نماذج اريما ARIMA وطريقة بوكس جيكينز:

مقدمة:

نماذج اريماARIMA

السكون

نماذج الانحدار الذاتي للسلسلة الزمنية

نماذج المتوسط المتحركMA

نماذج ارماARMA

السلسلة الزمنية المتكاملة ونماذجARIMA

اختيار النموذج باستخدام طريقة بوكس جينكينز Box-Jenkins :

مثال

مقدمة:

في هذا الفصل سنناقش تقدير معادلة واحدة بطريقة مختلفة عن الفصول السابقة. في تلك الفصول تم مناقشة سلوك المتغير التابع باستخدام عدد من المتغيرات المفسرة . في تحليل السلاسل الزمنية نبدأ تحليل المعلومات التي يمكن التحصيل عليها من المتغير نفسة . تحليل سلسلة زمنية واحدة يسمى سلسلة زمنية احادية المتغير univariate time series في هذا الموضوع الهدف من تحليل السلاسل الزمنية هو اسر واختبار ديناميكية البيانات، في الاقتصاد القياسي للسلاسل الزمنية يمكن ان يكون هناك نماذج سلسلة زمنية متعددة المتغيرات سوف يتم مناقشتها في فصول قادمة.

كما ذكر سابقا الاقتصاد القياسي التقليدي ركز على استخدام النظرية الاقتصادية ودراسة العلاقات المعاصرة من اجل شرح العلاقات بين المتغير التابع والمتغيرات المفسرة. من الآن فصاعدا نستخدم مسمى الاقتصاد القياسي التقليدي لتمييزه عن الاقتصاد القياسي الحديث.

المتغيرات المتباطئة يتم تعريفها بين آونة وأخرى ولكن ليس بمنهجية محددة او على الأقل ليس بطريقة تحاول تحليل الديناميكية او الهيكل الزماني للبيانات. هناك عدة جوانب لتحليل لسلسلة الزمنية ولكن هناك موضوع مشترك لها وهو استخدام كامل للهيكل الديناميكي للبيانات، المقصود هو استخراج كل ما يمكن من معلومات من المعلومات التاريخية للسلسلة الزمنية. المبدأين الأساسين لتحليل السلسلة الزمنية هي التنبؤ والنموذجة الديناميكية. التنبؤ مختلف عن بناء نموذج هيكلي وتفهم الاقتصاد او اختبار فرضية. هو مهتم ببناء نماذج تنبؤ فعالة. تعمل عادة باستغلال العلاقات المتبادلة التي وجدت عبر الزمن لمتغير واحد. النمذجة الديناميكية في الجانب الآخر مهتمة بالبناء الهيكلي للاقتصاد واختبار الفرضيات, وذلك لتفهم العملية يجب ان اسر عملية التكيف التي قد تكون طويلة ومعقدة. منذ بداية الثمانينات اساليب حديثة طورت في التنبؤ لذلك سوف نبدأ هذا الفصل بنماذج اريما ARIMA.

نماذج اريما ARIMA

BOX and Jenkins (1976) عرف نماذج اريما المصطلح يعني:

AR=autoregressive انحدار ذاتي.

I-integrated متكاملة.

MA=moving average المتوسط المتحرك.

الاجزاء التالية ستعرض اصدارات مختلفة لنماذج اريماARIMA وستقدم مفهوم السكون، سنبدأ بشرح ابسط نموذج نموذج الارتباط الذاتي من الدرجة 1 ثم نستمر لمسح نماذج اريماARIMA. واخيرا طريقة بوكس جيكينز لاختيار النموذج ثم يليه عرض للتنبؤ.

السكون:

أي سلسلة زمنية يمكن ان تولد من عملية عشوائية و مجموعه من البيانات مثل البيانات في الجدول التالي لاجمالي الناتج المحلي للسعودية:

GROSS DOMESTIC PRODUCT (GDP)							
السنة	gdp	السنة	gdp	السنة	gdp	السنة	gdp
1970	22.57	1980	546.6	1990	437.33	2000	706.66
1971	30.5	1981	622.18	1991	491.85	2001	686.3
1972	38.26	1982	524.2	1992	510.46	2002	707.07
1973	53.53	1983	445.21	1993	494.91	2003	804.65
1974	159.72	1984	420.39	1994	503.05	2004	938.77
1975	163.67	1985	376.32	1995	533.5	2005	1182.51
1976	225.35	1986	322.02	1996	590.75	2006	1335.58
1977	260.96	1987	320.93	1997	617.9	2007	1442.57
1978	272.27	1988	330.52	1998	546.65	2008	1786.14
1979	375.47	1989	357.06	1999	603.59	2009	1397.49

الجدول يعتبر خاص (عينة) كامنة وراء عملية عشوائية. الفرق بين عملية عشوائية وخاصتها هو مشابه للفرق بين المجتمع والعينة في دراسات مقطعية. كما نستخدم بيانات العينة لعمل استدلال عن المجتمع، في السلسلة الزمنية يستخدم البيانات الخاصة لبناء استدلال عن العملية العشوائية الكامنة وراءها. نوع من العملية العشوائية لاقى اهتماما كبيرا في تحليل السلاسل الزمنية هو ما يسمى العملية العشوائية الساكنة.

