다변량 통계적 분석 방법을 이용한 soft-sensor설계일반

홍 선 주, 한 종 훈

포항공대 화학공학과

1. 서론

실제 산업 공정에서 품질 규격(product specification)을 실시간으로 감시(monitoring), 진단 및 제어하는 것이 어렵거나 불가능할 때가 종종 있다. 예를 들어 고분자의 점성(viscosity)이 품질 변수인 고분자 회분식 공정은 품질 변수의 표본 채취(sampling) 자체가 제한 될 수 있는데, 이것은 한 batch가 완전히 거쳐야 측정할 수 있으므로 표본 채취는 근본적으로 운전 도중에는 할 수 없게 된다. 따라서 회분식 조업 도중에 품질 변수의 감시 및 진단이 이루어 질 수 없고, 결과적으로 더 양질의 품질 규격을 생산하기 위한 연속 제어가 불가능하게 된다. 또한, Gas Chromatography(GC)와 같은 실시간 분석기기는 가격이 비싸며, 채취 및 분석하는데 오래 걸려서 신뢰성이 낮고, 잦은 고장으로 인한 높은 보수비용 등 때문에 아직까지 품질 변수 감시 및 제어에 직접 사용되는 경우가 별로 없는 실정이다.

위와 같은 이유들은 품질 변수의 감시 및 제어를 적절한 시기에 할 수 없게 되면서 원하는 조업 조건에서 벗어나 회복되는데 시간이 오래 걸리는 외란(disturbance)의 영향을 받기 쉬운 공정의 운전을 초래하게 되는데, 이것은 결과적으로 설정 값(setpoint)에서 많이 벗어난 불안정한 공정의 운전 및 과도한 에너지 낭비를 가져온다. 이것은 때때로 고급 공정 제어(advanced process control)를 이용하여도 표본 채취가 어렵거나 드물게 채취되어지는 품질 변수의 경우는 만족할만한 규격 제어가 불가능하다.

이와 같이 자주 채취할 수 없거나 채취자진단 및 제어를 위해서, 쉽고 자주 측정되며 품질 변수의 거동과 비슷한 다른 공정의 운전 변수들로부터 간접적으로 품질 변수의 규격을 추정하는 추정기(estimator)에 대한 연구가 일찍부터 있어왔는데, 이것은 품질 규격을 실시간으로 직접 측정하는 하드 센서(hard sensor)와 구분하여 소프트 센서(soft sensor)로 불리 운다.

소프트 센서는 다양한 산업 공정에서 사용되고 있는데, 발효 중 biomass(P.A.Lant(1991)), 증류 공정의 탑상/탑저 생산물 조성(Thor Mejdell and Sigurd Skogestad(1991), 신준호(1995)), 산업용 고분자 공정에서의 melt flow index(M.T.Tham(1991)), lubricant production 공정에서 viscosity index (M.J.Willis(1997)), olefins pyrolysis 공정에서 tube-wall temperature(Manabu Kano(1996)) 등이 대표적으로 소프트 센서를 이용하여 감시, 진단 또는 추론 제어를 적용하고 있는 공정의 예다.

소프트 센서 설계는 일단 품질 변수를 가장 근사하게 추정할 수 있는 측정 가능한 공정 변수를 선정하는 것부터 시작하여 대상 공정에 적합한 모델을 찾아 그들 간의 관계를 찾아내는 것을 말하는데, 이 때 품질 변수와 선정된 공정 변수간의 관계가 수학적으로 단순하게 표현되는 경우 수학적인 이론적 모델을 사용한 정교한 소프트 센서의 설계가 가능하게 되지만, 일반적으로 이론적 모델은 실제 공정에서 구현하기 힘들기 때문에 실험에서 직접 얻은 입출력 데이터를 이용한 데이터 기반 소프트 센서가 소프트 센서의 기본이 되고 있다.

