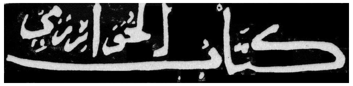
Résolution d'équations

- L'identification des points critiques revient à résoudre

$$\nabla f(x) = 0.$$

- Il s'agit d'un système de n équations à n inconnues.
- On verra que les conditions d'optimalité pour les problèmes avec contraintes se ramènent également à un système d'équations.
- Analysons d'abord les algorithmes permettant de résoudre ces systèmes d'équations.
- Ils seront ensuite adaptés pour les problèmes d'optimisation.

Algorithme



- Abu Ja’far Muhammad ibn Musa Al-Khwarizmi (780—840).
- Traité *al Kitab almukhtasar fi hisab al-jabr w’al muqabala*, qui est à l’origine de l’algèbre.
- traduction latine de cet ouvrage, intitulée *Algoritmi de numero Indorum*.

Méthode de Newton

Idée :

- Méthode itérative
- A chaque itération, simplification du problème
- Technique : linéarisation

Equation à une inconnue

$$f(x) = x^2 - 2, \quad \widehat{x} = 2$$

Théorème de Taylor

$$\begin{aligned} f(\widehat{x} + d) &= f(\widehat{x}) + df'(\widehat{x}) + o(|d|) \\ &= \widehat{x}^2 - 2 + 2\widehat{x}d + o(|d|) \\ &= 2 + 4d + o(|d|). \end{aligned}$$

Equation à une inconnue

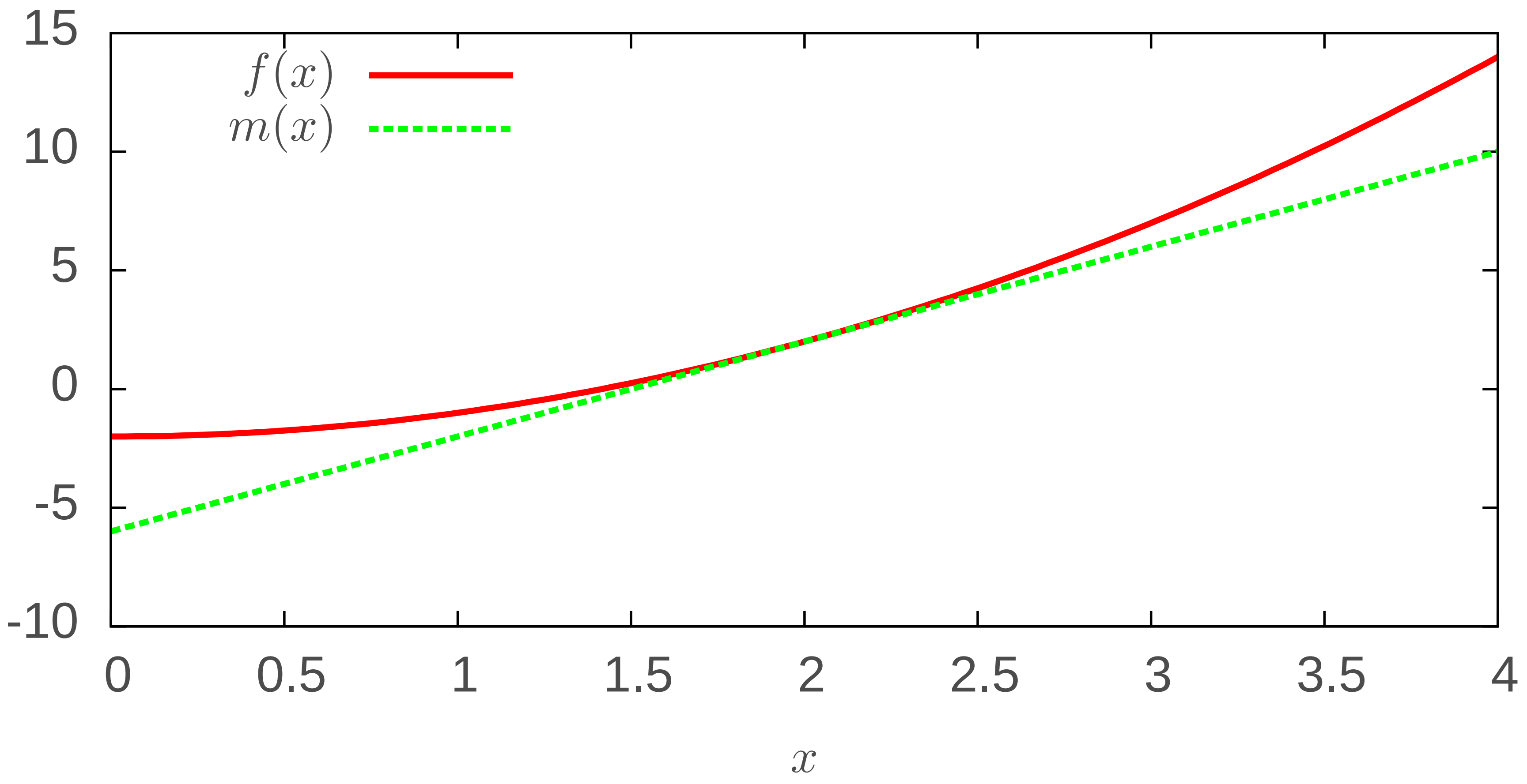
Ignorons le terme d'erreur pour obtenir un modèle :

$$m(\widehat{x} + d) = 2 + 4d.$$

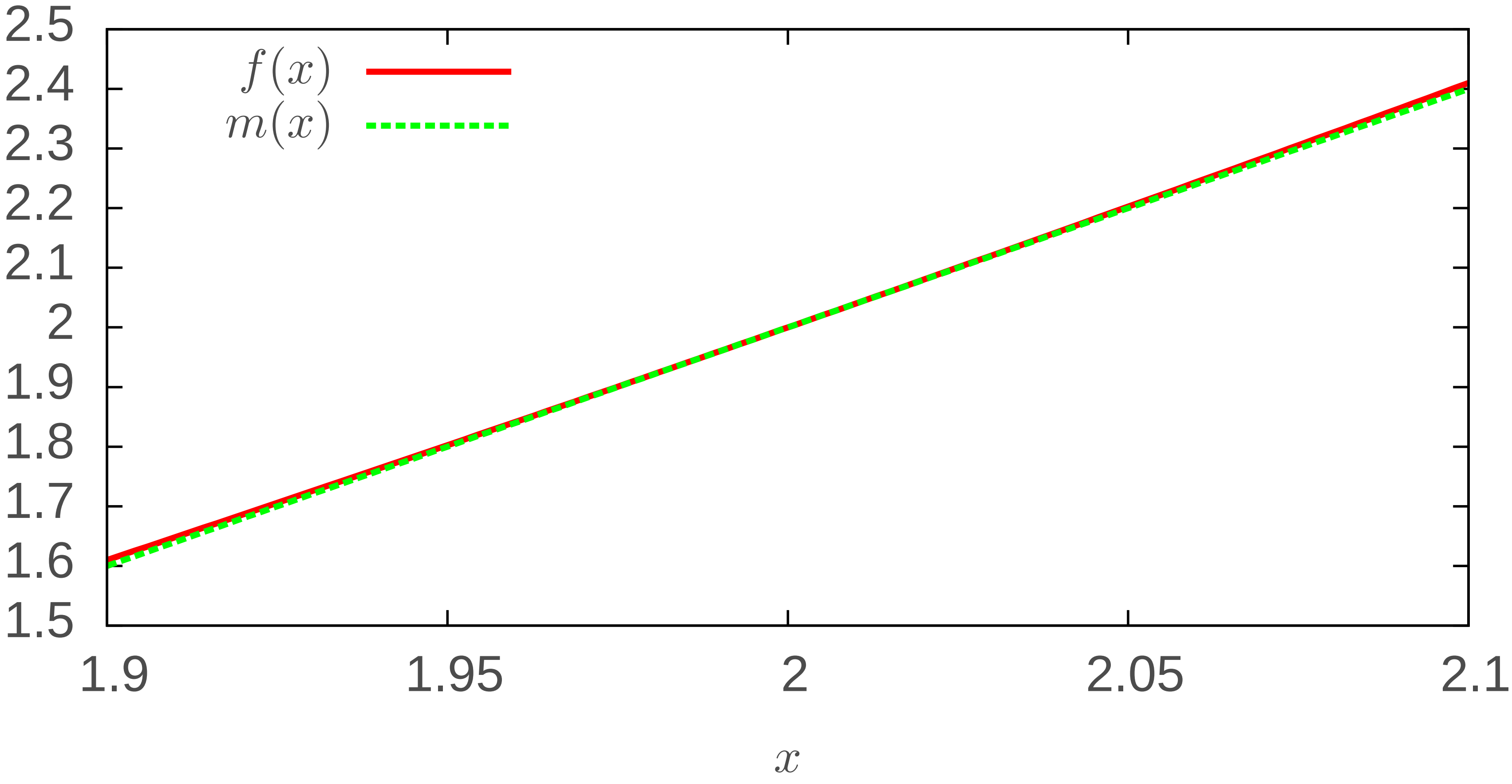
En posant $x = \widehat{x} + d$, nous obtenons

$$m(x) = 2 + 4(x - 2) = 4x - 6.$$

Equation à une inconnue



Equation à une inconnue



Equation à une inconnue

Modèle linéaire d'une fonction à une variable

Soit $f : \mathbb{R} \to \mathbb{R}$ *une fonction différentiable.*

Le modèle linéaire *de* f *en* $\widehat{x}$ *est une fonction* $m_{\widehat{x}} : \mathbb{R} \to \mathbb{R}$ *définie par*

$$m_{\widehat{x}}(x) = f(\widehat{x}) + (x - \widehat{x})f'(\widehat{x}).$$

Equation à une inconnue

Algorithme :

1. Calculer le modèle linéaire en $\widehat{x}$:

$$f(\widehat{x}) + (x - \widehat{x})f'(\widehat{x}) = 0,$$

2. Calculer sa racine x^+

$$x^+ = \widehat{x} - \frac{f(\widehat{x})}{f'(\widehat{x})},$$

3. Si x^+ n'est pas une racine du problème de départ, considérer x^+ comme nouvelle approximation, et recommencer.

Equation à une inconnue

Critère d'arrêt :

- En théorie $f(x^+) = 0$.
- En pratique, arithmétique finie.
- On définit une précision ε, et la condition est

$$|f(x^+)| \leq \varepsilon.$$

Algorithme : Méthode de Newton — une variable

Objectif

Trouver une approximation de la solution de l'équation

$$f(x) = 0.$$

Input

- La fonction $f : \mathbb{R} \to \mathbb{R}$;
- La dérivée de la fonction $f' : \mathbb{R} \to \mathbb{R}$;
- Une première approximation de la solution $x_0 \in \mathbb{R}$;
- La précision demandée $\varepsilon \in \mathbb{R}$, $\varepsilon > 0$.

