On the Statistics of Correlated Time Series

by Bill Menke in discussion with Jim Davis

April 9, 2025

Suppose that one has data $\mathbf{d}$ of length N and covariance $\mathbf{C}_d$. One can transform into a new set of variables $\mathbf{d}' = \mathbf{T}\mathbf{d}$, also of length N, that is uncorrelated and with uniform variance $\mathbf{C}_{d'} = \mathbf{I}$. The transformation $\mathbf{T}$ can be defined in a variety of ways. The choice $\mathbf{T} = \mathbf{C}_d^{-½}$ (where this is the symmetric square root) works, because by the usual rule of error propagation $\mathbf{C}_{d'} = \mathbf{T}\mathbf{C}_h\mathbf{T}^T = \mathbf{C}_d^{-½}\mathbf{C}_d\mathbf{C}_d^{-½T} = \mathbf{I}$. The choice $\mathbf{T} = \mathbf{\Lambda}^{-½}\mathbf{V}^T$ also works, where $\mathbf{\Lambda}$ and $\mathbf{V}$ are the eigenvalues and eigenvectors, respectively, of $\mathbf{C}_d$ because $\mathbf{C}_{d'} = \mathbf{T}\mathbf{C}_d\mathbf{T}^T = [\mathbf{\Lambda}^{-½}\mathbf{V}^T][\mathbf{V}\mathbf{\Lambda}\mathbf{V}^T][\mathbf{V}\mathbf{\Lambda}^{-½}] = \mathbf{I}$. Here, we have made use of the orthonormality of the eigenvectors; that is, $\mathbf{V}^T\mathbf{V} = \mathbf{V}\mathbf{V} = \mathbf{I}$. In this discussion, we focus on the second transformation.

The inverse transformation is $\mathbf{d} = \mathbf{T}^{-1}\mathbf{d}'$, where $\mathbf{T}^{-1} = \mathbf{C}_d^{½}$ or $\mathbf{T}^{-1} = \mathbf{V}\mathbf{\Lambda}^{½}$, depending on whether the first or second version of $\mathbf{T}$ is used. This formula provides a simple method of creating a realization of $\mathbf{d}$. One simply generates the uncorrelated, unit variance vector $\mathbf{d}'$ using a random number generator, and then transforms it to $\mathbf{d}$.

Only some of the eigenvalues of the covariance $\mathbf{C}_d$ are numerically large, representing the fact that a correlated timeseries has fewer degrees of freedom than an uncorrelated one. One could ask how $\mathbf{d}$ can be represented given a $\mathbf{d}'$ of length $N' < N$. We start partitioning the transformation matrix and the transformed data

$$\mathbf{T}^{-1} \equiv \left[\mathbf{V}_0, \mathbf{V}_p\right]\begin{bmatrix}\mathbf{\Lambda}_0^{½} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_p^{½}\end{bmatrix} = \left[\mathbf{V}_0\mathbf{\Lambda}_0^{½}, \mathbf{V}_p\mathbf{\Lambda}_\mathrm{p}^{½}\right] = \left[\mathbf{T}_0^{-1}, \mathbf{T}_p^{-1}\right] \quad \text{and} \quad \mathbf{d}' = \begin{bmatrix}\mathbf{d}_0' \\ \mathbf{d}_p'\end{bmatrix}$$

Here, the subscript p denotes the N' largest eigenvalue and the subscript 0 indicates the smaller ones. We now solve the equation $\mathbf{d} = \mathbf{T}_p^{-1}\mathbf{d}_p'$ by least squares:

$$\mathbf{d}_p' = \left[\mathbf{T}_p^{-1T}\mathbf{T}_p^{-1}\right]^{-1}\mathbf{T}_p^{-1T}\mathbf{d} = \left[\mathbf{\Lambda}_\mathrm{p}^{½}\mathbf{V}_p^T\mathbf{V}_p\mathbf{\Lambda}_\mathrm{p}^{½}\right]^{-1}\mathbf{\Lambda}_\mathrm{p}^{½}\mathbf{V}_p^T\mathbf{d} = \mathbf{\Lambda}_\mathrm{p}^{-½}\mathbf{V}_p^T\mathbf{d}$$

The least squares solution is $\mathbf{d}_p' = \mathbf{T}_p\mathbf{d}_p$ with $\mathbf{T}_p \equiv \mathbf{\Lambda}_\mathrm{p}^{-½}\mathbf{V}_p^T$, which is to say, the same result as one would achieve by simply setting $\mathbf{d}_0' = 0$. This behavior occurs in least squares problems whenever the columns of the data kernel are orthogonal, which is the case here for the data kernel $\mathbf{T}^{-1}$.

