Time Series Analysis by Higher Order Crossings

Benjamin Kedem
Department of Mathematics & ISR
University of Maryland
College Park, MD

Waseda University, Tokyo, Japan, March 2019

• Higher Order Crossings (HOC) is a fast time series analysis method based on zero-crossing counts obtained from sequentially filtered time series.

• Filter, count, filter, count, filter, count, etc.

• Filter, observe, filter, observe, filter, observe, etc.

• Some basic results about oscillation:
a. HOC Theorem, K- and Slud (1982).
b. Sinusoidal results, K- (1984), K- Li (1991).
c. Convergence of ZC rate, K- and Slud (1994).
d. Same spectrum, same oscillation??? Barnett and K- (1991).

• Applications:
a. Spectral Analysis and Discrimination by ZC.
K- (1986), (1994).
b. Estimation of frequencies in mixed spectra.
He and K- (1989), Li and K- (1993).

- Applications:

a. Non-destructive evaluation.

b. Discrimination analysis of discontinuous breath sounds.

c. Indirect validation problem of travel time data.

d. Magnetic anomaly detection.

e. Velocity estimation for cellular systems.

f. Emotion recognition from EEG.

g. Assessment of severity in diabetic autonomic neuropathy.

h. Identification of changes in the entropy of seismic signals preceding an event.

i. Distinguishing liver regeneration indices in rats.

j. At least 12 patents reference HOC.

Stochastic Process

A stochastic or random process $\{Z_t\}$, $\cdots, -1, 0, 1, \cdots$, is a collection of random variables, real or complex-valued, defined on the same probability space.

Gaussian Process: A real-valued process $\{Z_t\}$, $t \in T$, is called Gaussian process if for all $t_1, t_2, \cdots, t_n \in T$, the joint distribution of $(Z_{t_1}, Z_{t_2}, \cdots, Z_{t_n})$ is multivariate normal.

The finite dimensional distributions of a Gaussian process are completely determined from:

$$m(t) = E[Z_t]$$

and

$$R(s,t) = Cov[Z_s, Z_t].$$

Markov Process: For $t_1 < \cdots < t_{n-2} < t_{n-1} < t_n$

$$P(Z_{t_n} \leq z | Z_{t_{n-1}}, Z_{t_{n-2}}, ...Z_{t_1}) = P(Z_{t_n} \leq z | Z_{t_{n-1}})$$

Stationary Processes

A stochastic process $\{Z_t\}$ is said to be a *strictly stationary* process if its joint distributions are invariant under time shifts:

$$(Z_{t_1}, Z_{t_2}, \cdots, Z_{t_n}) \overset{Dist}{=} (Z_{t_1+\tau}, Z_{t_2+\tau}, \cdots, Z_{t_n+\tau})$$

for all $t_1, t_2, \cdots, t_n$, n, and τ.

When 2nd order moments exist, strict stationarity implies:

$$E[Z_t] = E[Z_{t+\tau}] = E[Z_0] = m \tag{1}$$

$$Cov[Z_t, Z_s] = R(t-s). \tag{2}$$

$\{Z_t\}$ is called *weakly stationary* when (1),(2) hold.

For simplicity, we shall assume all our processes are both strictly and weakly stationary and also real-valued.

Assume: $E(Z_t) = 0$.

Autocovariance:

$$R_k = E(Z_t Z_{t-k}) = \int_{-\pi}^{\pi} \cos(k\lambda) dF(\lambda) \tag{3}$$

Autocorrelation:

$$\rho_k = \frac{R_k}{R_0} = \int_{-\pi}^{\pi} \cos(k\lambda) d\overline{F}(\lambda)$$

Spectral Distribution:

$$F(\lambda), \quad -\pi \leq \lambda \leq \pi$$

When F is absolutely continuous, we have the *Spectral Density*:

$$f(\lambda) = F'(\lambda), \quad -\pi \leq \lambda \leq \pi$$

In general in practice:

$$F(\lambda) = F_c(\lambda) + F_d(\lambda)$$

where $F_c(\lambda)$ is absolutely continuous and $F_d(\lambda)$ is a step function, both monotone nondecreasing.

$R_k, \rho_k, f(\lambda)$ are symmetric.

spectral representation

A. Kolmogorov and H. Cramér in the early 1940's.

Let $\{Z_t\}$, $t = 0, \pm 1, \pm 2, \cdots$, *be a zero mean weakly stationary process. Then (3) implies*

$$Z_t = \int_{-\pi}^{\pi} e^{it\lambda} d\xi(\lambda), \quad t = 0, \pm 1, \cdots \tag{4}$$

where now the spectral distribution satisfies

$$E[d\xi(\lambda)\overline{d\xi(\omega)}] = \begin{cases} dF(\lambda), & \text{if } \lambda = \omega \\ 0, & \text{if } \lambda \neq \omega \end{cases} \tag{5}$$

We may interpret

$$dF(\lambda) = E|d\xi(\lambda)|^2$$

as the weight or "power" given to frequency λ.

Example: Sum of Random Sinusoids.

$$Z_t = \sum_{j=1}^{p} \{A_j \cos(\omega_j t) + B_j \sin(\omega_j t)\}, \quad t = 0, \pm 1, \cdots \quad (6)$$

$A_1, \cdots, A_p, B_1, \cdots, B_p$ uncorrelated. $E[A_j] = E[B_j] = 0$, $Var[A_j] = Var[B_j] = \sigma_j^2$, $\omega_j \in (0, \pi)$, for all j.

Then for all t, $E[Z_t] = 0$, and

$$\begin{aligned} R_k &= E[Z_t Z_{t-k}] = \sum_{j=1}^{p} \sigma_j^2 \cos(\omega_j k) \\ &= \sum_{j=1}^{p} \{\frac{1}{2}\sigma_j^2 \cos(\omega_j k) + \frac{1}{2}\sigma_j^2 \cos(-\omega_j k)\}, \quad k = 0, \pm 1, \cdots \end{aligned}$$

$$\rho_k = \frac{R_k}{R_0} = \frac{\sum_{j=1}^{p} \sigma_j^2 \cos(\omega_j k)}{\sum_{j=1}^{p} \sigma_j^2}, \quad k = 0, \pm 1, \cdots \quad (7)$$

Discrete spectrum:

$F(\omega)$ is a nondecreasing step function with jumps of size $\frac{1}{2}\sigma_j^2$ at $\pm\omega_j$, and $F(-\pi) = 0, F(\pi) = R_0 = \sum_{j=1}^{p} \sigma_j^2$.

Example: Stationary AR(1) Process.

Let $\{\epsilon_t\}$, $t = 0, \pm 1, \pm 2, \cdots$, be a sequence of uncorrelated real-valued random variables with mean zero and variance σ_ϵ^2. Define

$$Z_t = \phi_1 Z_{t-1} + \epsilon_t, \quad t = 0, \pm 1, \pm 2, \cdots \tag{8}$$

where $|\phi_1| < 1$. The process (8) is called a *first order autoregressive process* and is commonly denoted by $AR(1)$.

$$E[Z_t] = E\left\{\lim_{n\to\infty} \sum_{j=0}^{n} \phi_1^j \epsilon_{t-j}\right\} = \lim_{n\to\infty} E\left\{\sum_{j=0}^{n} \phi_1^j \epsilon_{t-j}\right\} = 0$$

$$R_k = \frac{\sigma_\epsilon^2 \phi_1^{|k|}}{1-\phi_1^2}, \quad k = 0, \pm 1, \pm 2, \cdots$$

$$\rho_k = \phi_1^{|k|}, \quad k = 0, \pm 1, \pm 2, \cdots$$

Continuous spectrum:

$$f(\lambda) = \frac{\sigma_\epsilon^2}{2\pi} \cdot \frac{1}{1 - 2\phi_1 \cos(\lambda) + \phi_1^2}, \quad -\pi \le \lambda \le \pi.$$

Example: Sum of Random Sinusoids Plus Noise.

