A Canonical Representation of Block Matrices with Applications to Covariance and Correlation Matrices[*]

Ilya Archakov[a] and Peter Reinhard Hansen[b†]

[a] *University of Vienna*

[b] *University of North Carolina & Copenhagen Business School*

July 24, 2020

Abstract

We obtain a canonical representation for block matrices. The representation facilitates simple computation of the determinant, the matrix inverse, and other powers of a block matrix, as well as the matrix logarithm and the matrix exponential. These results are particularly useful for block covariance and block correlation matrices, where evaluation of the Gaussian log-likelihood and estimation are greatly simplified. We illustrate this with an empirical application using a large panel of daily asset returns. Moreover, the representation paves new ways to regularizing large covariance/correlation matrices and to test block structures in matrices.

Keywords: Block Matrices, Block Covariance Matrix, Block Correlation Matrix, Equicorrelation, Covariance Regularization, Covariance Modeling, High Dimensional Covariance Matrices, Matrix Logarithm
JEL Classification: C10; C22; C58

---

[*]The second author would like to thank the Department of Statistics and Operations Research at University of Vienna for their hospitality during a visit in early 2020.

[†]Address: University of North Carolina, Department of Economics, 107 Gardner Hall Chapel Hill, NC 27599-3305

1 Introduction

We derive a canonical representation for a broad class of block matrices, that includes block covariance matrices. A special case of particular interest are block correlation matrices. The representation is a semi-spectral decomposition of block matrices, that are diagonalized with the exception of a single diagonal block, whose dimension is given by the number of blocks.

The canonical representation facilitates simple computations of several matrix functions, such as the matrix inverse, the matrix exponential, and the matrix logarithm. Consequently, the decomposition greatly simplifies the evaluation of Gaussian log-likelihood functions when the covariance matrix, or the correlation matrix, has a block structure.

We contribute to the literature on block correlation models by providing simple expressions for the inverse of any (invertible) block correlation matrix, as well as a simple expression for its determinant. This greatly eases the computational burden in the evaluation of a Gaussian (quasi-) log-likelihood function. The results apply to block correlation matrices with an arbitrary number of blocks. For block correlation matrices with two blocks, an expression for its inverse was obtained in Engle and Kelly (2012, lemma 2.3), and related results can be found in Viana and Olkin (1997).

As a preview of some of the results in this paper, one can consider the following $n \times n$ correlation matrix,

$$C = \begin{bmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho \\ \rho & \cdots & \rho & 1 \end{bmatrix}.$$

This correlation matrix is known as an equicorrelation matrix, and it is well known that its eigenvalues are $1+\rho(n-1)$ and $1-\rho$, where the latter has multiplicity $n-1$. This follows directly from the spectral decomposition,

$$Q'CQ = D = \begin{bmatrix} 1+\rho(n-1) & 0 \\ 0 & (1-\rho)I_{n-1} \end{bmatrix}, \tag{1}$$

where Q is an orthonormal matrix, so that $Q'Q = I_n$. Here I_n denotes the $n \times n$ identity matrix. The matrix Q is given by $Q = (v_n, v_{n\perp})$, where v_n is the n-dimensional vector, $v_n = (\frac{1}{\sqrt{n}}, \ldots, \frac{1}{\sqrt{n}})'$, and $v_{n\perp}$ is an $n \times (n-1)$ matrix that is orthogonal to v_n, i.e. $v_{n\perp}'v_n = 0$, and orthonormal, i.e. $v_{n\perp}'v_{n\perp} = I_{n-1}$.[1] It can now be verified that $QQ' = I$, so that Q is orthonormal and $C = QQ'CQQ' = QDQ'$. In this

[1]When $n = 1$, $v_{1\perp}$ is an 1×0 "matrix" and we use the conventions: $v_{1\perp}'v_{1\perp} = \emptyset$ (dimension 0×0) and $v_{1\perp}v_{1\perp}' = 0$ (dimension 1×1). This ensures that our expressions also hold in the special case, where one or more blocks has size one.

example, the matrix, D, is the canonical form of C, which is obtained via a rotation of C, where the rotation does not depend on ρ. In this example, where $K = 1$, D coincides with the diagonal matrix of eigenvalues in the spectral decomposition of C.

In this paper, we derive a similar decomposition for a broad class of block matrices that includes block covariance matrices and block correlation matrices. In the general case with multiple blocks, $K \geq 2$, the canonical representation does not fully disentangle all eigenvalues, and some eigenvalues may be complex-valued. The canonical representation decomposes any block matrix into a $K \times K$ matrix and $n - K$ real-valued eigenvalues, where K is the number of blocks. We can illustrate the general results with a 2×2 block correlation matrix,

$$C = \left[\begin{array}{c|c} C_{\rho_{11}} & \rho_{21}\mathbf{1}_{n_1 \times n_2} \\ \hline \bullet & C_{\rho_{22}} \end{array}\right]$$

where $C_{\rho_{11}}$ and $C_{\rho_{22}}$ are equicorrelation matrices with correlations ρ_{11} and ρ_{22}, respectively, and dimensions $n_1 \times n_1$ and $n_2 \times n_2$ respectively, and $\mathbf{1}_{n_1 \times n_2}$ is the $n_1 \times n_2$ whose elements are all equal to one. Now define

$$Q = \begin{bmatrix} v_{n_1} & 0 & v_{n_1\perp} & 0 \\ 0 & v_{n_2} & 0 & v_{n_2\perp} \end{bmatrix}.$$

For the equicorrelation matrix, we now have the following representation,

$$Q'CQ = \begin{bmatrix} 1 + \rho_{11}(n_1 - 1) & \rho_{12}\sqrt{n_1 n_2} & 0 & 0 \\ \rho_{12}\sqrt{n_1 n_2} & 1 + \rho_{22}(n_2 - 1) & 0 & 0 \\ 0 & 0 & (1 - \rho_{11})I_{n_1 - 1} & 0 \\ 0 & 0 & 0 & (1 - \rho_{22})I_{n_2 - 1} \end{bmatrix}. \quad (2)$$

We denote the upper-left 2×2 matrix by A. In general, A will be a $K \times K$ matrix, whose eigenvalues are also eigenvalues of C. The general result for block matrices with K blocks will be presented in Theorem 1, with a structure similar to that in (2). An important feature is that the matrix Q does not depend on the elements in block matrix, but is solely determined by the block partition, $(n_1, \ldots, n_K)$, where $n = n_1 + \cdots + n_K$.

The canonical representation is obtained for general block matrices that need not be symmetric, nor positive semidefinite. In fact, our results are applicable to non-square matrices. Block covariance matrices and block correlation matrices are interesting special cases. For block correlation matrices, the A-matrix, which emerges in (2), was previously established in Huang and Yang (2010) and Cadima et al. (2010), as we will discuss in Section 3. We derive additional results for block correlation matrices that simplify the evaluation of the log-likelihood function.

The rest of this paper is organized as follows. We present the main result in Section 2, where the canonical representation is established for a broad class of block matrices, along with some related results for the matrix exponential, matrix logarithm of block matrices, and matrix powers, including the matrix inverse. In Section 3, we consider the special case with block covariance matrices and block correlation matrices. Many of these results are useful for maximum likelihood estimation with a Gaussian log-likelihood function, as we show in Section 4. In Section 5, we apply the results to estimation of block covariance matrices for a very large panel of daily stock returns. We conclude in Section 6 and all proofs are presented in the Appendix.

