2005 International Symposium on Nonlinear
Theory and its Applications (NOLTA2005)
Bruges, Belgium, October 18-21, 2005

# System Identification using Constrained Kalman Filters

David M. Walker†

†Biomathematics and Statistics Scotland, Macaulay Institute, Craigiebuckler, Aberdeen AB15 8QH
Email: d.walker@bioss.ac.uk

**Abstract**—We suggest incorporating dynamical information such as locations of fixed points into parameter estimation algorithms in order to improve the method of reconstructing dynamics from time series. We show how the process of reconstruction using the extended or iterated Kalman filter can be easily modified to include the additional information. We demonstrate the methods using data from the Chua circuit operating in the chaotic regime. We find the models reconstructed by using constraints can better approximate the unstable fixed point structure of the underlying systems.

## 1. Introduction

This paper is concerned with a dynamical-systems approach to nonlinear systems identification [1]. The approach is based on Takens Embedding Theorem [2] which allows reconstruction of an equivalent phase space to the original (unknown) system from observations in the form of a time series. The dynamics in this reconstructed phase space can be approximated using nonlinear functions. The estimation of model parameters can result in nonlinear optimization problems.

The extended Kalman filter (EKF) algorithm has been suggested as an alternative to more traditional methods for solving these optimization problems. In [4] the benefits of Kalman filtering for parameter estimation when neural networks are the approximating functions are shown. Puskorius & Feldkamp [5] had earlier pointed out how the parameter-based EKF training algorithm can be naturally incorporated into nonlinear control architectures. Burgmeier [6] and Walker & Mees [7] explain how all parameters in a radial basis function model can be estimated using the EKF algorithm.

The problem of reconstructing dynamics of systems by estimating parameters of nonlinear functions can be thought of as "black-box" modelling. One can imagine that a more "grey-box" modelling approach would produce, in some sense, improved nonlinear models. Grey-box modelling can be thought of as an approach to reconstruction where additional information is used in the modelling process. For example, if the system is known to possess symmetry, one could select basis functions which preserve such symmetries. It might be expected that the resulting models would exhibit dynamical properties of the system more closely. Indeed in [8] it was shown that basis functions which mimicked a fading memory assumption resulted in more accurate free-running models over standard radial basis function models. An alternative "grey-box" approach to modelling is to include properties of the system as constraints in the parameter estimation optimization problems.

It is known that the (unstable) periodic orbits of chaotic systems play a significant rôle in the structure of attractors [9]. We suggest using the location of unstable fixed points (UFP), as additional constraints in the parameter estimation process[1][10]. The EKF is a natural environment in which to incorporate such constraints. The location of UFP's and low-order periodic orbits can be determined from data [11, 12, 13]. The purpose of this paper is to demonstrate the idea of using UFP's as constraints within the EKF framework. The errors in estimating the location of UFP's can be used to advise on levels of noise covariances within the EKF.

In Section 2 we discuss the nonlinear functions we use to reconstruct dynamics. In Section 3 we outline the EKF algorithm in the context of parameter estimation and introduce the modifications required in order to use the Kalman filter with constraints. Our use of constraints can also naturally be incorporated into the iterated Kalman Filter (IKF). The approach is demonstrated using data from an electronic circuit system which can exhibit chaotic dynamics; the well known Chua circuit [15].

## 2. Nonlinear models

A discrete time model of a system with state $\mathbf{z} \in \mathbf{R}^L$ is described by

$$\mathbf{z}_{k+1} = \mathbf{F}[\mathbf{z}_k, \mathbf{a}] \quad (1)$$

where $\mathbf{a}$ represents the system parameters. The problem of modelling the dynamics from scalar time series requires the state $\mathbf{z}$ to be reconstructed. Given a scalar time series $\{y_t\}_{t=0}^{N}$ the state $\mathbf{z}_k$ can be represented by

$$\mathbf{z}_k = (y_{k-(m-1)\tau}, \ldots, y_{k-\tau}, y_k) \quad (2)$$

where $m$ is called the embedding dimension and $\tau$ is the time delay lag. There are a number of methods which can be used to determine an appropriate $\tau$ with a good prescription being to choose the first minimum of the average

[1] The stability of the fixed point does not appear to be critical and so locations of stable fixed points, if known, could also be used.

mutual information function [1]. Similarly, a recognized method for selecting a suitable embedding dimension is false nearest neighbours [1].

