Département de mathématiques et de génie industriel
École Polytechnique de Montréal

MTH2210A-CALCUL SCIENTIFIQUE POUR INGÉNIEURS
AIDE-MÉMOIRE

Définitions, développement de Taylor et propagation d'erreurs

- Erreur absolue: $\Delta x = |x - x^\star|$
- Erreur relative: $e_r(x) = \frac{\Delta x}{|x|}$
- Chiffres significatifs:

  Le chiffre de $x^\star$ associé à la puissance de m et les chiffres associés aux puissances supérieures tels que $\Delta x \leq 0,5 \times 10^m$.

- $f(x_0 + h) = P_n(h) + R_n(h)$:

$$\begin{cases} P_n(h) = f(x_0) + f'(x_0)h + \frac{1}{2!}f''(x_0)h^2 + \frac{1}{3!}f'''(x_0)h^3 + \cdots + \frac{1}{n!}f^{(n)}(x_0)h^n \\ R_n(h) = \frac{1}{(n+1)!}f^{(n+1)}(\xi(h))h^{(n+1)} \quad \text{pour } \xi(h) \text{ entre } x_0 \text{ et } x_0 + h. \end{cases}$$

- $f(x) = P_n(x) + R_n(x)$:

$$\begin{cases} P_n(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2!}f''(x_0)(x - x_0)^2 + \cdots + \frac{1}{n!}f^{(n)}(x_0)(x - x_0)^n \\ R_n(x) = \frac{1}{(n+1)!}f^{(n+1)}(\xi(x))(x - x_0)^{(n+1)} \quad \text{pour } \xi(x) \text{ entre } x_0 \text{ et } x. \end{cases}$$

- $f(h) = \mathcal{O}(h^n)$:

  Il existe une constante $C > 0$ t.q. $\left|\frac{f(h)}{h^n}\right| \leq C$ pour h près de 0.

- $\Delta f \simeq \left|\frac{\partial f(x^*, y^*)}{\partial x}\right| \Delta x + \left|\frac{\partial f(x^*, y^*)}{\partial y}\right| \Delta y$

Systèmes d'équations algébriques

- La factorisation matricielle de Crout:

$$A = LU = \begin{pmatrix} \ell_{11} & 0 & 0 & \cdots & 0 \\ \ell_{21} & \ell_{22} & 0 & \ddots & \vdots \\ \ell_{31} & l_{32} & \ell_{33} & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \ell_{n1} & \ell_{n2} & \cdots & \ell_{nn-1} & \ell_{nn} \end{pmatrix} \begin{pmatrix} 1 & u_{12} & u_{13} & \cdots & u_{1n} \\ 0 & 1 & u_{23} & \cdots & u_{2n} \\ 0 & 0 & 1 & \cdots & u_{3n} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

- La méthode de Thomas pour la factorisation des matrices tridiagonales:

$$A = \begin{pmatrix} a_1 & c_1 & 0 & \cdots & 0 \\ b_1 & a_2 & c_2 & \ddots & \vdots \\ 0 & b_2 & a_3 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & c_{n-1} \\ 0 & \cdots & 0 & b_{n-1} & a_n \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ \beta_1 & 1 & 0 & \ddots & \vdots \\ 0 & \beta_2 & 1 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \beta_{n-1} & 1 \end{pmatrix} \begin{pmatrix} \alpha_1 & c_1 & 0 & \cdots & 0 \\ 0 & \alpha_2 & c_2 & \ddots & \vdots \\ 0 & 0 & \alpha_3 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & c_{n-1} \\ 0 & 0 & \cdots & 0 & \alpha_n \end{pmatrix}$$

