Calculs numériques approchés

IUT SGM

Y. Morel

https://xymaths.fr/

Exercice 1 : Une fonction simple...

On considère la fonction f définie par

$$f(x) = \frac{(1 + 10^{-2x}) - 1}{10^{-2x}}$$

1. Donner la valeur exacte de $f(x)$, pour tout réel x.
2. Tracer, avec un outil numérique graphique (calculatrice, ordinateur, smartphone, ...), la courbe de f pour $x \geqslant 0$
(et/ou remplir un tableau des valeurs $f(x)$ pour les valeurs entières de x entre 0 et 10.)
3. Commenter le phénomène observé.

Exercice 1 : Une fonction simple. . .

On considère la fonction f définie par

$$f(x) = \frac{(1+10^{-2x})-1}{10^{-2x}}$$

1. Donner la valeur exacte de $f(x)$, pour tout réel x.
2. Tracer, avec un outil numérique graphique (calculatrice, ordinateur, smartphone, . . .), la courbe de f pour $x \geqslant 0$
(et/ou remplir un tableau des valeurs $f(x)$ pour les valeurs entières de x entre 0 et 10.)
3. Commenter le phénomène observé.

Les calculs numériques sont effectués en **virgule flottante**, ou **"floating point operation"**, et donc avec un nombre fini et fixe de décimales.

Les erreurs d'arrondi peuvent remettrent en cause les propriétés des opérations, même les plus élémentaires : associativité, commutativité

Exercice 2 : Une équation simple. . .

1. Résoudre l'équation $P(x) = 0$ avec
$$P(x) = x^2 - 1634x + 2$$
avec 10 chiffres significatifs.
On note x_1 et x_2 ces solutions, avec $|x_1| < |x_2|$
2. Calculer $P(x_1)$ et $P(x_2)$.
3. En utilisant la relation entre les racines $x_1x_2 = \frac{c}{a}$, calculer x_1 avec 10 chiffres significatifs, puis $P(x_1)$.
4. Commenter et expliquer.

Exercice 2 : Une équation simple. . .

1. Résoudre l'équation $P(x) = 0$ avec
$$P(x) = x^2 - 1634x + 2$$
avec 10 chiffres significatifs.
On note x_1 et x_2 ces solutions, avec $|x_1| < |x_2|$
2. Calculer $P(x_1)$ et $P(x_2)$.
3. En utilisant la relation entre les racines $x_1x_2 = \frac{c}{a}$, calculer x_1 avec 10 chiffres significatifs, puis $P(x_1)$.
4. Commenter et expliquer.

Les phénomènes de **compensation** se produisent lors de la soustraction de valeurs proches.

$\hookrightarrow$ il en résulte une perte de chiffres significatifs

$\hookrightarrow$ besoin d'adaptation de l'algorithme numérique par rapport à celui algébrique

Exercice 3 : On considère le système linéaire :

$$\left\{\begin{array}{rcrcl} 12,5x & + & 43,7y & = & 181 \\ 2x & + & 7y & = & 29 \end{array}\right.$$

dont la solution est :

Exercice 3 : On considère le système linéaire :

$$\left\{\begin{array}{rcrcl} 12,5x & + & 43,7y & = & 181 \\ 2x & + & 7y & = & 29 \end{array}\right.$$

dont la solution est :

$$x = -3 \text{ et } y = 5$$

Exercice 3 : On considère maintenant le système linéaire :

$$\left\{\begin{array}{rcrcl} 12,5x & + & 43,7y & = & 181 \\ 2x & + & 7y & = & 29{,}\mathbf{1} \end{array}\right.$$

dont la solution est **?** :

$$x \simeq -3 \quad \text{et} \quad y \simeq 5$$

<u>Exercice 3 :</u> On considère le système linéaire :

$$\begin{cases} 12,5x + 43,7y = 181 \\ 2x + 7y = 29{,}\mathbf{1} \end{cases}$$

dont la solution est **?** :

$$x \simeq -3 \quad \text{et} \quad y \simeq 5$$

En fait, après calculs :

$$x = -46,7 \quad \text{et} \quad y = 17,5$$

Exercice 3 : On considère le système linéaire :

