JOURNAL OF THE
CHUNGCHEONG MATHEMATICAL SOCIETY
Volume 8, June 1995

LINEARIZED MODELLING TECHNIQUES

YOUNG-WOO CHANG AND KYONG-HO LEE

ABSTRACT. For analyzing systems of multi-variate nonlinear equations, the linearized modelling techniques are elaborated. The technique applies Newton-Raphson iteration, partial differentiation and matrix operation providing solvable solutions to complicated problems. Practical application examples are given in; determining the zero point of functions, determining maximum (or minimum) point of functions, nonlinear regression analysis, and solving complex coefficient polynomials. Merits and demerits of linearized modelling techniques are also discussed.

개 요

다변량 비선형으로 이루어진 함수군에 대한 해석 방법의 하나로, Newton-Raphson의 반복법, 편미분, 행렬연산등을 이용한 선형화 모델 기법의 접근방법을 제시하였다. 실제 활용이 가능한 경우의 예로써, 함수계의 제로점 계산, 최대 및 최소치의 계산, 비선형 회귀분석, 복소계수 고차방정식의 해법에의 응용에 관하여 고찰하였다. 선형화 모델 기법의 장단점과 위의 각 경우에 따른 활용예를 기술하였다.

1. 서 론

비선형의 모델은, 일반적으로 n개의 변수에 대해 n개의 서로 다른 비선형의 방정식 (경우에 따라서는 미분식)을 필요로 한다. 이 비선형계는 선형계보다 복잡하고 그 해석이 단순한 반복대입법으로는 쉽게 풀리지 않는 경우가 많다. 이러한 비선형계를 선형화하는 방법의 하나로 n개의 방

Received by the editors on June 30, 1995.
1991 *Mathematics subject classifications*: Primary 65H10.

정식을 만들어서 그 각각에 독립 변수에 의한 편미분을 적용하여 Newron-Raphson 반복법으로 이루어진 모델로 변환할 수 있다. 단일 독립변수의 단일 함수로 이루어진 비선형 모델의 가장 대표적인 경우는 고차방정식의 근을 구하는 것이고, 여러개의 독립변수로 이루어진 선형 함수계의 가장 대표적인 예는 선형계획법이다. 두개 또는 그 이상의 독립 변수를 가지는 비선형 모델의 네가지 경우에 대하여 선형화한 해석 방법을 고찰 하기로 한다.

2. 선형화 모델기법의 적용

2.1 다변량 함수의 제로점 추적

주어진 함수의 독립변수가 X와 Y이고, 함수를 $F(\equiv F(X,Y))$라 한다. 이때 $F(X,Y)=0$을 만족시키는 X와 Y를 찾는 과정을 기술하기로 한다. 선형화 기법을 적용하기 위한 방정식의 추가를 위해서 새로운 함수 $G(X,Y) = (F(X,Y))^2$를 정의한다. 이 때, $F(X,Y) = 0$이면 자연히 $G(X,Y) = 0$도 정이다. 또한 정의에 의해서 $G(X,Y) \geq 0$ 이 항상 성립한다. $(G(X,Y) = (F(X,Y))^2$ 로부 터) 따라서, G의 제로점에서는 G의 X 및 Y에 대한 편미분도 0이 되어야 한다. 위의 조건들을 정리하면 다음과 같다.

(1.1) 제로점에서의 함수조건,

1) $G(X,Y) = (F(X,Y))^2 = 0$
2) $\frac{\partial G}{\partial X} = H(X,Y) = 0$
3) $\frac{\partial G}{\partial Y} = I(X,Y) = 0$

위의 세가지 조건들을 동시에 만족시키는 선형 모델은 Newton-Raphson 공식을 적용하여 다음과 같이 표현할 수 있다.

$$(1.2) \qquad \begin{pmatrix} \frac{\partial G}{\partial X} & \frac{\partial G}{\partial Y} \\ \frac{\partial H}{\partial X} & \frac{\partial H}{\partial Y} \\ \frac{\partial I}{\partial X} & \frac{\partial I}{\partial Y} \end{pmatrix} \begin{pmatrix} dX \\ dY \end{pmatrix} = \begin{pmatrix} -G(X,Y) \\ -H(X,Y) \\ -I(X,Y) \end{pmatrix}$$

