MedCrave
Step into the World of Research

Biometrics & Biostatistics International Journal

Review Article

Open Access

CrossMark

A note on the application of Taylor expansions for vector- and matrix-valued functions in statistical inference

Abstract

Taylor expansion is a vital mathematical tool applied across various scientific disciplines. However, there is a widespread misuse of this popular technique in the statistical literature. While mean-value forms for real-valued functions rely on a scalar intermediate point, directly extending this single-intermediate-point representation (Lagrange form) to vector- and matrix-valued functions is mathematically invalid, as a single common evaluation point generally does not exist. This paper addresses the issue by presenting accurate forms of Taylor expansions and highlighting that the commonly used single-point versions by some statisticians are non-existent. We distinguish between exact finite-sample equality errors and the validity of theoretical asymptotic properties (such as consistency and asymptotic normality), demonstrating that theoretical conclusions typically remain sound when valid component wise or integral-form expansions are employed. By bringing clarity and exact mathematical precision to the proper use of Taylor expansions, the statistical community can enhance the integrity of scientific research and impart rigorous theoretical standards to future generations of statisticians.

Keywords: estimating equation, integral-form taylor expansion, matrix-valued function, mean-value theorem, vector-valued function

Volume 15 Issue 1 - 2026

Changyong Feng,[1,2] Tong Wu,[3] Hongyue Wang,[1] Honghong Liu[4]

[1]Department of Biostatistics and Computational Biology, University of Rochester, USA
[2]Department of Anesthesiology, University of Rochester Medical Center, USA
[3]Department of Management Science and Information Systems, College of Management, University of Massachusetts Boston, USA
[4]Department of Statistics, North Carolina State University, USA

Correspondence: Changyong Feng, Department of Biostatistics and Computational Biology, Department of Anesthesiology, University of Rochester, Rochester, NY 14642, USA

Received: October 5, 2026 | **Published:** November 13, 2026

Mathematics subject classification: 62E20, 62F12

Introduction

Taylor expansion is a fundamental result in mathematical analysis and finds widespread application in statistical sciences. It provides local polynomial approximations that serve as the foundational cornerstone for deriving central limit theorems, asymptotic normality, consistency, and limiting efficiency for M-estimators, generalized method of moments (GMM) estimators, maximum likelihood estimators, and penalized regression frameworks.

However, there is a significant oversight in the statistical community regarding a proper understanding of this tool, resulting in extensive misuse of single-point mean-value representations in multidimensional settings. Even well-known statisticians with esteemed reputations have inadvertently fallen victim to misusing this powerful technique. We examined articles across several highly-cited statistical journals and discovered a troubling trend. Between 1982 and 2023, over 100 papers (involving more than 250 PhD statisticians) utilized various forms of vector or matrix Taylor expansions evaluated at a single common intermediate point. However, these single-point mean-value expansions for vector- and matrix-valued functions do not exist in general mathematics. What is noteworthy is that these publications originated from esteemed journals such as *Annals of Statistics* (AOS), *Biometrika*, *Journal of the American Statistical Association* (JASA), *Journal of the Royal Statistical Society*, Series B (JRSSB), *Econometrica*, and *Journal of Econometrics* (JOE), rather than obscure sources.

While previous work by Feng et al.[1] highlighted the nonexistence of a single intermediate evaluation point in vector mean-value forms, it is essential to clarify the precise theoretical scope and practical consequences of this issue. Specifically, a rigorous distinction must be drawn between (i) an exact finite-sample equality error arising from the nonexistence of a universal scalar intermediate point, and (ii) the validity of the final asymptotic distributions. In the vast majority of statistical applications, substituting mathematically sound component-wise expansions or integral-form representations fully preserves the stochastic order bounds (e.g., $O_p(n^{-1/2})$ or $o_p(1)$) and limiting normal distributions, leaving the primary scientific and methodological conclusions of the cited literature intact.

This paper delves into a comprehensive examination of this long-standing mathematical result, specifically focusing on four types of nonexistent single-point Taylor expansions observed in statistical publications. Additionally, we propose exact corrective approaches using integral representations to rectify these issues while preserving theoretical proofs. The paper is structured as follows: Section 2 reviews the first- and second-order Taylor expansion for twice continuously differentiable real-valued functions. Section 3 explains and illustrates, using Examples and formal Proofs, the nonexistence of the four types of so-called single-point Taylor expansions for vector-valued and matrix-valued functions. Section 4 demonstrates how these nonexistent Taylor expansions have been incorporated into statistical publications and provides a reassessment of their asymptotic validity. Section 5 offers exact integral-form and component wise remedies, followed by concluding remarks in Section 6.

First- and second-order Taylor expansions for real-valued functions

In this paper, unless otherwise specified, all vectors are assumed to be column vectors. The transposes of a vector $\mathbf{x}$ and a matrix $\mathbf{A}$ are denoted as $\mathbf{x}^\top$ and $\mathbf{A}^\top$, respectively. Our discussion begins with the following definition.

Definition

Let

$$\mathbf{a} = (a_1, \ldots, a_p)^\top, \mathbf{b} = (b_1, \ldots, b_p)^\top, p \geq 1$$

©2026 Feng et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and build upon your work non-commercially.

