Available online at www.sciencedirect.com
SCIENCE DIRECT®
LINEAR ALGEBRA AND ITS APPLICATIONS
ELSEVIER Linear Algebra and its Applications 370 (2003) 185–191
www.elsevier.com/locate/laa

Scaling symmetric positive definite matrices to prescribed row sums

Dianne P. O'Leary

Department of Computer Science, University of Maryland, College Park, MD 20742, USA

Received 7 September 2001; accepted 7 January 2003

Submitted by H. Schneider

Abstract

We give a constructive proof of a theorem of Marshall and Olkin that any real symmetric positive definite matrix can be symmetrically scaled by a positive diagonal matrix to have arbitrary positive row sums. We give a slight extension of the result, showing that given a sign pattern, there is a unique diagonal scaling with that sign pattern, and we give upper and lower bounds on the entries of the scaling matrix.
© 2003 Elsevier Inc. All rights reserved.

Keywords: Positive definite matrices; Matrix scaling; Diagonal preconditioning; Homotopy

1. Introduction

We consider the following problem: Given a matrix $W \in R^{n\times n}$ and an n-vector $u > 0$, find a diagonal matrix X so that the scaled matrix XWX has row-sums equal to the elements of u. In other words, given W and u, solve the nonlinear equation

$$XWXe = u,$$

where e is an n-vector with all entries equal to one.

Scaling problems have been a topic of intense investigation. Brualdi [1] gave necessary and sufficient conditions for the existence of such a diagonal scaling when W is symmetric with nonnegative elements. Other authors have considered scalings of nonsymmetric matrices, allowing different diagonal matrices on the left and the right; see, for example, [2,12,13] and the references therein. The inverse problem of finding matrices of given sign patterns with given row and column sums has also been investigated, for example, in [3,5].

E-mail address: oleary@cs.umd.edu (D.P. O'Leary).

0024-3795/03/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/S0024-3795(03)00387-2

In this paper, we consider the problem when W is real symmetric and positive definite. Marshall and Olkin [10] proved that a positive scaling matrix exists when a symmetric matrix W is strictly copositive, and, as a special case, when W is positive definite. As noted by a referee, Kalantari investigated the positive semidefinite case from the point of view of a theorem of the alternative in [7] and via a separation theorem in [6, Sec. 11.3]. Algorithms for computing the scaling in polynomial time can be obtained by noting that our problem can be written as a *second-order cone program* (SOCP) [11]

$$\min_{x,y} \|Wx - y\|$$
$$u_i = x_i y_i, \quad i = 1, \dots, n,$$
$$x \geqslant 0,$$

where x is the diagonal of X and $\|\cdot\|$ denotes the 2-norm (for vectors or matrices). See [9] for techniques to convert this problem to standard SOCP form and for discussion of the polynomial time complexity.

In this paper, we give an alternate constructive existence proof of Marshall and Olkin's result. We generalize the result slightly, showing that there are 2^n solutions to the nonlinear equation $XWXe = u$, one for each sign pattern for X, and we give bounds on the size of the entries of the matrix X.

2. A constructive existence proof

Suppose we are given a symmetric positive definite matrix $W \in R^{n\times n}$ and an n-vector $u > 0$. We want to show that there exists a positive diagonal matrix X so that the scaled matrix XWX has row-sums equal to the elements of u.

We will prove this result by considering the matrix

$$V(t) = (1-t)I + tW.$$

Then $V(0) = I$, and $V(1) = W$.

We will study the mapping

$$H(t,x) = X(t)V(t)X(t)e - u = 0,$$

where $X(t)$ is a positive diagonal matrix with entries x_i. For $t = 0$, we have a unique positive solution $x(0) = \hat{x}$ with $\hat{x}_i = \sqrt{u_i}$.

If we can find a positive solution vector x for $t = 1$, then the solution to our scaling problem is the corresponding matrix $X(1)$.

Differentiating our mapping, we obtain

$$\partial_x H(t,x)x'(t) + \partial_t H(t,x) = 0, \quad x(0) = \hat{x}, \tag{1}$$

where

$$\partial_x H(t,x) = X(t)V(t) + \mathrm{diag}(V(t)X(t)e),$$
$$\partial_t H(t,x) = X(t)(W - I)X(t)e.$$

The matrix $V(t)$ is positive definite on some interval $(-\sigma, \tau)$ where $\sigma > 0$ and $\tau > 1$. Let $2\epsilon = \min(\sigma, \tau - 1)$. Then $V(t)$ is uniformly positive definite on the interval $(-\epsilon, 1 + \epsilon)$, with eigenvalues $(1 - t) + (t$ times the eigenvalues of $W)$, and we define the bounds on its eigenvalues to be $\lambda_{\min} > 0$ and $\lambda_{\max} < \infty$.

