QÜESTIIÓ, vol. 25, 2, p. 187-210, 2001

SOME APPLICATIONS OF THE MATRIX HAFFIAN IN CONNECTION WITH DIFFERENTIABLE MATRIX FUNCTIONS OF A CENTRAL WISHART VARIATE⋆

HEINZ NEUDECKER
Cesaro*

In this paper we revisit Haff's seminal work on the matrix Haffian as we proposed to call it. We review some results, and give new derivations. Use is made of the link between the matrix Haffian ∇F *and the differential of the matrix function,* dF.

Keywords: Kronecker product, commutation matrix, Hadamard product, matrix differentiation, matrix differentials, matrix partitioning

AMS Classification (MSC 2000): primary 62F0, secondary 62C99

⋆This research was supported by DGES.
*Oosterstraat, 13. 1741 GH Schagen. The Netherlands. E-mail: heinz@fee.uva.nl
–Received January 2001.
–Accepted April 2001.

1. INTRODUCTION

In the early eighties of last century Haff (1981, 1982) published seminal work on what I recently proposed to call the matrix Haffian. See Neudecker (2000*b*). Haff applied this matrix to various multivariate problems involving central Wishart variates. Relevant is a differentiable *square* matrix function $F(X)$, shortly F, which depends on a *symmetric* matrix X. Both matrices have the same dimension.

A strategic rôle is being played by a square matrix $\nabla = (d_{ij})$ of operators $d_{ij} := \frac{1}{2}(1+\delta_{ij}) \frac{\partial}{\partial x_{ij}}$, where δ_{ij} is the Kronecker delta ($\delta_{ii} = 1, \delta_{ij} = 0$ when $i \neq j$). Haff used the symbol D, not ∇. The matrix ∇ applied to F yields the matrix Haffian ∇F. In parallel work on the kindred *scalar* Haffian I proposed to use the symbol ∇ (Neudecker, 2000*a*) in order to avoid confusion with the so-called duplication matrix which naturally cropped up in that context. Neudecker (2000*b*) presented a link between ∇F and dF, the differential of F.

Haff (1981) gave a fundamental identity based on the matrix Haffian involving a differentiable, not necessarily square, matrix function whose argument was a central Wishart variate. This Fundamental Identity (FI) was used to find expected values of occasionally complicated functions of a central Wishart variate. See also Haff (1982) for further results.

In the present paper we shall revisit Haff's seminal oeuvres, review some of his results, and give new derivations using the link between ∇F and dF.

We shall also consider other applications, drawing heavily on work by Legault-Giguère (1974), Giguère & Styan (1978) and Styan (1989).

2. THE FUNDAMENTAL IDENTITY

Haff (1981, Section 2, (4)) presents the following Fundamental Identity (FI) which holds under mild conditions on the input matrix, viz

$$\mathcal{E} F_1 \Sigma^{-1} F_2 = 2\mathcal{E} F_1 \nabla F_2 + 2\left(\mathcal{E} F_2' \nabla F_1'\right)' + (n-m-1)\mathcal{E} F_1 S^{-1} F_2 \qquad (1)$$

with $S \sim W_m(\Sigma, n)$, $n > m+1$ and $F_i := F_i(S)$ $(i = 1,2)$. As usual $\mathcal{E}$ is the expectation operator.

In Haff's presentation $F_1(F_2)$ is of dimension $p \times m$ $(m \times q)$. We shall have $p = q = m$, hence F_1, F_2, S and ∇ are all square of dimension m. This will do for our purposes.

3. THE LINK BETWEEN ∇F AND dF

In Neudecker (2000*b*) the following theorem was proved.

Theorem 1

For the differentiable matrix function $F(X)$ of symmetric X:

$$dF = P'(dX)Q \quad \textit{implies} \quad \nabla F = \frac{1}{2}PQ + \frac{1}{2}(tr P)Q,$$

where dF and dX are differentials of F and X.

In the sections to follow we shall apply Haff's FI and our Theorem 1 to a wide collection of matrix functions of a central Wishart variate. We shall therefore use S instead of X to denote the argument matrix. See Magnus and Neudecker (1999) on matrix differentials.

4. APPLICATIONS I

In this section we reconsider results given by Haff (1981). We shall occasionally use partitioned matrix Haffians. These were also developed by Haff (1981, Section 2). For a survey see the Appendix of this paper.

Theorem 2

$$\mathcal{E}S_{11\cdot 2} = (n - m_2)\Sigma_{11\cdot 2} \quad \textit{and} \quad \mathcal{E}S_{22}^{-1}S_{21} = \Sigma_{22}^{-1}\Sigma_{21},$$

where $S_{11\cdot 2} := S_{11} - S_{12}S_{22}^{-1}S_{21}$, $S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}$, $m_1 \times m_1$ is the dimension of S_{11}, and $m_2 := m - m_1$.

$\Sigma_{11\cdot 2}$ is defined accordingly.

Proof

We take $F_1 = I_m$ and $F_2 = \begin{pmatrix} S_{11\cdot 2} & 0 \\ 0 & 0 \end{pmatrix}$. It is known that

$$S^{-1} = \begin{pmatrix} S_{11\cdot 2}^{-1} & -S_{11\cdot 2}^{-1}S_{12}S_{22}^{-1} \\ -S_{22}^{-1}S_{21}S_{11\cdot 2}^{-1} & S_{22\cdot 1}^{-1} \end{pmatrix}.$$

Further $S_{22\cdot 1}$ and Σ^{-1} are expressed analogously to $S_{11\cdot 2}$ and S^{-1}.

Haff's FI in partitioned form yields two equations, viz

(i) $\Sigma_{11\cdot 2}^{-1}\mathcal{E}S_{11\cdot 2} = (m_1+1)I_{m_1} + (n-m-1)I_{m_1}$

(ii) $\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11\cdot 2}^{-1}\mathcal{E}S_{11\cdot 2} = (m_1+1)\mathcal{E}S_{22}^{-1}S_{21} + (n-m-1)\mathcal{E}S_{22}^{-1}S_{21}$, as $\nabla F_2 = \begin{pmatrix} \nabla_{11}S_{11\cdot 2} & 0 \\ \nabla_{21}S_{11\cdot 2} & 0 \end{pmatrix}$,

$$\nabla_{11}S_{11\cdot 2} = \frac{1}{2}(m_1+1)I_{m_1}$$

and

$$\nabla_{21}S_{11\cdot 2} = -\frac{1}{2}(m_1+1)S_{22}^{-1}S_{21}.$$

For details see Corollary 4 (1 & 4) of the Appendix. Solving the two equations yields the result.

□

Theorem 3

$$C\left\{(S_{11\cdot 2})_{ij}, (S_{11\cdot 2})_{kl}\right\} = (n-m_2)\left\{(\Sigma_{11\cdot 2})_{ik}(\Sigma_{11\cdot 2})_{jl} + (\Sigma_{11\cdot 2})_{jk}(\Sigma_{11\cdot 2})_{il}\right\},$$

where $(S_{11\cdot 2})_{ij}$ *is the* ij^{th} *element of* $S_{11\cdot 2}$. *Further* $C(\cdot)$ *denotes the covariance.*

Proof

Take $F_1 = I_m$ and $F_2 = \begin{pmatrix} S_{11\cdot 2}E_{jk}S_{11\cdot 2}E_{li} & 0 \\ 0 & 0 \end{pmatrix}$, with E_{jk} being the jk^{th} basis matrix of dimension $m_1 \times m_1$. Haff's FI in partitioned form yields two equations of which we need only one, viz

$$\begin{aligned} \Sigma_{11\cdot 2}^{-1}\mathcal{E}S_{11\cdot 2}E_{jk}S_{11\cdot 2}E_{li} &= 2\mathcal{E}\nabla_{11}S_{11\cdot 2}E_{jk}S_{11\cdot 2}E_{li} + \\ &+ (n-m-1)\mathcal{E}E_{jk}S_{11\cdot 2}E_{li}. \end{aligned}$$

