

آشنایی با شبیه سازی سریهای زمانی خطی و غیر خطی

سید مهدی امیر جهانشاهی^۱ سید محمد امیر جهانشاهی^۲

چکیده

در این مقاله، ابتدا با روشهای شبیه سازی الگوهای سری زمانی خطی توسط نرم افزارهای S-PLUS و SAS آشنا می شویم و در ادامه به معرفی الگوهای سری زمانی غیرخطی از جمله الگوهای دوخطی^۳، اتورگرسیو تغییر وضعیت هموار^۴، اتورگرسیو مشروط به ناهمگنی واریانس^۵ و اتورگرسیو نمایی^۶ می پردازیم، و با روشهای شبیه سازی این الگوهای سریهای زمانی غیرخطی توسط نرم افزار SAS آشنا می شویم و در مورد هر الگو به بررسی کارایی توابع ACF، PACF و IACF در تشخیص الگوی مورد نظر می پردازیم. در انتهای مقاله نیز به اختصار چند روش از روشهای تشخیص سریهای زمانی غیرخطی را معرفی می کنیم.

۲. شبیه سازی سریهای زمانی خطی

در ابتدا به شبیه سازی صد مشاهده از یک الگوی $ARMA(1,1)$ می پردازیم و برای انجام این کار از شیوه IML و تابع ARMASIM در SAS به صورت زیر استفاده می کنیم:

```
* y(t) = 0.5y(t-1) + e(t) + 0.8e(t-1) ;
data arma ;
proc iml ;
  phi = {1 -0.5} ;
  theta = {1 0.8} ;
  y = armasim(phi, theta, 0, 1, 100, 1234321) ;
  print y ;
run ;
```

تکته - مقادیر ضرایب الگو طوری در نظر گرفته شده اند که

۱. مقدمه

تا به حال یک فرض مهم که در نظریه های سری زمانی به طور کلاسیک در نظر گرفته شده، این است که ساختمان سریها می تواند به وسیله یک مدل خطی توصیف شود. فرض خطی بودن، غالباً مشکوک است و فرایندهای زیادی که در طبیعت، در زمینه های مهندسی فیزیک اتفاق می افتند، نمایش شکلی از رفتار غیرخطی است. این حقیقت در سالهای اخیر موجب بسط و به وجود آمدن تعداد زیادی از مدل های جالب غیرخطی سری زمانی شده است. از جمله این الگوها می توان به الگوهای دوخطی، اتورگرسیو نمایی و اتورگرسیو آستانه ای و ... اشاره کرد. معمولاً از شبیه سازی برای آشنایی با خصوصیات و رفتارهای خاص این الگوهای غیرخطی استفاده می شود، و نتایج حاصل را، در برخورد با داده های واقعی غیرخطی به کار می برند.

^۴ Smooth Transition Autoregressive(STAR)

^۵ GARCH

^۶ Exponential Autoregressive(EAR)

^۱ عضو هیئت علمی، دانشگاه بیرجند

^۲ کارشناس ارشد آمار ریاضی، دانشگاه فردوسی مشهد

^۳ Bilinear

در شکل (۳) نمودار ACF به صفر میل می کند و نمودارهای PACF و IACF از مرتبه ۲ به بعد قطع شده اند. بنابراین کاملاً خصوصیات یک الگوی اتورگرسیو مرتبه ۲ را نشان می دهند [۱].
در ادامه با طریقه شبیه سازی الگوهای $ARIMA(0,1,1)$ و $ARMA(1,1)$ آشنا می شویم. الگوهایی که شبیه سازی خواهند شد، عبارت اند از:

$$ARIMA(0,1,1): X_t = X_{t-1} + \varepsilon_t + 0.6\varepsilon_{t-1}$$

$$ARMA(1,1): X_t = \varepsilon_t - 0.6\varepsilon_{t-1} + 0.5X_{t-1}$$

Simulate an ARMA(1,1)

```
> arma <- arima.sim(n=1000,model=list(ar=0.5, ma
=-0.6))
> print(arma)
> lag.plot(arma, lags=9, layout=c(3,3))
> par(mfrow=c(1,3))
> tsplot(arma)
> acf(arma, lag.max=40, type='correlation')
> acf(arma, lag.max=40, type='partial')
```

Simulate an ARIMA(0,1,1)

```
> arima <- arima.sim(n=1000,model=list(ndiff=1,
ma=0.6))
> print(arima)
> lag.plot(arima, lags=9, layout=c(3,3))
> par(mfrow=c(1,3))
> tsplot(arima)
> acf(arima, lag.max=150, type='correlation')
> acf(arima, lag.max=150, type='partial')
```

در شکل (۴) نمودارهای مربوط به سریهای زمانی شبیه سازی شده را ملاحظه می کنید. نمودارهای ACF و PACF هر دو به صفر میل می کنند. بنابراین خصوصیات یک الگوی اتورگرسیو میانگین متحرک مرتبه ۱ و ۱ را نشان می دهند [۱].

در شکل (۵) نمودارهای مربوط به الگوی $ARIMA(0,1,1)$ را ملاحظه می کنید. نمودار ACF به کندی به صفر میل می کند و نمودار PACF نیز با یک ضریب خود همبستگی جزئی بسیار نزدیک به یک در تأخیر ۱ به صفر میل می کند، که این خود دلیل بر نامانایی سری زمانی شبیه سازی شده است [۱].

شرطهای مانایی و وارون پذیری فرایند صادق باشند.

