JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 33, No. 1, JANUARY 1981

A Conjugate Directions Method for Subjective Assessment of Normal Distributions

S. S. OREN[1]

Communicated by R. A. Howard

Abstract. A procedure is developed that enables the encoding of a subjective n-dimensional joint normal probability density function through the assessment of its marginal means and variances and $n(n-1)/2$ conditional means. The new method is based on the theory of conjugate directions for quadratic forms, and it exploits the fact that normal distributions have quadratic equal-likelihood surfaces. Unlike previous approaches, this new method enables easy detection and resolution of inconsistencies in the assessments that could lead to an indefinite estimate of the covariance matrix.

Key Words. Conjugate directions, quadratic forms, normal distributions, probability assessment, decision analysis.

1. Introduction

Probability assessment plays a major role in the application of decision analysis. Consequently, a substantial portion of the literature in this area has been devoted to that topic. An exposition and further bibliography related to probability encoding are given in a paper by Spetzler and Stael von Holstein (Ref. 1). That paper also describes a methodology representing the state of the art in this area. As with most of the literature on this subject, the above paper focuses mainly on assessment of probabilities over discrete events or probability distributions over single random variables. In analyzing practical decision problems, one often encounters situations where variables are not independent and it is necessary to assess their joint probability distribution. A common approach in such cases is to discretize these variables and encode marginal and conditional distributions on single variables using methodologies of the type described in Ref. 1. This brute-force approach becomes impractical, however, even for a modest number of

[1] Research Scientist, Xerox Palo Alto Research Center, Palo Alto, California. Currently: Associate Professor, Department of Engineering Economic Systems, Stanford University, Stanford, California.

0022-3239/81/0100-0025$03.00/0 © 1981 Plenum Publishing Corporation

dependent variables, since the number of required assessments grows exponentially with the number of variables.

An alternative approach for dealing with decision problems involving dependent variables is the proximal decision analysis, proposed by Howard (Ref. 2). This approach is based on using a quadratic approximation of the utility function and requires only the means and the covariance matrix of the random variables. To encode this information, Howard (Ref. 2) proposes a method that, for n dependent variables, requires the assessment of n marginal distributions and $n(n-1)/2$ functions of the form $\langle x_i|x_j\rangle$, $i<j\leq n$, $i=1,\ldots,n$, that describe the conditional mean of the variables x_i for all values of x_j.

In a recent dissertation, Buede (Ref. 3) has developed a new technique for assessing joint probability distribution on time-dependent random variables. This approach is quite involved and requires a considerable amount of information on conditional fractiles. Furthermore, the approach is limited to the case where the dependent variables are essentially the values of a single time-dependent variable at different time instants.

The simplest and most direct approach for assessing joint probability distributions is based on using parametric multivariate distributions to represent subjective probabilities. Such a representation reduces the encoding task to that of estimating the parameters for the assumed distribution. Clearly, this approach imposes severe limitations on the shape of the encoded distributions and should not be used when such limitations are unacceptable. On the other hand, simplicity considerations make this approach extremely attractive whenever the implied approximation is reasonable. Particularly attractive is the normal distribution that is extensively used in control estimation theory and many other areas. In spite of its limitation, there are situations where the normal distribution can reasonably represent the decision maker's state of information. Normal approximations of the true density function can also be used to estimate the covariance matrix of a distribution needed for Howard's (Ref. 2) proximal decision analysis.

This paper develops an efficient procedure for eliciting subjective assessments of joint probability density functions represented by multivariate normal distributions. In particular, we exploit the special properties of normal distributions to simplify the encoding process.

