Stochastic processes III

The ASTA team

Contents

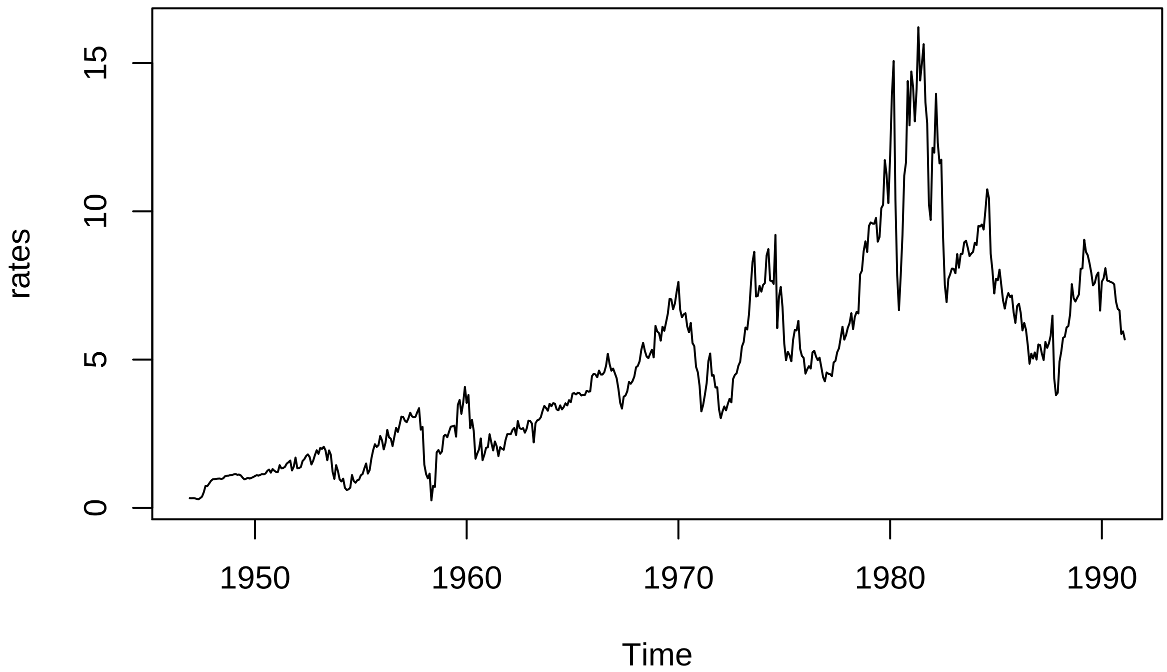
1 Continuous time stochastic processes

1.1 Data example

In this lecture we will study a type of data that on the surface looks like data from time series analysis as presented in the previous lectures. As an illustration, we shall study the data set `Irates`, that contains information about interest rates in the U.S. between 1946 and 1991.

We are interested in the relation between the two variables:

- `t=time`: The time point of the measurement.
- `X_t=rate`: The interest rate at the corresponding time point.



The plot should be understood as follows: For each time point (between 1946 and 1991) there is a value of the interest rate called X_t. So X_t could be seen as a function of t, and this function is plotted.

In principle we imagine that there are infinitely many data points, simply because there are infinitely many time points between 1946 and 1991. Of course this is never true: In practice we will always only have finitely many data points.

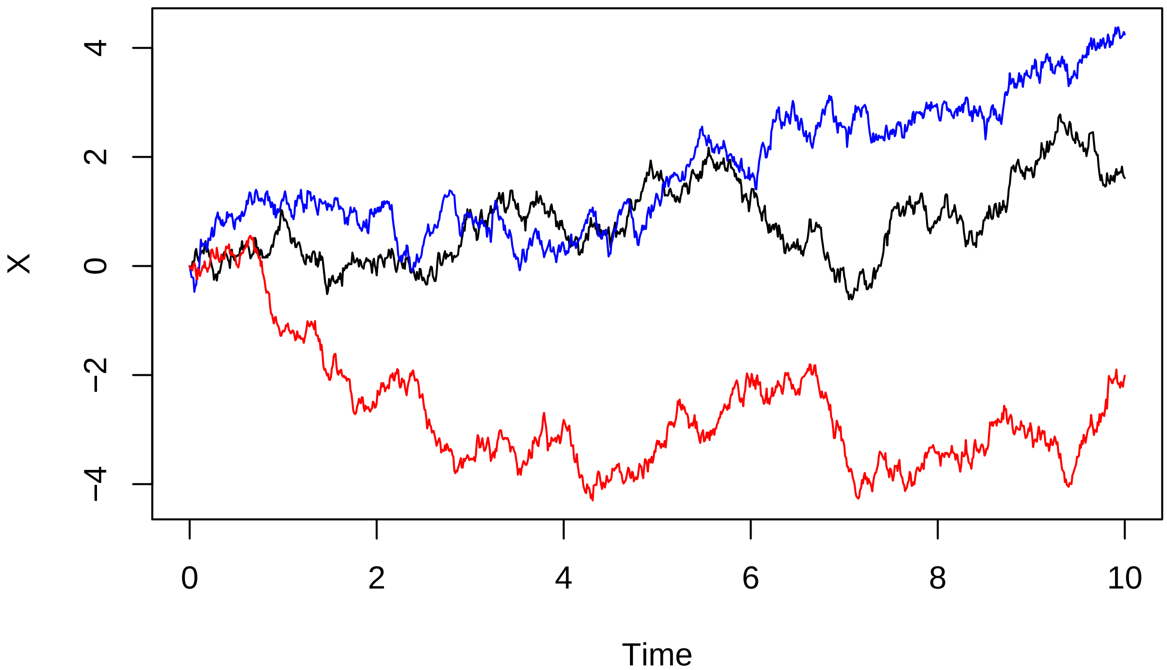
But it makes sense to believe that the real data actually contains all the data points. We are just not able to measure them (and to store them in a computer).