From Corollary 5(1) of the Appendix emerges that

$$\begin{aligned} 2\nabla_{11}S_{11\cdot 2}E_{jk}S_{11\cdot 2}E_{li} &= (m_1+1)(S_{11\cdot 2})_{kl}E_{ji} + (S_{11\cdot 2})_{jl}E_{ki} + \\ &+ (S_{11\cdot 2})_{kj}E_{li}. \end{aligned}$$

Taking expectations and using Theorem 2 (first part) yields

$$\begin{aligned} \mathcal{E}S_{11\cdot 2}E_{jk}S_{11\cdot 2}E_{li} &= (m_1+1)(n-m_2)(\Sigma_{11\cdot 2})_{kl}\Sigma_{11\cdot 2}E_{ji} + \\ &+ (n-m_2)(\Sigma_{11\cdot 2})_{jl}\Sigma_{11\cdot 2}E_{ki} + (n-m_2)(\Sigma_{11\cdot 2})_{kj}\Sigma_{11\cdot 2}E_{li} + \\ &+ (n-m-1)(n-m_2)(\Sigma_{11\cdot 2})_{kl}\Sigma_{11\cdot 2}E_{ji}, \end{aligned}$$

and finally after taking the trace

$$\begin{aligned}\mathcal{E}(S_{11\cdot 2})_{ij}(S_{11\cdot 2})_{kl} &= (n-m_2)^2\,(\Sigma_{11\cdot 2})_{ij}\,(\Sigma_{11\cdot 2})_{kl} + \\ &+ (n-m_2)\left\{(\Sigma_{11\cdot 2})_{ik}\,(\Sigma_{11\cdot 2})_{jl} + (\Sigma_{11\cdot 2})_{jk}\,(\Sigma_{11\cdot 2})_{il}\right\}\end{aligned}$$

from which the result follows.

□

Theorem 4

$$C\left(B'_{\cdot i}, B'_{\cdot j}\right) = (n-m_2-1)^{-1}\,(\Sigma_{11\cdot 2})_{ij}\,\Sigma_{22}^{-1},$$

where $B' := S_{22}^{-1}S_{21}$ and $B'_{\cdot i}$ is the i^{th} column of B'. Again $C(\cdot)$ denotes the covariance matrix.

Proof

Write $B'_{\cdot i} = S_{22}^{-1}S_{21}e_i$. Then $C(B'_{\cdot i}, B'_{\cdot j}) = \mathcal{E}S_{22}^{-1}S_{21}E_{ij}S_{12}S_{22}^{-1} - \Sigma_{22}^{-1}\Sigma_{21}E_{ij}\Sigma_{12}\Sigma_{22}^{-1}$, by virtue of Theorem 2 (second part).

Take then

$$F_2 = \begin{pmatrix} 0 & S_{11\cdot 2}E_{ij}S_{12}S_{22}^{-1} \\ 0 & 0 \end{pmatrix}.$$

Hence $\mathcal{E}S^{-1}F_2 = \mathcal{E}\begin{pmatrix} 0 & E_{ij}S_{12}S_{22}^{-1} \\ 0 & -S_{22}^{-1}S_{21}E_{ij}S_{12}S_{22}^{-1} \end{pmatrix}$.

With $F_1 = I_m$ the FI yields

$$\begin{aligned}&\mathcal{E}\begin{pmatrix} 0 & \Sigma_{11\cdot 2}^{-1}S_{11\cdot 2}E_{ij}S_{12}S_{22}^{-1} \\ 0 & -\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11\cdot 2}^{-1}S_{11\cdot 2}E_{ij}S_{12}S_{22}^{-1} \end{pmatrix} \\ &= 2\mathcal{E}\begin{pmatrix} 0 & \nabla_{11}S_{11\cdot 2}E_{ij}S_{12}S_{22}^{-1} \\ 0 & \nabla_{21}S_{11\cdot 2}E_{ij}S_{12}S_{22}^{-1} \end{pmatrix} + \\ &+(n-m-1)\mathcal{E}\begin{pmatrix} 0 & E_{ij}S_{12}S_{22}^{-1} \\ 0 & -S_{22}^{-1}S_{21}E_{ij}S_{12}S_{22}^{-1} \end{pmatrix}.\end{aligned}$$

We then get the following two equations:

(i)

$$\begin{aligned}\mathcal{E}\Sigma_{11\cdot 2}^{-1}S_{11\cdot 2}E_{ij}S_{12}S_{22}^{-1} &= 2\mathcal{E}\nabla_{11}S_{11\cdot 2}E_{ij}S_{12}S_{22}^{-1} + \\ &+ (n-m-1)\mathcal{E}E_{ij}S_{12}S_{22}^{-1},\end{aligned}$$

$$(ii) \qquad \mathcal{E}\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11\cdot 2}^{-1}S_{11\cdot 2}E_{ij}S_{12}S_{22}^{-1} = -2\mathcal{E}\nabla_{21}S_{11\cdot 2}E_{ij}S_{12}S_{22}^{-1} + (n-m-1)\mathcal{E}S_{22}^{-1}S_{21}E_{ij}S_{12}S_{22}^{-1}.$$

From (i) we derive

$$\mathcal{E}S_{11\cdot 2}E_{ij}S_{12}S_{22}^{-1} = (n-m_2)\Sigma_{11\cdot 2}E_{ij}\Sigma_{12}\Sigma_{22}^{-1}$$

by Lemma 1 (1) of the Appendix and Theorem 2 (second part).

Insertion in (ii) leads to

$$(iii) \qquad (n-m_2)\Sigma_{22}^{-1}\Sigma_{21}E_{ij}\Sigma_{12}\Sigma_{22}^{-1} + 2\mathcal{E}\nabla_{21}S_{11\cdot 2}E_{ij}S_{12}S_{22}^{-1} = (n-m-1)\mathcal{E}S_{22}^{-1}S_{21}E_{ij}S_{12}S_{22}^{-1}.$$

We use the approach used earlier to find now

$$\nabla_{21}S_{11\cdot 2}E_{ij}S_{12}S_{22}^{-1} = \frac{1}{2}(S_{11\cdot 2})_{ij}S_{22}^{-1} - \frac{1}{2}(m_1+1)S_{22}^{-1}S_{21}E_{ij}S_{12}S_{22}^{-1}.$$

In fact we applied Corollaries 4 (4) and 2(3) of the Appendix to split $\nabla_{21}S_{11\cdot 2}E_{ij}S_{12}S_{22}^{-1}$ into two portions.

Hence $2\mathcal{E}\nabla_{21}S_{11\cdot 2}E_{ij}S_{12}S_{22}^{-1} = \mathcal{E}(S_{11\cdot 2})_{ij}S_{22}^{-1} -$

$$-(m_1+1)\mathcal{E}S_{22}^{-1}S_{21}E_{ij}S_{12}S_{22}^{-1} = (n-m_2)(n-m_2-1)^{-1}(\Sigma_{11\cdot 2})_{ij}\Sigma_{22}^{-1} -$$

$$-(m_1+1)\mathcal{E}S_{22}^{-1}S_{21}E_{ij}S_{12}S_{22}^{-1},$$

by Corollary 7 of the Appendix and Theorem 6 in Section 5.

Substitution in (iii) leads to

$$\mathcal{E}S_{22}^{-1}S_{21}E_{ij}S_{12}S_{22}^{-1} = (n-m_2-1)^{-1}(\Sigma_{11\cdot 2})_{ij}\Sigma_{22}^{-1} + \Sigma_{22}^{-1}\Sigma_{21}E_{ij}\Sigma_{12}\Sigma_{22}^{-1},$$

from which the result follows immediately.

□

This completes Section 4.

5. APPLICATIONS II

In this section we shall consider results presented by various authors (including Haff), occasionally using *different* methods. We shall derive them by using matrix Haffians as advocated by us.

We shall start with some easy often well-known examples to exhibit the powerfulness of the method. They all involve $S \sim W_m(\Sigma, n)$, $n > m+1$.

Theorem 5

$$\mathcal{E}S = n\Sigma.$$

Proof

Take $F_1 = I_m, F_2 = S$ in the Fundamental Identity.

Clearly $dS = I_m(dS)I_m$, hence $\nabla S = \frac{1}{2}(m+1)I_m$ by Theorem 1. Then by the FI: $\Sigma^{-1}S = (m+1)I_m + (n-m-1)I_m = nI_m$, hence $\mathcal{E}S = n\Sigma$. We used $\nabla I_m = 0$.

