Almost minimax design of FIR filter using an IRLS algorithm without matrix inversion

Zhao, Ruijie; Lin, Zhiping; Toh, Kar-Ann; Sun, Lei; Lai, Xiaoping

2017

Zhao, R., Lin, Z., Toh, K.-A., Sun, L., & Lai, X. (2017). Almost minimax design of FIR filter using an IRLS algorithm without matrix inversion. Second International Workshop on Pattern Recognition, 10443, 104431F-. doi:10.1117/12.2280405

<https://hdl.handle.net/10356/102533>

<https://doi.org/10.1117/12.2280405>

© 2017 Society of Photo-optical Instrumentation Engineers (SPIE). This paper was published in Second International Workshop on Pattern Recognition and is made available as an electronic reprint (preprint) with permission of Society of Photo-optical Instrumentation Engineers (SPIE). The published version is available at: [<http://dx.doi.org/10.1117/12.2280405>]. One print or electronic copy may be made for personal use only. Systematic or multiple reproduction, distribution to multiple locations via electronic or other means, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper is prohibited and is subject to penalties under law.

Downloaded on 26 Apr 2023 22:28:59 SGT

Almost Minimax Design of FIR filter Using an IRLS Algorithm without Matrix Inversion

Ruijie Zhao^{a,b}, Zhiping Lin^b, Kar-Ann Toh^c, Lei Sun^d and Xiaoping Lai^e

^aSchool of Mechanical, Electrical and Information Engineering, Shandong University, Weihai, 264209, China;

^bSchool of Electrical and Electronic Engineering, Nanyang Technological University, 639798, Singapore;

^c School of Electrical and Electronic Engineering, Yonsei University, Seoul, 120-749, Korea;

^d School of Information and Electronics, Beijing Institute of Technology, Beijing, 100081, China;

^e Institute of Information and Control, Hangzhou Dianzi University, Hangzhou, 310018, China

ABSTRACT

An iterative reweighted least squares (IRLS) algorithm is presented in this paper for the minimax design of FIR filters. In the algorithm, the resulted subproblems generated by the weighted least squares (WLS) are solved by using the conjugate gradient (CG) method instead of the time-consuming matrix inversion method. An almost minimax solution for filter design is consequently obtained. This solution is found to be very efficient compared with most existing algorithms. Moreover, the filtering solution is flexible enough for extension towards a broad range of filter designs, including constrained filters. Two design examples are given and the comparison with other existing algorithms shows the excellent performance of the proposed algorithm.

Keywords: Conjugate gradient method, FIR filters, IRLS algorithm, Minimax design

1. INTRODUCTION

The design of FIR digital filters based on the minimax criterion has been investigated extensively because this design can yield an equiripple behavior and avoid the Gibbs phenomenon in magnitude. The Remez algorithm¹ is a powerful tool for the design of minimax linear-phase FIR filters. However, it is very sophisticated and cannot be used for designing nonlinear-phase filters or incorporating any constraints.

The iterative reweighted least squares (IRLS) technique²⁻⁸ is an alternative choice for the minimax filter design. The technique transforms the original problem into a series of weighted least squares (WLS) subproblems. By updating the weight functions and solving these WLS subproblems iteratively, an almost minimax solution can be obtained. The IRLS algorithms have been verified to be very efficient and flexible for designing various classes of digital filters.

The main computational burden in IRLS algorithms lies in solving those WLS subproblems. However, most of the IRLS algorithms^{2-4,6} developed so far are found to solve the WLS subproblems using the matrix inversion method. This would require a large memory space and considerable computation time, especially for high-order filters. On the other hand, several authors developed the IRLS algorithms^{5,7,8} in which matrix inversion is avoided. The algorithm⁵ introduces an extra frequency response implicitly including the weighting function so that no matrix inversion is needed. But the resulting filter is not precise enough. In the paper,⁷ the authors replace the matrix inversion with its power series expansion. However, practical design experience shows that the convergence cannot be achieved in some cases. The algorithm⁸ uses an iterative procedure to solve the WLS subproblems, and it turns out that a very high design efficiency can be achieved.

