Singularity Issues in Fixture Fault Diagnosis for Multi-Station Assembly Processes

Yu Ding
Abhishek Gupta

Department of Industrial Engineering, Texas A&M University, TAMU 3131
College Station, TX 77843-3131

Daniel W. Apley
Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL 60208-3119

1 Introduction

Dimensional variation is a major problem affecting product quality in discrete-part manufacturing. In automotive and aerospace industries, for example, dimensional problems contribute to roughly two-thirds of all quality-related problems during new product launch [1,2]. Dimensional quality of the finished product in panel assembly depends largely on the accuracy of the fixtures used to hold parts. Fixture locators are used extensively in multi-station assembly processes, such as automotive body assembly and aircraft fuselage assembly, in order to provide part support and dimensional reference within a given coordinate system, thereby determining the dimensional accuracy of the final assembly [2,3]. Fixture locators may fail to provide the desired positioning repeatability (relative to tolerances) during production due to gradual deterioration of locators and/or catastrophic events such as broken locators. Considerable efforts have been made in recent years to diagnose fixture errors (sometimes referred to as “fixture faults”) based on product dimensional measurements [4–10]. In these works, the effects of fixture errors on dimensional measurements are represented via the linear diagnostic model

\[
y(t) = Du(t) + v(t), \quad t = 1, \ldots, M, \tag{1}
\]

where \(y(t)\) is the vector of \(n\) measured product features, \(u(t)\) is the vector of \(p\) error sources, \(v(t)\) is the additive noise vector (e.g., sensor noise), \(D = [d_1, d_2, \ldots, d_p]\) is an \(n \times p\) diagnostic matrix linking fixture errors to measurements, \(t\) is the observation number index, and \(M\) is the sample size. Elements in \(u(t)\) are assumed independent random variables because fixture locators are assumed physically independent. It is normally assumed that the sensor system is such that the elements of the noise vector are independent and have equal variance \(\sigma_v^2\). Thus, the covariance matrix of \(v(t)\) is \(\sigma_v^2 I\). The matrix \(D\) can be determined from the relative positions of fixture locators and sensors using standard kinematics analyses [4–9]. Fixture errors manifest themselves as mean shifts and variance in the elements of \(u(t)\). Diagnosing fixture errors is equivalent to estimating the mean and variance components of \(u(t)\) based on the sample of measurement observations \(\{y(t)\}_{t=1}^M\). The focus of this paper is diagnosis of fixture error variance components, as opposed to mean shifts.

Most approaches are based on least-squares (LS) estimation, with a typical procedure as follows. The following procedure has been slightly modified to accommodate a nonzero mean in \(y\) (for the original presentation of this procedure, please refer to, e.g., Apley and Shi [5]): (DP1) Estimate \(\hat{u}(t) = (D^T D)^{-1} D^T (y(t) - \bar{y})\), where the sample mean \(\bar{y} = (\sum_{t=1}^M y(t))/M\); (DP2) Estimate \(\hat{\sigma}_v^2 = (1/(M-1)(n-p)) \sum_{t=1}^M \hat{v}(t)^T \hat{v}(t)\), where \(\hat{v}(t) = y(t) - \bar{y} - \hat{D} \hat{u}(t)\); (DP3) Estimate the variance components of \(u\): \(\hat{\sigma}_v^2 = (1/(M-1)) \sum_{t=1}^M \hat{u}(t)^T \hat{u}(t) - \hat{\sigma}_v^2 (D^T D)^{-1}\), where \(\hat{u}(t)\) represents the \(i\)th element of \(\hat{u}(t)\) and \((D^T D)^{-1}\) is the \((i,i)\)th element of \((D^T D)^{-1}\).

The interpretation is that we first estimate the random deviations \(\{\hat{u}(t)\}_{t=1}^M\), and then use the sample variance of their elements to estimate the variance components of \(u\). The quantity \(\hat{\sigma}_v^2 (D^T D)^{-1}\) is subtracted out in order to eliminate bias due to measurement noise. Because the deviations \(\{\hat{u}(t)\}_{t=1}^M\) are directly estimated via LS, this estimator will be referred to as a “deviation LS estimator.”

In order to implement procedure (DP1) to (DP3) and produce unique estimates, the following conditions are required: i) \(D^T D\) must be of full rank, or equivalently, the columns of \(D\) must be linearly independent; and ii) \(n > p\). These conditions are often satisfied for simple single-station assembly processes when a sufficient number of sensors are used to measure all degrees of freedom of each workpiece. System singularity (\(D^T D\) singular) is often encountered in complex multi-station assembly processes, however, where sensors can only be placed at a downstream station but variation sources are contributed from upstream stations. Singularity is also common in compliant-part assembly processes, where there are modes of rigid-body motion and compliant-part deformation [11]. Section 2 provides an example of singularity in multi-station assembly.

LS modifications using generalized inverses and singular value decomposition [12,13] and partial least squares [14] have been
developed to accommodate singularity and ill-conditioning in parameter estimation [15,16]. The basic idea behind these approaches is to transform the columns of $D$ into a smaller set of linearly independent basis vectors that span the column space of $D$, and then use LS on the reduced-dimensionality problem. Although these are appropriate for many parameter estimation problems, they may lead to erroneous conclusions in fixture diagnosis. The reason is that because the columns of $D$ and the associated variance components represent actual physical phenomena, any reduction in dimensionality and transformation of the columns of $D$ will void its physical meaning. Rong et al. [11] proposed a partial solution to this problem that they termed adjusted least squares. They partitioned $D=[D_1|D_2]$, where $D_1$ consists of the columns of $D$ that are linearly independent of all other columns. In other words, for any linear combination $\alpha_1d_1+\alpha_2d_2+\ldots+\alpha_nd_n$ that equals zero, the coefficients associated with the columns of $D_1$ must be zero. Assuming the set of linearly independent columns is nonempty, their method will provide a unique estimate of the subset of variance components associated with the columns of $D_1$. As in standard least squares, the estimates of the other variance components (associated with the columns of $D_2$) are nonuniquie.

This paper presents an approach that can provide unique estimates of all variance components in situations that satisfy certain diagnosability conditions, even if $D^TD$ is singular. We demonstrate that the deviation LS estimator ignores important information that can be utilized for this purpose and derive a diagnosability condition for the new estimator that is more relaxed than the diagnosability condition for the deviation LS estimator. The relationships between the various estimators are thoroughly discussed.