데이터 기반 소프트 센서의 기술로는 크게 통계적인 방법과 지식 기반(knowledge based) 방법이 있다.

통계적인 방법은 간단한 통계 기술을 이용하여 수집된 데이터로써 모델링을 하고 모델링이 끝난 후 공정에 대한 분석이 가능한 방법이며 지식 기반 방법에는 대상 공정에 대한 전문가의 지식을 바탕으로 하는 전문가 시스템 구축 방법과 비선형성이 강하고 복잡한 공정에서 수집된 데이터의 관계를 블랙박스 모델링 하는 신경 회로망을 이용한 방법이 있다. 이 때 전문가 시스템 구축 방법은 전문가의 주관적인 공정 지식으로써 모델이 설계되어지며, 신경 회로망은 공정에 대한 해석이 어려운 블랙박스 모델링 방법이기 때문에 실제로 소프트 센서의 기술로써 통계적인 방법에 대한 연구가 활발히 진행되고 있는 실정이다.

1980년대 후반부터 공정의 감시, 진단 및 제어의 방법으로 사용될 뿐만 아니라 소프트 센서의 기술로 꾸준히 연구되고 있는 다변량 통계적 분석 방법은 선형 회귀 분석 기술(linear regressiontechnique)을 이용하여 데이터간의 관계를 도출해내는 방법으로써 다음과 같은 식으로 간단하게 표현되어 질 수 있다.

$$\hat{y} = k\theta \quad (1)$$

이 때 y는 품질 변수이며 θ는 품질 변수를 추정하는 측정 변수를 가리키고, k는 구하고자 하는 행렬이다.

선형 회귀 분석 기술에서 가장 대표적인 방법으로서 MLR(Multiple Linear Regression)은 입출력 데이터들을 위의 식대로 최소 자승법(Least Squares)을 이용하여 해를 구하는 방법인데, 이 방법은 측정 변수들이 많고 서로 강하게 연관(correaltion)되어 있는 경우 모델의 over-parameterization현상이 일어날 수 있다. 또한 이 경우 측정 변수의 covariance matrix($\theta^T\theta$)의 inverse가 존재하지 않거나 singular가 되어 추정 모델이 ill-conditioned 되어performance가 현저하게 떨어질 수 있다.이 방법은 서로 연관성이 없는 독립적인(independent) 측정 변수를 추정 변수로써 선정하는 것이 하나의 해결 방법이 될 수 있으나, 일반적으로 공정 변수들끼리는 강한 상관관계를 가지고 있기 때문에 이러한 상관성을 효과적으로 다룰 수 있는 PCR, PLS 방법을 이용하면 모델의 추정 능력을 높일 뿐만 아니라 강건한 모델을 구성할 수 있다.

본 고에서는 PCR, PLS등의 다변량 통계적 분석 방법을 이용한 소프트 센서 설계 일반에 대해 기술하였는데, 제 2절에서는 변수 선정, 데이터 전처리 기술을 포함하는 소프트 센서 설계 일반에 대하여 다루었고, 3절에서는 PCR, PLS등을 비롯한 다변량 통계적 분석 방법에 대한 이론을 설명하였다.

2. 소프트 센서 설계 일반

소프트 센서 설계는 다음과 같이 크게 세 가지 과정을 거치게 된다.

① 공정 변수 선정
② 데이터 전처리
- 데이터 조절
- 특이한 데이터 제거

③ 모델링 및 검증

2-1 변수 선정(variable selection)

어떤 대상 공정의 품질 변수의 추정을 위한 소프트 센서 설계에서 가장 선행되어야 할 작업은 품질 변수를 가장 근사하게 추정할 수 있는 공정 변수들을 선정하는 것이다. 이것은 추정하는데 별로 영향을 주지 않는 변수들을 포함시키지 않음으로써 추론 모델(inferential model)의 예측 능력을 향상시킬 수 있을 뿐만 아니라 실제로 산업체에서 실행할 때 보수 및 유지도 쉽고 간편하게 해준다.

