Investigating the use of the Lomb-Scargle Periodogram for Heart Rate Variability Quantification

The use of the Lomb-Scargle Periodogram (LSP) for the analysis of biological signal rhythms has been well-documented. 1, 2

“The analysis of time-series of biological data often require special statistical procedures to test for the presence or absence of rhythmic components in noisy data, and to determine the period length of rhythms.” 3

“In the natural sciences, it is common to have incomplete or unevenly sampled time series for a given variable. Determining cycles in such series is not directly possible with methods such as Fast Fourier Transform (FFT) and may require some degree of interpolation to fill in gaps. An alternative is the Lomb-Scargle method (or least-squares spectral analysis, LSSA), which estimates a frequency spectrum based on a least squares fit of sinusoid.” 4

“The standard way to deal with such data is to compute a discrete Fourier transform of the data and then view the transform as an absorption spectrum, a power spectrum, a Schuster periodogram (Schuster 1905), or a Lomb-Scargle periodogram (Lomb 1976, and Scargle 1982 and 1989), see Priestley (1981) and Marple 1987 for a review of classical spectral estimation techniques. The problem with all such techniques is that they have not be derived from any single set of unifying principles that tell one what is the optimal way to estimate the period. In this paper we change that by using Bayesian probability theory to deriving the discrete Fourier transform, the power spectrum, the weighted power spectrum, the Schuster periodogram and the Lomb-Scargle periodogram as special cases of a generalized Lomb-Scargle periodogram, and show that the generalized Lomb-Scargle periodogram is a sufficient statistic for

4 Marc in the box, “Lomb-Scargle periodogram for unevenly sampled time series”
single frequency estimation, (a sufficient statistic summarizes all of the information in the data relevant to the question being asked)."

A key benefit of the LSP is in its ability to process unequally-spaced or missing data in terms of time. This is a distinct advantage over the discrete time Fourier Transform (DTFT). The normalized power is determined from:

\[
PN(\omega) = \frac{1}{2\pi} \left[ \frac{\sum_j (Y_j - \bar{Y}) \cos \omega(t_j - \tau)}{\sum_j \cos^2 \omega(t_j - \tau)} \right]^2 + \left[ \frac{\sum_j (Y_j - \bar{Y}) \sin \omega(t_j - \tau)}{\sum_j \sin^2 \omega(t_j - \tau)} \right]^2,
\]

and

\[
\tau = \left( \frac{1}{2\omega} \right) \tan^{-1} \left[ \frac{\sum_j \sin 2\omega t_j}{\sum_j \cos 2\omega t_j} \right],
\]

where, \( PN(\omega) \) is the normalized power as a function of angular frequency. The relationship between angular (or circular) frequency and frequency in Hz is given by:

\[
f = \frac{\omega}{2\pi}.
\]

The solution of these equations is rather straightforward. Let’s take a look at a simple example and then extrapolate to a more complex case.

Consider following sine wave given in the following diagram. This sine wave is described by the equation:

\[
y(t) = A \sin \omega t, \text{ where } A \text{ is amplitude (assumed to be 1), } \omega \text{ is the radian or circular frequency, and } t \text{ is time.}
\]

The relationship between circular frequency and frequency in Hertz is given above. In the plot of the following figure \( f = 10 \text{Hz} \). Hence, the sine equation becomes:

\[
y(t) = \sin(2\pi \times 10t) = \sin(20\pi t). \text{ This function is plotted in Figure 1.}
\]

---

5 G. Larry Brethorst, "Frequency Estimation And Generalized Lomb-Scargle Periodograms"
The Lomb-Scargle Periodogram for this simple waveform is given by the plot shown in Figure 2.

The maximum is centered about the principle frequency of 10Hz. The method used to calculate the LSP is written in Visual Basic and incorporated as a Macro into an Excel spreadsheet. The code used to create the LSP plot above is included in the table shown in Figure 3 on the next page.
Figure 3: Data table showing LSP output.

