The harmonic analysis of kernel functions

Mattia Zorzi

Department of Information Engineering
University of Padova

Cison di Valmarino, September 26th, 2016

Joint work with:

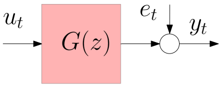
A. Chiuso (University of Padova)

Kernels in system identification

Kernels in system identification

Model class (e.g. OE models)

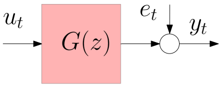
$$y_t = \sum_{s=1}^{\infty} g_s u_{t-s} + e_t \quad g_t \text{ impulse response}$$



Kernels in system identification

Model class (e.g. OE models)

$$y_t = \sum_{s=1}^{\infty} g_s u_{t-s} + e_t \quad g_t \text{ impulse response}$$



Gaussian linear regression model

$$y^N = \Phi\theta + e^N$$

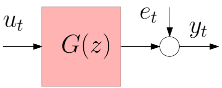
$$\theta \sim \mathcal{N}(0, K), \quad K \text{ kernel function}$$

$$y^N := \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}, \; \theta := \begin{bmatrix} g_1 \\ g_2 \\ \vdots \end{bmatrix}$$

Kernels in system identification

Model class (e.g. OE models)

$$y_t = \sum_{s=1}^{\infty} g_s u_{t-s} + e_t \quad g_t \text{ impulse response}$$



Gaussian linear regression model

$$y^N = \Phi\theta + e^N$$

$$\theta \sim \mathcal{N}(0, K), \quad K \text{ kernel function}$$

$$y^N := \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}, \; \theta := \begin{bmatrix} g_1 \\ g_2 \\ \vdots \end{bmatrix}$$

g_t is modeled as Gaussian process with zero mean and covariance function $\mathrm{Cov}[g_t g_s] = K(t, s)$

Kernels in system identification (cont'd)

Kernels in system identification (cont'd)

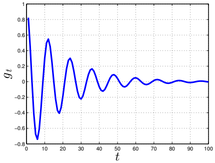
K encodes the a priori information on g_t

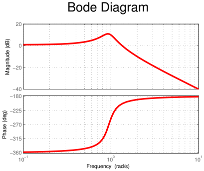
Kernels in system identification (cont'd)

K encodes the a priori information on g_t

Our a priori information on the impulse response:

- BIBO stable
- Frequency content



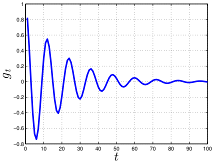


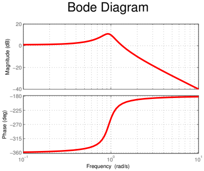
Kernels in system identification (cont'd)

K encodes the a priori information on g_t

Our a priori information on the impulse response:

- BIBO stable
- Frequency content





Question

How to embed this information in the kernel function?

Modeling the impulse response

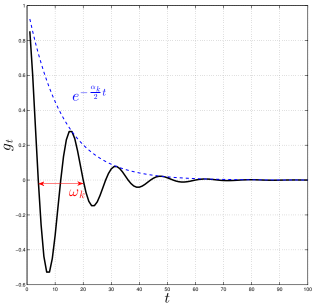
How?

Modeling the impulse response

How?

g_t as a sum of damped sinusoids

$$g_t = \sum_{k=1}^{M} |c_k| e^{-\frac{\alpha_k}{2}t} \cos(\omega_k t + \angle c_k)$$

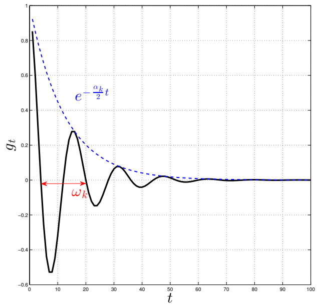


Modeling the impulse response

How?

g_t as a sum of damped sinusoids

$$g_t = \sum_{k=1}^{M} |c_k| e^{-\frac{\alpha_k}{2} t} \cos(\omega_k t + \angle c_k)$$