Let $\{Z_t\}$ be a "signal" as in (6), and let $\{\epsilon_t\}$ be a zero mean weakly stationary "noise" uncorrelated with $\{Z_t\}$ and with a spectrum which possesses a spectral density $f_\epsilon(\omega)$,

$$F_\epsilon(\omega) = \int_{-\pi}^{\omega} f_\epsilon(\lambda) d\lambda$$

Then the process

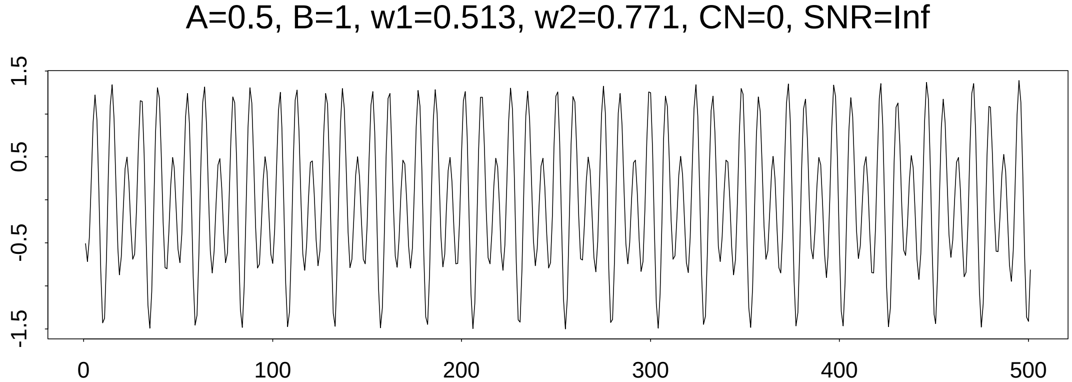
$$Y_t = \sum_{j=1}^{p} \{A_j \cos(\omega_j t) + B_j \sin(\omega_j t)\} + \epsilon_t, \quad t = 0, \pm 1, \cdots \quad (9)$$

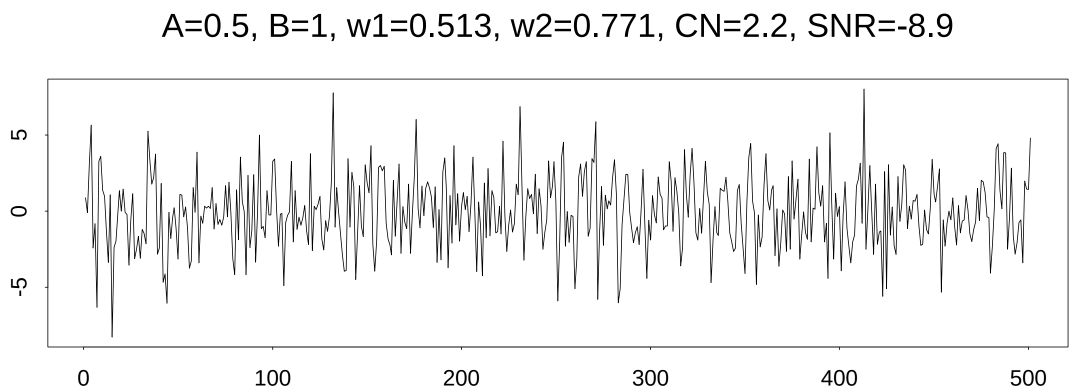
has a *mixed spectrum* of the form:

$$F_y(\omega) = F_z(\omega) + F_\epsilon(\omega) = \sum_{\lambda_j \leq \omega} \frac{1}{2}\sigma_j^2 + \int_{-\pi}^{\omega} f_\epsilon(\lambda) d\lambda$$

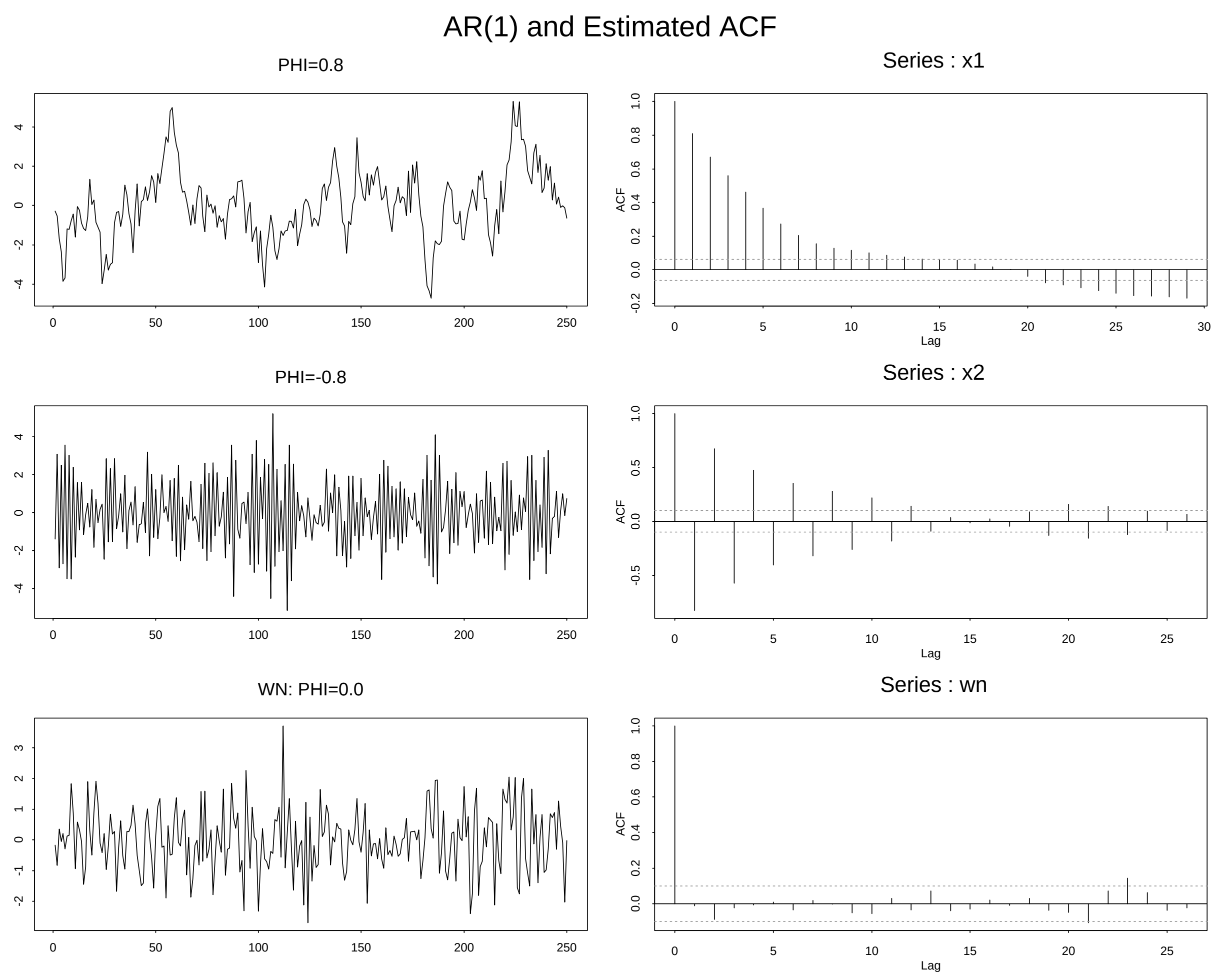
$$\lambda_j \in \{-\omega_p, -\omega_{p-1}, ..., \omega_{p-1}, -\omega_p\}$$

(a) Sum of two sinusoids.
(b) Sum of two sinusoids plus white noise.

