# 수공학에서 Broken Line Smoothing의 적용

김 진환*, 이 창준*, 김 민환**

## 1. 서 론

자연과학이나 공학의 대부분 현상과 일부 사회현상은 일정한 관계를 만족하는 수식으로 표현 될 수 있다. 그러나 실제로 측정 문제에서 여러 오차가 발생하거나 수식이 너무 이상화되어 있어서 예측하지 못했던 요소들이 결과에 영향을 줄 수 있다. 이 경우에 대안으로 여러 방법의 보완과 개발 및 수정에 대한 연구가 계속되어 오고 있다. 본 연구에서는 이러한 방법들 중 하나인 BLS(Broken Line Smoothing)를 소개하고 수공학 문제에 적용을 시도하였다. BLS방법은 보간과 회귀분석의 적용이 가능한 유연성을 가지고 있다. 이러한 특징을 가진 BLS의 검증을 위해서 임의의 단계함수와 지수함수에 BLS방법을 적용하여 검증을 수행하였고 수공학 분야 중 수위-유량곡선의 개발에 적용하였다.

## 2. BLS의 수학적 구성

일련의 데이터에 대한 일반화를 위해서 다음과 같은 조건을 수행한다. (Demetris Koutsoyiannis, 2000)
$xy$ 평면에서 $i=1,\cdots,n$ 에 대하여 ( $x_i, y_i$)를 $n$개의 점이라 할 때 모든 $x_i$를 포함한 구간 [ $c_0, c_m$]에서 $x$축 상에 ( $c_j, d_j$) $j=0,\cdots,m$는 $m+1$개의 점이라 하고 점들은 등간격 즉, $c_j-c_{j-1}=\delta$ 라고 가정하자. 이러한 관계조건은 그림1과 같다. 이것은 데이타( $x_i, y_i$)의 핏팅과 $m+1$개의 $d_j$를 찾기위한 것이다

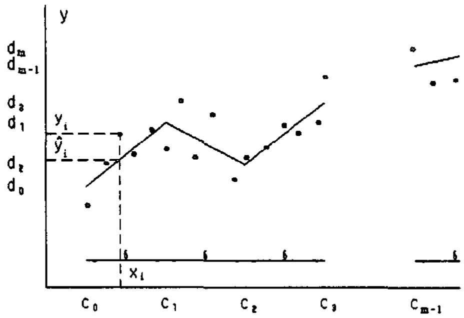


그림 1. BLS 개념도

이 핏팅은 데이타( $x_i, y_i$)와 핏팅된 파선(Broken Line; $c_j, d_j$)사이에 차의 제곱합을 최소화하는 항을 의미한다. 총 제곱오차의 최소화는 다음과 같이 표현된다.

$$p=\sum_{i=1}^{n}(y_i-\hat{y}_i)^2 \tag{1}$$

윗 식은 다음과 같이 간단히 나타낼 수 있다.

$$p=\| y-\hat{y}\|^2 \tag{2}$$

여기서, $\hat{y}$는 주어진 $x_i$에 대해 BL에 의해 구한 $y_i$의 추정값이며 다음과 같이 나타낼 수 있다.

$$\hat{y}=\Pi d \tag{3}$$

---

* 호남대학교 대학원 토목환경공학과 석사과정
** 호남대학교 토목환경공학과 교수

여기서, $d=[d_0,\cdots,d_m]^T$는 크기가 $[(m+1)\times 1]$인 벡터이고 $\Pi$는 $[n\times(m+1)]$크기의 매트릭스이며 $ij$ 번째 요소( $i=1,\cdots,n; j=1,\cdots,m+1$)는 다음 식과 같다.

$$\pi_{ij}=\begin{cases}\dfrac{x_i-c_{j-1}}{\delta}, & c_{j-2}<x_i\le c_{j-1}\\ \dfrac{c_j-x_i}{\delta}, & c_{j-1}<x_i\le c_j\\ 0, & otherwise\end{cases} \tag{4}$$

매우 거친 BL을 피하기 위한 목적과 핏팅 문제의 유일한 해를 보장하기 위한 것이다. BL의 거침 정도를 얻기 위하여 2개의 BL에 대한 연속적인 구간의 경사 차이를 구하면 다음과 같다.

$$q=\sum_{j=1}^{m-1}(2d_j-d_{j-1}-d_{j+1})^2 \tag{5}$$

윗 식을 매트릭스 형태로 나타내면 다음과 같다.

