NATIONAL RESEARCH UNIVERSITY
HIGHER SCHOOL OF ECONOMICS

Aleksei Minabutdinov, Ilia Manaev, Maxim Bouev

FINDING THE NEAREST VALID COVARIANCE MATRIX: A FX MARKET CASE

BASIC RESEARCH PROGRAM

WORKING PAPERS

SERIES: Financial Economics
WP BRP 32/FE/2014

Aleksei Minabutdinov[*] *Ilia Manaev*[†] *Maxim Bouev*[‡]

FINDING THE NEAREST VALID COVARIANCE MATRIX: A FX MARKET CASE

We consider the problem of finding a valid covariance matrix in the FX market given an initial non-PSD estimate of such a matrix. The standard no-arbitrage assumption implies additional linear constraints on such matrices, which automatically makes them singular. As a result, one cannot just take the given estimate plug it into the standard optimization problem and solve it by applying even the most advanced numerical techniques developed recently. The reason is that such a problem is not well-posed while the PSD-solution is not strictly feasible. In order to deal with this issue, we described a low-dimensional face of the PSD cone that contains the feasible set. After projecting the initial problem onto this face, we come out with a reduced problem, which turns out to be well posed and of a smaller scale. We show that after solving the reduced problem the solution to the initial problem can be uniquely recovered in one step. We run numerous numerical experiments to compare performance of different algorithms in solving the reduced problem and to demonstrate the advantages of dealing with the reduced problem as opposed to the original one. The smaller scale of the reduced problem implies that virtually any numerical method can be applied effectively to find its solution.

JEL Classification: C02, C63, C65

AMS subject classification: 49M29, 90C25, 90C22, 93E24

Keywords: covariance matrix, correlation matrix, foreign exchange

[*] National Research University Higher School of Economics, Department of Mathematics, 55/2 Sedova ul., St.Petersburg 192171, Russia; Corresponding author

[†] Saint Petersburg State Polytechnical University, Institute of Computing and Control, 29 Polytechnicheskaya ul., St.Petersburg 195251, Russia

[‡] European University at St.Petersburg, Department of Economics, 3 Gagarinskaya ul., St.Petersburg 191187, Russia

1 Introduction

The problem of correct specification of an implied covariance matrix is one of prominent significance in finance. A most typical situation in practice is when a covariance matrix is constructed with missing elements being replaced by their estimates obtained by various *ad hoc* methods, for example, extrapolation or interpolation of historical data, etc. That might render such a matrix estimate not positively semidefinite (PSD) and prevent it from being used in option pricing, risk management or portfolio allocation models.[1]

The focus of this paper is finding a valid covariance matrix of logarithmic returns in the foreign exchange (FX) market under distributional assumptions of the Black-Merton-Scholes (BMS) model. The FX market is beneficially distinct from other financial markets in providing for a direct way of monitoring violation of no arbitrage conditions. Namely, for any values of two given FX rates absence of arbitrage may be verified by looking at the values of the so-called cross-rates. This leads to additional functional (linear) constraints on the covariance matrix. In particular, when all the cross exchange rates are liquidly traded, the no-arbitrage assumption implies a structure of the implied covariance matrix which is fully determined by volatilities of the respective exchange rates (Wystup (2007)). Violating that structure puts a trader in danger of selling a mispriced trade and cause his books to bleed the money (Austing (2011)).

The literature on computational finance abounds in methods of correction of invalid covariance matrix estimates. However, while being generally useful, those methods do not take into account the specific constraints on matrix elements like those encountered for the FX market. Moreover, little recognition is given to that fact, that in practice financial professionals, while working with incorrectly specified covariance matrices, may restrict themselves to correcting some submatrix of the complete covariance matrix.[2] This is often done when a trader is not interested in the elements of the complete covariance matrix, outside the given submatrix, as the trade, which the trader is dealing with, does not explicitly include a reference to those elements. In practice, traders might employ specifically written "trader's tools" to fill-in and update the values of volatilities or covariances within the data system they use, and only then move on to pricing actual products. On the one hand, when correction of the submatrix is separated from actual pricing, making a submatrix of FX rates (used in updating the system data) PSD does not necessarily imply that the complete covariance matrix (used in pricing) becomes PSD, which in turn creates opportunities for arbitrage. On the other hand, making a wrongly specified complete covariance matrix PSD may in some cases be a numerically daunting task because, as we show below, it will always be singular – the fact that adversely affects convergence of dual numerical methods, as well as primal-dual algorithms of interior-point methods, including those used in industry (see, eg Qi and Sun (2006), Todd (2001) and also Bouev (2012)).

In search for an efficient correction method for a covariance matrix estimate we put the problem for the FX market in a broader context of finding a symmetric PSD matrix nearest to the given estimate of the covariance matrix, when a set of affine restrictions

[1] Using such a matrix in financial modelling might not immediately lead to, for example, negative prices, but will undoubtedly incorporate possibilities for arbitrage from the very start.

[2] By such a matrix we henceforth understand a matrix corresponding to the complete set of assets traded in the market. In the FX case the complete matrix for the given set of currencies will include all the so-called driving FX rates (ie most liquid, or of primary importance to the trader) and their corresponding crosses.

on the resulting matrix must be satisfied. Such a problem will typically be of medium scale as the size of the matrix to be corrected will hardly be more than 100×100. [3] A number of approaches can be invoked for such matrices, in particular the so-called primal, interior-point and dual numerical methods. The latter ones, being in fact the most efficient for the purpose as suggested by recent literature on numerical methods, are not directly applicable to the problem in the focus of this work, as the specific constraints derived from no-arbitrage conditions in the FX market (Wystup (2007)) make the problem ill-posed in the sense that the sufficient condition for strong duality (Slater's) is not satisfied. We show, however, that this medium scale problem can always be reduced to an equivalent but well-posed small-scale problem, which can be solved by various methods in hundredths of a second. In particular, we suggest a procedure of isometric rescaling of the original problem down to a new one. The idea is to build an orthogonal projection onto a face of a PSD cone that contains a feasible set of solutions. Compared to the original problem, the reduced problem has a number of advantages: more specifically, it satisfies the sufficient condition for strong duality, it enjoys a significantly reduced dimensionality, and, finally, the FX market no-arbitrage conditions need no longer be enforced. We prove a one-to-one correspondence between the solutions of the original and reduced problems, and provide ample evidence on a superb performance of various numerical methods when applied to the reduced problem.

The paper is organized as follows. The next section brings in notation. After that, we introduce the reader to the problem, briefly discuss its origins and provide a short survey of possible approaches to finding its solution. Section 4 provides more detail on the distinct structure of the covariance matrix of logarithmic returns in the FX market. In Section 5 we first state the independence of the original optimization problem from quotations of FX rates – a result of practical importance. We then move on to state a well-posed problem equivalent to the original one. That section contains the main result of the whole work. In Section 6 we briefly describe advantages and disadvantages of existing numerical methods in the context of the concrete re-formulated problem. In Section 7 we compare method's performance in low dimensions. The last few sections are devoted to generalizations, certain issues of application and conclusions.

[3] This is so as usually there are only about 20 currency pairs, actively traded in the market (see TCBS (2013)).

2 Notation

$\mathcal{S}^n,\ \mathcal{S}^n_+$	Respectively, the linear space of symmetric $(n \times n)$-matrices and the cone of positive semidefinite matrices in $\mathcal{S}^n$.
$\|\|v\|\|$	The Euclidean norm of a vector-column v, defined as $\|\|v\|\|^2 = \sum_i v_i^2$.
$\|\|X\|\|_n$	The Frobenius norm of a $(n \times n)$-matrix X, defined as $\|\|X\|\|_n^2 = trace(XX^T) = \sum_{i,j} x_{ij}^2$.
$X \succ (\succeq) 0$	Eigenvalues of matrix X are positive (nonnegative), ie the matrix X is positive (semi-) definite.
I_M	The identity $(M \times M)$-matrix.
$\mathbb{O}_{r,s}$	The zero $(r \times s)$-matrix.
$P_{\mathcal{S}^n_+} X$	The orthogonal projection (Higham (1988)) of $X \in \mathcal{S}^n$ onto $\mathcal{S}^n_+$ which equals $X_+ = \sum_{\lambda_i > 0} \lambda_i q_i q_i^T$, where q_i is the eigenvector corresponding to eigenvalue λ_i.
$P_{\mathcal{K}} X$	The orthogonal projector of matrix $X \in \mathcal{S}^n$ onto the affine subspace $\mathcal{K}$.
$\mathrm{diag}(X)$	A vector-column of diagonal elements of matrix X.
$A[\tau, \mu]$	The $(r \times s)$-submatrix of A lying at the intersection of rows $\tau_1, \ldots, \tau_r$ and columns $\mu_1, \ldots, \mu_s$. Where $\tau = \{\tau_1, \ldots, \tau_r\},\ \mu = \{\mu_1, \ldots, \mu_s\}$ denotes two sets of distinct positive integers, not exceeding n and m, respectively.
$A\{\tau, \mu\}$	The $(n \times m)$-matrix such that $A\{\tau, \mu\}_{\tau_i \mu_j} = A_{\tau_i \mu_j}$ for all $1 \le i \le r$, $1 \le j \le s$, and other elements being equal to zero.
$v[\tau]$	Let v be a n-dimensional row-vector, then $v[\tau]$ denotes a r dimensional subvector of v, $(v_{\tau_1}, \ldots, v_{\tau_r})$.
$v\{\tau\}$	A n-dimensional vector, such that $v\{\tau\}_{\tau_i} = v_{\tau_i}$ for all $1 \le i \le r$, with other elements equal to zero.
(x, y)	The scalar product of vectors $x, y \in \mathbb{R}^n$, equals $\sum_{i=1}^n x_i y_i$.

3 The Problem and Its Background

We are interested in finding a valid estimate X of a true covariance matrix of logarithmic returns in the FX market under assumptions of the BMS model[4].[5] The no arbitrage conditions, also known as triangular relationships (see Wystup (2007) and the following Section), imply a set of functional constraints on the elements of X. Generally, when first constructed in practice, a given estimate C of the true covariance matrix might not satisfy that set of functional constraints, nor it has to be PSD by default for the reasons mentioned in the introduction. In addition, practical considerations might impose even further constraints on the desired covariance matrix X. Such constraints are often affine in nature, as happens in problems of matching the implied volatility surface, portfolio optimization, etc. A typical example is when in building an estimate of the covariance

[4]The BMS model is invoked here to describe the probability space only, the following discussion is, however, true for any instantaneous covariance matrix of logarithms of exchange rates under no-arbitrage assumptions.

[5]Some market participants may prefer dealing with a correlation matrix rather than a covariance matrix, for correlations are often readily observable in the market. The problem of finding a valid covariance matrix estimate X and the problem of finding a valid correlation matrix estimate R go hand in hand, since $R = D^{-1/2} X D^{-1/2}$, where D is a diagonal matrix and $D_{ii} = X_{ii},\ i = 1 \ldots M$. Thus, X and R are PSD or non-PSD simultaneously.

matrix one relies on several volatility estimates representing "*ex-post*" volatilities of well-chosen (eg replicating) portfolios. The resulting valid covariance matrix estimate must take the volatilities of the portfolios into account. Let, for example, ξ be a random vector of size M with zero mean, and let $F^T\xi$ represent some benchmark well-chosen portfolio. Suppose the variance of portfolio $F^T\xi$ is given and equals b. This implies an affine restriction on the covariance matrix (see Boyd and Xiao (2005), Malick (2005)):

$$trace[F^T\, E(\xi\xi^T)\, F] = trace[(FF^T)\, X] = trace[A\, X] = b \tag{1}$$

where $A = FF^T$ is a symmetric $(M \times M)$-matrix.

Another example of an affine restriction is enforcement of conservation of the empirical total risk, that can be expressed as $trace(I_M \cdot X) = trace(C)$, where I_M is the identity matrix of size $M \times M$ (see Malick (2005)). In general, altering the whole matrix C in a controlled fashion requires the matrix to satisfy some affine restrictions (see Schöttle and Werner (2004), Rebonato and Jäckel (2000)).