□

Theorem 6

$$\mathcal{E}S^{-1} = (n-m-1)^{-1}\Sigma^{-1}.$$

Proof

Take $F_1 = F_2 = I_m$. This yields through the FI: $\Sigma^{-1} = (n-m-1)\mathcal{E}S^{-1}$, as $\nabla I_m = 0$.

□

Theorem 7

$$\mathcal{E}SAS = n^2\Sigma A\Sigma + n\Sigma A'\Sigma + n(\mathrm{tr}A\Sigma)\Sigma,$$

where A is a constant matrix.

Proof

Take $F_1 = I_m$ and $F_2 = SAS$. Hence by the FI: $\Sigma^{-1}\mathcal{E}SAS = (m+1)A\mathcal{E}S + A'\mathcal{E}S + \frac{1}{2}(\mathrm{tr}A\mathcal{E}S)I_m + (n-m-1)I_m + (n-m-1)A\mathcal{E}S = n^2A\Sigma + nA'\Sigma + \frac{1}{2}n(\mathrm{tr}A\Sigma)I_m$.

We applied $dSAS = I_m(dS)AS + SA(dS)I_m$, hence $\nabla SAS = \frac{1}{2}(m+1)AS + \frac{1}{2}A'S + \frac{1}{2}(\mathrm{tr}AS)I_m$. Also Theorem 5 was used.

□

Corollary 8

(1) $\mathcal{E}S^2 = n(n+1)\Sigma^2 + n(\mathrm{tr}\,\Sigma)\Sigma.$

(2) $\mathcal{E}(S \otimes S) = n^2(\Sigma \otimes \Sigma) + nK_{mm}(\Sigma \otimes \Sigma) + n(\mathrm{vec}\,\Sigma)(\mathrm{vec}\,\Sigma)'.$

(3) $\mathcal{E}(S \odot S) = n(n+1)\Sigma \odot \Sigma + \Sigma_d 1_m 1_m' \Sigma_d.$

Proof

(1) is obvious. (2) follows by vectorization, viz

$$\begin{aligned}
\mathcal{E}(S \otimes S)\mathrm{vec}A &= n^2(\Sigma \otimes \Sigma)\mathrm{vec}A + n(\Sigma \otimes \Sigma)\mathrm{vec}A' + n(\mathrm{vec}\,\Sigma)(\mathrm{vec}\,\Sigma)'\mathrm{vec}A \\
&= n^2(\Sigma \otimes \Sigma)\mathrm{vec}A + n(\Sigma \otimes \Sigma)K_{mm}\mathrm{vec}A + n(\mathrm{vec}\,\Sigma)(\mathrm{vec}\,\Sigma)'\mathrm{vec}A \\
&= n^2(\Sigma \otimes \Sigma)\mathrm{vec}A + nK_{mm}(\Sigma \otimes \Sigma)\mathrm{vec}A + n(\mathrm{vec}\,\Sigma)(\mathrm{vec}\,\Sigma)'\mathrm{vec}A,
\end{aligned}$$

where K_{mm} is a commutation matrix.

This equality holds for *any* A. We prove (3) by using the relation $S \odot S = W_m'(S \otimes S)W_m$, and the equalities $K_{mm}W_m = W_m$ and $W_m'\mathrm{vec}\,\Sigma = \Sigma_d 1_m$, where Σ_d is a diagonal matrix displaying the diagonal of Σ and 1_m is a column vector consisting of m ones.

For these and other properties of the Hadamard product see, e.g. Neudecker, Liu and Polasek (1995).

Theorem 9

$$\mathcal{E}SAS^{-1} = n(n-m-1)^{-1}\Sigma A \Sigma^{-1} - (n-m-1)^{-1}A' - (n-m-1)^{-1}(\mathrm{tr}A)I_m.$$

Proof

Take $F_1 = SA$ and $F_2 = I_m$. Hence by the FI $\mathcal{E}SA\Sigma^{-1} = 2(\mathcal{E}\nabla A'S)' + (n-m-1)\mathcal{E}SAS^{-1}$, which yields $n\Sigma A\Sigma^{-1} = A' + (\mathrm{tr}A)I_m + (n-m-1)\mathcal{E}SAS^{-1}$. We used $dA'S = A'(dS)I_m$ hence $\nabla A'S = \frac{1}{2}A + \frac{1}{2}(\mathrm{tr}A)I_m$.

□

Corollary 10

$$\mathcal{E}S^{-1}AS = n(n-m-1)^{-1}\Sigma^{-1}A\Sigma - (n-m-1)^{-1}A' - (n-m-1)^{-1}(\mathrm{tr}A)I_m.$$

Proof

Transpose the result of Theorem 9 and replace A' by A (A by A').

□

Corollary 11

(1) $\mathcal{E}(S \otimes S^{-1}) = n(n-m-1)^{-1}\Sigma \otimes \Sigma^{-1} - (n-m-1)^{-1}K_{mm} - (n-m-1)^{-1}(\text{vec}\, I_m)(\text{vec}\, I_m)'$.
(2) $\mathcal{E}(S \odot S^{-1}) = n(n-m-1)^{-1}\Sigma \odot \Sigma^{-1} - (n-m-1)^{-1}I_m - (n-m-1)^{-1}1_m 1_m'$.

Proof

As before. Use $W_m' W_m = I_m$.

□

Theorem 12

(1) $\mathcal{E} s_{ij} S = n^2 \sigma_{ij}\Sigma + n\Sigma(E_{ij} + E_{ji})\Sigma$
(2) $\mathcal{E} s_{ij} S^{-1} = n(n-m-1)^{-1}\sigma_{ij}\Sigma^{-1} - (n-m-1)^{-1}(E_{ij} + E_{ji})$
(3) $\mathcal{E} s^{ij} S = n(n-m-1)^{-1}\sigma^{ij}\Sigma - (n-m-1)^{-1}(E_{ij} + E_{ji})$
where $s^{ij} = \left(S^{-1}\right)_{ij}$.

Proof

(1) Premultiply in Corollary 8 (2) the expression $\mathcal{E}(S \otimes S)$ by $e_i' \otimes I_m$ and postmultiply by $e_j \otimes I_m$. Use $(e_i' \otimes I_m)K_{mm} = I_m \otimes e_i'$, $(e_i' \otimes I_m)\text{vec}\Sigma = \Sigma e_i$ and $\Sigma e_i \otimes e_j' \Sigma = \Sigma E_{ij} \Sigma$.

(2) Subject Corollary 11 (1) to the same treatment.

(3) Follows from (2) immediately.

□

Corollary 13

(1) $\mathcal{E}(\text{tr} AS)S = n^2(\text{tr} A\Sigma)\Sigma + n\Sigma(A + A')\Sigma$
(2) $\mathcal{E}(\text{tr} AS)S^{-1} = n(n-m-1)^{-1}(\text{tr} A\Sigma)\Sigma^{-1} - (n-m-1)^{-1}(A + A')$
(3) $\mathcal{E}(\text{tr} AS^{-1})S = n(n-m-1)^{-1}(\text{tr} A\Sigma^{-1})\Sigma - (n-m-1)^{-1}(A + A')$.

Proof

Use $\operatorname{tr} AS = \sum_{ij} a_{ij} s_{ij}, \sum_{ij} a_{ij} E_{ij} = A$.

□

Finding $\mathcal{E}(\operatorname{tr} AS^{-1})S^{-1}$ is not so easy. We need this for getting $\mathcal{E}S^{-1}AS^{-1}$.

We shall accomplish this in stages.

Theorem 14

For $\Sigma = I_m$:

$$\mathcal{E}(\operatorname{tr} S^{-1})S^{-1} = (n-m)^{-1}(n-m-1)^{-1}(n-m-3)^{-1}\{m(n-m-2)+2\}I_m.$$

Proof

We apply the FI with $F_1 = I_m$ and $F_2 = AS^{-1}$.

