Introduction à l'Analyse Numérique

Quentin Louveaux

ULg - Institut Montefiore

2011

Organisation du cours

Cours théorique

- Cours théorique tous les jeudis de 8h30 à 10h30 au 303/B7a
- Pas de cours le 13 octobre et le 3 novembre
- Dernier cours le 8 décembre

Répétitions

- 7 jeudis de 10h40 à 12h30
- Répartition en 7 groupes
  Répartition (purement alphabétique) donnée le jeudi 29 septembre.
  Veuillez respecter la répartition !
- Dans ce cours, pas de projet matlab vu la réduction des heures consacrées aux travaux pratiques.

Site web du cours

`http ://www.montefiore.ulg.ac.be/˜louveaux/analnum.html`
Contient les transparents du cours, les énoncés des répétitions et tous les autres renseignements utiles (examens des années précédentes, ...)

Organisation du cours

Cours théorique

- Cours théorique tous les jeudis de 8h30 à 10h30 au 303/B7a
- Pas de cours le 13 octobre et le 3 novembre
- Dernier cours le 8 décembre

Répétitions

- 7 jeudis de 10h40 à 12h30
- Répartition en 7 groupes
  Répartition (purement alphabétique) donnée le jeudi 29 septembre.
  Veuillez respecter la répartition !
- Dans ce cours, pas de projet matlab vu la réduction des heures consacrées aux travaux pratiques.

Site web du cours

`http ://www.montefiore.ulg.ac.be/˜louveaux/analnum.html`
Contient les transparents du cours, les énoncés des répétitions et tous les autres renseignements utiles (examens des années précédentes, . . .)

Organisation du cours

Cours théorique

- Cours théorique tous les jeudis de 8h30 à 10h30 au 303/B7a
- Pas de cours le 13 octobre et le 3 novembre
- Dernier cours le 8 décembre

Répétitions

- 7 jeudis de 10h40 à 12h30
- Répartition en 7 groupes
  Répartition (purement alphabétique) donnée le jeudi 29 septembre.
  Veuillez respecter la répartition !
- Dans ce cours, pas de projet matlab vu la réduction des heures consacrées aux travaux pratiques.

Site web du cours

`http ://www.montefiore.ulg.ac.be/~louveaux/analnum.html`
Contient les transparents du cours, les énoncés des répétitions et tous les autres renseignements utiles (examens des années précédentes, . . .)

Horaire

Date	Cours	Répétition
15 septembre	C1	×
22 septembre	C2	×
29 septembre	C3	R1 : Interpolation, Taylor et régressions
6 octobre	C4	R2 : Dérivation numérique
13 octobre	×	×
20 octobre	C5	R3 : Intégration numérique
27 octobre	C6	R4 : Algèbre linéaire : méthodes directes
3 novembre	×	×
10 novembre	C7	×
17 novembre	C8	R5 : Méthodes itératives et valeurs propres
24 novembre	C9	R6 : Optimisation linéaire
1 décembre	C10	R7 : Systèmes non linéaires
8 décembre	C11	×

Rappel : principe d'une méthode numérique

L'analyse numérique s'occupe principalement de deux aspects primordiaux

Trouver la solution de problèmes réels dont la solution analytique n'est pas conue

- En approximant les phénomènes, les équations, ...
- Souvent par des méthodes itératives qui s'approchent de plus en plus de la solution
- En approximant la solution...

Analyse du comportement des méthodes

- Efficacité, complexité
- Ordre de convergence
- Robustesse, sensibilité aux erreurs d'arrondi

Rappel : principe d'une méthode numérique

L'analyse numérique s'occupe principalement de deux aspects primordiaux

Trouver la solution de problèmes réels dont la solution analytique n'est pas conue

- En approximant les phénomènes, les équations, ...
- Souvent par des méthodes itératives qui s'approchent de plus en plus de la solution
- En approximant la solution...

Analyse du comportement des méthodes

- Efficacité, complexité
- Ordre de convergence
- Robustesse, sensibilité aux erreurs d'arrondi

Rappel : principe d'une méthode numérique

L'analyse numérique s'occupe principalement de deux aspects primordiaux

Trouver la solution de problèmes réels dont la solution analytique n'est pas conue

- En approximant les phénomènes, les équations, ...
- Souvent par des méthodes itératives qui s'approchent de plus en plus de la solution
- En approximant la solution...

Analyse du comportement des méthodes

- Efficacité, complexité
- Ordre de convergence
- Robustesse, sensibilité aux erreurs d'arrondi

Calcul du zéro d'une fonction sans forme analytique

Cas du projet de 1e bachelier ingénieur 2011
Contexte :

On shoote dans un ballon avec une certaine force et un certain angle.
Trouver le bon angle et la bonne force pour atteindre une cible.
Question Pour une force fixée, quel est l'angle qui permet d'atteindre la cible ?

Calcul du zéro d'une fonction sans forme analytique

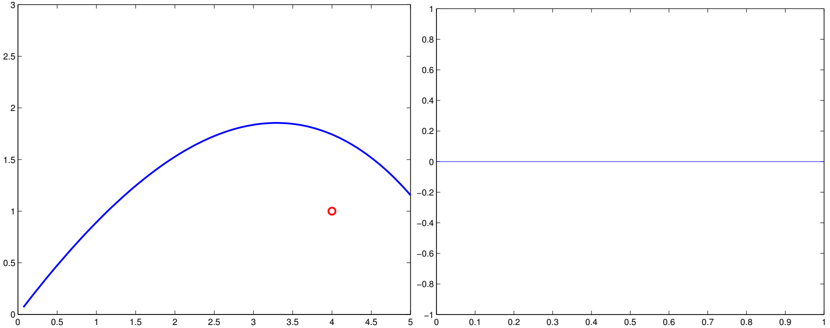
Cas du projet de 1e bachelier ingénieur 2011
Contexte :

On shoote dans un ballon avec une certaine force et un certain angle.
Trouver le bon angle et la bonne force pour atteindre une cible.
Question Pour une force fixée, quel est l'angle qui permet d'atteindre la cible ?

Calcul du zéro d'une fonction sans forme analytique

Cas du projet de 1e bachelier ingénieur 2011
Contexte :

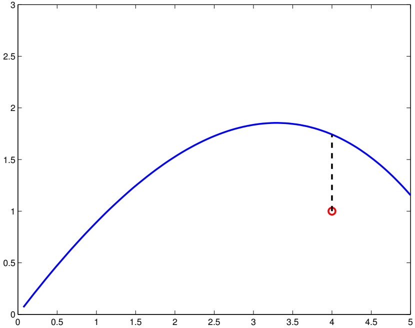
On shoote dans un ballon avec une certaine force et un certain angle.
Trouver le bon angle et la bonne force pour atteindre une cible.
Question Pour une force fixée, quel est l'angle qui permet d'atteindre la cible ?

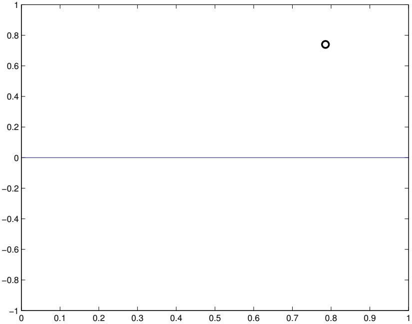


Calcul du zéro d'une fonction sans forme analytique

Cas du projet de 1e bachelier ingénieur 2011
Contexte :

On shoote dans un ballon avec une certaine force et un certain angle.
Trouver le bon angle et la bonne force pour atteindre une cible.
Question Pour une force fixée, quel est l'angle qui permet d'atteindre la cible ?

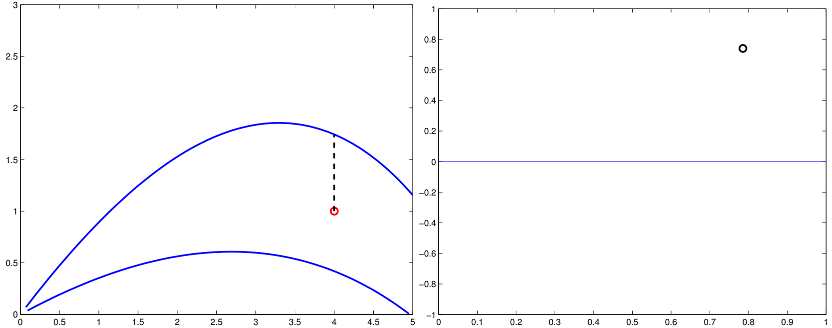




Calcul du zéro d'une fonction sans forme analytique

Cas du projet de 1e bachelier ingénieur 2011
Contexte :

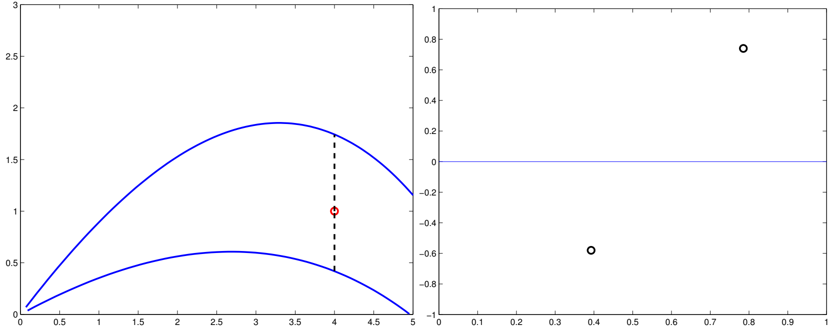
On shoote dans un ballon avec une certaine force et un certain angle.
Trouver le bon angle et la bonne force pour atteindre une cible.
Question Pour une force fixée, quel est l'angle qui permet d'atteindre la cible ?



Calcul du zéro d'une fonction sans forme analytique

Cas du projet de 1e bachelier ingénieur 2011
Contexte :

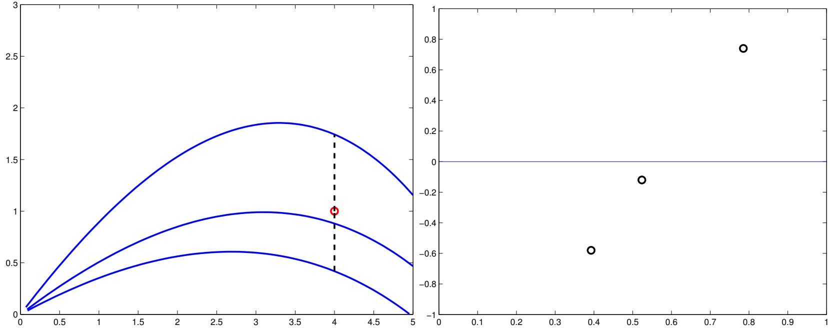
On shoote dans un ballon avec une certaine force et un certain angle.
Trouver le bon angle et la bonne force pour atteindre une cible.
Question Pour une force fixée, quel est l'angle qui permet d'atteindre la cible ?



