

Spectral estimation of the multivariate impulse response

Granville Tunncliffe Wilson*

October 17, 2016

Abstract

One of the applications of cross-spectral estimation of stationary time series, developed some five decades ago, is the estimation of the lagged response of an output series causally dependent upon an input series, in the absence of feedback from the output to the input. The direct application of cross-spectral analysis for this purpose is no longer appropriate in the presence of feedback or more general inter-dependence of the series. In that case vector autoregressive modeling has been used, particularly in the econometric context, to estimate the response of one series to a shock or impulse in the innovations of another series. To achieve the same end, cross-spectral analysis requires the application of spectral factorization, and in this paper we demonstrate this methodology, explaining how it may be used to construct impulse response function estimates and their statistical properties. Our presentation includes an information criterion for choosing the smoothing bandwidth to be used for cross-spectral estimation.

Keywords: Cross-spectral analysis, spectral factorization, impulse response estimation, information criterion, bandwidth choice

1 Lagged response estimation

In this introductory section we illustrate the application of classical cross-spectral analysis to the estimation of what we shall call the lagged response of one series to another. The coefficients of this response are estimates of those in the projection of the current values of one series on all past, present and future values of the others. Taking x_t to be a zero-mean multivariate stationary time series with elements $x_{i,t}$, this projection has the form

$$x_{i,t} = \sum_{j \neq i} \sum_{k=-\infty}^{\infty} v_{i,j,k} x_{j,t-k} + n_{i,t}, \quad (1)$$

where $n_{i,t}$ is a stationary error series orthogonal to the terms $x_{j,t-k}$ in the sum on the right of (1). Methods of cross-spectral estimation, as presented for example in Shumway and Stoffer (2000), may be used to estimate the projection coefficients $v_{i,j,k}$ over a finite span of lags k , given observations of x_1, \dots, x_n . As an illustration Figure 1 shows estimates of these coefficients for the input and output of the gas furnace series of length 296 given in Jenkins and Watts (1968).

*Department of Mathematics and Statistics, Lancaster University, Lancaster LA1 4YF, UK.

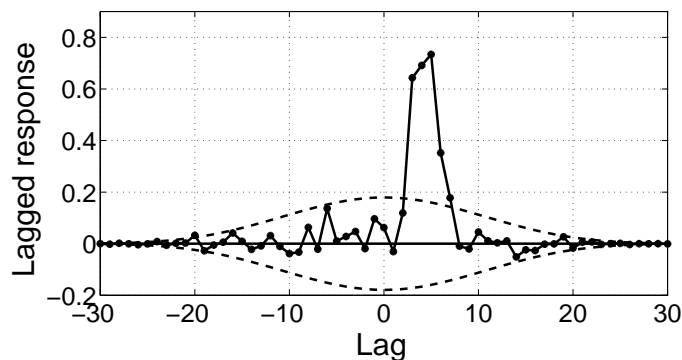


Figure 1: Spectral estimates for the Gas furnace series of the lagged response from Input to Output, using a smoothing bandwidth of 0.05. The lag is given as multiples of the sampling interval which is unknown.

These estimates are not significant at negative lags, confirming the absence of feedback from output to input in this open-loop system and allowing the response at positive lags to be interpreted as causal. In contrast, consider the spectral analysis of records of the respiratory system of a pre-term infant shown in Figure 2.

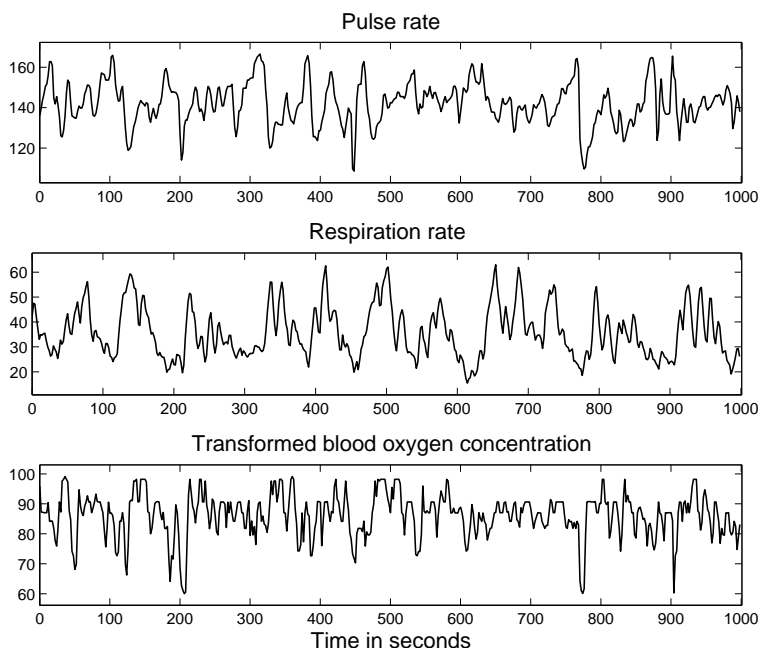


Figure 2: Records from monitoring the pulse rate, respiration rate and blood oxygen saturation of a pre-term infant, sampled at 2 second intervals. The blood oxygen saturation (oxygen level) is shown after applying an arcsine transformation.

