Intended Learning Outcomes:

On completion of this MATLAB exercise, you should be able to

- Analyze discrete-time signals in the frequency domain
- Implement and analyze several spectral estimation methods using synthetic and real-world signals

Deliverable:

- Each student is required to submit a hardcopy answer sheet which contains answers to the questions in this manual on or before 14 March 2012.

Background:

Spectral analysis [1]-[3] involves determining the distribution of amplitude or power over frequency, associated with a given signal. Periodogram [4] is a conventional method for it and the periodogram of a sequence \( x[n] \) of length \( N \) is defined as

\[
\hat{\phi}_{xx,p}(\omega) = \frac{1}{N} \left| \sum_{n=0}^{N-1} x[n] e^{-j\omega n} \right|^2 = \frac{1}{N} |X(e^{j\omega})|^2
\]

where \( X(e^{j\omega}) \) is the discrete-time Fourier transform (DTFT) of \( x[n] \) with a length of \( N \). The periodogram is a nonparametric spectral analysis approach because no assumptions are made on the observed data sequence. Since \( X(e^{j\omega}) \) is continuous in frequency, it is desirable to have its discrete version for digital signal processing. The discrete version is known as the discrete Fourier transform (DFT) and the DFT of \( x[n] \) is defined as

\[
X[k] = \sum_{n=0}^{N-1} x[n] e^{-j2\pi nk/N}, \quad k = 0, 1, \ldots, N - 1
\]

at frequencies \( \omega_k = 2\pi k / N \). Using the DFT, the sampled version of the periodogram is computed as

\[
\hat{\phi}_{xx,p}[k] = \frac{1}{N} \left| \sum_{n=0}^{N-1} x[n] e^{-j2\pi nk/N} \right|^2, \quad k = 0, 1, \ldots, N - 1
\]

An important topic in spectral analysis is to estimate the frequencies of signals and this problem has many applications such as wireless communications, audio and speech processing, biomedical engineering, power electronics, astronomy and instrumentation and measurement [5]. From the periodogram, the signal frequencies can be estimated from its peaks.
\[ \hat{\omega}_0 = \arg \max_\omega \{ \hat{\phi}_{xx, \rho} (\omega) \} \]

On the other hand, parametric methods, which assume that the signal satisfies a generating model with known functional form, such as modified covariance (MC) [2] and Pisarenko harmonic decomposition (PHD) [3] can be used when the discrete-time signal is a pure tone:

\[ x[n] = \alpha \cos(\omega_0 n + \phi), \quad n = 0, 1, \ldots, N - 1 \]

where \( \alpha > 0 \), \( \omega_0 \in (0, \pi) \) and \( \phi \in [0, 2\pi) \) denote the unknown tone amplitude, normalized radian frequency and phase, respectively. The MC method [2] makes use of the simple recurrence of \( x[n] \):

\[ x[n] = 2\cos(\omega_0) x[n-1] - x[n-2] \]

and the frequency estimate is given as:

\[ \hat{\omega}_0 = \cos^{-1} \left( \frac{\sum_{n=2}^{N-1} x[n-1][x[n] + x[n-2]]}{2 \sum_{n=2}^{N-1} x^2[n-1]} \right) \]

On the other hand, the PHD method utilizes the autocorrelation estimates of \( x[n] \):

\[ r_k = \frac{1}{N-k} \sum_{n=0}^{N-1-k} x[n] x[n+k], \quad k = 0, 1, 2 \]

and the frequency estimate is computed as:

\[ \hat{\omega}_0 = \cos^{-1} \left( \frac{r_2 + \sqrt{r_2^2 + 8r_1^2}}{4r_1} \right) \]

**Procedure:**

1. Let \( x(t) = 10\cos(200\pi t + 1.2) \) which is a continuous-time sinusoid. The \( x(t) \) is sampled every \( T_s \) sec. to obtain a sequence \( x[n] \) where \( x[n] = x(nT_s) \). Answer the following questions:

   (a) Produce \( x[n] \) for \( 0 \leq n \leq 100 \) with \( T_s = 0.001 \) sec. Use `stem` to plot \( x[n] \). Is \( x[n] \) a sinusoidal sequence which corresponds to \( x(t) \)? What is the frequency relationship between \( x(t) \) and \( x[n] \)?