همان طور که در شکل (۱) ملاحظه می شود نمودارهای ACF، PACF و IACF همگی از مرتبه یک به بعد قطع شده اند که این خود دلیلی بر پیروی مشاهدات از یک فرایند $ARMA(1,1)$ است [۱]. در زیر به طور مختصر با طریقه شبیه سازی الگوهای $MA(2)$ و $AR(2)$ به وسیله نرم افزار S-PLUS آشنا می شویم. به طور کلی برای شبیه سازی سریهای زمانی در S-PLUS از تابع `arma.sim` استفاده می کنیم. تابع `lag.plot` نمودار تأخیرهای مختلف سری را در مقابل یکدیگر رسم می کند، تابع `tsplot` نمودار سری زمانی را رسم می کند و تابع `acf` نمودارهای ACF و PACF را رسم می کند. الگوها به ترتیب عبارت اند از:

$$MA(2): X_t = -0.5\varepsilon_{t-1} - 0.25\varepsilon_{t-2} + \varepsilon_t$$

$$AR(2): X_t = -0.5X_{t-1} - 0.25X_{t-2} + \varepsilon_t$$

Simulate a MA(2)

```
> Ma2 <- arima.sim(n=1000,model=list(ma=c(-0.5,-
0.25)))
> print(Ma2)
> lag.plot(Ma2, lags=9, layout=c(3,3))
> par(mfrow=c(1,3))
> tsplot(Ma2)
> acf(Ma2, lag.max=40, type='correlation')
> acf(Ma2, lag.max=40, type='partial')
```

Simulate an AR(2)

```
> Ar2 <- arima.sim(n=1000,model=list(ar=c(-0.5,-
0.25)))
> print(Ar2)
> lag.plot(Ar2, lags=9, layout=c(3,3))
> par(mfrow=c(1,3))
> tsplot(Ar2)
> acf(Ar2, lag.max=40, type='correlation')
> acf(Ar2, lag.max=40, type='partial')
```

نمودارهای ACF، PACF و IACF مربوط به الگوهای مذکور را در شکل (۲) ملاحظه می کنید. نمودار ACF از مرتبه ۲ به بعد قطع شده و نمودارهای PACF و IACF به صفر میل می کنند. بنابراین کاملاً خصوصیات یک الگوی میانگین متحرک مرتبه ۲ را نشان می دهند [۱].

$$X_t = 0.5X_{t-1}Z_{t-2} + Z_t$$

$$X_t = 0.5X_{t-1}Z_{t-1} + Z_t$$

$$X_t = 0.5X_{t-2}Z_{t-1} + Z_t$$

در برنامه زیر به شبیه سازی صد مشاهده از الگوی دوخطی مارکوفی $X_t = 0.3X_{t-1} + 0.4X_{t-1}Z_t + Z_t$ به وسیله نرم افزار SAS می پردازیم:

* simulation: $X(t)=0.3X(t-1) + 0.4X(t-1)Z(t) + Z(t)$;

data bilin;

array xx{501} x1-x501;

xx{1} = 0;

array zz{501} z1-z501;

zz{1} = 0;

do t = 2 to 501;

zz{t} = normal(856509);

xx{t}=0.3*xx{t-1}+0.4*xx{t-1}*zz{t-1}+zz{t};

end;

output;

run;

data markov;

set bilin;

file 'd:\data.txt';

put x2-x501;

run;

data new;

infile 'd:\data.txt';

input x @@;

t+1;

run;

proc gplot;

plot x*t/frame vref = 0;

symbol i = join;

run;

proc arima;

identify var = x MINIC;

run;

همان طور که در شکل (۶) مشاهده می شود نمودار ACF به صفر میل می کند و نمودار توابع PACF و IACF از مرتبه ۱ به بعد قطع می شوند. بنابراین با توجه به توابع مذکور، الگو را به اشتباه $AR(1)$ تشخیص می دهیم [۱].

در این قسمت نیز به شبیه سازی الگوی بالاقطری $BL(0,0,2,1)$ می پردازیم:

۳. شبیه سازی سریهای زمانی غیرخطی

الگوی کلی میانگین متحرک اتورگرسیو دوخطی اولین بار توسط موهلر^۷ ارائه شد و به طور خلاصه با $BL(p,q,r,s)$ نشان داده می شود:

$$X_t = \sum_{i=1}^p \alpha_i X_{t-i} + \sum_{j=1}^q \beta_j Z_{t-j} + \sum_{k=1}^r \sum_{h=1}^s \gamma_{kh} X_{t-h} Z_{t-k} + Z_t$$

که در آن $\{Z_t\}$ ها متغیرهای تصادفی مستقل هموزیع با میانگین صفر و واریانس σ_z^2 هستند، $(Z_t \sim iid(0, \sigma_z^2))$. این الگو به چند حالت خاص، که در زیر معرفی می کنیم، تقسیم می شود:

۱- مارکوفی $MBL(p,q)$:

$$X_t = \sum_{i=1}^p a_i X_{t-i} + \sum_{j=1}^q b_j X_{t-j} Z_t + Z_t$$

$$X_t = \sum_{k=1}^r \sum_{h=1}^s \gamma_{kh} X_{t-h} Z_{t-k} + Z_t \quad \text{۲- دوخطی کامل:}$$

۳- دوخطی کامل بالا قطری^۸:

$$X_t = \sum_{k=1}^r \sum_{h=1}^s \gamma_{kh} X_{t-h} Z_{t-k} + Z_t \quad ; h > k$$

۴- دوخطی کامل قطری^۹:

$$X_t = \sum_{k=1}^r \sum_{h=1}^s \gamma_{kh} X_{t-h} Z_{t-k} + Z_t \quad ; h = k$$