Taking the above reasoning into consideration, suggests the following problem of finding a valid covariance matrix X in the FX market. Let C be a non-PSD matrix of M currency pairs for given N currencies. The M currency pairs will contain all the "driving" currency pairs of primary interest to the modeller, as well as *some*, but not necessarily *all* crosses.[6] The modeller then seeks to

$$\begin{cases} \min \dfrac{1}{2}||X - C||_M^2 & (2) \\ X \in \mathcal{C}_{FX}^M & (3) \\ trace(A_i\, X) = b_i,\ 1 \le i \le k_{eq} & (4) \\ trace(B_j\, X) \ge d_j,\ 1 \le j \le k_{ineq} & (5) \\ X \succeq 0 & (6) \end{cases}$$

where restriction (3) implies that X is a symmetric matrix consistent with the no arbitrage condition in the FX market; $\mathcal{C}_{FX}^M$ is a space of such $(M \times M)$-matrices (see, in particular, Theorem 2 below)[7]; restriction (4) represents a general set of k_{eq} affine equality constraints; restriction (5) represents a general set of k_{ineq} affine inequality constraints if some lower or upper bounds on portfolio variances are specified; and, finally, restriction (6) implies that X is a PSD matrix.

In literature on constrained optimization problem (2) above is often referred to as the so called SDLS problem (see Malick (2005)) or the LSCAP problem (see Boyd and Xiao (2005)) or the LSCM problem (see Gao and Sun (2009)). Its geometric interpretation is computing a (Frobenius norm) projection of C onto the intersection of the half-space, defined by affine equalities (4) and inequalities (5), and the cone of positive semidefinite matrices (6). Finding a solution to such a problem has been a focus of extensive research. The existing approaches could be divided into primal, dual and interior point methods,

[6]The results of this paper hold for this more general statement of the problem. However, in any practical situation, whenever it is necessary to update the values of variances or covariances, the modeller will have to work with a complete covariance matrix. In other words, to avoid introducing arbitrage into the data system, any correction to variances and covariances of a given set of currency pairs necessitates a further updating of the covariances of all the crosses, spanned by the given set of currency pairs.

[7]If C does not immediately belong to $\mathcal{C}_{FX}^M$ then one first needs to project it onto $\mathcal{C}_{FX}^M$ as shown in Corollary 2 to Theorem 2.

which can also be distinguished by the type of admissible constraints. Here we provide only a glimpse of relevant work.

Higham (Higham (1988)) described a (Frobenius norm) projector[8] on the cone of positive semidefinite matrices in an optimization problem without imposition of additional constraints on matrix elements. In the paper by Rebonato and Jaeckel (Rebonato and Jäckel (2000)), well-known among financial market practitioners, the authors proposed two methods, one of which is too only applicable to the problem of finding a nearest correlation matrix without imposition of additional constraints, while the other provides only a sub-optimal solution to the problem (2). Additional affine equality constraints have also featured in Higham (Higham (2002)), where the solution is sought via an alternating projections method with the Dykstra's correction (Dykstra (1983)).

Problem (2) can also be reformulated as a standard semidefinite-quadratic-linear program (SQLP), ie a linear problem with affine equality-inequality constraints, the PSD cone constraint and some additional quadratic constraints. That would open the door to using interior point numerical methods, such as, for example, the primal-dual interior-point algorithm based on the Gauss-Newton approach proposed in Alfakih *et al* (1999). An application of similar methods can also be found in Schoettle and Werner (2004), while Higham (2002), Gao and Sun (2009) welcomed using popular software (such as SEDUMI, SDPT3 and PENNON), based on interior-point methods, for solving problems of convex and nonconvex nonlinear programming and semidefinite programming. The difficulty with interior point methods, however, is that they are generally feasible only as long as the dimension of the problem is not large. Moreover, the methods are not applicable for problems with affine inequality constraints.

Another way to solve problem (2) is to re-state it as a convex quadratic semidefinite program (QSDP) as in (Toh *et al* (2007)). They invoke preconditioned symmetric quasi-minimal residual iterative solver with appropriate preconditioners. That approach is indeed applicable in larger dimensions but, at the same time, is memory-expensive as implemented on a PC (see numerical experiments in Gao and Sun (2009)).

Perhaps, the most recent development in the problem of finding a nearest valid correlation or covariance matrix estimate in finance is the use of the so-called dual methods. These have been employed since the mid 2000s and are capable of properly accounting for inequality constraints in a natural way. Applied to a singular complete covariance matrix the methods also allow to substantially reduce the dimensionality of the problem. Malick (2005) and Boyd and Xiao (2005) proposed using such methods for solving the original problem (2) almost at the same time. Specifically, in Malick (2005) the use of a quasi-Newton BFGS method was suggested, while Boyd and Xiao (2005) advocated a projected gradient method. Building on developments in understanding strongly semismooth matrix-valued functions, a quadratically convergent Newton semismooth method was introduced in Qi and Sun (2006), while Borsdorf and Higham (2010) suggested some modifications to it. As the method in Qi and Sun (2006) does not by default allow inequality constraints, Gao and Sun (2009) proposed another modification, namely a quadratically convergent inexact smoothing Newton method, applicable to problems with affine inequality constraints.

The success of any dual method, however, hinges upon the existence of a bounded dual solution as well as the zero dual gap, which require the dual function to satisfy the coercivity property (see Rockafellar (1996)). Problem (2) in its original specification for the market complete with crosses does not, generally, satisfy that property even in the

[8]However, well-known in statistics.

simplest case of a (3×3)-matrix C. That is due to the matrix being singular, which essentially renders the dual methods impracticable for our purpose.

In order to resolve this challenging issue we shall show that problem (2) can be isometrically reduced to a low dimensional one by getting rid of the zero eigen values introduced by triangular relationships (see Section 4). The idea of the reduction is in projecting the original problem on a face of the PSD cone that contains a feasible set. In so doing we obtain a reduced problem which is a well posed LSCAP or LSCM problem (Wolkowicz (1981)):

$$\begin{cases} \min \frac{1}{2}||P_{11}^T C_{N-1} P_{11} - Y||_{N-1}^2 \\ trace(P_1^T A_i P_1 Y) = b_i,\ 1 \le i \le k_{eq} \\ trace(P_1^T B_j P_1 Y) \ge d_j,\ 1 \le j \le k_{ineq} \\ Y \succeq 0 \end{cases} \tag{7}$$

where C_{N-1} is a $(N-1 \times N-1)$-submatrix of the original matrix estimate C; C_{N-1} contains variances and covariances for the driving exchange rates**,** which the modeller is preoccupied with; matrix P_1 and its submatrix P_{11} depend on the actual choice of the exchange rates, that "drive" elements of C (see Theorem 6 below for more detail).

The reduced problem (7) has a number of advantages compared to problem (2), namely

- the matrix estimate featuring in the reduced problem has a significantly lower size;
- the arbitrage conditions in the FX market need no longer be explicitly enforced;
- it satisfies the Slater sufficient condition[9] for strong duality in a convex optimization problem, which allows one to employ the most efficient dual methods for finding a solution.

Below we shall prove a one to one correspondence between solutions of the reduced and original optimization problems. Then, we shall show that the reduced problem can be solved extremely efficiently, practically in hundredths of a second on a standard PC. The performance of various numerical methods will be demonstrated in numerous numerical experiments. Before that we, however, describe the specific structure of the FX market that defines the set of constraints $X \in \mathcal{C}_{FX}^M$ in the original problem (2).

4 The Structure of a FX Covariance Matrix

The FX market is distinct from other financial asset markets in that arbitrage, if exists, reveals itself and is observable in the values of cross-exchange rates. Thus, the no arbitrage condition implies restrictions on the exchange rates known as triangular relationships (Wystup (2007)). To illustrate the main idea we consider a simple example of a market with three currencies (a so-called triangular FX market), namely USD, EUR and GBP. All three currency pairs USD/EUR, USD/GBP and GBP/EUR are liquidly traded. The standard market convention is to quote them as $EUR-USD$, $GBP-USD$ and $EUR-GBP$, respectively. The no arbitrage assumption then yields:

$$S_t^{EUR-USD} = S_t^{EUR-GBP} S_t^{GBP-USD}$$

[9] under standard assumption of linear independence of matrices $A_i,\ 1 \le i \le k_{eq}$ and matrices $B_j,\ 1 \le j \le k_{ineq}$ admitting strict feasibility

where S_t is the exchange rate for a given pair of currencies at time t.

Let us index the exchange rates by $1, \ldots, M$. Then the triangular relationship will, in general, take the following form

$$(S_t^{(1)})^{\alpha}(S_t^{(2)})^{\beta}(S_t^{(3)})^{\gamma} = 1 \tag{8}$$

where $\alpha, \beta, \gamma \in \{-1, 1\}$.

In the above example of the triangular FX market, let us assign the USD/EUR pair index 1, USD/GBP - index 2 and EUR/GBP - index 3. Then it follows that $\alpha = 1$, $\beta = -1$ and $\gamma = -1$.

Equivalently, for logarithms of the exchange rates, equation (8) can be expressed as

$$\alpha \ln(S_t^{(1)}) + \beta \ln(S_t^{(2)}) + \gamma \ln(S_t^{(3)}) = 0 \tag{9}$$

Consider a FX market with N currencies. We shall assume that all the currency pairs are quoted according to the market convention. We shall also assume that the assumptions of the BMS model[10] hold, in other words, dynamics of the exchange rates are modeled as a log-normal martingale under its respective forward measure:

$$dS_t^{(i)} = S_t^{(i)}[\mu_i dt + \sigma_i dW_t^{(i)}] \tag{10}$$

where $i = 1 \ldots M$, μ_i is the rate of appreciation, σ_i is the instantaneous volatility of i-th exchange rate, M is the number of exchange rates in focus, where, obviously, $M \leq \frac{N(N-1)}{2}$. Correlations $\rho_{i,j}$ and covariances of logarithmic returns are defined by the covariance matrix with elements $\mathbf{cov}(\ln(S_t^{(i)}), \ln(S_t^{(j)})) = t\rho_{i,j}\sigma_i\sigma_j$.

Then the covariance matrix of the process $\{\ln(S_t^{(i)})\}_{i=1}^{M}$ equals $t \cdot \Sigma$, where

$$\Sigma = (\rho_{i,j}\sigma_i\sigma_j)_{i,j}$$

is the matrix of instantaneous variances and covariances[11].

Consider such a relationship for logarithmic returns $\alpha \ln(S_t^{(i)}) + \beta \ln(S_t^{(j)}) + \gamma \ln(S_t^{(k)}) = 0$, with $\alpha, \beta, \gamma \in \{-1, 1\}$ as before. By computing the covariance of the left hand side with any other logarithmic return m, and using the linearity property of covariance we get

$$\alpha\,\mathbf{cov}(\ln(S_t^{(i)}), \ln(S_t^{(m)})) + \beta\,\mathbf{cov}(\ln(S_t^{(j)}), \ln(S_t^{(m)})) + \gamma\,\mathbf{cov}(\ln(S_t^{(k)}), \ln(S_t^{(m)})) = 0 \tag{13}$$
$$\forall\ 1 \leq m \leq M$$

[10]The BMS model is invoked here to describe the probability space only. What follows is, however, true for any instantaneous covariance matrix of logarithmic returns under no-arbitrage assumptions.

[11]Its elements off the main diagonal will generally look like

$$\Sigma_{i,j} = \frac{(-1)^{\delta_{i/j}}(\sigma_i^2 + \sigma_j^2 - \sigma_{i/j}^2)}{2} \tag{11}$$

if the coefficient is for covariance of two currency pairs which have a currency in common, or, alternatively,

$$\Sigma_{i,j} = \frac{(-1)^{\delta_{i/j}}(\sigma_k^2 - \sigma_l^2 - \sigma_m^2 + \sigma_n^2)}{2} \tag{12}$$

where $\delta_{i/j}$ equals ± 1 depending on a particular quotation selected. These expressions can easily be obtained via linear combinations of equations (13), generally, irrespective of M. In other words, matrix Σ is fully determined by *all* $M = \binom{N}{2}$ implied volatilities $\{\sigma_i^2\}_{i=1}^{M}$ of log-returns (see Wystup (2001)).

Theorem 1. *1. Let N be the number of different currencies represented in the set of M currency pairs. Assume there is no arbitrage. Then there exists $N-1$ currency pairs (the so called driving pairs) which through triangular relationships define exchange rates of all other $M-N+1$ currency pairs.*

2. The rank of matrix Σ is no greater than $N-1$.

Proof. Consider any set of currency pairs where all N currencies $\{c_1, \ldots, c_N\}$ are represented. The minimal cardinality of the set is $N-1$ and it is attained for the set without any currency triangles. For instance, $\frac{c_1}{c_2}, .., \frac{c_1}{c_N}$. Then, using Ito's formula (see Chesney and Elliot, 1995) any other exchange rate will be defined by triangular relationships.