The format of the remainder of the paper is as follows. Section 2 reviews the modeling procedure for fixture error propagation and explains the cause of singularity in multi-station models. Section 3 introduces a new variance estimator and two modifications that improve performance. Section 4 presents several examples of multi-station assembly processes corresponding to the situations discussed in Section 3. Section 5 concludes the paper.

2 A Variation Model and Singularity in Multi-Station Processes

Previous work has developed a fixture-error propagation model for general multi-station discrete-part manufacturing systems such as rigid-part assembly processes [17–19], compliant-part assembly processes [20], and machining processes [21,22]. In this section, we use a simplified two-station panel assembly process to illustrate the modeling procedure, and explain the cause of system singularity. Details on the modeling procedure can be found in Jin and Shi [17] and Ding et al. [18].

Fixtures in multi-station panel assembly processes generally use an $n$-2-1 layout, consisting of two locating pins and $n$ NC blocks to determine the part/subassembly location and orientation. A typical 3-2-1 (i.e., $n=3$) fixture is shown in Fig. 1. The two locating pins, $P_{4\text{way}}$ and $P_{2\text{way}}$, constrain the three degrees of freedom of a part in the $X$-$Z$ plane, where the 4-way pin restricts part motion in both the $X$- and $Z$-directions, and the 2-way pin restricts part motion in the $Z$-direction. The three NC blocks, $NC_i$, $i=1$, 2, 3, constrain the remaining degrees of freedom of the workpiece in the $Y$-direction. When a workpiece is non-rigid, more than three NC blocks may be needed in order to reduce part deformation. For simplicity, this section illustrates with a 2D example in the $X$-$Z$ plane (in which the part is rigid). More general modeling examples that result in the same linear model structure of Eq. (1) can be found in the aforementioned literature [17–22].

In a multi-station process, 3-2-1 fixtures are repeatedly used at every station to support parts/subassemblies. To illustrate, we refer to the following example throughout the paper. Figure 2 shows a two-station process, which is a segment of the simplified automotive body assembly process from [10]. Three workpieces are welded together at Station I. The first workpiece consists of two components and is a subassembly from the preceding assembly operation. After the welding operation is finished, the whole assembly is transferred to a dedicated measurement station (Station II) for inspection. This simple two-station segment involves all necessary assembly process operations, including positioning, joining, transferring, and inspection. A full-scale assembly process will simply repeat these operations when fabricating complex products.

In this process, each part or subassembly (consisting of several parts) is restrained by a 3-2-1 fixture. Locators being used are marked $P_{1\text{way}}-P_{6\text{way}}$ in Fig. 2 (note that $P_{1\text{way}}$ and $P_{6\text{way}}$ are used to position the whole subassembly in Station II). NC blocks are not shown since we are considering a 2D assembly process. The deviations of a 4-way locator in two directions or the deviation of a 2-way locator in the Z-direction could cause part deviation. They constitute 12 potential fixture errors, numbered 1–9 on Station I and 10–12 on Station II, with arrows indicating their deviation directions.

In such a 2D multi-station process, each part has three degrees of freedom. We use $x_{i,k}$ to denote the deviation state of part $i$ at station $k$,

\[
x_{i,k} = [\delta X_{i,k} \delta Z_{i,k} \delta \alpha_{i,k}]^T,
\]

where $\delta$ is the perturbation operator and $\alpha$ is the orientation angle. Thus the state of the product, which consists of $n_p$ ($n_p=4$ in this process) parts, is represented by

\[
x_k = [x_{1,k} \ldots x_{n_p,k}]^T,
\]

where $x_{i,k}=0$ if part $i$ has not yet appeared at station $k$. 

![Fig. 1 Illustration of a 3-2-1 fixture](image-url)
Random fixture errors on station \( k \) are represented by \( u_k \). Thus, we have \( \mathbf{u}_1=[\delta p_1 \ldots \delta p_9]^T \), and \( \mathbf{u}_2=[\delta p_{10} \delta p_{11} \delta p_{12}]^T \), where \( \delta p_j \) is the deviation associated with fixture error \( j \).

Nine coordinate sensors, denoted by \( m_1 \) through \( m_9 \) in Fig. 2(b), are installed in Station II. Each coordinate sensor measures the position of a part feature (e.g., a corner or hole) in two orthogonal directions (\( X \) and \( Z \)), so that the total number of measurements is \( n=18 \). We use \( y \) to represent the positional deviations detected by sensors at product features. In the above process, since sensors are only available at Station II, we have \( y_1=0 \) and \( y_2=[\delta m_t(X) \delta m_t(Z) \ldots \delta m_t(X) \delta m_t(Z)]^T \), where \( \delta m_t(X \text{ or } Z) \) is the deviation detected at product feature \( j \) in the \( X(\text{or }Z) \) direction.

For the two-station assembly process shown in Fig. 2, the state-space representation [10] of the fixture error propagation model becomes

\[
\begin{align*}
x_1 &= A_0 x_0 + B_1 u_1 + w_1 \\
x_2 &= A_1 x_1 + B_2 u_2 + w_2 \\
y_2 &= C_2 x_2 + v_2
\end{align*}
\] (4)

where the notation corresponds to that in Eq. (4). Note that the subscript represents the station index and the observation index \( t \) is not explicitly included. Matrices \( A_k \), \( B_k \), and \( C_k \) are determined by process design and sensor deployment and \( C_k=0 \) if no sensor is installed at station \( k \) (e.g., \( C_1=0 \) in the above example).

The first equation in (5) is called the state transition equation and \( A_{k-1} \) is accordingly called the state transition matrix because \( A_{k-1} \) links \( x_k \) to \( x_{k-1} \), the states of an assembly over two stations.