The Visual Basic code listing of the LSP algorithm is as follows in Table 1.

**Table 1: Listing of LSP algorithm (visual basic macro).**

```
Sub CalculateLS()
    ' Program to calculate the Lomb-Scargle Periodogram
    ' of a time-based signal in 1 dimension.
    ' Copyright(c) 2015, 2018
    ' John R. Zaleski, Ph.D., CAP, CPHIMS
    ' Free to use with proper attribution to the author.
    Dim Pi As Double
    Dim s As Double
    Dim c As Double
    Dim mean As Double
    Dim numRows As Long
    Dim I As Long
    Dim J As Long
    Dim Ybar As Double
```
Dim variance As Double
Dim n1 As Double
Dim n2 As Double
Dim d1 As Double
Dim d2 As Double
Dim tj As Double
Dim freqIncrements As Double
Dim omega(65535) As Double
Dim f(65535) As Double
Dim t(65535) As Double
Dim tau(65535) As Double
Dim y(65535) As Double
Dim PN(65535) As Double
Dim maxFCount As Long
Dim Maxf As Double
Dim Deltaf As Double
Dim fRow As Integer
Dim sheetName As String

' define frequency increment
freqIncrements = 1000#

' get sheet name
sheetName = ActiveSheet.Name

' Clear contents of output columns and set headers
Worksheets(sheetName).Columns(3).EntireColumn.Clear
Worksheets(sheetName).Columns(9).EntireColumn.Clear

' rotate cell header text upward
,
Worksheets(sheetName).Cells(1, 1).Orientation = xlUpward
Worksheets(sheetName).Cells(1, 2).Orientation = xlUpward
Worksheets(sheetName).Cells(1, 3).Orientation = xlUpward
Worksheets(sheetName).Cells(1, 4).Orientation = xlUpward
Worksheets(sheetName).Cells(1, 5).Orientation = xlUpward
Worksheets(sheetName).Cells(1, 6).Orientation = xlUpward
Worksheets(sheetName).Cells(1, 7).Orientation = xlUpward
Worksheets(sheetName).Cells(1, 8).Orientation = xlUpward
Worksheets(sheetName).Cells(1, 9).Orientation = xlUpward
Worksheets(sheetName).Cells(1, 10).Orientation = xlUpward

' label output cell headers

Worksheets(sheetName).Cells(1, 7) = "Frequency (Hz)"
Worksheets(sheetName).Cells(1, 8) = "w (rad/s)"
Worksheets(sheetName).Cells(1, 9) = "Tau"
Worksheets(sheetName).Cells(1, 10) = "Pn(f)"

' Set column widths for columns C through J

Worksheets(sheetName).Columns("A").ColumnWidth = 6
Worksheets(sheetName).Columns("B").ColumnWidth = 6
Worksheets(sheetName).Columns("C").ColumnWidth = 6
Worksheets(sheetName).Columns("D").ColumnWidth = 6
Worksheets(sheetName).Columns("E").ColumnWidth = 6
Worksheets(sheetName).Columns("F").ColumnWidth = 5
Worksheets(sheetName).Columns("G").ColumnWidth = 5
Worksheets(sheetName).Columns("H").ColumnWidth = 5
Worksheets(sheetName).Columns("I").ColumnWidth = 5
Worksheets(sheetName).Columns("J").ColumnWidth = 5

' Horizontal Center Align columns C through J

Worksheets(sheetName).Columns("A").HorizontalAlignment = xlHAlignCenter
Worksheets(sheetName).Columns("B").HorizontalAlignment = xlHAlignCenter
Worksheets(sheetName).Columns("C").HorizontalAlignment = xlHAlignCenter
Worksheets(sheetName).Columns("D").HorizontalAlignment = xlHAlignCenter
Worksheets(sheetName).Columns("E").HorizontalAlignment = xlHAlignCenter
Worksheets(sheetName).Columns("F").HorizontalAlignment = xlHAlignCenter
Worksheets(sheetName).Columns("G").HorizontalAlignment = xlHAlignCenter
Worksheets(sheetName).Columns("H").HorizontalAlignment = xlHAlignCenter
Worksheets(sheetName).Columns("I").HorizontalAlignment = xlHAlignCenter
Worksheets(sheetName).Columns("J").HorizontalAlignment = xlHAlignCenter

