

# FILTRES LINÉAIRES MINIMISANT LE DÉPHASAGE POUR L'AJUSTEMENT SAISONNIER

*Fabien GUGGEMOS(\*), Dominique LADIRAY(\*), Michel GRUN-REHOMME(\*\*)*

*(\*) Institut National de la Statistique et des Études Économiques, Département des Statistiques de Court Terme, Paris, France*

*(\*\*) École Nationale de la Statistique et de l'Administration Économique, Paris, France*

## Introduction

Les techniques mises en oeuvre sur des séries temporelles économiques pour l'ajustement saisonnier ou l'extraction du cycle des affaires reposent le plus souvent sur l'utilisation de filtres linéaires, parmi lesquels les moyennes mobiles sont très populaires. Le logiciel X-12-Arima, employé dans de nombreux instituts nationaux de statistiques, fait par exemple appel aux moyennes mobiles dites de Henderson, ou plus généralement à des combinaisons de ces dernières pour estimer les principales composantes d'une série. Cependant, si des filtres symétriques sont en général appliqués au 'centre' des séries, des méthodes alternatives sont nécessaires lorsque l'on est confronté au problème d'estimation des points les plus récents. Dans le cas des logiciels X-12-Arima ou Tramo-Seats, des moyennes symétriques sont appliquées à des prévisions des points futurs obtenus grâce à une modélisation Arima de la série étudiée. Ces prévisions étant par construction des combinaisons linéaires des valeurs observées présente et passées, il s'ensuit qu'une telle pratique est équivalente à l'application d'un filtre asymétrique sur la fin de série.

Certes, les filtres linéaires - notamment les moyennes mobiles - asymétriques possèdent en général de bonnes propriétés ; beaucoup sont notamment construits de façon à réduire l'importance des révisions à effectuer ultérieurement, lorsque de nouvelles observations seront disponibles (cf [2, 3, 8]). En contrepartie, leur utilisation introduit également des phénomènes de déphasage dont l'une des principales conséquences est d'affecter l'estimation en temps réel des points de retournement de l'activité économique. La figure 1 donne une illustration de ce problème, où l'on peut observer l'estimation finale, réalisée en juin 2010 avec X-12-Arima et représentée par la courbe noire, de la tendance-cycle de l'indice de la production industrielle en France. Un point de retournement apparaît ainsi en janvier ou février 2001. Les courbes de différentes couleurs figurant également sur le graphique mettent en évidence les estimations successives de la tendance-cycle de la série, toujours produites à l'aide de X-12-Arima mais en n'utilisant que les données observées jusqu'en mars 2001, avril 2001, ..., octobre 2001. Il faut attendre le mois de juin 2001 pour que le point de retournement soit effectivement détecté.

Dans cette communication, nous nous proposons de définir un critère de temporalité sur lequel se fonder pour construire des moyennes mobiles asymétriques minimisant les phénomènes de déphasage. En l'incluant dans un cadre théorique plus général, nous

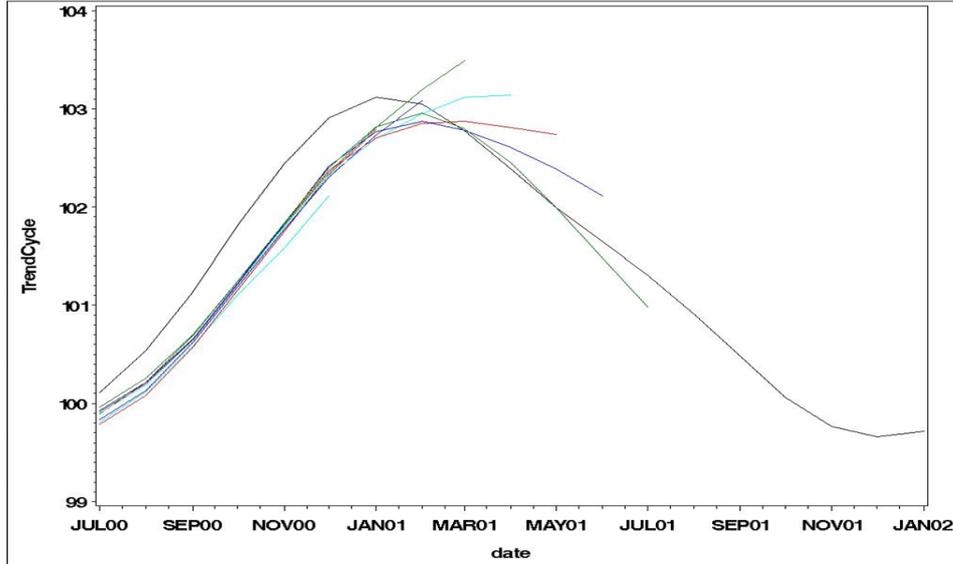


FIGURE 1 – Estimations successives en temps réel (courbes en couleur) et estimation finale (courbe noire) avec X-12-Arima de la tendance-cycle de l'indice de production industrielle en France.

pouvons ainsi établir un lien entre cette nouvelle méthode de construction de moyennes mobiles, les filtres fondés sur les critères de minimisation des révisions et ceux, dépendant des données, développés récemment dans le cadre de la méthode DFA généralisée (cf [11, 12, 13]). Ainsi, bien que l'on s'intéresse ici essentiellement à des moyennes mobiles dont les coefficients ne dépendent pas des caractéristiques de la série analysée, les résultats peuvent être aisément étendus à d'autres types de filtres linéaires.

La première partie introduit quelques définitions et met en évidence sur le plan théorique les principaux effets (gains et phases) induits par les moyennes mobiles. On rappelle également comment l'on peut construire des filtres linéaires préservant certaines formes de tendances ou supprimant certains types de saisonnalités. Dans un second temps est présenté le critère de révision développé par John Musgrave en 1964 ; celui-ci est ensuite mis en perspective avec l'approche DFA de Wildi. La troisième partie vise à montrer que tous ces types de filtres linéaires peuvent être perçus comme des cas particuliers d'un cadre théorique plus général. En jouant sur les hypothèses de modélisation sous-jacentes, ce dernier fournit une large classe de filtres linéaires, dépendant ou non des séries à analyser, et incluant notamment l'ensemble des moyennes mobiles précédemment introduites. L'approche DFA de wildi et ses extensions constitue alors un élément de comparaison pertinent pour le critère de temporalité que nous introduisons dans la quatrième partie afin de minimiser les effets de déphasage. Enfin, dans la cinquième partie, nous présentons la procédure globale de construction de moyennes mobiles asymétriques ainsi qu'un exemple sur données réelles.

## 1 Moyennes mobiles : définitions et propriétés

Une série temporelle discrète, que l'on notera  $(X_t)_{t \in \mathbb{Z}}$ , se décompose traditionnellement comme somme de trois termes, la tendance-cycle  $TC_t$ , la composante saisonnière  $S_t$  et la composante irrégulière  $I_t$  - avec  $X_t = TC_t + S_t + I_t$ . Dans le cas d'un modèle multiplicatif, la transformée logarithmique de  $(X_t)_{t \in \mathbb{Z}}$  permet de se ramener à ce cadre.

La méthode d'ajustement saisonnier X-12-Arima repose sur l'application de moyennes mobiles, lesquelles sont ainsi utilisées pour estimer ces trois composantes principales<sup>1</sup>.

## 1.1 Définitions

Considérons  $p$  et  $f$  deux entiers positifs et  $\boldsymbol{\theta} = (\theta_{-p}, \dots, \theta_{+f})'$  un vecteur de  $p + f + 1$  nombres réels. La moyenne mobile  $L_{\boldsymbol{\theta}}$  est alors définie comme l'endomorphisme linéaire sur l'espace vectoriel des séries temporelles discrètes qui, à une série quelconque  $(X_t)_{t \in \mathbb{Z}}$ , appelée également signal d'entrée, associe la série  $(Y_t)_{t \in \mathbb{Z}} = (L_{\boldsymbol{\theta}}X_t)_{t \in \mathbb{Z}}$ , appelée également signal de sortie, de sorte que :

$$\forall t \in \mathbb{Z}, Y_t = L_{\boldsymbol{\theta}}X_t = \sum_{k=-p}^{+f} \theta_k X_{t+k}.$$

La valeur à l'instant  $t$  de la série non-ajustée est par conséquent remplacée par une moyenne pondérée des  $p$  valeurs passées, de la valeur courante et des  $f$  valeurs futures de la série.

- ▷ L'entier  $p + f + 1$  est appelé **l'ordre de la moyenne mobile**.
- ▷ Lorsque  $p$  et  $f$  sont égaux, c'est-à-dire lorsque le nombre de points dans le passé et le nombre de points dans le futur impliqués dans l'estimation pour la date courante sont identiques, la moyenne mobile est dite **centrée**.
- ▷ Si en outre  $\theta_{-k} = \theta_k$  pour tout  $k$ , alors la moyenne mobile est dite **symétrique**.

