ORTHOGONALITY EVALUATED IN THE FREQUENCY
DOMAIN: A NEW AND SIMPLE TOOL FOR DERIVING
OPTIMAL IIR-FILTERS

Anders Ahlén and Mikael Sternad
Automatic Control and Systems Analysis Group
Department of Technology, Uppsala University
P O Box 534 S-751 21 Uppsala, Sweden

Abstract

A new method for deriving mean square optimal IIR - filters is presented. It is based on
transfer function parametrizations and rests on orthogonality, evaluated in the frequency
domain. In contrast to other, well known, methods it is simple to use. It is applicable to
both scalar and multivariable filtering and control problems. The method is illustrated for
the well known, scalar, output filtering problem. Beside this, a brief discussion on adaptive
input estimation and equalization is included.

1 Introduction

In the project “Adaptive Wiener filters for control and signal
processing”, the main goal is to develop and apply meth-
ods useful to both fields. The main efforts are focused on
equalization (or input estimation) and noise cancelling (or
feedforward control). The work is intended to be of both
theoretical and practical nature. In this paper, we will, in
Section 2, present a new, simple and useful technique for
deriving optimal IIR-filters. In Section 3, some problems
which actualized the need of deriving optimal IIR-filters
are discussed briefly. Possible future directions of research,
leading to adaptive filters, are indicated.

2 Derivation of optimal IIR-filters

In many signal processing and control problems, it is de-
sirable to obtain optimal filters or regulators. Normally
quadratic criteria are minimized, since second order statis-
tics is usually of interest. For filtering problems, using sta-
tionary input-output data, this is known as “Wiener filter-
ing”.

Using infinite future data, nonrealizable Wiener filters are
straightforward to obtain. The solution is also straightforward in many problems when a pre-specified FIR-filter
parametrization and a finite amount of future data is used.
However, it is often unsatisfactory to be restricted to work
either with optimal filters which are not exactly realizable, or
with realizable filters with a prespecified, and often subopt-
imal, FIR-structure. The problem of deriving realizable, stable and explicit solutions with more general IIR-filter
parametrizations is far from trivial. (For example, it must
be decided which filter structure to use for a particular prob-
lem.) For such problems, the following four methods are
known:

1. The problem may be transformed to state space form.
A (stationary) Kalman filter may then be designed.

2. The classical Wiener filtering approach is to use varia-
tional arguments, to obtain frequency functions whose
causal parts, \( 
\Sigma_{0,0} \), are used. These parts are evaluated by residue calculus. See e.g. [2] and [3], chapter
13.

3. The polynomial approach, pioneered by Kučera [1],
[10] provides filters directly in polynomial form. The
equations defining optimality are derived by "completing
the squares" in the quadratic criterion. Optimal
filters are then designed by solving these equations.
They usually consist of a spectral factorization and a
single or two coupled linear polynomial equations.

4. Optimal filters in polynomial form can also be derived
by differentiating the criterion with respect to the fil-
ter coefficients and ensuring that all sensitivity func-

In this context, the signal models are assumed to be given
in transfer function form. While straightforward in prin-
ciple, a detour via state space methods then tends to reduce
the physical insight. State space solutions are also rather
complicated, in particular when coloured noise is present
and smoothing filters are sought. Use of the frequency
domain/polynomial methods 2 - 4 above is unfortunately
rather cumbersome. This applies especially for multivari-
able problems. See [1], [2] and [3].

Here, we will present a new derivation technique, to be used
when optimal IIR-filters are sought. It leads to equations for
calculating the filters which are equivalent to those derived
from the polynomial approaches 3. and 4. above. The
derivation technique is, however, much simpler, especially
in the multivariable case. It basically rests on orthogon-
ality, evaluated in the frequency domain. While the technique
can be applied just as well in LQG control problems, we will,
for clarity, restrict the discussion to estimation problems in
this paper.

