

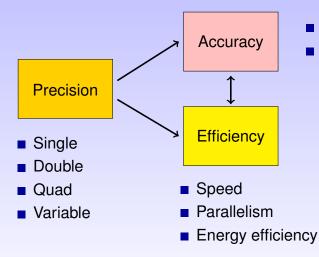
Rehabilitating Correlations, Avoiding Inversion, and Extracting Roots

Nick Higham School of Mathematics The University of Manchester

Nicholas.J.Higham@manchester.ac.uk @nhigham,nickhigham.wordpress.com

http://www.maths.manchester.ac.uk/~higham/talks

The Actuarial Profession March 20, 2013, London



EstimatesGuarantees

Outline

Rehabilitating Correlations

2 Matrix Inversion



Correlation Matrix

An $n \times n$ symmetric matrix A is a correlation matrix if

- It has ones on the diagonal.
- All its eigenvalues are nonnegative.

Correlation Matrix

An $n \times n$ symmetric matrix A is a correlation matrix if

- It has ones on the diagonal.
- All its eigenvalues are nonnegative.

Is this a correlation matrix?

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

Correlation Matrix

An $n \times n$ symmetric matrix A is a correlation matrix if

- It has ones on the diagonal.
- All its eigenvalues are nonnegative.

Is this a correlation matrix?

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$
. Spectrum: -0.4142 , 1.0000 , 2.4142 .

Question from London Fund Management Company (2000)

"Given a real symmetric matrix *A* which is almost a correlation matrix . . .

- What is the best approximating (in Frobenius norm?) correlation matrix?
- Is it unique?
- Can we compute it?

Typically we are working with 1400 \times 1400 at the moment, but this will probably grow to 6500 \times 6500."

How to Proceed

- Make ad hoc modifications to matrix: e.g., shift negative e'vals up to zero then diagonally scale.
- Plug the gaps in the missing data, then compute an exact correlation matrix.
- Compute the **nearest correlation matrix** in the weighted Frobenius norm ($||A||^2 = \sum_{i,j} w_i w_j a_{ij}^2$). Given approx correlation matrix A find correlation matrix C to minimize ||A C||.
 - Constraint set is a closed, convex set, so unique minimizer.

Development

Derived theory and algorithm:

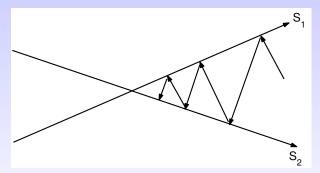
N. J. Higham, Computing the Nearest Correlation Matrix—A Problem from Finance, IMA J. Numer. Anal. 22, 329–343, 2002.

Extensions in:

Craig Lucas, Computing Nearest Covariance and Correlation Matrices, M.Sc. Thesis, University of Manchester, 2001.

Alternating Projections Algorithm

von Neumann (1933): for subspaces.Dykstra (1983): corrections for closed convex sets.

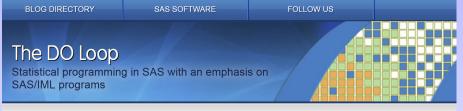


- Easy to implement.
- Guaranteed convergence, at a linear rate.
- Can add further constraints/projections.

Unexpected Applications

Some recent papers:

- Applying stochastic small-scale damage functions to German winter storms (2012)
- Estimating variance components and predicting breeding values for eventing disciplines and grades in sport horses (2012)
- Characterisation of tool marks on cartridge cases by combining multiple images (2012)
- Experiments in reconstructing twentieth-century sea levels (2011)



SAS BLOGS HOME > THE DO LOOP > COMPUTING THE NEAREST CORRELATION MATRIX

Computing the nearest correlation matrix



Rick Wicklin | NOVEMBER 28, 2012

Ο 1028 📮 0 😏 Tweet < 6



Frequently someone will post a question to the SAS Support Community that says something like this:

I am trying to do [statistical task] and SAS issues an error and reports that my correlation matrix is not positive definite. What is going on and how can I complete [the task]?

The statistical task varies, but one place where this problem occurs is in simulating multivariate normal data. I have previously written about why an estimated matrix of pairwise correlations is not always a valid correlation matrix. This article discusses what to do about it. The material in this article is taken from my forthcoming book, Simulating Data with SAS

Newton Method

Qi & Sun (2006): **Newton method** based on theory of strongly semismooth matrix functions.

- Applies Newton to **dual** (unconstrained) of $\min \frac{1}{2} ||A X||_F^2$ problem.
- Globally and quadratically convergent.

H & Borsdorf (2010) improve efficiency and reliability:

- use minres for Newton equation,
- Jacobi preconditioner,
- reliability improved by line search tweaks,
- extra scaling step to ensure unit diagonal.

Nick Higham

Applied mathematics, software and workflow.



Home About

← The Lanczos Tapes

SIAM Conference on Computational Science and Engineering 2013 Preview →

The Nearest Correlation Matrix

Posted on February 13, 2013

A correlation matrix is a symmetric matrix with unit diagonal and nonnegative eigenvalues. In 2000 I was approached by a London fund management company who wanted to find the nearest correlation matrix (NCM) in the Frobenius norm to an *almost correlation matrix*: a symmetric matrix having a significant number of (small) negative eigenvalues. This problem arises when the data from which the correlations are constructed is asynchronous or incomplete, or when models are stress-tested by artificially adjusting individual correlations. Solving the NCM problem (or obtaining a true correlation matrix some other way) is important in order to avoid subsequent calculations breaking down due to negative variances or volatilities, for example.



Recent Posts

- How Accurate Are Spreadsheets in the Cloud?
- SIAM Conference on Computational Science and Engineering 2013
- Tiny Relative Errors
- SIAM Conference on Computational Science and Engineering 2013 Preview
- The Nearest Correlation Matrix

Archives

- March 2013
- February 2013
- January 2013

Original NCM Problem

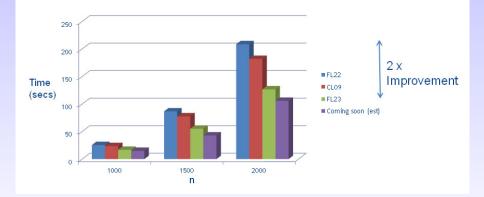
- Alternating Projections Method
 - MATLAB: Craig Lucas, <u>near cor</u>.
 - MATLAB: Erlend Ringstad, validcorr.
 - MATLAB: Nick Higham, randcorr (see below).
 - R: Jens Oehlschlaegel and R Matrix package authors, <u>nearPD</u>.
 - SAS: Rick Wicklin, <u>NearestCorr</u>.
- Newton Method
 - MATLAB: Defeng Sun, <u>various codes</u>.
 - NAG Library (Fortran/SMP, C, NAG Toolbox for MATLAB, <u>NET</u>, <u>Java</u>, <u>Excel</u>, <u>R</u>):
 - g02aaf (Fortran), g02aac (C), g02aa.m (MATLAB), g02aa (R),
 - <u>go2abf</u> (Fortran), go2abc (C), go2ab.m (MATLAB), <u>go2ab</u> (R): allows for weighted Frobenius norm based on two-side scaling and lower bounds on the eigenvalues.
 - g02ajf (Fortran), g02ajf.m (MATLAB) to appear in Mark 24: allows for componentwise weighted Frobenius norm and lower bounds on the eigenvalues.

Alternating Projections vs Newton

Matrix	tol	Code	Time (s)	Iters
1. Random (100)	1e-10	g02aa	0.023	4
		nearcorr	0.052	15
2. Random (500)	1e-10	g02aa	0.48	4
		nearcorr	3.01	26
3. Real-life (1399)	1e-4	g02aa	6.8	5
		nearcorr	100.6	68

Correlations Matrix Inversion Matrix Roots

Performance of NAG Codes



Factor Model (1)

$$\xi = \underbrace{X}_{n \times k} \underbrace{\eta}_{k \times 1} + \underbrace{F}_{n \times n} \underbrace{\varepsilon}_{n \times 1},$$

 $\eta_i, \varepsilon_i \in N(0, 1),$

where $var(\xi_i) \equiv 1$, $F = diag(f_{ii})$. Implies

$$\sum_{j=1}^{k} x_{ij}^2 \le 1, \qquad i = 1: \ n.$$

- "Multifactor normal copula model".
- Collateralized debt obligations (CDOs).
- Multivariate time series.

Factor Model (2)

Yields correlation matrix of form

$$egin{aligned} \mathcal{C}(X) &= \mathcal{D} + XX^{ op} = \mathcal{D} + \sum_{j=1}^k x_j x_j^{ op}, \ \mathcal{D} &= ext{diag}(I - XX^{ op}), \qquad X = [x_1, \dots, x_k]. \end{aligned}$$

C(X) has k factor correlation matrix structure.

$$C(X) = \begin{bmatrix} 1 & y_1^T y_2 & \dots & y_1^T y_n \\ y_1^T y_2 & 1 & \dots & \vdots \\ \vdots & & \ddots & y_{n-1}^T y_n \\ y_1^T y_n & \dots & y_{n-1}^T y_n & 1 \end{bmatrix}, \quad y_i \in \mathbb{R}^k.$$

Factor Structure

Nearest correlation matrix with factor structure.

 Principal factors method (Andersen et al., 2003) has no convergence theory and can converge to an incorrect answer.

Factor Structure

Nearest correlation matrix with factor structure.

- Principal factors method (Andersen et al., 2003) has no convergence theory and can converge to an incorrect answer.
- Algorithm based on spectral projected gradient method (Borsdorf, H & Raydan, 2010).
 - Respects the constraints, exploits their convexity, and converges to a feasible stationary point.
 - NAG routine g02aef (Mark 23, 2012).

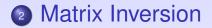
Variations

Specific rank or bound on correlations.

- Block structure.
- Bounds on individual correlations.
- Other requirements?

Outline

Rehabilitating Correlations



Matrix Roots

Avoiding Inversion

Fundamental Tenet of Numerical Analysis

Don't invert matrices.

Avoiding Inversion

Fundamental Tenet of Numerical Analysis

Don't invert matrices.

• Ax = b: use Gaussian elimination (with pivoting), not $x = A^{-1}b$.

$$x^T A^{-1} y = x^T (A^{-1} y).$$

$$(A^{-1})_{ii} = e_i^T A^{-1} e_i.$$

■ $\frac{\text{trace}(A^{-1})}{v_k} \approx m^{-1} \sum_{k=1}^m v_k^T A^{-1} v_k$, $v_k \sim \text{uniform}\{-1, 1\}$. (Bekas et al., 2007)

How to Invert (1)

 Use a condition estimator to warn when A is nearly singular.

>> A = hilb(16); >> X = inv(A); Warning: Matrix is close to singular or badly scaled. Results may be inaccurate. RCOND = 9.721674e-19.

How to Invert (2)

- To compute variance–covariance matrix (X^TX)⁻¹ of least squares estimator, use QR factorization X = QR ((X^TX)⁻¹ = R⁻¹R^{-T}) and *do not* explicitly form X^TX.
- Quality of computed inverse $\widehat{Y} \approx A^{-1}$ measured by

$$\frac{\|\widehat{Y}A - I\|}{\|\widehat{Y}\| \|A\|} \quad or \quad \frac{\|A\widehat{Y} - I\|}{\|\widehat{Y}\| \|A\|}$$

but not both.

Computing the Sample Variance

Sample variance of x_1, \ldots, x_n :

$$s_n^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$
, where $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$. (1)

Can compute using one-pass formula:

$$s_n^2 = \frac{1}{n-1} \left(\sum_{i=1}^n x_i^2 - \frac{1}{n} \left(\sum_{i=1}^n x_i \right)^2 \right).$$
 (2)

Computing the Sample Variance

Sample variance of x_1, \ldots, x_n :

$$s_n^2 = rac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$
, where $\bar{x} = rac{1}{n} \sum_{i=1}^n x_i$. (1)

Can compute using one-pass formula:

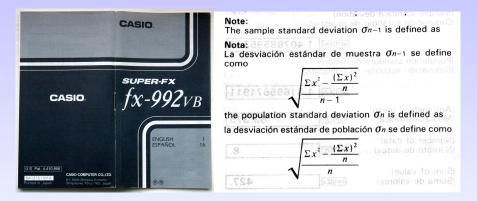
$$s_n^2 = \frac{1}{n-1} \left(\sum_{i=1}^n x_i^2 - \frac{1}{n} \left(\sum_{i=1}^n x_i \right)^2 \right).$$
 (2)

For x = (10000, 10001, 10002) using 8-digit arithmetic, (1) gives: 1.0, (2) gives 0.0.

(2) can even give negative results in floating point arithmetic!