With a model for all datapoints, we are - through simulation - able to describe the behaviour of data. Also between the observations.

2 Wiener process

A key example of a process in continuous time will be the so–called Wiener process.

Three simulated realizations (black, blue and red) of this process can be seen here



A Wiener process has the following properties:

- It starts in 0: $W_0 = 0$.
- It has independent increments: For $0 < s < t$ it holds that $W_t - W_s$ is independent of everything that has happened up to time s, that is W_u for all $u \leq s$.
- It has normally distributed increments: For $0 < s < t$ it holds that the increment $W_t - W_s$ is normally distributed with variance $t - s$:

$$W_t - W_s \sim N(\mu = 0, \sigma^2 = t - s).$$

The intuition of this process is that it somehow changes direction all the time: How the process changes after time s will be independent of what has happened before time s. So whether the process should increase or decrease after s will not be affected by how much it was increasing or decreasing before.

This gives the very bumpy behaviour over time.

3 Differential equation models

3.1 Ordinary differential equations

A common way to define a continuous time stochastic process model is through a stochastic differential equation (SDE) which we will turn to shortly, but before doing so we will recall some basic things about ordinary differential equations.

Suppose f is a differentiable function. Recall the mathematical description of a differential equation

$$\frac{df(t)}{dt} = -4f(t)$$

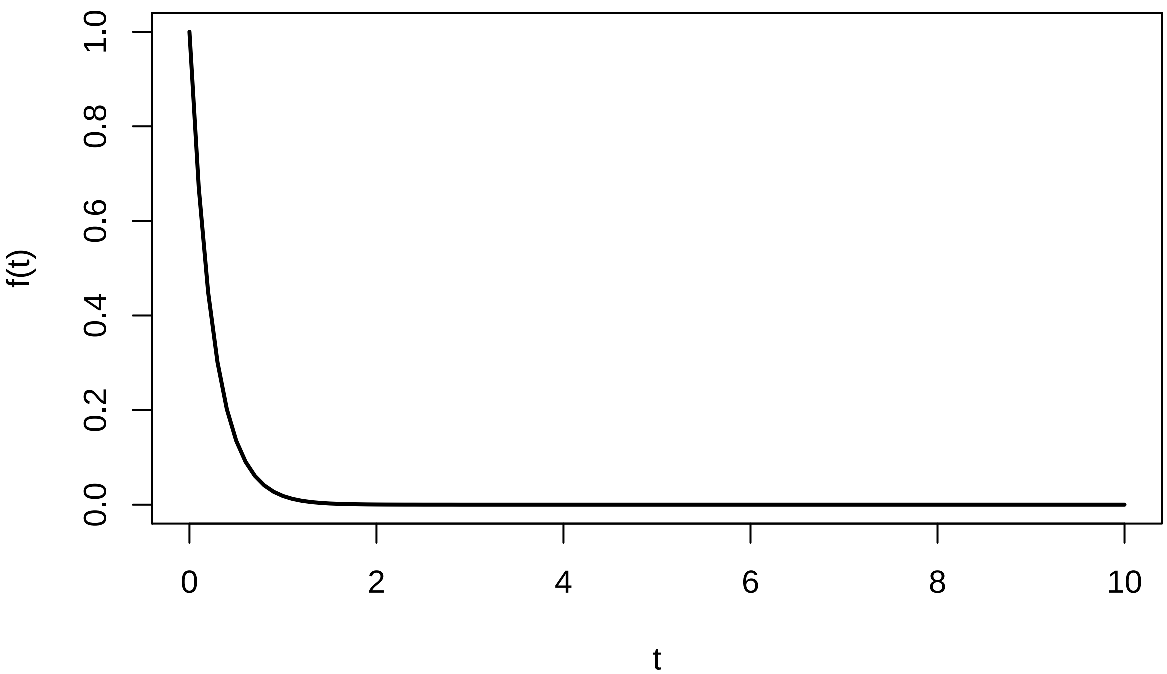
This equation has the solution

$$f(t) = C \cdot \exp(-4t)$$

for any constant C. If we furthermore know that $f(0) = 1$, then $C = 1$ and

$$f(t) = \exp(-4t)$$

The solution can be seen below



