Completing correlation matrices

by Horst Köhler, Thomas Streuer, Olaf Dreyer

1 Introduction

2 The completion criterium

Given an incomplete correlation matrix what is the right way to complete it? To motivate the answer we need to introduce some notation. Let

$$X = (x_1, \ldots, x_k)^T \in \mathbb{R}^k \tag{1}$$

be a k dimensional random variable with a density function given by a k dimensional normal distribution

$$\begin{aligned} f(X) &= N_k[\mu, H_X](X) & (2)\\ &= (2\pi)^{-k/2} \det(H_X)^{-1/2} \exp\left(-\frac{1}{2}(X-\mu)^T H_X^{-1}(X-\mu)\right), & (3) \end{aligned}$$

with mean μ and covariance matrix H. A crucial property of the normal distribution is that if we condition on some of the x_i, $i = 1, \ldots, k$, we again obtain a normal distribution. Assume that we want to condition on the last l components of X:

$$Z = (x_{k-l+1}, \ldots, x_k)^T \in \mathbb{R}^l. \tag{4}$$

We partition the matrix H in a way that reflects this partition of X:

$$H_X = \begin{pmatrix} A & B \\ B^T & C \end{pmatrix}, \tag{5}$$

with $A \in M_{k-l}$, $C \in M_l$, and $B \in M_{k-l,l}$. Let $\bar{X}$ be the first $k-l$ components of X:

$$\bar{X} = (x_1, \ldots, x_{k-l})^T \in \mathbb{R}^{k-l} \tag{6}$$

If we assume that X has zero mean (i.e. $\mu = 0$), the density function for X conditioned on Z is then given by

$$f(X|Z) = N_{k-l}[BC^{-1}Z, H_X/C](\bar{X}), \tag{7}$$

(see [1, chapter 8] and [2]) where H_X/C is the Schur complement of C in H_X:

$$H_X/C = A - BC^{-1}B^T \tag{8}$$

(Emilie Haynsworth introduced the name and highlighted its usefulness in [3]. Schur originally made use of it in [11]. For an overview of the properties of the Schur complement see [5].) We see that we again obtain a normal distribution. We will see the Schur complement again soon.

Now let us look at a second set of random variables

$$Y = (x_{k-l+1}, \ldots, x_k, x_{k+1}, \ldots, x_n)^T \in \mathbb{R}^{n-k}, \tag{9}$$

that has an overlap Z with the random variables X. Let the density function for Y be given by

$$f(Y) = N_{n-k}[0, H_Y](Y), \tag{10}$$

with

$$H_Y = \begin{pmatrix} C & D \\ D^T & E \end{pmatrix}. \tag{11}$$

Note that the submatrix C is shared with H_X. In our context the matrix might have been obtained by a previous calibration of a model that is part of both X and Y. If we were to condition on Z we would find

$$f(Y|Z) = N_{n-k-l}[D^T C^{-1} Z, H_Y/C](\bar{Y}), \tag{12}$$

with

$$H_Y/C = E - D^T C^{-1} D \tag{13}$$

and $\bar{Y}$ is the part of Y that is not Z. If we now want to describe X and Y together we are looking at the matrix

$$H = \begin{pmatrix} A & B & W \\ B^T & C & D \\ W^T & D^T & E \end{pmatrix}, \tag{14}$$

with a yet to be determined matrix $W \in M_{k-l,n-l}$. Because X and Y share the random variables in Z we can not make them independent. The next best thing that we can do is to demand that if we fix Z the remaining parts of X and Y are independent. The combined density for X and Y when we condition on Z is given by

$$f(X, Y|Z) = N_{n-l}[(B^T, D)^T C^{-1} Z, H/C](\bar{X}, \bar{Y}), \tag{15}$$

with

$$H/C = \begin{pmatrix} A - BC^{-1}B^T & W - BC^{-1}D \\ W^T - D^T C^{-1} B^T & E - D^T C^{-1} D \end{pmatrix} \tag{16}$$

$$= \begin{pmatrix} H_X/C & W - BC^{-1}D \\ W^T - D^T C^{-1} B^T & H_Y/C \end{pmatrix}. \tag{17}$$

