III

INTERPOLATION ET APPROXIMATION DE FONCTIONS

Analyse Numérique
Tronc Commun

POLYTECH CLERMONT-FERRAND

Un exemple

Évolution de la population en France

On considère l'évolution de la population française depuis 1936 :

1936	41 183 000
1946	39 848 000
1954	42 781 000
1962	46 459 000
1968	49 655 000
1975	52 599 000
1982	54 296 000
1990	56 652 000
1999	58 521 000
2005	60 825 000

Un exemple

Évolution de la population en France

On considère l'évolution de la population française depuis 1936 :

1936	41 183 000
1946	39 848 000
1954	42 781 000
1962	46 459 000
1968	49 655 000
1975	52 599 000
1982	54 296 000
1990	56 652 000
1999	58 521 000
2005	60 825 000

- Peut-on estimer le nombre d'habitants pendant les années où il n'y a pas eu de recensement ?

Un exemple

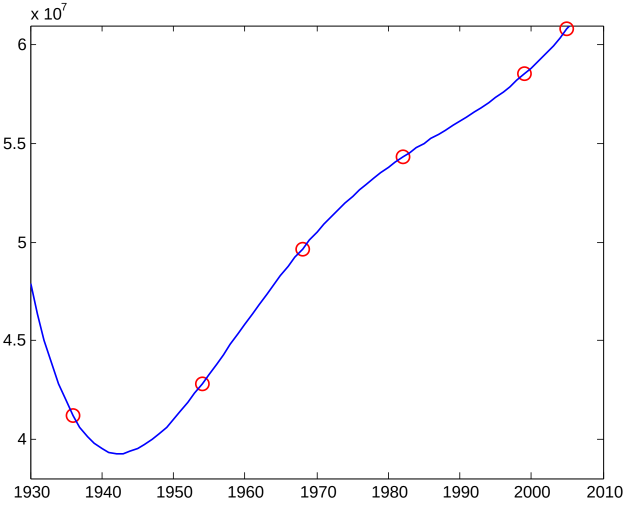
Évolution de la population en France

On considère l'évolution de la population française depuis 1936 :

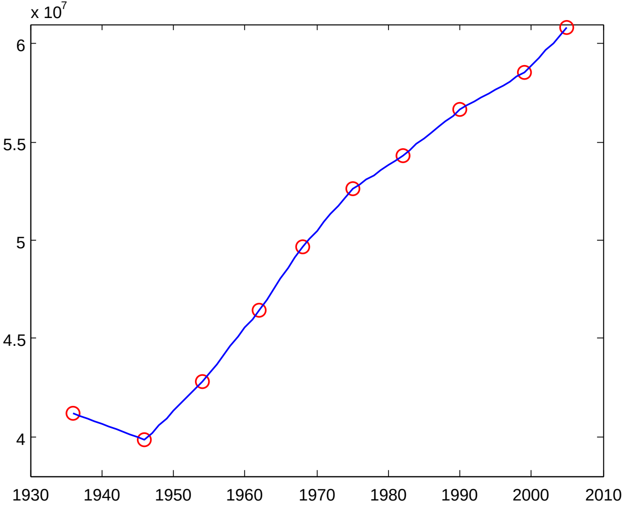
1936	41 183 000
1946	39 848 000
1954	42 781 000
1962	46 459 000
1968	49 655 000
1975	52 599 000
1982	54 296 000
1990	56 652 000
1999	58 521 000
2005	60 825 000

- Peut-on estimer le nombre d'habitants pendant les années où il n'y a pas eu de recensement ?
- Peut-on prédire le nombre d'habitants en 2010 ?

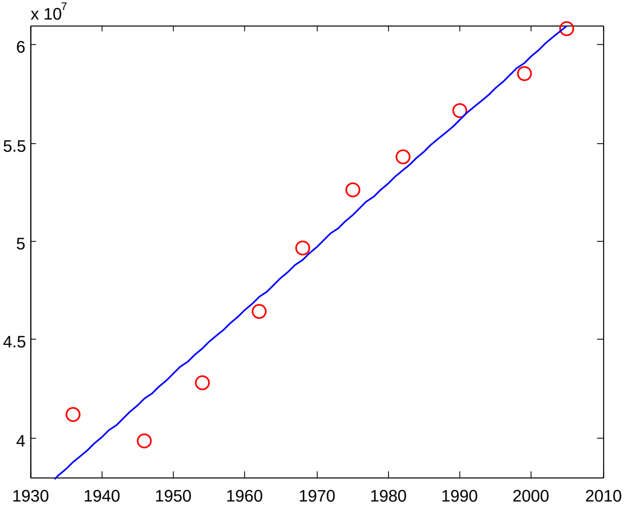
Interpolation polynomiale en 6 points



Interpolation linéaire par morceaux



Moindres carrés (Régression linéaire)



Interpolation de Lagrange

On considère $(n+1)$ points dans le plan (d'abscisses distincts)

$$(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$$

(données expérimentales, statistiques, ...)