(i) The line segment between $\mathbf{a}$ and $\mathbf{b}$ is the set

$$\overline{\mathbf{ab}} = \{t\mathbf{a} + (1-t)\mathbf{b} : 0 \le t \le 1\}$$

(ii) The hyper-rectangle enclosed by $\mathbf{a}$ and $\mathbf{b}$ is the set

$$[[\mathbf{a},\mathbf{b}]] = \{\mathbf{x} = (x_1,\ldots,x_p)^\top : \min(a_i,b_i) \le x_i \le \max(a_i,b_i), i = 1,\ldots,p\}$$

It is obvious that both $\overline{\mathbf{ab}}$ and $[[\mathbf{a},\mathbf{b}]]$ are convex, and $\overline{\mathbf{ab}}$ is a diagonal of $[[\mathbf{a},\mathbf{b}]]$. For any $\mathbf{x},\mathbf{y} \in [[\mathbf{a},\mathbf{b}]]$, $\|\mathbf{x}-\mathbf{y}\| \le \|\mathbf{a}-\mathbf{b}\|$, where $\|\cdot\|$ is the ℓ^2-norm.

Taylor expansion offers an approximation for functions satisfying specific continuity and differentiability conditions. We begin by revisiting the well-known first- and second-order Taylor expansions of real-valued functions. These results can be found in various mathematical analysis textbooks, such as Rudin,[2] Lang,[3] and Fitzpatrick.[4] While there are multiple ways to express the Taylor expansion, we adopt Lagrange's form of Taylor's Theorem. This form incorporates an 'intermediate' value to provide an exact expansion of the function.

Suppose $\mathcal{O}$ is a convex open subset of $\mathbb{R}^p\,(p \ge 1)$ and $f : \mathcal{O} \to \mathbb{R}$ is continuously differentiable with the derivative function

$$f'(\mathbf{x}) = \frac{\partial f(\mathbf{x})}{\partial \mathbf{x}^\top} = \left[\frac{\partial f(\mathbf{x})}{\partial x_1}, \frac{\partial f(\mathbf{x})}{\partial x_2}, \ldots, \frac{\partial f(\mathbf{x})}{\partial x_p}\right], \mathbf{x} \in \mathcal{O}$$

Then for any $\mathbf{a},\mathbf{b} \in \mathcal{O}$, there exists a $\theta \in \overline{\mathbf{ab}}$ such that

$$f(\mathbf{b}) = f(\mathbf{a}) + f'(\theta)(\mathbf{b}-\mathbf{a}) \tag{1}$$

If f is also twice continuously differentiable with the second-order derivative function $f''(\mathbf{x}) = \partial^2 f(\mathbf{x}) / \partial\mathbf{x}\partial\mathbf{x}^\top$, then for any $\mathbf{a},\mathbf{b} \in \mathcal{O}$, there always exists a $\theta \in \overline{\mathbf{ab}}$ such that

$$f(\mathbf{b}) = f(\mathbf{a}) + f'(\mathbf{a})\cdot(\mathbf{b}-\mathbf{a}) + \frac{1}{2}(\mathbf{b}-\mathbf{a})^\top f''(\theta)(\mathbf{b}-\mathbf{a}) \tag{2}$$

Mathematically nonexistent single-point Taylor expansions

The applicability of Taylor expansions in the form of (1) or (2) for real-valued functions cannot be directly extended to vector-valued or matrix-valued functions. This limitation has been well-known in mathematics for centuries but unfortunately has not received sufficient recognition in the statistical community. In the case of vector-valued or matrix-valued functions, the existence of a universal 'intermediate' value, as depicted in simple forms (1) and (2), is generally not guaranteed. Instead, it is essential to consider real-valued Taylor expansions applied to the individual components of the vector- or matrix-valued function.

Our examination of statistical publications reveals that this fundamental mathematical fact has occasionally been overlooked. Consequently, various nonexistent forms of single-point vector- or matrix-valued 'Taylor expansions' have persistently been employed for over 30 years. In this paper, we meticulously analyze four specific types of nonexistent Taylor expansions that have consistently appeared in statistical publications.

Nonexistent taylor expansions for vector-valued functions

Consider a convex open subset $\mathcal{O}$ of $\mathbb{R}^p$, where $p \ge 1$. Let $\mathbf{f}$ be a function mapping $\mathcal{O}$ to $\mathbb{R}^q$, where $q > 1$. We can express $\mathbf{f}$ as $f(x) = [f_1(x), f_2(x), \ldots, f_q(x)]^\top$ for. Here, $f_j(\cdot)$, $j = 1,\ldots,q$, are real-valued functions from $\mathcal{O}$ to $\mathbb{R}$.

The first-order expansion Suppose f is continuously differentiable. Then, for any $\mathbf{a},\mathbf{b} \in \mathcal{O}$ and $j \in \{1,\ldots,q\}$, according to Section 2, there always exists a $\theta^{(j)} \in \overline{\mathbf{ab}}$ such that

$$f_j(\mathbf{b}) = f_j(\mathbf{a}) + f_j'\left(\theta^{(j)}\right)\cdot(\mathbf{b}-\mathbf{a}) \tag{3}$$

Note that $\theta^{(j)}$ in (3) depends on the jth component f_j. Generally, $\theta^{(j)} \ne \theta^{(k)}$ if $j \ne k$. We can rewrite (3) in a compact form

$$\mathbf{f}(\mathbf{b}) = \mathbf{f}(\mathbf{a}) + \begin{bmatrix} f_1'\left(\theta^{(1)}\right) \\ f_2'\left(\theta^{(2)}\right) \\ \vdots \\ f_q'\left(\theta^{(q)}\right) \end{bmatrix}\cdot(\mathbf{b}-\mathbf{a}) \tag{4}$$

This represents a component wise first-order Taylor expansion of $\mathbf{f}$. This expansion is also referred to as the "Mean Value Theorem for General Mappings" in Fitzpatrick.[4] While we have $\mathbf{f}'(\mathbf{x}) = [f_1'(\mathbf{x}), f_2'(\mathbf{x}), \ldots, f_q'(\mathbf{x})]^\top$, it is important to note that in general there is no universal $\theta \in \overline{\mathbf{ab}}$ that satisfies the equation

$$\mathbf{f}(\mathbf{b}) = \mathbf{f}(\mathbf{a}) + \mathbf{f}'(\theta)\cdot(\mathbf{b}-\mathbf{a}) \tag{5}$$

The following Example demonstrates why (5) is generally incorrect if θ is required to lie on $\overline{\mathbf{ab}}$.

