I-DFA and I-MDFA: Companion Paper to R-Code Published on SEFBlog

Marc Wildi
Zurich University of Applied Sciences

I-DFA and I-MDFA: Companion Paper to R-Code Published on SEFBlog

Marc Wildi
Institute of Data Analysis and Process Design

March 18, 2011

Abstract

Keywords. Real-time signal extraction (RTSE), phase, amplitude, multivariate filter, timeliness-accuracy dilemma.

Disclaimer Parts of this paper were obtained by simply pasting (and slightly adapting) text pieces of some other (published or unpublished) papers - all of mine i.e. no plagiarism involved -. The appendix is mostly original text/formulas.

1 Introduction

The following document is an accompanying paper to R-code on I-DFA (univariate) and I-MDFA (multivariate) as published on SEF-Blog[1]. Section 2 briefly motivates the frequency-domain as a natural approach to the filtering problem. Section 3 reviews uni- and multivariate optimization criteria (DFA, I-DFA, I-MDFA). Closed-form solutions of optimal filter parameters are presented in the appendix.

2 Frequency Domain and Filter Effect

Let X_t, $t = 1, ..., T$ be a finite sample of observations and define Y_{T-r}, $r = 0, ..., T-1$ as the output of a filter with real coefficients γ_{kr}:

$$Y_{T-r} = \sum_{k=-r}^{T-r-1} \gamma_{kr} X_{T-r-k}$$

[1]http://blog.zhaw.ch/idp/sefblog: check the category ‘tutorial’ on the left-hand side of SEFBlog to access to code as well as to exercises.

For $r = 0$ a real-time or causal filter is obtained, namely a linear combination of present and past observations. For $r > 0$, Y_{T-r} relies on 'future' observations $X_{t-r+1}, ..., X_T$ (smoothing). In order to derive the important filter effect we assume a particular (complex) input series $X_t := \exp(i\omega t),\ t \in \mathbb{Z}$. The output signal is thus

$$\begin{aligned} Y_{T-r} &= \sum_{k=-r}^{T-r-1} \gamma_{kr} \exp(i\omega(T-r-k)) &(1)\\ &= \exp(i\omega(T-r)) \sum_{k=-r}^{T-r-1} \gamma_{kr} \exp(-i\omega k) &(2)\\ &= \exp(i\omega(T-r))\Gamma_r(\omega) &(3)\end{aligned}$$

The (generally complex) function

$$\Gamma_r(\omega) := \sum_{k=-r}^{T-r-1} \gamma_{kr} \exp(-i\omega k) \tag{4}$$

is called the transfer function of the filter. We can represent the complex number $\Gamma_r(\omega)$ in polar coordinates according to

$$\Gamma_r(\omega) = A_r(\omega) \exp(-i\Phi_r(\omega)) \tag{5}$$

where $A_r(\omega) = |\Gamma_r(\omega)|$ is called the amplitude of the filter and $\Phi_r(\omega)$ is its phase.
We deduce from 1 that $X_t, t \in \mathbb{Z}$ is a periodic eigensignal of the filter with eigenvalue $\Gamma_r(\omega)$. Linearity of the filter implies that real and imaginary parts of X_t are mapped into real and imaginary parts of Y_t and therefore

$$\begin{aligned} \cos(t\omega) &\to A_r(\omega) \left[\cos(t\omega)\cos(-\Phi_r(\omega)) - \sin(t\omega)\sin(-\Phi_r(\omega))\right] \\ &= A_r(\omega)\cos(t\omega - \Phi_r(\omega)) \\ &= A_r(\omega)\cos(\omega(t - \Phi_r(\omega)/\omega)) &(6)\end{aligned}$$

The amplitude function $A_r(\omega)$ can be interpreted as the weight (damping if $A_r(\omega) < 1$, amplification if $A_r(\omega) > 1$) attributed by the filter to a sinusoidal input signal with frequency ω. The function

$$\phi_r(\omega) := \Phi_r(\omega)/\omega \tag{7}$$

can be interpreted as the time shift function of the filter in ω[2]. As we shall see in section 3, real-time signal extraction, i.e. the case $r = 0$, aims at optimal simultaneous amplitude and time shift matchings or 'fits'. Amplitude and time-shift functions describe comprehensively the effect of the filter when applied to a simple trigonometric signal of frequency omega. In order to extend the

[2]The singularity in $\omega = 0$ is resolved by noting that $\Phi(0) = 0$ for filters satisfying $\Gamma_r(0) > 0$. As a result $\phi_r(0) := \dot{\Phi}_r(0)$. For $\Gamma_r(0) = 0$ the phase could be set to any arbitrary value, including zero, of course.

scope of the analysis and to found the validity of our approach we can rely on a well-known result stating that any sequence of numbers X_t, $t = 1, .., T$, sampled on an equidistant time-grid, can be decomposed uniquely into a weighted sum of mutually orthogonal complex exponential terms

$$X_t = \frac{1}{\sqrt{2\pi T}} \sum_{k=-T/2}^{T/2} DFT(\omega_k) \exp(-it\omega_k)$$

where $DFT(\omega_k)$ is the discrete Fourier transform of X_t and $\omega_k = \frac{k2\pi}{T}$, $k = -T/2, ..., 0, ..., T/2$ is a discrete frequency-grid in the interval $[-\pi, \pi]$. Linearity of the filter can then be invoked to extend the description of the filter effect in terms of amplitude and time-shift functions to arbitrary sequences of numbers X_t, $t = 1, ..., T$.

Note that this decomposition is a finite sample version of the fundamental spectral representation theorem and that it is fully compatible with the latter, asymptotically. However, the validity of the discrete finite-sample decomposition extends to any sequence of numbers, including realizations of non-stationary processes.

