Constrained Wiener Filtering

Bernard Picinbono, Michel Bouvet

Constrained Wiener Filtering

Bernard Picinbono, *Fellow, IEEE* and Michel Bouvet *Member, IEEE*
Laboratoire des Signaux et Systèmes (L2S), Division Signaux

The pdf copy of the final published text can be obtained from the first author at the following address : bernard.picinbono@lss.supelec.fr

Reference of this text : IEEE Trans. Inf. Theory, Vol. 33, Num. 1 pp. 160 - 166, 1987.

Abstract

As it is is well-known, the introduction of causality conditions in Wiener filtering problems completely changes their solution. A constraint such as causality can be presented as a particular reduction of the observation space, and the constrained filter can always be obtained by projection onto this space. However, it is sometimes simpler to use an indirect method which gives the impulse response of the constrained filter by an appropriate modification of the unconstrained response. This method is presented and applied to many examples. In particular, the structure of constrained prediction filters is analyzed, and it is shown that the constrained innovation can be expressed in terms of the unconstrained one by an appropriate filter.

Manuscript received

B. P. and M. B. are with the Laboratoire des Signaux et Systèmes (L2S), a joint laboratory of the C.N.R.S. and the École Supérieure d'Électricité, Plateau de Moulon, 3 rue Joliot–Curie 91192, Gif sur Yvette, France. The L2S is associated with the University of Paris–Sud, France. E-mail: bernard.picinbono@lss.supelec.fr, .

I. INTRODUCTION

Wiener filtering is a well-known technique for the extraction of a signal from a noisy observation [l]-[3]. In the discrete-time and stationary case, the output of a Wiener filter is

$$\hat{x}(k) = \sum_{l=-\infty}^{+\infty} h(l)y(k-l), \tag{1}$$

where $x(k)$ and $y(k)$ are the signal to be estimated and the observation respectively. The impulse response (IR) $h(k)$ of this filter is calculated in such a way that $\hat{x}(k)$ is the best linear mean square estimate of $x(k)$, and if $x(k)$ and $y(k)$ are zero mean and stationary, $h(k)$ is the solution of the Wiener-Hopf equation

$$\sum_{l=-\infty}^{+\infty} h(l)r_y(k-l) = r_{xy}(k), \tag{2}$$

where the r functions are the auto and cross correlation functions defined, respectively, by

$$r_y(k) = E[y(n)y(n-k)], \;\; r_{xy}(k) = E[x(n)y(n-k)]. \tag{3}$$

By using the z transforms of (2) and (3), we obtain the transfer function $H_0(z)$ of the Wiener filter $h_0(k)$ as

$$H_0(z) = [R_y(z)]^{-1}Rxy(z). \tag{4}$$

In many practical situations we have to solve the same problem *with a particular constraint* on the IR h(k). The best known case corresponds to the *causality constraint*, in which we impose in (1) that $h_c(k) = 0$ for $k < 0$. In this case the Wiener-Hopf equation (2) becomes

$$\sum_{l=0}^{+\infty} h_c(l)r_y(k-l) = r_{xy}(k), \;\; k \geq 0, \tag{5}$$

which can be solved by many means. The simplest one uses spectral factorization [4], [5], and the transfer function obtained is

$$H_c(z) = \alpha B(z)[B(z^{-1})R_{xy}(z)]_+, \tag{6}$$

where $B(z)$ is deduced from the strong factorization of $R_y(z)$, which means that $B(z)$ is a minimum-phase filter, α is a constant, and $[]_+$ means the causal part of the term in brackets. Two facts then appear immediately.

1) The constraint on the IR completely changes the Wiener filter.

2) The solution of the constrained problem can be obtained directly and without using the solution of the unconstrained problem.

These conclusions are quite general for other kinds of constraints. However, in many cases it appears more interesting, particularly from a computational point of view, to solve the constrained problem by using the unconstrained solution, which is often known.

Let us, for example, consider the interpolation problem in which the only constraint in (1) is $h(k) = 0$ for $k = 0$. The corresponding Wiener-Hopf equation is

$$\sum_{l \neq 0} h_c(l) r_y(k-l) = r_{xy}(k), \quad k \neq 0, \tag{7}$$

and the direct solution of this problem is not at all simple. By using the unconstrained solution (4) it can be reduced to one linear equation, and this indirect solution appears extremely simple and attractive.

Many other cases of such constraints exist for which the indirect solution is simpler to obtain, and the aim of this paper is to present and discuss some examples and applications of constrained Wiener filtering using indirect methods.