일반적으로 품질 변수를 추정하는 공정 변수로써 많이 사용되고 있는 변수는 압력, 온도, 유량 등과 같이 쉽고 간편하게 측정되며 측정 지연이 거의 없는 변수들이다.

통계 기반 지식으로 공정 변수를 선정하는 방법에는 Step wise regression procedures, Interactive variable selection procedure, Instructive variable selection procedure 등이 있는데, 측정 가능한 모든 공정 변수들을 하나씩 품질 변수와의 회귀(regression)에서 R^2 등을 평가하여, 추정에 적게 기여하는 변수는 데이터 집합에서 제거하는 Step wise regression procedures는 공정 변수의 수가 많아질 경우 조합(combination) 문제가 생길 수 있다.

Interactive variable selection procedure는 PLS를 이용하여 순차적으로 PC를 추출하는 매 단계마다 각 변수의 회귀 계수(regression coefficient)로써 추정 영향을 끼치지 않는 변수를 제거하면서 최종적으로 모델의 성능을 최대로 하는 변수를 선정하는 방법이다. 이 방법은 PC의 개수가 많은 복잡한 공정의 경우 초기 PC추출 단계에서는 별로 중요하지 않다가 나중 단계에서 품질 변수를 추정하는데 중요한 영향을 끼치는 변수가 있을 수 있다.

Instructive variable selection procedure는 일단 측정 가능한 모든 공정 변수와 품질 변수가 PLS방법을 사용하여 모델링을 한 후, 데이터의 상관 분포를 가장 많이 나타내고 있는 첫번째 두 번째 loading 값 중 0값 근방에 모여있는 변수들은 제외시키고, step wise regression procedures 등과 같은 다른 통계적인 방법을 함께 이용하여 최종적으로 품질 변수를 추정할 공정 변수를 선정하는 방법이다. 이 방법은 증류 탑의 탑정 생산물 조성을 추정하기 위해서 측정이 가능한 모든 공정 변수 중에서 탑정 조성에 영향을 주는 중요한 변수들을 선정하는 기준으로 사용하였다.

일반적으로 공정 변수는 서로 강하게 상관되어 있기 때문에 단변량(univariate) 통계 기법을 이용하여 선정하는 것보다는 PCA,PLS등과 같은 다변량 통계적 기법을 이용하는 것이 더 알맞다.

또한 변수의 상관성을 고려하여 변수 선정이 잘 되었더라도 품질 변수의 작은 변화에 민감하게 반응하는 민감도(sensitivity)와 외란의 영향이나 측정 오차 및 센서의 고장에도 강건한 강건성 (robustness)을 동시에 만족하는 것이 힘들기 때문에 대상 공정의 소프트 센서의 목적에 맞으면서 구현하기 쉽도록 적절하게 변수의 종류와 개수를 결정하는 것이 바람직하다.

2-2 데이터 전처리(Data preprocessing)

데이터 전처리 단계에서는 정상 조업 범위 차수

(order)가 다른 온도, 압력, 유량 등과 같은 여러 변수들의 크기 조절(scaling) 과정과 공정의 정상 조업에 해당되는 데이터의 범위를 벗어나는 특이한 데이터 제거 과정이 포함된다.

통계적인 방법으로써 소프트 센서를 설계할 때 데이터 조절은 중요한 과정이 된다. 이것은 적절하게 변수의 크기를 조절하지 않게 되면 통계적인 모델은 큰 차수를 갖는 변수에게 bias되는 경향이 있으며, 공정에 대한 정보가 있을 경우 선정된 변수 중에서도 품질 변수에 가장 많은 영향을 미치는 변수에게 가중치(weight)를 줌으로써 모델의 성능을 높일 수 있기 때문이다. 그러나, 일반적으로 통계적인 방법을 사용하여 소프트 센서를 설계할 경우 공정에 대한 사전 정보가 부족하기 때문에 가장 널리 사용되고 있는 데이터 조절 방법으로는 자료 행렬의 모든 데이터 값에서 일정한 평균 값을 빼고 그 값을 표준 편차로 나누는 자동조절(autoscaling)이다. 이 조절을 통해 전체 변수는 0~1의 범위의 값의 크기가 된다.

