Preprints.org

Article

Eliminating Iterative Methods: A Closed-Form Solution to Multivariate Quaternionic Least Squares

Edward Solomon [*] and Aiden DeGrace

Article

Eliminating Iterative Methods: A Closed-Form Solution to Multivariate Quaternionic Least Squares

Edward Solomon [1,*] and Aiden DeGrace [2]

[1] Stony Brook University
[2] Independent Researcher
* Correspondence: EdwardKingSolomon@gmail.com

Abstract: These Volumes present the generalized form of the cubic equation proposed by Li-Ping Huang and Wasin So for solving quaternionic quadratic equations. Utilizing a natural transformation from the standard orthogonal basis $\vec{i}, \vec{j}, \vec{k}$ to $\vec{\lambda}, \vec{\mu}, \vec{v}$, which maintains quaternion multiplication rules, we derive the general and depressed forms of the quadratic equation. The general form is expressed as $\vec{f} = \vec{x}^2 + \vec{x}\vec{b} + \vec{a}\vec{x}$, whilst the depressed form takes the structure $\vec{c} = \vec{t}^2 - \vec{t}\vec{v} + \vec{v}\vec{t}$, where $\vec{c} = \vec{f} + \vec{a}\vec{b}$ and $\vec{u} = \frac{1}{2}(\vec{a} + \vec{b})$, $\vec{v} = \frac{1}{2}(\vec{a} - \vec{b})$, and $\vec{t} = \vec{x} - \vec{u}$. Note that $\vec{u}$ and $\vec{v}$ are the Vector Mean Sum and Difference of $\vec{a}$ and $\vec{b}$, a crucial observation that also allows the definition of logarithms of quaternionic logics (logics that are associative, but not commutative). Considering complex roots of the cubic equation yields additional solutions in the form of complex quaternions (biquaternions), resulting in six unique solutions to $\vec{c} = \vec{t}^2 - \vec{t}\vec{v} + \vec{v}\vec{t}$. Furthermore, we derive the Closed Form Solution to Quaternionic Least Squares, both for real and complex quaternions. We also extend our analysis to Tessarine Least Squares, Euler's Formula for Reflectors, and establish the well-defined properties such as the conjugate, reciprocal, and logarithm of tessarines. The inaccuracies surrounding tessarines, quaternions, and other hypercomplex algebras in existing literature **necessitate** this Introductory Volume to rectify these misconceptions. Consequently, the introductory volume is rather lengthy and narrative-driven. In contrast, the subsequent volumes are relatively short, primarily presenting Theorems and Proofs with minimal narrative.

Keywords: quaternion; hypercomplex least squares; complex quaternion coefficients; biquaternions; logarithm of a quaternion; Rosenfeld Projective planes; tessarine; logarithm of a tessarine; tessarine conjugate and reciprocal

---

1. Introduction: Reduced Paper Concerning Quaternionic Least Squares Only

On behalf of my great friend, Aiden DeGrace—who was just seventeen when I first met him and who pioneered breakthroughs in commutative tessarine logic despite many challenges—I have chosen to first publish a paper on quaternion least squares independently. This will set the stage for bringing his work on tessarines to the forefront in future publications.

I first announced the Closed-Form Solution to Quaternionic Least Squares in November 2022, after which I was invited to lecture at the JMM 2023 Conference on January 7, 2023, at the Boston Sheraton Hotel.

My interest in the subject was piqued by the 2008 paper *An Iterative Algorithm for the Least Squares Problem in Quaternionic Quantum Theory*, which employed a recursive solution to find the local minima of error in the quaternionic regression:

$$\vec{z}_t = \vec{x}_t \vec{c} \vec{y}_t$$

where $\vec{c}$ is a "middle-handed" constant exhibiting opposite chirality to the quaternionic data points $\vec{x}_t$ and $\vec{y}_t$. This formulation explains an observed result $\vec{z}_t$, where t is the data point index.

Ultimately, the constant obtained by their method is not the optimal solution, but rather the best convergent solution achievable based on the initial parameters, similar to how a Neural Network seeks a functional solution rather than the best possible one.

They say necessity is the mother of invention, and in mathematics, the mother of discovery. I had a personal need not only to solve the middle-handed case of the univariate regression of $\vec{z}_t = \vec{x}_t \vec{c} \vec{y}_t$, but a multivariate system of quaternionic constants of left, right and middle-handed chirality, in particular, the bivariate quadratic regression:

$$\vec{z}_t = \vec{c}_0 + \vec{c}_1 \vec{x}_t + \vec{y}_t \vec{c}_2 + \vec{x}_t^2 \vec{c}_3 + \vec{x}_t \vec{c}_4 \vec{y}_t + \vec{y}_t \vec{c}_5 \vec{x}_t + \vec{c}_6 \vec{y}^2.$$

The above regression has two left-handed constants, $\vec{c}_1$ and $\vec{c}_6$; two right-handed constants, $\vec{c}_2$ and $\vec{c}_3$, and two middle-handed constants $\vec{c}_4$ and $\vec{c}_5$.

Clearly inaccurate recursive solutions could not be used to resolve such a thing, and thus I set out to solve the general closed form solution to such quandaries.

2. Fundamentals of Quaternionic Matrices and Chain Multiplication

Before proceeding with the derivation of quaternionic least squares, we must first review the general rules for quaternionic matrices in this section, and discuss how to solve a system of linear equations (with no error) that includes quaternionic entries as embedded 4x4 block submatrices in the following section.

2.1. Mother Nature's Matrix Form; and the Zero and Forward Vectors

Although we shall be using left-handed and right-handed matrix forms in this particular publication (the human forms) in the initial derivation of quaternionic least squares, I initially employed the Natural Form, where there is no distinction between left or right-handed matrices. Instead, there is only a single matrix form — the left-handed form, which is also the universal form.

The multiplication of $\vec{a}\vec{b}$ is simply the non-commutative multiplication of two universal matrices, AB.

Definition 1. *The Forward Vector*
There is no concept of a "Real Part" of a quaternion. Instead, there exists a forward part, denoted as $z\vec{q}$, where $\vec{q}$ is the forward vector. This forward vector represents the direction the observer is facing, extended by one unit of length, as defined by the observer.

Definition 2. *The Observer, the Zero Vector*
There is no concept of zero or $\vec{0}$ in the sense of "nothingness" or "absence of value." Instead, $\vec{0}$ represents the position of the observer, who can face any direction. The idea of the zero vector having an infinite number of potential facings may seem unusual, but it is precisely the facing of the zero vector that determines the direction of a derivative over the complex numbers or the quaternions.

We will explore the concepts of zero and the forward vector in greater detail in future publications. However, for the purposes of this reduced publication, let's proceed to the Universal Matrix Form. For now, the key takeaway from the above is that $\vec{q}$ is synonymous with the traditional definition of the "real part."

Definition 3. *The Universal Matrix form of a Quaternion*
Let $\vec{a} = a_0\vec{q} + a_1\vec{i} + a_2\vec{j} + a_3\vec{k}$.

Then the Universal Matrix Form of $\vec{a}$ is given by: $[\vec{a}]_{4\mathbb{H}} = \begin{pmatrix} +a_0 & -a_1 & -a_2 & -a_3 \\ +a_1 & +a_0 & +a_3 & -a_2 \\ +a_2 & -a_3 & +a_0 & +a_1 \\ +a_3 & +a_2 & -a_1 & +a_0 \end{pmatrix}$

Definition 4. *Natural Quaternionic Multiplication*

Given $\vec{z} = \vec{x}\vec{y}$, then $\vec{z}$ is a 4x4 matrix, given by:

$$
\begin{pmatrix}
+z_0 & -z_1 & -z_2 & -z_3 \\
+z_1 & +z_0 & +z_3 & -z_2 \\
+z_2 & -z_3 & +z_0 & +z_1 \\
+z_3 & +z_2 & -z_1 & +z_0
\end{pmatrix}
=
\begin{pmatrix}
+x_0 & -x_1 & -x_2 & -x_3 \\
+x_1 & +x_0 & +x_3 & -x_2 \\
+x_2 & -x_3 & +x_0 & +x_1 \\
+x_3 & +x_2 & -x_1 & +x_0
\end{pmatrix}
\begin{pmatrix}
+y_0 & -y_1 & -y_2 & -y_3 \\
+y_1 & +y_0 & +y_3 & -y_2 \\
+y_2 & -y_3 & +y_0 & +y_1 \\
+y_3 & +y_2 & -y_1 & +y_0
\end{pmatrix}
$$

2.2. Traditional Human Matrix Forms

Definition 5. *The Left-Handed Matrix Form (same as universal form)*

Let $\vec{a} = a_0\vec{q} + a_1\vec{i} + a_2\vec{j} + a_3\vec{k}$, then the Left-Handed Matrix of $\vec{a}$ is:

$$
[\vec{a}]_{4\mathbb{H}L} =
\begin{pmatrix}
+a_0 & -a_1 & -a_2 & -a_3 \\
+a_1 & +a_0 & +a_3 & -a_2 \\
+a_2 & -a_3 & +a_0 & +a_1 \\
+a_3 & +a_2 & -a_1 & +a_0
\end{pmatrix}
$$

Definition 6. *The Right-Handed Matrix Form*

Let $\vec{b} = b_0\vec{q} + b_1\vec{i} + b_2\vec{j} + b_3\vec{k}$, then the Right-Handed Matrix Form of $\vec{b}$ is:

$$
[\vec{b}]_{4\mathbb{H}R} =
\begin{pmatrix}
+b_0 & -b_1 & -b_2 & -b_3 \\
+b_1 & +b_0 & -b_3 & +b_2 \\
+b_2 & +b_3 & +b_0 & -b_1 \\
+b_3 & -b_2 & +b_1 & +b_0
\end{pmatrix}
$$

Observe that the lower 3x3 submatrix of the right-handed form is the transpose of the lower 3x3 submatrix of the left-handed form (entries in red retain the same signs, entries in purple, blue and green transpose).

Also observe the that Quaternionic Continuum of $\mathbb{H}$ is invoked in every matrix form and that the universal form (even though it is the same as the left-handed form) has no declaration of of "L" or "R," which signifies that the author is operating in the universal notation.

The invocation of $2^n\mathbb{C}, 4^n\mathbb{H}, 8^n\mathbb{O}$... tell us if we have a commutative tessarine logic, an associative quaternionic logic, or a non-associative octonion logic. The degree of the declaration tells us how many nested folds of the logic are present. Quaternions are $4^1\mathbb{H}$ and Biquaternions are $2^1\mathbb{C} \times 4^1\mathbb{H}$, and more generally $2^x\mathbb{C} \times 4^y\mathbb{H} \times 8^z\mathbb{O} \times 16^w\mathbb{S}$... defines the Rosenfeld Projective Plane which governs the algebra.

For this reduced publication, we are only concerned with the least squares regression of datalists in the form of $4^1\mathbb{H}$ and/or $2^1\mathbb{C} \times 4^1\mathbb{H}$ (the former is a subset of the latter).

2.3. Least Squares is a Misnomer: Least Conjugates Regression Is the Correct Terminology

The Human Forms of matrices are more suitable for deriving solutions, such as finding the roots of biquaternionic polynomials and performing least squares for both regular quaternions and biquaternions. This is because the first column of the adjugate matrix of a quaternion (or biquaternion) is equivalent to the conjugate of the quaternion (or biquaternion), while the first column of the inverse matrix corresponds to the reciprocal of the quaternion (or biquaternion).

This distinction is crucial, as not all biquaternions possess reciprocals, but they always have adjugate matrices and, therefore, always have conjugates.

Why does this matter? The term "Least Squares Regression" is a misnomer that originated from regressions over the real numbers. In reality, what we are performing is "Least Conjugates Regression." This means that Least Squares is always well-defined over the biquaternions, even when some data points involve zero divisors (non-invertible matrices), so long as one knows how to evaluate the limits of an indeterminate matrix (the inverse of a zero matrix) as the algebraic structure converges towards singularity.

2.4. Multiplication Rules for Three Variables

Definition 7. *Given $\vec{w} = \vec{x}\vec{z}\vec{y}$, this can be written in matrix form as:*

1. $\vec{w} = (\vec{x}\vec{z})\vec{y} = [\vec{x}]_{4\mathbb{H}L}[\vec{z}]_{4\mathbb{H}L}\vec{y} \neq [\vec{z}]_{4\mathbb{H}L}[\vec{x}]_{4\mathbb{H}L}\vec{y}.$
2. $\vec{w} = (\vec{x})\vec{z}(\vec{y}) = [\vec{x}]_{4\mathbb{H}L}[\vec{y}]_{4\mathbb{H}R}\vec{z}.$
3. $\vec{w} = (\vec{x})\vec{z}(\vec{y}) = [\vec{y}]_{4\mathbb{H}R}[\vec{x}]_{4\mathbb{H}L}\vec{z}.$
4. $\vec{w} = \vec{x}(\vec{z}\vec{y}) = [\vec{y}]_{4\mathbb{H}R}[\vec{z}]_{4\mathbb{H}R}\vec{x} \neq [\vec{z}]_{4\mathbb{H}R}[\vec{y}]_{4\mathbb{H}R}\vec{x}.$

That items two and three prove (trivially) that $[\vec{x}]_{4\mathbb{H}L}[\vec{y}]_{4\mathbb{H}R} = [\vec{y}]_{4\mathbb{H}R}[\vec{x}]_{4\mathbb{H}L}$, such that that product of a left-handed and right-handed matrix is commutative.

That the first item declares that $\left(\prod_{t=1}^{t=n}\vec{x}_t\right)\vec{c} = \left(\prod_{t=1}^{t=n}[\vec{x}_t]_{4\mathbb{H}L}\right)\vec{c}$ and that the product of two left-handed matrices is not commutative.

That the fourth item declares that $\vec{c}\left(\prod_{t=n}^{t=1}\vec{y}_t\right) = \left(\prod_{t=1}^{t=n}[\vec{y}_t]_{4\mathbb{H}R}\right)\vec{c}$ and that the product of two right-handed matrices is not commutative (also observe the reversal of the limits of the product expression on the left and right hand sides of the equation, as well as the the reversed placement of $\vec{c}$).

2.5. The Partial Commutativity of Product Chains Surrounding an Anchor

Definition 8. *Anchor, Left Chain and Right Chain*

1.   *Let $\vec{z}$ be the anchor of a multiplicative system, such that all other quaternions being multiped against it are perceived as external forces, and that if there are no quaternions being multiped against $\vec{z}$, then $\vec{z}$ is "at rest."*
2.   *Let set X be a set of quaternions $\{\vec{x}_1, \vec{x}_2...\vec{x}_m\}$ be the Left Chain, which are the left-handed forces that have acted on $\vec{z}$, with the indices from 1 to m reflecting the reverse temporal ordering of the left-handed actions taken (that is, $\vec{x}_1$ is the most recent left-handed action).*
3.   *Let set Y be a set of quaternions $\{\vec{y}_1, \vec{y}_2...\vec{y}_n\}$ be the Right Chain, which are the right-handed forces that have acted on $\vec{z}$, with the indices from 1 to n reflecting the reverse temporal ordering of the right-handed actions taken (that is, $\vec{y}_1$ is the most recent right-handed action).*

Lemma 1. *The Left Chain Lemma That the product of any two left-handed matrices, remains a left-handed matrix, such that $\vec{x}_1\vec{x}_2 = \vec{x}_{1,2}$, and more generally that any partial product of the left chain, from $a \geq 1$ to $(a+b) \leq m$:*

$$\left(\prod_{t=a}^{t=a+b}\vec{x}_t\right)\vec{z} = \vec{x}_{(a,a+1,a+2,...a+b)}\vec{z} = \left(\prod_{t=a}^{t=a+b}[\vec{x}_t]_{4\mathbb{H}L}\right)\vec{z}$$

Where $\vec{z}$ is the original anchor if $(a+b) = m$, otherwise $\vec{z} = \vec{x}_{a+b+1}$.

Lemma 2. *The Right Chain Lemma That the product of any two right-handed matrices, remains a right-handed matrix, such that $\vec{y}_2\vec{y}_1 = \vec{y}_{2,1}$, and more generally that any partial product of the right chain, from $c \geq 1$ to $(c+d) \leq n$:*

$$\vec{z}\left(\prod_{t=c+d}^{t=a}\vec{y}_t\right) = \vec{z}\vec{y}_{(c+d,-1+c+d,-2+c+d,...c)} = \vec{z}\left(\prod_{t=c}^{t=c+d}[\vec{y}_t]_{4\mathbb{H}R}\right)$$

Where $\vec{z}$ is the original anchor if $(a+b) = n$, otherwise $\vec{z} = \vec{y}_{c+d+1}$.

Theorem 1. *The Alternating Chain Theorem*
   That $\vec{w} = (\vec{x}_1\vec{x}_2\vec{x}_3...\vec{x}_m)\vec{z}(\vec{y}_n...\vec{y}_3\vec{y}_2\vec{y}_1)$ is given by:

1. $\vec{w} = \left(\prod_{s=1}^{s=m}[\vec{x}_s]_{4\mathbb{H}L}\right)\left(\prod_{t=1}^{t=n}[\vec{y}_t]_{4\mathbb{H}R}\right)\vec{z}$
2. $\vec{w} = \left(\prod_{t=1}^{t=n}[\vec{y}_t]_{4\mathbb{H}R}\right)\left(\prod_{s=1}^{s=m}[\vec{x}_s]_{4\mathbb{H}L}\right)\vec{z}$
3.   *Given $m = n$ that $\vec{w} = \left(\prod_{t=1}^{t=m}[\vec{x}_t]_{4\mathbb{H}L}[\vec{y}_t]_{4\mathbb{H}R}\right)\vec{z}$*
4.   *Given $m = n$ that $\vec{w} = \left(\prod_{t=1}^{t=m}[\vec{y}_t]_{4\mathbb{H}R}[\vec{x}_t]_{4\mathbb{H}L}\right)\vec{z}$*
5.   *Given that $m < n$ and $d = n - m$, that $\vec{w} = \left(\prod_{s=1}^{s=d}[\vec{y}_s]_{4\mathbb{H}R}\right)\left(\prod_{t=1}^{t=m}[\vec{x}_t]_{4\mathbb{H}L}[\vec{y}_{t+d}]_{4\mathbb{H}R}\right)\vec{z}$*
6.   *Given that $m < n$ and $d = n - m$, that $\vec{w} = \left(\prod_{s=1}^{s=d}[\vec{y}_s]_{4\mathbb{H}R}\right)\left(\prod_{t=1}^{t=m}[\vec{y}_{t+d}]_{4\mathbb{H}R}[\vec{x}_t]_{4\mathbb{H}L}\right)\vec{z}$*

7. *Given that $n < m$ and $d = m - n$, that $\vec{w} = \left(\prod_{s=1}^{s=d}[\vec{x}_s]_{4\mathbb{H}L}\right)\left(\prod_{t=1}^{t=m}[\vec{x}_{t+d}]_{4\mathbb{H}L}[\vec{y}_t]_{4\mathbb{H}R}\right)\vec{z}$*

8. *Given that $n < m$ and $d = m - n$, that $\vec{w} = \left(\prod_{s=1}^{s=d}[\vec{x}_s]_{4\mathbb{H}L}\right)\left(\prod_{t=1}^{t=m}[\vec{y}_t]_{4\mathbb{H}R}[\vec{x}_{t+d}]_{4\mathbb{H}L}\right)\vec{z}$*

Proof of Statement One: That the chain multiplication of left-handed matrices remains a left-handed matrix of a new quaternion (The Left Chain Lemma), and that the chain multiplication of right-handed matrices remains a right-handed matrix of a new quaternion (The Right Chain Lemma), such that $\vec{w} = (\vec{x}_1\vec{x}_2\vec{x}_3...\vec{x}_m)\vec{z}(\vec{y}_n...\vec{y}_3\vec{y}_2\vec{y}_1) = \vec{x}_0\vec{z}\vec{y}_0$, where the zeroth index is reserved for the full multiplication of an X or Y chain set.

Statement Two is a Corollary of the Proof of the first statement. Since we already know that left and right handed matrices, when multiplied, are commutative, that: $\vec{w} = \vec{x}_0\vec{z}\vec{y}_0 = (\vec{x}_0\vec{z})\vec{y}_0 = \vec{x}_0(\vec{z}\vec{y}_0)$, which upholds the associative property of the quaternions (more generally, the associativity of the quaternions compels the commutative property of a left-matrix times a right-matrix!).

Statement Three upholds the expected commutativity of nested left-right actions against $\vec{z}$, specifically that: $\vec{w} = \vec{x}_1(\vec{x}_2...(\vec{x}_{-2+n}(\vec{x}_{-1+n}(\vec{x}_n\vec{z}\vec{y}_n)\vec{y}_{-1+n})\vec{y}_{-2+n})...\vec{y}_2)\vec{y}_1$.

Statement Four is a corollary of Statement Three, since each nested iteration of a left-matrix times a right matrix is commutative with the right-matrix of the same iteration times the left-matrix of the same iteration. **Statements Three and Four are also what concern us in regards to Quaternionic Least Squares Regression when solving for a middle-handed constant.**

Statement Five is a corollary of Statement Three combined with the Right Chain Lemma. The span from $t = 1 + d$ to $t = m$ is treated as the symmetric case (statement three), whose left-right matrix alternation is preceded by the product of the initial s unlinked elements in Y.

Statement Six is corollary of Statement Five, simply exchanging the inner left-right matrix alternation with right-left due to the locally commutative structure within each t iteration.

Statement Seven is a corollary of Statement Three combined with the Left Chain Lemma. The span from $t = 1 + d$ to $t = m$ is treated as the symmetric case (statement three), whose left-right matrix alternation is preceded by the product of the initial s unlinked elements in X.

Statement Eight is corollary of Statement Seven, simply exchanging the inner left-right matrix alternation with right-left due to the locally commutative structure within each t iteration.

Corollary 1 (The Commutative Chaotic Chain Theorem). *More generally, the associative property of the quaternions enforces a chaotic list of commutative permutations of matrix multiplication with segmented chains of X and Y, such that:*

$$\vec{w} = \vec{x}_1\vec{x}_2\vec{x}_3\ldots\vec{x}_m\vec{z}\vec{y}_n\ldots\vec{y}_3\vec{y}_2\vec{y}_1$$

can be rewritten with multiple, unequal, and non-symmetric nested chains of parentheses, provided that the vectors themselves remain in the given order. This leads to a matrix multiplication paradigm, where each unequal chain within the set X_0 is now expressed as m subsets X_s, and each unequal chain within the set Y_0 is now expressed as n subsets Y_t, with no constraint governing the equality or inequality of m versus n:

$$\vec{w} = \left(\prod_{s=1}^{m} X_s\right)\left(\prod_{t=1}^{n} Y_t\right)\vec{z},$$

where:

1. *Any s iteration for some X_s can be placed anywhere between any t iteration for some Y_t and $Y_{t\pm1}$, provided that no X_s iteration is placed out of order with respect to all other X iterations.*

2. *Any t iteration for some Y_t can be placed anywhere between any s iteration for some X_s and $X_{s\pm1}$, provided that no Y_t iteration is placed out of order with respect to all other Y iterations.*

This corollary is particularly useful in Least Squares Regression when encountering a seemingly chaotic arrangement of known terms being multiplied against some constant anchor $\vec{c}$, which we seek to resolve using least squares.

Although this corollary will not feature in this publication, it is nevertheless important for those who wish to apply quaternionic least squares in the real world (a laboratory environment of real things).

2.6. The Conjugate and Reciprocal of a Hypercomplex Algebra are the Adjugate and Inverse Matrices

The determinant of a matrix represents the n-volume volume enclosed by an n-dimensional parallelepiped formed by the basis vectors. From the perspective of a given observer, this volume can be interpreted as that of an n-cube in some affine basis.

To understand this more deeply, consider a fundamental question: What is the definition of division? At its core, division measures one vector in terms of another.

For example, take two ordinary complex numbers. The expression $\vec{u}\vec{v}^{-1}$ means measuring the vector $\vec{u} = a\vec{q} + b\vec{i}$ relative to another vector $\vec{v} = c\vec{q} + d\vec{i}$. This operation represents a change of reference frame: Instead of taking $\vec{q}$ as the forward direction, the new forward vector for a secondary observer (sharing the same origin) is $\vec{v}$.

Now suppose $\vec{u}$ has a magnitude of 6 and a facing of 60 degrees, such that

$$\vec{u} = 6\left(\vec{q}\cos\frac{4\pi}{12} + \vec{i}\sin\frac{4\pi}{12}\right).$$

Similarly, let $\vec{v}$ have a magnitude of 1.5 and a direction of 45 degrees:

$$\vec{v} = 1.5\left(\vec{q}\cos\frac{3\pi}{12} + \vec{i}\sin\frac{3\pi}{12}\right).$$

Then, computing $\vec{w} = \vec{u}\vec{v}^{-1}$ yields: $\vec{w} = 4\left(\vec{q}\cos\frac{1\pi}{12} + \vec{i}\sin\frac{1\pi}{12}\right)$, which describes a vector of magnitude 4 with a direction of of 15 degrees.

Yet, fundamentally, nothing has changed. The vectors $\vec{u}$ and $\vec{w}$ are identical in their spatial meaning—only their representation differs. Vector $\vec{u}$ is described from the perspective of an observer who takes $\vec{q}$ as the forward direction at unit length, whilst $\vec{w}$ is expressed from the viewpoint of an observer for whom $\vec{v}$ is forward and at unit length (that is, both observers would walk to the exact same location, even though they perceive them differently!).

If standard division is the measuring of one isolated vector in terms of another isolated vector, which changes the frame of reference, what then is the definition of an inverse matrix (which is the general definition of division)?

