Quantifying Consistency between Series

Frank Masci

Version 2: 24/1/2006

This is a followup "summary" document from (Subject: Quantifying Consistency). A history of related work and other doclinks can be found under the original proposal document: (Subject: Coherence - project proposal).

1. Summary of popular methods and direction

This is nothing final. It only represents a glimmer of consensus picked up by the author from discussions with the team and refinements from Mark Zhang. The following four "consistency" diagnostics are proposed.

TEST 1 - "Equal Sign Test": As a first summary measure, check to see if a series pair move consistently in the same direction, i.e. overall movements have equal sign, irrespective of their magnitudes. This is very general but essential given the diversity of series we wish to test for consistency. This can be tested by comparing the signs of the movements at matched times and testing for the absence of structure in the run of sign differences:

- For example, suppose we have the sequence of movements for series 1 and series 2 at matched times: $\{\mathbf{M_1(t)}\}$, $\{\mathbf{M_2(t)}\}$, if sign[$\mathbf{M_1(t)}$] = sign[$\mathbf{M_2(t)}$] then we set an indicator variable $\mathbf{i(t)}$ = 1, or 0 if sign[$\mathbf{M_1(t)}$] .NE. sign[$\mathbf{M_2(t)}$].

- Given a sequence (for example) $\mathbf{i(t)}$ = 11111110111100111111111111, we can then test for complete uniformity (i.e. a structureless sequence of purely 1's) by computing the ACF at different lags and testing for:

$$H_0 : ACF[i(t)]_{lag\ h} = 1 \quad \text{versus} \quad H_1 : ACF[i(t)]_{lag\ h} = 0$$

- Rejection of the null hypothesis (at some significance level) implies that the series pair do not move consistently in the same direction.

- This test can be applied to movements in each of the three components: $\{\mathbf{O_1, O_2}\}$; $\{\mathbf{SA_1, SA_2}\}$; $\{\mathbf{T_1, T_2}\}$.

- Plots will be generated showing **i(t) versus t** for each of these components and also its **ACF** for various lags with confidence limits.

TEST 2 - "Cross-correlation Magnitude Tests": Supplement the above with a cross-correlation measure between the movements in each paired component: $\{\mathbf{O_1, O_2}\}$; $\{\mathbf{SA_1, SA_2}\}$; $\{\mathbf{T_1, T_2}\}$, i.e. we would compute $\rho(\mathbf{M_1, M_2})_h$ for the interesting lags h = -3, -2, -1, 0, 1, 2, 3. These lags are of interest because they reflect patterns in business cycles. More specifically,

- For lag h = 0: since we expect series pairs to be "conceptually" consistent (on apriori grounds) at lag 0 and since the maximum value of the absolute value of the cross-correlation coefficient is "1", we will be testing the following null hypothesis. This results in a stronger test for consistency.

$$H_0 : \rho(M_1, M_2)_{h=0} = 1 \quad \text{versus} \quad H_1 : \rho(M_1, M_2)_{h=0} < 1$$

Rejection of the null hypothesis (at some significance level) implies that the movement magnitudes in a series pair are not correlated or overall "inconsistent".

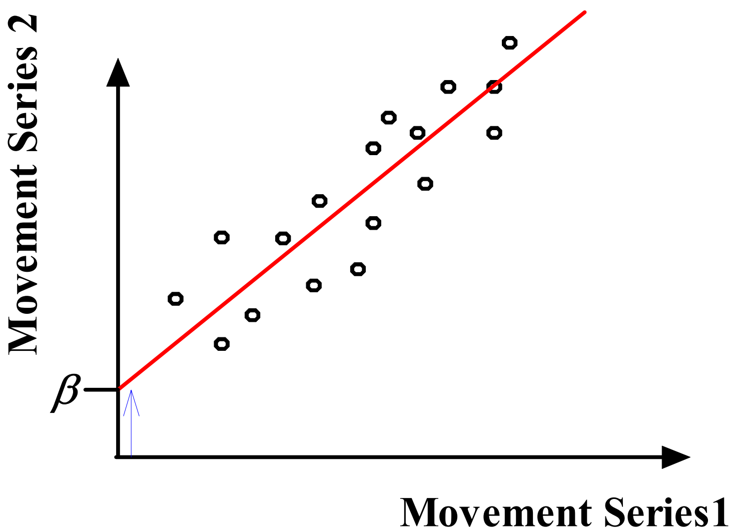
- For lags h .NE. 0, correlations are expected to be small (or insignificant most of the time). In this case we will test the following:

$$H_0 : \rho(M_1, M_2)_{h\neq0} = 0 \quad \text{versus} \quad H_1 : \rho(M_1, M_2)_{h\neq0} \neq 0$$

Rejection of this null hypothesis (at some significance level) implies that the movement magnitudes at the non-zero lag are correlated to some degree or overall, there is some level of "consistency" when the series are shifted with respect to each other.

- Probability (significance) levels will be computed for each correlation measure $\rho(\mathbf{M_1}, \mathbf{M_2})_h$. These measures will be plotted against the lag h.

TEST 3 - "Equal Movement Magnitude Test": The above cross-correlation measures, on their own, are independent of any relative differences in the overall movement magnitudes. This can be seen by noting that a correlation between any two variables does not imply that a straight line regression fit has zero intercept and unit slope:



The straight line regression fit can be parameterised as:

$$M_2 = \alpha M_1 + \beta,$$

where α is the slope and β is the intercept or distance from the horizontal axis along $\mathbf{M_1 = 0}$. A significantly non-zero value for β implies an overall difference (hence inconsistency) in the movement magnitudes for a series pair. Thus, we can test for the following:

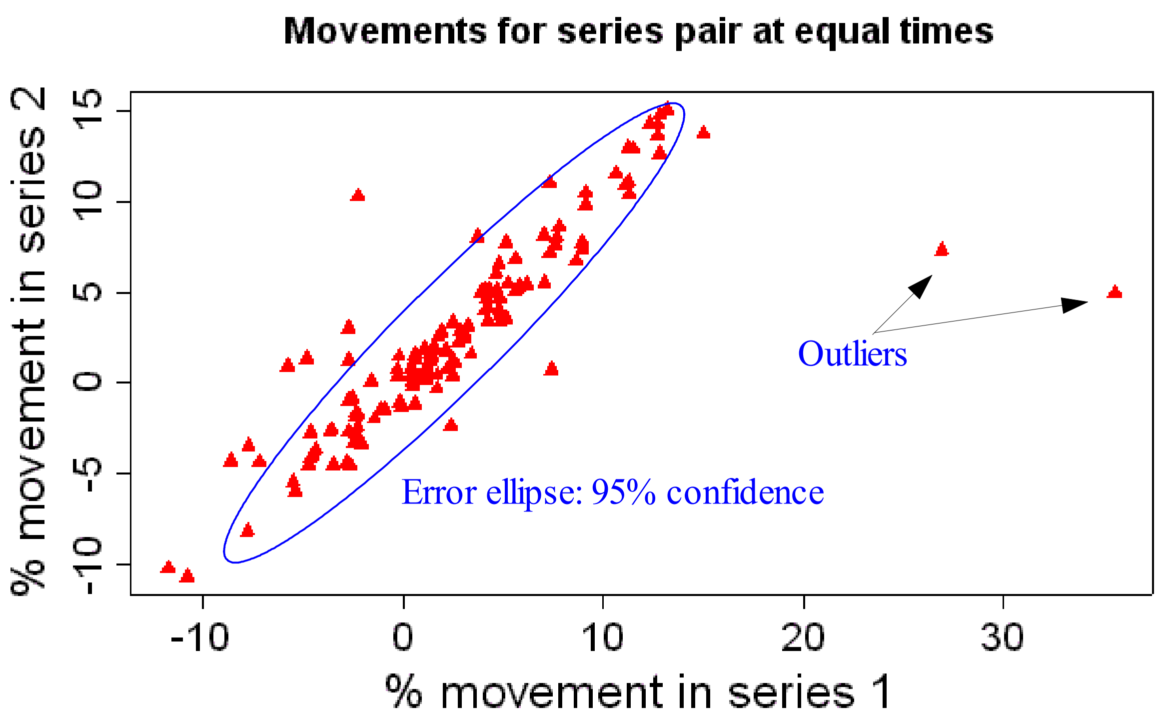
$$H_0 : \beta = 0 \quad \text{versus} \quad H_1 : \beta \neq 0$$

