



UFR des sciences

Université de Picardie Jules-Verne

Master Analyse Appliquée et Modélisation (AAM)

Mini Projet

Présenté par :

Hicham Abara Akdi

Encadrant :

Mr. Jean-Paul Chehab

Année universitaire 2025/2026

Table des matières

Introduction	2
0.1 Notions	4
1 B-Splines	6
1.1 Motivation	6
1.2 B-splines	7
1.2.1 Construction des B-splines	7
1.2.2 Quelques propriétés des B-splines	10
1.3 Courbes B-splines	11
1.3.1 Définition et propriétés	11
1.3.2 Exemple	11
2 Régularisation	15
2.1 Méthode des moindres carrés	15
2.2 Moindres carrés pondérés et localisés	17
2.2.1 Moindres carrés pondérés	17
2.2.2 Moindres carrés localisés	17
2.3 Moindre carrés régularisés	17
2.3.1 Exemple [4]	20

Introduction

Dans ce chapitre, nous nous appuyons en grande partie sur l'ouvrage : A Toolbox for Digital Twins : From Model-Based to Data-Driven de M. Asch et al., en particulier sur les pages 554–561 [1].

La caractéristique essentielle des données fonctionnelles est qu'elles proviennent de fonctions sous-jacentes (par exemple des données mesurées issues d'un système dynamique pour lequel il existe généralement une description en termes d'équations différentielles, et ces équations possèdent des solutions - possiblement inconnues - régulières et différentiables). Cependant, ces données peuvent être bruitées, multiscalaires et multiphysiques, et il se peut que nous n'ayons pas accès à la forme mathématique de leur génération. Dans ce cas, l'analyse de données fonctionnelles constitue un excellent outil de modélisation et d'analyse puisqu'elle tente de modéliser les données à l'aide de fonctions bien définies, généralement de forme polynomiale.

Supposons que la fonction x génère les données observées y , et que x est régulière. On considère alors $x_i(t)$ pour $i = 1, \dots, n$ comme des réalisations de cette fonction, et y_i comme les mesures correspondantes, affectées par du bruit.

$$y_i = x_i + \epsilon_i$$

ϵ_i est un terme de bruit. On a N mesures, ce qui donne (t_i, y_i) paires, avec

$$y_i = y(t_i) + \text{bruit}$$

Notre but est d'utiliser y_i pour estimer la fonction régulière x ainsi que ses dérivées avec le bruit.

Pour obtenir des données fonctionnelles à partir de données discrètes, il y a deux étapes :

-
1. Utiliser des méthodes d'expansion en base pour représenter des fonctions en temps continu.
 2. Réduire le bruit dans les mesures en effectuant une régularisation.

0.1 Notions

La moyenne d'échantillon fonctionnelle est définie comme la moyenne point par point des fonctions observées dans l'échantillon.

$$\bar{x}(t) = \frac{1}{n} \sum_{i=1}^n x_i(t)$$

La variance d'échantillon fonctionnelle :

$$s^2(t) = \frac{1}{n-1} \sum_{i=1}^n [x_i(t) - \bar{x}(t)]^2$$

covariance fonctionnelle :

$$v(s, t) = \frac{1}{n-1} \sum_{i=1}^n [x_i(t) - \bar{x}(t)][x_i(s) - \bar{x}(s)]$$

Bibliographie

- [1] M. Asch, C. G. Borzi, M. Chen, et al., *A Toolbox for Digital Twins : From Model-Based to Data-Driven*, SIAM, Philadelphia, 2022, pp. 554–561.

Chapitre 1

B-Splines

Dans ce chapitre, nous nous appuyons principalement sur deux ressources : le document de Pierre Pansu consacré aux courbes B-splines [1] et le support de cours sur l'approximation de fonctions proposé par Grenoble INP Pagora [2].

1.1 Motivation

Nous cherchons à approximer notre fonction $y(t)$. Cependant, une régression linéaire peut ne pas être adaptée à la forme de la courbe. Nous choisissons donc de l'approximer à l'aide d'un polynôme, plus flexible et capable de mieux représenter sa variabilité, on a alors :

$$x(t) \approx P_n$$

avec P_n un polynôme de degré n .

Phénomène de Runge [2]

Considérons la fonction :

$$f(x) = \frac{1}{1 + 25x^2}, x \in [-1, 1]$$

.

Runge (1856–1927) a montré que si cette fonction est interpolée aux points équidistants x_i entre -1 et 1

$$x_i = -1 + (i - 1)\frac{2}{n}, i = 0, \dots, n$$

par un polynôme P_n de degré $\leq n$ alors

$$\lim_{n \rightarrow \infty} \left(\max_{-1 \leq x \leq 1} |f(x) - P_n(x)| \right) = \infty$$

Lorsqu'on augmente le nombre de points, on constate que le polynôme se met à osciller fortement entre les points x_i avec une amplitude de plus en plus grande. Pour contourner le

