Scientific Bulletin of the „Petru Maior” University of Tîrgu Mureş
Vol. 10 (XXVII) no. 1, 2013
ISSN-L 1841-9267 (Print), ISSN 2285-438X (Online), ISSN 2286-3184 (CD-ROM)

WEIGHTED OVERDETERMINED LINEAR SYSTEMS

Béla Finta
”Petru Maior” University of Tg. Mureş, Romania
e-mail: fintab@science.upm.ro

ABSTRACT

The purpose of this paper is to introduce the weighted overdetermined linear systems and to solve them in the sense of the least squares method

Keywords: overdetermined linear system, weighted overdetermined linear system, least squares method

1 Introduction

Let us consider the real matrix $A = (a_{ij})_{\substack{i=\overline{1,m}\\ j=\overline{1,n}}}$, and the real, transposed arrays $x = (x_1, x_2, \ldots, x_n)^T \in \mathbb{R}^n$ and $b = (b_1, b_2, \ldots, b_m)^T \in \mathbb{R}^m$, respectively. The linear system $A \cdot x = b$ is called overdetermined linear system, if $m > n$. Generally, the overdetermined linear system is incompatible, i.e. doesn't exist an array $x^* = (x_1^*, x_2^*, \ldots, x_n^*)^T \in \mathbb{R}^n$ such that $A \cdot x^* = b$. It is well known to obtain the solution of the overdetermined linear system using the least squares method, see for example [1] and [2].

2 Main part

Let us consider the weight array $p = (p_1, p_2, \ldots, p_m)^T \in \mathbb{R}^m$, where $p_i > 0$ for every $i = \overline{1,m}$. We denote by $\sqrt{p} = (\sqrt{p_1}, \sqrt{p_2}, \ldots, \sqrt{p_m}) \in \mathbb{R}^m$ and for the vectors $x = (x_1, x_2, \ldots, x_m)^T \in \mathbb{R}^m$ and $y = (y_1, y_2, \ldots, y_m)^T \in \mathbb{R}^m$ we introduce the following product $x \otimes y = (x_1 \cdot y_1, x_2 \cdot y_2, \ldots, x_m \cdot y_m) \in \mathbb{R}^m$. Let us take the weighted matrix $A_p = (\sqrt{p_i} \cdot a_{ij})_{\substack{i=\overline{1,m}\\ j=\overline{1,n}}}$, and the weighted vector $b_p = \sqrt{p} \otimes b \in \mathbb{R}^m$. So to the overdetermined linear system $A \cdot x = b$ we can attache the weighted overdetermined linear system $A_p \cdot x = b_p$. It is immediatly that these two linear systems are equivalent. This means that, generally, the weighted overdetermined linear system is incompatible, i.e. doesn't exist an array $x^* = (x_1^*, x_2^*, \ldots, x_n^*)^T \in \mathbb{R}^n$ such that $A_p \cdot x^* = b_p$. For this reason, instead of the classical solution x^*, we consider such array $\overline{x} = (\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_n) \in \mathbb{R}^n$ for which the function $f : \mathbb{R}^n \to \mathbb{R}$, $f(x) = \|A_p \cdot x - b_p\|_m^2$ takes the minimal value, where $\| \cdot \|_m$ means the Euclidean norm on the space $\mathbb{R}^m$. The array $\overline{x} \in \mathbb{R}^n$, which minimizes the function f, it is accepted like the solution of the overdetermined linear system $A_p \cdot x = b_p$ in the sense of the least squares method. We can observe that $f(x) \geq 0$ for all $x \in \mathbb{R}^n$ and the minimal point $\overline{x} \in \mathbb{R}^n$ verifies the following system of equations with partial derivatives $\frac{\partial f}{\partial x_k}(\overline{x}) = 0$ for every $k = \overline{1,n}$. We calculate the partial derivatives:

$$\frac{\partial f}{\partial x_k}(x) = \sum_{i=1}^{n} 2 \cdot \left(\sum_{j=1}^{m} \sqrt{p_i} \cdot a_{ij} x_j - \sqrt{p_i} \cdot b_i \right) \cdot \sqrt{p_i} \cdot a_{ik} = 0,$$

and we obtain

$$\sum_{i=1}^{n} \sqrt{p_i} \cdot a_{ik} \cdot \left(\sum_{j=1}^{m} \sqrt{p_i} \cdot a_{ij} x_j - \sqrt{p_i} \cdot b_i \right) = 0$$

for every $k = \overline{1,m}$. This linear system has the matrix form $A_p^T \cdot (A_p \cdot x - b_p) = \theta_{\mathbb{R}^m}$. This is a Cramer's type linear system with n equations and n unknowns, so $\overline{x} \in \mathbb{R}^n$ will be the classical solution of this Cramer's linear system. Consequently the classical solution $\overline{x} \in \mathbb{R}^n$ of the linear system $(A_p^T \cdot A_p) \cdot x = A_p^T \cdot b_p$ it is accepted like solution of the weighted overdetermined linear system $A_p \cdot x = b_p$ in the sense of the least squares approach.

