Simultaneous Diagonalization of Positive Semi-definite Matrices

Jan de Leeuw

Version 21, May 21, 2017

Abstract

We give necessary and sufficient conditions for solvability of $A_j = XW_jX'$, with the A_j are m given positive semi-definite matrices of order n. The solution X is $n \times p$ and the m solutions W_j are required to be diagonal, positive semi-definite, and adding up to the identity. We do not require that $p \leq n$.

Contents

Note: This is a working paper which will be expanded/updated frequently. All suggestions for improvement are welcome. The directory deleeuwpdx.net/pubfolders/simul has a pdf version, the complete Rmd file, and the bib file.

1 Introduction

There is by now a large literature on simultaneous diagonalization of matrices. The topic is of importance in blind source separation, chemometrics, psychometrics, nonlinear optimization, multilinear algebra, fast matrix multiplication, and quantum physics. We have no intention

of giving a complete overview of the literature, our purpose is mainly to present and generalize the approach of De Leeuw and Pruzansky (1978).

The problem is the following. If $A_1, \cdots, A_m$ are real symmetric matrices of order n, then we want to solve the system of equations $A_j = XW_jX'$, where X is any $n \times p$ matrix and where the W_j are diagonal of order p. In this paper we will restrict our attention to the case in which the A_j are positive semi-definite and in which we require the W_j to be positive semi-definite as well. There are several important special cases that can be distinguished. If $p \leq n$ then Schönemann (1972) presented the main algebraic result, in the context of psychometric data analysis. His computational approach, however, left something to be desired, and consequently it was improved, generalized, and implemented by De Leeuw and Pruzansky (1978). We also note the special case in which $m = 2$, in which some additional results can be obtained (De Leeuw (1982)).

In this paper we generalize the approach of Schönemann (1972) and De Leeuw and Pruzansky (1978) to the case where we do not necessarily have $p \leq n$. As in the more restrictive case, we start by studying the decomposition $A_j = XW_jX'$ without the requirement that the W_j are diagonal. We give necessary and sufficient conditions for the set of solutions to be non-empty, and we describe the complete set of solutions in the case in which the system is solvable. We then characterize the solutions, if any, in the set of all solutions which have diagonal W_j.

2   l-rank

Start with a convenient defintion of the rank of a sequence of positive semi-definite matrices.

Defintion 1: [lrank] Suppose A_j are m positive semi-definite matrices of order n. The tuple $(A_1, \cdots, A_m)$ has *l-rank* less than or equal to p if there is an $n \times p$ matrix X and there are $p \times p$ positive semi-definite matrices W_j, with sum $W_\bullet = I$, such that $A_j = XW_jX'$. It has l-rank equal to p if it has l-rank less than or equal to p, but not less than or equal to $p - 1$. Note that we allow for $p > n$.

It may not be immediately obvious why we require $W_\bullet = I$. Suppose $W_\bullet$ has rank $r < p$. Its eigen decomposition is

$$W_\bullet = \begin{bmatrix} N & N_\perp \end{bmatrix} \begin{bmatrix} E & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} N' \\ N'_\perp \end{bmatrix}$$

with E diagonal and positive definite or order r. Write the W_j in the form

$$W_j = \begin{bmatrix} N & N_\perp \end{bmatrix} \begin{bmatrix} D_j & B_j \\ B'_j & C_j \end{bmatrix} \begin{bmatrix} N' \\ N'_\perp \end{bmatrix}$$

Thus $C_\bullet$, the sum of the C_j, is zero. And because the C_j are positive semi-definite this means all C_j are zero. Also, lemma 1 tells us that if a diagonal element of a positive semi-definite matrix is zero, then the whole corresponding row and column are zero. Thus the B_j are zero as well. It follows that $XW_jX' = (XN)D_j(N'X')$ and thus $(A_1, \cdots, A_m)$ has l-rank less than or equal to $r < p$. Thus it makes sense to require $W_\bullet$ to be at least positive

definite. But then, using for example the Cholesky decomposition $W_\bullet = SS'$, we see that $XW_jX' = (XS)(S^{-1}W_jS^{-T})(S'X')$, so we may as well require $W_\bullet = I$.

