The Journal of Marine Research, one of the oldest journals in American marine science, published important peer-reviewed original research on a broad array of topics in physical, biological, and chemical oceanography vital to the academic oceanographic community in the long and rich tradition of the Sears Foundation for Marine Research at Yale University.

An archive of all issues from 1937 to 2021 (Volume 1–79) are available through EliScholar, a digital platform for scholarly publishing provided by Yale University Library at https://elischolar.library.yale.edu/.

Requests for permission to clear rights for use of this content should be directed to the authors, their estates, or other representatives. The Journal of Marine Research has no contact information beyond the affiliations listed in the published articles. We ask that you provide attribution to the Journal of Marine Research.

Yale University provides access to these materials for educational and research purposes only. Copyright or other proprietary rights to content contained in this document may be held by individuals or entities other than, or in addition to, Yale University. You are solely responsible for determining the ownership of the copyright, and for obtaining permission for your intended use. Yale University makes no warranty that your distribution, reproduction, or other use of these materials will not infringe the rights of third parties.

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. https://creativecommons.org/licenses/by-nc-sa/4.0/
Complex Generalization of Canonical Correlation and its Application to a Sea-level Study

Motoyasu Miyata

Geophysical Institute
University of Tokyo
Tokyo, Japan

Abstract

The concept of multiple coherence is generalized to study the association between two sets of time series. "Intergroup coherence" is introduced as a measure of the linear relationship between the time histories at a collection of points and the time histories at a different collection of points. This theory is applied to a spectral analysis of marigrams from Canton Island.

1. Introduction. The theory of inter-relation between two sets of real-valued variables was developed by Hotelling (1936). The object of the present study is to extend his theory to the case for complex-valued variables and to define "integroup coherence", which corresponds to his "maximum canonical correlation". This is applied in the frequency domain to analyze power spectra of time series.

2. Theory. Let $x_k(t)$ be a component of a real-valued q-dimensional multivariate stationary process with mean zero. Let $X_k(f)$ be the finite Fourier transform of a realization of the process:

$$X_k(f) = \frac{1}{(n)^{1/2}} \sum_{t=1}^{n} x_k(t) e^{-2\pi j ft/n},$$

where the increment of time is taken to be the sampling interval. Let $U(f)$ be the q-dimensional random vector whose components are the $X_k(f)$. For large sample size, the covariance matrix $\Sigma(f)$ of $U(f)$ is approximately equal to the aliased spectral-density matrix of the stationary process:

1. Accepted for publication and submitted to press 22 January 1970.

202
$E[U(f) U^*(f)] \cong \Sigma(f) = \begin{pmatrix} S_{11}(f) & S_{12}(f) & \ldots & S_{1q}(f) \\ S_{21}(f) & S_{12}(f) & \ldots & S_{2q}(f) \\ \vdots & \vdots & \ddots & \vdots \\ S_{q1}(f) & S_{q2}(f) & \ldots & S_{qq}(f) \end{pmatrix}$, \hspace{1cm} (1)

where $E$ means ‘expected value’ and an * denotes the conjugate transpose. The elements of the covariance matrix are:

$S_{lm}(f) = E[X_l(f) X_m(f)]$

where — denotes the conjugate. Since only variances and covariances are of interest, it is assumed that $E[X_k(f)] = 0$. Note that $\Sigma(f)$ is Hermitian.

Let $U(f)$ be partitioned into two subvectors of $p$ and $(q-p)$ components (assuming $p \leq q$ for convenience) and let $\Sigma(f)$ be partitioned into four submatrices as described below; these become:

$U(f) = \begin{pmatrix} U_1 \\ U_2 \end{pmatrix}$, \hspace{1cm} $\Sigma(f) = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$, \hspace{1cm} (2)

where

$U_1 = \begin{pmatrix} X_1(f) \\ \vdots \\ X_p(f) \end{pmatrix}$, \hspace{1cm} $U_2 = \begin{pmatrix} X_{p+1}(f) \\ \vdots \\ X_q(f) \end{pmatrix}$