Algorithme : Méthode de Newton — une variable

Output

Une approximation de la solution $x^* \in \mathbb{R}$

Initialisation

$k = 0$

Itérations

1. $x_{k+1} = x_k - f(x_k)/f'(x_k)$,
2. $k = k + 1$.

Critère d'arrêt

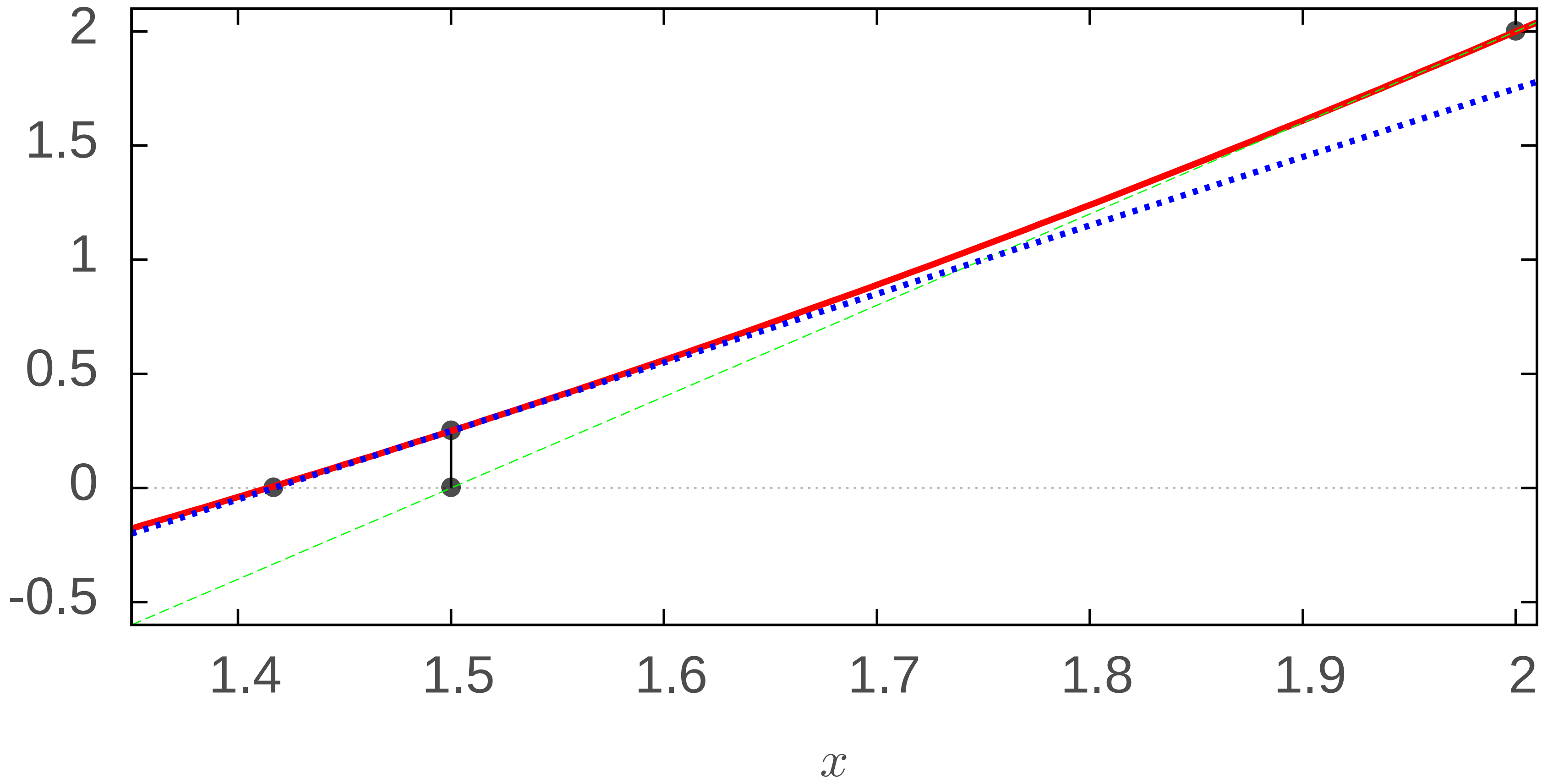
Si $|f(x_k)| \leq \varepsilon$, alors $x^* = x_k$.

Méthode de Newton — une variable

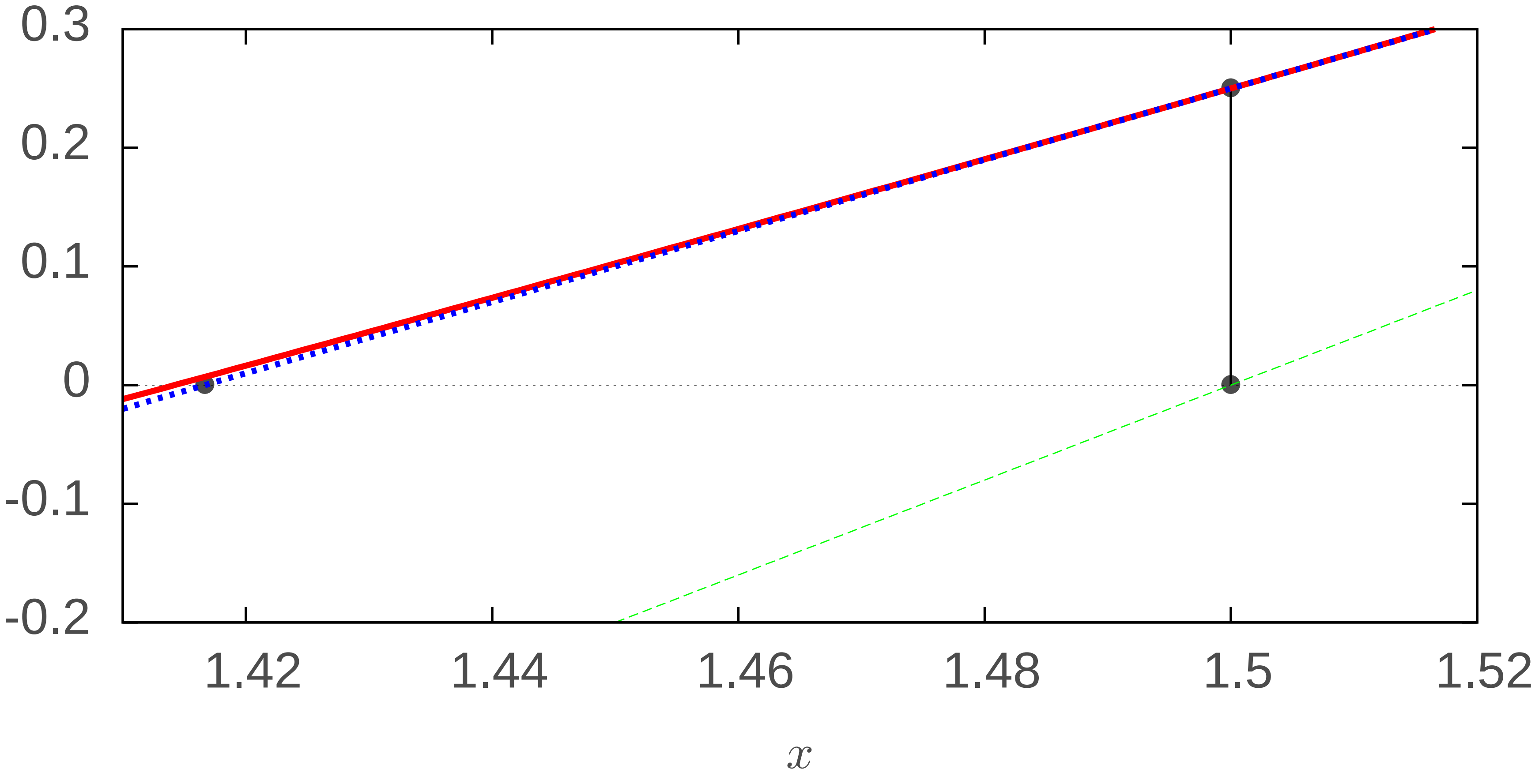
$$f(x) = x^2 - 2 = 0, \quad x_0 = 2, \quad \varepsilon = 10^{-15}$$

k	x_k	$f(x_k)$	$f'(x_k)$
0	+2.00000000E+00	+2.00000000E+00	+4.00000000E+00
1	+1.50000000E+00	+2.50000000E-01	+3.00000000E+00
2	+1.41666667E+00	+6.94444444E-03	+2.83333333E+00
3	+1.41421569E+00	+6.00730488E-06	+2.82843137E+00
4	+1.41421356E+00	+4.51061410E-12	+2.82842712E+00
5	+1.41421356E+00	+4.44089210E-16	+2.82842712E+00