- La factorisation matricielle de Cholesky pour les matrices symétriques définies positives:

$$A = LL^T = \begin{pmatrix} \ell_{11} & 0 & 0 & \cdots & 0 \\ \ell_{21} & \ell_{22} & 0 & \ddots & \vdots \\ \ell_{31} & \ell_{32} & \ell_{33} & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \ell_{n1} & \ell_{n2} & \cdots & \ell_{nn-1} & \ell_{nn} \end{pmatrix} \begin{pmatrix} \ell_{11} & \ell_{21} & \ell_{31} & \cdots & \ell_{n1} \\ 0 & \ell_{22} & \ell_{32} & \cdots & \ell_{n2} \\ 0 & 0 & \ell_{33} & \cdots & \ell_{n3} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \ell_{nn} \end{pmatrix}$$

  Les énoncés suivants sont équivalents:

  - A est une matrice symétrique et définie positive;
  - Toutes les valeurs propres de A sont réelles et strictement positives;
  - Le déterminant des sous-matrices principales de A est strictement positif.

- La résolution des systèmes linéaires:

$$A\vec{x} = \vec{b} \Leftrightarrow LU\vec{x} = P\vec{b} \Rightarrow \begin{cases} 1^\circ)\ L\vec{y} = P\vec{b}; \\ 2^\circ)\ U\vec{x} = \vec{y}. \end{cases}$$

  Note: On peut utiliser le vecteur de permutation $\vec{O}$ plutôt que la matrice de permutation P.

- Normes vectorielles:

$$\|\vec{x}\|_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} \qquad \|\vec{x}\|_1 = \sum_{i=1}^{n} |x_i| \qquad \|\vec{x}\|_\infty = \max_{i=1,2,\ldots,n} |x_i|$$

- Normes matricielles:

$$\|A\|_1 = \max_{j=1,2,\ldots,n} \sum_{i=1}^{n} |a_{ij}| \qquad \|A\|_\infty = \max_{i=1,2,\ldots,n} \sum_{j=1}^{n} |a_{ij}| \qquad \|A\|_{\mathrm{F}} = \sqrt{\sum_{i,j=1}^{n} a_{ij}^2}$$

- Conditionnement matriciel: $\operatorname{cond} A = \|A\| \|A^{-1}\|$
- Bornes de l'erreur: pour le résidu $\vec{r} = \vec{b} - A\vec{x}^\star$ et la perturbation E sur la matrice A,

$$\frac{1}{\operatorname{cond} A} \frac{\|\vec{r}\|}{\|\vec{b}\|} \le \frac{\|\vec{x} - \vec{x}^\star\|}{\|\vec{x}\|} \le \operatorname{cond} A \frac{\|\vec{r}\|}{\|\vec{b}\|} \qquad \text{et} \qquad \frac{\|\vec{x} - \vec{x}^\star\|}{\|\vec{x}^\star\|} \le \operatorname{cond} A \frac{\|E\|}{\|A\|}$$

Interpolation

- Interpolation polynômiale de Lagrange: étant donné $(n+1)$ points $(x_i, f(x_i))$ pour $i = 0, 1, \cdots, n$:

  $p_n(x) = \sum_{i=0}^{n} f(x_i) L_i(x)$,

$$\text{où} \quad L_i(x) = \frac{(x - x_0)(x - x_1)\cdots(x - x_{i-1})(x - x_{i+1})\cdots(x - x_n)}{(x_i - x_0)(x_i - x_1)\cdots(x_i - x_{i-1})(x_i - x_{i+1})\cdots(x_i - x_n)}$$

- Différences divisées: $f[x_i] = f(x_i)$,

$$f[x_i, x_{i+1}] = \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i}, \quad f[x_i, x_{i+1}, x_{i+2}] = \frac{f[x_{i+1}, x_{i+2}] - f[x_i, x_{i+1}]}{x_{i+2} - x_i}, \quad \text{etc.}$$

- Polynôme de Newton:

$$p_n(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1) + \cdots + a_n(x - x_0)(x - x_1)\cdots(x - x_{n-1}),$$

  où $a_i = f[x_0, x_1, x_2, \cdots, x_i]$ pour $i = 0, 1, \ldots, n$

- Erreur d'interpolation:

$$E_n(x) = f(x) - p_n(x) = \frac{f^{(n+1)}(\xi(x))}{(n+1)!}(x-x_0)(x-x_1)\cdots(x-x_n) \quad \text{pour} \quad \xi(x) \in]x_0, x_n[$$

