

비선형회귀분석을 위한 통계소프트웨어 NLIN2000

강근석¹⁾ 심규호²⁾

요약

Window 환경 하에서 사용이 간편하면서도 다양한 통계량을 제공하는 비선형회귀분석을 위한 통계소프트웨어 NLIN2000을 소개한다. 기존의 DOS용 프로그램을 업그레이드한 것으로 다른 통계팩키지들에 비하여 모형식의 설정 및 적합과정이 간편하고, 모형식 저장 및 삭제, 모형식 형태 보기 등의 기능을 제공한다. NLIN2000은 비선형회귀분석에 대한 통계적 이론을 연구하는 통계전공자들에게 필수적인 각종 통계량을 제공해줄 뿐만 아니라, 실제 현장에서 비선형모형을 사용하여 분석하는 다른 학문분야의 연구자들에게도 유용하게 사용될 수 있다.

주요용어: 비선형회귀분석, 비선형성, 통계팩키지, 곡률분석.

1. 서론

국내에서 널리 사용되고 있는 통계팩키지 SAS와 SPSS에는 비선형회귀분석을 위한 모듈이 포함되어 있다. 그러나, 그 기능이 한정되어 있어 저자들은 1990년에 다양한 기능을 제공하는 비선형회귀분석을 위한 소프트웨어 NLIN90을 개발하여 소개하였다(강근석, 1993). 하지만 이 NLIN90은 CUI(character user interface)에 기반을 둔 것으로 현재 Window를 사용하는 컴퓨터 환경 하에서는 여러 가지 사용상의 한계를 가지고 있었다. 예를 들어, 새로운 비선형 모형을 사용하고자 하는 경우에는 사용자가 직접 컴퓨터언어(Fortran)로 그 모형식과 모형식의 모수에 대한 도함수 계산을 위한 부함수(subroutine)를 작성하여 컴파일과 링크를 매번 다시 수행하여 실행파일을 작성하여야 하였다.

비선형회귀분석이 공학, 동·식물학, 약학 등의 통계 비전공자들에 의해 더 많이 사용된다는 사실은 이들이 더욱 간편하게 사용할 수 있는 소프트웨어의 개발 당위성을 더해준다. 그렇지만 기존의 통계팩키지들이 전체적인 환경을 GUI(graphic user interface) 기반으로 변환을 하였으면서도 아직 고급통계분석 부분들은 그 전의 CUI 기반에서 개발되었던 틀을 벗어나지를 못하고 있다. 더군다나 비선형회귀분석에서는 그래프의 사용이 필수적이므로 GUI 환경으로의 변환이 더욱 요구된다.

본 논문에서 소개되는 비선형회귀분석을 위한 소프트웨어 NLIN2000(a nonlinear regression analysis system for 2000s)은 NLIN90을 GUI기반으로 변환하고 새로운 기능을 첨가하여 업그레이드한 것이다. 기존의 소프트웨어에 비하여 모형의 설정 및 적합과정이 간편하

1) (156-743) 서울시 동작구 상도 5동 1-1, 숭실대학교 자연과학대학 정보통계학과, 교수

E-mail: gskang@stat.ssu.ac.kr

2) (156-020) 서울시 동작구 대방동 23-29, 에스링크(주), 전임연구원

E-mail: khshim@s-link.com

고, 모형식들의 입력, 저장 및 삭제가 가능한 모형식의 데이터베이스, 모수값의 변화에 따른 모형식의 형태 보기 등이 새로운 기능으로 추가되었다. 그리고, 모형식의 모수에 대한 도함수를 자동 계산하여 주어 사용자의 편리성을 크게 제고하였다. NLIN2000은 비선형회귀분석에 대한 통계적 이론을 연구하는 통계전공자들에게 필수적인 모수추정값, 표준오차, 상관계수, 비선형성 측도, 잔차, 영향력 측도 등의 각종 통계량을 제공해줄 뿐만 아니라, 실제 현장에서 비선형 회귀분석을 사용하는 다른 학문분야의 연구자들에게도 쉽고 유용하게 사용될 수 있는 기능들을 제공하여 준다.

2. NLIN2000의 소개

비선형회귀분석을 위한 소프트웨어 NLIN2000은 현재 독립적인 프로그램으로 운영되고 있으며 추후 기능을 보완하는 경우에도 사용자들에게 보다 간편하면서도 많은 정보를 제공하는 전문화된 모듈로 개발될 계획이다. NLIN2000은 Microsoft사의 Visual Studio(Ver. 6.0)를 이용하여 C++ 언어로 작성되었으며 현재 약 500KB의 용량을 가지고 있다. NLIN2000을 설치하여 실행하면 그림 2.1과 같은 초기 화면이 나오게 된다. NLIN2000에서는 비선형 회귀분석을 수행하기 위한 여러 과정을 단계적으로 '자료입력', '산점도', '모형식 설정과 보기', '모형적합', '잔차분석' 등의 메뉴로 나누어 제공하여 준다.