데이터 전처리 다음 단계는 수집된 데이터 중에서 특이한 데이터를 제거하는 것이다. 특이한 데이터는 측정기기의 부정확성으로 인한 측정 오차, 공정 잡음 및 장치 고장 등으로 인한 잘못된 정보인 경우가 많지만, 때로는 공정의 특별한 특성을 나타내기도 한다. 이 두 경우를 구별할 수 있는 방법은 대상공정을 가장 잘 이해하는 전문가가 판별하는 수 밖에는 없다.

특이한 데이터 제거 방법은 크게 Clustering Analysis와 Chemometrics 방법이 있다.

Clustering 은 전체 데이터 집합을 비슷한 특성(Characteristics)을 갖는 여러 개의 부분집합으로 나눔으로써 공정에 대한 정보를 좀더 명확하게 얻을 수 있는 방법인데, 어떤 목적으로 사용하느냐에 따라 조금씩 의미를 달리하기 때문에 cluster를 의미하는 용어도 type, group, class등 다양하다.

공정 데이터가 정상 조업을 벗어나거나, 센서의 고장 등으로 인하여 특이한 데이터를 포함하고 있을 경우 clustering analysis를 하면 정상 조업 데이터를 가진 큰 그룹과 매우 적은 데이터만을 포함하는 작은 그룹으로 나뉘어지는데 이때 적은 데이터를 가진 그룹을 잘못된 데이터로 분류한다. clustering analysis는 수집된 데이터가 공정의 가능한 모든 거동(behavior)을 포함하고 있으며 그 양이 방대할 때 신뢰성이 있으며, 데이터의 수가 적고 가능한 조업 영역을 모두 포함하고 있지 않은 데이터의 경우는 이 방법에 의하여 특이한 데이터를 가려내는 것은 신뢰도가 다소 떨어진다.

Chemometrics방법은 공정 변수의 상관성을 효과적으로 해결해주는 Principal Component Analysis (PCA), PLS등과 같은 다변량 데이터 분석(multivariate data analysis)방법을 이용하여 특이한 데이터를 가려내는 방법이다.

그림1은 데이터의 상관 분포가 가장 큰 첫번째 주성분과 다음으로 큰 두 번째 주성분으로 관측치가 정사영(t_1,t_2)되었을 때 왼쪽 상단 56번째 관측치가 정상 조업에서 벗어나고 있는 것을 볼 수 있다.

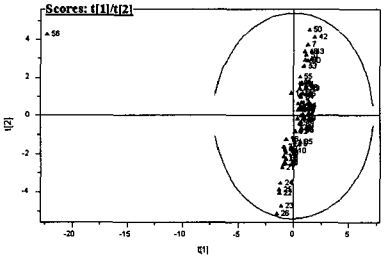

그림 1. PCA를 이용한 특이한 데이터 감지

실험을 통해 수집된 데이터는 센서를 통해 입력된 출력 값이므로 필연적으로 오차를 내포하게 되는데 이러한 오차를 적절하게 변환하는 데이터 변환 과정을 거침으로써 데이터 전처리 단계가 끝난다. 종래에 오차 변환은 오차모델을 기반으로 각각의 공정 변수에 대해서 독립적으로 변환과정을 거치는 single-scale 방법이 널리 사용되었으나, 최근에는 multiway-PCA를 이용하여 공정 변수들간의 상관성을 고려한 wavelet transform에 관한 연구가 활발히 진행중이다.