۵- دوخطی کامل زیر قطری^{۱۰}:

$$X_t = \sum_{k=1}^r \sum_{h=1}^s \gamma_{kh} X_{t-h} Z_{t-k} + Z_t \quad ; h < k$$

که به عنوان مثال، به الگوهای زیر قطری، قطری و بالا قطری زیر می توان اشاره کرد:

Mohler^۷

Super Diagonal^۸

Diagonal^۹

Sub Diagonal^{۱۰}

که در آن،

$$w_t = (X_{t-1}, \dots, X_{t-p}) \text{ و } E(Z_t W_t) = 0, Z_t \sim iid(0, \sigma^2)$$

و بردارهای پارامترهای ثابت الگو عبارت اند از:

$$\theta' = (\theta_1, \dots, \theta_p)'$$

$$\phi' = (\phi_1, \dots, \phi_p)'$$

غیرخطی بودن الگو نیز به وسیله تابع انتقال $F(\cdot)$ ایجاد می شود که

معمولاً به یکی از دو صورت زیر در نظر گرفته می شود:

۱- نمایی که در نتیجه الگو ESTAR نامیده می شود:

$$F(\gamma, X_{t-d}, c) = 1 - \exp\{-\gamma(X_{t-d} - c)^\gamma\}; \gamma > 0$$

۲- لوزستیک که در نتیجه الگو LSTAR نامیده می شود:

$$F(\gamma, X_{t-d}, c) = \frac{1}{1 + \exp\{-\gamma(X_{t-d} - c)\}}; \gamma > 0$$

الگو را به صورت غیربرداری زیر نیز می توان مشخص کرد:

$$X_t = \theta_0 + \sum_{i=1}^p \theta_i X_{t-i} +$$

$$F(\gamma, X_{t-d}, c)(\phi_0 + \sum_{i=1}^p \phi_i X_{t-i}) + Z_t$$

به عنوان مثال الگوی $ESTAR(2, 1, 0.2)$ به صورت زیر مشخص

می شود:

$$X_t = 1/8 X_{t-1} - 1/6 X_{t-2} + (1/2 - 1/9 X_{t-1} + 1/79 X_{t-2})[1 - \exp\{-1/3(X_{t-1} - 1/2)^\gamma\}] + Z_t$$

برای شبیه سازی این الگو از برنامه زیر می توان استفاده کرد:

```
data starp;
  x1=0.01;
  x2=0.02;
  do t=1 to 500;
    z=normal(76574598);
    x=1.8*x1-1.06*x2+(0.02-0.9*x1+0.79*x2)*(1-
    exp(-0.3*(x1-0.02)**2))+z;
  output;
  x2=x1;
  x1=x;
  end;
  proc gplot;
    plot x*t/frame;
    symbol i=join;
  run;
```

* simulation super diagonal : $X(t) = 0.5X(t-2)Z(t-1) + Z(t)$

option ps = 400;

data bilin;

array xx{502} x1-x502;

xx{1} = -0.26; xx{2} = 0.26;

array zz{502} z1-z502;

zz{1} = -0.45264; zz{2} = -0.31599;

do t=3 to 502;

zz{t} = normal(856509);

xx{t} = 0.5*xx{t-2}+zz{t-1}+zz{t};

end;

output;

run;

data sup;

set bilin;

file 'd:\data1.txt';

put x3-x502;

run;

data new;

infile 'd:\data1.txt';

input x @@;

t+1;

run;

proc gplot;

plot x*t/frame vref=0;

symbol i=join;

run;

proc arima;

identify var=x MINIC;

run;

همان طور که در شکل (۷) ملاحظه می شود در نمودار توابع

ACF، PACF و IACF ضریبی معنی دار مشاهده نمی شود. بنابراین

بر اساس توابع مذکور، الگو را به اشتباه جزو خانواده فرایند تصادفی

محض تشخیص می دهیم [۱].

۴. الگوی $STAR(p, d, c)$:

این الگو در ابتدا توسط باکن و واتر^{۱۱} ارائه شد و در ادامه توسط

تراسویرتا^{۱۲} تعمیم داده شد. الگوی عمومی $STAR(p, d, c)$ است

که به صورت برداری زیر تعریف می شود:

$$X_t \theta_0 + \theta' w_t + (\phi_0 + \phi' w_t) F(\gamma, X_{t-d}, c) + Z_t$$

^{۱۱} Baken and Wats

^{۱۲} Terasvirta

```

option nodate nonumber ps =300;
data jahan;
  warning = 'HT < 0';
  ht1 = 0;
  ht2 = 0;
  eps1 = 0;
  eps2 = 0;
  do time = 1 to 600;
    ht = 0.2 + 0.4*eps1+ 0.1*eps2*eps2 +
      0.3*ht1+ 0.1*ht2;
    if ht < 0 then put warning = ;
    eps=sqrt(abs(ht))*normal(529);
    y = eps;
    y2 = y**2;
    output;
    eps2 = eps1;
    eps1 = eps;
    ht2 = ht1;
    ht1 = ht;
  end;
run;
pro autoreg data = jahan;
  model y = / noint garch = (p=2,q=2) method = ml
archtest;
  model y = / noint garch = (p=1,q=1) method = uls;
run;

```

ضمناً برای برآورد پارامترهای الگوهای $GARCH(2,2)$ و $GARCH(1,1)$ که به سری شیه سازی شده برازش می‌دهیم از روشهای درستمایی ماکسیم (ML) و کمترین توانهای دوم غیرشرطی (ULS) استفاده شده است. همچنین آزمون بررسی وجود الگوی ARCHTEST مشروط به ناهمگنی واریانس نیز با استفاده از سوئیچ انجام می‌شود.

