Spectral Analysis of Stationary Stochastic Process

Hanxiao Liu
hanxiaol@cs.cmu.edu

February 20, 2016

Outline

- Stationarity
- The time-frequency dual
  - Spectral representation
  - Marginal/conditional dependencies
- Inference

Stationary Stochastic Process

Strong stationarity: $\forall t_1, \ldots, t_k, h$

$$(X(t_1), \ldots, X(t_k)) \stackrel{D}{=} (X(t_1 + h), \ldots, X(t_k + h)) \qquad (1)$$

Weak/2nd-order stationarity:

$$\mathbb{E}\left(X(t)X(t)^\top\right) < \infty \qquad \forall t \qquad (2)$$

$$\mathbb{E}\left(X(t)\right) = \mu \qquad \forall t \qquad (3)$$

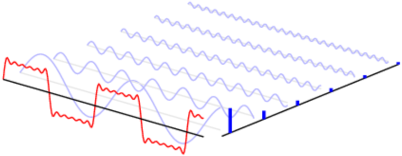
$$\mathrm{Cov}\left(X(t), X(t+h)\right) = \Gamma(h) \qquad \forall t, h \qquad (4)$$

The r.h.s. does not depend on t.

$\Gamma(h)$ autocovariance function (marginal dependencies)

$\Gamma(0)$ variance (power) of X

Spectral Representation Theorem



$$X(t) = \int_{-\pi}^{\pi} e^{iwt} dZ(\omega) \quad (5)$$

- $\mathbb{E}\left[dZ(\omega) dZ^*(\omega')\right] = 0$ if $\omega \neq \omega'$.
- $*$ denotes Hermitian (conjugate) transpose.

Compared to $X(t)$, we are more interested in $\Gamma(h)$—

[0] illustrative animation A and B.

Spectral Representation Theorem

$$\Gamma(h) = \mathbb{E}\left(X(0)X(h)^\top\right) \tag{6}$$
$$= \mathbb{E}\left\{\int_\omega e^0 dZ(\omega) \int_{\omega'} e^{iw'h} dZ^*(\omega')\right\} \tag{7}$$
$$= \int_\omega \int_{\omega'} e^{iw'h}\mathbb{E}\left[dZ(\omega)dZ^*(\omega')\right] \tag{8}$$
$$= \int_\omega e^{iwh}\mathbb{E}\left[dZ(\omega)dZ^*(\omega)\right] \tag{9}$$
$$= \int_\omega e^{iwh} s(\omega) d\omega \tag{10}$$

$\Gamma(h)$ - covariance with lag h (time domain)

$s(\omega)$ - covariance at frequency ω (freq domain)

Spectral Density Function

The Fourier transform pair

$$\Gamma(h) = \int_{\omega} e^{iwh} s(\omega) d\omega \tag{11}$$

$$s(\omega) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \Gamma(h) e^{-i\omega h} \tag{12}$$

We call s the *spectral density function*, since

$$\Gamma(0) = \int_{\omega} s(\omega) d\omega \tag{13}$$

$\Gamma(0) = \mathrm{Cov}(X(t), X(t)) =$ cumulative effect of $s(w)$

Marginal Dependencies

$\Gamma(h) \leftarrow$ sample autocovariance function

$$\hat{\Gamma}(h) = \frac{1}{N} \sum_{t=0}^{N-h-1} \left(X(t) - \bar{X}\right) \left(X(t+h) - \bar{X}\right)^\top \tag{14}$$

Asymptotic normality under mild assumptions.

$s(\omega) \leftarrow$ *periodogram.* Let $\omega_k = \frac{2\pi k}{N}$,

$$I(\omega_k) = d(k)d(k)^* \to \hat{s}(\omega) \tag{15}$$

where $d(k) := \frac{1}{N} \sum_{t=0}^{N-1} X(t) e^{-ikt}$ is obtained via DFT.

- bad estimator in general
- good estimator with appropriate smoothing

For time-series i and j

$$
\begin{aligned}
& X_i \perp\!\!\!\perp X_j \mid X_{V \setminus \{i,j\}} && (16)\\
\iff & \operatorname{Cov}\left(X_i(t), X_i(t+h) \mid X_{V \setminus \{i,j\}}\right) = 0, \quad \forall h && (17)\\
\iff & (\Gamma(h)^{-1})_{ij} = 0, \quad \forall h && (18)\\
\iff & (s(\omega)^{-1})_{ij} = 0, \quad \forall \omega \in [0, 2\pi] && (19)
\end{aligned}
$$

Inferring conditional dependences

- = inferring $\Gamma(h)^{-1}$
- = inferring $s(\omega)^{-1}$

Applicable to any stationary X

Autoregressive Gaussian Process

The Autoregressive (AR) process

$$X(t) = -\sum_{h=1}^{p} A_h X(t-h) + \epsilon(t) \tag{20}$$

$\epsilon(t)$ Gaussian white noise $\sim \mathcal{N}(0, \Sigma)$

We'd like to parametrize $s(\omega)^{-1}$ with A

- Inferring conditional dependences for AR can be cast as an optimization problem w.r.t. A

Filter Theorem

For any stationary X and $\{a_t\}$ s.t. $\sum_{t=-\infty}^{\infty} |a_t| < \infty$, process $Y(t) = \sum_{h=-\infty}^{\infty} a_h X(t-h)$ is stationary with

$$s_Y(\omega) = |\mathcal{A}(e^{i\omega})|^2 s_X(\omega) \tag{21}$$

where $\mathcal{A}(z) = \sum_{-\infty}^{\infty} a_h z^{-h}$

In 1-d AR,
$\epsilon(t) = x(t) + \sum_{h=1}^{p} a_h x(t-h) \implies s(\omega)^{-1} = \frac{|\mathcal{A}(e^{i\omega})|^2}{\sigma^2}$

Multi-dimensional analogy:

$$s(\omega)^{-1} = \mathcal{A}(e^{i\omega})\Sigma^{-1}\mathcal{A}(e^{i\omega})^* \tag{22}$$

where $\mathcal{A}(z) = \sum_{h=0}^{p} A_h z^{-h}$, $A_0 := I$.

