Exercise 15 – Fourier Analysis. Effects of linear trends. Significance of peaks.

1. Import data from the file 'CO2series.txt'. The series contains monthly measurements of atmospheric CO_2 at the Mauna Loa observatory in Hawaii. The record starts in January of 1959 and continues until the end of 1990. Plot the series and observe the trends. Notice the total number of points.

   The periodicity in the CO_2 concentration is caused by the seasonal cycle: the northern hemisphere has more land mass than the southern hemisphere and this causes the annual growth cycle of the land plants during the northern summer to affects the global CO_2 concentration.

2. Perform a Fourier analysis on the series. Identify the period corresponding to the dominant peak. When plotting the spectrum, omit the zero frequency, which is the time-independent component. Notice that the peak is buried in low-frequency noise.

Example: plotting a Fourier spectrum of series y:

```
% series is contained in variable y

Fs = 12;                    % Sampling frequency (times per year)
T = 1/Fs;                     % Sampling time (in years)
L = max(size(y));                     % Length of signal
t = (0:L-1)*T;                % Time vector

NFFT = 2^nextpow2(L); % Next power of 2 from length of y
Y = fft(y,NFFT)/L;
f = Fs/2*linspace(0,1,NFFT/2);

% Plot amplitude spectrum.
plot(f,2*abs(Y(1:NFFT/2)))
```

If you have the Signal Processing toolbox, you can compare your result with the output of the 'periodogram' command.

3. Make the series more stationary by removing the upward trend. For example, fit the series with a linear function and subtract linear trend. A polynomial fit will yield better results. Also subtract the mean of the series. This will eliminate the large first Fourier component, which results from the non-zero mean.

4. Repeat the Fourier analysis. Observe the changes in the Fourier spectrum.

5. The results of the Fourier analysis on most real-world series are noisy. To discriminate a signal from noise, one can use the following test for the significance of the Fourier peak.

Null hypothesis H_0: power at a given frequency f is attributable to randomness.
Alternative, H_1: The peak in the series is caused by a periodic signal, not noise.

The so-called *g*-test test looks at the Fourier power as variance; it compares the maximum variance (s^2_{max}) at frequency *f* to the *total* variance of the series (s^2), which is equivalent to the un-normalized sum of the powers at all frequencies. Calculate this quantity:

$$g = s^2_{max} / (2s^2)$$

That is, s_{max} is the value of the Fourier power spectrum at the frequency of interest. You can obtain it, for example, by clicking on the Matlab graph. The s^2 is the sum of all values in Y^2 (except the first).

Make sure that, in your calculation, the variances correspond to the Fourier *power* (Y^2) rather than amplitude ($|Y|$). The critical value of g, above which the peak should be considered significant, is

$$g_{crit} = 1 - \exp[\ (\ln p - \ln m) / (m - 1)\ ],$$

where p is the level of significance and m = 0.5*(number of data points in your signal). Because this test looks only at a single frequency, it is rather conservative, as it disregards other frequencies' contributions to the peak.

Perform this test and conclude on the significance of the peaks in your analyses.