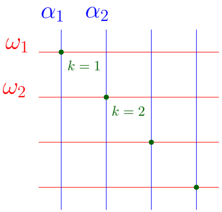
- c_k complex Gaussian random variable such that:
  - c_k is zero mean
  - $\mathrm{Cov}(c_k, \bar{c}_j) = p_k \delta_{k-j}$
  - $\mathrm{Cov}(c_k, c_j) = 0$

Modeling the impulse response (cont'd)

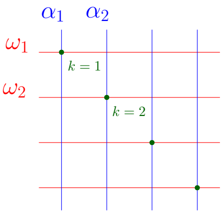
Modeling the impulse response (cont'd)

Sum of damped sinusoids in a grid

$$g_t = \sum_i \sum_j |c_{ij}| e^{-\frac{\alpha_j}{2} t} \cos(\omega_i t + \angle c_{ij})$$



Modeling the impulse response (cont'd)



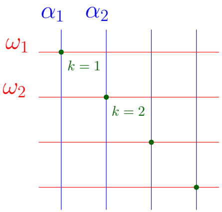
Sum of damped sinusoids in a grid

$$g_t = \sum_i \sum_j |c_{ij}| e^{-\frac{\alpha_j}{2} t} \cos(\omega_i t + \angle c_{ij})$$

"Infinite dense sum" of damped sinusoids

$$g_t = \int_0^{\infty} \int_{-\infty}^{\infty} |c(\alpha, \omega)| e^{-\frac{\alpha}{2} t} \cos(\omega t + \angle c(\alpha, \omega)) \mathrm{d}\omega \mathrm{d}\alpha$$

Modeling the impulse response (cont'd)



Sum of damped sinusoids in a grid

$$g_t = \sum_i \sum_j |c_{ij}| e^{-\frac{\alpha_j}{2} t} \cos(\omega_i t + \angle c_{ij})$$

"Infinite dense sum" of damped sinusoids

$$g_t = \int_0^\infty \int_{-\infty}^\infty |c(\alpha,\omega)| e^{-\frac{\alpha}{2} t} \cos(\omega t + \angle c(\alpha,\omega)) \mathrm{d}\omega \mathrm{d}\alpha$$

$c(\alpha,\omega)$ generalized Fourier transform of g_t

Harmonic analysis

Harmonic analysis

Let

$$g_t = \int_0^\infty \int_{-\infty}^\infty |c(\alpha,\omega)| e^{-\frac{\alpha}{2}t} \cos(\omega t + \angle c(\alpha,\omega)) \mathrm{d}\omega \mathrm{d}\alpha$$

Harmonic analysis

Let

$$g_t = \int_0^\infty \int_{-\infty}^\infty |c(\alpha,\omega)| e^{-\frac{\alpha}{2}t} \cos(\omega t + \angle c(\alpha,\omega)) \mathrm{d}\omega \mathrm{d}\alpha$$

Harmonic representation of the kernel function K

$$K(t,s) = \frac{1}{2}\int_0^\infty \int_{-\infty}^\infty p(\alpha,\omega) e^{-\alpha\frac{t+s}{2}} \cos(\omega(t-s)) \mathrm{d}\omega \mathrm{d}\alpha$$

Harmonic analysis

Let

$$g_t = \int_0^\infty \int_{-\infty}^\infty |c(\alpha,\omega)| e^{-\frac{\alpha}{2}t} \cos(\omega t + \angle c(\alpha,\omega)) \mathrm{d}\omega \mathrm{d}\alpha$$

Harmonic representation of the kernel function K

$$K(t,s) = \frac{1}{2} \int_0^\infty \int_{-\infty}^\infty p(\alpha,\omega) e^{-\alpha \frac{t+s}{2}} \cos(\omega(t-s)) \mathrm{d}\omega \mathrm{d}\alpha$$