Figure 3 shows the spectrum of the oxygen level series and the partial coherency between that series and the respiration rate. Although the variables are recorded at the rate of 10 samples per second we have used series sub-sampled at intervals of 2 seconds. From the

figure, this is sufficiently frequent to give an adequate representation of the power in the spectrum and the dependence between the series. The spectrum of the respiration series is similar to that of the oxygen level, but with a less prominent peak around frequency 0.05 cps.

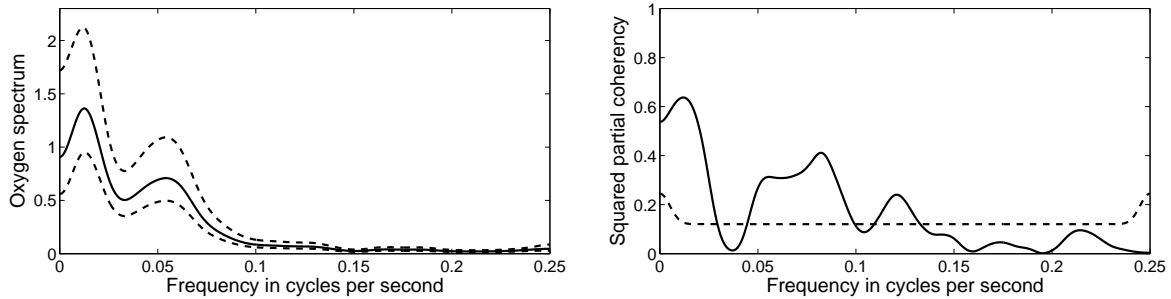


Figure 3: The estimated spectrum of the oxygen level series with 95% confidence limits and the estimated partial coherency between the respiration rate and oxygen level series with 95% significance level. The band width used was 0.025 cycles per second.

The partial coherency shown in Figure 3 implies a strongly significant (partial) dependence between the series. Spectral analysis of this three variable system taking the oxygen level as response gave the estimated lagged response to respiration rate shown on the left in Figure 4. This shows significant coefficients at both positive and negative lags, indicating the presence of feedback between these two series in the system. On the right of Figure 4 is the estimated lagged response with the role of these two series reversed, showing again clearly significant coefficients at both positive and negative lags, confirming the presence of feedback.

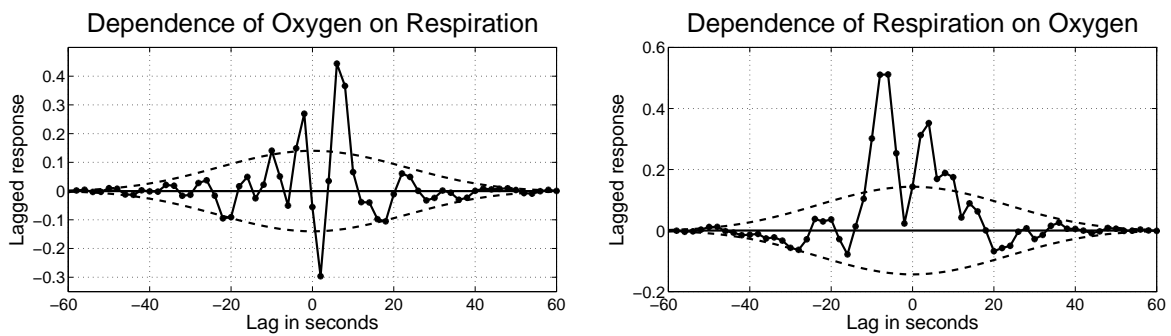


Figure 4: On the left is shown the estimated lagged response of the oxygen series to the respiration series with 95% significance limits set about zero, and on the right the estimated response with the series reversed.

For systems of interdependent series the causal response from one series to another has been estimated by fitting an autoregressive model and, from this, deriving the properties known as the impulse response function (IRF). This term has been used in somewhat different ways in the systems and econometrics literature. We will define and distinguish the usage

in these two contexts but illustrate only the systems usage. However, rather than autoregressive modeling, we will use spectral methods to derive these quantities, which requires the application of spectral factorization. A criterion for choice of the spectral smoothing bandwidth is also furnished by these methods.

2 Causal response estimation

We assume that x_t is a weakly stationary zero-mean multivariate time series having a Wold representation with no deterministic part, of the form

$$x_t = \sum_{k=0}^{\infty} \Psi_{xk} e_{t-k}. \quad (2)$$

Here, $\Psi_{x0} = I$, the identity matrix, and e_t is the linear innovation in x_t with variance matrix V_e . Let the covariance matrices of x_t be

$$\Gamma_{xk} = E x_t x_{t-k}' \quad (3)$$

where $'$ indicates transpose. Then, because e_t is multivariate white noise

$$\Gamma_{xk} = \sum_{i=0}^{\infty} \Psi_{xi} V_e \Psi_{x(i-k)}' \quad (4)$$

Define the (multivariate) spectrum of x_t over the frequency range $-0.5 \leq f \leq 0.5$ as

$$S_x(f) = \sum_{k=-\infty}^{\infty} \Gamma_k \exp(-2\pi i f k). \quad (5)$$

Then, corresponding to (4) we have

$$S_x(f) = \Psi_x \{ \exp(-2\pi i f) \} V_e \Psi_x \{ \exp(2\pi i f) \}' \quad (6)$$

where

$$\Psi_x \{ \exp(-2\pi i f) \} = \sum_{k=0}^{\infty} \Psi_{xk} \exp(-2\pi i k f). \quad (7)$$