Definition 9. *Physical Interpretation of an Inverse Matrix (Reduced Definition for Square Matrices)*
Let X be an $n \times n$ matrix, and let Y be an $n \times n$ identity matrix, then (in this publication we use the zero index and write the row index before the column index as is done in C++):

$$X = Y \iff \begin{pmatrix} (a_{0,0})\vec{x}_0 & (a_{0,1})\vec{x}_1 & (a_{0,2})\vec{x}_2 & \cdots & (a_{0,n-1})\vec{x}_{n-1} \\ (a_{1,0})\vec{x}_0 & (a_{1,1})\vec{x}_1 & (a_{1,2})\vec{x}_2 & \cdots & (a_{1,n-1})\vec{x}_{n-1} \\ (a_{2,0})\vec{x}_0 & (a_{2,1})\vec{x}_1 & (a_{2,2})\vec{x}_2 & \cdots & (a_{2,n-1})\vec{x}_{n-1} \\ \ldots\vec{x}_0 & \ldots\vec{x}_1 & \ldots\vec{x}_2 & \ldots\vec{x}_m\ldots & \ldots\vec{x}_{n-1} \\ (a_{n-1,0})\vec{x}_0 & (a_{n-1,1})\vec{x}_1 & (a_{n-1,2})\vec{x}_2 & \cdots & (a_{n-1,n-1})\vec{x}_{n-1} \end{pmatrix} = \begin{pmatrix} (1\vec{y}_0) & 0\vec{y}_1 & 0\vec{y}_2 & \cdots & 0\vec{y}_{n-1} \\ 0\vec{y}_0 & (1\vec{y}_1) & 0\vec{y}_2 & \cdots & 0\vec{y}_{n-1} \\ 0\vec{y}_0 & 0\vec{y}_1 & (1\vec{y}_2) & \cdots & 0\vec{y}_{n-1} \\ 0\vec{y}_0 & 0\vec{y}_1 & 0\vec{y}_2 & \ldots(1\vec{y}_m)\ldots & 0\vec{y}_{n-1} \\ 0\vec{y}_0 & 0\vec{y}_1 & 0\vec{y}_2 & \cdots & (1\vec{y}_{n-1}) \end{pmatrix}$$

The above states that there exists an observer, X, who perceives the vectors $1\vec{x}_0, 1\vec{x}_1, 1\vec{x}_2...1\vec{x}_{n-1}$ as being pairwise orthogonal, and of the same length, and another observer Y, who perceives the vectors $1\vec{y}_0, 1\vec{y}_1, 1\vec{y}_2...1\vec{y}_{n-1}$ as being pairwise orthogonal, and of the same length, and that Observer X and Y share the same origin.

However, Observer X perceives the Y basis vectors as affine, such that the X observer writes the Y basis vectors in terms of his X basis vectors, which is given by the above matrix of X.

Hence the inverse matrix of X is an affine change of basis, and tell us how the Y observer writes the X basis vectors in terms of his Y basis vectors:

$$X = Y \iff \begin{pmatrix} (1\vec{x}_0) & 0\vec{x}_1 & 0\vec{x}_2 & \cdots & 0\vec{x}_{n-1} \\ 0\vec{x}_0 & (1\vec{x}_1) & 0\vec{x}_2 & \cdots & 0\vec{x}_{n-1} \\ 0\vec{x}_0 & 0\vec{x}_1 & (1\vec{x}_2) & \cdots & 0\vec{x}_{n-1} \\ 0\vec{x}_0 & 0\vec{x}_1 & 0\vec{x}_2 & \ldots(1\vec{x}_m)\ldots & 0\vec{x}_{n-1} \\ 0\vec{x}_0 & 0\vec{x}_1 & 0\vec{x}_2 & \cdots & (1\vec{x}_{n-1}) \end{pmatrix} = \begin{pmatrix} (b_{0,0})\vec{y}_0 & (b_{0,1})\vec{y}_1 & (b_{0,2})\vec{y}_2 & \cdots & (b_{0,n-1})\vec{y}_{n-1} \\ (b_{1,0})\vec{y}_0 & (b_{1,1})\vec{y}_1 & (b_{1,2})\vec{y}_2 & \cdots & (b_{1,n-1})\vec{y}_{n-1} \\ (b_{2,0})\vec{y}_0 & (b_{2,1})\vec{y}_1 & (b_{2,2})\vec{y}_2 & \cdots & (b_{2,n-1})\vec{y}_{n-1} \\ \ldots\vec{y}_0 & \ldots\vec{y}_1 & \ldots\vec{y}_2 & \ldots\vec{y}_m\ldots & \ldots\vec{y}_{n-1} \\ (b_{n-1,0})\vec{y}_0 & (b_{n-1,1})\vec{y}_1 & (b_{n-1,2})\vec{y}_2 & \cdots & (b_{n-1,n-1})\vec{y}_{n-1} \end{pmatrix},$$

where $b_{u,v}$ are the coefficient of X inverse matrix. Hence why LU decomposition (with full pivoting) is the most efficient method of inverting a matrix, because the very process is an affine change of basis via 2-dimensional long division from $X = I_y$ to $X_I = Y$, where Y is the inverse of X (or, if you prefer $A\vec{x} = I\vec{y} \iff I\vec{x} = B\vec{y} \iff B = A^{-1}$).

And herein lies the difference between Multiplication (of single vectors or entire matrices) and Division.

Neither is the inherent inverse of the other (except for the Norm-Division algebras, limited to the Tessarines ($2^n\mathbb{C}$), Quaternions ($4^n\mathbb{H}$) and Octonions by Degen's Eight Square Identity ($8^n\mathbb{O}$)). Division means to measure some set of basis vectors in terms of another affine basis, whereas multiplication is to act upon some data-points in a particular basis by another set of vectors in that same basis.

When you divide, you're simply changing the perspective — nothing in the system itself moves. But when you multiply, you're actively transforming the system, causing things to shift or change.

Hence, a conjugate is an action, such that a vector times it's conjugate is a real number that is not equal to 1, whereas division is a change of reference, such multiplication of a vector and its reciprocal is equal to 1.

However, since this is the reduced publication, we will set aside the physical interpretation of the objects at hand and go straight to the core result in this section.

Definition 10 (Transpose Lemma: The Conjugate Quaternionic Matrix Form)**.** *The conjugate of a quaternion*

$$\vec{x} = a\vec{q} + b\vec{i} + c\vec{j} + d\vec{k}$$

is given by the first column of its adjugate matrix, scaled by the inverse of its norm squared, yielding:

$$\vec{x}^* = a\vec{q} - b\vec{i} - c\vec{j} - d\vec{k}.$$

In its universal matrix representation, this corresponds to the full 4×4 transpose of the matrix form of $\vec{x}$, whereas the right-handed form is only the transpose of the lower 3×3 block.

Furthermore, the conjugate matrices of both the left-handed and right-handed forms are precisely the full 4×4 transposes of their respective forms. Crucially, all four matrices (the left-handed, right-handed, and their conjugates) preserve the same diagonal, corresponding to the same real part.

We scale the first column of the adjugate matrix by $\frac{1}{a^2+b^2+c^2+d^2}$, because the original first column is not the conjugate of $\vec{x}$.

Definition 11 (First Column Lemma: The Reciprocal Quaternionic Matrix Form)**.** *The reciprocal of a quaternion*

$$\vec{x} = a\vec{q} + b\vec{i} + c\vec{j} + d\vec{k}$$

is given by the first column of its inverse matrix, with no scaling, which yields:

$$\vec{x}^{-1} = \frac{1}{a^2 + b^2 + c^2 + d^2}\left(a\vec{q} - b\vec{i} - c\vec{j} - d\vec{k}\right).$$

In its universal matrix representation, this corresponds to the first column of the inverse matrix of $\vec{x}$.

Furthermore, the reciprocal matrices of both the left-handed and right-handed forms are precisely the first columns of their inverse matrices. And that this holds for biquaternions (so long as the matrix entries remain complex numbers!).

The reason we don't scale the first column of the inverse matrix is because the original column of the inverse matrix is already scaled down by $\frac{1}{a^2+b^2+c^2+d^2}$ by default. As for physical interpretation of complex determinants and norms, that's a discussion for another day.

3. Quaternionic Linear Systems of Equations

A quaternion is not a collection of four independent components acting together; rather, it is a single entity — a vector with both magnitude and direction.

As a consequence, a system of linear equations involving two quaternionic variables requires two independent equations to fully determine the solution, and more generally, a system of n variables requires n independent equations to resolve.

3.1. The Power of Quaternionic Least Squares for Market Analysis

This is why quaternionic-valued neural networks outperform their real-valued counterparts in $\mathbb{R}^4$ and bicomplex-valued neural networks in $2^2\mathbb{C}$.

Quaternionic regressions enable the quantification of relationships between variables that were previously treated as independently correlated qualities. For example, instead of performing a real-number regression of four raw material stock prices (e.g., iron, nickel, copper, and aluminum) onto the four most consumed products that utilize them—such as for making an informed futures trade—it is more effective to treat the four raw material prices as a single quaternionic vector and the four consumed goods as another quaternionic vector.

The optimal approach is to compute the quaternionic least squares regression of the vector representing consumed goods in terms of the vector of raw material prices. This method treats the four qualities of metals and the four qualities of consumed goods as a single unified entity, rather than as four separate variables whose correlations one merely hopes to be strong.

I've had people approach me with the argument that an $\mathbb{R}^4$ regression often yields a higher R^2 than a quaternionic regression when applied to market data. My response? So what?

An $\mathbb{R}^4$ regression treats the problem as four independent input variables, performing a separate regression for each individual output variable. Naturally, this leads to higher R^2 values for each of the four consumed goods, but it fails to capture any intrinsic intercorrelation between the four raw materials and the four consumed goods. It's akin to a soothsayer claiming confidence in their prediction simply because they have four shiny crystal balls.

Here's the key point: If a quaternionic least squares regression returns a strong R^2, you can be certain it's a reliable predictor. Unlike an $\mathbb{R}^4$ regression, which treats variables as isolated entities, the quaternionic approach models the data as a unified quaternionic wave state, preserving the inherent relationships that exist within the system.

3.2. Are There Cases Where $\mathbb{R}^4$ Regression Might Still Be Preferred?

Yes, if a strong quaternionic fit is already established, then the corresponding real-numbered regressions are also reliable for all four consumed goods. This allows for investment in all four at once with higher precision.

Moreover, $\mathbb{R}^4$ regression can be leveraged as an efficient screening tool. Given the computational cost of quaternionic regression, one can first test a large number of consumed goods—potentially in the hundreds—using $\mathbb{R}^4$ regression. From this, a ranked list of the top ten highest-R^2 candidates can be compiled.

The next step is to test various subsets of four from this top ten list using quaternionic regression. If any subset of four retains a high R^2 in the quaternionic least squares regression, it indicates that these four goods form a structurally stable investment set. Unlike independent real-number regressions, which may capture statistical correlations without deeper interdependencies, a strong quaternionic fit confirms that the relationships between raw materials and goods are fundamentally linked rather than coincidental.

This hybrid approach balances computational efficiency with structural robustness:

1.   Use $\mathbb{R}^4$ regression to broadly identify top-performing consumed goods.
2.   Apply quaternionic regression to validate the most stable subsets.

3. Invest in goods that demonstrate high predictive power in quaternionic space, ensuring a more reliable and stable portfolio.
4. Since goods exhibit a latency of two to four weeks in responding to fluctuations in raw material prices, the changes in raw material prices effectively serve as a predictive model for future trades in the goods market.

Further extensions of this method could include applying octonionic regression ($8^n \mathbb{O}$) to analyze sets of eight goods or incorporating time-sequencing quaternionic models to track market stability across different economic cycles.

3.3. Solving a Purely Left-Handed System of Quaternionic Linear Equations

Given n quaternionic affine basis vectors $\vec{c}_0, \vec{c}_1, \ldots, \vec{c}_{n-1}$ and n quaternionic Y basis vectors $\vec{y}_0, \vec{y}_1, \ldots, \vec{y}_{n-1}$, solve for the constants $\vec{c}_s$ (affine set of basis vectors) such that:

$$\vec{y}_t = \vec{c}_0 \vec{x}_{t,0} + \vec{c}_1 \vec{x}_{t,1} + \cdots + \vec{c}_{n-1} \vec{x}_{t,n-1}$$

for all $0 \leq t \leq n - 1$, where X is a general $n \times n$ block matrix, with each embedded block being the 4×4 right-handed quaternionic matrix form of $\vec{x}_{t,s}$.

Using a system of three equations, we have:

1. $\vec{y}_0 = \vec{c}_0 \vec{x}_{0,0} + \vec{c}_1 \vec{x}_{0,1} + \vec{c}_2 \vec{x}_{0,2}$
2. $\vec{y}_1 = \vec{c}_0 \vec{x}_{1,0} + \vec{c}_1 \vec{x}_{1,1} + \vec{c}_2 \vec{x}_{1,2}$
3. $\vec{y}_2 = \vec{c}_0 \vec{x}_{2,0} + \vec{c}_1 \vec{x}_{2,1} + \vec{c}_2 \vec{x}_{2,2}$

Which is written as (in condensed 3x3 block matrix form):

$$\begin{pmatrix} [\vec{x}_{0,0}]_{4\mathbb{H}R} & [\vec{x}_{0,1}]_{4\mathbb{H}R} & [\vec{x}_{0,2}]_{4\mathbb{H}R} \\ [\vec{x}_{1,0}]_{4\mathbb{H}R} & [\vec{x}_{1,1}]_{4\mathbb{H}R} & [\vec{x}_{1,2}]_{4\mathbb{H}R} \\ [\vec{x}_{2,0}]_{4\mathbb{H}R} & [\vec{x}_{2,1}]_{4\mathbb{H}R} & [\vec{x}_{2,2}]_{4\mathbb{H}R} \end{pmatrix} \begin{pmatrix} \vec{c}_0 \\ \vec{c}_1 \\ \vec{c}_2 \end{pmatrix} = \begin{pmatrix} \vec{y}_0 \\ \vec{y}_1 \\ \vec{y}_2 \end{pmatrix}$$

Which, it its full 12x12 times 12x1 equals 12x1 form over the reals:

$$\begin{pmatrix} +a_{0,0} & -b_{0,0} & -c_{0,0} & -d_{0,0} & +a_{0,1} & -b_{0,1} & -c_{0,1} & -d_{0,1} & +a_{0,2} & -b_{0,2} & -c_{0,2} & -d_{0,2} \\ +b_{0,0} & +a_{0,0} & -d_{0,0} & +c_{0,0} & +b_{0,1} & +a_{0,1} & -d_{0,1} & +c_{0,1} & +b_{0,2} & +a_{0,2} & -d_{0,2} & +c_{0,2} \\ +c_{0,0} & +d_{0,0} & +a_{0,0} & -b_{0,0} & +c_{0,1} & +d_{0,1} & +a_{0,1} & -b_{0,1} & +c_{0,2} & +d_{0,2} & +a_{0,2} & -b_{0,2} \\ +d_{0,0} & -c_{0,0} & +b_{0,0} & +a_{0,0} & +d_{0,1} & -c_{0,1} & +b_{0,1} & +a_{0,1} & +d_{0,2} & -c_{0,2} & +b_{0,2} & +a_{0,2} \\ +a_{1,0} & -b_{1,0} & -c_{1,0} & -d_{1,0} & +a_{1,1} & -b_{1,1} & -c_{1,1} & -d_{1,1} & +a_{1,2} & -b_{1,2} & -c_{1,2} & -d_{1,2} \\ +b_{1,0} & +a_{1,0} & -d_{1,0} & +c_{1,0} & +b_{1,1} & +a_{1,1} & -d_{1,1} & +c_{1,1} & +b_{1,2} & +a_{1,2} & -d_{1,2} & +c_{1,2} \\ +c_{1,0} & +d_{1,0} & +a_{1,0} & -b_{1,0} & +c_{1,1} & +d_{1,1} & +a_{1,1} & -b_{1,1} & +c_{1,2} & +d_{1,2} & +a_{1,2} & -b_{1,2} \\ +d_{1,0} & -c_{1,0} & +b_{1,0} & +a_{1,0} & +d_{1,1} & -c_{1,1} & +b_{1,1} & +a_{1,1} & +d_{1,2} & -c_{1,2} & +b_{1,2} & +a_{1,2} \\ +a_{2,0} & -b_{2,0} & -c_{2,0} & -d_{2,0} & +a_{2,1} & -b_{2,1} & -c_{2,1} & -d_{2,1} & +a_{2,2} & -b_{2,2} & -c_{2,2} & -d_{2,2} \\ +b_{2,0} & +a_{2,0} & -d_{2,0} & +c_{2,0} & +b_{2,1} & +a_{2,1} & -d_{2,1} & +c_{2,1} & +b_{2,2} & +a_{2,2} & -d_{2,2} & +c_{2,2} \\ +c_{2,0} & +d_{2,0} & +a_{2,0} & -b_{2,0} & +c_{2,1} & +d_{2,1} & +a_{2,1} & -b_{2,1} & +c_{2,2} & +d_{2,2} & +a_{2,2} & -b_{2,2} \\ +d_{2,0} & -c_{2,0} & +b_{2,0} & +a_{2,0} & +d_{2,1} & -c_{2,1} & +b_{2,1} & +a_{2,1} & +d_{2,2} & -c_{2,2} & +b_{2,2} & +a_{2,2} \end{pmatrix} \begin{pmatrix} +c_{0,0} \\ +c_{0,1} \\ +c_{0,2} \\ +c_{0,3} \\ +c_{1,0} \\ +c_{1,1} \\ +c_{1,2} \\ +c_{1,3} \\ +c_{2,0} \\ +c_{2,1} \\ +c_{2,2} \\ +c_{2,3} \end{pmatrix} = \begin{pmatrix} +y_{0,0} \\ +y_{0,1} \\ +y_{0,2} \\ +y_{0,3} \\ +y_{1,0} \\ +y_{1,1} \\ +y_{1,2} \\ +y_{1,3} \\ +y_{2,0} \\ +y_{2,1} \\ +y_{2,2} \\ +y_{2,3} \end{pmatrix}$$

Now let X be the known 12x12 real number matrix, and let Y be the known 12x1 Matrix concatenation of the vectors $\vec{y}_t$ and let C be the **unknown** 12x1 Matrix concatenation of the vectors $\vec{c}_s$, then it follows that:

$$XC = Y \iff X^{-1}Y = C.$$

Q.E.D.

3.4. Solving a Purely Right-Handed System of Quaternionic Linear Equations

This section doesn't require as much introduction since it's fundamentally the same as the previous section, except we use Left-Handed X matrices to resolve the unknown right-handed C basis vectors.

Using a system of three equations, we have:

1. $\vec{y}_0 = \vec{x}_{0,0} \vec{c}_0 + \vec{x}_{0,1} \vec{c}_1 + \vec{x}_{0,2} \vec{c}_2$
2. $\vec{y}_1 = \vec{x}_{1,0} \vec{c}_0 + \vec{x}_{1,1} \vec{c}_1 + \vec{x}_{1,2} \vec{c}_2$
3. $\vec{y}_2 = \vec{x}_{2,0} \vec{c}_0 + \vec{x}_{2,1} \vec{c}_1 + \vec{x}_{2,2} \vec{c}_2$

Which is written as (in condensed 3x3 block matrix form):

$$\begin{pmatrix} [\vec{x}_{0,0}]_{4\mathbb{H}L} & [\vec{x}_{0,1}]_{4\mathbb{H}L} & [\vec{x}_{0,2}]_{4\mathbb{H}L} \\ [\vec{x}_{1,0}]_{4\mathbb{H}L} & [\vec{x}_{1,1}]_{4\mathbb{H}L} & [\vec{x}_{1,2}]_{4\mathbb{H}L} \\ [\vec{x}_{2,0}]_{4\mathbb{H}L} & [\vec{x}_{2,1}]_{4\mathbb{H}L} & [\vec{x}_{2,2}]_{4\mathbb{H}L} \end{pmatrix} \begin{pmatrix} \vec{c}_0 \\ \vec{c}_1 \\ \vec{c}_2 \end{pmatrix} = \begin{pmatrix} \vec{y}_0 \\ \vec{y}_1 \\ \vec{y}_2 \end{pmatrix}$$

3.5. Solving a Purely Middle-Handed System of Quaternionic Linear Equations

The key to this problem is realizing that a left matrix times a right matrix results in a single matrix. Using a system of three equations, we have:

1. $\vec{z}_0 = \vec{x}_{0,0}\vec{c}_0\vec{y}_{0,0} + \vec{x}_{0,1}\vec{c}_1\vec{y}_{0,1} + \vec{x}_{0,2}\vec{c}_2\vec{y}_{0,2}$
2. $\vec{z}_1 = \vec{x}_{1,0}\vec{c}_0\vec{y}_{1,0} + \vec{x}_{1,1}\vec{c}_1\vec{y}_{1,1} + \vec{x}_{1,2}\vec{c}_2\vec{y}_{1,2}$
3. $\vec{z}_2 = \vec{x}_{2,0}\vec{c}_0\vec{y}_{2,0} + \vec{x}_{2,1}\vec{c}_1\vec{y}_{2,1} + \vec{x}_{2,2}\vec{c}_2\vec{y}_{2,2}$

Which is written as (in condensed 3x3 block matrix form):

$$\begin{pmatrix} [\vec{x}_{0,0}]_{4\mathbb{H}L}[\vec{y}_{0,0}]_{4\mathbb{H}R} & [\vec{x}_{0,1}]_{4\mathbb{H}L}[\vec{y}_{0,1}]_{4\mathbb{H}R} & [\vec{x}_{0,2}]_{4\mathbb{H}L}[\vec{y}_{0,2}]_{4\mathbb{H}R} \\ [\vec{x}_{1,0}]_{4\mathbb{H}L}[\vec{y}_{1,0}]_{4\mathbb{H}R} & [\vec{x}_{1,1}]_{4\mathbb{H}L}[\vec{y}_{1,1}]_{4\mathbb{H}R} & [\vec{x}_{1,2}]_{4\mathbb{H}L}[\vec{y}_{1,2}]_{4\mathbb{H}R} \\ [\vec{x}_{2,0}]_{4\mathbb{H}L}[\vec{y}_{2,0}]_{4\mathbb{H}R} & [\vec{x}_{2,1}]_{4\mathbb{H}L}[\vec{y}_{2,1}]_{4\mathbb{H}R} & [\vec{x}_{2,2}]_{4\mathbb{H}L}[\vec{x}_{2,2}]_{4\mathbb{H}R} \end{pmatrix} \begin{pmatrix} \vec{c}_0 \\ \vec{c}_1 \\ \vec{c}_2 \end{pmatrix} = \begin{pmatrix} \vec{z}_0 \\ \vec{z}_1 \\ \vec{z}_2 \end{pmatrix}$$

Now let the 4x4 matrix $G_{s,t} = [\vec{x}_{s,t}]_{4\mathbb{H}L}[\vec{y}_{s,t}]_{4\mathbb{H}R}$.

$$\begin{pmatrix} G_{0,0} & G_{0,1} & G_{0,2} \\ G_{1,0} & G_{1,1} & G_{1,2} \\ G_{2,0} & G_{2,1} & G_{2,2} \end{pmatrix} \begin{pmatrix} \vec{c}_0 \\ \vec{c}_1 \\ \vec{c}_2 \end{pmatrix} = \begin{pmatrix} \vec{z}_0 \\ \vec{z}_1 \\ \vec{z}_2 \end{pmatrix}$$

Now let H be the inverse of the 12x12 real number matrix embedded in the $G_{s,t}$ blocks.

$$H \begin{pmatrix} \vec{z}_0 \\ \vec{z}_1 \\ \vec{z}_2 \end{pmatrix} = \begin{pmatrix} \vec{c}_0 \\ \vec{c}_1 \\ \vec{c}_2 \end{pmatrix}$$

Q.E.D.

3.6. Solving System of Quaternionic Linear Equations of Mixed Chirality

The key to this problem is realizing that we're ultimately dealing with a 12x12 real number matrix, such that the chirality of any block has no bearing on the problem at hend.

1. $\vec{z}_0 = \vec{x}_{0,0}\vec{c}_0 + \vec{x}_{0,1}\vec{c}_1\vec{y}_{0,1} + \vec{c}_2\vec{y}_{0,2}$
2. $\vec{z}_1 = \vec{x}_{1,0}\vec{c}_0 + \vec{x}_{1,1}\vec{c}_1\vec{y}_{1,1} + \vec{c}_2\vec{y}_{1,2}$
3. $\vec{z}_2 = \vec{x}_{2,0}\vec{c}_0 + \vec{x}_{2,1}\vec{c}_1\vec{y}_{2,1} + \vec{c}_2\vec{y}_{2,2}$

Which is written as (in condensed 3x3 block matrix form):

$$\begin{pmatrix} [\vec{x}_{0,0}]_{4\mathbb{H}L} & [\vec{x}_{0,1}]_{4\mathbb{H}L}[\vec{y}_{0,1}]_{4\mathbb{H}R} & [\vec{y}_{0,2}]_{4\mathbb{H}R} \\ [\vec{x}_{1,0}]_{4\mathbb{H}L} & [\vec{x}_{1,1}]_{4\mathbb{H}L}[\vec{y}_{1,1}]_{4\mathbb{H}R} & [\vec{y}_{1,2}]_{4\mathbb{H}R} \\ [\vec{x}_{2,0}]_{4\mathbb{H}L} & [\vec{x}_{2,1}]_{4\mathbb{H}L}[\vec{y}_{2,1}]_{4\mathbb{H}R} & [\vec{x}_{2,2}]_{4\mathbb{H}R} \end{pmatrix} \begin{pmatrix} \vec{c}_0 \\ \vec{c}_1 \\ \vec{c}_2 \end{pmatrix} = \begin{pmatrix} \vec{z}_0 \\ \vec{z}_1 \\ \vec{z}_2 \end{pmatrix}$$

Now let the 4x4 matrix $G_{s,1} = [\vec{x}_{s,1}]_{4\mathbb{H}L}[\vec{y}_{s,1}]_{4\mathbb{H}R}$.

$$\begin{pmatrix} [\vec{x}_{0,0}]_{4\mathbb{H}L} & G_{0,1} & [\vec{y}_{0,2}]_{4\mathbb{H}R} \\ [\vec{x}_{1,0}]_{4\mathbb{H}L} & G_{1,1} & [\vec{y}_{1,2}]_{4\mathbb{H}R} \\ [\vec{x}_{2,0}]_{4\mathbb{H}L} & G_{2,1} & [\vec{y}_{2,2}]_{4\mathbb{H}R} \end{pmatrix} \begin{pmatrix} \vec{c}_0 \\ \vec{c}_1 \\ \vec{c}_2 \end{pmatrix} = \begin{pmatrix} \vec{z}_0 \\ \vec{z}_1 \\ \vec{z}_2 \end{pmatrix}$$

Now let H be the inverse of the 12x12 real number matrix embedded in the $G_{s,1}$ blocks (and yes, it's actually that easy!):

$$H \begin{pmatrix} \vec{z}_0 \\ \vec{z}_1 \\ \vec{z}_2 \end{pmatrix} = \begin{pmatrix} \vec{c}_0 \\ \vec{c}_1 \\ \vec{c}_2 \end{pmatrix}$$

Q.E.D.

