Completing the correlation matrix in a hybrid model.

by Horst Köhler, Thomas Streuer, Olaf Dreyer

1 Introduction

2 The cross currency model

To make the issues more concrete we introduce the basic cross currency model as an example. This model consists of two interest rate models and a model for the exchange rate. We will denote the base currency by E and assume that we model a rate x_E together with a volatility v_E. The details of the model are not important since we are only concerned about the correlation of the stochastic processes. An example of such a model would be a Heath-JarrowÂĂ-Morton model with a stochastic volatility. The model for the foreign currency A will also consist of a model for the rate x_A and the volatility v_A. The remaining model will be for the exchange rate

$$X_E^A, \tag{1}$$

where the lower index E denotes the currency we are converting into and the upper index is the currency we are converting from. If m_A is an amount in currency A then

$$X_E^A m_A \tag{2}$$

is the corresponding amount in the currency E. We also assume that we are modeling the volatility v_{EA} of the exchange rate. An example of such a model would be a Heston model [1]. Altogether we thus have these six stochastic variables:

$$x_E, v_E, x_A, v_A, X_E^A, v_{EA} \tag{3}$$

Let us denote this set of variables by I. For two elements $a, b \in I$ let

$$(a, b) = \langle W_a, W_b \rangle \tag{4}$$

be the correlation coefficient[1] for the stochastic processes of a and b. For all $a, b \in I$ we have

$$(a, b) = (b, a) \tag{5}$$
$$(a, a) = 1 \tag{6}$$
$$|(a, b)| \leq 1 \tag{7}$$

---

[1]We introduce this non-standard notation here to improve the readability in later chapters.

In principle one could imagine that all entries of W are determined in one calibration using an assortment of financial products. This is usually not what one does though because the number of parameters is too large and hence the calibration is too unstable but also because one will have already performed calibrations for the interest models themselves and thus

$$(x_E, v_E) \text{ and } (x_A, v_A) \tag{8}$$

are already known. In addition to these two, one might decide to calibrate the coefficients

$$(x_E, X_E^A), (x_E, x_A), (x_A, X_E^A), \text{ and } (X_E^A, v_{EA}). \tag{9}$$

The question that arises is then how to determine the remaining coefficients (see figure 1) of the correlation matrix.

$$\begin{pmatrix} 1 & \boxed{(x_E, v_E)} & (x_E, x_A) & (x_E, v_A) & (x_E, X_E^A) & (x_E, v_{EA}) \\ & 1 & (v_E, x_A) & (v_E, v_A) & (v_E, X_E^A) & (v_E, v_{EA}) \\ & & 1 & \boxed{(x_A, v_A)} & (x_A, X_E^A) & (x_A, v_{EA}) \\ & & & 1 & (v_A, X_E^A) & (v_A, v_{EA}) \\ & & & & 1 & (X_E^A, v_{EA}) \\ & & & & & 1 \end{pmatrix}$$

Figure 1: The correlation matrix for the cross currency model. The coefficients (x_E, v_E) and (x_A, v_A) (the boxes with the solid lines) will be determined in two separate interest rate calibrations. The coefficients (x_E, X_E^A), (x_E, x_A), (x_A, X_E^A), and (X_E^A, v_{EA}) (the boxes with the dashed lines) will be determined in the cross currency calibration. The question then arises what to do with he remaining coefficients?

3   Independence, maximal determinant, and entropy

3.1   A simpler problem

Before we answer the question of the last section in full generality we look at a simpler case to determine how we want to complete the matrix M. As in the previous

section we denote by I be the set of stochastic processes. For this simple case we assume that there are two sets $\alpha, \beta \subset I$ with $\alpha \cup \beta = I$ and $\alpha \cap \beta = \gamma \neq \emptyset$. For any $\delta \subset I$ we will denote by

$$M[\delta] \tag{10}$$

the principal submatrix of M with row and column indices in δ. For two sets $\delta, \eta \subset I$ we denote by

$$M[\delta, \eta] \tag{11}$$

the submatrix of M with row indices in δ and column indices in η. We assume that

$$M[\alpha] \text{ and } M[\beta] \tag{12}$$

are known. The matrix M can then be written in the form

$$M = \begin{pmatrix} C & D & W \\ D^T & E & F \\ W^T & F^T & G \end{pmatrix}, \tag{13}$$

where we make the following identifications:

$$M[\alpha] = \begin{pmatrix} C & D \\ D^T & E \end{pmatrix} \tag{14}$$

$$M[\beta] = \begin{pmatrix} E & F \\ F^T & G \end{pmatrix} \tag{15}$$

$$M[\gamma] = E \tag{16}$$

$$M[\alpha - \gamma, \alpha - \gamma] = C \tag{17}$$

$$M[\alpha - \gamma, \gamma] = D \tag{18}$$

$$M[\alpha - \gamma, \beta - \gamma] = W \tag{19}$$

$$M[\gamma, \beta - \gamma] = F \tag{20}$$

$$M[\beta - \gamma, \beta - \gamma] = G \tag{21}$$

The part of the matrix that we need to determine is W. If the overlap E didn't exist we would just set $W = 0$ and thus ensure that α and β are independent. This is the desirable result because any other choice would assume information about the correlation between α and β that we do not have. Since the overlap E is present, we have to make a different choice.

