Stationary Stochastic Process

The behavior of a *stochastic process*, or simply a *process*, $z(t)$ on a domain $\mathcal{T}$ is characterized by the probability distributions of its finite dimensional restrictions $\big(z(t_1), \ldots, z(t_m)\big)$,

$$p\big(z(t_1), \ldots, z(t_m)\big),$$

for all $t_1, \ldots, t_m \in \mathcal{T}$.

A process is (strictly) stationary if

$$p\big(z(t_1), \ldots, z(t_m)\big) = p\big(z(t_1 + h), \ldots, z(t_m + h)\big),$$

for all $t_1, \ldots, t_m, h \in \mathcal{T}$.

In (discrete time) time series analysis, $\mathcal{T}$ is the integer lattice, so t_i and h will be integers. We will write $z(t)$ as z_t.

Moments of Stationary Process

For $m = 1$ with a stationary process, $p(z_t) = p(z)$ is the same for all t. Its *mean* and *variance* are

$$\mu = E[z_t] = \int z p(z) dz, \quad \sigma^2 = E\big[(z_t - \mu)^2\big] = \int (z - \mu)^2 p(z) dz.$$

The *autocovariance* of the process at *lag* k is

$$\gamma_k = \text{cov}[z_t, z_{t+k}] = E\big[(z_t - \mu)(z_{t+k} - \mu)\big].$$

The *autocorrelation* of the process is

$$\rho_k = \frac{E\big[(z_t - \mu)(z_{t+k} - \mu)\big]}{\sqrt{E\big[(z_t - \mu)^2\big] E\big[(z_{t+k} - \mu)^2\big]}} = \frac{\gamma_k}{\gamma_0},$$

where $\gamma_0 = \sigma^2$.

Autocovariance and Positive Definiteness

The covariance matrix of $(z_1, \ldots, z_n)$,

$$\begin{pmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{n-1} \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{n-2} \\ \vdots & \vdots & \ldots & \vdots \\ \gamma_{n-1} & \gamma_{n-2} & \cdots & \gamma_0 \end{pmatrix} = \sigma^2 \begin{pmatrix} \rho_0 & \rho_1 & \cdots & \rho_{n-1} \\ \rho_1 & \rho_0 & \cdots & \rho_{n-2} \\ \vdots & \vdots & \ldots & \vdots \\ \rho_{n-1} & \rho_{n-2} & \cdots & \rho_0 \end{pmatrix},$$

is positive definite. In general, a bivariate function $R(t,s)$ on $\mathcal{T}$ is nonnegative definite if for all l_i real and all $t_i \in \mathcal{T}$,

$$\textstyle\sum_{i,j} l_i l_j R(t_i, t_j) \geq 0.$$

For 1-D stationary process, the autocovariance $R(t,s) = R\big(|t-s|\big)$ is generated from some univariate function $R(h)$.

Weak Stationarity, Gaussian Process

A process is a *Gaussian process* if its restrictions $(z_{t_1}, \ldots, z_{t_m})$ follow normal distributions.

A process z_t on $\mathcal{T}$ is *weakly stationary* of second order if

$$E[z_t] = E[z_0] = \mu$$
$$\text{cov}[z_t, z_{t+h}] = \text{cov}[z_0, z_h] = \gamma_h,$$

for all $t, h \in \mathcal{T}$. A Gaussian process that is weakly stationary of second order is also strictly stationary.

For z_t stationary, the linear function with coefficients $l_1, \ldots, l_n$,

$$L_t = l_1 z_t + l_2 z_{t-1} + \cdots + l_n z_{t-n+1},$$

is stationary. These include $\nabla z_t = z_t - z_{t-1}$ and higher order differences $\nabla^d z_t$.

Examples: AR(1) and MA(1) Processes

Let a_t be independent with $E[a_t] = 0$ and $E[a_t^2] = \sigma_a^2$. The process a_t is called a *white noise* process.

