Jacobian Conjecture: Making Sense of the AI-Assisted Mathematical Discovery

Luisenrique Zepeda-Rodriguez, Vladik Kreinovich, and Nguyen Hoang Phuong

Abstract Recently, an AI helped to find a counterexample to a long-standing mathematical problem: Jacobian Conjecture about invertibility of polynomial mappings. The problem that AIs in general – and in this case in particular – do not provide any explanations for their answers. It is therefore desirable to provide some explanations for this counterexample. In this paper, we show that some simple properties explain the set of monomials used in this counterexample. How to explain the coefficients at these monomials is still an open problem.

1 Formulation of the Problem

What is a Jacobian and what is the Jacobian Conjecture. This conjecture is about smooth (= differentiable) mappings $y = f(x)$ from the space $\mathbb{R}^n$ of all n-element tuples of real numbers $x = (x_1, \ldots, x_n)$ to itself:

$$y_1 = f_1(x_1, \ldots, x_n), \ldots, y_n = f_n(x_1, \ldots, x_n).$$

Specifically, it is about mappings in which all the functions $f_i(x_1, \ldots, x_n)$ are polynomials.

The simplest case is when all these polynomials are linear, i.e.,

Luisenrique Zepeda-Rodriguez
Department of Computer Science, University of Texas at El Paso, 500 W. University
El Paso, Texas 79968, USA, e-mail: lzepedarodrig@miners.utep.edu¿

Vladik Kreinovich
Department of Computer Science, University of Texas at El Paso, 500 W. University
El Paso, Texas 79968, USA, e-mail: vladik@utep.edu

Nguyen Hoang Phuong
Artificial Intelligence Division, Information Technology Faculty, Thang Long University
Nghiem Xuan Yem Road, Hoang Mai District, Hanoi, Vietnam
e-mail: nhphuong2008@gmail.com

$$f_i(x_1,\ldots,x_n) = a_n + \sum_{j=1}^{n} b_{i,j} \cdot x_j.$$

In this case, according to linear algebra, there are two options:

- either the matrix b with components $b_{i,j}$ is non-degenerate, i.e., its determinant $D = \det(b)$ is different from 0; in this case, the mapping is reversible;
- or the determinant is equal to 0, in which case the mapping is not reversible, i.e., there exist tuples $x \neq x'$ for which $f(x) = f(x')$.

What happens in a non-linear case? In this case, in the vicinity of each point $x^{(0)} = \left(x_1^{(0)},\ldots,x_n^{(0)}\right)$, the mapping can be approximated by a linear transformation

$$y_i = f_i\left(x^{(0)}\right) + \sum_{j=1}^{n} b_{i,j}\left(x^{(0)}\right) \cdot \left(x_i - x_i^{(0)}\right),$$

where

$$b_{i,j}\left(x^{(0)}\right) \stackrel{\text{def}}{=} \frac{\partial f_i}{\partial x_j}_{|x=x^{(0)}}.$$

The matrix consisting of the elements b_{ij} is called the *Jacobian*.

In the linear case, for each point $x^{(0)}$, the Jacobian coincides with the matrix $b_{i,j}$: $b_{i,j}\left(x^{(0)}\right) = b_{i,j}$. So, in this case, the determinant $D\left(x^{(0)}\right)$ of the Jacobian is a constant – it has the same value for all points $x^{(0)}$. Interesting, it turned out that for the case when $n = 2$. every polynomial mapping with the constant non-zero determinant $D\left(x^{(0)}\right)$ is reversible.

For $n = 3$, for a long time, researchers could neither prove that every polynomial mapping with the constant non-zero determinant is reversible, not could they find a counter-example. The resulting hypothesis that every polynomial mapping with constant non-zero determinant is reversible is known as the *Jacobian Conjecture*.

AI helped to find a counter-example. In 2026, a researcher used AI to help find a counterexample – and such a counterexample was successfully found [1]. In the equivalent form presented in [3], this counterexample has the following form:

$$y_1 = x_1 - 3\cdot x_1^2\cdot x_2 - x_1^3\cdot x_3;$$

$$y_2 = x_2 + 3\cdot x_1\cdot x_3 + 24\cdot x_1\cdot x_2^2 + 12\cdot x_1^2\cdot x_2\cdot x_3 + 36\cdot x_1^2\cdot x_2^3 + 12\cdot x_1^3\cdot x_2^2\cdot x_3; \quad (1)$$

$$y_3 = x_3 + 8\cdot x_2^2 + 6\cdot x_1\cdot x_2\cdot x_3 + 28\cdot x_1\cdot x_2^3 + 12\cdot x_1^2\cdot x_2^2\cdot x_3 + 24\cdot x_1^2\cdot x_2^4 + 8\cdot x_1^3\cdot x_2^3\cdot x_3.$$

Natural question. As usual, AI does not provide any logical explanation of how it came up with this example. From this viewpoint, modern AI tools are black boxes: they produce results, but they do not explain how they got them.

