NESTED SIMILARITY IN MATRIX THEORY

Annraoi de Paor, MRIA
School of Electrical and Electronic Engineering, University College Dublin,
Belfield, Dublin 4

and

Tomás E Ward*
Insight SFI Research Centre for Data Analytics, Dublin City University,
Glasnevin, Dublin 9

[Accepted 17 December 2019.]

Abstract

We explore the problem of making two matrices similar while at the same time submatrices are similar. An approach is adopted based on ideas from state variable feedback and control theory resulting in an easily implementable solution to the problem. The method is perfectly general and puts no constraints on the matrices other than they are real.

1. Derivation

We consider the problem of nested similar matrices.[1] We begin with a given A and B, real $n \times n$, we want B similar to A, i.e.

$$B = NAN^{-1} \tag{1.1}$$

simultaneously if we partition A and B as follows:

$$A = \left[\begin{array}{c|c} a_{11} & a_1^T \\ \hline a_2 & A_{22} \end{array}\right], B = \left[\begin{array}{c|c} b_{11} & b_1^T \\ \hline b_2 & B_{22} \end{array}\right]$$

with a_{11} 1×1, a_1^T $1 \times (n-1)$, a_2 $(n-1) \times 1$, A_{22} $(n-1) \times (n-1)$, and similarly for B we want B_{22} similar to A_{22}, i.e.,

$$B_{22} = MA_{22}M^{-1} \tag{1.2}$$

We want to explore the conditions under which this is possible, i.e. that matrices

*Corresponding author, e-mail: tomas.ward@dcu.ie

[1] This problem was proposed to the first author by the late Professor J. O. Scanlan, long term editor of the International Journal of Circuit Theory and former President of the Royal Irish Academy.

and submatrices can be similar. To begin, we write (1.1) as $NA = BN$, i.e.

$$\left[\begin{array}{c|c} n_{11} & n_1^T \\ \hline n_2 & N_{22} \end{array}\right] \left[\begin{array}{c|c} a_{11} & a_1^T \\ \hline a_2 & A_{22} \end{array}\right] = \left[\begin{array}{c|c} b_{11} & b_1^T \\ \hline b_2 & MA_{22}M^{-1} \end{array}\right] \left[\begin{array}{c|c} n_{11} & n_1^T \\ \hline n_2 & N_{22} \end{array}\right] \tag{1.3}$$

We take it that A and M are given and that we have to find b_{11}, b_1^T, b_2 (and thus B_{22})and $n_{11}, n_1^T, n_2, N_{22}$. We begin by writing the equation for the $(2,1)$ blocks in (1.3):

$$n_2 a_{11} + N_{22} a_2 = b_2 n_{11} + MA_{22}M^{-1} n_2 \tag{1.4}$$

$$\therefore b_2 n_{11} = \left[a_{11} I - MA_{22}M^{-1}\right] n_2 + N_{22} a_2 \tag{1.5}$$

$$\therefore b_2 = \frac{1}{n_{11}} \{\left[a_{11} I - MA_{22}M^{-1}\right] n_2 + N_{22} a_2\} \tag{1.6}$$

This defines b_2 in terms of n_{11}, n_2 and N_{22}. Note that we choose

$$n_{11} \neq 0 \tag{1.7}$$

We now look at the $(2,2)$ block in (1.3)

$$n_2 a_1^T + N_{22} A_{22} = b_2 n_1^T + MA_{22}M^{-1} N_{22} \tag{1.8}$$

Using (1.6) to substitute for b_2, this gives

$$MA_{22}M^{-1}N_{22} - N_{22}\left[A_{22} - \frac{a_2 n_1^T}{n_{11}}\right] = n_2 a_1^T - \frac{1}{n_{11}}\left[a_{11} I - MA_{22}M^{-1}\right] n_2 n_1^T \tag{1.9}$$

$$\Rightarrow MA_{22}M^{-1}N_{22} I - I N_{22}\left[A_{22} - \frac{a_2 n_1^T}{n_{11}}\right] = n_2 a_1^T - \frac{1}{n_{11}}\left[a_{11} I - MA_{22}M^{-1}\right] n_2 n_1^T \tag{1.10}$$

To address the solution of this we write the N_{22} in terms of its columns.

$$N_{22} = [(N_{22})_1, (N_{22})_2, \ldots, (N_{22})_{n-1}] \tag{1.11}$$

and rearrange this as a single column of dimension $(n-1)^2 \times 1$:

$$n = \begin{bmatrix} (N_{22})_1 \\ (N_{22})_2 \\ \vdots \\ (N_{22})_{n-1} \end{bmatrix} \tag{1.12}$$

