ON A CHARACTERIZATION OF THE BEST ℓ_2 SCALING OF A MATRIX

BY

G. H. GOLUB
J. M. VARAH

STAN-CS-72-319
OCTOBER 1972

COMPUTER SCIENCE DEPARTMENT
School of Humanities and Sciences
STANFORD UNIVERSITY

On a Characterization of the Best ℓ_2 Scaling of a Matrix

G.H. Golub
Computer Science Department
Stanford University
Stanford, California, USA

and

J.M. Varah
Computer Science Department
University of British Columbia
Vancouver, B.C., Canada

Issued jointly as a Technical Report from Stanford University Computer Science Department and The University of British Columbia Computer Science Department. The work of the first author was supported by NSF and AEC grants; that of the second was supported by **NRC(Canada)** grant **#A8240.**

. Abstract

This paper is concerned with best two-sided scaling of a general square matrix, and in particular with a certain characterization of that best scaling: namely that the first and last singular vectors (on left and right) of the scaled matrix have components of equal modulus. Necessity, sufficiency, and its relation with other characterizations are discussed. Then the problem of best scaling for rectangular matrices is introduced and a conjecture made regarding a possible best scaling. The conjecture is verified for some special cases.

1. Introduction

Let A be an n x n nonsingular matrix.+ We are interested in the best row and column scaling of A in the ℓ_2 norm; that is

$$\min_{D,E \text{ diag}} (||DAE||_2 ||(DAE)^{-1}||_2).$$

Of course this is equivalent to

$$\min_{D,E \text{ diag}} (\sigma_1(DAE) / \sigma_n(DAE))$$

where $\sigma_1(A) \geq \sigma_2(A) \geq \ldots \geq o,(A) > 0$ are the singular values of A. In this paper we will discuss the following useful characterization of this best two-sided scaling: let $A = U \sum V$ be the singular value decomposition of A. Then A is best scaled in the ℓ_2 norm if $|u_i^{(1)}| = |u_i^{(n)}|, |v_i^{(1)}| = |v_i^{(n)}|$, for $i=1,\ldots,n$. That is, A is best scaled if the first and last columns of U and V have components of equal magnitude. We refer to this as the EMC property.

This characterization has had an interesting history: it was to our knowledge first discussed by Forsythe and Straus [3] in connection with one-sided scaling, or equivalently best symmetric scaling (DAD) of a positive definite matrix A. (For one-sided scaling, only one of U,V is involved in the EMC property.) They showed sufficiency of EMC for best one-sided scaling. It was also mentioned by Bauer [1] for one-sided scaling; he also gave an explicit representation of the best ℓ_2 scaling for matrices A with A and A^{-1} having a checkerboard sign pattern. More recently, McCarthy and Strang [4] have settled the question of necessity for one-sided scaling: for matrices A which when best scaled have σ_1 and σ_n distinct, the EMC

property must hold; however this is not always true if σ_1 or σ_n is multiple even using the inherent ambiguity in the singular vectors, and they give counterexamples.

The EMC property for two-sided scaling was first brought *to* our attention by C.L. Lawson (see also [6, pg 44]) in connection with the matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & -1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \qquad \sigma_1/\sigma_n \simeq 27.4$$

We found the best ℓ_2 scaling by minimizing $\sigma_1(DAE)/\sigma_n(DAE)$ as a function of D,E using a function minimizing procedure. This gave $D = \mathrm{diag}(1,\sqrt{3},3)$, $E = \mathrm{diag}(1,1/2,1/\sqrt{6})$, $\sigma_1/\sigma_n \simeq 13.9$,

$$DAE = \begin{pmatrix} 1 & 1 & \sqrt{6}/2 \\ \sqrt{3} & -\sqrt{3}/2 & 1/\sqrt{2} \\ 0 & 3/2 & \sqrt{6}/2 \end{pmatrix}$$

In this paper, we discuss the EMC property for best two-sided scaling and how it is related to the Bauer representation for checkerboard matrices. Then we discuss the problem of best scaling for a rectangular matrix.

