PARTICLE BASED CONFIDENCE INTERVALS

Fabien Campillo, INRIA/IRISA, France
Gilles Canales, INRIA/IRISA, France
Laurent Mevel, INRIA/IRISA, France


Abstract

This paper deals with the problem of real time identification of the linear characteristics of the linear system associated with a mechanical structure. More than the parameter itself, the algorithm presented here is interested in estimating the confidence intervals for the parameter of interest. The algorithm is based on the particular filtering theory, where statistical laws of hidden states are approximated by some evolving particle collection.

1 Introduction

A critical problem for mechanical structures exposed to unmeasured non stationary natural excitation (turbulence) is an instability phenomenon also known as flutter. It is formulated as the monitoring of the time varying complex eigenvalues associated to the discretized linear system corresponding to the monitored mechanical system. It has already been investigated through batch identification modal analysis using only output-only in-flight data has already been investigated. See Mevel et al [8] for a case study of monitored aircraft using subspace identification methods.

For improving the estimation of the parameters of interest, the collection of frequency and damping coefficients, and moreover for achieving this in real-time during flight tests, one possible route is to resort to tracking algorithms.

Frequency and damping coefficients can be monitored by a recursive maximum likelihood (RML) procedure [1]. The considered tracking procedure is a special case of adaptive algorithms where the gain is kept constant. The associated score function is evaluated by a joint Kalman filter and its derivative w.r.t. to the parameters (the tangent filter). In the nonlinear case, these filters can be approximated by particle filtering techniques [5]. Doucet & Tadic [6], Guyader et al [7], and Caylus et al [3] already applied these techniques to RML estimation.

Particle approximation for health monitoring was already proposed by Yoshida & Sato [10] in order to handle non–Gaussian noise. Modal characteristics monitoring is also considered by Ching et al [4]. In both cases authors use a state augmentation approach by including the unknown parameters in the state process. This is also the method investigated here.

2 Modeling

2.1 Dynamical model and structural parameters

Let us consider observations sampled at a rate $1/\delta$

$$y_k = L Z(k \delta)$$  \hspace{1cm} (1)

of the state $Z(t)$ of a $n$–degrees of freedom mechanical system. These measurements are gathered through $d$ sensors, i.e. $y_k$ takes values in $\mathbb{R}^d$. The matrix $L$ indicates which components of the state
vector are actually measured, i.e. where the sensors are located. The behavior of the mechanical system is described by the following linear dynamical system

$$M \ddot{Z}(t) + C \dot{Z}(t) + K Z(t) = \sigma \zeta(t)$$

(2)

where the (non measured) input force $\zeta$ is a non–stationary white Gaussian noise with time-varying covariance matrix $\mathcal{Q}_\zeta(t)$. $M$, $C$, $K$ are respectively the matrices of mass, damping and stiffness.

Now let us describe the structural characteristics of the system (2). The modes or eigenfrequencies $\mu$ and the associated eigenvectors $\Phi_\mu$ of the system (2) are solutions of

$$\det[\mu^2 M + \mu C + K] = 0,$$

$$[\mu^2 M + \mu C + K] \Phi_\mu = 0.$$  

(3)

Then the mode–shapes are $\Psi_\mu = L \Phi_\mu$. The frequency and damping coefficients are

$$f = \frac{b}{2\pi} \text{ (Hz)}, \quad d = \frac{|a|}{\sqrt{a^2 + b^2}} \in [0, 1]$$

(4)

with $a = \Re(\mu)$ and $b = \Im(\mu)$.

The monitored structure is defined by its modal characteristics: the collection of frequencies, dampings and mode shapes, as well as the covariances of the noises. The problem is to follow the slow evolutions of the structural characteristics of the mechanical system (2) by a recursive tracking method, whose starting values will be defined as the output of the data driven subspace method as described in Van Overschee & De Moor [9, Fig. 3.13 p. 90].

The tracking algorithm will focus on the frequencies and dampings, the mode shapes are assumed not to change significantly during the monitoring in regard to the changes in the eigenvalues. A change in the mode shapes would most likely be a local change in the structure, thus will indicate the presence of damage, whereas a change in the eigenvalues can still occur without presence of damage and not affect significantly the mode shapes (as for example the effect of temperature on the stiffness of the structure).

