Bounds for the Distance to the Nearest Correlation Matrix

Higham, Nicholas J. and Strabić, Nataša

2015

Manchester Institute for Mathematical Sciences

School of Mathematics

The University of Manchester

BOUNDS FOR THE DISTANCE TO THE
NEAREST CORRELATION MATRIX[*]

NICHOLAS J. HIGHAM[†] AND NATAŠA STRABIĆ[†]

Abstract. In a wide range of practical problems correlation matrices are formed in such a way that, while symmetry and a unit diagonal are assured, they may lack semidefiniteness. We derive a variety of new upper bounds for the distance from an arbitrary symmetric matrix to the nearest correlation matrix. The bounds are of two main classes: those based on the eigensystem and those based on a modified Cholesky factorization. Bounds from both classes have a computational cost of $O(n^3)$ flops for a matrix of order n but are much less expensive to evaluate than the nearest correlation matrix itself. For unit diagonal A with $|a_{ij}| \leq 1$ for all $i \neq j$ the eigensystem bounds are shown to overestimate the distance by a factor at most $1 + n\sqrt{n}$. We show that for a collection of matrices from the literature and from practical applications the eigensystem-based bounds are often good order of magnitude estimates of the actual distance; indeed the best upper bound is never more than a factor 5 larger than a related lower bound. The modified Cholesky bounds are less sharp but also less expensive, and they provide an efficient way to test for definiteness of the putative correlation matrix. Both classes of bounds enable a user to identify an invalid correlation matrix relatively cheaply and to decide whether to revisit its construction or to compute a replacement, such as the nearest correlation matrix.

Key words. correlation matrix, distance to the nearest correlation matrix, indefinite matrix, positive semidefinite matrix, shrinking, modified Cholesky factorization

AMS subject classifications. 15A45, 65F30

1. Introduction. In many applications involving statistical modelling the first step in the analysis is to compute the sample correlation matrix—a real symmetric positive semidefinite matrix with unit diagonal—from empirical or experimental data [28, p. 25]. Because of missing observations, indefinite approximations to the sample correlation matrix frequently arise [19, Sec. 2.2]. A lack of definiteness is also encountered in several other contexts. Stress testing in finance requires certain elements of a valid correlation matrix to be replaced by new values, which may result in the new matrix becoming indefinite [11], [24]. The sensitivity of sample correlation matrices to outliers in the data has led to the development of robust estimators. Devlin, Gnanadesikan, and Kettenring [9] propose several possibilities and note that some methods that compute the estimator in an elementwise manner can produce matrices with negative eigenvalues.

A further example of how indefiniteness can arise is in aggregation methods used in large-scale resource assessment, for example in geology [5] or finance [1]. These methods combine reliable estimates of correlation matrices for each group, say a geographical region or a market portfolio, into a global correlation matrix. The combination is achieved either by embedding small, "within group" correlation matrices as diagonal blocks into a crudely estimated global correlation matrix, or by constructing a block-diagonal matrix from the individual group correlation matrices and filling out the off-diagonal blocks by assigning the "between group" correlation coefficients according to expert judgement. Again, there is no guarantee that the newly constructed

[†]School of Mathematics, The University of Manchester, Manchester, M13 9PL, UK (nick.higham@manchester.ac.uk, http://www.maths.manchester.ac.uk/~higham, natasa.strabic@manchester.ac.uk, http://www.maths.manchester.ac.uk/~nstrabic)

matrix is in fact positive semidefinite.

A popular approach to correcting an indefinite approximation $A \in \mathbb{R}^{n \times n}$ to a correlation matrix is to replace A by the nearest correlation matrix in the Frobenius norm, that is, by

(1.1) $\qquad \mathrm{ncm}(A) = \mathrm{argmin}\{ \|A - X\|_F : X \text{ is a correlation matrix } \},$

where $\|A\|_F^2 = \sum_{i,j} a_{ij}^2$. The first method with guaranteed convergence for solving (1.1) was the alternating projections method proposed by Higham [15, Alg. 3.3], which iteratively projects onto the set $\mathcal{U}_n$ of matrices with unit diagonal and onto the convex cone S_n of symmetric positive semidefinite matrices, respectively, supplementing the projection onto S_n with a correction due to Dykstra [10]. The rate of convergence of this method is at best linear but we have recently shown that the convergence can be accelerated significantly using Anderson acceleration [16]. The alternating projections method remains widely used in applications (see, for example, the references in [16], [17]), though a faster, globally quadratically convergent Newton algorithm was later developed by Qi and Sun [23], to which practical improvements were made by Borsdorf and Higham [6]. The algorithm of Borsdorf and Higham requires an eigendecomposition of a symmetric matrix on each iteration and so it costs at least $10n^3/3$ flops per iteration. The Newton algorithm typically needs about 7 iterations, so the total cost to compute $\mathrm{ncm}(A)$ is at least $70n^3/3$ flops.

While methods for computing the nearest correlation matrix to a given symmetric matrix A are well developed, little attention has been given to estimating the distance

$$\mathrm{d}_{\mathrm{corr}}(A) = \|A - \mathrm{ncm}(A)\|_F$$

without computing $\mathrm{ncm}(A)$. Importantly, the iterates produced by the alternating projections method are not themselves correlation matrices as (with $\mathcal{P}$ denoting projection) the matrix $\mathcal{P}_{\mathcal{U}_n}(\mathcal{P}_{S_n}(X))$ might be indefinite and the matrix $\mathcal{P}_{S_n}(\mathcal{P}_{\mathcal{U}_n}(X))$ might not have an exactly unit diagonal. For the Newton method, the iterates do not satisfy the constraint of having a unit diagonal, as discussed in [6, Sec. 3.4]. Hence for both methods the iterates do not provide upper bounds on $\mathrm{d}_{\mathrm{corr}}(A)$. As the Newton method solves the dual problem of (1.1) [23], on each iteration the value of the dual function provides a lower bound for $\mathrm{d}_{\mathrm{corr}}$ [20].

Our goal is to obtain bounds on $\mathrm{d}_{\mathrm{corr}}(A)$ that are inexpensive to compute and are of the correct order of magnitude. Indeed bounds correct to within a small constant factor are entirely adequate for practical applications. In this work we summarize the few existing bounds for $\mathrm{d}_{\mathrm{corr}}(A)$ and derive several new upper bounds. While the best bounds have a computational cost of $O(n^3)$ flops they are significantly less expensive to compute than $\mathrm{ncm}(A)$ itself.

