A Factor-Analysis Method for Diagnosing Variability in Multivariate Manufacturing Processes

Daniel W. Apley

Department of Industrial Engineering
Texas A&M University
College Station, TX 77843-3131
(*apley@tamu.edu*)

Jianjun Shi

Department of Industrial and Operations Engineering
The University of Michigan
Ann Arbor, MI 48109-2117
(*shihang@engin.umich.edu*)

In many modern manufacturing processes, large quantities of multivariate process-measurement data are available through automated in-process sensing. This article presents a statistical technique for extracting and interpreting information from the data for the purpose of diagnosing root causes of process variability. The method is related to principal components analysis and factor analysis but makes more explicit use of a model describing the relationship between process faults and process variability. Statistical properties of the diagnostic method are discussed, and illustrative examples from autobody assembly are provided.

KEY WORDS: Fault diagnosis; Multivariate analysis; Principal components analysis; Quality control; Statistical process control.

Recent years have witnessed marked improvements in automated in-process sensing and data-collection technology. Such technology is widely implemented in manufacturing and often provides high-throughput measurement—100% inspection for discrete parts processes or high sampling rate for continuous flow processes—of a large number of product/process features. The resulting volume of multivariate data creates tremendous potential for monitoring and diagnosing quality-related problems in the manufacturing process.

Consider the autobody assembly process, which provides a prime example of the use of multivariate in-process sensing for quality improvement. Laser measurement stations built into the assembly line are now widespread. Figure 1 illustrates one such measurement station. At this stage, the autobody is referred to as the Body-in-White (BIW) and consists mainly of the bodysides joined to the underbody and roof. Each laser sensor is set up to measure a key point or feature on the autobody and can measure one-, two-, or three-dimensional coordinates of the point. To measure a point, the laser beam is fanned into a plane and projected onto the autobody. A charge coupled device (CCD) array built into the laser sensor then records an image of the laser on the autobody, and, from rudimentary image processing and triangulation principles, the dimensional coordinates can be extracted. Since the measurement station is built into the assembly line, every autobody is automatically measured. There can be multiple measurement stations per line, typically located after major subassembly stations, and as many as 150 measured coordinates per station. Consequently, a tremendous amount of multivariate data is available for monitoring and diagnosing root causes of dimensional variability.

The BIW is constructed from a large number of stamped sheet-metal panels. Usually one or two panels at a time are added to the subassemblies until major subassemblies (e.g., complete bodysides or the underbody) are formed. The BIW is formed by joining the bodysides to the underbody in a framing station and then adding the roof and reinforcement spot welds. Thousands of tooling elements are used throughout the process to join the components and subassemblies. The major tooling elements are fixture components (i.e., clamps, blocks, and pins, for locating parts) and robotic resistance spot-welding guns. When tooling elements fail, the dimensional integrity of the BIW can be severely affected. The most commonly encountered tooling faults are loose, worn, or broken locating elements; misaligned weld guns; and malfunctioning clamps.

A typical BIW has 150 to 200 sheet-metal components, and the assembly process may involve over 60 assembly stations, 2,000 locating elements, and 4,000 spot welds. A failure in any tooling element may be a root cause of dimensional variability, and several different tooling faults may be present simultaneously. Following the detection of a fault (or faults) through effective statistical process monitoring, the goal is to identify and eliminate the root cause(s) and bring the process back in control as soon as possible. The complexity of such production processes presents challenges for root-cause diagnosis, which is the focus of this article. We present a technique for extracting diagnostic information from the measurement data to facilitate identification and elimination of root causes of process variability.

Although we illustrate with examples from autobody assembly, similar scenarios exist in many other manufacturing industries. For example, laser-optical, X-ray, and vision systems are commonly used in printed-circuit assembly to measure the dimensional integrity of wet solder paste, solder joints, and chip features. In more traditional machined-parts manufacturing, machine-tool probing attachments for automatically

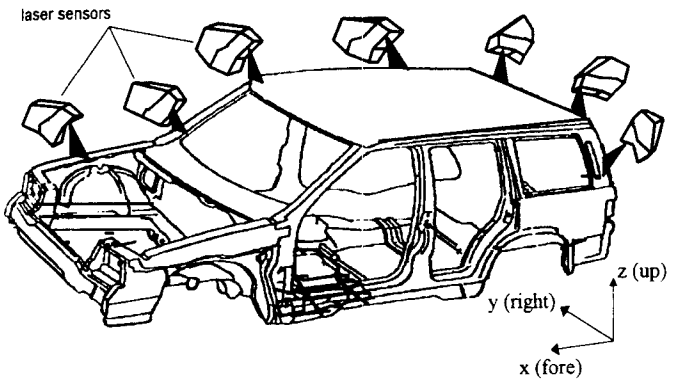

Figure 1. Laser Sensing at the BIW Stage in Autobody Assembly.

inspecting machined parts are increasing in popularity. In both of these situations, process data and multiple potential causes of process variability are abundant.

Most of the existing multivariate statistical process control (SPC) research (see Lowry and Montgomery 1995 and Mason, Champ, Tracy, Wierda, and Young 1997 for reviews) relates to process monitoring with limited diagnostic capabilities. A number of well-researched multivariate statistical analysis techniques, in particular principal components analysis (PCA) and factor analysis (Jackson 1980), relate more closely to the diagnostic problem this article addresses. Hu and Wu (1992) and Ceglarek and Shi (1996) applied PCA to variation diagnosis in autobody assembly. PCA works well when a single fault is present but yields results that may be noninterpretable when multiple faults are present (see Sec. 2). Since multiple faults are a common occurrence in many processes, including autobody assembly, PCA has severe limitations. Factor-rotation techniques (e.g., see Jackson 1981; Johnson and Wichern 1998) attempt to improve "interpretability" with multiple factors (i.e., multiple faults) by forming orthogonal linear combinations of the factors obtained from PCA. The method presented in this article bears resemblance to factor rotation. In contrast to existing methods, which use somewhat artificially contrived measures of interpretability [e.g., the varimax criterion (Johnson and Wichern 1998)], the proposed method uses a more physically meaningful interpretability criterion that produces results that relate more directly to the actual faults. This representation facilitates the ultimate goal of monitoring and diagnosis—to track down and eliminate root causes of variation.

There has been significant prior work on analysis of dimensional coordinate data. Dowling, Griffin, Tsui, and Zhou (1997), and the discussion that followed provided an excellent overview. Most of the existing research (e.g., Etesami 1988; Hulting 1992; Wang, Gupta, Hulting, and Fussell 1998) focuses on modeling geometric surface features of manufactured parts and fitting parameterized models to the measurement data. The model is fit to single parts, separately, and the model parameters are used to evaluate the conformance to tolerance of the geometric features for each part. The method presented in this article has significantly different emphasis—data from a *sequence* of parts are used collectively to analyze part-to-part variation patterns and to subsequently diagnose root causes of the variation.

The remainder of the article is organized as follows. Section 1 describes the model used to represent the effects of process faults on variability. Section 2 discusses the capabilities and limitations of PCA and factor rotation for process diagnosis. Section 3 discusses methods for estimating the number of faults that are simultaneously present. Section 4 presents the main results of the article, a technique for estimating the effect of each individual fault when multiple faults are present, and an example. In Section 5, the statistical properties of the estimates are discussed.

1. A FACTOR-ANALYSIS MODEL FOR PROCESS VARIABILITY

One of the keys to diagnosing variability faults is the incorporation of a suitable fault model. This section presents the model structure that is assumed throughout this article and illustrates with an example from autobody assembly.

Let $\mathbf{x} = [x_1, x_2, \ldots, x_n]^T$ be an $n \times 1$ random vector that represents a set of n measured features from the product or process. Let $\mathbf{x}_i$, $i = 1, 2, \ldots, N$, be a random sample of N observations of $\mathbf{x}$. It is assumed that $\mathbf{x}$ obeys the model

$$\mathbf{x} = \mathbf{C}\mathbf{v} + \mathbf{w}, \tag{1}$$

where $\mathbf{C} = [\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_p]$ is an $n \times p$ constant matrix, $\mathbf{v} = [v_1, v_2, \ldots, v_p]^T$ is a $p \times 1$ zero-mean random vector with covariance matrix $\Sigma_v = \mathbf{I}$, and $\mathbf{w}$ is an $n \times 1$ zero-mean random vector that is independent of $\mathbf{v}$ and has covariance matrix $\Sigma_w = \sigma^2\mathbf{I}$. $\mathbf{I}$ denotes the identity matrix of appropriate dimension.

The interpretation of the model is that there are p separate uncorrelated faults that affect the measurement vector $\mathbf{x}$. Each fault has a linear effect on $\mathbf{x}$ that is dictated by the corresponding column of $\mathbf{C}$. Together, $\mathbf{c}_i v_i$ describes the effect of the ith fault on $\mathbf{x}$. The direction of $\mathbf{c}_i$ indicates the nature of the variation pattern caused by the ith fault. Specifically, it indicates how the fault causes the different measured features to vary with respect to each other. $\mathbf{c}_i$ will be referred to as the "fault geometry" vector. Since the elements of $\mathbf{v}$ are scaled to have unit variance, $\mathbf{c}_i$ also indicates the magnitude or severity of the ith fault. $\mathbf{w}$ represents the aggregated effects of measurement noise and any inherent unmodeled variation in the manufacturing process. It is assumed that $p \leq n$ and that $\mathbf{C}$ has full rank p. The focus of this article is on faults that act as sources of variation, as opposed to mean shifts. All random variables are assumed to be zero mean. If not, the mean of $\mathbf{x}$ should be estimated and subtracted from the data.

The objective is to estimate the number of faults p that are contributing to process variability, as well as each of the p fault geometry vectors in $\mathbf{C}$, using the set of multivariate observations of $\mathbf{x}$. The presumption is that, if $\mathbf{C}$ can be accurately estimated, it may be used to diagnose each of the p faults and aid in the identification and elimination of their root causes.