We see that if we set

$$W = BC^{-1}D \tag{18}$$

the above matrix becomes block diagonal and the conditional densities obey

$$f(X,Y|Z) = f(X|Z)f(Y|Z) \tag{19}$$

because the blocks in the diagonal are H_X/C and H_Y/C from equations (8) and (13). Given that Z is part of both X and Y this is as much independence as we can ask for. The way to choose the matrix W and complete H is to demand that X and Y, when conditioned on their shared part Z, are independent.

This choice of W can be characterized in two more ways. Banachiewicz showed that the inverse of a matrix could be expressed using the Schur complement. Since the Schur complement H/C is block diagonal the inverse of H is given by

$$H^{-1} = \begin{pmatrix} (H_X/C)^{-1} & -(H_X/C)^{-1}BC^{-1} & 0 \\ -C^{-1}B^T(H_X/C)^{-1} & \Xi & -C^{-1}D(H_Y/C)^{-1} \\ 0 & -(H_Y/C)^{-1}D^TC^{-1} & (H_Y/C)^{-1} \end{pmatrix}, \tag{20}$$

with

$$\Xi = C^{-1} + C^{-1}B^T(H_X/C)^{-1}BC^{-1} + C^{-1}D(H_Y/C)^{-1}D^TC^{-1}. \tag{21}$$

We see that H^{-1} has zeroes in the places where W sits in H. The W from equation (18) is the unique choice with that property.

When looking at the determinant of H we find another characterization of W. Using the Schur complement again, we find

$$\det H = \det C \ \det H/C \tag{22}$$

$$\leq \det C \ \det H_X/C \ \det H_Y/C, \tag{23}$$

where we have equality in equation (23) if and only if the off-diagonal blocks vanish, i.e. if and only if W is given by equation (18). This choice of W is the unique choice that maximizes the determinant of H.

The entropy S of the normal distribution $N_n[\mu, H]$ is given by

$$S = \frac{1}{2}\log\det H + \text{ const.} \tag{24}$$

Because we are maximizing the determinant of H we are also maximizing the entropy S of the distribution described by H. The H chosen by the W from equation (18) is the one that puts the fewest constraints on the distribution of X and Y.

3 The completion procedure

In the last section we have completed the matrix H in the very simple case where there are two known matrices H_X and H_Y with just one shared part C. The general case is usually not this simple. Building on the result in [6] Smith [16] (see also [8] for a more general treatment of the semidefinite case) showed how to solve the general problem by reducing it to a repeated application of the result from the previous section. We will discuss the required steps in this section by looking at a particular example, the cross currency model.

Our model will consist of two interest rate models for the base currency E and the foreign currency A. These two currencies are linked by an exchange rate X. We will model the two interest rates and the exchange rate with stochastic volatility models so that we have a total of six stochastic variables:

$$E, \nu_E, A, \nu_A, X, \nu_X \tag{25}$$

We will denote the corresponding correlation coefficients by

$$(a, b), \tag{26}$$

where a and b are one of the stochastic variables from equation (25). In total there are 15 different correlation coefficients in the correlation matrix H of our model. We assume that the coefficients

$$(E, \nu_E) \text{ and } (A, \nu_A) \tag{27}$$

have already been determined in separate calibrations of the interest rate models. That leaves us with 13 free coefficients in H. This number of coefficients is too large for a stable calibration. Instead we will calibrate the coefficients

$$(E,A), (E,X), (A,X), (X, \nu_X), \tag{28}$$

and determine the remaining nine coefficients using the completion procedure (see figure 1).

The first step to complete the matrix H is to determine the graph corresponding to H. The vertices of the graph are given by the row (or column) indices of H. In our case these are identified with the stochastic variables in equation (25). We then obtain the graph for H by connecting the vertices for which we have coefficients. I n our example we obtain the graph in figure 2.