3 Data-Dependent Filters

We here propose optimization criteria which emphasize optimal properties of asymmetric filters. For this purpose, we assume that a particular signal or, equivalently, a symmetric filter has been defined by the user. The signal could be a trend, a cycle or a seasonally adjusted component and the definition could be either ad hoc or 'model-based'. In order to simplify notations we here emphasize the practically relevant real-time or concurrent filter which approximates the signal at the end $t = T$ of the sample. We propose criteria which emphasize the revision error as well as speed (timeliness) and reliability (noise suppression) issues. Criteria in the first group are able to replicate traditional model-based filters (X-12-ARIMA,TRAMO,Stamp) perfectly. Criteria in the second group are able to account for more complex real-time inferences (for example the detection of turning-points) and for more sophisticated user priorities (for example different levels of risk aversion).

3.1 Mean-Square Error Criterion

Let the target signal be defined by the output of a symmetric (possibly bi-infinite) filter

$$Y_T = \sum_{j=-\infty}^{\infty} \gamma_j X_{T-j} \tag{8}$$

and let $\hat{Y}_t$ denote its real-time estimate

$$\hat{Y}_T = \sum_{j=0}^{T-1} b_j X_{T-j} \tag{9}$$

Furthermore, let $\Gamma(\cdot) = \sum_{j=-\infty}^{\infty} \gamma_j \exp(-ij\cdot)$ and $\hat{\Gamma}(\cdot) = \sum_{j=0}^{T-1} b_j \exp(-ij\cdot)$ denote the corresponding transfer functions. For stationary processes X_t, the mean-square filter error can be expressed as

$$\int_{-\pi}^{\pi} |\Gamma(\omega) - \hat{\Gamma}(\omega)|^2 dH(\omega) = E[(Y_t - \hat{Y}_t)^2] \tag{10}$$

where $H(\omega)$ is the unknown spectral distribution of X_t. Consider now the following finite sample approximation of the above integral

$$\frac{2\pi}{T} \sum_{k=-[T/2]}^{[T/2]} w_k |\Gamma(\omega_k) - \hat{\Gamma}(\omega_k)|^2 S(\omega_k) \tag{11}$$

where $\omega_k = k2\pi/T$, $[T/2]$ is the greatest integer[3] smaller or equal to $T/2$ and the weights w_k are defined by

$$w_k = \begin{cases} 1 & , \quad |k| \neq T/2 \\ 1/2 & , \quad \text{otherwise} \end{cases} \tag{12}$$

In this expression, $S(\omega_k)$ can be interpreted as an estimate of the unknown spectral density of the process. Consistency of this estimate is not necessary because we are not interested in estimating the (unknown) spectral density but the filter mean-square error instead. In this perspective, we may take benefit of the smoothing effect provided by the summation operator in 11. So for example Wildi1998, Wildi2005, Wildi2008 and Wildi2010 propose to plug the periodogram into the above expression:

$$S(\omega_k) := I_{TX}(\omega_k) = \frac{1}{2\pi T} \left| \sum_{t=1}^{T} X_t \exp(-it\omega_k) \right|^2$$

Formal efficiency results applying to the resulting Direct Filter Approach (DFA) are presented in Wildi2008 and 2009. Real-world true out-of-sample performances are extensively documented and discussed in Wildi2008.

In other parts of on-going work we propose to extend the original DFA by considering alternative spectral estimates $S(\omega_k)$ derived from *models* of the Data Generating Process (DGP)[4]. The term 'model' makes reference to explicit representations of the DGP by X-12-ARIMA, TRAMO or STAMP, for example, as well as to 'ad hoc' implicit DGP-assumptions underlying classical filters, such as HP, CF or Henderson, for example. This way, a formal link between the original DFA

[3]In order to simplify the exposition we now assume that T is even. In our applications, the sample length is generally a multiple of 4 (quarterly data) or 12 (monthly data) in order that the important seasonal frequencies can be matched by ω_k.

[4]See: http://blog.zhaw.ch/idp/sefblog/index.php?/archives/165-Real-Time-Signal-Extraction-RTSE-an-Agnostic-Perspective-Plus-a-Frivolity-of-Mine.html

and traditional model-based approaches is established which allows to transpose the powerful customization principle of the latter to the former.

The case of non-stationary integrated processes can be handled by noting that 11 addresses the filter error $Y_t - \hat{Y}_t$, not the data X_t. The former is generally stationary even if the latter isn't[5] and therefore all spectral decomposition results are still valid in a formal mathematical perspective. In the case of integrated processes, stationarity of the filter error is obtained by imposing cointegration between the signal Y_t and the real-time estimate $\hat{Y}_t$. Formally, this amounts to impose suitable real-time filter constraints. A comprehensive treatment of the topic is given in Wildi2008 and 2010.

3.2 Controlling Speed and Reliability: the Classical DFA

Wildi1998, 2005, 2008 and 2010 propose a decomposition of the mean-square filter error into distinct components attributable to the amplitude and the phase functions of the real-time filter. We here briefly review this decomposition and derive customized criteria which emphasize explicitly speed and/or reliability aspects subject to particular user priorities, such as, for example, various degrees of risk aversion.

The following identity holds for general transfer functions Γ and $\hat{\Gamma}$:

$$\begin{aligned} |\Gamma(\omega) - \hat{\Gamma}(\omega)|^2 &= A(\omega)^2 + \hat{A}(\omega)^2 - 2A(\omega)\hat{A}(\omega)\cos\left(\hat{\Phi}(\omega) - \Phi(\omega)\right) \\ &= (A(\omega) - \hat{A}(\omega))^2 \\ &\quad +2A(\omega)\hat{A}(\omega)\left[1 - \cos\left(\hat{\Phi}(\omega) - \Phi(\omega)\right)\right] \end{aligned} \tag{13}$$

If we assume that Γ is symmetric and positive, then $\Phi(\omega) \equiv 0$. Inserting 13 into 11 and using $1 - \cos(\hat{\Phi}(\omega)) = 2\sin(\hat{\Phi}(\omega)/2)^2$ then leads to

$$\frac{2\pi}{T}\sum_{k=-T/2}^{T/2} w_k(A(\omega_k) - \hat{A}(\omega_k))^2 S(\omega_k) \tag{14}$$

$$+\frac{2\pi}{T}\sum_{k=-T/2}^{T/2} w_k 4A(\omega_k)\hat{A}(\omega_k)\sin(\hat{\Phi}(\omega)/2)^2 S(\omega_k) \tag{15}$$

The first summand 14 is the distinctive part of the total mean-square filter error which is attributable to the amplitude function of the real-time filter (the MS-amplitude error). The second summand 15 measures the distinctive contribution of the phase or time-shift to the total mean-square error (the MS-time-shift error). The term $A(\omega_k)\hat{A}(\omega_k)$ in 15 is a scaling factor which accounts for the fact that the phase function does not convey level information.