عملية عشوائية تسمى ساكنة عندما يكون المتوسط والتباين ثابت عبر الزمن وقيمة التغاير بين فترتين زمنيتين يعتمد على المسافة او المتباطئة بين فترتين زمنيتين وليس الوقت الزمني الحقيقي التي حسب فيها التغايير. لشرح هذا اذا كانت Y سلسلة زمنية عشوائية لها الخصائص التالية:

$$Mean:\ \ E(Y_t) = \mu \qquad \text{المتوسط}$$

$$Variance:\ var(Y_t) = E(Y_t - \pi)^2 = \sigma^2 \qquad \text{التباين}$$

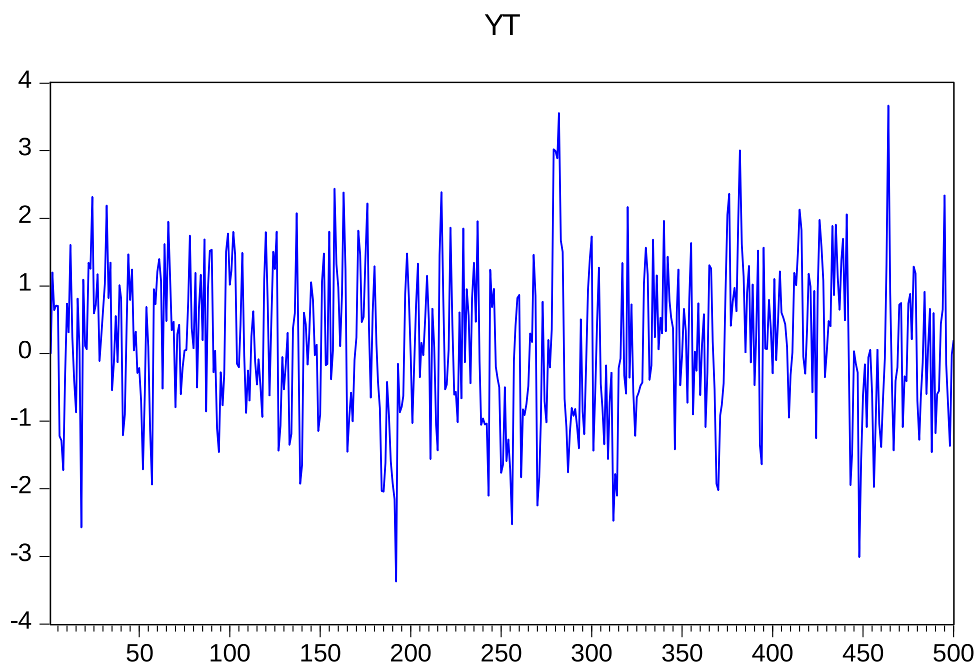
$$Covariance\ \gamma_k = E[(Y_t - \mu)(Y_{t+k} - \mu)] \qquad \text{التغاير}$$

حيث ترمز γ للتغاير او التغاير الذاتي عن المتباطئة k التغاير بين القيمة Yt و Y_{t+k} أي بين قيمتين للفترة بينهما. اذا كانت k=0 نحصل على γ_0 والتي تعني التباين للقيمة Y ويساوي$= \sigma^2$: أذا كانت k=1 و γ_1 هو التغاير بين قيمتين لـY نوع التغاير الذي تحدثنا عنة في فصل الارتباط الذاتي .

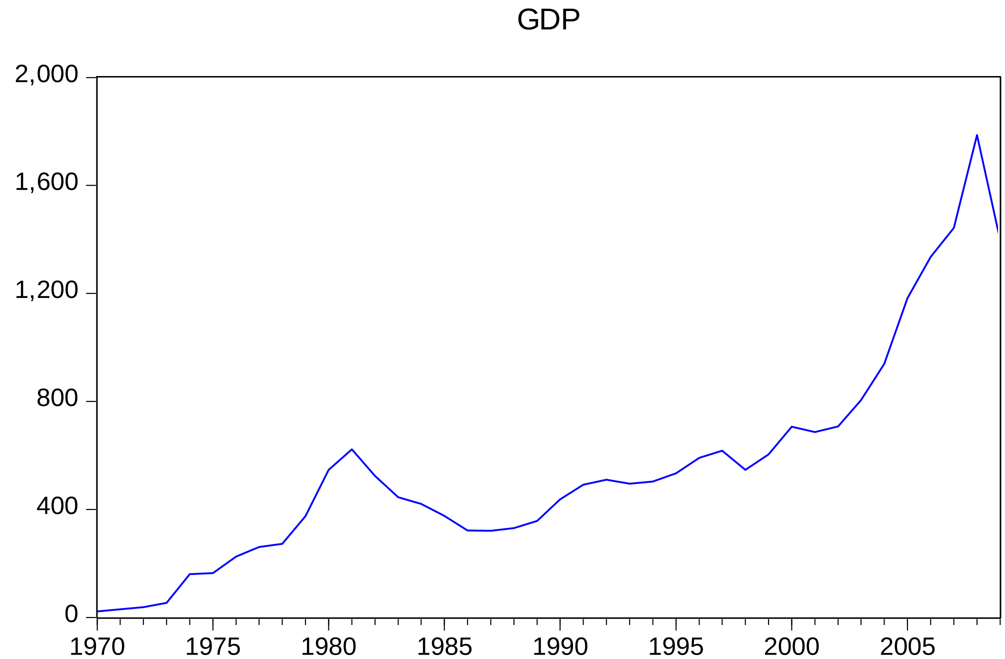
لنفرض اننا نقلنا اصل Y من Yt الى Y_{t+m} . الآن اذا كانت Y ساكنة، فأن المتوسط، التباين، والتغاير الذاتي للقيمة Y_{t+m} يجب ان يكون هو نفسة للقيمة Yt . بالاختصار اذا كانت السلسلة الزمنية ساكنة ، فأن المتوسط ، التباين والتغاير الذاتي سيبقون ثابتين عند أي فترة زمنية.

