Linear Algebra and its Applications 548 (2018) 293–304

Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa

Differentiating the pseudo determinant ☆

Andrew Holbrook

Department of Statistics, University of California, Irvine, United States

A R T I C L E I N F O

Article history:
Received 13 February 2018
Accepted 8 March 2018
Available online 13 March 2018
Submitted by R. Brualdi

MSC:
primary 15A15
secondary 62H12

Keywords:
Pseudo determinant
Pseudo inverse
Maximum likelihood
Degenerate Gaussian
Singular covariance

A B S T R A C T

A class of derivatives is defined for the pseudo determinant $\mathrm{Det}(A)$ of a Hermitian matrix A. This class is shown to be non-empty and to have a unique, canonical member $\nabla\mathbf{Det}(A) = \mathrm{Det}(A)A^+$, where A^+ is the Moore–Penrose pseudo inverse. The classic identity for the gradient of the determinant is thus reproduced. Examples are provided, including the maximum likelihood problem for the rank-deficient covariance matrix of the degenerate multivariate Gaussian distribution.

1. Introduction

We derive the class of derivatives of the pseudo determinant with respect to Hermitian matrices, placing an emphasis on understanding the forms taken by this class and their relationship to established results in linear algebra. In particular, care must be taken to address the discontinuous nature of the pseudo derivative. The contributions in

☆ This work was supported by the UC Irvine Graduate Dean's Dissertation Fellowship. I thank Professor Oliver Knill for his encouragement and helpful discussion, and I am grateful to an anonymous reviewer for further helpful recommendations.

E-mail address: aholbroo@uci.edu.

https://doi.org/10.1016/j.laa.2018.03.018
0024-3795/

this paper are primarily of a linear algebraic nature but are well motivated in fields of application.

The pseudo determinant arises in graph theory within Kirchoff's matrix tree theorem [1] and in statistics, in the definition of the degenerate Gaussian distribution. The degenerate Gaussian has been useful in image segmentation [2], communications [3], and as the asymptotic distribution for multinomial samples [4]. Despite these appearances, knowledge of how to differentiate the distribution's density function is conspicuously absent from the literature, and—since differentiation is often essential for maximization—the lack of this knowledge is a plausible barrier to the distribution's wider use.

Specifically, to obtain the maximum likelihood (ML) estimator for the singular covariance matrix of the degenerate Gaussian, one must be able to calculate the derivative of the log likelihood and hence the pseudo determinant of the covariance. Although [5] firmly establishes the subject of ML estimation for multivariate Gaussians, the authors never directly address singular covariance estimation. This problem is explored in Section 3. In Section 2, the pseudo determinant is introduced, and its derivative with respect to Hermitian matrices is derived.

2. The canonical derivative

We begin by introducing the pseudo determinant both as a product of eigenvalues and as a limiting form.

Definition 2.1. The pseudo determinant Det of a square matrix A is defined as the product of its non-zero eigenvalues. If a matrix has no non-zero eigenvalues, then we say $\mathrm{Det}(0) = 1$.

See [1] for an equivalent definition of the pseudo determinant in terms of the characteristic polynomial. In deriving its derivative, it will be useful to write the pseudo determinant as a limit.

Proposition 2.2. *If A is an $n \times n$ matrix of rank k, then* $\mathrm{Det}(A)$ *is the limit*

$$\mathrm{Det}(A) = \lim_{\delta \to 0} \frac{\det(A + \delta I)}{\delta^{n-k}} \tag{2.1}$$

for $\det(\cdot)$ *the regular determinant.*

Whereas this result is known [6], we were unable to find its proof, so it is given here in the spirit of completeness.

Proof. We use the identity

$$\det(X + ZYZ^*) = \det(Y^{-1} + Z^* X^{-1} Z)\, \det(Y)\, \det(X)\,. \tag{2.2}$$