II－301　年水文量に対する時系列モデルについて

群馬高専　正会員　山本好克

1．はじめに　時系列モデルは、水工計画上必要となる水文量の模擬発生あるいは将来予測をする上で有用となろう。ここでは、年単位の水文量に対する時系列モデルとその適用例について述べる。

2．ARMA(p, q)モデルとその一般的性質　定常過程とみなせる年水文量系列に対する時系列モデルには、p次自己回帰型［AR(p)］，q次移動平均型［MA(q)］，あるいは混合型［ARMA(p, q)］などが考えられる。ARMA(p, q)モデルにおいて，$q=0$とすればAR(p)モデルに，$p=0$とすればMA(q)モデルになるので，ここでは，ARMA(p, q)モデルについて述べる。

モデルの一般式は，$y_t=\mu+\sigma Z_t$ ---(1)，$Z_t=\sum_{j=1}^{p}\phi_j Z_{t-j}-\sum_{j=0}^{q}\theta_j\varepsilon_{t-j}$；$\theta_0=-1$ ----(2) と表わされる。ここで，y_t：水文量系列，μ；σ^2：y_tの平均および分散，ϕ_j；θ_j：j番目の自己回帰係数および移動平均係数，ε_t：平均0，分散σ_ε^2の独立正規変量である。パラメータは，μ，σ^2，$\phi_1\sim\phi_p$，$\theta_1\sim\theta_q$，σ_ε^2の$(p+q+3)$個となる。

モデルの自己共分散γ_kは，(2)式の両辺にZ_{t-k}を乗じ期待値をとると，$\gamma_k=\sum_{i=1}^{p}\phi_i\gamma_{k-i}+\gamma_{Z\varepsilon}(k)-\sum_{i=1}^{q}\theta_i\gamma_{Z\varepsilon}(k-i)$，$k<q+1$ ----(3)，$\gamma_k=\sum_{i=1}^{p}\phi_i\gamma_{k-i}$，$k\geq q+1$ ----(4) となる。ここで $\gamma_{Z\varepsilon}(k)=Cov[Z_{t-k}\cdot\varepsilon_t]$ であり，kは遅れ時間を表わす。こうしてモデルの自己相関係数ρ_kは，(3)，(4)式を分散$\gamma_0=Var[Z_t]$で除すことにより，$\rho_k=\sum_{i=1}^{p}\phi_i\rho_{k-i}+\sigma_\varepsilon\rho_{Z\varepsilon}(k)-\sigma_\varepsilon\sum_{i=1}^{q}\theta_i\rho_{Z\varepsilon}(k-i)$，$k<q+1$ ----(5)，$\rho_k=\sum_{i=1}^{p}\phi_i\rho_{k-i}$，$k\geq q+1$ ----(6) となる。また部分自己相関係数$\phi_k(k)$は，(2)式が無限次元次数のAR過程に変換できることより，$Z_t=\phi_1(k)Z_{t-1}+\cdots+\phi_j(k)Z_{t-j}+\cdots+\phi_k(k)Z_{t-k}$ ----(7) とし，Durbin (1960)の方法にて解くと，$\phi_1(1)=\rho_1$，$\phi_2(2)=(\rho_2-\rho_1^2)/(1-\rho_1^2)$，$\cdots\phi_k(k)=\{\rho_k-\sum_{j=1}^{k-1}\phi_j(k-1)\rho_{k-j}\}/\{1-\sum_{j=1}^{k-1}\phi_j(k-1)\rho_j\}$ ----(8) となる。

パラメータの定常性かつ反転性条件は，$\phi_1\sim\phi_p$および$\theta_1\sim\theta_q$を係数とする特性方程式 $u^p-\phi_1u^{p-1}-\cdots-\phi_p=0$ ---(9)　$v^q-\theta_1v^{q-1}-\cdots-\theta_q=0$ ---(10) の各々の根が単位円内に存在することである。こうして，根uについては，$|u_i|<1$，$i=1,\cdots,p$ ---(11)，根vについては，$|v_i|<1$，$i=1,\cdots,q$ ---(12) となる。

3．ARMA(p, q)モデルの確立手法　N年間の水文量系列データy_tを用いて，水工計画に供しうるARMA(p, q)モデルを確立するには，次の手順が考えられる。

(1) 生データの予備解析　a．生データy_tの正規性を検討する。これには，χ^2−検定，歪み度検定が考えられる　b．生データが非正規の場合には，正規化（対数変換など）を試み，再度正規性検定をする。

(2) モデルの型および次数p, qの決定　a．生データy_tの自己相関係数r_kを，$r_k=\sum_{t=1}^{N-k}(y_t-\hat{\mu})(y_{t+k}-\hat{\mu})/\sum_{t=1}^{N}(y_t-\hat{\mu})^2$；$\hat{\mu}=\sum_{t=1}^{N}y_t$ ---(13) にて計算する。　b．部分自己相関係数$\hat{\phi}_k(k)$を，(8)式のρ_kを(13)式のr_kとして求める。　c．r_k，$\hat{\phi}_k(k)$とモデルのρ_k，$\phi_k(k)$とを比較検討し，モデルの型と次数の決定を試みる。

