✓ 전체 세미나 순서

1. Smart Factory (스마트 팩토리)
2. Big Data란 무엇인가? / Big Data의 용도
3. 다변량(Multivariate, 多變量) Data 및 분석기법
4. 주성분/인자 분석이란?
5. Taguchi Robust Design (강건설계) & SN Ratio (신호 대 잡음비)의 개념
6. Minitab 및 MTS(Mahalanobis - Taguchi System : 마하라노비스-다구찌 시스템)를 활용한 사례 설명
   1) 에어컨 모터 이음 자동검사 시스템 구축 예
   2) 전수검사 LINE 에서의 양품/불량품 자동선별 예
   3) MD계산 간단실습 : Macro File 계산결과의 Excel Graph化
7. 마무리

1. Smart Factory(스마트 팩토리)

✓ 스마트공장은 제품 기획·설계·제조·공정·유통·판매 전 과정을 IT로 통합해 최소 비용·시간으로 고객맞춤형 제품을 생산하는 공장으로, 단순 설비자동화에 국한되지 않는다. 사물인터넷80(IoT270)을 기반으로 전 과정을 자동화·정보화하여 제조업 가치사슬 전체가 하나의 공장처럼 실시간 연동·통합된 생산체계를 지향한다.

✓ 필요기술로는

ⓐ 사물과 주위환경으로부터 정보를 얻는 “센싱 기술”

ⓑ 사물이 인터넷에 연결되도록 지원하는 “유무선 통신 및 Network Infra 기술”

ⓒ 각종 서비스 분야에 형태에 적합하게 정보를 가공하고 처리하거나 각종 기술을 융합하는 “서비스 Interface 기술”이 핵심이며,

ⓓ 대량의 데이터를 해석하는 “Big Data” 기술 등 사물인터넷 구성요소에 대한 해킹이나 정보 유출을 방지하는 “보안기술”도 필수적이다.

2. Big Data (빅데이터)란 무엇인가? /빅데이터의 용도

"빅데이터란 기존의 일반적인 기술로는 관리하기 곤란한 대량의 데이터 군이다."
예를 들어 '기존의 일반적인 기술로는 관리하기 곤란'하다는 것은 현재 기업의 데이터베이스의 주류를 차지하는 관계형 데이터베이스로는 관리할 수 없는 복잡한 구조의 데이터를 가리킨다. 또한 데이터 양이 증가해서 데이터에 대한 쿼리(질의) 응답시간이 허용범위를 넘는 상태를 가져오는 방대한 데이터를 가리킨다.

3V로 나타내는 빅 데이터의 특성

빅데이터의 특성은 다음처럼 V로 시작하는 **세 가지 키워드**로 나타낼 수 있다.



* 출처: TDWI Research, Big Data Analytics Report(2011)

〈그림 1〉 빅 데이터의 3가지 특성

1) 데이터 양(Volume)

빅데이터의 정의로 본다면 기존 기술로는 관리할 수 없는 데이터양으로 현재는 수십 테라바이트에서 수 페타바이트 정도라는 견해가 많다. 물론 기술이 진화함에 따라 이 수치는 증가할 것이다.

2) 다양성(Variety)

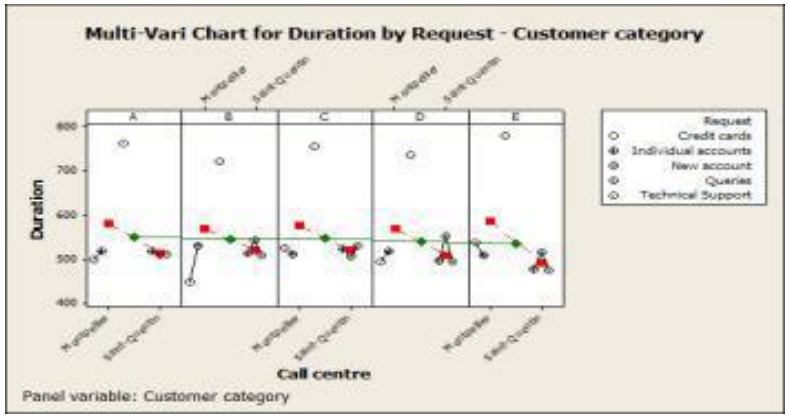
기존에 기업에서 사용하던 구조화된(판매 데이터, 재고 데이터 등) 데이터 외에 최근 급증하고 있는 인터넷의 텍스트 데이터, 위치정보, 센서 데이터, 동영상 등 현재 기업에서 주로 사용하는 관계형 데이터베이스로 다루기 어려운 구조화되지 않은 데이터(비구조화 데이터)를 분석하여 유용한 지식을 얻을 수 있도록 한다.

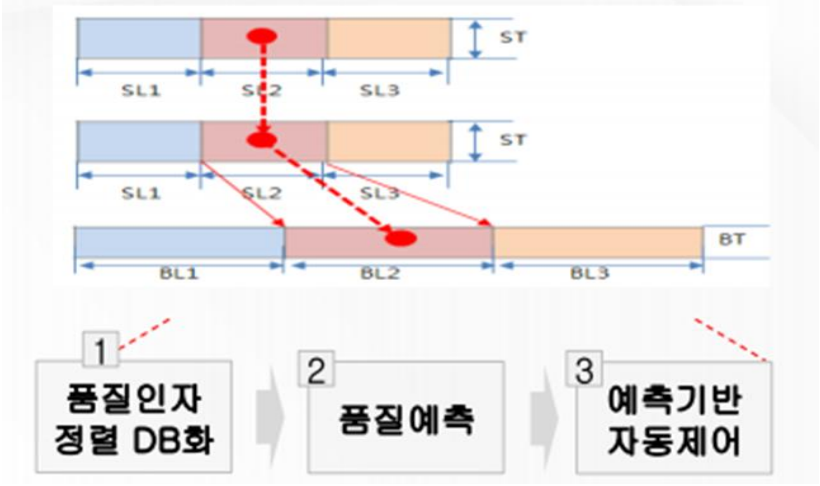
3)속도(Velocity)

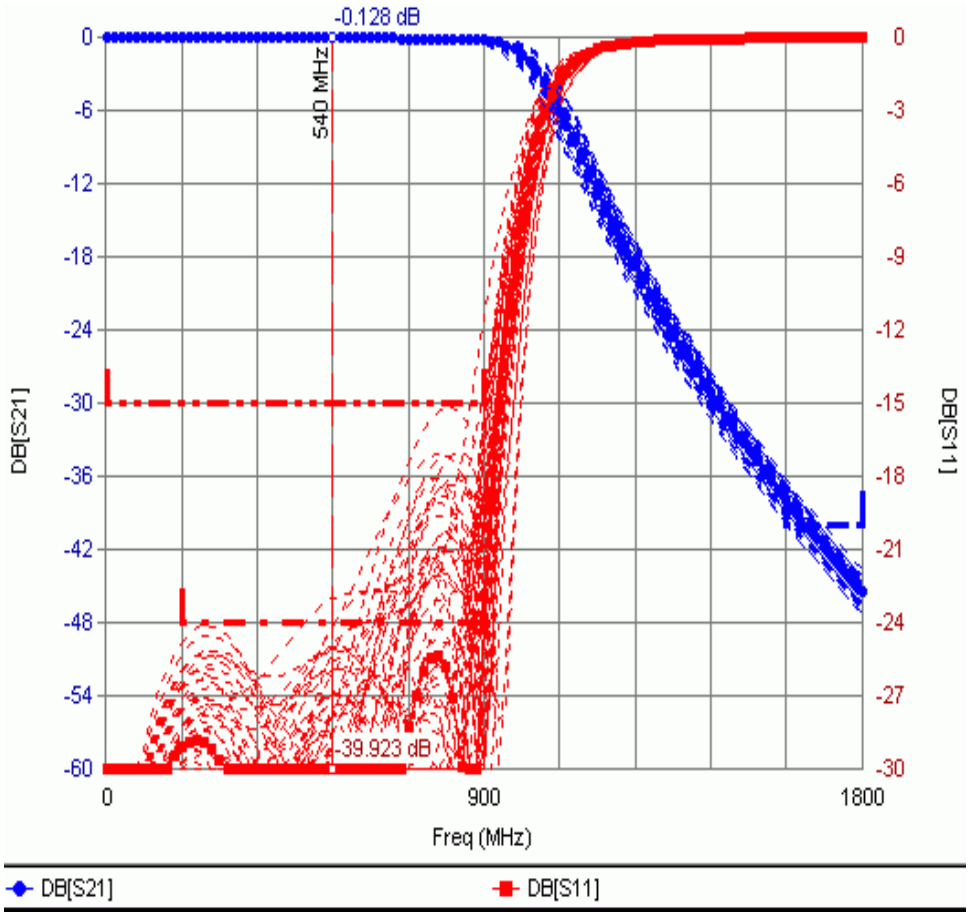
시시각각 발생하는 데이터에 대한 분석처리는 아주 도적적인 과제다. 스트림 데이터 처리 등 새로운 기술이 등장해 분석 처리가 가능해진 점도 최근의 빅데이터 붐을 뒷받침하는 요인이다.



