

Inférence statistique utilisant des transformées de Fourier pour des modèles de régression semi-paramétriques

Travail de recherche sous la direction de Fabrice Gamboa

Myriam Vimond

Institut de Mathématiques de Toulouse
Laboratoire de Statistique et Probabilités

Toulouse, le 12 juillet 2007



Introduction

Partie 1 : Alignement de signaux continus

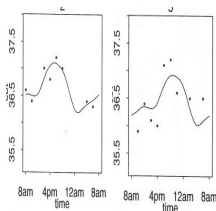
- 1 Estimation des paramètres
- 2 Test d'adéquation

Partie 2 : Alignement d'images

Partie 3 : Alignement de séries temporelles

Introduction

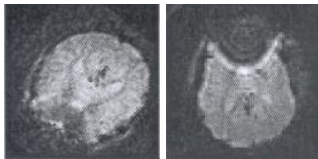
Signaux continus :
courbes de température



Wang and Brown

A flexible Model For Human Circadian
Rhythms
Biometrics 52 (1996)

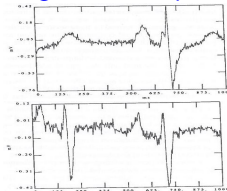
Images 2D :
IRM du cerveau



Keller, Shkolnisky et Averbuch

The angular difference function and its
application to image registration
Pattern Analysis and Machine
Intelligence, IEEE Transactions on,
27(6) (2005)

Signaux discrets :
signaux cardiaques



Jane, Rix, Caminal et Laguna

Alignment methods for averaging of
high-resolution cardiac signals
Biomedical Engineering, IEEE
Transactions on, 38(6) (1991)

Le **modèle de régression** :

$$Y_{i,j} = f_j^*(t_i) + \epsilon_{i,j}, \quad i = 1, \dots, n, \quad j = 1, \dots, J.$$

où

- f_j^* modélise le $j^{\text{ème}}$ signal observé (inconnu) ;
- t_i les points d'observation (connus).
- $\epsilon_{i,j}$ est un bruit blanc (inconnu),

Le modèle **paramétrique** d'opérateurs de déformation D_j ,

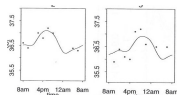
$$f_j^* = D_j f^*, \quad j = 1, \dots, J,$$

avec f^* **la forme commune du signal.**

Les familles d'opérateurs considérées

Signaux continus

$$f : \mathbb{R}^d \rightarrow \mathbb{R},$$

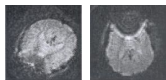


$$D : f(\cdot) \rightarrow af(\cdot - \theta) + v,$$

l'amplitude $a \in \mathbb{R}$,
la translation $\theta \in \mathbb{R}^d$
le niveau $v \in \mathbb{R}$

Images 2D

$$f : \mathcal{D} = [-d/2, d/2]^2 \rightarrow \mathbb{R},$$



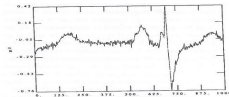
$$D : f(\cdot) \rightarrow f(R_\theta(\cdot - b)/a),$$

avec R_θ la matrice de rotation d'angle
 $\theta \in [0, 2\pi)$,

l'angle de rotation $\theta \in [0, 2\pi)$,
l'échelle $a \in \mathbb{R}_+^*$
la translation $b \in \mathbb{R}^2$

Signaux discrets

$$f = \{f_l, l = -n, \dots, n\}$$



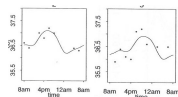
$$D : (f_l) \rightarrow (f_{l-\theta}),$$

le décalage $\theta \in \mathbb{Z}$,

Les familles d'opérateurs considérées

Signaux continus

$$f : \mathbb{R}^d \rightarrow \mathbb{R},$$



$$D : f(\cdot) \rightarrow af(\cdot - \theta) + v,$$

Objectifs : ✓ Estimer les paramètres suivants afin d'aligner les signaux,

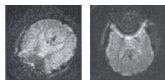
l'amplitude $a \in \mathbb{R}$,
la translation $\theta \in \mathbb{R}^d$
le niveau $v \in \mathbb{R}$

l'angle de rotation $\theta \in [0, 2\pi)$,
l'échelle $a \in \mathbb{R}_+$
la translation $b \in \mathbb{R}^2$

✓ Test d'adéquation pour le modèle de déformations paramétriques.

Images 2D

$$f : \mathcal{D} = [-d/2, d/2]^2 \rightarrow \mathbb{R},$$

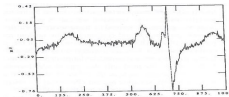


$$D : f(\cdot) \rightarrow f(R_\theta(\cdot - b)/a),$$

avec R_θ la matrice de rotation d'angle
 $\theta \in [0, 2\pi)$,

Signaux discrets

$$f = \{f_l, l = -n, \dots, n\}$$



$$D : (f_l) \rightarrow (f_{l-\theta}),$$

Transformées de Fourier

Signaux continus

Images 2D à niveau de gris

Signaux discrets

Transformées de Fourier :

$$c_l(f) = \int_{\mathcal{D}} f(t) e^{-il \cdot t} \frac{dt}{(2\pi)^d},$$

$$l \in \mathbb{Z}^d;$$

$$\text{où } \mathcal{D} = [0, 2\pi)^d$$

$$D_f(\lambda) = \frac{1}{\sqrt{n}} \sum_{|l| \leq n} f_l e^{-i\lambda l},$$

$$\lambda \in [-\pi, \pi];$$

Liens avec les familles d'opérateurs :

$$\text{Si } g(\cdot) = af(\cdot - \theta) + v,$$

$$c_l(g) = ac_l(f) e^{-il \cdot \theta}, \text{ si } l \neq 0$$

$$c_l(g) = ac_l(f) + v, \text{ si } l = 0.$$

$$\text{Si } g(\cdot) = f(R_\theta(\cdot - b)/a),$$

$$\text{Si } g = (f_{l-\theta})_{l \in \mathbb{Z}},$$

$$D_g(\lambda) \approx D_f(\lambda) e^{-i\lambda \theta}$$

Transformées de Fourier

Signaux continus

Images 2D à niveau de gris

Signaux discrets

Transformées de Fourier :

$$c_l(f) = \int_{\mathcal{D}} f(t) e^{-il \cdot t} \frac{dt}{(2\pi)^d},$$

$$l \in \mathbb{Z}^d;$$

$$\text{où } \mathcal{D} = [0, 2\pi)^d$$

$$\hat{f}(\omega) = \int_{\mathcal{D}} f(t) e^{-i\omega \cdot t} \frac{dt}{2\pi}, \quad \omega \in \mathbb{R}^2;$$

$$\mathcal{M}_{|\hat{f}|^2}(k, \nu) = \int_0^{2\pi} \int_0^{\infty} |\hat{f}|^2(r, \theta) r^{-i\nu} e^{-ik\theta} \frac{dr d\theta}{2\pi r}$$

$$(k, \nu) \in \mathbb{Z} \times \mathbb{R};$$

$$D_f(\lambda) = \frac{1}{\sqrt{n}} \sum_{|l| \leq n} f_l e^{-i\lambda l},$$

$$\lambda \in [-\pi, \pi];$$

Liens avec les familles d'opérateurs :

$$\text{Si } g(\cdot) = af(\cdot - \theta) + v,$$

$$c_l(g) = ac_l(f) e^{-il \cdot \theta}, \quad \text{si } l \neq 0$$

$$c_l(g) = ac_l(f) + v, \quad \text{si } l = 0.$$

$$\text{Si } g(\cdot) = f(R_\theta(\cdot - b)/a),$$

$$\hat{g}(\omega) = a^2 \hat{f}(aR_\theta \omega) e^{-i\omega \cdot b}$$

$$\mathcal{M}_{|\hat{g}|^2}(k, \nu) = \mathcal{M}_{|\hat{f}|^2}(k, \nu) e^{ik\theta} a^{4+i\nu}$$

$$\text{Si } g = (f_l)_{l \in \mathbb{Z}},$$

$$D_g(\lambda) \approx D_f(\lambda) e^{-i\lambda \theta}$$

Partie 1 : Alignement de signaux continus

- 1 Estimation des paramètres
- 2 Test d'adéquation

Partie 2 : Alignement d'images

Partie 3 : Alignement de séries temporelles

Partie 1 : Alignement de signaux continus

$$Y_{i,j} = a_j^* f^*(t_i - \theta_j^*) + v_j^* + \sigma_j^* \epsilon_{i,j}, \quad j = 1, \dots, J, \quad i \in \mathcal{I}_n = \{1, \dots, n\}^d.$$

Estimation des paramètres $\theta_j^* \in \mathbb{R}^d$, $\mathbf{a}_j^* \in \mathbb{R}^*$, $v_j^* \in \mathbb{R}$, $\mathbf{j} = 1, \dots, J$

Lawton, W.M., Sylvestre, E.A. et Maggio, M.G. (1972) : méthode itérative (**SEMOR**) basée sur l'approximation polynomiale de f^* .