3.7. Invertibility of the X Matrix and Construction of H

The core reason why the system admits a unique solution is the general invertibility of the X matrix. The structure of X is inherently a 12×12 real-valued matrix despite being composed of quaternionic sub-blocks. To see why invertibility is ensured in general cases, we observe the block decomposition:

$$X = \begin{pmatrix} [\vec{x}_{0,0}]_{4\mathbb{H}L} & G_{0,1} & [\vec{y}_{0,2}]_{4\mathbb{H}R} \\ [\vec{x}_{1,0}]_{4\mathbb{H}L} & G_{1,1} & [\vec{y}_{1,2}]_{4\mathbb{H}R} \\ [\vec{x}_{2,0}]_{4\mathbb{H}L} & G_{2,1} & [\vec{y}_{2,2}]_{4\mathbb{H}R} \end{pmatrix}, \tag{1}$$

where each block is a 4×4 real-valued matrix, ensuring that the full system embeds naturally into $\mathbb{R}^{12 \times 12}$. The only theoretical cases in which X would be singular require that its rows be linearly dependent in the real-valued sense, which does not occur under general conditions.

By standard results in linear algebra, the existence of an inverse for X follows whenever $\det(X) \neq 0$, which holds unless an explicitly contrived dependency is introduced. Since quaternionic matrices in this formulation do not introduce extra degeneracy, we conclude that X is generically invertible.

Thus, the inverse matrix H exists such that:

$$HX = I_{12}, \tag{2}$$

which allows us to explicitly solve for the unknown vector:

$$H \begin{pmatrix} \vec{z}_0 \\ \vec{z}_1 \\ \vec{z}_2 \end{pmatrix} = \begin{pmatrix} \vec{c}_0 \\ \vec{c}_1 \\ \vec{c}_2 \end{pmatrix}. \tag{3}$$

This confirms that the quaternionic system of mixed chirality reduces cleanly to a uniquely solvable linear system.

We are now ready to proceed to Quaternionic Least Squares Regression (well, after the following philosophical lecture!).

4. What is the Physical Interpretation of Least Squares Regression over the Reals and the Quaternions?

Although I had promised not to dwell on the philosophical and physical interpretations of hypercomplex numbers, angles, matrices, and logarithms, this paper was ultimately distilled to deliver the closed-form solution to Quaternionic Least Squares. However, unless you understand its physical meaning, you'll never truly use it — because you won't grasp what it is.

We'll begin with the real numbers because, if I'm being honest, I am fairly certain you do not fully understand their physical interpretation either!

4.1. Multiple Lists of Real Numbers All Exist in the Same Dimension

Yes, the title says it all. Suppose you present me with a list such that

$$\forall t: \quad y_t = x_t^2$$

and then proceed to draw a parabola in 2D, declaring that I am mistaken because y_t exists in a separate dimension from x_t.

I counter your argument by stating: No, what you have actually presented is the radius vector function

$$\vec{R}_t = t\vec{q} + t^2\vec{i}$$

which resides within the complex numbers.

I can even prove it to you using the derivative:

$$\vec{R}'_t = 1\vec{q} + 2t\vec{i}$$

which describes both the immediate direction and speed (velocity) of the projectile at any moment in time, t, whereas your derivative, $y' = 2x$, only tells me the lateral velocity component.

Yet, despite this, I still encounter people who continue to argue against this, employing one straw man after another. At this point, I have no choice but to throw down the gauntlet.

Rulers and Measurement in a Single Dimension: Consider a ruler of length 4 and a ruler of length 8.

You give me your single ruler of length 6. I measure it as:

1.   One-half the sum of either ruler.
2.   One-fourth the sum of the first ruler plus five-eighths the sum of the second.
3.   Or any one of the infinite linear combinations of 4 and 8 that yield 6, all of which exist in the same forward direction.

In fact, I only require one of my rulers to measure against yours to get unique solutions. Your ruler of length 6 is either 1.5 times my ruler of length 4 or 0.75 times my ruler of length 8.

Now, suppose you present to me a single ruler of length 10 and use it to measure two data points:

$$y_1 = 5, \quad y_2 = 8$$

which could also be written as

$$y_1 = 50, \quad y_2 = 80$$

if you define 10 as your unit length.

I, however, have my own two rulers of **unequal lengths**: one of length 3 and the other of length 8. There still exists an infinite number of solutions satisfying:

1. $3x_{0,0} + 8x_{0,1} = 50$
2. $3x_{1,0} + 8x_{1,1} = 80$

In fact, we can extend this list indefinitely, and there will still be an infinite number of solutions for each equation because there is no requirement that $x_{j,0} = x_{k,0}$ or that $x_{j,1} = x_{k,1}$:

1. $3x_{0,0} + 8x_{0,1} = y_1$
2. $3x_{1,0} + 8x_{1,1} = y_2$
3. $\vdots$
4. $3x_{n,0} + 8x_{n,1} = y_n$

Now let's re-frame the problem. You present me with a list of n measurements and give me some arbitrary X matrix. You then demand that I determine the unique set of n rulers (which are fixed and cannot change) that, when multiplied against the X matrix, yield the n measurements of y.

This is straightforward: Simply multiply the inverse of X against the column vector of y.

But now you change the problem. You demand that I find a set of m rulers with respect to these arbitrary X coefficients that solve for these n measurements, with the condition that $m > n$. Since we now have more rulers than there are data points in y, an infinite set of solutions exists.

Finally, you impose one last constraint: Find a set of m rulers with respect to these arbitrary X coefficients that solve for these n measurements, with $m < n$. Now, we have fewer rulers than there are data points in y, meaning no solution exists—assuming all y measurements are pairwise distinct.

And yet, throughout all of this, we have remained in a single dimension: The forward direction of $\vec{q}$. Regardless of how many n measurements of y were listed, we never once had to step into two or more dimensions.

Introducing a Second Dimension: Now, suppose you introduce a measurement that exists 2 units above one of the y data points,

$$y_z \vec{q} + 2\vec{i}.$$

Since the X matrix is real, its inverse, X^{-1}, must also be a real-number matrix. However, the Y column vector of measurements now contains at least one complex number. This means that at least one of our rulers must also be a complex number — lo!

Thus, scalars are not merely numbers with magnitude; they are numbers with direction. Direction is inherently scalar.

We'll return shortly to the case of $m < n$ and how that leads to least squares regression, but for the moment, let's discuss what it means for a scalar, or a determinant of matrix, or the magnitude of a vector, to have both length and direction.

To demonstrate this, we shall introduce the most basic form of Tessarine (other than an ordinary complex number) in the $2^2\mathbb{C}$ commutative logic, which is the same logic used to generate the Hyperbolic Rotations of Spacetime Intervals.

The multiplication rules are straightforward (each list is the successive Column of the Cayley Table):

1. Multiplication by the First Reflector (The Forward Vector):

    (a) $\vec{q}\vec{q} = +\vec{q}$
    (b) $\vec{w}\vec{q} = +\vec{w}$
    (c) $\vec{j}\vec{q} = +\vec{j}$
    (d) $\vec{\lambda}\vec{q} = +\vec{\lambda}$

2. Multiplication by the First Rotator:

    (a) $\vec{q}\vec{w} = +\vec{w}$
    (b) $\vec{w}\vec{w} = -\vec{q}$
    (c) $\vec{j}\vec{w} = +\vec{\lambda}$
    (d) $\vec{\lambda}\vec{w} = -\vec{j}$

3. Multiplication by the Second Reflector:

    (a) $\vec{q}\vec{j} = +\vec{j}$
    (b) $\vec{w}\vec{j} = +\vec{\lambda}$
    (c) $\vec{j}\vec{j} = +\vec{q}$
    (d) $\vec{\lambda}\vec{j} = +\vec{w}$

4. Multiplication by the Second Rotator:

    (a) $\vec{q}\vec{\lambda} = +\vec{\lambda}$
    (b) $\vec{w}\vec{\lambda} = -\vec{j}$
    (c) $\vec{j}\vec{\lambda} = +\vec{w}$
    (d) $\vec{\lambda}\vec{\lambda} = -\vec{q}$

5. Exponent Rules:

    (a) $e^{\vec{q}\theta} = \vec{q}\cosh\theta + \vec{q}\sinh\theta$
    (b) $e^{\vec{w}\theta} = \vec{q}\cosh\vec{w}\theta + \vec{q}\sinh\vec{w}\theta = \vec{q}\cos\theta + \vec{w}\sin\theta$
    (c) $e^{\vec{j}\theta} = \vec{q}\cosh\theta + \vec{j}\sinh\theta$
    (d) $e^{\vec{\lambda}\theta} = \vec{q}\cosh\vec{\lambda}\theta + \vec{q}\sinh\vec{\lambda}\theta = \vec{q}\cos\theta + \vec{\lambda}\sin\theta$

6. Vector Form: $\vec{z} = z_0\vec{q} + z_1\vec{w} + z_2\vec{j} + z_3\vec{\lambda} = z_0\vec{q} + z_1\vec{w} + \vec{j}(z_2\vec{q} + z_3\vec{w})$

7. Trigonometric Form: $\vec{z} = (z_0\vec{q} + z_1\vec{w})\operatorname{sech}\vec{\theta}\left(\vec{q}\cosh\vec{\theta} + \vec{j}\sinh\vec{\theta}\right)$

    with $\vec{\theta} = \operatorname{arctanh}\frac{z_0\vec{q} + z_1\vec{w}}{z_2\vec{q} + z_3\vec{w}}$

8.   Exponential Form: $\vec{z} = e^{\vec{u}} e^{\vec{j}\theta}$, where $e^{\vec{u}} = e^{\ln\left(z_0 \operatorname{sech}\vec{\theta}\right)}$.

9.   The "Magnitude" of $\vec{z}$ is the complex number $z_0 \operatorname{sech}\vec{\theta}$, which is a complex number **possessing direction**. The complex determinant of the 2x2 complex matrix form is precisely $(z_0\vec{q} + \vec{z_1}\vec{w})^2 \operatorname{sech}^2\vec{\theta}$. The determinant of the 4x4 real number matrix form is the square of the traditional magnitude of the complex number $(z_0\vec{q} + \vec{z_1}\vec{w})^2 \operatorname{sech}^2\vec{\theta}$.

4.2. Was the Previous Discussion about $2^2\mathbb{C}$ Tessarine Logic a Waste of Time for This Article?

No. Only to those who do not understand what hypercomplex numbers are and how all hypercomplex logics reduce to a locally bicomplex number in isolation.

An ordinary quaternion exists within the locally complex plane of $\vec{q} + \vec{i}_2$, such that there always exists a unique change of orthogonal basis—a rigid body rotation of the imaginary axes—that maps $\vec{i}_1, \vec{j}_1, \vec{k}_1$ to $\vec{i}_2, \vec{j}_2, \vec{k}_2$, satisfying:

$$\vec{z} = z_0\vec{q} + z_1\vec{i}_1 + z_2\vec{j}_1 + z_3\vec{k}_1 = z_0\vec{q} + M\vec{i}_2 + 0\vec{j}_2 + 0\vec{k}_2.$$

After this transformation, one is free to further rotate $\vec{j}_2$ and $\vec{k}_2$ within their own plane to

$$\vec{j}_3 = \vec{j}_2 \cos\theta + \vec{k}_2 \sin\theta, \quad \vec{k}_3 = -\vec{j}_2 \sin\theta + \vec{k}_2 \cos\theta,$$

if necessary to simplify additional problems, such as solving a quadratic equation over the biquaternions.

The magnitude of the quaternion is given by $\sqrt{z_0^2 + M^2}$. The square of this magnitude is the determinant of its 2×2 matrix form, and the square of that determinant corresponds to the 4×4 determinant of $\vec{z}$ in its original imaginary basis vectors.

Furthermore, in the reduced basis, the quaternion can be written as:

$$\vec{z} = N\left(\vec{q}\cos\theta + \vec{i}_2 \sin\theta\right),$$

where $\theta = \arctan\frac{M}{z_0}$ and $N = z_0 \sec\theta$, which can be a positive or negative real. This formulation avoids the cumbersome four-quadrant arctangent function. With the aid of $\operatorname{sech}\theta$ and $\sec\theta$, it allows us to bypass higher-dimensional convolutions of the four-quadrant arctangent function.

Similarly, a **Proper Biquaternion**, whose imaginary components square to $+1$ and whose coefficients are complex numbers, exists as a locally bicomplex number. Here, M and $\vec{z}_0$ are complex-valued magnitudes:

$$\vec{z} = \vec{z}_0\vec{q} + \vec{z}_1\vec{i}_1 + \vec{z}_2\vec{j}_1 + \vec{z}_3\vec{k}_1 = \vec{z}_0\vec{q} + \vec{M}\vec{i}_2 + 0\vec{j}_2 + 0\vec{k}_2.$$

Its determinant, when represented as a 2×2 Reflector Matrix of the form:

$$\begin{pmatrix} +\vec{z} & +\vec{M} \\ +\vec{M} & +\vec{z} \end{pmatrix},$$

is given by $\vec{z}^2 - \vec{M}^2$, which is the square of $\sqrt{\vec{z}^2 - \vec{M}^2}$. The determinant of this 2×2 matrix, when squared, gives the complex determinant of the 4×4 matrix representation of $\vec{z}$. The real magnitude of this determinant, when squared again, provides the real determinant of the 8×8 real matrix form of $\vec{z}$ in the original basis.

From this, we learn that determinants are not simply the volume of n-dimensional parallelepipeds—a common misconception. Rather, they represent the magnitude of the expressed hypercomplex vector, raised to the power n, where n is the total number of dimensions used to describe the hypercomplex structure.

4.2.1. The Physical Interpretation of the Determinant of a Real Number Matrix

In our previous example, all rulers existed in the same direction, meaning the real number matrix did not form an n-dimensional parallelepiped. Instead, the determinant of a real matrix represents the magnitude of a Hypocomplex Number raised to the power of n. For real numbers, a Hypocomplex Number is the vector sum of terms aligned in the same direction or, more generally, at least two or more terms sharing a common axis.

For hypercomplex numbers, the concept extends: A Hypocomplex Number is the vector sum of terms existing within the same subspace. To solve for a system of n hypocomplex variables in an m-dimensional hypercomplex space, one requires m hypocomplex rulers, also in m-dimensional hypercomplex space.

If the number of rulers does not match the number of variables, n, the system is not invertible, even if the rulers share the set of hypercomplex components as the y measurements.

Furthermore, if the hypercomplex components of the rulers (m) are neither an improper superset nor a proper superset of the hypercomplex components of the n measurements, then the system remains non-invertible — unless matrix multiplication is allowed to expand the set of ruler components (which is the normal course of action). This is analogous to our earlier example, where we added $2\vec{i}$ to one of the y-inputs. That is, if rulers are restricted to m hypercomplex components, then even if their count, m, matches the number of measurements, n, the system remains unsolvable if the rulers' m hypercomplex components form a proper subset of those present in the y measurements. Consider the two real numbers, 5 and 10, placed in reflector matrix form:

$$\begin{pmatrix} +5 & +10 \\ +10 & +5 \end{pmatrix}.$$

We compute the determinant using the following steps (we call it a Reflector Matrix because the first row is reflection of the second row over the 45 degree line, never again shall you call it a Hyperbolic Rotation Matrix!):

1. Express the sum in hyperbolic form:

$$5\vec{q} + 10\vec{q} = e^u e^{\vec{q}\theta} = e^u(\vec{q}\cosh\theta + \vec{q}\sinh\theta).$$

2. Compute θ as: $\vec{\theta} = \operatorname{arctanh}\frac{10}{5} = 0.549306\vec{q} - \frac{\pi}{2}\vec{i}.$
3. Rewrite using hyperbolic secant:

$$5\vec{q} + 10\vec{q} = 5\operatorname{sech}\vec{\theta}(\vec{q}\cosh\vec{\theta} + \vec{q}\sinh\vec{\theta}).$$

4. Recognizing that the determinant of the unit reflector matrix

$$\begin{pmatrix} +\cosh\theta & +\sinh\theta \\ +\sinh\theta & +\cosh\theta \end{pmatrix}$$

is 1, we conclude that the determinant of

$$\begin{pmatrix} +5 & +10 \\ +10 & +5 \end{pmatrix}$$

must be

$$(e^u)^2 = e^{2u} = \ln\left(5\operatorname{sech}\vec{\theta}\right).$$

Verifying the Calculation:

- The determinant of the original matrix is $5^2 - 10^2 = -75$. The square root of 75 is approximately 8.66025.

- Compute θ:

$$\theta = \operatorname{arctanh} \frac{10}{5} = 0.549306\vec{q} - \frac{\pi}{2}\vec{i}.$$

- Compute $\operatorname{sech}\theta$: $\operatorname{sech}\theta = \vec{i}\sqrt{3}$.
- Compute $5\operatorname{sech}\theta$: $5\operatorname{sech}\theta = 5\vec{i}\sqrt{3}$.
- Squaring this value:

$$(5\vec{i}\sqrt{3})^2 = -75.$$

This demonstrates that the determinant captures the structure of hypocomplex numbers — vectors aligned in the same direction.

For real number matrices, determinants represent uni-directional magnitudes, measuring the geometric mean reflective power of the matrix (i.e., whether it is a unit reflector matrix or not), rather than geometric areas or n-dimensional volumes.

The true parallelepiped interpretation of the determinant only comes into play when the matrix consists of something other than real numbers.

Case in point, let us return to the original question of this section: $y = x^2$ is not a parabola. It is a straight-line Radius Vector whose relative magnitude is equal to the relative magnitude of time, from the perspective of an observer who redefines the "unit length of time" as the current time. What!?

Let us replace x with t, then $\vec{y} = t^2\vec{q};\quad t \in \mathbb{R}$. This means that whatever time it is, we declare that as our new forward vector, then scale it by the current time. By definition, the length of that vector is t^2, and it remains in the same direction as time.

Now consider the traditional parabola, $\vec{R}_t = \vec{q}t + \vec{i}t^2$. What is the area under it from $t = 0$ to $t = 10$? Is it not:

$$\frac{1}{2}t^2 + \frac{1}{3}\vec{i}t^3 = 50\vec{q} + 333.33\vec{i}?$$

Suppose, for the sake of argument, that $\vec{R}_t$ represents velocity over time. Is the value $50\vec{q} + 333.33\vec{i}$ not the total change in the ball's displacement from $t = 0$ to $t = 10$?

You see, area — being two-dimensional — only makes sense in complex space. Volume, in turn, only makes sense in hypercomplex space. Thus, the determinant of a real number matrix cannot represent the volume of a parallelepiped; rather, it is simply the exponentiated logarithmic mean of some hypocomplex real number sum, masquerading as a tessarine in matrix form.

After some reviews, I still have people insisting that $y = x^2$ is a parabola when $x, y \in \mathbb{R}$. For the sake of God, it's time to throw down the gauntlet. What does $x = \sqrt{y}$ mean? In other words, let's examine this question through the inverse function.

It means that one travels from $\vec{q}$ to $\vec{P}_1 = \vec{x}$, then refines $\vec{x}$ as the new $\vec{q}$ and redefines the original $\vec{q}$ as $\frac{1}{\vec{x}}$ (i.e., the system is divided by $\vec{x}$ to establish a new frame of reference). Then, in this new frame, they travel to $\vec{P}_2 = \vec{x}$ again, which is equivalent to $\vec{x}^2$ in the new frame.

This is precisely why $\vec{i}^2 = -1$: you travel to $\vec{i}$, declare it as the new forward direction, and then travel to $\vec{i}$ again, which in the original frame corresponds to $-1\vec{q}$.

Thus, the definition of a square root is as follows:

"A relative action, which when repeated twice, brings the Specimen to $\vec{y}$ in the Observer's Reference Frame."

Similarly, the definition of a square is:

"A relative action, repeated twice."

From this, it becomes obvious that exponents, x^n, have nothing to do with squares, cubes, or volumes. Instead, they simply represent repeated actions (if n is a positive integer), repeated inverse actions (if n is a negative integer), roots (if n is the reciprocal of an integer, where the root means "find the n repeated action that produces this value"), or a combination of repeated actions and roots if n is a rational number. More generally, in the complex case, exponents represent a logarithmic scaling of n,

which—if n is complex—also implies a change in direction. More broadly, it corresponds to a shift in the equiangular spiral locus upon which the data point resides, a topic to be published in the near future.

So, the next time a junior high school student asks, "What is the square root of negative one?" tell them the truth:

"Two consecutive left turns, or two consecutive right turns."

And when they ask, "Then what's the square root of $\vec{i}$?" tell them it's two consecutive half-left turns or two consecutive three-half right turns. Better yet, demonstrate it physically — use your body and feet. Teach them this while they are young, and we won't need this much exposition in a so-called "reduced version" of a paper.

4.2.2. Practical use of Quaternionic Least Squares

I assume you have now agreed that $\vec{R}_t = t\vec{q} + t^2\vec{i}$ is the path of a thrown ball (with no wind). Now imagine you're piloting an airplane that experiences an engine failure in the middle of a hurricane. While gravity remains constant, the wind speed is a continuously fluctuating vector. Halfway through the descent, you've gathered enough data points—assuming you have a high-resolution clock—to perform a fast cubic quaternion regression on the trajectory.

This allows you to confidently predict where the plane will crash, despite the erratic and unpredictable fluctuations in wind speed and direction. Now, try to make the same prediction using only real number inputs. Good luck with that!

With quaternionic least squares, you don't need to account for wind speed (or even gravity!). The observed trajectory during a portion of the fall already encapsulates the "average effect" of the wind and gravity combined. All you need to do is analyze the existing data, and it will tell you where the plane is going to land at time t_{future} — whether you're on Earth with a mild breeze or on titanic gas giant with crushing gravity and massive wind speeds of 500km/h. Care to navigate Jupiter's Red Spot with me? I hear Elon Musk is planning a manned mission there!

If you were doing this with real numbers, you would actually have to account for both gravity and wind speed. A wind that changes in both direction and magnitude as a continuous and random function of time! You see, if $\mathbb{R}^n$ was actually multi-dimensional, then you wouldn't have a problem modeling the crash with real numbers, would you?

And this why you will fail if you try to predict four separate commodity prices in the future from four separate raw material prices in the present, because all such proceedings are strictly in the forward direction, you're using four rulers to predict one price; whereas in the quaternionic case we're using one ruler with four literal dimensions to predict one vector price.

This is why hypercomplex-valued neural networks (particular the quaternion valued variety) out perform the real and complex valued neural networks, because their rulers can change direction to quantify that which was originally thought unquantifiable...**qualitative differences**...now represented as different linearly independent directions in some affine basis.

Thus, quaternionic least squares is not limited to quantum theory or abstracted market data. In fact, it's often the simplest method for solving problems in a 3D+Time world, as long as you know how to perform the regression.

4.2.3. Rethinking Linear Independence

Linear independence is traditionally misunderstood. It is not about vectors existing in different directions. Instead, it should be reconsidered as **Ruler Independence**. A non-invertible matrix lacks sufficient rulers ($m < n$) or has an excess of rulers ($m > n$), requiring some to be set to zero to satisfy $m_2 = n$.

4.2.4. It seems like you're wasting more time explaining why this isn't a waste of time—let's just move on to Quaternionic Least Squares!

I assure you, everything stated in this section was not a waste of time. Quaternionic Least Squares requires multiplying the input data points by the conjugates of the data points.

Biquaternionic Least Squares is no different, but now the conjugate matrix has a complex number determinant in the form of

$$e^{\vec{u}} = z_0^2 \operatorname{sech}^2 \theta$$

and if you don't know this, then you cannot obtain the conjugate matrix, and by proxy, the conjugate vector of the data point.

More generally:

1. Given a quaternion, its reduced imaginary basis is:

$$\vec{z} = (z_{0,0}\vec{q} + \vec{z}_{01}\vec{w})\vec{q} + \vec{M}\vec{i_2} + 0\vec{j_2} + 0\vec{k_2} = \vec{z_0}\vec{q} + \vec{M}\vec{i_2} + 0\vec{j_2} + 0\vec{k_2} = e^{\vec{u}}e^{\vec{i_2}\theta}$$

   with:

   (a) $\vec{\theta} = \operatorname{arctanh} \frac{\vec{M}}{\vec{z_0}}$
   (b) $e^{\vec{u}} = \vec{z_0} \operatorname{sech} \vec{\theta}$

2. The conjugate of the quaternion, $\vec{z}$, is of the form:

$$e^{+\vec{u}}e^{-\vec{i_2}\theta}$$

   which always exists, even if either $\vec{u}$ or $\vec{v}$ (or both) are equal $\pm\infty\vec{d}$, where $\vec{d}$ is some finite unit vector. However, the zero vector still retains its complex direction, $e^{\vec{i\theta}}$, which is important for calculus (for instance, finding the derivative of the logarithmic spiral
   $\vec{R_t} = e^{t(\vec{q}\cos\phi + \vec{i}\sin\phi)} = e^{te^{\vec{i}\phi}}$, for some constant real ϕ).

3. The reciprocal of the quaternion, $\vec{z}$, is of the form:

$$e^{-\vec{u}}e^{-\vec{i_2}\theta}$$

   which is not guaranteed to exist, but still retains its direction, even though if it extends infinitely. This is still important for calculus. The appearance of infinity with direction is actually quite common once you dive into the logarithmic spirals of complex numbers, tessarines, and quaternions by inverting them through the inverse hyperbolic tangent or cotangent functions! Especially if one needs to invoke L'Hôpital's rule to remove competing zeros and/or infinities.

4. In the event that

$$\vec{z} = a\vec{q} + b\vec{w}\vec{i_1} + c\vec{w}j_1 + d\vec{w}\vec{k_1}$$

   prior to a basis reduction, then it is the ordinary quaternion with which you are familiar, and the conjugate and reciprocal simplify to the known constructs:

$$\left(a\vec{q} - b\vec{w}\vec{i_1} - c\vec{w}j_1 - d\vec{w}\vec{k_1}\right)$$

   and

$$\frac{1}{a^2 + b^2 + c^2 + d^2}\left(a\vec{q} - b\vec{w}\vec{i_1} - c\vec{w}j_1 - d\vec{w}\vec{k_1}\right)$$

   respectively.

4.3. There's no way that $u, v = \pm\infty$ can yield a valid conjugate!

Ok, so here's a tessarine $c\vec{q} + c\vec{j}$. There's nothing that stops us from multiplying by this tessarine. The determinant of its matrix form $C = \begin{pmatrix} c & c \\ c & c \end{pmatrix}$ is zero ($u = -\infty$, such that $e^u=0$). The direction, $\vec{\theta} = \text{arctanh} \frac{4}{4} = +\infty$.

Yet this tessarine has a conjugate $c\vec{q} - c\vec{j}$, and it also has an adjugate matrix: $C^* = \begin{pmatrix} c & -c \\ -c & c \end{pmatrix}$.