Equation (13), where $S_t^{(j)}$ and $S_t^{(k)}$ are two exchange rates from the selected set of currency pairs, implies that $M-N+1$ columns of matrix Σ are linearly dependent on the columns corresponding to exchange rates from the selected set. Thus $rank(\Sigma) = N-1$. □

For matrix Σ formula (13) can be rewritten as:

$$\alpha\Sigma_{i,m} + \beta\Sigma_{j,m} + \gamma\Sigma_{k,m} = 0 \qquad 1 \le m \le M \tag{14}$$

Therefore, vector $v^{(ijk)} = (v_1, \ldots, v_M) \in \mathbb{R}^{\mathbb{M}}$, such that $v_i = \alpha$, $v_j = \beta$, $v_k = \gamma$ with other elements being zero is the eigenvector of matrix Σ, corresponding to a zero eigenvalue. Now, denote the linear span of vectors $v^{(ijk)}$ by $\mathcal{N}_{FX}^M$. Theorem 1 implies that $\dim(\mathcal{N}_{FX}^M) + N - 1 = M$, then there are $M-N+1$ linearly independent vectors $v^{(ijk)}$.

Conversely if $\mathcal{N}_{FX}^M$ is a subspace of the null-space of a symmetric positive semi-definite matrix Σ, then $t\Sigma$ is a valid covariance matrix in the BMS world.

Let S_t be a random process with components satisfying (10), with μ_i such that at any moment $t \ge 0$ the equality $v^{(ijk)}\big(t(\mu - \frac{1}{2}\mathrm{diag}(\Sigma)) + \ln(S_0)\big) = 0$ holds for all $v^{(ijk)}$ from the basis of $\mathcal{N}_{FX}^M$. Assume, that the covariances of $\ln(S_t^{(i)})$ and $\ln(S_t^{(j)})$ for all i, j are given by matrix $t\Sigma$. Then it is enough to show that at any moment of time $t > 0$ the components of the process $\xi_t = \{\ln(S_t^{(i)})\}_{i=1}^M$, satisfy triangular relationships almost surely.

$$0 = v^{(ijk)}\Sigma v^{(ijk)^T} = v^{(ijk)}[E(\xi_t\xi_t^T) - E(\xi_t)\cdot E(\xi_t^T)]v^{(ijk)^T} = v^{(ijk)}E(\xi_t\xi_t^T)v^{(ijk)^T} =$$
$$= \quad E(v^{(ijk)}\xi_t)(v^{(ijk)}\xi_t)^T = E\Big[(v_i^{(ijk)}\xi_t^i + v_j^{(ijk)}\xi_t^j + v_k^{(ijk)}\xi_k)^2\Big]$$

Therefore, $v_i^{(ijk)}\xi_t^i + v_j^{(ijk)}\xi_t^j + v_k^{(ijk)}\xi_t^k = 0$ a.s. Thus, we have just proved the following theorem.

Theorem 2. *The following statements are equivalent:*

1. Matrix Σ defines an instantaneous covariance matrix in the BMS world.

2. Vectors $v^{(ijk)} \in \mathcal{N}_{FX}^M$ are the eigenvectors of positive-semidefinite matrix Σ, corresponding to a zero eigenvalue.

Theorem 2 reads that the essence of triangular relationships is in their imposing necessary restrictions (14) on matrix Σ.

Thus, for any matrix to be a valid covariance matrix in the BMS world two conditions must hold simultaneously: it is PSD and vectors $v^{(ijk)}$ are its eigenvectors corresponding to the zero eigenvalue. The space of matrices which satisfy only the second condition are denoted by $\mathcal{C}_{FX}^M$ (see, eg restriction (3) in (2)). Therefore the set of valid instantaneous covariance matrices in the BMS world is simply an intersection of the linear space $\mathcal{C}_{FX}^M$ and the cone $\mathcal{S}_+^M$ of PSD matrices.

The covariance matrix Σ is, in general, not full rank, hence the multivariate normal distribution of return increments is degenerate and does not have density with respect to M-dimensional Lebesgue measure. Thus, we can reduce the M-dimensional random process to no more than a $(N-1)$-dimensional one. This can be done using a change of variables in multiple ways. Here we consider two of them.

The first approach maps the $(M \times M)$-covariance matrix to its $(S \times S)$-submatrix $\Sigma[\tau, \tau]$ of the index set $\tau = \{\tau_1, \ldots, \tau_s\}$ assigned to the set of currency pairs which includes driving pairs. Matrix Σ may subsequently be restored from matrix $\Sigma[\tau, \tau]$.

Theorem 3. *For any M such that $S \leq M \leq \frac{N(N-1)}{2}$ and any set τ as described above, there exists a change-of-basis matrix Q_τ such that*

$$Q_\tau \Sigma\{\tau, \tau\} Q_\tau^T = \Sigma \tag{15}$$

Proof. Let $(\xi_t^1, \ldots, \xi_t^M) = \xi_t$ be a random $M-$ dimensional process satisfying (10) with covariance matrix $t\Sigma$. Let $\tilde{\xi}_t = \xi_t\{\tau\}$. Let matrix V be the $(M \times M)$-matrix, whose rows with indexes $k \in \{1, \ldots, M\} \backslash \tau$ are formed by linear independent vectors $(v^{(ijk)})$, where $i, j \in \tau$ and $v_k^{(ijk)} = -1$ and other elements being zero.

Consider an $(M \times M)$-matrix $Q_\tau = I\{\tau, \tau\} + V$, where I is the $(M \times M)$-identity matrix[12].

Then by Theorem 1 the random vector $\xi_t = Q_\tau \tilde{\xi}_t^{\,T}$ and formula (15) follow from the linear change-of-basis given above.

□

Unfortunately, such a matrix Q_τ is not an orthonormal change of basis, ie it does not preserve the Frobenius norm. Another, second approach thus may be considered, which is an isometrical projection in the Frobenius norm.

Theorem 4. *There exists the $(M \times N-1)$-matrix P_1 such that for any $(M \times M)$-matrix $C \in \mathcal{C}_{FX}^M$ the following decomposition holds*

$$K = P_1^T C P_1 \tag{16}$$

where K is the $(N-1 \times N-1)$-matrix, and, importantly, $||C|| = ||K||$. Conversely, given the matrix K and the matrix P_1 one can restore the matrix C by $C = P_1 K P_1^T$.

Proof. Let P be a matrix, which $M - N + 1$ first columns form an orthonormal basis in $\mathcal{N}_{FX}^M$, and the last $N-1$ columns form an orthonormal basis in the orthogonal complement of $\mathcal{N}_{FX}^M$. Since $\mathcal{N}_{FX}^M$ is a subset of the null-space of matrix C, the first $M - N + 1$ rows and columns of $P^T C P$ contain zeros only.

$$P^T C P = \begin{pmatrix} \mathbb{O}_{r,r} & \mathbb{O}_{r,s} \\ \mathbb{O}_{s,r} & K \end{pmatrix}, \text{ where } r = M - N + 1,\ s = N - 1. \tag{17}$$

Since matrix P is an orthonormal change of the basis matrix, it does not affect the eigenvalues of matrix C, in particular any eigenvalue λ_i, $(1 \leq i \leq N-1)$ of matrix K is eigenvalue of matrix C and $||C||_M^2 = ||K||_{N-1}^2 = \sum\limits_{1 \leq i \leq N-1} \lambda_i^2$.

Denote by P_1 the $(M \times N-1)$-submatrix consisting of the last $N-1$ columns of matrix P, ie

$$P_1 = P[1 \ldots M, M - N + 2 \ldots M].$$

[12] Obviously, Q_τ is invertible.

Then using (17), $P_1^T C P_1 = K$ and, conversely, $P_1 K P_1^T = C$.

□

Theorem 3 implies that matrix K (from Theorem 4) can be obtained from any submatrix $\Sigma[\tau,\tau]$ of matrix Σ, if the set of indices $\tau = \{\tau_1, \ldots, \tau_s\}$ is assigned to the set of currency pairs which contains any set of driving pairs.

Corollary 1. *For any M such that $S \leq M \leq \frac{N(N-1)}{2}$ and any set τ described above there exist matrices $\Sigma[\tau,\tau]$ and $P_1[\tau,\mu]$ such that:*

$$K = P_1[\tau,\mu]^T \Sigma[\tau,\tau] P_1[\tau,\mu] \tag{18}$$

where μ runs through all the indices from 1 to $N-1$.

Proof. Using the notation of the Theorem 3 we obtain:

$$Q_\tau P_1 = (I\{\tau,\tau\} + V)P_1 = I\{\tau,\tau\}P_1 + VP_1 = I\{\tau,\tau\}P_1 = P_1\{\tau,\mu\},$$

so

$$K = P_1^T \Sigma P_1 = P_1^T Q_\tau^T \Sigma\{\tau,\tau\} Q_\tau P_1 = P_1\{\tau,\mu\}^T \Sigma\{\tau,\tau\} P_1\{\tau,\mu\} = P_1[\tau,\mu]^T \Sigma[\tau,\tau] P_1[\tau,\mu].$$

□

Remark 1. *It is important to note the following result. Theorem 3 explicitly, and Corollary 1 implicitly assume positive-semidefiniteness of matrix Σ. However, once proven for matrices belonging to the intersection of cone $\mathcal{S}_+^M$ and linear subspace $\mathcal{C}_{FX}^M$, the theorem and the corollary can be extended to the whole space $\mathcal{C}_{FX}^M$.*

Now let us consider some matrix C that is a candidate for being a valid covariance matrix. For example, it can be an empirical estimate of the true covariance matrix.

Corollary 2. *Let C be a $(M \times M)$-matrix, then the projection of C onto $\mathcal{C}_{FX}^M$ equals*

$$P_{\mathcal{C}_{FX}^M}(C) = P_1(P_1^T C P_1)P_1^T \tag{19}$$

Proof. We need to verify that $P_{\mathcal{C}_{FX}^M}(C)$ belongs to $\mathcal{C}_{FX}^M$ and that $P^2_{\mathcal{C}_{FX}^M}(C)$ equals $P_{\mathcal{C}_{FX}^M}(C)$
Both of them follow directly from Theorem 4.

□

Summarizing the above arguments, matrix C is a valid instantaneous covariance matrix in the BMS world if and only if expression (19) holds as identity

$$P_{\mathcal{C}_{FX}^M}(C) = C$$

and the corresponding matrix C is PSD. Matrices C and $K = P_1^T C P_1$ are either PSD or non-PSD simultaneously, so it is sufficient to check the PSD property for a $(N-1 \times N-1)$-matrix K.

Thus, we have discussed both the PSD and FX market constraints, that have to be satisfied for matrix C to be a valid covariance matrix in the BMS world. Now we are ready to continue with finding a solution to problem (2).

5 Correctness: An Equivalent Well-Posed Problem

Without loss of generality let us suppose that the given estimate C of the true covariance matrix satisfies the no-arbitrage conditions in the FX market. Otherwise, using Corollary 2 we can substitute matrix C for the nearest to it matrix $P_{\mathcal{C}^M_{FX}}(C)$, which does satisfy the no arbitrage conditions by construction.

The structure of the covariance matrix definitely depends on the way that currency pairs are quoted. Since our goal is to find a PSD matrix nearest to the given one, it is vital to ensure correctness of the problem. Namely, we need to ensure that the PSD property of matrix C is independent of the chosen quotation.

Theorem 5. *The PSD property of matrix C is independent of the chosen quotation.*

Proof. see Appendix B. □

Assume that the restrictions in problem (2) do not contradict positive semi-definiteness of the matrix, ie the feasible set is nonempty. Since the problem is a strictly convex optimization problem with a bounded feasible set it has a unique solution X^* (see Rockafellar (1996)).

First of all, notice that Theorem 2 implies that restriction $X \in \mathcal{C}^M_{FX}$ is linear. Hence, it can be rewritten in a general affine form. For example, take v_i as any basis vector in $\mathcal{N}^M_{FX}$, then

$$trace(G_i X) = 0, \text{ where } M \times M\text{-matrices } G_i = v_i^T v_i, \ i = 1, \dots, M - N + 1 \tag{20}$$

Then problem (2) can equivalently be reformulated as

$$\begin{cases} \min \frac{1}{2}||X - C||^2_M \\ trace(G_k X) = 0, \ k = 1, \dots, M - N + 1 \\ trace(A_i\, X) = b_i, \ 1 \le i \le k_{eq} \\ trace(B_j\, X) \ge d_j, \ 1 \le j \le k_{ineq} \\ X \succeq 0 \end{cases} \tag{21}$$

Despite problem (21) being an ordinary strictly convex optimization problem, it does not satisfy the Slater condition when the size of the matrix exceeds the number of driving pairs (see Appendix C and also Gao and Sun (2009), Malick (2005)):

$$\begin{cases} \{A_i\}, \{G_k\} \text{ are linearly independent} \\ \text{There exists a feasible } X^0 \text{ such that } trace(B_j\, X^0) > d_j \text{ and } X^0 \succ 0 \end{cases} \tag{22}$$

This is because Theorem 2 implies that in the feasible set any matrix of such a size is singular.