2.1 Outline of the technique

Assume a linear stochastic system to be parametrized by
discrete time transfer functions and ARMA models. It gen-
erates a measurement signal \( y(t) \) and a desired response \( f(t) \). For simplicity, in the sequel, we assume these signals to be real-valued and scalar. Our aim is to optimize a linear filter which operates on \( y(t + m) \) and estimates \( f(t) \),

\[
f(t | t + m; \theta) = \frac{Q(q^{-1})}{R(q^{-1})} y(t + m)
\]

(2.1)

Depending on \( m \), (2.1) constitutes a predictor \( (m < 0) \), a filter \( (m = 0) \) or a fixed lag smoother \( m > 0 \). The estimator is designed to minimize the quadratic criterion

\[
V(\theta) = \mathbb{E}[e(t, \theta)^2]
\]

(2.2)

under the constraint of causality and stability of the filter \( Q(q^{-1})/R(q^{-1}) \), where

\[
e(t, \theta) \triangleq f(t) - \hat{f}(t | t + m; \theta)
\]

(2.3)

We require \( e(t, \theta) \) to be stationary. The polynomials \( Q(q^{-1}) = Q_0 + Q_1 q^{-1} + \cdots + Q_m q^{-m} \) and \( R(q^{-1}) = 1 + r_1 q^{-1} + \cdots + r_m q^{-m} \) have the backward shift operator \( q^{-1} y(t) = y(t - 1) \) as arguments. The coefficients are collected in the parameter vector \( \theta = (Q_0, \ldots, Q_m, r_1, \ldots, r_m) \). The minimizing argument of (2.2) is denoted \( \theta^* \). For any polynomial \( P = P(q^{-1}) \), we denote \( P_* = P(q) \). In the frequency domain, the complex variable \( z \) is substituted for the forward shift operator \( q \), defining the stability region to be located inside \( |z| < 1 \).

**Figure 1.** The filtering problem.

If we choose some filter \( Q/R \), will it be possible to find a better one? Introduce an alternative estimator candidate \( \hat{d}(t | t + m) \), such that

\[
\hat{d}(t | t + m) = \frac{Q(q^{-1})}{R(q^{-1})} y(t + m) + n(t)
\]

(2.4)

where \( n(t) \) is an arbitrary stationary term, free to choose. It may depend on a linear combination \(^1\) of measured data up to time \( t + m \). For example, \( n(t) \) may be described by \( n(t) = (C/H) y(t + m) \). Substituting \( \hat{d}(t | t + m) \) for \( f(t | t + m) \) in (2.3) gives the criterion

\[
\bar{V}(\theta) = \mathbb{E}[e(t)^2] - 2\mathbb{E}[e(t)n(t)] + \mathbb{E}[n(t)^2]
\]

(2.5)

Depending on the choice of \( n(t) \), \( \bar{V}(\theta) \) might attain a greater or lower value, compared to \( V(\theta) \). The key idea is now to choose the filter polynomials \( Q \) and \( R \) such that \( e(t) \) becomes orthogonal to any admissible additional signal \( n(t) \). This means that the mixed term in (2.5) is set to zero. This condition will be sufficient to determine the polynomial degrees and the coefficient values of \( Q \) and \( R \) in (2.1), (2.4)

\(^1\)Since we consider only linear filters, the additional term \( n(t) \) should be a linear combination of data. Since the filtering error must be stationary, nonstationary \( n(t) \) are excluded.

uniquely. With \( Q \) and \( R \) chosen such that \( E[e(t)n(t)] = 0 \), it is then obvious that the choice \( n(t) = 0 \) will minimize (2.5) and therefore also (2.2). Since no modification of the filter can improve the criterion value, the derived \( Q \) and \( R \) polynomials are optimal and \( E[e(t, \theta^*)^2] \) is the minimal value.