[5] Optimal real-time signalextraction aims precisely at a smallest possible mean-square filter error i.e. the filter error is neither trending nor unbounded in variance.

Now consider the following generalized version of the original mean-square criterion[6]:

$$
\begin{aligned}
&\frac{2\pi}{T}\sum_{k=-T/2}^{T/2} w_k(A(\omega_k)-\hat{A}(\omega_k))^2 W(\omega_k)S(\omega_k)\\
&+(1+\lambda)\frac{2\pi}{T}\sum_{k=-T/2}^{T/2} w_k 4A(\omega_k)\hat{A}(\omega_k)\sin(\hat{\Phi}(\omega_k)/2)^2 W(\omega_k)S(\omega_k)\\
=\;&\frac{2\pi}{T}\sum_{k=-[T/2]}^{[T/2]} w_k|\Gamma(\omega_k)-\hat{\Gamma}(\omega_k)|^2 W(\omega_k)S(\omega_k)\\
&+4\lambda\frac{2\pi}{T}\sum_{k=-T/2}^{T/2} w_k A(\omega_k)\hat{A}(\omega_k)\sin(\hat{\Phi}(\omega_k)/2)^2 W(\omega_k)S(\omega_k) \to \min
\end{aligned}
\tag{16}
$$

where $W(\cdot) := W(\omega_k, expweight, cutoff)$ is a two-parametric family of weighting functions

$$
W(\omega_k, expweight, cutoff) = \left\{ \begin{array}{c} 1\ ,\ \text{if } |\omega_k| < cutoff \\ (1+|\omega_k|-cutoff)^{expweight}\ ,\ \text{otherwise} \end{array} \right. \tag{17}
$$

The parameter $cutoff$ marks the transition between pass- and stop-bands ($cutoff = \pi/7$ in the empirical applications below) and positive values of the parameter *expweight* emphasize high-frequency components. Classical mean-square optimization is obtained for $\lambda = expweight = 0$: the revision error is addressed. For $\lambda > 0$ the user can emphasize the contribution of the MS-time-shift error. As a result, corresponding real-time filters (typically low-pass trend or cycle extraction) will convey less delayed signals: turning-points can be detected earlier. Note that the weighting $A(\omega_k)\hat{A}(\omega_k)$ in this expression implies that λ acts on the *pass-band* frequencies exclusively and that *expweight* does not alter the time-shift error. The latter parameter emphasizes the MS-amplitude error by magnifying 'noisy' high-frequency components in the *stop-band*. As a result, 'noise' is suppressed more effectively and the reliability of real-time estimates will improve accordingly.

It is generally admitted that reliability and timeliness (speed of detection) of real-time estimates are to some extent mutually exclusive requirements. It is not our intention to contradict this fundamental uncertainty principle, of course, but it seems obvious that the user can attempt to improve performances in both dimensions simultaneously by increasing λ as well as *expweight*. By doing so, control on the amplitude function is lost in the pass-band: mean-square performances are affected adversely. Risk-averse users can operationalize their preference by matching *expweight* to their specific needs. On the other hand, fast 'early' estimates are obtained by prioritizing the phase error over the amplitude error.

In the following section we review the I-DFA criterion.

[6]For notational simplicity it is assumed that $\Gamma(\omega) > 0$ for all ω such that $\Gamma(\omega) = A(\omega)$.

3.3 I-DFA

The mean-square error criterion is a quadratic function of the filter parameters and therefore the solution can be obtained analytically. The expression 16, however, is more tricky when $\lambda > 0$ because it involves non-linear functions of the filter parameters. Therefore, we here propose a new criterion which opens the way to an analytical approximation of 16 (for notational ease the additional weight w_k 12 has been dropped from all subsequent expressions). Consider the following expression:

$$\frac{2\pi}{T}\sum_{k=-[T/2]}^{[T/2]}\left|\Gamma(\omega_k)-\left\{\Re\left(\hat{\Gamma}(\omega_k)\right)+i\sqrt{1+4\lambda\Gamma(\omega_k)f(\omega_k)}\Im\left(\hat{\Gamma}(\omega_k)\right)\right\}\right|^2 W(\omega_k)S(\omega_k)\to\min \quad (18)$$

where $\Re(\cdot)$ and $\Im(\cdot)$ denote real and imaginary parts and $i^2=-1$ is the imaginary unit. We here assume throughout that $f(\omega_k)=\mathrm{Id}$ is an identity and call the resulting optimization criterion I-DFA[7]. Obviously, the above expression is quadratic in the filter coefficients. In analogy to 16, the weighting function $W(\omega_k)$ emphasizes the fit in the stop band. The term $\lambda\Gamma(\omega_k)$ emphasizes the imaginary part of the real-time filter in the pass band: for $\lambda > 0$ the imaginary part is artificially inflated and therefore the phase is affected. The following development allows for a direct comparison of 16 and 18:

$$\begin{aligned}
&\frac{2\pi}{T}\sum_{k=-[T/2]}^{[T/2]}\left|\Gamma(\omega_k)-\left(\Re\left(\hat{\Gamma}(\omega_k)\right)+i\sqrt{1+4\lambda\Gamma(\omega_k)}\Im\left(\hat{\Gamma}(\omega_k)\right)\right)\right|^2 W(\omega_k)S(\omega_k) \qquad (19)\\
=\;&\frac{2\pi}{T}\sum_{k=-[T/2]}^{[T/2]}\left\{\left(\Gamma(\omega_k)-\Re\left(\hat{\Gamma}(\omega_k)\right)\right)^2+\Im\left(\hat{\Gamma}(\omega_k)\right)^2\right\}W(\omega_k)S(\omega_k)\\
&+4\lambda\frac{2\pi}{T}\sum_{k=-[T/2]}^{[T/2]}\Gamma(\omega_k)\Im\left(\hat{\Gamma}(\omega_k)\right)^2 W(\omega_k)S(\omega_k)\\
=\;&\frac{2\pi}{T}\sum_{k=-[T/2]}^{[T/2]}|\Gamma(\omega_k)-\hat{\Gamma}(\omega_k)|^2W(\omega_k)S(\omega_k)\\
&+4\lambda\frac{2\pi}{T}\sum_{k=-[T/2]}^{[T/2]}A(\omega_k)\hat{A}(\omega_k)^2\sin(\hat{\Phi}(\omega_k))^2W(\omega_k)S(\omega_k) \qquad (20)
\end{aligned}$$