As we discuss in the following section, the most efficient numerical methods for solving convex optimization problems are the so-called Newton-type methods. Generally, they converge if the Slater condition is satisfied. Since in our case the condition is violated, problem (21) should be modified in such a way that it becomes well-posed.

In order to resolve this issue, we suggest a procedure of isometric rescaling of the original problem down to a new one:

$$\begin{cases} \min \frac{1}{2}||K-Y||^2_{N-1} \\ trace(\hat{A}_i Y) = b_i,\ 1 \le i \le k_{eq} \\ trace(\hat{B}_j Y) \ge d_j,\ 1 \le j \le k_{ineq} \\ Y \succeq 0 \end{cases} \tag{23}$$

where $\hat{A}_i = P_1^T A_i P_1$, $\hat{B}_j = P_1^T B_j P_1$, $K = P_1^T C P_1$ and P_1 is as given in Theorem 4.

Remark 2. *Conversely unreduced problem* (21) *can be obtained from reduced problem* (23) *by multiplication of all matrices* $K, \hat{A}_i, \hat{B}_j$ *from the left by the matrix* P_1 *and from the right by the matrix* P_1^T *and imposing constraints* (20)*. Obviously transitions from the original problem to the reduced one and back are equivalent in terms of computational costs involved.*

According to Corollary 1 and the Remark 1 matrix K can be written as $P_1[\tau,\mu]^T C[\tau,\tau] P_1[\tau,\mu]$, where $\tau = \{\tau_1, \ldots, \tau_S\}$ contains indices assigned to some driving currency pairs and μ runs through all the indices from 1 to $N-1$. Therefore, in order to state problem (23) one only needs to know some submatrix $C[\tau,\tau]$, corresponding to those driving currency pairs marked with $\{\tau_1, \ldots, \tau_S\}$.

The following theorem postulating the equivalence of the original and reduced problems is the main result of this paper.

Theorem 6. *For any choice of set* τ*, containing the indices assigned to driving currency pairs, solution* X^* *to problem* (2) *and solution* Y^* *to problem* (23) *are related as:*

$$X^* = P_1 Y^* P_1^T$$

Proof. Since the FX market constraints are linear and both $C, X \in \mathcal{C}^M_{FX}$, matrix $X - C$ is also in $\mathcal{C}^M_{FX}$. Theorem 4, Corollary 1 and the Remark 1 therein state that the norm of transformation does not change.

$$\begin{gathered} ||C-X||^2_M = ||P_1^T Q_\tau^T C\{\tau,\tau\} Q_\tau P_1 - P_1^T X P_1||^2_{N-1} = \\ = ||P_1[\tau,\mu]^T C[\tau,\tau] P_1[\tau,\mu] - Y||^2_{N-1} = ||K-Y||^2_{N-1}, \end{gathered}$$

where μ runs through all the indices from 1 to $N-1$. The following equalities hold:

$$\begin{gathered} trace(\hat{A}_i Y) = trace(P_1^T A_i P_1 Y) = trace(P_1^T A_i P_1 P_1^T X P_1) = \\ trace(P^T A_i P P^T X P) = trace(A_i X) = b_i. \end{gathered}$$

By analogy, for inequalities we have: $trace(\hat{B}_j Y) = trace(B_j X) \ge d_j$.

Therefore the affine constraints of problem (2) correspond to the affine constraints of problem (23).

Finally, denote the solution to problem (23) by Y^*. Then, by Theorem 4, the solution to the initial problem (2) is given by $X^* = P_1 Y^* P_1^T$.

□

The principal algorithm of using the Theorem 6 is given below in Appendix D. We provide an example of use of Theorem 6 in Appendix E, where we implement step by step the algorithm from the Appendix D for the market data.

6 Numerical Methods

In this section we briefly describe existing methods of finding nearest covariance matrix for different sets of equality and inequality constraints in the context of the re-formulated problem (23). We start off with the simplest case of no additional affine constraints.

6.1 A problem without additional affine constraints

It is well known (see Higham (1988)) that in the case of no additional constraints the projector $Proj_{\mathcal{S}_+^M}$ onto the cone of positive semi-definite matrices gives the optimal solution to the problem

$$\min_{X \in \mathcal{S}_+^M} ||X - C||_M$$

This projector simply returns the positive semi-definite part C_+ of matrix C and has the explicit form:

$$Proj_{\mathcal{S}_+^M} C = C_+ = Q\,\Lambda\,Q^{-1}$$

where columns of matrix Q is a set of orthonormal eigenvectors of C corresponding to eigenvalues $\lambda_1, \ldots, \lambda_M$ and Λ is a diagonal $(M \times M)$-matrix with $\Lambda_{ii} = \max\{0, \lambda_i\}$.

Considerations above clearly signify that the projector has exactly the same form also in the case when the FX market constraints (3) are taken into account

$$Proj_{\mathcal{S}_+^M} = Proj_{\mathcal{S}_+^M \cap \mathcal{C}_{FX}^M}$$

Thus, this implies that the solution in the case of no additional affine constraints is[13] $X^* = C_+ = P_1 K_+ P_1^T$.

6.2 A general problem

In the case of additional affine constraints

$$trace(\hat{A}_i Z) = b_i,\ i = 1 \ldots k_{eq}$$
$$trace(\hat{B}_j Z) \geq d_j,\ j = 1 \ldots k_{ineq}$$

in order to get the solution one needs to invoke a projection onto the intersection $K \cap \mathcal{S}_+^n$ of the PSD cone $\mathcal{S}_+^n$ and the affine cone K defined by the equalities and the inequalities above. Generally, however, the explicit expression for this projector is unknown. Thus, it is necessary to apply numerical methods. A number of researchers has paid a lot of attention to this very general problem. Methods thereby suggested can be divided into primal, dual and interior-point ones, as well as split by the types of admissible constraints.

It is not our goal to describe here various methods in detail. A good recent overview could be found in Malick (Henrion and Malick (2012)). We instead focus only on the principal differences between the methods, as well as the conditions that have a direct impact on the speed of convergence of these methods.

[13] Since for complete FX $(M \times M)$-matrix C we have $M = \frac{N(N-1)}{2}$ and computational complexity of eigendecomposition of $(n \times n)$-matrix is more than $O(n^2)$, computation of $K_+ = Proj_{\mathcal{S}_+^{N-1}} K$ and two matrix-matrix multiplications is generally more efficient and accurate, than computation of $C_+ = Proj_{\mathcal{S}_+^M} C$.

6.2.1 A primal algorithm

In order to solve problem (23) Higham suggested using an **alternating projection method with Dykstra's correction step** (see Higham (2002)). Despite its being designed for the correction of a correlation matrix, it can straightforwardly be modified for the problem of finding the nearest valid covariance matrix too. The main idea of the method is rather intuitive and consists in creating an alternating projection onto positive definite cone $\mathcal{S}_+^n$ and an affine hyperplane defined by affine (equality) constraint $\mathcal{K}$.

While the projection on the positive definite cone $P_{\mathcal{S}_+^n}$ was already described above, the projection $P_{\mathcal{K}}(X)$ onto the set defined by the affine constraints is, in fact, a standard quadratic programming problem. In the case of equality constraints alone it has an explicit expression for the solution (Rockafellar (1996)). In order to ensure convergence of the projection of matrix C to an optimal (not just feasible) point, a so-called Dykstra's correction step should be added.

The well-known properties speaking for and/or against the use of this algorithm for our purpose are:

1. The pros:
   (a) It is relatively simple to implement.
   (b) It could be applied directly to problem (2).
2. The cons:
   (a) It converges linearly[14].
   (b) If there is at least one affine inequality constraint (5) in the original problem then the method is not applicable (at least in its original implementation).

For details of algorithm implementation refer to (Higham (2002), Henrion and Malick (2012)).

6.2.2 Dual algorithms

The dual function θ is computed from the Lagrangian of the problem after minimizing in primal variable X. After some simplification we shall get

$$\theta(y) = -\frac{1}{2}||(C + \sum A_i y_i)_+||_M^2 + \langle b, y\rangle + \frac{1}{2}||C||_M^2$$

Function $\theta(y)$ is sufficiently well-behaved, namely it is concave and continuously differentiable (see Th3.2 in Malick (2005)). Its gradient mapping $\nabla\theta(y)$, given by

$$\nabla\theta(y) = \begin{bmatrix} b_1 - trace\,(A_1 \cdot (C + \sum A_j y_j)_+) \\ \vdots \\ b_k - trace\,(A_m \cdot (C + \sum A_j y_j)_+) \end{bmatrix} \tag{24}$$

is globally Lipschitz continuous with the Lipschitz constant equal to 1. Having said that, it may not have a second derivative, however.

Let $Q_+ = \mathbb{R}^{k_{eq}} \times \mathbb{R}_+^{k_{ineq}}$.The dual problem to problem (21) takes the form

[14] In the sense of an assymptotically exponential decline of the residue.

$$\begin{cases} \theta(y) \to \max \\ y \in Q_+ \end{cases} \tag{25}$$

To solve problem (25) one needs to solve the equation below (see Eaves (1971))

$$F(y) = 0 \tag{26}$$

where $F(y) = y - P_{Q_+}(y + \nabla\theta(y))$. Note that dual variable y just belongs to $Q_+ \subset \mathbb{R}^k$, while primal variable $X \in \mathbb{R}^M$. Therefore it is enough to solve relatively easy k-dimensional dual problem in the space $\mathbb{R}^k$ instead of the $\frac{M(M-1)}{2}$ dimensional primal problem in the space $\mathcal{S}^M$.

If the problem is well-posed, then there exists a solution $y^* \in Q_+$ to the equation (26). By Theorem 4.1 from (Malick (2005)), the solution is given by $Z^* = (C + \sum A_i y_i^*)_+$.One needs the so-called coercivity property, however, to provide for the existence of a (bounded) solution of dual problem $y^* \in Q_+$. The primal Slater condition (22) implies this property (see Gao and Sun (2009)). We shall provide more details on that in the numerical section below, but now we just note that the only difference between various dual methods is, in fact, in the method of solving equation (26).

For example, Boyd and Xiao (2005) proposed a **projected gradient method**. This method is a dual variant of the Dykstra's correction step algorithm above. For details see their work as well as Henrion and Malick (2012). The pros and cons of that algorithm are:

1. The pros:
   (a) It is easy to implement (in fact, it can be implemented in less than 10 lines of pseudocode).
   (b) It is readily applicable for the case with inequality affine constraints.
2. The cons:
   (a) Its convergence rate is linear.

In Malick (2005) a **quasi-Newton BFGS method** is proposed to solve the dual problem (25). For details the reader should refer to the original work as well as Chen (1995), while here we again just give the pros and cons of the algorithm:

1. The pros:
   (a) It is applicable for problems with inequality affine constraints(5).
2. The cons:
   (a) It is more expensive in terms of memory use than other methods (see numerical experiments Section in Gao and Sun (2009)).
   (b) In practice it may be difficult to implement when tolerance is rather low ($< 10^{-8}$).
   (c) Its convergence rate is at best quadratic, but in practice is often linear (see Qi and Sun (2006)).

Qi and Sun (2006) suggested a quadratically convergent **Newton semismooth method** for the case when inequality constraints are not present. The method builds on strong semismoothness of eigenvalues (see Sun and Sun (2002), Chen *et al* (2002)) and a generalization of the Jacobian matrix, namely the Clarke's generalized Jacobian in the sense of Sun and Sun (2002). As before, the pros and cons are:

1. The pros:
   (a) The algorithm is quadratically convergent.
2. The cons:
   (a) It is not suitable for problems with inequality affine constraints (5).
   (b) It requires strong non-degeneracy of the constraints for quadratic convergence.

Borsdorf and Higham put forward a number of modifications to the previous method to increase its performance (see Borsdorf and Higham (2010)). Finally, in Gao and Sun (2009) a quadratically convergent **inexact smoothing Newton method** is considered. The main idea is smoothing the function F by functions $G(\varepsilon, y)$, $G(\varepsilon, y) \to F(y)$, $\varepsilon \to 0$, and solving a smoothed-nonsmooth system of equations $\begin{bmatrix} \varepsilon \\ G(\varepsilon, y) \end{bmatrix} = 0$.