University of Manchester

Casio fx-992VB



Spreadsheets in the Cloud

Standard deviation of $x = [n, n+1, n+2]^T$.

n	Exact	Google Sheet
10 ⁷	1	1
10 ⁸	1	0

Spreadsheets in the Cloud

Standard deviation of $x = [n, n+1, n+2]^T$.

n	Exact	Google Sheet
10 ⁷	1	1
10 ⁸	1	0



Applied mathematics, software and workfloss.



-- SIAM Conference on Computational Science and Engineering 2013

How Accurate Are Spreadsheets in the Cloud?

Posted on Merch 13, 2013

For a voter s with - dement the sample variants is $\frac{1}{2} - \frac{1}{2}\sum_{i}(r_i - 2^{i})$, where the mode $i = \frac{1}{2}\sum_{i}(r_i - 2^{i})$, where the there denotes its is a simple near $i = 1 - \frac{1}{2}\sum_{i}(r_i - 2^{i})$. This second formula to the absolute that if an isotropic variant of the solution is $\frac{1}{2} - \frac{1}{2}\sum_{i}(r_i - 2^{i})$. This second formula has the absolute that if an isotropic variant of the solution is the solution

Search

Recent Posts

- How Accurate Are Spreadsheets in the Cloud?
- SIAM Conference on Computational Science and Engineer-
- ing 2013
- SIAM Conference on Computational Science and Engineering
- 2013 Preview The Newcest Correlation Matrix
- Archives
- March 2013
- February 2013

Nick Higham

Actuarial Matrix Computations

Spreadsheets in the Cloud

Standard deviation of $x = [n, n+1, n+2]^T$.

n	Exact	Google Sheet
10 ⁷	1	1
10 ⁸	1	0



Applied mathematics, software and



-- SIAM Conference on Computational Science and

How Accurate Are Spreadsheets in the Cloud?

Posted on Merch 13, 2013

For a vector x with n elements the sample variance is $s_n^2 = \frac{1}{-1} \sum_{i=1}^n (x_i - \overline{x})^2$, where the sample mean is $x = \frac{1}{2} \sum_{i=1}^{n} x_i$. An alternative formula often given in textbooks is $s_n^2 = \frac{1}{n-1} \left(\sum_{i=1}^n x_i^2 - \frac{1}{n} \left(\sum_{i=1}^n x_i \right)^2 \right)$. This second formula has the advantage that it can be computed with just one pass through the data, whereas the first formula requires two passes. However, the one-pass formula can suffer damaging subtractive cancellation, making it numerically unstable. When I wrote my book Accuracy and Stability of Numerical Algorithms I found that several pocket calculators appeared to use the one-pass

Search

Recent Posts

- · How Accurate Are Spoudsheets in
- · SIAM Conference on Computa-
- · SIAM Conference on Computational Science and Engineering
- . The Nearest Correlation Matrix

Archiver







Gallery The journal performed a series of standard statistical tests against three well-known cloud spreadsheets: Google Spreadsheet, Microsoft's Excel Web App and Zoho Sheet.

McCullough & Yalta (2013)

University of Manchester

Nick Higham

Actuarial Matrix Computations

Outline

Rehabilitating Correlations

2 Matrix Inversion

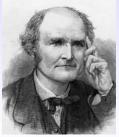


Cayley and Sylvester

Term "matrix" coined in 1850 by James Joseph Sylvester, FRS (1814–1897).

 Matrix algebra developed by Arthur Cayley, FRS (1821– 1895).
 Memoir on the Theory of Matrices (1858).





Cayley and Sylvester on Matrix Functions

 Cayley considered matrix square roots in his 1858 memoir.

Tony Crilly, Arthur Cayley: Mathematician Laureate of the Victorian Age, 2006.

Sylvester (1883) gave first definition of f(A) for general f.

Karen Hunger Parshall, James Joseph Sylvester. Jewish Mathematician in a Victorian World, 2006.



Matrix Roots in Markov Models

- Let vectors v₂₀₁₁, v₂₀₁₀ represent risks, credit ratings or stock prices in 2011 and 2010.
- Assume a Markov model $v_{2011} = P v_{2010}$, where *P* is a transition probability matrix.
- *P*^{1/2} enables predictions to be made at 6-monthly intervals.