Similarly we introduce the notation

$$F = n_2 a_1^T - \frac{1}{n_{11}}\left[a_{11} I - MA_{22}M^{-1}\right] n_2 n_1^T \tag{1.13}$$

and rewrite F in terms of its columns as

$$F = [(F)_1, (F)_2, \ldots, F_{n-1}] \tag{1.14}$$

defining the $(n-1)^2 \times 1$ column f as

$$f = \begin{bmatrix} (F)_1 \\ (F)_2 \\ \vdots \\ (F)_{n-1} \end{bmatrix} \tag{1.15}$$

We now convert (1.10) to Kronecker product notation:

$$\left[I \otimes MA_{22}M^{-1} - \left[A_{22} - \frac{a_2 n_1^T}{n_{11}} \right]^T \otimes I \right] n = f \tag{1.16}$$

Now, as will be shown later, the $(n-1)^2 \times (n-1)^2$ matrix on the left hand side of (1.16) is nonsingular if and only if $MA_{22}M^{-1}$ and $\left[A_{22} - \frac{a_2 n_1^T}{n_{11}} \right]^T$ have no eigenvalues in common, i.e. if and only if A_{22} and $A_{22} - \frac{a_2 n_1^T}{n_{11}}$ have no eigenvalue in common. Hence we invoke state variable feedback theory [1] provided (A_{22}, a_2) form a controllable pair, i.e. provided

$$\det \left[a_2, A_{22}a_2, (A_{22})^2 a_2, \ldots, (A_{22})^{n-2} a_2 \right] \neq 0 \tag{1.17}$$

It is possible to choose the $1 \times (n-1)$ row $\frac{n_1^T}{n_{11}}$ so that $\left[A_{22} - \frac{a_2 n_1^T}{n_{11}} \right]$ has any desired eigenvalues, subject only to complex pairing (since the matrices are real). We specify any $n_{11} \neq 0$ and select n_1^T accordingly.
Now lets look again at F.

$$F = n_2 a_1^T - \frac{1}{n_{11}} \left[a_{11} I - MA_{22}M^{-1} \right] n_2 n_1^T \tag{1.18}$$

Having specified n_{11} and completed n_1^T so that A_{22} and $A_{22} - \frac{a_2 n_1^T}{n_{11}}$ have no eigenvalues in common, all we need to do now is choose n_2,$[(n-1) \times 1]$, so that $F \neq 0$, i.e. $f \neq 0$ in (1.16). We then solve (1.16) for n, i.e. for N_{22} from (1.12).
We have now got n_{11}, n_1^T, n_2 and N_{22}, i.e. we have the $n \times n$ matrix N.
We must check that $\det N \neq 0$. We can write

$$\det N = n_{11}. \det \left[N_{22} - \frac{n_2 n_1^T}{n_{11}} \right]$$

, so that the extra condition here is

$$\det \left[N_{22} - \frac{n_2 n_1^T}{n_{11}} \right] \neq 0 \tag{1.19}$$

Note that we have b_2 from equation (1.6)
Finally, we solve for b_{11} and b_1^T from the equation

$$\left[b_{11}, b_1^T\right] N = \left[n_{11}, n_1^T\right] A \tag{1.20}$$

The first row[2] of BN is the first row of NA. This concludes the solution. It is worth highlighting that in many applications the matrix A may be assumed to have n distinct eigenvalues. In such a case it suffices that A and B= $\left[\begin{array}{c|c} b_{11} & b_1^T \\ \hline b_2 & MA_{22}M^{-1} \end{array}\right]$ have the same characteristic polynomial. Note we make use of the observation that $b_{11} = a_{11}$ when a solution exists. The characteristic polynomial condition in this case can then be expressed through the use of Schur complements as follows:

$$tr(a_1^T A_{22}^j a_2) = tr(b_1^T M A_{22}^j M^{-1} b_2) \tag{1.21}$$

for $j = 0, 1, ..., n-2$. From such a formulation a simple solution can be seen in the form of $B_{22} = MA_{22}M^{-1}, b_1^T = a_1^T M^{-1}, b_2 = Ma_2$.

2. EXAMPLE

We assume A and M are given. In this example

$$A = \begin{bmatrix} 1 & 5 & -3 \\ 5 & -4 & 15 \\ -7 & 2 & 9 \end{bmatrix}, M = \begin{bmatrix} -5 & 6 \\ 23 & -5 \end{bmatrix}$$

where M has been chosen to be invertible. Next we specify any $n_{11} \neq 0$ and select n_1^T accordingly.
So let us arbitrarily choose $n_{11} = 5$
n_1^T is now arbitrarily chosen subject to the constraint that A_{22} and $A_{22} - \frac{a_2 n_1^T}{n_{11}}$ have no eigenvalues in common, for example:

$$n_1^T = \begin{bmatrix} 3 & -7 \end{bmatrix}$$

Examination of the corresponding eigenvalues confirms the appropriateness of this choice of n_1^T. We now choose n_2,$[(n-1) \times 1]$, so that $F \neq 0$, for example with:

$$n_2 = \begin{bmatrix} 5 \\ 6 \end{bmatrix}$$

This yields

$$F = \begin{bmatrix} 16.1965 & 5.5416 \\ 224.3363 & -471.4513 \end{bmatrix}$$

We use F then to solve (1.16) for n, i.e. for N_{22} from (1.12).

