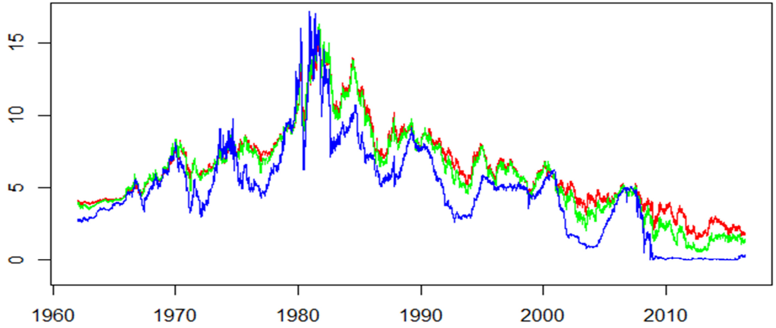
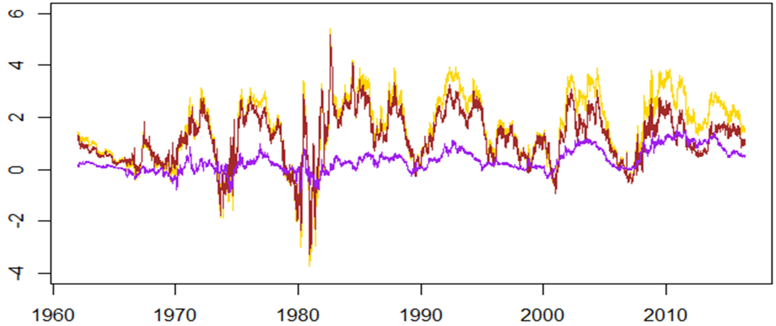
<u>Exercise:</u> Download historical yields of the 10-year treasury note (^TNX), the 5-year treasury note (^FVX), and the 13-week treasury bill (^IRX) from Yahoo!Finance and produce time series plots of the yields and the spreads.

```
TNX <- read.csv("^TNX.csv",head=T,na.strings="null")
FVX <- read.csv("^FVX.csv",head=T,na.strings="null")
IRX <- read.csv("^IRX.csv",head=T,na.strings="null")
  # 7 columns: Date,Open,High,Low,Close,AdjClose,Vol

Y <- TNX
Y <- merge(Y,FVX,by.x="Date",by.y="Date",all=F)
   # all=F: excl. incomplete cases
Y <- merge(Y,IRX,by.x="Date",by.y="Date",all=F)
Y <- na.omit(Y)   # rows with missing values are omitted

d <- as.Date(Y[,1]); N <- length(d)
Y <- cbind(Y[,6],Y[,12],Y[,18])   # Adj Close

par(mfrow=c(2,1),mar=c(2,2,0.5,0.5))
plot(d,Y[,1],type="l",col="red",ylim=c(-1,17))
lines(d,Y[,2],col="green"); lines(d,Y[,3],col="blue")
```

```
S <- cbind(Y[,1]-Y[,3],Y[,2]-Y[,3],Y[,1]-Y[,2])
plot(d,S[,1],type="l",col="gold",ylim=c(-4,6))
lines(d,S[,2],col="brown"); lines(d,S[,3],col="purple")
```





Let X be a vector-valued process with first differences

$$\Delta X_t = \Delta \begin{pmatrix} X_{t1} \\ \vdots \\ X_{tk} \end{pmatrix} = \begin{pmatrix} \Delta X_{t1} \\ \vdots \\ \Delta X_{tk} \end{pmatrix} = \begin{pmatrix} (1-L)X_{t1} \\ \vdots \\ (1-L)X_{tk} \end{pmatrix} = \begin{pmatrix} X_{t1} - X_{(t-1)1} \\ \vdots \\ X_{tk} - X_{(t-1)k} \end{pmatrix}$$

and assume that $\Delta X - E(\Delta X)$ is stationary.

If $X - E(X)$ is stationary, then X is said to be **integrated of order 0** or ***I*(0)**, otherwise X is said to be **integrated of order 1** or ***I*(1)**.