Since the normal distribution is completely characterized by the means and the covariance matrix of the variables, one may consider encoding these parameters directly. It appears, however, that a decision maker would find it quite difficult to assess the covariances of every pair of variables directly. Buede (Ref. 3) suggested encoding subjective assessments of normal distributions by using Howard's (Ref. 2) aforementioned technique for encoding covariance matrices of general distributions. This approach is

appealing from the point of view of the assessor's ability to provide the needed information. However, the method is quite inefficient, as it requires far more information than necessary for estimating the parameters of a normal distribution. Both the direct assessment approach and Howard's method estimate the covariance matrix from assessed pairwise relationships among the variables. The most serious drawback of such methods stems from the potential inconsistencies of the assessments, which can lead to indefinite estimates of the covariance matrix (which must be positive semidefinite). In general, one cannot guarantee that covariance estimates based on subjective assessments of this type form a positive semidefinite matrix. The definiteness of a covariance matrix estimate $\mathbf{Q}$ can be tested fairly easily by factoring it, for example (using standard algorithms), into a product of the form

$$\mathbf{Q} = \mathbf{D}^T \mathbf{\Lambda} \mathbf{D},$$

where $\mathbf{D}$ is an upper triangular matrix and $\mathbf{\Lambda}$ is diagonal. The nonnegativity of the elements of $\mathbf{\Lambda}$ is necessary and sufficient for positive semidefiniteness of $\mathbf{Q}$. The definiteness of $\mathbf{Q}$ can also be tested by checking whether its eigenvalues or alternatively the determinants of its principal minors are nonnegative. Unfortunately, none of these test quantities can be related in a simple fashion to assessments of the type mentioned above. Consequently, there is no intuitive way to demonstrate such inconsistencies to the assessor so that he can resolve them or prevent them.

In the following sections, we introduce a new approach that enables us to fully encode a joint normal distribution of n random variables using the means and variances of these variables together with the values of $n(n-1)/2$ conditional means. The proposed procedure is such that the covariance matrix estimate is obtained directly in an $\mathbf{D}^T\mathbf{\Lambda}\mathbf{D}$ product form. The diagonal matrix $\mathbf{\Lambda}$ results from solving a simple lower-triangular system of linear equations involving the assessed quantities. This system permits easy detection of inconsistencies (which show up as negative elements in $\mathbf{\Lambda}$). It also enables simple interpretation and resolution of such inconsistencies. The total number of assessed values required in the basic version of this method equals the number of parameters of the normal distribution. Thus, no information is being *wasted.* Additional assessed values may be needed, however, to assure robustness of the estimates and resolve inconsistencies.

2. Basic Idea and Theoretical Background

The multivariate normal distribution is defined as

$$N_x(\bar{\mathbf{x}}, \mathbf{Q}) = K_x \exp(-\tfrac{1}{2}(\mathbf{x}-\bar{\mathbf{x}})^T \mathbf{Q}^{-1}(\mathbf{x}-\bar{\mathbf{x}})), \tag{1a}$$

with

$$K_x = (2\pi)^{-n/2}|\mathbf{Q}|^{1/2}, \tag{1b}$$

where $\mathbf{x}$ is a random vector, i.e., a vector whose components are random variables and

$$\mathbf{Q} = \overline{(\mathbf{x}-\bar{\mathbf{x}})(\mathbf{x}-\bar{\mathbf{x}})^T}$$

is the covariance matrix of $\mathbf{x}$. To completely characterize $N_x(\bar{\mathbf{x}}, \mathbf{Q})$, it is clearly sufficient to specify $\bar{\mathbf{x}}$ and $\mathbf{Q}$. As mentioned in the Introduction, however, assessing the off-diagonal elements of $\mathbf{Q}$ [i.e., $\text{cov}(x_i, x_j)$ for every i and j] may be difficult.

Suppose that we had a matrix $\mathbf{D}$ whose rows $\mathbf{d}_i$ satisfied

$$\mathbf{d}_i^T\mathbf{Q}^{-1}\mathbf{d}_j = 0, \qquad \text{for } i \neq j;$$

i.e., the vectors $\mathbf{d}_i, \ldots, \mathbf{d}_n$ are conjugate with respect to $\mathbf{Q}^{-1}$ (see Luenberger, Ref. 4). Then,

$$\mathbf{D}\mathbf{Q}^{-1}\mathbf{D}^T = \mathbf{\Lambda}^{-1}, \tag{2}$$

where $\mathbf{\Lambda}$ is a diagonal matrix. Let λ_i and q_i denote the ith diagonal element of $\mathbf{\Lambda}$ and $\mathbf{Q}$, respectively; then, since

$$\mathbf{Q} = \mathbf{D}^T\mathbf{\Lambda}\mathbf{D},$$

we have

$$q_j = \sum_{i=1}^{n} d_{ij}^2\lambda_i, \qquad j = 1, \ldots, n. \tag{3}$$

Using the *box operator*,[2] we can rewrite (3) as

$$\mathbf{q} = (\mathbf{D}\square\mathbf{D})^T\boldsymbol{\lambda}, \tag{4}$$

where

$$\mathbf{q} = (q_1, \ldots, q_n) \qquad \text{and} \qquad \boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n).$$

