

서울지방 겨울철 기온의 확률모델

김해경* · 김태수**

요 약

본 논문의 목적은 서울지방 겨울기온의 예측을 위한 확률모델의 개발과 그 응용에 있다. 겨울기온의 회귀추세, 주기성 그리고 증속성들의 연중, 연간변동을 과거 30년(1959~1989) 일일자료를 기초로 하여 분석하였다. 기온예측을 위한 확률모델을 개발하고 그 응용을 위한 통계적 절차를 제안하였다. 겨울철기온의 특성인 이상기온의 출현과 삼한사온현상의 실체도 논하였다.

1. 서 론

겨울철의 가장 주목되는 氣象현상은 氣溫의 변화이다. 기온의 변화는 일상생활은 물론 산업 특히 농업과 경제활동에 많은 영향을 주게되며, 이로인한 피해를 최소화 하기 위한 정확한 예측방법이 요구된다.

기온은 늘 변하면서도 어떤 패턴을 가지고 있다. 이런 패턴의 존재는 기온이 현재는 과거에, 미래는 과거와 현재에 의하여 從屬되어 있고, 이 從屬性이 시간에 따라 일정한 변화를 하고 있음을 의미한다. 더욱이, 다양한 기상요인과 복잡한 물리현상들로 결정되는 이 패턴은 몇가지 간단한 물리적 법칙으로는 만족스럽게 설명되지 않으며, 오히려 같은 조건하에서 동일한 결과가 주어 지지는 않지만 確率현상으로 기술하는 것이 더 합리적이다.

본 연구의 목적은 우리나라 겨울철 기온현상의 確率 및 統計的 특성을 파악하고 겨울기온의 분석을 위한 確率모델을 확립하는데 있다. 이를 위해, 먼저 서울지방기온의 回歸趨勢, 週期性 및 從屬性 등 統計的 특성을 파악하고, 다음으로 과거 30년을 통하여 그 특성들이 어떻게 변화되고 있는가를 분석하였다. 마지막으로 이 결과들을 토대로하여 기온분석을 위한 確率모델을 완성하고, 그 適合度를 확인하였다. 유도된 確率모델을 이용한 기온예측의 統計的 절차도 제안하고 그 實例를 들었다. 또한, 30년 자료분석을 기초로 하여 겨울철 異常氣溫의 출현과 三寒四溫현상의 실체도 論하였다.

* (120-749) 서울특별시 서대문구 신촌동 134 연세대학교 이과대학 수학과 교수

** (120-749) 서울특별시 서대문구 신촌동 134 연세대학교 이과대학 수학과 박사과정

2. 자료와 분석

본 연구에 이용된 자료는 서울지역에서 중앙기상청이 1959부터 1989사이에 관측한 30년간의 겨울철(11월 1일부터 2월 28일까지) 일별기온이며, 특히 겨울철 기온의 대표로 最低기온을 사용하였다. 그림 2-1은 과거 30년 자료중 일부를 나타낸 것이다.

자료의 분석은 그림 2-1에서 관찰되는 외관상의 특징으로부터 시작된다. 먼저 겨울철 기온은 전반부부터 점차 하강하여 1월중순경에는 최저상태를 보이고, 그후 다시 점차 상승하는 어떤 趨勢를 내포하며, 동시에 기온은 趨勢주위를 상하진동하며 어떤 일정기간을 두고 반복되는 週期성이 나타난다. 또한 기온의 변화는 연속변화로서 순간적 跳躍을 갖지 않기 때문에 기온 상호간의 從屬性의 존재가 가정될 수 있다. 이와같은 관점에서 볼 때, 겨울철 기온의 관찰값 X_t (t 는 시간)는 다음식으로 가정된다.

$$X_t = D_t + C_t + S_t + E_t$$

여기에서, D_t 는 回歸趨勢성분, C_t 는 週期성분, S_t 는 確率성분 그리고 E_t 는 측정오차를 포함하는 確率誤差성분이다.

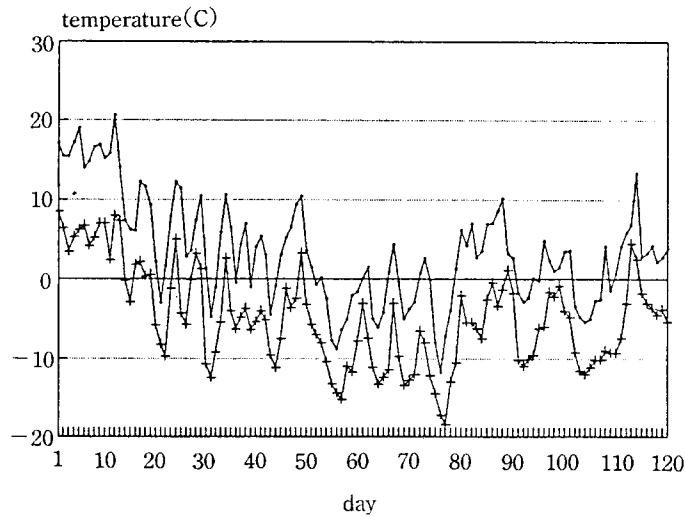


그림 2-1. 일별기온시계열 : 1966. 11-1967. 2. (최저기온(-+-); 최고기온(-.-)).

그러나, 이 가정식에서 확률성분 S_t 와 측정오차 E_t 의 합, $Z_t = S_t + e_t$ 은 보통 定常 또는 非定常 ARMA(p, q) 모델

$$\sum_{j=0}^p a_j Z_{t-j} = \sum_{k=0}^q b_k e_t \quad (p, q \text{는 양의 정수}) \quad (2.1)$$

로 설명된다. 여기서 a_j 와 b_k 는 적당한 상수(단 $a_1=b_1=1$)이고 e_t 는 (正規) 白色誤差이다. 따라서 관찰값 X_t 의 가정식은 (2. 1)을 만족하는 Z_t 를 이용하여

$$X_t = D_t + C_t + Z_t \tag{2. 2}$$

로 표시된다.

다음 부분節에서는, 가정된 모델 (2. 2)에 대하여 각 성분의 년별 고유의 최적함수 형태를 결정하고, 시간에 따른 특성의 변화를 30년 자료를 통하여 관찰한다. 특히, 趨勢성분 D_t 의 분석은 시계열 線型回歸분석을 통하여, 성분 C_t 에 대한 週期함수식은 週期圖분석과 非線型回歸분석에 의하여 결정된다. 그리고 확률성분 Z_t 의 결정에는 Box & Jenkins방법[4]을 이용한다. 특히, 1966. 11~1967. 2 겨울철 最低기온이 분석의 數的 事例로 이용된다.

a) 回歸趨勢의 분석

回歸趨勢성분 D_t 의 분석은 주어진 자료 $\{X_t\}$, $t=0, 1, \dots, T-1$,를 가장 적절히 맞추어 주는 다항식을 찾는 데 있다. 직관적으로 볼 때, 적당한 함수식은 1 또는 2차식으로 가정될 수 있고, 가장 보편화된 최소제곱법을 이용하여 미지수를 결정하고 그 타당성을 檢定할 수 있다. 그러나 이 경우 線型回歸모델의 確率誤差項에 해당하는 $C_t + Z_t$ (간단히 U_t 로 표시)가 從屬상태에 있기 때문에 推定값이 一致性을 위하여, U_t 의 分散共分散의 逆行列 Σ^{-1} 가 加重行列로 이용되는 一般化된 加重최소제곱법이 더 바람직하다. 즉, 다음식

$$\sum_{t=0}^{T-1} \sum_{s=0}^{T-1} \sigma^{ts} (X_t - \beta_0 - \beta_1 t - \beta_2 t^2) (X_s - \beta_0 - \beta_1 s - \beta_2 s^2) \tag{2. 3}$$

(단 σ^{ts} 는 Σ^{-1} 의 (t, s) -원소)를 최소화 하는 $\beta_0, \beta_1, \beta_2$ 를 결정한다. 이때 σ^{ts} 는 未知이므로 加重최소제곱추정값은 다음 두 단계로 계산된다. 먼저, 식 (2. 3)에 $\sigma^{ts}=1$ ($t, s=0, 1, \dots, T-1$)을 사용한 常최소제곱 추정값 $\hat{\beta}_i, i=0, 1, 2$ 를 구하고, 여기서 계산된 殘差 $\{R_t^{(1)}\}$, $t=0, 1, \dots, T-1$, (단, $R_t^{(1)} = X_t - \hat{\beta}_0 + \hat{\beta}_1 t + \hat{\beta}_2 t^2$)의 標本 分散共分散의 逆行列 $\hat{\Sigma}^{-1}$ 으로 부터 σ^{ts} 를 결정한다. 다음에, 식 (2. 3)에 추정된 $\hat{\sigma}^{ts}$ 를 사용하여 加重최소제곱추정값 $\hat{\beta}_{(0)}, \hat{\beta}_{(1)}, \hat{\beta}_{(2)}$ 를 계산한다.