It is therefore desirable to make sense and to try to explain at least some of its features.

2 What Is Known and What We Do in This Paper

What is known. It is known (see, e.g., [2]) that the transformation (1) is invariant with respect to a special scaling $T_\lambda(x_1,x_2,x_2) = \left(\lambda \cdot x_1, \lambda^{-1} \cdot x_2, \lambda^{-2} \cdot x_3\right)$. Namely, for every tuple x and for every λ, if $y = f(x)$, then $T_\lambda(y) = f(T_\lambda(x))$.

This helps, but additional help is needed. The scale-invariance condition helps explain why some monomial do not appear in the formula (1), but still there are infinitely many different monomials whose addition could preserve scale-invariance, e.g., $x_1^{3+n} \cdot x_2^n \cdot x_3$ in f_1 for $n = 1,2,3,\ldots$ We need additional explanations for the formulas (1).

Let us divide this problem into two. By explaining, we mean that we can have some simple conditions that would lead us to the formula (1). It is natural to consider two steps:

- first, it would be nice to explain why these specific monomials appear in the formula (1) and not any others;
- second, once we understand why only these monomials are present, it would be nice to understand the numerical coefficients wwith which these monomials occur in the formula (1).

What we do in this paper. In this paper, we provide a possible explanation for the first question – why these specific monomials. Specifically, we provide simple conditions that lead exactly to the monomials from (1). The second part is still an open problem.

3 Our Result

Proposition. *Let $f(x)$ be a scale-invariant polynomial mapping that has the following two properties:*

- *for each i, one of the monomials with the highest overall degree in $f_i(x)$ has the form $x_1^2 \cdot x_2^{\alpha_2} \cdot x_3$ for some α_2;*
- *for each i, the expression $f_i(x)$ is linear in x_3.*

Then, for each i, all the monomials in $f_i(x)$ are exactly the mononimals from (1).

Comment. Both conditions are necessary:

- if we do not have the first condition, then, e.g., in $f_1(x)$, we could have additional monomials $x_1^{3+n} \cdot x_2^n \cdot x_3$ for $n = 1,2,3,\ldots$;
- if we do not have the second condition, then, e.g., in $f_3(x)$, we can have an additional monomial $x_2^2 \cdot x_3^2$.

Proof of the Proposition. To prove this result, let us consider the cases $i = 1$, $i = 2$, and $i = 3$ one by one.

$1°$. A general monomial has the form $x_1^{\alpha_1} \cdot x_1^{\alpha_2} \cdot x_1^{\alpha_3}$ for some non-negative integers α_i. By definition, the overall degree d of the monomial is equal to the sum of the degrees α_i: $d = \alpha_1 + \alpha_2 + \alpha_3$. Scale-invariance implies that

$$(\lambda \cdot x_1)^{\alpha_1} \cdot \left(\lambda^{-1} \cdot x_2\right)^{\alpha_2} \cdot \left(\lambda^{-2} \cdot x_3\right)^{\alpha_3} = \lambda \cdot x_1^{\alpha_1} \cdot x_2^{\alpha_2} \cdot x_3^{\alpha_3}.$$

In the left-hand side, the dependence on λ has the form

$$\lambda^{\alpha_1} \cdot \lambda^{-\alpha_2} \cdot \lambda^{-2\alpha_3} = \lambda^{\alpha_1 - \alpha_2 - 2\alpha_3},$$

while on the right-hand side, we have $\lambda = \lambda^1$. Thus, we must have

$$\alpha_1 - \alpha_2 - 2\alpha_3 = 1. \tag{2}$$

In particular, this must be true for the term of the largest degree that has the form $x_1^3 \cdot x_2^{\alpha_2} \cdot x_3$. So for this term, we have $3 - \alpha_2 - 2 = 1$, so we must have $\alpha_2 = 0$. In this case, the overall degree of this term is $3 + 0 + 1 = 4$. Since this term has the largest overall degree, this means that all the monomials in $f_1(x)$ must have degree not larger than 4:

$$\alpha_1 + \alpha_2 + \alpha_3 \le 4. \tag{3}$$

Since the dependence of each term on x_3 is linear, we will always have $\alpha_3 = 0$ or $\alpha_3 = 1$. Let us consider these two cases one by one,

$1.0°$. Let us first consider the case when $\alpha_3 = 0$. In this case, (2) turns into $\alpha_1 - \alpha_2 = 1$, so $\alpha_1 = \alpha_2 + 1$. Substituting this expression for α_1 and 0 for α_3 into the inequality (3), we conclude that $2\alpha_2 + 1 \le 4$, i.e., $2\alpha_2 \le 3$ and $\alpha_2 \le 3/2$. Since α_2 is a non-negative integer, we have either $\alpha_2 = 0$ or $\alpha_2 = 1$. Let us consider these two cases one by one.

