Lecture 11: White and red noise

©Christopher S. Bretherton

Winter 2015

Reference: Hartmann Atm S 552 notes, Chapter 6.1-2.

11.1 White noise

A common way to statistically assess the significance of a broad spectral peak as in the Nino3.4 example is to compare with a simple noise process. **White noise** has zero mean, constant variance, and is uncorrelated in time. As its name suggests, white noise has a power spectrum which is uniformly spread across all allowable frequencies.

In Matlab, **w = randn(N)** generates a sequence of length N of $n(0,1)$ 'Gaussian' white noise (i.e. with a normal distribution of mean 0 and std 1). The upper two panels of Fig. 1 show a white noise sequence of length $N = 128$ and its periodogram, which shows that the power spectrum is uniformly spread across frequencies with a mean spectral power of $1/N$ per harmonic.

11.2 Red noise

Red noise has zero mean, constant variance, and is serially correlated in time, such that the lag-1 autocorrelation between two successive time samples has correlation coefficient $0 < r < 1$. As we will show shortly, red noise has a power spectrum weighted toward low frequencies, but has no single preferred period. Its 'redness' depends on r, which can be tuned to match the observed time series. For the Nino3.4 case, a reasonable statistical null hypothesis would be that the observed power spectrum could have been generated purely by red noise.

To sequentially generate a $n(0,1)$ red noise sequence x_j from a white noise sequence w_j, we set

$$\begin{aligned} x_1 &= w_1 \\ x_{j+1} &= r x_j + (1-r^2)^{1/2} w_{j+1}, \; j \geq 1 \end{aligned} \tag{11.2.1}$$

Using properties of normal distributions, it is easily shown that x_{j+1} is $n(0,1)$ (Gaussian) and that the lag-1 correlation coefficient of x_{j+1} and x_j is r. It is also easy to show by induction that the correlation coefficient of x_{j+p} and

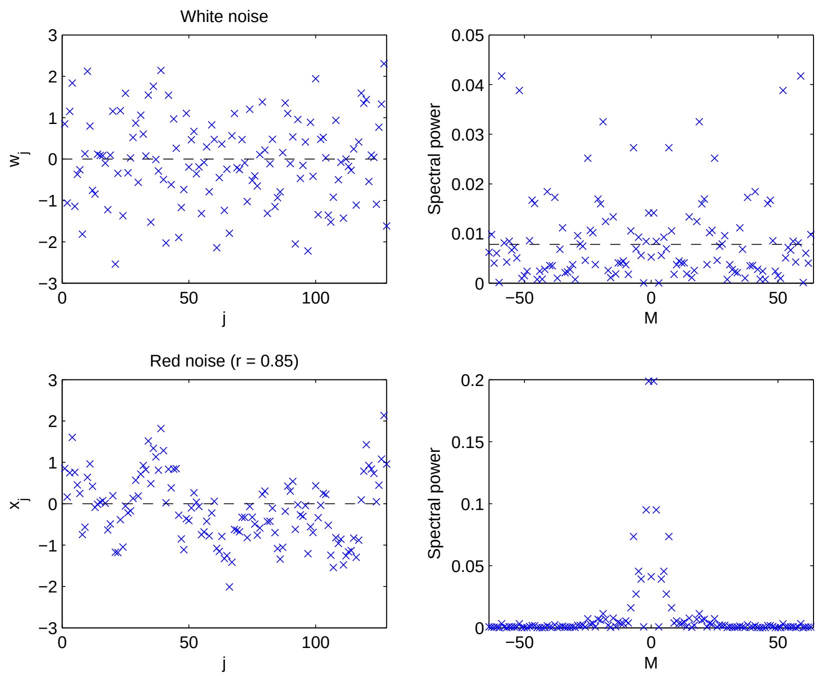

Figure 1: White and red noise time series (left) and their periodograms (right)

x_j for $p > 1$ is $r^p = \exp(-p \log r) = \exp(-Rp\Delta t)$, where $R = -\log r/\Delta t$ is the decorrelation rate. The autocovariance sequence of red noise thus decays exponentially with lag. The lag at which the autocorrelation drops to $1/e$ is $\tau = R^{-1}$.

The function **rednoise.m** (class web page) implements this algorithm, It was used with the white noise sequence on the upper left of Fig. 1 and $r = 0.85$ to generate the red noise time series on the lower left. The periodogram of this sequence, shown in the lower right, now has a predominance of spectral power in low harmonics M.