   (b) Create and save a file called “mc.m” which contains

   ```matlab
   function w = mc(x);
   % w = mc(x) is used to estimate frequency
   % the MC method from the vector x
   ```
% x is supposed to be a noisy single-tone sequence
% w is the estimated frequency in radian
N=max(size(x));
t1=0;
t2=0;
for n=3:N
  t1=t1+x(n-1)*(x(n)+x(n-2));
t2=t2+2*(x(n-1))^2;
end
r = t1/t2;
if (r>1)
  r=1;
end
if (r<-1)
  r=-1;
end
w=acos(r);

That is, mc is the MATLAB function for the MC estimator. What is purpose of setting r to 1 and -1 when its value is larger than 1 and smaller than -1, respectively? Determine the frequency estimate for x[n] using mc.

(c) Develop the MATLAB function for the PHD estimator, denoted by phd. Note that xcorr with “unbiased” can be utilized in implementing phd.m. Determine the frequency estimate for x[n] using phd.

(d) With the use of the fft command, develop the MATLAB function to compute the sampled periodogram \( \hat{f}_{xx} \), \( k = 0,1,\cdots,N-1 \), denoted by per. Based on the periodogram peak, determine the frequency estimate for x[n] using per. You may use the command max in the peak search.

(e) Determine the frequency estimates using MC, PHD and sampled periodogram when \( x(t) \) is changed to \( x(t) = 10\cos(267\pi t + 1.2) \).

(f) Based on (b)-(e), discuss the frequency estimation accuracy of MC, PHD and sampled periodogram according to the absolute difference between the true and estimated values.

(g) With the use of zeros command, construct a new sequence y[n] by padding M zeros at the end of x[n] with \( x(t) = 10\cos(267\pi t + 1.2) \). Compute the frequency estimates using the sampled periodogram of y[n] for M = 100, M = 1000 and M = 1000000. Discuss the effect of zero padding in frequency estimation based on the sampled periodogram.

(h) The frequency components of a continuous-time signal can be investigated using its sampled version x with the following MATLAB code:

```matlab
[freq_response,freq_index] = freqz(x,1,N,Fs);
plot(freq_index,abs(freq_response))
```

In doing so, you will see the magnitude plot of the Fourier transform for x versus frequency in Hz. The N and Fs represent the number of samples used in the plot and sampling frequency in Hz, respectively. Note that N should be chosen sufficiently large to ensure a smooth plot. Try the above MATLAB code with N equals 500 to analyze x[n] for \( 0 \leq n \leq 100 \) with
1. A continuous-time sinusoid with radian frequency $\Omega$ is sampled with sampling period $T_s$ to form a discrete-time sinusoid with radian frequency $\omega$. Determine a mathematical relationship between $\Omega$, $\omega$ and $T_s$.

2. Given a discrete-time sinusoid $x[n] = A\cos(\omega n + \phi)$, $n = 0, 1, 2, 3$. With the use of trigonometric identities, simplify the following expression:

$$\frac{x[1](x[2]+x[0])+x[2](x[3]+x[1])}{x^2[1]+x^2[2]}$$

3. In the sunspot cycle estimation, there are large differences in the frequency estimates of MC, PHD and sampled periodogram. Why?

2. An application for spectral analysis in the field of astronomy is to find the sunspot cycle [6]. Download “spot_num.txt” at [6]. Remove the first row and rows after Jan. 2012 and then overwrite the file. Try the following code:

```matlab
load spot_num.txt
ssn = spot_num(:,3);
ssn = ssn - mean(ssn);
```

(a) Describe the usage of each of the three commands.
(b) Compute the frequency estimates for ssn using the MC, PHD and sampled periodogram.
(c) Based on each estimate in (b), determine the sunspot cycle in terms of number of months. Note that the sunspot number sampling interval in “spot_num.txt” is one month.
(d) According to [6], which sunspot cycle estimate(s) is/are close to the expected value?

3. There are 10 data files at [7], namely, dataset1.dat, dataset2.dat,..., dataset10.dat, which contain independent noisy measurements of a single sinusoid.

(a) Download the data files and then compute the frequency estimates from the 10 sequences using the MC, PHD and sampled periodogram. Write down the 30 values with the use of a table.
(b) Then determine the mean and variance of the frequency estimate for each method. Which method gives the smallest variance? Note that var can be utilized.

Questions for Discussion:

1. A continuous-time sinusoid with radian frequency $\Omega$ is sampled with sampling period $T_s$ to form a discrete-time sinusoid with radian frequency $\omega$. From the plot or freq_index and abs(freq_response), determine the peak frequency. For more information regarding the usage of freqz, try help freqz. Try different values of N to study its effects.

$$T_s = 0.001 \text{ sec.}$$ which corresponds to $x(t) = 10\cos(267\pi t + 1.2)$, in the frequency domain.
References: