A Matrix Inverse Eigenvalue Problem and Its Application

Nian Li

School of Mathematical Sciences
Swinburne University of Technology
Hawthorn 3122, Australia

Submitted by Ludwig Elsner

ABSTRACT

The problem of generating a matrix $A$ with specified eigenvalues, which maps a given set of vectors of another given set, is presented. An existence theorem is given and proved. A stable algorithm for producing the matrix $A$ is discussed. The relation between this problem and the pole assignment problem in control theory is investigated. The application of this problem in the design of neural networks is discussed. © 1997 Elsevier Science Inc.

1. INTRODUCTION

In this paper we consider the following matrix inverse eigenvalue problem (MIEP): Given two sets of real $n$-vectors $(x_1, x_2, \ldots, x_p)$ and $(y_1, y_2, \ldots, y_p)$, $p \leq n$, and an arbitrary set of complex numbers $\mathcal{L} = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$, find a real $n \times n$ matrix $A$ such that

$$Ax_i = y_i, \quad i = 1, 2, \ldots, p,$$

and the spectrum of $A$

$$\rho(A) = \mathcal{L},$$

LINEAR ALGEBRA AND ITS APPLICATIONS 266:143-152 (1997)

© 1997 Elsevier Science Inc. All rights reserved.

0024-3795/97/$17.00

655 Avenue of the Americas, New York, NY 10010

PH 0024-3795(96)00639-8
where we assume the set \( \{ x_1, x_2, \ldots, x_p \} \) is linearly independent and \( \mathcal{L} \) is closed under complex conjugation, i.e.

\[
\lambda \in \mathcal{L} \iff \overline{\lambda} \in \mathcal{L}.
\]

The prototype of this problem initially arose in the design of Hopfield neural networks, where only the case of \( p = 1 \) was concerned [6]. In Section 2, an existence theorem is presented, which gives a sufficient condition for the existence of \( A \) using some system-stabilizing techniques from control theory. In the proof of this theorem, the following basic concepts and a theorem from control theory are needed.

A pair of matrices, an \( n \times n \) matrix \( A \) and a full-rank \( l \times n \) matrix \( C \), where \( l \leq n \), is said to be **completely observable** if the observability matrix

\[
\begin{bmatrix}
C \\
CA \\
\vdots \\
CA^{n-1}
\end{bmatrix}
\]

has full rank \( n \), which is equivalent to the matrix

\[
\begin{bmatrix}
A - sI_n \\
C
\end{bmatrix}
\]

being of full rank for each eigenvalue \( s \) of \( A \). We also call the matrix (5) the observability matrix. A pair of matrices, an \( n \times n \) matrix \( A \) and a full-rank \( n \times m \) matrix \( B \), where \( m \leq n \), is said to be **completely controllable** if and only if the pair of \( A' \) and \( B' \) is completely observable.

The following theorem is well known.

**Theorem 1** (Wonham [12]). For any given \( \mathcal{L} \), if the pair of \( A \in \mathbb{R}^{n \times n} \) and \( C \in \mathbb{R}^{l \times n} \) is completely observable, then there is a real \( n \times l \) matrix \( K \) such that

\[
\rho(A + KC) = \mathcal{L}.
\]

Also, if the pair \( A \in \mathbb{R}^{n \times n} \) and \( B \in \mathbb{R}^{n \times m} \) is completely controllable, then there is a real \( m \times n \) matrix \( K \) such that

\[
\rho(A + BK) = \mathcal{L}.
\]
The proof of the existence theorem also gives a method for generating the matrix $A$. In this method the MIEP is first converted to a pole assignment problem (PAP) in control theory, then several existing approaches may be applied to solve this problem.

The discussion in Section 4 shows that there is a close relation between the PAP problem and the MIEP problem. The application of the MIEP problem in the design of additive neural networks is discussed in Section 4.

In Section 5 two direct methods are presented. One of them can produce a robust solution (one which is not sensitive to changes in data), while the other needs less calculation for obtaining $A$.

2. AN EXISTENCE THEOREM

Let $X = [x_1, x_2, \cdots, x_p]$, $Y = [y_1, y_2, \cdots, y_p]$, and $X^\dagger$ be the Moore-Penrose generalized inverse of $X$ (see [2]). We now can state the main theorem of this paper.

**Theorem 2.** Suppose $X$ is of full rank. Then the MIEP problem is solvable for any $\mathcal{L}$ if the matrix $Y - sX$ is of full rank for $s \in \rho(YX^\dagger)$.

**Proof.** Equation (1) gives

$$AX = Y.$$  \hspace{1cm} (8)

Therefore

$$A = YX^\dagger + W,$$ \hspace{1cm} (9)

where $W$ is any $n \times n$ matrix such that $WX = 0$. We do QR decomposition for $X$,

$$X = \begin{bmatrix} O_1 & O_2 \end{bmatrix} \begin{bmatrix} R \\ 0 \end{bmatrix},$$ \hspace{1cm} (10)

where $[O_1 \ O_2]$ is an orthogonal matrix and $R$ is a $p \times p$ invertible matrix. It is easy to see that the columns of $O_2$ span the orthogonal complement of the subspace spanned by columns of $X$ and $W = ZO_2^t$ for some $n \times (n - p)$ matrix $Z$. Thus (9) can be written

$$A = YX^\dagger + ZO_2^t.$$ \hspace{1cm} (11)
If we can prove the matrix pair of $YX^+$ and $O_2^t$ is completely observable, then from Theorem 1, the proof is complete. Notice that the rank of the observability matrix

$$
\begin{bmatrix}
YX^+ - sI_n \\
O_2^t
\end{bmatrix}
$$

(12)
is the same as that of

$$
\begin{bmatrix}
Y - sX^* \\
0
\end{bmatrix} = 
\begin{bmatrix}
YX^+ - sI_n \\
O_2^t
\end{bmatrix} \begin{bmatrix} X & O_2 \end{bmatrix}.
$$

(13)
The conclusion now follows from the assumption of the theorem.

Since (see Horn and Johnson [4, 53–54])

$$
\det(sI_n, YX^*) = s^{n-p} \det(sI_p, X^*Y),
$$

(14)
we need to check the rank of $Y - sX$ only for $s = 0$ and for eigenvalues of a smaller matrix $X^*Y$.

For a matrix to have full rank is a "black-and-white" concept. It is oversimplifying to describe an MIEP problem as solvable if it is close to an unsolvable neighbor. In terms of the numerical determination of the rank of $Y - sX$, it is well known that different error tolerances will produce different answers. To overcome these problems we can introduce solvability measures for MIEP problems, which are similar to controllability measures in control theory. For details one can refer to [8] and [9].

There are several methods for generating $Z$. Some of them are not numerically stable. Kautsky, Nichols, and Van Dooren first introduced the robust pole assignment problem [5] and presented a stable method for producing $Z$. Later on Byers and Nash [1] extended this work and developed a more accurate and economical algorithm. Li and Chu [7] also developed a robust algorithm for the output feedback pole assignment problem.

When solving the MIEP problem it is not necessary to convert it to a PAP problem. Inspired by the ideas initiated by Kautsky et al. [5], in Section 5 we shall present a direct algorithm for generating the robust solution of the MIEP problem.
3. POLE ASSIGNMENT PROBLEMS

In the main theorem we convert the MIEP problem to a pole assignment problem, then derive a condition for the MIEP problem being solvable. Interestingly, a pole assignment problem can also be converted to an MIEP problem.

Let \( A \in \mathbb{R}^{n \times n} \) and \( C \in \mathbb{R}^{t \times n} \) be a completely observable pair. We assume that \( C \) is of full rank as before. The \( QR \) decomposition for \( C \)

\[
C = [S \quad 0] \begin{bmatrix} U_1^t \\ U_2^t \end{bmatrix}
\]  

(15)

gives \( U_2 \) whose columns form the orthogonal complement of the subspace spanned by the rows of \( C \). Also we have \( C^t \_ U_1 S^{-1} \). As the pair of \( A \) and \( C \) is completely observable, and \( [C^t \_ U_2] \) is of full rank, so too is the matrix

\[
\begin{bmatrix} * & AU_2 \_ sU_2 \\ I_t & 0 \end{bmatrix} = \begin{bmatrix} A \_ sI_n \\ C \end{bmatrix} [C^t \quad U_2].
\]

(16)

Thus the matrix \( AU_2 \_ sU_2 \) is of full rank, and from Theorem 2, there is a matrix \( \tilde{A} \) with

\[
\tilde{A}U_2 = AU_2 \quad \text{and} \quad \rho(\tilde{A}) = \mathcal{L}.
\]

(17)

The \( K \) matrix then can be obtained through the equation

\[
K = (\tilde{A} \_ A)C^t.
\]

(18)

The following direct calculations show that \( K \) is the required matrix

\[
A + KC = A + (\tilde{A} \_ A)C^tC
\]

\[
= A + (\tilde{A} \_ A)U_1U_1^t
\]

\[
= A + (\tilde{A} \_ A)(I_n \_ U_2U_2^t)
\]

\[
= \tilde{A} \_ (\tilde{A} \_ A)U_2U_2^t
\]

\[
= \tilde{A}.
\]

(19)
A similar result can be obtained for the state feedback PAP problem which finds a matrix $K$ such that $\rho(A + BK) = \mathcal{Z}$ given the matrices $A$ and $B$.

4. THE DESIGN OF THE NEURAL NETWORKS

An additive neural network can be described by a set of differential equations

$$
\frac{d u_i}{dt} = -a_i u_i + \sum_{j=1}^{n} \omega_{ij} g_j(u_j) + p_i, \quad i = 1, \ldots, n,
$$

(20)

where $g_j$ is a squashing function which is strictly increasing ($g'_j > 0$) and approaching fixed limits for large negative and positive values of $u_j$, $\omega_{ij}$ is the connection coefficient between the $i$th and $j$th neurons, and $p_i$ is the constant input to the $i$th neuron [3]. The equivalent matrix form of (20) is

$$
\frac{d u}{dt} = -Au + Wg(u) + p,
$$

(21)

where $u = [u_1, u_2, \ldots, u_n]'$, $W = (\omega_{ij})$, $g(u) = [g_1(u_1), g_2(u_2), \ldots, g_n(u_n)]'$, $A = \text{diag}(a_1, a_2, \ldots, a_n)$, and $p = [p_1, p_2, \ldots, p_n]'$. In a design problem, given $A$, $g_i$, $p$, and a point $u^* \in \mathbb{R}^n$, $W$ is to be found such that the network has the point $u^*$ as a stable equilibrium. As the derivative of $u$ with respect to $t$ at $u^*$ equals zero, we have

$$
-Au^* + Wg(u^*) + p = 0.
$$

(22)

Since $u = u^*$ is a stable solution, the Jacobian of the system at this point,

$$
J = -A + WG_d,
$$

(23)

where $G_d = \text{diag}(g_1'(u^*), g_2'(u^*), \ldots, g_n'(u^*))$, should be stable, that is, its eigenvalues should be in the left half plane. Combining with (22), one obtains

$$
JG^{-1}_d g(u^*) = Au^* - AG_d^{-1} g(u^*) - p,
$$

(24)

or simply

$$
Jx = y,
$$

(25)

where $x = G_d^{-1} g(u^*)$ and $y = Au^* - AG_d^{-1} g(u^*) - p$.
If $L^\prime$ is chosen so that each element has a negative real part, then determining $J$ with $\rho(J) = L^\prime$ becomes an MIEP problem. The required $W$ can be calculated through the equation

$$W = (J + A)G_d^{-1}. \quad (26)$$

5. DIRECT METHODS

5.1. An Explicit Formula for Calculating $A$

From the existence theorem in Section 2, if the matrix $Y - sX$ is of full rank, then the MIEP problem is solvable. A sufficient condition for $Y - sX$ to be of full rank is that the $2p$ vectors $x_i, y_i$, $i = 1, 2, \ldots, p$, form a linearly independent set in $R^n$, or $\text{rank}([X \ Y]) = 2p$.