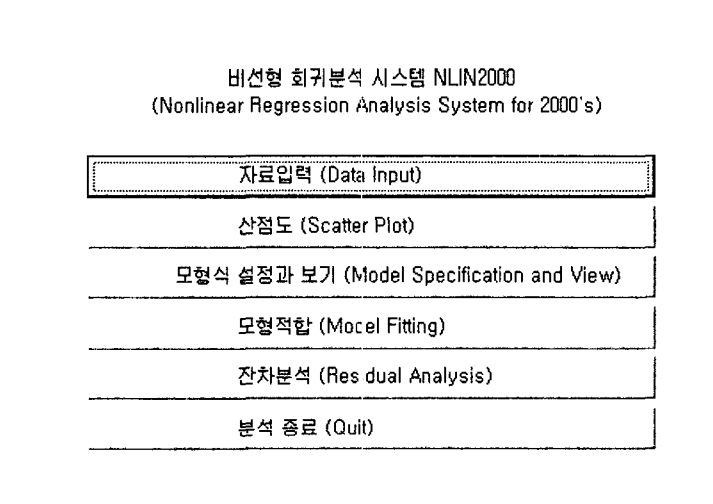


그림 2.1: NLIN2000의 초기화면

2.1. 자료입력

먼저 '자료입력'은 분석자료를 입력, 저장, 또는 불러오는 기능을 가지고 있다. 프로그램의 크기를 줄이기 위하여 필수적인 기능만을 갖춘 자료입력 시트를 제공하여 사용자가 편

	X	Y
1	0.02	76
2	0.02	47
3	0.06	97
4	0.06	107
5	0.11	123
6	0.11	139
7	0.22	159
8	0.22	152
9	0.56	181
10	0.56	201
11	1.10	207
12	1.10	200
13		
14		
15		
16		

그림 2.2: 자료입력 창

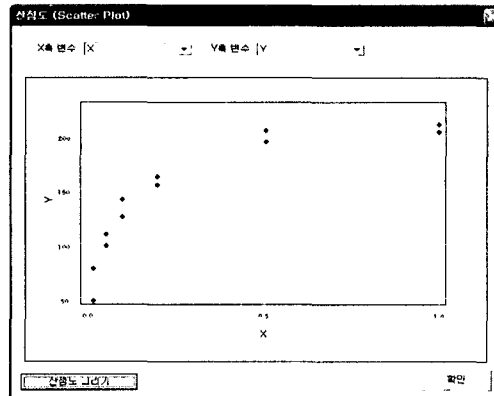


그림 2.3: 산점도

하게 자료를 입력할 수 있게 하였다(그림 2.2).

2.2. 산점도

‘산점도’는 회귀분석의 사전 분석단계로 필수적인 입력된 자료의 산점도를 보여주는 기능으로 이 산점도에 의거하여 타당한 모형식을 선정할 수 있게 하여준다(그림 2.3). 선형 회귀분석과 달리 비선형회귀분석에서는 자료분석에 사용할 모형식의 설정이 매우 중요하다. 일반적으로 비선형회귀모형을 사용하는 대부분의 문제들에서는 모형식의 형태가 과학적 이론이나 분석에 근거하여 알려져 있다. 예를 들어, 화학반응의 경우에는 반응과정이 미분방정식으로 표현되어 여기에서 필요한 식의 형태를 얻을 수 있다.

2.3. 모형식 설정과 보기

그림 2.4의 ‘모형식 설정과 보기’에서는 우선 많이 사용되는 비선형회귀모형식들을 데이터베이스 형태로 제공하는 한편 사용자가 필요한 모형식을 직접 설정하고 데이터베이스에 추가로 저장 및 삭제를 할 수 있는 기능을 제공한다. 모형 DB에는 각 모형에 대하여 일반적으로 사용되는 수학 연산자를 이용하여 표현되는 모형식, 모형식 이름 및 추가 설명, 설명변수 개수 및 이름, 모수 개수 및 이름 등의 정보를 가지고 있다. 예를 들어, 동·식물의 성장분석에 사용되는 Michaelis-Menten 모형의 경우에는 모형식: $p1 * x / (p2 + x)$, 모형식 이름: *m230*, 추가 설명: *Michaelis-Menten model 1*, 설명변수 개수: 1, 설명변수 이름: *x*, 모수 개수: 2, 모수 이름: *p1 p2* 로 설정할 수 있다. 모형식 표현에 사용할 수 있는 연산자로는 +, -, *, /, ^ 이 사용되며, 함수로는 exp, ln, log, sqrt, sin, cos, ... 등이 있다. 모형이름은 문자 또는 숫자로 길이에 제한이 없으나 처음 8자까지만 사용되며, 설명변수이름과 모수이름으로는 특수기호를 제외한 문자 또는 숫자로 6자 이내로 설정할 수 있다.

