Lecture Notes on Statistical Analysis of Environmental Data

G. Venuti - D. Triglione

Contents

Chapter 1

Splines

1.1 Definitions

Splines are piecewise polynomials with bounded support, i.e. they are zero outside a bounded closed set. A piecewise polynomial is a function that has a polynomial behavior everywhere, except for a finite number of junction points.

Functions with unbounded support and functions with bounded support, on the real axis R^1 are shown in Fig. 1.1.

In these notes we shall consider only splines with support in R^1 and in R^2 obtained by a linear combination of translated copies of a parent spline on the knots of a regular grid.

a) 1D splines

We shall use the following three types of parent splines in R:

- the zero order spline

$$\varphi_0(\tau) = \begin{cases} 1 & \tau \in [0,1) \\ \\ 0 & otherwise \end{cases} \tag{1.1.1}$$

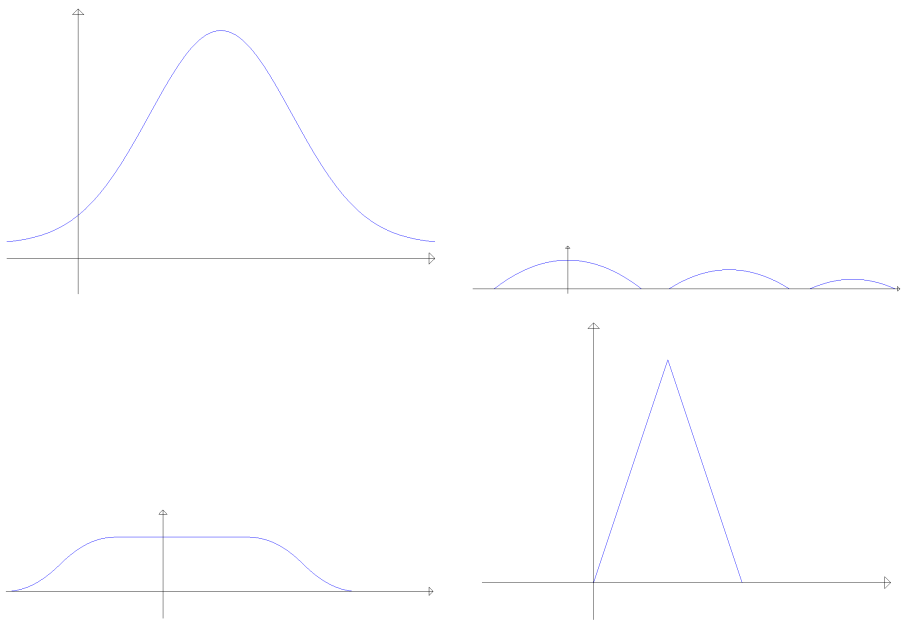

Figure 1.1: a) and b) are functions with unbounded support; c) is a bounded support function; d) is a Spline with support [0,2].

- the first order spline

$$\varphi_1(\tau) = \begin{cases} 1 - |\tau| & |\tau| < 1 \\ \\ 0 & otherwise \end{cases} \tag{1.1.2}$$

- the third order spline

$$\varphi_3(\tau) = \begin{cases} \frac{1}{6}\Big[(2-|\tau|)^3 - 4(1-|\tau|)^3\Big] & |\tau| \leq 1 \\ \frac{1}{6}(2-|\tau|)^3 & 1 < |\tau| < 2 \\ \\ 0 & otherwise \end{cases} \tag{1.1.3}$$

They are represented in Fig. 1.2.

We want to write a 1D spline over a given interval of the real axis.

The first step is to define a grid covering this interval, for instance by choosing the position of the starting knot, its lag size and the total number of knots. Let t_0 be the first grid knot abscissa and Δ the grid lag, all the grid knot abscissas will be identified by an integer number n such that $t_n = t_0 + n\Delta$. The intervals between two consecutive knots can be numbered as well: we shall call $I_n, n = 0, \cdots, N$ the interval between the knots n and $n + 1$.

