14.384 Time Series Analysis, Fall 2008
Recitation by Paul Schrimpf
Supplementary to lectures given by Anna Mikusheva
September 26, 2008

Recitation 4

Spectrum Estimation

We have a stationary series, $\{z_t\}$ with covariances γ_j and spectrum $S(\omega) = \sum_{j=-\infty}^{\infty} \gamma_j e^{-i\omega j}$. We want to estimate $S(\omega)$.

Using Covariances

As in lecture 5, we can estimate the spectrum in the same way that we estimate the long-run variance.

Naïve approach

We cannot estimate all the covariances from a finite sample. Let's just estimate all the covariances that we can

$$\hat{\gamma}_j = \frac{1}{T} \sum_{j=k+1}^{T} z_j z_{j-k}$$

and use them to form

$$\hat{S}(\omega) = \sum_{j=-(T-1)}^{T-1} \hat{\gamma}_j e^{-i\omega j}$$

This estimator is not consistent. It convergers to a distribution instead of a point. To see this, let $y_\omega = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} e^{-i\omega t} z_t$, so that

$$\hat{S}(\omega) = y_\omega \bar{y}_\omega$$

If $\omega \neq 0$

$$2\hat{S}(\omega) \Rightarrow S(\omega)\chi^2(2)$$

Kernel Estimator

$$\hat{S}(\omega) = \sum_{j=-S_T}^{S_T} \left(1 - \frac{|j|}{S_T}\right) \hat{\gamma}_j e^{-i\omega j}$$

Under appropriate conditions on S_T ($S_T \to \infty$, but more slowly than T), this estimator is consistent[1] This can be shown in a way similar to the way we showed the Newey-West estimator is consistent.

[1] In a uniform sense, *i.e.* $P\left(\sup_{\omega \in [-\pi,\pi]} |\hat{S}(\omega) - S(\omega)| > \epsilon\right) \to 0$

Proof. This is an informal "proof" that sketches the ideas, but isn't completely rigorous. It is nearly identical to the proof of HAC consistency in lecture 3.

$$
\begin{aligned}
|\hat{S}(\omega) - S(\omega)| =& | - \sum_{|j|>S_T} \gamma_j e^{-i\omega j} + \sum_{j=-S_T}^{S_T} (k_T(j) - 1)\gamma_j e^{-i\omega j} + \sum_{j=-S_T}^{S_T} k_T(j)(\hat{\gamma}_j - \gamma_j) e^{-i\omega j}| \\
\leq & | \sum_{|j|>S_T} \gamma_j| + | \sum_{j=-S_T}^{S_T} (k_T(j) - 1)\gamma_j| + | \sum_{j=-S_T}^{S_T} k_T(j)(\hat{\gamma}_j - \gamma_j)|
\end{aligned}
$$

We can interprete these three terms as follows;

1. $|\sum_{|j|>S_T} \gamma_j|$ is truncation error
2. $|\sum_{j=-S_T}^{S_T} (k_T(j) - 1)\gamma_j|$ is error from using the kernel
3. $|\sum_{j=-S_T}^{S_T} k_T(j)(\hat{\gamma}_j - \gamma_j)|$ is error from estimating the covariances

Terms 1 and 2 are non-stochastic. They represent bias. The third term is stochastic; it is responsible for uncertainty. We will face a bias-variance tradeoff.

We want to show that each of these terms goes to zero

1. Disappears as long as $S_T \to \infty$, since we assumed $\sum_{-\infty}^{\infty} |\gamma_j| < \infty$.
2. $\sum_{j=-S_T}^{S_T} (k_T(j) - 1)\gamma_j \leq \sum_{j=-S_T}^{S_T} |k_T(j) - 1||\gamma_j|$ This will converge to zero as long as $k_T(j) \to 1$ as $T \to \infty$ and $|k_T(j)| < 1 \ \forall j$.
3. Notice that for the first two terms we wanted S_T big enough to eliminate them. Here, we'll want S_T to be small enough.

