INVERSE POLYNOMIAL MATRIX APPROXIMATION

JERZY KISILEWICZ*

A computational algorithm for matrix C(z) which approximates the inverse of a given matrix Y(z) is described in this paper. Finite degree polynomials as the entries of the matrices C(z) and Y(z) are assumed. The sum of squares of coefficients of polynomials, which are the elements of difference $C(z)Y(z) - Iz^{-h}$ is assumed as a criterion of the approximation quality, where I is an adequate unit matrix.

1. Introduction

The physically realizable inverse function approximation problems are solved in many technical domains such robustness control design (Chen and Guo, 1990; Zhou and Gu, 1992), optimal control problems (Mutoh and Nikiforuk, 1992) and particularly optimal control problems in Hardy space \mathcal{H}^{∞} (Nyman, 1991; Özbay and Tannenbaum, 1990; Smith, 1990). The approximation problems of inverse matrices whose entries are finite degree polynomials are solved for instance in computation of a state-space realization of a linear dynamic system (Grasselli and Tornmbe, 1992) and in multichannel interference equalization problems in data transmission (Clark, 1977; Dąbrowski, 1979; 1982; Kisilewicz, 1990; 1991; 1992). If equalizer is realized as a transversal filter network, then elements of approximated matrix should be definite degree polynomials.

This equalization problem is described and solved for a single channel in (Clark, 1977; Dąbrowski, 1979; 1982) and for more than one channel in (Kisilewicz, 1990; 1991). Multichannel interference is given as Y(z) matrix of polynomials whose degrees are not greater than G. Transfer function of equalizer is C(z) matrix of polynomials whose degrees are equal to M. The C(z) matrix is computed so that product C(z)Y(z) approximates the unit matrix multiplied by z^{-h} .

Different criteria of approximation performance are described in (Clark, 1977; Dąbrowski, 1982). The mean square error is mainly used as such a criterion in channel equalization problems. The problem of multichannel equalizer C(z) synthesis giving minimum of mean square error is described in (Kisilewicz, 1990; 1991; 1992). If the noise level in transmission channels is low and can be omitted, then mean square error is equal to the sum of squares of coefficients of polynomials, which are the entries of difference $C(z)Y(z) - Iz^{-h}$, where I is the respective identity matrix.

An easy-to-compute algorithm giving the polynomials matrix C(z) minimizing the mean square error is presented in this paper. The C(z) matrix, which is computed

^{*} Institute of Control and Systems Engineering, Technical University of Wrocław, Wybrzeże Wyspiańskiego 27, 50–370 Wrocław, Poland

for assumed value h, is denoted later by $\overline{C}(z,h)$. Other matrices, polynomials and coefficients for assumed value h are represented in the same way.

2. Problem Formulation

Let N be the number of columns and rows in a square matrix Y(z), whose entries are not greater than polynomials of degree G

$$\overline{y}_{ij}(z) = \sum_{k=0}^{G} y_k^{ij} z^{-k} \tag{1}$$

where i, j = 1, 2, ..., N.

The matrix $\overline{C}(z,h)$ should be computed. The entries of C(z,h) are polynomials of degree M

$$\overline{c}_{ij}(z,h) = \sum_{k=0}^{M} c_k^{ij}(h) z^{-k}$$

$$\overline{C}(z,h) = [\overline{c}_{ij}(z,h)], \quad (i,j = 1, 2, ..., N)$$
(2)

and should satisfy $\overline{C}(z,h)Y(z) \simeq z^{-h}I_N$, where I_N is the identity matrix having N rows and columns. Matrix $z^h\overline{C}(z,h)$ is close to the matrix that is inverse to $\overline{Y}(z)$. Let $\overline{E}(z,h)$ be a matrix of polynomials

$$\overline{e}_{ij}(z,h) = \sum_{k=0}^{M+G} e_k^{ij}(h) z^{-k}$$

$$\overline{E}(z) = [\overline{e}_{ij}(z,h)], \quad (i,j=1,2,...,N)$$
(3)

such that

$$\overline{E}(z,h) = \overline{C}(z,h)\overline{Y}(z) - z^{-h}I_N$$
(4)

Computing matrix $\overline{C}(z,h)$ means computing such number h, (h = 0, 1, ..., M + G) and such coefficients $c_k^{ij}(h)$, (i, j = 1, 2, ..., N, k = 0, 1, ..., M + G) which minimize the following criterion

$$Q_{h} = \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=0}^{M+G} \left(e_{k}^{ij}(h) \right)^{2}$$
(5)

Criterion (5) approximates the mean square error if noise level in transmission channels is very low and can be omitted (Clark, 1977; Dąbrowski, 1979; 1982).