همان طور که در شکل (۹) ملاحظه می‌شود در نمودارهای توابع ACF، PACF و IACF ضریبی معنی دار مشاهده نمی‌شود. بنابراین بر اساس توابع مذکور الگو را به اشتباه جزو خانواده فرایند تصادفی محض تشخیص می‌دهیم [۱].

۶. الگوهای اتورگرسیو نمایی

مدلهای نمایی توسط اوزاکی و هاگان^{۱۵} برای بیان شکل‌های معینی از نوسانات تصادفی غیرخطی معرفی شده‌اند. به طور کلی یک الگوی $EAR(p,d)$ به صورت زیر مشخص می‌شود:

```

proc arima;
  identify var = x minic esacf;
run;

```

همان طور که در شکل (۸) ملاحظه می‌شود نمودار ACF به صفر میل می‌کند و نمودار توابع PACF و IACF از مرتبه ۲ به بعد قطع می‌شوند. بنابراین با توابع مذکور الگو را به اشتباه $AR(2)$ تشخیص می‌دهیم [۱].

۵. الگوی $GARCH(p,q)$:

الگوی اتورگرسیو مشروط به ناهمگنی واریانس تعمیم یافته توسط بولرسلف^{۱۳} ارائه شد. در این الگو فرض می‌شود که $X_t = \varepsilon_t$ باشد که در آن

$$\varepsilon_t | \psi_{t-1} \sim N(0, h_t), \quad \varepsilon_t = h_t^{1/2} Z_t; Z_t \sim iid(0,1)$$

$$Var(\varepsilon_t | \psi_{t-1}) = h_t, \quad E(\varepsilon_t | \psi_{t-1}) = 0$$

که ψ_{t-1} سیگما میدانی از اطلاعات تا لحظه $t-1$ است و h_t واریانس شرطی در لحظه t و تابعی همواره مثبت است که به صورت زیر مشخص می‌شود:

$$Var(\varepsilon_t | \psi_{t-1}) = h_t = \gamma + \sum_{j=1}^q \alpha_j \varepsilon_{t-j}^2 + \sum_{i=1}^p \beta_i h_{t-i}$$

که در آن، $\alpha_i \geq 0; i=1, \dots, q, \alpha_0 > 0, p > 0, q > 0$ و $\beta_j \geq 0; j=1, \dots, p$ و شرط مانایی فرایند $\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j < 1$ است و واریانس غیرشرطی نیز به صورت زیر است:

$$\sigma^2 = \alpha_0 / (1 - \sum_{i=1}^q \alpha_i - \sum_{j=1}^p \beta_j)$$

ARCH^{۱۴} به معنی اتورگرسیو مشروط به ناهمگنی واریانس است. برای شیه سازی و برآورد پارامترهای الگوی $GARCH(2,2)$ که به صورت زیر است:

$$X_t = \varepsilon_t, \quad \varepsilon_t = h_t^{1/2} Z_t, \quad Z_t \sim N(0,1)$$

$$h_t = 0.2 + 0.4\varepsilon_{t-1}^2 + 0.1\varepsilon_{t-2}^2 + 0.3h_{t-1} + 0.1h_t$$

از برنامه زیر می‌توان استفاده کرد:

^{۱۳} Bolerself

^{۱۴} Conditional Heteroscedasticity Autoregressive

^{۱۵} Osaki and Hogon

```
proc arima;
  identify var = x minic esacf;
run;
```

طبق شکل (۱۰) نمودار ACF به صفر میل می کند و نمودار توابع PACF و IACF از مرتبه ۲ به بعد قطع می شوند. بنابراین با توابع مذکور الگو را به اشتباه $AR(2)$ تشخیص می دهیم [۱].

تکته - در سری زمانی کلاسیک، توابع خود همبستگی نمونه (SACF)، خود همبستگی جزئی نمونه (SPACF)، و خود همبستگی وارون نمونه (SIACF) برای تشخیص مرتبه الگوهای خطی موارد استفاده گسترده ای دارند، ولی در تشخیص سریهای زمانی غیرخطی از آنها استفاده چندانی نمی شود، زیرا آنها فقط توانایی توصیف تغییرات در الگوهای خطی را دارند و بر این اساس در تشخیص سریهای زمانی غیرخطی، همان طور که در تحلیل نمودار توابع مذکور برای الگوهای غیرخطی مختلف ملاحظه کردید کارآمد نیستند.

برای تشخیص الگوهای غیرخطی روشهای دیگری به کار می روند که از آن جمله می توان به استفاده از گشتاورهای مراتب بالاتر (سوم و چهارم) [۵]، استفاده از سری توان دوم مشاهدات [۴]، استفاده از توانهای دوم مانده های الگوی برازش یافته، استفاده از معیار اطلاع آکائیک [۲]، استفاده از روشهای ناپارامتری [۶] (در این روش از تقریب توابع امید ریاضی شرطی و واریانس شرطی استفاده می شود) و به استفاده از توابع LDF، NLDF، PLDF و PRSF [۶] اشاره کرد.

