Une introduction aux preuves assistées par ordinateur pour l'étude d'équations non linéaires

Comment transformer une simulation numérique en théorème?

Maxime Breden

Séminaire du CMAP, 1er octobre 2019

Preuves assistées par ordinateur

- Objectifs : démontrer des théorèmes (existence de solutions + propriétés quantitatives) à l'aide de simulations numériques.
  - Typiquement, on s'intéresse aux états stationnaires, aux orbites périodiques, aux variétés invariantes, aux connexions homoclines/hétéroclines entre ces différentes solutions...

- Objectifs : démontrer des théorèmes (existence de solutions + propriétés quantitatives) à l'aide de simulations numériques.
  - Typiquement, on s'intéresse aux états stationnaires, aux orbites périodiques, aux variétés invariantes, aux connexions homoclines/hétéroclines entre ces différentes solutions...

- Motivations : Étudier des systèmes non linéaires hors d'un cadre perturbatif, dans des situations où de nombreuses solutions différentes peuvent co-exister.

Preuves assistées par ordinateur

- Objectifs : démontrer des théorèmes (existence de solutions + propriétés quantitatives) à l'aide de simulations numériques.
  - Typiquement, on s'intéresse aux états stationnaires, aux orbites périodiques, aux variétés invariantes, aux connexions homoclines/hétéroclines entre ces différentes solutions...

- Motivations : Étudier des systèmes non linéaires hors d'un cadre perturbatif, dans des situations où de nombreuses solutions différentes peuvent co-exister.
  - **D'un point de vue théorique :** ces méthodes peuvent permettre de démontrer certains théorèmes qu'on n'est pas capable de prouver de manière classique.

Preuves assistées par ordinateur

- Objectifs : démontrer des théorèmes (existence de solutions + propriétés quantitatives) à l'aide de simulations numériques.
  - Typiquement, on s'intéresse aux états stationnaires, aux orbites périodiques, aux variétés invariantes, aux connexions homoclines/hétéroclines entre ces différentes solutions...

- Motivations : Étudier des systèmes non linéaires hors d'un cadre perturbatif, dans des situations où de nombreuses solutions différentes peuvent co-exister.
  - **D'un point de vue théorique :** ces méthodes peuvent permettre de démontrer certains théorèmes qu'on n'est pas capable de prouver de manière classique.
  - **Pour les applications :** ces méthodes permettent de certifier les résultats de simulations numériques, et fournissent plus de garanties que les estimations d'erreur a posteriori traditionnelles.

Sommaire

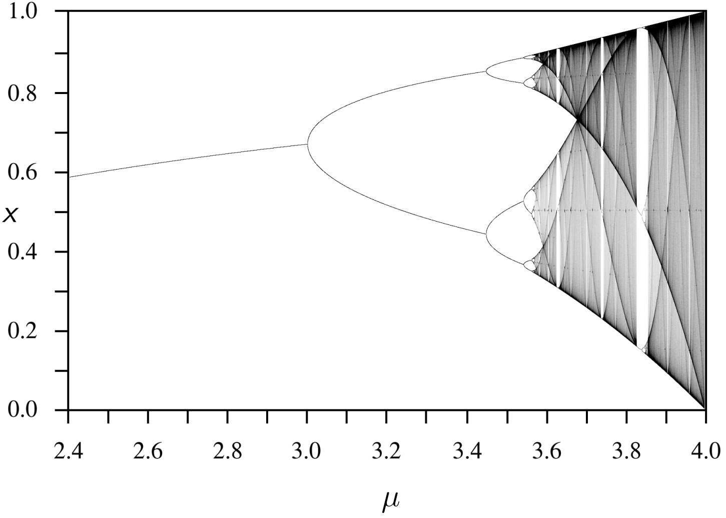
Sommaire

Motivation : orbites périodiques et chaos

Considérons la suite logistique : $x_{n+1} = \mu x_n(1 - x_n)$.