Since \( e(t) \) and \( n(t) \) are stationary, the mixed term in (2.5) can be expressed with the aid of Parseval’s formula

\[
E[e(t)n(t)] = \frac{1}{2\pi j} \int_{|z|=1} \phi_{en} \frac{dz}{z}
\]

(2.6)

where \( \phi_{en} \) is a rational function in \( z \) and \( z^{-1} \) which depends on the filter polynomials \( Q \) and \( R \). Let \( \phi_{en} \) be defined as

\[
\phi_{en} = \frac{T(z, z^{-1})}{S_+ (z, z^{-1}) S_-(z, z^{-1})}
\]

(2.7)

where the polynomials \( S_+ \) and \( S_- \) have all zeros strictly inside and outside the unit circle, respectively. In order to force (2.6) to zero, we require

\[
T(z, z^{-1}) = z L_0(z) S_-(z, z^{-1})
\]

(2.8)

where \( L_0(z) \) is an arbitrary polynomial only in \( z \) (not in \( z^{-1} \)). Then, the integrand \( L_0(z)/S_- (z, z^{-1}) \) in (2.6) has no poles inside the integration path, and the integral vanishes.

When the measurable signal (or signals) is a sum of independent stochastic sequences, for example,

\[
y(t) = C(q^{-1}) z(t) + M(q^{-1}) v(t)
\]

(2.9)
a spectral factorization\(^2\) must often be introduced:

\[
r \beta \beta^\top = CC^\top NN^\top + \rho DD^\top MM^\top \quad ; \quad \rho = \lambda_u / \lambda_1
\]

(2.10)

In (2.10), \( r \) is a scalar and \( \beta \) is the stable spectral factor. It is monic by construction. Expressing (2.7) in terms of the spectral factorization and (2.6), leads to one or two coupled linear polynomial equations in \( Q(q^{-1}) \), \( R(q^{-1}) \) and \( L(z) \). They will be polynomial equations in both \( z \) and \( z^{-1} \). If a solution exists, it will be unique. For open loop problems, such as filtering, equalization and noise cancelling problems, we will end up with a single polynomial equation. The filter denominator \( R \) will, in such cases, be equal to the spectral factor \( \beta \), or have \( \beta \) as a factor.

The derivation procedure can be summarized as follows.

1. Define (if needed) one or several spectral factorizations.

2. Calculate \( e(t) \) in terms of the involved polynomials and add a signal \( n(t) \) to the filter output. Use Parseval’s formula to express the mixed term in (2.5), using the spectral factor(s).

3. Cancel the denominator poles of (2.7) inside the unit circle by means of factors in the numerator. This leads to one or two coupled polynomial equations, to be solved for the estimator polynomials.

\(^2\)In some problems, such as in the derivation of the optimal DFE presented in [4, 5], the spectral factor is not needed. This situation appears when measurements are available in every branch driven by a single noise source.
2.2 Example: The scalar output filtering problem.

Let the output measurements be described by (2.9), where \( D \) and \( N \) are stable. The signal \( s(t) = (C/D)e(t) \), depicted in Figure 2, is to be estimated.

\[
\begin{align*}
\text{signal} & \quad \text{description} \quad \text{filter} \\
\varepsilon(t) & \quad M^{-1} \quad N^{-1} \quad Q(s^{-1}) \quad R(s^{-1}) \quad d(t) + m \\
\frac{C}{D} & \quad (s^{-1}) \quad \frac{G}{H} \quad \frac{N}{R} \quad v(t) \quad \frac{Q}{R} \\
\frac{C}{D} & \quad (s^{-1}) \quad \frac{G}{H} \quad \frac{N}{R} \quad v(t) \quad \frac{Q}{R} \\
\end{align*}
\]

Figure 2. The scalar output filtering problem \((m = 0)\). The signal \( s(t) \) is to be estimated from \( y(t) \). \( \rho = \lambda_1/\lambda_s \).

The spectral factorization (2.10) is easily derived from the innovations model of \( y(t) \). It is assumed to be stable, i.e. the right-hand side has no zeros on the unit circle. We seek a stable causal linear filter \( Q/R \) which minimizes \( Ee(t)^2 \).