Using the notation

$$\mathbf{\Lambda} = \text{diag}(\boldsymbol{\lambda}),$$

we can write now

$$\mathbf{Q} = \mathbf{D}^T[\text{diag}((\mathbf{D}^T\square\mathbf{D}^T)^{-1}\mathbf{q})]\mathbf{D}. \tag{5}$$

[2] $\mathbf{C} = \mathbf{A}\square\mathbf{B}$ implies that $c_{ij} = a_{ij}b_{ij}$ for all i, j where a_{ij}, b_{ij}, c_{ij} are the respective elements of the matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}$.

Recall that the elements of $\mathbf{q}$ are just the variances of the random variables x_i. It is assumed in this paper that these variances are available. Thus, to obtain $\mathbf{Q}$, it only remains to find a set of conjugate vectors $\mathbf{d}_i$.

The following propositions form the foundation of the method for obtaining the vectors $\mathbf{d}_i$.

Proposition 2.1. Let

$$f(\mathbf{x}) = \tfrac{1}{2}(\mathbf{x}-\mathbf{x}^*)^T H(\mathbf{x}-\mathbf{x}^*),$$

where $\mathbf{x} \in R^n$ and $\mathbf{H}$ is a positive-semidefinite matrix. Let M and N be linear subspaces, such that $N \subseteq M \subseteq R^n$ and $\mathbf{x}_0$ is a fixed point. Given that

$$\mathbf{x}_1 = \arg \min_{N+\mathbf{x}_0} f(\mathbf{x}), \tag{6}$$

$$\mathbf{x}_2 = \arg \min_{M+\mathbf{x}_0} f(\mathbf{x}), \tag{7}$$

then $\mathbf{x}_1 - \mathbf{x}_2$ is conjugate (with respect to H) to any vector in N; i.e.,

$$(\mathbf{x}_1-\mathbf{x}_2)^T \mathbf{H}\mathbf{d} = 0, \qquad \forall \mathbf{d} \in N. \tag{8}$$

Proof. The first-order necessary conditions for a minimum imply that

$$\nabla f(\mathbf{x}_1)\mathbf{d} = (\mathbf{x}_1-\mathbf{x}^*)^T \mathbf{H}\mathbf{d} = 0, \qquad \forall \mathbf{d} \in N, \tag{9}$$

$$\nabla f(\mathbf{x}_2)\mathbf{d} = (\mathbf{x}_2-\mathbf{x}^*)^T \mathbf{H}\mathbf{d} = 0, \qquad \forall \mathbf{d} \in M \supseteq N. \tag{10}$$

Thus,

$$(\nabla f(\mathbf{x}_1) - \nabla f(\mathbf{x}_2))\mathbf{d} = (\mathbf{x}_1-\mathbf{x}_2)\mathbf{H}\mathbf{d} = 0, \qquad \forall \mathbf{d} \in N. \qquad \square \tag{11}$$

The above result implies that a set of conjugate vectors can be constructed using information about the location of minimum points of the corresponding quadratic form over various linear manifolds. The following proposition suggests a way for eliciting such information from a probability assessor.