현재 NLIN2000에서는 설명변수의 개수는 6개까지, 모수의 개수는 20개까지 포함된 모형의 분석이 가능하다. 현재 모형 DB에는 대표적 비선형 회귀분석 관련 책(Bates & Watts,

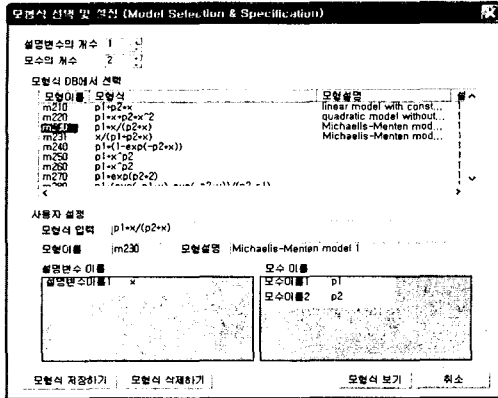


그림 2.4: 모형식 선택 및 설정 창

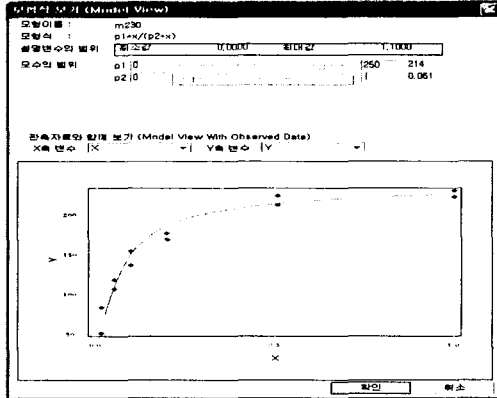


그림 2.5: 모형식 보기 창

1988; Gallant, 1987; Ratkowsky, 1983, 1988; Seber & Wild, 1989 등)에서 소개되고 있는 약 60개의 모형식이 저장되어 있으며 설명변수와 모수의 개수를 이용하여 분류되어 있다. DB에 없는 모형식의 경우에는 사용자가 필요한 정보를 그림 2.4의 하단에 보이는 ‘사용자 설정’창에서 입력하여 DB에 추가하여 사용할 수 있다.

그리고, 그림 2.4에서 사용하고자 하는 모형식을 모형이름을 클릭 하여 선택한 다음 하단 오른쪽의 ‘모형식 보기’를 누르면, 그림 2.5의 ‘모형식 보기(Model View)’ 기능창이 뜬다. 여기에서는 모수들의 값을 스크롤 바를 이용하여 변화시키면 그에 따른 모형식의 형태를 보여준다. 관측자료가 있는 경우에는 자료의 산점도와 함께 볼 수도 있다. 이와 같이 모형식의 형태를 사전에 파악할 수 있어 이를 이용하여 선택한 모형식이 사용 목적에 타당한지를 판단하는데 도움을 얻을 수 있고, 또한 모수 추정 시의 경우 관측자료를 잘 적합하는 모수들의 대략적인 값을 알 수 있다. 비선형 회귀추정에는 반복과정방법을 사용하므로 적당한 모수 초기값을 제공하는 것이 중요한데 여기에서 얻은 모수값들을 초기값으로 사용할 수 있을 것이다.

2.4. 모형적합

그림 2.6은 ‘모형적합(Model Fitting)’ 창으로 관측자료를 입력한 다음, 분석에 사용하고자 하는 모형식을 선택하고, 반응변수와 설명변수, 그리고 모수들의 초기값을 입력하면 적합 결과를 얻을 수 있다. 그림 2.7은 모형적합 시에 선택할 수 있는 몇 가지 사항들을 보여준다. 특히 곡률분석을 선택하면 다른 통계팩키지에서는 얻을 수 없는 모형의 비선형성(nonlinearity) 정도를 파악할 수 있는 여러 가지 통계량을 제공받을 수 있다.

