Least Square Approximation

by

Jens Hee
https://jenshee.dk

November 2020

Change log

10. November 2020

1. Rewritten

Contents

Chapter 1

Introduction

In the following Linear and non linear least square methods are described. If the problem is linear it can be solved by solving at set of linear equations. For non-linear problems iterative methods are are used. Several examples are given as well as C-code for solving the problems

Approximation is the problem of finding a function that in some sense fits a given data set. Sometimes a minimax approximation is required, but often a least square approximation is used because the minimax problem is too hard to solve. The minimax problem is not covered here. The least square problem can be formulated as finding a set of $\bar{\boldsymbol{a}}$= $(a_0..., a_{K-1})$ that minimizes:

$$m(\bar{\boldsymbol{a}}) = \sum_{j=0}^{N-1} w_j |f(\bar{\boldsymbol{a}}, x_j) - y_j|^2$$

y_j are the values to be approximated and w_j is a weighting function.

Letting:

$$\frac{\partial m(\bar{\boldsymbol{a}}, x_j)}{\partial a_k} = 0$$

we have K equations for finding the K values of $\bar{\boldsymbol{a}}$:

$$Re \sum_{j=0}^{N-1} w_j (f^*(\bar{\boldsymbol{a}}, x_j) - y_j^*) \frac{\partial f(\bar{\boldsymbol{a}}, x_j)}{\partial a_k} = 0, \qquad k = 0, ..., K-1$$

For real data the equations becomes:

$$\sum_{j=0}^{N-1} w_j (f(\bar{\boldsymbol{a}}, x_j) - y_j) \frac{\partial f(\bar{\boldsymbol{a}}, x_j)}{\partial a_k} = 0, \qquad k = 0, ..., K-1$$

Chapter 2

Linear approximation

If f is linear in a_i then:

$$f(\bar{\boldsymbol{a}}, x) = \sum_{i=0}^{K-1} a_i f_i(x)$$

and the K equations are given by:

$$\sum_{j=0}^{N-1} w_j (\sum_{i=0}^{K-1} a_i f_i(x_j) - y_j) f_k(x_j) = 0$$

or

$$\sum_{i=0}^{K-1} a_i \sum_{j=0}^{N-1} w_j f_k(x_j) f_i(x_j) = \sum_{j=0}^{K-1} w_j f_k(x_j) y_j$$

2.1 Polynomial fit

In this case:

$$f(x) = \sum_{i=0}^{K-1} a_i x^i$$

and the set of linear equations becomes:

$$\sum_{i=0}^{K-1} a_i \sum_{j=0}^{N-1} w_j x_j^k x_j^i = \sum_{j=0}^{N-1} w_j x_j^k y_j$$

2.1.1 Linear regression

If f can be written:

$$f(x) = a_0 + a_1 x$$

then

$$\sum_j (a_0 + a_1 x_j) = \sum_j y_j$$

$$\sum_j (a_0 x_j + a_1 x_j^2) = \sum_j x_j y_j$$

and

$$a_1 = \frac{N \sum_j x_j y_j - \sum_j x_j \sum_j y_j}{N \sum_j x_j^2 - (\sum_j x_j)^2}$$

$$a_0 = \frac{1}{N}(\sum_j y_j - a_1 \sum_j x_j)$$

where N is the number of values to be approximated.

2.2 Exponential fit example

If f can be written:

$$f(x) = a_0 x + a_1 e^{-0.2x}$$

then

$$\sum_j (a_0 x_j^2 + a_1 x_j e^{-0.2x_j}) = \sum_j x_j y_j$$

$$\sum_j (a_0 x_j e^{-0.2x_j} + a_1 e^{-0.4x_j}) = \sum_j e^{-0.2x_j} y_j$$

An example is given in Figure 2.1.

