A NOTE ON JACOBI BEING MORE ACCURATE THAN QR

By

Walter F. Mascarenhas

IMA Preprint Series # 956

April 1992

A note on Jacobi being more accurate than QR

Walter F. Mascarenhas
Institute for Mathematics and Its Applications
University of Minnesota, Minneapolis

March 10, 1992

Abstract

In [DV], Demmel and Veselić present a theoretical and experimental analysis to show that the Jacobi method is more accurate than the QR method when computing the eigenvalues of positive definite matrices. They show that the error caused by the Jacobi method depends on the size of a factor ρ, which is related to the singular values of certain matrices associated with the Jacobi iterates. Their experiments suggest that $\rho = O(1)$. However, in this note we present a family of matrices and orderings for which $\rho = O(N)$, where N is the dimension of the matrix.

key words. eigenvalues, Jacobi method, accuracy.

1. Introduction

In [DV], Demmel and Veselić present strong evidence that the Jacobi method for computing eigenvalues of symmetric matrices is optimally accurate when applied to positive a definite matrix. Their work starts by showing that if H is positive definite and P is such that $|P_{ij}| \leq \eta \, |H_{ij}|$ then

$$\max_i \frac{|\lambda_i(H) - \lambda_i(H+P)|}{\lambda_i(H)} \leq \frac{N\eta}{\sigma(H)}, \tag{1.1}$$

where λ_i denotes the ith eigenvalue and $\sigma(H)$ is the smallest singular value of the matrix $A = D^{-1}HD^{-1}$, where D is a diagonal matrix and $A_{ii} = 1$. We notice that although Demmel and Veselić use the condition number of A instead of $1/\sigma(H)$ in (1.1), a more carefull look at their analysis shows that the sharper bound in (1.1) holds.

The bound in (1.1) is better than what can be expected from the usual pertubation theory. The usual analysis leads to a bound of the form (1.1) but with $\sigma(H)$ replaced by the smallest singular value of H, which can be much smaller than $\sigma(H)$.

The bound (1.1) is sharp, in the sense that there exist P for which equality is almost achieved in (1.1). Therefore, the best accuracy we can hope for when using a finite arithmetic with rounding ϵ is obtained by replacing η by ϵ in equation (1.1).

In order to show that the Jacobi method has optimal accuracy, Demmel and Veselić obtain a bound

$$\max_i \frac{|\lambda_i(H) - \bar{\lambda}_i(H)|}{\lambda_i(H)} \leq \frac{cMN\epsilon}{\theta(H,O)}, \tag{1.2}$$

for the relative error in the eigenvalues $\bar{\lambda}$ computed by the Jacobi method using an ordering [*] O. In this last equation c is a constant, M is the number of rotations performed until convergence and

$$\theta(H,O) = \min_k \sigma(H^{(k)}), \tag{1.3}$$

where $H^{(k)}$ are the iterates produced by the Jacobi method with ordering O to H.

Therefore, the ratio

$$\rho(H,O) = \frac{\sigma(H)}{\theta(H,O)}$$

gives an idea of how the accuracy obtained by the Jacobi method compares to the best accuracy we can expect to obtain in finite arithmetic. Demmel and Veselić conjecture that ρ is small. In fact, their numerical experiments suggest that $\rho = O(1)$ for the ordering by rows. The purpose of this note is to present matrices and orderings for which $\rho = O(N)$, where N is the dimension of the matrix. More formally, at the end of the next section we prove the following result

Theorem 1. *Given* $N = 2^n$ *and* α, $0 < \alpha < \frac{1}{2}$, *there exists a* $N \times N$ *symmetric positive definite matrix* $H = H(N,\alpha)$ *and a family of orderings* O *such that* $\rho(H,O) \geq \frac{N}{4}$.

We emphasize, however, that a ρ of order N is not enough to invalidate the claims in [DV]. If ρ can be at most $O(N)$ then Jacobi is more accurate than QR. Our search for matrices and orderings with bigger $\rho's$ was unsucessful and leads as to believe that the growth above is optimal. As it was the case with Gaussian elimination, the use of clever optimazation techniques can lead to bigger values of ρ. Even if such examples are found, they are likely to be complicated and we believe that our example gives a nice and simple example of how ρ can grow.

Finally, in the appendix we present a very short Matlab program that simulates the behavior of the Jacobi method for the orderings and matrices from Theorem 1. The experiments show that the maximum relative error in this case is asymptoticly equal to $N\epsilon/2$, where N is the dimension of the matrix and ϵ is machine precision.

2. Description of H and O

We start this section by presenting the family $H(N,\alpha)$ and the orderings O. Then we present a lemma and finally a proof of theorem 1.