In the untransformed domain, the time series predicted by $\mathbf{d}_p'$ is

$$\mathbf{d}_p \equiv \mathbf{T}^{-1}\begin{bmatrix}\mathbf{0} \\ \mathbf{d}_p'\end{bmatrix} = \left[\mathbf{V}_0\mathbf{\Lambda}_0^{½}, \mathbf{V}_p\mathbf{\Lambda}_\mathrm{p}^{½}\right]\begin{bmatrix}\mathbf{0} \\ \mathbf{\Lambda}_\mathrm{p}^{-½}\mathbf{V}_p^T\mathbf{d}\end{bmatrix}$$

$$= \mathbf{V}_p\mathbf{\Lambda}_\mathrm{p}^{½}\mathbf{\Lambda}_\mathrm{p}^{-½}\mathbf{V}_p^T\mathbf{d}_p = \mathbf{V}_p\mathbf{V}_p^T\mathbf{d}_p \equiv \mathbf{R}\mathbf{d}$$

Here, $\mathbf{R} \equiv \mathbf{V}_p\mathbf{V}_p^T$ is a resolution matrix. Note as $N' \to N$ is $\mathbf{V}_p \to \mathbf{V}$ and $\mathbf{R} \to \mathbf{I}$.

The error E between $\mathbf{d}_p$ and $\mathbf{d}$ is

$$E = \left[\mathbf{d} - \mathbf{d}_p\right]^T\left[\mathbf{d} - \mathbf{d}_p\right] = [(\mathbf{I} - \mathbf{R})\mathbf{d}]^T(\mathbf{I} - \mathbf{R})\mathbf{d}$$

Note that $\mathbf{I} = \mathbf{V}_0\mathbf{V}_0^T + \mathbf{V}_p\mathbf{V}_p^T$, so that $(\mathbf{I} - \mathbf{R}) = \mathbf{V}_0\mathbf{V}_0^T$. Then,

$$E = [\mathbf{V}_0\mathbf{V}_0^T\mathbf{d}]^T[\mathbf{V}_0\mathbf{V}_0^T\mathbf{d}] = \mathbf{d}^T\mathbf{V}_0\mathbf{V}_0^T\mathbf{V}_0\mathbf{V}_0^T\mathbf{d} = \mathbf{d}^T\mathbf{V}_0\mathbf{V}_0^T\mathbf{d}^T = [\mathbf{V}_0^T\mathbf{d}]^T[\mathbf{V}_0^T\mathbf{d}]$$

$$= \left[\mathbf{\Lambda}_0^{½}\mathbf{d}_0'\right]^T\left[\mathbf{\Lambda}_0^{½}\mathbf{d}_0'\right] = \mathbf{d}_0'^T\mathbf{\Lambda}_0\mathbf{d}_0'$$

Let the largest eigenvalue in $\mathbf{\Lambda}_0$ be λ_0^{max}. The vector $\mathbf{d}_0'$ is of length $K = (N - N')$ and has elements with zero mean and unit variance, so the quantity $\mathbf{d}_0'^T\mathbf{d}_0'$ is a chi-squared distributed random variable with mean K. Consequently,

$$E = \mathbf{d}_0'^T\mathbf{\Lambda}_0\mathbf{d}_0' < \lambda_0^{max}\mathbf{d}_0'^T\mathbf{d}_0' \approx K\lambda_0^{max}$$

Presuming that the eigenvalues are ordered by increasing size, E decreases monotonically as $N' \to N$, because both λ_0^{max} and K decrease with increasing N'.

