Visualization of economic data

Bernt Arne Ødegaard

24 November 2021

1 Introduction

We now look at plotting of data. Remember the old adage that "A picture says more than a thousand words."

In economics, when we want to explore relationships between variables, coming up with the right picture to illustrate a point may be infinitely superior to a "kitchen sink" of regressions.

As part of applied econometrics, knowledge of how to generate pictures is therefore important. The "eyeball" metric is a well-know tool in data analysis.

I thus find it be useful to spend one lecture hour on some examples of the use of a graphical illustrations.

The "right picture" may actually be a do or die in terms of publishing the paper. The "killer picture" (like the killer app) may be what catches the reader (referee)'s eye.

2 Plotting the term structure of interest rates

Let us now look at an economic example that illustrates some of the issues in the display of information, and can lead to nontrivial choices in terms of pictures.

Consider the *term structure of interest rates.*

2.1 What is the "term structure"?

What is an interest rate? From basic economics, remember the interest rate as the "value of time," that is, we can interpret the interest rate as the price we are willing to be paid to get money *later* rather than *now*. This interpretation allows us to define the basic rules of investment analysis, that we look at the discounted value of future cashflows as the value *now*. For example, to calculate the value of an investment project, we calculate its *Net Present Value*

$$NPV = \sum_{i=1}^{T} \frac{C_t}{(1+r)^t} - C_0$$

If we had an investment with the cashflows:

$$\begin{array}{rccc} t = & 0 & 1 & 2 \\ C_t = & -100 & 10 & 100 \end{array}$$

we would calculate its value *now* as

$$NPV = -100 + \frac{10}{(1+r)} + \frac{110}{(1+r)^2}$$

If the interest rate were 5%, the value today of the investment is

$$NPV = 9.29705$$

Note that this calculation uses the *same* interest rate r to discount the cash flows in periods 1 and 2. In other words, the "price" per period for waiting two periods is the same as the "price" for waiting one period.

However, there is no reason to believe that this is the case. In general, the "price" of waiting will depend on the length of time you are waiting for. Hence, the interest rate r is a function of the time t that it is discounting, and we term this r_t, the *spot* rate for borrowing for t periods.[1] (or the spot rate at *maturity* t.

$$NPV = \sum_{i=1}^{T} \frac{C_t}{(1+r_t)^t} - C_0$$

This relation between the time t and the interest rate r_t is what we call the "term structure of interest rates."

In the investment example above, assume the term structure is

$$\begin{array}{rcc} t = & 1 & 2 \\ r_t = & 4\% & 8\% \end{array}$$

The present value would be calculated as

$$NPV = -100 + \frac{10}{(1+r_1)} + \frac{110}{(1+r_2)^2}$$

$$NPV = -100 + \frac{10}{(1+0.04)} + \frac{110}{(1+0.08)^2}$$

$$NPV = 3.92$$

Thus, in presenting the term structure of interest rates, the relationship we want to convey information about is the relation between the "maturity" and interest rate.

In practice, we find information about the term structure by looking at prices for *bonds* with various maturities. For US data, we estimate the term structure from Treasury Bills and Bonds, at various maturities. The Wall Street journal for 21 jan 94 gives the following information summarized in table 1:

Table 1 Interest rates implied in US treasury prices on 12 jan 1994

Maturity	yield 21 jan 94
1 week	2.875
1 month	3
3 months	3.1875
6 months	3.3125
1 year	3.6875
5 years	5.04
10 years	5.69

The information in this table summarizes the term structure on that particular date, but it is not in a "user-friendly" format.

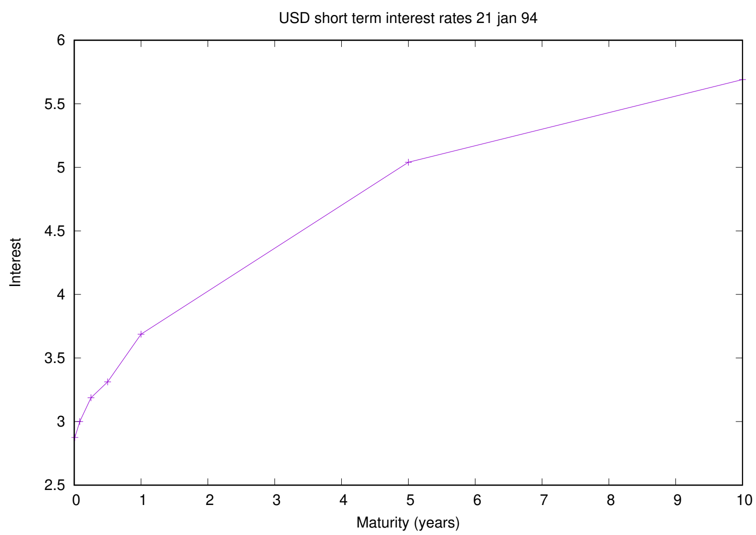
One way to illustrate the term structure is to plot the interest rate against maturity. This gives the following picture, shown in figure 1, which summarizes information about the term structure much better than the above table.

Seeing the data presented this way gives meaning to the concept of "shape" of the term structure. The term structure is "rising," that is, the short term interest rate is lower than the longer term interest rate.