Further author information: (Send correspondence to Zhiping Lin)

Ruijie Zhao: E-mail: zhao_rj@sdu.edu.cn; Zhiping Lin: E-mail: ezplin@ntu.edu.sg;
Kar-Ann Toh: E-mail: katoh@ieee.org; Lei Sun: E-mail: sunlei@bit.edu.cn;
Xiaoping Lai: E-mail: laixp@hdu.edu.cn

In this paper, an alternative approach to IRLS algorithm design without matrix inversion is presented, wherein the conjugate gradient (CG) method is used to solve the WLS subproblems. By using the solution of the current WLS subproblem as the initial iterative point in the CG method for the next WLS subproblem and setting a convergence tolerance, the iteration numbers required by the CG method will be reduced greatly. This makes the developed IRLS algorithm be more efficient. Finally, two design examples are given to demonstrate its efficiency.

2. MINIMAX DESIGN OF FIR FILTERS

The frequency response of a length- N FIR filter can be expressed as

$$H(e^{j\omega}, \mathbf{h}) = \sum_{n=0}^{N-1} h(n)e^{-jn\omega} = \Psi(e^{j\omega})^T \mathbf{h}, \quad (1)$$

where ω is the frequency in the range $|\omega| \leq \pi$ and $\mathbf{h}$ is the impulse response, which can be real or complex. In (1), both $\mathbf{h}$ and $\Psi(e^{j\omega})$ are vectors defined respectively by

$$\mathbf{h} = [h(0), h(1), h(2), \dots, h(N-1)]^T, \quad \Psi(e^{j\omega}) = [1, e^{-j\omega}, e^{-2j\omega}, \dots, e^{-(N-1)j\omega}]^T,$$

and the superscript “ T ” denotes the transpose operation.

Let $D(e^{j\omega})$ be the frequency response of the desired filter. Our objective is to find a length- N FIR filter $H(e^{j\omega}, \mathbf{h})$ to approximate the desired one $D(e^{j\omega})$. In this paper, we consider the minimax design problem of FIR filters, which can be described as

$$\min_{\mathbf{h}} \max_{|\omega| \leq \pi} \{|H(e^{j\omega}, \mathbf{h}) - D(e^{j\omega})|\}, \quad (2)$$

i.e., to find the impulse response $\mathbf{h}$ so that the maximum value of the frequency response error $|H(e^{j\omega}, \mathbf{h}) - D(e^{j\omega})|$ is minimized. For FIR filters, if $\mathbf{h}$ is real and satisfies the symmetry $h(n) = h(N-1-n)$ or $h(n) = -h(N-1-n)$, the frequency response $H(e^{j\omega}, \mathbf{h})$ has a linear phase, and the filter is the so-called linear-phase filter.

For simplicity, we describe the minimax design of FIR filters and the proposed algorithm only for the linear-phase filters with the symmetry $h(n) = h(N-1-n)$ and odd length N , even if the obtained algorithm can be used to design other classes of filters. In this case, the magnitude response can be written as

$$H(\omega, \mathbf{a}) = \Phi(\omega)^T \mathbf{a}, \quad (3)$$

where $\mathbf{a}$ is a real vector related to $\mathbf{h}$,⁸ and $\Phi(\omega)$ is given by

$$\Phi(\omega) = [1, \cos(\omega), \dots, \cos((N-1)\omega/2)]^T.$$

It is only needed to consider the approximation of a desired magnitude response $D(\omega)$ by $H(\omega, \mathbf{a})$ on the frequency band $\Omega \in [0, \pi]$ of interest. Assume Ω is represented by a set of discretized point as $\bar{\Omega} = \{\omega_i \in \Omega, i = 1, 2, \dots, M\}$, the design problem (2) becomes

$$\min_{\mathbf{a}} \max_{\omega \in \bar{\Omega}} \{|H(\omega, \mathbf{a}) - D(\omega)|\}. \quad (4)$$

3. PROPOSED IRLS ALGORITHM

Considering that IRLS algorithms solve the minimax problem by transforming it into a series of WLS subproblems, we first formulate the WLS design problem of FIR filters,

$$\min_{\mathbf{a}} E(\mathbf{a}) = 0.5 \sum_{i=1}^M w_i |H(\omega_i, \mathbf{a}) - D(\omega_i)|^2, \quad (5)$$

where w_i 's are the given nonnegative weighting coefficients.

Further, the weighted squared error $E(\mathbf{a})$ can be expressed in matrix-vector form as

$$E(\mathbf{a}) = 0.5(\mathbf{Pa} - \mathbf{d})^T \mathbf{W}(\mathbf{Pa} - \mathbf{d}),$$

where $\mathbf{d}$ and $\mathbf{P}$ are written as

$$\mathbf{d} = [D(\omega_1), D(\omega_2), \dots, D(\omega_M)]^T, \quad \mathbf{P} = [\Phi(\omega_1), \Phi(\omega_2), \dots, \Phi(\omega_{r-1})]^T,$$

and $\mathbf{W}$ is a diagonal matrix with its main diagonal entries being w_i 's.