$$\begin{cases} 12,5x & + & 43,7y & = & 181 \\ 2x & + & 7y & = & 29,\mathbf{1} \end{cases}$$

dont la solution est **?** :

$$x \simeq -3 \text{ et } y \simeq 5$$

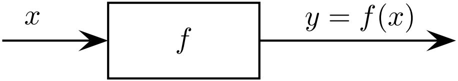
En fait, après calculs :

$$x = -46,7 \text{ et } y = 17,5$$

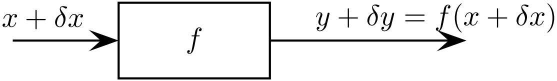
Ainsi, une modification de l'ordre de 1/100ème d'une donnée entraîne une modification d'un rapport de 10 du résultat :

La solution, calculée exactement !, est entachée d'une erreur relative 1000 fois supérieure ! !

On étudie un système qui a une valeur d'entrée x associe une valeur de sortie $y = f(x)$.



Comment se propage sur la valeur de sortie y une erreur, par exemple de mesure, sur x ?
Après perturbation, erreur(s), ...



En supposant la pertubation δx petite devant x, la sortie perturbée s'écrit

$$y + \delta y = f(x + \delta x) \simeq f(x) + \delta x \, f'(x)$$

et donc l'erreur en sortie est

$$\delta y = f\,(x + \delta x) - f(x) \simeq \delta x \, f'(x)$$

L'erreur relative est donc $\quad \dfrac{\delta y}{y} \simeq \delta x \, \dfrac{f'(x)}{f(x)}$

ou encore,

$$\underbrace{\left|\frac{\delta y}{y}\right|}_{\text{Erreur relative en sortie}} \quad \simeq \quad \underbrace{\left|\frac{\delta x}{x}\right|}_{\text{Erreur relative en entrée}} \quad \times \quad \left|x\frac{f'(x)}{f(x)}\right|$$

Le nombre $C = \left|x\dfrac{f'(x)}{f(x)}\right|$ est le conditionnement du problème, et on a donc

$$\left|\frac{\delta y}{y}\right| \simeq C \left|\frac{\delta x}{x}\right|$$

L'erreur relative est donc $\dfrac{\delta y}{y} \simeq \delta x \, \dfrac{f'(x)}{f(x)}$

ou encore,

$$\underbrace{\left|\frac{\delta y}{y}\right|}_{\text{Erreur relative en sortie}} \simeq \underbrace{\left|\frac{\delta x}{x}\right|}_{\text{Erreur relative en entrée}} \times \left|x\frac{f'(x)}{f(x)}\right|$$

Le nombre $C = \left|x\dfrac{f'(x)}{f(x)}\right|$ est le conditionnement du problème, et on a donc

$$\left|\frac{\delta y}{y}\right| \simeq C \left|\frac{\delta x}{x}\right|$$

- Si $C \simeq 1$, ou $C \leqslant 1$, l'erreur en sortie reste au plus du même ordre que celle en entrée : le système est stable
- Sinon, le conditionnement est mauvais et les erreurs sont amplifiées, surtout dans le cas de méthodes **itératives** : le système est instable

Exercice 4 : On considère le nombre réel

$$A = \left(\sqrt{2}-1\right)^6 = \left(3-2\sqrt{2}\right)^3 = 99-70\sqrt{2} \simeq 0.005050633$$

On peut dans ce calcul approcher $\sqrt{2} \simeq 1,414$ par $\sqrt{2} \simeq 1,4 = \dfrac{7}{5}$ ou, un peu mieux $\sqrt{2} \simeq 1,416 \simeq \dfrac{17}{12}$.

Compléter le tableau :

	$\left(\sqrt{2}-1\right)^6$	$\left(3-2\sqrt{2}\right)^3$	$99-70\sqrt{2}$
$\sqrt{2} \simeq \dfrac{7}{5}$			
$\sqrt{2} \simeq \dfrac{17}{12}$			

Exercice 4 : On considère le nombre réel

$$A = \left(\sqrt{2}-1\right)^6 = \left(3-2\sqrt{2}\right)^3 = 99-70\sqrt{2} \simeq 0.005050633$$

On peut dans ce calcul approcher $\sqrt{2} \simeq 1,414$ par $\sqrt{2} \simeq 1,4 = \dfrac{7}{5}$ ou, un peu mieux $\sqrt{2} \simeq 1,416 \simeq \dfrac{17}{12}$.

