Detecting and estimating the shape of a periodic component in short duration signals

Ali Mohammad-Djafari

Laboratoire des Signaux et Systèmes (L2S)

UMR8506 CNRS-CentraleSupélec-UNIV PARIS SUD

SUPELEC, 91192 Gif-sur-Yvette, France

http://lss.centralesupelec.fr

Email: djafari@lss.supelec.fr

http://djafari.free.fr

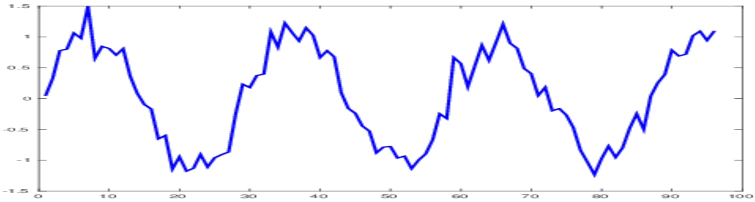
http://publicationslist.org/djafari

MaxEnt 2016 workshop, July 10-15, 2016, Gent, Belgium.

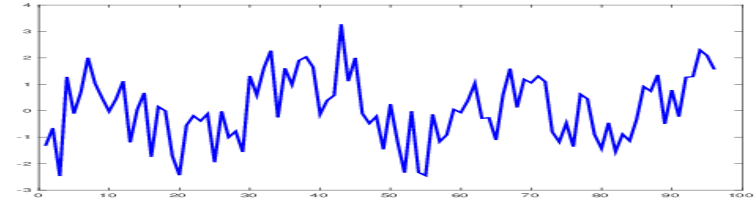
Contents

Description of the problem

- Detecting a periodic component in a short duration signal, estimating its period and its shape.
- Low noise example

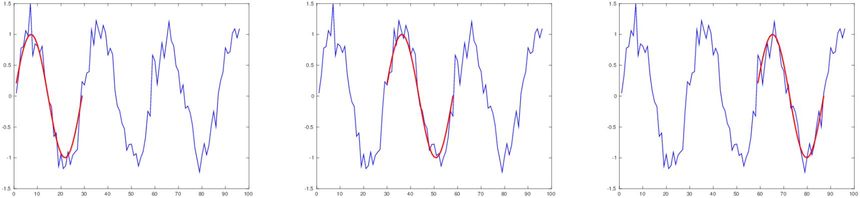


- High noise example

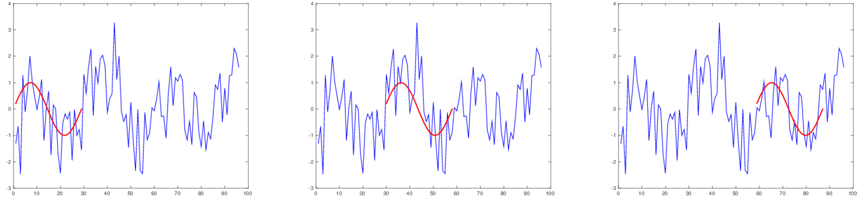


Description of the problem

- Detecting a periodic component in a short duration signal, estimating its period and its shape.
- Low noise example



- High noise example



Description of the problem

-

$$\begin{aligned} \mathbf{g} &= [g_1, \cdots, g_N, g_{N+1}, \cdots, \cdots, ...g_{KN+1}, \cdots, g_M]' \\ &= [f_1, \cdots, f_N, f_1 \quad, \cdots, f_N, \cdots, f_1, \cdots\cdots, f_r]' \end{aligned}$$

  where $M = KN + r$ where K is the number of complete repetition of the periodic shape and r is the rest.

- This relation can be written as a linear relation:

$$\mathbf{g} = \mathbf{H}_N \mathbf{f}$$

  where $\mathbf{H}$ has the following structure

$$\mathbf{H}_N = [\mathbf{I}_N | \mathbf{I}_N | ... | \mathbf{I}_N | \mathbf{I}(:, 1:r)]$$

  where $\mathbf{I}_N$ is the unitary matrix of size $N \times N$ and $\mathbf{I}(:, 1:r)$ is its first r columns.

$$\mathbf{g} = \mathbf{H}_N \mathbf{f} + \epsilon$$

  the vector ϵ represents those errors.