[1] Alternatively, we talk about the "discount factor" $d(t) = \frac{1}{(1+r_t)^t}$ for a t period investment, and we would calculate the NPV as

$$NPV = \sum_{i=1}^{T} d(t) \cdot C_t - C_0$$

Figure 1 Term structure 21 jan 1994



2.2 Evolution over time.

Let us now go a bit further. The one thing we all know about interest rates is that they tend to change as time goes by. Let us think about how we can convey information about changes in the interest rate. We now know about the whole term structure, and realize that as time passes, the "shape" of the term structure may change, not only the level of interest rate.

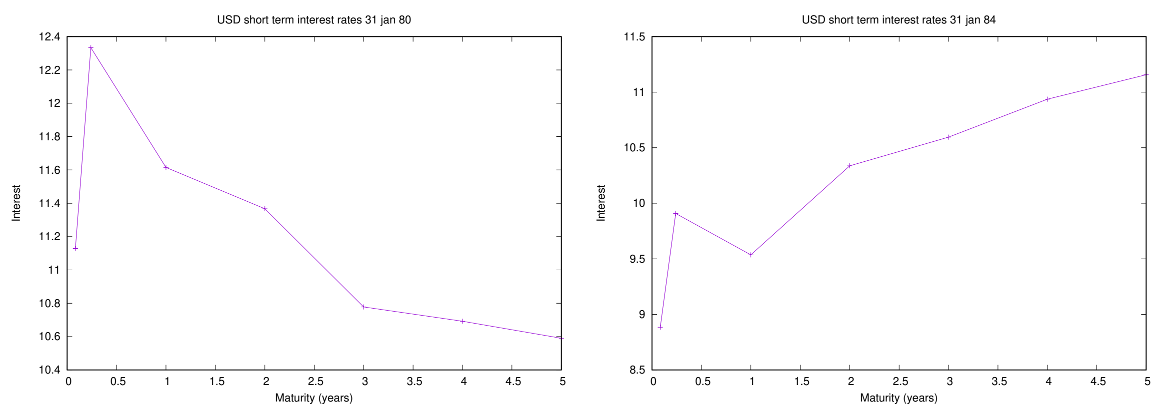
If the shape does not change, and an interest rate change only shifts the term structure up or down, we could convey all information about the evolution of the term structure by a time series plot of e.g. the short term interest rate.

To see if this is the case, let us plot the shape of the term structure for a couple of dates, using historical US data. Table 2 show interest rates fro two dates in 1980 and 1984. The implied term structures are plotted in figure 2.

Table 2 Interest rates implied in US treasury prices in 1980 and 1984

Maturity	yield	
	31 jan 80	31 jan 84
1 month	11.129	8.884
3 months	12.334	9.908
1 year	11.615	9.536
2 years	11.367	10.337
3 years	10.778	10.595
4 years	10.692	10.937
5 years	10.590	11.157

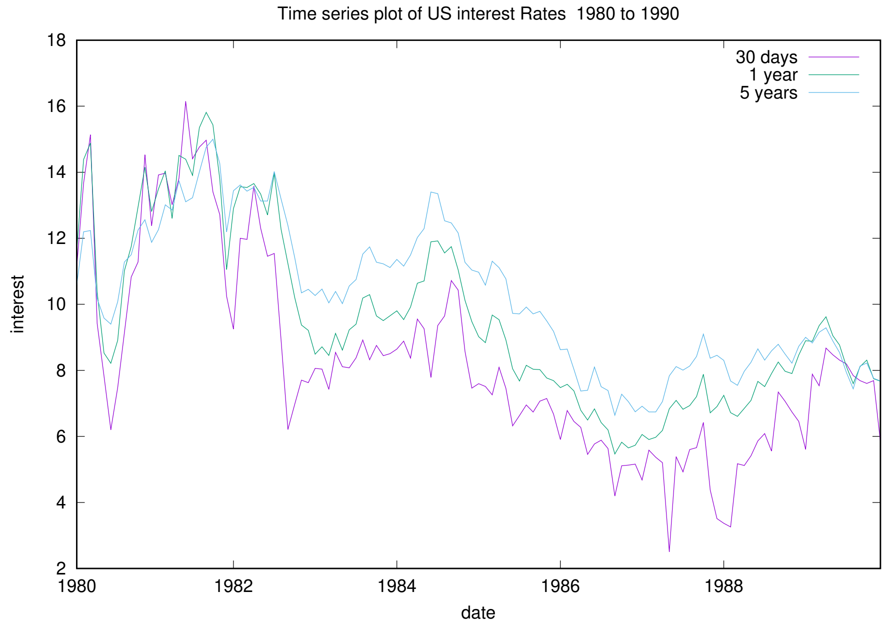
Figure 2 Example term structure in 1980 and 1984.



As we observe in these plots, the shape of the term structure is changing, as well as the level. One of the term structures is rising, the other falling most of the time, except for the very short end. We potentially need to convey this information about the changing "shape" of the term structure as well as the change in level. One solution is to make one plot of the term structure, as shown above, on each date we are interested in. But this may be "information overkill," it is hard to see how the term structure evolves over time.