أذا كانت السلسلة الزمنية غير ساكنة كما عرفناها، تسمى سلسلة زمنية غير ساكنة، تكون أحيانا ناتجة لنقلة في المتوسط.

الشكل 10.1



الشكل 10.2



الشكل 10.1و 10.2 مثال لسلسلة زمنية ساكنة وسلسلة زمنية غير ساكنة.

اختبار ات السكون باستخدام دالة الارتباط الذاتي مبني على correlogram

اختبار بسيط للسكون مبني على ما يسمى دالة الارتباط الذاتي Autocorrelation function (ACF)

دالة $$\rho_k = \frac{\gamma_k}{\gamma_0}$$

$$\frac{\text{التغاير عند المتباطئة } k}{\text{التباين}} = \frac{Covariance\ at\ lag\ k}{Variance}$$

عندما k=0 فأن ${}_0 = 1$

حيث ان كل من التغاير والتباين تقاس بنفس الوحدة فأن الارتباط الذاتي من غير وحدات وتتراوح قيمتة بين +1 و -1 كأي معامل ارتباط اذا تم رسم الشكل البياني لقيمة الارتباط الذاتي نحصل على مايعرف ب ارتباط الذاتي للمجتمع. حيث اننا في الواقع نحصل على عينة للعملية

العشوائية فأنة يمكن حساب دالة الارتباط الذاتي للعينة $\hat{\rho}_k$ لحساب دالة الارتباط الذاتي يحسب التغاير ومن ثم التباين

$$\hat{\gamma}_k = \frac{\sum(Y_t - \bar{Y})(Y_{t+k} - \overline{Y})}{n}$$

$$\hat{\gamma}_0 = \frac{\sum\left(Y_t - \hat{Y}\right)^2}{n}$$

حيث تشيرn الى حجم العينة و $\bar{Y}$ متوسط العينة.

رسم الدالة بيانيا مقابل المتباطئات يسمى Sample Correlogram إذا انحدرت قيمة ببطيء فهذا يدل على إن السلسلة الزمنية غير مستقره ويمكن فحص الرسم البياني ρk للتحقق من استقرار الدالة وكذلك يمكن استخدام اختبار إحصاء Q ماذا كان معامل الارتباط الذاتي ρk يساوي الصفر أي لا توجد علاقة بين المتباطئات

اختبار إحصاء Q لـ بوكس وبيرز Box and Pierce

$$Q = n\sum_{k=1}^{m} \hat{\rho}_k^2$$

nحجم العينة و m طول المتباطئات. يتوزع اختبار Q حسب توزيع كاي χ2 بدرجة حرية df =mإذا كانت اختبار يفوق القيمة الجد وليه نرفض فرضية العدم أن معاملات التباطىء تساوي الصفر.

اختبار آخر لاختبار ماذا كانت المعاملات تساوي الصفر هو اختبار إحصاء LJung-Box ((LB

$$LB = n(n+2)\sum_{k=1}^{m}\left(\frac{\hat{\rho}_k^2}{n-k}\right) \sim \chi_m^2$$

الشكل 10.3

Sample: 1970 2009 عند المستوى
Included observations: 40

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
. |******|	. |******|	1	0.880	0.880	33.384	0.000
. |***** |	**| . |	2	0.700	-0.332	55.070	0.000
. |**** |	. |*. |	3	0.554	0.124	69.002	0.000
. |*** |	.*| . |	4	0.412	-0.181	76.906	0.000
. |** |	. |*. |	5	0.301	0.121	81.254	0.000
. |** |	. | . |	6	0.222	-0.056	83.695	0.000
. |*. |	. | . |	7	0.169	0.074	85.152	0.000
. |*. |	. | . |	8	0.134	-0.034	86.100	0.000
. |*. |	. | . |	9	0.097	-0.039	86.615	0.000
. | . |	. | . |	10	0.068	0.022	86.871	0.000
. | . |	. |*. |	11	0.068	0.100	87.140	0.000
. | . |	. | . |	12	0.073	-0.045	87.457	0.000
. | . |	. | . |	13	0.056	-0.063	87.653	0.000
. | . |	. | . |	14	0.036	0.002	87.739	0.000
. | . |	. | . |	15	0.020	-0.005	87.765	0.000

Sample: 1970 2009الفروق الأولى
Included observations: 39

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
. | . |	. | . |	1	-0.016	-0.016	0.0111	0.916
. | . |	. | . |	2	0.067	0.067	0.2054	0.902
. | . |	. | . |	3	0.051	0.054	0.3230	0.956
.*| . |	.*| . |	4	-0.120	-0.123	0.9767	0.913
. | . |	.*| . |	5	-0.057	-0.069	1.1276	0.952
.*| . |	.*| . |	6	-0.093	-0.083	1.5478	0.956
.*| . |	.*| . |	7	-0.083	-0.067	1.8917	0.966
. | . |	. | . |	8	0.027	0.028	1.9282	0.983
. | . |	. | . |	9	-0.064	-0.060	2.1491	0.989
. | . |	.*| . |	10	-0.057	-0.084	2.3258	0.993
. |*. |	. |*. |	11	0.109	0.086	3.0066	0.991
. | . |	. | . |	12	0.005	0.016	3.0081	0.995
. | . |	.*| . |	13	-0.042	-0.076	3.1150	0.997
. | . |	. | . |	14	-0.015	-0.056	3.1294	0.999
. | . |	. | . |	15	0.039	0.054	3.2300	0.999