In order to express the state space variation model in the same format as Eq. (1), we reformulate Eq. (5) into an input-output linear model by eliminating all intermediate state variables \( x_k \). Assume that \( x_0=0 \) and sensors are placed at Station \( N \). We have

\[
y = \sum_{k=1}^{N} C_N \Phi_{N,k} B_k u_k + \sum_{k=1}^{N} C_N \Phi_{N,k} w_k \quad \text{and} \quad v = v_N,
\] (6)

where \( \Phi_{N,k} = A_{N-1} \cdots A_k \) for \( N>k \) and \( \Phi_{k,k} = I \). Further define \( \Gamma \) and \( \Psi \) as

\[
\Gamma = [C_N \Phi_{N,1} B_1 C_N \Phi_{N,2} B_2 \cdots C_N B_N]
\]
\[
\Psi = [C_N \Phi_{N,1} C_N \Phi_{N,2} \cdots C_N]
\] (7)

This input-output relationship is of the same form as Eq. (1),

\[
y = Du + v.
\] (8)

where \( u = [u_1 \cdots u_k \cdots w_N \cdots w_N] \), \( D = [\Gamma \quad \Psi] \), and the subscript \( N \) (a station index) is dropped from \( y \) and \( v \) without causing any ambiguity. When the higher order terms and process background noise represented by \( w_k \) are negligible, i.e., \( u' = [u_1' \cdots u_k'] \), Eq. (8) further simplifies to

\[
y = \Gamma u + v.
\] (9)

In this paper, we focus on the model in Eq. (9). The approaches developed for Eq. (9) can be easily extended to the model in Eq. (8), because they share the same model structure.

In our example of the two-station assembly process, the measurement station (Station II) is in a well-controlled environment, and we only consider variation sources associated with locators in Station I (i.e., fixture errors 1–9). Thus, \( u_2=0 \), and Eq. (4) becomes
\[ y_2 = \begin{bmatrix} C_2 A_1 B_1 & C_2 B_2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + v_2 = C_2 A_1 B_1 u_1 + v_2, \]  

(10)

where the diagnostic matrix \( D = \Gamma = C_2 A_1 B_1 \). The numerical expression for \( \Gamma \) for this two-station assembly process is provided in Appendix A1. It can be verified that the columns of \( \Gamma \) are linearly dependent (\( \Gamma^T \Gamma \) is singular) so that the deviation LS estimator is not applicable.

The singularity problem is quite common in multi-station systems, especially when we include a comprehensive set of fixture errors in the model, and may be unavoidable regardless of how many sensors are used. The simple example in Fig. 4 illustrates the reasons why. Suppose the assembly deviation shown in Fig. 4(a) is observed at measurement Station \( k+1 \). Any of the Station \( k \) fixture error scenarios illustrated in Figs. 4(b), 4(c), and 4(d) could have resulted in the assembly deviation in Fig. 4(a). Recall that the assembly deviation observed at measurement Station \( k+1 \) is related to the fixture errors incurred at the previous station via the model \( y_{k+1} = C_{k+1} A_k B_k u_k = \Gamma_k u_k \) if the measurement station is free of fixture errors. The preceding observation that any of the three fixture error scenarios could have resulted in the same observed assembly deviation means that, given \( y_{k+1} \), there is no unique solution for \( u_k \). Mathematically, this means that the columns of \( \Gamma_k \) are linearly dependent, so that \( \Gamma_k^T \Gamma_k \) is singular. Because these conclusions clearly hold regardless of how many sensors we add at the measurement Station \( k+1 \), the only way to avoid a nonsingular system in this case is to add sensors to the assembly Station \( k \). If this is not possible, then the deviation LS estimator cannot be applied.

This singularity problem in multi-station assembly processes was also illustrated in the examples presented by Carlson et al. [8]. With the fixture errors that are included in the error propagation models developed in [17,18,20–22], the diagnostic matrices (\( \Gamma_k \)'s) are all less than full rank.

### 3 Variation Diagnosis

#### 3.1 Variance LS Estimator

Section 2 presented an example for multi-station assembly in which \( A_k \) and \( \Gamma \) are both singular. Thus, the deviation LS estimator outlined in (DP1)–(DP3) cannot be applied. This section develops an alternative approach that circumvents this problem. Taking the covariance matrix for both sides of Eq. (9) gives

\[ \Sigma_g = \Gamma \Sigma \Gamma^T + \sigma_p^2 I, \]

(11)

where \( \Sigma_g \) is the covariance matrix of a random vector. Since fixture errors associated with different fixture locators are assumed physically independent, \( \Sigma_g = \text{diag} [\sigma_1^2, \sigma_2^2, \cdots, \sigma_p^2] \) is diagonal, where \( p \) is the number of fixture errors included in the model. Define

\[ \mathbf{\sigma}^2 = [\sigma_1^2, \sigma_2^2, \cdots, \sigma_p^2]^T \]

as the vector of variance components to be estimated.

Equation (11) can be written as

\[ \Sigma_g = \sum_{i=1}^{p} (\gamma_i \gamma_i^T) \sigma_i^2 + \sigma_p^2 I \]

(12)

where \( \gamma_i \) is the \( i \)th column vector of \( \Gamma \). In practice, the population covariance \( \Sigma_g \) is estimated by the sample covariance matrix

\[ S_g = \frac{1}{M-1} \sum_{i=1}^{M} (y(t) - \bar{y})(y(t) - \bar{y})^T = \Sigma_g + \mathbf{E}, \]

(13)

where \( \mathbf{E} \) denotes the estimation error matrix. If we define \( V_i = \gamma_i \gamma_i^T \) for \( i = 1, \ldots, p \), \( V_{p+1} = \mathbf{I} \), and \( \sigma_{p+1}^2 = \sigma_p^2 \), Eqs. (12) and (13) become

\[ S_g = \sum_{i=1}^{p+1} V_i \sigma_i^2 + \mathbf{E}. \]

(14)

In light of this, one approach for estimating the variance components is to choose \( \mathbf{\sigma}^2 \) (the “\( \sim \)” symbol denotes an estimate) to minimize the sum of the squares of the elements of the error matrix \( S_g = \sum_{i=1}^{p+1} V_i \sigma_i^2 \). For square matrices \( \mathbf{A} \) and \( \mathbf{B} \) of compatible dimension, define the matrix inner product \( \langle \mathbf{A}, \mathbf{B} \rangle = \text{tr}(\mathbf{A}^T \mathbf{B}) \) and the associated matrix norm \( \| \mathbf{A} \|^2 = \langle \mathbf{A}, \mathbf{A} \rangle \), which is exactly sum of the squares of the elements of \( \mathbf{A} \). Using standard results for least squares estimation in inner-product spaces [23], the estimates in this case must satisfy the so-called normal equations

\[ \mathbf{P}\mathbf{\sigma}^2 = \mathbf{b}, \]