One possible way of getting information across is to pick selected maturities, and plot the time series evolution of these. Let us plot the time series evolution of interest rates for 30 days, 1 year and 5 years. This plot, shown in figure 3 let us easily see how the level of interest rates has changed, but it is hard to judge how the shape of the term structure has changed by the relative positions of the three different maturities. We do see that the short interest rate is below the longer term interest rates most of the time (upward sloping term structure), but that this was reversed in the early part of the period.

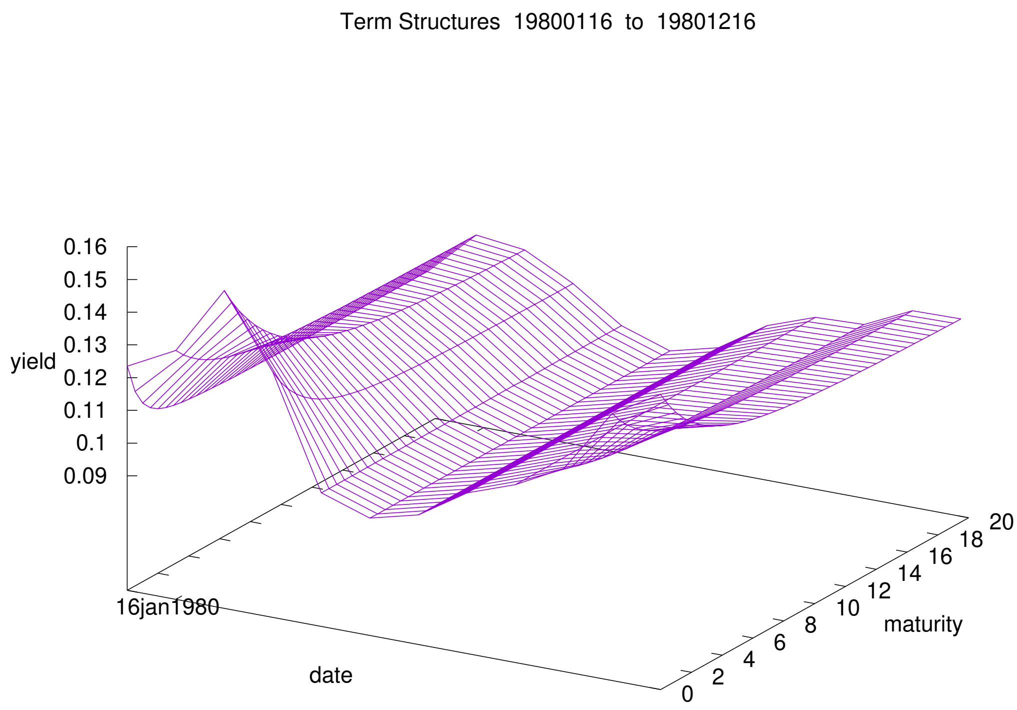
Figure 3 The time series evolution of selected US interest rates



The plots shows the time series evolution of 30 days, 1 year and 5 years interest rates.

As a final try at conveying the information about the changing shape of the term structure, consider the next plot, shown in figure 4

Figure 4 The time series evolution of US term structures (3d)



By going to three dimensions, we could show much of the information available. We see how the shape is changing with passing time, as well as changes in the level of the term structure. But this picture also illustrates that 3 dimensional pictures tend to be harder to "read" than two dimensional ones. We could possibly try to make this easier to read by rotating the picture etc. That is one of the challenges of displaying information visually.

3 Plotting Norwegian term structures

In the previous example we saw some selected US term structure information. Let us now do similar plots with term structure data from Norway. The interest rate date for Norway is on the course homepage. You are expected to play with that input data and attempt to reproduce (and improve on) the graphs illustrated.

We illustrate graph generation using the two types of programs we will be using throughout the course.

3.1 Using octave/matlab

Looking at the term structure data provided, we see that this is from 1991 and onwards, almost till today, with selected maturities. These interest rates are yields on various government securities, provided by Norges Bank. We will use data for the whole period.

We now will use Octave/Matlab to provide similar plots to the american ones illustrated before.

We start up the matrix handler and read the data in. The data as provided is in a semicolon separated text file, the first few lines are reproduced below

```
date;NIBOR T/N-nom;NIBOR 1U-nom;NIBOR 2U-nom;NIBOR 1M-nom;NIBOR 2M-nom;NIBOR 3M-nom;NIBOR 6M-nom;NIBOR 9M-nom;NIBOR 12M-nom;ST3;ST5;ST10
19911031;10.36;10.37;10.46;10.46;10.39;10.39;10.34;10.32;10.29;9.89;9.79;9.88
19911130;8.58;9.74;9.83;10.18;10.28;10.34;10.37;10.36;10.34;9.89;9.72;9.83
....
```

Such a file is read into a matrix by octave

```
> M = dlmread("term_stru_norway.txt",";",1,0);
```

The first column is the set of dates, and the rest is the interest rate yields. We split the two into a vector of dates, and a matrix of yields. We also specify the maturities.

```
> dates=M(:,1);
> yields=M(:,2:13);
> maturities=[1/250,1/52,2/52,1/12,2/12,3/12,6/12,9/12,1,3,5,10];
```

To for example plot the first date we simply give the command

```
> plot(maturities,yields(1,:))
```

The resulting term structure picture is illustrated in figure 5.

