

주성분 분석을 이용한 측정시스템의 경제적 평가

Economic Evaluation of Measurement System by Principal Component Analysis

강충오* · 변재현**

Chung-Oh Kang* · Jai-Hyun Byun**

Abstract

It is very important to have a satisfactory measurement system, since it is useless to try to improve the manufacturing process without an adequate measurement system. Therefore, evaluation of the measurement system is the first step for the quality improvement of the manufacturing process. To estimate the measurement error we must conduct a controlled gage repeatability and reproducibility(gage R&R) study. Many manufacturers use a gage or instrument to measure multiple dimensions for the overall quality of the manufactured parts. In this case, it is necessary to estimate the gage R&R for multiple dimensions. When a gage measures a large number of dimensions of a part, it is very time-consuming and costly to measure all the dimensions. In this paper we propose the use of the principal component analysis method to identify a few principal components out of the original multivariate measurement capability to explain most of the measurement system variation pattern.

1. 서론

갈수록 치열해지는 국경없는 경쟁사회에서 기업은 제품의 고급품질을 통해 경쟁력강화를 서두르고 있고, 통계적 공정관리는 품질향상을 위한 중요한 요소가 되고 있다. 측정된 데이터에 근거하여 공정을 관리하므로 정확한 데이터의 수집과 분석을 위한 측정관리는 통계적 공정관리의 선행조건이다([1]). 만일 측정과정(measurement process)이 충분히 관리되지 않아서 측정된 데이터가 제조공정(manufacturing process)의 변동을 정확하지 보여주지 못하면 측정결과로부터 공정능력등의 품질평가를 제

로 할 수 없다. 그러므로 품질 보증의 기본인 측정관리를 위하여 여러개의 계측기, 검사원, 검사과정 등을 포함한 측정시스템이 제품이나 공정의 성능특성을 정확히 측정하여 올바른 데이터를 산출할 수 있는지 반드시 평가되어야 한다. 만일 정확히 평가되지 않고 사용된다면, 잘못된 측정시스템으로 인하여 공정을 정확히 판단할 수 없으며 이로 인하여 품질비용은 상승하게 되고 지속적으로 품질을 향상시킨다는 것은 거의 불가능하다.

계측기의 반복성 및 재현성(gage repeatability and reproducibility, gage R&R) 연구는 두가지 경우를 생각할 수 있다. 첫째는 계측기를 범용으로 다수의 제품(또는 부

* 경상대학교 산업공학과

** 경상대학교 산업공학과, 생산기술연구소

뿔)을 측정하는 경우이고, 둘째는 계측기를 주로 어떤 특정 제품을 측정하는데 이용할 경우이다. 본 연구는 후자의 경우 gage R&R 연구를 대상으로 한다. 기존의 경험을 바탕으로 관리하여야 할 측정치수(measurement dimension)를 알고 있을 때, 즉, 어떤 부품의 여러 가지 치수항목을 갖고 있는데 치수품질이 특정한 주요측정치수(critical measurement dimension)에 의존할 때는 이 치수를 측정하여 치수의 측정변동이 공정의 전체변동에서 차지하는 비율을 파악하면 된다. 하지만 신제품개발의 경우나 기존생산제품이라 하더라도 주요측정치수가 알려져 있지 않은 경우에는 임의로 하나의 측정치수를 선정하여 이 치수의 측정변동이 공정의 전체변동에서 어느 정도를 차지하는가를 알아보기 위하여 사용되었다. 하지만 어떤 계측기로 다수의 치수를 측정해야 하는 경우에는 어느 한 측정치수의 gage R&R만을 가지고 그 계측기의 전반적 측정능력으로 판단하기는 어렵다. 예를 들어 버어니어 캘리퍼스를 가지고 그림 1과 같은 부품을 측정할 때, 어느 한 치수만을 대상으로 버어니어 캘리퍼스의 gage R&R을 파악하는 것은 버어니어 캘리퍼스의 전반적 계측능력을 판단하기에 미흡하다.

단일치수를 대상으로 계측기의 측정능력을 파악할 때

생기는 이러한 단점을 극복하기 위하여 모든 치수를 다 재어 개개의 gage R&R을 파악하는 것은 과도한 시간과 비용을 필요로 한다. 이러한 이유로, 다변량 통계분석방법중 하나인 주성분분석(principal component analysis)를 이용하여 변동의 형태를 대표하는 소수의 주성분을 파악한 다음, 그 주성분 자체를 관리하거나 또는 주성분을 크게 설명하는 개별 측정치수를 선정하여 집중관리하게 되면 측정관리의 효율을 제고할 수 있다. 본 연구에서는 먼저 계측기의 반복성 및 재현성 평가방법들을 정리하고, 신규로 개발된 제품 또는 한 부품 내에 다수의 측정 치수가 있어 변동이 심하거나 변동의 형태를 대표하는 소수의 주성분을 파악하여 집중관리할 수 있도록 주성분 분석방법을 사용하여 측정변동을 분석하는 일반적인 체계를 제시하고자 한다. 주성분 분석을 하면 원자료(raw data)에 내재하는 전체변동을 중요한 몇 개의 주성분으로 표현할 수 있게 된다. 따라서 본 연구에서는 주성분 분석방법을 이용한 측정변동분석의 일반적인 해석체계를 제안하고, 직접적인 측정실험을 실시하여 얻은 데이터를 가지고 측정변동을 분석한 사례를 제시하고자 한다.