Kneip, A. et Gasser, T.(1988) : consistance de la méthode SEMOR.

Härdle, W. et Marron, J.S.(1990) : estimateurs \sqrt{n} -consistants et asymptotiquement normaux, à partir d'un estimateur à noyau de f^* .

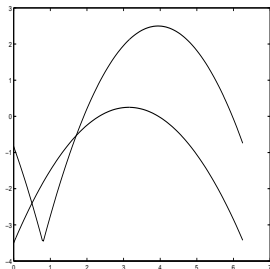
Gamboa, F., Loubes, J.M. et Maza, E. (2004) : estimateurs \sqrt{n} -consistants et asymptotiquement normaux, à partir de la **DFT** des données.

Le modèle

Théorie : pour $j = 1, \dots, J$,

$$f_j^*(t) = a_j^* f^*(t - \theta_j^*) + v_j^*, \quad t \in [0, 2\pi]^d.$$

- f^* $2\pi\mathbb{Z}^d$ -périodique (minimale)
continue.



Le modèle

Théorie : pour $j = 1, \dots, J$,

$$f_j^*(t) = a_j^* f^*(t - \theta_j^*) + v_j^*, \quad t \in [0, 2\pi]^d.$$

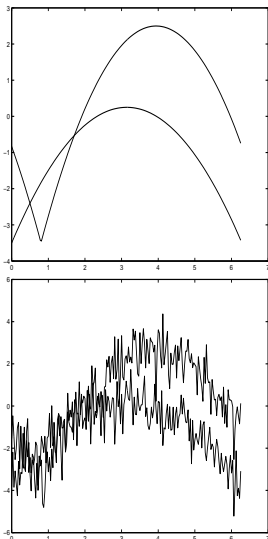
- f^* $2\pi\mathbb{Z}^d$ -périodique (minimale)
continue.

Les données : pour $j = 1 \dots J$,

$$Y_{i,j} = a_j^* f^*(t_i - \theta_j^*) + v_j^* + \sigma_j^* \epsilon_{i,j},$$

$$i \in I_n = \{1, \dots, n\}^d,$$

- $t_i = 2\pi(i)/n \in [0, 2\pi]^d$,
- $(\epsilon_{i,j})_{ij}$ sont i.i.d, $\mathbb{E}\epsilon_{i,j} = 0$, $\mathbb{E}\epsilon_{i,j}^2 = 1$,
- variance du bruit $\sigma_j^{*2} = \sigma^{*2}$,



Contraintes d'identifiabilité

Deux jeux de contraintes d'identifiabilité sont principalement considérés :

- 1 **Une paramétrisation naturelle** : on prend un des signaux comme "référence",

$$\theta_1^* = 0, \quad a_1^* = 1 \quad \text{et} \quad v_1^* = 0$$

- 2 **La paramétrisation la plus adéquate** :

$$\theta_1^* = 0, \quad \sum_{j=1}^J a_j^{*2} = J \quad \text{et} \quad a_1 > 0, \quad (1)$$

$$c_0(f^*) = \int_{\mathcal{D}} f(t) \frac{dt}{(2\pi)^d} = 0 \quad (2)$$

- (2) fournit des estimateurs asymptotiquement indépendants,
- (1) fournit des estimateurs efficaces (vraisemblance profilée) dans le cas gaussien,

Critère d'estimation (I)

Transformée de Fourier Discrète (DFT) des données : soit m_n une suite d'entiers telle que $m_n \leq n/2$.

$$d_{j,l} = \frac{1}{n^d} \sum_{i \in \mathcal{I}_n} Y_{i,j} e^{-il \cdot t_i} = c_l^n(f_j^*) + \sigma^* w_{j,l},$$

$$l = (l_1, \dots, l_r) \in \mathbb{Z}^d \text{ tel que } \|l\|_\infty = \max_{r=1, \dots, d} |l_r| \leq m_n$$

- où
- $w_{j,l}$ sont des variables complexes décorrélatées de variance $1/n^d$;
 - $n^{d/2} w_{j,l}$ convergent en loi vers une loi Gaussienne complexe $\mathcal{N}^{\mathbb{C}}(0, 1)$.

Critère profilé des moindres carrés

Critère d'estimation (I)

Transformée de Fourier Discrète (DFT) des données : soit m_n une suite d'entiers telle que $m_n \leq n/2$.

$$d_{j,l} = \frac{1}{n^d} \sum_{i \in \mathcal{I}_n} Y_{i,j} e^{-il \cdot t_i} = c_l^n(f_j^*) + \sigma^* w_{j,l},$$

$$l = (l_1, \dots, l_r) \in \mathbb{Z}^d \text{ tel que } \|l\|_\infty = \max_{r=1, \dots, d} |l_r| \leq m_n$$

- où
- $w_{j,l}$ sont des variables complexes décorrélées de variance $1/n^d$;
 - $n^{d/2} w_{j,l}$ convergent en loi vers une loi Gaussienne complexe $\mathcal{N}^{\mathbb{C}}(0, 1)$.

Critère profilé des moindres carrés

$$\alpha = (\theta_1, \dots, \theta_J, a_1, \dots, a_J, v_1, \dots, v_J) \rightarrow \inf_f \sum_{i \in \mathcal{I}_n} \sum_{j=1}^J (Y_{i,j} - a_j f(t_i - \theta_j) - v_j)^2.$$

Critère d'estimation (I)

Transformée de Fourier Discrète (DFT) des données : soit m_n une suite d'entiers telle que $m_n \leq n/2$.

$$d_{j,l} = \frac{1}{n^d} \sum_{i \in \mathcal{I}_n} Y_{i,j} e^{-il \cdot t_i} = c_l^n(f_j^*) + \sigma^* w_{j,l},$$

$$l = (l_1, \dots, l_r) \in \mathbb{Z}^d \text{ tel que } \|l\|_\infty = \max_{r=1, \dots, d} |l_r| \leq m_n$$

- où
- $w_{j,l}$ sont des variables complexes décorrélées de variance $1/n^d$;
 - $n^{d/2} w_{j,l}$ convergent en loi vers une loi Gaussienne complexe $\mathcal{N}^{\mathbb{C}}(0, 1)$.

Critère profilé des moindres carrés

$$\alpha = (\theta_1, \dots, \theta_J, a_1, \dots, a_J, v_1, \dots, v_J) \rightarrow \inf_f \sum_{i \in \mathcal{I}_n} \sum_{j=1}^J (Y_{i,j} - a_j f(t_i - \theta_j) - v_j)^2.$$

On minimise sur l'espace des polynômes trigonométriques de degré m_n .