- If $\alpha_2 = 0$, then $\alpha_1 = \alpha_2 + 1 = 1$ and, since $\alpha_3 = 0$, we have the term x_1.
- If $\alpha_2 = 1$, then $\alpha_1 = \alpha_2 + 1 = 2$, and, since $\alpha_3 = 0$, we have the term $x_1^2 \cdot x_2$.

Both terms are in $f_1(x)$.

$1.1°$. Let us now consider the case when $\alpha_3 = 1$. In this case, (2) turns into $\alpha_1 - \alpha_2 - 2 = 1$, so $\alpha_1 = \alpha_2 + 3$. Substituting this expression for α_1 and 1 for α_3 into the inequality (3), we conclude that $2\alpha_2 + 3 \le 4$, i.e., $2\alpha_2 \le 1$ and $\alpha_2 \le 1/2$. Since α_2 is a non-negative integer, we have $\alpha_2 = 0$. Then, $\alpha_1 = \alpha_2 + 3 = 3$ and, since $\alpha_3 = 0$, we have the term $x_1^3 \cdot x_2$. This term is also in $f_1(x)$.

We have thus explained all the terms in $f_1(x)$.

$2°$. Similar to Part 1 of this proof, for terms in $f_2(x)$, sca;e-invariance implies that

$$\alpha_1 - \alpha_2 - 2\alpha_3 = -1. \tag{4}$$

In particular, this must be true for the term of the largest degree that has the form $x_1^3 \cdot x_2^{\alpha_2} \cdot x_3$, so for this term, we have $3 - \alpha_2 - 2 = -1$, so we must have $\alpha_2 = 2$. In this case, the overall degree of this term is $3 + 2 + 1 = 6$. Since this term has the largest overall degree, this means that all the monomials in $f_2(x)$ must have degree not larger than 6:

$$\alpha_1 + \alpha_2 + \alpha_3 \leq 6. \tag{5}$$

Since the dependence of each term on x_3 is linear, we will always have $\alpha_3 = 0$ or $\alpha_3 = 1$. Let us consider these two cases one by one,

2.0°. Let us first consider the case when $\alpha_3 = 0$. In this case, (4) turns into $\alpha_1 - \alpha_2 = -1$, so $\alpha_2 = \alpha_1 + 1$. Substituting this expression for α_2 and 0 for α_3 into the inequality (5), we conclude that $2\alpha_1 + 1 \leq 6$, i.e., $2\alpha_1 \leq 5$ and $\alpha_1 \leq 5/2$. Since α_1 is a non-negative integer, we have either $\alpha_1 = 0$ or $\alpha_1 = 1$ or $\alpha_1 = 2$. Let us consider these three cases one by one.

- If $\alpha_1 = 0$, then $\alpha_2 = \alpha_1 + 1 = 1$ and, since $\alpha_3 = 0$, we have the term x_2.
- If $\alpha_1 = 1$, then $\alpha_2 = \alpha_1 + 1 = 2$, and, since $\alpha_3 = 0$, we have the term $x_1 \cdot x_2^2$.
- If $\alpha_1 = 2$, then $\alpha_2 = \alpha_1 + 1 = 3$, and, since $\alpha_3 = 0$, we have the term $x_1^2 \cdot x_2^3$.

All three terms are in $f_2(x)$.

2.1°. Let us now consider the case when $\alpha_3 = 1$. In this case, (4) turns into $\alpha_1 - \alpha_2 - 2 = -1$, so $\alpha_1 = \alpha_2 + 1$. Substituting this expression for α_1 and 1 for α_3 into the inequality (4), we conclude that $2\alpha_2 + 2 \leq 6$, i.e., $2\alpha_2 \leq 4$ and $\alpha_2 \leq 2$. Since α_2 is a non-negative integer, we can have $\alpha_2 = 0$, $\alpha_2 = 1$, or $\alpha_2 = 2$. Let us consider these three cases one by one.

- If $\alpha_2 - 0$, then $\alpha_1 = \alpha_2 + 1 = 1$ and, since $\alpha_3 = 1$, we have the term $x_1 \cdot x_3$.
- If $\alpha_2 = 1$, then $\alpha_1 = \alpha_2 + 1 = 2$, and, since $\alpha_3 = 1$, we have the term $x_1^2 \cdot x_2 \cdot x_3$.
- If $\alpha_2 = 2$, then $\alpha_1 = \alpha_2 + 1 = 3$, and, since $\alpha_3 = 1$, we have the term $x_1^3 \cdot x_2^2 \cdot x_3$.