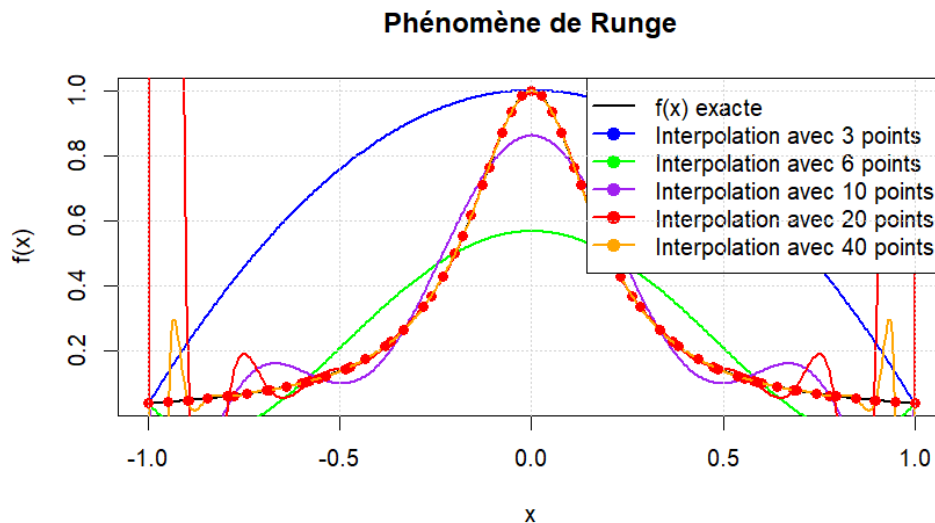


FIGURE 1.1 – Illustration du phénomène de Runge.

phénomène de Runge, on introduit la notion des B-splines.

1.2 B-splines

1.2.1 Construction des B-splines

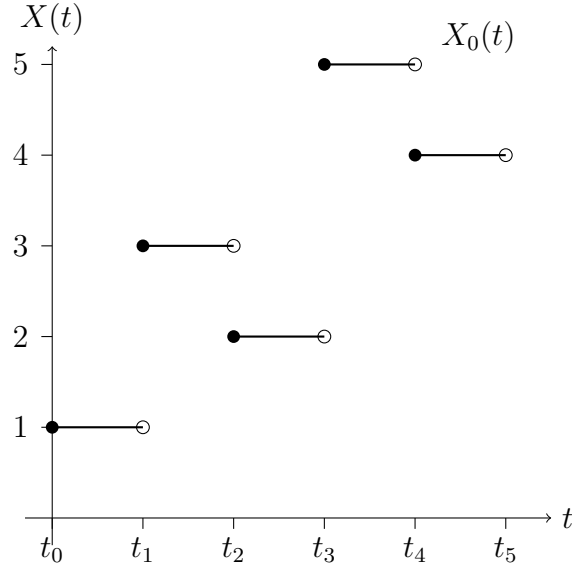
On se donne une suite de points $t_0 \leq \dots \leq t_m$ de la droite réelle, appelés **noeuds**. Le vecteur (t_1, \dots, t_m) s'appelle le vecteur des noeuds. Certains noeuds peuvent être confondus. Si r noeuds sont égaux à un réel τ , on dit que τ est de multiplicité r .

On se donne d'autre part des points P_0, \dots, P_m dans \mathbb{R}^n appelés **points de contrôle** qui forment ensemble le **polygone de contrôle**. On l'imagine comme une courbe $t \mapsto X_0(t)$ qui saute d'un point à l'autre aux temps t_i , i.e. :

$$X_0(t) = \sum_{i=1}^m \mathbf{1}_{[t_i, t_{i+1}[}(t) P_i$$

Si $t_i = t_{i+1}$ le sommet P_i est ignoré.

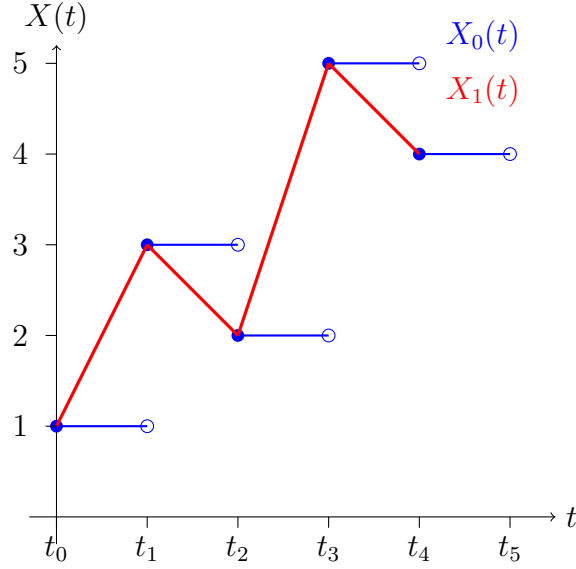
X_0 est donc une fonction en escalier comme montre la figure suivante.



On cherche à approcher cette courbe discontinue par une courbe plus régulière. La première étape consiste à faire passer une ligne polygonale par les points P_i , i.e. lorsque t varie entre deux noeuds t_i et t_{i+1} , $X_1(t)$ décrit le segment $[P_{i-1}, P_i]$ à vitesse constante. On trouve la formule :

$$X_1(t) = \sum_{i=1}^m \left(\left(1 - \frac{t - t_i}{t_{i+1} - t_i}\right) P_{i-1} + \frac{t - t_i}{t_{i+1} - t_i} P_i \right) \mathbf{1}_{[t_i, t_{i+1}[}(t)$$

Si $t_i = t_{i+1}$ la courbe saute de P_{i-1} à P_i en t_i . Si les noeuds sont tous distincts, la courbe obtenue est continue mais non dérivable en général. Ses composantes sont des fonctions linéaires par morceaux.



