Design matrix A orthogonalization

Gram-Schmidt method

The product of the design matrix A by the parameter vector $\underline{x}$, can be seen as a linear combination of the A matrix column vectors. If the columns are independent (as we assume) this combination produces a linear subspace of R^n, of dimension 2, equal to the number of columns of A (or number of parameters). If n=3, then we can represent its vectors (the observation vector, the column vectors of A, the residual vector) in the cartesian space, as arrows stemming from the origin of the cartesian system. The above linear subspace will be the plane generated by the linear combination of the arrow vectors representing the columns of A. If n is greater than 3 we cannot represent its vectors, but we can extend the concepts seen for the arrow vectors as well as the formulas, such as the scalar product, the norm,etc.

$$A\hat{\underline{x}} = \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \dots & \dots \\ \dots & \dots \\ 1 & t_n \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \dots \\ \dots \\ 1 \end{bmatrix} b_0 + \begin{bmatrix} t_1 \\ t_2 \\ \dots \\ \dots \\ t_n \end{bmatrix} b_1 = \underline{A}_1 x_1 + \underline{A}_2 x_2$$

The linear subspace $A\underline{x}$ can be obtained by combining whatever two vectors belonging to the subspace itself. In particular, it is useful to find two **orthogonal vectors** to represent it. To this aim we use the Gram-Schmidt method.

We leave the first column vector as it is, and look for a second column vector orthogonal to it. We then compute the component of the $\underline{A}_2$ vector parallel to $\underline{A}_1$:

$$\underline{A}_{2\|} = \left(\underline{A}_2 \times \frac{\underline{A}_1}{\left\| \underline{A}_1 \right\|} \right) \frac{\underline{A}_1}{\left\| \underline{A}_1 \right\|} = \frac{\underline{A}_2 \times \underline{A}_1}{\underline{A}_1 \times \underline{A}_1} \underline{A}_1$$

The orthogonal component is obtained by subtracting the parallel component to $\underline{A}_2$:

$$\underline{A}_{2\perp} = \underline{A}_2 - \underline{A}_{2\|} = \underline{A}_2 - \frac{\underline{A}_2 \times \underline{A}_1}{\underline{A}_1 \times \underline{A}_1} \underline{A}_1$$

In our specific case we have:

$$\underline{A}_{2||} == \frac{\sum_{i=1}^{n} t_i}{n} \underline{A}_1 = \bar{t} \begin{bmatrix} 1 \\ 1 \\ \dots \\ \dots \\ 1 \end{bmatrix}$$

$$\underline{A}_{2\perp} = \begin{bmatrix} t_1 \\ t_2 \\ \dots \\ \dots \\ t_n \end{bmatrix} - \begin{bmatrix} \bar{t} \\ \bar{t} \\ \dots \\ \dots \\ \bar{t} \end{bmatrix} = \begin{bmatrix} t_1 - \bar{t} \\ t_2 - \bar{t} \\ \dots \\ \dots \\ t_3 - \bar{t} \end{bmatrix} = \begin{bmatrix} \tau_1 \\ \tau_2 \\ \dots \\ \dots \\ \tau_n \end{bmatrix}$$

Where we have introduced the new orthogonal vector $\underline{\tau}$.

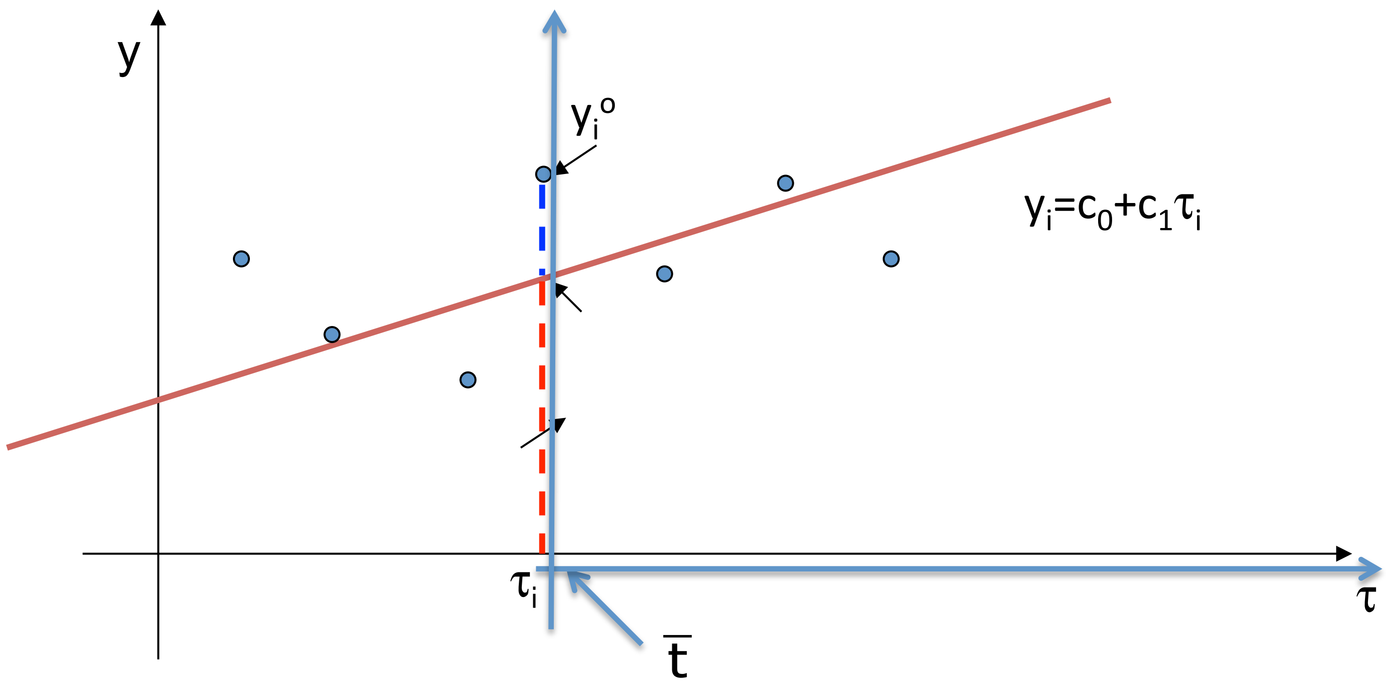
Coming back to the interpolation problem, it must be modified as follows:

$$\underline{y}^{o} = A\underline{x} + \underline{v} = A^{GS}\underline{x}^{GM} + \underline{v} = \begin{bmatrix} 1 & \tau_1 \\ 1 & \tau_2 \\ \dots & \dots \\ \dots & \dots \\ \dots & \tau_n \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} + \underline{v}$$

Where we have introduced the new orthogonal vector $\underline{\tau}$.

This change corresponds to a translation of the cartesian system (t,y) into (τ,y).

We look for the same interpolating straight line, now referred to another reference system,



whose equation is

$$y = b_0 + b_1\left(\tau + \bar{t}\right) = \left(b_0 + b_1\bar{t}\right) + b_1\tau =$$

$$= c_0 + c_1\tau$$

Where the relation between the old and the new parameters is:

$$c_1 = b_1$$
$$c_0 = b_0 + b_1 \overline{t}$$

The advantage in using the A^{GM} matrix is in the solution of the least squares problem, which now have a diagonal normal matrix:

$$\underline{\hat{x}}^{GM} = \begin{bmatrix} \hat{c}_0 \\ \hat{c}_1 \end{bmatrix} = \left(N^{GM}\right)^{-1}\left(A^{GM}\right)'\underline{y}^o = \begin{bmatrix} n & 0 \\ 0 & \sum_{i=1}^{n}\tau_i^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{i=1}^{1n} y_i^o \\ \sum_{i=1}^{n} y_i^o\tau_i \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{n} & 0 \\ 0 & \frac{1}{\sum_{i=1}^{n}\tau_i^2} \end{bmatrix} \begin{bmatrix} \sum_{i=1}^{1n} y_i^o \\ \sum_{i=1}^{n} y_i^o\tau_i \end{bmatrix} = \begin{bmatrix} \frac{\sum_{i=1}^{1n} y_i^o}{n} \\ \frac{\sum_{i=1}^{n} y_i^o\tau_i}{\sum_{i=1}^{n}\tau_i^2} \end{bmatrix}$$

The same orthogonalization method can be extended to a matrix with m>2 Independent columns, as we can have in the least squares interpolation with a polynomial function of degree larger than 1.

The general formulas becomes:

$$\underline{A}_1^{GS} = \underline{A}_1$$

$$\underline{A}_2^{GS} = \underline{A}_2 - \frac{\underline{A}_2 \times \underline{A}_1^{GS}}{\underline{A}_1^{GS} \times \underline{A}_1^{GS}} \underline{A}_1^{GS}$$

$$\underline{A}_3^{GS} = \underline{A}_3 - \frac{\underline{A}_3 \times \underline{A}_1^{GS}}{\underline{A}_1^{GS} \times \underline{A}_1^{GS}} \underline{A}_1^{GS} - \frac{\underline{A}_3 \times \underline{A}_2^{GS}}{\underline{A}_2^{GS} \times \underline{A}_2^{GS}} \underline{A}_2^{GS}$$

$$\underline{A}_k^{GS} = \underline{A}_k - \sum_{i=1}^{k-1} \frac{\underline{A}_k \times \underline{A}_i^{GS}}{\underline{A}_i^{GS} \times \underline{A}_i^{GS}} \underline{A}_i^{GS}$$