2.2 State–space model and canonical parameterization

We rewrite the preceding system (1)–(2) as a linear state–space model. Define

$$X_k \overset{\text{def}}{=} \begin{bmatrix} Z(k\delta) \\ \dot{Z}(k\delta) \end{bmatrix}$$

and $F = e^{\delta A}$ with $A \overset{\text{def}}{=} \begin{bmatrix} 0 & -M^{-1} K \\ -M^{-1} C & I \end{bmatrix} \in \mathbb{R}^{2n \times 2n}$. From (2) we get

$$X_{k+1} = F X_k + \sigma \zeta_k$$

(5)

where $\zeta_k \overset{\text{def}}{=} \int_{k\delta}^{(k+1)\delta} \dot{\zeta}(s) \, ds$ is a Brownian motion. Hence $\zeta_k$ is a (discrete–time) white Gaussian noise with covariance matrix

$$\int_{(k-1)\delta}^{k\delta} e^{(k\delta-u)A} \begin{bmatrix} 0 & 0 \\ 0 & M^{-1} Q_\zeta(u)(M^{-1})^T \end{bmatrix} e^{(k\delta-u)A^T} \, du$$

which is approximated by $\delta \mathcal{Q}_\zeta^k$ with

$$\mathcal{Q}_\zeta^k \overset{\text{def}}{=} \begin{bmatrix} 0 & 0 \\ 0 & M^{-1} Q_\zeta(k\delta)(M^{-1})^T \end{bmatrix}.$$
From (1) we get

\[ y_k = [L \ 0] X_k + \nu v_k \]  

where \([L \ 0] \in \mathbb{R}^{d \times 2n}\) and \(v_k\) is a \(N(0, Q_k)\) white Gaussian noise which allows to take into account of the errors of modeling and the measurement noise. We suppose that the Hermitian matrix \(Q_k\) is positive definite.

Let \((\lambda, \Phi_\lambda)\) be the eigenstructure of the state transition matrix \(F\), namely

\[ \det(F - \lambda I) = 0, \quad (F - \lambda I) \Phi_\lambda = 0. \]  

The parameters \((\mu, \Phi_{\mu})\) in (3) can be deduced from the \((\lambda, \Phi_\lambda)\)'s using \(e^{\delta \mu} = \lambda\) and \(\Phi_{\mu} = \Phi_\lambda\). The frequency and damping coefficients (4) are recovered from a discrete eigenvalue \(\lambda\) through

\[ a = \frac{1}{\delta} \log |\lambda|, \quad b = \frac{1}{\delta} \arctan \left( \frac{\Im(\lambda)}{\Re(\lambda)} \right). \]

**Hypothesis:** We suppose that \(F\) admits \(2n\) pairwise complex conjugate distinct eigenvalues \(\lambda_1: n, \bar{\lambda}_1: n\) with associated orthonormal set of eigenvectors \(\Phi_1: n, \bar{\Phi}_1: n\) \(^1\). We also suppose that these eigenvalues have modulus less than one.

It turns out that this collection of modes forms a very natural parameterization for structural analysis. It is invariant w.r.t. changes in the state basis of system (5)–(6). In other words, the \((\lambda, \Phi_{\lambda})\)'s form a canonical parameterization of the eigenstructure (or equivalently the pole part) of that system.

**Change of variables**

Define

\[ \Phi \alpha= [\Phi_1: n], \quad \Psi \alpha= [\Psi_1: n], \quad \Lambda \alpha= \text{diag}(\lambda_1: n). \]

We introduce the following linear transformation

\[ T \alpha= [\Phi \ \bar{\Phi}] \in \mathbb{C}^{2n \times 2n}, \]

i.e. the matrix whose columns are the eigenvectors of \(F\). It is a unitary matrix, i.e. \(T^{-1} = T^*\). Then

\[ \begin{bmatrix} \Lambda & 0 \\ 0 & \Lambda \end{bmatrix} = T^* F T \in \mathbb{C}^{2n \times 2n}. \]

