Digital Signal Processing - DSP
(PDS - Processamento Digital de Sinais)
Instituto Superior Técnico - 2º semestre 2008/2009
Jorge Salvador Marques

Noise Reduction and Interpolation

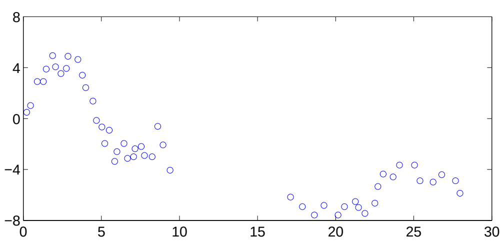
Many signals are corrupted by noise and it is often important to remove or atenuate the noise without destroying the information contained in the signal. One approach to this problem consists of assuming that the signal has a simple structure e.g., it is a linear combination of known basis functions

$$x_i = \sum_{k=1}^{M} c_k \phi_k(t_i) + e_i \tag{1}$$

where t_i are the observation instants, x_i are the noisy samples and e_i is an additive error. Basically we are assuming that the signal belongs to a vector space defined by the basis functions. A simple example is a polynomial fit of the noisy data.

The second problem discussed in this work is Interpolation. Interpolation is closely related to noise reduction. In this case, we want to reconstruct the signal from a set of observations with missing data. This is a noise reduction problem in which the noise is so large (infinite) that no information is obtained during a time interval. Interpolation can be solved by the same techniques.

This work discusses the approximation for the noisy data shown in the figure by two types of models: polymonials and spline functions. We first study the data fit in the interval $[0, 10]$ in which we have noisy observations. Then we compute a fit for the whole interval $[0, 30]$ to account for the interpolation problem as well.



We assume that we know two data sequences $t = [t_1, \dots t_N], x = [x_1, \dots, x_N]$ and estimate the coefficients c_k by the least squares method, i.e., by minimizing a quadratic error energy

$$E(c) = \sum_{i=1}^{N} e_i^2 \tag{2}$$

Since the model is linear in the unknwon coefficients c_k, his leads to a set of linear equations

$$Mc = m \tag{3}$$

which can be easily solved.

Data fit with polynomials and splines

1. Read the vectors t, x in the file *data.m*, using the Matlab command *load data*. Visualize the signals.

2. Consider only the first 30 samples of the vectors t, x (data before the gap) and approximate them by a 4rth order polynomial using the least squares method. You should compute a linear set of equations $Mc = m$ for the least squares estimate of the coefficients and solve it. Visualize the results.

3. Perform a polynomial fit of higer order, e.g., order 11, and check if the method still performs well.

4. Consider a set of spline functions (triangular pulses): $\phi_i(t) = (1-|t-i|)$ if $i-1 \leq t \leq i+1$, $\phi(t) = 0$ otherwise. Compute the least squares fit in the interval [0,10] approximating the data by a linear combination of spline functions. Consider the spline functions centered at $0, 1, \ldots, 10$. Visualize the results.

5. Repeat the previous fits for the whole t, x sequences. If you find any difficulties try to understand what is going on. Visualize the results.