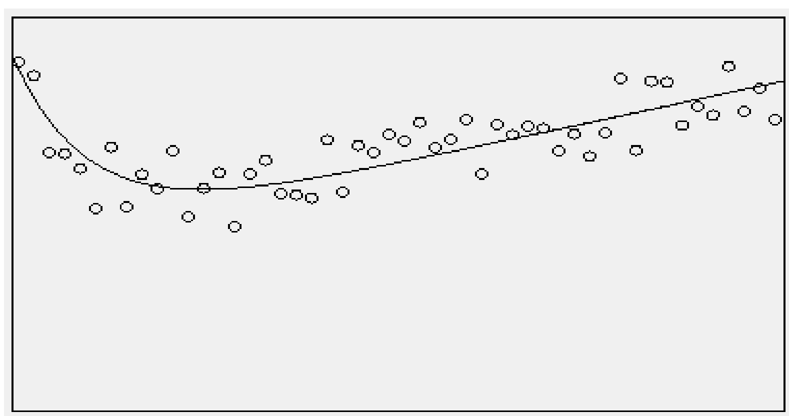

Figure 2.1: Exponential fit example

2.3 Low pass FIR filter design example

If a linear phase FIR filter with length L= 2P+1 is given by

$$H(z) = \sum_{i=0}^{L-1} b_i z^{-i}$$

$$b_i = b_{N-1-i}$$

then

$$H(e^{j\omega}) = \sum_{i=0}^{L-1} b_i e^{-j\omega i} = e^{-j\omega P}(b_P + 2\sum_{i=0}^{P-1} b_i \cos(\omega(i-P))) = \sum_{i=0}^{P} a_i f_i(\omega)$$

where

$$a_i = b_i$$

$$f_i(\omega) = \begin{cases} 2e^{-j\omega P}cos(\omega(i-P)) & \text{for } 0 \leq i \leq P-1 \\ e^{-j\omega P} & \text{for } i = P \end{cases}$$

If the desired response is given by:

$$y_j = \begin{cases} e^{-j\omega_j P} & \text{for } 0 \leq \omega \leq \frac{\pi}{3} \\ 0 & \text{for } \frac{2\pi}{3} \leq \omega \leq \pi \end{cases}$$

and $L = 25$
then the response is shown in Figure 2.2.

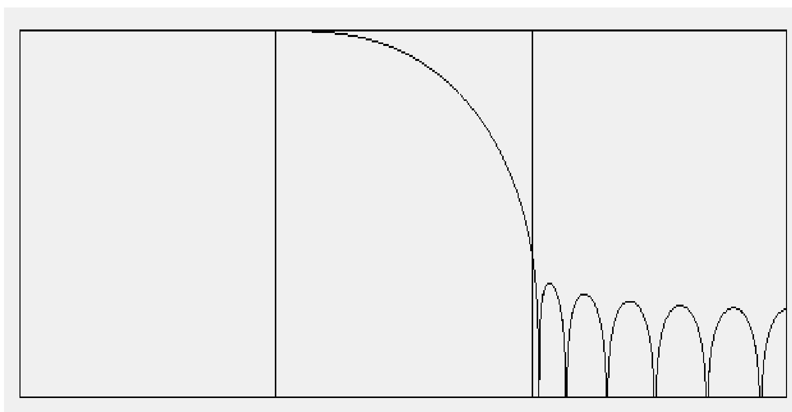

Figure 2.2: LP filter example

Chapter 3

Non-linear approximation

$$m(\bar{\boldsymbol{a}}) = \sum_j |f(\bar{\boldsymbol{a}}, x_j) - y_j)|^2$$

$$\frac{\partial m(\bar{\boldsymbol{a}}, x_j)}{\partial a_k} = \sum_j ((f^*(\bar{\boldsymbol{a}}, x_j) - y_j^*)\frac{\partial f(\bar{\boldsymbol{a}}, x_j)}{\partial a_k} + (f(\bar{\boldsymbol{a}}, x_j) - y_j)\frac{\partial f^*(\bar{\boldsymbol{a}}, x_j)}{\partial a_k}) =$$

$$2Re\sum_j (f^*(\bar{\boldsymbol{a}}, x_j) - y_j^*)\frac{\partial f(\bar{\boldsymbol{a}}, x_j)}{\partial a_k} = 0$$

f can be approximated by:

$$f(\bar{\boldsymbol{a}}, x) = f_0(x) + \sum_i \frac{\partial f(\bar{\boldsymbol{a}}, x)}{\partial a_i}\Delta a_i$$

$$\sum_j Re((f_0^*(x_j) + \sum_i \frac{\partial f^*(\bar{\boldsymbol{a}}, x_j)}{\partial a_i}\Delta a_i - y_j)\frac{\partial f(\bar{\boldsymbol{a}}, x_j)}{\partial a_k}) = 0$$

$$\sum_i \Delta a_i \sum_j Re(\frac{\partial f^*(\bar{\boldsymbol{a}}, x_j)}{\partial a_k}\frac{\partial f(\bar{\boldsymbol{a}}, x_j)}{\partial a_i}) = \sum_j Re(\frac{\partial f(\bar{\boldsymbol{a}}, x_j)}{\partial a_k}(y_j - f_0(x_j))$$

$$\bar{\boldsymbol{a}}_n = \bar{\boldsymbol{a}}_{n-1} + \alpha\Delta\bar{\boldsymbol{a}}$$

