Sādhanā (2026)51:221
https://doi.org/10.1007/s12046-026-03180-y

The row picture of least squares and symmedian coordinates

M HARIPRASAD[1,*], TANAY SAHA[2] and S GOKULRAJ[3]

[1]School of Mathematics and Natural Sciences, Chanakya University, Devanahalli, Bangalore 562 165, India
[2]Department of Mathematics and Computer Science, Mizoram University, Aizawl 796 004, India
[3]Department of Mathematics, National Institute of Technology Calicut, Kozhikode 673 601, India
*Corresponding Author.
E-mail: hariprasad.h@chanakyauniversity.edu.in; sahatanay50@gmail.com; gokulrajs@nitc.ac.in

MS received 29 January 2025; revised 25 October 2025; accepted 15 December 2025
© The Author(s), under exclusive licence to Indian Academy of Sciences 2026

Abstract. In this work, we study the least squares problem for an overdetermined system $Ax = b$, where $A \in \mathbb{R}^{m\times n}$ with $m > n$, which finds the vector x minimizing the Euclidean norm of the residual $b - Ax$. While the classical column picture provides a geometric interpretation in $\mathbb{R}^m$, the row picture offers an interpretation in $\mathbb{R}^n$. We show that, for a 3×2 system, the least squares solution lies within the corresponding triangle, and for an $(n + 1) \times n$ system, within the corresponding simplex. In the general case, the solution resides in the closed and bounded region defined by the system's hyperplanes. We further relate the least squares solution to the symmedian point in the simplex case and introduce symmedian coordinates, demonstrating their connection to trilinear and barycentric coordinates. These results establish explicit relationships between least squares solutions and classical coordinate systems, providing geometric insight into the structure of the solution.

Keywords. Least squares problem; column space; row space; symmedian point.

1. Introduction

With the matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$ $(m > n)$ and vector $\mathbf{b} \in \mathbb{R}^m$, consider an over-determined system of linear equations

$$\mathbf{Ax} = \mathbf{b}. \tag{1}$$

Then the least squares problem is formally stated as follows:

Problem 1 (Least squares problem). Given $\mathbf{A} \in \mathbb{R}^{m\times n}$ with $m > n$ and $\mathbf{b} \in \mathbb{R}^m$. Find a vector $x \in \mathbb{R}^n$ such that the function $r(x) := \|\mathbf{Ax} - \mathbf{b}\|_2$ is minimized.

The least squares problem arises in a wide range of scientific and engineering applications, including areas like data analysis [1–3], prediction [4–6], statistics [7], image processing [8, 9], signal processing and communication engineering [10, 11], and geoscience problems [12], etc.

The popular visualization of the problem is via the column picture (subsection 1.1), using the columns of matrix $\mathbf{A}$ and relating it to vector b, whereas the row interpretation of the problem has received little attention, which looks at hyperplanes corresponding to rows of $\mathbf{Ax} = \mathbf{b}$ and the geometric object formed by their intersection. So, here we develop the row picture perspective systematically.

First, we begin by recalling the column picture in subsection 1.1 and then present the relevant literature on the row picture in subsection 1.2.

1.1 *Column picture*

We can consider the matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$ as a linear mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$. The column space of $\mathbf{A}$, denoted as $\mathcal{R}(\mathbf{A})$, is defined as the collection of all vectors $\mathbf{Ax}$, for any $x \in \mathbb{R}^n$. In case the vector $\mathbf{b}$ does not lie in $\mathcal{R}(\mathbf{A})$, then we take an orthogonal projection of $\mathbf{b}$ onto $\mathcal{R}(\mathbf{A})$. Suppose that projection is denoted as $\widehat{\mathbf{b}}$, then there will always exist a vector $\mathbf{x}_0$ such that $\mathbf{Ax}_0 = \widehat{\mathbf{b}}$. That particular $\mathbf{x}_0$ is the least squares solution of eq. (1). The general picture has been explained in figure 1. There are several classic books and articles in the literature where this column interpretation of the least squares problem has been elaborately discussed (see [13, 14]).

Since $\mathbf{b} - \widehat{\mathbf{b}}$ is orthogonal to $\mathcal{R}(\mathbf{A})$, we have $\mathbf{A}^T(\mathbf{b} - \widehat{\mathbf{b}}) = \mathbf{0}$. Thus, $\mathbf{x}_0$ is a least squares solution of eq. (1) if and only if $\mathbf{x}_0$ satisfies the normal equation $\mathbf{A}^T\mathbf{Ax}_0 = \mathbf{A}^T\mathbf{b}$. If the columns are linearly independent, i.e., $\text{rank}(\mathbf{A}) = n$, then $\mathbf{x}_0 = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}$.

Published online: 12 August 2026

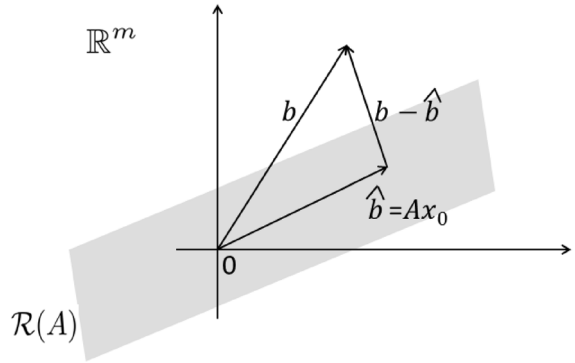


Figure 1. Geometrical interpretation of the least squares problem in the column picture.

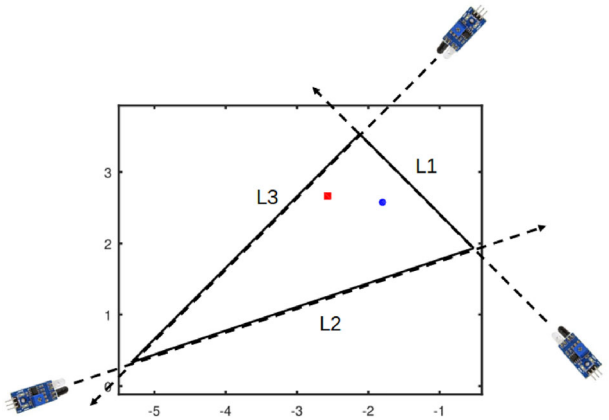


Figure 2. Least squares solution of the system $\mathbf{Ax} = \mathbf{b}$ is denoted by blue circle and $\mathbf{DAx} = \mathbf{Db}$ is denoted by red square, with $\mathbf{A} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{10}} & \frac{-3}{\sqrt{10}} \\ \frac{-2}{\sqrt{8}} & \frac{2}{\sqrt{8}} \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 1 \\ -2 \\ 4 \end{bmatrix}$ and $\mathbf{D} = \begin{bmatrix} 0.8 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$. Note that by increasing the weight of the third equation, in $\mathbf{DAx} = \mathbf{Db}$, the solution moves towards $L3$. Here, $L3$ denotes the line corresponding to the third equation of the system. As the weight of the third equation increases (through diagonal scaling), the least squares solution shifts closer to this line.

1.2 *Row picture*

The row picture is a visualization of the problem in $\mathbb{R}^n$, whereas the column picture is in $\mathbb{R}^m$. The row interpretation of the least squares problem is given briefly in Section 12.1 of [13].

In the case of a triangle, symmedians are the reflections of the medians on the angular bisectors. The symmedian point of a triangle is the intersection of symmedians [15–17]. Honsberger [18] called the symmedian point as 'one of the crown jewels of modern geometry'. He also illustrated several properties of the symmedian point in the second chapter of his book [18]. The symmedian point is also called the Lemoine point in England and France and the Grebe point in Germany. Smither [19] used the symmedian point of a triangle to find the location of mines in a harbor during the post-WWII period while working in the US Navy. The symmedian point is a least squares solution of a special system [20].

In eq. (1), each of the equations determined by rows of $\mathbf{A}$ can be treated as a hyperplane in $\mathbb{R}^n$. The least squares problem can also be geometrically interpreted using these hyperplanes. Throughout this article, we call this interpretation the row picture of the least squares problem. Note that for an invertible diagonal matrix $\mathbf{D}$, the systems $\mathbf{Ax} = \mathbf{b}$ and $\mathbf{DAx} = \mathbf{Db}$ represent the same geometric object in R^n, but their least squares solutions are different in general (since $(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b} \neq (\mathbf{A}^T\mathbf{D}^2\mathbf{A})^{-1}\mathbf{A}^T\mathbf{D}^2\mathbf{b}$ illustrated with an example in figure 2). So, it is necessary to specify $\mathbf{A}$ and $\mathbf{b}$ to interpret the problem in the row picture (not just the hyperplanes).

For a $n \times (n-1)$ matrix $\mathbf{A}$, assuming that all $(n-1) \times (n-1)$ the subsystems are solvable, the system $\mathbf{Ax} = \mathbf{b}$ corresponds to a simplex in $\mathbb{R}^{n-1}$. Specifically, each choice of $n-1$ hyperplanes intersects at a unique point, and these intersection points serve as n vertices of the simplex in $\mathbb{R}^{n-1}$. In this article, we prove that the least squares solution lies inside this simplex. If the system is scaled by a diagonal matrix $\mathbf{D}$, then $\mathbf{DAx} = \mathbf{Db}$ corresponds to the same simplex. If $\mathbf{D}$ is such that $\mathbf{DA}$ has all the rows of length ($\mathbf{L}_2$ norm) one, then the system has the symmedian point of the simplex as its least squares point [20].

1.3 *Research gap and outline of the article*

The literature on the row picture is minimal, and in this broader context, we undertake a systematic study of the row picture of the least squares problem.