The algorithm advantages and disadvantages are as follows:

1. The pros:
   (a) It is quadratically convergent.
   (b) It is applicable for problems with inequality affine constraints (5).
2. The cons:
   (a) It requires strong non-degeneracy of the constraints for quadratic convergence.

6.3 Interior-point algorithms

In general, two approaches can be followed as far as the interior-point methods are concerned.

A standard approach

Mathematically, problem (23) can be equivalently re-stated (see Higham (2002), Gao and Sun (2009)) as a linear optimization problem with affine equality and inequality constraints, quadratic constraints, and PSD cone constraints by pushing down the objective function with an additional variable t as follows:

$$\begin{cases} \min t \\ trace(A_i X) = b_i,\ 1 \le i \le k_{eq} \\ trace(B_j X) \ge d_j,\ 1 \le j \le k_{ineq} \\ t + 1 \ge \sqrt{(t-1)^2 + 2||X - C||^2} \\ X \succeq 0 \end{cases}$$

Then this problem can readily be passed over to many SDP solvers that use interior-point methods and are implemented in well-developed and publicly available software such as SeDuMi (Sturm (1999)) or SDPT3 (Toh *et al* (2007)). Unfortunately, standard primal-dual interior-point methods are prohibitively expensive in terms of memory usage and time required to find a solution. The reason is that at each iteration a linear system of equations of dimension more than $\frac{1}{2}M^2$ needs to be solved to find the direction of search at the following step. All in all, the pros and cons are:

1. The pros:

   (a) A publicly available software can be used.

2. The cons:

   (a) The methods are computationally inefficient and applicable for low-dimensional problems only[15].

An alternative approach

In Toh *et al* (2007) the authors proposed a primal-dual path-following Mehrotra-type predictor-corrector method. They solved the linear system of equations determining the search direction at the next step in a much more efficient way using a preconditioned symmetric quasi-minimal residual (PSQMR) iterative solver with appropriately constructed preconditioners. Their method's pros and cons are:

1. The pros:

   (a) Its implementation is publicly available.

   (b) For large-scale matrices it can compete with latest dual methods in terms of speed and efficiency.

2. The cons:

   (a) It is inapplicable for problems with inequality affine constraints (5).

   (b) In our experiments (see below) the method demonstrated relatively low tolerance.

   (c) The stopping criteria for this method are conditional, ie tolerance will depend on the matrix size.

7 Comparison of Methods in Solving Small-Scale Problems

This section is devoted to comparison of different algorithms, producing a solution to (2) and (23). Despite there being a general consensus in literature on superior performance of Newton-type methods in solving large scale problems (see the references below), to our knowledge no comparison has still been drawn for relative efficiency of numerical methods in solving small and medium scale problems. At the same time, Theorem 6 above shows

[15]Generally, for matrices of size less than 80×80 on a standard PC.

that the problem of finding a nearest valid estimate of the covariance matrix in the FX market with $N-1 \leq M \leq \frac{N(N-1)}{2}$ currency pairs can be reduced to a low-dimensional problem of size $N-1$. Since the overall number of currencies in the world is about 200, of which about 20 currencies are liquidly traded (see, eg TCBS (2013)), the reduced problem is usually a low-dimensional one. It is then interesting to ascertain quantitatively the gains of scaling the problem down. First, however, we would like to look at the original problem (2) and illustrate the consequences of lack of coercivity in it.

7.1 Solving the initial (unreduced) problem

From the previous section we know that the unreduced problem (2) can be directly tackled by primal methods[16]. Those, on the one hand, are less efficient than dual methods[17], and, on the other hand, converge rather slowly (at a linear rate). That becomes a serious shortcoming when dimensionality of the problem becomes too high. Moreover, some primal methods (like the alternating projections algorithm with Dykstra's correction used below) are capable of working only with a certain type of additional constraint, eg equality constraints. The more advanced dual methods, however, despite having been developed more recently are not duly applicable to the unreduced problem. Since the function $\theta(y)$ is not coercive, the finite solution of the equation (26) does not exist (see Appendix C for more detail).

This observation implies that in order to improve on the performance of numerical methods in tackling the problem typical of the FX market, one has to either reduce the dimensionality of the optimization problem, or transform the problem in such a way that it starts to satisfy the coercivity condition. In an ideal world both things would be achieved at once. That would *inter alia* open the door to using dual algorithms.

7.2 Solving the reduced problem

Thereby problem (2) should be scaled down. Obviously, one would want to see the effect of such a transformation of the original problem into the reduced form as well as the computational costs involved. We shall now demonstrate these using a number of numerical methods mentioned above.

7.2.1 Comparing method's performance

Formula (23) and Remark 2 show the relationship between the original unreduced and reduced problems. By relying on its premise, without loss of generality one can simply start working with matrices of size $N-1 \times N-1$ and obtaining unreduced problem when necessary.

We start by randomly generating non-PSD matrices. We consider here markets with $N = 5$, 7, 13 and 21 currencies. For each market we generate 100 matrices, which implies 400 matrices in total. Each matrix has elements from $[-1, 1]$ and is generated under (conditional) uniform distribution with 30% of eigenvalues being less than -0.001.

[16] As well as primal realization of interior point methods, but it is considered to be less efficient than interior point methods invoking primal-dual path following algorithms mentioned above, see for example, Todd (2001), therefore we don't consider them here.

[17] At least Newton-type methods, see convincing numerical experiments in Gao and Sun (2009) and Qi (2013).

To keep things simpler, assume for the moment that there are no inequality constraints at all.[18]

No inequality constraints. As regards the number of the additional affine equality constraints we, somewhat arbitrarily, choose it to be a quarter of the number of the degrees of freedom which any optimization problem of the type we consider here will have. In particular, that number of the degrees of freedom in a symmetric $(N-1 \times N-1)$-matrix equals $\frac{1}{2}N(N-1)$. Thus, specifying more than $\frac{1}{2}N(N-1)$ independent equality constraints will imply emptiness of the feasible set. Hence, we shall consider problems with $k_{eq} = \frac{1}{8}N(N-1)$ equality constraints that is a quarter of total number of degrees of freedom (which we believe is a realistic scenario in practical applications of the methods[19]). For equality constraints

$$trace(A_i X) = b_i,\ 1 \leq i \leq k_{eq},$$

we take A_i to be a random symmetric $(N-1 \times N-1)$-matrix consisting of zeros and ones. All such matrices are chosen to be linearly independent. The numbers b_i are chosen at random in such a way that on average 30% change of elements c_{ij} of matrix C is needed to satisfy these constraints.

In order to test the feasibility of the problem with given constraints we employ the algorithm suggested by Henrion and Malick (2007). The idea of the algorithm, is that in case of infeasibility, one can check Farkas' lemma, ie find numerically Farkas' dual vector. If the problem turns out to be infeasible, we generate new constraints. In general, knowing in advance which constraints enjoy the condition of constraint nondegeneracy (ie the sufficient condition for quadratic convergence of Newton-type methods) is very difficult. This is because the validity of the condition often becomes clear only *a posteriori*, when the solution to the maximization problem itself has been found ((Gao and Sun, 2009, p. 18), formula (54)). In the tests that we discuss here we came across the issue only a few times.

All the tests were coded in a 32-bit version of MATLAB R2011b and run on a Windows 7 desktop powered by Intel Core 2 Duo CPU of 3.00 GHz, with 6.0 GB of RAM.

The results of performance testing for the case with affine equality constraints alone are provided in Table 1 from Appendix A.

First of all it presents results for an unreduced problem (ie for unreduced $(\frac{N(N-1)}{2} \times \frac{N(N-1)}{2})$-matrix) obtained as in Remark 2. The only algorithm capable of solving the unreduced problem is the primal alternating projections method with Dykstra's correction step (referred as `dykstra`). Since transitions from the original problem to the reduced one and back are equivalent in terms of computational costs involved, we simply measure the time required to convert the reduced problem to unreduced one.

For the reduced problem we test for the following methods: Boyd and Xiao (2005) projected gradient method (referred to as `PGM` in the table), Malick (2005) quasi-Newton method (referred to as `BFGS`), the Newton semismooth method by Qi and Sun (2006)(referred to as `newt`), the modified version of the Newton semismooth method by Borsdorf

[18] Of the algorithms considered previously the only one capable of solving the initial (unreduced) problem is the primal alternating projections method with Dykstra's correction step. It can handle only *equality* constraints. Thus, considering the case without inequality constraints renders possible the direct comparison of the method's performance in solving the initial and reduced problems. For more details see this section below.

[19] As well such choice gives no "low number of constraints" advantage for the dual methods

and Higham (2010) (referred to as `newt_bh`), the infeasible path-following algorithm implemented by Tütüncü *et al* (2003)) in the SDPT3 solver[20], and, finally, the inexact interior-point method of the same authors implemented in the QSDP solver[21]. The PGM method was implemented as in (Boyd and Xiao (2005)), the Newton semismooth method was modified for general affine constraints on the basis of the source code provided by Qi and Sun (2006)[22]. For the quasi Newton method we used the implementation of the BFGS method for the dual problem, written by Henrion and Malick (2007).

The primal alternating projections method applied for the reduced problem is omitted since it is actually the PGM[23] for the dual problem with step size 1 (see Henrion and Malick (2012), Gao and Sun (2009)).

Also we do not present results for the inexact smoothing Newton method (referred below as `newt_gs`), since its performance, in the case of experiments with equality constraints only, seems coincide with performance of the Newton semismooth method.

We stop the run of dual algorithms[24] when $Res = ||\nabla\theta(y)|| < tol$, with $tol = 10^{-8}$. As for QSDP and SDPT3 solvers it should be noted that the residue used there is unfortunately relative rather than an absolute one (see Tütüncü *et al* (2003), Toh *et al* (2007)). As a result, given the same tolerance, the answer is usually less accurate than in the case of other methods. Therefore following Gao and Sun (2009) we use lower values of tolerance and put to stop those solvers when the relative gap (see Toh (2007)) was less than 10^{-9}. As for the other settings, we used the default parameter values. For primal alternating projections method with Dykstra's correction step we used relative stopping condition, that was initially suggested in Higham (2002)[25]

$$\frac{||X_k - X_{k-1}||}{||X_k||} \leq tol = 10^{-8}$$

where X_k stands for a matrix obtained at the k-th step. We find that this criterion gives in fact similar level of accuracy as the one used in QSDP and SDPT3 solvers.

All that matters for a practitioner in algorithm's performance is the time that CPU is engaged (`Time,` measured in seconds (s) or milliseconds (ms)) and the algorithm's residual (`Res`). However, time is closely linked to the number of iterations (`iter`) needed for the method to converge for the given tolerance parameter. However it is less dependent on particular implementation. Due to restrictions on space imposed by the format of the article we present here only the median figures. Nevertheless, they give the reader a complete idea of the comparative performance of the algorithms.

In the Table 1 in Appendix A, the median values of `Time`, `Res` and `Iter` obtained in our experiments are shown. Also it shows matrix sizes on which tests were carried on (`unred.size` for unreduced problem and `red.size` for reduced problem) as well as median time in milliseconds (referred as `conv.time`) needed for each problem to be

[20]The SDPT3 solver is availbale at `http://www.math.nus.edu.sg/ mattohkc/sdpt3.html`

[21]The QSDP solver is availbale at `http://www.math.nus.edu.sg/ mattohkc/qsdp.html`

[22]See http://www.math.nus.edu.sg/ matsundf

[23]under relevant constraints considered here

[24]The BFGS method uses inexact line search for finding appropriate step size. While the step size must satisfy the so called weak Wolfe's condition. To verify this condition numerically, we used the same settings as in Henrion and Malick (2007). Nevertheless in some cases this condition was violated and the method terminated earlier, namely, when Res reached values of the order 10^{-7}. In these cases, we were leaving the result unchanged, although note that in the case of using this algorithm for finding solution with higher accuracy, line search may require more fine-tuned parameters.

[25] $||\cdot||$ here stands for the matrix 1-norm

scaled from unreduced to reduced form (or *vice versa*). The column `restr.` stands for the number of equality constraints used for the problem of each size.

An important observation from (see Appendix A, Table 1) is that solving of unreduced problem is inefficient as compared with *any* dual method once number of currencies N is greater then 9 or 10. Since primal alternating projections method applied for the reduced problem is actually the PGM method to see the effect of transformation of the original problem into the reduced form it is instructive to compare firstly these two methods. For example, `dykstra` takes on average more than 3 minutes to solve 210×210 problem (ie N= 21) while PGM needed about half of a second to solve its reduced 20×20 version. Time needed to convert one problem into another (`conv.time`)), is less then one tenth of a second!