الشكل 10.3 يوضح دالة الارتباط الذاتي لأجمالي الناتج المحلي للسعودية من عام 1970-2009 ، تظهر قيمة الارتباط الذاتي الى 15متباطئة. كيف يوضح شكل الدالة ان السلسلة الزمنية ساكنة؟ من الملاحظ انا يبدأ بقيمة مرتفعه 0.880 عند المتباطئة واحد. ثم يبدأ يتناقص تدريجيا. هذا النوع من الارتباط الذاتي عموما مؤشر ان السلسلة الزمنية غير ساكنة. بالمقارنة السلسلة الزمنية تكون غير ساكنة هو احتمال ان يكون الارتباط الذاتي عند أي متباطئة من صفر هو صفر. بينما السلسلة الزمنية احتمال ان يكون صفر اكبر من 5%

<u>-اختبار ديكي فيلرDickey-Fuller : اختبار جذر الوحدة:</u>

u_t تتبع الفروض الخاصة بالنموذج الكلاسيكي، وسط صفري، وتباين ثابت والتغاير يساوي الصفر، هذه الخواص تجعل الخطأ العشوائي u_t أن يسمى White Noise

إذا كان معامل الانحدار بين Yt وY_{t-1} يساوي الواحد وهذا يسمى بجذر الوحدة. أي تكون

$$Y_t = \rho Y_{t-1} + u_t$$

غير ساكنة

إذا كانت تساوي الواحد فان السلسة الزمنية يقال أنها ذات جذر وحده أو ما يعرف بالمســار العشوائي random walk أي عندما تكون تتبع المســار العشوائي أي أن السلسلة الزمنية غير مستقره.

ويعبر عن معادلة جذر الوحدة بالتالي

$$\Delta Y_t = (Y_t - Y_{t-1}) = (\rho - 1)_1 Y_{t-1_t} + u_t$$

$$\Delta Y_t = \delta_1 Y_{t-1_t} + u_t$$

$$\delta = (\rho - 1)$$

نقوم باختبار احتواء المتغير على جذر الوحدة آي نقوم بأجراء الاختبار التالي:

السلسلة الزمنية غير ساكنة $H_0 : \delta_1 = 0$

السلسلة الزمنية ساكنة $H_A : \delta_1 < 0$

إذا كانت δ_1 اقل من الصفر نرفض فرضية العدم بعدم استقرار الدالة ونستنتج أن الدالة ساكنة

الاختبار الإحصائي هو t $t = \frac{\delta_1 - 0}{Se(\delta\ )}$

إلا أن قيم t لا تتبع جدولt بل هناك جدول خاص يسمى بجدول ديكي فيلر. Dickey Fuller (1979) والتي طورت من قبل ما كنون (1991) MacKinnon نقارن القيمة المحسوبة والقيمة الجدلية حيث نرفض فرضيه العدم إذا كانت القيمة المحسوبة أعلى من القيمة الجد وليه.
يجري اختبار ديكي فيلر بإجراء المعادلات الثلاث التالية:

اختبار ديكي فيلر DF $\Delta Y_t = \delta_1 Y_{t-1_t} + u_t$

اختبار ديكي فيلرDF بوجود قاطع $\Delta Y_t = \delta_0 + \delta_1 Y_{t-1_t} + u_t$

اختبار ديكي فيلر DF مع قاطع ومتجهة زمني T. $\Delta Y_t = \delta_0 + \delta_1 Y_{t-1_t} + \delta_2 T + u_t$

إذا كان الخطأ العشوائي يتصف بوجود الارتباط الذاتي فانه يمكن استخدام ديكي فيلر الموسع حيث يتضمن الاختبار متباطئات الفروق:

اختبار ديكي فيلر الموســع ADF $\Delta Y_t = \delta_0 + \delta_1 Y_{t-1_t} + \delta_2 T + \sum_{i=1}^{m} \Delta Y_{t-1} + u_t$

حيث تتضمن المعادلة قيم متباطئة للفروق بعدد يمكن الخطأ العشوائي بان يكون مستقل (لا يوجد ارتباط الذاتي).

Null Hypothesis: GDP has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=9)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		**-0.044794**	0.9484
Test critical values:	1% level	-3.610453	
	5% level	**-2.938987**	
	10% level	-2.607932	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(GDP)
Method: Least Squares

Sample (adjusted): 1971 2009
Included observations: 39 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GDP(-1)	-0.002123	0.047402	-0.044794	0.9645
C	36.38616	30.95879	1.175309	0.2474

R-squared	0.000054	Mean dependent var	35.25436
Adjusted R-squared	-0.026971	S.D. dependent var	110.2454
S.E. of regression	111.7223	Akaike info criterion	12.31983
Sum squared resid	461829.1	Schwarz criterion	12.40514
Log likelihood	-238.2367	Hannan-Quinn criter.	12.35044
F-statistic	0.002007	Durbin-Watson stat	1.639084
Prob(F-statistic)	0.964512		

Null Hypothesis: D(GDP) has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=9)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-4.786159	0.0004
Test critical values:	1% level	-3.615588	
	5% level	-2.941145	
	10% level	-2.609066	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(GDP,2)
Method: Least Squares

Sample (adjusted): 1972 2009
Included observations: 38 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(GDP(-1))	-1.028128	0.214813	-4.786159	0.0000
C	37.27884	20.88734	1.784757	0.0827