$p(\alpha,\omega)$ generalized power spectral density

Harmonic analysis

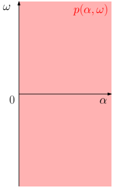
Let

$$g_t = \int_0^\infty \int_{-\infty}^\infty |c(\alpha,\omega)| e^{-\frac{\alpha}{2}t} \cos(\omega t + \angle c(\alpha,\omega)) \mathrm{d}\omega \mathrm{d}\alpha$$

Harmonic representation of the kernel function K

$$K(t,s) = \frac{1}{2} \int_0^\infty \int_{-\infty}^\infty p(\alpha,\omega) e^{-\alpha\frac{t+s}{2}} \cos(\omega(t-s)) \mathrm{d}\omega \mathrm{d}\alpha$$

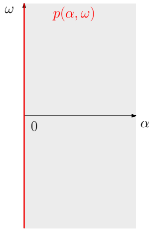
$p(\alpha,\omega)$ generalized power spectral density



Example 1: Stationary kernels

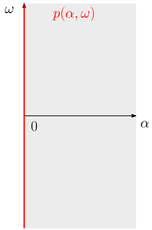
Example 1: Stationary kernels

Choice: $p(\alpha, \omega) = \delta(\alpha)q(\omega)$



Example 1: Stationary kernels

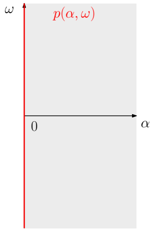
Choice: $p(\alpha,\omega) = \delta(\alpha)q(\omega)$



Stationary kernel

$$K(t-s) = \frac{1}{2}\int_{-\infty}^{\infty} q(\omega)\cos(\omega(t-s))\mathrm{d}\omega$$ $q(\omega)$ power spectral density

Example 1: Stationary kernels



Choice: $p(\alpha,\omega) = \delta(\alpha)q(\omega)$

Stationary kernel

$$K(t-s) = \frac{1}{2}\int_{-\infty}^{\infty} q(\omega)\cos(\omega(t-s))\mathrm{d}\omega$$ $q(\omega)$ power spectral density

Stationary process ("infinite dense sum" of sinusoids)

$$g_t = \int_{-\infty}^{\infty} |c(\omega)|\cos(\omega t + \angle c(\omega))\mathrm{d}\omega$$ $c(\omega)$ Fourier transform

Example 1: Stationary kernels (cont'd)

Example 1: Stationary kernels (cont'd)

- The power spectral density is even:

$$q(\omega) = \frac{\tilde{q}(\omega) + \tilde{q}(-\omega)}{2}$$

Example 1: Stationary kernels (cont'd)

- The power spectral density is even:

$$q(\omega) = \frac{\tilde{q}(\omega) + \tilde{q}(-\omega)}{2}$$

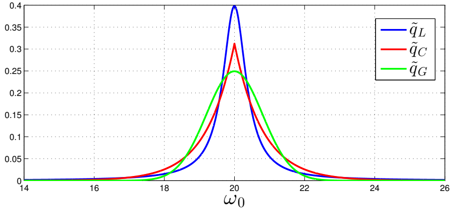
Shape	$\tilde{q}(\omega)$	Kernel
Laplacian	$\tilde{q}_L(\omega) = \frac{\beta/2}{\pi[(\omega-\omega_0)^2+(\beta/2)^2]}$	$K_L(t-s) = e^{-\frac{\beta}{2}\lvert t-s\rvert}\cos(\omega_0(t-s))$
Cauchy	$\tilde{q}_C(\omega) = \frac{1}{2\beta}e^{-\frac{\lvert\omega-\omega_0\rvert}{\beta}}$	$K_C(t-s) = \frac{1}{1+\beta^2(t-s)^2}\cos(\omega_0(t-s))$
Gaussian	$\tilde{q}_G(\omega) = \frac{1}{\sqrt{2\pi\beta^2}}e^{-\frac{(\omega-\omega_0)^2}{2\beta^2}}$	$K_G(t-s) = e^{-\frac{\beta^2(t-s)^2}{2}}\cos(\omega_0(t-s))$

Example 1: Stationary kernels (cont'd)