표 2.1은 Bates와 Watts(1988)에 있는 예제 자료(Example: Puromycin)를 분석한 결과물이다. (모수 추정치들간의 분산-공분산 행렬은 삭제됨.) 분석에 사용된 변수명과 모형식에 대한 간단한 정보 등을 먼저 출력하고, 모수 추정 결과(추정치, 표준오차, 오차분산 추정치, 추정치들 간의 상관계수행렬 등)를 보여준다. 이 값들은 SAS와 SPSS에서도 계산해주는 것

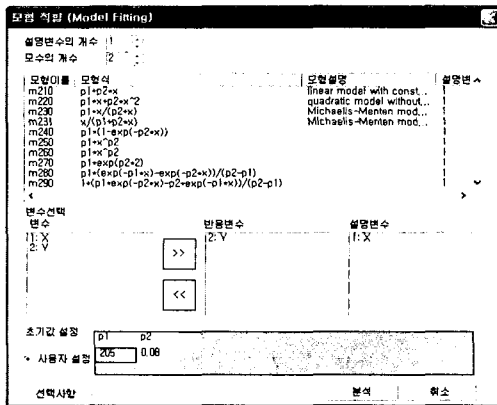


그림 2.6: 모형적합 창

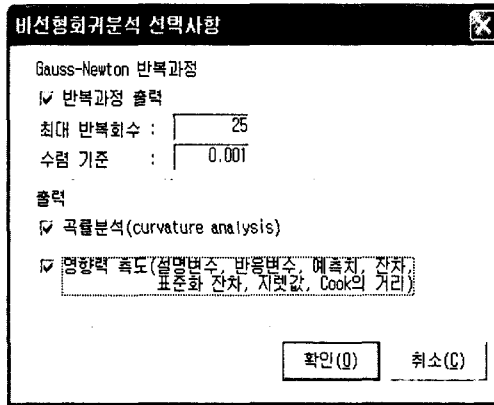


그림 2.7: 분석 선택사항 창

으로 각 패키지에서의 값들은 일치한다. 다음으로는 선택사항인 곡률분석 결과를 보여준다. 곡률분석에서 제공되는 여러 통계값들에 대한 해석은 참고문헌 Bates 와 Watts(1980), Clarke(1987), Ratkowsky(1983, 1988)를 참조하면 된다. 예를 들어, 전체곡률측도에서 고유곡률은 반응표면의 비평면 정도를 나타내는 측도이며, 모수효과곡률은 반응표면에서 좌표체계의 비선형 정도를 나타내는 측도이다. 각 측도의 크기 여부는 함께 주어진 참조기준값을 이용하여 판정할 수 있는데 이 값은 $1/(2\sqrt{F})$ 로 주어진다. 여기에서 모수의 개수를 p , 관측자료 개수를 n 이라 하면, F 는 자유도가 $(p, n - p)$ 인 F 분포의 상위 5% 백분위수를 나타낸다. 각 곡률값은 이 참조기준값보다 작으면 그 정도가 작다고 판정된다.

표 2.1: NLIN2000의 분석결과 출력문

```

◆ 선택사항
  반응변수 : V(Y)
  설명변수 : V(X)
  모형이름 : m230      Michaelis-Menten model 1
  모형식   : p1*x/(p2+x)
  모수이름 : 모수1=p1  모수1=p2

◆ 모수 추정 결과
<< Gauss-Newton 반복과정 >>

```

단계	오차제곱합	목적함수값	모수 추정값	
1	3155.0042	7.8658	205.0000	0.0800
2	1205.6618	0.0417	213.0289	0.0629
3	1195.4771	0.0001	212.6034	0.0640
4	1195.4491	0.0000	212.6754	0.0641
5	1195.4488	0.0000	212.6829	0.0641

```

<< 최종 추정 결과 >>

```

모수	모수명	추정치	근사표준오차	t값
1	p1	212.6829	6.9471	30.6146
2	p2	0.0641	0.0083	7.7432

요인	자유도	제곱합	평균 제곱	F값
회귀	2	270213.5512	135106.7756	1130.1762
오차	10	1195.4488	119.5449	
전체	12	271409.0000		
전제(수정)	11	30858.9167		

R-sq = 1 - (오차제곱합)/(수정된 전체제곱합) = 0.9613
 오차분산 추정치 : 119.5449 (자유도 : 10)
 모수 추정치들의 근사 상관계수 행렬

모수명	p1	p2
p1	1.0000	
p2	0.7	1.0000

◆ 곡률분석

<< 전체곡률(global curvature) >>
 고유곡률(intrinsic) : 0.0454
 모수모과곡률(parameter-effects) : 0.1047
 참조 기준값(유의수준=0.05) : 0.2468

