DISCRETE H_∞ FILTER DESIGN WITH APPLICATION TO SPEECH ENHANCEMENT

X. Shen

Department of Electrical and Computer Engineering
University of Waterloo
Waterloo, Ontario, Canada N2L 3G1

ABSTRACT

In this paper, discrete H_∞ filter design with a linear quadratic (LQ) game approach is presented. The exogenous inputs composed of the "hostile" noise signals and system initial condition are assumed to be finite energy signals with unknown statistics. The design criterion is to minimize the worst possible amplification of the estimation error signal, which is different from the classical minimum variance estimation error criterion for the modified Wiener or Kalman filter design. The approach can show that how far the estimation error can be reduced under an existence condition on the solution to a corresponding Riccati equation. The application of the discrete H_∞ filter to enhance speech contaminated by additive noise is then investigated. In the H_∞ estimation, the noise sources can be arbitrary signals with only requirement of bounded energy, this estimation is more appropriate in the practical speech enhancement.

1. INTRODUCTION

Noise contaminated speech results in reduction of speech discrimination. There have been numerous studies involving speech enhancement of noised speech [1]-[5]. All the studies have been based on the minimization of the variance of the estimation error of the speech, i.e. the celebrated Wiener and/or Kalman filtering approach. This type of estimation assumes that the speech generating process has a known dynamic, and that the noise sources have known statistical properties. However, these assumptions may limit the application of the estimators since in many situations only an approximate signal model is available and/or the statistics of the noise sources are not fully known or unavailable. Also, both Wiener and Kalman estimators may not be robust against uncertainty of the speech signal model. Recently, a new class of optimal filter has been developed using H_∞ minimum estimation error spectrum error criterion [6]-[9]. The optimal estimator is designed to guarantee that the operator relating the noise signals to the resulting estimation error should possess an H_∞ norm less than a prescribed positive value. In the H_∞ estimation, the noise sources can be arbitrary signals with only requirement of bounded noise. Since H_∞ estimation problem involves the minimization of the worst possible amplification of the error signal, it can be viewed as a dynamic, two-persons, zero sum game. In this paper, we define a difference game in which the state estimator and the disturbance signals (system noise, initial condition and measurement noise) have the conflicting objectives of respectively minimizing and maximizing the estimation error. The minimizer picks the optimal filtered estimate and the maximizer picks the worst-case disturbance and initial condition. We give a detail derivation to solve the game which directly produces the solution for the problems of the discrete H_∞ filtering. Similar design approach has been proposed in [6] for the continuous case. We then investigate a possibility of the application of the H_∞ filter to the speech enhancement. The speech source model and the observation model are first written into canonical state space models, the developed H_∞ filtering algorithm is applied to estimate(enhance) speech signal. The performance of the H_∞ filter is better compared with those of Wiener and Kalman filters in terms of estimation sensitivity to parameter uncertainty and the reduction of the peak of the error spectrum.

2. DISCRETE H_∞ FILTER DESIGN

Consider the following discrete-time system

$$\begin{aligned} x_{k+1} &= A_k x_k + B_k w_k \\ y_k &= C_k x_k + v_k \\ k &= 0, 1, \ldots, N-1; \quad x_0 = x_0 \end{aligned} \tag{1}$$

where $x_k \in \mathcal{R}^n$ is the state vector, $w_k \in \mathcal{R}^m$ is the system noise vector, $y_k \in \mathcal{R}^p$ is the measurement vector and $v_k \in \mathcal{R}^p$ is the measurement noise vector. A_k, B_k and C_k are matrices of the appropriate dimensions. Assume that (A_k, B_k) is controllable and (C_k, A_k) is detectable. Define the measurement history as Y_k=(y_k, $0 \le k \le N-1$). The estimate of the state $\hat{x}_k$ at time k, is computed based on the measurement history up to $N-1$. We are interested not necessarily in the estimation of x_k but in the estimation of a linear combination of x_k

$$z_k = L_k x_k \tag{2}$$

The H_∞ filter is required to provide an uniformly small estimation error, e_k=z_k-$\hat{z}_k$, for any w_k, $v_k \in l_2$ and $x_0 \in \mathcal{R}^n$. The measure of performance is then given by