3. Matrix Form of Criterion Q_h

To calculate the coefficients $c_k^{ij}(h)$ (i, j = 1, 2, ..., N, k = 0, 1, ..., M + G) minimizing Q_h it is suitable to have a formula giving dependence of Q_h on $c_k^{ij}(h)$. It is convenient to write this formula as a relationship of matrices as shown in (Clark, 1977).

Let

$$\delta_i^2(h) = \sum_{j=1}^N \sum_{k=0}^{M+G} \left(e_k^{ij}(h) \right)^2 \tag{6}$$

and single row vector

$$E_{ij}(h) = \left[e_0^{ij}(h), e_1^{ij}(h), \dots, e_{M+G}^{ij}(h)\right]$$
(7)

The criterion Q_h from (5) and (6) can be written as

$$Q_h = \sum_{i=1}^N \delta_i^2(h) \tag{8}$$

where

$$\delta_i^2(h) = \sum_{j=1}^N E_{ij}(h) E_{ij}^T(h)$$
(9)

and superscript T denotes the matrix transposition.

We can assume from (4) that the vector $E_{ij}(h)$ contains coefficients $e_k^{ij}(h)$ of the polynomials

$$\bar{e}_{ij}(z,h) = \sum_{r=1}^{N} \bar{c}_{ir}(z,h) \overline{y}_{rj}(z) - \Delta_{ij} z^{-h}$$
(10)

where

$$\Delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

To obtain the coefficients of polynomials $\bar{c}_{ir}(z,h)\bar{y}_{rj}(z)$, we create (M+1)-row convolution matrices Y_{rj} (Clark, 1977; Dabrowski, 1982), the *p*-th rows of which (p = 1, 2, ..., M + 1) contain M + G + 1 elements: p - 1 nulls, G + 1 coefficients $y_0^{rj}, ..., y_G^{rj}$, and M + 1 - p nulls. These rows of matrices Y_{rj} can be written as

$$\left[0,...,0,y_{0}^{rj},y_{1}^{rj},...,y_{G}^{rj},0,...,0\right]$$

Now we can write from (10)

$$E_{ij}(h) = \sum_{r=1}^{N} C_{ir}(h) Y_{rj} - \mathcal{E}_{ij}(h)$$
(11)

where $C_{ij}(h)$, $\mathcal{E}_{ij}(h)$ are single row vectors

$$C_{ij}(h) = \left[c_0^{ij}(h), c_1^{ij}(h), ..., c_M^{ij}(h)\right]$$
(12)

$$\mathcal{E}_{ij}(h) = \Delta_{ij} \left[\Delta_{0h}, \Delta_{1h}, \dots \Delta_{M+G,h} \right].$$
(13)

Using (11) in (9) we finally obtain

$$\delta_{i}^{2}(h) = \sum_{j=1}^{N} \left(\sum_{r=1}^{N} C_{ir}(h) Y_{rj} - \mathcal{E}_{ij}(h) \right) \left(\sum_{r=1}^{N} C_{ir}(h) Y_{rj} - \mathcal{E}_{ij}(h) \right)^{T}$$
(14)

4. Minimization of Criterion Q_h for assumed h

To find matrix C(z,h) minimizing criterion Q_h we have to find (M + 1)-component vectors $C_{ir}(h)$ (i, r = 1, 2, .., N), minimizing errors $\delta_i^2(h)$ giving (14) for each *i*. Comparing to zero derivatives of $\delta_i^2(h)$ with respect to $c_r^{ik}(h)$, we can obtain for each i, j = 1, 2, ..., N the following matrix equations

$$\sum_{j=1}^{N} \sum_{q=1}^{N} C_{iq}(h) Y_{qj} Y_{rj}^{T} = \sum_{j=1}^{N} \mathcal{E}_{ij}(h) Y_{rj}^{T}$$
(15)