We similarly plot for a couple more term structures, one in 2000 and one in 2008. Note how the shapes move from rising to declining term structures.

Figure 5 Term structure 1991

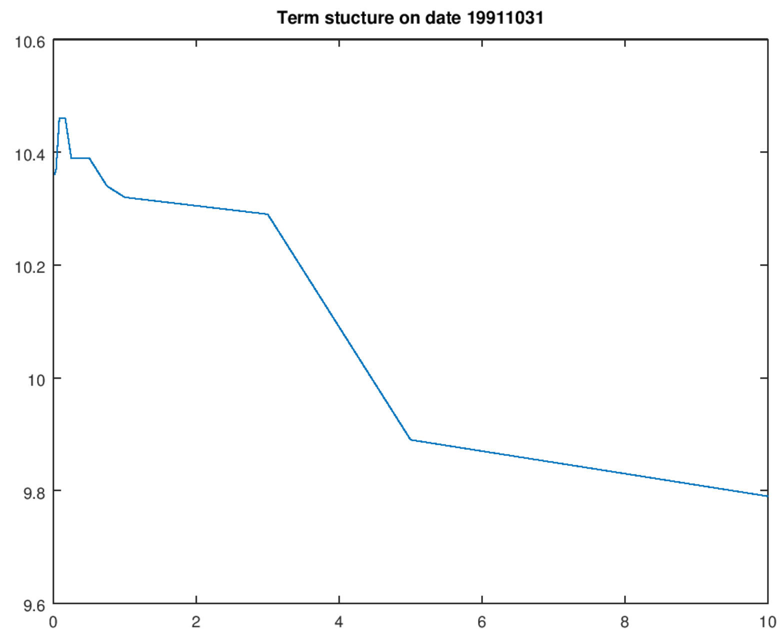
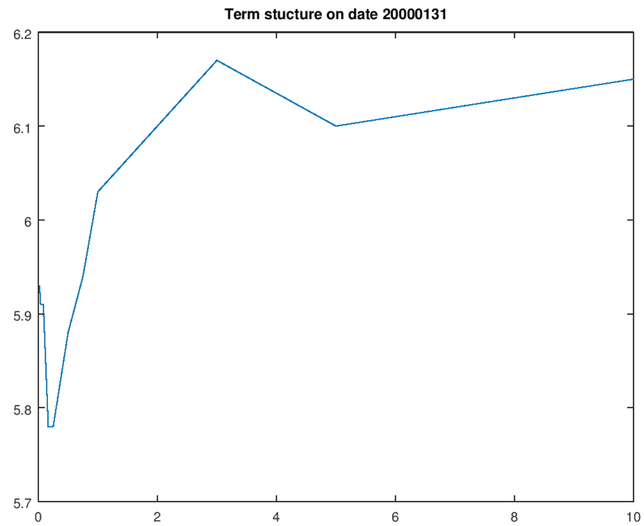
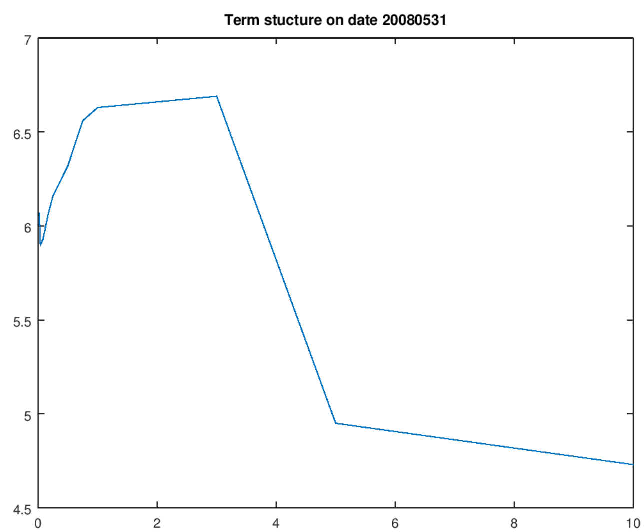


