Analyzing and modelling exchange rate data using VAR framework

ROYAL INSTITUTE
OF TECHNOLOGY

Rokas Serepka

Royal Institute of Technology

A thesis submitted for the degree of

Master in Mathematical Statistics and Financial Mathematics

2012

Acknowledgements

I would like to thank KTH for giving me valuable education in field of financial mathematics. Especially professor Tobias Rydn for his ideas, guidance and patience during the time of my theses project.

I also would like to thank my family Lina and Ula for being there for me and never stopping supporting me.

Abstract

In this report analysis of foreign exchange rates time series are performed. First, triangular arbitrage is detected and eliminated from data series using linear algebra tools. Then Vector Autoregressive processes are calibrated and used to replicate dynamics of exchange rates as well as to forecast time series. Finally, optimal portfolio of currencies with minimal Expected Shortfall is formed using one time period ahead forecasts.

Contents

1 Introduction **1**

2 Theoretical background **2**

2.1 Linear Algebra 2

2.2 Time series 4

2.3 Portfolio optimization 7

3 Time series under study **10**

4 Data projection **12**

5 VAR modelling **14**

6 Portfolio optimization **16**

7 Conclusions **18**

Bibliography **19**

1 Introduction

Foreign exchange market is the largest financial market in the world [5]. With daily turnover of \$4 trillion, which makes it also most liquid one. Market grew by 20% from 2007 to 2010 and 51% since 2004. Spot rate trades increased from \$1 trillion to \$1.5 trillion from 2007 to 2010 which makes it biggest part of total increase for that period and making it 37.4% of total turnover [1]. Therefore analysis of spot rates time series, arbitrage opportunities and portfolio optimization is an interesting and relevant topic. Daily data is easily accessible from the Internet and analysis can be carried out.

In Chapter 2 theoretical aspects, necessary for this work, are covered. Chapter will start with linear algebra, which is needed for data preprocessing. Later projection matrices and pesudoinvers are defined as they are of biggest interest for this work. Triangular arbitrage will be introduced and linear constraints matrix with respect to arbitrage will be set up. Vector Autoregressive (VAR) process and toolbox of its estimation, order selection and stability testing methods will be presented later. Chapter ending with risk measure and portfolio optimization with these measures.

Chaper 3 and 4 will present data for exchange rates under consideration and its main qualities (series plots and correlation heat map). Then arbitrage elimination using linear constraints matrix and linear projection will be done and results will be compared with data before projection. In addition variances of original and projected data will be compared.

Chapter 5 covers VAR analysis, estimation and forecasting. From this part, just USD exchange rates will be used. This will be done to have a reference currency and measure portfolio value in next chapter. The USD is chosen as reference currency as it accounts for as much as 84.9% (out of 200%) of daily trading turnover and can be still seen as the biggest currency in the world. Thus in Chapter 6 forecasts from Chapter 5 are used for portfolio optimization and trading simulations for currency portfolios.

2 Theoretical background

2.1 Linear Algebra

First, I will present some basic results from linear algebra. Most of them can be found, in more details and examples, in the book by Axler (1997 [8]). These results will be used to eliminate triangular arbitrage opportunities in downloaded data.

Definition 2.1 *A* ***vector space*** *is a set V along with an addition on V and a scalar multiplication on V such that the following properties hold:*

commutativity $u + v = v + u$, *for all* $u, v \in V$,

associativity $(u + v) + w = u + (v + w)$ *and* $(ab)v = a(bv)$ *for all* $u, v, w \in V$ *and all* $a, b \in \mathbb{R}$,

additive identity *there exists an element* $0 \in V$ *such that* $v + 0 = v$ *for all* $v \in V$,

additive inverse *for every* $v \in V$, *there exists* $w \in V$ *such that* $v + w = 0$,

multiplicative identity $1v = v$ *for all* $v \in V$,

distributive properties $a(u + v) = au + av$ *and* $(a + b)u = au + bu$ *for all* $a, b \in \mathbb{R}$ *and all* $u, v \in V$.

Definition 2.2 *A nonempty subset U of a vector space V is a* ***subspace*** *of V if*

- $x, y \in U \Rightarrow x + y \in U$,
- $x \in U \Rightarrow \alpha x \in U, \forall \alpha \in \mathbb{R}$.

Data for daily exchange rates can be seen as vectors and all these vectors form a vector space. Assume there are n currencies and let us denote them as a_i, $i = 1 \dots n$. Then there will be $\binom{n}{2} = n(n-1)/2$ exchange rates $\frac{a_i}{a_j}$, $i, j = 1 \dots n$ stored in vectors

$$\mathbf{x} = \left(\frac{a_1}{a_2}, \frac{a_1}{a_3}, \frac{a_1}{a_4}, \dots, \frac{a_1}{a_n}, \frac{a_2}{a_3}, \frac{a_2}{a_4}, \dots, \frac{a_{n-1}}{a_n} \right)^T. \tag{2.1}$$

Definition 2.3 *A list $(v_1, \dots, v_m)$ of vectors in V is called linearly independent if the only choice of $a_1, \dots, a_m \in \mathbb{R}$ that satisfies $a_1 v_1 + \dots + a_m v_m = 0$ is $a_1 = \dots = a_m = 0$.*

Definition 2.4 *A* ***basis*** *of V is a list of vectors in V that is linearly independent and spans V.*

Definition 2.5 *An **inner product** on V is a function that takes each ordered pair (u, v) of elements of V to a number $\langle u, v\rangle \in \mathbb{R}$ and has the following properties:*

positivity $\langle v, v\rangle \geq 0$ *for all* $v \in V$,

definiteness $\langle v, v\rangle = 0$ *if and only if* $v = 0$,

additivity in first slot $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ *for all* $u, v, w \in V$,

homogeneity in first slot $\langle av, w\rangle = a\langle v, w\rangle$ *for all* $a \in \mathbb{R}$ *and all* $v, w \in V$,

symmetry $\langle v, w\rangle = \langle w, v\rangle$ *or all* $v, w \in V$.