혹은 간략히 표현하여,

$$(1.3) \qquad A \circ \vec{d} = -\vec{f};$$

(1.3)은 (1.2)와 동일함

위의 식 (1.3)에서 A는 3×2의 행렬이므로, 양변에 A^T를 곱하여;

$A^T \circ A \circ \vec{d} = A^T \circ -\vec{f}$가 되고 $A^T \circ A$의 역행렬 $(A^T \circ A)^{-1}$를 양변에 곱하여,

$$(A^T \circ A)^{-1} \circ (A^T \circ A) \circ \vec{d} = (A^T \circ A)^{-1} \circ A^T \circ -\vec{f}$$

좌변을 정리하면,

$$\vec{d} = (A^T \circ A)^{-1} \circ A^T - \vec{f} \tag{1.4}$$

위 식 (1.4)를 X, Y의 임의의 초기치로부터 시작하여 반복적으로 우변에 계산치를 대입하여 다음 두 식을 계산한다.

$$\begin{aligned} X_{new} &= X_{old} + d_1 (d_1 = dX) \\ Y_{new} &= Y_{old} + d_2 (d_2 = dY) \end{aligned} \tag{1.5}$$

식 (1.4)와 식 (1.5)를 반복, 교대 계산하여 수렴하는 X와 Y를 구한다.

2.2 다변량 함수의 최대점과 최소점

주어진 실수함수의 최소치 (혹은 최대치)를 구하는 방법의 하나로서 선형화 기법 을 채택할 수가 있다. 예로써, 독립변수가 세개 (X, Y, Z) 인 이차함수 S가 주어졌 을 때 $(S \equiv S(X, Y, Z))$;

(2.1)

$$S(X, Y, Z) = a + bX + cY + dZ + eX^2 + fY^2 + gZ^2 + hXY + iXZ + jYZ$$

위 식에서 a 부터 j 까지는 주어진 각 항의 계수들이고, X, Y,Z는 독립변수들 이다. (이는 박 성현 교수의 저서 "회귀분석" [3] p.538에서 인용한 문제이다. 단, 계산의 편의상 이차항의 계수들은 모두 양수일 경우에 국한하였음.) 이 함수 S는 위로 열려 있으므로 최소치를 가짐을 알 수 있다. 따라서, 이 최소점에서의 X,Y,Z 각각에 대한 이 함수의 기울기 (즉,

편미분)는 모두 0이 되어야 한다. 또한 X,Y,Z에 대한 S의 일차 편미분은 각각 다음과 같이 주어진다.

$$
\begin{aligned}
\frac{\partial S}{\partial X} &= b + 2eX + hY + iZ = 0 = F_1 \\
\frac{\partial S}{\partial Y} &= c + 2fY + jZ + hX = 0 = F_2 \\
\frac{\partial S}{\partial Z} &= d + 2gX + iY + jZ = 0 = F_3
\end{aligned} \tag{2.2}
$$

F_1, F_2, F_3는 새로 정의됨. F_1, F_2, F_3의 세 함수를 X, Y, Z로 한 번 더 편미분을 행하면 Jacobian 행렬의 element 들이 다음과 같이 구해진다.

$$
\begin{array}{lll}
\dfrac{\partial F_1}{\partial X} = 2e & \dfrac{\partial F_1}{\partial Y} = h & \dfrac{\partial F_1}{\partial Z} = i \\
\dfrac{\partial F_2}{\partial X} = h & \dfrac{\partial F_2}{\partial Y} = 2f & \dfrac{\partial F_2}{\partial Z} = j \\
\dfrac{\partial F_3}{\partial X} = i & \dfrac{\partial F_3}{\partial Y} = j & \dfrac{\partial F_3}{\partial Z} = 2g
\end{array} \tag{2.3}
$$

위의 사항들을 종합적으로 정리하여 다음과 같은 행렬식으로 나타낼 수 있다.

$$
\begin{pmatrix} 2e & h & i \\ h & 2f & j \\ i & j & 2g \end{pmatrix} \begin{pmatrix} dX \\ dY \\ dZ \end{pmatrix} = \begin{pmatrix} -F_1 \\ -F_2 \\ -F_3 \end{pmatrix} \tag{2.4}
$$

위 행렬식 (2.4)에서의 행렬은 대칭형이나, 이것은 원함수 S가 X, Y, Z에 대 하여 대칭형이 되도록 정의되었기 때문이며, 일반적으로는 대칭형이 아니다. 또한, 이 문제의 경우에 한해서 a 부터 j까지의 계수들은 모두 상수이 므로 어떤 초기치에 의해서도 단 한번의 역행렬 연산으로 해가 구해진다. 원함수가 2차 이상일 경우에는 반복계산에 의해서 그 해가 얻어진다.