3. 다변량 통계적 모델링 기법(multivariate statistical modeling method)

3-1 PCR

서로 연관되어 있는 공정 변수와 품질 변수가 직접 관계를 맺기 이전에 공정 변수의 연관성을 해결해주는 데이터 전처리의 의미로써 공정 변수들이 PCA를 통해 의미 있는 k개의 독립적인 잠재 변수(latent variables)로 나타낸 후, 품질 변수와 관계를 맺는 방법이다(그림 2).

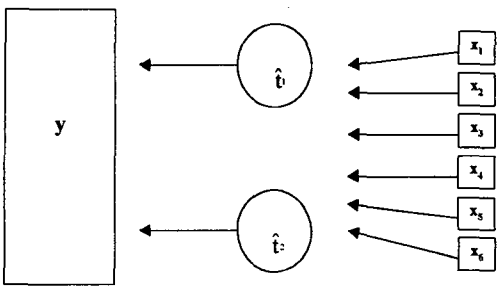

그림 2. PCR의 원리

식으로 나타내면,

$$\theta = t_1 p^T_1 + t_2 p^T_2 + \dots + t_m p^T_m \quad (2)$$

(θ : 공정 변수)

$$K_{PCR} = Y^T[\theta_k^T]^+ = Y^T T[T^T T]^{-1} P^T \quad (3)$$

(Y : 품질 변수)

위와 같고, 이 때 주성분(Principal Component)의 개수 k개는 Nonlinear Iterative Partial Least Squares (NIPALS) 알고리즘을 이용하여 F-test로써 결정하거나 공정 변수의 Singular Value Decomposition (SVD)를 통하여 condition number를 최소로 하는 k번째까지를 선택한다.

PCR은 공정 변수의 개수가 많고 서로 연관되어 있는 경우, 의미 있는 PC의 개수만큼 차수를 감소시킴으로써 모델의 over-parameterization을 극복하였지만, 품질 변수의 개수가 여러 개인 경우 데이터의 상관성은 공정 변수 및 품질 변수와의 상호적인 상관 관계로부터 비롯될 것이고, 이것은 품질 변수와는 독립적으로 공정 변수들간의 상관 관계로부터 결정된 주성분을 이용한 PCR의 정확성을 떨어뜨리는 결과를 가져 온다.

3-2 PLS

PLS는 PCR(Principal Component Regression)[5] 방법이 변형 개선된 것으로서 PCR이 공정 변수간의 연관성만 고려한 것에 반해 PLS는 공정 변수 행렬의 연관성의 방향(direction)이 품질 변수와 가장 큰 covariance를 갖도록 재배열하는 방법이다. 그림 3은 PLS의 원리를 나타내고 있으며 식으로 나타낼 경우에는,

$$\mathbf{X} = \mathbf{T}\mathbf{P}^T + \mathbf{E} \quad (4)$$

$$\mathbf{Y} = \mathbf{U}\mathbf{C}^T + \mathbf{F} \quad (5)$$

$$\mathbf{U} = \mathbf{B}\mathbf{T} + \mathbf{R} \quad (\mathbf{B} = (\mathbf{T}^T\mathbf{T})^{-1}\mathbf{T}^T\mathbf{U}) \quad (6)$$

이다. **X**는 공정 변수 행렬을 가리키며, **Y**는 추정하고자 하는 품질 변수에 속한다.

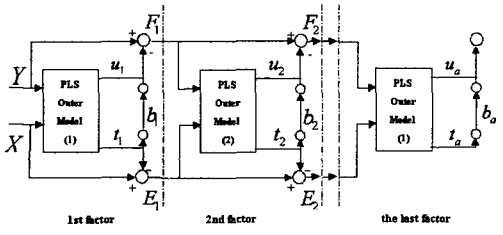

그림 3. PLS의 원리

PLS 알고리즘은 크게 외적변환(outer relation)과 내적모델(inner model) 단계를 거치게 되는데, 외적 변환은 Nonlinear Iterative Partial Least Squares (NIPALS) 알고리즘을 이용하여 다변수 입출력 데이터가 각각 가장 넓게 분포하고 있는 방향부터 순차적으로 새로운 벡터(**P**,**C** or principal component)를 정의하고 그 벡터에 정사영 된 값, score 벡터(**T**, **U**)를 만드는 과정이다.