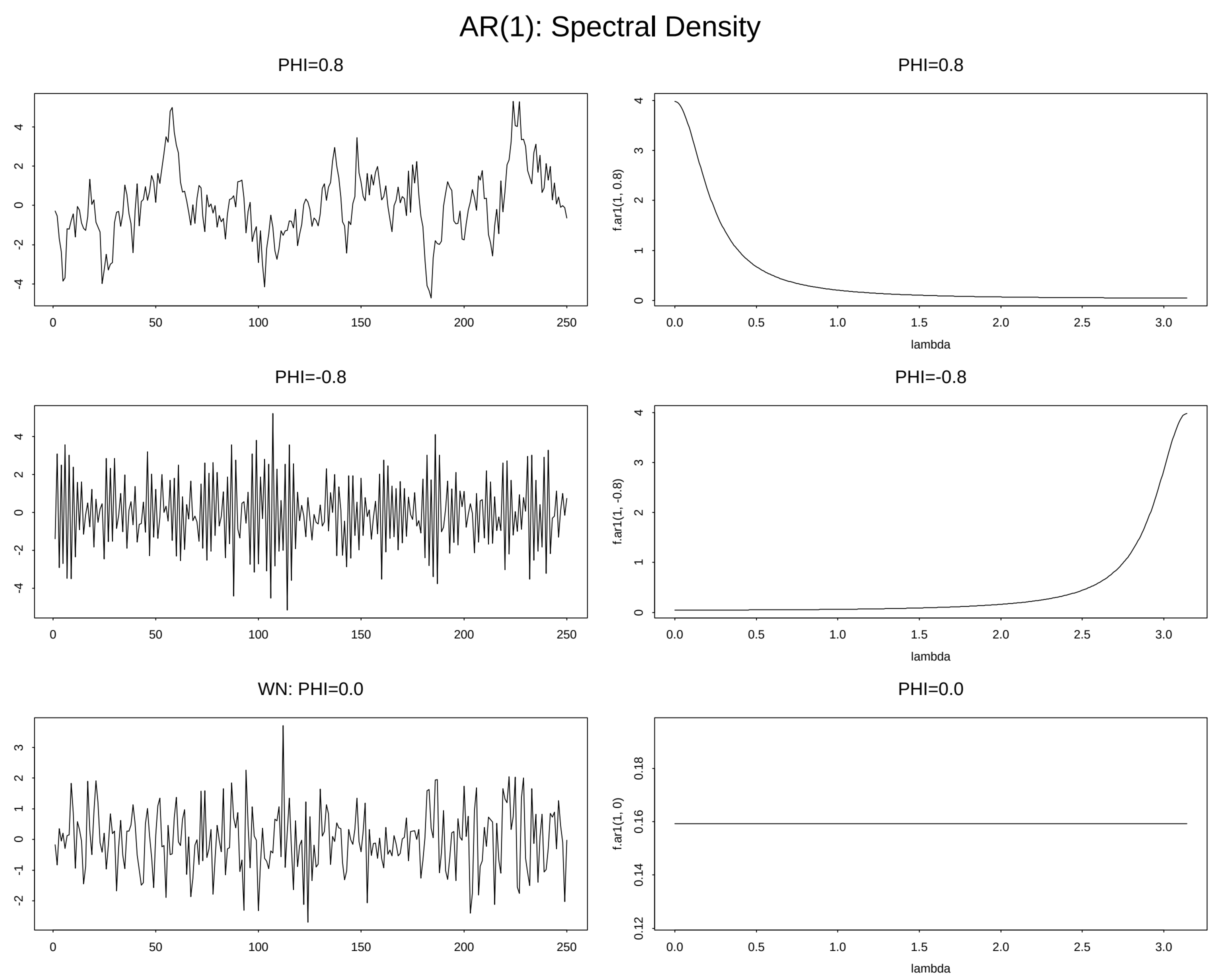




Plots of AR(1) time series and their estimated autocorrelation.



Plots of AR(1) time series and their spectral densities on $[0, \pi]$.



Linear Filtering

By a *time invariant linear filter* applied to a stationary time series $\{Z_t\}$, $t = 0, \pm 1, \ldots$, we mean the linear operation or convolution,

$$Y_t = \mathcal{L}(\{Z_t\}) = \sum_{j=-\infty}^{\infty} h_j Z_{t-j} \tag{10}$$

with

$$H(\lambda) \equiv \sum_{j=-\infty}^{\infty} h_j e^{-ij\lambda}, \quad 0 < \lambda \leq \pi$$

The function $H(\lambda)$ is called the *transfer function*.

$|H(\lambda)|$ is called the *gain*.

Fact:

$$dF_y(\lambda) = |H(\lambda)|^2 dF_z(\lambda) \tag{11}$$

In particular, when spectral densities exist we have

$$f_y(\lambda) \;\; = \;\; |H(\lambda)|^2 f_z(\lambda) \tag{12}$$

This is an important relationship between the input and output spectral densities.

The Difference Operator:

$$\nabla Z_t \equiv Z_t - Z_{t-1}$$

This is a linear filter with $h_0 = 1, h_1 = -1$, and $h_j = 0$ otherwise.

The transfer function is,

$$H(\lambda) = 1 - e^{-i\lambda}$$

and the squared gain is

$$|H(\lambda)|^2 = |1 - e^{-i\lambda}|^2 = 2(1 - \cos\lambda)$$

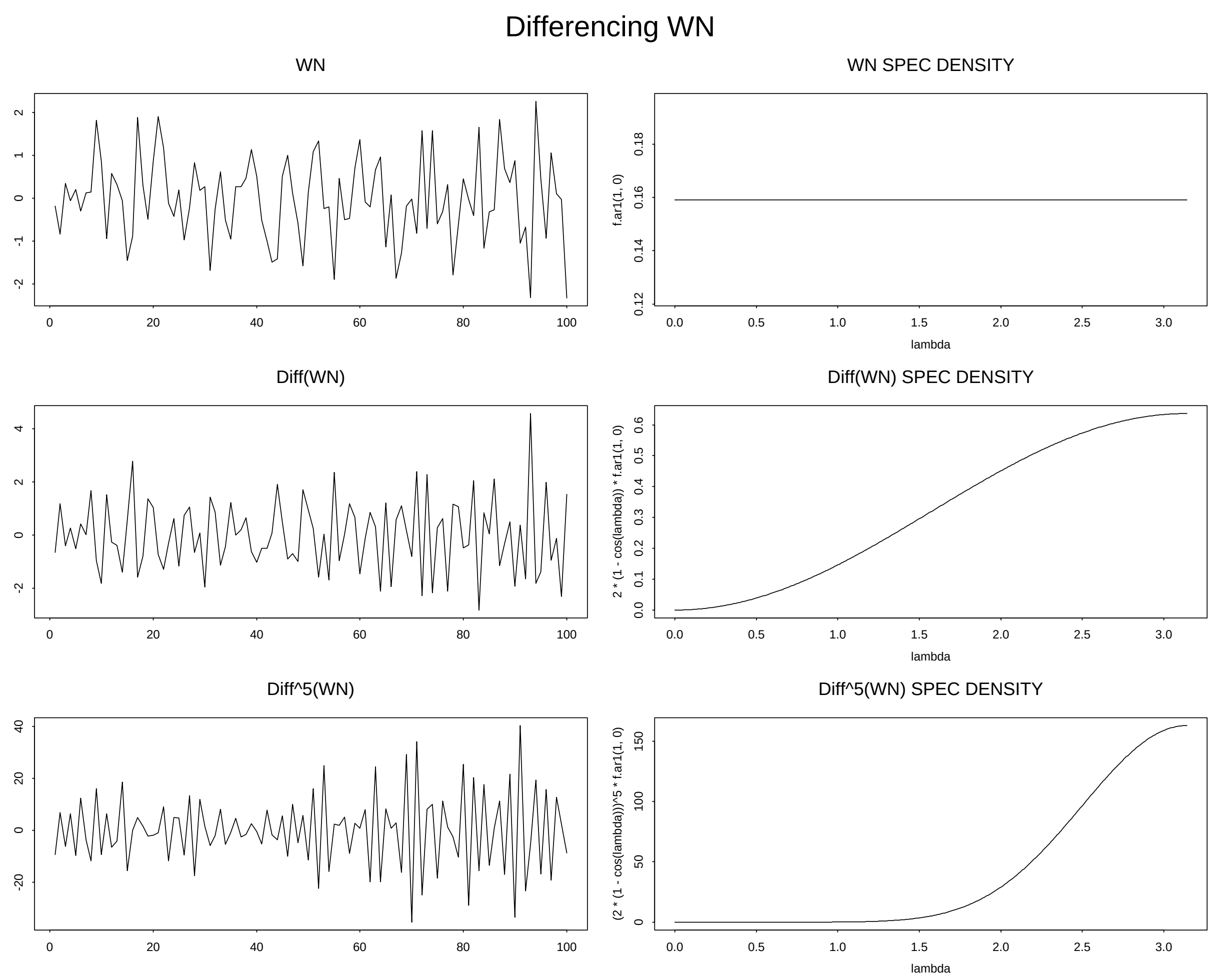
In $[0, \pi]$ the gain is monotone increasing and hence this is a high-pass filter.

The squared gain of the second difference ∇^2 is

$$4(1 - \cos\lambda)^2,$$

and hence this is a more pronounced high-pass filter. Repeated differencing is a simple way to obtain high-pass filters.

Differencing white noise: Higher frequencies get more power.



We define a parametric filter by the convolution,

$$Z_t(\theta) \equiv \mathcal{L}_\theta(Z)_t = \sum_n h_n(\theta) Z_{t-n} \tag{13}$$

In other words

$$Z_t(\theta) \equiv h_t(\theta) \otimes Z_t \tag{14}$$

where $\otimes$ denotes convolution.

The Parametric AR(1) Filter.

Let $|\alpha| < 1$.

The $AR(1)$ (or α) filter is the recursive filter

$$Y_t = \alpha Y_{t-1} + Z_t$$

or

$$Y_t = \mathcal{L}_\alpha(Z)_t = Z_t + \alpha Z_{t-1} + \alpha^2 Z_{t-2} + \cdots$$

