Exploring fMRI data for periodic signal components

Lars Kai Hansen*, Finn Årup Nielsen, Jan Larsen
Informatics and Mathematical Modelling, Technical University of Denmark B321, DK-2800 Lyngby, Denmark

Abstract

We use a Bayesian framework to detect periodic components in fMRI data. The resulting detector is sensitive to periodic components with a flexible number of harmonics and with arbitrary amplitude and phases of the harmonics. It is possible to detect the correct number of harmonics in periodic signals even if the fundamental frequency is beyond the Nyquist frequency. We apply the signal detector to locate regions that are highly affected by periodic physiological artifacts, such as cardiac pulsation.

Keywords: Exploratory data analysis; fMRI; Periodic signals; Bayesian framework

1. Introduction

In fMRI experiments we may want to local detect periodic components in the local hemodynamic activity. Much attention has been devoted to estimate the confounding signals generated by pulsatory activity at the cardiac frequency. The approaches range from simple digital filters [1] to sophisticated adaptive techniques based on the complex $k$-space MR signal [6], both methods are based on external monitoring of the cardiac activity. Dagli et al. concluded that regions near vessels have reduced sensitivity for detection of activation because of the signal variance induced by cardiac pulsation [2]. The typical acquisition frequencies (TR $\sim$3–4 s) in whole brain data sets preclude direct spectral filtering. Even for rapid acquisition single slice data sets the cardiac signal will have higher harmonics at frequencies beyond the Nyquist limit.

The quest for periodic components can be formulated as a local test for the presence of a periodic signal with flexible number of harmonics and basic frequency against a null-hypothesis under which the signal is white noise. We will review a Bayesian framework below that allows calculation of relative probabilities of such competing hypotheses; this framework was first applied in the context of fMRI data analysis in [5].

* Corresponding author. Tel.: +45-4525-3889; fax: +45-4587-2599.
E-mail address: lkh@imm.dtu.dk (L.K. Hansen).

0933-3657/02/$ – see front matter © 2002 Published by Elsevier Science B.V.
PIL: S0933-3657(02)00007-6
The Bayesian framework is of interest in this context because it gives a more complete picture of the interplay between the null-hypothesis and relevant alternatives and the framework has an embedded quantitative statement of the a priori knowledge that enters the formulation of hypotheses. Our approach is a Bayesian generalization of the so-called general linear model used frequently in fMRI analysis, see e.g. [4]; the main generalization is that we can eliminate the unknown amplitudes and phases of the harmonics as well as the noise variance, which in turn allows us to operate the general linear model as an explorative tool.

Frank et al. [3] recently reviewed a Bayesian framework for signal detection in fMRI data. Here, we expand on the application of the Bayesian framework based on so-called conjugate priors in the context of periodic component detection.

2. Bayes’ theory

We will be focus on models of the local hemodynamic activity in a region or as in this exposition, a single pixel. Let \( y \) be a fMRI signal measured at times \( t = 1, \ldots, T \), and represented as a \( T \times 1 \) vector with components \( y(t) \). The signal is modeled as a sum of harmonic components of the form,

\[
\hat{y}(t) = \sum_{k=1}^{K} x_k(t)b_k ,
\]

where \( x_k(t) \) is a set of periodic “basis functions” and \( b \) a set of \( K \) linear coefficients quantifying the content of signal of the given periodicity in the local hemodynamic activity. Introducing the \( T \times K \) matrix with components \( X(t,k) \equiv x_k(t) \) the linear model can be written in matrix form

\[
\hat{y} = Xb
\]

In an fMRI experiment, we expect that the actual measurement deviates from the “ideal” model output by various noise contributions that we will represent by a random white noise process so that \( y(t) = \hat{y}(t) + n(t) \), where \( n(t) \) is assumed zero mean normal with unknown variance \( (\sigma^2) \).

The basis functions will be chosen as \( x_{2k-1}(t) = \sin(kw_0t), x_{2k}(t) = \cos(kw_0t), k = 1, \ldots, K \), where \( w_0 \) is the fundamental frequency of the signal. By including a linear combination of both sines and cosines, we can generate harmonics with arbitrary phase relations. Note that a model with \( \kappa \) frequencies has \( K = 2\kappa \) basis functions, hence a \( 2\kappa \) dimensional coefficient vector \( b \).