외적 변환을 거쳐 나온 입출력 score 벡터는 최소자승법(least square)을 이용하여 최종적으로 입출력 데이터의 covariance 구조를 가장 잘 나타낼 수 있도록 재배열되는데 이 과정이 내적 모델링 과정이며, 입출력 데이터의 상관성을 동시에 나타내는 주성분이 최종적으로 결정된다.

식 4), 5)은 각각 X, Y블록에 대한 외적변환을 나타내며, 식 6)는 입출력 score t와 u간의 관계식으로써 내적 모델(inner relation)을 의미한다. 그리고 여기서 E, F, R은 잔차 행렬(residual matrix)들이고 E와F가 최소가 되어 거의 공정에 대한 정보가 없다고 판단되어 질 때까지(위 그림에서 last factor인 a가 계산되어 질 때까지), 위와 같은 과정이 반복 계산되어 진다. 이 때 몇 개의 주성분을 사용하여 모델링을 할 것인가는 cross validation을 이용하여 결정한다.

PLS는 공정 변수들간의 redundancy 를 해결하고 측정 잡음(measurement noise)을 없애 주는 강건한 선형 모델로써 근래 들어 많은 이론적 연구와 함께 실제 공정에 적용이 중요한 현안이 되고 있다. 그러나, 변수들간의 auto-correlation이 없는 공정, stationarity를 만족하는 공정의 운전으로부터 얻어진 데이터에 대해서만 만족할만한 해를 줄 수 있다는 것이 PLS의 한계로 지적되어 오고 있다.

3-3 비선형 PLS

3-3-1. Artificial neural network PLS (NNPLS)

신경회로망의 뛰어난 비선형 관계 근사와 변수의 상관성을 효과적으로 다루는 PLS를 복합한 비선형 강건 소프트 센서 설계에 대한 연구가 활발히 진행중이다.

이것은 크게 두 가지 접근 방식이 있는데 하나는 McAvoy group에 의해서 연구되어 지고 있는선형 강건 소프트 센서 설계에 대한 연구가 활발히 진행중이다. 것으로서, (4)식과 (5)식의 PLS 외적 변환은 그대로 이용하면서 (6)식, 즉 내적 모델을 다음과 같이 신경 회로망을 이용해 비선형 근사를 하는 것이다.

$$u_h = N(t_h) + r_h \quad (7)$$

$N(\cdot)$은 신경회로망에 의해 표현된 비선형 관계를 나타내며, 이때 사용되는 신경 회로망의 종류는 multilayer feedforward network, radial basis functions, 또는 recurrent networks이 될 수 있다.

NNPLS의 구조는 다음과 같다(그림 4).

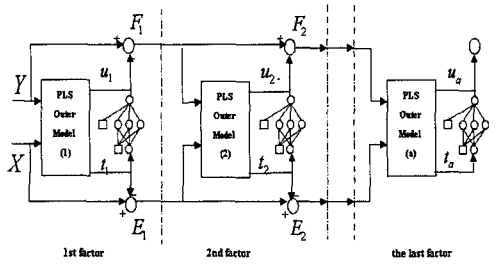

그림 4. NNPLS의 원리

이 방법은 공정 변수간의 상관성이 PLS 외적 변환을 통해 차원이 감소된 주성분으로써 정의된 후 주성분에 정사영된 값, score벡터의 관계를 신경 회로망에 의해 모델링하고 있으므로, 기존의 신경 회로망에 의해 직접적으로 데이터가 학습되는 것과는 차이가 있다. 직접적인 network학습일 경우 MIMO (multiple input multiple output) network이므로 무수히 많은 뉴런의 가중치를 결정해야 되는 반면 NNPLS의 network은 SISO이므로 가중치 수가 현저하게 줄어듦으로써 over-parameterization을 피할 수 있을 것이며 또한 국부 최소값의 수도 network의 크기가 작으므로 더 줄어들 것으로 기대할 수 있다.

