Finc 520
Time Series Analysis
Lawrence Christiano

Approximating the Normal Density Function Using the Asymptotic Diagonalization Result

Let $f_j$ denote the $1 \times T$ row vector:

$$f_j = \frac{1}{\sqrt{T}} \left[ \begin{array}{cccc} e^{-i\omega_{j-1}} & e^{-i\omega_{j-1}2} & \cdots & e^{-i\omega_{j-1}T} \end{array} \right],$$

for $j = 1, ..., T$, where

$$\omega_j = \frac{2\pi j}{T},\ j = 0, ..., T-1.$$

The row vectors, $f_1, ..., f_T$, have the following property:

$$f_j \bar{f}_k' = \left\{ \begin{array}{cc} 1 & k = j \\ 0 & k \neq j \end{array} \right. . \tag{1}$$

Here, a bar over a variable indicates complex conjugation, and a prime denotes transposition (without conjugation)[1]. Let the matrix, $F$, be defined by:

$$F = \left[ \begin{array}{c} f_1 \\ f_2 \\ \vdots \\ f_T \end{array} \right].$$

The $T \times T$ matrix, $F$, is called the Fourier matrix. Property (1) of the rows implies:

$$F\bar{F}' = I, \tag{2}$$

where $I$ denotes the $T$-dimensional identity matrix. A property like (1) also applies to the columns of $F$. In particular, let

$$F = \left[ \begin{array}{ccc} g_1 & \cdots & g_T \end{array} \right],$$

where $g_k$ is a $T \times 1$ column vector, $k = 1, ..., T$. Then,

$$\bar{g}_j' g_k = \left\{ \begin{array}{cc} 1 & k = j \\ 0 & k \neq j \end{array} \right. . \tag{3}$$

Property (3) can be written:

$$\bar{F}'F = I. \tag{4}$$

[1] Careful, MATLAB defines the prime to mean conjgation *and* transposition.

Let

$$y=\left(\begin{array}{c} y_1 \\ \vdots \\ y_T \end{array}\right),\ V=Eyy'.$$

Here, $y_t$ is a covariance stationary stochastic process. It can be shown that

$$f_j V \bar{f}_k \rightarrow_{T\rightarrow\infty} \left\{ \begin{array}{cc} S(\omega_j) & k=j \\ 0 & k\neq j \end{array} \right.,$$

where

$$S(\omega_j)=\sum_{l=-\infty}^{\infty}\gamma_l e^{-i\omega_j l},\ \gamma_l=Ey_t y_{t-l}.$$

Note that

$$V=\left[\begin{array}{cccc} \gamma_0 & \gamma_1 & \cdots & \gamma_{T-1} \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{T-2} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{T-1} & \gamma_{T-2} & \cdots & \gamma_0 \end{array}\right].$$

A different way to put the result is as follows:

$$FV\bar{F}' \rightarrow_{T\rightarrow\infty} = \left[\begin{array}{cccc} S(\omega_0) & 0 & \cdots & 0 \\ 0 & S(\omega_1) & \cdots & 0 \\ \vdots & \vdots & \ddots & \cdots \\ 0 & 0 & \cdots & S(\omega_{T-1}) \end{array}\right]. \tag{5}$$

Not surprisingly, this result is referred to as the 'asymptotic diagonalization' result. A discussion of this result can be found in Harvey (1989, section 4.3).

The asymptotic diagonalization result delivers an efficient algorithm for evaluating the Normal density function when doing maximum likelihood or Bayesian inference. In addition, the result gives rise to a useful set of diagnostics for model evaluation. Some of these are discussed in Christiano and Vigfusson (2003).

The multivariate Normal density function is:

$$L(y;\beta)=(2\pi)^{-\frac{T}{2}}|V|^{-\frac{1}{2}}\exp\left[-\frac{1}{2}y'V^{-1}y\right],$$

where $|A|$ is the determinant of the matrix, $A$, and $\beta$ denotes the parameters of the model. The parameters, $\beta$, determine $V$, but the dependence is suppressed to keep from cluttering the notation. Consider the quadratic form:

$$y'V^{-1}y=y'\bar{F}'FV^{-1}\bar{F}'Fy=\bar{\tilde{y}}'FV^{-1}\bar{F}'\tilde{y},$$

where $\tilde{y}$ is a $T\times 1$ column vector defined by:

$$\tilde{y}=Fy=\left(\begin{array}{c} y(\omega_0) \\ y(\omega_1) \\ \vdots \\ y(\omega_{T-1}) \end{array}\right),$$

where

$$y(\omega_j) \equiv f_j y = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} e^{-i\omega_j t} y_t, \; j = 0, ..., T-1.$$