$$q=d^T\Psi^T\Psi d \tag{6}$$

여기서, $\Psi$는 $(m-1)\times(m+1)$매트릭스이며 $ij$번째 요소는 다음과 같다.

$$\Psi_{ij}=\begin{cases}2, & j=i+1\\ -1, & |j-i-1|=1\\ 0, & otherwise\end{cases} \tag{7}$$

식(2)와 식(6)을 조합하고 $q$에 대해 무차원 곱수 ($\lambda\ge 0$)를 도입하여 목적함수를 구성하면 다음과 같다.

$$f(d):=(y-\prod d)^T(y-\prod d)+\lambda\, d^T\Psi^T\Psi d \tag{8}$$

3. BLS의 매개변수 선택

BLS는 두 개의 조정 가능한 변수를 가지고 있다. 즉 구간의 수 $m$과 매끈함의 변수 $\lambda$이다. 변수의 선택은 자료의 매끈함 크기를 어느 정도 부과하느냐에 따라 수행이 가능하다.

$\lambda$의 더욱 편리한 탐색에 대해 다른 변수 $\tau$의 항을 도입하여 $\lambda$의 변환변수로 사용할 수 있으며 그 값의 범위는 0과 1사이로 제한된다(Demetris Koutsoyiannis, 2000). $\tau$항을 도입한 변환은 다음과 같은 형태를 갖는다.

$$\lambda=(10m\frac{\ln\tau_m}{\ln\tau})^k \tag{9}$$

여기서, $\tau_m$=0.99는 $\lambda$의 상한계에 해당되는 $\tau$의 최대 허용값이며 상수 $k$는 다음과 같다.

$$k=\frac{\ln\lambda_m}{\ln(10m)},\qquad \lambda_m=\frac{trace(\prod^T\prod)}{trace(\Psi^T\Psi)}10^8 \tag{10}$$

4 BLS의 검증

BLS방법의 적용성을 시험하기 위하여 임의의 함수에 잡음(noise)이 첨가된 random data를 임의의 자료군

으로 형성하고 이에 대하여 BL의 핏팅을 수행하였다. 사용된 함수는 표 1과 같다(문영일, 2001).

이에 대한 결과를 그림 2~5에 제시하였다. m=99와 m=49에 대해 $\tau$에 따라서 그림 2는 $\tau$=0.30, $\tau$=0.04, 그림 3은 각각 $\tau$=0.24, 0.64에서 그림 4는 각각 $\tau$=0.47, 0.02에서 그림 5는 각각 $\tau$=0.60, 0.29에서 가장 매끄러운 형태를 나타냈다.

표 1. 적용함수

| 구 분 | True Function | Noise | 구간수(m) | Sample size |
|---|---|---|---|---|
| 1 | f(t)=2 $0.5<t\leq 1$<br>f(t)=1 $0\leq t\leq 0.5$ | N(0,0.04) | m=99<br>m=49 | 100 |
| 2 | f(t)=2 $0.5<t\leq 1$<br>f(t)=1 $0\leq t\leq 0.5$ | N(0,1) | m=99<br>m=49 | 100 |
| 3 | $f(t)=e^{-t}\sin(2\pi t)$ $0\leq t\leq 1$ | N(0,0.01) | m=99<br>m=49 | 100 |
| 4 | $f(t)=e^{-t}\sin(2\pi t)$ $0\leq t\leq 1$ | N(0,0.25) | m=99<br>m=49 | 100 |