We can notice from (13) that $\mathcal{E}_{ij}(h)$ for $i \neq j$ is a null matrix, so

$$\sum_{j=1}^{N} \mathcal{E}_{ij}(h) Y_{rj}^{T} = \mathcal{E}_{ii}(h) Y_{ri}^{T}$$
(16)

 and

$$\sum_{j=1}^{N} Y_{qj} Y_{rj}^{T} = \left[Y_{q1}, Y_{q2}, ..., Y_{qN} \right] \left[Y_{r1}, Y_{r2}, ..., Y_{rN} \right]^{T}$$
(17)

Defining matrix

$$Y = \begin{bmatrix} Y_{ij} \end{bmatrix} = \begin{bmatrix} Y_{11} \dots Y_{1N} \\ \dots \\ Y_{N1} \dots Y_{NN} \end{bmatrix}$$
(18)

and multiplying it by its transpose we obtain all sums of products $Y_{qj}Y_{rj}^T$ (17) necessary in formula (15). We can now write this formula as follows

$$\left[C_{i1}(h), C_{i2}(h), ..., C_{iN}(h)\right] \left[r - \text{th } M + 1 \text{ columns of } YY^T\right] = \mathcal{E}_{ii}(h)Y_{ri}^T$$
(19)

for i, r = 1, 2, ..., N. Combining (19) for all i and r, we finally obtain,

$$C(h)\left(YY^{T}\right) = \mathcal{E}(h)Y^{T}$$
⁽²⁰⁾

where $C(h) = [C_{ij}(h)]$ is the N row and N(M+1) column matrix, Y - the N(M+1) row and N(M+G+1) column matrix,

$$\mathcal{E}(h) = \begin{bmatrix} \mathcal{E}_{11}(h) & 0 & \dots & 0 \\ 0 & \mathcal{E}_{22}(h) & \dots & 0 \\ 0 & 0 & \dots & \mathcal{E}_{NN}(h) \end{bmatrix}$$
(20a)

 \mathcal{E}^h is the N row and N(M+G+1) column matrix since $\mathcal{E}_{ii}(h)$ are the (M+G+1)component row vectors.

Formula (20) represents the system of N(M+1) linear equations with N right hand side free values. This formula is suitable for computation of matrix C(h). Formally matrix C(h) is given by formula

$$C(h) = \mathcal{E}(h)Y^T \left(YY^T\right)^{-1}$$

Defining $E(h) = [E_{ij}(h)]$ (i, j = 1, ..., N), we can write from (11) an equalized interference matrix as

$$E(h) = C(h)Y - \mathcal{E}(h) \tag{21}$$

5. Computation Algorithm for Unknown h

Since value h is not given, it is necessary to solve equations (20) and compute Q_h from (8) and (9) for each possible h = 0, 1, ..., M + G, and then to choose such value h which minimizes criterion Q_h .

It is convenient to exclude matrix $\mathcal{E}(h)$ in practical computations. Rewriting (19) for h = 0, 1, ..., M + G and defining

$$C_{ij} = \begin{bmatrix} C_{ij}(0) \\ C_{ij}(1) \\ \vdots \\ C_{ij}(M+G) \end{bmatrix}, \quad \tilde{C} = \begin{bmatrix} C_{11} & \dots & C_{1N} \\ \vdots & \vdots & \vdots \\ C_{N1} & \dots & C_{NN} \end{bmatrix}$$
(22)

where C_{ij} are the M + G + 1 row and M + 1 column matrices and \tilde{C} is the N(M + G + 1) row and N(M + 1) column matrix, we obtain

$$\left[C_{i1}, C_{i2}, \dots, C_{iN}\right] \left[r\text{-th } M+1 \text{ columns of } YY^T\right] = Y_{ri}^T$$
(23)

for i, r = 1, 2, ..., N. Combining (23) for each i and r we obtain (like (20) from (19)) the following system of linear equations

$$\widetilde{C}\left(YY^T\right) = Y^T \tag{24}$$

Matrix \tilde{C} is the solution of (24). Successive rows $C_{ij}(h)$ of submatrices $C_{ij}(i, j = 1, ..., N)$ of matrix \tilde{C} contain the solutions for successive values h (h = 0, 1, ..., M + G). The solution of approximation problem is the solution for such value h which gives least value of the criterion Q_h . To find Q_h we compute $E_{ij}(h)$ (i, j = 1, 2, ..., N) for each h. Defining

