Stationarity

from Peter Bartlett

$\{X_t\}$ is **strictly stationary** if
for all $k, t_1, \ldots, t_k, x_1, \ldots, x_k$, and h,

$$P(X_{t_1} \leq x_1, \ldots, X_{t_k} \leq x_k) = P(X_{t_1+h} \leq x_1, \ldots, X_{t_k+h} \leq x_k).$$

i.e., shifting the time axis does not affect the distribution.

We shall consider **second-order properties** only.

Mean and Autocovariance

Suppose that $\{X_t\}$ is a time series with $\mathrm{E}[X_t^2] < \infty$.

Its **mean function** is

$$\mu_t = \mathrm{E}[X_t].$$

Its **autocovariance function** is

$$\begin{aligned}\gamma_X(s,t) &= \mathrm{Cov}(X_s, X_t) \\ &= \mathrm{E}[(X_s - \mu_s)(X_t - \mu_t)].\end{aligned}$$

Weak Stationarity

We say that $\{X_t\}$ is **(weakly) stationary** if

1. μ_t is independent of t, and
2. For each h, $\gamma_X(t+h,t)$ is independent of t.

In that case, we write

$$\gamma_X(h) = \gamma_X(h,0).$$

Stationarity

The **autocorrelation function (ACF)** of $\{X_t\}$ is defined as

$$\begin{aligned}\rho_X(h) &= \frac{\gamma_X(h)}{\gamma_X(0)} \\ &= \frac{\text{Cov}(X_{t+h}, X_t)}{\text{Cov}(X_t, X_t)} \\ &= \text{Corr}(X_{t+h}, X_t).\end{aligned}$$

Stationarity

Example: i.i.d. noise, $\mathrm{E}[X_t] = 0, \mathrm{E}[X_t^2] = \sigma^2$. We have

$$\gamma_X(t+h,t) = \begin{cases} \sigma^2 & \text{if } h = 0, \\ 0 & \text{otherwise.} \end{cases}$$

Thus,

1. $\mu_t = 0$ is independent of t.
2. $\gamma_X(t+h,t) = \gamma_X(h,0)$ for all t.

So $\{X_t\}$ is stationary.

Similarly for any white noise (uncorrelated, zero mean), $X_t \sim WN(0, \sigma^2)$.

Stationarity

Example: Random walk, $S_t = \sum_{i=1}^{t} X_i$ for i.i.d., mean zero $\{X_t\}$.
We have $\mathrm{E}[S_t] = 0, \mathrm{E}[S_t^2] = t\sigma^2$, and

$$\begin{aligned}
\gamma_S(t+h,t) &= \mathrm{Cov}(S_{t+h}, S_t) \\
&= \mathrm{Cov}\left(S_t + \sum_{s=1}^{h} X_{t+s}, S_t\right) \\
&= \mathrm{Cov}(S_t, S_t) = t\sigma^2.
\end{aligned}$$

1. $\mu_t = 0$ is independent of t, but
2. $\gamma_S(t+h,t)$ is not.

So $\{S_t\}$ is not stationary.