Define also

\[ H \alpha= [L \ 0] T = [L \ 0] [\Phi \ \Phi] = [\Psi \ \Psi] \in \mathbb{C}^{d \times 2n}. \]

Then after the change of variables

\[ \tilde{X}_k \alpha= T^* X_k, \]

the vector \(\tilde{X}_k\) is of the form \(\begin{bmatrix} x_k \\ \bar{x}_k \end{bmatrix}\) and (5) reduces to

\[ x_{k+1} = \Lambda x_k + \sigma \Phi^* \zeta_k, \quad \zeta_k \iid \sim N(0, \delta Q_k^\beta). \]

---

\(^1\)Notations: \(x^T\) is the transpose of \(x\), \(\bar{x}\) is the complex conjugate, \(x^*\) is the transpose/conjugate, \(|x|\) the modulus. \(j\) will denote \(\sqrt{-1}\).
Note that in practice we just have access to the mode shapes matrix \( \Psi_{1:n} \) and not to the eigenvectors matrix \( \Phi_{1:n} \), so in order to fully specify the state equation we suppose that the covariance matrix \( Q_k \) is of the form \([L \ 0]|Q_i|L \ 0\) for a given covariance matrix \( Q_i \). Hence \( w_k \equiv \Phi^* \zeta_k \) is a white Gaussian noise with covariance matrix \( Q_k^* = \delta \Psi^* Q_i \Psi \).

The observation equation (6) becomes
\[
y_k = \Psi x_k + \Psi \bar{x}_k + \nu_k, \quad \nu_k \overset{iid}{\sim} N(0, Q_i^k) .
\]

Note that \( \Psi x + \Psi \bar{x} = 2 \Re(\Psi x) \) is a linear operator.

The state/space system

One finally obtains the following system
\[
x_{k+1} = \Lambda x_k + \sigma w_k , \quad w_k \overset{iid}{\sim} N(0, Q_i^k), \quad (8)
y_k = 2 \Re(\Psi x_k) + \nu_k, \quad \nu_k \overset{iid}{\sim} N(0, Q_i^k) . \quad (9)
\]

In this model all parameters are assumed known, or previously estimated, except the eigenvalues matrix \( \Lambda \equiv \text{diag}(\lambda_{1:n}) \) and the noise intensities \( \sigma \) and \( \nu \). The mode shapes matrix \( \Psi = [\Psi_{1:n}] \), the sampling period \( \delta \), and the covariance matrices \( Q_k \) and \( Q_i^k \) are given (then \( Q_i^k = \delta \Psi^* Q_i \Psi \)). From now on we suppose that \( Q_i^k = I \).

3 Parameterization

Consider the following linear system
\[
x_{k+1} = F(\theta) x_k + G(\theta) w_k , \quad w_k \overset{iid}{\sim} N(0, Q_i^k), 
y_k = H(\theta) x_k + \Sigma(\theta) \nu_k , \quad \nu_k \overset{iid}{\sim} N(0, Q_i^k) .
\]

The state process \( x_k \) takes values in \( \mathbb{C}^n \), the observation process \( y_k \) in \( \mathbb{C}^d \), and \( \theta \in \mathbb{R}^p \). The state initial law is \( x_0 \sim N(x_0, R_0) \). Initial condition \( x_0 \), state noise \( w_k \) and observation noise \( \nu_k \) are mutually independent.

Suppose that the matrices \( F(\theta) \in \mathbb{C}^{n \times n} \), \( G(\theta) \in \mathbb{C}^{n \times d} \), \( H(\theta) \in \mathbb{C}^{d \times n} \) and \( \Sigma(\theta) \in \mathbb{C}^{d \times d} \) are differentiable w.r.t. \( \theta \). We also suppose that \( G(\theta) G(\theta)^* \geq M_1 > 0 \) and \( \Sigma(\theta) \Sigma(\theta)^* \geq M_2 > 0 \) where \( M_1 \) and \( M_2 \) are symmetric positive definite matrices.

Here we suppose that the unknown parameter is the realization of a random variable denoted \( \theta \) (the realizations of this variable are denoted \( \tilde{\theta} \)). Let \( \rho(d\theta) \) denotes the a priori law of this random, for example \( \rho = N(\tilde{\theta}_0, T_0) \).

