Climatic Fingerprints, Patterns, and Indices.

W. C. Thacker*

Atlantic Oceanographic and Meteorological Laboratory
4301 Rickenbacker Causeway
Miami FL 33149 USA

November 30, 1995

Abstract

The purpose of this note is to point out that Hasselmann's (1993) optimal fingerprints for detecting climatic change follow from the geometrical interpretation of covariance as an inner product.

1 Introduction.

If numerical models are used to simulate "normal" and "perturbed" climatic variability, the difference between two simulations can suggest a signal that might be sought in observations. The signal can be regarded as a "pattern of climatic change", *i.e.*, values of the model's variables characterizing the simulated response to the perturbation. To look for climatic change in data, it is necessary to discern the pattern suggested by the simulations within a background of many other patterns of variability. Hasselmann (1993) pointed out that, when projecting the data onto the pattern in order to determine what part of the variability is attributable to the signal, the computation of the projection requires the use of a second pattern, the "fingerprint", which is obtained by multiplying the original pattern by the inverse of the covariance matrix that characterizes the usual variability of the data.

Projecting data onto a pattern is an intrinsically geometrical concept. The geometry is fixed by the statistical relationships between the variables

* thacker@aoml.erl.gov

used to characterize the climatic state. Because these variables are correlated, the basis vectors used in representing the patterns are not orthogonal. Consequently, projection necessarily involves the adjoint or dual (Lanczos, 1988) of the pattern vector, *viz.*, its fingerprint. The fingerprint is a vector of coefficients defining a temporally varying climatic index[1] as a linear combination of the observed variables. The amplitude of this index measures the strength of the climatic pattern.

The geometrical view of statistics presented here, while not new and certainly recognized by Hasselmann (1993), is not widely appreciated and is often misunderstood. In particular, when observations of the climatic state are presented as a column vector, there is the tendency to confuse the dot product (matrix multiplication of one column vector by the transpose of another) with the inner product appropriate for projection. Because the geometry of the patterns is dictated by statistical relationships among random variables, the inner product of two vectors represents the covariance of two variables. This definition of the inner product distinguishes between the vector space of patterns and the dual space of fingerprints.

2 Geometry of random variables.

Suppose the collection of variables X_i, $i = 1, \ldots, i_{\max}$, are to be examined for climatic change[2]. For example, they might represent surface temperature, pressure, and humidity at a large number of stations distributed around the earth. These variables should be regarded as random variables with possible values distributed about a climatic mean state according to a probability distribution inferred from past observations. In general, pairs of these variables may be correlated, some pairs more than others. Accounting for these correlations is essential, when judging whether a pattern among the many variables is sufficiently prominent to qualify as evidence of climatic change.

To understand the role of correlations, it is useful to regard random variables as vectors in an abstract metric space. The inner product $\langle X_i, X_j \rangle$ of vectors X_i and X_j is the covariance of the two variables:

$$\langle X_i, X_j \rangle = \mathrm{cov}(X_i, X_j) = \Sigma_{i,j} \ . \tag{1}$$

[1] Hasselmann (1993) uses the term, "detector."

[2] The subscript i, like Hasselmann's (1993) subscript a, distinguishes between variable types, locations, and temporal lags.

This definition of inner product is easily motivated by considering how sample covariances might be computed: if time series of temperature and humidity are available, each variable's deviation from its mean could be represented as a column vector, and the sample covariance would involve the dot product of the two column vectors. All of the information about the relative orientation of the vectors is contained in the covariance matrix $\mathbf{\Sigma}$, which is estimated by the sample covariance matrix $\mathbf{S}$ (Johnson and Wichern, 1992).

The inner product of a vector with itself is the variable's variance. A vector's length is its standard deviation:

$$|X_i| = \langle X_i, X_i \rangle^{1/2} = \sqrt{\mathrm{var}(X_i)} \tag{2}$$

Thus, a standardized variable, *i.e.*, a variable measured in units of its standard deviation about its mean, is a unit vector. The angle $\theta_{i,j}$ between vectors, X_i and X_j can be computed in the usual way:

$$\cos\theta_{i,j} = \frac{\langle X_i, X_j \rangle}{|X_i||X_j|} = \mathrm{cor}(X_i, X_j)\ . \tag{3}$$

In other words, the cosine of the angle between two vectors is the correlation between the two variables. Two vectors are orthogonal when the variables are uncorrelated. The relative orientation of the vectors are the geometric manifestation of the correlations among the variables.

