LEAST PTH OPTIMIZATION FOR THE DESIGN OF 1-D FILTERS WITH ARBITRARY AMPLITUDE AND PHASE RESPONSES

W.-S. Lu, Y. Cui, R. L. Kirlin

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

ABSTRACT

This paper presents closed-form formulas for the evaluation of the gradient vector and the Hessian matrix of an IIR transfer function. These formulas are then used in a modified Newton optimization algorithm for the optimal design of IIR digital filters in the L_p-norm sense. The method is expected to be useful in designing digital filters with arbitrary amplitude and phase responses. Issues concerning the selection of initial point and stability of the filter designed are addressed. Design examples are included to illustrate the use of the proposed closed-form formulas.

I. Introduction

Recursive digital filters have been extensively used in a variety of applications in digital-signal-processing (DSP) related areas where high selectivity and efficeint processing of the input signals are required. For phase-sensitive applications such as image processing tasks, phase response of the filter is often an important part of design specifications. Moreover, both phase and amplitude responses in a given design problem could be fairly arbitrary [1]. Early work dealing with filter design problems include references [2]-[5]. References [2] and [3] investigated the design issues in the least-squares sense in the time domain, while reference [4] dealt with the same issues in the sense of L_p-norm in the frequency domain. The approach taken in [4] to design an IIR 1-D filter with specified amplitude and phase specifications was first to design a transfer function that satisfies the given amplitude response and then design an all-pass filter to meet the phase specification, while the approach taken in this paper is to perform a least L_p-norm optimization for an objective function that combines both amplitude and phase specifications. From the SVD-design point of view [1], this is a natural way to formulate our design problem since the generalized singular vectors are in general complex vectors that contain both amplitude and phase information.

In this paper, we present for a given L_p-norm type objective function two sets of closed-form formulas that can be used for evaluating the gradient and Hessian of the objective function, respectively. The availability of these formulas makes it possible to carry out the minimization efficiently by using a variety of established unconstrained optimization techniques [6][7]. Issues concerning the selection of initial point and stability of the filter designed are addressed. Design examples are included to illustrate the use of the proposed close-form formulas.

II. The Objective Function and Its Gradient

2.1 The Objective Function

Let $H_d(\omega)$ be the desired frequency response over a frequency range Ω which is a compact subset of $[0, \pi]$, and $H(\omega)$ be the mth-order transfer function given by

$$H(\omega) = h_0 \frac{1 + b^T f}{1 + a^T f} \quad (1)$$

$$f = [z^{-1} \ \dots \ z^{-m}]^T \quad (2)$$

$$a = [a_1 \ \dots \ a_m]^T \quad (3)$$

$$b = [b_1 \ \dots \ b_m]^T \quad (4)$$

and $z = e^{j\omega}$ with ω the normalized frequency, i.e., $0 \le \omega \le \pi$. The filter design problem is to find real-valued M-dimensional vector ($M = 2m + 1$)

$$x = [h_0 \ \ a_1 \ \ \dots \ \ a_m \ \ b_1 \ \ \dots \ \ b_m]^T \quad (5)$$

that minimizes the L_p–norm objective function

$$E(x) = \int_\Omega |H(\omega) - H_d(\omega)|^p d\omega \quad (6)$$

where $p \ge 2$ is an even integer and that the corresponding filter (1) is stable.

2.2 The Gradient

Writing

$$e(\omega) = H(\omega) - H_d(\omega) \quad (7)$$

$$|e(\omega)|^p = [\bar{e}(\omega)e(\omega)]^{p/2} \quad (8)$$

we have

$$\frac{\partial |e(\omega)|^p}{\partial x} = p|e(\omega)|^{p-2} \mathrm{Re}\left[\bar{e}(\omega)\frac{\partial H(\omega)}{\partial x}\right] \quad (9)$$

Therefore the gradient of $E(x)$ is given by

$$\nabla E(x) = p \int_\Omega |e(\omega)|^{p-2} \mathrm{Re}[\bar{e}(\omega) \nabla H(\omega)] d\omega \quad (10)$$

