CHAPTER 5

Matrix Factorizations and Jacobians

This chapter contains a collection of results concerning the factorization of matrices and the Jacobians of certain transformations on Euclidean spaces. The factorizations and Jacobians established here do have some intrinsic interest. Rather than interrupt the flow of later material to present these results, we have chosen to collect them together for easy reference. The reader is asked to mentally file the results and await their application in future chapters.

5.1. MATRIX FACTORIZATIONS

We begin by fixing some notation. As usual, R^n denotes n-dimensional coordinate space and $\mathcal{L}_{m,n}$ is the space of $n \times m$ real matrices. The linear space of $n \times n$ symmetric real matrices, a subspace of $\mathcal{L}_{n,n}$, is denoted by $\mathcal{S}_n$. If $S \in \mathcal{S}_n$, we write $S > 0$ to mean S is positive definite and $S \geqslant 0$ means that S is positive semidefinite.

Recall that $\mathcal{F}_{p,n}$ is the set of all $n \times p$ linear isometries of R^p into R^n, that is, $\Psi \in \mathcal{F}_{p,n}$ iff $\Psi'\Psi = I_p$. Also, if $T \in \mathcal{L}_{n,n}$, then $T = \{t_{ij}\}$ is *lower triangular* if $t_{ij} = 0$ for $i < j$. The set of all $n \times n$ lower triangular matrices with $t_{ii} > 0$, $i = 1, \ldots, n$, is denoted by G_T^+. The dependence of G_T^+ on the dimension n is usually clear from context. A matrix $U \in \mathcal{L}_{n,n}$ is upper triangular if U' is lower triangular and G_U^+ denotes the set of all $n \times n$ upper triangular matrices with positive diagonal elements.

Our first result shows that G_T^+ and G_U^+ are closed under matrix multiplication and matrix inverse. In other words, G_T^+ and G_U^+ are groups of matrices with the group operation being matrix multiplication.

Proposition 5.1. If $T = \{t_{ij}\} \in G_T^+$, then $T^{-1} \in G_T^+$ and the ith diagonal element of T^{-1} is $1/t_{ii}$, $i = 1,\dots, n$. If T_1 and $T_2 \in G_T^+$, then $T_1T_2 \in G_T^+$.

Proof. To prove the first assertion, we proceed by induction on n. Assume the result is true for integers $1, 2,\dots, n-1$. When T is $n \times n$, partition T as

$$T = \begin{pmatrix} T_{11} & 0 \\ T_{21} & t_{nn} \end{pmatrix}$$

where T_{11} is $(n-1)\times(n-1)$, T_{21} is $1 \times (n-1)$, and t_{nn} is the (n, n) diagonal element of T. In order to be T^{-1}, the matrix

$$A \equiv \begin{pmatrix} A_{11} & 0 \\ A_{21} & a_{nn} \end{pmatrix}$$

must satisfy the equation $TA = I_n$. Thus

$$\begin{pmatrix} T_{11} & 0 \\ T_{21} & t_{nn} \end{pmatrix}\begin{pmatrix} A_{11} & 0 \\ A_{21} & a_{nn} \end{pmatrix} = \begin{pmatrix} T_{11}A_{11} & 0 \\ T_{21}A_{11} + t_{nn}A_{21} & t_{nn}a_{nn} \end{pmatrix} = \begin{pmatrix} I_{n-1} & 0 \\ 0 & 1 \end{pmatrix}$$

so $A_{11} = T_{11}^{-1}$, $a_{nn} = 1/t_{nn}$, and

$$A_{21} = -\frac{T_{21}T_{11}^{-1}}{t_{nn}}.$$

The induction hypothesis implies that T_{11}^{-1} is lower triangular with diagonal elements $1/t_{ii}$, $i = 1,\dots, n-1$. Thus the first assertion holds. The second assertion follows easily from the definition of matrix multiplication. □

Arguing in exactly the same way, G_U^+ is closed under matrix inverse and matrix multiplication. The first factorization result in this chapter is next.

Proposition 5.2. Suppose $A \in \mathcal{L}_{p,n}$ where $p \leqslant n$ and A has rank p. Then $A = \Psi U$ where $\Psi \in \mathfrak{F}_{p,n}$ and $U \in G_U^+$ is $p \times p$. Further, Ψ and U are unique.

Proof. The idea of the proof is to apply the Gram–Schmidt orthogonalization procedure to the columns of the matrix A. Let $a_1,\dots, a_p$ be the columns of A so $a_i \in R^n$, $i = 1,\dots, p$. Since A is of rank p, the vectors $a_1,\dots, a_p$ are linearly independent. Let $\{b_1,\dots, b_p\}$ be the orthonormal set of vectors obtained by applying the Gram–Schmidt process to $a_1,\dots, a_p$ in the order $1, 2,\dots, p$. Thus the matrix Ψ with columns $b_1,\dots, b_p$ is an element of $\mathcal{F}_{p,n}$ as $\Psi'\Psi = I_p$. Since $\operatorname{span}\{a_1,\dots, a_i\} = \operatorname{span}\{b_1,\dots, b_i\}$ for $i = 1,\dots, p$, $b_j'a_i = 0$ if $j > i$, and an examination of the Gram–Schmidt Process shows that $b_i'a_i > 0$ for $i = 1,\dots, p$. Thus the matrix $U \equiv \Psi'A$ is an element of G_U^+, and

$$\Psi U = \Psi\Psi'A.$$

But $\Psi\Psi'$ is the orthogonal projection onto $\operatorname{span}\{b_1,\dots, b_p\} = \operatorname{span}\{a_1,\dots, a_p\}$ so $\Psi\Psi'A = A$, as $\Psi\Psi'$ is the identity transformation on its range. This establishes the first assertion. For the uniqueness of Ψ and U, assume that $A = \Psi_1 U_1$ for $\Psi_1 \in \mathcal{F}_{p,n}$ and $U_1 \in G_U^+$. Then $\Psi_1 U_1 = \Psi U$, which implies that $\Psi'\Psi_1 = UU_1^{-1}$. Since A is of rank p, U_1 must have rank p so $\mathfrak{R}(A) = \mathfrak{R}(\Psi_1) = \mathfrak{R}(\Psi)$. Therefore, $\Psi_1\Psi_1'\Psi = \Psi$ since $\Psi_1\Psi_1'$ is the orthogonal projection onto its range. Thus $\Psi'\Psi_1\Psi_1'\Psi = I_p$—that is, $\Psi'\Psi_1$ is a $p \times p$ orthogonal matrix. Therefore, $UU_1^{-1} = \Psi'\Psi_1$ is an orthogonal matrix and $UU_1^{-1} \in G_U^+$. However, a bit of reflection shows that the only matrix that is both orthogonal and an element of G_U^+ is I_p. Thus $U = U_1$ so $\Psi = \Psi_1$ as U has rank p. □

The main statistical application of Proposition 5.2 is the decomposition of the random matrix Y discussed in Example 2.3. This decomposition is used to give a derivation of the Wishart density function and, under certain assumptions on the distribution of $Y = \Psi U$, it can be proved that Ψ and U are independent. The above decomposition also has some numerical applications. For example, the proof of Proposition 5.2 shows that if $A = \Psi U$, then the orthogonal projection onto the range of A is $\Psi\Psi' = A(A'A)^{-1}A'$. Hence this projection can be computed without computing $(A'A)^{-1}$. Also, if $p = n$ and $A = \Psi U$, then $A^{-1} = U^{-1}\Psi'$. Thus to compute A^{-1}, we need only to compute U^{-1} and this computation can be done iteratively, as the proof of Proposition 5.1 shows.

Our next decomposition result establishes a one-to-one correspondence between positive definite matrices and elements of G_T^+. First, a property of positive definite matrices is needed.