The proof of our theorem relies on three lemmas, one establishing the boundedness of $X(t)$, one showing Lipschitz continuity of $f(t, x) = \partial_x H(t, x)^{-1} \partial_t H(t, x)$, and one rather standard result concerning existence of solutions to initial value problems.

Lemma 1. *There exist scalars* $\xi_\ell > 0$ *and* $\xi_u < \infty$, *independent of* t, *such that if* $x(t) > 0$ *satisfies* (1) *for some value of* $t \in [-\epsilon, 1 + \epsilon]$, *then*

$$\xi_\ell \leqslant \min_i x_i(t) \leqslant \max_i x_i(t) \leqslant \xi_u.$$

Proof. The matrix $X(t)$ satisfies $X(t)V(t)X(t)e - u = 0$, so

$$e^{\mathrm{T}} X(t)V(t)X(t)e = e^{\mathrm{T}} u > 0.$$

Since $Xe = x$, we know that

$$e^{\mathrm{T}} u \geqslant \lambda_{\min}(V)\|x\|^2 \geqslant \lambda_{\min} x_i^2, \quad i = 1, \ldots, n,$$

so

$$x_i^2 \leqslant \frac{e^{\mathrm{T}} u}{\lambda_{\min}} \equiv \xi_u^2.$$

This means that the elements of $X(t)$ are uniformly bounded above for $t \in [-\epsilon, 1 + \epsilon]$.

Now since

$$x_i(t)(V(t)X(t)e)_i = u_i, \quad i = 1, \ldots, n,$$

we have

$$x_i(t) = \frac{u_i}{(V(t)X(t)e)_i} \geqslant \frac{u_i}{\|V(t)\|\|X(t)\|\sqrt{n}},$$

so we can define

$$\xi_\ell = \frac{\min_i u_i}{\lambda_{\max} \xi_u \sqrt{n}}. \qquad \square$$

Lemma 2. *Let*

$$\Omega = \left\{(t, x) : -\epsilon < t < 1 + \epsilon, \tfrac{1}{2}\xi_\ell e < x(t) < 2\xi_u e, V(t)x(t) > 0\right\}.$$

For a fixed value of t, *the function* $f(t, x)$ *is Lipschitz continuous on* Ω, *where* f *is defined by*

$$\partial_x H(t, x) f(x) = -\partial_t H(t, x). \tag{2}$$

Proof. The matrix $\partial_x H(t, x)X(t)$ is symmetric and positive definite on Ω, so the inverse of $\partial_x H(t, x)$ must exist, and it is a continuous function of x and t. The right-hand side $-X(t)(W - I)X(t)e$ is continuous on Ω. Therefore, $f(x)$ is continuous.

Now, for a fixed $t \in [-\epsilon, 1+\epsilon]$, we show that $f(t,x)$ satisfies a Lipschitz condition in x.

Let (t,x) and $(t,\hat{x})$ be two points in Ω. Let $Y = X(W-I)Xe$ and $Z = XVX + \operatorname{diag}(XVx)$, and define $\hat{Y}$ and $\hat{Z}$ by substituting $\hat{X}$ for X in these expressions. Then we have these bounds:

$$\begin{aligned}
&\|\hat{X}\|, \|X\| \leqslant \xi_u,\\
&\|\hat{Y}\|, \|Y\| \ \leqslant \sqrt{n}\|W-I\|\xi_u^2,\\
&\|\hat{Z}^{-1}\|, \|Z^{-1}\| \ \leqslant \frac{1}{\xi_\ell^2\lambda_{\min}}.
\end{aligned}$$

We compute

$$\begin{aligned}
\|f(t,\hat{x}) - f(t,x)\| &= \|\hat{X}\hat{Z}^{-1}\hat{Y} - XZ^{-1}Y\|\\
&= \|(\hat{X}-X)\hat{Z}^{-1}\hat{Y} + X\hat{Z}^{-1}(\hat{Y}-Y) + X(\hat{Z}^{-1}-Z^{-1})Y\|\\
&\leqslant \|(\hat{X}-X)\|\,\|\hat{Z}^{-1}\|\,\|\hat{Y}\| + \|X\|\,\|\hat{Z}^{-1}\|\|\hat{Y}-Y\|\\
&\quad +\|X\|\|\hat{Z}^{-1}-Z^{-1}\|\|Y\|.
\end{aligned}$$

We already have bounds on many of these norms, so to conclude that f is Lipschitz continuous, it suffices to bound $\|\hat{Y}-Y\|$ and $\|\hat{Z}^{-1}-Z^{-1}\|$ in terms of $\|\hat{X}-X\|$, since $\|\hat{X}-X\| \leqslant \|\hat{x}-x\|$.