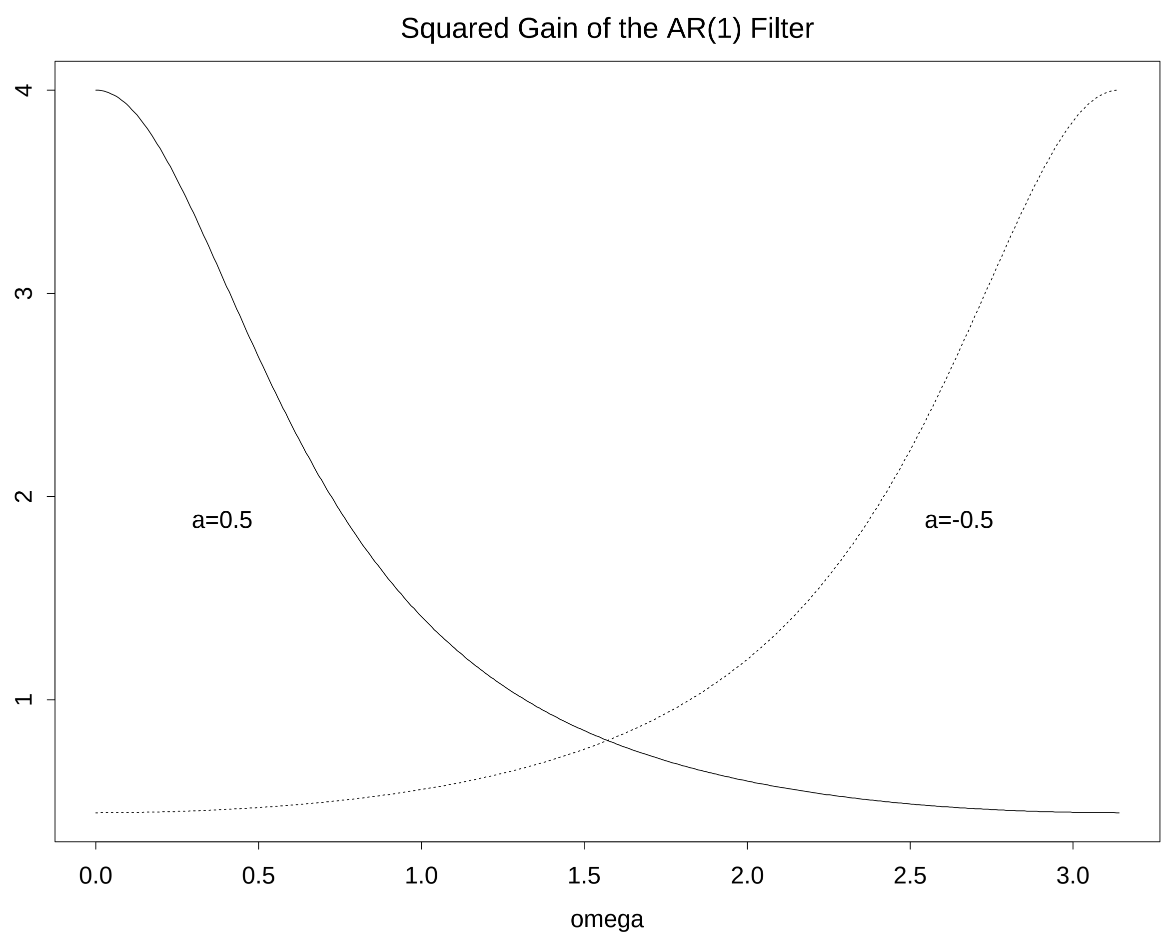
The transfer function for $\omega \in [0, \pi]$ is

$$H(\lambda) = \frac{1}{1 - \alpha e^{-i\lambda}}$$

The squared gain is

$$|H(\omega; \alpha)|^2 = \frac{1}{1 - 2\alpha \cos(\omega) + \alpha^2}, \quad \alpha \in (-1, 1). \quad (15)$$

For $\alpha > 0$ the AR(1) filter is a low-pass filter, and a high-pass for $\alpha < 0$.



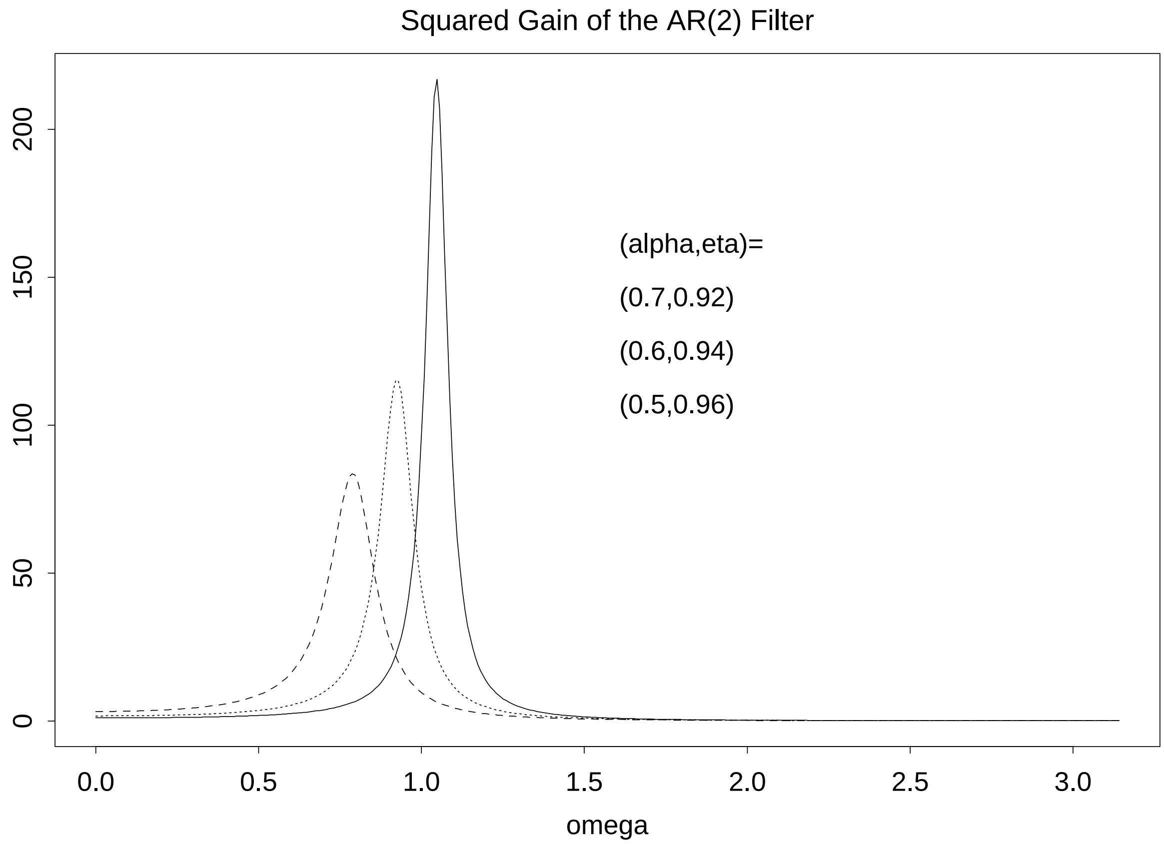
The Parametric AR(2) Filter.

With $\alpha \in (-1,1)$, define $Y_t(\alpha)$ by operating on Z_t,

$$Y_t(\alpha) = (1+\eta^2)\alpha Y_{t-1}(\alpha) - \eta^2 Y_{t-2}(\alpha) + Z_t \qquad (16)$$

where $\eta \in (0,1)$ is the **bandwidth parameter**.

Squared gains of the AR(2) filter centered approximately at $\cos^{-1}(\alpha)$ for η close to 1.



Zero-crossings in Discrete Time

Let $Z_1, Z_2, \cdots, Z_N$ be a zero-mean stationary time series.

The zero-crossing count in discrete time is defined as the number of symbol changes in the corresponding clipped binary time series.

First define the clipped binary time series:

$$X_t = \begin{cases} 1, & \text{if } Z_t \geq 0 \\ 0, & \text{if } Z_t < 0 \end{cases}$$

The number of zero-crossings, denoted by D, is defined in terms of $\{X_t\}$,

$$D = \sum_{t=2}^{N} [X_t - X_{t-1}]^2 \qquad (17)$$

$$0 \leq D \leq N - 1$$

Example:

Z:	-3	-4	6	7	8	-8	9	7	-1	2
X:	0	0	1	1	1	0	1	1	0	1

$N = 10$, $D = 5$.

(•) *Discovering the Cosine Formula*

$\{Z_t\}$ a stationary Gaussian AR(1) time series:

$$Z_t = \phi Z_{t-1} + \epsilon_t, \quad t = 1, 2, ..., N$$

$|\phi| < 1$, and $\{\epsilon_t\}$ iid N(0,σ^2).

Clipped binary time series:

$$X_t = \begin{cases} 1, & \text{if } Z_t \geq 0 \\ 0, & \text{if } Z_t < 0 \end{cases}$$

Define

$$\lambda_1 \equiv P(X_t = 1 | X_{t-1} = 1)$$

Then

$$\phi = -\cos(\pi \lambda_1)$$

"Assume" that X_t is a 0-1 Markov chain.

Define

$$s = \sum_{t=1}^{n} x_t, \ r = \sum_{t=2}^{n} x_t x_{t-1}, \ h = x_1 + x_N$$

Then

$$\begin{aligned}\log L(\lambda_1) &= [2r-2s+h+(N-1)]\log(\lambda_1) \\ &+ [2s-2r-h]\log(1-\lambda_1)-\log(2)\end{aligned}$$

and this is maximized for

$$\widehat{\lambda}_1 = \frac{2r-2s+h+(N-1)}{N-1}$$

But a closer look shows that

$$D = 2s-2r-h$$

is the *number of symbol changes* in $X_1, X_2, \cdots X_N$, or the number of "ZC" in $Z_1, Z_2, ..., Z_N$,

so that in fact

$$\widehat{\lambda}_1 = 1 - \frac{D}{N-1}$$

Hence:

$$\widehat{\phi} = -\cos\left(\pi - \frac{\pi D}{N-1}\right) = \cos\left(\frac{\pi D}{N-1}\right)$$

($\bullet$) *Cosine Formula*

This holds in general for any stationary Gaussian process $\{Z_t\}$.