0-7803-2431-5/95 $4.00

$$J = \frac{\sum_{k=0}^{N-1} ||z_k - \hat{z}_k||^2_{Q_k}}{||x_0 - \hat{x}_0||^2_{p_0^{-1}} + \sum_{k=0}^{N-1}\{||w_k||^2_{W_k^{-1}} + ||v_k||^2_{V_k^{-1}}\}} \quad (3)$$

where $((x_0 - \hat{x}_0), w_k, v_k) \neq 0$, $Q_k \geq 0$, $p_0^{-1} > 0$, $W_k > 0$ and $V_k > 0$ are the weighting matrices. The optimal estimate z_k among all possible $\hat{z}_k$(i.e. the worse-case performance measure) should satisfy

$$sup J < 1/\gamma \quad (4)$$

where "sup" stands for supremum and $\gamma > 0$ is a prescribed level of noise attenuation. The matrices $Q_k \geq 0$, $W_k > 0$, $V_k > 0$ and $p_0 > 0$ are left to the choice of the designer and depend on performance requirements. The discrete H_∞ filtering can be interpreted as a minimax problem where the estimator strategy $\hat{z}_k$ play against the exogenous inputs w_k and the initial state x_0. The performance criterion becomes

$$\begin{aligned} &\min_{\hat{z}_k} \max_{(v_k, w_k, x_0)} J \\ &= -\frac{1}{2\gamma}||x_0 - \hat{x}_0||^2_{p_0^{-1}} + \frac{1}{2}\sum_{k=0}^{N-1}[||z_k - \hat{z}_k||^2_{Q_k} \\ &\quad - \frac{1}{\gamma}(||w_k||^2_{W_k^{-1}} + ||v_k||^2_{V_k^{-1}})] \end{aligned} \quad (5)$$

where "min" stands for minimization and "max" maximization. Note that unlike the traditional minimum variance filtering approach(Wiener and/or Kalman filtering), the H_∞ filtering deals with deterministic disturbances and no *a priori* knowledge of the noise statistics is required. Since noisy speech (the observations) y_k is given, and $z_k = L_k x_k$, $\hat{z}_k = L_k \hat{x}_k$, we can rewrite the performance criterion (5) as

$$\begin{aligned} &\min_{\hat{x}_k} \max_{(y_k, w_k, x_0)} J \\ &= -\frac{1}{2\gamma}||x_0 - \hat{x}_0||^2_{p_0^{-1}} + \frac{1}{2}\sum_{k=0}^{N-1}[||x_k - \hat{x}_k||^2_{\bar{Q}_k} \\ &\quad - \frac{1}{\gamma}(||w_k||^2_{W_k^{-1}} + ||y_k - C_k x_k||^2_{V_k^{-1}})] \end{aligned} \quad (6)$$

where $\bar{Q}_k = L_k^T Q_k L_k$. The following theorem presents a complete solution to the H_∞ filtering problem for the system (1) with performance criterion (6).

Theorem : Let $\gamma > 0$ be a prescribed level of noise attenuation. Then, there exists an H_∞ filter for z_k if and only if there exists a stabilizing symmetric solution $P_k > 0$ to the following discrete-time Riccati equation

$$\begin{aligned} P_{k+1} &= A_k P_k A_k^T + B_k W_k B_k^T \\ &\quad + A_k P_k(\gamma\bar{Q}_k - C_k^T V_k^{-1} C_k) \\ &\quad \cdot [I - P_k(\gamma\bar{Q}_k - C_k^T V_k^{-1} C_k)]^{-1} P_k A_k^T \quad (7) \\ P_0 &= p_0 \quad \bar{Q}_k = L_k^T Q_k L_k \end{aligned}$$

Proof: By using a set of Lagrange multipliers $\lambda_1, \lambda_2, \ldots, \lambda_N$, we define a new performance index M as follows

$$\begin{aligned} M &= -\frac{1}{2\gamma}||x_0 - \hat{x}_0||_{p_0^{-1}} + \frac{1}{2}\sum_{k=0}^{N-1}\{[||z_k - \hat{z}_k||^2_{Q_k} \\ &\quad - \frac{1}{\gamma}(||w_k||^2_{W_k^{-1}} + ||y_k - C_k x_k||^2_{V_k^{-1}} \\ &\quad + \frac{\lambda_{k+1}^T}{\gamma}[A_k x_k + B_k w_k - x_{k+1}] \\ &\quad + [A_k x_k + B_k w_k - x_{k+1}]^T \frac{\lambda_{k+1}}{\gamma}\} \end{aligned} \quad (8)$$

Taking the first variation, the necessary conditions for a maximum are

$$\begin{aligned} x_0 &= \hat{x}_0 + p_0\lambda_0, \quad \lambda_N = 0 && (9) \\ w_k &= W_k B_k^T \lambda_{k+1} && (10) \\ \lambda_k &= A^T\lambda_{k+1} + \gamma\bar{Q}_k(x_k - \hat{x}_k) \\ &\quad + C_k^T V_k^{-1}(y_k - C_k x_k) && (11) \end{aligned}$$

where $\bar{Q}_k = L_k^T Q_k L_k$.

These first order necessary conditions result in a two point boundary value problem

$$\begin{aligned} \begin{pmatrix} x_{k+1} \\ \lambda_k \end{pmatrix} &= \begin{pmatrix} A_k & B_k W_k B_k^T \\ \gamma\bar{Q}_k - C_k^T V_k^{-1} C_k & A_k^T \end{pmatrix} \\ &\quad \cdot \begin{pmatrix} x_k \\ \lambda_{k+1} \end{pmatrix} + \begin{pmatrix} 0 \\ -\gamma\bar{Q}_k\hat{x}_k + C_k^T V_k^{-1} y_k \end{pmatrix}, \\ &\quad k = 0, 1, \ldots, N-1 \end{aligned} \quad (12)$$

with boundary conditions

$$x_0 = \hat{x}_0 + p_0\lambda_0, \quad \lambda_N = 0 \quad (13)$$

Since the two-point boundary value problem is linear, the solution is assumed to be of the form

$$x_k^* = \bar{x}_k + P_k\lambda_k^* \quad (14)$$

where $\bar{x}$ and P_k are undetermined variables. x_k^* and λ_k^* represent optimal value of x_k and λ_k, respectively, for any fixed admissible functions of $\bar{x}_k$ and y_k. The optimal values for w_k and x_0 are

$$w_k^* = W_k B_k^T \lambda_k^*, \quad x_0^* = \hat{x}_0 + p_0\lambda_0^* \quad (15)$$

Substituting (14) into (12) results in

$$\begin{aligned} &\bar{x}_{k+1} + P_{k+1}\lambda_{k+1}^* \\ &= A_k\bar{x}_k + A_kP_k\lambda_k^* + B_kW_kB_k^T\lambda_{k+1}^* \end{aligned} \quad (16)$$

and

$$\begin{aligned} \lambda_k^* &= (I - \gamma\bar{Q}_kP_k + C_k^TV_k^{-1}C_kP_k)^{-1} \\ &\cdot [\gamma\bar{Q}_k(\bar{x}_k - \hat{x}_k) + C_k^TV_k^{-1}(y_k - C_k\bar{x}_k) \\ &+ A_k^T\lambda_{k+1}^*) \end{aligned} \quad (17)$$

From (16)-(17) we have

$$\begin{aligned} \bar{x}_{k+1} &+ P_{k+1}\lambda_{k+1}^* \\ &= A_k\bar{x}_k + A_kP_k(I - \gamma\bar{Q}_kP_k + C_k^TV_k^{-1}C_kP_k)^{-1} \\ &\cdot [\gamma\bar{Q}_k(\bar{x}_k - \hat{x}_k) \\ &+ C_k^TV_k^{-1}(y_k - C_k\bar{x}_k) + A_k^T\lambda_{k+1}^*] \\ &+ B_kW_kB_k^T\lambda_{k+1}^* \end{aligned} \quad (18)$$

i.e.

$$\begin{aligned} \bar{x}_{k+1} &- A_k\bar{x}_k - A_kP_k(I - \gamma\bar{Q}_kP_k + C_k^TV_k^{-1}C_kP_k)^{-1} \\ &\cdot [\gamma\bar{Q}_k(\bar{x}_k - \hat{x}_k) + C_k^TV_k^{-1}(y_k - C_k\bar{x}_k)] \\ &= [-P_{k+1} + A_kP_k(I - \gamma\bar{Q}_kP_k + C_k^TV_k^{-1}C_kP_k)^{-1}A_k^T \\ &+ B_kW_kB_k^T]\lambda_{k+1}^* \end{aligned} \quad (19)$$

For equation (19) to hold true for arbitrary λ_k^*, both sides are set identically to zero, resulting in

$$\begin{aligned} \bar{x}_{k+1} &= A_k\bar{x}_k + A_kP_k[(I - (\gamma\bar{Q}_k - C_k^TV_k^{-1}C_k)P_k]^{-1} \\ &[\gamma\bar{Q}_k(\bar{x}_k - \hat{x}_k) + C_k^TV_k^{-1}(y_k - C_k\bar{x}_k)] \quad (20) \\ \bar{x}_0 &= \hat{x}_0 \end{aligned}$$

and

$$\begin{aligned} P_{k+1} &= A_kP_kA_k^T + B_kW_kB_k^T + A_kP_k(\gamma\bar{Q}_k - C_K^TV_k^{-1}C_k) \\ &\cdot [I - P_k(\gamma\bar{Q}_k - C_k^TV_k^{-1}C_k)]^{-1}P_kA_k^T, \\ &P_0 = p_0 \end{aligned} \quad (21)$$

Equation (21) is the well-known Riccati difference equation.