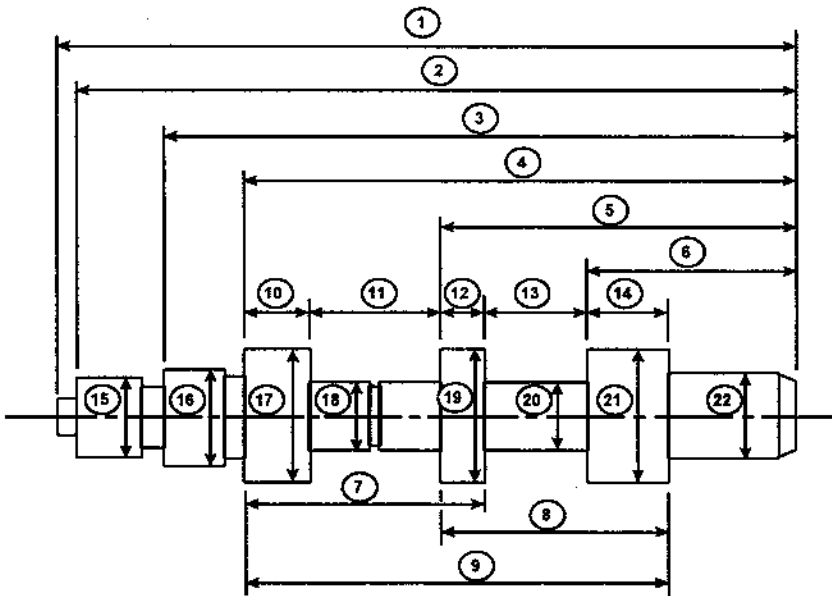


그림 1. 조절밸브의 측정치수 (버어니어 캘리퍼스 이용)

2. 측정시스템의 평가

2.1 계측기의 반복성 및 재현성 정의

계측기의 반복성(repeatability)은 한 명의 측정자가 동일한 계측기를 사용하여 동일한 부품을 여러번 반복하여 측정할 경우 나타나는 측정값의 변동으로 정의되며, 측정자의 재현성(reproducibility)은 여러명의 측정자가 동일한 계측기를 사용하여 동일한 부품을 측정할 경우 서로 다른 측정자들이 측정한 측정 평균값의 변동을 의미한다. 측정시스템의 평가는 정확성, 정밀도, 안정성 등의 기준도 포함하나, 본 연구에서는 계측기의 반복성과 재현성을 동시에 평가하는 것으로 한정한다. 즉 계측기에 의한 변동과 측정자에 의한 변동을 동시에 고려하여 측정시스템을 평가하는 것이다. 공정의 전체변동은 크게 부품간 변동(part-to-part variation)과 측정 과정 중에 생기는 변동(measurement variation)으로 볼 수 있는데, 이를 수식으로 표현하면

$$\sigma^2_{total} = \sigma^2_{part} + \sigma^2_{measurement} \quad (1)$$

이 되며, 여기서 측정변동은

$$\sigma^2_{measurement} = \sigma^2_{repeatability} + \sigma^2_{reproducibility} \quad (2)$$

로 표현된다.

2.2 계측기의 반복성 및 재현성 분석방법

계측기의 반복성 및 재현성 분석방법에는 크게 범위법(range method), 평균과 범위법(average and range method), 실험계획법(design of experiments)이 있는데, 범위법은 측정의 변동을 개략적으로 구하는 비교적 간단한 방법이다. 그러나 이 방법은 반복성과 재현성을 따로 구분하여 평가할 수 없는 한계점이 있다. 평균과 범위법은 측정시스템에 대한 반복성과 재현성 모두를 평가하는 방법으로 선진국의 산업계에서 널리 사용되고 있다([5], [6]). 본 논문에서는 통계소프트웨어의 보급에 부합하여 최근 자주 사용되고 있는 분산요소 추정방법(variance component estimation method)에 의거한 실험계획법을 이용한다. 이 방법은 다른 방법들에 비해 보다 정교한 측정시스템 평가방법이다. 실험계획법의 장점은 (1) 변동요소를 통계적인 의미에서 보다 효율적으로 추정할 수 있고, (2) 측정

자와 부품의 교호작용에 관한 정보등 평균과 범위법에서 얻을 수 없는 정보를 얻을 수 있으며, (3) 측정 데이터 분석시 제한사항이 적다.

분산요소는 SAS[9]의 PROC MIXED procedure를 이용하여 추정할 수 있는데, 계측시스템의 반복성 및 재현성 평가를 위하여 분산요소 추정방법을 사용할 경우, 데이터의 구조식은

$$X_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijk} \quad (3)$$

가 된다. 여기서

- X_{ijk} : i 번째 부품을 j 번째 측정자가 k 번째 반복 측정값
- α_i : i 번째 부품의 효과
- β_j : j 번째 측정자의 효과
- γ_{ij} : i 번째 부품을 j 번째 측정자가 잴 때, 부품과 측정자간의 교호작용효과
- ϵ_{ijk} : 동일한 측정자가 동일한 부품을 반복측정할 경우의 변동

이다. 따라서 측정시스템의 총변동은 다음과 같이 분해될 수 있다.

$$\sigma^2_{total} = \sigma^2_{PV} + \sigma^2_{AV} + \sigma^2_{IV} + \sigma^2_{EV} \quad (4)$$

여기서 σ^2_{PV} 는 부품간 변동(part-to-part variation), σ^2_{AV} 는 측정자에 의한 변동(appraiser variation), σ^2_{IV} 는 측정자와 부품의 교호작용에 의한 변동(appraiser-by-part interaction variation), σ^2_{EV} 는 계측기에 의한 변동(equipment variation)을 나타낸다. 통상적으로 gage R&R 연구에서는 각 변동을 이용한 신뢰구간의 크기로써 반복성, 재현성 및 측정시스템의 변동을 나타낸다. 즉, $Z_{\alpha/2}$ 를 표준정규분포의 $100(1-\alpha/2)\%$ 에 해당하는 값이라고 할 때, $\nu = 2 \cdot Z_{\alpha/2}$ 라 하면 (보통 $\alpha=0.01$ 을 사용할 때, $2 \cdot Z_{0.005} = 5.15$ 또는 0.05 에 해당하는 $2 \cdot Z_{0.025} = 3.92$ 를 이용), 반복성, 재현성 및 측정시스템의 변동은 99% 또는 95% 신뢰구간의 크기에 해당하는 ' $\nu \cdot \sigma$ (측정데이터로부터 구한 표준편차)'로 구한 다음의 식 (5), (6) 및 (7)로 정의된다 (King[4], SAS[8]).