Proposition 5.3. For $S \in \mathbb{S}_p$ and $S > 0$, partition S as

$$S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}$$

where S_{11} and S_{22} are both square matrices. Then S_{11}, S_{22}, $S_{11} - S_{12}S_{22}^{-1}S_{21}$, and $S_{22} - S_{21}S_{11}^{-1}S_{12}$ are all positive definite.

Proof. For $x \in R^p$, partition x into y and z to be comformable with the partition of S. Then, for $x \neq 0$,

$$0 < x'Sx = y'S_{11}y + 2z'S_{21}y + z'S_{22}z.$$

For $y \neq 0$ and $z = 0$, $x \neq 0$ so $y'S_{11}y > 0$, which shows that $S_{11} > 0$. Similarly, $S_{22} > 0$. For $y \neq 0$ and $z = -S_{22}^{-1}S_{21}y$,

$$0 < x'Sx = y'\left(S_{11} - S_{12}S_{22}^{-1}S_{21}\right)y,$$

which shows that $S_{11} - S_{12}S_{22}^{-1}S_{21} > 0$. Similarly, $S_{22} - S_{21}S_{11}^{-1}S_{12} > 0$. □

Proposition 5.4. If $S > 0$, then $S = TT'$ for a unique element $T \in G_T^+$.

Proof. First, we establish the existence of T and then prove it is unique. The proof is by induction on dimension. If $S \in \mathcal{S}_p$ with $S > 0$, partition S as

$$S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}$$

where S_{11} is $(p-1) \times (p-1)$ and $S_{22} \in (0, \infty)$. By the induction hypothesis, $S_{11} = T_{11}T'_{11}$ for $T_{11} \in G_T^+$. Consider the equation

$$\begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} = \begin{pmatrix} T_{11} & 0 \\ T_{21} & T_{22} \end{pmatrix}\begin{pmatrix} T_{11} & 0 \\ T_{21} & T_{22} \end{pmatrix}',$$

which is to be solved for $T_{21} : 1 \times (p-1)$ and $T_{22} \in (0, \infty)$. This leads to the two equations $T_{21}T'_{11} = S_{21}$ and $T_{21}T'_{21} + T_{22}^2 = S_{22}$. Thus $T_{21} = S_{21}(T'_{11})^{-1}$, so

$$\begin{aligned} S_{22} &= T_{22}^2 + S_{21}(T'_{11})^{-1}\left(S_{21}(T'_{11})^{-1}\right)' \\ &= T_{22}^2 + S_{21}\left(T_{11}T'_{11}\right)^{-1}S_{12} = T_{22}^2 + S_{21}S_{11}^{-1}S_{12}. \end{aligned}$$

Therefore, $T_{22}^2 = S_{22} - S_{21}S_{11}^{-1}S_{12}$, which is positive by Proposition 5.3. Hence, $T_{22} = (S_{22} - S_{21}S_{11}^{-1}S_{12})^{1/2}$ is the solution for $T_{22} > 0$. This shows that $S = TT'$ for some $T \in G_T^+$. For uniqueness, if $S = TT' = T_1T_1'$, then $T_1^{-1}TT'(T_1')^{-1} = I_p$ so $T_1^{-1}T$ is an orthogonal matrix. But $T_1^{-1}T \in G_T^+$ and the only matrix that is both orthogonal and in G_T^+ is I_p. Hence, $T_1^{-1}T = I_p$ and uniqueness follows. □

Let $\mathbb{S}_p^+$ denote the set of $p \times p$ positive definite matrices. Proposition 5.4 shows that the function $F: G_T^+ \to \mathbb{S}_p^+$ defined by $F(T) = TT'$ is both one-to-one and onto. Of course, the existence of $F^{-1}: \mathbb{S}_p^+ \to G_T^+$ is also part of the content of Proposition 5.4. For $T_1 \in G_T^+$, the uniqueness part of Proposition 5.4 yields $F^{-1}(T_1ST_1') = T_1F^{-1}(S)$. This relationship is used later in this chapter. It is clear that the above result holds for G_T^+ replaced by G_U^+. In other words, every $S \in \mathbb{S}_p^+$ has a unique decomposition $S = UU'$ for $U \in G_U^+$.

Proposition 5.5. Suppose $A \in \mathcal{L}_{p,n}$ where $p \leqslant n$ and A has rank p. Then $A = \Psi S$ where $\Psi \in \mathfrak{F}_{p,n}$ and S is positive definite. Furthermore, Ψ and S are unique.

Proof. Since A has rank p, $A'A$ has rank p and is positive definite. Let S be the positive definite square root of $A'A$, so $A'A = SS$. From Proposition 1.31, there exists a linear isometry $\Psi \in \mathfrak{F}_{p,n}$ such that $A = \Psi S$. To establish the uniqueness of Ψ and S, suppose that $A = \Psi S = \Psi_1 S_1$ where $\Psi, \Psi_1 \in \mathfrak{F}_{p,n}$, and S and S_1 are both positive definite. Then $\mathfrak{R}(A) = \mathfrak{R}(\Psi) = \mathfrak{R}(\Psi_1)$. As in the proof of Proposition 5.2, this implies that $\Psi'\Psi_1\Psi_1'\Psi = I_p$ since $\Psi_1\Psi_1'$ is the orthogonal projection onto $\mathfrak{R}(\Psi_1) = \mathfrak{R}(\Psi)$. Therefore, $SS_1^{-1} = \Psi'\Psi_1$ is a $p \times p$ orthogonal matrix so the eigenvalues of SS_1^{-1} are all on the unit circle in the complex plane. But the eigenvalues of SS_1^{-1} are the same as the eigenvalues of $S^{1/2}S_1^{-1}S^{1/2}$ (see Proposition 1.39) where $S^{1/2}$ is the positive definite square root of S. Since $S^{1/2}S_1^{-1}S^{1/2}$ is positive definite, the eigenvalues of $S^{1/2}S_1^{-1}S^{1/2}$ are all positive. Therefore, the eigenvalues of $S^{1/2}S_1^{-1}S^{1/2}$ must all be equal to one, as this is the only point of intersection of $(0, \infty)$ with the unit circle in the complex plane. Since the only $p \times p$ matrix with all eigenvalues equal to one is the identity, $S^{1/2}S_1^{-1}S^{1/2} = I_p$ so $S = S_1$. Since S is nonsingular, $\Psi = \Psi_1$. □

The factorizations established this far were concerned with writing one matrix as the product of two other matrices with special properties. The results below are concerned with factorizations for two or more matrices. Statistical applications of these factorizations occur in later chapters.

Proposition 5.6. Suppose A is a $p \times p$ positive definite matrix and B is a $p \times p$ symmetric matrix. There exists a nonsingular $p \times p$ matrix C and a

$p \times p$ diagonal matrix D such that $A = CC'$ and $B = CDC'$. The diagonal elements of D are the eigenvalues of $A^{-1}B$.