---

[*] i.e. the order in which the rotations are performed

The matrices $H = H(N,\alpha)$ are rather simple. They have ones on the diagonal and $1-\alpha$ in all the other entries. In other words:

$$H_{ii} = 1, \qquad H_{ij} = 1-\alpha \ \text{ if } \ i \neq j.$$

These H are positive definite and $\lambda_{min}(H) = \alpha$ if $\alpha < 1$.

We will say that an ordering O for applying the Jacobi method is **acceptable** if it can obtained by the following recursive procedure:

- if N = 2, then pivot (1,2).
- If $N = 2^n$, $n > 1$, partition the matrix as

  $$H = \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^T & H_{22} \end{pmatrix},$$

  where the H_{ij} are $2^{n-1} \times 2^{n-1}$ blocks, pivot the entries on the main diagonal of H_{12}, then rotate the remaining entries of H_{12} and the entries in H_{11} in any ordering. Finally, apply an acceptable ordering to H_{22}.

The iterates of the Jacobi method applied according to an acceptable ordering to the matrices $H(N,\alpha)$ above can be simply described, provided we assume that in the ambiguous case of repeated diagonal entries we choose the Jacobi rotation J as, for example, Golub and Van Loan, that leads to the example below

$$J \begin{pmatrix} 1 & b \\ b & 1 \end{pmatrix} J^T = \begin{pmatrix} 1-b & 0 \\ 0 & 1+b \end{pmatrix}. \tag{2.4}$$

Let us call $\tilde{H}$ the matrix obtained after applying the rotations in the $\frac{N}{2} \times \frac{N}{2}$ blocks H_{12} and H_{22} of H. We have then the following lemma:

Lemma 1. *If the ordering O is acceptable and in the case of repeated diagonal entries we rotate as in* (2.4) *then*

$$\tilde{H}(N,\alpha) = \begin{pmatrix} \alpha I_{\frac{N}{2}} & 0 \\ 0 & (2-\alpha) H(\frac{N}{2}, \frac{\alpha}{2-\alpha}) \end{pmatrix},$$

where I_k is the $k \times k$ identity matrix.

Proof of lemma 1 We analyze here the case $N = 4$. The general case is analogous and is left for the reader. In this case we have

$$H = H(N,\alpha) = \begin{pmatrix} 1 & 1-\alpha & 1-\alpha & 1-\alpha \\ & 1 & 1-\alpha & 1-\alpha \\ & & 1 & 1-\alpha \\ & & & 1 \end{pmatrix}$$

The first rotation is at $(1,3)$ and leads to the matrix

$$H^{(1)} = \begin{pmatrix} \alpha & 0 & 0 & 0 \\ & 1 & \frac{\sqrt{2}(1-\alpha)}{2} & 1-\alpha \\ & & 2-\alpha & \frac{\sqrt{2}(1-\alpha)}{2} \\ & & & 1 \end{pmatrix}.$$

Then we pivot at $(3,4)$, getting

$$H^{(2)} = \begin{pmatrix} \alpha & 0 & 0 & 0 \\ & \alpha & 0 & 0 \\ & & 2-\alpha & 2(1-\alpha) \\ & & & 2-\alpha \end{pmatrix}.$$

The rest of the pivots in the block H_{12}, (1,4) and (2,3), are zero, so we do not rotate. The only pivot in the block H_{11}, (1,2), is also zero and again no rotation is performed. In order to complete the proof, notice that

$$\begin{pmatrix} 2-\alpha & 2(1-\alpha) \\ & 2-\alpha \end{pmatrix} = (2-\alpha)\begin{pmatrix} 1 & \frac{2(1-\alpha)}{2-\alpha} \\ & 1 \end{pmatrix} = (2-\alpha)\begin{pmatrix} 1 & 1-\frac{\alpha}{2-\alpha} \\ & 1 \end{pmatrix} \bullet$$

Proof of theorem 1 Take O to be an acceptable ordering. Lemma 1 shows that applying the Jacobi method to $H(N,\alpha)$ according to O reduces to applying the Jacobi method to $(2-\alpha)H(\frac{N}{2}, \frac{\alpha}{2-\alpha})$, according to another acceptable ordering O'. Since the Jacobi method, and the smallest singular value of $A = D^{-1}HD^{-1}$, are scale invariant, this implies that

$$\theta(\, H(N,\alpha),\, 0\,) \leq \theta(\, H(\frac{N}{2}, \frac{\alpha}{2-\alpha}),\, O'\,) \tag{2.5}$$