We end this introduction with a warning: although these best scalings are attractive and theoretically interesting, it may be quite improper to scale a particular problem this way; this can cause inaccurate data and unimportant variables to assume too much influence. Such is the case for example in solving ill-posed problems using the singular value decomposition (see [5]). Normally several of the equations are ignored and a reasonable solution is constructed solving the remaining ones; however "best" scaling can cause the whole matrix to become quite well-conditioned, with its (well-determined) solution bearing no relation to the solution of the original problem.

2. Aspects of the EMC Property

First we show the sufficiency of the EMC property for best two-sided scaling. The proof is a slight extension of Forsythe and Straus [3].

Theorem 2.1:

Let A be an n x n nonsingular matrix. Then A is best scaled in the ℓ_2 norm with respect to diagonal scalings DAE if the EMC property holds.

Proof:

We have for any nonsingular diagonal D,E,

$$\mathrm{cond}_2(DAE) = \sigma_1/\sigma_n = \max_{p,q,r,s} \left(\frac{\dfrac{|p^T DAEq|}{\|p\|_2\|q\|_2}}{\dfrac{|r^T DAEs|}{\|r\|_2\|s\|_2}} \right)$$

$$= \max_{p,q,r,s} \left[\frac{\dfrac{|p^T Aq|}{\|D^{-1}p\|_2\|E^{-1}q\|_2}}{\dfrac{|r^T As|}{\|D^{-1}r\|_2\|E^{-1}s\|_2}} \right] \geq \mathrm{cond}_2(A) \left[\frac{(u^{(n)T}D^{-2}u^{(n)})(v^{(n)T}E^{-2}v^{(n)})}{(u^{(1)T}D^{-2}u^{(1)})(v^{(1)T}E^{-2}v^{(1)})} \right]^{1/2}$$

where $u^{(1)}, u^{(n)}, v^{(1)}, v^{(n)}$ are the appropriate singular vectors of A. Now if $|(u^{(n)})_i| = |(u^{(1)})_i|$ and $|(v^{(n)})_i| = |(v^{(1)})_i|$ for $i = 1,\ldots,n$, i.e. if the EMC property holds, the term in square brackets is 1 and $\mathrm{cond}_2(A) \leq \mathrm{cond}_2(DAE)$ for all D,E. QED

For the EMC property to be also necessary for best two-sided ℓ_2 scaling, we must show the existence of a D,E with DAE having the EMC property. However as we mentioned earlier, McCarthy and Strang [4] gave examples of one-sided best scaled matrices for which the corresponding one-sided EMC property failed

to hold. These examples hoever had multiple σ_1 or σ_n in best scaled form; for matrices with distinct σ_1 and σ_n in best scaled form, they showed that EMC was attainable. From this we easily obtain:

Corollary 2.2:

Let A have distinct σ_1 and σ_n in best scaled form; then the EMC property is necessary and sufficient for best two-sided ℓ_2 scaling,

Thus the existence of an EMC scaling is assured with this restriction of distinct extreme singular values. Of course it need not be unique: for example if A has a special symmetry so that PAQ=A for P, Q permutation matrices, then if DAE is best scaled, so is $(PDP^T)A(Q^TEQ)$. . (This is P(DAE)Q with singular value decomposition $(PU)\sum(V^TQ)$ and this has EMC if $U \sum V^T$ does.)

Now we discuss the relation between EMC and Bauer's characterization for best ℓ_2 scaling of a real irreducible checkerboard matrix A. We must also assume, although it normally follows from the irreducibility of A, that $|A|\,|A^{-1}|$, $|A^{-1}|\,|A|$, $|A|\,|A^T|$, $|A^T|\,|A|$ are irreducible. Recall the Bauer characterization (see [1]): if A, A^{-1} have checkerboard sign patterns, that is if there exist diagonal orthogonal matrices, J_1, J_2, J_3, J_4 so that $J_1AJ_2 = |A| \geq 0$ and $J_3A^{-1}J_4 = |A^{-1}| \geq 0$, and if we let $y^{(1)}, x^{(1)}$ be the left and right Perron eigenvectors of $|A|\,|A^{-1}|$ (and similarly $y^{(2)}, x^{(2)}$ for $|A^{-1}|\,|A|$), then the best ℓ_2 scaling DAE for A is given by $d_i^2 = y_i^{(1)} / x_i^{(1)}$,

