By Kenneth S. Miller

Multidimensional Gaussian Distributions

Multidimensional Gaussian Distributions

by Kenneth S. Miller

Professor of Mathematics
New York University
and Visiting Professor of Engineering Mathematics
Columbia University, 1962–64

John Wiley and Sons, Inc., New York · London · Sydney

Library of Congress Catalog Card Number: 64-23853
Printed in the United States of America

Preface

Our objective has been to present the basic facts concerning multidimensional Gaussian distributions in a concise, crisp, and we hope elegant form. We have restricted ourselves to properties that depend on the full generality of a nondiagonal covariance matrix and do not consider, *a priori*, problems where the n-dimensional distribution is the product of n one-dimensional distributions.

The language of matrix algebra and linear vector spaces has been used throughout. We have attempted to be extremely consistent in notation. Capital letters are always used for vectors and matrices, and lower-case letters are always used for scalars. Subscripts on vectors indicate the dimension. We have resisted almost all temptations to abrogate our conventions and yield to the exigencies of the moment, except at some points, especially in Chapter 4, where custom dictates otherwise.

We assume that the reader is familiar with the elementary facts concerning linear algebra and has some acquaintance with advanced calculus and probability theory.

Chapter 1 is devoted to a discussion of quadratic forms, certain special theorems in the theory of matrices, the covariance matrix, generalized spherical coordinates, and certain integrals of quadratic forms. In Chapter 2 we define the joint normal distribution and prove various theorems concerning Gaussian variates and linear combinations of Gaussian variates. This is followed by a lengthy treatment of Rayleigh distributions and distributions associated with the Rayleigh. We conclude with an envelope-type distribution. In Chapter 3 we consider certain associated functions such as the moment generating function and characteristic function. We also develop the theory of regression functions and linear least squares. The chapter concludes with a treatment of multidimensional singular Gaussian distributions. Applications to Gaussian random noise form the content of Chapter 4. The problem of estimating signal parameters in the presence of additive and multiplicative noise is considered. We also examine the statistics of the output noise resulting from the passage of a Gaussian process through a linear filter.

Equations are numbered by sections. In Section x of Chapter A we refer to equation α simply as (α). In Section x of Chapter A we refer to equation β of Section y of Chapter A (with $x \neq y$) as ($y.\beta$). In Section x of Chapter A we refer to equation γ of Section z of Chapter B (with $A \neq B$) as ($B.\gamma.z$).

K. S. MILLER

New York, New York
April 1963

Contents

CHAPTER 1

Quadratic Forms

1. INTRODUCTION

The study of multidimensional Gaussian distributions is essentially the study of a certain class of functions of n variables. These functions involve a quadratic form q in n variables and the determinant of the symmetric matrix associated with q. It is therefore appropriate that we begin our discussion with a brief review of certain salient properties of quadratic forms. In Section 2 we outline and sketch some proofs of the necessary critical formulas directly related to our future needs. Certain results in elementary matrix theory will also be required. Some of these theorems are perhaps not quite standard. We shall state and prove the requisite properties in a form immediately applicable to our subsequent developments (Section 3).

If $x_1, x_2, \cdots, x_n$ are random variables, then the mean or expected value of x_i is denoted by $\mathbf{E}x_i$, $1 \leqq i \leqq n$. The covariance of x_i and x_j is

$$\text{Cov}\,(x_i, x_j) = \mathbf{E}(x_i - a_i)(x_j - a_j),$$

where $\mathbf{E}x_i = a_i$ and $\mathbf{E}x_j = a_j$. The variance of x_i is

$$\text{Var}\,x_i = \mathbf{E}(x_i - a_i)^2 = \text{Cov}\,(x_i, x_i)$$

and is always nonnegative. If we write $\text{Cov}\,(x_i, x_j) = m_{ij}$, then the covariance matrix of $x_1, x_2, \cdots, x_n$ is the $n \times n$ (symmetric) matrix

$$M = \| m_{ij} \|_{1 \leqq i, j \leqq n}$$

where m_{ij} is the element in the ith row and jth column of M. Certain definitions and properties associated with covariance matrices are treated in Section 4.