On suppose tout les noeuds distincts, la prochaine étape est de construire un courbe X_2 de C^1 , elle est quadratique par morceaux (le prix pour qu'elle soit de C^1). Elle ne passe plus par les sommets P_i , mais conserve une proximité au polygone en un sens différent : si t est compris entre les noeuds t_i et t_{i+1} , $X_2(t)$ est dans l'encloppe convexe des sommets P_{i-2} , P_{i-1} et P_i .

Question : Comment trouver X_2 et plus généralement X_k pour $k \geq 2$?

Supposons les points P_i affinement indépendants. Alors X_{k-1} s'écrit de manière unique

$$X_{k-1}(t) = \sum_i B_{i,k-1}(t) P_i$$

$$— B_{i,k-1} \geq 0.$$

$$— \sum_i B_{i,k-1} = 1.$$

Pour gagner un degré de différentiabilité, on remplace P_i par $t \mapsto P_i(t)$, qui peut être vu comme un point mobile le long du polygone de P_{i-1} et P_i pendant l'intervalle de temps $[t_i, t_{i+k}]$. On pose alors :

$$X_k(t) = \sum_i B_{i,k-1}(t) \left(\left(1 - \frac{t - t_i}{t_{i+k} - t_i}\right) P_{i-1} + \frac{t - t_i}{t_{i+k} - t_i} P_i \right)$$

Cela donne pour les fonctions $B_{i,k}$ la relation de récurrence

$$B_{i,k}(t) = \frac{t - t_i}{t_{i+k} - t_i} B_{i,k-1}(t) + \left(1 - \frac{t - t_{i+1}}{t_{i+k+1} - t_{i+1}}\right) B_{i+1,k-1}(t)$$

On se donne une suite de noeuds $t_0 \leq t_1 \leq \dots \leq t_m$ sur la droite réelle.

Notation : Soit $j = 1, \dots, m+1-i$. Si $t_i < t_{i+1}$, on note

$$\omega_{i,j}(t) = \frac{t - t_i}{t_{i+j} - t_i}$$

Si $t_i = t_{i+1}$, on pose $\omega_{i,j} = 0$

Définition 1.2.1. On définit par récurrence sur k les fonctions B-splines $B_{i,k}$ pour $i = 0, \dots, m-k+1$ par les relations suivantes.

$$B_{i,0}(x) = \begin{cases} 1, & t_i \leq x < t_{i+1} \\ 0, & \text{sinon} \end{cases}$$

et pour $k \geq 1$

$$B_{i,k}(x) = \omega_{i,k}(t)B_{i,k-1}(x) + (1 - \omega_{i+1,k}(t))B_{i+1,k-1}(x)$$

1.2.2 Quelques propriétés des B-splines

Proposition 1.2.1. *Propriétés générales de fonctions B-splines.*

1. La fonction $B_{j,k}$ est sur chaque intervalle $[t_i, t_{i+1}[$ un polynôme de degré $\leq k$.
2. La fonction $B_{i,k}$ s'annule en dehors de l'intervalle $[t_i, t_{i+k}[$
3. La fonction $B_{i,k}$ s'annule en t_i , sauf cas : $t_i = t_{i+1} = \dots = t_{i+k} < t_{i+k+1}$ auquel cas $B_{i,k}(t_i) = 1$
4. $0 < B_{i,k}(t) \leq 1$ pour $t \in [t_i, t_{i+k}[$.
5. Sur l'intervalle $]t_i, t_{i+k}[$, la fonction $B_{i,k}$ ne prend la valeur 1 que si $t_{i+1} = \dots = t_{i+k}$ et en ce point seulement.
6. $\sum_{i=0}^{m-k-1} B_{i,k} \equiv 1$, sur l'intervalle $[t_k, t_{m-k}[$.
7. La fonction $B_{i,k}$ est C^∞ à droite de chaque point.
8. Au voisinage d'un noeud de multiplicité r , la fonction $B_{i,k}$ est de C^{k-r} .

Démonstration. Voir [1]. □

1.3 Courbes B-splines

1.3.1 Définition et propriétés

Définition 1.3.1. On se donne un vecteur de noeuds $t = (t_1 = \dots = t_m)$ et des points P_1, \dots, P_m appelés points de contrôle qui forment ensemble le polygone de contrôle. La courbe B-spline de degré k associée au vecteur de noeuds t et au polygone de contrôle $P = (P_1, \dots, P_m)$ est

$$t \mapsto X_k(t) = \sum_i B_{i,k}(t) P_i$$

Remarque. On dit qu'une courbe B-spline est visé aux extrémités (clamped) si elle est de degré k et ses noeuds t_0 et t_m sont de multiplicité $k + 1$. L'utilité de cette notion vient du fait qu'on doit imposer que $X_k(t_0) = P_0$ et $X_k(t_m) = P_m$.

Théorème 1.3.1. La k -ième B-spline $t \mapsto X_k(t)$ à les propriétés suivantes.

