

PCA 를 이용한 세탁기 진동 · 소음 데이터 분석 및 특성 파악

Analysis about vibration and noise of washing machine using PCA

정동현*· 박상길*· 강귀현*· 이유엽***· 오재응†

Dong-Hyun Jung, Sang-Gil Park, Kwi-Hyun Kang, You-Yub Lee and Jae-Eung Oh

1. 서론

최근 다변량 분석방법은 통계학, 경영학, 사회학, 심리학, 생물학 등 여러 전공분야에서 복잡한 형태의 자료를 분석하는데 매우 활발히 사용되고 있다. 하지만 많은 변수들이 서로 상관이 있는 경우에는 분석에서 해석상의 복잡한 구조적 문제가 발생하기 때문에 복잡한 구조를 쉽게 이해하고 설명하기 위해서 PCA 가 많이 사용된다. 또한 변수들 간의 단위가 다를 경우 상대적 비교 및 분석에 많은 어려움이 있으므로 데이터들 간의 단위를 보정해줄 수 있는 개념이 들어가야 한다. 따라서 본 연구에서는 PCA 의 기본적인 이론을 이해하고 진동·소음 데이터를 적용하여 의미를 파악하고 특성 및 경향을 파악한다. 이를 위해서 세탁기 진동·소음 데이터를 측정하여 서로 비교 및 분석을 위하여 단위 보정을 한다. 이렇게 보정된 진동·소음 데이터의 공분산 행렬, 상관계수 행렬, 고유치, 고유벡터를 구하고 PCA 를 통하여 측정 변수들과 주성분(PC)와의 상관관계를 고려하여 영향도를 정량적으로 분석한다. 결과적으로 세탁기 진동·소음 측정데이터의 PCA 를 통하여 특성 및 경향을 분석한다.

2. PCA 이론

주성분 분석(PCA)은 서로 상관관계가 있는 변수들 사이의 복잡한 구조를 좀더 간편하고, 이해하기 쉽게 설명하기 위하여 사용되는 분석 기법이다. 즉, 변수들의 선형결합을 통하여 변수들이 가지고 있는 전체정보를 최대한 설명할 수 있는 서로 독립적인 새로운 인공변수들을 유도하여 해석하는 다변량 분석방법이다. 이러한 인공변수는 변수들의 선형결합으로 표시되며 이를 주성분(Principal Component)라 한다. P 개의 변수를 가진 N 개의 데이터를 matrix

로 나타내기 위해 x_{ij} 개념을 도입하여 i^{th} 변수와 j^{th} 샘플 데이터를 나타낸다. 여기서 $j=1,2,\dots,N$ 이고 $i=1,2,\dots,P$ 가 된다. P 개의 변수와 N 개의 데이터 샘플을 모두 고려하여 matrix 로 나타내면 식 (1)과 같이 나타난다.

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1j} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2j} & \dots & x_{2n} \\ x_{31} & x_{32} & \dots & x_{ij} & \dots & x_{3n} \\ \vdots & \vdots & & \vdots & & \vdots \\ \vdots & \vdots & & \vdots & & \vdots \\ x_{p1} & x_{p2} & \dots & x_{pj} & \dots & x_{pn} \end{bmatrix} \quad (1)$$

데이터 matrix 의 샘플 평균은 식(2)과 같이 나타난다.

$$\bar{x}_i = \frac{1}{n} \sum_{j=1}^n x_{ij} \quad (2)$$

데이터 matrix 의 공분산은 식 (3)와 식(4)과 같이 나타난다.

$$s_{ik} = \frac{1}{n-1} \sum_{j=1}^n (x_{ij} - \bar{x}_i)(x_{kj} - \bar{x}_k) \quad (3)$$

$$s_{ii} = \frac{1}{n-1} \sum_{j=1}^n (x_{ij} - \bar{x}_i)^2 \quad (4)$$

데이터들의 상관계수는 식 (5)를 통하여 알 수 있다.

$$r_{ik} = \frac{s_{ik}}{\sqrt{s_{ii}} \times \sqrt{s_{kk}}} = \frac{\sum_{j=1}^n (x_{ij} - \bar{x}_i)(x_{kj} - \bar{x}_k)}{\sum_{j=1}^n (x_{ij} - \bar{x}_i)^2 \sum_{j=1}^n (x_{kj} - \bar{x}_k)^2} \quad (5)$$