R-squared	0.388870	Mean dependent var	-10.43632
Adjusted R-squared	0.371895	S.D. dependent var	142.7647
S.E. of regression	113.1455	Akaike info criterion	12.34642
Sum squared resid	460868.4	Schwarz criterion	12.43261
Log likelihood	-232.5820	Hannan-Quinn criter.	12.37709
F-statistic	22.90731	Durbin-Watson stat	1.621576
Prob(F-statistic)	0.000029		

درجة التكامل Degree of integration :

درجة التكامل تختبر ماذا كانت السلسلة الزمنية مستقره في المستويات I(0) أو مستقره في الاختلاف الأول I(1) أو في الاختلاف الثاني I(2). ويتم معرفة درجة التكامل بأجراء اختبار ديكي فيلر على الاختلاف الأول $\Delta Y_t = Y_t - Y_{t-1}$ و الاختلاف الثاني $\Delta\Delta Y_t = \Delta Y_t - \Delta Y_{t-1}$ فإذا كان الاختلاف الأول مستقر والدالة غير مستقره في المستويات يقال أنها متكاملة من الدرجة الأولى.. I(1) واغلب السلاسل الزمنية الاقتصادية الغير مستقره تكون متكاملة من الدرجة الأولى

هل اجمالي الإنتاج المحلي للمملكة العربية السعودية مستقر؟

	المستوى		الفروق الأولى	
	قاطع	قاطع ومتجة	قاطع	قاطع ومتجة
القيم الحرجية	2.938987-	3.529758-	2.941145-	3.533083-
GDP	0.044794-	1.398059-	4.786159-	4.710758-

عند المستوى قيمة اختبار ديكي قيلر اقل من القيمة الحرجة أي ان السلسلة الزمنية غير مستقرة، عند الفروق الأولى قيمة الاختبار الاحصائي اعلى من القيمة الحرجة نرفض فرضية العدم ان السلسلة الزمنية غير مستقرة ونستنتج انها مستقرة عند الفروق الأولى.

نماذج الانحدار الذاتي للسلاسل الزمنية من الدرجة الأولى.

ابسط نموذج للسلسلة الزمنية هو الانحدار الذاتي من الدرجة الأولى AR(1)

$$Y_t = \theta Y_{t-1} + u_t$$

للتبسيط لاتتضمن قاطع و تمثل $|\theta| < 1$ والعشوائي يمثل ضجيج ابيض*(White Noise)* الافتراض خلف نموذج الانحدار الذاتي من الدجة الأولى ان سلوك السلسلة الزمنية Y_t يحدد غالبا من قبل قيمها للفترة الزمنية السابقة. أي ان ماسوف يحدث في الفترة T يعتمد على مايحدث في الفترة *t-1* . وكذلك ماسوف يحدث في الفترة *T+1* سوف يتحدد بسلوك السلسلة الزمنية في الفترة الحالية.

نماذج الانحدار الذاتي من الدرجة اعلى من الواحد:*AR(P)*

لتعميم نموذج الانحدار من الدرجة الأولى*AR(1)* نستخدم*AR(p)* الرقم داخل القوس يمثل درجة عملية الانحدار الذاتي. على سبيل المثال *AR(2)* سيكون من الدرجة الثانية

$$Y_t = \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + u_t$$

وكذلك *AR(p)* سيكون انحدار ذاتي من الدرجة P كما يلي:

$$Y_t = \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \cdots \ldots + \theta_p Y_{t-p} + u_t$$

أو باستخدام رمز الجمع:

$$Y_t = \sum_{i=1}^{p} \theta_i Y_{t-i} + u_t$$

واخيرا باستخدام متباطئة المشغل*Lag Operator* والذي يمتلك الخاصية يمك $L^n Y_t = Y_{t-n}$

يمكن كتابة نموذج الانحدار الذاتي من الدرجة p كما يلي

$$L^0 Y_t = \theta_1 L^1 Y_t + \theta_2 L^2 Y_t + \cdots \ldots + \theta_p L^p Y_t + u_t$$

$$Y_t\left(1 - \theta_1 L - \theta_2 L^2 - \cdots - \theta_p L^p\right) = u_t$$

$$\emptyset(L) Y_t = u_t$$

السكون في نموذج الانحدار الذاتي:

شرط كون *AR(p)* ساكنة هو اذا كان جذر *P* للمعادلة كثيرة الحدود

$$\emptyset(z) = 0$$

يكون اكبر من الواحد في القيمة المطلقة حيث تشير *Z* للمتغير الحقيقي. من الممكن التعبير عنها بالمصطلحات التالية حل معادلة كثيرة الحدود يجب ان يكون خارج دائرة جذر الوحدة. لاثبات ذلك باستخدام *AR(1)*

$$(1 - \theta z) = 0$$

حيث ان الجذر أعلى من الواحد اذا

$$|\gamma| = \left|\frac{1}{\theta}\right| > 1$$

اذا $|\theta| < 1$

نماذج المتوسط المتحرك*(MA)* *Moving Average*:

نموذج المتوسط المتحرك في ابسط أشكاله هو من الدرجة الأولى وهو بالشكل التالي:

$$Y_t = u_t + \theta u_{t-1}$$

MA(1) نموذج المتوسط المتحرك من الدرجة الأولى يتضمن أن *Yt* تعتمد على قيمة المتغير العشوائي الحالي ويعتبر *ut* ضجيج ابيض>

نموذج المتوسط المتحرك من درجة *(q)*

$$Y_t = u_t + \vartheta_1 u_{t-1} + \vartheta_2 u_{t-2} + \cdots \ldots + \vartheta_q u_{t-q}$$

$$Y_t = u_t + \sum_{j=1}^{q} \vartheta_i u_{t-j}$$

وباستخدام متباطئة المشغل

$$Y_t = \left(1 - \vartheta_1 L - \vartheta_2 L^2 - \cdots - \vartheta_p L^p\right) u_t$$

$$Y_t = \varphi(L) u_t$$

لأن $MA(q)$ تعرف انها متوسط متحرك ثابت ومن ذلك يتبع ان المتوسط المتحرك ساكن مادامت q محدودة.