We then get by employing Theorem 1:

$$(n-m-1)^{-1}A = -\mathcal{E}S^{-1}A'S^{-1} - \mathcal{E}(\operatorname{tr} AS^{-1})S^{-1} + (n-m-1)\mathcal{E}S^{-1}AS^{-1}.$$

Expected values of the expressions $(s^{ii})^2$, $s^{ii}s^{ij}$, $s^{ii}s^{jj}$, $s^{ii}s^{jk}$, $(s^{ij})^2$, $s^{ij}s^{ik}$ and $s^{ij}s^{kl}$ have to be determined, where i, j, k and l are distinct.

This will be done by choosing appropriate values of A.

(i) $A = E_{ii}$ yields the equation

$$(n-m-1)^{-1}E_{ii} = -\mathcal{E}S^{-1}E_{ii}S^{-1} - \mathcal{E}s^{ii}S^{-1} + (n-m-1)\mathcal{E}S^{-1}E_{ii}S^{-1}.$$

Pre(post)multiplication by $e_i'(e_i), e_i'(e_j)$, $e_j'(e_j)$ and $e_k'(e_l)$ yields

$$\mathcal{E}(s^{ii})^2 = (n-m-1)^{-1}(n-m-3)^{-1}$$

$$\mathcal{E}s^{ii}s^{ij} = 0$$

(1) $$\mathcal{E}(s^{ij})^2 = (n-m-2)^{-1}\mathcal{E}s^{ii}s^{jj}$$

(2) $$\mathcal{E}s^{ik}s^{il} = (n-m-2)^{-1}\mathcal{E}s^{ii}s^{kl}$$

(ii) $A = E_{ij}$ leads to the equation

$$(n-m-1)^{-1}E_{ij} = -\mathcal{E}S^{-1}E_{ji}S^{-1} - \mathcal{E}s^{ij}S^{-1} + (n-m-1)\mathcal{E}S^{-1}E_{ij}S^{-1}.$$

Pre(post)multiplication by $e_i'(e_j)$ yields

$$\mathcal{E}(s^{ij})^2 = \frac{1}{2}(n-m-1)\mathcal{E}s^{ii}s^{jj} - \frac{1}{2}(n-m-1)^{-1},$$

which in combination with (1) gives

$$\begin{aligned}\mathcal{E}(s^{ij})^2 &= (n-m)^{-1}(n-m-1)^{-1}(n-m-3)^{-1}, \\ \mathcal{E}s^{ii}s^{jj} &= (n-m)^{-1}(n-m-1)^{-1}(n-m-2)(n-m-3)^{-1}.\end{aligned}$$

Pre(post)multiplication by $e_i'(e_k)$ yields

$$\mathcal{E}s^{ij}s^{ik} = \frac{1}{2}(n-m-1)\mathcal{E}s^{ii}s^{jk}$$

which in combination with (2) gives

$$\mathcal{E}s^{ii}s^{kl} = \mathcal{E}s^{ik}s^{il} = 0.$$

Finally, pre(post)multiplication by $e_k'(e_l)$ leads to

$$\mathcal{E}s^{jk}s^{il} + \mathcal{E}s^{ij}s^{kl} = (n-m-1)\mathcal{E}s^{ik}s^{jl}$$

which implies

$$\mathcal{E}s^{ij}s^{kl} = 0,$$

as all these terms are identical.

We conclude that

$$\begin{aligned}\mathcal{E}s^{ii}S^{-1} &= d_1 I_m + 2d_2 E_{ii} \\ \mathcal{E}s^{ij}S^{-1} &= d_2(E_{ij} + E_{ji})\end{aligned}$$

with

$$\begin{aligned}d_1 &:= (n-m)^{-1}(n-m-1)^{-1}(n-m-2)(n-m-3)^{-1} \\ d_2 &:= (n-m)^{-1}(n-m-1)^{-1}(n-m-3)^{-1}.\end{aligned}$$

As $\sum_{i=1}^{m}(d_1 I_m + 2d_2 E_{ii}) = (md_1 + 2d_2)I_m$, the theorem has been proved.

□

Theorem 15

For $\Sigma = I_m$:
$\mathcal{E}(\mathrm{tr} AS^{-1})S^{-1} = (n-m)^{-1}(n-m-1)^{-1}(n-m-3)^{-1}\left[A + A' + (n-m-2)(\mathrm{tr}A)I_m\right]$.

Proof

Write $\mathrm{tr} AS^{-1} = \sum_i a_{ii}s^{ii} + \sum_{i \neq j} a_{ij}s^{ij}$.

$$\begin{aligned}
\text{Hence}\quad \mathcal{E}(\mathrm{tr} AS^{-1})S^{-1} &= \sum_i a_{ii}\mathcal{E}s^{ii}S^{-1} + \sum_{i\neq j} a_{ij}\mathcal{E}s^{ij}S^{-1} \\
&= \sum_i a_{ii}(d_1 I_m + 2d_2E_{ii}) + \sum_{i\neq j} a_{ij}d_2(E_{ij}+E_{ji}) \\
&= d_1(\mathrm{tr}A)I_m + 2d_2\sum_i a_{ii}E_{ii} + d_2\sum_{i\neq j} a_{ij}E_{ij} + d_2\sum_{i\neq j} a_{ij}E_{ji} \\
&= d_1(\mathrm{tr}A)I_m + d_2\sum_{ij} a_{ij}E_{ij} + d_2\sum_{ij} a_{ij}E_{ji} \\
&= d_1(\mathrm{tr}A)I_m + d_2(A+A').
\end{aligned}$$

□

Having found $\mathcal{E}(\mathrm{tr} AS^{-1})S^{-1}$ with $\Sigma = I_m$ we can finally determine $\mathcal{E}(\mathrm{tr} AS^{-1})S^{-1}$ for scale parameter $\Sigma \neq I_m$.

Theorem 16

When $S \sim W_m(\Sigma, n)$ *then*

$$\begin{aligned}
\mathcal{E}(\mathrm{tr} AS^{-1})S^{-1} &= (n-m)^{-1}(n-m-1)^{-1}(n-m-3)^{-1}\cdot \\
&\cdot\left[\Sigma^{-1}(A+A')\Sigma^{-1} + (n-m-2)(\mathrm{tr}A\Sigma^{-1})\Sigma^{-1}\right].
\end{aligned}$$

Proof

When $S \sim W_m(\Sigma, n)$ then $\tilde{S} \equiv \Sigma^{-\frac{1}{2}}S\Sigma^{-\frac{1}{2}} \sim W_m(I_m, n)$.

$$\begin{aligned}
\text{Hence}\quad \mathcal{E}(\mathrm{tr} AS^{-1})S^{-1} &= \mathcal{E}(\mathrm{tr}\,\Sigma^{-\frac{1}{2}}A\Sigma^{-\frac{1}{2}}\tilde{S}^{-1})\Sigma^{-\frac{1}{2}}\tilde{S}^{-1}\Sigma^{-\frac{1}{2}} \\
&= d_2\Sigma^{-\frac{1}{2}}\left[\Sigma^{-\frac{1}{2}}A\Sigma^{-\frac{1}{2}} + \Sigma^{-\frac{1}{2}}A'\Sigma^{-\frac{1}{2}} + (n-m-2)(\mathrm{tr}\Sigma^{-\frac{1}{2}}A\Sigma^{-\frac{1}{2}})I_m\right]\Sigma^{-\frac{1}{2}} \\
&= d_2\left[\Sigma^{-1}(A+A')\Sigma^{-1} + (n-m-2)(\mathrm{tr}A\Sigma^{-1})\Sigma^{-1}\right], \quad \text{by Theorem 15.}
\end{aligned}$$

□

Having obtained this result we now present

Theorem 17

$$\mathcal{E}(S^{-1}AS^{-1}) = d_1\Sigma^{-1}A\Sigma^{-1} + d_2\left[\Sigma^{-1}A'\Sigma^{-1} + (\operatorname{tr}A\Sigma^{-1})\Sigma^{-1}\right],$$

where

$$d_1 := (n-m)^{-1}(n-m-1)^{-1}(n-m-2)(n-m-3)^{-1}$$
$$d_2 := (n-m)^{-1}(n-m-1)^{-1}(n-m-3)^{-1}.$$