With a slightly unusual notation we can rewrite this as

$$df(t) = -4 \cdot f(t)dt$$

This equation has the following (hopefully intuitive) interpretation:

- We imagine that we increase the time point from t to $t + dt$, where dt is something small. So the time is increased by dt.
- Then the value of f is (approximately) changed from $f(t)$ to $f(t) - 4f(t)dt$. So actually the value of f is decreased by $4f(t)\,dt$

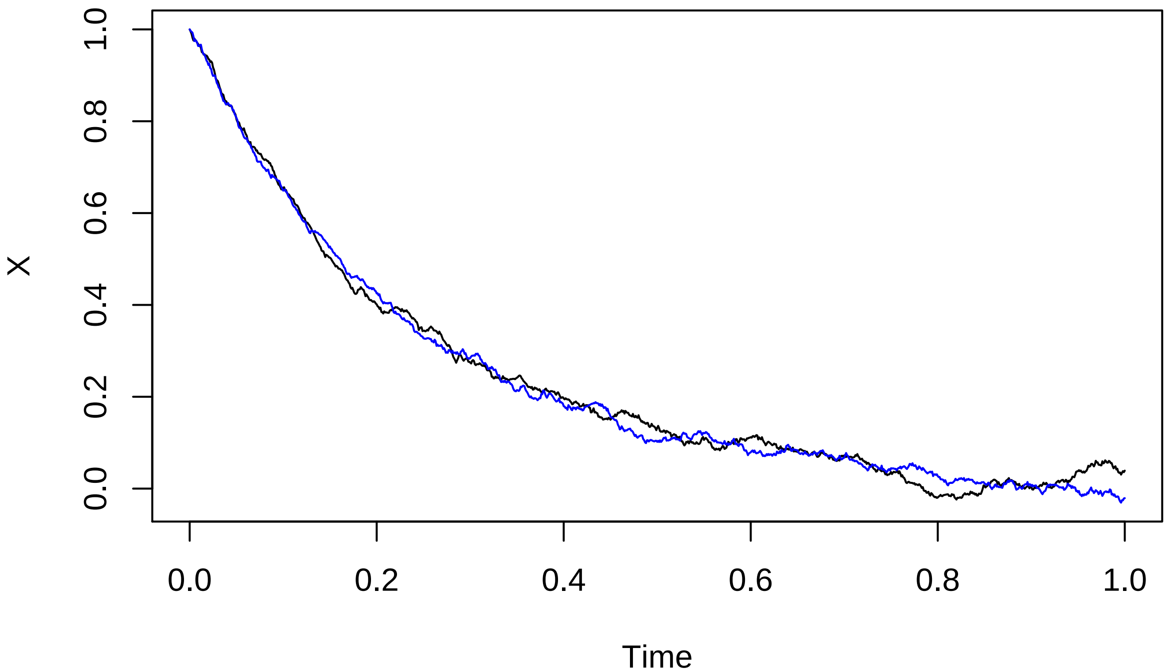
So when t is increased, then $f(t)$ is decreased. And the decrease is determined by the value of $f(t)$. That is why f decreases slower and slower, when t is increased.

We say that the function has a **drift** towards zero, and this drift is determined by the value of the function.

3.2 Stochastic differential equations

It will probably never be true that data behaves exactly like the exponentially decreasing curve on the previous slide.

Instead we will consider a model, where some random noise from a Wiener process has been added. Two different (black/blue) simulated realizations can be seen below



The type of process that is simulated above is most often described formally by the equation

$$dX_t = -4X_t dt + 0.1 dW_t$$

This is called a **Stochastic Differential Equation** (SDE), and the processes simulated above are called solutions of the stochastic differential equation.

The SDE $dX_t = -4X_t dt + 0.1dW_t$ has two terms:

- $-4X_t dt$ is the **drift term**.
- $0.1dW_t$ is the **diffusion term**.

The intuition behind this notation is very similar to the intuition in the equation $df(t) = -4 \cdot f(t)\, dt$ for an ordinary differential equation. When the time is increased by the small amount dt, then the process X_t is increased by $-4X_t\, dt$ AND by how much the process $0.1W_t$ has increased on the time interval $[t, t + dt]$.

