CLOSED FORM EXPRESSIONS FOR OPTIMUM FIR POLYNOMIAL PREDICTIVE FILTERING

Emmanouil Z. Psarakis[⋆] *and George V. Moustakides*[†]
[⋆]Dept. of Business Planning & Information Systems, TEI of Patras, 26334 Patras, Greece
and Computer Technology Institute (CTI), 26110 Patras, Greece
E-mail: `psarakis@cti.gr`
[†]IRISA-INRIA, Campus de Beaulieu, 35042 Rennes Cedex, France
E-mail: `George.Moustakides@inria.fr`

ABSTRACT

We propose two different approaches for finding optimum FIR polynomial predictive filters. The first reveals an important property of these filters and can lead to a numerical solution that is of the size of the number of constraints imposed instead of the size of the filter. With the second method, using a different parametrization of the filter coefficients, it is possible to obtain closed form expressions for the optimum filter coefficients and for the minimum output noise power, in the general case. Previous approaches were successful in providing closed forms expressions only for special cases.

1 INTRODUCTION AND BACKGROUND

The problem we are interested in can be described as follows. We are given sequentially observations $x(n)$ with $x(n) = p(n) + w(n)$ where $p(n)$ is a deterministic polynomial signal of the form

$$p(n) = a_0 + a_1 n + \cdots + a_{M-1} n^{M-1}, \tag{1}$$

with unknown coefficients a_i and $w(n)$ is white noise with zero mean and variance σ_w^2. We would like to design a linear predictive filter that can estimate $p(n)$ from the past available observations $x(n-1), x(n-2), \ldots$. Such problems arise in video and image processing [1], satellite communications [2], industrial electronics, instrumentation and measurements [3] and biomedical signal processing [5].

If we consider FIR filters of length L and denote by $h_1, \ldots, h_{L-1}$ the impulse response and by $\hat{p}(n)$ the output of the filter then

$$\begin{aligned} \hat{p}(n) &= \sum_{j=1}^{L} h_j x(n-j) \\ &= \sum_{j=1}^{L} h_j p(n-j) + \sum_{j=1}^{L} h_j w(n-j). \end{aligned}$$

As performance measure we consider the mean square error $\mathbb{E}\{[\hat{p}(n) - p(n)]^2\}$ between the estimate $\hat{p}(n)$ and and the desired signal $p(n)$. We would therefore like to select the coefficients in the MMSE sense.

Since the desired signal $p(n)$ is deterministic, time varying and tends to infinity, in order to be able to obtain a measure that does not depend on time and does not diverge, it is necessary to consider estimates $\hat{p}(n)$ that are unbiased, that is,

$$\mathbb{E}\{\hat{p}(n)\} = \sum_{j=1}^{L} h_j p(n-j) = p(n). \tag{2}$$

In fact this condition is also equivalent to requiring the estimate to be exact under zero noise power. It can be shown [1], [5], that the constraint in (2) is equivalent to the following system of linear equality constraints

$$\sum_{j=1}^{L} h_j = 1 \tag{3}$$

$$\sum_{j=1}^{L} h_j j^k = 0, \quad k = 1, \ldots, M-1. \tag{4}$$

Furthermore, under (2) our performance measure takes the simple form

$$\begin{aligned} \mathbb{E}\{[\hat{p}(n) - p(n)]^2\} &= \mathbb{E}\left\{\left[\sum_{j=1}^{L} h_j w(n-j)\right]^2\right\} \\ &= \sigma_w^2 \sum_{j=1}^{L} h_j^2. \end{aligned} \tag{5}$$

We can now define the optimization problem that will determine the optimality of our FIR filter. As we said, we are interested in the MMSE between $\hat{p}(n)$ and $p(n)$, under the constraint that $\mathbb{E}\{\hat{p}(n)\} = p(n)$. This, from (5), translates into

$$\min_{h_1, \ldots, h_L} \sum_{j=1}^{L} h_j^2 \tag{6}$$

under constraints (3) and (4).

The optimization defined by (6), (3) and (4) has been considered in [3], [5]. However here we propose two different approaches, that can lead to a better understanding of the filter structure and provide closed form expressions for the filter coefficients and the final MMSE.