Suppose z_t satisfies $z_t = \phi z_{t-1} + a_t$, a first order *autoregressive* (AR) process, with $|\phi| < 1$ and z_{t-1} independent of a_t. It is easy to verify that $E[z_t] = 0$ and

$$\gamma_0 = \sigma_a^2/(1-\phi^2), \qquad \rho_k = \phi\rho_{k-1}, \qquad \rho_k = \phi^{|k|}.$$

Let $z_t = a_t - \theta a_{t-1}$, a first order *moving average* (MA) process. It is easy to verify that $E[z_t] = 0$ and

$$\gamma_0 = \sigma_a^2(1+\theta^2), \qquad \gamma_1 = \sigma_a^2(-\theta), \qquad \gamma_k = 0, \quad k > 1.$$

Examples of Nonstationary Processes

Consider a random walk with drift, $z_t = z_{t-1} + \delta + a_t$, $t > 0$, with $z_0 = 0$, δ a constant, and a_t white noise. It is easy to calculate

$$E[z_t] = \delta t, \quad \text{var}[z_t] = t\sigma_a^2,$$

so z_t is nonstationary. The difference $w_t = z_t - z_{t-1}$ however is stationary.

Consider $z_t = \mu_t + y_t$, with μ_t a deterministic function and y_t a stationary process. $E[z_t] = \mu_t + E[y]$ depends on t, so z_t is nonstationary.

For $\mu_t = \delta t$, $w_t = z_t - z_{t-1}$ is stationary.

For $\mu_t = A\cos(2\pi t/k) + B\sin(2\pi t/k)$, $w_t = z_t - z_{t-k}$ is stationary.

Estimation of Mean

Observing $z_1, \ldots, z_N$, one estimates $\mu = E[z_t]$ using $\bar{z} = \sum_{t=1}^{N} z_t / N$, with the variance

$$\text{var}[\bar{z}] = \frac{1}{N^2} \sum_{t=1}^{N} \sum_{s=1}^{N} \gamma_{t-s} = \frac{\gamma_0}{N} \big[1 + 2 \sum_{k=1}^{N-1} (1 - \frac{k}{N}) \rho_k \big].$$

Assuming $\sum_k |\rho_k| < \infty$, it can be shown that as $N \to \infty$,

$$N \, \text{var}[\bar{z}] \to \gamma_0 \big(1 + 2 \sum_{k=1}^{\infty} \rho_k \big),$$

which yields the "large sample" variance $(\gamma_0/N)\big(1 + 2 \sum_{k=1}^{\infty} \rho_k\big)$. Compare with the familiar *i.i.d.* result $\text{var}[\bar{z}] = \sigma^2/N$, the effective sample size becomes $N/\big(1 + 2 \sum_{k=1}^{\infty} \rho_k\big)$ due to autocorrelation.

The factor $\big(1 + 2 \sum_{k=1}^{\infty} \rho_k\big)$ is 1 for white noise, $(1+\phi)/(1-\phi)$ for AR(1), and $(1-\theta)^2/(1+\theta^2)$ for MA(1).

Estimation of Autocorrelation Function

To estimate γ_k, one uses the *sample covariance function*

$$c_k = \frac{1}{N}\sum_{t=1}^{N-k}(z_t - \bar{z})(z_{t+k} - \bar{z}).$$

To estimate ρ_k, one uses the *sample autocorrelation function*

$$r_k = c_k/c_0.$$

Both $c_{|t-s|}$ and $r_{|t-s|}$ are nonnegative definite.

It can be shown that

$$E[c_k] \approx \gamma_k - \frac{k}{N}\gamma_k - \frac{1}{N}\gamma_0\big(1 + 2\sum_{v=1}^{\infty}\rho_v\big) = \gamma_k + O(N^{-1}),$$

and that

$$E[r_k] \approx \rho_k - \frac{k}{N}\rho_k - \frac{1}{N}\big(1 + 2\sum_{v=1}^{\infty}\rho_v\big) = \rho_k + O(N^{-1}),$$

so c_k and r_k are asymptotically unbiased estimates of γ_k and ρ_k.