If each component of X is integrated of order 1 but some nontrivial linear combination

$$\lambda_1 X_{t1} + \ldots + \lambda_k X_{tk}$$

of its components is $I(0)$, then X is called **cointegrated**.

The vector containing the coefficients of the linear combination is called **cointegrating vector**. The number of linearly independent cointegrating vectors is called the **cointegrating rank** and the space spanned by these vectors is called the **cointegrating space**.

<u>Exercise:</u> Let U be univariate white noise. Show that the process X defined by

$$X_t = a + bt + U_t$$

is $I(0)$. C1

<u>Exercise:</u> Let U and V be the components of bivariate white noise. Show that

(i) the process W defined by

$$W_0 = 0,\ W_1 = U_1,\ W_2 = U_1 + U_2,\ W_3 = U_1 + U_2 + U_3, \ldots$$

and

$$W_{-1} = -U_{-1},\ W_{-2} = -U_{-1} - U_{-2},\ W_{-3} = -U_{-1} - U_{-2} - U_{-3}, \ldots$$

is $I(1)$,

(ii) the bivariate process X defined by

$$X_t = \begin{pmatrix} W_t \\ W_t + V_t \end{pmatrix}$$

is cointegrated. C2

Exercise: Check whether the yield series are cointegrated.

```
library(tseries)
adf.test(Y[,1],k=0) # const&trend, H0=UR, HA=station.
Dickey-Fuller=-1.8192, Lag order=0, p-value=0.655
adf.test(Y[,1],k=5)
Dickey-Fuller=-1.9341, Lag order=5, p-value=0.6064
adf.test(Y[,1],k=20)
Dickey-Fuller=-2.0851, Lag order=20, p-value=0.5424
adf.test(Y[,2],k=0)
Dickey-Fuller=-1.9427, Lag order=0, p-value=0.6027
adf.test(Y[,2],k=5)
Dickey-Fuller=-2.0841, Lag order=5, p-value=0.5428
adf.test(Y[,2],k=20)
Dickey-Fuller=-2.323, Lag order=20, p-value=0.4415
adf.test(Y[,3],k=0)
Dickey-Fuller=-2.3699, Lag order=0, p-value=0.4216
adf.test(Y[,3],k=5)
Dickey-Fuller=-2.684, Lag order=5, p-value=0.2885
adf.test(Y[,3],k=20)
Dickey-Fuller=-2.9804, Lag order=20, p-value=0.1629
```

The spreads are linear combinations of the yields.

```
adf.test(S[,1],k=0)
Dickey-Fuller=-4.291, Lag order=0, p-value=0.01*
adf.test(S[,1],k=5)
Dickey-Fuller=-4.8241, Lag order=5, p-value=0.01*
adf.test(S[,1],k=20)
Dickey-Fuller=-4.9111, Lag order=20, p-value=0.01*
adf.test(S[,2],k=0)
Dickey-Fuller=-4.9963, Lag order=0, p-value=0.01*
adf.test(S[,2],k=5)
Dickey-Fuller=-5.5293, Lag order=5, p-value=0.01*
adf.test(S[,2],k=20)
Dickey-Fuller=-5.3636, Lag order=20, p-value=0.01*
adf.test(S[,3],k=0)
Dickey-Fuller=-4.7901, Lag order=0, p-value=0.01*
adf.test(S[,3],k=5)
Dickey-Fuller=-4.7769, Lag order=5, p-value=0.01*
adf.test(S[,3],k=20)
Dickey-Fuller=-4.8341, Lag order=20, p-value=0.01*

* warning: p-value smaller than printed p-value
```