Méthode de Newton — une variable



Méthode de Newton — une variable



Méthode de Newton — une variable

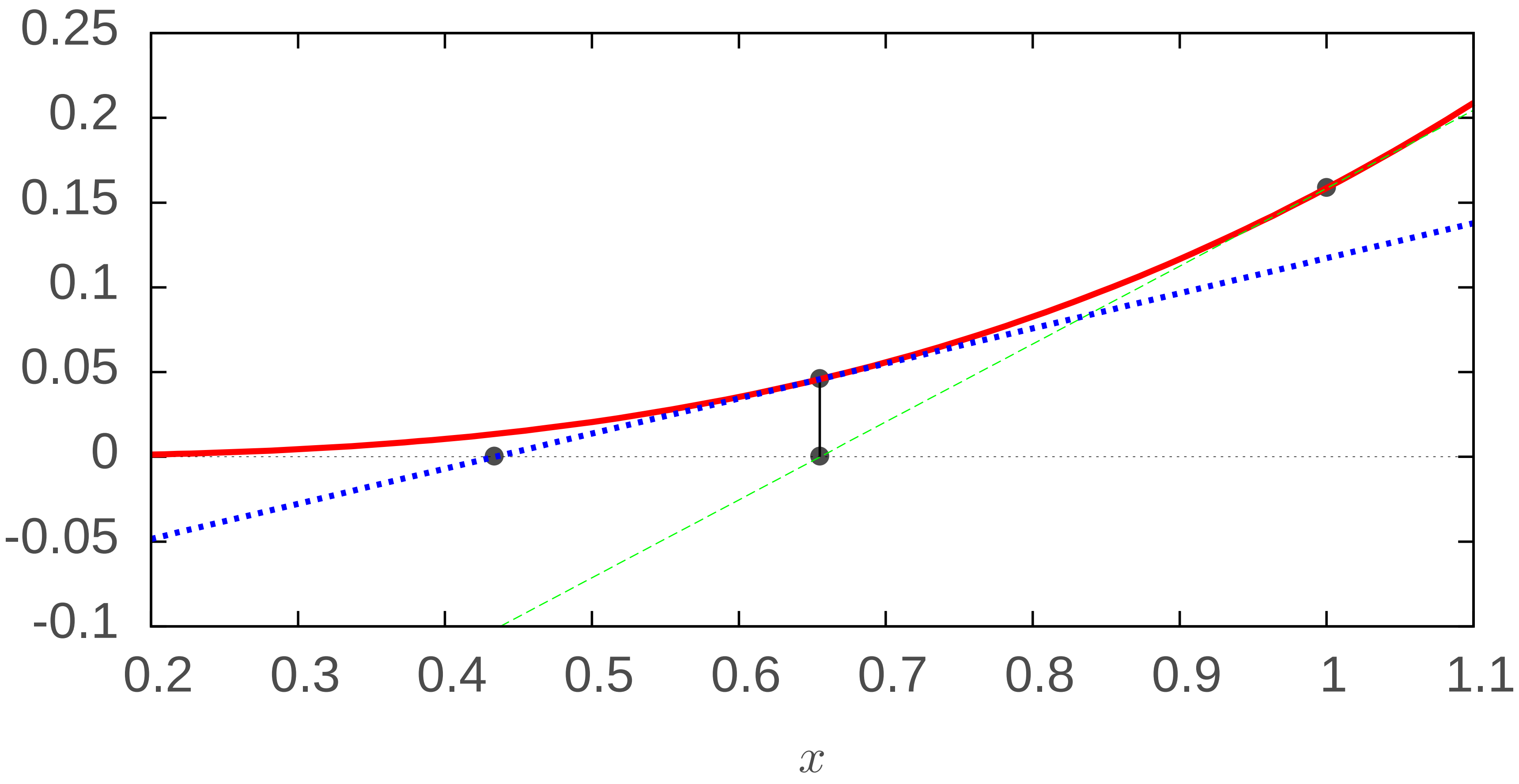
- La méthode semble très rapide
- Malheureusement, elle ne fonctionne pas toujours bien

Méthode de Newton — une variable

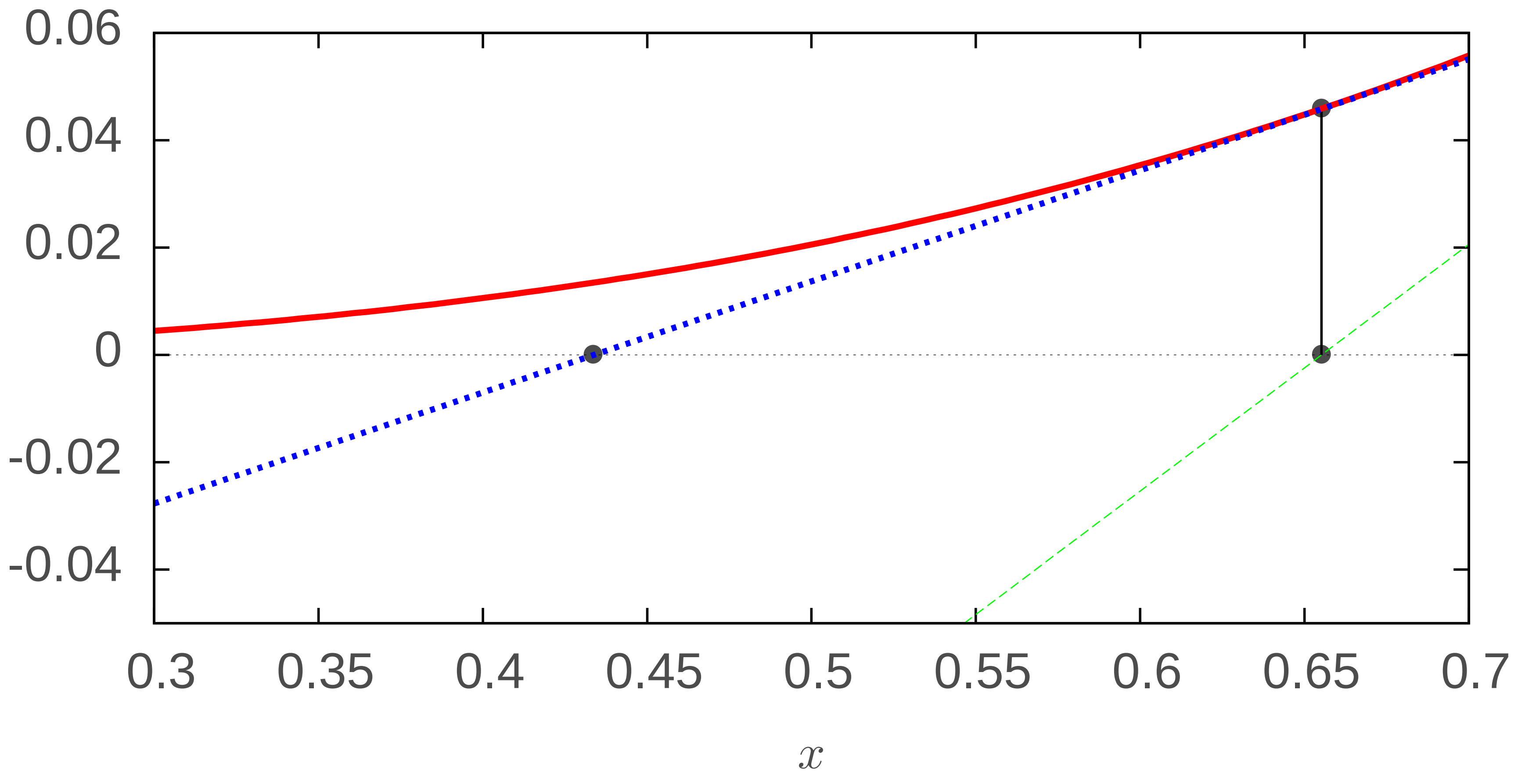
$$f(x) = x - \sin x = 0, \quad x_0 = 1, \quad \varepsilon = 10^{-15}$$

k	x_k	$f(x_k)$	$f'(x_k)$
0	+1.00000000E+00	+1.58529015E-01	+4.59697694E-01
1	+6.55145072E-01	+4.58707860E-02	+2.07040452E-01
2	+4.33590368E-01	+1.34587380E-02	+9.25368255E-02
3	+2.88148401E-01	+3.97094846E-03	+4.12282985E-02
4	+1.91832312E-01	+1.17439692E-03	+1.83434616E-02
⋮			
23	+8.64386893E-05	+1.07639890E-13	+3.73582354E-09
24	+5.76257947E-05	+3.18933045E-14	+1.66036607E-09
25	+3.84171966E-05	+9.44986548E-15	+7.37940486E-10
26	+2.56114682E-05	+2.79996227E-15	+3.27973648E-10
27	+1.70743119E-05	+8.29617950E-16	+1.45766066E-10