- ✓ 다변량분석, 실험계획법, 다구찌 품질공학을 이용한 품질불량/ 설비고장 주요인자 감지 → 주요 인자에 대한 영향도 및 인과관계 확인
  → 알고리즘 개발 및 품질예측을 통한 자동제어
- ✓ 전체 Data (Big Data)를 활용한 사전 품질 예측 및 품질 사전 최적화
  → 몬테카를로 시뮬레이션 (Monte Carlo Simulation)



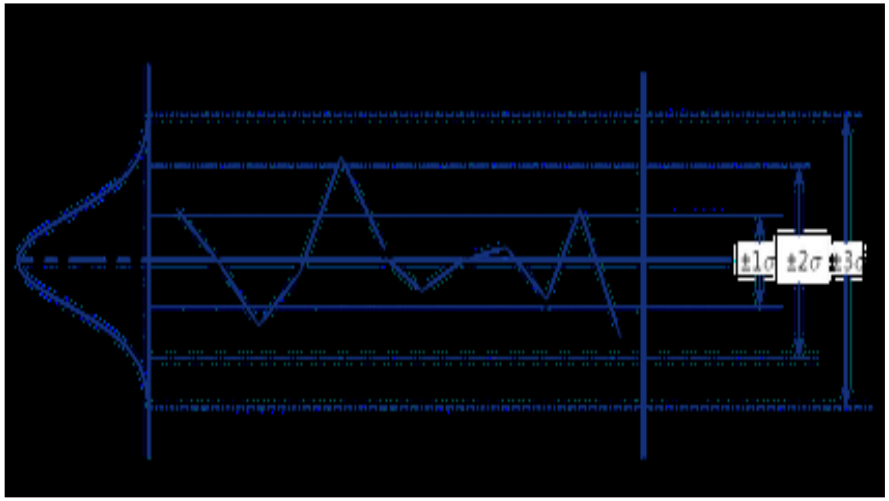




✓머신비전(Machine Vision)을 이용한 품질검증 – 사례) 포장 후 검사기
: 불량판단 인공지능 알고리즘

✓IOT (및 Big Data)를 통계적 품질관리에 적용한 품질패턴 예측 및 불량방지 알고리즘



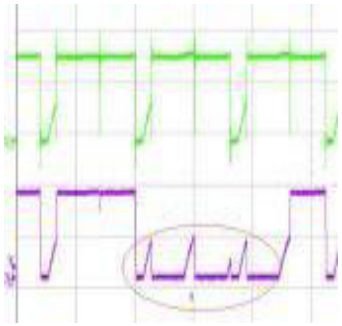


➔ 비접촉식 3D 측정 센서/ 장비를 포함한 제품의 품질을 측정할 수 있는,
In-line 가능한 다양한 센서류의 개발 및 측정 알고리즘 개발이 필요함.

✓스마트 팩토리의 핵심영역인 예지보전(Predictive Maintenance)과 품질 최적화 (Quality Optimization)구현을 위한 더 많은 초저가, 초정밀 Sensor의 개발과 이 Sensor를 활용한 통신 가능한 Device의 개발 및 표준화가 요구됨.



더 많은 초저가, 초정밀 Sensor 및 통신 가능한 Device



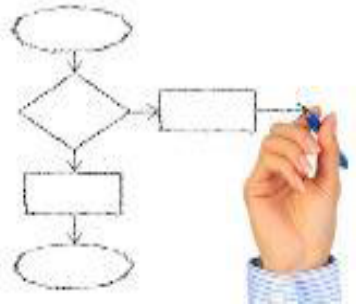
이상신호 감지



문제해결 Solution

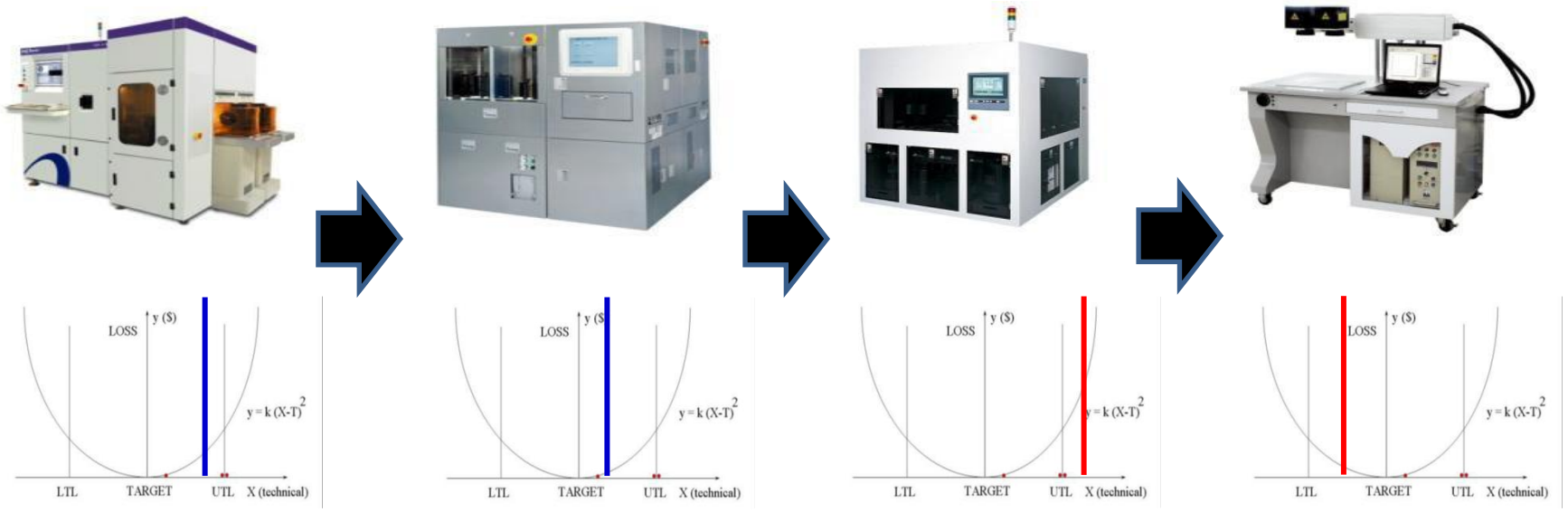


예지/ 예방 보전

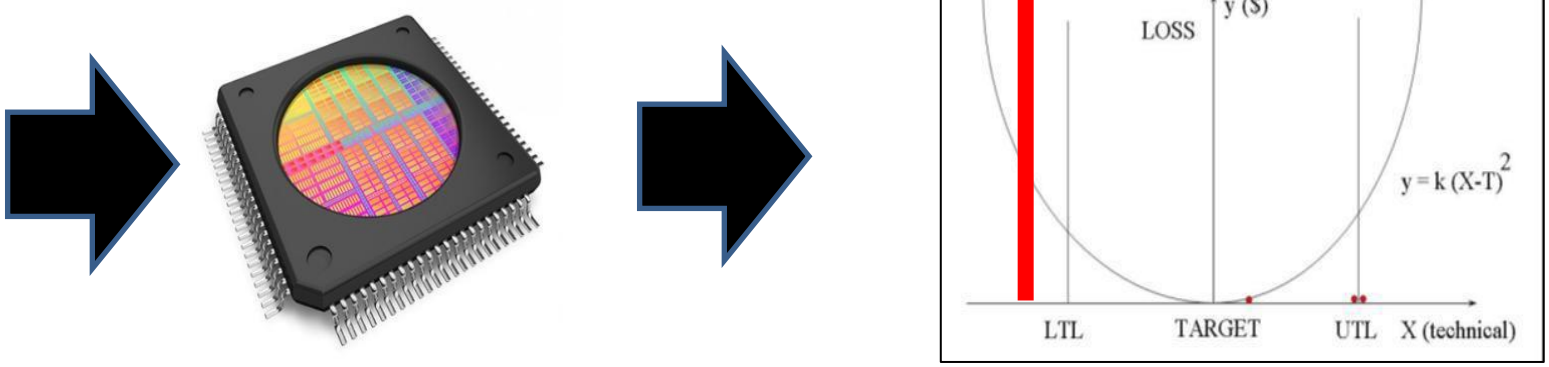


사전조치
알고리즘 개발

✓여러 개의 공정/ 장비의 Process Indicator와 이 공정을 통해 나오는 제품 (산출물)의 Output Indicator의 빅데이터 분석을 통해 이상 공정/ 설비를 발견

