ROBUST FILTERING BASED ON PROBABILISTIC DESCRIPTIONS OF MODEL ERRORS

Mikael Sternad and Anders Ahlén
Automatic Control Group, Uppsala University
P.O. Box 27, S–751 03 Uppsala, Sweden

Abstract. A new approach to robust estimation of signals and prediction of time–series is considered. Signal and system parameter deviations are represented as random variables, with known covariances. A robust design is obtained by minimizing the squared estimation error, averaged both with respect to model errors and noise. A polynomial solution, based on averaged spectral factorizations and averaged Diophantine equations, is derived. The robust estimator is called a cautious Wiener filter. It turns out to be no more complicated to design than an ordinary Wiener filter.

Keywords: Robustness; Optimal estimation; Wiener filtering; Deconvolution; Prediction.

1. INTRODUCTION

Kalman and Wiener filters are in frequent use. With many computer programs for design available, filter parameters are easy to obtain. Perhaps too easy, since we are not forced to specify the validity of our models. It is also hard to take a whole range of expected system behaviour into account. The goal of *robust filter* synthesis is to overcome these drawbacks.

We here propose a novel approach to robust design for signal estimation. It is based on a stochastic description of model errors, related to the stochastic embedding concept of Goodwin and Salgado (1989). A single robust filter, for the whole class of possible models, is obtained by minimizing the squared estimation error, averaged both with respect to model errors and the noise.

Most previous suggestions for robust filter design have been based on the minimax approach. See e.g. D'Appolito and Hutchinson (1972) and Kassam and Poor (1985). Apart from leading to a much simpler design methodology, the approach proposed here avoids two drawbacks of robust minimax design. First, the descriptions of model uncertainties may have *soft bounds.* These are more readily obtainable in a noisy environment than the hard bounds required for minimax design. Secondly, not only the range of uncertainties, but also their likelihood is taken into account; Probable model errors will have a greater impact on an estimator design than do very rare "worst cases". The conservativeness is thus reduced.

A polynomial solution, based on averaged spectral factorizations and averaged Diophantine equations, will be presented. The design procedure constitutes a generalization of the polynomial equations approach, which was pioneered by Kučera (1979). Mild solvability conditions guarantee the existence of stable optimal filters. The robust design turns out to be no more complicated than the design of an ordinary Wiener filter. The methodology is here exemplified on a scalar discrete–time deconvolution problem. Derivation of design equations for this and other problems, such as state estimation and feedforward control, can be found in Sternad and Ahlén (1993).

Remarks on the notation. For any complex polynomial of degree np, in the backward shift operator

$$P(q^{-1}) = p_o + p_1 q^{-1} + \ldots + p_{np} q^{-np}$$

the *conjugate polynomial* is defined as $P_*(q) \triangleq p_o^* + p_1^* q + \ldots + p_{np}^* q^{np}$. For convenience, polynomial arguments will often be omitted. We call $P(q^{-1})$ *stable* if all zeros of $P(z^{-1})$ are in $|z| < 1$.

2. THE ESTIMATION PROBLEM

A scalar deconvolution problem will be considered, to illustrate the design principles. It includes eg. ordinary output filtering and prediction of ARMA–processes as special cases. It also includes the design of linear recursive equalizers for digital communications. Measurements are described as

$$y(t) = \mathcal{G}(q^{-1})u(t-k) + w(t) \quad . \tag{1}$$

The linear, causal and possibly uncertain transfer function $\mathcal{G}(q^{-1})$ may eg. represent a transducer. If the delay is uncertain, k denotes its minimum value. The input $u(t)$ and the measurement noise $w(t)$ are described by possibly uncertain ARMA–models

$$u(t) = \mathcal{F}(q^{-1})e(t) \quad ; \quad w(t) = \mathcal{H}(q^{-1})v(t) \tag{2}$$

$$E|e(t)|^2 = 1 \quad ; \quad E|v(t)|^2 \triangleq \sigma_v^2 \quad .$$

The time–series $e(t)$ and $v(t)$ are assumed mutually uncorrelated. They are stationary white noises or impulse sequences, with zero mean. The variance σ_v^2 may be uncertain and has nominal value ρ_0. All transfer functions are assumed time–invariant. Signals may be complex–valued; this is the case eg. in digital communications applications.

A stable, linear and time–invariant estimator of $u(t)$, given $y(t+m)$, is sought. Depending on m, it could be a predictor $(m < 0)$, a filter $(m = 0)$ or a fixed lag smoother $(m > 0)$.

