IEEE Pacific Rim Conference on Communications, Computers and Signal Processing, May 9-10, 1991

On the Minimization of System Matrix for Two-Dimensional Digital Filters

Q.-H. Meng, A. Hasan, W.-S. Lu, and V. K. Bhargava

Department of Electrical and Computer Engineering
University of Victoria
Victoria, B.C., Canada V8W 3P6

Abstract

In this paper, an optimization algorithm is developed to minimize the 2-norm of the transformed system matrix for a 2-D state-space digital filter. The method is actually a two-loop Fletcher-Reeves type method, where the line search is carried out using a combination of the golden section ratio search and the method of false position. In general, the proposed method can be used to solve nonquadratic optimization problems where analytic version of the objective function is not available.

I. Introduction

Two-dimensional digital filters have been widely used in many areas such as image processing and geophysical data processing. Desirable properties for digital filters include stability, lower parameter sensitivity, and freedom from limit circles.

It has been known that for a stable 2-D filter satisfying the constant 2-D Lyapunov equation, there exists a 2-D state transformation such that the 2-norm of the transformed system matrix is strictly less than one. It is also known that a stable 2-D state-space digital filter with the norm of its system matrix less than one has low parameter sensitivity as well as the freedom from limit circles. On the other hand, finding such a state transformation is numerically not straightforward [1].

In this paper, an optimization method is developed to minimize the 2-norm of the transformed system matrix for a stable 2-D state-space digital filter. The method is actually a two-loop Fletcher-Reeves type method, where the line search is carried out using a combination of the golden section ratio search and the method of false position. In general, the proposed method can be used to solve nonquadratic optimization problems where analytic version of the objective function is not available.

The problem statement is presented in next section. In section III, the optimization algorithm is described in detail. In particular, two numerical differentiation methods are discussed and compared each other, and a line search scheme suitable for our optimization problem is proposed. An example studied in [1][2] is used in section IV to illustrate the proposed algorithm.

II. Problem Statement

A single-input, single-output 2-D digital filter (or a discrete system) can be represented by the Roesser model [1][2] as

$$\begin{aligned}\begin{bmatrix} h(m+1,n) \\ v(m,n+1)\end{bmatrix} &= \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4\end{bmatrix}\begin{bmatrix} h(m,n) \\ v(m,n)\end{bmatrix} + \begin{bmatrix} b_1 \\ b_2\end{bmatrix} u(m,n) \\ &\equiv A\begin{bmatrix} h(m,n) \\ v(m,n)\end{bmatrix} + bu(m,n) \qquad (1) \\ y(m,n) &= [c_1 \;\; c_2]\begin{bmatrix} h(m,n) \\ v(m,n)\end{bmatrix} \equiv c\begin{bmatrix} h(m,n) \\ v(m,n)\end{bmatrix}\end{aligned}$$

where h and v are real vectors of dimensions M and N, respectively. The system described by (1) is in this paper assumed to be stable, and is referred to as a state-space digital filter, which admits equivalent realizations by block-diagonal similarity transformation T where

$$T = \begin{bmatrix} T_1 & 0 \\ 0 & T_2\end{bmatrix} \equiv T_1 \oplus T_2 \qquad (2)$$

The minimum norm of the state-space realization is achieved by minimizing the 2-norm of the system matrix TAT^{-1}, i.e.

$$\mu = \min_T \|TAT^{-1}\| \qquad (3)$$

Although finding a nonsingular minimizer $T = \tilde{T}$ such that

$$\mu = \|\tilde{T}A\tilde{T}^{-1}\|$$

represents a fairly complex problem [2], an approximate solution of the minimization problem is always possible if matrices T_1 and T_2 are further chosen as upper triangular matrices with positive diagonal elements [1][2]. It is known [2] that if the state-space filter satisfies the 2-D constant Lyapunov equation, then the approximate solution of the optimization problem (3) will lead to a state-space realization

$$\tilde{A} = TAT^{-1}, \quad \tilde{b} = Tb, \quad \tilde{c} = cT^{-1}$$

such that $\|\tilde{A}\| < 1$. Such a realization is free of limit circles and have fairly low parameter sensitivity [4][5].

In what follows, we concentrate on the optimization problem defined in (3).