Calcul du zéro d'une fonction sans forme analytique

Cas du projet de 1e bachelier ingénieur 2011
Contexte :

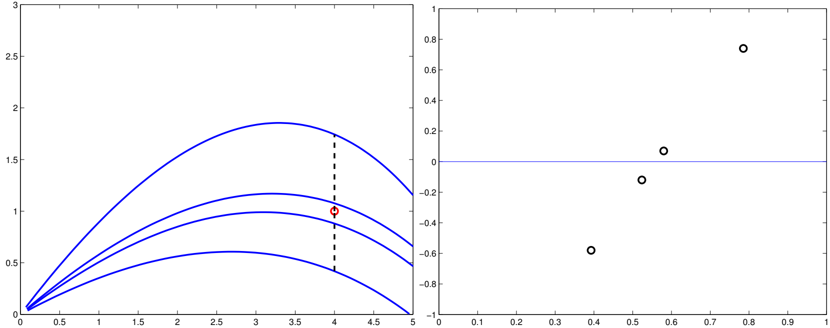
On shoote dans un ballon avec une certaine force et un certain angle.
Trouver le bon angle et la bonne force pour atteindre une cible.
Question Pour une force fixée, quel est l'angle qui permet d'atteindre la cible ?



Calcul du zéro d'une fonction sans forme analytique

Cas du projet de 1e bachelier ingénieur 2011
Contexte :

On shoote dans un ballon avec une certaine force et un certain angle.
Trouver le bon angle et la bonne force pour atteindre une cible.
Question Pour une force fixée, quel est l'angle qui permet d'atteindre la cible ?



Rappel : ordre de convergence

- Mesure très importante de la qualité de convergence d'une méthode itérative
- Plus l'ordre est élevé, plus la méthode converge vite

Définition

Soit (x_n) une suite telle que $\lim_{n\to\infty} x_n = \bar{x}$, si il existe $c_0, p > 0$ et $0 < c < 1$ tels que

$$|x_n - \bar{x}| \leq c_0 c^{pn}$$

alors la suite (x_n) converge avec un ordre p

Importance de l'ordre de convergence

Cas du projet ingénieur 2011 :
Question 2 : Trouver la force initiale minimale qui permette d'atteindre la cible

Proposition d'un algorithme

- Partir d'une force initiale
- Utiliser la méthode précédente pour trouver un angle qui atteigne la cible
- Diminuer ou augmenter la force en fonction du résultat (cible atteinte ou non)
- Itérer

$\longrightarrow$ On utilise un grand nombre de fois l'algorithme de recherche de racine

Importance de l'ordre de convergence

Cas du projet ingénieur 2011 :
Question 2 : Trouver la force initiale minimale qui permette d'atteindre la cible

Proposition d'un algorithme

- Partir d'une force initiale
- Utiliser la méthode précédente pour trouver un angle qui atteigne la cible
- Diminuer ou augmenter la force en fonction du résultat (cible atteinte ou non)
- Itérer

$\longrightarrow$ On utilise un grand nombre de fois l'algorithme de recherche de racine

Importance de l'ordre de convergence

Cas du projet ingénieur 2011 :
Question 2 : Trouver la force initiale minimale qui permette d'atteindre la cible

Proposition d'un algorithme

- Partir d'une force initiale
- Utiliser la méthode précédente pour trouver un angle qui atteigne la cible
- Diminuer ou augmenter la force en fonction du résultat (cible atteinte ou non)
- Itérer

$\longrightarrow$ On utilise un grand nombre de fois l'algorithme de recherche de racine

Importance de l'ordre de convergence

Cas du projet ingénieur 2011 :
Question 2 : Trouver la force initiale minimale qui permette d'atteindre la cible

Proposition d'un algorithme

- Partir d'une force initiale
- Utiliser la méthode précédente pour trouver un angle qui atteigne la cible
- Diminuer ou augmenter la force en fonction du résultat (cible atteinte ou non)
- Itérer

$\longrightarrow$ On utilise un grand nombre de fois l'algorithme de recherche de racine

Différents ordres de convergence

Comparons trois méthodes de recherche de racine !

Méthode ordre 1	Méthode ordre 1.618	Méthode ordre 2
$e_n := \lvert x_n - \bar{x} \rvert$	$e_n := \lvert x_n - \bar{x} \rvert$	$e_n := \lvert x_n - \bar{x} \rvert$
0.3176721962	0.3176721962	0.3176721962
0.1823278038	1.3176721962	0.0676721962
0.1176721962	0.1926721962	0.0037187078
0.0725717063	0.1201669987	0.0000117788
0.0466401060	0.0175578415	0.0000000001
0.0293280790	0.0017100613	0.0000000000
0.0187335691	0.0000254625	0.0000000000
0.0118560080	0.0000000372	0.0000000000
0.0075498284	0.0000000000	0.0000000000
0.0047894229	0.0000000000	0.0000000000

Conclusion

- Dans de nombreux cas, on utilisera une sous-routine à de nombreuses reprises
- Pour de telles sous-routines, il est très important de gagner un maximum de temps sur la complexité
- Il est important de savoir quelle est la précision requise
- D'autres sous-routines ne sont utilisées qu'une fois !
  Dans ce cas, gagner une seconde n'est peut-être pas si important !

Calcul du zéro d'une fonction sans forme analytique

Cas du projet de 1e bachelier ingénieur 2010

Contexte :

On dispose d'un modèle climatique d'une planète.
Le modèle dépend de : température, C_a, C_s, W_a, W_s.
Grâce à la résolution d'une EDO, on calcule l'évolution au cours du temps.
Question : comment faire varier un paramètre pour obtenir une températue d'équilibre entre 15 et 17 degrés ?

Calcul du zéro d'une fonction sans forme analytique

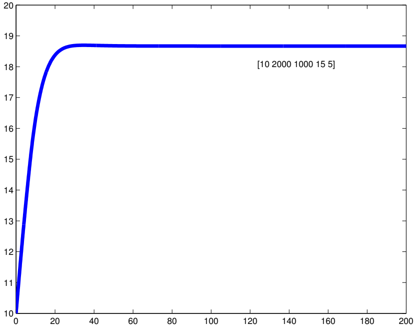
Cas du projet de 1e bachelier ingénieur 2010
Contexte :

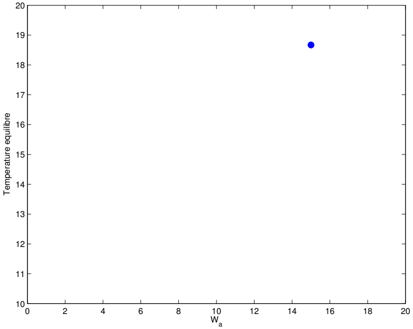
On dispose d'un modèle climatique d'une planète.
Le modèle dépend de : température, C_a, C_s, W_a, W_s.
Grâce à la résolution d'une EDO, on calcule l'évolution au cours du temps.
Question : comment faire varier un paramètre pour obtenir une températue d'équilibre entre 15 et 17 degrés ?

Calcul du zéro d'une fonction sans forme analytique

Cas du projet de 1e bachelier ingénieur 2010
Contexte :

On dispose d'un modèle climatique d'une planète.
Le modèle dépend de : température, C_a, C_s, W_a, W_s.
Grâce à la résolution d'une EDO, on calcule l'évolution au cours du temps.
Question : comment faire varier un paramètre pour obtenir une températue d'équilibre entre 15 et 17 degrés ?

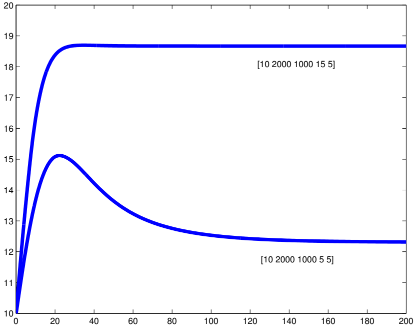


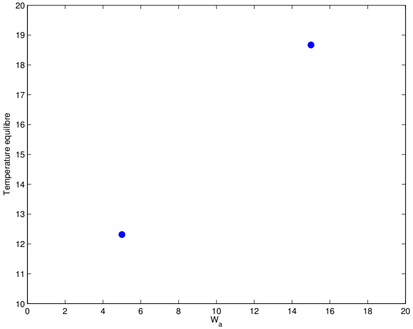


Calcul du zéro d'une fonction sans forme analytique

Cas du projet de 1e bachelier ingénieur 2010
Contexte :

On dispose d'un modèle climatique d'une planète.
Le modèle dépend de : température, C_a, C_s, W_a, W_s.
Grâce à la résolution d'une EDO, on calcule l'évolution au cours du temps.
Question : comment faire varier un paramètre pour obtenir une températue d'équilibre entre 15 et 17 degrés ?

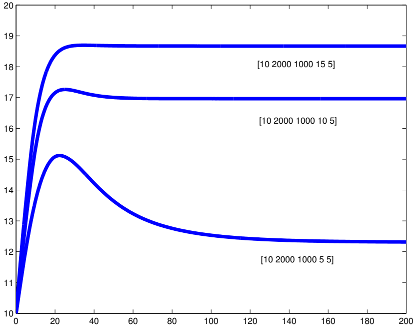


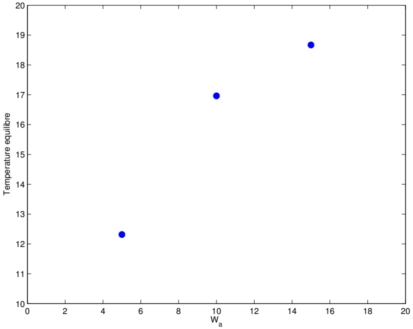


Calcul du zéro d'une fonction sans forme analytique

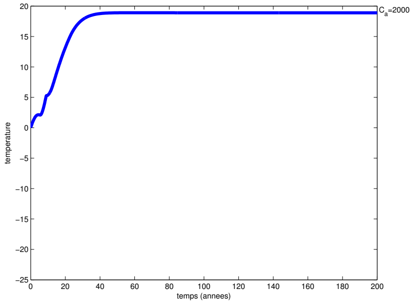
Cas du projet de 1e bachelier ingénieur 2010
Contexte :

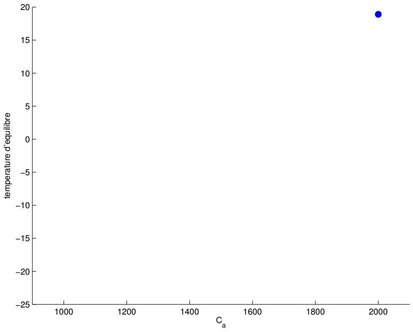
On dispose d'un modèle climatique d'une planète.
Le modèle dépend de : température, C_a, C_s, W_a, W_s.
Grâce à la résolution d'une EDO, on calcule l'évolution au cours du temps.
Question : comment faire varier un paramètre pour obtenir une températue d'équilibre entre 15 et 17 degrés ?