Proposition 2.2. Let $E_x\{f(\mathbf{x}) | \mathbf{x} \in L\}$ denote the conditional expectation of a random vector $\mathbf{x}$ having density function $f(\mathbf{x})$, given that $\mathbf{x}$ is restricted to the set L; and let $N_{\mathbf{x}}(\bar{\mathbf{x}}, \mathbf{Q})$ denote the normal distribution with mean $\bar{\mathbf{x}}$ and covariance matrix $\mathbf{Q}$. Then, for a linear subspace $M \subseteq R^n$ and a fixed $\mathbf{x}_0 \in R^n$,

$$\arg \min_{\mathbf{y} \in M+\mathbf{x}_0} (\mathbf{y}-\bar{\mathbf{x}})^T \mathbf{Q}^{-1} (\mathbf{y}-\bar{\mathbf{x}}) = E_x\{N_x(\bar{\mathbf{x}}, \mathbf{Q}) | \mathbf{x} \in M + \mathbf{x}_0\}. \tag{12}$$

Proof. Let $\mathbf{G}$ be a full-rank matrix whose columns span M. Then, any $\mathbf{x} \in M + \mathbf{x}_0$ can be written as

$$\mathbf{x} = \mathbf{G}\boldsymbol{\alpha} + \mathbf{x}_0.$$

Let

$$\mathbf{y}^* = \arg \min_{\mathbf{y} \in M + \mathbf{x}_0} (\mathbf{y} - \bar{\mathbf{x}})^T \mathbf{Q}^{-1} (\mathbf{y} - \bar{\mathbf{x}}). \tag{13}$$

Then, there exist an $\boldsymbol{\alpha}^*$ such that

$$\mathbf{y}^* = \mathbf{G}\boldsymbol{\alpha}^* + \mathbf{x}_0. \tag{14}$$

Furthermore, the gradient of the quadratic function at $\mathbf{y}^*$ is orthogonal to M, so that

$$\mathbf{G}^T \mathbf{Q}^{-1} (\mathbf{y}^* - \bar{\mathbf{x}}) = \mathbf{G}^T \mathbf{Q}^{-1} (\mathbf{G}\boldsymbol{\alpha}^* + \mathbf{x}_0 - \bar{\mathbf{x}}) = 0. \tag{15}$$

Thus,

$$\boldsymbol{\alpha}^* = (\mathbf{G}^T \mathbf{Q}^{-1} \mathbf{G})^{-1} \mathbf{G}^T \mathbf{Q}^{-1} (\bar{\mathbf{x}} - \mathbf{x}_0), \tag{16}$$

$$\mathbf{y}^* = \mathbf{G}(\mathbf{G}^T \mathbf{Q}^{-1} \mathbf{G})^{-1} \mathbf{G}^T \mathbf{Q}^{-1} (\bar{\mathbf{x}} - \mathbf{x}_0) + \mathbf{x}_0. \tag{17}$$

On the other hand,

$$\begin{aligned} E_x\{N_x(\bar{\mathbf{x}}, \mathbf{Q}) \,|\, \mathbf{x} \in M + \mathbf{x}_0\} &= E_x\{K_x \exp[\tfrac{1}{2}(\mathbf{x} - \bar{\mathbf{x}})^T \mathbf{Q}^{-1} (\mathbf{x} - \bar{\mathbf{x}})] \,|\, \mathbf{x} = \mathbf{G}\boldsymbol{\alpha} + \mathbf{x}_0\} \\ &= \mathbf{x}_0 + \mathbf{G} E_\alpha\{K_x \exp[\tfrac{1}{2}(\mathbf{G}\boldsymbol{\alpha} + \mathbf{x}_0 - \bar{\mathbf{x}})^T \mathbf{Q}^{-1} (\mathbf{G}\boldsymbol{\alpha} + \mathbf{x}_0 - \bar{\mathbf{x}})]\} \\ &= \mathbf{x}_0 + \mathbf{G} E_\alpha\{K_\alpha \exp[\tfrac{1}{2}(\boldsymbol{\alpha} - \boldsymbol{\alpha}^*)^T \mathbf{G}^T \mathbf{Q}^{-1} \mathbf{G}(\boldsymbol{\alpha} - \boldsymbol{\alpha}^*)]\} \\ &= \mathbf{x}_0 + \mathbf{G} E_\alpha\{N_\alpha(\boldsymbol{\alpha}^*, [\mathbf{G}^T \mathbf{Q}^{-1} \mathbf{G}])\} = \mathbf{x}_0 + \mathbf{G}\boldsymbol{\alpha}^* = \mathbf{y}^*. \end{aligned} \tag{18}$$ □

The above proposition implies that the location of the minimum of the quadratic form $(\mathbf{x} - \bar{\mathbf{x}})^T \mathbf{Q}^{-1} (\mathbf{x} - \bar{\mathbf{x}})$ over the linear manifold $M + \mathbf{x}_0$ can be obtained by assessing the expected value of a random vector $\mathbf{x}$ having distribution $N_x(\bar{\mathbf{x}}, \mathbf{Q})$, conditional on $\mathbf{x}$ being restricted to $M + \mathbf{x}_0$.