Méthode de Newton — une variable



Méthode de Newton — une variable

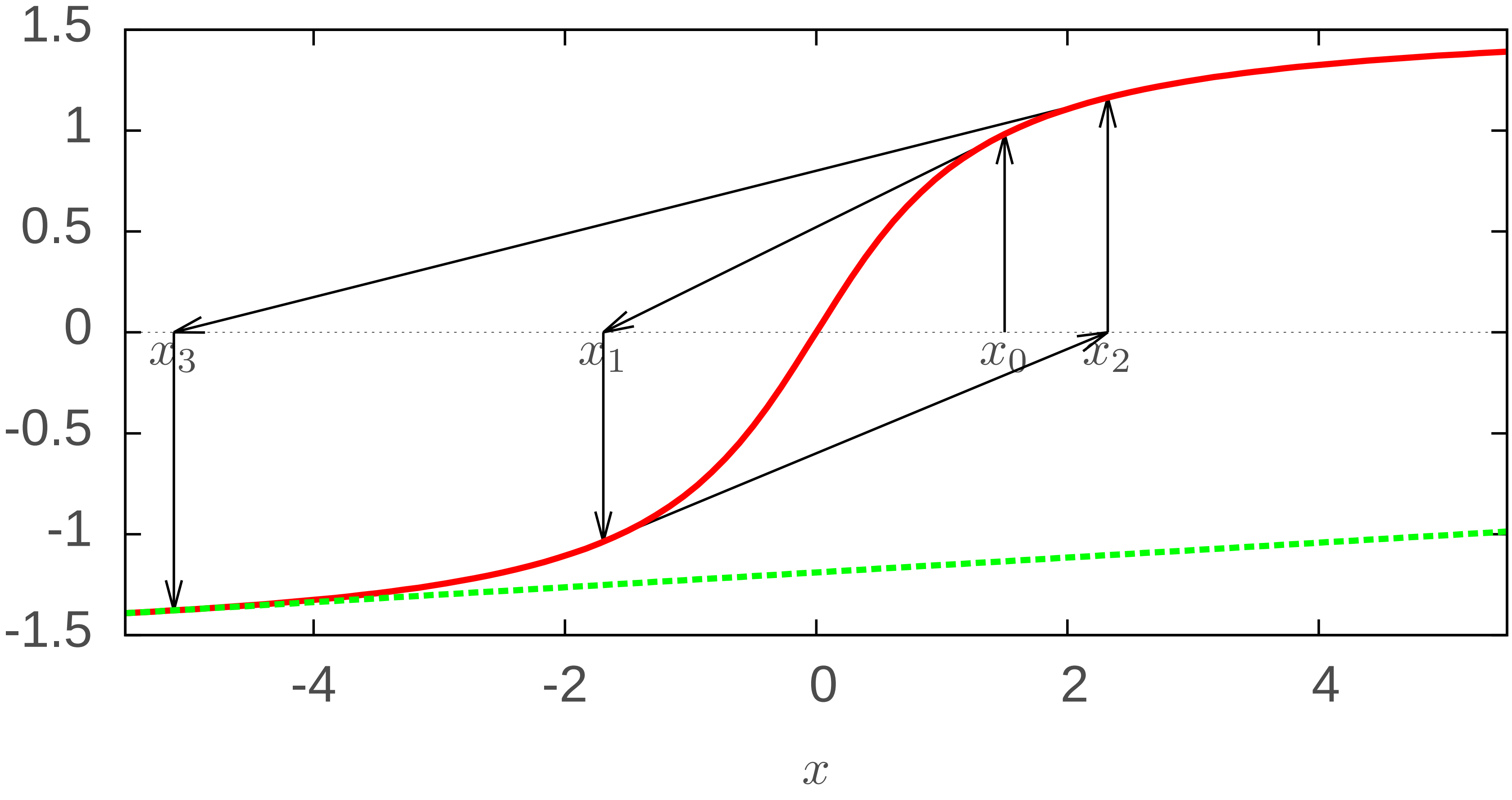


Méthode de Newton — une variable

$$f(x) = \arctan x = 0, \quad x_0 = 1.5, \quad \varepsilon = 10^{-15}$$

k	x_k	$f(x_k)$	$f'(x_k)$
0	+1.50000000E+00	+9.82793723E-01	+3.07692308E-01
1	-1.69407960E+00	-1.03754636E+00	+2.58404230E-01
2	+2.32112696E+00	+1.16400204E+00	+1.56552578E-01
3	-5.11408784E+00	-1.37769453E+00	+3.68271300E-02
4	+3.22956839E+01	+1.53984233E+00	+9.57844131E-04
5	-1.57531695E+03	-1.57016153E+00	+4.02961851E-07
6	+3.89497601E+06	+1.57079607E+00	+6.59159364E-14
7	-2.38302890E+13	-1.57079633E+00	+1.76092712E-27
8	+8.92028016E+26	+1.57079633E+00	+1.25673298E-54
9	-1.24990460E+54	-1.57079633E+00	+6.40097701E-109
10	+2.45399464E+108	+1.57079633E+00	+1.66055315E-217

Méthode de Newton — une variable



TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Méthode de Newton — une variable

Performance de la méthode de Newton

- Si la fonction n'est pas trop non-linéaire;
- Si la dérivée de f à la solution n'est pas trop proche de 0;
- Si x_0 n'est pas trop éloigné de la racine;
- Alors la méthode de Newton converge très vite vers la solution.

Méthode de Newton — une variable

Erreur du modèle linéaire — une variable *Soit un intervalle ouvert $X \subseteq \mathbb{R}$, et une fonction f dont la dérivée est continue au sens de Lipschitz sur X, la constante de Lipschitz étant M.*
Alors, pour tout $\widehat{x}$, $x^+ \in X$,

$$|f(x^+) - m_{\widehat{x}}(x^+)| \leq M \frac{(x^+ - \widehat{x})^2}{2}.$$

(sans preuve)

Rappel

Borne pour l'intégrale *Soit $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}^{m\times n}$, où X est un ensemble ouvert convexe, et soient x et $x+d$ dans X. Alors, si f est intégrable sur $[x, x+d]$,*

$$\begin{aligned} \left\| \int_0^1 f(x+td)d \, dt \right\| &\leq \int_0^1 \|f(x+td)d\| \, dt \\ &\leq \int_0^1 \|f(x+td)\| \, \|d\| \, dt \end{aligned}$$

où $\|\cdot\|$ est une norme sur $\mathbb{R}^{m\times n}$.

(sans preuve)

Méthode de Newton — une variable

Convergence de la méthode de Newton — une variable *Soit un intervalle ouvert $X \subseteq \mathbb{R}$, et une fonction f dont la dérivée est continue au sens de Lipschitz sur X, la constante de Lipschitz étant M.*
Supposons qu'il existe $\rho > 0$ tel que $|f'(x)| \geq \rho \quad \forall x \in X$. Supposons qu'il existe $x^* \in X$ tel que $f(x_*) = 0$.
Alors, il existe $\eta > 0$ tel que, si $|x_0 - x^*| < \eta$ avec $x_0 \in X$, la suite $(x_k)_k$ définie par

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \quad k = 0, 1, \ldots$$

est bien définie et converge vers x^*.
De plus, $|x_{k+1} - x^*| \leq \frac{M}{2\rho} |x_k - x^*|^2$.

(p. 197)

Méthode de Newton — une variable

- Fonction pas trop non-linéaire M petit
- Dérivée pas trop proche de 0 $|f'(x)| \geq \rho \quad \forall x \in X.$
- x_0 n'est pas trop éloigné de la racine $|x_0 - x^*| < \eta$, avec

$$\eta = \min(r, \tau \frac{2\rho}{M}).$$

M petit $\Rightarrow \eta$ grand

Méthode de Newton — une variable

- Convergence quadratique

$$|x_{k+1} - x^*| \leq \frac{M}{2\rho} |x_k - x^*|^2.$$

Méthode de Newton — une variable

Convergence q-quadratique

Soit une suite $(x_k)_k$ dans $\mathbb{R}^n$ convergeant vers x^*.
On dit que la suite converge q-quadratiquement vers x^* s'il existe $c \geq 0$ et $\widehat{k} \in \mathbb{N}$ tels que

$$\|x_{k+1} - x^*\| \leq c\|x_k - x^*\|^2 \quad \forall k \geq \widehat{k}.$$

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Méthode de Newton — n variables

Modèle linéaire d'une fonction à n variables

Soit $f : \mathbb{R}^n \to \mathbb{R}^m$ *une fonction différentiable.*

Le modèle linéaire *de* f *en* $\widehat{x}$ *est une fonction* $m_{\widehat{x}} : \mathbb{R}^n \to \mathbb{R}^m$ *définie par*

$$m_{\widehat{x}}(x) = f(\widehat{x}) + \nabla f(\widehat{x})^T (x - \widehat{x}) = f(\widehat{x}) + J(\widehat{x})(x - \widehat{x})$$

où $\nabla f(\widehat{x})$ *est la matrice gradient de* f *en* $\widehat{x}$ *et* $J(\widehat{x}) = \nabla f(\widehat{x})^T$ *la matrice Jacobienne.*

Méthode de Newton — n variables

Erreur du modèle linéaire — n variables *Soit $f : \mathbb{R}^n \to \mathbb{R}^m$ une fonction différentiable sur un ensemble ouvert convexe $X \subset \mathbb{R}^n$, dont la matrice jacobienne est continue au sens de Lipschitz sur X, où M est la constante de Lipschitz.*
Alors, pour tout $\widehat{x}$, $x^+ \in X$,

$$\|f(x^+) - m_{\widehat{x}}(x^+)\| \leq M \frac{\|x^+ - \widehat{x}\|^2}{2}.$$

(sans preuve)

Algorithme : Méthode de Newton — n variables

Objectif

Trouver une approximation de la solution du système d'équations

$$f(x) = 0.$$

Input

- La fonction $f : \mathbb{R}^n \to \mathbb{R}^n$;
- La matrice jacobienne de la fonction $J : \mathbb{R}^n \to \mathbb{R}^{n \times n}$;
- Une première approximation de la solution $x_0 \in \mathbb{R}^n$;
- La précision demandée $\varepsilon \in \mathbb{R}$, $\varepsilon > 0$.

Output

Une approximation de la solution $x^* \in \mathbb{R}^n$

Algorithme : Méthode de Newton — n variables

Initialisation

$k = 0$

Itérations

1. Calculer d_{k+1} solution de $J(x_k)d_{k+1} = -f(x_k)$.
2. $x_{k+1} = x_k + d_{k+1}$.
3. $k = k + 1$.

Critère d'arrêt

Si $\|f(x_k)\| \leq \varepsilon$, alors $x^* = x_k$.

Méthode de Newton — n variables

$$\begin{aligned}(x_1+1)^2+x_2^2 &= 2\\ e^{x_1}+x_2^3 &= 2.\end{aligned}$$

$$f(x)=\begin{pmatrix}(x_1+1)^2+x_2^2-2\\ e^{x_1}+x_2^3-2\end{pmatrix}\quad J(x)=\begin{pmatrix}2(x_1+1) & 2x_2\\ e^{x_1} & 3x_2^2\end{pmatrix}$$

Si $x_0=(1\ 1)^T$, nous avons

$$f(x_0)=\begin{pmatrix}3\\ e-1\end{pmatrix}\approx\begin{pmatrix}3\\ 1.7182\end{pmatrix},\quad J(x_0)=\begin{pmatrix}4 & 2\\ e & 3\end{pmatrix}$$

Méthode de Newton — n variables

k	x_k	$f(x_k)$	$\|f(x_k)\|$
0	1.00000000e+00	3.00000000e+00	3.45723768e+00
	1.00000000e+00	1.71828182e+00	
1	1.52359213e-01	7.56629795e-01	1.15470870e+00
	1.19528157e+00	8.72274931e-01	
2	-1.08376809e-02	5.19684443e-02	1.14042557e-01
	1.03611116e+00	1.01513475e-01	
3	-8.89664601e-04	1.29445248e-03	3.94232975e-03
	1.00153531e+00	3.72375572e-03	
4	-1.37008875e-06	3.13724882e-06	8.07998556e-06
	1.00000293e+00	7.44606181e-06	
5	-5.53838974e-12	1.05133679e-11	2.88316980e-11
	1.00000000e+00	2.68465250e-11	
6	-1.53209346e-16	-2.22044604e-16	2.22044604e-16
	1.00000000e+00	0.00000000e+00	

Méthode de Newton — n variables

Convergence de la méthode de Newton — n variables *Soit un ensemble convexe ouvert $X \subseteq \mathbb{R}^n$, et une fonction $f : X \to \mathbb{R}^n$. Supposons qu'il existe $x^* \in X$, une boule $B(x^*, r)$ centrée en x^* de rayon r, et une constante $\rho > 0$ tels que $f(x^*) = 0$, $B(x^*, r) \subset X$, $J(x^*)$ est inversible,*

$$\|J(x^*)^{-1}\| \leq \frac{1}{\rho},$$

et J est continue au sens de Lipschitz sur $B(x^*, r)$, la constante de Lipschitz étant M.

(suite...)