Importantly, our numerical results show that after scaling the problem down some of the methods show even higher performance than PGM. Of course, for small-scale problems ($N < 7$) the PGM algorithm outperforms the others thanks to low initialization costs and low step complexity for small N (see Boyd and Xiao (2005)). As N increases, however, PGM requires a growing number of iterations to converge. For problems of larger sizes Newton methods start to outperform other methods. The reason is a quadratic rate of convergence of Newton methods in a certain neighborhood of the solution. The BFGS method demonstrates the same performance for low-dimensional problems as the Newton methods, and is slightly outperformed by them for larger matrix sizes. This is consistent with the results of (Qi and Sun (2006); Gao and Sun (2009)). The performance of `newton_bh` is similar to the performance of `newt.` mostly because the problems we consider in (see Appendix A, Table 1) are too small.[26] The benefits of using the Newton-type methods (and the differences between them) are more evident when the size of the matrices used in calculations increases (see below). We find that the SDPT3 solver as well as the QSDP solver work pretty well, but it seems that high initialization costs suppress their performance for the low scale problems.

Adding inequality constraints In the Table 2 in Appendix A, we present results for reduced problems with additional inequality constraints. Following the scheme above we consider $0.35 \cdot \frac{N(N-1)}{2}$ constraints, where as above there are $0.25 \cdot \frac{N(N-1)}{2}$ equality constraints, and additionally there are $0.10 \cdot \frac{N(N-1)}{2}$ inequality constraints. For inequality constraints

$$trace(B_j X) = d_j,\ 1 \leq j \leq k_{ineq}$$

we take B_j random symmetric $(N-1 \times N-1)$-matrix consisting of zeros and ones and minus ones, but not necessary linearly independent. The numbers d_j are chosen at random in the same way as for equality constraints.

In Table 2 in Appendix A we present results for the Newton smoothing method by Gao and Sun (2009) and the PGM method by Boyd and Xiao (2005) carried out on the same sample of 300 random matrices, as for equality constrained cases. The results support conclusion derived for equality constrained methods. If sizes of matrices (ie the number of currency pairs) are less than 6 or 7 PGM method slightly outperforms the Newton smoothing method (`newt_gs`). For problems of larger sizes Newton methods start to outperform PGM.

[26] We implemented the Newton semismooth method with a minres solver, a modified line searcher and a Jacobi preconditioner as suggested by Borsdorf anf Higham (Borsdorf and Higham (2010)). Contrary to them, however, we did not modify the eigenvalue solver. All in all, we do not think that this adversely affects the outcome at the problem scale we are dealing with.

Specifics of the problem of finding the valid covariance matrix for FX-market implies that while initial problem is medium scaled one, the reduced problem is usually low-dimensional. However, most authors of algorithms tested their methods with medium and large-scale problems, presenting no numerical experiments with low dimensional matrices. In order to fill this gap and to ensure that our experiments are inline with the results obtained by authors of algorithms we proceed with some more numerical experiments. In the tests above the number of constraints was proportional to the number of the degrees of freedom, while testing their algorithms Gao and Sun (2009), Qi (2013), Borsdorf and Higham (2010) proceeded somewhat differently. For a $(n \times n)$-matrix they considered $c \cdot n$ of constraints for some $c > 0$. Also they considered slightly higher values of tolerance.

Following Gao and Sun (2009), (see Example 5.9) we test all the methods with equality constraints on the diagonal of the matrix X: $x_{ii} = a + (1-a)\omega$, $i = 1 \dots N$ where $a = 0.1$ and ω is a randomly generated number in $[0, 1]$. Additionally for equality constrained methods we impose $N-1$ restrictions of the type $x_{ij} = 0$, where $i = 1 \dots N-1$ and $j > i$ chosen at random. For inequality constrained methods we also impose $N-1$ additional restrictions $x_{ij} \leq 0.1$, where $i = 1 \dots N-1$ and $j > i$ chosen at random. Also, following Gao and Sun (2009) and Qi and Sun (2006), we set the level of tolerance equals 10^{-6} for dual methods and 10^{-7} for interior point methods[27].

Our tests were carried out on 300 random matrices with coefficients in $[-1, 1]$ of sizes 20×20, 50×50, 100×100 generated under (conditional) uniform distribution with 30% of eigenvalues being less than -0.001.

In the Table 3 in Appendix A, we report our numerical experience with PGM, BFGS, Newton Semismooth, algorithms as well as QSDP and SDPT3 solvers on equality constrained problems. It demonstrates clearly that Newton-type methods (BFGS, Newton Semismooth) outperform the alternating projection method as well as QSDP and SDPT3 solvers. The modified Newton Semismooth method (`newt_bh`) looks beneficial in comparison with the original Newton Semismooth (`newt.`) as matrix size increases.

In the Table 4 in Appendix A, we report our numerical experience with PGM and Smoothing Newton algorithms on problems with both equality and inequality constraints. As expected Smoothing Newton algorithm outperforms PGM for medium-scaled problems.

8 A Remark on the Choice of Norm

The modeller who is to employ the method suggested in this work in practice will most often face the following two extreme situations, which imply two different choices of the norm when solving problem (2). In particular, she can be interested in correcting either the matrix of drivers of size $N-1$, or the complete covariance matrix, i.e. the one including both the drivers and the crosses, of size $\frac{N(N-1)}{2}$. In the former case, the modeller would not in fact be interested in what happens after correction with the elements of the complete covariance matrix located beyond the principal submatrix lying at the intersection of rows and columns with indices of selected driving pairs. It then may turn out that after the correction some elements of the complete matrix are significantly off their market values, which is by default acceptable as the focus of the modeller is the principal submatrix of drivers. In that case the norm $\|\cdot\|_M^2$ in Theorem 6 will be chosen with $M = N - 1$.

[27]Performance of general purpose solver (QSDP) was taken into consideration only by Gao and Sun (2009)

When the focus of attention is the complete covariance matrix, however, the modeller would want by design to have some control over elements beyond the principal matrix of drivers. Hence, the norm $\|\cdot\|_M^2$ in Theorem 6 will be chosen with $M > (N-1)$, or, in fact in most situations with $M = \frac{N(N-1)}{2}$. After isometrically scaling down the initial matrices in both limiting cases[28], and solving the resulting optimization problems, the modeller will have obtained a well-behaved matrices of size $N-1$, from which the complete covariance matrices will be restored using the no-arbitrage equations (triangular relationships). However, obviously, the two solutions will generally differ as one will have been obtained in norm $\|\cdot\|_{N-1}^2$, and the other - in norm $\|\cdot\|_{\frac{N(N-1)}{2}}^2$. This is important to keep in mind when implementing the method in practice. For consistency between many possible solutions we recommend using norm $\|\cdot\|_{\frac{N(N-1)}{2}}^2$ always, regardless of the primary focus of the modeller.

9 Generalizations and Applications

In this part we provide some generalizations to the proposed method of finding a valid covariance matrix that might be useful in practice.

9.1 Weighted problems

First of all, instead of the original problem one may need to solve its weighted version:

$$\begin{cases} \min \frac{1}{2}||X-C||_W^2 \\ X \in \mathcal{C}_{FX}^M \\ trace(A_i\, X) = b_i,\ 1 \le i \le k_{eq} \\ trace(B_j\, X) \ge d_j,\ 1 \le j \le k_{ineq} \\ X \succeq 0 \end{cases} \tag{27}$$

where $W \in \mathcal{S}^M$ is a positive definite matrix of weights and for any $B \in \mathcal{S}^M$

$$||B||_W = ||W^{1/2} B W^{1/2}||_M$$

The main reason one might want to consider a weighted version of the model is a necessity to adjust covariances of different currency pairs individually. Namely, the matrix coefficients with larger weights would be adjusted less. To take a hypothetical example, in practice one can generally be more confident about the value of the estimate of covariance between $EUR-USD$ and $GBP-USD$ than about the value of the estimate of covariance between $BWP-JPY$ and $AUD-RUB$, as the former couple is traded more liquidly then the latter. To take this into account, one can assign the pairs different weights as appropriate. More similar reasons are mentioned in Grubisic and Pietersz (2005), Higham (2002) and Rebonato (2005).

All the methods considered above can easily be generalized for the case of a W-weighted norm. In some algorithms such a modification has already been implemented

[28] In case $M = N-1$ rescaling is just multiplication by identity matrices and therefore leaves problem unchanged.

(QSDP, SDPT3). In others it can straightforwardly be incorporated by change of variables:

$$C_{new} = W^{1/2} C W^{1/2}{}^{29}$$
$$A_i^{new} = W^{1/2} C W^{1/2}$$

As before the problem can be reduced to size $(N-1 \times N-1)$ via change of basis $P^{new} = W^{1/2}P$, where P is given by Theorem 4.

Since for the market with $M = \frac{N(N-1)}{2}$ liquidly traded exchange rates the covariance matrix is completely determined by the rate implied volatilities (see, eg Wystup (2007) or Bouev (2012)), it is especially interesting to consider the so-called H-weighted problem with diagonal matrix H. The minimization function for such a problem takes the form:

$$\min \frac{1}{2}||X-C||_H^2 = trace\left(H^2 \circ (X-C)\cdot(X-C)^T\right) \tag{28}$$

where $\circ$ is Hadamart product $B \circ D = (b_{ij}d_{ij})$. This approach weighs variances only and makes their choice intuitive and comfortable. To our knowledge, there is still no reliable (ie globally convergent and fast) method for solving such a problem, unfortunately. Research by Qi (2013), Qi and Sun (2011), Toh *et al* (2007) illustrates some algorithms already developed.

9.2 Positively definite matrices

In finance in general and in portfolio theory in particular, one often needs a strictly *positive* definite covariance matrix. The reason is that the procedure of finding an optimal portfolio needs the covariance matrix to be invertible (and, moreover, well-conditioned, see Kwan (2010)).

Unfortunately, the covariance matrix in the FX market is usually singular, except for the case of a $(N-1 \times N-1)$-covariance (sub)matrix of driving pairs[30]. In fact market practitioners often work with even smaller submatrices of the covariance matrix of driving pairs (Kwan (2010)). In order to obtain a well-conditioned covariance submatrices one can use the following trick (see Malick (2005), Qi and Sun (2006)). For the reduced problem additional constraint $Y \succ \alpha I_{N-1}$ is considered

$$\begin{cases} \min \frac{1}{2}||Y-K||_{N-1}^2 \\ Y \succ \alpha I_{N-1} \\ trace(\hat{A}_i Y) = b_i,\ 1 \le i \le k_{eq} \\ trace(\hat{B}_j Y) \ge d_j,\ 1 \le j \le k_{ineq} \\ Y \succeq 0 \end{cases}$$

where K, $\hat{A}_i$ and $\hat{B}_j$ are as in (23). Coefficient α guarantees that the required matrix be not less than a positive diagonal matrix. This problem can be reduced to the standard one by subtracting matrix αI_{N-1} at the beginning. Then the problem can be solved via any numerical method. In the end the matrix Y^* is obtained by adding matrix αI_{N-1} to the output matrix (see Qi and Sun (2006) for details). Finally the correct covariance matrix X^* with well-conditioned pricipal submatrices of drivers is obtained by $X^* = P_1 Y^* P_1^T$.

[29] The use of matrix W is similar to introducing a vector of importance weights in the literature on the intrinsic currency framework (see, e.g., Golts, 2010).

[30] As well as its principal submatrices.

9.3 Correlation matrices

This article has primarily focused on FX covariance matrices. However, in stress testing (Qi and Sun (2010)) and some other applications often a correlation matrix is more preferable. The FX market constraints for such matrices are nonlinear, so the optimization problem needs to be re-stated. The solution of such problem is generally unknown. However, in case of a $(N-1\times N-1)$-matrix corresponding to currency drivers, the matrix is free of any additional FX market constraints, so any method considered in this paper or other methods adapted to such problem (see Qi and Sun (2010)) could be applied without change.

9.4 Homotheticity

Finally we mention here that the problem is homothetic. That is for any $T > 0$ one can multiply both the initial matrix C and the vector of affine constraints b by T, and then solve the resulting modified problem. The solution to the initial problem will then be recovered by dividing by T the solution to the modified problem. The benefit of such an approach is that by choosing T large enough one may solve the problem in single precision arithmetics and avoid numeric singularities.