Bounds on $\mathrm{d}_{\mathrm{corr}}(A)$ that can be easily evaluated using standard computational tools will certainly be of interest to practitioners, as illustrated by the discovery by Xu and Evers [35] that several matrices thought to be correlation matrices in the work of Tyagi and Das [33] actually had some negative eigenvalues. While attempting to compute the Cholesky factorization is sufficient to determine whether a matrix is positive semidefinite, we propose using a modified Cholesky factorization instead. The standard and modified Cholesky factorizations have the same computational cost, but for an indefinite matrix modified Cholesky factorization provides additional information that can be used to construct an upper bound on $\mathrm{d}_{\mathrm{corr}}(A)$; this bound can help the user to decide whether to revisit the construction of the matrix, perhaps by acquiring more data or by refining the statistical analysis. In our experiments,

the best modified Cholesky bound is at most two orders of magnitude larger than $\mathrm{d_{corr}}(A)$.

Sharper bounds are available based on spectral information. We present several bounds based only on the knowledge of the eigenvalues of A, but the best bound in this class, which in our experiments is at most one order of magnitude larger than $\mathrm{d_{corr}}(A)$, uses the nearest positive semidefinite matrix to A and so a knowledge of eigenvectors is also required.

The paper is organized as follows. In section 2 we summarize existing upper and lower bounds on the distance to the nearest correlation matrix. In section 3 we derive our new upper bounds and give a result bounding the overestimation by a factor that does not exceed $1 + n\sqrt{n}$. We analyze the computational cost of the bounds in section 4. In section 5 we illustrate the quality of the bounds on several small examples of invalid correlation matrices found in the literature and on larger, real-life matrices from practical applications. Conclusions are given in section 6.

2. Existing bounds. We first summarize currently available bounds for the distance to the nearest correlation matrix. We will need the following result on the nearest positive semidefinite matrix found in, for example, [14, Thm. 2.1], [29, Thm. 1].

LEMMA 2.1. *Let $A \in \mathbb{R}^{n \times n}$ be symmetric with spectral decomposition $A = Q\Lambda Q^T$, where $Q = [q_1, \ldots, q_n]$ is orthogonal and $\Lambda = \mathrm{diag}(\lambda_i)$. Then the unique solution to $\min\{\, \|A - X\|_F : X \text{ is symmetric positive semidefinite}\, \}$ is*

$$(2.1) \qquad A_+ = Q \, \mathrm{diag}(\max(\lambda_i, 0)) Q^T.$$

The next result is [15, Lem. 1.1].

LEMMA 2.2 (Higham). *For symmetric $A \in \mathbb{R}^{n \times n}$ with eigenvalues λ_i,*

$$\max\{\alpha_1, \alpha_2\} \le \mathrm{d_{corr}}(A) \le \min\{\beta_1, \beta_2, \beta_3\},$$

where

$$(2.2) \qquad \alpha_1^2 = \sum_{i=1}^{n} (a_{ii} - 1)^2 + \sum_{\substack{|a_{ij}| > 1 \\ i \ne j}} (1 - |a_{ij}|)^2,$$

$$(2.3) \qquad \alpha_2 = \|A - A_+\|_F = \left(\sum_{\lambda_i < 0} \lambda_i^2 \right)^{1/2},$$

$$(2.4) \qquad \beta_1 = \|A - I\|_F,$$

$$(2.5) \qquad \beta_2 = \min\{\, \|A - zz^T\|_F : z_i = \pm 1, \ i = 1\colon n \, \},$$

$$(2.6) \qquad \beta_3 = \min_{-1 \le \rho \le 1} \|A - T(\rho)\|_F, \quad where \quad (T(\rho))_{ij} = \rho^{|i-j|}.$$

The lower bound α_1 follows from the fact that the elements of a correlation matrix are bounded in modulus by 1. The equivalence of the two formulas for α_2 is shown by Lemma 2.1. The upper bounds in Lemma 2.2 are obtained as the distance to certain classes of correlation matrices. In particular, β_3 arises from the matrices with (i, j) element $\rho^{|i-j|}$, known as Kac-Murdock-Szegő Toeplitz matrices [31], which are positive semidefinite for $-1 \le \rho \le 1$.

Travaglia [30, Prop. 3.1] obtained a further lower bound on $\mathrm{d}_{\mathrm{corr}}(A)$ using the circulant mean A_c, defined as the circulant matrix with first row $(c_0, c_1, \ldots, c_{n-1})$, where

$$c_0 = \frac{1}{n}\operatorname{trace}(A),$$

$$c_k = \frac{1}{n}\left(\sum_{i=1}^{n-k} a_{i,i+k} + \sum_{i=1}^{k} a_{i,i+n-k}\right), \quad k = 1, 2, \ldots, n-1.$$

This lower bound and a trivial upper bound are combined in the next result.

LEMMA 2.3 (Travaglia). *For symmetric $A \in \mathbb{R}^{n \times n}$,*

$$(2.7) \qquad \mathrm{d}_{\mathrm{corr}}(A_c) \le \mathrm{d}_{\mathrm{corr}}(A) \le \mathrm{d}_{\mathrm{corr}}(A_c) + \|A - A_c\|_F.$$

3. New bounds. In this section we derive new upper bounds on the distance to the nearest correlation matrix that do not require the solution to a minimization problem, unlike the bounds β_2 and β_3 from Lemma 2.2 and the upper bound in Lemma 2.3. Our first bound is the distance to the correlation matrix obtained by scaling A_+ from (2.1) to have unit diagonal.

THEOREM 3.1. *Let $A \in \mathbb{R}^{n \times n}$ be symmetric with positive diagonal elements. Then*

$$(3.1) \qquad \|A - A_+\|_F \le \mathrm{d}_{\mathrm{corr}}(A) \le \|A - \widetilde{A}_+\|_F,$$

where $\widetilde{A}_+ = D^{-1/2} A_+ D^{-1/2}$, with $D = \operatorname{diag}((A_+)_{ii})$.

Proof. The lower bound is α_2 in (2.3). The upper bound is immediate if we can show that $\widetilde{A}_+$ is a correlation matrix. The only question is whether it is defined, that is, whether the positive semidefinite matrix A_+ has positive diagonal elements, so that D is nonsingular and positive definite. From Lemma 2.1 we see that $A_+ - A$ is positive semidefinite and it follows that the diagonal elements of A_+ are at least as large as the corresponding diagonal elements of A, and hence they are positive. □

In the next result we obtain an alternative upper bound that, while weaker than that in (3.1), is less expensive to compute, as we explain in section 4. Note that the theorem is valid for $t = n$, that is, for a positive semidefinite matrix A.