常최소제곱 추정값은 상용되는 統計팩키지(예, SAS 또는 BMDP)를 이용하여 계산된다. 그러나, 일반적으로 一般化 加重최소제곱 추정값의 계산은 새로운 알고리즘의 개발을 필요로 한다. 그것은 加重값의 推定값 $\hat{\sigma}^{ts}$ 의 계산이 표본크기 T (본 논문에서는 $T=120$)에 의해 결정되는 계수 $T \times T$ 의 標本 分散共分散行列 $\hat{\Sigma}$ 의 역행열계산을 필요로 하기 때문이다.

그러나, 본 분석에서는 常최소제곱방법의 적용으로 얻어진 殘差들이 비교적 간단하고 특수한 확률구조를 갖는 사실로 부터 회귀계수 $\beta_0, \beta_1, \beta_2$ 의 일반화된 加重최소 제곱추정값은 비교적 간단히 계산된다. 계산방법과 그 절차는 보다 일반적인 경우를 포함하여 부록에서 다룬다. 계산된 常최소제곱 추정값은

$$\hat{\beta}_0 = 7.5471 (1.2145) \quad \hat{\beta}_1 = -0.4602 (0.0463) \quad \hat{\beta}_2 = 0.0032 (0.0003)$$

이고, 一般化된 加重최소제곱추정값은

$$\hat{\beta}_{(0)}=7.6629 \quad (0.9091) \quad \hat{\beta}_{(1)}=-0.4620 \quad (0.0347) \quad \hat{\beta}_{(2)}=0.0032 \quad (0.0003)$$

으로 약간의 차이를 나타낸다(괄호안은 추정값의 標準誤差).

가중최소제곱추정치는 회귀추세 성분 D₁의 분석에 그리고 상최소제곱 추정치는 주기성분 C₁의 분석에 이용된다.

계수 β_2 의 값은 기온회귀추세의 曲率효과를 의미하므로 특히 중요하다. 추정값중 $\hat{\beta}_{(2)}$ 는 비교적 작은값을 보이고 있어 실제로 이 값이 零인지 아닌지에 관한 有意性검정이 요구된다. 그러나, 對立假說 " $H_1 : \beta_2 \neq 0$ "에 대한 歸無假說 " $H_0 : \beta_2 = 0$ "은 t-검정(또는 t^2 되는 F-검정)의 값

$$t^* = \frac{\hat{\beta}_{(2)} - 0}{SE(\hat{\beta}_{(2)})} = \frac{0.0032}{0.0003} = 10.67$$

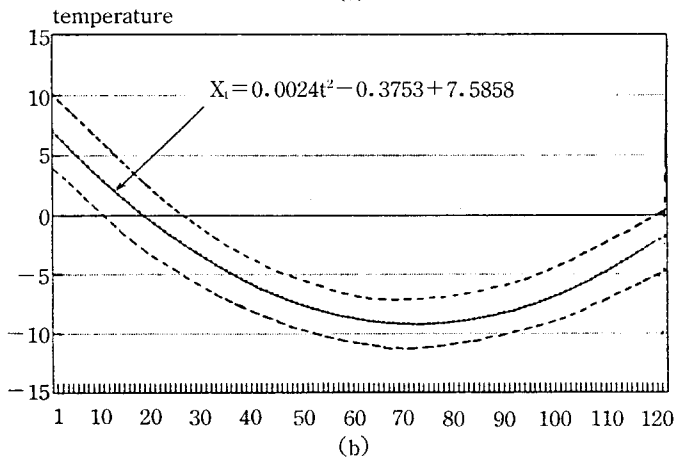
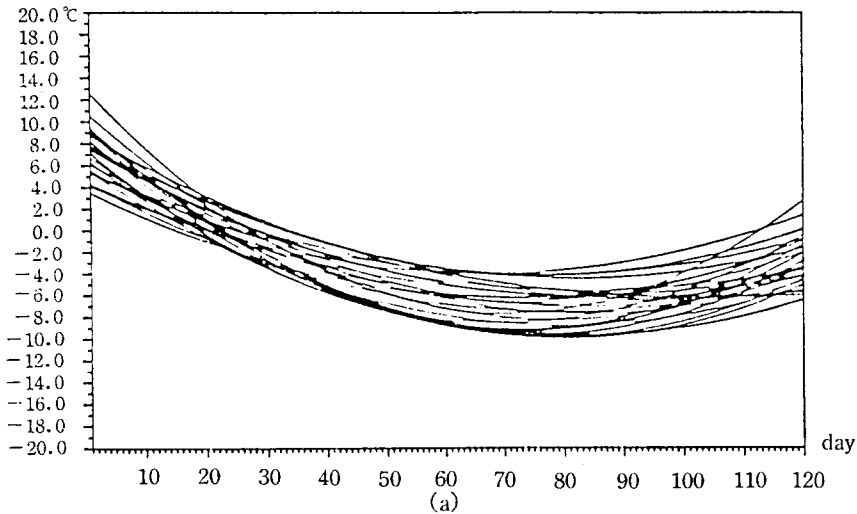


그림 2-2. (a) : 30년 년도별 회귀곡선군 :
(b) : 30년 회귀곡선과 회귀곡선군의 95% 신뢰구간.

으로 t -분포($d. f. = 117$)에서 계산되는 p -값=0.0001으로 쉽게 棄却된다.

그림 2-2(a)는 위와같은 방법으로 결정된 30년에 대한 년도별 回歸曲線群을 나타내고 있다. 어느정도 曲率의 차이는 있지만 서울지역의 겨울철 기온은 2차곡선의 일정한 回歸趨勢를 가지며, 回歸曲線群의 분석으로 볼때 回歸趨勢상의 年最低氣溫은 매년 1월 17일 전후 5일 내에서 85% 이상, 10일 내에서 90% 이상이 나타나고 있다. 그러나, 실제기온에 있어서 년 최저기온의 발생일은 1월 11일을 중심으로 回歸趨勢의 경우보다 더 큰 分散을 갖는다. 실제기온은 回歸趨勢성분만이 아니라, 回歸趨勢와 週期 그리고 確率성분의 총화로 결정되기 때문이다. 1966. 11~1967. 2의 경우 回歸曲線상의 최저기온 일자는 1월 11일, 실제 최저기온 일자는 1월 16일이다. 과거 30년 서울지방 최저기온은 平均 -15.7C 그리고 標準偏差는 2.5C이다. 실제 30년 최저기온은 1970년 1월 5일의 -20.2C이다.

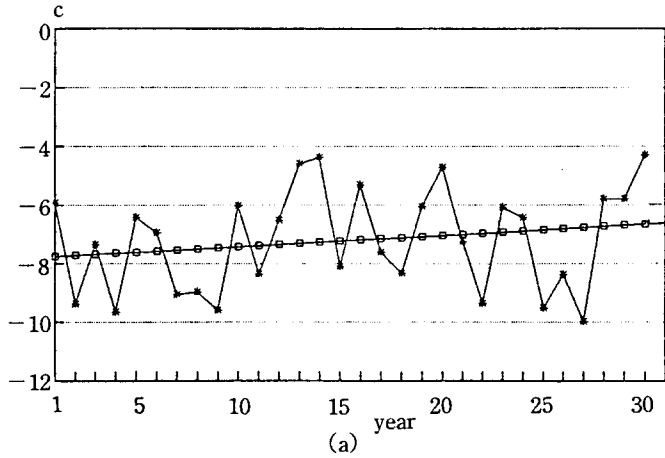
그림 2-2(b)는 30년기온의 回歸曲線과 回歸曲線群의 95% 信賴區間을 동시에 표시한 것이다. 여기서, 30년 기온의 회귀곡선이란 30년기온 관찰값(365×30개) 전체를 대응되는 시점($t=1, \dots, 365$)상에 중복배열하여 얻은 전체자료의 최소제곱포물선을 말하며, 회귀곡선군의 신뢰구간은 독립적으로 계산된 30개년의 각 회귀곡선들을 이용하여 얻어진 30개 예측값의 95%가 포함되는 각 시점의 구간을 말한다. 한 시점에서 계산된 30년 회귀곡선군의 신뢰구간은 독립적으로 계산된 30개의 회귀곡선군상의 값들의 산술평균이므로, 이 곡선들은 異常氣溫현상등 기온의 統計의 특성을 파악하는데 도움이 된다. 보통, 異常氣溫이란 월 또는 년 평균기온이 과거 30년 평균의 $(\pm 2) \times$ 標準誤差를 초과하는 상태를 말한다. 그러나, 기온자료상에 나타나는 異常氣溫은 자연환경의 변화 등에 기인한 순수한 기온의 이상현상 뿐 아니라, 기온에 내존하는 回歸趨勢, 週期, 確率 및 誤差성분들의 중첩 및 상쇄현상에 의해서 나타날 수도 있음을 감안할 때, 異常氣溫의 분석은 回歸趨勢曲線을 이용하는 것이 바람직하다. 이러한 관점에서 異常氣溫의 출현은 그림 2-2(a)에서 각 回歸곡선의 전부 또는 일부가 그림 2-2(b)의 $(\pm 2) \times$ 標準誤差 上, 下界線을 이탈한 정도를 가지고 설명된다.