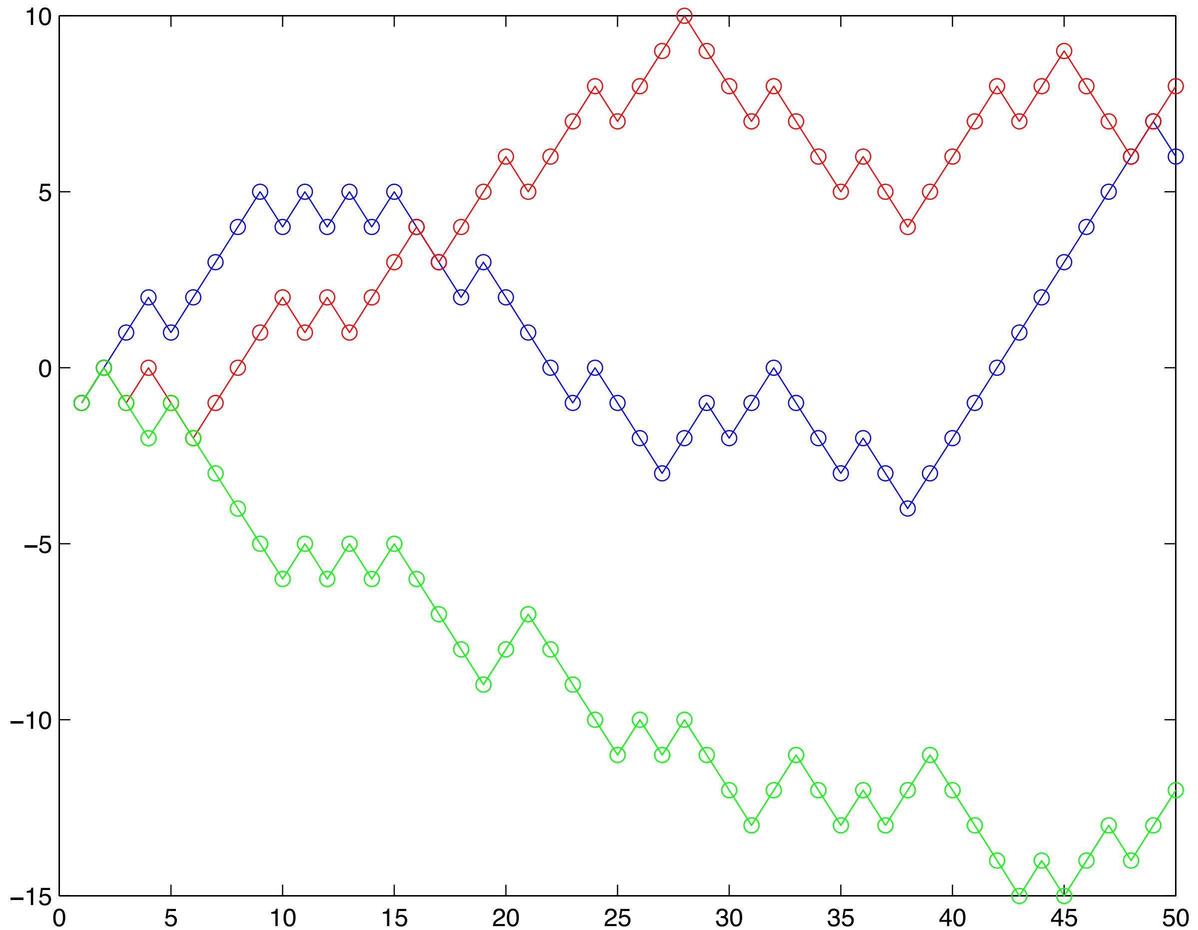
An aside: covariances

$$\mathrm{Cov}(X+Y,Z) = \mathrm{Cov}(X,Z) + \mathrm{Cov}(Y,Z),$$
$$\mathrm{Cov}(aX,Y) = a\,\mathrm{Cov}(X,Y),$$

Also if X and Y are independent (e.g., $X = c$), then

$$\mathrm{Cov}(X,Y) = 0.$$

Random walk



Stationarity

Example: MA(1) process (**Moving Average**):

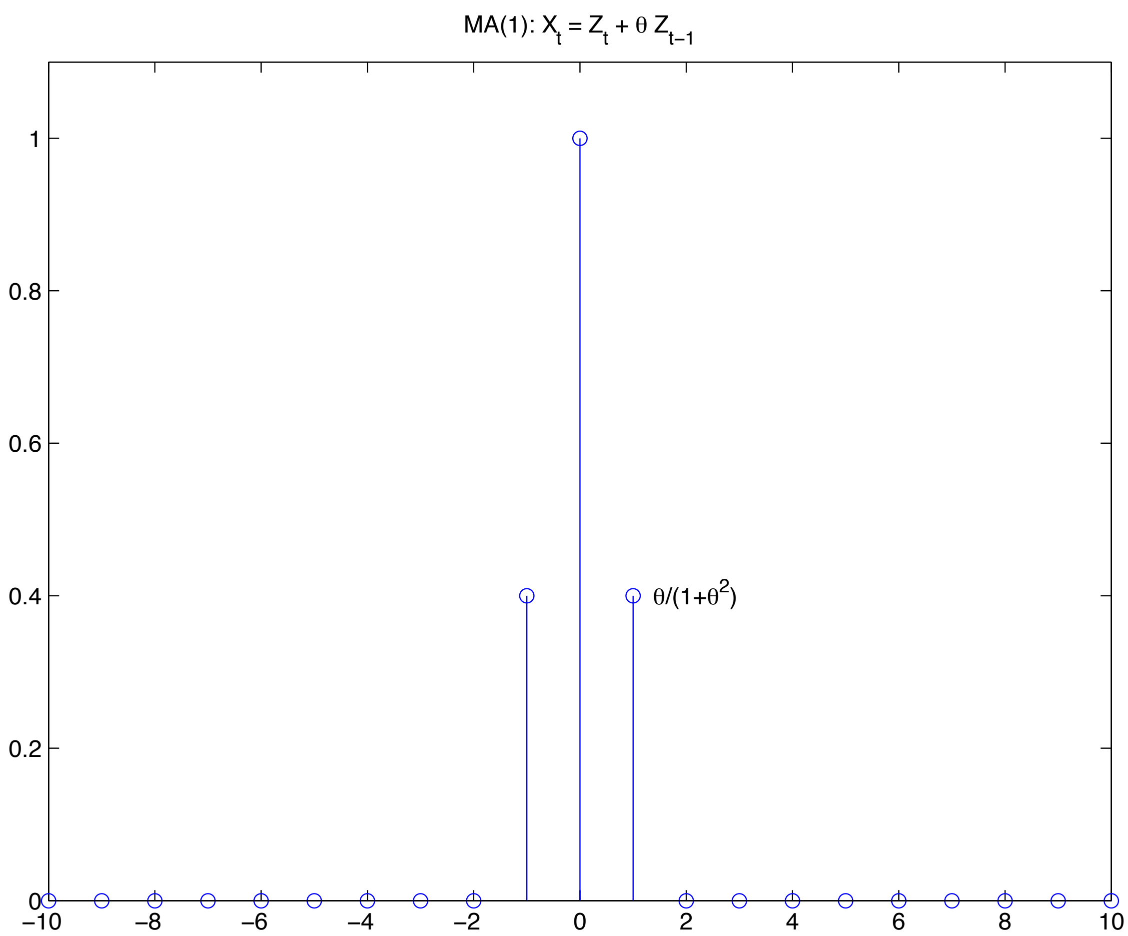
$$X_t = W_t + \theta W_{t-1}, \qquad \{W_t\} \sim WN(0, \sigma^2).$$

