WHITTLE'S APPROXIMATION TO THE LIKELIHOOD FUNCTION

Suppose we wish to fit a parametric model (such as ARIMA, or Fractional ARIMA) to data $x = (x_0, \ldots, x_{n-1})'$ from the zero-mean weakly stationary Gaussian time series $\{x_t\}$. Let θ denote the vector of model parameters. Under the model θ, suppose that $\{x_t\}$ has spectral density $f_\theta(\omega)$, autocovariance sequence $\{c_{r,\theta}\}$, and suppose that x has $n \times n$ covariance matrix $\Sigma_{n,\theta}$. Then the likelihood for θ is

$$lik(\theta) = (2\pi)^{-n/2} \frac{1}{\sqrt{|\Sigma_{n,\theta}|}} \exp\{-\frac{1}{2} x' \Sigma_{n,\theta}^{-1} x\} \ .$$

The MLE, $\hat{\theta}$, is the value of θ which maximizes $lik(\theta)$, or equivalently, which minimizes

$$-2 \log lik(\theta) = n \log(2\pi) + \log |\Sigma_{n,\theta}| + x' \Sigma_{n,\theta}^{-1} x \ .$$

In general, the cost of inverting an $n \times n$ matrix is $O(n^3)$. Thus, in principle, each evaluation of the likelihood function will require $O(n^3)$ operations. Using Levinson's algorithm (described later), we can bring the cost of the inversion, and therefore the cost of each evaluation of the likelihood function, down to $O(n^2)$. Here, we will present **Whittle's Approximation** to $-2 \log lik(\theta)$, which has the advantage that it can be evaluated in $O(n \log n)$ operations.

The matrix $\Sigma_{n,\theta}$ is said to be a **Toeplitz matrix,** since all diagonals of $\Sigma_{n,\theta}$ are constant. (This follows since $\Sigma_{n,\theta}(j, k) = c_{j-k,\theta}$). It can be shown that, for large n, *all* $n \times n$ symmetric Toeplitz matrices have complex orthonormal eigenvectors which can be well approximated by

$$V_j = n^{-1/2} \{\exp(-i\,\omega_j t)\}_{t=0}^{n-1} \ , \ (j = 0, \ldots, n-1) \ .$$

It can also be shown that the corresponding eigenvalues of $\Sigma_{n,\theta}$ are well approximated by $2\pi f_\theta(\omega_j)$. Thus, if $V = (V_0, \ldots, V_{n-1})$, and Λ is an $n \times n$ diagonal matrix with $\{2\pi f_\theta(\omega_j)\}_{j=0}^{n-1}$ on the main diagonal and zero elsewhere, then $\Sigma_{n,\theta} \approx V \Lambda V^*$, where V^* is the conjugate transpose of V. Note that $VV^* = V^* V = I$, the $n \times n$ identity matrix, so that $\Sigma_{n,\theta}^{-1} \approx V \Lambda^{-1} V^*$. In addition, $|V| = 1$, since V is a unitary matrix. Thus,

$$-2 \log lik(\theta) \approx n \log 2\pi + \log |V \Lambda V^*| + x' V \Lambda^{-1} V^* x$$

$$= n \log 2\pi + \log |\Lambda| + (x'V\Lambda^{-1/2})(x'V\Lambda^{-1/2})^* \quad .$$

Now, the j'th entry of $x'V$ is

$$x'V_j = \frac{1}{\sqrt{n}} \sum_{t=0}^{n-1} x_t \exp(-i\,\omega_j\, t) = \sqrt{n}\; J_j \quad ,$$

and the j'entry of $x'V\Lambda^{-1/2}$ is $\sqrt{n}\,J_j/\sqrt{2\pi f_\theta(\omega_j)}$. Thus,

$$-2 \log lik(\theta) \approx n \log 2\pi + \log \prod_{j=0}^{n-1} (2\pi f_\theta(\omega_j)) + \sum_{j=0}^{n-1} \frac{n}{2\pi f_\theta(\omega_j)} |J_j|^2$$

$$= 2n \log 2\pi + \sum_{j=0}^{n-1} [\log f_\theta(\omega_j) + I_j / f_\theta(\omega_j)] \quad . \qquad (1)$$

Formula (1) is Whittle's approximation to $-2 \log lik(\theta)$. Since $\{I_j\}_{j=0}^{n-1}$ can be evaluated in $O(n \log n)$ using the Fast Fourier Transform, and since $\{\log f_\theta(\omega_j)\}_{j=0}^{n-1}$ can be evaluated in $O(n)$, we can evaluate the righthand side of (1) in $O(n \log n)$ operations. The value of θ which minimizes the righthand side of (1) is called the **Whittle Estimator,** $\hat{\theta}_W$.

Fox and Taqqu have shown that for a Gaussian fractional ARIMA model, $\hat{\theta}_W$ is asymptotically normal, and is asymptotically efficient, so that $\hat{\theta}_W$ is asymptotically equivalent to the exact MLE, $\hat{\theta}$. It follows that $\hat{\theta}_W$ also provides an asymptotically efficient estimate of an ARMA model, since the ARMA models are a subclass of the fractional ARIMA models.