So this process has a **drift** towards zero, but it is also pushed in a random direction (either up or down) by the Wiener process (more precisely, the process $0.1W_t$)

3.2.1 Simulation examples

Firstly we simulate the SDE from before

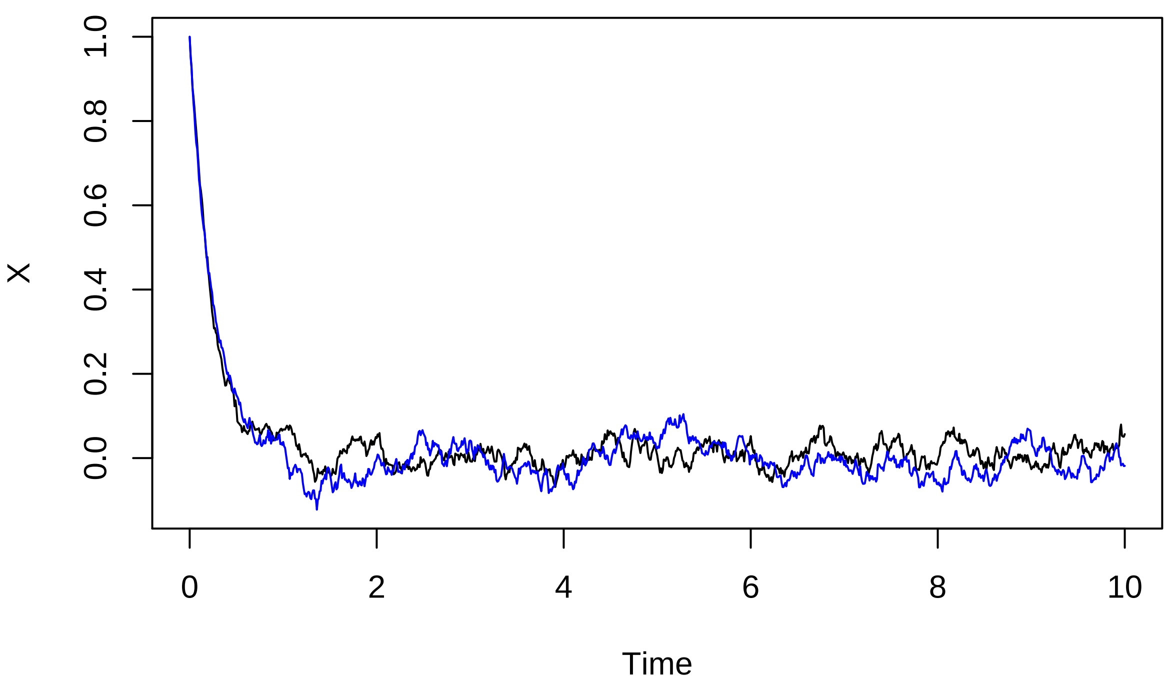
$$dX_t = -4X_t dt + 0.1dW_t$$

For this, we need the package `Sim.DiffProc`. We use the function `snssde1` for which we have to specify the drift term and the diffusion term as R-functions of x. In this case, the function for the diffusion term is constantly equal to 0.1.

The parameters needed in the function input are:

- `drift` is the function determining the drift term.
- `diffusion` is the function determining the diffusion term.
- `M` is the desired number of realizations of the process.
- `N` is the number of simulation steps (`R` does not simulate a continuous curve but a lot of connected dots, and this is the number of dots).
- `t0` is the initial time of the simulated process.
- `T` is the ending time of the simulated process.
- `x0` is the initial value of the process (the value of X at time `t0`).