Hence, there is an explicit formula connecting ρ_1 and $E[D]$,

$$(\star) \quad \rho_1 = \cos\left(\frac{\pi E[D]}{N-1}\right) \tag{18}$$

An inverse relationship:

$$E(D) \to 0 \Longleftrightarrow \rho_1 \to 1$$

$$E(D) \to N-1 \Longleftrightarrow \rho_1 \to -1$$

(•) *zero-crossing spectral representation*

$$\cos\left(\frac{\pi E[D]}{N-1}\right) = \frac{\int_{-\pi}^{-\pi} \cos(\omega) dF(\omega)}{\int_{-\pi}^{-\pi} dF(\omega)} \qquad (19)$$

Assume F is continuous at the origin,

$$\cos\left(\frac{\pi E[D]}{N-1}\right) = \frac{\int_{0}^{\pi} \cos(\omega) dF(\omega)}{\int_{0}^{\pi} dF(\omega)} \qquad (20)$$

(•) **Dominant Frequency Principle**: For $\omega_0 \in (0, \pi)$

$$\begin{aligned} F(\omega+) - F(\omega-) &> 0, \ \omega = \omega_0 \\ &= 0, \ \omega \neq \omega_0 \end{aligned}$$

then

$$\cos\left(\frac{\pi E[D]}{N-1}\right) = \cos(\omega_0)$$

or, by the monotonicity of $\cos(x)$ in $[0, \pi]$,

$$\frac{\pi E[D]}{N-1} = \omega_0$$

(•) *Higher Order Crossings (HOC)*

1.
$$\{Z_t\},\ t = 0, \pm 1, \pm 2, \cdots$$

2.
$$\{\mathcal{L}_\theta(\cdot),\ \theta \in \Theta\}$$

3.
$$\mathcal{L}_\theta(Z)_1, \mathcal{L}_\theta(Z)_2, \cdots, \mathcal{L}_\theta(Z)_N$$

4.
$$X_t(\theta) = \begin{cases} 1, & \text{if } \mathcal{L}_\theta(Z)_t \geq 0 \\ 0, & \text{if } \mathcal{L}_\theta(Z)_t < 0 \end{cases}$$

5. HOC $\{D_\theta,\ \theta \in \Theta\}$:
$$D_\theta = \sum_{t=2}^{N} [X_t(\theta) - X_{t-1}(\theta)]^2$$

HOC Combines ZC counts and linear operations (filters.)

Connection Between ZC and Filtering

$H(\omega;\theta)$ the transfer function corresponding to $\mathcal{L}_\theta(\cdot)$, and assume $\{Z_t\}$ is a zero-mean stationary Gaussian process.

$$\rho_1(\theta) \equiv \cos\left(\frac{\pi E[D_\theta]}{N-1}\right) = \frac{\int_{-\pi}^{\pi} \cos(\omega)|H(\omega;\theta)|^2 dF(\omega)}{\int_{-\pi}^{\pi} |H(\omega;\theta)|^2 dF(\omega)} \quad (21)$$

Assuming F is continuous at 0,

$$\rho_1(\theta) \equiv \cos\left(\frac{\pi E[D_\theta]}{N-1}\right) = \frac{\int_{0}^{\pi} \cos(\omega)|H(\omega;\theta)|^2 dF(\omega)}{\int_{0}^{\pi} |H(\omega;\theta)|^2 dF(\omega)} \quad (22)$$

The representation (21) and (22) help to understand the effect of filtering on zero-crossings through the spectrum even in the general non-Gaussian case.

HOC now refers to both $\rho_1(\theta)$ and D_θ.

Note: $\rho_1(\theta)$ can be defined directly and in general. There is no need for the Gaussian assumption.

$$\rho_1(\theta) \equiv \frac{\int_{-\pi}^{\pi} \cos(\omega)|H(\omega;\theta)|^2 dF(\omega)}{\int_{-\pi}^{\pi} |H(\omega;\theta)|^2 dF(\omega)} \quad (23)$$

HOC From Differences

$$\bigtriangledown Z_t \equiv Z_t - Z_{t-1}$$

and define

$$\mathcal{L}_j \equiv \bigtriangledown^{j-1}, \quad j \in \{1, 2, 3, \cdots\}$$

with $\mathcal{L}_1 \equiv \bigtriangledown^0$ being the identity filter. The corresponding HOC

$$D_1, D_2, D_3, \cdots$$

are called the *simple* HOC.

Thus,

D_1	from ZC of Z_t
D_2	from ZC of $(\bigtriangledown Z)_t$
D_3	from ZC of $(\bigtriangledown^2 Z)_t$
D_4	from ZC of $(\bigtriangledown^3 Z)_t$
.	.
.	.
.	.

Properties of Simple HOC:

(a) *Monotonicity:*

$$D_j - 1 \le D_{j+1}$$

which implies under strict stationarity,

$$0 \le E[D_1] \le E[D_2] \le E[D_3] \le \cdots \le N - 1$$

Example: $E[D_k]$ of Gaussian White Noise. $N = 1000$:

$$E[D_k] = (N-1)\{\frac{1}{2} + \frac{1}{\pi}\sin^{-1}(\frac{k-1}{k})\}$$

k	$E[D_k]$
1	499.50
2	666.00
3	731.55
4	769.18
5	794.37
6	812.76
7	826.93
8	838.30
9	847.67
10	855.58

Problem: Thus, $\{E[D_j]\}$ is a monotone bounded sequence. What does it converge to as $j \to \infty$???

Problem: What happens when $E[D_1] = E[D_2]$???

Problem: As $j \to \infty$, $\{X_t(j)\} \Rightarrow$???

Problem: As $N \to \infty$, $\frac{D_1}{N} \to$ Constant ???

(b) <u>*K (1984)*</u>: If $\{Z_t\}$ is Gaussian, then with prob. 1

$$\frac{E[D_1]}{N-1} = \frac{E[D_2]}{N-1} \Longleftrightarrow Z_t = A\cos(\omega_0 t + \varphi)$$

$$\omega_0 = \pi E[D_1]/(N-1).$$

(c) Suppose $\{Z_t\}$ is Gaussian, and let ω^* be the highest positive frequency in the spectral support

$$\omega \in [0, \omega^*].$$

Then, regardless of spectrum type,

$$\frac{\pi E[D_j]}{N-1} \to \omega^*, \quad j \to \infty \tag{24}$$

(d) Suppose $\{Z_t\}$ is Gaussian.

$$\frac{D_1}{N} \to \text{Constant ???}$$

<u>*K–Slud (1994):*</u>
($\star$) Yes in the continuous spectrum case.
($\star$) Sometime if the spectrum contains 1 jump.
($\star$) No if the spectrum contains 2 or more jumps.

Concerning "No if the spectrum contains 2 or more jumps":

Write the ZC rate as:

$$\hat{\gamma} = \frac{1}{N-1} \sum_{t=2}^{N} d_t$$

Let σ be the normalized spectral measure of $\{Z_t\}$.
Let σ_d be the spectral measure of $\{d_t\}$.

Observe that

$$\lim_{N \to \infty} Var[\hat{\gamma}] = \sigma_d(\{0\})$$

We need to determine if $\sigma_d(\{0\})$ vanishes.

When the spectrum is of a mixed type, $\sigma_d(\{0\})$ is in general positive, and therefore the asymptotic variance does not vanish, except for some particular cases. This means that as an estimator of its expected value, $\hat{\gamma}$ need not be a consistent estimator under the Gaussian assumption.

(K- Slud 1994, Slud 1991). Let $\{Z_t;\ t = 0, \pm 1, \pm 2, \cdots\}$ be a real-valued zero-mean stationary Gaussian process with normalized spectral measure σ. The discrete spectral masses corresponding to the points of jump in the spectrum of $\{d_t\}$ are given by

$$\sigma_d(\{\omega\}) = \frac{1}{\pi^2} \sum_{m=1}^{\infty} \frac{1}{(2m)!} E\left\{ \left(\sum_{j=0}^{2m} S_{2m}^j C_{m,j}(\rho_1) \right)^2 I_{[\Lambda_1 + \cdots + \Lambda_{2m} = \omega]} \right\} \tag{25}$$

where the random variables Λ_j are independent with distribution σ on $[-\pi, \pi]$, and addition is modulo 2π; where $\{C_{m,j}(\cdot)\}$ denotes a special family of hypergeometric functions determined for $0 \le j \le 2m$ by

$$C_{m,j}(x) = C_{m,2m-j}(x)$$

$$C_{m+1,j}(x) = 2(m-j)C_{m,j}(x) + C_{m+1,j+2}(x)$$

$$C_{m,1}(x) = -\frac{(2m-2)!}{(m-1)!2^{m-1}}(1-x^2)^{-m+1/2}$$

$$C_{m,2}(x) = \frac{(2m-2)!}{(m-1)!2^{m-1}}x(1-x^2)^{-m+1/2} = -xC_{m,1}(x)$$

and where S_{2m}^j is a function of $\Lambda_1, \cdots, \Lambda_{2m}$, defined by

$$S_{2m}^0 = 1$$

$$S_{2m}^j = \sum \exp[i(\Lambda_{k_1} + \cdots + \Lambda_{k_j})], \qquad 1 \le j \le 2m$$

and the summation is over all distinct j-element subsets $\{k_1, \cdots, k_j\} \subset \{1, \cdots, 2m\}$.