Proof. Let $A^{1/2}$ be the positive definite square root of A and $A^{-1/2} = (A^{1/2})^{-1}$. By the spectral theorem for matrices, there exists a $p \times p$ orthogonal matrix Γ such that $\Gamma' A^{-1/2} B A^{-1/2} \Gamma \equiv D$ is diagonal (see Proposition 1.45), and the eigenvalues of $A^{-1/2}BA^{-1/2}$ are the diagonal elements of D. Let $C = A^{1/2}\Gamma$. Then $CC' = A^{1/2}\Gamma\Gamma' A^{1/2} = A$ and $CDC' = B$. Since the eigenvalues of $A^{-1/2}BA^{-1/2}$ are the same as the eigenvalues of $A^{-1}B$, the proof is complete. □

Proposition 5.7. Suppose S is a $p \times p$ positive definite matrix and partition S as

$$S = \begin{pmatrix} S_{11} & S_{12} \\ S'_{12} & S_{22} \end{pmatrix}$$

where S_{11} is $p_1 \times p_1$ and S_{22} is $p_2 \times p_2$ with $p_1 \leqslant p_2$. Then there exist nonsingular matrices A_{ii} of dimension $p_i \times p_i$, $i = 1, 2$, such that $A_{ii}S_{ii}A'_{ii} = I_{p_i}$, $i = 1, 2$, and $A_{11}S_{12}A'_{22} = (D0)$ where D is a $p_1 \times p_1$ diagonal matrix and 0 is a $p_1 \times (p_2 - p_1)$ matrix of zeroes. The diagonal elements of D^2 are the eigenvalues of $S_{11}^{-1}S_{12}S_{22}^{-1}S_{21}$ where $S_{21} = S'_{12}$, and these eigenvalues are all in the interval $[0, 1]$.

Proof. Since S is positive definite, S_{11} and S_{22} are positive definite. Let $S_{11}^{1/2}$ and $S_{22}^{1/2}$ be the positive definite square roots of S_{11} and S_{22}. Using Proposition 1.46, write the matrix $S_{11}^{-1/2}S_{12}S_{22}^{-1/2}$ in the form

$$S_{11}^{-1/2}S_{12}S_{22}^{-1/2} = \Gamma D \Psi$$

where Γ is a $p_1 \times p_1$ orthogonal matrix, D is a $p_1 \times p_1$ diagonal matrix, and Ψ is a $p_1 \times p_2$ linear isometry. The p_1 rows of Ψ form an orthonormal set in R^{p_2} and $p_2 - p_1$ orthonormal vectors can be adjoined to Ψ to obtain a $p_2 \times p_2$ orthogonal matrix Ψ_1 whose first p_1 rows are the rows of Ψ. It is clear that

$$D\Psi = (D0)\Psi_1$$

where 0 is a $p_1 \times (p_2 - p_1)$ matrix of zeroes. Set $A_{11} = \Psi' S_{11}^{-1/2}$ and $A_{22} = \Psi_1 S_{22}^{-1/2}$ so $A_{ii}S_{ii}A'_{ii} = I_{p_i}$ for $i = 1, 2$. Obviously, $A_{11}S_{12}A'_{22} = (D0)$. Since $S_{11}^{-1/2}S_{12}S_{22}^{-1/2} = \Gamma D \Psi$,

$$S_{11}^{-1/2}S_{12}S_{22}^{-1}S_{21}S_{11}^{-1/2} = \Gamma D^2 \Gamma'$$

so the eigenvalues of $S_{11}^{-1/2}S_{12}S_{22}^{-1}S_{21}S_{11}^{-1/2}$ are the diagonal elements of D^2. Since the eigenvalues of $S_{11}^{-1/2}S_{12}S_{22}^{-1}S_{21}S_{11}^{-1/2}$ are the same as the eigenvalues of $S_{11}^{-1}S_{12}S_{22}^{-1}S_{21}$, it remains to show that these eigenvalues are in [0, 1]. By Proposition 5.3, $S_{11} - S_{12}S_{22}^{-1}S_{21}$ is positive definite so $I_{p_1} - S_{11}^{-1/2}S_{12}S_{22}^{-1}S_{21}S_{11}^{-1/2}$ is positive definite. Thus for $x \in R^{p_1}$,

$$0 \leqslant x'S_{11}^{-1/2}S_{12}S_{22}^{-1}S_{21}S_{11}^{-1/2}x \leqslant x'x,$$

which implies that (see Proposition 1.44) the eigenvalues of $S_{11}^{-1/2}S_{12}S_{22}^{-1}S_{21}S_{11}^{-1/2}$ are in the interval [0, 1]. □

It is shown later that the eigenvalues of $S_{11}^{-1}S_{12}S_{22}^{-1}S_{21}$ are related to the angles between two subspaces of R^p. However, it is also shown that these eigenvalues have a direct statistical interpretation in terms of correlation coefficients, and this establishes the connection between canonical correlation coefficients and angles between subspaces. The final decomposition result in this section provides a useful result for evaluating integrals over the space of $p \times p$ positive definite matrices.

Proposition 5.8. Let $\mathbb{S}_p^+$ denote the space of $p \times p$ positive definite matrices. For $S \in \mathbb{S}_p^+$, partition S as

$$S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}$$

where S_{ii} is $p_i \times p_i$, $i = 1, 2$, S_{12} is $p_1 \times p_2$, and $S_{21} = S_{12}'$. The function f defined on $\mathbb{S}_p^+$ to $\mathbb{S}_{p_1}^+ \times \mathbb{S}_{p_2}^+ \times \mathcal{L}_{p_2, p_1}$ by

$$f(S) = \left(S_{11} - S_{12}S_{22}^{-1}S_{21}, S_{22}, S_{12}S_{22}^{-1}\right)$$

is a one-to-one onto function. The function h on $\mathbb{S}_{p_1}^+ \times \mathbb{S}_{p_2}^+ \times \mathcal{L}_{p_2, p_1}$ to $\mathbb{S}_p^+$ given by

$$h(A_{11}, A_{22}, A_{12}) = \begin{pmatrix} A_{11} + A_{12}A_{22}A_{12}' & A_{12}A_{22} \\ A_{22}A_{12}' & A_{22} \end{pmatrix}$$

is the inverse of f.

Proof. It is routine to verify that $f \circ h$ is the identity function on $\mathbb{S}_{p_1}^+ \times \mathbb{S}_{p_2}^+ \times \mathcal{L}_{p_2, p_1}$ and $h \circ f$ is the identity function on $\mathbb{S}_p^+$. This implies the assertions of the proposition. □

5.2. JACOBIANS

Jacobians provide the basic technical tool for describing how multivariate integrals over open subsets of R^n transform under a change of variable. To describe the situation more precisely, let B_0 and B_1 be fixed open subsets of R^n and let g be a one-to-one onto mapping from B_0 to B_1. Recall that the differential of g, assuming the differential exists, is a function D_g defined on B_0 that takes values in $\mathcal{L}_{n,n}$ and satisfies

$$\lim_{\delta \to 0} \frac{\|g(x+\delta) - g(x) - D_g(x)\delta\|}{\|\delta\|} = 0$$

for each $x \in B_0$. Here δ is a vector in R^n chosen small enough so that $x + \delta \in B_0$. Also, $D_g(x)\delta$ is the matrix $D_g(x)$ applied to the vector δ, and $\|\cdot\|$ denotes the standard norm on R^n. Let $g_1, \dots, g_n$ denote the coordinate functions of the vector valued function g. It is well known that the matrix $D_g(x)$ is given by

$$D_g(x) = \left\{ \frac{\partial g_i}{\partial x_j}(x) \right\}, \qquad x \in B_0.$$

In other words, the (i, j) element of the matrix $D_g(x)$ is the partial derivative of g_i with respect to x_j evaluated at $x \in B_0$. The Jacobian of g is defined by

$$J_g(x) = |\det D_g(x)|, \qquad x \in B_0$$

so the Jacobian is the absolute value of the determinant of D_g. A formal statement of the change of variables theorem goes as follows. Consider any real valued Borel measurable function f defined on the open set B_1 such that

$$\int_{B_1} |f(y)|\, dy < +\infty$$

where dy means Lebesgue measure. Introduce the change of variables $y = g(x)$, $x \in B_0$ in the integral $\int_{B_1} f(y)\, dy$. Then the change of variables theorem asserts that

$$\int_{B_1} f(y)\, dy = \int_{B_0} f(g(x)) J_g(x)\, dx. \tag{5.1}$$

An alternative way to express (5.1) is by the formal expression

$$d(g(x)) = J_g(x)\,dx, \qquad x \in B_0. \tag{5.2}$$

To give a precise meaning to (5.2), proceed as follows. For each Borel measurable function h defined on B_0 such that $\int_{B_0} |h(x)| J_g(x)\,dx < +\infty$, define

$$I_1(h) \equiv \int_{B_0} h(x) J_g(x)\,dx,$$

and define

$$I_2(h) \equiv \int_{B_0} h(x)\,d(g(x)) \equiv \int_{g(B_0)} h(g^{-1}(x))\,dx.$$

Then (5.2) means that $I_1(h) = I_2(h)$ for all h such that $I_1(|h|) < +\infty$. To show that (5.1) and the equality of I_1 and I_2 are equivalent, simply set $f = h \circ g^{-1}$ so $f \circ g = h$. Thus $I_1(h) = I_2(h)$ iff

$$\int_{B_0} f(g(x)) J_g(x)\,dx = \int_{B_1} f(x)\,dx$$

since $B_1 = g(B_0)$.