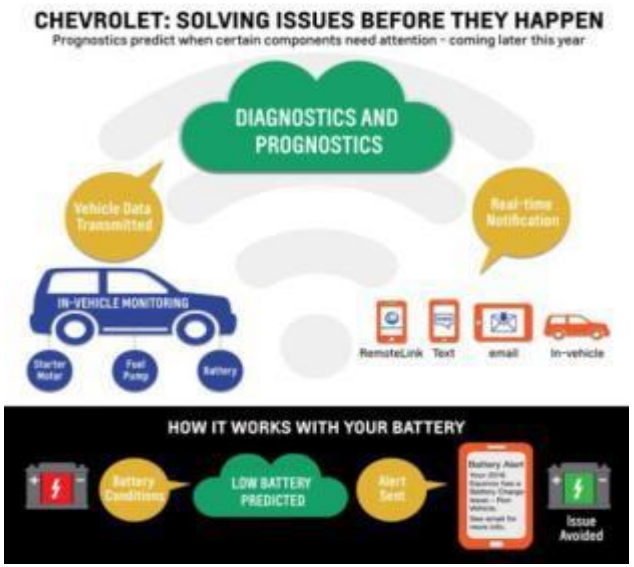


Process Indicator 1 Process Indicator 2 Process Indicator 3 Process Indicator 4



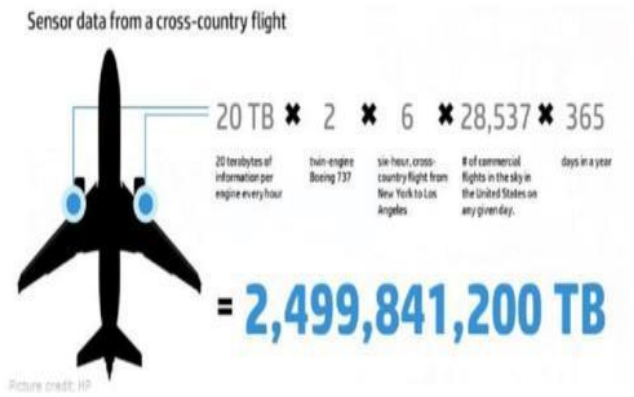
Output Indicator

✓각종 센서, Big Data를 이용한 예지Data 정보확보 후 예지보전, 고장이상 신호에 따른 원격 서비스, 실제 고장 시 (사전부품준비)통한 정비 시간 단축, 고객사 설비 常時 모니터링



Big Data, Sensor 고장신호 유형 파악

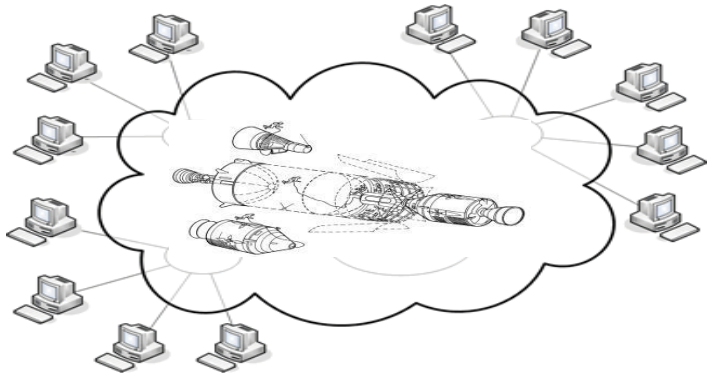
✓예지보전 및 사전조치
✓(고객설비) 실시간 모니터링
✓경고 신호에 따른 원격 조치
✓고장 (예지)신호 유형에 따른 사전조치 알고리즘 및 자동 조치
✓고장신호 유형에 따른 신속 정비
← 부품 사전 준비



✓항공산업에서 **Big Data를 활용 MRO(Maintenance, Repair, and Overhaul)에 활용하고 있음.**
✓테슬러는 자동차 이상 시 조치의 **70%~80%를 원격 서비스와 펌 웨어 업그레이드를 통해 수행**

✓R&D 분야에서는 IOT를 이용한 (SCM 내) Concurrent Engineering 및 제품/ 공정의 사전 시뮬레이션을 통한 검증 및 최적화
✓ Industry 4.0 시대를 대비한 비즈니스 모델, 제품/ 서비스 개발

✓R&D/ 제품개발 Process
1) 클라우드 이용한 도면공유
2) 개발프로세스 상 Output 공유 :
Proto 성능 Test 결과, Pilot 결과 등
3) SCM 상 개발진행현황 실시간 공유
4) Design 결과 사전 시뮬레이션/ 최적화





✓제품개발
1) Smart M/C, 첨단소재 (3D 등) 개발 :
친환경, 초경량, 초소형, 초고강도
2) Digital Sensor, 측정장치/ 장비 개발
3) Embedded Parts 개발
4) S/W 개발 : CPS, 응용S/W 등
5) 스마트 팩토리용 보안 솔루션

3. 다변량(Multivariate, 多變量) Data 및 분석기법

"여러 개의 변수에 의해서 다양한 관점에서 관찰된 데이터를 말하며, 각 변수별 서로간의 미치는 영향이나 샘플간의 상호 관계를 모든 품질요소의 관점에서 살펴보는 것이 목적으로, 이를 위해서는 다변량 자료분석 기법이 요구된다. p개의 변수에 대해 n개의 개체를 관측한 경우의 자료는 행렬 (n×p)로 표현된다.

1) R방법 : 변수들간 상관관계를 통하여 유사한 변수의 선별이 주목적

ⓐ 주성분분석 : Scree(기울기가 급격히 변하는 지점을 주성분의 개수로 판단) & Score Plot
ⓑ 인자분석

2) Q방법 : 변수공간에서 개체간의 거리(Distance)를 이용, 유사성(Similarity) 판별

ⓐ 군집분석 : 그룹이 알려지지 않은 경우 개체들을 서로 가까이 있는 유사한 것끼리 군집화 하거나 상관관계가 큰 변수들끼리 집단으로 묶어주는 방법이다.
ⓑ 판별분석 : 알려진 그룹이 있는 표본이 있을 경우 관측치를 둘 이상의 그룹으로 분류한다. 판별분석을 사용하면 그룹에 대한 예측변수의 기여 정도를 알 수 있다.
ⓒ 대응분석 : 범주형 변수 간의 관계를 탐색하기 위한 방법이다.

3) 다변량 추정과 검정에 관련된 기법

ⓐ Hotelling의 T^2 검정, 다변량 분산분석(MANOVA), 다중회귀분석(Logistic포함)

4. 주성분/인자 분석이란?

원래 변수의 공분산 구조를 이해하고 이 구조를 사용하여 보다 적은 수의 변수를 만드는 데 유용하다. 즉, 여러 개의 변수로부터 전체 변동을 잘 설명할 수 있는 중요한 몇 개의 성분을 얻는데 그 목적이 있으며, 여러 개의 변수를 서로 상관관계가 없는 주 성분으로 줄이기 때문에 원 자료를 훨씬 간단한 형태의 자료로 축약시키며, 원 자료의 복잡성 때문에 찾을 수 없었던 異常値를 찾게 하므로 주성분은 그 자체가 목적이기 보다는 다른 분석을 위한 중간과정으로 사용되는 경우도 많다. 예를 들면 다중공선성(VIF : Variance Inflation Factor)이 있는 변수들을 회귀분석하기 위한 방법으로 사용되기도 한다.

서로 연관이 있는 변수들이 관측되었을 때, 이 변수들이 가지고 있는 정보들을 최대한 확보하는 적은 수의 새로운 변수들을 생성하는 방법이다.
즉, 서로 연관이 있는 p개의 변수들에 대하여 주요한 관심사 중의 하나는 이 변수들이 가지고 있는 변이를 측정하는 것이다. 변이란 이 변수들이 가지고 있는 정보의 양이라고 할 수 있는데, 주성분분석에서는 원래의 변수들이 가지고 있는 변이의 양을 가장 많이 확보하는 순서대로 변수들의 선형결합을 이용하여 새로운 변수를 구하는 과정이다.