Example 1. Consider the vector-valued function

$$\mathbf{f}(x,y) = \begin{bmatrix} f_1(x,y) \\ f_2(x,y) \end{bmatrix} = \begin{bmatrix} x^3 - 3 \\ y^4 - 1 \end{bmatrix}, (x,y) \in \mathcal{O} = (0,\infty)\times(0,\infty)$$

It is easy to see that $\mathbf{f}$ is continuously differentiable on $\mathcal{O}$ with the derivative function

$$\mathbf{f}'(x,y) = \begin{bmatrix} 3x^2 & 0 \\ 0 & 4y^3 \end{bmatrix}, x > 0, y > 0$$

Which is invertible everywhere on $\mathcal{O}$. Let $\mathbf{a} = (1,1)^\top$ and $\mathbf{b} = (2,2)^\top$

We find that (i) the only $\theta^{(1)}$ and $\theta^{(2)}$ on $\overline{\mathbf{ab}}$ and satisfying (4) are

$\theta^{(1)} = \left[\sqrt{21}/3, \sqrt{21}/3\right]^\top$ and $\theta^{(2)} = \left[\sqrt[3]{20}/2, \sqrt[3]{20}/2\right]^\top$;

(ii) the only θ satisfying (5) is

$\theta = \left[\sqrt{21}/3, \sqrt[3]{20}/2\right]^\top$

Which is not on $\overline{\mathbf{ab}}$. However, $\theta \in [[\mathbf{a},\mathbf{b}]]$.

This example clearly demonstrates that (5) does not hold in general for continuously differentiable vector-valued functions when θ is required to be on $\overline{\mathbf{ab}}$. However, the incorrect result (5), commonly referred to as the "first-order Taylor expansion for vector-valued functions" by many statisticians, has been utilized by authors in numerous statistical publications for over 30 years. Since this so-called "first-order Taylor expansion" (5) is non-existent, we categorize it as the "*nonexistent first-order Taylor expansion for vector-valued functions*" (N1TEV).

Citation: Feng C, Wu T, Wang H, et al. A note on the application of Taylor expansions for vector- and matrix-valued functions in statistical inference. *Biom Biostat Int J.* 2026;15(1):74–79. DOI: 10.15406/bbij.2026.15.00451

Copyright: ©2026 Feng et al.

Generalized first-order expansion Suppose $\mathbf{x}$ is partitioned into two parts $\mathbf{x}^\top = \left[\mathbf{x}_1^\top, \mathbf{x}_2^\top\right]$. Let $\mathbf{a}, \mathbf{b} \in \mathcal{O}$ and be partitioned accordingly. We want to know if there exists $t_1, t_2 \in [0,1]$ such that

$$\mathbf{f}(\mathbf{b}_1, \mathbf{b}_2) = \mathbf{f}(\mathbf{a}_1, \mathbf{a}_2) + \mathbf{f}'\left[t_1\mathbf{a}_1 + (1-t_1)\mathbf{b}_1, t_2\mathbf{a}_2 + (1-t_2)\mathbf{b}_2\right] \cdot (\mathbf{b} - \mathbf{a}) \tag{6}$$

The following example again shows that (6) is wrong in general.

Example 2. Consider the vector-valued function

$$\mathbf{f}(x,y,z) = \begin{bmatrix} x^2 - 1 \\ y^3 - 2 \\ z^4 - 3 \end{bmatrix}, (x,y,z) \in \mathcal{O} = (0,\infty) \times (0,\infty) \times (0,\infty)$$

It is easy to see that $\mathbf{f}$ is continuously differentiable on $\mathcal{O}$, and

$$\mathbf{f}'(x,y,z) = \begin{bmatrix} 2x & 0 & 0 \\ 0 & 3y^2 & 0 \\ 0 & 0 & 4z^3 \end{bmatrix}$$

which is invertible everywhere on $\mathcal{O}$. Let $a = (1,1,2)^\top$ and $\mathbf{b} = (2,2,3)^\top$, with $\mathbf{a}_1 = (1,1)^\top, \mathbf{a}_2 = 2, \mathbf{b}_1 = (2,2)^\top$ and $\mathbf{b}_2 = 3$. If (6) were true, then we should have

$$\mathbf{f}(\mathbf{b}) = \mathbf{f}(\mathbf{a}) + \begin{bmatrix} 2(2-t_1) & 0 & 0 \\ 0 & 3(2-t_1)^2 & 0 \\ 0 & 0 & 4(3-t_2)^3 \end{bmatrix} \cdot (\mathbf{b} - \mathbf{a})$$

However, there is no solution to this system of equations.

In (6), the expression $\left[t_1\mathbf{a}_1 + (1-t_1)\mathbf{b}_1, t_2\mathbf{a}_2 + (1-t_2)\mathbf{b}_2\right]$ does not belong to $\overline{\mathbf{ab}}$ unless $t_1 \neq t_2$. Therefore, we refer to (6) as the "*generalized nonexistent first-order Taylor expansion for vector-valued functions*" (GN1TEV). Similar to N1TEV, this GN1TEV has also been observed in numerous publications, as indicated by our survey.