<< 주변곡률(marginal curvature) >>

모수	모수명	주변곡률	주변곡률*(t)
1	p1	0.0160	0.0290
2	p2	0.0534	0.0969

<< 모수 추정치의 편향(bias)과 왜도(skewness) >>

모수	모수명	추정값	편향	%편향	왜도
1	p1	212.6829	0.1900	0.089	0.0961
2	p2	0.0641	0.0004	0.690	0.3207

◆ 모형진단

<< 모수에 대한 95% 신뢰구간 >>

모수	모수명	추정값	선형근사와 아한	신뢰구간 상한	곡률조정된 아한	신뢰구간 상한
1	p1	212.6829	197.2038	228.1621	197.7560	228.7142
2	p2	0.0641	0.0457	0.0826	0.0479	0.0848

◆ 회귀진단 통계량

설명변수 X	반응변수 Y	예측치	잔차	표준화잔차	지렛값	Cook의 D
0.0200	76.0000	50.5665	25.4335	2.4866	0.1249	0.4411
0.0200	47.0000	50.5665	-3.5665	-0.3487	0.1249	0.0087
0.0600	97.0000	102.8116	-5.8116	-0.5917	0.1931	0.0419
0.0600	107.0000	102.8116	4.1884	0.4264	0.1931	0.0218
0.1100	123.0000	134.3620	-11.3620	-1.1230	0.1437	0.1058
0.1100	139.0000	134.3620	4.6380	0.4584	0.1437	0.0176
0.2200	159.0000	164.6848	-5.6848	-0.5494	0.1043	0.0176
0.2200	152.0000	164.6848	-12.6848	-1.2258	0.1043	0.0875
0.5600	191.0000	190.8326	0.1674	0.0169	0.1766	0.0000
0.5600	201.0000	190.8326	10.1674	1.0248	0.1766	0.1127
1.1000	207.0000	200.9683	6.0317	0.6402	0.2574	0.0710
1.1000	200.0000	200.9683	-0.9683	-0.1028	0.2574	0.0018

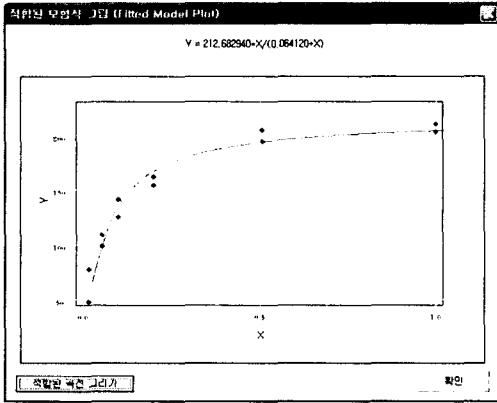


그림 2.8: 적합된 모형식 보기 창

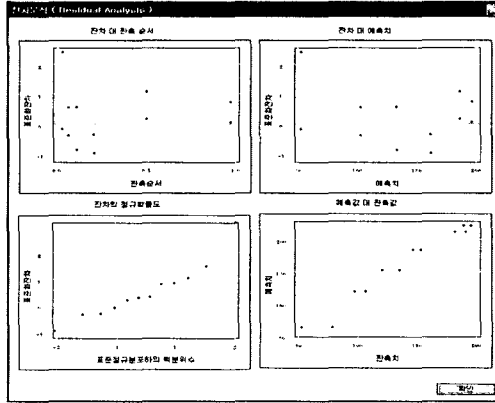


그림 2.9: 잔차의 산점도

그림 2.8은 적합된 모형식을 관측자료의 산점도와 같이 보여주는 ‘적합된 모형식 그림(Fitted Model Plot)’ 기능이다.

2.5. 잔차분석

‘잔차분석’에서는 잔차분석을 위하여 잔차 또는 표준화잔차의 여러 가지 산점도들을 보여준다. 현재는 잔차 대 예측값, 잔차 대 관측순서의 산점도와 잔차의 정규확률도를 보여주며, 그 외에 관측값과 예측값의 산점도를 볼 수 있다(그림 2.9).