인자분석도 주성분 분석과 마찬가지로 데이터 공분산 구조를 보다 적은 수의 차원으로 요약하는 방법인데, 인자분석에서 중요시하는 것은 데이터의 변동성의 많은 부분과 연관된 차원을 설명할 수 있는 기본적인 “인자(Factor)”를 식별하는 것이다.

Minitab - 제목 없음

파일(F) 편집(E) 데이터(A) 계산(C) 통계분석(S) 그래프(G) 편집기(D) 도구(T) 창(W) 도움말(H) 보조 도구(N)

세션

08-20 10:16:54

Minitab을 시작합니다. 도움말을 보려면 F1 키를 누르십시오.

```
MTB > print eigenvec
```

데이터 표시

행렬 eigenvec

```
-0.558403  -0.131014   0.007947   0.550481  -0.606572
-0.313454  -0.628612  -0.549130  -0.452781   0.006230
-0.568275  -0.004581   0.117257   0.268143   0.769021
-0.486519   0.309976   0.454965  -0.647947  -0.201116
 0.174055  -0.701169   0.691123   0.014963   0.013846
```

```
MTB > print eigenval
```

데이터 표시

```
eigenval
   3.02906   1.29115   0.57246   0.09543   0.01192
```

```
MTB >
```

워크시트 1 ***

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20	C21	C22	C23	C24
	지역번호	총인구수(천명)	교육중위수(년)	총고용인수(천명)	보건의료 종사자수(백명)	소득중위수(만불)	eigenval																	
1	1	5.935	14.2	2.265	2.27	2.91	3.02906																	
2	2	1.523	13.1	0.597	0.75	2.62	1.29115																	
3	3	2.599	12.7	1.237	1.11	1.72	0.57246																	
4	4	4.009	15.2	1.649	0.81	3.02	0.09543																	
5	5	4.687	14.7	2.312	2.50	2.22	0.01192																	
6	6	8.044	15.6	3.641	4.51	2.36																		
7	7	2.766	13.3	1.244	1.03	1.97																		
8	8	6.538	17.0	2.618	2.39	1.85																		
9	9	6.451	12.9	3.147	5.52	2.01																		
10	10	3.314	12.2	1.606	2.18	1.82																		
11	11	3.777	13.0	2.111	2.83	1.80																		
12	12	1.530	13.8	0.798	0.84	4.25																		
13	13	2.768	13.6	1.336	1.75	2.64																		
14	14	6.585	14.9	2.763	1.91	3.17																		
15																								
16																								
17																								
18																								
19																								
20																								

Proj...

일정 - dyu@lsis.c... Lync Minitab - 제목 없... 프레젠테이션1 - ...

오전 10:23
2015-08-20

● 주성분 분석: 총인구수(천명), 교육중위수(년), 총고용인수(천명), 보건의료 종사자수(백명), 소득중위수(만불)

상관 행렬에 대한 고유 분석

고유값	3.0291	1.2911	0.5725	0.0954	0.0119
비율	0.606	0.258	0.114	0.019	0.002
누적	0.606	0.864	0.979	0.998	1.000

변수	PC1	PC2	PC3	PC4	PC5
총인구수(천명)	-0.558	0.131	-0.008	-0.550	0.607
교육중위수(년)	-0.313	0.629	0.549	0.453	-0.006
총고용인수(천명)	-0.568	0.005	-0.117	-0.268	-0.769
보건의료 종사자수(백명)	-0.487	-0.310	-0.455	0.648	0.201
소득중위수(만불)	+0.174	0.701	-0.691	-0.015	-0.014

<고유벡터>

해설) 5개 변수들의 상관행렬에 대한 고유값, 설명비율 및 누적 설명비율을 나타낸다. 1보다 큰 고유값의 개수는 2개이며, 처음 두 개 고유값에 의해 전체 변동의 86.4%가 설명된다. 제1열은 제1주성분의 계수를 나타내는 고유벡터이며, 따라서 제1주성분은

PC1 = -0.558*총인구수-0.303*교육-0.568*총고용인수-0.487*보건의료 종사자수+0.174*소득

와 같이 표현할 수 있고, 나머지 다른 주성분에 대해서도 같은 요령으로 표현할 수 있다.

결과적으로 주성분분석은 상관행렬 또는 공분산 행렬의 고유값 및 고유벡터를 구하는 문제로 귀착된다고 볼 수 있다. 처음 두 개 주성분만으로도 표본변동을 매우 잘 설명하고 있으며, 5개 변수와 14개 관측치로 구성된 자료를 2개 주성분에 대한 14개 관찰 자료로 축약(reduced)하게 된다.

● 데이터 표시

1) Eigenvalue(固有값)

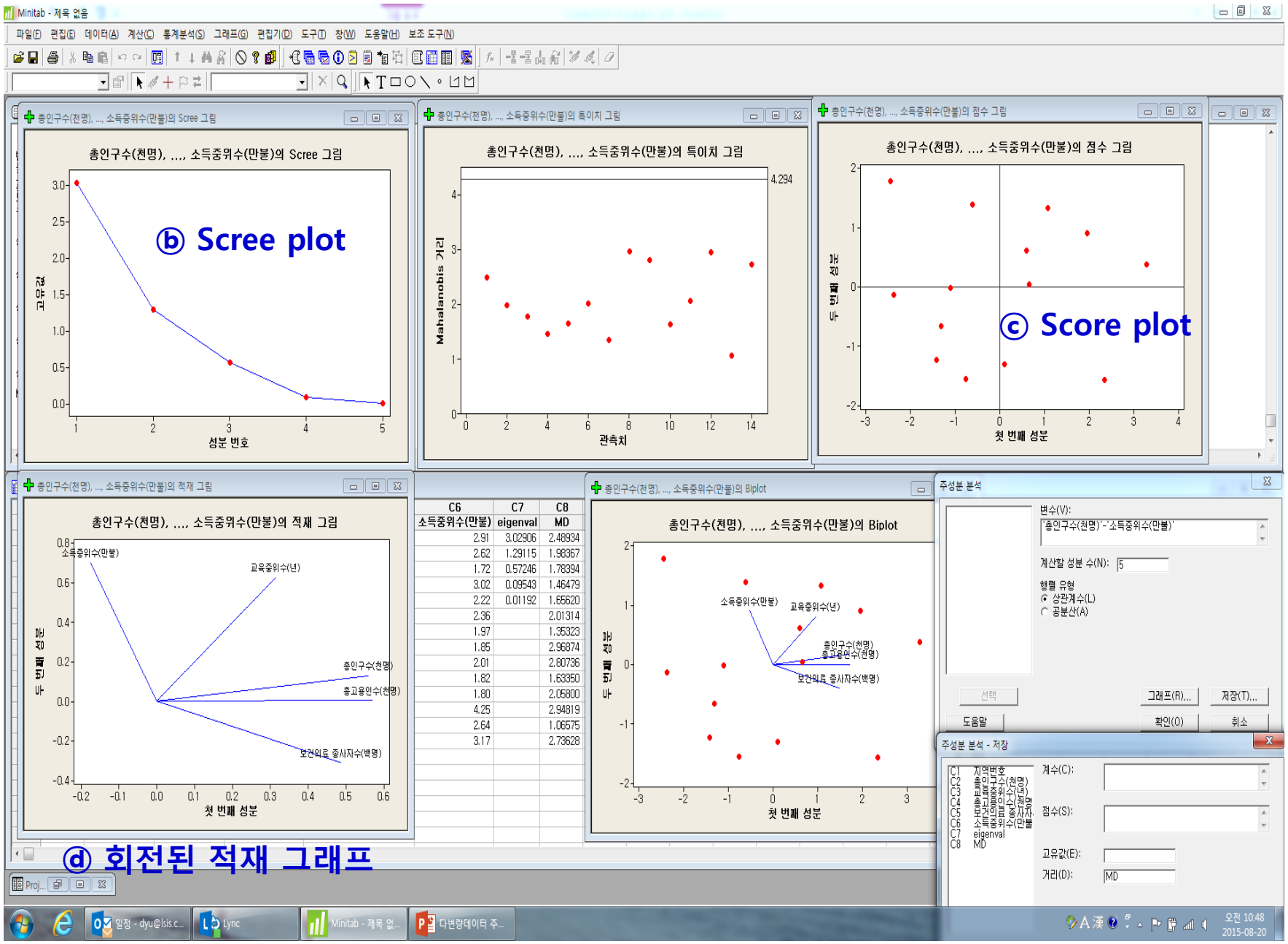
ⓐ 3.02906 1.29115 0.57246 0.09543 0.01192