A global negative result for the first-order expansion

In Section 3.1.1, we have demonstrated that the first-order Taylor expansion in the form of (5) does not exist for continuously differentiable vector-valued functions if θ is required to be on $\overline{\mathbf{ab}}$. Part (ii) of Example 1 may suggest the possibility of finding a $\theta \in \left[\left[\mathbf{a}, \mathbf{b}\right]\right]$ that satisfies (5). However, the following theorem establishes that this is not generally true.

Theorem 1. *Suppose $\mathcal{O}$ is a convex open subset of $\mathbb{R}^p$, $(p \geq 1)$ and $\mathbf{f}: \mathcal{O} \to \mathbb{R}^q$ $(q > 1)$ is a continuously differentiable function with the derivative function $\mathbf{f}'$. For any $\mathbf{a}$ and $\mathbf{b}$ in $\mathcal{O}$, there does not exist a $\theta \in \mathcal{O}$ that satisfies equation (5) in general.*

Proof. Consider the function

$$\mathbf{f}(x,y) = \begin{bmatrix} x - \dfrac{1}{y} \\ y^2 - 1 \end{bmatrix}, (x,y) \in \mathcal{O} = (0,\infty) \times (0,\infty)$$

It is easy to see that $\mathbf{f}$ is continuously differentiable on $\mathcal{O}$ with derivative function

$$\mathbf{f}'(x,y) = \begin{bmatrix} 1 & 1/y^2 \\ 0 & 2y \end{bmatrix}, x > 0, y > 0$$

Let $\mathbf{a} = (1,1)^\top$ and $b = (2,1/2)^\top$. If $\theta = (\theta_1, \theta_2)\top$ satisfied equation (5), then we should have

$$\begin{bmatrix} 0 \\ -3/4 \end{bmatrix} = \begin{bmatrix} 1 & 1/\theta_2^2 \\ 0 & 2\theta_2 \end{bmatrix} \begin{bmatrix} 1 \\ -1/2 \end{bmatrix}$$

However, there is no solution to this system of equations. Q.E.D.

Remark 1. Theorem 1 provides a stronger result compared to the findings in Sections 3.1.1 and 3.1.2. It demonstrates that it is not possible to find any θ within the entire domain $\mathcal{O}$, let alone within $\left[\left[\mathbf{a}, \mathbf{b}\right]\right]$, that satisfies equation (5). Therefore, any assertion or utilization of the first-order Taylor expansion in the form of (5) is incorrect.

The second-order expansion

Suppose $\mathbf{f}$ is twice continuously differentiable. For any $\mathbf{a}, \mathbf{b} \in \mathcal{O}$, and $j \in \{1, \ldots, q\}$, there exists a $\theta^{(j)} \in \overline{\mathbf{ab}}$ such that

$$f_j(\mathbf{b}) = f_j(\mathbf{a}) + f_j'(\mathbf{a}) \cdot (\mathbf{b} - \mathbf{a}) + \frac{1}{2}(\mathbf{b} - \mathbf{a})^\top \cdot f_j''\left(\theta^{(j)}\right) \cdot (\mathbf{b} - \mathbf{a}). \tag{7}$$

Similarly, $\theta^{(j)}$ in (7) depends on f_j. Generally, $\theta^{(j)} \neq \theta^{(k)}$ if $j \neq k$.

Let $\mathbf{I}_{q \times q}$ be the $q \times q$ identity matrix. We can also rewrite (7) in a compact form

$$\mathbf{f}(\mathbf{b}) = \mathbf{f}(\mathbf{a}) + \mathbf{f}'(\mathbf{a}) \cdot (\mathbf{b} - \mathbf{a}) + \frac{1}{2}\left(\mathbf{I}_{q \times q} \otimes (\mathbf{b} - \mathbf{a})^\top\right) \cdot \begin{bmatrix} f_1''\left(\theta^{(1)}\right) \\ f_2''\left(\theta^{(2)}\right) \\ \vdots \\ f_q''\left(\theta^{(q)}\right) \end{bmatrix} \cdot (\mathbf{b} - \mathbf{a}) \tag{8}$$

Where $\mathbf{A} \otimes \mathbf{B}$ is the Kronecker product for matrices $\mathbf{A}$ and $\mathbf{B}$. Expression (8) is a *componentwise second-order expansion* of $\mathbf{f}$. Let

$$\mathbf{f}''(\mathbf{x}) = \frac{\partial^2 \mathbf{f}(\mathbf{x})}{\partial \mathbf{x} \partial \mathbf{x}^\top} = \begin{bmatrix} f_1''(\mathbf{x}) \\ f_2''(\mathbf{x}) \\ \vdots \\ f_q''(\mathbf{x}) \end{bmatrix}_{qp \times p}$$

Expression (8) clearly demonstrates that, in general, there is no universal θ belonging to $\overline{\mathbf{ab}}$ such that

$$\mathbf{f}(\mathbf{b}) = \mathbf{f}(\mathbf{a}) + \mathbf{f}'(\mathbf{a}) \cdot (\mathbf{b} - \mathbf{a}) + \frac{1}{2}\left(\mathbf{I}_{q \times q} \otimes (\mathbf{b} - \mathbf{a})^\top\right) \cdot \mathbf{f}''(\theta) \cdot (\mathbf{b} - \mathbf{a}) \tag{9}$$

