نرم افزار R : محیط برنامه نویسی برای تحلیلهای اقتصادسنجی و سریهای زمانی

کامبیز هژبر کیانی[1]

E-Mail : khkiani@yahoo.com

[1] – استاد تمام دانشکدهٔ علوم اقتصادی و سیاسی شهید بهشتی

علیرضا مرادی [1]

تاریخ ارسال : تاریخ پذیرش:

چکیده :

شاید نرم افزارهای محاوره ای [2] اقتصادسنجی در برخی امور نیاز محققان اقتصادی را برآورده سازند ولی این نرم افزارها با تمامی امکاناتشان در مواردی نمیتوانند برای تخمین و پیش بینی الگوهای اقتصادسنجی نمی توانند بطور ایده آل عمل نمایند ، این سبب شده تا نیاز به زبان و محیط های برنامه نویسی و بکارگیری آنها در پژوهشهای اقتصادی روز به روز با اهمیت تر گردد . امروزه نرم افزار R در بین محققان اقتصادی دنیا که بدنبال استفاده از یک نرم افزار با قابلیتهای برنامه نویسی هستند ، از جایگاه خاصی برخوردار است . دلایل این امر توانائیهای بالای این نرم افزار است که سبب شده مورد اقبال محققان قرار گیرد . هدف این مقاله معرفی کردن بخشی از این توانائیها برای محققان و دانشجویان کشور است که در مطالعات خود با برخی محدودیتهای محاسباتی نرم افزارهای محاوره ای مواجه اند . یکی از ویژگیهای قابل توجه نرم افزار R رایگان بودن آنست ، ولی این تنها ویژگی R نیست . در این مقاله سعی شده با ارائه مثالهایی گوشه ای از تواناییهای این نرم افزار در حوزه اقتصادسنجی و تحلیل سریهای زمانی نشان داده شود.

طبقه بندی JEL : c88, c87

واژه های کلیدی : نرم افزارهای اقتصادسنجی؛ اقتصادسنجی کاربردی ؛ تحلیل سریهای زمانی

۱ -مقدمه

امروزه به میزان چشمگیری بکارگیری الگوهای کمی و محاسباتی در مطالعات اقتصادی ، متداول شده است . جهت تخمین و محاسبات این الگوها از نرم افزارهای اقتصادسنجی و آماری استفاده می شود . نرم افزارهایی چون Ox , TSP, EViews, SAS, STATA, RATS, MATLAB, S-PLUS, Microfit و ... مثالهایی از این نرم افزارها هستند. عموماً این نرم افزارها در دو فرم محاوره ای و برنامه نویسی [3] طراحی میشوند . یکی از نرم افزارهای مهم در حوزۀ اقتصادسنجی نرم افزار R است، که بر اساس زبان برنامه نویسی S طراحی شده است .

۲ – تاریخچۀ پیدایش زبان R:

سابقۀ پیدایش زبان R از زبان برنامه نویسی S جدا نیست . S یک زبان سطح بالا برای دستکاری ، محاسبات ، تحلیل و نمایش داده هاست . زبان S ابتدا در سال ۱۹۷٦ توسط جان چامبرز [4] و همکارانش در

۲ - عضو هیات علمی دانشگاه آزاد کرمانشاه و دانشجوی دکتری اقتصاد دانشگاه آزاد واحد علوم و تحقیقات
E-Mail : info@radinworks.com

[2] Interactive
[3] Programming
[4] Chambers Jhon

دانشگاه (A&AT) تهیه شد . این نرم افزار ابتدا با زبان فرترن تهیه گردید و سپس در سال ۱۹۸۸ با زبان C بازنویسی شد که این نسخه از زبان S با نام نسخهٔ ۳ وارد بازار شد . نسخهٔ ٤ این نرم افزار در سال ۱۹۹۳ در موسسهٔ بل[1] تولید گردید. در سال ۲۰۰٤ موسسهٔ یاد شده تمامی حقوق مادی و معنوی این نرم افزار را به موسسهٔ اینسایتفول [2] به مبلغ ۲ میلیون دلار واگذار کرد . این موسسه نسخهٔ تجاری این نرم افزار را تحت عنوان S^{+} [3]وارد بازار کرد که هم اکنون نسخهٔ ۸ این نرم افزار آخرین نسخهٔ این سری می باشد . به موازات تحولات در نرم افزار S ، تولید نرم افزاری با خصوصیات بسیار مشابه با نرم افزار S و با نام R در سال ۱۹۹۱ توسط رابرت جنتلمن [4] و رز لاکا[5] آغاز شد . در سال ۲۰۰۰ نسخهٔ ۱ این نرم افزار بطور رایگان ارائه گردید . سپس این نرم افزار تا به امروز توسط محققین و برنامه نویسان نرم افزاری در سراسر دنیا تکمیل گردیده و در اختیار عموم قرار گرفته است .آخرین نسخهٔ این نرم افزار در زمان نگارش این مقاله نسخهٔ 2.10.0 است که در اکتبر ۲۰۰۹ ارائه شد .

اقتصادسنجی و تحلیل سریهای زمانی و موضوعات پیرامونی به استفاده از این نرم افزار سبب شده تا کاربران بسیار زیادی به استفاده از این نرم افزار روی آورند.

۳ -چرا نرم افزار R را میتوان به برخی از نرم افزارهای دیگر ترجیح داد؟

زبان R بسیار انعطاف پذیرتر از بسیاری از نرم افزارهای متداول است . زیرا نرم افزار R یک زبان برنامه نویسی ریاضی بسیار قوی است که صرفاً رگرسیونهای خاص و آزمونهای مشخصی را انجام نمی دهد . بلکه توانایی برنامه نویسی کاربر به او این اجازه را میدهد که برنامه ای را با مشخصات متفاوت نسبت به مدلها و رگرسیونهای خاص ارائه کند . ویژگی مطلوب دیگر این نرم افزار این است که هر محققی که یک کتابخانه[6] از دستورات R را خلق کند با ارسال این توابع به تارنمای[7] نرم افزار R این توابع کتابخانه ای را در اختیار سایر محققان قرار می دهد . البته زمانی این دستورات و توابع کتابخانه ای در اختیار عموم قرار می گیرد که صحت آنها مورد تایید کارشناسان قرار گیرد . ویژگی بسیار مهم زبان R اینست که این همه توانایی و امکانات بصورت رایگان در اختیار عموم قرار می گیرد و به راحتی بر روی هر پلتفرمی[8] اعم از ویندوز (۹۸ ، XP ، ویستا) و لینوکس و مکینتاش قابل نصب است .

٤ -دستیابی به نرم افزار R :

این نرم افزار را میتوان از آدرس ذیل دانلود کرد .