$e_i^2 = x_i^{(2)}/y_i^{(2)}$. (Because of the irreducibility, the Perron vectors have positive components.) Thus A is best scaled if the left and right Perron vectors of $|A|\,|A^{-1}|$ and $|A^{-1}|\,|A|$ are equal. But such a matrix A satisfies the conditions of Corollary 2.2, so the above must be equivalent to the EMC property. We expand on this as follows:

Theorem 2.3:

Let A be a real irreducible matrix with a checkerboard sign pattern. Suppose $|A| = J_1 A J_2$, $|A^{-1}| = J_3 A^{-1} J_4$ and let $A = U \Sigma V^T$ be its singular value decomposition.

(i) Suppose the EMC property holds. Then $|u^{(1)}|$ is the left and right Perron vector of $|A|\,|A^{-1}|$, and $|v^{(1)}|$ is the left and right Perron vector of $|A^{-1}|\,|A|$.

(ii) Suppose the left and right Perron vectors of $|A|\,|A^{-1}|$ are equal (call it u), and similarly for $|A^{-1}|\,|A|$ (call it v). Then $u^{(1)} = J_1 u$, $u^{(n)} = J_4 u$, $v^{(1)} = J_2 v$, $v^{(n)} = J_3 v$.

Proof:

(i) We have $|A| = J_1 A J_2 = (J_1 U)\,\Sigma\,(V^T J_2)$, and this must be the singular value decomposition for $|A|$. Hence $J_1 u^{(1)} > 0$, $J_2 v^{(1)} > 0$ (positive because of the irreducibility of A). Similarly $|A^{-1}| = J_3 A^{-1} J_4 = (J_3 V)\Sigma^{-1}(U^T J_4)$ and we must have $J_3 v^{(n)} > 0$, $J_4 u^{(n)} > 0$. Now the EMC property and orthogonality of the $\{u^{(i)}\}, \{v^{(i)}\}$ gives

$$
\begin{aligned}
J_4 u^{(n)} &= J_1 u^{(1)} & \qquad J_3 v^{(n)} &= J_2 v^{(1)} \\
J_1 u^{(n)} &= J_4 u^{(1)} & \qquad J_2 v^{(n)} &= J_3 v^{(1)} \\
J_1 u^{(1)} &\perp J_4 u^{(1)} & \qquad J_2 v^{(1)} &\perp J_3 v^{(1)}
\end{aligned}
\qquad (1)
$$

Now $|A|\,|A^{-1}| = J_1\, U\, \Sigma\, (V^T J_2 J_3 V)\, \Sigma^{-1}\, U^T J_4 = J_1\, U\, (\Sigma\, Q\, \Sigma^{-1})\, U^T J_4$.

Consider Q; it is orthogonal and symmetric, and from (1) we see that

$Q_{1n}=Q_{n1}=1$ and the rest of the first and last rows and columns of Q are zero.

Thus $$\sum Q \sum^{-1} = \begin{pmatrix} 0 & \cdots & 0 & \sigma_1/\sigma_n \\ \vdots & & & 0 \\ \vdots & & B & \vdots \\ 0 & & & \vdots \\ \sigma_n/\sigma_1 & 0 & \cdots & 0 \end{pmatrix}$$

Thus

$$|A||A^{-1}|(J_1 u^{(1)}) = J_1 U (\sum Q \sum^{-1}) U^T J_4 (J_4 u^{(n)})$$

$$= J_1 U (\sum Q \sum^{-1}) e_n$$

$$= J_1 U (\frac{\sigma_1}{\sigma_n}) e_1 = \frac{\sigma_1}{\sigma_n} (J_1 u^{(1)}).$$

So $J_1 u^{(1)} = |u^{(1)}|$ is the unique positive right Perron eigenvector of $|A||A^{-1}|$ corresponding to the eigenvalue σ_1/a_n. A similar computation shows it is also the left Perron vector. Likewise, $J_2 v^{(1)} = |v^{(1)}|$ can be shown to be the left and right Perron vector for $|A^{-1}||A|$.