10 Conclusion

In this paper we considered the problem of finding a valid covariance matrix in the FX market. The standard no-arbitrage assumption implies additional linear constraints on such matrices, which automatically makes the null space of such matrices non-empty. As a result, one cannot just take a given, non-PSD estimate of the covariance matrix for driving currency pairs and their crosses, plug it into the standard optimization problem and solve it by applying the most advanced numerical methods developed recently (namely, the dual algorithms). The reason is that such a problem is not well-posed while the PSD-solution is not strictly feasible. In order to deal with this issue, we described a low-dimensional face of the PSD cone that contains the feasible set. After projecting the initial problem onto this face, we come out with a reduced problem, which turns out to be well posed and of a smaller scale. As the reduced problem was solved, the solution to the initial problem can be uniquely recovered in one step. We ran numerous numerical experiments to compare performance of different algorithms in solving the reduced problem and to demonstrate the advantages of dealing with the reduced problem as opposed to the original one. The smaller scale of the reduced problem implies that effectively any numerical method can be applied to find its solution. We find that Newton-type methods start to outperform other methods as soon as the dimension of the problem increases. We mentioned some generalizations that must be considered in stress testing and mean-variance portfolio theory. Our findings are up-to-date and ready for use by market practitioners.

11 Appendix

Appendix A: Numerical experiment results

Table 1: Comparison of unreduced and reduced problem: equality constraints

		Unreduced problem			**Reduced problem**								
N	restr.	unred.size	Med.val	dykstra	conv.time (ms)	red. size	Med.val	PGM	bfgs	newt.	newt._bh	SDTP3	QSDP
5	k_{eq}=2	10x10	Time (ms)	8	<1	4x4	Time (ms)	5	11.5	6	7.5	149	510
			Iter	25.5			Iter	32	8	3	3	12	14
			Res	7.07E-09			Res	7.98E-09	-9.12E-09	2.99E-11	2.01E-11	6.05E-10	7.07E-10
7	k_{eq}=5	21x21	Time (ms)	37	2	6x6	Time (ms)	12	15	8	11	152.5	901.5
			Iter	46			Iter	60	13	4	4	12	15
			Res	8.71E-09			Res	8.79E-09	-1.90E-07	1.02E-11	9.66E-13	9.83E-10	1.15E-09
13	k_{eq}=19	78x78	Time (ms)	1700	6	12x12	Time (ms)	51	23	20	18	178	2102.5
			Iter	104			Iter	108	23	6	5	13	16
			Res	9.53E-09			Res	9.57E-09	-3.82E-08	7.10E-11	6.04E-12	2.06E-09	4.13E-09
21	k_{eq}=52	210x210	Time (ms)	209 788	63	20x20	Time (ms)	685.5	39	71.5	32	301.5	2239
			Iter	638			Iter	556	33	7	7	13	16
			Res	9.9E-09			Res	9.88E-09	-9.7E-09	1.72E-10	1.09E-11	3.4E-09	7.13E-09

Table 2: Comparison of algorithms with equality and inequality constraints

Method	Med. val.	N=5 red.size =4x4 k_{eq}=2 k_{ineq}=1	N=7 red.size=6x6 k_{eq}=5 k_{ineq}=2	N=13 red.size=12x12 k_{eq}=19 k_{ineq}=8	N=21 red.size=20x20 k_{eq}=52 k_{ineq}=21
PGM	Time(ms)	6	17.5	322.5	4277.5
	Iter	36	79	569	2638
	Res	8.17E-09	9.19E-09	9.93E-09	9.99E-09
newt._gs	Time(ms)	14	16	24	35
	Iter	7	8	10	11
	Res	1.56E-09	1.68E-09	1.98E-09	1.52E-09

Table 3: Comparison of algorithms for small and medium-scaled problems: equality constraints

mat.size, restr.	Med.val	PGM	bfgs	newt.	newt._bh	SDTP3	QSDP
20x20 k_{eq}=39	Time(s)	0.0635	0.1441	0.0295	0.0485	0.2625	2.395
	Iter	57	28	5	5	11	12
	Res	9.38E-07	5.75E-08	1.85E-09	2.68E-09	5.44E-07	4.63E-06
50x50 k_{eq}=99	Time(s)	0.8400	0.1603	0.1815	0.066	3.5165	2.8745
	Iter	97	33	5	5	12	13
	Res	9.43E-07	1.51E-08	1.08E-07	3.68E-10	2.21E-06	2.52E-05
100x100 k_{eq}=199	Time(s)	6.1055	0.7389	1.3445	0.3090	73.5335	5.1740
	Iter	119	22	6	5	13	13
	Res	9.69E-07	3.37E-09	7.7E-11	1.08E-08	5.84E-06	8.63E-05

Table 4: Comparison of algorithms for small and medium-scaled problems: equality and inequality constraints

		size=20x20 k_{eq}=20, k_{ineq}=19	size=50x50 k_{eq}=50, k_{ineq}=49	size= 100x100 k_{eq}=100, k_{ineq}=99
PGM	Time(s)	0.2455	3.2385	30.4515
	Iter	252	422.5	632.5
	Res	9.83E-07	9.84E-07	9.88E-07
newt._gs	Time(s)	0.0235	0.041	0.0905
	Iter	9	9	10
	Res	7.81E-10	3.29E-09	1.23E-09

Appendix B: Quotation independence

As in the main text, $\xi_t = (\ln(S_t^{(1)}), \ldots, \ln(S_t^{(M)}))$.

Denote by ξ the normally distributed random vector ξ_1. Therefore the instantaneous covariance matrix Σ equals $E(\xi - E\xi)^T(\xi - E\xi)$. The change of quotation order for the i-th currency pair implies the change of sign for the i-th element of random vector ξ. Hence, the new random vector ζ can be obtained by multiplying the random vector ξ by the unitary matrix U

$$(\zeta_1, \ldots, \zeta_M)^T = U \cdot (\xi_1, \ldots, \xi_M)^T$$

where U is a diagonal $(M \times M)$-matrix satisfying $U_{jj} = 1,\ j \neq i$, and $U_{ii} = -1$. Thereby the covariance matrix of vector ζ :

$$\Sigma_\zeta = E\big(\zeta - E(\zeta)\big)^T\big(\zeta - E(\zeta)\big) = U\,\Sigma_\xi\,U^{-1}$$

Using the Remark 1, $C_\zeta = U\,C_\xi\,U^{-1}$ for any estimate covariance matrices in different quotations C_ξ, C_ζ.

The PSD property of any estimate covariance matrix C_ξ is invariant under quotation change. Namely, the characteristic polynomial of matrix C_ξ is invariant under multiplication by matrix U and its inverse:

$$det(C_\zeta - \lambda I) = det(U\,C_\xi\,U^{-1} - \lambda I) = det(U^{-1}(U\,C_\xi\,U^{-1} - \lambda I)U) =$$
$$det(C_\xi - U^{-1}\,\lambda I\,U) = det(C_\xi - \lambda I)$$

Appendix C: Violating the coercivity property

Let us for simplicity assume that there are no inequailty constraints and consider dual problem (25) that would be the target of application of dual methods. Since it is a maximization problem with a concave objective function the coercivity property of function $\theta(\cdot)$, namely $\theta(y) \to -\infty$ as $||y|| \to \infty$, is of great importance. If the problem satisfies this property then the problem has no duality gap (Rockafellar (1996)). Then the solution to the primal problem can easily be obtained via

$$Z^* = (C + \sum_{i=1}^{k_{eq}} A_i y_i^*)_+ \tag{29}$$

If, however, the problem does not possess that property then (29) holds in the limit at best for some divergent sequences y_i^k:

$$Z^* = \lim_{||y^k|| \to \infty} (C + \sum_{i=1}^{k_{eq}} A_i y_i^k)_+ \tag{30}$$

Unreduced problem (2) does not satisfy the coercivity property even for the (3×3)-matrices consistent with the no arbitrage condition in the FX market. For illustration, consider an example of the FX market spanned by three currencies. Suppose the additional constraints in the initial problem imply that variances of the first and second exchange rates must stritcly be equal to c_1 and c_2, correspondingly. The space $\mathcal{N}_{FX}^M$ is one dimensional by Theorem 2, and vector $v = (1, -1, 1)^T$ belongs to that space. Then by

(20) the constraint $X \in \mathcal{C}_{FX}^M$ is equivalent to $trace(GX) = 0$, $G = vv^T$. The constraints on the variances are $trace(A_1X) = c_1$ and $trace(A_2X) = c_2$, where A_1 and A_2 are the matrices with $A_1(1,1) = 1$ and $A_2(2,2) = 1$, and other elements being equal to zero.

The coercivity property (see Qi and Sun (2006) sect.4.2 or Qi (2013) Proposition 2.1) implies that matrix $y_1 \cdot G + y_2 \cdot A_1 + y_3 \cdot A_2$ is not negative semi-definite for any $y_1, y_2, y_3 \in \mathbb{R}$ such that $y_2c_1 + y_3c_2 \geq 0$. Set $y_1 = -1$, $y_2 = y_3 = 0$, then this condition is violated since matrix $-G$ is negative semi-definite.

From a computational point of view, lack of coercivity implies bad behavior of all the dual algorithms. For illustration we consider the Newton semismooth and PGM algorithms. In the Table 5 (see Appendix A, Table 5) we report performance of the algorithms for two unreduced problems, size 3×3 and 78×78 for two levels of tolerance equal to 10^{-3} and 10^{-8}. For both problems there is linear FX market constraint (3) as well as two affine equality constraints $x_{11} = 1$ and $x_{22} = 1$. Then using (20) we rewrite problem (2) as problem (21) with $k_{FX} = M - N + 1$ general linear constraints and two general affine constraints $trace(A_1X) = 1$ and $trace(A_2X) = 1$, where A_1 and A_2 are the $(M \times M)$-matrices with $A_1(1,1) = 1$, $A_2(2,2) = 1$, and other elements being equal to zero. In addition to time, the number of iterations[31] and the residual, we also report the value of the norm of the dual solution at the last iteration as well as its accuracy evaluated as the distance to the "exact" solution. This "exact" solution was obtained by the (primal) alternating projections algorithm with Dykstra's correction with the tolerance parameter equal to 10^{-14}. One can see that as the *tol* parameter decreases from 10^{-3} to 10^{-8} the value of the norm of the dual solution increases dramatically. Despite the fact that the solution obtained by dual methods does "tend" to the "exact" solution, the rate of convergence seems to be extremely low whereas the number of iterations is very high. That points to convergence in the limit at best as in (30) above.

Appendix D: The principle of the reduction: the algorithm

1. Select any $N-1$ driving pairs among M currency pairs under consideration. Let their indexes be $\tau_1, \ldots, \tau_{N-1}$. Construct $M - N + 1$ triangular relations, expressing the remaining $M - N + 1$ spot exchange rates;

2. For $j = 1$ to $M - N + 1$

   (a) Using triangular relations, as shown below the formulae (14), construct a $M-$ dimensional vector-column v_j in $\{-1, 0, 1\}^M$

3. Construct a $(M \times M - N + 1)$-matrix V whose columns are vectors $v_1, \ldots, v_{M-N+1}$;

4. Compute the $(M \times N - 1)$-matrix P_1 in three steps. First compute $(M \times N - 1)$-matrix S with the j-th column given by

$$S(:, j) := e_j - V(V^TV)^{-1}V^Te_j \qquad j = 1 \ldots N - 1$$

   where e_j is the M dimensional vector-column with τ_j component being equal to one and other components being equal to zero. Next, apply Gram-Schmidt process to columns of S:

[31] * means that an algorithm reaches the set maximum number of iterations before the accuracy is achieved

Table 5: Solving unreduced problem with primal and dual methods

Tol	Method	Time	Iter	Res	norm of dual	accuracy
Case 1: $N = 3, M = 3, k_{eq} = 2, k_{FX} = 1$						
		Solution with primal method:				
tol=E-14	dykstra	0.019	4	7.6E-16	NaN	0
		Solution with dual methods:				
tol=E-03	newt.	0.136	20	0.000709	62.5417	0.029172
	PGM	1.257	8718	0.001	45.85299	0.040065
tol=E-08	newt.	0.173	35	1.75E-10	567184.1	3.38E-06
	PGM	14.402	100000*	0.000197	104.8872	0.017908
Case 2: $N = 13, M = 78, k_{eq} = 2, k_{FX} = 66$						
		Solution with primal method:				
tol=E-14	dykstra	0.811	59	7.54E-15	NaN	0
		Solution with dual methods:				
tol=E-03	newt.	0.998	29	0.000814	3628.239	0.607694
	PGM	1217	100000*	0.003334	1511.924	1.312189
tol=E-08	newt.	28.191	41	0.000475	322679.2	0.000921
	PGM	1236	100000*	0.001353	694.0605	0.29836

$$P_1(:,j) := S(:,j) - \sum_{k=1}^{j-1} \frac{S(:,j)^T P_1(:,k)}{||S(:,k)||^2} S(:,k)$$

and finally normalize the result $P_1(:,j) := \frac{P_1(:,j)}{||P_1(:,j)||^2}$;

5. Set $K = P_1^T C P_1,\ \hat{A}_i = P_1^T A_i P_1$ for $1 \le i \le k_{eq},\ \hat{B}_l = P_1^T B_l P_1$ for $1 \le l \le k_{ineq}$;
6. Obtain the solution Y^* of the problem (22) by any recommended method;
7. The solution X^* of the problem (2) is given by $X^* = P_1 Y^* P_1^T$.

