Spectral estimation for non-stationary signal classes

Adrien Meynard

Aix-Marseille Université

SampTa 2017, Tallin

Motivations

In the context of audio signal processing, the analyzed sounds are generally non stationary.

Example:

Classical spectral estimation cannot be applied to this type of signals because the spectrum is only defined for stationary signals.
The questions are:

- How to model such signals?
- Can we define a spectrum and can we estimate it?

Stationarity

Stationarity

A stochastic process X is said to be stationary if its statistical properties are translation invariant.

A less general context concerns second order stationary processes i.e

- $\mathbb{E}\{X(t)\} = \mathbb{E}\{X(0)\} = m$,
- $\mathbb{E}\{X(t)X^*(\tau)\} = \mathbb{E}\{X(t-\tau)X^*(0)\} = k_X(t-\tau)$.

$\Rightarrow$ The power spectrum of a second order stationary process X is given by the Fourier transform of the autocorrelation k_x (Wiener-Khinchin theorem).

Goal

We consider classes of non-stationary processes that can be written as

$$Y(t) = \mathcal{T} X(t) \ ,$$

where X is a stationary process and $\mathcal{T}$ is a stationarity breaking operator.

Goal

Spectral estimation can be viewed as the joint estimation of $\mathcal{T}$ and the spectrum of the underlying stationary process X.
$\Rightarrow$ From a single realization y of Y, we estimate both $\mathcal{T}$ and $\mathscr{S}_X$.

Remark: This problem is far too general. Which types of operators $\mathcal{T}$ can we consider?

Considered deformations[1, 2]

- Frequency modulation

$$\mathscr{M}_\alpha : \qquad \mathscr{M}_\alpha x(t) = e^{2i\pi\alpha(t)} x(t) \ ,$$

where $\alpha \in C^2$ is a smooth function.

- Time warping

$$\mathscr{D}_\gamma : \qquad \mathscr{D}_\gamma x(t) = \sqrt{\gamma'(t)}\, x(\gamma(t)) \ ,$$

where γ is a smooth, strictly increasing function, assumed to fulfill the control conditions

$$\forall t,\ 0 < c_\gamma < \gamma'(t) < C_\gamma < \infty,\ \text{for some constant } c_\gamma, C_\gamma .$$

- Any combination of the two above deformations.

[1] M. Clerc and S. Mallat, "Estimating deformations of stationary processes," *Ann. Statist.*, vol. 31, no. 6, pp. 1772–1821, 12 2003.

[2] H. Omer, B. Torrésani, Time-frequency and time-scale analysis of deformed stationary processes, with application to non-stationary sound modeling, *Applied and Computational Harmonic Analysis*, Vol. 43, no. 1, July 2017, pp. 1-22

Observation Model

Model

Assume X is a zero mean, circular complex Gaussian stationary generalized random process. The observation is of the form

$$Y = \mathscr{M}_\alpha \mathscr{D}_\gamma X + W \ ,$$

where W is a white noise generalized random process.

Idea: Construct a representation of Y where the relation between the transform of Y and the transform of X is characterized by simple geometric transformation (translation, ...) depending on α and γ.

Let us first consider translation, modulation and rescaling operators:

$$T_\tau x(t) = x(t-\tau)\,,\; M_\nu x(t) = e^{2i\nu t}x(t)\,,\, D_s x(t) = q^{-s/2}x\left(q^{-s}t\right)$$

Adapted transform

$$\mathcal{V}_x(\nu, s, \tau) = \langle x, T_\tau M_\nu D_s \psi\rangle \;,$$

where ψ is a fixed analyzing waveform (concentrated around the origin).

Remarks:

- If $\nu = 0$, $\mathcal{W}_x(s,\tau) = \mathcal{V}_x(0, s, \tau)$ is the wavelet transform of x.
- If $s = 0$, $\mathcal{G}_x(\nu,\tau) = \mathcal{V}_x(\nu, 0, \tau)$ is the short time Fourier transform of x.

Example

Here, γ' is sinusoidal. Assuming that $\alpha(t) = \eta(t - \gamma(t))$, the adapted transform is the modified wavelet transform $\mathcal{V}_x(\eta, \cdot, \cdot)$.