Although the symmedian point is a geometric concept naturally related to the row picture, its connection to the least squares solution of a linear system has not been fully explored in the literature. We extend earlier work on symmedians by introducing symmedian coordinates and their transformations to barycentric and trilinear systems. This establishes a coherent link between classical geometry and the linear algebraic viewpoint of the least squares problem.

Our contributions are summarized as follows:

- In Section 2, we consider three equations and two unknowns where the row picture is in a 2D plane. We prove that the least squares solution lies within the triangle when all the subsystems are solvable.
- In Section 3, we extend this result to $(n-1)$ dimensions and prove that the least squares solution lies within the simplex formed by n hyperplanes.

- In Section 3.1, we define a symmedian coordinate system that arises by varying the entries of the diagonal matrix $\mathbf{D}$ in the system $\mathbf{DAx} = \mathbf{Db}$. We illustrate how the least squares solution varies as the entry of the diagonal matrix $\mathbf{D}$ varies. We give the coordinate conversion from the symmedian coordinate to the well-known and widely used barycentric coordinates and trilinear coordinates. The coordinate conversion theorems are proved in Appendix.
- In Section 4, we further generalize and prove that if every subsystem is solvable, then the least squares solution lies inside a closed and bounded region formed by the hyperplanes.

These geometric results provide a new perspective on the least squares problem.

1.4 *Notations*

The boldface letters $\mathbf{A}$ and $\mathbf{x}$ are used to denote the matrices and vectors. The letters A, B etc. denote the points in $\mathbb{R}^n$. The decomposition $\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^T$ denotes the singular value decomposition (SVD) of the matrix $\mathbf{A}$, with Σ denoting the diagonal matrix with singular values and $\mathbf{U}$ and $\mathbf{V}$ are matrices with the left and right singular vectors. $\|\cdot\|$ denotes the $\mathbf{L}_2$ norm.

2. Three equations and two unknowns in a 2D plane

In this section, we study the row picture of a 3×2 system. First, we prove Lemmas 1–3. Using these lemmas, we prove that the least squares solution lies within a triangle formed by the system in Theorem 1. We conclude the section with an example.

In $\mathbb{R}^2$, when we consider three non-parallel straight lines, pairwise intersection points form a triangle. Let the system of lines be given by the relation

$$\mathbf{Ax} = \mathbf{b}. \tag{2}$$

Here, $\mathbf{A}$ is a 3×2 real matrix, $\mathbf{x}$ is a 2×1 real vector and $\mathbf{b}$ is a 3×2 real vector. Given that all 2×2 subsystems of $\mathbf{A}$ are uniquely solvable, solution points form a triangle in $\mathbb{R}^2$. It is well known that the least squares solution of the given system $\mathbf{Ax} = \mathbf{b}$ is given by

$$\mathbf{x}_0 = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}. \tag{3}$$

Next, we prove Lemma 1, in a general setting of a simplex, in which the triangle becomes a special case when $n = 3$. Consider the singular value decomposition of the given $n \times (n-1)$ real matrix $\mathbf{A}$ as $\mathbf{U}\Sigma\mathbf{V}^T$ where $\Sigma, \mathbf{V}$ are of $n-1$ dimensions and $\mathbf{U}$ is an $n \times (n-1)$ matrix.

For $\mathbf{q} \in \mathcal{R}(\mathbf{U})^{\perp}$ (orthogonal complement of the range of $\mathbf{U}$), we have the following lemma:

Lemma 1. *Let $\mathbf{x}_0$ be the least squares solution of a given $n \times (n-1)$ linear system $\mathbf{Ax} = \mathbf{b}$. Then, $\mathbf{x}_0$ lies inside the simplex in the row picture if and only if the simplex in the row picture of the least squares problem $\mathbf{Ux} = \mathbf{q}$ contains the origin (here, $[\mathbf{U} \quad \mathbf{q}]$ is an orthogonal matrix).*

Proof. For a given system $\mathbf{Ax} = \mathbf{b}$, assume that the least squares solution $\mathbf{x}_0$ lies inside the simplex in the row picture. Now, consider $\mathbf{y} := \mathbf{x} - \mathbf{x}_0$ and $\tilde{\mathbf{b}} := \mathbf{Ax}_0 = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}$, (i.e., we move the least squares solution to the origin in $\mathbb{R}^n$). Then

$$\begin{aligned}\mathbf{Ay} &= \mathbf{b} - \tilde{\mathbf{b}},\\ &= \mathbf{b} - \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b},\\ &= (\mathbf{I} - \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T)\mathbf{b}.\end{aligned}$$

Since $(\mathbf{I} - \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T) = \mathbf{I} - \mathbf{U}\mathbf{U}^T$ is of rank one, we have $\mathbf{Ay} = \mathbf{qq}^T\mathbf{b}$ for $\mathbf{q} \in \mathcal{R}(\mathbf{U})^{\perp}$. Then, considering $\mathbf{z}:=\mathbf{\Sigma V}^T\frac{\mathbf{y}}{\mathbf{q}^T\mathbf{b}}$, we have $\mathbf{U\Sigma V}^T\frac{\mathbf{y}}{\mathbf{q}^T\mathbf{b}} = \mathbf{Uz}$. This corresponds to the system $\mathbf{Uz} = \mathbf{q}$ (i.e., A translated, stretched, and rotated version of the original system). Since $[\mathbf{U} \quad \mathbf{q}]$ is orthogonal, we can see that the origin is a least squares solution of this system. Thus, we have the simplex corresponding to this system encompassing the origin if the original system's least squares solution lies in the simplex. The converse follows immediately from the fact that $\Sigma\mathbf{V}^T$ is invertible. □

Thus, if we can prove that the origin of an orthogonalized system lies inside a triangle or a simplex, we can apply the reverse transformations of stretching, rotation, and translation to the original problem and show that the least squares solution lies inside the simplex.

Here, we prove that for any orthogonal matrix system corresponding to a triangle, the origin lies inside the simplex. Thus, it implies that the least squares solution lies within the corresponding initial simplex.

Lemma 2. *Let $[a \quad b \quad c]^T$ and $[d \quad e \quad f]^T$ be two orthonormal column vectors and $(ae - bd)(bf - ce) \neq 0$ (denominator of LHS of (4)). Then*

$$\frac{ac + df}{(ae - bd)(bf - ce)} = -1. \tag{4}$$

Proof. Given two orthonormal column vectors $[a \quad b \quad c]^T$ and $[d \quad e \quad f]^T$, we claim $ac + df + aebf + bdce - aece - bdbf = 0$ (rearrangement of the claim of the

lemma). Since the given vectors are orthonormal, we have the following equalities:

$$\begin{aligned} & ac + df + aebf + bdce - aece - bdbf \\ &= ac(1 - e^2) + df(1 - b^2) + aebf + bdce \\ &= ac(d^2 + f^2) + df(a^2 + c^2) + aebf + bdce \\ &= dc(da + be + cf) + af(cf + be + da) \\ &= 0. \end{aligned}$$

Hence, the lemma is proved. □

We consider the system $\mathbf{Pz} = \mathbf{q}$, where

$$\mathbf{P} := \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}^T,$$

$$\mathbf{z} := [x \quad y]^T,$$

and

$$\mathbf{q} := [q_1 \quad q_2 \quad q_3]^T.$$

The vertices of the triangle formed by the lines $ax + dy = q_1, bx + ey = q_2$, and $cx + fy = q_3$ are denoted as v_1, v_2, and v_3. Then we have the following lemma:

Lemma 3. *For the system* $\mathbf{Pz} = \mathbf{q}$ *with* $[\mathbf{P}\ \mathbf{q}]$ *being orthogonal, the origin lies within the triangle formed by* $\mathbf{v}_1, \mathbf{v}_2$ *and* $\mathbf{v}_3$.

Proof. Let

$$\mathbf{A} := \begin{bmatrix} a & d \\ b & e \end{bmatrix},$$

$$\mathbf{B} := \begin{bmatrix} a & d \\ c & f \end{bmatrix},$$

and

$$\mathbf{C} := \begin{bmatrix} b & e \\ c & f \end{bmatrix}.$$

Then, we have

$$\mathbf{v_1} = \frac{1}{\det \mathbf{A}} \begin{bmatrix} e & -d \\ -b & a \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \end{bmatrix}, \tag{5}$$

$$\mathbf{v_2} = \frac{1}{\det \mathbf{B}} \begin{bmatrix} f & -d \\ -c & a \end{bmatrix} \begin{bmatrix} q_1 \\ q_3 \end{bmatrix}, \tag{6}$$

$$\mathbf{v_3} = \frac{1}{\det \mathbf{C}} \begin{bmatrix} f & -e \\ -c & b \end{bmatrix} \begin{bmatrix} q_2 \\ q_3 \end{bmatrix}. \tag{7}$$

Now, we decompose $\mathbf{v_3}$ in terms of $\mathbf{v_1}$ and $\mathbf{v_2}$. In particular, we express $\mathbf{v_3}$ using the basis $\{\hat{\mathbf{v}}_1, \hat{\mathbf{v}}_2\}$ obtained using the Gram–Schmidt orthogonalization procedure on $\mathbf{v_1}, \mathbf{v_2}$. Here,

$$\hat{\mathbf{v}}_1 = \frac{\mathbf{v_1}}{\|\mathbf{v_1}\|}$$

and

$$\hat{\mathbf{v}}_2 = \frac{\mathbf{r}}{\|\mathbf{r}\|}$$

with

$$\mathbf{r} := \mathbf{v_2} - \frac{\mathbf{v_2}^T \mathbf{v_1}}{\|\mathbf{v_1}\|} \hat{\mathbf{v}}_1.$$