Under very wide conditions, given $S_x(f)$, (6) may be solved uniquely for $\Psi_x \{ \exp(-2\pi i f) \}$ and V_e . Equivalently, given Γ_{xk} , (4) may be solved uniquely for the coefficients Ψ_{xk} and V_e . This is demonstrated in Tunncliffe Wilson (1978) which is also the basis of an efficient numerical procedure for constructing the solution. The procedure is called spectral factorization and $\Psi_x \{ \exp(-2\pi i f) \}$ the spectral factor.

From the Wold representation (2) we find that the element $\Psi_{xi,j,k}$ of Ψ_{xk} expresses the dependence of the future series value $x_{i,t+k}$ upon the present innovation, or shock, $e_{j,t}$. If we interpret (2) as a linear systems equation whereby the output x_t depends on the input

e_t , with this input able to be specified without restriction, then setting all values of $e_t = 0$ except $e_{j,t} = 1$ at $t = 0$ (an impulse) gives $x_{i,t} = \Psi_{x_{i,j,t}}$ for $t = 0, 1, 2, \dots$. This motivates the term impulse response function (IRF) for the sequence Ψ_{x_k} .

There is, however, a somewhat different usage of the term IRF. Define the power series

$$\Psi_x(z) = \sum_{k=0}^{\infty} \Psi_{x_k} z^k, \quad (8)$$

which is an extension of (7) from the unit circle to a general complex argument. A consequence of the Wold representation is that $\Psi_x(z)$ has no zeros for $|z| < 1$. Assume a slight strengthening of this property, that $\Psi_x(z)$ has no zeros for $|z| < \rho$ for some $\rho > 1$. Then we may also express

$$x_t - \sum_{k=1}^{\infty} \Pi_{x_k} x_{t-k} = e_t, \quad (9)$$

or in terms of the backward shift operator B

$$\Pi_x(B)x_{t-k} = e_t, \quad (10)$$

where, for $|z| \leq 1$, $\Pi_x(z) = \Psi_x(z)^{-1}$ is the power series with coefficients $\Pi_{x_0} = I$ and $-\Pi_{x_k}$ for $k > 0$. This is the infinite autoregressive representation of the series, which expresses explicitly the interdependence of the series, including any feedback between them, by re-writing (9) as

$$x_{i,t} = \sum_{j \neq i} \sum_{k=1}^{\infty} \Pi_{x_{i,j,k}} x_{j,t-k} + e_{i,t}. \quad (11)$$

The response in which we are now interested is the effect of an impulse given to one series (the input), say $x_{j,t}$, on subsequent values of another series (the output) $x_{i,t}$ *in the absence of feedback to the input*. This is constructed using (11) by setting $e_t = 0$ for all t , $x_{j,t} = 0$ for $t < 0$, and $x_{j,t} = 1$ for $t = 0$. Furthermore, set coefficients $\Pi_{x_{j,l,k}} = 0$ for the specified j and all l and $k > 0$, so that there is no feedback to the input from its own past or that of any other series, resulting also in $x_{j,t} = 0$ for $t > 0$. Thus $x_{j,t}$ is the impulse series. We will distinguish the resulting response sequence $x_{i,0}, x_{i,1}, x_{i,2}, \dots$, by calling it the open loop impulse response function (OLIRF) of series i to series j . It is given formally by the coefficients in the expansion of

$$-\Pi_{x_{i,j}}(B)/\Pi_{x_{i,i}}(B). \quad (12)$$

There is in general no guarantee that this response is stable, in the sense that it converges to zero. It may diverge; indeed, the lack of stability of this response is a major reason why feedback is required in many systems.

Figure 5 shows, on the left, the OLIRF estimated for the Gas furnace series by this means, following factorization of its sample spectrum. The bandwidth of 0.0278 used for spectral smoothing is determined by minimizing a criterion defined in the next section, where the properties of the estimates are also presented. On the right of Figure 5 is the estimated lagged response estimated in the same manner as that in Figure 1 but using the same bandwidth of 0.0278 as for the plot on the left of the figure and plotted only for non-negative lags. There is, not surprisingly, a very close similarity between the two estimates, because there is no

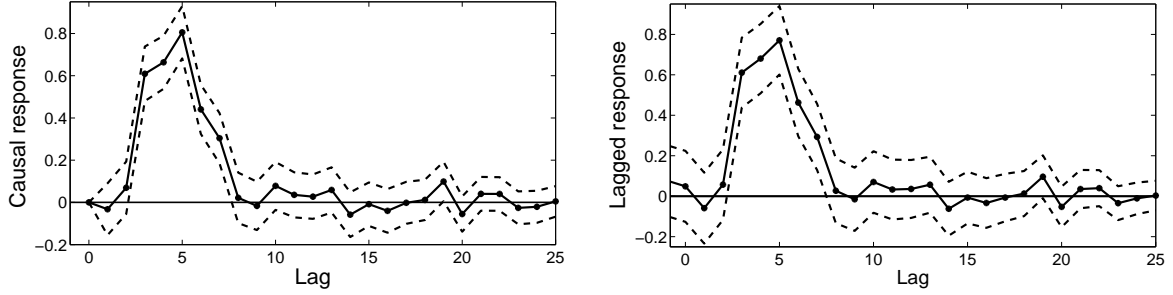


Figure 5: On the left is shown the estimated causal response of the output series to the input series for the Gas Furnace series, with 95% error limits about the estimates. A smoothing bandwidth of 0.0278 was used in this analysis and, for comparison, on the right is shown the estimated lagged response using the same bandwidth and similar error limits.

feedback in this system. For both of these plots the two standard error limits are centered on the estimates.

