Journal of the Korean Institute of Industrial Engineers
Vol. 16, No. 2, Dec., 1990

임의의 수준변화에 적절히 반응할 수 있는 지수이동가중평균법

전 덕 빈*

Exponential Smoothing with an Adaptive Response to Random Level Changes

Duk Bin Jun*

Abstract

Exponential smoothing methods have enjoyed a long history of successful applications and have been used in forecasting for many years. However, it has been long known that one of the deficiencies of the method is an inability to respond quickly to interventions, to interruptions, or to large changes in level of the underlying process. An exponential smoothing method adaptive to repeated random level changes is proposed using a change-detection statistic derived from a simple dynamic linear model. The results are compared with Trigg and Leach's and the exponential smoothing methods.

1. 서 론

지수이동가중평균법(Exponential Smoothing)은 1961년 Brown과 Meyer[2]가 그 방법을 제시한 이래 현재까지 여러 분야에 걸쳐 간단한 방법임에도 불구하고 대단히 효과적임이 입증되었다. 또한 이 방법이 최적인 시계열의 underlying process에 관한 연구들이 Harrison[4], Box와 Jenkins[1] 및 Harvey[6] 등에서 행하여졌다.

그러나 이 방법은 시계열의 underlying process가 급격한 수준의 변화(level change) 또는 기울기(slope)를 포함할 때, 그러한 변화에 신속히 적응하지 못하는 것 또한 널리 알려진 사실이다. 그 중 기울기를 포함한 시계열의 경우 지수이동가중평균법 적용시 해결방안들은 Brown[3], Holt[7] 및 Harrison[4]에서 충분히 설명되었다. 단 급격한 수

* 한국과학기술원 경영과학과

** 본 연구는 1988년 한국과학재단의 일반연구비 지원에 의한 것임

준의 변화에 적응하는 방법으로서는 현재까지 Trigg과 Leach[12]의 방법이 우수한 것으로서 나타나고 있다. 그들의 방법은 매우 간단하면서도 수준의 변화 또는 기울기의 변화에 적응하는 속도가 지수이동가중평균법에 비하여 대단히 빠르지만, 그 방법의 이론적인 근거가 제시되지 못한 것이 지적되어 왔다.

전덕빈[8], 전덕빈과 Oliver[10]는 Harrison과 Stevens[5]의 Dynamic Linear Model에서의 수준 또는 기울기의 급격한 변화시 그 변화시기 및 변화량의 추정법을 제시하면서 지수이동가중평균법 적용시 수준변화를 감지하는 통계치를 유도하였다. 본 연구에서는 그 통계치를 활용하여 수준변화가 거듭 발생하는 경우에도 신속히 적응할 수 있는 지수이동가중평균법을 제시하며, 수준변화 발생횟수 및 변화량이 증가할 때 본 연구에서 제시하는 방법이 Trigg과 Leach의 방법보다 우수함을 실험예로서 보이고자 한다.

2. 수준에서의 변화를 고려한 구조모형

예측하고자 하는 확률변수 Z_t는 관찰할 수 없는 수준 L_t에 의존하며 관측 오차 a_t가 추가되어 시간 t가 지나면서 관찰되며, L_t는 random walk process를 따르는 Harrison과 Stevens[5]가 제안한 가장 간단한 형태의 구조모형에서, 시간 M에 예기치 않은 수준에서의 변화 Δ가 일어난 상황을 모형화하면 다음과 같다.

$$Z_{t+1}=L_{t+1}+a_{t+1},\ t=0,1,\cdots,n-1 \quad \cdots\cdots (1a)$$

$$L_{t+1}=L_t+b_{t+1},\quad t\neq M \quad \cdots\cdots (1b)$$

$$=L_t+\Delta+b_{t+1},\quad t=M \quad \cdots\cdots (1c)$$

관측오차$\{a_t\}$와 교란변수(perturbation noise) $\{b_t\}$는 serially uncorrelated한 정규분포를 따르며 기대값 0과 σ_a^2와σ_b^2의 분산을 가지며 σ_a^2와 σ_b^2는 알고 있다고 가정한다. 변화시기를 나타내는 확률변수 M은 1부터 n−1까지의 값을 취할 수 있는 일반적인 이산분포 p(M)을 따르며, L_t와 Δ의 시간 0에서의 사전 확률 분포는 다음과 같이 정규분포를 따르고 M, L_0, Δ, $\{a_t\}$, $\{b_t\}$는 서로 독립이라 가정한다.