Now, using eqs. (5) and (7), we see that

$$\begin{aligned} \mathbf{v_3}^T \mathbf{v_1} &= \frac{1}{\det \mathbf{C} \det \mathbf{A}} [q_2 \quad q_3] \begin{bmatrix} f & -c \\ -e & b \end{bmatrix} \begin{bmatrix} e & -d \\ -b & a \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} \\ &= \frac{1}{\det \mathbf{C} \det \mathbf{A}} [q_2 \quad q_3] \begin{bmatrix} -q_2 q_3 & q_1 q_3 \\ -1 + q_2^2 & -q_2 q_1 \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} \\ &= \frac{-q_1 q_3}{\det \mathbf{C} \det \mathbf{A}}. \end{aligned}$$

Further, using Lemma 2, we have $\mathbf{v_3}^T \mathbf{v_1} = -1$. This implies

$$\mathbf{v_3}^T \hat{\mathbf{v}}_1 = \frac{-1}{\|\mathbf{v_1}\|}.$$

Similarly, we observe that

$$\mathbf{v_3}^T \mathbf{v_2} = \mathbf{v_2}^T \mathbf{v_1} = -1.$$

Thus, the component on $\hat{\mathbf{v}}_2$ is given by

$$\mathbf{v_3}^T \hat{\mathbf{v}}_2 = -\beta \|\mathbf{r}\|,$$

where

$$\beta = \frac{1}{\|\mathbf{r}\|^2} \left(1 + \frac{1}{\|\mathbf{v_1}\|^2}\right).$$

Thus,

$$\mathbf{v_3} = \frac{-1}{\|\mathbf{v_1}\|^2} \mathbf{v_1} - \beta \left(\mathbf{v_2} + \frac{1}{\|\mathbf{v_1}\|^2} \mathbf{v_1}\right).$$

This implies

$$\mathbf{v_3} + \left(\frac{1 + \beta}{\|\mathbf{v_1}\|^2}\right) \mathbf{v_1} + \beta \mathbf{v_2} = 0.$$

This says that the origin is given by a positive linear combination of the vertices $\mathbf{v_1}, \mathbf{v_2}$, and $\mathbf{v_3}$. Hence, the origin lies within the triangle formed by the given linear system $\mathbf{Pz} = \mathbf{q}$. □

Theorem 1. *Let* $\mathbf{Ax} = \mathbf{b}$ *be a* 3×2 *system with all its subsystems solvable. Then, the least squares solution of* $\mathbf{Ax} = \mathbf{b}$ *lies within the triangle that has the solution of the subsystems as vertices.*

Proof. From Lemma 1, we observe that the system has the least squares solution inside the triangle if the orthogonalized system has the origin inside the corresponding triangle. By Lemma 3, we know that the origin lies within the triangle for every 3×2 orthogonalized system. Hence, any 3×2 system has the least squares solution inside the triangle. □

Example 1 (Estimating the position of an object). We draw inspiration from the work of Smither [19] to illustrate a practical application of our framework. Consider a scenario in which three sensors, each providing a directional estimate (bearing) towards an object. Due to measurement inaccuracies, the lines defined by these bearings do not intersect at a single point; rather, they form a triangle when taken in pairs. The problem then is to estimate the most likely position of the object within this triangular region. A similar situation arose during post-WWII naval operations, where three observation stations in a harbor were used to estimate the location of falling landmines. In this example, illustrated in figure 2, we have three bearing lines given by the system

$$\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{10}} & \frac{-3}{\sqrt{10}} \\ \frac{-2}{\sqrt{8}} & \frac{2}{\sqrt{8}} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 4 \end{bmatrix}.$$

The least squares solution of the system is given by $\begin{bmatrix} -1.8012 \\ 2.5750 \end{bmatrix}$ (denoted by a blue circle in figure 2). When the third sensor has high confidence compared to the other two, denoted by the diagonal matrix

$$\mathbf{D} = \begin{bmatrix} 0.8 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix},$$

then system $\mathbf{DAx} = \mathbf{Db}$ has the least squares solution given by $\begin{bmatrix} -2.5711 \\ 2.6623 \end{bmatrix}$ (denoted by red squares in figure 2). Note that the least squares solution has moved towards the line corresponding to the higher confidence sensor.

3. $n \times (n-1)$ system: n simplex in $\mathbb{R}^{n-1}$

In this section, we generalize the result of the triangle to a simplex formed by the $n \times (n-1)$ system. We use the same methodology as in the previous section. We generalize Lemmas 2 and 3 to 4 and 5, respectively. This enables us to prove Theorem 2 using these Lemmas.

Given any $n \times (n-1)$ system with all its subsystems being solvable, we have a simplex formed by the solutions of n subsystems in $\mathbb{R}^{n-1}$. The following lemma can be considered as a generalization of Lemma 2.

Lemma 4. *Let*

$$\mathbf{A} := \begin{bmatrix} \mathbf{I} - \mathbf{u}\mathbf{u}^T & -u_1\mathbf{u} \\ -u_n\mathbf{u}^T & -u_1u_n \end{bmatrix}$$

be an invertible matrix where $\mathbf{I}$ *is an* $(n-1)$*-dimensional identity matrix,* $\mathbf{u}$ *is an* $(n-1)$*-dimensional column vector, and* u_i *is an* ith*-coordinate of* $\mathbf{u}$*. Then we have the following.*

1. $\mathbf{A}^{-1} = \begin{bmatrix} \mathbf{I} & \alpha\mathbf{u} \\ \beta\mathbf{u}^T & \gamma \end{bmatrix}$, *for some* $\alpha, \beta, \gamma \in \mathbb{R}$.
2. $\begin{bmatrix} \mathbf{u}^T & u_1 \end{bmatrix} \mathbf{A}^{-1} \begin{bmatrix} \mathbf{u} \\ u_n \end{bmatrix} = -1.$

Proof. We have $u_1, u_n \neq 0$, since $\mathbf{A}$ is a non-singular matrix. Now, it is easy to verify that for some $\alpha, \beta, \gamma \in \mathbb{R}$, $\begin{bmatrix} I & \alpha\mathbf{u} \\ \beta\mathbf{u}^T & \gamma \end{bmatrix} = \mathbf{A}^{-1}$. That is, $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}_n$. This implies that $-1 - u_1\beta = 0$ and $-\alpha u_n\mathbf{u}^T\mathbf{u} - u_1u_n\gamma = 1$. Now, evaluating the product, we have

$$\begin{bmatrix} \mathbf{u}^T & u_1 \end{bmatrix} \mathbf{A}^{-1} \begin{bmatrix} \mathbf{u} \\ u_n \end{bmatrix} = \begin{bmatrix} \mathbf{u}^T + u_1\beta\mathbf{u}^T & \alpha\mathbf{u}^T\mathbf{u} + \gamma u_1 \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ u_n \end{bmatrix}$$
$$= \mathbf{u}^T\mathbf{u} + u_1\beta\mathbf{u}^T\mathbf{u} + \alpha u_n\mathbf{u}^T\mathbf{u} + \gamma u_1u_n.$$

This implies

$$\begin{bmatrix} \mathbf{u}^T & u_1 \end{bmatrix} \mathbf{A}^{-1} \begin{bmatrix} \mathbf{u} \\ u_n \end{bmatrix} = -1.$$

Hence proved. □

Consider the system of equations $\mathbf{Az} = \mathbf{q}$ where $\mathbf{A}$ is a $n \times (n-1)$ real matrix and $\mathbf{q}$ is a $n \times 1$ column vector such that $[\mathbf{A}\ \mathbf{q}]$ is an orthogonal matrix. Taking $\mathbf{a}_i$ as the ith row of the matrix $\mathbf{A}$ and q_j as the jth coordinate of the vector $\mathbf{q}$, we have the vertex of the simplex in $n-1$ dimension of the form

$$\mathbf{v_1} = \left(\begin{bmatrix} \mathbf{a}_2 \\ \mathbf{a}_3 \\ \vdots \\ \mathbf{a}_{n-1} \\ \mathbf{a}_n \end{bmatrix}\right)^{-1} \begin{bmatrix} q_2 \\ q_3 \\ \vdots \\ q_{n-1} \\ q_n \end{bmatrix};$$

$$\ldots \mathbf{v}_i = \left(\begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_{i-1} \\ \mathbf{a}_{i+1} \\ \vdots \\ \mathbf{a}_n \end{bmatrix}\right)^{-1} \begin{bmatrix} q_1 \\ \vdots \\ q_{i-1} \\ q_{i+1} \\ q_n \end{bmatrix};$$

$$\ldots \mathbf{v}_n = \left(\begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \mathbf{a}_3 \\ \vdots \\ \mathbf{a}_{n-1} \end{bmatrix}\right)^{-1} \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ \vdots \\ q_{n-1} \end{bmatrix}.$$

Considering

$$\mathbf{w}_n = [\, q_1 \quad q_2 \quad \ldots \quad q_{n-1} \,]^T$$

and

$$\mathbf{w}_1 = [\, q_2 \quad q_3 \quad \ldots \quad q_{n-1} \quad q_n \,]^T,$$

we have

$$\mathbf{v_1}^T\mathbf{v}_n = \mathbf{w}_1^T \left(\left(\begin{bmatrix} \mathbf{a}_2 \\ \mathbf{a}_3 \\ \vdots \\ \mathbf{a}_{n-1} \\ \mathbf{a}_n \end{bmatrix}\right)^{-1}\right)^T \left(\begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_{n-1} \end{bmatrix}\right)^{-1} \mathbf{w}_n,$$

$$= \mathbf{w}_1^T \left(\left(\begin{bmatrix} \mathbf{a}_2 \\ \mathbf{a}_3 \\ \vdots \\ \mathbf{a}_{n-1} \\ \mathbf{a}_n \end{bmatrix}\right)^T\right)^{-1} \left(\begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_{n-1} \end{bmatrix}\right)^{-1} \mathbf{w}_n,$$

$$= \mathbf{w}_1^T \left(\begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_{n-1} \end{bmatrix} \begin{bmatrix} \mathbf{a}_2^T & \mathbf{a}_3^T & \ldots & \mathbf{a}_n^T \end{bmatrix}\right)^{-1} \mathbf{w}_n.$$