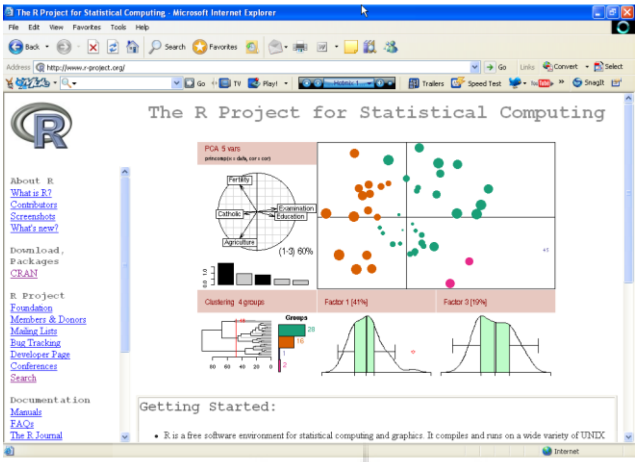
$http://www.r-project.org$

از آنجائیکه نرم افزار R یک نرم افزار با منوهای کرکره ای و گزینه های از قبل مشخص نیست . پس بایستی توابع و کتابخانه های مورد علاقهٔ خود را از تارنمای نرم افزار R دانلود کنید . در تارنمای نرم افزار

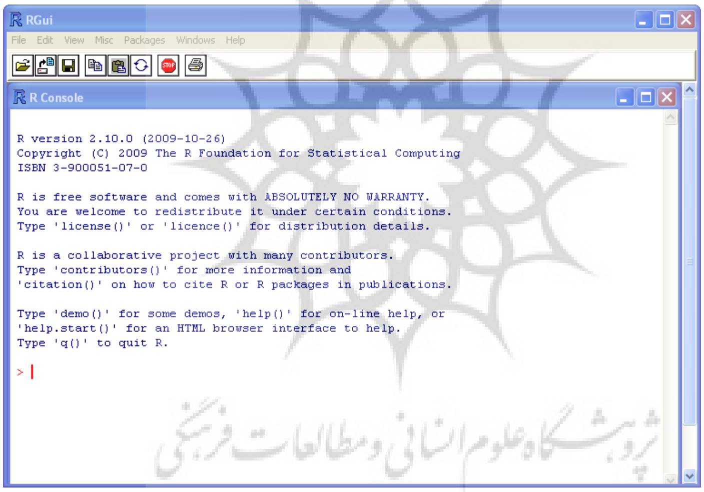
---

1 Bell Institute
2 Insightful
3 S-plus
4 Robert Gentleman
5 Ross Lhaka
6 Library
7 WebSite
8 Platform

R یک لینک با عنوان *CRAN*[1] وجود دارد (تصویر (۱) را ببینید .) که بایستی کتابخانه ها را از آن دانلود کرد .در لینک *contrib* در تارنمای نرم افزار R میتوان لیست کاملی از کتابخانه های R را یافت . پس از اتمام دانلود میتوان نرم افزار R را به راحتی بر روی کامپیوتر خود نصب کرد . مراحل نصب بسیار ساده بوده و نیازی به تشریح ندارد ولی پس از نصب برنامه با اجرای فایل اجرائی R پنجره ای بصورت تصویر (۲) باز می شود .



تصویر (۱) صفحهٔ اصلی تارنمای نرم افزار R



تصویر (۲) صفحهٔ دستور نرم افزار R

کتابخانه های مورد نظر خود را پس از دانلود بایستی در فلدر library در دایرکتوری R کپی کنید .

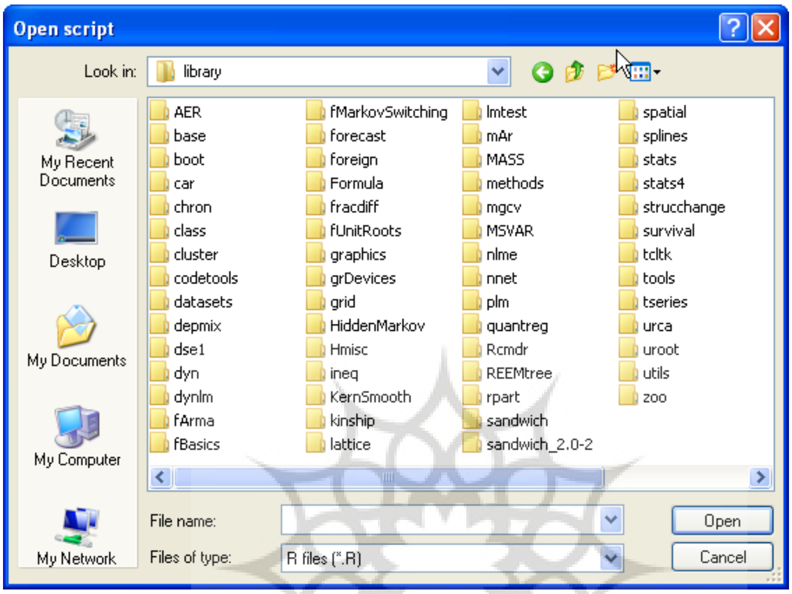
یک روش دیگر که مرحلهٔ دانلود و نصب را بطور همزمان انجام میدهد ، اینست که بایستی پس از وصل شدن به اینترنت در پنجرهٔ دستور نرم افزار R (پنجره ای که در تصویر (۲) دیده میشود) ، فرمان ذیل را اجرای کرد.

به عنوان مثال اگر چنین دستوری را تایپ و اجرا کنید .

```
>install.packages(c("car","systemfit"),repo="http://cran.st
at.ucla.edu",dep=TRUE)
```

[1] Comprehensive R Archive Network (CRAN).

این دستور توابع کتابخانه ای مربوط بـه جعبـه ابزارهـای `"car","systemfit"` را از آدرس اینترنتـی `"http://cran.stat.ucla.edu"` دانلـود و نصـب مـی کنـد. منظـور از `"ucla.edu"` دانشگاه کالیفرنیا است . علاوه بر این دانشگاه ، دانشگاههای دیگری هسـتند کـه میتوانیـد توابع کتابخانه ای را از تارنمای آنها دانلود کنید . فهرست دقیقی از این دانشگاهها و آدرس تارنمـای آنهـا را میتوان از آدرس $http://www.r-project.org$ بدست آورد . تصویر (۳) فلـدر library رایانـهٔ شخصی نویسندهٔ این مقاله را نشان میدهد .



تصویر(۳) محتویات فلدر library نرم افزار R نصب شده در رایانهٔ نویسندهٔ مقاله

برخی از توابع کتابخانه ای در نوشتن این مقاله استفاده شده اند. عبارتند از

AER	تحلیلهای رگرسیون خطی و غیر خطی
Car	آزمونهای رگرسیونی
Dse1	الگوهای فضای حالت و فیلترینگ کالمن و ARMA برداری
Fseries	تخمین و پیش بینی الگوهای GARCH
Fracdiff	الگوهای اتورگرسیو میانگین متحرک جمع بسته جزئی ARFIMA
foreign *	بار کردن و ذخیره کردن داده ها از سایر برنامه های کاربردی
graphics *	گرافیک منحنیهای تراز و تصاویر بردارها
Lmtest	آزمونهای بریوش – گادفری و بریوش – پاگان
ineq	محاسبهٔ نابرابری در توزیع درآمد
MASS *	رگرسیونهای Robust در الگوهای لاجیت و پروبیت
MCMCpack	معکوس تابع توزیع گاما
MNP	الگوی پروبیت چند جمله ای از طریق MCMC
nlme *	تخمین الگوهای اثرات ثابت و تصادفی غیرخطی
nls *	حداقل مربعات غیر خطی

Nnet	الگوهای لاجیت و پروبیت چند جمله ای
Plm	الگوهای داده های تابلویی (اثرات مشترک ، ثابت و تصادفی)
Quantreg	رگرسیونهای کوانتیل
R.matlab	خواندن داده های ذخیره شده در نرم افزار MATLAB
Sandwicch	محاسبهٔ واریانس و کوواریانس در شرایط اختلال اتورگرسیو و واریانس ناهمسانی
survival *	رگرسیون داده های سانسور شده و توبیت
Systemfit	تخمینهای SUR زلنر و 2SLS برای سیستم معادلات همزمان
ts *	توابع دستکاری در سریهای زمانی
Tseries	توابع سریهای زمانی GARCH , ARIMA
urca	تحلیل سریهای زمانی جمع بسته و همجمع بسته[1]

توابع ستاره دار بر توابع کتابخانه ای دلالت دارد که پس از دانلود کردن R بطور پیش فرض در نرم افزار R وجود دارد .