Definition 2.6 *For $v \in V$ the norm of v, denoted $\|v\|$, by*

$$\|v\| = \sqrt{\langle v, v\rangle}.$$

Definition 2.7 *Two vectors $u, v \in V$ are said to be **orthogonal** if $\langle u, v\rangle = 0$.*

Definition 2.8 *A list of vectors is called **orthonormal** if the vectors in it are pairwise orthogonal and each vector has norm 1. In other words, a list $(e_1, \ldots, e_m)$ of vectors in V is orthonormal if $\langle e_j, e_k\rangle$ equals 0 when $j \neq k$ and equals 1 when $j = k$ (for $j, k = 1, \ldots, m$).*

Definition 2.9 *An **orthonormal basis** of V is an orthonormal list of vectors in V that is also a basis of V.*

Now let's go back to exchange rates. Exchange rates can be expressed in terms of each other $\left(\frac{a_i}{a_j}\right) = \left(\frac{a_i}{a_k}\right) \cdot \left(\frac{a_k}{a_j}\right)$. If one of these relationships does not hold, there will be triangular arbitrage opportunity. Which basically means, that with taking no risk, money can be made, by buying and instantly selling currencies [2]. In practice, such opportunities exist, even though they can be neglected due to transactions costs. Nevertheless, it would be nice to take them away form data before modelling. Here is where linear algebra steps in. Relationships $\left(\frac{a_i}{a_j}\right) = \left(\frac{a_i}{a_k}\right) \cdot \left(\frac{a_k}{a_j}\right)$ can be transformed to log-scale to make them linear $\ln\left(\frac{a_i}{a_j}\right) = \ln\left(\frac{a_i}{a_k}\right) + \ln\left(\frac{a_k}{a_j}\right)$. Linear relationships can be rewritten in matrix form

$$\mathbf{A}\mathbf{x}^T = \mathbf{0} \tag{2.2}$$

where

$$\mathbf{A} = \begin{pmatrix} -1_{\{1,2\}} & 1_{\{1,3\}} & 0 & 0 & \cdots & -1_{\{2,3\}} & 0 & \cdots & \cdots & 0 \\ -1_{\{1,2\}} & 0 & 1_{\{1,4\}} & 0 & \cdots & 0 & -1_{\{2,4\}} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & -1_{\{i,k\}} & \cdots & 0 & 1_{\{i,j\}} & 0 & \cdots & -1_{\{k,j\}} & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} \tag{2.3}$$

and $\mathbf{x}$ is defined as in (2.1). After having constructed matrix A it is possible to project data onto the subspace where the linear constraints do hold.

Definition 2.10 *Let A be an $m \times n$ matrix. The null and range space of A are the set of vectors*

$$N(A) := \left\{x \in \mathbb{R}^n | Ax^T = \mathbf{0}\right\}$$
$$R(A) := \left\{Ax^T \in \mathbb{R}^m | x \in \mathbb{R}^n\right\}$$

Next definition is taken from [3].

Definition 2.11 *For $v \in V$, let $v = m + n$, where $m \in U$ and $n \in U^{\perp}$.*

- *m is called the **orthogonal projection** of v onto U.*
- *The projector P_U onto U along $U^{\perp}$ is called the **orthogonal projector** onto U.*
- *P_U is the unique linear operator such that $P_U = m$*

Theorem 2.1 *Let A be and $m \times n$ matrix. The pseudo inverse A^+ is an $n \times m$ matrix satisfying:*

i. *$A^+A = P$ is the orthogonal projector $P : \mathbb{R}^n \to N(A)^{\perp}$ and $AA^+ = \bar{P}$ is the orthogonal projector $\bar{P} : \mathbb{R}^m \to R(A)$.*

ii. *The following formulas hold:*

- *(a)* $A^+A = (A^+A)^T$
- *(b)* $AA^+ = (AA^+)^T$
- *(c)* $AA^+A = A$
- *(d)* $A^+AA^+ = A^+$

Proof of **Theorem 2.1** can be found in [9]. For my purposes I will need the orthonormal basis of matrix A and its pseudo-inverse.

2.2 Time series

Modelling the big amount of related time series individually is not an easy task if possible at all. Using multivariate models makes it much easier. It also takes into account the correlation between variables, which is very important with any amount of time series under analysis. I will consider **Vector Autoregressive** (VAR) models for my data. Most of the multivariate time series theory is taken from Lutkepohl (2005 [4]).

A VAR model of order p (VAR(p)) can be written as

$$y_t = \nu + A_1 y_{t-1} + \cdots + A_p y_{t-p} + u_t, \quad t = 0, 1, \ldots \tag{2.4}$$

where $y_t = (y_{1t}, \ldots, y_{Kt})'$ is a random vector, K is the process dimension, the A_i are coefficient matrices, $\nu = (\nu_1, \ldots, \nu_K)'$ is fixed vector of terms allowing for the possibility of nonzero mean and $u_t = (u_{1t}, \ldots, u_{Kt})'$ is white noise with covariance matrix Σ_u.

First thing to do, when analysing the process itself, is to check its **stability**. If the process is not stable it can explode or produce irrelevant results. The process in (2.4) is stable, if its reverse characteristic polynomial has no roots in and on the complex unit circle

$$\det(I_K - A_1 z - \cdots - A_p z^p) \neq 0 \quad \text{for } |z| \leq 1. \tag{2.5}$$

For computational purposes I have used another (equivalent) definition of stability from Tsay (2002 [10]). The VAR(p) can be written down as a VAR(1)

$$Y_t = \boldsymbol{\nu} + \mathbf{A} Y_{t-1} + U_t \tag{2.6}$$

where

$$Y_t = \begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p+1} \end{bmatrix}, \quad \nu = \begin{bmatrix} \nu \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

$$\boldsymbol{A} = \begin{bmatrix} A_1 & A_2 & \dots & A_{p-1} & A_p \\ I_K & 0 & \dots & 0 & 0 \\ 0 & I_K & \dots & & 0 \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & I_K & 0 \end{bmatrix}, \quad U_t = \begin{bmatrix} u_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

The process (2.6) will be stable when all eigenvalues of $\mathbf{A}$ are less than 1 in the modulus.