Figure 6 Term structure on example dates

31 jan 2000



31 may 2008

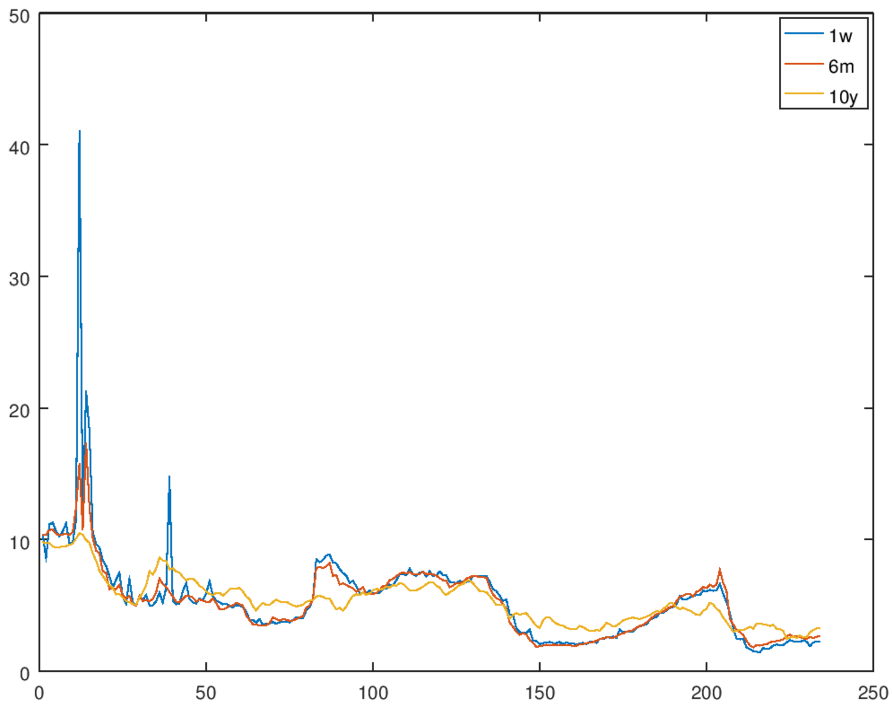


Now, to see the time series behaviour of Norwegian interest rates we plot 3 selected series, the o/n rate, the 6 month rate and the 10 year rate.

The octave command to do this is

```
> t=1:rows(yields);
> plot(t,yields(:,1),";o/n;",t,yields(:,7),";6m;",t,yields(:,12),";10y;");
```

Figure 7 Yield time series



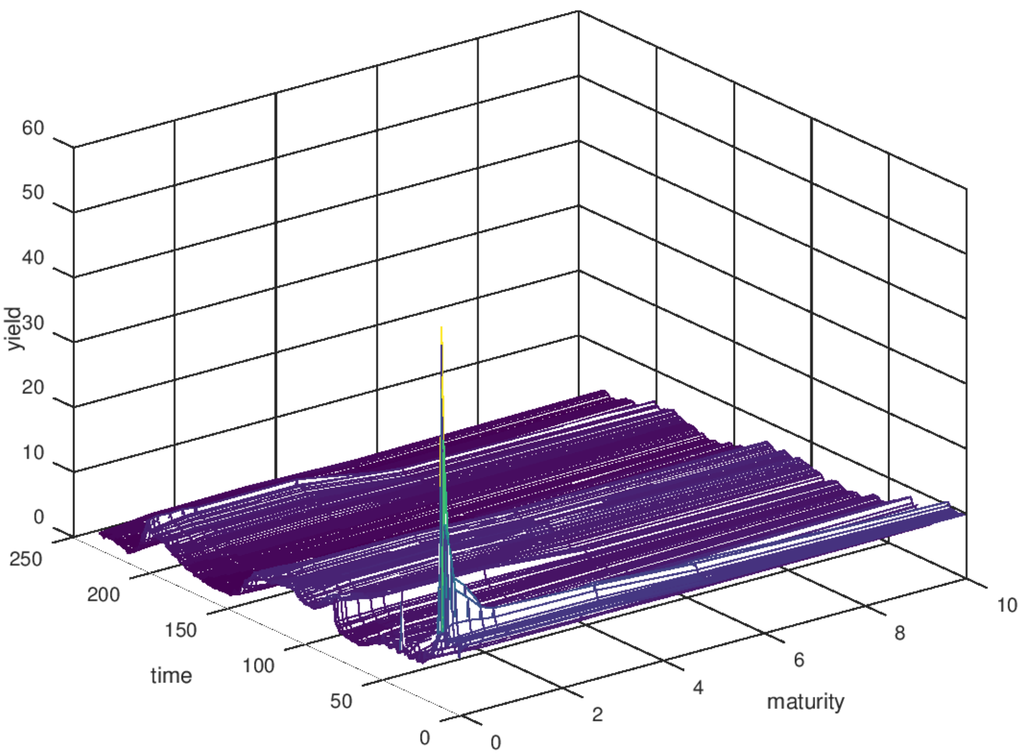
The result is shown in figure 7. Let me point out a few items of interest.

1. The "extreme" observation in the beginning of the period, particularly of the o/n rate, may look like an error, but it is not. This was a currency crisis, with accompanying devaluation.
2. Note that the long and short term rates are moving back and forth across each other. The shape of the term structure is therefore changing.

The final picture is a 3d picture of the evolution of the term structure.

```
> mesh (maturities,t,yields)
> zlabel("yield");
> xlabel("maturity");
> ylabel("time");
> print("3dyields_1.eps",'-depsc');
```