2 POLYNOMIAL STRUCTURE OF OPTIMUM FILTER COEFFICIENTS

In this section we are going to show an important property of the filter coefficients. If this property is employed in the computation of the optimum filter structure it reduces the size L linear system required to find the optimum coefficients into a size M system. In view however of the results of the next section, where we present closed form expressions, the results of the present section are mainly of qualitative interest.

Although the filter coefficients can lie anywhere in an L-dimensional space we will show that the optimum ones will necessarily lie inside a significantly smaller and well defined linear subspace. More precisely we will show that the optimum filter coefficients have a similar polynomial structure as the deterministic signal $p(n)$, that is,

$$h_k = f_0 + f_1 k + \cdots + f_{M-1} k^{M-1}, \quad k = 1, \ldots, L. \quad (7)$$

Note that this relation was used in [4], however without any proof of its general validity. Before proving this result, let us redefine our optimization problem using matrix notation.

Let $H = [h_1\ h_2\ \cdots\ h_L]^t$ denote the filter coefficient vector; $E_M = [1\ \underbrace{0\ \cdots\ 0}_{M-1}]^t$ and consider the following (Vandermonde) matrix

$$\mathbf{V}_{L,N} = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 2^2 & \cdots & 2^{N-1} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & L & L^2 & \cdots & L^{N-1} \end{bmatrix}.$$

With the above definitions, the optimization problem we like to solve becomes

$$\min_H \|H\|^2, \quad \text{under} \quad \mathbf{V}_{L,M}^t H = E_M. \quad (8)$$

Notice now that since the elements of the second column of $\mathbf{V}_{L,L}$ are different the square matrix $\mathbf{V}_{L,L}$ is invertible. This suggests that any vector H can be written as a linear combination of the L columns of $\mathbf{V}_{L,L}$. As far as the optimum solution of the optimization problem (8) is concerned we have the following theorem.

Theorem 1: The optimum filter coefficient vector H_o that solves the constrained minimization problem in (8) can be written as a linear combination of the columns of $\mathbf{V}_{L,M}$ and in particular has the form

$$H_o = \mathbf{V}_{L,M}(\mathbf{V}_{L,M}^t \mathbf{V}_{L,M})^{-1} E_M. \quad (9)$$

Proof: The proof is interesting and simple. Let the matrix $\mathcal{V}_{L,L} = [V_1\ V_2\ \cdots\ V_L]$, with $V_i,\ i = 1, \ldots, L$, denoting the corresponding columns, be the *orthonormalized* version of $\mathbf{V}_{L,L}$ obtained by orthonormalizing successively the columns of $\mathbf{V}_{L,L}$. It is then clear that the i-th column V_i of $\mathcal{V}_{L,L}$ is a linear combination of the first i columns of $\mathbf{V}_{L,L}$.

Any vector H can be written as a linear combination of the columns of $\mathcal{V}_{L,L}$. Therefore let

$$H = \mathcal{V}_{L,L} F$$

where $F = [f_0\ \cdots\ f_{L-1}]^t$. Since $\mathcal{V}_{L,L}$ is orthonormal we have that its columns V_i for $i > M$ are orthogonal to its columns V_j with $j \leq M$ and therefore orthogonal to the matrix $\mathbf{V}_{L,M}$. This in return suggests that the constraint can be written

$$\mathbf{V}_{L,M}^t \mathcal{V}_{L,L} F = \mathbf{V}_{L,M}^t \mathcal{V}_{L,M} F_M = E_M,$$

where $\mathcal{V}_{L,M}$ contains the first M columns of $\mathcal{V}_{L,L}$ and F_M is a vector of length M containing the first M elements of F. More precisely we conclude that any vector satisfying the constraint in (8), if it is expressed as a linear combination of the columns of $\mathcal{V}_{L,L}$, then its first M elements must satisfy the above equation. This means that this first part must always be the same and therefore necessarily equal to

$$F_M = (\mathbf{V}_{L,M}^t \mathcal{V}_{L,M})^{-1} E_M.$$

The norm on the other hand of H, since $\mathcal{V}_{L,L}$ is orthonormal, can be written as

$$\|H\|^2 = \|F\|^2 \geq \|F_M\|^2,$$

with equality if and only if $f_i = 0$, for $i > M$. From this we conclude that the optimum H_o is a linear combination of the first M columns of $\mathcal{V}_{L,L}$ or equivalently of all columns of $\mathbf{V}_{L,M}$.

