SPECTRAL ANALYSIS

Any stationary Time-Series sequence will normally exhibit random cycles of various frequencies. Depending on the model, some frequencies are more likely to occur that others. The likelihood of generating a random cycle of a given frequency can be expressed in terms of the so called SPECTRAL DENSITY, which is computed by matching the X_i sequence to an *exact* sin/cos sequence of a specific (angular) frequency β, i.e. $Y_n \equiv \sin(\beta\, n + \theta)$, and adjusting the phase shift θ to yield a maximum value of the following quantity:

$$\omega(\beta) \equiv \frac{\left[\sum\limits_{n=1}^{N} X_n \cdot \sin(\beta\, n + \theta)\right]^2}{\sum\limits_{n=1}^{N} X_n^2}$$

(reminiscent of computing $\frac{N}{2}\rho^2$ between the two sequences). Note that it is sufficient to take β from the $[0, \pi]$ interval, since $\omega(\beta) = \omega(\beta + 2\pi) = \omega(-\beta)$.

To maximize

$$\begin{aligned}&\left[\sum_{n=1}^{N} X_n \cdot \sin(\beta\, n + \theta)\right]^2 = \\ &\left[\cos\theta \sum_{n=1}^{N} X_n \cdot \sin(\beta\, n) + \sin\theta \sum_{n=1}^{N} X_n \cdot \cos(\beta\, n)\right]^2 \equiv \\ &(\cos\theta \cdot S + \sin\theta \cdot C)^2\end{aligned}$$

we differentiate withrespect to θ, getting

$$\begin{aligned}&2(\cos\theta \cdot S + \sin\theta \cdot C)(-\sin\theta \cdot S + \cos\theta \cdot C) = \\ &2\sin\theta\cos\theta(C^2 - S^2) + 2SC(\cos^2\theta - \sin^2\theta)\end{aligned}$$

Setting this expression to zero yields four solutions (first two maximize, last two minimize):

$$\begin{aligned}\sin\theta &= \frac{C}{\sqrt{S^2+C^2}}, \frac{-C}{\sqrt{S^2+C^2}}, \frac{S}{\sqrt{S^2+C^2}}, \frac{-S}{\sqrt{S^2+C^2}} \\ \cos\theta &= \frac{S}{\sqrt{S^2+C^2}}, \frac{-S}{\sqrt{S^2+C^2}}, \frac{-C}{\sqrt{S^2+C^2}}, \frac{C}{\sqrt{S^2+C^2}}\end{aligned}$$

The maximum value of the original expression is thus

$$\begin{aligned}&\left(\frac{S}{\sqrt{S^2+C^2}} \cdot S + \frac{C}{\sqrt{S^2+C^2}} \cdot C\right)^2 = \\ &S^2 + C^2 = \left[\sum_{=1}^{N} X_n \cdot \sin(\beta\, n)\right]^2 + \left[\sum_{i=1}^{N} X_n \cdot \cos(\beta\, n)\right]^2\end{aligned}$$

The spectral density can thus be redefined as

$$\omega(\beta) \equiv \frac{\left[\sum\limits_{n=1}^{N} X_n \cdot \sin(\beta\, n)\right]^2 + \left[\sum\limits_{n=1}^{N} X_n \cdot \cos(\beta\, n)\right]^2}{\sum\limits_{n=1}^{N} X_n^2}$$

which is more convenient computationally (no messing up with θ), but loses the transparency of its original meaning.

So far, we have made only an 'empirical' definition of $\omega(\beta)$, providing, effectively, only its *estimator*. The 'theoretical' definition of spectral density uses the *expected value* of the previous expression in the $N \to \infty$ limit. Realizig that $\sin(\beta\, n) \equiv \frac{\exp(i\, n\, \beta) - \exp(-i\, n\, \beta)}{2\, i}$ and $\cos(\beta n) \equiv \frac{\exp(i\, n\, \beta) + \exp(-i\, n\, \beta)}{2}$, the numerator of the previous formula can be written as

$$\begin{aligned}
&\sum_{n=1}^{N} X_n \frac{\exp(i\, n\, \beta) - \exp(-i\, n\, \beta)}{2\, i} \cdot \\
&\sum_{k=1}^{N} X_k \frac{\exp(i\, k\, \beta) - \exp(-i\, k\, \beta)}{2\, i} + \\
&\sum_{n=1}^{N} X_n \frac{\exp(i\, n\, \beta) + \exp(-i\, n\, \beta)}{2} \cdot \\
&\sum_{k=1}^{N} X_k \frac{\exp(i\, k\, \beta) + \exp(-i\, k\, \beta)}{2} = \\
&2 \sum_{n=1}^{N} \sum_{k=1}^{N} X_n \cdot X_k \frac{\exp[i\,(n-k)\,\beta] + \exp[-i\,(n-k)\,\beta]}{4} = \\
&\sum_{n=1}^{N} \sum_{k=1}^{N} X_n \cdot X_k \cdot \cos[\beta(n-k)] = \\
&(N - |\, j\,|) \cdot Var(X) \sum_{j=-N+1}^{N-1} \rho_j \cos(\beta\, j)
\end{aligned}$$