Therefore if α_k is given by the recurrence relation

$$\alpha_0 = \alpha \tag{2.6}$$

$$\alpha_{k+1} = \frac{\alpha_k}{2-\alpha_k} \tag{2.7}$$

we have

$$\theta(\, H(N,a),\, O\,) \leq \theta(\, H(\frac{N}{2^k}, \alpha_k),\, O'\,) \leq \theta(\, H(2, \alpha_{\log N-1}),\, O''\,) = \alpha_{\log N-1}.$$

Fortunately, the recurence relation (2.6) has a simple closed form solution:

$$\alpha_k = \frac{\alpha}{2^k(1-\alpha)+\alpha}.$$

Thus,

$$\theta(\, H(N,\alpha),\, O) = \frac{\alpha}{2^{\log N-1}(1-\alpha)+\alpha} = \frac{\alpha}{\frac{N}{2}(1-\alpha)+\alpha}.$$

Since we are assuming $0 < \alpha < \frac{1}{2}$, this implies that

$$\theta(\, H(N,\alpha),\, O) \leq \frac{4\alpha}{N}.$$

Therefore, since $\sigma(\, H(N,\alpha)\,) = \alpha$,

$$\rho(\, H(N,\alpha),\, O\,) \geq \frac{N}{4},$$

and the proof of Theorem 1 is complete.

3. Appendix

In this appendix we present Matlab code to simulate the operations realized by the Jacobi method when applied to the matrices H and orderings O in theorem 1. Given the special structure of these matrices and orderings, the operations performed by the Jacobi method can be exactly described by the Matlab program

```
a=.5; d=1; ev=[];  %  d = diagonal, ev = eigenvalues.
for i=1:n,         %  N = 2^n is the dimension, to be given.
  ev(i)=d-a;
  d=d+a;
  a=a*sqrt(.5)*sqrt(.5)*4; % or a = 2*a*(1+eps);
end
max(abs((ev-.5)./.5))/(eps*2^(n-1))
```

In this program

$$a = a_k = (1 - \alpha_k) \prod_{i=0}^{k-1} (2 - \alpha_i)$$

is the nonzero off diagonal element of the matrices H computed using the recursion in lemma 1 and d is the diagonal entry of H.

References

[DV] J. Demmel and K. Veselić, *Jacobi's method is more accurate than QR*, Tech. Rep. 486, Comp. Sci. Dept., New York University, October 1989.

[M] W. Mascarenhas, *On the convergence of the Jacobi method for arbitrary orderings*, Phd. disseration, Numerical Analysis Report 91-2, Dept. of Math. M.I.T. Jan, 91., submitted to SIMAX.