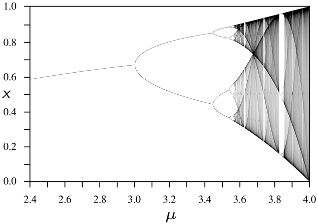
Motivation : orbites périodiques et chaos

Considérons la suite logistique : $x_{n+1} = \mu x_n(1 - x_n)$.



Motivation : orbites périodiques et chaos

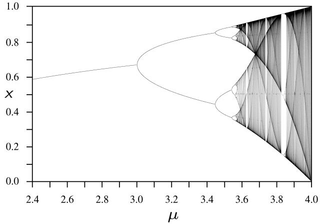
Considérons la suite logistique : $x_{n+1} = \mu x_n(1 - x_n)$.



- Comment caractériser la dynamique de plus en plus complexe qu'on observe pour les valeurs de μ proches de 4 ? Notion de *chaos*.

Motivation : orbites périodiques et chaos

Considérons la suite logistique : $x_{n+1} = \mu x_n(1 - x_n)$.



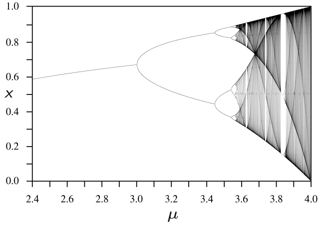
- Comment caractériser la dynamique de plus en plus complexe qu'on observe pour les valeurs de μ proches de 4 ? Notion de *chaos*.

Théorème [Sharkovsky '64, Li-York '75]

"L'existence d'une orbite de période 3 implique le chaos"

Motivation : orbites périodiques et chaos

Considérons la suite logistique : $x_{n+1} = \mu x_n(1 - x_n)$.



- Comment caractériser la dynamique de plus en plus complexe qu'on observe pour les valeurs de μ proches de 4 ? Notion de *chaos*.

Théorème [Sharkovsky '64, Li-York '75]

"L'existence d'une orbite de période 3 implique le chaos"

- Pour une valeur du paramètre μ donnée, comment démontrer l'existence d'une telle orbite, afin d'appliquer le théorème ci-dessus ?

A la recherche d'une orbite de période 3

$$x_{n+1} = \mu x_n(1 - x_n)$$

A la recherche d'une orbite de période 3

$$x_{n+1} = \mu x_n(1 - x_n)$$

- On commencer par chercher numériquement une orbite de période 3.

A la recherche d'une orbite de période 3

$$x_{n+1} = \mu x_n(1 - x_n)$$

- On commencer par chercher numériquement une orbite de période 3.
- Pour ce faire, on peut considérer $F : \mathbb{R}^3 \to \mathbb{R}^3$ définie par

$$F(x_0, x_1, x_2) = \begin{pmatrix} \mu x_0(1 - x_0) - x_1 \\ \mu x_1(1 - x_1) - x_2 \\ \mu x_2(1 - x_2) - x_0 \end{pmatrix}.$$

Si on trouve un zéro de F (tel que $x_0 \neq x_1 \neq x_2$), on a bien une orbite de période 3.

A la recherche d'une orbite de période 3

$$x_{n+1} = \mu x_n(1 - x_n)$$

- On commencer par chercher numériquement une orbite de période 3.
- Pour ce faire, on peut considérer $F : \mathbb{R}^3 \to \mathbb{R}^3$ définie par

$$F(x_0, x_1, x_2) = \begin{pmatrix} \mu x_0(1 - x_0) - x_1 \\ \mu x_1(1 - x_1) - x_2 \\ \mu x_2(1 - x_2) - x_0 \end{pmatrix}.$$

  Si on trouve un zéro de F (tel que $x_0 \neq x_1 \neq x_2$), on a bien une orbite de période 3.
- Numériquement, on trouve facilement une "solution numérique" $\bar{X} = (\bar{x}_0, \bar{x}_1, \bar{x}_2)$ telle que $F(\bar{X}) \approx 0$.