Claim 1: If the solution P_k to the Riccati equation (21) exists $\forall k \in [0, N-1]$, then $P_k > 0\ \forall k \in [0, N-1]$.

The optimal strategies (15) are substituted into the performance (6), results in the min-max problem

$$\begin{aligned} &\min_{\hat{x}_k} \max_{y_k} J \\ &= \frac{1}{2}\sum_{k=0}^{N-1}[||\bar{x}_k - \hat{x}_k||_{\bar{Q}_k}^2 - \frac{1}{\gamma}||y_k - C_k\bar{x}_k||_{V_k^{-1}}^2] \end{aligned} \quad (22)$$

subject to the dynamic constraints (20) and (21). Let

$$r_k = \bar{x}_k - \hat{x}_k, \quad q_k = y_k - C_k\bar{x}_k \quad (23)$$

equation (22) becomes

$$\begin{aligned} &\min_{r} \max_{q} J \\ &= \frac{1}{2}\sum_{k=0}^{N-1}[||r_k||_{\bar{Q}_k}^2 - \frac{1}{\gamma}||q_k||_{V_k^{-1}}^2] \end{aligned} \quad (24)$$

The two independent players r_k and q_k in (24) affects the variables $\bar{x}_k$, but $\bar{x}_k$ does not appear in the performance index, therefore the optimal strategies of r_k and q_k are

$$r_k^* = 0, \quad q_k^* = 0 \quad (25)$$

i.e.

$$\bar{x}_k = \hat{x}_k^*, \quad y_k^* = C_k\bar{x}_k \quad (26)$$

The value of the game is the value of the cost function (6), when all the players use their optimal strategies. When the optimal strategies $\hat{x}_k^*$, y_k^*, w_k^* and x_0^* in (15) and (26) are substituted into the (6)

$$J(\hat{x}_k^*, y_k^*, w_k^*, x_0^*) = 0 \quad (27)$$

giving a zero value game.

Moreover, with the Theorem and equations (20) and (26), the optimal H_∞ filter is given by

$$\hat{z}_k^* = L_k\hat{x}_k^*, \quad k = 0, 1, \ldots, N-1 \quad (28)$$

where

$$\begin{aligned} \hat{x}_{k+1}^* &= A_k\hat{x}_k^* + K_k(y_k - C_k\hat{x}_k^*), \quad \hat{x}_0 = x_0 \quad (29) \\ K_k &= A_kP_k(I - \gamma\bar{Q}_kP_k + C_k^TV_k^{-1}C_kP_k)^{-1} \\ &C_k^TV_k^{-1} \quad (30) \end{aligned}$$

It is important to note that the optimal H_∞ estimator depends on the weighting on the estimation error in the performance criterion, i.e. the designer choses the weighting matrices based on the performance requitements, while both Wiener and Kalman estimators are dependents on the variance of the noise.

3. APPLICATION TO THE SPEECH ENHANCEMENT

Assume that a windowed noise signal v_k has been added to a windowed speech signal x_k, with their sum denoted by s_k

$$s_k = x_k + v_k \quad (31)$$

Let the speech signal x_k be generated according to the nth order autoregressive (AR) model

$$x_{k+1} = \sum_{j=1}^{n} a_j x_{k-j} + w_k \tag{32}$$

where w_k is a zero mean, white Gaussian process. The speech generating mechanism is illustrated in Figure 1.

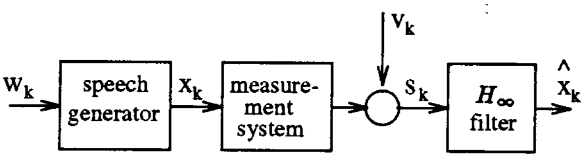

Figure 1. The speech generating mechanism.