2 Canonical Representation of Block Matrices

Let B be a square $n \times n$ matrix. The extension to rectangular matrices, which is trivial, will be addressed in the end of this section. The matrix, B, is called a block matrix with block partition, $n_1, \ldots, n_K$, if it can be expressed as:

$$B = \begin{bmatrix} B_{[1,1]} & B_{[1,2]} & \cdots & B_{[1,K]} \\ B_{[2,1]} & B_{[2,2]} & & \\ \vdots & & \ddots & \\ B_{[K,1]} & & & B_{[K,K]} \end{bmatrix},$$

where $B_{[i,j]}$ is an $n_i \times n_j$ matrix with the following structure

$$B_{[i,i]} = \begin{bmatrix} d_i & b_{ii} & \cdots & b_{ii} \\ b_{ii} & d_i & \ddots & \\ \vdots & \ddots & \ddots & \\ b_{ii} & & & d_i \end{bmatrix} \quad \text{and} \quad B_{[i,j]} = \begin{bmatrix} b_{ij} & \cdots & b_{ij} \\ \vdots & \ddots & \\ b_{ij} & & b_{ij} \end{bmatrix} \quad \text{if } i \neq j, \tag{3}$$

for some constants, d_i and b_{ij}, $i, j = 1, \ldots, K$. So the diagonal elements of the diagonal blocks, $B_{[i,i]}$, can take a different value than the off-diagonal elements, whereas all elements in an off-diagonal block, $B_{[i,j]}$, $i \neq j$, are identical.

We introduce the following notation that relates to orthogonal projections. Let $P_{[i,j]} = v_{n_i} v_{n_j}'$ be the $n_i \times n_j$ matrix whose elements are all equal to $\frac{1}{\sqrt{n_i n_j}}$. It is simple to verify that $P_{[i,k]} P_{[k,j]} = P_{[i,j]}$, and with $i = k = j$ it follows that $P_{[i,i]}P_{[i,i]} = P_{[i,i]}$, so that $P_{[i,i]}$ is a projection matrix. It then follows that $P_{[i,i]}^{\perp} = I_{n_i} - P_{[i,i]}$ is a projection matrix, and it can be verified that $P_{[i,i]}^{\perp} = v_{n_i\perp}v_{n_i\perp}'$, where the matrix, $v_{n\perp}$, was characterized in the introduction.

Finally, we define the $n \times n$ matrix

$$Q = \begin{bmatrix} v_{n_1} & 0 & \cdots & & v_{n_1\perp} & 0 & \cdots & 0 \\ 0 & v_{n_2} & & & 0 & v_{n_2\perp} & & \vdots \\ \vdots & & \ddots & & & & \ddots & \\ 0 & \cdots & & v_{n_K} & 0 & \cdots & & v_{n_K\perp} \end{bmatrix},$$

and observe that Q is an orthonormal matrix, characterized by the identity $Q'Q = I$. The first K columns of Q can be used to form averages within each of the K blocks, whereas the remaining columns of Q capture "differences" within each block. The two sets of columns span orthogonal subspaces that correspond to distinct components of the block decomposition. Note that Q is solely defined by the block partition, $n_1, \ldots, n_K$, and it is therefore invariant to the actual values taken by the elements in the block matrix.

Theorem 1. *Suppose that B is a block matrix with block partition $n_1, \ldots, n_K$. Then*

$$B_{[i,j]} = a_{ij}P_{[i,j]} + 1_{\{i=j\}}\lambda_i P_{[i,i]}^{\perp}, \qquad \text{for} \quad i, j = 1, \ldots, K,$$

where $a_{ij} = b_{ij}\sqrt{n_i n_j}$, for $i \neq j$, $a_{ii} = d_i + (n_i - 1)b_{ii}$, and $\lambda_i = d_i - b_{ii}$. Moreover,

$$B = QDQ', \qquad \text{with} \quad D = \begin{bmatrix} A & 0 & \cdots & 0 \\ 0 & \lambda_1 I_{n_1-1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_K I_{n_K-1} \end{bmatrix}. \tag{4}$$

The matrix Q rotates B into its canonical form, D. The first K columns of Q span an eigenspace of B, associated with the eigenvalues that A and B have in common. The last $n - K$ columns of Q are the remaining eigenvectors of B.

Theorem 1 can be used to characterize properties of B and simplifies the computation of some transformations of B, including the matrix logarithm of B, which is denoted by $\log B$. These results are stated in the following Corollary:

Corollary 1. *Suppose that B is a block matrix as defined above. (i) The eigenvalues of B are given by those of A as well as $\lambda_i = d_i - b_{ii}$, $i = 1, \ldots, K$, and $\det B = \det(A)\lambda_1^{n_1-1} \cdots \lambda_K^{n_K-1}$. (ii) B is invertible, if and only if A is invertible and $d_i \neq b_{ii}$, for all $i = 1, \ldots, K$. (iii) The q-th power of the block matrix, B^q, is well-defined whenever A^q and λ_i^q, $i = 1, \ldots, K$, are well-defined, in which case B^q has the same block structure as B, with blocks given by*

$$B^q_{[i,j]} = a_{ij}^{(q)} P_{[i,j]} + 1_{\{i=j\}} \lambda_i^q P^{\perp}_{[i,i]},$$

where a^q_{ij} is the ij-th element of A^q, for $i, j = 1, \ldots, K$. (iv) The matrix exponential of B has the same block structure as B, with blocks given by

$$\exp(B)_{[i,j]} = a_{ij}^{\exp} P_{[i,j]} + 1_{\{i=j\}} e^{\lambda_i} P^{\perp}_{[i,i]},$$

where $a_{ij}^{\exp}$ is the ij-th element of $\exp A$, for $i, j = 1, \ldots, K$. (v) If $\log A$ and $\log \lambda_i$, $i = 1, \ldots, K$, exist, then $\log B$ has the same block structure as B, with blocks given by

$$\log(B)_{[i,j]} = a_{ij}^{\log} P_{[i,j]} + 1_{\{i=j\}} \log \lambda_i P^{\perp}_{[i,i]},$$

where $a_{ij}^{\log}$ is the ij-th element of $\log A$.

It follows that B^q is well-defined for all positive integers of q, and the matrix inverse, B^{-1}, exists whenever A is invertible and $\lambda_i \neq 0$, for all $i = 1, \ldots, K$, in which case B^d is also well-defined for other negative integers of d. The logarithms, $\log A$ and $\log(d_k - b_{kk})$, exist provided that A is invertible and $d_k - b_{kk} \neq 0$. This may result in a complex-valued solution to the matrix logarithm. If a real-valued solution is required, then the conditions are that A is positive definite and that $d_k - b_{kk} > 0$ for all $k = 1, \ldots, K$.

2.1 Block Matrices with Kronecker Representation

Many of the expressions can be simplified further, in the special case, where all block sizes are identical, so that $n_1 = n_2 = \ldots = n_K = n$, with $n = N/K$. In this situation, we have $B = A \otimes P + \Lambda \otimes P_{\perp}$, where P is the $n \times n$ matrix with $1/n$ in all entries, $P_{\perp} = I_n - P$, and $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_K)$. In this case, it follows that $h(B) = h(A)P + h(\Lambda)P_{\perp}$, where $h(\cdot)$ represents the matrix inverse, the matrix exponential, or the matrix logarithm, provided these are well-defined.

2.2 Rectangular Block Matrices

Suppose that B has blocks, $B_{[i,j]} \in \mathbb{R}^{n_i \times n_j}$, as specified in (3), where $i = 1, \ldots, K_1$ and $j = 1, \ldots, K_2$, and $K_1 \neq K_2$, so that B is a non-square matrix. Set $K = \max(K_1, K_2)$ and suppose that $K_1 > K_2$. Then, by appending blocks with zero elements to B, we obtain a square matrix, $\tilde{B} = (B, 0)$, which is a block matrix with block partition, $n_1, \ldots, n_K$. Our results apply to $\tilde{B}$, so that it has the canonical form $\tilde{B} = QDQ'$ and $B = QD\tilde{Q}'$, where $\tilde{Q}$ is made up of the first $n_1 + \cdots + n_{K_2}$ columns of Q. If $K_2 > K_1$, we can instead define $\tilde{B} = (B', 0)'$, and the results follow similarly.