۵ – عملیات ماتریسی :

زیر بنای محاسبات مربوط به تخمین ، پیش بینی و شبیه سازی در مطالعات اقتصادسنجی محاسبات و دستکاری در آرایه های ماتریسی است . نرم افزار R دارای این توانایی است که بتواند محاسبات ماتریسی را با سرعت و دقت بسیار زیاد انجام دهد . تصور کنید که میخواهید بردار x را با عناصر ذیل خلق کنید ، کافیست دستور ذیل را در پنجرهٔ دستور R اجراء کنید .

```
x <- c(10.4, 5.6, 3.1, 6.4, 21.7)
```

با تایپ مجدد x در پنجرهٔ دستور و اجراء آن ، نرم افزار R بردار x را به شما باز میگرداند . این بردار میتواند یک دنباله باشد مثلاً اگر بخواهیم دنباله ای از اعداد شامل ۵۱ مشاهده که با عدد (-5) آغاز شده و با نموی (0.2) افزایش می یابد را با عنوان S4 تولید کنیم ، داریم :

```
s4 <- seq(length=51, from=-5, by=.2)
```

با اجرای S4 داریم :

```
> S4
[1]  -5.0 -4.8 -4.6 -4.4 -4.2 -4.0 -3.8 -3.6 -3.4 -3.2 -3.0
[12] -2.8 -2.6 -2.4 -2.2 -2.0 -1.8 -1.6 -1.4 -1.2 -1.0 -0.8
[23]  0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  1.0  1.2  1.4
[34]  1.6  1.8  2.0  2.2  2.4  2.6  2.8  3.0  3.2  3.4  3.6
[44]  3.84  .0  4.2  4.4  4.6  4.8  5.0
```

برای ایجاد یک ماتریس (۵ × ٤)با عنوان X میتوان دستور ذیل را تایپ و اجراء کرد .

```
> x <- array(1:20, dim=c(4,5))
```

با تایپ X و اجرای آن در صفحهٔ دستور نرم افزار، ماتریس X نمایش داده میشود .

```
> x
     [,1] [,2] [,3] [,4] [,5]
[1,]    1    5    9   13   17
[2,]    2    6   10   14   18
[3,]    3    7   11   15   19
[4,]    4    8   12   16   20
```

[1] Integrated & Cointegrated

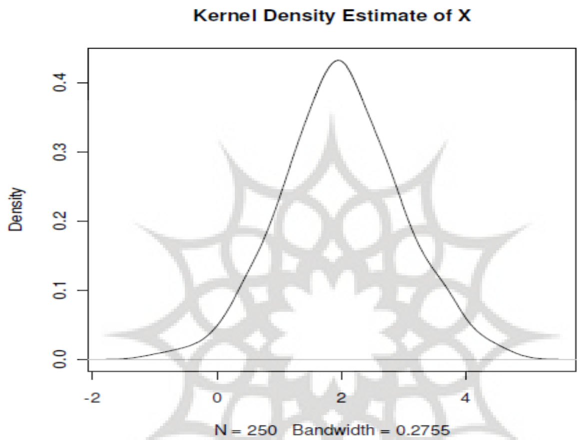
٦- گرافیک دو و سه بعدی :

یکی از قابلیتهای بسیار مهم نرم افزار R توانایی این نرم افزار در خلق گرافیکهای بـا کیفیـت دو و سـه بعدی است ، نرم افزار R به کاربر این توانایی را میدهد که بتواند در جزئیات تصاویر خلق شده دستکاری و ویراستاری کند . با اجرای دستور **plot(.)** در پنجرۀ دستور نرم افزار ، یک پنجره برای خلق یک نمودار ایجاد میگردد ، و با این دستور میتوان یک آرگومان که یک بردار از پیش تعریف شـده باشـد را رسـم کـرد . انواع نمودارهای پراکنش ، خطوط و هیستوگرام را میتوان در آرگومان **type** در دستور زیر معرفی کرد .

```
> plot(x,y,type="l", main="X and Y example",ylab="y values",xlab="x values")
```

گزینه های **ylab** و **xlab** هم برچسب هر محور را تعیین میکند . دستورات زیر متغیر تصادفیX را بـا حجم ۲۵۰ مشاهده با توزیع نرمال با میانگین ۲ و انحراف معیار ۱ خلق کرده و تابع چگالی احتمال این متغیر را بصورت کرنل[1] ترسیم میکند.[2]

```
x <- rnorm(250,mean=2,sd=1)
d<-density(x)
plot(d,main="Kernel Density Estimate of X")
```



تصویر(٤) تابع چگالی احتمال یک متغیر تصادفی نرمال

در نرم افزار R میتوان تصاویر را بصورت پارتیشن بندی شده با هم ادغام کرد . بـرای ایـن منظـور از تـابع **par(mfrow=c(a,b),mar=c(d1,d2,d3,d4))** استفاده میشـود کـه a نشـان دهنـده تعداد تصاویر بصورت سطری و b تعداد تصاویر بصورت ستونی است . **d4,d3,d2,d1** هم بـه ترتیـب حاشیۀ تصویر با پائین ، سمت چپ ، بالا و سمت راست را نشان میدهد . برنامۀ ذیل با داشتن ضریب جینـی چهار سال متوالی (۱۳۸۳-۱۳۸۰) برای مناطق شهری کشور ابتدا منحنی لـورنز هـر سـال را بدسـت آورده و سپس چهار تصویر را در یک تصویر خلاصه میکند .

```
################################################################
#        Plot Lorenz Curves with R 2.10.0                      #
#              Date : 26.04.1388                               #
################################################################
gini<- 0.4324
```