Let's assume that time series $y_1, \dots, y_T$ of currency rates are available and are generated by a stable VAR(p) process as in (2.4), but now coefficients ν, A_i, and Σ_u are unknown. By following [4], estimates can be achieved using **multivariate least squares** (**LS**) and **maximum likelihood** (**ML**) methods. There will be p presample vectors $y_{-p+1}, \dots, y_0$ of data needed and we define

$$\begin{array}{ll} \mu = \dfrac{1}{T}\displaystyle\sum_{t=1}^{T} y_t & \\ Y := (y_1, \dots, y_T) & (K \times T), \\ B := (\nu, A_1, \dots, A_p) & (K \times (Kp+1)), \\ Z_t := \begin{bmatrix} 1 \\ y_t \\ \vdots \\ y_{t-p+1} \end{bmatrix} & ((Kp+1) \times 1), \\ U := (Z_0, \dots, Z_{T_1}) & (K \times T), \\ A := (A_1, \dots, A_p) & (K \times Kp), \\ Y^0 := (y_1, \dots, y_T - \mu) & (K \times T) \\ Y_t^0 := \begin{bmatrix} y_t - \mu \\ \vdots \\ y_{t-p+1} - \mu \end{bmatrix} & (Kp+1), \\ X := (Y_0^0, \dots, Y_{T-1}^0) & \end{array} \tag{2.7}$$

Using this notation (2.6) can be rewritten as

$$Y = BZ + U.$$

LS and **ML** estimates of B and Σ_u can be written down in a simple and nice matrix way

$$\begin{aligned}\hat{B} &= YZ^T(ZZ^T)^{-1}\\ \tilde{\Sigma}_u &= \frac{1}{T}(Y - \hat{B}Z)(Y - \hat{B}Z)^T.\end{aligned} \tag{2.8}$$

The log-likelihood function can be written as

$$\tilde{\Sigma}_u = -\frac{KT}{2}\ln 2\pi - \frac{T}{2}\ln|\Sigma_u| - \frac{1}{2}\mathrm{tr}[(Y^0 - AX)'\Sigma_u^{-1}(Y^0 - AX)] \tag{2.9}$$

where tr is matrix trace. Maximization of the log-likelihood function gives ML estimates. It is shown in [4], that these estimates are asymptotically equivalent to the least squares estimates. For my estimates I will use (2.8) for $\tilde{B}$ and (2.9) for $\tilde{\Sigma}_u$.

In the discussion up until now, it was assumed that the VAR order p is known. In practice, however, it is almost always unknown. I will consider a few methods for order indication and selection. The principle behind all of them is: choose an upper limit of VAR order, fit model to the data and compare fitted models.

The likelihood ratio test compares the maxima of the log-likelihood function over the unrestricted and restricted parameter space. Test statistics have asymptotic $\chi^2(K^2)$ distribution. In VAR case, given upper order bound M the sequence of null and alternative hypotheses are tested:

$$H_0^i: \;\; A_{M-i+1} = 0 \;\; \text{versus} \;\; H_1^i: \;\; A_{M-i+1} \neq 0 \;\; |A_M = \cdots = A_{M-i+2} = 0, \;\; \text{for } i = 1 : M$$

The sequence is terminated when H_0^i is rejected and order $\hat{p} = M - i + 1$ is chosen. The test statistic for the i'th null hypothesis is

$$\lambda_{LR}(i) = T[\ln|\tilde{\Sigma}_u(M - i)| - \ln|\tilde{\Sigma}_u(M - i + 1)|]$$

For prediction purposes one would like to choose such an order that prediction error is minimized. As one step ahead prediction error is determined by the covariance matrix of residuals $\tilde{\Sigma}_u$ (see [4]) quite a few criteria have been developed based on it. The **Akaikes Information Criterion** (AIC), **final prediction error** (FPE), **Hanna-Quinn criterion** (HQ) and **Schwarz criterion** (SC) are few of the most popular ones in practice.

$$\begin{aligned}\mathrm{AIC}(m) &= \ln|\tilde{\Sigma}_u(m)| + \frac{2mK^2}{T},\\ \mathrm{FPE}(m) &= \left[\frac{T + Km + 1}{T - Km - 1}\right]^K |\tilde{\Sigma}_u(m)|,\\ \mathrm{HQ}(m) &= \ln|\tilde{\Sigma}_u(m)| + \frac{2\ln\ln T}{T} mK^2,\\ \mathrm{SC}(m) &= \ln|\tilde{\Sigma}_u(m)| + \frac{\ln T}{T} mK^2.\end{aligned}$$

AIC and FPE tend to overestimate the process order with positive probability while HQ and SC seem to be consistent. Nevertheless it does not necessarily mean that AIC and FPE are worst. On the contrary, let's say, for predicting purposes they suit better.

The last task to do in modelling part is to check model adequacy. One of the ways is to check the whiteness of the residuals. Residuals are estimated by $\hat{U} = (\hat{u_1}, \dots, \hat{u_T}) = Y - \hat{B}Z$, autocovariances of estimated residuals are calculated

$$\hat{C}_i = \frac{1}{T} \sum_{t=i+1}^{T} \hat{u}_t \hat{u}_{t-i}^T$$

and a test statistic is built

$$\bar{Q}_h = T^2 \sum_{i=1}^{h} (T-i)^{-1} \mathrm{tr}(\hat{C}_i^T \hat{C}_i^{-1} \hat{C}_i \hat{C}_i^{-1}).$$

$\bar{Q}_h$ is asymptotically $\chi^2(K^2(h-p))$ distributed if residuals are uncorrelated. This test is called **portmanteau test** (Box-Pierce version).

2.3 Portfolio optimization

Before defining portfolio choice, some risk measures will be introduced. Probably the most popular risk measure in practice is Value-at-Risk (**VaR**). Consider a portfolio $w^T y$, where w is vector of portfolio weights and y is vector of risky assets. Let's denote the portfolio loss function as $L(w, y) = -w^T y$ and its distribution function as $F_L(l) = P(L(w, y) \leq l)$. Then $\mathrm{VaR}_\alpha(w)$ can be defined as in [6]

Definition 2.12 *Given some confidence level $\alpha \in (0, 1)$, **VaR** of our portfolio at that confidence level α is given by the smallest number l such that the probability that the loss $L(w, y)$ exceeds l is no larger than $(1 - \alpha)$*

$$\mathrm{VaR}_\alpha(w) = \inf\{l \in \mathbb{R} : P(L(w, y) > l) \leq 1 - \alpha\} = \inf\{l \in \mathbb{R} : F_L(l) \geq \alpha\}.$$

From a probabilistic point of view, it is nothing more than quantile of loss distribution $q_\alpha(F_L(l))$.