Theorem 1: [**lrank**] $(A_1, \cdots, A_m)$, with sum $A_\bullet$, has l-rank $r := \mathbf{rank}(A_\bullet)$.

Proof: Suppose the l-rank of $(A_1, \cdots, A_m)$ is $p < r$. Then $A_\bullet = XX'$, for some $n \times p$ matrix X, and $\mathbf{rank}(A_\bullet) = \mathbf{rank}(X) \leq p < r$, a contradiction. Thus l-rank$(A_1, \cdots, A_m) \geq r$.

To show l-rank$(A_1, \cdots, A_m) \leq r$ we will actually construct the set of all solutions to $A_j = XW_jX'$ with $W_\bullet = I$. Start with the eigenvalue decomposition of $A_\bullet$.

$$A_\bullet = \begin{bmatrix} K & K_\perp \end{bmatrix} \begin{bmatrix} \Lambda^2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} K' \\ K'_\perp \end{bmatrix},$$

where Λ^2 is diagonal and positive definite of order r.

Because $A_\bullet = XX'$ it follows from lemma 2 that $X = K\Lambda L'$, with a $p \times r$ orthonormal L. It remains to solve $A_j = K\Lambda L'W_jL\Lambda K'$. Write

$$A_j = \begin{bmatrix} K & K_\perp \end{bmatrix} \begin{bmatrix} P_j & Q_j \\ Q'_j & R_j \end{bmatrix} \begin{bmatrix} K' \\ K'_\perp \end{bmatrix}$$

Now $A_\bullet K_\perp = 0$ implies that $A_jK_\perp = 0$ for all j, and thus $Q_j = 0$ and $R_j = 0$ for all j. Consequently

$$\Lambda^{-1}K'A_jK\Lambda^{-1} = L'W_jL$$

This means

$$W_j = \begin{bmatrix} L & L_\perp \end{bmatrix} \begin{bmatrix} \Lambda^{-1}K'A_jK\Lambda^{-1} & U_j \\ U'_j & V_j \end{bmatrix} \begin{bmatrix} L' \\ L'_\perp \end{bmatrix} \tag{1}$$

where the U_j and V_j are arbitrary, except for the fact that we must have W_j positive semi-definite, and we must have $U_\bullet = 0$ and $V_\bullet = I$. Remember that L and $L_\perp$ are also arbitrary, except for their orthonormality. One convenient solution for W_j, with minimal Frobenius norm, is $W_j = L\Lambda^{-1}K'A_jK\Lambda^{-1} = X^+A_j(X^+)'$, where X^+ is the Moore-Penrose inverse of X. **QED**

3   d-rank

We now turn to the case in which the W_j are required to be diagonal, and they still have to add up to the identity.

Defintion 2: [**drank**] The sequence $(A_1, \cdots, A_m)$ has *d-rank* less than or equal to p if there is an $n \times p$ matrix X and there are $p \times p$ **diagonal** positive semi-definite matrices W_j, with sum $W_\bullet = I$, such that $A_j = XW_jX'$. It has d-rank equal to p if it has d-rank less than or equal to p, but not less than or equal to $p - 1$. Note that again we allow for $p > n$.

Clearly l-rank is less than or equal to d-rank, which implies by theorem 1 that d-rank is larger than or equal to $r = \mathbf{rank}(A_\bullet)$. In order to proceed we need another definition.

Defintion 3: [pextension] Any $p \times p$ symmetric matrix of order p which has the $r \times r$ matrix A as its leading principal submatrix is a *p-extension* of A.

Theorem 2: [drank] The d-rank of $(A_1, \cdots . A_m)$ is less than or equal to p if and only if there exist for each $j = 1, \cdots, m$ pairwise commuting positive semi-definite p-extensions C_j of $B_j = \Lambda^{-1} K' A_j K \Lambda^{-1}$ that add up to one.