The weighted energy $\Phi = \mathbf{d}^T\mathbf{C}_d^{-1}\mathbf{d}$ is invariant under the transformation $\mathbf{T} = \mathbf{\Lambda}^{-½}\mathbf{V}^T$

$$\Phi = \mathbf{d}^T\mathbf{C}_d^{-1}\mathbf{d} = \mathbf{d}^T\mathbf{V}\mathbf{\Lambda}^{-1}\mathbf{V}^T\mathbf{d} = [\mathbf{\Lambda}^{-½}\mathbf{V}^T\mathbf{d}]^T[\mathbf{\Lambda}^{-½}\mathbf{V}^T\mathbf{d}] = \mathbf{d}'^T\mathbf{d}' = \Phi'$$

As $\mathbf{d}'$is of length N with elements that are uncorrelated and of unit variance, Φ' is chi-squared distributed with N degrees of freedom. As $\Phi = \Phi'$, the untransformed quantity Φ is chi-squared distributed with N degrees of freedom, too.

Similarly, the weighted energy

$$\Phi_p = \mathbf{d}_p^T\mathbf{V}_p\mathbf{\Lambda}_p^{-1}\mathbf{V}_p^T\mathbf{d}_p = \left[\mathbf{\Lambda}_p^{-½}\mathbf{V}_p^T\mathbf{d}_p\right]^T\left[\mathbf{\Lambda}_p^{-½}\mathbf{V}_p^T\mathbf{d}_p\right] = \mathbf{d}_p'^T\mathbf{d}_p' = \Phi_p'$$

is invariant under the transformation $\mathbf{T}_p = \mathbf{\Lambda}_p^{-½}\mathbf{V}_p^T$. As $\mathbf{d}_p'$ is a length N' time series with elements that are uncorrelated and of unit variance, $\Phi_p = \Phi_p'$ is chi-squared distributed with N' degrees of freedom.

The elements of $\mathbf{d}_p$ are not linearly independent. A total of $K = (N - N')$ linear combinations are prescribed by the condition $\mathbf{V}_0^T\mathbf{d}_p = 0$. This condition can be demonstrated by multiplying the equation $\mathbf{d}_p = \mathbf{V}_p\mathbf{\Lambda}_p^{½}\mathbf{d}_p'$ by $\mathbf{V}_0^T$ to yields $\mathbf{V}_0^T\mathbf{d}_p = \mathbf{V}_0^T\mathbf{V}_p\mathbf{\Lambda}_p^{½}\mathbf{d}_p' = 0$ (as $\mathbf{V}_0$ and $\mathbf{V}_p$ are orthogonal so $\mathbf{V}_0^T\mathbf{V}_p = 0$). In contrast, the elements of $\mathbf{d}$ are linearly independent, for this is just the limiting case of $N' = N$ (in which case $K = 0$). An exception is when $\mathbf{C}_d$ has, say, $L > 0$ identically-zero eigenvalues (in which case, the transformation $\mathbf{T}$ does not exist, but $\mathbf{T}_p$ does exist). These eigenvalues must be included in $\mathbf{\Lambda}_0$, in which case $K = L$.

Consider the special case where $\mathbf{C}_d$ has unit variance and a correlation length of s, meaning that $J \approx 2s$ neighboring points in the time series are approximately equal. Then, one would expect that the time series $\tilde{\mathbf{d}}$ of length $M \approx N/J$ and consisting of every Jth element of $\mathbf{d}$ would be approximately uncorrelated. Consequently, the unweighted energy $\tilde{\Psi} = \tilde{\mathbf{d}}^T\tilde{\mathbf{d}}$ is chi-squared distributed with M degrees of freedom. The unweighted energy $\Psi = \mathbf{d}^T\mathbf{d}$ of the original time series $\mathbf{d}$ is J times larger. Consequently, Ψ/J is approximately chi-squared distributed with M degrees of freedom.

As an example, we consider a time series of length $N = 100$ and exponential covariance (Figure 1)

$$[\mathbf{C}_d]_{ij} = \sigma_d^2 \exp\{-|i-j|/s\}$$

Here, $\sigma_d^2 = 1$ is a variance and $s = 5$ is a scale factor that governs the degree of correlation of neighboring points in the time series.