VaR is good risk measure, but it has few major drawbacks. It is not sub-additive, which means it does not reward diversification and it ignores losses in the tail higher than confidence level. Expected shortfall (**ES**) is another risk measure closely related to $\mathrm{VaR}_\alpha(w)$, which is sub-additive.

Definition 2.13 *For a loss $L(w, y)$ with $E(|L(w, y)|) < \infty$ and distribution function $F_L(l)$ the expected shortfall at confidence level $\alpha \in (0, 1)$ is defined as*

$$\mathrm{ES}_\alpha(w) = \frac{1}{1-\alpha} \int_\alpha^1 q_u(F_L) \mathrm{d}u$$

where $q_u(F_L)$ is the quantile function of F_L. As mentioned above $\mathrm{VaR}_\alpha(w) = q_\alpha(F_L)$, so $\mathrm{ES}_\alpha(w)$ can be rewritten as

$$\mathrm{ES}_\alpha(w) = \frac{1}{1-\alpha} \int_\alpha^1 \mathrm{VaR}_u(w) \mathrm{d}u$$

and thus $\mathrm{ES}_\alpha(w)$ averages on $\mathrm{VaR}_u(w)$ with levels $u \geq \alpha$. In this way, it looks further into the tail of loss distribution. If loss distribution is continuous then one more definition of $\mathrm{ES}_\alpha(w)$ can be derived

Theorem 2.2 *For an integrable loss $L(w, y)$ with countinuous distribution function $F_L(l)$ and any $\alpha \in (0, 1)$*

$$\mathrm{ES}_\alpha(w) = E\big[L(w, y) | L(w, y) \geq \mathrm{VaR}_\alpha(w)\big] =$$
$$(1-\alpha)^{-1} E\big[L(w, y) I\{L(x, y) \geq \mathrm{VaR}_\alpha(w)\}\big] =$$
$$(1-\alpha)^{-1} \int\limits_{L(w,y) \geq \mathrm{VaR}_\alpha(w)} L(w, y) p(y) dy$$

where $p(y)$ is density of the asset returns y and $I(A)$ is indicator function.

Proof of this theorem can be found in [6].

In my portfolio choice I would like to get such weights w that $\mathrm{ES}_\alpha(w)$ would be as small as possible. Main idea behind that is to replace $\mathrm{ES}_\alpha(w)$ with a convex function $\phi_\alpha(w, \mathrm{VaR})$ defined as follows:

$$\phi_\alpha(w, \mathrm{VaR}) = \mathrm{VaR} + (1-\alpha)^{-1} \int\limits_{\mathbb{R}^m} [L(w, y) - \mathrm{VaR}]^+ p(y) dy \tag{2.10}$$

where

$$[t]^+ = \begin{cases} t, & \text{when } t > 0, \\ 0, & \text{when } t \leq 0. \end{cases}$$

Next theorem enables us to use this function for $\mathrm{ES}_\alpha(w)$ minimization.

Theorem 2.3 *Minimizing the* ES_α *of the loss associated with w is equivalent to minimizing $\phi_\alpha(w, \mathrm{VaR})$ in the sense that*

$$\min_w \mathrm{ES}_\alpha(w) = \min_{w, \mathrm{VaR}} \phi_\alpha(w, \mathrm{VaR}).$$

The proof of this theorem can be found in [6] or [7]. The integral in (2.10) can be approximated by sampling y from a VAR(p) model. With generated vectors $y_1, \ldots, y_N$, an approximation of $\phi_\alpha(w, \mathrm{VaR})$ is

$$\bar{\phi}_\alpha(w, \mathrm{VaR}) = \mathrm{VaR} + \frac{1}{(1-\alpha)} \frac{1}{N} \sum_{n=1}^{N} [L(w, y_n) - \mathrm{VaR}]^+$$

The minimization of $\bar{\phi}_\alpha(w, \mathrm{VaR})$ can be reduced to convex programming. It is equivalent to minimizing

$$\min_{\mathrm{VaR}, w, d_1, \ldots, d_N} \quad \mathrm{VaR} + \frac{1}{N} \frac{1}{1-\alpha} \sum_{n=1}^{N} d_n$$

$$\begin{aligned} s.t. \quad & d_n + \mathrm{VaR} + L(w, y_n) \geq 0 \\ & d_n \geq 0 \end{aligned}$$

where $n = 1, \ldots, N$.

For my optimization I will use rates of returns of currencies r_n as risky assets, budget constraint and a no-short-selling constraint. Hence, portfolio optimization problem is formulated as

$$
\begin{aligned}
\min_{\mathrm{VaR}, w, d_1, \ldots, d_N} \quad & \mathrm{VaR} + \frac{1}{N} \frac{1}{1-\alpha} \sum_{n=1}^{N} d_n \\
s.t. \quad & d_n + \mathrm{VaR} + w^T r_n \geq 0 \\
& d_n \geq 0 \\
& w_i \geq 0 \\
& w^T Y_0 = w_0^T Y_0
\end{aligned}
\tag{2.11}
$$

for $n = 1, \ldots, N$ and $i = 1, \ldots, K$. Here N is number of simulations, K is number of exchange rates, r_n are simulated returns, Y_0 are today exchange rates, w_0 are yesterdays portfolio weights and d_t are dummy variables.