THEOREM 3.2. *Let $A \in \mathbb{R}^{n \times n}$ be symmetric with positive diagonal elements and eigenvalues $\lambda_1 \ge \cdots \ge \lambda_t \ge 0 > \lambda_{t+1} \ge \cdots \ge \lambda_n$. Then*

$$(3.2) \qquad \left(\sum_{i=t+1}^{n} \lambda_i^2\right)^{1/2} \le \mathrm{d}_{\mathrm{corr}}(A) \le \left(\sum_{i=t+1}^{n} \lambda_i^2\right)^{1/2} + \theta\left(\sum_{i=1}^{t} \lambda_i^2\right)^{1/2},$$

where

$$\theta = \max\left\{\left|1 - \frac{1}{\max_i a_{ii} - \min(\lambda_n, 0)}\right|, \left|1 - \frac{1}{\min_i a_{ii}}\right|\right\}.$$

Proof. The lower bound is (2.3).

By Theorem 3.1 and the triangle inequality we have

$$(3.3) \qquad \mathrm{d}_{\mathrm{corr}}(A) \le \|A - \widetilde{Y}\|_F \le \|A - Y\|_F + \|Y - \widetilde{Y}\|_F,$$

4

where $Y = A_+$ in (2.1) is positive semidefinite and $\widetilde{Y} = D^{-1/2}YD^{-1/2}$ is a correlation matrix, with $D = \mathrm{diag}(d_i)$ and $d_i = y_{ii} > 0$ for all i. We now bound $|y_{ij} - \widetilde{y}_{ij}|$.

Let $m = \min_i a_{ii}$ and $M = \max_i a_{ii}$. With e_i the ith unit vector and $\Lambda_- = \mathrm{diag}(\min(\lambda_i, 0))$, we have

$$d_i = e_i^T Y e_i = e_i^T(A - Q\Lambda_- Q^T)e_i = a_{ii} - e_i^T Q\Lambda_- Q^T e_i = a_{ii} - \delta_i,$$

with $\delta_i = e_i^T Q\Lambda_- Q^T e_i$, and so

$$m - \delta_i \le d_i \le M - \delta_i.$$

With $y = Q^T e_i$, we have $\|y\|_2 = 1$ and $\delta_i = y^T \Lambda_- y \le 0$, and we can bound δ_i below by

$$\min(\lambda_n, 0) = \min_{x \ne 0} \frac{x^T \Lambda_- x}{x^T x} \le \frac{y^T \Lambda_- y}{y^T y} = y^T \Lambda_- y = \delta_i \le 0.$$

Note that we must have $\min(\lambda_n, 0)$ for the first equality in the previous line to hold, since λ_n could be positive. It follows that $m \le d_i \le M - \min(\lambda_n, 0)$ for every i and so

$$\left(\frac{1}{M - \min(\lambda_n, 0)}\right)^{1/2} \le d_i^{-1/2} \le m^{-1/2}.$$

Since $\widetilde{y}_{ij} = d_i^{-1/2} d_j^{-1/2} y_{ij}$ we have

$$|y_{ij} - \widetilde{y}_{ij}| = \left|(1 - d_i^{-1/2} d_j^{-1/2})y_{ij}\right| = \left|1 - d_i^{-1/2} d_j^{-1/2}\right| |y_{ij}|.$$

Finally, from

$$1 - \frac{1}{m} \le 1 - d_i^{-1/2} d_j^{-1/2} \le 1 - \frac{1}{M - \min(\lambda_n, 0)}$$

we have $|y_{ij} - \widetilde{y}_{ij}| \le \theta |y_{ij}|$ and therefore $\|Y - \widetilde{Y}\|_F \le \theta\|Y\|_F$. The upper bound in (3.2) then follows from (3.3). $\square$

For $t = n$, Theorem 3.2 yields the following corollary, which quantifies the effect on the distance $\mathrm{d}_{\mathrm{corr}}(A)$ of the departure of the diagonal elements of a positive semidefinite matrix A from 1.

COROLLARY 3.3. *Let $A \in \mathbb{R}^{n \times n}$ be symmetric positive semidefinite with positive diagonal elements. Then*

$$(3.4) \qquad \mathrm{d}_{\mathrm{corr}}(A) \le \max\left\{\left|1 - \frac{1}{\max_i a_{ii}}\right|, \left|1 - \frac{1}{\min_i a_{ii}}\right|\right\} \|A\|_F.$$

If all the diagonal elements of a positive semidefinite matrix A are at most 1 then $\mathrm{ncm}(A)$ is easy to compute directly as it is obtained from A by replacing each diagonal element by 1. However, (3.4) applies more generally.

In many applications the invalid approximation to a correlation matrix has unit diagonal and at least one negative eigenvalue. In this case Theorem 3.2 simplifies as follows.

COROLLARY 3.4. *Let $A \in \mathbb{R}^{n \times n}$ be symmetric with unit diagonal and eigenvalues $\lambda_1 \ge \cdots \ge \lambda_t \ge 0 > \lambda_{t+1} \ge \cdots \ge \lambda_n$, where $\lambda_n < 0$. Then*

$$(3.5) \qquad \left(\sum_{i=t+1}^{n} \lambda_i^2\right)^{1/2} \le \mathrm{d}_{\mathrm{corr}}(A) \le \left(\sum_{i=t+1}^{n} \lambda_i^2\right)^{1/2} + \frac{|\lambda_n|}{1 + |\lambda_n|}\left(\sum_{i=1}^{t} \lambda_i^2\right)^{1/2}.$$

The next result gives a sharper upper bound than (3.5). The proof uses the idea of shrinking from Higham, Strabić, and Šego [17].

THEOREM 3.5. *Let $A \in \mathbb{R}^{n \times n}$ be symmetric with unit diagonal and smallest eigenvalue $\lambda_n < 0$. Then*

$$(3.6) \qquad \mathrm{d_{corr}}(A) \leq \frac{|\lambda_n|}{1 + |\lambda_n|} \|A - I\|_F,$$

and this bound is no larger than the upper bound in (3.5).

Proof. Let $S(\alpha) = \alpha I + (1 - \alpha)A$, which has unit diagonal. We have $\mathrm{d_{corr}}(A) \leq \|A - S(\alpha_*)\|_F$, where $\alpha_* = \min\{\alpha \in [0, 1] : S(\alpha) \text{ is positive semidefinite}\}$. It is easy to see that $\alpha_* = -\lambda_n/(1 - \lambda_n)$ and $A - S(\alpha_*) = \alpha_*(A - I)$, which gives (3.6).