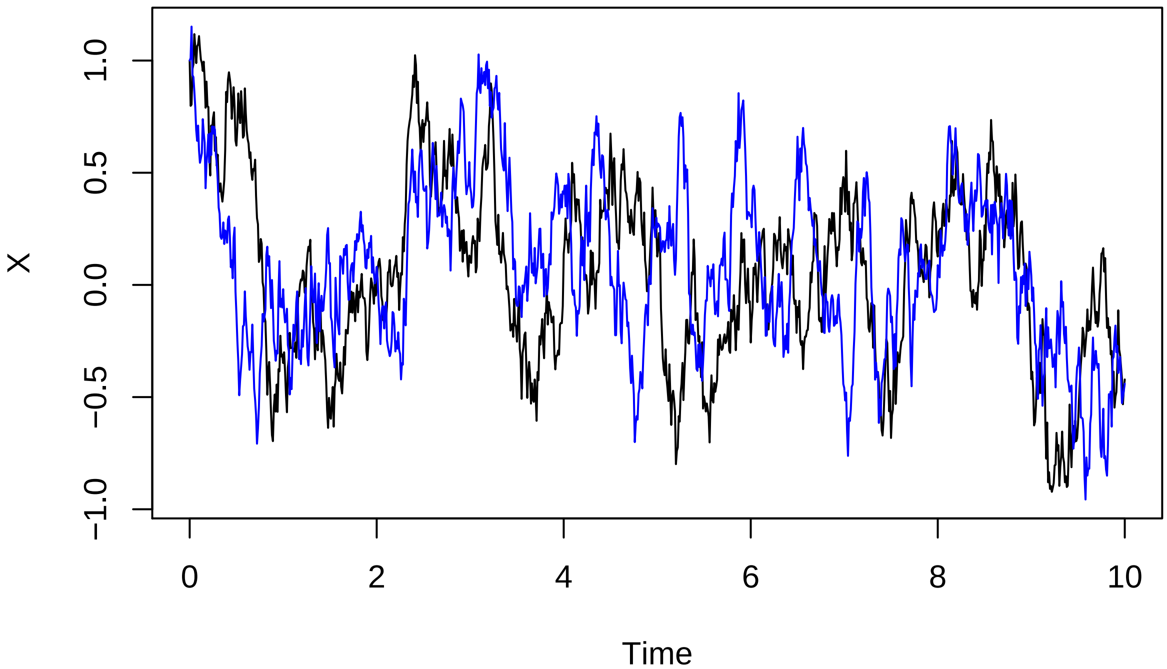
```
library(Sim.DiffProc)
f <- expression(-4*x)     ## the drift term as a function of x
g <- expression(0.1)      ## the diffusion term as a function of x
res <- snssde1d(drift=f, diffusion=g, M=2, N=1000, t0=0, T=10, x0=1)
plot(res, plot.type = 'single', col = c('black','blue'))
```



With the ending time being larger than before, we see that the process stabilizes around 0: There is a drift towards 0, but also some noise pushing the process away from 0.

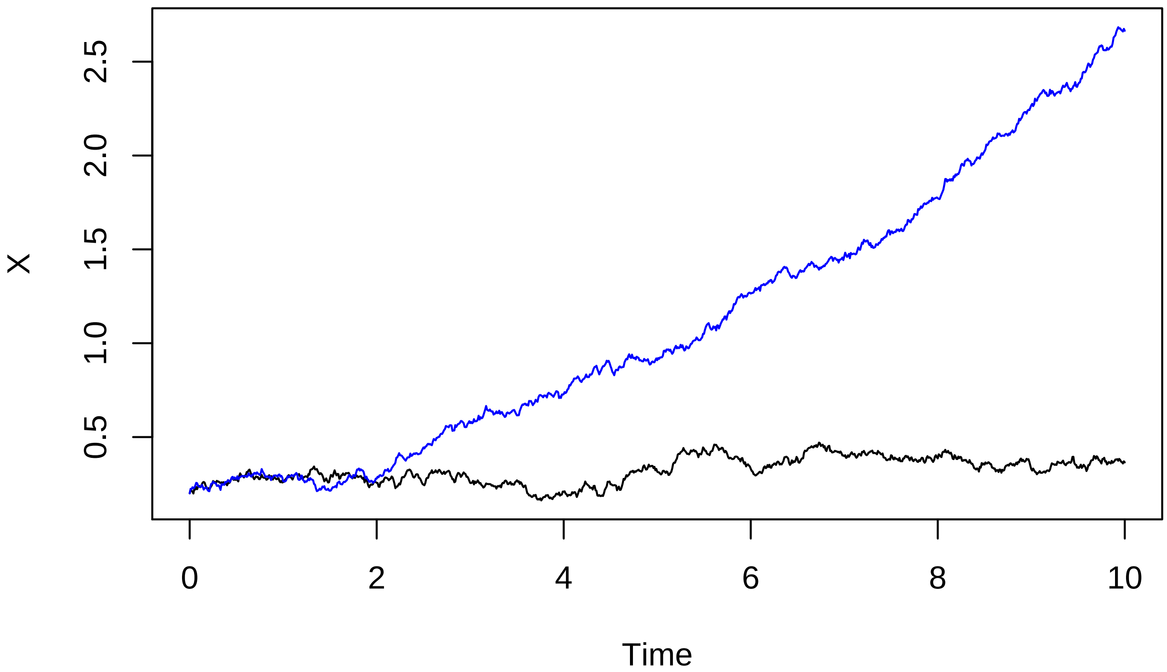
Increasing the diffusion parameter from 0.1 to 1 (i.e. $dX_t = -4X_t dt + 1dW_t$) makes the process more varying:

```
f <- expression(-4*x)
g <-expression(1)
res <- snssde1d(drift=f, diffusion=g, M=2, N=1000, t0=0, T=10, x0=1)
plot(res, plot.type='single', col=c('black','blue'))
```



