단기수요예측 분석방법론
- ARIMA모형 -

정 동 빈*

Methodology on Short-Term Demand Forecasting
- ARIMA model -

< 목차 >

개 요

본 논문에서는 단기 수요예측방법으로 ARIMA모형을 소개하며, 실전에서 이 모형을 설정하기 위해서 필요한 경험적인 전략에 초점을 두었다. 이 실전전략을 통해 프로세서 내에 복잡하게 얽혀있는 메커니즘을 단순용이하게 설명 및 규명할 수 있으며, 더 나아가 기업의 판매계획, 생산계획, 인사계획, 재무계획 등에 필요한 의사결정을 하는데 도움이 될 수 있다.

주제어 : ARIMA모형, 가역성, 단기수요예측, 이동평균구조, 자기회귀구조

---

* 강릉대학교 자연과학대학 정보통계학전공
접수일자 : 2003-4-1 게재확정일자 : 2003-10-15

I. 서 론

모든 기업(제조업체 또는 구매/유통업체)은 수요예측을 통해 생산계획 및 구매계획을 설정한다. 수요예측을 하는 근본적인 이유는 제품을 만드는데 소요되는 시간이 길기 때문이다. 고객으로부터 주문을 받고 나서 생산 또는 구매 활동을 할 수도 있지만, 수요예측을 통해 제조 리드타임과 유통 리드타임을 줄임으로써 고객이 제품을 필요로 하는 시점보다 일찍 생산을 시작해야 할 것이다. 수요예측에서 고려할 수 있는 사항은, 무엇을 언제까지 생산(또는 구입)해야 할까? 계획을 세운다면, 무엇에 근거하는 것일까? 고려하는 집단이 또 다른 특성을 가진 집단과 어떻게 비교될 수 있는가? 앞으로의 수요가 어떤 방향으로 변화할 것인가? 등을 들 수 있다.

수요예측방법은 정성적(qualitative)방법과 정량적(quantatitive)방법으로 구분할 수 있다. 정성적인 방법은 과거자료가 없거나, 예측대상이 아직 도입되지 않거나, 수리적 모델링이 불가능한 상황에서 사용되며, 이 경우 수요자의 의향 및 태도를 파악하는 설문조사방법 또는 전문가들의 지식과 의견에 따라 예측하는 델파이(delphi) 방법이 보편적으로 적용된다. 반면에 정량적인 방법은 크게 시계열분석(time series analysis)방법과 인과분석(causal analysis) 방법이 있다. 시계열분석방법이란 과거자료가 존재할 경우 변수 하나를 선정한 후에, 해당 변수의 과거 자료를 근거로 해당 변수의 미래값을 예측하는 방법이다. 반면에 인과분석은 어떤 변수의 값이 다른 변수들에 의해 영향을 받아 결정될 때, 다른 변수들의 과거값과 해당 변수의 관계를 모델링하여 원하는 변수의 미래값을 추정하는 방법이다. 본 논문에서는 여러 시계열 예측방법들 중 단기예측에 많이 적용하는 모형중 하나인 ARIMA모형을 살펴보고, 실현값에 근거하여 적절한 ARIMA모형을 구축하기 위해 필요한 경험적인 실전전략을 피력하기로 한다. 더 나아가 단순화 전략을 통해 계절적 요소와 비계절적 요소에 각각 포함된 자기회귀구조와 이동평균구조를 보다 용이하게 파악하여, 프로세서 내의 존재하는 메커니즘을 이해하여 미래수요를 예측하는데 그 의의가 있다.

2장에서는 경제시계열의 전통적인 분석방법으로 단기 수요예측모형으로 사용할 수 있는 ARIMA모형에 대해 소개한다. 3장에서는 실제 자료에 ARIMA모형을 적합시킬 때 적절한 모형을 구축하는 단순화 전략에 대해 살펴보도록 한다. 이 실전전략을 적용하여 기업의 판매계획, 생산계획, 인사계획, 재무계획 등에서 의사결정을 할 수 있다. 끝으로 4장에서 결론을 내린다.

II. ARIMA모형

대부분의 경제 및 경영분야의 실현값은 계절적 패턴 및 비계절적 패턴을 동시에 포함하고 있다. 물론 이 두 가지 요소를 논리적이면서도 시각적으로 분리할 수 있다. 즉, ARIMA (p, d, q)

$(P, D, Q)s$모형을 일반화하여 다음과 같이 표현할 수 있다.

$$\phi_p(B)\,\Phi_P(B^s)\,\nabla^d \nabla_s^D \tilde{y}_t = \Theta_Q(B^s)\,\theta_q(B)\,e_t$$

여기에서,

$$\Phi_P(B^s) = (1 - \Phi_s B^s - \Phi_{2s} B^{2s} - \cdots - \Phi_{Ps} B^{Ps}),$$

$$\nabla_s^D = (1 - B^s)^D,$$

$$\Theta_Q(B^s) = (1 - \Theta_s B^s - \Theta_{2s} B^{2s} - \cdots - \Theta_{Qs} B^{Qs}),$$

$$\nabla^d = (1 - B)^d,$$

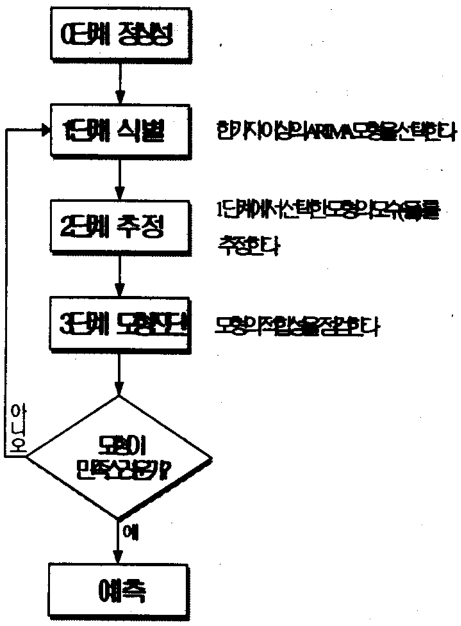
$$\phi(B) = (1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p),$$

$$\theta(B) = (1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q)$$

이다. 참고로 (p, d, q)는 비계절성 차수, $(P, D, Q)_s$는 계절성 차수, e_t는 백색잡음(white noise)을 의미한다.

III. 실전분석전략

Box와 Jenkins는 좋은 모형을 찾기 위해 실전적인 3단계 절차를 제안하였다. 다음은 ARIMA 모형을 설정을 위해 필요한 절차를 도표로 나타내고, 각 단계별로 요약하였다.



0단계 : 계열의 평균과 분산을 정상적으로 만든다.

일변량 ARIMA모형은 정상적인 자료계열에만 적용하므로 비정상적인 계열의 분산과 평균에 대하여 각각 변수변환 및 차분을 행하여 계열의 평균과 분산을 정상적으로 만든다. 즉, d와 D를 결정한다. 정상성(stationarity)은 확률과정에서 n개의 시계열 $y_{t_1+h}, y_{t_2+h}, \cdots, y_{t_n+h}$의 결합확률분포 $P(y_{t_1+h}, y_{t_2+h}, \cdots, y_{t_n+h})$이 $t_1, t_2, \cdots, t_n$ 및 h (임의의 상수)의 어떠한 유한집합에 대해서도 h와 관계없이 동일하다는 성질을 의미한다.

1단계 : 식별단계

식별단계에서는 일변량 자료계열 내 관측

값들 사이의 상관관계를 측정하는 두 개의 도구(자기상관함수, 편자기상관함수)를 이용한다. 이 추정된 두 상관함수는 덜 효율적인 방법이지만 자료계열 내의 존재하는 통계적 관계를 측정하여 자료에 대한 패턴을 이해한다.

다음 단계는 두 개의 측정된 상관함수를 통해 되도록 간단히 자료계열 내에 존재하는 통계적 관계를 요약한다. 두 상관함수를 이용하여 적절해 보이는 하나 이상의 ARIMA모형을 고려한다. 이때 자료로부터 계산된 추정된 두 상관함수와 이론적인 두 상관함수를 각각 비교하여 이론적인 자기상관함수와 편자기상관함수 형태와 비슷한 모형(p, P, q, Q)을 임시로 선택한다. 식별단계에서 선택한 모형은 다만 임시적으로 고른 것이며 최종모형을 설정하기 위한 후보에 지나지 않는다. 최종 모형을 선택하기 위해서 다음 두 단계를 거쳐야 한다.

2단계 : 추정단계

이 단계에서는 식별단계에서 선택한 모형의 계수(들)를 정확히 추정하며, 모형적합성에 관한 몇 가지 경고신호를 탐지할 수 있다. 특히 추정된 계수들이 주어진 수학적인 부등식 조건을 만족하지 못하면 그 모형은 기각된다.

3단계 : 모형진단단계

추정된 모형이 통계적으로 적절한지를 결정하는 단계이며, 이 진단검증에 부합하는 모형은 기각된다. 이 단계에서 나타난 결과들을 통해 고려한 모형이 얼마나 좋은지를 알 수 있다. 최종모형을 찾을 때까지 식별, 추정, 모형진단의 3단계를 되풀이하여 반복한다. 특히 추정 및 모형진단 단계를 통해 언제, 어떻게 다른 후보 모형(들)을 고려할지를 알 수 있다. 몇 가지 기준들을 만족하는 모형을 찾을 때까지 재식별, 재추정, 재진단을 계속 행한다. 이 3단계를 반복적으로 적용하더라도 가장 좋은 ARIMA모형을 찾는다는 보장은 없다.

4단계 : 예측단계

추정된 ARIMA모형에 근거하여 예측시점(들)에 대해 예측한다.

앞에서 논의한 ARIMA모형을 비교적 쉽게 구축하기 위해 필요한 경험적인 31개의 실전전략을 알아보기로 한다. 다음은 일변량 ARIMA모형을 설정하기 위한 실전전략이다.

(1) 모든 자료를 분석하기 전에 그 정확성을 점검해야 한다.

(2) 이상적인 ARIMA모형을 설정하기 위해서는 최소 50개 이상의 관측값이 필요하다. 차분으로 말미암아 관측값(들)이 손실되더라도 적절한 식별과 추정을 위해 충분한 자유도가 필요하기 때문이다. 특히 계절적 변동이 존재할 때 50개 이상의 관측값이 필요한 것은 당연하다.

(3) 최초 실현값에 대한 시도표를 그려 시각적으로 조사하는 것은 중요한 선결단계이다. 이는 실현값 분산의 정상성 여부를 판정하는데 매우 중요하다. 몇 가지 통계적 검증방법을 사용할 수 있지만, 시도표를 이용하여

실현값을 조사하는 것은 다른 통계적 검증만큼 유용하다(Granger와 Newbold, 1977). 변수의 대수변환은 경제학과 경영학에서 다뤄지는 자료에 흔히 사용된다. 이 변환은 실현값의 표준편차가 평균에 비례할 때 사용된다. 만일 대수변환이 적절하지 않다고 생각될 때는 Box-Cox 멱변환 $(y_t^{\lambda}-1)/\lambda$을 사용한다. 계열수준이 증가함에 따라 분산이 증가하면 $\lambda<1$, 반대로 분산이 감소하면 $\lambda>1$을 만족하는 λ값이 필요하다.

(4) 모형설정에 앞서 시도표는 실현값 평균의 정상성 여부를 감지하는데 도움을 줄 수 있다. 만약 계열의 평균이 시간에 따라 변한다면, 추정된 자기상관함수를 통해 정상성 여부를 재차 확인할 수 있다. 추정된 자기상관계수 (r_k)는 시계열에서 k시차만큼 떨어져 있는 관측값과의 상관계수이다.

(5) 또한 시도표는 실현값에 대한 계절적 패턴의 존재여부(명백한 계절적 변동, 희미한 계절성, 비계절적 패턴 중 하나)를 아는데 도움이 된다. 자기상관분석과 추정결과를 통해 모형 내에 계절적 요소를 포함시킬지 여부를 최종 결정한다. 그러나, 시도표를 통한 선결조사는 계절적 패턴의 존재여부를 탐지하는데 보충정보를 줄 수 있다.

(6) 자료분석에 필요한 추정된 자기상관함수의 개수는 약 "$n/4$"이다. 즉, 관측값 개수의 25%이다.

(7) 경제학과 경영학에서 다루는 대부분의 시계열은 계절적 패턴을 갖는다. 설사 계절성을 없애기 위해 자료를 조정하더라도 여전히 계절적 패턴이 존재하기도 한다. 이 때문에 항상 계절성에 대한 근거를 염두해둔다.

(8) 평균을 정상적으로 만들기 위해서는 실현값을 차분해야 한다. 다음은 정상성을 점검할 수 있는 도구를 소개하기로 한다.

(a) 시도표(timeplot)

이는 차분여부를 결정하는 유일한 척도가 아니라 아래에서 소개할 다른 방법들과 더불어 정상성의 단서를 찾는데 도움을 준다. 시도표를 통해서 수준의 뚜렷한 변화(강한 상승경향 또는 강한 하강경향) 즉, 눈에 띄는 기울기가 발견되면 차분해야 된다는 징조다. 두드러진 수준의 변화가 탐지되면 비계절형 1차 차분을, 기울기가 변화하면 비계절형 2차 차분을 취해야 한다. 일반적으로, 강한 계절적 변동이 감지되면 계절적 1차 차분만을 한다.

(b) 추정된 자기상관함수

만일 추정된 자기상관함수가 긴 시차에 걸쳐서 서서히 감소할 때, 자료의 평균은 비정상적이므로 자료를 차분해야 된다. 이런 사항은 계절적 시차에 해당하는 자기상관함수의 경우에도 별도로 적용한다. 예를 들면, 분기별 자료에는 4의 배수, 월별자료에는 12의 배수에 해당하는 자기상관함수를 고려한다. 즉, 계절적 차분을 고려할 때 계절성의 주기의 배수에만 해당하는 자기상관계수(들)를 조사하여 일시적으로 계절적 시차를 별도의 구조로써 다룬다.

(c) 추정된 자기회귀계수의 부호와 크기

추정된 자기회귀계수의 부호와 크기는 정상성 조건을 점검하는 또 다른 도구이다. 승법 자기회귀계절형 모형을 추정할 때 비계절적 자기회귀계수와 계절적 자기회귀계수에 이 조건들을 별도로 적용한다. 추정된 자기상관함수를 통해 만일 평균이 비정상적이라는 근거가 명백하다면, 비차분된 자료에 대한 자기회귀계수를 추정하지 않고 차분을 행한다(다음 쪽 표를 참고).

(9) 만일 자료가 물리적 단위나 기준시점 화폐단위라기보다는 현 시세 화폐단위로 측정되면 비계절적 1차 차분(d=1)이 더 자주 필요하다. 비계절적 2차 차분(d=2)은 상대적으로 흔히 사용하지 않는다.

[정상성조건과 가역성 조건]

모형	정상성 조건	가역성 조건
단순 AR	p에 의존	항상 가역적
단순 MA	항상 정상적	q에 의존
AR(1)	$\vert\phi_1\vert < 1$	항상 가역적
AR(2)	$\vert\phi_2\vert < 1$ $\phi_2 + \phi_1 < 1$ $\phi_2 - \phi_1 < 1$	항상 가역적
MA(1)	항상 정상적	$\vert\theta_1\vert < 1$
MA(2)	항상 정상적	$\vert\theta_1\vert < 1$ $\theta_2 + \theta_1 < 1$ $\theta_2 - \theta_1 < 1$
ARMA(p,q)	p에 의존	q에 의존

(10) 계절적 차분은 실제 상황에서 자주 하게 되며, 만일 계절적 차분을 하게 될 때 보통 한번(D=1)만 하면 충분하다.

(11) 불필요한 차분을 하면 덜 만족스런 모형을 만들기 때문에 이를 피해야 한다. 필요 이상의 차분을 하면, 자료 내에 인위적인 패턴이 만들어지며 예측오차분산은 불필요하게 커지게 된다. 그러나, 의심스러우면 차분을 해야 하며 판단하기 애매한 경우에 차분하게 되면 뜻밖에 좋은 예측값을 얻을 수 있다. 왜냐하면 예측값은 고정된 평균에 얽매이지 않기 때문이다.

(12) 추정된 모수들의 상관행렬을 통해 평균의 정상성을 판별할 수 있다. 즉, 추정된 평균으로 관측값 평균보다는 오히려 자기상관계수(들)와 이동평균계수(들)를 동시에 고려

한 평균을 추정한다. 몇 개의 추정된 모형에 근거하여 추정된 자기회귀계수(들)가 정상성 조건을 만족하지만, 추정된 하나 이상의 자기회귀계수(들)는 추정된 평균과 크게 상관될 수 있다 (|ㅋ>0.9)). 이런 경우 평균의 추정값은 정확할 수 없으며, 가장 현명한 치유책은 자료를 차분하는 것이다.

(13) 보통 좋은 모형은 평범하며 모수의 수를 절약하는 원칙에 근거하여 얻을 수 있다. 경제학과 경영학에서 다뤄지는 대부분의 자료는 적절한 차분 후에 차수가 2 또는 그 이하의 ARMA모형으로 적절히 표현될 수 있다. 일반적으로 차수가 1인 모형이 계절적 부분에 적절하다. 평범한 모형과 모수절약의 원칙을 염두에 두고, 식별단계에서 비계절적 시차 1과 2, 첫 두 개의 계절적 시차들(s, $2s$)을 제외한 시차에 해당하는 추정된 자기상관함수가 유의하지만 잔차 자기상관함수는 유의하지 않을 수 있다. 차수가 1 또는 2인 모형을 이용하여 만족할 만한 모형을 찾지 못하면, 차수가 큰(high-order)모형을 시도한다.

(14) 먼저 계절적 요소를 식별 및 추정한 후 비계절적 패턴에 관한 잔차 자기상관함수를 조사하면 때때로 더 좋은 결과들은 얻게 된다. 이는 원자료에 근거하여 추정된 자기상관함수에서 두드러진 계절성을 발견할 때 특히 유효하다. 즉, 계절적 요소와 비계절적 요소가 혼합된 ARIMA모형을 설정하는데 어려운 상황에서는, 계절적 요소(들)만을 추정하는 전략을 채택하고 난 뒤, 계절적 요소만 포함된 모형에 근거한 잔차 자기상관함수와 편자기상관함수를 이용하여 비계절적 부분을 식별한다. 이 전략은 특히 계절적인 주기가 작은 시계열(예 : 반기별, 분기별) 자료를 분석하는데 도움이 된다.

(15) 계절적 패턴은 흔히 원자료에 근거한 자기상관함수보다는 비계절적 1차 차분 후 계산된 자기상관함수에서 더 명확히 나타나기 때문에, 차분을 통해 계열의 평균을 정상적으로 만들 필요가 없다고 할지라도 비계절적 1차 차분 후에 추정된 자기상관함수를 조사하는 것이 좋다.

(16) 때때로 강한 계절성 때문에 적절한 일변량 ARIMA모형을 설정하는데 어려움이 있다. 우선 자료의 계절성을 없애고 계절성이 제거된 계열에 대해 일변량 ARIMA모형을 설정한 후, 계절적으로 조정된 요소를 계절성이 제거된 예측값들에 적용할 수는 있다. 그러나, 계절적인 패턴을 ARIMA모형에 통합하는 것이 차라리 훨씬 좋은 방법이다.

(17) 승법 계절형 모형은 일반적으로 만족할 만한 결과를 얻는다. 의심스러울 때 승법형태와 가법형태의 결과를 비교해야만 한다. 승법형태가 갖는 장점은 정상성 조건과 가역성의 조건을 보다 용이하게 점검할 수 있다는 점이다. 왜냐하면 이 두 조건을 모형의 계절적 부분과 비계절적 부분에 개별적으로 적용할 수 있기 때문이다.

(18) 경영학 또는 경제학에서 나타나는 대부분의 계절적인 프로세서는 AR모형 또는 혼합 ARMA모형이다. 이때, 계절적 시차에 해당하는 자기회귀계수가 "1.0"이면 계절적

차분을 해야 하는 상황이 자주 나타난다.

(19) 정상적 자기회귀프로세서가 갖는 이론적인 자기상관함수는 지수적으로 감소하거나 진폭이 감소하는 사인곡선형태이다. 반면에 이론적 편자기상관함수는 두드러진 스파이크 다음에 "0"으로 절단되는 형태이며, 이때 두드러진 스파이크를 갖는 마지막 시차가 자기회귀모형의 차수와 같게 된다.

(20) 이동평균프로세서를 나타내는 이론적인 자기상관함수는 두드러진 스파이크 다음에 "0"으로 절단되는 형태이며, 스파이크를 갖는 마지막 시차가 이동평균모형의 차수와 같게 된다. 이론적인 편자기상관함수는 자기회귀모형이 갖는 자기상관함수와 마찬가지로 지수적으로 감소하거나 또는 진폭이 감소하는 사인곡선형태를 갖는다.

(21) 정상적 ARMA프로세서를 나타내는 이론적 자기상관함수와 편자기상관함수는 모두 "0"으로 점차 감소하는 형태이다. 자기상관함수는 첫 $(q-p)$시차 뒤부터 지수적으로 감소하거나 진폭이 감소하는 사인곡선형태이며, 편자기상관함수는 첫 $(p-q)$시차 후부터 감소한다.

(22) 계절적 자기회귀프로세서나 계절적 이동평균프로세서는 [실전전략 19-21]에서 언급한 비계절적 프로세서와 똑같은 자기상관함수와 편자기상관함수의 특징을 갖는다. 유일한 차이점은 계절적 프로세서에 대한 자기상관함수와 편자기상관함수는 계절적 주기의 배수 시차들만을 고려한다는 점이다. 예를 들면, 사분기별 자료의 정상적 계절형 자기회귀프로세서의 경우, 시차 4, 8, 12,...에 해당하는 이론적 자기상관함수는 지수적으로 감소하는 형태이다. 반면에 분기별 자료의 계절적 이동평균프로세서의 경우, 이론적 자기상관함수는 시차 4에만 뚜렷한 스파이크를 갖고 나머지 4의 배수 시차 8, 12, 16, ...에서는 "0"의 값을 갖는 형태이다.

(23) 때때로 초기 식별단계에서 모수가 절약된 혼합모형을 식별하기 어렵다. 종종 자기회귀모형을 우선적으로 추정하고, 잔차 자기상관함수에 근거하여 관련된 계수들을 덧붙이는 방법이 효과적이다. 이렇게 함으로써 모수가 절약된 모형을 찾는데 더 용이하며, 불필요한 계수를 모형에 삽입하는 것을 피할 수 있다.

(24) 모형의 차수는 총 계수의 수가 아닌 추정할 계수의 최대 수로 결정된다. 예를 들면, 항상 MA(2)모형은 항상 계수 θ_2를 포함하지만, 계수 θ_1을 모형에 포함하거나 하지 않을 수도 있다. 이에 관련된 두 개의 MA모형은 다음과 같다(단, C는 상수항).

$$y_t = C + (1 - \theta_1 B - \theta_2 B^2)e_t \text{와}$$
$$y_t = C + (1 - \theta_2 B^2)e_t$$

(25) 추정단계에서 추정된 계수의 t절대값이 "2.0"보다 훨씬 작을 때 이 계수를 모형에 포함시키지 않으며, 특히 시차 1과 2 또는 계절적 시차를 제외한 시차들에 해당하는 계수들에는 더욱 주의를 기울여야 한다.

(26) 차분한 시계열의 평균이 "0"이 아니며 이 시계열이 결정적 추세일 때만, 상수항을

차분한 자료에 포함시킨다. 사회과학에서 다뤄지는 자료는 흔히 적절한 차분 후에 통계적으로 "0"인 평균을 갖는다. 평균이 "0"이면 모형 안의 상수항은 0이 되며, 예측값에 대한 추세는 단순히 확률적이 된다. 자기회귀계수 및 이동평균계수와 동시에 평균을 추정할 때, 추정된 평균과 상수의 통계적 유의성을 점검 후에 경험에 입각하여 결정을 내려야 한다. 결정적 추세는 해석이 가능하다.

(27) 다음 쪽에 소개된 표는 추정된 자기상관함수의 t절대값에 대한 실전적 경고수준을 요약한 것이다.

(a) 식별단계에서 t절대값이 대략 "1.6"을 초과하는 비계절적 자기상관계수에 주목한다. 이들 시차에 해당하는 계수들은 추정단계에 거의 통계적으로 유의하다.

(b) 식별단계와 모형검진단계에서 t절대값이 대략 "1.25"를 초과하는 계절적 자기상관계수에 주목한다. 이에 대응하는 추정된 계절형 자기회귀계수 또는 계절형 이동평균계수는 추정단계에서 거의 통계적으로 매우 유의하다. 만일 계절적 시차(s, $2s$, ...)에 대응하는 잔차 자기상관함수가 통계적으로 "0"이면, 계절적 주기의 반에 해당하는 시차($0.5s$, $1.5s$, ...)와 계절적 시차들에 근접하는 시차($s+1$, $s-1$, $2s+1$, $2s-1$, ...)의 잔차 자기상관함수에 해당하는 t절대값이 "1.25"를 초과하는지를 주목하여 앞에서와 같이 이를 적용한다.

(c) 모형검진단계에서 단기시차(1, 2, 3)에 대응하는 잔차 자기상관계수의 t절대값이 "1.25"를 초과하면, 이 단기시차(1, 2, 3)에 대응하는 계수를 추정한다. 단기시차에 대응하는 잔차 자기상관계수의 t값은 "Bartlett 근사(Bartlett, 1946)"때문에 때때로 과소평가되기도 한다.

(d) 앞 (a)(b)(c)경우 편자기상관함수에 대응하는 t절대값의 실전적 경고수준은 "2.0"이다.

(28) 잔차 자기상관함수에 근거하여 모형을 재설정할 때 모수가 더 절약된 모형을 얻기 위해 계수를 한번에 하나씩 모형에 포함시킨다. 이렇게 함으로써 불필요한 계수를 모형에 삽입시키는 문제를 피할 수 있다. 때때로 단지 하나의 계수를 모형에 포함시켜 복잡한 잔차 자기상관함수를 쉽게 해석할 수 있다.

(29) 최종 ARIMA모형을 설정한 후 잔차들을 플롯한다. 이는 잘못 기재된 자료를 찾을 수 있는 단서가 되며, 또한 자료에 내재된 변동의 원인을 더 자세하게 파악할 수 있기 때문이다.

[추정된 자기상관함수에 대응하는 t절대값의 실전적 경고수준]

자기상관함수 시차	식별단계 (조기 추정된 자기상관함수)	검진단계 (잔차 자기상관함수)
단기시차(1, 2, 3)	1.6	1.25
계절적 시차(s, 2s, ...)	1.25	1.25
계절적 주기에 근접한 시차들 (s−1, s+1, 2s−1, 2s+1, ...)**와 계절적 주기의 반에 해당하는 시차들** (0.5s, 1.5s, ...)	−	1.25
기타	1.6	1.6

유의사항 계절적 차분을 하거나 모형에 계절적 자기회귀계수 또는 이동평균계수를 포함시킨 후 계절적 시차(s, 2s, ...)에 대응하는 자기상관함수가 통계적으로 유의하지 않을 때만 계절적 시차에 근접한 시차(들) 또는 계절적 주기의 반에 해당하는 시차(들)에 주의를 기울인다.

(30) 추정된 계수간의 상관계수 절대값이 “0.9”이상이면, 이 ARIMA모형을 채택하지 않는 것이 좋다. 이런 경우 추정된 계수들이 불안전한 경향이 있다. 즉, 추정된 계수(들)가 특정적인 실현값에 크게 영향을 받으며, 자료 패턴에 약간의 변화만 있더라도 계수의 추정값이 크게 변한다.