Frequently we have occasion to consider n-fold integrals of functions of quadratic forms. In order to evaluate these integrals it will be convenient to introduce certain changes of variables. Specifically, the analog of classical three-dimensional spherical coordinates in n dimensions will be a particularly useful transformation. Following Blumenson [A derivation of n-dimensional spherical coordinates, *American Mathematical Monthly*, **67**, No. 1, 63–66 (1960)] we shall deduce the appropriate coordinate changes using the methods of finite-dimensional vector spaces. This is done in Section 5. Finally, in Section 6, we shall evaluate certain integrals involving quadratic forms.

2. QUADRATIC FORMS

If P is a matrix, we shall denote its transpose by a prime: P'. In particular, unprimed vectors will always be assumed to be column vectors; and hence primed vectors are row vectors. We shall consistently use capital letters for vectors and matrices and lower-case letters for scalars. Let $M = \|m_{ij}\|_{1 \leq i,j \leq n}$ be an $n \times n$ symmetric matrix and $X = \{x_1, x_2, \cdots, x_n\}$ a column vector. Then

$$q(X) = X'MX = \sum_{i,j=1}^{n} m_{ij}x_ix_j$$

is called the *quadratic form associated with M.*

Let P be a nonsingular $n \times n$ matrix and consider the transformation

$$X = PY \quad \text{or} \quad Y = P^{-1}X.$$

Then

$$q(X) = (PY)'M(PY) = (Y'P')M(PY) = Y'(P'MP)Y = q^*(Y),$$

and $q^*(Y)$ is the quadratic form associated with $M^* = P'MP$. Clearly M^* is symmetric since $(P'MP)' = P'M'P = P'MP = M^*$.

We shall now assume that all entries in our matrices belong to the field of real numbers. Let M be a symmetric $n \times n$ matrix and $q(X) = X'MX$ the associated quadratic form. Then one of the principal results in the elementary theory of quadratic forms states that $q(X)$ may be reduced to diagonal quadratic form by nonsingular linear transformations. Precisely worded, if $q(X)$ is not identically zero, then there exists a nonsingular matrix P such that if $X = PY$, then the quadratic form $q(X)$ may be written

$$q(X) = Y'(P'MP)Y = \sum_{i=1}^{r} d_iy_i^2, \qquad d_i \neq 0, \quad 1 \leqq i \leqq r \leqq n, \tag{1}$$

where $Y = \{y_1, y_2, \cdots, y_n\}$. We shall sketch a proof of this important result.

Suppose first that the coefficients of the quadratic terms x_i^2 are not all zero. Without loss of generality suppose $m_{11} \neq 0$. Then

$$q(X) = m_{11}x_1^2 + 2x_1 \sum_{i=2}^{n} m_{1i}x_i + \sum_{i,j=2}^{n} m_{ij}x_ix_j.$$

Completing the square,

$$q(X) = m_{11}\left(x_1 + m_{11}^{-1} \sum_{i=2}^{n} m_{1i}x_i\right)^2 + q'(X) \tag{2}$$

where $q'(X)$ is a quadratic form involving only $x_2, x_3, \cdots, x_n$. If we make the nonsingular linear transformation

$$y_1 = x_1 + m_{11}^{-1} \sum_{i=2}^{n} m_{1i}x_i$$

$$y_j = x_j, \qquad 2 \leqq j \leqq n,$$

then (2) becomes

$$q(X) = m_{11}y_1^2 + q'(Y), \qquad m_{11} \neq 0,$$

where $q'(Y)$ is a quadratic form in the $n - 1$ variables $y_2, y_3, \cdots, y_n$.

If $m_{ii} = 0$, $1 \leqq i \leqq n$, then there must be some $m_{ij} \neq 0$ with $i \neq j$ since $q(X) \not\equiv 0$. Suppose without loss of generality that $m_{12} \neq 0$. Then

$$q(X) = 2m_{12}x_1x_2 + q^*(X), \tag{3}$$

where $q^*(X)$ does not involve x_1^2, x_2^2 or x_1x_2. The nonsingular linear transformation

$$y_1 = x_1 + x_2$$

$$y_2 = x_1 - x_2$$

$$y_j = x_j, \qquad 3 \leqq j \leqq n,$$

reduces (3) to

$$q(X) = \tfrac{1}{2}m_{12}y_1^2 + q^{\dagger}(Y)$$

where $q^{\dagger}(Y)$ does not contain y_1^2. One can then proceed as before.