Interpolation de Lagrange

On considère $(n+1)$ points dans le plan (d'abscisses distincts)

$$(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)$$

(données expérimentales, statistiques, ...)
On cherche une fonction *simple*, $p(x)$ facile à évaluer, passant par ces points :

$$p(x_i) = y_i \qquad i = 0, 1, \ldots, n$$

Interpolation de Lagrange

On considère $(n+1)$ points dans le plan (d'abscisses distincts)

$$(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$$

(données expérimentales, statistiques, ...)
On cherche une fonction *simple*, $p(x)$ facile à évaluer, passant par ces points :

$$p(x_i) = y_i \qquad i = 0, 1, \dots, n$$

Par exemple, la fonction p peut être polynomiale :

$$p(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n$$

Remarque

On obtient le système linéaire d'ordre $(n+1)$ suivant :

$$
\begin{array}{lllllll}
a_0 & + x_0\, a_1 & + x_0^2\, a_2 & + \ldots & + x_0^n\, a_n & = y_0 \\
a_0 & + x_1\, a_1 & + x_1^2\, a_2 & + \ldots & + x_1^n\, a_n & = y_1 \\
\ldots\ldots & \ldots & \ldots & \ldots & \ldots & \\
a_0 & + x_n\, a_1 & + x_n^2\, a_2 & + \ldots & + x_n^n\, a_n & = y_n
\end{array}
$$

Remarque

On obtient le système linéaire d'ordre $(n+1)$ suivant :

$$\begin{array}{lllll} a_0 & + x_0\, a_1 & + x_0^2\, a_2 & + \ldots + x_0^n\, a_n & = y_0 \\ a_0 & + x_1\, a_1 & + x_1^2\, a_2 & + \ldots + x_1^n\, a_n & = y_1 \\ \ldots\ldots & \ldots & \ldots & \ldots & \ldots \\ a_0 & + x_n\, a_1 & + x_n^2\, a_2 & + \ldots + x_n^n\, a_n & = y_n \end{array}$$

La résolution de ce système peut être en général coûteuse et n'est pas recommandée.

Bases de Lagrange

Prenons l'exemple d'une interpolation linéaire $n = 1$. On veut :

$$a_0 + a_1 x_0 = y_0$$
$$a_0 + a_1 x_1 = y_1$$

Bases de Lagrange

Prenons l'exemple d'une interpolation linéaire $n = 1$. On veut :

$$
\begin{aligned}
a_0 + a_1 x_0 &= y_0 \\
a_0 + a_1 x_1 &= y_1
\end{aligned}
$$

On définit la fonction

$$
\begin{aligned}
p(x) =& y_0 \, \frac{x - x_1}{x_0 - x_1} + y_1 \, \frac{x - x_0}{x_1 - x_0} \\
=& y_0 \, L_0(x) \quad + y_1 \, L_1(x)
\end{aligned}
$$

Bases de Lagrange

Prenons l'exemple d'une interpolation linéaire $n = 1$. On veut :

$$
\begin{aligned}
a_0 + a_1 x_0 &= y_0 \\
a_0 + a_1 x_1 &= y_1
\end{aligned}
$$

On définit la fonction

$$
\begin{aligned}
p(x) &= y_0 \frac{x - x_1}{x_0 - x_1} + y_1 \frac{x - x_0}{x_1 - x_0} \\
&= y_0 L_0(x) \quad + y_1 L_1(x)
\end{aligned}
$$

On vérifie que p est un polynôme de degré 1 et que

$$p(x_0) = y_0, \; p(x_1) = y_1$$

Donc p est le polynôme recherché.

Prenons l'exemple d'une interpolation linéaire $n = 1$. On veut :

$$a_0 + a_1 x_0 = y_0$$
$$a_0 + a_1 x_1 = y_1$$

On définit la fonction

$$\begin{aligned} p(x) &= y_0 \frac{x - x_1}{x_0 - x_1} + y_1 \frac{x - x_0}{x_1 - x_0} \\ &= y_0\, L_0(x) \quad + y_1\, L_1(x) \end{aligned}$$

On vérifie que p est un polynôme de degré 1 et que

$$p(x_0) = y_0, \;\; p(x_1) = y_1$$

Donc p est le polynôme recherché.

Définition

On dit que p est le polynôme d'interpolation de Lagrange associé aux points (x_0, y_0), (x_1, y_1).