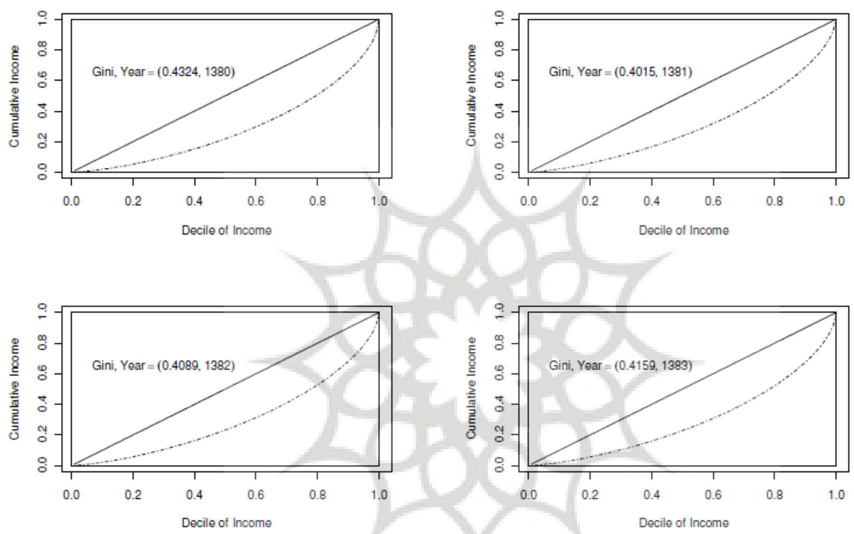
[1] Kernel

[2] برای اجرای هر برنامه در نرم افزار R پس از فراخوانی فایل متنی مورد نظر ، بایستی مسیر زیر طی شود .
Edit / Run all

```
gini1<- 0.4015
gini2<- 0.4089
gini3 <- 0.4159
year <- 1380
year1 <- 1381
year2 <- 1382
year3 <- 1383

a<- 1+gini;
b<- 1-gini;
Incomedicile<- array(0,dim=c(100,1));
Lorenz <- array(0, dim=c(100,1));
Equiltyline <- array(0, dim=c(100,1));
for (i in 1 : 100){
 Incomedicile[i] <- i/100 ;
 Lorenz[i]<- pbeta(i/100,a,b);
 Equiltyline[i]<-i/100;
}

a1<- 1+gini1;
b1<- 1-gini1;
Incomedicile<- array(0,dim=c(100,1));
Lorenz1 <- array(0, dim=c(100,1));
Equiltyline <- array(0, dim=c(100,1));
for (i in 1 : 100){
 Incomedicile[i] <- i/100 ;
 Lorenz1[i]<- pbeta(i/100,a1,b1);
 Equiltyline[i]<-i/100;
}

a2<- 1+gini2;
b2<- 1-gini2;
Incomedicile<- array(0,dim=c(100,1));
Lorenz2 <- array(0, dim=c(100,1));
Equiltyline <- array(0, dim=c(100,1));
for (i in 1 : 100){
 Incomedicile[i] <- i/100 ;
 Lorenz2[i]<- pbeta(i/100,a2,b2);
 Equiltyline[i]<-i/100;
}

a3<- 1+gini3;
b3<- 1-gini3;
Incomedicile<- array(0,dim=c(100,1));
Lorenz3 <- array(0, dim=c(100,1));
Equiltyline <- array(0, dim=c(100,1));
for (i in 1 : 100){
 Incomedicile[i] <- i/100 ;
 Lorenz3[i]<- pbeta(i/100,a3,b3);
 Equiltyline[i]<-i/100;
}
reza<- c(1,0,0,1,1);
ali<- c(1,1,0,0,1);
par(mfrow=c(2,2),mar=c(5,5,4,4))
plot(Incomedicile,Lorenz,type="l",axes=T, xlab="Decile of
Income",ylab="Cumulative Income",lty=4,col=1)
lines(Incomedicile,Equiltyline,lty=1, col=1)
lines(reza,ali,type="l",col=1);
text(0.3,0.65, substitute(list(Gini,Year) == group("(",list(x,y),")"),
list(x=gini, y=year)))
#title (" Fig A :Lorenz Curve  ")
plot(Incomedicile,Lorenz1,type="l",axes=T, xlab="Decile of
Income",ylab="Cumulative Income",lty=4,col=1)
lines(Incomedicile,Equiltyline,lty=1, col=1)
lines(reza,ali,type="l",col=1);
text(0.3,0.65, substitute(list(Gini,Year) ==
group("(",list(x,y),")"),list(x=gini1, y=year1)))
#title (" Fig B :Lorenz Curve ")
plot(Incomedicile,Lorenz2,type="l",axes=T, xlab="Decile of
Income",ylab="Cumulative Income",lty=4,col=1)
lines(Incomedicile,Equiltyline,lty=1, col=1)
lines(reza,ali,type="l",col=1);
```

```
text(0.3,0.65, substitute(list(Gini,Year) ==
group("(",list(x,y),")"),list(x=gini2, y=year2)))
#title (" Fig C :Lorenz Curve")
plot(Incomedicile,Lorenz3,type="l",axes=T, xlab="Decile of
Income",ylab="Cumulative Income",lty=4,col=1)
lines(Incomedicile,Equiltyline,lty=1, col=1)
lines(reza,ali,type="l",col=1);
text(0.3,0.65, substitute(list(Gini,Year) ==
group("(",list(x,y),")"),list(x=gini3, y=year3)))
#title (" Fig D :Lorenz Curve ")
```

نرم افزار R توانایی افزودن متن به تصویر را دارد . دستور **text(a,b," ...")** متن داخل گیومه را در مختصات a برای محور **x** و **b** برای محور **y** قرار میدهد . با بکارگیری دستور **title ("........")** میتوان متن داخل گیومه را به عنوان "تیتر تصویر " به نمودار اضافه کرد . این مثال فقط حاوی توانایی نرم افزار R در ترسیم نمودار نیست بلکه به لحاظ محاسباتی نیز بسیار حائز اهمیت است ، توضیح این بخش را در قسمت "توانایی محاسباتی نرم افزار R " ببینید



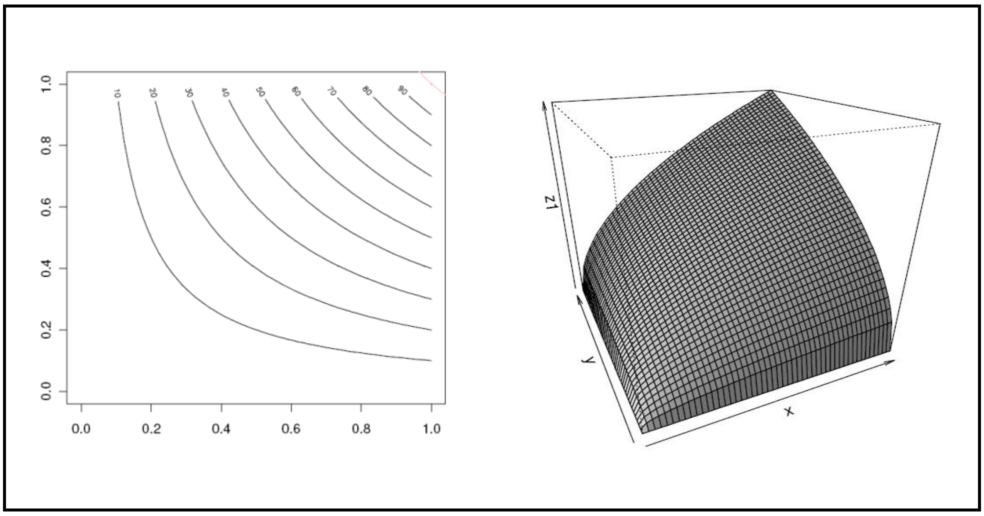
تصویر (۵) منحنیهای لورنز متفاوت در یک تصویر

نرم افزار R توانایی خلق گرافهای سه بعدی را دارد . بطور مثال تصور کنید که میخواهیم تابع تولید کاب داگلاس $z = x^{0.5}y^{0.5}$ را به همراه منحنی تراز[1] آن ترسیم کنیم . دستورات ذیل این نمودارها را در پنجرهٔ تصویر نرم افزار ایجاد میکند .

```
x <- seq(0,10,l=50)
y <- x
op <- par(mfrow = c(2, 1))
myf1 <- function(x,y){
x^0.5*y^0.5
}
par(mfrow(=c(2,2))
z1 <- outer(x,y, FUN = myf1)
persp(x,y,z1, theta=-25, phi=25, shade = 0.45)
contour(outer(x, y), method = "edge", vfont = c("sans serif", "plain"))
contour(x,y, z1, col = "pink", add = TRUE, method = "edge",
vfont = c("sans serif", "plain"))
par(op)
```

[1] Contour Curve

خروجی نرم افزار R برای ترسیم تابع تولید کاب داگلاس و منحنیهای تراز آن را در تصویر (٦) میتوان دید.



