leq b.$

Globala felet i $t = t_n$ är $\boxed{g_n = y(t_n) - y_n}$

Lokala felet i $t = t_{n+1}$ är $\boxed{l_{n+1} = u_n(t_{n+1}) - y_{n+1}}$ där $u_n(t)$ lösningen till

$$\begin{cases} u_n' = f(t,u_n), t_n \leq t \leq t_{n+1} \\ u_n(t_n) = y_n. \end{cases}$$

En metod har noggrannhetsordning p om det lokala felet är $\mathcal{O}(h^{p+1})$.

Euler framåt- och bakåtmetoden $p = 1$. Mittpunkts- och Trapetsmetoden $p = 2$.

Heuns metod ($p = 2$) $\boxed{\begin{array}{l} y_{n+1} = y_n + \frac{h}{2}\{k_1 + k_2\}, \\ k_1 = f(t_n, y_n), \quad k_2 = f(t_n + h, y_n + hk_1) \end{array}}$

RK3 ($p = 3$) $\boxed{\begin{array}{l} y_{n+1} = y_n + \frac{h}{6}\{k_1 + k_2 + 4k_3\}, \\ k_1 = f(t_n, y_n), \quad k_2 = f(t_n + h, y_n + hk_1), \\ k_3 = f(t_n + \frac{h}{2}, y_n + \frac{h}{4}\{k_1 + k_2\}) \end{array}}$

RK4 ($p = 4$) $\boxed{\begin{array}{l} y_{n+1} = y_n + \frac{h}{6}\{k_1 + 2k_2 + 2k_3 + k_4\}, \\ k_1 = f(t_n, y_n), \qquad k_2 = f(t_n + \frac{h}{2}, y_n + \frac{h}{2}k_1), \\ k_3 = f(t_n + \frac{h}{2}, y_n + \frac{h}{2}k_2), \quad k_4 = f(t_n + h, y_n + hk_3) \end{array}}$

En metod är absolut stabil om den för givna h och λ ger begränsade approximationer av lösningen till testekvationen; $\boxed{y' = \lambda y, t \geq 0}$

Metodens absoluta stabilitetsområde är mängden av alla $h\lambda$ i komplexa planet för vilka metoden är absolut stabil.

Metoden är $\boxed{\text{A-stabil}}$ om stabilitetsområdet innehåller vänstra halvplanet.