On définit les polynômes

$$L_i(x) = \frac{(x-x_0)(x-x_1)\dots(x-x_{i-1})(x-x_{i+1})\dots(x-x_n)}{(x_i-x_0)(x_i-x_1)\dots(x_i-x_{i-1})(x_i-x_{i+1})\dots(x_i-x_n)} = \prod_{j\neq i} \frac{x-x_j}{x_i-x_j}$$

Bases de Lagrange

On définit les polynômes

$$L_i(x) = \frac{(x - x_0)(x - x_1)\dots(x - x_{i-1})(x - x_{i+1})\dots(x - x_n)}{(x_i - x_0)(x_i - x_1)\dots(x_i - x_{i-1})(x_i - x_{i+1})\dots(x_i - x_n)} = \prod_{j \neq i} \frac{x - x_j}{x_i - x_j}$$

L_i est un polynôme de degré n et on a

$$L_i(x_j) = \begin{cases} 0 & \text{si } i \neq j, \\ 1 & \text{si } i = j \end{cases}$$

On définit les polynômes

$$L_i(x) = \frac{(x-x_0)(x-x_1)\dots(x-x_{i-1})(x-x_{i+1})\dots(x-x_n)}{(x_i-x_0)(x_i-x_1)\dots(x_i-x_{i-1})(x_i-x_{i+1})\dots(x_i-x_n)} = \prod_{j\neq i} \frac{x-x_j}{x_i-x_j}$$

L_i est un polynôme de degré n et on a

$$L_i(x_j) = \begin{cases} 0 & \text{si } i \neq j, \\ 1 & \text{si } i = j \end{cases}$$

Montrons que les polynômes L_i sont linéairement indépendants :
Soient $b_0, b_1, \dots, b_n$ $(n+1)$ réels tels que

$$\sum_{j=0}^{n} b_j L_j(x) = 0 \qquad \forall\, x \in \mathbb{R}.$$

On définit les polynômes

$$L_i(x) = \frac{(x-x_0)(x-x_1)\dots(x-x_{i-1})(x-x_{i+1})\dots(x-x_n)}{(x_i-x_0)(x_i-x_1)\dots(x_i-x_{i-1})(x_i-x_{i+1})\dots(x_i-x_n)} = \prod_{j\neq i} \frac{x-x_j}{x_i-x_j}$$

L_i est un polynôme de degré n et on a

$$L_i(x_j) = \begin{cases} 0 & \text{si } i \neq j, \\ 1 & \text{si } i = j \end{cases}$$

Montrons que les polynômes L_i sont linéairement indépendants :
Soient $b_0, b_1, \dots, b_n$ $(n+1)$ réels tels que

$$\sum_{j=0}^{n} b_j L_j(x) = 0 \qquad \forall\ x \in \mathbb{R}.$$

On a, pour $x = x_i$

$$0 = \sum_{j=0}^{n} b_j L_j(x_i) = b_i.$$

Bases de Lagrange

On définit les polynômes

$$L_i(x) = \frac{(x-x_0)(x-x_1)\ldots(x-x_{i-1})(x-x_{i+1})\ldots(x-x_n)}{(x_i-x_0)(x_i-x_1)\ldots(x_i-x_{i-1})(x_i-x_{i+1})\ldots(x_i-x_n)} = \prod_{j\neq i} \frac{x-x_j}{x_i-x_j}$$

L_i est un polynôme de degré n et on a

$$L_i(x_j) = \begin{cases} 0 & \text{si } i \neq j, \\ 1 & \text{si } i = j \end{cases}$$

Montrons que les polynômes L_i sont linéairement indépendants :
Soient $b_0, b_1, \ldots, b_n$ $(n+1)$ réels tels que

$$\sum_{j=0}^{n} b_j L_j(x) = 0 \qquad \forall\ x \in \mathbb{R}.$$

On a, pour $x = x_i$

$$0 = \sum_{j=0}^{n} b_j L_j(x_i) = b_i.$$

Donc $b_i = 0$ pour tout $i = 0, \ldots, n$.

Ainsi la famille $(L_i)_{i=0}^n$ forme une base de l'espace $\mathbb{P}_n$ (espace des polynômes de degré $\leq n$), appelée Base de Lagrange.

Ainsi la famille $(L_i)_{i=0}^n$ forme une base de l'espace $\mathbb{P}_n$ (espace des polynômes de degré $\leq n$), appelée Base de Lagrange.

On définit le polynôme de $\mathbb{P}_n$:

$$p(x) = \sum_{j=0}^{n} y_j L_j(x) \qquad x \in \mathbb{R}.$$

Ainsi la famille $(L_i)_{i=0}^n$ forme une base de l'espace $\mathbb{P}_n$ (espace des polynômes de degré $\leq n$), appelée Base de Lagrange.

On définit le polynôme de $\mathbb{P}_n$:

$$p(x) = \sum_{j=0}^{n} y_j L_j(x) \qquad x \in \mathbb{R}.$$

On vérifie

$$p(x_i) = \sum_{j=0}^{n} y_j L_j(x_i) = y_i, \qquad i = 0, \ldots, n$$

Ainsi la famille $(L_i)_{i=0}^n$ forme une base de l'espace $\mathbb{P}_n$ (espace des polynômes de degré $\leq n$), appelée Base de Lagrange.