Proof: Equation (1) in the proof of theorem 1 shows there must exist U_j and V_j such that

$$\begin{bmatrix} \Lambda^{-1} K' A_j K \Lambda^{-1} & U_j \\ U_j' & V_j \end{bmatrix} = \begin{bmatrix} L' \\ L_\perp' \end{bmatrix} W_j \begin{bmatrix} L & L_\perp \end{bmatrix}$$

But this is the same as saying there must be p-extensions of the B_j that commute. **QED**

It seems that in general for $p > n$ and $m > 2$ commuting p-extensions are difficult to work with. But in some cases theorem 2 simplifies.

Corollary 1: [small_d_rank] The d-rank of $(A_1, \cdots, A_m)$ is $r = \mathbf{rank}(A_\bullet)$ if and only if $A_j A_\bullet^+ A_\ell = A_\ell A_\bullet^+ A_j$ for all j, ℓ.

Proof: The $\Lambda^{-1} K' A_j K \Lambda^{-1}$ must commute without any p-extension. This translates to the condition in the theorem. **QED**

Finally, we also have as a corollary a version of the basic result of De Leeuw (1982).

Corollary 2: [m=2] The d-rank of (A_1, A_2) is $\mathbf{rank}(A_1 + A_2)$.

Proof: Because $\Lambda^{-1} K' A_1 K \Lambda^{-1} + \Lambda^{-1} K' A_2 K \Lambda^{-1} = I$ we see from lemma 4 that $\Lambda^{-1} K' A_1 K \Lambda^{-1}$ and $\Lambda^{-1} K' A_2 K \Lambda^{-1}$ commute. **QED**

4 Appendix: Some Lemmas

Lemma 1: [diagonal] If A is positive semi-definite and $a_{ii} = 0$ then $a_{ij} = a_{ji} = 0$ for all j.

Proof: Suppose

$$A = \begin{bmatrix} 0 & r' \\ r & S \end{bmatrix}$$

is positive semi-definite. Define $z = \begin{bmatrix} 1 & -\epsilon r \end{bmatrix}$ with $\epsilon > 0$. Then $z' A z = -2\epsilon r' r + \epsilon^2 r' S r$. If $r' S r = 0$ and $r' r > 0$ then $z' A z < 0$ for all $\epsilon > 0$, which contradicts that A is positive semi-definite. If $r' S r > 0$ and $r' r > 0$ then

$$\min_{\epsilon > 0} z' A z = -\frac{(r' r)^2}{r' S r} < 0,$$

which again contradicts that A is positive semi-definite. Thus $r' r = 0$, i.e. $r = 0$. **QED**

Lemma 2: [crossprod] Suppose the positive semi-definite matrix A of order n has eigenvalue decomposition

$$A = \begin{bmatrix} K & K_\perp \end{bmatrix} \begin{bmatrix} \Lambda^2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} K' \\ K_\perp' \end{bmatrix},$$

with Λ^2 a positive definite diagonal matrix of order $r = \mathbf{rank}(A)$. The equation $A = XX'$, with X an $n \times p$ matrix has a solution if and only if $p \geq r$. All solutions are of the form $X = K\Lambda L'$, with L is a $p \times r$ orthonormal matrix.

Proof: Write X as

$$X = \begin{bmatrix} K & K_\perp \end{bmatrix} \begin{bmatrix} U \\ V \end{bmatrix},$$

which gives

$$XX' = \begin{bmatrix} K & K_\perp \end{bmatrix} \begin{bmatrix} UU' & UV' \\ VU' & VV' \end{bmatrix} \begin{bmatrix} K' \\ K'_\perp \end{bmatrix}.$$

Thus $XX' = A_\bullet$ if and only if $V = 0$ and $UU' = \Lambda$. It follows that $X = K\Lambda L'$, with a $p \times r$ orthonormal L. Also $\mathbf{rank}(X) = \mathbf{rank}(A_\bullet) = r$ and $p \geq r$. **QED**

Lemma 3: [simultaneous] Suppose $(A_1, \cdots, A_m)$ is a sequence of real symmetric matrices of order n. Then there exist a square orthonormal X and diagonal W_j such that $A_j = XW_jX'$ if and only if the A_j commute in pairs, i.e. if and only if $A_jA_\ell = A_\ell A_j$ for all $j \neq \ell$.