(3) パラメータ$\phi_1\sim\phi_p$，$\theta_1\sim\theta_q$の推定および適合条件の満足性の検討　$\phi_1\sim\phi_p$，$\theta_1\sim\theta_q$の推定は，まず予備推定を行ない，その結果を初期値とする最小二乗法推定により精度向上を計る。なお分散σ^2の推定は，$\hat{\sigma}^2=\frac{1}{N-1}\sum_{t=1}^{N}(y_t-\hat{\mu})^2$ ----(14)による。　a．予備推定法は，データ$Z_t[=(y_t-\hat{\mu})/\hat{\sigma}]$の自己共分散$C_k=\frac{1}{N}\sum_{t=1}^{N-k}(Z_t-\bar{Z}_t)(Z_{t+k}-\bar{Z}_t)$；$\bar{Z}_t=\frac{1}{N}\sum_{t=1}^{N}Z_t$ ---(15) を計算し，(4)式のγ_kをC_kとおきかえて，$\hat{\phi}_1\sim\hat{\phi}_p$を算定する。次に(2)式を，$Z_t'=Z_t-\hat{\phi}_1Z_{t-1}-\cdots-\hat{\phi}_pZ_{t-p}$ ---(16) とし，このZ_t'をMA過程とみなし，自己共分散$C_k'=\sigma_\varepsilon^2\sum_{j=0}^{q-k}\hat{\theta}_j\hat{\theta}_{j+k}$，$k\leq q$ ---(17)，$C_k'=0$，$k>q$ ---(18) および分散$C_0'=\sigma_\varepsilon^2\sum_{j=0}^{q}\hat{\theta}_j^2$ ---(19) とより，$\hat{\theta}_1\sim\hat{\theta}_q$を算定する。　b．最小二乗推定法は，a．にて求めたパラメータを用いて，(2)式の残差ε_tを次の様に計算する。$\varepsilon_j=0$，$j=1,\cdots\max(p,q)$ ---(20)，たとえば$p>q$ならば，$\varepsilon_{p+j}=-\sum_{i=1}^{p}\hat{\phi}_iZ_{p+j-i}+\sum_{i=1}^{q}\hat{\theta}_i\varepsilon_{p+j-i}$，$j=1,\cdots,(N-p)$ ---(21)。次に二乗和$S=\sum_{t=1}^{N}\varepsilon_t^2$ ---(22) を計算する。この様にして，Sが最小となる$\hat{\phi}$，$\hat{\theta}$の組を探索する。　c．求めた$\hat{\phi}_1\sim\hat{\phi}_p$，$\hat{\theta}_1\sim\hat{\theta}_q$による特性方程式(9)，(10)の根が，(11)，(12)式を満足するかどうかを調べる。満足しない場合には，次数あるいはモデルそのもの

を検討する必要がある.

(4) モデルの適合性の検討　a. 確立されたモデルの残差 ε_t の独立性の検定を, Porte Monteau Test により行なう. すなわち, ε_t の自己相関係数 $r_k(\varepsilon)$ より $Q=N\sum_{k=1}^{L}[r_k(\varepsilon)^2]$, $L=N/10+p+q$ ---(23) を計算し, もし $Q<\chi^2(L-p-q)$ ならば, ε_t の独立性が成立する.　b. 同様に残差 ε_t の正規性について, χ^2-検定あるいは歪み度検定を行なう. c. 確立されたモデルが, 他の次数をもつモデルに対し適切なものであるかを, 赤池情報基準 $AIC(p,q)=N\ln(\hat{\sigma}_\varepsilon^2)+2(p+q)$ ---(24) にて検討する. すなわち, 異なる p, q に対して, AICの値が最小値を示すモデルを最適とする.　d. 決定されたモデルにより, データを模擬発生させ, このデータの統計量の諸性質と, 生データのそれらとを比較し, モデルの有意性を検討する.

4. 適用例

1954～1976年(23年間)の石淵ダム(胆沢川)および1953～1976年(24年間)の相模ダム(相模川)への年流入量データを用いて, 3. での手法による時系列モデルの確立を試みる. 表-1に流入量データを, 表-2には, 流入量データの自己相関係数 r_k, 部分自己相関係数 $\hat{\phi}_{kk}(k)$ および Anderson(1941) による独立性のための95%確率範囲 $r_k(95\%)$ $(\approx\pm1.96/N)$ を示す.