One property of Jacobians that is often useful in simplifying computations is the following. Let B_0, B_1, and B_2 be open subsets of R^n, suppose g_1 is a one-to-one onto map from B_0 to B_1, and suppose D_{g_1} exists. Also, suppose g_2 is a one-to-one onto map from B_1 to B_2 and assume that D_{g_2} exists. Then, $g_2 \circ g_1$ is a one-to-one onto map from B_0 to B_2 and it is not difficult to show that

$$D_{g_2 \circ g_1}(x) = D_{g_2}(g_1(x)) D_{g_1}(x), \qquad x \in B_0.$$

Of course, the right-hand side of this equality means the matrix product of $D_{g_2}(g_1(x))$ and $D_{g_1}(x)$. From this equality, it follows that

$$J_{g_2 \circ g_1}(x) = J_{g_2}(g_1(x)) J_{g_1}(x), \qquad x \in B_0.$$

In particular, if $B_2 = B_0$ and $g_2 = g_1^{-1}$, then $g_2 \circ g_1 = g_1^{-1} \circ g_1$ is the identity function on B_0 so its Jacobian is one. Thus

$$1 = J_{g_2 \circ g_1}(x) = J_{g_2}(g_1(x)) J_{g_1}(x), \qquad x \in B_0$$

and

$$J_{g_1^{-1}}(y) = \frac{1}{J_{g_1}(g_1^{-1}(y))}, \qquad y \in B_1.$$

We now turn to the problem of explicitly computing some Jacobians that are needed later. The first few results present Jacobians for linear transformations.

Proposition 5.9. Let A be an $n \times n$ nonsingular matrix and define g on R^n to R^n by $g(x) = A(x)$. Then $J_g(x) = |\det(A)|$ for $x \in R^n$.

Proof. We must compute the differential matrix of g. It is clear that the ith coordinate function of f is g_i where

$$g_i(x) = \sum_{k=1}^{n} a_{ik} x_k.$$

Here $A = \{a_{ij}\}$ and x has coordinates $x_1, \ldots, x_n$. Thus

$$\frac{\partial g_i}{\partial x_j}(x) = a_{ij}$$

so $D_g(x) = \{a_{ij}\}$. Thus $J_g(x) = |\det(A)|$. □

Proposition 5.10. Let A be an $n \times n$ nonsingular matrix and let B be a $p \times p$ nonsingular matrix. Define g on the np-dimensional coordinate space $\mathcal{L}_{p,n}$ to $\mathcal{L}_{p,n}$ by

$$g(X) = AXB' = (A \otimes B)X.$$

Then $J_g(X) = |\det A|^p |\det B|^n$.

Proof. First note that $A \otimes B = (I_n \otimes B)(A \otimes I_p)$. Setting $g_1(X) = (A \otimes I_p)X$ and $g_2(X) = (I_n \otimes B)X$, it is sufficient to verify that

$$J_{g_1}(X) = |\det A|^p$$

and

$$J_{g_2}(X) = |\det B|^n.$$

Let $x_1, \ldots, x_p$ be the columns of the $n \times p$ matrix X so $x_i \in R^n$. Form the np-dimensional vector

$$[X] = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}.$$

Since $(A \otimes I_p)X$ has columns $Ax_1, \ldots, Ax_n$, the matrix of $A \otimes I_p$ as a linear transformation on $[X]$ is

$$\begin{pmatrix} A & & & \\ & A & & \\ & & \ddots & \\ & & & A \end{pmatrix} : (np) \times (np)$$

where the elements not indicated are zero. Clearly, the determinant of this matrix is $(\det A)^p$ since A occurs p times on the diagonal. Since the determinant of a linear transformation is independent of a matrix representation, we have that

$$\det(A \otimes I_p) = (\det A)^p.$$

Applying Proposition 5.9, it follows that

$$J_{g_1}(X) = |\det A|^p.$$

Using the rows instead of the columns, we find that

$$\det(I_n \otimes B) = (\det B)^n,$$

so

$$J_{g_2}(X) = |\det B|^n. \qquad \square$$

Proposition 5.11. Let A be a $p \times p$ nonsingular matrix and define the function g on the linear space $\mathbb{S}_p$ of $p \times p$ real symmetric matrices by

$$g(S) = ASA' = (A \otimes A)S.$$

Then $J_g(S) = |\det A|^{p+1}$.

Proof. The result of the previous proposition shows that $\det(A \otimes A) = (\det A)^{2p}$ when $A \otimes A$ is regarded as a linear transformation on $\mathcal{L}_{p,p}$. However, this result is not applicable to the current case since we are considering the restriction of $A \otimes A$ to the subspace $\mathcal{S}_p$ of $\mathcal{L}_{p,p}$.

To establish the present result, write $A = \Gamma_1 D \Gamma_2$ where Γ_1 and Γ_2 are $p \times p$ orthogonal matrices and D is a diagonal matrix with positive diagonal elements (see Proposition 1.47). Then,

$$ASA' = (A \otimes A)S = (\Gamma_1 \otimes \Gamma_1)(D \otimes D)(\Gamma_2 \otimes \Gamma_2)S$$

so the linear transformation $A \otimes A$ has been decomposed into the composition of three linear transformations, two of which are determined by orthogonal matrices.

We now claim that if Γ is a $p \times p$ orthogonal matrix and g_1 is defined on $\mathcal{S}_p$ by

$$g_1(S) = \Gamma S \Gamma' = (\Gamma \otimes \Gamma)S,$$

then $J_{g_1} = 1$. To see this, let $\langle \cdot, \cdot \rangle$ be the natural inner product on $\mathcal{L}_{p,p}$ restricted to $\mathcal{S}_p$, that is, let

$$\langle S_1, S_2 \rangle = \operatorname{tr} S_1 S_2.$$

Then

$$\begin{aligned}\langle (\Gamma \otimes \Gamma)S_1, (\Gamma \otimes \Gamma)S_2 \rangle &= \operatorname{tr} \Gamma S_1 \Gamma' \Gamma S_2 \Gamma' = \operatorname{tr} \Gamma S_1 S_2 \Gamma' \\ &= \operatorname{tr} \Gamma' \Gamma S_1 S_2 = \operatorname{tr} S_1 S_2 = \langle S_1, S_2 \rangle.\end{aligned}$$

Therefore, $\Gamma \otimes \Gamma$ is an orthogonal transformation on the inner product space $(\mathcal{S}_p, \langle \cdot, \cdot \rangle)$, so the determinant of this linear transformation on $\mathcal{S}_p$ is ± 1. Thus g_1 is a linear transformation that is also orthogonal so $J_{g_1} = 1$ and the claim is established.