Now we compare the bound with (3.5). The triangle inequality gives

$$(3.7) \qquad \|A - I\|_F \leq \|A - A_+\|_F + \|A_+ - I\|_F,$$

where A_+ from (2.1) is the nearest positive semidefinite matrix to A. For the second term, we have

$$\|A_+ - I\|_F^2 = \mathrm{trace}\big((A_+ - I)^T (A_+ - I)\big) = \|A_+\|_F^2 - 2\,\mathrm{trace}(A_+) + n.$$

We noted in the proof of Theorem 3.1 that $A_+ - A$ is positive semidefinite, and since A has unit diagonal it follows that $\mathrm{trace}(A_+) \geq n$. Therefore, $-2\,\mathrm{trace}(A_+) + n \leq -n < 0$ and so $\|A_+ - I\|_F^2 \leq \|A_+\|_F^2$. Then from (3.7) it follows that $\|A - I\|_F \leq \|A - A_+\|_F + \|A_+\|_F$, so, since $\alpha_* \leq 1$,

$$\alpha_* \|A - I\|_F \leq \alpha_* \|A - A_+\|_F + \alpha_* \|A_+\|_F \leq \|A - A_+\|_F + \alpha_* \|A_+\|_F.$$

This completes the proof, since the right-hand side of the latter inequality is the upper bound in (3.5). $\square$

Note that the bound (3.6) is also sharper than β_1 given in (2.4).

We now have several upper bounds and a natural question is "how sharp are they?" For the most practically important case of A with unit diagonal, the next result gives a limit on the overestimation for the upper bound of Theorem 3.1.

THEOREM 3.6. *Let $A \in \mathbb{R}^{n \times n}$ be symmetric with unit diagonal, t nonnegative eigenvalues, largest eigenvalue λ_1, and smallest eigenvalue $\lambda_n < 0$. Then, in the notation of Theorem 3.1,*

$$(3.8) \qquad \frac{\|A - \widetilde{A}_+\|_F}{\|A - A_+\|_F} \leq 1 + \frac{\sqrt{t}\,\lambda_1}{1 + |\lambda_n|}.$$

If, in addition, $|a_{ij}| \leq 1$ for $i \neq j$, then

$$(3.9) \qquad \frac{\|A - \widetilde{A}_+\|_F}{\|A - A_+\|_F} \leq 1 + n\sqrt{t}.$$

Proof. Using the triangle inequality and the relation $A = A_+ + A_-$ we have

$$\frac{\|A - \widetilde{A}_+\|_F}{\|A - A_+\|_F} = \frac{\|A_- + A_+ - \widetilde{A}_+\|_F}{\|A_-\|_F} \leq 1 + \frac{\|A_+ - \widetilde{A}_+\|_F}{\|A_-\|_F}.$$

As in the proof of Theorem 3.2, we have $\|A_+ - \widetilde{A}_+\|_F \le \theta\|A_+\|_F$, where $\theta = |\lambda_n|/(1+|\lambda_n|)$, since A has unit diagonal and $\lambda_n < 0$. Therefore

$$(3.10) \qquad \frac{\|A - \widetilde{A}_+\|_F}{\|A - A_+\|_F} \le 1 + \frac{|\lambda_n|}{1 + |\lambda_n|}\frac{\|A_+\|_F}{\|A_-\|_F}.$$

Since $\|A_+\|_F^2 = \sum_{i=1}^t \lambda_i^2 \le t\lambda_1^2$ and $\|A_-\|_F \ge \|A_-\|_2 = |\lambda_n|$, it follows that (noting that the assumptions of the theorem imply $\lambda_1 > 0$)

$$\frac{\|A_+\|_F}{\|A_-\|_F} \le \frac{\sqrt{t}\,\lambda_1}{|\lambda_n|}.$$

Substituting this bound into (3.10) yields (3.8). The bound (3.9) follows because $\lambda_1 \le n$ for any matrix with elements bounded in modulus by 1. $\quad\square$

It is easy to see that the upper bound (3.8) also holds for the ratio of the upper and lower bounds from Corollary 3.4, and hence also for the ratio of the shrinking bound (3.6) and (2.3), by Theorem 3.5. Moreover, we have $\|A - A_+\|_F \le \mathrm{d_{corr}}(A) \le \psi\|A - A_+\|_F$, where $\psi = 1 + \sqrt{t}\lambda_1/(1 + |\lambda_n|)$.

Another way to obtain an upper bound on the distance to the nearest correlation matrix is to modify Theorem 3.1 by replacing the nearest positive semidefinite matrix A_+ by a more cheaply computable approximation to A_+. To construct such an approximation we will use modified Cholesky factorizations, originally developed in the context of nonlinear optimization as a way of dealing with indefinite Hessians. Given a symmetric, possibly indefinite matrix A these algorithms construct a factorization

$$(3.11) \qquad P^T(A + E)P = LDL^T,$$

where P is a permutation matrix, L is unit lower triangular, and E and $A + E$ are positive semidefinite. The algorithms of Gill, Murray, and Wright [12, Sec. 4.4.2.2] and Eskow and Schnabel [26], [27] produce a diagonal D and a diagonal E, while the algorithm of Cheng and Higham [8] produces a block diagonal D with diagonal blocks of order 1 or 2 and an E that is generally full. The cost of all three algorithms is the same as the cost of computing the Cholesky factorization to highest order terms, which is substantially less than the cost of computing A_+. Note that bounds based on modified Cholesky factorizations provide an efficient way to determine whether the matrix is positive semidefinite to start with, as in this case $E = 0$.

THEOREM 3.7. *Let $A \in \mathbb{R}^{n \times n}$ be symmetric with positive diagonal elements. Then*

$$(3.12) \qquad \mathrm{d_{corr}}(A) \le \|A - \widetilde{A}_{\mathrm{mc}}\|_F,$$

where $\widetilde{A}_{\mathrm{mc}} = D^{-1/2}A_{\mathrm{mc}}D^{-1/2}$ is a correlation matrix with $A_{\mathrm{mc}} = A + E$ from the modified Cholesky factorization (3.11) and $D = \mathrm{diag}(A_{\mathrm{mc}})$.

Proof. The proof is essentially the same as that of the upper bound in Theorem 3.1 and uses the fact that the diagonal elements of A_{mc} are at least as large as those of A. $\quad\square$

As a final new upper bound on $\mathrm{d_{corr}}(A)$ we make use of one of the rare explicitly known solutions to the nearest correlation matrix problem, for the so-called one parameter model. Here, a matrix $C(w) \in \mathbb{R}^{n \times n}$ is defined for a real parameter w as a unit diagonal matrix with all off-diagonal elements equal to w:

$$C(w) = (1 - w)I + wee^T = I + w(ee^T - I),$$

where $e = [1, 1, \ldots, 1]^T$. As shown in [7, Lem. 2.1] the matrix $C(w)$ is a correlation matrix if and only if $-1/(n-1) \le w \le 1$.