A la recherche d'une orbite de période 3

$$x_{n+1} = \mu x_n(1 - x_n)$$

- On commencer par chercher numériquement une orbite de période 3.
- Pour ce faire, on peut considérer $F : \mathbb{R}^3 \to \mathbb{R}^3$ définie par

$$F(x_0, x_1, x_2) = \begin{pmatrix} \mu x_0(1 - x_0) - x_1 \\ \mu x_1(1 - x_1) - x_2 \\ \mu x_2(1 - x_2) - x_0 \end{pmatrix}.$$

  Si on trouve un zéro de F (tel que $x_0 \neq x_1 \neq x_2$), on a bien une orbite de période 3.
- Numériquement, on trouve facilement une "solution numérique" $\bar{X} = (\bar{x}_0, \bar{x}_1, \bar{x}_2)$ telle que $F(\bar{X}) \approx 0$.
- Comment démonter rigoureusement l'existence d'un tel zéro ?

On veut des preuves !

$$F : \mathbb{R}^3 \to \mathbb{R}^3 \qquad F(\bar{X}) \approx 0.$$

On veut des preuves !

$$F : \mathbb{R}^3 \to \mathbb{R}^3 \qquad F(\bar{X}) \approx 0.$$

- On veut démontrer **a posteriori** qu'il existe un zéro de F proche de $\bar{X}$.

On veut des preuves !

$$F : \mathbb{R}^3 \to \mathbb{R}^3 \qquad F(\bar{X}) \approx 0.$$

- On veut démontrer **a posteriori** qu'il existe un zéro de F proche de $\bar{X}$.

Théorème *à la Newton-Kantorovich*

Soient $\varepsilon, K, L > 0$ tels que

$$\begin{aligned} \|F(\bar{X})\| &\leq \varepsilon \\ \|DF(\bar{X})^{-1}\| &\leq K \\ \|DF(X) - DF(\bar{X})\| &\leq L\|X - \bar{X}\| \qquad \forall\, X \in \mathbb{R}^3. \end{aligned}$$

S'il existe $r > 0$ tel que

$$KLr^2 - r + K\varepsilon < 0,$$

alors F possède un unique zéro X^* satisfaisant $\|X^* - \bar{X}\| \leq r$.

On veut des preuves !

$$F : \mathbb{R}^3 \to \mathbb{R}^3 \qquad F(\bar{X}) \approx 0.$$

- On veut démontrer **a posteriori** qu'il existe un zéro de F proche de $\bar{X}$.

Théorème *à la Newton-Kantorovich*

Soient $\varepsilon, K, L > 0$ tels que

$$\begin{aligned} \|F(\bar{X})\| &\leq \varepsilon \\ \|DF(\bar{X})^{-1}\| &\leq K \\ \|DF(X) - DF(\bar{X})\| &\leq L\|X - \bar{X}\| \qquad \forall\, X \in \mathbb{R}^3. \end{aligned}$$

S'il existe $r > 0$ tel que

$$KLr^2 - r + K\varepsilon < 0,$$

alors F possède un unique zéro X^* satisfaisant $\|X^* - \bar{X}\| \leq r$.

Démonstration : $T : X \mapsto X - DF(\bar{X})^{-1}F(X)$ est une contraction sur la boule fermée de centre $\bar{X}$ et de rayon r.

Un exemple effrayant

- Peut-on vraiment faire confiance aux calculs en flottants ?

Un exemple effrayant

- Peut-on vraiment faire confiance aux calculs en flottants ?
- Considérons l'exemple suivant [Rump '94]

$$g(a,b) = 333.75b^6 + a^2(11a^2b^2 - b^6 - 121b^4 - 2) + 5.5b^8 + \frac{a}{2b},$$

évalué avec $a = 77617$ et $b = 33096$, pour différentes précisions.

Un exemple effrayant

- Peut-on vraiment faire confiance aux calculs en flottants ?
- Considérons l'exemple suivant [Rump '94]

$$g(a,b) = 333.75b^6 + a^2(11a^2b^2 - b^6 - 121b^4 - 2) + 5.5b^8 + \frac{a}{2b},$$

  évalué avec $a = 77617$ et $b = 33096$, pour différentes précisions.

- Il faut se méfier des erreurs d'arrondis, surtout si on prétend démontrer un théorème en utilisant des calculs en flottants !