' Vertical Center Align columns C through J
'
Worksheets(sheetName).Columns("A").VerticalAlignment = xlVAlignCenter
Worksheets(sheetName).Columns("B").VerticalAlignment = xlVAlignCenter
Worksheets(sheetName).Columns("C").VerticalAlignment = xlVAlignCenter
Worksheets(sheetName).Columns("D").VerticalAlignment = xlVAlignCenter
Worksheets(sheetName).Columns("E").VerticalAlignment = xlVAlignCenter
Worksheets(sheetName).Columns("F").VerticalAlignment = xlVAlignCenter
Worksheets(sheetName).Columns("G").VerticalAlignment = xlVAlignCenter
Worksheets(sheetName).Columns("H").VerticalAlignment = xlVAlignCenter
Worksheets(sheetName).Columns("I").VerticalAlignment = xlVAlignCenter
Worksheets(sheetName).Columns("J").VerticalAlignment = xlVAlignCenter

' specify number of decimal places on a column basis
'
Worksheets(sheetName).Range("A:A").NumberFormat = "0.00"
Worksheets(sheetName).Range("B:B").NumberFormat = "0"
Worksheets(sheetName).Range("C:C").NumberFormat = "0.00"
Worksheets(sheetName).Range("D:D").NumberFormat = "0.00"
Worksheets(sheetName).Range("E:E").NumberFormat = "0"
Worksheets(sheetName).Range("F:F").NumberFormat = "0"
Worksheets(sheetName).Range("G:G").NumberFormat = "0.000"
Worksheets(sheetName).Range("H:H").NumberFormat = "0.000"
Worksheets(sheetName).Range("I:I").NumberFormat = "0.00"
Worksheets(sheetName).Range("J:J").NumberFormat = "0.00"

' Declare value for Pi
'
Pi = Application.WorksheetFunction.Pi()

' Define the number of processing rows
' as excluding the header, and assuming
' the first value begins at count of zero,
' not one.
'
numRows = worksheets(sheetName).Cells(Rows.Count, 2).End(xlUp).Row - 1
' Calculate mean of signal, column 2. Assign value to column 4, row 1.

mean = Application.WorksheetFunction.Average(Range("B:B"))
Ybar = mean
Worksheets(sheetName).Cells(1, 3) = "Ybar"
Worksheets(sheetName).Cells(2, 3) = Ybar

' Calculate variance of signal, column 2. Assign value to column 4, row 2.

variance = Application.WorksheetFunction.Var_S(Range("B:B"))
Worksheets(sheetName).Cells(1, 4) = "Signal Var"
Worksheets(sheetName).Cells(2, 4) = variance

' Model frequency is contained in column 4, row 3.

' Assign number of data points to column 4, row 4.

Worksheets(sheetName).Cells(1, 5) = "Nrows"
Worksheets(sheetName).Cells(2, 5) = numRows

' Create the frequency column, defined by the sample frequency of the model in cell (3,4). Make the maximum frequency twice the frequency in this cell to provide a better visual spectrum. Make the increment in frequency something large so as to provide better resolution for graphing.