Les observations  $X_t$  sont en général disponibles pour un ensemble fini d'indices  $t = 1, \dots, T$ . En conséquence, il est impossible d'appliquer une moyenne mobile d'ordre  $p + f + 1$  pour produire des estimations relatives aux  $p$  premiers points et aux  $f$  derniers points de la série. Dans les méthodes X-11 puis X-12, les moyennes mobiles symétriques jouent un rôle important, mais pour éviter de perdre de l'information en fin de série, on leur substitue des moyennes mobiles asymétriques adéquates. Lorsque l'on souhaite par exemple obtenir une estimation de la tendance-cycle à partir d'un modèle de décomposition des séries tel que celui mentionné précédemment,  $X_t = TC_t + S_t + I_t$ , en utilisant la moyenne mobile  $L_{\boldsymbol{\theta}}$ , on peut légitimement souhaiter que  $L_{\boldsymbol{\theta}}$  soit conçue pour supprimer la saisonnalité  $S_t$ , réduire la composante irrégulière  $I_t$  autant que possible et préserver au mieux la tendance-cycle  $TC_t$ .

## 1.2 Préservation des tendances et suppression des composantes saisonnières

Il est possible de déterminer les coefficients  $\{\theta_k\}_{-p \leq k \leq +f}$  d'une moyenne mobile de façon à ce que cette dernière préserve des formes simples de tendances, notamment de type polynômial. Pour qu'une moyenne mobile conserve par exemple les séries constantes, i.e. de la forme  $X_t = a$ , il est nécessaire et suffisant que  $a = L_{\boldsymbol{\theta}}X_t = \sum_k \theta_k X_{t+k} = a \cdot (\sum_k \theta_k)$ , c'est-à-dire que la somme des coefficients  $\sum_{k=-p}^{+f} \theta_k$  soit égale à 1. De façon similaire, la conservation de tendances linéaires  $X_t = at + b$  revient à ce que l'on ait, pour tout  $t$ ,  $X_t = at + b = L_{\boldsymbol{\theta}}X_t = \sum_k \theta_k [a(t+k) + b] = (at + b) \cdot (\sum_k \theta_k) + a \cdot (\sum_k k\theta_k)$ , ce qui

---

1. Dans X-12-Arima et Tramo-seats, les autres composantes comme les outliers ou les effets de jours ouvrables sont détectées et estimées avant la procédure d'ajustement saisonnier elle-même (cf [7] pour une présentation détaillée de l'algorithme de X-12-Arima).

se traduit par  $\sum_{k=-p}^{+f} \theta_k = 1$  et  $\sum_{k=-p}^{+f} k\theta_k = 0$ . Plus généralement, on peut montrer que la préservation de tendances polynômiales de degré  $d$  est strictement équivalente à l'ensemble suivant de contraintes portant sur les coefficients de la moyenne mobile :

$$\sum_{k=-p}^{+f} \theta_k = 1 \quad \text{et} \quad \sum_{k=-p}^{+f} k^j \theta_k = 0, \quad j = 1, \dots, d.$$

Il est également possible de construire une moyenne mobile qui supprime une composante saisonnière pouvant éventuellement varier polynômialement en fonction du temps  $t$ . Notons à cet effet  $S_t = (a_0 + a_1 t^1 + a_2 t^2 + \dots + a_d t^d) u_t$  une telle composante saisonnière où  $u_t$  est une fonction périodique de période  $\ell$  et définissons  $n$  comme l'entier satisfaisant  $p + f + 1 = n\ell$  où l'ordre  $p + f + 1$  de la moyenne mobile est déterminé en conséquence. On peut alors montrer (cf [4] pour une preuve complète) qu'une moyenne mobile supprime une composante saisonnière évoluant avec le temps comme un polynôme de degré  $d$  si et seulement si ses coefficients satisfont le système suivant de contraintes :

$$\begin{cases} \sum_{j=0}^{n-1} \theta_{k-j\ell} - \sum_{j=0}^{n-1} \theta_{f-j\ell} & = 0 \\ \sum_{j=0}^{n-1} (k - j\ell) \theta_{k-j\ell} - \sum_{j=0}^{n-1} (f - j\ell) \theta_{f-j\ell} & = 0 \\ \dots & \\ \sum_{j=0}^{n-1} (k - j\ell)^d \theta_{k-j\ell} - \sum_{j=0}^{n-1} (f - j\ell)^d \theta_{f-j\ell} & = 0 \end{cases}$$

Ainsi, la préservation des tendances et la suppression des saisonnalités consiste à imposer un certain nombre de contraintes linéaires que les coefficients du filtre doivent respecter. De façon plus générale, l'ensemble des vecteurs de coefficients admissibles  $\boldsymbol{\theta}$  dans  $\mathbb{R}^{p+f+1}$  - c'est-à-dire l'ensemble des vecteurs satisfaisant toutes les contraintes requises - sera supposé au moins **convexe** (ce qui est bien évidemment le cas si les contraintes sur les coefficients sont linéaires) et sera si besoin noté  $\Theta$ .

### 1.3 Effets d'amplitude et de déphasage

Afin d'identifier les deux principaux effets liés à l'application de moyennes mobiles, considérons à présent une série harmonique de fréquence  $\omega$ ,  $X_t(\omega) = e^{i\omega t}$ . La transformée de  $X_t$  par une moyenne mobile quelconque  $L_{\boldsymbol{\theta}}$  s'écrit alors :

$$Y_t = L_{\boldsymbol{\theta}} X_t = \sum_{k=-p}^{+f} \theta_k X_{t+k} = \sum_{k=-p}^{+f} \theta_k e^{i\omega(t+k)} = \left( \sum_{k=-p}^{+f} \theta_k e^{i\omega k} \right) \cdot X_t$$

Les séries harmoniques apparaissent donc naturellement comme les vecteurs propres de l'opérateur linéaire  $L_{\boldsymbol{\theta}}$ , dont les valeurs propres sont respectivement les valeurs prises aux différentes fréquences par la fonction de transfert du filtre,  $\widehat{L}_{\boldsymbol{\theta}}(e^{i\omega}) = \sum_{k=-p}^{+f} \theta_k e^{i\omega k}$ . Si, pour toute fréquence  $\omega$ ,  $\rho_{\boldsymbol{\theta}}(\omega)$ ,  $\varphi_{\boldsymbol{\theta}}(\omega)$  désignent respectivement le module et l'argument dans  $] - \pi; \pi ]$  de cette fonction de transfert, alors  $L_{\boldsymbol{\theta}} e^{i\omega t} = \rho_{\boldsymbol{\theta}}(\omega) e^{i[\omega t + \varphi_{\boldsymbol{\theta}}(\omega)]}$ , soit encore, pour rester dans le domaine des séries à valeurs réelles,

$$L_{\boldsymbol{\theta}} \cos(\omega t) = \rho_{\boldsymbol{\theta}}(\omega) \cos[\omega t + \varphi_{\boldsymbol{\theta}}(\omega)].$$

Appliquer une moyenne mobile à une série harmonique modifie donc cette dernière de deux façons distinctes : d'une part en la multipliant par un facteur d'amplitude  $\rho_{\boldsymbol{\theta}}(\omega)$ ,

d'autre part en la 'décalant' temporellement d'une période  $\varphi_{\theta}(\omega)/\omega$ , ce qui affecte notamment la détection des points de retournement de la série. Dans la théorie du signal,  $\rho_{\theta}(\omega)$  et  $\varphi_{\theta}(\omega)$  sont généralement appelés le gain et la phase de la fonction de transfert  $\widehat{L}_{\theta}(e^{i\omega})$ . Si l'on considère à présent une série temporelle *stationnaire* quelconque  $(X_t)_{t \in \mathbb{Z}}$ , celle-ci peut se décomposer en composantes harmoniques de différentes fréquences, chacune de ces dernières étant alors affectée par les deux effets mentionnés précédemment.

## 2 Moyennes mobiles optimales pour les révisions

### 2.1 L'approche de Musgrave

Les filtres utilisés fréquemment en pratique pour l'ajustement saisonnier sont généralement conçus pour répondre à des objectifs de nature statistique très précis, comme par exemple minimiser les révisions qu'induirait fatalement la prise en compte d'observations futures de la série lorsque de telles observations seront disponibles. En particulier, Doherty ([2]) a remis au goût du jour la méthode de Musgrave ([8]) pour le calcul dans X-12-Arima de la tendance-cycle à l'aide de moyennes mobiles asymétriques.