(31) 모형의 최종점검은 예측력(ability to forecast)이다. 평범하고 모수가 절약된 모형을 선호하지만, 그렇지 않은 모형이 더 좋은 예측값을 산출한다면 이 모형을 선택해야 할 것이다.

Ⅳ. 결 론

본 논문에서는 시계열자료에 대한 단기 수요예측방법으로 ARIMA모형을 소개하였고, 경험적인 접근방법으로 31개의 실전전략을 제시하였다. 이 전략을 실제 시계열자료에 적용시켜 보다 효율적이며 단순화된 즉, 최적화된 ARIMA 모형을 구축할 수 있으며, 기업의 판매계획, 생산계획, 인사계획, 재무계획 등에 필요한 의사결정을 하는데 도움이 될 수 있다.

모델링을 함에 있어서 잡음과 시그널을 구분하는 일은 최적화의 문제와 깊이 관련되어 있으며, 이를 설명력이 있

고 정확도 있는 모형으로 표현하는 일은 분명 단순한 작업은 아닐 것이다. 무엇보다도 이러한 배경에서 충실히 이론적인 면을 적절히 결부시켜 많은 시행착오를 경험하고 직관적인 분석능력을 배양해야 할 것이다. 최종모형 선정에 대한 가장 중요한 척도는 예측력과 모수가 절약된 모형이 될 것이다.

참 고 문 헌

정동빈과 원태연, SPSS를 활용한 시계 열자료와 단순화분석 I, SPSS 아카데미, 2001.

정동빈과 원태연, SPSS를 활용한 시계 열자료와 단순화분석 II, SPSS 아카데미, 2003.

Bartlett, M.S., "On the theoretical specification of sampling properties of autocorrelated time series", *Journal of the Royal Statistical Society*, B8, 27, 1946.

Bowerman, B.L and O'Connell, R.T., *Forecasting and time sereis*, 3rd. ed, Duxbury Press,1993.

Box, G.E.P., Jenkins, G. M. and Reinsel.l G. C. *Time series analysis: forecasting and control*, 3rd. ed., NJ. Prentice Hall, 1994.

Cleary, J. P. and Levenbach, H., *The professional forecaster*, Bel mont., Calif., Lifetime Learning Publications, 1982.

Granger, C.W.J. and Newbold, P., *Forecasting Economic Time Series*, New York, Academic Press, 1977.

McCleary, R. and Hay, R.A., *Practical experiences with modelling and forecasting time series*, Jersey, Gwilym Jenkins and Partners, 1979.

McLeod, G., *Box Jenkins in Practice*, Gwilym Jenkins & Partners Ltd, 1983.

ABSTRACT

Methodology on Short-Term Demand Forecasting
- ARIMA model -

Jeong, Dong-Bin*

In this study, ARIMA model is introduced, which is one of the widely used short-term demand forecasting methods and empirical strategies which can be applied to the real fields are focused. We can easily and simply explain and/or diagnose the mechanism of process and furthermore obtain the important information to make decisions by applying these 31 strategies.

$$\phi_p(B)\Phi_P(B^s)\nabla^d\nabla_s^D\tilde{y}_t = \Theta_Q(B^s)\theta_q(B)e_t$$

where

$$\begin{aligned}
\Phi_P(B^s) &= (1 - \Phi_s B^s - \Phi_{2s}B^{2s} - \cdots - \Phi_{Ps}B^{Ps}),\\
\nabla_s^D &= (1 - B^s)^D,\\
\Theta_Q(B^s) &= (1 - \Theta_s B^s - \Theta_{2s}B^{2s} - \cdots - \Theta_{Qs}B^{Qs}),\\
\nabla^d &= (1 - B)^d,\\
\phi(B) &= (1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p),\\
\theta(B) &= (1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q).
\end{aligned}$$

Keywords : *ARIMA, Autoregressive, Demand Forecasting, Invertible, Moving Average, Stationary*

* Professor, Division of Information Statistics, kangnung National University