Let $\mathbf{a} = (a_1, \ldots, a_p)^\top$, $\mathbf{b} = (b_1, \ldots, b_p)^\top$, $x = (x_1, \ldots, x_p)\top$, and

$$\frac{\partial^2 \mathbf{f}(\mathbf{x})}{\partial x_i \partial \mathbf{x}^\top} = \frac{\partial}{\partial x_i}\left[\mathbf{f}'(\mathbf{x})\right], i = 1, \ldots, p$$

Simple algebra shows that (9) is the same as

$$\mathbf{f}(\mathbf{b})=\mathbf{f}(\mathbf{a})+\mathbf{f}'(\mathbf{a})\cdot(\mathbf{b}-\mathbf{a})+\frac{1}{2}\sum_{i=1}^{p}(b_i-a_i)\cdot\left.\frac{\partial^2\mathbf{f}(\mathbf{x})}{\partial x_i\partial\mathbf{x}^\top}\right|_{\mathbf{x}=\theta}\cdot(\mathbf{b}-\mathbf{a}) \tag{10}$$

We need to mention that (10) cannot be written as

$$\mathbf{f}(\mathbf{b})=\mathbf{f}(\mathbf{a})+\mathbf{f}'(\mathbf{a})\cdot(\mathbf{b}-\mathbf{a})+\frac{1}{2}(\mathbf{b}-\mathbf{a})^\top\cdot\left.\frac{\partial^2\mathbf{f}(\mathbf{x})}{\partial\mathbf{x}\partial\mathbf{x}^\top}\right|_{\mathbf{x}=\theta}\cdot(\mathbf{b}-\mathbf{a}) \tag{11}$$

In fact, the matrix multiplication in the second term on the right-hand side of (11) is not possible due to the mismatch in matrix dimensions. The following example serves to illustrate why (9) [and (10)] generally does not hold when θ is required to be within $\mathbf{ab}$.

Example 3. Consider the function $\mathbf{f}$ defined in Example 1, which is twice continuously differentiable on $\mathcal{O}$ with

$$f_1''(x,y)=\begin{bmatrix}6x & 0\\ 0 & 0\end{bmatrix},f_2''(x,y)=\begin{bmatrix}0 & 0\\ 0 & 12y^2\end{bmatrix},x>0,y>0$$

Let $\mathbf{a}=(1,1)^\top$ and $\mathbf{b}=(2,2)^\top$. It is easy to see that (i) the only $\theta^{(1)}$ and $\theta^{(2)}$ on $\mathbf{ab}$ and satisfying (8) are

$\theta^{(1)}=[5/3,5/3]^\top$ and $\theta^{(2)}=\left[\sqrt{102}/6,\sqrt{102}/6\right]^\top$

(ii) the only θ satisfying (9) is

$\theta=\left[5/3,\sqrt{102}/6\right]^\top$

which is not on $\overline{\mathbf{ab}}$. However, $\theta\in\left[[\mathbf{a},\mathbf{b}]\right]$.

Although (9) may resemble (2), they are fundamentally different. In the case of a real-valued differentiable function, (2) is always true, while (9) does not generally hold if θ is required to be within $\overline{\mathbf{ab}}$. Hence, we refer to (9) as the *nonexistent second-order Taylor expansion for vector-valued functions* (N2TEV).

A global negative result for the second-order expansion

Theorem 2. *Suppose $\mathcal{O}$ is a convex open subset of R^p $(p\geq 1)$, and $\mathbf{f}:\mathcal{O}\to\mathbb{R}^q$ $(q>1)$ is a continuously differentiable function with twice continuously differentiable components. Let a and b be points in $\mathcal{O}$. In general, there does not exist any $\theta\in\mathcal{O}$ that satisfies equation (9) for the second-order Taylor expansion.*

Proof. Consider the function

$$\mathbf{f}(x,y)=\begin{bmatrix}x-\dfrac{1}{y}\\ y^3-1\end{bmatrix},(x,y)\in\mathcal{O}=(0,\infty)\times(0,\infty)$$

It is easy to see that $\mathbf{f}$ satisfies the assumptions. Let $a=[1,1]\top$ and $\mathbf{b}=[2,1/2]^\top$. We cannot find any $\theta\in\mathcal{O}$ to make (9) hold. Q.E.D.

Nonexistent first-order taylor expansions for matrix--valued functions

In this section, we examine the nonexistent first-order Taylor expansions for matrix-valued functions, which have also been encountered in statistical literature.

Let $\mathcal{O}$ be a convex open subset of $\mathbb{R}^p$ $(p\geq 1)$, and consider a function $\mathbf{A}:\mathcal{O}\to\mathcal{M}_{m\times n}$, where $\mathcal{M}_{m\times n}$ denotes the set of matrices of size $m\times n$.

Definition 2. The above matrix-valued function $\mathbf{A}$ is said to be differentiable on $\mathcal{O}$ with the derivative function $\mathbf{A}'_{m\times np}$ if

$$\lim_{\backslash\ddot{A}\mathbf{x}\ \backslash\to 0}\frac{\backslash\mathbf{A}(\mathbf{x}+\ddot{A}\mathbf{x})-\mathbf{A}(\mathbf{x})-\mathbf{A}'(\mathbf{x})(\mathbf{I}_{n\times n}\otimes\ddot{A}\mathbf{x})\ \backslash}{\backslash\ddot{A}\mathbf{x}\ \backslash}=0,\mathbf{x}\in\mathcal{O} \tag{12}$$

This definition aligns with that of vector-valued functions (when $n=1$) as discussed in mathematical analysis textbooks, such as Rudin (1976).