$$\text{반복성} = \nu \sigma_{EV} \quad (5)$$

$$\text{재현성} = \nu \hat{\sigma}_{AV} \tag{6}$$

$$\text{측정시스템 변동} = \nu \sqrt{\hat{\sigma}_{EV}^2 + \hat{\sigma}_{AV}^2 + \hat{\sigma}_{IV}^2} \tag{7}$$

3. 다변량 데이터 해석을 위한 주성분 분석방법

3.1 주성분 분석의 개념과 목적

주성분 분석은 여러개의 양적변수(quantitative variable)들 사이의 관계를 분석하여 이 변수들의 선형결합으로 표시되는 새로운 주성분(principal components)을 찾고, 이 중에서 중요한 몇 개의 주성분으로 전체변동을 설명하고자 하는 다변량 통계분석법이다. 자료의 요약이나 선형관계식을 통하여 차원(dimension)을 감소시켜 해석을 용이하게 하는데 목적이 있다. 자료가 갖고 있는 전체적인 변동의 대부분을 원래변수의 수(p개)보다 적은 수(m개)의 주성분으로 설명할 수 있다고 하면, p개의 변수가 갖고 있는 정보의 대부분을 m개의 주성분으로 대체(p>m)할 수 있다. 이렇게 함으로써 변수의 차원을 감소시킬 수 있으며, 상관관계가 있는 변수들의 경향이나 변수들이 상관관계를 이루며 나타내는 변동을 몇개의 주성분으로 파악할 수 있게 한다.

3.2 주성분 분석의 모형 및 특성

서로 상관되어 있는 p(≥2)개의 확률변수 X_1, X_2, \dots, X_p 를 원소로 하는 확률벡터 X 가 모평균벡터 μ 와 모공분산행렬 Σ 를 가진다고 하고, 이들을 다음의 기호를 나타내자.

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_i \\ \vdots \\ X_p \end{bmatrix} \quad \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_i \\ \vdots \\ \mu_p \end{bmatrix} \quad \Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1j} & \dots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2j} & \dots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \sigma_{i1} & \sigma_{i2} & \dots & \sigma_{ij} & \dots & \sigma_{ip} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \dots & \sigma_{pj} & \dots & \sigma_{pp} \end{bmatrix}$$

단, $\sigma_{ij} = E[(X_i - \mu_i)(X_j - \mu_j)] = \sigma_{ji}$ 는 X_i 와 X_j 의 모공분산이고, $\mu_i = E[X_i]$ 는 X_i 의 모평균이다. 여기서 Σ 가 고유값(eigenvalues) $\lambda_1, \lambda_2, \dots, \lambda_p$ ($\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$)와 이에 대응하는 고유벡터(eigenvectors) l_1, l_2, \dots, l_p 를 갖는다고 가정하면, 주성분 Y_1, Y_2, \dots, Y_p 는 다음과 같은 모형을 이루고 있다고 가정할 수 있다.

$$\begin{aligned} Y_1 &= [l_1]'X = l_{11}X_1 + l_{12}X_2 + \dots + l_{1p}X_p \\ Y_2 &= [l_2]'X = l_{21}X_1 + l_{22}X_2 + \dots + l_{2p}X_p \\ &\vdots \\ Y_p &= [l_p]'X = l_{p1}X_1 + l_{p2}X_2 + \dots + l_{pp}X_p \\ [l_i]' &= (l_{i1}, l_{i2}, \dots, l_{ip}), i = 1, 2, \dots, p \\ [l_i]'l_j &= \sum_{j=1}^p l_{ij}^2 = 1 \end{aligned} \tag{8}$$

주성분들의 분산과 공분산은 $\text{Var}(Y) = l' \Sigma l$, $\text{Cov}(Y_i, Y_j) = l_i' \Sigma l_j = 0$ ($i \neq j$)가 된다.

첫번째 주성분 Y_1 의 고유벡터 $l_1' = (l_{11}, l_{12}, \dots, l_{1p})$ 는 Y_1 이 최대분산을 갖게 한다. 두번째 주성분 Y_2 의 고유벡터 $l_2' = (l_{21}, l_{22}, \dots, l_{2p})$ 는 l_1' 과는 직교하며, 최대분산을 갖게 하며, 이러한 과정을 통해 나머지 주성분들의 고유벡터가 차례로 구해진다. 이렇게 구해진 주성분들은 다음과 같은 성질을 가진다.

- ① 주성분들은 서로 상관관계가 없다.
- ② $\text{Var}(Y_i) = \lambda_i, i = 1, 2, \dots, p$
- ③ $\sum_{i=1}^p \text{Var}(Y_i) = \sum_{i=1}^p \lambda_i = \sum_{i=1}^p \text{Var}(X_i)$

Y_1, Y_2, \dots, Y_m ($m < p$)의 분산의 합 $\lambda_1 + \lambda_2 + \dots + \lambda_m$ 이 전체 분산의 대부분을 설명한다고 판단되면, p개의 변수 X_1, X_2, \dots, X_p 대신에 m개의 새로운 변수 Y_1, Y_2, \dots, Y_m 으로 전체 변동을 설명할 수 있다. $\text{Var}(Y) = \lambda_i$ 이므로 i번째 주성분에 의해서 설명되는 전체변동(분산)의 비율은 다음과 같다.

$$\frac{\lambda_i}{\lambda_1 + \lambda_2 + \dots + \lambda_p}, i = 1, 2, \dots, p \tag{9}$$

그리고 처음m개의 주성분에 의해서 설명되는 분산의 비율은 다음과 같다.

$$\frac{\lambda_1 + \lambda_2 + \dots + \lambda_m}{\lambda_1 + \lambda_2 + \dots + \lambda_p}, \quad (m < p) \quad (10)$$

식 (10)의 값이 80% 이상 된다면 전체변동의 대부분을 m개의 주성분으로 설명할 수 있다고 생각한다. 그러므로, 크기순으로 나열된 처음 m개의 주성분으로 원래자료(p개의 변수)에 내재하는 전체적 변동의 형태를 대부분 설명함으로써 변동에 관한 정보의 손실을 작게 하면서 차원의 축약을 기할 수 있다.