Critère d'estimation (I)

Le processus de contraste

$$M_n(\alpha) = \frac{1}{J} \sum_{j=1}^J |d_{j,0} - v_j|^2 + \sum_{1 \leq |l|_\infty \leq m_n} \left| d_{j,l} - a_j \hat{c}_l(\alpha) e^{-il \cdot \theta_j} \right|^2.$$

où $\hat{c}_l(\alpha) = \frac{1}{J} \sum_{j=1}^J a_j d_{j,l} e^{il \cdot \theta_j}$ est le coefficient de Fourier associé à la fonction qui minimise les moindres carrés :

$$t \in \mathbb{R}^d \rightarrow \sum_{|l|_\infty \leq m_n} \hat{c}_l(\alpha) e^{il \cdot t}.$$

Théorème (consistance)

Soit

$$\hat{\alpha}_n = \arg \min_{\alpha} M_n(\alpha).$$

Si $\sum_{|l|_\infty > m_n} |c_l(f^*)| = o\left(\frac{1}{m_n^{d/2}}\right)$ et $\frac{m_n^d}{n^d} = o(1)$,

alors $\hat{\alpha}_n$ converge en probabilité vers $\alpha^* = (\theta_1^*, \dots, \theta_J^*, a_1^*, \dots, a_J^*, v_1^*, \dots, v_J^*)$.

Théorème (Efficacité semi-paramétrique et normalité asymptotique (d=1))

Si $\sum_{|l|_\infty > m_n} |c_l(f^*)| = o\left(\frac{1}{m_n^{d/2}}\right)$, $\sum_{l \in \mathbb{Z}^d} |l_1| + \dots + |l_d| |c_l(f^*)| < \infty$, et $\frac{m_n^{d+1}}{n^{d/2}} = o(1)$,

Alors $\hat{\alpha}_n = (\hat{\theta}_n, \hat{a}_n, \hat{v}_n)$ est asymptotiquement normal, c'est-à-dire

$$n^{d/2}(\hat{\theta}_n - \theta^*, \hat{a}_n - a^*, \hat{v}_n - v^*) \rightsquigarrow \mathcal{N}_{3J-2}(0, \sigma^{*2} H^{-1}),$$

où :

$$H = \begin{pmatrix} \|\partial f^*\|_{\mathbb{L}^2}^2 (D^2 - \frac{1}{J} V_a V_a^T) & 0 & 0 \\ 0 & \|f^*\|_{\mathbb{L}^2}^2 \left(I_{J-1} + \frac{1}{a_1^{*2}} a^* a^{*T} \right) & 0 \\ 0 & 0 & I_J \end{pmatrix}.$$

et $D = \text{diag}(a_2^* \dots a_j^*)$, $V_a = (a_2^{*2} \dots a_j^{*2})^T$, $a^* = (a_2^* \dots a_j^*)^T$.

De plus, si $\epsilon_{ij} \sim \mathcal{N}(0, 1)$, $\hat{\alpha}_n$ est asymptotiquement efficace.

Idée de la preuve (I), (McNeney et Wellner, 2000)

Soit $\mathbb{P}_{\alpha^*, \sigma^*, f^*}^{(n)}$ la loi du modèle des observations.

Normalité Asymptotique Locale du modèle :

Différentiabilité du paramètre $\nu_n(\mathbb{P}_{\alpha^*, \sigma^*, f^*}^{(n)}) = \alpha^*$ sur les voisinages locaux :

$\hat{\alpha}_n$ est asymptotiquement linéaire

Idée de la preuve (I), (McNeney et Wellner, 2000)

Soit $\mathbb{P}_{\alpha^*, \sigma^*, f^*}^{(n)}$ la loi du modèle des observations.

Normalité Asymptotique Locale du modèle : soit $(\alpha_n(h), \sigma_n(h), f_n(h))$ dans un voisinage de $(\alpha^*, \sigma^*, f^*)$ (dans la direction $h \in \mathcal{H}$, où \mathcal{H} un espace pré-hilbertien),

$$\log \frac{d\mathbb{P}_{\alpha_n(h), \sigma_n(h), f_n(h)}^{(n)}}{d\mathbb{P}_{\alpha^*, \sigma^*, f^*}^{(n)}} = \Delta_n(h) - \frac{1}{2} \|h\|_{\mathcal{H}}^2 + o_{\mathbb{P}}(1),$$

où $\Delta_n(h) \rightsquigarrow \mathcal{N}(0, \|h\|_{\mathcal{H}}^2)$.

Différentiabilité du paramètre $\nu_n(\mathbb{P}_{\alpha^*, \sigma^*, f^*}^{(n)}) = \alpha^*$ sur les voisinages locaux :

$\hat{\alpha}_n$ est asymptotiquement linéaire

Ideée de la preuve (I), (McNeney et Wellner, 2000)

Soit $\mathbb{P}_{\alpha^*, \sigma^*, f^*}^{(n)}$ la loi du modèle des observations.

Normalité Asymptotique Locale du modèle : soit $(\alpha_n(h), \sigma_n(h), f_n(h))$ dans un voisinage de $(\alpha^*, \sigma^*, f^*)$ (dans la direction $h \in \mathcal{H}$, où \mathcal{H} un espace pré-hilbertien),

$$\log \frac{d\mathbb{P}_{\alpha_n(h), \sigma_n(h), f_n(h)}^{(n)}}{d\mathbb{P}_{\alpha^*, \sigma^*, f^*}^{(n)}} = \Delta_n(h) - \frac{1}{2} \|h\|_{\mathcal{H}}^2 + o_{\mathbb{P}}(1),$$

où $\Delta_n(h) \rightsquigarrow \mathcal{N}(0, \|h\|_{\mathcal{H}}^2)$.

Différentiabilité du paramètre $\nu_n(\mathbb{P}_{\alpha^*, \sigma^*, f^*}^{(n)}) = \alpha^*$ **sur les voisinages locaux** : il existe $\dot{\nu} \in \text{cl}(\mathcal{H})^{3J-2}$ (théorème de Riesz) tel que

$$\lim_{n \rightarrow \infty} n^{d/2} \left(\nu_n(\mathbb{P}_{\alpha^*, \sigma^*, f^*}^{(n)}) - \nu_n(\mathbb{P}_{\alpha_n(h), \sigma_n(h), f_n(h)}^{(n)}) \right) = \langle \dot{\nu}, h \rangle_{\mathcal{H}}$$

$\hat{\alpha}_n$ est asymptotiquement linéaire

Ideée de la preuve (I), (McNeney et Wellner, 2000)

Soit $\mathbb{P}_{\alpha^*, \sigma^*, f^*}^{(n)}$ la loi du modèle des observations.

Normalité Asymptotique Locale du modèle : soit $(\alpha_n(h), \sigma_n(h), f_n(h))$ dans un voisinage de $(\alpha^*, \sigma^*, f^*)$ (dans la direction $h \in \mathcal{H}$, où \mathcal{H} un espace pré-hilbertien),

$$\log \frac{d\mathbb{P}_{\alpha_n(h), \sigma_n(h), f_n(h)}^{(n)}}{d\mathbb{P}_{\alpha^*, \sigma^*, f^*}^{(n)}} = \Delta_n(h) - \frac{1}{2} \|h\|_{\mathcal{H}}^2 + o_{\mathbb{P}}(1),$$

où $\Delta_n(h) \rightsquigarrow \mathcal{N}(0, \|h\|_{\mathcal{H}}^2)$.

Différentiabilité du paramètre $\nu_n(\mathbb{P}_{\alpha^*, \sigma^*, f^*}^{(n)}) = \alpha^*$ **sur les voisinages locaux** : il existe $\dot{\nu} \in \text{cl}(\mathcal{H})^{3J-2}$ (théorème de Riesz) tel que

$$\lim_{n \rightarrow \infty} n^{d/2} \left(\nu_n(\mathbb{P}_{\alpha^*, \sigma^*, f^*}^{(n)}) - \nu_n(\mathbb{P}_{\alpha_n(h), \sigma_n(h), f_n(h)}^{(n)}) \right) = \langle \dot{\nu}, h \rangle_{\mathcal{H}}$$

$\hat{\alpha}_n$ est asymptotiquement linéaire si $n^{d/2}(\hat{\alpha}_n - \alpha^*) = \Delta_n(\tilde{h}) + o_{\mathbb{P}}(1)$, $\tilde{h} \in \mathcal{H}^{3J-2}$

Idée de la preuve (I), (McNeney et Wellner, 2000)

Soit $\mathbb{P}_{\alpha^*, \sigma^*, f^*}^{(n)}$ la loi du modèle des observations.