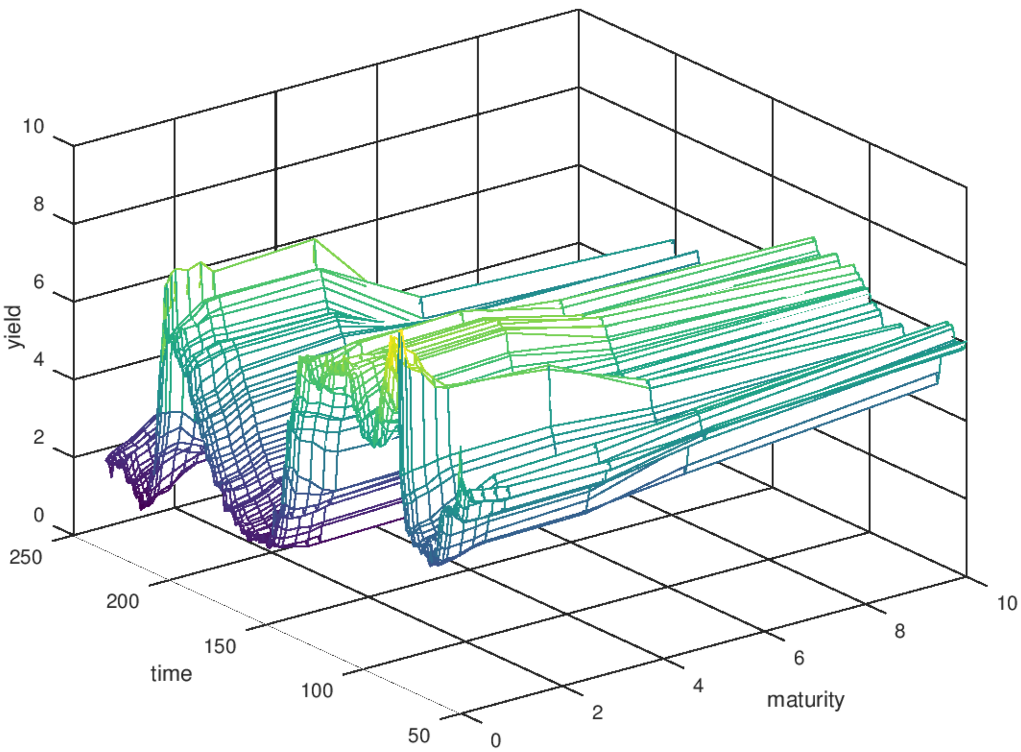
Figure 8 3d plot of time series evolution of term structure



The first picture show the whole period. The crisis ruins the scaling of the other periods, so let us do this without the crisis period, by leaving out the first fifty observations. In the resulting picture it is much easier to see the changing shapes of the term structure in the period.

```
> mesh (maturities,t(50:T),yields(50:T,:))
```

Figure 9 3d plot of time series evolution of term structure, leaving out first 50 dates



3.2 Redoing this using R

The R program is both similar and different from octave. They are similar in the sense that in both cases we interact with the by writing commands, and the data is read into variables which we can access. Octave is however more limited, the data needs to be accessed as vectors or matrices, and manipulations will follow the rules of linear algebra (of which more later), while R is written as a statistics handler from the outset, and allows more flexibility in accessing the data.

When one works with both packages, it is easy to be bitten by the small differences between them, but the rewards to having both available are large. For example, the graphics in R are superior to those offered by Octave/Matlab.

We read the same file into R as follows

```
> termstru <- read.table("term_stru_norway.txt",header=TRUE,sep=";")
```

and do some definitions

```
> nrows<-dim(termstru)[1]
> ncols<-dim(termstru)[2]
> times<-1:nrows
> mat<-c(1/360,1/52,2/52,1/12,2/12,3/12,6/12,9/12,1,3,5,10)
```

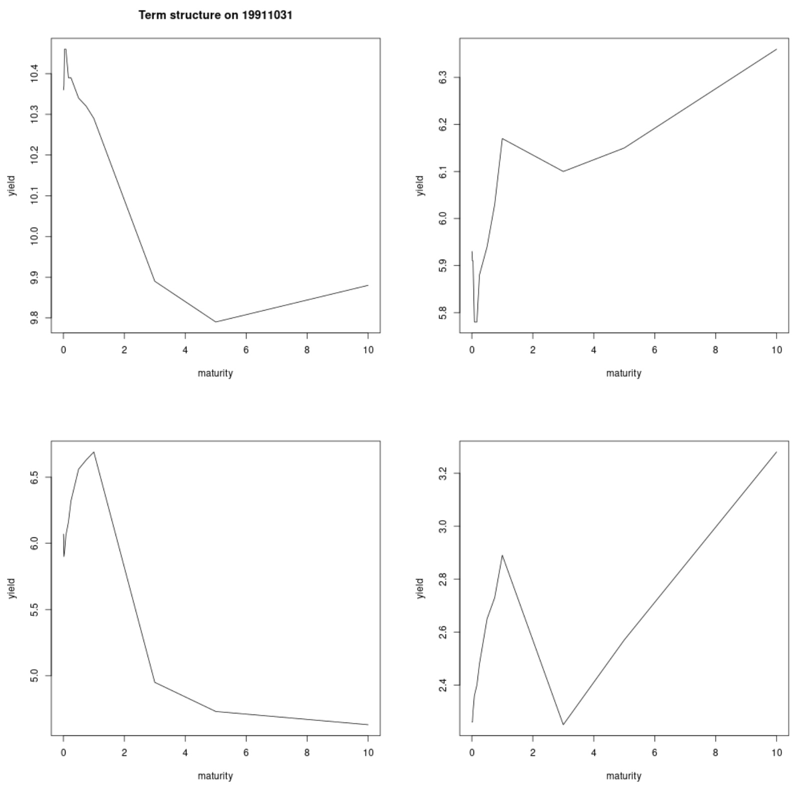
Note that the assignment operator $<-$ is different from the more usual equality sign used in octave.