- Interpolation par splines cubiques:

  En posant $h_i = x_i - x_{i-1}$, l'équation de la spline dans l'intervalle $[x_{i-1}, x_i]$ est donnée par:

$$\begin{aligned} p_i(x) = &-\frac{f''_{i-1}}{6h_i}(x-x_i)^3 + \frac{f''_i}{6h_i}(x-x_{i-1})^3 \\ &- \left\{\frac{f(x_{i-1})}{h_i} - \frac{f''_{i-1}h_i}{6}\right\}(x-x_i) + \left\{\frac{f(x_i)}{h_i} - \frac{f''_i h_i}{6}\right\}(x-x_{i-1}), \end{aligned}$$

  pour $i = 1, 2, \ldots, n$.
  Les f''_i, sont solutions d'un système de $n-1$ équations données par:

$$\frac{h_i}{h_i + h_{i+1}} f''_{i-1} + 2f''_i + \frac{h_{i+1}}{h_i + h_{i+1}} f''_{i+1} = 6f[x_{i-1}, x_i, x_{i+1}],$$

  pour $i = 1, 2, \ldots, n-1$.

Différentiation et intégration numériques

- Différentiation numérique:

formule de différence finie	terme d'erreur
$f'(x) \simeq \frac{f(x+h)-f(x)}{h}$	$-\frac{f''(\xi)}{2}h$
$f'(x) \simeq \frac{f(x)-f(x-h)}{h}$	$\frac{f''(\xi)}{2}h$
$f'(x) \simeq \frac{-f(x+2h)+4f(x+h)-3f(x)}{2h}$	$\frac{f'''(\xi)}{3}h^2$
$f'(x) \simeq \frac{f(x+h)-f(x-h)}{2h}$	$-\frac{f'''(\xi)}{6}h^2$
$f'(x) \simeq \frac{3f(x)-4f(x-h)+f(x-2h)}{2h}$	$\frac{f'''(\xi)}{3}h^2$
$f''(x) \simeq \frac{f(x+2h)-2f(x+h)+f(x)}{h^2}$	$-f'''(\xi)h$
$f''(x) \simeq \frac{f(x+h)-2f(x)+f(x-h)}{h^2}$	$-\frac{f''''(\xi)}{12}h^2$
$f''(x) \simeq \frac{f(x-2h)-2f(x-h)+f(x)}{h^2}$	$f'''(\xi)h$
$f''(x) \simeq \frac{-f(x+2h)+16f(x+h)-30f(x)+16f(x-h)-f(x-2h)}{12h^2}$	$\frac{1}{90}f^{(6)}(\xi)h^4$

- Extrapolation de Richardson:

  Soit
$$Q_{\text{exa}} = Q_{\text{app}}(h) + c_n h^n + c_{n+1}h^{n+1} + c_{n+2}h^{n+2} + \cdots$$
  alors pour $p > 1$
$$Q_{\text{exa}} = \frac{p^n Q_{\text{app}}(\frac{h}{p}) - Q_{\text{app}}(h)}{p^n - 1} + \frac{\left(\frac{1}{p} - 1\right)c_{n+1}h^{n+1} + \left(\frac{1}{p^2} - 1\right)c_{n+2}h^{n+2} + \cdots}{p^n - 1}$$