Description of the problem

- Also, we require some regularity in the shape of $\mathbf{f}$. This regularity can be modelled as

$$\mathbf{f} = \mathbf{D}\mathbf{f} + \xi \rightarrow (\mathbf{I} - \mathbf{D})\mathbf{f} = \xi \rightarrow \mathbf{C}\mathbf{f} = \xi$$

where $\mathbf{C}$ can be of the form

$$\mathbf{C}_N = \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 \\ 0 & 1 & -1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & -1 \\ -1 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

- A criterion which measures the regularity can be $\|\mathbf{C}\mathbf{f}\|^2$

Deterministic regularization method

$$\mathbf{g} = \mathbf{H}_N \mathbf{f} + \epsilon, \quad \mathbf{C}\mathbf{f} = \xi$$

With these two equations, we have at least two possibilities:

- Deterministic regularization:

$$\widehat{\mathbf{f}} = \arg\min_{\mathbf{f}} \{J(\mathbf{f})\} \text{ with } J(\mathbf{f}) = \|\mathbf{g} = \mathbf{H}_N \mathbf{f}\|_2^2 + \lambda \|\mathbf{C}_N \mathbf{f}\|_2^2$$

  The solution is given by:

$$\widehat{\mathbf{f}} = [\mathbf{H}_N' \mathbf{H}_N + \lambda \mathbf{C}_N' \mathbf{C}_N]^{-1} \mathbf{H}_N' \mathbf{g}$$

- The above criterion was given for a given value of N. We can now try to define a criterion which depends explicitly on N :

$$J(N, \mathbf{f}) = \|\mathbf{g} = \mathbf{H}_N \mathbf{f}\|_2^2 + \lambda \|\mathbf{C}_N \mathbf{f}\|_2^2$$

  and try to optimize it to find both the seeked period N and the shape $\mathbf{f}$:

$$(\widehat{N}, \widehat{\mathbf{f}}) = \arg\min_{(N, \mathbf{f})} \{J(N, \mathbf{f})\}$$

Bayesian approach

$$\mathbf{g} = \mathbf{H}_N \mathbf{f} + \epsilon, \quad \mathbf{C}\mathbf{f} = \boldsymbol{\xi}$$

- Likelihood

$$p(\mathbf{g}|\mathbf{f}) = \mathcal{N}(\mathbf{g}|\mathbf{H}_N \mathbf{f}, v_\epsilon \mathbf{I}),$$

- Prior

$$p(\mathbf{f}) = \mathcal{N}(\mathbf{f}|0, v_\xi (\mathbf{C}'_N \mathbf{C}_N)^{-1})$$

- Poserior

$$p(\mathbf{f}|\mathbf{g}) \propto p(\mathbf{g}|\mathbf{f}) p(\mathbf{f}) = \mathcal{N}(\mathbf{f}|\widehat{\mathbf{f}}, \widehat{\boldsymbol{\Sigma}})$$

with

$$\widehat{\mathbf{f}} = \arg\max_{\mathbf{f}} \{p(\mathbf{f}|\mathbf{g})\} = \arg\min_{\mathbf{f}} \{J(\mathbf{f})\}$$

which is equivalent to the quadratic regularization as before

$$\widehat{\mathbf{f}} = [\mathbf{H}'_N \mathbf{H}_N + \lambda \mathbf{C}'_N \mathbf{C}_N]^{-1} \mathbf{H}'_N \mathbf{g}$$

with $\lambda = v_\epsilon / v_\xi$.

Bayesian approach

- Poserior
$$p(\mathbf{f}|\mathbf{g}) \propto p(\mathbf{g}|\mathbf{f})p(\mathbf{f}) = \mathcal{N}(\mathbf{f}|\widehat{\mathbf{f}}, \widehat{\boldsymbol{\Sigma}})$$
with
$$\widehat{\mathbf{f}} = \arg\max_{\mathbf{f}} \{p(\mathbf{f}|\mathbf{g})\} = \arg\min_{\mathbf{f}} \{J(\mathbf{f})\}$$
and
$$\widehat{\boldsymbol{\Sigma}} = [\mathbf{H}_N'\mathbf{H}_N + \lambda \mathbf{C}_N'\mathbf{C}_N]^{-1}$$
which can be used to put error bars on the solution.
- Joint posterior $p(N, \mathbf{f}|\mathbf{g})$ and and try to optimize it to find the seeked solution
$$(\widehat{N}, \widehat{\mathbf{f}}) = \arg\max_{(N,\mathbf{f})} \{p(N, \mathbf{f}|\mathbf{g})\}$$
- We can do better.

