Stat 221, Time Series Analysis. Day 3.

Outline for the day:

1. Random Walks (RW) with drift.

2. Signal plus noise.

3. Mean function.

4. Autocovariance function.

1. Random Walk (RW) with drift.

Suppose $x_0 = 0$, and for t = 1, 2, ...,

$x_t = x_{t-1} + \delta + w_t$.

Then x_t is called a Random Walk (RW) with drift.

The constant δ is called the *drift*, and when $\delta = 0$, x_t is called a RW.

An example with and without drift is shown in Fig 1.10.

It is called a random walk because when $\delta = 0$, the value of x_t at time t is equal to x_{t-1} plus the completely random term w_t .

We can write x_t as a sum of the first t white noise terms. For a RW with drift,

$x_t = x_{t-1} + \delta + w_t = [x_{t-2} + \delta + w_{t-1}] + \delta + w_t = x_{t-2} + 2\delta + w_{t-1} + w_t$, and continuing to plug in $x_{t-2} = x_{t-3} + \delta + w_{t-2}$, etc., we eventually get

$x_t = \delta t + \sum_{i=1}^{t} w_i$.

Note that the code in the book has the line set.seed(154) .

This is useful, because it sets the random seed so that your results will look exactly like those in the book. This is very handy if you want to be able to reproduce your results exactly over and over. You don't have to use the number 154 unless you want to reproduce the results they happen to have generated in the book.

For instance, try

```
set.seed(123)
```

```
x = rnorm(1); print(x)

x = rnorm(1); print(x)

set.seed(123)

x = rnorm(1); print(x)
```

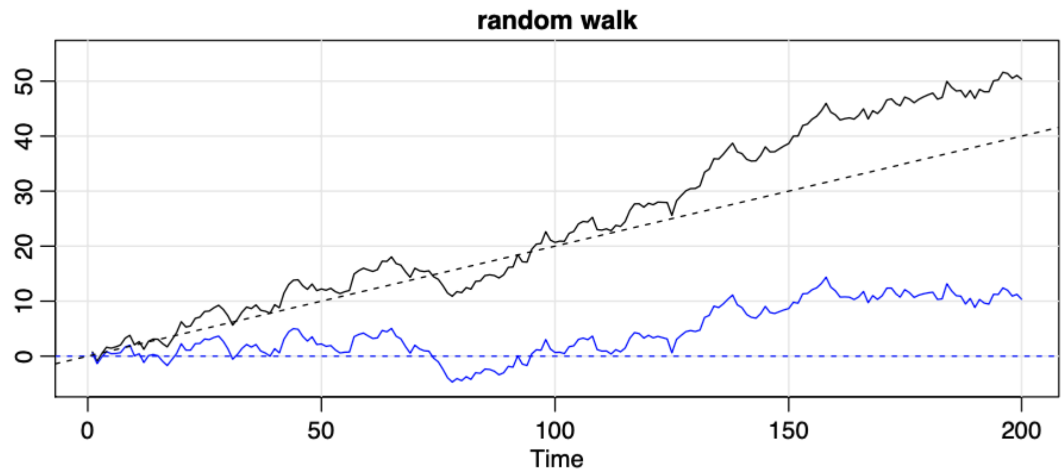

Fig. 1.10. *Random walk, $\sigma_w = 1$, with drift $\delta = .2$ (upper jagged line), without drift, $\delta = 0$ (lower jagged line), and straight (dashed) lines with slope δ.*

Suppose we want to simulate $x_t = 0.7\ x_{t-1} + 0.6\ x_{t-2} + w_t$, where $s^2 = v(w_t)=1$.