The number r of nonzero diagonal terms in (1) is called the *rank* of the form q and is invariant under nonsingular linear transformations. From (1) we see that $P'MP$ is a diagonal matrix. Since $P'MP$ is nonsingular if and only if M is nonsingular, it follows that $q(X)$ has rank n if and only if M is nonsingular. It also follows from (1) that, by trivial, real, nonsingular linear transformations, $q(X)$ may be reduced to the form

$$q(X) = z_1^2 + z_2^2 + \cdots + z_p^2 - (z_{p+1}^2 + z_{p+2}^2 + \cdots + z_r^2).$$

The number p of positive squares is an invariant of the quadratic form.

The difference between the number of positive and negative terms in q is $s = p - (r - p) = 2p - r$, where r is the rank of q. We call s the *signature* of q.

If the signature and rank of a quadratic form are equal, then we say that the form is *nonnegative definite* or *positive semidefinite*. Thus if $q(X)$ is nonnegative definite,

$$q(X) = z_1^2 + z_2^2 + \cdots + z_r^2, \qquad r \leqq n. \tag{4}$$

If $r = n$, then we call q *positive definite*. A symmetric matrix M is called nonnegative definite if and only if $X'MX$ is nonnegative definite and is called positive definite if and only if $X'MX$ is positive definite. It easily follows that $q(X) = X'MX = \Sigma m_{ij}x_ix_j$ is nonnegative definite if and only if $q(X) \geqq 0$ for all X. If $q(X) = 0$ if and only if $X = 0$, then it follows that $q(X)$ is positive definite. Thus a nonsingular nonnegative definite symmetric matrix is positive definite.

If a symmetric matrix M is positive definite, then it follows from (4) that there exists a nonsingular linear transformation P such that $P'MP$ is the identity matrix I,

$$P'MP = I. \tag{5}$$

Thus a symmetric matrix M is positive definite if and only if there exists a nonsingular $n \times n$ matrix Q such that

$$M = Q'Q. \tag{6}$$

If M is a symmetric matrix, then there exists an orthogonal matrix R (that is, $R' = R^{-1}$) such that

$$R'MR = D, \tag{7}$$

where D is a diagonal matrix. In fact

$$D = \|\lambda_i \, \delta_{ij}\|_{1 \leqq i, j \leqq n}$$

where the λ_i, $1 \leqq i \leqq n$, are the (necessarily real but not necessarily distinct) characteristic roots of M (and δ_{ij} is the Kronecker delta). Thus we may state that M is positive definite if and only if $\lambda_i > 0$, $1 \leqq i \leqq n$.

For a further discussion of the results of this section we refer the reader to Birkhoff and MacLane (*A survey of modern algebra*, The Macmillan Company, New York, 1953, Chapter 9); Bôcher (*Introduction to higher algebra*, The Macmillan Company, New York, 1907, Chapters 10 and 11); Finkbeiner (*Introduction to matrices and linear transformations*, W. H. Freeman and Company, San Francisco, 1960, Chapter 9); and Hohn (*Elementary matrix algebra*, The Macmillan Company, New York, 1958, Chapter 9).

3. SOME MATRIX THEOREMS

Certain results from matrix algebra will be needed from time to time in our future work. In order not to interrupt the continuity of our development we collect the pertinent formulas in this section. They include two cases of Shur's identity and a special form of Jacobi's theorem. We also discuss the important problem of differentiating a scalar with respect to a matrix.

