第4回メモ

**ハッシュタグ：#ybkstat**

時系列データの解析

# 時系列データ

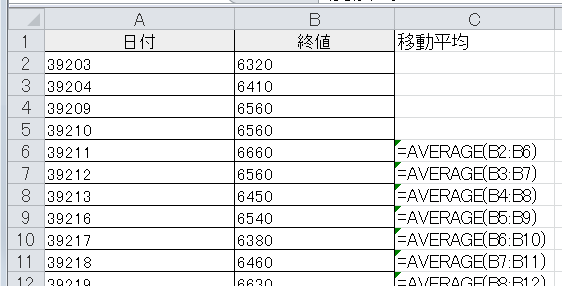
* **傾向変動（トレンド）**
* **周期変動（季節変動）**
* **不規則変動**

# 移動平均

時系列データの傾向的な変動を見るのに使うのが移動平均である。

## 関数で計算する方法

短期の傾向変化を調べるために、5区間の**移動平均**を求めてみる。

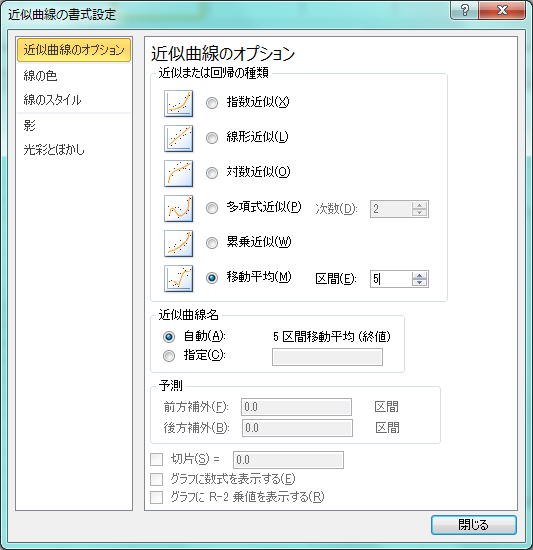


## 折れ線グラフに移動平均を重ねる方法

株価（終値）のデータを読み込み、折れ線グラフを描いてみよう。このような時系列データを可視化するときは、折れ線グラフを用いる。

短期の傾向変化を調べるために、5区間の**移動平均**を求めてみる。

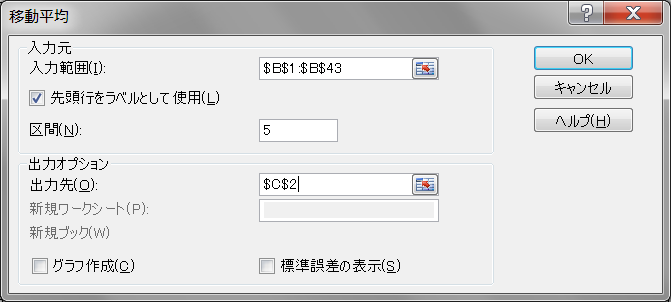
回帰直線を描いた時と同じ要領で、グラフの折れ線をクリック→右クリック→近似直線の追加。



長期の傾向変化を調べるために、10区間の移動平均を求めてみる。

短期の移動平均が長期の移動平均を下から抜いた後は株価が上昇する。上から抜いたとき、株価は下降する。

## 分析ツールで計算する方法



# 時系列データの分解

時系列データ＝トレンド＋周期変動＋残差

## 準備

GDPデータをExcelで開き、GDPの書式を「標準」にしてからCSV形式で保存する。

setwd("c:/cit")

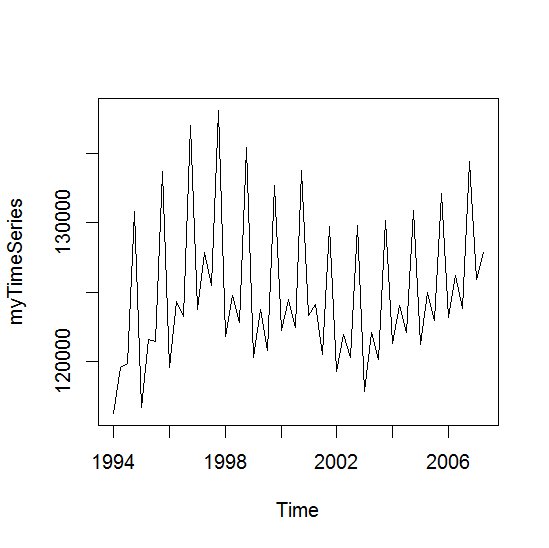
myData <- read.csv("GDPデータ.csv")

head(myData)

#生データの3列目を使って、1994年の1番目の四半期からの、時系列データを作る

myTimeSeries <- ts(myData[, 3], start = c(1994,1), frequency = 4)