Un exemple effrayant

- Peut-on vraiment faire confiance aux calculs en flottants ?
- Considérons l'exemple suivant [Rump '94]

$$g(a,b) = 333.75b^6 + a^2(11a^2b^2 - b^6 - 121b^4 - 2) + 5.5b^8 + \frac{a}{2b},$$

évalué avec $a = 77617$ et $b = 33096$, pour différentes précisions.

- Il faut se méfier des erreurs d'arrondis, surtout si on prétend démontrer un théorème en utilisant des calculs en flottants !
- Dans notre "preuve" d'existence d'une orbite de période 3, comment peut-on être certain que la quantité ε qu'on a calculée numériquement majore bien $\|F(\bar{X})\|$, ou qu'on a bien $KLr^2 - r + K\varepsilon < 0$?

L'arithmétique d'intervalles à la rescousse

L'arithmétique d'intervalles à la rescousse

- Soit $\mathbb{F}$ l'ensemble des nombres flottants, c'est à dire l'ensemble (fini !) de nombres que peut représenter l'ordinateur pour une précision donnée, et

$$\bigtriangledown, \bigtriangleup : \mathbb{R} \to \mathbb{F},$$

les opérations d'arrondi par défaut et par excès.

L'arithmétique d'intervalles à la rescousse

- Soit $\mathbb{F}$ l'ensemble des nombres flottants, c'est à dire l'ensemble (fini !) de nombres que peut représenter l'ordinateur pour une précision donnée, et
$$\bigtriangledown, \bigtriangleup : \mathbb{R} \to \mathbb{F},$$
les opérations d'arrondi par défaut et par excès.
- Exemple : on considère $x = 0.1$. En base 2, x s'écrit
$$x = (1.1001100110011001100...)_2 \times 2^{-4}.$$
Si on utilise 8 bits de précision, on a
$$\bigtriangledown(x) = (1.1001100)_2 \times 2^{-4} \quad \text{et} \bigtriangleup(x) = (1.1001101)_2 \times 2^{-4}.$$

L'arithmétique d'intervalles à la rescousse

- Soit $\mathbb{F}$ l'ensemble des nombres flottants, c'est à dire l'ensemble (fini !) de nombres que peut représenter l'ordinateur pour une précision donnée, et
$$\triangledown, \triangle : \mathbb{R} \to \mathbb{F},$$
les opérations d'arrondi par défaut et par excès.
- A la place d'utiliser un flottant, on représente désormais chaque nombre réel par un intervalle qui le contient :
$$x \in \mathbb{R} \quad \to \quad [x] := [\triangledown(x)\,,\, \triangle(x)]\,.$$

- Soit $\mathbb{F}$ l'ensemble des nombres flottants, c'est à dire l'ensemble (fini !) de nombres que peut représenter l'ordinateur pour une précision donnée, et
$$\triangledown, \triangle : \mathbb{R} \to \mathbb{F},$$
les opérations d'arrondi par défaut et par excès.
- A la place d'utiliser un flottant, on représente désormais chaque nombre réel par un intervalle qui le contient :
$$x \in \mathbb{R} \quad \to \quad [x] := [\triangledown(x)\, ,\, \triangle(x)]\,.$$
- On peut redéfinir les opérations élémentaires $(+, -, \times, \div)$, de telle sorte que le résultat contienne toujours la vraie valeur :
$$x + y \quad \to \quad [x]\ [+]\ [y],$$
où l'opération $[+]$ est définie de la manière suivante
$$[x]\ [+]\ [y] := [\triangledown\left(\triangledown(x) + \triangledown(y)\right)\, ,\, \triangle\left(\triangle(x) + \triangle(y)\right)]\,.$$