(15)

where the notation is as follows. \( \mathbf{P} \) is the Gram matrix, defined so that the \( i,j \)-th row, \( j \)-th column element is \( \langle \mathbf{V}_i, \mathbf{V}_j \rangle \) for \( 1 \leq i, j \leq p+1 \). The \( (p+1) \)-length column vector \( \mathbf{b} \) is defined so that its \( i \)-th element is \( \langle \mathbf{V}_i, \mathbf{S}_g \rangle \). For the particular inner product defined above, it can be verified that

\[
\mathbf{P} = \begin{bmatrix}
(\gamma_1^T \gamma_1)^2 & \cdots & (\gamma_1^T \gamma_p)^2 & \gamma_1^T \gamma_n \\
\vdots & \ddots & \vdots & \vdots \\
(\gamma_p^T \gamma_1)^2 & \cdots & (\gamma_p^T \gamma_p)^2 & \gamma_p^T \gamma_n \\
\gamma_1^T \gamma_1 & \cdots & \gamma_p^T \gamma_p & n
\end{bmatrix}
\]

and

\[
\mathbf{b} = \begin{bmatrix}
\gamma_1^T \mathbf{S}_g \gamma_1 \\
\vdots \\
\gamma_p^T \mathbf{S}_g \gamma_p \\
\text{tr}(\mathbf{S}_g)
\end{bmatrix}.
\]

(16)
When $\Pi$ is nonsingular, or equivalently when the matrices $\gamma_1 \gamma_1^T, \ldots, \gamma_n \gamma_n^T$, and $I$ are linearly independent, $\hat{\mathbf{b}} = \Pi^{-1}\mathbf{b}$ is a unique solution to Eq. (15). We refer to this approach as the “variance LS estimator.”

### 3.2 Diagnosability of Deviation LS Estimator and Variance LS Estimator

The diagnosability condition required for a variance LS estimator is different from that required for a deviation LS estimator. For the variance LS estimator, the matrix $\Pi$ must be of full rank in order for Eq. (15) to yield a unique solution. For the deviation LS estimator, the matrix $\Pi^T \Pi$ must be of full rank and also $n > p$ in order to yield a unique solution. To more clearly illustrate the difference, consider the simplified example shown in Fig. 5, in which each part has only one degree of freedom and can only translate (no rotation) in the Z-direction.

Three locators are used to position the three panels at Station I, and their instantaneous position errors are denoted as $\mathbf{u}_1 = [\delta p_1, \delta p_3, \delta p_1]^T$. After the joining operation are finished, the three parts become one subassembly and it is transferred to Station II for measurement. The state vectors are $\mathbf{x}_k = [\delta z_{1,k}, \delta z_{2,k}, \delta z_{3,k}]^T (k = 1, 2)$ and the measurement vectors are $\mathbf{y}_1 = \mathbf{0}$ and $\mathbf{y}_2 = [\delta m_1(Z), \delta m_2(Z), \delta m_3(Z)]^T$. At Station II, the locating hole on part 2 is used to position the whole assembly. The locator on Station II is assumed to be free of positioning errors (i.e., $\mathbf{u}_2 = \mathbf{0}$).

When this three-panel assembly is transferred to Station II, it undergoes a translation by the amount $-\delta z_{2,1}$, which can be represented as

$$
\mathbf{x}_2 = \mathbf{x}_1 + \begin{bmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{x}_1 = \mathbf{A}_1 \mathbf{x}_1.
$$

Because $\mathbf{y}_2 = \mathbf{A}_1 \mathbf{u}_1 + \mathbf{v}$ at Station II, the state space model becomes $\mathbf{x}_2 = \mathbf{A}_1 \mathbf{x}_1$ and $\mathbf{y}_2 = \mathbf{A}_1 \mathbf{u}_1 + \mathbf{v}$. The linear diagnostic model for this two-station process becomes

$$
\mathbf{y}_2 = \mathbf{A}_1 \mathbf{u}_1 + \mathbf{v}.
$$

Relating this to the model in Eq. (1), we have $\mathbf{D} = \Pi = \mathbf{A}_1$, where $\mathbf{A}_1$ shown above is clearly singular. However,

$$
\Pi = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 4 & 1 & 2 \\ 0 & 1 & 1 & 1 \\ 1 & 2 & 1 & 3 \end{bmatrix}
$$

is of full rank. Consequently, the variance vector $\mathbf{a}_2$ can be diagnosed using the variance LS estimator in Eq. (15), but not using the deviation LS estimator.

An explanation for diagnosability of variance vector $\mathbf{a}_2$ using the variance LS estimator is apparent from the covariance matrix

$$
\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_1^2 & 0 & \sigma_2^2 \\ 0 & \sigma_2^2 & 0 & 0 \\ \sigma_2^2 & 0 & \sigma_2^2 + \sigma_3^2 + \sigma_4^2 \end{bmatrix}
$$

for this single degree of freedom assembly. The diagonal elements in $\Sigma$ provide information regarding the summation of the fixture error variance components $\sigma_1^2$, $\sigma_2^2$, and $\sigma_3^2$, as well as the noise variance $\sigma_4^2$. The non-zero off-diagonal element, which is the covariance between $\delta m_1(z)$ and $\delta m_2(z)$, provides extra information. In Station II, $\delta m_1(z) = \delta p_1 - \delta p_2$ and $\delta m_2(z) = \delta p_3 - \delta p_2$, so that $\text{cov}(\delta m_1(z), \delta m_2(z)) = \text{var}(\delta p_2) = \sigma_2^2$ (recall that $\delta p_1$ and $\delta p_3$ are assumed independent). This extra piece of information is utilized by the variance LS estimator so that variance components are diagnosable. The discussion so far did not include the fixture error $\mathbf{u}_2$ in Station II. A straightforward extension that includes $\mathbf{u}_2$ would result in the same conclusion.

Another way of viewing the difference between these two estimators is the following: a variance LS estimator first calculates covariance matrices of $\mathbf{u}$ and $\mathbf{y}$, and then applies the LS criterion on the sample covariance matrices, whereas a deviation LS estimator first calculates the LS estimates for the individual error vectors $\{\hat{\mathbf{u}}(i)\}_{i=1}^M$, and then calculated the variances of $\mathbf{u}$ from $\{\hat{\mathbf{u}}(i)\}_{i=1}^M$. Because estimating $\{\hat{\mathbf{u}}(i)\}_{i=1}^M$ requires more information than simply estimating its covariance matrix, it is not surprising that the deviation LS estimator requires a stronger diagnosability condition than the variance LS estimator. This is stated in the following theorem, the proof of which is included in Appendix A2.