نماذج $ARMA$:

جمع نماذج الانحدار الذاتي ونماذج المتوسط المتحرك نتحصل على سلسلة زمنية جديدة تسمى $ARMA(p,q)$

$$Y_t = \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \cdots .. \theta_p Y_{t-p} + u_t + \vartheta_1 u_{t-1} + \vartheta_2 u_{t-2} + \cdots ... + \vartheta_q u_{t-q}$$

وتكتب باستخدام صيغة الجمع

$$Y_t = \sum_{i=1}^{p} \theta_i Y_{t-i} + u_t + \sum_{j=1}^{q} \vartheta_i u_{t-j}$$

او باستخدام متباطئة المشغل

$$Y_t\left(1 - \theta_1 L - \theta_2 L^2 - \cdots - \theta_p L^p\right) = \left(1 - \vartheta_1 L - \vartheta_2 L^2 - \cdots - \vartheta_p L^p\right) u_t$$

شرط السكون يتعامل مع جزء $AR(p)$. بناء على ذلك على كون

$$\emptyset(z) = 0$$

تكامل السلسلة الزمنية ونماذج $ARIMA$:

نماذج $ARMA$ تكون فقط مع سلاسل زمنية Yt ساكنة. هذا يعني ان يكون المتوسط والتباين والتغاير ثابت عبر الزمن. ولكن معظم السلاسل الزمنية والمالية تمتلك متجه عبر الزمن وكذلك المتوسط لـYt خلال سنة واحدة سيختلف عن المتوسط في سنة أخرى. هكذا المتوسط لمعظم السلاسل الزمنية الاقتصادية والمالية غير ثابت عبر الزمن. مما يشير ان السلاسل الزمنية غير ساكنة لتجنب هذه المشكلة وللحصول على سلاسل زمنية ساكنة نحتاج لإزالة المتجه من البيانات الأصلية ويتم ذلك من خلال استخدام الفروق

$$\Delta Y_t = Y_t - Y_{t-1}$$

معظم السلاسل الزمنية عند الفروق الأولى. فاذا كانت ساكنة في الفروق الأولى تسمى متكاملة من الدرجة الأولى *I(1)* وهذا يكمل المصطلح *ARIMA* اذا كانت السلسلة الزمنية غير ساكنة في الفروق الأولى يجب أخذ الفروق الثانية.

$$\Delta^2 Y_t = \Delta\Delta Y_t = \Delta Y_t - \Delta Y_{t-1}$$

اذا كانت السلسلة الزمنية ساكنة في الفروق الثانية تسمى متكاملة من الدرجة الثانية *I(2)*

وبصفة عامة اذا كانت السلسلة الزمينة اخذت لها الفروق من الدرجة d لتكون ساكنة فنه يقال انها متكاملة من الدرجة d أي *I(d)* لذا يسمى نموذج *ARIMA(p,d,q)* حيث تشير p الى عدد متباطئات المتغير التابع *(AR)* و d عدد المرات التي تؤخذ فيها الفروق للحصول على سكون السلسلة الزمنية و q عدد متباطئات حد الخطأ.

مثال لنموذج *ARIMA*

اختيار النموذج باستخدام طريقة بوكس جينكينز *Box-Jenkins* :

بكوكس جبنكينز *(1976)* اقترحوا طريقة الثلاث مراحل لنموذجة السلسلة الزمنية. الثلاث مراحل تتضمن، التمييز (التعريف)، التقدير، فحص النموذج.

مرحلة التمييز اختبار *correlogram*:

يتم باستخدام الرسم البياني لدالة الارتباط الذاتي *ACF* ودالة الارتباط الذاتي الجزئي *PACF* باستخدام الرسم البياني لتسلسل *Yt* مع الزمن *T* يقدم معلومات مفيدة بخصوص القيم المتطرفة والقيم المفقودة والتغيرات الهيكلية للبيانات. كما ذكر سابقا معظم السلاسل الاقتصادية والمالية غير ساكنة وفي الغالب المتغيرات تمتلك متجه قد يكون متزايدا او متناقصا يتسكع بدون قيمة متوسط او تباين ثابت. يمكن تصحيح القيم المتطرفة او المفقودة في هذه المرحلة، متعارف علية ان يتم استخدام الفروق الأولى.

مقارنة لدالة الارتباط الذاتي *ACF* ودالة الارتباط الذاتي الجزئي*PACF* قد يؤدي الى اقتراح نماذج ممكنة. نظريا اذا كانت السلسلة الزمنية غير ساكنة لن تتناقص بقوة او اظهار علامات الاضمحلال (التناقص). لتحويلها الى دالة ساكنة كما ذكر سابقا باستخدام الفروقات الأولى.

عند الحصول على سلسلة زمنية مستقرة الخطوة الثانية هي تعريف p,q لنموذج $ARIMA$ لعملية $MA(q)$ فان دالة الارتباط الذاتي ACF ستظهر مقدرات مختلفة معنويا عن الصفر الى متباطئة q ثم تنخفض فجأة. اما دالة الارتباط الذاتي الجزئي $PACF$ لسلسلة $MA(q)$ سوف تنخفض بسرعة اما بطريقة اسي او بطريقه الظل.