상관행렬의 고유값은 그 크기 순으로 3.0291>1.29115>0.5725>0.0954>0.0119 와 같다. **I번째 크기의 固有값은 i번째 주성분을 결정하며, 固有값의 상대적인 크기는 주성분의 설명비율과 관련된다.** 예를 들어 최대 고유값 3.0291은 고유값의 총합 5의 약 0.606이므로 제1주성분은 총변동의 60.6%를 설명한다고 해석한다. 처음 두 개 고유값에 의한 누적설명 비율은 (3.291+1.2911)/5 = 0.864이므로 약 86.4%를 설명하게 된다

2) 행렬(Matix) eigenvector(고유벡터)

-0.558403	-0.131014	0.007947	0.550481	-0.606572
-0.313454	-0.628612	-0.549130	-0.452781	0.006230
-0.568275	-0.004581	0.117257	0.268143	0.769021
-0.486519	0.309876	0.454965	-0.647947	-0.201116
0.174055	-0.701169	0.691123	0.014963	0.013846

固有값에 대응하는 고유벡터를 나타내는 행렬이다. **제1열은 최대 固有값 3.0291에 대응하는 固有벡터 e1=(-0.558, -0.313, -0.568, -0.487, 0.174)이다.** 똑같은 요령으로 제2열은 두 번째 고유값 1.2911에 대응하는 고유벡터이다.고유벡터의 요소들은 5개의 원 변수 총인구수~소득에 의한 일차결합(“주성분 ‘)의 계수를 나타낸다.


Minitab - 제목 없음
파일(F) 편집(E) 데이터(A) 계산(C) 통계분석(S) 그래프(G) 편집기(D) 도구(T) 창(W) 도움말(H) 보조 도구(N)
총인구수(천명), ..., 소득중위수(만불)의 Scree 그림
총인구수(천명), ..., 소득중위수(만불)의 Scree 그림
고유값
성분 번호
ⓑ Scree plot
총인구수(천명), ..., 소득중위수(만불)의 특이치 그림
총인구수(천명), ..., 소득중위수(만불)의 특이치 그림
4.294
Mahalanobis 거리
관측치
총인구수(천명), ..., 소득중위수(만불)의 점수 그림
총인구수(천명), ..., 소득중위수(만불)의 점수 그림
두 번째 성분
첫 번째 성분
ⓒ Score plot
총인구수(천명), ..., 소득중위수(만불)의 적재 그림
총인구수(천명), ..., 소득중위수(만불)의 적재 그림
소득중위수(만불)
교육중위수(년)
총인구수(천명)
총고용인수(천명)
보건의료 종사자수(백명)
두 번째 성분
첫 번째 성분
C6 C7 C8
소득중위수(만불) eigenval MD
2.91 3.02906 2.48934
2.62 1.29115 1.98367
1.72 0.57246 1.78394
3.02 0.09543 1.46479
2.22 0.01192 1.65620
2.36 2.01314
1.97 1.35323
1.85 2.96874
2.01 2.80736
1.82 1.63350
1.80 2.05800
4.25 2.94819
2.64 1.06575
3.17 2.73628
총인구수(천명), ..., 소득중위수(만불)의 Biplot
총인구수(천명), ..., 소득중위수(만불)의 Biplot
소득중위수(만불)
교육중위수(년)
총인구수(천명)
총고용인수(천명)
보건의료 종사자수(백명)
두 번째 성분
첫 번째 성분
주성분 분석
변수(V):
'총인구수(천명)' '소득중위수(만불)'
계산할 성분 수(N): 5
행렬 유형
상관계수(L)
공분산(A)
선택
그래프(R)...
저장(T)...
도움말
확인(O)
취소
주성분 분석 - 저장
C1 지역번호
C2 총인구수(천명)
C3 교육중위수(년)
C4 총고용인수(천명)
C5 보건의료 종사자수
C6 소득중위수(만불)
C7 eigenval
C8 MD
계수(C):
점수(S):
고유값(E):
거리(D): MD
ⓓ 회전된 적재 그래프
Proj...
일정 - dyu@lsis.c...
Lync
Minitab - 제목 없...
다변량데이터 주...
오전 10:48
2015-08-20

● 그림 설명

Ⓐ 제1주성분 : PC1

PC1(제1주성분) = -0.558*총인구수-0.303*교육-0.568*총고용인수-0.487*보건의료 종사자수
+0.174*소득

에서 처음 네 개의 변수의 계수는 음수 값인 반면 소득의 계수는 양수(+)이므로 두 변수 집합 간의 대비를 나타낸다고 볼 수 있지만, 소득의 계수는 상대적으로 작다. 따라서 제1주성분 PC1은 인구수, 교육, 총고용인수, 보건의료 종사자수와 밀접한 관련을 나타낸다고 볼 수 있다. 반면에 제2주성분 PC2는 교육 및 소득 두 변수와 보건의료 종사자간의 대비를 보여준다. 처음 두 개 주성분만을 사용 하더라도 자료의 변동 중 86%를 설명하며, 나머지 3개 주성분에 의해서는 약 14%밖에 설명하지 못함을 알 수 있다.

Ⓑ Scree plot은 고유값의 크기에 따라 Plot한 것으로써 고유값의 상대적 크기를 비교하는데 편리하다. 주성분 개수를 정할 때 상관행렬의 고유값이 1보다 큰 개수를 보통 택하지만 Scree 도표에서 고유값의 상대적 크기가 급격하게 작아지는 고유값 번호를 기준으로 주성분 개수를 정하기도 한다.

Ⓒ Score Plot은 처음 두 개 주성분에 대한 주성분 점수의 산점도이다. 14개 지역에 대한 5개 변수의 측정값을 2개 주성분점수로 대신할 수 있다. 이러한 주성분 점수의 산점도를 이용하면 자료의 정규성이나 이상값(Outlier)을 찾는데 유용하다. 산점도의 전체적 모양이 타원형이면 대략 정규분포를 따른다고 볼 수 있는데, 이 Score Plot는 타원에 가깝다.

5. Taguchi Robust Design (강건설계) & SN Ratio (신호 대 잡음비)의 개념

1) 강건설계 (Robust Design)

인자가 품질특성의 평균에 대해 가지는 효과뿐만 아니라 분산에 미치는 효과에 대해서도 초점을 맞추고 있다. 어떤 인자가 품질특성의 분산에 영향을 미치는지 결정함으로써 분산을 줄일 수 있는 통제 가능한 인자에 대한 조건을 설정할 수 있으며, 품질특성이 통제 불가능한 인자의 변화에 대해서는 둔감하게끔 한다.

주목적은 먼저 **분산을 줄이는 인자에 대한 최적조건을 발견하고(1St Step)**, 그 다음에 **품질특성의 평균치를 목표치(Target)로 이동(2nd step)**하여 **품질 손실을 최소화**하는데 있다.

2) 품질특성의 종류

(1) 정특성(Static Characteristic) : 신호인자 無

a) 망소특성(Smaller, the Better) : 적을수록 좋은 특성 ~ 주로 이롭지 못한 폐해항목
자동차 배기 가스량(공해물질), 마모량 , 불순물 량, 저항치, 에너지 소모량, 약의 부작용

