ICSV13 - Vienna
The Thirteenth International Congress on Sound and Vibration
Vienna, Austria, July 2-6, 2006

DYNAMIC, NOISY CHANNEL DECONVOLUTION: A MODEL BASED APPROACH

Michael J Roan[1] and Mark Gramann[2]

[1]Department of Mechanical Engineering, Virginia Polytechnic Institute and State University
141 Durham Hall, Blacksburg, VA 24060
[2] Graduate Program in Acoustics, The Pennsylvania State University
mroan@vt.edu

Abstract
Blind Deconvolution (BDC) algorithms typically assume a noiseless signal model, and a stationary signal propagating through a static channel. In real world systems this is almost never the case (e.g. the problem of interest in this paper: a noisy multipath propagation environment where the source and receiver are moving). In such cases, it is proposed that model-based techniques be applied to incorporate further a priori information about the system into the existing blind processing framework. The significant original contributions of this work are as follows: First, a modified formulation of the extended Kalman filter (EKF) is developed that allows incorporation of a priori information into gradient-based blind processing algorithms. This formulation is then applied to the existing Natural Gradient BDC algorithm. Finally, simulation results are presented that suggest significant improvement in signal recovery performance through application of the modified EKF to the NG BDC algorithm for dynamic noisy channels

INTRODUCTION

It is common in Acoustics and communications to measure a signal that has been degraded by propagation through an unknown channel prior to measurement. While only the degraded measured signal may be available for processing, the actual data of interest may be the original signal or perhaps the filtering properties of the channel itself. In these situations, it may be desirable to reverse the filtering process through the application of an inverse filter to recover the original signal. In a situation where neither the input signal properties nor the channel properties are deterministically known, this problem is known as blind deconvolution (BDC).Typically, BDC algorithms assume that the system is noiseless, that the propagation channel is static, and that the source signal is stationary. However, these assumptions are often violated in many real world cases. One class of examples that violate the former two assumptions consists of cases including noisy multipath propagation environments where the source and/or receiver locations are time varying. In such cases, it is proposed that model-based techniques be applied to incorporate further a priori information about the system into the existing blind processing framework. The significant original contributions of this work are as follows. First, a modified formulation of the extended Kalman filter (EKF) is proposed which allows incorporation of a priori information into gradient-based blind processing algorithms. This formulation is then applied to the existing Natural Gradient (NG) [1] BDC algorithm. Finally, simulation results are presented which suggest significant improvement in signal recovery performance through application of the modified EKF to the NG BDC algorithm for dynamic noisy channels.

BACKGROUND

BDC Problem Formulation

The BDC problem is typically considered as a discrete time process, where the discrete sequences result from sampling the signals at the symbol rate Therefore, consider a system with an unknown discrete input signal s(k) that propagates through some channel with impulse response **h**(k). The actual filtering process of the channel may be described by a number of forms, however it is convenient to consider the filtering in terms of its vector (bold) impulse response. It is assumed that the signal available for measurement has been filtered by this channel and can be represented as the convolution of the input signal s(k) with the channel impulse response vector **h**(k). This measured signal is then given by

$$x(k)=\sum_{l=0}^{M}h_{l}(k)s(k-l) \tag{1}$$

For the case where h(k) is of finite length, M+1.

The term *deconvolution* refers to the process by which the effects of this convolution are reversed by convolving the measured signal x(k) with an estimated inverse impulse response $w(k)=[w_0(k),...,w_L(k)]$. The desired result is then to find $w(k)\approx h^{-1}(k)$ such that $w(k)*h(k)\approx \delta(k-m)$ where m is an arbitrary constant delay. Then the input signal can be approximated as

$$s(k-m)\approx y(k)=\sum_{l=0}^{L}w_{l}(k)x(k-l) \tag{2}$$

Note that in keeping with the notation of [1,3], the length of the inverse filter is specified as L+1 taps.

The designation of this process as *blind* refers to the assumption that neither the source signal nor the channel impulse response are known deterministically. In the case either of these is known a priori, then the problem may still be non-trivial, but may be approached using techniques such as Least Mean Squares (LMS) .

Natural Gradient BDC Algorithm

The NG BDC algorithm was presented by Amari, et al [1,3] and is a gradient based technique that maximizes the entropy of the signal y(k) with respect to the estimated inverse filter **w**(k). The resulting weight updating equation is then described by

$$\Delta w_l(k) = w_l(k) - \phi(y(k-L))u(k-l) \tag{3}$$

Where

$$u(k) = \sum_{q=0}^{L} w_{L-q}(k)\, y(k-q) \tag{4}$$

and $\phi(.)$ is a nonlinear activation function [1,3]. The choice of the nonlinear function depends largely on the probability density function (PDF) of the input signal s(k) [4]. For a super-Gaussian (positive kurtosis) input, such as human speech, it is common to use $\phi(y) = \tanh(y)$ or. For a sub-Gaussian (negative kurtosis) input, such as a digital communication signal, it is common to use $\phi(y) = y^3$.

