Some Decompositions of Spherical Distributions

Certain rotationally symmetric or spherical multivariate distributions can be factored into independent components that produce in a natural manner various uniform, chi, triangular chi, root F, triangular root F, t, disguised t, triangular root beta, Wishart and disguised Wishart distributions.

AMS classification: Primary 62E10; Secondary 62H10
Key words and phrases: Matrix distributions, chi distributions, disguised distributions, root beta distributions, root F distributions, Student distributions.

1. Introduction

The standard normal distribution for a \( p \times n \) matrix \( Z \) has probability element \((2\pi)^{-pn/2} \exp(-\frac{1}{2}ZZ')dZ\) on \( \mathbb{E}^{pn} \).

The transformed variable \( Y = \Theta + AZB \) has the matrix normal \( N_{p \times n}(\Theta; \Sigma_2 \otimes \Sigma_1) \) distribution with probability element

\[
(2\pi)^{-pn/2} |\Sigma_1|^{-n/2} |\Sigma_2|^{-p/2} \exp(-\frac{1}{2}(Y-\Theta)\Sigma_2^{-1}(Y-\Theta)'dY
\]

where \( \Sigma_1 = AA' \) is \( p \times p \) and \( \Sigma_2 = B'B \) is \( n \times n \). In particular, if \( \Theta = 0 \) and \( B \) is orthogonal, then \( Y \) has the \( N_{p \times n}(0; I \otimes \Sigma_1) \); this is the distribution of a sample of \( n \) from the central normal distribution with variance matrix \( \Sigma_1 \), and various Wishart, Student, and beta distributions arise naturally as marginal distributions.

The \( N_{p \times n}(0; I \otimes \Sigma_1) \) distribution for a \( p \times n \) matrix \( Y \) is invariant under right multiplication, \( YB \sim Y \), by any orthogonal \( n \times n \) matrix \( B \). In this paper we examine various decompositions and marginal distributions that devolve from the invariance of a \( p \times n \) matrix distribution under right multiplication by any orthogonal \( n \times n \) matrix.

---

* Research supported by Canadian N.R.C. Grant A3011; the present affiliation of L. Bishop is the University of Wisconsin.
We record some specialized notation to be used in the sequel. For a $p \times p$ matrix $A$ let $A^{(k)}$ and $A_{(k)}$ be the upper-left and lower-right $k \times k$ submatrices and define

$$|A|_{\nabla} = \Pi_1^{p} |A^{(k)}|, \quad |A|_{\Delta} = \Pi_1^{p} |A_{(k)}|.$$  

The decreasing determinant $|\cdot|_{\nabla}$ and increasing determinant $|\cdot|_{\Delta}$ have some simple properties in part relative to a positive lower triangular (PLT) $p \times p$ matrix $L$:

$$|A|_{\nabla} = |A'|_{\nabla}, \quad |A|_{\Delta} = |A'|_{\Delta} \quad (1.1)$$

$$|L|_{\nabla} = \xi_1^{p} \xi_{22}^{p-1} \cdots \xi_{pp} \quad |L|_{\Delta} = \xi_1^{p} \xi_{22}^{2} \cdots \xi_{pp} \quad (1.2)$$

$$|LA|_{\nabla} = |AL'|_{\nabla} = |L|_{\nabla} |A|_{\nabla}, \quad |AL|_{\Delta} = |L'A|_{\Delta} = |A|_{\Delta} |L|_{\Delta} \quad (1.3)$$

$$|L^{-1}|_{\nabla} = |L^{-1}|_{\nabla} \quad (|L| \neq 0), \quad |L^{-1}|_{\Delta} = |L^{-1}|_{\Delta} \quad (|L| \neq 0) \quad (1.4)$$

$$|L|_{\Delta} |L|_{\nabla} = |L|_{p+1} \quad (1.5)$$

$$\frac{|A|_{\Delta}}{|A^{-1}|_{\nabla}} = |A|_{p+1} = \frac{|A|_{\nabla}}{|A^{-1}|_{\Delta}} \quad (|A| \neq 0) \quad (1.6)$$

$$|A|_{\nabla} = |A_{11}|_{\nabla} |A_{11}|_{p-k} |A_{22}^{-1} A_{21} A_{11}^{-1} A_{12}|_{\nabla} \quad (|A_{11}| \neq 0) \quad (1.7)$$

$$|A|_{\Delta} = |A_{22}|_{\Delta} |A_{22}|_{k} |A_{11}^{-1} A_{12} A_{22}^{-1} A_{21}|_{\Delta} \quad (|A_{22}| \neq 0).$$

where $A$ is partitioned into the first $k$ rows and columns and last $p-k$ rows and columns.

We also record some Jacobians of matrix transformations. The volume element for a $p \times n$ matrix is obtained typically by inclusion in $E^{pn}$; an exception is a symmetric matrix $S$ where the simpler $dS = \Pi_{i<j} ds_{ij}$ is commonly used. Let $X$ be $p \times n$, $L$ be $p \times p$ PLT and $S$ be $p \times p$ positive definite and as transformation matrices let $A$ be $p \times p$, $B$ be $n \times n$, $T$ be $p \times p$ PLT, and $W$ be $p \times p$ symmetric. Then:

$$dAXB = |A|^n |B|^p \, dX \quad (1.8)$$

$$dASA' = |A|^{p+1} \, ds \quad (1.9)$$
\[ dT L = \left| T \right|_\Delta dL, \quad dLT = \left| T \right|_\vee dL \tag{1.10} \]
\[ dL L' = 2^p |L|_\vee dL, \quad dL' L = 2^p |L|_\Delta dL \tag{1.11} \]
\[ ds^{-1} = |s|^{-(p+1)} ds \tag{1.12} \]
\[ dL W L' = 2^p |L|_\vee |W|_\vee dL \tag{1.13} \]

2. Preliminaries: the Triangular Chi Distribution

Let \( Z \) be a \( p \times n \) matrix \((p \leq n)\) and consider the unique factorization \( Z = LR \) where \( L \) is \( p \times p \) PLT and \( R \) is \( p \times n \) semi-orthogonal \((RR' = I)\); we delete the measure zero portion of \( \mathbb{E}^{pn} \) having \( Z \) less than full rank \( p \).

Let \( P_p = \{ L \} \) be the set of \( p \times p \) PLT matrices. The set \( P_p \) can be presented as a positive \( 1/2^p \) portion of \( \mathbb{E}^{p(p+1)/2} \) with Euclidean volume \( dL \). Under matrix multiplication \( P_p \) is a group.