b) 망대특성(Larger the Better) : 클수록 좋은 특성 ~ 연비, 효율, 인장/접착강도, 사용수명

c) 망목특성(Nominal the Better) : 일정할수록 좋은 특성 ~ 길이, 중량, 코팅 두께

d) 기능창(Operating Window) : 망소와 망대특성이 혼합 ~ 용접, 복사기, 스크류 체결

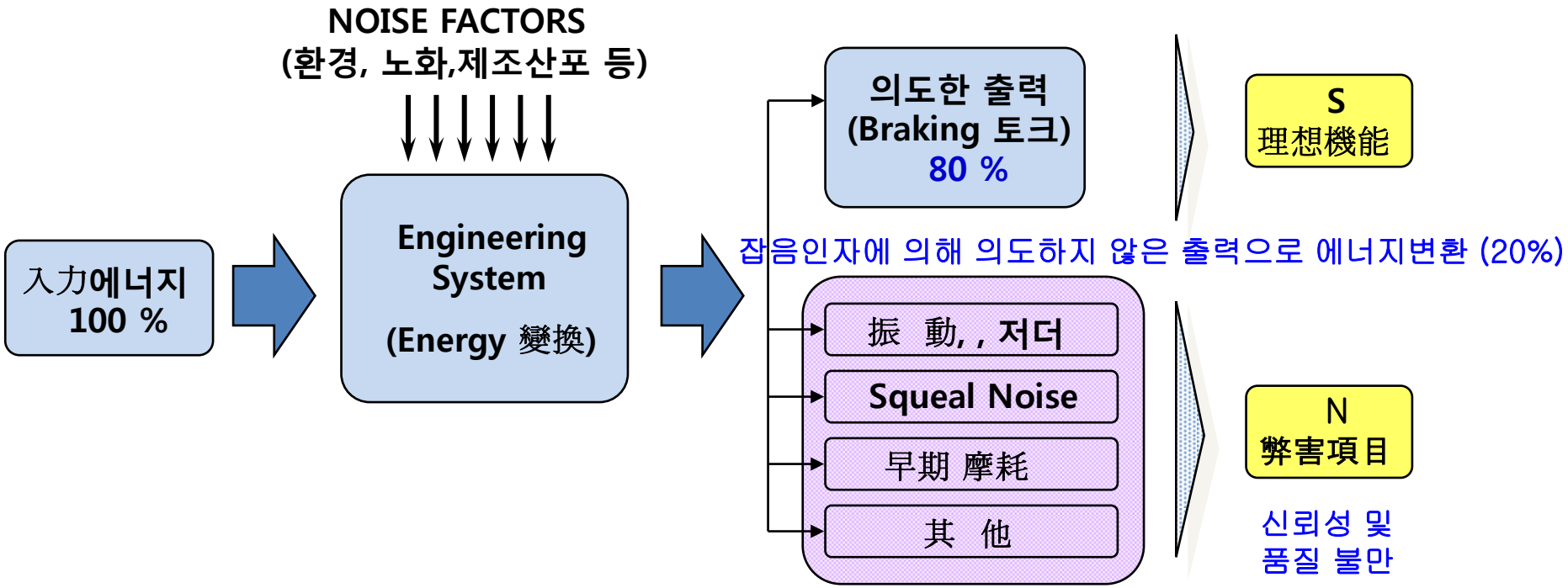
e) 계수치특성 : PCB 缺陷, 缺點數, 合格/不合格, 良品/不良品, 코팅품질 등 오감관련 특성 多
: a) 단순계수치, b) 0, 1 양품/불량데이터, c) 언어데이터의 점수화, d)가중치 계수치

(2) 동특성(Dynamic Characteristic) : 신호인자(입력) 변화에 대응하여 출력이 변화

~ 계측기의 정밀도, 통신시스템의 성능, 가속Pedal의 압력차이에 따른 차 속도 변화

★ S/N比 (Signal to Noise Ratio) 란 ?

- 信號(Signal) 와 雜音(Noise) 출력의 상대적인 크기
- SN比를 크게 하는 의미는 원하는 목적기능은 크게 하고, 폐해항목은 작게 하여 품질손실(Quality Loss)을 작게 한다는 의미임 ➔ 또한 신뢰성을 함께 높일 수 있음



$$SN = \frac{S \quad (\text{바라는 목적으로 변환된 Energy量})}{N \quad (\text{바라지 않는 것으로 변환된 Energy量})} = \frac{\text{Effect of Mean}}{\text{Variability Around Mean}}$$

6. Minitab 및 MTS(Mahalanobis - Taguchi System)를 활용한 전수검사 LINE 에서의 양품/불량품 자동선별 및 설비고장 진단 및 예측

1) 과정요약

MTS/MTGS는 마하라노비스(Mahalanobis)의 다차원 공간에서의 거리 측정방법과 Taguchi 품질공학의 SN비 개념을 적용한 다변량 데이터의 패턴인식 시스템 개발방법

2) 기대 효과

-컴퓨터공학 분야의 문자인식
-음성인식 방법개발
-의료분야의 환자 건강도 진단과 평가
-생산현장의 설비이상 진단과 예측 등에 활용

3) 학습 목표

MTS/MTGS의 기본 아이디어는 다변량 특성치를 하나의 종합적 지표인 마하라노비스 거리로 나타내고, 마하라노비스 거리 (Mahalanobis Distance) 변화에 민감도가 높은 특성치를 패턴분류에 중요한 특성치로 선정하는데 있다.
MTS에서 정상그룹(기준집단)에 속한 시료들은 마하라노비스 거리(MD)가 1에 가까운 값을 갖고 개별시료들의 마하라노비스 거리는 큰 차이가 나지 않는 동질적인 그룹이지만, 정상그룹(기준집단)에 속하지 않는 시료(비정상그룹)의 마하라노비스 거리는 큰 값을 가지며, 개별시료들의 거리 차이도 커서 동질적이지 않은 경우가 대부분이다. 정상그룹과 비정상그룹의 이러한 차이를 활용하여 Taguchi 품질공학의 파라메타 설계 방법으로 패턴인식에 중요한 측정변수와 중요하지 않은 측정변수를 선정, 공정 및 설계에 적용한 후 재현성을 확인한다.

1 단계 : “정상 그룹 (Normal Group)” 또는 참고 그룹(reference Group) 정의

2 단계 : k개의 변수 선정. $2 < k < 1000$

3 단계 : 정상 그룹으로부터 n개의 표본(sample)을 추출 하여 각 표본에 대한
k개 변수의 데이터 측정 (n>k 이어야 하지만, 아닌 경우 몬테카를로 시뮬레이션 이용함)

4 단계 : 정상 그룹의 각 표본 데이터로부터 MD 계산

4.1 : 데이터의 표준화(Normalize the Data), 변수들간 VIF가 크면 직교화 先진행

4.2 : k x k의 correlation matrix (상관행렬) 작성

4.3 : Correlation matrix의 Inverse matrix (逆행렬) 구함

4.4 : Inverse matrix를 사용 하여 n개의 표본의 각각에 대하여 MD, D^2을 계산
하여 Mahalanobis Space 구축 ➔ Excel을 이용, Graph化

●1단계

5 단계 : r개의 검사 대상 표본(Abnormal Group)에서 k개 변수의 데이터 측정 ➔ Excel 이용, Graph化

6 단계 : 각 검사 대상 표본의 MD, D^2을 계산 및 문턱 값(Threshold Value, 역치, 한계치)결정

7 단계 : 각 대상 표본 검정(정상 여부 판단)

8 단계 : 시스템 최적화

8.1 : k개 변수를 직교표에 할당, 1수준: 변수 사용, 2수준: 변수 사용 안 함.
직교표의 각행은 MD 계산을 위한 k개 변수의 subset(일부 변수만 사용)