Figure 6 similarly shows, on the left, the OLIRF for the response of oxygen to respiration from the analysis of the three infant monitoring series by factorization of their sample spectrum smoothed with a bandwidth of 0.05. For comparison, on the right is shown the estimated lagged response estimated in the same manner as that in Figure 4 but using the same bandwidth of 0.05 as for the plot on the left of the figure and plotted only for non-negative lags. These two plots are, again, quite similar, apart from the much less pronounced negative response at lag 1 in the causal response, which we ascribe to distortion by feedback in the relationship between the series.

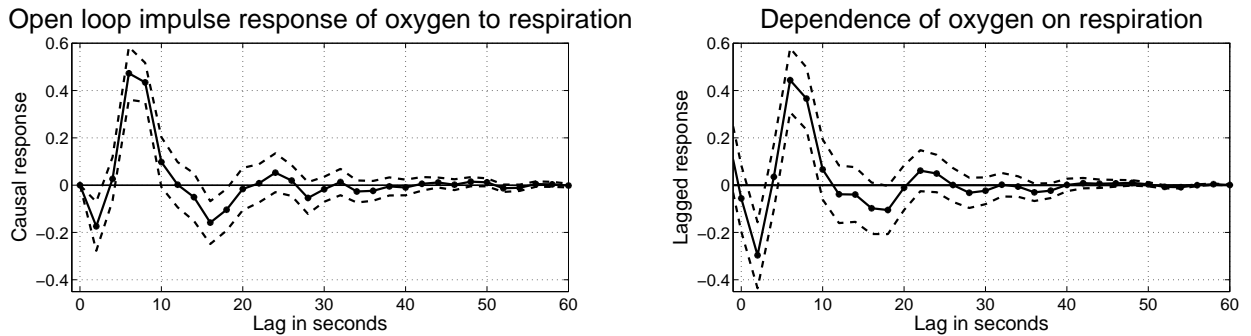


Figure 6: On the left is shown the estimated causal response of the oxygen series to the respiration series for the infant monitoring series, with 95% error limits about the estimates. A smoothing bandwidth of 0.05 was used in this analysis and, for comparison, on the right is shown the estimated lagged response using the same bandwidth and similar error limits.

There is one further comment appropriate at this point, concerning how any contemporaneous dependence between the series should be treated. This dependence would be evident in significant off-diagonal coefficients of the estimated innovation matrix V_e in (6). If it could be assumed that this coefficient arises from a lag zero dependence of the output response on the input, then the OLIRF given in (12) should be modified by correcting the term $\Pi_{x i,j}(B)$

in its numerator by a simultaneous regression term of the form $\alpha \Pi_{x,j,j}(B)$. For both of our examples this regression is not significant and we show no lag zero response in our results.

3 Selection of the smoothing bandwidth

For the two examples we presented in the previous section, the smoothing bandwidth b used for multivariate spectral estimation of the series was given by finding the minimum of a non-parametric information criterion (NIC) defined by

$$\text{NIC}(b) = \gamma n \log \det \hat{V}_e + 2m^2 p \quad (13)$$

where n is the series length, \hat{V}_e is the estimate of V_e obtained by applying the factorization (6) to the smoothed spectrum estimate $\hat{S}_x(f)$ of the series, m is the dimension of the series and p is the equivalent order of the smoother. We define this as

$$p = \sum_{k>0} w_k^2 \quad (14)$$

where w_k is the lag window used for spectral smoothing. For the window we are using, the equivalent order is given to an excellent approximation by $p = 0.5(b^{-1} - 1)$. The multiplier γ is a tapering efficiency factor defined in terms of a normalized data taper h_t , that may have been applied to the series, as

$$\gamma = n / \sum_{t=1}^n h_t^4. \quad (15)$$

In order to evaluate the NIC it is not necessary to carry out the complete factorization (6) because

$$\log \det \hat{V}_e = \int_{-0.5}^{0.5} \log \det \hat{S}_x(f) df \quad (16)$$

which may be evaluated numerically to give the required term in (13). To clarify, for each of a range of proposed values of the bandwidth b , we construct the smoothed spectrum estimate $\hat{S}_x(f)$ using that bandwidth, we use (16) to give the first term on the RHS of (13) and (14) for the second term, to give NIC as a function of b . The selected bandwidth is given at the minimum of $\text{NIC}(b)$.

Figure 7 shows plots of the NIC against the equivalent model order for both of our example series. For the Gas furnace series the selected order and bandwidth are, respectively, 17 and 0.0278. For the respiratory system series they are 10 and 0.05.