$$L_0\sim N(\mu_0,\ p_0)\ ;\ \Delta\sim N(\lambda_0,q_0) \quad \cdots\cdots (2)$$

2.1 수준만을 포함한 구조 모형

$\Delta=0$인 경우 즉, 수준에서의 예기치 않은 변화가 없는 원래의 구조모형에서 L_t의 사후분포에 관한 반복계산식은 Harrison과 Stevens[5]에 다음과 같이 유도되었으며

$$L_t\mid z_t,\cdots,z_1\sim N(\mu_t,\ p_t),\qquad t=1,2,\cdots,n \qquad (3a)$$

$$p_{t+1}=\frac{\sigma_a^2(p_t+\sigma_b^2)}{p_t+\sigma_a^2+\sigma_b^2},\qquad t=0,1,\cdots,n-1 \quad (3b)$$

$$\mu_{t+1}=\mu_t+\frac{p_{t+1}}{\sigma_a^2}(Z_{t+1}-\mu_t),\qquad t=0,1,\cdots,n-1 \quad (3c)$$

시간 t에서 일단계 미래값의 예측치(one-step-ahead forecast), 즉$\{Z_t,\cdots,Z_1\}$를 관찰하여 Z_{t+1}의 조건부 확률 분포를 계산했을 때의 기대치를 f_t, 실제값과 그 예측치와의 오차를 $e_t=Z_{t+1}-f_t$, 그의 분산을 ν_t라 표시할 때

$$f_t=\mu_t\ ;\ \nu_t=p_t+\sigma_b^2+\sigma_a^2 \qquad t=0,1,\cdots,n \qquad (4)$$

2.2 변화시기를 알고있는 경우

변화량 Δ는 주어진 기간중 단, 일회, 시간 m에서 발생하여 Z_{m+1}에 처음으로 관찰되므로 그 이전의 관찰변수들의 분포는 $\Delta=0$인 위의 경우와 동일하다.

$$Z_{t+1}\mid z_t,\cdots,z_1\sim N(f_t,\nu_t)\qquad t=1,\cdots,m-1 \qquad (5)$$

변화시기 m이후에 관찰변수들이 다음과 같은 분포를 따르게 됨을 유사한 방법을 적용하여 보일 수 있다.

$$Z_{t+1} \mid z_t, \cdots, z_1;\ m;\ \Delta \sim N(f_t + \kappa_t \Delta, v_t) \quad t=m, \cdots, n \quad \cdots\cdots (6a)$$

$$\kappa_t = 1, \qquad t=m \quad (6b)$$

$$= \pi_{i=m+1}^{t} \frac{\sigma_a^2}{v_{i-1}}, \qquad t=m+1, \cdots, n \quad (6c)$$

시간 0에서의 변화량의 사전확률분포 $N(\lambda_0, q_0)$를 고려하였을 때 시간 t에서 Δ의 사후확률분포는 다음과 같다.

$$\Delta \mid z_t, \cdots, z_1;\ m \sim N(\lambda_t, q_t), \quad t=1, \cdots, n \quad (7a)$$

$$q_t = q_0, \qquad t=1, \cdots, m \quad (7b)$$

$$= \frac{1}{(1 \mid q_0) + \Sigma_{i=m}^{t-1}(\kappa_i^2 \mid v_i)}, \quad t=m+1, \cdots, n \quad (7c)$$

$$\lambda_t = \lambda_0, \qquad t=1, \cdots, \mathrm{m} \quad (7d)$$

$$= q_t\{\frac{\lambda_0}{q_0} + \sum_{i=m}^{t-1} \frac{\kappa_i}{v_i} e_i\}, \qquad t=m+1, \cdots, n \quad (7e)$$

Z_{m+1}이전의 관찰치에는 Δ가 포함되지 않기 때문에, 변화가 발생하기 전 Δ의 확률분포는 시간 0에 주어진 사전확률분포와 같고 Jun과 Oliver[9]에서의 λ_t, g_t의 반복계산식과 위의 결과는 동일하다.

식(3b, 4)의 p_t와 v_t는 관찰치와 무관하여 시간이 경과한 후 p^*, v^*로 수렴하며(Oliver[11]참조), 식(7c)의 q_t도 역시 수렴함을 Jun[8]에서 볼 수 있다.

2.3 미지의 변화시기의 경우

변화시기 M을 알 수 없는 상황하에서 시간 0에서 사전확률분포로서 p(M)을 사용한 경우 M과 Δ의 사후확률분포는 다음과 같다.

$$\mathrm{p}(\mathrm{M} \mid z_n, \cdots, z_1) \propto \mathrm{p}(\mathrm{M}) \sqrt{q_n} \exp(\frac{\lambda_n^2}{2q_n}),$$

$$\mathrm{M}=1, \cdots, n-1 \quad \cdots\cdots (8a)$$

$$f(\Delta \mid z_n, \cdots, z_1) \propto \sum_{M=1}^{n-1} \mathrm{p}(\mathrm{M} \mid z_n, \cdots, z_1)$$

$$\times \Phi(\Delta \mid z_n, \cdots, z_1;\ \mathrm{M}) \quad \cdots\cdots (8b)$$

식(1)의 모형에서 H_0: Δ=0 과 H_1: Δ≒0의 두가지 가설을 비교하여 예기치 않았던 수준변화의 감지가 가능하며, M=m인 경우, H_1과 H_0의 log-likelihood ratio의 p(m)의 가중평균 중 관찰치와 관계된 항목은 다음과 같다.

$$\frac{1}{2}\sum_{M=1}^{t-1} \mathrm{p}(\mathrm{M}) \frac{\lambda_n^2}{q_n} \quad \cdots\cdots (9)$$

3. 지수이동가중평균법에서의 수준 변화의 감지 및 적응

식(1)의 모형에서 Δ=0인 경우 $p_0=p^*$를 사용하면

$$f_t = f_{t-1} + (1-\alpha)e_{t-1};\quad \alpha = \frac{\sigma_e^2}{v^*}, \quad t=1, 2, \cdots, n \quad (10)$$

로서 Brown[3]이 제안한 지수이동가중평균법을 사용한 예측반복계산식이, 수준만을 포함한 구조모형의 최적예측 계산방법임을 알 수 있다(Harrison[4], Oliver[11] 참조). Brown[3]의 지수이동가중평균법을 사용할 때 일단계 미래값의 예측치는

$$f_t = (1-\alpha)\sum_{k=0}^{\infty} \alpha^k Z_{t-k}, \qquad t=1, 2, \cdots, n \quad (11)$$

로서 현시점까지 관찰한 값에 1, α, α^2, …의 지수적으로 감소하는 비중을 부여하여 그 가중평균이 예측치가 된다.

단기예측시 지수이동가중평균법을 적용하는 것이 적절한 상황하에서 전체적인 경제상황의 변화, 신제품의 개발 및 유입에 따른 수요량의 변화등이 수준변화로서 나타날 때, 수준변화가 일어난 것인지 아닌지를 감지하여, 그 변화시기를 판별하고

변화량을 추정하며, 그 변화에 적절히 적응할 수 있도록 자동예측 시스템을 설계하는 것은 중요한 문제 중의 하나이다. $p_0=p^*$를 사용하고 변화시기와 변화량에 대한 전문가의 의견을 구할 수 없을 때 즉, $p(m)=1/(n-1)$, $q_0\to\infty$경우 식(9)중 다음의 값만이 관찰치에 의존한다.

$$S_n=e_{n-1}^2 + \frac{(e_{n-2}+\alpha e_{n-1})^2}{1+\alpha^2}+\cdots + \frac{(e_1+\alpha e_2+\cdots+\alpha^{n-2}e_{n-1})^2}{1+\alpha^2+\cdots+\alpha^{2n-4}} \qquad (12)$$

따라서 지수이동가중평균법 적용시 수준변화가 일어났는지 아닌지의 감지는 식(12)의 통계치 S_n를 사용하여 판단할 수 있다. 즉 매 시기마다 수준변화가 발생할 수 있다고 가정하여, 그 시기부터 현재까지의 관찰된 값에 미래로 갈수록 지수감소하는 가중치를 예측오차들에 적용하여, 모든 시기에 관하여 합한것 으로서 감지한다. 수준변화가 발생하지 않은 경우 S_n은 작은값을 가지며, 관련된 시기중 어느 한 시기에서 수준변화가 발생하면 예측치가 그 변화에 적응할 때까지 지수감소하는 예측오차들의 지수이동가중평균값이 증가함으로서 S_n은 대단히 큰 값을 갖게되어 그 변화의 감지가 가능하다.

지수이동가중평균법은 시계열 자체가 random walk를 따르는 수준을 유지하고 있을 때 적용하면 minimum mean squared error forecast를 제공한다. 그러나 예측 하고자 하는 시계열에 예기치 않은 급격한 수준변화가 발생하면 지수이동가중평균법은 그 수준변화에 적응하는데 상당한 시간을 요구하게 되며 그 기간동안의 예측은 실제 관측치로부터 상당히 벗어나게 된다. 이러한 상황을 극복하기 위하여 다음의 방법을 제시한다.

$$f_t=f_{t-1}+\frac{S_t}{AS_t}e_{t-1} \quad \cdots\cdots\cdots\cdots\cdots\cdots (13)$$

여기서 AS_t는 식(12)의 S_t에 $e_1,\cdots,e_{t-1}$ 대신 그 절대값을 사용한 통계치이다. 따라서 $e_1,\cdots,e_{t-1}$의 부호가 모두 동일할 때를 제외하고는 S_t와 AS_t의 비율은 항상 0과 1사이에 존재한다. 지수이동가중평균법이 적절한 예측치를 제공하고 있을 때 예측오차들은 0에 근접한 양수와 음수의 적절한 조합이 될 것이므로 S_t/AS_t의 비율은 0에 가까와져서 현단계의 예측치가 예측오차보다 다음단계의 예측치에 지대한 영향을 미치게 된다. 반면 급격한 수준변화의 발생시 지수이동가중평균법에 의한 예측치들은 매우 큰 예측오차로부터 시작하여 예측오차들을 지수적으로 감소시키면서 새로운 수준에 적응하여 가기 때문에 S_t가 일시적으로 매우 큰 값을 띄게 되며 따라서 그 비율은 1에 가까워지고 결국 다음단계의 예측치는 현단계의 예측치보다는 최근발생한 예측오차에 의하여 좌우된다. 결국 지수이동가중평균법에서는 현 단계의 예측치와 예측오차를 고정된 비율로 다음 단계의 예측치에 반영하지만, 본 연구에서 제시된 방법은 급격한 수준변화의 발생시 그 수준변화를 감지할 수 있는 통계치를 사용하여 그 통계치의 증감정도가 직접적으로 예측오차의 반영에 영향을 줄 수 있도록 설계 되었다.

4. Trigg과 Leach의 방법과의 비교

지수이동가중평균법의 적용시 수준 및 기울기 변화에 적응하기 위하여 Trigg과 Leach[12]는 다음과 같은 방법을 사용하였다.

$$p_t=(1-\xi)z_t-f_{t-1}+\xi p_{t-1},\ 0<\xi<1, p_0 \text{ given}, \qquad (14a)$$

$$q_t=(1-\xi)\,|z_t-f_{t-1}|\,+\xi q_{t-1},\ q_0 \text{ given}, \qquad (14b)$$

$$k_t=\left|\frac{p_t}{q_t}\right|<1, \qquad (14c)$$

$$f_t=f_{t-1}+k_t(Z_t-f_{t-1}),\ f_0 \text{ given}. \qquad (14d)$$

비교를 위하여 Box와 Jenkins[1]의 Series A의 처음 100개에 $\sqrt{5}$를 곱한 자료가 사용되었다(표 1 참조). 그 중 처음 60개의 자료는 관련된 모수 추정을 위하여 사용하였으며, 그 결과 지수이동가중

(표 1) 비교에 사용된 시계열

1	38.013	26	29.355	51	39.355	76	38.013
2	37.119	27	38.908	52	37.790	77	37.790
3	36.448	28	38.684	53	37.342	78	38.013
4	36.001	29	38.013	54	37.566	79	37.119
5	38.237	30	39.802	55	37.566	80	37.342
6	37.790	31	39.131	56	38.460	81	37.566
7	37.566	32	40.473	57	37.566	82	37.342
8	38.908	33	39.131	58	39.355	83	36.672
9	38.237	34	38.908	59	38.460	84	36.895
10	38.013	35	38.908	60	37.119	85	36.672
11	37.342	36	38.237	61	38.237	86	37.119
12	38.908	37	39.355	62	37.790	87	36.895
13	38.460	38	39.578	63	37.119	88	37.342
14	38.908	39	38.908	64	40.249	89	36.672
15	38.908	40	39.802	65	38.460	90	36.672
16	38.013	41	39.355	66	38.684	91	36.224
17	38.684	42	39.131	67	38.013	92	36.672
18	38.460	43	36.895	68	37.790	93	36.448
19	38.908	44	39.802	69	38.684	94	36.672
20	37.566	45	38.684	70	37.566	95	38.013
21	38.237	46	38.684	71	38.684	96	37.790
22	38.908	47	38.237	72	38.908	97	38.237
23	38.908	48	38.908	73	39.578	98	38.237
24	39.131	49	37.790	74	37.566	99	37.342
25	38.908	50	38.684	75	37.790	100	37.790

평균법의 모수 $\alpha=0.775$로서 나타났고(상세한 내용은 Jun[8] 참조), Trigg과 Leach의 방법과 관련되어 $\xi=0.9$, $p_0=q_0=0.1$이 가장 적절한 것으로 판명되었다.

