Linear interpolation

Outline

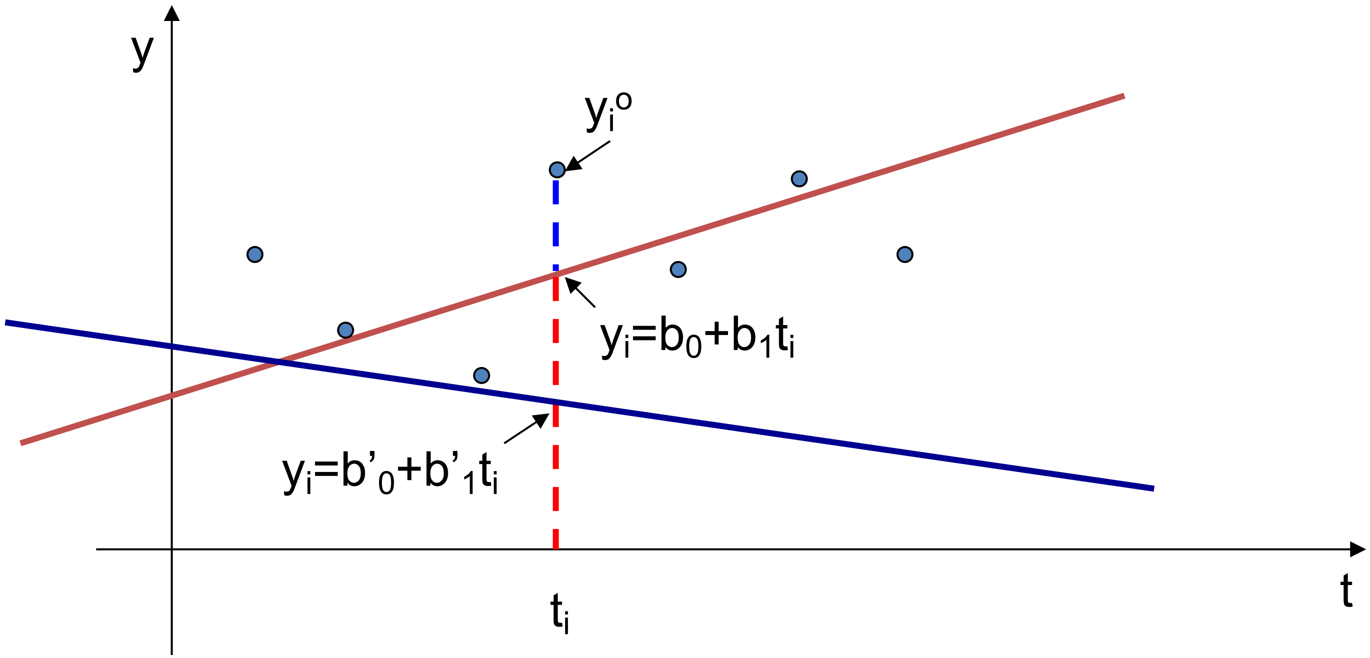
1. The curve fitting problem
2. The parametric model: observation equations
3. The Least Squares principle
4. The Least Squares estimator

Introduction

We want to find a model able to describe, in an approximate way, the behaviour of a certain *variable*, function of time (what we call a signal).
To this aim we sample the variable at a number n of epochs.

By looking at the data we decide to use a straight line (a first degree polynomial). This means that we have to determine ‘in some way’ two parameters.

The observations and two possible linear models



In the picture we show a set of observations and
two possible straight lines (corresponding to two different sets of parameters):
the red one is close to the observations we want to model,
the blue one is far from them.

Notation

We call:
- t_i the known generic observation epoch (i=1,…,n),
- y_i^o the observed value at the epoch t_i of the variable we want to model,
- y_i the unknown value assumed at the epoch t_i by the linear model we are looking for.