Let \( R_{p,n} = \{ R \} \) be the set of \( p \times n \) semiorthogonal matrices. The set \( R_{p,n} \) forms a Stiefel manifold in \( \mathbb{E}^{pn} \) which has dimension \( pn - p(p+1)/2 \). (A \( p \)-frame is a sequence of \( p \) orthonormal vectors in \( \mathbb{E}^n \); the Stiefel manifold \( S_{p,n} \) is the set of \( p \) frames in \( \mathbb{E}^n \); \( R_{p,n} \) is an imbedding of \( S_{p,n} \) in \( \mathbb{E}^{pn} \).)

For the full-measure retained portion of \( \mathbb{E}^{pn} \), the unique factorization \( Z = LR \) shows that \( R_{p,n} \) forms a cross-section to the orbits of the group \( P_p \) applied by left matrix multiplication. Consider volume measures in the neighbourhood of the cross-section \( R_{p,n} \). The volume measure tangent to the orbit \( P_p R \) at \( R \) coincides with the volume measure \( dL \) in the neighbourhood of the identity. A natural measure for the cross-section \( R_{p,n} \) is given locally at \( R \) by the projection on the orthogonal space to the orbit \( P_p R \). Let \( dR \) designate this Euclidean volume orthogonal to the orbits. Then in the neighborhood of the cross-section \( R_{p,n} \), the full measure \( dZ \) can be factored \( dZ = dL dR \) into orthogonal Euclidean components. The alternative in which \( dL \) is taken to be volume measured tangent to the manifold \( R_{p,n} \) involves, in terms of appropriate local coordinates, \( p(p-1)/2 \) duplications of the orthogonal coordinates, each duplication giving a \( \sqrt{2} \) inflation of length measure; the resulting factorization in the neighbourhood of \( R_{p,n} \) would then be \( dZ = dL^{2-p(p-1)/4} dR \) in terms of the non-orthogonal Euclidean components.
Now let \( Z \) be a \( p \times n \) matrix with a standard normal distribution and consider the joint distribution of \((L,R)\) obtained from the preceding factorization \( Z = LR \). The invariant \( (z^p) \) measure on \( EP^n \) that agrees with Euclidean volume at the \( p \) cross-section \( \Omega_{p,n} \) is \( |L|^{-p} dZ \); the invariant measure on \( \mathbb{R}^p \) that agrees with tangent volume at the point \( R \) is \( |L|^{-1}_\Delta dL \); accordingly we have

\[
dZ = |L|^n |L|^{-1}_\Delta dLdR. \tag{2.1}
\]

The probability element for \( Z \) factors so that the variables \( L \) and \( R \) separate

\[
(2\pi)^{-pn/2} \text{etr} \{-\frac{1}{2} ZZ'\}dZ = (2\pi)^{-pn/2} \text{etr} \{-\frac{1}{2} LL'\} \frac{|L|^n}{|L|_\Delta} dLdR \tag{2.2}
\]

where \( A_n^{(p)} = A_n^A A_{n-1} \cdots A_{n-p+1} \) and \( A_k = 2\pi^{k/2}/\Gamma(k/2) \) is the volume of the unit sphere in \( E_k \); the normalizing constants are obtained by noting that the \( \ell_{ij} \) (\( i > j \)) are independent \( N(0,1) \) and that the \( \ell_{ii} \) are independent \( \chi(n-i+1) \); the chi density notation follows from

\[
\frac{1}{\Gamma(k/2)} e^{-\chi^2/2} \left( \frac{\chi^2}{2} \right)^{k/2-1} d\chi = \frac{A_k}{(2\pi)^{k/2}} e^{-\chi^2/2} \chi^{k-1} d\chi.
\]

Thus \( L \) and \( R \) are statistically independent; the Stiefel component \( R \) is uniform on \( \mathbb{R}_{p,n}^{(p)} \) (with total volume \( A_n^{(p)} \) measured orthogonal to \( \mathbb{R}_p^{(p)} \) orbits); and \( L \) is said to have the triangular chi distribution \( \Delta \chi_p(n) \) (this is the Bartlett decomposition matrix from the Wishart distribution \( \mathcal{W}_p(n, I) \) for \( S = LL' \)).

Now let \( Z \) be a \( p \times n \) matrix with a rotationally symmetric distribution under right multiplication by an \( n \times n \) orthogonal matrix: \( f(Z) = f(ZB) \) for \( B \) in \( R_{n,n}^- \). Using the factorization \( Z = LR \) we can then write

\[
f(Z) = f(LR) = f(ILI) = f_p(ILI') = f_p(ZZ')
\]

and call \( f_p \) the kernel of the distribution \( f \). The probability element for \( Z \) factors so that the variables \( L \) and \( R \) separate
\[ f(z)dz = \frac{f_p(\Delta)}{\sqrt{\frac{L}{\Delta}}} \frac{dL}{dL} \frac{dL}{\Delta} \]

\[ = A_n^{(p)} f_p(\Delta) \frac{|L|^{n/2}}{|\Delta|^{n/2}} \frac{dL}{\Delta} \cdot \frac{d\Delta}{A_n^{(p)}} \]

(2.3)

where the normalizing constants are obtained by noting the integration properties of \( d\Delta \) derived from (2.2). Thus \( L \) and \( R \) are statistically independent; the Stiefel component \( R \) is uniform on \( \mathbb{R}_{p,n} \), and \( L \) is said to have the generalized triangular chi distribution \( \Delta \chi_p^{(f_p;n)} \) with kernel \( f_p \).