On définit le polynôme de $\mathbb{P}_n$:

$$p(x) = \sum_{j=0}^{n} y_j L_j(x) \qquad x \in \mathbb{R}.$$

On vérifie

$$p(x_i) = \sum_{j=0}^{n} y_j L_j(x_i) = y_i, \qquad i = 0, \dots, n$$

i.e. p est un polynôme de Lagrange associé aux points (x_i, y_i). Montrons que p est unique.

Ainsi la famille $(L_i)_{i=0}^n$ forme une base de l'espace $\mathbb{P}_n$ (espace des polynômes de degré $\leq n$), appelée Base de Lagrange.

On définit le polynôme de $\mathbb{P}_n$:

$$p(x) = \sum_{j=0}^{n} y_j L_j(x) \qquad x \in \mathbb{R}.$$

On vérifie

$$p(x_i) = \sum_{j=0}^{n} y_j L_j(x_i) = y_i, \qquad i = 0, \dots, n$$

i.e. p est un polynôme de Lagrange associé aux points (x_i, y_i). Montrons que p est unique.

Soit $q \in \mathbb{P}_n$ un autre polynôme de Lagrange associé aux points (x_i, y_i). Le polynôme $r = p - q$ est dans $\mathbb{P}_n$ et vérifie $r(x_i) = 0$ pour $i = 0, \dots, n$.

Ainsi la famille $(L_i)_{i=0}^n$ forme une base de l'espace $\mathbb{P}_n$ (espace des polynômes de degré $\leq n$), appelée Base de Lagrange.

On définit le polynôme de $\mathbb{P}_n$:

$$p(x) = \sum_{j=0}^{n} y_j L_j(x) \qquad x \in \mathbb{R}.$$

On vérifie

$$p(x_i) = \sum_{j=0}^{n} y_j L_j(x_i) = y_i, \qquad i = 0, \dots, n$$

i.e. p est un polynôme de Lagrange associé aux points (x_i, y_i). Montrons que p est unique.

Soit $q \in \mathbb{P}_n$ un autre polynôme de Lagrange associé aux points (x_i, y_i). Le polynôme $r = p - q$ est dans $\mathbb{P}_n$ et vérifie $r(x_i) = 0$ pour $i = 0, \dots, n$.
Puisque les polynômes L_j sont linéairement indépendants, on trouve $r = 0$. D'où $p = q$.

Erreur d'interpolation

Théorème

Soit f une fonction de classe $\mathcal{C}^{n+1}$ ($(n+1)$ fois continûment dérivable) et soit p le polynôme d'interpolation de Lagrange de f, aux points $x_0, x_1, \dots, x_n$ (polynôme de degré $\leq n$). Alors on a l'inégalité

$$|f(x) - p(x)| \leq \frac{M_{n+1}}{(n+1)!} |\pi_n(x)| \qquad \forall\, x \in \mathbb{R},$$

où

$$M_{n+1} := \sup_{x \in \mathbb{R}} |f^{(n+1)}(x)|,$$

$$\pi_n(x) = (x - x_0)\,(x - x_1)\,\dots\,(x - x_n).$$

Erreur d'interpolation

Théorème

Soit f une fonction de classe $\mathcal{C}^{n+1}$ ($(n+1)$ fois continûment dérivable) et soit p le polynôme d'interpolation de Lagrange de f, aux points $x_0, x_1, \dots, x_n$ (polynôme de degré $\leq n$). Alors on a l'inégalité

$$|f(x) - p(x)| \leq \frac{M_{n+1}}{(n+1)!} |\pi_n(x)| \qquad \forall\, x \in \mathbb{R},$$

où

$$M_{n+1} := \sup_{x \in \mathbb{R}} |f^{(n+1)}(x)|,$$
$$\pi_n(x) = (x - x_0)\,(x - x_1)\,\dots\,(x - x_n).$$

Remarque

On n'a pas nécessairement convergence lorsque $n \to +\infty$! !

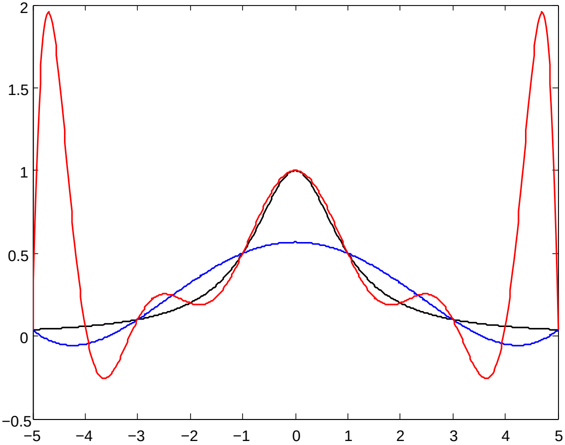
Inconvénient

Considérons par exemple la fonction suivante (fonction de Runge) :

$$f(x) = \frac{1}{1+x^2} \qquad x \in [-5, 5]$$

Considérons par exemple la fonction suivante (fonction de Runge) :

$$f(x) = \frac{1}{1+x^2} \qquad x \in [-5, 5]$$



Soit $a = x_0 < x_1 < \ldots < x_n = b$ une subdivision de l'intervalle $[a, b]$.