Now, using the fact that $\mathbf{a}_i\mathbf{a}_i^T = 1 - q_i^2$ and $\mathbf{a}_i\mathbf{a}_j^T = -q_iq_j$, we get

$$\mathbf{v_1}^T\mathbf{v}_n = \mathbf{w}_1^T \begin{bmatrix} -q_1q_2 & -q_1q_3 & \cdots & -q_1q_{n-1} & -q_1q_n \\ 1-q_2^2 & -q_2q_3 & -q_2q_4 & \cdots & -q_2q_n \\ -q_3q_2 & 1-q_3^3 & -q_3q_4 & \cdots & -q_3q_n \\ \vdots & \ldots & \ddots & \ddots & \vdots \\ -q_{n-1}q_2 & -q_{n-1}q_3 & \cdots & 1-q_{n-1}^2 & -q_{n-1}q_n \end{bmatrix}^{-1} \mathbf{w}_n.$$

Note that the resulting expression is similar to the one in Lemma 4. So, we get $\mathbf{v}_i^T\mathbf{v}_j = -1$, for all $i \neq j$.

Lemma 5. *For any unitary matrix* $\mathbf{Q} = [\mathbf{P} \quad \mathbf{q}]$ *of dimension n, the least squares solution lies within the simplex* $\mathbf{Px} = \mathbf{q}$.

Proof. From the previous arguments, we have $\mathbf{v}_i^T\mathbf{v}_j = -1$, for vertices $\mathbf{v}_i$ of the simplex in the $(n-1)$ dimension. Now, using the Gram–Schmidt procedure, let us decompose the vector $\mathbf{v}_n$ in terms of $\{\mathbf{v}_i\}_{i=1}^{n-1}$.

Let

$$\pi_{\mathbf{v}_i}(\mathbf{v}_j) = \frac{\mathbf{v}_i^T\mathbf{v}_j}{\mathbf{v}_i^T\mathbf{v}_i}\mathbf{v}_i = \frac{-\mathbf{v}_i}{\|\mathbf{v}_i\|^2}.$$

Then, the orthogonal basis formed is given by

$$\begin{aligned} \mathbf{u}_1 &= \alpha_{11}\mathbf{v}_1, \\ \mathbf{u}_2 &= \alpha_{22}(\mathbf{v_2} + \alpha_{21}\mathbf{v_1}), \\ &\vdots \qquad \vdots \\ \mathbf{u}_i &= \alpha_{i,i}\left(\mathbf{v}_i + \sum_{j=1}^{i-1}\alpha_{ij}\mathbf{v}_j\right). \end{aligned}$$

The coefficients $\alpha_{ij} > 0$ follow from

$$\pi_{\mathbf{v}_i}(\mathbf{v}_j) = \frac{-\mathbf{v}_i}{\|\mathbf{v}_i\|^2} < 0.$$

Then, we have

$$\mathbf{v}_n = \sum_{i=1}^{n-1}(\mathbf{v}_n^T\mathbf{u}_i)\mathbf{u}_i.$$

Since $\mathbf{v}_n^T\mathbf{v}_j = -1$ for $j \neq n$, we have

$$\mathbf{v}_n = \sum_{i=1}^{n-1}(-\beta_i)\mathbf{v}_i$$

for $\beta_i > 0$. Thus,

$$\mathbf{v}_n + \sum_{i=1}^{n-1}\beta_i\mathbf{v}_i = 0.$$

This implies that the origin is in the convex hull of the simplex. □

Theorem 2. *When every subsystem is solvable, the least squares solution of the $n \times (n-1)$ system $\mathbf{Ax} = \mathbf{b}$ lies inside the simplex whose vertices are the solutions of the subsystems.*

Proof. From Lemma 5, we have an orthogonalized system containing the origin inside the simplex. So, from Lemma 1, we have the original system that has the least squares solution inside the simplex. □

Figure 3 shows the least squares solution of $\mathbf{Ax} = \mathbf{b}$ for 4×3 systems within the simplex.

Theorem 3. *The least squares solution does not lie on the boundary of the simplex.*

Proof. If the least squares solution $\mathbf{x}_0$ lies on the boundary, then there is some equation j such that

$$(\mathbf{Ax}_0)_j = \mathbf{b}_j,$$
$$(\mathbf{UU}^T\mathbf{b})_j = \mathbf{b}_j.$$

Decompose the vector $\mathbf{b}$ into $\mathbf{u} + \mathbf{w}$, with $\mathbf{u}$ in the range of $\mathbf{U}$ and $\mathbf{w}$ in its orthogonal complement.

Then we have $\mathbf{UU}^T(\mathbf{u} + \mathbf{w}) = \mathbf{UU}^T\mathbf{u} = \mathbf{u}$. So its jth coordinate should be $\mathbf{u}(j)$, whereas the vector $\mathbf{b}$ has jth coordinate $\mathbf{u}(j) + \mathbf{w}(j)$. So, we must have $\mathbf{w}(j) = 0$. This implies that the vector in the null space of the matrix $\mathbf{U}$ has zero in the jth entry. Now, consider the transformed system $\mathbf{Px} = \mathbf{q}$, where $[\mathbf{P} \quad \mathbf{q}]$ forms an orthogonal matrix. Then we have $\mathbf{q}(j) = 0$. Wlog let us take $j = 1$. Then the subsystem $\mathbf{P}(2 : n, 1 : n-1)$ has a vector in a null space, which is $\mathbf{P}(1, 1 : n-1)$. Thus, the system does not form a simplex. Hence, a contradiction. □

3.1 *Symmedian coordinate system*

In this section, we first prove in Theorem 4 that the least squares solution moves along a line when a particular entry in the diagonal scaling is varied. This motivates us to define the symmedian coordinates 1. This provides a simple procedure for finding the least squares solution. Further, we give formulae for coordinate conversion between symmedian, barycentric, and trilinear coordinates. The section concludes with an example.

Consider the system $\mathbf{Ax} = \mathbf{b}$ where $\mathbf{A}$ of dimension $(n+1) \times n$ corresponds to a simplex. Let

$$\mathbf{A} = \begin{bmatrix} \mathbf{a} \\ \mathbf{B}_{n\times n} \end{bmatrix}_{(n+1)\times n}$$

and

$$\mathbf{b} = [\, b_{11} \quad \mathbf{b}_t \,]^T$$

with $\mathbf{b}_t$ being $n \times 1$. Let the corresponding vertex of the simplex be $\mathbf{v}_0 = \mathbf{B}^{-1}\mathbf{b}_t$.

Let

$$\mathbf{D} = \begin{bmatrix} \alpha & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_n \end{bmatrix},$$

with $\mathbf{I}_n$ being the $n \times n$ identity matrix. Then, we have the following theorem.

Theorem 4. *As α varies, the least squares solution of $\mathbf{DAx} = \mathbf{Db}$ moves along a straight line passing through $\mathbf{v}_0$.*

Proof. From the least squares solution of the problem $\mathbf{DAx} = \mathbf{Db}$, we have

$$\begin{aligned}\mathbf{x}_0 &= (\mathbf{A}^T\mathbf{D}^2\mathbf{A})^{-1}\mathbf{A}^T\mathbf{D}^2\mathbf{b} \\ &= \left([\, \alpha^2\mathbf{a}^T \quad \mathbf{B}^T \,] \begin{bmatrix} \mathbf{a} \\ \mathbf{B} \end{bmatrix} \right)^{-1} (\alpha^2\mathbf{a}^T b_{11} + \mathbf{B}^T\mathbf{b}_t) \\ &= (\alpha^2\mathbf{a}^T\mathbf{a} + \mathbf{B}^T\mathbf{B})^{-1} (\alpha^2\mathbf{a}^T b_{11} + \mathbf{B}^T\mathbf{b}_t).\end{aligned}$$

Now, applying the Sherman–Morrison formula and denoting $\mathbf{B}^T\mathbf{B} = \mathbf{G}$, we get

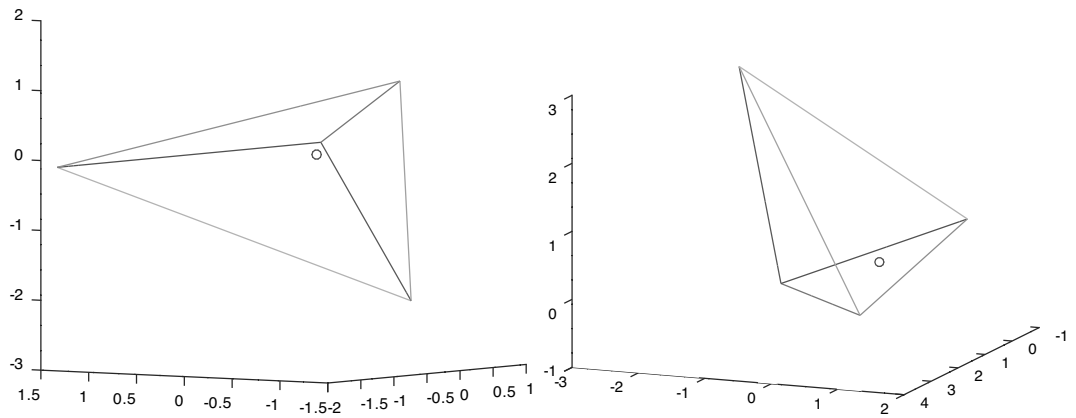


Figure 3. Least squares solution lies inside the simplex formed by a 4×3 system. Vertices of the simplex are the solution of 3×3 subsystems. Here, entries of A and b are chosen from the unit Gaussian distribution $\mathcal{N}(0, 1)$.