Le principe de la méthode de Musgrave est le suivant. Notons  $\{w_1, \dots, w_N\}$  (avec  $N = 2p + 1 = 2f + 1$ ) les poids de la moyenne mobile symétrique de Henderson – dont la définition est rappelée en partie 3.2.2. Par ailleurs  $\{\theta_1, \dots, \theta_M\}$  – avec  $p + 1 \leq M \leq 2p$  et  $\sum_{i=1}^M \theta_i = 1$  – désigne l'un des  $p$  vecteurs de coefficients des moyennes mobiles asymétriques de Musgrave. Il existe en effet  $p$  moyennes mobiles à déterminer, une pour chacun des  $p$  derniers points de la série. Pour une valeur de  $M$  fixé, supposons que les derniers points de la série suivent une tendance linéaire de la forme suivante,

$$X_t = a + bt + \varepsilon_t, \quad (1)$$

où les  $\varepsilon_t$  sont des variables aléatoires décorrélées de moyenne nulle et de variance constante  $\sigma^2$  (Noter que les paramètres  $a, b$  et  $\sigma^2$  dépendent de  $M$  et caractérisent un modèle appliqué aux points en fin de série, ils peuvent varier lorsque de nouvelles observations deviennent disponibles). Les poids  $\{\theta_1, \dots, \theta_M\}$  qui minimisent l'espérance du carré des révisions,

$$R(\theta) = E \left( \sum_{i=1}^M \theta_i X_i - \sum_{i=1}^N w_i X_i \right)^2,$$

sous la contrainte  $\sum_{i=1}^M \theta_i = 1$ , sont alors donnés par

$$\theta_i = w_i + \frac{1}{M} \sum_{j=M+1}^{2p+1} w_j + \frac{\left(i - \frac{M+1}{2}\right) D}{1 + \frac{M^3 - M}{12} D} \sum_{j=M+1}^{2p+1} \left(j - \frac{M+1}{2}\right) w_j, \quad i = 1, \dots, M$$

où  $D = b^2/\sigma^2$ .

Dans X-12-Arima comme dans X-11, l'ordre de la moyenne mobile de Henderson utilisée pour obtenir un estimateur de la tendance-cycle est fondée sur l'estimation du ratio suivant,

$$I/C = \frac{\sum |I_t - I_{t-1}|}{\sum |TC_t - TC_{t-1}|}$$

où  $TC_t$  et  $I_t$  représentent les composantes tendances-cycles et irrégulières de la série désaisonnalisée. Sous l'hypothèse supplémentaire que les  $\varepsilon_t$  sont normalement distribués,

les valeurs de  $D$  et  $I/C$  sont liées par l'égalité suivante,  $D = 4/(\pi(I/C)^2)$  (cf Doherty [2]). Dans X-12-Arima, les valeurs  $I/C = 1.0, 3.5$  et  $4.5$  sont utilisées avec les moyennes de Henderson d'ordre 9,13 et 23 respectivement puis pour les moyennes asymétriques de Musgrave correspondantes.

En réalité, l'approche adoptée par Musgrave peut être aisément généralisée de différentes façons, tout particulièrement en considérant des modélisations plus souples des fins de série et en choisissant d'autres types de poids symétriques de référence  $w_i$ . Ainsi, supposons désormais par exemple que les  $2p+1$  derniers points de la série suivent une tendance polynômiale de degré  $d$  et que, pour  $t = -p, \dots, +p$ ,  $X_t = a_0 + a_1 t^1 + \dots + a_d t^d + \varepsilon_t$ , où  $(\varepsilon_t)_t$  est un processus aléatoire stationnaire de moyenne nulle. En reprenant les notations introduites en partie 1.1, ces hypothèses se réécrivent matriciellement de la façon suivante :  $\mathbf{X} = \mathbf{A}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ , avec  $\mathbf{X} = [X_{-p}, X_{-p+1}, \dots, X_{+p}]'$ ,  $\boldsymbol{\beta} = [a_0, a_1, \dots, a_d]'$ ,  $\boldsymbol{\varepsilon} = [\varepsilon_{-p}, \varepsilon_{-p+1}, \dots, \varepsilon_{+p}]'$  et

$$\mathbf{A} = \begin{pmatrix} 1 & -p & (-p)^2 & \dots & (-p)^d \\ 1 & -p+1 & (-p+1)^2 & \dots & (-p+1)^d \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & +p & (+p)^2 & \dots & (+p)^d \end{pmatrix}.$$

Si  $\mathbf{w} = (w_{-p}, \dots, w_{+p})'$  désigne le vecteur des coefficients de la moyenne mobile symétrique de référence et  $\boldsymbol{\theta} = (\theta_{-p}, \dots, \theta_{+p})'$  le vecteur de la moyenne mobile asymétrique (avec  $\theta_k$  valant 0 si  $f+1 \leq k \leq p$ ), le critère de révision  $R$  à minimiser peut alors être écrit de la sorte,

$$R(\boldsymbol{\theta}) = \mathbb{E} \left[ \sum_{i=-p}^{+p} \theta_i X_i - \sum_{i=-p}^{+p} w_i X_i \right]^2 = (\boldsymbol{\theta} - \mathbf{w})' \mathbf{R} (\boldsymbol{\theta} - \mathbf{w})$$

où, comme  $\mathbb{E}(\boldsymbol{\varepsilon}) = \mathbf{0}$  et  $\text{Var}(\boldsymbol{\varepsilon}) = \boldsymbol{\Sigma}$ , la matrice  $\mathbf{R}$  est donnée par :

$$\mathbf{R} = \mathbb{E}(\mathbf{X}\mathbf{X}') = \mathbb{E}[(\mathbf{A}\boldsymbol{\beta} + \boldsymbol{\varepsilon})(\mathbf{A}\boldsymbol{\beta} + \boldsymbol{\varepsilon})'] = \mathbf{A}\boldsymbol{\beta}\boldsymbol{\beta}'\mathbf{A}' + \boldsymbol{\Sigma}.$$

## 2.2 Vers la méthode *Direct Filter Approach* généralisée

La façon de construire les moyennes mobiles de Musgrave ou leurs généralisations repose jusqu'à présent sur des modélisations paramétriques, qui impliquent que des hypothèses plus ou moins contraignantes soient effectuées sur le signal d'entrée. A l'inverse, on peut également considérer une approche non-paramétrique. En supposant seulement que l'erreur induite par le filtre asymétrique – i.e. la différence entre la série obtenue par application du filtre asymétrique,  $L_{\boldsymbol{\theta}}X_t$ , et celle obtenue par application du filtre symétrique,  $L_{\mathbf{w}}X_t$  – est un processus stationnaire, la méthode DFA (*Direct Filter Approach*, cf [11, 12]) consiste tout d'abord à transcrire le critère de révision dans le domaine fréquentiel. Pour un processus stationnaire  $X_t$ , l'erreur quadratique induite par le filtre asymétrique s'écrit en effet de la façon suivante :

$$\mathbb{E} (L_{\boldsymbol{\theta}}X_t - L_{\mathbf{w}}X_t)^2 = \int_0^{2\pi} \left| \widehat{L}_{\boldsymbol{\theta}}(e^{i\omega}) - \widehat{L}_{\mathbf{w}}(e^{i\omega}) \right|^2 dH_X(\omega), \quad (2)$$

où  $H_X(\omega)$  est la densité spectrale, inconnue, de  $X_t$ . Les poids (vecteur  $\boldsymbol{\theta}$ ) du filtre asymétrique peuvent alors être déterminés en minimisant une estimation du terme de droite dans l'équation (2), fondée sur une discrétisation de l'intervalle  $[0; 2\pi]$ . Le cas de processus intégrés non stationnaires peut se traiter de façon similaire, dans la mesure où la

cointégration entre les deux signaux (filtrés par les filtres symétriques et asymétriques) peut être obtenue en imposant des contraintes linéaires sur le vecteur  $\boldsymbol{\theta}$ , telles que celles présentées en partie 1.2.