3. Constructing Conjugate Vectors

In view of Propositions 2.1 and 2.2, it is evident that one can construct a set of conjugate vectors (with respect to $\mathbf{Q}^{-1}$) by encoding a series of conditional means. To describe such a method, we consider a three-dimensional case, i.e., three dependent random variables x_1, x_2, x_3. It is assumed that these variables have a joint normal distribution with unknown covariance matrix $\mathbf{Q}$ and mean $\bar{\mathbf{x}}$. The method assumes the availability of the following information:

$\bar{x}_1, \bar{x}_2, \bar{x}_3 \equiv$ expected values of x_1, x_2, x_3;
$\bar{x}_{2|1}, \bar{x}_{3|1} \equiv$ expected values of x_2, x_3, given $x_1 = \bar{x}_1 + \Delta_1$;
$\bar{x}_{3|2} \equiv$ expected value of x_3, given $x_1 = \bar{x}_1 + \Delta_1$, $x_2 = \bar{x}_{2|1} + \Delta_2$.

According to Proposition 2.2, the vectors

$$\mathbf{v}_0 = (\bar{x}_1, \bar{x}_2, \bar{x}_3),$$

$$\mathbf{v}_1 = (\bar{x}_1 + \Delta_1, \bar{x}_{2|1}, \bar{x}_{3|1}),$$

$$\mathbf{v}_2 = (\bar{x}_1 + \Delta_1, \bar{x}_{2|1} + \Delta_2, \bar{x}_{3|2}),$$

$$\mathbf{v}_3 = (\bar{x}_1 + \Delta_1, \bar{x}_{2|1} + \Delta_1, \bar{x}_{3|2} + \Delta_3)$$

are the respective minimum points of $(\mathbf{x} - \bar{\mathbf{x}})\mathbf{Q}^{-1}(\mathbf{x} - \bar{\mathbf{x}})$ over the whole space R^3, the plane

$$M_1 = \{\mathbf{x} | x_1 = \bar{x}_1 + \Delta_1\},$$

the line

$$M_2 = \{\mathbf{x} | x_1 = \bar{x}_1 + \Delta_1, x_2 = \bar{x}_{2|1} + \Delta_2\},$$

and the point

$$M_3 = (\bar{x}_1 + \Delta_1, \bar{x}_{2|1} + \Delta_2, \bar{x}_{3|2} + \Delta_3).$$

Since $M_3 \subset M_2 \subset M_1 \subset R^3$, it follows from Proposition 2.1 that the vectors $\mathbf{v}_3 - \mathbf{v}_2$, $\mathbf{v}_2 - \mathbf{v}_1$, $\mathbf{v}_1 - \mathbf{v}_0$ are conjugate with respect to $\mathbf{Q}^{-1}$. Thus, we obtain the set of conjugate vectors

$$\mathbf{d}_1 = (\Delta_1, \bar{x}_{2|1} - \bar{x}_2, \bar{x}_{3|1} - \bar{x}_3),$$

$$\mathbf{d}_2 = (0, \Delta_2, \bar{x}_{3|2} - \bar{x}_{2|1}),$$

$$\mathbf{d}_3 = (0, 0, \Delta_3).$$

Figure 1 illustrates the above construction.

The following theorem formalizes the above procedure for n-dimensions.

Theorem 3.1. Let x be a random vector with joint probability density function $N_x(\bar{\mathbf{x}}, \mathbf{Q})$. Let $\bar{x}_{i|j}$ be the expected value of variable x_i, conditional on $x_k = \bar{x}_{k|k-1} + \Delta_k$, for $k = 1, \ldots, j$, where Δ_k are arbitrary nonzero constants and $\bar{x}_{i|0} = \bar{x}_i$.

