Three proofs of an identity involving derivatives of a positive definite matrix and its determinant

Bai-Ni Guo
Department of Applied Mathematics and Informatics
Jiaozuo Institute of Technology,
Jiaozuo City, Henan 454000
The People's Republic of China
e-mail: guobaini@jzit.edu.cn

ABSTRACT
In the paper, three proofs for an identity involving derivatives of a positive definite matrix and its determinant are given using technique of linear algebra. The identity is basic in differential geometry.

Key words and phrases: Identity, positive definite matrix, determinant, derivative

1 Introduction
Let $M$ be an $n$-dimensional, $n \leq 1$, connected, $C^\infty$, Riemannian manifold. For definition of manifold, please refer to standard texts [1, 4]. The Riemannian metric on $M$ associates to each $p \in M$ an inner product on $M_p$, which we denote by $(\cdot, \cdot)$. The associated norm will be denoted by $|\cdot|$. The Riemannian metric is $C^\infty$ in the sense that if $X, Y$ are $C^\infty$ vector fields on $M$, then $(X, Y)$ is a $C^\infty$ real-valued function on $M$. 

*The author was supported in part by NNSF (#10001016) of China, SF for the Prominent Youth of Henan Province, NSF of Henan Province (#004051800), SF for Pure Research of Natural Science of the Education Department of Henan Province (#1999110004), Doctor Fund of Jiaozuo Institute of Technology, China.
Let \( U \) be an open set in \( M \), and \( x : U \to \mathbb{R}^n \) a diffeomorphism of \( U \) into \( \mathbb{R}^n \), that is, a chart on \( M \). Then associated to the chart are \( n \) coordinate vector fields, written as \( \partial / \partial x^j \) or as \( \partial_j \), \( j = 1, \ldots, n \).

For the given Riemannian metric, define

\[
g_{jk} = \langle \partial_j, \partial_k \rangle, \quad g = \det G, \quad G^{-1} = (g^{jk})_{1 \leq j, k \leq n},
\]

where \( j, k = 1, \ldots, n \), \( \det G \) and \( G^{-1} \) denote the determinant and the inverse of \( G \) respectively. It is well-known that \( G \) is a positive definite matrix. See [2, pp. 3–7].

The following identity involving derivatives of a positive definite matrix and its determinant is fundamental in differential geometry.

**Theorem 1** For \( 1 \leq j \leq n \), we have

\[
\text{tr}(G^{-1} \partial_j G) = \partial_j (\ln g). \tag{1}
\]

In this short note, we will give three proofs of the identity (1) using different technique of linear algebra. For concepts of linear algebra, please refer to [3].

## 2 Three proofs of identity (1)

**First proof.** Since the metric matrix \( G = (g_{ij}) \) is a positive definite matrix, then we can assume its eigenvalues of \( G \) are \( \lambda_i > 0 \), \( i = 1, \ldots, n \). From theory of linear algebra, we have

\[
g = \det G = |G| = \prod_{i=1}^{n} \lambda_i, \tag{2}
\]

\[
\ln g = \sum_{i=1}^{n} \ln \lambda_i, \tag{3}
\]

\[
\partial_j (\ln g) = \sum_{i=1}^{n} \frac{\partial_j \lambda_i}{\lambda_i}, \tag{4}
\]

where \( j = 1, \ldots, n \).

Further, there is an orthogonal matrix \( P \) such that

\[
P^{-1}GP = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \Lambda, \tag{5}
\]
Three proofs of an identity involving ...

therefore, we have \( G = P\Lambda P^{-1}, \ G^{-1} = P\Lambda^{-1}P^{-1}, \) and

\[
\partial_j G = \partial_j(P\Lambda P^{-1}) = (\partial_j P)\Lambda P^{-1} + P(\partial_j \Lambda)P^{-1} + P\Lambda(\partial_j(P^{-1})),
\]

\[
G^{-1}(\partial_j G) = (P\Lambda^{-1}P^{-1})(\partial_j P)\Lambda P^{-1} + (P\Lambda^{-1}P^{-1})(P(\partial_j \Lambda)P^{-1})
\]

\[
+ (P\Lambda^{-1}P^{-1})(P\Lambda(\partial_j(P^{-1}))) = P\Lambda^{-1}(P^{-1}\partial_j P)\Lambda P^{-1} + P(\Lambda^{-1}\partial_j \Lambda)P^{-1} + P\partial_j(P^{-1}).
\]

From \( P^{-1}P = E, \) it follows that \((\partial_j(P^{-1}))P + P^{-1}(\partial_j P) = 0, \) thus

\[
G^{-1}(\partial_j G) = -(P\Lambda^{-1})[(\partial_j(P^{-1}))P](P\Lambda^{-1})^{-1} + P(\Lambda^{-1}\partial_j \Lambda)P^{-1} + P\partial_j(P^{-1})
\]

\[
= -(P\Lambda^{-1}P^{-1})P[(\partial_j(P^{-1}))P]P^{-1}(P\Lambda^{-1}P^{-1})^{-1}
\]

\[
+ P(\Lambda^{-1}\partial_j \Lambda)P^{-1} + P\partial_j(P^{-1})
\]

\[
= -G(P\partial_j(P^{-1}))G^{-1} + P(\Lambda^{-1}\partial_j \Lambda)P^{-1} + P\partial_j(P^{-1}).
\]