A direct comparison of 16 and 20 reveals that $\hat{\Phi}(\omega_k)/2$ is replaced by $\hat{\Phi}(\omega_k)$ and a supernumerary weighting-term $\hat{A}(\omega_k)$ appears in the latter expression. Expression 20 can be solved analytically for arbitrary λ and/or weighting functions $W(\omega_k)$ because $\hat{A}(\omega_k)^2\sin(\hat{\Phi}(\omega_k))^2$ is simply the squared imaginary part of the real-time filter. For $\lambda=0$ the original (DFA) mean-square criterion 11 is

[7] Tweaking of $f(\omega_k)$ will be treated in a separate paper.

obtained. Overemphasizing the imaginary part of the real-time filter in the pass-band by augmenting $\lambda > 0$ results in filters with smaller phase (time-shifts). It should be noted, however, that the analytic I-DFA criterion 18/20 is 'less effective' in controlling the time-shift than the original DFA-criterion 16 (this is where the function $f(\omega_k)$ in 18 comes into play...). Published I-DFA code on SEFBlog relies on 18. A derivation of the analytic solution is provided in the appendix.

3.4 I-MDFA: Mean-Square

The above (univariate) DFA has been generalized to a general multivariate framework (MDFA) in Wildi (2008.2). Specifically, theorem 7.1 proposes optimization criteria by highlighting the cointegration rank, ranging from full-rank (stationary case) to zero-rank. We here briefly summarize the main results: for ease of exposition we restrict the discussion to the stationary case. Let Y_t be defined by 8 and assume the existence of m additional explaining variables W_{jt}, $j = 1, ..., m$ enrichening the information universe. We here rewrite 11 by adopting the traditional DFA-framework based on the periodogram ($S(\omega_k) := I_{TX}(\omega_k)$) and the DFT $\Xi_{TX}(\omega_k)$:

$$\frac{2\pi}{T} \sum_{k=-[T/2]}^{[T/2]} |\Gamma(\omega_k) - \hat{\Gamma}(\omega_k)|^2 I_{TX}(\omega_k) = \frac{2\pi}{T} \sum_{k=-[T/2]}^{[T/2]} |\Gamma(\omega_k)\Xi_{TX}(\omega_k) - \hat{\Gamma}(\omega_k)\Xi_{TX}(\omega_k)|^2 \tag{21}$$

Consider the following generalization of the univariate real-time filter expression:

$$\hat{\Gamma}_X(\omega_k)\Xi_{TX}(\omega_k) + \sum_{n=1}^{m} \hat{\Gamma}_{W_n}(\omega_k)\Xi_{TW_n}(\omega_k) \tag{22}$$

where

$$\begin{aligned} \hat{\Gamma}_X(\omega_k) &= \left(\sum_{j=0}^{L} b_{Xj} \exp(-ij\omega_k)\right) \Xi_{TX}(\omega_k) \\ \hat{\Gamma}_{W_n}(\omega_k) &= \left(\sum_{j=0}^{L} b_{w_n j} \exp(-ij\omega_k)\right) \Xi_{TW_n}(\omega_k) \end{aligned}$$

are the (one-sided) transfer functions applying to the 'explaining' variables and $\Xi_{TX}(\omega_k)$, $\Xi_{TW_n}(\omega_k)$ are the corresponding DFT's. Theorem 7.1 in Wildi (2008.2) shows that the following straightforward extension of 21

$$\frac{2\pi}{T} \sum_{k=-T/2}^{T/2} \left| \left(\Gamma(\omega_k) - \hat{\Gamma}_X(\omega_k)\right) \Xi_{TX}(\omega_k) - \sum_{n=1}^{m} \hat{\Gamma}_{W_n}(\omega_k)\Xi_{TW_n}(\omega_k) \right|^2 \to \min_{\mathbf{B}} \tag{23}$$

inherits all efficiency properties of the (univariate) DFA and therefore the whole customization principle can be carried over to a general multivariate framework ($\mathbf{B}$ denotes the matrix of unknown filter parameters).

3.5 I-MDFA: Customization

A generalization of the customized criterion 16 to the multivariate case can be obtained by a simple transformation applying to 23:

$$\begin{aligned}
&\frac{2\pi}{T}\sum_{k=-T/2}^{T/2}\left|\Gamma(\omega_k)\Xi_{TX}(\omega_k)-\hat{\Gamma}_X(\omega_k)\Xi_{TX}(\omega_k)-\sum_{n=1}^{m}\hat{\Gamma}_{W_n}(\omega_k)\Xi_{TW_n}(\omega_k)\right|^2\\
=&\frac{2\pi}{T}\sum_{k=-T/2}^{T/2}\left|\Gamma(\omega_k)-\hat{\Gamma}_X(\omega_k)-\sum_{n=1}^{m}\hat{\Gamma}_{W_n}(\omega_k)\frac{\Xi_{TW_n}(\omega_k)}{\Xi_{TX}(\omega_k)}\right|^2|\Xi_{TX}(\omega_k)|^2\\
=&\frac{2\pi}{T}\sum_{k=-T/2}^{T/2}\left|\Gamma(\omega_k)-\tilde{\Gamma}(\omega_k)\right|^2|\Xi_{TX}(\omega_k)|^2 \qquad (24)
\end{aligned}$$

where

$$\tilde{\Gamma}(\omega_k):=\hat{\Gamma}_X(\omega_k)+\sum_{n=1}^{m}\hat{\Gamma}_{W_n}(\omega_k)\frac{\Xi_{TW_n}(\omega_k)}{\Xi_{TX}(\omega_k)} \qquad (25)$$