과거 30개 겨울철중 異常氣溫 현상은 3개년에서 부분적으로 4회 나타나고 있다. 그중 1회는 異常低溫 그리고 3회는 異常高溫현상으로, 異常低溫은 1967년 2월, 異常高溫은 1975년 11월, 1976년 2월 그리고 1988년 2월의 기온이다.

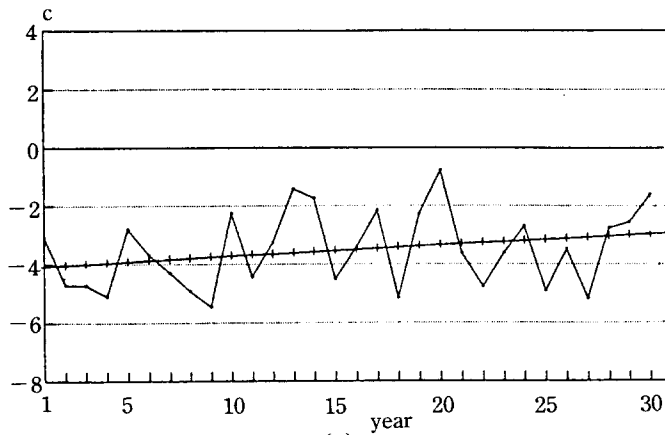
서울지방 겨울철 기온이 매년 상승하고 있는가의 문제는 그림 2-2(a)의 각 回歸曲線상에 나타나는 최소값의 回歸趨勢분석으로 설명될 수 있다. 그림 2-3(a)에서 보는 바와 같이, 각 회귀포물선에서 얻어진 이 최소값들은 비교적 일정한 分散을 가지며 또 어떠한 직선을 중심으로 상하 확률진동을 하고 있다. 따라서, 각 년도 $t(=1, \dots, 30)$ 에 대응하는 최소값 m_t 들은 백색오차 w_t 를 이용한 선형모델

$$m_t = \alpha_0 + \alpha_1 t + w_t, \quad t=1, \dots, 30$$

으로 가정된다. 여기서 얻어진 최소제곱회귀직선은 서울지방의 겨울기온이 30년동안에 1.10C 정도가 상승하고 있음을 설명하고 있다. 한편, 分散分析(표 2. 1(a))에서는 “겨울 기온은 장기적으로 상승한다(즉, $H_1 : \alpha_1 > 0$)”의 (대립)가설을 크게 뒷받침 하지는 못하고 있다(p -값



(a)



(b)

그림 2-3. (a) : 30년 회귀곡선상의 최소값시계열과 최소제공직선 :
(b) : 30년 년평균시계열과 최소제공직선.

=0.3109). 이러한 결과는 30년의 연평균 겨울기온의 회귀추세의 분석에서도 나타나고 있다(그림 2-3(b), 표 2.1(b)). 그러나, 이 분산분석의 결과로 부터 서울지역 겨울기온의 상승추세를 부정할 수는 없다. 김해경 · 한성민[1]에 의하면 서울을 포함한 우리나라 도시지역의 연평균 기온에는 1960년대 부터 시작되는 유의적인 상승추세(서울은 100년에 약 2C)가 존재하며, 이러한 결과가 확인되기 위해서는 일정기간이상의 관측값(서울지역은 유의수준 0.01이하에서 적어도 40년)을 필요로 함이 함이 확인됐기 때문이다.

표 2-1. 분산분석표 (a) : 최소값시계열 : (b) : 년평균시계열.

Source	DF	Sum of Square	Mean Square	F-value	Pf>F
Model	1	3.3074	3.3074	1.06	0.3109
Error	28	86.9632	3.1058		
Total	29	90.2706			

(a)

Source	DF	Sum of Square	Mean Square	F-value	Pf>F
Model	1	3.2837	3.2837	2.00	0.1684
Error	28	45.9875	1.6424		
Total	29	49.27126			

(b)

b) 週期性 분석

모델 (2. 2)의 週期성분 C_t 의 분석은, 시계열 $\{X_t\}$ 에서 (상최소제곱방법에 의하여 얻어진) 回歸趨勢성분이 제거된 殘差 $\{R_t^{(1)}\}$, $R_t^{(1)} = X_t - (\hat{\beta}_0 + \hat{\beta}_1 t + \hat{\beta}_2 t^2)$ 의 주기성 분석에서 시작된다. 그림 2-4에서 보는 바와 같이 시계열 $\{R_t^{(1)}\}$ 는 어떤 週期성분을 내포하고 있다. 이 사실은 스펙트럼의 추정값인 週期圖(그림 2-5(a))와 그의 平滑週期圖(그림 2-5(b))에서 좀 더 구체적으로 관찰된다. 이 때 平滑週期圖은 가장 적당한 $L=9$ 로 식 (2. 4)에 의하여 계산되었다.

$$f(v_h) = \frac{1}{L} \sum_{h=-p}^p w_h P(v_h + \frac{1}{T}) \tag{2. 4}$$

여기서 $P(v_k)$ 는 $R_t^{(1)}$ 의 주파 v_k 에서 週期圖의 값, $W_h = p - |h| + 1$ 그리고 $L = \sum W_h$ 이다.

표 2-2은 과거 30년 동안의 週期상태를 파악하여 연별 主 및 副週期 요소를 나타낸 것이다.

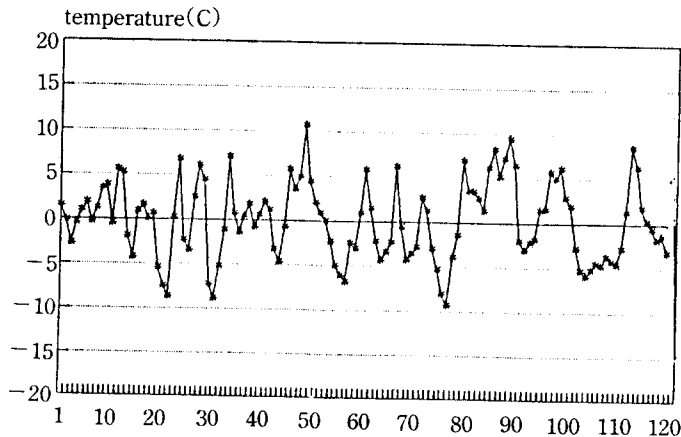


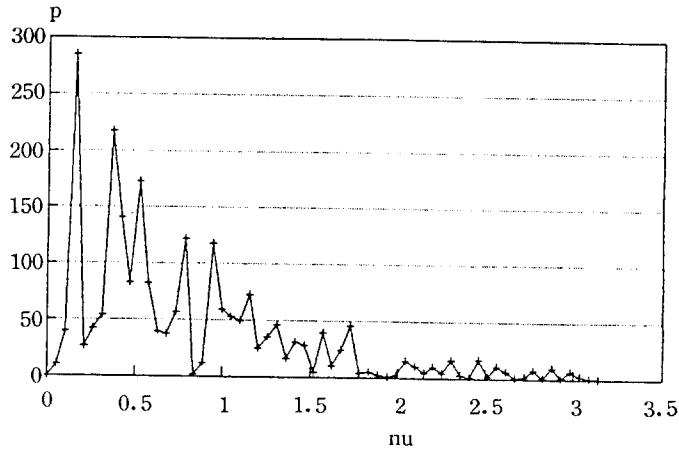
그림 2-4. 殘差 $\{R_t^{(1)}\}$ 의 시계열.

표 2-2. 시계열 $\{R_t^{(1)}\}$ 에서 계산된 30년(1959-1989: 11.1 -2.28) 겨울기온의 主, 副週期.