3. ADDITIVE PROBABILISTIC ERROR MODELS

The transfer functions $\mathcal{G}$, $\mathcal{F}$ and $\mathcal{H}$ may be uncertain. An *error model* is a quantification of the model error class. Together with a nominal model, it constitutes the *extended design model* on which a robust design is based. As error models, we will utilize additive transfer functions $\Delta\mathcal{F}, \Delta\mathcal{G}, \Delta\mathcal{H}$, with unknown stochastic numerators and pre–specified denominators. This choice is crucial for obtaining a simple solution to the filtering problem. The extended design models are specified as

$$\begin{aligned}
\mathcal{F} &= \frac{C_0}{D_0} + \frac{C_1 \Delta C}{D_1} = \frac{C_0 D_1 + D_0 C_1 \Delta C}{D_0 D_1} \triangleq \frac{C}{D} \\
\mathcal{G} &= \frac{B_0}{A_0} + \frac{B_1 \Delta B}{A_1} = \frac{B_0 A_1 + A_0 B_1 \Delta B}{A_0 A_1} \triangleq \frac{B}{A} \\
\mathcal{H} &= \frac{M_0}{N_0} + \frac{M_1 \Delta M}{N_1} = \frac{M_0 N_1 + N_0 M_1 \Delta M}{N_0 N_1} \triangleq \frac{M}{N}
\end{aligned} \quad (3)$$

Above, C_0/D_0 etc. represent nominal models, with degrees nc_0, nd_0 etc. They are assumed known and stable. Stable "error denominators" D_1, A_1 and N_1, of degrees nd_1, na_1 and nn_1, as well as numerator factors C_1, B_1, M_1, of degrees nc_1, nb_1 and nm_1, may be specified by the designer or obtained from data. Coefficients of numerator polynomials

$$\Delta P(q^{-1}) = \Delta p_0 + \Delta p_1 q^{-1} + \ldots + \Delta p_{\delta p} q^{-\delta p} \quad (4)$$

are stochastic variables. They have zero means and parameter covariances $\bar{E}\Delta p_i \Delta p_j^*$, collected in the covariance matrices $\mathbf{P}_{\Delta P}$. These coefficients are constant in time, so they are independent of the time series $e(t)$ and $v(t)$. Except for first and second order moments, their distributions need not be known, since they will not affect the design.

All polynomial degrees are assumed known to (or specified by) the designer [1]. Note also that denominators polynomials are *all assumed stable.* Model uncertainty more or less forces us to restrict attention to stable extended design models [2].

In the following, we utilize three further assumptions:

A1. Coefficients of ΔC and of ΔB are independent.

A2. The covariance matrices $\mathbf{P}_{\Delta C}, \mathbf{P}_{\Delta B}$ and $\mathbf{P}_{\Delta M}$ are Hermitian and positive semidefinite.

A3. The standard deviation of $v(t), \sigma_v$, is regarded as a stochastic variable, independent of ΔM, with mean $\sqrt{\rho_0}$ and variance $E\sigma_v^2 - (E(\sigma_v))^2 = \rho_v$. Thus, $E\sigma_v^2 = \rho_0 + \rho_v \triangleq \rho$.

The uncertainty in ρ_0 will be taken into account by using a higher equivalent variance ρ. It is necessary to assure A2 when the covariance matrices are used pragmatically, as "robustness tuning knobs". Design equations could be derived for situations with correlations between ΔC and ΔB. Assumption A1 does, however, simplify the solution, and seems reasonable.

Model error covariances may be obtained from identification experiments, or from frequency domain data on system variability. See Sternad and Ahlén (1993), Goodwin and Salgado (1989) and Goodwin, Gevers and Ninnes (1991). If a fixed filter is to be designed for a large number of systems, the statistics may be obtained from a representative sample of systems.

Probabilistic error models remain useful also when statistics is hard to obtain. Those who prefer a Bayesian view could then interprete error distributions as subjective probabilities. Others may just use them pragmatically, as robustness "tuning knobs". The covariances are then altered until satisfactory spectral properties of the filter are obtained.

4. DESIGN OF ROBUST FILTERS

We proceed from the model (0.1), (0.2) and (0.3). The coefficients of ΔC, ΔB and ΔM are random variables, whose possible values parametrize a set of systems. We will minimize the *averaged MSE criterion*

$$\bar{E}(E|z(t)|^2) = \bar{E}E|f(t) - \hat{f}(t|t+m)|^2 \quad (5)$$

where E represents expectation over noise and $\bar{E}$ is an expectation over the model error distribution. We thus seek a single estimator which provides the best MSE

[1] Note that we are talking about an extended design *model.* In practice, it will only be an approximation of a class of possibly infinite dimensional and time–varying true systems.