Next we show that the stationary point $\overline{x} \in \mathbb{R}^n$, which verifies the relation $(A_p^T \cdot A_p) \cdot x = A_p^T \cdot b_p$ will be the minimal point of the function f.

Theorem 1. *If A_p is a matrix of type $n \times m$, b_p is a column matrix of type $m \times 1$, and $\overline{x} \in \mathbb{R}^n$ is the classical solution of the linear system $A_p^T \cdot (A_p \cdot x - b_p) = \theta_{\mathbb{R}^m}$, then for every $y \in \mathbb{R}^m$ we obtain $\|b_p - A_p \cdot \overline{x}\|_2 \leq \|b_p - A_p \cdot y\|_2$.*

Proof. We denote the residual vectors with $r_{\overline{x}} = b_p - A_p \cdot \overline{x}$ and $r_y = b_p - A_p \cdot y$. Then we have $r_y = b_p - A_p \cdot y = b_p - A_p \cdot \overline{x} + A_p \cdot \overline{x} - A_p \cdot y = r_{\overline{x}} + A_p \cdot (\overline{x} - y)$. Now we use the well known formulas for the transposed matrices and we get: $r_y^T = (r_{\overline{x}} + A_p(\overline{x} - y))^T = r_{\overline{x}}^T + (\overline{x} - y)^T A_p^T$. Consequently taking the scalar product on the space $\mathbb{R}^n$ we obtain:

$$\begin{aligned} r_y^T \cdot r_y &= (r_{\overline{x}}^T + (\overline{x} - y)^T A_p^T) \cdot (r_{\overline{x}} + A_p(\overline{x} - y)) = \\ &= r_{\overline{x}}^T \cdot r_{\overline{x}} + (\overline{x} - y)^T \cdot A_p^T \cdot r_{\overline{x}} + \\ &+ r_{\overline{x}}^T \cdot A_p(\overline{x} - y) + (\overline{x} - y)^T A_p^T \cdot A_p(\overline{x} - y). \end{aligned}$$

But $A_p^T \cdot r_{\overline{x}} = \theta_{\mathbb{R}^m}$ and $r_{\overline{x}}^T A_p = (A_p^T r_{\overline{x}})^T = \theta_{\mathbb{R}^m}$. So we can deduce that $r_y^T \cdot r_y = r_{\overline{x}}^T \cdot r_{\overline{x}} + (\overline{x} - y)^T A_p^T \cdot A_p(\overline{x} - y)$. Using the definition of the Euclidean norm it is easy to deduce that $\|r_y\|_2^2 = r_y^T \cdot r_y$. Consequently $\|r_y\|_2^2 = \|r_{\overline{x}}\|_2^2 + \|A_p(\overline{x} - y)\|_2^2 \geq \|r_{\overline{x}}\|_2^2$, what we must it to prove. □

The linear system $(A_p^T \cdot A_p) \cdot x = A_p^T \cdot b_p$ can be solved by Cramer's rule from theory of determinants for $n \in \mathbb{N}^*$ small natural numbers, but in other cases, for $n \in \mathbb{N}^*$ great natural numbers we can use numerical methods of linear algebra.

Consequence 1. *If we choose the weights $p_1 = p_2 = \ldots = p_m = 1$, then from our theorem we reobtain the well known result for overdetermined linear systems.*

Example 1. *Let us consider the following overdetermined linear system:* $\begin{cases} x + y = 2 \\ x + 2y = 3 \\ 2x + y = 4 \end{cases}$