Normalité Asymptotique Locale du modèle : soit $(\alpha_n(h), \sigma_n(h), f_n(h))$ dans un voisinage de $(\alpha^*, \sigma^*, f^*)$ (dans la direction $h \in \mathcal{H}$, où \mathcal{H} un espace pré-hilbertien),

$$\log \frac{d\mathbb{P}_{\alpha_n(h), \sigma_n(h), f_n(h)}^{(n)}}{d\mathbb{P}_{\alpha^*, \sigma^*, f^*}^{(n)}} = \Delta_n(h) - \frac{1}{2} \|h\|_{\mathcal{H}}^2 + o_{\mathbb{P}}(1),$$

où $\Delta_n(h) \rightsquigarrow \mathcal{N}(0, \|h\|_{\mathcal{H}}^2)$.

Différentiabilité du paramètre $\nu_n(\mathbb{P}_{\alpha^*, \sigma^*, f^*}^{(n)}) = \alpha^*$ **sur les voisinages locaux** : il existe $\dot{\nu} \in \text{cl}(\mathcal{H})^{3J-2}$ (théorème de Riesz) tel que

$$\lim_{n \rightarrow \infty} n^{d/2} \left(\nu_n(\mathbb{P}_{\alpha^*, \sigma^*, f^*}^{(n)}) - \nu_n(\mathbb{P}_{\alpha_n(h), \sigma_n(h), f_n(h)}^{(n)}) \right) = \langle \dot{\nu}, h \rangle_{\mathcal{H}}$$

$\hat{\alpha}_n$ est asymptotiquement linéaire si $n^{d/2}(\hat{\alpha}_n - \alpha^*) = \Delta_n(\tilde{h}) + o_{\mathbb{P}}(1)$, $\tilde{h} \in \mathcal{H}^{3J-2}$

$\hat{\alpha}_n$ est régulier et asymptotiquement efficace ssi

$$\tilde{h} = \dot{\nu}$$

Idée de la preuve (II)

Normalité asymptotique :

Par un développement de Taylor, il existe $\bar{\alpha}_n$ tel que :

$$\begin{aligned}n^{d/2}(\hat{\alpha}_n - \alpha^*) &= - [\nabla^2 M_n(\bar{\alpha}_n)]^{-1} n^{d/2} \nabla M_n(\alpha^*), \\ &= \left(H/\sigma^{*2}\right)^{-1} G_n/\sigma^* + o_{\mathbb{P}}(1),\end{aligned}$$

où $G_n \rightsquigarrow \mathcal{N}_{3J-2}(0, H)$, et $\nabla^2 M_n(\bar{\alpha}_n) \xrightarrow[n \rightarrow \infty]{\mathbb{P}} -\frac{2}{J}H$. □

Idée de la preuve (II)

Normalité asymptotique :

Par un développement de Taylor, il existe $\bar{\alpha}_n$ tel que :

$$\begin{aligned} n^{d/2}(\hat{\alpha}_n - \alpha^*) &= - [\nabla^2 M_n(\bar{\alpha}_n)]^{-1} n^{d/2} \nabla M_n(\alpha^*), \\ &= \underbrace{\left(H / \sigma^{*2} \right)^{-1}}_{\Delta_n(\dot{\nu}) + o_{\mathbb{P}}(1)} G_n / \sigma^* + o_{\mathbb{P}}(1), \end{aligned}$$

où $G_n \rightsquigarrow \mathcal{N}_{3J-2}(0, H)$, et $\nabla^2 M_n(\bar{\alpha}_n) \xrightarrow[n \rightarrow \infty]{\mathbb{P}} -\frac{2}{J} H$. □

Ainsi si $(\epsilon_{ij} \sim \mathcal{N}(0, 1))$: H / σ^{*2} est la matrice d'information efficace, G_n / σ^* est le score efficace.

"la vraisemblance profilée se comporte comme la vraisemblance classique"
(Murphy et Van der Vaart (2000))

Vers la selection de modèle ...

Étudions le critère à $m = m_n$ fixé pour $J = 2$.

La variance asymptotique de l'estimateur de θ_2^* est

$$\frac{2\sigma^{*2}}{n \sum_{|l| \leq m} l^2 |c_l(f)|^2}.$$

On génère des données à partir de la forme commune :

$$f(t) = 20t/(2\pi)(1 - t/(2\pi)).$$

| | $m = 5$ | $m = 15$ |
|-----------|-------------------------------------|-------------------------------------|
| $n = 51$ | $\hat{\theta} = 0.9460$ (0.1612) | $\hat{\theta} = 0.5722$ (0.5519) |
| $n = 101$ | $\hat{\theta} = 0.9454$ (0.1055) | $\hat{\theta} = 0.8828$ (0.2385) |
| $n = 201$ | $\hat{\theta} = 0.9496$ (0.0706) | $\hat{\theta} = 0.9499$ (0.1086) |

Tab.: Moyennes et écarts-type (entre parenthèses) pour l'estimation du paramètre $\theta_2^* = 0.95$ calculés à partir 1000 jeux de données.

Vers la selection de modèle ...

Étudions le critère à $m = m_n$ fixé pour $J = 2$.

La variance asymptotique de l'estimateur de θ_2^* est

$$\frac{2\sigma^{*2}}{n \sum_{|l| \leq m} l^2 |c_l(f)|^2}$$

La variance empirique (heuristique) est :

$$\frac{2\sigma^{*2}}{n \sum_{|l| \leq m} l^2 |c_l(f)|^2} + \sigma^{*4} \frac{m(m+1)(2m+1)}{3n^2 \left(\sum_{|l| \leq m} l^2 |c_l(f)|^2 \right)^{-2}}$$

En cours ...

estimation de la variance empirique

On génère des données à partir de la forme commune :

$$f(t) = 20t/(2\pi)(1 - t/(2\pi)).$$

| | $m = 5$ | $m = 15$ |
|-----------|-------------------------------------|-------------------------------------|
| $n = 51$ | $\hat{\theta} = 0.9460$ (0.1612) | $\hat{\theta} = 0.5722$ (0.5519) |
| $n = 101$ | $\hat{\theta} = 0.9454$ (0.1055) | $\hat{\theta} = 0.8828$ (0.2385) |
| $n = 201$ | $\hat{\theta} = 0.9496$ (0.0706) | $\hat{\theta} = 0.9499$ (0.1086) |

Tab.: Moyennes et écarts-type (entre parenthèses) pour l'estimation du paramètre $\theta_2^* = 0.95$ calculés à partir 1000 jeux de données.

Partie 1 : Alignement de signaux continus

Le modèle général : $d = 1$,

$$Y_{i,j} = f_j^*(t_i) + \sigma^* \epsilon_{i,j}, \quad j = 1, \dots, J, \quad i = 1, \dots, n.$$

L'hypothèse nulle H_0 :

$$Y_{i,j} = a_j^* f^*(t_i - \theta_j^*) + v_j^* + \sigma^* \epsilon_{i,j}, \quad j = 1, \dots, J, \quad i = 1, \dots, n.$$

Test d'adéquation du modèle

Möcks, J., Dinh, T. et Gasser, T.(1984) : deux procédures de test pour tester $f_j^* = f^*$.

Härdle, W. et Marron, J.S.(1990) : procédures de test pour tester $f_j^* = f^*(\cdot - \theta_j^*) + v_j^*$.

Wang, Y. et Brown, M. (1996) : test d'ajustement sur les paramètres θ_j^* .