We still have to investigate the properties of spectral estimation using this bandwidth selection criterion, but we expect that they are similar to the optimal properties of autoregressive spectral estimation using the AIC as presented by Shibata (1981) and, in particular, we assume that p grows sufficiently slowly with n to ensure consistent estimation of the spectrum.

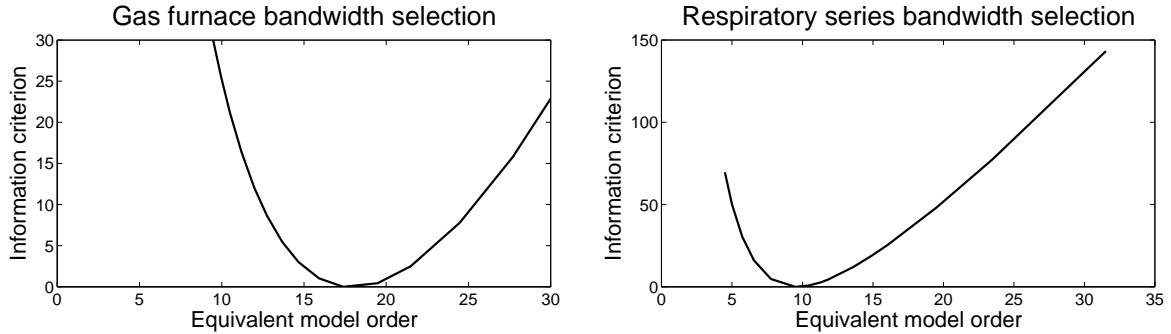


Figure 7: The non-parametric information criteria used to select the smoothing bandwidth for the Gas furnace series and the Infant respiratory system series. The criterion is plotted against the equivalent model order of the bandwidth.

4 Derivation of the statistical properties

The methodology and results presented in the previous sections are based on the statistical properties of the factorized sample spectrum. Rather than delegate the derivation of these properties to an appendix, we present in this and the following section an outline which we hope will be reasonably accessible and convey to the reader the underlying ideas. Though informal, the derivations are valid and original and could be extended to formal proofs. One subsidiary formal theorem is presented in an appendix.

Our starting point is the factorization, in the same manner as (6), of the estimated spectrum of the series:

$$\hat{S}_x(f) = \hat{\Psi}_x\{\exp(-2\pi if)\}\hat{V}_e\hat{\Psi}_x\{\exp(2\pi if)\}'. \quad (17)$$

The derivation of all statistical properties is then based on the result that the estimated factor is related to the true factor by the approximation

$$\hat{\Psi}_x\{\exp(-2\pi if)\}\hat{R}_e \approx \Psi_x\{\exp(-2\pi if)\}R_e\hat{\Psi}_a\{\exp(-2\pi if)\}\hat{R}_a, \quad (18)$$

where $\hat{\Psi}_a\{\exp(-2\pi if)\}$ is the factor of the estimated spectrum of the normalized innovations series $a_t = (R_e)^{-1}e_t$, and R_e , \hat{R}_e and \hat{R}_a are, respectively, the left Choleski factors of V_e , \hat{V}_e and \hat{V}_a . The estimated spectrum, $\hat{S}_a(f)$, is tapered and smoothed in exactly the same manner as $\hat{S}_x(f)$.

To make use of this result we need also the further approximation that the coefficients of $\hat{\Psi}_a\{\exp(-2\pi if)\}$ are given in terms of the windowed sample autocorrelations of a_t as

$$\hat{\Psi}_{a_k} \approx w_k r_{a_k}. \quad (19)$$

Note that we do not form the innovation series a_t and the estimated quantities $\hat{\Psi}_a\{\exp(-2\pi if)\}$ and \hat{R}_a on the RHS of (18), but we shall use simulated values of these quantities, together with (18) and (19), to generate confidence intervals for estimated impulse responses.

Before we justify (18) and 19 we explain their application to deriving the order selection criterion NIC and the construction of the error limits of the OLIRF causal response.

For the NIC we first consider the bias in the estimate of $\log \det \hat{V}_e$ given by (16). This has been investigated by Mohanty and Pourahmadi (1996) who give a formula based on the number of degrees of freedom of the spectral estimates calculated from the frequency domain smoothing window. This can also be expressed in terms of the lag window; in fact as a function of the equivalent order p that we defined in (14), the series dimension m and the sample size n . When expanded, the lowest order term of the formula is $m^2 p/n$ and it can be checked that this alone gives an accurate approximation to the formula for realistic values of m , p and n . Although Mohanty and Pourahmadi (1996) do not consider series tapering, the effect of this on all the statistical properties of spectral estimates is very well approximated as equivalent to a reduction in the series length from n to γn .

With our definition of p in (14), the form of the NIC (13) is the same as that of the well known AIC of Akaike (1973), in which p corresponds to the order of a multivariate Autoregressive model. This explains why we call p the equivalent order. To complete the justification of the NIC we appeal to the principle of minimum final prediction error variance for model order selection, or FPE. The FPE variance for our smoothed spectrum factor, as approximated in (18), is the determinant of the variance of the one step ahead error that would result from applying the factor, as a model, to an *independent realization* of the series x_t . This variance includes the increase in prediction error due to the uncertainty in using the estimated model coefficients.