We can write $\mathbf{A}(\mathbf{x})$ as $\mathbf{A}(\mathbf{x})=\left[\mathbf{A}_1(\mathbf{x}),\mathbf{A}_2(\mathbf{x}),\ldots,\mathbf{A}_n(\mathbf{x})\right]$, where $\mathbf{A}_j,j=1,\ldots,n$ are $m\times 1$ vector-valued functions. It is easy to prove that

$$\mathbf{A}'(\mathbf{x})=\left[\frac{\partial\mathbf{A}_1(\mathbf{x})}{\partial\mathbf{x}^\top},\ldots,\frac{\partial\mathbf{A}_n(\mathbf{x})}{\partial\mathbf{x}^\top}\right]=\left[\mathbf{A}_1'(\mathbf{x}),\ldots,\mathbf{A}_n'(\mathbf{x})\right]$$

Theorem 3. *Suppose $\mathbf{A}$ as defined above is continuously differentiable with the derivative function $\mathbf{A}'$. Let $\mathbf{a},\mathbf{b}\in\mathcal{O}$. In general, there does not exist a $\theta\in\mathcal{O}$ such that*

$$\mathbf{A}(\mathbf{b})=\mathbf{A}(\mathbf{a})+\mathbf{A}'(\theta)\left(\mathbf{I}_{n\times n}\otimes(\mathbf{b}-\mathbf{a})\right) \tag{13}$$

Proof. Consider the function

$$\mathbf{A}(x,y)=\begin{bmatrix}x-\dfrac{1}{y} & 0\\ 0 & y^3\end{bmatrix},(x,y)\in\mathcal{O}=(0,\infty)\times(0,\infty)$$

It is easy to see that $\mathbf{A}$ is continuously differentiable on $\mathcal{O}$. Let $a=[1,1]\top$ and $\mathbf{b}=[2,1/2]^\top$. We cannot find a $\theta\in\mathcal{O}$ satisfying (13). Q.E.D.

Theorem 3 provides a global negative result for the first-order Taylor expansion of matrix-valued functions. It demonstrates that even if $\mathbf{A}$ is continuously differentiable, it is generally impossible to find any θ in the entire domain $\mathcal{O}$, let alone on $\overline{\mathbf{ab}}$, that satisfies equation (13). Therefore, any assertion of utilizing the first-order Taylor expansion in the form of (13) is incorrect. We refer to (13) as the "*nonexistent first-order Taylor expansion for matrix-valued functions*" (N1TEM).

Suppose $\mathbf{A}_{m\times m}(\mathbf{x})$ is invertible and continuously differentiable on $\mathcal{O}$. Let $\mathbf{A}^{-1'}$ denote the derivative function of $\mathbf{A}^{-1'}$. Using the fact that $\mathbf{A}^{-1}_{m\times m}(\mathbf{x}+\ddot{A}\mathbf{x})\mathbf{A}_{m\times m}(\mathbf{x}+\ddot{A}\mathbf{x})=\mathbf{I}_{m\times m}$, we can prove that $\mathbf{A}^{-1'}$ and $\mathbf{A}'$ satisfy

$$\lim_{\backslash\ddot{A}\mathbf{x}\ \backslash\to 0}\frac{\backslash\mathbf{A}^{-1'}(\mathbf{x})\left(\mathbf{I}_{m\times m}\otimes\ddot{A}\mathbf{x}\right)+\mathbf{A}^{-1}(\mathbf{x})\mathbf{A}'(\mathbf{x})\left(\mathbf{I}_{m\times m}\otimes\ddot{A}\mathbf{x}\right)\mathbf{A}^{-1}(\mathbf{x})\ \backslash}{\backslash\ddot{A}\mathbf{x}\ \backslash}=0 \tag{14}$$

Theorem 4. *Suppose $\mathbf{A}_{m\times m}$ is invertible and continuously differentiable on $\mathcal{O}$ with the derivative function $\mathbf{A}'$. Let $\mathbf{a},\mathbf{b}\in\mathcal{O}$. There does not in general exist a $\theta\in\mathcal{O}$ such that*

$$\mathbf{A}^{-1}(\mathbf{b})=\mathbf{A}^{-1}(\mathbf{a})-\mathbf{A}^{-1}(\theta)\mathbf{A}'(\theta)\left(\mathbf{I}_{m\times m}\otimes(\mathbf{b}-\mathbf{a})\right)\mathbf{A}^{-1}(\theta) \tag{15}$$

Proof. Consider the function

$$\mathbf{A}(x,y)=\begin{bmatrix}x & y\\ 0 & y^2\end{bmatrix},(x,y)\in\mathcal{O}=(0,\infty)\times(0,\infty)$$

It is easy to see that $\mathbf{A}$ is invertible and continuously differentiable on $\mathcal{O}$. Let $a=[1,1]^\top$ and $\mathbf{b}=[2,1/2]^\top$. We cannot find a $\theta\in\mathcal{O}$ satisfying (15). Q.E.D.

Theorem 4 presents a global negative result for the first-order expansion of the inverse of continuously differentiable matrix-valued functions. It demonstrates that even if all the conditions are satisfied by $\mathbf{A}$, it is generally impossible to find any θ in the entire domain $\mathcal{O}$, let alone on $\mathbf{ab}$, that fulfills equation (15). Therefore, any assertion of using the first-order Taylor expansion in the form of (15) is incorrect. We refer to (15) as the "*second nonexistent first-order Taylor expansion for matrix-valued functions*" (N1TEM2).

Nonexistent taylor expansions in statistical publications and asymptotic reassessment

In Section 3, we examined various versions of Taylor expansions that do not exist mathematically. Regrettably, these nonexistent single-point Taylor expansions have been utilized in the theoretical frameworks of numerous publications featured in reputable statistical journals. Below we review representative examples and reassess their theoretical conclusions.