The marginal distribution of \( S = LL' \) is a one-one map of the \( \Delta \chi_p^{(f_p;n)} \) distribution. Using (1.11) and (1.5) we obtain

\[ A_n^{(p)} f_p(S) |S|^{(n-p-1)/2} 2^{-p} dS \]

(2.4)

as a generalized Wishart \( W_p^{(f_p;n)} \) with kernel \( f_p \) (see Anderson, 1958, Lemma 13.3.1); in the standard normal case (from (2.2)) this becomes the ordinary Wishart \( W_p^{(n)} \)

\[ \frac{A_n^{(p)}}{(2\pi)^{pn/2}} \text{etr}(-S/2) |S|^{(n-p-1)/2} \cdot \frac{ds}{2^p} . \]

(2.5)

Alternatively if \( S \) is defined to be \( LL' \) we obtain the disguised Wishart \( \hat{W}_p^{(n)} \) in Tan and Guttmann (1971):

\[ \frac{A_n^{(p)}}{(2\pi)^{pn/2}} \text{etr}(-S/2) |S|^{(n-p-1)/2} \cdot \frac{|S|^{(p+1)/2} |\Delta|^{n/2}}{\Delta} \cdot \frac{ds}{2^p} \]

(2.6)

which involves a 'disguising' factor \( |S|^{(p+1)/2} / |\Delta|^{n/2} \). A generalized disguised Wishart is also available. Note the alternative forms for the constant in the Wishart Distribution (2.4):

\[ A_n^{(p)} = \frac{\pi^{pn/2}}{2^p \Gamma_p^{(n/2)}} = \frac{\pi^{pn/2}}{\Gamma_p^{(n/2)} \Gamma_p^{(n-1/2)} \ldots \Gamma_p^{(n-p+1/2)}} \]

\[ \Gamma_p^{(n/2)} = \frac{\pi^{pn/2}}{2^p \Gamma_p^{(n/2)} \Gamma_p^{(n-1/2)} \ldots \Gamma_p^{(n-p+1/2)}} \]
3. Preliminaries: the Chi Distribution

Let $Z$ be a $p \times n$ matrix with the factorization $Z = LR$ introduced in Section 2. We now consider a further factorization involving a splitting of the matrix $R$. The orthonormal row vectors in $R$ generate a $p$ dimensional subspace $L(R)$ in $E_n$. Let $D$ be a $p \times n$ semiorthogonal matrix that provides a basis for $L(R)$, a basis that depends on $L(R)$ and not otherwise on $R$. Such an orthonormal basis $D$ can be chosen as a smooth function (a.e.) on $R_{p,n}$ by various procedures; for example, by successively orthonormalizing the projection into $L(R)$ of the first $p$ usual basis vectors of $E^n$ (see, for example, Fraser (1968), p. 228). Then let $D$ be the $p \times p$ orthogonal matrix that produces the special orthonormal basis $R$ from the standard basis $D : R = OD$ and

$$Z = LOD = CD$$

(3.1)

where $LO = C$ is a $p \times p$ matrix with values arbitrary in $E^{p^2}$.

Let $O_p$ be the orthogonal group $R_{p,p}$ and let $D_{p,n}$ be the set of matrices $D$ formed by a smooth procedure as discussed in the preceding paragraph; $D_{p,n}$ provides a representation for almost all of the Grassman manifold $G_{p,n}$. (The set of $p$-dimensional subspaces of $E^n$ is called the Grassman manifold $G_{p,n}$.) The unique factorization $R = OD$ shows that $D_{p,n}$ forms a cross-section to the orbits of the orthogonal group $O_p$ applied to $R_{p,n}$ by the left matrix multiplication. The dimension of $O_p$ is $p(p-1)/2$ and the dimension of $D_{p,n}$ is $pn-p^2$.

Now let $C_p$ be the general linear group of $p \times p$ matrices $C = LO$. For the full-measure retained portion of $E^{pn}$ the unique factorization $Z = CD$ shows that $D_{p,n}$ forms a cross-section to the orbits of the group $C_p$ applied by left matrix multiplication. The volume measure tangent to the orbit $C_p D$ at $D$ coincides with the volume measure $dC$ in the neighbourhood of the identity. The natural measure for the cross-section $D_{p,n}$ is given locally at $D$ by the projection on the orthogonal space to the orbit $C_p D$ at $D$; let $dD$ designate this Euclidean volume orthogonal to the $C_p$ orbits. Then in the neighbourhood of the cross-section $D_{p,n}$ the full measure $dZ$ can be factored as $dZ = dCdD$ into orthogonal Euclidean components.

Now let $Z$ be a $p \times n$ matrix with a standard normal distribution and consider the joint distribution of $(C, D)$ obtained from the preceding factorization $Z = CD$. The invariant $(C_p)$ measure on $E^{pn}$ that agrees with Euclidean volume at the
cross-section $D_{p,n}$ is $|C|^{-n} dZ$; the invariant measure on $C_p$ that agrees with tangent volume at any point $D$ is $|C|^{-p} dC$; accordingly we have

$$dZ = |C|^{n-p} dC dD. \quad (3.1)$$

We now take care of the further factorization $C = LO$ by using (2.1) with $n = p$:

$$dC = |L|^p |L|^{-1}_\Lambda dL dO, \quad (3.2)$$

where the formula assumes that $dO$ measures Euclidean volume orthogonal to the orbits $P_0^p$ at $O$. Accordingly we obtain

$$dZ = |L|^n |L|^{-1}_\Lambda dL dO dD. \quad (3.3)$$

The probability element for $Z$ factors so that the variables $L$, $O$, $D$ separate

$$(2\pi)^{-pn/2} \text{etr}(\frac{-1}{2} ZZ') dZ = (2\pi)^{-pn/2} \text{etr}(\frac{-1}{2} LL') \frac{|L|^n}{|L|_\Lambda} dL dO dD$$

$$= \frac{A^{(p)}_n}{(2\pi)^{pn/2} \text{etr}(\frac{-1}{2} LL')} \frac{|L|^n}{|L|_\Lambda} \frac{dL}{A^{(p)}_p} \frac{dO}{A^{(p)}_n} \frac{dD}{A^{(p)}_n}, \quad (3.4)$$

where the first normalizing factor comes from (2.2) and the second from (2.2) with $n = p$ and where $A^{(p)}_{n:p} = A^{(p)}_n / A^{(p)}_p$ is the volume of the manifold $D_{p,n}$ calculated orthogonal to the orbits $C_D$. Thus $L$, $O$, and $D$ are independent; $L$ is $\Delta(x)_n$; $O$ is uniform on $O_0^p$ (with total volume $A^{(p)}_p$ measured orthogonal to $P_0^p$ orbits); and the Grassman component $D$ is uniform on $D_{p,n}$ with volume measured orthogonal to the orbits $C_D$.