**Theorem 1.** If $\Pi^T \Pi$ is of full rank and $n > p$, then $\Pi$ is of full rank.

The significance of Theorem 1 is that a unique variance LS estimator exists whenever a unique deviation LS exists. The converse, however, is not necessarily true. As illustrated in the preceding example, there are situations where the variance LS estimator is unique but the deviation LS estimator is not.

### 3.3 Effect of System Structure Modeled by $\Gamma$ on Variance Estimation

The performance of the deviation LS estimator will deteriorate for “ill-conditioned” systems, even if $\Pi^T \Pi$ is not exactly singular. The most common criteria [24] used to quantify how ill-conditioned a system is include $\text{tr}((\Gamma^T \Gamma)^{-1})$, $\text{cond}(\Gamma^T \Gamma)$, and $\text{det}(\Gamma^T \Gamma)^{-1}$, where cond() and det() are the condition number and the determinant of a matrix, respectively. These three measures are related to each other through the eigenvalues of $\Gamma^T \Gamma$, which we denote $\{\lambda_i\}_{i=1}^n$. The relationship is...
The larger these measures are, the more ill-conditioned the system is. Since $\Gamma^T \Gamma$ is non-negative definite, all eigenvalues must be non-negative. The system is singular when these measures are infinite, or equivalently, when one or more eigenvalues are exactly zero. Throughout the remainder of the paper, we use $\text{tr}(\Gamma^T \Gamma)^{-1}$ as the measure of how ill-conditioned a system is.

The following simulations were used to investigate the extent to which both estimators are affected as the system changes from being well-conditioned to being ill-conditioned. The following parameters were used: $n=6$, $p=3$, $M=50$, $\sigma^2_0=0.25$, $\{\sigma^2_i\}_{i=1}^5=[1,4,9]$, and

$$\Gamma^T = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & -1 & a & a & -1 & -1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix},$$

where $a$ was varied from 1 to 0.1 so that the value of $\text{tr}(\Gamma^T \Gamma)^{-1}$ changes accordingly. For each $a$, a Monte Carlo simulation with $K=5,000$ replicates was conducted for each estimator.

The performance measure is the mean square error (MSE) of the estimates

$$\text{MSE} = \frac{1}{p+1} \sum_{j=1}^{p+1} \left( \frac{1}{K} \sum_{i=1}^K (\hat{\sigma}^2_j - \sigma^2_j)^2 \right),$$

where $\hat{\sigma}^2_{j,k}$ is the estimate of $\sigma^2_j$ for the $k$th replicate.

Figure 6 shows the MSEs of two estimators vs. the values of $\text{tr}(\Gamma^T \Gamma)^{-1}$ as $a$ is varied. From this, the following observations can be made: (a) The performance of the deviation LS estimator deteriorates rapidly (the MSE increases) as $\text{tr}(\Gamma^T \Gamma)^{-1}$ increases. In contrast, the MSE of the variance LS estimator is largely independent of $\text{tr}(\Gamma^T \Gamma)^{-1}$. Clearly, the variance LS estimator is less sensitive to linear dependencies in the system. (b) Although the variance LS estimator performs better for ill-conditioned systems, the deviation LS estimator has a smaller MSE value when $\text{tr}(\Gamma^T \Gamma)^{-1}$ is small (e.g., less than 10). We may conclude that the deviation LS estimator can outperform a variance LS estimator for a well-conditioned system. For an ill-conditioned system, however, the variance LS estimator will perform substantially better. In the following subsection, we present a modified version of the variance LS estimator that performs uniformly better than the deviation LS estimator.

Although the performances of these two estimators will generally differ, the estimators are actually equivalent in the special case that all columns of $\Gamma$ are orthogonal, i.e., when $\gamma_i \gamma_j = 0$, $\forall i \neq j$. This obviously requires that $\Gamma^T \Gamma$ is of full rank. This is stated as Theorem 2, the proof of which is included in Appendix A3.

**Theorem 2.** If $n>p$, and $\gamma_i \gamma_j = 0$, $\forall i \neq j$, the variance LS estimator in Eq. (15) is the same as the deviation LS estimator described in (DP1)–(DP3).

### 3.4 Modified Procedures to Enhance the Performance of the Variance LS Estimator

One observation in the previous section was that the variance LS estimator may perform worse than the deviation LS estimator for a well-conditioned system. This motivates the following algorithm for improving the performance of the variance LS estimator. A more general version of the algorithm was originally proposed by Anderson [25] as an approximate maximum likelihood method. For the present case, the expressions required in Step 2 of the algorithm simplify considerably to those shown below. The algorithm iterates over the following two steps until convergence.

**Modified Procedure 1 (MP1)**

1. Based on the estimate $\hat{\sigma}^2$ at the previous iteration, calculate the following estimate of the covariance matrix [see Eq. (12)]

$$\hat{\Sigma}_\gamma = \sum_{i=1}^p (\gamma_i \gamma_i^T) \hat{\sigma}^2 + \hat{\sigma}^2 I$$

2. Solve the equation $\Pi^* \hat{\sigma}^2 = b^*$ for the new estimate $\hat{\sigma}^2$ at the current iteration, where

$$\Pi^* = \begin{bmatrix} (\gamma_1 \hat{\Sigma}_\gamma^{-1} \gamma_1) & \cdots & (\gamma_1 \hat{\Sigma}_\gamma^{-1} \gamma_p) \\ \vdots & \ddots & \vdots \\ (\gamma_p \hat{\Sigma}_\gamma^{-1} \gamma_1) & \cdots & (\gamma_p \hat{\Sigma}_\gamma^{-1} \gamma_p) \\ (\gamma_1 \hat{\Sigma}_\gamma^{-1} \gamma_1) & \cdots & (\gamma_p \hat{\Sigma}_\gamma^{-1} \gamma_p) \\ \gamma_1 \hat{\Sigma}_\gamma^{-2} \gamma_1 & \cdots & \gamma_p \hat{\Sigma}_\gamma^{-2} \gamma_p & \text{tr}(\hat{\Sigma}_\gamma^{-2}) \end{bmatrix}$$

and

$$b^* = \begin{bmatrix} \gamma_1 \hat{\Sigma}_\gamma^{-1} \gamma_1 \\ \vdots \\ \gamma_p \hat{\Sigma}_\gamma^{-1} \gamma_1 \\ \gamma_1 \hat{\Sigma}_\gamma^{-1} \gamma_1 \\ \text{tr}(\hat{\Sigma}_\gamma^{-2}) \end{bmatrix}.$$
At each iteration, the MP1 algorithm is therefore the least squares solution that minimizes the norm of the transformed error matrix \(
abla_{i,j} = S_{i/j}^{-1/2} S_{i,j}^{-1/2} V_{i,j}^{-1/2} \). The transformation by \( S_{i,j}^{-1/2} \) can be viewed as weighted least squares [26], since the elements of \( S_{i,j}^{-1/2} V_{i,j}^{-1/2} \) are uncorrelated with equal variance. Note that the elements of the untransformed error matrix \( E \) are neither uncorrelated nor have equal variance.