Using the formulae \( \text{tr}(AB) = \text{tr}(BA), \text{tr}(P^{-1}AP) = \text{tr} A, \) and \( \text{tr}(A+B) = \text{tr} A + \text{tr} B, \) we obtain

\[
\text{tr}[G^{-1}(\partial_j G)] = \text{tr}(P\partial_j(P^{-1})) + \text{tr}[P(\Lambda^{-1}\partial_j \Lambda)P^{-1}] - \text{tr}[G(P\partial_j(P^{-1}))G^{-1}]
\]

\[
= \text{tr}(\Lambda^{-1}\partial_j \Lambda)
\]

\[
= \sum_{i=1}^{n} \frac{\partial_j \lambda_i}{\lambda_i}
\]

\[
= \partial_j(\ln g).
\]

The proof is complete.

\[\text{Remark 1} \quad \text{In fact, we have obtained the following}
\]

\[
\text{tr}(G^{-1}\partial_j G) = \text{tr}[(\partial_j G)G^{-1}] = \partial_j(\ln |G|) = \partial_j(\ln g).
\]

\[\text{Second proof.} \quad \text{We partition the matrix} \ G \ \text{by columns, that is}
\]

\[
G = (\alpha_1, \ldots, \alpha_n),
\]

\[
\alpha_i = \begin{pmatrix} g_{1i} \\ \vdots \\ g_{ni} \end{pmatrix}
\]
where $1 \leq i \leq n$. Then we have

$$\partial_j (\ln g) = \partial_j \ln |G| = \frac{\partial_j |G|}{|G|},$$

where

$$\partial_j |G| = \partial_j |\alpha_1, \ldots, \alpha_n| = \sum_{i=1}^{n} |\alpha_1, \ldots, \alpha_{i-1}, \partial_j \alpha_i, \alpha_{i+1}, \ldots, \alpha_n|,$$

$$\partial_j \alpha_i = \begin{pmatrix} \partial_j g_{1i} \\ \vdots \\ \partial_j g_{ni} \end{pmatrix}, \quad i = 1, 2, \ldots, n.$$  

The Laplace expansion yields

$$\partial_j |G| = \sum_{i=1}^{n} \sum_{k=1}^{n} (\partial_j g_{ki}) G_{ki},$$

where $G_{ki} = G_{ik}$ is the cofactor of the element $g_{ik} = g_{ki}$ in symmetric matrix $G^T = G$. Hence

$$\partial_j (\ln g) = \frac{1}{|G|} \sum_{i,k=1}^{n} (\partial_j g_{ki}) G_{ki},$$  

Moreover, since $\partial_j G = (\partial_j g_{ik})$ and $G^{-1} = \frac{G^*}{|G|} = \frac{(G_{ik})}{|G|}$, where $G^*$ denotes the adjoint of $G$, we have

$$\text{tr}(G^{-1} \partial_j G) = \text{tr} \left( \frac{(G_{ik})(\partial_j g_{ik})}{|G|} \right) = \frac{1}{|G|} \sum_{i=1}^{n} (G_{i1}, \ldots, G_{in}) \begin{pmatrix} \partial_j g_{1i} \\ \vdots \\ \partial_j g_{ni} \end{pmatrix}$$

$$= \frac{1}{|G|} \sum_{i,k=1}^{n} G_{ik} (\partial_j g_{ki}),$$

the identity $\text{tr}(G^{-1} \partial_j G) = \partial_j (\ln |G|)$ follows.

**Remark 2** For arbitrary square matrix $A$ of order $n$, if $|A| > 0$, its element $a_{ij}$ is a function of $x$, then

$$\frac{d(\ln |A|)}{dx} = \text{tr} \left( A^{-1} \frac{dA}{dx} \right) = \text{tr} \left( \frac{dA}{dx} A^{-1} \right).$$
Remark 3 Let $A(t)$ is an invertible differentiable matrix, then
\[
(\det A)' = (\det A) \, \text{tr}(A^{-1}A'),
\]
where $A'$ denotes the derivative of matrix $A$ with respect to $t$.

**Third proof.** Let $G^* = (G_{ij})$ denote the adjoint of the positive definite matrix $G$, then $G_{ij} = G_{ji}$, and
\[
\text{tr}(G^{-1} \partial_j G) = \frac{\text{tr}(G^* \partial_j G)}{|G|} = \frac{1}{g} \sum_{i,k=1}^{n} G_{ik}(\partial_j g_{ki}),
\]
\[
\partial_j (\ln g) = \frac{\partial_j g}{g} = \frac{\partial_j |G|}{g} = \frac{1}{g} \partial_j \sum_{t=1}^{n} g_{1t}G_{1t} = \frac{1}{g} \sum_{t=1}^{n} [(\partial_j g_{1t})G_{1t} + g_{1t} \partial_j G_{1t}].
\]

The proof reduces to prove that
\[
\sum_{t=1}^{n} g_{1t}(\partial_j G_{1t}) = \sum_{i=2}^{n} \sum_{t=1}^{n} (\partial_j g_{it})G_{it}.
\]
In fact, we have
\[
\sum_{t=1}^{n} g_{kt}(\partial_j G_{kt}) = \sum_{i \neq k}^{n} \sum_{t=1}^{n} (\partial_j g_{it})G_{it}, \quad k = 1, 2, \ldots, n.
\]
This completes the proof.

**References**