The next claim is that if D is a $p \times p$ diagonal matrix with positive diagonal elements and g_2 is defined on $\mathcal{S}_p$ by

$$g_2(S) = DSD,$$

then $J_{g_2} = (\det D)^{p+1}$. In the $[p(p+1)/2]$-dimensional space $\mathcal{S}_p$, let s_{ij}, $1 \leqslant j \leqslant i \leqslant p$, denote the coordinates of S. Then it is routine to show that the (i, j) coordinate function of g_2 is $g_{2,ij}(S) = \lambda_i \lambda_j s_{ij}$ where $\lambda_1, \ldots, \lambda_p$ are the diagonal elements of D. Thus the matrix of the linear transformation g_2 is a $[p(p+1)/2] \times [p(p+1)/2]$ diagonal matrix with diagonal entries

$\lambda_i\lambda_j$ for $1 \leqslant j \leqslant i \leqslant p$. Hence the determinant of this matrix is the product of the $\lambda_i\lambda_j$ for $1 \leqslant j \leqslant i \leqslant p$. A bit of calculation shows this determinant is $(\Pi\lambda_i)^{p+1}$. Since $\det D = \Pi\lambda_i$, the second claim is established.

To complete the proof, note that

$$g(S) = ASA' = (\Gamma_1 \otimes \Gamma_1)(D \otimes D)(\Gamma_2 \otimes \Gamma_2)S = h_1(h_2(h_3(S)))$$

where $h_1(S) = (\Gamma_1 \otimes \Gamma_1)S$, $h_2(S) = (D \otimes D)S$, and $h_3(S) = (\Gamma_2 \otimes \Gamma_2)S$. A direct argument shows that

$$\begin{aligned} J_{h_1 \circ h_2 \circ h_3}(S) &= J_{h_1 \circ h_2}(h_3(S))J_{h_3}(S) \\ &= J_{h_1}(h_2(h_3(S)))J_{h_2}(h_3(S))J_{h_3}(S). \end{aligned}$$

But $J_{h_1} = 1 = J_{h_3}$ and $J_{h_2} = (\det D)^{p+1}$. Since $A = \Gamma_1 D \Gamma_2$, $|\det A| = \det D$, which entails $J_g = |\det A|^{p+1}$. □

Proposition 5.12. Let M be the linear space of $p \times p$ skew-symmetric matrices and define g on M to M by

$$g(S) = ASA'$$

where A is a $p \times p$ nonsingular matrix. Then $J_g(S) = |\det A|^{p-1}$.

Proof. The proof is similar to that of Proposition 5.11 and is left to the reader. □

Proposition 5.13. Let G_T^+ be the set of $p \times p$ lower triangular matrices with positive diagonal elements and let A be a fixed element of G_T^+. The function g defined on G_T^+ to G_T^+ by

$$g(T) = AT, \qquad T \in G_T^+$$

has a Jacobian given by $J_g(T) = \prod_1^p a_{ii}^i$ where $a_{11}, \ldots, a_{pp}$ are the diagonal elements of A.

Proof. The set G_T^+ is an open subset of $[\frac{1}{2}p(p+1)]$-dimensional coordinate space and g is a one-to-one onto function by Proposition 5.1. For $T \in G_T^+$, form the vector $[T]$ with coordinates $t_{11}, t_{21}, t_{22}, t_{31}, \ldots, t_{pp}$ and write the coordinate functions of g in the same order. Then the matrix of partial derivatives is lower triangular with diagonal elements

$a_{11}, a_{22}, a_{22}, a_{33}, \ldots, a_{pp}$ where a_{ii} occurs i times on the diagonal. Thus the determinant of this matrix of partial derivatives is $\prod_1^p a_{ii}^i$ so $J_g = \prod_1^p a_{ii}^i$. □

Proposition 5.14. In the notation of Proposition 5.13, define g on G_T^+ to G_T^+ by

$$g(T) = TB, \qquad T \in G_T^+$$

where B is a fixed element of G_T^+. Then $J_g(T) = \prod_1^p b_{ii}^{p-i+1}$ where $b_{11}, \ldots, b_{pp}$ are the diagonal elements of B.

Proof. The proof is similar to that of Proposition 5.13 and is omitted. □

Proposition 5.15. Let G_U^+ be the set of all $p \times p$ upper triangular matrices with positive diagonal elements. For fixed elements A and B of G_U^+, define g by

$$g(U) = AUB, \qquad U \in G_U^+.$$

Then,

$$J_g(U) = \prod_1^p a_{ii}^{p-i+1} \prod_1^p b_{ii}^i$$

where $a_{11}, \ldots, a_{pp}$ and $b_{11}, \ldots, b_{pp}$ are diagonal elements of A and B.

Proof. The proof is similar to that given for lower triangular matrices and is left to the reader. □

Thus far, only Jacobians of linear transformations have been computed explicitly, and, of course, these Jacobians have been constant functions. In the next proposition, the Jacobian of the nonlinear transformation described in Proposition 5.8 is computed.

Proposition 5.16. Let p_1 and p_2 be positive integers and set $p = p_1 + p_2$. Using the notation of Proposition 5.8, define h on $\mathbb{S}_{p_1}^+ \times \mathbb{S}_{p_2}^+ \times \mathfrak{L}_{p_2, p_1}$ to $\mathbb{S}_p^+$ by

$$h(A_{11}, A_{22}, A_{12}) = \begin{pmatrix} A_{11} + A_{12}A_{22}A_{12}' & A_{12}A_{22} \\ A_{22}A_{12}' & A_{22} \end{pmatrix}.$$

Then $J_h(A_{11}, A_{22}, A_{12}) = (\det A_{22})^{p_1}$.

Proof. For notational convenience, set $S = h(A_{11}, A_{22}, A_{12})$ and partition S as

$$S = \begin{pmatrix} S_{11} & S_{12} \\ S_{12} & S_{22} \end{pmatrix}$$

where S_{ij} is $p_i \times p_j$, $i, j = 1, 2$. The partial derivatives of the elements of S, as functions of the elements of A_{11}, A_{12} and A_{22}, need to be computed. Since $S_{11} = A_{11} + A_{12}A_{22}A'_{12}$, the matrix of partial derivatives of the $p_1(p_1+1)/2$ elements of S_{11} with respect to the $p_1(p_1+1)/2$ elements of A_{11} is just the $[p_1(p_1+1)/2]$-dimensional identity matrix. Since $S_{12} = A_{12}A_{22}$, the matrix of partial derivatives of the $p_1 p_2$ elements of S_{12} with respect to the elements of A_{11} is the $p_1 p_2 \times p_1 p_2$ zero matrix. Also, since $S_{22} = A_{22}$, the partial derivative of elements of S_{22} with respect to the elements of A_{11} or A_{12} are all zero and the matrix of partial derivatives of the $p_2(p_2+1)/2$ elements of S_{22} with respect to the $p_2(p_2+1)/2$ elements of A_{22} is the identity matrix. Thus the matrix of partial derivatives has the form

$$\begin{array}{cc} & \begin{array}{ccc} A_{11} & A_{12} & A_{22} \end{array} \\ \begin{array}{c} S_{11} \\ S_{12} \\ S_{22} \end{array} & \begin{pmatrix} I_1 & - & - \\ 0 & B & - \\ 0 & 0 & I_2 \end{pmatrix} \end{array}$$

so the determinant of this matrix is just the determinant of the $p_1 p_2 \times p_1 p_2$ matrix B, which must be found. However, B is the matrix of partial derivatives of the elements of S_{12} with respect to the elements of A_{12} where $S_{12} = A_{12}A_{22}$. Hence the determinant of B is just the Jacobian of the transformation $g(A_{12}) = A_{12}A_{22}$ with A_{22} fixed. This Jacobian is $(\det A_{22})^{p_1}$ by Proposition 5.10. $\square$

As an application of Proposition 5.16, a special integral over the space $\mathbb{S}_p^+$ is now evaluated.