The AR source model and the observation model can be written into canonical state space models (1) [2]-[3]. The optimal H_∞ filtering algorithm (28)-(30) and (21) is applied to the state space model to estimate the state vector where the last component of which gives the estimated (enhanced) speech signal. The algorithm is given by

$$\hat{z}_k^* = L_k \hat{x}_k^*, \quad k = 0, 1, \ldots, N-1, \quad L_k = 1 \tag{33}$$

$$\hat{x}_{k+1}^* = A_k \hat{x}_k^* + K_k (y_k - C_k \hat{x}_k^*), \quad \hat{x}_0 = x_0 \tag{34}$$

$$K_k = A_k P_k (I - \gamma \bar{Q}_k P_k + C_k^T V_k^{-1} C_k P_k)^{-1} C_k^T V_k^{-1} \tag{35}$$

$$P_{k+1} = A_k P_k A_k^T + B_k W_k B_k^T + A_k P_k (\gamma \bar{Q}_k - C_K^T V_k^{-1} C_k) \cdot [I - P_k (\gamma \bar{Q}_k - C_k^T V_k^{-1} C_k)]^{-1} P_k A_k^T, \quad P_0 = p_0 \tag{36}$$

Since the H_∞ filtering algorithms require the knowledge of model parameters, application of the H_∞ filtering for speech enhancement consists of two steps: at each iteration, (a) apply the Expectation-Maximization algorithm [10] for training the parameters of the dynamical model; (b) filter the signals with the H_∞ filter. Repeats these two steps until convergence.

The discrete H_∞ estimator is designed based on an upper bound on the noise signals, it is more robust and more appropriate in the practical speech enhancement. All estimates quantities are subjects to errors. The classical methods of estimation (Wiener and Kalman filtering approach) usually provide a good estimate with minimum least squares error. The proposed estimation algorithm minimizes the H_∞-norm of the map from exogenous inputs (noise) to the estimation error of the linear combination of the states. The strength of the H_∞-estimation results lies in its superior performance at the peak error range, and in its impressive robustness to parameter uncertainty.

4. CONCLUSIONS

A difference game has been formulated and solved. The existence of a solution to the difference Riccati equation, over the time interval, is a necessary and sufficient condition for the existence of the optimal discrete H_∞ filter. Since the design criterion is based on the worst case disturbance, the H_∞-filter is less sensitive to uncertainty in the exogenous signals statistics and dynamical model. The application of the H_∞ filtering to the speech enhancement is also discussed.

ACKNOWLEDGEMENT

The authors would like to thank the Natural Science and Engineering Research Council (NSERC) of Canada for the financial support of this work.

REFERENCES

[1] J.S. Lim and A.V. Oppenheim, "All-Pole Modeling of Degraded Speech", *IEEE Trans. Acoust., Speech, Signal Processing*, Vol. ASSP-26, 197-210, 1978.

[2] K.K. Paliwal and A. Basu, "A Speech Enhancement Method Based on Kalman Filtering", Proc. IEEE ICASSP, 177-180, 1987.

[3] J.D. Gibson, B. Koo and S.D. Gray, "Filtering of Colored Noise for Speech Enhancement and Coding", *IEEE Trans. Signal Processing*, Vol. SP-39, 1732-1742, 1991.

[4] Y. Ephraim, "A Minimum Mean Square Error Approach for Speech Enhancement", *Proc. IEEE ICASSP*, 829-832, 1990.

[5] H. Sameli, H. Sheikhzaheh, L. Deng and R. L. Brennan, "Statistical Speech Enhancement for Nonstationary Noise", submitted for publication.

[6] R.N. Banavar and J.L. Speyer, "A Linear Quadratic Game Theory Approach to Estimation and Smoothing", *Proc. IEEE ACC*, 2818-2822, 1991.

[7] M.J. Grimble and A. Elsayed, "Solution of the H_∞ Optimal Linear Filtering Problem for Discrete-Time Systems", *IEEE Trans. Acoust., Speech, Signal Processing*, Vol. ASSP-38, 1092-1104, 1990.

[8] U. Shaked and Y. Theodor, "H_∞-Optimal Estimation: A Tutorial", *Proc. 31st IEEE CDC*, 2278-2286, 1992.

[9] K.M. Nagpal and P.P. Khargonekar, "Filtering and Smoothing in an H_∞ Setting", *IEEE Trans. Auto. Control*, Vol. AC-36, 152-166, 1991.

[10] V. Digalakis, J.R. Rohlicek, and M. Ostendorf, "ML Estimation of a Stochastic Linear System with the EM Algorithm and Its Application to Speech Recognition", *IEEE Trans. Speech, Audio Processing*, Vol. SAP-1, 431-442, 1993.