Compléter le tableau :

	$\left(\sqrt{2}-1\right)^6$	$\left(3-2\sqrt{2}\right)^3$	$99-70\sqrt{2}$
$\sqrt{2} \simeq \dfrac{7}{5}$			
$\sqrt{2} \simeq \dfrac{17}{12}$			

$\hookrightarrow$ Toutes les cases contiennent une valeur approchée : $A \simeq 0,005$!

$\hookrightarrow$ Attention aux puissances $\Longrightarrow$ **propagation** des erreurs d'arrondi

Les méthodes itératives sont des méthodes qui recherchent la solution x d'un problème (solution d'une équation, solution d'un système, ...) en construisant une suite (x_k) qui converge vers la solution x.

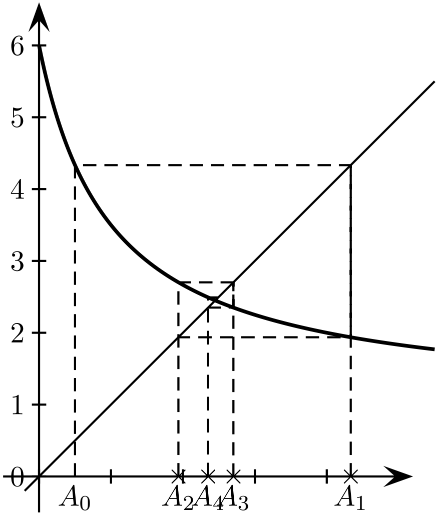
Par exemple, pour déterminer un point fixe x d'une fonction f, i.e. x tel que $x = f(x)$, on peut construire une suite (x_k) par la relation de récurrence

$$x_{n+1} = f\left(x_n\right)$$

Sous certaines hypothèses, cette suite converge vers le point fixe x.
Par contre, sous ces mêmes hypothèses, qu'en est-il du conditionnement de f ? des erreurs, par exemple d'arrondi, fait à chaque itération sur x_k ?

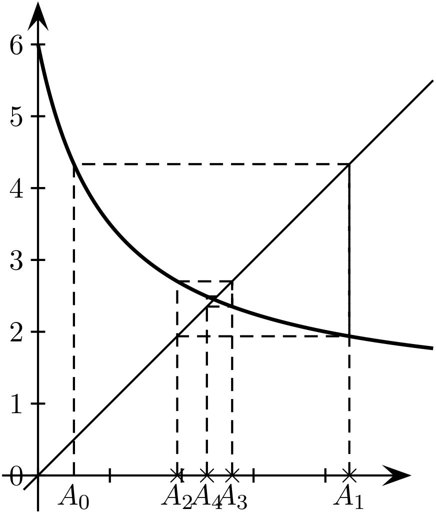
Point fixe : $x = f(x)$ avec $f(x) = 1 + \dfrac{5}{x+1}$.

On définit et construit (x_k) par récurrence $x_0 = \dfrac{1}{2}$ puis $x_{k+1} = f(x_k)$.



Point fixe : $x = f(x)$ avec $f(x) = 1 + \dfrac{5}{x+1}$.

On définit et construit (x_k) par récurrence $x_0 = \dfrac{1}{2}$ puis $x_{k+1} = f(x_k)$.



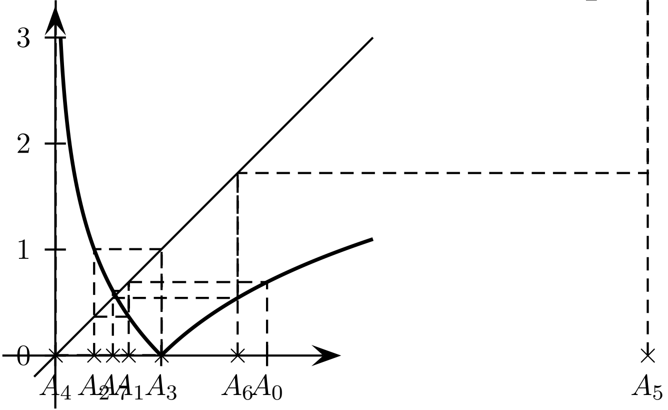
Le point fixe est $l = \sqrt{6}$ pour lequel on a le conditionnement

$$C = \left| l \frac{f'(l)}{f(l)} \right| \simeq 0,3$$

Ainsi, les erreurs relatives d'arrondi sont environ divisées par trois à chaque itération (à proximité de l)