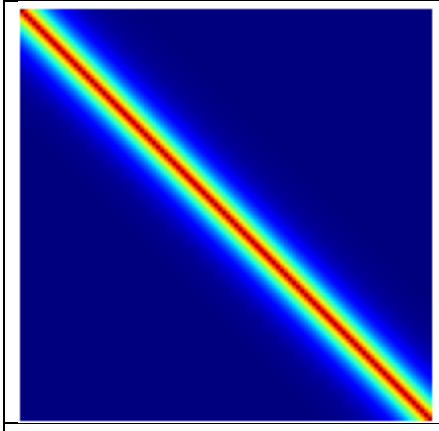

Fig. 1. Exemplary covariance matrix $\mathbf{C}_d$.

Only about 20 of the 100 eigenvalues of this covariance matrix are numerically large (Figure 2).

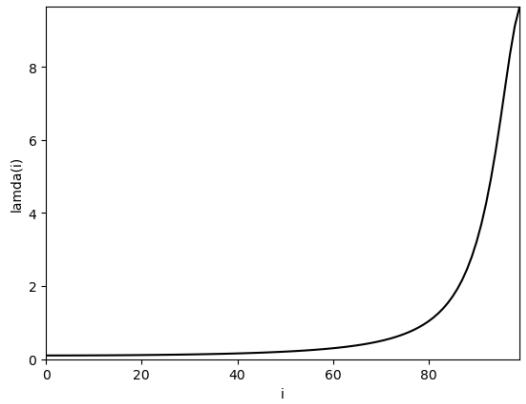

Fig. 2. Eigevalues λ_i of the exemplary covariance matrix $\mathbf{C}_d$.

A realization of this time series is created by randomly generating $\mathbf{d}'$ and then transforming to $\mathbf{d}$ (Figure 3, black curve). The approximation $\mathbf{d}_p$ then can be generated by for an arbitrary value of N' (Figure 3, red curves). The larger the value of N', the better the approximation.

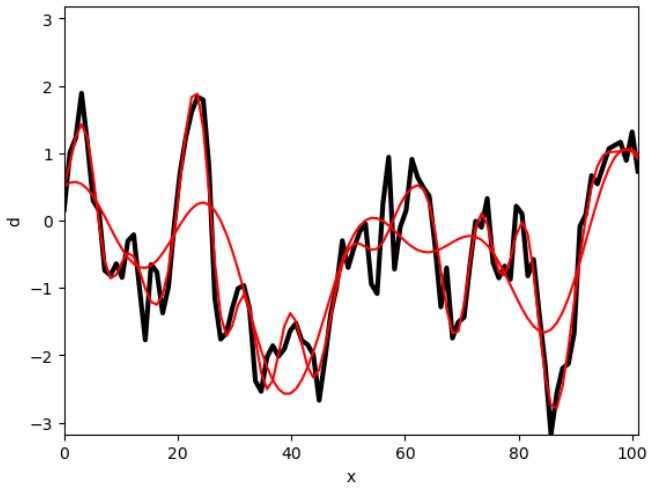

Fig. 3. A realization of the time series $\mathbf{d}$ (black curve) and two approximations $\mathbf{d}_p$ of it (red curves) for $N' = 10$ (smoother curve) and $N' = 30$ (rougher curve).

The empirical p.d.f. of the weighted energy Φ_p (for $N' = 30$) is generated by binning values derived from 10,000 randomly-generated realizations of $\mathbf{d}_p$. It compares well with the predicted chi-squared p.d.f. (Figure 4).