The model and the objective of estimating $\mathbf{C}$ are similar to the standard linear orthogonal factor-analysis problem (Johnson and Wichern 1998), in which the elements of $\mathbf{v}$ are referred to as the factors, $\mathbf{C}$ is referred to as the factor loading

matrix, and $\Sigma_{u'}$ is typically only assumed to be diagonal. For the method presented in this article to produce estimates of the fault geometry vectors that have the desired interpretability, it is necessary to assume $\Sigma_{u'} = \sigma^2\mathbf{I}$. This can be assumed without loss of generality if $\Sigma_{u'}$ is known up to a constant scalar multiple. Although this assumption may not be reasonable for typical applications of factor analysis, it is often appropriate for SPC of manufacturing processes, because essentially it is equivalent to assuming that a sufficient quantity of data has been collected when the process was known to be in control (i.e., when there were no faults present so that $\mathbf{x} = \mathbf{w}$). In this event, the sample covariance matrix of the in-control data would provide an estimate of $\Sigma_{u'}$. Before applying the method of this article, the data would first be transformed via $\Sigma_{u'}^{-1/2}\mathbf{x}$, where $\Sigma_{u'}^{-1/2}$ is any positive definite square root of $\Sigma_{u'}$. The noise covariance matrix for the transformed data would then be the identity matrix. After estimating $\mathbf{C}$ for the transformed data, the estimates should be transformed back by premultiplying by $\Sigma_{u'}^{1/2}$ before interpreting the results.

Although factor analysis is widely used in the social sciences and other fields, it is seldom used as a tool for diagnosing variability in manufacturing processes. The applicability of the model is illustrated with the following example from autobody assembly. Figure 2 shows the layout of 26 measurement points (labeled 1–26 in the figure) taken at the BIW stage of the assembly process. All three coordinates ($\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$ in the figure) are measured for each point, except for points 10 and 23, for which only the $\mathbf{x}$ and $\mathbf{z}$ coordinates are measured. Thus, the measurement vector $\mathbf{x}$ has a total of $n = 76$ elements. The cross-member, roof header, and plenum (in addition to the roof, underbody, cowl, and additional roof bows, which have all been omitted from the figure for clarity) join the left and right bodysides to form the BIW.

Before the bodysides and connecting members are welded together, they must be accurately located with respect to one another with fixtures and then clamped into place. The bodysides are positioned in the x–z plane with pins rigidly attached to the fixtures. Each pin mates with either a hole or slot in the autobody panel so that the panel position is constrained but not overly constrained. Through repeated use (a thousand panels per day may be placed into each fixture), the pins frequently become worn or loose. In this event, the panel will no longer be constrained to lie in its proper location when it is placed into the fixture. For example, suppose the pin that constrains the right bodyside in the x direction becomes loose. When a bodyside is placed into the fixture, it may be positioned too far forward (in the positive x direction) by, say, one millimeter. When it is subsequently clamped into place and welded to the rest of the BIW, it will retain the incorrect position. For the next autobody, the right bodyside may be positioned too far aftward (in the negative x direction). From autobody to autobody, the loose pin will cause a distinct variation pattern in the BIW dimensions. The elements of $\mathbf{x}$ that represent measurements in the x direction on the right bodyside will reflect this pattern when the BIW is measured. If we refer to this loose pin as fault 1, v_1 is the random variable representing by how much the right bodyside translates in the x direction for each autobody (scaled to have unit variance). All elements of $\mathbf{c}_1$ that do not correspond to x direction measurements on the right bodyside would be 0. All elements that do correspond to x direction measurements on the right bodyside would be the same constant value, equal to the standard deviation of the bodyside translations.

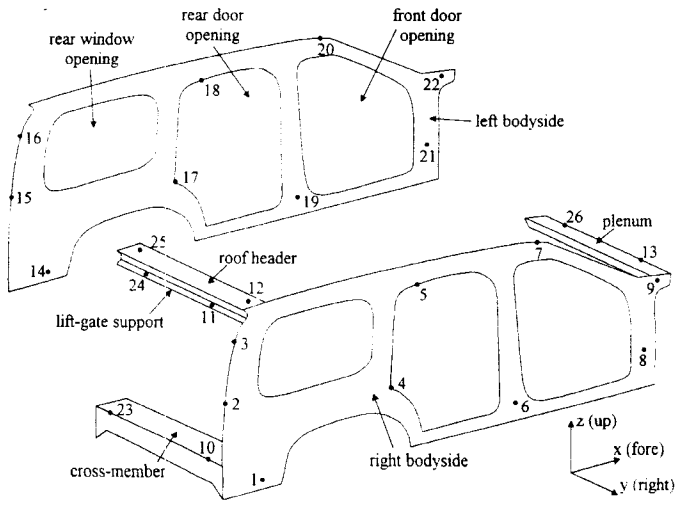

Figure 2. Measurement Layout at the BIW Stage.

Likewise, suppose a second pin became loose or worn and caused the right bodyside to rotate in the x–z plane before being welded to the BIW. The second fault would also cause a distinct variation pattern in the BIW measurements that is well represented by Model (1). The $\mathbf{c}_2$ vector would be determined by the measurement layout and the geometry of the panel and fixture. Apley and Shi (1998) provided details on how to analytically model the effects of tooling faults and a justification for the linear structure of Model (1). Technically, the rotation of a panel is nonlinear in the parameters but is closely approximated as linear when small angles of rotation are involved. The linear model structure is quite versatile and provides a good representation of a variety of commonly encountered faults, including those that are not rigid body translations and rotations. This includes variability introduced by stamping, welding, and material-handling faults, in which case Model (1) can be viewed as a linearization of a more exact nonlinear model (Apley and Shi 1998).

We point out that it is not necessary to analytically model the faults to use the method presented in this article. This is the main distinction between this work and that of Apley and Shi (1998), which assumed that an exhaustive set of potential faults can be analytically modeled off-line to obtain the $\mathbf{C}$ matrix. The problem then reduces to one of fault classification, where, based on the on-line measurements, one seeks to identify which of the modeled faults are present. The method of this article attempts to estimate $\mathbf{C}$ directly from the data, with no a priori knowledge of the faults, and use the results to gain insight into the root causes of the variability.

To uniquely estimate $\mathbf{C}$ with no a priori knowledge of the faults, an additional assumption must be made regarding the structure of $\mathbf{C}$. The reason this assumption is necessary is discussed in Section 2. It is assumed that $\mathbf{C}$ has the ragged lower

riangular form

$$\mathbf{C} = \begin{bmatrix} \mathbf{c}_{1,1} & & & & \\ \mathbf{c}_{2,1} & \mathbf{c}_{2,2} & & & \\ \mathbf{c}_{3,1} & \mathbf{c}_{3,2} & \mathbf{c}_{3,3} & & \\ \vdots & \vdots & \vdots & \ddots & \\ \mathbf{c}_{p,1} & \mathbf{c}_{p,2} & \mathbf{c}_{p,3} & \cdots & \mathbf{c}_{p,p} \end{bmatrix}. \tag{2}$$

where $\mathbf{c}_{i,j}$ is an $n_i \times 1$ vector with $n_i \geq 2$ and $\sum_{i=1}^{p} n_i = n$. The interpretation of this structure is that there exists a subgroup of n_1 measurements $\{x_1, x_2, \ldots, x_{n_1}\}$ that are affected by only one fault and not by the remaining $p-1$ faults. Furthermore, there must exist a second subgroup of n_2 measurements $\{x_{n_1+1}, x_{n_1+2}, \ldots, x_{n_1+n_2}\}$ that are affected by only one of the *remaining* $p-1$ faults. Note that these measurements may also be affected by the first fault. There must also exist a third subgroup of measurements affected by only one of the remaining $p-2$ faults, and so on. In the previous example, $\{x_1, x_2, \ldots, x_{n_1}\}$ could be taken to be any set of z-direction measurements on the right bodyside. Such a group would be affected by fault 2, which causes the right bodyside to rotate, but not by fault 1, which causes only an x-direction translation of the bodyside and does not affect the z-direction measurements. Upon appropriate reordering of the measurements and faults, $\mathbf{C}$ has the structure in (2). If the faults are such that $\mathbf{C}$ does not have the assumed structure (2), the method of this article cannot be applied. When n is large relative to p, this structure is often satisfied in autobody assembly. An additional example in which this is the case will be given in Section 4.3.

2. LIMITATIONS OF PCA AND FACTOR ROTATION

From the model structure and assumptions, the covariance matrix of $\mathbf{x}$ is

$$\Sigma_x = E[(\mathbf{Cv}+\mathbf{w})(\mathbf{Cv}+\mathbf{w})'] = \mathbf{CC}' + \sigma^2\mathbf{I}. \tag{3}$$

Let $\mathbf{z}_i, i = 1, 2, \ldots, n$, denote an orthogonal set of unit norm eigenvectors of Σ_x. Let λ_i, $i = 1, 2, \ldots, n$, denote the corresponding eigenvalues, arranged in descending order. It follows from (3) that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p > \sigma^2 = \lambda_{p+1} = \lambda_{p+2} = \cdots = \lambda_n$, and $\operatorname{span}\{\mathbf{z}_i\}_{i=1}^{p} = \operatorname{span}\{\mathbf{c}_i\}_{i=1}^{p}$.

PCA can also be used to decompose Σ_x in terms of its eigenvectors and eigenvalues as

$$\begin{aligned} \Sigma_x &= \sum_{i=1}^{n} \lambda_i \mathbf{z}_i \mathbf{z}_i' = \sum_{i=1}^{p} (\lambda_i - \sigma^2)\mathbf{z}_i\mathbf{z}_i' + \sigma^2 \sum_{i=1}^{n} \mathbf{z}_i\mathbf{z}_i' \\ &= \mathbf{Z}_p[\Lambda_p - \sigma^2\mathbf{I}]\mathbf{Z}_p' + \sigma^2\mathbf{I}, \end{aligned} \tag{4}$$

where $\mathbf{Z}_p = [\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_p]$, and $\Lambda_p = \operatorname{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_p\}$. Comparing (3) and (4), it is obvious that one possible estimate of $\mathbf{C}$ is $\mathbf{Z}_p[\Lambda_p - \sigma^2\mathbf{I}]^{1/2}$. Other possible estimates that would preserve the covariance structure are $\mathbf{Z}_p[\Lambda_p - \sigma^2\mathbf{I}]^{1/2}\mathbf{Q}$, where $\mathbf{Q}$ is any $p \times p$ orthogonal matrix.

When there is a single fault present, PCA is an effective tool for diagnosing process variability. Σ_x has a single dominant eigenvalue and a unique estimate of $\mathbf{C}(=\mathbf{c}_1)$ is $\mathbf{z}_1[\lambda_1 - \sigma^2]^{1/2}$. σ^2 can be taken to be any of the smallest $n-1$ eigenvalues. In practice, one must work with the sample covariance matrix. In this case it may not be clear how many eigenvalues are "dominant," and thus how many faults are present. Methods for estimating the number of faults from the sample covariance matrix are discussed in Section 3.