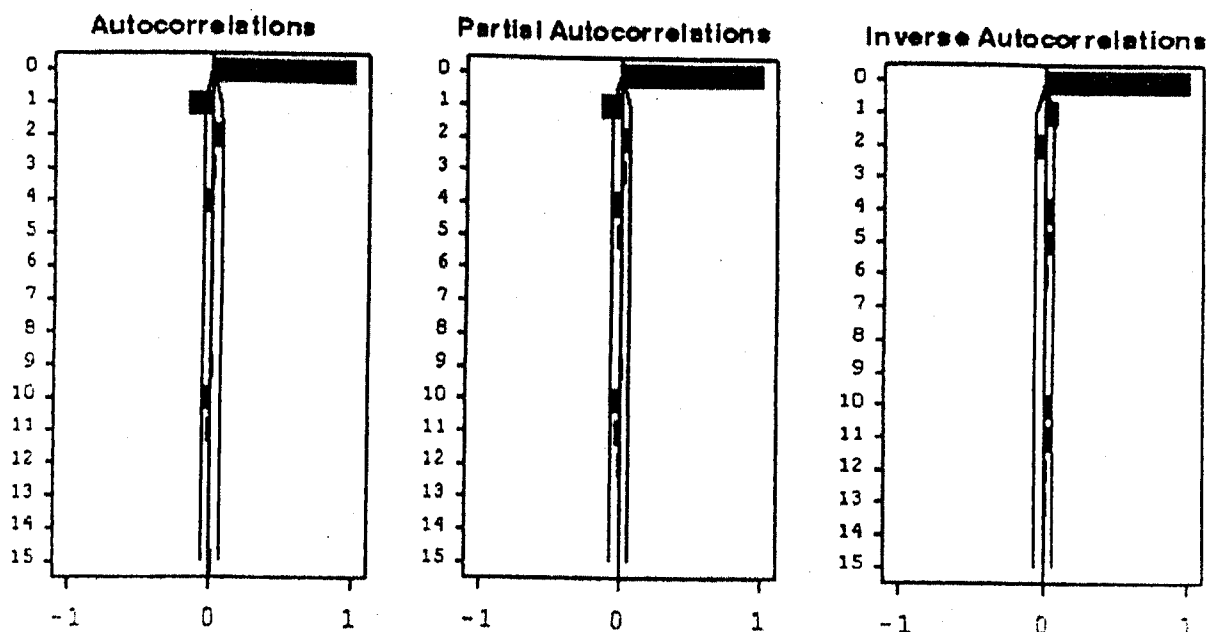
$$X_t = (\phi_1 + \pi_1 e^{-\gamma X_{t-1}^2}) X_{t-1} + \dots + (\phi_p + \pi_p e^{-\gamma X_{t-p}^2}) X_{t-p} + Z_t$$

که در آن، $\phi_1, \dots, \phi_p$ و $\theta_1, \dots, \theta_p$ مقادیر حقیقی هستند و $\gamma > 0$ است. مرتبه p ی الگو با جستجوی مقادیر مینیمم کننده مقدار آکائیک [۱] برای الگوی فوق به دست می آید. به عنوان مثال الگوی $EAR(2,1)$ به صورت زیر است:

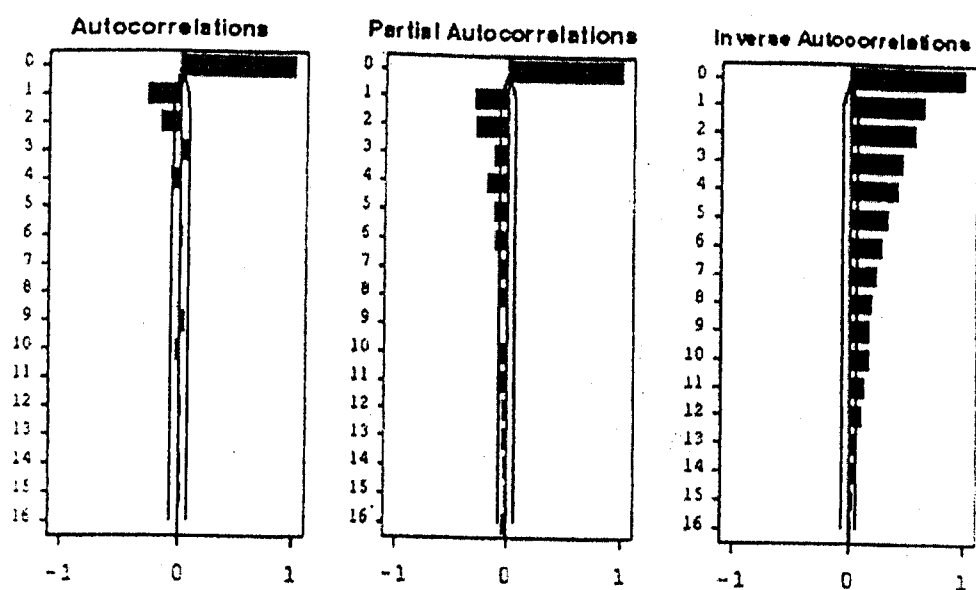
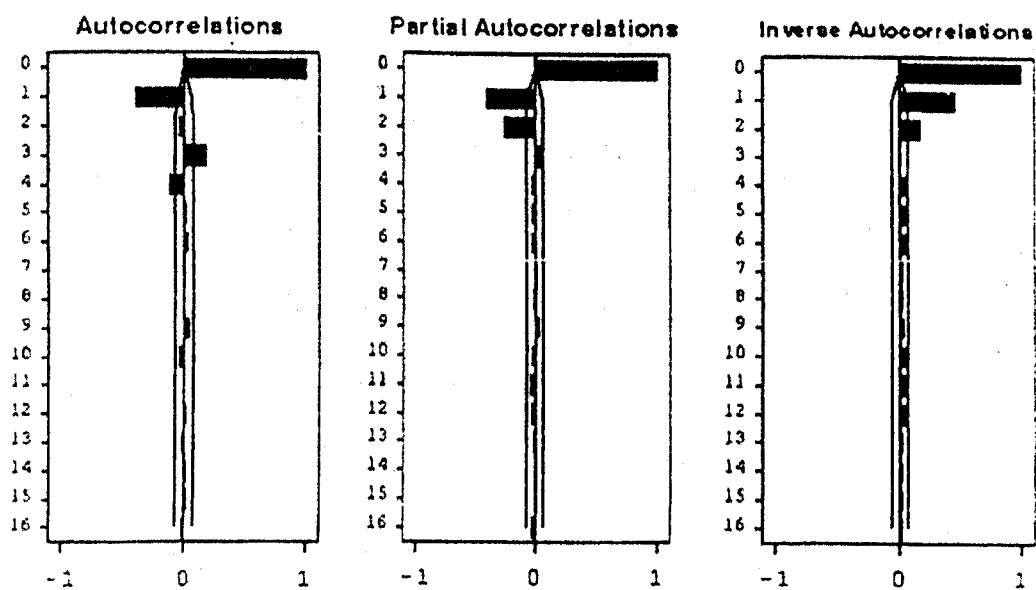
$$X_t = (0.2 + 0.4 e^{-0.1 X_{t-1}^2}) X_{t-1} + (0.3 + 0.5 e^{-0.1 X_{t-1}^2}) X_{t-2} + Z_t$$

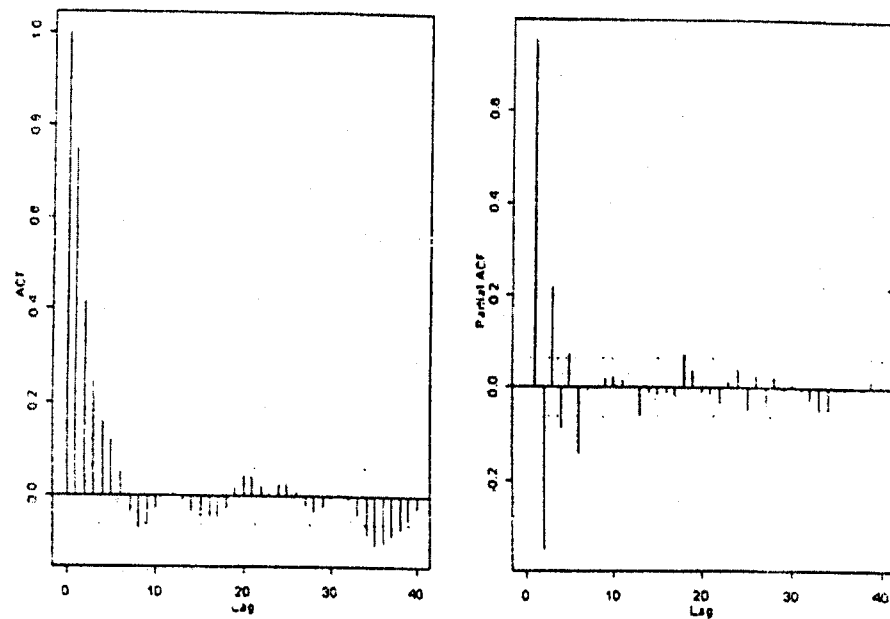
این الگو را می توان توسط برنامه زیر شبیه سازی کرد:

```
data ear2;
  x1=0.01;
  x2=0.02;
  do t=1 to 500;
    z=normal(76574598);
    x=(0.2+0.4*exp(-0.1*x1**2))*x1+
      (0.3+0.5*exp(-0.1*x1**2))*x2+z;
  output;
  x2=x1;
  x1=x;
  end;
proc gplot;
  plot x*t/frame;
  symbol i=join;
run;
```

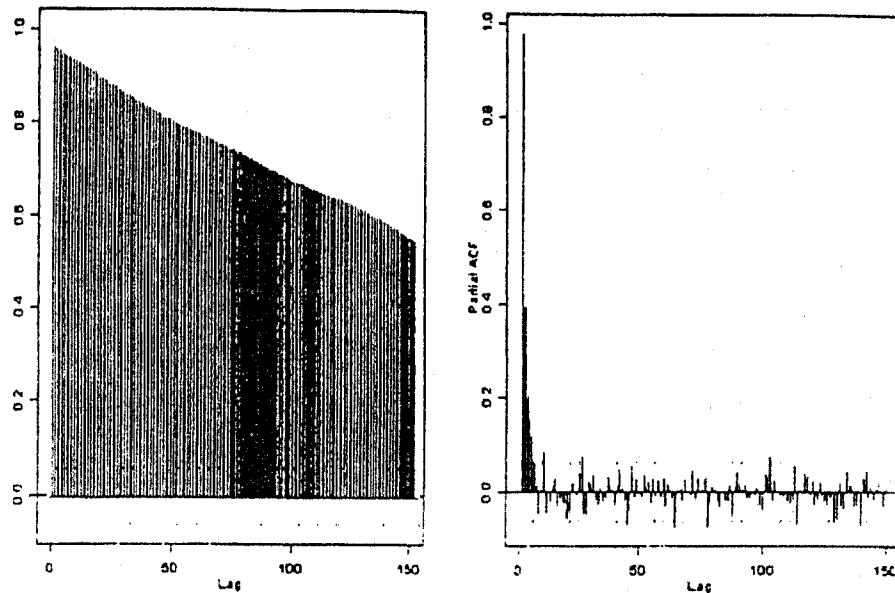


شکل ۱ (مدل $ARMA(1,1)$)