méthode	formule de quadrature	terme d'erreur	h
trapèze composée	$\sum_{i=0}^{n-1} \frac{h}{2}[f(x_i)+f(x_{i+1})] =$ $\frac{h}{2}(f(x_0)+2[f(x_1)+f(x_2)+\cdots+f(x_{n-1})]+f(x_n))$	$-\frac{(b-a)}{12} f''(\eta)h^2$	$\frac{b-a}{n}$
Simpson $\frac{1}{3}$ composée	$\sum_{i=0}^{n-1} \frac{h}{3}[f(x_{2i})+4f(x_{2i+1})+f(x_{2i+2})] =$ $\frac{h}{3}[f(x_0)+4f(x_1)+2f(x_2)+4f(x_3)+2f(x_4)+\cdots$ $\cdots+4f(x_{2n-3})+2f(x_{2n-2})+4f(x_{2n-1})+f(x_{2n})]$	$-\frac{(b-a)}{180} f^{(4)}(\eta)h^4$	$\frac{b-a}{2n}$
Simpson $\frac{3}{8}$ composée	$\sum_{i=0}^{n-1} \frac{3h}{8}[f(x_{3i})+3f(x_{3i+1})+3f(x_{3i+2})+f(x_{3i+3})] =$ $\frac{3h}{8}[f(x_0)+3f(x_1)+3f(x_2)+2f(x_3)+3f(x_4)+\cdots$ $\cdots+2f(x_{3n-3})+3f(x_{3n-2})+3f(x_{3n-1})+f(x_{3n})]$	$-\frac{(b-a)}{80} f^{(4)}(\eta)h^4$	$\frac{b-a}{3n}$
Boole composée	$\sum_{i=0}^{n-1} \frac{2h}{45}[7f(x_{4i})+32f(x_{4i+1})+12f(x_{4i+2})$ $+32f(x_{4i+3})+7f(x_{4i+4})] =$ $\frac{2h}{45}[7f(x_0)+32f(x_1)+12f(x_2)+32f(x_3)+14f(x_4)\cdots+$ $\cdots+32f(x_{4n-5})+14f(x_{4n-4})+32f(x_{4n-3})+12f(x_{4n-2})$ $+32f(x_{4n-1})+7f(x_{4n})]$	$-\frac{2(b-a)}{945} f^{(6)}(\eta)h^6$	$\frac{b-a}{4n}$

- Quadratures de Newton-Cotes:

- Intégration de Gauss:

$$\int_a^b f(x)\,dx = \frac{(b-a)}{2}\int_{-1}^{1} f\left(\frac{(b-a)t+(a+b)}{2}\right)dt = \frac{(b-a)}{2}\int_{-1}^{1} g(t)\,dt \simeq \frac{(b-a)}{2}\sum_{i=1}^{n}\omega_i\, g(t_i)$$

nb de pts (n)	points de Gauss t_i	poids de Gauss ω_i	degré de précision ($2n-1$)
1	0	2	1
2	$-\sqrt{3}/3$ $+\sqrt{3}/3$	1 1	3
3	$-\sqrt{15}/5$ 0 $+\sqrt{15}/5$	$5/9$ $8/9$ $5/9$	5
4	$-0,861\,136\,312$ $-0,339\,981\,044$ $+0,339\,981\,044$ $+0,861\,136\,312$	$0,347\,854\,845$ $0,652\,145\,155$ $0,652\,145\,155$ $0,347\,854\,845$	7

Équations algébriques non linéaires

- Problème « de racine »: chercher r t.q. $f(r) = 0$
- Borne supérieure de l'erreur pour la méthode de la bissection: $|x_n - r| \leq \frac{b-a}{2^n}$
- Problème de point fixe: chercher r t.q. $r = g(r)$
- Algorithme de point fixe: pour x_0 donné, $x_{n+1} = g(x_n)$ pour $n = 0, 1, 2, \ldots$
- Développement pour l'analyse de convergence de la méthode de points fixes:

  $e_{n+1} := x_{n+1} - r = g'(r)e_n + \frac{1}{2}g''(r)e_n^2 + \frac{1}{6}g'''(r)e_n^3 + \cdots$
- Méthode de Newton: pour x_0 donné,

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad \text{pour} \quad n = 0, 1, 2, \ldots$$