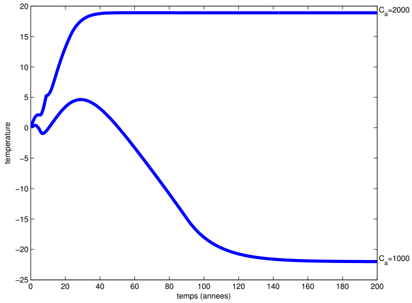


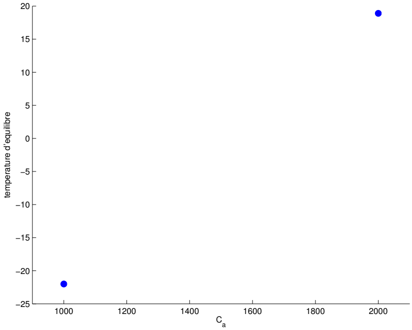
Des fonctions discontinues, ça existe !



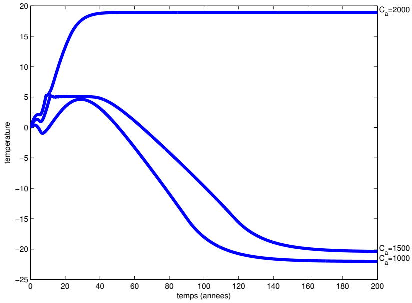


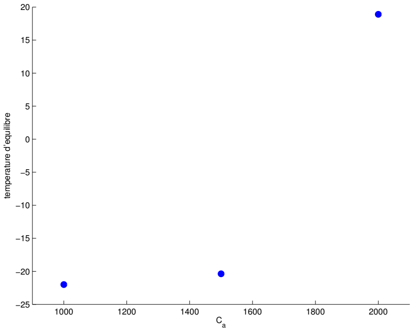
Des fonctions discontinues, ça existe !



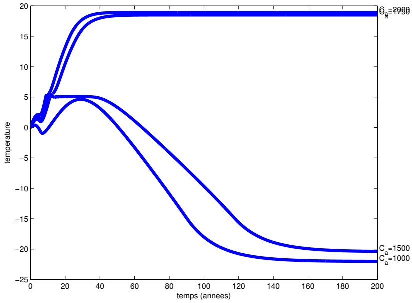


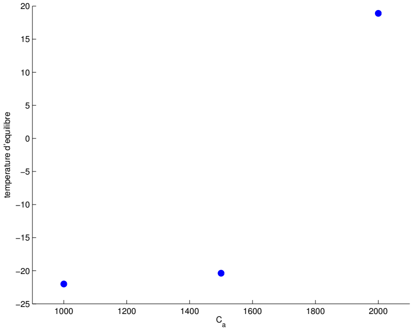
Des fonctions discontinues, ça existe !



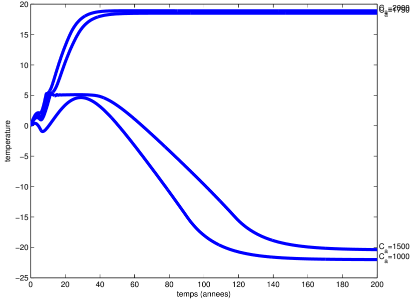


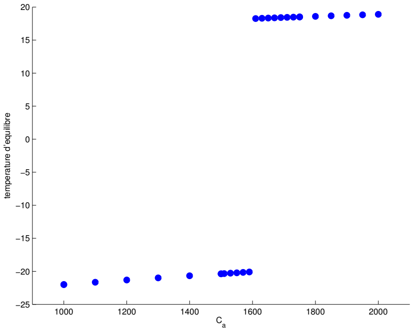
Des fonctions discontinues, ça existe !





Des fonctions discontinues, ça existe !





Autre rappel : le développement de Taylor

Théorème

Soit f, une fonction possédant ses $(n+1)$ premières dérivées continues sur un intervalle fermé $[a, b]$, alors pour chaque $c, x \in [a, b]$, f peut s'écrire comme

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(c)}{k!}(x-c)^k + E_{n+1}$$

où le terme d'erreur E_{n+1} peut s'écrire sous la forme

$$E_{n+1} = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-c)^{n+1},$$

et ξ est un point situé entre c et x.

Dernier rappel : l'interpolation polynomiale par la formule de Lagrange

$$
\begin{aligned}
l_i(x) &= \frac{(x-x_1)(x-x_2)\cdots(x-x_{i-1})(x-x_{i+1})\cdots(x-x_n)}{(x_i-x_1)(x_i-x_2)\cdots(x_i-x_{i-1})(x_i-x_{i+1})\cdots(x_i-x_n)} \\
&= \frac{\prod_{k\neq i}(x-x_k)}{\prod_{k\neq i}(x_i-x_k)}.
\end{aligned}
$$

$$
\begin{aligned}
l_i(x_i) &= 1 \\
l_i(x_k) &= 0 \quad \text{pour tout } k \neq i.
\end{aligned}
$$

Théorème

$$P(x) = \sum_{i=1}^{n} u(x_i) l_i(x)$$

Dernier rappel : l'interpolation polynomiale par la formule de Lagrange

$$
\begin{aligned}
l_i(x) &= \frac{(x-x_1)(x-x_2)\cdots(x-x_{i-1})(x-x_{i+1})\cdots(x-x_n)}{(x_i-x_1)(x_i-x_2)\cdots(x_i-x_{i-1})(x_i-x_{i+1})\cdots(x_i-x_n)} \\
&= \frac{\prod_{k\neq i}(x-x_k)}{\prod_{k\neq i}(x_i-x_k)}.
\end{aligned}
$$

$$
\begin{aligned}
l_i(x_i) &= 1 \\
l_i(x_k) &= 0 \quad \text{pour tout } k \neq i.
\end{aligned}
$$

Théorème

$$P(x) = \sum_{i=1}^{n} u(x_i) l_i(x)$$

Dernier rappel : l'interpolation polynomiale par la formule de Lagrange

$$l_i(x) = \frac{(x-x_1)(x-x_2)\cdots(x-x_{i-1})(x-x_{i+1})\cdots(x-x_n)}{(x_i-x_1)(x_i-x_2)\cdots(x_i-x_{i-1})(x_i-x_{i+1})\cdots(x_i-x_n)}$$
$$= \frac{\prod_{k\neq i}(x-x_k)}{\prod_{k\neq i}(x_i-x_k)}.$$

$$l_i(x_i) = 1$$
$$l_i(x_k) = 0 \quad \text{pour tout } k \neq i.$$

Théorème

$$P(x) = \sum_{i=1}^{n} u(x_i) l_i(x)$$

Formule de Lagrange

Exemple

Soit le polynôme $u(x) = x^2$ (“inconnu”) donné en $(0,0), (2,4)$ et $(3,9)$.

$$
\begin{aligned}
l_1(x) &= \frac{(x-2)(x-3)}{(0-2)(0-3)} && = \frac{1}{6}(x^2 - 5x + 6) \\
l_2(x) &= \frac{(x-0)(x-3)}{(2-0)(2-3)} && = -\frac{1}{2}(x^2 - 3x) \\
l_3(x) &= \frac{(x-0)(x-2)}{(3-0)(3-2)} && = \frac{1}{3}(x^2 - 2x)
\end{aligned}
$$

Le polynôme d'interpolation est

$$
\begin{aligned}
p(x) &= \frac{0}{6}(x^2 - 5x + 6) - \frac{4}{2}(x^2 - 3x) + \frac{9}{3}(x^2 - 2x) \\
&= x^2.
\end{aligned}
$$

Formule de Lagrange

Exemple

Soit le polynôme $u(x) = x^2$ ("inconnu") donné en $(0,0),(2,4)$ et $(3,9)$.

$$
\begin{aligned}
l_1(x) &= \frac{(x-2)(x-3)}{(0-2)(0-3)} &&= \frac{1}{6}(x^2-5x+6) \\
l_2(x) &= \frac{(x-0)(x-3)}{(2-0)(2-3)} &&= -\frac{1}{2}(x^2-3x) \\
l_3(x) &= \frac{(x-0)(x-2)}{(3-0)(3-2)} &&= \frac{1}{3}(x^2-2x)
\end{aligned}
$$

Le polynôme d'interpolation est

$$
\begin{aligned}
p(x) &= \frac{0}{6}(x^2-5x+6) - \frac{4}{2}(x^2-3x) + \frac{9}{3}(x^2-2x) \\
&= x^2.
\end{aligned}
$$

Formule de Lagrange

Exemple

Soit le polynôme $u(x) = x^2$ ("inconnu") donné en $(0,0), (2,4)$ et $(3,9)$.

$$\begin{aligned}
l_1(x) &= \frac{(x-2)(x-3)}{(0-2)(0-3)} &&= \frac{1}{6}(x^2-5x+6)\\
l_2(x) &= \frac{(x-0)(x-3)}{(2-0)(2-3)} &&= -\frac{1}{2}(x^2-3x)\\
l_3(x) &= \frac{(x-0)(x-2)}{(3-0)(3-2)} &&= \frac{1}{3}(x^2-2x)
\end{aligned}$$

Le polynôme d'interpolation est

$$\begin{aligned}
p(x) &= \frac{0}{6}(x^2-5x+6) - \frac{4}{2}(x^2-3x) + \frac{9}{3}(x^2-2x)\\
&= x^2.
\end{aligned}$$

Erreur d'interpolation

Quelques réflexions sur l'interpolation

- Si on approxime u par un polynôme, quelle est l'erreur ?
- Quel est le comportement général d'une interpolation ?

Théorème

Soit $u : \mathbb{R} \mapsto \mathbb{R}$ et $P(x)$ un polynôme interpolant les points $(x_1, u(x_1)), \ldots, (x_n, u(x_n))$. Si on suppose que $x_i \in [a, b]$ pour tout i, et que $u(x)$ est n fois continûment dérivable sur $[a, b]$, alors pour tout $x \in [a, b]$, il existe $\xi \in [a, b]$ tel que

$$e(x) = u(x) - P(x) = \frac{u^{(n)}(\xi)}{n!}(x - x_1) \cdots (x - x_n).$$

Erreur d'interpolation

Quelques réflexions sur l'interpolation

- Si on approxime u par un polynôme, quelle est l'erreur ?
- Quel est le comportement général d'une interpolation ?