**Theorem 3.** If $p \leq n/2$ and the combined set of vectors $x_i$ and $y_i$, $i = 1, 2, \ldots, p$, is linearly independent, then the MIEP is solvable and a required matrix is explicitly given by Equations (27)–(30).

**Proof.** We construct a block-diagonal matrix $A_d \in R^{n \times n}$ as follows: if the total number $2q$ of complex eigenvalues is not greater than $2p$,

$$A_d = \text{diag}\left(\begin{bmatrix} 0 & -\lambda_1 \lambda_2 \\ 1 & \lambda_1 + \lambda_2 \end{bmatrix}, \ldots, \begin{bmatrix} 0 & -\lambda_{2p-1} \lambda_{2p} \\ 1 & \lambda_{2p-1} + \lambda_{2p} \end{bmatrix}, \lambda_{2p+1}, \ldots, \lambda_n\right); \quad (27)$$

otherwise,

$$A_d = \text{diag}\left(\begin{bmatrix} 0 & -\lambda_1 \lambda_2 \\ 1 & \lambda_1 + \lambda_2 \end{bmatrix}, \ldots, \begin{bmatrix} 0 & -\lambda_{2q-1} \lambda_{2q} \\ 1 & \lambda_{2q-1} + \lambda_{2q} \end{bmatrix}, \lambda_{2q+1}, \ldots, \lambda_n\right). \quad (28)$$

Without loss of generality we can assume that any conjugate pair of numbers in $L^\prime$ appears in the same $2 \times 2$ block. It should be noted that the matrix $A_d$ is real under this arrangement and $\rho(A_d) = L^\prime$.

Since the set of vectors $\{x_1, y_1, \ldots, x_p, y_p\}$ is linearly independent, there are $n - 2p$ $n$-vectors $z_1, z_2, \ldots, z_{n-2p}$ in $R^n$ such that the extended set of $n$ $n$-vectors

$$\{x_1, y_1, \ldots, x_p, y_p, z_1, z_2, \ldots, z_{n-2p}\}$$

is still linearly independent. Thus the matrix

$$T = \begin{bmatrix} x_1 & y_1 & \cdots & x_p & y_p & z_1 & z_2 & \cdots & z_{n-2p} \end{bmatrix} \quad (29)$$

is nonsingular.
We claim that the matrix

$$A = T A_d T^{-1}$$

(30)
is the required one. As $A$ is similar to $A_d$, we have $\rho(A) = \rho(A_d) = \mathcal{L}$. Direct calculations give that

$$A x_i = y_i, \quad i = 1, 2, \ldots, p. \quad \blacksquare$$

(31)

Although this theorem provides a rather simple method to obtain the matrix $A$, unfortunately it needs the strong assumption $\text{rank}([X \ Y]) = 2p$. Furthermore, it may produce numerically unstable solutions. If two eigenvalues in the same block, say $\lambda_1$ and $\lambda_2$, are very close, then the eigenvalues of the computed $2 \times 2$ matrix block

$$A_{d1} = \begin{bmatrix} 0 & -\lambda_1 \lambda_2 \\ 1 & \lambda_1 + \lambda_2 \end{bmatrix}$$

will be very sensitive to the perturbations in the data. In fact

$$\text{diag}\{\lambda_1, \lambda_2\} = P^{-1} A_{d1} P,$$

(32)

where $P = \begin{bmatrix} \lambda_2 & \lambda_1 \\ -1 & -1 \end{bmatrix}$, and the Frobenius condition number of $P$,

$$k_F(P) = \frac{2 + |\lambda_1|^2 + |\lambda_2|^2}{|\lambda_2 - \lambda_1|},$$

which is an upper bound of the rate of change of the eigenvalues by the Bauer-Fike theorem [1], is large. Also, when the angles between the column vectors in $[X \ Y]$ are too small, $T$ will be ill conditioned (see [11]), regardless of the selection of $z_i$, $i = 2p + 1, \ldots, n$. Let