0-7803-0946-4/93 $3.00

where $\nabla H(\omega)$ denotes the gradient of $H(\omega)$, i.e.

$$\nabla H(\omega) = \left[(\frac{\partial H}{\partial h_0})^T \ (\frac{\partial H}{\partial a})^T \ (\frac{\partial H}{\partial b})^T\right]^T \qquad (11)$$

From (1) it follows that

$$\begin{aligned}
\frac{\partial H}{\partial h_0} &= \frac{1+b^Tf}{1+a^Tf} = \frac{H(\omega)}{h_0} \\
\frac{\partial H}{\partial a} &= -h_0\frac{1+b^Tf}{(1+a^Tf)^2} = -\frac{H(\omega)f}{1+a^Tf} \\
\frac{\partial H}{\partial b} &= h_0\frac{1}{1+a^Tf} = \frac{H(\omega)f}{1+b^Tf}
\end{aligned}$$

$$\nabla H(\omega) = H(\omega)\left[\frac{1}{h_0} \quad -\frac{f^T}{1+a^Tf} \quad \frac{f^T}{1+b^Tf}\right]^T \qquad (12)$$

Eqns. (7), (10) and (12) provide a set of closed-form formulas for the evaluation of $\nabla E(\omega)$. As a special case, we obtain the gradient of the least-squares objective function (i.e. function $E(x)$ in Eq.(6) with $p = 2$) as

$$\nabla E(x) = 2\int_\Omega \mathrm{Re}[\bar{e}(\omega)\nabla H(\omega)]d\omega \qquad (13)$$

III. The Hessian Matrix

Assume $p \geq 4$ (the general case) and re-write (9) as

$$\nabla |e(\omega)|^p = \frac{p}{2}|e(\omega)|^{p-2}(\bar{e}\nabla H + e\nabla\bar{H}) \qquad (14)$$

It can readily be shown that the Hessian of $|e(\omega)|^p$ is given by

$$\begin{aligned}
\nabla^2|e(\omega)|^p = p(p-2)|e(\omega)|^{p-4}\mathrm{Re}[\bar{e}(\omega)\nabla H]\mathrm{Re}[\bar{e}(\omega)\nabla^T H] \\
+p|e(\omega)|^{p-2}\mathrm{Re}[\bar{e}(\omega)\nabla^2 H(\omega) + (\nabla\bar{H})(\nabla H)^T]
\end{aligned}$$

where $\nabla H(\omega)$ is given by (11), $\nabla\bar{H}$ is the complex conjugate of ∇H,

$$\nabla^2 H(\omega) = \begin{bmatrix}
\frac{\partial^2 H}{\partial h_0^2} & \frac{\partial}{\partial h_0}(\frac{\partial H}{\partial a})^T & \frac{\partial}{\partial h_0}(\frac{\partial H}{\partial b})^T \\
\frac{\partial}{\partial a}(\frac{\partial H}{\partial h_0}) & \frac{\partial}{\partial a}(\frac{\partial H}{\partial a})^T & \frac{\partial}{\partial a}(\frac{\partial H}{\partial b})^T \\
\frac{\partial}{\partial b}(\frac{\partial H}{\partial h_0}) & \frac{\partial}{\partial b}(\frac{\partial H}{\partial a})^T & \frac{\partial}{\partial b}(\frac{\partial H}{\partial b})^T
\end{bmatrix} \qquad (15)$$

The submatrices in (15) can be calculated as follows

$$\begin{aligned}
\frac{\partial^2 H}{\partial h_0^2} &= 0 \\
\frac{\partial}{\partial a}(\frac{\partial H}{\partial h_0}) &= -\frac{(1+b^Tf)}{(1+a^Tf)^2}f = -\frac{H(\omega)f}{h_0(1+a^Tf)} \\
\frac{\partial}{\partial b}(\frac{\partial H}{\partial h_0}) &= \frac{1}{(1+a^Tf)}f = \frac{H(\omega)f}{h_0(1+b^Tf)} \\
\frac{\partial}{\partial a}(\frac{\partial H}{\partial a})^T &= 2h_0\frac{(1+b^Tf)}{(1+a^Tf)^3}ff^T = \frac{2H(\omega)F}{h_0(1+a^Tf)^2} \\
\frac{\partial}{\partial b}(\frac{\partial H}{\partial a})^T &= -\frac{h_0 F}{(1+a^Tf)} = -\frac{H(\omega)F}{(1+a^Tf)(1+b^Tf)} \\
\frac{\partial}{\partial b}(\frac{\partial H}{\partial b})^T &= 0 \\
F &= ff^T
\end{aligned}$$

For the case of $p = 2$,

$$\nabla^2|e(\omega)|^2 = 2\mathrm{Re}[\bar{e}(\omega)\nabla^2 H(\omega) + (\nabla\bar{H})(\nabla H)^T]$$

The Hessian matrix of $E(x)$ is now given by

$$\nabla^2 E(x) = \int_\Omega \nabla^2(|e(\omega)|^p)d\omega \qquad (16)$$

IV. Selection of A Good Initial Point

It has been known that the objective function $E(x)$ defined by (6) has many local minimum points and it is of critical importance to select a "good" initial point in an iterative optimization procedure. By good initial point we mean a point x_0 in the $(2m+1)$−dimensional parameter space that corresponds to a *stable* transfer function $H(\omega)$ and that $E(x_0)$ is not exceedingly large. An effective approach to finding such a good initial point is to first design an (possibly high-order) FIR filter using a conventional method such as the window method to approximate the frequency response of the desired filter and to apply the well-known balanced approximation technique [8] to obtain a $(2m+1)$th-order IIR filter. As the original (FIR) filter is always stable and the balanced approximation preserves its stability, the IIR filter obtained is stable, and it has a good amplitude response.

V. The Stability Issue

Although the initial point corresponds to a stable transfer function, most of unconstrained optimization algorithms do not include a mechanism to ensure that the minimum point corresponds to a stable IIR filter. It is also noted that the stabilization technique using a reciprocal substitution for the unstable factors in the filter obtained does not apply if phase response is a part of the design specifications because the reciprocal substitution changes the phase response of the filter.

A simple yet effective approach to incorporating a stability monitoring mechanism into an optimization algorithm is that one checks the stability of the filter obtained after each iteration and, if the filter is found unstable, reduces the adjustment parameter vector Δx, say by half in length to $\Delta x/2$. The algorithm continues if the filter is then found stable, otherwise reduce the length of the new adjustment vector again, and so on.

VI. A Case Study

Consider the problem of designing a lowpass IIR filter of order 8 with normalized cutoff frequency $\omega_c = \pi/2$ rad/s (the normalized frequency is the one with the Nyquist frequency= π rad/s) and a 0.1π transition band (i.e. $\omega_p = 0.45\pi, \omega_a = 0.55\pi$). As both gradient and Hessian formulas are available in closed form, the modified Newton method [6][7] is employed to minimize the objective function $E(x)$, where the parameter vector x defined in (5) is uptaded using $x_{k+1} = x_k + \alpha^* s$, where $s = -(\nabla^2 E_k + \beta I)^{-1} g_k$ with g_k the gradient at x_k and $\beta > 0$ a scalar that is chosen to guarantee the possitive defeniteness of $\nabla^2 E_k + \beta I$, and $\alpha^* > 0$ is a scalar that minimizes $E(x_k + \alpha s)$. To obtain a stable design we begin with an initial point $x_0 = [h_0 \ \ a^T \ \ b^T]^T$

where $a = 0$ and h_0 and b are chosen randomly. It is required that the phase response of the IIR filter be nearly linear with group delay $= -0.2356$. With $p = 2$ and the initial point

$$x_0 = \begin{bmatrix} 1.3353303e-01 \\ 0.0000000e+00 \\ 0.0000000e+00 \\ 0.0000000e+00 \\ 0.0000000e+00 \\ 0.0000000e+00 \\ 0.0000000e+00 \\ 0.0000000e+00 \\ 0.0000000e+00 \\ 5.0583328e-01 \\ 2.4793412e-01 \\ -3.6705445e-01 \\ 6.7904697e-01 \\ 1.2669970e+00 \\ 2.3582406e-01 \\ 1.4866899e-02 \\ -7.2458709e-01 \end{bmatrix}$$