3 Time series under study

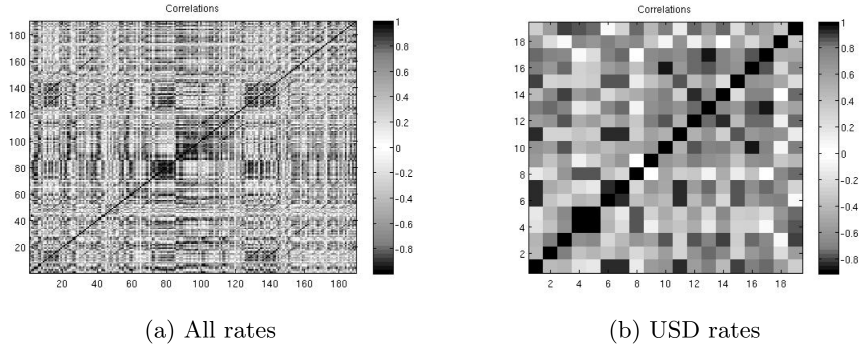
For modelling purposes multivariate data of exchange rates will be needed. 20 currencies have been chosen and daily data of their exchange rates have been downloaded.

USD	US Dollar	INR	Indian Rupee
CAD	Canadaian Dollar	IDR	Indonesian Rupee
EUR	European Euro	KRW	South Korean Won
GBP	Great Britain Pound	MXN	Mexican Peso
JPY	Japanese Yen	NOK	Norwegian Krone
ARS	Argentine Peso	RUB	Russia Ruble
AUD	Australian Dollar	ZAR	South Africa Rand
BRL	Brazilian Real	SEK	Swedish Krona
CNY	Chinese Yuan	CHF	Swiss Frank
HKD	Hong Kong Dollar	TRY	Turkish Lira

Table 3.1: Currencies

Data is organised in the following manner to match the vector in (2.1)

$$\left(\frac{\text{USD}}{\text{CAD}}, \frac{\text{USD}}{\text{EUR}}, \frac{\text{USD}}{\text{GBP}}, \ldots, \frac{\text{USD}}{\text{TRY}}, \frac{\text{CAD}}{\text{EUR}}, \frac{\text{CAD}}{\text{GBP}}, \ldots, \frac{\text{CHF}}{\text{TRY}}\right)^T.$$



(a) All rates (b) USD rates

Figure 3.1: Correlation between exchange rates. White area represents uncorrelated parts, while black area represents strong (both positively and negatively) correlation.

This vector includes 190 exchange rates. Time period for data is between 2007-10-16 and 2011-10-16, which is 1004 consecutive trading days. Thus a 190-by-1004 matrix of data is formed. After projection (next chapter) data will be divided into two samples: one for model estimation, and another for forecasting and portfolio performance comparison. Data will be also scaled to weekly (every 5'th trading day) and bi-weekly (every 10'th trading day) samples, to see how exchange rate dynamics change for longer time periods. Some examples of time series from the data set are plotted in Figure 3.2. The correlation matrix of all 190 exchange rates and of the 19 first ones $\left(\frac{\text{USD}}{\text{CAD}}, \ldots, \frac{\text{USD}}{\text{TRY}},\right)$ are plotted in Figure 3.1. It can be seen that data is strongly correlated, which supports the choice of multivariate modelling.

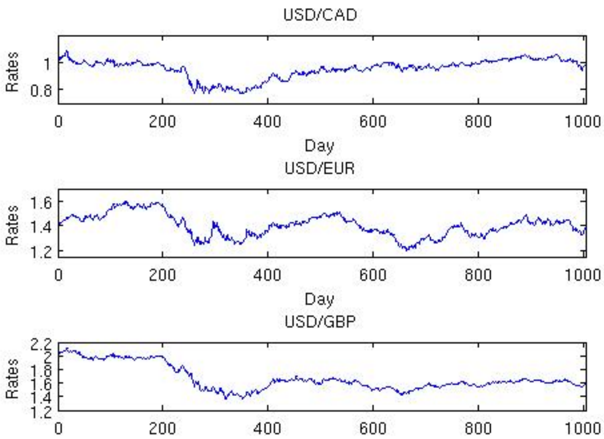
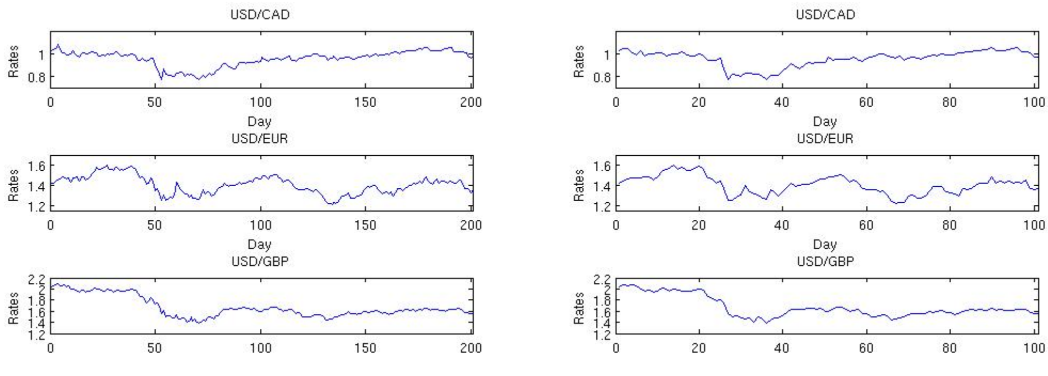


Figure 3.2: Daily data of exchange rates: $\frac{\text{USD}}{\text{CAD}}, \frac{\text{USD}}{\text{EUR}}, \frac{\text{USD}}{\text{GBP}}$



(a) Weekly data

(b) Bi-weekly

Figure 3.3: Weekly and bi-weekly data of exchange rates: : $\frac{\text{USD}}{\text{CAD}}, \frac{\text{USD}}{\text{EUR}}, \frac{\text{USD}}{\text{GBP}}$

4 Data projection

As it was indicated before, a linear projection will be used to eliminate triangular arbitrage opportunities. In Chapter 2 the matrix A of linear constraints (2.3) was constructed. In our case (with 190 exchange rates) $A \in \mathbb{R}^{1140 \times 190}$. We would like to project data to the space where the linear constraints hold, hence the null space of A. One way of achieving it is to find orthonormal basis (N) for the null space of A and its pseudo inverse N^+. Now matrix N is in $\mathbb{R}^{190 \times 19}$, $N^+ \in \mathbb{R}^{19 \times 190}$ and projection $\mathbf{P}_{N(A)} = NN^+ \in \mathbb{R}^{190 \times 190}$.