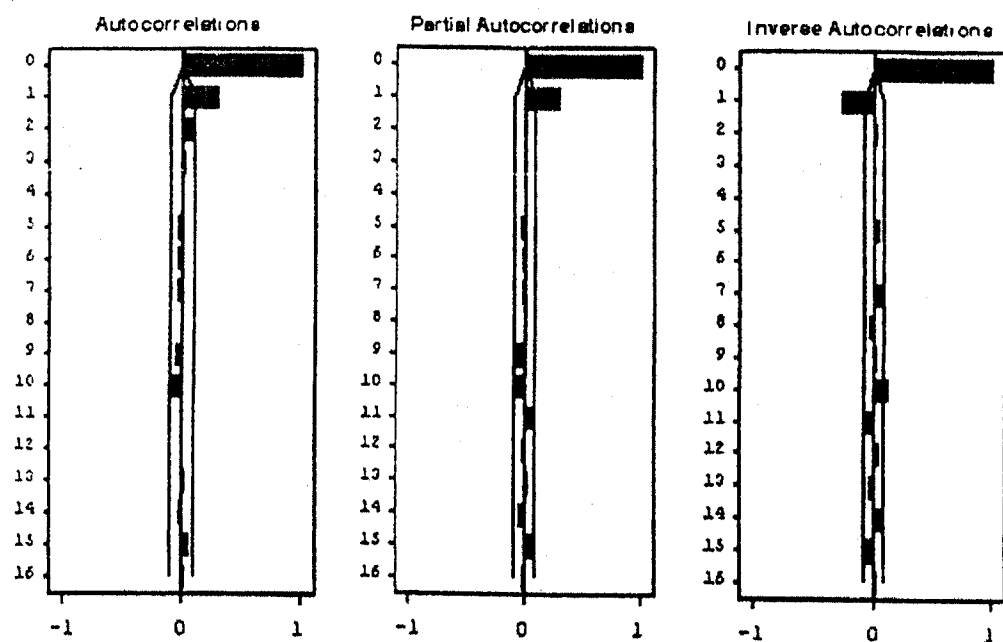
شکل ۲ (مدل $MA(2)$)شکل ۳ (مدل $AR(2)$)



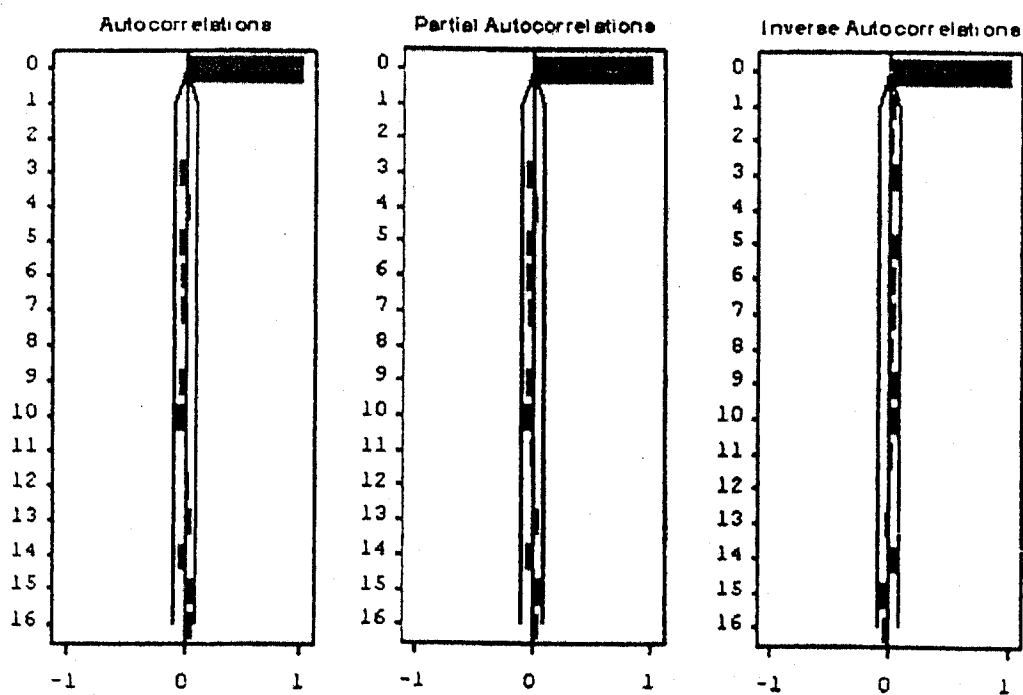
شکل ۴ (مدل $ARMA(1,1)$)



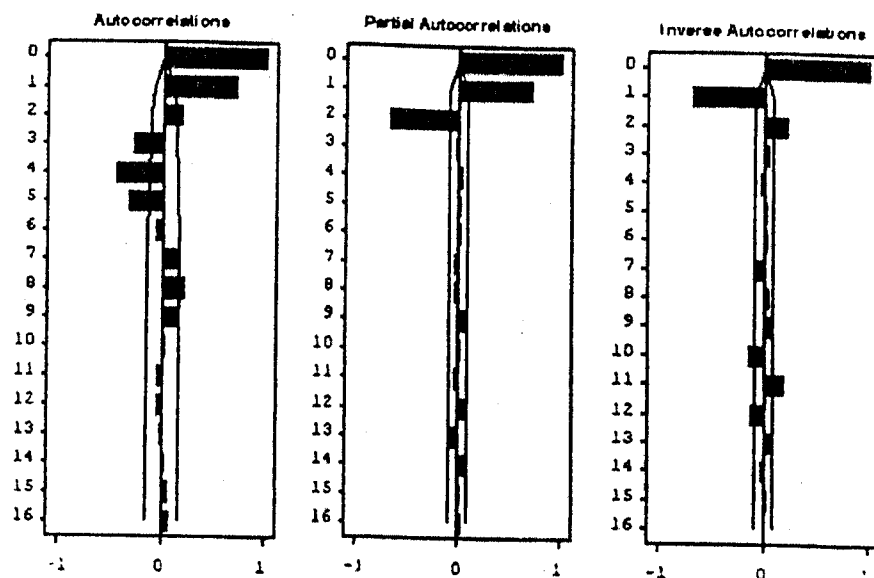
شکل ۵ (مدل $ARMA(1,1)$)