[2] We would like to thank an anonymous reviewer for this insight

We have now got n_{11}, n_1^T, n_2 and N_{22}, i.e. we have the $n \times n$ matrix N having checked that $\det N \neq 0$. N in this case is revealed as

$$N = \begin{bmatrix} 5 & 3 & -7 \\ 5 & -0.63703 & -1.66247 \\ 6 & 11.07589 & -15.44987 \end{bmatrix}$$

We calculate b_2 from equation (1.6) and the remaining unknown elements of B from (1.20)
This reveals B as

$$B = \begin{bmatrix} 1 & 12.379 & 0.35083 \\ 4.62494 & -2.85841 & 0.76991 \\ -32.07305 & 56.54867 & 7.85841 \end{bmatrix}$$

and concludes the calculation.

In summary this worked example demonstrates how for a given A and M the method described in this paper provides a means for calculating B and N such that:

$$B = NAN^{-1} \tag{2.1}$$

and simultaneously with A and B partitioned as follows:

$$A = \left[\begin{array}{c|c} a_{11} & a_1^T \\ \hline a_2 & A_{22} \end{array}\right], B = \left[\begin{array}{c|c} b_{11} & b_1^T \\ \hline b_2 & B_{22} \end{array}\right]$$

with a_{11} 1×1, a_1^T $1 \times (n-1)$, a_2 $(n-1) \times 1$, A_{22} $(n-1) \times (n-1)$, that B_{22} is similar to A_{22}, i.e.,

$$B_{22} = MA_{22}M^{-1} \tag{2.2}$$

3. APPENDIX

Nonsingularity of the $(n-1)^2 \times (n-1)^2$ matrix, $\left[I \otimes MA_{22}M^{-1} - \left[A_{22} - \frac{a_2 n_1^T}{n_{11}}\right]^T \otimes I\right]$.
To simplify the notation let's write this as $[I \otimes P - Q \otimes I]$.
Now let V_i be an eigenvector of P belonging to the eigenvalue λ_i and W_j be an eigenvector of Q belonging to the eigenvalue ρ_i. Now consider

$$\begin{aligned} [I \otimes P - Q \otimes I]\,[W_j \otimes V_i] &= [I \otimes P]\,[W_j \otimes V_i] - [Q \otimes I]\,[W_j \otimes V_i] \\ &= [IW_j \otimes PV_i] - [QW_j \otimes IV_i] \\ &= [W_j \otimes \lambda_i V_i] - [\rho_i W_j \otimes V_i] \\ &= (\lambda_i - \rho_i)\,[W_j \otimes V_i] \end{aligned}$$

Therefore $W_j \otimes V_i$ is an eigenvector of $[I \otimes P - Q \otimes I]$ belonging to the eigenvalue $\lambda_i - \rho_i$. A matrix is singular if and only if it has 0 as an eigenvalue. Therefore

$[I \otimes P - Q \otimes I]$ is nonsingular if and only if P and Q have no eigenvalues in common, i.e. if and only if $MA_{22}M^{-1}$ and $\left[A_2 - \frac{a_2 n_1^T}{n_{11}}\right]^T$ have no eigenvalues in common, i.e. if and only if A_{22} and $A_2 - \frac{a_2 n_1^T}{n_{11}}$ have no eigenvalues in common.
As mentioned, provided (A_{22}, a_2) is a controllable pair, $\frac{n_1^T}{n_{11}}$ can always be chosen to ensure this. Interestingly and relevant,[3] a recent result [2] examines conditions under which the Sylvester equation $PX - XQ = Z$ where P and Q may have some common eigenvalues, is solvable.

References

[1] H. M. Power, "A new result on eigenvalue assignment by means of dyadic output feedback," *International Journal of Control*, vol. 21, pp. 149–158, Jan 1975.

[2] N. Č. Dinčić, "Solving the sylvester equation ax-xb=c when $\sigma(a) \cap \sigma(b) \neq \emptyset$," *Electronic Journal of Linear Algebra*, vol. 35, pp. 1–23, 2019.

[3] Again we acknowledge the valuable contribution of an anonymous reviewer for drawing our attention to this result