تصویر (٦) نمودار رویهٔ سه بعدی تابع تولید کاب داگلاس همراه با منحنی تراز آن

٧ – توانایی محاسباتی نرم افزار R :

برای این بخش باز به منحنی لورنز بخش (٦) باز میگردیم که در تصویر (٥) نشان داده شده است . این منحنهای لورنز را منحنی لورنز بتا می گویند ، زیرا فرض شده که منحنی لورنز از توزیع احتمالی بتا تبعیت میکند . تابع بتای ناقص را میتوان چنین تعریف کرد :

$$B(x;\alpha,\beta)=\frac{1}{B(\alpha\,,\,\beta)}\int_0^x t^{(\alpha-1)}.(1-t)^{\beta-1}\,dt \qquad 0\le x\le 1\,,\quad \alpha>0\quad,\quad \beta>0$$ رابطهٔ (١)

که منظور از $B(\alpha,\beta)$ تابع بتا است که بصورت ذیل تعریف میگردد :

$$B(\alpha\,,\,\beta)=\int_0^1 t^{\alpha-1}.(1-t)^{\beta-1}dt \qquad , \qquad \alpha\,,\,\beta>0$$ رابطهٔ (٢)

در ضمن تابع بتا با تابع گاما مرتبط است ، بطوریکه داریم :

$$B(\alpha\,,\,\beta)=\frac{\Gamma(\alpha)\,\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$ رابطهٔ (٣)

تابع بتای ناقص دلالت بر تابع بتای معرفی شده در رابطهٔ (١) دارد با این وصف که محدودیتهای خاصی بر پارامترهای α و β تحمیل میگردد ، تا منحنی لورنز مناسبی برای تخمین ضریب جینی بدست آید .

$$L(p)=\frac{1}{B(\alpha\,,\,\beta)}\int_0^p t^{(\alpha-1)}.(1-t)^{\beta-1}\,dt \qquad , \quad \alpha\ge 1 \quad , \quad 0<\beta\le 1$$ رابطهٔ (٤)

با این مختصر مشاهده میشود که برای ترسیم منحنی لورنز نیاز است که نرم افزار R قابلیتهای برنامه نویسی و توابع مربوط به توزیعهای احتمالی پیوسته استاندارد را داشته باشد . در برنامهٔ ای که محاسبات

منحنی لورنز و ترسیم آن منحنی را پردازش میکند ، از قابلیتهای خلق حلقه هـای محاسـباتی اسـتفاده شـده است ، که از مهمترین این حلقه های کنترل دستور for است که در حالت عمومی حلقهٔ for بصـورت زیـر است .

```
for (i in 1 : n){
 commands
}
```

که محاسباتی را رویcommands برایn بار تکرار صورت میدهد . ممکنست در برخی از موارد نیاز باشد تا حلقه ها بصورت متداخل باشند .

[illegible]

رگرسیونی دارای مشکل عدم تحقق فروض کلاسیک باشد ، در پاره ای از این موارد گفته میشـود کـه جملـهٔ اختلال ناهمواریست[1] ، مثـلاً وجـود اخـتلال اتورگرسـیو . در ایـن وضـعیت بایسـتی بـرای بدسـت آوردن تخمینهای حداقل مربعات تعمیم یافته (GLS) ماتریس Ω را که به مقدار ρ مرتبط اسـت، بدسـت آورد . همانطوریکه در کمنتا (۱۹۹۰) نشان داده شده است ماتریس Ω در شرایط وجود اختلال اتورگرسیو از مرتبهٔ اول AR(1) عبارتست از :

$$\Omega = \frac{1}{\left(1-\rho^2\right)} \cdot \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{T-1} \\ \rho & 1 & \rho & \cdots & \rho^{T-2} \\ \vdots & \vdots & & & \vdots \\ \vdots & \vdots & & & \vdots \\ \rho^{T-1} & \rho^{T-2} & \rho^{T-3} & \cdots & 1 \end{bmatrix}_{\Omega_{(T\times T)}} \qquad \text{رابطهٔ (۵)}$$

نتیجهٔ فوق از این رابطه بدست آمده است که کوواریانس بین دو جملهٔ اختلال عبارتست از :

$$E\left(U_{t-s} \cdot U_t\right) = \rho^{|t-(t-s)|} \cdot \sigma_u^2 = \rho^s \cdot \sigma_u^2 \qquad \text{رابطهٔ (٦)}$$

در اینصورت برای محاسبهٔ معکوس ماتریس Ω داریم :

$$\Omega^{-1} = \begin{bmatrix} 1 & -\rho & 0 & 0 & \cdots & 0 & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & 0 & \cdots & 0 & 0 & 0 \\ 0 & -\rho & 1+\rho^2 & -\rho & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -\rho & 1+\rho^2 & -\rho \\ 0 & 0 & 0 & 0 & \cdots & 0 & -\rho & 1 \end{bmatrix} \qquad \text{رابطهٔ (۷)}$$

برای بدست آوردن Ω , Ω^{-1} میتوان از دستورات زیر که از دو حلقهٔ for متداخل[2] تشکیل شـده اسـتفاده کرد . در این تابع n تعداد مشاهدات است که برای سهولت ٤ در نظر گرفته شده است و مقـدار $\rho = 0.5$ در نظر گرفته شده است .

```
n=4
rho<- .5
omega<-array(0,dim=c(n,n))
for(i in 1:n){
for(j in 1:n){
omega[i,j]=(rho^abs(i-j))*(1/(1-rho^2))
}
}
omega
omegainv=solve(omega)
```

---

[1] Nonsphrical
[2] Nonnested

```
omegainv
```

با اجرای آن داریم :

```
> omega
          [,1]      [,2]      [,3]      [,4]
[1,] 1.3333333 0.6666667 0.3333333 0.1666667
[2,] 0.6666667 1.3333333 0.6666667 0.3333333
[3,] 0.3333333 0.6666667 1.3333333 0.6666667
[4,] 0.1666667 0.3333333 0.6666667 1.3333333
> omegainv=solve(omega)
> omegainv
     [,1]  [,2]  [,3] [,4]
[1,]  1.0 -0.50  0.00  0.0
[2,] -0.5  1.25 -0.50  0.0
[3,]  0.0 -0.50  1.25 -0.5
[4,]  0.0  0.00 -0.50  1.0
```

علاوه بر این حلقهٔ کنترلی ، سایر حلقه های کنترلی دیگر را میتوان در کتابچهٔ راهنمـای الکترونیکـی نـرم افزار – که بصورت رایگان در تارنمای نرم افزار قرار داده شده است– مشاهده کرد .

۸ – شبیه سازی مونت کارلو [1]

مثال ۱ : فرض کنید که میخواهیم خواص تخمین زننده های حداقل مربعات معمولی الگوی رگرسیونی را بطور تجربی و با استفاده از شبیه سازی مونت کارلو بررسی کنیم . تصور کنید که میخواهیم الگوی رگرسیونی چند جمله ای از درجهٔ دوم را با پارامترهای معلوم بصورت رابطهٔ (۸) که دارای ۳۰ مشاهده است را به تعداد ۱۰۰۰۰ بار با خلق بردار تصادفی جملهٔ اختلال که دارای توزیع نرمال استاندارد است را تولید کرده و سـپس به تعداد ۱۰۰۰۰ بار رگرسیون نمونه را برازش کرده و آماره های $\hat{\alpha}, \hat{\beta}_2, \hat{\beta}_3$ را بررسی کنیم . با استفاده از رایانه پنتیوم ٤ و با CPU ۳/۲ گیگا هرتز ، برنامهٔ R این محاسبات را در ٥٤ ثانیه انجام میدهد ، حـال آنکـه در نرم افزار EViews نسخهٔ ٤/۱ این محاسبات ٥٦٠ ثانیه زمان خواهد برد .