Analysis of literature examples

Example 4. Fan, Xue, and Zou.[5] AOS, Vol 42, pp819–849.

The following statement appears in the proof of their Theorem 5.4 on p.840.

"By a Taylor expansion,

$$\nabla\ell_n\left(\hat{\beta}^{\text{oracle}}\right)=\nabla\ell_n\left(\hat{\beta}^{*}\right)+\nabla^2\ell_n\left(\hat{\beta}^{*}\right)\cdot\hat{\mathrm{A}}+\left(\nabla^2\ell_n\left(\tilde{\beta}\right)-\nabla^2\ell_n\left(\hat{\beta}^{*}\right)\right)\cdot\hat{\mathrm{A}} \tag{16}$$

Where $\tilde{\beta}$ is on the line segment joining $\hat{\beta}^{\text{oracle}}$ and $\hat{\beta}^{*}$."

Since $\nabla\ell_n(\beta)$ is a $p\times 1$ vector-valued function and $\hat{\mathrm{A}}=\hat{\beta}^{\text{oracle}}-\hat{\beta}^{*}$, equation (16) is equivalent to N1TEV (5).

Example 5. Ma.[6] AOS, Vol 40, pp2943–2972.

In the proof of Theorem 2 on p.2963, equation (18) uses GN1TEV (6):

$$\mathbf{g}_n\left(\hat{\beta}_n,\hat{\gamma}_n\right)-\mathbf{g}_n\left(\beta_0,\gamma_0\right)=-\mathcal{D}_n\left(\beta^{*},\gamma^{*}\right)\begin{pmatrix}\hat{\beta}_n-\beta_0\\ \hat{\gamma}_n-\gamma_0\end{pmatrix} \tag{18}$$

where $\beta^{*}=t_1\hat{\beta}_n+(1-t_1)\beta_0$ and $\gamma^{*}=t_2\hat{\gamma}_n+(1-t_2)\gamma_0$ for some $t_1,t_2\in(0,1)$.

Example 6. Wang, Xue, Qu, and Liang.[7] AOS, Vol 42, pp592–624.

In the proof of Theorem 2 on p.616, equation (19) utilizes N2TEV (9)

$$\dot{Q}_n\left(\hat{\theta}^{\text{QIF}}\right)-\dot{Q}_n\left(\tilde{\theta}_0\right)=\ddot{Q}_n\left(\tilde{\theta}_0\right)\left(\hat{\theta}^{\text{QIF}}-\tilde{\theta}_0\right)+\frac{1}{2}\left(\hat{\theta}^{\text{QIF}}-\tilde{\theta}_0\right)^{\top}\left.\frac{\partial\dot{Q}_n(\theta)}{\partial\theta\partial\theta^{\top}}\right|_{\theta=\theta^{*}}\left(\hat{\theta}^{\text{QIF}}-\tilde{\theta}_0\right) \tag{19}$$

Example 7. Choi, Jeong, and Park.[8]

In formula (B.14) on p.553, N1TEM2 (15) is used to expand matrix inverse Hessians $\mathcal{H}_j^{-1}\left(\tilde{\theta}_j\right)$.

Reassessment of asymptotic results

A critical question raised during peer review is whether these single-point notation errors invalidate the main asymptotic theorems of the cited studies. We emphasize that in asymptotic statistical theory, the primary goal of local linearization is to establish limiting distributions for $n_{1/2}(\theta_j-\tilde{\theta}_0)$.When authors write single-point mean-value expansions, they are typically using Lagrange notation as shorthand for controlling remainder terms. When these single-point expressions are replaced with valid integral-form Taylor expansions (Section 5), the integral operators integrate continuous derivatives over consistent neighborhoods. Because estimators such as $\hat{\beta}^{\text{oracle}}$ or $\hat{\theta}^{\text{QIF}}$ are consistent ($\|\tilde{\theta}-\tilde{\theta}_0\|=O_p(1)$), the integrated Hessian or score derivative matrices converge in probability to their expected limit matrices $E[H_j^{(-1)}(\theta_0)]$ Consequently, the stochastic bounds $O_p(n^{-1/2})$ and limiting normal distributions remain completely valid. Thus, while single-point mean-value notation represents a technical mathematical flaw in finite-sample statements, it does not alter the core scientific conclusions of these published works.

Fixing the mistakes: exact component wise and integral-form expansions

To maintain exact mathematical rigor without introducing nonexistent single-point mean values, researchers can utilize two well-established formulations from mathematical analysis.[3,4]

Component wise expansions

For a continuously differentiable vector function $f:\mathcal{O}\to\mathbb{R}^q$, exact equality is achieved by evaluating each component at its own intermediate point as in Equation (4).

Integral-form taylor expansions

Integral representations[3] avoid multiple intermediate points entirely by integrating along the line segment connecting a and b:

1. First-order expansion for vector-valued functions:

$$\mathbf{f}(\mathbf{b})=\mathbf{f}(\mathbf{a})+\int_0^1\mathbf{f}'\left[\mathbf{a}+t(\mathbf{b}-\mathbf{a})\right]dt\cdot(\mathbf{b}-\mathbf{a})$$

2. Second-order expansion for vector-valued functions:

$$\mathbf{f}(\mathbf{b})=\mathbf{f}(\mathbf{a})+\mathbf{f}'(\mathbf{a})\cdot(\mathbf{b}-\mathbf{a})+\frac{1}{2}\left(\mathbf{I}_{q\times q}\otimes(\mathbf{b}-\mathbf{a})^{\top}\right)\int_0^1\mathbf{f}''\left[\mathbf{a}+t(\mathbf{b}-\mathbf{a})\right]dt\cdot(\mathbf{b}-\mathbf{a})$$