3. 개발 내용

NLIN2000은 DOS 환경에서 사용되었던 NLIN90을 Window 환경으로 업그레이드시킨 것이다. 따라서 많은 기능들이 개선되어 사용자가 더욱 쉽게 사용할 수 있다. 두 프로그램간의 차이점은 표 3.1에 정리되어 있다. 개선된 기능 중 제일 대표적인 것은 symbolic computation을 이용한 도함수식의 자동 생성이라 할 수 있다. NLIN90을 사용할 때 제일 어려웠던 점은 분석하고자하는 모형식의 각 모수에 대한 일차도함수 또는 이차도함수 식을 계산하는 부프로그램을 사용자가 직접 작성하는 것이었다. 이와 같은 작업은 귀찮을 뿐만 아니라 작성과정에서도 많은 실수가 발생하여 사용자를 곤혹스럽게 한다. 특히 사용자가 통계전공자가 아닌 경우에는 이와 같은 작업을 요구하기가 무리이다. 이를 위해 도함수를 수치적으로 근사화시키는 방법도 있으나, 정확한 분석을 위해서는 도함수를 사용하는 것이 중요하고, 또한 곡률분석의 경우에는 이차도함수가 필수적이다. NLIN2000에서는 symbolic computation을 이용하여 모형식의 모수에 대한 일차 및 이차도함수식을 계산함으로써 이러한 불편을 없앴다.

모형식의 DB 역시 크게 개선된 부분 중의 하나이다. NLIN2000 자체에서 제공하는 모형식의 개수가 훨씬 많아진 것 외에 사용자가 필요한 모형식을 DB에 손쉽게 추가할 수 있

표 3.1: NLIN90과 NLIN2000의 차이점 비교

항목	NLIN90	NLIN2000
운용 시스템 환경	DOS에서의 CUI 환경	WINDOW에서의 GUI 환경
개발 언어	Fortran	C++
모형식 사용	사용자가 Fortran 함수를 작성하여 컴파일과 링크를 매번 수행	모형식 DB에서 선택을 하거나, 입력창에서 연산식으로 표현
모형식의 도함수	사용자가 Fortran 함수로 작성	자동생성
모형식 DB	Fortran 함수로 구성되어 관리 및 확장이 어려움	모형식에 대한 필수정보 입력만으로 가능하며 관리가 쉽고 무한정 확장이 가능
그래프	제공하지 않음	제공함 (모형식 그래프, 잔차 그래프)

어 관측자료와 모형식만 주어지면 분석 결과를 최단시간 내에 얻을 수 있다.

NLIN2000에 통계적 측면에서 새로 추가된 기능은 모수값의 변화에 따른 모형식의 형태를 보여주는 '모형식 보기'가 있고, 새로운 통계량으로 모수추정치에 대한 편향(bias)과 왜도(skewness)가 추가되었다(Box, 1971; Hougaard, 1985; Ratkowsky, 1988). 모수 추정에는 Gauss-Newton 방법을 사용하였으며, 수렴판단 기준으로는 Bates & Watts(1981)가 제안한 방법(SAS에서의 R 기준)을 사용하였다. 추론을 위한 기초통계량 및 곡률측도들의 계산은 Bates 와 Watts(1988)에 기술된 도함수행렬의 QR 분해를 이용한 계산방법을 사용하였다. 일반적으로 QR 분해는 계산오차를 줄이는 한편 계산속도는 증가시키는 계산방법으로 알려져 있다(Dongarra 외, 1979).

4. 다른 패키지들과의 비교

비선형 회귀분석을 위한 소프트웨어들 중에서 본 논문에서는 현재 대표적으로 사용되고 있는 SAS(version 8.02), SPSS(version 10), nls2(Bouvier & Huet, 1994)와 NLIN2000의 기능들을 비교하였다.(Minitab은 비선형 회귀분석을 위한 모듈이 없는 관계로 제외됨.)

SAS의 PROC NLIN과 SPSS의 NLR에서는 분석결과에서 공통적으로 기초 통계량이라 할 수 있는 모수 추정값, 근사 표준오차, 모수에 대한 근사 95% 신뢰구간, 모수 추정값의 근사 상관행렬, 오차분산 추정값들과 그 외 반응변수의 예측값, 잔차 등을 제공하고 있다. 그러나, NLIN2000에서와 같은 곡률분석을 위한 통계량들은 제공을 하고 있지 않으며 단지 SAS에서는 Hougaard의 왜도측도를 옵션으로 계산하여 주고 있다.

nls2는 S-PLUS에 기반하여 작성된 것으로 위의 기초통계량 외에 S-PLUS의 기능을 이용하여 다양한 그래프를 제공하여 주며, 모수에 대한 가설검정 과정을 수행할 수 있다. 그러나, 이를 위해서는 각 문제마다 S-PLUS 명령어를 이용한 프로그래밍이 필요하게 되어 S-PLUS 사용에 매우 익숙한 경우에만 이용할 수 있다.