Consider the object, $FV^{-1}\bar{F}'$. Note that by the definition of an inverse,

$$\left(FV\bar{F}'\right)^{-1} FV\bar{F}' = I.$$

Post-multiply both sides of this expression by $F$ and use (4) to obtain:

$$\left(FV\bar{F}'\right)^{-1} FV = F.$$

Now, post-multiply both sides of the above expression by $V^{-1}$ :

$$\left(FV\bar{F}'\right)^{-1} F = FV^{-1}.$$

Finally, post-multiply both sides by $\bar{F}'$ and use (2):

$$\left(FV\bar{F}'\right)^{-1} = FV^{-1}\bar{F}'.$$

Thus, in considering the quadratic form in the Normal density function, we are led to consider the inverse of $FV\bar{F}'$. By (5), this matrix is (for large enough $T$) diagonal. Thus,

$$\begin{aligned} & y'V^{-1}y \qquad (6) \\ \simeq & \begin{pmatrix} y(-\omega_0) & y(-\omega_1) & \cdots & y(-\omega_{T-1}) \end{pmatrix} \begin{bmatrix} \frac{1}{S(\omega_0)} & 0 & \cdots & 0 \\ 0 & \frac{1}{S(\omega_1)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \cdots \\ 0 & 0 & \cdots & \frac{1}{S(\omega_{T-1})} \end{bmatrix} \begin{pmatrix} y(\omega_0) \\ y(\omega_1) \\ \vdots \\ y(\omega_{T-1}) \end{pmatrix} \\ = & \sum_{j=0}^{T-1} \frac{I(\omega_j)}{S(\omega_j)}, \; I(\omega_j) \equiv y(-\omega_j)\, y(\omega_j). \end{aligned}$$

Here, $I(\omega_j)$ is the sample periodogram. It is an estimator of $S(\omega_j)$. The other term in the likelihood is the determinant:

$$|V| = \left|\bar{F}'FV\bar{F}'F\right| = \left|\bar{F}'\right| \left|FV\bar{F}'\right| |F|.$$

By a property of determinants, $\left|\bar{F}'\right| |F| = \left|\bar{F}'F\right|$. By (4), $\left|\bar{F}'F\right| = |I| = 1$. Thus, we conclude

$$|V| = \left|FV\bar{F}'\right| \simeq \det\left(\begin{bmatrix} S(\omega_0) & 0 & \cdots & 0 \\ 0 & S(\omega_1) & \cdots & 0 \\ \vdots & \vdots & \ddots & \cdots \\ 0 & 0 & \cdots & S(\omega_{T-1}) \end{bmatrix}\right) = \prod_{j=0}^{T-1} S(\omega_j). \qquad (7)$$

We conclude that, for sufficiently large $T$,

$$\begin{aligned} L(y;\beta) &= (2\pi)^{-\frac{T}{2}} |V|^{-\frac{1}{2}} \exp\left[-\frac{1}{2} y' V^{-1} y\right] \\ &\simeq (2\pi)^{-\frac{T}{2}} \left[\prod_{j=0}^{T-1} S(\omega_j)\right]^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\sum_{j=0}^{T-1} \frac{I(\omega_j)}{S(\omega_j)}\right]. \end{aligned}$$

Using the asymptotic diagonalization result, we have converted the problem of evaluating the Norma density from one that requires inverting a $T \times T$ matrix into a problem of inverting $T$ scalars. Finding the $\beta$ that maximizes $L$ involves fairly straightforward calculations. The periodogram, $I(\omega_j)$, must be computed only once. The object, $S(\omega_j)$, is also straightforward to evaluate. Suppose, for example, that we are estimating an ARMA(p,q) model for $y_t$ :

$$\phi(L) y_t = \theta(L) \varepsilon_t,$$

where $\varepsilon_t$ is a mean zero, non-autocorrelated disturbance that is orthogonal to $y_{t-s}$ for $s > 0$. Then,

$$S(\omega) = \phi\left(e^{-i\omega}\right)^{-1} \theta\left(e^{-i\omega}\right) \Omega \theta\left(e^{i\omega}\right)' \left[\phi\left(e^{i\omega}\right)^{-1}\right]'.$$

This object has been written for the case in which $y_t$ is an $n \times 1$ vector. In this case the likelihood is written

$$(2\pi)^{-\frac{nT}{2}} \left[\prod_{j=0}^{T-1} |S(\omega_j)|\right]^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\sum_{j=0}^{T-1} tr\left[S(\omega_j)^{-1} I(\omega_j)\right]\right),$$

where $tr(A)$ is the sum of the diagonal elements of the matrix, $A$.

# References


[1] Harvey, Andrew C., 1989, Forecasting, Structural Time Series Models and the Kalman Filter, Cambridge University Press.

[2] Christiano, Lawrence J., and Robert Vigfusson, 2003, 'Maximum Likelihood in the Frequency Domain: the Importance of Time to Plan', Journal of Monetary Economics 50, pp. 789-815.