The MP1 algorithm can substantially improve the performance of the variance LS estimator, especially for a well-conditioned system. In fact, Anderson [25] has shown that the MP1 estimator is asymptotically efficient. The improvement is illustrated via simulation using the same \( \Gamma \) as in Section 3.3 (except that we now use a larger sample size of \( M = 100 \)). The simulation results are shown in Fig. 7(a), which shows that the MP1 estimator has uniformly smaller MSE values than the other estimators.

If \( S_{i,j} \) is positive definite at each step of the iteration, \( \mathbf{I}^\pi \) will be full rank if and only if \( \mathbf{I} \) is full rank. Thus, the conditions required for the MP1 algorithm to produce a unique least squares estimate are identical to the conditions required by the variance LS algorithm, provided that \( \mathbf{S}_{i,j} \) remains positive definite at each iteration. This is not guaranteed, however, because elements of \( \hat{\sigma}^2 \) may take on negative values. This is more likely to occur when the sample size is relatively small and the true variance components are close to zero. The negativity of estimates is a problem in almost all of the variance estimation algorithms (the deviation LS estimator cannot avoid negative estimates, either). The most popular approach to enforce non-negativity is to replace the negative elements in \( \hat{\sigma}^2 \) with zeros. As long as \( \sigma^2_{i,j} > 0 \), \( \hat{\sigma}_{i,j} \) is then guaranteed to be positive definite. If \( \sigma^2_{i,j} \) becomes negative and is replaced by 0, the pseudoinverse of \( \hat{\Sigma}^2 \) can be used in step 2.

Rao and Kleffe developed a different variance estimation approach [[27], Eq. 9.1.8] that avoids negative estimates. Their iterative algorithm will give positive variance component estimates, provided that the initial estimates are positive. This procedure is less intuitive and its development is rather mathematically involved. Consequently, we simply present the final form of the algorithm for practical use. Note that the form of the algorithm in Rao and Kleffe’s Eq. 9.1.8 is much more complex than the algorithm below. The reason is that for the particular model in Eq. (1), the expressions simplify considerably to those shown below.

**Modified Procedure 2 (MP2)**

(M1) Select an initial \( \hat{\sigma}_{i,j}^2 = [\hat{\sigma}_{1,0}^2 \cdots \hat{\sigma}_{p,0}^2] \) with all positive values.

(M2) Calculate \( \hat{\Sigma}_{e,0} = \Gamma \hat{\Sigma}_{a,0} \Gamma^T + \hat{\sigma}_{e,0}^2 \mathbf{I} \), where \( \hat{\Sigma}_{a,0} = \text{diag}(\hat{\sigma}_{1,0}^2 \cdots \hat{\sigma}_{p,0}^2) \).

(M3) Solve the following set of linear equations for \( j = 0 \) and 1.

\[
\hat{\sigma}_{i,j}^2 = \hat{\sigma}_{i,j}^2 \left( (\Gamma^T \Gamma)^{-1} \right)^{-1} \cdot \left( \Gamma^T \hat{\Sigma}_{e,0}^{-1} \Gamma \right)^{-1} \cdot \left( \hat{\Sigma}_{e,j}^{-1} \right)^{-1} \cdot \left( \hat{\Sigma}_{e,j}^{-1} \right) \cdot \left( \hat{\Sigma}_{e,j}^{-1} \right)^{-1} \cdot \left( \hat{\Sigma}_{e,j}^{-1} \right)^{-1} \cdot \hat{\sigma}_{e,0}^2 \text{ for } \hat{\sigma}_{e,0}^2 > 0.
\]

The \( \hat{\sigma}_{i,j}^2 \) and \( \hat{\sigma}_{e,0}^2 \) in MP2 will remain positive as long as the initial values of \( \hat{\sigma}_{i,j}^2 \) are chosen positive. The usual choice is to let \( \hat{\sigma}_{i,0}^2 = \hat{\sigma}_{e,0}^2 = 1 \), for \( i = 1, \ldots, p \). The solution \( \left[ \hat{\sigma}_{i,0}^2 \cdots \hat{\sigma}_{p,0}^2 \right]^T \) is the final estimate. The results for the MP2 estimator in the situation described in the preceding simulation were also included in Fig. 7. For the relatively small sample size of \( M = 25 \), the MP2 estimator outperforms the other estimators. For the more typical sample size of \( M = 100 \), however, the MP1 estimator performs better than the MP2 estimator. The reason is that the MP2 estimator forces a bias in order to make \( \hat{\sigma}^2 \) positive, and this bias does not disappear as sample size increases. In contrast, the MP1 estimator is unbiased and consistent, meaning that its variance approaches zero as sample size increases. Consequently, the MP2 estimator is only recommended if sample size is very small.

**4 Examples**

In this section, the estimators are applied to fixture error diagnosis in various automotive body assembly problems. Monte Carlo simulations with 5,000 replicates were conducted in a MATLAB environment, and fixture errors were assumed to follow a normal distribution in all cases. For detailed descriptions of the processes, the reader is referred to the various references cited below. For convenience, the diagnostic matrix \( \Gamma \) is provided below for each situation.

**4.1 Assembly System With an Orthogonal Diagnostic Matrix.** The automotive assembly process was described in considerable detail in Apley and Shi [5]. In Section 5 of their paper, they apply the deviation LS estimator to diagnosing errors in fixtures that locate the side frames of a car body. They assumed the linear structured model of Eq. (9) to represent the effects of fixture errors on dimensional measurements. There were 14 measurements \( (n = 14) \) and two potential fixture errors \( (p = 2) \). The matrix \( \Gamma \) (which is the \( C \) matrix in their paper) was through kinematics analysis determined to be...
\[ \Gamma^2 = \begin{bmatrix} .354 & .354 & .354 & .354 & .354 & .354 & .354 & .354 & .354 & 0 & 0 & 0 & 0 & 0 & 0 \\ -.057 & -.026 & 0 & -.004 & .046 & -.087 & -.024 & .043 & .187 & .361 & .053 & .495 & .536 \end{bmatrix}. \] (25)

This \( \Gamma \) matrix is of full column rank and \( n > p \), suggesting that the variance LS estimator and the deviation LS estimator can both be applied. We also have \( \text{tr}(\Gamma^2 \Gamma^{-1}) = 1.99 \), indicating that the system is well-conditioned.