표 4.1: 여러 통계패키지에서의 비선형 회귀분석을 위한 모듈들의 특징

패키지	특징
SAS/NLIN	<ul style="list-style-type: none"> · 모수추정을 위한 다양한 옵션을 제공 · 계산된 통계값들을 변수로 저장하여 2차 분석이 용이함
SPSS/NLR	<ul style="list-style-type: none"> · 손실 함수의 사용자 지정과 모수 제약조건 지정이 가능 · 모형식 입력창 제공 · 모수추정치의 표준오차의 붓스트랩 추정값 제공
nls2	<ul style="list-style-type: none"> · S-PLUS를 기반으로 한 것으로 다양한 그래프 제공 · S-PLUS 명령어를 사용한 프로그래밍이 필요 · 모수추정치의 표준오차와 모수 신뢰구간의 붓스트랩 추정값 제공
NLIN2000	<ul style="list-style-type: none"> · 곡률분석 통계량, 영향력 측도 등 가장 다양한 통계량 제공 · 모형식들의 DB와 모형식 사전보기 기능 제공 · 다양한 그래프를 직접 제공

SAS와 SPSS는 NLIN2000보다 적은 개수의 통계량을 계산하여 주지만 모수 추정을 위한 다양한 옵션을 제공해준다는 장점을 가지고 있다. 예를 들어, 가중치 설정 기능, 모수 추정치에 대한 제한구간 설정 기능, 오차제곱합이 아닌 임의의 목적함수를 사용자가 지정할 수 있는 기능 등을 제공한다.

SPSS와 nls2의 특징 중의 하나는 모수 추정치에 대한 표준오차의 붓스트랩(bootstrap) 추정값을 제공한다는 것이다. 곡률이 큰 경우에는 이 추정값이 모형식의 선형근사화를 이용한 값보다 더욱 정확할 수 있다. nls2에서는 붓스트랩을 이용한 모수에 대한 95% 신뢰구간도 제공하여 주는데 이는 NLIN2000에서 제공하는 곡률조정된 신뢰구간과 비슷한 역할을 한다.

NLIN2000에서는 잔차분석을 위한 그래프들 외에 각 관측치들의 영향력 분석을 위한 통계량인 지렛값(leverage)과 Cook의 거리값(distance)을 계산하여 준다. 비선형 회귀분석에서 이들 통계량은 소위 일 단계 근사(one-step approximation) 방법에 의하여 정의된다(Cook & Weisberg, 1982).

모수추정을 위한 반복과정으로 NLIN2000과 nls2에서는 Gauss-Newton 방법을 사용하며, SPSS에서는 Levenberg-Marquardt 방법과 축차 2차 계획법(sequential quadratic programming)을, SAS에서는 이 외에 Newton 방법을 추가로 제공하고 있다. 이와 같은 여러 추정 방법은 임의의 손실함수를 허용하는 SAS와 SPSS의 경우에는 필요하다고 할 수 있다. 그러나, 오차제곱합을 최소화하는 최소제곱법만을 허용하는 NLIN2000과 nls2에서는 저자들의 실제 자료를 이용한 분석경험에서 보면 Gauss-Newton 방법만으로 충분하다고 본다. 그러나, 다양성을 제공하기 위하여 이들 방법을 NLIN2000에 추가할 계획으로 있다. 계산에 필요한 도함수의 자동생성 기능은 모든 패키지가 가지고 있어 과거보다 비선형 회귀분석이 더욱 쉬워졌다고 할 수 있다. 이러한 각 패키지들의 특성은 표 4.1에 정리되어 있다.

5. 결론 및 추후 개발내용

본 논문에서는 통계전문가 뿐만 아니라 다양한 응용분야에서의 분석자들이 실제 문제 해결에 사용할 수 있는 비선형회귀분석용 통계소프트웨어 NLIN2000을 소개하였다. 이 소프트웨어는 기존의 통계팩키지보다 많은 분석 통계량과 그래프를 제공하여주는 한편, 비선형 회귀분석을 수행하는 과정에 맞게 단계별로 필요한 메뉴를 설정하여 기존의 팩키지와 차별성을 두었다. 또한 용량이 작으면서 MS 윈도우 환경에서 설치가 쉬워 많은 활용이 있으리라 기대한다. 앞으로 미비한 많은 기능들을 보완하고, 사용자들의 의견을 수렴하여 NLIN2000을 더욱 확장 발전시켜 나갈 계획이다. NLIN2000의 현재 버전은 인터넷 홈페이지 <http://stat.ssu.ac.kr/~khshim> 에서 다운받아 사용할 수 있다.