3 Block Correlation Matrices

A block correlation matrix is characterized by the correlation coefficients that form a block structure, so that the correlation between two variables is solely determined by the blocks to which the two variable belong. This results in a correlation matrix with a common correlation coefficient within each block.

Block correlation matrices offer a way to parameterize large covariance matrices in a parsimonious manner, and can be used to impose economically relevant structures that reduce the complexity of the covariance matrix. This structure is used in some multivariate GARCH models, see Engle and Kelly (2012) and Archakov et al. (2020).

An $n \times n$ block correlation matrix, C, with K blocks, is a symmetric block matrix with blocks,

$$C_{[i,i]} = \begin{bmatrix} 1 & \rho_{ii} & \cdots & \rho_{ij} \\ \rho_{ii} & 1 & \ddots & \\ \vdots & \ddots & \ddots & \\ \rho_{ii} & & & 1 \end{bmatrix} \quad \text{and for } i \neq j, \quad C_{[i,j]} = \begin{bmatrix} \rho_{ij} & \cdots & \rho_{ij} \\ \vdots & \ddots & \\ \rho_{ij} & & \rho_{ij} \end{bmatrix}, \tag{5}$$

where ρ_{ii} is within-block correlations, and $\rho_{ij} = \rho_{ji}$, $i \neq j$, are between-block correlations, for $i, j = 1, \ldots, K$. For C to be a correlation matrix, we obviously need $\rho_{ij} \in [-1, 1]$ for all $i, j = 1, \ldots, K$. However, this alone is not sufficient to produce a valid correlation matrix, because negative eigenvalues can arise with some combinations of correlation coefficients, even if these area all strictly smaller than one.

The case with block equicorrelation matrices corresponds to the case where the diagonal elements of all diagonal blocks, $B_{[kk]}$ equal $d_k = 1$, for all $k = 1, \ldots, K$. So Theorem 1 fully characterizes the set of correlation coefficients that yields a positive (semi-) definite correlation matrix. We formulate this result as a separate Corollary. Note that the canonical form, (4), for C in (5), is such that A is symmetric with elements given by $a_{ij} = \rho_{ij}\sqrt{n_i n_j}$, for $i \neq j$, $a_{ii} = 1 + \rho_{ii}(n_i - 1)$, and $\lambda_i = 1 - \rho_{ii}$.

Corollary 2 (Block correlation matrices)**.** *Let C be a block correlation matrix. Then*

$$\det C = \det A \cdot \prod_{i=1}^{K}(1-\rho_{ii})^{n_i - 1},$$

so that C is a non-singular block correlation matrix if and only if A is positive definite and $|\rho_{ii}| < 1$. In this case, both the inverse correlation matrix, C^{-1}, and the matrix logarithm, $\log C$, have the same block structure as C, with blocks given by

$$C^{-1}_{[i,j]} = a^{\#}_{ij} P_{[i,j]} + 1_{\{i=j\}} \tfrac{1}{1-\rho_{ii}} P^{\perp}_{[i,i]},$$

and

$$\log(C)_{[i,j]} = \tilde{a}_{ij} P_{[i,j]} + 1_{\{i=j\}} \log(1-\rho_{ii}) P^{\perp}_{[i,i]},$$

respectively, where $a^{\#}_{ij}$ is the ij-th element of A^{-1} and $\tilde{a}_{ij}$ is the ij-th element of $\log A$.

The conditions for C in (5) to be a (possibly singular) correlation matrix is that A is positive semi-definite and $|\rho_{ii}| \leq 1$. So, Corollary 2 characterizes the set of positive definite block equicorrelation matrices, where the additional requirements are that A is positive definite and $|\rho_{ii}| < 1$.

In this context with block correlation matrices, the expression for A was previously obtained in Huang and Yang (2010, proposition 5) and in Cadima et al. (2010, theorem 3.1). The focus in Huang and Yang (2010) was on computational issues, which might explain that their paper is overlooked in much of the literature.[2] Their results add valuable insight about the block-DECO model by Engle and Kelly (2012). For instance, their results provide a simple way to evaluate if a block matrix a positive definite (or semidefinite) correlation matrix. The expression for the determinant of a correlation matrix in Corollary 2 is a simple implication of the eigenvalues derived in Huang and Yang (2010) and Cadima et al. (2010), whereas the expressions for the inverse and logarithmically transformed correlation matrices are new.

[2] We were, until recently, also unaware of the results in Huang and Yang (2010) and Cadima et al. (2010). An anonymous referee (on a different paper than the present one) directed us to Roustant and Deville (2017) and we subsequently discovered the more detailed results in Huang and Yang (2010) and Cadima et al. (2010). Some of their results, e.g. Huang and Yang (2010, eq. 6), were rediscovered in Roustant and Deville (2017), who do not cite Huang and Yang (2010) or Cadima et al. (2010). In fact, none of the papers, Cadima et al. (2010), Huang and Yang (2010), Engle and Kelly (2012), and Roustant and Deville (2017) cite any of the other papers listed here.

3.1 Parametrization of Block Correlation Matrices

A new parametrization of correlation matrices was introduced in Archakov and Hansen (2020). The new parametrization consists of the elements below the diagonal of $\log C$ (matrix logarithm of C). Let ϱ denote the vector with these $n(n-1)/2$ elements, where n is the dimension of C ($n \times n$). Archakov and Hansen (2020) showed that the $\varrho = \varrho(C)$ is a one-to-one mapping between the set of non-singular correlation matrices and $\mathbb{R}^{n(n-1)/2}$.

For a block equicorrelation matrix, C, it follows (from Corollary 2) that $\log C$ has the same block structure as C. So, for $i \neq j$, all elements in $[\log C]_{i,j}$ are identical and given by $\frac{\tilde{a}_{ij}}{\sqrt{n_i n_j}}$, and the off-diagonal elements of the diagonal blocks, $[\log C]_{k,k}$, $k = 1, \ldots, K$, are all equal to $\frac{\tilde{a}_{kk} - \log(1-\rho_{kk})}{n_k}$, where $\tilde{a}_{ij}$ are the elements of $\log A$. Thus, the unique elements of $[\log C]$ are

$$\Lambda_n^{-1}[\log A - \log \Lambda_{1-\rho}]\Lambda_n^{-1} = \Lambda_n^{-1}[\log(\Lambda_n R \Lambda_n + \Lambda_{1-\rho}) - \log \Lambda_{1-\rho}]\Lambda_n^{-1},$$

where

$$R = \begin{bmatrix} \rho_{11} & \cdots & \rho_{1K} \\ \vdots & \ddots & \\ \rho_{K1} & & \rho_{KK} \end{bmatrix} \qquad \Lambda_n = \begin{bmatrix} \sqrt{n_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{n_K} \end{bmatrix}, \quad \Lambda_{1-\rho} = \begin{bmatrix} 1-\rho_{11} & & 0 \\ & \ddots & \\ 0 & & 1-\rho_{KK} \end{bmatrix}.$$

4 Applications of the Canonical Representation to Gaussian Log-Likelihood

In this section, we focus on covariance and correlation matrices for normally distributed random variables. We derive simplified expressions for the corresponding log-likelihood functions, that greatly reduce the computational burden when n is large relative to K. We derive the maximum likelihood estimators, and provide a simple expression for the first derivatives of the log-likelihood function with respect to the unknown parameters (the scores).