THEOREM 3.8. *For $A \in \mathbb{R}^{n \times n}$ symmetric with $n \ge 2$,*

$$\mathrm{d_{corr}}(A) \le \min\{\, \|A - C(w)\|_F : C(w) \text{ is a correlation matrix}\,\}$$

(3.13)
$$= \|A - C(w_{\mathrm{opt}})\|_F,$$

where w_{opt} is the projection of $w = \left(e^T A e - \mathrm{trace}(A)\right)/(n^2 - n)$ onto the interval $[-1/(n-1), 1]$.

Proof. The equality (3.13) is from [7, Thm. 2.2]. $\qquad\square$

4. Computing the bounds. The main criteria for judging a bound are its cost and its accuracy. In this section we discuss the cost of the bounds presented above and in the next section we carry out numerical experiments to test their accuracy.

Table 4.1 summarizes the bounds, their applicability, and their cost. We will comment only on the nontrivial entries in the table.

We can evaluate the lower bound α_2 in (2.3) and the upper bounds in (3.2) and (3.5) without computing A_+ explicitly, but rather by computing all the positive eigenvalues or all the negative eigenvalues—whichever are fewer in number—and then using $\sum_{i=1}^{t} \lambda_i^2 + \sum_{i=t+1}^{n} \lambda_i^2 = \|A\|_F^2$. We can assume $t \ge n/2$ without loss of generality and therefore we compute the $n - t$ negative eigenvalues by tridiagonalizing A at the cost of $4n^3/3$ flops [13, p. 459] and then computing the $n - t$ negative eigenvalues of the tridiagonal matrix by the bisection method at a cost of $O(n(n-t))$ flops [3, p. 50], which makes the total cost for the bounds α_2, (3.2), and (3.5) at most $4n^3/3$ flops. The cost of (3.6) is the same.

As noted in [15], computing the upper bound β_2 from Lemma 2.2 is equivalent to maximizing $z^T A z$ over all vectors z with elements ± 1, which is an NP-hard problem [25]. For a matrix A of size n there are 2^n positive semidefinite matrices $z z^T$ for such z, which makes an exhaustive search algorithm unfeasible unless n is very small.

A formula for the distance $\mathrm{d_{corr}}(A_c)$ in Lemma 2.3 is given in [30, Thm. 4.1]. However, it requires not only all the eigenvalues but also their multiplicities, which are not reliably computable in floating point arithmetic. We therefore have to regard $\mathrm{d_{corr}}(A_c)$ as no more easily computable in general than $\mathrm{d_{corr}}(A)$, and so the bounds of Lemma 2.3 are of limited interest.

Next we turn to the bound (3.1), which requires $\widetilde{A}_+$, and hence A_+. Recall that we order the eigenvalues $\lambda_1 \ge \cdots \ge \lambda_t \ge 0 > \lambda_{t+1} \ge \cdots \ge \lambda_n$, and assume without loss of generality that $t \ge n/2$ so that the majority of eigenvalues are nonnegative. We first compute the tridiagonalization $A = Q T Q^T$ and do not form Q explicitly but keep it in factored form. By applying bisection and inverse iteration to the tridiagonal matrix T we compute $\lambda_{t+1}, \ldots, \lambda_n$ and the corresponding eigenvectors, which are placed in the columns of $Z = [z_{t+1}, \ldots, z_n]$. We then compute the matrix $W = Z \, \mathrm{diag}(|\lambda_{t+1}|, \ldots, |\lambda_n|)^{1/2} \in \mathbb{R}^{n \times (n-t)}$ and apply Q to get $B = QW$. Finally, $A_+ = A + BB^T$. The total cost is $4n^3/3$ flops for T, $O(n^2)$ flops to compute $\lambda_{t+1}, \ldots, \lambda_n$ and form Z and W, $2n^2(n-t)$ flops[1] to form B [13, Sec. 5.1.6], and $n^2(n-t)$ flops to form A_+, exploiting symmetry throughout. The total cost is therefore $4n^3/3 + 3n^2(n-t) \le 4n^3/3 + 3n^3/2 = 17n^3/6$ flops.

Three different bounds are obtained from (3.12), corresponding to the three different modified Cholesky algorithms. While E in (3.11) is explicitly produced by the

---

[1] The errata web page http://www.cs.cornell.edu/cv/GVL4/Errata.htm for the fourth edition of [13] notes that the book incorrectly omits the leading 2 on page 238 from this operation count.

TABLE 4.1
Approximate cost in flops of the bounds for a symmetric $A \in \mathbb{R}^{n \times n}$. For the bound α_1, k is the number of elements $|a_{ij}| > 1$, $i \neq j$. For the bound (3.4), $m = \min_i a_{ii}$ and $M = \max_i a_{ii}$.

	Definition	Cost (flops)	Restrictions
Lower bounds			
α_1	(2.2)	$3(n+k)$	
$\alpha_2 = \|A - A_+\|_F$	(2.3)	$4n^3/3$	
$\mathrm{d}_{\mathrm{corr}}(A_c)$	(2.7)	As $\mathrm{d}_{\mathrm{corr}}(A)$	
Upper bounds			
$\beta_1 = \|A - I\|_F$	(2.4)	n^2	
$\beta_2 = \min\{\|A - zz^T\|_F : z_i = \pm 1\}$	(2.5)	$O(n^2 2^n)$	
$\beta_3 = \min_{-1 \leq \rho \leq 1} \|A - T(\rho)\|_F$	(2.6)	$O(n^2)$	
$\mathrm{d}_{\mathrm{corr}}(A_c) + \|A - A_c\|_F$	(2.7)	As $\mathrm{d}_{\mathrm{corr}}(A)$	
$\|A - \widetilde{A}_+\|_F$	(3.1)	$17n^3/6$	$a_{ii} > 0$
$\|A - A_+\|_F + \theta\|A_+\|_F$	(3.2)	$4n^3/3$	$a_{ii} > 0$
$\max\{|1 - 1/M|, |1 - 1/m|\}\|A\|_F$	(3.4)	n^2	$a_{ii} > 0, \lambda_n \geq 0$
$\|A - A_+\|_F + \|A_+\|_F |\lambda_n|/(1 + |\lambda_n|)$	(3.5)	$4n^3/3$	$\lambda_n < 0, a_{ii} \equiv 1$
$\|A - I\|_F |\lambda_n|/(1 + |\lambda_n|)$	(3.6)	$4n^3/3$	$\lambda_n < 0, a_{ii} \equiv 1$
$\|A - \widetilde{A}_{\mathrm{mc}}\|_F$	(3.12)	$2n^3/3$	$a_{ii} > 0$
$\|A - C(w_{\mathrm{opt}})\|_F$	(3.13)	$2n^2$	

algorithms of Gill, Murray and Wright, and Eskow and Schnabel, the algorithm of Cheng and Higham does not explicitly produce E, so this algorithm requires an extra matrix multiplication $L \cdot DL^T$. The cost stated in Table 4.1 includes the latter step. Alternatively, $\|A - PLDL^TP^T\|$ can be estimated in $O(n^2)$ flops, using the algorithm of [18] for the 1-norm or the power method for the 2-norm.