Le principal intérêt de cette procédure réside dans le fait qu'il est à présent possible de décomposer le critère de révision en deux, de façon à isoler l'effet lié à la phase de celui lié au gain de la fonction de transfert. En effet, comme le montre Wildi (cf [13]) :

$$\begin{aligned} \left| \widehat{L}_{\boldsymbol{\theta}}(e^{i\omega}) - \widehat{L}_{\mathbf{w}}(e^{i\omega}) \right|^2 &= (\rho_{\boldsymbol{\theta}}(\omega) - \rho_{\mathbf{w}}(\omega))^2 + 2\rho_{\mathbf{w}}(\omega) \rho_{\boldsymbol{\theta}}(\omega) [1 - \cos(\varphi_{\boldsymbol{\theta}}(\omega))] \\ &= \underbrace{(\rho_{\boldsymbol{\theta}}(\omega) - \rho_{\mathbf{w}}(\omega))^2}_{\text{Effet lié au gain}} + \underbrace{4\rho_{\mathbf{w}}(\omega) \rho_{\boldsymbol{\theta}}(\omega) \sin^2(\varphi_{\boldsymbol{\theta}}(\omega)/2)}_{\text{Effet lié à la phase}}. \end{aligned} \quad (3)$$

Ces équations résultent du fait que la fonction de transfert d'un filtre symétrique est réelle et que par conséquent sa phase  $\varphi_{\mathbf{w}}(\omega)$  est toujours égale à 0 ou  $\pi$ . En pratique, on fera naturellement l'hypothèse qu'elle est toujours nulle (i.e. la fonction de transfert est positive), au moins pour les basses fréquences. Les contraintes usuelles auxquelles sont soumis les filtres, comme celle leur imposant de préserver les tendances constantes, garantissent qu'une telle hypothèse est satisfaite. En effet, par exemple, la condition  $\sum_k w_k = 1$  implique que  $\widehat{L}_{\mathbf{w}}(0)$  est égal à 1 et, comme la fonction de transfert est continue, que la phase  $\varphi_{\mathbf{w}}(\omega)$  est nulle aux basses fréquences.

Grâce à la décomposition du critère de révision établie précédemment, il est alors possible de moduler l'objectif de réduction du déphasage en regard de celui de préservation des composantes fréquentielles de la série initiale figurant dans la bande passante du filtre. C'est ainsi que Wildi suggère de modifier le critère de révision de la façon suivante,

$$\begin{aligned} \int_0^{2\pi} \left| \widehat{L}_{\boldsymbol{\theta}}(e^{i\omega}) - \widehat{L}_{\mathbf{w}}(e^{i\omega}) \right|^2 W(\omega) dH_X(\omega) \\ + 4\lambda \int_0^{2\pi} \rho_{\mathbf{w}}(\omega) \rho_{\boldsymbol{\theta}}(\omega) \sin^2(\varphi_{\boldsymbol{\theta}}(\omega)/2) W(\omega) dH_X(\omega), \end{aligned} \quad (4)$$

où  $\lambda$  est un facteur d'ajustement, permettant de pondérer différemment l'effet lié au gain de celui lié à la phase de la fonction de transfert, et  $W(\omega)$  est une fonction attribuant des poids distincts aux différentes fréquences de la bande passante du filtre. Augmenter  $\lambda$  revient à privilégier la réduction des effets de déphasage et donc à donner une importance accrue aux objectifs de détection en temps réel des points de retournement – ou plus généralement des changements de tendances.

Comme indiqué plus haut, la présence d'effets de déphasage est essentiellement due au caractère asymétrique du filtre employé ou, de façon équivalente, au fait que la fonction de transfert de celui-ci est à valeurs complexes. Par conséquent, une autre façon de réduire ces phénomènes peut consister à accorder un poids plus important à la partie imaginaire de la fonction de transfert dans le critère que l'on cherche à minimiser. Wildi a ainsi proposé récemment une généralisation de l'approche DFA, remplaçant  $\widehat{L}_{\boldsymbol{\theta}}(e^{i\omega})$  par  $\Re\left(\widehat{L}_{\boldsymbol{\theta}}(e^{i\omega})\right) + i\sqrt{1 + 4\lambda\widehat{L}_{\mathbf{w}}(e^{i\omega})} \cdot \Im\left(\widehat{L}_{\boldsymbol{\theta}}(e^{i\omega})\right)$  dans (2), où  $\Re$  et  $\Im$  désignent respectivement la partie réelle et la partie imaginaire. Si l'on développe l'expression (2) en tenant compte de cette modification et en pondérant les différentes fréquences de la bande passante du filtre, la méthode DFA généralisée de Wildi revient alors à minimiser un équivalent discret du

critère suivant :

$$\int_0^{2\pi} \left| \widehat{L}_{\boldsymbol{\theta}}(e^{i\omega}) - \widehat{L}_{\mathbf{w}}(e^{i\omega}) \right|^2 W(\omega) dH_X(\omega) + 4\lambda \int_0^{2\pi} \rho_{\mathbf{w}}(\omega) \rho_{\boldsymbol{\theta}}^2(\omega) \sin^2(\varphi_{\boldsymbol{\theta}}(\omega)) W(\omega) dH_X(\omega). \quad (5)$$

A la différence du précédent critère introduit en (4), ce dernier est une forme quadratique des coefficients du filtre, si bien que le problème de minimisation est simple à résoudre d'un point de vue opérationnel.

### 3 Un cadre théorique général englobant les principaux filtres linéaires utilisés dans les logiciels de désaisonnalisation

#### 3.1 Du critère de révision à un problème d'optimisation très général

Le critère de révision, sur lequel l'approche de Musgrave est fondée, de même que la méthode DFA reposent sur une mesure de distance entre le signal de sortie (résultant de l'application d'un filtre asymétrique) et celui que l'on obtiendrait si les observations futures de la série étaient disponibles en temps réel. En fait, le choix de la mesure de distance tout comme celui de la série à laquelle le signal de sortie est comparé peuvent être élargis. Les méthodes de construction de filtres linéaires présentées en section 2 peuvent ainsi être vues comme des cas particuliers du problème d'optimisation plus général suivant,

$$\begin{cases} \min_{\boldsymbol{\theta}} \text{E} [|\nabla^q(L_{\boldsymbol{\theta}}X_t - u_t)|^2] \\ \text{sous contrainte } \mathbf{C}\boldsymbol{\theta} = \mathbf{a}, \end{cases} \quad (6)$$

où  $\nabla$  est l'opérateur de différenciation,  $\nabla X_t = X_t - X_{t-1}$ , et, pour tout entier  $q \geq 1$ ,  $\nabla^q = \nabla \circ \nabla^{q-1} = \underbrace{\nabla \circ \dots \circ \nabla}_{q \text{ facteurs}}$  (avec  $\nabla^0$  désignant par convention l'opérateur identité).

Les coefficients du filtre sont choisis de sorte qu'ils minimisent la différence d'ordre  $q$  de l'écart entre le signal de sortie et une série dite de référence  $(u_t)_{t \in \mathbb{Z}}$ . En outre, ils doivent satisfaire des contraintes linéaires préétablies, représentées sous forme matricielle par l'équation  $\mathbf{C}\boldsymbol{\theta} = \mathbf{a}$ . Le nombre de lignes de la matrice  $\mathbf{C}$  et du vecteur  $\mathbf{a}$  est égal au nombre de ces contraintes. Celles-ci étant supposées linéairement indépendantes,  $\mathbf{C}$  est une matrice de plein rang ligne.

La série de référence  $u_t$  est choisie de façon à ce qu'elle constitue un estimateur robuste de la tendance-cycle de la série initiale  $X_t$ . Elle peut être déterministe ou stochastique et peut dépendre elle-même des coefficients du filtre. L'ordre de différenciation et les contraintes linéaires sont déterminés quant à eux pour que le filtre satisfasse un certain nombre de propriétés souhaitées, comme la préservation de certaines formes de tendances ou encore la suppression de certains types de saisonnalités (cf partie 1.2), de sorte qu'in fine le processus  $\nabla^q(L_{\boldsymbol{\theta}}X_t - u_t)$  soit stationnaire.

Au-delà des considérations précédentes, les hypothèses de modélisation sur la série initiale  $X_t$  peuvent être librement choisies. En faisant varier ces hypothèses, on obtient

ainsi une large classe de filtres linéaires. Les moyennes mobiles de Musgrave et la méthode DFA sont des cas particuliers pour lesquels  $q = 0$  et  $u_t = L_{\mathbf{w}}X_t$ . Dans le premier cas, un modèle paramétrique est utilisé tandis que, dans le second cas, les coefficients du filtre dépendent directement d'une estimation non-paramétrique de la densité spectrale des données. Dans les paragraphes suivants, nous verrons que des contraintes additionnelles permettent de retrouver des filtres classiques ne dépendant pas des données.