From (18), this is the variance of $R_e \tilde{a}_{t+1}$, where

$$\tilde{a}_{t+1} = a_{t+1} - \hat{a}_{t+1} = a_{t+1} - \sum_{k>0} \hat{\Psi}_{a_k} a_{t+1-k} \quad (20)$$

is the error of predicting a_{t+1} in a new realization of the series a_t . Now using (19) and the property that in large samples the elements of r_{a_k} are all uncorrelated with mean zero and variance $1/n$, we have the variance of this term given by

$$\text{Var } \tilde{a}_{t+1} = I_m (1 + m \sum_{k>0} w_k^2/n) = I_m (1 + mp/n). \quad (21)$$

To the same order of approximation the logarithm of the determinant of $\text{Var } R_e \tilde{a}_{t+1}$ is

$$\log \det V_e + m^2 p/n. \quad (22)$$

Now substituting the estimate of $\log \det V_e$ from (16), correction for bias by addition of the term $m^2 p/n$, gives the NIC in (13).

To obtain the large sample standard errors of the estimated causal responses used to construct the error limits shown on the left in each of the Figures 5 and 6, we consider the expression (18) which relates the estimated spectral factor of the observed series x_t , to that of its orthogonalized innovation series a_t . The distributional properties of this second factor are determined by (19) and the known properties of the sample autocorrelations r_{a_k} of a_t . By applying the linearization procedure of the delta-method this relationship could be used to express the large sample variance matrix of the whole set of estimated coefficients $\hat{\Psi}_{x_k}$ and \hat{R}_e of the required spectral factor, in terms of the true coefficients Ψ_{x_k} and R_e of the unknown spectral factor. Substituting the values of the estimated coefficients for the true coefficients would then give a consistent estimate of this variance matrix.

To implement the delta method we use a simulation technique by the following steps:

1. Generate a sample from r_{ak} which are taken as independent normal with mean zero and variances $1/n$.
2. Scale this whole sample by a small factor, say $\alpha = 0.1$, to ensure local linearity of these perturbations.
3. Similarly generate a sample from \hat{R}_a by Choleski factorization of a sample from \hat{V}_a with expected value I and scaled variances α^2/n for off-diagonal and $2\alpha^2/n$ for diagonal terms.
4. Use these to evaluate the windowed coefficients $\hat{\Psi}_{ak}$ in (19) and thence the RHS of (18), substituting the estimated factor $\hat{\Psi}_x\{\exp(-2\pi if)\}$ in place of the true unknown factor $\Psi_x\{\exp(-2\pi if)\}$.
5. The LHS of (18) is then a perturbed value of the estimate $\hat{\Psi}_x\{\exp(-2\pi if)\}$, from which the set of perturbed estimated coefficients $\hat{\Psi}_{xk}$ and \hat{R}_e is derived.
6. Calculate the perturbations of these coefficients and (re)scale them by $1/\alpha$. This is a sample from the variance matrix of these coefficients.
7. Replicate this process with sufficient further samples to form a good estimate of the variance matrix.

This scheme is simply extended to construct the variance matrix of the estimated responses on the left of Figures 5 and 6 and any other function of the estimated coefficients.

5 Further derivation of properties

The starting point for deriving the central expression (18), is the large sample relationship between the smoothed spectrum of the series and that of its innovation series, smoothed in the same way:

$$\hat{S}_x(f) \approx \Psi_x\{\exp(-2\pi if)\} R_e \hat{S}_a(f) R_e' \Psi_x\{\exp(2\pi if)\}'. \quad (23)$$

Factorizing both sides then gives (18). Essentially, (23) states that the operations of filtering and spectral smoothing approximately commute, but for this to be true the filter must be slowly varying over the width of the smoothing window. Translated into our notation, Corollary 5 of Hannan (1970, p. 214), states the result (23) between the un-smoothed sample spectra. The Proof of Theorem 11, Hannan (1970, p. 318) gives further detail. The error in the approximation between the un-smoothed sample spectra is of the order $n^{-0.5}$. Smoothing will reduce this by a factor of the order $(2bn)^{-0.5} \approx (n/p)^{-0.5} = p^{0.5}/n$, giving an error of order n^{-a} where $a > 0.5$.

The further approximation (19) can be written, for $z = \exp(-2\pi if)$, in terms of the sample covariances C_{ak} of a_t as

$$\hat{S}_a(f) = \sum_k w_k C_{ak} z^k \approx (I + \sum_{k>0} w_k r_{ak} z^k) \hat{V}_a (I + \sum_{k>0} w_k r_{ak} z^{-k})'. \quad (24)$$

We demonstrate the validity of this approximation to terms of order $n^{-0.5}$, neglecting terms of higher order, using the property that in large samples the estimates r_{ak} are uncorrelated with mean 0 and variance n^{-1} . From this, on expanding the RHS of (24) the product term $(\sum_{k>0} w_k r_{ak} z^k) \hat{V}_a (\sum_{k>0} w_k r_{ak} z^{-k})'$ has root mean square error of order $(\sum w_k^2)^{0.5}/n = p^{0.5}/n = n^{-a}$. The constant terms $C_{a0} = \hat{V}_a$, leaving the approximation $C_{ak} \approx r_{ak} \hat{V}_a$ to be demonstrated for $k \neq 0$.