Méthode de Newton — n variables

Convergence de la méthode de Newton — n variables (suite)

Alors, il existe $\eta > 0$ tel que si

$$x_0 \in B(x^*, \eta),$$

alors la suite $(x_k)_k$ définie par

$$x_{k+1} = x_k - J(x_k)^{-1} f(x_k) \quad k = 0, 1, \ldots$$

est bien définie et converge vers x^*. De plus,

$$\|x_{k+1} - x^*\| \leq \frac{M}{\rho} \|x_k - x^*\|^2.$$

(sans preuve)

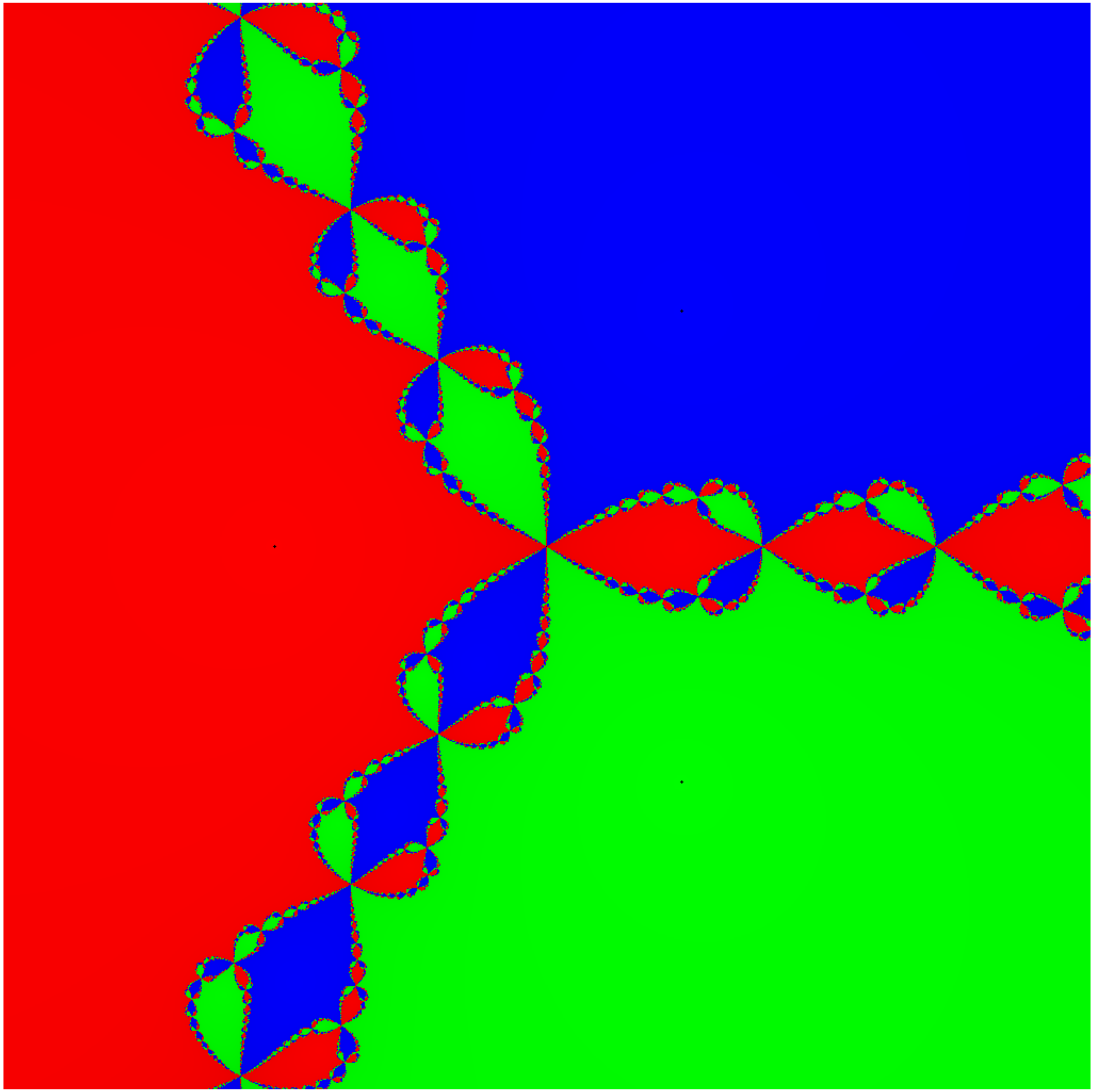
Exemple

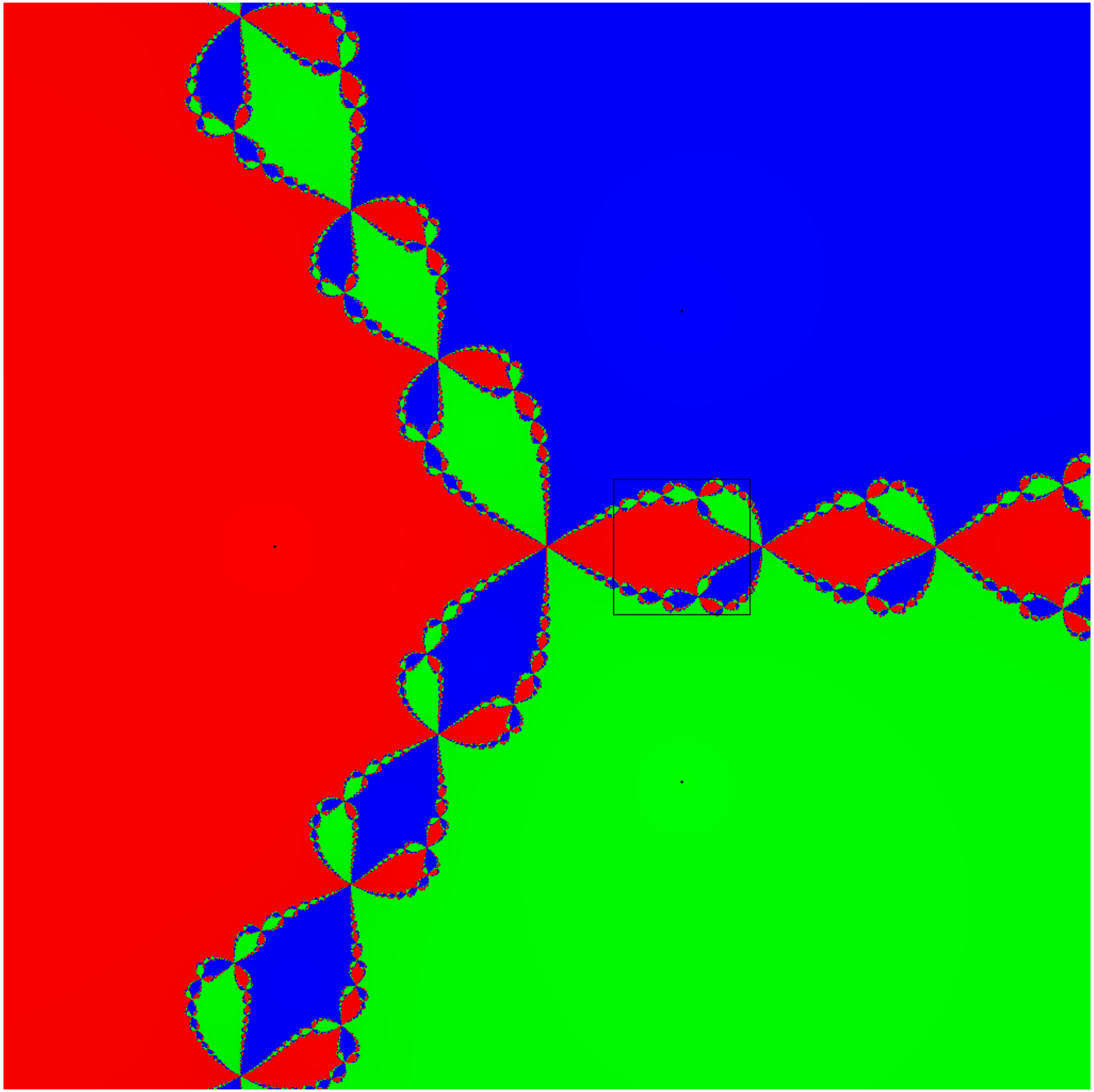
$$f(x,y) = \begin{pmatrix} x^3 - 3xy^2 - 1 \\ y^3 - 3x^2y \end{pmatrix} = 0$$

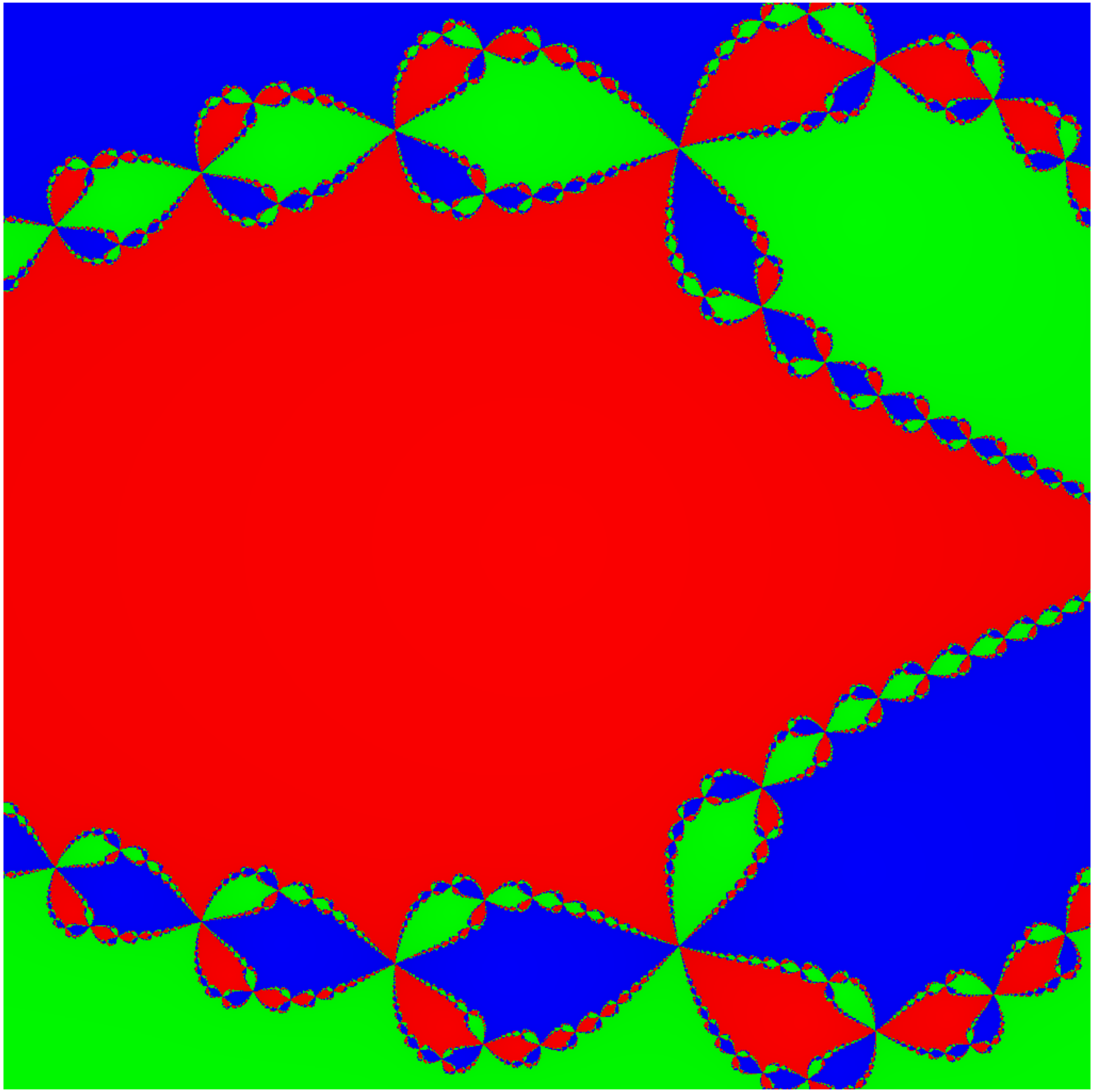
Solutions :