All three terms are also in $f_2(x)$.

We have thus explained all the terms in $f_2(x)$.

3°. Similar to Part 1 of this proof, for terms in $f_3(x)$, scale-invariance implies that

$$\alpha_1 - \alpha_2 - 2\alpha_3 = -2. \tag{6}$$

In particular, this must be true for the term of the largest degree that has the form $x_1^3 \cdot x_2^{\alpha_2} \cdot x_3$, so for this term, we have $3 - \alpha_2 - 2 = -2$, so we must have $\alpha_2 = 3$. In this case, the overall degree of this term is $3 + 3 + 1 = 7$. Since this term has the largest overall degree, this means that all the monomials in $f_3(x)$ must have degree not larger than 7:

$$\alpha_1 + \alpha_2 + \alpha_3 \leq 7. \tag{7}$$

Since the dependence of each term on x_3 is linear, we will always have $\alpha_3 = 0$ or $\alpha_3 = 1$. Let us consider these two cases one by one,

3.0°. Let us first consider the case when $\alpha_3 = 0$. In this case, (6) turns into $\alpha_1 - \alpha_2 = -2$, so $\alpha_2 = \alpha_1 + 2$. Substituting this expression for α_2 and 0 for α_3 into the inequality (7), we conclude that $2\alpha 1 + 2 \leq 7$, i.e., $2\alpha_1 \leq 5$ and $\alpha_1 \leq 5/2$. Since α_1 is a non-negative integer, we have either $\alpha_1 = 0$ or $\alpha_1 = 1$ or $\alpha_1 = 2$. Let us consider these three cases one by one.

- If $\alpha_1 = 0$, then $\alpha_2 = \alpha_1 + 2 = 2$ and, since $\alpha_3 = 0$, we have the term x_2^2.
- If $\alpha_1 = 1$, then $\alpha_2 = \alpha_1 + 2 = 3$, and, since $\alpha_3 = 0$, we have the term $x_1 \cdot x_2^3$.
- If $\alpha_1 = 2$, then $\alpha_2 = \alpha_1 + 2 = 4$, and, since $\alpha_3 = 0$, we have the term $x_1^2 \cdot x_2^4$.

All three terms are in $f_3(x)$.

3.1°. Let us now consider the case when $\alpha_3 = 1$. In this case, (6) turns into $\alpha_1 - \alpha_2 - 2 = -2$, so $\alpha_1 = \alpha_2$. Substituting this expression for α_1 and 1 for α_3 into the inequality (7), we conclude that $2\alpha_2 + 1 \le 7$, i.e., $2\alpha_2 \le 6$ and $\alpha_2 \le 3$. Since α_2 is a non-negative integer, we can have $\alpha_2 = 0$, $\alpha_2 = 1$, $\alpha_2 = 2$, or $\alpha_2 = 3$. Let us consider these three cases one by one.

- If $\alpha_2 - 0$, then $\alpha_1 = \alpha_2 = 0$ and, since $\alpha_3 = 1$, we have the term x_3.
- If $\alpha_2 = 1$, then $\alpha_1 = \alpha_2 = 1$, and, since $\alpha_3 = 1$, we have the term $x_1 \cdot x_2 \cdot x_3$.
- If $\alpha_2 = 2$, then $\alpha_1 = \alpha_2 = 2$, and, since $\alpha_3 = 1$, we have the term $x_1^2 \cdot x_2^2 \cdot x_3$.
- If $\alpha_2 = 3$, then $\alpha_1 = \alpha_2 = 3$, and, since $\alpha_3 = 1$, we have the term $x_1^3 \cdot x_2^3 \cdot x_3$.

All three terms are also in $f_3(x)$.

We have thus explained all the terms in $f_3(x)$. The proposition is proven.

Acknowledgments

This work was supported by the AT&T Fellowship in Information Technology, by the Institute for Risk and Reliability, Leibniz Universitaet Hannover, Germany, by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) Focus Program SPP 100+ 2388, Grant Nr. 501624329, by the European Union under the project ROBOPROX (No. CZ.02.01.01/00/22 008/0004590), by the Center of Excellence in Econometrics, Faculty of Economics, Chiang Mai University, Thailand, by the Ho Chi Minh City University of Banking, Vietnam, and by Thang Long University, Hanoi, Vietnam.

References

1. L. Alpöge, *Post announcing an explicit counterexample to the Jacobian Conjecture*, X, July 20, 2026, https://x.com/__alpoge__/status/2079028340955197566 (accessed July 20, 2026)
2. S. Gao, *Counterexamples to the Jacobian conjecture in dimensions greater than two*, arXiv:2608.00222v1, 2026.
3. C. D. Long, *Small Counterexamples to the Gaussian Moments Conjecture*, arXiv:2607.1818v1, 2026.