Before applying the projection, our data matrix is log-scaled $D = \log(Data)$ and premultiplied with matrix A to see, if there exists any arbitrage at all. Biggest arbitrage is 1.0532×10^{-4} which probably can be neglected due to transaction costs, but still exists. Average arbitrage would be 1.3072×10^{-5}.

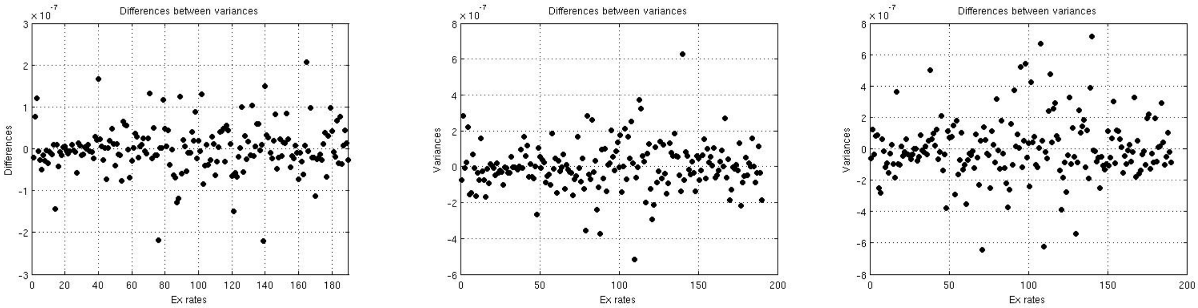
After projection is applied ($\mathbf{P}_{N(A)}D$), the resulting matrix is premultiplied with A. The biggest arbitrage now is 4.3521×10^{-14} and average would be 5.7346×10^{-15}. Such a small number is approximately equal to zero in both economical and mathematical sense and is due to rounding errors.

As P is an orthogonal projection and by using **Definition 2.11** the following is true

$$\begin{aligned} &||v||^2 = ||Pv + n||^2 = ||Pv||^2 + \langle Pv, n\rangle + \langle n, Pv\rangle + ||n||^2 \\ &= ||Pv||^2 + ||n||^2 \geq ||Pv||^2 \Rightarrow ||Pv||^2 \leq ||v||^2 \Rightarrow \langle Pv, Pv\rangle \leq \langle v, v\rangle. \end{aligned}$$

If we define $\langle X, Y\rangle = E[XY]$ and $v = X - \mu$ we get

$$\begin{aligned} &E[(P(X-\mu))^2] \leq E[(X-\mu)^2] \Rightarrow \\ &PE[(X-\mu)^2] \leq E[(X-\mu)^2] \Rightarrow \\ &PVar(X) \leq Var(X). \end{aligned}$$



(a) Daily data (b) Weekly data (c) Bi-weekly data

Figure 4.1: Differemce between variances of projected data and original ones.

So the variance of the projected data is smaller than the original one. To check that, variance of projected data is subtracted from variance of original data and results are plotted in Figure 4.1. It seems that some of variances has increased and some decreased. But on average it has decreased by: 9.1793×10^{-11} (daily data), 9.2305×10^{-11} (week data), 9.1641×10^{-11} (bi-week data).

The residuals of projection with respect to the original data might be of some interest as it would represent series of price mismatch. Some of them are plotted below.

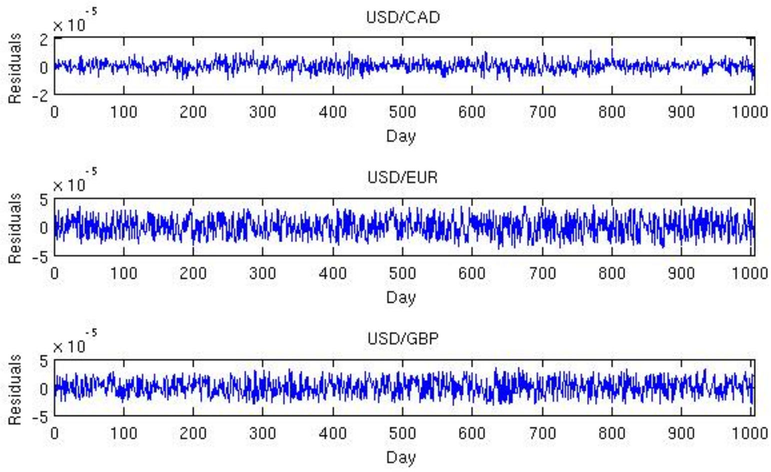
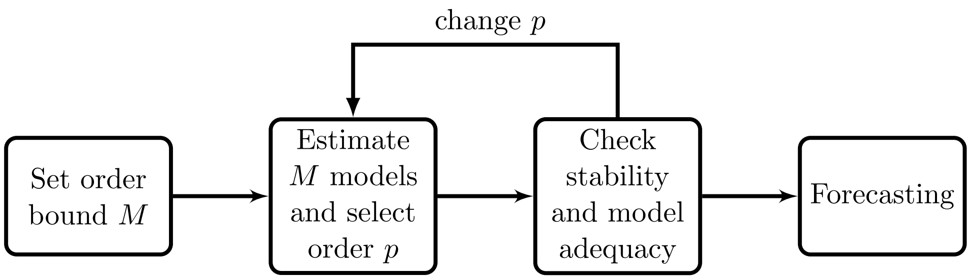


Figure 4.2: Residuals of projetion of exchange rates for: $\frac{\text{USD}}{\text{CAD}}, \frac{\text{USD}}{\text{EUR}}, \frac{\text{USD}}{\text{GBP}}$.

5 VAR modelling

For this part data of USD exchange rates will be used. In this way, a portfolio with respect to USD can be formed and forecasts will be in great interest (see Section 6). Hence the 19 first exchange rates for the projected data are taken. Some data has been put aside for out of the sample comparison and (in next chapter) portfolio optimization. Modelling procedure is set up based on theory in Chapter 2. It will be carried out according to chart below.