When multiple faults are present ($p > 1$), the fault geometry vectors will not correspond one-to-one with the eigenvectors unless the fault geometry vectors happen to be orthogonal. PCA does provide an estimate of the number of faults that are present (refer to Sec. 3), but estimation of $\mathbf{C}$ is less straightforward. As discussed previously the estimate is not unique. Treating the eigenvectors as (scaled) estimates of the fault geometry vectors may yield noninterpretable results and provide little useful diagnostic information. To improve interpretability, a commonly used procedure is factor rotation (Jackson 1981; Johnson and Wichern 1998). The standard factor-rotation problem is to find the $p \times p$ orthogonal matrix $\mathbf{Q}$ so that the resulting estimate of C (typically of the form $\mathbf{Z}_p\Lambda_p^{1/2}\mathbf{Q}$) provides the clearest interpretability. What is meant by "interpretability" is subjective, but the most widely used criterion is the *varimax rotation* (Jackson 1981; Johnson and Wichern 1998). The varimax criterion for best interpretability is that each column of $\mathbf{Z}_p\Lambda_p^{1/2}\mathbf{Q}$ should consist of elements that are either very large in magnitude or very small in magnitude with as few moderate-sized elements as possible.

For diagnosis of manufacturing variability, there is little justification for the varimax criterion. The $\mathbf{Q}$ matrix that results in the clearest interpretability would be such that $\mathbf{Z}_p[\Lambda_p - \sigma^2\mathbf{I}]^{1/2}\mathbf{Q} = \mathbf{C}$, whose columns are the physical fault geometry vectors themselves. That is, knowing the fault geometry vectors for the actual faults will surely provide the most effective root-cause diagnosis. As shown in Section 4, $\mathbf{C}$ may be uniquely identified if it has the structure assumed in (2).

3. ESTIMATING THE NUMBER OF FAULTS

The method for estimating $\mathbf{C}$ that will be presented in Section 4 requires an estimate of the number of faults present. If p faults are present, $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p > \sigma^2 = \lambda_{p+1} = \lambda_{p+2} = \cdots = \lambda_n$. Thus, a natural means of estimating p is to look at the eigenvalues $\{\hat{\lambda}_i\}_{i=1}^{n}$ of the sample covariance matrix $\hat{\Sigma}_x$. Throughout, the symbol "$\hat{\ }$" will be used to denote an estimate of a parameter. A number of methods for estimating p have been suggested, the most popular of which can be classified as either maximum likelihood based or information based. Strictly speaking, these methods require that the data are multivariate Gaussian.

Anderson (1963) developed results for the asymptotic distribution of eigenvectors and eigenvalues of a sample covariance matrix for Gaussian data, which yields an asymptotically valid (for large enough N) likelihood ratio test of the null hypothesis that $\lambda_{m+1} = \lambda_{m+2} = \cdots = \lambda_n$, for some fixed m, versus the alternative hypothesis that not all of the $n-m$ smallest eigenvalues are equal. Note that when $m \geq p$ the null hypothesis holds. For $m = 0, 1, \ldots, n-1$, one calculates the test statistics

$$\Lambda(m) = N(n-m)\log\left(\frac{a_m}{g_m}\right),$$

where a_m and g_m are the arithmetic and geometric means, respectively, of the $n-m$ smallest eigenvalues of $\hat{\Sigma}_x$. Under the null hypothesis, $\Lambda(m)$ is asymptotically chi-squared distributed with $(n-m)(n-m+1)/2 - 1$ df. A set of thresholds $\eta(m)$, $m = 0, 1, \ldots, n-1$, are specified, typically based

on the null distribution. The suggested procedure for estimating p is, increasing m sequentially from 0, to choose $\hat{p}$ to be the first m for which $\Lambda(m) < \eta(m)$.

To improve the chi-squared approximation to the null distribution for finite N, Lawley (1956) introduced the modified statistic

$$\Lambda_{\text{mod}}(m) = \left[1 - \frac{m}{N} - \frac{2(n-m)^2 + (n-m) + 2}{6N(n-m)} + \frac{1}{N}\sum_{i=1}^{m}\left(\frac{a_m}{\hat{\lambda}_i - a_m}\right)^2\right]\Lambda(m),$$

which has the same asymptotic chi-squared null distribution with $(n-m)(n-m+1)/2-1$ df. The procedure for estimating p is the same.

Alternative procedures for estimating p are based on the Akaike (AIC) and minimum description length (MDL) information criteria introduced by Akaike (1971a), Schwartz (1978), and Rissanen (1978). For the problem of estimating the number of significant factors in PCA, the AIC was applied by Akaike (1971b, 1987) and Wax and Kailath (1985) and the MDL criterion by Wax and Kailath (1985). The AIC and MDL tests require calculation of, for $m = 0, 1, \ldots, n-1$,

$$\text{AIC}(m) = N(n-m)\log\left(\frac{a_m}{g_m}\right) + m(2n-m),$$

and

$$\text{MDL}(m) = N(n-m)\log\left(\frac{a_m}{g_m}\right) + m(2n-m)\log(N)/2.$$

Using the AIC or MDL criterion, $\hat{p}$ is chosen to be the m that minimizes AIC(m) or MDL(m), respectively.

One major advantage of using either the AIC test or the MDL test is simplicity. For the likelihood ratio tests one must select, somewhat arbitrarily, the set of thresholds. The information-based tests do not require this.

Certain properties of the various methods of estimating p are well known. The likelihood ratio methods often lead to overestimating p and choosing a higher number of factors than can be interpreted (Basilevsky 1994). The MDL and AIC criteria are less dependent on the validity of the asymptotic chi-squared approximation. On the other hand, the AIC estimate is not consistent (in the sense of yielding the true number of faults with probability 1 as N approaches infinity), whereas the other three methods do produce consistent estimates (Wax and Kailath 1985). The AIC method slightly overestimates p asymptotically.

For some problems of typical scale (i.e., typical n, N, and p) in autobody assembly, we present a Monte Carlo comparison of the different methods. Ten thousand Monte Carlo trials were used for all simulations, and the generated data were Gaussian. Note that, as shown in Appendix A, the distribution of each of the four test statistics depends on only n, N, and $\{\lambda_i/\sigma^2\}_{i=1}^{p}$ and not on the individual eigenvectors. Tables 1 through 3 show the probability mass functions of $\hat{p}$ for various cases when at least one fault is present, using AIC, MDL, and Λ_{mod}, respectively. The values are rounded to three decimal places. The results using the unmodified likelihood ratio statistic Λ are not shown. Except for the case in which $n = 40$ and $N =$

Table 1. Probability Mass Function of $\hat{p}$ From AIC for Various n, N, and $\{\lambda_i/\sigma^2\}_{i=1}^{p}$

			$p_j \equiv$ probability$\{\hat{p} = j\}$						
n	N	$\{\lambda_i/\sigma^2\}_{i=1}^{p}$	p_0	p_1	p_2	p_3	p_4	p_5	p_6
40	50	{11,11,11}	0	0	0	**.951**	.046	.003	0
40	50	{5,5,5}	0	.002	.067	**.892**	.037	.002	0
40	50	{3,3,3}	.084	.334	.394	**.18**	.008	0	0
40	50	{11}	0	**.975**	.024	.001	0	0	0
40	50	{5}	.006	**.973**	.02	.001	0	0	0
40	50	{3}	.365	**.621**	.014	0	0	0	0
40	100	{11,11,11}	0	0	0	**.941**	.058	.001	0
40	100	{5,5,5}	0	0	0	**.947**	.052	.001	0
40	100	{2,2,2}	.294	.446	.22	**.039**	.001	0	0
40	100	{11}	0	**.962**	.038	0	0	0	0
40	100	{5}	0	**.961**	.039	0	0	0	0
40	100	{2}	.608	**.379**	.013	0	0	0	0
40	500	{11,11,11}	0	0	0	**.906**	.091	.003	0
40	500	{3,3,3}	0	0	0	**.914**	.083	.003	0
40	500	{2,2,2}	0	0	0	**.927**	.071	.002	0
40	500	{11}	0	**.914**	.085	.001	0	0	0
40	500	{3}	0	**.917**	.082	.001	0	0	0
40	500	{2}	0	**.918**	.081	.001	0	0	0
200	250	{11,11,11}	0	0	0	**1**	0	0	0
200	250	{5,5,5}	0	0	0	**1**	0	0	0
200	250	{3,3,3}	.003	.122	.567	**.308**	0	0	0
200	250	{11}	0	**1**	0	0	0	0	0
200	250	{5}	0	**1**	0	0	0	0	0
200	250	{3}	.181	**.819**	0	0	0	0	0
200	800	{5,5,5}	0	0	0	**1**	0	0	0
200	800	{3,3,3}	0	0	0	**1**	0	0	0
200	800	{2,2,2}	0	.009	.319	**.672**	0	0	0
200	800	{3}	0	**1**	0	0	0	0	0
200	800	{2}	.047	**.953**	0	0	0	0	0

NOTE: The numbers corresponding to the actual value of p are indicated by bold type.

Table 2. Probability Mass Function of $\hat{p}$ From MDL for Various n, N, and $\{\lambda_i/\sigma^2\}_{i=1}^{p}$

			$p_j \equiv$ probability$\{\hat{p} = j\}$						
n	N	$\{\lambda_i/\sigma^2\}_{i=1}^{p}$	p_0	p_1	p_2	p_3	p_4	p_5	p_6
40	50	{11,11,11}	0	0	.008	**.992**	0	0	0
40	50	{5,5,5}	.39	.405	.168	**.037**	0	0	0
40	50	{3,3,3}	.995	.005	0	**0**	0	0	0
40	50	{11}	0	**1**	0	0	0	0	0
40	50	{5}	.577	**.423**	0	0	0	0	0
40	50	{3}	.997	**.003**	0	0	0	0	0
40	100	{11,11,11}	0	0	0	**1**	0	0	0
40	100	{5,5,5}	.005	.029	.233	**.733**	0	0	0
40	100	{2,2,2}	1	0	0	**0**	0	0	0
40	100	{11}	0	**1**	0	0	0	0	0
40	100	{5}	.047	**.953**	0	0	0	0	0
40	100	{2}	1	**0**	0	0	0	0	0
40	500	{11,11,11}	0	0	0	**1**	0	0	0
40	500	{3,3,3}	0	0	0	**1**	0	0	0
40	500	{2,2,2}	.931	.067	.002	**0**	0	0	0
40	500	{11}	0	**1**	0	0	0	0	0
40	500	{3}	0	**1**	0	0	0	0	0
40	500	{2}	.96	**.04**	0	0	0	0	0
200	250	{11,11,11}	0	0	0	**1**	0	0	0
200	250	{5,5,5}	.995	.005	0	**0**	0	0	0
200	250	{3,3,3}	1	0	0	**0**	0	0	0
200	250	{11}	0	**1**	0	0	0	0	0
200	250	{5}	1	**0**	0	0	0	0	0
200	250	{3}	1	**0**	0	0	0	0	0
200	800	{5,5,5}	0	0	0	**1**	0	0	0
200	800	{3,3,3}	1	0	0	**0**	0	0	0
200	800	{2,2,2}	1	0	0	**0**	0	0	0
200	800	{3}	1	**0**	0	0	0	0	0
200	800	{2}	1	**0**	0	0	0	0	0

NOTE: The numbers corresponding to the actual value of p are indicated by bold type.