3.2 Conditioning & the Schur complement

For $\delta \subset I$ let us assume that $M[\delta]$ is invertible (we will weaken this assumption later). Then let

$$M/M[\delta] = M[I - \delta, I - \delta] - M[I - \delta, \delta]M[\delta]^{-1}M[\delta, I - \delta] \tag{22}$$

be the Schur complement[2] (see [3] for a an overview of the properties of the Schur complement) of $M[\delta]$ in M. In our context the Schur complement arises naturally when we ask for the correlation matrix of a subset of the stochastic variables in I conditioned on the values of the remaining variables. If the variables that we are conditioning on are in γ then $M/M[\gamma]$ is the correlation matrix for the variables in $I-\gamma$ (for a proof see [4, chapter 8] and [5]).

3.3 The completion

With the result from the last section we have a way to proceed. We can look at the correlation matrix for the elements in I that we obtain when we condition on the elements in $\gamma = \alpha \cap \beta$. This matrix is given by the Schur complement

$$M/M[\gamma]. \tag{23}$$

We can use equation (13) to write

$$M/M[\gamma] = M/E \tag{24}$$

$$= \begin{pmatrix} C - DE^{-1}D^T & W - DE^{-1}F \\ W^T - F^T E^{-1} D^T & G - F^T E^{-1} F \end{pmatrix}. \tag{25}$$

A special choice for W now presents itself. If we choose

$$W = DE^{-1}F \tag{26}$$

the off-diagonal matrices vanish and the matrix becomes block-diagonal. We can even interpret the remaining blocks. We note that

$$M[\alpha]/M[\gamma] = C - DE^{-1}D^T \tag{27}$$

$$M[\beta]/M[\gamma] = G - F^T E^{-1} F. \tag{28}$$

Again, using the result from the previous section, we see that the remaining blocks give the correlation matrix for α and β respectively, when we condition on γ. The block-diagonal form in equation (26) says that the variables α conditioned on γ are independent from the variables β conditioned on γ. If f is the corresponding density function we have

$$f(x_{I-\gamma}|\gamma) = f(x_{\alpha-\gamma}|\gamma) f(x_{\beta-\gamma}|\gamma). \tag{29}$$

Given that γ is shared by both α and β this is the most independence we can hope for.

3.4 Maximal determinant

Another way to characterize the choice of W in equation (26) is by looking at the determinant of M. In general we have

$$\det(M) = \det(M[\gamma]) \det(M/M[\gamma]) \tag{30}$$

$$\leq \det(M[\gamma]) \det(M[\alpha]/M[\gamma]) \det(M[\beta/M[\gamma]), \tag{31}$$

(see [3] for the first equality), where we have used the Fischer inequality in the second line. With W as in equation (26) the Schur complement $M/M[\gamma]$ becomes block-diagonal and we obtain

$$\det(M) = \det(M[\gamma]) \det(M[\alpha]/M[\gamma]) \det(M[\beta/M[\gamma]). \tag{32}$$

The determinant thus reaches the upper bound of equation (31). Choosing W as in equation (26) thus maximizes the determinant of M.

3.5 Zeroes & the inverse matrix

A remarkable formula due to Babachiewicz [3] expresses the inverse of a matrix in terms of the Schur complement. For a matrix

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \tag{33}$$

we have

$$\Sigma^{-1} = \begin{pmatrix} \Sigma_{11}^{-1} + \Sigma_{11}^{-1}\Sigma_{12}(\Sigma/\Sigma_{11})^{-1}\Sigma_{21}\Sigma_{11}^{-1} & -\Sigma_{11}^{-1}\Sigma_{12}(\Sigma/\Sigma_{11})^{-1} \\ -(\Sigma/\Sigma_{11})^{-1}\Sigma_{21}\Sigma_{11}^{-1} & (\Sigma/\Sigma_{11})^{-1} \end{pmatrix}. \tag{34}$$