To find the final form of H_o, since now we know that $H_o = \mathbf{V}_{L,M} F_o$ and H_o needs to satisfy the constraint in (8), we conclude that $F_o = (\mathbf{V}_{L,M}^t \mathbf{V}_{L,M})^{-1} E_M$ which yields the H_o appearing in (9). This concludes the proof. ■

We would like to stress once more that this result basically reveals the polynomial structure (7) of the optimum filter and can also be used for the numerical computation of the corresponding coefficients. In the next section, however, following a different parametrization of the filter coefficients we will be able to come up with closed form expressions for the coefficients and the optimum output noise power (MMSE).

3 CLOSED FORM EXPRESSIONS

In order to be able to compute analytically the optimum filter coefficients it is more convenient to employ the transfer function of the filter and, as was stated previously, to use a different parametrization of the filter coefficients. Specifically if $h(z)$ is the transfer function then it must have the form

$$h(z) = 1 - (1 - z^{-1})^M \left(1 + \sum_{j=1}^{L-M} \alpha_j z^{-j} \right).$$

Indeed we can see that $h(z)$ is a polynomial of order L in z^{-1} therefore it corresponds to a finite impulse response of the form $h_0, h_1, \ldots, h_L$. Notice also that $\lim_{z\to\infty} h(z) = 0$ therefore we conclude that $h_0 = 0$, corresponding to a one step ahead predictor. Finally the filter coefficients satisfy the M constraints defined in (3) and (4), because $h(z)|_{z=1} = 1$ and $h^{(k)}(z)|_{z=1} = 0,\ k = 1, \ldots, M-1$, where $h^{(k)}(z)$ denotes the k-th derivative of $h(z)$. Therefore the proposed transfer function belongs to the FIR filter class we are interested in and has the necessary degrees of freedom expressed by the $L-M$ parameters α_i.

It is convenient at this point to introduce the following function $\tilde{h}(z)$

$$\begin{aligned}\tilde{h}(z) &= 1 - h(z) = (1-z^{-1})^M \left(1 + \sum_{j=1}^{L-M} \alpha_j z^{-j}\right) \\ &= \left(\sum_{l=0}^{M} \mathcal{C}_l^M (-1)^l z^{-l}\right)\left(1 + \sum_{j=1}^{L-M} \alpha_j z^{-j}\right) \quad (10)\end{aligned}$$

where $\mathcal{C}_b^a$ denotes the binomial coefficient

$$\mathcal{C}_b^a = \begin{cases} \binom{a}{b} & \text{for } a \geq b \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Notice that $\tilde{h}(z)$ corresponds to the following impulse response $1, -h_1, \ldots, -h_L$. Since the first term is fixed to one, this suggests that

$$\|H\|^2 = \|\tilde{H}\|^2 - 1$$

where $\tilde{H} = [1\ -h_1\ \cdots\ -h_L]^t$ is the vector containing the modified impulse response. We therefore conclude that minimizing $\|H\|^2$ is equivalent to minimizing $\|\tilde{H}\|^2$.

Consider now the particular form of $\tilde{h}(z)$ defined in (10). Multiplication in the $\mathcal{Z}$-transform domain corresponds to convolution in the time domain, therefore we can write

$$\tilde{H} = \mathbf{T}\begin{bmatrix} 1 \\ A \end{bmatrix}$$

where $A = [\alpha_1\ \cdots\ \alpha_{L-M}]^t$ is the parameter vector and $\mathbf{T}$ is a convolution matrix of dimensions $(L+1)\times(L-M+1)$ of the form

$$\mathbf{T} = \begin{bmatrix} \mathcal{C}_0^M & 0 & \cdots & 0 \\ -\mathcal{C}_1^M & \mathcal{C}_0^M & \cdots & \cdots \\ \vdots & \vdots & \cdots & \vdots \\ (-1)^M\mathcal{C}_M^M & \vdots & \cdots & \vdots \\ 0 & (-1)^M\mathcal{C}_M^M & \cdots & \vdots \\ \vdots & \vdots & \cdots & (-1)^M\mathcal{C}_M^M \end{bmatrix}.$$