- Acceptance of the null hypothesis therefore implies that the intercept is consistent with zero and hence movement magnitudes are overall of the same magnitude.

- The significance for zero slope will come from the standard errors on the parameters of a linear regression fit.

TEST 4 - "Spot-Outlier Movement Test": The basis of this test involves searching for outliers in a two-dimensional movement versus movement plot (see figure below).

- More specifically, we will attempt to seach for outlying, incompatible movements at specific time points. These are represented by points lying outside some standard error ellipse defined by some pre-determined confidence level in $\mathbf{M_1}$ **and** $\mathbf{M_2}$ jointly:



- Outliers in this plot indicate unequal movement magnitudes and could arise from intrinsic differences in the original series (e.g., outliers, breaks etc..) or, if they occur in movements between the seasonally adjusted or trend component series, could indicate differences in prior correction factors.

- This test can be automated in the sense that if any outliers are detected, a flag is set to indicate that this test "failed" and that the user should review the plot.

- In the production environment (e.g., under SEASABS), the goal would be to click on an outlier in the movment plot and all prior factors and series values would pop up for the two series.

Tying it all together..?

- So, for a given pair of series that we suspect to be "consistent", we can apply the above four tests to potentially each of the component pairs:

  Originals: **{O_1, O_2}**;
  Prior-corrected B1 originals: **{$B1_1$, $B1_2$}**;
  Seasonally adjusted: **{SA_1, SA_2}**;
  Trends: **{T_1, T_2}**

The prior-correct B1 pair was just recently added to this list as it represents a different but important step in the processing chain. An inconsistency in this component pair implies a possible inconsistency in the prior corrections early on in processing.

- Consistency checks in each of the above component pairs can be used to quantify the accuracy of different assumptions and methodologies. E.g., revisions and conceptual differences (originals - good for the client), priors and corrections (B1 and Trends), seasonal factors and filter properties (SA) and trend estimation parameters/properties (Trends).

- Note that if the originals are out-right inconsistent, this does not necessarily mean that everything downstream is also inconsistent. Why?

- How do we represent all of the above in a "consistency summary" table? A suggestion is below.

- On the group level, we can have a table stored in the series knowledge that lists all potentially correlated series pairs (first column). Clicking on any of these entries will bring up a consistency summary table with priors and parameters.

- In the subsequent columns, we summarise information on the above four tests for each of series component. "P" stands for "pass" and under this column, we list the above tests 1,2,3 or 4 that passed in confirming consistency. "F" stands for "failed" and under this column, we list those tests that failed.

- Clicking on any of these "test" numbers will bring up results of statistical tests and associated diagnotic plots.

- The last column may list a quality flag that assesses the overall level of consistency for the series pair, i.e., a number which encompasses the results of all tests across all four series components. This is still to be devised and we may leave it out since the "red danger flags" will be represented by the "test" numbers in the "F" columns.

	Originals P	Originals F	B1's P	B1's F	SA's P	SA's F	T's P	T's F	Final consistency quality flag?
pair 1	1, 2	3, 4	etc...	etc...					?
pair 2	2, 4	1, 2							
pair 3	1,2,3,4	.							
pair 4	3	1, 2, 3							
.									
.									
.									

A simple tool is currently under development for TSA to use under production. Example output is as follows:
S:\data\BOP\coherence\Frank'sCode\results1.html