The dynamics $\mathbf{F}[\cdot]$ for a time delay embedding consist of a simple shift operator together with a (nonlinear) scalar-valued function $f$, i.e.,

$$y_{k+1} = f(\mathbf{z}_k). \tag{3}$$

There are a number of choices to use as the function $f$. A class of models which have been particularly successful in capturing nonlinear dynamics from data is the so-called pseudo-linear models [3] which include a linear combination of linear and nonlinear basis functions. This is the class of models we will consider in this study, namely,

$$y_{k+1} = \sum_{i=1}^{K} \omega_i \phi(\|\mathbf{c}_i - \mathbf{z}_k\|) + \sum_{i=0}^{m} \alpha_i y_{k-i\tau} + \beta. \tag{4}$$

The function $\phi(\cdot)$ is a radial basis function and $\omega_i$, $\alpha_i$ and $\beta$ are weight parameters to be determined. The $\mathbf{c}_i$ are the radial basis centres whose locations must also be determined. In this paper we use the following basis function

$$\phi(u) = \frac{1}{1 + \cosh(-\delta u)}, \quad u \geq 0. \tag{5}$$

The scale factor $\delta$ is another parameter requiring determination but we set it equal to 1 here. We will explain how the EKF can be used to estimate all unknown parameters of a pseudo-linear model but for ease of exposition we restrict the problem to determining the weight parameters in the examples.

The steady-state or long term behaviour of a dynamical system is of particular interest. The properties of unstable fixed points and unstable periodic orbits of a system play an important rôle in determining this behaviour. An unstable fixed point of a discrete time dynamical system satisfies

$$\bar{\mathbf{z}} = \mathbf{F}[\bar{\mathbf{z}}, \mathbf{a}] \tag{6}$$

where $\bar{\mathbf{z}}$ denotes the UFP. The fixed points of a pseudo-linear model can be seen to satisfy

$$p = f(\mathbf{p}) \tag{7}$$

where $\mathbf{p} = (p, p, \ldots, p) \in R^m$ is the fixed point location in reconstructed phase space. There have been a number of methods proposed in the literature for extracting dynamical information such as UFP locations. These include studying close returns [11], or investigating properties of locally reconstructed models [12, 13]. We use a modification of the method in [12] to determine UFP's where neighbours in reconstructed space rather than time are used to construct local dynamics.

## 3. Extended Kalman filter and parameter estimation

The Kalman filter and its variants are statistical state estimators. The algorithms are typically used to achieve noise reduction in signal processing but can be applied to the problem of parameter estimation. The system model required by the Kalman filter for this particular application is

$$\begin{aligned} \mathbf{a}_t &= \mathbf{a}_{t-1} \\ y_t &= f(\mathbf{z}_t, \mathbf{a}_t) + n_t \end{aligned} \tag{8}$$

where $\mathbf{a}_t$ represents the model parameters. These can be the pseudo-linear weights, the location of the centres and the values of radial basis function scale parameters. There is no additive noise term in the evolution of $\mathbf{a}_t$ as the parameters are assumed to be stationary. The second equation is the pseudo-linear model used to predict the time series values $y_t$ and $\mathbf{z}_t$ is the embedded time series data. The noise term $n_t$ can be thought of as the fitting error and is assumed to be from a Gaussian distribution, i.e., $n_t \sim N(0, R)$. The parameters are initialized with $\mathbf{a}_0 \sim N(\hat{\mathbf{a}}_0, \mathbf{P}_{\mathbf{a}_0})$ where $\hat{\mathbf{a}}_0 = E[\mathbf{a}_0]$ and $\mathbf{P}_{\mathbf{a}_0} = E[(\mathbf{a}_0 - \hat{\mathbf{a}}_0)^T(\mathbf{a}_0 - \hat{\mathbf{a}}_0)]$. The EKF update equations for the system (8) takes the form [16, 7, 4]

**time-update equations**

$$\begin{aligned} \hat{\mathbf{a}}_t^- &= \mathbf{a}_{t-1} \\ \mathbf{P}_{\mathbf{a}_t}^- &= \mathbf{P}_{\mathbf{a}_{t-1}} \end{aligned}$$