4. 주성분 분석방법을 이용한 측정변동 분석

신규로 개발된 제품 또는 한 부품내에 다수의 측정치수가 있어서 큰 변동을 나타내는 몇 개의 치수나 변동의 형태를 대표하는 소수의 주성분을 파악하여 집중관리함으로써 시간과 비용을 절약하고자 할 때 주성분분석을 이용하면 효과적이다. 만일 개개의 측정치수의 변동을 일일이 관리하고자 한다면, 다수의 측정치수간의 상관관계에 의거한 변동을 대표하는 형태를 파악할 수 없어 측정변동의 관리가 비효율적이다. 주성분 분석을 이용하면 다수의 측정치수간의 상관관계를 파악하여 측정변동의 형태를 대표하는 소수의 주성분을 찾아낼 수 있다. 이렇게 구한 소수의 주성분을 이용하면 측정변동을 경제적으로 관리할 수 있다. 따라서 본 논문에서는 주성분분석을 통하여 측정변동의 형태를 대표하는 소수의 주성분을 집중관리할 수 있는 방법을 다음과 같이 제시한다.

(1) 측정변동의 추정치 계산

측정변동의 분석방법중 하나인 실험계획법을 사용하여 각 측정변동의 추정치를 계산한다. 이 때 분산요소 추정을 위해 SAS의 PROC MIXED procedure를 사용할 수 있다.

(2) 주성분 분석

측정변동의 추정치를 가지고 주성분 분석을 수행한다. 측정변동 추정치들의 단위가 동일하면 공분산행렬을 이용하고 동일하지 않으면 상관행렬을 이용한다. 주성분 분석을 통해 자료가 갖고 있는 전체적인 변동의 대부분을 원래변수의 수(p개)보다 적은 수(m개)의 주성분으로 설명할 수 있으면, 원자료에서 파악하기 힘들었던 상관관계가 있는 변수들의 변동경향을 파악하여 그들을 집중관리할

수 있게 된다. 본 논문에서는 주성분에 의해서 설명되는 분산의 양이 전체변동의 80% 이상 되는 주성분들을 선정한다.

(3) 분석결과의 해석

주성분 분석으로 구한 고유값(eigenvalues)의 누적비율이 80% 이상 될 때까지의 주성분을 구하면 각 주성분은 변수(측정치수)들의 선형관계식으로 표현된다. 일반적인 측정변동 관리에서는 각 주성분의 선형관계식에서 변수의 계수(고유벡터 상수)를 제공하여 누적한 값이 0.8 이상 될 때까지 변수를 선정하여 관리하고, 측정변동을 보다 엄격하게 관리해야할 경우에는 변수의 계수를 제공하여 누적한 값이 더 큰 값(예를 들어, 0.9) 이상이 될 때까지 변수를 선정하여 관리할 수도 있다. 이때 주성분에서 한 변수의 계수제공이 0.8 이상일 경우는 하나의 변수만 선정하여 관리하면 되지만, 주성분에서 여러변수의 계수제공이 누적되어 0.8 이상될 경우는 누적된 여러변수로 이루어지는 주성분 자체를 관리해야 한다. 5절의 사례연구에서도 나타나듯이 각 주성분이 어떤 변수들로 구성되는가에 따라서 측정변동의 경향을 분석결과 및 문제상황에 의거하여 파악할 수 있다. 이는 일반적으로 주성분분석시 결과의 해석이 분석결과를 토대로 문제상황, 경험 및 직관에 상당히 의존한다는 사실과 부합한다([2], [3]).

5. 측정변동분석 사례연구

두 가지 사례연구를 통하여 한 부품내에 다수의 측정치수가 있는 부품을 버니어 캘리퍼스(vernier calipers)와 하이트게이지(height gage)를 사용하여 측정하는 과정, 측정변동 계산과 주성분 분석을 통하여 측정변동이 심한 소수의 주성분을 파악하는 과정, 그리고 주성분에 의미를 부여하는 과정을 소개한다. 또한 측정시스템의 평가가 계측기의 정밀도에 따라 어떠한 차이가 있는지 살펴 보고자 한다.

5.1 사례연구 1

한 부품내 다수의 측정 치수가 있는 실험대상으로 자동차용 자동변속장치(automatic transmission)에 조립되는 부품중 하나인 조절밸브(regulator valve)를 선택하였다(그

림 1). 이 밸브는 오일펌프(oil pump)에서 발생한 유압을 일정하게, 즉, 라인압으로 조절하며, 그리고 라인압을 모든 운전조건에 적용하도록 조정하는 기능을 가지고 있다. 랜덤하게 추출한 조절밸브 30개를 대상으로 버어니어 캘리퍼스(눈금:0.01mm)를 사용하여 3명의 측정자가 5회 반복측정한 데이터를 가지고, 한 부품내에 있는 22개의 측정치수중에서 측정변동이 심하거나 변동을 대표하는 주성분을 파악하여 집중적으로 관리해야 할 소수의 주성분 또는 측정치수를 선정하고자 한다.

(1) 측정치수

그림1에 조절밸브의 22개 측정치수를 표현하였다. 측

정치수 1에서 14까지는 길이를 나타내며, 15에서 22까지는 지름을 나타낸다.

(2) 측정변동 계산

본 사례에서는 실험계획법(분산요소 추정방법)을 이용하여 측정변동을 분석하였으며, 분산요소 추정을 위해 SAS의 PROC MIXED procedure를 사용하였다. 실험계획법을 이용한 구체적인 측정변동의 계산절차는 다음과 같다.

- ① 부품(조절밸브) 30개를 랜덤으로 샘플링하고, 식별 번호를 1~30까지 부여한다.