Présentation du test

Le modèle général : $d = 1$,

$$Y_{ij} = f_j^*(t_i) + \sigma_j^* \epsilon_{ij}, \quad j = 1 \dots J, \quad i = 1 \dots n.$$

L'hypothèse nulle H_0 :

il existe α^* et f^* tels que : $\forall j = 1 \dots J, f_j^*(\cdot) = \alpha_j^* f^*(\cdot - \theta_j^*) + v_j^*$,

L'hypothèse alternative H_1 :

pour tout α et pour toute fonction f , $\exists j = 1 \dots J, f_j^*(\cdot) \neq \alpha_j f(\cdot - \theta_j) + v_j$,

Quelques outils

Sous H_0

✓ Estimateur des paramètres θ_j^* , a_j^* ,

v_j^* :

$$\hat{\alpha}_n = (\hat{\theta}_{1,n}, \dots, \hat{\theta}_{J,n}, \hat{a}_{1,n}, \dots, \hat{v}_{J,n})$$

✓ Estimateur de la forme commune f^* :

$$\hat{f}_n(t) = \sum_{1 \leq |l| \leq m_n} \hat{c}_l(\hat{\alpha}_n) e^{ilt}, \quad \forall t \in \mathbb{R}.$$



✓ Estimateur des fonctions de régression f_j^* :

$$\hat{a}_{j,n} \hat{f}_n(t - \hat{\theta}_{j,n}) + \hat{v}_{j,n}, \quad \forall t \in \mathbb{R}.$$

Sous le modèle général

Rappel

$d_{j,l} = \frac{1}{n} \sum_{i=1}^n Y_{i,j} e^{ilt_i}$ est la DFT des données,

et $\hat{c}_l(\alpha)$ sont les coefficients de Fourier de la fonction qui minimise le critère profilé des moindres carrés

Quelques outils

Sous H_0

✓ Estimateur des paramètres θ_j^* , a_j^* ,

v_j^* :

$$\hat{\alpha}_n = (\hat{\theta}_{1,n}, \dots, \hat{\theta}_{J,n}, \hat{a}_{1,n}, \dots, \hat{v}_{J,n})$$

✓ Estimateur de la forme commune f^* :

$$\hat{f}_n(t) = \sum_{1 \leq |l| \leq m_n} \hat{c}_l(\hat{\alpha}_n) e^{ilt}, \quad \forall t \in \mathbb{R}.$$



✓ Estimateur des fonctions de régression f_j^* :

$$\hat{a}_{j,n} \hat{f}_n(t - \hat{\theta}_{j,n}) + \hat{v}_{j,n}, \quad \forall t \in \mathbb{R}.$$

Sous le modèle général

✓ Estimateur des fonctions de régression

f_j^* :

$$\hat{f}_{j,n}(t) = \sum_{1 \leq |l| \leq m_n} d_{j,l} e^{ilt}, \quad \forall t \in \mathbb{R}.$$

✓ Estimateur de la variance σ^{*2} :

$$\hat{\sigma}_n^2 = \frac{1}{nJ} \sum_{j=1}^J \sum_{i=1}^n \left(Y_{ij} - \hat{f}_{j,n}(t_i) \right)^2.$$

Rappel

$d_{j,l} = \frac{1}{n} \sum_{i=1}^n Y_{i,j} e^{ilt_i}$ est la DFT des données,

et $\hat{c}_l(\alpha)$ sont les coefficients de Fourier de la fonction qui minimise le critère profilé des moindres carrés

Statistiques de test

Corollaire

On considère la statistique de test :

$$T_n = Jn \inf_{\alpha} M_n(\alpha) / \hat{\sigma}_n^2,$$

où $M_n(\alpha) = \frac{1}{J} \sum_{j=1}^J |d_{j0} - v_j|^2 + \sum_{|l| \leq m_n} |d_{jl} - a_j \hat{c}_l(\alpha) e^{-il\theta_j}|^2$. Supposons que pour tout $j = 1, \dots, J$:

$$\sum_{l \in \mathbb{Z}} |l| |c_l(f_j^*)| < \infty, \text{ et } m_n^4/n = o(1),$$

Alors

- sous H_0 , T_n/m_n converge en probabilité vers $(2J-2)$
- sous H_1 , $T_n/m_n \rightarrow +\infty$

Statistiques de test

Corollaire

On considère la statistique de test :

$$T_n^\delta = Jn \inf_{\alpha} M_n^\delta(\alpha) / \hat{\sigma}_n^2,$$

où $M_n^\delta(\alpha) = \frac{1}{J} \sum_{j=1}^J |d_{j0} - v_j|^2 + \sum_{|l| \leq m_n} \delta_l^2 |d_{jl} - a_j \hat{c}_l(\alpha) e^{-il\theta_j}|^2$, et $(\delta)_{l \in \mathbb{Z}}$ une série positive de carré sommable. Supposons que pour tout $j = 1, \dots, J$:

$$\sum_{|l| > m} |c_l(f_j^*)| = o\left(\frac{1}{\sqrt{m}}\right), \quad m_n^2/n = o(1), \quad \text{et} \quad \sum_{l \in \mathbb{Z}} l^2 \delta_l^2 < \infty,$$

Alors

- sous H_0 , T_n^δ converge en loi
- sous H_1 , $T_n^\delta \rightarrow +\infty$

Performances numériques du test Bootstrap

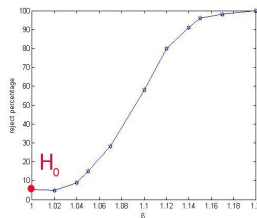
Génération des données pour différentes valeurs de $\beta \in [1, 1.2]$:

$$\begin{cases} Y_{i,1} = a_1^* f^*(t_i - \theta_1^*) + v_1^* + \sigma^* \epsilon_{i1} & i = 1 \dots n, \\ Y_{i,2} = a_2^* f^*(\beta t_i - \theta_2^*) + v_2^* + \sigma^* \epsilon_{i2} & i = 1 \dots n, \end{cases}$$

On applique la procédure Bootstrap de test au niveau 5% (chaque point est calculé à partir de 1000 jeux de données) :

✓ Quand $\beta = 1$, H_0 est vraie ;

Puissance du test en fonction de β



Performances numériques du test Bootstrap

Génération des données pour différentes valeurs de $\beta \in [1, 1.2]$:

$$\begin{cases} Y_{i,1} = a_1^* f^*(t_i - \theta_1^*) + v_1^* + \sigma^* \epsilon_{i1} & i = 1 \dots n, \\ Y_{i,2} = a_2^* f^*(\beta t_i - \theta_2^*) + v_2^* + \sigma^* \epsilon_{i2} & i = 1 \dots n, \end{cases}$$

On applique la procédure Bootstrap de test au niveau 5% (chaque point est calculé à partir de 1000 jeux de données) :

- ✓ Quand $\beta = 1$, H_0 est vraie ;
- ✓ Quand $\beta = 1.1$, H_0 est fausse.

Ouverture

Étude de l'erreur de première espèce

Puissance du test en fonction de β

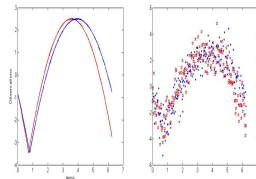
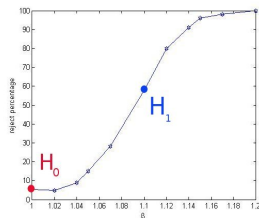
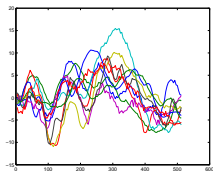


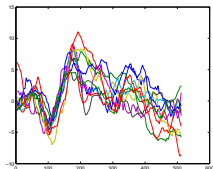
Illustration avec des données réelles : les ondes P300

Les données

✓ 11 ondes placebo



✓ 11 ondes traitement

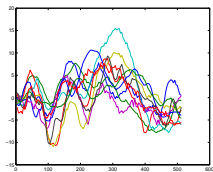


- ✓ Effets Placébo et Traitement étudiés simultanément
- ✓ Effets Placébo et Traitement étudiés séparément

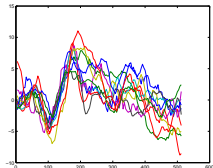
Illustration avec des données réelles : les ondes P300

Les données

✓ 11 ondes placebo



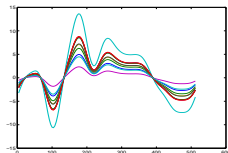
✓ 11 ondes traitement



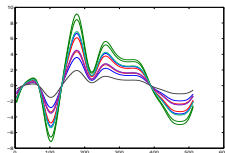
✓ Effets Placébo et Traitement étudiés simultanément

Le modèle : $Y_{i,j} = a_j^* f^*(t_i - \theta_j^*) + \epsilon_{i,j}$, $j=1..22$, $i=1..127$

Estimateurs des ondes placebo



Estimateurs des ondes traitement



Statistique de test $T_n^* = 2652$,

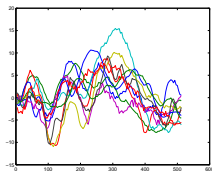
Estimateur Bootstrap du quantile d'ordre 95% : 27.230.