**measurement-update equations**

$$\begin{aligned} \mathbf{K}_t &= \mathbf{P}_{\mathbf{a}_t}^- \mathbf{C}_t^T (\mathbf{C}_t \mathbf{P}_{\mathbf{a}_t}^- \mathbf{C}_t^T + \mathbf{R})^{-1} \\ \hat{\mathbf{a}}_t &= \hat{\mathbf{a}}_t^- + \mathbf{K}_t[\mathbf{y}_t - f(\hat{\mathbf{z}}_t, \hat{\mathbf{a}}_t)] \\ \mathbf{P}_{\mathbf{a}_t} &= (\mathbf{I} - \mathbf{K}_t \mathbf{C}_t)\mathbf{P}_{\mathbf{a}_t}^- \end{aligned}$$

where

$$\mathbf{C}_t = \frac{\partial f(\mathbf{z}, \mathbf{a})}{\partial \mathbf{a}}\Big|_{\hat{\mathbf{a}}_t}$$

The matrix $\mathbf{K}_t$ is referred to as the *Kalman gain*. The terms $[\mathbf{y}_t - f(\hat{\mathbf{z}}_t, \hat{\mathbf{a}}_t)]$ are known as the innovations.

### 3.1. Constrained Kalman Filtering

The extended Kalman filter has proved to be successful in estimating the parameters of nonlinear models which capture the dynamical behaviour of nonlinear systems [7]. We suggest treating important properties of the system such as UFP locations as constraints to be met as accurately as possible. The setup of the Kalman filter is readily amenable to including such constraints. We do this by modifying the observation model of (8) to include the model's approximation to the UFP's. The system model for applying the Kalman Filter with constraints, referred to in the sequel as the KFC algorithm, becomes

$$\begin{aligned} \mathbf{a}_t &= \mathbf{a}_{t-1} \\ y_t &= f(\mathbf{z}_t, \mathbf{a}_t) + \mathbf{n}_t \\ 0 &= p - f(\mathbf{p}, \mathbf{a}_t) + m_t \end{aligned}$$

where the new term includes $m_t \sim N(0, R_p)$ where $R_p$ reflects the accuracy of the knowledge of the fixed point location. The value $R_p$ is set to can be guided by the error in estimates of the UFP's. The filter update equations are essentially unchanged but the matrices and vectors in the measurement update equations now account for $N_p + 1$ observation functions, where $N_p$ are the number of UFP constraints we try to meet. Explicitly the measurement update equations become

$$\begin{aligned}
\mathbf{K}_t &= \mathbf{P}_{\mathbf{a}_t}^- \mathbf{C}_t^T (\mathbf{C}_t \mathbf{P}_{\mathbf{a}_t}^- \mathbf{C}_t^T + \mathbf{R}')^{-1} \\
\hat{\mathbf{a}}_t &= \hat{\mathbf{a}}_t^- + \mathbf{K}_t [\begin{pmatrix} \mathbf{y}_t - f(\hat{\mathbf{z}}_t, \hat{\mathbf{a}}_t) \\ 0 - (p - f(\mathbf{p}, \mathbf{a}_t)) \end{pmatrix}] \\
\mathbf{P}_{\mathbf{a}_t} &= (\mathbf{I} - \mathbf{K}_t \mathbf{C}_t) \mathbf{P}_{\mathbf{a}_t}^-
\end{aligned}$$

where

$$\mathbf{C}_t = \begin{pmatrix} \frac{\partial f(\mathbf{z},\mathbf{a})}{\partial \mathbf{a}}|_{\hat{\mathbf{a}}_t} \\ \frac{\partial f(\mathbf{p},\mathbf{a})}{\partial \mathbf{a}}|_{\hat{\mathbf{a}}_t} \end{pmatrix} \text{ and } R' = \begin{pmatrix} R & 0 \\ 0 & R_p \end{pmatrix}$$

The time update equations remain unchanged.

### 3.2. Iterated Kalman Filter

The iterated Kalman Filter is a modification of the Kalman filter conceived by the intuition that using the new state estimate to improve the "innovation" should lead to a better update of the state estimate. That is, we re-linearize the observation function about the current state estimate in an iterative fashion at each measurement update step before advancing to the next time update step. The updated equations are

$$\begin{aligned}
\mathbf{K}_t^i &= \mathbf{P}_{\mathbf{a}_t}^- \mathbf{C}_t^{iT} (\mathbf{C}_t^i \mathbf{P}_{\mathbf{a}_t}^- \mathbf{C}_t^{iT} + \mathbf{R}')^{-1} \\
\hat{\mathbf{a}}_t^{i+1} &= \hat{\mathbf{a}}_t^- + \mathbf{K}_t^i (y_t - c(\hat{\mathbf{a}}_t^i) - \mathbf{C}_t^i (\hat{\mathbf{a}}_t^- - \hat{\mathbf{a}}_t^i))
\end{aligned}$$

The IKF works by iterating over $i$ $M$-times and then updating the covariance by $\mathbf{P}_{\mathbf{a}_t} = (\mathbf{I} - \mathbf{K}_t^M \mathbf{C}_t^M) \mathbf{P}_{\mathbf{a}_t}^-$ before advancing to the next time update step. The IKF gives greater weight to the observations over the dynamics as $M$ increases (see [14] for an example). The KFC can clearly be modified in a like manner by including the constraints in an IKF update step.