The first two factors in (3.4) can be combined giving the probability element for $(C, D)$:

$$\frac{A^{(p)}_{n:p}}{(2\pi)^{pn/2} \text{etr}(\frac{-1}{2} CC')} |C|^{n-p} dC \cdot \frac{dD}{A^{(p)}_n}. \quad (3.5)$$

Thus $C$ and $D$ are statistically independent; and $C$ is said to have the chi distribution $\chi_p(n)$ on $E^p$. Note that $C$ and $C'$ have the same distribution.
The link between (3.4) and (3.5) shows that \( C = \lambda_0 \) is \( \chi_p(n) \) if and only if \( L \) is \( \Delta \chi_p(n) \) and \( O \) is uniform on \( O_p \).

Now let \( Z \) be a \( p \times n \) matrix with a rotationally symmetric distribution under right multiplication by an \( n \times n \) orthogonal matrix:

\[
f(Z) = f_p(ZZ').
\]

(3.6)

The probability element for \( Z = CD \) can be factored so the variables \( C \) and \( D \) separate

\[
f(Z) dZ = A_p^{(p)} f_p(\text{CC'}) |C|^{n-p} dC \cdot \frac{dD}{A_p^{(p)}}
\]

(3.7)

where the normalizing constants are obtained by noting the integration properties of \( dD \) derived from (3.5). Thus \( C \) and \( D \) are statistically independent; the Grassman component \( D \) is uniform on \( D_{p,n} \) as described above; and \( C \) is said to have the generalized chi distribution \( \chi_p(f_p^n; n) \) with kernel \( f_p^n \). Note that \( C = \lambda_0 \) is \( \chi_p(f_p^n; n) \) if and only if \( L \) is \( \Delta \chi_p(f_p^n; n) \) and \( O \) is uniform on \( O_p \). Also note that the \( \chi_p(f_p^n; n) \) distribution of \( C \) produces the Wishart \( W_p(f_p^n; n) \) distribution for \( S = \text{CC'} \), or \( S = C'C \).

4. Triangular Chi Decompositions

Let \( Z = (Z_1 : Z_2) \) be a \( p \times (n_1 + n_2) \) matrix and consider the triangular chi factorization in Section 2 for \( Z, Z_1, \) and \( Z_2 \):

\[
Z = LR = L(\tilde{R}_1 : \tilde{R}_2) = (L_1 R_1 : L_2 R_2) = (Z_1 : Z_2).
\]

(4.1)

We now investigate derived distributions for a root \( F \)-type variable \( E = L_2^{-1} L_1 \), root-beta type variables \( A_1 = L_1^{-1} L_1 \), and a \( t \)-type variable \( H = L_2^{-1} Z_1 \). Note that the root-beta variables \( A_1, A_2 \) provide the split of the Stiefel component \( R \):

\[
R = (\tilde{R}_1 : \tilde{R}_2) = (A_1 R_1 : A_2 R_2)
\]

(4.2)

and also note that \( E = L_2^{-1} L_1 = A_2^{-1} A_1 \).

A distribution for \( Z \) can be expressed in terms of \( L, R_1, R_2 \) and a further variable of dimension \( p(p+1)/2 \) such as the root \( F \) variable \( E \) or a root beta variable \( A_1 \) or \( A_2 \).
For the F factorization we have
\[
\begin{align*}
\frac{dz_1 dz_2}{dL_1 dL_2 dR_1 dR_2} &= |L_1|^{n_1} |L_1|^{-1} |L_2|^{n_2} |L_2|^{-1} dL_1 dL_2 dR_1 dR_2 \\
&= \frac{|E|^{n_1} |L_2|^{n_1+n_2}}{|L_2| \Delta} dE dL_2 dR_1 dR_2 \\
&= \frac{|E|^{n_1} |L_2|^{n_1+n_2}}{|L_2| \Delta} |A_2| \sqrt{E} dE dL dR_1 dR_2 \\
&= \frac{|L_2|^{n_1+n_2}}{|L| \Delta} \frac{|E|^{n_1} |A_2|^{n_1+n_2}}{|A_2| \sqrt{E}} dE dL dR_1 dR_2 \\
&= \frac{|L_1|^{n_1} |L_2|^{n_2}}{|L_2| \Delta} \frac{|E|^{n_1}}{|(n_1+n_2)/2\sqrt{E}|} \frac{|I+EE'|^{(p+1)/2}}{|I+EE'|^{1/2}} dE dL dR_1 dR_2
\end{align*}
\]

where the last step uses $A_2^{-1}(A_1 : A_2) = (E : I)$ with the consequence $|A_2|^{-1} = |I+EE'|^{1/2}$. For the beta factorization we have
\[
\begin{align*}
\frac{dz_1 dz_2}{dL_1 dL_2 dR_1 dR_2} &= |L_1|^{n_1} |L_1|^{-1} |L_2|^{n_2} |L_2|^{-1} dL_1 dL_2 dR_1 dR_2 \\
&= \frac{|L_1|^n |L_2|^n}{|L_1| \Delta} \frac{|L_2|}{|L_2| \sqrt{E}} dL_1 dL_2 dR_1 dR_2 \\
&= \frac{|L_1|^n}{|L_1| \Delta} \frac{|L_2|^n}{|L_2| \Delta} \frac{|L_1|}{|L_2| \Delta} \frac{|L_1|}{|L_2| \Delta} dL_1 dL_2 dR_1 dR_2 \\
&= \frac{|L_1|^n}{|L_1| \Delta} \frac{|L_1|^n}{|L_1| \Delta} \frac{|L_1|^n}{|L_1| \Delta} dL_1 dR_1 dR_2 \\
&= \frac{|L_1|^n}{|L_1| \Delta} \frac{|L_1|^n}{|L_1| \Delta} \frac{|L_1|^n}{|L_1| \Delta} \frac{|L_1|^n}{|L_1| \Delta} dL_1 dR_1 dR_2
\end{align*}
\]

Now consider $Z = (Z_1 : Z_2)$ with a rotationally symmetric density $f(Z) = f_0(ZZ')$ under right multiplication by an orthogonal matrix. We first obtain the joint distribution for $L$, $R_1$, $R_2$, $E$ by using the change of variable relation (4.3):
\[ f(z)dz = f_p(zz')dz_1dz_2 \]

\[ = f_p(LL') \cdot \frac{|L|^{n_1+n_2}}{|L|_\Delta} \cdot \frac{|E|^{n_1}}{(n_1+n_2)/2} \cdot \frac{|I+EE'|^{(p+1)/2}}{|I+EE'|\_\nu} \cdot \frac{dE}{|E|_\Delta} \cdot dL \cdot dR_1dR_2 \]

\[ = A^{(p)}_{n_1} A^{(p)}_{n_2} \cdot \frac{|L|^{n_1+n_2}}{|L|_\Delta} \cdot \frac{dL}{A^{(p)}_{n_1}} \cdot \frac{dR_1}{A^{(p)}_{n_1}} \cdot \frac{dR_2}{A^{(p)}_{n_2}} \cdot \frac{|E|^{n_1}}{(n_1+n_2)/2} \cdot \frac{|I+EE'|^{(p+1)/2}}{|I+EE'|\_\nu} \]