This formula has the remarkable property that only the inverses

$$\Sigma_{11}^{-1} \text{ and } (\Sigma/\Sigma_{11})^{-1} \tag{35}$$

appear in it. With our choice of W in equation (26) the Schur complement M/E is block diagonal. This means that also its inverse is block diagonal. Indeed, we have

$$(M/E)^{-1} = \begin{pmatrix} (M[\alpha]/E)^{-1} & 0 \\ 0 & (M[\beta]/E)^{-1} \end{pmatrix}. \tag{36}$$

This allows us to write the inverse of M explicitly:

$$M^{-1} = \begin{pmatrix} (M[\alpha]/E)^{-1} & -(M[\alpha]/E)^{-1}DE^{-1} & 0 \\ -E^{-1}D^T(M[\alpha]/E)^{-1} & \Xi & -E^{-1}F(M[\beta]/E)^{-1} \\ 0 & -(M[\beta]/E)^{-1}F^TE^{-1} & (M[\beta]/E)^{-1} \end{pmatrix}, \tag{37}$$

with

$$\Xi = E^{-1} + E^{-1}\left(D^T(M[\alpha]/E)^{-1}D + F(M[\beta]/E)^{-1}F^T\right)E^{-1}. \tag{38}$$

We see that the inverse of M has zeroes in exactly those places where the undetermined matrix W used to be.

3.6 Entropy

The determinant of M is directly related to the entropy of the distribution that is described by M:

$$S = \frac{1}{2}\big(\log(\det(M)) + n\log(2\pi) + n\big) \tag{39}$$

Because we are maximizing the determinant we are also maximizing the entropy.

3.7 The semidefinite case

In the previous sections we have assumed that $E = M[\gamma]$ is non-singular so that E^{-1} exists. In this section we will only assume that E is positive semi-definite. It is in this case that our results differ from those in [6]. Because of this we give new proves for the following results that do not make a distinction between the definite and semi-definite case. We start by noting some properties of symmetric matrices that are also positive semi-definite.

Lemma 3.1 *Let Σ be a symmetric matrix that is partitioned as in equation (33). If Σ is positive semidefinite then we have*

1. $\mathscr{C}(\Sigma_{12}) \subset \mathscr{C}(\Sigma_{11})$

2. $\mathscr{R}(\Sigma_{21}) \subset \mathscr{R}(\Sigma_{11})$

3. $r(\Sigma) = r(\Sigma_{11}) + r(\Sigma_{22})$

Here $\mathscr{C}(\cdot)$ denotes the span of the column vectors of a matrix, $\mathscr{R}(\cdot)$ denotes the row space, and $r(\cdot)$ denotes the rank of a matrix.

Proof The proof is given in appendix A. ∎

In the following we no longer assume that the matrices are non-singular. It turns out that we can still define the Schur complement of a matrix. For a matrix E we will denote by E^- a generalized inverse of E. E^- is a matrix that satisfies

$$EE^-E = E. \tag{40}$$

A particular example of a generalized inverse is the Moore-Penrose inverse E^+ [8]. For now, we will not assume that we are dealing with E^+ but keep the discussion general. Let us look at a general matrix Σ as in equation (33). Given a particular choice of generalized inverse we define

$$\Sigma/\Sigma_{11} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^-\Sigma_{12}. \tag{41}$$

If Σ is a symmetric matrix that is also positive semi-definite it turns out that the Schur complement does not depend on the choice of the particular generalized inverse used in its definition.

Lemma 3.2 *Let Σ be a symmetric matrix that is partitioned as in equation (33). If Σ is positive semidefinite then*

$$\Sigma/\Sigma_{11} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-}\Sigma_{12} \tag{42}$$

does not depend on the choice of the generalized inverse Σ_{11}^{-}.

Proof The proof is given in appendix A. ∎

The next theorem characterizes positive symmetric matrices in terms of their Schur complement:

Theorem 3.3 *Let Σ be a symmetric matrix that is partitioned as in equation (33) where Σ_{11} and Σ_{22} are square. Then $\Sigma \geq 0$ if and only if $\Sigma_{11} \geq 0$, $\mathscr{C}(\Sigma_{12}) \subset \mathscr{C}(\Sigma_{11})$, and $\Sigma/\Sigma_{11} \geq 0$.*

Proof See [3, Theorem 1.20] and [9]. ∎

In equation (26) we defined W using the inverse E^{-1}. We are now going to define W using a general inverse instead. The next theorem shows that this again defines a positive semidefinite completion of M. The following theorems then show that the other properties also carry over to the semidefinite case.

Theorem 3.4 *Let M be as in equation (13). Let both $M[\alpha]$ and $M[\beta]$ be positive semidefinite. Let E^{-} be a generalized inverse of E that is also symmetric. Setting*

$$W = DE^{-}F \tag{43}$$
$$= M[\alpha - \gamma, \gamma]M[\gamma]^{-}M[\gamma, \beta - \gamma] \tag{44}$$

turns M into a positive semidefinite matrix.