And suppose we want to MA filter it generating

$y_t = 0.5\ x_{t-1} + 0.25\ x_{t-2} + 0.25\ x_{t-3}$.

```
w = rnorm(150,0,1)

x = filter(w, filter=c(0.7,0.6), method="recursive")[-(1:50)]

y = filter(x, c(0.5,0.25,0.25), sides = 1) plot.ts(x)

lines(y,lty=2)
```

2. Signal plus noise.

It is often convenient to view time series data as consisting of an underlying signal plus some random noise. The signal part of the model is usually modelled as deterministic, rather than random. It is simply some function of t. For example, suppose, for each t,

$x_t = 2 \cos\{2\pi (t+15)/15\} + w_t$.

The first term is viewed as the signal, and w_t is the noise. An example is shown in Fig 1.11.

Any cosine wave can always be written in the form

$A \cos(2\pi \omega t + \phi)$,

where A is called the *amplitude*, ω is called the *frequency* indicating how many cycles are occurring per time unit, and ϕ is called the *phase shift*.

Here $A = 2$, $\omega = 1/50$ because there is one cycle every 50 time points, and $\phi = (2\pi)\ 15/50 = 0.6\pi$.

In the bottom two panels of Fig 1.11, we can see what the cosine looks like after white noise has been added to it. Here there are two different WN processes in the middle and bottom panel, with two different variances, or sizes, σ_W , of the white noise. In both cases the WN is Gaussian. The amplitude of the signal and the variance of the WN determine how clearly the signal appears. The ratio of the amplitude of the signal to σ_W is called the *signal-to-noise ratio (SNR)*. Larger SNR means the signal is easier to discern and lower SNR means it is more obscured.