표 1. 조절밸브의 측정데이터 샘플 (단위 : mm) (버어니어 캘리퍼스 이용)

측정치수	반복 1			반복 2			반복 3			반복 4			반복 5		
	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C
1	60.48	60.50	60.49	60.51	60.52	60.49	60.51	60.51	60.49	60.50	60.51	60.49	60.51	60.51	60.51
2	59.00	59.03	58.99	59.01	59.06	59.00	59.02	59.02	58.99	59.06	59.07	59.00	59.03	59.04	59.04
3	51.48	51.45	51.53	51.53	51.47	51.45	51.48	51.49	51.48	51.47	51.52	51.50	51.48	51.50	51.52
4	44.01	44.03	43.98	44.07	44.02	43.99	44.09	44.07	44.01	44.01	44.02	44.00	44.01	44.04	44.03
5	30.50	30.55	30.50	30.52	30.49	30.57	30.50	30.49	30.54	30.49	30.52	30.48	30.50	30.53	30.49
6	19.50	19.47	19.45	19.49	19.46	19.47	19.48	19.49	19.47	19.50	19.48	19.45	19.49	19.53	19.49
7	16.04	16.01	16.01	16.03	16.01	16.02	16.03	16.01	16.01	16.02	16.01	16.00	16.03	16.01	16.01
8	18.52	18.51	18.51	18.54	18.53	18.52	18.53	18.51	18.51	18.53	18.51	18.50	18.52	18.53	18.52
9	32.03	31.98	32.00	32.02	32.01	32.01	32.03	32.03	32.01	32.03	32.01	32.00	32.02	32.04	32.02
10	6.00	5.99	5.99	6.01	6.00	5.99	6.00	5.98	5.98	6.01	6.00	5.98	6.00	6.00	5.99
11	7.44	7.45	7.46	7.43	7.43	7.45	7.47	7.44	7.47	7.44	7.47	7.46	7.45	7.45	7.46
12	2.51	2.49	2.53	2.52	2.52	2.50	2.52	2.49	2.49	2.52	2.51	2.50	2.51	2.51	2.50
13	8.45	8.45	8.45	8.45	8.45	8.45	8.45	8.45	8.48	8.44	8.46	8.46	8.44	8.48	8.46
14	7.47	7.50	7.50	7.50	7.49	7.51	7.52	7.48	7.49	7.51	7.49	7.49	7.50	7.49	7.48
15	8.18	8.20	8.19	8.20	8.20	8.18	8.20	8.19	8.19	8.19	8.20	8.19	8.20	8.20	8.20
16	9.89	9.89	9.87	9.89	9.90	9.87	9.88	9.87	9.87	9.89	9.89	9.87	9.89	9.89	9.88
17	15.94	15.96	15.94	15.96	15.96	15.95	15.96	15.95	15.94	15.95	15.96	15.94	15.96	15.96	15.94
18	8.00	8.02	8.00	8.01	8.01	8.00	8.00	8.01	7.99	8.00	8.01	7.99	8.00	8.01	7.99
19	15.95	15.96	15.95	15.96	15.96	15.95	15.97	15.95	15.94	15.96	15.95	15.94	15.96	15.96	15.94
20	7.98	7.99	7.99	8.01	8.01	8.06	8.00	7.98	7.98	8.00	8.00	7.97	8.00	8.00	7.98
21	15.96	15.96	15.95	15.97	15.96	15.96	15.96	15.95	15.94	15.95	15.96	15.94	15.96	15.96	15.95
22	11.76	11.77	11.76	11.78	11.79	11.77	11.76	11.78	11.76	11.79	11.77	11.77	11.79	11.78	11.78

- ② 측정작업을 수행할 측정자 3명을 선택한다.
- ③ 계측기는 버어니어 캘리퍼스(눈금 : 0.01mm)를 사용한다.
- ④ 측정자의 측정순서는 랜덤하게 결정한다.
- ⑤ 랜덤하게 결정된 첫번째 측정자는 30개의 부품을 랜덤한 순서로 부품내 22개의 치수를 한번씩 측정한다.
- ⑥ 나머지 측정자도 30개의 부품을 랜덤한 순서로 22개의 치수를 한번씩 측정한다.
- ⑦ 3명의 측정자가 30개의 부품을 5회 반복 측정한다. 이때 ④, ⑤, ⑥의 과정을 반복한다.
- ⑧ 분산요소 추정방법을 이용하여 반복성, 재현성 및 측정시스템의 변동을 추정한다.

표 1은 하나의 벨브에 대하여 버어니어 캘리퍼스를 사용하여 3명의 측정자(A, B, C)가 5회 반복해서 측정한 데이터이다. 표 1과 같은 30개의 측정데이터세트를 이용, 분산요소 추정을 위해 SAS의 PROC MIXED 절차를 사용하여 22개의 측정 치수별 재현성(측정자 변동), 반복성(계측기 변동), 측정시스템 변동(반복성 및 재현성의 합)의 추정치를 구한 결과를 표 2에 나타내었다. 식 (5)-(7)에 언급한 ν 는 $\alpha=0.01$ 일 때의 값 5.15를 적용한다.

(3) 주성분 분석

한 부품내 여러 측정 치수가 있는 경우, 각 측정 치수 간의 상관관계에 의거하여 측정변동의 형태를 대표하는 소수의 주성분을 파악하기 위하여 주성분 분석을 실시한다. 측정변동 추정치들의 단위가 동일하므로 공분산행렬에 대하여 SAS[7]의 PROC PRINCOMP 절차를 수행하여 주성분 분석을 한다. 본 사례에서는 측정시스템의 변동에 대한 분석을 한다.