Define $\mathbf{D}$ to be an $n \times n$ upper-triangular matrix whose elements d_{ij} are

$$d_{ij} = \begin{cases} 0, & \text{for } i < j, \\ \Delta_j, & \text{for } i = j, \\ x_{i|j} - x_{i|j-1}, & \text{for } i > j. \end{cases} \tag{19}$$

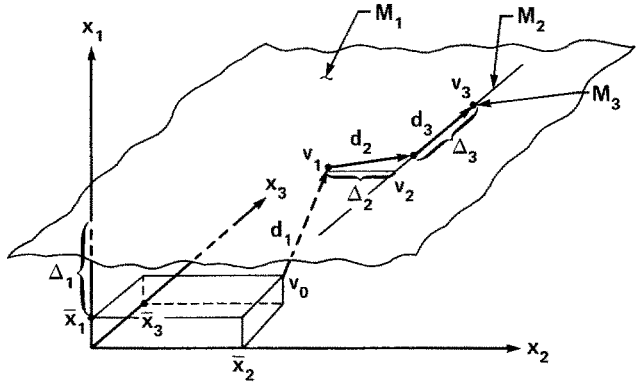

Fig. 1. Constructing conjugate directions in a three-dimensional case.

Then, the rows of $\mathbf{D}$, $\mathbf{d}_1, \ldots, \mathbf{d}_n$, are mutually conjugate with respect to $\mathbf{Q}^{-1}$; i.e.,

$$\mathbf{d}_i^T \mathbf{Q}^{-1} \mathbf{d}_j = 0, \qquad \text{for } i \neq j.$$

Proof. Let

$$\mathbf{v}_0 = (\bar{x}_1, \ldots, \bar{x}_n); \tag{20a}$$

and, for $i = 1, \ldots, n$, let

$$\mathbf{v}_i = (\bar{\mathbf{x}}_1 + \Delta_1, \bar{x}_{2|1} + \Delta_2, \ldots, \bar{x}_{i|i-1} + \Delta_i, \bar{x}_{i+1|i}, \bar{x}_{i+2|i}, \ldots, \bar{x}_{n|i}). \tag{20b}$$

Then,

$$\mathbf{d}_i = \mathbf{v}_i - \mathbf{v}_{i-1}, \qquad \text{for } i = 1, \ldots, n. \tag{21}$$

But,

$$\mathbf{v}_i = E_x\{N_x(\bar{\mathbf{x}}, Q) | \mathbf{x} \in M_i\}, \tag{22}$$

where

$$M_0 = R^n,$$

$$M_i = \{\mathbf{x} | x_k = \bar{x}_{k|k-1} + \Delta_k, k = 1, \ldots, i\}, \qquad \text{for } i = 1, \ldots, n.$$

By Proposition 2.2, we have

$$\mathbf{v}_i = \arg \min_{\mathbf{x} \in M_i} (\mathbf{x} - \bar{\mathbf{x}}) Q^{-1} (\mathbf{x} - \bar{\mathbf{x}}). \tag{23}$$

But,

$$M_0 \supset M_1 \supset M_2 \supset \cdots M_n.$$

Thus, the sequence $\{M_i - \mathbf{v}_n\}$ is a sequence of nested subspaces in R^n, hence, by Proposition 2.1, we have

$$\mathbf{d}_i^T \mathbf{Q}^{-1} \mathbf{y} = (\mathbf{v}_i - \mathbf{v}_{i-1})^T \mathbf{Q}^{-1} \mathbf{y} = 0, \qquad \forall \mathbf{y} \in M_i - \mathbf{v}_n,\ i = 1, \ldots, n. \tag{24}$$

However,

$$\mathbf{v}_j \in M_i, \qquad \text{for } j \geq i.$$

Thus,

$$\mathbf{d}_j = \mathbf{v}_j - \mathbf{v}_{j-1} \in M_i, \qquad \text{for } n \geq j \geq i+1, \tag{25}$$

so

$$\mathbf{d}_i^T \mathbf{Q}^{-1} \mathbf{d}_j = 0, \qquad \text{for } n \geq j \geq i+1 \text{ and } i = 1, \ldots, n-1. \quad \square \tag{26}$$

4. Assessment Procedure

The results presented in the previous sections are summarized in the following procedure for assessing the parameters of an n-dimensional normal density function.