In the signal processing or automatic control area, P-step prediction is a good example of the constrained Wiener filtering problem. In this case the constraint on the IR $h(k)$ is characterized by$h(k) = 0$ for $0 \leq k \leq P-1$. This problem is important in many areas of adaptive control [6] and has been partially solved by Astrom [7, p. 165]. This solution does not use a systematic approach of constrained estimation problems, and the result is obtained after a long calculation and only under limited assumptions. In the following we will present a general solution to this problem and discuss some interesting properties of the general constrained predictor. Other constrained problems are studied in [8] with a completely different approach.

Note that constrained estimation problems have been considered in areas other than signal processing or information theory. Similar problems appear in mathematical geology under the name kriging, and an excellent review can be found in [9]. More recently, a paper was presented discussing the problem of linear constraints on least square methods from a computational point of view [10].

Finally, we note that many kinds of constraints can be imposed on the IR $h(k)$ of (1), and we will restrict our study to linear constraints, as defined in the next section. For example, an amplitude constraint such as $|h(k)| < M$, which is necessary for some applications, is nonlinear and cannot be studied by the same procedure.

II. Statement and Solution of the Problem

A. Notations

Let us call F the space consisting of functions $h(k)$ (representing IR's of filters) with z transforms

$$H(z) = \sum_{k=-\infty}^{+\infty} h(k) z^{-k} = \sum_k h(k) z^{-k}, \tag{8}$$

which are convergent on the unit circle (u.c.). Two filters of F are orthogonal, denoted by the symbol $\perp$, if

$$\sum_k h(k)g(k) = 0 \text{ or, equivalently, if } [F(z)G(z^{-1})]_0 = 0, \tag{9}$$

where $[A(z)]_0$ means $a(0)$, coefficient of z^0 in the z expansion of $A(z)$.

Let us denote by H the Hilbert space of zero-mean second-order random variables with the scalar product $E[uv]$. In this space the orthogonality, specified by the symbol $\perp\!\!\!\perp$, means uncorrelatedness. We also denote by H_y the observation space which is the Hilbert subspace of H generated by the observation $y(k)$. A filter of F is subject to a linear constraint if its IR $h(k)$ belongs to a linear subspace S of F. This linear constraint generates a Hilbert subspace $H_y(S)$ of the observation space H_y defined by

$$H_y(S) = \left\{ s(k)\ | s(k) = \sum_l g(l)y(k-l), g(k) \in S \right\}. \tag{10}$$

Let us now introduce the concept of constrained z transforms. If $h(k) \in S$, its z transform is denoted $H_S(z)$. From the projection theorem any IR $h(k)$ can be decomposed into the sum

$$h(k) = h_S(k) + h_{\not S}(k),\ h_S(k) \perp h_{\not S}(k) \tag{11}$$

and this decomposition is unique. Of course, the subspace $\not S$ is orthogonal to S. In terms of z transforms, this decomposition becomes

$$H(z) = H_S(z) + H_{\not S}(z),\ [H_S(z)H_{\not S}(z^{-1})]_0 = 0. \tag{12}$$

B. Direct Solution of the Constrained Wiener Filtering

The problem is to find $h(k) \in S$ such that $E[x(k) - \hat{x}(k)]^2$, where $\hat{x}(k)$ is given by (1), is minimum. This problem can be solved by variational methods, but it is simpler to use the orthogonality principle [2, p. 3361. From this principle $h(k)$ is deduced from the orthogonality equation

$$x(k) - \hat{x}(k) \perp\!\!\!\perp H_y(S). \tag{13}$$

By using the definition of $\hat{x}(k)$ from (1) and of the orthogonality in H, we immediately obtain an equivalent condition

$$\sum_k g(k) \left[\sum_l h(l)r_y(k-l) - r_{xy}(k) \right] = 0,\ \forall g(k) \in S. \tag{14}$$

In other words, the constrained Wiener filter has an IR $h(k)$ belonging to S and such that

$$\sum_l h(l)r_y(k-l) - r_{xy}(k) \perp S. \tag{15}$$

If S is of finite dimension, this equation can be easily solved. Otherwise, the problem is much more complex, and no general method exists for its solution. In the z domain the problem is to find a function $H_S(z)$ such that

$$[H_S(z)R_y(z) - R_{xy}(z)]_S = 0. \tag{16}$$

and this is the equation we use to find the causal solution (6).

