

Lissages et espaces de Hilbert

1 Lissages de séries chronologique par moyenne mobile

En statistiques on a l'habitude de régulariser une série chronologique pour mieux interpréter les données et réduire les fluctuations. Soit $(x_i)_{i=0}^{N+1}$ une série chronologique (i représente le temps). On construit une nouvelle série $(y_j)_{j=1}^N$ par moyenne mobile d'ordre 3 à l'aide de la formule suivante pour $1 \leq j \leq N$: $y_j := \frac{x_{j-1} + x_j + x_{j+1}}{3}$. Si cette nouvelle série n'est pas assez régulière on peut itérer le procédé.

1. Calculez $\bar{y} := \frac{1}{N} \sum_{j=1}^N y_j$ en fonction de $\bar{x} := \frac{1}{N+2} \sum_{i=0}^{N+1} x_i$ et des termes de bords: x_0, x_1, x_N, x_{N+1} .
2. Comparez $Var(y)$ et $Var(x)$. On pourra commencer par le cas où $\bar{x} = 0 = \bar{y}$.

2 Suites lissées

Pour simplifier l'étude du lissage, on élimine les termes de bords en considérant des suites indicées dans \mathbb{Z} . Soit

$$\mathcal{M}: \quad \mathbb{C}^{\mathbb{Z}} \quad \rightarrow \quad \mathbb{C}^{\mathbb{Z}}$$

$$u = (u_n)_{n \in \mathbb{Z}} \mapsto v = (v_n)_{n \in \mathbb{Z}} \quad \text{avec} \quad v_n := \frac{u_{n-1} + u_n + u_{n+1}}{3}$$

1. On note $l^\infty(\mathbb{Z})$ l'ensemble des suites bornées de $\mathbb{C}^{\mathbb{Z}}$. On muni $l^\infty(\mathbb{Z})$ de sa norme naturelle: $\|u\|_\infty := \sup_{n \in \mathbb{Z}} |u_n|$.
 - (a) Vérifiez que \mathcal{M} est un endomorphisme continue de $l^\infty(\mathbb{Z})$. Calculer $\|\mathcal{M}\|_\infty := \sup_{u \in l^\infty(\mathbb{Z}) - \{0\}} \frac{\|\mathcal{M}(u)\|_\infty}{\|u\|_\infty}$.
 - (b) Donnez une condition nécessaire et suffisante sur u pour que $\|\mathcal{M}(u)\|_\infty = \|u\|_\infty$.
2. Soit $p > 3$. On note $l_p^2(\mathbb{Z})$ l'ensemble des suites de $\mathbb{C}^{\mathbb{Z}}$ p -périodiques: $u \in l_p^2(\mathbb{Z}) \iff u_{n+p} = u_n$ pour tout n . On notera $\langle u, v \rangle_p := \frac{1}{p} \sum_{n=1}^p u_n \bar{v}_n$, $\|u\|_{2,p} := \sqrt{\langle u, u \rangle_p}$ et $\langle u \rangle_p := \langle u, 1 \rangle_p$.
 - (a) Montrez que $\langle u, v \rangle_p$ munie $l_p^2(\mathbb{Z})$ d'une structure hermitienne. Donnez une base orthogonale et $\dim l_p^2(\mathbb{Z})$.
 - (b) Vérifiez que \mathcal{M} , restreint à $l_p^2(\mathbb{Z})$, est un endomorphisme symétrique défini positif de $l_p^2(\mathbb{Z})$.
 - (c) Calculer $\|\mathcal{M}\|_2 := \sup_{u \in l_p^2(\mathbb{Z}), \langle u, u \rangle_p = 1} \|\mathcal{M}(u)\|_{2,p}$.
 - (d) En déduire que la plus grande valeur propre de \mathcal{M} sur $l_p^2(\mathbb{Z})$ est 1. Calculer la dimension de l'espace propre associé à la valeur propre 1 dans $l_p^2(\mathbb{Z})$.
 - (e) En déduire que $\lim_{k \rightarrow +\infty} \mathcal{M}^k(u) = \langle u \rangle_p$, pour tout $u \in l_p^2(\mathbb{Z})$.
3. On note $l^2(\mathbb{Z})$ l'ensemble des suites de $\mathbb{C}^{\mathbb{Z}}$ dont le carré des modules sont sommables: $u \in l^2(\mathbb{Z}) \iff \sum_{n \in \mathbb{Z}} |u_n|^2$ converge. On note: $\langle u, v \rangle := \sum_{n \in \mathbb{Z}} u_n \bar{v}_n$, $\|u\|_2^2 := \langle u, u \rangle$.
 - (a) Vérifiez que $\langle u, v \rangle$ est un produit scalaire sur $l^2(\mathbb{Z})$ qui le muni d'une structure hilbertienne.
 - (b) Montrez que $\|\mathcal{M}(u)\|_2 \leq \|u\|_2$. A quelle condition a-t-on l'égalité?
 - (c) Trouver la $\lim_{k \rightarrow +\infty} \mathcal{M}^k(u)$ pour tout $u \in l^2(\mathbb{Z})$.
4. Pour $u \in l_p^2(\mathbb{Z})$, on note $\bar{u} := \langle u \rangle_p$ et $Var(u) := \overline{(u - \bar{u})^2}$.

(a) Démontrez et commentez cette égalité: $\overline{\mathcal{M}(u)} = \bar{u}$.

(b) Démontrez et commentez cette inégalité: $Var(\mathcal{M}(u)) \leq Var(u)$. Quand a-t-on égalité?