Worksheets(sheetName).Cells(1, 6) = "Max freq (Hz)"
If Worksheets(sheetName).Cells(2, 11) = "" Then
    Maxf = 150
Else
    Maxf = Worksheets(sheetName).Cells(2, 11)
End If
Worksheets(sheetName).Cells(2, 6) = Maxf
Deltaf = Maxf / freqIncrements
Maxf = Maxf * 2#

' Create the frequency column. Initialize the first frequency as
' the frequency increment--cannot make zero, else singularity.
'
maxFCount = Int(Maxf / Deltaf)
If maxFCount > 65535 Then
    maxFCount = 65000
End If

For I = 0 To maxFCount
    f(I) = (I + 1) * Deltaf
Next ' I

' read time & data columns into arrays

For I = 2 To numRows
    t(I) = Worksheets(sheetName).Cells(I, 1)
    y(I) = Worksheets(sheetName).Cells(I, 2)
Next ' I

' Calculate periodogram

For I = 0 To maxFCount
    ' Compute omega.
    omega(I) = 2 * Pi * f(I)
    s = 0
    c = 0
    ' Compute denominators
    For J = 2 To numRows
        tj = t(J)
        s = s + Sin(2 * omega(I) * tj)
        c = c + Cos(2 * omega(I) * tj)
    Next
    ' Compute tau.
\[
\text{tau}(I) = \frac{1}{2 \cdot \omega(I)} \cdot \text{Atn}(s / c)
\]

' Initialize PN(w) summations.

\[
\begin{align*}
n_1 &= 0 \\
n_2 &= 0 \\
d_1 &= 0 \\
d_2 &= 0
\end{align*}
\]

' Compute PN(w). Assign to column 10.

For \( J = 2 \) To numRows
\[
\begin{align*}
t_j &= t(J) \\
y_j &= y(J) \\
n_1 &= n_1 + (y_j - \overline{y}) \cdot \cos(\omega(I) \cdot (t_j - \text{tau}(I))) \\
n_2 &= n_2 + (y_j - \overline{y}) \cdot \sin(\omega(I) \cdot (t_j - \text{tau}(I))) \\
d_1 &= d_1 + (\cos(\omega(I) \cdot (t_j - \text{tau}(I))))^2 \\
d_2 &= d_2 + (\sin(\omega(I) \cdot (t_j - \text{tau}(I))))^2
\end{align*}
\]

Next

\[
\text{PN}(I) = \frac{1}{2 \cdot \text{variance}} \cdot \left( \frac{n_1 \cdot n_1}{d_1} + \frac{n_2 \cdot n_2}{d_2} \right)
\]

Next ' I

' write output to output columns

For \( I = 2 \) To maxFCount
\[
\begin{align*}
\text{Worksheets(sheetName).Cells}(I, 7) &= f(I) \\
\text{Worksheets(sheetName).Cells}(I, 8) &= \omega(I) \\
\text{Worksheets(sheetName).Cells}(I, 9) &= \text{tau}(I) \\
\text{Worksheets(sheetName).Cells}(I, 10) &= \text{PN}(I)
\end{align*}
\]

Next ' I

End Sub

The program may be used to study behavior of time-varying signals, such as heart rate (i.e., R-R interval) or other signals which may possess uneven time intervals or missing data.

Heart rate is calculated from two identical points on the standard electrocardiogram (ECG or EKG). The calculation is such:
where $\Delta t_{RR}$ is the R-R interval. For a heart rate of 80 beats per minute, the R-R interval is computed to be:

$$\Delta t_{RR} = \frac{60}{80} = 0.750s = 750\text{ms}$$

While normal sinuous rhythm is usually between 60 and 100 beats/minute, in patients with problems, the R-R interval can vary. For example, an acceleration of heart rate with inspiration and a slowing with expiration can be representative of an arrhythmia.

Heart Rate Variability (HRV) has been used as an assessment of the autonomic nervous system, based on sympathetic and parasympathetic tone (SNS versus PSNS). High HRV is indicative of parasympathetic tone. Low HRV is indicative of sympathetic tone. Low HRV has been associated with coronary heart disease and those who have had heart attacks. A low HRV among those who have had heart attacks places those individuals at higher risk of death within 3 years.