A fundamental problem posed by Eq. (1) and the definitions above is that the model order \( K \), the fundamental frequency \( w_0 \), and the noise variance \( \sigma^2 \) are unknown. Here, we will develop a Bayesian scheme that will allow us to make inferences about these parameters, independent of the amplitude and phase of the harmonics and independent of the noise variance. This is achieved by invoking a simple, yet flexible prior distribution of these parameters so that we may eliminate these parameters by explicit integration. In particular, we will aim at estimating the probability \( P(w_0, K | y) \) of the “hypothesis”
specified by $w_0, K$, using Bayes’ theorem,

$$P(w_0, K|y) = \frac{P(y|w_0, K)P(w_0, K)}{P(y)}, \quad (3)$$

where $P(y|w_0, K)$ is the likelihood function, $P(w_0, K)$ the prior probability, whereas $P(y)$ is a normalization constant.

For a fixed set of parameters $b, \sigma^2$ we can use Eq. (2) to establish the likelihood function, i.e. the probability density of the observations given the parameters,

$$P(y|\sigma^2, b, X, K) = \left(\frac{1}{2\pi \sigma^2}\right)^{T/2} \exp\left(-\frac{1}{2\sigma^2}(y - Xb)^2\right). \quad (4)$$

Since, however, $b$ and $\sigma^2$ are unknown too we need to eliminate them using a prior distribution $P(b, \sigma^2)$ which quantifies the general knowledge we have on the domain and which potentially depends on the given basis set and model order,

$$P(y|w_0, K) = \int d\sigma^2 \int db P(b, \sigma^2)P(y|\sigma^2, b, X, K)$$

$$= \int d\sigma^2 \int db P(b, \sigma^2)\left(\frac{1}{2\pi \sigma^2}\right)^{T/2} \exp\left(-\frac{(y - Xb)^2}{2\sigma^2}\right). \quad (5)$$

We will use the principle of **conjugate priors** to establish a convenient prior $P(b, \sigma^2)$. The subsequent analysis is described comprehensively in most textbooks on Bayesian data analysis, see e.g. [7] for a particularly clear treatment. The idea is to employ a prior density so that the posterior (proportional to the product of the prior and the likelihood, c.f. Eq. (3)) is of the same form as the prior but with “updated”, i.e. data dependent parameters. The conjugate prior for the earlier linear model with additive Gaussian noise is the so-called normal–inverse–gamma or $\text{NIG}(a, d, m, V)$, distribution,

$$P(b, \sigma^2|a, d, K, m, V) = \frac{(a/2)^{d/2}(\sigma^2)^{-(d+K+2)/2}}{(2\pi)^{K/2}|V|^{1/2}\Gamma(d/2)} \times \exp\left(-(b - m)'(2\sigma^2 V)^{-1}(b - m) - \frac{a}{2\sigma^2}\right). \quad (6)$$

The new (hyper-) parameters $d, a, m, V$ have the following meaning.

$$P(b|a, d, K, m, V) = \int d\sigma^2 P(b, \sigma^2|m, K) = \frac{(a/2)^{-K/2}\Gamma((d + K)/2)}{(2\pi)^{K/2}|V|^{1/2}\Gamma(d/2)} \times (1 + (b - m)'(aV)^{-1}(b - m))^{-(d+K+2)/2} \quad (7)$$

The marginal prior distribution of $b$, is a multivariate $t$-distribution with mean $m$ and covariance determined by $(a/(d - 2)V$. This distribution is unimodally centered at $m$, with heavier “tails” than a normal distribution, see Fig. 1. The marginal prior distribution of $\sigma^2$ is given by

$$P(\sigma^2|a, d) = \frac{(a/2)^{-d/2}(\sigma^2)^{-(d+2)/2}}{\Gamma(d/2)} \exp\left(-\frac{a}{2\sigma^2}\right). \quad (8)$$
Hence, an inverse gamma distribution\(^1\) of mean \(\frac{a}{d - 2}, d > 2\).

The next step of the inference is then to set the parameters of the prior. In general, we prefer to give the parameters values so that they have minimal influence on results. In particular, we should check that for long time series their effects should vanish completely.