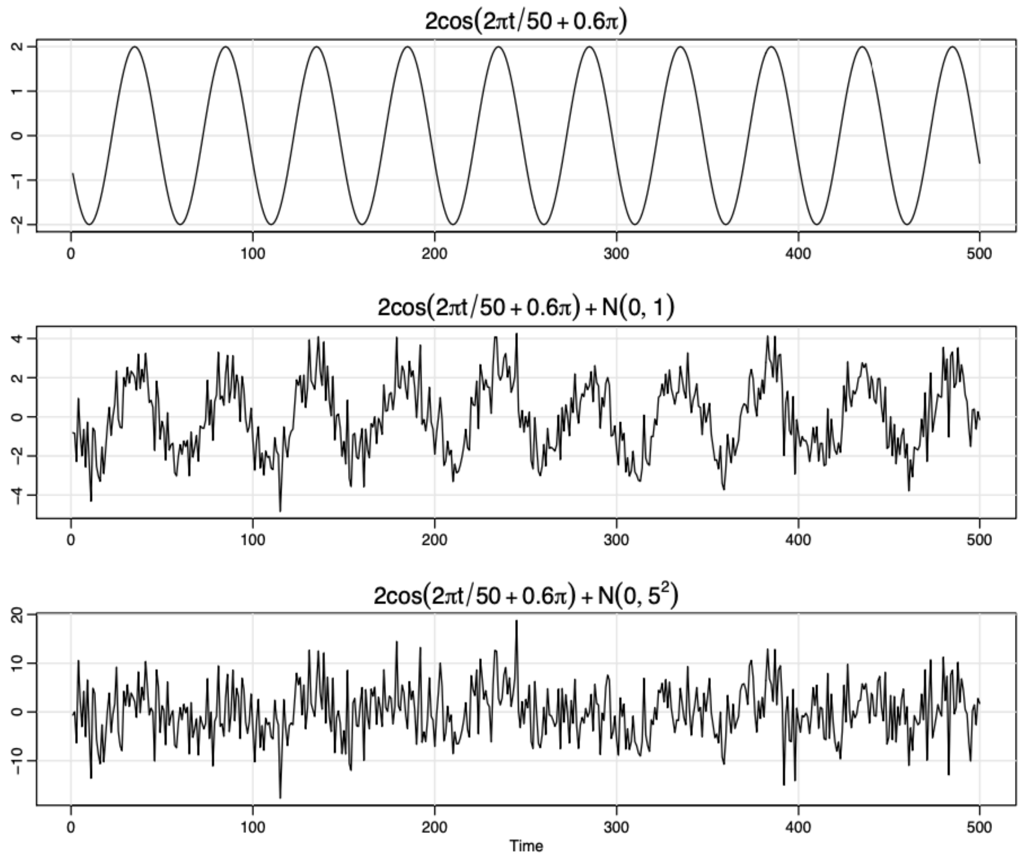

Fig. 1.11. *Cosine wave with period 50 points (top panel) compared with the cosine wave contaminated with additive white Gaussian noise, $\sigma_w = 1$ (middle panel) and $\sigma_w = 5$ (bottom panel); see (1.5).*

```
## yt = 2yt-1 + wt
## so 2yt-1 = yt - wt
## and yt-1 = .5yt - .5wt.

set.seed(184)
alpha = .5
sigma = .5
par(mfrow=c(1,2))
n = 100
w = rnorm(n,mean=0,sd=sigma)
x = rep(0,n)
x[1] = w[1]
for(i in 2:n) x[i] = alpha*x[i-1] + w[i]
```

```
plot(1:n,x,xlab="t",ylab=expression(x[t]),type="l")
y = rep(0,n)
y[n] = w[1]
for(i in 2:n) y[n+1-i] = alpha*y[n+2-i] - alpha * w[n+2-i]
plot(1:n,y,xlab="t",ylab=expression(y[t]),type="l")

## Verify that yt = 2yt-1 + wt.
y[1:3]
w[1:3]
2*y[1]+w[2]
2*y[2]+w[3]

library(astsa)
par(mfrow=c(3,2))
n = 500
t = (1:n)/n
x = 10*cos(t*2*pi*100) + rnorm(n)*12 ## 100 cycles of amplitude 10, in 500 pts,
so freq = 1/5.
y = 0*cos(t*2*pi*100) + rnorm(n)*12 ## 100 cycles of amplitude 0, in 500 pts, so
freq = 1/5.
plot(t,x,type="l",main="x")
plot(t,y,type="l",main="y")
spectrum(x,ylim=c(1,10000),main="periodogram of x")
spectrum(y,ylim=c(1,10000),main="periodogram of y")
spec.ar(x,ylim=c(1,10000),main="spectral density estimate of x")
spec.ar(y,ylim=c(1,10000),main="spectral density estimate of y")
## For the periodogram, see the bottom of p172.
```

3. Mean function.

If x_t is a time series, then $E(x_t)$ is called the *mean function*. It is also sometimes written $\mu(t)$ or μ_t , or sometimes just μ in the case where it is constant for all t.

If x_t is the MA process $x_t = w_{t-1} + w_t + w_{t+1}$,

then $E(X_t) = 1/3\ [E(w_{t-1})+E(w_t)+E(w_{t+1})]=0$.

For a RW with drift, $E(X_t) = \delta t$.

4. Autocovariance.

The *autocovariance function* is defined for any s and t as

$\gamma(s,t) = \text{cov}(x_s, x_t) = E[(x_s - \mu_s)(x_t - \mu_t)]$.

Note that $\gamma(s,t) = \gamma(t,s)$. Also, note that when s = t,

$\gamma(s,t) = \text{cov}(x_t, x_t) = \text{var}(x_t)$.

Often when the time series is smooth, the autocovariance is large and decays very very slowly to zero, i.e. it is large even for s and t rather far apart in time. When the time series is rough, the autocovariance tends to be close to 0 for t-s exceeding some threshold. This difference t-s is called the *lag* or *time lag*.

If x_t = WN, then what is $\gamma(s,t)$?

Here $\gamma(s,t) = 0$ whenever s ≠ t,

but when s=t, $\gamma(s,t) = \text{var}(x_t) = \sigma_W^2$.

So this is the autocovariance of WN. Now we will look at the autocovariance of a MA process.

See p17 of ch1.