Variance of Sample ACF

For Gaussian process with N large, it can be shown that

$$\text{cov}[c_k, c_{k+s}] \approx \frac{1}{N} \sum_{v=-\infty}^{\infty} (\gamma_v \gamma_{v+s} + \gamma_v \gamma_{v+2k+s}),$$

so $N \,\text{var}[c_k] \approx \sum_{v=-\infty}^{\infty} (\gamma_v^2 + \gamma_v \gamma_{v+2k})$. Similarly, one has

$$\text{cov}[r_k, r_{k+s}] \approx \frac{1}{N} \sum_{v=-\infty}^{\infty} (\rho_v \rho_{v+s} + \rho_v \rho_{v+2k+s} + 2\rho_k \rho_{k+s} \rho_v^2 - 2\rho_k \rho_v \rho_{v+k+s} - 2\rho_{k+s} \rho_v \rho_{v+k}),$$

and $N \,\text{var}[r_k] \approx \sum_{v=-\infty}^{\infty} (\rho_v^2 + \rho_v \rho_{v+2k} + 2\rho_k^2 \rho_v^2 - 4\rho_k \rho_v \rho_{v+k})$.

For k large, ρ_k often "dies out", leaving only the first terms contributing to the large-lag (co)variance.

Sample ACF in R

The `acf` function in the R library `ts` can be used to calculate and plot c_k and r_k, with standard errors superimposed for r_k. Part of the arguments and default options are listed below.

```
acf(x,lag.max=NULL,type=c("cor","cov","partial"),
    plot=TRUE,demean=TRUE,...)
plot.acf(acf.obj,ci=0.95,ci.col="blue",
    ci.type=c("white","ma"),...)
```

The default `lag.max` is $10\log_{10} N$. The estimated variances for r_k are $(1/N)$ for `ci.type="white"` (white noise model, with $\rho_v = 0$, $v > 0$), or $(1/N)(1 + 2\sum_{v=1}^{k-1} r_k^2)$ for `ci.type="ma"` (MA(k-1) model, with $\rho_v = 0$, $v > k-1$).

The option `type="partial"` concerns partial autocorrelation, to be discussed later.

Some Elementary Fourier Analysis

A complex number $a+ib$ can be written as $Ae^{i\theta} = A(\cos\theta + i\sin\theta)$, where $A = |a+ib| = \sqrt{a^2+b^2}$, $\cos\theta = a/A$, and $\sin\theta = b/A$. $|e^{i\theta}| = 1$.

For any $f(x)$ on $(-1/2, 1/2)$ satisfying $\int_{-1/2}^{1/2} \big|f(x)\big|^2 dx < \infty$, one has the *Fourier series expansion*,

$$f(x) = \sum_{v=-\infty}^{\infty} f_v e^{-i2\pi vx},$$

where $f_v = \int_{-1/2}^{1/2} f(x) e^{i2\pi vx} dx$ are the *Fourier coefficients*. The *Parseval's identity* asserts that $\sum_{v=-\infty}^{\infty} |f_v|^2 = \int_{-1/2}^{1/2} \big|f(x)\big|^2 dx$.

For vector $(z_1, \ldots, z_N)$, the *discrete Fourier transform* (DFT) is given by

$$\zeta_v = \frac{1}{\sqrt{N}} \sum_{t=1}^{N} z_t e^{-i2\pi tv/N}, \quad v = 1, \ldots, N,$$

and the inverse DFT is given by

$$z_t = \frac{1}{\sqrt{N}} \sum_{v=1}^{N} \zeta_v e^{i2\pi tv/N}, \quad t = 1, \ldots, N.$$

Spectral Density of Stationary Process

Herglotz's Theorem. A necessary and sufficient condition for ρ_k, $k = 0, \pm 1, \pm 2, \dots$ to be the autocorrelation function for some stationary process z_t is that there exists a probability function (cdf) $F(\omega)$ on $(-1/2, 1/2)$ such that

$$\rho_k = \int_{-1/2}^{1/2} e^{i2\pi k\omega} dF(\omega).$$

When $F(\omega)$ has a density $f(\omega)$, ρ_k are the Fourier coefficients of $f(\omega)$. The *spectral density* $f(\omega)$ has an expression