주성분분석의 결과 고유값을 표 3에 나타내었는데, 9개의 주성분을 이용하여 측정시스템 변동의 84% 정도를 설명할 수 있다. 또한 각 주성분의 고유벡터를 나타낸 표 4로부터 9개 주성분을 원래 변수(측정치수)의 선형식으로 표현할 수 있다. 예를 들어,

$$Y_1 = 0.187231X_1 + 0.311346X_2 + \dots + 0.101537X_{22} \quad (11)$$

표 2. 측정변동의 추정치(버어니어 캘리퍼스 이용)

측정치수	반복성 (계측기 변동)	재현성 (측정자 변동)	측정시스템 변동 (계측기 및 측정자 변동의 합)
1	0.05234	0.03029	0.06047
2	0.11554	0.09684	0.15076
3	0.13195	0.00000	0.13195
4	0.14697	0.08499	0.16977
5	0.13228	0.00000	0.13228
6	0.09622	0.05156	0.10917
7	0.02928	0.05804	0.06501
8	0.04925	0.04033	0.06366
9	0.07012	0.03868	0.08008
10	0.03686	0.04125	0.05532
11	0.06246	0.03304	0.07066
12	0.06497	0.02345	0.06908
13	0.05860	0.02357	0.06316
14	0.06140	0.00000	0.06140
15	0.03806	0.01424	0.04064
16	0.03408	0.03970	0.05232
17	0.02988	0.03366	0.04501
18	0.02630	0.03093	0.04060
19	0.03205	0.03703	0.04898
20	0.11012	0.00000	0.11012
21	0.03429	0.02990	0.04550
22	0.05312	0.00440	0.05330

각 주성분의 고유벡터로부터 계수제곱이 누적되어 0.8 이상되는 변수의 계수를 골라 표 4에 굵게 표시한다. 예를 들어, 제1주성분의 경우는

$$Y_1 \approx 0.311346X_2 + 0.495180X_{12} + 0.282788X_{13} + 0.417711X_{14} + 0.322358X_{16} - 0.239024X_{20} \quad (12)$$

결국 그림1의 22개 측정 치수 각각에 대하여 측정시스템 변동을 관리하는 대신에, 주성분분석을 통하여 변수의 계수제곱이 누적되어 0.8 이상되는 9개 주성분의 선형관계를 토대로 측정시스템 변동을 관리한다. 여기서 특기할 사항은 측정시스템 변동을 설명하는 주성분을 구성하는

표 3. 고유값 (버어니어 캘리퍼스 이용)

주성분	고유값	%	누계%
1	0.001282	15.72	15.72
2	0.001183	14.50	30.22
3	0.000978	12.00	42.22
4	0.000816	10.01	52.23
5	0.000667	8.17	60.40
6	0.000576	7.07	67.47
7	0.000531	6.51	73.98
8	0.000471	5.78	79.76
9	0.000343	4.20	83.96
:	:	:	:
22	0.000004	0.05	100

치수의 번호가 1부터 14까지, 즉, 대부분 밸브의 길이치수를 나타내고 있다는 점이다

5.2 사례연구 2

버어니어 캘리퍼스를 사용하여 측정실험을 실시한 사례1의 경우는 전반적으로 길이치수의 측정변동이 크게 나타난다. 따라서 보다 정밀한 계측기인 하이트게이지 (height gage, 눈금:0.001mm)를 이용하여 길이치수를 나타내는 14개의 측정치수중에서 측정변동이 심하거나 변동을 대표하는 주성분을 파악하기 위한 실험을 하였다. 사례1에서 사용했던 조절밸브중 10개를 랜덤 샘플링하고, 사례1과 동일한 3명의 측정자가 3회 반복측정하여, 10개 부품에 대한 측정변동의 추정치를 구한후 사례1과 동일한 방법으로 주성분 분석을 실시한다.

(1) 측정치수

사례연구1과 동일한 조절밸브의 길이치수를 하이트게이지로 측정한다. 그림2에 하이트게이지를 이용한 측정치

표 4. 고유벡터 (버어니어 캘리퍼스 이용)

측정치수	제1 주성분	제2 주성분	제3 주성분	제4 주성분	제5 주성분	제6 주성분	제7 주성분	제8 주성분	제9 주성분
1	0.187231	-1.78836	-1.46601	-1.28846	-2.55447	-1.70007	0.429630	0.414780	-1.72883
2	0.311346	-5.67839	-3.29177	-1.76082	0.183763	0.322824	0.095737	-2.33882	0.227062
3	0.103819	0.356470	-5.75330	0.403709	-1.62726	0.193834	0.327016	-0.96535	-2.00361
4	0.173447	0.268647	-3.05835	0.037257	-1.03138	0.236569	-4.38259	0.146793	0.395685
5	0.175525	0.148823	0.414688	-1.52511	-2.69804	0.326338	0.269386	0.171842	0.393235
6	0.180316	-2.03057	0.271084	0.677826	0.277187	0.069758	-0.69841	0.299800	-0.17564
7	-0.054752	0.200343	0.192205	0.125831	-0.64229	0.137758	0.099541	-2.04139	-2.51191
8	-1.45956	0.376555	-0.64199	-2.89223	0.211023	-0.15626	-0.06094	0.338152	0.027551
9	0.057371	0.020145	0.106081	-1.02074	0.328358	0.365080	0.008348	-0.037864	-1.94849
10	-1.99432	0.140864	0.219165	0.037828	-0.47222	0.182100	0.234184	-5.44144	0.140575
11	0.099232	0.203289	-1.02611	-1.06232	0.612006	-1.08965	0.018535	-1.01941	0.059854
12	0.495180	0.057175	0.216889	0.031384	-0.69418	0.198999	-1.63059	-0.40141	-3.00855
13	0.282788	0.160499	0.105089	-0.048665	0.279610	-3.54252	0.287511	-0.27407	-0.27528
14	0.417711	0.076749	0.017326	0.076783	-1.63520	-5.16157	-0.66225	-3.49403	0.257501
15	0.038807	0.173966	-0.77214	0.165790	0.098238	0.028560	-0.55619	-0.28262	0.025724
16	0.322358	0.218848	0.112952	-0.051753	-0.71308	0.054317	-0.73336	0.085729	0.054798
17	-0.00245	0.066667	0.028513	-0.057467	0.000251	-0.11685	0.023255	0.004215	0.036103
18	0.111151	0.111259	-0.20676	-1.20523	0.176151	0.152558	0.404760	0.073645	0.044088
19	0.041037	0.071509	-0.53373	-0.02611	0.033491	0.038562	0.058732	0.011882	0.178070
20	-2.39024	-0.045437	0.021005	0.351984	0.112231	-0.60259	0.258959	0.128332	0.490543
21	0.007348	0.048588	0.086740	0.031061	0.058059	-0.01832	-0.27768	-0.13771	0.007876
22	0.101537	0.002967	0.010959	-0.24357	0.091571	-0.25166	0.057105	0.064870	0.069579

수를 나타내었다.