시계열에 수준변화가 발생한 경우의 결과를 비교하기 위하여 62번째부터 100번째까지의 싯점에 수준변화가 랜덤(random)하게 발생하게 하고 그 변화량은 서로 독립인 정규분포로부터 추출하여 실험하였다. 관련된 모수는 앞에서 추정한 값을 적용하였고 $f_{60}=37.6$을 임의로 사용하였다. 수준변화 발생횟수는 1회부터 9회까지, 변화량의 크기는 기대치 0인 정규분포를 따를 때 그 분산이 1, 10, 20인 경우에 각각 10회씩 실시하였다.

표 2의 결과는 수준변화 발생횟수 및 변화량의 분포가 주어졌을 때 Trigg과 Leach의 방법대비 본 연구에서 제시한 방법에서 얻어진 일단계 미래값의

(표 2) Trigg과 Leach의 방법 대비 본 연구에서 제시한 방법에 의한 일단계 미래값의 예측오차의 제곱들의 합의 10회 평균 비율

변화횟수 / 변화량	1	3	5	7	9
N(0,1)	0.905	0.941	1.011	0.952	0.979
N(0,10)	0.941	1.036	1.078	1.163	1.282
N(0,20)	0.950	1.076	1.231	1.271	1.190

(표 3) 지수 이동가중 평균법 대비 본 연구에서 제시한 방법에 의한 일단계 미래값의 예측오차의 제곱들의 합의 10회 평균 비율

변화횟수 / 변화량	1	3	5	7	9
N(0,1)	0.939	0.986	1.214	1.050	1.105
N(0,10)	1.318	1.582	1.603	1.861	2.016
N(0,20)	1.655	2.166	1.635	2.049	1.988

예측오차의 제곱들의 합(sum of the squared one-step-ahead forecast errors)의 10회 실시한 평균비율이다. 수준변화 발생횟수 또는 변화량이 증가함에 따라 본 연구에서 제시한 방법이 비교우위를 나타내고 있다.

표 3의 결과는 표 2와 동일한 경우에 지수이동가중평균법과 본 연구에서 제시한 방법에 의한 결과들의 평균 비율이다. 수준변화의 양이 매우 적을 때를 제외하고 상당한 적응효과가 있음을 볼 수 있다.

참고문헌

[1] Box, G.E.P. and Jenkins, G.M., *Time Series Analysis ; forecasting and control*, Holden-Day, Oakland, 1976

[2] Brown, R.G. and Meyer, R.F., "The fundamental theorem of exponential smoothing," Operations Research, vol. 9, pp. 673-685, 1961

[3] Brown, R.G., *Smoothing, Forecasting and Prediction of Discrete Time Series*, Prentice-Hall : New Jersey, 1962

[4] Harrison, P.J., "Exponential smoothing and short-term sales forecasting," Management Science, vol. 13, no. 11, pp. 821-842, 1967

[5] Harrison, P.J., and Stevens, C.F., "Bayesian forecasting," Journal of Royal Statistical Society(B), vol, 38, no.3, pp.205-247, 1976

[6] Harvey, A.C., "A unified view of statistical forecasting procedures," Journal of Forecasting, vol. 3, no.3, pp.245-275, 1984

[7] Holt, C.C., "Forecasting trends and seasonals by exponentially weighted moving averages," O.N.R. Memorandum, no.52, Carnegie Institute of Technology, 1957

[8] 전덕빈, "On detecting and estimating a major level or slope change in general exponential smoothing," Journal of Forceasting, vol.8, no.1, pp.55-64, 1989

[9] 전덕빈, and Oliver, R.M. "An adaptive response rate to monitor exponetial smoothing," Asian-Pacific Journal of Operational Research vol. 4, no.2, pp.112-121, 1987

[10] 전덕빈, and Oliver, R.M, "Bayesian forecasts following a major level change in exponential smoothing," Journal of Forecasting, vol. 4, no.3, pp.293-302, 1985

[11] Oliver, R.E., "Exponential Smoothing-I," Operations Research Center Report, UCB/ORC-84/5, Berkeley, 1984 : also Operations Research Letters, vol.3, no.3, pp.111-117, 1984

[12] Trigg, D.W. and Leach, A.G. "Exponential smoothing with an adaptive response," Operations Research Quarterly, vol. 18, pp.53-59, 1967