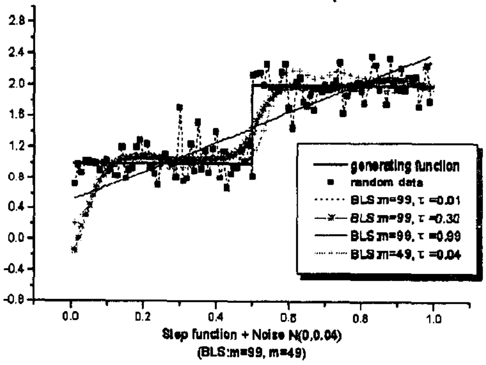


그림 2. 함수 1의 결과

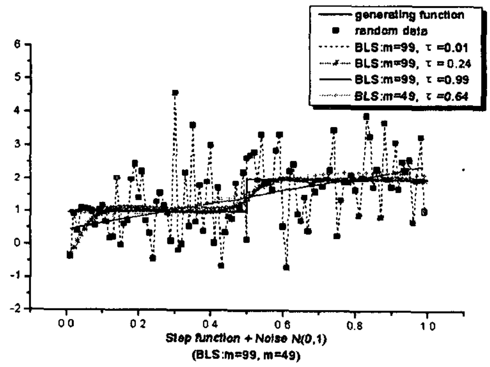


그림 3. 함수 2의 결과

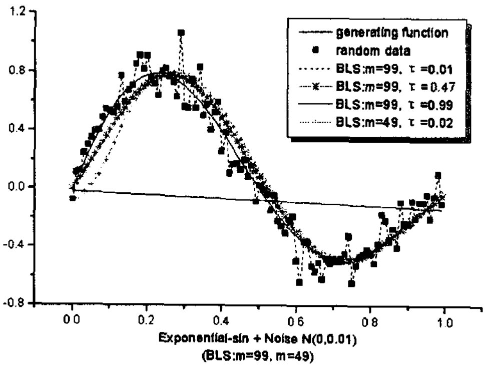


그림 4. 함수 3의 결과

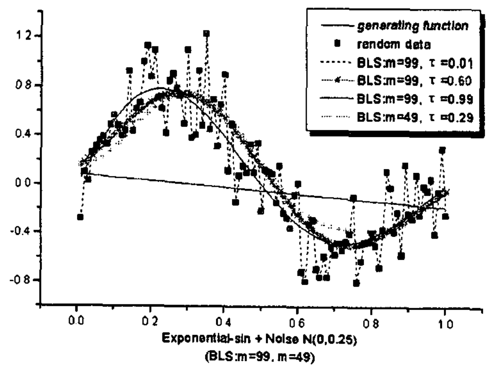


그림 5. 함수 4의 결과

## 5. 수위-유량곡선에의 적용

기존의 수위-유량곡선 개발에 사용되는 매개변수적 기법과 BLS를 적용한 수위-유량곡선을 비교함으로써

이에 대한 적용성을 시험하였다. 이에 사용된 자료는 동진강 수계 태인 수위표의 수위-유량자료이다. 이 자료에 대해 상관계수 값을 구한 결과 매개변수적 방법은 0.991, BLS 방법은 0.996이었다. 매개변수적 방법에 의해 개발된 수위-유량곡선과 BLS를 적용하여 개발된 수위-유량곡선을 그림 6에 제시하였다.

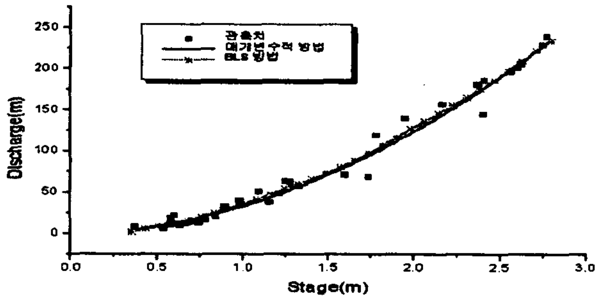


그림 6. 매개변수적 방법과 BLS 방법에
의한 수위-유량곡선의 비교

## 4. 결 론

본 연구에서 제안한 BLS(Broken line smoothing)는 펏팅에 따른 오차와 BL(broken line)의 거침을 최소화하기 위한 것이며 δ(broken line의 연속적인 구간의 길이)가 주어진 자료 점의 구간과 반드시 일치하지 않고 임의의 대상영역에 대해 곡선의 펏팅을 수행할 수 있다. 이러한 특징을 가진 BLS를 검증하고 수공학 문제에 적용해 본 결과 다음과 같은 결론을 얻었다.

(1) BLS를 단계함수와 지수함수에 의하여 검증한 결과, 원함수에 근접한 매끄러운 펏팅의 결과를 얻을 수 있었다.

(2) BLS를 검증한 후에 수위-유량곡선의 개발에 적용한 결과 매개변수적 방법보다 BLS방법에 의한 상관계수가 더 양호하였다.

## 5. 참고문헌


이병해, 심종성. 1996. 수치해석, 구미서관, pp. 309~339.

문영일. 2001. General cross-validation과 Least squares method에 의한 비매개변수적 회귀모형의 특성, 2001 학술발표회 논문집(Ⅰ), 한국수자원학회, pp. 92~97.

Craven, P. 1979. Smoothing noisy data with spline functions. Numerische Mathematik 31, pp, 377~ 403

Bellman, R., Roth, R. 1969. Curve fitting by segmented straight lines. Journal of the American Statistical Association 64, pp. 1079~1084.

Demetris Koutsoyiannis. 2000. Broken line smoothing: a simple method for interpolation and smoothing data series, Environmental Modeling & Software 15, pp. 139~149.