## 3.2 Cas particuliers

### 3.2.1 Fidélité

Une première application du cadre général donné en partie 3.1 fournit les moyennes mobiles suggérées dans la littérature par Bongard (cf [1]). En effet, l'approche de Bongard consiste à choisir les moyennes mobiles qui réduisent au maximum la composante irrégulière d'une série une fois la tendance préservée et les saisonnalités supprimées. Dans le modèle de décomposition de la série  $X_t$  introduit en partie 1, la composante irrégulière résiduelle est souvent modélisée par un bruit blanc  $\varepsilon_t$ , de moyenne nulle et de variance constante  $\sigma^2$ . Ce bruit blanc est alors transformé en une séquence de variables aléatoires  $\varepsilon_t^*$ , avec une variance constante égale à  $\sigma^2 \sum_{k=-p}^{+f} \theta_k^2$ . Réduire la composante irrégulière, (sa variance puisqu'elle de moyenne nulle), revient alors à minimiser le critère  $\sum_{k=-p}^{+f} \theta_k^2$ . Le signal de sortie étant en quelque sorte 'aussi proche que possible' du signal d'entrée pour lequel le bruit a été supprimé, ce critère sera appelé critère de Fidélité et noté  $F(\boldsymbol{\theta})$ . Si l'on considère des bruits blancs additifs, le critère de Fidélité est simplement une forme quadratique définie positive,

$$F(\boldsymbol{\theta}) = \sum_{k=-p}^{+f} \theta_k^2 = \boldsymbol{\theta}'\mathbf{F}\boldsymbol{\theta},$$

où  $\mathbf{F}$  est la matrice identité d'ordre  $p+f+1$ . L'approche de Bongard peut être directement obtenue en considérant le problème général (6), où  $q = 0$  et où  $u_t$  est la série de référence déterministe  $u_t = E(L_{\boldsymbol{\theta}}X_t)$ .

### 3.2.2 Lissage

X-12-Arima utilise les moyennes mobiles de Henderson pour extraire la tendance-cycle à partir d'une estimation de la série déjà désaisonnalisée (Tables B7, C7, D7, D12 dans X-12-Arima; en d'autres termes, l'ajustement saisonnier est effectué dans un premier temps). Dans le cas d'un modèle de décomposition additif de la série, celui-ci se simplifie alors de la sorte,  $X_t = TC_t + I_t$ . Henderson suggère alors d'utiliser une moyenne mobile obtenue à partir d'un critère qui garantisse une estimation robuste et suffisamment lisse de la tendance-cycle. Si l'on considère la série de Dirac  $\delta_t^{t_0}$ , valant 1 à l'instant  $t = t_0$  et 0 à tout autre instant, l'application d'une moyenne mobile  $L_{\boldsymbol{\theta}}$  d'ordre  $p+f+1$  et de coefficients  $\{\theta_k\}$  la transforme alors en la série suivante :

$$L_{\boldsymbol{\theta}}\delta_t^{t_0} = \begin{cases} \theta_{t_0-t} & \text{si } t_0 - f \leq t \leq t_0 + p \\ 0 & \text{sinon} \end{cases} .$$

Cette série transformée sera par conséquent 'lisse' si la courbe des coefficients de la moyenne mobile elle-même ne présente pas trop d'irrégularités. En remarquant que l'espace des séries est engendré par les séries de Dirac, puisque toute série temporelle  $X_t$

peut s'écrire  $X_t = \sum_{t_0 \in \mathbb{Z}} X_{t_0} \delta_t^{t_0}$ , Henderson [5, 6] propose alors d'utiliser la quantité  $S = \sum_k (\nabla^3 \theta_k)^2$  pour mesurer la 'régularité' de la courbe des coefficients. La notation  $\nabla$ , introduite dans la partie 3.1 comme opérateur linéaire sur les séries temporelles, est ici étendue, avec un sens similaire, comme opérateur linéaire de différenciation sur l'ensemble des vecteurs  $\boldsymbol{\theta}$ , i.e.  $(\nabla \boldsymbol{\theta})_k = \nabla \theta_k = \theta_k - \theta_{k-1}$ . La quantité  $S$  s'annule dès lors que les coefficients  $\{\theta_k\}$  sont situés le long d'une droite ou d'une parabole. Henderson détermine alors les moyennes mobiles centrées qui préservent les poynômes de degré au plus 2 et qui minimisent  $S$ .

Minimiser le critère d'Henderson, ou plus généralement le critère de 'Lissage',

$$S(\boldsymbol{\theta}) = \sum_k (\nabla^q \theta_k)^2,$$

visé donc à obtenir des signaux de sortie relativement réguliers. On peut noter que  $S(\boldsymbol{\theta})$ , comme le critère de Fidélité, est une forme quadratique définie positive sur  $\mathbb{R}^{p+f+1}$ . Par exemple, dans le cas particulier d'Henderson  $q = 3$ , on obtient  $S(\boldsymbol{\theta}) = \boldsymbol{\theta}' \mathbf{S} \boldsymbol{\theta}$ , où  $\mathbf{S}$  est la matrice symétrique de Toeplitz suivante :

$$\mathbf{S} = \begin{pmatrix} 20 & -15 & 6 & -1 & 0 & \dots & 0 \\ -15 & 20 & -15 & 6 & \ddots & \ddots & \vdots \\ 6 & -15 & 20 & \ddots & \ddots & \ddots & 0 \\ -1 & 6 & \ddots & \ddots & \ddots & 6 & -1 \\ 0 & \ddots & \ddots & \ddots & 20 & -15 & 6 \\ \vdots & \ddots & \ddots & 6 & -15 & 20 & -15 \\ 0 & \dots & 0 & -1 & 6 & -15 & 20 \end{pmatrix}.$$

En fait, en posant  $\theta_k = 0$  pour tout entier  $k \notin [-p; +f]$ , on peut facilement vérifier que  $\nabla^q (L_{\boldsymbol{\theta}} X_t) = (-1)^q \sum_{k \in \mathbb{Z}} (\nabla^q \theta_k) X_{t+k-q}$ . En conséquence, avec des hypothèses similaires à celles faites pour le critère de Fidélité en partie 3.2.1 sur la composante irrégulière de la série, l'approche de Henderson apparaît alors comme un cas particulier du problème (6) où  $u_t$  est de nouveau la série de référence déterministe  $u_t = E(L_{\boldsymbol{\theta}} X_t)$ , mais où  $q$  est désormais un entier strictement positif.

D'autres types de filtres linéaires classiques, comme le filtre d'Hodrick-Prescott notamment, peuvent être vus comme cas particuliers du problème général (6).

## 4 Le critère de Temporalité

### 4.1 Minimiser le déphasage

Comme cela a été évoqué dans la partie 1.3, une façon de réduire le déphasage observé entre signaux d'entrée et de sortie consiste à déterminer des moyennes mobiles qui n'induisent qu'un faible effet de déphasage lorsqu'elles sont appliquées à des signaux harmoniques  $(e^{i\omega t})_{t \in \mathbb{Z}}$ . Autrement dit, il s'agit de rechercher des vecteurs  $\boldsymbol{\theta}$  tels que la phase  $\varphi_{\boldsymbol{\theta}}(\omega)$  de la fonction de transfert du filtre  $L_{\boldsymbol{\theta}}$  soit proche de 0. Une telle approche devrait fournir des moyennes mobiles possédant *a priori* des propriétés satisfaisantes en termes de déphasage, quelles que soient les séries temporelles auxquelles elles sont appliquées.

Si un vecteur  $\boldsymbol{\theta}$  est tel que  $\varphi_{\boldsymbol{\theta}}(\omega_0) = 0$ , alors les deux signaux  $(e^{i\omega_0 t})_{t \in \mathbb{Z}}$  and  $(L_{\boldsymbol{\theta}} e^{i\omega_0 t})_{t \in \mathbb{Z}}$  sont en phase, mais cette propriété n'est plus satisfaite pour tout couple de signaux

$(e^{i\omega t})_{t \in \mathbb{Z}}$  et  $(L_{\theta} e^{i\omega t})_{t \in \mathbb{Z}}$  dès lors qu'on considère une fréquence  $\omega$  distincte de  $\omega_0$ . En fait, comme la fonction de transfert est une fonction régulière (en l'occurrence, infiniment différentiable) et que sa phase est choisie dans l'intervalle  $] - \pi; \pi ]$ , on peut seulement affirmer que  $\varphi_{\theta}(\omega)$  est également une fonction régulière, donc est proche de 0, dans un voisinage de la fréquence  $\omega_0$ . Puisque tout signal stationnaire peut se décomposer en une combinaison linéaire de signaux harmoniques caractérisés par des fréquences distinctes, les coefficients  $\theta$  du filtre doivent être choisis parmi ceux pour lesquels la fonction de phase est proche de 0 quelle que soit la fréquence  $\omega$  considérée. En pratique, nous considérons plutôt l'intervalle de fréquences  $] \omega_1; \omega_2 [$  qui figurent dans le spectre des composantes stationnaires des séries temporelles auxquelles la moyenne mobile  $L_{\theta}$  sera appliquée, avec  $0 \leq \omega_1 \leq \omega_2 \leq 2\pi$ .

De plus, lors de l'application du filtre, la composante harmonique du signal d'entrée de fréquence  $\omega$  est multipliée par un facteur d'amplitude égal au module de la fonction de transfert. En conséquence, il est assez naturel de considérer que prendre en compte l'impact des effets de déphasage sur la composante harmonique de fréquence  $\omega$  est d'autant plus important que la valeur de  $\rho_{\theta}(\omega)$  est élevée.