For this side-frame assembly system, both deviation LS estimator and variance LS estimator are used to estimate the variance components associated with fixture errors. Five different sample sizes were used (\( M = 5,10,25,50,100 \)) in the simulation. From the MSE values shown in Fig. 8, it can be seen that the two estimators have almost identical performance in this example. The reason is that the two columns of \( \Gamma \) are almost orthogonal (\( \gamma_1 \gamma_2 = 0.018, \gamma_1 \gamma_1 = 1.0025, \) and \( \gamma_2 \gamma_2 = 0.9999 \)). This agrees with Theorem 2, which states that the two estimators are equivalent when the columns of \( \Gamma \) are orthogonal.

### 4.2 Assembly System With a Non-Orthogonal Diagnostic Matrix.

Many engineering systems do not result in an orthogonal \( \Gamma \) matrix, in which case the performance of the deviation LS and variance LS estimators will differ. For example, the \( \Gamma \) matrix used in Section 4 of Apley and Shi [5] is

\[ \Gamma^2 = \begin{bmatrix} .93 & 0 & -.93 & .93 & 0 & .647 & -.370 & 0 & .647 \\ .577 & 0 & 0 & .577 & 0 & 0 & .577 & 0 & 0 \\ -.120 & .843 & -.120 & 0 & -.120 & .482 & 0 & -.120 \end{bmatrix}. \] (26)

The columns of which are not orthogonal. In this case, \( n = 9, p = 3 \), and \( \text{tr}(\Gamma^2 \Gamma^{-1}) = 3.5 \), implying the system is relatively well-conditioned.

Monte Carlo simulations were again conducted, but this time with a sample size of \( M = 15 \). A comparison of the deviation LS estimator, the variance LS estimator, and the MP2 estimator (due to the small sample size) is shown in Table 1. The quantity \( \Sigma_{g+i}^a \) represents the average sample variance of the estimators for comparison with the MSE. We found that the estimator from MP2 demonstrates slightly more bias than the other two, but has smaller dispersion. Based on the MSE criterion, the MP2 estimator performed the best, followed by the deviation LS estimator.

### 4.3 Assembly System With a Singular Diagnostic Matrix

We next apply the variance estimators to the two-station example introduced in Section 2.2. The \( \Gamma \) matrix for this model is given in Appendix A1. Because the system is singular with \( \text{tr}(\Gamma^2 \Gamma^{-1}) = \infty \), the deviation LS estimator cannot be used here. It can be verified that \( \Gamma^* \) is full rank, so that the variance LS estimator and its modified versions are applicable.

As discussed in Section 2, we only consider fixture errors associated with Station 1. Hence \( p = 9 \) and \( n = 18 \). Simulations were carried out using a sample size of \( M = 100 \). The variance LS estimator of Eq. (15) and the MP1 and MP2 estimators were compared in this example, and the results are shown in Table 2. In this example, the variance LS estimator and MP1 estimator perform comparably, although the latter has slightly smaller MSE and dispersion. This is consistent with the results shown in Fig. 7(a) as \( \text{tr}(\Gamma^* \Gamma^{-1}) \) increases. The MP2 procedure has the smallest MSE and dispersion among the three. But it also has quite noticeable bias.

Although the MP2 estimator outperformed the MP1 estimator in this example, our experience indicates this is more an exception than the norm. For example, in Fig. 7(a), the MP1 has a smaller MSE value. As another example, suppose we modify the two-station assembly example considered in the preceding paragraph so that we are now only interested in diagnosing the Z-direction fixture errors. The \( \Gamma \) matrix in this case is the same as the \( \Gamma \) matrix given in Appendix A1, except that we remove column 1, 4, and 7. It can be verified that the new \( \Gamma \) matrix is also less than full rank. Repeating the above simulations but with the new \( \Gamma \) matrix, the MSEs for the MP1 and MP2 estimators are 1.197 and 1.215. Thus, the MP1 estimator is slightly more effective than the MP2 estimator in this case.

### Table 1: Comparison of three estimators for the linear system with \( \Gamma^* \) as in Eq. (26)

<table>
<thead>
<tr>
<th>( \sigma^2 )</th>
<th>Deviation LS estimator</th>
<th>Variance LS estimator</th>
<th>MP2</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \bar{\sigma}^2 )</td>
<td>[1.0347 4.0276]</td>
<td>[1.0015 4.0419]</td>
<td>[1.0050 3.4421]</td>
</tr>
<tr>
<td>( \Sigma_{g+i}^a )</td>
<td>[3.9965 0.9861]</td>
<td>[3.9925 0.9854]</td>
<td>[3.3303 1.0358]</td>
</tr>
</tbody>
</table>

The "true" value of \( \sigma^2 \) used in the simulation is [1.4, 4.1].

![Fig. 8](image-url) MSE for the linear system with \( \Gamma^* \) as in Eq. (25)
5 Concluding Remarks

Singularity is a common problem in engineering systems, in which case the traditional least-squares estimation method cannot be applied effectively. This paper presents a new diagnosability condition and a variance LS estimator that takes into account the covariance between error terms and results in diagnosability for systems that are not diagnosable using traditional LS methods. Two modified versions of the algorithm were also presented to improve the performance of the variance LS estimator.

We note that the presented methods typically require a random sample of 25–50 units. For a dynamic process with tool wear, the process data are inherently autocorrelated. However, since 25–50 units typically translates to production periods of one hour or less, the sampling period will generally be too small to observe any noticeable tool wear effects. Consequently, the methods should still be applicable to diagnosing other types of fixture errors in processes that also experience relatively slow tool wear dynamics. For processes with faster tool wear dynamics, recursive estimation methods would need to be developed.