500, this method greatly overestimated the number of faults and will not be considered further. For the Λ_{mod} method, all thresholds were chosen to be the upper .001 percentile of the appropriate chi-squared distribution. The values of $\{\lambda_i/\sigma^2\}_{i=1}^{p}$ were chosen to span the range of what can be considered small (in terms of the tests having difficulty detecting the faults) to large (in terms of the tests detecting the faults with high probability). Note that for the case $p = 1$, $\mathbf{c}_1'\mathbf{c}_1/\sigma^2 = \lambda_1/\sigma^2 - 1$. $\mathbf{c}_1'\mathbf{c}_1/\sigma^2$ can be viewed as the total variance (summed over all n measurements) due to the fault, divided by the average variance (averaged over all n measurements) due to the noise.

Table 4 shows the probability of correctly concluding that $p = 0$ when there are truly no faults present. This situation is of particular interest if an eigenvalue test is used to detect the presence of a fault. In this event, the probabilities shown in Table 4 are equal to 1 minus the Type I (α) error. The values in Table 4 are also rounded to three decimal places.

The results in Tables 1 through 3 agree with the fact that the MDL and Λ_{mod} methods provide consistent estimates of p, whereas the AIC method asymptotically overestimates p (consider the $n = 40$, $N = 500$ cases). On the other hand, the AIC method was generally able to detect much smaller magnitude faults than either the MDL or Λ_{mod} methods. For example, consider the case when $n = 40$, $N = 100$, and three faults are present with $\{\lambda_i/\sigma^2\}_{i=1}^{p} = \{5, 5, 5\}$. The AIC method correctly estimates p with .947 probability, whereas MDL and Λ_{mod} correctly estimate p with only .733 and .639 probability, respectively. Consider also the cases in which $n = 200$, $N = 250$, and $\{\lambda_i/\sigma^2\}_{i=1}^{p} = \{5, 5, 5\}$, $\{3, 3, 3\}$, $\{5\}$, or $\{3\}$. In these situations, the AIC method provides much more accurate estimates of p. Furthermore, for n large and N not sufficiently large, Λ_{mod} greatly overestimates p. Although not shown in Table 3, for $n = 200$ and $N = 250$, Λ_{mod} often estimated p to be greater than 20.

These observations also extend to the case in which there are no faults present, as shown in Table 4. The alpha error is generally very low for the MDL method and slightly higher for the AIC method. Note that for all cases the Type I α error is 0 to three decimal places for the MDL method. For large N/n (e.g., $n = 20$ or 50, $N = 500$), Λ_{mod} (which is consistent) has a lower Type I α error than AIC (which asymptotically overestimates p). For small N/n (e.g., $n = 100$, $N = 120$), AIC still has a very small Type I α error, whereas that for Λ_{mod} is close to 1.

In light of this, we suggest using either the AIC or MDL methods. If small-magnitude faults are expected and need to be detected, the AIC method is more suitable. Otherwise, the MDL method is preferable.

4. UNIQUE IDENTIFICATION OF MULTIPLE FAULTS

4.1 Estimating the Fault Geometry Vectors

Assume that $\mathbf{x}$ follows Model (1), where $\mathbf{C}$ has the structure in (2), and that Σ_x is known. Assume also that one has identified a subgroup of measurements that are affected by a single fault, a procedure for which is given in Section 4.2. Define the "latent" covariance matrix Σ_ϵ to be the component of Σ_x

Table 3. Probability Mass Function of $\hat{p}$ From Λ_{mod} for Various n, N, and $\{\lambda_i/\sigma^2\}_{i=1}^{p}$

			$p_j \equiv$ probability$\{\hat{p}=j\}$						
n	N	$\{\lambda_i/\sigma^2\}_{i=1}^{p}$	p_0	p_1	p_2	p_3	p_4	p_5	p_6
40	50	{11,11,11}	0	0	.033	**.867**	.034	.021	.021
40	50	{5,5,5}	.002	.111	.462	**.319**	.04	.022	.016
40	50	{3,3,3}	.289	.31	.202	**.099**	.038	.022	.011
40	50	{11}	.005	**.878**	.047	.027	.014	.011	.005
40	50	{5}	.315	**.573**	.047	.021	.016	.007	.01
40	50	{3}	.678	**.212**	.045	.02	.016	.008	.011
40	100	{11,11,11}	0	0	0	**.998**	.001	.001	0
40	100	{5,5,5}	0	0	.358	**.639**	.002	.001	0
40	100	{2,2,2}	.858	.12	.017	**.004**	.001	0	0
40	100	{11}	0	**.999**	0	.001	0	0	0
40	100	{5}	.108	**.891**	0	.001	0	0	0
40	100	{2}	.985	**.013**	0	.002	0	0	0
40	500	{11,11,11}	0	0	0	**.998**	.001	.001	0
40	500	{3,3,3}	0	0	0	**.999**	.001	0	0
40	500	{2,2,2}	0	.012	.669	**.318**	.001	0	0
40	500	{11}	0	**.998**	.001	0	0	.001	0
40	500	{3}	0	**.998**	.001	0	.001	0	0
40	500	{2}	.419	**.578**	.002	0	.001	0	0
200	250	{11,11,11}	0	0	0	**0**	0	0	0
200	250	{5,5,5}	0	0	0	**0**	0	0	0
200	250	{3,3,3}	0	0	0	**0**	0	0	0
200	250	{11}	0	**0**	0	0	0	0	0
200	250	{5}	0	**0**	0	0	0	0	0
200	250	{3}	0	**0**	0	0	0	0	0
200	800	{5,5,5}	0	0	0	**.995**	.005	0	0
200	800	{3,3,3}	0	.003	.461	**.523**	.004	.003	.002
200	800	{2,2,2}	.289	.475	.208	**.025**	.003	0	0
200	800	{3}	.301	**.693**	.005	.001	0	0	0
200	800	{2}	.939	**.055**	.005	.001	0	0	0

NOTE: The numbers corresponding to the actual value of p are indicated by bold type.

that is due to the underlying latent variables that represent the faults; that is,

$$\Sigma_c = \mathbf{Z}_p\left[\Lambda_p - \sigma^2\mathbf{I}\right]\mathbf{Z}_p' = \mathbf{C}\mathbf{C}' = \Sigma_{i=1}^{p}\mathbf{c}_i\mathbf{c}_i' = \begin{bmatrix}\mathbf{c}_{1,1}\\ \mathbf{c}_{2,1}\\ \vdots\\ \mathbf{c}_{p,1}\end{bmatrix}\left[\mathbf{c}_{1,1}'\ \mathbf{c}_{2,1}'\cdots\mathbf{c}_{p,1}'\right] + \begin{bmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \Sigma_c^{(2)}\end{bmatrix}. \tag{5}$$

The last equality follows from the assumed structure of $\mathbf{C}$. From (5), the upper left $n_1 \times n_1$ block of Σ_c is the rank-1 matrix $\mathbf{c}_{1,1}\mathbf{c}_{1,1}'$, which has a single nonzero eigenvalue $\lambda_{1,1} = \mathbf{c}_{1,1}'\mathbf{c}_{1,1}$ with eigenvector $\mathbf{z}_{1,1} = \mathbf{c}_{1,1}\lambda_{1,1}^{-1/2}$. Thus, the first fault geometry vector can be obtained from Σ_c via

$$\Sigma_c\begin{bmatrix}\mathbf{z}_{1,1}\lambda_{1,1}^{-1/2}\\ \mathbf{0}\end{bmatrix} = \left[\begin{bmatrix}\mathbf{c}_{1,1}\\ \mathbf{c}_{2,1}\\ \vdots\\ \mathbf{c}_{p,1}\end{bmatrix}\left[\mathbf{c}_{1,1}'\ \mathbf{c}_{2,1}'\cdots\mathbf{c}_{p,1}'\right] + \begin{bmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \Sigma_c^{(2)}\end{bmatrix}\right] \times \begin{bmatrix}\mathbf{c}_{1,1}(\mathbf{c}_{1,1}'\mathbf{c}_{1,1})^{-1}\\ \mathbf{0}\end{bmatrix} = \begin{bmatrix}\mathbf{c}_{1,1}\\ \mathbf{c}_{2,1}\\ \vdots\\ \mathbf{c}_{p,1}\end{bmatrix} = \mathbf{c}_1.$$

After identifying $\mathbf{c}_1$, deflate the latent covariance matrix via

$$\Sigma_c - \mathbf{c}_1\mathbf{c}_1' = \sum_{i=2}^{p}\mathbf{c}_i\mathbf{c}_i' = \begin{bmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \Sigma_c^{(2)}\end{bmatrix},$$

where

$$\Sigma_c^{(2)} = \begin{bmatrix}\mathbf{c}_{2,2}\\ \mathbf{c}_{3,2}\\ \vdots\\ \mathbf{c}_{p,2}\end{bmatrix}\left[\mathbf{c}_{2,2}'\ \mathbf{c}_{3,2}'\cdots\mathbf{c}_{p,2}'\right] + \begin{bmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \Sigma_c^{(3)}\end{bmatrix}.$$

The upper left $n_2 \times n_2$ block of $\Sigma_c^{(2)}$ is the rank-1 matrix $\mathbf{c}_{2,2}\mathbf{c}_{2,2}'$, which has a single nonzero eigenvalue $\lambda_{2,2} = \mathbf{c}_{2,2}'\mathbf{c}_{2,2}$

Table 4. Probability That $\hat{p} = 0$ When No Fault Is Present, for Various n and N