- The power spectral density is even:

$$q(\omega) = \frac{\tilde{q}(\omega) + \tilde{q}(-\omega)}{2}$$

Shape	$\tilde{q}(\omega)$	Kernel
Laplacian	$\tilde{q}_L(\omega) = \frac{\beta/2}{\pi[(\omega-\omega_0)^2+(\beta/2)^2]}$	$K_L(t-s) = e^{-\frac{\beta}{2}\lvert t-s\rvert}\cos(\omega_0(t-s))$
Cauchy	$\tilde{q}_C(\omega) = \frac{1}{2\beta} e^{-\frac{\lvert\omega-\omega_0\rvert}{\beta}}$	$K_C(t-s) = \frac{1}{1+\beta^2(t-s)^2}\cos(\omega_0(t-s))$
Gaussian	$\tilde{q}_G(\omega) = \frac{1}{\sqrt{2\pi\beta^2}} e^{-\frac{(\omega-\omega_0)^2}{2\beta^2}}$	$K_G(t-s) = e^{-\frac{\beta^2(t-s)^2}{2}}\cos(\omega_0(t-s))$



ω_0 center frequency

β bandwidth

Example 2: Locally stationary kernels

Example 2: Locally stationary kernels

Choice: $p(\alpha, \omega) = q_1(\alpha) q_2(\omega)$

Example 2: Locally stationary kernels

Choice: $p(\alpha, \omega) = q_1(\alpha) q_2(\omega)$

Locally stationary kernel (Silverman, 1957)

$$K(t,s) = K_1\left(\frac{t+s}{2}\right) \underbrace{K_2(t-s)}_{\text{stationary kernel}}$$

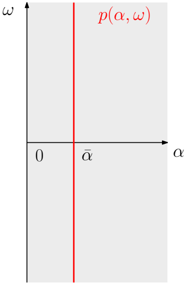
Example 2: Locally stationary kernels

Choice: $p(\alpha, \omega) = q_1(\alpha) q_2(\omega)$

Locally stationary kernel (Silverman, 1957)

$$K(t,s) = K_1\left(\frac{t+s}{2}\right) \underbrace{K_2(t-s)}_{\text{stationary kernel}}$$

Choice: $q_1(\alpha) = \delta(\alpha - \bar{\alpha})$



Example 2: Locally stationary kernels

Choice: $p(\alpha, \omega) = q_1(\alpha) q_2(\omega)$

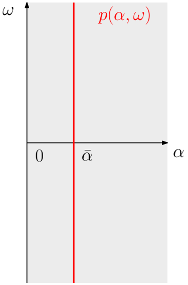
Locally stationary kernel (Silverman, 1957)

$$K(t, s) = K_1\left(\frac{t+s}{2}\right) \underbrace{K_2(t-s)}_{\text{stationary kernel}}$$

Choice: $q_1(\alpha) = \delta(\alpha - \bar{\alpha})$

ECLS kernel

$$K(t, s) = e^{-\bar{\alpha}\frac{t+s}{2}} K_2(t-s)$$



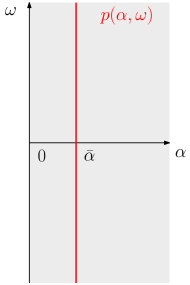
Example 2: Locally stationary kernels

Locally stationary kernel (Silverman, 1957)

$$K(t,s) = K_1\left(\frac{t+s}{2}\right) \underbrace{K_2(t-s)}_{\text{stationary kernel}}$$

Choice: $p(\alpha,\omega) = q_1(\alpha)q_2(\omega)$

Choice: $q_1(\alpha) = \delta(\alpha - \bar{\alpha})$



ECLS kernel

$$K(t,s) = e^{-\bar{\alpha}\frac{t+s}{2}} K_2(t-s)$$

- Examples of ECLS kernels (Chen-Ljung, 2015): stable-spline, dynamic-correlated, tuned-correlated...