[2] If unstable poles were exactly known, a finite estimation error could be obtained, by a filter which cancels unstable poles by zeros in the total signal path to the estimation error. Such a strategy is, of course, highly non–robust to mis–modelling of unstable poles. With *uncertain* unstable poles, the design problem becomes unsolvable, in the open–loop context considered here. Therefore, a general solution involving two coupled Diophantine equations will not be of interest here.

performance, on average, when applied on randomly selected systems within the specified class.

This type of criterion has been used in connection to other filtering problems, e.g. by Chung and Bélanger (1976), Speyer and Gustafson (1975) and by Grimble (1984). These works were based on the assumption of small uncertainties and on series expansion of uncertain poles. In our design philosophy, we start from a model structure (0.3), and adjust it to the uncertainty directly. Large uncertainties can then be described in a much better way. See Sternad and Ahlén (1993).

4.1 The averaged spectral factorization

An averaged spectral factor $\beta(q^{-1})$ is defined as the stable and monic solution to

$$r\beta\beta_* \triangleq \bar{E}\{CC_*BB_*NN_* + \sigma_v^2 MM_*AA_*DD_*\} \qquad (6)$$

with scalar r. Define double–sided polynomials

$$\tilde{C}\tilde{C}_* \triangleq \bar{E}(CC_*),\ \tilde{B}\tilde{B}_* \triangleq \bar{E}(BB_*),\quad \tilde{M}\tilde{M}_* \triangleq \bar{E}(MM_*) \qquad (7)$$

Then, use of (0.3) gives

$$\tilde{C}\tilde{C}_* = C_0C_{0*}D_1D_{1*} + D_0D_{0*}C_1C_{1*}\bar{E}(\Delta C\Delta C_*)$$
$$\tilde{B}\tilde{B}_* = B_0B_{0*}A_1A_{1*} + A_0A_{0*}B_1B_{1*}\bar{E}(\Delta B\Delta B_*)$$
$$\tilde{M}\tilde{M}_* = M_0M_{0*}N_1N_{1*} + N_0N_{0*}M_1M_{1*}\bar{E}(\Delta M\Delta M_*) \qquad (8)$$

We can now simplify (0.6).

Lemma 1. Let assumptions A1 and A3 hold. Then, (0.6) can be expressed as

$$r\beta\beta_* = \tilde{C}\tilde{C}_*\tilde{B}\tilde{B}_*NN_* + \rho\tilde{M}\tilde{M}_*AA_*DD_* \qquad (9)$$

■

Proof: The coefficients of a polynomial ΔP are zero mean stochastic variables. Coefficients of $\Delta P\Delta P_*$ will also be stochastic variables, having expected values given by (0.11) below. The noise standard deviation is independent of ΔM and the coefficients of (ΔB, ΔC), are independent, and so are the coefficients of $\Delta B\Delta B_*$, $\Delta C\Delta C_*$. Using independence for complex parameters, the right–hand side of (0.6) becomes

$$\bar{E}(CC_*)\bar{E}(BB_*)NN_* + \bar{E}(\sigma_v)^2\bar{E}(MM_*)AA_*DD_*$$

which, utilizing (0.7) and A3, is (0.9) ■

The averaged factors in (0.8) can be evaluated as follows. For a stochastic error model numerator $\Delta P(q^{-1})$, as in (0.4), let the Hermitian parameter covariance matrix be

$$\mathbf{P}_{\Delta P} = \begin{bmatrix} \bar{E}|\Delta p_0|^2 & \dots & \bar{E}(\Delta p_0\Delta p_{\delta p}^*) \\ \vdots & \ddots & \vdots \\ \bar{E}(\Delta p_{\delta p}\Delta p_0^*) & \dots & \bar{E}|\Delta p_{\delta p}|^2 \end{bmatrix} \qquad (10)$$

Denote the sum of the diagonal elements h_0, the sum of elements in the i'th super–diagonal h_i, and the sum of elements in the i'th subdiagonal h_{-i}. Note that $h_{-i} = h_i^*$. Then it becomes evident, by direct multiplication of $\Delta P(q^{-1})\Delta P_*(q)$, and taking expectations, that

$$\bar{E}(\Delta P\Delta P_*) =$$
$$h_{dp}^*q^{-dp} + \dots + h_1^*q^{-1} + h_0 + h_1q + \dots + h_{dp}q^{dp} \quad . \ (11)$$

Thus, the averaged factors in (0.8) are readily obtained. (Above, $dp \leq \delta p$, with $dp = 0$ if coefficients are uncorrelated.)