Because the variables are correlated, the vectors cannot be expected to be mutually orthogonal. In fact, there is the possibility that they might be linearly dependent, *i.e.*, that some can be represented as a linear combination of others, or equivalently, that the value of some variables can be reliably estimated from the values of other variables. In practice, the true covariances are not known, so it is necessary to work with estimates based on finite samples, and the number of vectors that can be judged to be linearly independent is never greater than the number of realizations in the sample. To avoid biasing our view toward the idiosyncrasies of the particular record, it is wise to choose the smallest possible number of variables, consistent with the requirement that they explain the bulk of the variability of the remainder. Principal components are an attractive choice, because sample correlations suggest that they are orthogonal. A subset of the original variables can also be used. Below it is assumed that the vectors X_i are not linearly dependent, so $\mathbf{\Sigma}$ is not singular.

3 Patterns, fingerprints, and climatic indices.

A climatic pattern is a relationship that might be observed among the variables. For example, the observed values $x_i(t)$ at time t constitute a pattern, as do the averages of the values observed over an interval of time. A pattern can be represented as a column vector β with ith element being the relative amount of the ith variable in the pattern. The set of all possible patterns determines a vector space spanned by the unit column vectors $(1\ 0 \ldots 0)^T$, $(0\ 1 \ldots 0)^T$, *etc.* There is a natural tendency to think of these unit vectors as being orthogonal and to compute the inner product of two patterns as the dot product of the corresponding column vectors, but this would be wrong. The geometry of the patterns is determined by the geometry of the random variables.

A fingerprint vector α is a column vector of coefficients that define a random variable Γ as a linear combination of the random variables X_i. If the random variables are collectively denoted by the column vector X, then

$$\Gamma = \alpha^T X \ . \tag{4}$$

The coefficient vectors are easy to confuse with the pattern vectors. The easiest way to distinguish them is by their units: the elements of β should have the same units as the corresponding variable, while the units of the elements of α should be their inverses so that the linear combination (4) is dimensionally homogeneous. It is clear that there is a one-to-one correspondence between the patterns and the fingerprints. In fact, (see equation 8 below) the fingerprints are duals or adjoints[3] of the patterns.

The observed value

$$\gamma(t) = \alpha^T x(t) \tag{5}$$

of the random variable Γ at time t is an index measuring the strength of the pattern β having dual α. Thus, the three entities — fingerprint α, pattern β, and index γ — are intimately related.

There is a one-to-one relationship between all possible coefficient vectors and all-possible random-variable vectors: each possible Γ is a vector in the random-variable space, and conversely, each random variable in the space is a linear combination of the X_i specified by some α. The inner product between

[3] The concept of adjoint variables has recently been exploited by Farrell (1989), among others, within the context of "fastest growing meteorological disturbances."

coefficient vectors α_1 and α_2 should be the covariance of the corresponding random variables Γ_1 and Γ_2:

$$\langle \alpha_1, \alpha_2 \rangle = \mathrm{cov}(\Gamma_1, \Gamma_2) = \alpha_1^T \mathbf{\Sigma} \alpha_2 \ . \tag{6}$$

Note that the inner product is *not* the dot product $\alpha_1^T \alpha_2$; in fact, if the units of the coefficients are chosen so that the linear combinations Γ_1 and Γ_2 are dimensionally homogeneous, *i.e.*, (temperature)$^{-1}$ and (humidity)$^{-1}$, then $\alpha_1^T \alpha_2$ is not dimensionally homogeneous. With this inner product, the basis vectors $(1\,0 \ldots 0)^T$, $(0\,1 \ldots 0)^T$, *etc.*. are not orthogonal; the angles between them are given by the correlations between the corresponding variables.

To understand why (6) is the correct inner product, consider the linear algebra involved in computing sample covariances. The vector X of random variables becomes the data matrix $\mathbf{X}$, each row of which is a time series of observations of one of the climatic variables, and $\alpha^T \mathbf{X} = g^T$ is the row vector containing the time series of the index γ. Thus, g is in the row space of $\mathbf{X}$, and α is in the column space. The sample covariance is:

$$\frac{1}{n} g_1^T g_2 = \frac{1}{n} \alpha^T \mathbf{X} \mathbf{X}^T \alpha = \alpha^T \mathbf{S} \alpha \ , \tag{7}$$

where $\mathbf{S}$ is the sample covariance matrix of the original climatic variables. The geometry of the row space is reflected in the column space, with orthogonality of time series dictating the inner product of the fingerprints.