✓ Effets Placébo et Traitement étudiés séparément

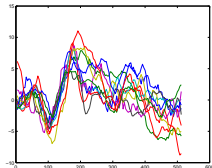
Illustration avec des données réelles : les ondes P300

Les données

✓ 11 ondes placebo



✓ 11 ondes traitement



✓ Effets Placébo et Traitement étudiés simultanément

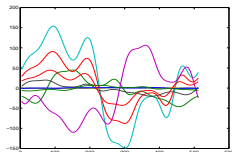
Statistique de test $T_n^* = 2652$,

Estimateur Bootstrap du quantile d'ordre 95% : 27.230.

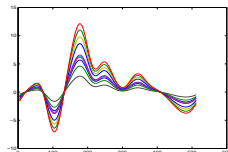
✓ Effets Placébo et Traitement étudiés séparément

Le modèle : $Y_{i,j} = a_j^* f^*(t_i - \theta_j^*) + \epsilon_{i,j}$, $j=1..11$, $i=1..127$.

Estimateurs des ondes placebo



Estimateurs des ondes traitement



Statistique de test pour les ondes traitement $T_n^* = 784.4$,

Estimateur Bootstrap du quantile d'ordre 95% : 25.61.

Partie 2 : Alignement d'images 2D

$$dY_j(x) = f^*(R_{\theta_j^*}(x - b_j^*)/a_j^*) + \epsilon dW_j(x), x \in \mathcal{D}$$

Estimation des paramètres $a_j^* > 0$, $\theta_j^* \in [0, 2\pi)$, $b_j^* \in \mathbb{R}^2$, $j = 1, \dots, J$

Travail en collaboration avec Jérémie Bigot et Fabrice Gamboa

Kuglin, et Hines, (1975) : technique de corrélation de phase (translations).

Reddy et Chatterji, (1996) : estimation par la technique de corrélation de phase (coordonnées log-polaires), en 3 étapes.

Keller, Averbuch, Israeli (2005) : estimation par la technique de corrélation de phase (translations), et par la fonction de différence angulaire (rotations).

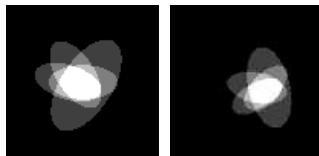
Derrode, et Ghordel, (2004) : transformée de Fourier-Mellin analytique (rotations et échelles).

Le modèle

Théorie : pour $j = 1, \dots, J$,

$$f_j^*(x) = f^*\left(\frac{1}{a_j^*} A_{\theta_j^*}(x - b_j^*)\right), \quad x \in \mathbb{R}^2.$$

- f^* a un support compact $\mathcal{D} \subset \mathbb{R}^2$.



Le modèle

Théorie : pour $j = 1, \dots, J$,

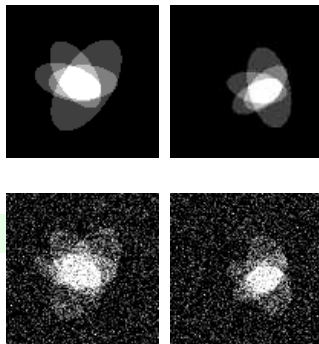
$$f_j^*(x) = f^*\left(\frac{1}{a_j^*} A_{\theta_j^*}(x - b_j^*)\right), \quad x \in \mathbb{R}^2.$$

- f^* a un support compact $\mathcal{D} \subset \mathbb{R}^2$.

Le modèle de bruit blanc : pour $j = 1 \dots J$,

$$dY_j(x) = f_j^*(x)dx + \epsilon dW_j(x), \quad x \in \mathcal{D}.$$

- W_j , $j = 1 \dots J$, sont des mouvements browniens indépendants.
- ϵ est un paramètre de niveau de bruit inconnu.



Le modèle

Théorie : pour $j = 1, \dots, J$,

$$f_j^*(x) = f^*\left(\frac{1}{a_j^*} A_{\theta_j^*}(x - b_j^*)\right), \quad x \in \mathbb{R}^2.$$

- f^* a un support compact $\mathcal{D} \subset \mathbb{R}^2$.

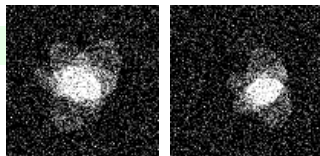
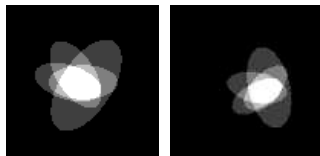
Le modèle de bruit blanc : pour $j = 1 \dots J$,

$$dY_j(x) = f_j^*(x)dx + \epsilon dW_j(x), \quad x \in \mathcal{D}.$$

- W_j , $j = 1 \dots J$, sont des mouvements browniens indépendants.
- ϵ est un paramètre de niveau de bruit inconnu.

Contraintes d'identifiabilité

$$\theta_1^* = 0, \quad a_1^* = 1, \quad b_1^* = (0, 0)$$



Un exemple réel :



Idée de départ

Notons $\alpha = (a_1, \dots, a_J, \theta_1, \dots, \theta_J) \in [a_{\min}, a_{\max}]^J \times [0, 2\pi[$ ($a_1 = 1, \theta_1 = 0$) et $\beta = (b_1, \dots, b_J) \in [b_{\min}, b_{\max}]^{2J}$ ($b_1 = (0, 0)$).

But : définir un estimateur du minimum de contraste de :

$$\alpha^* = (a_1^*, \dots, a_J^*, \theta_1^*, \dots, \theta_J^*) \quad \text{et} \quad \beta^* = (b_1^*, \dots, b_J^*).$$

✓ On cherche à estimer la fonction de contraste suivante :

$$(\alpha, \beta) \rightarrow \frac{1}{J} \sum_{j=1}^J \left\| g_j(\alpha, \beta, x) - \frac{1}{J} \sum_{j'=1}^J g_{j'}(\alpha, \beta, x) \right\|_{L^2(\mathbb{R}^2)}^2,$$

avec $g_j(\alpha, \beta, x) = f_j^* \left(\frac{a_j}{a_j^*} R_{\theta_j^* - \theta_j} x + \frac{1}{a_j^*} R_{\theta_j^*} (b_j - b_j^*) \right)$.

Idee de départ

Notons $\alpha = (a_1, \dots, a_J, \theta_1, \dots, \theta_J) \in [a_{\min}, a_{\max}]^J \times [0, 2\pi[$ ($a_1 = 1, \theta_1 = 0$) et $\beta = (b_1, \dots, b_J) \in [b_{\min}, b_{\max}]^{2J}$ ($b_1 = (0, 0)$).

But : définir un estimateur du minimum de contraste de :

$$\alpha^* = (a_1^*, \dots, a_J^*, \theta_1^*, \dots, \theta_J^*) \quad \text{et} \quad \beta^* = (b_1^*, \dots, b_J^*).$$

✓ On cherche à estimer la fonction de contraste suivante :

$$(\alpha, \beta) \rightarrow \frac{1}{J} \sum_{j=1}^J \left\| g_j(\alpha, \beta, x) - \frac{1}{J} \sum_{j'=1}^J g_{j'}(\alpha, \beta, x) \right\|_{L^2(\mathbb{R}^2)}^2,$$

avec $g_j(\alpha, \beta, x) = f_j^* \left(\frac{a_j}{a_j^*} R_{\theta_j^* - \theta_j} x + \frac{1}{a_j^*} R_{\theta_j^*} (b_j - b_j^*) \right)$.

✗ Très complexe :

- calcul du minimum,
- choix des paramètres de lissage du critère.