after 34 iterations the optimization algorithm reduces $E(x)$ from 64.3896 to 0.0355, and the parameter vector converges to

$$x^* = \begin{bmatrix} -2.6935863e-02 \\ -2.5215390e+00 \\ 4.4119570e+00 \\ -5.4390598e+00 \\ 5.0072698e+00 \\ -3.4599730e+00 \\ 1.7413254e+00 \\ -5.8620247e-01 \\ 1.0298105e-01 \\ -2.1221500e+00 \\ 2.3046177e+00 \\ -1.6913922e+00 \\ 1.9665881e+00 \\ -1.8181378e+00 \\ -3.3802596e+00 \\ -2.7697074e+00 \\ -2.7637528e+00 \end{bmatrix}$$

that crresponds to a stable 8th-order IIR filter whose poles are at $0.0355 \pm 0.9794j$, $0.1872 \pm 0.7566j$, $0.4371 \pm 0.5068j$, and $0.6010 \pm 0.1812j$. It is found that the maximum amplitude ripples over passband and stopband are 0.0474 and 0.0165, respectively. The largest relative deviation of the group delay over passband is 0.0382.

Next we use the same initial point and the same optimization algorithm with $p = 4$. After 120 iterations $E(x)$ is reduced from 96.0855 to 1.3750×10^{-5}. The maximum ripples over passband and stopband are 0.0350 and 0.0215, respectively, and the largest relative deviation of the group delay over passband is 0.0299. The stable filter obtained is characterized by the point x^* where

$$x^* = \begin{bmatrix} -2.3713190e-02 \\ -2.3758093e+00 \\ 4.0621735e+00 \\ -4.8932758e+00 \\ 4.4083373e+00 \\ -2.9826650e+00 \\ 1.4693185e+00 \\ -4.8331438e-01 \\ 8.2611150e-02 \\ -1.8603566e+00 \\ 1.7699303e+00 \\ -1.2841709e+00 \\ 2.1161220e+00 \\ -2.0996904e+00 \\ -3.8910917e+00 \\ -3.8107935e+00 \\ -3.7601722e+00 \end{bmatrix}$$

Finally, we use an initial point x_0 obtained from the balanced approximation method, where

$$x_0 = \begin{bmatrix} -2.4000000e-03 \\ -1.3289000e+00 \\ 1.5139000e+00 \\ -1.2916000e+00 \\ 8.9750000e-01 \\ -5.0780000e-01 \\ 2.2410000e-01 \\ -7.0600000e-02 \\ 1.3000000e-02 \\ 4.1666667e-01 \\ -4.7500000e+00 \\ -1.7083333e+00 \\ 1.9458333e+01 \\ 1.7083333e+00 \\ -6.5625000e+01 \\ -9.1375000e+01 \\ -4.6333333e+01 \end{bmatrix}$$

It is easy to verify that x_0 corresponds to a stable transfer function of order 8 that approximates the desired lowpass characteristics with good linear phase response over the passband. With 50 iterations $E(x)$ is reduced from 1.2199 to 0.0401 and the algorithm leads to a parameter vector

$$x^* = \begin{bmatrix} -1.4389789e-03 \\ -2.2575908e+00 \\ 3.7455771e+00 \\ -4.3780299e+00 \\ 3.8072623e+00 \\ -2.4620460e+00 \\ 1.1395524e+00 \\ -3.4102604e-01 \\ 4.9090480e-02 \\ 1.4803702e-01 \\ -1.2183445e+01 \\ 1.3706396e+01 \\ 1.5564879e+01 \\ -2.5910250e+01 \\ -6.4122167e+01 \\ -8.0143141e+01 \\ -5.9112535e+01 \end{bmatrix}$$

The filter's amplitude response and group delay over passband are depicted in Fig. 1(a) and (b), respectively. Its maximum ripples over the passband and stopband are 0.0366 and 0.0302, respectively. The largest relative deviation of the group delay over passband is 0.0265, and the filter is stable with poles at $0.0095 \pm 0.9845j$, $0.1959 \pm 0.7100j$, $0.4274 \pm 0.4226j$, and $0.4961 \pm 0.1112j$.