Théorème

Soit $u : \mathbb{R} \mapsto \mathbb{R}$ et $P(x)$ un polynôme interpolant les points $(x_1, u(x_1)), \ldots, (x_n, u(x_n))$. Si on suppose que $x_i \in [a, b]$ pour tout i, et que $u(x)$ est n fois continûment dérivable sur $[a, b]$, alors pour tout $x \in [a, b]$, il existe $\xi \in [a, b]$ tel que

$$e(x) = u(x) - P(x) = \frac{u^{(n)}(\xi)}{n!}(x - x_1) \cdots (x - x_n).$$

Erreur d'interpolation

Quelques réflexions sur l'interpolation

- Si on approxime u par un polynôme, quelle est l'erreur ?
- Quel est le comportement général d'une interpolation ?

Théorème

Soit $u : \mathbb{R} \mapsto \mathbb{R}$ et $P(x)$ un polynôme interpolant les points $(x_1, u(x_1)), \ldots, (x_n, u(x_n))$. Si on suppose que $x_i \in [a, b]$ pour tout i, et que $u(x)$ est n fois continûment dérivable sur $[a, b]$, alors pour tout $x \in [a, b]$, il existe $\xi \in [a, b]$ tel que

$$e(x) = u(x) - P(x) = \frac{u^{(n)}(\xi)}{n!}(x - x_1) \cdots (x - x_n).$$

Erreur d'interpolation

Quelques réflexions sur l'interpolation

- Si on approxime u par un polynôme, quelle est l'erreur ?
- Quel est le comportement général d'une interpolation ?

Théorème

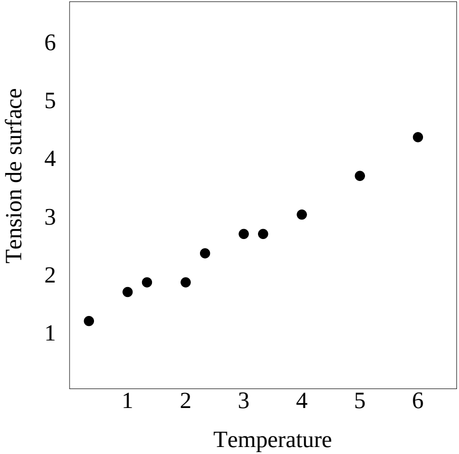
Soit $u : \mathbb{R} \mapsto \mathbb{R}$ et $P(x)$ un polynôme interpolant les points $(x_1, u(x_1)), \ldots, (x_n, u(x_n))$. Si on suppose que $x_i \in [a, b]$ pour tout i, et que $u(x)$ est n fois continûment dérivable sur $[a, b]$, alors pour tout $x \in [a, b]$, il existe $\xi \in [a, b]$ tel que

$$e(x) = u(x) - P(x) = \frac{u^{(n)}(\xi)}{n!}(x - x_1) \cdots (x - x_n).$$

Problème : On sait que la tension de surface d'un liquide est fonction linéaire de la température. Comment, à partir de mesures, déterminer la relation linéaire ?

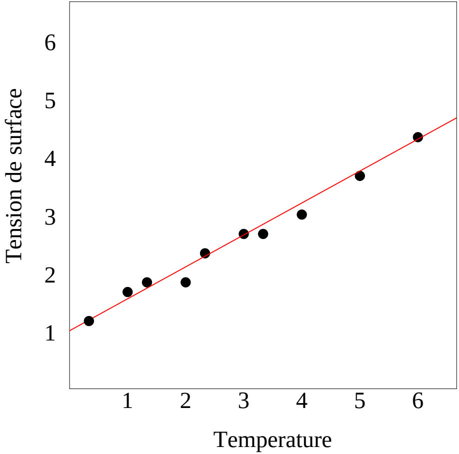
$$Tension \approx 1 + 0.55\, Temp$$

Problème : On sait que la tension de surface d'un liquide est fonction linéaire de la température. Comment, à partir de mesures, déterminer la relation linéaire ?



$$Tension \approx 1 + 0.55\, Temp$$

Problème : On sait que la tension de surface d'un liquide est fonction linéaire de la température. Comment, à partir de mesures, déterminer la relation linéaire ?



$$Tension \approx 1 + 0.55\,Temp$$

Et l'interpolation ?

- Outils précédents pas appropriés
- Interpolation polynomiale est mauvaise :
  - Grande erreur
  - Oscillations démesurées
  - Aucun sens en dehors de l'intervalle de mesure
- Interpolation par splines moins mauvaise mais
  - Trop dépendante de mesures particulières
  - Impossibilité d'avoir 2 ordonnées pour la même abscisse
  - Aucune prédiction en dehors de l'intervalle de mesure
  - Impossibilité d'obtenir des coefficients mesurés par ex. si on sait que la dépendance est linéaire

Et l'interpolation ?

- Outils précédents pas appropriés
- Interpolation polynomiale est mauvaise :
  - Grande erreur
  - Oscillations démesurées
  - Aucun sens en dehors de l'intervalle de mesure
- Interpolation par splines moins mauvaise mais
  - Trop dépendante de mesures particulières
  - Impossibilité d'avoir 2 ordonnées pour la même abscisse
  - Aucune prédiction en dehors de l'intervalle de mesure
  - Impossibilité d'obtenir des coefficients mesurés par ex. si on sait que la dépendance est linéaire

Et l'interpolation ?

- Outils précédents pas appropriés
- Interpolation polynomiale est mauvaise :
  - Grande erreur
  - Oscillations démesurées
  - Aucun sens en dehors de l'intervalle de mesure
- Interpolation par splines moins mauvaise mais
  - Trop dépendante de mesures particulières
  - Impossibilité d'avoir 2 ordonnées pour la même abscisse
  - Aucune prédiction en dehors de l'intervalle de mesure
  - Impossibilité d'obtenir des coefficients mesurés par ex. si on sait que la dépendance est linéaire

Régression linéaire

Données : $(x_1, u(x_1)), (x_2, u(x_2)), \ldots, (x_n, u(x_n))$
Question : Trouver a et b tels

$$u(x_i) \approx ax_i + b \qquad \text{pour tout } i$$

Définissons l'erreur

$$e_i = u(x_i) - (ax_i + b).$$

Minimisons la somme des carrés des erreurs

$$E(a, b) := \sum_{i=1}^{n} e_i^2$$

Régression linéaire

Données : $(x_1, u(x_1)), (x_2, u(x_2)), \ldots, (x_n, u(x_n))$
Question : Trouver a et b tels

$$u(x_i) \approx ax_i + b \qquad \text{pour tout } i$$

Définissons l'erreur

$$e_i = u(x_i) - (ax_i + b).$$

Minimisons la somme des carrés des erreurs

$$E(a, b) := \sum_{i=1}^{n} e_i^2$$

Régression linéaire

Données : $(x_1, u(x_1)), (x_2, u(x_2)), \dots, (x_n, u(x_n))$
Question : Trouver a et b tels

$$u(x_i) \approx ax_i + b \qquad \text{pour tout } i$$

Définissons l'erreur

$$e_i = u(x_i) - (ax_i + b).$$

Minimisons la somme des carrés des erreurs

$$E(a, b) := \sum_{i=1}^{n} e_i^2$$

Pourquoi minimiser la somme des carrés des erreurs ?

- Erreurs positives et négatives sont comptabilisées positivement (pas de "compensation" douteuse)
- Forte pénalisation des grandes erreurs (discutable)
- Produit une fonction erreur continue et dérivable
  $\longrightarrow$ Coefficients faciles à calculer
- Dans certains cas, un autre choix pourrait être judicieux
- Dans la plupart des cas : très satisfaisant

Pourquoi minimiser la somme des carrés des erreurs ?

- Erreurs positives et négatives sont comptabilisées positivement (pas de "compensation" douteuse)
- Forte pénalisation des grandes erreurs (discutable)
- Produit une fonction erreur continue et dérivable
  $\longrightarrow$ Coefficients faciles à calculer
- Dans certains cas, un autre choix pourrait être judicieux
- Dans la plupart des cas : très satisfaisant

Pourquoi minimiser la somme des carrés des erreurs ?

- Erreurs positives et négatives sont comptabilisées positivement (pas de "compensation" douteuse)
- Forte pénalisation des grandes erreurs (discutable)
- Produit une fonction erreur continue et dérivable
  $\longrightarrow$ Coefficients faciles à calculer
- Dans certains cas, un autre choix pourrait être judicieux
- Dans la plupart des cas : très satisfaisant

Pourquoi minimiser la somme des carrés des erreurs ?

- Erreurs positives et négatives sont comptabilisées positivement (pas de “compensation” douteuse)
- Forte pénalisation des grandes erreurs (discutable)
- Produit une fonction erreur continue et dérivable
  $\longrightarrow$ Coefficients faciles à calculer
- Dans certains cas, un autre choix pourrait être judicieux
- Dans la plupart des cas : très satisfaisant

Pourquoi minimiser la somme des carrés des erreurs ?

- Erreurs positives et négatives sont comptabilisées positivement (pas de "compensation" douteuse)
- Forte pénalisation des grandes erreurs (discutable)
- Produit une fonction erreur continue et dérivable
  $\longrightarrow$ Coefficients faciles à calculer
- Dans certains cas, un autre choix pourrait être judicieux
- Dans la plupart des cas : très satisfaisant

Calcul des coefficients

$$E(a,b) = \sum_{i=1}^{n} (ax_i + b - u(x_i))^2.$$

Minimisation $\Rightarrow$ Annuler le Jacobien

$$\frac{\partial E(a,b)}{\partial a} = 0, \qquad \frac{\partial E(a,b)}{\partial b} = 0.$$

$$\sum_{i=1}^{n} 2x_i(ax_i + b - u(x_i)) = 0$$

$$\sum_{i=1}^{n} 2(ax_i + b - u(x_i)) = 0.$$

a et b sont les variables $\Rightarrow$ Système linéaire

$$\begin{pmatrix} \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & n \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} x_i u(x_i) \\ \sum_{i=1}^{n} u(x_i). \end{pmatrix}.$$

Equations normales

Calcul des coefficients

$$E(a,b) = \sum_{i=1}^{n} (ax_i + b - u(x_i))^2.$$

Minimisation $\Rightarrow$ Annuler le Jacobien

$$\frac{\partial E(a,b)}{\partial a} = 0, \qquad \frac{\partial E(a,b)}{\partial b} = 0.$$

$$\sum_{i=1}^{n} 2x_i(ax_i + b - u(x_i)) = 0$$

$$\sum_{i=1}^{n} 2(ax_i + b - u(x_i)) = 0.$$

a et b sont les variables $\Rightarrow$ Système linéaire

$$\begin{pmatrix} \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & n \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} x_i u(x_i) \\ \sum_{i=1}^{n} u(x_i). \end{pmatrix}.$$