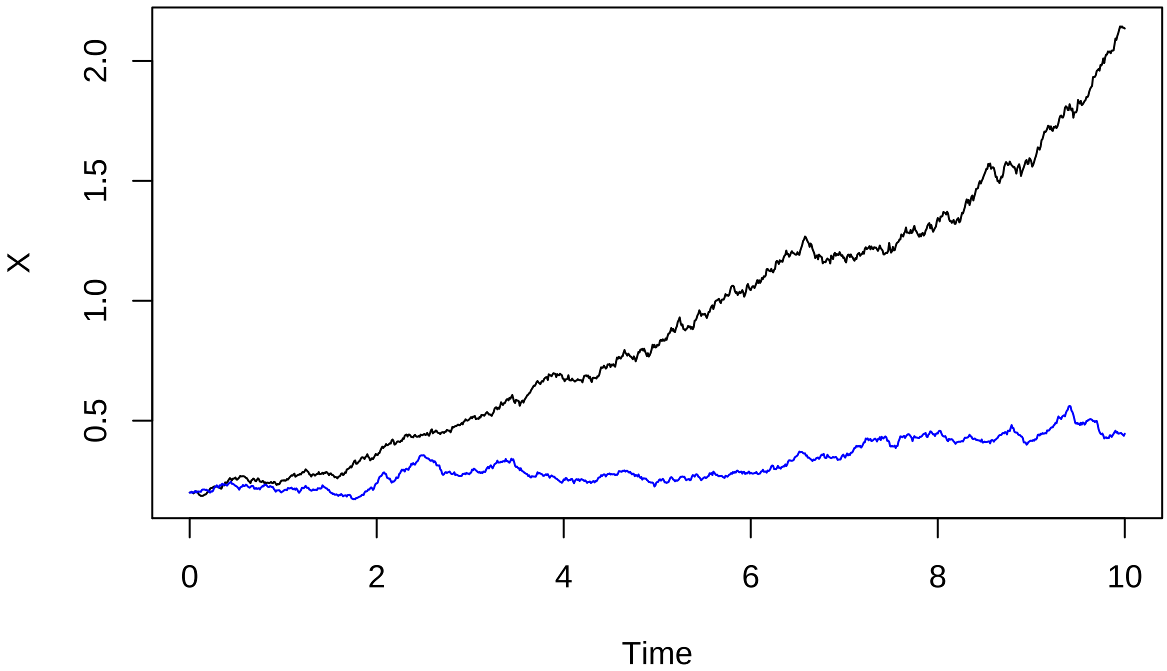
Making the drift parameter positive (e.g. $dX_t = 0.2X_t dt + 0.1dW_t$) drives the process away from 0:

```
f <- expression(0.2*x)
g <-expression(0.1)
res <- snssde1d(drift=f, diffusion=g, M=2, N=1000, t0=0, T=10, x0=0.2)
plot(res,plot.type='single',col=c('black','blue'))
```



Sometimes the noise depends on the value of the process itself. In this case the diffusion term includes X_t. E.g. $dX_t = 0.2X_t dt + 0.1\sqrt{X_t} dW_t$:

```
f <- expression(0.2*x)
g <- expression(0.1*sqrt(x))
res <- snssde1d(drift=f,diffusion=g,M=2,N=1000,t0=0,T=10,x0=0.2)
plot(res,plot.type='single',col=c('black','blue'))
```



3.3 Fitting SDE models to data

When we want to fit a model to data, we will work with the following more general stochastic differential equation