K- Slud 1994: Let $\{Z_t;\ t = 0, \pm 1, \pm 2, \cdots\}$ be a real-valued zero-mean stationary Gaussian process with normalized spectral measure σ. If σ has at least two points of jump (atoms) in $[0, \pi]$, then $\sigma_d(\{0\}) > 0$. That is, the spectrum of $\{d_t\}$ has a jump at 0.

That is, with two or more jumps in the spectrum

$$\frac{D_1}{N} \nrightarrow \text{Constant}$$

(e) *K–Slud (1982): Higher Order Crossings Theorem*

$\{Z_t\}$, $t = 0, \pm 1, \cdots$, be a zero-mean stationary process, and assume that π is included in the spectral support. Define

$$X_t(j) = \begin{cases} 1, & \text{if } \nabla^{j-1} Z_t \geq 0 \\ 0, & \text{if } \nabla^{j-1} Z_t < 0 \end{cases}$$

Then,

(i)

$$\{X_t(j)\} \Rightarrow \begin{cases} \cdots 01010101 \cdots, & \text{wp } 1/2 \\ \cdots 10101010 \cdots, & \text{wp } 1/2 \end{cases}$$

as $j \to \infty$.

(ii) $\lim_{j\to\infty} \lim_{N\to\infty} \frac{D_j}{N-1} = 1$, wp 1.

Demonstration of the HOC Theorem using AR(1) with parameter $\phi = 0.0, 0.8$. $N = 1000$. 15 inferences.

j	$X_t(j)$	D_j	$X_t(j)$	D_j
1	100100000111111	207	011101010111111	503
2	001100110111111	515	011001010101101	617
3	001100110101101	659	010001010101101	695
4	011001100101101	715	010111010101101	729
5	011001001101001	745	010101010101001	743
6	010011001101001	773	010101010101001	761
7	010011011101011	807	010101010101011	781
8	010110011001011	823	010101010101010	795
9	010110010001010	829	010101010101010	813
10	010100110101010	849	010101010101010	821
11	010100110101010	855	010101010101010	827
12	010101110101010	865	110101010101010	831
13	010101000101010	875	110101010101010	837
14	010101010101010	883	100101010101010	841
15	010101010101010	885	100101010101010	843
16	010101010101010	893	101101010101010	849
	$\phi = 0.8$		$\phi = 0.0$ (WN)	

The ...010101010101... state is approached quite fast, and

$$D_1 < D_2 < D_3 < \cdots < D_{16}$$

Only the first few D_k's are useful in discrimination between processes.

(f) For a zero-mean stationary Gaussian process, the sequence of expected simple HOC $\{E[D_k]\}$ determines the spectrum up to a constant:

$$\{E[D_k]\} \Leftrightarrow \{\rho_k\} \Leftrightarrow \overline{F(\omega)} \tag{26}$$

(g) <u>*K (1980):*</u> Let $\{Z_t\}$ be a zero mean stationary Gaussian process with acf ρ_j. If $\sum_{j=-\infty}^{\infty} |\rho_j| < \infty$, then

$$\sum_{j=-\infty}^{\infty} |\kappa_x(1, -j, 1-j)| < \infty$$

and

$$\frac{D_1 - E[D_1]}{\sqrt{N}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_1^2), \quad N \to \infty$$

where

$$\sigma_1^2 =$$

$$\frac{1}{\pi^2} \sum_{j=-\infty}^{\infty} \left\{ (\sin^{-1} \rho_j)^2 + \sin^{-1} \rho_{j-1} \sin^{-1} \rho_{j+1} + 4\pi^2 \kappa_x(1, -j, 1-j) \right\}$$

(h) *Slud (1991):* Let $\{Z_t\}$, $t = 0, \pm 1, \cdots$, be a zero mean stationary Gaussian process with acf ρ_j, $Var[Z_0] = 1$, and square integrable spectral density f. Then

$$\frac{D_1 - E[D_1]}{\sqrt{N}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_1^2), \quad N \to \infty$$

where σ_1^2 satisfies,

$$\sigma_1^2 \geq \frac{4}{\pi(1 - \rho_1^2)} \int_{-\pi}^{\pi} |\rho_1 - \cos(\omega)|^2 f^2(\omega) d\omega > 0$$

Proof: Use Itô–Wiener calculus.

(i) **Problem:** Given two stationary processes with the *same* spectrum of which one is Gaussian and the other is not. Is it true that the Gaussian process has a higher expected ZCR ???

Answer: Assume continuous time. $\{Z(t)\}$, stationary Gaussian process, $Z(t) \sim N(0, 1)$, $-\infty < t < \infty$.

Consider the interval $[0, 1]$.

Divide $[0,1]$ into $N-1$ intervals of size Δ.

Define the sampled time series,

$$Z_k \equiv Z((k-1)\Delta),\ \ k = 1, 2, \cdots, N$$

Then

$$\rho_1 = \rho(\Delta) \tag{27}$$

From the cosine formula we obtain: <u>*Rice (1944) formula*</u>

$$\begin{aligned} (\star) \quad E[D_c] &\equiv \lim_{N\to\infty} E[D_1] = \lim_{\Delta\to 0} \frac{1}{\pi\Delta}\cos^{-1}(\rho(\Delta)) \\ &= \frac{1}{\pi}\sqrt{-\rho''(0)} \end{aligned} \tag{28}$$

<u>*Barnett–K (1998):*</u> There are non-Gaussian processes such that

$$(\star) \quad E[D_c] = \frac{\kappa}{\pi}\sqrt{-\rho''(0)}$$

with $\kappa < 1$ and $\kappa > 1$.

If $Z_1(t), Z_2(t)$ are independent copies of $Z(t)$, then the product $Z_1(t)Z_2(t)$ has $\kappa = \sqrt{2}$.

For $Z^3(t)$, $\kappa = \sqrt{5/9}$.

Application: Discrimination by Simple HOC

The ψ^2 Statistic:

When N is sufficiently large (e.g. $N \geq 200$), then with a high probability

$$0 < D_1 < D_2 < D_3 < \cdots < (N-1)$$

To capture the **rate of increase** in the first few D_k, consider the *increments*

$$\Delta_k \equiv \begin{cases} D_1, & \text{if } k = 1 \\ D_k - D_{k-1}, & \text{if } k = 2, \cdots, K-1 \\ (N-1) - D_{K-1}, & \text{if } k = K \end{cases}$$

Then

$$\sum_{k=1}^{K} \Delta_k = N - 1$$

Let $m_k = E[\Delta_k]$. We define a general similarity measure

$$\psi^2 \equiv \sum_{k=1}^{K} \frac{(\Delta_k - m_k)^2}{m_k} \tag{29}$$

When Δ_k, m_k are from the same process, and $K = 9$:

$$P(\psi^2 > 30) < 0.05$$

Application: Frequency Estimation

Recall the $AR(1)$ filter (α-filter),

$$Z_t(\alpha) = \mathcal{L}_\alpha(Z)_t = Z_t + \alpha Z_{t-1} + \alpha^2 Z_{t-2} + \cdots$$

with squared gain

$$|H(\omega;\alpha)|^2 = \frac{1}{1 - 2\alpha\cos(\omega) + \alpha^2}, \quad \alpha \in (-1,1),\ \omega \in [0,\pi].$$

Consider:

$$Z_t = A_1\cos(\omega_1 t) + B_1 \sin(\omega_1 t) + \zeta_t,\ t = 0, \pm 1, \cdots$$

$\omega_1 \in (0,\pi)$.
A_1, B_1 are uncorrelated $N(0,\sigma_1^2)$.
$\{\zeta_t\}$ $N(0,\sigma_\zeta^2)$ white noise independent of A_1, B_1.

