ERROR BANDS FOR IMPULSE RESPONSES, GCP

1. ERROR BANDS FOR IMPULSE RESPONSES

- In one sense, this is straightforward: make draws from the posterior pdf of \(\Sigma\) and \(A\) (or \(\beta\)) and for each draw calculate impulse responses \(c_{ij}(t)\).
- Plot the \(c_{ij}\) corresponding to the MLE (and/or the mean or median of the draws), together with the 5% upper and lower tails of the draws, for example. Could also plot HPD intervals.

2. THESE ARE NOT CONFIDENCE INTERVALS

- It’s a Bayesian calculation. It gives intervals with a clear interpretation, by a straightforward procedure, but they are not confidence intervals (except asymptotically!).
- True confidence intervals for individual \(c_{ij}\)’s are not possible.
- This is a special case of a general point. If \(\theta\) is the complete parameter vector for the distribution of the data \(X\), then we can always (in principle) produce a 90% (say) confidence region for \(\theta\) by constructing a 90% significance level test for \(\theta\) as \(H_0\) for each \(\theta\) in the parameter space. The set of \(\theta\)’s that are accepted in a given sample is an exact 90% confidence set.
- But if we try to construct a confidence set for an individual \(\theta_i\), the problem is that the distribution of any test statistic generally depends on all the parameters, not just \(\theta_i\).
- The normal linear regression model is a special case where there is a set of test statistics that depend on single \(\theta_i\)’s.

3. SO WHAT ARE FREQUENTIST CI’S?

- There are “asymptotic” confidence intervals. These have coverage probabilities that converge to the correct ones for parameter values in some neighborhood of the true parameter value. They can be constructed by linearizing the mapping from \(B_{ij}\) (the AR coefficients) to \(C_{ij}\), then transforming the normal asymptotic distribution for \(B\) to the corresponding approximate normal distribution for \(C\).
- But these have no more frequentist asymptotic justification than do the Bayesian intervals.
- As the forecast horizon expands, the nonlinearities rapidly become more extreme, so the intervals based on linearization are always inaccurate at distant horizons.
- The Bayesian intervals are accurate whether or not \(A\) might have roots of one or larger in absolute value. The frequentist intervals are uninterpretable if that is true.

Date: October 26, 2010.
©2010 by Christopher A. Sims. This document may be reproduced for educational and research purposes, so long as the copies contain this notice and are retained for personal use or distributed free.
4. Pointwise bands, alternatives

- These bands are constructed from marginals for each $c_{ij}$.
- They cannot answer a question like “How likely is a hump-shaped response?”
- One could directly get a distribution for “up-life” from the MCMC sample.
- SZ alternative: display $1 - \alpha$ probability bands for principal components of $\text{Var}(c_{ij})$.
- Like displaying bands for principal components of the posterior of $\beta$ in a regression, but with a more direct intuitive interpretation of the bounds.

5. Testing GCP

Suppose we have a system of the form

$$
\begin{bmatrix}
y_1(t) \\
y_2(t)
\end{bmatrix} =
\begin{bmatrix}
X_1(t) \\
X_2(t)
\end{bmatrix}
\begin{bmatrix}
B_{11} & B_{12} \\
B_{21} & B_{22}
\end{bmatrix} +
\begin{bmatrix}
\epsilon_1(t) \\
\epsilon_2(t)
\end{bmatrix}.
\tag{1}
$$

We can choose a $\gamma$ to make $\nu(t) = \epsilon_1(t) - \epsilon_2(t)\gamma$ orthogonal to $\epsilon_2(t)$ and define $C_1 = B_{11} - B_{12}\gamma, C_2 = B_{21} - B_{22}\gamma, \Omega = \text{Var}([\nu(t) \epsilon_2(t)])$ to allow us to rewrite the system as

$$
\begin{bmatrix}
y_1(t) - y_2(t)\gamma \\
y_2(t)
\end{bmatrix} =
\begin{bmatrix}
X_1(t) \\
X_2(t)
\end{bmatrix}
\begin{bmatrix}
C_1 & B_{12} \\
C_2 & B_{22}
\end{bmatrix} +
\begin{bmatrix}
\nu(t) \\
\epsilon_2(t)
\end{bmatrix}.
\tag{2}
$$

Because the Jacobian of the transformation of parameters is the identity, and because the disturbances in the two blocks of equations in the transformed system are orthogonal, the likelihood factors into two pieces, one involving the parameters of the second equation only, the other involving the transformed first equation’s parameters. Thus if our prior beliefs about the parameters of the transformed system make the parameters of the two blocks of equations independent, inference about the parameters of the second block of equations can be conducted by considering the likelihood for that block alone, as if there were no other equations in the system.

Of course in general, if we had started with prior beliefs expressed in terms of the original system’s parameters, it would be a rare accident if our beliefs about the parameters of the two blocks in the transformed system were unrelated.

But testing for Granger causal priority (GCP) is a case where these results apply. GCP is the condition that, in a system in which $X(t)$ consists entirely of lagged values of $y$ and we have grouped all lagged values of the first block $y_1$ of $y$ into $X_1(t)$, $B_{12} = 0$. Thus a classical likelihood ratio test of the hypothesis that $y_2$ is GCP to $y_1$ is obtained by constructing twice the difference in log likelihoods and treating it as $\chi^2(df)$, where $df$, the degrees of freedom, is the number of elements in $B_{12}$. To be more specific, the classical LR test statistic is

$$
T(\log(\left|S_{22}^R\right|) - \log(\left|S_{22}^U\right|)),
\tag{3}
$$

where $S_{22}^R$ and $S_{22}^U$ are the restricted and unrestricted cross-product matrices of residuals for the second block of equations.