This matrix fulfills the definition that

$$CC^* = (Det_C)I = \begin{pmatrix} Det_C & 0 \\ 0 & Det_C \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

It also so happens that $\left(c\vec{q} + c\vec{j}\right)\left(c\vec{q} - c\vec{j}\right) = c^2\vec{q} + c^2\vec{j} - c^2\vec{j} - c^2\vec{q} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

Which means that when such a horror comes up in your least squares regression, that all we did was add a zero matrix to the existing sum matrices for the vector input that triggered it, and still multiped by something, $C^* = \begin{pmatrix} c & -c \\ -c & c \end{pmatrix}$, against the remaining input vectors for that data point, before lumping it into the design matrix sum.

Yes, if an overwhelming portion of your data points consist of what James Cockle (mid nineteenth century) called 'impossible numbers' — values with near zero magnitude but near infinite direction — you would indeed end up with very **ill-conditioned** error matrix, because the inverse of the block would be something tending towards infinity in all entries (and if all of the data-points corresponding to that vector input were impossible numbers, you'd have a totally non-invertible matrix). However, this would require a highly contrived dataset to begin with.

Not to mention that this isn't even a possibility for ordinary quaternions, no matter how contrived the dataset is, because ordinary quaternions always have reciprocals and conjugates (all of the coefficients before their imaginary reflectors are pure rotators)!

4.3.1. The Digital Definition of Ill-Conditioned Matrices

Forget everything you know about the standard academic definition of ill-conditioned matrices — throw it in the trash.

In this publication, a matrix is considered ill-conditioned if the mantissa lengths of each entry in the inverse matrix (prior to the bit-shift) are less than half the total length (meaning more than one-half of the digits are leading zeros with no nonzero digit in front). This results in a catastrophic loss of floating-point precision when the inverse matrix is multiplied by the response vector column.

To test for an ill-conditioned matrix, perform an LU decomposition with **full pivoting** and multiply the resulting factors by the original matrix. Then, compare each entry of the product with the corresponding entry in the identity matrix. If the error in any part of the matrix exceeds $r^{-\frac{1}{2}n}$, where r is the radix (typically base 2) and n is the number of mantissa digits (32-bit or 64-bit in most applications), we declare the matrix ill-conditioned.

Of course, the threshold of $\frac{1}{2}n$ can be adjusted based on the application — air conditioners require minimal precision, whereas long-term cosmic orbit calculations demand extreme accuracy. However, this is the definition we adopt in this publication.

Overall, we do not explicitly measure the leading or trailing zeros before or after the bit shift. Instead, we analyze the matrix product XX^{-1}. If the ones on the diagonal or the zeros elsewhere in the matrix product exhibit an error of $\pm r^{\frac{1}{2}n}$, this indicates the presence of a significant number of leading zeros prior to multiple bit shifts in the mantissa during the LU decomposition.

4.4. What Is Least Squares Regression? How Does It Physically Work?

Least Squares Regression describes what happens when I attempt to measure n data points using fewer than n rulers, regardless of the number of hypercomplex components in either the rulers or the data points.

Since I lack a sufficient number of rulers to make an exact measurement, we instead agree to approximate your measurements with some error, quantified as

$$\Psi = \ln \frac{RSS}{TSS} = \ln \left(1 - R^2 \right)$$

where RSS is the residual sum of squares and TSS is the total sum of squares.

- If Ψ is negative, our model (as measured by RSS) performs better than the baseline hypercomplex average of the y-measurements (given by TSS). - If Ψ is positive, our model is worse than the baseline average.

To ensure that our model is always at least as good as the baseline, we can introduce a hypercomplex constant to all regressions. This guarantees that

$$\Psi \leq 0 \quad \Rightarrow \quad R^2 = \left(1 - \frac{RSS}{TSS} \right) \geq 0$$

Thus, by construction, our model will never perform worse than the baseline. I personally measure Ψ in log base two, instead of base e, so that I understand exactly what it means: $2^{-\Psi}$ more precise than the baseline average.

4.4.1. Low R^2 Values Don't Always Indicate a Bad Fit!

It is important to recognize that an R^2 value near zero (or equivalently, a Ψ value near zero) does not necessarily indicate a poor model.

Consider a dataset of real-number percentages with an average of 70% and a standard deviation of only 0.05%. In this case, the baseline average is already an excellent predictor. However, performing a linear regression on these percentages would yield an R^2 value close to zero, despite the regression still being a slightly better predictor than the baseline — an already strong model.

To better analyze such data, one can apply a 45-degree rotation — transforming both the quantile index and the percentage data into a new basis—before rerunning the linear regression. Since the original data has minimal variance (only 0.05%), this rotation will align the percentages along a near-perfect 45-degree line, resulting in an extremely high R^2 value. This confirms that the baseline average was already a strong predictor in the original, unrotated dataset.

Of course, a 45-degree rotation is just one possible transformation, but it serves as a useful example to illustrate the concept.

In fact, even if the error matrix is ill-conditioned by our own previous definition, it doesn't actually matter — as long as the real number magnitude of the error of the expected regression vector has a mean and variance error remain within an acceptable tolerance compared to the actual regressed vector. Ultimately, that tolerance is entirely up to you! If necessary, you can simply adjust the ill-conditioning threshold by decreasing the $\frac{1}{2}n$ ill-conditioned threshold to smaller fraction, such as $\frac{1}{4}n$ or $\frac{1}{8}n$. Remember that lower fractional exponents result in larger numbers for $r^{-\frac{1}{k}n}$, which means higher error tolerance, since it's the kth root of the inverse radix!

4.5. What If I need Astronomical Precision...Like 13 Billion Light Years Level?

Here a radix $r = 65,536 = 2^{16}$, which cover every number from 0 to 65,535 per digit. Do you need 8 digits of precision which is one part in 340282366920938463463374607431768211456?

Ok then, write every real number in the real number matrix form as $a_7 r^7 + a_6 r^6 + \ldots a_0 r^0$. Make a precomputed table multiplication table that is 2^{16} by 2^{16} and reference the table as an instant lookup processor. Also compute the reciprocals of each number from 1 to 65536 out to 128+16 places (a total of

144 places), which is a table possessing only 9.4 million digits total (the pre-computed reciprocal table grows linearly with n digits).

Now multiply a and b as an 8x8 submatrix of A in hypocomplex tessarine form for each of the z^3 steps of the LU decomposition, where z is the matrix size (since I'm quite sure you don't know what the tessarine form is of a radix of n digits, you can see it below).

$$ab = AB = \begin{pmatrix} a_7r^7 & a_6r^6 & a_5r^5 & a_4r^4 & a_3r^3 & a_2r^2 & a_1r^1 & a_0r^0 \\ a_6r^6 & a_7r^7 & a_4r^4 & a_3r^3 & a_2r^2 & a_3r^3 & a_0r^0 & a_1r^1 \\ a_5r^5 & a_4r^4 & a_7r^7 & a_6r^6 & a_1r^1 & a_0r^0 & a_3r^3 & a_2r^2 \\ a_4r^4 & a_5r^5 & a_6r^6 & a_7r^7 & a_0r^0 & a_1r^1 & a_2r^2 & a_3r^3 \\ a_3r^3 & a_2r^2 & a_1r^1 & a_0r^0 & a_7r^7 & a_6r^6 & a_5r^5 & a_4r^4 \\ a_2r^2 & a_3r^3 & a_0r^0 & a_1r^1 & a_6r^6 & a_7r^7 & a_4r^4 & a_3r^3 \\ a_1r^1 & a_0r^0 & a_3r^3 & a_2r^2 & a_5r^5 & a_4r^4 & a_7r^7 & a_6r^6 \\ a_0r^0 & a_1r^1 & a_2r^2 & a_3r^3 & a_4r^4 & a_5r^5 & a_6r^6 & a_7r^7 \end{pmatrix} \begin{pmatrix} b_7r^7 \\ b_6r^6 \\ b_5r^5 \\ b_4r^4 \\ b_3r^3 \\ b_2r^2 \\ b_1r^1 \\ b_0r^0 \end{pmatrix}$$

There's no need to worry about the r^m portion, that's the local exponent of the radix. When you need to multiply ab during LU decomposition, use the precomputed table on the 16 digit a_m mantissas. This results in a total of 64 look-ups in this example and 64 additions. Now just apply the global floating point exponent of $(ar^x)(br^y)$ to get the final result $(ar^x)(ar^y) = r^{x+y}(AB)$, where AB is the fixed length result, and r^{x+y} yields the mobility to an exponent of any arbitrary size

If you need to divide b by a, multiply by the inverse matrix of A, because $ab = c \iff AB = C \iff b = \frac{c}{a} \iff B = A^{-1}C$. This requires $\frac{4}{3}8^3$ operations with full pivoting. Utilize the pre-computed reciprocal tables for the division in each step of the LU decomposition of the inverse 8x8 sub-matrix. Then multiply $A^{-1}B$, which requires n lookup per entry, which is a total of n^3 lookups. The result is a column matrix like B, telling you the digit of each radix power. Assuming your pre computed reciprocal table went out to 144 binary digits (for this example), you should have no noticeable precision loss. Those tessarines are damn useful, aren't they?

After you multiply $A^{-1}B$, round off the last 16 digits of the fixed-length r^{-1} portion into the r^0 portion to maintain the eight-digit format. Retaining those extra 16 binary digits would create a nine-digit representation, which is unacceptable! This is the only extra step from direct multiplication of AB, which always gives the exact answer.

If, for some extraordinary reason, you require even higher precision (to minimize error propagation) when inverting the 8x8 matrix, simply extend your precomputed reciprocal tables — covering values from 1 to 65,535 — by an additional $16k$ binary digits. This ensures that rounding to r^0 (eliminating the excess digits beyond the fixed-length mantissa) remains as precise as necessary. Don't hesitate — since the precomputed reciprocal table scales **linearly** with k, you can expand precision as far as needed without exponential growth in storage.

And if you think you're about to get rich patenting hardware to pull this off—sorry to break it to you, but my friend Rodger Fuller and I already beat you to it! Although I'd definitely argue that my boy, Aiden Degrace, should get half my royalties.

4.5.1. What is the Physical Mechanism Behind Least Squares Regression?

Good news! This is the final part of our philosophical discussion on the subject in this otherwise condensed paper. Feel free to skip this section if you're not interested in the underlying mechanics.

So what is it? We'll start with the Pseudo-Interpretation, which assumes that $\mathbb{R}^n$ behaves like a multi-dimensional space.

Pseudo-Interpretation: Flat Plane Regression in $\mathbb{R}^3$

1.   We are given the equation:

$$z_t = c_0 + c_1x_t + c_2y_t$$

where $x, y, z \in \mathbb{R}$ and t indexes each data point.

2. If there are exactly three data points, we obtain a perfect fit—a unique flat plane that passes through all three (x_t, y_t, z_t) data points.

3. If there are fewer than three data points:

    *   With just one data point, an infinite number of planes can pass through it.
    *   With two data points, an infinite number of planes can pass through the line connecting them.
    *   With zero data points, the solution is trivial (the zero vector).

    In these cases, solving for a unique regression plane is impossible.

4. If there are more than three data points, we construct an error matrix, invert it, and multiply it by the response vector. Despite the presence of error, Mother Nature treats the multiplication of the 3×3 matrix with the 3×1 response vector as if no error exists.

5. This is because, no matter how many data points we provide, Mother Nature only "sees" three ghost points—points that do not correspond to any of the original data points.

6. She then fits a perfect plane through these three ghost points, since whatever plane passes through them must be the one that minimizes the residual error across all of the actual data points. And that's it!

7. By the way, for this to be a legitimate plane regression in 3D space, we're actually calculating $z\vec{j} = c_0\vec{j} + c_1\vec{qj} + c_2\vec{ik} = (c_0 + c_1 + c_2)\vec{j}$ over the quaternions.

Actual: Ruler Regression in $\mathbb{R}^3$

1. We are given the equation:

$$z_t = c_0 + c_1 x_t + c_2 y_t$$

    where $x, y, z \in \mathbb{R}$ and t indexes each data point.

2. If there are exactly three data points, we obtain a perfect fit— three rulers that can measure all three (x_t, y_t, z_t) data points.

3. If there are fewer than three data points:

    *   With less than three data points, an infinite number of x, y ruler-trios can measure each of the z's. In these cases, solving for a unique set of rulers is impossible.

4. If there are more than three data points, we construct an error matrix, invert it, and multiply it by the response ruler. Despite the presence of error, Mother Nature treats the multiplication of the 3×3 matrix with the 3×1 response ruler as if no error exists.

5. This is because, no matter how many data points we provide, Mother Nature only "sees" three ghost rulers — rulers that do not correspond to any of the original data points.

6. She then fits a perfect set of three ghost rulers through the response ruler, since whatever rulers measure the z's must be the one that minimizes the residual error across all of the actual data points.

7. Nothing changes when our rulers or z measurements are hypercomplex numbers, other than the rulers and measurements having directions beyond the forward vector.

8. My My two cents on the apparent probability-driven nature of Quantum Mechanics? I believe that Nature collapses systems onto ghost points — abstract constructs that don't actually exist in our universe — in order to reduce computational complexity. If true (even though untestable, as it doesn't conflict with Hidden-But-Knowable Variables, because the ghost points are inherently Unknowable, they are Nature's fiction during the number crunch made real post-rendering!), this would support Wolfram's hypothesis that the speed of light is the speed of computation. To maintain this speed, Mother Nature collapses onto these ghost points when the computation time expires. While this remains untestable directly, if rival theories (Bohmian Hidden Variables, Many Worlds, Objective Collapse, etc.) are eliminated, and if anything lends credence to Wolfram's speed-of-computation hypothesis, I believe it could be confirmed indirectly — invoking the wisdom of Sherlock Holmes: *"When you have eliminated the impossible, whatever remains, however improbable, must be the truth."*

5. Univariate Least Squares, with and without a Constant

5.1. Univariate Least Squares, No Constant

We're given two quaternionic lists, X (and/or Y) and Z, of length m. We are then asked find the best fit, left or right handed, such that the constant, $\vec{c}$, produces the least error.

5.1.1. Human Readable-Forms of the Univariate Case

The subsequent form are intended for programmers, since one must always use a machine to perform calculations of such magnitude in a timely fashion. However, such things are not kind on the eyes, especially for a referee reviewing this paper, so below are the human-readable versions:

1. Let t be the total number of data points. $R^2 = 1 - \frac{RSS}{TSS}$ and $\Psi = \ln \frac{RSS}{TSS}$.

$$\bar{z} = \frac{1}{t} \sum_{b=0}^{b=t-1} \vec{z}_t$$

$$TSS = \sum_{b=0}^{b=t-1} (\vec{z}_t - \bar{z})(\vec{z}_t - \bar{z})^*$$

2. The Left-Handed Regression of $\vec{z}_b = \vec{c}\vec{y}_b$

$$\vec{c} = \left(\sum_{b=0}^{b=t-1} [\vec{y}_t^*]_{4\mathbb{H}R} [\vec{y}_t]_{4\mathbb{H}R} \right)^{-1} \left(\sum_{b=0}^{b=t-1} [\vec{y}_t^*]_{4\mathbb{H}R} \vec{z}_t \right)$$

$$RSS = \sum_{b=0}^{b=t-1} (\vec{z}_t - \vec{c}\vec{y}_t)(\vec{z}_t - \vec{c}\vec{y}_t)^*$$

3. The Right-Handed Regression of $\vec{z}_b = \vec{x}_b\vec{c}$

$$\vec{c} = \left(\sum_{b=0}^{b=t-1} [\vec{x}_t^*]_{4\mathbb{H}L} [\vec{x}_t]_{4\mathbb{H}L} \right)^{-1} \left(\sum_{b=0}^{b=t-1} [\vec{x}_t^*]_{4\mathbb{H}L} \vec{z}_t \right)$$

$$RSS = \sum_{b=0}^{b=t-1} (\vec{z}_t - \vec{x}_t\vec{c})(\vec{z}_t - \vec{x}_t\vec{c})^*$$

4. The Middle-Handed Regression of $\vec{z}_b = \vec{x}_b\vec{c}\vec{y}_b$

$$\vec{c} = \left(\sum_{b=0}^{b=t-1} (([\vec{y}_t^*]_{4\mathbb{H}R} [\vec{x}_t^*]_{4\mathbb{H}R}) [\vec{x}_t]_{4\mathbb{H}L} [\vec{y}_t]_{4\mathbb{H}R}) \right)^{-1} \left(\sum_{b=0}^{b=t-1} [\vec{y}_t^*]_{4\mathbb{H}R} [\vec{x}_t^*]_{4\mathbb{H}L} \vec{z}_t \right)$$

$$RSS = \sum_{b=0}^{b=t-1} (\vec{z}_t - \vec{x}_t\vec{c}\vec{y}_t)(\vec{z}_t - \vec{x}_t\vec{c}\vec{y}_t)^*$$

Theorem 2. *Quaternionic Left-Handed Regression*

1. *Let Z be a $4 \times t$ real number matrix containing the four components of each $\vec{z}_t$ quaternion, such that each element of this matrix is of the form $Z_{a,b}$, with $0 \le p \le 3$ and $0 \le b \le t-1$.*
2. *Let X be a $4 \times t$ real number matrix containing the four components of each $\vec{x}_t$ quaternion, such that each element of this matrix is of the form $X_{a,b}$, with $0 \le p \le 3$ and $0 \le b \le t-1$.*
3. *Let X^* be a $4 \times t$ real number matrix containing the four components of the conjugate of each $\vec{x}_t$ quaternion, such that each element of this matrix is of the form $X_{a,b}^*$, with $0 \le p \le 3$ and $0 \le b \le t-1$, where:*

    (a) $X_{0,b}^* = X_{0,b}$
    (b) $X_{1,b}^* = -X_{1,b}$

(c) $\quad X_{2,b}^* = -X_{2,b}$

(d) $\quad X_{3,b}^* = -X_{3,b}$

(e) $\quad$ The above saves time for the regular quaternions, since the alternative would be to calculate the first column of the adjugate matrix form. This method does not extend to the General Case of Biquaternionic Regression.

4. We seek the quaternionic regression of $\vec{z}_t = \vec{c}\vec{x}_t$. The left-handed constant of $\vec{c}_t$ is unknown, which means we need the right-handed matrix of $\vec{x}_t$.

5. Let $[X]_{4\mathbb{H}R}$ be a three dimensional tensor that stores the Right-Handed Matrix Form of each $\vec{x}_t$ from the two-dimensional array of X. And let an element of this tensor be $[X]_{4\mathbb{H}R,u,v,b}$, with $0 \le u \le 3$, $0 \le v \le 3$ and $0 \le b \le t-1$, such that:

   (a) $\quad$ From $u = 0$ to $u = 3$ and from $v = 0$ to $v = 3$:

$$
\begin{pmatrix}
[X]_{4\mathbb{H}R,0,0,b} = +X_{0,b} & [X]_{4\mathbb{H}R,0,1,b} = -X_{1,b} & [X]_{4\mathbb{H}R,0,2,b} = -X_{2,b} & [X]_{4\mathbb{H}R,0,3,b} = -X_{3,b} \\
[X]_{4\mathbb{H}R,1,0,b} = +X_{1,b} & [X]_{4\mathbb{H}R,1,1,b} = +X_{0,b} & [X]_{4\mathbb{H}R,1,2,b} = -X_{3,b} & [X]_{4\mathbb{H}R,1,3,b} = +X_{2,b} \\
[X]_{4\mathbb{H}R,2,0,b} = +X_{2,b} & [X]_{4\mathbb{H}R,2,1,b} = +X_{3,b} & [X]_{4\mathbb{H}R,2,2,b} = +X_{0,b} & [X]_{4\mathbb{H}R,2,3,b} = -X_{1,b} \\
[X]_{4\mathbb{H}R,3,0,b} = +X_{3,b} & [X]_{4\mathbb{H}R,3,1,b} = -X_{2,b} & [X]_{4\mathbb{H}R,3,2,b} = +X_{1,b} & [X]_{4\mathbb{H}R,3,3,b} = +X_{0,b}
\end{pmatrix}
$$

   (b) $\quad$ And let a Matrix Element of this Tensor be referenced as $[X]_{4\mathbb{H}R,b}$.

6. Let $[X^*]_{4\mathbb{H}R}$ be a three dimensional tensor that stores the Right-Handed Matrix Form of each $\vec{x}_t^*$ from the two-dimensional array of X^*. And let an element of this tensor be $[X^*]_{4\mathbb{H}R,u,v,b}$, with $0 \le u \le 3$, $0 \le v \le 3$ and $0 \le b \le t-1$, such that:

   (a) $\quad$ From $u = 0$ to $u = 3$ and from $v = 0$ to $v = 3$:

$$
\begin{pmatrix}
[X^*]_{4\mathbb{H}R,0,0,b} = +X_{0,b}^* & [X^*]_{4\mathbb{H}R,0,1,b} = -X_{1,b}^* & [X^*]_{4\mathbb{H}R,0,2,b} = -X_{2,b}^* & [X^*]_{4\mathbb{H}R,0,3,b} = -X_{3,b}^* \\
[X^*]_{4\mathbb{H}R,1,0,b} = +X_{1,b}^* & [X^*]_{4\mathbb{H}R,1,1,b} = +X_{0,b}^* & [X^*]_{4\mathbb{H}R,1,2,b} = -X_{3,b}^* & [X^*]_{4\mathbb{H}R,1,3,b} = +X_{2,b}^* \\
[X^*]_{4\mathbb{H}R,2,0,b} = +X_{2,b}^* & [X^*]_{4\mathbb{H}R,2,1,b} = +X_{3,b}^* & [X^*]_{4\mathbb{H}R,2,2,b} = +X_{0,b}^* & [X^*]_{4\mathbb{H}R,2,3,b} = -X_{1,b}^* \\
[X^*]_{4\mathbb{H}R,3,0,b} = +X_{3,b}^* & [X^*]_{4\mathbb{H}R,3,1,b} = -X_{2,b}^* & [X^*]_{4\mathbb{H}R,3,2,b} = +X_{1,b}^* & [X^*]_{4\mathbb{H}R,3,3,b} = +X_{0,b}^*
\end{pmatrix}
$$

   (b) $\quad$ And let a Matrix Element of this Tensor be referenced as $[X^*]_{4\mathbb{H}R,b}$.

7. Let D be a three dimensional tensor that stores the product of $[X^*]_{4\mathbb{H}R,b}[X]_{4\mathbb{H}R,b}$. And let an element of this tensor be $D_{u,v,b}$, with $0 \le u \le 3$, $0 \le v \le 3$ and $0 \le b \le t-1$, such that:

   (a) $\quad D_{u,0,b} = [X^*]_{4\mathbb{H}R,u,0,b}[X]_{4\mathbb{H}R,0,0,b} + [X^*]_{4\mathbb{H}R,u,1,b}[X]_{4\mathbb{H}R,1,0,b} + [X^*]_{4\mathbb{H}R,u,2,b}[X]_{4\mathbb{H}R,2,0,b} + [X^*]_{4\mathbb{H}R,u,3,b}[X]_{4\mathbb{H}R,3,0,b}$

   (b) $\quad D_{u,1,b} = [X^*]_{4\mathbb{H}R,u,0,b}[X]_{4\mathbb{H}R,0,1,b} + [X^*]_{4\mathbb{H}R,u,1,b}[X]_{4\mathbb{H}R,1,1,b} + [X^*]_{4\mathbb{H}R,u,2,b}[X]_{4\mathbb{H}R,2,1,b} + [X^*]_{4\mathbb{H}R,u,3,b}[X]_{4\mathbb{H}R,3,1,b}$

   (c) $\quad D_{u,2,b} = [X^*]_{4\mathbb{H}R,u,0,b}[X]_{4\mathbb{H}R,0,2,b} + [X^*]_{4\mathbb{H}R,u,1,b}[X]_{4\mathbb{H}R,1,2,b} + [X^*]_{4\mathbb{H}R,u,2,b}[X]_{4\mathbb{H}R,2,2,b} + [X^*]_{4\mathbb{H}R,u,3,b}[X]_{4\mathbb{H}R,3,2,b}$

   (d) $\quad D_{u,3,b} = [X^*]_{4\mathbb{H}R,u,0,b}[X]_{4\mathbb{H}R,0,3,b} + [X^*]_{4\mathbb{H}R,u,1,b}[X]_{4\mathbb{H}R,1,3,b} + [X^*]_{4\mathbb{H}R,u,2,b}[X]_{4\mathbb{H}R,2,3,b} + [X^*]_{4\mathbb{H}R,u,3,b}[X]_{4\mathbb{H}R,3,3,b}$

   (e) $\quad$ And let a Matrix Element of this Tensor be referenced as D_b.

8. And let D_t be the Direct Matrix Sum of all D_b, such that an element of $D_{u,v,t}$ is equal to:

$$
D_{u,v,t} = \sum_{b=0}^{b=t-1} D_{u,v,b}
$$

9. Let R be a two dimensional tensor that stores the product of $[X^*]_{4\mathbb{H}R,b}\vec{z}_t$. And let an element of this tensor be $R_{u,b}$, with $0 \le u \le 3$ and $0 \le b \le t-1$, such that:

   (a) $\quad R_{u,b} = [X^*]_{4\mathbb{H}R,u,0,b}Z_{0,b} + [X^*]_{4\mathbb{H}R,u,1,b}Z_{1,b} + [X^*]_{4\mathbb{H}R,u,2,b}Z_{2,b} + [X^*]_{4\mathbb{H}R,u,3,b}Z_{3,b}$

   (b) $\quad$ And let a Column Matrix Element of this Tensor be reference as R_b.

10. *And let R_t be the Direct Matrix Sum of all R_b, such that an element of $R_{u,t}$ is equal to:*

$$R_{u,t} = \sum_{b=0}^{b=t-1} R_{u,b}$$

11. *Finally let G be the inverse matrix of D_t, such that each element of $G_{u,v}$ is equal to each entry of D_t^{-1}.*

12. *Then each component of the initially unknown $\vec{c}$ is given by each entry (respectively) of the product G and the column vector of R_t, such that:*

$$c_u = G_{u,0}Z_{0,t} + G_{u,1}Z_{1,t} + G_{u,2}Z_{2,t} + G_{u,3}Z_{3,t}$$

13. *The Residual Sum of Squares is a real number equal to $\sum_{b=0}^{b=t-1} \vec{g}_b\vec{g}_b^*$, where $\vec{g}_t = \vec{z}_t - \vec{c}\vec{x}_t$.*

14. *The Total Sum of Squares is a real number equal to $\sum_{b=0}^{b=t-1} \vec{h}_b\vec{h}_b^*$, where $\vec{h}_t = \vec{z}_t - \vec{f}$, where $\vec{f} = \frac{1}{t}\sum_{b=0}^{b=t-1} \vec{z}_b$.*

15. $R^2 = 1 - \frac{RSS}{TSS}$ *and* $\Psi = \ln\frac{RSS}{TSS}$.