Définition

On appelle interpolant linéaire par morceaux d'une fonction f une fonction p satisfaisant les conditions :

1. $p_{|(x_{i-1},x_i)}$ est un polynôme de degré 1, pour tout $i = 0, \ldots, n$

Interpolation linéaire par morceaux

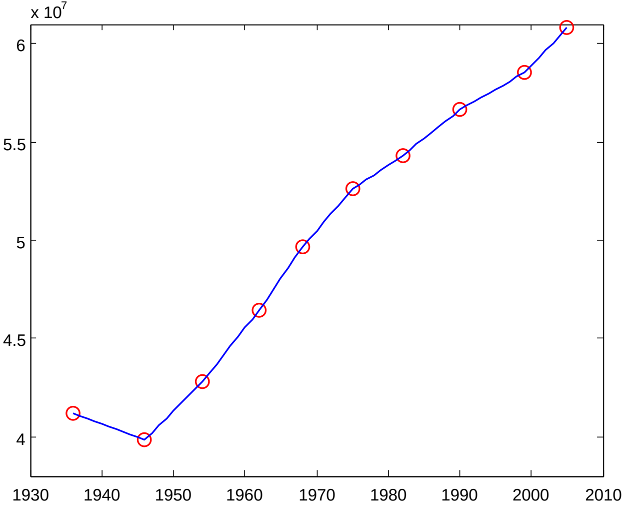
Soit $a = x_0 < x_1 < \ldots < x_n = b$ une subdivision de l'intervalle $[a, b]$.

DÉFINITION

On appelle interpolant linéaire par morceaux d'une fonction f une fonction p satisfaisant les conditions :

1. $p_{|(x_{i-1}, x_i)}$ est un polynôme de degré 1, pour tout $i = 0, \ldots, n$
2. $p(x_i) = f(x_i)$ pour tout $i = 0, \ldots, n$

Interpolation linéaire par morceaux



Soit $a = x_0 < x_1 < \ldots < x_n = b$ une subdivision de l'intervalle $[a, b]$.

Définition

On appelle spline cubique interpolant f une fonction s satisfaisant les conditions :

1. $s_{|(x_{i-1},x_i)}$ est un polynôme de degré 3, pour tout $i = 1, \ldots, n$

Soit $a = x_0 < x_1 < \ldots < x_n = b$ une subdivision de l'intervalle $[a, b]$.

DÉFINITION

On appelle spline cubique interpolant f une fonction s satisfaisant les conditions :

1. $s_{|(x_{i-1},x_i)}$ est un polynôme de degré 3, pour tout $i = 1, \ldots, n$
2. $s(x_i) = f(x_i)$ pour tout $i = 0, \ldots, n$

Interpolation par fonctions splines

Soit $a = x_0 < x_1 < \ldots < x_n = b$ une subdivision de l'intervalle $[a, b]$.

Définition

On appelle spline cubique interpolant f une fonction s satisfaisant les conditions :

1. $s_{|(x_{i-1},x_i)}$ est un polynôme de degré 3, pour tout $i = 1, \ldots, n$
2. $s(x_i) = f(x_i)$ pour tout $i = 0, \ldots, n$
3. s est une fonction de classe $\mathcal{C}^2$.

Ceci veut dire qu'on a les conditions :

$$\begin{aligned}
&s(x_i^-) = f(x_i) && i = 1, \dots, n-1, \\
&s(x_i^+) = f(x_i) && i = 1, \dots, n-1, \\
&s(x_0) = f(x_0), && \\
&s(x_n) = f(x_n), && \\
&s'(x_i^-) = s'(x_i^+) && i = 1, \dots, n-1, \\
&s''(x_i^-) = s''(x_i^+) && i = 1, \dots, n-1
\end{aligned}$$

Ceci veut dire qu'on a les conditions :

$$\begin{aligned}
&s(x_i^-) = f(x_i) && i = 1, \dots, n-1, \\
&s(x_i^+) = f(x_i) && i = 1, \dots, n-1, \\
&s(x_0) = f(x_0), && \\
&s(x_n) = f(x_n), && \\
&s'(x_i^-) = s'(x_i^+) && i = 1, \dots, n-1, \\
&s''(x_i^-) = s''(x_i^+) && i = 1, \dots, n-1
\end{aligned}$$