We will follow the conventional notation for covariances and variances, we write σ_{ij} in place of b_{ij}, $i, j = 1, \ldots, K$, and σ_k^2 in place of d_k, $k = 1, \ldots, K$. Similarly, for correlation matrices we write ρ_{ij} in place of b_{ij}, and observe that $d_k = 1$.

The density function for the multivariate Gaussian distribution with mean zero and an $n \times n$ covariance matrix, Σ, is $f(x) = (2\pi)^{-\frac{n}{2}}(\det \Sigma)^{-\frac{1}{2}} \exp(-\frac{1}{2}x'\Sigma^{-1}x)$. Suppose that Σ has the block structure given by $(n_1, \ldots, n_K)$, so that it can be expressed as $\Sigma = QDQ'$, using the canonical representation.

The corresponding log-likelihood function (multiplied by -2) can now be expressed as

$$-2\ell = n \log 2\pi + \log\det D + X'QD^{-1}Q'X,$$

where $D = \mathrm{diag}(A, \lambda_1 I_{n_1-1}, \ldots, \lambda_K I_{n_K-1})$, with $\lambda_i = \sigma_i^2 - \sigma_{i,i}$.

$$a_{ij} = \begin{cases} \sigma_i^2 + (n_i - 1)\sigma_{i,i} & \text{for } i = j, \\ \sigma_{i,j}\sqrt{n_i n_j} & \text{for } i \neq j. \end{cases}$$

So, if we define $Y = (y_0', y_1', \ldots, y_K')' = Q'X$, where y_0 is K-dimensional and y_k is $n_k - 1$ dimensional, $k = 1, \ldots, K$, then it follows that

$$-2\ell = n \log 2\pi + \log\det A + y_0' A^{-1} y_0 + \sum_{k=1}^{K} \left((n_k - 1) \log \lambda_k + \tfrac{y_k' y_k}{\lambda_k} \right). \tag{6}$$

This expression of the log-likelihood function shows that the block structure yields a considerable simplification in the evaluation of the log-likelihood. Instead of inverting the $n \times n$ matrix Σ and computing $\det \Sigma$, it suffices to invert the smaller $K \times K$ matrix, A, and evaluate its determinant. Moreover, the maximum likelihood estimator based on a random sample, $X_1, \ldots, X_N$, is easily expressed in terms of the transformed variables, $Y_1 = Q'X_1, \ldots, Y_N = Q'X_N$, as formulated in the following Theorem.

Theorem 2. *Suppose that $X_1, \ldots, X_N$ are independent and identically distributed as $N(0, \Sigma)$, where Σ is a block covariance matrix with block partition, $n_1, \ldots, n_K$. Define the transformed variables, $Y_s = Q'X_s$, $s = 1, \ldots N$, where $Y_s = (y_{0,s}', y_{1,s}', \ldots, y_{K,s}')'$ and where $y_{0,s}$ is K-dimensional and $y_{k,s}$ is $n_k - 1$ dimensional, $k = 1, \ldots, K$.*

The maximum likelihood estimator of Σ is given by $\hat{\Sigma} = Q\hat{D}Q'$, where $\hat{D} = \mathrm{diag}(\hat{A}, \hat{\lambda}_1 I_{n_1-1}, \ldots,$ $\hat{\lambda}_K I_{n_K-1})$ with

$$\hat{A} = \frac{1}{N} \sum_{s=1}^{N} y_{0,s} y_{0,s}' \quad \textit{and} \qquad \hat{\lambda}_k = \frac{1}{N} \sum_{s=1}^{N} \tfrac{y_{k,s}' y_{k,s}}{n_k - 1}, \quad k = 1, \ldots, K.$$

The maximum likelihood estimates of the individual parameters can be obtained directly from $\hat{A}$ and $\hat{\lambda}_k$, $k = 1, \ldots, K$. For $i \neq j$, it follows from the definition of A that $\hat{\sigma}_{i,j} = \hat{a}_{ij}/\sqrt{n_i n_j}$. For $i = j$, we have $\hat{\sigma}_{i,i} = (\hat{a}_{ii} - \hat{\lambda}_i)/n_i$ and $\hat{\sigma}_i^2 = \hat{\lambda}_i + \hat{\sigma}_{i,i} = \frac{1}{n_i}\hat{a}_{ii} + \frac{n_i - 1}{n_i}\hat{\lambda}_i$.

In the special case where a block has size one, we have $\Sigma_{[k,k]} = \sigma_k^2$ and $\sigma_{k,k}$ is obviously undefined.

In this situation, the corresponding variables, $y_{k,s}$, $s = 1, \dots, N$, are also undefined, and hence, so is $\hat{\lambda}_k$. Yet the expressions for the maximum likelihood estimators continue to be valid, including the expression for $\hat{\Sigma}$ in Theorem 2. If $n_k = 1$, then $\hat{\sigma}_k^2 = \hat{a}_{kk}$, while the expression for $\hat{\sigma}_{k,k}$ is undefined and can be ignored.

Estimation when the correlation matrix is assumed to have a block structure, as opposed to the covariance matrix, is similar. However, a block correlation matrix is entirely given by the A-matrix, and computation of the eigenvalues $\lambda_1, \dots, \lambda_K$ is redundant.

Corollary 3. *Suppose that $X_1, \dots, X_N$ are independent and identically distributed as $N_n(0, \Sigma)$, where $\Sigma = \Lambda_\sigma C \Lambda_\sigma$ with $\Lambda_\sigma = \mathrm{diag}(\sigma_1, \dots, \sigma_n)$ and C is a block correlation matrix with block partition, $n_1, \dots, n_K$. The maximum likelihood estimates of the diagonal elements of Σ are given by $\hat{\sigma}_i^2 = N^{-1} \sum_{s=1}^N X_{i,s}^2$, for $i = 1, \dots, n$. Next define $\tilde{X}_{i,s} = X_{i,s}/\hat{\sigma}_i$, and introduce the transformed variables, $\tilde{Y}_s = Q' \tilde{X}_s$, $s = 1, \dots N$, where $\tilde{Y}_s = (\tilde{y}'_{0,s}, \tilde{y}'_{1,s}, \dots, \tilde{y}'_{K,s})'$ and where $\tilde{y}_{0,s}$ is K-dimensional and $\tilde{y}_{k,s}$ is $n_k - 1$ dimensional, $k = 1, \dots, K$.*

The maximum likelihood estimator of C is given by $\hat{C} = Q\tilde{D}Q'$, where $\tilde{D} = \mathrm{diag}(\tilde{A}, \tilde{\lambda}_1 I_{n_1 - 1}, \dots, \tilde{\lambda}_K I_{n_K - 1})$ with

$$\tilde{A} = \frac{1}{N} \sum_{s=1}^{N} \tilde{y}_{0,s} \tilde{y}'_{0,s} \quad \text{and} \qquad \tilde{\lambda}_k = \frac{n_k - \tilde{a}_{kk}}{n_k - 1}, \quad k = 1, \dots, K.$$

So the estimate of D can be obtained solely from $\tilde{A}$. For the individual correlations we have $\hat{\rho}_{i,j} = \tilde{a}_{ij}/\sqrt{n_i n_j}$, for $i \neq j$, and for $i = j$, we have $\hat{\rho}_{i,i} = (\tilde{a}_{ii} - 1)/n_i$.

The score of the log-likelihood function is often of separate interest. For instance, the score is used for the computation of robust standard errors, in Lagrange multiplier tests, in tests for structural breaks, see e.g. Nyblom (1989), and in dynamic models with time-varying parameters (the so-called score-drive models), see Creal et al. (2013). So we provide the expressions for the score in this context with a block covariance matrix.