This model is equivalent to the following extended state-space system:
\[
\theta_{k+1} = \theta_k , \quad \theta_0 \sim N(\tilde{\theta}_0, T_0) , \quad (10)
x_{k+1} = F(\theta_k) x_k + G(\theta_k) w_k , \quad x_0 \sim N(x_0, R_0) , \quad (11)
y_k = H(\theta_k) x_k + \Sigma(\theta_k) \nu_k . \quad (12)
\]
4 Convolution particle filter

The second approach takes place in a classical Bayesian framework: θ is the unknown realization of a random variable θ with prescribed a priori probability law \( N(\hat{\theta}_0, T_0) \). Let:

\[
M^\theta_k(x, dx') \equiv \mathbb{P}(x_k \in dx'| \theta = \theta, x_{k-1} = x),
\]
\[
\Psi^\theta_k(x, dy) \equiv \mathbb{P}(y_k \in dy| \theta = \theta, x_k = x).
\]

An extended state variable \((x_k, \theta_k)\) joining all the unknown quantities is considered and the posterior law \( \mathbb{P}(x_k \in dx, \theta_k \in d\theta|y_{1:k}) \) is then approximated using a convolution particle filter.

**Kernel estimation**

A kernel \( K : \mathbb{R}^n \rightarrow \mathbb{R} \) is a bounded, positive, symmetric application such that \( \int K(x) \, dx = 1 \). We denote

\[
K_{h_0}(x) \equiv \frac{1}{(2\pi)^{n/2}} e^{-|x|^2/2}.
\]

Let \( x^1 \cdots x^N \) be i.i.d. random variables with common density \( f \). The kernel estimator \( f_N \) of \( f \) associated with the kernel \( K \) is given by

\[
f_N(x) = (K_{h_N} * \mu_N)(x) \equiv \frac{1}{N} \sum_{i=1}^{N} K(\frac{x-x^i}{h_N})
\]

for \( x \in \mathbb{R}^n; \mu_N = \frac{1}{N} \sum_{i=1}^{N} \delta_{x^i} \) is the empirical measure associated with \( x^1 \cdots x^N \).

**The convolution filter**

Let \( \{x^i_0\}_{i=1}^{i=N} \) be a sample of size \( N \) of \( N(\bar{x}_0, \mathcal{R}_0) \), and \( \{\theta^i_0\}_{i=1}^{i=N} \) be a sample of size \( N \) of \( N(\bar{\theta}_0, T_0) \).

We describe now the iteration \( k-1 \rightarrow k \) of the algorithm. Starting from \( \{x^i_{k-1}\}_{i=1}^{i=N} \), we can simulate the following samples

\[
x^i_k - \sim M_k(x^i_{k-1}, \cdot), \quad \theta^i_k - \sim \theta^i_{k-1}, \quad y^i_k - \sim \Psi_k(x^i_k, \cdot)
\]

for \( i = 1 \cdots N \). We deduce the following empirical estimate of the joint conditional law of \((x_k, y_k)\) given \( y_{1:k-1} \)

\[
\mathbb{P}(x_k \in dx, \theta \in d\theta, y_k \in dy|y_{1:k-1}) \approx
\]
\[
\frac{1}{N} \sum_{i=1}^{N} \left\{ K^\theta_{h_N}(x - x^i_k) K^\theta_{h_N}(\theta - \theta^i_k) K^y_{h_N}(y - y^i_k) \right\} \, dx \, d\theta \, dy
\]

(13)

where \( K^\theta_{h_N}, K^\theta_{h_N} \) and \( K^y_{h_N} \) are suitably chosen kernels of appropriate dimensions. Here we propose to use Parzen-Rosenblatt kernels, see Appendix 4. Note that in \( K^\theta_{h_N} \) (resp. \( K^\theta_{h_N} \)) \( h_N \) could...
implicitly depend on $N$, $d$ and $x_{k+1}^j$ (resp. $N$, $q$ and $y_{k+1}^j$) (resp. $N$, $p$ and $\theta_{k+1}^j$). From (13) we deduce the following convolution approximation of the optimal filter:

$$\pi_k(dx, d\theta) = \pi_k^0(dx, d\theta) = \frac{1}{N} \sum_{i=1}^N K_{k,0}^i(x - x_k) K_{k,0}^i(\theta - \theta_k) K_{k,0}^i(y_k - y_k^i) \ dx \ d\theta.$$  \label{eq:approx}