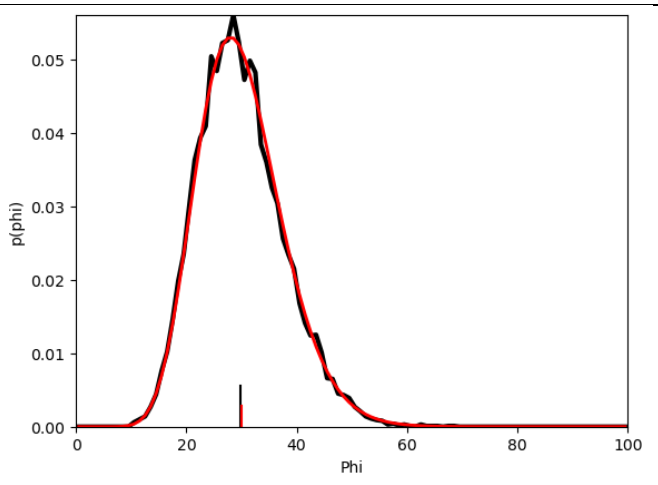

Fig. 4. Empirical p.d.f. (black curve) of the weighted energy Φ_p for 10,000 realizations of the time series $\mathbf{d}_p$ (with $N' = 30$). It agrees well with the chi-squared p.d.f. $p(\chi^2_{N'})$

(red curve). Vertical tick marks are the mean of the empirical (black) and chi-squared (red) p.d.f.s, respectively.

The empirical p.d.f. of the unweighted energy Ψ is generated by binning 10,000 values derived from randomly-generated realizations of $\mathbf{d}$. It compares fairly well with the predicted chi-squared p.d.f. (Figure 5). The means of the two p.d.f.s match well, but the empirical p.d.f. is somewhat narrower than the chi-squared p.d.f.

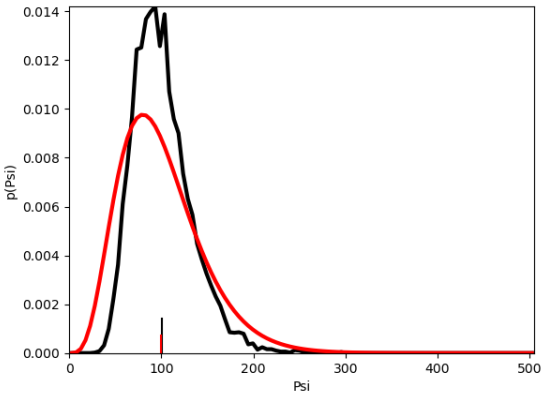

Fig. 5. Empirical p.d.f. (black curve) of the unweighted energy Ψ for 10,000 realizations of the time series $\mathbf{d}$ and the scaled chi-squared p.d.f. (red curve) described in the text. Vertical tick marks are the mean of the empirical (black) and chi-squared (red) p.d.f.s, respectively.

The most important part of the Pythin code is

```
# x-axis
Nx = 100;
xmin = 0.0;
xmax = 1.0*Nx;
Dx = (xmax-xmin)/(Nx-1);
x = gda_cvec( np.linspace( xmin, xmax, Nx ) );

# number of data and their x-positions
Nd = Nx;
xd = gda_cvec(np.linspace(xmin,xmax,Nd));

# covariance of data
Cd = np.zeros( (Nd,Nd) );
g2 = 1.0;  # diagonal
sx = 5.0;  # scale factor
for i in range(Nd):
    for j in range(i,Nd):
        Cd[i,j] = g2*exp( - np.abs(xd[i,0]-xd[j,0]) / sx );
        Cd[j,i] = Cd[i,j];

# eigenvale decomposition of covariance
lam, v = la.eigh(Cd);
lammax = np.max(lam);
```

```
# transformations for d = Tp dp   and   dp = T d
Tp = np.matmul( v, np.diag(np.sqrt(lam)) );

# primed data, uncorrelated with unit variance
dp = np.random.normal( loc=0.0, scale=1.0, size=(Nd,1) );

# transform to unprimed data
d = np.matmul( Tp, dp );
dabsmax = np.max( np.abs(d) );

# loop over reducion in the number of primed data
Ndnew = 30;
dpnew = dp[Nd-Ndnew:Nd,0:1];

# reduce size of eigen-matices
lamnew = np.copy( lam[Nd-Ndnew:Nd] );
vnew = np.copy( v[0:Nd,Nd-Ndnew:Nd] );

# now build transformation that operates on Ndnew data
# dnew = Tpnew dpnew = sqrt(Cd) dpnew = vnew * sqrt(lamnew)
Tpnew = np.matmul( vnew, np.diag(np.sqrt(lamnew)) );
dnew = np.matmul( Tpnew, dpnew ) ;
```