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \begin{pmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix} \quad \begin{pmatrix} -\frac{1}{2} \\ -\frac{\sqrt{3}}{2} \end{pmatrix}$$

En fonction du point de départ, la méthode de Newton converge vers une racine ou l’autre







Méthode de Newton — n variables

Inconvénients de la méthode de Newton :

- Importance du point de départ. Solution : techniques de globalisation (seront vues dans le cadre de l'optimisation)
- Nécessite le calcul de la matrice des dérivées. Solution : méthodes quasi-Newton.

Quasi-Newton — une variable

$$f'(x) = \lim_{s \to 0} \frac{f(x+s) - f(x)}{s}.$$

Si l'on prend s petit

$$a_s(x) = \frac{f(x+s) - f(x)}{s}.$$

Reprenons les idées de la méthode de Newton, en utilisant cette approximation

Quasi-Newton — une variable

Modèle linéaire sécant d'une fonction à une variable

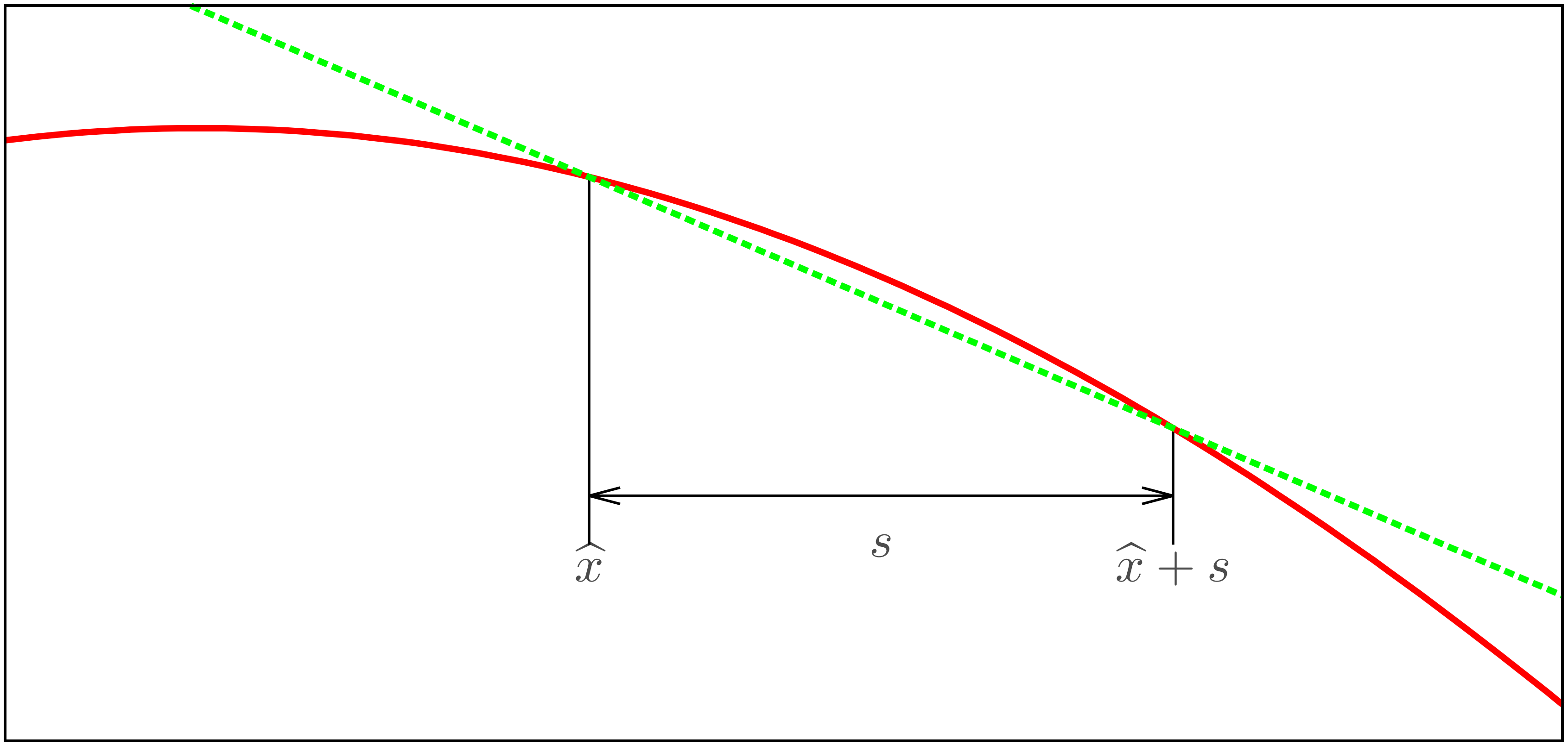
Soit $f : \mathbb{R} \to \mathbb{R}$ *une fonction différentiable. Le* modèle linéaire sécant *de* f *en* $\widehat{x}$ *est une fonction* $m_{\widehat{x};s} : \mathbb{R} \to \mathbb{R}$ *définie par*

$$\begin{aligned} m_{\widehat{x};s}(x) &= f(\widehat{x}) + \frac{f(\widehat{x}+s)-f(\widehat{x})}{s}(x-\widehat{x}) \\ &= f(\widehat{x}) + a_s(\widehat{x})(x-\widehat{x}), \end{aligned}$$

où $s > 0$*.*

Itération quasi-Newton : $x^+ = \widehat{x} - \frac{f(\widehat{x})}{a_s(\widehat{x})}$.

Quasi-Newton — une variable



Quasi-Newton — une variable

- Pour avoir une bonne approximation, choisir s petit.

$$s = \begin{cases} \tau \widehat{x} & \text{si } |\widehat{x}| \geq 1, \\ \tau & \text{sinon} \end{cases}$$

  où $\tau = 10^{-7}$.
- On obtient la méthode de Newton par différence finie

Algorithme : Différence finie — une variable

Objectif

Trouver une approximation de la solution de l'équation

$$f(x) = 0.$$

Input

- La fonction $f : \mathbb{R} \to \mathbb{R}$;
- Une première approximation de la solution $x_0 \in \mathbb{R}$;
- Un paramètre $\tau > 0$.
- La précision demandée $\varepsilon \in \mathbb{R}$, $\varepsilon > 0$.

Algorithme : Différence finie — une variable

Output

Une approximation de la solution $x^* \in \mathbb{R}$

Initialisation

$k = 0$

Algorithme : Différence finie — une variable

Itérations

1. $$s = \begin{cases} \tau x_k & \text{si } |x_k| \geq 1, \\ \tau & \text{sinon} \end{cases}$$

2. $$x_{k+1} = x_k - \frac{s f(x_k)}{f(x_k + s) - f(x_k)}$$

3. $k = k + 1$.

Critère d'arrêt

Si $|f(x_k)| \leq \varepsilon$, alors $x^* = x_k$.