Expression 24 'looks like' 11 and therefore the same customization can be applied, in principle, as in the latter expression (introducing λ and *expweight*). Specifically, we here rely on the analytically tractable customized criterion 20

$$\begin{aligned}
&\frac{2\pi}{T}\sum_{k=-[T/2]}^{[T/2]}|\Gamma(\omega_k)-\tilde{\Gamma}(\omega_k)|^2W(\omega_k)\,|\Xi_{TX}(\omega_k)|^2\\
&+4\lambda\frac{2\pi}{T}\sum_{k=-T/2}^{T/2}A(\omega_k)\tilde{A}(\omega_k)^2\sin(\tilde{\Phi}(\omega_k))^2W(\omega_k)\,|\Xi_{TX}(\omega_k)|^2\rightarrow\min \qquad (26)
\end{aligned}$$

Note that potential singularities introduced by small values of $\Xi_{TX}(\omega_k)$ in the denominator of 25 could be ignored because they are canceled by the outer-product with $|\Xi_{TX}(\omega_k)|^2$. However, numerical routines don't like this kind of canceling. A numerically robust alternative, implemented in the published R-code, is presented in the appendix. Note also that 24 in combination with 20, as shown in 26, is a 'clever' solution in the sense that one does not need to define series specific $\lambda's$ or $expweight's$. The idea goes as follows:

- We are not interested in controlling for time-shifts or smoothness of series specific filter outputs of the multivariate filter.
- Instead, we are interested in having a timely (fast) and accurate (smooth) **aggregate**: indeed, the output of the multivariate filter is obtained by aggregating cross-sectionally the individual series' outputs.
- The parameters λ and *expweight* acting on $\tilde{\Gamma}(\omega_k)$, as defined by 25 (and plugged into 26), determine properties of the aggregate output, as desired, and therefore the resulting criterion 26 matches our intention.

Appendix

I here briefly derive analytic formulas for the parameters of the real-time MA-filters based on I-DFA and I-MDFA. These formulas are used in my R-code. **Warning: the published R-code is deliberately ineffective, in computational terms, because I wanted to follow as closely as possible the explicit formula expressions derived here!**[8].

I-DFA

The following calculations apply to the univariate filter criterion 18 (for simplicity of exposition the constant multiplication term 4 has been concatenated into λ):

$$\begin{aligned}
&\sum_k |\Gamma(\omega_k) - Re(\hat{\Gamma}(\omega_k)) - i * \sqrt{1+\lambda\Gamma(\omega_k)} Im(\hat{\Gamma}(\omega_k))|^2 I_{TX}(\omega_k) \\
= &\sum_k \left(\left[\Gamma(\omega_k) - Re(\hat{\Gamma}(\omega_k))\right]^2 + (1+\lambda\Gamma(\omega_k)) Im\left(\hat{\Gamma}(\omega_k)\right)^2 \right) I_{TX}(\omega_k)
\end{aligned}$$

We now differentiate this expression with respect to filter parameters:

$$\sum_k \left((\Gamma(\omega_k) - Re(\hat{\Gamma}(\omega_k)))(-d/db_j(Re(\hat{\Gamma}(\omega_k)))) + (1+\lambda\Gamma(\omega_k)) Im(\hat{\Gamma}(\omega_k)) d/db_j(Im(\hat{\Gamma}(\omega_k)))) \right) I_{TX}(\omega_k) = 0$$

Now $-d/db_j(Re(\hat{\Gamma}(\omega_k))) = -\cos(j\omega_k)$ and $d/db_j(Im(\hat{\Gamma}(\omega_k))) = \sin(j\omega_k)$[9]. Therefore we obtain

$$\sum_k \left((\Gamma(\omega_k) - Re(\hat{\Gamma}(\omega_k)))(-\cos(j\omega_k)) + (1+\lambda\Gamma(\omega_k)) Im(\hat{\Gamma}(\omega_k)) \sin(j\omega_k)) \right) I_{TX}(\omega_k) = 0$$

Or

$$\sum_k (\Gamma(\omega_k)\cos(j\omega_k)) I_{TX}(\omega_k) = \sum_k \left(Re(\hat{\Gamma}(\omega_k))\cos(j\omega_k) + (1+\lambda\Gamma(\omega_k)) Im(\hat{\Gamma}(\omega_k)) \sin(j\omega_k)) \right) I_{TX}(\omega_k) \quad (27)$$

[8] The usual trade-off applies: the logic of numerically more effective formulations are generally less straightforward to understand (from an abstract formal point of view).

[9] Please note that I inverted the sign of the complex exponential functions i.e. one should read $\cos(-j\omega_k)$ instead of $\cos(j\omega_k)$ (which does not change anything...) and $\sin(-j\omega_k)$ instead of $\sin(j\omega_k)$. This is because my code was implemented accordingly. Obviously, this arbitrary change of sign is completely irrelevant to the derivation of parameters. The only 'relevant' modification concerns the phase of the real-time filter whose sign must be inverted in order to allow for a meaningful interpretation of filter diagnostics.