Recent IMA Preprints

\# Author/s Title

866 **Gui–Qiang Chen and Tai–Ping Liu**, Zero relaxation and dissipation limits for hyperbolic conservation laws
867 **Gui–Qiang Chen and Jian–Guo Liu**, Convergence of second-order schemes for isentropic gas dynamics
868 **Aleksander M. Simon and Zbigniew J. Grzywna**, On the Larché–Cahn theory for stress-induced diffusion
869 **Jerzy Łuczka, Adam Gadomski and Zbigniew J. Grzywna**, Growth driven by diffusion
870 **Mitchell Luskin and Tsorng-Whay Pan**, Nonplanar shear flows for nonaligning nematic liquid crystals
871 **Mahmoud Affouf**, Unique global solutions of initial-boundary value problems for thermodynamic phase transitions
872 **Richard A. Brualdi and Keith L. Chavey**, Rectangular L-matrices
873 **Xinfu Chen, Avner Friedman and Bei Hu**, The thermistor problem with zero–one conductivity II
874 **Raoul LePage**, Controlling a diffusion toward a large goal and the Kelly principle
875 **Raoul LePage**, Controlling for optimum growth with time dependent returns
876 **Marc Hallin and Madan L. Puri**, Rank tests for time series analysis a survey
877 **V.A. Solonnikov**, Solvability of an evolution problem of thermocapillary convection in an infinite time interval
878 **Horia I. Ene and Bogdan Vernescu**, Viscosity dependent behaviour of viscoelastic porous media
879 **Kaushik Bhattacharya**, Self-accommodation in martensite
880 **D. Lewis, T. Ratiu, J.C. Simo and J.E. Marsden**, The heavy top: a geometric treatment
881 **Leonid V. Kalachev**, Some applications of asymptotic methods in semiconductor device modeling
882 **David C. Dobson**, Phase reconstruction via nonlinear least-squares
883 **Patricio Aviles and Yoshikazu Giga**, Minimal currents, geodesics and relaxation of variational integrals on mappings of bounded variation
884 **Patricio Aviles and Yoshikazu Giga**, Partial regularity of least gradient mappings
885 **Charles R. Johnson and Michael Lundquist**, Operator matrices with chordal inverse patterns
886 **B.J. Bayly**, Infinitely conducting dynamos and other horrible eigenproblems
887 **Charles M. Elliott and Stefan Luckhaus**, 'A generalised diffusion equation for phase separation of a multi-component mixture with interfacial free energy'
888 **Christian Schmeiser and Andreas Unterreiter**, The derivation of analytic device models by asymptotic methods
889 **LeRoy B. Beasley and Norman J. Pullman**, Linear operators that strongly preserve the index of imprimitivity
890 **Jerry Donato**, The Boltzmann equation with lie and cartan
891 **Thomas R. Hoffend Jr., Peter Smereka and Roger J. Anderson**, A method for resolving the laser induced local heating of moving magneto-optical recording media
892 **E.G. Kalnins, Willard Miller, Jr. and Sanchita Mukherjee**, Models of q-algebra representations: the group of plane motions
893 **T.R. Hoffend Jr. and R.K. Kaul**, Relativistic theory of superpotentials for a nonhomogeneous, spatially isotropic medium
894 **Reinhold von Schwerin**, Two metal deposition on a microdisk electrode
895 **Vladimir I. Oliker and Nina N. Uraltseva**, Evolution of nonparametric surfaces with speed depending on curvature, III. Some remarks on mean curvature and anisotropic flows
896 **Wayne Barrett, Charles R. Johnson, Raphael Loewy and Tamir Shalom**, Rank incrementation via diagonal perturbations
898 **Mingxiang Chen, Xu-Yan Chen and Jack K. Hale**, Structural stability for time-periodic one-dimensional parabolic equations
899 **Hong-Ming Yin**, Global solutions of Maxwell's equations in an electromagnetic field with the temperature-dependent electrical conductivity
900 **Robert Grone, Russell Merris and William Watkins**, Laplacian unimodular equivalence of graphs
901 **Miroslav Fiedler**, Structure-ranks of matrices
902 **Miroslav Fiedler**, An estimate for the nonstochastic eigenvalues of doubly stochastic matrices
903 **Miroslav Fiedler**, Remarks on eigenvalues of Hankel matrices
904 **Charles R. Johnson, D.D. Olesky, Michael Tsatsomeros and P. van den Driessche**, Spectra with positive elementary symmetric functions
905 **Pierre-Alain Gremaud**, Thermal contraction as a free boundary problem
906 **K.L. Cooke, Janos Turi and Gregg Turner**, Stabilization of hybrid systems in the presence of feedback delays
907 **Robert P. Gilbert and Yongzhi Xu**, A numerical transmutation approach for underwater sound propagation
908 **LeRoy B. Beasley, Richard A. Brualdi and Bryan L. Shader**, Combinatorial orthogonality
909 **Richard A. Brualdi and Bryan L. Shader**, Strong hall matrices
910 **Håkan Wennerström and David M. Anderson**, Difference versus Gaussian curvature energies; monolayer versus bilayer curvature energies applications to vesicle stability
911 **Shmuel Friedland**, Eigenvalues of almost skew symmetric matrices and tournament matrices
912 **Avner Friedman, Bei Hu and J.L. Velazquez**, A Free Boundary Problem Modeling Loop Dislocations in Crystals
913 **Ezio Venturino**, The Influence of Diseases on Lotka-Volterra Systems
914 **Steve Kirkland and Bryan L. Shader**, On Multipartite Tournament Matrices with Constant Team Size