Suppose that Σ is a block covariance matrix, and consider its canonical representation $\Sigma = QDQ'$. Since Q is entirely given by the block partition $(n_1, \dots, n_K)$, and does not depend on the unknown parameters in Σ, the expressions for the partial derivatives are relatively simple.

Proposition 1. *Let $\Sigma = QDQ$ be the canonical representation of Σ. Then $\partial(-2\ell)/\partial A = M =$*

$A^{-1} - A^{-1}y_0y_0'A^{-1}$ *and for,* $k = 1, \ldots, K$, *we have*

$$\begin{aligned}
\frac{\partial(-2\ell)}{\partial\sigma_k^2} &= M_{k,k} + \left(\tfrac{n_k-1}{\lambda_k} - \tfrac{y_k'y_k}{\lambda_k^2}\right) \\
\frac{\partial(-2\ell)}{\partial\sigma_{kk}} &= (n_k-1)M_{k,k} - \left(\tfrac{n_k-1}{\lambda_k} - \tfrac{y_k'y_k}{\lambda_k^2}\right),
\end{aligned}$$

and, for $i \neq j$, *we have* $\frac{\partial(-2\ell)}{\partial\sigma_{ij}} = 2\sqrt{n_i n_j}M_{i,j}$.

The hessian could be derived similarly. In some applications, it might be preferable to parametrize the block covariance matrix with A and $(\lambda_1, \ldots, \lambda_K)$. In this case, one can use $\partial(-2\ell)/\partial A = M$, and $\partial(-2\ell)/\partial\lambda_k = \frac{n_k-1}{\lambda_k} - \frac{y_k'y_k}{\lambda_k^2}$, for $k = 1, \ldots, K$.

5 Empirical Estimation of Block Correlation Matrices

We proceed to illustrate how high-dimensional covariance matrices with a block structure are straightforward to estimate in practice. We estimate block structures for a large panel of assets for two calendar years, 2008 and 2013, using daily returns. We included all stocks in the CRSP database that could be matched with a unique ticker symbol, and which did not have any missing observations. This resulted in 3958 assets in 2008 and 2998 assets in 2013. The objective of this empirical application is demonstrate that high-dimensional covariance matrices can be estimated with relatively few observation once block structures are imposed, and that the canonical representation makes it simple to evaluate the log-likelihood function and to obtain the maximum likelihood estimates. Given the well-known variation in conditional variances and covariances, our estimated covariance matrices should be viewed as estimates of the average covariance matrix for 2008 and 2013, rather than an accurate description of the data generating process.

We impose five nested structures on the correlation matrix, where the equicorrelation structure ($K = 1$) is the simplest and most restrictive model. The remaining four correlation models use block structures defined by the Sector, Group, Industry, and Sub-Industry categories, as classified by the Global Industry Classification Standard (GICS) in 2013. The five specifications correspond to $K = 1$, 10, 24, 67, and 151, respectively, in 2008, and the same number of blocks in 2013, except for Sub-Industry categories, which had $K = 146$.

We estimated the canonical correlation matrix using the results in Corollary 3. Thus, we first compute the estimates of the variances, $\hat{\sigma}_i^2 = \sum_{t=1}^N X_{i,t}^2$, for each of the individual assets. Then we define the standardized variables, $\tilde{X}_{i,t} = X_{i,t}/\hat{\sigma}_i$, and $\tilde{y}_{0,t} \in \mathbb{R}^K$, whose k-th element is given by

$\frac{1}{\sqrt{n_k}}\sum_{i=n_1+\cdots+n_{k-1}+1}^{n_1+\cdots+n_k}\tilde{X}_{i,t}$. From Corollary 3, we have $\hat{\rho}_{i,j}=\tilde{a}_{ij}/\sqrt{n_in_j}$, for $i\neq j$, and $\hat{\rho}_{i,i}=(\tilde{a}_{ii}-1)/n_i$, where $\tilde{A}=\frac{1}{N}\sum_{t=1}^{N}\tilde{y}_{0,t}\tilde{y}_{0,t}'$. Thus, the entire $n\times n$ covariance matrix with a block correlation structure is estimated by computing the estimates of the n variances, and the $K\times K$ matrix $\tilde{A}$. Given the high number of assets and just over 250 daily returns, the unrestricted sample covariance matrix would be singular, because most of its eigenvalues will be zero. Once a block structure is imposed, we can compute the inverse covariance matrix from the invertible $K\times K$ matrix, $\tilde{A}$, and it becomes simple to evaluate the log-likelihood function.

The empirical results are summarized in Table 1. We report the range of estimated correlations for each of the block structures, with the results for 2008 and 2013 are reported separately. The range of estimated correlations. i.e. the interval between the smallest and the largest coefficient in the correlation matrix, obviously increases with the number of blocks in the correlation matrix, and the correlations are generally higher in 2008 than in 2013. These estimates are likely biased, because they entail cross-sectional averaging within each sector/group/industry and time averaging, over a full calendar year. We also report the value of the maximized log-likelihood function (scaled by $-2/(nN)$) and the corresponding value of the Bayesian Information Criterion (BIC). The minimum BIC is obtained with a block structures based on Groups in both 2008 and 2013.[3] The last column reports the number $K(K+1)/2$ of unique correlations within a block structure with K blocks, and while this number increases rapidly with K, the gains in the log-likelihood are relatively modest. Consequently, the BIC increases substantially once the number of blocks are defined by Industries and Sub-industries.

The estimated block structures are illustrated in Figures 1 and 2. The upper panels of Figure 1 show the estimated correlation coefficients of assets within and between sectors. The lower panels are the estimates for the 24 groups, with the actual estimates indicated by color coding. A darker shade of red denotes a stronger correlation. The left panels are for 2008 and the right panels are for 2013. Figure 2 presents the estimated correlations using a block structure based on industries and sub-industries. We observe that the correlations were generally higher in 2008 than in 2013, in part because of the turmoil period leading up to, and following, the collapse of Lehman Brothers in late 2008. The block structure is perhaps more visible in 2008, which might be explained by the Global Financial Crises having a differentiated impact on different sections. For instance, Figure 1 shows that the correlations between the Energy (10), Materials (15), and Utilities (55) sectors were relatively high, while Financials (40) were relatively uncorrelated with other sectors in 2008. The partition by the GICS groups in the lower

[3] The BIC adds the penalty $p\log(nN)$ to -2ℓ, where p is the number of free parameters. For comparison, the AIC, which uses the penalty $2p$, selects the most general specification in both years. It is well known that the AIC tends to favor more heavily parametrized models.

panels of Figure 1 reveals additional details about the correlations. For financials (40), we observe that the low correlations with other sectors are largely driven by Banks (4010).[4] Figure 2 presents the corresponding results for industries and sub-industries. The number of blocks are too plentiful to be listed individually, however the industries and sub-industries are placed chronologically in the ascending order in Figure 2 according to their GICS code.

6 Concluding Remarks

We have derived a canonical representation of block matrices, that is particularly useful for covariance and correlation matrices. We derived a number of expressions that greatly simplify the computation of the log-likelihood function. We illustrated this in an empirical application, where we estimated the covariance matrix for nearly 4000 stocks returns using daily returns from a single calendar year, i.e. just over 250 observations. Inverting the covariance matrix, and evaluating the log-likelihood is straightforward once a block structure is imposed, where we used as many as $K = 151$ blocks, motivated by the Global Industry Classification Standard.