(4.5)

Thus, L, R_1, R_2, E are statistically independent; L is a generalized triangular chi \( \Delta_{p}\left(f_p; n_1+n_2\right) \); the Stiefel components R_1 and R_2 are uniform on \( \mathbb{R}_{p,n_1} \) and \( \mathbb{R}_{p,n_2} \) respectively; and E has a triangular root F distribution \( \Delta_{F_{p}}(n_1,n_2) \) distribution. Note that the Stiefel component R with a uniform distribution on \( \mathbb{R}_{p,n_1+n_2} \) has been decomposed into two Stiefel components with \( n_1 \) and \( n_2 \) degrees of freedom and a root F component E that describes the "ratio" of the "coefficients" of the two components.

If we set \( F = EE' \) (\( dF = 2^{p}|E|_\nu dE \)) then F has a disguised F distribution \( \mathbb{V}F_{p}(n_1', n_2) \) with element

\[ \frac{A^{(p)}_{n_1} A^{(p)}_{n_2}}{A^{(p)}_{n_1+n_2}} \cdot \frac{|F|^{n_1/2}}{|I+F|^{(p+1)/2}} \cdot \frac{|I+F|^{(p+1)/2}}{|I+F|\_\nu} \cdot \frac{dF}{2^{p}|F|^{(p+1)/2}} \]

(4.6)

where the disguising factor \( |I+F|^{(p+1)/2}/|I+F|\_\nu \) is superimposed on the usual matrix F distribution. Note the constant.
\[ \frac{A_{n_1+n_2}^{(p)}}{A_{n_1}^{(p)} A_{n_2}^{(p)}} = 2^{-p} B_p \left( \frac{n_1}{2}, \frac{n_2}{2} \right) = 2^{-p} \frac{\Gamma_p(n_1/2) \Gamma_p(n_2/2)}{\Gamma_p(n_1+n_2/2)}. \]

We next obtain the joint distribution for \( L, R_1, R_2 \), and say \( A_1 \). For this note that \( E = \nu_1^{-1}_2 \nu_1^{-1}_1, A_2^{-1} R = A_2^{-1}(A_1 R_1 : A_2 R_2) = (E R_1 : R_2' \nu_1^{-1}_2) \), and thus that \( A_2^{-1} \) is the PLT root of \( I + EE' \) and similarly \( A_1^{-1} \) is the PLT root of \( I + E^{-1}E'^{-1} \). It follows that \( E, A_1, A_2 \) are one-one equivalent. By using (4.4) in place of (4.3) we obtain a separation similar to (4.5); the difference is that the following element for \( A_1 \) replaces the earlier probability element for \( B \):

\[ \frac{A_{n_1}^{(p)} A_{n_2}^{(p)}}{A_{n_1+n_2}^{(p)}} \left| A_1 \right|^{n_1} \left| A_2 \right|^{n_2} \frac{dA_1}{A_1^\Delta A_2^\Delta} \]

\[ = \frac{A_{n_1}^{(p)} A_{n_2}^{(p)}}{A_{n_1+n_2}^{(p)}} \left| A_1 \right|^{n_1} \left| I - A_1 A_1' \right|^{(n_2-p-1)/2} \frac{dA_1}{A_1^\Delta}. \]

(4.7)

this is a triangular root beta distribution \( B_p^{1/2}(n_1, n_2) \).

The symmetric positive definite matrix \( B = A_1 A_1' \) subject to \( 0 < B < I \) has the matrix beta distribution \( B_p(n_1/2, n_2/2) \) with probability element

\[ \frac{A_{n_1}^{(p)} A_{n_2}^{(p)}}{A_{n_1+n_2}^{(p)}} \left| B \right|^{n_1/2} \left| I - B \right|^{(n_2-p-1)/2} \frac{dB}{2^p |B|^{(p+1)/2}}. \]

(4.8)

On the other hand, if we put \( B = (I+F)^{-1}F \) we have

\[ dF = d((I-B)^{-1}-I) = d(I-B)^{-1} = \left| I - B \right|^{-(p+1)} d(I-B) = \left| I - B \right|^{-p-1} dB \]

and obtain the probability element

\[ \frac{A_{n_1}^{(p)} A_{n_2}^{(p)}}{A_{n_1+n_2}^{(p)}} \left| B \right|^{n_1/2} \left| I - B \right|^{p+1/2} \frac{dB}{\left| I - B \right|^{\Delta} 2^p |B|^{2}}. \]

(4.9)
over the space of positive definite matrices $B$ satisfying $0 < B < I$. This is a disguised beta distribution

$$\mathcal{V}_{B_p}(n_1/2, n_2/2),$$

with a disguising factor $\frac{p+1}{\sqrt{|I-B|}}$. We now derive the decomposition involving the $t$-type variable $H = L_2^{-1}Z_1$. Specifically we find the joint probability element for $L, R_2, H$ using the differential relation (4.3) together with (2.1) reexpressed as $dH = dR_1 = |E|^{n_1} |E|^{-1}_\Delta dEdR_1$:

$$f(Z)dz = f_p(ZZ')dz_1dz_2$$

$$= f_p(LL') \frac{|L|^{n_1+n_2}}{|L\Delta|} \cdot \frac{1}{(n_1+n_2)/2} \cdot \frac{|I+HH'|^{(p+1)/2}}{|I+HH'|^v} dLdHdR_2$$

$$= A^{(p)}_{n_1+n_2} f_p(LL') \frac{|L|^{n_1+n_2}}{|L\Delta|} \cdot \frac{dR_2}{A^{(p)}_{n_2}}$$

$$\cdot \frac{A^{(p)}_{n_2}}{A^{(p)}_{n_1+n_2}} \cdot \frac{1}{(n_1+n_2)/2} \cdot \frac{|I+HH'|^{(p+1)/2}}{|I+HH'|^v} dH. \quad (4.10)$$

Thus $L, R_2, H$ are statistically independent; $L$ is $\Delta_{p}^{(p)}(f_p, n_1+n_2)$; the Stiefel component $R_2$ is uniform on $\mathcal{R}_{p,n_2}$, and $H$ has a disguised $t$-distribution $\mathcal{V}_{T_p \times n_1}(n_2)$, for example, see Fraser and Ng (1977). Also note that $F = HH' = L_2^{-1}Z_1Z_1' L_2^{-1}$

$$= L_2^{-1}L_1L_1' L_2^{-1} = EE'$$

is $\mathcal{V}_{F_p}(n_1, n_2)$.