$$E_{ij} = \begin{bmatrix} E_{ij}(0) \\ E_{ij}(1) \\ \vdots \\ E_{ij}(M+G) \end{bmatrix}, \quad \widetilde{E} = \begin{bmatrix} E_{11} & \dots & E_{1N} \\ \vdots & \vdots & \vdots \\ E_{N1} & \dots & E_{NN} \end{bmatrix}$$
(25a)

$$\mathcal{E}_{ij} = \begin{bmatrix} \mathcal{E}_{ij}(0) \\ \mathcal{E}_{ij}(1) \\ \vdots \\ \mathcal{E}_{ij}(M+G) \end{bmatrix}, \quad \widetilde{\mathcal{E}} = \begin{bmatrix} \mathcal{E}_{11} & \dots & \mathcal{E}_{1N} \\ \vdots & \vdots & \vdots \\ \mathcal{E}_{N1} & \dots & \mathcal{E}_{NN} \end{bmatrix}$$
(25b)

where E_{ij} are the M + G + 1 row and M + G + 1 column matrices and \tilde{E} is the N(M + G + 1) row and N(M + G + 1) column matrix. If we notice from (13), (20a) and (25b) that $\tilde{\mathcal{E}} = I_{N(M+G+1)}$, then we obtain from (21) and (25)

$$E = CY - I_{N(M+G+1)}.$$
 (26)

The inverse polynomial matrix approximation problem can be solved using the following algorithm.

Algorithm:

- 1. Choose degree M of polynomials of approximating matrix (recommended $M \geq 3G$ should be greater for greater computational precision).
- 2. Create matrix Y given by (18) and containing convolution matrices Y_{ij} (i, j = 1, 2, ..., N). Each matrix Y_{ij} contains M+1 rows. Each p-th (p = 1, 2, ..., M+1) row contains successive M+G+1 components: p-1 nulls, G+1 values $y_0^{ij}, ..., y_G^{ij}, M-p+1$ nulls. These rows are as follows

$$\left[0, ..., 0, y_0^{ij}, y_1^{ij}, ..., y_G^{ij}, 0, ..., 0\right]$$

3. Compute N(M + G + 1) row and N(M + 1) column matrix \tilde{C} (22) solving the system of N(M + 1) linear equations, each with N(M + G + 1) right hand side free values

$$\left(YY^{T}\right)^{T}\left(\widetilde{C}\right)^{T} = Y \tag{27}$$

Successive rows $C_{ij}(h)$ (h = 0, 1, ..., M + G) of submatrices C_{ij} of matrix \tilde{C} contain solutions for successive values h.

4. Compute matrix \tilde{E} from

$$\widetilde{E} = \widetilde{C}Y - I_{N(M+G+1)}$$

where $I_{N(M+G+1)}$ is the N(M+G+1) row and column identity matrix.

5. Compute mean square errors $\delta_i^2(h)$ (h = 0, 1, ..., M + G) for each i = 1, 2, ..., N from

$$\delta_{i}^{2}(h) = \sum_{k=1}^{N(M+G+1)} \left(\tilde{e}_{rk}\right)^{2}$$
(28)

where \tilde{e}_{rk} are the components of matrix \tilde{E} and

$$r = (i-1)(M+G+1) + h$$

- 6. Compute the value of criterion Q_h from (8) for each value h and choose h^* minimizing this criterion, that is, $Q_{h^*} = \min\{Q_h\}$.
- 7. Rows

$$C_{ij}^{h^*} = \left[c_0^{ij}(h^*), c_1^{ij}(h^*), ..., c_M^{ij}(h^*)\right] \qquad (i, j = 1, 2, ..., N)$$

contain the searched solution of inverse matrix approximation problem.

8. If the obtained value of criterion Q_{h^*} is too large, then increase M by G and go to point 2 of this algorithm unless last increase of M gives unsatisfactory decrease of criterion value.