Now that we have the graph we need to determine if the graph is chordal (see e.g. [9]). A graph is chordal if every loop of length four or more has a chord (i.e. an edge connecting two non-consecutive vertices in the loop). In our example this is easy to check. The largest loop is the central loop consisting of the vertices E, A,

$$
\begin{pmatrix}
1 & \boxed{(x_E, \nu_E)} & (x_E, x_A) & (x_E, \nu_A) & (x_E, X_E^A) & (x_E, \nu_{EA}) \\
 & 1 & (\nu_E, x_A) & (\nu_E, \nu_A) & (\nu_E, X_E^A) & (\nu_E, \nu_{EA}) \\
 & & 1 & \boxed{(x_A, \nu_A)} & (x_A, X_E^A) & (x_A, \nu_{EA}) \\
 & & & 1 & (\nu_A, X_E^A) & (\nu_A, \nu_{EA}) \\
 & & & & 1 & (X_E^A, \nu_{EA}) \\
 & & & & & 1
\end{pmatrix}
$$

Figure 1: The correlation matrix for the cross currency model. The coefficients (x_E, ν_E) and (x_A, ν_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, ν_{EA}) (the boxes with the dashed lines) will be determined in the cross currency calibration itself. The rest will be determined using the completion procedure.

and X. This loop has just three elements and thus does not require a chord. The procedure that we are describing in this section requires a chordal graph.

The next step is to identify the cliques in our graph. A clique is a maximal set of vertices that are all connected to each other (the subgraph induced by the set of vertices is complete). In our example there are four such cliques given by the central loop and the three arms of the graph:

$$\alpha_C = \{E, A, X\} \tag{29}$$
$$\alpha_E = \{E, \nu_E\} \tag{30}$$
$$\alpha_A = \{A, \nu_A\} \tag{31}$$
$$\alpha_X = \{X, \nu_X\} \tag{32}$$

Note that the cliques have non-trivial intersections. We can use these intersections to construct a new graph. The vertices of this new graph are given by the cliques $\alpha_C, \ldots, \alpha_X$ and we connect two cliques if their intersection is not empty.

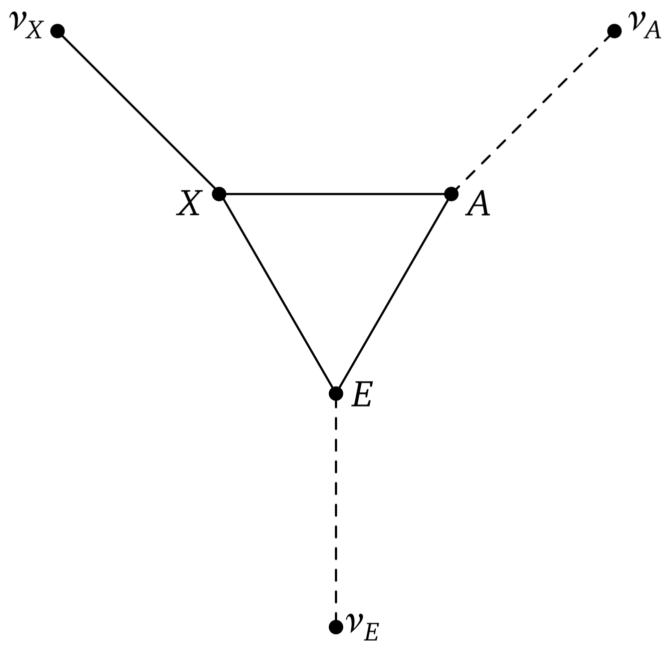

Figure 2: The initial graph for the cross currency calibration. The coefficients (E, v_E) and (A, v_A) (dashed edges) are determined in separate interest rate calibrations. The other edges are determined in the cross currency calibration itself.

4 Conclusion

References

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

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

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

[4] J. Schur, Über Potenzreihen, die im Innern des Einheitskreises beschränkt sind, Journal für die reine und angewandte Mathematik **147**, 205 – 232, 1917.