915 **Richard A. Brualdi and Jennifer J.Q. Massey**, More on Structure-Ranks of Matrices
916 **Douglas B. Meade**, Qualitative Analysis of an Epidemic Model with Directed Dispersion
917 **Kazuo Murota**, Mixed Matrices Irreducibility and Decomposition
918 **Richard A. Brualdi and Jennifer J.Q. Massey**, Some Applications of Elementary Linear Algebra in Combinations
919 **Carl D. Meyer**, Sensitivity of Markov Chains
920 **Hong-Ming Yin**, Weak and Classical Solutions of Some Nonlinear Volterra Integrodifferential Equations
921 **B. Leinkuhler and A. Ruehli**, Exploiting Symmetry and Regularity in Waveform Relaxation Convergence Estimation
922 **Xinfu Chen and Charles M. Elliott**, Asymptotics for a Parabolic Double Obstacle Problem
923 **Yongzhi Xu and Yi Yan**, An Approximate Boundary Integral Method for Acoustic Scattering in Shallow Oceans
924 **Yongzhi Xu and Yi Yan**, Source Localization Processing in Perturbed Waveguides
925 **Kenneth L. Cooke and Janos Turi**, Stability, Instability in Delay Equations Modeling Human Respiration
926 **F. Bethuel, H. Brezis, B.D. Coleman and F. Hélein**, Bifurcation Analysis of Minimizing Harmonic Maps Describing the Equilibrium of Nematic Phases Between Cylinders
927 **Frank W. Elliott, Jr.**, Signed Random Measures: Stochastic Order and Kolmogorov Consistency Conditions
928 **D.A. Gregory, S.J. Kirkland and B.L. Shader**, Pick's Inequality and Tournaments
929 **J.W. Demmel, N.J. Higham and R.S. Schreiber**, Block *LU* Factorization
930 **Victor A. Galaktionov and Juan L. Vazquez**, Regional Blow-Up in a Semilinear Heat Equation with Convergence to a Hamilton-Jacobi Equation
931 **Bryan L. Shader**, Convertible, Nearly Decomposable and Nearly Reducible Matrices
932 **Dianne P. O'Leary**, Iterative Methods for Finding the Stationary Vector for Markov Chains
933 **Nicholas J. Higham**, Perturbation theory and backward error for $AX - XB = C$
934 **Z. Strakos and A. Greenbaum**, Open questions in the convergence analysis of the lanczos process for the real symmetric eigenvalue problem
935 **Zhaojun Bai**, Error analysis of the lanczos algorithm for the nonsymmetric eigenvalue problem
936 **Pierre-Alain Gremaud**, On an elliptic-parabolic problem related to phase transitions in shape memory alloys
937 **Bojan Mohar and Neil Robertson**, Disjoint essential circuits in toroidal maps
938 **Bojan Mohar**, Convex representations of maps on the torous and other flat surfaces
939 **Bojan Mohar and Svatopluk Poljak** Eigenvalues in combinatorial optimization
940 **Richard A. Brualdi, Keith L. Chavey and Bryan L. Shader**, Conditional sign-solvability
941 **Roger Fosdick and Ying Zhang**, The torsion problem for a nonconvex stored energy function
942 **René Ferland and Gaston Giroux**, An unbounded mean-field intensity model: Propagation of the convergence of the empirical laws and compactness of the fluctuations
943 **Wei-Ming Ni and Izumi Takagi**, Spike-layers in semilinear elliptic singular Perturbation Problems
944 **Henk A. Van der Vorst and Gerard G.L. Sleijpen**, The effect of incomplete decomposition preconditioning on the convergence of conjugate gradients
945 **S.P. Hastings and L.A. Peletier**, On the decay of turbulent bursts
946 **Apostolos Hadjidimos and Robert J. Plemmons**, Analysis of p-cyclic iterations for Markov chains
947 **ÅBjörck, H. Park and L. Eldén**, Accurate downdating of least squares solutions
948 **E.G. Kalnins, William Miller, Jr. and G.C. Williams**, Recent advances in the use of separation of variables methods in general relativity
949 **G.W. Stewart**, On the perturbation of LU, Cholesky and QR factorizations
950 **G.W. Stewart**, Gaussian elimination, perturbation theory and Markov chains
951 **G.W. Stewart**, On a new way of solving the linear equations that arise in the method of least squares
952 **G.W. Stewart**, On the early history of the singular value decomposition
953 **G.W. STewart**, On the perturbation of Markov chains with nearly transient states
954 **Umberto Mosco**, Composite media and asymptotic dirichlet forms
955 **Walter F. Mascarenhas**, The structure of the eigenvectors of sparse matrices
956 **Walter F. Mascarenhas**, A note on Jacobi being more accurate than QR
957 **Raymond H. Chan, James G. Nagy and Robert J. Plemmons**, FFT-based preconditioners for Toeplitz-Block least squares problems
958 **Zhaojun Bai**, The CSD, GSVD, their applications and computations
959 **D.A. Gregory, S.J. Kirkland and N.J. Pullman**, A bound on the exponent of a primitive matrix using Boolean rank
960 **Richard A. Brualdi, Shmuel Friedland and Alex Pothen**, Sparse bases, elementary vectors and nonzero minors of compound matrices
961 **J.W. Demmel**, Open problems in numerical linear algebra
962 **James W. Demmel and William Gragg**, On computing accurate singular values and eigenvalues of acyclic matrices
963 **James W. Demmel**, The inherent inaccuracy of implicit tridiagonal QR
964 **J.J.L. Velázquez**, Estimates on the $(N-1)$-dimensional Hausdorff measure of the blow-up set for a semilinear heat equation