Proof

Take $F_1 = I_m$ and $F_2 = AS^{-1}$. We get $dF_2 = -AS^{-1}(dS)S^{-1}$ which implies

$$\nabla F_2 = -\frac{1}{2}S^{-1}A'S^{-1} - \frac{1}{2}(\operatorname{tr}AS^{-1})S^{-1}.$$

Applying the FI we get

$$\Sigma^{-1}\mathcal{E}AS^{-1} = -\mathcal{E}S^{-1}A'S^{-1} - \mathcal{E}(\operatorname{tr}AS^{-1})S^{-1} + (n-m-1)\mathcal{E}S^{-1}AS^{-1}.$$

Using Theorems 6 and 16 we arrive at

$$\begin{aligned}&(n-m-1)^{-1}\Sigma^{-1}A\Sigma^{-1} + d_2\Sigma^{-1}(A+A')\Sigma^{-1} + d_1(\operatorname{tr}A\Sigma^{-1})\Sigma^{-1}\\ &= (n-m-1)\mathcal{E}S^{-1}AS^{-1} - \mathcal{E}S^{-1}A'S^{-1}.\end{aligned}$$

Hence by transposition:

$$\begin{aligned}&(n-m-1)^{-1}\Sigma^{-1}A'\Sigma^{-1} + d_2\Sigma^{-1}(A+A')\Sigma^{-1} + d_1(\operatorname{tr}A\Sigma^{-1})\Sigma^{-1}\\ &= (n-m-1)\mathcal{E}S^{-1}A'S^{-1} - \mathcal{E}S^{-1}AS^{-1}.\end{aligned}$$

The first equation we rewrite as

$$\begin{aligned}(n-m-1)\mathcal{E}S^{-1}AS^{-1} &= (n-m-1)^{-1}\Sigma^{-1}A\Sigma^{-1} + d_2\Sigma^{-1}(A+A')\Sigma^{-1}+\\ &+d_1(\operatorname{tr}A\Sigma^{-1})\Sigma^{-1} + \mathcal{E}S^{-1}A'S^{-1} = (n-m-1)^{-1}\Sigma^{-1}A\Sigma^{-1}+\\ &+d_2\Sigma^{-1}(A+A')\Sigma^{-1} + d_1(\operatorname{tr}A\Sigma^{-1})\Sigma^{-1} + (n-m-1)^{-2}\Sigma^{-1}A'\Sigma^{-1}+\\ &+d_2(n-m-1)^{-1}\Sigma^{-1}(A+A')\Sigma^{-1} + d_1(n-m-1)^{-1}(\operatorname{tr}A\Sigma^{-1})\Sigma^{-1}+\\ &+(n-m-1)^{-1}\mathcal{E}S^{-1}AS^{-1}.\end{aligned}$$

Hence

$$\begin{aligned}(n-m)(n-m-1)^{-1}(n-m-2)\mathcal{E}S^{-1}AS^{-1} &= d_1(n-m)(n-m-1)^{-1}(\operatorname{tr}A\Sigma^{-1})\Sigma^{-1}+\\ &+d_1(n-m)(n-m-1)^{-1}(n-m-2)\Sigma^{-1}A\Sigma^{-1}+\\ &+d_2(n-m)(n-m-1)^{-1}(n-m-2)\Sigma^{-1}A'\Sigma^{-1},\end{aligned}$$

which proves the theorem as $d_1 = (n-m-2)d_2$.

$\square$

Corollary 18

$$
\begin{aligned}
&(1)\quad \mathcal{E}(S^{-1}\otimes S^{-1}) &=& \; d_1\Sigma^{-1}\otimes\Sigma^{-1} + d_2 K_{mm}(\Sigma^{-1}\otimes\Sigma^{-1}) + \\
& &+& \; d_2(\operatorname{vec}\Sigma^{-1})(\operatorname{vec}\Sigma^{-1})' \\
&(2)\quad \mathcal{E}(S^{-1}\odot S^{-1}) &=& \; (d_1+d_2)\Sigma^{-1}\odot\Sigma^{-1} + d_2(\Sigma^{-1})_d 1_m 1_m'(\Sigma^{-1})_d. \\
&(3)\quad \mathcal{E}S^{-2} &=& \; (d_1+d_2)\Sigma^{-2} + d_2(\operatorname{tr}\Sigma^{-1})\Sigma^{-1},
\end{aligned}
$$

with $d_1 := (n-m)^{-1}(n-m-1)^{-1}(n-m-2)(n-m-3)^{-1}$
$d_2 := (n-m)^{-1}(n-m-1)^{-1}(n-m-3)^{-1}$, hence
$d_1 + d_2 := (n-m)^{-1}(n-m-3)^{-1}$.

Proof

As before.

$\square$

Theorem 19

$$
\begin{aligned}
\mathcal{E}SASBS &= n[\Sigma(nA+A') + (\operatorname{tr}A\Sigma)I_m]\Sigma[(nB+B')\Sigma + (\operatorname{tr}B\Sigma)I_m] + \\
&+ n\Sigma B\Sigma(A+A')\Sigma + n\Sigma B'\Sigma(nA'+A)\Sigma + \\
&+ n\{\operatorname{tr}A\Sigma(nB+B')\Sigma\}\Sigma
\end{aligned}
$$

Proof

Take $F_1 = I_m$ and $F_2 = SASBS$. The FI yields the equality

$$\Sigma^{-1}\mathcal{E}SASBS = 2\mathcal{E}\nabla SASBS + (n-m-1)\mathcal{E}ASBS.$$

It is easy to see that

$$
\begin{aligned}
2\nabla SASBS + (n-m-1)ASBS &= nASBS + A'SBS + \\
&+ B'SA'S + (\operatorname{tr}AS)BS + (\operatorname{tr}ASBS)I_m.
\end{aligned}
$$

Its expected value is equal to

$$\begin{aligned}
&n^3 A\Sigma B\Sigma + n^2 A\Sigma B'\Sigma + n^2(\operatorname{tr} B\Sigma)A\Sigma + n^2 A'\Sigma B\Sigma +\\
&+nA'\Sigma B'\Sigma + n(\operatorname{tr} B\Sigma)A'\Sigma + n^2 B'\Sigma A'\Sigma + nB'\Sigma A\Sigma +\\
&+n(\operatorname{tr} A\Sigma)B'\Sigma + n^2(\operatorname{tr} A\Sigma)B\Sigma + nB\Sigma(A+A')\Sigma +\\
&+n[n\operatorname{tr} A\Sigma B\Sigma + \operatorname{tr} A\Sigma B'\Sigma + (\operatorname{tr} A\Sigma)(\operatorname{tr} B\Sigma)]I_m.
\end{aligned}$$

We used Theorems 1 and 7, and Corollary 13 (1). Premultiplication by Σ and some rearranging yields the result.

□

Corollary 20

(1) $\mathcal{E}SAS^2 = n[\Sigma(nA + A') + (\operatorname{tr} A\Sigma)I_m]\Sigma[(n+1)\Sigma + (\operatorname{tr}\Sigma)I_m] +$
$+n(n+1)(\operatorname{tr} A\Sigma^2)\Sigma + n\Sigma^2[(n+1)A' + 2A]\Sigma.$

(2) $\mathcal{E}S^2AS = n[(n+1)\Sigma + (\operatorname{tr}\Sigma)I_m]\Sigma[(nA + A')\Sigma + (\operatorname{tr} A\Sigma)I_m] +$
$+n(n+1)(\operatorname{tr} A\Sigma^2)\Sigma + n\Sigma[(n+1)A' + 2A]\Sigma^2$

(3) $\mathcal{E}S^3 = n(n^2 + 3n + 4)\Sigma^3 + 2n(n+1)(\operatorname{tr}\Sigma)\Sigma^2 +$
$+n\left[(\operatorname{tr}\Sigma)^2 + (n+1)\operatorname{tr}\Sigma^2\right]\Sigma.$

(4) $\mathcal{E}(S \otimes S^2) = n^2(n+1)\Sigma \otimes \Sigma^2 + n^2(\operatorname{tr}\Sigma)\Sigma \otimes \Sigma + n(n+1)(\operatorname{vec}\Sigma)(\operatorname{vec}\Sigma^2)' +$
$+n(n+1)\left(\Sigma \otimes \Sigma^2 + \Sigma^2 \otimes \Sigma\right)K_{mm} + n(\operatorname{tr}\Sigma)(\Sigma \otimes \Sigma)K_{mm} + n(n+1)(\operatorname{vec}\Sigma^2)(\operatorname{vec}\Sigma)' +$
$+n(\operatorname{tr}\Sigma)(\operatorname{vec}\Sigma)(\operatorname{vec}\Sigma)' + 2n\Sigma^2 \otimes \Sigma.$

(5) $\mathcal{E}(S \odot S^2) = n(n^2 + 3n + 4)\Sigma \odot \Sigma^2 + n(n+1)(\operatorname{tr}\Sigma)\Sigma \odot \Sigma +$
$+n(\operatorname{tr}\Sigma)\Sigma_d 1_m 1_m' \Sigma_d + n(n+1)\Sigma_d 1_m 1_m' \Sigma_d^2 + n(n+1)\Sigma_d^2 1_m 1_m' \Sigma_d.$

Proof

(1) Replace B by I_m in Theorem 19.