On a donc

$$2(n-1) + 2 + 2(n-1) = 4n-2 \quad \text{conditions}$$

Ceci veut dire qu'on a les conditions :

$$
\begin{aligned}
& s(x_i^-) = f(x_i) && i = 1, \dots, n-1, \\
& s(x_i^+) = f(x_i) && i = 1, \dots, n-1, \\
& s(x_0) = f(x_0), && \\
& s(x_n) = f(x_n), && \\
& s'(x_i^-) = s'(x_i^+) && i = 1, \dots, n-1, \\
& s''(x_i^-) = s''(x_i^+) && i = 1, \dots, n-1
\end{aligned}
$$

On a donc

$$2(n-1) + 2 + 2(n-1) = 4n - 2 \quad \text{conditions}$$

Or il faut trouver sur chaque intervalle 4 coefficients

Ceci veut dire qu'on a les conditions :

$$
\begin{aligned}
&s(x_i^-) = f(x_i) && i = 1, \dots, n-1, \\
&s(x_i^+) = f(x_i) && i = 1, \dots, n-1, \\
&s(x_0) = f(x_0), && \\
&s(x_n) = f(x_n), && \\
&s'(x_i^-) = s'(x_i^+) && i = 1, \dots, n-1, \\
&s''(x_i^-) = s''(x_i^+) && i = 1, \dots, n-1
\end{aligned}
$$

On a donc

$$2(n-1) + 2 + 2(n-1) = 4n - 2 \quad \text{conditions}$$

Or il faut trouver sur chaque intervalle 4 coefficients . . . Il manque donc 2 conditions

- Si on impose

- Si on impose

$$s''(x_0^+) = s''(x_n^-) = 0$$

- Si on impose

$$s''(x_0^+) = s''(x_n^-) = 0$$

alors la spline s est complètement déterminée et s'appelle Spline naturelle.

- Si on impose

$$s''(x_0^+) = s''(x_n^-) = 0$$

alors la spline s est complètement déterminée et s'appelle Spline naturelle.

- On peut aussi imposer

- Si on impose

$$s''(x_0^+) = s''(x_n^-) = 0$$

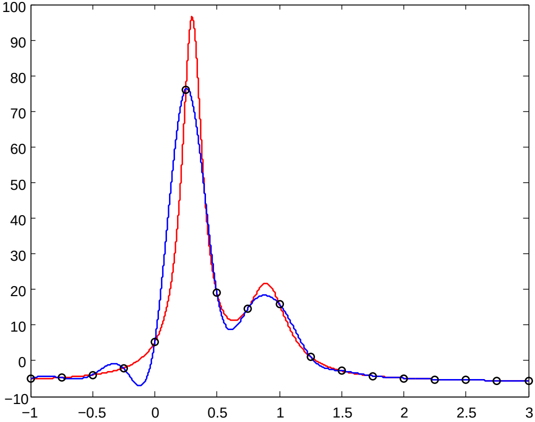
alors la spline s est complètement déterminée et s'appelle Spline naturelle.

- On peut aussi imposer

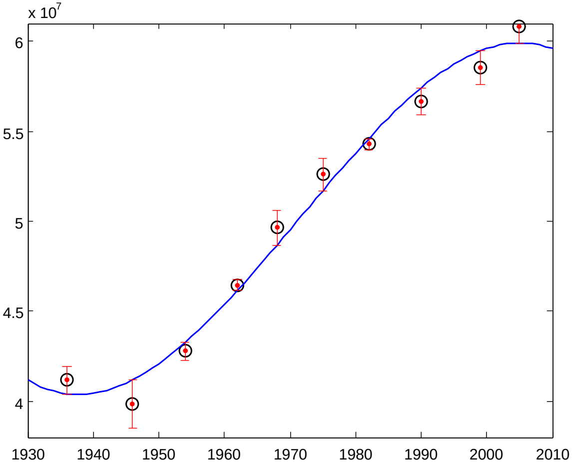
$$s'''(x_2^-) = s'''(x_2^+) \quad \text{et} \quad s'''(x_{n-1}^-) = s'''(x_{n-1}^+)$$

Interpolation par spline de la fonction

$$f(x) = \frac{1}{(x-0.3)^2+0.01} + \frac{1}{(x-0.9)^2+0.04} - 6$$



On reprend l'exemple de la population française. On veut décrire l'évolution de cette population à l'aide d'un polynôme de degré 3.



Approximation au sens des moindres carrés

Soient m points distincts (résultats de mesures par exemple)

$$\begin{matrix} x_1 & x_2 & \dots & x_m \\ y_1 & y_2 & \dots & y_m \end{matrix}$$

On veut construire une courbe qui passe aussi près que possible des valeurs y_i.

Approximation au sens des moindres carrés

Soient m points distincts (résultats de mesures par exemple)

$$\begin{array}{cccc} x_1 & x_2 & \dots & x_m \\ y_1 & y_2 & \dots & y_m \end{array}$$

On veut construire une courbe qui passe aussi près que possible des valeurs y_i.
On peut, par exemple, chercher une représentation des données en cherchant un polynôme de degré $n < m - 1$ qui approche au mieux les données.