As a final decomposition we consider the variables $(L, R_2, \tilde{R}_1)$ where $\tilde{R}_1 = A_1 R_1$ is the "first sample" portion of the Stiefel component. For this we use the differential relation (4.4) together with (2.1) expressed as $d\tilde{R}_1 = |A_1|^{n_1} |A_1|^{-1}_\Delta dA_1 dR_1$

and obtain the probability element

$$f(Z)dz = f_p(ZZ')dz_1dz_2$$

$$= f_p(LL') \frac{|L|^{n_1+n_2}}{|L\Delta|} \cdot \frac{A^{n_2}_2}{|A_2|^{(p+1)}} dLd\tilde{R}_1dR_2$$

13
\[ = A_{n_1+n_2}^{(p)} f_p(\text{LL}') \left| \frac{L_{1+n_2}}{L_\Delta} \right| \text{dL} \cdot \frac{\text{dR}_2}{A_{n_2}^{(p)}} \]

\[ \frac{A_{n_2}^{(p)}}{A_{n_1+n_2}^{(p)}} \left| I - \tilde{R}_2 \tilde{R}_1 \right|^{(n_2-p-1)/2} \tilde{R}_1^{-1} \left( I - \tilde{R}_2 \tilde{R}_1 \right) > 0 \]

(4.11)

where the last step uses \( I = RR' = \tilde{R}_1 \tilde{R}_1 + \tilde{R}_2 \tilde{R}_2 = A_1 A_1' + A_2 A_2' \).

Thus \( L, R_2, \tilde{R}_1 \) are statistically independent, and the new variable \( \tilde{R}_1 \) has a distribution similar to the inverted \( t \) distribution discussed by Dickey (1967). In particular we see that the uniform distribution for \( R \) can be represented as the product of two component distributions: the inverted \( t \) distribution for \( \tilde{R}_1 \) and a uniform distribution for \( R_2 = ((\tilde{R}_2 \tilde{R}_1)^T)^-1 \tilde{R}_2 \) where the superscript \( T \) denotes the PLT square root matrix. In conclusion note the connection between the variables \( H \) and \( \tilde{R}_1 : H = L_2^{-1} Z_1 = L_2^{-1} L_1 R_1 = A_2^{-1} A_1' R_1 = A_2^{-1} \tilde{R}_1 \).

5. Chi Decompositions

Let \( Z = (Z_1 : Z_2) \) be a \( p \times (n_1+n_2) \) matrix and consider the chi factorization in Section 3 for \( Z, Z_1, Z_2 \):

\[ Z = CD = C(D_1 : D_2) = (C_1 D_1 : C_2 D_2) = (Z_1 : Z_2). \]

(5.1)

We now investigate derived distributions for a root \( F \) type variable \( E = C_2^{-1} C_1 \) and a \( t \) type variable \( H = C_2^{-1} Z_1 \). Note that root beta variables \( A_i = C^{-1} C_i \) provide a split of the Grassman component \( D \):

\[ D = (\tilde{D}_1 : \tilde{D}_2) = (A_1 D_1 : A_2 D_2) \]

and also note that \( E = C_2^{-1} C_1 = A_2^{-1} A_1' \).

A distribution for \( Z \) can be expressed in terms of \( C, D_1, D_2 \) and the root \( F \) type variable \( E = C_2^{-1} C_1 \) which has dimension \( p^2 \). For the differentials we have
\[
\frac{dZ_1 dZ_2}{dC} = |C|^{n_1-p} |A_1|^{n_2-p} \frac{dE}{dC} \frac{dD_1 dD_2}{dE},
\]

where the last step uses \( A_2^{-1}(A_1 : A_2) = (E : I) \) with the consequence \( |A_2|^{-1} = |I+EE'|^{1/2} \).

Now consider \( Z = (Z_1 : Z_2) \) with a rotationally symmetric density \( f(Z) = f_p(ZZ') \) under right multiplication by an orthogonal matrix. We first obtain the joint distribution for \( L, D_1, D_2, E \) using the preceding change of variable for the differential:

\[
f(Z) = f_p(ZZ') dZ_1 dZ_2
\]

\[
= f_p(CC') |C|^{n_1+n_2-p} \frac{dE}{dC} \frac{|E|^{n_1-p}}{(n_1+n_2)^2/2} \frac{dE dD_1 dD_2}{dE},
\]

Thus \( C, D_1, D_2, E \) are statistically independent; \( C \) is generalized chi \( \chi_p(f_p, n_1+n_2) \), the Grassman components \( D_1 \) and
D_2 are uniform on \( D_{p,n_2} \) and \( D_{p,n_1} \), respectively, and \( E \) has a root F distribution \( F_{p}^{1} (n_1, n_2) \) on \( E \). Note that the Grassman component \( D \) on \( V_{p,n_1+n_2} \) has been decomposed into two Grassman components with \( n_1 \) and \( n_2 \) degrees of freedom and a root F component describing the ratio of the two Grassman components. Also note that \( E \) and \( E' \) have the same distribution.

If we set \( F = EE' \) or \( F = E'E \) we obtain the following probability element for \( F \)

\[
\frac{A_{n_1}^{(p)}A_{n_2}^{(p)}}{A_{n_1+n_2}^{(p)}} |I+F|^\frac{n_1}{2} |I+F|^\frac{n_2}{2(n_1+n_2)} (n_1+n_2)/2 \frac{dF}{2^p|F|^{(p+1)/2}} \tag{5.4}
\]

over the \( p \times p \) positive definite matrices; this is called the F distribution \( F_{p}^{1} (n_1, n_2) \). It follows easily that if \( F \) is \( V_{p}^{1} (n_1, n_2) \) and \( O \) is uniform on the \( p \times p \) orthogonal matrices, then \( O'FO \) is \( F_{p}^{1} (n_1, n_2) \).