11.3 Theoretical power spectrum of red noise

The true power spectrum of $n(0, 1)$ red noise is most easily deduced as the DFT of its autocovariance sequence. Rather than grinding through discrete sums, it is more helpful to interpret the DFT as a Riemann sum that approximates the continuous integral for the complex Fourier coefficients of the continuous function $a(t) = e^{-R|t|}$, L-periodically extended for $|t| > L/2$):

$$\begin{aligned} S_m &= N^{-1}\mathrm{DFT}(\mathbf{a}) \\ &\approx c_M[a(t)]] = L^{-1}\int_{-L/2}^{L/2} e^{-R|t|-i\omega_M t}dt \end{aligned}$$

So far, the approximation is good if $R\Delta t \ll 1$ and $\omega_M^{-1}\Delta t \ll 1$, so that $a(t)\exp(-i\omega_m t)$ is well resolved by the grid of spacing Δt. If in addition $RL/2 \gg 1$, the integrand becomes very small for $|t| > L/2$. Then, with negligible error we can extend the range of integration to infinity:

$$\begin{aligned} S_m &\approx L^{-1}\int_{-\infty}^{\infty} e^{-R|t|-i\omega_M t}dt \\ &= L^{-1}\left\{\int_{-\infty}^{0} e^{t[R-i\omega_M]}dt + \int_{0}^{\infty} e^{-t[R+i\omega_M]}\right\} \\ &= L^{-1}\left\{\frac{1}{R-i\omega_M} + \frac{1}{R+i\omega_M}\right\} \\ &= L^{-1}\frac{2R}{R^2+\omega_M^2}. \\ &= \frac{\Delta\omega}{\pi}\frac{R}{R^2+\omega_M^2} \qquad (\Delta\omega = \frac{2\pi}{L}) \end{aligned} \tag{11.3.1}$$

The power spectrum of red noise has a maximum value for low frequencies $\omega_M \ll R$, and decreases at high frequencies - a 'red' spectrum, as claimed. If we sum the power spectrum across the harmonics, and think of it as a Riemann sum approximation to a continuous integral

$$\begin{aligned} \sum_{m=1}^{N} S_m &\approx \sum_{M=-N/2}^{N/2-1} \frac{\Delta\omega}{\pi}\frac{R}{R^2+\omega_M^2} \\ &\approx \frac{1}{\pi}\int_{-N\Delta\omega/2}^{(N/2-1)\Delta\omega} d\omega\frac{R}{R^2+\omega^2} \\ &\approx \frac{1}{\pi}\int_{-\infty}^{\infty} d\omega\frac{R}{R^2+\omega^2} \\ &= \frac{1}{\pi}\tan^{-1}(\omega/R)\Big|_{-\infty}^{\infty} \\ &= 1 \end{aligned}$$

Consistent with Parseval's theorem, we have deduced that the power spectrum sums to 1, the variance that we constructed our red noise to have. In this derivation, extending the limits of the integral to infinity is a good approximation if $N\Delta\omega/2 \gg R$. Since $N\Delta\omega/2 = N\pi/L = \pi/\Delta t$, this is equivalent to $R\Delta t \ll 1$, which was the assumption we made in deriving the discrete red noise spectrum (that the red noise is well resolved).

11.4 Fitting red noise to data

One common way of fitting the autocorrelation sequence is a *red noise fit*, as an exponentially decreasing function of lag. This fit is shown in plot as chain dash, using an e-folding time of $\tau = 6.1$ months. Roughly speaking, measurements closer together than τ will be significantly correlated and those further apart will be only weakly correlated. This can be cast in terms of a *effective lag-1 autocorrelation* $r = \exp(\Delta t/\tau)$ (= 0.85 in our case). Because the actual autocorrelation is not exactly an exponentially decreasing function of lag, r is not exactly the same as the true lag-1 autocorrelation of 0.9.

Script **nino2** adds a red-noise fit to the SSTA power spectrum based on $\tau = 6.1$ months and scaled to match the observed variance of SSTA. There are four harmonics in the 0.2-0.4 yr^{-1} range that clearly stand above the red noise spectrum.

To test how likely this is to be a chance occurrence, we now look for a less noisy way to estimate the power spectrum.