CH2954-6/91-0000-396

III. The Optimization Method

The above mentioned optimization problem can be stated as

$$\text{Minimize } f(X)$$

where the objective function $f(X)$ is given by

$$f(X) = \|T(X)AT^{-1}(X)\|_2 \tag{4}$$

with

$$X = [x_{1i}, x_{1j}]^T, \quad \begin{aligned} i &= 1,2,\ldots,M(M-1)/2 \\ j &= 1,2,\ldots,N(N-1)/2 \end{aligned}$$

$$T(X) = T_1(X) \oplus T_2(X) = \begin{bmatrix} T_1(X) & 0_{M\times N} \\ 0_{N\times M} & T_2(X) \end{bmatrix}_{(M+N)\times(M+N)}$$

and

$$T_1(X) = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1M} \\ 0 & x_{1(M+1)} & \cdots & x_{1(2M-1)} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & x_{1(M(M-1)/2)} \end{bmatrix}_{M\times M}$$

$$T_2(X) = \begin{bmatrix} x_{21} & x_{22} & \cdots & x_{2N} \\ 0 & x_{2(N+1)} & \cdots & x_{2(2N-1)} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & x_{2(N(N-1)/2)} \end{bmatrix}_{N\times N}$$

A two-loop Fletcher-Reeves type method is chosen for the optimization, the reason being that it is globally convergent and requires only the first order derivative of the object function.

Derivative Approaximation

Obviously the objective function (4) is a nonquadratic function and the analytic expression can not be found in general. So the derivative of the object function can only be approximated using numerical derivatives.

The following two formulas are considered in this paper:

(i) **Forward-difference Formula**

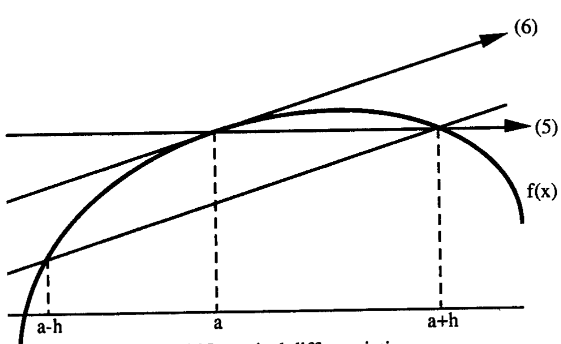

Fig.1 Numerical differenciation

$$f'(x_0) \approx f[x_0, x_0+\delta] = \frac{f(x_0+\delta)-f(x_0)}{\delta} \tag{5}$$

with the approximation error

$$E(f) = -\frac{1}{2}\delta f''(\eta)$$

for x_0, $\eta \in (x_1, x_2)$ and a sufficiently small perturbation δ.

(ii) **Central-difference Formula**

$$f'(x_0) \approx f[x_0-\delta, x_0+\delta] = \frac{f(x_0+\delta)-f(x_0-\delta)}{2\delta} \tag{6}$$

with the approximation error

$$E(f) = -\frac{\delta_2}{6}\delta f'''(\eta)$$

for x_0, $\eta \in (x_1, x_2)$ and a sufficiently small perturbation δ.

The above two methods are illustrated in Fig.1. It follows that the first method appears to be more dependent on the value at point x_0 and requires very small δ to achieve an accurate approximation while the second one is more dependent on the values in the neighborhood of x_0. It is found that both methods give very good results but the first one results in a slightly better μ in the optimization problem (3) if the perturbation δ is chosen within $[10^{-3}, 10^{-5}]$. Therefore, the first method is adopted in our optimization algorithm.

Line Search

To carry out the line search, methods such as cubic fit, quadratic fit, three point pattern, modified three point pattern, and method of false position have been tried but they all fail to work. This is primarily because the function involved in the line search is usually non-convex with multiple local minimums, as is shown in Fig.2.

It is found that golden section ratio (GSR) search [3] can be used during the first stage of line search. For example, when the GSR is applied to the function in Fig.2, we obtained $a = 0.75386$ as its minimum point, which is obviously a good approximation of the global minimum point. However, as is shown in Fig.3, $a = 0.75386$ is still "far" away from the true global minimum point of the function. It is therefore necessary to combine GSR with another line search method to do the fine tuning. The method of our choice during the second state of line search is the method of false position [3]. Fig.4 shows a satisfactory result $a = 0.614915$ which is obtained by using the method of false position after the GSR search.

Optimization Algorithm

The above mentioned Fletcher-Reeves method with numerical derivative and a modified line search method forms the following optimization algorithm:

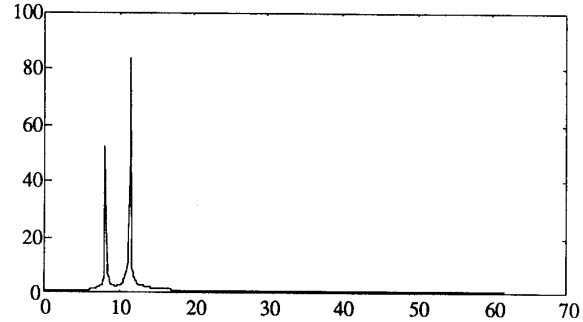

Fig.2 $f(a)$ in [0.01*a 100*a] with a=0.614915

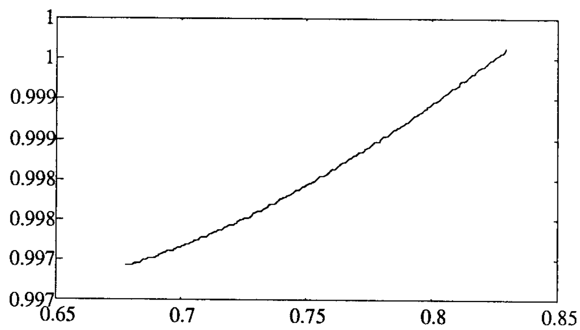

Fig.3 $f(a)$ in [0.9*a 1.1*a] with a=0.75386 (GSR only)

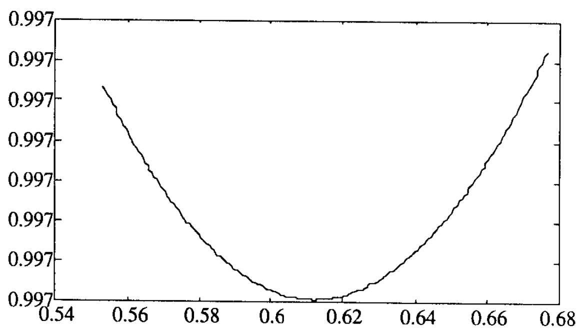

Fig.4 $f(a)$ in [0.9*a 1.1*a] with a=0.614915 (GSR+MFP)

Algorithm

Step 1. Choose initial value $X_0 \in R^n$, where $n = M + N$;

Step 2. Using numerical differentiation method (5) to obtain the gradient

$$g_0 \approx \nabla^T f(X_0)$$

and then set

$$d_0 = -g_0$$

Step 3. Golden section ratio line search

3.1 Set initial left side point of the uncertainty interval as $\alpha_l = 0$ and the initial right side point $\alpha_r = \epsilon$, where ϵ is an arbitrary positive number with initial width of uncertainty $h = \alpha_r - \alpha_l = \epsilon$;

3.2 Evaluate

$$\begin{aligned} l(\alpha_l) &= f(X_k + \alpha_l d_k) \\ l(\alpha_r) &= f(X_k + \alpha_r d_k) \end{aligned}$$

3.3 If $l(\alpha_r) < l(\alpha_l)$, then $\alpha_l = \alpha_r$, $h = \gamma h$, where $\gamma = 1.618$, $\alpha_r = \alpha_r + h$. Go back to 3.2. Otherwise, set the uncertainty interval as (α_l, α_r) and $h_1 = \alpha_r - \alpha_l$;

3.4 Set the width of uncertainty after r iterations at

$$h_r = 0.618^{r-1} h_1$$

and evaluate the updated α_l and α_r until $h_r < \nu$, where ν is predefined error of the width of uncertainty.

3.5 Output the optimal α_l and α_r.

Step 4. Using the method of false position to find the minimum point $\alpha_{min} \in [\alpha_l \ \ \alpha_r]$ and set $\alpha_k = \alpha_{min}$;

Step 5.

$$X_{k+1} = X_k + \alpha_k d_k$$

and

$$g_{k+1} \approx \nabla^T f(X_{k+1})$$

Step 6.

$$\begin{aligned} \beta_k &= \frac{g_{k+1}^T g_{k+1}}{g_k^T g_k} \\ d_{k+1} &= -g_{k+1} + \beta_k d_k \end{aligned}$$

Step 6. If $\|g_{k+1}\| < \xi$, where ξ is prescribed tolerance, then output X_{k+1} as the solution, otherwise if $k < n - 1$, update k and back to Step 3, else replace X_0 by X_n and go back to Step 1.

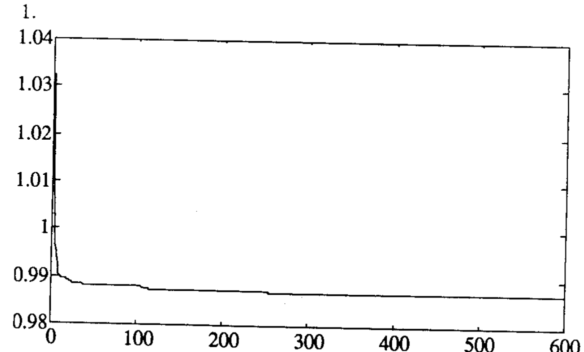

Fig.5 Convergence of the objective function

IV. An Illustrative Example

The example discussed in [1][2] is used to illustrate the proposed optimization algorithm. Consider a 2-D state-space digital filter with system matrix given by

$$A_1 = \begin{bmatrix} 1.88899 & -0.91219 & -1 & 0 \\ 1 & 0 & 0 & 0 \\ 0.02771 & -0.0258 & 1.88899 & 1 \\ -0.0258 & 0.02431 & -0.91219 & 0 \end{bmatrix}$$