3-3-2. Autoassociative neural network PLS

비선형 PLS의 다른 접근 방법으로써 E. Malthouse (1997)가 Autoassociative neural network을 기본으로 하는 비선형PLS 방법을 제안하였다. 이 방법은 Kramer의Autoassociative neural network이 공정 변수의 전처리기로 사용되어, 공정 변수의 차원 감소 및 센서의 큰 측정 잡음, 바이어스와 고장 등에도 품질 변수의 추정성능과 모델의 강건성을 잃지 않도록 개발된 방법(그림 5)인데, 품질 변수가 여러 개일 경우는 품질 변수들도 autoassociative neural network으로써 전처리를 거친 후 모델링을 하게 된다.

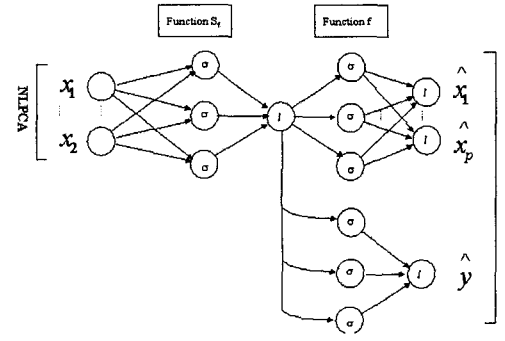

그림 5. Autoassociative neural network PLS(MISO)

이 방법은 앞에서 언급한 비선형PLS가 기존의 PLS의 선형 외적 변환은 그대로 고수하고 내적 모델링 단계에서 비선형 회귀를 사용한 반면, Autoassociative neural network PLS 는 측정되는 공정 변수들 및 품질 변수들끼리의 비선형 연관성을 autoassociative neural network으로서 다루고자 하는 것이 특기할 만할 사실이다.

이와 같은 Kramer의 5층 autoassociative neural network은 비선형 PCA에 적합하다고 보고되었는데, 위 그림에서 병목층(bottleneck layer)을 통과한 출력은 비선형 PC를 의미하며, 다음과 같이 목적함수를 최소화하는 곡선 또는 초평면(hyperplane) f와 sf를 추정하는 주성분 또는 주곡선(principal curve)을 사용하고 있다.

$$J = \sum_{i=1}^{N} \left\| X_i - f(s_f(X_i)) \right\|^2 \quad (N: \text{관측수}, X_i: \text{데이터 벡터}) \tag{8}$$

이처럼 평면을 추정하는 비선형 PC는 종전의 방법들보다 적은 PC를 가지고 보다 더 좋은 성능을 보일 수 있다. 하지만, 신경회로망을 기본으로 한 비선형 PCA는 모두 5개의 층을 가지고 있어서 이에 대한 학습이 어려울 뿐더러, mapping층의 개수가 모호하며, 기존의 이론적 탄탄한 구조에 바탕을 둔 방법들에 비해 병목층의 의미를 규정하기가 힘드는 등의 몇 가지 문제점을 보이고 있다.

4. 결론

본 고는 데이터 기반 소프트 센서 설계에서 품질 변수와 구체적인 연관성을 갖는 변수의 선정과 양질의 데이터를 얻기 위한 데이터 전처리 단계로써 데이터 조절, 특이한 데이터 제거와 데이터 변환에 관해 개괄적으로 설명하였고, 적절한 데이터 전처리 단계를 거쳐 공정 변수와 품질 변수의 구체적인 함수 관계를 맺는 모델링 방법으로써 최근에 공정 감시, 진단 및 제어에 뿐만 아니라 소프트 센서의 기술로써 활발하게 연구되고 있는 대표적인 다변량 통계적 분석 방법인 PCR, PLS, 비선형 PLS에 대해서 기술하였다.