Proof: It is obvious that simultaneously diagonalizability implies that the A_j commute in pairs.

The interesting part of the proof is to show that pairwise commuting implies simultaneous diagonalizability. The standard proof, repeated most recently in Jiang and Li (2016), uses induction, starting from the fact that the lemma is trivially true for $m = 1$. We give the proof in our notation and make it perhaps a bit more explicit and computational.

So let us suppose the real symmetric matrices $(A_1, \cdots, A_m)$ commute in pairs. And suppose A_m has eigenvalue decomposition

$$A_m = \begin{bmatrix} K_1 & K_2 & \cdots & K_r \end{bmatrix} \begin{bmatrix} \lambda_1 I & 0 & \cdots & 0 \\ 0 & \lambda_2 I & \cdots & \cdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_r I \end{bmatrix} \begin{bmatrix} K'_1 \\ K'_2 \\ \vdots \\ K'_r \end{bmatrix},$$

with all λ_s different. Set $K := \begin{bmatrix} K_1 & K_2 & \cdots & K_r \end{bmatrix}$. Then for all $j = 1, \cdots, m-1$

$$A_m A_j K_s = A_j A_m K_s = \lambda_1 A_j K_s$$

and thus A_jK_s are eigenvectors of A_m with eigenvalue λ_s, i.e. $A_jK_s = K_s(K'_sA_jK_s)$. Write this as

$$K'A_jK = \begin{bmatrix} K'_1A_jK_1 & 0 & \cdots & 0 \\ 0 & K'_2A_jK_2 & \cdots & \cdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & K'_rA_jK_r \end{bmatrix}.$$

Now obviously the matrices $K'A_jK$ commute in pairs, which implies that that the $m-1$ matrices $(K'_sA_1K_s, \cdots, K'_sA_{m-1}K_s)$ commute in pairs for each s. By the induction hypothesis there are square orthonormal L_s and diagonal Φ_{js} such that $K'_sA_jK_s = L_s\Phi_{js}L'_s$. Define

$L := \begin{bmatrix} L_1 & L_2 & \cdots & L_r \end{bmatrix}$. Then

$$L'K'A_jKL = \begin{bmatrix} \Phi_{j1} & 0 & \cdots & 0 \\ 0 & \Phi_{j2} & \cdots & \cdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Phi_{jr} \end{bmatrix},$$

for $j = 1, \cdots, m - 1$, while of course

$$L'K'A_jKL = \begin{bmatrix} \lambda_1 I & 0 & \cdots & 0 \\ 0 & \lambda_2 I & \cdots & \cdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_r I \end{bmatrix}.$$

QED

Lemma 4: [commute] If A and B are symmetric matrices with $A + B = I$ then A and B commute.

Proof: $AB = A(I - A) = A - A^2$, which is symmetric. Another way to see this is that A and $B = I - A$ have the same eigenvectors L, and L diagonalizes both matrices. **QED**

References

De Leeuw, J. 1982. "Generalized Eigenvalue Problems with Positive Semidefinite Matrices." *Psychometrika* 47: 87–94.

De Leeuw, J., and S. Pruzansky. 1978. "A New Computational Method to Fit the Weighted Euclidean Distance Model." *Psychometrika* 43: 479–90.

Jiang, R., and D. Li. 2016. "Simultaneous Diagonalization of Matrices and its Applications in Quadratically Constrained Quadratic Programming." *SIAM Journal ofnOptimization* 26 (3): 1649–68.

Schönemann, P. H. 1972. "An Algebraic Solution for a Class of Subjective Metrics Models." *Psychometrika* 37: 441–51.