(ii) If the hypothesis of (ii) holds, then from Bauer [1] we have that the spectral radius of $|A||A^{-1}|$ and $|A^{-1}||A|$ is σ_1/σ_n. Thus $|A||A^{-1}| u = \frac{\sigma_1}{\sigma_n} u$, which gives

$$\sigma_n J_3 A^{-1} J_4 u = \sigma_1 J_2 A^{-1} J_1 u.$$

Now let $J_4 u = \sum_1^n \alpha_i u^{(i)}$, $J_1 u = \sum_1^n \beta_i u^{(i)}$, with $\sum_1^n \alpha_i^2 = \sum_1^n \beta_i^2 = 1$. Then the above can be written

$$\sigma_n \sum_1^n \frac{\alpha_i}{\sigma_i} v^{(i)} = (J_3 J_2) \sigma_1 \sum_1^n \frac{\beta_i}{\sigma_i} v^{(i)}$$

Now take ℓ_2 norms:

$$1 \geq \sum_i \left(\frac{\sigma_n \alpha_i}{\sigma_i}\right)^2 = \sum_i \left(\frac{\sigma_1 \beta_i}{\sigma_i}\right)^2 \geq 1$$

and equality must hold, implying that $\alpha_n = 1$, $\beta_1 = 1$ with the other components zero, giving $J_4 u = u^{(n)}$, $J_T u = u^{(1)}$. By a similar argument, one can show $J_2 v = v^{(1)}$, $J_3 v = v^{(n)}$. QED.

We should also remark that the equivalence of these two characterizations can be used to check the accuracy of A^{-1}, when it is known that both A and A^{-1} have checkerboard sign patterns. For a given A and computed A^{-1}, one can compute the best scaling via the Perron vectors of $|A|\,|A^{-1}|$ and $|A^{-1}|\,|A|$; then one can test--the EMC criterion on the singular vectors of the scaled matrix.

3. Best Scaling for Rectangular Matrices

Let A be m x n with m > n and rank n. Then we can still ask for the best scaled DAE in the sense of minimizing $\sigma_1/\sigma_n(DAE)$. It is clear that for the best scaling on the right, the EMC property on V is still sufficient, since $A^T A$ is still a nonsingular n x n matrix and the Forsythe-Straus argument still holds. However this is not the case for scaling on the left, since in particular we could take any n linearly independent rows of A and best scale the resulting n x n matrix; this will then have the EMC property (assuming σ_1 and σ_n are distinct) but will not necessarily give the best scaling for A. There are in fact $\binom{m}{n}$ such choices of n x n submatrices, so a leading contender for the best scaled A would be that n x n submatrix which when best scaled gives the minimal condition. This leads to the intriguing

Conjecture: There exists an n x n submatrix of A which, when best scaled, gives the best scaling for A also.

It would be better to say one of the best scalings because it is not necessarily unique. We cannot prove this in general, only in some special cases which we discuss below. We have also verified it numerically on several examples.

Case I: $A = \begin{pmatrix} B \\ FB \end{pmatrix}$ where B is n x n, nonsingular, and F^TF is diagonal.

Then $A^TA = B^TB + B^TF^TFB$

$$= B^T(I + F^TF)B$$

$$= B^TG^TGB$$

so the nonzero singular values of A and GB are the same. Now if F^TF is diagonal, G is diagonal, and thus the best scaling for A occurs when GB (or B) is best scaled. So one best scaling for A is $DAE = \begin{pmatrix} D_1BE \\ 0 \end{pmatrix}$ where D_1BE is best scaled. However this is not necessarily unique: let B be best scaled, and consider

$$DAE = \begin{pmatrix} D_1BE \\ D_2FBE \end{pmatrix}.$$

Then

$$(DAE)^T(DAE) = EB^T(D_1^2 + F^TD_2^2F)BE = (GBE)^T(GBE).$$

Now if F is such that $F^TD_2^2F$ is diagonal for all D_2 diagonal (e.g. if F has at most one nonzero element in each row and column), then G is also diagonal for all choices of Dl and D_2 and the best scaling of A occurs for E = I and any D_1, D_2 such that G = I (since B Is best scaled). That is, we must have

$$D_1^2 + F^TD_2^2F = I.$$

Of course this will occur for Dl = I, $D_2 = 0$, but there can be many other solutions.