Pour ces différentes raisons, nous suggérons – pour limiter les phénomènes de déphasage entre signaux d'entrée et de sortie – d'utiliser une moyenne mobile dont les coefficients minimisent le critère suivant,

$$\int_{\omega_1}^{\omega_2} f[\rho_{\theta}(\omega); \varphi_{\theta}(\omega)] d\omega, \quad (7)$$

où la fonction  $f$ , définie sur  $[0; +\infty[ \times ] - \pi; \pi ]$  est choisie au préalable et satisfait ces six conditions :

1.  $f \geq 0$
2.  $f(\rho, 0) = 0$
3.  $f(0, \varphi) = 0$
4.  $f(\rho, \varphi) = f(\rho, -\varphi)$
5.  $\frac{\partial f}{\partial \rho} \geq 0$
6.  $\varphi \cdot \frac{\partial f}{\partial \varphi} \geq 0$

La fonction  $f$  sera appelée fonction de pénalité dans le reste de ce papier. Les conditions 1 et 2 garantissent que cette fonction est minimisée lorsque la phase est nulle. Les conditions 4 et 6 signifient que, dans le critère (7), la phase est pénalisée d'autant plus qu'elle est éloignée de 0. De plus, la condition 5 est la traduction mathématique du fait que le déphasage induit pour une fréquence donnée doit être d'autant plus pénalisé que cette fréquence est amplifiée par le filtre. Enfin, dans le cas extrême d'une fréquence coupée par le filtre, la condition 3 rappelle qu'il est inutile d'accorder dans le critère une quelconque importance à la phase correspondante.

## 4.2 Choix de fonctions de pénalités adéquates

À présent, il faut choisir une fonction de pénalité convenable. Il serait souhaitable de trouver une fonction satisfaisant les six conditions précédentes et telle que le problème consistant à minimiser (7) puisse être résolu analytiquement ou, si ce n'est pas possible, numériquement. Dans ce dernier cas, il s'agit de s'assurer de l'existence d'une solution théorique et d'un algorithme numérique convergeant rapidement vers cette solution. En

particulier, une fonction de pénalité  $f$  produisant un critère convexe serait tout à fait appropriée puisque minimiser la quantité donnée par (7) ne serait qu'un problème classique d'optimisation convexe pour la résolution duquel on dispose d'algorithmes bien connus (méthodes du gradient par exemple).

**Proposition 4.2.1** (Une première famille de fonctions de pénalité). *Pour tout  $k > 0$  et  $\ell > 0$ , la fonction  $(\rho, \varphi) \mapsto \rho^k |\varphi|^\ell$  satisfait les six conditions auxquelles sont soumises les fonctions de pénalité apparaissant dans le critère (7).*

Cela étant, on recommandera de choisir le second paramètre,  $\ell$ , plus grand que 1. En effet, dans le cas contraire, la dérivée partielle  $|\partial f / \partial \varphi|$  tendrait vers  $+\infty$  lorsque  $\varphi$  tend vers 0, si bien que les phases non nulles mais très proches de 0 risqueraient d'être trop fortement pénalisées.

**Proposition 4.2.2** (Une seconde famille de fonctions de pénalité). *Pour tout  $k > 0$  et  $\ell > 0$ , la fonction  $(\rho, \varphi) \mapsto \rho^k \left| \sin\left(\frac{\varphi}{2}\right) \right|^\ell$  satisfait les six conditions auxquelles sont soumises les fonctions de pénalité apparaissant dans le critère (7).*

A nouveau, il est préférable d'utiliser une telle fonction avec un paramètre  $\ell$  supérieur ou égal à 1. Remarquons que les fonctions de cette seconde famille sont  $2\pi$ -périodiques par rapport à la variable  $\varphi$  (si bien qu'il n'est plus nécessaire d'imposer à la phase d'être dans l'intervalle  $]-\pi; \pi]$ ).

Ici, nous suggérons de choisir la fonction de pénalité donnée par la proposition 4.2.2 avec les paramètres  $k = 1$  et  $\ell = 2$ . En d'autres termes,

$$f[\rho_\theta(\omega); \varphi_\theta(\omega)] = \rho_\theta(\omega) \sin^2\left(\frac{\varphi_\theta(\omega)}{2}\right) = \frac{1}{2} \left[ \rho_\theta(\omega) - \Re\left(\widehat{L}_\theta(e^{i\omega})\right) \right] \quad (8)$$

Comme cela sera établi plus loin, une telle fonction de pénalité est particulièrement intéressante dans la mesure où elle rend le critère (7) convexe. On peut également souligner qu'elle correspond précisément à l'effet de déphasage apparaissant dans l'équation (3), qui est ensuite pondéré dans le critère de révision (4) pour la méthode DFA de Wildi. Toutefois, dans cette partie, le filtre est supposé être indépendant des données, ses propriétés en termes de déphasage sont analysées directement sans être comparées à celles d'un filtre symétrique de référence.

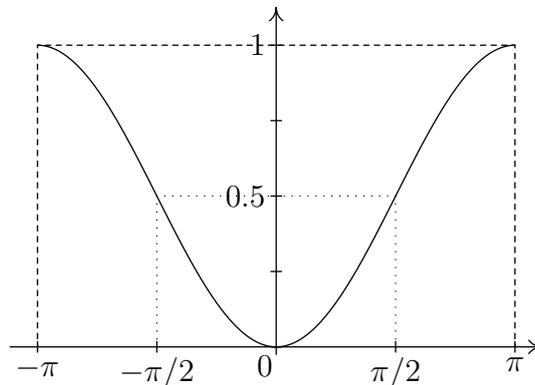


FIGURE 2 – Fonction de pénalité  $\varphi \mapsto f(1, \varphi) = \sin^2\left(\frac{\varphi}{2}\right)$

Désignons par  $\mathbf{Co}(\omega)$  et  $\mathbf{Si}(\omega)$  les vecteurs d'ordre  $p + f + 1$  définis respectivement par  $[\cos(-p\omega), \dots, \cos(f\omega)]'$  et  $[\sin(-p\omega), \dots, \sin(f\omega)]'$ . Notons également  $\mathbf{\Omega}(\omega)$

la matrice carrée d'ordre  $p + f + 1$  dont l'élément à l'intersection de la  $k$ -ième ligne et de la  $\ell$ -ième colonne est donné par  $\Omega_{k\ell} = \cos[(k - \ell)\omega]$ , pour tout couple d'indices  $-p \leq k, \ell \leq +f$ . On a alors  $\Re(\widehat{L}_\theta(e^{i\omega})) = \mathbf{Co}(\omega)' \boldsymbol{\theta}$ ,  $\Im(\widehat{L}_\theta(e^{i\omega})) = \mathbf{Si}(\omega)' \boldsymbol{\theta}$  et  $\Omega(\omega) = \mathbf{Co}(\omega) \mathbf{Co}(\omega)' + \mathbf{Si}(\omega) \mathbf{Si}(\omega)'$ , si bien que  $\rho_\theta^2(\omega) = \boldsymbol{\theta}' \Omega(\omega) \boldsymbol{\theta}$ . En conséquence, la matrice  $\Omega(\omega)$  est symétrique positive (mais non inversible, de rang au plus égal à 2).

Si l'on considère la fonction de pénalité donnée par (8), le critère (7) peut alors être réécrit de la façon suivante :

$$T_0(\boldsymbol{\theta}) = \int_{\omega_1}^{\omega_2} \sqrt{\boldsymbol{\theta}' \Omega(\omega) \boldsymbol{\theta}} d\omega - \left( \int_{\omega_1}^{\omega_2} \mathbf{Co}(\omega) d\omega \right)' \boldsymbol{\theta} \quad (9)$$

**Proposition 4.2.3.** *Le critère  $T_0(\boldsymbol{\theta})$  donné par (9) est une fonction convexe du vecteur des coefficients  $\boldsymbol{\theta}$ .*

L'objectif mentionné au début de cette partie est donc atteint : la proposition 4.2.3 implique que le problème de minimisation du déphasage pour la fonction de pénalité donnée par (8), en tant que problème d'optimisation convexe, possède une solution et peut être résolu numériquement en utilisant par exemple la méthode du gradient. Dans cette optique, signalons l'expression du gradient du critère  $T_0(\boldsymbol{\theta})$ .