EKF-BDC FORMULATION

To apply model based processing to the BDC problem, we make several modifications to the EKF algorithm. The result is an unconventional formulation of the state space model and the EKF for blind parameter estimation. This formulation will be referred to as EKF-BDC1.

Development

In the blind deconvolution problem, there are two variables that must be estimated: the input signal s(k) and the inverse filter **w**(k) Because we are concerned with the case where the inverse signal has stationary statistics and the inverse filter varies according to a known model, we assume that the unknown state vector that we are attempting to estimate is the inverse filter $\mathbf{w}(k)$. Because of the blind nature of the problem, and the assumption of a random input signal s(k), there is insufficient information to predict the actual measurement x(k) from this state vector alone without deterministic knowledge of s(k). Because the true measurement cannot be predicted, we then assume that our ``measurement'' is the output of the gradient update equations.

In this case, we use the NG BDC gradient update equation given in Eq. (3). These equations comprise a function of the last estimate $\mathbf{w}(k-1)$, the current actual measurement x(k), and past actual measurements and intermediate calculations x(k-j), y(k-j) and u(k-j), as given in Eqs. (1), (2), and (4), respectively, for $j \in \{1, ...L\}$, where L+1 is the length of the inverse filter **w**(k) We therefore augment the state vector with these values, so that it takes the form

$$\theta(k) = [w_0(k),..., w_L(k), \quad u(k),...,u(k-L+1), \\ y(k),...y(k-L+1), \quad x(k),...,x(k-L+1)]^T \tag{5}$$

Or

$$\theta(k) = [w_0(k),..., w_L(k), \quad u_0(k),...,u_{L-1}(k), \\ y(k),...y_{L-1}(k), \quad x_0(k),...,x_{L-1}(k)]^T \tag{6}$$

where the notation $u_j(k) = u(k-j)$ and likewise for \$y\$ and \$x\$, is introduced for convenience. This state vector defined above contains the required values for the NG BDC update, however it does not contain any additional necessary values for the state prediction. Any such variables must be appended to the state vector defined in Eq (6).

The *measurement calculation* then takes the form (7)

$$\mathbf{w}(k) = g(\hat{\theta}(k-1 \mid k-1, x(k))$$

where $g(\hat{\theta}(k-1 \mid k-1), x(k))$, x(k))\$ is a vector form of the nonlinear Natural Gradient deconvolution weight updating equation given in Eq.3. The predicted measurement is then given by

$$\mathbf{w}(k \mid k-1) = c(\hat{\theta}(k \mid k-1) \tag{8}$$

where in this case the function $c(\hat{\theta}(k \mid k-1)$ simply extracts the vector $[w_0(k),..., w_L(k)]$ from the state estimate and can take the form of a linear matrix operator.

The innovation is then calculated as the vector difference

$$e(k) = w(k) - \hat{w}(k \mid k-1) \qquad (9)$$

While the measurement function in Eq. (8) is linear, the measurement calculation function in Eq(8) is a nonlinear transform of the state variable $\hat{\theta}(k-1 \mid k-1)$ and of the noisy variable x(k). Therefore, the innovation covariance calculation must then account for the uncertainty of these variables by propagating the covariances of x(k) and $\hat{\theta}(k-1 \mid k-1)$ using linearizations of Eq. (8) with respect to x(k) and $\hat{\theta}(k-1 \mid k-1)$.

Because the processing task is blind, there is no reference signal u(k) on which to base the state prediction. We therefore use the noisy measurement x(k) in its place, and the state prediction takes the form

$$\hat{\theta}(k \mid k-1) = f(\hat{\theta}(k-1 \mid k-1), x(k)) \qquad (10)$$

Where $f(\hat{\theta}(k-1 \mid k-1), x(k))$, is a function that describes the dynamics of the system. It is noted that $f(\hat{\theta}(k-1 \mid k-1), x(k))$ may be linear or nonlinear. Because the state prediction is now a function of the noisy measurement x(k), the covariance R_v of the noise in this measurement must be accounted for in the state covariance prediction by propagating the covariance of x(k) through a linearization of Eq. (10)