6. An Example

Let N = 2, G = 1, M = 5, and

$$Y(z) = \begin{bmatrix} 1 - 0.5z^{-1} & 0.5 + 0.125z^{-1} \\ 0.5 - 0.125z^{-1} & 1 + 0.5z^{-1} \end{bmatrix}$$

Matrix \widetilde{C} is the 14 row and 12 column matrix, and matrix \widetilde{E} is the 14 row and 14 column matrix. The obtained \widetilde{C} contains

$$C_{11} = \begin{bmatrix} 1.332 & 0.665 & 0.413 & 0.201 & 0.118 & 0.042 \\ -0.001 & 1.329 & 0.660 & 0.402 & 0.188 & 0.081 \\ -0.003 & -0.006 & 1.321 & 0.644 & 0.376 & 0.135 \\ -0.004 & -0.014 & -0.020 & 1.285 & 0.602 & 0.261 \\ -0.009 & -0.021 & -0.040 & -0.072 & 1.203 & 0.433 \\ -0.014 & -0.044 & -0.063 & -0.155 & -0.208 & 0.834 \\ -0.029 & -0.066 & -0.127 & -0.232 & -0.416 & -0.748 \end{bmatrix}$$

$$C_{12} = \begin{bmatrix} -0.666 & -0.166 & -0.206 & -0.049 & -0.059 & -0.005 \\ 0.000 & -0.664 & -0.165 & -0.198 & -0.047 & -0.031 \\ 0.002 & 0.003 & -0.660 & -0.155 & -0.188 & -0.015 \\ 0.002 & 0.010 & 0.005 & -0.633 & -0.150 & -0.100 \\ 0.005 & 0.010 & 0.020 & 0.037 & -0.602 & -0.048 \\ 0.005 & 0.030 & 0.016 & 0.107 & 0.052 & -0.321 \\ 0.018 & 0.033 & 0.064 & 0.117 & 0.208 & 0.378 \end{bmatrix}$$

]	-0.666	0.166	-0.206	0.049	-0.059	0.005
	-0.000	-0.664	0.165	-0.198	0.047	-0.031
	0.002	-0.003	-0.660	0.155	-0.188	0.015
$C_{21} =$	-0.002	0.010	-0.005	-0.633	0.150	-0.100
	0.005	-0.010	0.020,	-0.037	-0.602	0.048
	-0.005	0.030	-0.016	0.107	-0.052	-0.321
	0.018	-0.033	0.064	-0.117	0.208	-0.378
	- 1.332	-0.665	0.413	-0.201	0.118	-0.042
	0.001	1.329	-0.660	0.402	-0.188	0.081
	-0.003	0.006	1.321	-0.644	0.376	-0.135
$C_{22} =$	0.004	-0.014	0.020	1.285	-0.602	0.261
	-0.009	0.021	-0.040	0.072	1.203	-0.433
	0.014	-0.044	0.063	-0.155	0.208	0.834
	-0.029	0.066	-0.127	0.232	-0.416	0.748
	-					

The values of Q_h are given in Table 1.

Tab. 1.

Q_0	Q_1	Q_2	Q_3	Q_4	Q_5	Q_6
0.0013	0.0044	0.0128	0.0448	0.1315	0.4585	1.3468

As we can see, the least value of criterion Q_h is for h = 0. The solution is then given by first rows of matrices C_{ij} . Then $C(z) = \overline{C}(z, 0)$ contains polynomials

$$c_{11}(z) = 1.332 + 0.665z^{-1} + 0.413z^{-2} + 0.201z^{-3} + 0.118z^{-4} + 0.042z^{-5}$$

$$c_{12}(z) = -0.666 - 0.166z^{-1} - 0.206z^{-2} - 0.049z^{-3} - 0.059z^{-4} - 0.005z^{-5}$$

$$c_{21}(z) = -0.666 + 0.166z^{-1} - 0.206z^{-2} + 0.049z^{-3} - 0.059z^{-4} + 0.005z^{-5}$$

$$c_{22}(z) = 1.332 - 0.665z^{-1} + 0.413z^{-2} - 0.201z^{-3} + 0.118z^{-4} - 0.042z^{-5}$$

