Problem set #4

Least Squares

1. *[line fit]* We wish to fit a straight line to a set of data points $(x_1, y_1), \ldots, (x_n, y_n)$. Assume that the straight line is defined by

$$y_i = ax_i + b + e_i$$

where $\theta = (a, b)$ are the parameters to be estimated and e_i is the error of the $i - th$ point.

a) Define the least squares energy $E(\theta)$ and
b) determine a linear set of equations for the estimate of θ

2. *[polynomial fit]* Extend the previous problem to a polynomial fit of order p

$$y_i = c_p x_i^p + c_{p-1} x_i^{p-1} + \cdots + c_0 + e_i$$

with coefficients $c_p, c_{p-1}, \ldots, c_0$. Determine the least squares energy and a linear set of equations for the least squares estimate of the coefficients.

3. *[polynomial fit]* Express the error of the polynomial approximation as (see previous problem)

$$e = y - Ac$$

where $e = [e_1 \ldots e_n]^T, y = [y_1 \ldots y_n]^T, c = [c_0 \ldots c_p]^T$ and A is a $n \times p$ matrix. Determine

a) matrix A
b) the least squares energy $E(c)$ as a function of A and c
b) the linear set of equations obtained by minimizing $E(c)$

4. * *[line fit]* The line model $y_i = ax_i + b$ is non symmetric. It assumes that all the uncertainty is due to errors in the y component. A better model is

$$x_i \cos\theta + y_i \sin\theta - r = e_i$$

where $[\cos\theta \ \sin\theta]^T$ is a unit vector orthogonal to the line and r is the distance of the line to the origin. Determine the model which minimizes the least squares criterion by solving a linear set of equations.

Hint: the unknown variables can be reparametrized as $a = \cos\theta, b = sin\theta, r$.

5. *[sinusoidal fit]* Consider a set of data points $(x_1, y_1), \ldots, (x_n, y_n)$. We wish to estimate the parameters of a sinusoidal model

$$y_i = A\cos(x_i + \phi) + e_i$$

using the least squares method. Obtain a linear set of equations for this problem.

6. *[prediction]* Consider n samples of the signal $y(1), \dots, y(N)$. Determine the coefficients of the linear predictor

$$y(n) = a_1 y(n-1) + \cdots + a_p y(n-p) + e(n)$$

using the least squares method.

7. *[system identification]* Consider n samples of the input and output $x(1), \dots, x(N), y(1), \dots, y(N)$ of a linear time invariant system described by a difference equation

$$y(n) - a_1 y(n-1) - \cdots - a_p y(n-p) = b_0 x(n) + b_1 x(n-1) + \cdots + b_q x(n-q) + e(n)$$

Determine the unknown parameters $\theta = [a_1, \dots a_p, b_0, \dots, b_q]^T$ using the least squares method.

* more difficult.

Divirtam-se!
Jorge Salvador Marques