Exercise: Apply ADF tests to yield subseries.

```
library(tseries); h <- which(d<"1980-01-01")
adf.test(Y[h,1],k=0) # const&trend, H0=UR, stationary
Dickey-Fuller=-1.6795, Lag order=0, p-value=0.7141
adf.test(Y[h,1],k=5)
Dickey-Fuller=-2.4581, Lag order=5, p-value=0.3844
adf.test(Y[h,1],k=20)
Dickey-Fuller=-2.9934, Lag order=20, p-value=0.1577
adf.test(Y[h,2],k=0)
Dickey-Fuller=-1.7915, Lag order=0, p-value=0.6666
adf.test(Y[h,2],k=5)
Dickey-Fuller=-2.5334, Lag order=5, p-value=0.3525
adf.test(Y[h,2],k=20)
Dickey-Fuller=-3.0593, Lag order=20, p-value=0.1298
adf.test(Y[h,3],k=0)
Dickey-Fuller=-0.96289, Lag order=0, p-value=0.9448
adf.test(Y[h,3],k=5)
Dickey-Fuller=-1.6594, Lag order=5, p-value=0.7226
adf.test(Y[h,3],k=20)
Dickey-Fuller=-1.4352, Lag order=20, p-value=0.8175
```

```
h <- which(d>="1980-01-01")
adf.test(Y[h,1],k=0)
Dickey-Fuller=-3.3442, Lag order=0, p-value=0.06292
adf.test(Y[h,1],k=5)
Dickey-Fuller=-3.5553, Lag order=5, p-value=0.03686
adf.test(Y[h,1],k=20)
Dickey-Fuller=-3.8664, Lag order=20, p-value=0.01559
adf.test(Y[h,2],k=0)
Dickey-Fuller=-3.0885, Lag order=0, p-value=0.1172
adf.test(Y[h,2],k=5)
Dickey-Fuller=-3.3721, Lag order=5, p-value=0.0581
adf.test(Y[h,2],k=20)
Dickey-Fuller=-3.8881, Lag order=20, p-value=0.01451
adf.test(Y[h,3],k=0)
Dickey-Fuller=-2.8143, Lag order=0, p-value=0.2334
adf.test(Y[h,3],k=5)
Dickey-Fuller=-3.2357, Lag order=5, p-value=0.08162
adf.test(Y[h,3],k=20)
Dickey-Fuller=-3.9408, Lag order=20, p-value=0.01187
```

The yield series might actually be $I(0)$ and appear $I(1)$ only because of a structural break in the trend.

To test whether a vector-valued process X is cointegrated we may proceed as follows. First we use an ADF test to test whether each component is $I(1)$ individually. If the unit root hypothesis cannot be rejected for any component, we next estimate the cointegrating vector $\lambda=(\lambda_1,\ldots,\lambda_k)^T$ subject to some normalization condition like

$$\lambda=(1,-\xi_2,\ldots,-\xi_k)^T,$$

e.g., by minimizing the sum of squared residuals

$$\sum_{t=1}^{n}(X_{t1}-\xi_2 X_{t2}-\ldots-\xi_k X_{tk})^2,$$

$$\sum_{t=1}^{n}(X_{t1}-a-\xi_2 X_{t2}-\ldots-\xi_k X_{tk})^2,$$

or

$$\sum_{t=1}^{n}(X_{t1}-a-bt-\xi_2 X_{t2}-\ldots-\xi_k X_{tk})^2.$$

Finally, we apply an ADF test to the residuals. Rejection of the unit root hypothesis is interpreted as evidence in favor of cointegration.

Unfortunately, the (asymptotic) distribution of the final ADF test depends not only on the specification of the deterministic trend and the number of included lags but also on k. We therefore cannot use the same critical values for a series of residuals as for a series of observations. Of course, the whole test procedure depends also on the specification of the individual ADF tests applied to the components of X.

If there are two or more cointegrating relationships, it is a priori not clear, which one is estimated. For example, if the three components X_{t1}, X_{t2}, X_{t3} of the 3-dimensional process X represent 3-month, 5-year, and 10-year government bond yields, all three spreads $X_{t2}-X_{t1}$, $X_{t3}-X_{t2}$, and $X_{t3}-X_{t1}$ as well as all linear combinations

$$\nu_1(X_{t2}-X_{t1})+\nu_2(X_{t3}-X_{t2})+\nu_3(X_{t3}-X_{t1})$$

might appear stationary.