반복계산에서는;

$X_{new} = X_{old} + dX$ (Y 와 Z에 대해서도 마찬가지 공식을 적용함)

와 같은 요령으로 X, Y, Z의 수렴치를 구해 나간다.

2.3 비선형 다중회귀분석

어떤 시스템을 선형식이 아닌 일반함수로 다중회귀분석을 하고자 할 때, 구하려는 식을

(3.1) $$Y = F(\vec{\theta}, \vec{X})$$

(위 식에서, $\vec{\theta}$: 구하는 식의 계수 벡터 $\vec{X}$: 독립변수 벡터) 라 놓으면, 관측치군의 각 경우에 k 라는 첨자를 붙여,

(3.2) $$Y_k = F(\theta, X_k)$$

; (k 는 관측 순번 (Observation serial number)) 로 표기할 수 있다.

다중회귀분석의 정의는 식 (3.2)에 독립변수(X 벡터)와 종속변수 (Y 벡터)의 각 관측치의 모든 경우($k = 1$ to n ; n is the total number of observations)에 대한 자승합을 최소로 하는 계수 벡터를 구하기 위하여 "계수에 의한 편미분" 을 행하는 과정이다. 따라서, 이 정의를 식으로 나타내면 다음과 같다.

(3.3) $$\frac{\partial}{\partial\theta_i}\{\Sigma_k(Y_k - F_k)^2\} = 0 \quad \text{for each} \quad i$$

(이 식에서 "i" 는 각 계수를 나타내는 첨자임.) 식 (3.3)을 풀고 2로 나눈 것을 새로 정의 하여 i에 대하여 각각,

(3.4) $$G_i = \Sigma_k\{(Y_k - F_k) \circ \frac{\partial F_k}{\partial\theta_i}\} = 0$$

라 한다.

위 식 (3.4)에서 F_k 의 일차 편미분 함수를 $\frac{\partial F_k}{\partial\theta_i} = R_{ki}$ 라 정의하고 각 G_i 에 편미분을 한번씩 더 행하면,

(3.5) $$\frac{\partial G_i}{\partial\theta_j} = \Sigma_k\left\{-R_{kj} \circ R_{ki} + (Y_k - F_k) \circ \frac{\partial R_{kj}}{\partial\theta_j}\right\}$$

식 (3.5)는 Jacobi 행렬의 각 원소를 나타내는 것이 되고, 이 문제를 풀기 위한 행렬식은 다음과 같다. (다음의 행렬식은 편의상 계수와 독립변수가 각각 세개인 경우를 보여줌.)

$$\begin{pmatrix} \frac{\partial G_1}{\partial \theta_1} & \frac{\partial G_1}{\partial \theta_2} & \frac{\partial G_1}{\partial \theta_3} \\ \frac{\partial G_2}{\partial \theta_1} & \frac{\partial G_2}{\partial \theta_2} & \frac{\partial G_2}{\partial \theta_3} \\ \frac{\partial G_3}{\partial \theta_1} & \frac{\partial G_3}{\partial \theta_2} & \frac{\partial G_3}{\partial \theta_3} \end{pmatrix} \begin{pmatrix} d\theta_1 \\ d\theta_2 \\ d\theta_3 \end{pmatrix} = \begin{pmatrix} -G_1 \\ -G_2 \\ -G_3 \end{pmatrix} \tag{3.6}$$

위 식 (3.6)에 θ_i $(i = 1, 2, 3)$ 의 초기치를 대입한 후에 반복적인 계산에 의하여 계수들을 구하고, 이 계수들을 원식 (3.1)에 대입하면 구하는 함수식이 주어지고 회귀분석이 끝나게 된다.

2.4 고차방정식의 해석

이 절에서는 고차방정식, 특히 일반적인 경우의 예로서 계수가 복소수인 경우에 근을 삼각함수 변환으로 치환하는 해석방법에 관하여 고찰하기로 한다. 이러한 삼각함수 변환은 행렬의 Eigenvalue 계산에 있어서 실수부와 허수부를 분리하는 계산 방법을 제시한 Murthy[5]의 방법과 유사하며, 저자의 다른 논문[4]에서 실수 계수의 경우와 함께 자세히 설명되어 있다.