$$\mathbf{x}_0 = \left(\mathbf{G}^{-1} - \alpha^2 \frac{\mathbf{G}^{-1}\mathbf{a}^T\mathbf{a}\mathbf{G}^{-1}}{1+\alpha^2\mathbf{a}\mathbf{G}^{-1}\mathbf{a}^T}\right)\left(\alpha^2\mathbf{a}^T b_{11} + \mathbf{B}^T\mathbf{b}_t\right) \quad (8)$$

$$= \mathbf{v}_0 + \left(\alpha^2 b_{11}\left(1 - \frac{\alpha^2\mathbf{a}\mathbf{G}^{-1}\mathbf{a}^T}{1+\alpha^2\mathbf{a}\mathbf{G}^{-1}\mathbf{a}^T}\right) - \frac{\alpha^2\mathbf{a}\mathbf{v}_0}{1+\alpha^2\mathbf{a}\mathbf{G}^{-1}\mathbf{a}^T}\right) \times \mathbf{G}^{-1}\mathbf{a}^T. \quad (9)$$

Thus, we have the least squares solution as $\mathbf{v}_0 + f(\alpha)\mathbf{G}^{-1}\mathbf{a}^T$, which is an equation of a straight line in $\mathbf{R}^n$. Furthermore, as $\alpha \to 0$, we have $\mathbf{x}_0 \to \mathbf{v}_0$. Also, when $\alpha \to \infty$, we have

$$\mathbf{x_0} \to \mathbf{p}_\infty = \mathbf{v_0} - \frac{(\mathbf{a}\mathbf{v_0} - b_{11})}{\mathbf{a}\mathbf{G}^{-1}\mathbf{a}^T}\mathbf{G}^{-1}\mathbf{a}.$$

Note that $\mathbf{p}_\infty$ lies on the hyperplane $\mathbf{a}\mathbf{x} = b_{11}$. □

The symmedian coordinate lines are illustrated for the case of a triangle in figure 4.

Definition 1 (Symmedian coordinates.) Consider the symplex in $\mathbb{R}^{n-1}$ formed by the normalized $n \times (n-1)$ system $\tilde{\mathbf{A}}\mathbf{x} = \tilde{\mathbf{b}}$ (normalization is such that every row of $\tilde{\mathbf{A}}$ has $\mathbf{L}_2$ norm one). Then the least squares point $\tilde{\mathbf{x}}_\mathbf{0}$ (is the symmedian point) is denoted by $(1:1:1:\ldots:1)$ in the symmedian coordinates. In general, the solution of $\mathbf{D}\tilde{\mathbf{A}}\mathbf{x} = \mathbf{D}\tilde{\mathbf{b}}$ with diagonal matrix $\mathbf{D}$ having positive diagonal entries $d_1, d_2, d_3, \ldots, d_n$, the least squares solution is denoted by $(d_1 : d_2 : d_3 : \cdot : d_n)$.

Here ' : ' is used instead of ', ' to denote homogeneity, similar to the barycentric coordinates. Also note that, in barycentric coordinates, the point $(1:1:1\cdots:1)$ corresponds to the centroid, whereas here it is the symmedian point.

Next, we give a procedure for locating the least squares solution, just like the point (1, 3, 4) in $\mathbb{R}^3$ can be reached by moving from the origin to one step in the X direction, then moving 3 steps along the Y direction, followed by moving 4 steps in the Z direction.

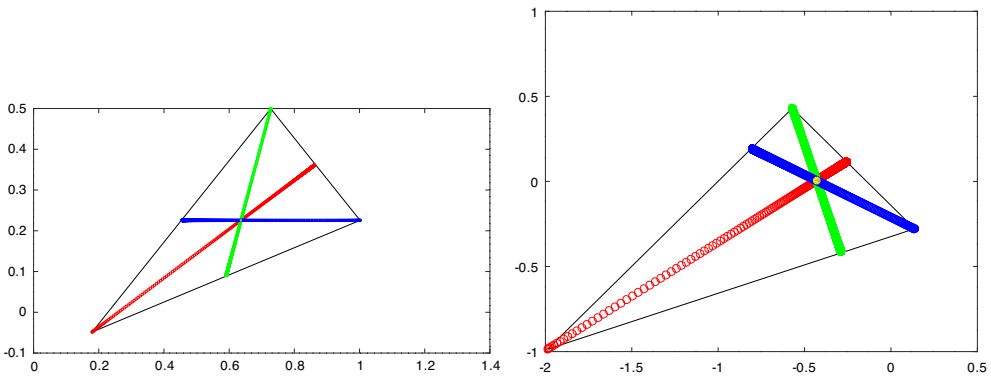


Figure 4. The systems $\mathbf{DAx} = \mathbf{Db}$, with diagonal entries $\mathbf{D}(i, i)$ varying while keeping the other two diagonal values fixed to be one. When $\mathbf{D}(i,i) \to 0$, the least squares point moves to the vertex (i.e., the intersection of the other two lines). When $\mathbf{D}(i,i) \to \infty$, the least squares point moves to the opposite edge (i.e., no weightage is given to the other two lines).

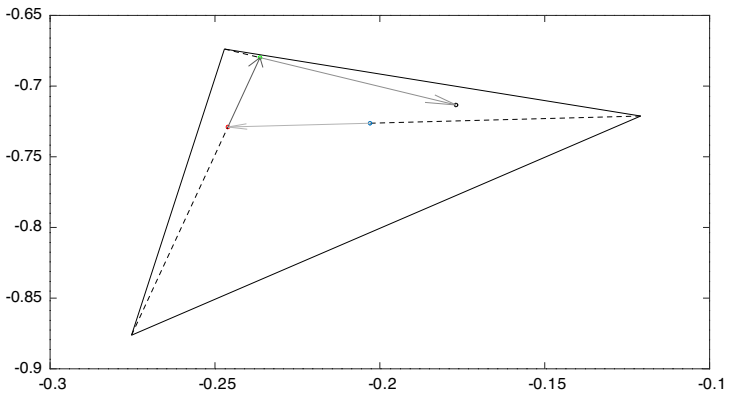


Figure 5. Starting with the symmedian point, locating the least squares solution according to the procedure in Algorithm 1, with $\mathbf{A} = \begin{bmatrix} -0.9905 & 0.1378 \\ 0.3516 & 0.9361 \\ -0.7086 & 0.7056 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 0.1519 \\ -0.7176 \\ -0.4232 \end{bmatrix}$.

Algorithm 1. Locating the least squares solution.

1: **Initialization:** $\mathbf{Ax} = \mathbf{b}$. Normalized system $\tilde{\mathbf{A}}\mathbf{x} = \tilde{\mathbf{b}}$, such that $\mathbf{D}\tilde{\mathbf{A}} = \mathbf{A}$ and $\mathbf{D}\tilde{\mathbf{b}} = \mathbf{b}$. Symmedian point S (the solution of the normalized system). $d_i = \mathbf{D}(i,i)$. The matrices $\mathbf{D}_i$ such that $\mathbf{D}_i(j,j) = \mathbf{d}_j$ for $j \leq i$ and $\mathbf{D}_i(j,j) = \mathbf{1}$ for $n \geq j > i$. Vertices of the subsystem $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$.
2: $P_0 \leftarrow S$
3: $P \leftarrow P_0$
4: **for** $k = 1$ **to** n **do**
5: Move the point P along the line joining $\mathbf{v}_k$ and P_{k-1}.
6: Stop at the point P_k which is the solution of the system $\mathbf{D}_k\tilde{\mathbf{A}}\mathbf{x} = \mathbf{D}_k\tilde{\mathbf{b}}$. (Theorem 4 implies that the solution lies on this line.
7: **end for**
8: **Output:** The final point P_n is the solution of the original system $\mathbf{Ax} = \mathbf{b}$.

Let $\mathbf{A} = \mathbf{D}\tilde{\mathbf{A}}$ and $\mathbf{b} = \mathbf{D}\tilde{\mathbf{b}}$ with diagonal matrix $\mathbf{D}$ having positive diagonal entries $d_1, d_2, d_3, \ldots, d_n$. Let $\mathbf{D}_i$ be the matrices with $\mathbf{D}_i(j,j) = d_j$ for $j \leq i$ and $\mathbf{D}_i(j,j) = 1$ for $j > i$. Let $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_n$ denote the vertices of the simplex formed by solving the $n-1 \times n-1$ subsystems obtained by removing row 1, row 2, etc. Let S denote the symmedian point. Then Algorithm 1 gives a procedure for locating the least squares solution in the simplex. This procedure is illustrated in figure 5 in the case of a triangle.

The next remark provides formulae for the coordinate conversion between symmedian, barycentric, and trilinear coordinates. These results are extended for the general simplex in Theorem 7.