$$f(\omega) = \sum_{k=-\infty}^{\infty} \rho_k e^{-i2\pi k\omega}.$$

For z_t real with $\rho_k = \rho_{-k}$, one has

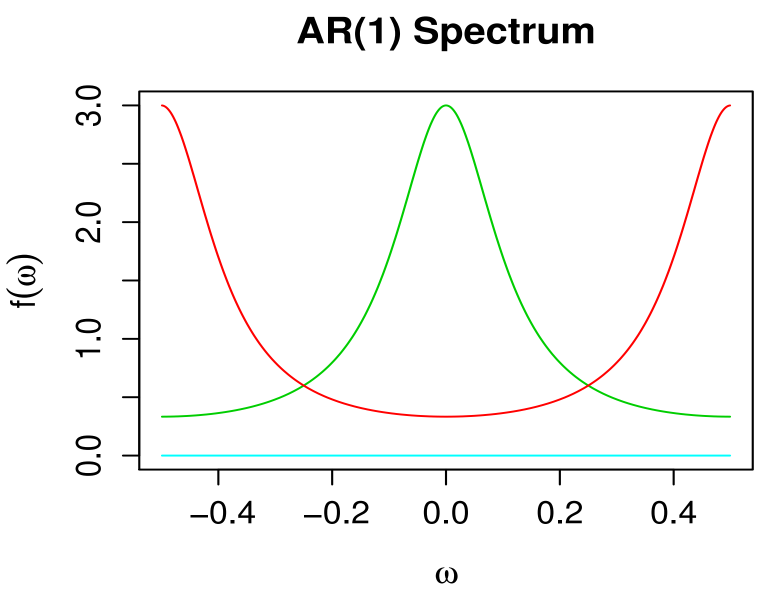
$$f(\omega) = 1 + 2 \sum_{k=1}^{\infty} \rho_k \cos 2\pi k\omega.$$

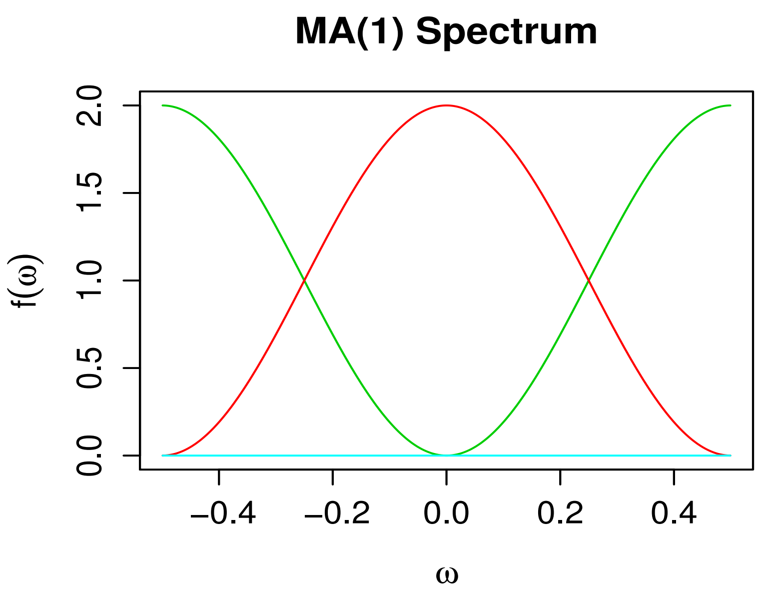
Examples of Spectral Density

For white noise, $f(\omega) = 1$ is uniform.

For AR(1), $f(\omega) = \dfrac{1-\phi^2}{1-2\phi\cos 2\pi\omega+\phi^2}$. Plots are with $\phi = \pm .5$.

For MA(1), $f(\omega) = 1 - 2\,\dfrac{\theta\cos 2\pi\omega}{1+\theta^2}$. Plots are with $\theta = \pm .7$.





Continuous and Discrete Spectrum

Consider $z_t = a_1 \cos 2\pi\lambda t + a_2 \sin 2\pi\lambda t$, where $a_1, a_2 \sim N(0, \sigma^2)$. One has $E[z_t] = 0$ and

$$\begin{aligned} \text{cov}[z_t, z_s] &= \sigma^2 \{\cos 2\pi\lambda t \cos 2\pi\lambda s + \sin 2\pi\lambda t \sin 2\pi\lambda s\} \\ &= \sigma^2 \cos 2\pi\lambda |t - s|, \end{aligned}$$

so z_t is stationary with $\rho_k = \cos 2\pi\lambda k$. The spectral distribution $F(\omega)$ is discrete with mass at $\omega = \pm\lambda$.

The above example shows that a discrete spectrum corresponds to a sinusoidal deterministic process, thus a purely random process should have a spectral density. In general, a stationary process may have both deterministic and purely random components.