By a change of p, it is meant that p is decreased and the procedure repeated. The scheme was run on daily, weekly and bi-weekly data. In Table 1 the result of order indicators are presented. Likelihood-ratio test seems to give inconsistent results. It can be due to the normality assumption which may not hold with exchange rate data. AIC and FPR order estimates are greater or equal to HQ and SC which is coherent with theory.

Table 1.

Suggested order					
Data	AIC	FPE	HQ	SC	LRT
Daily	2	2	1	1	10
Weekly	8	9	8	8	10
Bi-weekly	3	10	3	3	9

Unfortunately, models of order suggested for weekly and bi-weekly data are not stable. The order is decreased by one and procedure repeated. Finally order 5 for weekly and 2 for bi-weekly is selected, which fulfil both stability and adequacy test requirements. After the models have been calibrated, forecast have been made. The historical and simulated data are plotted below.

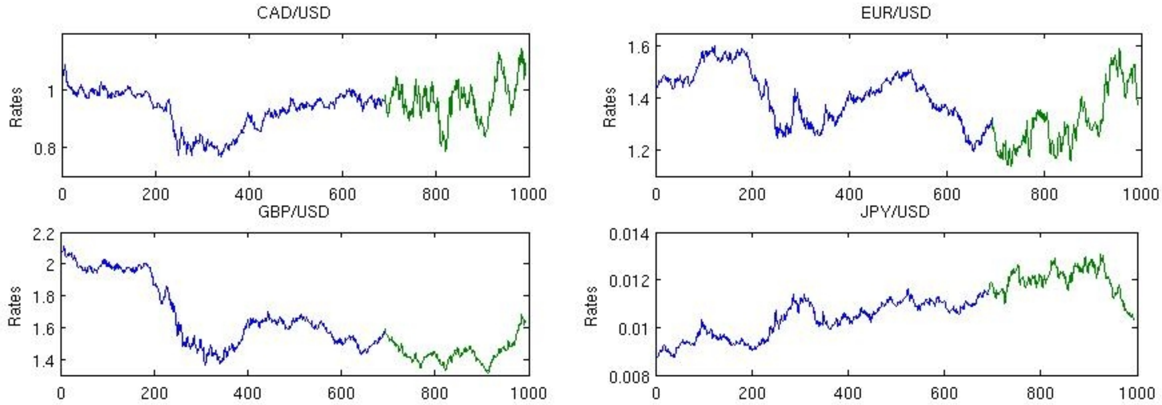


Figure 5.1: Plot of historical and foretasted daily data

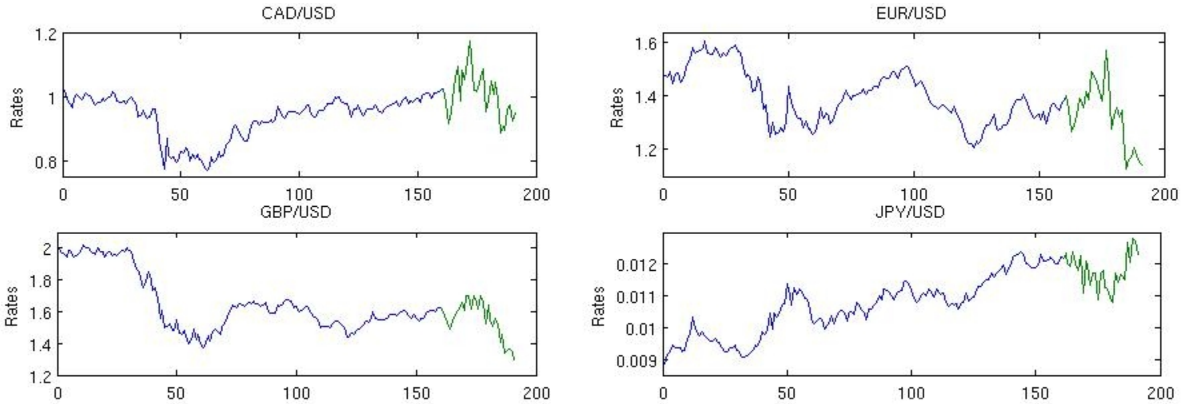


Figure 5.2: Plot of historical and foretasted weekly data

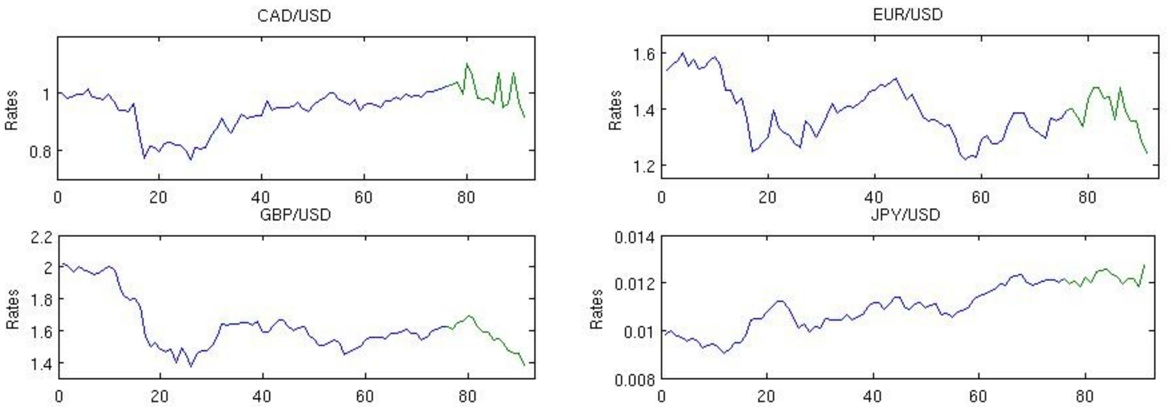


Figure 5.3: Plot of historical and foretasted bi-weekly data

From Figures 5.1, 5.2, 5.3 it seems that, simulated data for bi-weekly data resembles original process best, while for daily data worst.

6 Portfolio optimization

A portfolio of foreign currencies with respect to USD is formed. Scenarios one day ahead are simulated from the VAR(p) model as in Chapter 5. Then optimal portfolio of min-ES as in (2.11) are obtained. For the first optimization we set $w_0^T Y_0 = 100$ which means, that the initial budget is 100\$. Then data used for VAR model simulations is updated with historical data (adding one day to existing data). Simulations one day ahead are made again and optimization is repeated. For second and further optimizations the budget is what is left form yesterday's investment $w_0^T Y_0$.