이를 통하여 데이터의 공분산 matrix 와 상관계수 matrix 로부터 p 개의 고유치와 고유벡터를 구하면 식 (6)과 식 (7)과 같이 나타난다.

$$\lambda_1 \geq \lambda_2 \geq \dots \lambda_p \geq 0 \quad (6)$$

오재응: 한양대학교 기계공학과
E-mail : jeoh@hanyang.ac.kr
Tel : (02) 2294-8294, Fax : (02) 2299-3153

* 한양대학교 대학원 기계공학과

** 호원대학교 기계·자동차공학부

$$U = \begin{bmatrix} u_{11} & u_{12} & \dots & u_{1p} \\ u_{21} & u_{22} & \dots & u_{2p} \\ u_{31} & u_{32} & \dots & u_{3p} \\ \vdots & \vdots & \ddots & \vdots \\ u_{p1} & u_{p2} & \dots & u_{pp} \end{bmatrix} \quad (7)$$

고유벡터와 기존의 변수와의 관계로부터 새로운 주 성분인 PC를 식 (8)을 통하여 나타낼 수 있다.

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_p \end{bmatrix} = \begin{bmatrix} u_{11} & u_{12} & \dots & u_{1p} \\ u_{21} & u_{22} & \dots & u_{2p} \\ u_{31} & u_{32} & \dots & u_{3p} \\ \vdots & \vdots & \ddots & \vdots \\ u_{p1} & u_{p2} & \dots & u_{pp} \end{bmatrix}^T \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_p \end{bmatrix} \quad (8)$$

3. PCA 적용

PCA를 이용한 진동·소음 데이터는 Fig. 1에 나타난 실험을 통하여 측정된 데이터를 분석한 것이다.

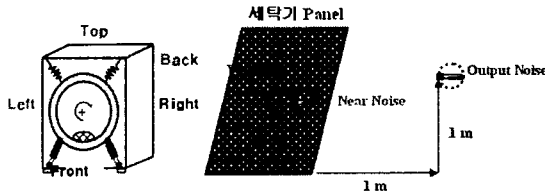


Fig. 1 PCA Measurement set-up

데이터의 변수는 전방 1m의 목적주파수이고 측정 위치는 각 면 별 3개의 측정점으로 총 15개 측정점의 데이터를 분석하였다. 측정된 데이터의 단위를 보정하여 공분산과 상관관계수 matrix는 Fig. 2 같다.

Covariance (Z)										
1.00000	0.45028	0.62225	0.69938	0.20328	0.52281	0.55983	0.20726	0.38016	0.28638	0.28228
0.45028	1.00000	0.76691	0.57816	0.51182	0.94019	0.44707	0.51906	0.30583	0.26220	0.26220
0.62225	0.76691	1.00000	0.60724	0.08724	0.97674	0.36743	0.67203	0.73624	0.57462	0.57462
0.69938	0.57816	0.60724	1.00000	0.32297	0.64120	0.35219	0.95423	0.51729	0.57971	0.57971
0.20328	0.51182	0.08724	0.32297	1.00000	0.41367	0.22219	0.26967	0.43164	0.26421	0.26421
0.52281	0.94019	0.97674	0.64120	0.41367	1.00000	0.56731	0.65332	0.34476	0.46171	0.46171
0.55983	0.44707	0.36743	0.35219	0.22219	0.56731	1.00000	0.39413	0.4206	0.27251	0.27251
0.20726	0.51906	0.67203	0.95423	0.26967	0.65332	0.39413	1.00000	0.43684	0.43684	0.43684
0.38016	0.30583	0.73624	0.51729	0.43164	0.34476	0.4206	0.43684	1.00000	0.39522	0.39522
0.28638	0.26220	0.57462	0.57971	0.26421	0.46171	0.27251	0.43684	0.39522	1.00000	1.00000