Theorem 3. *Quaternionic Right-Handed Regression*

1. *Let Z be a $4 \times t$ real number matrix containing the four components of each $\vec{z}_t$ quaternion, such that each element of this matrix is of the form $Z_{a,b}$, with $0 \le p \le 3$ and $0 \le b \le t-1$.*

2. *Let X be a $4 \times t$ real number matrix containing the four components of each $\vec{x}_t$ quaternion, such that each element of this matrix is of the form $X_{a,b}$, with $0 \le p \le 3$ and $0 \le b \le t-1$.*

3. *Let X^* be a $4 \times t$ real number matrix containing the four components of the conjugate of each $\vec{x}_t$ quaternion, such that each element of this matrix is of the form $X_{a,b}^*$, with $0 \le p \le 3$ and $0 \le b \le t-1$, where:*

    (a) $X_{0,b}^* = X_{0,b}$

    (b) $X_{1,b}^* = -X_{1,b}$

    (c) $X_{2,b}^* = -X_{2,b}$

    (d) $X_{3,b}^* = -X_{3,b}$

    (e) *The above saves time for the regular quaternions, since the alternative would be to calculate the first column of the adjugate matrix form. This method does not extend to the General Case of Biquaternionic Regression.*

4. *We seek the quaternionic regression of $\vec{z}_t = \vec{x}_t\vec{c}$. The right-handed constant of $\vec{c}_t$ is unknown, which means we need the left-handed matrix of $\vec{x}_t$.*

5. *Let $[X]_{4\mathbb{H}L}$ be a three dimensional tensor that stores the Left-Handed Matrix Form of each $\vec{x}_t$ from the two-dimensional array of X. And let an element of this tensor be $[X]_{4\mathbb{H}L,u,v,b}$, with $0 \le u \le 3$, $0 \le v \le 3$ and $0 \le b \le t-1$, such that:*

    (a) *From $u = 0$ to $u = 3$ and from $v = 0$ to $v = 3$:*

$$\begin{pmatrix} [X]_{4\mathbb{H}L,0,0,b} = +X_{0,b} & [X]_{4\mathbb{H}L,0,1,b} = -X_{1,b} & [X]_{4\mathbb{H}L,0,2,b} = -X_{2,b} & [X]_{4\mathbb{H}L,0,3,b} = -X_{3,b} \\ [X]_{4\mathbb{H}L,1,0,b} = +X_{1,b} & [X]_{4\mathbb{H}L,1,1,b} = +X_{0,b} & [X]_{4\mathbb{H}L,1,2,b} = +X_{3,b} & [X]_{4\mathbb{H}L,1,3,b} = -X_{2,b} \\ [X]_{4\mathbb{H}L,2,0,b} = +X_{2,b} & [X]_{4\mathbb{H}L,2,1,b} = -X_{3,b} & [X]_{4\mathbb{H}L,2,2,b} = +X_{0,b} & [X]_{4\mathbb{H}L,2,3,b} = +X_{1,b} \\ [X]_{4\mathbb{H}L,3,0,b} = +X_{3,b} & [X]_{4\mathbb{H}L,3,1,b} = +X_{2,b} & [X]_{4\mathbb{H}L,3,2,b} = -X_{1,b} & [X]_{4\mathbb{H}L,3,3,b} = +X_{0,b} \end{pmatrix}$$

    (b) *And let a Matrix Element of this Tensor be referenced as $[X]_{4\mathbb{H}L,b}$.*

6. *Let $[X^*]_{4\mathbb{H}L}$ be a three dimensional tensor that stores the Left-Handed Matrix Form of each $\vec{x}_t^*$ from the two-dimensional array of X^*. And let an element of this tensor be $[X^*]_{4\mathbb{H}L,u,v,b}$, with $0 \le u \le 3$, $0 \le v \le 3$ and $0 \le b \le t-1$, such that:*

(a)   *From $u = 0$ to $u = 3$ and from $v = 0$ to $v = 3$:*

$$\begin{pmatrix} [X^*]_{4\mathbb{H}L,0,0,b} = +X_{0,b}^* & [X^*]_{4\mathbb{H}L,0,1,b} = -X_{1,b}^* & [X^*]_{4\mathbb{H}L,0,2,b} = -X_{2,b}^* & [X^*]_{4\mathbb{H}L,0,3,b} = -X_{3,b}^* \\ [X^*]_{4\mathbb{H}L,1,0,b} = +X_{1,b}^* & [X^*]_{4\mathbb{H}L,1,1,b} = +X_{0,b}^* & [X^*]_{4\mathbb{H}L,1,2,b} = +X_{3,b}^* & [X^*]_{4\mathbb{H}L,1,3,b} = -X_{2,b} \\ [X^*]_{4\mathbb{H}L,2,0,b} = +X_{2,b}^* & [X^*]_{4\mathbb{H}L,2,1,b} = -X_{3,b}^* & [X^*]_{4\mathbb{H}L,2,2,b} = +X_{0,b}^* & [X^*]_{4\mathbb{H}L,2,3,b} = +X_{1,b}^* \\ [X^*]_{4\mathbb{H}L,3,0,b} = +X_{3,b}^* & [X^*]_{4\mathbb{H}L,3,1,b} = +X_{2,b}^* & [X^*]_{4\mathbb{H}L,3,2,b} = -X_{1,b}^* & [X^*]_{4\mathbb{H}L,3,3,b} = +X_{0,b}^* \end{pmatrix}$$

(b)   *And let a Matrix Element of this Tensor be referenced as $[X^*]_{4\mathbb{H}L,b}$.*

7.   *Let D be a three dimensional tensor that stores the product of $[X^*]_{4\mathbb{H}L,b}[X]_{4\mathbb{H}L,b}$. And let an element of this tensor be $D_{u,v,b}$, with $0 \le u \le 3$, $0 \le v \le 3$ and $0 \le b \le t-1$, such that:*

(a) $D_{u,0,b} = [X^*]_{4\mathbb{H}L,u,0,b}[X]_{4\mathbb{H}L,0,0,b} + [X^*]_{4\mathbb{H}L,u,1,b}[X]_{4\mathbb{H}L,1,0,b} + [X^*]_{4\mathbb{H}L,u,2,b}[X]_{4\mathbb{H}L,2,0,b}$
$+ [X^*]_{4\mathbb{H}L,u,3,b}[X]_{4\mathbb{H}L,3,0,b}$

(b) $D_{u,1,b} = [X^*]_{4\mathbb{H}L,u,0,b}[X]_{4\mathbb{H}L,0,1,b} + [X^*]_{4\mathbb{H}L,u,1,b}[X]_{4\mathbb{H}L,1,1,b} + [X^*]_{4\mathbb{H}L,u,2,b}[X]_{4\mathbb{H}L,2,1,b}$
$+ [X^*]_{4\mathbb{H}L,u,3,b}[X]_{4\mathbb{H}L,3,1,b}$

(c) $D_{u,2,b} = [X^*]_{4\mathbb{H}L,u,0,b}[X]_{4\mathbb{H}L,0,2,b} + [X^*]_{4\mathbb{H}L,u,1,b}[X]_{4\mathbb{H}L,1,2,b} + [X^*]_{4\mathbb{H}L,u,2,b}[X]_{4\mathbb{H}L,2,2,b}$
$+ [X^*]_{4\mathbb{H}L,u,3,b}[X]_{4\mathbb{H}L,3,2,b}$

(d) $D_{u,3,b} = [X^*]_{4\mathbb{H}L,u,0,b}[X]_{4\mathbb{H}L,0,3,b} + [X^*]_{4\mathbb{H}L,u,1,b}[X]_{4\mathbb{H}L,1,3,b} + [X^*]_{4\mathbb{H}L,u,2,b}[X]_{4\mathbb{H}L,2,3,b}$
$+ [X^*]_{4\mathbb{H}L,u,3,b}[X]_{4\mathbb{H}L,3,3,b}$

(e)   *And let a Matrix Element of this Tensor be referenced as D_b.*

8.   *And let D_t be the Direct Matrix Sum of all D_b, such that an element of $D_{u,v,t}$ is equal to:*

$$D_{u,v,t} = \sum_{b=0}^{b=t-1} D_{u,v,b}$$

9.   *Let R be a two dimensional tensor that stores the product of $[X^*]_{4\mathbb{H}L,b}\vec{z}_t$. And let an element of this tensor be $R_{u,b}$, with $0 \le u \le 3$ and $0 \le b \le t-1$, such that:*

(a) $R_{u,b} = [X^*]_{4\mathbb{H}L,u,0,b}Z_{0,b} + [X^*]_{4\mathbb{H}L,u,1,b}Z_{1,b} + [X^*]_{4\mathbb{H}L,u,2,b}Z_{2,b} + [X^*]_{4\mathbb{H}L,u,3,b}Z_{3,b}$

(b)   *And let a Column Matrix Element of this Tensor be referenced as R_b.*

10.   *And let R_t be the Direct Matrix Sum of all R_b, such that an element of $R_{u,t}$ is equal to:*

$$R_{u,t} = \sum_{b=0}^{b=t-1} R_{u,b}$$

11.   *Finally let G be the inverse matrix of D_t, such that each element of $G_{u,v}$ is equal to each entry of D_t^{-1}.*

12.   *Then each component of the initially unknown $\vec{c}$ is given by each entry (respectively) of the product G and the column vector of R_t, such that:*

$$c_u = G_{u,0}Z_{0,t} + G_{u,1}Z_{1,t} + G_{u,2}Z_{2,t} + G_{u,3}Z_{3,t}$$

13.   *The Residual Sum of Squares is a real number equal to $\sum_{b=0}^{b=t-1} \vec{g}_b \vec{g}_b^*$, where $\vec{g}_t = \vec{z}_t - \vec{x}_t \vec{c}$.*

14.   *The Total Sum of Squares is a real number equal to $\sum_{b=0}^{b=t-1} \vec{h}_b \vec{h}_b^*$, where $\vec{h}_t = \vec{z}_t - \vec{f}$, where $\vec{f} = \frac{1}{t}\sum_{b=0}^{b=t-1}\vec{z}_b$.*

15. $R^2 = 1 - \frac{RSS}{TSS}$ *and* $\Psi = \ln \frac{RSS}{TSS}$.

Theorem 4. *Quaternionic Middle-Handed Regression*

1.   *Let Z be a $4 \times t$ real number matrix containing the four components of each $\vec{z}_t$ quaternion, such that each element of this matrix is of the form $Z_{a,b}$, with $0 \le p \le 3$ and $0 \le b \le t-1$.*

2.   *Let X be a $4 \times t$ real number matrix containing the four components of each $\vec{x}_t$ quaternion, such that each element of this matrix is of the form $X_{a,b}$, with $0 \le p \le 3$ and $0 \le b \le t-1$.*

3. Let Y be a $4 \times t$ real number matrix containing the four components of each $\vec{y}_t$ quaternion, such that each element of this matrix is of the form $Y_{a,b}$, with $0 \leq p \leq 3$ and $0 \leq b \leq t-1$.

4. Let X^* be a $4 \times t$ real number matrix containing the four components of the conjugate of each $\vec{x}_t$ quaternion, such that each element of this matrix is of the form $X^*_{a,b}$, with $0 \leq p \leq 3$ and $0 \leq b \leq t-1$, where:

5. Let Y^* be a $4 \times t$ real number matrix containing the four components of the conjugate of each $\vec{y}_t$ quaternion, such that each element of this matrix is of the form $Y^*_{a,b}$, with $0 \leq p \leq 3$ and $0 \leq b \leq t-1$, where:

    (a) $X^*_{0,b} = +X_{0,b}$ and $Y^*_{0,b} = +Y_{0,b}$

    (b) $X^*_{1,b} = -X_{1,b}$ and $Y^*_{1,b} = -Y_{1,b}$

    (c) $X^*_{2,b} = -X_{2,b}$ and $Y^*_{2,b} = -Y_{2,b}$

    (d) $X^*_{3,b} = -X_{3,b}$ and $Y^*_{3,b} = -Y_{3,b}$

    (e)   The above saves time for the regular quaternions, since the alternative would be to calculate the first column of the adjugate matrix form. This method does not extend to the General Case of Biquaternionic Regression.

6. We seek the quaternionic regression of $\vec{z}_t = \vec{x}_t \vec{c} \vec{y}$. The middle-handed constant of $\vec{c}_t$ is unknown, which means we need the left-handed matrix of $\vec{x}_t$ and the right-handed matrix of $\vec{y}_t$.

7. Let $[X]_{4\mathbb{H}L}$ be a three dimensional tensor that stores the Left-Handed Matrix Form of each $\vec{x}_t$ from the two-dimensional array of X. And let an element of this tensor be $[X]_{4\mathbb{H}L,u,v,b}$, with $0 \leq u \leq 3$, $0 \leq v \leq 3$ and $0 \leq b \leq t-1$, such that:

    (a)   From $u = 0$ to $u = 3$ and from $v = 0$ to $v = 3$:

$$
\begin{pmatrix}
[X]_{4\mathbb{H}L,0,0,b} = +X_{0,b} & [X]_{4\mathbb{H}L,0,1,b} = -X_{1,b} & [X]_{4\mathbb{H}L,0,2,b} = -X_{2,b} & [X]_{4\mathbb{H}L,0,3,b} = -X_{3,b} \\
[X]_{4\mathbb{H}L,1,0,b} = +X_{1,b} & [X]_{4\mathbb{H}L,1,1,b} = +X_{0,b} & [X]_{4\mathbb{H}L,1,2,b} = +X_{3,b} & [X]_{4\mathbb{H}L,1,3,b} = -X_{2,b} \\
[X]_{4\mathbb{H}L,2,0,b} = +X_{2,b} & [X]_{4\mathbb{H}L,2,1,b} = -X_{3,b} & [X]_{4\mathbb{H}L,2,2,b} = +X_{0,b} & [X]_{4\mathbb{H}L,2,3,b} = +X_{1,b} \\
[X]_{4\mathbb{H}L,3,0,b} = +X_{3,b} & [X]_{4\mathbb{H}L,3,1,b} = +X_{2,b} & [X]_{4\mathbb{H}L,3,2,b} = -X_{1,b} & [X]_{4\mathbb{H}L,3,3,b} = +X_{0,b}
\end{pmatrix}
$$

    (b)   And let a Matrix Element of this Tensor be referenced as $[X]_{4\mathbb{H}L,b}$.

8. Let $[X^*]_{4\mathbb{H}L}$ be a three dimensional tensor that stores the Left-Handed Matrix Form of each $\vec{x}_t^*$ from the two-dimensional array of X^*. And let an element of this tensor be $[X^*]_{4\mathbb{H}L,u,v,b}$, with $0 \leq u \leq 3$, $0 \leq v \leq 3$ and $0 \leq b \leq t-1$, such that:

    (a)   From $u = 0$ to $u = 3$ and from $v = 0$ to $v = 3$:

$$
\begin{pmatrix}
[X^*]_{4\mathbb{H}L,0,0,b} = +X^*_{0,b} & [X^*]_{4\mathbb{H}L,0,1,b} = -X^*_{1,b} & [X^*]_{4\mathbb{H}L,0,2,b} = -X^*_{2,b} & [X^*]_{4\mathbb{H}L,0,3,b} = -X^*_{3,b} \\
[X^*]_{4\mathbb{H}L,1,0,b} = +X^*_{1,b} & [X^*]_{4\mathbb{H}L,1,1,b} = +X^*_{0,b} & [X^*]_{4\mathbb{H}L,1,2,b} = +X^*_{3,b} & [X^*]_{4\mathbb{H}L,1,3,b} = -X^*_{2,b} \\
[X^*]_{4\mathbb{H}L,2,0,b} = +X^*_{2,b} & [X^*]_{4\mathbb{H}L,2,1,b} = -X^*_{3,b} & [X^*]_{4\mathbb{H}L,2,2,b} = +X^*_{0,b} & [X^*]_{4\mathbb{H}L,2,3,b} = +X^*_{1,b} \\
[X^*]_{4\mathbb{H}L,3,0,b} = +X^*_{3,b} & [X^*]_{4\mathbb{H}L,3,1,b} = +X^*_{2,b} & [X^*]_{4\mathbb{H}L,3,2,b} = -X^*_{1,b} & [X^*]_{4\mathbb{H}L,3,3,b} = +X^*_{0,b}
\end{pmatrix}
$$

    (b)   And let a Matrix Element of this Tensor be referenced as $[X^*]_{4\mathbb{H}L,b}$.

9. Let $[Y]_{4\mathbb{H}R}$ be a three dimensional tensor that stores the Right-Handed Matrix Form of each $\vec{y}_t$ from the two-dimensional array of Y. And let an element of this tensor be $[Y]_{4\mathbb{H}R,u,v,b}$, with $0 \leq u \leq 3$, $0 \leq v \leq 3$ and $0 \leq b \leq t-1$, such that:

    (a)   From $u = 0$ to $u = 3$ and from $v = 0$ to $v = 3$:

$$
\begin{pmatrix}
[Y]_{4\mathbb{H}R,0,0,b} = +Y_{0,b} & [Y]_{4\mathbb{H}R,0,1,b} = -Y_{1,b} & [Y]_{4\mathbb{H}R,0,2,b} = -Y_{2,b} & [Y]_{4\mathbb{H}R,0,3,b} = -Y_{3,b} \\
[Y]_{4\mathbb{H}R,1,0,b} = +Y_{1,b} & [Y]_{4\mathbb{H}R,1,1,b} = +Y_{0,b} & [Y]_{4\mathbb{H}R,1,2,b} = -Y_{3,b} & [Y]_{4\mathbb{H}R,1,3,b} = +Y_{2,b} \\
[Y]_{4\mathbb{H}R,2,0,b} = +Y_{2,b} & [Y]_{4\mathbb{H}R,2,1,b} = +Y_{3,b} & [Y]_{4\mathbb{H}R,2,2,b} = +Y_{0,b} & [Y]_{4\mathbb{H}R,2,3,b} = -Y_{1,b} \\
[Y]_{4\mathbb{H}R,3,0,b} = +Y_{3,b} & [Y]_{4\mathbb{H}R,3,1,b} = -Y_{2,b} & [Y]_{4\mathbb{H}R,3,2,b} = +Y_{1,b} & [Y]_{4\mathbb{H}R,3,3,b} = +Y_{0,b}
\end{pmatrix}
$$

    (b)   And let a Matrix Element of this Tensor be referenced as $[Y]_{4\mathbb{H}R,b}$.

10. Let $[Y^*]_{4\mathbb{H}R}$ be a three dimensional tensor that stores the Right-Handed Matrix Form of each $\vec{y}_t^*$ from the two-dimensional array of Y^*. And let an element of this tensor be $[Y^*]_{4\mathbb{H}R,u,v,b}$, with $0 \leq u \leq 3$, $0 \leq v \leq 3$ and $0 \leq b \leq t-1$, such that:

(a) From $u = 0$ to $u = 3$ and from $v = 0$ to $v = 3$:

$$\begin{pmatrix} [Y^*]_{4\mathbb{H}R,0,0,b} = +Y^*_{0,b} & [Y^*]_{4\mathbb{H}R,0,1,b} = -Y^*_{1,b} & [Y^*]_{4\mathbb{H}R,0,2,b} = -Y^*_{2,b} & [Y^*]_{4\mathbb{H}R,0,3,b} = -Y^*_{3,b} \\ [Y^*]_{4\mathbb{H}R,1,0,b} = +Y^*_{1,b} & [Y^*]_{4\mathbb{H}R,1,1,b} = +Y^*_{0,b} & [Y^*]_{4\mathbb{H}R,1,2,b} = -Y^*_{3,b} & [Y^*]_{4\mathbb{H}R,1,3,b} = +Y_{2,b} \\ [Y^*]_{4\mathbb{H}R,2,0,b} = +Y^*_{2,b} & [Y^*]_{4\mathbb{H}R,2,1,b} = +Y^*_{3,b} & [Y^*]_{4\mathbb{H}R,2,2,b} = +Y^*_{0,b} & [Y^*]_{4\mathbb{H}R,2,3,b} = -Y^*_{1,b} \\ [Y^*]_{4\mathbb{H}R,3,0,b} = +Y^*_{3,b} & [Y^*]_{4\mathbb{H}R,3,1,b} = -Y^*_{2,b} & [Y^*]_{4\mathbb{H}R,3,2,b} = +Y^*_{1,b} & [Y^*]_{4\mathbb{H}R,3,3,b} = +Y^*_{0,b} \end{pmatrix}$$

(b) And let a Matrix Element of this Tensor be referenced as $[Y^*]_{4\mathbb{H}R,b}$.

11. Let D be a four-dimensional tensor.

12. Let D_1 be a three dimensional tensor, which is the first 3D layer of D, that stores the product of $[X]_{4\mathbb{H}L,b}[Y]_{4\mathbb{H}R,b}$. And let an element of this tensor be $D_{1,u,v,b}$, with $0 \leq u \leq 3$, $0 \leq v \leq 3$ and $0 \leq b \leq t-1$, such that:

(a) $D_{1,u,0,b} = [X]_{4\mathbb{H}L,u,0,b}[Y]_{4\mathbb{H}R,0,0,b} + [X]_{4\mathbb{H}L,u,1,b}[Y]_{4\mathbb{H}R,1,0,b} + [X]_{4\mathbb{H}L,u,2,b}[Y]_{4\mathbb{H}R,2,0,b} + [X]_{4\mathbb{H}L,u,3,b}[Y]_{4\mathbb{H}R,3,0,b}$

(b) $D_{1,u,1,b} = [X]_{4\mathbb{H}L,u,0,b}[Y]_{4\mathbb{H}R,0,1,b} + [X]_{4\mathbb{H}L,u,1,b}[Y]_{4\mathbb{H}R,1,1,b} + [X]_{4\mathbb{H}L,u,2,b}[Y]_{4\mathbb{H}R,2,1,b} + [X]_{4\mathbb{H}L,u,3,b}[Y]_{4\mathbb{H}R,3,1,b}$

(c) $D_{1,u,2,b} = [X]_{4\mathbb{H}L,u,0,b}[Y]_{4\mathbb{H}R,0,2,b} + [X]_{4\mathbb{H}L,u,1,b}[Y]_{4\mathbb{H}R,1,2,b} + [X]_{4\mathbb{H}L,u,2,b}[Y]_{4\mathbb{H}R,2,2,b} + [X]_{4\mathbb{H}L,u,3,b}[Y]_{4\mathbb{H}R,3,2,b}$

(d) $D_{1,u,3,b} = [X]_{4\mathbb{H}L,u,0,b}[Y]_{4\mathbb{H}R,0,3,b} + [X]_{4\mathbb{H}L,u,1,b}[Y]_{4\mathbb{H}R,1,3,b} + [X]_{4\mathbb{H}L,u,2,b}[Y]_{4\mathbb{H}R,2,3,b} + [X]_{4\mathbb{H}L,u,3,b}[Y]_{4\mathbb{H}R,3,3,b}$

(e) And let a Matrix Element of this Tensor be referenced as $D_{1,b}$.

13. Let D_2 be a three dimensional tensor, which is the second 3D layer of D, that stores the product of $[Y^*]_{4\mathbb{H}R,b}[X^*]_{4\mathbb{H}L,b}$. And let an element of this tensor be $D_{2,u,v,b}$, with $0 \leq u \leq 3$, $0 \leq v \leq 3$ and $0 \leq b \leq t-1$, such that:

(a) $D_{2,u,0,b} = [X^*]_{4\mathbb{H}R,u,0,b}[X^*]_{4\mathbb{H}L,0,0,b} + [Y^*]_{4\mathbb{H}R,u,1,b}[X^*]_{4\mathbb{H}L,1,0,b} + [Y^*]_{4\mathbb{H}R,u,2,b}[X^*]_{4\mathbb{H}L,2,0,b} + [Y^*]_{4\mathbb{H}R,u,3,b}[X^*]_{4\mathbb{H}L,3,0,b}$

(b) $D_{2,u,1,b} = [Y^*]_{4\mathbb{H}R,u,0,b}[X^*]_{4\mathbb{H}L,0,1,b} + [Y^*]_{4\mathbb{H}R,u,1,b}[X^*]_{4\mathbb{H}L,1,1,b} + [Y^*]_{4\mathbb{H}R,u,2,b}[X^*]_{4\mathbb{H}L,2,1,b} + [Y^*]_{4\mathbb{H}R,u,3,b}[X^*]_{4\mathbb{H}L,3,1,b}$

(c) $D_{2,u,2,b} = [Y^*]_{4\mathbb{H}R,u,0,b}[X^*]_{4\mathbb{H}L,0,2,b} + [Y^*]_{4\mathbb{H}R,u,1,b}[X^*]_{4\mathbb{H}L,1,2,b} + [Y^*]_{4\mathbb{H}R,u,2,b}[X^*]_{4\mathbb{H}L,2,2,b} + [Y^*]_{4\mathbb{H}R,u,3,b}[X^*]_{4\mathbb{H}L,3,2,b}$

(d) $D_{2,u,3,b} = [Y^*]_{4\mathbb{H}R,u,0,b}[X^*]_{4\mathbb{H}L,0,3,b} + [Y^*]_{4\mathbb{H}R,u,1,b}[X^*]_{4\mathbb{H}L,1,3,b} + [Y^*]_{4\mathbb{H}R,u,2,b}[X^*]_{4\mathbb{H}L,2,3,b} + [Y^*]_{4\mathbb{H}R,u,3,b}[X^*]_{4\mathbb{H}L,3,3,b}$

(e) And let a Matrix Element of this Tensor be referenced as $D_{2,b}$.