$\Sigma_{11} = \begin{pmatrix} S_{11}(f) & \ldots & S_{1p}(f) \\ \vdots & \ddots & \vdots \\ S_{p1}(f) & \ldots & S_{pp}(f) \end{pmatrix}$, \hspace{1cm} $\Sigma_{12} = \begin{pmatrix} S_{1p+1}(f) & \ldots & S_{1q}(f) \\ \vdots & \ddots & \vdots \\ S_{p+p+1}(f) & \ldots & S_{pq}(f) \end{pmatrix}$

$\Sigma_{21} = \begin{pmatrix} S_{p+11}(f) & \ldots & S_{p+p}(f) \\ \vdots & \ddots & \vdots \\ S_{q1}(f) & \ldots & S_{qq}(f) \end{pmatrix}$, \hspace{1cm} $\Sigma_{22} = \begin{pmatrix} S_{p+p+11}(f) & \ldots & S_{p+p+q}(f) \\ \vdots & \ddots & \vdots \\ S_{q+p+1}(f) & \ldots & S_{qq}(f) \end{pmatrix}$

The functional notation $(f)$ is to be understood. Note that $\Sigma_{11}^* = \Sigma_{11}$, $\Sigma_{22}^* = \Sigma_{22}$, and $\Sigma_{12}^* = \Sigma_{21}$.

Now we define two linear functions, $z_1 = L^* U_1$ and $z_2 = M^* U_2$, where $L$ and $M$ are arbitrary complex-valued column vectors. Each function is normalized so that its variance is equal to unity. Both $z_1$ and $z_2$ are scalers:

$E[z_1 \bar{z}_1] = E[L^* U_1 U_1^* L] = L^* \Sigma_{11} L = 1$, \hspace{1cm} (3)
Similarly,
\[ E[z_1 \bar{z}_2] = E[M^* U_2 U_2^* M] = M^* \Sigma_{22} M = 1. \] (4)

Similarly,
\[ E[z_1 \bar{z}_2] = L^* \Sigma_{12} M, \] (5)
\[ E[z_2 \bar{z}_1] = M^* \Sigma_{21} L. \] (6)

We now consider the problem to be one of prediction of \( z_1 \) from \( z_2 \). The error, in the least-squares sense, will be designated \( e \):

\[
e = z_1 - z_2,
E[ee] = E[(z_1 - z_2)(\bar{z}_1 - \bar{z}_2)],
= E[z_1 \bar{z}_1 + z_2 \bar{z}_2 - z_1 \bar{z}_2 - z_2 \bar{z}_1],
= L^* \Sigma_{11} L + M^* \Sigma_{22} M - L^* \Sigma_{12} M - M^* \Sigma_{21} L.
\] (7)

\( L \) and \( M \) must now be chosen so as to minimize the mean-square error \( E[ee] \) subject to the normalization conditions (3) and (4). Introducing Lagrangean multipliers \( \lambda \) and \( \mu \) (both \( \lambda \) and \( \mu \) being real),

\[
g = E[ee] + \lambda (L^* \Sigma_{11} L - 1) + \mu (M^* \Sigma_{22} M - 1),
= (1 + \lambda) L^* \Sigma_{11} L + (1 - \mu) M^* \Sigma_{22} M - L^* \Sigma_{12} M - M^* \Sigma_{21} L - \lambda - \mu; \] (8)

\( g \) is a real-valued function of complex variables. To obtain the minimum value for \( g \), a conventional technique is used (see, for instance, Margenau and Murphy 1943). Let \( w_j = u_j + iv_j \) be an element of \( L \). The function \( g \) is not analytic with respect to \( w_j \) (or \( w_j^* \)), although it is differentiable with respect to \( u_j \) and \( v_j \). However, instead of taking \( u_j \) and \( v_j \) as independent variables, we consider \( w_j \) and \( w_j^* \) as independent variables and differentiate both sides of (8) in the formal sense. Symbolically,

\[
\frac{\partial g}{\partial L} = (1 + \lambda) L^* \Sigma_{11} M - M^* \Sigma_{21} = 0.
\] (9)

In the same way,
\[ (1 + \mu) M^* \Sigma_{22} - L^* \Sigma_{12} = 0. \] (10)

Combining (3) and (9), and (4) and (10),
\[ I + \lambda = M^* \Sigma_{21} L, \] (11)
\[ I + \mu = L^* \Sigma_{12} M. \] (12)