1. Les composants de $X_k(t)$ sont sur chaque intervalle $[t_i, t_{i+1}[$ des polynômes de degré k .
2. En un noeud de multiplicité r , la courbe est de classe C^{k-r} .
3. Si $t \in [t_i, t_{i+1}[$, $X_k(t)$ ne dépend que des points de contrôle P_{i-k}, \dots, P_i et se trouve dans l'enveloppe convexe de ces points.
4. Si t_i est un noeud simple et $k \geq 1$, $X_k(t_i)$ ne dépend que des points de contrôle P_{i-k}, \dots, P_{i-1} et se trouve dans l'enveloppe convexe de ces points.
5. Si $t_i = \dots = t_{i+k} < t_{i+k+1}$ est un noeud de multiplicité $k + 1$, alors $X_k(t_i) = P_i$ et $X'_k(t_i) = \frac{k}{t_{i+k+1} - t_i} (P_{i+1} - P_i)$.

Démonstration. Voir [1]. □

1.3.2 Exemple

On retourne à l'exemple

$$f(x) = \frac{1}{1 + 25x^2}, x \in [-1, 1]$$

et avec des points intérieurs. On l'approche par une fonction de B-splines de degré ($k = 0$), des B-splines linéaires ($k = 1$), des B-splines quadratiques ($k = 2$) et cubiques ($k = 3$).

B-splines de degré ($k = 0$)

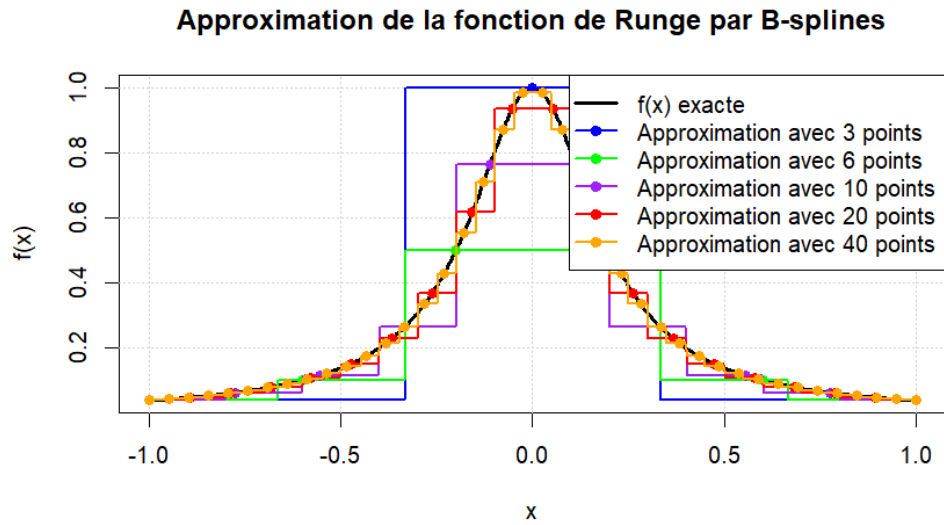


FIGURE 1.2 – B-splines de degré 0.

B-splines linéaires

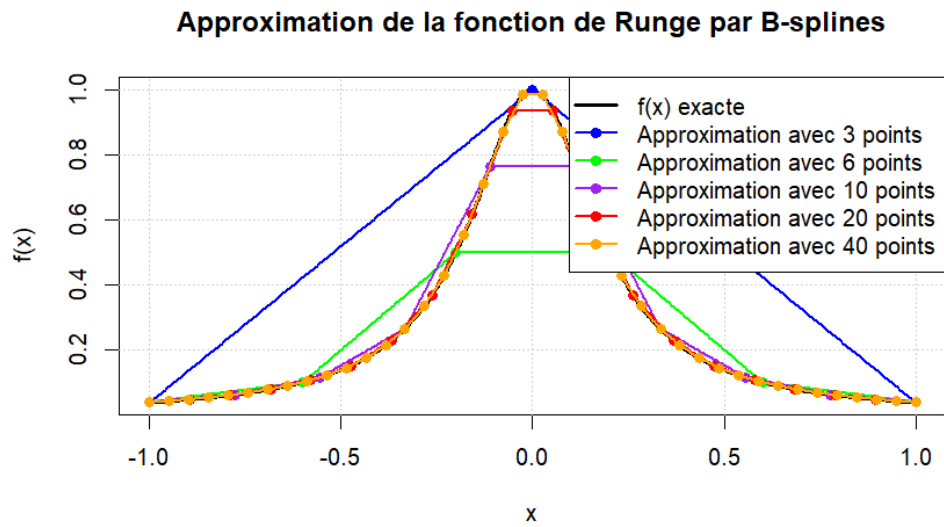


FIGURE 1.3 – B-splines de degré 1.