Equations normales

Calcul des coefficients

$$E(a,b) = \sum_{i=1}^{n} (ax_i + b - u(x_i))^2.$$

Minimisation $\Rightarrow$ Annuler le Jacobien

$$\frac{\partial E(a,b)}{\partial a} = 0, \qquad \frac{\partial E(a,b)}{\partial b} = 0.$$

$$\sum_{i=1}^{n} 2x_i(ax_i + b - u(x_i)) = 0$$

$$\sum_{i=1}^{n} 2(ax_i + b - u(x_i)) = 0.$$

a et b sont les variables $\Rightarrow$ Système linéaire

$$\begin{pmatrix} \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & n \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} x_i u(x_i) \\ \sum_{i=1}^{n} u(x_i). \end{pmatrix}.$$

Equations normales

Calcul des coefficients

$$E(a, b) = \sum_{i=1}^{n} (ax_i + b - u(x_i))^2.$$

Minimisation $\Rightarrow$ Annuler le Jacobien

$$\frac{\partial E(a,b)}{\partial a} = 0, \qquad \frac{\partial E(a,b)}{\partial b} = 0.$$

$$\sum_{i=1}^{n} 2x_i(ax_i + b - u(x_i)) = 0$$
$$\sum_{i=1}^{n} 2(ax_i + b - u(x_i)) = 0.$$

a et b sont les variables $\Rightarrow$ Système linéaire

$$\begin{pmatrix} \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & n \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} x_i u(x_i) \\ \sum_{i=1}^{n} u(x_i). \end{pmatrix}.$$

Equations normales

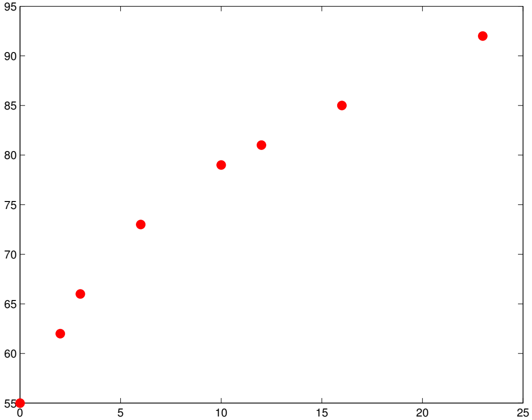
Régression non linéaire

Nouveau problème : Mesure d'un enfant à des moments précis
Peut-on avoir une expression de la taille en fonction de l'âge ?

Probablement pas une fonction linéaire
Régression linéaire peu appropriée
Possibilité d'inclure des fonctions non linéaires ?

Régression non linéaire

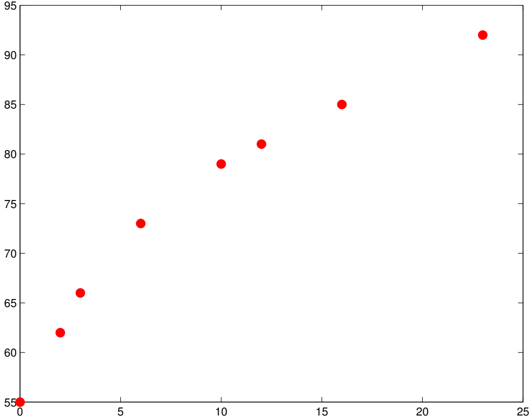
Nouveau problème : Mesure d'un enfant à des moments précis
Peut-on avoir une expression de la taille en fonction de l'âge ?



Probablement pas une fonction linéaire
Régression linéaire peu appropriée
Possibilité d'inclure des fonctions non linéaires ?

Régression non linéaire

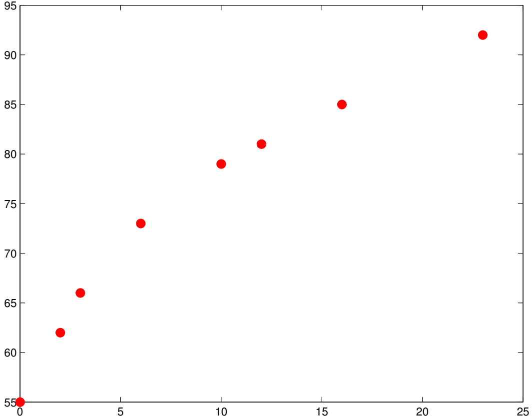
Nouveau problème : Mesure d'un enfant à des moments précis
Peut-on avoir une expression de la taille en fonction de l'âge ?



Probablement pas une fonction linéaire
Régression linéaire peu appropriée
Possibilité d'inclure des fonctions non linéaires ?

Régression non linéaire

Nouveau problème : Mesure d'un enfant à des moments précis
Peut-on avoir une expression de la taille en fonction de l'âge ?



Probablement pas une fonction linéaire
Régression linéaire peu appropriée
Possibilité d'inclure des fonctions non linéaires ?

Régression non linéaire

On utilise des fonctions de base.

Ex. : 1, $\ln x$: on veut $u(x_i) \approx a \ln x_i + b$.

$$E(a,b) = \sum_{i=1}^{n} (u(x_i) - (a \ln x_i + b))^2$$

$$\frac{\partial E(a,b)}{\partial a} = 0, \qquad \frac{\partial E(a,b)}{\partial b} = 0.$$

$$\sum_{i=1}^{n} 2 \ln x_i (a \ln x_i + b - u(x_i)) = 0$$

$$\sum_{i=1}^{n} 2 (a \ln x_i + b - u(x_i)) = 0.$$

a et b sont les variables $\Rightarrow$ Système linéaire

$$\begin{pmatrix} \sum_{i=1}^{n} (\ln x_i)^2 & \sum_{i=1}^{n} \ln x_i \\ \sum_{i=1}^{n} \ln x_i & n \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} \ln x_i u(x_i) \\ \sum_{i=1}^{n} u(x_i). \end{pmatrix}.$$

Régression non linéaire

On utilise des fonctions de base.

Ex. : 1, $\ln x$: on veut $u(x_i) \approx a \ln x_i + b$.

$$E(a,b) = \sum_{i=1}^{n} (u(x_i) - (a \ln x_i + b))^2$$

$$\frac{\partial E(a,b)}{\partial a} = 0, \qquad \frac{\partial E(a,b)}{\partial b} = 0.$$

$$\sum_{i=1}^{n} 2 \ln x_i (a \ln x_i + b - u(x_i)) = 0$$

$$\sum_{i=1}^{n} 2(a \ln x_i + b - u(x_i)) = 0.$$

a et b sont les variables $\Rightarrow$ Système linéaire

$$\begin{pmatrix} \sum_{i=1}^{n} (\ln x_i)^2 & \sum_{i=1}^{n} \ln x_i \\ \sum_{i=1}^{n} \ln x_i & n \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} \ln x_i u(x_i) \\ \sum_{i=1}^{n} u(x_i). \end{pmatrix}.$$

Régression non linéaire

On utilise des fonctions de base.

Ex. : 1, $\ln x$: on veut $u(x_i) \approx a \ln x_i + b$.

$$E(a,b) = \sum_{i=1}^{n} (u(x_i) - (a \ln x_i + b))^2$$

$$\frac{\partial E(a,b)}{\partial a} = 0, \qquad \frac{\partial E(a,b)}{\partial b} = 0.$$

$$\sum_{i=1}^{n} 2 \ln x_i (a \ln x_i + b - u(x_i)) = 0$$

$$\sum_{i=1}^{n} 2(a \ln x_i + b - u(x_i)) = 0.$$

a et b sont les variables $\Rightarrow$ Système linéaire

$$\begin{pmatrix} \sum_{i=1}^{n} (\ln x_i)^2 & \sum_{i=1}^{n} \ln x_i \\ \sum_{i=1}^{n} \ln x_i & n \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} \ln x_i u(x_i) \\ \sum_{i=1}^{n} u(x_i). \end{pmatrix}.$$

Régression non linéaire

On utilise des fonctions de base.

Ex. : 1, $\ln x$: on veut $u(x_i) \approx a \ln x_i + b$.

$$E(a, b) = \sum_{i=1}^{n} (u(x_i) - (a \ln x_i + b))^2$$

$$\frac{\partial E(a, b)}{\partial a} = 0, \qquad \frac{\partial E(a, b)}{\partial b} = 0.$$

$$\sum_{i=1}^{n} 2 \ln x_i (a \ln x_i + b - u(x_i)) = 0$$

$$\sum_{i=1}^{n} 2(a \ln x_i + b - u(x_i)) = 0.$$

a et b sont les variables $\Rightarrow$ Système linéaire

$$\begin{pmatrix} \sum_{i=1}^{n} (\ln x_i)^2 & \sum_{i=1}^{n} \ln x_i \\ \sum_{i=1}^{n} \ln x_i & n \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} \ln x_i u(x_i) \\ \sum_{i=1}^{n} u(x_i). \end{pmatrix}.$$

Régression non linéaire

On utilise des fonctions de base.
Ex. : 1, $\ln x$: on veut $u(x_i) \approx a \ln x_i + b$.

$$E(a,b) = \sum_{i=1}^{n} \left(u(x_i) - (a \ln x_i + b)\right)^2$$

$$\frac{\partial E(a,b)}{\partial a} = 0, \qquad \frac{\partial E(a,b)}{\partial b} = 0.$$

$$\sum_{i=1}^{n} 2 \ln x_i (a \ln x_i + b - u(x_i)) = 0$$

$$\sum_{i=1}^{n} 2 (a \ln x_i + b - u(x_i)) = 0.$$

a et b sont les variables $\Rightarrow$ Système linéaire

$$\begin{pmatrix} \sum_{i=1}^{n} (\ln x_i)^2 & \sum_{i=1}^{n} \ln x_i \\ \sum_{i=1}^{n} \ln x_i & n \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} \ln x_i u(x_i) \\ \sum_{i=1}^{n} u(x_i). \end{pmatrix}.$$

Régression non linéaire sur l'exemple

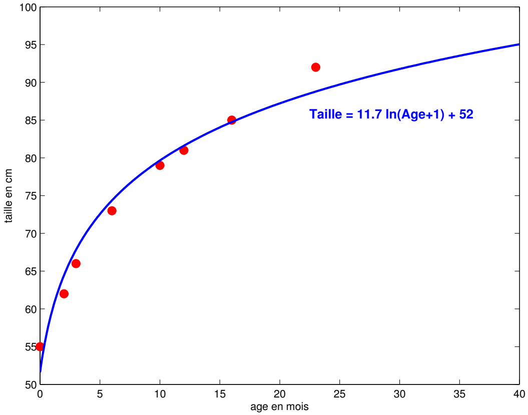
Dans notre cas, on a

$$a = 11.7 \text{ et } b = 52.$$

Régression non linéaire sur l'exemple

Dans notre cas, on a

$$a = 11.7 \text{ et } b = 52.$$



Choix des fonctions de base

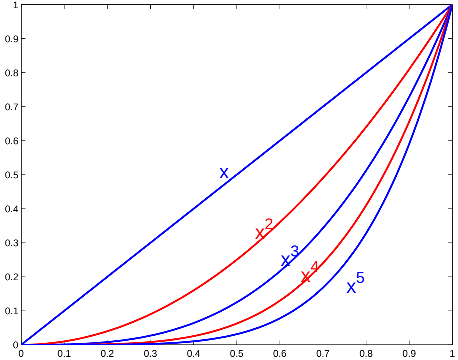
Choix courant : $1, x, x^2, x^3, x^4, \ldots$
Est-ce un bon choix ?