복소계수의 고차방정식은 실수계수의 그것과 비교하여 다음과 같은 특성이 있다. 즉, 실계수의 경우에는 복소수근이 있을 경우 그의 Conjugate Pair 도 자동으로 근이 되지만 복소계수의 경우에는 일반적으로 그렇지 않다는 것이다.

다음에 설명하는 복소계수 고차방정식의 해법에 의한 알고리즘은 그 부분집합에 해당되는 실계수의 고차방정식의 해석에도 적용될 수 있다.

3차의 복소계수 고차방정식을 예로 들어 설명하기로 한다. 고차항으로부터 상수 항까지의 계수들이 $a_k + i \circ b_k$ $(k = 1$ to 4 in this example) 로 주어져 있다고 할 때, 다음과 같이 근 X 에 대해 삼각함수로 변환을 한다.

어떤 고차방정식 Y 가,

$$
\begin{aligned}
Y &= Y(X) \\
&= (a_1 + i \circ b_1)X^3 + (a_2 + i \circ b_2)X^2 + (a_3 + i \circ b_3)X^1 + (a_4 + i \circ b_4) \\
&= 0
\end{aligned}
$$

로 주어져 있을 때 X 를 다음과 같이 치환한다,

$$X = r \circ (\cos h + i \circ \sin h)$$

그러면 X 의 함수 Y 는 r 과 h 의 함수로 재정의된다. (편의상 새로운 함수 Y 에 대하여 같은 기호를 쓰기로 함.)

$$
\begin{aligned}
(4.1) \qquad Y(r,h) &= (a_1 + b_1 \circ i)r^3(\cos 3h + i \circ \sin 3h) \\
&+ (a_2 + b_2 \circ i)r^2(\cos 2h + i \circ \sin 2h) \\
&+ (a_3 + b_3 \circ i)r(\cos h + i \circ \sin h) \\
&+ (a_4 + b_4 \circ i) = 0 (\equiv F_1 + i \circ F_2 : \text{Def.})
\end{aligned}
$$

위의 식 (4.1)을 정리하고 실수부와 허수부를 분리해내면, 독립변수 r 과 h 로 정의 되는 두 개의 새로운 함수 F_1 과 F_2 가 다음과 같이 유도된다.

$$
\begin{aligned}
(4.2) \qquad F_1 &= r^3(a_1 \cos 3h - b_1 \sin 3h) \\
&+ r^2(a_2 \cos 2h - b_2 \sin 2h) \\
&+ r(a_3 \cos h - b_3 \sin h) + a_4 = 0 \\
F_2 &= r^3(a_1 \sin 3h + b_1 \cos 3h) \\
&+ r^2(a_2 \sin 2h + b_2 \cos 2h) \\
&+ r(a_3 \sin h + b_3 \cos h) + b_4 = 0
\end{aligned}
$$

여기에서 F_1 과 F_2 는 각각 r 과 h 의 함수이고 또한 실수함수임을 알 수 있다. 선형화 기법을 도입하기 위하여 두 함수의 r 과 h 에 대한 편미분을

구하면;

(4.3)
$$\begin{aligned}
\frac{\partial F_1}{\partial r} &= 3r^2(a_1\cos 3h - b_1 \sin 3h) \\
&\quad + 2r(a_2\cos 2h - b_2\sin 2h) \\
&\quad + (a_3 \cos h - b_3 \sin h) \\
\frac{\partial F_1}{\partial h} &= -3r^3(a_1 \sin 3h + b_1 \cos 3h) \\
&\quad - 2r^2(a_2 \sin 2h + b_2 \cos 2h) \\
&\quad - r(a_3 \sin h + b_3 \cos h) \\
&\quad\quad [\equiv -r \circ \frac{\partial F_2}{\partial r}] \\
\frac{\partial F_2}{\partial r} &= 3r^2(a_1 \sin 3h + b_1 \cos 3h) \\
&\quad + 2r(a_2 \sin 2h + b_2 \cos 2h) \\
&\quad + (a_3 \sin h + b_3 \cos h) \\
\frac{\partial F_2}{\partial h} &= 3r^3(a_1\cos 3h - b_1 \sin 3h) \\
&\quad + 2r^2(a_2\cos 2h - b_2\sin 2h) \\
&\quad + r(a_3 \cos h - b_3 \sin h) \\
&\quad\quad [\equiv r \circ \frac{\partial F_1}{\partial r}]
\end{aligned}$$