반복성, 재현성 및 측정시스템 변동의 추정치를 구한 결

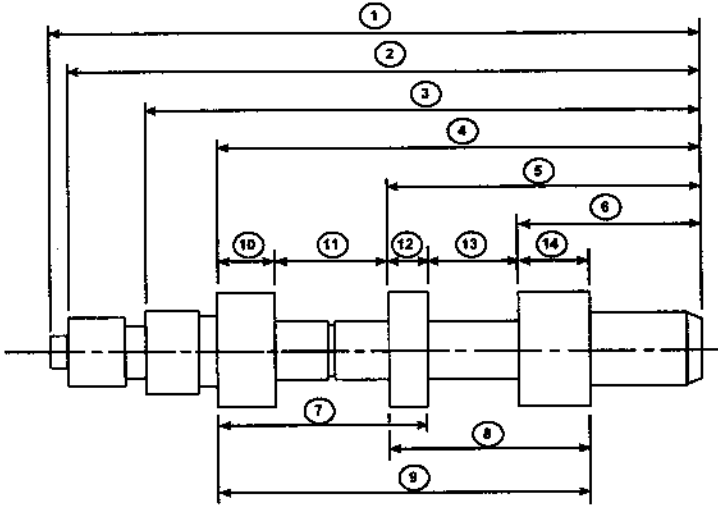


그림 2. 조절밸브의 측정치수 (하이트게이지 이용)

(2) 측정변동 계산

과가 표 6과 같다.

사례1과 동일한 방법으로 계산하였다. 표 5는 한개의 밸브에 대하여 하이트게이지를 사용하여 사례1과 동일한 3명의 측정자(A, B, C)가 3회 반복해서 측정한 데이터이다. 표 5의 측정 데이터를 이용하여 14개의 측정 치수별

(3) 주성분 분석

측정시스템 변동에 대한 주성분 분석을 실시한 결과, 표 7로부터 측정시스템 변동의 81% 정도를 설명할 수 있

표 5. 조절밸브의 측정데이터 샘플 (단위: mm) (하이트게이지 이용)

측정 치수	반복 1			반복 2			반복 3		
	A	B	C	A	B	C	A	B	C
1	60.502	60.467	60.500	60.507	60.487	60.514	60.480	60.517	60.514
2	58.999	58.969	58.993	58.990	58.989	58.999	58.976	59.008	59.003
3	51.474	51.442	51.473	51.475	51.461	51.473	51.457	51.481	51.480
4	44.001	43.975	44.006	44.002	44.000	44.012	43.978	44.014	44.015
5	30.498	30.472	30.496	30.501	30.486	30.496	30.475	30.471	30.498
6	19.480	19.464	19.481	19.490	19.476	19.474	19.453	19.459	19.473
7	16.025	16.038	16.041	16.024	16.016	16.046	16.043	16.049	16.047
8	18.523	18.537	18.534	18.522	18.523	18.535	18.543	18.518	18.537
9	32.033	32.038	32.046	32.022	32.009	32.054	32.050	32.053	32.050
10	6.008	6.020	6.022	6.007	5.997	6.022	6.022	6.028	6.019
11	7.490	7.483	7.487	7.491	7.510	7.499	7.491	7.503	7.495
12	2.525	2.536	2.537	2.527	2.527	2.525	2.532	2.527	2.534
13	8.491	8.482	8.476	8.487	8.496	8.487	8.491	8.504	8.490
14	7.508	7.521	7.521	7.509	7.511	7.521	7.521	7.514	7.515

표 6. 측정변동의 추정치 (하이트게이지 이용)

측정 치수	반복성 (계측기 변동)	재현성 (측정자 변동)	측정시스템 변동 (계측기 및 측정자 변동의 합)
1	0.08875	0.00000	0.08875
2	0.06498	0.00000	0.06498
3	0.06545	0.00000	0.06545
4	0.07257	0.02356	0.07630
5	0.04922	0.04626	0.06755
6	0.05994	0.00000	0.05994
7	0.05999	0.01409	0.06162
8	0.04510	0.00000	0.04510
9	0.07952	0.01126	0.08032
10	0.05095	0.00000	0.05095
11	0.04332	0.00000	0.04332
12	0.02456	0.00000	0.02456
13	0.04030	0.00898	0.04128
14	0.02831	0.00103	0.02833

는 2개의 주성분을 구할 수 있다. 또한 표 8로 부터 2개 주성분을 원래 측정치수의 선형식으로 나타낼 수 있다. 각 주성분에 대하여 계수제곱의 누적값이 0.8 이상되는 치수의 계수를 골라 표 8에 진하게 표시하였다. 결국 2개의 주성분은 다음과 같은 근사식으로 표현된다.