B-splines quadratiques

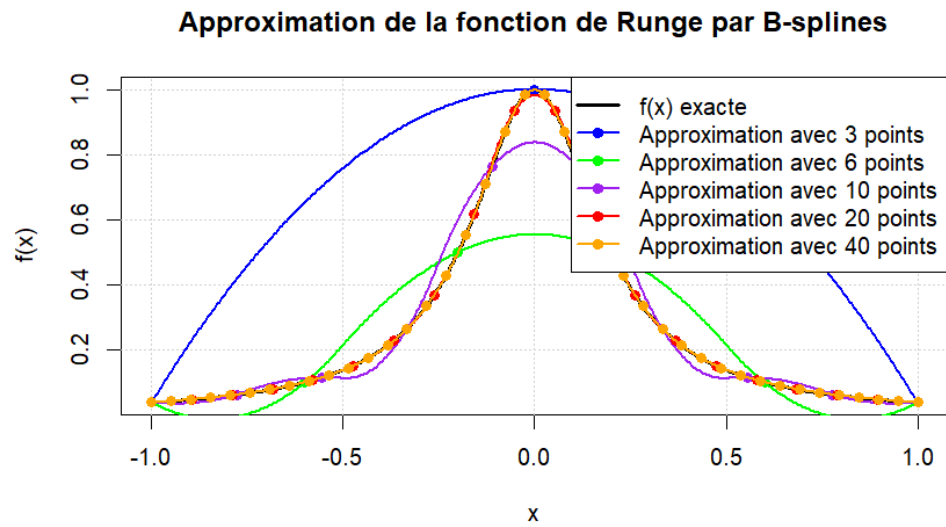


FIGURE 1.4 – B-splines de degré 2.

B-splines cubiques

Le minimum de points dans ce cas doivent être 4.

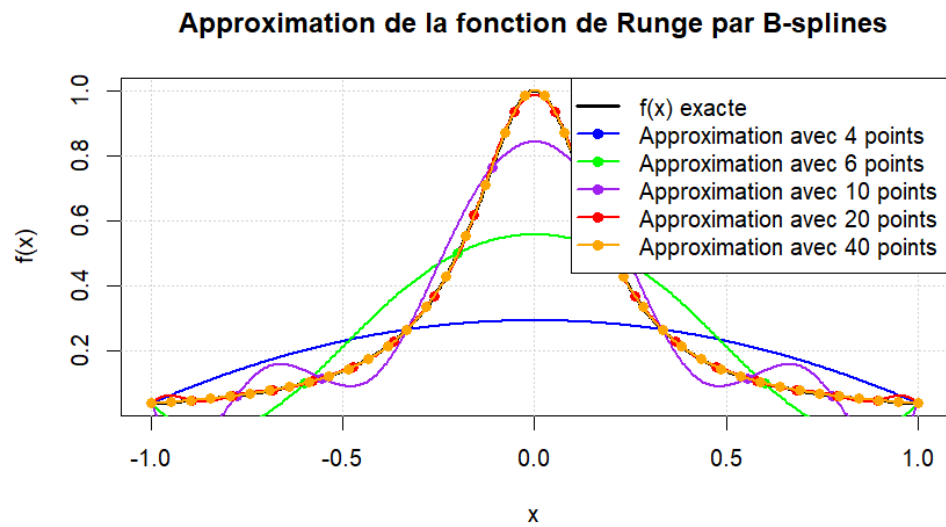


FIGURE 1.5 – B-splines de degré 3.

Bibliographie

- [1] Pierre Pansu, *Courbes B-splines*, Document du Laboratoire de Mathématiques d'Orsay,
https://www.imo.universite-paris-saclay.fr/~pierre.pansu/web_maitrise/bsplines.pdf.
- [2] Analyse numérique : Approximation de fonctions, Grenoble INP Pagora, Support de cours au format Beamer, 2013, https://team.inria.fr/airsea/files/2013/01/Cours_approx_fonc.pdf.
- [3] A. Popier, *Mouvement Brownien, Intégrale Stochastique*, Séminaire des doctorants, Bordeaux, mercredi 12 mars 2003.
<https://perso.univ-lemans.fr/~apopier/documents/resumesemindoct.pdf>

Chapitre 2

Régularisation

Dans ce chapitre, nous nous appuyons en grande partie sur l'ouvrage : A Toolbox for Digital Twins : From Model-Based to Data-Driven de M. Asch et al., en particulier sur les pages 555–557 [1].

Maintenant qu'on a construit les B-splines, on a la fonction suivante

$$x(t) \approx \sum_{k=1}^K c_k \phi_k(t)$$

Avec $(\phi_k)_{k \geq 1}$ des fonctions B-splines, reste à déterminer les paramètres c_k .

2.1 Méthode des moindres carrés

On utilise la méthode des moindres carrés (Least Squared Methode), on rappelle que $(y_i)_{i \geq 1}$ sont les points qu'on veut approcher.

$$RSS(x) = \sum_{j=1}^n [y_j - x(t_j)]^2$$

En posant $y^T = (y_1 \ \dots \ y_n) \in \mathcal{M}_{1,n}(\mathbb{R})$ et $\phi_c^T = (x(t_1) \ \dots \ x(t_n)) \in \mathcal{M}_{1,n}(\mathbb{R})$ on trouve

$$RSS(x) = (\phi_c - y)^T (\phi_c - y) \tag{RSS}$$

Soit la matrice $\phi \in \mathcal{M}_{n,K}(\mathbb{R})$

$$\phi = \begin{pmatrix} \phi_1(t_1) & \dots & \phi_K(t_1) \\ \phi_1(t_2) & \dots & \phi_K(t_2) \\ \dots & \dots & \dots \\ \phi_1(t_n) & \dots & \phi_K(t_n) \end{pmatrix}$$

et $c^T = (c_1 \dots c_K) \in \mathcal{M}_{1,n}(\mathbb{R})$, on a alors

$$\phi = \phi_c c$$

Remarque. La valeur optimale du vecteur c (qu'on note \hat{c}) pour minimiser (RSS) est

$$\hat{c} = (\phi^{-1} \phi)^{-1} \phi^T y$$

Démonstration.