(2) Replace A by I_m and B by A in Theorem 19 *or* transpose Corollary 20 (1) and interchange A and A' in the result.

(3) Replace A by I_m in Corollary 20 (1) or (2).

(4) Vectorize Corollary 20 (2) and omit $\operatorname{vec} A$. This goes as follows. Vectorization of the LHS expression leads to $\mathcal{E}(S \otimes S^2)\operatorname{vec} A$.

Vectorization of the RHS expressions yields

$$\left\{nI_m\otimes\left[(n+1)\Sigma^2+(\mathrm{tr}\,\Sigma)\Sigma\right]\right\}\{(\Sigma\otimes I_m)\,(nI_{m^2}+K_{mm})+$$

$$+n(\mathrm{tr}\,\Sigma)(\mathrm{vec}\,\Sigma)(\mathrm{vec}\,\Sigma)'\}\,\mathrm{vec}A+n(n+1)(\mathrm{vec}\Sigma)(\mathrm{vec}\,\Sigma^2)'\mathrm{vec}A+$$

$$+n(n+1)(\mathrm{vec}\Sigma^2)(\mathrm{vec}\,\Sigma)'\mathrm{vec}A+$$

$$+n(\Sigma^2\otimes\Sigma)\{(n+1)K_{mm}+2I_{m^2}\}\,\mathrm{vec}A.$$

We then cancel $\mathrm{vec}A$.

(5) Follows from (4) immediately, see e.g. Corollary 8 (3).

□

Theorem 21

(1) $\mathcal{E}s_{ij}S^2=n^2(n+1)\sigma_{ij}\Sigma^2+n^2(\mathrm{tr}\Sigma)\sigma_{ij}\Sigma+n(n+1)\Sigma E_{ij}\Sigma^2+$
$+n(\mathrm{tr}\Sigma)\Sigma E_{ij}\Sigma+n(n+1)\left(\Sigma^2E_{ji}\Sigma+\Sigma E_{ji}\Sigma^2\right)+$
$+n(\mathrm{tr}\Sigma)\Sigma E_{ji}\Sigma+n(n+1)\Sigma^2E_{ij}\Sigma+2n\left(\Sigma^2\right)_{ij}\Sigma$

(2) $\mathcal{E}\left(S^2\right)_{ij}S=n^2(n+1)\left(\Sigma^2\right)_{ij}\Sigma+n^2(\mathrm{tr}\Sigma)\sigma_{ij}\Sigma+n(n+1)\Sigma E_{ij}\Sigma^2+$
$+n(n+1)(\Sigma E_{ji}\Sigma^2+\Sigma^2E_{ji}\Sigma)+n(\mathrm{tr}\Sigma)\Sigma E_{ji}\Sigma+n(n+1)\Sigma^2E_{ij}\Sigma+$
$+n(\mathrm{tr}\Sigma)\Sigma E_{ij}\Sigma+2n\sigma_{ij}\Sigma^2$

Proof

(1) Premultiply in Corollary 20 (4) the expression $\mathcal{E}(S\otimes S^2)$ by $e_i'\otimes I_m$ and postmultiply by $e_j\otimes I_m$. Use $K_{mm}(e_i\otimes I_m)=I_m\otimes e_i$ and $a'\otimes b=ba'$.

(2) Pre(post)multiply in Corollary 20 (4) the expression $\mathcal{E}(S\otimes S^2)$ by $I_m\otimes e_i'$ $(I_m\otimes e_j)$. Use $a\otimes b'=ab'$.

□

Corollary 22

(1) $\mathcal{E}(\mathrm{tr}AS)S^2=n^2(n+1)(\mathrm{tr}A\Sigma)\Sigma^2+n^2(\mathrm{tr}A\Sigma)(\mathrm{tr}\,\Sigma)\Sigma+$
$+n(n+1)\Sigma A\Sigma^2+n(\mathrm{tr}\Sigma)\Sigma A\Sigma+n(n+1)\left(\Sigma^2A'\Sigma+\Sigma A'\Sigma^2\right)+n(\mathrm{tr}\Sigma)\Sigma A'\Sigma+$
$+n(n+1)\Sigma^2A\Sigma+2n(\mathrm{tr}A\Sigma^2)\Sigma$

(2) $\mathcal{E}(\text{tr}AS^2)S = n^2(n+1)(\text{tr}A\Sigma^2)\Sigma + n^2(\text{tr}A\Sigma)(\text{tr}\,\Sigma)\Sigma +$
$+n(n+1)\Sigma A\Sigma^2 + n(n+1)\left(\Sigma A'\Sigma^2 + \Sigma^2 A'\Sigma\right) + n(\text{tr}\Sigma)\Sigma A'\Sigma +$
$+n(n+1)\Sigma^2 A\Sigma + n(\text{tr}\Sigma)\Sigma A\Sigma + 2n(\text{tr}A\Sigma)\Sigma^2$

Proof

Use $\text{tr}AS = \sum_{ij} a_{ij}s_{ij}$ and $\sum_{ij} a_{ij}E_{ij} = A$.

□

This has brought us to the end of the article. We want to mention that Theorem 6 and Corollary 18 (3) have been given by Haff (1982). Legault-Giguère (1974) derived Theorems 5, 6, 7, 9, 15, 17 and Corollary 18 (3) in a completely different way.

For Theorems 5, 6, 7, 17 (for $\Sigma = I$) see also Giguère and Styan (1978).

Corollary 10 and Theorems 9 and 17 can also be found in Styan (1989).

For completely different proofs of Theorem 7 see Ghazal and Neudecker (2000) and Neudecker (2000*c*).

Corollaries 8(1) and 20(3) have been established by de Waal and Nel (1973) using a different method.

6. REFERENCES

Ghazal, G. A. and Neudecker, H. (2000). «On second-order and fourth-order moments of jointly distributed random matrices: a survey». *Linear Algebra Appl.*, 321, 61-93.

Giguère, M. A. and Styan, G. P. H. (1978). «Multivariate normal estimation with missing data on several variates». Transactions of the Seventh Prague Conference on Information Theory, Statistical Decision Functions, Random Processes, and of the Eighth European Meeting of Statisticians. *Academia, Prague and D. Reidel, Dordrecht*, vol. B, 129-139.

Haff, L. R. (1981). «Further identities for the Wishart distribution with applications in regression». *Canad. J. Statist.*, 9, 215-224.

Haff, L. R. (1982). «Identities for the inverse Wishart distribution with computational results in linear and quadratic discrimination». *Sankhyã*, B, 44, 245-258.

Legault-Giguère, M. A. (1974). «Multivariate normal estimation with missing data». MSc Thesis, McGill University, Montréal, Québec, Canada.

Magnus, J. R. and Neudecker, H. (1999). *Matrix Differential Calculus with Applications in Statistics and Econometrics*, revised edition. John Wiley, Chichester, England.