The proofs of the lemmas are most conveniently carried out using partitioned matrices. Let $M_n = \|m_{ij}\|_{1 \leq i,j \leq n}$ be a positive definite $n \times n$ symmetric matrix. Partition the matrix M_n as

$$M_n = \left\| \begin{array}{c|c} M_p & T_p \\ \hline T_p' & S_{n-p} \end{array} \right\| \tag{1}$$

where $M_p = \|m_{ij}\|_{1 \leq i,j \leq p}$ is the $p \times p$ matrix in the upper left-hand corner of M_n, S_{n-p} is the $(n-p) \times (n-p)$ matrix in the lower right-hand corner of M_n, T_p is the $p \times (n-p)$ matrix in the upper right-hand corner of M_n, and T_p', the transpose of T_p, is the $(n-p) \times p$ matrix in the lower left-hand corner of M_n. Since M_n is symmetric, $M_p = M_p'$ and $S_{n-p} = S_{n-p}'$.

The matrix M_n is nonsingular by hypothesis, and hence M_n^{-1} exists. Write M_n^{-1} as the partitioned matrix

$$M_n^{-1} = \left\| \begin{array}{c|c} Q_p^{-1} & R_p \\ \hline R_p' & P_{n-p}^{-1} \end{array} \right\| \tag{2}$$

where Q_p^{-1} is the $p \times p$ matrix in the upper left-hand corner of M_n^{-1}, P_{n-p}^{-1} is the $(n-p) \times (n-p)$ matrix in the lower right-hand corner of M_n^{-1}, R_p is the $p \times (n-p)$ matrix in the upper right-hand corner of M_n^{-1}, and R_p', the transpose of R_p, is the $(n-p) \times p$ matrix in the lower left-hand corner of M_n^{-1}. Since M_n^{-1} is symmetric, $Q_p' = Q_p$ and $P_{n-p}' = P_{n-p}$.

Using this notation we shall prove the following three lemmas.

LEMMA 1. $P_{n-p}R_p' = -T_p'M_p^{-1}$.

LEMMA 2. $Q_p^{-1} = M_p^{-1} + R_pP_{n-p}R_p'$.

LEMMA 3. $|P_{n-p}| = |M_n|\,|M_p|^{-1}$.

The first two lemmas are called Shur's identities, and the last is a special case of Jacobi's theorem. (The bars in Lemma 3 denote determinants. Thus, if A is a square matrix we write det A or $|A|$ to indicate the determinant of A.)

Proof of Lemma 1. We have

$$M_n M_n^{-1} = I_n, \tag{3}$$

where $I_n = \|\delta_{ij}\|_{1 \leq i,j \leq n}$ is the $n \times n$ identity matrix. Using (1) and (2) we may write (3) in expanded form as

$$\left\| \begin{array}{c|c} M_p Q_p^{-1} + T_p R_p' & M_p R_p + T_p P_{n-p}^{-1} \\ \hline T_p' Q_p^{-1} + S_{n-p} R_p' & T_p' R_p + S_{n-p} P_{n-p}^{-1} \end{array} \right\| = \left\| \begin{array}{c|c} I_p & O_{p,n-p} \\ \hline O_{n-p,p} & I_{n-p} \end{array} \right\| \tag{4}$$

where $O_{rs} = \|0_{ij}\|_{1 \leq i \leq r, 1 \leq j \leq s}$ is the $r \times s$ matrix with zeros everywhere.

Equating the terms in the upper right-hand corners of the matrices of (4) we obtain

$$M_p R_p + T_p P_{n-p}^{-1} = O_{p,n-p} \tag{5}$$

or

$$M_p R_p = -T_p P_{n-p}^{-1}. \tag{6}$$

Multiply on the right by P_{n-p} and on the left by M_p^{-1}. The transpose of this resulting equation is Lemma 1.

Proof of Lemma 2. Equating the matrices in the upper left-hand corners of the partitioned matrices of (4) we may write

$$M_p Q_p^{-1} + T_p R_p' = I_p. \tag{7}$$

Multiply on the left by M_p^{-1} to obtain

$$Q_p^{-1} = M_p^{-1} - M_p^{-1} T_p R_p'. \tag{8}$$

But by Lemma 1, $M_p^{-1} T_p = -R_p P_{n-p}$. Substituting in (8) yields Lemma 2.