Step (a). For each variable x_i, obtain the mean $\bar{x}_i$ and standard deviation σ_i of its marginal distribution. These values may be assessed directly or through any technique for assessing single-variable density functions.

Step (b). Set $\bar{x}_{i|0} = \bar{x}_i$, for $i = 1, \ldots, n$. Then, for $j = 1, \ldots, n$, obtain the conditional expectations $\bar{x}_{j+1|j}, \bar{x}_{j+2|j}, \ldots, \bar{x}_{n|j}$ of the last $n-j$ variables, given the values of the first j variables $x_i = \bar{x}_{i|i-1} + \Delta_i$, $i = 1, \ldots, j$, with arbitrary nonzero increments Δ_i.

Step (c). Obtain the matrix

$$\mathbf{D} = \begin{bmatrix} \Delta_1 & \bar{x}_{2|1} - \bar{x}_1 & \bar{x}_{3|1} - \bar{x}_3 & \cdots & \bar{x}_{j|1} - \bar{x}_j & \cdots & \bar{x}_{n|1} - \bar{x}_n \\ & \Delta_2 & \bar{x}_{3|2} - \bar{x}_{3|1} & \cdots & \bar{x}_{j|2} - \bar{x}_{j|1} & \cdots & \bar{x}_{n|2} - \bar{x}_{n|1} \\ & & \Delta_3 & \vdots & & \vdots & \\ & & & \Delta_i & \cdots \quad \bar{x}_{j|i} - \bar{x}_{j|i-1} & \cdots & \bar{x}_{n|i} - \bar{x}_{n|i-1} \\ & & & & \vdots & & \vdots \\ & & & & \Delta_j & & \Delta_n \end{bmatrix} \tag{27}$$

Step (*d*). Obtain $\lambda_1, \ldots, \lambda_n$ by solving the system

$$\begin{bmatrix} d_{11}^2 & & & \\ d_{12}^2 & d_{22}^2 & & \\ \vdots & \vdots & \ddots & \\ d_{1n}^2 & d_{2n}^2 & \cdots & d_{nn}^2 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{bmatrix} = \begin{bmatrix} \sigma_1^2 \\ \sigma_2^2 \\ \vdots \\ \sigma_n^2 \end{bmatrix}, \tag{28}$$

where d_{ij} is the ijth element of the matrix $\mathbf{D}$.

Step (*e*). Obtain the covariance matrix

$$\mathbf{Q} = \mathbf{D}^T \begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \\ & & & \lambda_n \end{bmatrix} \boldsymbol{D}. \tag{29}$$

As pointed out earlier, the matrix $\mathbf{Q}$ given by (29) is positive semidefinite if the values $\lambda_1, \lambda_2, \ldots, \lambda_n$ are nonnegative. Obtaining a negative λ_i in solving (28) would indicate that some assessments are inconsistent. More specifically, Eq. (28) suggests that a negative λ_i may result from some right-hand component σ_i^2 being too small, relative to the terms σ_j^2 and d_{ji}^2, for $j = 1, \ldots, i-1$. This may be interpreted as an indication that the assessed variance of the variable x_i is too small as compared to the assessed variances of the other variables and their influence on x_i. Thus, the inconsistency of such an assessment stems from the fact that the uncertainty in x_i due to the influence of the other uncertain variables exceeds the total uncertainty in x_i as reflected by the assessed variance σ_i^2. Clearly, such inconsistencies may be easily circumvented by altering the relative magnitude of the appropriate variances on the right-hand side of (28). The simple form of (28) enables the analyst to easily verify whether the adjustment resolves the inconsistency (i.e., yields a nonnegative $\boldsymbol{\Lambda}$) and guide the assessor toward a consistent assessment of the variances. Resolving inconsistencies in this manner is based on the premise that the conditional-means assessment (from which the d_{ij} components are derived) are correct. Adjusting the components of the $\mathbf{D}$-matrix might be quite difficult, since the rows of $\mathbf{D}$ must remain conjugate. To avoid such adjustments, one should attempt to obtain more robust estimates of $\mathbf{D}$ by employing redundant assessments of the conditional expectations $\bar{x}_{i|j}$. In particular, we can use several values of Δ_j in Step (b) of the procedure and, for each j, encode the corresponding assessments of $\bar{x}_{i|j}$, with $i = j+1, \ldots, n$. Exploiting the theoretical linearity of $\bar{x}_{i|j}$ in Δ_j, we then use least squares to fit a straight line through the encoded points and obtain representative values of $\bar{x}_{i|j}$ and Δ_j on that line. Specifically, let us use Δ_j^k, $k = 1, \ldots, r$, and denote the corresponding