Once selected the mother spline, one has to modify its support length according to the grid lag size (scale variation) by putting

$$\tau = \frac{t - t_0}{\Delta} \tag{1.1.4}$$

and to create a family of translated copies of the modified mother spline, each related to a different grid knot, to cover the whole grid by putting

$$\tau = \frac{t - t_n}{\Delta} = \frac{t - t_0}{\Delta} - n. \tag{1.1.5}$$

We remark here that all the polynomials involved in the construction of a spline are splines as well. In the sequel, we shall call spline ϕ the single

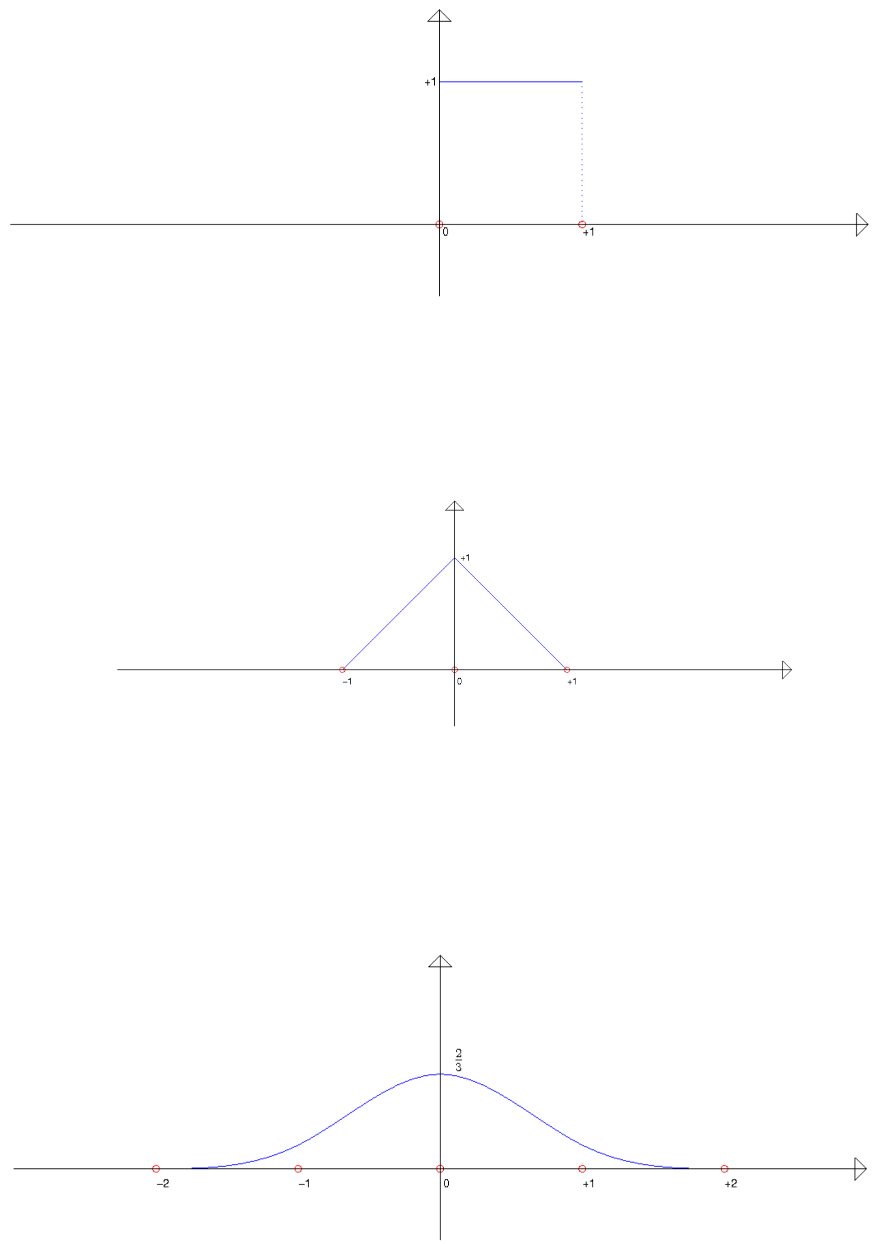

Figure 1.2: The parent splines $\varphi_0(\tau)$, $\varphi_1(\tau)$, $\varphi_3(\tau)$.

copy of the parent spline φ and spline signal S the linear combination of the translated copies.

The family of splines of order $\alpha = 0, 1, 3$ associated to a given grid writes:

$$\varphi_\alpha\left(\frac{t-t_n}{\Delta}\right) = \varphi_\alpha\left(\frac{t-t_0}{\Delta} - n\right). \tag{1.1.6}$$

The spline signal is then obtained by the following linear combination:

$$S_\alpha(t; t_0, \Delta, N(\alpha)) = \sum_{n=0}^{n=N(\alpha)} a_n \varphi_\alpha\left(\frac{t-t_0}{\Delta} - n\right) \tag{1.1.7}$$

where $N(\alpha = 0) = N - 1$ while $N(\alpha = 1) = N(\alpha = 3) = N$.