Matrix
$$E(z) = \overline{E}(z,0) = C(z)Y(z) - I_2$$
 contains polynomials
 $e_{11}(z) = -0.0006 - 0.001z^{-1} - 0.002z^{-2} - 0.004z^{-3} - 0.006z^{-4} - 0.01z^{-5} - 0.02z^{-6}$
 $e_{12}(z) = 0.0002z^{-1} + 0.0003z^{-2} + 0.0005z^{-3} + 0.0009z^{-4} + 0.0017z^{-5} + 0.003z^{-6}$
 $e_{21}(z) = -0.0002z^{-1} + 0.0003z^{-2} - 0.0005z^{-3} + 0.0009z^{-4} - 0.0017z^{-5} + 0.003z^{-6}$
 $e_{22}(z) = -0.0006 + 0.001z^{-1} - 0.002z^{-2} + 0.004z^{-3} - 0.006z^{-4} + 0.01z^{-5} - 0.02z^{-6}$

7. Final Remarks

The algorithm presented in this paper is simple and suitable for computations. The optimal solution is chosen from among M + G + 1 solutions of N(M + 1) linear equations of system (27), with N(M + G + 1) right hand side free values. The solution is then obtained using the well known methods after the matrices Y, \tilde{C} and \tilde{E} are defined. For the problem described above, the computational algorithm needs memory for $N^2(M + 1)(2M + G + 2)$ real numbers. For instance, assuming N = 9, G = 9, M = 23, the required memory is $384 \, kB$, if one real number occupies 4 bytes. The number of equations is 216 in this case.

In the example described in this paper, the optimal solution is for h = 0. This is true when all the polynomials in Y(z) have maximal module coefficients y_0^{ij} (i, j = 1, 2, ..., N) standing by z^0 . Many real examples of multichannel interference are described by matrix Y(z) whose polynomials have maximal module coefficients y_k^{ij} standing by common z^{-k} , mostly with k > 0. In this case the optimum solution will be expected for h > 0.

References

- Chen H. and Guo L. (1990): Continuous-time stochastic adaptive tracking-robustness and asymptotic properties. SIAM J. Control and Optimization, v.28, No.3, pp.513-527.
- Clark A.P. (1977): Advanced Data-Transmission Systems. London: Pentech Press.
- Dąbrowski A. (1982): Detection of a data-transmission signals in the presence of intersymbol interference and noise. - Chapter 4, Edited by Z. Baran: Principles of Data-Transmission.
 Warszawa: WKL, (in Polish).
 - Dąbrowski A. (1979): Signal equalization. Chapter 5, Edited by Z. Baran: Problems of Data-Transmission. Warszawa: WKL, (in Polish).
 - Grasselli O.M. and Tornmbe A. (1992): On obtaining a realization of a polynomial matrix description of a system. — IEEE Trans. Automatic Control, v.37, No.6, pp.852-856.
 - Kisilewicz J. (1990): Multidimensional interference equalizer giving the minimum meansquare error. — Electronics and Telecommunications Quarterly, v.36, No.3-4, pp.663-671 (in Polish).
 - Kisilewicz J. (1991): Data transmission fidelity improvement problems in computer networks. — Sci. Papers of the Inst. of Control and Systems Engng., Technical University of Wrocław, No.9, Monographs No.6, PWr, Wrocław (in Polish).
 - Kisilewicz J. (1992): An algorithm for design multichannel equalizer giving the minimum meansquare error. — Electronics and Telecommunications Quarterly, v.38, No.1, pp.67-76 (in Polish).
 - Mutoh Y. and Nikiforuk P.N. (1992): Inversed interactoring and triangularizing with an arbitrary pole assignment using the state feedback. — IEEE Trans. Automatic Control, v.37, No.5, pp.630-633.
 - Nyman P.O. (1991): Equalizer principle of \mathcal{H}^{∞} optimal control and its application to superoptimization. — Int. J. Control, v.54, No.2, pp.393-415.

- Özbay H. and Tannenbaum A. (1990): A skew Toeplitz approach to the H[∞] optimal control of multivariable distributed systems. — SIAM J. Control and Optimization, v.28, No.3, pp.653-670.
- Smith M.C. (1990): Well-posedness of \mathcal{H}^{∞} optimal control problems. SIAM J. Control and Optimization, v.28, No.2, pp.342-358.
- Zhou K. and Gu G. (1992): Robust stability of multivariable systems with both real parametric and norm bounded uncertainties. — IEEE Trans. Automatic Control, v.37, No.10, pp.1533-1537.

Received March 12, 1993 Revised December 15, 1993