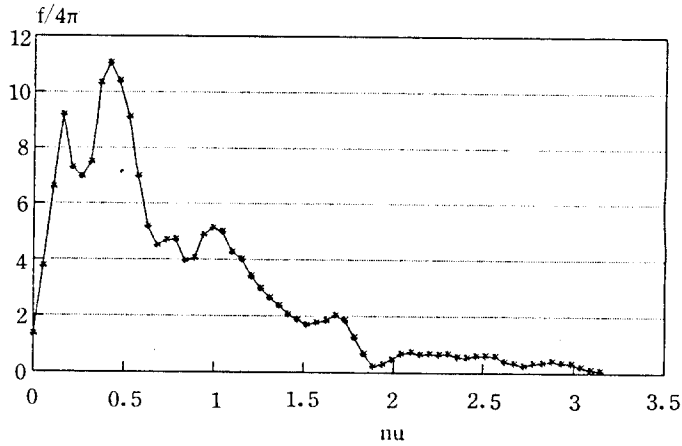
Years	Period(days)				Years	Period(days)			
	Main	Minor				Main	Minor		
1959-1960	30.3	20.2	13.4	40.3	1974-1975	10.9	30.0	40.0	10.0
1960-1961	15.0	10.9	60.0	7.5	1975-1976	40.3	15.1	30.3	20.2
1961-1962	40.0	10.9	7.5	12.0	1976-1977	17.1	8.6	40.0	24.0
1962-1963	60.0	40.0	120.0	6.3	1977-1978	20.0	13.3	15.0	24.0
1963-1964	9.3	24.2	40.3	8.1	1978-1979	15.0	24.0	10.9	6.7
1964-1965	10.9	10.0	40.0	30.0	1979-1980	13.4	60.5	40.3	30.3
1965-1966	8.6	17.1	13.3	7.1	1980-1981	60.0	40.0	30.0	8.6
1966-1967	40.0	17.1	12.0	8.0	1981-1982	60.0	24.0	30.0	13.3
1967-1968	20.2	40.3	24.2	9.3	1982-1983	40.0	30.0	10.9	10.0
1968-1969	60.0	17.1	40.0	24.0	1983-1984	40.3	11.0	10.1	13.4
1969-1970	20.0	24.0	5.5	4.6	1984-1985	30.0	60.0	20.0	8.0
1970-1971	30.0	6.7	8.6	40.0	1985-1986	24.0	30.0	17.1	15.0
1971-1972	60.5	7.6	121.0	40.0	1986-1987	60.0	10.9	30.0	12.0
1972-1973	24.0	40.0	13.3	120.0	1987-1988	60.5	30.3	13.4	20.2
1973-1974	17.1	30.0	13.0	20.0	1988-1989	40.0	10.0	15.0	60.0

1966. 11~1967. 2 자료에서는 40일 主週期요소로, 17일 12일 그리고 8일이 副週期요소로 나타나고 있다. 그러나, 표 2-2에서 보는 바와같이 이 週期는 매년 반복하여 나타나지 않는다. 30년 전체를 통하여 볼 때 겨울철의 週期상태는 11일, 24일, 30일, 40일 그리고 60일로 요약된다. 더욱이, 주기의 調和상태 분석을 통해, 같은 해에서 어떤 週期는 그것의 倍數週期和 독립적으로 발생하고 있음이 확인된다. 이런 관점에서 보면 특히 40일과 30일 그리고 11일은 우리나라 겨울철 기온의 대표週期이며, 평균적으로 40일 週期는 년 0.63회, 30일 週期는 0.57회, 그리고 11일 週期는 년 0.50회 정도 나타난다.

日日 기온자료로부터 재생되는 最低週期가 2일일을 감안하여 6일에서 11일 사이의 週期를 短週期라 하면, 短週期는 평균적으로 년 1.1회 발생한다. 우리나라 겨울철기온의 특성인 三寒四溫은 이 短週期の 출현을 의미한다. 평균적으로 3년에 1회 정도는 이 短週期가 강한 週期로 나타나며 이 경우에는 비교적 뚜렷한 三寒四溫 현상을 갖는다.



(a)



(b)

그림 2-5. 시계열 $\{R_t^{(1)}\}$ 의 (a) : 피리오도그램 ; (b) 평활(L=9)로 추정된 스펙트럼.

위에서 밝혀진 殘差 $\{R_t^{(1)}\}$ 의 主·副週期(단, 40.0과 17.1 週期만 사용)로부터 週期 성분 C_t 의 함수형태는 삼각다항식

$$C_t = \beta_{10} \sin(\beta_{11} + \beta_{12} t) + \beta_{20} \sin(\beta_{21} + \beta_{22} t) \quad (2.5)$$

으로 가정된다. 특히, 일반화된 가중최소제곱 삼각다항식을 결정하기 위하여 다음식

$$\sum_{t=0}^{T-1} \sum_{s=0}^{T-1} \sigma_0^{ts} (R_t^{(1)} - C_t) (R_s^{(1)} - C_s)$$

또는

$$\sum_{t=0}^{T-1} \sum_{s=0}^{T-1} \sigma_0^{ts} [X_t - (D_t + C_t)] [X_s - (D_s + C_s)]$$

을 최소화 하는 $\beta_{i0}, \beta_{i1}, \beta_{i2}, i=1, 2$ 을 찾는다(이 두 결과는 일치함). 여기서 σ_0^{ts} 는 확률성분 Z_t 의 逆分散共分散行列의 (t, s)-원소로 식 (2.3)에서와 같이 이 단계로 추정될 수 있다. 이 계산과정은 부록에서 다룬다. $\{R_t^{(1)}\}$ 의 主, 副週期는 가중최소제곱법의 反復계산에서 β_{12} 와 β_{22} 의 初期값으로 사용한다.

그러나, 다음절에서 수행되는 從屬性분석을 위해서는 $\sigma_0^{ts}=1$ 인 常최소제곱추정값과 대응하는 \hat{C}_t 를 구하고 여기서 계산되는 두번째 殘差 $\{R_t^{(2)}\}, R_t^{(2)}=R_t^{(1)}-\hat{C}_t$ 의 계산이 필요하다.

1966.11-1967.2의 경우, 상최소제곱 추정치는

$$\begin{aligned} \hat{\beta}_{10} &= 2.3276 & \hat{\beta}_{11} &= 0.3175 & \hat{\beta}_{12} &= 0.1576 \\ \hat{\beta}_{20} &= -1.0242 & \hat{\beta}_{21} &= -0.5347 & \hat{\beta}_{22} &= 0.3050 \end{aligned}$$

이고, 대응되는 週期성분 C_t 의 함수식은

$$\hat{C}_t = 2.3276 \sin(0.317 + 0.1576 t) - 1.0242 \sin(-0.5347 + 0.3050 t)$$

이다. 그림 2-6은 殘差 $\{R_t^{(1)}\}$ 와 推定式 \hat{C}_t 를 동시에 나타낸 것이다. 殘差 $\{R_t^{(2)}\}$ 는 그림 2-7에 나타내었다. 殘差 $\{R_t^{(1)}\}$ 에 내존하는 40일, 17.1일의 主, 副週期성분에 대응하는 平滑週期圖의 값이 각각 284.905, 217.375인데 반하여, 殘差 $\{R_t^{(2)}\}$ 의 경우에는 이 값들이 각각 0.087, 0.008로 감소되므로 이 과정에서 이 두 주기성분이 어느정도 제거됨을 확인할 수 있다. 또, 이 사실은 推定된 삼각다항식이 週期성분의 함수식으로 적합함을 의미한다.

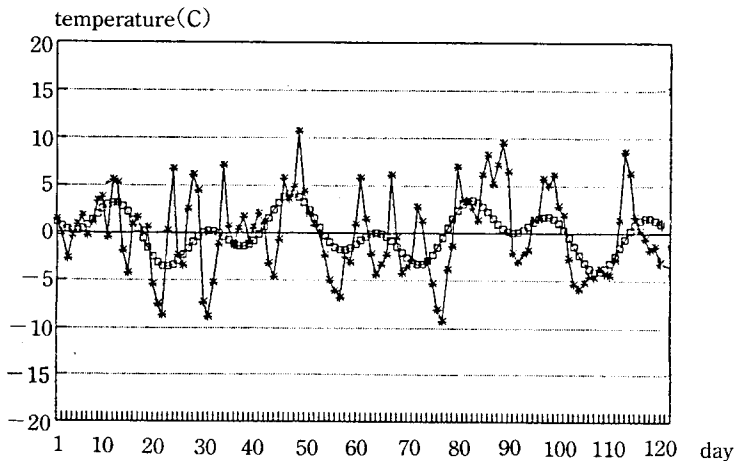


그림 2-6. 시계열 $\{R_t^{(1)}\}$ (-*-)과 추정된 주기성분함수 $\{\hat{C}_t\}$ (-□-).

c) 從屬性 분석

분석모델 (2. 2)에서 確率성분 Z_t 는 시계열 X_t 에 내존하는 從屬性을 설명한다. 確率 성분 Z_t 의 確率모델은 X_t 에서 回歸성분 D_t 와 週期성분 C_t 가 모두 제거된 殘差 $R_t^{(2)}$ 의 분석으로 결정된다. 殘差 $R_t^{(2)}$ 는 앞절에서와 같이 $R_t^{(1)} - \hat{C}_t$ 또는 $X_t - (\hat{D}_t + \hat{C}_t)$ 으로 정의될 수 있으나, 이 두 잔차는 같은 시계열이 된다. 그림 2-7에서 보는 바와같이 시간에 따라 나타나는 일정한 平均과 分散등은 시계열 $\{R_t^{(2)}\}$ 의 定常 또는 弱定常상태를 암시하며, 이 사실은 구체적으로 時差에 따른 標本自己相關함수와 標本部分自己相關함수로 확인된다.

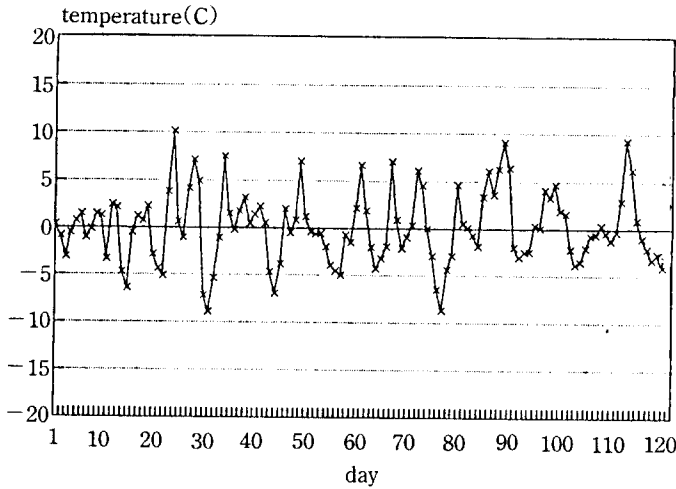


그림 2-7. 殘差 $\{R_t^{(2)}\}$ 의 시계열.

그림 2-8(a)와 (b)에서 보는 바와같이, 標本自己相關함수는 時差가 증가할수록 급속히 쪼아 가까워지고, 標本部分自己相關함수는 時差 2 이후에서 切捨하고 있다. 이와같은 상태는 1966. 11~1967. 2의 경우, Z_t 가 2계의 自己回歸過程(AR(2)로 표시)

$$Z_t + a_1 Z_{t-1} + a_2 Z_{t-2} = e_t \tag{2. 6}$$

을 따르고 있음을 의미한다. (조건) 最小제곱법(Abraham & Ledolter, 1983 참조)을 이용한 a_1, a_2 의 推定값는 다음과 같다.

$$\hat{a}_1 = -0.8206 (0.04979), \hat{a}_2 = 0.3171 (0.04982),$$

따라서, 確率성분 Z_t 의 2계 自己回歸모델은

$$Z_t - 0.8206 Z_{t-1} + 0.3171 Z_{t-2} = e_t$$

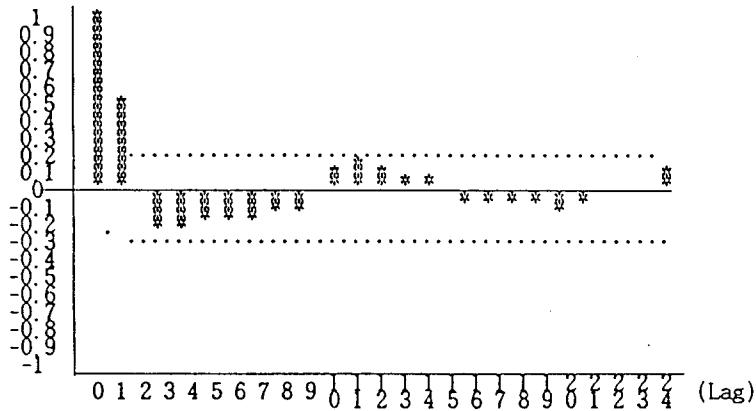
로 주어진다. 이 推定값들은 2계 自己回歸過程의 定常조건인 부등식 $a_1 + a_2 \geq -1, a_1 - a_2 \leq 1, |a_2| < 1$ 을 만족하고 있다. 그림 2-9은 殘差 $\{R_t^{(2)}\}, t=0, 1, \dots, 119,$ 와 모델 (2. 6)에 의한 추정값 $\{\hat{Z}_t\}$ 를 비교한 것이다. 이 때, \hat{Z}_t 는

$$\hat{Z}_t = \begin{cases} R_0^{(2)} & t=0 \\ 0.8206 \hat{Z}_0 & t=1 \\ 0.8206 \hat{Z}_{t-1} - 0.3171 Z_{t-2} & t \geq 2 \end{cases}$$

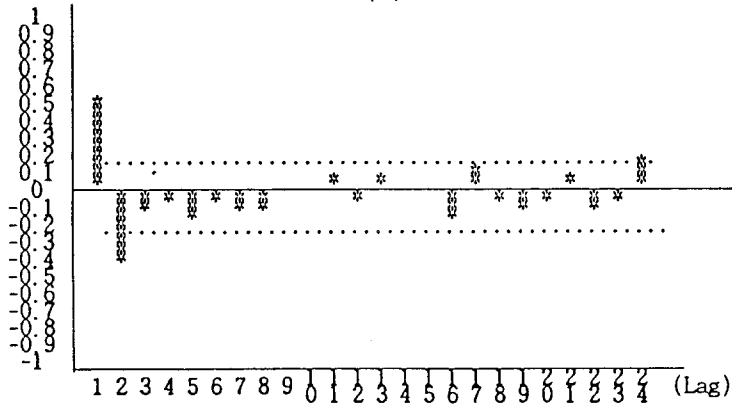
로 주어진다. 새로운 殘差 $\{R_t^{(3)}\}$, $R_t^{(2)} - \hat{Z}_t$ 의 正規白色誤差의 성립은 식 (2.6)이 해당년(1966. 11~1967. 2) 기온의 확률성분 모델로써 적합함을 의미하며 이것은 4節에서 확인된다.

위와같은 방법으로, 30년 기온자료의 年도별 확률성분 모델들이 결정된다. 30년 전체로 볼 때, 기온의 확률성분은 약 80%가 AR(2)로 나타나고 나머지는 AR(1)로 설명된다.

확률성분이 대부분 AR(2)로 설명됨은 겨울기온의 “2日從屬性”을 의미한다. 즉, 오늘의 기온은



(a)



(b)

그림 2-8. 시계열 $\{R_t^{(2)}\}$ 의 (a) : 표본자기상관함수 ; (b) : 표본부분자기상관함수.

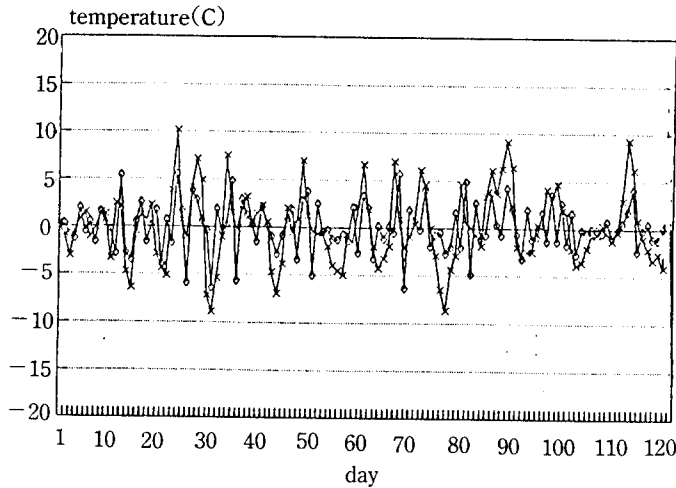


그림 2-9. 시계열 $\{R_t^{(2)}\}$ ($- \times -$)와 추정된 확률성분함수 $\{\hat{Z}_t\}$ ($- \diamond -$).

내일과 그리고 그 다음날의 기온에만 직접적인 영향을 주고 있음을 말한다. 더욱이, 확률성분의 계수가 식 (2. 7)에서와 같이 주어질 때 현재의 기온이 2일을 넘어서 미래의 기온에 미치는 영향은 급속히 감소됨은 쉽게 설명된다. 이러한 현상은 30년 기온자료 전체에서 관찰된다. 겨울기온의 이와같은 현상은 과거의 일일 기온자료를 기초로한 2일을 넘어서 미래기온의 예측에는 적어도 이 확률성분의 표준편차 정도의 예측오차가 필연적으로 동반됨을 말해준다. 이와같은 3절의 實例분석을 통해 확인된다.

3. 기온의 確率모델과 適合度

기본가정식 (2. 2)와 앞 節의 분석들을 통하여 얻은 각 성분 D_t , C_t , Z_t 등의 確率모델들을 이용하여 시계열 X_t 의 確率모델이 완성된다. 즉, 기온시계열 X_t 의 確率모델은

$$\begin{aligned}
 X_t = & \beta_0 + \beta_1 t + \beta_2 t^2 \\
 & + \beta_{10} \sin(\beta_{11} + \beta_{12} t) + \beta_{20} \sin(\beta_{21} + \beta_{22} t) \\
 & + a_1 Z_{t-1} + a_2 Z_{t-2} \\
 & + e_t
 \end{aligned} \tag{3. 1}$$

여기서 Z_t 는 殘差 $R_t^{(2)}$ 로 近似化 될 수 있는 確率變數이고, e_t 는 正規($0, \sigma^2$) 確率變數이다. 특히 Z_t 가 AR(1)이면 $a_2=0$.

測定誤差의 分散 σ^2 은 殘差 $\{R_t^{(3)}\}$ 의 分散으로 推定될 수 있다.

確率모델 (3. 1)은 기온의 豫測방정식을 구하는데 이용 될 수 있다. 그러나, 앞 節에서 밝혀진

바와같이 回歸趨勢 D_t 는 異常氣溫 현상 등 연도에 따른 특성을, 週期성분 C_t 는 해당년 고유의 週期를, 自己回歸確率 성분 Z_t 는 자기의 階數를 각각 갖게 되므로, 豫測방정식은 기본적으로 豫測年の 기온자료만으로 결정되는 것이 바람직하다.

豫測年 전반기 일일 기온자료 $\{X_t\}$, $t=0, 1, \dots, T-1$ 을 사용하여, 모델 (3. 1)를 통한 후반기의 기온의 豫測방정식은 다음 다섯단계로 결정된다.

1단계: 回歸趨勢성분 $D_t = \beta_0 + \beta_1 t + \beta_2 t^2$ 의 係數 $\beta_0, \beta_1, \beta_2$ 들에 대한 尙最劣供給 추정값을 구한다. 이 때, 이 趨勢성분은 非確率성분으로 간주하여 과거(30년) 기온자료 모두를 이용하여 초기값을 결정한다.

2단계: 殘差 $\{R_t^{(1)}, R_t^{(1)} = X_t - \hat{D}_t\}$ 의 分散共分散 逆行列을 구하고, 식 (2. 2)에 의한 $\beta_0, \beta_1, \beta_2$ 의 尙最劣供給 추정값들을 구한다.

3단계: 殘差 $\{R_t^{(1)}\}$ 의 主, 副週期를 平滑週期圖을 통해 구하고, 이것을 초기값으로 하여 週期성분

$$C_t = \sum_{i=1}^p \beta_{i0} \sin(\beta_{i1} + \beta_{i2}t)$$

의 未知係數 $\beta_{i0}, \beta_{i1}, \beta_{i2}$, ($i=i=1, \dots, p$)에 대한 尙最劣供給추정값을 구한다. 이때의 尙最劣供給 殘差 $\{R_t^{(1)}\}$, $R_t^{(2)} = X_t - (\hat{D}_t + \hat{C}_t)$ 의 分散共分散逆行列이고, p 는 主, 副週期の 갯수이다.

4단계: 尙最劣供給 D_t, C_t 의 尙最劣供給 $D_t + C_t$ 을 X_t 에 적용하여 추정값을 재조정하고 여기서 얻은 殘差 $\{\bar{R}_t^{(2)}\}$, $\bar{R}_t^{(2)} = X_t - (\hat{D}_t + \hat{C}_t)$ 의 標本 自己相關 및 部分自己相關함수를 관찰하여 自己回歸過程의 階數 q 를 결정하고, 또 이 殘差를 이용하여 식 (2. 6)에서와 같은 방법으로 未知係數의 尙最劣供給추정값을 구한다.

5단계: 위에서 결정된 각 성분들로 부터 豫測방정식을 완성한다. 즉, $t \geq T+1$ 에 대하여

$$\hat{X}_t = \hat{\beta}_0 + \hat{\beta}_1 t + \hat{\beta}_2 t^2 + \sum_{i=1}^p \hat{\beta}_{i0} \sin(\hat{\beta}_{i1} + \hat{\beta}_{i2}t) + \sum_{u=1}^q \hat{a}_u \hat{Z}_{t-u} \quad (3. 2)$$

이다. 여기서, \hat{Z}_t , $t = -q, \dots, -1, 0, 1, 2, \dots$ 는 다음식으로 계산한다.

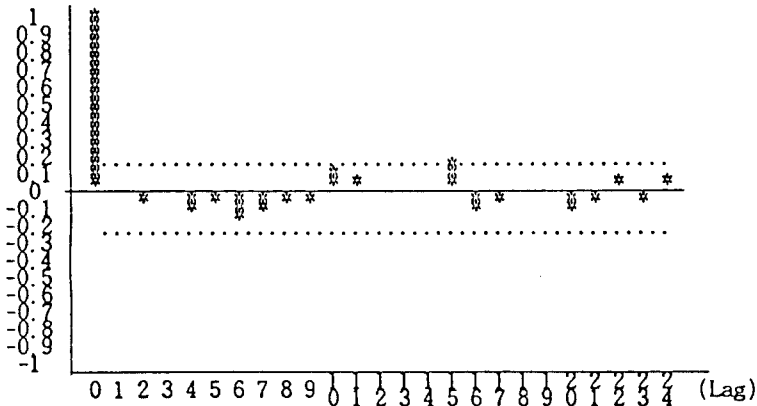
$$\hat{Z}_t = \begin{cases} 0 & t = -q, -q+1, \dots, -1 \\ \bar{R}_t^{(2)} & t = 1, 2, \dots, T \\ \sum_{u=1}^q \hat{a}_u \hat{Z}_{t-u} & t = T+1, T+2, \dots \end{cases}$$

모델 (3. 1)이 기온시계열의 모델로써 적합한지의 확인은 이 모델을 실제자료에 적용한 후 얻어지는 殘差 $\{R_t^{(3)}\}$, $R_t^{(3)} = X_t - (\hat{D}_t + \hat{C}_t) - \hat{Z}_t$ 의 분석으로 이루어진다. 확률모델 (3. 1)에서 가정된 바와같이 誤差項 $\{e_t\}$ 의 正規白色誤差의 성립은 모델의 適合을 의미한다.

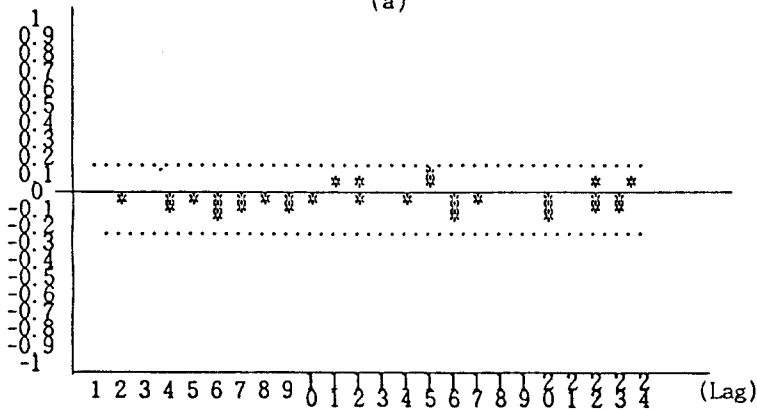
그림 3-1(a), (b)는 1966. 11~1967. 2 자료에 대한 殘差 시계열의 標本 自己 및 部分自己相關함수를 나타낸 것으로 모두 白色誤差의 특성을 보여준다. 標本 自己相關係數들로 계산되는 Box-Pierce Chi-square 統計量[3] 등 다른 檢定統計量들도 이 시계열이 白色誤差임을 확인해 준다. 또, 이 시계열의 正規性은 正規確率紙 또는 適合度檢定을 통해 쉽게 확인된다.

확률모델의 적합성이 곧 관련 예측방정식의 정확성을 의미하지는 않는다. 일반적으로 예측방정식이 미래의 기온을 얼마나 정확하게 예측할 수 있는나하는 문제는 우선(일반화된 가중최소제곱 추정값의 계산을 포함한) 모든 알고리즘들이 얼마나 성공적으로 개발되느냐에 달려 있다.

그러나, 앞절에서 언급한 바와같이 예측방정식에 포함된 (自己回歸) 從屬성분으로부터 발생하는 예측오차는 알고리즘의 성공적인 개발로도 해결될 수 없다. 이러한 현상은 從屬성분을 갖는 기상 또는 기온현상의 특징이기도 하다. 이 從屬성분에서 발생하는 오차의 크기는 기본적으로 예측 간격의 크기에 달려 있다. 서울지방 겨울기온의 경우(오차의 크기를 비교적 적게하는) 적당한 예측간격은 대략 4, 5일로 관찰되지만 기본적으로 自己回歸성분의 階數에 의존되는 예측년의 고유값으로 볼 수 있다. 해당년 適正豫測 간격을 초과한 기온예측은 回歸趨勢와 週期성분만을 나타내는 예측으로 한정된다.



(a)

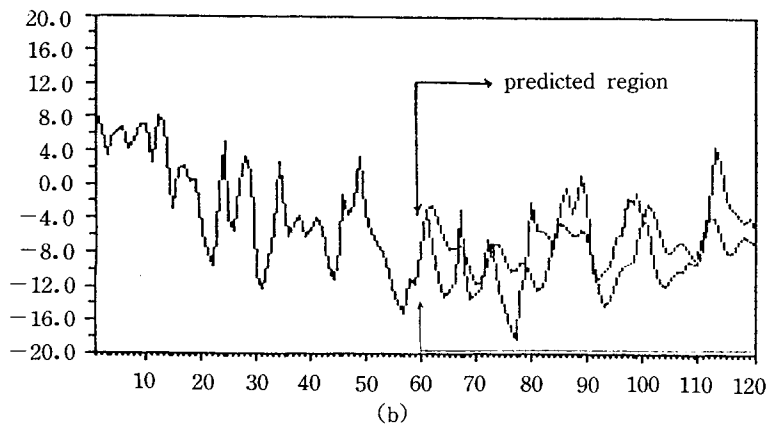
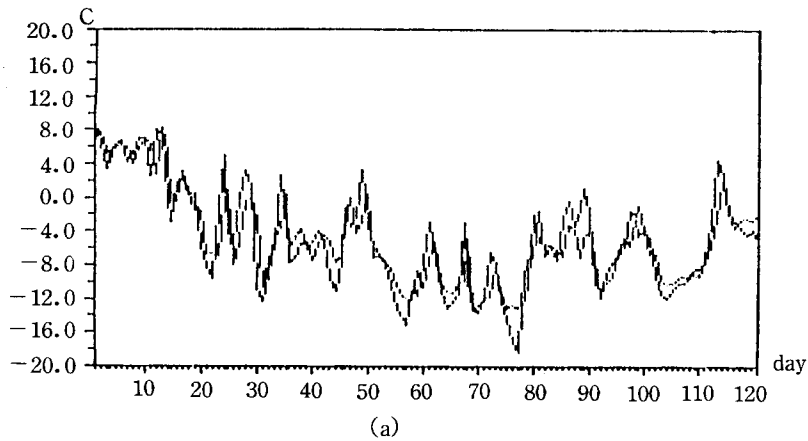


(b)

그림 3-1. 시계열 $\{R_t^{(3)}\}$ 의 (a) : 표본자기상관함수 : (b) : 표본부분자기상관함수.

그림 3-2는 1966. 11~1967. 2 자료를 사용하여 얻는 豫測값과 실제관찰값을 동시에 나타낸 것이다. 그림 3-2(a)에서는 120일 전체자료를 사용하여 계산된 예측값과 관찰값을 먼저 비교하였다. 그림 3-2(b), (c) 그리고 (d)는 이 자료중 전반부 60일, 70일 그리고 80일 관찰값만을 각각 사용해 얻은 후반부의 예측값과 실제관찰값을 동시에 나타낸 것이다. 여기서 보는 바와 같이 세 경우 모두 예측간격 5, 6일까지는 비교적 정확한 예측이 이루어진다.

기온예측에 있어서, 위와 같은 點豫測 뿐 아니라 豫測信賴區間의 설정은 보다 유용할 때가 있다. 그러나, 확률모델 (3. 1)을 기초로 하는 信賴區間의 개발에는 많은 수학적인 어려움이 있다. 특히, 기온모델이 線型 및 非線型的 回歸함수 그리고 시계열의 確率함수를 동시에 갖는 형태로, 관련 統計量에 관한 聯合確率分布의 결정에 어려움이 있다.



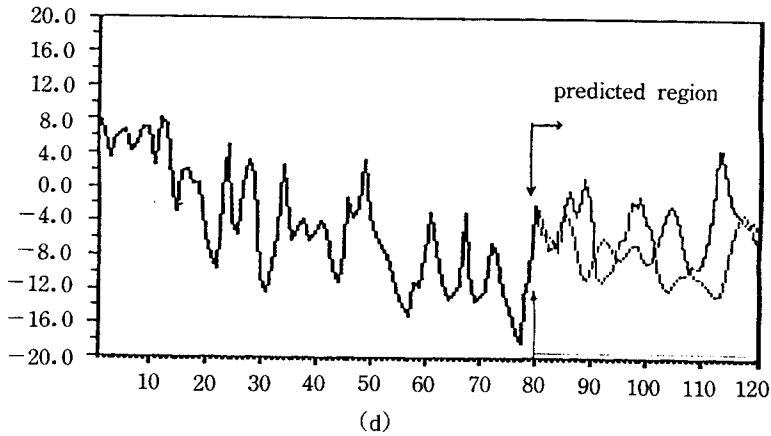
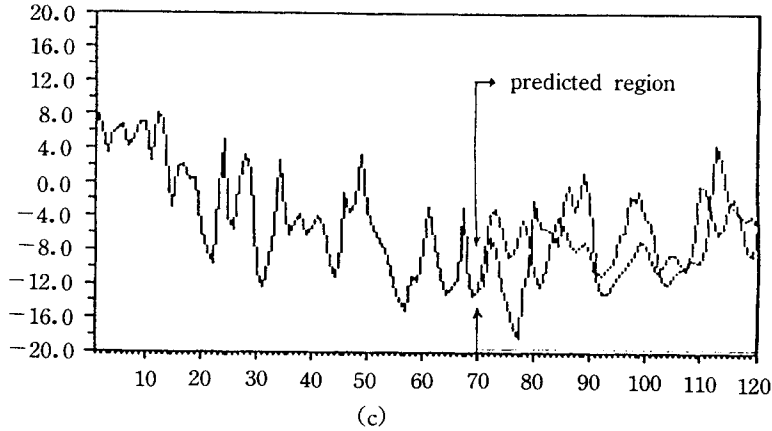


그림 3-2. 기온시계열(—)과 예측시계열(---); 예측에 사용된 부분 :
 (a) : 1966. 11. 1-1967. 2. 28 ; (b) : 1966. 11. 1.-1966. 12. 31 ;
 (c) : 1966. 11. 1-1967. 1. 9 ; (d) : 1966. 11. 1.-1967. 1. 19.

4. 결 론

겨울철 일일 최저기온과 최고기온 그리고 최저기온과 평균기온의 (시계열)相關係數가 모두 0.93 이상임을 감안할 때, 서울지방 겨울철 기온의 統計 및 確率의 특징은 다음과 같다.

첫째, 겨울철 기온은 2차곡선의 回歸趨勢를 가지며, 回歸趨勢상의 최저기온은 매년 1월 17일전후 10일 내에서 나타난다. 異常低溫 또는 高溫 현상은 평균적으로 10년에 1회정도 나타나고 있으며,

연평균 겨울기온에는 30년에 1.11C 정도의 일정한 선형의 회귀추세를 가지고 상승하는 경향이 있다.

둘째, 겨울철 기온의 대표 週期는 40일, 30일 그리고 11일이며, 평균적으로 40일 週期는 연 0.63회, 30일 週期는 연 0.57회 그리고 11일 週期는 연 0.50회 정도 나타난다. 三寒四溫현상은 연평균 1.1회 발생하는 短週期(6~11일 週期)의 출현을 의미하며, 비교적 뚜렷한 三寒四溫은 평균적으로 3년에 1회 정도 나타난다.

셋째, 겨울철 기온은 階數가 2인 自己回歸過程으로 설명되는 “2日從屬性”을 갖는다. 이것은 겨울기온의 영향 지속기간이 48시간 임을 의미하며, 이 때 적당한 예측간격은 4, 5일이다.

넷째, 서울지방 겨울철 기온의 確率모델은 (3. 2)로 주어지며, 이 모델을 이용한 기온예측의 統計的 절차는 3節에서와 같다.

5. 附 錄

((一般化된 加重최소제곱 추정값의 계산))

일반적으로, 線型 또는 非線型회귀함수 $f(\theta, t)$ 를 갖는 회귀모델

$$X_t = f(\theta, t) + Z_t, \quad t=0, \dots, T-1$$

에 관한 θ 의 일반화된 加重최소제곱 추정값은 이차식

$$[Y - F(\theta)]\Sigma^{-1}[Y - F(\theta)] \quad (A. 1)$$

을 최소화하는 상수벡터 $\theta \in R^p$ 이다. 이때, $Y = [X_0, \dots, X_{T-1}]'$, $F(\theta) = [f(\theta, 0), \dots, f(\theta, T-1)]'$ 이고 Σ 는 관찰되지 않는 확률오차 Z_t 의 분산공분산행렬이다. 따라서, 이 추정값의 계산에는 상최소제곱방법에 의하여 얻어지는 (Z_t 의 추정값으로) 殘差의 표본 분산공분산 행렬을 이용한 加重行列의 逆行列 Σ^{-1} 의 추정이 선행되어야 한다. 그러나, 이 과정에서 큰 T값에 대한 역행렬 계산은 효과적인 알고리즘의 개발을 요구하고 있다.

본 분석의 경우에는 상최소제곱방법의 적용으로 얻어진 殘差들이 비교적 간단하고 또 특수한 확률구조를 갖는 사실로부터, 회귀함수식에 나타난 모수의 일반화된 加重최소제곱 추정값은 비교적 쉽게 계산된다.