$$\begin{aligned} J(c) &= (\phi c - y)^T (\phi c - y) \\ &= c^T \phi^T \phi c - 2y^T \phi c + y^T y \end{aligned}$$

On applique le gradient sur J par rapport à c , on a

$$\nabla J(c) = 2\phi^T \phi c - 2y^T \phi$$

Pour $c = \hat{c}$, $\nabla J(\hat{c}) = 0$, on trouve alors

$$2\phi^T \phi \hat{c} - 2y^T \phi = 0$$

$$\implies \hat{c} = (\phi^T \phi)^{-1} \phi^T y$$

□

Par conséquent, notre fonction régularisé est

$$\begin{aligned} \hat{y} &= \phi \hat{c} \\ &= \phi (\phi^T \phi)^{-1} \phi^T y \\ &= Hy \end{aligned}$$

2.2 Moindres carrés pondérés et localisés

La méthode précédente est utile quand les bruits ϵ_i sont indépendants, identiquement distribué de moyenne nulle et de variance constante, ce qui n'est toujours pas le cas.

2.2.1 Moindres carrés pondérés

On suppose maintenant que chaque observation y_i a une importance différente, on lui attribut un poids $w_i > 0$. On définit la matrice diagonale des poids

$$W = \text{diag}(w_1, \dots, w_n)$$

On a alors

$$RSS_W(x) = (\phi_c - y)^T W (\phi_c - y) \quad (\text{RSSW})$$

On reprenant les mêmes calculs de (RSS) sur (RSSW) on trouve

$$\hat{c} = (\phi^T W \phi)^{-1} \phi^T W y$$

On pose $w_i = \frac{1}{\sigma_i^2}$, inverse de la variance.

2.2.2 Moindres carrés localisés

Pour une régularisation locale autour d'un point t_i , on attribut des poids plus élevés aux points proches de t_i , càd plus t_j proche de t_i , plus le poids $w(t_j)$ est élevé, cela permet une meilleur estimation locale. Les poids sont de la forme

$$w_j(t) = g\left(\frac{t_i - t_j}{h}\right)$$

Avec $h > 0$ et g une fonction noyau (fonction positive et $x \mapsto g(x)$ décroît quand x s'éloigne du 0).

2.3 Moindre carrés régularisés

Revenons à (RSSW), on l'écrit sous forme d'une somme, on a

$$RSS_W(x) = \sum_{j=1}^n w_i [y_j - x(t_j)]^2$$

$$= \sum_{j=1}^n \frac{1}{\sigma_i^2} [y_j - x(t_j)]^2$$

On remarque que si l'un des variances est très petit, une grande perturbation arrive. Il y'a aussi le fait qu'il y a trop de paramètres c_k par rapport au nombre d'observations.

On introduit alors un terme qui pénalise le fait d'ajouter beaucoup de données, l'équation (RSSW) devient alors

$$RRSS_W(x) = (\phi_c - y)^T W (\phi_c - y) + \lambda \int |L(x)(t)|^2 dt \quad (\text{RRSSW1})$$

- L est la fonction de perte, c'est un opérateur linéaire (de la forme $\frac{d^m}{ds^m}$).
- λ paramètre de régularisation.
- $\lambda \rightarrow \infty$: Pas assez d'informations, il faut ajouter des données.
- $\lambda \rightarrow 0$: Trop d'informations, les coefficients sont potentiellement instables.

On doit donc balancer entre le degré de régularisation et l'insertion d'informations.

Déterminer L

On veut pénaliser la dérivation m -ième, on peut alors écrire la fonction de perte sous la forme

$$\begin{aligned} L_m(x) &= \int \left| \frac{d^m}{ds^m} x(s) \right|^2 ds \\ &= \int \left| \frac{d^m}{ds^m} c^T \phi(s) \right|^2 ds \\ &= c^T R_m c \end{aligned}$$

Avec $\phi(s)^T = (\phi_1(s) \dots \phi_K(s))$. L'équation (RRSSW1) devient alors

$$RRSS_m = (\phi_c - y)^T W (\phi_c - y) + \lambda c^T R_m c \quad (\text{RRSSm})$$

Avec

$$R_m = \int \frac{d^m}{ds^m} \phi(s) \frac{d^m}{ds^m} \phi^T(s) ds$$

La valeur optimale \hat{c} pour minimiser (RRSSm) est

$$\begin{aligned} \hat{c} &= (\phi^T W \phi + \lambda R_m)^{-1} \phi^T W y \\ &= S_{\phi, \lambda} y \end{aligned}$$

Déterminer λ

Remarque. On veut trouver une solution pour le problème matriciel suivant

$$\phi^T W \phi + \lambda R$$

Déterminer λ est un problème qui peut être décomposé en deux parties.