This, divided by N, yields $Var(X) \sum\limits_{j=-\infty}^{\infty} \rho_j \cos(\beta\, j)$ in the $N \to \infty$ limit. Similarly $\frac{\sum\limits_{n=1}^{N} X_n^2}{N} \to Var(X)$. Thus, we get

$$\omega(\beta) = \sum_{j=-\infty}^{\infty} \rho_j \cos(\beta\, j) \equiv \sum_{j=-\infty}^{\infty} \rho_j \exp(i\, \beta\, j)$$

For the white-noise model, this results in $\omega(\beta) \equiv 1$ (one is always the *average* value of any model).

For our autoregression models, there is yet another way of expressing $\omega(\beta)$, this time directly in terms of the model's α coefficients. First we rewrite

$$X_n = \alpha_1 X_{n-1} + \alpha_2 X_{n-2} + \alpha_3 X_{n-3} + ... + \epsilon_n$$

as

$$X_n = \alpha_1 \mathcal{B} X_n + \alpha_2 \mathcal{B}^2 X_n + \alpha_3 \mathcal{B}^3 X_n + ... + \epsilon_n$$

where $\mathcal{B}$ is an `operator` which *decreases* the index of its argument by one. In the formal manner, the previous equation can be rearranged as follows

$$X_n - \alpha_1 \mathcal{B} X_n - \alpha_2 \mathcal{B}^2 X_n - \alpha_3 \mathcal{B}^3 X_n + ... = \epsilon_n$$

and 'solved' (rather symbolically) for X_n, thus:

$$X_n = \frac{1}{1 - \alpha_1 \mathcal{B} - \alpha_2 \mathcal{B}^2 - \alpha_3 \mathcal{B}^3 + ...}\, \epsilon_n \equiv f(\mathcal{B})\, \epsilon_n$$

To find an *explicit* solution for X_n, all we have to do is to expand $f(x)$ into the usual Taylor series $f_0 + f_1 x + f_2 x^2 + f_3 x^3 + ...$ and write

$$\begin{aligned} X_n &= (f_0 + f_1 \mathcal{B} + f_2 \mathcal{B}^2 + f_3 \mathcal{B}^3 + ...)\, \epsilon_n = \\ & \quad f_0 \epsilon_n + f_1 \epsilon_{n-1} + f_2 \epsilon_{n-2} + f_3 \epsilon_{n-3} + ... \end{aligned}$$

Now, since

$$\begin{aligned} &\rho_k \equiv \frac{\mathbb{E}(X_n \cdot X_{n+k})}{\mathrm{Var}(X_n)} = \frac{1}{\mathrm{Var}(X_n)} \cdot \\ &\mathbb{E}[(f_0 \epsilon_n + f_1 \epsilon_{n-1} + f_2 \epsilon_{n-2} + f_3 \epsilon_{n-3} + ...) \cdot \\ &(f_0 \epsilon_{n+k} + f_1 \epsilon_{n+k-1} + f_2 \epsilon_{n+k-2} + f_3 \epsilon_{n+k-3} + ...)] \\ &= \frac{(f_0 f_k + f_1 f_{k+1} + f_2 f_{k+2} + f_3 f_{k+3} + ...)\sigma^2}{\mathrm{Var}(X_n)} = \\ &\frac{\sigma^2}{\mathrm{Var}(X_n)} \sum_{j=0}^{\infty} f_j\, f_{j+k} \end{aligned}$$

for k non-negative. We can actually extend the previous formula to cover any k by writing

$$\rho_k = \frac{\sigma^2}{\mathrm{Var}(X_n)} \sum_{j=-\infty}^{\infty} f_j\, f_{j+k}$$

with the understanding that $f_{-1} = f_{-2} = ... \equiv 0$.