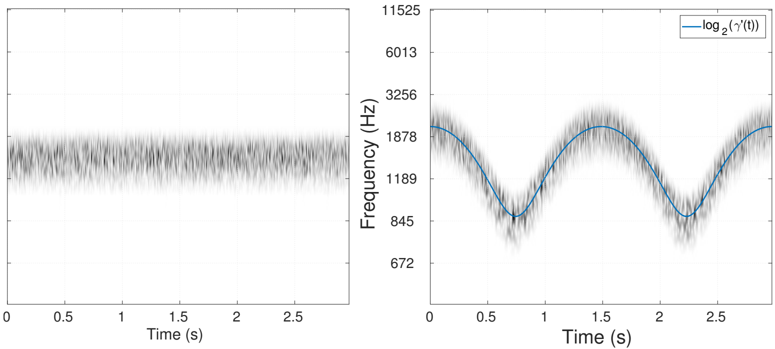

Figure: Scalograms of stationary and deformed signals.

$\Rightarrow$ The deformation generates a displacement of the coefficients in the adapted transform domain.

Approximation theorem

$$\mathcal{V}_Y(\nu, s, \tau) \approx e^{2i\pi\alpha(\tau)} \mathcal{V}_X\left(\frac{\nu - \alpha'(\tau)}{\gamma'(\tau)}, s + \log_q(\gamma'(\tau)), \gamma(\tau)\right)$$

The error $\epsilon = \mathcal{V}_Y - \widetilde{\mathcal{V}}_Y$ is a zero mean Gaussian random field with variance $\mathbb{E}\left\{|\epsilon(\nu, s, \tau)|^2\right\}$ that depends on the smoothness of α' and γ', and on the waveform decay rate.

$\Rightarrow$ How can we estimate these displacements in the time-scale-frequency space?

Estimation strategy

The joint estimation of the deformation operator and the spectrum of the underlying stationary signal is separated into two steps:

- Estimation of the deformation assuming that the spectrum is known,
- Estimation of the spectrum assuming that the deformation is known.

$\Rightarrow$ These two steps are computed alternatively until convergence of the estimators.

Step 1: Deformation estimation

Assume that the spectrum of the underlying stationary signal $\mathscr{S}_Z$ is known.

At fixed time τ, considering that the parameters Θ to estimate are

$$\Theta = (\theta_1, \theta_2) := (\alpha'(\tau), \log_q(\gamma'(\tau))).$$

Denote by $\mathbf{V}_y$ the restriction of $\mathcal{V}_y(\cdot, \cdot, \tau)$ to a finite sampling subset of the frequency-scale space. $\mathbf{V}_y$ is a zero mean circular, Gaussian random vector. This yields to the following log-likelihood

$$\mathscr{L}(\Theta) = -\ln\left|\det \mathbf{C}(\Theta)\right| - \mathbf{C}(\Theta)^{-1}\mathbf{V}_y \cdot \mathbf{V}_y \ ,$$

where

$$(\mathbf{C}(\Theta))_{ij} = q^{(s_i+s_j)/2} \int_0^\infty \mathscr{S}_Z(q^{-\theta_2}u)\overline{\hat{\psi}}\left[q^{s_i}\left(u + \theta_1 - \nu_i\right)\right] \hat{\psi}\left[q^{s_j}\left(u + \theta_1 - \nu_j\right)\right] du.$$

Step 2: Spectrum estimation

Assume that the deformation operators α and γ are known.

- The underlying stationary signal z is derived from the application of the inverse deformation to y

$$z := \mathscr{D}_{\gamma}{}^{-1}\mathscr{M}_{\alpha}{}^{-1}y = x + \mathscr{D}_{\gamma}^{-1}\mathscr{M}_{\alpha}^{-1}w \ .$$

- A spectral estimation $\hat{\mathscr{S}}_Z$ is performed on z using standard tools of spectral estimation on stationary signals (for example: Welch estimator).

Joint estimation scheme

Joint spectrum and deformation estimation algorithm

Initialization: provide an initial guess $\hat{\mathscr{S}}_Z^{(0)}$ for the spectrum.
while stopping criterion is false **do**

- As $\hat{\mathscr{S}}_Z^{(k-1)}$ is known, the deformation function estimators $\hat{\alpha}^{(k)}$ and $\hat{\gamma}^{(k)}$ are obtained by computing the approximated maximum likelihood estimator.
- Construct a "stationarized" signal $\hat{z}^{(k)}$ from y using $\hat{\alpha}^{(k)}$ and $\hat{\gamma}^{(k)}$, and estimate the corresponding power spectrum.
- k:= k+1

end while

Cramér-Rao lower bound

The maximum likelihood estimator being asymptotically unbiased and consistent, the Cramér-Rao lower bound provides relevant information regarding the achievable precision of the estimator.