We can look at some simulated data to gain a better understanding of the measurement of HRV and its visualization. The following plot shows a constant heart rate of 80 beats/minute, as represented by the constant RR interval of 750 milliseconds. The plot of Figure 4 shows the RR interval (vertical axis) as a function of measurement number (horizontal axis).

![ECG waveform R-R interval LSP plot with no periodic behavior exhibited due to a constant signal.]

---


For a constant signal, as per above, the LSP is non-existent (i.e., no periodic behavior). When an artificial amount of variability, corresponding to sinusoidal perturbative behavior, having frequency equal to 10 Hz, is added onto the signal, as shown in the Figure 5, the oscillatory behavior is overlaid on the constant offset:

![Raw Data](image)

**Figure 5: Constant plus 10 Hz signal.**

When the mean signal is subtracted from the oscillatory component, the following LSP is computed, shown in Figure 6:

![Pn(f)](image)

**Figure 6: LSP of 10 Hz sinusoid.**

The plot in this figure illustrates that the oscillatory component is isolated and shown to be 10 Hz, verifying the actual oscillatory component of the perturbative signal.
As the frequency is reduced, this is visible in the raw data plot, wherein $f = 0.5$ Hz, per the plot of Figure 7 below:

![Figure 7: 0.5 Hz sinusoid.](image)

The corresponding LSP is shown in Figure 8:

![Figure 8: LSP of 0.5 Hz sinusoid.](image)

Note, again, that the constant (or D.C.) component has been removed by subtracting the mean value of the raw signal. Hence,

$$y(t) = y_R(t) - \bar{y}, \text{ where } \bar{y} \text{ is the mean or average value of the signal. If we leave the constant component in, we will see a large value of the power expressed at a frequency}$$
of zero. If the amplitude of the perturbation is far less in magnitude than the constant component, then the power of the constant component will be very large when depicted in the power plot as above and the frequency component will not be as visible. For example, suppose a model of the heart rate is represented as follows:

\[ y(t) = HR_{\text{Mean}} + HRV_{\text{Per}} \sin(2\pi t), \]

where \( HR_{\text{Mean}} \) is the average heart rate and \( HRV_{\text{Per}} \) is the amplitude of the variable or perturbative component. If \( HR_{\text{Mean}} >> HRV_{\text{Per}} \), then the LSP will be dominated at zero frequency and the frequency of the perturbative component will be less easy to see in the plot. Hence, the mean component is subtracted off to reveal only the perturbation.

We can also add a small amount of white noise to the signal, as shown in the plot of Figure 9. This will have the effect of simulating artifact or randomness.

![Figure 9: 0.5 Hz sinusoid with gaussian white noise added.](image)

Mathematically, the addition of the white noise to the signal is given by:

\[ y(t) = HR_{\text{Mean}} + HRV_{\text{Per}} \sin(2\pi t) + \nu_{WN}, \]

where \( \nu_{WN} \) is a random variate drawn from a Gaussian distribution having zero mean and some variance. In the case of the plot above, the variance is 0.01. The white noise is additive on top of the original signal. A corresponding plot of the LSP is as shown in Figure 10, illustrating rather well that the white noise has little effect on the determination of the underlying perturbation frequency:
Figure 10: LSP of 0.5 Hz sinusoid with white noise added.

Even when the noise is increased to 0.05, as shown in the following plot, Figure 11, with a concomitant increase in the noise of the signal, the underlying LSP is still unaffected:

Figure 11: 0.5 Hz sinusoid with white noise increased to 0.05.

The corresponding LSP is shown in the following plot, shown in Figure 12:
Figure 12: LSP of 0.5 Hz sinusoid with white noise increased to 0.5.

Summary
The preceding analysis demonstrates a viable method for researching the effects of signal variability. The next step is to determine thresholds in variability that correspond to physiological events, and to identify these events relative to the changes in variability determination.