Remark 1. Relation to trilinear and barycentric coordinates:

- Consider the triangle $\mathcal{ABC}$ formed by points $\mathcal{A}, \mathcal{B}$, and $\mathcal{C}$, let the corresponding opposite sides have length a, b, and c. Let $\tilde{\mathbf{A}}\mathbf{x} = \tilde{\mathbf{b}}$ be the corresponding normalized system. Let $\mathbf{D} = \text{diag}(\alpha, \beta, \gamma)$. Then the solution of the least squares problem $\mathbf{D}\tilde{\mathbf{A}}\mathbf{x} = \mathbf{D}\tilde{\mathbf{b}}$, $\mathbf{x}_0$ is given by $\alpha : \beta : \gamma$ in symmedian coordinates.
- The solution $\mathbf{x}_0$ is also given in trilinear coordinates as $\frac{a}{\alpha^2} : \frac{b}{\beta^2} : \frac{c}{\gamma^2}$. (This relation and its generalization are proved using the points given in the blog https://www.cut-the-knot.org/triangle/symmedians.shtml. The proof is given in Theorem 6).
- Thus, for a point P inside a triangle, if the symmedian coordinates are $(\alpha : \beta : \gamma)$, then the trilinear coordinates are $\frac{a}{\alpha^2} : \frac{b}{\beta^2} : \frac{c}{\gamma^2}$ and the barycentric coordinates are $\frac{a^2}{\alpha^2} : \frac{b^2}{\beta^2} : \frac{c^2}{\gamma^2}$.

Next, we present an example that uses the symmedian coordinates.

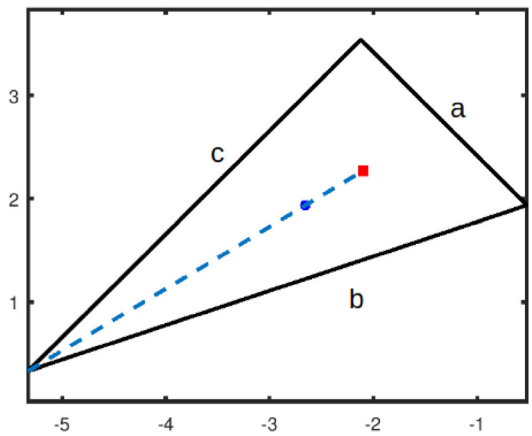


Figure 6. Centroid of the triangle given by the system $\mathbf{Ax} = \mathbf{b}$ is denoted by blue circle. It is $(1:1:1)$ in barycentric coordinates, also it is $\left(\frac{1}{a} : \frac{1}{b} : \frac{1}{c}\right)$ in trilinear and $(a:b:c)$ in symmedian coordinates with $a = 2.2639$, $b = 2.2639$, $c = 4.5279$ as respective edge lengths. When the symmedian coordinate changes to $(a+1:b:c)$, the point, shown as a red square, moves towards the opposite side. Here $\mathbf{A} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{10}} & \frac{-3}{\sqrt{10}} \\ \frac{-2}{\sqrt{8}} & \frac{2}{\sqrt{8}} \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 1 \\ -2 \\ 4 \end{bmatrix}$ and $\mathbf{D} = \begin{bmatrix} 1+a & & \\ & b & \\ & & c \end{bmatrix}$.

Example 2 (This example is inspired by computer graphics). Consider a point P inside a triangle ABC, and we are working in Barycentric coordinates given by $g_1 : g_2 : g_3$. Now, if we want to move the point along the

direction of AP towards the edge BC then we need to decrease g_1 and increase g_2 and g_3. But we do not know how much we have to increase these values. However, if we convert to the symmedian coordinates, we get $(\alpha : \beta : \gamma) = \left(\frac{a}{\sqrt{g_1}} : \frac{b}{\sqrt{g_2}} : \frac{c}{\sqrt{g_3}}\right)$ and by increasing α, we will move the point in the desired direction. In this example (illustrated in figure 6), we consider triangle formed by $\mathbf{Ax} = \mathbf{b}$ with

$$\mathbf{A} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{10}} & \frac{-3}{\sqrt{10}} \\ \frac{-2}{\sqrt{8}} & \frac{2}{\sqrt{8}} \end{bmatrix},$$

$$\mathbf{b} = \begin{bmatrix} 1 \\ -2 \\ 4 \end{bmatrix}.$$

The centroid is denoted by $(1 : 1 : 1)$ in barycentric coordinates and $(a : b : c)$ in symmedian coordinates (with $a = 2.2639$, $b = 2.2639$, $c = 4.5279$). It is indicated by a blue circle in figure 6. To move this point towards the opposite edge, colinear with the vertex of interest, we convert symmedian coordinate to $(a + 1 : b : c)$. Now the point moves in the desired direction, denoted by red squares in figure 6. The corresponding point is obtained with

$$\mathbf{D} = \begin{bmatrix} a+1 & & \\ & b & \\ & & c \end{bmatrix}.$$

4. General system

In this section, we study the row picture for general $m > n$. We prove that the least squares solution lies in a closed and bounded region formed by hyperplanes in Theorem 5. We use Lemma 6 proved in Appendix A to prove this theorem. After the theorem is given, we provide a brief illustration and discuss its implications, concluding the section with an example.

Theorem 5. *The least squares solution for a general system lies within a closed and bounded region when all the subsystems of* $\mathbf{A}$ *are solvable.*

Proof. Consider the general system $\mathbf{Ax} = \mathbf{b}$. Let $\mathbf{x}_0$ be the least squares solution. For any vector $\mathbf{y} \in \mathbb{R}^m$, if we can find a positive k, such that $(\mathbf{A}(\mathbf{x}_0 + k\mathbf{y}))_j = \mathbf{b}_j$, for some j. This implies that moving from $\mathbf{x}_0$ in the direction of $\mathbf{y}$, we reach a hyperplane boundary given by the corresponding j. It also proves that the region is closed and bounded if $k < \infty$. The implications of closed and boundedness from this ray hit property is provided in Lemma 6 (given in the Appendix). Taking into account the vector $\mathbf{Ax}_0 - \mathbf{b}$ and $-\mathbf{Ay}$, we have

$$\begin{aligned} \mathbf{Ax}_0 - \mathbf{b} &= \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b} - \mathbf{b} \\ &= (\mathbf{UU}^T - \mathbf{I})\mathbf{b}, \end{aligned}$$

and the vector $-\mathbf{Ay} = \mathbf{Uz}$ with $\mathbf{z} = -\Sigma\mathbf{V}^T\mathbf{y}$. Now, we define $\mathbf{v} := (\mathbf{UU}^T - \mathbf{I})\mathbf{b}$ and $\mathbf{w} := \mathbf{Uz}$; note that $\mathbf{v}^T\mathbf{w} = 0$ since $\mathbf{v} \in \mathit{Range}(U^{\perp})$ and $\mathbf{w} \in \mathit{Range}(U)$ are in orthogonal sub-spaces. This implies that there exist some j for which we have $\mathbf{v}(j)\mathbf{w}(j) > 0$ or the vectors $\mathbf{v}$ and $\mathbf{w}$ are complementary (i.e., they have zeros in complementary positions such that they are orthogonal). Suppose for all j if $\mathbf{v}(j)\mathbf{w}(j) \leq 0$ and vectors are not complementary, then $\mathbf{v}^T\mathbf{w} < 0$, which gives the contradiction. Suppose that vectors $\mathbf{v}$ and $\mathbf{w}$ are not complementary, then there exists a j such that $\mathbf{v}(j)\mathbf{w}(j) > 0$. We can choose that particular j to find the value of finite positive k, which is $\mathbf{v}(j)/\mathbf{w}(j)$ to satisfy $(\mathbf{A}(\mathbf{x}_0 + k\mathbf{y}))_j = \mathbf{b}_j$. Thus, proving that if we move from the least squares solution $\mathbf{x}_0$ along any direction $\mathbf{y}$, we reach a hyperplane given by the corresponding j.

What remains to prove is that any such $\mathbf{v}$ and $\mathbf{w}$ are not complementary. Here, we prove this using a contradiction. First, recall $\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^T$ and since every subsystem of $\mathbf{A}$ is solvable, $\Sigma\mathbf{V}^T$ is invertible. So, every subsystem of $\mathbf{U}$ is solvable. Now suppose $\mathbf{v}$ and $\mathbf{w}$ are complementary. We can consider the column spaces $\mathbf{U}$ and $\mathbf{U}^{\perp}$ and divide the condition into two cases.

- When $\mathbf{w}$ has more than or equal to n zeros: This gives a subsystem of $\mathbf{U}$ which has a vector in the null space, thus giving the contradiction (since $\mathrm{rank}(U) = \mathrm{rank}(A) = n$, any n rows of U must be linearly independent).
- When $\mathbf{w}$ has less than n zeros: In this case, $\mathbf{v}$ has more than $m - n$ zeros. We can have a block partition of $Q = \begin{bmatrix} \mathbf{U} & \mathbf{U}^{\perp} \end{bmatrix}$ by rearranging the rows into

$$\tilde{\mathbf{Q}} = \begin{bmatrix} \tilde{\mathbf{A}}_{(m-n)\times n} & \tilde{\mathbf{B}}_{(m-n)\times(m-n)} \\ \tilde{\mathbf{C}}_{n\times n} & \tilde{\mathbf{D}}_{n\times(m-n)} \end{bmatrix}$$

with $\tilde{\mathbf{B}}\mathbf{w_1} = \mathbf{0}$ for some $\mathbf{w_1}$. Therefore, $\tilde{\mathbf{B}}$ has a vector in the left null space. Let it be $\mathbf{w}_2^T$. Then we have

$$\mathbf{w}_2^T \begin{bmatrix} \tilde{\mathbf{A}} & \tilde{\mathbf{B}} \end{bmatrix} = \begin{bmatrix} \mathbf{w}_2^T\tilde{\mathbf{A}} & 0 \end{bmatrix}.$$

However, the vector $\mathbf{w}_2^T\tilde{\mathbf{A}}$ should be in the null space of $\tilde{\mathbf{C}}$ due to the orthogonality of the rows of the matrix $\tilde{\mathbf{Q}}$. Thus, we have a subsystem of $\mathbf{U}$ that is not solvable. This gives a contradiction. □