Soient m points distincts (résultats de mesures par exemple)

$$\begin{array}{cccc} x_1 & x_2 & \dots & x_m \\ y_1 & y_2 & \dots & y_m \end{array}$$

On veut construire une courbe qui passe aussi près que possible des valeurs y_i.
On peut, par exemple, chercher une représentation des données en cherchant un polynôme de degré $n < m - 1$ qui approche au mieux les données.

Définition

On appelle polynôme aux moindres carrés de degré n, tout polynôme $f(x)$ vérifiant

$$\sum_{i=1}^{m} |y_i - f(x_i)|^2 \le \sum_{i=1}^{m} |y_i - p(x_i)|^2 \qquad \forall\ p \in \mathbb{P}_n.$$

Approximation au sens des moindres carrés

Soient m points distincts (résultats de mesures par exemple)

$$\begin{array}{cccc} x_1 & x_2 & \dots & x_m \\ y_1 & y_2 & \dots & y_m \end{array}$$

On veut construire une courbe qui passe aussi près que possible des valeurs y_i.
On peut, par exemple, chercher une représentation des données en cherchant un polynôme de degré $n < m - 1$ qui approche au mieux les données.

DÉFINITION

On appelle polynôme aux moindres carrés de degré n, tout polynôme $f(x)$ vérifiant

$$\sum_{i=1}^{m} |y_i - f(x_i)|^2 \leq \sum_{i=1}^{m} |y_i - p(x_i)|^2 \qquad \forall\, p \in \mathbb{P}_n.$$

Le polynôme aux moindres carrés est donc le polynôme de degré n qui, parmi tous les polynômes de degré n, minimise la distance aux données.

Notons que si $n = m - 1$ alors f est le polynôme d'interpolation de Lagrange aux points $x_1, \dots, x_m$.

Notons que si $n = m - 1$ alors f est le polynôme d'interpolation de Lagrange aux points $x_1, \ldots, x_m$.

Notons

$$\Phi(a_0, a_1, \ldots, a_n) = \sum_{i=1}^{m} \left(y_i - (a_0 + a_1 x_i + \ldots a_n x_i^n)\right)^2.$$

Les coefficients du polynôme solution vérifient les relations :

$$\frac{\partial \Phi}{\partial a_k} = 0 \qquad 0 \leq k \leq n.$$

Cela revient à dire que nous cherchons un minimum de la fonction Φ.

Notons que si $n = m - 1$ alors f est le polynôme d'interpolation de Lagrange aux points $x_1, \dots, x_m$.

Notons

$$\Phi(a_0, a_1, \dots, a_n) = \sum_{i=1}^{m} \left(y_i - (a_0 + a_1 x_i + \dots a_n x_i^n)\right)^2.$$

Les coefficients du polynôme solution vérifient les relations :

$$\frac{\partial \Phi}{\partial a_k} = 0 \qquad 0 \leq k \leq n.$$

Cela revient à dire que nous cherchons un minimum de la fonction Φ.

Ceci nous donne un système de $n+1$ relations linéaires entre les a_k. Notons

$$\mathbf{B} = \begin{pmatrix} 1 & x_1 & \dots & x_1^n \\ 1 & x_2 & \dots & x_2^n \\ \vdots & & & \vdots \\ 1 & x_m & \dots & x_m^n \end{pmatrix}$$

Ainsi $\mathbf{B}$ est une matrice rectangulaire à m lignes et $n+1$ colonnes.

Théorème

Une condition nécessaire et suffisante pour que $\mathbf{a} \in \mathbb{R}^{n+1}$ réalise le minimum de la fonction Φ est que

$$\mathbf{B}^T\mathbf{B}\,\mathbf{a} = \mathbf{B}^T\mathbf{y}.$$

Théorème

Une condition nécessaire et suffisante pour que $\mathbf{a} \in \mathbb{R}^{n+1}$ réalise le minimum de la fonction Φ est que

$$\mathbf{B}^T\mathbf{B}\,\mathbf{a} = \mathbf{B}^T\mathbf{y}.$$

Ainsi l'approximation par moindres carrés se ramène à la résolution d'un système linéaire.

Théorème

Une condition nécessaire et suffisante pour que $\mathbf{a} \in \mathbb{R}^{n+1}$ réalise le minimum de la fonction Φ est que

$$\mathbf{B}^T\mathbf{B}\,\mathbf{a} = \mathbf{B}^T\mathbf{y}.$$

Ainsi l'approximation par moindres carrés se ramène à la résolution d'un système linéaire.

Remarques

- La matrice $\mathbf{B}^T\mathbf{B}$ est symétrique définie positive

Théorème

Une condition nécessaire et suffisante pour que $\mathbf{a} \in \mathbb{R}^{n+1}$ réalise le minimum de la fonction Φ est que

$$\mathbf{B}^T\mathbf{B}\,\mathbf{a} = \mathbf{B}^T\mathbf{y}.$$

Ainsi l'approximation par moindres carrés se ramène à la résolution d'un système linéaire.