If we form the norm of $\tilde{H}$, we obtain

$$\|\tilde{H}\|^2 = [1\ A^t]\mathbf{T}^t\mathbf{T}\begin{bmatrix} 1 \\ A \end{bmatrix} = [1\ A^t]\mathbf{Q}\begin{bmatrix} 1 \\ A \end{bmatrix},$$

where $\mathbf{Q} = \mathbf{T}^t\mathbf{T}$ has dimensions $(L-M+1)\times(L-M+1)$, it is symmetric, nonnegative definite and Toeplitz. The last property is easy to see due to the convolution matrix structure of $\mathbf{T}$. If the first column of the matrix $\mathbf{Q}$ is the vector $[q_0\ q_1\ \cdots\ q_{L-M}]^t$, then we have that

$$q_k = \sum_{n=k}^{M} (-1)^n \mathcal{C}_n^M (-1)^{n-k} \mathcal{C}_{n-k}^M = (-1)^k \mathcal{C}_{M-k}^{2M}$$

where we used the identities

$$\sum_{k=0}^{b} \mathcal{C}_k^a \mathcal{C}_{b-k}^c = \mathcal{C}_b^{a+c}$$

and $\mathcal{C}_b^a = \mathcal{C}_{a-b}^a$. Due to the symmetric Toeplitz structure of $\mathbf{Q}$, it is clear that the first column defines completely this matrix.

Solving now the minimization problem

$$\delta = \min_A [1\ A^t]\mathbf{Q}\begin{bmatrix} 1 \\ A \end{bmatrix},$$

can be seen to be equivalent to solving the following linear system of equations

$$\mathbf{Q}\begin{bmatrix} 1 \\ A \end{bmatrix} = \begin{bmatrix} \delta \\ \mathbf{0} \end{bmatrix} \quad (11)$$

As we can see (11) provides simultaneously the optimum vector A and the minimum squared norm δ.

Since $\mathbf{Q}$ is symmetric Toeplitz this is a classical system that can be solved by applying the *Levinson* algorithm. In particular we recognize that our problem corresponds to the "prediction" part of the Levinson scheme. It is exactly this method we propose to exploit to show the next theorem.

Theorem 2: The optimum coefficients α_k are given by the following formula

$$\alpha_k = \frac{\mathcal{C}_{M-1}^{M+k-1}\mathcal{C}_M^{L-k}}{\mathcal{C}_M^L}, \quad k = 1, \ldots, L-M. \quad (12)$$

Furthermore the corresponding MMSE, or minimum output noise power, is equal to

$$\mathrm{MMSE} = \sigma_w^2 \left(\frac{\mathcal{C}_M^{L+M}}{\mathcal{C}_M^L} - 1\right). \quad (13)$$

Proof: The proof is slightly involved, we will therefore only highlight its basic steps without presenting all details. We can prove our theorem using induction in the size of the problem. Let us therefore define the problem of size m. Denote with $\mathbf{Q}_{m+1}$ the size $m+1$ upper-left part of the $\mathbf{Q}$ matrix and with A_m and δ_m the corresponding optimum vector and power of the size m optimization problem,

$$\delta_m = \min_{A_m} [1\ A_m^t]\mathbf{Q}_{m+1}\begin{bmatrix} 1 \\ A_m \end{bmatrix},$$

which is equivalent to the linear system

$$\mathbf{Q}_{m+1}\begin{bmatrix}1\\A_m\end{bmatrix}=\begin{bmatrix}\delta_m\\\mathbf{0}\end{bmatrix},$$

then $\delta=\delta_{L-M}$ and $A=A_{L-M}$.

For the induction we will show the following general formula for the optimum vector and the optimum power of the size m problem. If $\alpha_k^m,\ k=1,\ldots,m$, denote the elements of the optimum vector A_m then

$$\alpha_k^m=\frac{\mathcal{C}_{M-1}^{M+k-1}\mathcal{C}_M^{M+m-k}}{\mathcal{C}_M^{M+m}},\quad k=1,\ldots,m, \tag{14}$$

whereas for the power δ_m we have

$$\delta_m=\frac{\mathcal{C}_M^{2M+m}}{\mathcal{C}_M^{M+m}}. \tag{15}$$

Relations (14) and (15) can be shown by induction on the order m. Key point in the proof constitutes the Levinson order recursion