Same framework is repeated for weekly data (every 5'th trading day) and bi-weekly data (every 10'th trading day) and plots of portfolio value changes can be seen in Figure 6.1 and Figure 6.2.

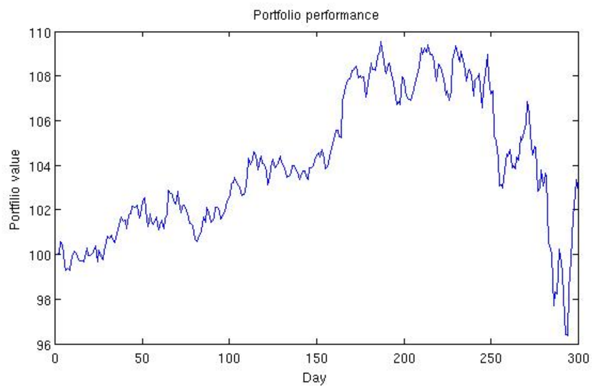
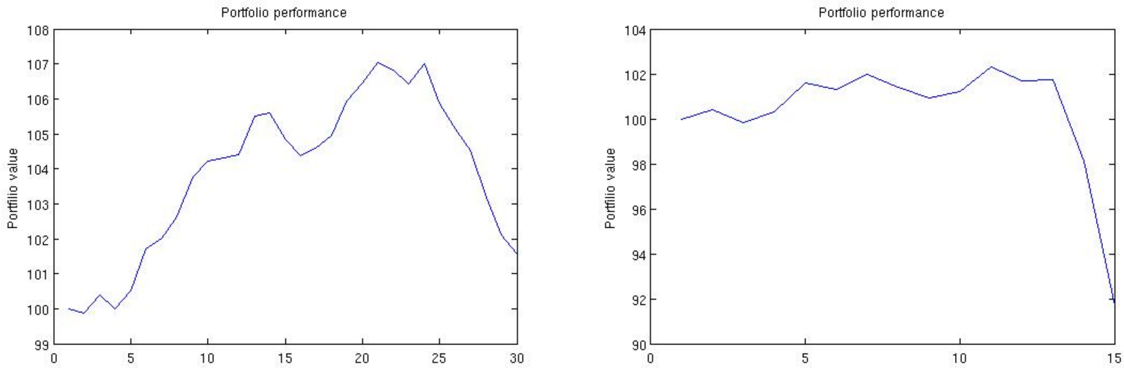


Figure 6.1: Portfolio value change (daily trading)

All portfolios perform similarly start quite well, but end up with a drop. It is due to bigger negative move in a all 19 USD exchange rates. The average changes of all 19 assets in last 50 trading days is -5.97%. Naturally the portfolio value will fall as all the assets fall. In this case, the purpose of a portfolio with minimized risk is to make this fall as small as possible. From the figures it can be seen, that the portfolio of weekly reallocations has done best and bi-weekly worst.



(a) Weekly trading

(b) Bi-weekly trading

Figure 6.2: Portfolio value change

7 Conclusions

Modelling multivariate date, especially when the dimension is as high as 19, is a quite complicated task. Main purpose of this report was to implement methods from different mathematical fields to analyse and model currency data. After doing that some conclusions on methods performance can be drawn.

First, and probably the most interesting part, is the elimination of triangular arbitrage from data using linear projection. The matrix representation of linear constraints can be used not only to transform the data to arbitrage free space, but also to identify arbitrage opportunities and their values. The arbitrages discovered in this work seem to be too small to make use of, due to transaction costs. However, currency trading is open 24 hours for five days a week and is continuous in time. It can be suggested to look into higher frequency data and apply the same methods. Arbitrages there could be big enough to make profit of. It might be aslo interesting to analyse residuals of projection.

Biggest difference between modelling univariate and multivariate time series is that in the univariate case one looks at partial autocorrelation or autocorrelation function for order indication and in multivariate case it is not possible. Order selection relies on tests and these differ in purposes and consistency. And even after deciding on which one to use, the estimated coefficients can produce an unstable model. In the end two out of three model order have been chosen just by reducing order from indicated one, until a stable process is obtained. Nevertheless, the estimated models seemed to replicate dynamics of time series quite well, especially for data with longer time steps. More complicated models like VARMA, multivariate ARCH or GARCH could potentially improve replication of currency rates dynamics and should be considered for future studies.

Finally, portfolio of min-ES was estimated using one time step forecasts. All three portfolios performed quite similarly. The fall in its value is due to general fall in USD value in world market. Portfolio of weekly reallocations has done best, as it made this fall smoother, while bi-weekly has done worst. This type of portfolio optimization can searve as suggestion for trader. Other types of portfolio choices could be considered, combined with other models for scenarios generation.

Bibliography

[1] *2010 Triennial Central Bank Survey.* Bank for International Settlements, 2010.

[2] Y. Aiba, N. Hatano, H. Takayasu, K. Marumo, and T. Shimizu. Triangular arbitrage as an interaction among foreign exchange rates. *Physica A*, 310:467–479, 2002.

[3] G.H. Golub and C.F. van Loan. *Matrix computations (3rd ed.).* JHU, 1996.

[4] Luetkepohl H. *New Introduction to Multiple Time Series Analysis.* Springer, 2005.

[5] M.R. King and D. Rime. *The $4 trillion question: what explains FX growth since the 2007 survey?* Bank for International Settlements, 2010.

[6] Alexander J. McNeil, Rdiger Frey, and Paul Embrechts. *Quantitative Risk Management: Concepts, Techniques, and Tools.* Princeton Series in Finance, 2005.

[7] R.T. Rockafellar and S. Uryasev. Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance*, 26:1443–1471, 2002.

[8] Axler S. *Linear Algebra Done Right (2nd ed.).* Springer, 1997.

[9] Bulirsch R. Stoer J. *Introduction to Numerical Analysis.* Springer, 2002.

[10] Ruey S. Tsay. *Analysis of Financial Time Series.* A Wiley-Interscience publication, 2002.