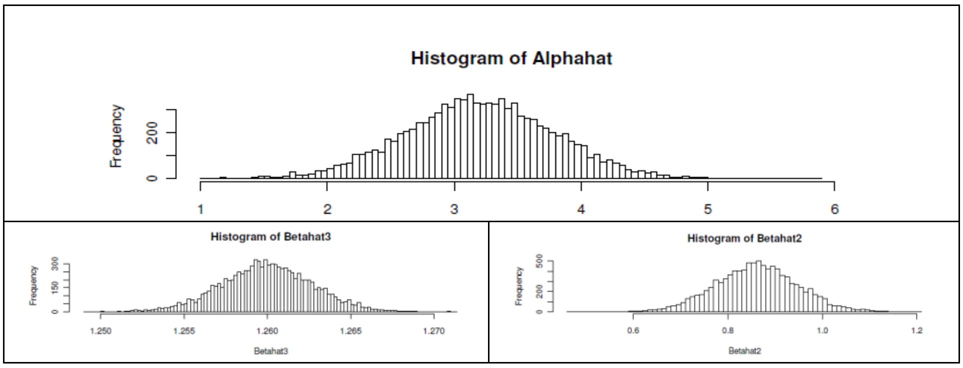
$$Y_i = 3.2 + 0.86 X_i + 1.26 X_i^2 + U_i \qquad \text{رابطهٔ (۸)}$$

```
######################################################################
#    Monte Carlo Simulation Experiment With R 2.10.2                #
#        For PRM : Yi=3.2+0.86*X2i+1.26*X3i+Ui                       #
######################################################################
Alphahat <- array(0, dim=c(10000,1))
Betahat2 <- array(0, dim=c(10000,1))
Betahat3 <- array(0, dim=c(10000,1))
x2<-1:30
x3<- x2^2
for(i in 1:10000){
y <- rnorm(30, mean=(3.2+0.86*x2+1.26*x3), sd=1)
beta <- lm(y~x2+x3)
Alphahat[i]<- beta$coef[1]
Betahat2[i] <- beta$coef[2]
Betahat3[i] <- beta$coef[3]
}
par(mfcol = c(3, 1))
hist(Alphahat ,100)
hist(Betahat2 ,100)
hist(Betahat3 ,100)
AlphaExp <- mean(Alphahat)
```

[1] Monte Carlo Simulation

```
BiasAlpha <- 3.2-AlphaExp
varAlpha <- var(Alphahat)
Beta2Exp <- mean(Betahat2)
BiasBeta2 <- 0.86-Beta2Exp
varBeta <- var(Betahat2)
Beta3Exp <- mean(Betahat3)
BiasBeta3 <- 1.26-Beta3Exp
varBeta3 <- var(Betahat3)
```

خروجی این برنامه را میتوان در تصویر (۷) و خروجی (۱) مشاهده کرد .



تصویر (۷) هیستوگرام تخمین زننده های الگوی رگرسیونی در یک آزمایش شبیه سازی مونت کارلو

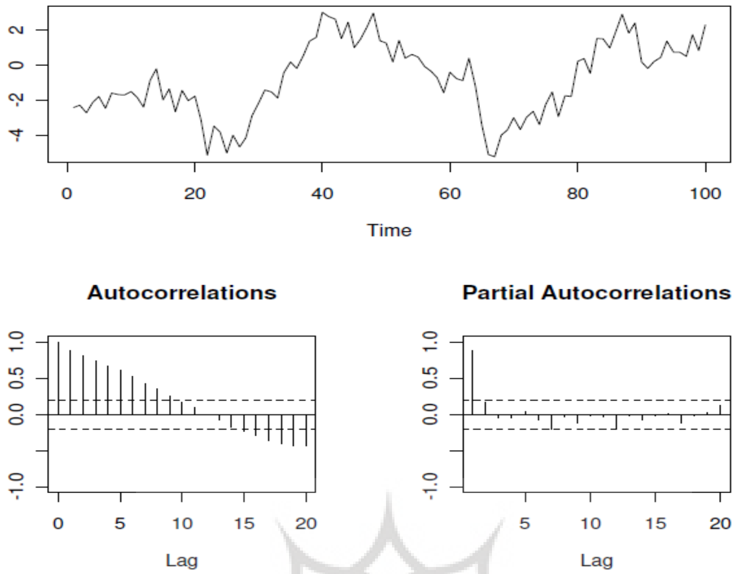
> AlphaExp	امید ریاضی تخمین زنندۀعرض از مبداء
[1] 3.192430	
> BiasAlpha	میزان اریب عرض از مبداء
[1] 0.007570475	
> varAlpha	واریانس تخمین زنندۀ عرض از مبداء
[1,] 0.3469190	
> Beta2Exp	امیدریاضی $\hat{\beta}_2$
[1] 0.8607964	
> BiasBeta2	اریب $\hat{\beta}_2$
[1] -0.0007963932	
> varBeta	واریانس $\hat{\beta}_2$
[1,] 0.007825871	
> Beta3Exp	امیدریاضی $\hat{\beta}_3$
[1] 1.259981	
> BiasBeta3	اریب $\hat{\beta}_3$
[1] 1.913949e-05	
> varBeta3	واریانس $\hat{\beta}_3$
[1,] 7.683248e-06	

خروجی (۱) خروجی نرم افزارR برای آزمایش مونت کارلو

مثال ۲ : تابع زیر به شبیه سازی یک فرآیند **AR(1)** با ۱۰۰ مشاهده و با ضریب اتورگرسیو ۰/۹ اشاره دارد که برای شناسایی این فرآیند توابع خودهمبستگی (ACF) و خودهمبستگی جزئی (PACF) آنرا ترسیم می کند.

```
set.seed(123456)
y <- arima.sim(n = 100, list(ar = 0.9), innov=rnorm(100))
```

```
op <- par(no.readonly=TRUE)
layout(matrix(c(1, 1, 2, 3), 2, 2, byrow=TRUE))
plot.ts(y, ylab='')
acf(y, main='Autocorrelations', ylab='',
    ylim=c(-1, 1), ci.col = "black")
pacf(y, main='Partial Autocorrelations', ylab='',
     ylim=c(-1, 1), ci.col = "black")
par(op)
```

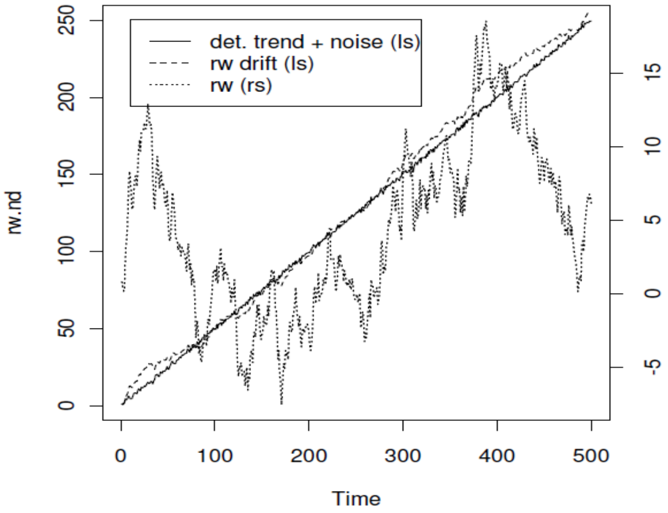


تصویر(۸) یک فرآیند شبیه سازی شدهٔ AR(1) همراه با توابع PACF, ACF

مثال ۳ : توابع ذیل تفاوت بین یک فرایند گام تصادفی خالص[1]، گام تصادفی همراه با رانش[2] و یک روند زمانی قطعی[3] جمع شده با فرآیند نوفهٔ سفید[4] برای ۵۰۰ مشاهده شبیه سازی میکند .

```
set.seed(123456)
e <- rnorm(500)
## pure random walk
rw.nd <- cumsum(e)
## trend
trd <- 1:500
## random walk with drift
rw.wd <- 0.5*trd + cumsum(e)
## deterministic trend and noise
dt <- e + 0.5*trd
## plotting
par(mar=rep(5,4))
plot.ts(dt, lty=1, ylab='', xlab='')
lines(rw.wd, lty=2)
par(new=T)
plot.ts(rw.nd, lty=3, axes=FALSE)
axis(4, pretty(range(rw.nd)))
lines(rw.nd, lty=3)
legend(10, 18.7, legend=c('det. trend + noise (ls)', 'rw drift (ls)',
'rw (rs)'), lty=c(1, 2, 3))
```