		Probability$\{\hat{p}=0\}$		
n	N	AIC	MDL	λ_{mod}
20	50	.926	1.000	.998
20	100	.911	1.000	.999
20	200	.879	1.000	.999
20	500	.858	1.000	.999
50	75	.991	1.000	.952
50	150	.978	1.000	.998
50	250	.968	1.000	.999
50	500	.947	1.000	.016
100	120	1.000	1.000	.016
100	200	1.000	1.000	.979
100	500	.992	1.000	1.000
200	250	1.000	1.000	.000
200	500	1.000	1.000	.961

with eigenvector $\mathbf{z}_{2,2} = \mathbf{c}_{2,2}\lambda_{2,2}^{-1/2}$. Thus, $\mathbf{c}_2$ can be obtained via

$$\left[\Sigma_e - \mathbf{c}_1\mathbf{c}_1'\right]\begin{bmatrix}\mathbf{0}\\ \mathbf{Z}_{2,2}\lambda_{2,2}^{-1/2}\\ \mathbf{0}\end{bmatrix}$$

$$= \left[\begin{bmatrix}\mathbf{0}\\ \mathbf{c}_{2,2}\\ \vdots\\ \mathbf{c}_{p,2}\end{bmatrix}\begin{bmatrix}\mathbf{0}\,\mathbf{c}_{2,2}'\cdots\mathbf{c}_{p,2}'\end{bmatrix} + \begin{bmatrix}\mathbf{0} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \Sigma_e^{(3)}\end{bmatrix}\right]$$

$$\times\begin{bmatrix}\mathbf{0}\\ \mathbf{c}_{2,2}(\mathbf{c}_{2,2}'\mathbf{c}_{2,2})^{-1}\\ \mathbf{0}\end{bmatrix} = \mathbf{c}_2.$$

The entire process can be repeated, each time deflating the latent covariance matrix with the most recently identified fault geometry vector until $\mathbf{c}_p$ is identified. At the final stage, $\mathbf{C}$ is completely determined.

In practice, Σ_x will not be known, and the sample covariance matrix $\widehat{\Sigma}_x$ must be used instead. We suggest using the preceding procedure with all quantities replaced by their estimates. Let $\hat{p}$ denote an estimate of p. Form $\widehat{\mathbf{Z}}_{\hat{p}}$ and $\widehat{\Lambda}_{\hat{p}}$ from the $\hat{p}$ largest eigenvalues/eigenvectors of $\widehat{\Sigma}_x$, and estimate σ^2 via

$$\hat{\sigma}^2 = (n-\hat{p})^{-1}\sum_{i=\hat{p}+1}^{n}\hat{\lambda}_i. \tag{6}$$

Anderson (1963) showed that (6) is asymptotically the maximum likelihood estimate of σ^2. An estimate of Σ_e would then be $\widehat{\Sigma}_e = \widehat{\mathbf{Z}}_{\hat{p}}[\widehat{\Lambda}_{\hat{p}} - \hat{\sigma}^2\mathbf{I}]\widehat{\mathbf{Z}}_{\hat{p}}'$. At each stage, the estimate of $\mathbf{c}_i$ is

$$\hat{\mathbf{c}}_i = \left[\widehat{\Sigma}_e - \sum_{j=0}^{i-1}\hat{\mathbf{c}}_j\hat{\mathbf{c}}_j'\right]\begin{bmatrix}\mathbf{0}\\ \hat{\mathbf{z}}_{i,i}\hat{\lambda}_{i,i}^{-1/2}\\ \mathbf{0}\end{bmatrix}. \tag{7}$$

$\hat{\lambda}_{i,i}$ and $\hat{\mathbf{z}}_{i,i}$ denote the dominant eigenvalue/eigenvector pair of the $n_i \times n_i$ block of the deflated latent covariance matrix, corresponding to the subgroup of measurements affected by only one of the remaining faults. The subgroups can be identified by the procedure given in Section 4.2.

Although the upper left $n_1 \times n_1$ submatrix of Σ_e has rank 1, the corresponding submatrix of $\widehat{\Sigma}_e$ will not necessarily. Its rank will be close to 1, however, if $\widehat{\Sigma}_e$ is sufficiently close to Σ_e. One eigenvalue of the upper left $n_1 \times n_1$ block of $\widehat{\Sigma}_e$ will then be much larger than the others, and it should be clear how to choose $\hat{\lambda}_{1,1}$. Likewise for $\hat{\lambda}_{i,i}$, $i = 2, 3, \ldots, \hat{p}$. This issue is related to whether the subgroups affected by only one fault have been appropriately selected (see Sec. 4.2).

4.2 Identifying Subgroups

The accuracy in estimating the fault geometry vectors depends on, among other factors, whether the subgroups of measurements affected by only one of the remaining faults can be identified. This section discusses a simple procedure to accomplish this.

If $\{x_1, x_2, \ldots, x_{n_1}\}$ is a subgroup affected by only one fault, the submatrix of Σ_e corresponding to this subgroup is the matrix $\mathbf{c}_{1,1}\mathbf{c}_{1,1}'$ having rank 1. Thus, to identify the first subgroup it is necessary to find a set of measurements whose latent covariance matrix has rank 1. Define $\mathbf{R}_e$ to be the latent correlation matrix associated with Σ_e, obtained in the usual way by scaling each element of Σ_e by the square roots of the corresponding diagonal elements. A submatrix of Σ_e will have unit rank if and only if all elements of the corresponding submatrix of $\mathbf{R}_e$ have unit magnitude. Consequently, finding a subgroup of measurements whose latent covariance matrix has unit rank reduces to finding a subgroup whose latent correlation matrix has all elements equal to either 1 or -1.

In the event that Σ_e is known, this subgroup can be found by inspecting the elements of $\mathbf{R}_e$. If Σ_e is estimated, no submatrix of $\widehat{\mathbf{R}}_e$ will have all unit magnitude elements with positive probability. Therefore, $\widehat{\mathbf{R}}_e$ must be searched for submatrices whose elements are all close to 1 in magnitude. If n is large, we recommend using a clustering technique with the magnitude of the latent correlation coefficients as the similarity measure. We have found agglomerative hierarchical methods with complete linkage to work well in practice (Everitt 1993).

The result of the clustering will be a set of candidate subgroups, one of which must be selected as the first subgroup $\{x_1, x_2, \ldots, x_{n_1}\}$. To accomplish this, we recommend calculating the eigenvalues of the submatrices of $\widehat{\Sigma}_e$ that correspond to each candidate subgroup. For notational convenience, denote the eigenvalues for a given subgroup (of size n_1) as $\{\lambda_i\}_{i=1}^{n_1}$, arranged in descending order. We suggest choosing the subgroup that maximizes the criterion

$$\frac{\lambda_1}{(n_1-1)^{-1}\sum_{i=2}^{n_1}\lambda_j}. \tag{8}$$

This criterion is the ratio of the largest eigenvalue to the average of the remaining eigenvalues, and helps to ensure that the rank of the corresponding submatrix of $\widehat{\Sigma}_e$ is close to 1. After identifying the first subgroup, the first fault geometry vector is estimated and $\widehat{\Sigma}_e$ is deflated as described in Section 4.1. The preceding procedure can then be repeated on the deflated latent covariance matrix to identify the second subgroup, and so on.

4.3 Fault Interpretation and Illustrative Example

The estimated fault geometry vectors can provide powerful diagnostic tools for identifying root causes of process variability, as illustrated with the following example. Consider again the measurement layout on the BIW described in Section 1 and shown in Figure 2. The sample consists of $N = 200$ autobodies, produced and measured over a four-hour period. The measurements are in units millimeters. All three methods discussed in Section 3 indicated that $\hat{p} = 4$. $\widehat{\Sigma}_e$ was formed from the four dominant eigenvalue/eigenvector pairs, with $\hat{\sigma}^2 = .040$ estimated from the remaining 72 eigenvalues. Using the clustering procedure outlined in Section 4.2, several candidate subgroups were found. The subgroup {11X, 13X, 15X, 16X, 17X, 18X, 24X, 25X} was selected first since it maximized (8). The eigenvalues of the corresponding

submatrix of $\hat{\Sigma}_e$ were {.2873, .0008, .0004, .0004, 0, 0, 0, 0}, and the corresponding submatrix of $\hat{\mathbf{R}}_e$ was

$$\begin{bmatrix} 1 & .994 & .997 & .991 & .995 & .985 & .993 & .998 \\ .994 & 1 & .994 & .985 & .997 & .982 & .998 & .991 \\ .997 & .994 & 1 & .996 & .998 & .994 & .997 & .999 \\ .991 & .985 & .996 & 1 & .996 & .989 & .993 & .995 \\ .995 & .997 & .998 & .996 & 1 & .987 & .999 & .995 \\ .985 & .982 & .994 & .989 & .987 & 1 & .989 & .994 \\ .993 & .998 & .997 & .993 & .999 & .989 & 1 & .994 \\ .998 & .991 & .999 & .995 & .995 & .994 & .994 & 1 \end{bmatrix}.$$

Thus, it appears that subgroup 1 is affected by only a single fault. $\hat{\mathbf{c}}_1$, estimated via (7), is shown graphically in Figure 3. The length of each arrow is the "6-sigma" value, due to fault 1, for that coordinate; that is, the arrow length is six times the magnitude of the corresponding element of $\hat{\mathbf{c}}_1$. To make the plot less cluttered, an arrow was omitted if the 6-sigma level was less than .25 mm (deemed insignificant from a practical viewpoint). Fault 1 appears to affect only fore/aft direction coordinates, except for a relatively minor effect on point 1Y. Moreover, it appears to translate all points on the vehicle by approximately the same amount in the fore/aft direction. Since the fault does not appear to cause the autobody panels to vary relative to each other, measurement error was suspected to be the cause. Investigation revealed that, in the measurement station, the vehicle's position is fixed via a pin/hole/slot combination in the underbody, as described in Section 1. The same underbody hole is used to locate the vehicle throughout the assembly process, prior to measurement. Through repeated use, the hole in each underbody was significantly enlarged in the fore/aft direction, allowing the vehicles to translate in the fore/aft direction when placed into the measurement station fixture and resulting in the measurement error represented by fault 1. The problem was solved by changing the design of the pin/hole.