Our spectral density $\omega(\beta)$ can now be expressed as

$$\begin{aligned} &\frac{\sigma^2}{\mathrm{Var}(X_n)} \sum_{k=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} f_j\, f_{j+k}\, \exp(i\, k\, \beta) = \\ &\frac{\sigma^2}{\mathrm{Var}(X_n)} \sum_{j=-\infty}^{\infty} f_j\, \exp(-i\, j\, \beta) \sum_{k=-\infty}^{\infty} f_{j+k}\, \exp(i\, \beta)^{j+k} = \\ &\frac{\sigma^2}{\mathrm{Var}(X_n)}\, f(e^{i\,\beta}) \sum_{j=-\infty}^{\infty} f_j\, \exp(-i\, j\, \beta) = \frac{\sigma^2 f(e^{i\,\beta}) f(e^{-i\,\beta})}{\mathrm{Var}(X_n)} \end{aligned}$$

Examples:

1. Markov model: Since $Var(X_n) = \frac{\sigma^2}{1-\rho^2}$ and $f(x) = \frac{1}{1-\rho\,x}$, we get
$$\omega(\beta) = \frac{1-\rho^2}{(1-\rho\,e^{i\,\beta})\,(1-\rho\,e^{-i\,\beta})} = \frac{1-\rho^2}{1-2\,\rho\cos\beta+\rho^2}$$
(try to visualize the shape of the corresponding function).

2. Yule model:
$$\omega(\beta) = \frac{(1+\alpha_2)(1-\alpha_1-\alpha_2)(1-\alpha_2+\alpha_1)}{(1-\alpha_2)(1-\alpha_1 e^{i\,\beta}-\alpha_2 e^{2i\,\beta})(1-\alpha_1 e^{-i\,\beta}-\alpha_2 e^{-2i\,\beta})} =$$
$$\frac{(1+\alpha_2)(1-\alpha_1-\alpha_2)(1-\alpha_2+\alpha_1)}{(1-\alpha_2)[1+\alpha_1^2+\alpha_2^2-2\alpha_1(1-\alpha_2)\cos\beta-2\alpha_2\cos 2\beta]}$$

3. The $X_i = 0.3\,X_{i-1} + 0.1\,X_{i-2} - 0.2\,X_{i-3} + \epsilon_i$ model: We already know that $\frac{Var(X_i)}{\sigma^2} = 1.15$. Thus,
$$\omega(\beta) =$$
$$\frac{1}{1.15\,(1-0.3e^{i\,\beta}-0.1e^{2i\,\beta}+0.2e^{3i\,\beta})(1-0.3e^{-i\beta}-0.1e^{-2i\,\beta}+0.2e^{-3i\,\beta})}$$
$$= \frac{1}{1.15\,(1.14-0.58\,\cos\beta-0.32\,\cos 2\beta+0.4\,\cos 3\beta)}$$

Moving Averages (Filtering)

It is sometimes desirable to modify a sequence generated by any of the previous (autoregressive, e.g. AR) models to (for example) smooth out its random oscillations. This can be achieved by converting it into a new sequence of a so called `moving averages`, thus
$$Y_i \equiv \frac{X_i + X_{i-1}}{2}$$
or
$$Y_i = \frac{X_i + X_{i-1} + X_{i-2}}{3}$$
etc.
$$Y_i = \gamma_0 X_i + \gamma_1 X_{i-1} + \gamma_2 X_{i-2} + ...$$
in general (this will correspond to an 'average' only when the sum of all γ s equals to 1, but one can eventually allow anything else, e.g. $Y_i = X_i - X_{i-1}$ – exploring the daily increases/decreases, etc). Note that the new Y_i sequence is no longer Markovian. Nevertheless, since
$$\begin{aligned} Y_i = (\gamma_0 + \gamma_1\mathcal{B} + \gamma_2\mathcal{B}^2 + ...)X_i \quad &= \\ \frac{\gamma_0 + \gamma_1\mathcal{B} + \gamma_2\mathcal{B}^2 + ...}{1-\alpha_1\mathcal{B}-\alpha_2\mathcal{B}^2-\alpha_3\mathcal{B}^3+...}\,\epsilon_n \quad &\equiv \quad f(\mathcal{B})\,\epsilon_n \end{aligned}$$
the formulas of the previous section for finding ρ_k and constructing the corresponding spectrum still hold.

We can also *design* the $\gamma_0+\gamma_1\mathcal{B}+\gamma_2\mathcal{B}^2+...$ operator to deliberately suppress (up to total elimination) or emphasize certain frequencies, this procedure is then called `filtering`.

The corresponding $\omega(\beta)$ function is then called the filter's `response`, clearly indicating which frequencies will get suppressed (or emphasized), and by what factor.

Or, we can simply use

$$Y_i = \frac{\gamma_0 + \gamma_1 \mathcal{B} + \gamma_2 \mathcal{B}^2 + ...}{1 - \alpha_1 \mathcal{B} - \alpha_2 \mathcal{B}^2 - \alpha_3 \mathcal{B}^3 + ...} \, \epsilon_n$$

as a general model of a random process, making it thus more flexible than the autoregressive model alone (the new model is usually called ARMA, for obvious reasons).