◆ **Example 5.1.** Let dS denote Lebesgue measure on the set $\mathbb{S}_p^+$. The integral below arises in our discussion of the Wishart distribution. For a positive integer p and a real number $r > p - 1$, let

$$c(r, p) = \int_{\mathbb{S}_p^+} |S|^{(r-p-1)/2} \exp\left[-\tfrac{1}{2}\operatorname{tr} S\right] dS.$$

In this example, the constant $c(r, p)$ is calculated. When $p = 1$,

$\mathbb{S}_p^+ = (0, \infty)$ so for $r > 0$,

$$c(r,1) = \int_0^\infty s^{(r/2)-1}\exp\left[-\frac{s}{2}\right] ds = 2^{r/2}\Gamma\left(\frac{r}{2}\right)$$

where $\Gamma(r/2)$ is the gamma function evaluated at $r/2$. The first claim is that

$$c(r, p+1) = (2\pi)^{p/2} c(r-1, p) c(r, 1),$$

for $r > p$ and $p \geqslant 1$. To verify this claim, consider $S \in \mathbb{S}_{p+1}^+$ and partition S as

$$S = \begin{pmatrix} S_{11} & S_{12} \\ S_{12}' & S_{22} \end{pmatrix}$$

where $S_{11} \in \mathbb{S}_p^+$, $S_{22} \in (0, \infty)$, and S_{12} is $p \times 1$. Introduce the change of variables

$$\begin{pmatrix} S_{11} & S_{12} \\ S_{12}' & S_{22} \end{pmatrix} = \begin{pmatrix} A_{11} + A_{12}A_{22}A_{12}' & A_{12}A_{22} \\ A_{22}A_{12}' & A_{22} \end{pmatrix}$$

where $A_{11} \in \mathbb{S}_p^+$, $A_{22} \in (0, \infty)$, and $A_{12} \in R^p$. By Proposition 5.16, the Jacobian of this transformation is A_{22}^p. Since $\det S = \det(S_{11} - S_{12}S_{22}^{-1}S_{12}')\det S_{22} = (\det A_{11})A_{22}$, we have

$$c(r, p+1) = \int_{\mathbb{S}_{p+1}^+} |S|^{(r-p-2)/2}\exp\left[-\tfrac{1}{2}\operatorname{tr} S\right] dS$$

$$= \int_0^\infty \int_{R^p}\int_{\mathbb{S}_p^+} |A_{11}|^{(r-p-2)/2} A_{22}^{(r-p-2)/2}$$

$$\times \exp\left[-\tfrac{1}{2}\operatorname{tr} A_{11} - \tfrac{1}{2}A_{22}A_{12}'A_{12} - \tfrac{1}{2}A_{22}\right]$$

$$\times A_{22}^p \, dA_{11}\, dA_{12}\, dA_{22}.$$

Integrating with respect to A_{12} yields

$$\int_{R^p} \exp\left[-\tfrac{1}{2}A_{22}A_{12}'A_{12}\right] dA_{12} = (2\pi)^{p/2} A_{22}^{-p/2}.$$

Substituting this into the second integral expression for $c(r, p+1)$

and then integrating on A_{22} shows that

$$c(r, p+1) = (2\pi)^{p/2} c(r,1) \int_{\mathbb{S}_p^+} |A_{11}|^{(r-p-2)/2} \exp\left[-\tfrac{1}{2} \operatorname{tr} A_{11}\right] dA_{11}$$

$$= (2\pi)^{p/2} c(r,1) c(r-1, p).$$

This establishes the first claim. Now, it is an easy matter to solve for $c(r, p)$. A bit of manipulation shows that with

$$c(r, p) = \pi^{p(p-1)/4} 2^{rp/2} \prod_{j=1}^{p} \Gamma\left(\frac{r-j+1}{2}\right),$$

for $p = 1, 2, \ldots,$ and $r > p - 1$, the equation

$$c(r, p+1) = (2\pi)^{p/2} c(r,1) c(r-1, p)$$

is satisfied. Further,

$$c(r,1) = 2^{r/2} \Gamma\left(\frac{r}{2}\right).$$

Uniqueness of the solution to the above equation is clear. In summary,

$$\int_{\mathbb{S}_p^+} |S|^{(r-p-1)/2} \exp\left[-\tfrac{1}{2} \operatorname{tr} S\right] dS = \pi^{p(p-1)/4} 2^{rp/2} \prod_{j=1}^{p} \Gamma\left(\frac{r-j+1}{2}\right)$$

and this is valid for $p = 1, 2, \ldots$ and $r > p - 1$. The restriction that r be greater than $p - 1$ is necessary so that $\Gamma[(r-p+1)/2]$ be well defined. It is not difficult to show that the above integral is $+\infty$ if $r \leqslant p - 1$. Now, set $\omega(r, p) = 1/c(r, p)$ so

$$f(S) \equiv \omega(r, p) |S|^{(r-p-1)/2} \exp\left[-\tfrac{1}{2} \operatorname{tr} S\right]$$

is a density function on $\mathbb{S}_p^+$. When r is an integer, $r \geqslant p$, f turns out to be the density of the Wishart distribution. ◆

Proposition 5.4 shows that there is a one-to-one correspondence between elements of $\mathbb{S}_p^+$ and elements of G_T^+. More precisely, the function g defined on G_T^+ by

$$g(T) = TT', \qquad T \in G_T^+$$

is one-to-one and onto $\mathbb{S}_p^+$. It is clear that g has a differential since each coordinate function of g is a polynomial in the elements of T. One way to find the Jacobian of g is to simply compute the matrix of partial derivatives and then find its determinant. As motivation for some considerations in the next chapter, a different derivation of the Jacobian of g is given here. The first observation is as follows.

Proposition 5.17. Let dS denote Lebesgue measure on $\mathbb{S}_p^+$ and consider the measure μ on $\mathbb{S}_p^+$ given by $\mu(dS) = dS/|S|^{(p+1)/2}$. For each Borel measurable function f on $\mathbb{S}_p^+$, which is integrable with respect to μ, and for each nonsingular matrix A,

$$\int_{\mathbb{S}_p^+} f(S)\mu(dS) = \int_{\mathbb{S}_p^+} f(ASA')\mu(dS).$$

Proof. Set $B = ASA'$. By Proposition 5.11, the Jacobian of this transformation on $\mathbb{S}_p^+$ to $\mathbb{S}_p^+$ is $|\det A|^{p+1}$. Thus

$$\begin{aligned}\int_{\mathbb{S}_p^+} f(ASA')\mu(dS) &= \int_{\mathbb{S}_p^+} f(ASA')\frac{dS}{|S|^{(p+1)/2}}\\ &= \int_{\mathbb{S}_p^+} \frac{f(ASA')|\det A|^{p+1}}{|ASA'|^{(p+1)/2}}\,dS\\ &= \int_{\mathbb{S}_p^+} \frac{f(B)}{|B|^{(p+1)/2}}\,dB = \int_{\mathbb{S}_p^+} f(S)\mu(dS). \qquad \square\end{aligned}$$

The result of Proposition 5.17 is often paraphrased by saying that the measure μ is invariant under each of the transformations g_A defined on $\mathbb{S}_p^+$ by $g_A(S) = ASA'$. The following calculation gives a heuristic proof of this result:

$$\begin{aligned}\mu(dg_A(S)) &= \frac{d(g_A(S))}{|ASA'|^{(p+1)/2}} = \frac{J_{g_A}(S)\,dS}{|ASA'|^{(p+1)/2}}\\ &= \frac{|\det A|^{p+1}}{|AA'|^{(p+1)/2}}\frac{dS}{|S|^{(p+1)/2}} = \frac{dS}{|S|^{(p+1)/2}} = \mu(dS).\end{aligned}$$