5. utilisation de modèle linéaire: En pratique, on peut considérer que la suite représente une échantillon bruité d'une fonction modèle f , i.e. $u_n = f(n\delta t) + \varepsilon_n$. Prenons $\delta t = 1$ sans perte de généralité et, pour simplifier, le bruit $(\varepsilon_n)_n \in l_p^2(\mathbb{Z})$ de moyenne nulle: $\bar{\varepsilon} = 0$. On se propose de voir sur quelques exemples dans quelle mesure le lissage affecte la fonction modèle. On notera $v := \mathcal{M}(u)$

(a) modèle linéaire: $f(x) := ax + b$, vérifiez que $v_n = f(n) + \eta_n$, avec, $\eta \in l_p^2(\mathbb{Z})$, $\bar{\eta} = 0$, $Var(\eta) \leq Var(\varepsilon)$.

(b) modèle polynomiale: étudiez l'effet du lissage si $f(x) = ax^2 + bx + c$, $ax^3 + bx^2 + cx + d$ ou un polynôme.

(c) modèle de croissance exponentielle: Commentez l'intérêt de ce lissage si $f(x) = \exp(x)$.

3 Lissage de fonction

Notations: Pour $T > 0$, on note $C_{T\#}^0$ l'espace des fonctions de $C^0(\mathbb{R}, \mathbb{C})$ périodique de période T ; et

$C_{T\#}^1 = C_{T\#}^0 \cap C^1(\mathbb{R}, \mathbb{C})$. Pour $k \in \mathbb{Z}$, on note: $e_k(x) := \exp(2ik\pi x)$ où $i^2 = -1$,

$\langle f, g \rangle := \int_0^1 f(x)\overline{g(x)}dx$, $\langle f \rangle := \langle f, 1 \rangle$, $\|f\|_2^2 := \langle f, f \rangle$, $c_k[f] := \langle f, e_k \rangle$. Soit $0 < h < 1$, on considère l'opérateur:

$$\begin{aligned} \mathcal{M}: C_{1\#}^0 &\rightarrow C_{1\#}^0 \\ f &\mapsto g \text{ avec } g(x) := \frac{f(x-h) + f(x) + f(x+h)}{3} \end{aligned}$$

1. Vérifiez que \mathcal{M} , restreint à $C_{1\#}^0$, est un endomorphisme symétrique défini positif de $C_{1\#}^0$.

2. Montrez que $\mathcal{M}(e_k) = \lambda_k e_k$ où λ_k est un réel que l'on encadrera.

3. En déduire que $\mathcal{M}(f) := \sum_{k \in \mathbb{Z}} \lambda_k c_k[f] e_k$, d'abord pour $f \in C_{1\#}^1$ puis aussi pour $f \in C_{1\#}^0$.

4. On vérifie que $\mathcal{M}^\infty(f) := \lim_{n \rightarrow +\infty} \mathcal{M}^n(f)$ est bien définie pour tout $f \in C_{1\#}^0$.

(a) Montrez que si $h \notin \mathbb{Q}$ alors $\mathcal{M}^\infty(f) = \langle f \rangle$.

(b) En revanche, si $h = \frac{p}{q}$, $p, q \in \mathbb{N} - \{0\}$, $\text{pgcd}(p, q) = 1$,

alors, démontrez que \mathcal{M}^∞ est le projecteur orthogonal de $C_{1\#}^0$ sur $C_{1/q\#}^0$.

4 Un modèle probabiliste pour l'étude du lissage par moyenne mobile

On suppose qu'une suite $(X_n)_{n \in \mathbb{Z}}$ de variable aléatoire satisfait, $X_n = an + b + \varepsilon_n$, pour tout n , où le "bruit" $(\varepsilon_n)_{n \in \mathbb{Z}}$ est une famille de v.a. indépendantes de moyenne nulle et de variance $\sigma^2 > 0$. Soit $m \in \mathbb{N} - \{0\}$, on se donne

$2m + 1$ réels positifs: $(\alpha_k)_{k=-m}^m$ dont la somme est égale à 1. On considère le lissage suivant $Y_n := \sum_{k=-m}^m \alpha_k X_{n+k}$

pour tout n .

1. Vérifiez que le lissage préserve le modèle linéaire bruité, i.e.: $Y_n = an + b + \eta_n$, $E(\eta_n) = 0, \forall n$. Et montrez que $\tau^2 := Var(\eta_n) = \sigma^2 \sum_{k=-m}^m \alpha_k^2 \leq Var(\varepsilon_n)$.

2. On va chercher à réduire au maximum la variance. Pour cela, on peut se ramener à une fonction de $2m$ variables: $f(\alpha_{-m}, \dots, \alpha_0, \dots, \alpha_{m-1}) := \sum_{k=-m}^m \alpha_k^2$ sachant que $\sum_{k=-m}^m \alpha_k = 1$.

Pour résoudre ce problème d'optimisation avec contrainte, proposez une élégante, simple, et rapide solution géométrique ou, tentez les calculs algébriques suivants:

(a) Vérifiez que que f n'admet qu'un extrémum local lorsque tous les coefficients sont égaux.

(b) Calculez la Hessienne de f , en déduire que l'on a un minimum global stricte et conclure le problème de réduction de la variance. On pourra utiliser la matrice de projecteur: $P \in M_{2m}(\mathbb{R})$ dont tous les coefficients sont égaux à $1/(2m)$.