Again following the pattern from Section 4 we have that \( B = (I+F)^{-1}F \) has the density

\[
\frac{A_{n_1}^{(p)}A_{n_2}^{(p)}}{A_{n_1+n_2}^{(p)}} |B|^\frac{(n_1-p-1)/2}{|I-B|^\frac{(n_2-p-1)/2}{2^p} db} \tag{5.5}
\]

over the \( p \times p \) positive definite matrices satisfying \( 0<B<I \). This is the beta distribution \( B_{p}^{1} (n_1/2, n_2/2) \) in \( (4.8) \). As we found earlier for the \( F \), chi, and Wishart distributions, the beta distribution is an average of rotated disguised betas: if \( B \) is \( V_{p}^{1} (n_1/2, n_2/2) \) and \( O \) is uniform on \( O_{p} \), then \( O'BO \) is \( B_{p}^{1} (n_1/2, n_2/2) \).

A root-beta type distribution can also be derived in the pattern followed in Section 4. Specifically we derive the distribution for the variables \( C, D_1, D_2 \) and \( A = (A_1 : A_2) \).

For this note that \( A \) provides the coordinates for \( D = (A_1 D_1 : A_2 D_2) \) in \( D_{p,n_1+n_2} \) relative to \( D_1 \) in \( D_{p,n_1} \) and \( D_2 \) in \( D_{p,n_2} \); the coordinates will of course depend on the
representations used for the Grassman manifolds. For the differentials we have

\[ dz_1 dz_2 = |c_1|^{n_1-p} |c_2|^{n_2-p} dc_1 dc_2 dD_1 dD_2 \]

\[ = |c|^{n_1+n_2-p} |a_1|^{n_1-p} |a_2|^{n_2-p} dC dA dD_1 dD_2, \]  

(5.6)

where we have used that \(|c|^{-p} dc_1 |c|^{-p} dc_2\) is invariant under the group \(C_p\) and agrees with Euclidean volume at \((a_1 : a_2)\), that \(|c|^{-p} dc\) is invariant on the orbits \(C_p(a_1 : a_2)\) and agrees with Euclidean volume on the orbit at \((a_1 : a_2)\), and that \(dA\) is taken to be Euclidean volume orthogonal to the preceding orbits.

Now consider \(Z = (Z_1 : Z_2)\) with a rotationally symmetric density \(f(Z) = f_p(ZZ')\) under right multiplication by an orthogonal matrix. Using the integration results from (5.3) we then obtain

\[ f(Z) dz = f_p(ZZ') dZ_1 dZ_2 \]

\[ = A_{n_1+n_2:p}^{(p)} f_p(CC') |c|^{n_1+n_2-p} dc \cdot \frac{dD_1}{A_{n_1:p}^{(p)}} \cdot \frac{dD_2}{A_{n_2:p}^{(p)}} \]

\[ = \frac{A_{n_1:p}^{(p)}}{A_{n_1+n_2:p}^{(p)}} \cdot \frac{A_{n_2:p}^{(p)}}{A_{n_1+n_2:p}^{(p)}} |a_1|^{n_1-p} |a_2|^{n_2-p} dA; \]

this is a root beta distribution \(\beta_{n_1/2, n_2/2}^{1/2}\) with dimension \(p^2\). The differential \(dA\) does not seem to lend itself to a more convenient representation without a specification of the Grassman manifold representations.

We now derive a decomposition involving the \(t\)-type variable \(H = C_2^{-1} Z_1\). Specifically we find the joint probability element for \(C, D, H\) and use the differential relation (5.2) together with (3.1) reexpressed as \(dH = dE dD_1 = |E|^{n_1-p} dE dD_1;\)
\[ f(z)dz = f_p(zz')dz_1dz_2 = f_p(cc')|c|^{-n_1+n_2-p}dc \cdot |I+HH'|^{-\frac{n_1+n_2}{2}}dHd_2 \]
\[ = \frac{A^{(p)}_{n_1+n_2}}{A^{(p)}_{n_2}A^{(p)}_{n_1+n_2}}f_p(cc')|c|^{-n_1+n_2-p}dc \cdot \frac{d_2}{A^{(p)}_{n_2}} \cdot |I+HH'|^{-\frac{n_1+n_2}{2}}dH. \]

Thus \( C, D_2, \) and \( H \) are statistically independent; and the new variable \( H \) has a t-distribution \( T_{p \times n_1}^{n_2}(n_2) \). Also we have that this t-distribution is an average of rotated disguised t-variables: if \( H \) is \( V_T^{p \times n_1} \), then \( 0'HO \) is \( T_{p \times n_1}^{n_2}(n_2) \). Also if \( H \) is \( T_{p \times n_1}^{n_2}(n_2) \), then \( HH' \) is \( F_p(n_1, n_2) \).

6. Some Scaled Distributions

Scaled versions of the distributions in Sections 2 to 5 are easily derived using the Jacobians available from Section 1.

Suppose that \( I_0 \) is \( \Delta \chi_p^{(n)} \), then the new variable \( L = TL_0 \) where \( T \) is \( p \times p \) PLT is triangular chi \( \Delta \chi_p^{(n; T)} \) and \( LL' \) is Wishart \( T_p^{(n; \Sigma)} \) with \( \Sigma = TT' \).

Now suppose that \( F_0 \) is \( \Delta F_p^{1/2}(n_1, n_2) \), then the new variable \( E = TE_0 \) where \( T \) is \( p \times p \) PLT is the triangular root F distribution \( \Delta F_p^{1/2}(n_1, n_2; T) \) with distribution that depends only on \( \Sigma = TT' \).

Similar scaled versions are available for the disguised F distribution and the disguised beta distribution.

Now suppose that \( H_0 \) is \( V_T^{p \times n_1}(n_2) \), then the new variable \( H = TH_0B \) where \( T \) is PLT and \( B \) is \( n_1 \times n_1 \) non-singular, has probability element...
\[ \frac{A_{n_2}^{(p)}}{n_2} \cdot \frac{|\Sigma_1|^{(n_2-p-1)/2}}{|\Sigma_1 + H\Sigma_2^{-1}H'|^{(n_1+n_2-p-1)/2}} \frac{|\Sigma_2|^{-p/2}}{|\Sigma_1 + H\Sigma_2^{-1}H'|^{n_1/2}} \mathrm{d}H \]

over \( E \); this is called the disguised T distribution \( \mathcal{V}_{T_{p \times n_1}}(n_2; \Sigma_1; \Sigma_2) \) with \( \Sigma_1 = TT', \Sigma_2 = B'B \).