   First, note that $\hat{\gamma}_j \equiv \frac{1}{T} \sum_{k=1}^{T-j} z_k z_{k+j}$ is not unbiased. $E\hat{\gamma}_j = \frac{T-j}{T}\gamma_j = \tilde{\gamma}_j$. However, it's clear that this bias will disappear as $T \to \infty$.

   Let $\xi_{t,j} = z_t z_{t+j} - \gamma_j$, so $\hat{\gamma}_j - \tilde{\gamma}_j = \frac{1}{T} \sum_{\tau=1}^{T-j} \xi_{\tau,j}$. We need to show that the sum of $\xi_{t,j}$ goes to zero.

$$
\begin{aligned}
E(\hat{\gamma}_j - \tilde{\gamma}_j)^2 =& \frac{1}{T^2} \sum_{k=1}^{T-j} \sum_{t=1}^{T-j} \operatorname{cov}(\xi_{k,j}, \xi_{t,j}) \\
\leq & \frac{1}{T^2} \sum_{k=1}^{T-j} \sum_{t=1}^{T-j} |\operatorname{cov}(\xi_{k,j}, \xi_{t,j})|
\end{aligned}
$$

   We need an assumption to guarantee that the covariances of ξ disappear. The assumption that $\xi_{t,j}$ are sationary for all j and $\sup_j \sum_k |\operatorname{cov}(\xi_{t,j}, \xi_{t+k,j})| < C$ for some constant C implies that

$$
\frac{1}{T^2} \sum_{k=1}^{T-j} \sum_{t=1}^{T-j} |\operatorname{cov}(\xi_{k,j}, \xi_{t,j})| \leq \frac{C}{T}
$$

   By Chebyshev's inequality we have:

$$
P(|\hat{\gamma}_j - \tilde{\gamma}_j| > \epsilon) \leq \frac{E(\hat{\gamma}_j - \tilde{\gamma}_j)^2}{\epsilon^2} \leq \frac{C}{\epsilon^2 T}
$$

Then adding these together:

$$P(\sum_{-S_T}^{S_T} |\hat{\gamma}_j - \tilde{\gamma}_j| > \epsilon) \leq \sum_{-S_T}^{S_T} P(|\hat{\gamma}_j - \tilde{\gamma}_j| > \frac{\epsilon}{2S_T + 1})$$
$$\leq \sum_{-S_T}^{S_T} \frac{E(\hat{\gamma}_j - \gamma_j)^2}{\epsilon^2} (2S_T + 1)^2$$
$$\leq \sum_{-S_T}^{S_T} \frac{C}{T} (2S_T + 1)^2 \approx C_1 \frac{S_T^3}{T}$$

so, it is enough to assume $\frac{S_T^3}{T} \to 0$ as $T \to \infty$.

□

Using the Sample Periodogram

The sample periodogram (or sample spectral density) is the square of the finite Fourier transform of the data, *i.e.*

$$I(\omega) = \frac{1}{T} \left| \sum_{t=1}^{T} z_t e^{-i\omega t} \right|^2$$

The sample periodogram is the same as the naive estimate of the spectrum that uses all the sample covariances.

$$\begin{aligned} I(\omega) =& \frac{1}{T} \left(\sum_{t=1}^{T} z_t e^{-i\omega t} \right) \left(\sum_{t=1}^{T} z_t e^{i\omega t} \right) \\ =& \frac{1}{T} \sum_{t=1}^{T} \sum_{s=1}^{T} e^{i\omega(t-s)} z_t z_s \\ =& \sum_{j=-(T-1)}^{T-1} e^{i\omega j} \frac{1}{T} \sum_{t=|j|}^{T} z_t z_{t-|j|} \\ =& \sum_{j=-(T-1)}^{T-1} e^{i\omega j} \hat{\gamma}_j \end{aligned}$$