Define:

$$C(\alpha) = \frac{Var(\zeta_t(\alpha))}{Var(Z_t(\alpha))}. \tag{30}$$

Then for $\alpha \in (-1,1)$,

$$0 < C(\alpha) < 1.$$

He–K (1989) Algorithm

Let $\{D_\alpha\}$ be the HOC from the AR(1) filter.
Fix $\alpha_1 \in (-1, 1)$. Define

$$(\star) \quad \alpha_{k+1} = \cos\left(\frac{\pi E[D_{\alpha_k}]}{N-1}\right), \; k = 1, 2, \cdots \tag{31}$$

Then, as $k \to \infty$,

$$\alpha_k \to \cos(\omega_1)$$

and

$$\frac{\pi E[D_{\alpha_k}]}{N-1} \to \omega_1 \tag{32}$$

Proof:

Note the **fundamental property** of the AR(1) filter gain:

$$(\star) \quad \alpha = \rho_{1,\zeta}(\alpha) = \frac{\int_{-\pi}^{\pi} \cos(\omega)|H(\omega;\alpha)|^2 d\omega}{\int_{-\pi}^{\pi} |H(\omega;\alpha)|^2 d\omega} \tag{33}$$

We have

$$\rho_1(\alpha) = \cos\left(\frac{\pi E[D_\alpha]}{N-1}\right)$$
$$= \frac{\sigma_1^2 |H(\omega_1;\alpha)|^2 \times \cos(\omega_1) + \int_{-\pi}^{\pi} |H(\omega;\alpha)|^2 dF_\zeta(\omega) \times \alpha}{\sigma_1^2 |H(\omega_1;\alpha)|^2 + \int_{-\pi}^{\pi} |H(\omega;\alpha)|^2 dF_\zeta(\omega)}$$

A weighted average of $\alpha^* \equiv \cos(\omega_1)$ and α (!)

Thus, we have a **contraction mapping**

$$(\star) \quad \rho_1(\alpha) = \alpha^* + C(\alpha)(\alpha - \alpha^*) \tag{34}$$

and the recursion (31) becomes,

$$(\star) \quad \alpha_{k+1} = \rho_1(\alpha_k) \tag{35}$$

with **fixed point** α^*:

$$\alpha^* = \rho_1(\alpha^*)$$

or

$$\cos(\omega_1) = \cos\left(\frac{\pi E[D_{\alpha^*}]}{N-1}\right)$$

By the monotonicity of $\cos(x),\ x \in [0, \pi]$,

$$\omega_1 = \frac{\pi E[D_{\alpha^*}]}{N-1} \qquad \triangle$$

Extension: Contractions From Bandpass Filters (CM)

Yakowitz (1991): We do not need the Gaussian assumption, and can speed up the contraction convergence.

1. Parametric family of band-pass filters indexed by $r \in (-1, 1)$, and by a bandwidth parameter M:

$$\{\mathcal{L}_{r,M}(\cdot),\ r \in (-1,1),\ M = 1, 2, \cdots\}$$

2. Let $h(n; r, M)$ and $H(\omega; r, M)$, be the corresponding complex impulse response and transfer function, respectively.

3. It is required that as $M \to \infty$, $|H(\omega; r, M)|^2$ converges to a Dirac delta function centered at $\theta(r) \equiv \cos^{-1}(r)$.

4. Assume further that the *filter passes only the (positive) discrete frequency to be detected*; suppose it is ω_1. Also, observe that

$$\Re\left\{\frac{E[\zeta_t(r,M)\overline{\zeta_{t-1}(r,M)}]}{E|\zeta_t(r,M)|^2}\right\} = \frac{\int_{-\pi}^{\pi}\cos(\omega)|H(\omega;r,M)|^2 dF_\zeta(\omega)}{\int_{-\pi}^{\pi}|H(\omega;r,M)|^2 dF_\zeta(\omega)}$$

where the overbar denotes "complex conjugate".

5. Suppose that for *any* M the *fundamental property* takes the form

$$r = \Re \left\{ \frac{E[\zeta_t(r,M)\overline{\zeta_{t-1}(r,M)}]}{E|\zeta_t(r,M)|^2} \right\} \tag{36}$$

6. Define,

$$\rho_1(r,M) \equiv \Re \left\{ \frac{E[Z_t(r,M)\overline{Z_{t-1}(r,M)}]}{E|Z_t(r,M)|^2} \right\} \tag{37}$$

7. Clearly,

$$\rho_1(r,M) = \frac{\frac{1}{2}\sigma_1^2|H(\omega_1;r,M)|^2 \times \cos(\omega_1) + \int_{-\pi}^{\pi} |H(\omega;r,M)|^2 dF_\zeta(\omega) \times r}{\frac{1}{2}\sigma_1^2|H(\omega_1;r,M)|^2 + \int_{-\pi}^{\pi} |H(\omega;r,M)|^2 dF_\zeta(\omega)}$$

8. Let

$$C(r,M) = \frac{E|\zeta_t(r,M)|^2}{E|Z_t(r,M)|^2}$$

9. The contraction has now an extra parameter M:

$$\rho_1(r,M) = r^* + C(r,M)(r - r^*) \tag{38}$$

where $r^* = \cos(\omega_1)$, and ω_1 is the true frequency.

10. The CM algorithm takes now the form

$$r_{k+1} = \rho_1(r_k, M_k) \tag{39}$$

CM With the Parametric AR(2) Filter in Practice

With $\alpha \in (-1, 1)$, $\eta \in (0, 1)$,

$$Y_t(\alpha) = (1 + \eta^2)\alpha Y_{t-1}(\alpha) - \eta^2 Y_{t-2}(\alpha) + Y_t \qquad (40)$$

Initial guess of ω_1: θ_0.
Start with $\alpha_0 = \cos(\theta_0)$.
Start with η close to 1, for example $\eta = 0.98$.
Increment of η: e.g. 0.0015.

Define the sample autocorrelation

$$\hat{\rho}_1(\alpha, \eta) = \frac{\sum_{t=1}^{N-1} Y_t(\alpha) Y_{t-1}(\alpha)}{\sum_{t=0}^{N-1} Y_t^2(\alpha)} \qquad (41)$$

Then the CM algorithm is given by,

$$(\star) \quad \alpha_{k+1} = \hat{\rho}_1(\alpha_k, \eta_k), \quad k = 0, 1, 2, \cdots \qquad (42)$$

where η_k increases with each iteration.

Li (1992), Song–Li (2000), Li–Song (2002):

$$Y_t = \beta \cos(\omega_1 t + \phi) + \epsilon_t \tag{43}$$

β is a positive constant.
$\omega_1 \in (-\pi, \pi)$.
$\phi \sim \text{Unif}(0, \pi]$
$\{\epsilon_t\}$ i.i.d. with mean 0 and variance σ_ϵ^2, independent of ϕ.

If $(1-\eta)^2 N \to 0$ as $N \to \infty$, then

$$(1-\eta)^{-1/2} N(\widehat{\omega}_1 - \omega_1) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \gamma^{-1})$$

$$\gamma = \frac{1}{2}\beta^2/\sigma_\epsilon^2$$

The implication of this is that by a judicious choice of η, the precision of the CM estimate can be made arbitrarily close to that achieved by periodogram maximization and nonlinear least squares.

More properties and discussions can be found in Li–K (1993a, 1993b, 1994, 1998), K (1994), Li–Song (2002).

S-Plus Code for the CM Algorithm

```
KY.AR2 <- function(z,theta0,eta,inc,niter){
y <- rep(0,length(z))
r <- rep(0,niter); OMEGA <- rep(0,niter)
r [1] <- cos(theta0); OMEGA[1] <- theta0
cat(c("Initial frequency guess is", OMEGA[1]),fill=T)
cat(c("eta","         r(k)","         Omega(k)",
"          Var(y)"), fill=T)
for(k in 2:niter){# eta increments by inc
eta <- eta+inc
if((eta < 0)|(eta >1))
  stop("eta must be between 0 and 1")
FiltCoeff <- c((1+eta^2)*r[k-1],-(eta^2))
y <- filter(z,FiltCoeff, "rec")
# CM Iterations-------------------
rrr <- acf(y)      # motif() must be on
r[k] <- rrr$acf[2] # Gives acf(1)!!!
# ------------------------------
OMEGA[k] <- acos(r[k])
cat(c(eta,r[k],OMEGA[k],var(y)),fill=T)}}
```

Example:

$$Y_t = 0.5\cos(0.513t + \phi_1) + \cos(0.771t + \phi_2) + 2.2\epsilon_t$$

$t = 1, ..., 1500.$
ϵ_t i.i.d. $\mathcal{N}(0, 1)$.
SNR $= 10\log_{10}((.5^2/2 + 1^2/2)/2.2^2) = -8.890$

Starting at $\theta_0 = 0.48$, $\eta = 0.98$. Increment of η 0.0015.
Final estimate is $\widehat{\omega} = 0.5135$. Error: 0.0005.

η	$\alpha(k)$	$\omega(k)$	Var$(Y_t(\alpha))$
0.9815	0.8807	0.4932	425.958
0.9830	0.8755	0.5042	574.342
0.9845	0.8723	0.5107	856.165
0.9860	0.8711	0.5131	1134.483
0.9875	0.8708	0.5138	1365.735
0.9890	0.8707	0.5139	1666.432
0.9905	0.8708	0.5138	2106.870
0.9920	0.8709	0.5136	2783.643
0.9935	0.8710	0.5135	3892.713

We are going to apply CM to the data without centering.