In the following, without losing generality we shall assume $\Delta = 1$ and $t_0 = 0$.

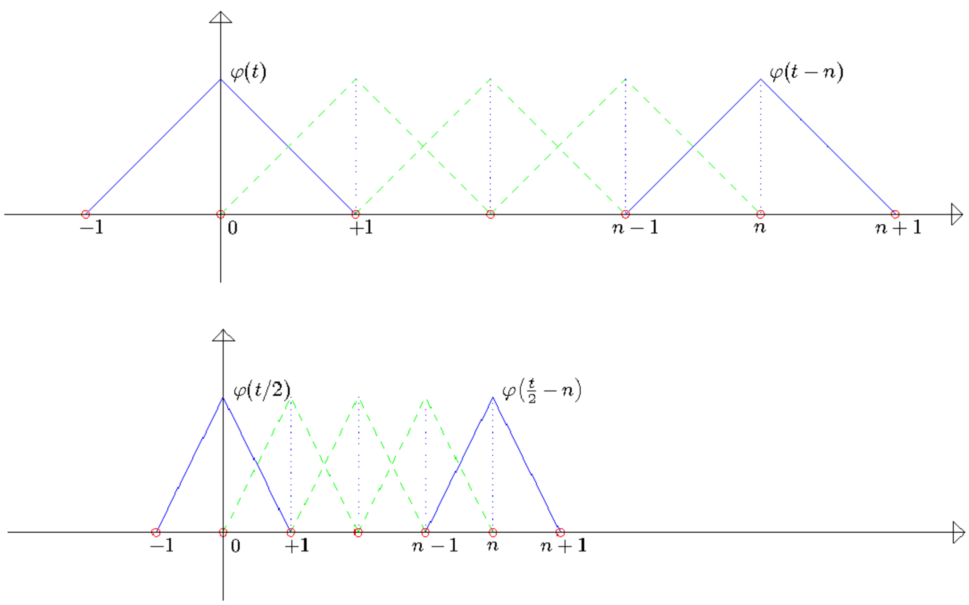

Figure 1.3: a) the family of linear splines $\varphi_1(t-n)$; b) the family of linear splines $\varphi_1(2t-n), n = 0, \cdots, N$

.

b) Spline properties

Zero order splines: Zero order splines $\varphi_0(t-n),\ n=0,\cdots,N-1$ have their support on a single grid interval $[n,n+1)=I_n$, where they are constant $\varphi_0(t-n)=1,\ \forall t\in I_n$. Each polynomial has two discontinuities, one at n and one at $n+1$. Moreover, the family of such polynomials have no overlapping. It results that if we linearly combine those spline copies

$$S_0(t)=\sum_{n=0}^{N-1} a_n\varphi_0(t-n) \tag{1.1.8}$$

we get a piecewise constant function, discontinuous across the grid knots, assuming the value a_n in the corresponding interval I_n, $n=0,\cdots,N-1$ (cfr. Fig. 1.4). It has to be noted that due to the fact that we tag the intervals with an index representing the left extreme, our N functions $\varphi(t-n)$ are indexed from 0 to $N-1$.

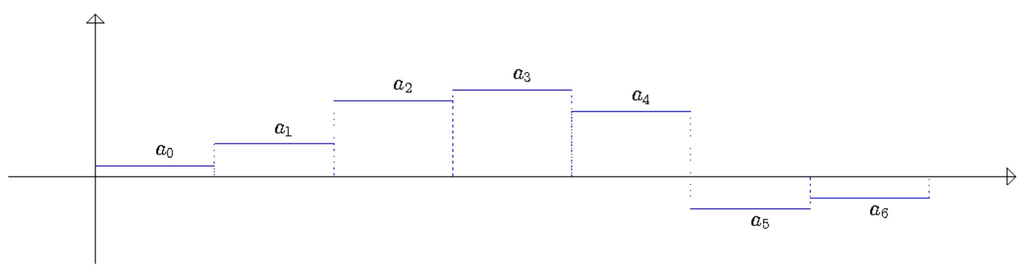

Figure 1.4: A staircase signal generated from $\varphi_0(\tau)$.