表-1 石淵ダム(I)および相模ダム(S)への年流入量(m^3/s)

年＼ダム	I	S	年＼ダム	I	S	年＼ダム	I	S	年＼ダム	I	S
1953		11763.5	1959	5135.53	23005.96	1965	5836.32	14222.02	1971	4207.26	15263.24
54	3904.7	13902.2	60	4309.46	14627.90	66	5336.27	18593.05	72	4952.51	18589.27
55	4721.6	13366.3	61	5173.03	15454.53	67	4231.21	11285.85	73	3868.21	11333.04
56	4837.1	16890.2	62	4717.57	12623.17	68	4677.42	15443.33	74	6120.22	19532.53
57	4564.86	13641.20	63	3970.06	12398.65	69	4451.20	14594.85	75	4408.91	15640.20
58	4885.23	18084.35	64	4788.11	10786.67	70	4217.98	14463.27	76	4403.27	13926.10

表-2 r_k, $\hat{\phi}_{kk}(k)$ および $r_k(95\%)$

		I		S		
k	95% 下限値	r_k	$\hat{\phi}_{kk}(k)$	r_k	$\hat{\phi}_{kk}(k)$	95% 上限値
1	-0.409	-0.193	-0.193	-0.063	-0.063	0.409
2	〃	-0.143	-0.188	0.079	0.075	〃
3	〃	0.022	-0.051	-0.067	-0.058	〃
4	〃	-0.053	-0.093	-0.200	-0.215	〃
5	〃	-0.168	-0.222	-0.340	-0.377	〃

(1) 流入量データの歪み係数 $\hat{\gamma}$ は, I = 2.840, S = 1.030 である. 有意水準 $\alpha=0.02$ に対する歪み度検定表値 $\gamma_\alpha(N)$ (Snedecor & Corchran, 1967) = 1.061 となり, $\hat{\gamma}<\gamma_\alpha(N)$ より両データは正規分布をなしているとみなす.　(2) 表-2の結果および種々の次数 p, q 値に対する ARMA(p,q) モデルの残差分散 $\hat{\sigma}_\varepsilon^2$ を比較検討し, ここでは, I, S とも時系列モデルとして ARMA(0,1) モデルを選定する.

(3) パラメータ θ_1 の推定値と残差分散値 $\hat{\sigma}_\varepsilon^2$ を表-3に示す. I, Sどちらの $\hat{\theta}_1$ も転換性条件を満足している.

表-3 $\hat{\theta}_1$ と $\hat{\sigma}_\varepsilon^2$

	$\hat{\theta}_1$	$\hat{\sigma}_\varepsilon^2$
I	0.192	0.946
S	0.061	0.996

表-4 独立性検定

計算値	$\chi^2_{0.95}$ 値
1.879	9.490
4.863	11.071

表-5 歪み度検定

計算値	$\gamma_{0.02}$ 値
0.929	1.061
1.016	1.061

表-6 AIC-値

ARMA(0,1)	ARMA(1,1)	ARMA(0,2)
0.7560	1.8370	2.3373
1.9120	4.7955	4.6942

表-7 模擬データと生データの諸統計量

	データ＼統計	平均値	標準偏差	歪み係数	r_1	r_2
I	模擬	4679.230	586.451	0.137	-0.144	-0.011
	生	4657.293	591.687	0.840	-0.193	-0.143
S	模擬	14957.460	2977.232	0.091	-0.029	-0.006
	生	14830.450	2965.217	1.030	-0.063	0.079

(4) モデルの残差 ε_t の独立かつ正規性検定として, (23)式の計算値と $\chi^2(95\%)$ 値を表-4に, 歪み係数値と $\gamma_{0.02}$ 値を表-5に示す. 両表より ε_t は独立かつ正規であるとみなせる. 表-6には, ARMA(0,1)モデルとその近傍次数のそれらの計算結果を示す. 選定したモデルのAIC値が最小となっている. ARMA(0,1)モデルにより, Iでは23年間の, Sでは24年間の模擬データを各々30個発生させ, 平均値, 標準偏差, 歪み係数および $k=1,2$ 年の自己相関係数に関しての平均値を, 生データのそれらとを比較して表-7に示してある. 歪み係数の再現性に難はあるが, その他は概ね良好である.

5. おわりに

ダムへの年流入量を定常過程とみなし, ARMAモデルを適用した結果, 統計的にはほぼ満足すべき結論を得た. 実際的には, このモデルによる水文量の将来予測の可能性にあろう. このことと, 水工計画上一層有用となる月, 日単位のARMAモデルについては, 今後検討していきたい.
最後に, 流入量データは, 建設省土木研究所の今村, 横道両氏が整備提供して下さった. 記して深謝致します.

〔参考文献〕 J.D.Salas, et al, "Applied Modeling of Hydrologic Time Series", WRP, 1980 ; Box & Jenkins, "Time Series Analysis", H-Day, 1976 ; 山本, "定常時系列モデルの選定および次数の決定", 才9回自然災害研究発表会