The optimality condition of problem (5) is that the derivative of $E(\mathbf{a})$ with respect to $\mathbf{a}$ equals zero, i.e.,

$$\nabla E(\mathbf{a}) = \mathbf{P}^T \mathbf{W} \mathbf{P} \mathbf{a} - \mathbf{P}^T \mathbf{W} \mathbf{d} = \mathbf{0}.$$

If the matrix $\mathbf{P}^T \mathbf{W} \mathbf{P}$ is invertible, which condition could generally be satisfied in practical filter design, the solution of the WLS problem (5) can be given by

$$\mathbf{a} = (\mathbf{P}^T \mathbf{W} \mathbf{P})^{-1} \mathbf{P}^T \mathbf{W} \mathbf{d}. \quad (6)$$

In most of those IRLS algorithms developed so far, the formulation (6) is used to solve the WLS subproblems involved in each iteration, which leads to a large amount of computation. To reduce the computational complexity, we present the following CG method to solve the WLS design problem (5).

Proposed WLS Algorithm

Step 1: Given filter length N , $D(\omega)$, $\bar{\Omega}$, and an error tolerance $\varepsilon_1 > 0$, evaluate $\mathbf{d}$ and $\mathbf{P}$. Given a weight vector $\mathbf{w}$, choose an initial iterative vector $\mathbf{a}_0$. Let $\mathbf{g}_0 = -\nabla E(\mathbf{a}_0)$ and set $k = 0$.

Step 2: Compute $\mathbf{r}_k = \mathbf{P}^T[(\mathbf{P}\mathbf{g}_k) \circ \mathbf{w}]$ and $den = \mathbf{r}_k^T \mathbf{g}_k$. Let $\mathbf{a}_{k+1} = \mathbf{a}_k + \beta_k \mathbf{g}_k$, where $\beta_k = -\frac{\mathbf{g}_k^T \nabla E(\mathbf{a}_k)}{den}$.

Step 3: Let $\mathbf{g}_{k+1} = -\nabla E(\mathbf{a}_{k+1}) + \lambda_k \mathbf{g}_k$, where $\nabla E(\mathbf{a}_{k+1}) = \nabla E(\mathbf{a}_k) + \beta_k \mathbf{r}_k$, $\lambda_k = \frac{\beta_k \|\nabla E(\mathbf{g}_{k+1})\|_2^2}{den}$.

Step 4: If $\|\mathbf{a}_{k+1} - \mathbf{a}_k\|_2 / \|\mathbf{a}_k\|_2 < \varepsilon_1$ or $k > (N+1)/2$, terminate the algorithm. Otherwise, let $k = k+1$ and go back to Step 2.

In the above WLS algorithm, $\mathbf{w}$ is the weight vector with its components being w_i 's. “ $\circ$ ” and $\|\cdot\|_2$ denote the Hadamard product and the vector L_2 norm, respectively. In addition, we use an error tolerance ε_1 to reduce the number of iterations.

Now, we combine the above WLS algorithm with the IRLS technique to obtain the following IRLS algorithm, which can solve the minimax problem of FIR filter (2) very efficiently.

Proposed IRLS Algorithm

Step 1: Given filter length N , $D(\omega)$, $\bar{\Omega}$, and two error tolerances $\varepsilon_1 > 0$, $\varepsilon_2 > 0$, evaluate $\mathbf{d}$ and $\mathbf{P}$. Let the initial weight vector $\mathbf{w}_0 = [1, 1, \dots, 1]^T$ and initial iterative vector $\mathbf{a}_0 = (\mathbf{P}^T \mathbf{P})^{-1} \mathbf{P}^T \mathbf{d}$. Let $\mathbf{g}_0 = -\nabla E(\mathbf{a}_0)$ and set $k = 0$.

Step 2: Update $\mathbf{w}_{k+1}$ using $\mathbf{w}_k$ and $\mathbf{e}_k$, where $\mathbf{e}_k$ is the error vector in the k -th iteration, given by $\mathbf{e}_k = \mathbf{Pa}_k - \mathbf{d}$.

Step 3: Solve the WLS subproblem

$$\min_{\mathbf{a}} (\mathbf{Pa} - \mathbf{d})^T \mathbf{W}_{k+1} (\mathbf{Pa} - \mathbf{d}),$$

by using the proposed WLS algorithm with the error tolerance ε_1 and initial iterative vector $\mathbf{a}_k$. The solution is denoted as $\mathbf{a}_{k+1}$.