Cramér-Rao lower bound and Slepian-Bangs formula

For any unbiased estimator $\hat{\theta}$ of a component θ of the multivariate parameter Θ,

$$\mathbb{E}\left\{(\hat{\theta}-\theta)^2\right\} \geq \mathrm{CRLB}(\theta).$$

When the observation is zero mean complex Gaussian

$$\mathrm{CRLB}(\theta) = \frac{1}{\mathrm{Trace}\left\{\left(\mathbf{C}(\Theta)^{-1}\frac{\partial \mathbf{C}(\Theta)}{\partial\theta}\right)^2\right\}}.$$

Toy example

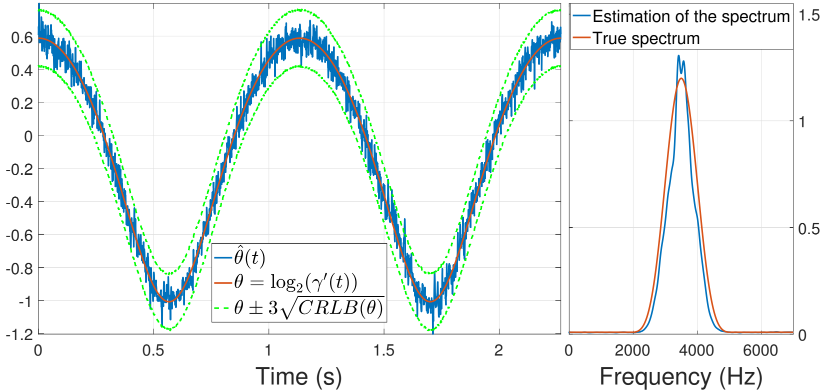

Figure: Joint time warping/spectrum estimation on a synthetic signal. Left: time warping function estimate (full, blue), ground truth (full, red) and Cramér-Rao bound (dotted, green); Right: spectrum of the underlying stationary signal.

Racing car engine

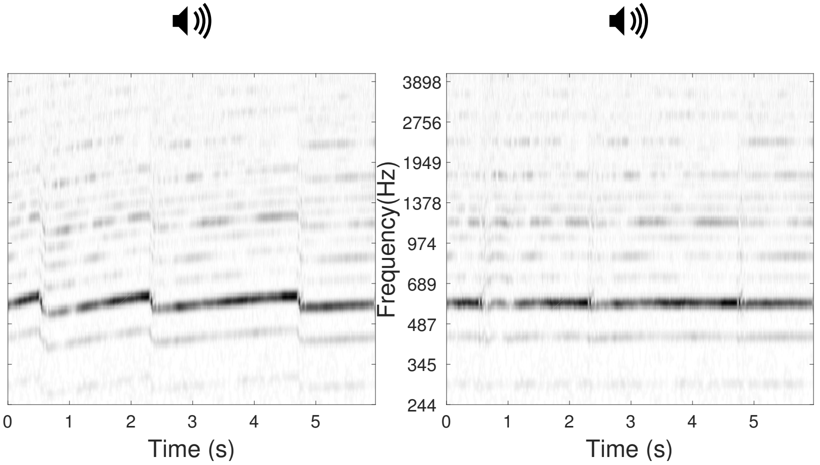

Figure: Joint time warping/spectrum estimation on an accelerating car engine: scalograms of the original signal and the estimated underlying stationary signal.

Racing car engine

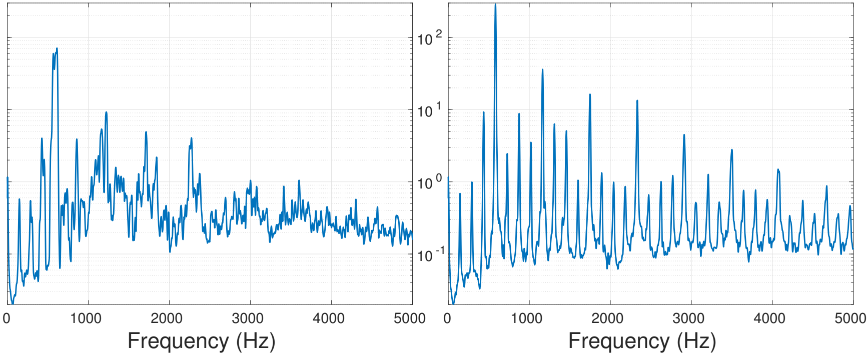

Figure: Estimation of the power spectra of the input signal (left) and the processed signals (right).