If we solve the linear system $\begin{cases} x + y = 2 \\ x + 2y = 3 \end{cases}$ then we receive the solution $x = y = 1$, which doesn't verify the last equation: $2x + y = 3 \neq 4$. We obtain the same conclusion if we calculate the characteristic determinant: $\begin{vmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 2 & 1 & 4 \end{vmatrix} = 1 \neq 0$. So our overdetermined linear system is incompatible and does not have classical solution. We attache to this overdetermined linear system the following weighted overdetermined linear system: $\begin{cases} \sqrt{p_1} \cdot x + \sqrt{p_1} \cdot y = 2 \cdot \sqrt{p_1} \\ \sqrt{p_2} \cdot x + 2\sqrt{p_2} \cdot y = 3 \cdot \sqrt{p_2} \\ 2\sqrt{p_3} \cdot x + \sqrt{p_3} \cdot y = 4 \cdot \sqrt{p_3} \end{cases}$

We consider the function $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = p_1 \cdot (x+y-2)^2 + p_2 \cdot (x+2y-3)^2 + p_3 \cdot (2x+y-4)^2$ and for the linear system $\begin{cases} \frac{\partial f}{\partial x}(x, y) = 0 \\ \frac{\partial f}{\partial y}(x, y) = 0 \end{cases}$ we obtain:
$\begin{cases} (p_1 + p_2 + 4p_3) \cdot x + (p_1 + 2p_2 + 2p_3) \cdot y = 2p_1 + 3p_2 + 8p_3 \\ (p_1 + 2p_2 + 2p_3) \cdot x + (p_1 + 4p_2 + p_3) \cdot y = 2p_1 + 6p_2 + 4p_3. \end{cases}$
In another way we get $m = 3$, $n = 2$, $A_p = \begin{pmatrix} \sqrt{p_1} & \sqrt{p_1} \\ \sqrt{p_2} & 2 \cdot \sqrt{p_2} \\ 2 \cdot \sqrt{p_3} & \sqrt{p_3} \end{pmatrix}$, $b_p = \begin{pmatrix} 2 \cdot \sqrt{p_1} \\ 3 \cdot \sqrt{p_2} \\ 4 \cdot \sqrt{p_3} \end{pmatrix}$, so $A_p^T \cdot A_p = \begin{pmatrix} p_1 + p_2 + 4p_3 & p_1 + 2p_2 + 2p_3 \\ p_1 + 2p_2 + 2p_3 & p_1 + 4p_2 + p_3 \end{pmatrix}$ and $A_p^T \cdot b_p = \begin{pmatrix} 2p_1 + 3p_2 + 8p_3 \\ 2p_1 + 6p_2 + 4p_3 \end{pmatrix}$. Hence our system $A_p^T \cdot A_p \cdot x = A_p^T \cdot b_p$ is the same: $\begin{pmatrix} p_1 + p_2 + 4p_3 & p_1 + 2p_2 + 2p_3 \\ p_1 + 2p_2 + 2p_3 & p_1 + 4p_2 + p_3 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2p_1 + 3p_2 + 8p_3 \\ 2p_1 + 6p_2 + 4p_3 \end{pmatrix}$. We denote

$$\Delta = \begin{vmatrix} p_1 + p_2 + 4p_3 & p_1 + 2p_2 + 2p_3 \\ p_1 + 2p_2 + 2p_3 & p_1 + 4p_2 + p_3 \end{vmatrix},$$

$$\Delta_x = \begin{vmatrix} 2p_1 + 3p_2 + 8p_3 & p_1 + 2p_2 + 2p_3 \\ 2p_1 + 6p_2 + 4p_3 & p_1 + 4p_2 + p_3 \end{vmatrix},$$

$$\Delta_y = \begin{vmatrix} p_1 + p_2 + 4p_3 & 2p_1 + 3p_2 + 8p_3 \\ p_1 + 2p_2 + 2p_3 & 2p_1 + 6p_2 + 4p_3 \end{vmatrix},$$

We have the classical solution $x = \frac{\Delta_x}{\Delta}$ and $y = \frac{\Delta_y}{\Delta}$. So our weighted overdetermined linear system admits the solution $x = \frac{\Delta_x}{\Delta}$ and $y = \frac{\Delta_y}{\Delta}$ in the sense of the least squares approach. For the particular case $p_1 = p_2 = p_3 = 1$ we obtain $x = \frac{18}{11}$ and $y = \frac{7}{11}$, which is the solution of the initial overdetermined linear system in the sense of the least squares method.

3 Conclusions

In this paper we extended the least squares method from overdetermined linear systems to weighted overdetermined linear systems.

References

[1] B. Finta, Analiză numerică, Editura Universităţii ”Petru Maior”, Tg. Mureş, 2004.

[2] S. S. Rao, Applied Numerical Methods for Engineers and Scientists, Prentice Hall, Upper Saddle River, New Jersey, 2002.