The canonical representation and the related results are potentially useful for regularizing large covariance matrices. For instance, one could shrink the sample correlation matrix towards a block correlation matrix, analogous to the way Ledoit and Wolf (2004) proposed to shrink towards the equicorrelation matrix. The canonical representation also paves new way to testing block structures in covariance and correlation matrices. This predominantly amounts to testing a large number of zero-restrictions in the canonical representation. We identified a number of transformations that preserves the block structures, so testing of block structures could be based on any of the transformations, rather than the original matrix. For instance, block structures in a correlation matrix C would be tested on the canonical representation for $\log C$. This is potentially interesting, because the connection between logarithmically transformed correlation matrix and the Fisher transformation, see Archakov and Hansen (2020). Finally, the group assignments, and hence K, will be unknown in many empirical applications. The literature has therefore proposed various techniques that aim to determine the most appropriate block structure. It is possible that the canonical representation will be useful for this type of model selection problem.

[4] Note that we are using the GICS classification as it was in 2013, where Real Estate (4040) was a part of the Financials Sector. A separate Real Estate Sector (50) was added to the GICS classification in 2016.

Appendix of Proofs

Proof of Theorem 1. For $i \neq j$, we have $B_{[i,j]} = a_{ij}P_{[i,j]}$ if $a_{ij} = b_{ij}\sqrt{n_i n_j}$, since the elements of $P_{[i,j]}$ are all equal to $\frac{1}{\sqrt{n_i n_j}}$. For $i = j$, the diagonal elements differs from off-diagonal elements by $\lambda_i = d_i - b_{ii}$, so that $B_{[i,i]} = b_{ii}n_i P_{[i,i]} + (d_i - b_{ii})I_{n_i}$. Since $I_{n_i} = P_{[i,i]} + P_{[i,i]}^{\perp}$, we have $B_{[i,i]} = (b_{ii}n_i + d_i - b_{ii})P_{[i,i]} + (d_i - b_{ii})P_{[i,i]}^{\perp} = a_{ii}P_{[i,i]} + \lambda_i P_{[i,i]}^{\perp}$. The canonical representation, (4), follows by verifying that $Q'BQ$ is equal to the block-diagonal matrix in (4). This follows from the identities: $v_{n_i}' P_{[i,j]} v_{n_j} = 1$, $v_{n_i}' P_{[i,j]} v_{n_j\perp} = 0$, $v_{n_i\perp}' P_{[i,j]} v_{n_j\perp} = 0$, $v_{n_i}' P_{[i,i]}^{\perp} v_{n_i} = 0$, $v_{n_i}' P_{[i,i]}^{\perp} v_{n_i\perp} = 0$, and $v_{n_i\perp}' P_{[i,i]}^{\perp} v_{n_i\perp} = I_{n_i-1}$, and the fact that $Q'Q = I_n$, so that $Q^{-1} = Q'$, and hence $B = QQ'BQQ'$. This proves (4). □

Proof of Corollary 1. The first result for the eigenvalues of B and the determinant of B, follows immediately from (4). The results for $f(B)$, where f denotes the q-th power of a matrix, the matrix exponential, or the matrix logarithm, follow by $f(B) = Qf(D)Q'$ and using the structure in Q, such as $v_{n_i} v_{n_j}' = P_{[ij]}$ and $v_{n_{i\perp}} v_{n_i\perp}' = P_{[ii]}^{\perp}$. This completes the proof. □

Proof of Corollary 2. It follows from Theorem 1 and Corollary 1 by setting $d_k = 1$ for all k. Some expressions can also be verified directly. For instance, one can verify the expression for C^{-1} , by noting that diagonal blocks of C^{-1} are given by

$$(C^{-1})_{[i,i]} = \sum_{k=1}^{K} a_{ik} P_{[ik]} a_{ki}^{\#} P_{[k,i]} + (1 - \rho_{ii}) P_{[i,i]}^{\perp} \tfrac{1}{1-\rho_{ii}} P_{[i,i]}^{\perp} = \sum_{k=1}^{K} a_{ik} a_{ki}^{\#} P_{[i,i]} + P_{[i,i]}^{\perp} = I,$$

where we used that $a_{ki}^{\#}$ are the elements of the A^{-1} so we have $\sum_{k=1}^{K} a_{ik} a_{ki}^{\#} = 1$. Next, for $i \neq j$, we have

$$(C^{-1})_{[i,j]} = \sum_{k=1}^{K} a_{ik} P_{[i,k]} a_{kj}^{\#} P_{[k,j]} + \frac{a_{ij}}{b_j} P_{[i,j]} P_{[j,j]}^{\perp} + b_i a_{ij}^{\#} P_{[i,i]}^{\perp} P_{[i,j]} = \sum_{k=1}^{K} a_{ik} a_{kj}^{\#} P_{[i,j]} = 0,$$

where we used that $P_{[i,k]} P_{[k,j]} = P_{[i,j]}$ and $P_{ij} P_{[j,j]}^{\perp} = P_{[i,j]}(I_{s_j} - P_{[j,j]}) = 0$, and that $\sum_{k=1}^{K} a_{ik} a_{kj}^{\#} = 0$, for $i \neq j$. This completes the proof. □

Proof of Theorem 2. The expression, (6), shows that the log-likelihood function is made up of two terms:

$$-2N \left[\log \det A + \mathrm{tr}\{A^{-1} \tfrac{1}{N} \sum_{s=1}^{N} Y_{0,s} Y_{0,s}'\} \right],$$

and

$$-2N \sum_{k=1}^{K} (n_k - 1) \left(\log \lambda_k + \frac{\frac{1}{N} \sum_{s=1}^{N} \frac{Y_{k,s}' Y_{k,s}}{n_k - 1}}{\lambda_k} \right).$$

It is well known that $\hat{A} = \frac{1}{N}\sum_{s=1}^{N} Y_{0,s}Y_{0,s}'$ maximizes the first term and that $\hat{\lambda}_k = \frac{1}{N}\sum_{s=1}^{N}\frac{Y_{k,s}'Y_{k,s}}{n_k-1}$ maximizes the elements of the second term. Since $(A, \lambda_1, \ldots, \lambda_K)$ is merely a reparametrization of the elements of the block covariance matrix Σ, it follows that $\hat{\Sigma} = Q\hat{D}Q'$ is the maximum likelihood estimator of Σ. It is easy to verify that this result is also valid in the special case, where one or more of the blocks are 1-dimensional. In this case, σ_{ii} is undefined, and so is $\hat{\lambda}_k$. In this case, σ_i^2 is identified from the corresponding diagonal element of A, since $\hat{a}_{ii} = \sigma_i^2$, when $n_i = 1$. □

Proof of Corollary 3. We have $\det\Sigma = \det(\Lambda_\sigma C\Lambda_\sigma) = \prod_{i=1}^{n}\sigma_i^2 \det C = \prod_{i=1}^{n}\sigma_i^2 \det D$. So the log-likelihood function for the observation, $X \in \mathbb{R}^n$, is

$$-2\ell(\sigma_1^2, \ldots, \sigma_n^2, C) = n\log 2\pi + \sum_{j=1}^{n}\log\sigma_j^2 + \log\det C + X'\Lambda_\sigma^{-1}C\Lambda_\sigma^{-1}X,$$

and given a sample $X_1, \ldots, X_N$, the first order condition for σ_j^2:

$$0 = \tfrac{N}{\sigma_j^2} - \sum_{t=1}^{N}\tfrac{1}{\sigma_j^4}X_t'e_je_j'Ce_je_j'X_t = \tfrac{N}{\sigma_j^2} - \tfrac{1}{\sigma_j^4}\sum_{t=1}^{N}X_t'e_j1e_j'X_t = \tfrac{N}{\sigma_j^2} - \tfrac{1}{\sigma_j^4}\sum_{t=1}^{N}X_{j,t}^2,$$