Bayesian approach with more appropriate priors

- Forward and prior model equations:

$$\mathbf{g} = \mathbf{H}_N \mathbf{f} + \boldsymbol{\epsilon}, \quad \mathbf{C}\mathbf{f} = \boldsymbol{\xi}$$

- ϵ_i and ξ_j are Gaussian but with unknown variances that we want to estimate to.

$$p(\epsilon_i|v_{\epsilon_i}) = \mathcal{N}(\epsilon_i|0, v_{\epsilon_i}), \quad p(v_{\epsilon_i}|\alpha_{\epsilon_0}, \beta_{\epsilon_0}) = \mathcal{IG}(v_{\epsilon_i}|\alpha_{\epsilon_0}, \beta_{\epsilon_0})$$

$$p(\xi_i|v_{\xi_i}) = \mathcal{N}(\xi_i|0, v_{\xi_i}), \quad p(v_{\xi_i}|\alpha_{\xi_0}, \beta_{\xi_0}) = \mathcal{IG}(v_{\epsilon_i}|\alpha_{\xi_0}, \beta_{\xi_0})$$

- This can also be interpreted as a wish to model them by a heavier tailed probability laws such as Student-t:

$$\mathcal{S}t(\epsilon_i|\alpha_{\epsilon_i}, \beta_{\epsilon_i}) = \int_0^\infty \mathcal{N}(\epsilon_i|, 0, v_{\epsilon_i})\, \mathcal{IG}(v_{\epsilon_i}|\alpha_{\epsilon_0}, \beta_{\epsilon_0})\, \mathrm{d}v_{\epsilon_i}$$

$$\mathcal{S}t(\xi_j|\alpha, \beta) = \int_0^\infty \mathcal{N}(\xi_j|, 0, v_{\xi_j})\, \mathcal{IG}(v_{\xi_j}|\alpha_{\xi_0}, \beta_{\xi_0})\, \mathrm{d}v_{\xi_j}$$

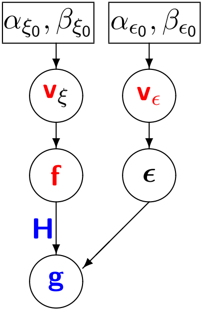
Non stationary noise and sparsity enforcing model

– Forward model:

$$\mathbf{g} = \mathbf{H}\mathbf{f} + \boldsymbol{\epsilon},\ \epsilon_i \sim \mathcal{N}(\epsilon_i|0, v_{\epsilon_i}) \rightarrow \boldsymbol{\epsilon} \sim \mathcal{N}(\boldsymbol{\epsilon}|0, \mathbf{V}_\epsilon), \mathbf{V}_\epsilon = \text{diag}\,[v_{\epsilon 1}, \cdots, v_{\epsilon M}])$$

– Prior model:

$$\mathbf{C}_N\mathbf{f} = \boldsymbol{\xi},\ \xi_j \sim \mathcal{N}(\xi_i|0, \xi_j) \rightarrow \boldsymbol{\xi} \sim \mathcal{N}(\boldsymbol{\xi}|0, \mathbf{V}_\xi), \mathbf{V}_\xi = \text{diag}\left[v_{\xi_1}, \cdots, v_{\xi_N}\right])$$



$$\begin{cases} p(\mathbf{g}|\mathbf{f}, \mathbf{v}_\epsilon) = \mathcal{N}(\mathbf{g}|\mathbf{H}\mathbf{f}, \mathbf{V}_\epsilon), & \mathbf{V}_\epsilon = \text{diag}\,[\mathbf{v}_\epsilon] \\ p(\mathbf{f}|\mathbf{v}_f) = \mathcal{N}(\mathbf{f}|0, \mathbf{V}_\xi\mathbf{C}\mathbf{C}'), & \mathbf{V}_\xi = \text{diag}\,[\mathbf{v}_\xi] \end{cases}$$

$$\begin{cases} p(\mathbf{v}_\epsilon) = \prod_i \mathcal{IG}(v_{\epsilon_i}|\alpha_{\epsilon_0}, \beta_{\epsilon_0}) \\ p(\mathbf{v}_\xi) = \prod_i \mathcal{IG}(v_{\xi_j}|\alpha_{\xi_0}, \beta_{\xi_0}) \end{cases}$$

$$p(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_\xi|\mathbf{g}) \propto p(\mathbf{g}|\mathbf{f}, \mathbf{v}_\epsilon)\, p(\mathbf{f}|\mathbf{v}_\xi)\, p(\mathbf{v}_\epsilon)\, p(\mathbf{v}_\xi)$$

Objective: Infer $(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_\xi)$

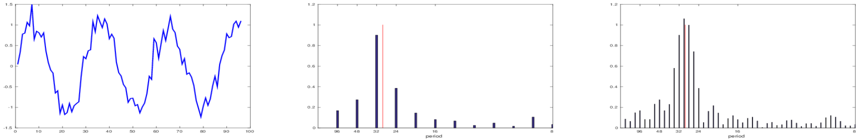
– VBA: Approximate $p(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_\xi|\mathbf{g})$
by $q_1(\mathbf{f})\, q_2(\mathbf{v}_\epsilon)\, q_3(\mathbf{v}_\xi)$