The error signal is given by

\[
e(t) = \frac{C(R - Q)}{DR} e(t) - \frac{MQ}{NR} v(t)
\]

and the additional, arbitrary signal is

\[
n(t) = \frac{G}{H} v(t) = \frac{G}{H} \left( \frac{C}{D} e(t) + \frac{M}{N} v(t) \right)
\]

with \( G \) and \( H \) undetermined, but \( H \) stable. The mixed term in (2.5) is then directly found to be

\[
Ee(t)n(t) =
\]

\[
= E \left( \frac{C(R - Q)}{DR} e(t) \frac{CG}{DH} e(t) \right) - E \left( \frac{QM}{RN} v(t) \frac{MG}{NH} v(t) \right)
\]

\[
= \frac{\lambda_s}{2\pi i} \oint_{|z|=1} \left( \frac{R - Q}{RDD,NN} - \rho QMM,DD \cdot \frac{G, dz}{H, z} \right) + \frac{\lambda_s}{2\pi i} \oint_{|z|=1} \left( \frac{RCC,NN - Qr\beta \beta}{RDD,NN} \cdot \frac{G, dz}{H, z} \right)
\]

In the second equality, Parseval's formula was used and in the last, the spectral factorization (2.10) was inserted. Thus, we have completed steps 1 and 2 and now proceed with step 3. Since \( R, D, N \) and \( H \) are assumed to be stable polynomials, they will have all their zeros inside the unit circle \((D, N, H, \text{will have all zeros outside the unit circle})\). The numerator of the integrand (2.7) then fulfills (2.8) if and only if

\[
RCC,NN - Qr\beta \beta = zL, RDN
\]

Obviously, \( Q \) must have \( N \) as a factor, \( Q = Q_1, N \). Rearranging the factors in (2.13) gives

\[
R(CC, N, - zL, D) = Q_1r\beta \beta.
\]

Since \( R \) must be stable and \( z^{-\delta} \beta, (z) \) is unstable while \( Q_1 \) is a filter polynomial, choose \( R = \beta \). Then, \( Q_1 \), together with \( L, \) can be found as the solution to the linear polynomial equation

\[
CC, N = tr, Q_1 + zDL,
\]

with degrees

\[
nQ_1 = \max(nc, nd + 1) \quad nL_r = n\beta - 1
\]

Equation (2.15) is solvable since \( D(z^{-1}) \) (stable) and \( z^{-\delta} \beta, (z) \) (unstable) cannot have common factors. The solution \((Q_1, L_r)\) is unique.

Since the second term in (2.5) becomes zero and the remaining, third term, is quadratic, the optimal choice of \( n(t) \) is zero. Obviously the solutions above is then optimal. The filter is calculated by solving (2.10) for \( \beta \), and then solving (2.15) for \( Q_1 \) and \( L_r \).

2.3 Remarks and interpretations

Note that causality \(^3\) of the filter is guaranteed by requiring \( Q \) and \( R \) to be polynomials in the backward shift operator \( q^{-1} \). The stability of the filter is simple to assure, with the unique choice \( R = \beta \) in (2.14). This choice is unique in the sense that any other choice \((\beta = \beta_1, \beta_2, R = \beta_1, \deg \beta_2 > 0)\) would make the polynomial equation resulting from (2.14) unsolvable, unless \( \beta_2 \) were a common factor of \( C \) and \( D \).

The polynomial \( L_r \) does not appear in the filter. It is, however, important that \( L_r \) should be a polynomial in \( z \) (or in \( \bar{z} \)), not in \( q^{-1} \) (or \( z^{-1} \)). If \( L_r \) contained the argument \( z^{-1} \), the integrand (2.7) would have poles in the origin, and the integral would not vanish. The requirements that \( L_r \) should be a polynomial only in \( z \), while \( Q_1 \) is a polynomial only in \( z^{-1} \), lead to a unique solution to the linear polynomial equation (2.15), with degrees (2.16). (These degrees are determined by the requirement that \( L_r(z) \) should cover the maximal occuring power of \( z \), while \( Q_1(z^{-1}) \) covers the maximal power of \( z^{-1} \).) This gives a linear system of equations, with equal number of equations and unknowns.