Remarques

- La matrice $\mathbf{B}^T\mathbf{B}$ est symétrique définie positive
- On peut utiliser d'autres fonctions d'approximation que des polynômes : Toute combinaison linéaire de fonctions linéairement indépendantes est possible

Théorème

Une condition nécessaire et suffisante pour que $\mathbf{a} \in \mathbb{R}^{n+1}$ réalise le minimum de la fonction Φ est que

$$\mathbf{B}^T\mathbf{B}\,\mathbf{a} = \mathbf{B}^T\mathbf{y}.$$

Ainsi l'approximation par moindres carrés se ramène à la résolution d'un système linéaire.

Remarques

- La matrice $\mathbf{B}^T\mathbf{B}$ est symétrique définie positive
- On peut utiliser d'autres fonctions d'approximation que des polynômes : Toute combinaison linéaire de fonctions linéairement indépendantes est possible
- Dans la pratique, on utilise des solveurs spécifiques pour le système linéaire, celui-ci étant mal conditionné.

Un exemple de calcul

On considère les données suivantes :

$$\begin{aligned} x_1 &= 0, & x_2 &= 1, & x_3 &= 2, \\ y_1 &= 2, & y_2 &= 2, & y_3 &= 8. \end{aligned}$$

Un exemple de calcul

On considère les données suivantes :

$$x_1 = 0, \qquad x_2 = 1, \qquad x_3 = 2,$$
$$y_1 = 2, \qquad y_2 = 2, \qquad y_3 = 8.$$

On recherche le polynôme aux moindres carrés de degré 1 associé à ces points (Régression linéaire). On recherche donc les coefficients a_0, a_1 du polynôme $a_0 + a_1 x$.

Un exemple de calcul

On considère les données suivantes :

$$x_1 = 0, \qquad x_2 = 1, \qquad x_3 = 2,$$
$$y_1 = 2, \qquad y_2 = 2, \qquad y_3 = 8.$$

On recherche le polynôme aux moindres carrés de degré 1 associé à ces points (Régression linéaire). On recherche donc les coefficients a_0, a_1 du polynôme $a_0 + a_1 x$.
Ainsi

$$\begin{aligned}\Phi(a_0, a_1) &= \sum_{i=1}^{3} \big(y_i - (a_0 + a_1 x_i)\big)^2 \\ &= (2 - a_0)^2 + \big(2 - (a_0 + a_1)\big)^2 + \big(8 - (a_0 + 2a_1)\big)^2\end{aligned}$$

POLYTECH CLERMONT-FERRAND

Un exemple de calcul

On considère les données suivantes :

$$x_1 = 0, \qquad x_2 = 1, \qquad x_3 = 2,$$
$$y_1 = 2, \qquad y_2 = 2, \qquad y_3 = 8.$$

On recherche le polynôme aux moindres carrés de degré 1 associé à ces points (Régression linéaire). On recherche donc les coefficients a_0, a_1 du polynôme $a_0 + a_1 x$. Ainsi

$$\begin{aligned}\Phi(a_0, a_1) &= \sum_{i=1}^{3} \big(y_i - (a_0 + a_1 x_i)\big)^2 \\ &= (2 - a_0)^2 + \big(2 - (a_0 + a_1)\big)^2 + \big(8 - (a_0 + 2a_1)\big)^2\end{aligned}$$

On calcule

$$\frac{\partial \Phi}{\partial a_0} = 6a_0 + 6a_1 - 24$$
$$\frac{\partial \Phi}{\partial a_1} = 6a_0 + 10a_1 - 36$$

En imposant

$$\frac{\partial \Phi}{\partial a_0} = \frac{\partial \Phi}{\partial a_1} = 0,$$

on obtient le système linéaire :

$$\begin{aligned} 6a_0 + 6a_1 &= 24 \\ 6a_0 + 10a_1 &= 36 \end{aligned}$$

En imposant

$$\frac{\partial \Phi}{\partial a_0} = \frac{\partial \Phi}{\partial a_1} = 0,$$

on obtient le système linéaire :

$$\begin{aligned} 6a_0 + 6a_1 &= 24 \\ 6a_0 + 10a_1 &= 36 \end{aligned}$$

La solution est alors $a_0 = 1$, $a_1 = 3$.

En imposant

$$\frac{\partial \Phi}{\partial a_0} = \frac{\partial \Phi}{\partial a_1} = 0,$$

on obtient le système linéaire :

$$\begin{aligned} 6a_0 + 6a_1 &= 24 \\ 6a_0 + 10a_1 &= 36 \end{aligned}$$

La solution est alors $a_0 = 1$, $a_1 = 3$. La droite est donc d'équation

$$y = 1 + 3x$$