Note also that if B is an orthogonal matrix, a best scaling is certainly obtained with Dl = I, $D_2 = 0$, no matter what F is.

Case II: $n = 2$

We have $A = \begin{pmatrix} u_1 & v_1 \\ \vdots & \vdots \\ u_n & v_n \end{pmatrix}$, $D = \mathrm{diag}(d_1,\ldots,d_n)$, $E = \mathrm{diag}(e_1,e_2)$,

and we seek $\min_{D,E} \mathrm{cond}_2(B=DAE) = \min_{D,E} \left(\frac{\lambda_1(B^TB)}{\lambda_2(B^TB)}\right)^{1/2} = \sqrt{g(D,E)}$

Let $B^TB = \begin{pmatrix} p & r \\ r & s \end{pmatrix} = \begin{pmatrix} e_1^2\Sigma d_i^2 u_i^2 & e_1e_2\Sigma d_i^2 u_i v_i \\ e_1e_2\Sigma d_i^2 u_i v_i & e_2^2\Sigma d_i^2 v_i^2 \end{pmatrix}$

Then $g(D,E) = \frac{1+\sqrt{f(D,E)}}{1-\sqrt{f(D,E)}}$ where $f(D,E) = \frac{(p-s)^2+4r^2}{(p+s)^2}$

. Since g is a monotone function of f, we need only find $\min f(D,E)$. As a function of $e = e_2/e_1$, we can write

$$f(D,E) = \frac{(\alpha-\gamma e^2)^2 + 4e^2\beta^2}{(\alpha+\gamma e^2)^2}$$

where a, β, γ are constants. This is minimized as a function of e for $e^2 = \gamma/\alpha$, making $p = s$ and thus $B^TB = \begin{pmatrix} p & r \\ r & p \end{pmatrix}$ which has eigenvector matrix $\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, possessing the EMC property.

With this E,

$$f(D) = \frac{r^2}{p^2} = \frac{(\Sigma d_i^2 u_i v_i)^2}{(\Sigma d_i^2 u_i^2)(\Sigma d_i^2 v_i^2)} = \cos^2\theta(Du,Dv).$$

To minimize this, we need to examine three cases.

(i) some u_i or $v_i = 0$. Suppose $u_i = 0$, $v_i \neq 0$. Taking $d_i \to \infty$ gives $f(D) = 0$ for any choice of the other d_j. If $u_i = v_i = 0$, the problem

reduces to one of lower dimension. So assume all $u_i, v_i \neq 0$.

(ii) $\{u_i\}, \{v_i\}$ not all of the same sign. Suppose $u1 > 0$, $u_2 > 0$, $v1 > 0$, $v_2 < 0$ for example. Then we can make $(Du) \perp (Dv)$ and $f(D) = 0$ by choosing

$$d1 = \frac{1}{\sqrt{u_1 v_1}}, \quad d_2 = \frac{1}{\sqrt{-u_2 v_2}}, \quad d_i = 0, \ i \neq 1,2.$$

If $r = u_1/v_1$, $R = -u_2/v_2$, this gives $e^2 = rR$ and

$$\text{best } B = DAE = \begin{pmatrix} \sqrt{r} & \sqrt{R} \\ \sqrt{R} & -\sqrt{r} \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{pmatrix}$$

its eigenvector matrix with the EMC property.