In case of a cointegrating relationship

$$y_t=\beta x_t+v_t$$

between the non-stationary processes x and y, the simple model

$$\Delta y_t=\gamma\Delta x_{t-1}+w_t$$

for the stationary differences could be augmented by an error-correction term that counterbalances the deviation from the long-term equilibrium, i.e.,

$$\Delta y_t=\gamma\Delta x_{t-1}+\alpha(y_{t-1}-\beta x_{t-1})+u_t.$$

The parameter α determines the speed of the return to the equilibrium. Normally, we would expect that this parameter is negative.

The estimation of the parameters can easily be carried out by OLS in two steps (first β and then γ and α), or in only one step if the model is rewritten as

$$\Delta y_t=\gamma\Delta x_{t-1}+\pi_1 y_{t-1}+\pi_2 x_{t-1}+u_t.$$

More generally, a **vector error correction model (VECM)** for the $I(1)$ process X is given by

$$\Delta X_t=\Pi X_{t-1}+\sum_{j=1}^{p-1}\Gamma_j\Delta X_{t-j}+U_t,$$

where ΠX_{t-1} is the only potential source of non-stationarity. If there are no cointegrating relationships, the rank r of Π must be zero.

Rewriting the VECM as an AR model for the levels, i.e.,

$$\begin{aligned}X_t &= X_{t-1}+\Pi X_{t-1}+(\Gamma_1 X_{t-1}-\Gamma_1 X_{t-2})+(\Gamma_2 X_{t-2}-\Gamma_2 X_{t-3})+\ldots\\ &\qquad +(\Gamma_{p-2}X_{t-(p-2)}-\Gamma_{p-2}X_{t-(p-1)})+(\Gamma_{p-1}X_{t-(p-1)}-\Gamma_{p-1}X_{t-p})+U_t\\ &= (I+\Pi+\Gamma_1)X_{t-1}+(-\Gamma_1+\Gamma_2)X_{t-2}+\ldots+(-\Gamma_{p-2}+\Gamma_{p-1})X_{t-(p-1)}\\ &\qquad -\Gamma_{p-1}X_{t-p}+U_t\\ &= \Phi_1 X_{t-1}+\ldots+\Phi_p X_{t-p}+U_t,\end{aligned}$$

we find a unit root if r=0 because

$$z=1 \Rightarrow \det(I-\Phi_1 z-\ldots-\Phi_p z^p)=\det(I-\Phi_1-\ldots-\Phi_p)=\det(-\Pi)=0.$$

CE

Johansen's test procedure:

Consider now the more general VECM

$$\Delta X_t = C + \Pi X_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \Delta X_{t-j} + U_t,$$

which contains the deterministic term C.

To estimate the number r of cointegrating relationships we may proceed as follows. First we use a likelihood ratio test to test the null hypothesis H_0: $r=0$ against the alternative hypothesis H_A: $r>0$. If the null hypothesis is rejected, we test H_0: $r=1$ against H_A: $r>1$. If the null hypothesis is rejected again, we continue in this fashion until the null hypothesis is not rejected. This sequential testing procedure is carried out without assuming that C satisfies the restriction $C=-\Pi EX_{t-1}$. Hence, the differences may have non-zero means and the levels may exhibit linear trends.

<u>Exercise:</u> Estimate the number r of cointegrating relationships for the yield dataset (^IRX, ^FVX, ^TNX).

```
library(urca)
h <- ca.jo(ts(Y),K=10,ecdet="const")
                         # with constant and lag order=10
summary(h)

Values of teststatistic and critical values of test:
          test 10pct  5pct  1pct
r <= 2 |  1.97  7.52  9.24 12.97
r <= 1 | 19.04 13.75 15.67 20.20
r = 0  | 56.90 19.77 22.00 26.81
```

Apparently, there are two linearly independent cointegrating relationships because the hypotheses $r=0$ and $r\leq 1$ can be safely rejected.

In general, if there are r cointegrating relationships, the k-dimensional process X is driven by r $I(0)$ components and k-r $I(1)$ components **(Granger representation)**. In our case, there is only one $I(1)$ component, i.e., all three yield series have the same stochastic trend.