Neudecker, H. (2000*a*). «A note on the scalar Haffian». *Qüestiió*, 24, 243-249.
Neudecker, H. (2000*b*). «A note on the matrix Haffian». *Qüestiió*, 24, 419-424.
Neudecker, H. (2000*c*). «On expected values of fourth-degree matrix products of a multinormal matrix variate». *New Trends in Probability and Statistics. Proceedings of the 6th Tartu Conference on Multivariate Statistics*, TEV Vilnius/Utrecht.
Neudecker, H., Liu, S. and Polasek, W. (1995). «The Hadamard product and some of its applications in statistics». *Statistics*, 26, 365-373.
Styan, G. P. H. (1989). «Three useful expressions for expectations involving a Wishart matrix and its inverse». *Statistical Data Analysis and Inference*. Elsevier Science Publishers, Amsterdam, 203-296.
de Waal, D. J. and Nel, D. G. (1973). «On some expectations with respect to Wishart matrices». *S. Afr. Statist. J.*, 7, 61-68.

APPENDIX

Partitioned matrix Haffians

Occasionally we meet with lower-dimensional (not necessarily square) matrix functions of a symmetric matrix X.

Examples are $X_{11}^{-1}, X_{11\cdot 2} := X_{11} - X_{12}X_{22}^{-1}X_{21}, X_{22}^{-1}X_{21}$ and $X_{11\cdot 2}\ E_{jk}\ X_{11\cdot 2}E_{li}$, where E_{jk} is the jk^{th} unit matrix of appropriate dimension, and

$$X = \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix}.$$

The submatrices X_{11} and X_{12} are usually of dimension $m_1 \times m_1$ and $m_2 \times m_2$ respectively with $m_1 + m_2 = m$.

The application of the Fundamental Identity and of Theorem 1 is then not clear-cut. It is obvious that X_{11}^{-1} depends on X_{11}, $X_{11\cdot 2}$ depends on X_{11}, X_{12} and X_{22} (with $X_{12} = X_{21}'$) etc.

We can immediately find $\nabla_{11}X_{11}^{-1}, \nabla_{11}X_{11\cdot 2}$ and $\nabla_{11}X_{11\cdot 2}E_{jk}X_{11\cdot 2}E_{li}$ (when E_{li} is square), because operator and operand have equal dimensions in all these cases, viz. $m_1 \times m_1$.

Finding e.g. $\nabla_{12}PX_{11\cdot 2}Q$, $\nabla_{22}PX_{11\cdot 2}Q$ and $\nabla_{21}PX_{11\cdot 2}Q$ (where the generic constant matrices P and Q have such dimensions that operators and operands fit and the products are square) is not trivial.

The application of the FI and of Theorem 1 will be greatly facilitated by partitioning of the operator ∇, viz as

$$\nabla = \begin{pmatrix} \nabla_{11} & \nabla_{12} \\ \nabla_{21} & \nabla_{22} \end{pmatrix}.$$

As ∇ is symmetric, the off-diagonal block matrices ∇_{12} and ∇_{21} satisfy $\nabla_{21} = \nabla_{12}'$. The symmetry of ∇ follows from the circumstance that the ij^{th} scalar element of ∇ is

$$\frac{1}{2}(1 + \delta_{ij})\frac{\partial}{\partial x_{ij}} \qquad (i, j = 1, \ldots, m)$$

Haff (1981, Lemma 3) presented a collection of useful results on partitioned Haffians. We shall summarize these, in a streamlined and sometimes generalized form. The proofs will be very similar to those of Haff's Lemma 3.

Lemma 1

1. $\nabla_{11} P' X_{11} Q = \frac{1}{2} PQ + \frac{1}{2} (\operatorname{tr} P) Q$

2. $\nabla_{12} P' X_{12} Q = \frac{1}{2} PQ$

3. $\nabla_{12} P' X_{21} Q = \frac{1}{2} (\operatorname{tr} P) Q,$

where P and Q are generic constant matrices.

Proof

1. Apply Theorem 1 with X and F replaced by X_{11} and $P' X_{11} Q$ respectively.
2. Take

$$F = \begin{pmatrix} 0 & 0 \\ P' & 0 \end{pmatrix} X \begin{pmatrix} 0 & 0 \\ Q & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ P' X_{12} Q & 0 \end{pmatrix}.$$

   Clearly $\nabla F = \frac{1}{2} \begin{pmatrix} PQ & 0 \\ 0 & 0 \end{pmatrix}$.

   As $\nabla F = \begin{pmatrix} \nabla_{12} P' X_{12} Q & 0 \\ 0 & 0 \end{pmatrix}$, the result follows.
3. Take

$$F = \begin{pmatrix} 0 & 0 \\ 0 & P' \end{pmatrix} X \begin{pmatrix} Q & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ P' X_{21} Q & 0 \end{pmatrix}.$$

   Then $\nabla F = \frac{1}{2} (\operatorname{tr} P) \begin{pmatrix} Q & 0 \\ 0 & 0 \end{pmatrix}$ and

$$\nabla_{12} P' X_{21} Q = \frac{1}{2} (\operatorname{tr} P) Q.$$

□

Corollary 2

1. $\nabla_{22} P' X_{22} Q = \frac{1}{2} PQ + \frac{1}{2} (\operatorname{tr} P) Q$

2. $\nabla_{21} P' X_{21} Q = \frac{1}{2} PQ$

3. $\nabla_{21} P' X_{12} Q = \frac{1}{2} (\operatorname{tr} P) Q.$

Proof

2. Take $F = \begin{pmatrix} 0 & P' \\ 0 & 0 \end{pmatrix} X \begin{pmatrix} 0 & Q \\ 0 & 0 \end{pmatrix}$.
3. Take $F = \begin{pmatrix} P' & 0 \\ 0 & 0 \end{pmatrix} X \begin{pmatrix} 0 & 0 \\ 0 & Q \end{pmatrix}$.

□

We shall now consider some special results.

Corollary 3

1. $\nabla_{11} P' X_{11}^{-1} Q = -\frac{1}{2} X_{11}^{-1} P X_{11}^{-1} Q - \frac{1}{2} (\operatorname{tr} P X_{11}^{-1}) X_{11}^{-1} Q$

2. $\nabla_{12} P' X_{12} X_{22}^{-1} Q = \frac{1}{2} P X_{22}^{-1} Q$

3. $\nabla_{12} P' X_{21} X_{11}^{-1} Q = \frac{1}{2} (\operatorname{tr} P) X_{11}^{-1} Q$

4. $\nabla_{11} P' X_{21} X_{11}^{-1} Q = -\frac{1}{2} X_{11}^{-1} X_{12} P X_{11}^{-1} Q - \frac{1}{2} (\operatorname{tr} P X_{11}^{-1} X_{12}) X_{11}^{-1} Q$

5. $\nabla_{12} P' X_{22}^{-1} X_{21} Q = \frac{1}{2} (\operatorname{tr} P X_{22}^{-1}) Q$

6. $\nabla_{12} P' X_{11}^{-1} X_{12} Q = \frac{1}{2} X_{11}^{-1} P Q$

7. $\nabla_{22} P' X_{22}^{-1} X_{21} Q = -\frac{1}{2} X_{22}^{-1} P X_{22}^{-1} X_{21} Q - \frac{1}{2} (\operatorname{tr} P X_{22}^{-1}) X_{22}^{-1} X_{21} Q$

Proof

1. Consider $dP' X_{11}^{-1} Q = P'(dX_{11}^{-1}) Q = -P' X_{11}^{-1} (dX_{11}) X_{11}^{-1} Q$.
   Replace then P' by $-P' X_{11}^{-1}$ and Q by $X_{11}^{-1} Q$ in Lemma 1 (1).
2. Replace Q by $X_{22}^{-1} Q$ in Lemma 1 (2).
3. Replace Q by $X_{11}^{-1} Q$ in Lemma 1 (3).
4. Replace P' by $P' X_{21}$ in 1 of this corollary.
5. Replace P' by $P' X_{22}^{-1}$ in Lemma 1 (3).
6. Replace P' by $P' X_{11}^{-1}$ in Lemma 1 (2).

7. Replace ∇_{11} by ∇_{22}, X_{11}^{-1} by X_{22}^{-1} and Q by $X_{21}Q$ in 1 of this corollary.

□

Note. Haff's Lemma 3 (e) is a special case of 5 in this corollary.