We compute the Y bound by noting that

$$\begin{aligned}
\hat{Y} - Y &= \hat{X}(W-I)\hat{X}e - X(W-I)Xe\\
&= (\hat{X}-X)(W-I)\hat{X}e + X(W-I)(\hat{X}-X)e,
\end{aligned}$$

so

$$\|\hat{Y}-Y\| \leqslant 2\|W-I\|\xi_u\sqrt{n}\|\hat{X}-X\|.$$

Now we bound the Z term. Let $D = \operatorname{diag}(XVx)$, and similarly for $\hat{D}$, and note that

$$\begin{aligned}
\hat{Z}^{-1} - Z^{-1} &= (\hat{X}V\hat{X}+\hat{D})^{-1} - (XVX+D)^{-1}\\
&= (\hat{X}V\hat{X}+\hat{D})^{-1}[-\hat{X}V(\hat{X}-X) + (\hat{X}-X)VX - \hat{D} + D]\\
&\quad \times(XVX+D)^{-1}.
\end{aligned}$$

The norms of the first and last factors are bounded, so we just need to bound the norm of the middle expression:

$$\begin{aligned}
&\|-\hat{X}V(\hat{X}-X) + (\hat{X}-X)VX - \hat{D} + D\|\\
&\quad \leqslant 2\xi_u\lambda_{\max}\|\hat{X}-X\| + \|\hat{D}-D\|.
\end{aligned}$$

Focusing on the last term gives

$$(\hat{D} - D)_i = \hat{x}_i \sum_j w_{ij}\hat{x}_j - x_i \sum_j w_{ij} x_j$$
$$= (\hat{x}_i - x_i) \sum_j w_{ij}\hat{x}_j + x_i \sum_j w_{ij}(\hat{x}_j - x_j),$$

so

$$|(\hat{D} - D)_i| \leqslant \lambda_{\max}\xi_u|\hat{x}_i - x_i| + \xi_u\lambda_{\max}\|\hat{x}_j - x_j\|$$

and thus we have a bound on every term in terms of $\|\hat{x} - x\|$, yielding a conclusion of Lipschitz continuity for f. □

Lemma 3. *Let Ω be a bounded domain in R^{n+1} with $(0, x_0) \in \Omega$. If f is continuous in Ω and locally satisfies a Lipschitz condition in the x variables, then there exists a solution of the initial value problem*

$$x'(t) = f(t, x), \quad x(0) = x_0$$

that can be uniquely extended arbitrarily close to the boundary of Ω.

Proof. See, for example, Hurewicz [4, Theorem 11]. □

Now we use our three lemmas to prove that the scaling matrix exists.

Theorem 1. *Given a symmetric positive definite matrix $W \in R^{n\times n}$ and an n-vector $u > 0$, there exists a positive diagonal matrix X so that the scaled matrix XWX has row-sums equal to the elements of u.*

Proof. To construct our scaling X, we use Lemma 3 to show that (1) has a solution at $t = 1$.

It is clear that $(0, x_0) \in \Omega$, and Lemma 2 assures us that the function f defined by (2) is Lipschitz continuous on Ω. Thus, the assumptions of Lemma 3 are satisfied, so a solution to (1) can be extended to the boundary of Ω.

Now, consider any solution point $(t, x(t))$ for $t \in [-\epsilon, 1 + \epsilon]$ with $x > 0$. By Lemma 1, $\xi_\ell e \leqslant x \leqslant \xi_u e$, and thus, since $XV(t)x = u > 0$, we must have

$$V(t)x \geqslant \frac{1}{\xi_u} u > 0.$$

Therefore, any solution point $(t, x(t))$ with $t \in [-\epsilon, 1 + \epsilon]$ has x bounded away from the constraints

$$\frac{1}{2}\xi_\ell e < x(t) < 2\xi_u e,\ V(t)x(t) > 0$$

that define Ω. Therefore, we must be able to extend the solution from $t = 0$ to the boundary $t = 1 + \epsilon$, and thus the solution exists for $t = 1$. □

By replacing $V(t)$ by the positive definite matrix $EV(t)E$, where E is a diagonal matrix with entries ± 1, we can see that there are actually 2^n scaling matrices, one for each orthant, that give the prescribed row sums. For $t = 1$, the equation $XVXe = u$ is a polynomial system of degree 2^n, so this accounts for all possible solutions.

Corollary 1. *The equation $XWXe = u$, with W symmetric positive definite and X a diagonal matrix, has 2^n solutions, one per orthant, so we can scale the matrix W by a diagonal matrix with arbitrary signs, so that it has prescribed row sums.*

3. Bounds on the entries in the scaling matrix

Khachiyan and Kalantari [8] gave upper and lower bounds on the Frobenius norm of the scaling matrix X in the special case when $u = e$ and $X > 0$. We generalize their result to an arbitrary positive vector u and an arbitrary orthant, deriving an alternate lower bound expression.