بالمقارنة $AR(P)$ لدالة الارتباط الذاتي ACF سوف تتناقص بسرعة، اما بتناقص اسي او بجيب . بينما الارتباط الذاتي الجزئي $PACF$ ستظهر تموج (ارتباط ذاتي معنوي) للمتباطئات حتى قيمة p ثم تتناقص فجأة.

اذا كلاهما ACF و $PACF$ لم يحددون نقطه ينقطع فيها، سيكون هناك سلسلة مختلطة في هذه الحالة سيكون من الصعب او المستحيل تحديد درجة AR و MA ستكون ACF و $PACF$ متراكبة *(superimposed)*. مثالا اذا اظهرت كلا ACF و $PACF$ علامات تناقص اسي متباطئ فان $ARMA(1,1)$ قد نعرف. اذا اظهرت ACF ثلاث موجات معنوية عند المتباطئة 1 ،2 و 3 اذا تناقص اسي $AMA(3,1)$ الجدول 10.1 يقدم مجموعات من ACF و $PACF$ تسمح بتحري درجة $ARMA$. بصفة عامة من الصعب تحري عملية مخلوطة او اكثر من 1 لذا فانة من المهم تقدير وفحص النموذج.

جدول 10.1 ACF و $PACF$ الممكنة لـ $ARMA(p,q)$

	$PACF$	ACF
$MA(1)$	تناقص سواء ظل او اسي	موجة عند المتباطئة واحد
$AR(1)$	موجة عند المتباطئة واحد	تناقص سواء جيب او اسي
$ARMA(1,1)$	موجة عند المتباطئة واحد يتبعها تناقص سواء ظل او اسي	موجة عند المتباطئة واحد يتبعها تناقص سواء ظل او اسي
$ARMA(1,2)$	موجة عند المتباطئة واحد يتبعها تناقص سواء ظل او اسي	موجة عند المتباطئتين واحد واثنين يتبعها تناقص سواء ظل او اسي
$(ARMA(2,1)$	موجة عند المتباطئتين واحد واثنين يتبعها تناقص سواء ظل او اسي	موجة عند المتباطئة واحد يتبعها تناقص سواء ظل او اسي
$ARMA(2,2)$	موجة عند المتباطئتين واحد واثنين يتبعها تناقص سواء ظل او اسي	موجة عند المتباطئتين واحد واثنين يتبعها تناقص سواء ظل او اسي

التقدير:

عند هذه المرحلة يتم تقدير معاملات النموذج ويتم اختبارها. ثم يتم مقارنة النماذج باستخدام *Akaika Information Creteria(AIC)* وكذلك *Schwarz Bayesian criterion(SBC)* في هذه المرحلة يجب التأكد ان النموذج ساكن.

حيث ان السلسلة الزمنية متكاملة من الدرجة الأولى *I(1)* سنحاول ان نعرف (نميز) السلسلة الزمنية يجب النظر الى قيم *ACF* و *PACF* اذا كانت تختلف عن الصفر باستخدام فترة الثقة

$$0 \mp 2 \times se(\widehat{\rho_k})$$

الخطأ المعياري لدالة الارتباط الذاتي *ACF* هو

$$se(\hat{\rho}_k) = \sqrt{\frac{1}{n}\left(1 + 2\sum_{i=1}^{k-1} \widehat{\rho_i^2}\right)}$$

تحت فرضية العدم ان $for\ i \geq k \qquad \rho_i = 0$ سيكون الخطأ المعياري

$$se(\hat{\rho}_k) = \sqrt{\frac{1}{n}} \quad for\ all\ k$$

النقاط في الشكل حول معامل الارتباط هو فترة الثقة والتي تشير الى $\pm 2\sqrt{\frac{1}{n}}$ الشكل 10.4 للفروق الاولى يشير الى *ARIMA(3,1,2)*

فحص النموذج:

اختبار جودة النموذج

مثال لطريقة بوكس جينكينز:

Date: 03/17/12   Time: 18:36
Sample: 1970 2009
Included observations: 35

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
**| . |	**| . |	1	-0.310	-0.310	3.6672	0.055
.*| . |	**| . |	2	-0.170	-0.294	4.7947	0.091
. |** |	. |*. |	3	0.268	0.131	7.6966	0.053
. | . |	. |*. |	4	0.069	0.204	7.8928	0.096
.*| . |	. |*. |	5	-0.098	0.101	8.3058	0.140
. | . |	. | . |	6	0.006	-0.009	8.3075	0.216
. | . |	. | . |	7	0.044	-0.052	8.3987	0.299
. | . |	. | . |	8	-0.013	-0.041	8.4066	0.395
. | . |	. | . |	9	-0.020	-0.016	8.4265	0.492
. | . |	. | . |	10	0.003	-0.002	8.4270	0.587
. | . |	. | . |	11	-0.001	-0.002	8.4270	0.675
. | . |	. | . |	12	-0.011	-0.009	8.4336	0.750
. | . |	. | . |	13	-0.008	-0.014	8.4375	0.814
. | . |	. | . |	14	-0.005	-0.016	8.4394	0.865
. | . |	. | . |	15	-0.008	-0.016	8.4437	0.905
. | . |	. | . |	16	-0.010	-0.016	8.4500	0.934

الشكل 10.4 دالة Correlgrom للفروق الأولى.

وبعد حساب *ACF* و*PACF* والاطلاع على الشكل نجد *ACF* تنقطع عند المتباطئة 4 مما يقترح ان السلسلة ساكنة عند الفروق الأولى كما أن الاحتمالية *prop.* اكبر من 5% مما يشير الى قبول فرضية العدم ان معامل الارتباط الذاتي يساوي الصفر.