It should be noted that in substituting x(k) for u(k-1), the index is changed such that in Eq. (10) the prediction actually uses the current, or k^{th} measurement to predict the value of $\theta(k)$. While this would seem to conflict with the notion of prediction and correction, the values of $\theta(k \mid k-1)$ that are calculated using x(k) are not actually used in predicting the state variables of interest, namely the estimation of $\hat{\mathbf{w}}(k \mid k-1)$. Instead, the values that depend on x(k) are used only in the calculation of the following ``measurement" that occurs at time k+1. As such, these values and their corresponding entries in the state covariance matrix are not changed by the *a posteriori* state correction. This implementation is made for the sake of convenience, but it is possible to reformulate the algorithm such that the ``prediction" step depends only on measurements from time k-1 without affecting the performance of the system. This is done in the formulation given in SecV. There, the formulation avoids this issue by restructuring the state space such that the prediction step does not depend on x(k).

The proposed algorithm is summarized below in Eqs. (11)-(24). The signal path described by these equations can be seen in block diagram form in Figs. 1 and 2.

Prediction

$$\hat{\theta}(k \mid k-1) = f(\hat{\theta}(k-1 \mid k-1), x(k)) \tag{11}$$

$$\hat{P}(k \mid k-1) = A(k-1)\hat{P}(K-1 \mid k-1)A^T(k-1) + B(k)R_v B^T(k) + R_r \tag{12}$$

where

$$A(k) = \left[\frac{\partial f_i}{\partial \theta_j}\right]_{ij}\Bigg|_{\theta=\hat{\theta}(k-1|k-1)} \tag{13}$$

$$B(k) = \left[\frac{\partial f_i}{\partial x}\right]_{i}\Bigg|_{x=x(k)} \tag{14}$$

$$\mathbf{w}(k \mid k-1) = c(\hat{\theta}(k \mid k-1) \tag{15}$$

Measurement

$$\mathbf{w}(k) = g(\hat{\theta}(k-1 \mid k-1, x(k)) \tag{16}$$

Innovation

$$\mathbf{e}(k) = \mathbf{w}(k) - \hat{\mathbf{w}}(k \mid k-1) \tag{17}$$

$$R_e = C(k)\hat{P}(k \mid k-1)C^T(k) + D(k)\hat{P}(K-1 \mid k-1)D^T(k) + E(k)R_v E^T(k) \tag{18}$$

Where

$$C(k) = \left[\frac{\partial c_i}{\partial \theta_j}\right]_{ij}\Bigg|_{\theta=\hat{\theta}(k|k-1)} \tag{19}$$

$$D(k) = \left[\frac{\partial g_i}{\partial \theta_j}\right]_{j}\Bigg|_{\theta=\hat{\theta}(k-1|k-1)} \tag{20}$$

$$E(k) = \left[\frac{\partial g_i}{\partial x}\right]_{i}\Bigg|_{x=x(k)} \tag{21}$$

Gain

$$K(k) = \hat{P}(K \mid k-1)C^T(k)R_e^{-1}(k) \tag{22}$$

Correction

$$\hat{\theta}(k \mid k) = \hat{\theta}(k \mid k-1) + K(k)\mathbf{e}(k) \tag{23}$$

$$\hat{P}(k \mid k) = [I - K(k)C(k)]\hat{P}(k \mid k-1) \tag{24}$$

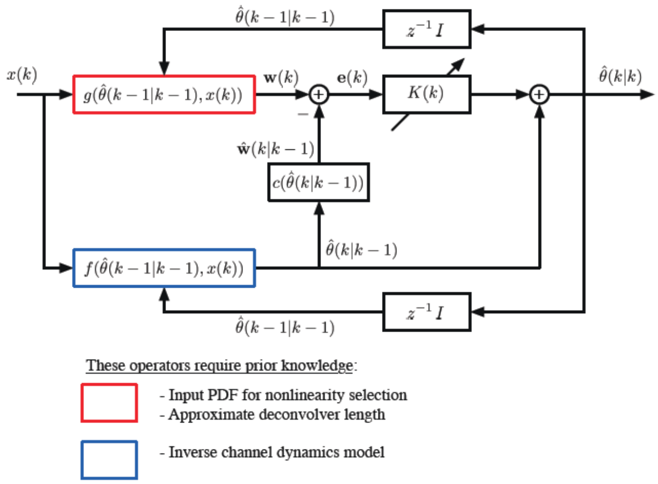

Figure 1: Block diagram of the signal path for proposed EKF-BDC1 algorithm. The function $g(\hat{\theta}(k-1|k-1), x(k))$ is the natural gradient weight updating equation. The function $f(\hat{\theta}(k-1|k-1), x(k))$ is the state predictor based on the channel dynamics model. The Kalman gain matrix K(k) is updated as shown in Fig. 2.