After estimating the first fault, $\hat{\Sigma}_x$ was deflated and the procedure was repeated. The clustering procedure produced as a second subgroup {1X, 2X, 3X, 4X, 5X, 6X, 7X, 8X, 9X}. The eigenvalues of the corresponding submatrix of the deflated latent covariance matrix were {.5205, .0018, .0008, 0, 0, 0, 0, 0, 0}, and the corresponding correlation matrix mostly contained elements greater than .99 in magnitude. $\hat{\mathbf{c}}_2$, estimated via (7), is shown graphically in Figure 4. It appears that the second fault also affects only fore/aft direction points. In contrast to fault 1, fault 2 affects only points on the right bodyside, causing each point to translate by approximately the same amount. Thus, it appears that fault 2 results in the right bodyside being incorrectly positioned with respect to the rest of the vehicle, identical to the variation pattern described in Section 1. This observation led to the suspicion that, in the station in which the bodysides are joined to the roof and underbody, the pin that locates the right bodyside in the fore/aft direction was loose or worn. Further investigation revealed that it was loose.

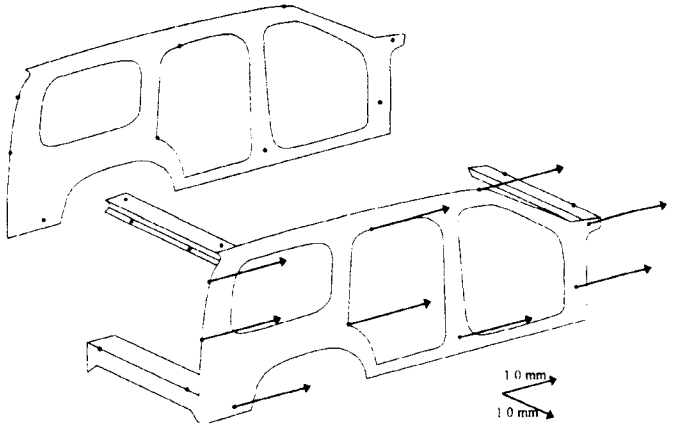

Figure 4. Graphical Illustration of $\mathbf{c}_2$ for the Example.

The latent covariance matrix was deflated again, and the procedure was repeated twice more to estimate the two remaining fault geometry vectors. Although not shown, faults 3 and 4 affected the y-direction measurements predominantly and also had clear interpretations. One of the faults arose from a misaligned robotic weld gun. Together, the four fault geometry vectors resulted in a $\mathbf{C}$ matrix with the structure (2).

These results can be compared with standard PCA, in which the four dominant eigenvectors of $\hat{\Sigma}_x$ are directly plotted. Two of the eigenvectors, $\hat{\mathbf{z}}_1$ and $\hat{\mathbf{z}}_3$, are shown in Figures 5 and 6.

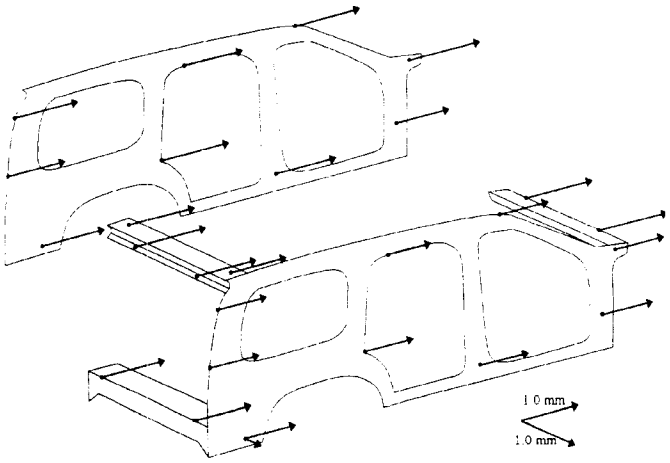

Figure 3. Graphical Illustration of $\mathbf{c}_1$ for the Example.

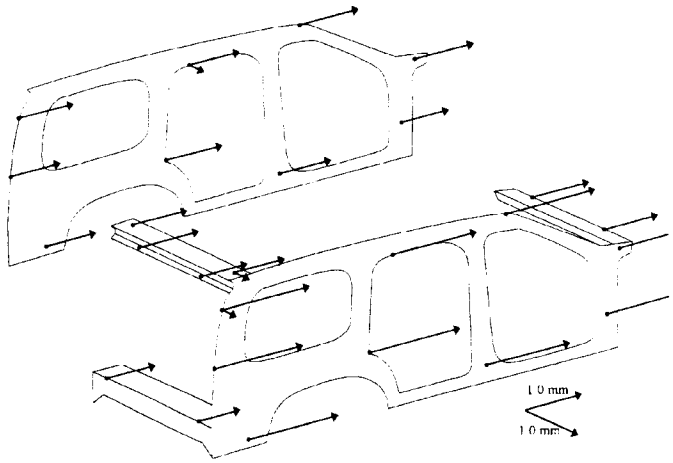

Figure 5. Graphical Illustration of $\mathbf{z}_1$ in the Example.

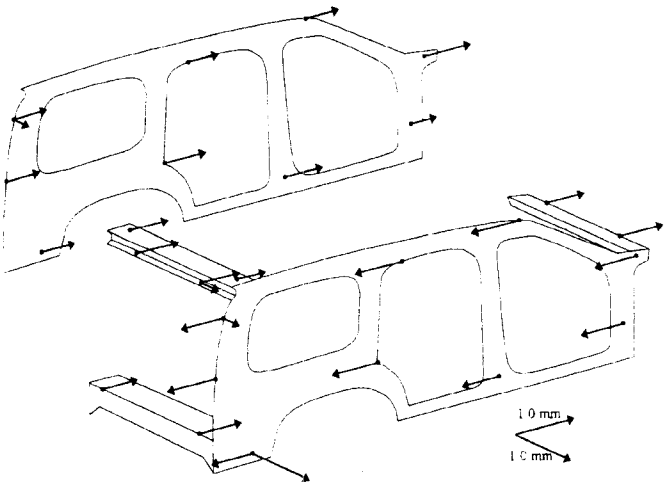

Figure 6. Graphical Illustration of $\mathbf{z}_3$ in the Example.

Interpretation of $\hat{\mathbf{z}}_1$ could possibly have led to the same conclusion as $\hat{\mathbf{c}}_1$, although $\hat{\mathbf{z}}_1$ makes it appear that the fault causes the right bodyside to translate by larger amounts than the rest of the vehicle. $\hat{\mathbf{c}}_1$ more clearly represents the actual fault. $\hat{\mathbf{z}}_3$ would be difficult to interpret, since it appears that the fault causes the right and left sides of the vehicle to move in opposite directions.

In this example, $\widehat{\mathbf{C}}$ can be written in the form $\widehat{\mathbf{Z}}_p[\widehat{\Lambda}_p - \hat{\sigma}^2\mathbf{I}]^{1/2}\mathbf{Q}$, where $\mathbf{Q}$ is a 4×4 orthogonal matrix, since the method of this article provides a $\widehat{\mathbf{C}}$ that preserves the latent covariance structure. Specifically, the way $\widehat{\mathbf{C}}$ is formed, $\widehat{\mathbf{C}}\widehat{\mathbf{C}}'$ is always equal to $\widehat{\Sigma}_c$. Thus, the method of this article represents a form of factor rotation, in which the criterion for "best" interpretability is that the columns of $\widehat{\mathbf{C}}$ are the actual fault geometry vectors.

As in PCA, a breakdown of the percentage of total sample variation due to each fault can be easily obtained. The estimated variation due to the ith fault is $\mathbf{c}_i'\mathbf{c}_i$. The total variation is the trace of $\widehat{\Sigma}_c$, or, equivalently, the sum of its eigenvalues. For the preceding example, the total variation is 5.18 mm^2. The percentages due to faults 1 through 4 are 15.5%, 10.3%, 10.4%, and 5.6%, respectively. Together, the four faults account for 41.8% of the total sample variability.

5. STATISTICAL PROPERTIES

The statistical properties of the estimated fault geometry vectors are complicated, in general, and depend on a number of factors that include $n, N, p, \sigma^2, \{\mathbf{c}_i\}_{i=1}^p$. The fact that the estimates depend on which measurements are selected as the subgroups, which involves user subjectivity, further complicates the analysis. Some results can be obtained, however, in special situations. It is assumed that the data are normally distributed.

Of particular interest is the case of only one fault present and the number of faults correctly estimated. Then $\widehat{\Sigma}_c$ has unit rank, and $\hat{\mathbf{c}}_1 = \sqrt{\hat{\lambda}_1 - \hat{\sigma}^2}\hat{\mathbf{z}}_1$. Although the distribution of eigenvalues and eigenvectors of a sample covariance matrix is, in general, complicated, Anderson (1963) provided some asymptotic (for large N) results when the data are normally distributed. Using these results, the following asymptotic expression for the accuracy of $\hat{\mathbf{c}}_1$ is derived in Appendix B.

$$\kappa \equiv E\left[\frac{\|\hat{\mathbf{c}}_1 - \mathbf{c}_1\|^2}{\|\mathbf{c}_1\|^2}\right] \cong \frac{1}{N}\left\{(n-1)\gamma(1+\gamma) + \frac{(1+\gamma)^2}{2} + \frac{\gamma^2}{2(n-1)}\right\}, \tag{9}$$

where $\gamma \equiv \sigma^2/\mathbf{c}_1'\mathbf{c}_1$ can be viewed as the inverse of a "signal-to-noise ratio," since $\mathbf{c}_1'\mathbf{c}_1$ is the total variance due to the fault. For large n, the third term can be neglected, and (9) reduces to

$$\kappa \cong \frac{1+\gamma}{N}\left\{n\gamma + \frac{1-\gamma}{2}\right\}. \tag{10}$$

Clearly, increasing sample size N reduces κ and improves accuracy. It can be verified that the partial derivative of (10) with respect to γ is always positive. Consequently, holding n and N fixed and increasing the signal-to-noise ratio (decreasing γ) always reduces κ, which is intuitively obvious. The effects of n are less obvious. Equation (10) seems to imply that increasing n increases κ and decreases accuracy. γ, however, may also change with n, since $\mathbf{c}_1$ will increase in dimension and $\mathbf{c}_1'\mathbf{c}_1$ may change. As n increases, if the additional measurements are not affected by the fault, then γ remains constant, and κ does increase. Thus, adding irrelevant measurements will decrease accuracy.

If the additional measurements are affected by the fault, it is convenient to define $\bar{\gamma} \equiv \sigma^2/(n^{-1}\mathbf{c}_1'\mathbf{c}_1) = n\gamma$, which can be viewed as the average variance due to the noise, divided by the *average* variance due to the fault. As n increases, suppose the additional measurements are affected by the fault in such a way that $\bar{\gamma}$ remains constant. Substituting $\gamma = \bar{\gamma}/n$ into (10) reveals that the partial derivative of κ with respect to n (holding $\bar{\gamma}$ and N fixed) is always negative for $n > 1$. Consequently, adding measurements that are affected by the fault (to the point that $\bar{\gamma}$ does not decrease) will reduce κ and improve accuracy.