The i-th observation (see the previuos picture) can be expressed as the sum of the value assumed by the model plus a residual v_i :

$$y_i^o = y_i + v_i = b_0 + b_1 t_i + v_i$$

Here the chosen model is a straight line depending on the two parameters b_0 and b_1.

Notation

We can write the single observation equation in a vector form as the product of a raw vector containing the known coefficients and a column vector containing the unknown parameters

$$y_i^o = y_i + v_i = b_0 + b_1 t_i + v_i = \begin{bmatrix} 1 & t_i \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} + v_i$$

We then put all the observation equations in a linear system of n equations in m=2 parameters and write the system in a vector form.

Notation

We introduce the following vectors:

- The observation vector y^o containing all the given observations (n rows x 1 column) ,
- the parameter vector $\underline{x}$ containing the m=2 unknown parameters (mx1=2x1),
- the coefficient matrix A (nxm=nx2) containing the coefficients of the parameters in the observation equations
- the residual vector containing the unknown residuals (nx1)

$$\underline{y}^o = \begin{bmatrix} y_1^o \\ y_2^o \\ \cdots \\ y_i^o \\ \cdots \\ y_n^o \end{bmatrix} \qquad \underline{x} = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} \qquad A = \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \cdots & \cdots \\ 1 & t_i \\ \cdots & \cdots \\ 1 & t_n \end{bmatrix} \qquad \underline{v} = \begin{bmatrix} v_1 \\ v_2 \\ \cdots \\ v_i \\ \cdots \\ v_n \end{bmatrix}$$

Notation

So that the system of observation equations can be written as:

$$\underline{y}^{o} = A\underline{x} + \underline{v}$$

Note that, if you observe just two values, then the straight line will be univocally determined as there is only one straight line passing through 2 points (you solve a linear system of n=2 equations in m=2 parameters).

If the number of observations is greater than 2, and the observations do not lie on a straight line, then a criterion must be invoked in order to select one of the infinite possible lines.

Here we use the Least Squares (LS) principle.

The LS principle

We select the straight line (namely the parameters) which makes the sum of the square observation residuals (target function) minimal:

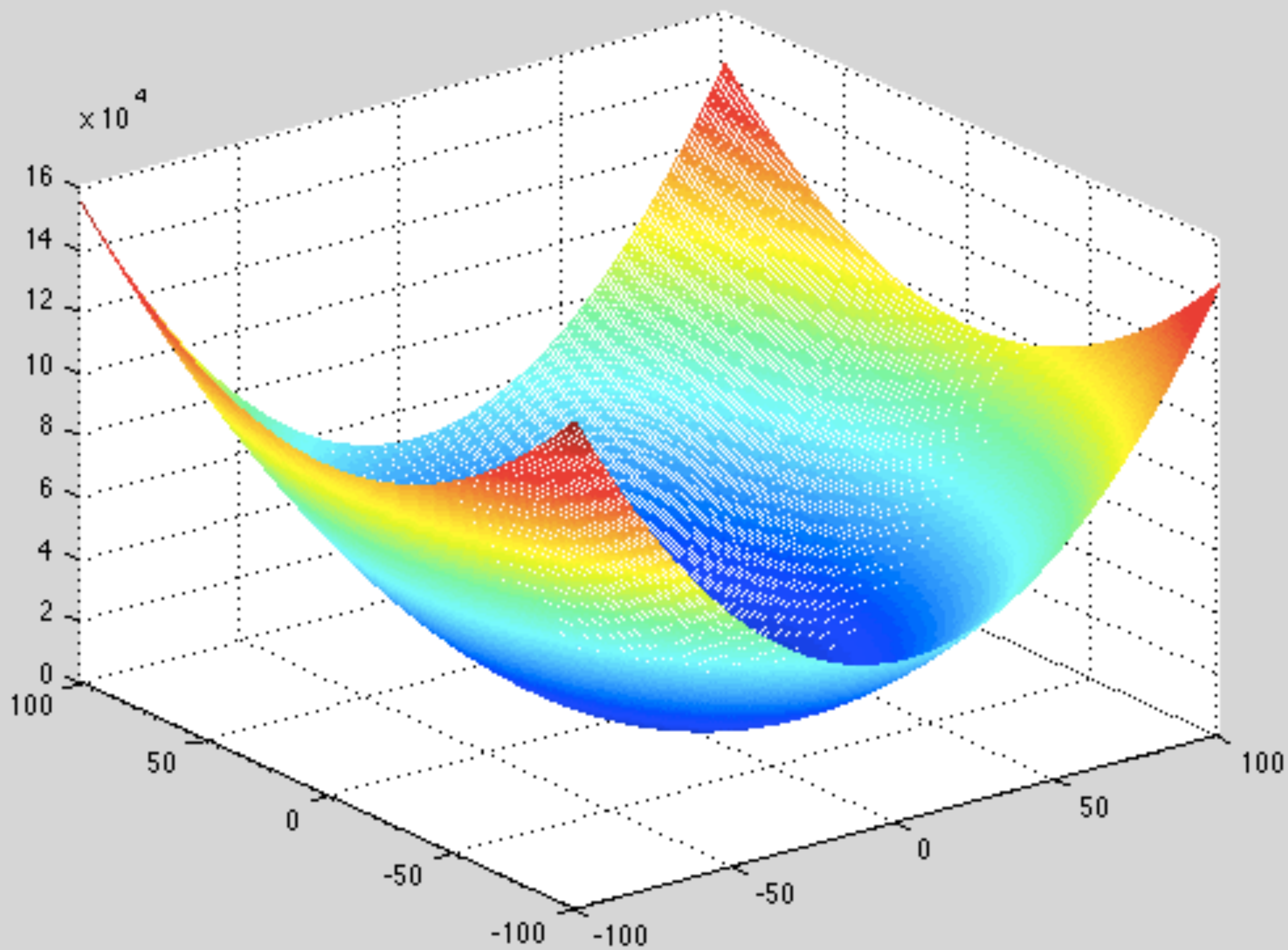
$$\Phi\left(b_0, b_1\right) = \sum_{i=1}^{n} v_i^2 = \sum_{i=1}^{n}\left(y_i^o - y_i\right)^2 = \sum_{i=1}^{n}\left(y_i^o - b_0 - b_1 t_i\right)^2 = \min_{b_0, b_1}$$

In vector form it writes:

$$\Phi\left(\underline{x}\right) = \sum_{i=1}^{n} v_i^2 = \underline{v}'\underline{v} = \left(\underline{y}^o - \underline{y}\right)'\left(\underline{y}^o - \underline{y}\right) =$$

$$= \left(\underline{y}^o - A\underline{x}\right)'\left(\underline{y}^o - A\underline{x}\right) = \underline{y}^{o\,\prime}\underline{y}^o - 2\underline{y}^{o\,\prime}A\underline{x} + \underline{x}'A'A\underline{x} = \min_{\underline{x}}$$

The target function Phi(b_0,b_1)



The LS solution

To find out the minimum of the target function we look for point $\left(\hat{b}_0, \hat{b}_1\right)$ where the plane tangent to the surface is horizontal.
The plane tangent to the target function in $\left(\hat{b}_0, \hat{b}_1\right)$ has equation:

$$p\_t\left(b_0, b_1; \hat{b}_0, \hat{b}_1\right) =$$

$$= \Phi\left(\hat{b}_0, \hat{b}_1\right) + \left.\frac{\partial \Phi\left(b_0, b_1\right)}{\partial b_0}\right|_{\left(\hat{b}_0, \hat{b}_1\right)} \left(b_0 - \hat{b}_0\right) + \left.\frac{\partial \Phi\left(b_0, b_1\right)}{\partial b_1}\right|_{\left(\hat{b}_0, \hat{b}_1\right)} \left(b_1 - \hat{b}_1\right)$$