Matrix Roots in Markov Models

- Let vectors v₂₀₁₁, v₂₀₁₀ represent risks, credit ratings or stock prices in 2011 and 2010.
- Assume a Markov model $v_{2011} = P v_{2010}$, where *P* is a transition probability matrix.
- *P*^{1/2} enables predictions to be made at 6-monthly intervals.

 $P^{1/2}$ is matrix X such that $X^2 = P$. What are $P^{2/3}$, $P^{0.9}$?

 $P^s = \exp(s \log P).$

Problem: log *P*, $P^{1/k}$ may have wrong sign patterns \Rightarrow "regularize".

Chronic Disease Example

- Estimated 6-month transition matrix.
- Four AIDS-free states and 1 AIDS state.
- 2077 observations (Charitos et al., 2008).

Want to estimate the 1-month transition matrix.

$$\Lambda(\mathbf{P}) = \{1, 0.9644, 0.4980, 0.1493, -0.0043\}.$$

H & Lin (2011).

Lin (2011, for survey of regularization methods.

Correlations Matrix Inversion Matrix Roots

MATLAB: Arbitrary Powers

```
>> A = [1 \ 1e-8; \ 0 \ 1]
A =
  1.0000e+000 1.0000e-008
             0 1.0000e+000
>> A^0.1
ans =
     1
            0
     0
            1
>> expm(0.1*loqm(A))
ans =
  1.0000e+000 1.0000e-009
             0 1.0000e+000
```

MATLAB Arbitrary Power

- New Schur algorithm (H & Lin, 2011, 2013) reliably computes A^p for any real p.
- New backward-error based inverse scaling and squaring alg for matrix logarithm (Al-Mohy, H & Relton, 2012)—faster and more accurate.
- Alternative Newton-based algorithms available for A^{1/q} with q an integer, e.g., for

$$X_{k+1} = rac{1}{q} [(q+1)X_k - X_k^{q+1}A], \qquad X_0 = A,$$

 $X_k o A^{-1/q}.$

Knowledge Transfer Partnership

- University of Manchester and NAG (2010–2013) funded by EPSRC, NAG and TSB.
- Developing suite of NAG Library codes for matrix functions.
- Extensive set of new codes included in Mark 23 (2012), Mark 24 (2013).
- Improvements to existing state of the art: faster and more accurate.

My work also supported by $\in 2M$ ERC Advanced Grant.

nag blog

WEDNESDAY, 11 JULY 2012

U

The Matrix Square Root, Blocking and Parallelism

NAG recently embarked on a 'Knowledge Transfer Partnership' with the University of Manchester to introduce matrix function capabilities into the NAG Library. As part of this collaboration, Nick Higham (University of Manchester), Rui Ralha (University of Minho, Portugal) and I have been investigating how blocking can be used to speed up the computation of matrix square roots.

There is plenty of interesting mathematical theory concerning matrix square roots, but for now we'll just use the definition that a matrix X is a square root of A if X^2 =A. Matrix roots have applications in finance and population modelling, where transition matrices are used to describe the evolution of a system from over a certain time interval, t. The square root of a transition matrix can be used to describe the evolution for the interval t/2. The matrix square root also forms a key part of the algorithm used to compute other matrix functions.

To find a square root of a matrix, we start by computing a Schur decomposition. The square root U of the resulting upper triangular matrix T can then be found via a simple recurrence over the elements U_{μ} and T_{μ} :

$$U_{ii} = \sqrt{T_{ii}},$$

 $U_{ii}U_{ij} + U_{ij}U_{jj} = T_{ij} - \sum_{k=l+1}^{j-1} U_{ik}U_{kj}.$

FOLLOW BY EMAIL

Email address... Submit

PRODUCTS & SERVICES

- Numerical Libraries
- Compilers
- GPU Software & Consulting
- HPC Applications Support
- HPC Technology Consulting
- Software Engineering Services
- Training
- Partner Program
- Research & Development



WEDNESDAY, 16 JANUARY 2013

Matrix Functions in Parallel

Last year I wrote a blog post about NAG's work on parallelising the computation of the matrix square root. More recently, as part of our Matrix Functions Knowledge Transfer Partnership with the University of Manchester, we've been investigating parallel implementations of the Schur-Parlett algorithm [1].

blog

Most algorithms for computing functions of matrices are tailored for a specific function, such as the matrix exponential or the matrix square root. The Schur-Parlett algorithm is much more general; it will work for any "well behaved" function (this general term can be given a more mathematically precise meaning). For a function such as

$$f(A) = e^A + \sin 2A - \cosh 4A,$$

FOLLOW BY EMAIL

Email address...