14. Let D_0 be a three dimensional tensor, which is the zeroth 3D layer of D, that stores the product of $D_{1,b}D_{2,b}$. And let an element of this tensor be $D_{0,u,v,b}$, with $0 \leq u \leq 3$, $0 \leq v \leq 3$ and $0 \leq b \leq t-1$, such that:

(a) $D_{0,u,0,b} = D_{1,u,0,b}D_{2,0,0,b} + D_{1,u,1,b}D_{2,1,0,b} + D_{1,u,2,b}D_{2,2,0,b} + D_{1,u,3,b}D_{2,3,0,b}$

(b) $D_{0,u,1,b} = D_{1,u,0,b}D_{2,0,1,b} + D_{1,u,1,b}D_{2,1,1,b} + D_{1,u,2,b}D_{2,2,1,b} + D_{1,u,3,b}D_{2,3,1,b}$

(c) $D_{0,u,2,b} = D_{1,u,0,b}D_{2,0,2,b} + D_{1,u,1,b}D_{2,1,2,b} + D_{1,u,2,b}D_{2,2,2,b} + D_{1,u,3,b}D_{2,3,2,b}$

(d) $D_{0,u,3,b} = D_{1,u,0,b}D_{2,0,3,b} + D_{1,u,1,b}D_{2,1,3,b} + D_{1,u,2,b}D_{2,2,3,b} + D_{1,u,3,b}D_{2,3,3,b}$

(e) And let a Matrix Element of this Tensor be referenced as $D_{0,b}$.

15. And let $D_{0,t}$ be the Direct Matrix Sum of all $D_{0,b}$, such that an element of $D_{0,u,v,t}$ is equal to:

$$D_{0,u,v,t} = \sum_{b=0}^{b=t-1} D_{0,u,v,b}$$

16. And that the entries of $D_{1,t}$ and $D_{2,t}$ remain the empty set, $\varnothing$, never to be used (you can technically set it to anything you want since it not used, so set default the value to zero upon Tensor Creation).

17. Let R be a two dimensional tensor that stores the product of $D_{2,b}\vec{z}_t$. And let an element of this tensor be $R_{u,b}$, with $0 \leq u \leq 3$ and $0 \leq b \leq t-1$, such that:

   (a) $R_{u,b} = D_{2,u,0,b}Z_{0,b} + D_{2,u,1,b}Z_{1,b} + D_{2,u,2,b}Z_{2,b} + D_{2,u,3,b}Z_{3,b}$

   (b)   *And let a Column Matrix Element of this Tensor be referenced as R_b.*

18.   *And let R_t be the Direct Matrix Sum of all R_b, such that an element of $R_{u,t}$ is equal to:*

$$R_{u,t} = \sum_{b=0}^{b=t-1} R_{u,b}$$

19.   *Finally let G be the inverse matrix of $D_{0,t}$, such that each element of $G_{u,v}$ is equal to each entry of $D_{0,t}^{-1}$.*

20.   *Then each component of the initially unknown $\vec{c}$ is given by each entry (respectively) of the product G and the column vector of R_t, such that:*

$$c_u = G_{u,0}Z_{0,t} + G_{u,1}Z_{1,t} + G_{u,2}Z_{2,t} + G_{u,3}Z_{3,t}$$

21.   *The Residual Sum of Squares is a real number equal to $\sum_{b=0}^{b=t-1} \vec{g}_b \vec{g}_b^*$, where $\vec{g}_t = \vec{z}_t - \vec{x}_t \vec{c} \vec{y}_t$.*

22.   *The Total Sum of Squares is a real number equal to $\sum_{b=0}^{b=t-1} \vec{h}_b \vec{h}_b^*$, where $\vec{h}_t = \vec{z}_t - \vec{f}$, where $\vec{f} = \frac{1}{t}\sum_{b=0}^{b=t-1} \vec{z}_b$.*

23. $R^2 = 1 - \frac{RSS}{TSS}$ *and* $\Psi = \ln \frac{RSS}{TSS}$.

   And this completes the all of the univariate cases for the Quaternions. For Bilinear instances of Octonions and other non-associative logics, the order of left and right handed multiplications of x_t^* and y_t^* is opposite the specified order of $\vec{x}_t$ and y_t against $\vec{c}$.

   This is because $(\vec{x}\vec{y})^* = \vec{y}^*\vec{x}^*$ for all hypercomplex logics, for this guarantees the product of two matrices that perceive the other as the adjugate matrix. By extension, this also applies to a middle-handed case of the octonions or other non-associative logics:

$$(\vec{x}_t\vec{c})\vec{y} \implies \vec{x}_t^*(((\vec{x}_t\vec{c})\vec{y}_t)\vec{y}_t^*)\vec{c} = [\vec{x}_t^*]_{8\circledcirc L}\left([\vec{y}_t^*]_{8\circledcirc R}[\vec{y}_t]_{8\circledcirc R}\right)[\vec{x}_t]_{8\circledcirc L}\vec{c} = [\vec{x}_t^*]_{8\circledcirc L}(real_y)[\vec{x}_t]_{8\circledcirc L} = real_y real_x \vec{c}$$

Q.E.D

   There was a time when I referred to D as the Divine Matrix (the product of left- and right-handed matrices!) and called the middle-handed case the "Divine Chirality." In many ways, I still view this construct as a divine solution to several unresolved questions. For instance: Where is all the antimatter? Could there be a middle-handed particle that interacts with matter from the left and with antimatter from the right? Is this the key to understanding the sterile neutrino? Are neutrinos and anti-neutrinos middle-handed, such that either interacts with matter in the same chirality?

5.2. A Constant is Variable and Rulers have Density and the Physical Interpretation of a Derivative

   "Please, no more philosophy." Sorry, but this is necessary for you to understand how to include a constant.

Definition 12. *Ruler and Ruler Density*

   *Let $f(\vec{x})$ be the $\vec{f}$ **Ruler** with m axial hypercomplex components.*

   Regardless of the number of hypercomplex components, that ruler has two indices:

1.   **Ruler Inputs:** $\vec{x}_t$. *The vector difference $\vec{x}_u - \vec{x}_v$ is the same as the distance between the inputs on the Ruler's Continuum.*

2.   **Ruler Output:** $f(\vec{x})$. *The vector difference $\vec{x}_u - \vec{x}_v$ is the distance between outputs on the ruler, regardless of the difference between the outputs of $f(\vec{x}_u)$ and $f(\vec{x}_v)$.*

3.   *This is the same definition as a Slide Ruler, but in multiple dimensions. Imagine a thin double-sided plate for $\vec{x} \in \mathbb{C}$ that reveals the opposite side $f(\vec{x}_t)$ when you press on the visible side of $\vec{x}_t$. In short, these are Continuous Vector Maps.*

4. **Real-valued Rulers** *exist solely in one direction. For $f(x) = x^2$, there are two parallel lists: $1, 2, 3, 4, \ldots, n$ and $1, 4, 9, 16, \ldots, n^2$. The distance between consecutive integer squares on the ruler is still equal to 1, no matter the size of n.*

5. **Complex-valued Rulers** *exist in two directions. Looking in only one unique complex direction on the ruler, one could indeed graph $f(\vec{x}) = \vec{x}^2$ alongside it with a color gradient for the input and output vectors, yielding some quadratic swirl (in this particular case). However, the true form is a two-sided map, with the input vector on the front side and the output vector on the bottom side.*

6. **Hypercomplex-valued Rulers** *extend this idea to more than two dimensions. The input exists on the "front side of space" and the output on the "back side of space." For three dimensions, you can envision stacks of two-sided parallel planes.*

7. *For four dimensions and beyond, you can envision an entire 3D palette of 2D doubled-sided planes as the first entry of the fourth dimension, a second palette as the second entry in the fourth dimension, a third palette (all palettes in the same straight horizontal line) as the third entry in the fourth dimension. Then, for the fifth dimension, you translate these palettes laterally, and for the sixth dimension, you stack these palettes vertically. Then call this a super-6D palette, and begin stacking them again in the same three dimensions to express the seventh, eighth, and ninth dimensions, continuing this process until you have no dimensions left to express.*

*The **Density** of the ruler equals $f'(\vec{x})$ such that the ruler itself is given by*

$$\int_{\vec{x}=\vec{0}}^{\vec{x}=\vec{c}} f'(\vec{x}) d\vec{x} = \vec{0} + f(\vec{c}) - f\left(\vec{0}\right)$$

where the integral is taken over the straight-line trajectory from $\vec{x} = \vec{0}$ to $\vec{x} = \vec{c}$. For a real number, $f(x) = x^2 + 4x - 3$, successive values $f(x_1 = y)$ and $f(x_2 = z)$ are spaced exactly $z - y$ apart upon the ruler of $f(x)$, regardless of the actual difference between $\left((z^2 + 4z - 3) - (y^2 + 4y - 3)\right)$.

In short, take the word "function" and replace it with "ruler." I only use the letter f for the sake of tradition, but had I invented the notation centuries prior, it would have been $r(\vec{x})$. You can also ditch the word "derivative" and replace it with "ruler density."

Definition 13. *Bilinear Normalized Ruler and Normalized Ruler Density*
    *Let $f(\vec{x}, \vec{y})$ be the $\vec{f}$ **Ruler** with m axial hypercomplex components.*
    Regardless of the number of hypercomplex components, that ruler has three indices:

1. **Primary Input**: *$\vec{x}_b$. The vector difference $\vec{x}_u - \vec{x}_v$ represents the distance between the primary inputs on the Ruler's Continuum.*

2. **Secondary Input**: *$\vec{y}_b$. The vector difference $\vec{y}_u - \vec{y}_v$ is the normalized distance relative to the primary input distance on the Ruler's Continuum, such that:*

$$(\vec{y}_u - \vec{y}_v)_{rulerdistance} = \frac{\vec{y}_u}{\vec{x}_u}\vec{x}_u - \frac{\vec{y}_v}{\vec{x}_v}\vec{x}_v = \vec{x}_u - \vec{x}_v$$

since $\vec{y}_u$ and $\vec{y}_v$ map to the same continuum indices as $\vec{x}_u$ and $\vec{x}_v$, respectively.

3. *If $\vec{x} = \vec{0}$, then the Ruler of X is normalized to $\vec{y}$ instead. If they are both $\vec{0}$, then they equal, which means they don't need to be normalized, such that $f(0,0)$ can be computed as is.*

4. **Ruler Output:** *$f(\vec{x}, \vec{y})$. The vector difference $\vec{x}_u - \vec{x}_v$ determines the distance between outputs on the ruler, irrespective of the actual output values of $f(\vec{x}_u)$ and $f(\vec{x}_v)$.*

5. *This definition extends the concept of a Slide Ruler to multiple dimensions, with auxiliary input rulers dynamically scaled to the primary input ruler. Imagine a thin double-sided plate for $\vec{x}, \vec{y} \in \mathbb{C}$, where pressing on the visible side of $\vec{x}_b$ reveals the opposite side $f(\vec{x}_b, \vec{y}_b)$. In essence, these are Continuous Vector Maps.*

6. ***Real-valued Bilinear Rulers*** *exist solely in one direction. For $f(x,y) = x^2 + xy + y^2$, there are three parallel lists:*

$$x_1, x_2, x_3, \ldots, x_t, \quad y_1, y_2, y_3, \ldots, y_t$$

where y_m scales as $\frac{y_m}{x_m}$ to ensure alignment with x_m on the X ruler. The computed outputs are:

$$\left(x_1^2 + x_1 y_1 + y_1^2\right), \left(x_2^2 + x_2 y_2 + y_2^2\right), \ldots, \left(x_t^2 + x_t y_t + y_t^2\right).$$

The distance between consecutive integer increments of x remains 1, independent of $f(x,y)$ or the value of input y, such that a Ruler List of $\vec{x}_t$ always has the corresponding list of $\vec{y}_t$ below it and slide-rulers output on the reverse side (a multi-valued function has multiple output lists, such that the relationship between neighboring entities is continue in the same output list and the other output list!).

7. *That the term "list" is not cheating the idea of Ruler and replacing it with a function. The ruler of T is the datalists themselves (from which we get our data for least squares!). T is an **Index** of discrete ticks with finite length, t, with lists of x_b and y_b on its opposite side (and each tick of the index is $\vec{q} = 1$)! That is, using traditional terminology, x_b and y_b are themselves functions of T from $0 \leq b \leq t$, which then invoke the continuum rulers. Hence why the mathematical domains of quantile analysis and PCA analysis are so different, even though both branches are ultimately analyzing the same data! The Quantile Ruler is not the same Ruler as the x, y inputs and z outputs!*

8. *The Index Ruler: T_0 is a list of consecutive integers from 0 to $t-1$, existing only in the forward direction. This temporal order of the data points as they were measured. Any other ordering of this list must be expressed as an Ordinal Ruler $t_{1,b} = f(t_{0,b})$, and there must always exist a bijection between T_0 and T_{new}, and that if all elements of T_0 are exhausted (which exhausts all remaining ordinal rulers by default), that $T_0 = \varnothing$, which is distinctly different from $\vec{0}$, which is the position of the observer.*

9. *The Quantile Ruler: Q is a list of consecutive integer multiples of $\frac{1}{t}$ from $\vec{q}_0 = \frac{1}{t}\vec{q}$ to $\vec{q}_{t-1}\vec{q} = \frac{t}{t}\vec{q}$, existing only in the forward direction. Hypercomplex Index Rulers and Quantiles Rulers do indeed exist, but they are beyond the scope of this publication.*

10. *Example of a real y ruler scaling: Given $x = 5$ and $y = 15$, the y ruler must be scaled by $\frac{1}{3}$ so that $y = 15$ aligns with $x = 5$. This allows one to read:*

$$f(x = 5, y = 15) = x^2 + xy + y^2$$

on the ruler, which transforms into:

$$f(x = 5, 3x = 15) = 5^2 + 5(3(5)) + (3(5))^2 = x^2 + x\frac{y}{x}x + \left(\frac{y}{x}x\right)^2.$$

11. ***Complex-valued Bilinear Rulers*** *extend in two directions. The y ruler both scales and rotates by $\frac{\vec{x}}{\vec{y}}$, ensuring a dynamically oriented bijection between $\vec{y}$ inputs and the complex continuum of $\vec{x}$.*

12. ***Hypercomplex-valued Bilinear Rulers*** *generalize this concept beyond two dimensions, incorporating non-commutative and non-associative structures while preserving the chirality of the ruler. For example, in:*

$$f(\vec{x}, \vec{y}) = \vec{x}^2 + \vec{x}\vec{y} + \vec{y}^2$$

the $\vec{y}$ ruler scales as:

$$f(\vec{x}, \vec{y}) = \vec{x}^2 + \vec{x}\left(\vec{x}\frac{1}{\vec{x}}\vec{y}\right) + \left(\vec{x}\frac{1}{\vec{x}}\vec{y}\right)^2.$$

However, for:

$$f(\vec{x}, \vec{y}) = \vec{x}^2 + \vec{y}\vec{x} + \vec{y}^2,$$

the $\vec{y}$ ruler scales differently:

$$f(\vec{x}, \vec{y}) = \vec{x}^2 + \vec{y}\frac{1}{\vec{x}}\vec{x}\vec{x} + \left(\vec{y}\frac{1}{\vec{x}}\vec{x}\right)^2.$$

13. *Observe that $\vec{y}\frac{1}{\vec{x}}\vec{x}$ is the reverse of $\vec{x}\frac{1}{\vec{x}}\vec{y}$. However, this order only affects the bilinear terms $\vec{x}\vec{y}$ in the former and $\vec{y}\vec{x}$ in the latter. It does not impact the substitution for $\vec{y}^2$ in either ruler.*

5.2.1. The Physical Mechanism which resolves $\vec{z}_t = \vec{x}_t\vec{c}\vec{y}_t$

From the above definition we can understand the physical mechanics of the bilinear regression $\vec{z}_t = \vec{x}_t\vec{c}\vec{y}_t$ which confounded the physicist that initially tried to tackle this problem in 2008. Allow me to explain how I initially dealt with the problem:

1. Microsoft Excel is the digital incarnation of Ruler Space.
2. Let T be the set of data points from $b = 0$ to $b = t - 1$ written in an Excel column.
3. Let the vectors of $\vec{x}_b$ and $\vec{y}_b$ be appended column-wise, using as many columns per vector as each vector has components.
4. Let the set of columns expressing $\vec{x}_t$ be the Ruler of X, and let the Ruler Y "be seen" from the perspective of X, such that $f(\vec{x}) = \vec{w}_b = \vec{y}_b\frac{1}{\vec{x}_b}$. Remember earlier in the first chapter of the paper when I defined division as change of reference frame! We now have the F ruler.
5. Then $g(\vec{x}_b) = f(\vec{x}_b)\vec{x}_b$. This is the G Ruler, which is ultimately attuned to the T Index Ruler, is it not!?
6. Then $\vec{z}_b = h(\vec{x}_b) = \vec{x}_b\vec{c}f(\vec{x}_b)\vec{x}_b = \vec{x}_b\vec{c}\vec{w}_b\vec{x}_b$. This is the H ruler. This is when I realized that if vectors can be inputs and outputs on a ruler, then so can matrices!
7. Let redefine $h(\vec{x}_b) = [\vec{x}_b]_{4\mathbb{H}L}[\vec{x}_b]_{4\mathbb{H}R}[\vec{w}_b]_{4\mathbb{H}L}\vec{c}$. This is the actual H ruler.
8. Now let's simply into bilinear form:

$$h(\vec{x}_b) = [\vec{x}_b]_{4\mathbb{H}L}[\vec{y}_b]_{4\mathbb{H}R}\vec{c}$$

9. The problem is that $\vec{c}$ is unknown! So we need one last ruler (and it's conjugate), the Ruler of Divine Chirality and its Conjugate Ruler (Adjugate Matrix)!

$$D(\vec{x}_b) = D_b = [\vec{x}_b]_{4\mathbb{H}L}[\vec{y}_b]_{4\mathbb{H}R}$$

$$D^*(\vec{x}_b) = D_b^* = [\vec{y}_b]^*_{4\mathbb{H}R}[\vec{x}_b]^*_{4\mathbb{H}L}$$

10. Then clearly

$$\vec{c} = \left(\sum_{b=0}^{b=t-1} D^*D\right)^{-1}\left(\sum_{b=0}^{b=t-1} D^*\vec{z}\right)$$

11. How does one measure the R^2 of this regression? For this we need the Quantile Ruler from $\frac{1}{t}$ to $\frac{t}{t}$ to yield the Total Sum of Squares: $\bar{\vec{z}} = \frac{1}{t}\sum_{b=0}^{b=t-1}\vec{z}_t$, which states that $\vec{z}_t$ is the Riemann Sum of $\vec{z}_t$ over the indices of T normalized from zero to one (a series of 5D rectangular prisms added together, all having the same width of $\frac{1}{t}$). From which is follows that $TSS = \sum_{b=1}^{b=t-1}(\vec{u}_b\vec{u}_b^*)$, where $\vec{u}_b = (\vec{z}_t - \bar{\vec{z}})$.
12. From which it also follows that $RSS = \sum_{b=0}^{b=t-1}\vec{v}_b\vec{v}_b^*$, where $\vec{v}_b = (\vec{z}_t - D_b\vec{c})$, since $D_b\vec{c}$ is a column vector output, as is $\vec{v}_b$ (naming variable types, such as "integer", "real", "column vector" or "matrix" in C++, Mathlab and/or in Excel's column headers really helps!).
13. And therewithin Excel was the physical incarnation of Ruler Space, was the closed form solution to the regression of $\vec{z}_t = \vec{x}_t\vec{c}\vec{y}_t$, along with its real numbered R^2 value. All I required was a bunch of rulers. Some with only one direction (T and Q) and some with four directions X, Y, Z, and some as 4x4 tables to mimic a single spot on a ruler (D and $D*$). This is because Excel is a pre-built compiler for ruler space!

The confusion in LU Decomposition with Full Pivoting arises when people attempt to unify disjoint ruler spaces — Quantile and Continuous — without properly managing their distinctions. The challenge lies in handling PCA components within the matrix while simultaneously remapping Ordinals during row and column swaps.

For an Excel user like me, this isn't an issue— I intuitively track these transformations. But for someone who programs exclusively in C++ or Python, the process often becomes unclear. In practice, they usually reach out to someone like me, ask for a spreadsheet version, and then translate it into formal code in their own language.

This is precisely why I don't see a fundamental problem with General Relativity lacking unification with quantum mechanics. GR is PCA; quantum is quantile. The universe doesn't force them into an artificial union — it naturally separates them into two disjoint problem spaces, each governed by its own "Wolfram Ruliad" (Ruler Set). When a problem involves both domains, the universe simply takes their union, much like how LU decomposition with full pivoting handles mixed spaces by unifying ruler sets as needed.

As for the Density of a Bilinear Ruler, it's the derivative of $f(\vec{x}, \vec{y})$, with $\vec{y}$ normalized to $\vec{x}$ via rotation and scaling. Explaining this way makes it far easier to explain to students what it means to take a derivative of a multi-valued input (and I'd argue, the only natural way).

Of course, I've had people all besiege me with the same counter-example. They draw a quadratic manifold $z = x^2 + xy + y^2$ with $x, y \in \mathbb{R}$ in 3D space and insist that the derivative of the function cannot be explained with rulers (this because they think that x, y, z exist in a 3D space, which we already debunked earlier).

I then point to a spot on the 3D manifold and ask them: So what's the derivative at this exact spot? They then answer: "Well, it depends upon the direction."

I then say, "Ok, declare a direction for me." They almost always go along the direction of $y = x$ at 45 degrees.

I then say, "Then how is that any different from $f(x, y) = f(x, x) = x^2 + x^2 + x^2 = 3x^2$?" They then say it isn't any different, because y is now normalized as a function of x. In other words, they just normalized the damn ruler.

They tend to get angry and use a different direction (let's say 60 degrees for the sake of the argument). Then $y = x \tan \frac{\pi}{3} = mx$. I then say:

"Then how is that any different from $x^2 + mx^2 + m^2x^2 = (1 + m + m^2)x^2$?" They again admit, "It's not...," and then the realization hits them, "because this locus of y was normalized to x."

And this when they realize that both x and y and z all exist in the same one-dimensional space, and all we're doing is the expansion or contraction of the Y ruler by $\frac{y}{x}$, and that z is just a slide rule output, then can't care less about the normalized that was applied.

Of course the exceptionally clever mathematician tries one more approach. They draw the surface again in 3D, point to a location in the x,y plane that's at some angle θ and then tell me to find the derivative at that height of the manifold in some other direction ϕ. All I do is take their $x, y = m, n$ coordinate, subtract m from the x ruler and n from the y ruler and rewrite $f(x, y)$ as $g(x, y) = f(x - m, y - n)$. Problem solved. All we did was refine the position of the Observer, and therefore the position of $\vec{0}$, which is exactly how we're going to add a constant to our quaternionic regression.

The physical interpretation of a derivative, especially in the context we're discussing, is grounded in a fundamental and tangible concept: Change and the rate of change. It reflects how things actually behave or evolve in space. By explaining the derivative using rulers and physical objects, we transform an abstract mathematical idea into something intuitive and observable (and in the context of real numbers, rulers spanning the same singular dimension).

While skeptics may argue about the complexity of a function, the physical interpretation of the derivative remains unchanged. One can only "stump" it by misinterpreting or misapplying the concept itself. Grounding the concept in reality (as done with rulers and shifts in space) makes it impossible to dismiss or confuse — because the answer is always right there, waiting to be seen. The math is not simply an abstract equation; it models real-world behavior.

This is the key: By making the math something that can be seen, felt, and experienced, it turns into a tangible representation of what would otherwise be an abstract concept. When skeptics try to

argue against the physical interpretation of a derivative, they miss the fact that it is the rate of change in the real world, not something that can be tricked or twisted. You can't stump it because it's as real as the ruler you're using to measure it.

Theorem 5. *The Physical Definition of a Derivative The rate of compression between the Domain of the Ruler of X and the Range of the Ruler $f(\vec{x})$. If $\vec{y} = f(\vec{x}) = \vec{x}$, then the Rulers of X and Y exist in phase; otherwise, the ruler of Y compresses (or expands) at a rate of $f'(\vec{x})$ from the Observer wielding the Ruler of X (which in layman's terms means that the Y Ruler starts at $\vec{y} = f'(0)$ because the Observer wielding the X Ruler exists at $\vec{x} = \vec{0}$).*

Theorem 6. *The Physical Definition of Partial Derivative Let the Ruler of Y be normalized to the Ruler of X, such that $f(\vec{x}, \vec{y})$ is transformed to the Single Ruler of U through the strictly linear transformation of $\vec{u} = \vec{x} - \vec{m}$ and $\vec{v} = \frac{\vec{v} - \vec{n}}{\vec{u}}$, yielding the function $f(\vec{u}, \vec{u}\vec{v}) = g(\vec{u})$, then $f'(\vec{x}, \vec{y}) = g'(\vec{u})$.*

And this this holds for any arbitrary number of Domains, regardless of the number of hypercomplex components for any of those domains (just remember to mind chirality when multiplying and dividing).

As for multi-linear rulers, we just normalized all of the extra inputs to one of non-zero inputs (as inferred by the end of the previous theorem). If they are all zero, then again, there's no need to normalize. At the end of the day we have a slide rule T that produces an $\vec{x}$ that produces an $f(x)$, regardless of the number of distinct x multi-linear variables.

Of course, there's some that still protest that $\mathbb{R}^n$ represents multi-dimensional space. If you're still amongst that crowd, then I ask you to consider why $\vec{z} = x\vec{q} + y\vec{i} \neq x + y$.

For example, $\vec{z} = 6\vec{q} + 2\vec{i} \neq 6 + 2 = 8$. If you really think that $\mathbb{R}^2 = \mathbb{C}$ or that $\mathbb{R}^4 = \mathbb{H}$, then why isn't $6\vec{q} + 2\vec{i}$ equal to 8?

This is why the determinant of a real number matrix has nothing to do with the volume of a parallepiped, but rather the scaling of rulers. Using the same example from earlier (re-posted verbatim): $\begin{pmatrix} +5 & +10 \\ +10 & +5 \end{pmatrix}$. We compute the determinant using the following steps (we call it a Reflector Matrix because the first row is reflection of the second row over the 45 degree line, never again shall you call it a Hyperbolic Rotation Matrix!):

1. Express the sum in hyperbolic form:

$$5\vec{q} + 10\vec{q} = e^u e^{\vec{q}\theta} = e^u(\vec{q}\cosh\theta + \vec{q}\sinh\theta).$$

2. Compute θ as: $\vec{\theta} = \operatorname{arctanh}\frac{10}{5} = 0.549306\vec{q} - \frac{\pi}{2}\vec{i}.$
3. Rewrite using hyperbolic secant:

$$5\vec{q} + 10\vec{q} = 5\operatorname{sech}\vec{\theta}(\vec{q}\cosh\vec{\theta} + \vec{q}\sinh\vec{\theta}).$$

4. Recognizing that the determinant of the unit reflector matrix

$$\begin{pmatrix} +\cosh\theta & +\sinh\theta \\ +\sinh\theta & +\cosh\theta \end{pmatrix}$$

is 1, we conclude that the determinant of

$$\begin{pmatrix} +5 & +10 \\ +10 & +5 \end{pmatrix}$$

must be

$$(e^u)^2 = e^{2u} = \ln\left(5\operatorname{sech}\vec{\theta}\right).$$

Hence, when we have a bicomplex number $\vec{z} = \left(a\vec{q} + b\vec{j} \right)$, and we multiply by $\vec{j}$, whose square is $+1 = +\vec{q}$, we get

$$\vec{z}\vec{j} = \left(b\vec{j} + a\vec{j} \right),$$

, because we are swapping rulers (which appears to be a reflection of the 45 degree line when falsely drawn in 2D). That is what $\vec{j}$ actually does, and it is why multiplication by $\vec{q}\cosh\theta + \vec{j}\sinh\theta$ preserves apparently 2D linearity in spacetime diagrams (though poorly named hyperbolic rotations because they are falsely rendered in 2D instead of 1D), because they invoke a partial swapping of rulers, both of which remain in the same straight line in the same direction, regardless of how they are phased. In fact is would be better to call the above matrix of 5 and 10 a Ruler Phase Matrix.