C. *Indirect Solution of the Problem*

To any function $H(z)$ let us associate

$$\bar{H}(z) = R_{xy}(z) - R_y(z)H(z). \tag{17}$$

Assuming that $R_y(z) \neq 0$, at least in the vicinity of the u.c., we can write

$$H(z) = C(z)[R_{xy}(z) - \bar{H}(z)], \tag{18}$$

where

$$C(z) = [R_y(z)]^{-1}, \tag{19}$$

We deduce immediately that

$$H(z) = H_0(z) - C(z)\bar{H}z), \tag{20}$$

where $H_0(z)$ is the transfer function of the unconstrained Wiener filter given by (4).

With these notations it appears that our problem consists of finding a function $H_S(z)$ such that $\bar{H}_{\not S}(z) = 0$. This can also be expressed in the form: find a function $\bar{H}_{\not S}(z)$ such that

$$[\bar{H}_{\not S}(z)C(z) - H_0(z)]_{\not S} = 0. \tag{21}$$

It is clear that this equation is very similar to (16). However, as we will verify in the following, it appears that in many problems it is much simpler, particularly from a computational point of view, to solve (21) than (16). To calculate $H(z)$ it is sufficient to insert the solution of (21) into (20), which shows clearly that the indirect method needs the *a priori* knowledge of the unconstrained solution $H_0(z)$.

D. *Finite Case*

If either S or $\not S$ are of finite dimension, (16) or (21) can be converted to a finite set of linear equations of the same dimension. As an example, that is the case of interpolation giving (7). Here the dimension of S is unity, and clearly, the indirect solution must be applied.

The case where S has a finite dimension is very well-known, and the direct solution can be applied immediately. Accordingly, let us consider the case where $\not S$ has a finite dimension and can be spanned by a set of P orthonormal functions $v_i(k)$. If we expand $\bar{H}_{\not S}$ in terms of the z transforms of these functions, we have

$$\bar{H}_{\not S} = \sum_{i=1}^{P} \bar{h}[i] V_i(z). \tag{22}$$

It is shown in Appendix I that the P coefficients $\bar{h}[i]$ are the solutions of the system

$$\sum_{i=1}^{P} c[i,j]\bar{h}[j] = h_0[i],\ 1 \leq i \leq P, \tag{23}$$

where

$$c[i,j] = \sum_{k,l} v_i(k)c(k-l)v_j(l), \tag{24}$$

and

$$h_0[i] = \sum_{k} h_0(k)v_i(k). \tag{25}$$

In (24) $c(k)$ is, of course, the inverse z transform of $C(z)$ defined by (19). In the time domain, the IR of the constrained Wiener filter is deduced from (20), which gives

$$h(k) = h_0(k) - \sum_{l} c(k-l)\bar{h}(l) = h_0(k) - \sum_{i=1}^{P} \bar{h}[i] \sum_{l} c(j-l)v_i(l), \tag{26}$$

where $h_0(k)$ is deduced from (4) and (19) by

$$h_0(k) = \sum_{l} c(k-l)r_{xy}(l). \tag{27}$$

E. Expression of the Minimum Mean Square Error

The mean square error is the variance of$x(k) - \hat{x}(k)$, and from (13) it follows that

$$\epsilon^2 = E\{[x(k) - \hat{x}(k)]^2\} = \sigma_x^2 - \sigma_{\hat{x}}^2, \tag{28}$$

where σ^2 denotes variance. From (1) we deduce that

$$\sigma_{\hat{x}}^2 = [R_y(z)H(z)H(z^{-1})]_0\ , \tag{29}$$

where $H(z)$ is given by (20). Bt using the fact that $R_y(z)C(z) = 1$ and that $H(z) = H_S(z)$, $\bar{H}(z) = \bar{H}_{\not S}(z)$ and then are orthogonal, we deduce easily that

$$\epsilon^2 = \epsilon_0^2 + \epsilon_c^2, \tag{30}$$

where ϵ_0^2 is the minimum error without constraint and ϵ_c^2 an additional error due to the constraint and given by

$$\epsilon_c^2 = [C(z)\bar{H}(z)\bar{H}(z^{-1})]_0. \tag{31}$$

In the time domain this additional error is, of course,

$$\epsilon_c^2 = \sum_{k,l} \bar{h}(k)c(k-l)\bar{h}(l). \tag{32}$$

and in the finite case considered in Section II-D it becomes

$$\epsilon_c^2 = \sum_{k,l=1}^{P} \bar{h}(i)c(i,j)\bar{h}(j). \tag{33}$$

III. Constrained Causal Wiener Filters

The problem of Wiener filtering with a causality constraint has been widely studied, and in the stationary case it is solved by the spectral factorization technique. A constrained causal filter is a causal filter which satisfies some additional constraints. We have seen that the filter for the P step prediction must be causal but also satisfy the condition $h(k) = 0$ for $0 \leq k < P$. Of course many possible other constraints exist. For example, in detection problems it is sometimes necessary to make an estimation with the noise alone reference (NAR) property [ll], [12], which means that no contamination of the observation by the signal occurs. Then the corresponding causal filter-must satisfy the causality and the NAR constraints.