Step 4: If $\|\mathbf{a}_{k+1} - \mathbf{a}_k\|_2 / \|\mathbf{a}_k\|_2 < \varepsilon_2$, terminate the algorithm. Otherwise, let $k = k+1$ and return to Step 2.

In Step 2, we use the weight update method from reference³ to yield $\mathbf{w}_{k+1}$. In Step 3, $\mathbf{W}_{k+1} = \text{diag}(\mathbf{w}_{k+1})$ represents the diagonal matrix with vector $\mathbf{w}_{k+1}$ on the main diagonal. It can be seen that the solution vector $\mathbf{a}_k$ in the k -th iteration is used as the initial iterative vector for the WLS problem in the $(k+1)$ -th iteration, which can greatly reduce the number of iterations required by the WLS algorithm in practical designs.

Figure 1. Magnitude response of the length-101 flatness-constrained filter

Figure 2. Magnitude error of the complex filter for $\tau = 15$

4. DESIGN EXAMPLES

Two design examples are provided to show the high efficiency of the proposed IRLS algorithm. Comparisons with some existing methods are given. All the algorithms are implemented in Matlab programs on a PC with an Intel (R) Core (TM) i5-4210U CPU (2.4-GHz) and an 8-Gb memory.

Example 1: Minimax design of flatness-constrained linear-phase FIR filter with passband $[0, 0.3\pi]$ and stopband $[0.34\pi, \pi]$. The flatness constraint is described as

$$d^k H(\mathbf{a}, \omega) / d\omega^k \big|_{\omega=\omega_m} = d^k D(\omega) / d\omega^k \big|_{\omega=\omega_m}, \quad \text{for } k = 0, 1, \dots, K-1.$$

In this example, let $\omega_m = 0.15\pi$ and $K = 3$.

The flatness constraints are linear equality constraints with respect to $\mathbf{a}$, which can be incorporated into the cost function of the WLS problem. We use the algorithms^{5,8,9} and the proposed IRLS algorithm to design such filters. The error tolerances are taken to be $\varepsilon_1 = 10^{-7}$ and $\varepsilon_2 = 10^{-6}$ in the algorithm⁸ and the proposed one.

The design results, including the design time (T_{cpu}), iteration number (N_i) and maximum magnitude error (E_{max}), are shown in Table 1. It can be seen that, compared with the algorithm⁹, the proposed one is more efficient except for a small deviation in E_{max} . In addition, the algorithm⁹ fails in high-order filter design. Moreover, there is an improvement in design time and design precision comparing with the algorithms.^{5,8}

Further, we find that the average iteration numbers required by the CG method are 14.2, 26.5, 35 and 40 for $N=101, 151, 201$ and 251 , respectively, which explains why our algorithm is very efficient. Figure 1 shows the magnitude response of the flatness-constrained filter designed by our algorithm for $N=101$.

Table 1. Comparisons with the algorithms in the papers^{5,8,9}

N	Algorithm ⁹			Algorithm ⁵			Algorithm ⁸			Proposed IRLS Algorithm		
	T_{cpu} (s)	E_{max}	N_i	T_{cpu} (s)	E_{max}	N_i	T_{cpu} (s)	E_{max}	N_i	T_{cpu} (s)	E_{max}	
101	0.15	9.8502×10^{-3}	1448	0.1	1.0837×10^{-2}	30	0.5	9.8896×10^{-3}	29	0.09	9.8912×10^{-3}	
151	0.58	1.7006×10^{-3}	395	0.1	1.9874×10^{-3}	7	0.46	1.7589×10^{-3}	12	0.06	1.7219×10^{-3}	
201	1.4	3.0760×10^{-4}	264	0.23	3.7745×10^{-4}	7	0.42	3.2174×10^{-4}	7	0.08	3.2046×10^{-4}	
251	— [*]	— [*]	228	0.32	7.2927×10^{-5}	6	0.79	5.9052×10^{-5}	9	0.14	5.8970×10^{-5}	

* Algorithm fails.

Example 2: Minimax design of length-71 complex FIR band-pass filter with passband $[-0.4\pi, 0.4\pi]$ and stopbands $[-\pi, -0.5\pi]$ and $[0.5\pi, \pi]$. The desired frequency response in the passband is given by $e^{-\tau\omega}$.