## 4. Examples

In this section we examine the above ideas with application to Chua's circuit [15] operating in the regime which produces a double scroll attractor. There are three fixed points given by $(x^*, y^*, z^*) = (0, 0, 0)$ and $(x^*, y^*, z^*) = (r, 0, -r)$ with $r = \pm\frac{m_0 - m_1}{1 + m_1}$ where $m_0 = -8/7$ and $m_1 = -5/7$. We integrated the Chua circuit equations with initial condition $(-1, 0, 0)$ from $t = 0$s to $t = 2000$s sampling every 0.1s and then retained the last 3000 data points. We observe the $x_t$ component of the state subject to observational noise distributed as $N(0, 0.1)$ and reconstruct predictive models of the circuit from this time series.

We circumvent the problem of model selection by an initial screening process. We reconstruct radial basis models with increasing numbers of centres chosen by k-means. For each model we estimate the parameters (weights) using least squares and calculate the Schwarz information criterion (SIC) [17] and select a model size which results in a minimum. A calculation of mutual information suggests a time delay lag of 6 and false nearest neighbours with the selected lag suggests a 3-dimensional embedding. The screening process using SIC suggested a model size somewhere around 35.

A portion of the time series is used to estimate UFP's of the system to be used in the KFC algorithm. The algorithm in [12] modified for neighbours in space suggests two definite fixed points with the possibility of a third at the origin coinciding with the true fixed points of the Chua circuit. We apply the KFC algorithm in four cases: (i) only the UFP at the origin is used as a constraint (ii) the two definite UFP's at $-r$ and $r$ suggested by the algorithm (iii) all three UFP's are used as constraints and (iv) the IKF with constraints for $M = 50$ – an IKFC – is applied using all three UFP's. The noise distribution representing the uncertainty in the UFP determination is given by the distribution $N(p, 0.1)$. The noise term associated with the models fit error is set to $N(0, 2\sigma_y^2)$ where $\sigma_y$ is the standard deviation of the time series data. The initial parameter weights are set to zero except for one so that the initial model is given by $y_{t+1} = y_t$.

The results are summarized in Table 1. The accuracy of the LS and EKF models with respect to approximating the UFP's is rather poor. The performance of the KFC models with respect to this measure clearly demonstrates the benefits of incorporating the UFP's as constraints. The KFC(0) model accurately models the fixed point at the origin but not the other two. The KFC$(-r, r)$ model approximates the $\pm r$ fixed points very well but the origin is predicted poorly. Meanwhile the KFC$(-r, 0, r)$ model is able to approximate all UFP's extremely accurately with only a small loss in overall RMS error. In this example there does not appear to be an advantage of using the iterated form of the filter, however, for more complicated systems, or additional constraints it may be a worthwhile option. Figure 1 gives a view of the fixed point accuracy of the local accuracy of the LS, EKF and KFC$(-r, 0, r)$ models.

## 5. Summary

We have described the idea of using the fixed point structure of a dynamical system to aid parameter estimation of nonlinear models. This information can take the form of constraints in the optimization problem associated with reconstruction. The extended Kalman filter algorithm provides a natural framework in which to incorporate such constraints; one simply includes a measure of model fixed point accuracy as an extra observation function in the EKF system model. We showed that the KFC models capture the

| | LS | EKF | KFC (0) |
|---|---|---|---|
| 0 | 0.01 | 0.006 | $9.53 \times 10^{-6}$ |
| $(-r)$ | 0.019 | 0.041 | 0.041 |
| $r$ | 0.013 | 0.037 | 0.037 |

| | KFC $(-r, r)$ | $KFC(-r, 0, r)$ | $IKFC(-r, 0, r)$ |
|---|---|---|---|
| 0 | $5.67 \times 10^{-3}$ | $9.68 \times 10^{-6}$ | $9.68 \times 10^{-6}$ |
| $(-r)$ | $1.15 \times 10^{-4}$ | $1.16 \times 10^{-4}$ | $1.16 \times 10^{-4}$ |
| $r$ | $1.30 \times 10^{-4}$ | $1.30 \times 10^{-4}$ | $1.30 \times 10^{-4}$ |

Table 1: Model accuracy of fixed point estimation for the Chua circuit.