일반적으로 다변량 통계적 모델의 goodness of fit은 R2 등을 이용하여 검증하고 모델의 예측 능력은 EPV(Explained Prediction Variances), FVU (Fraction of Variance Unexplained) 등을 통해 가늠되어지며, 여러 모델과 성능(performance)을 비교하고자 할 때는 PRESS(Prediction Error of Sum of Squares), MSEP(Mean Squares Error of Prediction) 등을 이용한다.

소프트 센서는 공정의 운전 도중에 가끔씩 측정되거나 또는 측정이 불가능한 품질 변수를 측정하기 쉽고 자주 측정되는 공정의 운전 변수들을 이용

하여 운전 도중에 추정함으로써 안정된 공정 운전 조건을 기대할 수 있으며, 또한 품질 규격(product specification) 미달 방지를 위해 실제보다 높게 운전 목표치를 설정하여 운전하고 있는 여러 산업 공정의 조업 비용을 절감할 수 있을 것이다.

최근에는 산업 공정의 환경규제가 강화되면서 부산물로 생 산될 수 있는 SOx, NOx등의 유해한 가스를 최대한 줄이고자, 공정의 운전 중 이러한 부산물을 신속하고 정확하게 추정할 수 있는 소프트 센서의 설계에 대한 연구가 활발히 진행중인데, 이처럼 소프트 센서는 후처리 공정의 추가에 따른 투자 비용 상승을 방지하는 목적으로도 사용될 수 있다.

감 사

본 연구를 위해 포항공대 공정산업의 지능자동화 연구센터를 통해 재정적 지원을 해주신 한국과학재단에 감사를 드립니다.

참고문헌

[1] Ming T. Tham, Gary A. Montague, A. Julian Morris, and Paul A. Lant, "Soft-sensors for process estimation and inferential control", J. *Proc.Cont,* vol 1, pp. 3-14, 1991.

[2] J. V. Kresta, T. E. Marlin and J. F. MacGregor, "Development of inferential process models using PLS", *Computers chem. Engng.*, vol. 18, no. 7, pp. 597-611, 1994.

[3] Mark A. Kramer, "Nonlinear principal component analysis using autoassociative neural networks", *AIChE,* vol. 37, no. pp. 233-243, 1991.

[4] S. J. Qin and T. J. McAvoy, "Nonlinear PLS modeling using neural networks", *Computers chem. Engng.,* vol. 16, no. 4, pp. 379-391, 1992.

[5] Lindgren, F. Geladi P. Rannar S. and S Wold, "Interactive variable selection for pls(part 1)", *J. of Chemometrics*, no. 8, pp. 349-363.

[6] J. Neter. W. Wasserman, and Kutner, M. H. : "Applied linear statistical models", *3rd ed., Richard,* D. IRWIN, INC.(1990)

[7] E. C. Malth-ouse, A. C. Tamhane and R. S. H. Mah, "Nonlinear partial least squares", *Computers chem. Engng.,* no. 8, pp. 875-890.

저자소개

홍 선 주

1973년 3월 17일생
1997년 한양대학교 화학공학과 졸업(공학사)
1998년 포항공대 화학공학과 석사과정
<관심분야>
소프트 센서, 품질 변수 모니터링 및 감시

한 종 훈

1961년 9월 1일생
1984년 서울대학교 화학공학과 졸업(공학사)
1986년 서울대학교 대학원 화학공학과 졸업(공학석사)
1993년 M.I.T.졸업(공학박사)
1995년-현재 포항공대 화학공학과 조교수
<관심분야>
- On-line process monitoring, diagonosis and optimization
- Plantwide optimization and control
- Application of real-time information technology to process systems engineering
- Intelligent systems