L'arithmétique d'intervalles à la rescousse

- Soit $\mathbb{F}$ l'ensemble des nombres flottants, c'est à dire l'ensemble (fini !) de nombres que peut représenter l'ordinateur pour une précision donnée, et
$$\triangledown, \vartriangle : \mathbb{R} \to \mathbb{F},$$
les opérations d'arrondi par défaut et par excès.
- A la place d'utiliser un flottant, on représente désormais chaque nombre réel par un intervalle qui le contient :
$$x \in \mathbb{R} \quad \to \quad [x] := [\triangledown(x)\,,\, \vartriangle(x)]\,.$$
- On peut redéfinir les opérations élémentaires $(+, -, \times, \div)$, de telle sorte que le résultat contienne toujours la vraie valeur :
$$x + y \quad \to \quad [x]\ [+]\ [y],$$
où l'opération $[+]$ est définie de la manière suivante
$$[x]\ [+]\ [y] := [\triangledown\left(\triangledown(x) + \triangledown(y)\right)\,,\, \vartriangle\left(\vartriangle(x) + \vartriangle(y)\right)]\,.$$
- On a alors $\boldsymbol{x} + \boldsymbol{y} \in [\boldsymbol{x}]\ [+]\ [\boldsymbol{y}]$.

- L'arithmétique d'intervalles permet donc de faire des calculs numériques certifiés.

L'arithmétique d'intervalles à la rescousse

- L'arithmétique d'intervalles permet donc de faire des calculs numériques certifiés.
- Ses deux inconvénients principaux sont :
  - l'augmentation du cout de calcul,
  - une possible surestimation des erreurs (*wrapping effect*), notamment si on accumule trop d'opérations.

L'arithmétique d'intervalles à la rescousse

- L'arithmétique d'intervalles permet donc de faire des calculs numériques certifiés.
- Ses deux inconvénients principaux sont :
  - l'augmentation du cout de calcul,
  - une possible surestimation des erreurs (*wrapping effect*), notamment si on accumule trop d'opérations.
- Utilisée de manière judicieuse, par exemple pour des estimations d'erreur *a posteriori*, l'arithmétique d'intervalles peut cependant être très utile !

Une preuve assistée par ordinateur de chaos, résumé

$$x_{n+1} = \mu x_n(1 - x_n)$$

Une preuve assistée par ordinateur de chaos, résumé

$$x_{n+1} = \mu x_n(1 - x_n)$$

1. On reformule la recherche d'une orbite de période 3 sous la forme d'un problème $\boldsymbol{F}(\boldsymbol{X}) = \mathbf{0}$.

Une preuve assistée par ordinateur de chaos, résumé

$$x_{n+1} = \mu x_n(1 - x_n)$$

1. On reformule la recherche d'une orbite de période 3 sous la forme d'un problème $\boldsymbol{F}(\boldsymbol{X}) = \mathbf{0}$.

2. On trouve numériquement une **solution approchée**.

Une preuve assistée par ordinateur de chaos, résumé

$$x_{n+1} = \mu x_n(1 - x_n)$$

1. On reformule la recherche d'une orbite de période 3 sous la forme d'un problème $\boldsymbol{F}(\boldsymbol{X}) = \boldsymbol{0}$.

2. On trouve numériquement une **solution approchée**.

3. On **estime a posteriori**

$$\|F(\bar{X})\|, \quad \|DF(\bar{X})^{-1}\| \quad \text{et} \quad \|D^2F(X)\|,$$

   à l'aide de l'**arithmétique d'intervalles**.

Une preuve assistée par ordinateur de chaos, résumé

$$x_{n+1} = \mu x_n(1 - x_n)$$

1. On reformule la recherche d'une orbite de période 3 sous la forme d'un problème $\boldsymbol{F}(\boldsymbol{X}) = \mathbf{0}$.

2. On trouve numériquement une **solution approchée**.

3. On **estime a posteriori**

$$\|F(\bar{X})\|, \quad \|DF(\bar{X})^{-1}\| \quad \text{et} \quad \|D^2F(X)\|,$$

   à l'aide de l'**arithmétique d'intervalles**.

4. On cherche à en déduire que l'opérateur

$$\boldsymbol{T} : \boldsymbol{X} \mapsto \boldsymbol{X} - \boldsymbol{DF}(\bar{\boldsymbol{X}})^{-1}\boldsymbol{F}(\boldsymbol{X})$$

   est contractant sur un petit voisinage de $\bar{X}$.