The LS solution

It is horizontal when

$$\begin{cases} \left.\dfrac{\partial \Phi\left(b_0, b_1\right)}{\partial b_0}\right|_{\left(\hat{b}_0, \hat{b}_1\right)} = 0 \\ \left.\dfrac{\partial \Phi\left(b_0, b_1\right)}{\partial b_1}\right|_{\left(\hat{b}_0, \hat{b}_1\right)} = 0 \end{cases}$$

Namely when

$$\begin{cases} \sum_{i=1}^{n}\left(y_i^o - b_0 - b_1 t_i\right) = 0 \\ \sum_{i=1}^{n}\left(y_i^o - b_0 - b_1 t_i\right) t_i = 0 \end{cases}$$

The LS solution

Which can be written as

$$\begin{cases} \sum_{i=1}^{n} b_0 + \sum_{i=1}^{n} b_1 t_i = \sum_{i=1}^{n} y_i^o \\ \sum_{i=1}^{n} b_0 t_i + \sum_{i=1}^{n} b_1 t_i^2 = \sum_{i=1}^{n} y_i^o t_i \end{cases}$$

$$\begin{cases} b_0 n + b_1 \sum_{i=1}^{n} t_i = \sum_{i=1}^{n} y_i^o \\ b_0 \sum_{i=1}^{n} t_i + b_1 \sum_{i=1}^{n} t_i^2 = \sum_{i=1}^{n} y_i^o t_i \end{cases}$$

$$\begin{bmatrix} n & \sum_{i=1}^{n} t_i \\ \sum_{i=1}^{n} t_i & \sum_{i=1}^{n} t_i^2 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} y_i^0 \\ \sum_{i=1}^{n} y_i^0 t_i \end{bmatrix}$$

The LS solution

This system is the so-called normal system of the LS problem

$$A'A\underline{\hat{x}} = A'\underline{y}^{o}$$

A'A=N is a square symmetric matrix of dimension mxm=2x2

$A'y_o=t_n$ is the normal known terms vector of dimension mx1=2x1

The system solution is

$$\underline{\hat{x}} = \left(A'A\right)^{-1} A'\underline{y}^{o} = N^{-1}A'\underline{y}^{o}$$

A geometrical interpretation

Let's start from the observation equations in vector form:

$$\underline{y}^o = A\underline{\hat{x}} + \underline{v} = \begin{bmatrix} \underline{A}_1 & \underline{A}_2 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} + \underline{v} = \underline{A}_1 b_0 + \underline{A}_2 b_1 + \underline{v}$$