assessments by $\bar{x}_{i|j}^k$. The line minimizing the square errors and satisfying

$$\bar{x}_{i|j} = \bar{x}_{i|j-1}, \qquad \text{for } \Delta_j = 0,$$

is given by

$$\bar{x}_{i|j} = \bar{x}_{i|j-1} + \Delta_j \left\{ \sum_{k=1}^{r} [(\bar{x}_{i|j}^k - \bar{x}_{i|j-1})] \Delta_j^k \Big/ \sum_{k=1}^{r} [(\Delta_j^k)^2] \right\}. \tag{30}$$

Since we need only one point on this line, it is convenient to choose

$$\Delta_j = \sum_{k=1}^{r} [(\Delta_j^k)^2] \Big/ \sum_{k=1}^{r} (\Delta_j^k). \tag{31}$$

Then,

$$\bar{x}_{i|j} = \sum_{k=1}^{r} (\Delta_j^k \bar{x}_{i|j}^k) \Big/ \sum_{k=1}^{r} (\Delta_j^k). \tag{32}$$

5. Conclusions

In the preceding sections, we have developed a procedure for assessing a joint normal probability density function over n variables. By exploiting the properties of the normal distribution, we have reduced the task to that of specifying the n marginal means and variances and $n(n-1)/2$ conditional means. The assessment of the conditional means is done sequentially in groups where each subsequent group involves the fixing of one additional variable. In the context of a dynamic process, where the random variables are successive values of a time-dependent entity, the procedure can be interpreted as successively revealing subsequent values of the entity and assessing its expected values at the remaining time periods conditional on the revealed values.

An important issue in deriving an assessment procedure is the computational requirements. In view of the properties of the normal distribution, one could obtain a general set of linear equations relating the parameters of the normal distribution with arbitrary conditional means. Extracting these parameters from such a system may require, however, the solution of a system of $n(n+1)/2$ equations. In our procedure, the only system of equations that has to be solved is the one in Step (d). That system is solvable by forward substitution, since it is in lower-triangular form. It should also be pointed out that the covariance matrix obtained in Step (e) is already in decomposed form, which enables easy computation of $\mathbf{Q}^{-1}$, often required in various applications. Furthermore, as mentioned earlier, obtaining $\mathbf{Q}$ in this factored form facilities the detection and resolution of inconsistencies in the assessments.

The proposed procedure enables efficient encoding of a joint probability density function whenever there is reason to believe that it might be adequately approximated by a normal distribution. This procedure may also be used to obtain *local approximations* (around the mean) to the covariance matrix of more general distributions for use in proximal decision analysis. Such analysis is aimed at analyzing the effect of uncertain, but small, changes in the variables under consideration. It seems, therefore, more appropriate to employ in such analysis covariance information based on a local approximation to the joint density function, rather than the actual covariance matrix. Such an approach will exclude from the proximal analysis the effects of dependence among the variables outside the domain of interest.

References

1. SPETZLER, C. S., and STAEL VON HOLSTEIN, C. S., *Probability Encoding in Decision Analysis*, Management Science, Vol. 22, No. 3, 1975.
2. HOWARD, A. R., *Proximal Decision Analysis*, Management Science, Vol. 17, No. 9, 1971.
3. BUEDE, D., *The Assessment of a Joint Probability Distribution on an Evolving Dynamic Uncertain Process*, Stanford University, PhD Thesis, 1977.
4. LUENBERGER, D. G., *Introduction to Linear and Nonlinear Programming*, Addison-Wesley Publishing Company, Reading, Massachusetts, 1973.