Example 2: Locally stationary kernels (cont'd)

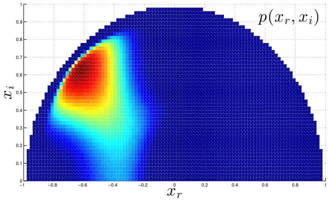
Example 2: Locally stationary kernels (cont'd)

Probability density function of a 2^{nd} order stable model

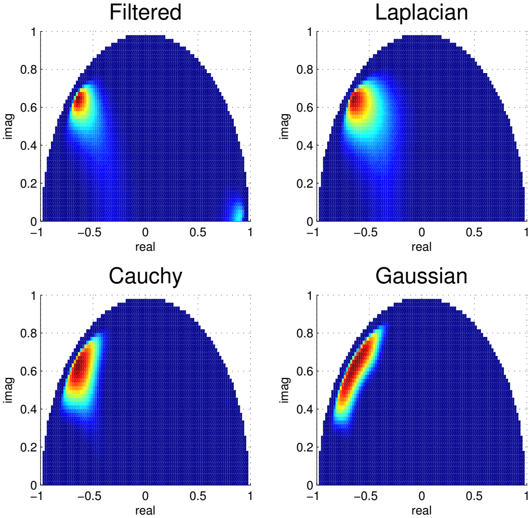
- Transfer function

$$G(z) = \frac{1}{(z-\rho)(z-\bar{\rho})} = \sum_{k=1}^{\infty} g_k z^{-k} \quad |\rho| < 1$$

- $\rho = x_r + jx_i$ pole of the model
- g_k process with kernel $K \rightarrow g_k$ with pdf $p(x_r, x_i)$



Example 2: Locally stationary kernels (cont'd)



- $\bar{\alpha} = 0.9$
- $\omega_0 = \frac{3}{4}\pi$
- β small

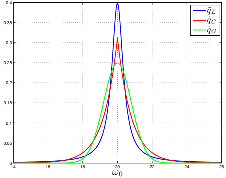
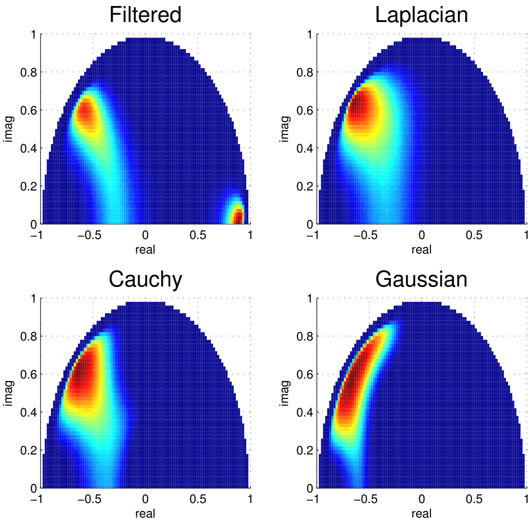


Figure: pdf of a 2^{nd} order stable model with ECLS

Example 2: Locally stationary kernels (cont'd)



- $\bar{\alpha} = 0.9$
- $\omega_0 = \frac{3}{4}\pi$
- β large

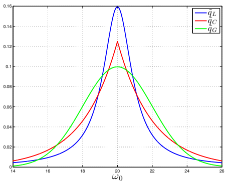


Figure: pdf of a 2$^{\text{nd}}$ order stable model with ECLS

Example 3: Integrated kernels

Example 3: Integrated kernels

Choice:

$$p(\alpha,\omega) = 1_{[\alpha_m,\alpha_M]}(\alpha)\frac{\alpha/2}{\pi[\omega^2+(\alpha/2)^2]}$$

Example 3: Integrated kernels

Choice:

$p(\alpha, \omega) = 1_{[\alpha_m, \alpha_M]}(\alpha) \frac{\alpha/2}{\pi[\omega^2 + (\alpha/2)^2]}$

Integrated TC kernel (Pillonetto *et. al.*)

$$K(t, s) = \frac{e^{-\alpha_m \max\{t,s\}} - e^{-\alpha_M \max\{t,s\}}}{\max\{t, s\}}$$