The derivation technique provides a solution with an optimal filter structure. Above, it is an IIR-filter \( Q_1, N/\beta \). It is important to postulate a filter with the most general structure from the outset. If we had specified the filter to be FIR \((R = 1)\), equation (2.14) would have been unsolvable.

A filter attains the minimal estimation error if and only if it contains the same coprime factors as the filter derived above. The solution is thus unique, modulo possible stable common factors in the filter. (If, for example, \( D \) and \( N \) have common factors, these factors will appear in \( \beta \). They are thus common factors of the filter \( Q_1, N/\beta \).)

\(^3\)Causality" does not exclude consideration of smoothing problems. It only implies that optimal filters do not require data further into the future than defined by the prescribed smoothing lag \( m \).
The technique can also be applied to systems where $y(t)$ and/or $f(t)$ are nonstationary, under some technical conditions which correspond to detectability. The important point is that the signals $e(t)$ and $n(t)$, which appear in the criteria (2.2) and (2.3), must be stationary.

The idea of adding an arbitrary term, $n(t)$, to the filter $Q/R$ in (2.4), is useful also when candidates for optimal filters ($Q/R$) have been derived by other means. The idea is then to prove, by contradiction, that $n(t)$ must be zero. This approach, originally used in [12], has been applied in some of our earlier work [4,5,7,9,11]. In contrast to this, the derivation technique presented above is a constructive way to find optimal IIR-filters.

The suggested derivation procedure also applies to multivariable filtering and control problems. As an example, the multivariable filtering result given in [6] is derived in just a few lines with the new method. The need for a simple derivation technique, as e.g. described above, has been actualized when investigating explicit and optimal solutions to some filtering problems described below.

3 Adaptive input estimation and equalization

In [4] and [5], an optimal and explicit IIR - decision feedback equalizer was derived for a general channel structure with coloured noise. For practical reasons, it is desirable to create an adaptive scheme. In order to do this, an adaptive input estimator has been derived and tested in [7]. Under certain identifiability conditions, given in [8], the linear deconvolution problem described in [9], has been solved on-line. The results in [7], based on an explicit scheme, are promising for adaptive DFE's and LFE's under certain circumstances.

Adaptive equalization based on the DFE described in [4], [5] may be developed along the following lines.

1. This step is a common part to 2a, 2b. The channel structure introduced in [4], [5] is depicted in Figure 2. In "reference mode", $e(t)$ is a known sequence. During this period, $(C, D, M, N)$ are estimated by means of system identification. Based on this channel estimate, the DFE described in [4], [5] is calculated. This DFE is used as an initial estimate to step 2, the "equalization mode".

2a. The results in [7] is adopted. Under certain conditions on the channel and measurement noise [8], the $(C, D, M, N)$ polynomials can be estimated adaptively, from the received sequence only. Based on these estimates, the DFE is updated.

2b. Direct adaptation of the filters in the IIR-DFE structure in [4], [5]. In contrast to step 2a, the equalizer update is then based on decisioned data.

Which line to follow is a topic for further research. It can be noted, though, that the basic principle to be used in 2a. works well, as reported in [7]. It is thus reasonable to assume that it will work for the equalization problem if, loosely speaking, the channel and measurement noise is not too extreme.

4 Conclusions

A new method for deriving mean square optimal IIR-filters has been presented. It is transfer-function based and utilizes orthogonality evaluated in the frequency domain, constructively. The method is simple and straightforward to use and applies to scalar and multivariable prediction, filtering or smoothing problems. The optimal filters are obtained by solving a spectral factorization and a polynomial equation. A simple example illustrates the main ideas. The need for deriving optimal IIR-filters has been actualised when investigating e.g. optimal equalizers. We have briefly discussed future directions of research on adaptive equalizers.

References