Smoothed Periodogram

Above, we showed that

$$2I(\omega) \Rightarrow S(\omega)\chi^2(2)$$

It's also true that,

$$\lim_{T \to \infty} cov(I(\omega_1), I(\omega_2)) = 0$$

The sample periodogram is uncorrelated at adjacent frequencies. This suggests that we could estimate the spectrum at ω by taking an average over frequencies near ω. That is,

$$\hat{S}_{sp}(\omega) = \int_{-\pi}^{\pi} h_T(\omega - \lambda) I(\lambda) d\lambda$$

where $h_T()$ is a kernel function that peaks at 0. It turns out that this estimator is equivalent to a kernel covariance estimator.

$$\begin{aligned}\hat{S}_{sp}(\omega) &= \int_{-\pi}^{\pi} h_T(\omega-\lambda) I(\lambda) d\lambda \\ &= \int_{-\pi}^{\pi} \sum_{j=-(T-1)}^{T-1} \hat{\gamma}_j e^{i\lambda j} h_T(\omega-\lambda) d\lambda \\ &= \sum_{j=-(T-1)}^{T-1} \hat{\gamma}_j \int_{-\pi}^{\pi} e^{i\lambda j} h_T(\omega-\lambda) d\lambda \\ &= \sum_{j=-(T-1)}^{T-1} \hat{\gamma}_j \int_{-\pi}^{\pi} e^{i(\lambda-\omega)j} h_T(\lambda) d\lambda \\ &= \sum_{j=-(T-1)}^{T-1} \hat{\gamma}_j e^{-i\omega j} k_T(j)\end{aligned}$$

where $k_T(j) = \int_{-\pi}^{\pi} e^{i\lambda j} h_T(\lambda) d\lambda$. $k_T(j)$ is the inverse Fourier transform of $h_T(\lambda)$. Conversely, it must be that $h_T(\lambda)$ is the Fourier transform of $k_T(j)$, *i.e.*

$$h_T(\lambda) = \frac{1}{2\pi} \sum_j k_T(j) e^{-i\lambda j}$$

Conditions on $h_T()$ for consistency can be derived from the conditions on k_T in the lecture on HAC estimation, but it does not look entirely straightforward.

VAR ML

In lecture 7, we said that for a VAR, MLE (with normal distribution) is equivalent to OLS equation by equation. We'll prove that now. The argument can be found in Chapter 11 of Hamilton.

Proof. Let's say we have a sample of y_t from $t = 0...T$, and we estimate a VAR of order p, $A(L)$. The model is

$$y_t = \sum_{k=1}^{p} A_k y_{t-k} + e_t \ , \ e_t \sim N(0, \Omega)$$

The likelihood of $y_p, ..., y_T$ conditional on $y_0, ..., y_{p-1}$ is

$$\begin{aligned}f(y_p, ..., y_T | y_0, ..., y_{p-1}) =& f(y_{p+1}, ..., y_T | y_0, ..., y_{p-1}) f(y_p | y_0, ..., y_{p-1}) \\ & \vdots \\ =& \pi_{t=p}^{T} f(y_t | y_{t-1}, ..., y_{t-p})\end{aligned}$$

Each $f(y_t|y_{t-1}, ..., y_{t-p})$ is simply a normal distribution with mean $\sum_{k=1}^{p} A_k y_{t-k}$ and variance Ω, so

$$f(y_p, ..., y_T | y_0, ..., y_{p-1}) = \pi_{t=p}^{T} \sqrt{\frac{|\Omega^{-1}|}{(2\pi)^n}} \exp\left(-\frac{1}{2}(y_t A(L))' \Omega^{-1} (y_t A(L))\right)$$

So the conditional log likelihood is

$$\mathcal{L}(A, \Omega) = -\frac{(T-p)n}{2} \log(2\pi) + \frac{T-p}{2} \log|\Omega^{-1}| - \frac{1}{2} \sum_{t=p}^{T} (y_t A(L))' \Omega^{-1} (y_t A(L))$$

Let $\hat{A}(L)$ be the equation by equation OLS estimate of $A(L)$. We want to show that $\hat{A}(L)$ minimizes $\mathcal{L}(A, \Omega)$. To show this we only need to worry about the last term.