3. First-order expansion for matrix-valued functions:

$$\mathbf{A}(\mathbf{b})=\mathbf{A}(\mathbf{a})+\int_0^1\mathbf{A}'\left[\mathbf{a}+t(\mathbf{b}-\mathbf{a})\right]dt\cdot\left(\mathbf{I}_{n\times n}\otimes(\mathbf{b}-\mathbf{a})\right)$$

4. First-order expansion for the inverse of matrix-valued functions:

$$\mathbf{A}^{-1}(\mathbf{b})=\mathbf{A}^{-1}(\mathbf{a})-\int_0^1\mathbf{A}^{-1}\left[\mathbf{a}+t(\mathbf{b}-\mathbf{a})\right]\mathbf{A}'\left[\mathbf{a}+t(\mathbf{b}-\mathbf{a})\right]\left(\mathbf{I}_{m\times m}\otimes(\mathbf{b}-\mathbf{a})\right)\mathbf{A}^{-1}\left[\mathbf{a}+t(\mathbf{b}-\mathbf{a})\right]dt$$

All these formulas provide valid, exact Taylor expansions. By utilizing these correct integral expansions, statistical proofs derive asymptotic distributions with flawless mathematical precision.

Concluding remarks

Precision in mathematical notation is essential for maintaining scientific rigor in statistical methodology. Over the past three decades, numerous influential works in top-tier journals have utilized single-point vector or matrix Taylor expansions. In addition to the studies cited in Examples 1–7, additional foundational contributions appear across leading statistical and econometric journals: Hansen,[9] Hu et al.,[10] Kong, Cai, and Sen,[11] Liang et al.,[12] Ma, Genton, and Tsiatis,[13] McCullagh,[14] Murphy et al.,[15] Park and Phillips,[16] Qu, Lee, and Lindsay,[17] Song, Fan, and Kalbfleisch,[18] Yang and Prentice,[19] and Zhang, Duchi, and Wainwright.[20]

Although direct single-point Lagrange expansions do not exist in general for vector- and matrix-valued functions, adopting integral-form expansions or explicit component-wise evaluations fully

resolves these theoretical gaps. Reassessing the literature confirms that the asymptotic properties established in these works remain robust under rigorous formulations. Promoting precise mathematical representations in research and graduate education will further strengthen methodological clarity and safeguard research integrity across the statistical sciences.

Acknowledgments

None.

Conflicts of interest

The authors declare that there are no conflicts of interest.

References

1. Feng C, Wang H, Chen T, et al. On exact forms of Taylor's theorem for vector-valued functions. *Biometrika*. 2014;101(4):1003.
2. Rudin W. *Principles of Mathematical Analysis*. 3rd edition. McGraw-Hill; 1976.
3. Lang S. *Real and Functional Analysis*. 3rd edition. Springer-Verlag; 1993.
4. Fitzpatrick PM. *Advanced Calculus: A Course in Mathematical Analysis*. 2nd edition. Brooks Cole; 2006.
5. Fan J, Xue L, Zou H. Strong oracle optimality of folded concave penalized estimation. *Ann Stat*. 2014;42(2):819–849.
6. Ma S. Two-step spline estimating equations for generalized additive partially linear models with large cluster sizes. *Ann Stat*. 2012;40(6):2943–2972.
7. Wang L, Xue L, Qu A, et al. Estimation and model selection in generalized additive partial linear models for correlated data with diverging number of covariates. *Ann Stat*. 2014;42(2):592-624.
8. Choi H, Jeong M, Park JY. An asymptotic analysis of likelihood-based diffusion model selection using high frequency data. J Econom. 2014;178:539–557.
9. Hansen LP. Large sample properties of generalized method of moments estimators. *Econometrica*. 1982;50(4):1029–1053.
10. Hu T, Nan B, Lin X, et al. Time-dependent cross-ratio estimation for bivariate failure times. Biometrika. 2011;98(2):341–354.
11. Kong L, Cai J, Sen PK. Weighted estimating equations for semiparametric transformation models with censored data from a case-cohort design. *Biometrika*. 2004;91(2):305–319.
12. Liang H, Liu X, Li R, et al. Estimation and testing for partial linear single-index models. Ann Stat. 2010;38(6):3811–3836.
13. Ma Y, Genton MG, Tsiatis AA. Locally efficient semiparametric estimators for generalized skew-elliptical distributions. *J Am Stat Assoc*. 2005;100(471):980–989.
14. McCullagh P. Quasi-likelihood functions. *Ann Stat*. 1983;11(1):59–67.
15. Murphy SA, van der Laan MJ. Marginal mean models for dynamic regimes. *J Am Stat Assoc*. 2001;96(456):1410–1423.
16. Park JY, Phillips PCB. Nonlinear regression with integrated time series. *Econometrica*. 2001;69(1):117–161.
17. Qu A, Lee JJ, Lindsay BG. Model diagnostic tests for selecting informative correlation structure in correlated data. *Biometrika*. 2008;95(4):891–905.
18. Song PXK, Fan Y, Kalbfleisch JD. Maximization by parts in likelihood inference. *J Am Stat Assoc*. 2005;100(472):1145–1158.
19. Yang S, Prentice R. Semiparametric analysis of short-term and long-term hazard ratios with two-sample survival data. *Biometrika*. 2005;92(1):1–17.
20. Zhang Y, Duchi JC, Wainwright MJ. Communication-efficient algorithms for statistical optimization. *J Mach Learn Res*. 2013;14:3321–3363.