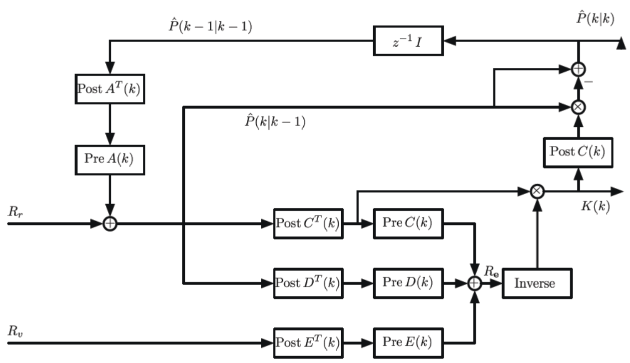

Figure 2: Block diagram of the Kalman gain and state covariance estimate calculation. The state and measurement noise covariances R_r and R_v are assumed to be known through prior knowledge. The matrices A(k), B(k), C(k), D(k), and E(k) are calculated at each time k according to Eqs. (13), (14) (19), (20), and (21), The notations ``Pre A" and ``Post A" indicate premultiplication and postmultiplication, respectively, by the matrix A.

Preliminary investigations with this algorithm focused on the case of a stationary channel in the presence of noise. Under this assumption, the channel dynamic model is assumed to be described by a first order Gauss-Markov process. As a result, the state variable of interest, namely the vector of inverse filter coefficients, $\mathbf{w}(k)$ is assumed to evolve according to

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mathbf{r}(k) \tag{25}$$

where **r**(k) is a vector of zero mean Gaussian white noise with covariance R_r. Noting that the calculation of values $[u_1(k),...,u_{L-1}(k)]$, $[y_1(k),...,y_{L-1}(k)]$, $[x_1(k),...,x_{L-1}(k)]$ are merely a shift of the previous values, the state prediction function then takes the relatively simple form

$$\hat{\theta}(k\,|\,k-1)=\begin{bmatrix} w_0(k)\\ \vdots \\ w_L(k)\\ \hline u_0(k)\\ u_1(k)\\ \vdots\\ u_{L-1}(k)\\ \hline y_0(k)\\ y_1(k)\\ \vdots\\ y_{L-1}(k)\\ \hline x_0(k)\\ x_0(k)\\ \vdots\\ x_{L-1}(k)\end{bmatrix}=f(\hat{\theta}(k-1\,|\,k-1),x(k)=\begin{bmatrix} w_0(k-1)\\ \vdots\\ w_L(k-1)\\ \hline \sum_{q=0}^{L} w_{L-q}(k-1)y_q(k-1)\\ u_0(k-1)\\ \vdots\\ u_{L-2}(k-1)\\ \hline \sum_{q=0}^{L} w_q(k-1)x_q(k-1)\\ y_0(k)\\ \vdots\\ y_{L-2}(k-1)\\ \hline x\ (k)\\ x_0(k-1)\\ \vdots\\ x_{L-2}(k-1)\end{bmatrix} \quad (26)$$

C. Discussion

This formulation is unconventional in several ways. First, the primary part of the state vector is the vector of inverse filter coefficients. Second, the state vector contains a large number of intermediate variables. The result is an unwieldy vector of length 4L+1 (not including variables to describe the state dynamics) requiring a (4L+1)x(4L+1) Jacobian matrix A(k). Third, the noisy true measurement x(k) is used as a reference signal for the state prediction in addition to being used as an input to the measurement calculation.

REFERENCES

[1| Amari, S. ., "Natural Gradient Works Efficiently in Learning", *Neural Computation,* Vol. 10, pp.251-276, 1998.

[2] Amari, S. Cichochi, A., and Yang, H.H., "Multi-Channel Blind Deconvolution and Equalization Using the Natural Gradient," *Proc. First IEEE Signal Processing Workshop on Signal Processing Advances in Wireless Communications*, April 1997, pp. 101-104.

[3] Amari, S., Douglas, S., Cichochi, A., and Yang, H.H., "Novel Online Adaptive Learning Algorithms for Blind Deconvolution Using the Natural Gradient Approach," *Proc. Proc. 11th IFAC Symposium on System Identification*, July 1997, pp. 1057-1062.

[4] Bell, A.J., and Sejnowski, T., J., Blind Source Separation and Blind Deconvolution, and Information Theoretic Approach, Proc. IEEE Sonf. On Speech, Acoustics, And Signal Processing, vol.5. 1995, pp. 3415-3418.