5. Karhunen-Loève Theorem and Singular Spectrum Analysis

The $N \times N$, $\mathbf{R}$ matrix in Eq. (1.16) is square and symmetric. It is an important result of linear algebra that such matrices have an orthogonal decomposition

$$\mathbf{R} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^T \tag{5.1}$$

where $\boldsymbol{\Lambda}$ is a diagonal matrix, $diag\{\lambda_i\}$ of the eigenvalues of $\mathbf{R}$ and $\mathbf{V}$ is the matrix of eigenvectors $\mathbf{V} = \{\mathbf{v}_i\}$, such that

$$\mathbf{R}\mathbf{v}_i = \lambda_i \mathbf{v}_i, \tag{5.2}$$

and they are orthonormal $\mathbf{v}_i^T \mathbf{v}_j = \delta_{ij}$. (It follows that $\mathbf{V}^{-1} = \mathbf{V}^T$.)

Let us write a time-series as an expansion in the $\mathbf{v}_q$ in the form

$$x_n = \text{element } n \text{ of } \left[\sum_{q=1}^{N} \alpha_q \sqrt{\lambda_q} \mathbf{v}_q\right] \tag{5.3}$$

or more succinctly, if we regard x_n as an $N-$element vector, $\mathbf{x}$,

$$\mathbf{x} = \left[\sum_{q=1}^{N} \alpha_q \sqrt{\lambda_q} \mathbf{v}_q\right]. \tag{5.4}$$

Here a_q are unit variance, uncorrelated random variates, e.g., $G(0,1)$. We assert that such a time-series has covariance matrix $\mathbf{R}$, and therefore must have the corresponding (by the Wiener-Khinchin Theorem) power density spectrum. Consider

$$\begin{aligned} R_{ij} = < x_i x_j > &= \sum_{q=1}^{N} \sum_{r=1}^{N} < \alpha_q \alpha_r > \sqrt{\lambda_q \lambda_r} v_{iq} v_{jr} \\ &= \sum_{q=1}^{N} \lambda_q v_{iq} v_{jq} \end{aligned} \tag{5.5}$$

by the covariance properties of α_q. But this last equation is just

$$R_{ij} = \left\{\mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^T\right\}_{ij} \tag{5.6}$$

which is what is required.

Exercise, Confirm (5.6).

Thus (5.4) gives us another way of synthesizing a time series from its known covariance. That the decomposition (5.4) can always be constructed (in continuous time) for a stationary time series is called the Karhunen-Loève Theorem (see Davenport and Root, 1958). Because it is based upon a decomposition of the covariance matrix, it is evidently a form of empirical orthogonal function synthesis, or if one prefers, an expansion in principal components (see e. g., Jolliffe, 1986).

The relative importance of any orthogonal structure, $\mathbf{v}_i$, to the structure of x_n, is controlled by the magnitude of $\sqrt{\lambda_i}$. Suppose one has been successful in obtaining a physical interpretation of one or more of the important $\mathbf{v}_i$. Each of these vectors is itself a time series. One can evidently compute a power density spectrum for them, either by Fourier or more exotic methods. The idea is that the spectra of the dominant $\mathbf{v}_i$ are meant to be informative about the spectral properties of processes underlying the full time series. This hypothesis is a plausible one, but will only be as valid as the reality of the underlying physical structure attributed to $\mathbf{v}_i$. (At least two issues exist, determining when λ_i is significantly different from zero, and obtaining a physical interpretation of the components. Principal components are notoriously difficult to relate to normal modes and other physically generated structures.) This subject has come to be known as "singular spectrum analysis" (e.g. Vautard and Ghil, 1989) and its powers (and statistical information such as confidence limits) are still murky.