First order (or linear) splines: First order splines $\varphi_1(t-n),\ n=0,\cdots,N$ have their support between $[n-1,n+1)=I_{n-1}\cup I_n$. As we translate the mother spline just of one grid lag, two consecutive splines have one lag in common. Therefore, in each interval I_n, the spline signal will be the combination of the two splines $\varphi(t-n)$ and $\varphi(t-n-1)$. More precisely we have:

$$\begin{aligned} t\in I_n,\ S_1(t) &= a_n\ \varphi_1(t-n)+a_{n+1}\ \varphi_1(t-n-1)= \\ &= a_n(1-t+n)+a_{n+1}(1+t-n-1)= \\ &= (a_{n+1}-a_n)(t-n)+a_n\ , \end{aligned} \tag{1.1.9}$$

which is the straight line passing through the points (n, a_n) and $(n+1, a_{n+1})$. Moreover, in each knot n of the grid, only the spline $\varphi_1(t-n)$ is different from 0 and as it is

$$\varphi_1(n-n) = 1 - |n-n| = 1 \quad n = 1, \cdots \ N-1 \tag{1.1.10}$$

here the signal will be equal to

$$S(n) = a_n \varphi_1(n-n) = a_n \ . \tag{1.1.11}$$

Accordingly, $S_1(t)$ is a piecewise linear function connecting the values a_n (cfr. Fig. 1.5).

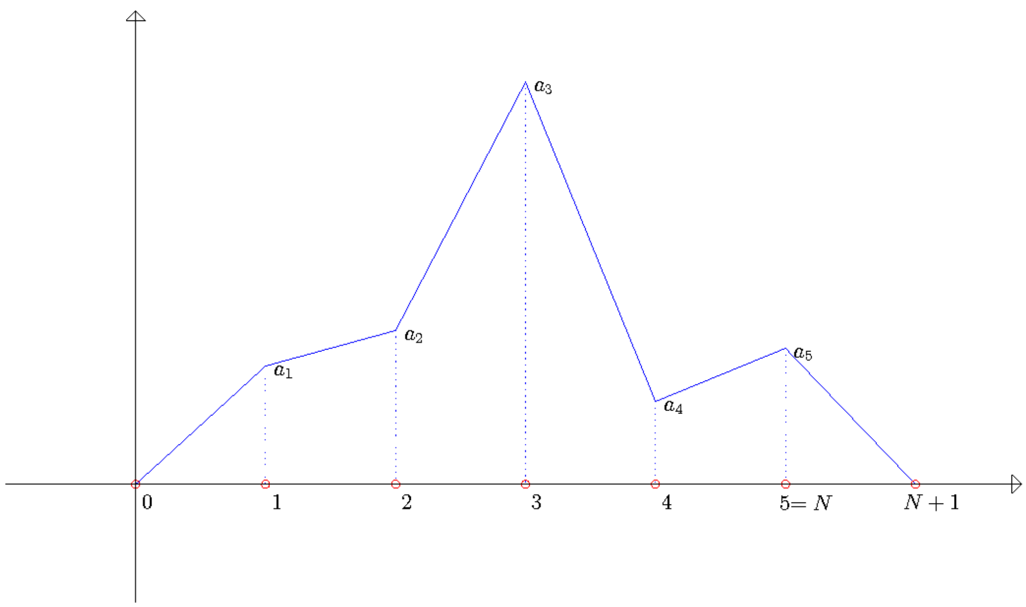

Figure 1.5: A piecewise linear spline obtained by a combination of N linear splines $\varphi_1(t-n), \ n = 1, \cdots, N$

Third order (or cubic) splines: Third order or cubic splines $\varphi_3(t-n), \ n = 0, \cdots, N$ have their support between $[n-2, n+2) = I_{n-2} \cup I_{n-1} \cup I_n \cup I_{n+1}$. By translating the mother spline just of one grid lag, in each grid interval I_n four consecutive splines are different from zero, namely $\varphi_3(t-n+1), \varphi_3(t-n), \varphi_3(t-n-1), \varphi_3(t-n-2)$, as shown in Fig. 1.6.

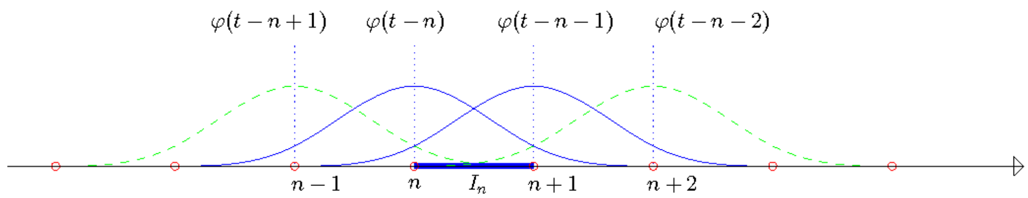

Figure 1.6: The 4 cubic splines different from zero on I_n.