Proof Since both $M[\alpha]$ and $M[\beta]$ are positive semidefinite it follows from theorem 3.3 that

$$E = M[\gamma] \geq 0 \tag{45}$$
$$\mathscr{C}(D^{T}) = \mathscr{C}(M[\alpha - \gamma, \gamma]^{T}) \subset \mathscr{C}(E) \tag{46}$$
$$\mathscr{C}(F) = \mathscr{C}(M[\gamma, \beta - \gamma]) \subset \mathscr{C}(E) \tag{47}$$
$$M[\alpha]/E \geq 0 \tag{48}$$
$$M[\beta]/E \geq 0 \tag{49}$$

From this it follows that

$$E = M[\gamma] \geq 0 \tag{50}$$

$$\mathscr{C}(D^T\ F) = \mathscr{C}(M[\gamma, I - \gamma]) \subset \mathscr{C}(E) \tag{51}$$

$$M/E = \begin{pmatrix} M[\alpha]/E & 0 \\ 0 & M[\beta]/E \end{pmatrix} \geq 0 \tag{52}$$

Using theorem 3.3 in the other direction now yields the result. ∎

We next look at the generalized inverse M^+ of M. It turns out that it can be uniquely expressed in terms of the Schur complement. This is a generalization of the Babachiewicz's result from equation (34) to the semidefinite case.

Theorem 3.5 *Let Σ be a symmetric matrix partitioned as in equation (33). Then we have*

$$\Sigma^+ = \begin{pmatrix} \Sigma_{11}^+ + \Sigma_{11}^+ \Sigma_{12}(\Sigma/\Sigma_{11})^+ \Sigma_{21}\Sigma_{11}^+ & -\Sigma_{11}^+ \Sigma_{12}(\Sigma/\Sigma_{11})^+ \\ -(\Sigma/\Sigma_{11})^+ \Sigma_{21}\Sigma_{11}^+ & (\Sigma/\Sigma_{11})^+ \end{pmatrix} \tag{53}$$

if and only if

$$r(\Sigma) = r(\Sigma_{11}) + r(\Sigma_{22}). \tag{54}$$

Proof See [7, section 4] and [3]. ∎

For a symmetric positive semidefinite matrix Σ the condition on the rank of Σ is automatically satisfied. We thus have the following corollary:

Corollary 3.1 *Let Σ be symmetric and positive semidefinite. Then Σ^+ is given by equation (53).*

We can now state a unique property of our completion:

Theorem 3.6 *The completion defined by setting $W = DE^+F$ is the unique completion that turns M into a symmetric positive semidefinite matrix and for which M^+ has zeroes in the places where W used to be.*

The uniqueness result stated in [6] gives this result only for the positive definite case. This result shows that it is indeed also true for the positive semidefinite case.

Proof Because of the previous corollary 3.1 we see that for any completion W that turns M into a symmetric positive semidefinite matrix the Moore-Penrose inverse is given by equation (53). This inverse has zeroes in exactly the places where W used to be if and only if $W = DE^+F$. ∎

4 The completion procedure

5 The general hybrid model

References

[1] Steven L. Heston, A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options, The Review of Financial Studies, **6**(2), April 1993, 327–343.

[2] Emilie Virginia Haynsworth, Determination of the inertia of a partitioned Hermitian matrix, Linear Algebra and its Applications **1**, 1968, 73–81.

[3] Roger A. Horn, Basic Properties of the Schur Complement. In Fuzhen Zhang (Ed.) The Schur complement and its application (17–46). Springer, 2005.

[4] C. Radharkishna Rao, Linear Statistical Inference and its Applications, 2nd edition, John Wiley & Sons, Inc, 2002.

[5] Richard W. Cottle, Manifestations of the Schur complement, Linear Algebra and its Applications **8**, 1974, 189–211.

[6] Smith, R. (2008). The positive definite completion problem revisited Linear Algebra and its Applications 429(7), 1442-1452. https://dx.doi.org/10.1016/j.laa.2008.04.020

[7] Ouellette, D. (1981). Schur complements and statistics Linear Algebra and its Applications 36, 187-295. https://dx.doi.org/10.1016/0024-3795(81)90232-9

[8] Penrose, R. (1955). A generalized inverse for matrices Mathematical Proceedings of the Cambridge Philosophical Society 51(3), 406-413. https://dx.doi.org/10.1017/s0305004100030401

[9] Albert, A. (1969). Conditions for Positive and Nonnegative Definiteness in Terms of Pseudoinverses SIAM Journal on Applied Mathematics 17(2), 434-440. https://dx.doi.org/10.1137/0117041

A Symmetric matrices & the Schur complement