We also point out that the methods are for variance component estimation, as opposed to mean component estimation. Our experience has been that the autobody industry views fixture error variance as more problematic than mean shifts. A sustained, consistent deviation from nominal (i.e., a mean shift) can often be compensated quite easily by process engineers via shimming and other adjustments. In contrast, variation is much more difficult to compensate and requires either some form of on-line feedback control or the removal of the variation root cause. The methods presented in this paper are intended to be a tool to aid in detecting, identifying, and, ultimately, eliminating root causes of random variation.

The examples in this study have been exclusively for fixture error diagnosis in multi-station assembly processes. However, all of the results and conclusions should also hold for other types of error sources and multi-station manufacturing processes, provided that the linear structured model adequately represents the effects of the error sources on the process and product measurements.

Acknowledgment

This research was partially supported by the NSF grants DMI-0217481 and DMI-0093580. The authors also gratefully acknowledge the valuable comments and suggestions from the associate editor and referees.

Appendices

A1 Expression of Matrices for Example in Section 2.

\[
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0.0005 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -0.0005 & -0.2392 \\
-1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0.5550 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & -0.4450 & -222.49 \\
0 & 0.0005 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & -0.0005 & -0.2392 \\
-1 & -0.2153 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0.2153 & 107.655 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & -0.7608 & -380.38 \\
0 & 0.0005 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.0005 & -0.2392 \\
-1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -0.0005 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.2392 & -380.38 \\
0 & 0.0005 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.0005 & 0.7608 \\
\end{pmatrix}
\]

\[A_1\] \[12 \times 12\]
a set of scalars $a_g$'s, and at least one of the $g$s is full rank. It follows that $g_1 g_2 g_2 + \ldots + g_p g_p = 0$, and at least one of the $a_g$s (say $a_g$) is nonzero. Post-multiplying the preceding equation by $g$ gives $g_i (g_1 g_2 \ldots g_p) + g_2 (g_2 g_2 + \ldots + g_p g_p) = 0$. Because at least one of the coefficients $g_i (g_2 \ldots g_p)$ is nonzero, this implies that the vectors $(g_1 g_2, \ldots, g_p)$ are linearly dependent. Their Gram matrix $A$ as its $(i, i)$-th element and $\Sigma_{i=1}^p a_i g_i = \Gamma \Omega \Gamma^T$, where $\Omega = \text{diag}(\omega_1, \ldots, \omega_p)$ and $\omega_i, i = 1, \ldots, p$, is an arbitrary real number. Then, we have

$$\sum_{i=1}^p \frac{1}{g_i^2} \Gamma \gamma_i = \Gamma (\Gamma^T \Gamma)\gamma_i = \Gamma \Omega \Gamma^T$$

(a4)

Given all these results, we can write Eq. (a3) as

$$\hat{\sigma}_e^2 = \frac{1}{(n-p)} \text{tr}(I - \Gamma \Gamma^T) S_p.$$  

(a5)

It can be further shown that this $\hat{\sigma}_e^2$ is the same as the one in the deviation LS estimator. Substitute $\hat{v}(t) = y(t) - \bar{y} - \Gamma \hat{u}(t)$ and $\hat{u}(t) = \Gamma^T (y(t) - \bar{y})$ into $\hat{\sigma}_e^2 = \sum_{i=1}^M \hat{v}(t) \hat{v}(t)/((M-1)(n-p))$. It turns out that

$$\hat{\sigma}_e^2 = \frac{1}{(M-1)(n-p)} \sum_{i=1}^M (y(t) - \bar{y}) (I - \Gamma \Gamma^T) (y(t) - \bar{y})$$

$$= \frac{1}{(M-1)(n-p)} \text{tr} \left( \sum_{i=1}^M (y(t) - \bar{y}) (I - \Gamma \Gamma^T) (y(t) - \bar{y})^T \right)$$

$$= \frac{1}{(n-p)} \text{tr} \left( I - \Gamma \Gamma^T \right) \cdot \frac{1}{M-1} \sum_{i=1}^M (y(t) - \bar{y})(y(t) - \bar{y})^T$$

$$= \frac{1}{(n-p)} \text{tr}((I - \Gamma \Gamma^T) S_p).$$

(a6)

We can solve $\sigma_i^2$'s in terms of $\hat{\sigma}_e^2$ from the first equation and substitute it into the second equation. Then, we have

$$\hat{\sigma}_e^2 = \frac{1}{(n-p)} \text{tr}(I - \Gamma \Gamma^T) S_p.$$  

(a5)

After obtaining the solution of $\hat{\sigma}_e^2$, we can substitute it into (a2) to solve for $\sigma_i^2$ as
\[ \hat{\sigma}_i^2 = \frac{1}{(y_i')^2} \operatorname{tr}(\gamma_i' S_i) - \hat{\sigma}_i^2, \quad i = 1, 2, \ldots, p. \]  

(a7)

Recall that \( \operatorname{tr}(\gamma_i' S_i) = \gamma_i' S_i \gamma_i \) and \( 1/\gamma_i' \gamma_i \) is the \((i, i)\) element of \((I^T \Gamma)^{-1}\). Then, \( \gamma_i' \gamma_i \) is the \(i\)th row of \((I^T \Gamma)^{-1} I^T\). We can further write (a7) as

\[ \hat{\sigma}_i^2 = \frac{1}{\gamma_i' \gamma_i} S_i - \hat{\sigma}_i^2, \quad i = 1, 2, \ldots, p, \]  

(a8)

where \( \gamma_i' \gamma_i \) is the \(i\)th row of \((I^T \Gamma)^{-1} I^T\). The second term \( \hat{\sigma}_i^2 (I^T \Gamma)^{-1} \) in the right-hand side of the above equation is the same as the one in (DP3) in Section 1. We shall show that \( \sum_{i=1}^{M} u_i(t)^2 / (M - 1) \) is the same as \( \gamma_i' S_i (\Gamma_i^+)^T \). In fact, \( u_i(t) = ((I^T \Gamma)^{-1} I^T)_{i} (y(t) - \bar{y}) = \Gamma_i^+ (y(t) - \bar{y}) \). Then,

\[ \frac{1}{M - 1} \sum_{i=1}^{M} \hat{u}_i(t)^2 = \frac{1}{M - 1} \sum_{i=1}^{M} \left( \gamma_i' S_i (\Gamma_i^+)^T \right) \]  

\[ \Gamma_i^+ (y(t) - \bar{y})(y(t) - \bar{y})^T \]  

\[ = \Gamma_i^+ \cdot S_i \cdot (\Gamma_i^+)^T. \]  

(a9)

This completes the proof. \( \diamond \)

References


