CONVOLUTION and CORRELATION

Due date: 4/12/2016

You may have noticed from last week's exercise how long it took to perform the discrete Fourier transform of an array with ∼200,000 elements. This would greatly limit the applicability of the Fourier transforms, as the time cost scales as $\mathcal{O}(N^2)$. Luckily, the Fast Fourier Transform (FFT) comes to rescue with its $\mathcal{O}(N \log_2 N)$ cost.

Because of FFT, a range of operations that would otherwise be prohibitively slow can now be done. Two most notable examples are convolution:

$$r(t) * s(t) = \int_{-\infty}^{\infty} r(t)s(\tau - t)dt,$$

and correlation:

$$r(t) \star s(t) = \int_{-\infty}^{\infty} r(t)s(t + \tau)dt,$$

where τ is the lag parameter. The discrete versions of these two equations are:

$$(r * s)_j = \sum_{k=-N/2+1}^{N/2} s_{j-k} r_k, \quad (r \star s)_j = \sum_{k=0}^{N-1} r_{j+k} s_k.$$

If the signal function s_j is periodic with period N and the response function r_k is *finite* on the $[-N/2, N/2]$ interval, then their convolution and correlation can be computed using FFT:

$$\mathcal{F}(r * s) = \mathcal{F}(r)\mathcal{F}(s), \qquad \mathcal{F}(r \star s) = \mathcal{F}(r)\mathcal{F}(s)^*.$$

Assignment:

a) A study of car density as function of time on the turnpike exit to the Dodgers stadium in LA was performed in a period of roughly 25 weeks, with a 5-min sampling. In addition, the start time, end time and the number of visitors was recorded for the events at the stadium. The data are in files `dodgers.cars.data` and `dodgers.events.data`. Analyze the data and interpret the results.

b) Compute the autocorrelation function of the sound of boiling water. If you are so inclined, acquire your own data, otherwise you can use `boiling.data` from the course homepage. *Hint:* the autocorrelation function may drop rapidly and you should use the log scale to visualize it.

c) Sunspots are closely correlated with the Sun's magnetic activity. Their number has been recorded since 1700 on a yearly basis and since 1749 on a monthly basis. Using FFT, find any periodicity in the data (`sunspots.yearly.data` and `sunspots.monthly.data`). Compute autocorrelation functions for both data-sets and compare them. Is higher sampling of sunspots warranted?

d) Using convolution, predict the shape of a spectral line out of a diffraction-limited spectrograph. *Hint:* show (or assume) that the response function is a gaussian.

e) **Extra credit:** file `google.txt` lists the value of stocks between 2004-2007. Analyze the data. Could you have made any predictions about the coming recession?