Fonctions de base trop similaires
Souvent un mauvais choix numériquement.

On préférera souvent des polynômes orthogonaux

Choix des fonctions de base

Choix courant : $1, x, x^2, x^3, x^4, \ldots$
Est-ce un bon choix ?

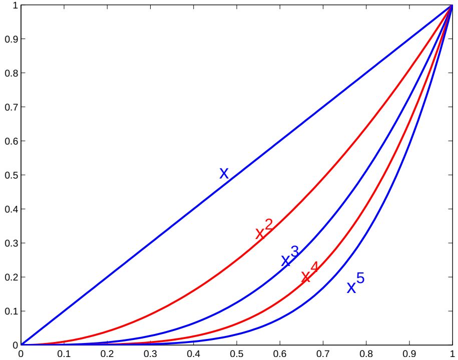


Fonctions de base trop similaires
Souvent un mauvais choix numériquement.

On préférera souvent des polynômes orthogonaux

Choix des fonctions de base

Choix courant : $1, x, x^2, x^3, x^4, \ldots$
Est-ce un bon choix ?

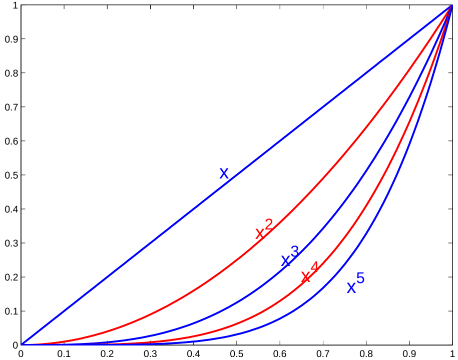


Fonctions de base trop similaires
Souvent un mauvais choix numériquement.

On préférera souvent des polynômes orthogonaux

Choix des fonctions de base

Choix courant : $1, x, x^2, x^3, x^4, \ldots$
Est-ce un bon choix ?



Fonctions de base trop similaires
Souvent un mauvais choix numériquement.

On préférera souvent des polynômes orthogonaux

Que sont les polynômes orthogonaux ?

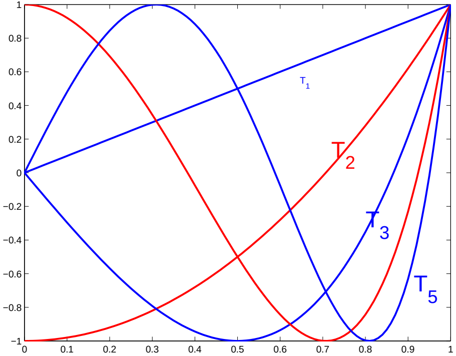
On a déjà rencontré une famille de polynômes orthogonaux : les polynômes de Chebyshev.

Des polynômes plus "différents".
Plus stable numériquement

Facilités pour le calcul des coefficients

Que sont les polynômes orthogonaux ?

On a déjà rencontré une famille de polynômes orthogonaux : les polynômes de Chebyshev.



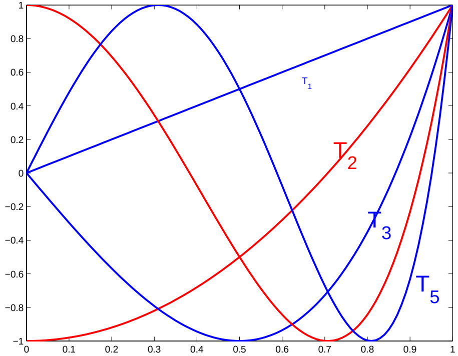
Des polynômes plus "différents".

Plus stable numériquement

Facilités pour le calcul des coefficients

Que sont les polynômes orthogonaux ?

On a déjà rencontré une famille de polynômes orthogonaux : les polynômes de Chebyshev.

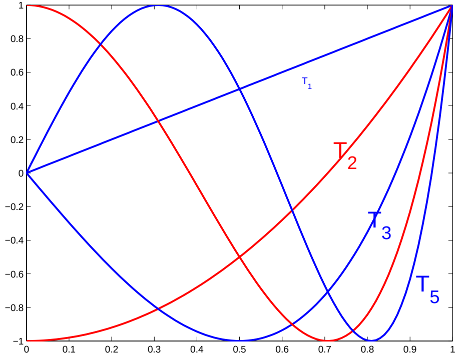


Des polynômes plus "différents".
Plus stable numériquement

Facilités pour le calcul des coefficients

Que sont les polynômes orthogonaux ?

On a déjà rencontré une famille de polynômes orthogonaux : les polynômes de Chebyshev.



Des polynômes plus "différents".
Plus stable numériquement

Facilités pour le calcul des coefficients

Les polynômes orthogonaux

Etape 1 : Définition d'un produit scalaire sur les polynômes.

Un produit scalaire

$\langle f, g\rangle$ est un produit scalaire si les propriétés

(i) $\langle f, g\rangle = \langle g, f\rangle$

(ii) $\langle f, f\rangle \geq 0$ et $\langle f, f\rangle = 0 \implies f = 0$

(iii) $\langle af, g\rangle = a\langle f, g\rangle$ pour tout $a \in \mathbb{R}$

(iv) $\langle f, g+h\rangle = \langle f, g\rangle + \langle f, h\rangle$

sont satisfaites.

Exemples de produit scalaire sur les polynômes :

- $\int_a^b f(x)g(x)dx$
- $\int_a^b w(x)f(x)g(x)dx$ avec $w(x) \geq 0$ sur $[a, b]$
- Les polynômes de Chebyshev sont la famille orthogonale pour

$$\langle f, g\rangle = \int_{-1}^{1} \frac{f(x)g(x)}{\sqrt{1-x^2}}dx$$

- $\sum_{i=1}^{n} f(x_i)g(x_i)$ Attention : précaution pour la propriété (ii)

Les polynômes orthogonaux

Etape 1 : Définition d'un produit scalaire sur les polynômes.

Un produit scalaire

$\langle f, g\rangle$ est un produit scalaire si les propriétés

(i) $\langle f, g\rangle = \langle g, f\rangle$

(ii) $\langle f, f\rangle \geq 0$ et $\langle f, f\rangle = 0 \implies f = 0$

(iii) $\langle af, g\rangle = a\langle f, g\rangle$ pour tout $a \in \mathbb{R}$

(iv) $\langle f, g + h\rangle = \langle f, g\rangle + \langle f, h\rangle$

sont satisfaites.

Exemples de produit scalaire sur les polynômes :

- $\int_a^b f(x)g(x)dx$
- $\int_a^b w(x)f(x)g(x)dx$ avec $w(x) \geq 0$ sur $[a, b]$
- Les polynômes de Chebyshev sont la famille orthogonale pour

$$\langle f, g\rangle = \int_{-1}^{1} \frac{f(x)g(x)}{\sqrt{1 - x^2}} dx$$

- $\sum_{i=1}^n f(x_i)g(x_i)$ Attention : précaution pour la propriété (ii)

Les polynômes orthogonaux

Etape 1 : Définition d'un produit scalaire sur les polynômes.

Un produit scalaire

$\langle f, g\rangle$ est un produit scalaire si les propriétés

(i) $\langle f, g\rangle = \langle g, f\rangle$

(ii) $\langle f, f\rangle \geq 0$ et $\langle f, f\rangle = 0 \implies f = 0$

(iii) $\langle af, g\rangle = a\langle f, g\rangle$ pour tout $a \in \mathbb{R}$

(iv) $\langle f, g + h\rangle = \langle f, g\rangle + \langle f, h\rangle$

sont satisfaites.

Exemples de produit scalaire sur les polynômes :

- $\int_a^b f(x)g(x)dx$
- $\int_a^b w(x)f(x)g(x)dx$ avec $w(x) \geq 0$ sur $[a, b]$
- Les polynômes de Chebyshev sont la famille orthogonale pour

$$\langle f, g\rangle = \int_{-1}^{1} \frac{f(x)g(x)}{\sqrt{1-x^2}}dx$$

- $\sum_{i=1}^n f(x_i)g(x_i)$ Attention : précaution pour la propriété (ii)

Les polynômes orthogonaux

Etape 1 : Définition d'un produit scalaire sur les polynômes.

Un produit scalaire

$\langle f, g\rangle$ est un produit scalaire si les propriétés

(i) $\langle f, g\rangle = \langle g, f\rangle$

(ii) $\langle f, f\rangle \geq 0$ et $\langle f, f\rangle = 0 \implies f = 0$

(iii) $\langle af, g\rangle = a\langle f, g\rangle$ pour tout $a \in \mathbb{R}$

(iv) $\langle f, g + h\rangle = \langle f, g\rangle + \langle f, h\rangle$

sont satisfaites.

Exemples de produit scalaire sur les polynômes :

- $\int_a^b f(x)g(x)dx$
- $\int_a^b w(x)f(x)g(x)dx$ avec $w(x) \geq 0$ sur $[a, b]$
- Les polynômes de Chebyshev sont la famille orthogonale pour

$$\langle f, g\rangle = \int_{-1}^{1} \frac{f(x)g(x)}{\sqrt{1 - x^2}} dx$$

- $\sum_{i=1}^{n} f(x_i)g(x_i)$ Attention : précaution pour la propriété (ii)

Les polynômes orthogonaux

Etape 1 : Définition d'un produit scalaire sur les polynômes.

Un produit scalaire

$\langle f, g\rangle$ est un produit scalaire si les propriétés

(i) $\langle f, g\rangle = \langle g, f\rangle$

(ii) $\langle f, f\rangle \geq 0$ et $\langle f, f\rangle = 0 \implies f = 0$

(iii) $\langle af, g\rangle = a\langle f, g\rangle$ pour tout $a \in \mathbb{R}$

(iv) $\langle f, g + h\rangle = \langle f, g\rangle + \langle f, h\rangle$

sont satisfaites.

Exemples de produit scalaire sur les polynômes :

- $\int_a^b f(x)g(x)dx$
- $\int_a^b w(x)f(x)g(x)dx$ avec $w(x) \geq 0$ sur $[a, b]$
- Les polynômes de Chebyshev sont la famille orthogonale pour

$$\langle f, g\rangle = \int_{-1}^{1} \frac{f(x)g(x)}{\sqrt{1-x^2}}dx$$

- $\sum_{i=1}^{n} f(x_i)g(x_i)$ Attention : précaution pour la propriété (ii)

Les polynômes orthogonaux

Définition

La famille de polynômes $(p_0, \dots, p_t)$ est orthogonale si

$$\langle p_i, p_j \rangle = 0$$

pour tout $i \neq j$.

Possibilité de construire une famille orthogonale à partir d'un produit scalaire.
(Voir notes de cours)

Les polynômes orthogonaux

Définition

La famille de polynômes $(p_0, \ldots, p_t)$ est orthogonale si

$$\langle p_i, p_j \rangle = 0$$

pour tout $i \neq j$.

Possibilité de construire une famille orthogonale à partir d'un produit scalaire.
(Voir notes de cours)

Utilisation de polynômes orthogonaux pour régression polynomiale

Les équations normales s'écrivent

$$\sum_{j=1}^{m}(\sum_{i=1}^{n}\phi_1(x_i)\phi_j(x_i))a_j = \sum_{i=1}^{n}\phi_1(x_i)u(x_i)$$

$$\vdots$$

$$\sum_{j=1}^{m}(\sum_{i=1}^{n}\phi_m(x_i)\phi_j(x_i))a_j = \sum_{i=1}^{n}\phi_m(x_i)u(x_i)$$

Si on utilise le produit scalaire

$$\langle f, g\rangle = \sum_{i=1}^{n} f(x_i)g(x_i),$$

et la famille correspondante orthogonale $(p_0, p_1, \ldots, p_k)$ elles se réécrivent

$$\begin{pmatrix} \langle p_0, p_0\rangle & \langle p_0, p_1\rangle & \cdots & \langle p_0, p_k\rangle \\ \langle p_1, p_0\rangle & \langle p_1, p_1\rangle & \cdots & \langle p_1, p_k\rangle \\ \vdots & & \ddots & \vdots \\ \langle p_k, p_0\rangle & \langle p_k, p_1\rangle & \cdots & \langle p_k, p_k\rangle \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_k \end{pmatrix} = \begin{pmatrix} \langle u, p_0\rangle \\ \langle u, p_1\rangle \\ \vdots \\ \langle u, p_k\rangle \end{pmatrix}.$$

Utilisation de polynômes orthogonaux pour régression polynomiale

Les équations normales s'écrivent

$$\sum_{j=1}^{m}(\sum_{i=1}^{n}\phi_1(x_i)\phi_j(x_i))a_j = \sum_{i=1}^{n}\phi_1(x_i)u(x_i)$$

$$\vdots$$

$$\sum_{j=1}^{m}(\sum_{i=1}^{n}\phi_m(x_i)\phi_j(x_i))a_j = \sum_{i=1}^{n}\phi_m(x_i)u(x_i)$$

Si on utilise le produit scalaire

$$\langle f, g\rangle = \sum_{i=1}^{n} f(x_i)g(x_i),$$

et la famille correspondante orthogonale $(p_0, p_1, \ldots, p_k)$ elles se réécrivent

$$\begin{pmatrix} \langle p_0,p_0\rangle & \langle p_0,p_1\rangle & \cdots & \langle p_0,p_k\rangle \\ \langle p_1,p_0\rangle & \langle p_1,p_1\rangle & \cdots & \langle p_1,p_k\rangle \\ \vdots & & \ddots & \vdots \\ \langle p_k,p_0\rangle & \langle p_k,p_1\rangle & \cdots & \langle p_k,p_k\rangle \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_k \end{pmatrix} = \begin{pmatrix} \langle u,p_0\rangle \\ \langle u,p_1\rangle \\ \vdots \\ \langle u,p_k\rangle \end{pmatrix}.$$

Utilisation de polynômes orthogonaux pour régression polynomiale

Les équations normales s'écrivent

$$\sum_{j=1}^{m}\left(\sum_{i=1}^{n}\phi_1(x_i)\phi_j(x_i)\right)a_j = \sum_{i=1}^{n}\phi_1(x_i)u(x_i)$$

$$\vdots$$

$$\sum_{j=1}^{m}\left(\sum_{i=1}^{n}\phi_m(x_i)\phi_j(x_i)\right)a_j = \sum_{i=1}^{n}\phi_m(x_i)u(x_i)$$

Si on utilise le produit scalaire

$$\langle f, g\rangle = \sum_{i=1}^{n} f(x_i)g(x_i),$$

et la famille correspondante orthogonale $(p_0, p_1, \ldots, p_k)$ elles se réécrivent

$$\begin{pmatrix} \langle p_0, p_0\rangle & \langle p_0, p_1\rangle & \cdots & \langle p_0, p_k\rangle \\ \langle p_1, p_0\rangle & \langle p_1, p_1\rangle & \cdots & \langle p_1, p_k\rangle \\ \vdots & & \ddots & \vdots \\ \langle p_k, p_0\rangle & \langle p_k, p_1\rangle & \cdots & \langle p_k, p_k\rangle \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_k \end{pmatrix} = \begin{pmatrix} \langle u, p_0\rangle \\ \langle u, p_1\rangle \\ \vdots \\ \langle u, p_k\rangle \end{pmatrix}.$$

Utilisation de polynômes orthogonaux pour régression polynomiale

Si on utilise le produit scalaire

$$\langle f, g \rangle = \sum_{i=1}^{n} f(x_i) g(x_i),$$

et la famille correspondante orthogonale $(p_0, p_1, \ldots, p_k)$ elles se réécrivent

$$\begin{pmatrix} \langle p_0, p_0 \rangle & \langle p_0, p_1 \rangle & \cdots & \langle p_0, p_k \rangle \\ \langle p_1, p_0 \rangle & \langle p_1, p_1 \rangle & \cdots & \langle p_1, p_k \rangle \\ \vdots & & \ddots & \vdots \\ \langle p_k, p_0 \rangle & \langle p_k, p_1 \rangle & \cdots & \langle p_k, p_k \rangle \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_k \end{pmatrix} = \begin{pmatrix} \langle u, p_0 \rangle \\ \langle u, p_1 \rangle \\ \vdots \\ \langle u, p_k \rangle \end{pmatrix}.$$

Par la propriété d'orthogonalité, on a donc

$$\begin{pmatrix} \langle p_0, p_0 \rangle & 0 & \cdots & 0 \\ 0 & \langle p_1, p_1 \rangle & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \langle p_k, p_k \rangle \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_k \end{pmatrix} = \begin{pmatrix} \langle u, p_0 \rangle \\ \langle u, p_1 \rangle \\ \vdots \\ \langle u, p_k \rangle \end{pmatrix}.$$

Utilisation de polynômes orthogonaux pour régression polynomiale

$$\begin{pmatrix} \langle p_0, p_0\rangle & \langle p_0, p_1\rangle & \cdots & \langle p_0, p_k\rangle \\ \langle p_1, p_0\rangle & \langle p_1, p_1\rangle & \cdots & \langle p_1, p_k\rangle \\ \vdots & & \ddots & \vdots \\ \langle p_k, p_0\rangle & \langle p_k, p_1\rangle & \cdots & \langle p_k, p_k\rangle \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_k \end{pmatrix} = \begin{pmatrix} \langle u, p_0\rangle \\ \langle u, p_1\rangle \\ \vdots \\ \langle u, p_k\rangle \end{pmatrix}.$$

Par la propriété d'orthogonalité, on a donc

$$\begin{pmatrix} \langle p_0, p_0\rangle & 0 & \cdots & 0 \\ 0 & \langle p_1, p_1\rangle & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \langle p_k, p_k\rangle \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_k \end{pmatrix} = \begin{pmatrix} \langle u, p_0\rangle \\ \langle u, p_1\rangle \\ \vdots \\ \langle u, p_k\rangle \end{pmatrix}.$$

Polynôme des moindres carrés de degré k : $q_k(x) = \sum_{i=0}^{k} a_i p_i(x)$ où

$$a_i = \frac{\langle u, p_i\rangle}{\langle p_i, p_i\rangle}.$$

a_i ne dépend pas du nombre de polynômes dans la base !

Utilisation de polynômes orthogonaux pour régression polynomiale

$$\begin{pmatrix} \langle p_0, p_0 \rangle & \langle p_0, p_1 \rangle & \cdots & \langle p_0, p_k \rangle \\ \langle p_1, p_0 \rangle & \langle p_1, p_1 \rangle & \cdots & \langle p_1, p_k \rangle \\ \vdots & & \ddots & \vdots \\ \langle p_k, p_0 \rangle & \langle p_k, p_1 \rangle & \cdots & \langle p_k, p_k \rangle \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_k \end{pmatrix} = \begin{pmatrix} \langle u, p_0 \rangle \\ \langle u, p_1 \rangle \\ \vdots \\ \langle u, p_k \rangle \end{pmatrix}.$$

Par la propriété d'orthogonalité, on a donc

$$\begin{pmatrix} \langle p_0, p_0 \rangle & 0 & \cdots & 0 \\ 0 & \langle p_1, p_1 \rangle & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \langle p_k, p_k \rangle \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_k \end{pmatrix} = \begin{pmatrix} \langle u, p_0 \rangle \\ \langle u, p_1 \rangle \\ \vdots \\ \langle u, p_k \rangle \end{pmatrix}.$$

Polynôme des moindres carrés de degré k : $q_k(x) = \sum_{i=0}^{k} a_i p_i(x)$ où

$$a_i = \frac{\langle u, p_i \rangle}{\langle p_i, p_i \rangle}.$$

a_i ne dépend pas du nombre de polynômes dans la base !

Quel degré pour une régression polynomiale ?

Pour une régression donnée, on définit la variance

$$\sigma_j^2 = \frac{1}{n}\sum_{i=1}^{n}(u(x_i) - q_j(x_i))^2.$$

On peut prouver

$$\sigma_0^2 > \sigma_1^2 > \sigma_2^2 > \cdots > \sigma_{n-1}^2 = 0.$$

Degré optimal : Quand la variance ne décroit plus significativement

On peut prouver que le calcul est simplifié par une base orthogonale

$$\begin{aligned}
\sigma_k^2 &= \frac{1}{n}(\langle u - q_k, u - q_k\rangle) \\
&= \frac{1}{n}(\langle u - q_k, u\rangle - \langle u - q_k, q_k\rangle) \quad \text{(partie rouge nulle)} \\
&= \frac{1}{n}(\langle u, u\rangle - \langle q_k, u\rangle) \\
&= \frac{1}{n}(\langle u, u\rangle - \sum_{i=0}^{k} a_i\langle p_i, u\rangle) \\
&= \frac{1}{n}(\langle u, u\rangle - \sum_{i=0}^{k} \frac{\langle p_i, u\rangle^2}{\langle p_i, p_i\rangle}).
\end{aligned}$$

Quel degré pour une régression polynomiale ?

Pour une régression donnée, on définit la variance

$$\sigma_j^2 = \frac{1}{n}\sum_{i=1}^{n}(u(x_i) - q_j(x_i))^2.$$

On peut prouver

$$\sigma_0^2 > \sigma_1^2 > \sigma_2^2 > \cdots > \sigma_{n-1}^2 = 0.$$

Degré optimal : Quand la variance ne décroit plus significativement

On peut prouver que le calcul est simplifié par une base orthogonale

$$\begin{aligned}
\sigma_k^2 &= \frac{1}{n}(\langle u - q_k, u - q_k\rangle) \\
&= \frac{1}{n}(\langle u - q_k, u\rangle - \langle u - q_k, q_k\rangle) \quad \text{(partie rouge nulle)} \\
&= \frac{1}{n}(\langle u, u\rangle - \langle q_k, u\rangle) \\
&= \frac{1}{n}(\langle u, u\rangle - \sum_{i=0}^{k} a_i\langle p_i, u\rangle) \\
&= \frac{1}{n}(\langle u, u\rangle - \sum_{i=0}^{k} \frac{\langle p_i, u\rangle^2}{\langle p_i, p_i\rangle}).
\end{aligned}$$

Quel degré pour une régression polynomiale ?

Pour une régression donnée, on définit la variance

$$\sigma_j^2 = \frac{1}{n}\sum_{i=1}^{n}(u(x_i) - q_j(x_i))^2.$$

On peut prouver

$$\sigma_0^2 > \sigma_1^2 > \sigma_2^2 > \cdots > \sigma_{n-1}^2 = 0.$$

Degré optimal : Quand la variance ne décroit plus significativement

On peut prouver que le calcul est simplifié par une base orthogonale

$$\begin{aligned}
\sigma_k^2 &= \frac{1}{n}(\langle u - q_k, u - q_k\rangle) \\
&= \frac{1}{n}(\langle u - q_k, u\rangle - \langle u - q_k, q_k\rangle) \quad \text{(partie rouge nulle)} \\
&= \frac{1}{n}(\langle u, u\rangle - \langle q_k, u\rangle) \\
&= \frac{1}{n}(\langle u, u\rangle - \sum_{i=0}^{k} a_i\langle p_i, u\rangle) \\
&= \frac{1}{n}(\langle u, u\rangle - \sum_{i=0}^{k} \frac{\langle p_i, u\rangle^2}{\langle p_i, p_i\rangle}).
\end{aligned}$$

Quel degré pour une régression polynomiale ?

Pour une régression donnée, on définit la variance

$$\sigma_j^2 = \frac{1}{n}\sum_{i=1}^{n}(u(x_i) - q_j(x_i))^2.$$

On peut prouver

$$\sigma_0^2 > \sigma_1^2 > \sigma_2^2 > \cdots > \sigma_{n-1}^2 = 0.$$

Degré optimal : Quand la variance ne décroit plus significativement

On peut prouver que le calcul est simplifié par une base orthogonale

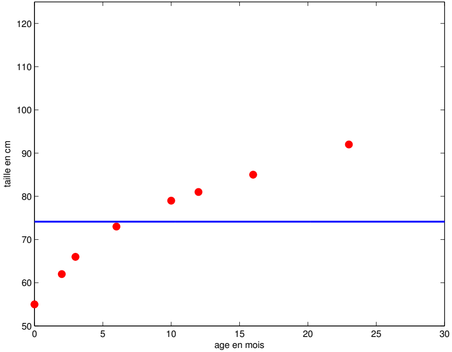
$$\begin{aligned}
\sigma_k^2 &= \frac{1}{n}(\langle u - q_k, u - q_k\rangle) \\
&= \frac{1}{n}(\langle u - q_k, u\rangle - \langle u - q_k, q_k\rangle) \quad \text{(partie rouge nulle)} \\
&= \frac{1}{n}(\langle u, u\rangle - \langle q_k, u\rangle) \\
&= \frac{1}{n}(\langle u, u\rangle - \sum_{i=0}^{k} a_i \langle p_i, u\rangle) \\
&= \frac{1}{n}(\langle u, u\rangle - \sum_{i=0}^{k} \frac{\langle p_i, u\rangle^2}{\langle p_i, p_i\rangle}).
\end{aligned}$$

Calcul des variances successives :

Degré	Variance
0	8711
1	509.5
2	77.8
3	3.4
4	1.8
5	1

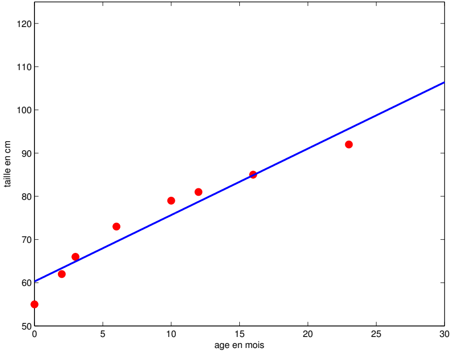
Calcul des variances successives :

Degré	Variance
0	8711
1	509.5
2	77.8
3	3.4
4	1.8
5	1



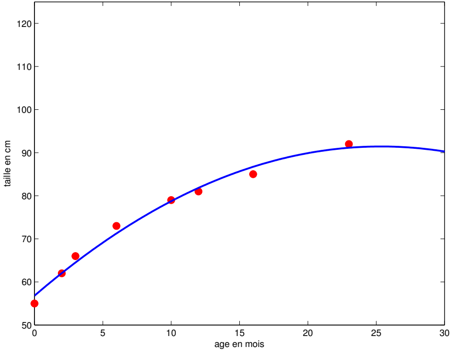
Calcul des variances successives :

Degré	Variance
0	8711
1	509.5
2	77.8
3	3.4
4	1.8
5	1



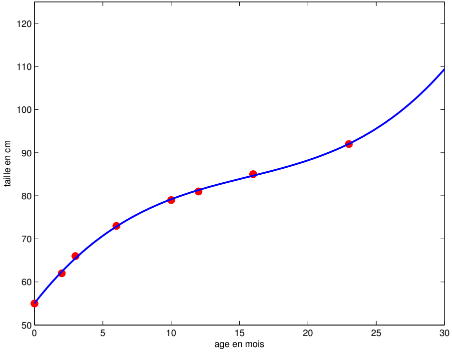
Calcul des variances successives :

Degré	Variance
0	8711
1	509.5
2	77.8
3	3.4
4	1.8
5	1



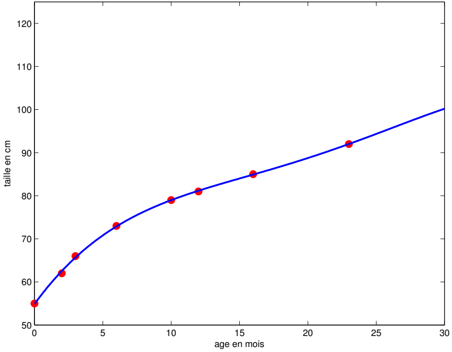
Calcul des variances successives :

Degré	Variance
0	8711
1	509.5
2	77.8
3	3.4
4	1.8
5	1



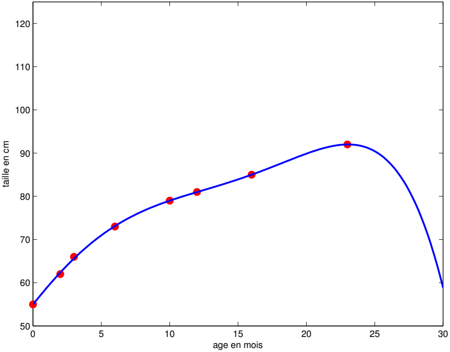
Calcul des variances successives :

Degré	Variance
0	8711
1	509.5
2	77.8
3	3.4
4	1.8
5	1



Calcul des variances successives :

Degré	Variance
0	8711
1	509.5
2	77.8
3	3.4
4	1.8
5	1



Pas satisfait ?

Une régression polynomiale ne semble pas très appropriée.
On voudrait une puissance de x inférieure à 1.

Solution : Procéder à la régression linéaire sur les logarithmes

$$(\ln(x_1), \ln(y_1)), \ldots, (\ln(x_n), \ln(y_n))$$

On va écrire

$$\ln(y) \approx a \ln(x) + b$$

$$y \approx e^{a \ln(x) + b}$$

$$y \approx e^b x^a = Cx^a$$

Pas satisfait ?

Une régression polynomiale ne semble pas très appropriée.
On voudrait une puissance de x inférieure à 1.

Solution : Procéder à la régression linéaire sur les logarithmes

$$(\ln(x_1), \ln(y_1))), \ldots, (\ln(x_n), \ln(y_n))$$

On va écrire

$$\ln(y) \approx a \ln(x) + b$$

$$y \approx e^{a \ln(x) + b}$$

$$y \approx e^b x^a = C x^a$$

Pas satisfait ?

Une régression polynomiale ne semble pas très appropriée.
On voudrait une puissance de x inférieure à 1.

Solution : Procéder à la régression linéaire sur les logarithmes

$$(\ln(x_1), \ln(y_1))), \ldots, (\ln(x_n), \ln(y_n))$$

On va écrire

$$\ln(y) \approx a \ln(x) + b$$

$$y \approx e^{a \ln(x) + b}$$

$$y \approx e^b x^a = Cx^a$$

Pas satisfait ?

Une régression polynomiale ne semble pas très appropriée.
On voudrait une puissance de x inférieure à 1.

Solution : Procéder à la régression linéaire sur les logarithmes

$$(\ln(x_1), \ln(y_1))), \ldots, (\ln(x_n), \ln(y_n))$$

On va écrire

$$\ln(y) \approx a \ln(x) + b$$

$$y \approx e^{a \ln(x) + b}$$

$$y \approx e^b x^a = Cx^a$$

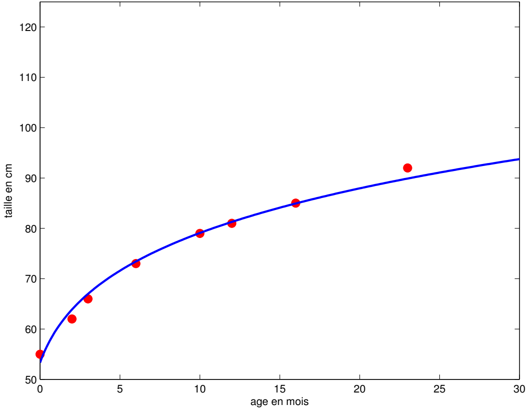
Dans notre exemple

Attention : Une des valeurs est nulle, il faudra décaler : $\ln(x+1)$

$$Taille = 53.3(Age+1)^{0.16}$$

Dans notre exemple

Attention : Une des valeurs est nulle, il faudra décaler : $\ln(x+1)$



$$Taille = 53.3(Age+1)^{0.16}$$