A Diagnostic for Seasonality Based Upon Autoregressive Roots

**Keywords:** Autoregressive estimator, Seasonal adjustment, Spectral peaks, Visual Significance

# Introduction

The problem of identifying seasonality in published time series is of enduring importance. Many official time series -- such as gross domestic product (GDP) and unemployment rate data -- have an enormous impact on public policy, and are heavily

scrutinized by economists and journalists. Obscuring the debate is the lack of universally

agreed-upon criteria for detecting seasonality, as well as the different behavior of seasonal patterns in raw versus seasonally adjusted data.

We propose the following verbal definition of seasonality: persistency in a time series over seasonal periods that is not explainable by intervening time periods. For a monthly series with a seasonal period equal to twelve, seasonality is indicated by persistency from year to year that is not explained by month-to-month changes. Note that both parts of this definition are crucial: without seasonal persistency from year to year, no seasonal pattern will be apparent, so this facet is clearly necessary; however, any trending time series also has persistency from year to year, which comes through the intervening months -- we need to screen out such cases. If a time series is covariance stationary, it is natural to parse persistency in terms of autocorrelation. The paper shows that we can adapt persistency to non-integer lags of the autocorrelation function via its decomposition in terms of autoregressive (AR) roots, and examine seasonality of arbitrary frequency

through the modulus and phase of the root. Whereas under-adjustment would be indicated by the presence of AR roots of near-unit magnitude and seasonal phase, over-adjustment corresponds to a negative form of persistency (i.e., negative seasonal autocorrelations) termed anti-persistency, and can be measured through moving average (MA) roots computed from the inverse autocorrelations, i.e., the autocorrelations of the reciprocal of the spectral density.

# Methods

## Foundations

We first derive the connection between a fitted AR model and the facets of the autocovariance function (acf). Our analysis indicates that calculation of the roots of an approximating AR process' autoregressive polynomial yields important information about persistencies in the acf. If there are roots with a magnitude close to unity and with phase close to a seasonal frequency, then seasonal persistence will be present in the acf, and hence the process will exhibit seasonality. Conversely, the absence of such strong seasonal roots indicates that seasonal persistence is either not present at all in the acf, or is masked by other effects, and hence no seasonality will be apparent in the process.

We also present a Moving Average (MA) approximation to the spectral density function (sdf), where unit MA roots correspond to the spectral zeroes. MA roots of magnitude close to unity correspond to troughs in the sdf, or peaks in the inverse spectrum, and therefore are featured in the inverse autocorrelation function (iacf) in mathematically the same way as the AR case. Thus, whereas substantial AR roots describe persistence in a process, substantial MA roots correspond to anti-persistence in a process, i.e., persistence in the iacf as opposed to the acf.

Seasonal adjustment attempts to remove identified seasonality in a process, often through linear filtering. Under-adjustment is an error whereby residual seasonality is left, and this could be identified via looking for seasonal persistence (i.e., substantial seasonal AR roots) in the seasonally adjusted component. Over-adjustment is a less serious error, whereby too much seasonality has been extracted (in fact, other dynamic effects in the time series are fallaciously associated with the seasonal component), and could be identified via looking for seasonal anti-persistence (i.e., substantial seasonal MA roots).

## Estimating the Autoregressive Roots

We derive a central limit theory for the complex roots of an AR estimated via MLE or OLS, with accompanying results for the roots’ magnitude and phase. These results are harvested to obtain a limit theory for the diagnostic, which is based on an examination of the magnitude of roots with phase in the vicinity of a seasonal frequency. The diagnostic can allow for stable or dynamic seasonality, by setting a tuning parameter delta, which corresponds to the magnitude of the seasonal root. Stable seasonality can be tested by setting delta equal to zero, whereas dynamic seasonality can be tested by taking positive values of delta, such as 0.10. We also have analogous results for MA roots based on estimation via MLE or the empirical iacf.