The general case in which $p > 1$ is more difficult to analyze. Suppose that the MDL method, which is known to yield a consistent estimate of p, is used so that asymptotically

$$\widehat{\Sigma}_c = \widehat{\mathbf{Z}}_p[\widehat{\Lambda}_p - \hat{\sigma}^2\mathbf{I}]\widehat{\mathbf{Z}}_p' = \sum_{i=1}^{p}(\hat{\lambda}_i - \hat{\sigma}^2)\hat{\mathbf{z}}_i\hat{\mathbf{z}}_i'.$$

For the case that $\lambda_1 > \lambda_2 > \cdots > \lambda_p$, the results of Anderson (1963) apply directly, yielding for $i = 1, 2, \ldots, p$,

$$\hat{\lambda}_i \sim \mathrm{N}\left(\lambda_i, \frac{2\lambda_i^2}{\mathrm{N}}\right),$$

$$\hat{\sigma}^2 \sim \mathrm{N}\left(\sigma^2, \frac{2\sigma^4}{(n-p)\mathrm{N}}\right),$$

and

$$\hat{\mathbf{z}}_i \sim \mathrm{N}\left(\mathbf{z}_i, \frac{\lambda_i}{\mathrm{N}}\sum_{\substack{k=1 \\ k\neq i}}^{n}\frac{\lambda_k}{(\lambda_k - \lambda_i)^2}\mathbf{z}_k\mathbf{z}_k'\right).$$

$\{\hat{\lambda}_i, \hat{\mathbf{z}}_i\}_{i=1}^p$ and $\hat{\sigma}^2$ are clearly consistent estimates. Therefore, so is $\widehat{\Sigma}_c$. Since $\{\hat{c}_i\}_{i=1}^p$ are calculated entirely from $\widehat{\Sigma}_c$, they also are consistent estimates, providing that the measurement subgroups have been correctly identified.

For the case in which $\lambda_j = \lambda_{j+1}$ for some $1 \le j < p$, Anderson's results do not apply. In fact, $\mathbf{z}_j$ and $\mathbf{z}_{j+1}$ are not even uniquely defined when their eigenvalue has multiplic-

ity 2, although the projection operator onto their eigenspace is. Consider the extreme case in which $\lambda_1 = \lambda_2 = \cdots = \lambda_p \equiv \lambda$. Note that $\mathbf{Z}_p\mathbf{Z}_p'$, the projection operator onto the eigenspace of λ, is well defined. Tyler (1981) showed that

$$\sqrt{N}\left[\widehat{\mathbf{Z}}_p\widehat{\mathbf{Z}}_p' - \mathbf{Z}_p\mathbf{Z}_p'\right] \to \frac{1}{\lambda - \sigma^2}\left\{\mathbf{Z}_p\mathbf{Z}_p'\mathbf{B}\left[\mathbf{I} - \mathbf{Z}_p\mathbf{Z}_p'\right] + \left[\mathbf{I} - \mathbf{Z}_p\mathbf{Z}_p'\right]\mathbf{B}\mathbf{Z}_p\mathbf{Z}_p'\right\},$$

where the convergence is in distribution as $N \to \infty$. Here, $\mathbf{B}$ is an $n \times n$ multivariate normal matrix with zero mean and covariance given by $\text{cov}([\mathbf{B}]_{i,j}, [\mathbf{B}]_{k,l}) = [\Sigma_x]_{i,k}[\Sigma_x]_{j,l} + [\Sigma_x]_{i,l}[\Sigma_x]_{j,k}$. Thus, $\widehat{\mathbf{Z}}_p\widehat{\mathbf{Z}}_p'$ is a consistent estimate of $\mathbf{Z}_p\mathbf{Z}_p'$. Since $\widehat{\Lambda}_p$ and $\hat{\sigma}^2$ are consistent estimates of $\Lambda_p = \lambda\mathbf{I}$ and σ^2, respectively, $\widehat{\Sigma}_c$ is also consistent. Similar arguments can be applied when some, but not all, of $\{\lambda_i\}_{i=1}^p$ are equal.

6. CONCLUSIONS

This article presents a method to facilitate the identification and elimination of root causes of variability in manufacturing processes. The method extracts diagnostic information from typically large quantities of multivariate process measurement data. As technology for automated in-process measurement becomes less expensive, more accurate, more reliable, and more widely used by manufacturing industries, there are many applications in which the proposed method could be used to reduce variability. The effectiveness of the method has been demonstrated with an example from autobody assembly.

The problem formulation and objectives are closely related to the factor-analysis problem. The variability patterns (the fault geometry vectors) must be estimated from process data with no a priori knowledge of the faults, except for an assumed model structure. In contrast to traditional factor-analysis methods, much greater emphasis has been placed on the underlying physical model that describes the effects of the faults on process variability. The criterion for "rotating" the factors is that the results should correspond closely to the actual faults, as opposed to artificial criteria that seek to aggregate the maximum amount of variability into the minimum number of variables. This allows more physically meaningful interpretations that better facilitate the overall goal of identifying and eliminating the root causes of variability.

The applicability of the method depends predominantly on whether the linear (or linearized) Model (1), with the assumed structure (2) for $\mathbf{C}$, adequately represents the effects of process faults. Ultimately, this depends on the underlying physics of the process and the faults, although we believe that the model provides a reasonable representation of many manufacturing processes. Additionally, the method requires some degree of subjective judgment by the user in identifying the measurement subgroups. The development of more "black-box" methods that remove the subjectivity and relax the model assumptions would be a valuable contribution.

ACKNOWLEDGMENTS

We thank the editor, associate editor, and anonymous referees for their many insightful comments, which have helped to improve the quality of this article.

APPENDIX A: INVARIANCE OF THE ORDER-SELECTION CRITERIA WITH RESPECT TO THE EIGENVECTORS

This appendix proves that the distributions of $\Lambda(m)$, $\Lambda_{\text{mod}}(m)$, AIC(m), and MDL(m) depend only on n, N, and $\{\lambda_i/\sigma^2\}_{i=1}^p$ and not on the individual eigenvectors. Assume that $\mathbf{x}$ is Gaussian. Consider the eigenvector decomposition of Σ_x in (4). Expanding $\mathbf{x}$ in terms of its principal components, it follows that $\mathbf{x}$ is distributed as $\mathbf{Zu}$, where $\mathbf{u}$ follows a multivariate normal distribution with covariance matrix Λ, and $\mathbf{Z} = [\mathbf{Z}_1, \mathbf{Z}_2, \ldots, \mathbf{Z}_n]$. Then,

$$\widehat{\Sigma}_x \equiv \frac{1}{N-1}\sum_{i=1}^{N}(\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}}) = \mathbf{Z}\left[\frac{1}{N-1}\sum_{i=1}^{N}(\mathbf{u}_i - \bar{\mathbf{u}})(\mathbf{u}_i - \bar{\mathbf{u}})'\right]\mathbf{Z}',$$

where $\bar{\mathbf{x}}$ and $\bar{\mathbf{u}}$ are the sample averages for $\mathbf{x}$ and $\mathbf{u}$, respectively. Since $\mathbf{Z}$ is an orthogonal matrix, the eigenvalues of $\widehat{\Sigma}_x$ are the eigenvalues of the matrix in brackets, the distribution of which depends only on n, $\mathbf{N}$, and $\{\lambda_i\}_{i=1}^n$ and not on $\mathbf{Z}$.

Now suppose that $\mathbf{y} \equiv a^{1/2}\mathbf{x}$ for some positive constant a, so that the covariance matrix of $\mathbf{y}$ is $\mathbf{Za\Lambda Z}'$. The eigenvalues of $\widehat{\Sigma}_y$ are distributed as the eigenvalues of $\widehat{\Sigma}_x$ multiplied by the constant a. By inspection of each of the four test statistics, it is clear that the constant drops out of the distribution. Consequently, if all eigenvalues of Σ_x are divided by a constant, say σ^2, the distribution of each of the test statistics remains unchanged. Therefore, the distribution of the test statistics depends only on n, N, and $\{\lambda_i/\sigma^2\}_{i=1}^n$. Noting that $\lambda_i/\sigma^2 = 1$, for $i = p+1, p+2, \ldots, n$, completes the proof.

APPENDIX B: DERIVATION OF THE ACCURACY MEASURE κ

This appendix derives the asymptotic expression for κ in (9) when $\hat{p} = p = 1$ and the data are Gaussian. Note that $\hat{\mathbf{c}}_1 = \sqrt{\hat{\lambda}_1 - \hat{\sigma}^2}\,\hat{\mathbf{z}}_1$. From Theorem 1 and equation (3.10) of Anderson (1963), the following results hold asymptotically:

$$\begin{aligned}
\hat{\lambda}_1 &\sim N\left(\lambda_1, \frac{2\lambda_1^2}{N}\right)\\
\hat{\sigma}^2 &\sim N\left(\sigma^2, \frac{2\sigma^4}{(n-1)N}\right)\\
\hat{\mathbf{z}}_1 &\sim N\left(\mathbf{z}_1, \frac{\lambda_1}{N}\sum_{k=2}^{n}\frac{\lambda_k}{(\lambda_k - \lambda_1)^2}\mathbf{z}_k\mathbf{z}_k'\right)\\
&= N\left(\mathbf{z}_1, \frac{\sigma^2 + \mathbf{c}_1'\mathbf{c}_1}{N}\sum_{k=2}^{n}\frac{\sigma^2}{(\mathbf{c}_1'\mathbf{c}_1)^2}\mathbf{z}_k\mathbf{z}_k'\right)\\
&= N\left(\mathbf{z}_1, \frac{\gamma(1+\gamma)}{N}\sum_{k=2}^{n}\mathbf{z}_k\mathbf{z}_k',\right), \qquad \text{(B.1)}
\end{aligned}$$

where $N(\bullet, \bullet)$ denotes the (scalar or multivariate) normal distribution. It also follows from theorem 1 of Anderson (1963) that $\hat{\lambda}_1$, $\hat{\sigma}^2$, and $\hat{\mathbf{z}}_1$ are asymptotically independent. Consequently,

$$\hat{\lambda}_1 - \hat{\sigma}^2 \sim N\left(\mathbf{c}_1'\mathbf{c}_1, \frac{2}{N}\left[(\mathbf{c}_1'\mathbf{c}_1 + \sigma^2)^2 + \frac{\sigma^4}{n-1}\right]\right) \qquad \text{(B.2)}$$

and is independent of $\hat{\mathbf{z}}_1$. The result $\lambda_1 = \mathbf{c}_1'\mathbf{c}_1 + \sigma^2$ has been used to obtain (B.2).