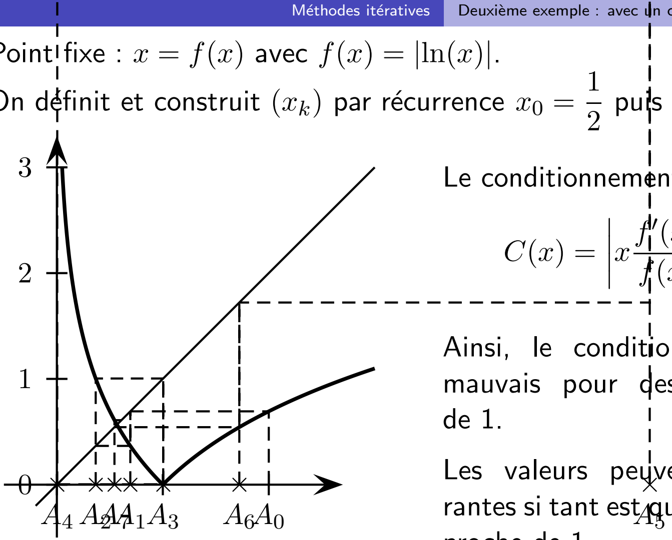
Point fixe : $x = f(x)$ avec $f(x) = |\ln(x)|$.

On définit et construit (x_k) par récurrence $x_0 = \dfrac{1}{2}$ puis $x_{k+1} = f(x_k)$.



Point fixe : $x = f(x)$ avec $f(x) = |\ln(x)|$.

On définit et construit (x_k) par récurrence $x_0 = \dfrac{1}{2}$ puis $x_{k+1} = f(x_k)$.



Le conditionnement est cette fois

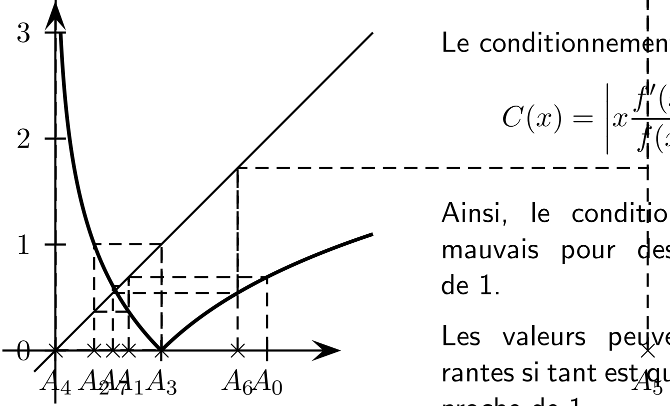
$$C(x) = \left| x \frac{f'(x)}{f(x)} \right| = \frac{1}{|\ln(x)|}$$

Ainsi, le conditionnement est très mauvais pour des valeurs proches de 1.

Les valeurs peuvent devenir aberrantes si tant est qu'une valeur se rapproche de 1.

Point fixe : $x = f(x)$ avec $f(x) = |\ln(x)|$.

On définit et construit (x_k) par récurrence $x_0 = \dfrac{1}{2}$ puis $x_{k+1} = f(x_k)$.



Le conditionnement est cette fois

$$C(x) = \left| x\frac{f'(x)}{f(x)} \right| = \frac{1}{|\ln(x)|}$$

Ainsi, le conditionnement est très mauvais pour des valeurs proches de 1.

Les valeurs peuvent devenir aberrantes si tant est qu'une valeur se rapproche de 1.

On peut s'en rendre clairement compte en calculant les valeurs successives de u_n pour $n \leqslant 30$ dans les 3 cas :

1) $u_0 = 2$ 2) $u_0 = 2 + 10^{-9}$ 3) $u_0 = 2 - 10^{-9}$.