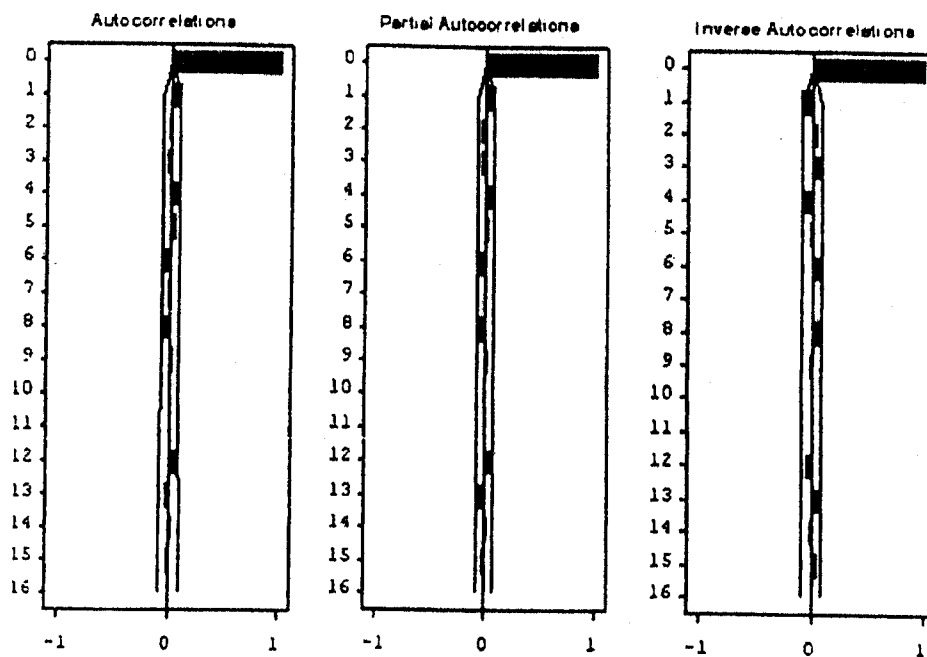
شکل ۶



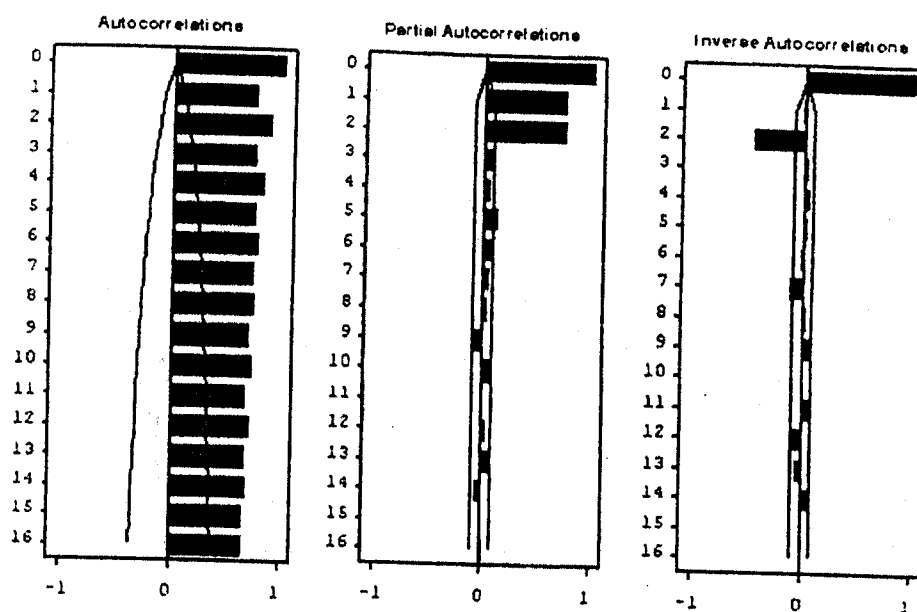
شکل ۷



شکل ۸



شکل ۹



شکل ۱۰

مراجع

[۱] ویلیام دبلیو.اس.وی.، تحلیل سریهای زمانی، روشهای یک متغیری و چند متغیری، ترجمه حسینی نیرمند، ۱۳۷۶.

[2] Subba Rao, T. and Gabr, M.M., (1984), *An Introduction to Bispectral and Bilinear Time Series Models*, Lecture Note in Statistics, New York: Springer-Verlag.

[3] Anderson, M. K., Eklund, B. and Lyhagen, J., (1999), *An ARCH Robust STAR Test*, Working Paper Series in Economics and Finance, No. 317, 42-53.

[4] Granger, C.W.J. and Anderson, A.P., (1978), *An Introduction to Bilinear Time Series Models*, Gottingen: Vandenhoeck and Ruprecht.

[5] Kumar, K., *On the Identification of Some Nonlinear Time Series Models*, J. Time Series Analysis, Vol. 7, No. 2.

[6] Auestad, B., Tjostheim, D., (1990), *Identification of Nonlinear Time Series: First Order Characterization and Order Determination*, Biometrika, Vol. 77, Issue 4, 669-687.