$$\sum_{t=p}^{T}(y_t A(L))'\Omega^{-1}(y_t A(L)) \tag{1}$$

Some different notation will help. Let $x_t = [y_{t-1} \dots y_{t-p}]'$, and let $\Pi = [A_1\ A_2\ \dots\ A_p]$. Similarly define $\hat{\Pi}$. We can rewrite (1) as

$$\begin{aligned}
\sum_{t=p}^{T}(y_t A(L))'\Omega^{-1}(y_t A(L)) =& \sum_{t=p}^{T}(y_t - \Pi' x_t)'\Omega^{-1}(y_t - \Pi' x_t) \\
=& \sum_{t=p}^{T}(y_t - \hat{\Pi}' x_t + (\hat{\Pi}' - \Pi')x_t)'\Omega^{-1}(y_t - \hat{\Pi}' x_t + (\hat{\Pi}' - \Pi')x_t) \\
=& \sum_{t=p}^{T}(\hat{\epsilon}_t + (\hat{\Pi}' - \Pi')x_t)'\Omega^{-1}(\hat{\epsilon}_t + (\hat{\Pi}' - \Pi')x_t) \\
=& \sum_{t=p}^{T}\hat{\epsilon}_t'\Omega^{-1}\hat{\epsilon}_t' + 2\hat{\epsilon}_t'\Omega^{-1}(\hat{\pi}' - \pi)x_t + x_t'(\hat{\Pi} - \Pi)\Omega^{-1}(\hat{\Pi}' - \Pi')x_t
\end{aligned}$$

The middle term is a scalar, so it is equal to its trace.

$$\begin{aligned}
\sum 2\hat{\epsilon}_t'\Omega^{-1}(\hat{\pi}' - \pi)x_t =& \sum \operatorname{trace}\left(2\hat{\epsilon}_t'\Omega^{-1}(\hat{\pi}' - \pi)x_t\right) \\
=& 2\operatorname{trace}\left(\Omega^{-1}(\hat{\pi}' - \pi)\sum x_t\hat{\epsilon}_t'\right) = \qquad 0
\end{aligned}$$

$\sum x_t\hat{\epsilon}_t = 0$ because $\hat{\epsilon}_t$ are OLS residuals and must be orthogonal to x_t. So, we're left with

$$\sum_{t=p}^{T}(y_t A(L))'\Omega^{-1}(y_t A(L)) = \sum_{t=p}^{T}\hat{\epsilon}_t'\Omega^{-1}\hat{\epsilon}_t' + x_t'(\hat{\Pi} - \Pi)\Omega^{-1}(\hat{\Pi}' - \Pi')x_t$$

Only the second term depends on Π. Ω^{-1} is positive definite, so $x_t'(\hat{\Pi} - \Pi)\Omega^{-1}(\hat{\Pi}' - \Pi')x_t$ is minimized when $x_t'(\hat{\Pi} - \Pi) = 0$ for all t, *i.e.* when $\Pi = \hat{\Pi}$. Thus, the OLS estimates are the maximum likelihood estimates.

To find the MLE for Ω, just consider the first order condition evaluated at $\Pi = \hat{\Pi}$:

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \Omega^{-1}} =& \frac{\partial \mathcal{L}}{\partial \Omega^{-1}}\left(-\frac{(T-p)n}{2}\log(2\pi) + \frac{T-p}{2}\log|\Omega^{-1}| - \frac{1}{2}\sum_{t=p}^{T}\hat{\epsilon}_t'\Omega^{-1}\hat{\epsilon}_t'\right) \\
=& \frac{T-p}{2}\Omega - \frac{1}{2}\sum \hat{\epsilon}_t\hat{\epsilon}_t' \qquad = 0 \\
\hat{\Omega} =& \frac{1}{T-p}\sum \hat{\epsilon}_t\hat{\epsilon}_t'
\end{aligned}$$

□