Appendix E: Numerical example

In this section we present an example of the step-by-step realization of our algorithm for a small-scale problem. We start from selecting random values of the implied volatilities of four currencies(EUR, GBP, USD and JPY) cross-rates. This data is presented in Table 6.

Table 6: Input volatilities

Quotation	Impl. Vol. in %	Impl. Var
EURJPY	9.70%	0.00940900
EURGBP	7.40%	0.00547600
EURUSD	15.70%	0.02464900
GBPJPY	9.60%	0.00921600
USDJPY	9.80%	0.00960400
GBPUSD	8.70%	0.00756900

The corresponding matrix C is presented below[32]:

$$\begin{pmatrix} 0.00940900 & 0.00283450 & 0.01222700 & 0.00657450 & -0.00281800 & 0.00939250 \\ 0.00283450 & 0.00547600 & 0.01127800 & -0.00264150 & -0.00844350 & 0.00580200 \\ 0.01222700 & 0.01127800 & 0.02464900 & 0.00094900 & -0.01242200 & 0.01337100 \\ 0.00657450 & -0.00264150 & 0.00094900 & 0.00921600 & 0.00562550 & 0.00359050 \\ -0.00281800 & -0.00844350 & -0.01242200 & 0.00562550 & 0.00960400 & -0.00397850 \\ 0.00939250 & 0.00580200 & 0.01337100 & 0.00359050 & -0.00397850 & 0.00756900 \end{pmatrix}$$

Its minimal eigenvalue $\lambda_{\min}$ equals -0.0027 and therefore matrix C is non PSD.

Initial volatility of $EUR-USD$ equals 15.7%. We assume that trader wants to enforce a constraint (for example, for stress-testing purposes) for the corrected matrix: volatility of $EUR-USD$ should equal to $0.12 = 12\%$. The corresponding variance should equal to 0.0144.

Therefore we consider the following initial problem:

$$\begin{cases} \min \frac{1}{2}||X - C||_6^2 \\ X \in \mathcal{C}_{FX}^6 \\ trace(A\,X) = 0.0144 \\ X \succeq 0 \end{cases}$$

where $A_{3,3} = 1$ and the other components of matrix A are equal to zero.

According to the Theorem 6 we can isometrically reduce this 6×6 problem to 3×3 problem. To do this we need to construct the matrix P_1 from Theorem 6. The matrix P_1 columns are orthonormal basis vectors of orthogonal complement of the space $\mathcal{N}_{FX}^6$. The space $\mathcal{N}_{FX}^6$ is given by (linearly independent) eigenvectors $v^{(ijk)}$. Since

$$\frac{EUR}{JPY} \cdot \left(\frac{EUR}{GBP}\right)^{-1} \cdot \left(\frac{GBP}{JPY}\right)^{-1} = 1 \qquad \frac{EUR}{JPY} \cdot \left(\frac{EUR}{USD}\right)^{-1} \cdot \left(\frac{USD}{JPY}\right)^{-1} = 1$$
$$\frac{EUR}{GBP} \cdot \left(\frac{EUR}{USD}\right)^{-1} \cdot \frac{GBP}{USD} = 1$$

these eigenvectors can be taken as follows

[32]The covariance coefficients are implied from the volatilities based on equations (11) and (12).

$$\begin{bmatrix} 1 \\ -1 \\ 0 \\ -1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \\ -1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

The matrix P_1 could be taken as follows

$$\begin{pmatrix} 0.70710678 & 0.00000000 & 0.00000000 \\ 0.35355339 & 0.61237244 & 0.00000000 \\ 0.35355339 & 0.20412415 & 0.57735027 \\ 0.35355339 & -0.61237244 & 0.00000000 \\ 0.35355339 & -0.20412415 & -0.57735027 \\ 0.00000000 & -0.40824829 & 0.57735027 \end{pmatrix}$$

The reduced matrix $K = P_1^T C P_1$ equals

$$\begin{pmatrix} 0.01881800 & -0.00431858 & 0.01995314 \\ -0.00431858 & 0.01331667 & 0.01254384 \\ 0.01995314 & 0.01254384 & 0.03378833 \end{pmatrix}$$

The corresponding constraints matrix $\hat{A} = P_1^T A P_1$ equals

$$\begin{pmatrix} 0.12500000 & 0.07216878 & 0.20412415 \\ 0.07216878 & 0.04166667 & 0.11785113 \\ 0.20412415 & 0.11785113 & 0.33333333 \end{pmatrix}$$

By Theorem 6 the initial problem is equivalent to the following one:

$$\begin{cases} \min \frac{1}{2}||Y - K||_3^2 \\ trace(\hat{A}\, Y) = 0.0144 \\ Y \succeq 0 \end{cases}$$

We find the solution of this problem using dual method by Boyd and Xiao (2005). Its solution Y^* equals

$$\begin{pmatrix} 0.01476531 & -0.00625943 & 0.01069353 \\ -0.00625943 & 0.01257461 & 0.00686846 \\ 0.01069353 & 0.00686846 & 0.02084801 \end{pmatrix}$$

Using formula $X^* = P_1 Y^* P_1^T$ we obtain the solution X^* of the initial problem. It equals

$$\begin{pmatrix} 0.00738265 & 0.00098092 & 0.00715347 & 0.00640174 & 0.00022918 & 0.00617256 \\ 0.00098092 & 0.00385073 & 0.00622172 & -0.00286982 & -0.00524081 & 0.00237099 \\ 0.00715347 & 0.00622172 & 0.01440000 & 0.00093175 & -0.00724653 & 0.00817828 \\ 0.00640174 & -0.00286982 & 0.00093175 & 0.00927156 & 0.00546999 & 0.00380156 \\ 0.00022918 & -0.00524081 & -0.00724653 & 0.00546999 & 0.00747571 & -0.00200572 \\ 0.00617256 & 0.00237099 & 0.00817828 & 0.00380156 & -0.00200572 & 0.00580728 \end{pmatrix}$$

Answer, presented in the input data format is:

Table 7: Output volatilities

Quotation	Impl. Vol. in %	Impl. Var
EURJPY	8.59%	0.00738265
EURGBP	6.21%	0.00385073
EURUSD	**12.00%**	0.01440000
GBPJPY	9.63%	0.00927156
USDJPY	8.65%	0.00747571
GBPUSD	7.62%	0.00580728

References

Alfakih, A., Khandani, A., and Wolkowicz, H. (1999). Solving Euclidean Distance Matrix Completion Problems Via Semidefinite Programming. *Computational Optimization* (Jong-Shi Pang, ed.), 13–30, Springer US.

Austing, P. (2011). Repricing the cross smile: an analytic joint density. *Journal of Risk*.

Borsdorf, R. and Higham, N. J. (2010). A Preconditioned Newton Algorithm for the Nearest Correlation Matrix. *j-IMAJNA* 30, 94–107.

Bouev, M. (2012). Finding a Valid FX Covariance Matrix in the BS World. EUSP Deparment of Economics Working Paper Series Ec-03/12 European University at St. Petersburg, Department of Economics.

Boyd, S. and Xiao, L. (2005). Least-Squares Covariance Matrix Adjustment. *SIAM J. Matrix Anal. Appl.* 27, 532–546.

Chen, X. (1995). Convergence of the BFGS Method for LC^1 Convex Constrained Optimization.

Chen, X., Qi, H., and Tseng, P. (2002). Analysis of Nonsmooth Symmetric-Matrix-Valued Functions with Applications to Semidefinite Complementarity Problems. *SIAM J. on Optimization* 13, 960–985.

Dykstra, R. (1983). An algorithm for restricted least squares regression. *J. Amer. Stat. Assoc.*

Eaves, B.C. (1971). On the basic theorem of complementarity. *Mathematical Programming* 1, 68–75.

Gao, Y. and Sun, D. (2009). Calibrating Least Squares Semidefinite Programming with Equality and Inequality Constraints. *SIAM J. Matrix Anal. Appl.* 31, 1432–1457.

Grubisic, I. and Pietersz, R. (2005). Efficient Rank Reduction of Correlation Matrices. Finance 0502007 EconWPA.

Henrion, D. and Malick, J. (2007). SDLS: a Matlab package for solving conic least-squares problems.

Henrion, D. and Malick, J. (2012). Projection Methods in Conic Optimization. *Handbook on Semidefinite, Conic and Polynomial Optimization* (M. F. Anjos and J. B. Lasserre, eds.) volume 166 of *International Series in Operations Research And Management Science*, 565–600, Springer US.

Higham, N. J. (1988). Computing a Nearest Symmetric Positive Semidefinite Matrix. *j-LAA* 103, 103–118.

Higham, N. J. (2002). Computing the nearest correlation matrix—a problem from finance. *IMA Journal of Numerical Analysis* 22, 329–343.

Kwan, C. C. Y. (2010). The Requirement of a Positive Definite Covariance Matrix of Security Returns for Mean-Variance Portfolio Analysis: A Pedagogic Illustration.

Malick, J. (2005). A Dual Approach to Semidefinite Least-Squares Problems. *SIAM J. Matrix Anal. Appl.* 26, 272–284.

Qi, H. (2013). A Semismooth Newton Method for the Nearest Euclidean Distance Matrix Problem. *SIAM Journal on Matrix Analysis and Applications* 34, 67–93.

Qi, H. and Sun, D. (2006). A Quadratically Convergent Newton Method for Computing the Nearest Correlation Matrix. *SIAM J. Matrix Anal. Appl.* 28, 360–385.

Qi, H. and Sun, D. (2010). Correlation stress testing for value-at-risk: an unconstrained convex optimization approach. *Comput. Optim. Appl.* 45, 427–462.

Qi, H. and Sun, D. (2011). An augmented Lagrangian dual approach for the H-weighted nearest correlation matrix problem. *IMA J. Numer. Anal.* 31, 491–511.

Rebonato, R. (2005). *Volatility and Correlation: The Perfect Hedger and the Fox*. The Wiley Finance Series Wiley.

Rebonato, R. and Jäckel, P. (2000). The most general methodology for creating a valid correlation matrix for risk management and option pricing purposes. *Journal of Risk*.

Rockafellar, R. T. (1996). *Convex Analysis (Princeton Landmarks in Mathematics and Physics)*. Princeton University Press.

Schöttle, K. and Werner, R. (2004). Improving the "Most General Methodology to Create a Valid Correlation Matrix for Risk Management and Option Pricing Purposes". *Risk Analysis IV, Management Information Systems* 9, 701–712.

Sturm, J. F. (1999). Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software* 11–12, 625–653.

Sun, D. and Sun, J. (2002). Semismooth Matrix-Valued Functions. *Math. Oper. Res.* 27, 150–169.

TCBS (2013). Triennial Central Bank Survey, Foreign exchange turnover in April 2013: preliminary global results.

Todd, M.J. (2001). Semidefinite Optimization. *Acta Numerica* 10, 515–560.

Toh, K. C. (2007). An inexact primal-dual path following algorithm for convex quadratic SDP. *Math. Program.* 112, 221–254.

Toh, K.C., Tütüncü, R.H., and Todd, M.J. (2007). Inexact primal-dual path-following algorithms for a special class of convex quadratic SDP and related problems. *Pacific Journal on Optimization* 3, 135–164.

Tütüncü, R. H., Toh, K. C., and Todd, M. J. (2003). Solving semidefinite-quadratic-linear programs using SDPT3. *Mathematical Programming* 95, 189–217.

Wolkowicz, H. (1981). Some applications of optimization in matrix theory. *Linear Algebra and its Applications* 40, 101 – 118.

Wystup, U. (2001). How the Greeks would have hedged correlation risk of foreign exchange options. *Wilmott Research Report, August.*

Wystup, U. (2007). *FX Options and Structured Products.* The Wiley Finance Series Wiley.

Aleksei Minabutdinov;

National Research University Higher School of Economics (Saint-Petersburg, Russia), Department of Mathematics, Senior Lecturer;

E-mail: aminabutdinov@gmail.com;

mob. 8 904 334 65 91; 8 921 634 83 72