A. Causality Constraint Only

In this case the subspace S is the subspace S_+ of functions $a(k)$ vanishing for $k < 0$. These functions and their constrained z transforms are denoted $a_+(k)$ and $A_+(z)$ respectively. Then the causal Wieser filtering problem consists of finding a pair of functions $|H, \bar{H}]$ such that $H = H_+$, and $\bar{H} = H_-$.

For this purpose let us introduce the innovation filter of $y(k)$ and its transfer function $B(z)$. It is the filter calculating $\tilde{y}(k) = y(k) - \hat{y}(k)$, where $\hat{y}(k)$ is the one step prediction with infinite past of $y(k)$. The variance of this innovation is the prediction error η^2, and, as the innovation is a white noise, we have

$$\eta^2 = R_y(z)B(z)B(s^{-1}). \tag{34}$$

Furthermore, note that the innovation filter is a minimum-phase filter, and its IR $b(k)$ satisfies $b(0) = 1$.

By using (18), (19), and (34), we can write

$$H(z) = (\eta^2)^{-1}B(z)T(z), \tag{35}$$

where $T(z)$ is defined by

$$T(z) = B(z^{-1})[R_{xy}(z) - \bar{H}(z)]. \tag{36}$$

To determine $H(z)$, it is equivalent to work with $\bar{H}$ or with T, provided that $H_{-}(z) = 0$. From (35) this condition means that $[BT]_{-} = 0$. If we decompose T in its causal and anticausal parts, $T = T_{-} + T_{+}$, and if we notice that $[BT_{+}]_{-} = 0$ because B is causal, we get

$$[B(z)T(z)]_{-} = [B(z)T_{-}(z)]_{-} = 0. \quad (37)$$

This means that BT_{-} is a causal filter. However, as $B(z)$ is a strictly minimum-phase filter, its inverse is also causal, and it follows that T_{-} is causal, which is impossible. Then the condition $H_{-}(z) = 0$ gives $T(z) = T_{+}(z)$. Using (36) and noting that $B(z^{-1})$ and $\bar{H}(z)$ are anticausal ($\bar{H} = H_{-}$ is the condition imposed), we get

$$T(z) = [B(z^{-1})R_{xy}(z)]_{+} \ , \quad (38)$$

which defines $H(z)$ by (35). That is the classical result which can be obtained by many other approaches and is indicated in (6).

B. Constrained Causal Filter

Denoting S_{+} and S_{-} as the subspaces of causal and anticausal functions, respectively, i.e., functions $a(k)$ vanishing for $k < 0$ or for $k \geq 0$, we impose on the impulse response $h(k)$ of the Wiener filter the constraint $h(k) \in S \subset S_{+}$. We can then define a subspace e_{+} of S_{+} such that

$$\not{S} = S_{-} \bigoplus e_{+} \ ; e_{+} \perp S_{-} \ . \quad (39)$$

As a consequence we have for any function $A(z)$

$$A_{\not{S}}(z) = A_{-}(z) + A_{e_{+}}(z), \quad (40)$$

and the condition $A_{\not{S}} = 0$ requires simultaneously

$$A_{-}(z) = 0 \ ; \ A_{e_{+}}(z) = 0. \quad (41)$$

We apply these properties to $H(z)$ which must satisfy $H_{\not{S}}(z) = 0$.

I) Consequences of $H_{-}(z) = 0$: They are the same as in Section III-A, but now $\bar{H} = H_{\not{S}} = \bar{H}_{-} + \bar{H}_{e+}$, which gives

$$T(z) = T_{+}(z) = [B(z^{-1})R_{xy}(z)]_{+} - [B(z^{-1})\bar{H}_{e_{+}}(z)]_{+} \ , \quad (42)$$

2) Consequences of $\bar{H}_{e_{+}}(z) = 0$: From (35) it is sufficient to write

$$[B(z)T(z)]_{e_{+}} = 0, \quad (43)$$

which gives with (42)

$$\{B(z)\{[B(z^{-1})\bar{H}_{e_{+}(z)}] + [B(z^{-1})R_{xy}(z)]_{+}\}\}_{e_{+}} = 0. \quad (44)$$

This equation defines $\bar{H}_{e_{+}}(z)$ and then $H(z)$ by using (34) and (42).