# Results

## Simulations

We simulated an AR(3) process that has a mingling of seasonal and transient dynamics, and investigated the size properties of the AR diagnostic for various sample sizes. We found that the variability in the real root was extremely high for 5 years of data, and there is a high degree of non-normality in the histogram. In contrast, the complex roots have fairly normal-shaped histograms at this small sample size. With 15 years of data the normality has improved, and the size of the test statistics is approaching the nominal level; there seems to be little additional improvement with 20 years of data. In summary, there is some confirmation of the theoretical results, and indication that testing may be conducted with 10 years of data.

## Empirical Illustrations

We studied a time series of the Advance Monthly Retail Trade Report, covering the sample period of January 1992 through December 2012. Focusing on series 442 (Furniture and Home Furnishings Stores), we examined the raw series, its seasonal adjustment (performed by X-12-ARIMA), and its seasonal component. While we expected the raw series to exhibit both seasonal and nonseasonal AR roots, we hoped that the seasonal and nonseasonal components, respectively, possess only such roots, i.e., the seasonal component has only seasonal AR roots, and the nonseasonal component

has only nonseasonal AR roots.

We fitted an AR(23) model to the differenced logged data, and found that many of the roots have a seasonal phase. Because we are testing for seasonality in raw data, we set delta = 0, being interested in the presence of stable seasonality. The results indicate eleven strong roots corresponding to the eleven of the twelve roots of unity; the high p-values indicate a failure to reject the null of seasonality, i.e., these roots are seasonal. There are another twelve roots, some of which have substantial magnitudes, but none of which are seasonal (their p-values are small) excepting the last two -- these are real negative roots associated with the sixth seasonal frequency. However, their moduli are too weak to permit them to be categorized as seasonal.

We considered a seasonal adjustment obtained via application of X-12-ARIMA, which involves identification of outliers, holiday effects, and a SARIMA model for forecast extension of the data. First we examined the seasonal component obtained from X-12-ARIMA, called the seasonal factors when expressed in the original scale. Logging the seasonal factors, we obtain a time series centered around zero, and fit an AR(12) model;

all the eleven seasonal roots were obtained, but in addition found a twelfth root associated with frequency zero. For these tests we took delta = 0, as we expected deterministic seasonality to be present. Failure to reject was indicated for each seasonal root; for the non-seasonal root at frequency zero, the null of seasonality is strongly rejected.

Next, we examined differenced logs of the seasonal adjustment, and fitted an AR(11): none of the roots had phase that was obviously seasonal. All but the first root, which lies at frequency zero, seemed to correspond to roots with similar phase in the original process, although some have been shifted somewhat. Since we are testing seasonally adjusted data, we do not expect deterministic seasonality to be present, and it is appropriate to take a higher value of delta, such as 0.10. One pair of roots, near to the fourth seasonal frequency, exhibit a weak form of seasonality because their phase and moduli sufficiently resemble a strong seasonal root. This identifies a possible concern with the quality of adjustment.

# Conclusions

We have demonstrated a new approach to data analysis and seasonality detection. Inferential theory for AR roots, in terms of magnitude and phase, has been established for stationary processes. We have extensively argued that seasonality corresponds to seasonal persistency, which is metrized through seasonal AR roots. By examining different phase components of the roots, different periodic effects can be simultaneously investigated. For instance, in daily time series we can examine weekly effects (phase of 2 pi/7, 4 pi/7, or 6 pi/7) together with annual effects (phase of 2 pi/365.25, among other integer multiples). When a series is down-sampled, or flow-aggregated, one can easily adapt the phase criteria for seasonality, although we have not mathematically derived how AR roots are altered by such distortions of sampling frequency (this is left for future research).

The AR diagnostic tests are principally useful for detecting under-adjustment in a seasonal adjustment, as well as detecting whether seasonality exists in a raw series. The former exercise is designated as testing for adequacy of seasonally adjusted data, whereas the latter is referred to as pre-testing, i.e., determining whether a given series is a candidate for seasonal adjustment. (Because seasonal adjustment procedures are not idempotent in general, there is a cost associated with seasonally adjusting data that does not warrant such a procedure.) In order to detect over-adjustment, we propose use of the MA diagnostic tests.