In fact, a similar calculation suggests that μ is the only invariant measure in $\mathbb{S}_p^+$ (up to multiplication of μ by a positive constant). Consider a measure ν

of the form $\nu(dS) = h(S)\,dS$ where h is a positive Borel measurable function and dS is Lebesgue measure. In order that ν be invariant, we must have

$$\begin{aligned} h(S)\,dS = \nu(dS) = \nu(dg_A(S)) &= h(g_A(S))d(g_A(S)) \\ &= h(g_A(S))|\det A|^{p+1}\,dS \end{aligned}$$

so h should satisfy the equation

$$h(S) = h(ASA')|AA'|^{(p+1)/2},$$

since $g_A(S) = ASA'$ and $|\det A|^{p+1} = |AA'|^{(p+1)/2}$. Set $S = I_p$, $B = AA'$, and $c = h(I_p)$. Then

$$h(B) = \frac{c}{|B|^{(p+1)/2}}, \qquad B \in \mathbb{S}_p^+$$

so

$$\nu(dS) = c\mu(dS)$$

where c is a positive constant. Making this argument rigorous is one of the topics treated in the next chapter.

The calculation of the Jacobian of g on G_T^+ to $\mathbb{S}_p^+$ is next.

Proposition 5.18. For $g(T) = TT'$, $T \in G_T^+$,

$$J_g(T) = 2^p \prod_{i=1}^{p} t_{ii}^{p-i+1}$$

where $t_{11}, \ldots, t_{pp}$ are the diagonal elements of T.

Proof. The Jacobian J_g is the unique continuous function defined on G_T^+ that satisfies the equation

$$\int_{\mathbb{S}_p^+} f(S) \frac{dS}{|S|^{(p+1)/2}} = \int_{G_T^+} \frac{f(g(T))J_g(T)}{|g(T)|^{(p+1)/2}}\,dT$$

for all Borel measurable functions f for which the integral over $\mathbb{S}_p^+$ exists. But the left-hand side of this equation is invariant under the replacement of $f(S)$ by $f(ASA')$ for any nonsingular $p \times p$ matrix. Thus the right-hand side must have the same property. In particular, for $A \in G_T^+$, we have

$$\int_{G_T^+} \frac{f(TT')}{|TT'|^{(p+1)/2}} J_g(T)\, dT = \int_{G_T^+} \frac{f(ATT'A')}{|TT'|^{(p+1)/2}} J_g(T)\, dT.$$

In this second integral, we make the change of variable $T = A^{-1}B$ for $A \in G_T^+$ fixed and $B \in G_T^+$. By Proposition 5.12, the Jacobian of this transformation is $1/\prod_1^p a_{ii}^i$ where $a_{11}, \ldots, a_{pp}$ are the diagonal elements of A. Thus

$$\int_{G_T^+} \frac{f(TT')}{|TT'|^{(p+1)/2}} J_g(T)\, dT = \int_{G_T^+} \frac{f(BB')}{|BB'|^{(p+1)/2}} \frac{J_g(A^{-1}B)}{|A^{-1}|^{p+1}} \frac{1}{\prod_1^p a_{ii}^i} dB.$$

Since this must hold for all Borel measurable f and since J_g is a continuous function, it follows that for all $T \in G_T^+$ and $A \in G_T^+$,

$$J_g(T) = J_g(A^{-1}T) \frac{|A|^{p+1}}{\prod_1^p a_{ii}^i}.$$

Setting $A = T$ and noting that $|T| = \prod_1^p t_{ii}$, we have

$$J_g(T) = J_g(I_p) \prod_1^p t_{ii}^{p-i+1}.$$

Thus $J_g(T)$ is a constant k times $\prod_1^p t_{ii}^{p-i+1}$. Hence

$$\int_{\mathbb{S}_p^+} f(S) \frac{dS}{|S|^{(p+1)/2}} = \int_{G_T^+} k \frac{f(TT')}{|T|^{p+1}} \prod_{i=1}^p t_{ii}^{p-i+1}\, dT = \int_{G_T^+} kf(TT') \prod_{i=1}^p t_{ii}^{-i}\, dT.$$

To evaluate the constant k, pick

$$f(S) = |S|^{r/2} \exp[-\tfrac{1}{2} \operatorname{tr} S], \qquad r > p - 1.$$

But

$$\int_{\mathbb{S}_p^+} |S|^{r/2} \exp[-\tfrac{1}{2} \operatorname{tr} S] \frac{dS}{|S|^{(p+1)/2}} = c(r, p)$$

where $c(r, p)$ is defined in Example 5.1. However,

$$k\int_{G_T^+}|TT'|^{r/2}\exp[-\tfrac{1}{2}\operatorname{tr} TT']\prod_1^p t_{ii}^{-i}\,dT$$

$$= k\int_{G_T^+}\prod_1^p t_{ii}^{r-i}\exp\left[-\tfrac{1}{2}\sum_{j\leqslant i} t_{ij}^2\right]dT = k2^{-p}c(r, p)$$

so $k = 2^p$. The evaluation of the last integral is carried out by noting that t_{ii} ranges from 0 to ∞ and t_{ij} for $j < i$ ranges from $-\infty$ to ∞. Thus the integral is a product of $p(p+1)/2$ integrals on R, each of which is easy to evaluate. □

A by-product of this proof is that

$$h(T) = \frac{\prod_1^p t_{ii}^{r-i}}{2^p c(r, p)}\exp\left[-\tfrac{1}{2}\sum_{j\leqslant i} t_{ij}^2\right]$$

is a density function on G_T^+. Since the density h factors into a product of densities, the elements of T, t_{ij} for $j \leqslant i$, are independent. Clearly,

$$\mathcal{L}(t_{ij}) = N(0, 1) \quad \text{for } j < i$$

and

$$\mathcal{L}(t_{ii}^2) = \chi_{n-i+1}^2$$

when r is the integer $n \geqslant p$.

Proposition 5.19. Define g on G_U^+ to $\mathbb{S}_p^+$ by $g(U) = UU'$. Then $J_g(U)$ is given by

$$J_g(U) = 2^p\prod_{i=1}^p u_{ii}^i$$

where $u_{11}, \ldots, u_{pp}$ are the diagonal elements of U.

Proof. The proof is essentially the same as the proof of Proposition 5.18 and is left to the reader. □

The technique used to prove Proposition 5.18 is an important one. Given g on G_T^+ to $\mathbb{S}_p^+$, the idea of the proof was to write down the equation the

Jacobian satisfies, namely,

$$\int_{\mathbb{S}_p^+} \frac{f(S)}{|S|^{(p+1)/2}} dS = \int_{G_T^+} \frac{f(g(T))}{|T|^{p+1}} J_g(T)\, dT$$

for all integrable f. Since this equation must hold for all integrable f, J_g is uniquely defined (up to sets of Lebesgue measure zero) by this equation. It is clear that any property satisfied by the left-hand integral must also be satisfied by the right-hand integral and this was used to characterize J_g. In particular, it was noted that the left-hand integral remained the same if $f(S)$ was replaced by $f(ASA')$ for an nonsingular A. For $A \in G_T^+$, this led to the equation

$$J_g(T) = J_g(A^{-1}T) \frac{|A|^{p+1}}{\prod_1^p a_{ii}^i},$$

which determined J_g. It should be noted that only Jacobians of the linear transformations discussed in Propositions 5.11 and 5.13 were used to determine the Jacobian of the nonlinear transformation g. Arguments similar to this are used throughout Chapter 6 to derive invariant integrals (measures) on matrix groups and spaces that are acted upon by matrix groups.