Writing s_{ai} for $\hat{V}_{ai,i}^{0.5}$,

$$\left(r_{ak} \hat{V}_a\right)_{i,j} = \sum_{\ell} \{C_{ak,i,\ell}/(s_{ai}s_{a\ell})\} \hat{V}_{a\ell,j} \quad (25)$$

in which the terms for $\ell \neq j$ in the sum on the RHS are of order n^{-1} leaving

$$C_{ak,i,j} s_{aj}/s_{ai} = C_{ak,i,j} + C_{ak,i,j}(s_{aj} - s_{ai})/s_{ai}. \quad (26)$$

The second term on the RHS is again of order n^{-1} . These bounds on the approximation error of the coefficients translate again to a bound of the order $(\sum w_k^2)^{0.5}/n = n^{-a}$ on the error in the spectral factor. We set

$$\tilde{\Psi}_a(z) = (I + \sum_{k>0} w_k r_{ak} z^k). \quad (27)$$

Substituting from (24) into (23), and using further the notation $z = \exp(-2\pi if)$, we now have the approximation

$$\hat{S}_x(f) = \hat{\Psi}_x(z) \hat{V}_e \hat{\Psi}_x(z^{-1})' \approx \Psi_x(z) R_e \tilde{\Psi}_a(z) \hat{V}_a \tilde{\Psi}_a(z^{-1})' R_e' \Psi_x(z^{-1})'. \quad (28)$$

The relative error bound on these combined approximations remains as n^{-a} where $a > 0.5$ and hence

$$\left\{ \hat{\Psi}_x(z) \hat{R}_e \right\}^{-1} \Psi_x(z) R_e \tilde{\Psi}_a(z) \hat{V}_a \tilde{\Psi}_a(z^{-1})' R_e' \Psi_x(z^{-1}) \left\{ \hat{R}_e' \hat{\Psi}_x(z^{-1})' \right\}^{-1} = I + g(z) \quad (29)$$

where $\|g(z)\| < \epsilon$ is of order n^{-a} . From Theorem 1 of the appendix, we therefore have

$$\left\{ \hat{\Psi}_x(z) \hat{R}_e \right\}^{-1} \Psi_x(z) R_e \tilde{\Psi}_a(z) \hat{R}_a = I + h(z) \quad (30)$$

where the coefficients of $h(z)$ are of the same order. This justifies our use of (18) and (19), because our estimates of impulse responses use only a finite number of the coefficients of $\hat{\Psi}_x(z)$, and the error of approximation is of higher order than $n^{-0.5}$, which is that of the sampling variability of our estimates.

6 Not the whole truth

One of the valuable aspects of spectral analysis is how it reveals the information available for modeling time series data - or the lack of such information. The estimated lagged response of Oxygen to Respiration shown on the left of Figure 4 is obtained by Fourier transformation

of the estimated frequency domain response. The gain of this response is shown on the left of Figure 8, from which it is seen not be significant for most of the frequency range above 0.3, though highly significant at most frequencies below this. This significance level is the same as that apparent for the partial coherency shown in Figure 3. The lack of significance of the gain does not necessarily cause a problem when estimating the lagged response, but does so in this case because the estimates are large and subject to considerable uncertainty. This is due to substantial noise in the Respiration series at these higher frequencies. On transforming back to the lagged responses this would lead to a corresponding substantial increase in the variability of the estimates.

We now admit that we avoided this adverse effect on the estimated lagged response by the simple device of zeroing the estimated gain over the higher frequency range as shown on the right in Figure 4. This is done smoothly over a narrow band just above the frequency 0.15 cps. We used the same device when generating both the estimated lagged response in Figure 4 and the two responses show in Figure 6. For the causal response on the left of this figure, the zeroing was applied to the frequency domain representation of (12) which is transformed to give this response. It should not be applied before the spectral factorization step.

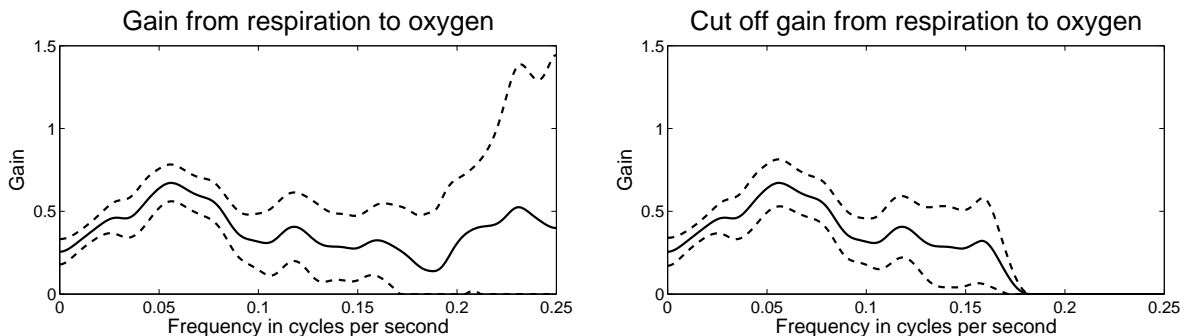


Figure 8: The estimated gain from Respiration to Oxygen and the truncated gain used to construct the estimates in Figure 6.