We have $\mathrm{E}[X_t] = 0$, and

$$\begin{aligned}
\gamma_X(t+h, t) &= \mathrm{E}(X_{t+h} X_t) \\
&= \mathrm{E}[(W_{t+h} + \theta W_{t+h-1})(W_t + \theta W_{t-1})] \\
&= \begin{cases} \sigma^2(1+\theta^2) & \text{if } h = 0, \\ \sigma^2 \theta & \text{if } h = \pm 1, \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}$$

Thus, $\{X_t\}$ is stationary.

ACF of the MA(1) process



Stationarity

Example: AR(1) process (**AutoRegressive**):

$$X_t = \phi X_{t-1} + W_t, \qquad \{W_t\} \sim WN(0, \sigma^2).$$

Assume that X_t is stationary and $|\phi| < 1$. Then we have

$$\begin{aligned} \mathrm{E}[X_t] &= \phi \mathrm{E} X_{t-1} \\ &= 0 \qquad \text{(from stationarity)} \\ \mathrm{E}[X_t^2] &= \phi^2 \mathrm{E}[X_{t-1}^2] + \sigma^2 \\ &= \frac{\sigma^2}{1 - \phi^2} \qquad \text{(from stationarity),} \end{aligned}$$

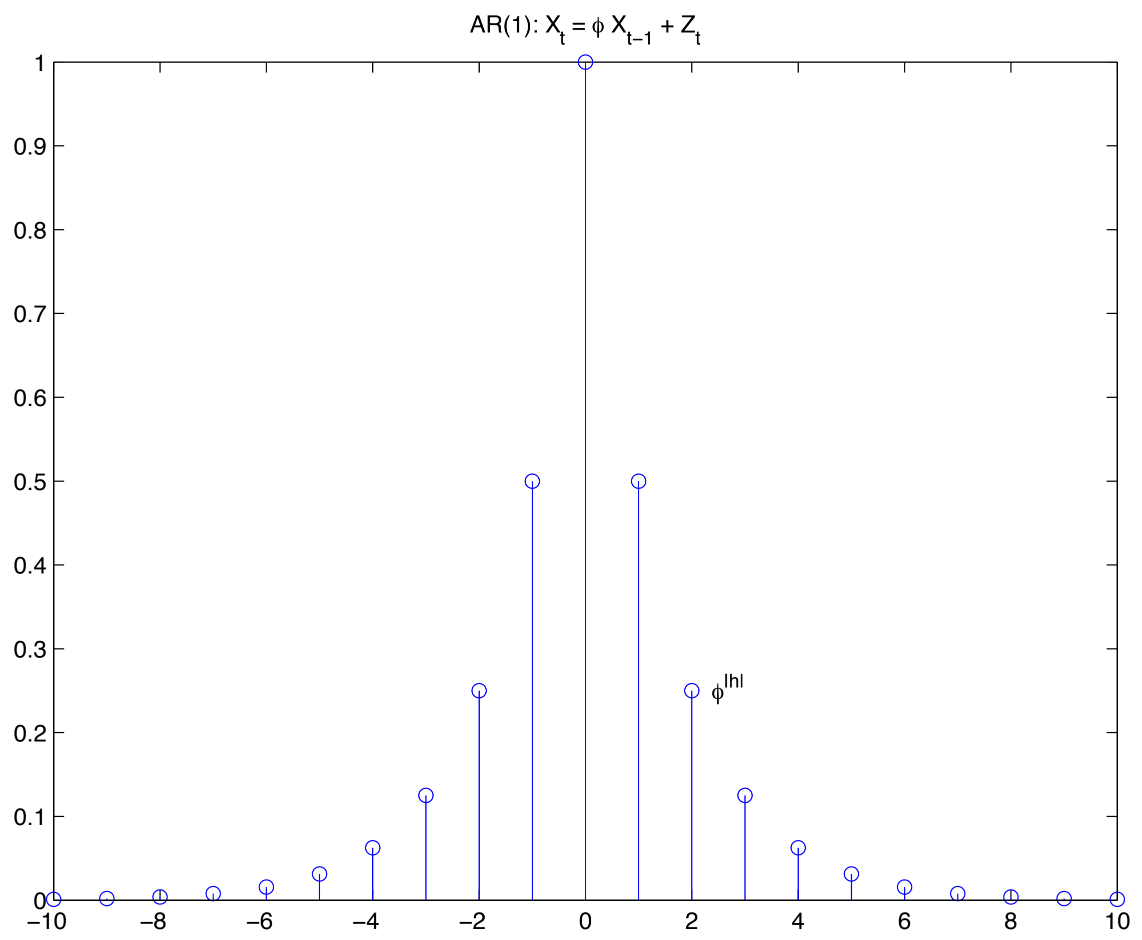
Stationarity

Example: AR(1) process, $X_t = \phi X_{t-1} + W_t, \qquad \{W_t\} \sim WN(0, \sigma^2)$.
Assume that X_t is stationary and $|\phi| < 1$. Then we have

$$\mathrm{E}[X_t] = 0, \qquad \mathrm{E}[X_t^2] = \frac{\sigma^2}{1-\phi^2}$$

$$\begin{aligned}
\gamma_X(h) &= \mathrm{Cov}(\phi X_{t+h-1} + W_{t+h}, X_t) \\
&= \phi \mathrm{Cov}(X_{t+h-1}, X_t) \\
&= \phi \gamma_X(h-1) \\
&= \phi^{|h|} \gamma_X(0) \qquad \text{(check for } h > 0 \text{ and } h < 0\text{)} \\
&= \frac{\phi^{|h|} \sigma^2}{1-\phi^2}.
\end{aligned}$$

ACF of the AR(1) process



Linear Processes

An important class of stationary time series:

$$X_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j W_{t-j}$$

where $\{W_t\} \sim WN(0, \sigma_w^2)$

and μ, ψ_j are parameters satisfying

$$\sum_{j=-\infty}^{\infty} |\psi_j| < \infty.$$

Linear Processes

$$X_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j W_{t-j}$$

We have

$$\mu_X = \mu$$

$$\gamma_X(h) = \sigma_w^2 \sum_{j=-\infty}^{\infty} \psi_j \psi_{h+j}. \qquad \text{(why?)}$$

Examples of Linear Processes: White noise