The resulting spline signal in the interval I_n is therefore

$$\begin{aligned} t \in I_n,\ S_3(t) &= a_{n-1}\ \varphi_3(t-n+1) + a_n\ \varphi_3(t-n) + \\ &+ a_{n+1}\ \varphi_3(t-n-1) + a_{n+2}\ \varphi_3(t-n-2)\ . \end{aligned} \tag{1.1.12}$$

Since each of the $\phi_3(t,n)$ is a third order polynomial, formula (1.1.12) shows that $S(t)$ is, on each interval I_n, a third order polynomial, obviously changing analytical form from n to n.

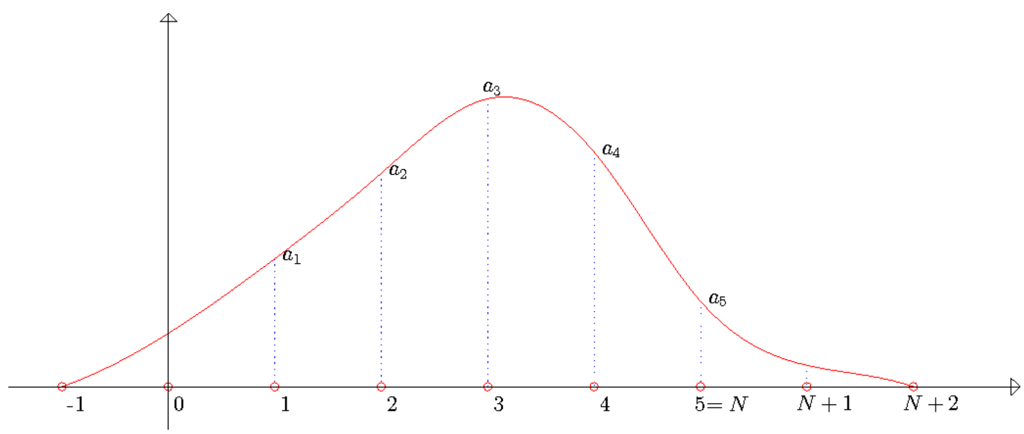

Figure 1.7: A piecewise cubic spline.

Moreover, we observe that, since in each grid knot n

$$\varphi_3(n,n) = \frac{4}{6} \tag{1.1.13}$$

$$\varphi_3(n\pm 1,n) = \frac{1}{6} \tag{1.1.14}$$

we have

$$S_3(n) = \frac{4}{6}\ a_n + \frac{1}{6}\ (a_{n-1} + a_{n+1})\ . \tag{1.1.15}$$

We notice as well that, outside the interval $[1, N]$, $S_3(t)$ goes to zero automatically after 2 knots.

1.2 Least squares spline interpolation

We want to use a spline signal to interpolate a sparse set of observations $\{y_m^o\}, m = 1, \cdots, M$ via least squares. The observations, which are not necessarily regularly distributed, are associated to a set of abscissas $\{t_m^o\}, m = 1, 2, \cdots, M$ referred to as *observation points*.

The first step to solve this problem is to associate a grid to the interval where the observation points are located. The first knot abscissa will be smaller than or equal to the first observation point $t_0 \leq t_1^o$. The choice of the grid lag size is a key issue. From a mere algebraic point of view, as an unknown coefficient is associated to each knot, the number of knots must be smaller than or equal to the number of observations $N \leq M$; moreover, to determine all the parameters from the available observations, one has to guarantee that at least one observation falls in each spline copy $\varphi_\alpha(t - n)$ support.

On the other hand, the grid size constitutes the resolution at which we represent the interpolating signal. This cannot be independent from the sampling rate of the signal itself, that is to an average distance between the observation points, as well as to the accuracies of the observations. Call

$$\delta = \frac{1}{M-1} \sum_{i=1}^{M-1} t_{i+1}^o - t_i^o \tag{1.2.1}$$

the mean distance between the observation points, a grid size too large with respect to this δ, will result in a smooth signal, far form the observations; a grid size comparable with δ (always satisfying the above algebraic constraint) on the contrary, will allow the interpolating signal to be close to the observations, filtering out a small error.

Analogous reasoning can be done on the choice of the spline signal degree.

We will come back again to this point at the end of this chapter, when we will introduce the Tykonov interpolation. For the moment we assume that the grid has been defined and therefore the number of unknown parameters and describe the least squares model associated to the three different spline signal interpolation: the zero, the first and the third order spline signal respectively.