This leads to slow convergence. Levenberg-Marquard solves the problem:

$$\sum_i \Delta a_i \sum_j Re(\frac{\partial f^*(\bar{\boldsymbol{a}}, x_j)}{\partial a_k}\frac{\partial f(\bar{\boldsymbol{a}}, x_j)}{\partial a_i}) + \gamma I = \sum_j Re(\frac{\partial f(\bar{\boldsymbol{a}}, x_j)}{\partial a_k}(y_j - f_0(x_j))$$

3.1 Power values approximation

$$\frac{d}{da_k}\sum_j W(j)f(j;\bar{\boldsymbol{a}})^2 = 0$$

$$\sum_j W(j)f(j;\bar{\boldsymbol{a}})\frac{\partial f(j;\bar{\boldsymbol{a}})}{\partial a_k} = 0$$

f can be approcimated by:

$$f(j;\bar{\boldsymbol{a}}) \approx f_0(j) + \sum_i \frac{\partial f(j;\bar{\boldsymbol{a}})}{\partial a_i}\Delta a_i$$

$$\sum_j W(j)(f_0(j) + \sum_i \frac{\partial f(j;\bar{\boldsymbol{a}})}{\partial a_i}\Delta a_i)\frac{\partial f(j;\bar{\boldsymbol{a}})}{\partial a_k} = 0$$

$$\sum_j W(j) \sum_i \frac{\partial f(j;\bar{\boldsymbol{a}})}{\partial a_i}\Delta a_i \frac{\partial f(j;\bar{\boldsymbol{a}})}{\partial a_k} = -\sum_j W(j) f_0(j) \frac{\partial f(j;\bar{\boldsymbol{a}})}{\partial a_k}$$

$$\sum_i \Delta a_i \sum_j W(j) \frac{\partial f(j;\bar{\boldsymbol{a}})}{\partial a_i} \frac{\partial f(j;\bar{\boldsymbol{a}})}{\partial a_k} = -\sum_j W(j) f_0(j) \frac{\partial f(j;\bar{\boldsymbol{a}})}{\partial a_k}$$

Let:

$$f(j;\bar{\boldsymbol{a}}) = H(j;\bar{\boldsymbol{a}})H^*(j;\bar{\boldsymbol{a}}) - D(j)D(j)^*$$

Then:

$$\frac{\partial f(j;\bar{\boldsymbol{a}})}{\partial a_i} = H(j;\bar{\boldsymbol{a}})\frac{\partial H^*(j;\bar{\boldsymbol{a}})}{\partial a_k} + H^*(j;\bar{\boldsymbol{a}})\frac{\partial H(j;\bar{\boldsymbol{a}})}{\partial a_k} = 2Re(H^*(j;\bar{\boldsymbol{a}})\frac{\partial H(j;\bar{\boldsymbol{a}})}{\partial a_k})$$

If H can be written:

$$H(z) = \frac{B(z)}{A(z)} = \frac{\sum_{i=0}^{N} b_i z^{-i}}{1 + \sum_{i=1}^{M} a_i z^{-i}}$$

then:

$$\frac{\partial H(z;\bar{\boldsymbol{b}},\bar{\boldsymbol{a}})}{\partial c_k} = \begin{cases} \frac{1}{A(z)} z^{-k}, for\ c_k = b_k \\ \\ \frac{-H(z)}{A(z)} z^{-k}, for\ c_k = a_k \end{cases}$$

If H can be written:

$$H(z) = \frac{B(z)}{A(z)} = s\frac{\prod_{i=0}^{N} z^i - z_i}{\prod_{i=1}^{N} z^i - p_i}$$

where $z_i$and p_i are real, then:

$$\frac{\partial H(z;\mathrm{s},\bar{\boldsymbol{z}},\bar{\boldsymbol{p}})}{\partial c_k} = \begin{cases} \frac{H(z)}{s}, for\ c_k = s \\ \\ \frac{-H(z)}{z-z_k}, for\ c_k = z_k \\ \\ \frac{H(z)}{z-p_k}, for\ c_k = p_k \end{cases}$$