4 Output from R

Selected term structures.

```
> row<-1
> plot(mat,termstru[row,2:13],xlab="maturity",ylab="yield",type="l",main=titl,)
```

Figure 10 Term structures on selected dates

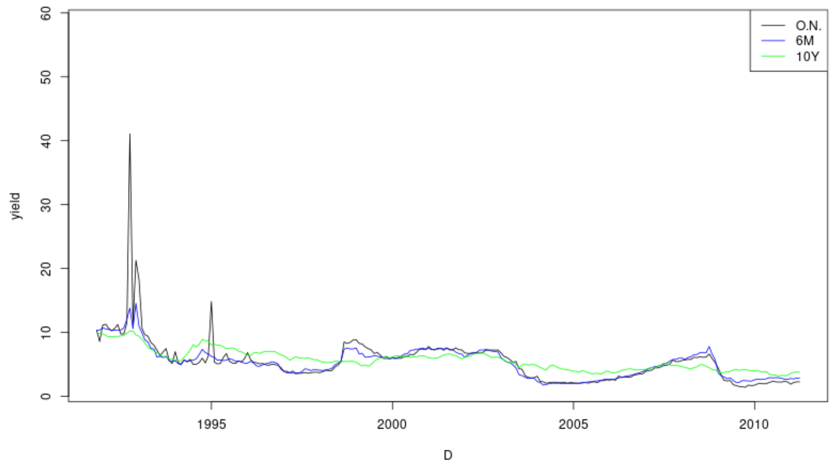


Time series plot of selected maturities.

```
> plot(D,termstru$NIBOR.T.N.nom,ylab="yield",type="l",ylim=ydim)
> lines(D,termstru$NIBOR.6M.nom)
> lines(D,termstru$NIBOR.ST10.nom)
```

Observe here how more easy it is to specify the plots when the columns are named.

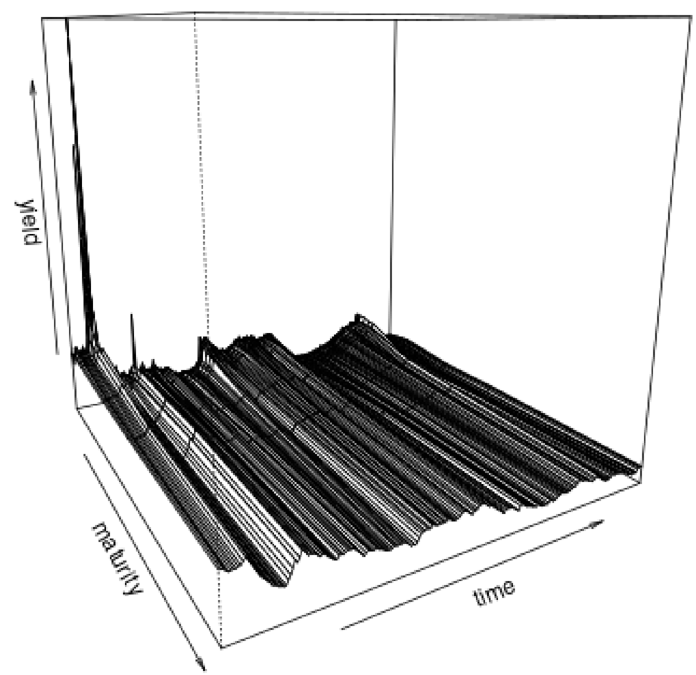
Figure 11 Time series plot of selected maturities



R has some extremely nice 3d routines.
Let us first look at the default version

```
> persp(mat,times,t(termstru[,2:13]),ylab="time",xlab="maturity",zlab="yield",theta=60)
```

Figure 12 3d plots of Norw term structures

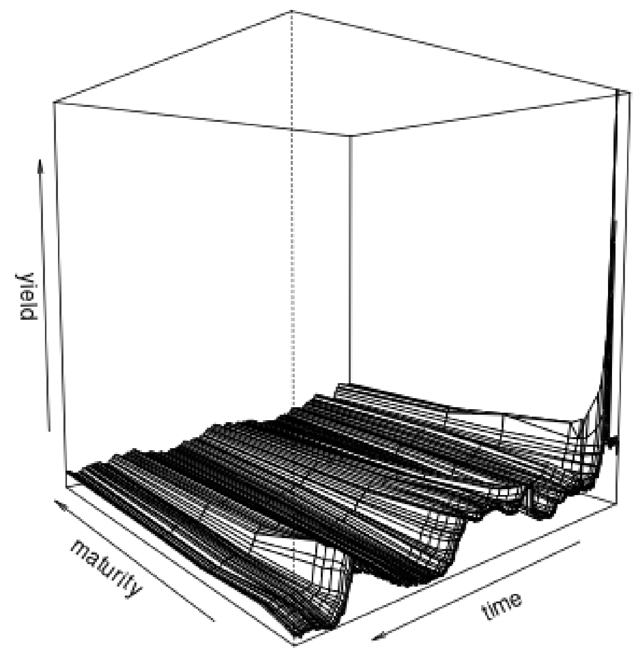


Note that it is hard to read such 3d pictures, one may get confused.