1 Pure Random walk
2 Random walk with drift
3 Determinestic Time Trend
4 White Noise

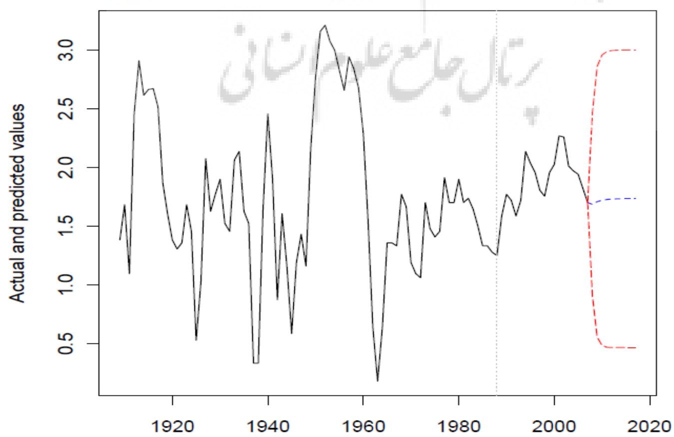


تصویر (۹) فرآیندهای گام تصادفی ساده و همراه با رانش و روند قطعی جمع شده با نوفهٔ سفید

۹ – پیش بینی در الگوهای باکس و جنکینز :

دستورات زیر برای پیش بینی یک فرآیند ARMA (1,1) نوشته شده است که پس از تخمین الگوی یاد شده سری زمانی را برای ۱۰ دوره پیش بینی میکند . در این تصویر کران بالایی و پائینی برای ± 2 انحراف معیار محاسبه و ترسیم شده است .

```
## Forecasts
arma11.pred <- predict(arma11, n.ahead = 10)
predict <- ts(c(rep(NA, length(y) - 1), y[length(y)],
                arma11.pred$pred), start = 1909, frequency = 1)
upper <- ts(c(rep(NA, length(y) - 1), y[length(y)], arma11.pred$pred + 2
* arma11.pred$se),  start = 1909, frequency = 1)
lower <- ts(c(rep(NA, length(y) - 1), y[length(y)], arma11.pred$pred - 2
* arma11.pred$se), start = 1909, frequency = 1)
observed <- ts(c(y, rep(NA, 10)), start=1909, frequency = 1)
## Plot of actual and forecasted values
plot(observed, type = "l",ylab = "Actual and predicted values", xlab =
"")
lines(predict, col = "blue", lty = 2)
lines(lower, col = "red", lty = 5)
lines(upper, col = "red", lty = 5)
abline(v = 1988, col = "gray", lty = 3)
```



تصویر (۱۰) پیش بینی یک فرآیند ARMA (1,1)

۱۰ – آزمون ریشهٔ واحد (دیکی فولرتعمیم یافته) :

در نرم افزار **R** با بهره گیری از توابع کتابخانه ای **URCA** براحتی میتوان آزمون ریشه های واحد را با تنوع بسیار زیاد انجام دهد . در اینجا با ارائهٔ مثالی از کتاب " هم جمعبستگی برای اقتصاددانان کاربردی[1]" که برگرفته از مقاله ای از "هولدن و پرمن (۱۹۹٤)[2]" است ، داده های فصلی لگاریتم مصرف **(LC)** ، برای پادشاهی بریتانیا در خلال دورهٔ (۱۹۹۱:۲ – ۱۹٦٦:٤) مورد آزمون ریشهٔ واحد قرار میگیرد .

```
library(urca)
data(Raotbl3)
attach(Raotbl3)
lc.df <- ur.df(y=lc, lags=3, type='trend')
summary(lc.df)
```

با اجرای این دستورات نتایج ذیل را بدست می آوریم .

```
###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################
Test regression trend
Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.7976591  0.3547775   2.248   0.0270 *
z.lag.1     -0.0758706  0.0338880  -2.239   0.0277 *
tt           0.0004915  0.0002159   2.277   0.0252 *
z.diff.lag1 -0.1063957  0.1006744  -1.057   0.2934
z.diff.lag2  0.2011373  0.1012373   1.987   0.0500 .
z.diff.lag3  0.2998586  0.1020548   2.938   0.0042 **

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01307 on 89 degrees of freedom
Multiple R-squared: 0.1472,     Adjusted R-squared: 0.09924
F-statistic: 3.071 on 5 and 89 DF,  p-value: 0.01325

Value of test-statistic is: -2.2389 3.7382 2.5972

Critical values for test statistics:
      1pct  5pct 10pct
tau3 -4.04 -3.45 -3.15
phi2  6.50  4.88  4.16
phi3  8.73  6.49  5.47
```

همانطوریکه از نتایج پیداست ، آمارهٔ آزمون محاسباتی ۲/۲۳۹– و مقادیر بحرانی در سطح ۵٪ معادل ۳/٤٥– است (اعدادی که در خروجی فوق بصورت **underline** گزارش شده اند) که این حکایت از وجود یک ریشهٔ واحد در سری زمانی یاد شده دارد .

[1] Cointegration for the Applied Economists.

[2] Holden, Darryl and Roger Perman (1994) .

۱۱- آزمون هم جمعبستگی[1] :

در این بخش با استفاده از داده های کشور دانمارک طی دورهٔ زمانی (۱۹۸۷:۳-۱۹۷٤:۱) ارائه شده توسط جوهانسن و جوسیلیوس[2] برای سریهای زمانی لگاریتم حجم نقدینگی (LRM) ، لگاریتم درآمد واقعی (LRY) ، نرخ بهرهٔ اوراق قرضه (IB0) و نرخ سپردهٔ بانکی (IDE) آزمون حداکثر راستنمایی هم جمعبستگی جوهانسن و جوسیلیوس را توسط نرم افزار R ارائه کرده و آزمون ماتریس اثر و حداکثر مقدار ویژه برای تعیین تعداد بردارهای هم جمعبستگی را ارائه میکنیم .

```
library(urca)
data(denmark)
sjd <- denmark[, c("LRM", "LRY", "IBO", "IDE")]
sjd.vecm <- ca.jo(sjd, ecdet = "const", type="eigen", K=2,
spec="longrun", season=4)
summary(sjd.vecm)
```