The prior mean of the noise variance can, e.g. be set to the observed signal variance, \(a/(d - 2) = \sigma^2 \equiv y'y/T\), meaning that we do not expect a noise variance larger than the total observed variance. Further, we will let \(d = 3\) leading to a prior as shown in Fig. 1, this choice of \(d\) is the smallest integer for which the prior noise variance is finite, hence, a “weak” prior. We will not express any prior knowledge about the mean amplitude of the periodic components, hence, \(m = 0\). The form of the prior covariance structure is chosen for simplicity to be \(V = vI\), where \(I\) is a unit matrix. The parameter \(v\) will be determined essentially by data by the following argument. The prior variance of the

---

\(^1\) That is, \(1/\sigma^2\) is gamma distributed
fitted signal $\hat{y}$, is given by

$$
\frac{\langle \hat{y}' \rangle_{\text{prior}}}{T} = Tr \left[ \frac{XX' \langle bb' \rangle_{\text{prior}}}{T} \right] = \frac{va}{d-2} Tr \left[ XX' \right] / T
$$

(9)

As earlier, for the noise variance hyperparameter selection, we can let variance be equal to the variance of the measured signal, i.e. let $v = 1 / (Tr[XX'] / T)$.

Comparing Eqs. (4) and 6 we see that by conjugacy they are of the same exponential form, so when we multiply them together, the integrand in Eq. (5) is again an NIG distribution, hence the integral is simply the NIG normalization integral, thus we find

$$
P(y|w_0, K) = \left( \frac{|V_p| d^d}{|V| (a_p)^d \pi T} \right)^{1/2} \frac{\Gamma(d_p/2)}{\Gamma(d/2)},
$$

(10)

Fig. 2. Simulation experiment created to illustrate the ability of the Bayesian approach for picking up the correct model order. The upper panel shows the true periodic signal (heavy dashed line) composed by the fundamental and the second harmonic (of relative amplitude 0.6). The noisy signal is obtained by adding Gaussian noise of the same standard deviation as the "signal". The middle panel shows the true signal (heavy dashed line) and the maximum a posteriori reconstructed signal ($K = 2$). The lower panel shows the probability of having $k = 1, \ldots, 10$ harmonics. As expected, this probability is strongly peaked at the true value $k = 2$. 
with the following definitions

\[ V_P^{-1} = V^{-1} + X'X, \]  
\[ m_P = V_P(V^{-1} m + X'y), \]  
\[ a_P = a + m'V^{-1}m + y'y' - m'PvP^{-1}mP, \]  
\[ d_P = d + T. \]

(11)  
(12)  
(13)  
(14)

Using our specifications of the prior parameters, we obtain the following simplifications:

\[ V_P^{-1} = v1 + X'X, \]  
\[ m_P = V_PX'y, \]

(15)  
(16)  

Fig. 3. Simulation experiment created to show the ability of the Bayesian approach for picking up the correct model order even when the fundamental frequency is undersampled. The upper panel shows the true periodic signal (heavy dashed line) composed by the fundamental and the second harmonic (of relative magnitude 0.6). The angular frequency is \( \omega_{true} = 6.4 \), this signal is aliased to appear low frequency. The noisy signal is obtained by adding noise of the same standard deviation as the signal (\( K = 2 \)). The middle panel shows the true signal (heavy dashed line) and the maximum a posteriori reconstructed signal (\( K = 2 \)). The lower panel shows the probability of having \( k = 1, \ldots, 10 \) harmonics. As expected, this probability is strongly peaked at the true value \( k = 2 \).
\[ a_P = (T + 1) \sigma_y^2 - y'XV_PX'y, \]  
\[ d_P = 3 + T. \]  

We can see explicitly that the influence of the prior choice of \( a \) and \( d \) is weak for \( T \gg 1 \), because the prior contributions are of order 1 relative to \( T \) in Eqs. (17) and (18), respectively.

Testing the earlier linear system hypotheses, a natural null-hypothesis is that the signal is Gaussian noise of unknown variance. The corresponding probability density \( P(y|0) \) is given by the \( X = 0 \) limit of the earlier expressions.

Fig. 4. A single slice holding \( 128 \times 128 \) pixels and cutting through primary visual cortex was acquired with a time interval between successive scans of \( TR = 333 \) ms. A window of \( M = 82 \times 68 \) pixel covering all of the brain of the particular slice was extracted for this analysis. The sampling frequency of this acquisition mode is high enough to allow faithful representation of the heart signal. Visual stimulation in the form of a flashing checkerboard pattern was interleaved with periods of fixation. A run consisting of 30 scans of fixation, 30 scans of stimulation, and 60 scans of post-stimulus fixation was repeated 10 times, here we analyze a single run. The figure shows the distribution of pixel-wise most probable model order \( (\kappa) \). Black pixels are assigned to the null-hypothesis, while the brighter pixels have up to four harmonics. Two locations are singled out for further analysis. The \"vessel\" pixel and the \"primary visual\" pixel. The most probable frequency is high corresponding to cardiac pulsation in the former and low, corresponding to the presentation frequency of the stimulus in the latter region.
The probabilities of the set of complete set of hypotheses (parameterized by \( w_0 \) and \( K \)) including the null-hypothesis are then given by

\[
P(w_0, K | y) = \frac{P(y | w_0, K)}{P(y | 0) + \sum_{w_0, K} P(y | w_0, K)}
\]

\[
P(0 | y) = \frac{P(y | 0)}{P(y | 0) + \sum_{w_0, K} P(y | w_0, K)}
\]