8.2 : 직교표의 각행에 대한 r개의 검사 대상 표본의 MD(=D^2)계산

●2단계

8.3 : 각 행에서 r개의 MD로 SN비를 계산

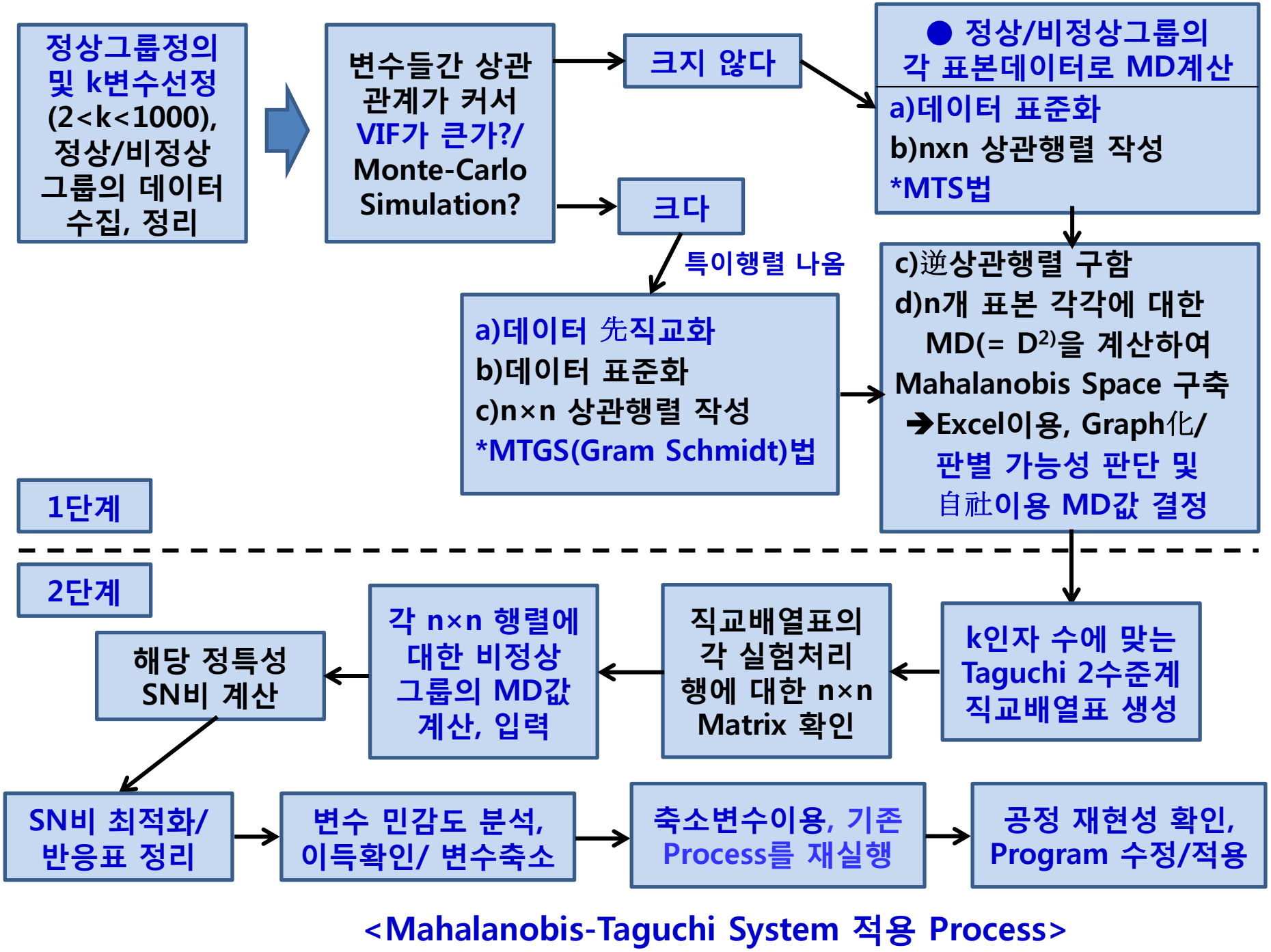
8.4 : SN비에 영향이 적은 변수 제거 : 이득이 큰 변수를 선정함(정확도 향상)

8.5 : 감소된 변수들로 확인실험 시행, 주요변수 확정 ➔ 문턱값 재검토, 제반 여건고려/선정

●3단계

9단계 : 설계개선 작업
최종 선정된 주요 변수들로 적절한 다구찌 직교배열표 선정 및 수준할당, 잡음인자를 고려해서 실험을 진행 ➔ MD값을 계산 후에 최적화 ➔ 재현성 여부 판정 ➔ 설계에 반영

<Mahalanobis-Taguchi System 적용 Process>



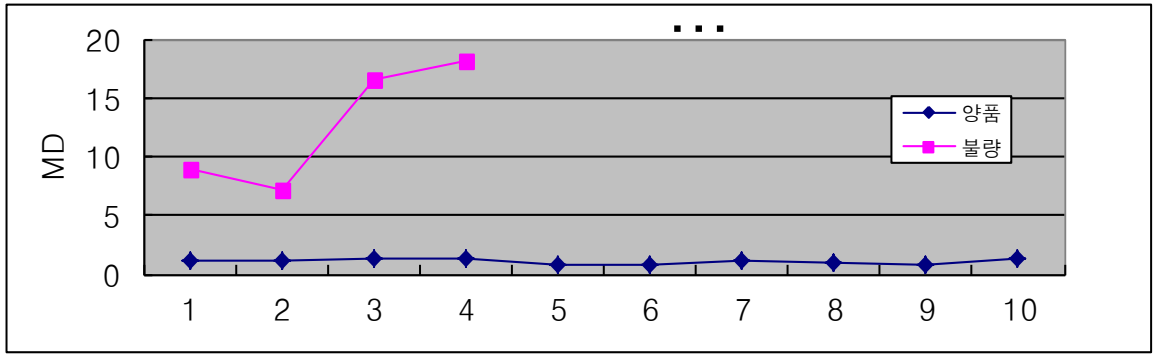
<Mahalanobis-Taguchi System 적용 Process>

적용사례 – 1 : 에어컨用 모터 이음 자동검사시스템 구축

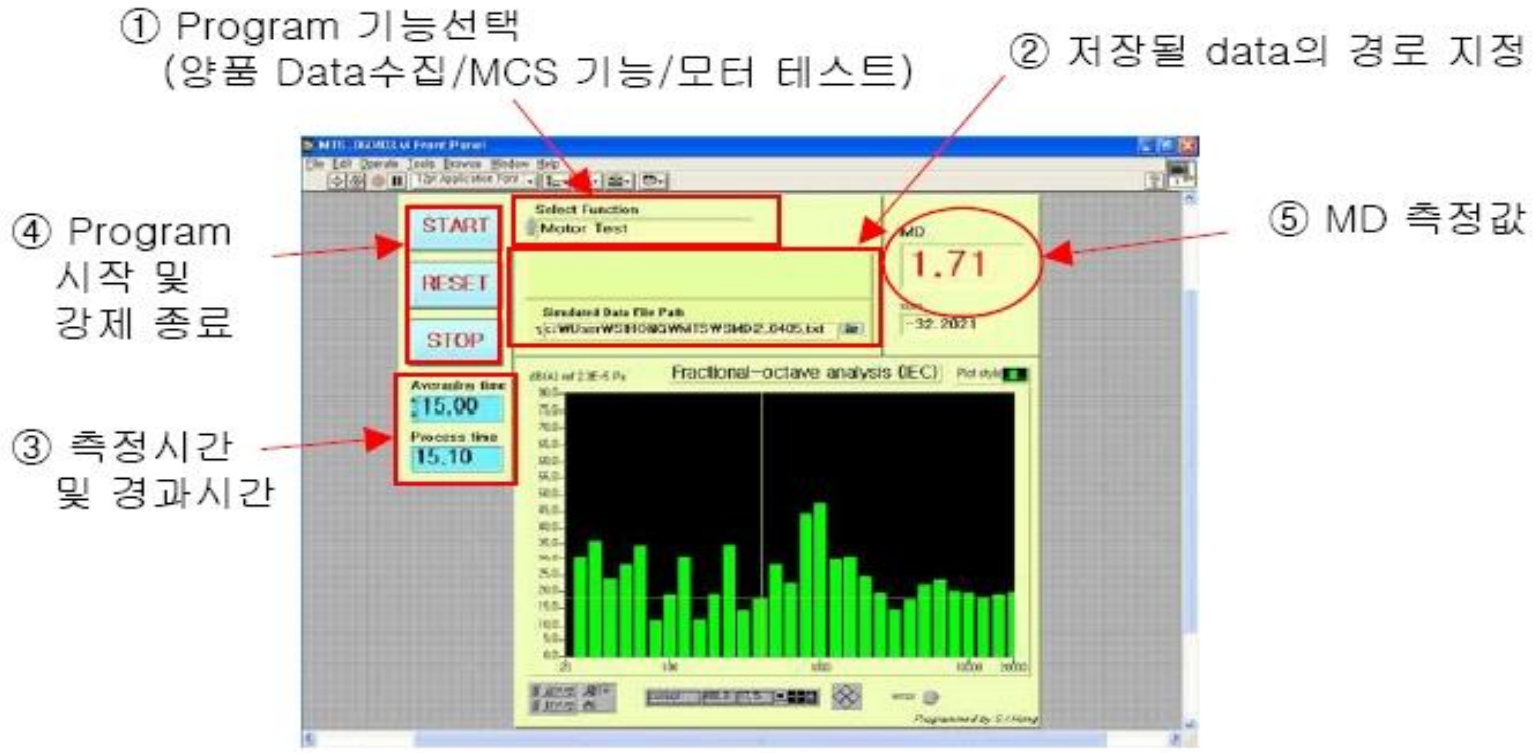
MTS(Mahalanobis-Taguchi System)는 多次元시스템에 관한 의사결정을 위한 **패턴인식(Pattern Recognition)**방법으로 간단하게 저비용/진단시간절감/오판감소를 할 수 있는 방법으로 알려져 있다. 이 MTS는 **병원에서의 환자진단, 음성/지문/홍채인식, 화재감지, 자동차의 Clutch Disc 및 엔진소음 자동검사, 車間距離 근접인식, 신용판단, 소프트웨어 결함판정 Program, 위성추진체 결함검사, Thermal Inkjet Image 품질 검사시스템** 등 미국과 일본 등지에서 수많은 분야에 적용되고 있는 중인 **진단시스템 및 패턴인식 시스템의 최적화를 위한 첨단기법**이다.