$$X_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j W_{t-j}$$

Choose μ,

$$\psi_j = \begin{cases} 1 & \text{if } j = 0, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\{X_t\} \sim WN(\mu, \sigma_W^2)$. (why?)

Examples of Linear Processes: MA(1)

$$X_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j W_{t-j}$$

Choose $\mu = 0$

$$\psi_j = \begin{cases} 1 & \text{if } j = 0, \\ \theta & \text{if } j = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Then $X_t = W_t + \theta W_{t-1}$. (why?)

Examples of Linear Processes: AR(1)

$$X_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j W_{t-j}$$

Choose $\mu = 0$

$$\psi_j = \begin{cases} \phi^j & \text{if } j \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Then for $|\phi| < 1$, we have $X_t = \phi X_{t-1} + W_t$. (why?)

Estimating the ACF: Sample ACF

Recall:

Suppose that $\{X_t\}$ is a stationary time series.

Its **mean** is

$$\mu = \mathrm{E}[X_t].$$

Its **autocovariance function** is

$$\begin{aligned}\gamma(h) &= \mathrm{Cov}(X_{t+h}, X_t) \\ &= \mathrm{E}[(X_{t+h} - \mu)(X_t - \mu)].\end{aligned}$$

Its **autocorrelation function** is

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)}.$$

Estimating the ACF: Sample ACF

For observations $x_1, \ldots, x_n$ of a time series,

the **sample mean** is $$\bar{x} = \frac{1}{n} \sum_{t=1}^{n} x_t.$$

The **sample autocovariance function** is

$$\hat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-|h|} (x_{t+|h|} - \bar{x})(x_t - \bar{x}), \qquad \text{for } -n < h < n.$$

The **sample autocorrelation function** is

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}.$$

Estimating the ACF: Sample ACF

Sample autocovariance function:

$$\hat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-|h|} (x_{t+|h|} - \bar{x})(x_t - \bar{x}).$$

$\approx$ the sample covariance of $(x_1, x_{h+1}), \ldots, (x_{n-h}, x_n)$, except that

- we normalize by n instead of $n - h$, and
- we subtract the full sample mean.