The Moore-Penrose Inverse in Art

D Huylebrouck
LUCA Department for Architecture KULeuven
Paleizenstraat 65-67
1030 Brussels BELGIUM
Huylebrouck@gmail.com

Abstract

The "Moore-Penrose inverse" of a matrix A corresponds to the (unique) matrix solution X of the system AXA=A, XAX=X, $(AX)^T=AX$, $(XA)^T=XA$. This generalized inverse has many applications, ranging from Gauss' historical prediction for finding Ceres to modern electrical engineering problems. The present paper provides some applications related to art: one about mathematical color theory, and one about curve fitting in architectural drawings or paintings.

The generalized inverse

In 1955 Roger Penrose (1931-) independently described a type of matrix pseudoinverse that had been proposed earlier by the American mathematician Eliakim Hastings Moore (1862 –1932). As was the case in other fields on which Penrose laid hands (such as the Penrose tiles or the Penrose triangle), this pseudoinverse then became widely used and popular.

Consider a system of equations

$$\begin{cases} a_{11}x_1 + a_{12}x_2 \; ... \; a_{1n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 \; ... \; a_{mn}x_n = b_m \end{cases}$$

written in matrix form as $\mathbf{Ax}=\mathbf{b}$. Here, $\mathbf{A}\in M_{m\times n}(\mathbb{R})$ is a real m×n matrix and $\mathbf{b}\in M_{m\times 1}(\mathbb{R})$ a real vector. The system has a unique solution if and only if the matrix **A** is invertible, in which case it is given by $\mathbf{A}^{-1}\mathbf{b}$. If **A** is not invertible, the so-called "Moore-Penrose inverse" or "MP"-inverse, can be used to get a convenient approximation, the word "convenient" being motivated by a least squares interpretation given below. This "generalized inverse" of a matrix $\mathbf{A}\in M_{m\times n}(\mathbb{R})$ is the (unique) solution $\mathbf{X}\in M_{n\times m}(\mathbb{R})$ of the system of matrix equations (^{T} denotes transposition): $\mathbf{AXA}=\mathbf{A}$, $\mathbf{XAX}=\mathbf{X}$, $(\mathbf{AX})^T=\mathbf{AX}$ and $(\mathbf{XA})^T=\mathrm{X}\mathbf{A}$. A common notation for this **X** is $\mathbf{A}^\dagger$, and this matrix is also called "the" generalized inverse of **A**. If the matrix **A** has an inverse matrix, it coincides with this generalized inverse. The link to the least squares method is important here, and we follow the algebraic formulation by Campbell and Meyer (see [1]).

Suppose $\mathbf{A}\in M_{m\times n}(\mathbb{R})$ and $\mathbf{b}\in M_{m\times 1}(\mathbb{R})$; a vector $\mathbf{u}\in M_{n\times 1}(\mathbb{R})$ is called "a least squares solution" to $\mathbf{Ax}=\mathbf{b}$ if $\|\mathbf{Au}-\mathbf{b}\| \le \|\mathbf{Av}-\mathbf{b}\|$ for all $\mathbf{v}\in M_{n\times 1}(\mathbb{R})$. A vector **u** is called "the minimal least squares solution" to $\mathbf{Ax}=\mathbf{b}$ if **u** is a least squares solution to $\mathbf{Ax}=\mathbf{b}$ and $\|\mathbf{u}\| < \|\mathbf{w}\|$ for all other least squares solutions **w**. The denomination "least squares" refers to the squares in the common Euclidean norm $\|\mathbf{x}\| = \sqrt{\sum_{i=1}^{m} \mathbf{x}_i^2}$ of a vector $\mathbf{x}\in M_{m\times 1}(\mathbb{R})$. The link to the Moore-Penrose generalized inverse is that $\mathbf{A}^\dagger\mathbf{b}$ is the minimal least squares solution to $\mathbf{Ax}=\mathbf{b}$, for given $\mathbf{A}\in M_{m\times n}(\mathbb{R})$ and $\mathbf{b}\in M_{m\times 1}(\mathbb{R})$ (see [2]).

Other generalizations of the notion of inverse of a matrix exist, and one such notion needed here is the so-called "constraint generalized inverse". The set $\mathscr{S}$ of least square solutions of $\mathbf{Cx}=\mathbf{f}$ is called the set of constraints. The problem of interest now is to find the unique minimum in $\mathscr{S}$ among the points of $\mathscr{S}$ for which $\|\mathbf{A}x - \mathbf{b}\|$ is minimal. It appears to be the vector $\mathbf{x_m}=\mathbf{A_C}^{\dagger}\mathbf{b}+(\mathbf{1}- \mathbf{A_C}^{\dagger}\mathbf{A}).\mathbf{C}^{\dagger}\mathbf{f}$, where $\mathbf{A_C} = \mathbf{A}(\mathbf{1}-\mathbf{C}^{\dagger}\mathbf{C})$. When there are no constraints, $\mathbf{C}=\mathbf{0}$, and thus $\mathbf{A_C}^{\dagger} = \mathbf{A}^{\dagger}$, whence again as above $\mathbf{x_m}=\mathbf{A}^{\dagger}\mathbf{b}$.

Elements of color theory

The R(ed)-G(reen)-B(lue) values of a color C are denoted in a vector: $\mathbf{C} = [\mathrm{R\ G\ B}]^{\mathrm{T}}$. A color with an R-value of 0.30 has 70% of red, 30% being white. Three arbitrary but identical RGB values yield **Gray** = $[\mathrm{w\ w\ w}]^{\mathrm{T}}$, the value indicating the amount of white in this gray. Special cases are **blacK** = $[0\ 0\ 0]^{\mathrm{T}}$, where the letter K is used for "blacK" to avoid confusion with “Blue”, and **White** = $[1\ 1\ 1]^{\mathrm{T}}$, or, more accurately, “pure light”.

There are two main ways to mix colors: “additive” and “subtractive”. In the most usual examples, the first corresponds to a projection of light, thus “adding” light for each projection, while the second puts transparent layers in front of a projection, thus “subtracting” light.

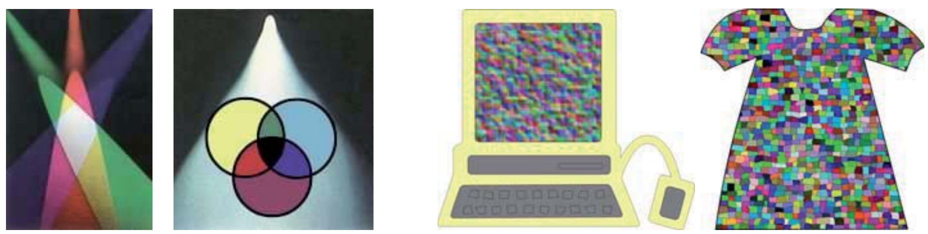

Figure 1: *Additive (far left), subtractive (middle left) and weighted color mixing (right; the images on a computer screen or a dress are best observed from a distance).*

In the additive case, the RGB-values are added, but in the subtractive case they are multiplied:

$$\begin{bmatrix} \mathrm{R_{add}} \\ \mathrm{G_{add}} \\ \mathrm{B_{add}} \end{bmatrix} = \begin{bmatrix} \mathrm{R_1} \\ \mathrm{G_1} \\ \mathrm{B_1} \end{bmatrix} + \begin{bmatrix} \mathrm{R_2} \\ \mathrm{G_2} \\ \mathrm{B_2} \end{bmatrix} = \begin{bmatrix} \mathrm{R_1+R_2} \\ \mathrm{G_1+G_2} \\ \mathrm{B_1+B_2} \end{bmatrix} \text{ and } \begin{bmatrix} \mathrm{R_{substr}} \\ \mathrm{G_{substr}} \\ \mathrm{B_{substr}} \end{bmatrix} = \begin{bmatrix} \mathrm{R_1} \\ \mathrm{G_1} \\ \mathrm{B_1} \end{bmatrix} \otimes \begin{bmatrix} \mathrm{R_2} \\ \mathrm{G_2} \\ \mathrm{B_2} \end{bmatrix} = \begin{bmatrix} \mathrm{R_1R_2} \\ \mathrm{G_1G_2} \\ \mathrm{B_1B_2} \end{bmatrix}.$$

Computer screens or plaid fabrics provide typical examples of additive color mixing. Given colors $\mathbf{C}_1, \ldots \mathbf{C}_n$, each used for an amount $\rho_1, \rho_2, \ldots, \rho_n$, with $\rho_1 + \rho_2 + \ldots + \rho_n = 1$ and $\rho_i \in \mathbb{R}$ (in fact, $\rho_i \in \mathbb{Q}$ and $0<\rho_i<1$), the RGB-coordinates of the composed color will be given by:

$$\mathbf{C}_{\mathrm{mixture}} = \begin{bmatrix} \rho_1\mathrm{R_1} \\ \rho_1\mathrm{G_1} \\ \rho_1\mathrm{B_1} \end{bmatrix} + \begin{bmatrix} \rho_2\mathrm{R_2} \\ \rho_2\mathrm{G_2} \\ \rho_2\mathrm{B_2} \end{bmatrix} + \ldots + \begin{bmatrix} \rho_n\mathrm{R_n} \\ \rho_n\mathrm{G_n} \\ \rho_n\mathrm{B_n} \end{bmatrix} = \rho_1\mathbf{C}_1 + \rho_2\mathbf{C}_2 + \ldots + \rho_n\mathbf{C}_n.$$

Example 1. A “complementary” color of a given color is such that the simultaneous projection of a light beam of the given color and the complementary color yields white: $\mathbf{C}_1 = [\mathrm{R_1\ \ G_1\ \ B_1}]^{\mathrm{T}}$ and $\mathbf{C}_2 = [\mathrm{R_2\ \ G_2\ \ B_2}]^{\mathrm{T}}$ are complementary iff $[\mathrm{R_1\ \ G_1\ \ B_1}]^{\mathrm{T}} + [\mathrm{R_2\ \ G_2\ \ B_2}]^{\mathrm{T}} = [1\ \ 1\ \ 1]^{\mathrm{T}}$. For instance, the “main color” **Cyan** = $[0\ \ 1\ \ 1]^{\mathrm{T}}$ and a double layer of transparent disks of the other main colors **Magenta** = $[1\ \ 0\ \ 1]^{\mathrm{T}}$ and **Yellow** = $[1\ \ 1\ \ 0]^{\mathrm{T}}$ are complementary: $[0\ \ 1\ \ 1]^{\mathrm{T}} + [1\ \ 0\ \ 1]^{\mathrm{T}} \otimes [1\ \ 1\ \ 0]^{\mathrm{T}} = [1\ \ 1\ \ 1]^{\mathrm{T}}$.

Thus cyan and the other main colors magenta and yellow have the property that the additive mixture of the first with the subtractive mixture of the second and the third is white.

Example 2. Suppose colors **White** = $[1\ 1\ 1]^{\mathrm{T}}$, **Green** = $[0\ 1\ 0]^{\mathrm{T}}$, **Cyan** = $[0\ 1\ 1]^{\mathrm{T}}$ and **black** = $[0\ 0\ 0]^{\mathrm{T}}$ are mixed at rates of respectively 30%, 40%, 10%, and 20%. The mixing can be considered as a simple additive mixing of a 30% gray, a dark green, a very dark cyan and black (adding nothing at all, of course), or it can be interpreted as a weighted mean:

$$\begin{bmatrix}.3\\.3\\.3\end{bmatrix}+\begin{bmatrix}0\\.4\\0\end{bmatrix}+\begin{bmatrix}0\\.1\\.1\end{bmatrix}+\begin{bmatrix}0\\0\\0\end{bmatrix}=0.3\begin{bmatrix}1\\1\\1\end{bmatrix}+0.4\begin{bmatrix}0\\1\\0\end{bmatrix}+0.1\begin{bmatrix}0\\1\\1\end{bmatrix}+0.2\begin{bmatrix}0\\0\\0\end{bmatrix}=\begin{bmatrix}0.3\\0.8\\0.4\end{bmatrix}.$$

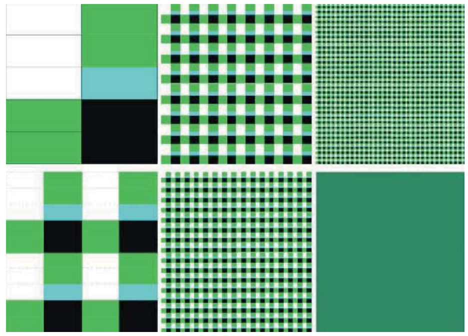

Figure 2: *A dark green surface illustrating the expression* $0.3[1, 1, 1]^T + 0.4[0, 1, 0]^T + 0.1[0, 1, 1]^T + 0.2[0, 0, 0]^T = [0.3, 0.8, 0.4]^T$.

Application of the Moore-Penrose inverse

A given color can be obtained in many ways, but on the other hand it is not always possible to form a desired color when only some given colors are available. Here, a least squares technique may help, since the situation corresponds to a matrix equation $\mathbf{A}.\rho = \mathbf{b}$ or

$$\rho_1\begin{bmatrix}R_1\\G_1\\B_1\end{bmatrix}+\rho_2\begin{bmatrix}R_2\\G_2\\B_2\end{bmatrix}+\ldots+\rho_n\begin{bmatrix}R_n\\G_n\\B_n\end{bmatrix}=\begin{bmatrix}R\\G\\B\end{bmatrix},$$

where the columns of **A** contain the available colors and **b** the desired color. The constraint $\rho_1+\rho_2+\ldots+\rho_n = 1$, becomes, in matrix notation, $\mathbf{C}.\rho = \mathbf{f}$, with $\mathbf{C} = [1, 1, \ldots, 1]$, $\rho = [\rho_1, \rho_2, \ldots, \rho_n]^T$ and $\mathbf{f} = [1]$. The "best solution" in terms of least squares corresponds to $\boldsymbol{\rho}_\mathbf{m}=\mathbf{A_C}^\dagger\mathbf{b}+(\mathbf{1}-\mathbf{A_C}^\dagger\mathbf{A}).\mathbf{C}^\dagger\mathbf{f}$, where $\mathbf{A_C} = \mathbf{A}(\mathbf{1}-\mathbf{C}^\dagger\mathbf{C})$.

Example 3. Suppose the green color of the above example is given, $[0.3, 0.8, 0.4]^T$, but that is has to be made using **red** $= [1, 0, 0]^T$, **green** $= [0, 1, 0]^T$, **blue** $= [0, 0, 1]^T$ and **dark yellow** $= [0.5, 0.5, 0]^T$. There are no $\rho_1, \rho_2, \rho_3, \rho_4$ values such that $[1, 0, 0]^T\rho_1 + [0, 1, 0]^T\rho_2 + [0, 0, 1]^T\rho_3 + [0.5, 0.5, 0]^T\rho_4 = [0.3, 0.8, 0.4]^T$, with $\rho_1 + \rho_2 + \rho_3 + \rho_4 = 1$, but the Moore-Penrose inverse allows one to obtain 'the best possible solution'. Now, $\mathbf{A}=\begin{bmatrix}1&0&0&0.5\\0&1&0&0.5\\0&0&1&0\end{bmatrix}$, $\mathbf{b}=\begin{bmatrix}0.3\\0.8\\0.4\end{bmatrix}$, $\mathbf{C} = [1, 1, 1, 1]$ and $\mathbf{f} = [1]$. Thus we consider $\mathbf{A_C} =$

$\mathbf{A}(\mathbf{1}-\mathbf{C}^\dagger\mathbf{C}) = \dfrac{1}{8}\begin{bmatrix}5&-3&-3&1\\-3&5&-3&1\\-2&-2&6&-2\end{bmatrix}$, since $\mathbf{C}^\dagger=\dfrac{1}{4}[1\ \ 1\ \ 1\ \ 1]^T$. MATHEMATICA$_{TM}$ shows $\mathbf{A_C}^\dagger$

$=\dfrac{1}{18}\begin{bmatrix}11&-7&-4\\-7&11&-4\\-6&-6&12\\2&2&-4\end{bmatrix}$. Thus, $\boldsymbol{\rho}_\mathbf{m} = \dfrac{1}{180}[1\ \ 91\ \ 42\ \ 46]^T$ and $\mathbf{A}.\boldsymbol{\rho}_\mathbf{m} = \dfrac{1}{30}\begin{bmatrix}4\\19\\7\end{bmatrix}=\begin{bmatrix}0.133...\\0.633...\\0.233....\end{bmatrix}$.

If, for instance, 180 threads are used, the computation proposes 1 red, 91 green, 42 blue and 46 dark yellow threads for making a plaid piece of textile that resembles the “best” (in the “least squares” meaning) to the desired green, though the obtained color, $[0.133\ldots, 0.633\ldots, 0.233\ldots]^T$, does not seem a very good solution when compared to the desired $[0.3, 0.8, 0.4]^T$. This “best” approximation, using a single thread of red for every 180 threads, would have been hard to guess without the mathematical approach.

Space here does not permit presenting an additional application such as the often occurring problem for color specialists of finding out if given colors are “well balanced”, that is, finding out if combinations provide perfectly complementary colors or, at least, gray without shades of a color. Some critique to the method will be discussed too, since the (subjective) observation of colors probably involves other parameters. For sure, color theory as taught in many art schools seems in need of a more mathematical approach. Steve Campbell formulated it as follows in an e-mail contact about this topic: *“The color theory application is one that many in mathematics may not be aware of, is timely, and would be fun to present to students.”*

Curve fitting

Another application of the Moore-Penrose inverse can be found in the common practice to draw all kinds of geometric figures on images of artworks and buildings. Usually, simple triangles, rectangles, pentagons or circles suffice and still recognizing these mathematical shapes is seen as an “interpretation” of an architectural edifice or painting. Today, research on the borderline of mathematics and architecture tends to reject these “geometric readings in art” (see [3]).

However, a way to meet the shortcomings of this seemingly arbitrary drawing method is to use a similarity between these geometric studies in art and celestial mechanics. In the latter field, use of the least squares method is well established, ever since Carl Friedrich Gauss found Ceres, the "lost planet" astronomer G. Piazzi had briefly observed in January of 1801. Of course, it can be opposed that, unlike for celestial mechanics with its involved applications, such a mathematical method would present a serious overkill with respect to the intended straightforward artistic applications. This was true in the past, but nowadays software allows reducing the computational aspects to a few computer clicks. One can only wonder why this widespread and straightforward mathematical technique has not been applied before in art.

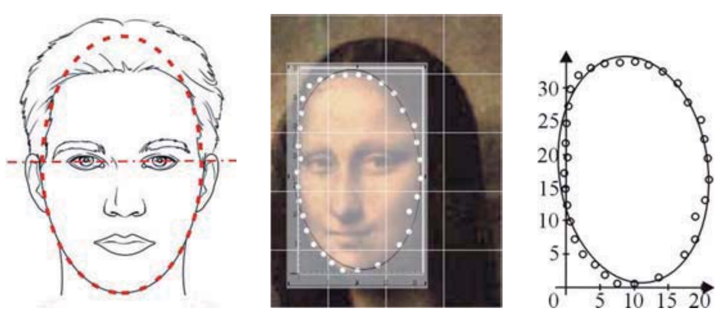

Figure 3: *As any sketch class learns how to draw a human face starting from an ellipse (left), fitting an ellipse around the Mona Lisa (middle, right) makes sense, more than for, for instance, a rectangle.*

Consider for instance the case of the Mona Lisa and finding the closed ellipse $x^2b_1 + y^2b_2 + xb_3 + yb_4 + xyb_5 - 1 = 0$ used to draw its face. Now, the data are not based on astronomical observations through a telescope, but on measurements made on a reproduction of the Mona Lisa as given in the illustration:

$x1 = 17.496$; $y1 = 4.871$; $x2 = 19.129$; $y2 = 7.143$;… ; $x29 = 13.805$; $y29 = 1.541$.

These will be filled in the matrices

$$\mathbf{X} = \begin{bmatrix} x_1^{\;2} & y_1^{\;2} & x_1 & y_1 & x_1 y_1 \\ x_2^{\;2} & y_2^{\;2} & x_2 & y_2 & x_2 y_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{29}^{\;2} & y_{29}^{\;2} & x_{29} & y_{29} & x_{29} y_{29} \end{bmatrix}, \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{bmatrix}, \mathbf{j} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}.$$

In this matrix notation, the curve fitting problem implies **X.b-j** should be minimal, and the numerical values can be found by computing the product of the Moore-Penrose inverse of **X** and **j**. This is easily done using Mathematica$_{\text{TM}}$:

X = {{x1^2, y1^2, x1, y1, x1*y1}, … {x29^2, y29^2, x29, y29, x29*y29}};
j = {1, … 1};
PseudoInverse[X].j
Norm[X.PseudoInverse[X].j]^2/Norm[j]^2

It results in an ellipse

$- 0.0078x^2 - 0.0033y^2 + 0.1686x + 0.1241y - 0.0009xy = 1.$

We can conclude that a rectangle with a proportion of 1.54... provides the closest fit and the word “closest” can be substantiated: the fit is “99.55%” (and 1.54… is closer to, say, 1.5, than to the golden number of 1.618...).