In [15] an approximation for β_3 from Lemma 2.2 was computed as the approximate local minimum obtained with the MATLAB `fminbnd` minimizer. We propose an alternative. Note that the function we are minimizing for the bound β_3 is a polynomial in the variable ρ:

$$f(\rho) = \|A - T(\rho)\|_F^2 = 2 \sum_{1 \leq i < j \leq n} (a_{ij} - \rho^{j-i})^2 + \sum_{i=1}^{n} (a_{ii} - 1)^2.$$

We compute the stationary points of f, that is, the zeros of the derivative

$$f'(\rho) = -4 \sum_{1 \leq i < j \leq n} \left[(j-i)a_{ij}\rho^{j-i-1} - (j-i)\rho^{2(j-i)-1} \right],$$

which has degree $2n - 3$. Then we obtain β_3 as the minimum value of f over all stationary points in $[-1, 1]$ along with the endpoints ± 1. The dominant cost for this bound is computing the stationary points, which are the real eigenvalues of a companion matrix of order $2n - 3$ in $[-1, 1]$; these can be computed in $O(n^2)$ flops [2].

To summarize, we can separate the new upper bounds into three main categories. The most expensive bound to compute is (3.1) and it uses $\widetilde{A}_+$; the less expensive bounds (3.2), (3.5), and (3.6) are based on the knowledge of eigenvalues only; and the least expensive bound is the modified Cholesky bound (3.12), which has half the cost of the eigenvalue-only based bounds. In the next section we analyze the accuracy of the bounds on examples collected from the literature and real-life applications.

5. Numerical experiments. Our tests were carried out in MATLAB R2014a on a machine with an Intel Core i7-4910MQ 2.90GHz processor and 16GB RAM. As test matrices we use the following indefinite symmetric matrices with unit diagonal.

high02 A matrix of order 3 from Higham [15, p. 334].

tec03 A matrix of order 4 from Turkay, Epperlein, and Christofides [32, $\widehat{\Omega}$ on p. 86,].

bhwi01 A matrix of order 5 from Bhansali and Wise [4, Sec. 2, second matrix].

mmb13 A matrix of order 6 constructed from foreign exchange trading data supplied by the Royal Bank of Scotland [21, p. 36].

fing97 A matrix of order 7 from Finger [11, Table 4].

tyda99R1–tyda99R3 The matrices R_1, R_2, and R_3 of order 8 from Tyagi and Das [33, Table 1]. Although thought by those authors to be correlation matrices, as pointed out by Xu and Evers [35] they have some negative eigenvalues.

usgs13 A matrix of order 94 corresponding to carbon dioxide storage assessment units for the Rocky Mountains region of the United States that was generated during the national assessment of carbon dioxide storage resources [34].

RiskMetrics1–RiskMetrics6 Six matrices from the RiskMetrics database, as used in [6] and [16]. Each matrix has dimension 387.

cor1399, cor3210 Two large matrices constructed from stock data, the first of order 1399 and the second of order 3120. The matrices were provided by investment company Orbis.

The nearest correlation matrix required to determine the true distance $d_{corr}(A)$ is computed by the code `nag_correg_corrmat_nearest (g02aa)` from the NAG Toolbox for MATLAB Mark 24 [22], which implements the preconditioned Newton algorithm of [6], and we chose tolerance tol $= 10^{-10}$.

In our first test we analyze the performance of the modified Cholesky algorithms used for the bound (3.12) for all the indefinite matrices listed above. In Table 5.1 the matrix $\widetilde{A}_{mc}$ from (3.12) corresponding to the algorithms of Gill, Murray and Wright [12, Sec. 4.4.2.2], Schnabel and Eskow [26] and [27], and Cheng and Higham [8] is denoted by GMW, SE90, SE99 and CH, respectively. The results show two main features. First, the four modified Cholesky algorithms provide bounds of similar quality for all but the RiskMetrics matrices, and these bounds are often of the correct order of magnitude. Second, for all the RiskMetrics matrices except RiskMetrics4, $d_{corr}(A)$ is relatively small and the revised Schnabel and Eskow [27] algorithm and the Cheng and Higham algorithm provide bounds three or four orders of magnitude smaller than those from the other two algorithms. Since the Cheng and Higham algorithm gives the best bounds overall, we use it in the remaining experiments.

We next compute all our bounds. The results are given in Table 5.2 and Table 5.3. The ordering of the bounds is the same as in Table 4.1, but note that we exclude the bound (3.5) as for our test matrices it is the same as (3.2). The bound α_1 is zero for all examples where $a_{ii} \equiv 1$ and $|a_{ij}| \leq 1$. Moreover, the circulant mean A_c of several of these matrices turns out to be a correlation matrix and so $d_{corr}(A_c) = 0$ in (2.7).

Ex.	$\|A - \text{GMW}\|_F$	$\|A - \text{SE90}\|_F$	$\|A - \text{SE99}\|_F$	$\|A - \text{CH}\|_F$	$d_{\text{corr}}(A)$
high02	8.45e-1	6.19e-1	6.19e-1	5.86e-1	5.28e-1
tec03	8.17e-2	9.47e-1	6.33e-2	5.19e-2	3.74e-2
bhwi01	6.31e-1	2.50e-1	2.50e-1	4.30e-1	1.51e-1
mmb13	3.13e1	3.14e1	3.14e1	3.04e1	3.03e1
fing97	1.50e-1	7.66e-2	7.90e-2	9.24e-2	4.91e-2
tyda99R1	2.18	2.35	2.17	2.36	1.40
tyda99R2	1.53	2.19	1.57	1.71	7.75e-1
tyda99R3	1.47	1.46	1.49	1.09	6.72e-1
usgs13	9.69e-1	8.02e-1	5.93e-1	1.92	5.51e-2
RiskMetrics1	1.11e2	1.12e1	1.72e-2	8.81e-3	3.88e-5
RiskMetrics2	1.38e2	6.21	1.90e-2	9.71e-3	4.75e-5
RiskMetrics3	9.33e1	3.14	7.10e-3	4.20e-3	1.81e-5
RiskMetrics4	1.27e2	6.66e1	6.62e1	1.22	8.40e-2
RiskMetrics5	1.32e2	1.44	2.01e-2	9.64e-3	4.46e-5
RiskMetrics6	8.53e1	1.30	5.85e-3	2.85e-3	1.59e-5
cor1399	3.59e2	3.57e2	3.57e2	4.52e1	2.10e1
cor3120	7.83e1	4.18e2	4.16e2	4.40e2	5.44