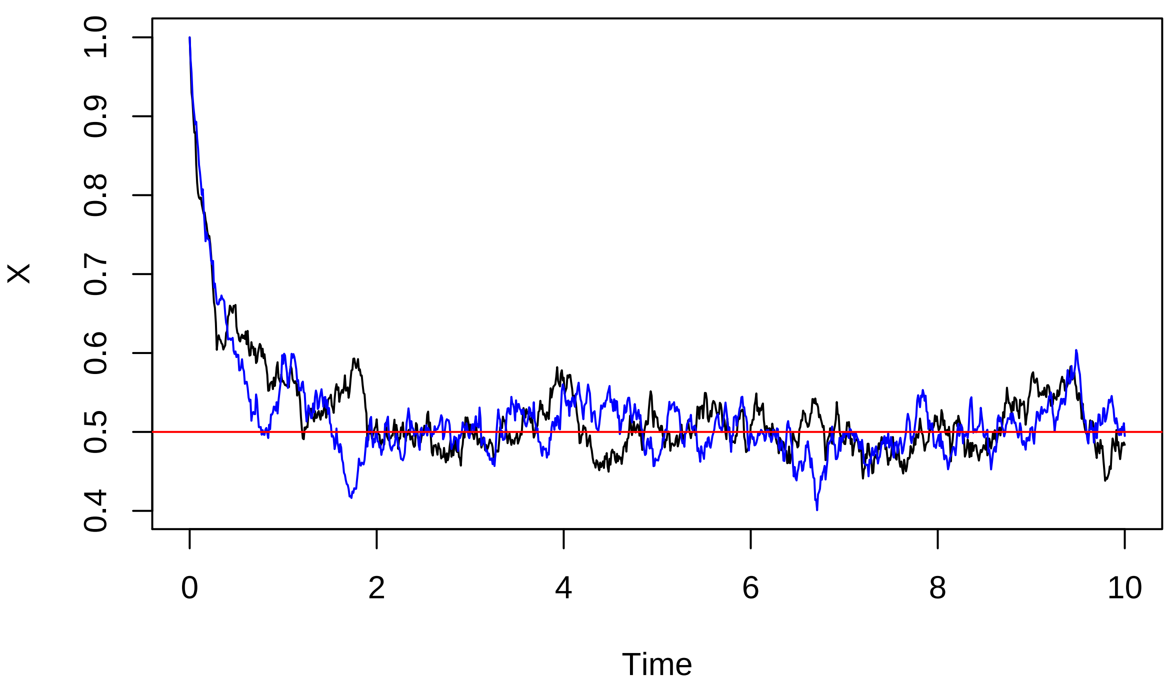
$$dX_t = (\theta_1 + \theta_2 X_t)dt + \theta_3 X_t^{\theta_4} dW_t$$

where $\theta_1, \theta_2, \theta_3, \theta_4$ are parameters, such that $\theta_3, \theta_4 \geq 0$. For a given dataset the goal is then to find (estimate) parameter values such that the model describes the data as well as possible.

We note that:

- The SDE above given by $dX_t = -4X_t dt + 0.1 dW_t$ is the special case with $\theta_1 = 0$, $\theta_2 = -4$, $\theta_3 = 0.1$ and $\theta_4 = 0$ (recall the mathematical convention that $x^0 = 1$).
- The Wiener process $X_t = W_t$ is the special case with $\theta_1 = 0$, $\theta_2 = 0$, $\theta_3 = 1$ and $\theta_4 = 0$.
- The ordinary differential equation is the special case with $\theta_1 = 0$, $\theta_2 = -4$, $\theta_3 = 0$ and $\theta_4 = 0$ (in principle, θ_4 could be anything, when $\theta_3 = 0$).
- The parameters θ_1 and θ_2 control the drift, and in fact it can be shown that the process will drift towards $-\frac{\theta_1}{\theta_2}$ if this is positive, and otherwise it will drift away from this. For example if $dX_t = (2 - 4X_t)dt + 0.1 dW_t$ then X_t will drift towards $-\frac{2}{-4} = 0.5$:

```
f <- expression(2-4*x)
g <- expression(0.1)
res <- snssde1d(drift=f, diffusion=g, M=2, N=1000, t0=0, T=10, x0=1)
plot(res, plot.type='single', col=c('black','blue'))
abline(h = 0.5, col = "red")
```



3.3.1 Fitting an SDE model to interest rate data

Recall the data for U.S. interest rates between 1946 and 1991.

This could look like a stochastic differential equation with $\theta_1 = 0$ and θ_4 being positive, since the process varies more, when it has high values. We can use the function `fitsde` to find the best choice of parameters. Note that in the function we have to give a (good) guess on the parameters (here we use $\theta_1 = 0$, $\theta_2 = 0$, $\theta_3 = 0.5$ and $\theta_4 = 0.5$)

```
data(Irates)
rate <- Irates[ , "r1"]
fx <- expression(theta[1]+theta[2]*x)
## the drift term as a function of x
gx <- expression(theta[3]*x^theta[4])
## the diffusion term as a function of x
fitmod <- fitsde(rate, drift = fx, diffusion = gx,
                 start = list(theta1=0, theta2=0, theta3=0.5, theta4=0.5))
summary(fitmod)
```

```
## Pseudo maximum likelihood estimation
##
## Method:  Euler
## Call:
## fitsde(data = rate, drift = fx, diffusion = gx, start = list(theta1 = 0,
##     theta2 = 0, theta3 = 0.5, theta4 = 0.5))
##
## Coefficients:
##          Estimate Std. Error
## theta1  0.8862133 0.24588655
## theta2 -0.1591198 0.08044809
## theta3  0.7138713 0.03379098
## theta4  0.5926193 0.02765048
##
## -2 log L: 648.049
```