This problem of unduly high variability in the impulse response estimates is not confined to spectral methods; it can also occur in estimates derived from autoregressive modeling. In that case some form of shrinkage can be applied to the estimates of the autoregressive coefficients to overcome the problem, as in Litterman (1986). However, the problem has the same cause whether it arises in the time or spectral domain, and it is in the spectral domain that this can be more directly detected and solved, as in our example.

Some of the methodology presented in this work is also described in Tunnicliffe Wilson et al. (2015), where an extension to very high frequency sampled series is given, including the example of the infant respiration system sampled at 0.1 second intervals. This paper complements this book by presenting original informal derivations of this methodology for the lower frequency sampled series as applied to the foregoing examples.

Appendix

We state and prove that a small variation of the spectrum leads to a corresponding variation of the same order in the coefficients of its spectral factor. We use the spectral norm of matrices.

Theorem 1 *Let $S(z) = \sum_k \Gamma_k z^k$, for $z = \exp(-2\pi if)$, be a multivariate spectral density such that $S(z) = I + g(z)$ where $\|g(z)\| \leq \epsilon$ and let $H(z) = H + \sum_{k>0} h_k z^k$ be its spectral factor, such that $S(z) = H(z)H(z)^*$, where $*$ indicates conjugate transpose, and H is lower triangular. Then $\|H - I\| \leq K\epsilon$ and $\|h_k\| \leq K\epsilon$ for fixed K , on neglecting higher orders of ϵ .*

Proof. From $\Gamma_0 = \int_{-0.5}^{0.5} S(z)df = HH' + \sum_{k>0} h_k h_k' \in (1 \pm \epsilon)I$ we have $HH' \leq (1 + \epsilon)I$. The eigenvalues λ_i of HH' are all therefore less than $1 + \epsilon$ and $\log \lambda_i < \epsilon$.

Further constraints on H derive from $\int_{-0.5}^{0.5} \log \det S(z)df = \log \det HH'$. The first order term in the expansion of $\log \det S(z)$ about I is $\text{trace}\{S(z) - I\}$ which is bounded by $m\epsilon$. From this we have $|\log \det(HH')| = |\sum \log \lambda_i| \leq m\epsilon$.

Then for any i , $\log \lambda_i > -m\epsilon - \sum_{j \neq i} \log \lambda_j > -(2m - 1)\epsilon$ giving, to first order, $[1 - (2m - 1)\epsilon]I < HH' < (1 + \epsilon)I$. We can then bound $\sum_{k>0} h_k h_k' \leq (1 + \epsilon)I - HH' \leq 2m\epsilon I$. A similar argument utilizing $\log \det(HH') = \sum \log H_{i,i}^2$ shows that $1 - (2m - 1)\epsilon < H_{i,i}^2 < 1 + \epsilon$.

Now let $H = I + h_0$ so that h_0 is lower triangular, and from the above we can determine that $h_0 h_0' < (4m - 2)\epsilon I$. Also let $h(z) = \sum_{k>0} h_k z^k$, which we can now bound by $\|h(z)h(z)\|^2 < (6m - 2)\epsilon$. Then $I + g(z) = H(z)H(z)^* = I + h(z) + h(z)^* + h(z)h(z)^*$ gives

$$h(z) + h(z)^* = \sum_{k \geq 0} h_k \exp(-2\pi ifk) + h_k' \exp(+2\pi ifk) = g(z) - h(z)h(z)^*.$$

The norm of the RHS is bounded by $(6m - 1)\epsilon$ which is therefore a bound on the norm of each Fourier coefficient h_k .

Remark. From the above, $\|h(z)\|$ can only be bounded by $\sqrt{(6m - 2)\epsilon}$, which is not of the order required. A bound of the form $K\epsilon$ on $\|h(z)\|$ requires smoothness conditions on $S(z)$.

References

- H. Akaike. A new look at statistical model identification. *IEEE Transactions on Automatic Control*, AC-19(2):716–723, 1973.
- E. J. Hannan. *Multiple time series*. New York: Wiley, 1970.
- G. M. Jenkins and D. G. Watts. *Spectral analysis and its applications*. San Francisco: Holden-Day, 1968.
- R. B Litterman. Forecasting with Bayesian vector autoregressions - five years of experience. *Journal of Business and Economic Statistics*, 4:25–38, 1986.

- R. Mohanty and M Pourahmadi. Estimation of the generalized prediction error variance of a multiple time series. *Journal of the American Statistical Association*, 91:294–299, 1996.
- R. Shibata. An optimal autoregressive spectral estimate. *The Annals of Statistics*, 8:300–306, 1981.
- R. H. Shumway and D. S. Stoffer. *Time series analysis and its applications*. New York: Springer-Verlag, 2000.
- G. Tunncliffe Wilson. A convergence theorem for spectral factorization. *Journal of Multivariate Analysis*, 8:222–232, 1978.
- G. Tunncliffe Wilson, M. Reale, and J. Haywood. *Models for dependent time series*. New York, CRC Press, 2015.