In (0.8), $\tilde{C}\tilde{C}_*$ will contain powers up to $q^{\pm n\tilde{c}}$, where $n\tilde{c} = \max\{nc_0 + nd_1, nd_0 + nc_1 + dc\}$, with analogous expressions for $n\tilde{b}, n\tilde{m}$. Since $N = N_0N_1$ etc, the averaged spectral factor in (0.9) has degree

$$n\beta = \max\{n\tilde{c}+n\tilde{b}+nn_0+nn_1, n\tilde{m}+na_0+na_1+nd_0+nd_1\}$$

The factorization (0.9) is solvable with respect to a unique stable $\beta(z^{-1})$ iff its right–hand side is positive on $|z| = 1$. Introduce the assumptions

A4. C_0, $C_1\bar{E}(\Delta C\Delta C_*)$, ρM_0 and $\rho M_1\bar{E}(\Delta M\Delta M_*)$ have no common zeros on $|z| = 1$

A5. B_0, $B_1\bar{E}(\Delta B\Delta B_*)$, ρM_0 and $\rho M_1\bar{E}(\Delta M\Delta M_*)$ have no common zeros on $|z| = 1$.

Lemma 2. Let D, A and N be stable and A2 hold. Then, a unique stable spectral factor β, satisfying (0.9), exists, if and only if both of A4 and A5 are true ■

Proof: See Sternad and Ahlén (1993).

The conditions A4 and A5 are mild. They will almost always be fulfilled, even if C_0, B_0 and M_0 have zeros on the unit circle. In fact, the conditions are more relaxed than for the nominal case, due to the presence of averaged factors $\bar{E}(\cdot)$.

4.2 The cautious Wiener filter

Theorem 1. Assume an extended design model (0.1)–(0.3) to be given, with known covariances of the stochastic polynomial coefficients. Assume A1–A5 to hold. An estimator of $u(t)$ then minimizes (0.5), among all linear time–invariant estimators based on $y(t+m)$, if and only if it has the same coprime factors as

$$\hat{u}(t|t+m) = \frac{Q}{R}y(t+m) \ \ ; \ \ \frac{Q}{R} = \frac{Q_1N_0N_1A_0A_1}{\beta} \qquad (12)$$

Here, $\beta(q^{-1})$ is obtained from (0.9), while $Q_1(q^{-1})$, together with $L_*(q)$, is the unique solution to

$$q^{-m+k}\tilde{C}\tilde{C}_*B_{0*}A_{1*}N_{0*}N_{1*} = r\beta_*Q_1 + qD_0D_1L_* \qquad (13)$$

with polynomial degrees

$$\begin{aligned} nQ_1 &\leq \max(n\tilde{c} - k + m \ , \ nd_0 + nd_1 - 1) \\ nL &\leq \max(n\tilde{c} + nb_0 + na_1 + nn_0 + nn_1 \\ &\quad +k - m \ , \ n\beta) - 1 \ . \end{aligned} \tag{14}$$

For an ensemble of systems, the minimal criterion value becomes

$$\bar{E}E|z(t)|^2_{\min} = \frac{1}{2\pi i}\oint \frac{LL_*}{r\beta\beta_*} + \rho\frac{\tilde{C}\tilde{C}_*\tilde{M}\tilde{M}_*AA_*}{r\beta\beta_*} + $$
$$+ \frac{\tilde{C}\tilde{C}_*\tilde{C}\tilde{C}_*\bar{E}(\Delta\mathcal{G}\Delta\mathcal{G}_*)AA_*NN_*}{DD_*r\beta\beta_*}\frac{dz}{z} \tag{15}$$

■

Proof: See Sternad and Ahlén (1993).

Remarks. The equations for minimizing (0.5) are (0.9), (0.11) and (0.13). The only new type of computation, as compared to a nominal solution, is trivial: summation of covariance matrix elements, diagonalwise.

Note that N_1 and A_1 affect the filter (0.12) directly. If $1/N_1$ or $1/A_1$ in the error models have resonance peaks, indicating large uncertainty, the filter (0.12) will have low gain at those frequencies. With increasing model uncertainty, the zeros of β are moved inward in the unit circle. Resonance peaks of the estimator are lowered and broadened.

Equation (0.13) will have a unique solution, with degrees (0.14). Note that $\beta_*(z)$ (unstable) and $D_0(z^{-1})D_1(z^{-1})$ (stable) have no common factors.