plot(myTimeSeries) #結果の確認

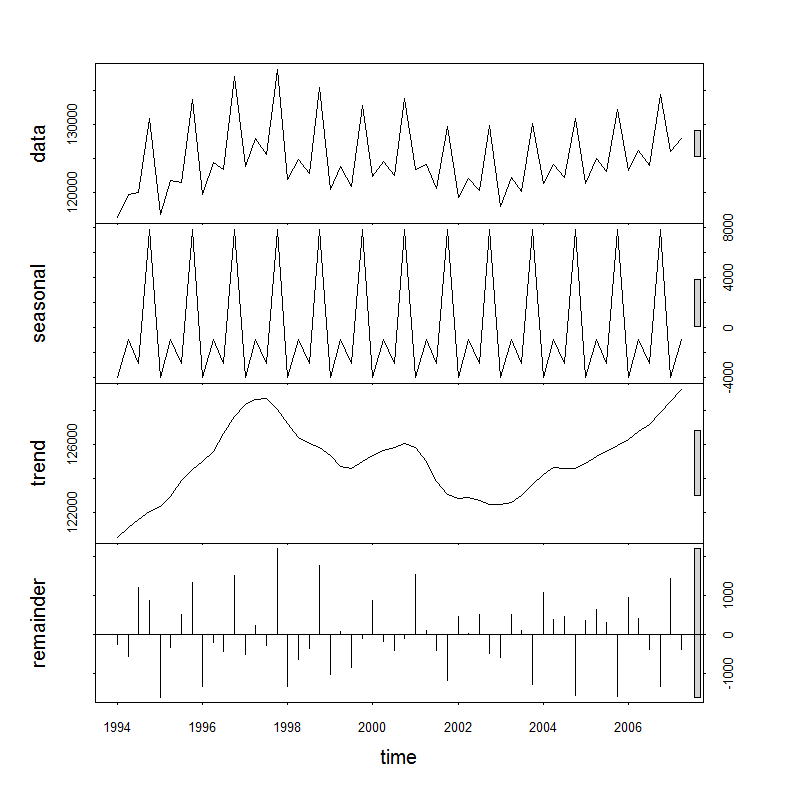


## 分解

#時系列データを、トレンドと周期変動、残差に分解する。

myResult <- stl(myTimeSeries, s.window = "periodic")

plot(myResult) #結果の確認



# 時系列データの予測

## forecastパッケージのインストール

次のコマンドで、予測のためのパッケージ（forecast）をインストールする（この例では、ダウンロードサイトをhttp://cran.ism.ac.jpに設定しているが、設定を省略すると、ダウンロードサイトを選択するためのダイアログが現れる）。

options(repos=c(CRAN="http://cran.ism.ac.jp")) #省略可

install.packages("forecast")

## 予測

パッケージをロードする。ここで「 ‘forecast’ という名前のパッケージはありません」というエラーになる場合は、パッケージのインストールに失敗している（参考：パッケージインストール時の注意）。

library(forecast)

時系列データをモデル化し、50%と90%の予測区間で、10ステップ分を予測する。

setwd("c:/cit")

myData <- read.csv("GDPデータ.csv")

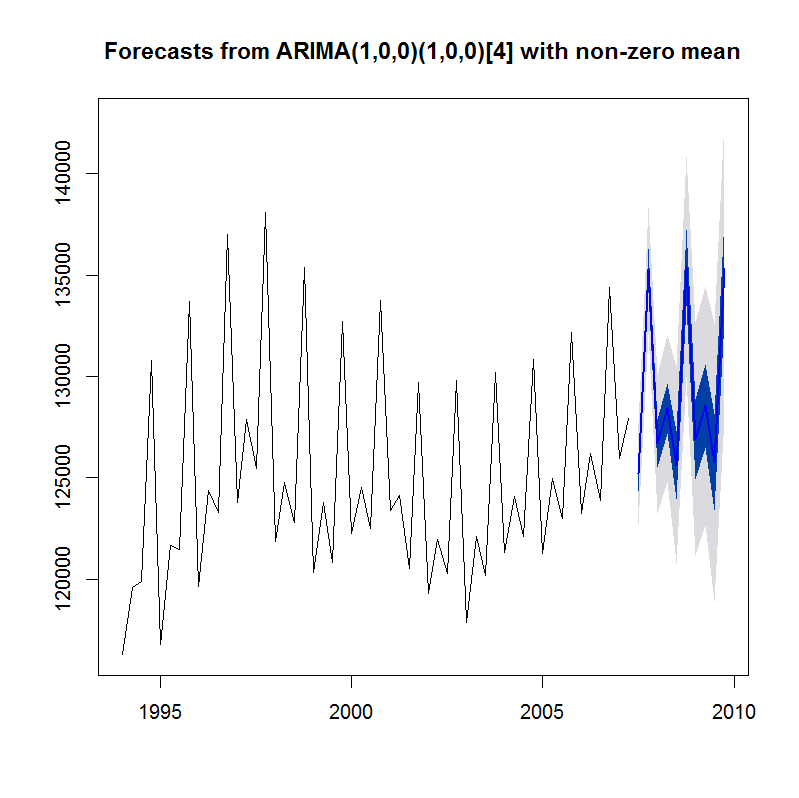
head(myData)

myTimeSeries <- ts(myData[, 3], start = c(1994,1), frequency = 4)

#ここまでは前と同じ

myModel <- auto.arima(myTimeSeries) #モデルの構築

(myForecast <- forecast(myModel, level = c(50, 95), h = 10)) #予測

plot(myForecast) #予測結果の表示

# レポート課題

「東京の平均気温データ」（1961年から2006年）を使い、2013年の東京の平均気温を予測せよ。

ヒント：データをRで読み込み、プロットするコードは次のようになる。

setwd("c:/cit")

myData <- read.csv("東京の平均気温データ.csv")

head(myData)

#生データの2列目を使って、1961年からの時系列データを作る

myTimeSeries <- ts(myData[, 2], start = c(1961))

plot(myTimeSeries)