(19)

3. Evaluation on simulated and real fMRI data

3.1. Simulation experiments

In order to illustrate the viability of the Bayesian approach for exploring time series for periodic components we have set up a simulation experiment. A signal was created by mixing a periodic signal (two harmonics) and Gaussian white noise signal of the same standard deviation. In Fig. 2, we show the true and the noisy signals, in the case the

Fig. 5. Analysis of the “primary visual” pixel of Fig. 4. This region has a strong positive activation trace and the optimal fundamental frequency is \( \sim 0.03 \) Hz. The measured signal is represented by the thin line, the MAP reconstructed signal is rendered with the heavy line.
fundamental frequency is lower than the Nyquist frequency. The middle panel shows the
signal reconstructed from the maximum a posteriori parameters ($m_P$). In the lower panel of
Fig. 2 we show the Bayesian probabilities for hypotheses with $k = 1, \ldots, 10$ frequencies,
which is strongly focused at the true value of $k = 2$, even in the presence of sizable noise
contamination. In Fig. 2, we show a similar setup as in Fig. 2, except now the fundamental
frequency is beyond the Nyquist. In this case, the periodic signal is aliased to appear as a
low frequency signal, but importantly, we may still detect its presence and can still
correctly detect that it consists of two harmonics.

3.2. fMRI experiment

A single fMRI slice holding $128 \times 128$ pixels and cutting through primary visual cortex
was acquired with a time interval between successive scans of $TR = 333$ ms. A window of
$M = 82 \times 68$ pixel covering all of the brain of the particular slice was extracted for this
analysis. This sampling frequency is high enough to allow faithful representation of the
heart signal. Visual stimulation in the form of a flashing annular checkerboard pattern was
interleaved with periods of fixation. A run consisting of 30 scans of fixation, 31 scans of
stimulation, and 60 scans of post-stimulus fixation was repeated 10 times. The data set was
acquired by Dr. Egill Rostrup at the Danish Center for Magnetic Resonance Research.

![MEASURED AND MAP RECONSTRUCTED SIGNALS](image)

![PROBABILITY](image)

Fig. 6. Analysis of the “vessel” pixel of Fig. 4. This region consists mainly of a high-frequency cardiac
pulsation signal, well described by the periodic signal model with four harmonics.
In each voxel we test the hypotheses: The white noise null-hypothesis, and a set of hypotheses parameterized by fundamental frequencies and number of components. Fig. 4 shows the distribution of the pixel-wise most probable number of frequencies ($k$).

Black pixels are assigned to the null-hypothesis, whereas the brightest pixels have four harmonics. Two locations are singled out for further analysis. The “vessel” pixel has four harmonics and is dominated by a high-frequency component corresponding to the cardiac frequency. The “primary visual” pixel has only one low-frequency component. In Fig. 5, we show the analysis of the primary visual area hemodynamic activity. This region has a strong positive stimulus response corresponding to a very low-frequency component. In Fig. 6, we show the analysis of the vessel pixel. This region is dominated by cardiac pulsation and is well approximated by the reconstructed MAP high-frequency signal.

4. Conclusion

We have outlined a Bayesian framework for signal detection in noisy linear systems. We used weak conjugate priors and as a result we obtain closed form expressions for the relative probabilities over competing hypotheses. Formulating a model based on periodic signals with a variable number of harmonics, we obtain a periodic component detector, this detector can detect periodic signals with arbitrary amplitude and phase relations between the harmonics, and was shown able to detect the correct number of harmonics even if the fundamental frequency is beyond the Nyquist frequency.

Acknowledgements

We thank Dr. Egill Rostrup for access to the fMRI data set used for illustration in this paper. Our work is funded by the Danish Research Councils through the THOR Center for Neuroinformatics, the EU Commission through project MAPAWAMO, and by the US National Institutes of Health Human Brain Project grant P20 MH57180 “Spatial and Temporal Patterns in Functional Neuroimaging”.

References