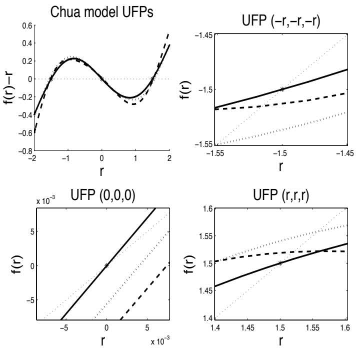


Figure 1: Accuracy of Chua model fixed points. The top left panel shows a plot of $f(\mathbf{r}) - r$ verses $r$ for the models obtained using least squares (dashed), EKF (dotted) and KFC algorithm with three fixed point constraints (solid). The IKFC is a dot-dash line indistinguishable in the plot from the KFC line. The remaining three panels are close up views of the fixed point accuracy (UFP's are indicated by stars).

local dynamics structure better than models reconstructed using least squares and ordinary extended Kalman filters.

**Acknowledgments**

This work was funded by the Scottish Executive Environment and Rural Affairs Department (SEERAD).

## References

[1] H. D. I. Abarbanel. *Analysis of Observed Chaotic Data*. Springer, 1995.

[2] F. Takens. Detecting strange attractors in turbulence. In D. Rand and L. S. Young, editors, *Dynamical Systems and Turbulence, Warwick 1980*, page 366. Springer, 1981.

[3] K. Judd and A. Mees. On selecting models for nonlinear time series. *PHYSICA D*, 82:426–444, 1995.

[4] S. Haykin, editor. *Kalman filtering and neural networks*. Wiley, 2001.

[5] G. V. Puskorius and L. A. Feldkamp. Neuro-control of nonlinear dynamical systems with Kalman filter trained recurrent networks. *IEEE Transactions on Neural Networks*, 5(2):279–297, 1994.

[6] M. Burgmeier. A fully Kalman-trained radial basis function network for nonlinear speech modeling. In *Proceedings of IEEE International Conference on Neural Networks*, pages 259–264, 1995.

[7] D. M. Walker and A. I. Mees. Recontructing nonlinear dynamics by extended Kalman filtering. *International Journal of Bifurcation and Chaos*, 8(3):557–569, 1998.

[8] D. M. Walker, N. B. Tuffilaro, and P. Gross. Radial-basis models for feedback systems with fading memory. *IEEE Transactions on Circuits and Systems–I: Fundamental Theory and Applications*, 48(9):1147–1151, 2001.

[9] R. Artuso, E. Aurell, and P. Cvitanović. Recycling of strange sets: I. cycle expansions. *Nonlinearity*, 3:325, 1990.

[10] D. M. Walker. Parameter estimation using Kalman filters with constraints. *International Journal of Bifurcation and Chaos*, 2006 (in press).

[11] D. Auerbach, P. Cvitanović, J-P. Eckmann, and G. Gunaratne. Exploring chaotic motion through periodic orbits. *Physical Review Letters*, 58(23):2387–2389, 1987.

[12] J. Glover and A. Mees. Reconstructing the dynamics of Chua's circuit. *Journal of Circuits, Systems, and Computers*, 3(1):201–214, 1993.

[13] P. So, E. Ott, S. J. Schiff, D. T. Kaplan, T. Sauer, and C. Grebogi. Detecting unstable periodic orbits in chaotic experimental data. *Physical Review Letters*, 76(25):4705–4708, 1996.

[14] D. M. Walker and A. I. Mees. Noise reduction of chaotic systems by Kalman filtering and by shadowing. *International Journal of Bifurcation and Chaos*, Vol. 7, No. 3 1997.

[15] T. Matsumoto, L. O. Chua, and M. Komoro. Birth and death of the double scroll. *PHYSICA D*, 24:97, 1987.

[16] O. L. R. Jacobs. *Introduction to Control Theory*. Oxford Science Publications, 2nd edition, 1993.

[17] G. Schwarz. Estimating the dimension of a model. *Annals of Statistics*, 6:461–464, 1978.