Proof of Lemma 3. Let K_p be a $p \times (n - p)$ matrix which we shall specify later. Then

$$\left\| \begin{array}{c|c} M_p & K_p \\ \hline O_{n-p,p} & P_{n-p} \end{array} \right\| \cdot \left\| \begin{array}{c|c} Q_p^{-1} & R_p \\ \hline R_p' & P_{n-p}^{-1} \end{array} \right\| = \left\| \begin{array}{c|c} M_p Q_p^{-1} + K_p R_p' & M_p R_p + K_p P_{n-p}^{-1} \\ \hline P_{n-p} R_p' & I_{n-p} \end{array} \right\|. \tag{9}$$

Choose K_p such that the matrix in the upper right-hand corner of the partitioned matrix on the right-hand side of (9) vanishes:

$$M_p R_p + K_p P_{n-p}^{-1} = O_{p,n-p}.$$

Then

$$K_p = -M_p R_p P_{n-p},$$

and the term in the upper left-hand corner of the matrix on the right-hand side of (9) becomes

$$M_pQ_p^{-1} + K_pR_p' = M_pQ_p^{-1} - M_pR_pP_{n-p}R_p'$$
$$= M_p(Q_p^{-1} - R_pP_{n-p}R_p') = M_pM_p^{-1} = I_p$$

by Lemma 2. If we take the determinant of both sides of (9), we obtain

$$|M_p|\,|P_{n-p}|\,|M_n|^{-1} = 1,$$

which proves Lemma 3.

We now turn to the problem of differentiating a scalar function of a matrix with respect to that matrix. Let $v = q(X)$ be a scalar function of the $m \times n$ matrix $X = \|x_{ij}\|_{1 \leqq i \leqq m, 1 \leqq j \leqq n}$. We define dv/dX as the $m \times n$ matrix

$$\left\| \frac{\partial v}{\partial x_{ij}} \right\|_{1 \leqq i \leqq m, 1 \leqq j \leqq n},$$

where the x_{ij}, $1 \leqq i \leqq m$, $1 \leqq j \leqq n$, are to be thought of as mn independent variables and the $\partial v/\partial x_{ij}$ are assumed to exist.

The most common situation arises when X is a column vector,

$$X = \{x_1, x_2, \cdots, x_n\},$$

and $q(X)$ is of the form $X'MX$ or $X'A$ where M is an $n \times n$ symmetric matrix and A is an n-dimensional vector (both independent of X). It is easy to verify that in these cases

$$\frac{d}{dX}(X'MX) = 2MX \tag{10}$$

and

$$\frac{d}{dX}(X'A) = \frac{d}{dX}(A'X) = A. \tag{11}$$

4. COVARIANCE MATRICES

Let $x_1, x_2, \cdots, x_n$ be n random variables with means $a_1, a_2, \cdots, a_n$ respectively. Then their *covariance matrix* M_n is defined as

$$M_n = \mathbf{E}(X_n - A_n)(X_n - A_n)' = \|\mathbf{E}(x_i - a_i)(x_j - a_j)\|_{1 \leqq i,j \leqq n}$$
$$= \|\mathrm{Cov}\,(x_i, x_j)\|_{1 \leqq i,j \leqq n} = \|m_{ij}\|_{1 \leqq i,j \leqq n},$$

where $X_n = \{x_1, x_2, \cdots, x_n\}$ and $A_n = \{a_1, a_2, \cdots, a_n\}$ are column vectors. Clearly a covariance matrix is symmetric. It is also nonnegative definite, since

$$0 \leqq \mathbf{E}[Z_n'(X_n - A_n)]^2 = \mathbf{E}\{Z_n'(X_n - A_n)[Z_n'(X_n - A_n)]'\}$$
$$= Z_n'\mathbf{E}(X_n - A_n)(X_n - A_n)'Z_n = Z_n'M_nZ_n$$

for any vector $Z_n = \{z_1, z_2, \cdots, z_n\}$. A nonsingular covariance matrix M_n is therefore positive definite. In such cases M_n^{-1} is also a symmetric positive definite matrix.