Quasi-Newton — une variable

$$f(x) = x^2 - 2 = 0 \quad \tau = 10^{-7}$$

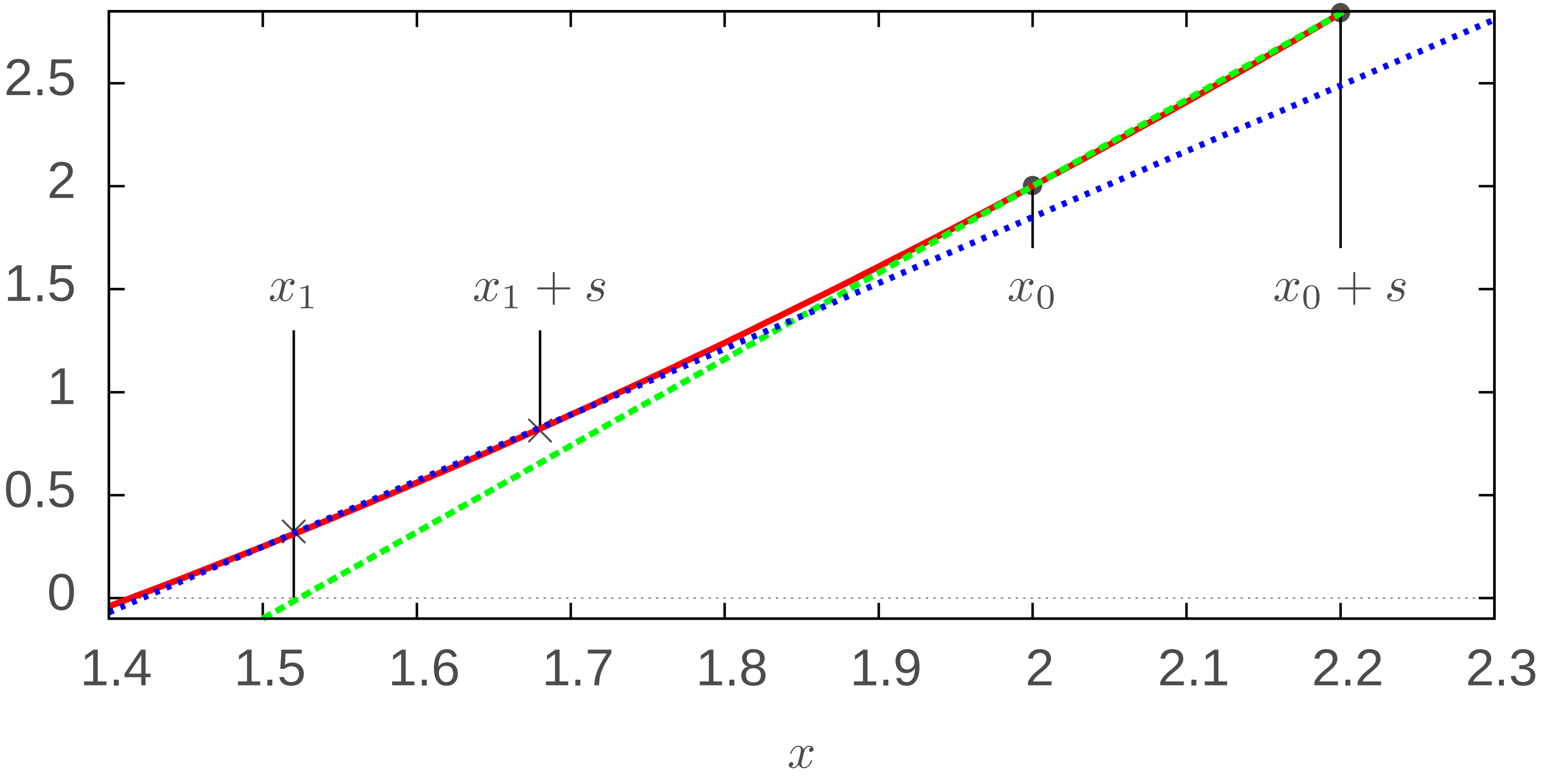
Diff. finie		Newton	
	$f(x_k)$	x_k	$f(x_k)$
.00000000E+00	+2.00000000E+00	+2.00000000E+00	+2.00000000E+00
.50000003E+00	+2.50000076E-01	+1.50000000E+00	+2.50000000E-01
.41666667E+00	+6.94446047E-03	+1.41666667E+00	+6.94444444E-03
.41421569E+00	+6.00768206E-06	+1.41421569E+00	+6.00730488E-06
.41421356E+00	+4.81081841E-12	+1.41421356E+00	+4.51061410E-12
.41421356E+00	+4.44089210E-16	+1.41421356E+00	+4.44089210E-16

Quasi-Newton — une variable

$$f(x) = x^2 - 2 = 0 \quad \tau = 0.1$$

k	x_k	$f(x_k)$
0	+2.00000000E+00	+2.00000000E+00
1	+1.52380952E+00	+3.21995465E-01
2	+1.42318594E+00	+2.54582228E-02
3	+1.41466775E+00	+1.28485582E-03
⋮		
9	+1.41421356E+00	+1.50284230E-11
10	+1.41421356E+00	+7.15427717E-13
11	+1.41421356E+00	+3.41948692E-14
12	+1.41421356E+00	+1.33226763E-15
13	+1.41421356E+00	+4.44089210E-16

Quasi-Newton — une variable



Quasi-Newton — une variable

- On peut se permettre une valeur de s relativement grande
- Il y a perte d'efficacité
- Idée : économiser un calcul de $f(x_k + s)$ en choisissant

$$s = x_{k-1} - x_k$$

- Ainsi,

$$a_s(x_k) = \frac{f(x_{k-1}) - f(x_k)}{x_{k-1} - x_k}.$$

- On obtient la méthode de Newton sécante

Algorithme : Méthode sécante — une variable

Objectif

Trouver une approximation de la solution de l'équation

$$f(x) = 0.$$

Input

- La fonction $f : \mathbb{R} \to \mathbb{R}$;
- Une première approximation de la solution $x_0 \in \mathbb{R}$;
- Une première approximation de la dérivée a_0 (par défaut: $a_0 = 1$).
- La précision demandée $\varepsilon \in \mathbb{R}$, $\varepsilon > 0$.

Algorithme : Méthode sécante — une variable

Output

Une approximation de la solution $x^* \in \mathbb{R}$

Initialisation

$k = 0$.

Algorithme : Méthode sécante — une variable

Itérations

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

2.
$$a_{k+1} = \frac{f(x_k) - f(x_{k+1})}{x_k - x_{k+1}}$$

3. $k = k + 1$.

Critère d'arrêt

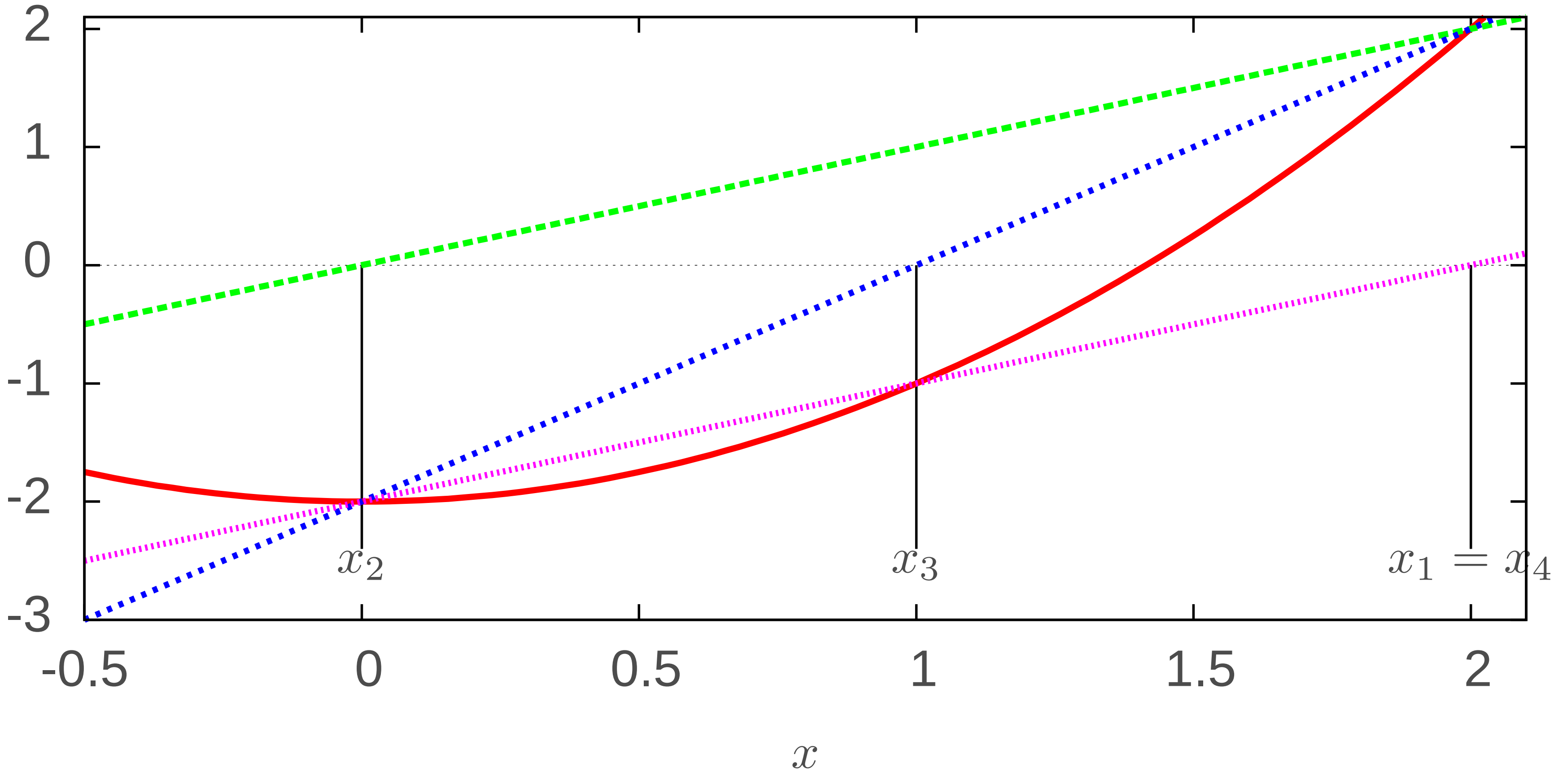
Si $|f(x_k)| \leq \varepsilon$, alors $x^* = x_k$.

Quasi-Newton — une variable

$$f(x) = x^2 - 2 = 0 \quad a_0 = 1$$

k	x_k	$f(x_k)$	$a_s(x_k)$
0	+2.00000000E+00	+2.00000000E+00	+1.00000000E+00
1	+0.00000000E+00	-2.00000000E+00	+2.00000000E+00
2	+1.00000000E+00	-1.00000000E+00	+1.00000000E+00
3	+2.00000000E+00	+2.00000000E+00	+3.00000000E+00
4	+1.33333333E+00	-2.22222222E-01	+3.33333333E+00
5	+1.40000000E+00	-4.00000000E-02	+2.73333333E+00
6	+1.41463415E+00	+1.18976800E-03	+2.81463415E+00
7	+1.41421144E+00	-6.00728684E-06	+2.82884558E+00
8	+1.41421356E+00	-8.93145558E-10	+2.82842500E+00
9	+1.41421356E+00	+8.88178420E-16	+2.82842706E+00

Quasi-Newton — une variable



Quasi-Newton — $n$variables

Modèle linéaire sécant d'une fonction à n variables

Soit $f : \mathbb{R}^n \to \mathbb{R}^m$ *une fonction différentiable, et* $A \in \mathbb{R}^{m \times n}$.

Le modèle linéaire sécant *de* f *en* $\widehat{x}$ *est une fonction* $m_{\widehat{x};A} : \mathbb{R}^n \to \mathbb{R}^m$ *définie par*

$$m_{\widehat{x};A}(x) = f(\widehat{x}) + A(x - \widehat{x}).$$

Algorithme : Différence finie — n variables

Objectif

Trouver une approximation de la solution du système d'équations

$$f(x) = 0.$$

Input

- La fonction $f : \mathbb{R}^n \to \mathbb{R}^n$;
- Une première approximation de la solution $x_0 \in \mathbb{R}^n$;
- Un paramètre $\tau > 0$;
- La précision demandée $\varepsilon \in \mathbb{R}$, $\varepsilon > 0$.