- Une racine r de la fonction $f(x)$ est de *multiplicité* m si $f(x) = (x-r)^m h(x)$ pour une fonction $h(x)$ qui vérifie $h(r) \neq 0$ ou encore si:

$$f(r) = f'(r) = f''(r) = \cdots = f^{(m-1)}(r) = 0 \quad \text{et} \quad f^{(m)}(r) \neq 0$$

- Taux de convergence de la méthode de Newton dans le cas d'une racine de *multiplicité* m: $1 - \frac{1}{m}$
- Méthode de la sécante: pour x_0 et x_1 donnés,

$$x_{n+1} = g(x_n, x_{n-1}) = x_n - \frac{f(x_n)(x_n - x_{n-1})}{(f(x_n) - f(x_{n-1}))} \quad \text{pour} \quad n = 1, 2, 3, \ldots$$

- Problème de points fixes pour les systèmes d'équations algébriques:

$$\vec{x} = \vec{G}(\vec{x}) \Longleftrightarrow \begin{cases} x_1 &= g_1(x_1, x_2, \ldots, x_n) \\ x_2 &= g_2(x_1, x_2, \ldots, x_n) \\ &\vdots \\ x_n &= g_n(x_1, x_2, \ldots, x_n) \end{cases}$$

- La matrice *jacobienne* associée à $\vec{G}(\vec{x})$:

$$J(\vec{x}) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(\vec{x}) & \cdots & \frac{\partial g_1}{\partial x_n}(\vec{x}) \\ \vdots & & \vdots \\ \frac{\partial g_n}{\partial x_1}(\vec{x}) & \cdots & \frac{\partial g_n}{\partial x_n}(\vec{x}) \end{pmatrix}$$

- Le rayon spectral d'une matrice A est défini par:

$$\rho(A) = \max_{i=1,2,\ldots,n} |\lambda_i|,$$

  où $|\lambda_i|$ est le module complexe de la valeur propre λ_i de A
- Propriété du rayon spectral pour une matrice A: $\rho(A) \leq \|A\|$
- Critère pour la convergence de l'algorithme des points fixes pour les systèmes:

$$\rho(J(\vec{r})) < 1 \quad \text{où} \quad \vec{r} \quad \text{est tel que} \quad \vec{r} = \vec{G}(\vec{r})$$

  et $J(\vec{r})$ est la matrice jacobienne associée à $\vec{G}(\vec{x})$ évaluée en $\vec{r}$.

- Matrice à diagonale strictement dominante A:

$$|a_{ii}| > \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}| \quad \text{pour} \quad i = 1, 2, \ldots, n$$

- La méthode de Jacobi: soit $A\vec{x} = (D + T_I + T_S)\vec{x} = \vec{b}$ et pour $\vec{x}^0$ donné,

$$\vec{x}^{k+1} = \vec{G}(\vec{x}^k) = T_{\mathrm{J}}\,\vec{x}^k + \vec{c}_{\mathrm{J}} = -D^{-1}(T_I + T_S)\,\vec{x}^k + D^{-1}\vec{b}$$

$$\Longleftrightarrow$$

$$x_i^{k+1} = \frac{1}{a_{ii}}(b_i - \sum_{\substack{j=1 \\ j\neq i}}^{n} a_{ij}x_j^k) \quad \text{pour} \quad \begin{cases} i\ = 1, 2, \ldots, n; \\ k = 0, 1, 2, \ldots \end{cases}$$

- La méthode de Gauss-Seidel: soit $A\vec{x} = (D + T_I + T_S)\vec{x} = \vec{b}$ et pour $\vec{x}^0$ donné,

$$\vec{x}^{k+1} = \vec{G}(\vec{x}^k) = T_{\mathrm{GS}}\,\vec{x}^k + \vec{c}_{\mathrm{GS}} = -(T_I + D)^{-1}T_S\,\vec{x}^k + (T_I + D)^{-1}\vec{b}$$

$$\Longleftrightarrow$$

$$x_i^{k+1} = \frac{1}{a_{ii}}(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{k+1} - \sum_{j=i+1}^{n} a_{ij}x_j^k) \quad \text{pour} \quad \begin{cases} i\ = 1, 2, \ldots, n; \\ k = 0, 1, 2, \ldots \end{cases}$$