**Proposition 4.2.4.** *Le gradient du critère  $T_0(\boldsymbol{\theta})$  figurant dans l'équation (9) est donné par :*

$$\frac{\partial T_0(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \left( \frac{1}{2} \int_{\omega_1}^{\omega_2} \frac{\Omega(\omega) \boldsymbol{\theta}}{\sqrt{\boldsymbol{\theta}' \Omega(\omega) \boldsymbol{\theta}}} d\omega \right) - \left( \int_{\omega_1}^{\omega_2} \mathbf{Co}(\omega) d\omega \right)$$

### 4.3 Variantes du critère de Temporalité

Dans cette partie, nous proposons de relâcher certaines contraintes sur la fonction de pénalité pour obtenir un critère de Temporalité plus aisé à utiliser. Comme cela a déjà été mentionné en partie 2.2, une autre façon de limiter les phénomènes de déphasage est de réduire la partie imaginaire de la fonction de transfert. Agir de la sorte revient à considérer la fonction de phase *modulo*  $\pi$  et donc à tolérer les cas d'opposition de phase entre signal d'entrée et signal de sortie. Géométriquement, cela signifie simplement que l'on cherche à réduire autant que possible la distance séparant, dans le plan complexe, les valeurs de la fonction de transfert de la droite réelle. Formellement, cela se traduit par le fait que la fonction de pénalité  $f$  doit désormais satisfaire les six conditions exposées en partie 4.1, mais avec les quelques modifications suivantes : tout d'abord, la sixième condition doit être à présent valable seulement pour les phases  $\varphi$  de l'intervalle  $[-\pi/2; \pi/2]$  ; ensuite une septième condition doit être satisfaite, stipulant que, pour tout phase  $\varphi \in ]\pi/2; \pi]$ ,  $f(\rho, \varphi) = f(\rho, \pi - \varphi)$ . Une famille de fonctions de pénalité satisfaisant ces sept conditions est donnée ci-dessous :

**Proposition 4.3.1.** *Pour tout  $k > 0$  et  $\ell > 0$ , la fonction  $(\rho, \varphi) \mapsto \rho^k |\sin(\varphi)|^\ell$  satisfait les sept conditions auxquelles sont soumises les fonctions de pénalité apparaissant dans le problème où les phénomènes de déphasage sont minimisés modulo  $\pi$ .*

Parmi ces fonctions, la suivante, obtenue pour  $k = \ell = 2$ ,

$$f[\rho_\theta(\omega); \varphi_\theta(\omega)] = \rho_\theta^2(\omega) \sin^2(\varphi_\theta(\omega)) = \left[ \Im(\widehat{L}_\theta(e^{i\omega})) \right]^2, \quad (10)$$

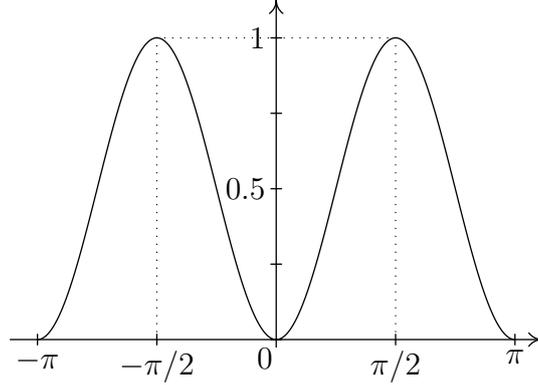


FIGURE 3 – Fonction de pénalité  $\varphi \mapsto f(1, \varphi) = \sin^2(\varphi)$

est particulièrement intéressante puisque, cette fois-ci, elle rend le problème de minimisation du critère de Temporalité résoluble analytiquement. En effet, ce nouveau critère, que l'on notera  $T(\boldsymbol{\theta})$ , est une forme quadratique positive :

$$T(\boldsymbol{\theta}) = \int_{\omega_1}^{\omega_2} y_{\boldsymbol{\theta}}^2(\omega) d\omega = \boldsymbol{\theta}' \mathbf{T} \boldsymbol{\theta}, \quad (11)$$

où  $\mathbf{T} = \int_{\omega_1}^{\omega_2} \mathbf{Si}(\omega) \mathbf{Si}(\omega)' d\omega$  est une matrice carrée d'ordre  $p + f + 1$ . Après quelques menus calculs, on obtient en effet que les éléments de cette matrice sont donnés par les formules suivantes, pour  $-p \leq k, \ell \leq f$  :

$$\mathbf{T}_{kl} = \begin{cases} \frac{\sin((k-\ell)\omega_2) - \sin((k-\ell)\omega_1)}{2(k-\ell)} - \frac{\sin((k+\ell)\omega_2) - \sin((k+\ell)\omega_1)}{2(k+\ell)} & \text{si } |k| \neq |\ell| \text{ et } k\ell \neq 0 \\ \frac{\omega_2 - \omega_1}{2} - \frac{\sin(2k\omega_2) - \sin(2k\omega_1)}{4k} & \text{si } k = \ell \text{ et } k\ell \neq 0 \\ \frac{\sin(2k\omega_2) - \sin(2k\omega_1)}{4k} - \frac{\omega_2 - \omega_1}{2} & \text{si } k = -\ell \text{ et } k\ell \neq 0 \\ 0 & \text{si } k\ell = 0 \end{cases}$$

Sans surprise, contrôler les effets de déphasage via la partie imaginaire de la fonction de transfert du filtre conduit à un critère (11) qui est l'équivalent, indépendant des données à analyser, du second membre du critère (5) sur lequel s'appuie la méthode DFA généralisée. Son principal inconvénient réside dans le fait que les cas où le filtre induit une opposition de phase ne sont pas pénalisés. En pratique, cela n'est guère gênant puisque, pour les mêmes raisons que celles invoquées en partie 2.2, la fonction de transfert est proche de 1 et par conséquent la phase proche de 0 pour les basses fréquences. Ainsi, les oppositions de phases ne peuvent réellement survenir que pour des fréquences suffisamment élevées ; or les contraintes relatives à la préservation des tendances et à la suppression des saisonnalités et auxquelles sont usuellement soumis les coefficients du filtre visent précisément à éliminer ou au moins à réduire fortement ce type de fréquences dans la série temporelle que l'on analyse.

## 5 Construction de moyennes mobiles asymétriques : procédure opérationnelle

### 5.1 Un critère combinant divers objectifs

Il est important de souligner qu'il y aura toujours un nombre infini de solutions  $\boldsymbol{\theta}$  au problème de minimisation du critère de Temporalité. Par exemple, le critère  $T_0$  est convexe, mais pas strictement convexe, de même que le critère  $T$  est seulement une forme quadratique *semi-définie* positive (la matrice sous-jacente n'est pas inversible). Ce dernier est notamment nul, donc minimal, sur tout le sous-espace vectoriel des vecteurs symétriques, i.e. pour tous les vecteurs  $\boldsymbol{\theta}$  tels que  $\theta_k = \theta_{-k}$  si  $|k| \leq \min(p, f)$  et  $\theta_k = 0$  sinon.

Les procédures classiques pour choisir les coefficients d'une moyenne mobile sont généralement fondées sur la minimisation des critères de Fidélité, de Lissage ou de Révision. Une idée assez naturelle consisterait maintenant à 'mixer' ces différents critères avec le critère de Temporalité pour tenir compte de l'ensemble des déformations ou biais induits par l'utilisation d'une moyenne mobile. En utilisant les notations introduites en parties 3.1, 3.2 et 4.3 pour décrire ces différents critères ainsi que les contraintes relatives aux tendances et saisonnalités, nous suggérons de suivre la procédure opérationnelle suivante pour construire des moyennes mobiles minimisant les phénomènes de déphasage :

**Problème 5.1.1.** *La procédure opérationnelle FST (Fidelity - Smoothness - Timeliness)*

$$\left\{ \begin{array}{l} \min_{\boldsymbol{\theta}} J(\boldsymbol{\theta}) \\ \text{sous contrainte } \mathbf{C}\boldsymbol{\theta} = \mathbf{a} \end{array} \right., \text{ où } J(\boldsymbol{\theta}) = \alpha \cdot F(\boldsymbol{\theta}) + \beta \cdot S(\boldsymbol{\theta}) + \gamma \cdot T(\boldsymbol{\theta}),$$

les trois nombres réels  $\alpha, \beta, \gamma$  étant fixés dans l'intervalle  $[0; 1]$  et satisfaisant deux conditions,  $\alpha + \beta + \gamma = 1$  et  $\alpha\beta \neq 0$ .

Un tel problème d'optimisation admet une unique solution.