$$S = \text{diag}\left\{ \begin{bmatrix} \lambda_2 & \lambda_1 \\ -1 & -1 \end{bmatrix}, \ldots, \begin{bmatrix} \lambda_{2q} & \lambda_{2q-1} \\ -1 & -1 \end{bmatrix} \right\}, \quad q < p, \quad (33)$$

or

$$S = \text{diag}\left\{ \begin{bmatrix} \lambda_2 & \lambda_1 \\ -1 & -1 \end{bmatrix}, \ldots, \begin{bmatrix} \lambda_{2q} & \lambda_{2q-1} \\ -1 & -1 \end{bmatrix} \right\}, \quad q > p. \quad (34)$$
Then we have $T A_d T^{-1} - T S \Lambda (T S)^{-1}$, and the condition number of the transformation matrix $T S$ gives an upper bound on the rate of change of the eigenvalues with respect to the perturbations in the data.

The following algorithm is preferable.

### 5.2. A Robust Algorithm

In this subsection the conditions for the MIEP to be solvable, as given in Theorem 1, are assumed to hold. Also it is assumed that $A$ is diagonalizable and

$$A = V \Lambda V^{-1}, \tag{35}$$

where $\Lambda = \text{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ and $\lambda_i \in \Sigma$. Equation (8) gives

$$\Lambda V^{-1} X = V^{-1} Y. \tag{36}$$

Let

$$V^{-1} = [v_1, v_2, \ldots, v_n]^T. \tag{37}$$

Then we have $v_i^T (\lambda_i X - Y) = 0$, or $v_i \in \text{Null}((\lambda_i X - Y)^T)$. By the Bauer-Fike theorem the insensitivity of the eigenvalues of $A$ to perturbations in the data is related to the condition number of $V$ or $V^{-1}$. A smaller condition number of $V$ gives greater insensitivity. So the robust solution can be obtained by minimizing the condition number of $V^{-1}$. The algorithm consists of three basic steps:

**Step 1.** Find an orthogonal basis $s_1^i, s_2^i, \ldots, s_k^i$ of the null space of $(\lambda_i X - Y)^T$ for each $\lambda_i$.

**Step 2.** Select vectors $v_i = \sum w_{ik} s_k^i$ such that $V^{-1} = [v_1, v_2, \ldots, v_n]^T$ is well conditioned.

**Step 3.** Calculate $A$ through $A = V \Lambda V^{-1}$.

For detailed numerical methods to implement the above algorithm one can refer to the papers of Kautsky et al. [5], Byers et al. [1], and Li et al. [7].

One thing should be pointed out. In step 2, in order to get a real matrix $A$, when $\lambda_i = \bar{\lambda}_j$ one has to select $v_i = \bar{v}_j$. Let $E(i, j)$ be a matrix generated by interchanging the $i$th and the $j$th row of the $n \times n$ identity matrix $I_n$. If there is only one complex pair $\{\lambda_i, \lambda_j\}$ with $\lambda_i = \bar{\lambda}_j$ in $\Sigma$, then one obtains

$$\bar{A} = \bar{V} \bar{\Lambda} \bar{V}^{-1} = VE(i, j) \bar{\Lambda} E(i, j) V^{-1} = V \Lambda V^{-1} - A. \tag{38}$$
So $A$ is real. One can show that, using this rule for the selection of $v_i$ for the case of multiple pairs of complex numbers, this is also true.

Interestingly, the procedure discussed in this section is equivalent to solving the MIEP by first converting it to a PAP problem and then applying the robust algorithm of Kautsky et al. [5].

6. CONCLUSIONS

In this paper the problem of generating a matrix $A$ with specified eigenvalues, which maps a given set of vectors to another given set, has been presented. An existence theorem was given and proved. Methods for obtaining $A$ with their respective advantages and disadvantages were presented and illustrated with applications to the design of neural networks.

REFERENCES


Received 3 January 1994; final manuscript accepted 23 September 1996