Parametrized Spectral Density

Parametrize $s(\omega)^{-1}$ by AR parameters

$$s(\omega)^{-1} = \left[\sum_{h=0}^{p} A_h e^{-ih\omega}\right]\Sigma^{-1}\left[\sum_{h=0}^{p} A_h e^{-ih\omega}\right]^* \tag{23}$$

$$= Y_0 + \frac{1}{2}\sum_{h=1}^{p}\left(e^{-ih\omega}Y_h + e^{ih\omega}Y_h^\top\right) \tag{24}$$

where $Y_0 = \sum_{h=0}^{p} A_h^\top \Sigma^{-1} A_h$, $Y_h = 2\sum_{i=0}^{p-h} A_i^\top \Sigma^{-1} A_{i+h}$

$B_h \overset{def}{=} \Sigma^{-\frac{1}{2}} A_h \implies Y_0 = \sum_{h=0}^{p} B_h^\top B_h$, $Y_h = 2\sum_{i=0}^{p-h} B_i^\top B_{i+h}$

$(s(\omega)^{-1})_{ij} = 0 \iff (Y_h)_{ij} = (Y_h)_{ji} = 0$, $\forall 0, \dots, p$, i.e.
linear constraints over $Y \iff$ quadratic constraints over B

Conditional MLE

Simplification: fix $x(1), \dots x(p)$

$$\epsilon(t) = \sum_{h=0}^{p} A_h x(t-h) \tag{25}$$

$$= [A_0, \dots, A_h] \begin{bmatrix} x(t) \\ x(t-1) \\ \vdots \\ x(t-p) \end{bmatrix} := A\mathbf{x}(t) \sim \mathcal{N}(0, \Sigma) \tag{26}$$

A least-squares estimate. Likelihood =

$$\frac{e^{-\frac{1}{2}\sum_{t=p+1}^{N} \mathbf{x}(t)^\top A^\top \Sigma^{-1} A\mathbf{x}(t)}}{(2\pi)^{\frac{m(N-p)}{2}} (\det \Sigma)^{\frac{N-p}{2}}} \overset{B=\Sigma^{-\frac{1}{2}}A}{=\!=\!=} \frac{e^{-\frac{1}{2}\sum_{t=p+1}^{N} \mathbf{x}(t)^\top B^\top B\mathbf{x}(t)}}{(2\pi)^{\frac{m(N-p)}{2}} (\det B_0)^{p-N}} \tag{27}$$

Regularized ML

Maximize log-likelihood

$$\min_{B} \quad -2\log\det B_0 + \operatorname{tr}\left(CB^\top B\right) \tag{28}$$

Solution given by Yule-Walker equations.

Enforcing sparsity over $s(\omega)^{-1}$

$$\min_{B} \quad -2\log\det B_0 + \operatorname{tr}\left(CB^\top B\right) + \gamma\|D(B^\top B)\|_1 \tag{29}$$

Convex relaxation:

$$\min_{Z\succeq 0} \quad -\log\det Z_{00} + \operatorname{tr}\left(CZ\right) + \gamma\|D(Z)\|_1 \tag{30}$$

- Exact if $\operatorname{rank}(Z^*) \leq m$
- Bregman divergence + ℓ_1-regularization. Well studied.

Non-stationary Extensions

With stationarity

$$s\,(\omega) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \Gamma(h) e^{-i\omega h} \tag{31}$$

No stationarity? The Wigner-Ville spectrum

$$s(t, \omega) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \Gamma\left(t + \frac{h}{2}, t - \frac{h}{2}\right) e^{-i\omega h} \tag{32}$$

Other types of power spectra

- Rihaczek spectrum
- (Generalized) Evolutionary spectrum

Reference I

Bach, F. R. and Jordan, M. I. (2004).
Learning graphical models for stationary time series.
Signal Processing, IEEE Transactions on, 52(8):2189–2199.

Basu, S., Michailidis, G., et al. (2015).
Regularized estimation in sparse high-dimensional time series models.
The Annals of Statistics, 43(4):1535–1567.

Matz, G. and Hlawatsch, F. (2003).
Time-varying power spectra of nonstationary random processes.
na.

Pereira, J., Ibrahimi, M., and Montanari, A. (2010).
Learning networks of stochastic differential equations.
In *Advances in Neural Information Processing Systems*, pages 172–180.

Songsiri, J., Dahl, J., and Vandenberghe, L. (2010).
Graphical models of autoregressive processes.
Convex Optimization in Signal Processing and Communications, pages 89–116.

Reference II

Songsiri, J. and Vandenberghe, L. (2010).
Topology selection in graphical models of autoregressive processes.
The Journal of Machine Learning Research, 11:2671–2705.

Tank, A., Foti, N. J., and Fox, E. B. (2015).
Bayesian structure learning for stationary time series.
In *Uncertainty in Artificial Intelligence, UAI 2015, July 12-16, 2015, Amsterdam, The Netherlands*, pages 872–881.