Submit

PRODUCTS & SERVICES

- Numerical Libraries
- Compilers
- GPU Software & Consulting
- HPC Applications Support
- HPC Technology Consulting
- Software Engineering Services
- Training
- Partner Program
- Research & Development

Final Remarks

- Fast moving developments in numerical linear algebra algorithms.
- Numerical reliability is essential.
- Partnership with NAG enables rapid inclusion of our algorithms in the NAG Library.

Keen to hear about your matrix problems.

A. H. Al-Mohy and N. J. Higham. Improved inverse scaling and squaring algorithms for the matrix logarithm. SIAM J. Sci. Comput., 34(4):C153–C169, 2012.

 A. H. Al-Mohy, N. J. Higham, and S. D. Relton.
 Computing the Fréchet derivative of the matrix logarithm and estimating the condition number.
 MIMS EPrint 2012.72, Manchester Institute for Mathematical Sciences, The University of Manchester, UK, July 2012.
 17 pp.
 Bevised December 2012.

References II

- L. Anderson, J. Sidenius, and S. Basu. All your hedges in one basket. *Risk*, pages 67–72, Nov. 2003. |www.risk.net|.
- C. Bekas, E. Kokiopoulou, and Y. Saad. An estimator for the diagonal of a matrix. *Appl. Numer. Math.*, 57:1214–1229, 2007.
- R. Borsdorf, N. J. Higham, and M. Raydan. Computing a nearest correlation matrix with factor structure.

SIAM J. Matrix Anal. Appl., 31(5):2603–2622, 2010.

References III

T. Charitos, P. R. de Waal, and L. C. van der Gaag. Computing short-interval transition matrices of a discrete-time Markov chain from partially observed data.

Statistics in Medicine, 27:905–921, 2008.

T. Crilly.

Arthur Cayley: Mathematician Laureate of the Victorian Age.

Johns Hopkins University Press, Baltimore, MD, USA, 2006. ISBN 0-8018-8011-4. xxi+610 pp.

References IV

- P. Glasserman and S. Suchintabandid.
 Correlation expansions for CDO pricing.
 Journal of Banking & Finance, 31:1375–1398, 2007.
- C.-H. Guo and N. J. Higham.
 A Schur–Newton method for the matrix *p*th root and its inverse.
 SIAM J. Matrix Anal. Appl., 28(3):788–804, 2006.

References V

N. J. Higham. Accuracy and Stability of Numerical Algorithms. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, second edition, 2002. ISBN 0-89871-521-0. xxx+680 pp.

N. J. Higham.

Computing the nearest correlation matrix—A problem from finance.

IMA J. Numer. Anal., 22(3):329-343, 2002.

References VI

- N. J. Higham and L. Lin. On *p*th roots of stochastic matrices. *Linear Algebra Appl.*, 435(3):448–463, 2011.
- N. J. Higham and L. Lin.
 An improved Schur–Padé algorithm for fractional powers of a matrix and their Fréchet derivatives.
 MIMS EPrint 2013.1, Manchester Institute for Mathematical Sciences, The University of Manchester, UK, Jan. 2013.
 20 pp.

References VII

L. Lin.

- Roots of Stochastic Matrices and Fractional Matrix Powers.
- PhD thesis, The University of Manchester, Manchester, UK, 2010. 117 pp. MIMS EPrint 2011.9, Manchester Institute for Mathematical Sciences.
- B. D. McCullough and A. T. Yalta. Spreadsheets in the cloud—not ready yet. Journal of Statistical Software, 52:1–14, 2013.

K. H. Parshall.

James Joseph Sylvester. Jewish Mathematician in a Victorian World.

Johns Hopkins University Press, Baltimore, MD, USA, 2006. ISBN 0-8018-8291-5. xiii+461 pp.

H.-D. Qi and D. Sun.

A quadratically convergent Newton method for computing the nearest correlation matrix. *SIAM J. Matrix Anal. Appl.*, 28(2):360–385, 2006.