In a parallel way for the chi type distributions we can define scaled versions. Suppose that \( C_0 \) is chi \( \chi_p(n) \); then the new variable \( C = AC_0 \) with \( A \in \mathbb{R}^{p \times p} \) nonsingular is \( \chi_p(n; A) \) and has a distribution that depends only on \( \Sigma = AA' \). Suppose \( Z \) is \( N_{p \times n}(0; I \otimes \Sigma) \), and \( Z = CD \) as in Section 3, then \( C \) is \( \chi_p(n; A) \).

Now suppose that \( E_0 \) is \( F_{p \times n_1}^{1/2}(n_1, n_2) \) then the new variable \( E = AE_0 \) where \( A \in \mathbb{R}^{p \times p} \) nonsingular is \( F_{p \times n_1}^{1/2}(n_1, n_2; A) \) with distribution that depends only on \( E = AA' \).

Similar scaled versions are available for the F distribution and beta distribution.

Now suppose that \( H_0 = T_{p \times n_1}(n_2) \). Then the new variable \( H = AH_0B \), with \( A \) and \( B \) nonsingular, has probability element

\[ \frac{A_{n_1}^{(p)}}{n_1 + n_2} \cdot \frac{|\Sigma_1|^{-n_1/2}}{|I + \Sigma_1^{-1}H\Sigma_2^{-1}H'|^{(n_1+n_2)/2}} |\Sigma_2|^{-p/2} \mathrm{d}H \]

over \( E \) called the T distribution \( T_{p \times n_1}(n_2; \Sigma_1, \Sigma_2) \) with \( \Sigma_1 = AA' \) and \( \Sigma_2 = B'B \).

References


Zusammenfassung

L. Bishop, D.A.S. Fraser, K.W. Ng:

In dieser Arbeit wird die Zerlegung multivariater Verteilungen, welche Invarianzeigenschaften gegenüber einer Rotationsgruppe besitzen, in unabhängige Komponenten untersucht. Es wird gezeigt, daß man so auf natürliche Weise als Faktoren eine Vielzahl von Wahrscheinlichkeitsgesetzen erhält, die durch Zufallsgrößen mit uniformer, normaler, Chi-quadrat, Wishart-, matrizieller F-, t- oder Beta-Verteilungen erzeugt sind oder bei denen es sich um Verallgemeinerungen der genannten Verteilungen handelt. Spezielle Beachtung wird dem Fall unterteilter Matrizen geschenkt.

Summary

L. Bishop, D.A.S. Fraser, K.W. Ng: Some Decompositions of Spherical Distributions

In this article the factorization of multivariate distributions exhibiting invariance properties with respect to a group of rotations into independent components is investigated. It is shown that one obtains in a natural manner a great number of probability laws, which are uniform, normal, chi-square, Wishart, matrix-F, matrix-t or matrix-Beta-distributions or which are generalizations of these. In particular the case involving partitioned matrices is treated.

Résumé

L. Bishop, D.A.S. Fraser, K.W. Ng:

Dans ce travail la partition des distributions multivariées possédant des propriétés d'invariance par rapport à un groupe de rotations en composantes indépendantes est analysée. Il est démontré, qu'on obtient de cette façon d'une manière naturelle un grand nombre de lois de probabilités, qui sont engendrées par des quantités aléatoires à distributions uniformes, normales, de chi-carré, de Wishart ou du type matriciel F, t ou beta, ou qui sont des généralisations des distributions mentionnées ci-dessus. Le cas résultant de matrices partagées est examiné de plus près.
В этой работе рассматривается разложение мультивариантных распределений, имеющих свойства инвариантности по отношению к группе вращений, на независимые компоненты. Показывается, что таким собственным образом получается множество вероятностных законов, которые образуются случайными величинами с однородными, нормальными, \( \chi^2 \) распределением, с распределением Вискеarta и матричными \( F, t, \) или \( \beta \)-распределением или они являются обобщениями вышеназванных распределений. Особое внимание обращается на случай подразделенных матриц.
Adaptive Mustererkennung
(Adaptive pattern recognition)*

1. Vorbemerkung

Wenn eine Frau zu ihrem Mann "Du Scheusal" sagt, kann es bedeuten, daß sie damit ihren Abscheu ausdrücken will; es kann aber auch sein, daß das Wort zärtlich gemeint ist. Daß sie weiß, wie es gemeint ist, ist unerheblich. Aber für den Empfänger der Nachricht interessieren wir uns in der Mustererkennung. Ist der Mann in der Lage und wenn ja bei Benutzung welcher Hilfsmittel, die Nachricht zu semantisieren, zu verstehen, einzuordnen? Vielleicht kann er aus der Situation, aus dem Tonfall heraus usw. das "gemeinte Muster" erkennen. Er hat zwei Musterklassen, sagen wir "Liebe" und "Abscheu", und die Mustererkennung bedeutet, daß er die Nachricht in eine der beiden Klassen einordnet, d.h. das Wort (mit allen Begleitumständen) als Muster der Klasse "Liebe" oder der Klasse "Abscheu" identifiziert.

Allgemein besteht die Aufgabe der Mustererkennung darin, eine ankommende Information oder Nachricht als Muster kohärenter Klassen zu identifizieren.

\[ \text{Information (Zeichen)} \rightarrow \text{Muster der Klasse } k_1 \rightarrow \text{Muster der Klasse } k_2 \cdots \rightarrow \text{Muster der Klasse } k_m \]

In den Verhaltenswissenschaften, auch in den Wirtschaftswissenschaften, wird der Mustererkennung heute große Beachtung geschenkt. In der Psychologie weiß man z.B., daß Individuen nicht auf Zeichen selbst in Form von Aggression oder Freund schaft etc. reagieren, sondern auf die von ihnen identifizierten ("erkennnten") Muster. Z.B. kann der Ehemann in unserem Eingangsbeispiel solange nicht eigentlich reagieren, wie er das Muster nicht erkannt (identifiziert) hat, d.h. nicht weiß, wie der Zuruf seiner Gattin gemeint war. Übrigens kann er sich dabei täuschen. Darauf gehen wir später detaillierter ein. Oder in der Ökonomie werden z.B. Konsumentenverhalten nach ihrem