먼저, 본문 2節 a)에서와 같이

$$f(\theta, t) = D_t = \beta_0 + \beta_1 t + \beta_2 t^2$$

으로 주어질 때, 식 (A. 1)는

$$[Y - X\theta]'\Sigma^{-1}[Y - X\theta] = [\Gamma Y - \Gamma X\theta]'\Gamma^{-1}[\Gamma Y - \Gamma X\theta]$$

로 표시되고, $\theta = (\beta_0, \beta_1, \beta_2)$ 의 일반화된 加重최소제곱 추정량은

$$\theta = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}Y \quad (A. 2)$$

이다. 이 때, X 는

$$X = \begin{pmatrix} 1 & 1 & 1^2 \\ 1 & 2 & 2^2 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ 1 & 119 & 119^2 \\ 1 & 120 & 120^2 \end{pmatrix}$$

으로 $\Sigma^{-1} = \Gamma' \Gamma$ 으로 표현된다(행렬 Γ 는 항상 존재한다). 한편, 2절 b)의 $\{R_t^{(2)}\}$ 분석에서와 같은 방법으로, $\{R_t^{(1)}\}$ 의 분석을 기호로 할때 오차확률성분 Z_t 는 $\hat{a} = -0.8206$, $\hat{a}_2 = 0.3171$ 를 추정값으로 갖는 2계의 자기회귀과정

$$Z_t + a_1 Z_{t-1} + a_2 Z_{t-2} = e_t$$

으로 설명된다. 더욱이, AR(2)를 따르는 Z_t 의 공분산행렬 Σ^{-1} 이

$$\Sigma^{-1} = \begin{pmatrix} 1 & a_1 & a_2 & 0 & \cdots & 0 & 0 \\ a_1 & 1+a_1^2 & a_1+a_1a_2 & a_2 & \cdots & 0 & 0 \\ a_2 & a_1+a_1a_2 & 1+a_1^2+a_2^2 & \cdot & \cdots & 0 & 0 \\ 0 & a_2 & a_1+a_1a_2 & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & a_1+a_1a_2 & a_2 & 0 \\ 0 & 0 & \cdots & \cdot & 1+a_1^2+a_2^2 & a_1+a_1a_2 & a_2 \\ 0 & 0 & \cdots & a_2 & a_1+a_1a_2 & 1+a_1^2 & a_1 \\ 0 & 0 & \cdots & 0 & a_2 & a_1 & 1 \end{pmatrix}$$

이 되는 사실([5], [6])을 이용하면, 일반화된 가중최소제곱추정치는 (A, 2)로 부터 쉽게

$$\hat{\beta}_{(0)} = 7.6629 \quad \hat{\beta}_{(1)} = -0.4620 \quad \hat{\beta}_{(2)} = 0.0032$$

로 주어진다.

다음으로, 2절 b)에서와 같이 가정식 (2. 2)의 두 성분 D_t , C_t 의 합 $D_t + C_t$ 가 회귀함수식

$$f(\theta, t) = D_t + C_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_{10} \sin(\beta_{11} + \beta_{12} t) + \beta_{20} \sin(\beta_{21} + \beta_{22} t)$$

일 때, (A. 1)의 성분표현형은

$$\sum_{t=0}^{T-1} \sum_{s=0}^{T-1} \sigma^{ts} [X_t - (D_t + C_t)] [X_s - (D_s + C_s)]$$

이며, σ^{ts} 는 잔차 $\{R_t^{(2)}\}$ 의 역 분산공분산 행렬의 (t, s)-원소로 추정될 수 있는 오차성분 Z_t 의 역분산공분산행렬의 대응원소이다. 그리고, 관련된 모수벡터

$$\theta = (\beta_0, \beta_1, \beta_2, \beta_{10}, \beta_{11}, \beta_{12}, \beta_{20}, \beta_{21}, \beta_{22})$$

의 추정치는 다음과 같다.

위에서와 같이, $\{R_t^{(2)}\}$ 분석에 의하면 확률성분 Z_t 는 근사적으로 2계의 자기회귀過程

$$Z_t + a_1 Z_{t-1} + a_2 Z_{t-2} = e_t$$

으로 설명되고, 여기서 얻어진 계수의 추정값 $\hat{a}_1 = -0.7618$, $\hat{a}_2 = 0.3731$ 이 σ^2 의 추정에 이용된다.

AR(2)의 $\Sigma^{-1} = \Gamma' \Gamma$ 인 Γ 를 이용하면 이차식 (A. 1)은

$$[Y - F(\theta)]' \Sigma^{-1} [Y - F(\theta)] = [\Gamma Y - \Gamma F(\theta)]' [\Gamma Y - \Gamma F(\theta)] \quad (A. 3)$$

으로 표현된다. 여기서 $f(\theta, t) = D_t + C_t$ 이다. 특히, 충분히 큰 T에 대해서는 다음에 정의되는 $T \times T$ 행렬 Γ^* 를 Γ 의 근사행렬로 사용될 수 있다.

$$\Gamma^* = \begin{pmatrix} 1 & a_1 & a_2 & 0 & \cdots & 0 \\ 0 & 1 & a_1 & a_2 & \cdots & 0 \\ \cdot & \cdot & \cdot & & \cdots & \cdot \\ \cdot & \cdot & \cdot & & \cdots & \cdot \\ \cdot & \cdot & \cdot & & \cdots & \cdot \\ 0 & 0 & \cdots & 1 & a_1 & a_2 \\ 0 & 0 & \cdots & & 1 & a_1 \\ 0 & 0 & \cdots & & & 1 \end{pmatrix}$$

식 (A. 3)의 좌변을 최소화하는 미지모수벡터 θ 의 일반화된 가중최소제곱추정값은 우변을 최소화하는 상최소제곱추정값으로 구해진다. a_1, a_2 의 추정값 그리고 비선형회귀모델의 상최소제곱 추정팩키지(예, SAS의 PROC NLIN)를 이용한 θ 의 일반화된 가중최소제곱추정값은

$$\begin{aligned} \hat{\beta}_{(0)} &= 6.5631 & \hat{\beta}_{(1)} &= -0.4313 & \hat{\beta}_{(2)} &= 0.0030 \\ \hat{\beta}_{(10)} &= 2.4829 & \hat{\beta}_{(11)} &= 0.3858 & \hat{\beta}_{(12)} &= 0.1562 \\ \hat{\beta}_{(20)} &= 1.2768 & \hat{\beta}_{(21)} &= -0.3193 & \hat{\beta}_{(22)} &= 0.3026 \end{aligned}$$

으로 주어지며, 다음의 상최소제곱 추정값

$$\begin{aligned} \hat{\beta}_0 &= 6.4856 & \hat{\beta}_1 &= -0.4342 & \hat{\beta}_2 &= 0.0030 \\ \hat{\beta}_{10} &= 2.3276 & \hat{\beta}_{11} &= 0.3175 & \hat{\beta}_{12} &= 0.1576 \\ \hat{\beta}_{20} &= 1.0242 & \hat{\beta}_{21} &= -0.5347 & \hat{\beta}_{22} &= 0.3050 \end{aligned}$$

과 약간의 차이를 보인다. 이 때 상최소제곱방법에 의한 잔차제곱합(SEE)은 1862.42이고, 이에 반해 가중최소제곱방법에 의해 만들어지는 잔차제곱합은 1122.09이다.

참 고 문 헌

- [1] 김해경, 한성민(1992). 太陽의 黑點이 地球氣溫에 미치는 影響. 大氣保全 : 韓國大氣保全學會誌(제출중)
- [2] Adbrham, B. & Ledolter, J.(1983), *Statistical Methods for Forecasting*. John Wiley & Sons, 27-28.
- [3] Box, G.E.P. & Jenkins, G.M.(1970), *Time Series Analysis : Forecasting and Control*, 2nd ed. San Francisco : Holdon-Day, 290.
- [4] Nelson, C.R.(1973), *Applied Time Series Analysis for Manageral Forecasting*. San Francisco : Holden-Day
- [5] Siddiqui, M.M.(1958), "On the inversion of the sample covariance matrix in a stationary autoregressive process". *Annals of mathematical Statistics* 29, 58-588.
- [6] Wise, J.(195), "The autocorrelation function and the spectral density function", *Biometrika*. 42, 151-159.

A Stochastic Model for Winter Air-Temperature of Seoul Area

Hae-Kyung Kim* and Tae-Soo Kim**

ABSTRACT

This paper is concerned with the development and application of a stochastic model for winter air-temperature of Seoul area. The annual and interannual fluctuations of the regression trend, periodicity and dependence of the air-temperature are analyzed based on the data during the past 30 years (1959–1989). A statistical procedure for using the stochastic model to predict the air-temperature is proposed. Some statistical characteristics of winter air-temperature including unusual air-temperature and Samhansaon are also discussed.

* Dept. of Mathematics, Yonsei University, Shinchon-dong, Seoul 120-749, Korea
** Dept. of Mathematics, Yonsei University, Shinchon-dong, Seoul 120-749, Korea