$$\begin{aligned}A_m &= \begin{bmatrix}A_{m-1}\\0\end{bmatrix}+\kappa_m\begin{bmatrix}\bar{A}_{m-1}\\1\end{bmatrix}\\ \delta_m &= \delta_{m-1}(1-\kappa_m^2),\end{aligned}$$

where $\bar{A}_{m-1}$ denotes a vector with the same elements as A_{m-1} but in reverse order and κ_m the reflection coefficient defined as

$$\kappa_m=-\frac{q_m+\sum_{k=1}^{m-1}\alpha_k^{m-1}q_{m-k}}{\delta_{m-1}}.$$

After some mathematical manipulation κ_m can be shown to satisfy the following simple formula

$$\kappa_m=\frac{M}{M+m}.$$

It is in fact this latter expression that facilitates considerably the proof of the induction. The desired formulas of the theorem are obtained by substituting $m=L-M$ and remembering that $\|H\|^2=\|\tilde{H}\|^2-1$. This concludes the proof. ■

In Figure 1 we can see the normalized (with respect to σ_w^2) MMSE in db, as a function of M and filter length L. We observe that for large filter lengths there exists a constant relative difference in performance for filters of the same length and different values of M. This is true because one can show that for large L the MMSE behaves as

$$\text{MMSE}\sim\sigma_w^2\frac{M^2}{L};$$

therefore ratios of MMSEs corresponding to the same L depend only on the orders of the polynomials.

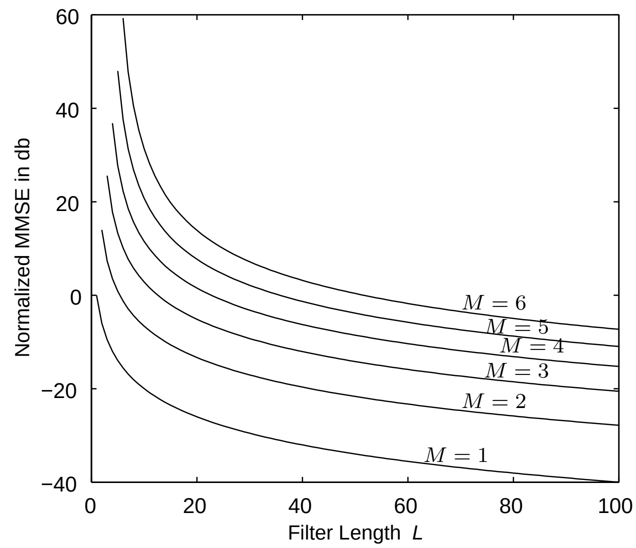

Figure 1: Normalized MMSE in db for polynomial orders $M-1=0,\ldots,5$ and filter lengths $L=M,\ldots,100$.

4 CONCLUSION

We have presented two methods for finding optimum FIR polynomial predictive filters. The first reveals an important qualitative structure of this filtering class that can be used for an efficient numerical computation of the optimum filter coefficients. The second, and most interesting one, leads to closed form expressions for the optimum filter coefficients and the final minimum output noise power. It should be noted that existing results containing closed form expressions consider only special (small) polynomial orders. On the other hand existing numerical solutions have complexity of the size of the filter length as compared to our method whose complexity is of the size of the order of the polynomial. The latter is usually significantly smaller than the former resulting in a significant computational gain.

References

[1] P. T. Harju and S. J. Ovaska, "Optimization of IIR polynomial predictive filter magnitude response," *Sign. Process.*, vol. 56, pp. 219-232, May 1997.

[2] P. T. Harju, T. I. Laasko and S. J. Ovaska, "Applying IIR predictors on Rayleight fading signal," *Sign. Process.*, vol. 48, no. 1, pp. 91-96, Jan. 1996.

[3] O. Vainio, "Design and efficient implementations of fredictive FIR filters," *IEEE Trans. on Instr. and Meas.*, vol. 44, no. 4, pp. 864-868, Aug. 1995.

[4] T. G. Campbell and Y. Neuvo, "Predictive FIR filters with low computational complexity," *IEEE Trans. on Circ. and Syst.*, vol. 38, no. 9, pp. 1067-1071, Sep. 1991.

[5] P. Heinonen and Y. Neuvo, "FIR-median hybrid filters with predictive FIR substructures," *IEEE Trans. on Acoust. Speech and Sign. Process.*, vol. 36, no. 6, pp. 892-899, June 1988.