is invariant to C. So that $\hat{\sigma}_j^2 = \frac{1}{N}\sum_{t=1}^{N}X_{j,t}^2$ is the maximum likelihood estimator of σ_j^2 regardless of the structure imposed on C. The concentrated log-likelihood function for the observation $X \in \mathbb{R}^n$ can be expressed as

$$-2\ell(C) \equiv -2\ell(\hat{\sigma}_1^2, \ldots, \hat{\sigma}_n^2, C) = n\log 2\pi + \sum_{j=1}^{n}\log\hat{\sigma}_j^2 + \log\det D + \tilde{X}'Q'DQ\tilde{X},$$

and it follows from the proof of Theorem 2 that minimizing $N\log\det D + \sum_{t=1}^{N}\tilde{X}_s'Q'DQ\tilde{X}_t$ is solved by the D-matrix whose elements are given by $\tilde{A} = \frac{1}{N}\sum_{t=1}^{N}\tilde{y}_{0,t}\tilde{y}_{0,t}'$ and $\tilde{\lambda}_k = \frac{1}{N}\sum_{t=1}^{N}\frac{\tilde{y}_{k,t}'\tilde{y}_{k,t}}{n_k-1}$, for $k = 1, \ldots, K$. For $Q\tilde{D}Q'$ to be a correlation matrix (have ones along its diagonal), we need $\tilde{a}_{ii} = 1+(n_i-1)\hat{\rho}_{ii}$ and $\tilde{\lambda}_i = 1 - \hat{\rho}_{ii}$, which implies $\tilde{\lambda}_i = \frac{n_i - \tilde{a}_{ii}}{n_i-1}$. □

Proof of Proposition 1. Recall that $a_{kk} = \sigma_k^2 + (n_k-1)\sigma_{kk}$, $a_{ij} = \sigma_{ij}\sqrt{n_in_j}$, for $i \neq j$, and $\lambda_k = \sigma_k^2 - \sigma_{kk}$. It follows that

$$\frac{\partial(\log\det A + y_0'A^{-1}y_0)}{\partial a_{ij}} = \mathrm{tr}\{A^{-1}(e_ie_j')(I - A^{-1}y_0y_0') = e_j'(I - A^{-1}y_0y_0')A^{-1}e_i = M_{j,i},$$

where $M = A^{-1} - A^{-1} y_0 y_0' A^{-1}$. From the expression (6), we find

$$\begin{aligned}
\frac{\partial(-2\ell)}{\partial\sigma_k^2} &= \frac{\partial(\log\det A + y_0' A^{-1} y_0)}{\partial a_{kk}} + (\tfrac{n_k - 1}{\lambda_k} - \tfrac{y_k' y_k}{\lambda_k^2}) = M_{k,k} + (\tfrac{n_k - 1}{\lambda_k} - \tfrac{y_k' y_k}{\lambda_k^2}) \\
\frac{\partial(-2\ell)}{\partial\sigma_{kk}} &= (n_k - 1)\frac{\partial(\log\det A + y_0' A^{-1} y_0)}{\partial a_{kk}} - \left(\tfrac{n_k - 1}{\lambda_k} - \tfrac{y_k' y_k}{\lambda_k^2}\right) = n_k M_{k,k} - \frac{\partial(-2\ell)}{\partial\sigma_k^2},
\end{aligned}$$

and, for $i \neq j$, we find that

$$\frac{\partial(-2\ell)}{\partial\sigma_{ij}} = \sqrt{n_i n_j}\left(\frac{\partial(\log\det A + y_0' A^{-1} y_0)}{\partial a_{ij}} + \frac{\partial(\log\det A + y_0' A^{-1} y_0)}{\partial a_{ji}}\right) = 2\sqrt{n_i n_j} M_{i,j},$$

where we used that M is symmetric. $\square$

References

Archakov, I. and Hansen, P. R. (2020), 'A new parametrization of correlation matrices', *https://sites.google.com/site/peterreinhardhansen/* .

Archakov, I., Hansen, P. R. and Lunde, A. (2020), 'A Multivariate Realized GARCH Model', *Working Paper* .

Cadima, J., Calheiros, F. L. and Preto, I. P. (2010), 'The eigenstructure of block-structured correlation matrices and its implications for principal component analysis', *Journal of Applied Statistics* **37**, 577–589.

Creal, D. D., Koopman, S. J. and Lucas, A. (2013), 'Generalized autoregressive score models with applications', *Journal of Applied Econometrics* **28**, 777–795.

Engle, R. and Kelly, B. (2012), 'Dynamic Equicorrelation', *Journal of Business & Economic Statistics* **30**(2), 212–228.

Huang, J. and Yang, L. (2010), 'Correlation matrix with block structure and efficient sampling methods', *Journal of Computational Finance* **14**, 81–94.

Ledoit, O. and Wolf, M. (2004), 'Honey, i shrunk the sample covariance matrix', *Journal of Portfolio Management* **30**, 110–119.

Nyblom, J. (1989), 'Testing for the constancy of parameters over time', *Journal of the American Statistical Association* **84**, 223–230.

Roustant, O. and Deville, Y. (2017), 'On the validity of parametric block correlation matrices with constant within and between group correlations', *arXiv* **math.ST/1705.09793**.

Viana, M. and Olkin, I. (1997), Correlation analysis of ordered observations from a block-equicorrelated multivariate normal distribution, *in* B. N. Panchapakesan S., ed., 'Advances in Statistical Decision Theory and Applications', Birkhäuser, Boston.

Tables and Figures

	Summary statistics of estimated block correlations									# blocks	
Block structure	Mean	Std.	Min	$Q_{10\%}$	$Q_{50\%}$	$Q_{90\%}$	Max	$-\frac{2\ell}{nN}$	$\frac{1}{nN}$BIC	K	$\frac{K(K+1)}{2}$
	U.S. market in 2008 (3958 stocks and 253 days)										
Equicorrelation	0.228	0	0.228	0.228	0.228	0.228	0.228	2.57829	2.57830	1	1
Sectors	0.269	0.073	0.162	0.191	0.252	0.381	0.521	2.53748	2.53824	10	55
Groups	0.253	0.059	0.119	0.177	0.253	0.321	0.521	2.52519	2.52933	24	300
Industries	0.263	0.074	0.088	0.172	0.258	0.362	0.659	2.51123	2.54266	67	2278
Sub-industries	0.273	0.095	-0.036	0.157	0.268	0.394	0.886	2.48261	2.64086	151	11476
	U.S. market in 2013 (2998 stocks and 252 days)										
Equicorrelation	0.165	0	0.165	0.165	0.165	0.165	0.165	2.66583	2.01128	1	1
Sectors	0.181	0.059	0.107	0.126	0.168	0.242	0.507	2.63742	1.99057	10	55
Groups	0.174	0.045	0.092	0.126	0.166	0.231	0.507	2.62850	1.98715	24	300
Industries	0.183	0.060	0.065	0.118	0.173	0.260	0.712	2.61095	2.00065	67	2278
Sub-industries	0.182	0.067	-0.040	0.108	0.175	0.268	0.810	2.58182	2.09290	146	10731

Table 1: Summary statistics for the estimated block correlation matrices.