Correlation coefficients (Z)										
1	0.52	0.57	0.5	0.2	0.52	0.3	0.27	0.38	0.28	0.28
0.52	1	0.76	0.47	0.51	0.94	0.45	0.52	0.31	0.26	0.26
0.57	0.76	1	0.09	0.08	0.98	0.37	0.67	0.68	0.74	0.67
0.5	0.47	0.09	1	0.32	0.64	0.22	0.95	0.52	0.58	0.58
0.2	0.51	0.08	0.32	1	0.42	0.22	0.27	0.43	0.26	0.26
0.52	0.94	0.98	0.64	0.42	1	0.57	0.65	0.35	0.46	0.46
0.3	0.45	0.37	0.22	0.22	0.57	1	0.34	0.42	0.41	0.37
0.27	0.52	0.67	0.52	0.43	0.65	0.34	1	0.52	0.47	0.47
0.38	0.31	0.74	0.58	0.43	0.35	0.42	0.42	1	0.39	0.39
0.28	0.26	0.67	0.58	0.26	0.46	0.37	0.47	0.39	1	1

Fig. 2 Covariance and Correlation coefficient matrix 측정된 진동·소음데이터의 상관관계수 matrix에 의한 고유치는 Fig. 3과 같이 나타난다.

PC	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
Eigenvalue	8.9940	1.8130	0.7142	0.5114	0.1889	0.1616	0.0112	0.0022	0.0009	0.0004
비율%	0.8994	0.1813	0.0714	0.0511	0.0189	0.0162	0.0011	0.0002	0.0001	0.0000
누적 영향도	0.8994	0.8817	0.9531	0.9845	0.9974	0.9997	0.9999	0.9999	1.0000	1.0000

Fig. 3 Eigenvalue of Correlation coefficient matrix 누적 영향도를 보면 3개의 주 성분만으로 93.3%의 충분한 표현력은 가지고 있음을 알 수 있다. 변수들과 주 성분 사이의 관계를 나타내면 Fig. 4과 같이 나타난다.

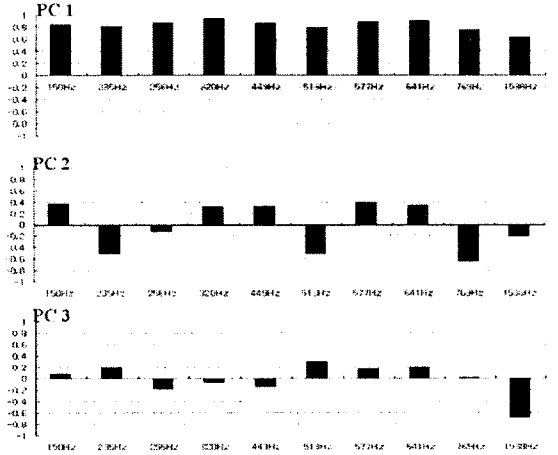


Fig. 4 Correlation between PC and variables PC1은 변수들과 양의 관계를 나타내며 진동과 소음이 양의 관계를 가짐을 알 수 있다. 또한 PC2는 235Hz, 513Hz, 769Hz가 주요한 영향을 미치며 PC3는 1538Hz가 주요한 영향을 미친다. 2개의 주 성분에 따른 측정위치 별 특성은 Fig. 5와 같다.

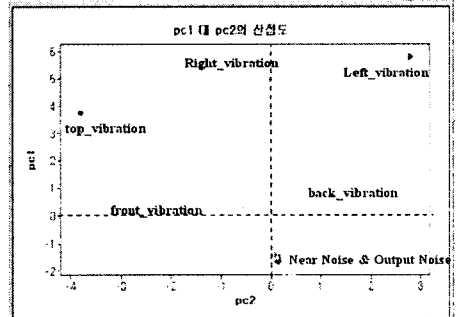


Fig. 4 Plot of PC and Measurement positions 세탁기 각 면에서의 가속도 값이 각 면의 근접 소음과 1m 출력 소음과는 다른 특성을 가진다. 특히 앞면, 뒷면의 진동은 상대적으로 소음과 유사하지만 윗면과 좌/우면은 상반된 특성을 가진다.

4. 결론

PCA의 이론적 정립 및 의미 파악하고 이를 토대로 진동·소음 데이터의 Normalizing 하여 PCA에 적용하였다. 이를 통하여 세탁기 각 면의 진동·소음 특성 및 경향을 분석하였다.