Algorithme : Différence finie — n variables

Output

Une approximation de la solution $x^* \in \mathbb{R}^n$

Initialisation

$k = 0$

Algorithme : Différence finie — n variables

Itérations

- Former le vecteur $s \in \mathbb{R}^n$ en posant

$$s_j = \begin{cases} \tau (x_k)_j & \text{si } |(x_k)_j| \geq 1, \\ \tau & \text{si } 0 \leq (x_k)_j \leq 1 \\ -\tau & \text{si } -1 \leq (x_k)_j \leq 0 \end{cases} \qquad j = 1, \ldots, n,$$

- Former la matrice A_k dont la colonne j est

$$(A_k)_j = \frac{f(x_k + s_j e_j) - f(x_k)}{s_j} \quad j = 1, \ldots, n,$$

Algorithme : Différence finie — n variables

Itérations (suite)

- Calculer d_{k+1} solution de $A_k d_{k+1} = -f(x_k)$,
- $x_{k+1} = x_k + d_{k+1}$,
- $k = k + 1$.

Critère d'arrêt

Si $\|f(x_k)\| \leq \varepsilon$, alors $x^* = x_k$.

Quasi-Newton — n variables

- Désavantage : nécessite $n+1$ évaluations de la fonction par itération.
- Préférable d'utiliser une méthode sécante, c'est-à-dire utiliser un modèle qui interpole la fonction en x_k et x_{k-1}.
- Par définition du modèle linéaire, on a toujours

$$m_{x_k;A_k}(x_k) = f(x_k)$$

- Il faut donc imposer

$$m_{x_k;A_k}(x_{k-1}) = f(x_k) + A_k(x_{k-1} - x_k) = f(x_{k-1}).$$

Quasi-Newton — n variables

Equation sécante

*Un modèle linéaire vérifie l'*équation sécante *en* x_k *et* x_{k-1} *si la matrice* A *le définissant est telle que*

$$A(x_{k-1} - x_k) = f(x_{k-1}) - f(x_k).$$

En posant

$$\begin{aligned} d_{k-1} &= x_k - x_{k-1} \\ y_{k-1} &= f(x_k) - f(x_{k-1}), \end{aligned}$$

elle s'écrit

$$Ad_{k-1} = y_{k-1}.$$

Quasi-Newton — n variables

- L'équation sécante est un système de n équations à n^2 inconnues.
- Il y a un nombre infini de solutions.
- Idée de Broyden : privilégier, parmi ce nombre infini de modèles linéaires, celui qui est le plus proche du modèle établi lors de l'itération précédente.

Quasi-Newton — n variables

Lemme *Soient $m_{x_k;A_k}(x)$ et $m_{x_{k-1};A_{k-1}}(x)$ les modèles linéaires sécants d'une fonction $f:\mathbb{R}^n \to \mathbb{R}^n$ en x_k et x_{k-1}, respectivement.*
Alors,

$$m_{x_k;A_k}(x) - m_{x_{k-1};A_{k-1}}(x) = (A_k - A_{k-1})(x - x_{k-1}).$$

(p. 219)

Quasi-Newton — n variables

Mise à jour de Broyden *Soit* $m_{x_{k-1};A_{k-1}}(x)$ *le modèle linéaire sécant d'une fonction* $f : \mathbb{R}^n \to \mathbb{R}^n$ *en* x_{k-1}*, et soit* $x_k \in \mathbb{R}^n$*,* $x_k \neq x_{k-1}$*.*
Le modèle linéaire sécant de f *en* x_k *qui vérifie l'équation sécante, et qui est le plus semblable au modèle* $m_{x_{k-1};A_{k-1}}(x)$ *est*

$$m_{x_k;A_k}(x) = f(x_k) + A_k(x - x_k),$$

avec

$$A_k = A_{k-1} + \frac{(y_{k-1} - A_{k-1}d_{k-1})d_{k-1}^T}{d_{k-1}^T d_{k-1}}.$$

(p. 220)

Quasi-Newton — n variables

Optimalité de Broyden *Soit* $A_{k-1} \in \mathbb{R}^{n\times n}$*,* d_{k-1} *et* $y_{k-1} \in \mathbb{R}^n$*,* $d_{k-1} \neq 0$*.*
Soit $\mathcal{S} = \{A | Ad_{k-1} = y_{k-1}\}$ *l'ensemble des matrices vérifiant l'équation sécante.*
Alors

$$A_k = A_{k-1} + \frac{(y_{k-1} - A_{k-1}d_{k-1})d_{k-1}^T}{d_{k-1}^T d_{k-1}}$$

est la solution de $\min_{A\in\mathcal{S}} \|A - A_{k-1}\|_2$ *et l'unique solution de* $\min_{A\in\mathcal{S}} \|A - A_{k-1}\|_F$ *.*

(p. 221)

Rappels

Norme de Frobenius

Soit $A \in \mathbb{R}^{m \times n}$*. La* norme de Frobenius *de* A *est*

$$\|A\|_F = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2}$$

Rappels

Norme induite

Soit $\|\cdot\|$ *une norme vectorielle sur* $\mathbb{R}^n$*. La norme matricielle* $\|\cdot\|_{m\times n}$ *sur* $\mathbb{R}^{m\times n}$ *définie par*

$$\|A\|_{m\times n} = \max_{x\in\mathbb{R}^n, x\neq 0} \frac{\|Ax\|}{\|x\|}$$

est appelée norme induite *par la norme vectorielle.*

Rappels

Normes matricielles *Les normes matricielles* $\|\cdot\|_2$ *et* $\|\cdot\|_F$ *vérifient les propriétés suivantes :*

1. *Soient* A, $B \in \mathbb{R}^{n\times n}$. *Alors* $\|AB\|_F \leq \|A\|_F\|B\|_F$
2. *Soient* A, $B \in \mathbb{R}^{n\times n}$. *Alors* $\|AB\|_2 \leq \|A\|_2\|B\|_2$
3. *Soient* A, $B \in \mathbb{R}^{n\times n}$. *Alors* $\|AB\|_F \leq \min\left(\|A\|_2\|B\|_F, \|A\|_F\|B\|_2\right)$
4. *Soient* $A \in \mathbb{R}^{n\times n}$ *et* $x \in \mathbb{R}^n$. *Alors* $\|Ax\|_2 \leq \|A\|_F\|x\|_2$.
5. *Soient* $x, y \in \mathbb{R}^n$. *Alors* $\|xy^T\|_F = \|xy^T\|_2 = \|x\|_2\|y\|_2$.

(sans preuve)

Algorithme : Méthode sécante — n variables

Objectif

Trouver une approximation de la solution du système d'équations

$$f(x) = 0.$$

Input

- La fonction $f : \mathbb{R}^n \to \mathbb{R}^n$;
- Une première approximation de la solution $x_0 \in \mathbb{R}^n$;
- Une première approximation de la matrice jacobienne A_0 (par défaut $A_0 = I$);
- La précision demandée $\varepsilon \in \mathbb{R}$, $\varepsilon > 0$.

Algorithme : Méthode sécante — n variables

Output

Une approximation de la solution $x^* \in \mathbb{R}^n$

Initialisation

1. $x_1 = x_0 - A_0^{-1} f(x_0)$
2. $d_0 = x_0 - x_1$, $y_0 = f(x_0) - f(x_1)$
3. $k = 1$

Algorithme : Méthode sécante — n variables

Itérations

1. Mise à jour de Broyden :

$$A_k = A_{k-1} + \frac{(y_{k-1} - A_{k-1}d_{k-1})d_{k-1}^T}{d_{k-1}^T d_{k-1}}$$

2. Calculer d_k solution de $A_k d_k = -f(x_k)$,
3. Calculer $y_k = f(x_{k-1}) - f(x_k)$
4. $x_{k+1} = x_k + d_k$, $k = k + 1$.

Algorithme : Méthode sécante — n variables

Critère d'arrêt

Si $\|f(x_k)\| \leq \varepsilon$, alors $x^* = x_k$.

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Méthode sécante — n variables

$$\begin{aligned}(x_1+1)^2 + x_2^2 &= 2\\ e^{x_1} + x_2^3 &= 2.\end{aligned}$$

k	x_k	$f(x_k)$	$\|f(x_k)\|$
0	1.00000000e+00	3.00000000e+00	3.45723768e+00
	1.00000000e+00	1.71828182e+00	
1	-2.00000000e+00	-4.84071214e-01	2.28706231e+00
	-7.18281828e-01	-2.23524698e+00	
2	-1.66450025e+00	-8.68008706e-01	1.51117836e+00
	8.30921595e-01	-1.23702099e+00	
3	-2.42562564e-01	2.72598221e+00	7.74156513e+00
	2.03771213e+00	7.24574714e+00	
4	-1.24155582e+00	-1.34676047e+00	1.83898030e+00
	7.71291329e-01	-1.25223192e+00	

Méthode sécante — n variables

k	x_k	$f(x_k)$	$\|f(x_k)\|$
15	-1.71374738e+00	-1.15696833e-07	1.85422885e-07
	1.22088678e+00	-1.44899582e-07	
16	-1.71374741e+00	-2.43091768e-10	3.89249065e-10
	1.22088682e+00	-3.04008596e-10	
17	-1.71374741e+00	8.17124146e-14	1.30803685e-13
	1.22088682e+00	1.02140518e-13	
18	-1.71374741e+00	-2.22044604e-16	2.22044604e-16
	1.22088682e+00	0.00000000e+00	

Méthode sécante — n variables

$J(x_k)$		A_k	
00000000e+00	2.00000000e+00	1.00000000e+00	0.00000000e+00
71828182e+00	3.00000000e+00	0.00000000e+00	1.00000000e+00
00000000e+00	-1.43656365e+00	1.12149881e+00	6.95897342e-02
35335283e-01	1.54778635e+00	5.61032855e-01	1.32133752e+00
32900051e+00	1.66184319e+00	1.00559588e+00	-4.65603572e-01
39285227e-01	2.07129209e+00	3.95856681e-01	5.58620104e-01
51487487e+00	4.07542427e+00	2.12000014e+00	4.80185068e-01
34614657e-01	1.24568122e+01	3.35797853e+00	3.07255629e+00
33111643e-01	1.54258265e+00	2.63710364e+00	1.13571564e+00
38934337e-01	1.78467094e+00	3.83878676e+00	3.68207545e+00

Méthode sécante — n variables

$J(x_k)$		A_k	
42749482e+00	2.44177364e+00	-1.06423011e+00	2.70386672e+00
80189282e-01	4.47169389e+00	6.34731480e-01	4.79964153e+00
42749482e+00	2.44177364e+00	-1.06870996e+00	2.71006685e+00
80189282e-01	4.47169389e+00	6.34731480e-01	4.79964153e+00