Theorem 7. *Least Squares Acts on the Spacing of the Rulers Given the ruler $f(\vec{x})$, the regression of*

$$\vec{z} = \vec{c}_1 f(\vec{x}) + \vec{c}_2 f(\vec{x})$$

acts upon the Ruler Space of X itself, not the Range of X.

Given $z = ax + bx^2 = c_1 f(x) + c_2 g(x)$, Least Squares Regression cannot distinguish between $g(x) = x^2$ and $g(x) = x^3$, for either variant of $g(x)$ is normalized to the ruler X (via density normalization).

In the case of $g(x) = x^2$, the Range of X is compacted, such that the sequence $1, 4, 9, 16, \ldots, x^2$ is as evenly spaced as $1, 2, 3, 4, \ldots, n$, and it is this evenly spaced tick upon which the constant c_2 acts, regardless of the density fluctuations in the Range of X.

Hence, Least Squares Regression reads the Ruler of $X = x^2$ as a new ruler $Y = g(y) = y$ in complete ignorance that $x_b^2 = y_b$, and derives the constant b, as if all $1, 4, 9, 16, \ldots, x^2$ are evenly spaced as $1, 2, 3, 4 \ldots y$.

That is, using the Pseudo-Interpretation of $\mathbb{R}^2$ being a multi-dimensional space, that given the regression of $z = c_1 x + c_2 x^2$, that one simply rewrites this as $z = c_1 x + c_2 y$, a flat plane regression, and then substitutes y out for x^2 later.

Theorem 8. *Ruler of Unity: A Constant is the Zeroth Power Ruler*

Let $f(\vec{x}_t) = \vec{x}_t^0$.

*This is the **Ruler of Unity**. It has zero density (its derivative) and evaluates to $1\vec{q}$ at zero via the limit as $\vec{x}$ goes to zero, for it's the number of ways you can choose nothing from the empty set, and its logarithm returns the position of the observer.*

Using complex numbers as an example, we have a ruler that spans the entire 2D complex continuum. When we tap upon $\vec{x} = a\vec{q} + b\vec{i}$ on the front side of the plate, it reveals the backside of the plate, and the answer is always $1 = \vec{q}$.

Thus, we add a constant to our regression as follows:

$$\vec{z} = \vec{c}_0 \vec{x}_t^0 + \vec{c}_1 \vec{x}_t^1 + \cdots + \vec{c}_n \vec{x}_t^m$$

In the bilinear case (or multi-linear case), we add:

$$\vec{z} = \vec{c}_0 f(\vec{x}, \vec{y}) + \ldots$$

with

$$f(\vec{x}, \vec{y}) = \left(\vec{x}^0 \vec{y}^0 \right),$$

*which forms the **crown at the top of Pascal's Pyramid** for binomial expansion, and more generally Pascal's Simplex for multinomial expansion.*

*Hypercomplex Least Squares Regression then scales and **rotates** the range of the ruler, which is always equal to $1\vec{q}$, to find the best offset from the origin to explain the observed data $\vec{z}$. That is, **Mother Nature finds the best location to position the Observer before actually observing the system**. This can be proven by*

the fact that Covariance Matrices subtract the means of $\vec{x}$ and $\vec{y}$ from the data list. Further support comes from the practice of programmers subtracting the means of x, y, z in 3D graphics before rotating a 3D object.

The constant term of a regression can only be explained within the Ruler Space—there is no other way to interpret it! The moment we substitute $y = x^0 = 1$ and perform a 3D flat-plane regression of $z = x + y$ no explanation other than the dual-sided Slide Rule could possibly clarify how this works, whether to a student, a professor, or any other individual. Hence, there's no way that the aforementioned example could ever be a 3D planar regression to begin with!

5.3. Univariate Least Squares, With a Constant

We're given two quaternionic lists, X (and/or Y) and Z, of length m. We are then asked find the best fit, left or right handed, such that the constant, $\vec{c}$, produces the least error.

5.3.1. Human Readable-Forms of the Univariate Case with a Constant

1. Let t be the total number of data points. $R^2 = 1 - \frac{RSS}{TSS}$ and $\Psi = \ln \frac{RSS}{TSS}$.

$$\bar{\vec{z}} = \frac{1}{t} \sum_{b=0}^{b=t-1} \vec{z}_t; \; TSS = \sum_{b=0}^{b=t-1} \left(\vec{z}_t - \bar{\vec{z}}\right)\left(\vec{z}_t - \bar{\vec{z}}\right)^*$$

2. Let C be the real number column vector of height 8 that reads the constants of $\vec{c}_0 = c_{0,0}\vec{q} + c_{0,1}\vec{i} + c_{0,2}\vec{j} + c_{0,3}\vec{k}$ and $\vec{c}_1 = c_{1,0}\vec{q} + c_{1,1}\vec{i} + \vec{c}_{0,2}\vec{j} + \vec{c}_{0,3}\vec{k}$ by their eight combined components, in the order

just listed, such that $C = \begin{pmatrix} c_{0,0} \\ c_{0,1} \\ c_{0,2} \\ c_{0,3} \\ c_{1,0} \\ c_{1,1} \\ c_{1,2} \\ c_{1,3} \end{pmatrix}$

3. The Left-Handed Regression of $\vec{z}_b = \vec{c}_0\vec{y}^0 + \vec{c}_1\vec{y}_b^1$ implies the linear of system of equations:
    (a) $\left(\vec{c}_0\vec{y}^0 + \vec{c}_1\vec{y}^1\right)\vec{y}^{0*} = \vec{z}\vec{y}^{0*}$
    (b) $\left(\vec{c}_0\vec{y}^0 + \vec{c}_1\vec{y}^1\right)\vec{y}^{1*} = \vec{z}\vec{y}^{1*}$
    (c) Which compels:

$$C = \left(\begin{pmatrix} \sum_{b=0}^{b=t-1} [\vec{y}_b^{0*}]_{4\mathbb{H}R} [\vec{y}_b^0]_{4\mathbb{H}R} \\ \sum_{b=0}^{b=t-1} [\vec{y}_b^{1*}]_{4\mathbb{H}R} [\vec{y}_b^0]_{4\mathbb{H}R} \end{pmatrix} \begin{pmatrix} \sum_{b=0}^{b=t-1} [\vec{y}_b^{0*}]_{4\mathbb{H}R} [\vec{y}_b]_{4\mathbb{H}R} \\ \sum_{b=0}^{b=t-1} [\vec{y}_b^{1*}]_{4\mathbb{H}R} [\vec{y}_b]_{4\mathbb{H}R} \end{pmatrix} \right)^{-1} \begin{pmatrix} \sum_{b=0}^{b=t-1} [\vec{y}_b^{0*}]_{4\mathbb{H}R} \vec{z}_b \\ \sum_{b=0}^{b=t-1} [\vec{y}_b^{1*}]_{4\mathbb{H}R} \vec{z}_b \end{pmatrix}$$

$$RSS = \sum_{b=0}^{b=t-1} (\vec{z}_b - \vec{c}_0 - \vec{c}_1\vec{y}_b)(\vec{z}_b - \vec{c}_0 - \vec{c}_1\vec{y}_b)^*$$

4. The Right-Handed Regression of $\vec{z}_b = \vec{x}^0\vec{c}_0 + \vec{x}_b^1\vec{c}_1$ implies the linear of system of equations:
    (a) $\vec{x}^{0*}\left(\vec{x}^0\vec{c}_0 + \vec{x}^1\vec{c}_1\right) = \vec{x}^{0*}\vec{z}$
    (b) $\vec{x}^{1*}\left(\vec{x}^0\vec{c}_0 + \vec{x}^1\vec{c}_1\right) = \vec{x}^{1*}\vec{z}$
    (c) Which compels:

$$C = \left(\begin{pmatrix} \sum_{b=0}^{b=t-1} [\vec{x}_b^{0*}]_{4\mathbb{H}L} [\vec{x}_b^0]_{4\mathbb{H}L} \\ \sum_{b=0}^{b=t-1} [\vec{x}_b^{1*}]_{4\mathbb{H}L} [\vec{x}_b^0]_{4\mathbb{H}L} \end{pmatrix} \begin{pmatrix} \sum_{b=0}^{b=t-1} [\vec{x}_b^{0*}]_{4\mathbb{H}L} [\vec{x}_b]_{4\mathbb{H}L} \\ \sum_{b=0}^{b=t-1} [\vec{x}_b^{1*}]_{4\mathbb{H}L} [\vec{x}_b]_{4\mathbb{H}L} \end{pmatrix} \right)^{-1} \begin{pmatrix} \sum_{b=0}^{b=t-1} [\vec{x}_t^{0*}]_{4\mathbb{H}R} \vec{z}_t \\ \sum_{b=0}^{b=t-1} [\vec{x}_t^{1*}]_{4\mathbb{H}R} \vec{z}_t \end{pmatrix}$$

$$RSS = \sum_{b=0}^{b=t-1} (\vec{z}_b - \vec{c}_0 - \vec{x}_b\vec{c}_1)(\vec{z}_b - \vec{c}_0 - \vec{x}_b\vec{c}_1)^*$$

5.   The Middle-Handed Regression of $\vec{z}_b = \vec{x}_b^0 \vec{c}_0 \vec{y}_b^0 + \vec{x}_b^1 \vec{c}_0 \vec{y}_b^1$ implies the linear of system of equations:

    (a) $\vec{x}^{0*}\left(\vec{x}^0 \vec{c}_0 \vec{y}^0 + \vec{x}^1 \vec{c}_0 \vec{y}^1\right)\vec{y}^{0*} = \vec{x}^{0*}\vec{z}_b \vec{y}^{0*}$

    (b) $\vec{x}^{1*}\left(\vec{x}^0 \vec{c}_0 \vec{y}^0 + \vec{x}^1 \vec{c}_0 \vec{y}^1\right)\vec{y}^{1*} = \vec{x}^{1*}\vec{z}_b \vec{y}^{1*}$

    (c)   Which compels:

$$C = \left(\begin{matrix} \left(\sum_{b=0}^{b=t-1} D_{0,0,b}^* D_{0,0,b}\right) & \left(\sum_{b=0}^{b=t-1} D_{0,0,b}^* D_{0,1,b}\right) \\ \left(\sum_{b=0}^{b=t-1} D_{1,0,b}^* D_{0,0,b}\right) & \left(\sum_{b=0}^{b=t-1} D_{1,0,b}^* D_{1,1,b}\right) \end{matrix}\right)^{-1} \left(\begin{matrix} \sum_{b=0}^{b=t-1} D_{0,0,b}^* \vec{z}_t \\ \sum_{b=0}^{b=t-1} D_{1,0,b}^* \vec{z}_t \end{matrix}\right)$$

$$RSS = \sum_{b=0}^{b=t-1} (\vec{z}_b - \vec{c}_0 - \vec{x}_b \vec{c}_1 \vec{y}_b)(\vec{z}_b - \vec{c}_0 - \vec{x}_b \vec{c}_1 \vec{y}_b)^*$$

6.   Where:

    (a) $D_{0,0,b}^* = D_{0,0,b} = I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

    (b) $D_{1,0,b}^* = \left[\vec{x}_b^*\right]_{4\mathbb{H}L}\left[\vec{y}_b^*\right]_{4\mathbb{H}R}$ and $D_{1,0,b} = \left[\vec{x}_b\right]_{4\mathbb{H}L}\left[\vec{y}_b\right]_{4\mathbb{H}R}$

There's no need to discuss programming instructions for the univariate case at this moment, as it is already encompassed within the General Case of Multivariate Least Squares of Mixed Chirality with a constant (which is the next section!).

Rather let's discuss the binomial and trinomial expansions of a quaternion and how they differ from the Pascal Simplex.

5.4. From Pascal's Simplex to Pascal's Cube

When we execute the layers of $(x + y)^n$ over the reals for x and y, we get the typical binomial expansion pattern, whose successive ranks appear as Pascal's Triangle.

However, suppose x and y are quaternions? What happens then?

1. $(x + y)^0 = \vec{q}$

2. $(x + y)^1 = x + y$

3. $(x + y)^2 = x^2 + xy + yx + y^2$

4. $(x + y)^3 = \left(x^3 + x^2 y + xyx + xy^2\right) + \left(yx^2 + yxy + y^2 x + y^3\right)$

5.   We just went from one term, to two terms, to four terms, to eight terms. Then the full bivariate cubic regression of $\vec{z}$ in terms of $\vec{x}$ and $\vec{y}$ contains a total of 15 terms:

$$\vec{z} = \vec{c}_{0,0} + (\vec{c}_{1,0}\vec{x} + \vec{c}_{1,1}\vec{y}) + \left(\vec{c}_{2,0}\vec{x}^2 + \vec{c}_{2,1}\vec{x}\vec{y} + \vec{c}_{2,2}\vec{y}\vec{x} + \vec{c}_{2,3}\vec{y}^2\right) + \dots$$

$$\dots + \left(\vec{c}_{3,0}\vec{x}^3 + \vec{c}_{3,1}\vec{x}^2 y + \vec{c}_{3,2}\vec{x}\vec{y}\vec{x} + \vec{c}_{3,3}\vec{x}\vec{y}^2 + \vec{c}_{3,4}\vec{y}\vec{x}^2 + \vec{c}_{3,5}\vec{y}\vec{x}\vec{y} + \vec{c}_{3,6}\vec{y}^2\vec{x} + \vec{c}_{3,7}\vec{y}^3\right)$$

6.   More generally, the fully bivariate regression of degree n contains $\left(-1 + 2^{n+1}\right)$ terms.

7.   Even worse all linear terms have two possible chirality ($\vec{c}_{1,0}\vec{x}$ versus $\vec{x}\vec{c}_{1,0}$), and the quadratic terms have three chiralities ($\vec{c}_{2,1}\vec{x}\vec{y}$ versus $\vec{x}\vec{c}_{2,1}\vec{y}$ versus $\vec{x}\vec{y}\vec{c}_{2,1}$) and the Cubic Terms, while having only three chiralities on paper, have four different placements of $c_{3,v}$.

8.   Thus the number of chiral permutations that needs to be executed to find the highest R^2 is precisely $\left(1^1\right)\left(2^2\right)\left(3^4\right)\left(4^8\right)\dots(n+1)^{2^n}$.

9.   For the bivariate quadratic introduced at the start of this paper, we have a total of 324 permutations since $324 = \left(1^1\right)\left(2^2\right)\left(3^4\right)$

10.   Now suppose this was trivariate. We'd go from Pascal's Pyramid over the reals and complex numbers, to Pascal's Cube, containing $1 + 3^n$ non-commutative polynomial terms. The number of permutations is now: $\left(1^1\right)\left(2^3\right)\left(3^9\right)\left(4^{27}\right)\dots(n+1)^{3^n}$.

Clearly, we need a system that reduces the number of permutations, regardless of the degree of the expansion or the number of input variables prior to the expansion.

The answer is quite simple: Each unique term is its own ruler. The moment an addition sign appears between $x + y$, even if $y = x^2$, it constitutes a unique term as far as Mother Nature is concerned. This implies that the contribution of y is independent of the contribution of x.

For example $\vec{z} = \vec{c}_0 + \vec{x}\vec{c}_1$ produces the higher R^2 than $\vec{z} = \vec{c}_0 + \vec{c}_1\vec{x}$, which means that the former is always the best, no matter how many additional terms we add to the regression.

When we proceed to $\vec{z} = \vec{c}_0 + \vec{x}\vec{c}_1 + \vec{y}$, no matter which chirality we assign to $\vec{c}_2$, the R^2 of the overall regression either remains the same as the previous, or improves, and whichever term ($\vec{c}_2\vec{y}$ or $\vec{y}\vec{c}_2$) performs better, is the answer.

Now let's consider the bivariate quadratic case:

$$\vec{z} = \vec{c}_{0,0} + (\vec{c}_{1,0}\vec{x} + \vec{c}_{1,1}\vec{y}) + \left(\vec{c}_{2,0}\vec{x}^2 + \vec{c}_{2,1}\vec{x}\vec{y} + \vec{c}_{2,2}\vec{y}\vec{x} + \vec{c}_{2,3}\vec{y}^2\right)$$

We only need to test $\sum_{h=0}^{h=n}(h+1)2^h = 1 + n2^n$ permutations, which is far better than $\prod_{h=0}^{h=n}(h+1)2^h$. Yes, $1 + n2^n$ grows exponentially, but it's certainly far better than product which evaluates to $\left(2^{\frac{n^2+n}{2}}\right)\Gamma(n+2)$, which is probably the fastest growing thing prior to Tetration.

The combinatorial solution space to the bivariate quadratic remains 324, but we only require 17 checks (technically 16 since the constant vector is a the real number 1).

1. $\vec{x}\vec{c}$ versus $\vec{c}\vec{x}$ is two checks.
2. $\vec{y}\vec{c}$ versus $\vec{c}\vec{y}$ is two checks.
3. $\vec{x}^2\vec{c}$ vs $\vec{x}\vec{c}\vec{x}$ vs $\vec{c}\vec{x}^2$ is three checks.
4. $\vec{x}\vec{y}\vec{c}$ vs $\vec{x}\vec{c}\vec{y}$ vs $\vec{c}\vec{x}\vec{y}$ is three checks.
5. $\vec{y}\vec{x}\vec{c}$ vs $\vec{y}\vec{c}\vec{x}$ vs $\vec{c}\vec{y}\vec{x}$ is three checks.
6. $\vec{y}^2\vec{c}$ vs $\vec{y}\vec{c}\vec{y}$ vs $\vec{c}\vec{y}^2$ is three checks.
7. A total of 12+4=16 checks.

For the non-associative logics, like the octonions, the number of permutations per term grows individually, but since each term is independent of one another, the growth is still checked. Let's see that the bivariate quadratic looks like for the octonions (this also covers all non-associative logics to the sedenions and beyond).

1. $\vec{x}\vec{c}$ versus $\vec{c}\vec{x}$ is two checks. For all logics, a left matrix times a vector is equal to the the right matrix of the vector times the column vector of the original matrix.
2. $\vec{y}\vec{c}$ versus $\vec{c}\vec{y}$ is two checks. or all logics, a left matrix times a vector is equal to the the right matrix of the vector times the column vector of the original matrix.
3. $\vec{x}^2\vec{c}$ has the six possible forms:

   (a) $(\vec{x}\vec{x})\vec{c}$ vs $\vec{x}(\vec{x}\vec{c})$
   (b) $(\vec{x}\vec{c})\vec{x}$ vs $\vec{x}(\vec{c}\vec{x})$
   (c) $(\vec{c}\vec{x})\vec{x}$ vs $\vec{c}(\vec{x}\vec{x})$

4. $\vec{x}\vec{y}\vec{c}$ has the six possible forms:

   (a) $(\vec{x}\vec{y})\vec{c}$ vs $\vec{x}(\vec{y}\vec{c})$
   (b) $(\vec{x}\vec{c})\vec{y}$ vs $\vec{x}(\vec{c}\vec{y})$
   (c) $(\vec{c}\vec{x})\vec{y}$ vs $\vec{c}(\vec{x}\vec{y})$

5. $\vec{y}\vec{x}\vec{c}$ has the six possible forms:

   (a) $(\vec{y}\vec{x})\vec{c}$ vs $\vec{y}(\vec{x}\vec{c})$
   (b) $(\vec{y}\vec{c})\vec{x}$ vs $\vec{y}(\vec{c}\vec{x})$
   (c) $(\vec{c}\vec{y})\vec{x}$ vs $\vec{c}(\vec{y}\vec{x})$

6. $\vec{y}^2\vec{c}$ has the six possible forms:

   (a) $(\vec{y}\vec{y})\vec{c}$ vs $\vec{y}(\vec{y}\vec{c})$

    (b) $(\vec{y}\vec{c})\vec{y}$ vs $\vec{y}(\vec{c}\vec{y})$

    (c) $(\vec{c}\vec{y})\vec{x}$ vs $\vec{c}(\vec{y}\vec{y})$

7. This results in a total of $24 + 4 = 28$ checks against the theoretical $5184 = 2^2 6^4$ permutations (which is not significantly worse than the 16 checks for the associative logics). However, if we extend this to a degree-three cubic or a three-term trinomial, the complexity escalates rapidly. Fortunately, most natural phenomena follow inverse-square laws governing interactions between two distinct entities, meaning the bivariate quadratic suffices for the vast majority of laboratory experiments.

5.4.1. The Least Squares Chirality

Nevertheless, I have not been able to formally prove that a discarded chirality for a previous term remains forever dormant upon the addition of new terms. Rather, I only know it to be true is regards to the Real and Complex Numbers (that each term has no effect on the veracity of the others) and through millions of simulations for the quaternions.

Definition 14. *The Chirality Conjecture That given the Hypercomplex Least Squares Regression of $\vec{z} = \vec{c}_1\vec{x}$ and $\vec{z} = \vec{x}\vec{c}_1$ yielding $R^2 = a_1$ and $R^2 = a_2$, respectively:*

1. *If $a_1 > a_2$, then let:*

    (a) b_1 be the R^2 of $\vec{z} = \vec{c}_1\vec{x} + \vec{c}_2\vec{y}$

    (b) b_2 be the R^2 of $\vec{z} = \vec{c}_1\vec{x} + \vec{y}\vec{c}_2$

    (c) b_3 be the R^2 of $\vec{z} = \vec{x}\vec{c}_1 + \vec{c}_2\vec{y}$

    (d) b_4 be the R^2 of $\vec{z} = \vec{x}\vec{c}_1 + \vec{y}\vec{c}_2$

2. *That both b_1 or b_2 is greater than or equal to a_1, and let the great of them be b_5 (this statement is automatically true).*

3. *That both b_3 or b_4 is greater than or equal to a_2, and let the great of them be b_6 (this statement is automatically true).*

4. *Such that $b_6 \leq b_5$ (this is the conjecture!), which implies that left-handed $\vec{c}_1\vec{x}$ always outperforms the right-handed $\vec{x}\vec{c}_1$, regardless of the contribution of $\vec{y}$.*

1. *If $a_2 > a_1$, then let:*

    (a) b_1 be the R^2 of $\vec{z} = \vec{c}_1\vec{x} + \vec{c}_2\vec{y}$

    (b) b_2 be the R^2 of $\vec{z} = \vec{c}_1\vec{x} + \vec{y}\vec{c}_2$

    (c) b_3 be the R^2 of $\vec{z} = \vec{x}\vec{c}_1 + \vec{c}_2\vec{y}$

    (d) b_4 be the R^2 of $\vec{z} = \vec{x}\vec{c}_1 + \vec{y}\vec{c}_2$

2. *That both b_1 or b_2 is greater than or equal to a_1, and let the great of them be b_5 (this statement is automatically true).*

3. *That both b_3 or b_4 is greater than or equal to a_2, and let the great of them be b_6 (this statement is automatically true).*

4. *Such that $b_5 \leq b_6$ (this is the conjecture!), which implies that right-handed $\vec{x}\vec{c}_1$ always outperforms the left-handed $\vec{c}_1\vec{x}$, regardless of the contribution of $\vec{y}$.*

5. *That if this Conjecture is proven true, it means that the placement of $\vec{c}$ for any and all additional terms can be checked independently of the other terms, reducing the number of required checks from factorial time to exponential time.*

6. *That a partial solution for associative logics is acceptable.*

7. *That a full solution for non-associative logics is desired.*

8. *And that it matters not if this conjecture is proven to be true or untrue, for the Closed Form Solution to Hypercomplex Least Squares only acts upon a particular permutation, and yields the best fit C fit to the data for the given permutation.*

6. General Closed Form Solution to Multivariate Quaternionic Least Squares of Mixed Chirality

6.1. Human Readable Version

Let T be the Index Ruler from $t_b = 0$ to $t_b = t - 1$ such that $b = t_b$, where t is the total number of Data Points and b is the Data Point Index.

Let Q be the Quantile Ruler from $q_0 = \frac{1}{t}$ to $q_{t-1} = \frac{t}{t}$ such that $q_b = \frac{b+1}{t}$

We are given the Quaternionic List Z, such that an element of $\vec{z}_b$ is the column vector $\vec{z} = \begin{pmatrix} +z_{b,0}\vec{q} \\ +z_{b,1}\vec{i} \\ +z_{b,2}\vec{j} \\ +z_{b,3}\vec{k} \end{pmatrix}$, which means our T ruler repeats the index b four times per data point, in the form

$$
T, \vec{z} = \begin{pmatrix}
0 & z_{0,0} \\
0 & z_{0,1} \\
0 & z_{0,2} \\
0 & z_{0,3} \\
1 & z_{1,0} \\
1 & z_{1,1} \\
1 & z_{1,2} \\
1 & z_{1,3} \\
2 & z_{2,0} \\
2 & z_{2,1} \\
2 & z_{2,2} \\
etc & etc
\end{pmatrix}
$$

And that Z shall be measured by the Rulers of X_m and Y_n, which are composite rulers of some Multinomial Expansion of the Original p distinct inputs

$$
\sum_{g=0}^{g=h} \left(\vec{w}_{1,b} + \vec{w}_{2,b} + ...\vec{w}_{p,b} \right)^g
$$

of degree h resulting in a total of γ hypercomplex terms.

The Left and Right Chain Lemmas state that any $\vec{a}_1 \vec{a}_2 ... \vec{a}_n \vec{c}_\beta \vec{b}_n \vec{b}_{n-1} ... \vec{b}_1$ term of the multinomial $\vec{w}$ expansion can be collapsed to $\vec{x}_p \vec{c}_\beta \vec{y}_p$ for all $0 \le \beta \le \gamma - 1$ due to quaternionic associativity (this does not hold for non-associative logics!).