The integrand of (0.15) consists of three terms. Term 1 represents the effect of finite smoothing lag m. It can be shown that $L_*(q) \to 0$ when $m \to \infty$. The second term mainly represents the effect of noise. It vanishes for $\rho = 0$. Finally, the third term represents degradation caused by errors $\Delta\mathcal{G} = B_1\Delta B/A_1$ in the transducer model. It vanishes only when $\Delta\mathcal{G} = 0$.

In situations with little noise and sufficiently large smoothing lag m, term 3 in (0.15) will dominate the error. This is not surprising; a deconvolution smoother then essentially inverts $\mathcal{G}$. This operation is sensitive to model errors there.

4.3 Analytical expressions for performance evaluation

Theorem 2. Let a nominal estimator Q_0/R_0 be designed based on a nominal model, as in eg. Ahlén and Sternad (1989),(1991). Applying it, instead of (0.12), on an ensemble of systems results in an increase, compared to (0.15), of the mean MSE $\bar{E}E|z(t)|^2$. The increase is given by

$$\bar{E}E|z(t)|^2_0 - \bar{E}E|z(t)|^2_{\min} =$$

$$= \frac{r}{2\pi i}\oint \left|\frac{\beta}{DAN}\right|^2 \left|\frac{Q_0}{R_0} - \frac{Q}{R}\right|^2 \frac{dz}{z} \tag{16}$$

where r, β is defined by (0.6)-(0.9) and Q/R is the optimal robust filter, given by (0.12). ■

Proof: See Sternad and Ahlén (1993).

Theorem 3. Let a robust estimator Q/R be designed by (0.9)–(0.13). When applying it on a system equal to the nominal model, the increased MSE, compared to the minimum, is

$$E|z(t)|^2 - E|z(t)|^2_0 \ =$$

$$\frac{r_0}{2\pi i}\oint \left|\frac{\beta_0}{D_0A_0N_0}\right|^2 \left|\frac{Q}{R} - \frac{Q_0}{R_0}\right|^2 \frac{dz}{z} \tag{17}$$

where r_0 and β_0 are obtained from the spectral factorization in a nominal design. ■

Proof: See Sternad and Ahlén (1993).

Remarks. The expression in (0.16) can be used for arbitrary linear estimators Q_0/R_0, for example minimax–designs. We thus do not have to evaluate the mean performance of alternative designs by Monte–Carlo simulation. The mean innovations model β/DAN in (0.16), and the nominal innovations model $\beta_0/D_0A_0N_0$ in (0.17) can be seen as weighting functions. In frequency regions where their magnitude is large, differences between the two estimators will have a large impact on the performance.

A simulation study of the filtering performance is presented in Sternad and Ahlén (1993).

REFERENCES

Ahlén, A and M Sternad (1989). Optimal deconvolution based on polynomial methods. *IEEE Trans. Acoustics, Speech and Signal Processing.*, vol 37, pp 217–226.

Ahlén, A and M Sternad, (1991). Wiener filter design using polynomial equations. *IEEE Trans. Signal Processing*, vol 39, pp 2387–2399.

Chung, R C and P R Bélanger (1976). Minimum–sensitivity filter for linear time–invariant stochastic systems with uncertain parameters. *IEEE Trans. Automatic Control*, vol 21, pp 98–100.

D'Appolito, J A and C E Hutchinson (1972). A minimax approach to the design of low sensitivity state estimators. *Automatica*, vol 8, pp 599–608.

Goodwin, G C and M E Salgado (1989). A stochastic embedding approach for quantifying uncertainty in the estimation of restricted complexity models. *International Journal of Adaptive Control and Signal Processing* vol 3, pp 333–356.

Goodwin, GC, M Gevers and B Ninness (1991). Quantifying the error in estimated transfer functions with application to model order selection. *Submitted.*

Grimble, M J (1984). Wiener and Kalman filters for systems with random parameters. *IEEE Trans. Automatic Control*, vol AC–29, pp 552–554.

Kassam, S A and H V Poor (1985). Robust techniques for signal processing: a survey. *Proc. IEEE*, vol 73, pp 433–481.

Kučera, V (1979). *Discrete Linear Control. The Polynomial Equations Approach.* Wiley, New York.

Speyer, J L and D E Gustafson (1975). An approximation method for estimation in linear systems with parameter uncertainty. *IEEE Trans. Automatic Control*, vol AC–20, pp 354–359.

Sternad, M and A Ahlén (1993). Robust filtering and feedforward control based on probabilistic descriptions of model errors. *Automatica*, to appear. (Also available as Report UPTEC 91071R, Dept. of Technology, Uppsala University, Sweden.)