One way to clear up such ambiguities is to rotate the picture, change the perspective. In R that is very easy, one merely spesifies another placing of the perspective by the parameters `theta` and `phi`. Doing so one changes completely the perspective. and see the short end of the term structure much better.

```
> persp(mat,times,t(termstru[,2:13]),ylab="time",xlab="maturity",zlab="yield",theta=-130,phi=5)
```

Figure 13 3d plots of Norw term structures, alternative perspective



4.1 Plotting the Black Sholes option price

Consider the Black Scholes option price

$$c = SN(d_1) - Ke^{-r(T-t)}N(d_2)$$

$$d_1 = \frac{\ln\left(\frac{S}{K}\right) + \left(r + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}$$

$$d_2 = d_1 - \sigma\sqrt{(T-t)}$$

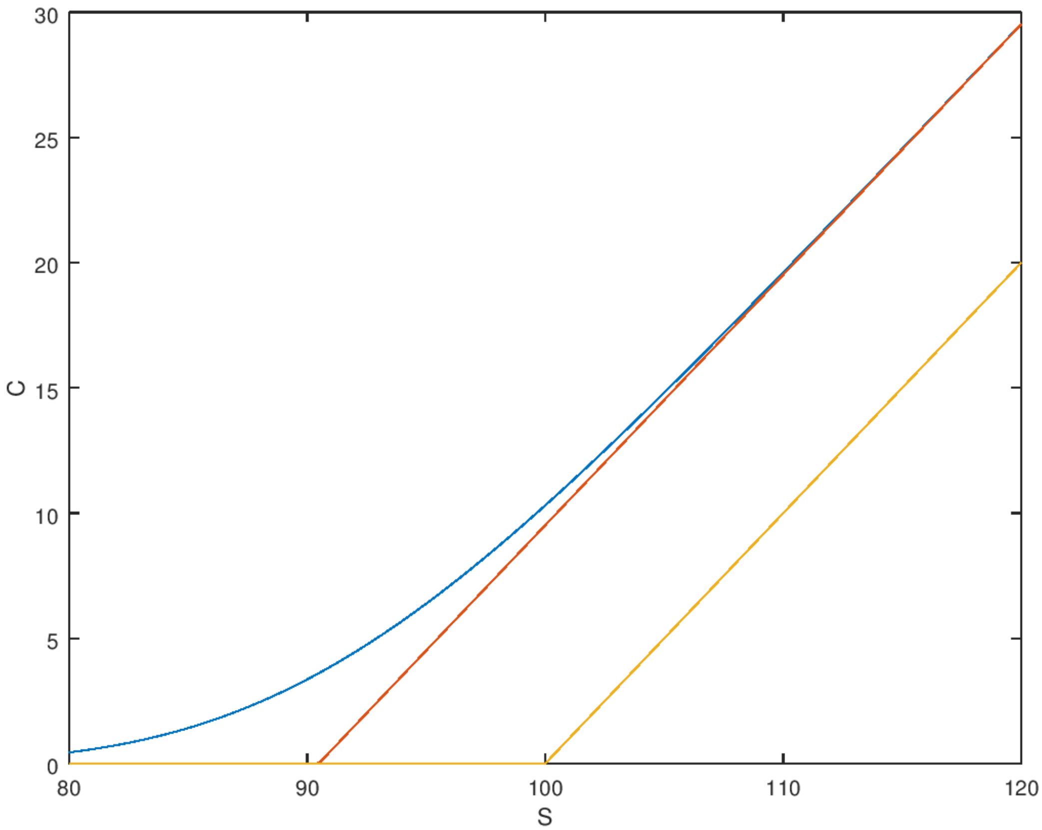
where S: price underlying. K: exercise price. r: interest rate. σ volatility of underlying. T: maturity date. $(T-t)$: time to maturity.

Let $S = 100$, $X = 100$, $\sigma = 0.1$, time=1, and $r = 0.1$; Plot the Black Scholes value varying S from 80 to 120.

```
function c = black_scholes_call(S,K,r,sigma,time)
   time_sqrt = sqrt(time);
   d1 = (log(S/K)+r*time)/(sigma*time_sqrt)+0.5*sigma*time_sqrt;
   d2 = d1-(sigma*time_sqrt);
   c = S * normcdf(d1) - K * exp(-r*time) * normcdf(d2);
endfunction

X=100;
sigma=0.1;
time=1;
r=0.1;
Srange=[80:0.2:120];
C=[];
Tight=[];
Exercise=[];
for S=Srange
    C  = [C black_scholes_call(S,X,r,sigma,time)];
    Tight = [Tight max(S-exp(-r*time)*X,0)];
    Exercise = [Exercise max(S-X,0)];
endfor;
plot(Srange,C,Srange,Tight,Srange,Exercise)
xlabel("S");
ylabel("C");
print("black_scholes_function_s.png","-dpng");
```

Results in the following picture


30
25
20
15
10
5
0
C
80
90
100
110
120
S

5 Readings

Schwabish (2014)

References

Jonathan A Schwabish. An economist's guide to visualizing data. *Journal of Economic Perspectives*, 28(1):209–234, 2014.