Figure 7 shows the example of the 4×2 system and figure 8 shows the example of the 5×2 system. In both figures, matrices $\mathbf{A}$ and $\mathbf{b}$ have random $\mathcal{N}(0, 1)$ entries. It

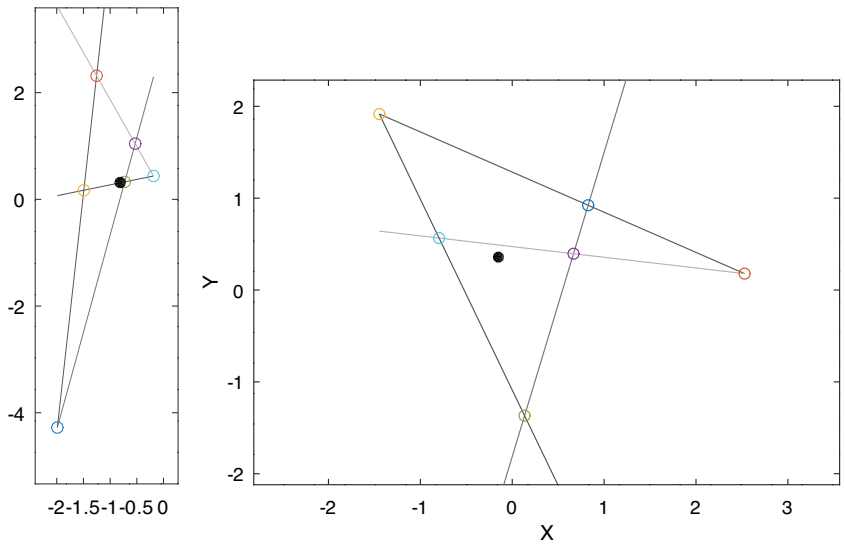


Figure 7. For 4×2 systems, we have four lines intersecting each other. Each intersection point is a solution of a 2×2 subsystem. The black point represents the least squares solution. Note that the least squares solution lies inside the quadrilateral.

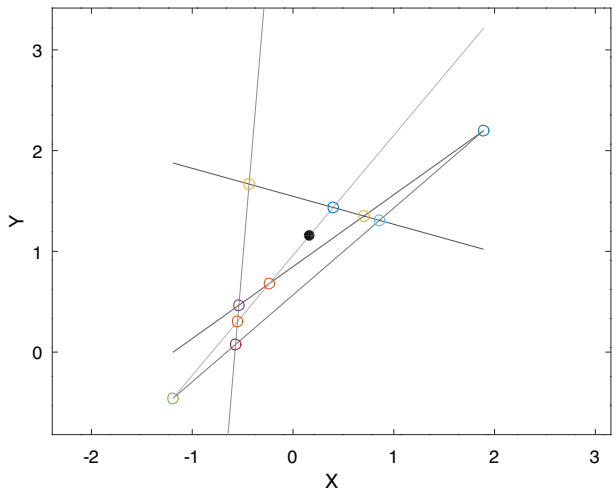


Figure 8. The 5×2 systems consist of five intersecting lines, every line intersecting another line at four points. The least squares solution, denoted by a black point, lies inside the pentagon.

can be seen that the least squares solution lies within a closed and bounded region.

Recent works ([21] and references therein) for solving least squares problems of large systems solve a smaller system by selecting the subset of rows of the matrix $\mathbf{A}$. Intuitively, from the above theorem, these rows should be formed by selecting the hyperplanes that do not change the least squares solutions much. Example 3 illustrates a special case, where an effective small system can be formed.

Example 3 (Multisensor fusion). Suppose that multiple sensors are used to take measurements and we know the set of core measurements $\mathbf{A}, \mathbf{b}$. Consider the least squares solution to the system $\mathbf{A}_B\mathbf{x} = \mathbf{b}_B$ with

$$\mathbf{A}_B = \begin{bmatrix} \mathbf{A} \\ D_1\mathbf{A} \\ D_2\mathbf{A} \\ \vdots \\ D_k\mathbf{A} \end{bmatrix}$$

and

$$\mathbf{b}_B = \begin{bmatrix} \mathbf{b} \\ D_1\mathbf{b} \\ D_2\mathbf{b} \\ \vdots \\ D_k\mathbf{b} \end{bmatrix}.$$

The system $\mathbf{A}_B\mathbf{x} = \mathbf{b}_B$ corresponds to multiple measurements with varying confidence levels. Then, the direct computation of the solution gives

$$(\mathbf{A}_B^T\mathbf{A}_B)^{-1}\mathbf{A}_B^T\mathbf{b}_B = (\mathbf{A}^T\mathbf{D}_{eff}^2\mathbf{A})^{-1}\mathbf{A}^T\mathbf{D}_{eff}^2\mathbf{b},$$

where

$$\mathbf{D}_{eff} = \sqrt{\mathbf{I} + \sum_{i=1}^{k} \mathbf{D}_i^2}.$$

5. Conclusion

The study of the row picture of the least squares problem $\mathbf{Ax} = \mathbf{b}$, for $\mathbf{A} \in \mathbb{R}^{m\times n}$, provides a complementary understanding of the solution structure in $\mathbb{R}^n$. The

analysis demonstrates that, under the assumption of solvable subsystems, the least squares solution resides within a simplex – or more generally, within a closed and bounded polytope defined by the system's hyperplanes. This formulation clarifies the structure of the problem in the row picture and complements the classical column picture. By connecting least squares with coordinate systems such as symmedian, barycentric, and trilinear, it highlights the link between linear algebra and coordinate geometry.

6. Limitations and future scope

The present work focuses on analyzing the least squares problem of minimizing the residual in the $\mathbf{L}_2$ norm. Constrained formulations are not considered in this study. Beyond the $\mathbf{L}_2$ case, extending the geometry of the row picture to other norms such as $\mathbf{L}_1$ and $\mathbf{L}_\infty$ remains an open problem and may reveal fundamentally different structures. Another promising direction is to investigate constrained and nonlinear least squares problems from the row picture perspective, where positivity, sparsity, or nonlinearities introduce additional analytical and computational challenges. A systematic study of the stability and statistical properties of row picture methods under perturbations, noise, and ill-conditioning would further enhance their applicability. Beyond the current study, the proposed framework has potential relevance to randomized algorithms [21] and optimization, particularly in the development of geometry-informed techniques for high-dimensional data analysis.

Appendices

Appendix A: Lemma on ray hit condition implies closed and boundedness

Lemma 6. *(Ray-hit condition implies closed and bounded containment). Let* $H_j = \{x \in \mathbb{R}^n : a_j^T x = b_j\}$ *be finitely many hyperplanes with* $a_j \neq 0$ *for* $j = 1, \ldots, m$, *and fix* $x_\star \in \mathbb{R}^n$. *For each j, let*

$$S_j := \{x : (a_j^T x - b_j)\,(a_j^T x_\star - b_j) \geq 0\}$$

be the closed half-space bounded by H_j *and containing* $x_\star$, *and set*

$$P := \bigcap_{j=1}^{m} S_j.$$

Assume the ray-hit condition: for every unit vector $u \in \mathbb{S}^{n-1}$ *there exist j and* $t > 0$ *with* $x_\star + t\,u \in H_j$.

Then

1. $x_\star \in P$ *and P is a non-empty closed convex set.*
2. (*Radial boundedness*). *For each unit u, define*

$$T(u) := \inf\{\, t > 0 : \ \exists j \text{ with } a_j^T(x_\star + t\,u) = b_j\}.$$

Then $T(u) \in (0, \infty)$ *for every u, and*

$$P \cap \{x_\star + t\,u : t \geq 0\} = \{x_\star + t\,u : 0 \leq t \leq T(u)\}.$$

Consequently,

$$P \subseteq x_\star + M\,\mathbb{B}, \qquad M := \max_{\|u\|=1} T(u) < \infty.$$

So P is bounded.

3. *If* $a_j^T x_\star \neq b_j$ *for all j, then* $x_\star \in \text{int}P$ (int *denotes the interior).*

Proof. (1) If $x_\star$ and x are on the same side of H_j, then $\text{sign}(a_j^T x_\star - b_j) = \text{sign}(a_j^T x - b_j)$. Hence, $(a_j^T x_\star - b_j)\,(a_j^T x - b_j) \geq 0$. Thus, S_j is the closed half-space containing $x_\star$ and $x_\star \in P$. Since the finite intersections of closed half-spaces are closed and convex, P is non-empty, closed, and convex.

(2) Fix $u \in \mathbb{S}^{n-1}$. Along $x(t) = x_\star + tu$,

$$\phi_j(t) = a_j^T x(t) - b_j = (a_j^T x_\star - b_j) + t(a_j^T u).$$

The ray meets H_j when $\phi_j(t) = 0$, i.e.

$$t_j(u) = \frac{b_j - a_j^T x_\star}{a_j^T u}, \qquad (a_j^T u \neq 0),$$

and we set $t_j(u) = +\infty$ whenever $a_j^T u = 0$ and $a_j^T x_\star \neq b_j$. By the ray-hit assumption, $T(u) := \min_j t_j(u) \in (0, \infty)$ for every u. For $0 \leq t < T(u)$, the ray is strictly inside all S_j and at $t = T(u)$ it is on some H_j. Hence,

$$P \cap \{x_\star + tu : t \geq 0\} = \{x_\star + tu : 0 \leq t \leq T(u)\}.$$

Each t_j is continuous whenever finite. Fix $u_0 \in \mathbb{S}^{n-1}$ and let $J_0 = \{j : t_j(u_0) = T(u_0)\}$. If $|J_0| = 1$, then near u_0 the same index remains minimal, so T is locally continuous. If $|J_0| > 1$, then $\{t_j : j \in J_0\}$ coincide at u_0, and continuity ensures that the indices outside J_0 remain strictly larger in a neighborhood. Thus, in the neighborhood of u_0, T is the minimum of finitely many continuous maps. Since $\min\{f, g\} = \frac{1}{2}(f + g - |f - g|)$, the pointwise minimum of finitely many continuous functions is continuous. Therefore, T is continuous at u_0. So T is continuous on the compact sphere $\mathbb{S}^{n-1}$. By the extreme value theorem ([22] Rudin, Theorem 4.16), T attains a finite maximum M, proving $P \subseteq x_\star + M\mathbb{B}$.