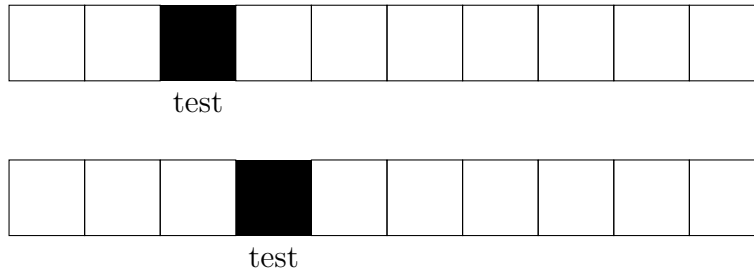
1. Déterminer un intervalle pour λ : $\lambda_{min} \leq \lambda \leq \lambda_{max}$.
2. Dans cet intervalle, trouver la valeur λ^* optimale.

Le choix de l'intervalle dépend du contexte, mais, en général, l'ampleur de λR doit être inférieure à 10^{10} fois celle de $\phi^T W \phi$ et dix fois plus grande que l'ampleur de $\phi^T W \phi$.

Dans les cas déterministes, on utilise la méthode *L-curve* [2]. Pour le cas statistique, on utilise la méthode de validation croisée.

Idée[3] :

On partitionne les données en données d'entraînement et de test, dans le cas extrême, pour n données, on prend $n - 1$ pour entraînement et une seule pour le test, on calcul l'erreur E_k , $k \in 1, 2 \dots n$ et on calcule la moyenne $CV(\lambda) = \frac{1}{n} \sum_{k=1}^n E_k$.



On refait cette méthode pour plusieurs valeurs de λ , on trouve alors la valeur optimale λ^* .

Problème de cette méthode :

- Très coûteuse au niveau du calcul.
- Elle a tendance à sous-régulariser les données, car elle privilégie des ajustements oscillants que l'on préférerait ignorer.

On utilise alors la méthode de validation croisée généralisée (Generalised Cross Validation)[1].

$$GCV(\lambda) = \frac{\frac{1}{n} RRS_{m,\lambda}}{|1 - n^{-1} \text{trace}(S_{\phi,\lambda})|^2}$$

2.3.1 Exemple [4]

En 1828, le botaniste Robert Brown observa le mouvement "aléatoire" d'un pollen face aux choc avec les molécules d'eau, le terme mouvement Brownien est donné aux trajectoires irrégulières de ce pollen. Mais le premier traitement mathématique rigoureux est dû à N. Wiener(1923,1924) qui a prouvé l'existence du mouvement Brownien.

On a besoin de quelques notions pour pouvoir définir un mouvement Brownien, soit alors un espace probabilisé (Ω, \mathcal{F}, P) .

Ω un ensemble, \mathcal{F} une tribu contenue dans l'ensemble des parties de Ω et P une probabilité sur la tribu \mathcal{F} .

Définition 2.3.1. Une **filtration** $\{\mathcal{F}_t; 0 \leq t < +\infty\}$ est une famille croissante de sous-tribus de \mathcal{F} : pour $0 \leq s \leq t < +\infty$, $\mathcal{F}_s \subseteq \mathcal{F}_t$.

Définition 2.3.2. Un **processus stochastique** X est la donnée de $\{X_t; 0 \leq t < +\infty\}$, où, à t fixé, X_t est une variable aléatoire définie sur (Ω, \mathcal{F}) à valeurs dans $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. Avec $\mathcal{B}(\mathbb{R}^d)$ la tribu borélienne de \mathbb{R}^d .

Définition 2.3.3. Un processus X est dit **mesurable** si l'application suivante

$$\begin{aligned} ([0, +\infty[\times \Omega, \mathcal{B}([0, +\infty[) \otimes \mathcal{F}) &\longrightarrow (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)) \\ (t, \omega) &\longmapsto X_t(\omega) \end{aligned}$$

est mesurable.

Un processus est dit **continu** si pour presque tout $\omega \in \Omega$, $t \mapsto X_t(\omega)$ est continue.

Définition 2.3.4. Un processus est dit **adapté** à la filtration $\{\mathcal{F}_t; 0 \leq t < +\infty\}$ si pour tout t , X_t est \mathcal{F}_t -mesurable.

On donne maintenant la définition du mouvement Brownien(MB).

Définition 2.3.5. Un **mouvement brownien de dimension k** $\{B_t, \mathcal{F}_t; 0 \leq t < +\infty\}$ est la donnée d'un processus mesurable B à valeurs dans \mathbb{R}^k , et d'une filtration, tel que B est adapté à $(\mathcal{F}_t)_{t \geq 0}$ est continu, et vérifie :

1. $B_0 = 0$ presque sûrement.

-
2. Pour $0 \leq s < t$, l'acrosissement $B_t - B_s$ est indépendant de \mathcal{F}_s .
 3. Pour $0 \leq s < t$, l'acrosissement $B_t - B_s$ suit une loi normale centrée, de matrice de covariance $\sqrt{t-s}Id_k$, et Id_k désigne la matrice identité de dimension k .