La première condition sur les nombres réels  $\alpha, \beta, \gamma$  signifie que le critère mixte  $J(\boldsymbol{\theta})$  est une combinaison linéaire convexe des critères de Fidélité, Lissage et Temporalité (On ne considère pas ici le critère de révision dont on peut en fait prouver qu'il est très fortement corrélé à une combinaison fixée, donc non ajustable par l'utilisateur, des critères de Fidélité et de Temporalité). La seconde condition garantit que  $J(\boldsymbol{\theta})$  est une fonction strictement convexe, ce qui justifie que la solution du problème global 5.1.1 est unique. Au passage, notons que cette unique solution est donnée par  $\boldsymbol{\theta} = \mathbf{J}^{-1}\mathbf{C}'(\mathbf{C}\mathbf{J}^{-1}\mathbf{C}')^{-1}\mathbf{a}$ , où  $\mathbf{J} = \mathbf{F} + \mathbf{S} + \mathbf{T}$ . Evidemment, le critère de Temporalité  $T(\boldsymbol{\theta})$  pourrait être remplacé par le critère  $T_0(\boldsymbol{\theta})$  de la formule (9), mais la résolution du problème global 5.1.1 ne pourrait plus être obtenue analytiquement. La question du choix des pondérations optimales  $\alpha, \beta$  et  $\gamma$  et de la matrice des contraintes linéaires à conserver fera l'objet d'un papier ultérieur.

### 5.2 Exemple applicatif

On présente pour finir un exemple applicatif d'utilisation de moyennes mobiles sur une série temporelle construite à partir de données réelles. Sur la figure 4 est représentée la série brute (courbe noire) ainsi que la série lissée (courbe violette) par une moyenne mobile asymétrique de Henderson classique d'ordre 13, pour laquelle  $p = 9$  et  $f = 3$ . Sur la figure 5 est représentée la même série brute (courbe noire) et la série lissée (courbe violette)

cette fois-ci par une moyenne mobile asymétrique d'Henderson (toujours d'ordre 13 avec  $p = 9$  et  $f = 3$ ) minimisant le déphasage. On attirera tout particulièrement l'attention du lecteur sur les zones cerclées où l'utilisation d'une moyenne mobile minimisant le déphasage apporte des gains significatifs, notamment en ce qui concerne les objectifs de Fidélité et de Temporalité, par rapport aux moyennes de Henderson usuelles, pourtant appliquées fréquemment dans les logiciels de désaisonnalisation.

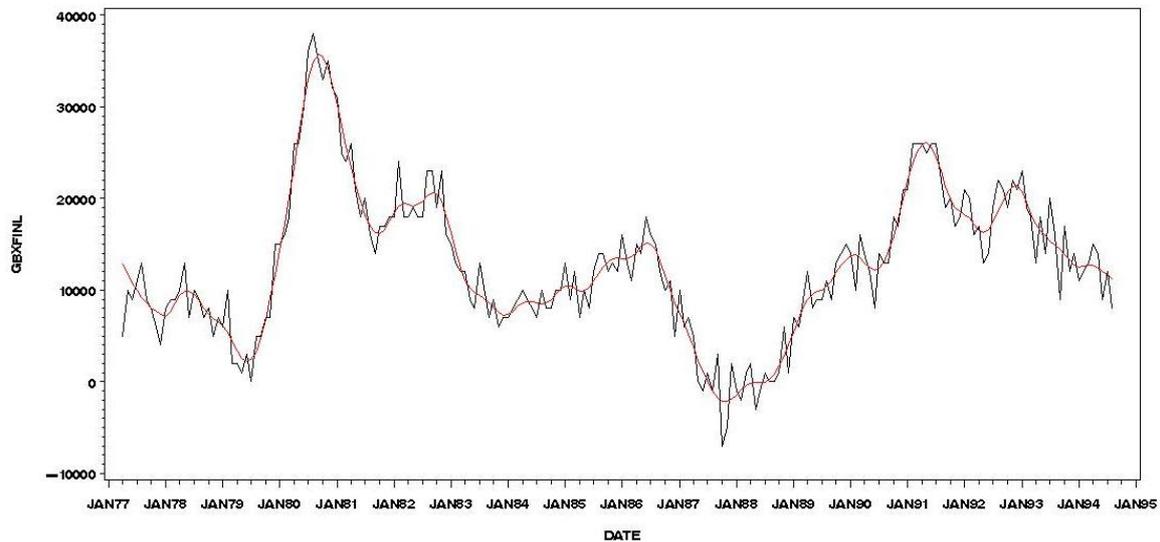


FIGURE 4 – Série brute (en noir) et série lissée par une moyenne de Henderson 9-3 classique (en violet)

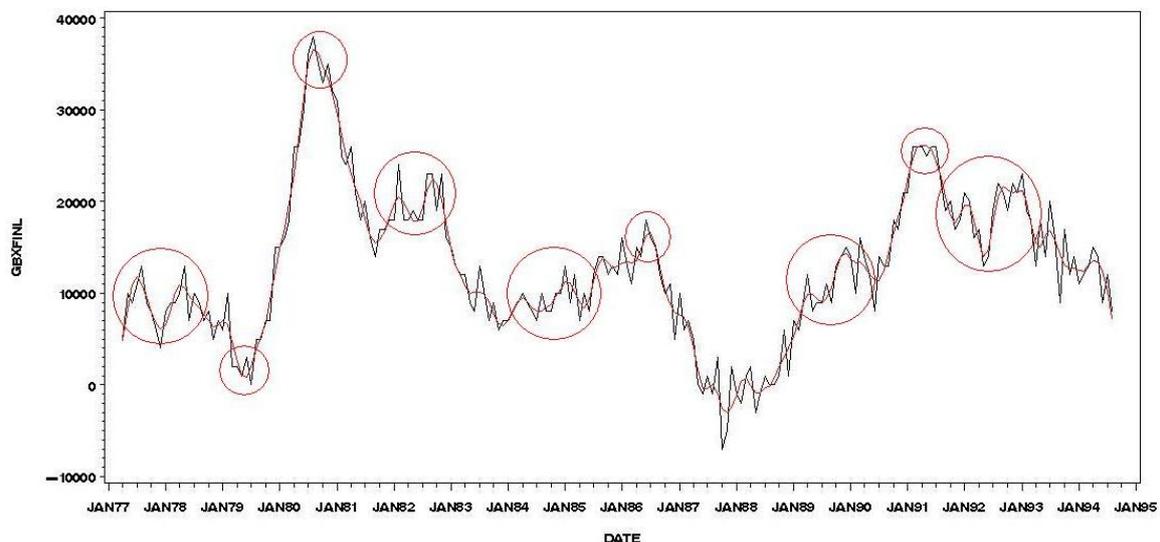


FIGURE 5 – Série brute (en noir) et série lissée par une moyenne de Henderson 9-3 minimisant le déphasage (en violet)

## Références

- [1] Bongard, J. (1962), Quelques remarques sur les moyennes mobiles, *L'élimination des variations saisonnières à l'aide de calculatrices électroniques*, OCDE, 1962, 389-427,

Paris.

- [2] Doherty, M. (2001), Surrogate Henderson Filters in X-11, *Australian and New Zealand Journal of Statistics*, 43, 385-392. (Initially published as a working paper in 1992).
- [3] Gray, A., Thomson, P. (2002), On a Family of Moving Average Trend Filters for the Ends of Series. *Journal of Forecasting*, 21, 125-149. (Initially published as a working paper in 1996.)
- [4] Grun-Rehomme, M., Ladiray, D. (1994), "Moyennes mobiles centrées et non centrées : construction et comparaison", *Revue de Statistique Appliquée*, vol XLII, 3, 33-61, Paris.
- [5] Henderson, R. (1916), Note on Graduation by Adjusted Average, *Transactions of the Actuarial Society of America*, 17, 43-48.
- [6] Henderson, R. (1924), A New Method of Graduation, *Transactions of the Actuarial Society of America*, 25, 29-40.
- [7] Ladiray, D., Quenneville, B. (2001), "*Understanding X-11*", Springer-Verlag, Lecture Notes in Statistics 158, New York.
- [8] Musgrave, J.C. (1964), A Set of End Weights to End All End Weights, Unpublished internal note. Washington : US Bureau of the Census.
- [9] Quenneville, B., Ladiray, D. (2000), Locally Adaptive Trend-Cycle Estimation for X-11, *Proceedings of the International Conference on Establishment Surveys II*, Buffalo, NY, June 2000.
- [10] Quenneville, B., Ladiray, D., Le François, B. (2003), A Note on Musgrave Asymmetrical Trend-Cycle Filters, *International Journal of Forecasting*, 19, 727-734.
- [11] Wildi, M. (1998), Detection of Compatible Turning Points and Signal Extraction for Non-Stationary Time Series, *Operation Research Proceedings*, Springer.
- [12] Wildi, M. (2011), Signal Extraction : Efficient Estimation, 'Unit-Root' Tests and Early Detection of Turning Points, Springer-Verlag, Lecture Notes in Economic and Mathematical Systems, 547, Berlin Heidelberg.
- [13] Wildi, M. (2011), Real Time Trend Extraction and Seasonal Adjustment : a Generalized Direct Filter Approach, *To be published*.