In the sequel, we will first talk of spline signal S without specifying the spline order to derive the interpolating formulas; then, we will look at the shape of the involved matrices in the three different cases.

We start modeling each observation y_m^o as the sum of the spline signal in the corresponding observation point t_m^o plus a random error ν_m:

$$y_m^o = S(t_m^o) + \nu_m = \sum_{n=0}^{n_{max}} a_n \varphi(t_m^o - n) + \nu_m, \quad m = 1, M \tag{1.2.2}$$

where $n_{max} = N - 1$ for the zero order signal, otherwise $n_{max} = N$.

Eq. 1.2.2 can be written in the following vector form:

$$\underline{y}^o = S(\underline{t}^o) + \underline{\nu} = \sum_{n=0}^{n_{max}} a_n \varphi(\underline{t}^o - n) + \underline{\nu} \tag{1.2.3}$$

where

$$\underline{y}^o = \begin{bmatrix} y_1^o \\ \dots \\ y_M^o \end{bmatrix}, \tag{1.2.4}$$

$$\underline{t}^o = \begin{bmatrix} t_1^o \\ cdots \\ t_M^o \end{bmatrix} \tag{1.2.5}$$

and

$$\varphi(\underline{t}^o - n) = \begin{bmatrix} \varphi(t_1^o - n) \\ \dots \\ \varphi(t_M^o - n) \end{bmatrix}. \tag{1.2.6}$$

Or, by setting:

$$\Phi = \begin{bmatrix} \varphi(\underline{t}^o - 0) & \varphi(\underline{t}^o - 1) & \cdots & \varphi(\underline{t}^o - n_{max}) \end{bmatrix} =$$

$$= \begin{bmatrix} \varphi(t_1^o - 0) & \varphi(t_1^o - 1) & \cdots & \varphi(t_1^o - n_{max}) \\ \dots & \dots & \dots & \dots \\ \varphi(t_M^o - 0) & \varphi(t_M^o - 1) & \cdots & \varphi(t_M^o - n_{max}) \end{bmatrix}, \tag{1.2.7}$$

and

$$\underline{a} = \begin{bmatrix} a_0 \\ \dots \\ a_{n_{max}} \end{bmatrix}. \tag{1.2.8}$$

Eq. 1.2.3 becomes:

$$\underline{y}^o = S(\underline{t}^o) + \underline{\nu} = \Phi\underline{a} + \underline{\nu} \tag{1.2.9}$$

Here we assume the that the noise random vector has zero mean, and a covariance matrix proportional to the identity matrix, that is errors have the same variance and are uncorrelated:

$$E\{\underline{\nu}\} = \underline{0},\ C_{\nu\nu} = \sigma_0^2 I. \tag{1.2.10}$$

The least squares solution of our problem, namely the parameter vector $\underline{\hat{a}}$ (cfr. Eq. 1.2.8) which minimizes the following objective function:

$$G(\underline{a}) = (\underline{y}^o - \Phi\underline{a})'(\underline{y}^o - \Phi\underline{a}), \tag{1.2.11}$$

writes

$$\underline{\hat{a}} = (\Phi'\Phi)^{-1}\Phi'\underline{y}^o = N^{-1}\Phi'\underline{y}^o. \tag{1.2.12}$$

where we have introduced the normal matrix:

$$N = \Phi'\Phi \tag{1.2.13}$$

Before looking at the shape of this normal matrix in the specific zero, first and third spline interpolation problems, we remind that the Φ matrix is composed by $n_{max}+1$ column vectors spanning a space of dimension $n_{max}+1$.Moreover, the solution of our least squares problem is a vector belonging to this span, representing our interpolating signal in the observation points:

$$\underline{\hat{y}} = \Phi\underline{\hat{a}} \in Span\left[\varphi(\underline{t}^o - n)\right]. \tag{1.2.14}$$

By construction the element (i, j) of the normal matrix is the scalar product of the vector $\varphi(\underline{t}^o - i)$ by the vector $\varphi(\underline{t}^o - k)$; so that, the diagonal elements $(i = j)$ of the matrix are the square norm of those vectors and are positive, while the elements out of the diagonal $(i \neq j)$ are equal to zero if the two corresponding vectors are orthogonal.

Consider now the Phi_0 matrix, namely the matrix associated to the zero order spline interpolation problem. Consider the generic column $\varphi(\underline{t}^o - k)$, this column will be different from zero only if $t^o_m \in I_k$.