Procédure d'estimation à 2 pas

Transformées de Fourier des données ($j=1, \dots, J$) :

$$\hat{Y}_j(\omega) = \int_{\mathcal{D}} e^{-i\omega \cdot x} \frac{dY_j(x)}{2\pi} = a_j^{*2} \hat{f}(a_j^* R_{\theta_j^*} \omega) e^{i\omega \cdot b_j^*} + \epsilon \hat{W}_j(\omega), \quad \omega \in \mathbb{R}^2$$

1^{er} pas : estimation des rotations et des échelles $\alpha^* = (a_1^*, \dots, a_J^*, \theta_1^*, \dots, \theta_J^*)$:

L'estimateur $\hat{\alpha}_\epsilon = \arg \min_{\alpha} M_{1,\epsilon}$ est ϵ -consistant.

2^{ème} pas : estimation des translations $\beta^* = (b_1^*, \dots, b_J^*)$

L'estimateur $\hat{\beta}_\epsilon = \arg \min_{\alpha} M_{2,\epsilon}$ est consistant.

Procédure d'estimation à 2 pas

Transformées de Fourier des données ($j=1, \dots, J$) :

$$\hat{Y}_j(\omega) = \int_{\mathcal{D}} e^{-i\omega \cdot x} \frac{dY_j(x)}{2\pi} = a_j^{*2} \hat{f}(a_j^* R_{\theta_j^*} \omega) e^{i\omega \cdot b_j^*} + \epsilon \hat{W}_j(\omega), \quad \omega \in \mathbb{R}^2$$

1^{er} pas : estimation des rotations et des échelles $\alpha^* = (a_1^*, \dots, a_J^*, \theta_1^*, \dots, \theta_J^*)$:
à partir des estimateurs de la *Transformée de Fourier-Mellin* des $r^\sigma |\hat{f}_j|^2$:

$$\mathcal{M}_{|\hat{f}_j|^2}(k, \nu) = a_j^{*4-2\sigma+iv} \mathcal{M}_{|\hat{f}_j|^2}(k, \nu) e^{ik\theta_j^*} \quad (j=1, \dots, J)$$

L'estimateur $\hat{\alpha}_\epsilon = \arg \min_{\alpha} M_{1,\epsilon}$ est ϵ -consistant.

2^{ème} pas : estimation des translations $\beta^* = (b_1^*, \dots, b_J^*)$

L'estimateur $\hat{\beta}_\epsilon = \arg \min_{\alpha} M_{2,\epsilon}$ est consistant.

Procédure d'estimation à 2 pas

Transformées de Fourier des données ($j=1, \dots, J$) :

$$\hat{Y}_j(\omega) = \int_{\mathcal{D}} e^{-i\omega \cdot x} \frac{dY_j(x)}{2\pi} = a_j^{*2} \hat{f}(a_j^* R_{\theta_j^*} \omega) e^{i\omega \cdot b_j^*} + \epsilon \hat{W}_j(\omega), \quad \omega \in \mathbb{R}^2$$

1^{er} pas : estimation des rotations et des échelles $\alpha^* = (a_1^*, \dots, a_J^*, \theta_1^*, \dots, \theta_J^*)$:
à partir des estimateurs de la *Transformée de Fourier-Mellin* des $r^\sigma |\hat{f}_j|^2$:

$$\mathcal{M}_{|\hat{f}_j|^2}(k, \nu) = a_j^{*4-\sigma+iv} \mathcal{M}_{|\hat{f}_j|^2}(k, \nu) e^{ik\theta_j^*} \quad (j=1, \dots, J)$$

L'estimateur $\hat{\alpha}_\epsilon = \arg \min_{\alpha} M_{1,\epsilon}$ est ϵ -consistant.

2^{ème} pas : estimation des translations $\beta^* = (b_1^*, \dots, b_J^*)$
à partir des transformées de Fourier alignées ($j=1, \dots, J$) :

$$\hat{Z}_j(\beta, \omega) = \hat{a}_j^{-2} \hat{Y}_j(\hat{a}_j^{-1} A_{-\hat{\theta}_j} \omega) e^{i\omega \cdot \hat{a}_j^{-1} A_{\hat{\theta}_j} b_j}$$

L'estimateur $\hat{\beta}_\epsilon = \arg \min_{\alpha} M_{2,\epsilon}$ est consistant.

Applications numériques

- Les erreurs standard sont proportionnelles à $\frac{1}{s2n}$.

| | $s2n$ | $j = 2$ |
|------------------|-------|-----------------|
| a_j^* | | 1 |
| \hat{a}_j | 1 | 1.02 (0.00835) |
| | 3 | 1.01 (0.00373) |
| | 5 | 1.02 (0.00215) |
| $b_j^{1,*}$ | | 10 |
| \hat{b}_j^1 | 1 | 9.64 (1.92) |
| | 3 | 9.69 (1.01) |
| | 5 | 9.89 (0.49) |
| b_j^2 | | 0 |
| \hat{b}_j^2 | 1 | 0.107 (1.02) |
| | 3 | 0.00603 (0.508) |
| | 5 | 0.0217 (0.448) |
| θ_j^* | | 0 |
| $\hat{\theta}_j$ | 1 | 0.197 (0.641) |
| | 3 | 0.27 (0.935) |
| | 5 | 0.0642 (0.501) |

Tab.: Moyennes empiriques et écarts-type (entre parenthèses) des estimateurs $\hat{a}_j, \hat{b}_j = (\hat{b}_j^1, \hat{b}_j^2), \hat{\theta}_j$ sur $M = 50$ jeux de données pour différents rapports signal-sur-bruit $s2n$. Les images de départs sont de taille 100×100 pixels

Applications numériques

- Les erreurs standard sont proportionnelles à $\frac{1}{s2n}$.
- Certains estimateurs semblent être biaisés

| | $s2n$ | $j = 2$ |
|------------------|-------|-----------------|
| a_j^* | | 1 |
| \hat{a}_j | 1 | 1.02 (0.00835) |
| | 3 | 1.01 (0.00373) |
| | 5 | 1.02 (0.00215) |
| $b_j^{1,*}$ | | 10 |
| \hat{b}_j^1 | 1 | 9.64 (1.92) |
| | 3 | 9.69 (1.01) |
| | 5 | 9.89 (0.49) |
| b_j^2 | | 0 |
| \hat{b}_j^2 | 1 | 0.107 (1.02) |
| | 3 | 0.00603 (0.508) |
| | 5 | 0.0217 (0.448) |
| θ_j^* | | 0 |
| $\hat{\theta}_j$ | 1 | 0.197 (0.641) |
| | 3 | 0.27 (0.935) |
| | 5 | 0.0642 (0.501) |

Tab.: Moyennes empiriques et écarts-type (entre parenthèses) des estimateurs $\hat{a}_j, \hat{b}_j = (\hat{b}_j^1, \hat{b}_j^2), \hat{\theta}_j$ sur $M = 50$ jeux de données pour différents rapports signal-sur-bruit $s2n$. Les images de départs sont de taille 100×100 pixels

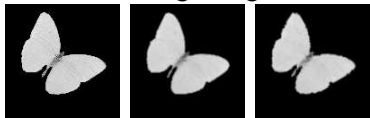
Applications numériques

- Les erreurs standard sont proportionnelles à $\frac{1}{s2n}$.
- Certains estimateurs semblent être biaisés

Un exemple réel :



Les mêmes images alignées :



| | $s2n$ | $j = 2$ |
|------------------|-------|-----------------|
| a_j^* | | 1 |
| \hat{a}_j | 1 | 1.02 (0.00835) |
| | 3 | 1.01 (0.00373) |
| | 5 | 1.02 (0.00215) |
| $b_j^{1,*}$ | | 10 |
| \hat{b}_j^1 | 1 | 9.64 (1.92) |
| | 3 | 9.69 (1.01) |
| | 5 | 9.89 (0.49) |
| b_j^2 | | 0 |
| \hat{b}_j^2 | 1 | 0.107 (1.02) |
| | 3 | 0.00603 (0.508) |
| | 5 | 0.0217 (0.448) |
| θ_j^* | | 0 |
| $\hat{\theta}_j$ | 1 | 0.197 (0.641) |
| | 3 | 0.27 (0.935) |
| | 5 | 0.0642 (0.501) |

Tab.: Moyennes empiriques et écarts-type (entre parenthèses) des estimateurs \hat{a}_j , $\hat{b}_j = (\hat{b}_j^1, \hat{b}_j^2)$, $\hat{\theta}_j$ sur $M = 50$ jeux de données pour différents rapports signal-sur-bruit $s2n$. Les images de départ sont de taille 100×100 pixels

Partie 3 : Alignement de séries temporelles

$$Y_{j,l} = f_{l-\theta_j^*} + \sigma_l \epsilon_{j,l}$$

Estimation des $\theta_j^* \in \mathbb{Z}$, $j=1, \dots, J$.