Results

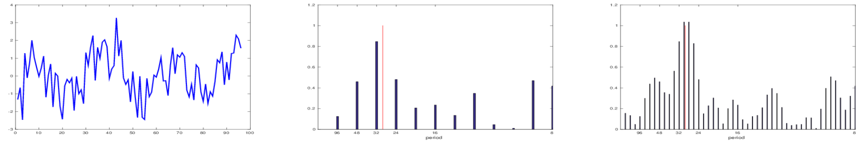
N=96, period=29,

- By using Fourier Transform technic, no way to find the right value of period.
- It is also difficult to estimate the shape of this repeating scheme.

Low noise case:

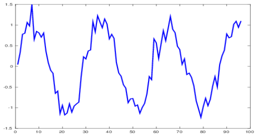
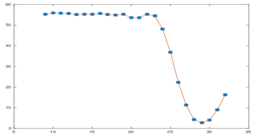
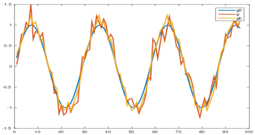


High noise case:

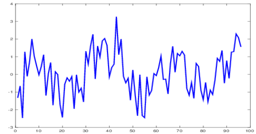
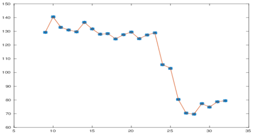
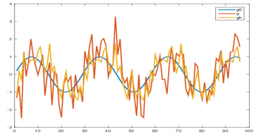


Results

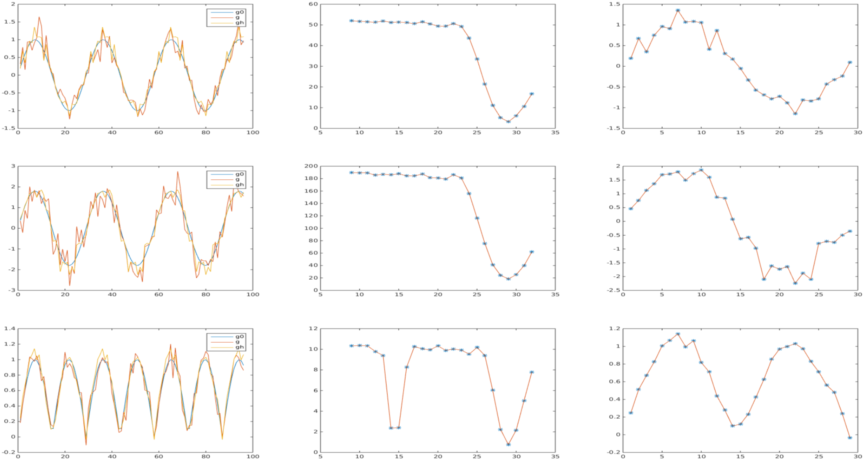
N=96, period=29,
Low noise case:

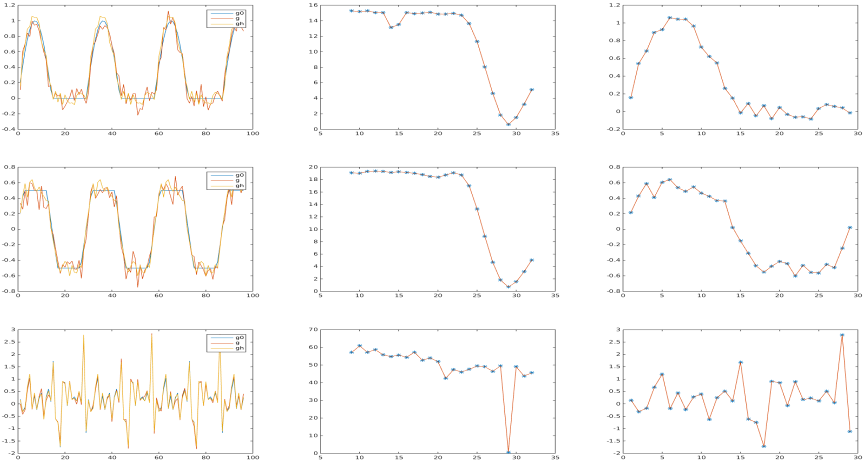
High noise case:

Results



Results



Conclusions

- the first step in any inference is to write down the relation between what you observe (data $\mathbf{g}$) and the unknowns $\mathbf{f}$.
- The second step is to model and assign priors to account for all uncertainties
- The third step is to use the Bayes rule to find the expression of the joint probability law of all the unknowns given the data and all the hyper parameters.
- Do the Bayesian computation, show the results
- Interpret your results
- Enjoy