- La méthode de Newton: pour $\vec{x}^k$ donné, résoudre

$$J(\vec{x}^k)\vec{\delta}x = -\vec{R}(\vec{x}^k) \Longleftrightarrow \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\vec{x}^k) & \frac{\partial f_1}{\partial x_2}(\vec{x}^k) & \cdots & \frac{\partial f_1}{\partial x_n}(\vec{x}^k) \\ \frac{\partial f_2}{\partial x_1}(\vec{x}^k) & \frac{\partial f_2}{\partial x_2}(\vec{x}^k) & \cdots & \frac{\partial f_2}{\partial x_n}(\vec{x}^k) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(\vec{x}^k) & \frac{\partial f_n}{\partial x_2}(\vec{x}^k) & \cdots & \frac{\partial f_n}{\partial x_n}(\vec{x}^k) \end{pmatrix} \begin{pmatrix} \delta x_1 \\ \delta x_2 \\ \vdots \\ \delta x_n \end{pmatrix} = - \begin{pmatrix} f_1(\vec{x}^k) \\ f_2(\vec{x}^k) \\ \vdots \\ f_n(\vec{x}^k) \end{pmatrix}$$

  puis $\vec{x}^{k+1} = \vec{x}^k + \vec{\delta}x$ pour $k = 0, 1, 2, \ldots$

Équations différentielles

- Pour le problème avec condition initiale $y'(t) = f(t, y(t))$ où $y(t_0) = y_0$:

  pour (t_0, y_0) et h donnés, et pour $n = 0, 1, 2, \ldots$:

  Euler explicite (ordre 1):

$$\begin{cases} y_{n+1} = y_n + hf(t_n, y_n) \\ t_{n+1}\ = t_n + h \end{cases}$$

  Euler modifié (Runge-Kutta d'ordre 2):

$$\begin{cases} \hat{y} = y_n + hf(t_n, y_n) \\ y_{n+1} = y_n + \frac{h}{2}(f(t_n, y_n) + f(t_n + h, \hat{y})) \\ t_{n+1}\ = t_n + h \end{cases}$$

  Point milieu (Runge-Kutta d'ordre 2):

$$\begin{cases} k_1 = hf(t_n, y_n) \\ y_{n+1} = y_n + hf(t_n + \frac{h}{2}, y_n + \frac{k_1}{2}) \\ t_{n+1}\ = t_n + h \end{cases}$$

Runge-Kutta d'ordre 4:

$$\begin{cases} k_1 = hf(t_n, y_n) \\ k_2 = hf(t_n + \frac{h}{2}, y_n + \frac{k_1}{2}) \\ k_3 = hf(t_n + \frac{h}{2}, y_n + \frac{k_2}{2}) \\ k_4 = hf(t_n + h, y_n + k_3) \\ y_{n+1} = y_n + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4) \\ t_{n+1} \ = t_n + h \end{cases}$$

Euler implicite (ordre 1):

$$\begin{cases} y_{n+1} = y_n + hf(t_{n+1}, y_{n+1}) \\ t_{n+1} \ = t_n + h \end{cases}$$

Crank-Nicholson (ordre 2):

$$\begin{cases} y_{n+1} = y_n + \frac{h}{2}(f(t_n, y_n) + f(t_{n+1}, y_{n+1})) \\ t_{n+1} \ = t_n + h \end{cases}$$

- Erreur de troncature locale:

  pour $y_{n+1} = y_n + h\varphi(t_n, y_n, t_{n+1}, y_{n+1})$:

$$\tau_{n+1}(h) = \frac{y(t_{n+1}) - y(t_n)}{h} - \varphi(t_n, y(t_n), t_{n+1}, y(t_{n+1}))$$

Steven Dufour