Since both \( I + \lambda \) and \( I + \mu \) are real (with \( \lambda \) and \( \mu \) being real Lagrangian multipliers) and \( (M^* \Sigma_{21} L)^* = L^* \Sigma_{12} M, \) \( I + \lambda = I + \mu = r. \) Minimizing (7) is equivalent to maximizing \( 2r = E[z_1 \bar{z}_2] + E[z_2 \bar{z}_1]. \) Thus, \( r \) should give a measure of the association between the two sets of variables, \( U_1 \) and \( U_2. \)
Taking the complex conjugates of (9) and (10) and rewriting them,

\[ r \Sigma_{11} \mathbf{L} - \Sigma_{12} \mathbf{M} = 0, \]  
\[ -\Sigma_{21} \mathbf{L} + r \Sigma_{22} \mathbf{M} = 0, \]

which can be solved simultaneously:

\[ (\Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} - r^2 \Sigma_{11}) \mathbf{L} = 0. \]

Solution of eq. (15) produces the eigenvalues \( r_1^2, r_2^2, \ldots, r_p^2 \) and their corresponding eigenvectors \( \mathbf{L}_1, \mathbf{L}_2, \ldots, \mathbf{L}_p \). Since both \( \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \) and \( \Sigma_{11} \) are Hermitian and positive definite, all the eigenvalues are positive. Among the \( 2p \) possible values for \( r \), the positive square root of the largest eigenvalue should be selected as a solution to minimize the mean-square error. Now the “intergroup coherence” \( \varrho \) is defined as

\[ \varrho = [\max (r_1^2, r_2^2, \ldots, r_p^2)]^{-1/2}. \]

Let the eigenvector for which \( r^2 = \varrho^2 \) be called \( \mathbf{A} \), where

\[ \mathbf{A} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_p \end{pmatrix}. \]

Then \( \alpha_1, \alpha_2, \ldots, \alpha_p \) are complex-valued coefficients for each variable in the first group \([X_1(f), \ldots, X_p(f)]\). The coefficients for the second group \([X_{p+1}(f), \ldots, X_q(f)]\) can be derived by using (14). \( \mathbf{B} \) represents the corresponding vector:

\[ \mathbf{B} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_{q-p} \end{pmatrix} = \frac{1}{\varrho} \Sigma_{22}^{-1} \Sigma_{21} \mathbf{A}. \]

The elements of these vectors are similar to the regression coefficients in ordinary complex multiple-regression analysis, but here only the magnitudes are determined uniquely and the phases are relative within the group. The quantities \( \mathbf{A}^* \mathbf{U}_1 \) and \( \mathbf{B}^* \mathbf{U}_2 \) represent normalized linear combinations of \( \mathbf{U}_1 \) and \( \mathbf{U}_2 \) with maximum coherency. \( \mathbf{A}, \mathbf{B}, \) and \( \varrho \) are related to each other as

\[ \varrho = \mathbf{B}^* \Sigma_{21} \mathbf{A} = \mathbf{A}^* \Sigma_{12} \mathbf{B}. \]
From (3), (4), (7), and (19),

\[
E[e \tilde{e}] = 2(1 - \varrho) \geq 0;
\]

\[
\Rightarrow 0 \leq \varrho \leq 1,
\]

(20)
since the positive square root was taken in (16).

For a special case when \( \varrho = 1 \), intergroup coherence becomes the same as multiple coherence, though the values of \( \beta_1 \ldots \beta_{q-1} \) are different from those of ordinary regression coefficients.

3. Application to Sea-level Spectra. The lunar gravitational potential function was generated by a computer program [Munk 1966], and the function was expanded into spherical harmonics, retaining terms of the second and third degree. Then the long-period equilibrium tides, \( a_2^o \) and \( a_3^o \), were computed from the corresponding coefficients. Records of sea level (z), surface atmospheric pressure (p), north-south component (u), and west-east component (v) of the surface wind from Canton Island (2°48'S, 171°43'W) were edited (see Groves and Miyata 1967). These four time series, together with the two series of equilibrium tides for the period from November 1949 through October 1964, were appropriately filtered and decimated to intervals of 48 hours.