Acknowledgement

The authors wish to thank Mr. S. Xiao for his help in preparing the manuscript.

1. REFERENCES

[1] W.-S. Lu and A. Antoniou, "Design of 2-D filters with arbitrary amplitude and phase response using singular value decomposition", in *Prco. IEEE Int. Symp. Circuits Syst.*, pp. 618-621, Singapore, Jun. 1991.

[2] E.G. Evans and R. Fischl, "Optimal least squares time-domain synthesis of recursive digital filters", *IEEE Trans. Audio Electroacoust.*, vol. 21, pp. 61-65, Feb. 1973.

[3] J.A. Cadzow, "Recursive digital filter synthesis via gradient based algorithms", *IEEE Trans. Acoust. Speech Signal Processing*, vol. 24, pp. 349-355, Oct. 1976.

[4] C. Charalambous, "Acceleration of the least *p*th algorithm for minimax optimization with engineering applications", *Math. Programming*, vol. 17, pp. 270-279, 1979.

[5] C. Charalambous and A. Antoniou, "Equalisation of recursive digital filters", *IEE Proc.*, vol. 127, pt. G, pp. 219-225, Oct. 1980.

[6] R. Fletcher, *Practical Methods of Optimization*, John Wiley: Chichester, 1987.

[7] D.G. Luenberger, *Linear and Nonlinear Programming*, Reading, MA: Addison-Wesley, 1984.

[8] B.C. Moore, "Principal component analysis in linear systems: Controllability, observability, and model reduction", *IEEE Trans. Automat. Contr.*, vol. 26, pp. 17-23, Jan. 1981.

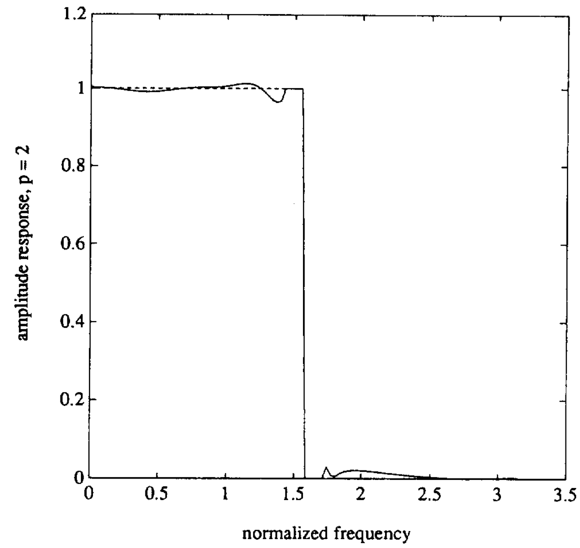

Fig. 1(a)

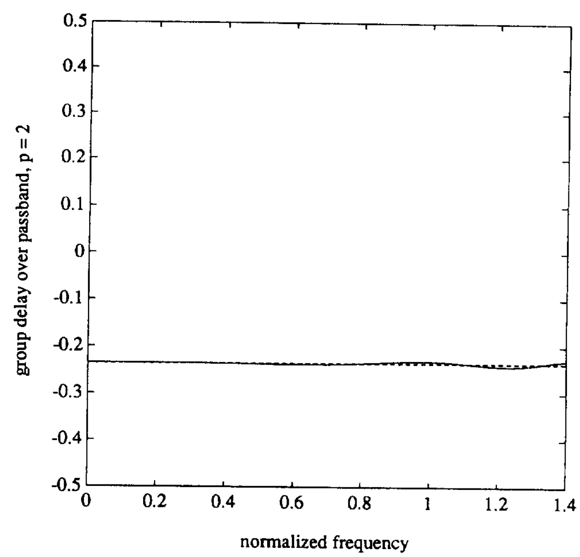

Fig. 1(b)