In [1], it is found that $\|A_1\|_2 = 0.7609$ (Actually it should be $\sqrt{7.609} = 2.7584$) and matrix A_1 is then transformed to a realization with a smaller norm for the system matrix $A_2 = TA_1T^{-1}$ given by

$$A_2 = \begin{bmatrix} 0.94467 & -0.10962 & -0.01808 & 0.13838 \\ 0.18354 & 0.94432 & 0 & 0 \\ 0.03507 & 0.09372 & 0.93169 & 0.16216 \\ -0.19567 & -0.00221 & -0.12508 & 0.9573 \end{bmatrix}$$

The 2-norm of this matrix $\|A_2\|_2 = 1.066$ (Again it should be $\sqrt{1.066} = 1.0324$). The optimization algorithm proposed in this paper can be used to perform 2-norm minimization for either the transformed matrix A_2 or the original matrix A_1. As indicated in [2], it is very hard to get the exact optimal solution of (3), but an approximate solution to the optimization problem always possible. The convergence of the object function to its optimal value is shown in Fig.5. It appears that the proposed algorithm will run for for an indefinite period of time to achieve its optimal value as the decreasing rate approaching to zero. The results obtained using the proposed algorithm are compared with those of [1] as follows:

	Proposed Algorithm	**Algorithm of [1]**
$\|\tilde{A}_1\|_2$	0.98820	N/A
$\|\tilde{A}_2\|_2$	0.98667	$\sqrt{0.982} = 0.99079$

where

$$\tilde{A}_2 = \tilde{T}_2 A_2 \tilde{T}_2^{-1} = \begin{bmatrix} 0.9607 & -0.1496 & -0.0241 & 0.1641 \\ 0.1362 & 0.9283 & 0 & 0 \\ 0.0209 & 0.0923 & 0.9271 & 0.1447 \\ -0.1658 & 0.0169 & -0.1411 & 0.9619 \end{bmatrix}$$

with

$$\tilde{T}_2 = \begin{bmatrix} 1.1955 & 0.1042 & 0 & 0 \\ 0 & 0.8873 & 0 & 0 \\ 0 & 0 & 0.8979 & 0.0333 \\ 0 & 0 & 0 & 1.0130 \end{bmatrix}$$

Note the condition number of matrix $\tilde{T}_2$ is $K(\tilde{T}_2) = 1.3696$. Thus the upper bound of the parameter variations of the system matrix as defined in [2] is

$$\frac{1-\mu}{K(\tilde{T}_2)} = \frac{1 - 0.986675}{1.3696} = 9.73 \times 10^{-3}$$

which is significantly higher than the upper bound 1.651×10^{-5} given in [1].

This algorithm has also been tested to solve a minimum norm realization problem of an 8×8 system matrix with a transformation matrix $T = T_{1(2\times2)} \oplus T_{2(6\times6)}$. The results show that the algorithm proposed in this paper is very efficient. In general, we believe that the method developed in this paper could be useful in solving a large class of nonquadratic optimization problems where the analytic version of the objective function is not available.

V. Conclusions

In this paper, an optimization algorithm for the minimization of a nonquadratic and numerical objective function has been proposed. The proposed algorithm uses a Fletcher-Reeves type method where a forward-difference numerical differentiation and a two stage line search algorithm based on the combination of the golden section ratio search and the method of false position are adopted. The proposed algorithm is fairly efficient and can be used to solve nonquadratic optimization problems where analytic version of the objective function is not available.

References

[1] J. H. Lodge and M. M. Fahmy, "Stability and overflow oscillations in 2-D state-space digital filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.29, pp. 1161-1171, 1981.

[2] W.-S. Lu, A. Antoniou, and P. Agathoklis, "Stability of 2-D digital filter under parameter variation," *IEEE Trans. Circuits and Systems*, vol.33, pp. 476-482, May 1986.

[3] D. G. Luenberger, *Linear and Nonlinear Programming*, New York: Addison-Wesley, 1984.

[4] W.-S. Lu and A. Antoniou, "Synthesis of 2-D State-space fixed-point digital filter structures with minimum roundoff noise," *IEEE Trans. Circuits and Systems*, vol.33, pp.965-973, Oct. 1986.

[5] M. Kanamata and T. Higuchi, "Synthesis of 2-D separable denominator digital filters with minimum roundoff noise and no overflow oscillations," Int. Symp. Circuits Syst., Kyoto, pp.1087-1091, June 1985.