Thus, the estimated model is

$$dX_t = (0.89 - 0.16X_t)dt + 0.71X_t^{0.59}dW_t.$$

We can have the confidence intervals by

```
confint(fitmod)
```

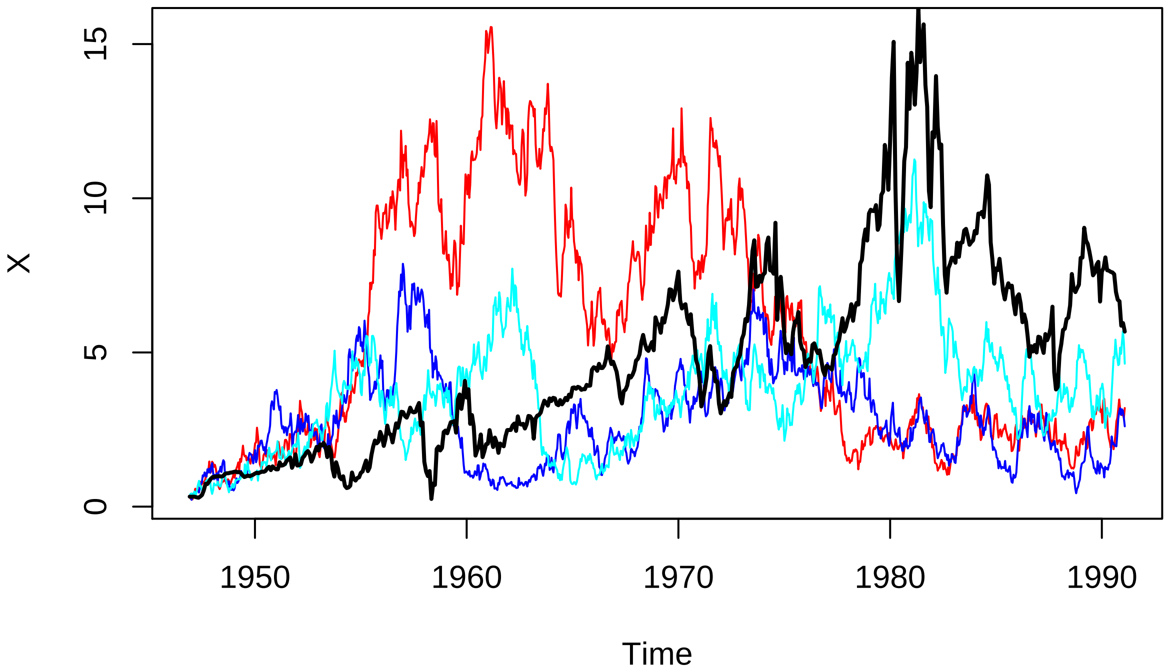
```
##             2.5 %       97.5 %
## theta1  0.4042845  1.368142087
## theta2 -0.3167952 -0.001444485
## theta3  0.6476422  0.780100349
## theta4  0.5384254  0.646813280
```

To draw random realizations from the fitted model we just have to extract the fitted parameters and then use `snssde1` as before:

```r
theta <- coef(fitmod)
t <- time(Irates)
s <- snssde1d(drift=fx, diffusion = gx, M = 3, t0 = min(t), T = max(t), x0 = 0.325)
```

A plot of three realizations overlayed the original data:

```r
plot(s, plot.type = "single", col = c("red", "blue", "cyan"))
lines(Irates[,'r1'], col = "black", lwd = 2)
```



3.4 Comparing fitted SDE models

If we believe that the data can be better (or equally well) described by another model we can compare the model using the AIC as we did previously for discrete processes.

If we propose a model with no drift (i.e. $\theta_1 = 0$ and $\theta_2 = 0$) we get the following fitted model:

```r
f2 <- expression(0)
## No drift term
g2 <- expression(theta[1]*x^theta[2])
## the diffusion term as a function of x
fitmod2 <- fitsde(rate, drift = f2, diffusion = g2,
                  start = list(theta1=0.5, theta2=0.5))
summary(fitmod2)
```

```
## Pseudo maximum likelihood estimation
##
## Method:  Euler
## Call:
```

```
## fitsde(data = rate, drift = f2, diffusion = g2, start = list(theta1 = 0.5,
##     theta2 = 0.5))
##
## Coefficients:
##         Estimate Std. Error
## theta1 0.7400495 0.03454472
## theta2 0.5743530 0.02698523
##
## -2 log L: 661.0032
```

The AIC of the original model with four parameters is lower than the AIC of this new model so we prefer the original model:

```
AIC(fitmod)
```

```
## [1] 656.049
```

```
AIC(fitmod2)
```

```
## [1] 665.0032
```