Institute of Statistical Mathematics (2010.06.26)

Stein-Haff identity for singular Wishart matrices

KONNO Yoshihiko
Japan Women's University

June 30, 2010

Real Wishart distributions

- n is a positive integer; $\Sigma > 0$ is a $p \times p$ postive definite matrix ;
- A $p \times p$ semi-positive matrix W follows the Wishart distribution $W_p(n, \Sigma)$, i.e.
  -
    $$\begin{aligned} dW_p(n,\,\Sigma)(dw) &= \frac{\mathrm{Det}\,\Sigma^{n/2}}{2^{np/2}\Gamma_p(n/2)} e^{-\mathrm{Tr}\,(w\Sigma^{-1})/2} \mathrm{Det}\,w^{n/2} \frac{dw}{\mathrm{Det}\,w^{(p+1)/2}}; \\ \Gamma_p(n/2) &= \Pi_{i=1}^{p}\Gamma((n-i+1)/2), \end{aligned}$$

    provided $n > p - 1$;
  - For $n \leq p - 1$, no density exists. For a $p \times p$ positive definite matrix s,

    $$\mathbb{E}[e^{-\mathrm{Tr}\,(sW)}] = \frac{1}{\mathrm{Det}\,(I_p - s\Sigma)^{n/2}}.$$

Stein-Haff identity for nonsingular Wishart matrices

- Decompose $W = OLO'$ where L is a diagonal matrix with the ordered eignevalues $\ell_1 \geq \ell_2 \geq \cdots \geq \ell_p$ and O is $p \times p$ orthogonal matrix.

- For $n > p + 1$, under some regularity conditons,

$$\mathbb{E}[\Sigma^{-1} O \mathrm{Diag}(\psi_1, \ldots, \psi_p) O'] = \mathbb{E}[O \mathrm{Diag}(\psi_1^{(1)}, \ldots, \psi_p^{(1)}) O']$$

where ψ_i's($i = 1, 2, \ldots, p$) are differential real-valued functions on $\ell_1 \geq \ell_2 \geq \cdots \geq \ell_p$ and

$$\psi_i^{(1)} = 2\frac{\partial \psi_i}{\partial \ell_i} + (n - p - 1)\frac{\psi_i}{\ell_i} + \sum_{b \neq i} \frac{\psi_i - \psi_b}{\ell_i - \ell_b}.$$

Approaches to derive the identity

- Stein's approach: Use Integraton-by-parts for centered Normal distributions and convert it to Wishart distributions;

- Haff's approach: Use Integraton-by-parts for Wishart distributions. Need a Wishart density.

- Sheena's approach: Use Integraton-by-parts for eigenvalues' distribution of Wishart matrix for trace style formula, i.e.,

$$\mathbb{E}[\mathrm{Tr}\,(\Sigma^{-1}O\mathrm{Diag}(\psi_1,\,\ldots,\,\psi_p)O')] = \mathbb{E}\left[\sum_{i=1}^{p}\left\{(n-p-1)\frac{\psi_i}{\ell_i} + 2\frac{\partial\psi_i}{\partial\ell_i} + \sum_{b\neq i}\frac{\psi_i-\psi_b}{\ell_i-\ell_b}\right\}\right].$$

Singular Wishart case

- Decompose $W = O_1 \mathrm{Diag}(\ell_1, \ldots, \ell_n) O_1'$, where O_1 is $p \times n$ semi-orhogonal matrix such that $O_1' O_1 = I_n$.

- For $n < p - 1$, under some regularity conditons,

$$\mathbb{E}[\Sigma^{-1} O_1 \mathrm{Diag}(\psi_1, \ldots, \psi_n) O_1'] = \mathbb{E}[O_1 \mathrm{Diag}(\psi_1^{(1r)}, \ldots, \psi_n^{(1r)}) O_1' + \sum_{i=1}^{n} \frac{\psi_i}{\ell_i}(I_p - O_1 O_1')],$$

where $\psi_i^{(1r)} = 2\frac{\partial \psi_i}{\partial \ell_i} - \frac{\psi_i}{\ell_i} + \sum_{b \neq i} \frac{\psi_i - \psi_b}{\ell_i - \ell_b}$.

- Kubokawa and Srivastava's formula via Sheena's approach:

$$\mathbb{E}[\mathrm{Tr}\,(\Sigma^{-1} O_1 \mathrm{Diag}(\psi_1, \ldots, \psi_n) O_1')] = \mathbb{E}\left[\sum_{i=1}^{n} \left\{ (p - n - 1)\frac{\psi_i}{\ell_i} + 2\frac{\partial \psi_i}{\partial \ell_i} + \sum_{b \neq i} \frac{\psi_i - \psi_b}{\ell_i - \ell_b} \right\}\right].$$

Corollary

- To obtain an unbiased risk estimate for $\mathbb{E}[\mathrm{Tr}\,(\hat{\Sigma}\Sigma^{-1} - I_p)^2]$ for an estimator $\hat{\Sigma} = O_1 \mathrm{Diag}(\psi_1, \ldots, \psi_n) O_1'$, we need

$$\mathbb{E}[\mathrm{Tr}\,(\Sigma^{-1}\hat{\Sigma}\Sigma^{-1}\hat{\Sigma})] = \mathbb{E}[\mathrm{Tr}\,(\Sigma^{-1} O_1 \mathrm{Diag}(\psi_1^{(1s)}, \ldots, \psi_n^{(1s)}) O_1')],$$

where

$$\psi_i^{(1s)} = (p - n - 1)\frac{\psi_i^2}{\ell_i} + 4\psi_i \frac{\partial \psi_i}{\partial \ell_i} + 2\psi_i \sum_{b \neq i} \frac{\psi_i - \psi_b}{\ell_i - \ell_i}.$$

Sketch of the proof

- Let Z follow the standard normal distribution. For real valued-function g, we have
$$\mathbb{E}[g(Z)] = \mathbb{E}[Zg'(Z)].$$

- Let $X = (x_{ij}) : n \times p$ be $N_{n\times p}(0, I_n \otimes \Sigma)$. For $G(W) = (g_{ij}) : p \times p$ and $W = X'X$,
$$\mathbb{E}[\Sigma^{-1}WG] = \mathbb{E}[nG + (X'\nabla_X)'G],$$
where $\nabla_X = (\partial/\partial x_{ij})$ and $X = (x_{ij})$. Here
$$(X'\nabla_X)'G = (\sum_{k_1, k_2} x_{k_2 k_1}(\partial g_{k_2 j}/\partial x_{k_2 i}))$$
.

- By chain-rule,

$$
\begin{aligned}
\frac{\partial \ell_m}{\partial x_{ij}} &= 2\sum_{c_1} o_{c_1 m} x_{i c_1} o_{jm}; \\
\frac{\partial o_{ak}}{\partial x_{ij}} &= \sum_{b \neq i}^{n} \sum_{c_1}^{n} \frac{o_{ab}\{o_{jb} o_{c_1 k} + o_{c_1 b} o_{jk}\} x_{i c_1}}{\ell_k - \ell_b} \\
&\quad + \sum_{b=n+1}^{p} \sum_{c_1}^{p} \frac{o_{ab}\{o_{jb} o_{c_1 k} + o_{c_1 b} o_{jk}\} x_{i c_i}}{\ell_k};
\end{aligned}
$$

where $O_1 : p \times n = (o_{ij})$, $W = O_1 \mathrm{Diag}(\ell_1, \ldots, \ell_n) O_1'$.