(iii) $u_i > 0, v_i > 0$ for all i. Then from a result of Cassels (see Beckenbach and Bellman [2, p. 45]), we have

$$\min_D f(D) = \frac{4rR}{(r+R)^2} = \frac{4}{2 + \frac{r}{R} + \frac{R}{r}}$$

where $r = \min_i u_i/v_i = u_m/v_m$ (say) and $R = \max_i u_i/v_i = u_M/v_M$. The corresponding D has $d_m = \frac{1}{\sqrt{u_m v_m}}$, $d_M = \frac{1}{\sqrt{u_M v_M}}$, $d_i = 0$, $i \neq m, M$. This gives $e^2 = rR$ and

$$\text{best } B = DAE = \begin{pmatrix} 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ \sqrt{r} & \sqrt{R} \\ \sqrt{R} & -\sqrt{r} \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{pmatrix} \begin{matrix} \\ \\ \\ m^{th} \text{ row} \\ \\ M^{th} \text{ row} \\ \\ \\ \\ \end{matrix}$$

Again BB^T has its eigenvector matrix with the EMC property.

Finally, one might think that for rectangular matrices with a checkerboard sign pattern, the best scaling could be achieved using Bauer's algorithm with A and A^ψ, the pseudo-inverse. We give the following counterexample:

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 4 & 1 \end{pmatrix} .$$

Best scaling: $D = \mathrm{diag}(1,0,1/2)$, $E = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$. Then $B = DAE = \begin{pmatrix} 1 & 2 \\ 0 & 0 \\ 2 & 1 \end{pmatrix}$, with $\mathrm{cond}_2(DAE) = 3$.

Now $B^\psi = \begin{pmatrix} -1/3 & 0 & 2/3 \\ 2/3 & 0 & -1/3 \end{pmatrix}$ and

$$|B^\psi||B| = \begin{pmatrix} 5/3 & 4/3 \\ 4/3 & 5/3 \end{pmatrix} , \quad |B|\,|B^\psi| = \begin{pmatrix} 5/3 & 0 & 4/3 \\ 0 & 0 & 0 \\ 4/3 & 0 & 5/3 \end{pmatrix}$$

Both of these are symmetric so both have equal left and right Perron vectors. Thus the Bauer ℓ_2 scaling leaves B unchanged, if we call $0/0 = 0$ (notice $|B|\,|B^\psi|$ is reducible). However if we try to derive B from A using Bauer's algorithm, it fails:

$$A^\psi = \frac{1}{14} \begin{pmatrix} -4 & -1 & 5 \\ 14 & 7 & -7 \end{pmatrix} , \quad |A^\psi||A| = \begin{pmatrix} 13/7 & 5/7 \\ 4 & 2 \end{pmatrix} ,$$

and this has spectral radius $= \rho \simeq 3.62 > 3 = \mathrm{cond}_2(B)$. Moreover the left and right Perron vectors of $|A^\psi||A|$ are

$$\begin{pmatrix} \rho - 2 \\ 5/7 \end{pmatrix} , \quad \begin{pmatrix} 5/7 \\ \rho - 13/7 \end{pmatrix}$$

giving a right-hand scaling matrix $E \simeq \begin{pmatrix} 1 & 0 \\ 0 & 2.4 \end{pmatrix}$, not optimal.

We might also remark that if the conjecture is valid for arbitrary m x n matrices, it would indicate the folly of trying to best scale a rectangular matrix arising from a least squares problem for example; only n of the observations would be retained!

References

1. F.L. Bauer, Optimally scaled matrices. Num. Math. 5 (1963), 73-87.
2. Beckenbach and Bellman, Inequalities. Springer-Verlag, Berlin, 1965.
3. G.E. Forsythe and E.G. Straus, On best conditioned matrices. Proc. AMS 6 (1955), 340-345.
4. C. McCarthy and G. Strang, Optimal conditioning of matrices. Siam. J. Num. Anal. (to appear).
5. J.M. Varah, On the numerical solution of ill-conditioned linear systems with applications to ill-posed problems. Siam. J. Num. Anal. (to appear).
6. G.E. Forsythe and C.B. Moler, Computer Solution of Linear Algebraic Systems. Prentice Hall, New York, 1967.