(3) If $a_j^T x_\star \neq b_j$ for all j, then each S_j contains an open neighborhood of $x_\star$ on the same side of H_j as $x_\star$.

Their intersection P therefore contains a neighborhood of $x_\star$. So $x_\star \in \text{int}P$. □

Appendix B: Coordinate conversions

Theorem 6. *The point $\alpha : \beta : \gamma$ in the symmedian coordinates is given by the point $\frac{a}{\alpha^2} : \frac{b}{\beta^2} : \frac{c}{\gamma^2}$ in the trilinear coordinates.*

Proof. The point P given by $\alpha : \beta : \gamma$ minimizes $x^2\alpha^2 + y^2\beta^2 + z^2\gamma^2$ with x, y, z being the distance from the sides $\mathcal{BC}, \mathcal{AC}, \mathcal{AB}$. Let the corresponding side lengths be a, b, c. Consider the Lagrange identity,

$$\begin{aligned}\left(\frac{a^2}{\alpha^2}+\frac{b^2}{\beta^2}+\frac{c^2}{\gamma^2}\right)\left(x^2\alpha^2+y^2\beta^2+z^2\gamma^2\right)\\ = (ax+by+cz)^2+\left(\frac{\gamma}{\beta}bz-cy\frac{\beta}{\gamma}\right)^2 \\ +\left(\frac{\alpha}{\gamma}cx-\frac{\gamma}{\alpha}az\right)^2+\left(\frac{\beta}{\alpha}ay-\frac{\alpha}{\beta}bx\right)^2.\end{aligned} \tag{A1}$$

Now, on the left-hand side of eq. (A1), $a, b, c, \alpha, \beta, \gamma$ are constants. On the right-hand side, $(ax + by + cz)$ is a constant. Moreover, it is twice the area of the triangle. Now, the minimum is achieved when $(x : y : z) = \left(\frac{a}{\alpha^2} : \frac{b}{\beta^2} : \frac{c}{\gamma^2}\right)$. This gives the desired result. □

Similar arguments can be used for a general simplex. In general, let $\{a_i\}_{i=1}^{n+1}$ be the surface areas ($n-1$-volume) of the faces of the n simplex. The point P in a simplex has simplex coordinates $y_1 : y_2 : \cdots : y_{n+1}$ when the distance from the jth hyperplane is proportional to y_j for $1 \le j \le n+1$. Then, we have the following theorem.

Theorem 7. *The point $\alpha_1 : \alpha_2 : \cdots : \alpha_{n+1}$ in the symmedian coordinates of the n-simplex is given by the point $\frac{a_1}{\alpha_1^2} : \frac{a_2}{\alpha_2^2} : \cdots : \frac{a_{n+1}}{\alpha_{n+1}^2}$ in the simplex coordinates.*

Proof. The point P given by $\alpha_1 : \alpha_2 : \cdots : \alpha_{n+1}$ minimizes $\sum_{i=1}^{n+1} x_i^2\alpha_i^2$ with $\{x_i\}_{i=1}^{n+1}$ being the distance from the surfaces. Consider the general form of the Lagrange identity,

$$\begin{aligned}\left(\sum_{i=1}^{n+1}\frac{a_i^2}{\alpha_i^2}\right)\left(\sum_{i=1}^{n+1}x_i^2\alpha_i^2\right) = \left(\sum_{i=1}^{n+1}a_ix_i\right)^2 \\ +\frac{1}{2}\sum_{i=1}^{n+1}\sum_{j\neq i}\left(\frac{a_ix_j\alpha_j}{\alpha_i}-\frac{a_jx_i\alpha_i}{\alpha_j}\right)^2.\end{aligned} \tag{A2}$$

Now, on the left-hand side of eq. (A2), a_i, α_i are constants. On the right-hand side, $\sum_{i=1}^{n+1} a_ix_i$ is a constant. Moreover, it is n times the volume of the simplex. Now, the minima are achieved when $x_1 : x_2 : \cdots : x_{n+1} = \frac{a_1}{\alpha_1^2} : \frac{a_2}{\alpha_2^2} : \cdots : \frac{a_{n+1}}{\alpha_{n+1}^2}$. This gives the desired result. □

Remark 2. Given the three lines that form the triangle as $\mathbf{Ax} = \mathbf{b}$ ($\mathbf{A}$ having unit length rows), and the side lengths a, b, c and angles of the triangle $\mathcal{ABC}$, any point $g_1 : g_2 : g_3$ in barycentric coordinate can be converted to global coordinates by solving a least squares problem.

The corresponding least squares problem is given by $\mathbf{DAx} = \mathbf{Db}$, with the diagonal matrix

$$\mathbf{D} = \text{diag}\left(\frac{a}{\sqrt{g_1}}, \frac{b}{\sqrt{g_2}}, \frac{c}{\sqrt{g_3}}\right).$$

To compute the global coordinates, we find the least squares solution $\mathbf{x} = (\mathbf{A}^T\mathbf{D}^2\mathbf{A})^{-1}\mathbf{A}^T\mathbf{D}^2\mathbf{b}$.

Remark 3. In general, if we are given the surface volumes of the simplex and the distance of the point to the surfaces, the point can be converted from the simplex coordinate to the global coordinate by solving a least squares problem.

In the general case, we intend to solve $\mathbf{x} = (\mathbf{A}^T\mathbf{D}^2\mathbf{A})^{-1}\mathbf{A}^T\mathbf{D}^2\mathbf{b}$. This requires a complexity of $\mathcal{O}(n^3)$.

Acknowledgements

Tanay Saha acknowledges the financial support from the DST Science and Engineering Research Board (SERB) India under grant SRG/2023/000183.

References

[1] Wolberg J 2006 Data analysis using the method of least squares: Extracting the most information from experiments. Springer. New York. Chap. 2: 31–64

[2] Liu Q, Jia Z and Wei Y 2022 Multidimensional total least squares problem with linear equality constraints. *SIAM J. Matrix Anal. Appl.* 43: 124–150

[3] Lee K and Stöger D 2023 Randomly initialized alternating least squares: Fast convergence for matrix sensing. *SIAM J. Math. Data Sci.* 5: 774–799

[4] Yeniay Ö and Göktaş A 2002 A comparison of partial least squares regression with other prediction methods. *Hacet. J. Math. Stat.* 31: 99–111

[5] Tanikella K R and Suryanarayan K 2024 Polynomial approximation of the vector field in normal form analysis of power system. *Sādhanā* 49: 306

[6] Choudhary A and Martha S C 2017 Approximate solutions of some problems of scattering of surface water waves by vertical barriers. *Sādhanā* 42: 759–768

[7] Moritz H 1978 Least-squares collocation. *Rev. Geophys.* 16: 421–430

[8] Chen G-Y, Gan M, Wang S and Chen C L P 2020 Insights into algorithms for separable nonlinear least squares problems. *IEEE Trans. Image Process.* 30: 1207–1218
[9] Gautam S and Jha K 2025 A new region-growing-based fast hybrid segmentation technique for 3D point clouds. *Sādhanā* 50: 136
[10] Lee S J 2023 Performance enhancement of codebook-based beamforming using least squares approach. *IEEE Wireless Commun. Lett.* 13: 338–342
[11] Narayana N, Sure P and Bhuma C M 2024 Experimental investigations on the applicability of LASSO-based sparse recovery for compressively sensed OFDM channel estimation. *Sādhanā* 49: 133
[12] Li Y, Jing Y, Wang Y, Yue Q and Fang Y 2025 Optimization and verification of 3D continuum-discrete coupled algorithm. *Sādhanā* 50: 1–12
[13] Boyd S and Vandenberghe L 2018 Introduction to applied linear algebra: Vectors, matrices, and least squares. Cambridge University Press, Cambridge. Chaps. 12–13, pp. 114–140
[14] Watkins D S 2004 Fundamentals of matrix computations. John Wiley & Sons. New York 3: 181–249.
[15] Bani-Yaghoub M, Rhee N H and Sadek J 2016 An algebraic method to find the symmedian point of a triangle. *Math. Mag.* 89: 197–200
[16] Weisstein E W 2004 Symmedian point. MathWorld – A Wolfram Web Resource. Available at https://mathworld.wolfram.com
[17] Kimberling C 2010 Trilinear distance inequalities for the symmedian point, the centroid, and other triangle centres. *Forum Geom.* 10: 135–139
[18] Honsberger R 1995 Episodes in nineteenth and twentieth-century Euclidean geometry. Cambridge University Press, Cambridge. Vol. 37, Chap. 7, pp. 53–79
[19] Smither R K 2011 The symmedian point: Constructed and applied. *College Math. J.* 42: 115–117
[20] Sadek J 2017 Isogonal conjugates in an n-simplex. *Forum Geom.* 17: 425–432
[21] Epperly E N 2024 Fast and forward stable randomized algorithms for linear least-squares problems. *SIAM J. Matrix Anal. Appl.* 45: 1782–1804
[22] Rudin W 1976 Principles of mathematical analysis, 3rd ed. McGraw-Hill, New York. Chap. 4, Thm. 16

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.