7: Time Series

John H Maindonald

June 18, 2018

Ideas and issues illustrated by the graphs in this vignette

In a time series, the value at the current time can be correlated with the value at the previous time, perhaps with the values several time points back. Such correlations lead to a series that exhibits what is known as "serial dependence".

If possible, the analyst will want to find covariates that largely or partly explain that dependence. At best, such covariates will commonly explain part only of the dependence, and there will remain dependence that requires to be modeled.

A result of positive serial dependence is that large changes take time to build up. Smoothing terms can be fitted to the pattern thus generated, leaving *errors* that are pretty much uncorrelated. Such a pattern is in general, however, unrepeatable. It gives little clue of what may happen the future.

What sorts of patterns are unlikely to be generated by serial correlation, and may be repeatable? Indications that a pattern may be repeatable include:

- A straight line trend is a good starting point for some limited extrapolation. But think: Is it plausible that the trend will continue more than a short distance into the future?
- There may be a clear pattern of seasonal change, e.g., with seasons of the year or (as in the case of airborne pollution) with days of the week. If yearly seasonal changes persist over different years, or weekly day-of-the-week changes persist over different weeks, these effects can perhaps be extrapolated with some reasonable confidence.
- There is a regression relationship that seems likely to explain future as well as current data.

```
# To include the figures, change `showFigs <- FALSE`
# to `showFigs <- TRUE` in the source `.Rnw` file,
# and regenerate the PDF.
#
showFigs <- FALSE
```

1 Code for the figures

```
fig7.1 <- function(){
Erie <- greatLakes[,"Erie"]
plot(Erie, xlab="",
     ylab="Level (m)")
}
```

```
fig7.2 <- function(){
    Erie <- greatLakes[,"Erie"]
    opar <- par(oma=c(0,0,4,0))
    lag.plot(Erie, lags=3,
             do.lines=FALSE,
             layout=c(2,3), main="")
    mtext(side=3, line=3, adj=-0.155,
          "A: Lag plots, for lags 1, 2 and 3 respectively", cex=1)
    par(fig=c(0,1,0,0.6), new=TRUE)
    par(mar=c(2.75, 3.1, 3.6, 1.6))
    acf(Erie, main="", xlab="")
    mtext(side=3, line=0.5, "B: Autocorrelation estimates at successive lags",
          adj=-0.35, cex=1)
    mtext(side=1, line=1.75, "Lag", cex=1)
    par(fig=c(0,1,0,1))
    par(opar)
}
```

```
fig7.3 <- function(){
    Erie <- greatLakes[,"Erie"]
    df <-  data.frame(height=as.vector(Erie), year=time(Erie))
    obj <- gam(height ~ s(year), data=df)
    plot(obj, shift=mean(df$height), residuals=T, pch=1, xlab="")
}
```

```
fig7.4 <- function(){
    if(!require(forecast))return("Package 'forecast' must be installed")
    Erie <- greatLakes[,"Erie"]
    assign('Erie', Erie, pos=1)
    erie.ar <- ar(Erie)
    plot(forecast(erie.ar, h=15), ylab="Lake level (m)")
}
```

```
fig7.5 <- function(mf=3,nf=2){
    opar <- par(mfrow=c(mf,nf), mar=c(0.25, 4.1, 0.25, 1.1))
    npanel <- mf*nf
    for(i in 1:npanel){
        df <- data.frame(x=1:200, y=arima.sim(list(ar=0.7), n=200))
        df.gam <- gam(y ~ s(x), data=df)
        plot(df.gam, residuals=TRUE)
    }
    par(opar)
}
```

```
fig7.6 <- function(){
    mdbRain.gam <- gam(mdbRain ~ s(Year) + s(SOI), data=bomregions2012)
    plot(mdbRain.gam, residuals=TRUE, se=2, pch=1, cex=0.5, select=1)
    plot(mdbRain.gam, residuals=TRUE, se=2, pch=1, cex=0.5, select=2)
}
```

2 Show the Figures

```
pkgs <- c("DAAG","mgcv","splines","forecast")
z <- sapply(pkgs, require, character.only=TRUE, warn.conflicts=FALSE)
if(any(!z)){
  notAvail <- paste(names(z)[!z], collapse=", ")
  print(paste("The following packages require to be installed:", notAvail))
}
```

```
fig7.1()
```

```
fig7.2()
```

```
fig7.3()
```

```
fig7.4()
```

```
fig7.5()
```

```
fig7.6()
```