Corollary 4

1. $\nabla_{11}P'X_{11\cdot 2}Q = \frac{1}{2}PQ + \frac{1}{2}(\operatorname{tr} P)Q$

2. $\nabla_{12}P'X_{11\cdot 2}Q = -\frac{1}{2}PX_{22}^{-1}X_{21}Q - \frac{1}{2}(\operatorname{tr} PX_{22}^{-1}X_{21})Q$

3. $\nabla_{22}P'X_{11\cdot 2}Q = \frac{1}{2}X_{22}^{-1}X_{21}PX_{22}^{-1}X_{21}Q + \frac{1}{2}(\operatorname{tr} PX_{22}^{-1}X_{21})X_{22}^{-1}X_{21}Q$

4. $\nabla_{21}P'X_{11\cdot 2}Q = -\frac{1}{2}X_{22}^{-1}X_{21}PQ - \frac{1}{2}(\operatorname{tr} P)X_{22}^{-1}X_{21}Q$

Proof

1. As only X_{11} varies this result equals that of Lemma 1 (1).
2. This follows from Lemma 1 (2 & 3 combined).

   The reason is that now $dX_{11\cdot 2} = -(dX_{12})X_{22}^{-1}X_{21} - X_{12}X_{22}^{-1}dX_{21}$. Hence we replace Q by $-X_{22}^{-1}X_{21}Q$ in 2 and P' by $P'X_{12}X_{22}^{-1}$ in 3 and add the resulting two expressions together.
3. This follows from Corollary 2 (1). Now

$$dX_{11\cdot 2} = -X_{12}(dX_{22}^{-1})X_{21} = X_{12}X_{22}^{-1}(dX_{22})X_{22}^{-1}X_{21}.$$

   Hence we replace P' by $P'X_{12}X_{22}^{-1}$ and Q by $X_{22}^{-1}X_{21}Q$ in 1.
4. This follows from Lemma 1 (4 & 5 combined).

   The reason is that $dX_{11\cdot 2} = -(dX_{12})X_{22}^{-1}X_{21} - X_{12}X_{22}^{-1}dX_{21}$. Hence we substitute $-X_{22}^{-1}X_{21}Q$ for Q in 4 and $-P'X_{12}X_{22}^{-1}$ for P' in 5 and add the resulting two expressions together.

□

Corollary 5

1. $\nabla_{11}P'X_{11\cdot 2}QX_{11\cdot 2}R = \frac{1}{2}PQX_{11\cdot 2}R + \frac{1}{2}(\operatorname{tr}P)QX_{11\cdot 2}R + \frac{1}{2}Q'X_{11\cdot 2}PR + \frac{1}{2}(\operatorname{tr}P'X_{11\cdot 2}Q)R$
2. $\nabla_{12}P'X_{11\cdot 2}QX_{11\cdot 2}R = -\frac{1}{2}PX_{22}^{-1}X_{21}QX_{11\cdot 2}R - \frac{1}{2}Q'X_{11\cdot 2}PX_{22}^{-1}X_{21}R -$
   $-\frac{1}{2}(\operatorname{tr}PX_{22}^{-1}X_{21}Q'X_{11\cdot 2})R - \frac{1}{2}(\operatorname{tr}PX_{22}^{-1}X_{21})QX_{11\cdot 2}R$
3. $\nabla_{22}P'X_{11\cdot 2}QX_{11\cdot 2}R = \frac{1}{2}X_{22}^{-1}X_{21}PX_{22}^{-1}X_{21}QX_{11\cdot 2}R + \frac{1}{2}(\operatorname{tr}PX_{22}^{-1}X_{21})X_{22}^{-1}X_{21}QX_{11\cdot 2}R +$
   $+\frac{1}{2}X_{22}^{-1}X_{21}Q'X_{11\cdot 2}PX_{22}^{-1}X_{21}R + \frac{1}{2}(\operatorname{tr}PX_{22}^{-1}X_{21}Q'X_{11\cdot 2})X_{22}^{-1}X_{21}R$
4. $\nabla_{21}P'X_{11\cdot 2}QX_{11\cdot 2}R = -\frac{1}{2}X_{22}^{-1}X_{21}PQX_{11\cdot 2}R - \frac{1}{2}(\operatorname{tr}P)X_{22}^{-1}X_{21}QX_{11\cdot 2}R -$
   $-\frac{1}{2}X_{22}^{-1}X_{21}Q'X_{11\cdot 2}PX_{11\cdot 2}R - \frac{1}{2}(\operatorname{tr}Q'X_{11\cdot 2}P)X_{22}^{-1}X_{21}X_{11\cdot 2}R$

Proof

1. Using Theorem 1 we conclude from

   $$dP'X_{11\cdot 2}QX_{11\cdot 2}R = P'(dX_{11})QX_{11\cdot 2}R + P'X_{11\cdot 2}Q(dX_{11})R$$

   that the identity holds.
2. This is proved in the same way as Corollary 4 (2). The expression $P'X_{11\cdot 2}QX_{11\cdot 2}R$ is split into $P'X_{11\cdot 2}(QX_{11\cdot 2}R)$ and $(P'X_{11\cdot 2}Q)X_{11\cdot 2}R$. We then make the following substitutions in Corollary 4 (2): *(i)* P remains P, Q becomes $QX_{11\cdot 2}R$ and *(ii)* P becomes $Q'X_{11\cdot 2}P$, Q becomes R.

   This yields the result.
3. This is proved in the same way as Corollary 4 (3). We make the same substitutions as previously.
4. The proof is similar to that of Corollary 4 (4).

   The same substitutions are used as above.

□

Lemma 6

$$\mathcal{E}\left(S_{22}^{-1}\right)_{ij} S_{11\cdot 2} = (n-m_2)\mathcal{E}\left(S_{22}^{-1}\right)_{ij}\Sigma_{11\cdot 2}.$$

Proof

Take

$$F_1 = I_{m_1} \quad \text{and} \quad F_2 = \begin{pmatrix} \left(S_{22}^{-1}\right)_{ij} S_{11\cdot 2} & 0 \\ 0 & 0 \end{pmatrix}.$$

Then

$$\Sigma^{-1} F_2 = \begin{pmatrix} \left(S_{22}^{-1}\right)_{ij} \Sigma_{11\cdot 2}^{-1} S_{11\cdot 2} & 0 \\ -\left(S_{22}^{-1}\right)_{ij} \Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11\cdot 2}^{-1} S_{11\cdot 2} & 0 \end{pmatrix},$$

$$\nabla F_2 = \begin{pmatrix} \nabla_{11} \left(S_{22}^{-1}\right)_{ij} S_{11\cdot 2} & 0 \\ \nabla_{21} \left(S_{22}^{-1}\right)_{ij} S_{11\cdot 2} & 0 \end{pmatrix}$$

and

$$S^{-1} F_2 = \begin{pmatrix} \left(S_{22}^{-1}\right)_{ij} I_{m_1} & 0 \\ -\left(S_{22}^{-1}\right)_{ij} S_{22}^{-1} S_{21} S_{11\cdot 2} & 0 \end{pmatrix}.$$

Hence by the FI we get

$$\begin{aligned} \mathcal{E} \left(S_{22}^{-1}\right)_{ij} \Sigma_{11\cdot 2}^{-1} S_{11\cdot 2} &= (m_1+1) \mathcal{E} \left(S_{22}^{-1}\right)_{ij} I_{m_1} + (n-m-1) \left(\mathcal{E} S_{22}^{-1}\right)_{ij} I_{m_1} \\ &= (n-m_2) \mathcal{E} \left(S_{22}^{-1}\right)_{ij} I_{m_1}, \end{aligned}$$

by virtue of Lemma 1 (1). This yields

$$\mathcal{E} \left(S_{22}^{-1}\right)_{ij} S_{11\cdot 2} = (n-m_2) \mathcal{E} \left(S_{22}^{-1}\right)_{ij} \Sigma_{11\cdot 2}.$$

□

Corollary 7

$$\mathcal{E} \left(S_{11\cdot 2} \otimes S_{22}^{-1}\right) = (n-m_2) \Sigma_{11\cdot 2} \otimes \mathcal{E} S_{22}^{-1}.$$

Proof

Immediate from Lemma 6.

□