표 7. 고유값 (하이트게이지 이용)

주성분	고유값	%	누계%
1	0.005902	62.06	62.06
2	0.001835	19.29	81.35
:	:	:	:
9	0.000025	0.26	100
:	:	:	:
14	0.000000	0.00	100

$$Y_1 \approx 0.352540X_1 + 0.389376X_4 + 0.416636X_5 + 0.346678X_6 - 0.368734X_8 - 0.360776X_{14} \quad (13)$$

$$Y_2 \approx 0.336631X_2 + 0.317958X_3 + 0.537210X_7 +$$

표 8. 고유벡터 (하이트게이지 이용)

측정 치수	제 1 주성분	제 2 주성분
1	0.158091	0.200176
2	0.352540	0.336631
3	0.336880	0.317958
4	0.389376	0.198726
5	0.416636	0.144893
6	0.346678	0.057883
7	0.032206	0.112097
8	-0.368734	0.537210
9	-0.137970	0.339488
10	-0.045757	0.092647
11	0.048574	0.026950
12	-0.023648	0.181382
13	-0.025291	0.134763
14	-0.360776	0.452650

$$0.339488X_9 + 0.452650X_{14} \quad (14)$$

두 사례연구에서 조절밸브의 길이에 대한 측정변동은 버어니어 캘리퍼스보다 하이트게이지를 이용할 때가 더 작은 것으로 나타났다. 이는 길이치수를 쉽게 잴 수 있고 정밀도가 높은 계측기를 사용하면 측정변동이 작아지는 실례를 보여주는 것이다. 표 2와 표 6의 측정변동 추정치로 부터 계측기 변동이 측정자 변동보다 크게 나타남을 볼 수 있다. 이는 측정변동의 원인 규명에 중요한 정보가 된다. 그리고 계측기로 인한 변동이 심한 측정치수가 나타나면 측정시 사용된 계측기의 교정상태와 정도를 다시 확인하고, 정도가 좋지 않으면 계측기의 정도를 높이기 위하여 수리하거나 새롭게 제작 혹은 구입하여야 하며, 측정시 계측기의 교정방법 및 계측위치에 문제가 생기지 않도록 개선조치 하여야 한다. 또한 측정자로 인한 변동이 심한 측정치수가 나타나면 측정자가 계측기의 눈금을 읽는 방법이나 사용방법에 대한 지식이 부족하므로 교육을 실시하여야 한다. 주성분의 선형관계식에서 한 번

수의 계수제곱이 0.8이상으로 큰 값을 가지는 경우는 그 변수에서 큰 측정변동을 나타내므로 그 변수자체를 중점적으로 관리하면 되지만, 주성분에서 한 변수의 계수제곱이 0.8 이상되는 경우가 없는 경우는 계수제곱이 누적되어 0.8 이상되는 여러변수의 선형적인 주성분을 관리하여야 한다. 이때 사례연구1과 같이 측정변동 관리를 위해 9개의 주성분을 관리해야 하는 경우도 있고, 사례연구2와 같이 2개의 주성분만을 관리해도 되는 경우가 발생함을 볼 수 있다.

6. 결 론

본 논문에서는 하나의 부품내에 다수의 측정치수가 있어 변동의 형태를 대표하는 소수의 주성분을 파악해서 집중관리할 수 있도록 하는 주성분 분석방법을 제시한다. 두가지 연구사례를 통하여 하나의 부품에 여러 측정치수가 있을 경우 변동의 형태를 대표하는 소수의 집중관리 대상인 주성분을 선정하는 과정을 버어니어 캘리퍼스를 사용할 경우와 하이트게이지를 사용할 경우로 나누어 살펴 보았다. 하나의 부품에 다수의 측정 치수가 존재할 경우에 측정변동이 어떤 형태로 크게 나타나는지를 빠른 시간내에 파악하기란 쉬운 일이 아니다. 본 연구에서 제시하는 측정시스템의 평가방법을 이용하면 다수의 측정치수간 상관관계에 근거한 측정치수의 변동경향을 신속히 파악하여 경제적 측정관리를 할 수 있다. 각 주성분의 선형관계식에서 변수의 계수를 제공하여 누적한 값이 0.8 이상 되는 변수들로 표현된 주성분을 관리하고, 측정변동을 엄격하게 관리해야할 경우에는 변수선정을 계수를 제공하여 누적한 값이 0.9이상 될 때까지 변수를 누적선정하여야 한다. 단일 어떤 주성분에서 한 변수의 계수제곱이 0.8 이상일 경우는 그 변수자체만 선정하여 gage R&R을 구하여 관리하면 된다. 반면, 주성분이 여러개 치수의 선형식으로 표현될 때는 관계된 치수데이터가 입력이 되면 각각의 gage R&R을 구하고, 주성분을 계산하여 그 주성분을 관리하여야 할 것이다. 본 연구의 결과는 한 부품내에 다수의 측정치수가 있거나 신규로 개발된 제품에

서 측정변동의 경향을 대표하는 소수의 주성분을 파악해서 집중관리할 수 있도록 하여 향후 측정시스템의 평가 및 관리를 경제적으로 하는데 도움이 되리라 기대된다.

참고문헌

- [1] Floyd, D. A. and Laurent, C. J., "Gauging: An Underestimated Consideration in the Application of Statistical Process Control", *Quality Engineering*, Vol. 8, No.1, pp.13-29, 1995.
- [2] Gnanadesikan, R., *Methods for Statistical Data Analysis of Multivariate Observations*, 2nd ed., pp. 5-15, 1997.
- [3] Johnson, R. A. and Wichern, D. W., *Applied Multivariate Statistical Analysis*, 3rd ed., pp.356-395, 1992.
- [4] King, D.W., *Statistical Quality Control Using the SAS System*, pp.288-301, 1995.
- [5] Montgomery, D. C. and Runger, G. C., "Gauge Capability Analysis and Designed Experiments, Part I", *Quality Engineering*, Vol.6, No.1, pp.115-135, 1993.
- [6] Montgomery, D. C. and Runger, G. C., "Gauge Capability Analysis and Designed Experiments, Part II", *Quality Engineering*, Vol.6, No.2, pp.289-305, 1993.
- [7] SAS, *SAS/STAT User's Guide*, SAS Institute Inc., ver. 6.0, 4th ed., Vol.2, pp.1241 -1264, 1990.
- [8] SAS, *SAS/QC Software: Usage and Reference*, SAS Institute Inc., ver.6.0, 1st ed., Vol.2, pp. 1582-1585, 1995.
- [9] SAS, *SAS/STAT Software: Changes and Enhancements through Release 6.12*, pp.571-702, 1997.