In this example, we design the complex FIR filters using the complex Remez algorithm,¹⁰ Lang's IRLS algorithm⁶ (matrix inversion method), Zhu's algorithm⁷ and the proposed IRLS algorithm ($\varepsilon_1 = \varepsilon_2 = 10^{-6}$) for different group delay τ .

The design results are shown in Table 2. From the table, we can see that the proposed IRLS algorithm is more numerically stable than the algorithms,^{7,10} which cannot obtain the optimal solutions for $\tau = 15, 5$ and for $\tau = 20, 10$. Comparing with the algorithm,⁶ the proposed one has a similar design precision, but enjoys much less design time. Figure 2 depicts the magnitude error of the complex filter designed by our algorithm for $\tau = 15$.

Table 2. Comparisons with the algorithms in the papers^{6,7,10}

τ	Algorithm ¹⁰			Algorithm ⁶			Algorithm ⁷			Proposed IRLS Algorithm		
	$T_{cpu}(s)$	E_{max}	N_i	$T_{cpu}(s)$	E_{max}	N_i	$T_{cpu}(s)$	E_{max}	N_i	$T_{cpu}(s)$	E_{max}	
20	0.028	9.0867×10^{-4}	14	2.8	9.1226×10^{-4}	—*	—*	—*	7	0.077	9.1597×10^{-4}	
15	1.2	3.3566×10^{-3}	9	1.8	1.2384×10^{-3}	10	1.72	1.2382×10^{-3}	6	0.055	1.2424×10^{-3}	
10	0.031	1.9410×10^{-3}	14	2.8	1.9733×10^{-3}	—*	—*	—*	7	0.088	1.9772×10^{-3}	
5	2.1	5.3250×10^{-3}	11	2.2	3.4925×10^{-3}	11	1.8	3.4927×10^{-3}	8	0.072	3.5006×10^{-3}	

* Algorithm fails.

5. CONCLUSION

An IRLS algorithm which avoided the matrix inversion was developed. This algorithm can be used to design a wide range of almost minimax FIR filters. The proposed one needs much less design time and at the same time gives comparable or better design precision compared with several existing algorithms. Moreover, it is numerically stable and flexible enough for various filter design extensions. The two design examples presented validated the good performance in both processing speed and filtering precision.

ACKNOWLEDGMENTS

This research was supported partially by the National Nature Science Foundation of China under Grants 61304142, 61573123 and 61673059, and partially by the Singapore Academic Research Fund (AcRF) Tier 1 under Project RG 31/16.

REFERENCES

- [1] Parks, T. W. and McClellan, J. H., "Chebyshev approximation for nonrecursive digital filters with linear phase," *IEEE Trans. Circuit Theory* **19**, 189–194 (1972).
- [2] Lawson, C. L., *Contributions to the theory of linear least maximum approximations*, PhD thesis, University of California (1961).
- [3] Lim, Y. C., Lee, J. H., Chen, C. K., and Yang, R. H., "A weighted least squares algorithm for quasi-equiripple FIR and IIR digital filter design," *IEEE Trans. Signal Process.* **40**, 551–558 (1992).
- [4] Burrus, C. S., Barreto, J. A., and Selesnick, I. W., "Iterative reweighted least-squares design of FIR filters," *IEEE Trans. Signal Process.* **42**, 2926–2936 (1994).
- [5] Jou, Y. D., Hsieh, C. H., and Kou, C. M., "Efficient weighted least-squares algorithm for the design of FIR filters," in *[IEE Proceedings-Vision, Image and Signal Processing]*, **144**, 244–248 (1997).
- [6] Lang, M., *Algorithms for the constrained design of digital filters with arbitrary magnitude and phase responses*, PhD thesis, Vienna University of Technology (1999).
- [7] Zhu, W. P., Ahmad, M. O., and Swamy, M. N. S., "Weighted least-square design of FIR filters using a fast iterative matrix inversion algorithm," *IEEE Trans. Circuits and Systems-I* **49**, 1620–1627 (2002).
- [8] Zhao, R. and Lai, X., "A new iterative reweighted least squares algorithm for the design of FIR filters," in *[2013 9th International Conference on Information, Communications and Signal Processing (ICICSP)]*, 10–13 (2013).
- [9] Zhao, R. and Lai, X., "Chebyshev design of linear-phase FIR filters with linear equality constraints," *IEEE Transactions on Circuits and Systems-II* **54**, 494–498 (2007).
- [10] Karam, L. J. and McClellan, J. H., "Complex chebyshev approximation for FIR filter design," *IEEE Transactions on Circuits and Systems-II* **42**, 207–216 (1995).