Example 4. Gaudi is known for his use of hyperbolic cosine functions. He used bags suspended by ropes which he inverted to get chain curves. Clearly, this way a not so continuous curve is obtained, yet nevertheless it is a classic statement among Gaudi specialists the architect did indeed use the catenary. Still, even if this is the case, for instance because further approximations were used, the question remains if a non-informed spectator can actually see this in Gaudi’s work, that is, if a spectator can notice, by mere observation, that a Gaudi gate, entrance or window, has a hyperbolic cosine shape.

Figure 4: *Fitting curves on the Paelle Guell (left) and on the Collegio Teresiano (right).*

We first tried this for the Paelle Guell: we wrote down coordinates of as many points we accurately could distinguish. Using the Moore-Penrose inverse, it turned out the parabola with equation $y = 1.84 - 52.12x^2$, fits at 96.75%, while the hyperbolic cosine $y = 1.34 - 0.36 \cdot \text{Cosh}[9.7x]$ and it fits at 99.88%. This 3% difference in closeness of fit can indeed be noticed, as the illustration shows. We executed a similar computation for a picture of Gaudi’s Collegio Teresiano, but there no difference can be noticed between the closed fit parabola and hyperbolic cosine. During the “Mathematics and Design 2007” conference in Blumenau, Brazil, I was most happy to learn Barcelona Prof. Amadeo Monreal shared this view, based on his knowledge of Gaudi’s work (see [4]).

Example 5. Suppose we want to determine the shape of a nuclear power plant.

Surprisingly, an ellipse fits well here:

$-0.11x^2 - 0.0105y^2 + 0.67x + 0.19y - 0.06xy = 1$ fits at 99.9996%.

Still, if we lift the x-axis over 5 units, we can propose the standard form equation for a hyperbola: $0.508687x^2 - 0.108291y^2 = 1$, and this is 99.8% close. Thus, we can faithfully claim the shape of a nuclear plant is seen as a hyperbola, as the adaptations on the top of the building, for reasons of resistance to wind, can hardly be detected.

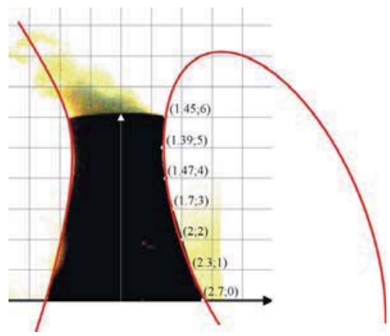

Figure 5: *Determining the profile of a nuclear plant. The profile of a nuclear plant, and the approximating shapes, a hyperbola and an ellipse.*

Further 'research'

The color theory approach could be applied to any list of 'preferences' that have to be combined. For instance, consider the 'art' of gastronomy: there are 4 main tastes, salt (Sa), bitter (Bi), sweet (Sw), sour (So). Food without salt gets a 0 on the salt scale, and an imaginary plate consisting of only salt a 1, and so on. The Sa-Bi-Sw-So values of some food are denoted in a vector: $\mathbf{T} = [\text{Sa Bi Sw So}]^T$. Tastes can be mixed in an additive way, when spices are simply thrown one upon another. Subtractive mixing is also possible, when the spices go through two layers of disks with holes like in many common spice containers. And so, the question now becomes how to get the taste that comes close to a given taste? Surely more work in the kitchen is needed to explore whether the Moore-Penrose inverse gives the most delicious answer.

Figure 6: *In gastronomy, it isn't light beams going through a filter that are considered, but spices going through holes (left). The principles of additive (middle) and subtractive (right) mixing are similar.*

References

[1] Campbell S. L. and Meyer, C. D. Jr, *Generalized Inverses of Linear Transformations*, Pitman Publishing Limited, London, 1979.
[2] Dirk Huylebrouck, Generalized Inverses of a Sum with Radical Element, Applications, *Linear Algebra and its Applications*, 246 (1996), 159-175.
[3] Huylebrouck D., Curve Fitting in Architecture, *Nexus Network Journal,* Vol. 9, no. 1, Spring, Ed. Kim Williams 2007 p. 59 – 70.
[4] Monreal A., Gaudi en el Siglo XXI, *Journal of Mathematics & Design*, Volume 7, Number 1, Proceedings of the Mathematics and Design conference, Blumenau, Brazil, 2007.