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

[6] R. Grone, C.R. Johnson, E. Sa, H. Wolkowicz, Positive definite completion of partial Hermitian matrices, Linear Algebra Appl. 58, 109–124, 1984.

[7] R. L. Smith, The positive definite completion problem revisited, Linear Algebra and its Applications 429 (2008) 1442 – 1452.

[8] our other paper ...

[9] M. C. Golumbic, Algorithmic Graph Theory and Perfect Graphs, Elsevier, 2004.

[10] Roger A. Horn, Charles R. Johnson, *Matrix Analysis*, 2nd Edition, Cambridge University Press, 2013.

[11] J. Schur, Über Potenzreihen, die im Innern des Einheitskreises beschränkt sind, Journal für die reine und angewandte Mathematik **147**, 205 – 232, 1917.

[12] T. Banachiewicz, Zur Berechnung der Determinanten, wie auch der Inversen, und zur darauf basierten Auflösung der Systeme linearer Gleichungen, Acta Astronom. Sér. C 3, 41–67, 1937.

[13] Frazer, R., Duncan, W., Collar, A. (1938). Elementary Matrices And Some Applications To Dynamics And Differential Equations, Cambridge University Press, 1938.

[14] 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

[15] 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

[16] 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

[17] 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

A Maybe ...

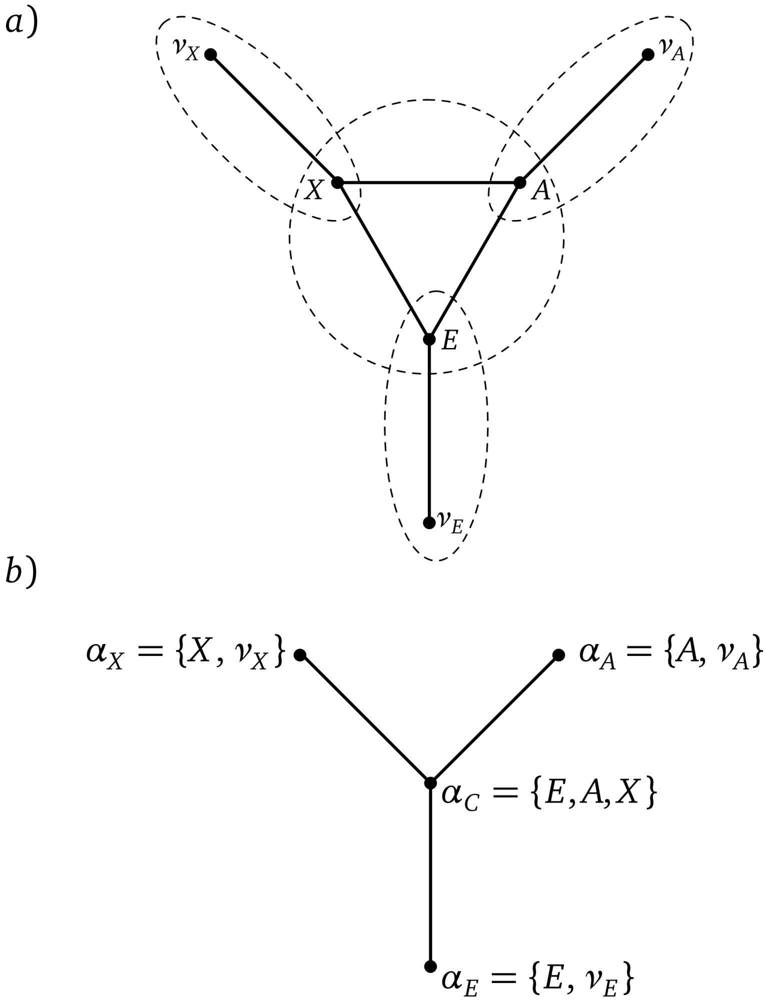

Figure 3: a) The graph for the cross currency has four cliques. b) These cliques are the vertices of a new graph. An edge in this new graph connects two cliques that have a non-empty intersection.