Laciar, E., Jane, R., Brooks, D.H. (2003) : alignement d'électrocardiogramme à partir de la "corrélation croisée" (cross-correlation) :

$$\theta \in \mathbb{Z} \rightarrow \frac{\sum_{l=\max(-n, -n-\theta)}^{\min(n, n-\theta)} Y_{1,l} Y_{2,(l-\theta)}}{\left(\sum_{|l| \leq n} |Y_{1,l}|^2\right)^{1/2} \left(\sum_{|l| \leq n} |Y_{2,l}|^2\right)^{1/2}}$$

- On développe la théorie duale de celle des signaux continus ;
- On généralise le critère de corrélation croisée pour un nombre J de signaux.

Le modèle

$$Y_{j,l} = f_{l-\theta_j^*} + \sigma_l \epsilon_{j,l}, \quad |l| \leq n, \quad j = 1, \dots, J$$

Le signal d'intérêt $(f_l)_{l \in \mathbb{Z}}$

La transformée de Fourier de (f_l) :

$$J_n(\lambda) = \frac{1}{\sqrt{n}} \sum_{|l| \leq n} f_l e^{-il\lambda}, \quad \lambda \in [-\pi, \pi].$$

L'autopériodogramme associé :

$$I_n(\lambda) = \frac{1}{2\pi} |J_n(\lambda)|^2, \quad \lambda \in [-\pi, \pi].$$

Propriété d'ergodicité : il existe une mesure F sur $[-\pi, \pi]$ telle que pour toute fonction bornée à valeurs réelles g sur $[-\pi, \pi]$, on a :

$$\lim_{n \rightarrow +\infty} \int_{-\pi}^{\pi} g(\lambda) I_n(\lambda) d\lambda = \int_{-\pi}^{\pi} g(\lambda) dF(\lambda).$$

La mesure F est la **distribution spectrale** associée au signal (f_l) .

Le bruit

La contrainte d'identifiabilité

Le modèle

$$Y_{j,l} = f_{l-\theta_j^*} + \sigma_l \epsilon_{j,l}, \quad |l| \leq n, \quad j = 1, \dots, J$$

Le signal d'intérêt $(f_l)_{l \in \mathbb{Z}}$

La transformée de Fourier de (f_l) :

$$J_n(\lambda) = \frac{1}{\sqrt{n}} \sum_{|l| \leq n} f_l e^{-il\lambda}, \quad \lambda \in [-\pi, \pi].$$

L'autopériodogramme associé :

$$I_n(\lambda) = \frac{1}{2\pi} |J_n(\lambda)|^2, \quad \lambda \in [-\pi, \pi].$$

Propriété d'ergodicité : il existe une mesure F sur $[-\pi, \pi]$ telle que pour toute fonction bornée à valeurs réelles g sur $[-\pi, \pi]$, on a :

$$\lim_{n \rightarrow +\infty} \int_{-\pi}^{\pi} g(\lambda) I_n(\lambda) d\lambda = \int_{-\pi}^{\pi} g(\lambda) dF(\lambda).$$

La mesure F est la **distribution spectrale** associée au signal (f_l) .

Le bruit

$(\epsilon_{ij})_{ij}$ sont i.i.d, $\mathbb{E}\epsilon_{ij} = 0$, $\mathbb{E}\epsilon_{ij}^2 = 1$.

Hypothèse sur la variance :

$$\frac{1}{n} \sum_{|l| \leq n} \sigma_l^2 = o(1).$$

La contrainte d'identifiabilité

$$\theta_1^* = 0$$

Le critère d'estimation

Transformées de Fourier des données, $j = 1, \dots, J$:

$$d_j(\lambda) = \frac{1}{\sqrt{n}} \sum_{|l| \leq n} Y_{j,l} e^{-i\lambda l}, \quad \lambda \in [-\pi, \pi].$$

Le processus de contraste, $\theta \in [-\theta_{\max}, \theta_{\max}]^J$:

$$M_n(\theta) = \int_{-\pi}^{\pi} \left| \frac{1}{J} \sum_{j=1}^J d_j(\lambda) e^{i\lambda \theta_j} \right|^2 \frac{d\lambda}{2\pi}.$$

Heuristique :

$$d_j(\lambda) \approx d_1(\lambda) e^{-i\lambda \theta_j^*}.$$

La fonction de contraste associée :

$$M(\theta) = \int_{-\pi}^{\pi} \left| \frac{1}{J} \sum_{j=1}^J e^{i\lambda(\theta_j - \theta_j^*)} \right|^2 dF(\lambda).$$

Le critère d'estimation

Transformées de Fourier des données, $j = 1, \dots, J$:

$$d_j(\lambda) = \frac{1}{\sqrt{n}} \sum_{|l| \leq n} Y_{j,l} e^{-i\lambda l}, \quad \lambda \in [-\pi, \pi].$$

Le processus de contraste, $\theta \in [-\theta_{\max}, \theta_{\max}]^J$:

$$M_n(\theta) = \int_{-\pi}^{\pi} \left| \frac{1}{J} \sum_{j=1}^J d_j(\lambda) e^{i\lambda \theta_j} \right|^2 \frac{d\lambda}{2\pi}.$$

Heuristique :

$$d_j(\lambda) \approx d_1(\lambda) e^{-i\lambda \theta_j^*}.$$

La fonction de contraste associée :

$$M(\theta) = \int_{-\pi}^{\pi} \left| \frac{1}{J} \sum_{j=1}^J e^{i\lambda(\theta_j - \theta_j^*)} \right|^2 dF(\lambda).$$

Théorème (Consistance)

Soit $\hat{\theta}_n$ l'estimateur :

$$\hat{\theta}_n = \arg \max \{ M_n(\theta), \theta \in [-\theta_{\max}, \theta_{\max}]^J \}.$$

Sous les hypothèses du modèle et si l'intérieur du support de la mesure spectrale F est supposé non vide, alors $\hat{\theta}_n$ converge en probabilité vers θ^* .

Commentaires

- Généralisation de la corrélation croisée en optimisant sur :

$$[-\theta_{\max}, \theta_{\max}]^J \cap \mathbb{Z}^J.$$

- Problème des maxima locaux : insertion d'un poids ψ ?

$$M_n(\theta) = \int_{-\pi}^{\pi} \left| \frac{1}{J} \sum_{j=1}^J d_j(\lambda) e^{i\lambda\theta_j} \right|^2 \psi(\lambda) \frac{d\lambda}{2\pi}.$$

- Robustesse du critère.

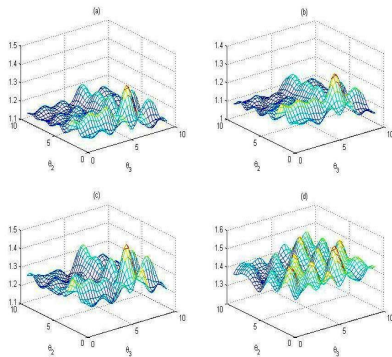


Fig.: Graphique du processus de contraste pour des variables i.i.d. ($J = 3$ et $n = 62$) et pour différentes valeurs de $\sigma = \sigma_I$: (a) $\sigma = 0$, (b) $\sigma = 0.2$, (c) $\sigma = 0.4$, (d) $\sigma = 0.6$.

Merci de votre attention