A sixth-order spectral matrix was formed from these variables in the frequency range from 0.00 cpd to 0.25 cpd, with an interval of 0.0027 cpd (93 frequency bands). Since the frequencies of interest were the lunar fortnightly tide (Mf) at 0.073 cpd and the lunar monthly tide (Mm) at 0.036 cpd, the resolution was designed so that the possible energy from the lunisolar synodic fortnightly tide (MSf) at 0.068 cpd would not contaminate the Mf frequency.

Since the effects of atmospheric tides may appear in the three weather parameters, it is most reasonable to group \( a_2^o \) and \( a_3^o \) in one set, with \( z, p, u, \) and \( v \) in the other. Intergroup coherence between the two sets was calculated and is shown in Fig. 1. It is seen that the group of sea level and the weather variables is highly coherent with equilibrium tides at Mf and Mm frequencies, but it is impossible to determine from this which member(s) of the group contributes most to this coherence.

Suppose we define “coherent power density” as

\[
C_l(f) = P_l(f)q^2(f) \cdot \frac{\beta_l(f)P_1(f)}{\sum_{k=1}^{4} \beta_k(f)P_k(f)},
\]

where \( P_l(f) \) is the original power density of \( z, p, u, \) or \( v \); \( q(f) \) is intergroup coherence; and \( \beta_l(f) \) are the coefficients defined in the previous section. Original, coherent, and original-minus-coherent power densities of \( z, p, u, \) and \( v \) are plotted in Figs. 2a, 2b, 2c, and 2d. The most prominent features are the large energy peaks at Mf and Mm frequencies of the coherent power density.
Figure 1. Intergroup coherence of $d_2^0$ and $d_3^0$ vs. $u$, $v$, and $w$. 
Figure 2a. Original (dots), coherent (solid line), and original-minus-coherent (crosses) power density for ω.
Figure 2b. Original (dots), coherent (solid line), and original-minus-coherent (crosses) power density for $\rho$. 
Figure 2c. Original (dots), coherent (solid line), and original-minus-coherent (crosses) power density for u.
Figure 3a. Multiple coherence of $z$ vs. $\rho, u, v, a_2^0$, and $a_3^0$. 
Figure 3b. Original (dots), coherent (solid line), and original-minus-coherent (crosses) power density for \( \pi \).
of sea level in Fig. 2a. None of the corresponding figures for the weather variables shows such high energy at those frequencies. Note that the peaks at Mf and Mm in the original sea-level spectrum are flattened in the original-minus-coherent spectrum whereas the peak at 0.20 cpd (weather-induced waves of five-day period; Groves and Miyata 1967) keeps the original shape. The lack of high-energy peaks at Mf and Mm frequencies in the coherent weather spectra implies that corresponding long-period atmospheric tides are negligible at Canton Island.

For the sake of comparison, similar calculations were carried out between sea level and the other five variables. The results are shown in Figs. 3a and 3b. The intergroup coherence (multiple coherence) in this case is generally higher than that in the previous case except at the Mf and Mm frequencies. This increase in coherence indicates that sea level is coherent with the weather over this frequency range, but the Mf and Mm components are dominated by the equilibrium tides. This agrees with a previous conclusion that no measurable atmospheric tides exist. In Fig. 3b there are again peaks in the coherent sea-level spectrum, but the interpretation for this grouping is not as clear as in the first case, since weather and equilibrium tides appear in a single group.

A comparison of Figs. 3a and 3b with Figs. 2a and 2b serves to demonstrate that the versatility of arbitrary grouping may lead to greater insight in certain cases of time-series analysis. Statistical reliability of this method is yet to be studied.

Acknowledgments. The author is grateful to Richard Jones, Gordon Groves, Harold Loomis, and Robert Harvey for their discussions and suggestions, to Joan Cooley for preparing the figures, and to ESSA for making available the data. This work was supported by the National Science Foundation through Grant GA-1117. The computations were carried out at the Statistical and Computing Center, University of Hawaii. This work was done while the author was staying at Hawaii Institute of Geophysics, University of Hawaii.

REFERENCES

Groves, G. W., and Motoyasu Miyata

Hotelling, Harold

Margenau, Henry, and G. M. Murphy

Munk, W. H.