Theorem 2. *The entries in the scaling X are bounded as*

$$\frac{u^{\mathrm{T}}e}{\hat{\mu}} \leqslant \|x\|^2 \leqslant \frac{u^{\mathrm{T}}e}{\mu},$$

where

$$\mu = \min_{y \in S} \frac{y^{\mathrm{T}}Wy}{y^{\mathrm{T}}y}, \quad \hat{\mu} = \max_{y \in S} \frac{y^{\mathrm{T}}Wy}{y^{\mathrm{T}}y},$$

and S is the orthant under consideration.

Proof. Since $XWXe = u$, we see that $Wx - X^{-1}u = 0$, so

$$x^{\mathrm{T}}Wx = u^{\mathrm{T}}e.$$

Therefore,

$$\mu \leqslant \frac{u^{\mathrm{T}}e}{x^{\mathrm{T}}x}.$$

Similarly,

$$\hat{\mu} \geqslant \frac{u^{\mathrm{T}}e}{x^{\mathrm{T}}x},$$

and combining these results gives upper and lower bounds on $x^{\mathrm{T}}x$. □

Note that the largest eigenvalue of W is an upper bound on $\hat{\mu}$.

For the special case $u = e$ and the positive orthant, our upper bound matches that of Khachiyan and Kalantari, but our lower bound is different from their bound of $1/(n\mu)$.

4. Conclusions and remarks

We have presented an existence proof showing that any symmetric positive definite matrix can be scaled by a unique positive diagonal matrix, or by a unique diagonal matrix with a given sign pattern, to have arbitrary positive row sums.

The proof is constructive in that it leads to algorithms for computing such a scaling: apply an ordinary differential equation solver to (1). This is one particular homotopy method applied to the solution of the nonlinear equation $XWXe - u = 0$; other methods for solution of nonlinear equations or SOCPs could also be applied.

If the matrix is not positive definite, then the homotopy breaks down at values t for which $(1 - t)I + tW$ is singular.

We have also given upper and lower bounds on the entries of the scaling matrix.

Acknowledgements

I am grateful to Roger Horn, John E. Osborn, Hans Schneider, G.W. Stewart, and Ilya Zavorin for helpful discussions. This work was supported in part by the National Science Foundation under Grant CCR-97-32022.

References

[1] R.A. Brualdi, The DAD theorem for arbitrary row sums, Proc. AMS 45 (1974) 189–194.
[2] R.A. Brualdi, S.V. Parter, H. Schneider, The diagonal equivalence of a nonnegative matrix to a stochastic matrix, J. Math. Anal. Appl. 16 (1966) 31–50.
[3] D. Hershkowitz, A.J. Hoffman, H. Schneider, On the existence of sequences and matrices with prescribed partial sums of elements, Linear Algebra Appl. 265 (1997) 71–92.
[4] W. Hurewicz, Lectures on Ordinary Differential Equations, John Wiley and Sons, New York, 1958.
[5] C.R. Johnson, S.A. Lewis, D.Y. Yau, Possible line sums for a qualitative matrix, Linear Algebra Appl. 327 (2001) 53–60.
[6] B. Kalantari, Scaling dualities and self-concordant homogeneousprogramming in finite dimensional spaces, Tech. Rep. LCSR-TR-359, Department of ComputerScience, Rutgers Univiersity, New Brunswick, NJ, 1999.
[7] B. Kalantari, A theorem of the alternative for multihomogenous functions and its relationship to diagonal scaling of matrices, Linear Algebra Appl. 236 (1996) 1–24.
[8] L. Khachiyan, B. Kalantari, Diagonal matrix scaling and linear programming, SIAM J. Optim. 2 (1992) 668–672.
[9] M.S. Lobo, L. Vandenberghe, S. Boyd, H. Lebret, Applications of second-order cone programming, Linear Algebra Appl. 284 (1996) 193–228.
[10] A.W. Marshall, I. Olkin, Scaling of matrices to achieve specified row and column sums, Numer. Math. 12 (1968) 83–90.
[11] Y. Nesterov, A. Nemirovskii, Interior-Point Polynomial Algorithms in Convex Programming, SIAM Press, Philadelphia, Pennsylvania, 1994.
[12] U.G. Rothblum, H. Schneider, Scalings of matrices which have prespecified row sums and column sums via optimisation, Linear Algebra Appl. 114/115 (1989) 737–764.
[13] R. Sinkhorn, Diagonal equivalence to matrices with prescribed row and column sums. ii, Proc. AMS 45 (1974) 195–198.