Singer

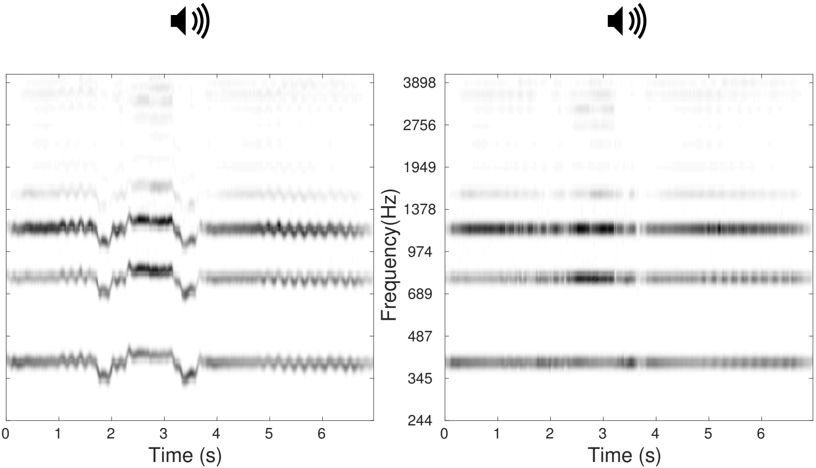

Figure: Joint time warping/spectrum estimation on a female voice singing: scalograms of the original signal and the estimated underlying stationary signal.

Singer

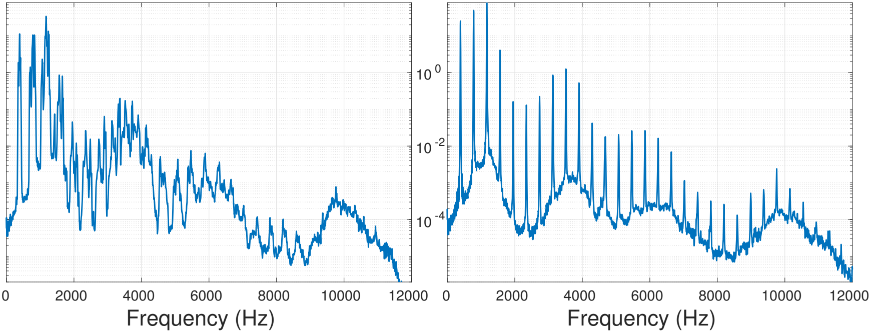

Figure: Estimation of the power spectra of the input signal (left) and the processed signals (right).

Summary

- We consider classes of non-stationary signals that are modeled as stationary signals deformed by linear operators (time warping and frequency modulation).
- The spectral estimation for non-stationary signal classes amounts to estimate the deformation operator together with the spectrum of the underlying stationary signal.
- An alternate algorithm is implemented.
- The formulation of the maximum likelihood problem as a continuous parameter estimation problem allows to get information about the precision of the estimator (Cramér-Rao bound).

Applications

- Analysis of non stationary signals in order to extract spectral properties.
- Synthesis of new signals applying any deformation to "stationarized" signals:

Perspectives

- The Gaussian assumption on the underlying stationary signal is not always relevant. ⇒ A new model adapted to time sparse signals must be able to synthesize such signals.
- Estimating other types of deformations like amplitude modulations.

Example:

THE END

Approximation error

If $|\psi(t)| \leq 1/(1+|t|^{\beta})$ for some $\beta > 2$, and that for all $u, v \in \mathbb{R}_+$,

$$I(u,v) := \sqrt{\left\langle \mathscr{S}_X, f_{u,v}^{(\beta)} \right\rangle} < \infty\,, \text{ with } f_{u,v}^{(\beta)}(\xi) = (u\xi + v)^{2\frac{\beta-1}{\beta+2}}.$$

Then

$$\mathbb{E}\left\{|\epsilon(s,\nu,\tau)|^2\right\} \leq q^{3s}\left(K_1\|\gamma''\|_\infty + K_2 q^{s\frac{\beta-4}{\beta+2}} I(\|\gamma''\|_\infty, \|\alpha''\|_\infty)\right)^2$$

where

$$K_1 = \frac{\beta\sigma_X}{2(\beta-2)\sqrt{c_\gamma}}\,, \quad K_2 = \left(\frac{\pi}{2}\right)^{\frac{\beta-1}{\beta+2}} \sqrt{C_\gamma}\,\frac{4(\beta+2)}{3(\beta-1)}$$

https://www.latp.univ-mrs.fr/~omer/SounDef/