Example 3: Integrated kernels

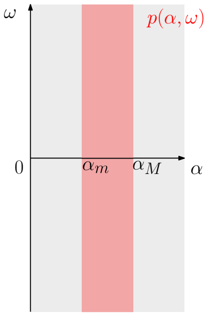
Integrated TC kernel (Pillonetto *et. al.*)

$$K(t,s) = \frac{e^{-\alpha_m \max\{t,s\}} - e^{-\alpha_M \max\{t,s\}}}{\max\{t,s\}}$$

Choice:

$p(\alpha,\omega) = 1_{[\alpha_m,\alpha_M]}(\alpha)\frac{\alpha/2}{\pi[\omega^2+(\alpha/2)^2]}$

Choice: $p(\alpha,\omega) = 1_{[\alpha_m,\alpha_M]}(\alpha)q(\omega)$



Example 3: Integrated kernels

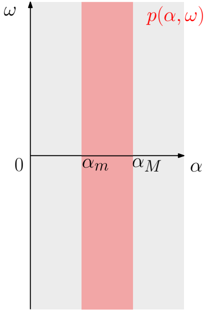
Choice:

$p(\alpha,\omega) = 1_{[\alpha_m,\alpha_M]}(\alpha)\frac{\alpha/2}{\pi[\omega^2+(\alpha/2)^2]}$

Integrated TC kernel (Pillonetto *et. al.*)

$$K(t,s) = \frac{e^{-\alpha_m \max\{t,s\}} - e^{-\alpha_M \max\{t,s\}}}{\max\{t,s\}}$$

Choice: $p(\alpha,\omega) = 1_{[\alpha_m,\alpha_M]}(\alpha)q(\omega)$



Integrated kernel

$$K(t,s) = 2\frac{e^{-\alpha_m \frac{t+s}{2}} - e^{-\alpha_M \frac{t+s}{2}}}{t+s} \underbrace{K(t-s)}_{\text{stationary kernel}}$$

Example 3: Integrated kernels (cont'd)

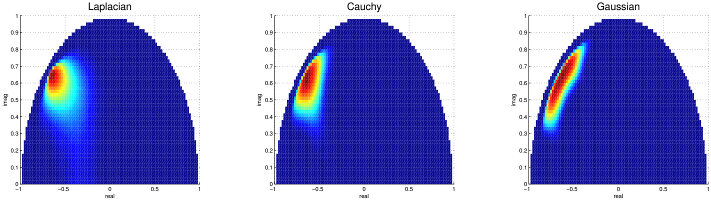


Figure: pdf of a 2^{nd} order model with ECLS kernel

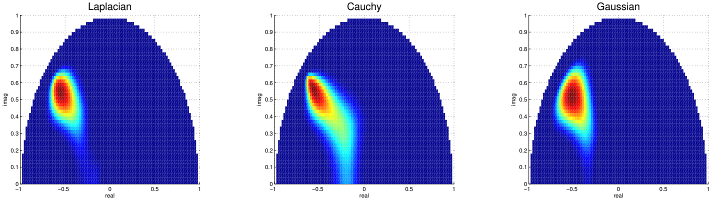


Figure: pdf of a 2^{nd} order model with INTEGRATED kernel

Conclusions

Key points:

Conclusions

Key points:

- Harmonic analysis for non-stationary Gaussian processes

Conclusions

Key points:

- Harmonic analysis for non-stationary Gaussian processes
- How the statistical power is distributed? generalized psd

Conclusions

Key points:

- Harmonic analysis for non-stationary Gaussian processes
- How the statistical power is distributed? generalized psd
- Relation between the generalized psd and the pdf of the process?

Conclusions

Key points:

- Harmonic analysis for non-stationary Gaussian processes
- How the statistical power is distributed? generalized psd
- Relation between the generalized psd and the pdf of the process?

Kernel design through the generalized psd

Conclusions

Key points:

- Harmonic analysis for non-stationary Gaussian processes
- How the statistical power is distributed? generalized psd
- Relation between the generalized psd and the pdf of the process?

Kernel design through the generalized psd

Thank you!