	PACF	*ACF*
ARMA(3,3)	موجات عند المتباطئه 3 و 4 يتبعها اضمحلال	موجات عند المتباطئه 3 يتبعها اضمحلال

تقدير النموذج *ARMA(2,1)*

Dependent Variable: DLGDPF
Method: Least Squares
Date: 03/17/12   Time: 21:59

Sample (adjusted): 1976 2009
Included observations: 34 after adjustments
Convergence achieved after 17 iterations
MA Backcast: 1975

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.057881	0.012318	4.699012	0.0001
AR(1)	0.347043	0.144489	2.401870	0.0227
AR(2)	-0.534594	0.133966	-3.990514	0.0004
MA(1)	0.997817	0.008294	120.3075	0.0000

R-squared	0.752688	Mean dependent var	0.062465
Adjusted R-squared	0.727957	S.D. dependent var	0.081906
S.E. of regression	0.042720	Akaike info criterion	-3.358157
Sum squared resid	0.054751	Schwarz criterion	-3.178585
Log likelihood	61.08866	Hannan-Quinn criter.	-3.296917
F-statistic	30.43473	Durbin-Watson stat	1.293293
Prob(F-statistic)	0.000000		

Inverted AR Roots	.17+.71i	.17-.71i
Inverted MA Roots	-1.00	

تقدير النموذج *ARMA(1,1)*

Dependent Variable: DLGDPF
Method: Least Squares
Date: 03/17/12 Time: 21:58
Sample (adjusted): 1975 2009
Included observations: 35 after adjustments
Convergence achieved after 11 iterations
MA Backcast: 1974

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.061347	0.021332	2.875884	0.0071
AR(1)	0.175358	0.159529	1.099224	0.2799
MA(1)	0.996470	0.017038	58.48362	0.0000

R-squared	0.621278	Mean dependent var	0.059714
Adjusted R-squared	0.597608	S.D. dependent var	0.082317
S.E. of regression	0.052217	Akaike info criterion	-2.985001
Sum squared resid	0.087252	Schwarz criterion	-2.851685
Log likelihood	55.23751	Hannan-Quinn criter.	-2.938980
F-statistic	26.24738	Durbin-Watson stat	1.619606
Prob(F-statistic)	0.000000		

Inverted AR Roots	.18
Inverted MA Roots	-1.00

نماذج *ARIMA* هي نماذج احصائية لها علاقة بالنماذج الاقتصادية ولكنهم ليسوا نماذج اقتصادية هذا يجعلة من الصعب الاختيار بين نماذج مختلفة خصوصا اذا كان التحديد متقارب والمقدرات متقاربة من الادوات المساعدة في اختيار النموذج معيار اكيكا *Akika information criterion (AIC)*ومعيار سشوارز *Schwarz information criterion (SIC)* كلا المعيارين مبنية على تباين البواقي$\tilde{\sigma}_u^2$. يفضل الحصول على نموذج يتضمن اصغر تباين للبواقي. ولكن من المعروف ان تباين البواقي يتناقص بزيادة عدد المتغيرات المفسرة. لنموذجين مبنين على نفس السلسلة الزمنية يتم اختيار النموذج الذي يمتلك اقل قيمة من *AIC, SIC* . قيم المعيار ممكن شرحها نسبيا. لأن السلسة الزمنية تستخدم باطوال مختلفة يتم تطبيع المعيار بقسمتة على عدد المشاهدات المستخدمة بتقدير النموذج. معيار *AIC, SIC* يعرف كالتالي:

$$AIC = \ln\left(\tilde{\sigma}_u^2\right) + \frac{2k}{n}$$

$$SIC = \ln\left(\tilde{\sigma}_u^2\right) + \frac{k.\ln(n)}{n}$$

معيار*SIC* يعتبر خيار بدلا من *AIC* لها نفس المعنى ولكن و لكن تعطي ثقل لعدد المعاملات*k* لهذا السبب*SIC* سوف تعطي نموذج ابسط من *AIC* وهذه ميزة هذه المعايير لا تستخدم لمقارنة نماذج التي تستخدم مستوى مختلف من الفروقات.

المثال استخدم نموذجين *ARIMA(1,1) ARIMA(2,1)* مبدئيا تم تحديد *p,q* بالاطلاع على دالة *ACF,PACF* ثانيا يتم اختيار *p,q* على ضوء نتائج تقدير النموذج من الاطلاع على تقدير النموذجين نجد ا كلا المعيارين*AIC, SIC* ذا قيمة اصغر في النموذج*ARIMA(1,1* كما أن *inverted roots* الجذر المقلوب يجب ان يكون ضمن دائرة الوحدة.

فحص النموذج يتضمن اختبار *Q-statistic* يت اختبار ماذا كانت البواقي تتبع عملية الضجيج الأبيض المعادلة كالتالي:

$$Q = n(n+2)\sum_{k=1}^{k}\frac{\tilde{\rho}_k^2}{n-k}$$

حيث تمثل $\tilde{\rho}^2$ مربع بواقي معامل الارتباط الذاتي عند المتباطئة k . عند فرضية عدم الضجيج الأبيض للبواقي للنموذج$ARMA(p,q)$ فأن اختبار Q يتبع توزيع كاي$\chi(K-p-q)$ χ^2

بالاطلاع على نتائج اختبار Q في الشكل 10.4 نجد انم فرضية العدم للبواقي بانها ذات ضجيج ابيض لا يمكن رفضها.