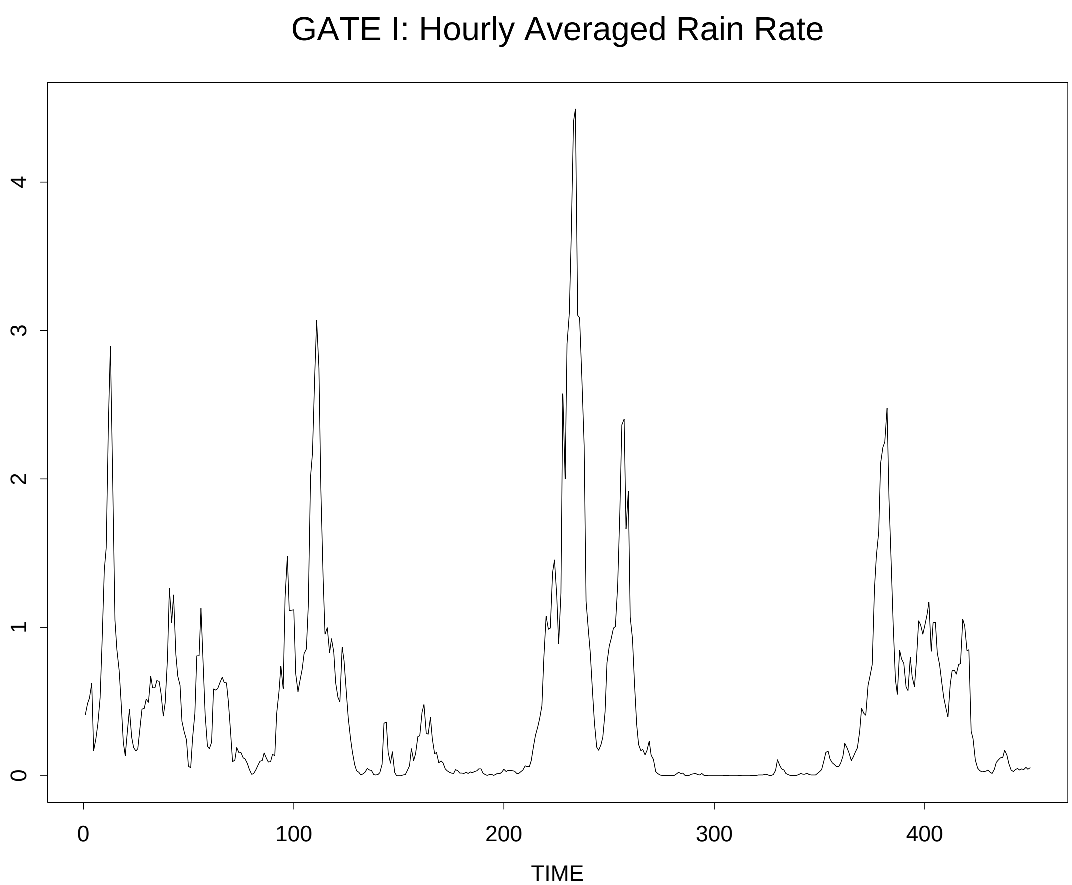
Starting at $\theta_0 = 0.88$, $\eta = 0.98$. Increment of η 0.001. Final estimate is $\hat{\omega} = 0.7709$. Error: 0.0001.

η	$\alpha(k)$	$\omega(k)$	$\mathrm{Var}(Y_t(\alpha))$
0.981	0.6518	0.8607	102.987
0.982	0.6672	0.8403	128.128
0.983	0.6822	0.8199	162.341
0.984	0.6973	0.7990	215.022
0.985	0.7104	0.7806	371.580
0.986	0.7162	0.7723	988.001
0.987	0.7171	0.7710	1555.238
0.988	0.7172	0.7708	1817.274
0.989	0.7172	0.7709	2130.545

Algorithm	$\hat{\omega}$	ω	Error
CM	0.5135	0.513	10^{-4}
FFT	0.5152		10^{-3}
CM	0.7709	0.771	10^{-4}
FFT	0.7749		10^{-3}

Example: Detection of a Diurnal Cycle in GATE I.

The CM algorithm with the AR(2) parametric filter was applied to a time series of length $N = 450$ of hourly rain rate from GATE I (early 1970s) averaged over a region of $280 \times 280\ km^2$.



Starting at 0.29 to the **right of 0.2617994:**
$\eta = 0.99$. Increment of η 0.0015.

η	$\alpha(k)$	$\omega(k)$	Var($Y_t(\alpha)$)
0.9915	0.962263	0.275596	554.29
0.9930	0.966709	0.258754	751.11
0.9945	0.968074	0.253367	2413.23
0.9960	0.966535	0.259434	3025.10
0.9975	0.967380	0.256120	4556.35
0.9990	0.966261	0.260501	7252.38

$$\frac{2\pi}{0.260501} = 24.11962$$

Starting at 0.25 to the **left of 0.2617994:**
$\eta = 0.995$. Increment of η 0.001.

η	$\alpha(k)$	$\omega(k)$	Var($Y_t(\alpha)$)
0.996	0.967167	0.256960	2353.12
0.997	0.967349	0.256243	4229.05
0.998	0.966722	0.258706	5453.12
0.999	0.967329	0.256323	7057.93
1.000	0.966053	0.261308	9989.52

$$\frac{2\pi}{0.261308} = 24.04513$$

References

Barnett and Kedem (1998). Zero-crossing rates of mixtures and products of Gaussian processes. *IEEE Tr. Information Theory*, 44, 1672-1677.

Hadjileontiadis, L.J. (2003). Discrimination analysis of discontinuous breath sounds using higher-order crossings. *Med. Biol. Eng. Comput.*, 41, 445-455.

He and Kedem (1989). Higher Order crossings Spectral Analysis of an Almost Periodic Random Sequence in Noise. *IEEE Tr. Information Theory*, 35, 360-370.

Kedem (1976). Exact Maximum Likelihood Estimation of the Parameter in the AR(1) Process After Hard Limiting. *IEEE Tr. Information Theory* IT-22, No. 4, 491-493.

Kedem (1984). On the Sinusoidal Limit of Stationary Time Series. *Annals of Statistics*, 12, 665-674.

Kedem (1986). Spectral Analysis and Discrimination by Zero-Crossings. *Proceedings of the IEEE*, 74, 1475-1493.

Kedem (1987). Detection of hidden periodicities by means of higher order crossings. *Jour. of Time Series Analysis*, 8, 39-50.

Kedem (1987). Higher Order Crossings in Time Series Model Identification. *Technometrics*, 19, 193-204.

Kedem (1992). Contraction mappings in mixed spectrum estimation. In *New Directions in Time Series Analysis*, D. Brillinger et al. eds., Springer-Verlag, NY, IMA Vol. 45, pp. 169-191

Kedem (1994). *Time Series Analysis by Higher Order Crossings*, IEEE Press, NJ. (•)

Kedem and Li (1991). Monotone gain, first-order autocorrelation, and expected zero-crossing rate. *Annals of Statistics*, 19, 1672-1676.

Kedem and Lopez (1992). Fixed points in mixed spectrum analysis. In *Probabilistic and Stochastic Methods in Analysis, With Applications*, J. S. Byrnes et al., eds., Kluwer, Mass., pp. 573-591.

Kedem and Slud (1981). On Goodness of Fit of Time Series Models: An Application of Higher Order Crossings. *Biometrika*, 68, 551-555.

Kedem and Slud (1982). Time Series Discrimination by Higher Order Crossings. *Annals of Statistics*, 10,786-794.

Kedem and Slud (1994). On autocorrelation estimation in mixed spectrum Gaussian processes. *Stochastic Processes and Their Applications*, 49, 227-244.

Kedem and Troendle (1994). An iterative filtering algorithm for non-Fourier frequency estimation. *Jour. of Time Series Analysis*, 15, 45-63.

Kedem and Yakowitz (1994). Practical aspects of a fast algorithm for frequency detection. *IEEE Tr. Communications*, 42, 2760-2767.

Kong and Qiu (2000). Injury detection for central nervous system via EEG with higher order crossing-based methods. *Method. Inform. Med.*, 39, 171-174.

Li (1997). Time series characterization, Poisson integral, and robust divergence measures. em Technometrics, 39, 357-371.

Li (1998). Time-correlation analysis of nonstationary time series. *Journal of Time Series Analysis*, 19, 47-64.

Li and Kedem (1993). Strong consistency of the contraction mapping method for frequency estimation. *IEEE Tr. Information Theory*, 39, 989-998.
Li and Kedem (1993). Asymptotic analysis of a multiple frequency estimation method. *Jour. Multivariate Analysis*, 46, 214-236.

Li and Kedem (1994). Iterative filtering for multiple frequency estimation. *IEEE Tr. Signal Processing*, 42, 1120-1132. Li and Kedem (1998). Tracking abrupt frequency changes. *Jour. of Time Series Analysis*, 19, 69-82.

Li and Song (2002). Asymptotic analysis of a contraction mapping algorithm for multiple frequency estimation. *IEEE Tr. Information Theory*, 48, No. 10.

Song and Li (2000). A statistically and computationally efficient method for frequency estimation. *Stochastic Processes and Their Applications*, 86, 29-47.

Yakowitz (1991). Some contributions to a frequency location method due to He and Kedem. *IEEE Tr. Information Theory*, 37, 1177-1182.