Lemma 1. Let x be a scalar random variable with mean μ_x and variance σ_x^2. Let $\mathbf{y}$ be a random vector, independent of x, with mean $\boldsymbol{\mu}_y$ and covariance Σ_y, and define $\mathbf{z} \equiv x\mathbf{y}$. Then $E[\mathbf{z}] = \mu_x\boldsymbol{\mu}_y$, and $\text{cov}(\mathbf{z}) = E[x^2]\Sigma_y + \sigma_x^2\boldsymbol{\mu}_y\boldsymbol{\mu}_y'$, where $\text{cov}(\mathbf{z})$ denotes the covariance matrix of $\mathbf{z}$. The proof is straightforward.

Lemma 2. Let x be a scalar random variable with mean μ and variance σ^2, where $\mu \neq 0$ and σ is sufficiently small. Then $E[\sqrt{x}] \cong \mu^{1/2} - \sigma^2/(8\mu^{3/2})$ and $\text{var}[\sqrt{x}] \cong \sigma^2/(4\mu)$, where $\text{var}(\bullet)$ denotes the variance of a random variable. The proof follows by taking a second-order Taylor expansion of $\sqrt{x}$ around $\sqrt{\mu}$ to approximate $E[\sqrt{x}]$. The expression for $\text{var}[\sqrt{x}]$ follows by noting $\text{var}[\sqrt{x}] = E[x] - (E[\sqrt{x}])^2$.

From (B.2) and Lemma 2,

$$\text{var}\left[\sqrt{\hat{\lambda}_1 - \hat{\sigma}^2}\right] \cong \frac{\text{var}(\hat{\lambda}_1 - \sigma^2)}{4E(\hat{\lambda}_1 - \sigma^2)} \cong \frac{\mathbf{c}_1'\mathbf{c}_1}{2N}\left[(1+\gamma)^2 + \frac{\gamma^2}{n-1}\right]. \tag{B.3}$$

Applying Lemma 1 to $\hat{\mathbf{c}}_1 \equiv \sqrt{\hat{\lambda}_1 - \hat{\sigma}^2}\,\hat{\mathbf{z}}_1$, and using (B.1) through (B.3),

$$\begin{aligned}
\text{trace}(\text{cov}(\hat{\mathbf{c}}_1)) &\cong E[\hat{\lambda}_1 - \hat{\sigma}^2]\text{trace}(\text{cov}(\hat{\mathbf{z}}_1)) \\
&\quad + \text{var}\left[\sqrt{\hat{\lambda}_1 - \hat{\sigma}^2}\right]\text{trace}(E[\hat{\mathbf{z}}_1]E[\hat{\mathbf{z}}_1']) \\
&\cong \mathbf{c}_1'\mathbf{c}_1\frac{\gamma(1+\gamma)(n-1)}{N} \\
&\quad + \frac{\mathbf{c}_1'\mathbf{c}_1}{2N}\left[(1+\gamma)^2 + \frac{\gamma^2}{n-1}\right] \\
&= \frac{\mathbf{c}_1'\mathbf{c}_1}{N}\left[(n-1)\gamma(1+\gamma)\right. \\
&\quad \left. + \frac{(1+\gamma)^2}{2} + \frac{\gamma^2}{2(n-1)}\right],
\end{aligned}$$

since $\text{trace}(\mathbf{z}_i\mathbf{z}_i') = 1$, $i = 1, 2, \ldots, n$. Thus,

$$\begin{aligned}
\kappa \equiv E\left[\frac{\|\hat{\mathbf{c}}_1 - \mathbf{c}_1\|^2}{\|\mathbf{c}_1\|^2}\right] &= \frac{E[\text{trace}\{(\hat{\mathbf{c}}_1 - \mathbf{c}_1)(\hat{\mathbf{c}}_1 - \mathbf{c}_1)'\}]}{\mathbf{c}_1'\mathbf{c}_1} \\
&= \frac{\text{trace}(\text{cov}(\hat{\mathbf{c}}_1))}{\mathbf{c}_1'\mathbf{c}_1} + \frac{\text{trace}\{(E[\mathbf{c}_1] - \hat{\mathbf{c}}_1)(E[\hat{\mathbf{c}}_1] - \mathbf{c}_1)'\}}{\mathbf{c}_1'\mathbf{c}_1} \\
&\cong \frac{1}{N}\left[(n-1)\gamma(1+\gamma) + \frac{(1+\gamma)^2}{2} + \frac{\gamma^2}{2(n-1)}\right] \\
&\quad + \frac{\text{trace}\{(E[\hat{\mathbf{c}}_1] - \mathbf{c}_1)(E[\hat{\mathbf{c}}_1] - \mathbf{c}_1)'\}}{\mathbf{c}_1'\mathbf{c}_1}.
\end{aligned} \tag{B.4}$$

Using Lemmas 1 and 2 and the fact that $\mathbf{c}_1 = (\mathbf{c}_1'\mathbf{c}_1)^{1/2}\mathbf{z}_1$,

$$\begin{aligned}
E[\hat{\mathbf{c}}_1] - \mathbf{c}_1 &\cong E\left[\sqrt{\hat{\lambda}_1 - \hat{\sigma}^2}\right]E[\hat{\mathbf{z}}_1] - (\mathbf{c}_1'\mathbf{c}_1)^{1/2}\mathbf{z}_1 \\
&\cong \left[(\mathbf{c}_1'\mathbf{c}_1)^{1/2} - \frac{\text{var}(\hat{\lambda}_1 - \hat{\sigma}^2)}{8(\mathbf{c}_1'\mathbf{c}_1)^{3/2}}\right]\mathbf{z}_1 - (\mathbf{c}_1'\mathbf{c}_1)^{1/2}\mathbf{z}_1 \\
&= \left[\frac{\text{var}(\hat{\lambda}_1 - \hat{\sigma}^2)}{8(\mathbf{c}_1'\mathbf{c}_1)^{3/2}}\right]\mathbf{z}_1
\end{aligned}$$

so that

$$\frac{\text{trace}\{(E[\hat{\mathbf{c}}_1] - \mathbf{c}_1)(E[\hat{\mathbf{c}}_1] - \mathbf{c}_1)'\}}{\mathbf{c}_1'\mathbf{c}_1} \cong \frac{1}{\mathbf{c}_1'\mathbf{c}_1}\left[\frac{\text{var}(\hat{\lambda}_1 - \hat{\sigma}^2)}{8(\mathbf{c}_1'\mathbf{c}_1)^{3/2}}\right]^2. \tag{B.5}$$

From (B.2) and (B.5) it follows that the second term in (B.4) is inversely proportional to N^2, whereas the first term is inversely proportional to N. Asymptotically, the second term can be neglected, which gives (9).

[Received May 1999. Revised May 2000.]

REFERENCES

Akaike, H. (1971a), "Autoregressive Model Fitting for Control," *Annals of the Institute of Statistical Mathematics*, 23, 163–180.

——— (1971b), "Determination of the Number of Factors by an Extended Maximum Likelihood Principle," Research Memorandum 44, Institute Statistical Mathematics, Tokyo.

——— (1987), "Factor Analysis and AIC," *Psychometrika*, 52, 317–332.

Anderson, T. W. (1963), "Asymptotic Theory for Principal Components Analysis," *The Annals of Mathematical Statistics*, 34, 122–148.

Apley, D. W., and Shi, J. (1998), "Diagnosis of Multiple Fixture Faults in Panel Assembly," *ASME Journal of Manufacturing Science and Engineering*, 120, 793–801.

Basilevsky, A. (1994), *Statistical Factor Analysis and Related Methods*, New York: Wiley.

Ceglarek, D., and Shi, J. (1996), "Fixture Failure Diagnosis for Autobody Assembly Using Pattern Recognition," *ASME Journal of Engineering for Industry*, 118, 55–66.

Dowling, M. M., Griffin, P. M., Tsui, K. L., and Zhou, C. (1997), "Statistical Issues in Geometric Feature Inspection Using Coordinate Measuring Machines," *Technometrics*, 39, 3–17.

Etesami, F. (1988), "Tolerance Verification Through Manufactured Part Modeling," *Journal of Manufacturing Systems*, 7, 223–232.

Everitt, B. S. (1993), *Cluster Analysis* (3rd ed.), London: Edward Arnold.

Hu, S., and Wu, S. M. (1992), "Identifying Root Causes of Variation in Automobile Body Assembly Using Principal Component Analysis," *Transactions of the NAMRI*, 20, 311–316.

Hulting, F. L. (1992), "Methods for the Analysis of Coordinate Measurement Data," *Computing Science and Statistics*, 24, 160–169.

Jackson, J. E. (1980), "Principal Components and Factor Analysis: Part I—Principal Components," *Journal of Quality Technology*, 12, 201–213.

——— (1981), "Principal Components and Factor Analysis: Part II—Additional Topics Related to Principal Components," *Journal of Quality Technology*, 13, 46–58.

Johnson, R. A., and Wichern, D. W. (1998), *Applied Multivariate Statistical Analysis* (4th ed.), Upper Saddle River, NJ: Prentice Hall.

Lawley, D. N. (1956), "Tests of Significance for the Latent Roots of Covariance and Correlation Matrices," *Biometrika*, 43, 128–136.

Lowry, C. A., and Montgomery, D. C. (1995), "A Review of Multivariate Control Charts," *IIE Transactions*, 27, 800–810.

Mason, R. L., Champ, C. W., Tracy, N. D., Wierda, S. J., and Young, J. C. (1997), "Assessment of Multivariate Process Control Techniques," *Journal of Quality Technology*, 29, 140–143.

Rissanen, J. (1978), "Modeling by Shortest Description Length," *Automatica*, 14, 465–471.

Schwartz, G. (1978), "Estimating the Dimensions of a Model," *The Annals of Statistics*, 6, 461–464.

Tyler, D. E. (1981), "Asymptotic Inference for Eigenvectors," *The Annals of Statistics*, 9, 725–736.

Wang, Y., Gupta, S., Hulting, F. L., and Fussell, P. S. (1998), "Manufactured Part Modeling for Characterization of Geometric Variations of Automotive Spaceframe Extrusions," *ASME Journal of Manufacturing Science and Engineering*, 120, 523–531.

Wax, M., and Kailath, T. (1985), "Detection of Signals by Information Theoretic Criteria," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 33, 387–392.