Then, we define

$$(x_k^i, \theta_k^i) \sim \pi_k^0(dx, d\theta) \quad i = 1 \cdots N.$$  \label{eq:approx}

**A posteriori law analysis**

There are two possibilities: we can define the support of the particles $(\theta_k^i)_{i=1:N}$

$$\theta_{k,\ell}^{\text{min}} \overset{\text{def}}{=} \min_{i=1:N} \theta_k^i, \quad \theta_{k,\ell}^{\text{max}} \overset{\text{def}}{=} \max_{i=1:N} \theta_k^i.$$  \label{eq:supp}

Then, at time $k$, for each component $\ell = 1 \cdots p$, we have the confidence interval $[\theta_{k,\ell}^{\text{min}}, \theta_{k,\ell}^{\text{max}}]$.

Another possibility is to make a Gaussian approximation:

$$\bar{\theta}_{k,\ell} \overset{\text{def}}{=} \frac{1}{N} \sum_{i=1}^N \theta_k^i, \quad \sigma_{k,\ell}^2 \overset{\text{def}}{=} \frac{1}{N-1} \sum_{i=1}^N (\theta_k^i - \bar{\theta}_{k,\ell})^2$$

Then a confidence interval could be $[\bar{\theta}_{k,\ell} - 2\sigma_{k,\ell}, \bar{\theta}_{k,\ell} + 2\sigma_{k,\ell}]$.

**Choice of the parameters** $h_{x,N}^i, h_{\theta,N}^i, h_{y,N}^i$

We choose: $h_{x,N}^i = C_x \times N^{-1/(4+q)}$, $h_{\theta,N}^i = C_\theta \times N^{-1/(4+p)}$, $h_{y,N}^i = C_y \times N^{-1/(4+q)}$ and

$$C_x = c_x \times [\text{Cov}(x_k^1, \ldots, x_k^N)]^{1/2},$$
$$C_\theta = c_\theta \times [\text{Cov}(\theta_k^1, \ldots, \theta_k^N)]^{1/2},$$
$$C_y = c_y \times [\text{Cov}(y_k^1, \ldots, y_k^N)]^{1/2}$$

with $c_x, c_\theta, c_y \approx 1$ gives good results. For the simulations in the last section, on taking a Gaussian kernel, we will see that the $c$'s are easily adjusted.

Simulations based on flutter evolution will be shown later.

## 5 References


filter initialization: $x_0^i \sim N(\bar{x}_0, R_0)$ and $\theta_0^i \sim N(\bar{\theta}_0, T_0)$ for $i = 1 : N$

for $k = 1, 2, \ldots$ do

prediction

$x_k^i \sim M_{x_k}^{i-1}(x_{k-1}^i, dx)$ for $i = 1 : N$ {state sampling}

$y_k^i \sim \Psi_{x_k}^{i-1}(x_{k-1}^i, dy)$ for $i = 1 : N$ {observation sampling}

$\theta_k^i = \theta_{k-1}^i$ for $i = 1 : N$ {parameter sampling}

compute the parameters $h^y_N$, $h^x_N$, $h^\theta_N$

filter updating:

$$\pi^N_k(dx, d\theta) = \frac{\sum_{i=1}^N K^y_{h^y_N}(y_k - y^i_k) \cdot K^x_{h^x_N}(x - x^i_k) \cdot K^\theta_{h^\theta_N}(\theta - \theta^i_k)}{\sum_{i=1}^N K^y_{h^y_N}(y_k - y^i_k)} \cdot dx \cdot d\theta$$

$(x_k^i, \theta_k^i) \sim \pi^N_k(dx, d\theta)$ {resampling}

end for

Table 1: The convolution filter for Bayesian estimation.