Now $Re(\hat{\Gamma}(\omega_k)) = \sum_l b_l \cos(l\omega_k)$ and $Im(\hat{\Gamma}(\omega_k)) = \sum_l b_l \sin(l\omega_k)$. Therefore we obtain the following set of equations on the right-hand side of 27:

$$\begin{aligned}
&b_0 \sum_k (\cos(j\omega_k)\cos(0\omega_k) + (1+\lambda\Gamma(\omega_k))\sin(j\omega_k)\sin(0\omega_k)) I_{TX}(\omega_k) + \\
&b_1 \sum_k (\cos(j\omega_k)\cos(1\omega_k) + (1+\lambda\Gamma(\omega_k))\sin(j\omega_k)\sin(1\omega_k)) I_{TX}(\omega_k) + \\
&... + \\
&b_L \sum_k (\cos(j\omega_k)\cos(L\omega_k) + (1+\lambda\Gamma(\omega_k))\sin(j\omega_k)\sin(L\omega_k)) I_{TX}(\omega_k)
\end{aligned}$$

Let $\mathbf{b} = \mathbf{X'X}^{-1}\mathbf{X'Y}$. Then the right-hand side of 27 implies

$$\mathbf{X'X} = \left(\sum_k (\cos(j\omega_k)\cos(m\omega_k) + (1+\lambda\Gamma(\omega_k))\sin(j\omega_k)\sin(m\omega_k)) I_{TX}(\omega_k) \right)_{jm} \tag{28}$$

where both indices $0 \leq j, m \leq L$. Finally, the left-hand side of 27 implies that

$$\mathbf{X'Y} = \left(\sum_k (\Gamma(\omega_k)\cos(j\omega_k)) I_{TX}(\omega_k) \right)_j \tag{29}$$

where $j = 0, ..., L$. These formulas are used in my R-code.

I-MDFA

The criterion is:

$$\sum_k \left| \Gamma(\omega_k) - \tilde{\Gamma}(\omega_k) \right|^2 |\Xi_{TX}(\omega_k)|^2 \to \min_{\mathbf{B}} \tag{30}$$

where

$$\tilde{\Gamma}(\omega_k) := \hat{\Gamma}_X(\omega_k) + \sum_{n=1}^{m} \hat{\Gamma}_{W_n}(\omega_k) \frac{\Xi_{TW_n}(\omega_k)}{\Xi_{TX}(\omega_k)} \tag{31}$$

Instead of the numerically potentially unstable ratio of DFT's appearing in 31, we here re-write the criterion such that the potential singularity is avoided (the R-code relies on this re-formulated expression). Let us first consider the following multivariate generalization of 18 (for notational simplicity the constant multiplier 4 is concatenated into λ and the weighting function $W(\omega_k)$ has been ignored):

$$\begin{aligned}
\sum_k \Bigg| &\Gamma(\omega)\Xi_{TX}(\omega_k) - \Re\left(\hat{\Gamma}_X(\omega_k)\Xi_{TX}(\omega_k) + \sum_{n=1}^{m} \hat{\Gamma}_{W_n}(\omega_k)\Xi_{TW_n}(\omega_k) \right) \\
&-i * \sqrt{1+\lambda\Gamma(\omega_k)} \left\{ \Im\left(\hat{\Gamma}_X(\omega_k)\Xi_{TX}(\omega_k) + \sum_{n=1}^{m} \hat{\Gamma}_{W_n}(\omega_k)\Xi_{TW_n}(\omega_k) \right) \right\} \Bigg|^2 \to \min
\end{aligned}$$

where i is the imaginary unit. Although this is not the right way to proceed we here first follow this line of attack (the necessary modification is provided below). One obtains:

$$\sum_k \left|\Gamma(\omega)\Xi_{TX}(\omega_k) - \Re\left(\hat{\Gamma}_X(\omega_k)\Xi_{TX}(\omega_k) + \sum_{n=1}^m \hat{\Gamma}_{W_n}(\omega_k)\Xi_{TW_n}(\omega_k)\right)\right. \tag{32}$$

$$\left. -i * \sqrt{1+\lambda\Gamma(\omega_k)}\left\{\Im\left(\hat{\Gamma}_X(\omega_k)\Xi_{TX}(\omega_k) + \sum_{n=1}^m \hat{\Gamma}_{W_n}(\omega_k)\Xi_{TW_n}(\omega_k)\right)\right\}\right|^2 \tag{33}$$

$$= \sum_k \left(\Re()^2 + \Im()^2\right)$$

$$= \sum_k \left(\Gamma(\omega)\Re\left(\Xi_{TX}(\omega_k)\right) - \Re\left(\hat{\Gamma}_X(\omega_k)\Xi_{TX}(\omega_k)\right) - \sum_n \Re\left(\hat{\Gamma}_{W_n}(\omega_k)\Xi_{TW_n}(\omega_k)\right)\right)^2$$

$$+\sum_k \left(\Gamma(\omega)\Im\left(\Xi_{TX}(\omega_k)\right) - \sqrt{1+\lambda\Gamma(\omega_k)}\left\{\Im\left(\hat{\Gamma}_X(\omega_k)\Xi_{TX}(\omega_k)\right) + \sum_{n=1}^m \Im\left(\hat{\Gamma}_{W_n}(\omega_k)\Xi_{TW_n}(\omega_k)\right)\right\}\right)^2 \tag{34}$$

We can recognize/identify two problems related to this expression:

- A nuisance term $\Gamma(\omega)\Im\left(\Xi_{TX}(\omega_k)\right)$ appears in the imaginary part 34.
- Requiring a smaller imaginary part (reduced phase) of the aggregate filter

  $$\Im\left(\hat{\Gamma}_X(\omega_k)\Xi_{TX}(\omega_k) + \sum_{n=1}^m \hat{\Gamma}_{W_n}(\omega_k)\Xi_{TW_n}(\omega_k)\right)$$

  by augmenting λ in 33 would not necessarily lead to the expected improvement because the target signal $\Gamma(\omega)\Xi_{TX}(\omega_k)$ in 32 is a complex number with a non-vanishing imaginary part too.