PROBLEMS

1. Given $A \in \mathcal{L}_{p,n}$ with rank$(A) = p$, show that $A = \Psi T$ where $\Psi \in \mathcal{F}_{p,n}$ and $T \in G_T^+$. Prove that Ψ and T are unique.

2. Define the function F on $\mathbb{S}_p^+$ to G_T^+ as follows. For each $S \in \mathbb{S}_p^+$, $F(S)$ is the unique element in G_T^+ such that $S = F(S)(F(S))'$. Show that $F(TST') = TF(S)$ for $T \in G_T^+$ and $S \in \mathbb{S}_p^+$.

3. Given $S \in \mathbb{S}_p^+$, show there exists a unique $U \in G_U^+$ such that $S = UU'$.

4. For $S \in \mathbb{S}_p^+$, partition S as

$$S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}$$

where S_{ij} is $p_i \times p_j$, $i, j = 1,2$. Assume for definiteness that $p_1 \leqslant p_2$.

Show that S can be written as

$$\begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} \begin{pmatrix} I_{p_1} & (D0) \\ (D0)' & I_{p_2} \end{pmatrix} \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}'$$

where A_i is $p_i \times p_i$ and nonsingular, D is $p_1 \times p_1$ and diagonal with diagonal elements in $[0, 1)$.

5. Let $\mathcal{L}_{p,n}^0$ be those elements in $\mathcal{L}_{p,n}$ that have rank p. Define F on $\mathfrak{F}_{p,n} \times G_U^+$ to $\mathcal{L}_{p,n}^0$ by $F(\Psi, U) = \Psi U$.
   (i) Show that F is one-to-one onto, and describe the inverse of F.
   (ii) For $\Gamma \in \mathcal{O}_n$ and $T \in G_T^+$, define $\Gamma \otimes \mathrm{T}$ on $\mathcal{L}_{p,n}^0$ to $\mathcal{L}_{p,n}^0$ by $(\Gamma \otimes \mathrm{T})A = \Gamma A \mathrm{T}'$. Show that $(\Gamma \otimes T)F(\Psi, U) = F(\Gamma\Psi, UT')$. Also, show that $F^{-1}((\Gamma \otimes T)A) = (\Gamma\Psi, UT')$ where $F^{-1}(A) = (\Psi, U)$.

6. Let B_0 and B_1 be open sets in R^n and fix $x_0 \in B_0$. Suppose g maps B_0 into B_1 and $g(x) = g(x_0) + A(x - x_0) + R(x - x_0)$ where A is an $n \times n$ matrix and $R(\cdot)$ is a function that satisfies

$$\lim_{u \to 0} \frac{\|R(u)\|}{\|u\|} = 0.$$

   Prove that $A = D_g(x_0)$ so $J_g(x_0) = |\det(A)|$.

7. Let V be the linear coordinate space of all $p \times p$ lower triangular real matrices so V is of dimension $p(p+1)/2$. Let $\mathcal{S}_p$ be the linear coordinate space of all $p \times p$ real symmetric matrices so $\mathcal{S}_p$ is also of dimension $p(p+1)/2$.
   (i) Show that G_T^+ is an open subset of V.
   (ii) Define g on G_T^+ to $\mathcal{S}_p$ by $g(T) = TT'$. For fixed $T_0 \in G_T^+$, show that $g(T) = g(T_0) + L(T - T_0) + (T - T_0)(T - T_0)'$ where L is defined on V to $\mathcal{S}_p$ by $L(x) = xT_0' + T_0x'$, $x \in V$. Also show that $R(T - T_0) = (T - T_0)(T - T_0)'$ satisfies

$$\lim_{x \to 0} \frac{\|R(x)\|}{\|x\|} = 0.$$

   (iii) Prove by induction that $\det L = 2^p\prod_1^p t_{ii}^{p-i+1}$ where $t_{11}, \ldots, t_{pp}$ are the diagonal elements of T_0.
   (iv) Using (iii) and Problem 6, show that $J_g(T) = 2^p\prod_1^p t_{ii}^{p-i+1}$. (This is just Proposition 5.18).

8. When S is a positive definite matrix, partition S and S^{-1} as

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

Show that

$$S^{11} = \left(S_{11} - S_{12}S_{22}^{-1}S_{21}\right)^{-1}$$

$$S^{12} = -S^{11}S_{12}S_{22}^{-1}$$

$$S^{22} = \left(S_{22} - S_{21}S_{11}^{-1}S_{12}\right)^{-1}$$

$$S^{21} = -S^{22}S_{21}S_{11}^{-1}$$

and verify the identity

$$S_{22}^{-1}S_{21}S^{11} = S^{22}S_{21}S_{11}^{-1}.$$

9. In coordinate space R^p, partition x as $x = \binom{y}{z}$, and for $\Sigma > 0$, partition $\Sigma : p \times p$ conformably as

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}.$$

Define the inner product $(\cdot, \cdot)$ on R^p by $(u, v) = u'\Sigma^{-1}v$.

(i) Show that the matrix

$$P = \begin{pmatrix} I & -\Sigma_{12}\Sigma_{22}^{-1} \\ 0 & 0 \end{pmatrix}$$

defines an orthogonal projection in the inner product $(\cdot, \cdot)$. What is $\mathfrak{R}(P)$?

(ii) Show that the identity

$$\begin{pmatrix} y \\ z \end{pmatrix}'\Sigma^{-1}\begin{pmatrix} y \\ z \end{pmatrix} = \left(y - \Sigma_{12}\Sigma_{22}^{-1}z\right)'\Sigma^{11}\left(y - \Sigma_{12}\Sigma_{22}^{-1}z\right) + z'\Sigma_{22}^{-1}z$$

is the same as the identity

$$\|x\|^2 = \|Px\|^2 + \|(I - P)x\|^2$$

where $(x, x) = \|x\|^2$ and $x = \binom{y}{z}$.

(iii) For a random vector

$$X = \begin{pmatrix} Y \\ Z \end{pmatrix} \in R^p$$

with $\mathcal{L}(X) = N(0, \Sigma)$, $\Sigma > 0$, use part (ii) to give a direct proof via densities that the conditional distribution of Y given Z is $N(\Sigma_{12}\Sigma_{22}^{-1}Z, \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})$.

10. Verify the equation

$$\int_{G_T^+} \prod_1^p t_{ii}^{r-i} \exp\left[-\tfrac{1}{2}\sum_{j \leq i}^{P} t_{ij}^2\right] dT = 2^{-p}c(r, p)$$

where $c(r, p)$ is given in Example 5.1. Here, r is real, $r > p - 1$.

NOTES AND REFERENCES

1. Other matrix factorizations of interest in statistical problems can be found in Anderson (1958), Rao (1973), and Muirhead (1982). Many matrix factorizations can be viewed as results that give a maximal invariant under the action of a group—a topic discussed in detail in Chapter 7.

2. Only the most elementary facts concerning the transformation of measures under a change of variable have been given in the second section. The Jacobians of other transformations that occur naturally in statistical problems can be found in Deemer and Olkin (1951), Anderson (1958), James (1954), Farrell (1976), and Muirhead (1982). Some of these transformations involve functions defined on manifolds (rather than open subsets of R^n) and the corresponding Jacobian calculations require a knowledge of differential forms on manifolds. Otherwise, the manipulations just look like magic that somehow yields answers we do not know how to check. Unfortunately, the amount of mathematics behind these calculations is substantial. The mastery of this material is no mean feat. Farrell (1976) provides one treatment of the calculus of differential forms. James (1954) and Muirhead (1982) contain some background material and references.

3. I have found Lang (1969, Part Six, Global Analysis) to be a very readable introduction to differential forms and manifolds.