	high02	tec03	bhwi01	mmb13	fing97	tyda99R1	tyda99R2	tyda99R3
	Lower bounds							
(2.2)	0.00	0.00	0.00	3.01e1	0.00	0.00	0.00	0.00
(2.3)	4.14e-1	2.78e-2	1.28e-1	2.15e1	3.83e-2	1.15	6.24e-1	5.59e-1
(2.7)	0.00	0.00	0.00	1.17e1	0.00	1.60e-1	0.00	0.00
	True distance							
$d_{\text{corr}}(A)$	5.28e-1	3.74e-2	1.51e-1	3.03e1	4.91e-2	1.40	7.75e-1	6.72e-1
	Upper bounds							
(2.4)	2.00	2.35	2.43	3.29e1	3.09	4.02	4.02	3.74
(2.6)	9.15e-1	2.03	2.21	3.04e1	2.32	3.98	2.81	3.73
(2.7)	1.15	2.08	2.35	3.98e1	2.50	3.24	2.11	3.28
(3.1)	5.38e-1	3.93e-2	1.61e-1	3.04e1	5.33e-2	1.45	8.41e-1	7.02e-1
(3.2)	1.18	1.11e-1	5.00e-1	4.54e1	1.88e-1	3.55	2.39	2.11
(3.6)	5.86e-1	6.35e-2	2.75e-1	3.14e1	1.14e-1	2.02	1.46	1.25
(3.12)	5.86e-1	5.19e-2	4.30e-1	3.04e1	9.24e-2	2.36	1.71	1.09
(3.13)	1.15	2.08	2.35	3.04e1	2.60	3.71	2.20	3.70

<div align="center">
TABLE 5.3

Real-life examples.
</div>

	usgs13	RiskMetrics1	RiskMetrics2	RiskMetrics3	RiskMetrics4	RiskMetrics5	RiskMetrics6	cor1399	cor3120
					Lower bounds				
(2.2)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.36e-1	1.06e-1
(2.3)	5.02e-2	3.39e-5	3.97e-5	1.63e-5	8.25e-2	3.76e-5	1.46e-5	1.32e1	2.29
(2.7)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
					True distance				
$d_{corr}(A)$	5.51e-2	3.88e-5	4.75e-5	1.81e-5	8.40e-2	4.46e-5	1.59e-5	2.10e1	5.44
					Upper bounds				
(2.4)	2.29e1	1.28e2	1.54e2	1.45e2	1.30e2	1.45e2	1.41e2	3.59e2	4.28e2
(2.6)	2.04e1	1.20e2	1.42e2	1.33e2	1.18e2	1.33e2	1.28e2	3.58e2	4.28e2
(2.7)	6.75	1.10e2	1.33e2	1.23e2	1.04e2	1.22e2	1.15e2	2.05e2	2.92e2
(3.1)	6.55e-2	1.19e-4	1.90e-4	5.88e-5	9.21e-2	1.71e-4	4.83e-5	2.37e1	1.11e1
(3.2)	1.15	1.00e-3	1.27e-3	7.35e-4	1.01e1	1.16e-3	9.98e-4	3.36e2	1.83e2
(3.6)	1.01	9.58e-4	1.22e-3	7.12e-4	9.91	1.11e-3	9.74e-4	3.21e2	1.79e2
(3.12)	1.92	8.81e-3	9.71e-3	4.20e-3	1.22	9.64e-3	2.85e-3	4.52e1	4.40e2
(3.13)	7.64	1.16e2	1.43e2	1.32e2	1.10e2	1.28e2	1.21e2	2.06e2	2.93e2

Several observations can be made about the results.

(a) Of the lower bounds, only (2.3) provides useful information. Moreover, in all examples this bound is within a factor 2.4 of $\mathrm{d_{corr}}$ (the worst case being cor3120).

(b) Of the upper bounds, (3.1)—the most expensive bound to compute—is the most accurate and is always within a factor 4 of $\mathrm{d_{corr}}$ (the worst case being RiskMetrics2).

(c) Over all the test matrices, the upper bound (3.1) exceeded the lower bound (2.3) by at most a factor 4.9 (the worst case being cor3120).

(d) Of the other eigenvalue-based upper bounds, the bound from shrinking (3.6) is better than the bound (3.2), as we already know from Theorem 3.5. The shrinking bound (3.6) is typically an order of magnitude larger than (3.1) on real-life examples.

(e) The upper bounds (3.6) and (3.12) based on shrinking and the modified Cholesky factorizations, respectively, are of similar quality and they overestimate $\mathrm{d_{corr}}$ at most by one or two orders of magnitude. The modified Cholesky bound has the advantage of being computable in half the number of operations as the bound based on shrinking.

(f) The upper bounds (2.4), (2.6), and (3.13), which are computable in $O(n^2)$ operations, are poor in these tests, the more so when the distance is small.

6. Conclusions. This is the first thorough treatment of upper and lower bounds for the distance of a symmetric matrix A to the nearest correlation matrix. For the most common case in practice, in which A is indefinite with unit diagonal and $|a_{ij}| \leq 1$ for $i \neq j$, we have obtained upper bounds (3.1), (3.5), and (3.6) that differ from the lower bound (2.3) by a factor at most $1 + n\sqrt{n}$. For the sharpest bound (3.1) we found the ratio to be always less than 5 in our experiments with matrices of dimension up to 3120, so the upper bound was demonstrably of the correct order of magnitude in every case. The cost of computing the pair (2.3) and (3.1) is $17n^3/6$ flops, which is substantially less than the $70n^3/3$ or more flops required to compute ncm(A) by the preconditioned Newton algorithm of [6].

The upper bound (3.6) based on shrinking has about half the cost of (3.1) and, while less sharp than (3.1), it still performed well in our tests.

The modified Cholesky bound (3.12) has the attraction that it provides an inexpensive test for definiteness ($n^3/3$ flops) along with an upper bound (costing another $n^3/3$ flops) that, while sometimes two orders of magnitude larger than (3.1), can still provide useful information.

We conclude that our bounds are well suited to gauging the size of $\mathrm{d_{corr}}(A)$. The information they provide enables a user to identify an invalid correlation matrix relatively cheaply and to decide whether to revisit its construction or proceed to compute a replacement directly from it. A natural replacement is the nearest correlation matrix itself; an alternative is to use shrinking [17].

Our test matrices can be downloaded in MATLAB form from `https://github.com/higham/matrices-correlation-invalid`, with the exception of the RiskMetrics matrices, which we do not have permission to distribute. The MATLAB implementation of the modified Cholesky algorithm from [8] that we used is available from `https://github.com/higham/modified-cholesky`.