Both problems could be avoided in 30, by isolating $\Xi_{TX}(\omega_k)$ outside of the filter expression. In doing this, we note that we don't need to 'isolate' the whole DFT: its argument would be sufficient. So let's have a look at the following modified expression

$$\sum_k \left|\Gamma(\omega)\left|\Xi_{TX}(\omega_k)\right| - \Re\left(\hat{\Gamma}_X(\omega_k)\left|\Xi_{TX}(\omega_k)\right| + \sum_{n=1}^m \hat{\Gamma}_{W_n}(\omega_k)\Xi_{TW_n}(\omega_k)\exp\left(-i\arg\left(\Xi_{TX}(\omega_k)\right)\right)\right)\right.$$

$$\left. -i * \sqrt{1+\lambda\Gamma(\omega_k)}\left\{\Im\left(\hat{\Gamma}_X(\omega_k)\left|\Xi_{TX}(\omega_k)\right| + \sum_{n=1}^m \hat{\Gamma}_{W_n}(\omega_k)\Xi_{TW_n}(\omega_k)\exp\left(-i\arg\left(\Xi_{TX}(\omega_k)\right)\right)\right)\right\}\right|^2$$

$$\left|\exp\left(i\arg\left(\Xi_{TX}(\omega_k)\right)\right)\right|^2$$

where we isolate $\exp\left(i\arg\left(\Xi_{TX}(\omega_k)\right)\right)$ 'only'. Since $\left|\exp\left(i\arg\left(\Xi_{TX}(\omega_k)\right)\right)\right|^2 = 1$ we can simplify the above expression to obtain:

$$\sum_k \left|\Gamma(\omega)\left|\Xi_{TX}(\omega_k)\right| - \Re\left(\hat{\Gamma}_X(\omega_k)\left|\Xi_{TX}(\omega_k)\right| + \sum_{n=1}^m \hat{\Gamma}_{W_n}(\omega_k)\Xi_{TW_n}(\omega_k)\exp\left(-i\arg\left(\Xi_{TX}(\omega_k)\right)\right)\right)\right.$$

$$\left. -i * \sqrt{1+\lambda\Gamma(\omega_k)}\left\{\Im\left(\hat{\Gamma}_X(\omega_k)\left|\Xi_{TX}(\omega_k)\right| + \sum_{n=1}^m \hat{\Gamma}_{W_n}(\omega_k)\Xi_{TW_n}(\omega_k)\exp\left(-i\arg\left(\Xi_{TX}(\omega_k)\right)\right)\right)\right\}\right|^2 \tag{35}$$

As can be seen, all previous problems are solved: the nuisance term has vanished in the imaginary part; imposing a smaller imaginary part of the aggregate multivariate filter (by augmenting λ) would meet our target signal $\Gamma(\omega)\,|\Xi_{TX}(\omega_k)|$ which is now *real*; finally, this expression and the resulting criterion are stable numerically. In order to simplify notations let us denote the rotated DFT's in 35 by:

$$\begin{aligned}\tilde{\Xi}_{TX}(\omega_k) &= |\Xi_{TX}(\omega_k)|\\ \tilde{\Xi}_{TW_n}(\omega_k) &= \Xi_{TW_n}(\omega_k)\exp\left(-i\arg\left(\Xi_{TX}(\omega_k)\right)\right)\end{aligned}$$

We can now proceed to the formal solution by differentiating expression 35 with respect to b_j^m (the $j-$th MA-coefficient of the filter applied to W_{mt}) and equating to zero:

$$\begin{aligned}&\sum_k\left(\Gamma(\omega)\tilde{\Xi}_{TX}(\omega_k)-\Re\left(\hat{\Gamma}_X(\omega_k)\tilde{\Xi}_{TX}(\omega_k)\right)-\sum_{n=1}^{m}\Re\left(\hat{\Gamma}_{W_n}(\omega_k)\tilde{\Xi}_{TW_n}(\omega_k)\right)\right)(-1)\Re\left(\exp(ij\omega_k)\tilde{\Xi}_{TW_m}(\omega_k)\right)\\ &-\sum_k\left(\sqrt{1+\lambda\Gamma(\omega_k)}\left\{\Im\left(\hat{\Gamma}_X(\omega_k)\tilde{\Xi}_{TX}(\omega_k)\right)+\sum_{n=1}^{m}\Im\left(\hat{\Gamma}_{W_n}(\omega_k)\tilde{\Xi}_{TW_n}(\omega_k)\right)\right\}\right)(-1)\Im\left(\exp(ij\omega_k)\tilde{\Xi}_{TW_m}(\omega_k)\right)\\ =\;& 0\end{aligned}$$

where we assumed $\tilde{\Xi}_{TW_m}(\omega_k)=\tilde{\Xi}_{TX}(\omega_k)$ if $m=0$. One then obtains

$$\begin{aligned}&\sum_k\Gamma(\omega)\tilde{\Xi}_{TX}(\omega_k)\Re\left(\exp(ij\omega_k)\tilde{\Xi}_{TW_m}(\omega_k)\right)\\ =\;&\sum_k\left\{\Re\left(\hat{\Gamma}_X(\omega_k)\tilde{\Xi}_{TX}(\omega_k)\right)+\sum_{n=1}^{m}\Re\left(\hat{\Gamma}_{W_n}(\omega_k)\tilde{\Xi}_{TW_n}(\omega_k)\right)\right\}\Re\left(\exp(ij\omega_k)\tilde{\Xi}_{TW_m}(\omega_k)\right)\\ &+\sum_k(1+\lambda\Gamma(\omega_k))\left\{\Im\left(\hat{\Gamma}_X(\omega_k)\tilde{\Xi}_{TX}(\omega_k)\right)+\sum_{n=1}^{m}\Im\left(\hat{\Gamma}_{W_n}(\omega_k)\tilde{\Xi}_{TW_n}(\omega_k)\right)\right\}\Im\left(\exp(ij\omega_k)\tilde{\Xi}_{TW_m}(\omega_k)\right)\end{aligned}\qquad(36)$$