When using a non-orthogonal basis, the concept of adjoint (or dual) vectors (Lanczos, 1988) is important. The relationship between a vector and its dual is via the metric defining the inner product. In equation (6) the covariance matrix $\mathbf{\Sigma}$ is the metric[4]; it relates the fingerprint α to a pattern β:

$$\beta = \mathbf{\Sigma} \alpha \ . \tag{8}$$

Note that $\beta = \mathrm{cov}(X, \Gamma)$; *i.e.*, it is the pattern of covariances of the individual variables X_i and the variable Γ. Because of the duality relationship, the inner product for the patterns is also the covariance of the corresponding random variables:

$$\langle \beta_1, \beta_2 \rangle = \beta_1^T \mathbf{\Sigma}^{-1} \beta_2 = \mathrm{cov}(\Gamma_1, \Gamma_2) \ , \tag{9}$$

as is the dot product of a fingerprint and pattern,

$$\alpha_1^T \beta_2 = \mathrm{cov}(\Gamma_1, \Gamma_2) \ . \tag{10}$$

[4]It is the metric used to define the Mahalanobis distance, which underlies Hotelling's T^2 test (Mardia *et al.*, 1979).

Because the fingerprints α_i and α_j defining uncorrelated variables Γ_i and Γ_j are bi-orthogonal to the corresponding patterns β_i and β_j,

$$\alpha_i^T \beta_j = 0 \ , \ i \neq j \ , \tag{11}$$

fingerprints are used to project patterns onto data. Such projections are simplest when the fingerprints are chosen to be unit vectors, *i.e.*, so that the Γ's have unit variance and $\alpha_i^T \beta_j = \cos \theta_{i,j}$.

Observations $x(t)$ at time t can be thought of as a pattern β of hypothesized climatic change of strength $\gamma(t)$ in a background $\epsilon(t)$ of other patterns:

$$x(t) = \gamma(t)\beta + \epsilon(t) \ . \tag{12}$$

The fact that the fingerprint α can be used to project the data onto the pattern reflects the fact that the background corresponds to variables uncorrelated with Γ, *i.e.*, $\alpha^T \epsilon(t) = 0$.

The index $\gamma(t) = \alpha^T x(t)$ measures the strength of the pattern β in the observations at time t. With α and β normalized so that $\mathrm{var}(\Gamma) = 1$, $|\gamma(t)| \gg 1$ suggests a climatic change.

4 A related problem.

A closely related problem is that of determining the amplitude of a pattern defined by covariances with a specified signal, *e.g.*, the sun-spot cycle. If the intensity of sun spots is denoted by the random variable Y, then $\beta = \mathrm{cov}(XY)$, and the variable $\Gamma = \alpha^T X$ should be the projection of Y onto the space spanned by the X_i's. In other words Γ should be maximally correlated with Y. The question is how to define α. The answer can be found variationally by maximizing the correlation coefficient:

$$\rho = \mathrm{cor}(Y, \Gamma) = \frac{\mathrm{cov}(Y, \Gamma)}{|Y||\Gamma|} = \frac{\alpha^T \beta}{|Y|(\alpha^T \Sigma \alpha)^{\frac{1}{2}}} \ . \tag{13}$$

Differentiating with respect to α gives the relationship between fingerprint and pattern:

$$\beta = \rho_{\max} \left(\frac{\mathrm{var} Y}{\mathrm{var} \Gamma} \right)^{\frac{1}{2}} \Sigma \alpha \ . \tag{14}$$

Even if Y and Γ are standardized to unit variance, $\beta \neq \Sigma\alpha$ as in (8); there is an extra factor of the canonical correlation $\rho_{\max} = \cos \theta_{Y,X}$ measuring the angle between Y and the space spanned by the X_i's.

5 Conclusion.

Correlations between climatic variables govern the projection of observations onto climatic patterns. When patterns are thought of as vectors, the appropriate inner product is not the usual dot product; instead it is computed with the covariance matrix of the climatic variables serving as a metric. This inner product distinguishes between the pattern and its dual, the fingerprint, which is used for projecting the data onto the pattern. The simple dot product of data and fingerprint is equivalent to the inner product of data and pattern; it defines an index measuring the strength of the pattern in the data. Thus, the three entities — fingerprint, index, and pattern — mutually reflect the correlations that exist between the climatic variables.

References

Farrell, B. F. (1989). Optimal excitation of baroclinic waves. *Journal of the Atmospheric Sciences*, **46**, 1193–1206.

Hasselmann, K. (1993). Optimal fingerprints for the detection of time-dependent climate change. *Journal of Climate*, **6**, 1957–1971.

Johnson, R. A., and D. W. Wichern (1992). *Applied Multivariate Statistical Analysis.* Prentice Hall, Englewood Cliffs NJ.

Lanczos, C. (1988). *Applied Analysis.* Dover, New York.

Mardia, K. V., J. T. Kent, and J. M. Bibby (1979). *Multivariate Analysis.* Academic Press, Reading MA.