If M_n is positive definite, we can always choose n linear combinations of the x_i which are uncorrelated (that is, have a diagonal covariance matrix). For suppose M_n is positive definite. Then there exists a nonsingular $n \times n$ matrix Q_n such that

$$M_n = Q_n'Q_n,$$

[compare (2.6)]. Let

$$Y_n = Q_n'^{-1}X_n$$

and

$$B_n = \mathbf{E}Y_n = Q_n'^{-1}\mathbf{E}X_n = Q_n'^{-1}A_n.$$

Thus if N_n is the covariance matrix of Y_n,

$$N_n = \mathbf{E}(Y_n - B_n)(Y_n - B_n)' = \mathbf{E}(Q_n'^{-1}X_n - B_n)(Q_n'^{-1}X_n - B_n)'$$
$$= Q_n'^{-1}[\mathbf{E}(X_n - A_n)(X_n - A_n)']Q_n^{-1} = Q_n'^{-1}M_nQ_n^{-1}$$
$$= Q_n'^{-1}Q_n'Q_nQ_n^{-1} = I_n.$$

A further refinement of this result is obtained by using orthogonal transformations. Suppose $Y_n = \{y_1, y_2, \cdots, y_n\}$ is a random vector with mean vector $B_n = \mathbf{E}Y_n = \{b_1, b_2, \cdots, b_n\}$ and diagonal covariance matrix $\psi_0 I_n = \|\psi_0\delta_{ij}\|_{1 \leq i,j \leq n}$ where ψ_0 is a positive constant. Let

$$Z_n = R_nY_n$$

where R_n is an orthogonal $n \times n$ matrix. Then

$$\mathbf{E}Z_n = R_n\mathbf{E}Y_n = R_nB_n,$$

and the covariance matrix P_n of Z_n is

$$P_n = \mathbf{E}(Z_n - R_nB_n)(Z_n - R_nB_n)' = R_n[\mathbf{E}(Y_n - B_n)(Y_n - B_n)']R_n'$$
$$= R_n\psi_0 I_nR_n' = \psi_0R_nR_n' = \psi_0R_nR_n^{-1} = \psi_0 I_n.$$

Thus if Y_n and Z_n are related by an orthogonal transformation, $Z_n = R_nY_n$, and if Y_n has the diagonal covariance matrix $\psi_0 I_n$; then Z_n also has $\psi_0 I_n$ as its covariance matrix.

5. GENERALIZED SPHERICAL COORDINATES

In certain multiple integrals involving functions of quadratic forms in n variables we shall have occasion to make changes of variables. One important transformation is the analog of spherical coordinates in three dimensions. The generalization of spherical coordinates to n dimensions is by no means unique. For example, the following three transformations are all examples of "generalized" spherical coordinates in five-dimensional Euclidean space $\mathscr{V}^5$:

$$(S_1)\qquad \begin{aligned} x_1 &= r\cos\phi_1 \\ x_2 &= r\sin\phi_1\cos\phi_2 \\ x_3 &= r\sin\phi_1\sin\phi_2\cos\phi_3 \\ x_4 &= r\sin\phi_1\sin\phi_2\sin\phi_3\cos\theta \\ x_5 &= r\sin\phi_1\sin\phi_2\sin\phi_3\sin\theta \end{aligned}$$

$$(S_2)\qquad \begin{aligned} x_1 &= r\cos\xi_1 \\ x_2 &= r\sin\xi_1(\cos\eta\cos\zeta) \\ x_3 &= r\sin\xi_1(\cos\eta\sin\zeta) \\ x_4 &= r\sin\xi_1(\sin\eta\cos v) \\ x_5 &= r\sin\xi_1(\sin\eta\sin v) \end{aligned}$$

$$(S_3)\qquad \begin{aligned} x_1 &= r\cos\alpha_1(\cos\beta) \\ x_2 &= r\cos\alpha_1(\sin\beta) \\ x_3 &= r\sin\alpha_1[\cos\alpha_2] \\ x_4 &= r\sin\alpha_1[\sin\alpha_2\cos\gamma] \\ x_5 &= r\sin\alpha_1[\sin\alpha_2\sin\gamma]. \end{aligned}$$