Approxiamtion du mouvement Brownien en utilisant la méthode L -curve[2]

Soit un mouvement Brownien, prenons alors n donées réels et soit $y \in \mathcal{M}_{n,1}$ vecteur de ces donées. On veut approcher ces donées par une fonction contrainte par p B-splines cubiques.

$$x(t) = \phi.c$$

- $\phi \in \mathcal{M}_{n,p}$ matrice des B-splines.
- $c \in \mathcal{M}_{p,1}$ vecteur de coefficients.

Prenons $W = I$ et supposons qu'on shouaite pénaliser la dérivée seconde, on alors

$$c^T R c = \int (x''(t))^2 dt$$

Trouver la valeur optimale de λ consiste à trouver la solution du problème suivant

$$\phi^T \phi + \lambda R$$

La première étape consiste à tracer, pour la variable λ , les deux fonctions suivante

$$f(\lambda) = \|\phi.c_\lambda - y\| \text{ et } g(\lambda) = \sqrt{c_\lambda^T . R . c_\lambda}$$

La valeur λ optimale se trouve au point où la courbure est maximale.

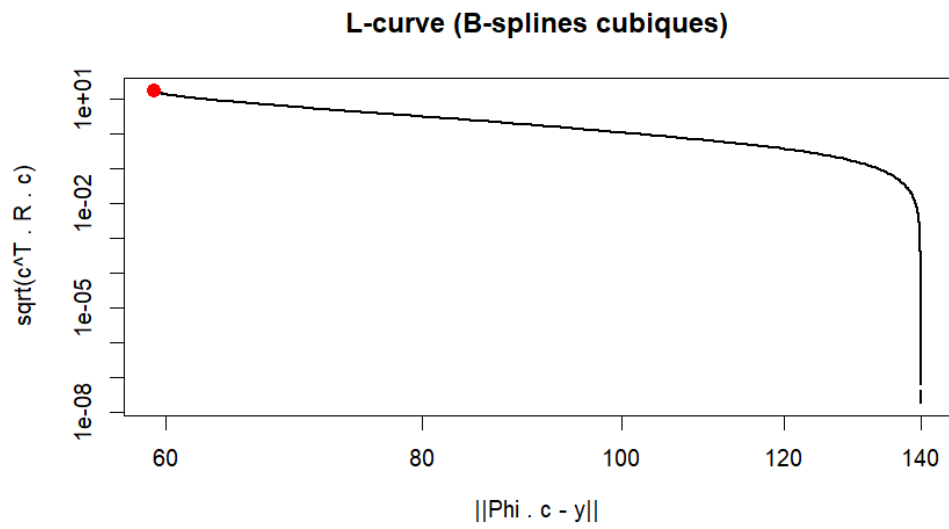


FIGURE 2.1 – Point rouge indique la valeur optimale de λ

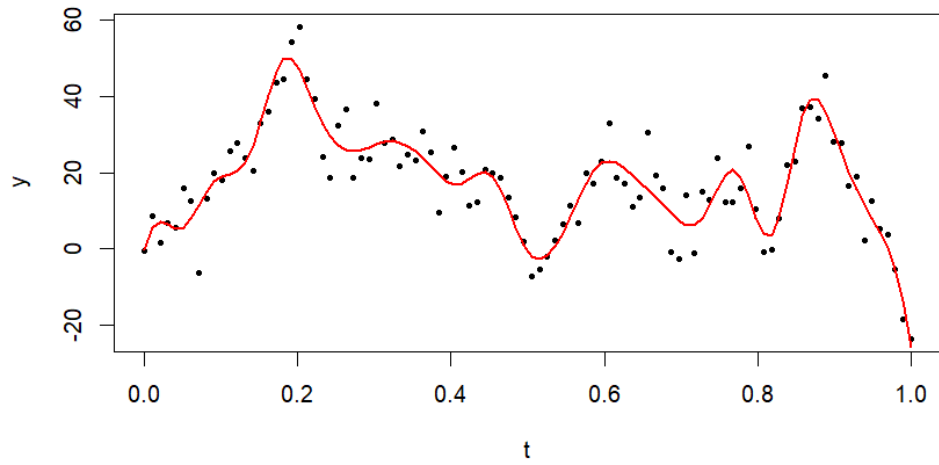


FIGURE 2.2 – Nuage de points des données et la courbe pénalisée

Bibliographie

- [1] M. Asch, C. G. Borzi, M. Chen, et al., *A Toolbox for Digital Twins : From Model-Based to Data-Driven*, SIAM, Philadelphia, 2022, pp. 555–557.
- [2] P. C. Hansen, *The L-curve and its use in the numerical treatment of inverse problems*, Technical University of Denmark, Lyngby. URL : <https://www.sintef.no/globalassets/project/evitameeting/2005/lcurve.pdf>
- [3] D. Kobak, *Machine Learning I : Regularization and cross-validation*. Eberhard Karls Universität Tübingen. Lecture handout, slide 23. Disponible en ligne : https://dkobak.github.io/teaching/ml1/kobak_ml1_lecture04_regularization_handout.pdf
- [4] A. Popier, *Mouvement Brownien, Intégrale Stochastique*, Séminaire des doctorants, Bordeaux, mercredi 12 mars 2003.
<https://perso.univ-lemans.fr/~apopier/documents/resumesemindoct.pdf>