금번 MTS를 이용하여 **에어컨 모터이음검사 시스템** 을 구축 시 **모터 이음검사 SPEC.을 정립함과 동시에 이음에 대한 자동화 검사시스템을 구축,** 품질비용을 절감하는데도 대단히 효과적임을 입증하였으며 이를 바탕으로 **모든 모터 類를 비롯한 회전기계에 적용가능 함을 우선 확인하였고, 광범위한 산업분야에 적용가능함을 확신하였다.**

1) MTS 적용결과

<table>
<tr><th>PROCESS</th><th>적용 결과</th></tr>
<tr><td>1.모터샘플 선정</td><td>● 분리형 에어컨(DSB-240)용 실외기 팬모터 양품 10대, 불량품 4대 선정
→ 관능검사 3인 이상 실시로 검증</td></tr>
<tr><td>2. 변수정의</td><td>● 모터단품에 대해 Overall 주파수(20~20,000Hz)의 1/3 옥타브 측정
- Overall 소음값의 검출력에 주파수의 검색의 장점 보강
<table>
<tr><td>변 수 명</td><td>X1</td><td>X2</td><td>…</td><td>X29</td><td>X30</td></tr>
<tr><td>주파수(Hz)</td><td>25</td><td>31.5</td><td>…</td><td>16,000</td><td>20,000</td></tr>
</table></td></tr>
<tr><td>3. 데이터 수집</td><td>● 데이타측정(양품 10대*6회, 불량품 4대) [단위:dB(A)]
<table>
<tr><td>모터시료</td><td>25Hz</td><td>31.5Hz</td><td>…</td><td>16,000Hz</td><td>20,000Hz</td></tr>
<tr><td>양품1</td><td>10.75</td><td>-0.79</td><td>…</td><td>6.30</td><td>-0.57</td></tr>
</table></td></tr>
<tr><td>4. MD(마하라노비스 거리) 계산
$MD = \frac{1}{k} ZA^{-1}Z^{T}$</td><td>
</td></tr>
<tr><td>5. MD규정 및 식별능력확인</td><td>● MD 규정: 2.2 (식스시그마 Level 적용: 양품 MD의 TrMean + 6σ)
● MD 2.2 적용시 모든 시료에 대해 모터 샘플 선정과 일치하는 결과 도출</td></tr>
</table>

2) LabView를 이용한 모터 異音 검출 프로그램 Layout



적용사례 – 2 : 인버터 Line 적용(대외 보안사항으로 미공개)

간단 실습 : Minitab File을 이용한 마하라노비스 거리 계산, Macro File 이용

< 마하라노비스 거리 계산 예제 : 2단계 중 1단계 MD 계산>

Z1	Z2	Z3	Z4	AZ1	AZ2	AZ3	AZ4
-0.67803	0.02741	0.16013	0.32132	1.68061	0.312661	-0.338078	-2.16742
-0.29711	0.15671	0.65830	0.15387	2.25095	0.400517	-0.559395	-0.53846
1.37891	1.39020	1.19207	1.42556	1.56654	0.829457	-0.320285	-0.47059
-1.05894	-1.42077	-1.22765	-1.08161				
0.65517	-0.15360	-0.78285	-0.81913				

Step 1) 정상그룹의 표준화 데이터 행렬(M1) 지정
- ▶ 데이터>복사>열을 행렬로
  - 복사될 열 : Z1-Z4
  - 복사된 데이터 저장 : M1 > 확인

Step 2) 비정상그룹의 표준화 데이터 행렬(M10) 지정
- ▶ 데이터>복사>열을 행렬로
  - 복사될 열 : AZ1-AZ4
  - 복사된 데이터 저장 : M10 > 확인

MTS.MAC 프로그램 내의 변수 명을 설명하면 다음과 같다.

1) K1 = 정상그룹(Normal Group)의 시료수
2) K2 = 측정항목의 개수(측정변수의 개수)
3) M1 = 정상그룹의 표준화 행렬(Standarized Matrix)
4) M10 = 비정상그룹의 표준화 행렬(Standarized Matrix)
5) M4 = 상관행렬(R) : Correlation Matrix
6) M5 = 역행렬(R^{-1}) : Inverse Matrix
7) M40 = 정상그룹의 MD(Mahalanobis Distance)
8) M80 = 비정상그룹의 MD(Mahalanobis Distance)

Step 3) K1(=시료수), K2(=측정 항목수), 입력 및 매크로 파일 실행
- ▶ 편집기>명령 사용
- ▶ MTB>LET K1=5 , ▶ MTB>LET K2=4, ▶ MTB>%MTS >확인 ➔ Macro File 을 실행함.

Step 4) MTS.MAC 파일 실행결과 출력 ➔ 세션 창에 정상 및 비정상그룹의 MD가 실행결과가 출력됨.

Step 5) 적절한 2수준계 다구찌 직교배열표를 생성하고, SN비를 계산하여 불필요한 측정항목을 제거

Step 6) 역치(한계치, Threshold Value)계산, 손실비용 체크 ➔ 최종결론 도출

7. 마무리

★ MTS를 이용하여 에어컨 모터이음검사 시스템을 구축 시 모터 이음검사 SPEC을 정립함과 동시에 이음에 대한 자동화 검사 시스템을 구축, 품질비용을 절감하는데도 대단히 효과적임을 입증하였으며, 이를 바탕으로 모든 모터 類를 비롯한 회전기계에 적용이 가능함을 우선 확인하였고 광범위한 산업분야에 적용가능함을 알 수 있었다. 향후 신 산업분야에도 접목하여 많은 발전을 할 수 있을 것으로 생각된다.

★ 양품과 불량품을 공정에서 효율적으로 정도가 높게 선별하고 걸러내기 위해 현재 운영되고 있는 Line에 직접 적용하여 이론의 타당성을 확인하였고, 재현성을 확인하였다. 타 사례(카메라 모듈에 사용되는 정밀사출 Lens 검사)에도 적용하여 동일 결과를 얻었으므로, 다양한 산업군에 적용하여 설계 및 제조생산성을 높여서 향후 기업의 장기적인 경쟁력을 확보하는데 도움을 주리라 생각한다.

★ 판별분석 방법인 Mahalanobis Distance를 구하는 방법과 다구찌 강건설계의 SN비를 적용하는 Process를 확립하여, 향후 여러 엔지니어들이 담당업무에 쉽게 적용할 수 있도록 S/W화 하는 방안이 필요하다.

● 참고 : MTS 교육일정(2일간) : 2015. 10.22(목) ~ 10.23(금)

• MTS 교육 커리큘럼 (2일간)
/"Macro File"이용 실제 개별기업
사례를 위주로 한 실습과정 운영

일자	교육내용	세부내용	시간
1일차	다구찌기법(강건설계)소개	강건설계의 특장점	09:00~10:00
	직교배열표 1	선점도 및 기본 직교배열표 실습	10:00~11:00
		다구찌 직교배열표 실습	11:00~12:00
	점심식사		12:00~13:00
	직교배열표 2	의인자법, 누적법 소개 및 실습	13:00~14:00
	정특성의 SN비 계산	망소특성의 SN비	14:00~15:00
		망대특성의 SN비	15:00~16:00
		SN비의 최적조건 도출	16:00~17:00
		Minitab에 의한 재현성 확인	17:00~18:00
2일차	MTS 개론	MTS의 이해	09:00~10:00
	MTS 1단계	정상그룹 마하라노비스 공간구성	10:00~11:00
		비정상그룹 마하라노비스공간구성	11:00~12:00
	점심시간		12:00~13:00
	마하라노비스 거리(MD)계산	MD 계산 및 한계치 계산	13:00~14:00
	MTS 2단계	직교배열표 구성	14:00~15:00
		직교배열표에 의한 MD배치	15:00~16:00
	SN비 계산 및 재현성 판정	배치된 MD값에 의한 SN비 계산	16:00~17:00
		재현실험에 의한 재현성 판정 및 결론 / 자사에의 적용 검토	17:00~18:00 Minitab 활용한 마하라노비스-다구찌 시스템