13

REFERENCES

[1] Greg Anderson, Lisa Goldberg, Alec N. Kercheval, Guy Miller, and Kathy Sorge. On the aggregation of local risk models for global risk management. *Journal of Risk*, 8(1):25–40, 2005.

[2] Jared L. Aurentz, Raf Vandebril, and David S. Watkins. Fast computation of the zeros of a polynomial via factorization of the companion matrix. *SIAM J. Sci. Comput.*, 35(1): A255–A269, 2013.

[3] Zhaojun Bai, James W. Demmel, Jack J. Dongarra, Axel Ruhe, and Henk A. Van der Vorst, editors. *Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2000. xxix+410 pp. ISBN 0-89871-471-0.

[4] Vineer Bhansali and Mark B. Wise. Forecasting portfolio risk in normal and stressed markets. *Journal of Risk*, 4(1):91–106, 2001.

[5] Madalyn S. Blondes, John H. Schuenemeyer, Ricardo A. Olea, and Lawrence J. Drew. Aggregation of carbon dioxide sequestration storage assessment units. *Stochastic Environmental Research and Risk Assessment*, 27(8):1839–1859, 2013.

[6] Rüdiger Borsdorf and Nicholas J. Higham. A preconditioned Newton algorithm for the nearest correlation matrix. *IMA J. Numer. Anal.*, 30(1):94–107, 2010.

[7] Rüdiger Borsdorf, Nicholas J. Higham, and Marcos Raydan. Computing a nearest correlation matrix with factor structure. *SIAM J. Matrix Anal. Appl.*, 31(5):2603–2622, 2010.

[8] Sheung Hun Cheng and Nicholas J. Higham. A modified Cholesky algorithm based on a symmetric indefinite factorization. *SIAM J. Matrix Anal. Appl.*, 19(4):1097–1110, 1998.

[9] Susan J. Devlin, R. Gnanadesikan, and J. R. Kettenring. Robust estimation and outlier detection with correlation coefficients. *Biometrika*, 62(3):531–545, 1975.

[10] Richard L. Dykstra. An algorithm for restricted least squares regression. *J. Amer. Statist. Assoc.*, 78:837–842, 1983.

[11] Christopher C. Finger. A methodology to stress correlations. *RiskMetrics Monitor*, Fourth Quarter:3–11, 1997.

[12] Philip E. Gill, Walter Murray, and Margaret H. Wright. *Practical Optimization*. Academic Press, London, 1981.

[13] Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. Fourth edition, Johns Hopkins University Press, Baltimore, MD, USA, 2013. xxi+756 pp. ISBN 978-1-4214-0794-4.

[14] Nicholas J. Higham. Computing a nearest symmetric positive semidefinite matrix. *Linear Algebra Appl.*, 103:103–118, 1988.

[15] Nicholas J. Higham. Computing the nearest correlation matrix—A problem from finance. *IMA J. Numer. Anal.*, 22(3):329–343, 2002.

[16] Nicholas J. Higham and Nataša Strabić. Anderson acceleration of the alternating projections method for computing the nearest correlation matrix. *Numer. Algorithms*, 2015. Online First.

[17] Nicholas J. Higham, Nataša Strabić, and Vedran Šego. Restoring definiteness via shrinking, with an application to correlation matrices with a fixed block. MIMS EPrint 2014.54, Manchester Institute for Mathematical Sciences, The University of Manchester, UK, November 2014. 20 pp. Revised June 2015. To appear in SIAM Rev.

[18] Nicholas J. Higham and Françoise Tisseur. A block algorithm for matrix 1-norm estimation, with an application to 1-norm pseudospectra. *SIAM J. Matrix Anal. Appl.*, 21(4):1185–1201, 2000.

[19] Craig Lucas. Computing nearest covariance and correlation matrices. M.Sc. Thesis, University of Manchester, Manchester, England, October 2001. 68 pp.

[20] Jerome Malick. A dual approach to solve semidefinite least-squares problems. *SIAM J. Matrix Anal. Appl.*, 26(1):272–284, 2004.

[21] Aleksei Minabutdinov, Ilia Manaev, and Maxim Bouev. Finding the nearest valid covariance matrix: A FX market case. Working paper Ec-07/13, Department of Economics, European University at St. Petersburg, St. Petersburg, Russia, 2013. Revised June 2014.

[22] NAG Toolbox for MATLAB. NAG Ltd., Oxford, UK. http://www.nag.co.uk.

[23] Hou-Duo Qi and Defeng Sun. A quadratically convergent Newton method for computing the nearest correlation matrix. *SIAM J. Matrix Anal. Appl.*, 28(2):360–385, 2006.

[24] HouDuo Qi and Defeng Sun. Correlation stress testing for value-at-risk: an unconstrained convex optimization approach. *Comput Optim Appl*, 45(2):427–462, 2010.

[25] Jiří Rohn. Computing the norm $\|A\|_{\infty,1}$ is NP-hard. *Linear and Multilinear Algebra*, 47(3): 195–204, 2000.

[26] Robert B. Schnabel and Elizabeth Eskow. A new modified Cholesky factorization. *SIAM J.*

      Sci. Statist. Comput., 11(6):1136–1158, 1990.

[27] Robert B. Schnabel and Elizabeth Eskow. A revised modified Cholesky factorization algorithm.
      SIAM J. Optim., 9(4):1135–1148, 1999.

[28] James R. Schott. *Matrix Analysis for Statistics*. Wiley, New York, 1997.

[29] N. C Schwertman and D. M Allen. Smoothing an indefinite variance-covariance matrix. *J.
      Statist. Comput. Simul.*, 9(3):183–194, 1979.

[30] M. V. Travaglia. A lower bound for the nearest correlation matrix problem based on the
      circulant mean. *Comp. Appl. Math.*, 33:27–44, 2014.

[31] William F. Trench. Numerical solution of the eigenvalue problem for Hermitian Toeplitz ma-
      trices. *SIAM J. Matrix Anal. Appl.*, 10(2):135–146, 1989.

[32] Saygun Turkay, Eduardo Epperlein, and Nicos Christofides. Correlation stress testing for value-
      at-risk. *Journal of Risk*, 5(4):75–89, 2003.

[33] Rajesh Tyagi and Chandrasekhar Das. Grouping customers for better allocation of resources to
      serve correlated demands. *Computers & Operations Research*, 26(10-11):1041–1058, 1999.

[34] U.S. Geological Survey Geologic Carbon Dioxide Storage Resources Assessment Team. *National
      Assessment of Geologic Carbon Dioxide Storage Resources—Results (Ver. 1.1, September
      2013)*, September 2013.

[35] Kefeng Xu and Philip T. Evers. Managing single echelon inventories through demand aggrega-
      tion and the feasibility of a correlation matrix. *Computers & Operations Research*, 30(2):
      297–308, 2003.