식 (4.3)의 항들을 사용하여 Newton-Raphson 반복계산을 하기 위한 행렬식을 나타내면;

(4.4)
$$\begin{pmatrix} \frac{\partial F_1}{\partial r} & -r \circ \frac{\partial F_2}{\partial r} \\ \frac{\partial F_2}{\partial r} & r \circ \frac{\partial F_1}{\partial r} \end{pmatrix} \begin{pmatrix} dr \\ dh \end{pmatrix} = \begin{pmatrix} -F_1 \\ -F_2 \end{pmatrix}$$

위 식 (4.4)로 표현되는 시스템은 r과 h의 임의의 초기치로 시작하여 풀어낼 수 있으며, 많이 알려진 Jenkins-Traub 방법보다 훨씬 더 정확히 근들을 계산해낼 수 있음이 확인되었다.

근을 구해내는 각 단계에서의 반복계산을 중단하기 위하여 두가지의 판정기준이 쓰인다. 한가지는 r과 h의 Tolerance (예: $1.e$-8)를 확인하고, 또한 다음의 determinant 값이 너무 작지 않은 한도 내에서 반복계산을 수행한다.

$$r \circ \{(\frac{\partial F_1}{\partial r})^2 + (\frac{\partial F_2}{\partial r})^2\}$$

Determinant의 값이 너무 작아지면 반복계산을 중지한다. 위에 설명한 방법으로 근이 하나씩 계산되며, 방정식의 차수를 하나씩 내려간다. 실수부와 허수부가 미리 분리되어 있으므로 행렬연산에서는 실수계산만이 수행되고 방정식의 차수를 내리는 계산 (Synthetic Division)에서는 복소수 계산이 필요하다.

4. 결 론.

본 논문에서는 2차 이상의 고차방정식을 포함한 일반적인 다변량 실수함 수의 여러 형태에 대한 해석방법의 하나로서 선형화 모델의 기법을 고찰 하였다. 다변량 함수계에서 함수의 제로점, 최대 (혹은 최소)치의 계산, 일반 함수로의 비선형 회귀분석, 그리고 복소계수 고차방정식의 해법에 대하 여 이 방법이 응용될 수 있는 실례를 설명하였다. 이 논문에 소개된 이외에도, 단순 필산으로는 불가능하고 컴퓨터를 이용한 복잡한 계산이 필요한 해석기하학적인 문제들의 해법에도 적용될 수 있을 것으로 보인다. Murthy[5]의 경우에는, 복소수를 원소로 갖는 행렬에서의 고유치 (Eigenvalues) 계산에 저자[4]와 유사한 방법으로 실수부와 허수부를 분리한 다음, 이를 선형화한 모델로 만들어 실수함수만을 취급하는 반복해법에 의해 풀어나가는 것을 보여주었다. 이 방법의 단점은 해의 근방에서 수렴이 느린 경우가 발생하기도 하고[3], 초기치의 추산이 잘못되면 다른 사분면에서 초기치를 취하여야 할 때가 있다[4]. 그러나 복소계수 고차방정식 해법[4]의 경우는 알고리즘 보완으 로 극히 정밀한 해를 구할 수 있었다.

참고 문헌

1. Anthony Ralston and Philip Rabinovitz, *A First Course in Numerical Analysis", 2nd Ed.*, McGraw-Hill.

2. J. Stoner and R. Bulirsch, *Introduction to Numerical Analysis*, Springer-Verlag.
3. 박성현, 회귀분석, 대영사.
4. 장영우, 삼각함수 변환을 이용한 복소계수 고차 방정식의 해법, 정보과학회 논문지 **7** (1992).
5. Durba V. Murthy, University of Toledo, Ohio, *Solution and Sensitivity Analysis of Complex Trascendental Eigenvalues with Pairs of Real Eigenvalues*, Int. J. for Num. Math in Engineering **33** (1992), 115-129.

KOREA ATOMIC ENERGY RESEARCH INSTITUTE
TAEJON 305-353, KOREA