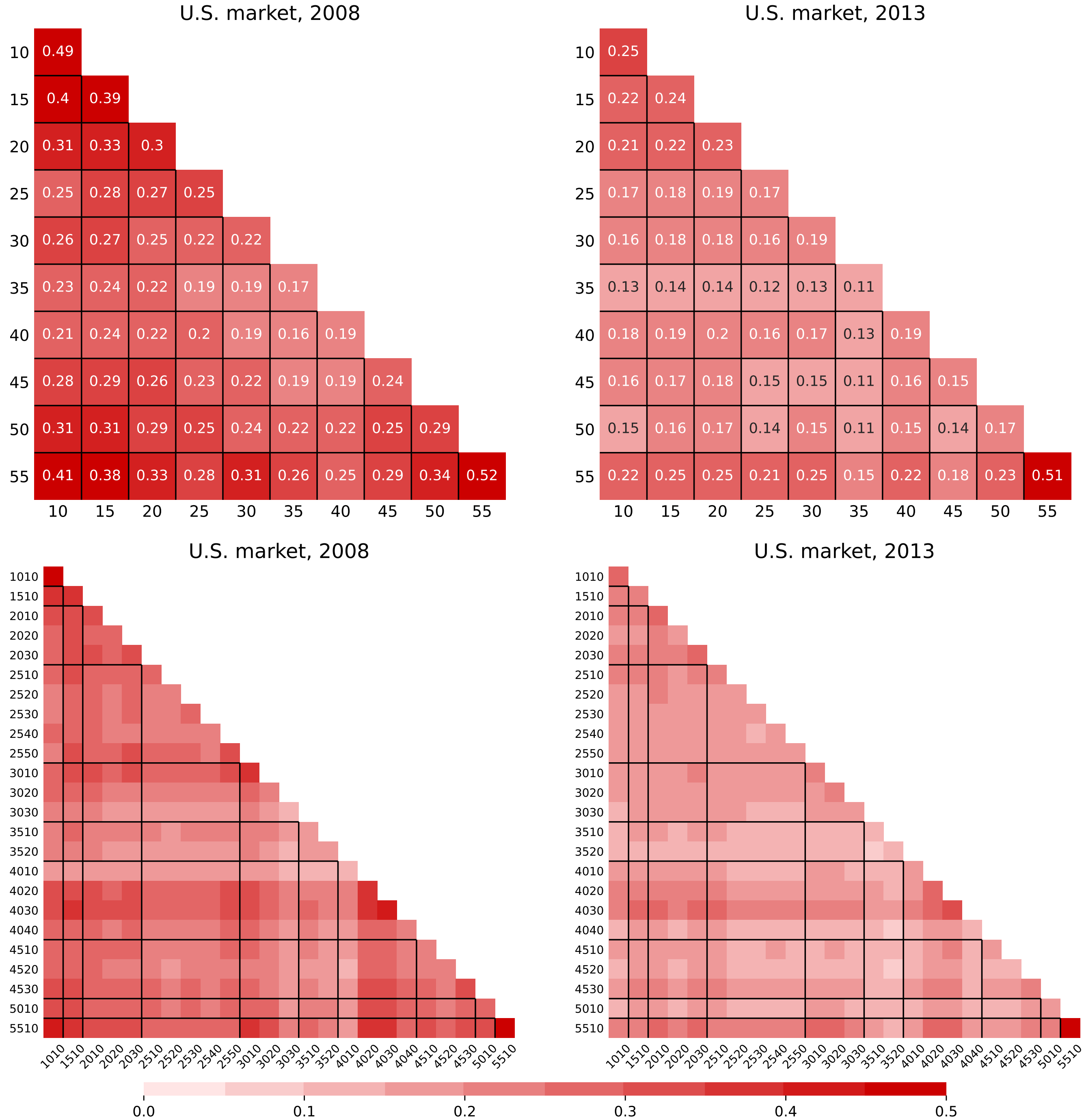

Figure 1: Estimated correlations for a block structure based on GICS sectors (upper panels) and GICS groups (lower panels). Left panels are the estimates based on 253 daily returns in 2008, and right panels are the estimated based on 252 daily returns from 2013. The numbers to the left and below each plot are GICS codes for sectors or groups.

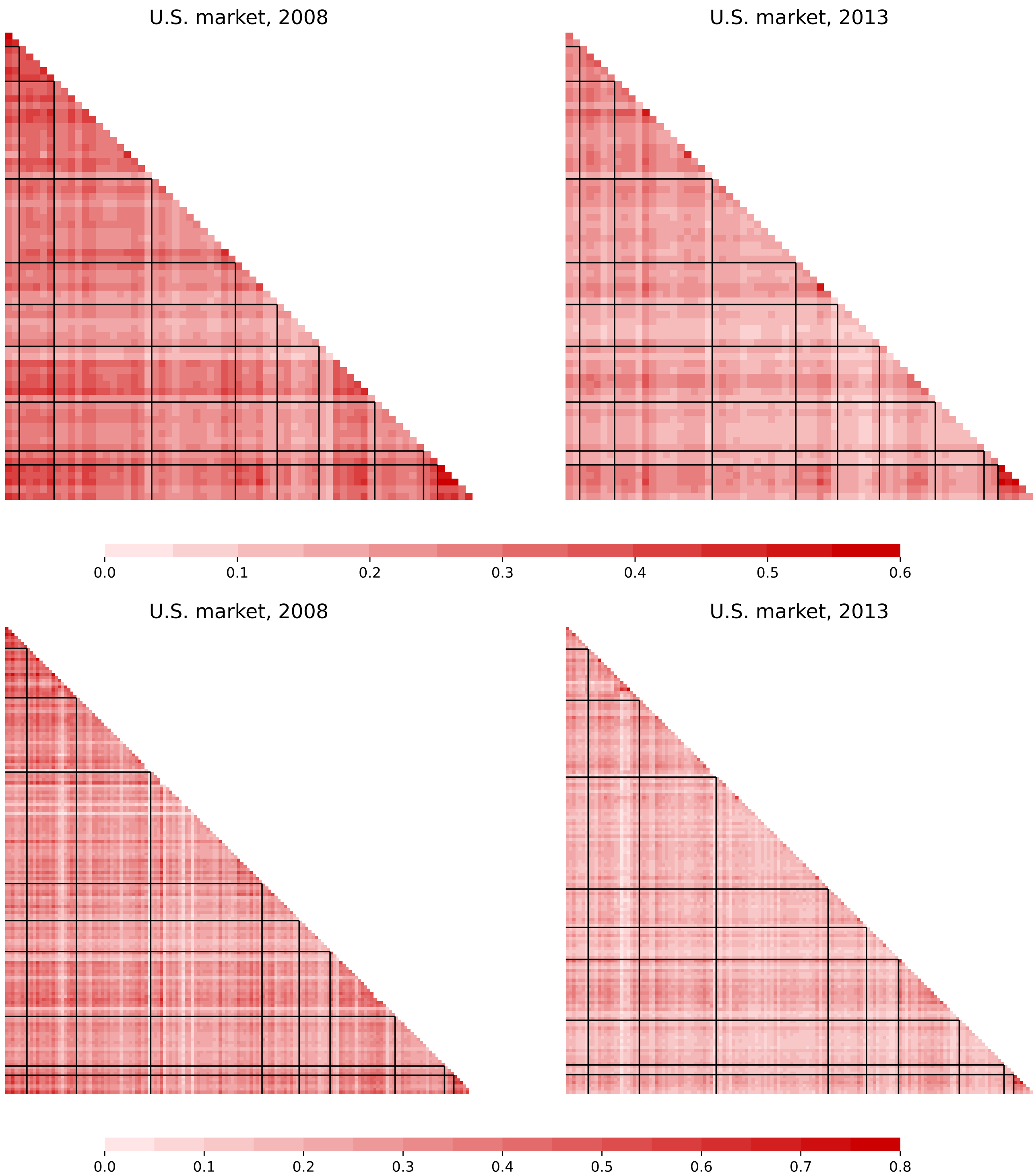

Figure 2: Estimated correlations for a block structure based on GICS industries (upper panels) and GICS sub-industries (lower panels). Left panels are the estimates based on 253 daily returns in 2008, and right panels are the estimated based on 252 daily returns from 2013. Industries and sub-industries are listed chronologically according to their GICS code, that are too plentiful to list individually.