با اجرای این دستورات نتایج ذیل را میتوان بدست آورد . نتایج حکایت از آن دارد که یک بردار هم جمعبستگی را میتوان برای ترکیب این سریهای متصور شد .

```
#################################################################
# Johansen-Procedure                                            #
#################################################################
Test type: maximal eigenvalue statistic (lambda max) , without linear
trend and constant in cointegration
Eigenvalues (lambda):
[1] 4.331654e-01 1.775836e-01 1.127905e-01 4.341130e-02 5.012875e-16
Values of teststatistic and critical values of test:

          test 10pct  5pct  1pct
r <= 3 |  2.35  7.52  9.24 12.97
r <= 2 |  6.34 13.75 15.67 20.20
r <= 1 | 10.36 19.77 22.00 26.81
r = 0  | 30.09 25.56 28.14 33.24

Eigenvectors, normalised to first column:
(These are the cointegration relations)

            LRM.l2     LRY.l2     IBO.l2     IDE.l2   constant
LRM.l2    1.000000  1.0000000  1.0000000   1.000000  1.0000000
LRY.l2   -1.032949 -1.3681031 -3.2266580  -1.883625 -0.6336946
IBO.l2    5.206919  0.2429825  0.5382847  24.399487  1.6965828
IDE.l2   -4.215879  6.8411103 -5.6473903 -14.298037 -1.8951589
constant -6.059932 -4.2708474  7.8963696  -2.263224 -8.0330127

Weights W:
(This is the loading matrix)
           LRM.l2      LRY.l2      IBO.l2       IDE.l2      constant
LRM.d -0.21295494 -0.00481498  0.035011128  2.028908e-03 -3.604015e-13
LRY.d  0.11502204  0.01975028  0.049938460  1.108654e-03  4.282155e-13
IBO.d  0.02317724 -0.01059605  0.003480357 -1.573742e-03  3.409081e-14
IDE.d  0.02941109 -0.03022917 -0.002811506 -4.767627e-05 -9.927144e-15
```

۱۲) الگوی سوئچینگ مارکف :

انگل و هامیلتون (۱۹۹۰)

چرخشهای بلند مدت[3] نرخ ارز (دلار در مقابل مارک آلمان ، فرانک فرانسه و پوند انگلیس) را به الگو

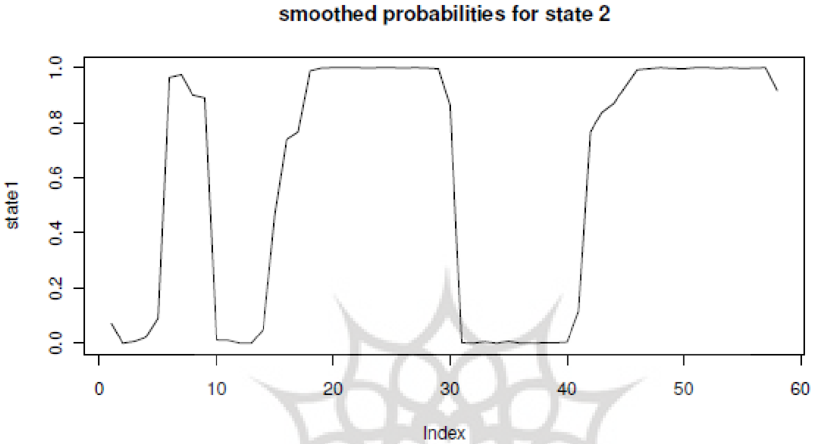
[1] Cointegration Test

[2] Johansen, S. and Juselius, K. (1990)

[3] Long Swings

درآورند . آنها با استفاده از این مدل اثبات کرده اند که نرخ دلار از یک فرآیند گام تصادفی تبعیت نمیکند . آنها نشان داده اند که احتمال وقوع این چرخشها ، چه زمانی بوقوع می پیوندد . تصویر (۱۱) زمان گـذاری چرخشهای دلار در مقابل ارزهای اروپایی آن دوره را نشان میدهد . دسـتورات ایـن بخـش در تـابع ذیـل آورده شده است ، اجرای آن مستلزم استفاده از توابع کتابخانه ای MSVAR است .

```
library(MSVAR)
Use the data-set 'Hamilton_Data.csv' to reproduce Hamilton's results
from "Long Swings in the Dollar: Are They in the Data and
#Do Markets Know It?", American Economic Review, Sept. 1990
data("EX_R")
data("INT_D")
dep<-cbind(EX_R,INT_D)
res<-MS_Var(dep)
```



تصویر (۱۱) احتمال وقوع تغییر در ارزش دلار امریکا

۱۳ – جمع بندی :

با توجه به گسترش مطالعات کاربردی در اقتصاد و همچنین نیاز مبرم به پردازش داده هـا و الگوهـای پیچیده ، تقاضای روز افزونی برای نرم افزارهای اقتصادسنجی بچشم میخورد . در این میان زبانهـای برنامـه نویسی که با محدودیت نرم افزارهای محاوره ای مواجه نیستند، از سوی محققان، دانشجویان رشتهٔ اقتصـاد و سایر رشته های پیرامونی مورد استقبال قرار گرفته اند. نرم افزار R به واسطهٔ قابلیـت هـای بـالا، رایگـان بودن آن و امکان دسترسی به توابع کتابخانه ای سایر محققان اقصی نقاط دنیا، در بین سـایر نـرم افزارهـای برنامه نویسی از جایگاه خاصی برخوردار استلین مقاله با توجه به محدودیت حجم صر فاً به گوشه هایی از این امکانات پرداخته است .

منابع :

Banerjee, A , J & J. Dolado, J & Galbraith, and D, F. Hendry. (2003). " Cointegration, Error Correction, and The Econometric Analysis Of Nonstationary Data " , Oxford University Press.

Braun, W, J . and D, J, Murdoch, (2007), " A First Course in Statistical Programming with R" Cambridge University Press .

Chambers, J, M. (2008)."Software for Data Analysis :Programming with R ". Springer.

Crawley, M, J.(2007). " The R Book" , John Wiley & Sons Ltd, New York .

Cribari-Neto, F. and Mark, J,Jensen. (1997). " MATLAB as an Econometric Programming Environment ". Journal of Applied Econometrics. Vol. 11, No. 6 , PP. 735-744.

Cribari-Neto, F. (1997), " Econometric Programming Environments : GAUSS, Ox and S-PLUS " Journal of Applied Econometrics, Vol. 12 , No. 1, PP. 77-89 .

Cribani-Neto, F. and S,G, Azrkos. (1999). "R: Yet Another Econometric Programming Environment " . Journal of Applied Econometrics. Vol. 14 . PP, 319-329 .

Engel, C. and J, D, Hamilton. (1990)," Lomg swings in the Dollar: are they in the data and Do Markets Know it? " The American Economic Review, Vol, 80, No.4 , PP. 689-713 .

Holden, Darryl and Roger Perman (1994), Unit Roots and Cointegration for the Economist, in: Cointegration for the Applied Economist, ed. B. Bhaskara Rao, chapter 3, Data Appendix, Table D.3.

Johansen, S. and Juselius, K. (1990), Maximum Likelihood Estimation and Inference on Cointegration – with Applications to the Demand for Money, Oxford Bulletin of Economics and Statistics, **52, 2**, 169–210.

Kleiber, C. and A, Zeileis. (2008). " Applied Econometrics With R ". Springer .

Kmenta , J.(1990). " Elements of econometrics " , 2nd Edition. Maxwell Macmillan International Editions.

Krause, A. and M,Olson . (2004)," The Basic of S-PLUS", 4th Edition, Springer.

Pfaff, B .(2008). " Analysis of Integrated and Cointegrated Time Series With R ", 2nd Edition. Springer .

Swayne,D. F. (2007), " Interactive and dynamic Graphics for Data Analysis with R " . Springer .