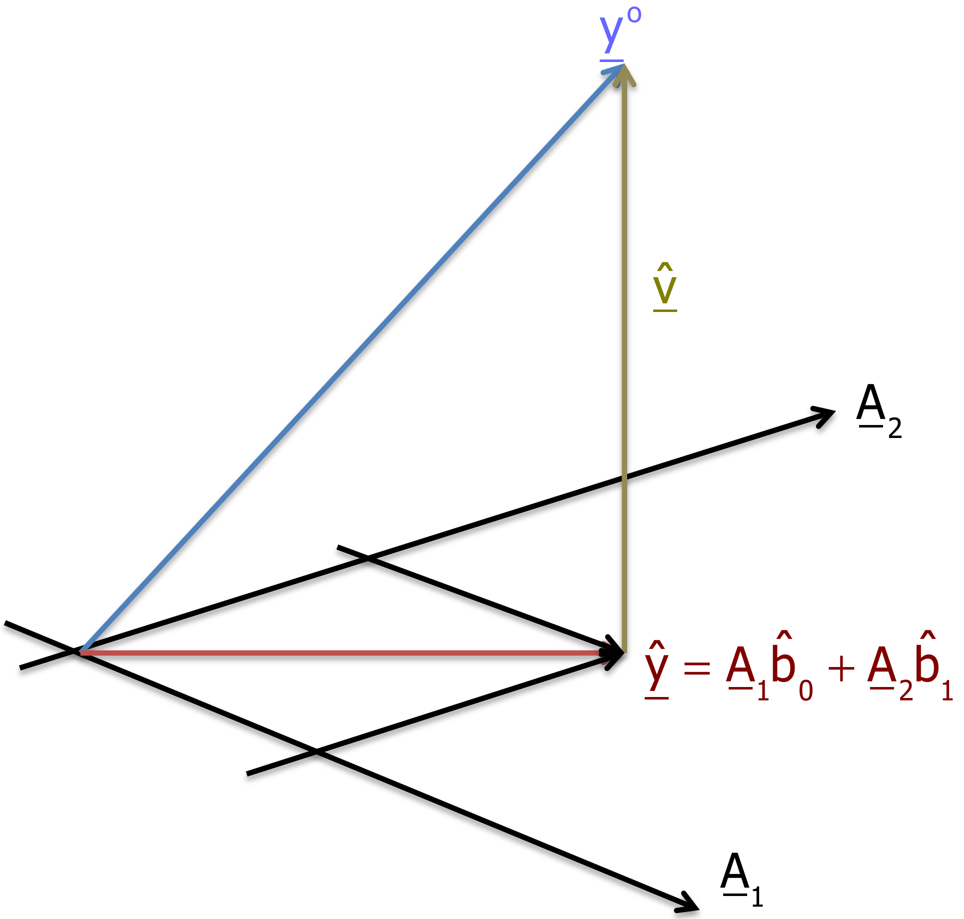
The observation vector is a vector of R^n, obtained as the sum of a vector lying on an hyper-plane, i.e., the linear combination of the column vectors of the A matrix, and another vector of R^n , $\underline{v}$, not contained in the plane.

All vectors contained in the plane are admissible models (admissible straight lines). We look for the vector lying on this plane, which is the closest to the observation vector.

A geometrical interpretation

The observation vector is a vector of R^n, obtained as the sum of a vector lying on an hyper-plane, i.e., the linear combination of the column vectors of the A matrix, and another vector of R^n , $\underline{v}$, not contained in the plane.

All vectors contained in the plane are admissible models (admissible straight lines). We look for the vector lying on this plane, which is the closest to the observation vector.


$\underline{y}^o$
$\hat{\underline{v}}$
$\underline{A}_2$
$\hat{\underline{y}} = \underline{A}_1\hat{b}_0 + \underline{A}_2\hat{b}_1$
$\underline{A}_1$

The vector satisfying the LS condition:

1- a combination of the A matrix column vectors,
2 - such that the residual vector is of minimal norm.

Is in fact the orthogonal projection of the observation vector onto the hyper-plane generated by the column vectors of A.

This corresponds to

$$A'\underline{y}^o = A'\left(A\hat{\underline{x}}\right)$$

exercises

1. Write the scalar product between two vectors u and v. What does it represent from a geometrical point of you?

2. Compute the scalar product between the u=(1,3,4) and v=(1,4,7).

3. What is u’ v?

4. What is u’u?

5. What happens of u’v when u is orthogonal to v? What happens when they are parallel?

6. Write the unit vector lying along u=(1,3,4)

7. What do you obtain from the linear combination of two vectors which are not parallel?

exercises

8. Sketch the following two vectors of R^n : $\underline{u}$=(1,1,1) and $\underline{v}$=(-1,0,1). Are they orthogonal?

9. Represent the plane obtained by a linear combination of the two vectors of exercise 8.
Consider the vector $\underline{y}$=(0,1,2). Does it belong to the above plane? Sketch it.
Consider the vector $\underline{y^o}$=(0.2,0.9,2.1). It does not belong to the above plane. Compute the vector representing the orthogonal projection of this vector on the plane.

10. Compute the residual vector between $\underline{y}^o$ (of exercise 9) and its orthogonal projection $\underline{y}$^ onto the plane generated by u and v. Compute the norm of the residual vector.

11. Why do we look for a model able to approximate our observations?