C. Calculation of the Constrained Causal Filter

As indicated earlier, this calculation requires the solution of (44) which can be written

$$\{B(z)[B(z^{-1})\bar{H}_{e_+}(z)]\}_{e_+} = \{B(z)[B(z^{-1})R_{xy}(z)]_+\}_{e_+}. \tag{45}$$

D. Constrained Prediction Filters

Prediction problems are generally much simpler than estimation ones because $R_{xy} = R_y$, which greatly simplifies the resolution of (45). Moreover, it is important to notice that because $x = y$ the subspace e_+ contains necessarily the function $\delta[k]$, which means that the impulse response $h(k)$ of any predictor must satisfy $h(0) = 0$, otherwise, the prediction problem would become singular. Using $R_{xy}(z) = R_y(z) = \eta^2[B(z)B(z^{-1})]$, (45) becomes

$$\{B(z)[B(z^{-1})\bar{H}_{e_+}(z)]_+\}_{e_+} = \eta^2, \tag{46}$$

$$H(z) = 1 - G((z), \tag{47}$$

$$G(z) = (\eta^2)^{-1}B(z)[B(z^{-1})\bar{H}_{e_+}]_+ \ . \tag{48}$$

We can make three comments about these expressions. First, the filter does not depend on η^2, because this variance disappears between (46) and (48). This fact is quite natural because the structure of any predictor does not depend on any scaling factor in the observation. Then to solve (46), we can assume that $\eta^2 = 1$. Second, if we call $\hat{y}_c(n)$ and $\tilde{y}_c(n)$ the outputs of the filters H and G, respectively, we deduce

$$\tilde{y}_c(n) = y(n) - \hat{y}_c(n), \tag{49}$$

which is the definition of the constrained innovation. However, due to the constraint, this innovation has no reason to be a white noise. To find the true innovation of the signal $y(n)$, we must use the one-step predictor. In this case the subspace e_+ is the subspace of constant functions, which means that $\bar{H}_{e_+} = c$. By using (46) and (48); we easily find $G(z) = B(z)$, which is effectively the filter giving the innovation of $y(n)$. Third, we check that $h(0) = 0$, as indicated for any predictor.

References

[1] L. A. Wainstein and V. D. Zubakov, *Extraction of Signal from Noise*. Englewood Cliffs, NJ: Prentice-Hall, 1962.

[2] A. Papoulis, *Signal Analysis*. New York: McGraw-Hill, 1977.

[3] T. Kailath, *Lectures on Linear Least-Squares Estimation*, Wien, Austria, Springer Verlag, Int. Center for Mechanical Sciences, Courses and Lectures, 140, Udine, 1976.

[4] A. Blanc-Lapierre and B. Picinbono, *Fonctions Aléatoires.* Paris, France: Masson, 1981.

[5] D. Sakrison, *Communication Theory; Transmission of Waveforms and Digital Information.* New York: Wiley, 1968.

[6] G. Favier, “Etude comparative de prédicteurs auto-ajustables,’ *Colloque GRETSI*, 1983, p. 525-533.

[7] K. J. Astrom, *Introduction to Stochastic Control Theory.* New York: Academic Press, 1970.

[8] D. Kazakos, “Optimal constrained representation and filtering of signals,” *Signal Processing*, vol. 5, pp. 347-353, 1983.

[9] A. G. Journel, “Kriging in terms of projections,” *Math. Geol.*, vol. 9, pp. 563-586,1977.

[10] W. D. Stirling, “Least squares subject to linear constraints,” *J. Roy. Sot. Appl. Stat.*, vol. 30, pp. 204-212, 1981

[11] B. Picinbono, “A geometrical interpretation of signal detection and estimation,” *IEEE Trans. Inform. Theory*, IT-26, pp. 493-497, 1980.

[12] H. El Ayadi and B. Picinbono, “NAR AGC adaptive detection of nonoverlapping signals in noise with fluctuating power,” *IEEE Trans. Acoust., Speech, Sig. Proc.*, ASSP-29, pp. 952-963, 1981.