Even when a permutation places $\vec{c}_\beta$ strictly on the left or right, such as $\vec{w}_{a_1}^{b_1} \vec{w}_{a_2}^{b_2} \vec{w}_{a_3}^{b_3} \vec{c}_\beta$, we still treat this as the middle-handed case of $\vec{x}_\beta \vec{c}_\beta \vec{y}_\beta$ with $\vec{x}_\beta = \vec{w}_{a_1}^{b_1} \vec{w}_{a_2}^{b_2} \vec{w}_{a_3}^{b_3}$ and $\vec{y}_\beta = \vec{q}$ (that is, we set the missing side to the Ruler of Unity!). This yields the Quaternionic Lists X_β and Y_β, such that an element of $\vec{x}_{b,\beta}$ and $\vec{y}_{b,\beta}$ are the column vectors

$$
\vec{x}_\beta = \begin{pmatrix} +x_{b,0}\vec{q} \\ +x_{b,1}\vec{i} \\ +x_{b,2}\vec{j} \\ +x_{b,3}\vec{k} \end{pmatrix} ; \vec{y}_\beta = \begin{pmatrix} +x_{b,0}\vec{q} \\ +x_{b,1}\vec{i} \\ +x_{b,2}\vec{j} \\ +x_{b,3}\vec{k} \end{pmatrix} , \forall \beta : 0 \le \beta \le \gamma - 1,
$$

which expands our T ruler to:

$$
T, \vec{z}, \vec{x}_0...\vec{x}_{\gamma-1}, \vec{y}_0...\vec{y}_{\gamma-1} =
\begin{pmatrix}
0 & z_{0,0} & x_{0,0,0}...x_{0,\gamma-1,0} & y_{0,0,0}...y_{0,\gamma-1,0} \\
0 & z_{0,1} & x_{0,0,1}...x_{0,\gamma-1,1} & y_{0,0,1}...y_{0,\gamma-1,1} \\
0 & z_{0,2} & x_{0,0,2}...x_{0,\gamma-1,2} & y_{0,0,2}...y_{0,\gamma-1,2} \\
0 & z_{0,3} & x_{0,0,3}...x_{0,\gamma-1,3} & y_{0,0,3}...y_{0,\gamma-1,3} \\
1 & z_{1,0} & x_{1,0,0}...x_{1,\gamma-1,0} & y_{1,0,0}...y_{1,\gamma-1,0} \\
1 & z_{1,1} & x_{1,0,1}...x_{1,\gamma-1,1} & y_{1,0,1}...y_{1,\gamma-1,1} \\
1 & z_{1,2} & x_{1,0,2}...x_{1,\gamma-1,2} & y_{1,0,2}...y_{1,\gamma-1,2} \\
1 & z_{1,3} & x_{1,0,3}...x_{1,\gamma-1,3} & y_{1,0,3}...y_{1,\gamma-1,3} \\
2 & z_{2,0} & x_{2,0,0}...x_{2,\gamma-1,0} & y_{2,0,0}...y_{2,\gamma-1,0} \\
2 & z_{2,1} & x_{2,0,1}...x_{2,\gamma-1,1} & y_{2,0,1}...y_{2,\gamma-1,1} \\
2 & z_{2,2} & x_{2,0,2}...x_{2,\gamma-1,2} & y_{2,0,2}...y_{2,\gamma-1,2} \\
2 & z_{2,3} & x_{2,0,3}...x_{2,\gamma-1,3} & y_{2,0,3}...y_{2,\gamma-1,3} \\
3 & z_{3,0} & x_{3,0,0}...x_{3,\gamma-1,0} & y_{3,0,0}...y_{3,\gamma-1,0} \\
3 & z_{3,1} & x_{3,0,1}...x_{3,\gamma-1,1} & y_{3,0,1}...y_{3,\gamma-1,1} \\
3 & z_{3,2} & x_{3,0,2}...x_{3,\gamma-1,2} & y_{3,0,2}...y_{3,\gamma-1,2} \\
3 & z_{3,3} & x_{3,0,3}...x_{3,\gamma-1,3} & y_{3,0,3}...y_{3,\gamma-1,3} \\
etc & etc & etc & etc
\end{pmatrix}
$$

Which compels the Quaternionic Lists X_β^* and Y_β^*, such that an element of $\vec{x}_{b,\beta}^*$ and $\vec{y}_{b,\beta}^*$ are the column vectors (for ordinary quaternions only, for biquaternions you must compute the conjugate)

$$
\vec{x}_\beta^* =
\begin{pmatrix}
+x_{b,0}\vec{q} \\
-x_{b,1}\vec{i} \\
-x_{b,2}\vec{j} \\
-x_{b,3}\vec{k}
\end{pmatrix}
; \vec{y}_\beta^* =
\begin{pmatrix}
+y_{b,0}\vec{q} \\
-y_{b,1}\vec{i} \\
-y_{b,2}\vec{j} \\
-y_{b,3}\vec{k}
\end{pmatrix}
, \forall \beta : 0 \le \beta \le \gamma - 1.
$$

This does not extend our Ruler Set (yet), since we are going to invoke the Variable Objects of Left and Left Conjugate Matrices for all X and Right and Right Conjugate Matrices for Y. We now the following set of four matrices to our Ruler for each β from $0 \le \beta \le \gamma - 1$

$$
X_{b,\beta,4\mathbb{H}L} =
\begin{pmatrix}
+x_{b,\beta,0} & -x_{b,\beta,1} & -x_{b,\beta,2} & -x_{b,\beta,3} \\
+x_{b,\beta,1} & +x_{b,\beta,0} & +x_{b,\beta,3} & -x_{b,\beta,2} \\
+x_{b,\beta,2} & -x_{b,\beta,3} & +x_{b,\beta,1} & +x_{b,\beta,2} \\
+x_{b,\beta,3} & +x_{b,\beta,2} & -x_{b,\beta,1} & +x_{b,\beta,0}
\end{pmatrix}
$$

$$
Y_{b,\beta,4\mathbb{H}R} =
\begin{pmatrix}
+y_{b,\beta,0} & -y_{b,\beta,1} & -y_{b,\beta,2} & -y_{b,\beta,3} \\
+y_{b,\beta,1} & +y_{b,\beta,0} & -y_{b,\beta,3} & +y_{b,\beta,2} \\
+y_{b,\beta,2} & +y_{b,\beta,3} & +y_{b,\beta,1} & -y_{b,\beta,2} \\
+y_{b,\beta,3} & -y_{b,\beta,2} & +y_{b,\beta,1} & +y_{b,\beta,0}
\end{pmatrix}
$$

$$
X_{b,\beta,4\mathbb{H}L}^* =
\begin{pmatrix}
+x_{b,\beta,0} & +x_{b,\beta,1} & +x_{b,\beta,2} & +x_{b,\beta,3} \\
-x_{b,\beta,1} & +x_{b,\beta,0} & -x_{b,\beta,3} & +x_{b,\beta,2} \\
-x_{b,\beta,2} & +x_{b,\beta,3} & +x_{b,\beta,1} & -x_{b,\beta,2} \\
-x_{b,\beta,3} & -x_{b,\beta,2} & +x_{b,\beta,1} & +x_{b,\beta,0}
\end{pmatrix}
= X_{b,\beta,4\mathbb{H}L}{}^T
$$

$$
Y_{b,\beta,4\mathbb{H}R}^* =
\begin{pmatrix}
+y_{b,\beta,0} & +y_{b,\beta,1} & +y_{b,\beta,2} & +y_{b,\beta,3} \\
-y_{b,\beta,1} & +y_{b,\beta,0} & +y_{b,\beta,3} & -y_{b,\beta,2} \\
-y_{b,\beta,2} & -y_{b,\beta,3} & +y_{b,\beta,1} & +y_{b,\beta,2} \\
-y_{b,\beta,3} & +y_{b,\beta,2} & -y_{b,\beta,1} & +y_{b,\beta,0}
\end{pmatrix}
= Y_{b,\beta,4\mathbb{H}R}{}^T
$$

which expands our *T* ruler to:

$$T, \vec{z}, \vec{x}_0...\vec{x}_{\gamma-1}, \vec{y}_0..\vec{y}_{\gamma-1}, X_{b,0,4\mathbb{H}L}{}^{*}..X_{b,\gamma-1,4\mathbb{H}L}{}^{*}, Y_{b,0,4\mathbb{H}R}{}^{*}..Y_{b,\gamma-1,4\mathbb{H}R}{}^{*}, X_{b,0,4\mathbb{H}L}..X_{b,\gamma-1,4\mathbb{H}L}, Y_{b,0,4\mathbb{H}R}..Y_{b,\gamma-1,4\mathbb{H}R}$$

Each matrix column element adds an extra column to the *T* Ruler. The *T* Ruler was already expanded fourfold per data point vertically, so the row entries of the matrix naturally fit into the existing structure. At every step, we only add more columns to the *T* Ruler, without altering its vertical expansion.

The Ruler undergoes vertical expansion only once at the very beginning of the Hypercomplex Least Squares process. This expansion follows a modulus of $2^{n_1}4^{n_2}8^{n_3}\ldots$ depending on the Rosenfeld logic $2^{n_1}\mathbb{C} \times 4^{n_2}\mathbb{H} \times 8^{n_3}\mathbb{O} \times \ldots$ that was invoked. All we need now are the Design and Response Matrices, after which, we need a new Ghost Ruler.

From $u = 0$ to $u = \gamma - 1$ and from $v = 0$ to $v = \gamma - 1$, for Ruler Index b:

$$D_{b,u,v} = (Y_{b,u,4\mathbb{H}L}{}^{*}X_{b,u,4\mathbb{H}R}{}^{*})(X_{b,v,4\mathbb{H}L}Y_{b,v,4\mathbb{H}R})$$

And let any (θ, ϕ) matrix element of $D_{b,u,v}$ be denoted $d_{b,u,v,\theta,\phi}$

And from $\beta = 0$ to $\beta = \gamma - 1$ $R_{b,\beta} = \left(X_{b,\beta,4\mathbb{H}L}{}^{*}Y_{b,\beta,4\mathbb{H}R}{}^{*}\right)\vec{z}_b$. And let any vector element of $R_{b,\beta}$ be denoted $r_{b,\beta,\theta}$

Now let *G* be the Ghost Ruler of height $4(\gamma)$ with $2 + 8(\gamma)$ columns. Append the integers 0 to 3 modulo 4 alongside the Ghost Ruler, representing θ mod 4 for each integer u from $u = 0$ to $u = \gamma - 1$, and append the integers 0 to 3 modulo 4 above the Ghost Ruler, representing ϕ mod 4 for each integer v from $v = 0$ to $v = \gamma - 1$.

We now fill the first 28 columns with the element-wise sum of the design matrices such that

$$G_{i,j} = G_{u,v,\theta,\phi} = \sum_{b=0}^{t-1} d_{b,u,v,\theta,\phi}$$

where $u = \lfloor\frac{i}{\gamma}\rfloor$ and $v = \lfloor\frac{j}{\gamma}\rfloor$ and $\theta = i - 4u$ and $\phi = j - 4u$. Using the 28x28 Ruler for the seven variable bivariate quadratic as an example (with $\gamma = 7$):

$$\begin{pmatrix}
 & & v=0 & v=0 & v=0 & v=0 & ...v=6 & v=6 & v=6 & v=6 \\
u & \theta & \phi=0 & \phi=1 & \phi=2 & \phi=3 & ...\phi=0 & \phi=1 & \phi=2 & \phi=3 \\
0 & 0 & \sum_{b=0}^{b=t-1} d_{b,u,v,\theta,\phi} & & & & & & & \sum_{b=0}^{b=t-1} d_{b,0,6,0,3} \\
0 & 1 & \sum_{b=0}^{b=t-1} d_{b,0,0,1,0} & & & & & & & \sum_{b=0}^{b=t-1} d_{b,0,6,1,3} \\
0 & 2 & \sum_{b=0}^{b=t-1} d_{b,0,0,2,0} & & & & & & & \sum_{b=0}^{b=t-1} d_{b,0,6,2,3} \\
0 & 3 & \sum_{b=0}^{b=t-1} d_{b,0,0,3,0} & & & & & & & \sum_{b=0}^{b=t-1} d_{b,0,6,3,3} \\
1 & 0 & \sum_{b=0}^{b=t-1} d_{b,1,0,0,0} & & & & & & & \sum_{b=0}^{b=t-1} d_{b,1,6,0,3} \\
1 & 1 & \sum_{b=0}^{b=t-1} d_{b,1,0,1,0} & & & & & & & \sum_{b=0}^{b=t-1} d_{b,1,6,1,3} \\
1 & 2 & \sum_{b=0}^{b=t-1} d_{b,1,0,2,0} & & & & & & & \sum_{b=0}^{b=t-1} d_{b,1,6,2,3} \\
1 & 3 & \sum_{b=0}^{b=t-1} d_{b,1,0,3,0} & & & & & & & \sum_{b=0}^{b=t-1} d_{b,1,6,3,3} \\
... & ... & ... & & & & & & & \\
6 & 0 & \sum_{b=0}^{b=t-1} d_{b,6,0,0,0} & & & & & & & \sum_{b=0}^{b=t-1} d_{b,6,6,0,3} \\
6 & 1 & \sum_{b=0}^{b=t-1} d_{b,6,0,1,0} & & & & & & & \sum_{b=0}^{b=t-1} d_{b,6,6,1,3} \\
6 & 2 & \sum_{b=0}^{b=t-1} d_{b,6,0,2,0} & & & & & & & \sum_{b=0}^{b=t-1} d_{b,6,6,2,3} \\
6 & 3 & \sum_{b=0}^{b=t-1} d_{b,6,0,3,0} & & & & & & & \sum_{b=0}^{b=t-1} d_{b,6,6,3,3}
\end{pmatrix}$$

We now fill next 4γ columns with the inverse matrix of Ruler *G* from $i = 0$ to $i = 27$ and $j = 0$ to $j = 27$. We do not need modular references here, this is the ordinary inversion of a 28x28 real number matrix. Call this portion of the Ruler G^{-1}.

We now fill the penultimate columns with the element-wise sum of the response columns such that

$$H_i = H_{\beta,\theta} = \sum_{b=0}^{t-1} r_{b,\beta,\theta}$$

where $\beta = \lfloor \frac{i}{\gamma} \rfloor$ and $\theta = i - 4\beta$. Using the 28x28 Ruler for the seven variable bivariate quadratic as an example (with $\gamma = 7$):

$$
\begin{pmatrix}
i & u & \theta & H_i \\
0 & 0 & 0 & \sum_{b=0}^{b=t-1} r_{b,u,\theta} \\
1 & 0 & 1 & \sum_{b=0}^{b=t-1} r_{b,0,1} \\
2 & 0 & 2 & \sum_{b=0}^{b=t-1} r_{b,0,2} \\
3 & 0 & 3 & \sum_{b=0}^{b=t-1} r_{b,0,3} \\
4 & 1 & 0 & \sum_{b=0}^{b=t-1} r_{b,1,0} \\
5 & 1 & 1 & \sum_{b=0}^{b=t-1} r_{b,1,1} \\
6 & 1 & 2 & \sum_{b=0}^{b=t-1} r_{b,1,2} \\
7 & 1 & 3 & \sum_{b=0}^{b=t-1} r_{b,1,3} \\
... \\
24 & 6 & 0 & \sum_{b=0}^{b=t-1} r_{b,6,0} \\
25 & 6 & 1 & \sum_{b=0}^{b=t-1} r_{b,6,1} \\
26 & 6 & 2 & \sum_{b=0}^{b=t-1} r_{b,6,2} \\
27 & 6 & 3 & \sum_{b=0}^{b=t-1} r_{b,6,3}
\end{pmatrix}
$$

Let the above be known as Column R. We finally calculate our regression constants $C = G^{-1}R$ and append this column element-wise into the final column of the G Ruler.

Then let $RSS = \sum_{b=0}^{b=t-1} \vec{\epsilon}_b \vec{\epsilon}_b^*$ where $\vec{\epsilon}_b = \vec{z}_b - \sum_{\beta=0}^{\beta=\gamma-1} X_{b,\beta,4\mathbb{H}L} Y_{b,\beta,4\mathbb{H}R} \vec{C}$.

And let $TSS = \sum_{b=0}^{b=t-1} \vec{\delta}_b \vec{\delta}_b^*$ where $\vec{\delta}_b = \vec{z}_b - \bar{\vec{z}}$ and $\bar{\vec{z}} = \frac{1}{t} \sum_{b=0}^{b=1} \vec{z}_t$.

Such that $R^2 = 1 - \frac{RSS}{TSS}$ and $\Psi = \ln \frac{RSS}{TSS}$.

6.2. Direct Human Readable Example

We now wish to resolve the particular quadratic equation specified in the first section of the article: $\vec{z}_b = \vec{c}_0 + \vec{c}_1 \vec{x}_b + \vec{y}_b \vec{c}_2 + \vec{x}_b^2 \vec{c}_3 + \vec{x}_b \vec{c}_4 \vec{y}_b + \vec{y}_b \vec{c}_5 \vec{x}_b + \vec{c}_6 \vec{y}_b^2$.

We shall rewrite this in it's full middle-handed form:

$$\vec{z}_b = \vec{x}_{b,0} \vec{c}_0 \vec{y}_{b,0} + \vec{x}_{b,1} \vec{c}_1 \vec{y}_{b,1} + \vec{x}_{b,2} \vec{c}_2 \vec{y}_{b,2} ... \vec{x}_{b,6} \vec{c}_6 \vec{y}_{b,6} = \sum_{\beta=0}^{\beta=\gamma-1} \vec{x}_{b,\beta} \vec{c}_\beta \vec{y}_{b,\beta}$$

1. $\vec{x}_{b,0} = \vec{q}$ and $\vec{y}_{b,0} = \vec{q}$
2. $\vec{x}_{b,1} = \vec{q}$ and $\vec{y}_{b,1} = \vec{x}_b^1$.
3. $\vec{x}_{b,2} = \vec{y}_b^1$ and $\vec{y}_{b,2} = \vec{q}$.
4. $\vec{x}_{b,3} = \vec{x}_b^2$ and $\vec{y}_{b,3} = \vec{q}$.
5. $\vec{x}_{b,4} = \vec{x}_b^1$ and $\vec{y}_{b,4} = \vec{y}_b^1$.
6. $\vec{x}_{b,5} = \vec{y}_b^1$ and $\vec{y}_{b,5} = \vec{x}_b^1$.
7. $\vec{x}_{b,6} = \vec{q}$ and $\vec{y}_{b,6} = \vec{y}_b^2$.
8. $D_{b,\beta} = \left[\vec{x}_{b,\beta}\right]_{4\mathbb{H}L} \left[\vec{y}_{b,\beta}\right]_{4\mathbb{H}R}$ and $D_{b,\beta}^* = \left[\vec{y}_{b,\beta}^*\right]_{4\mathbb{H}L} \left[\vec{x}_{b,\beta}^*\right]_{4\mathbb{H}R}$.

From which is follows that the Ruler Matrix, G, is of the form (I am running from $b = 1$ to $b = t$ in order to save space in the notation, the sums still runs from $b = 0$ to $b = t - 1$ in code):

$$
G = \begin{pmatrix}
\sum_{b=1}^{b=t} D_{b,0}^* D_{b,0} & \sum_{b=1}^{b=t} D_{b,0}^* D_{b,1} & \sum_{b=1}^{b=t} D_{b,0}^* D_{b,2} & \sum_{b=1}^{b=t} D_{b,0}^* D_{b,3} & \sum_{b=1}^{b=t} D_{b,0}^* D_{b,4} & \sum_{b=1}^{b=t} D_{b,0}^* D_{b,5} & \sum_{b=1}^{b=t} D_{b,0}^* D_{b,6} \\
\sum_{b=1}^{b=t} D_{b,1}^* D_{b,0} & \sum_{b=1}^{b=t} D_{b,1}^* D_{b,1} & \sum_{b=1}^{b=t} D_{b,1}^* D_{b,2} & \sum_{b=1}^{b=t} D_{b,1}^* D_{b,3} & \sum_{b=1}^{b=t} D_{b,1}^* D_{b,4} & \sum_{b=1}^{b=t} D_{b,1}^* D_{b,5} & \sum_{b=1}^{b=t} D_{b,1}^* D_{b,6} \\
\sum_{b=1}^{b=t} D_{b,2}^* D_{b,0} & \sum_{b=1}^{b=t} D_{b,2}^* D_{b,1} & \sum_{b=1}^{b=t} D_{b,2}^* D_{b,2} & \sum_{b=1}^{b=t} D_{b,2}^* D_{b,3} & \sum_{b=1}^{b=t} D_{b,2}^* D_{b,4} & \sum_{b=1}^{b=t} D_{b,2}^* D_{b,5} & \sum_{b=1}^{b=t} D_{b,2}^* D_{b,6} \\
\sum_{b=1}^{b=t} D_{b,3}^* D_{b,0} & \sum_{b=1}^{b=t} D_{b,3}^* D_{b,1} & \sum_{b=1}^{b=t} D_{b,3}^* D_{b,2} & \sum_{b=1}^{b=t} D_{b,3}^* D_{b,3} & \sum_{b=1}^{b=t} D_{b,3}^* D_{b,4} & \sum_{b=1}^{b=t} D_{b,3}^* D_{b,5} & \sum_{b=1}^{b=t} D_{b,3}^* D_{b,6} \\
\sum_{b=1}^{b=t} D_{b,4}^* D_{b,0} & \sum_{b=1}^{b=t} D_{b,4}^* D_{b,1} & \sum_{b=1}^{b=t} D_{b,4}^* D_{b,2} & \sum_{b=1}^{b=t} D_{b,4}^* D_{b,3} & \sum_{b=1}^{b=t} D_{b,4}^* D_{b,4} & \sum_{b=1}^{b=t} D_{b,4}^* D_{b,5} & \sum_{b=1}^{b=t} D_{b,4}^* D_{b,6} \\
\sum_{b=1}^{b=t} D_{b,5}^* D_{b,0} & \sum_{b=1}^{b=t} D_{b,5}^* D_{b,1} & \sum_{b=1}^{b=t} D_{b,5}^* D_{b,2} & \sum_{b=1}^{b=t} D_{b,5}^* D_{b,3} & \sum_{b=1}^{b=t} D_{b,5}^* D_{b,4} & \sum_{b=1}^{b=t} D_{b,5}^* D_{b,5} & \sum_{b=1}^{b=t} D_{b,5}^* D_{b,6} \\
\sum_{b=1}^{b=t} D_{b,6}^* D_{b,0} & \sum_{b=1}^{b=t} D_{b,6}^* D_{b,1} & \sum_{b=1}^{b=t} D_{b,6}^* D_{b,2} & \sum_{b=1}^{b=t} D_{b,6}^* D_{b,3} & \sum_{b=1}^{b=t} D_{b,6}^* D_{b,4} & \sum_{b=1}^{b=t} D_{b,6}^* D_{b,5} & \sum_{b=1}^{b=t} D_{b,6}^* D_{b,6}
\end{pmatrix}
$$

And the Response Column is of the Form:

$$R = \begin{pmatrix} \sum_{b=1}^{b=t} D_{b,0}^* \vec{z}_b \\ \sum_{b=1}^{b=t} D_{b,1}^* \vec{z}_b \\ \sum_{b=1}^{b=t} D_{b,2}^* \vec{z}_b \\ \sum_{b=1}^{b=t} D_{b,3}^* \vec{z}_b \\ \sum_{b=1}^{b=t} D_{b,4}^* \vec{z}_b \\ \sum_{b=1}^{b=t} D_{b,5}^* \vec{z}_b \\ \sum_{b=1}^{b=t} D_{b,6}^* \vec{z}_b \end{pmatrix} \; ; GC = R \iff G^{-1} R = C = \begin{pmatrix} \vec{c}_0 \\ \vec{c}_1 \\ \vec{c}_2 \\ \vec{c}_3 \\ \vec{c}_4 \\ \vec{c}_5 \\ \vec{c}_6 \end{pmatrix}$$

Then $C = G^{-1}R$

Each u, v 4x4 Block of the Ruler Matrix, G, can be expressed succinctly as: $G_{u,v} = \sum_{b=1}^{b=t} D_{b,u}^* D_{b,v}$. Each β 4x1 Rectangular Block of the Response Matrix, R, can be expressed succinctly as: $R_\beta = \sum_{b=1}^{b=t} D_{b,\beta}^* \vec{z}_b$.

Remember that Matrix Multiplication is not commutative, hence $D_{b,u}^*$ must always be the left product and $D_{b,v}$ the right product. Even though the case of $u = v$ commutes, that remaining cases do not, especially since $D_{b,u}^*$ must be multiplied from left against $\vec{z}_b$ in the response column!

Q.E.D.

We shall continue with Least Squares over the Tessarines (thanks to Aiden Degrace) and the Biquaternions in the next publication.

Until, it is recommended that you read Aram Boyijian's thesis titled *The Physical Interpretation of Complex Angles* to prepare yourself for the next installment.

Final Note: A potential critique might assert that $C = G^{-1}R$ does not guarantee minimal error, suggesting it yields a suboptimal C, such as one convergent to a local minimum akin to the 2008 iterative solution or neural network optimization. However, this is equivalent to claiming—using the 28x28 matrix example from the preceding section—that ordinary least squares (OLS) regression fails in $\mathbb{R}^{28}$. Yet, this reduction to a real-valued OLS problem ensures C globally minimizes the residual sum of squares.

To complement this theoretical assurance, empirical validation confirms the method's efficacy: Generate quaternionic lists X and Y with uniformly random components, set a known C, compute $\vec{z}_b = \sum_{\beta=0}^{\beta=\gamma-1} \vec{x}_{b,\beta} \vec{c}_\beta \vec{y}_{b,\beta}$ and add Gaussian noise to each $\vec{z}_b$ component using, for instance, Excel's $NORMINV(RAND(), 0, \sigma)$ function after an error-free initial computation.

This approach consistently recovers the original C, with deviations attributable to the added noise, affirming the solution's robustness. Before introducing noise, I also suggest generating $\vec{z}_b$ **no error** to directly observe the result where $R^2 = 1$.

References

1. Minghui Wang, Mushseng Wei, Yan Feng
   An Iterative Algorithm for Least Squares in Quaternionic Quantum Theory
   Computer Physics Communications, Volume 179, Issue 4, Pages 203-207
2. https://youtu.be/FOhWGq9KExE?si=SpI9kjdcg-WIO_yI
   JMM2023 Conference, January 7th, 2023: Closed Form Solution to Quaternionic Least Squares