Let me show that this expression reduces to the classical univariate mean-square DFA-criterion when $m=0$ and $\lambda=0$. The left-hand side becomes

$$\Gamma(\omega)\tilde{\Xi}_{TX}(\omega_k)\Re\left(\exp(ij\omega_k)\tilde{\Xi}_{TW_m}(\omega_k)\right)=\Gamma(\omega)\cos(j\omega_k)I_{TX}(\omega_k)$$

which corresponds to 29. The right-hand side simplifies to

$$\begin{aligned}&\Re\left(\hat{\Gamma}_X(\omega_k)\tilde{\Xi}_{TX}(\omega_k)\right)\Re\left(\exp(ij\omega_k)\tilde{\Xi}_{TX}(\omega_k)\right)+\Im\left(\hat{\Gamma}_X(\omega_k)\tilde{\Xi}_{TX}(\omega_k)\right)\Im\left(\exp(ij\omega_k)\tilde{\Xi}_{TX}(\omega_k)\right)\\ =\;&\Re\left(\hat{\Gamma}_X(\omega_k)\tilde{\Xi}_{TX}(\omega_k)\right)\Re\left(\overline{\exp(ij\omega_k)\tilde{\Xi}_{TX}(\omega_k)}\right)-\Im\left(\hat{\Gamma}_X(\omega_k)\tilde{\Xi}_{TX}(\omega_k)\right)\Im\left(\overline{\exp(ij\omega_k)\tilde{\Xi}_{TX}(\omega_k)}\right)\\ =\;&\Re\left(\hat{\Gamma}_X(\omega_k)\tilde{\Xi}_{TX}(\omega_k)\overline{\tilde{\Xi}_{TX}(\omega_k)\exp(ij\omega_k)}\right)\\ =\;&\Re\left(\hat{\Gamma}_X(\omega_k)\overline{\exp(ij\omega_k)}\right)I_{TX}(\omega_k)\end{aligned}$$

which corresponds to the data-matrix 28.

The right-hand side of equation 36 (the criterion differentiated with respect to b_j^m) attributes the following weight to the filter coefficient b_l^u:

$$\Re\Bigg(\exp(il\omega_k)\Xi_{TW_u}(\omega_k)\Bigg)\Re\Bigg(\exp(ij\omega_k)\Xi_{TW_m}(\omega_k)\Bigg)$$
$$+(1+\lambda\Gamma(\omega_k))\Im\Bigg(\exp(il\omega_k)\Xi_{TW_u}(\omega_k)\Bigg)\Im\Bigg(\exp(ij\omega_k)\Xi_{TW_m}(\omega_k)\Bigg)$$

where, once again, we assume that $\Xi_{TW_0}(\omega_k) = \Xi_{TX}(\omega_k)$ for $u = 0$. This generalized $\mathbf{X}'\mathbf{X}$-matrix reduces to 28 in the univariate case. This expression is used in the published R-code.

Filter Constraints

Let us briefly address the case of imposing first and second order filter restrictions. The first order restriction is

$$b_1^n + b_2^n + ... + b_L^n = w^n$$

Specifically: the restriction imposes a *level constraint* according to $\hat{\Gamma}_{W_n}(0) = w^n$ (we again assume that $\hat{\Gamma}_{W_0} = \hat{\Gamma}_X$ if $m = 0$). The second order restriction imposes vanishing time-shifts in frequency zero. For this purpose the derivative of the transfer function in frequency zero must vanish: $\frac{\partial}{\partial\omega}\Big|_{\omega=0}\sum_{j=1}^{L} b_j^n \exp(i(j-1)\omega) = 0$. This condition results in the following coefficient constraint:

$$b_2^n + 2b_3^n + 3b_4^n + ... + (L-1)b_L^n = 0$$

Imposing both constraints simultaneously leads to:

$$b_{L-1}^n = -(L-1)b_1^n - (L-2)b_2^n - ... - 2b_{L-2}^n + (L-1)w^n$$

$$b_L^n = (L-2)b_1^n + (L-3)b_2^n + (L-4)b_3^n + ... + b_{L-2}^n - (L-2)w^n$$

These restrictions are imposed by specifying i1$< -$T and/or i2$< -$T in the published R-code.

Cointegration and Varia

VERY PROTOTYPICAL!

Signal is bi-infinite univariate filter applied to X_t: if bi-infinite data is available, then a univariate framework is all we need. In a RT-perspective, however, redundancy conveyed in additional data may be helpful. Interestingly, the additional data could be completely uninformative about the interesting component! Let's illustrate the topic:

$$x_t = rw_t + c_t; y_t = rw_t$$

if one would like to extract the cycle c_t in x_t then y_t can be helpful although it dies not convey information about c_t. so for example $x_t - y_t = c_t$ i.e. estimation is perfect!

Need a more general concept than cointegration when interested in turning-points. Series can convey identical information about interesting turning-points (of cycle for example) without being co-integrated. Ordinary level-filter restrictions are not necessarily helpful in such a case.

References

[1] Wildi, M. (1998) Detection of Compatible Turning Points and Signal Extraction for Non-Stationary Time Series. *Operation Research Proceedings, Springer.*

[2] Wildi, M. (2004) Signal Extraction: How (In)efficient Are Model-Based Approaches? An Empirical Study Based on TRAMO/SEATS and Census X-12-ARIMA. *KOF-Working Paper Nr. 96*, ETH-Zurich.

[3] Wildi, M. (2005) Signal Extraction: Efficient Estimation, 'Unit-Root'-Tests and Early Detection of Turning-Points. *Lecture Notes in Economic and Mathematical Systems, 547*, Springer-Verlag Berlin Heidelberg.

[4] Wildi, M. (2008) Real-Time Signal-Extraction: Beyond Maximum Likelihood Principles. http://www.idp.zhaw.ch/fileadmin/user_upload/engineering/_Institute_und_Zentren/IDP/forschungsschwerp

[5] Wildi, M. (2008.2) Efficient Multivariate Real-Time Filtering and Cointegration. *IDP-working paper, IDP-WP-08Sep-01.* http://www.idp.zhaw.ch/fileadmin/user_upload/engineering/_Institute_und_Zentren/IDP/forschungsschwerp WP-08Sep-01.pdf

[6] Wildi, M. (2009) Real-Time US-Recession Indicator (USRI): A Classical Cycle Perspective with Bounceback. *IDP-working paper, IDP-WP-09Jun-06.* http://www.idp.zhaw.ch/en/engineering/idp-institute-of-data-analysis-and-process-design/research/finance-risk-management-and-econometrics/economic-indices/us-economic-recession-indicator.html

[7] Wildi, M. (2011) Real-Time Signalextraction: a Shift of Perspective. *Estudios de Economia Aplicada* [forthcoming]. Working paper: http://blog.zhaw.ch/idp/sefblog/uploads/customized_final.pdf