The first two represent a decomposition of the five-dimensional sphere of radius r into a one-dimensional and a four-dimensional sphere. In S_1 the four-dimensional sphere is decomposed into a one-dimensional and a three-dimensional sphere; in S_2 the four-dimensional sphere is decomposed into two two-dimensional spheres. The third transformation S_3 is a decomposition of the five-dimensional sphere of radius r into a two-dimensional and a three-dimensional sphere. [See, for example, Shelupsky, An introduction of spherical coordinates, *American Mathematical Monthly*, **69**, No. 7, 644–646 (1962). There seems to be an error in his formula for the number, s_n, of distinct coordinate systems in $\mathscr{V}^n$.] For our purposes the first, S_1, is the most natural and most useful generalization. We shall derive this case in n-dimensional Euclidean space $\mathscr{V}^n$ by the techniques of linear algebra.

We begin our derivation by recalling a few definitions. If $V_n = \{v_1, v_2, \cdots, v_n\}$ and $W_n = \{w_1, w_2, \cdots, w_n\}$ are column vectors, then

$$V_n'W_n = W_n'V_n = \sum_{i=1}^{n} v_i w_i$$

is called the *inner product* of V_n and W_n. We write

$$|V_n|^2 = V_n'V_n = \sum_{i=1}^{n} v_i^2$$

and call $|V_n|$, the nonnegative square root of $\Sigma_{i=1}^n v_i^2$, the *norm* of V_n. The angle ϕ between two nonzero vectors V_n and W_n is defined by

$$\cos\phi = \frac{V_n'W_n}{|V_n|\,|W_n|}, \qquad 0 \leqq \phi \leqq \pi.$$

Let $E_n^{(1)}, E_n^{(2)}, \cdots, E_n^{(n)}$ be an orthonormal basis in an n-dimensional unitary space $\mathscr{V}^n$ (that is, $E_n^{(i)\prime}E_n^{(j)} = \delta_{ij}$, $1 \leqq i, j \leqq n$). Let X_n be any vector on the n-dimensional sphere of radius r with center at the origin. Then $|X_n| = r$, and if

$$X_n = \sum_{i=1}^{n} x_i E_n^{(i)}, \tag{1}$$

we have

$$|X_n|^2 = \sum_{i=1}^{n} x_i^2 = r^2.$$

Let θ_i be the angle between X_n and $E_n^{(i)}$. Then

$$\cos\theta_i = \frac{X_n'E_n^{(i)}}{|X_n|\,|E_n^{(i)}|} = \frac{x_i}{r}, \qquad 0 \leqq \theta_i \leqq \pi,$$

since $|E_n^{(i)}| = 1$ and $X_n'E_n^{(i)} = x_i$ by virtue of the orthogonality of the $E_n^{(i)}$. Hence (1) implies

$$X_n = r\sum_{i=1}^{n} \cos\theta_i\, E_n^{(i)}. \tag{2}$$

Thus X_n may be specified by giving its length r and the n angles θ_i. Since

$$r^2 = X_n'X_n = r^2\sum_{i=1}^{n}\cos^2\theta_i$$

we see that the θ_i, $1 \leqq i \leqq n$, are not independent. Our derivation of spherical coordinates will show how to choose $n-1$ angles $\phi_1, \phi_2, \cdots, \phi_{n-2}, \theta$ which are independent of each other and which, when combined with the norm r, completely describe the vector X_n with respect to the given orthonormal basis, $E_n^{(1)}, E_n^{(2)}, \cdots, E_n^{(n)}$.

Let $E_n^{(i)}$, $1 \leqq i \leqq n$, be an orthonormal basis. Let X_n be a vector on the sphere of radius r with center at the origin. Let ϕ_1 be the angle between X_n and $E_n^{(1)}$. Then [compare (1) and (2)],

$$X_n = r\cos\phi_1\, E_n^{(1)} + \sum_{i=2}^{n} x_i E_n^{(i)}, \qquad 0 \leqq \phi_1 \leqq \pi. \tag{3}$$