현재 추가로 개발되고 있는 부분은 설명변수가 2개인 경우의 3차원 반응표면 그림, 변환된 모수에 관한 곡률분석과 실험설계분석이다. 변환된 모수에 관한 곡률분석(Kang & Rawlings, 1998)은 모수들의 함수로 주어지는 비선형모형의 여러 특성들의 사후분석에 중요한 정보를 제공해 줄 수 있다. 그리고, 곡률분석을 위한 그래프로 profile curve(Bates & Watts, 1988)와 confidence curve(Cook & Weisberg, 1990; Kahng, 2003) 등을 시스템에 추가할 계획으로 있다.

참고문헌

- [1] 강근석 (1993). 비선형 회귀분석을 위한 소프트웨어 NLIN90의 소개, <응용통계연구>, 제6권 1호, 163-172.
- [2] Bates, D. M. and Watts, D. G. (1980). Relative Curvature Measures of Nonlinearity (with discussion), *Journal of the Royal Statistical Society, Ser. B*, Vol. 42, 1-25.
- [3] Bates, D. M. and Watts, D. G. (1981). A Relative Offset Orthogonality Convergence Criterion for Nonlinear Least Squares, *Technometrics*, Vol. 23, 179-183.
- [4] Bates, D. M. and Watts, D. G. (1988). *Nonlinear Regression Analysis and its Applications*, Wiley, New York.
- [5] Bouvier, A. and Huet, S. (1994). nls2. nonlinear regression by s-plus functions, *Computational Statistics and Data Analysis*, Vol. 18, 187-190.
- [6] Box, M. J. (1971). Bias in nonlinear estimation (with discussion), *Journal of the Royal Statistical Society, Ser. B*, Vol. 33, 171-201.
- [7] Clarke, G. P. Y. (1987). Marginal Curvatures and Their Usefulness in the Analysis of Nonlinear Regression Models, *Journal of the American Statistical Association*, Vol. 82, 844-850.
- [8] Cook, R. D. and Weisberg, S. (1982). *Residuals and Influence in Regression*, Chapman

and Hall, New York.

- [9] Cook, R. D. and Weisberg, S. (1990). Confidence Curves in Nonlinear Regression, *Journal of the American Statistical Association*, Vol. 85, 544-551.
- [10] Dongarra, J. J., Bunch, J. R., Moler, C. B., and Stewart, G. W. (1979), *Linpack Users' Guide*. Philadelphia:SIAM.
- [11] Gallant, A. R. (1987). *Nonlinear Statistical Models*, Wiley, New York.
- [12] Hougaard, P. (1985). The appropriateness of the asymptotic distribution in a nonlinear regression model in relation to curvature, *Journal of the Royal Statistical Society, Ser. B*, Vol. 47, 103-114.
- [13] Kahng, Myung-Wook (2003). Confidence Curves for a Function of Parameters in Nonlinear Regression, *Journal of the Korean Statistical Society*, Vol. 32, 1-10.
- [14] Kang, G. and Rawlings, J. (1998). Marginal Curvatures for Functions of Parameters in Nonlinear Regression, *Statistica Sinica*, Vol. 8, 467-476.
- [15] Ratkowsky, D. A. (1983). *Nonlinear Regression Modelling: A Unified Practical Approach*, Marcel Dekker, New York.
- [16] Ratkowsky, D. A. (1988). *Handbook of Nonlinear Regression Models*, Marcel Dekker, New York.
- [17] Seber, G. and Wild, C. (1989). *Nonlinear Regression*, Wiley, New York.

[2003년 7월 접수, 2003년 10월 채택]

Introduction of a Nonlinear Regression Analysis System NLIN2000

Gunseog Kang ¹⁾ Kyu-Ho Shim ²⁾

ABSTRACT

A statistical software for nonlinear regression analysis, NLIN2000, is introduced. This software, operated under the Window systems, has many user-friendly functions and provides various statistics. As an upgraded version of the previous program operated under the DOS system, NLIN2000 provides easier steps for model specification and fitting process than any other statistical packages. Also it has a database system for model functions which has addition and deletion options. While it can be a useful research tool for statisticians, NLIN2000 can be used practically also by researchers in many other scientific fields, who needs nonlinear regression analysis for their study.

Keywords: nonlinear regression analysis; nonlinearity; statistical package; curvature.

1) Professor, Department of Statistics, Soongsil University. E-mail: gskang@stat.ssu.ac.kr
2) Researcher, S-Link. E-mail: khshim@s-link.com