- Now we will introduce some useful parametric models for time series that are stationary processes.
- We begin by defining the *General Linear Process*.
- Let $\{Y_t\}$ be our observed time series and let $\{e_t\}$ be a white noise process (consisting of iid zero-mean r.v.'s).
- $\{Y_t\}$ is a general linear process if it can be represented by

$$Y_t = e_t + \psi_1 e_{t-1} + \psi_2 e_{t-2} + \cdots$$

where $e_t, e_{t-1}, \ldots$ are white noise.
- So this process is a weighted linear combination of present and past white noise terms.

More on General Linear Process

- When the number of terms is actually infinite, we need some regularity condition on the coefficients, such as $\sum_{i=1}^{\infty} \psi_i^2 < \infty$.
- Often we assume the weights are exponentially decaying, i.e.,

$$\psi_j = \phi^j$$

  where $-1 < \phi < 1$.
- Then $Y_t = e_t + \phi e_{t-1} + \phi^2 e_{t-2} + \cdots$

Properties of the General Linear Process

- Since the white noise terms all have mean zero, clearly $E(Y_t) = 0$ for all t.
- Also,

$$
\begin{aligned}
var(Y_t) &= var(e_t + \phi e_{t-1} + \phi^2 e_{t-2} + \cdots) \\
&= var(e_t) + \phi^2 var(e_{t-1}) + \phi^4 var(e_{t-2}) + \cdots \\
&= \sigma_e^2(1 + \phi^2 + \phi^4 + \cdots) \\
&= \frac{\sigma_e^2}{1 - \phi^2}
\end{aligned}
$$

More Properties of the General Linear Process

$$
\begin{aligned}
cov(Y_t, Y_{t-1}) &= cov(e_t + \phi e_{t-1} + \phi^2 e_{t-2} + \cdots, \\
&\quad e_{t-1} + \phi e_{t-2} + \phi^2 e_{t-3} + \cdots) \\
&= cov(\phi e_{t-1}, e_{t-1}) + cov(\phi^2 e_{t-2}, \phi e_{t-2}) + \cdots \\
&= \phi\sigma_e^2 + \phi^3\sigma_e^2 + \phi^5\sigma_e^2 + \cdots \\
&= \phi\sigma_e^2(1 + \phi^2 + \phi^4 + \cdots) \\
&= \frac{\phi\sigma_e^2}{1-\phi^2}
\end{aligned}
$$

- Hence $corr(Y_t, Y_{t-1}) = \phi$.
- Similarly, $cov(Y_t, Y_{t-k}) = \dfrac{\phi^k \sigma_e^2}{1-\phi^2}$ and $corr(Y_t, Y_{t-k}) = \phi^k$.

Stationarity of the General Linear Process

- Thus we see this process is stationary.
- The expected value is constant over time, and this covariance depends only only the lag k and not the actual time t.
- To obtain a process with some (constant) nonzero mean, we can just add some term μ to the definition.
- This does not affect the autocovariance structure, so such a process is still stationary.

Moving Average Processes

- This is a special case of the general linear process.
- A moving average process of order q, denoted $MA(q)$, is defined as:

$$Y_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}$$

- The simplest moving average process is the (first-order) $MA(1)$ process:

$$Y_t = e_t - \theta e_{t-1}$$

- Note that we do not need the subscript on the θ since there is only one of them in this model.

Properties of the $MA(1)$ Process

- It is easy to see that $E(Y_t) = 0$ and $var(Y_t) = \sigma_e^2(1 + \theta^2)$.
- Furthermore, $cov(Y_t, Y_{t-1}) =$ $cov(e_t - \theta e_{t-1}, e_{t-1} - \theta e_{t-2}) = cov(-\theta e_{t-1}, e_{t-1}) = -\theta\sigma_e^2$.
- And $cov(Y_t, Y_{t-2}) = cov(e_t - \theta e_{t-1}, e_{t-2} - \theta e_{t-3}) = 0$, since we see there are no subscripts in common.
- Similarly, $cov(Y_t, Y_{t-k}) = 0$ for any $k \geq 2$, so in the $MA(1)$ process, the observations farther apart than 1 time unit are uncorrelated.
- Clearly, $corr(Y_t, Y_{t-1}) = \rho_1 = -\theta/(1 + \theta^2)$ for the $MA(1)$ process.

Lag-1 Autocorrelation for the $MA(1)$ Process

- We see that the value of $corr(Y_t, Y_{t-1}) = \rho_1$ depends on what θ is.
- The largest value that ρ_1 can be is 0.5 (when $\theta = -1$) and the smallest value that ρ_1 can be is -0.5 (when $\theta = 1$).
- Some examples: When $\theta = 0.1, \rho_1 = -0.099$; when $\theta = 0.5, \rho_1 = -0.40$; when $\theta = 0.9, \rho_1 = -0.497$ (see R example plots).
- Just reverse the signs when θ is negative: When $\theta = -0.1, \rho_1 = 0.099$; when $\theta = -0.5, \rho_1 = 0.40$; when $\theta = -0.9, \rho_1 = 0.497$.
- Note that the lag-1 autocorrelation will be the same for the reciprocal of θ as for θ itself.
- Typically we will restrict attention to values of θ between -1 and 1 for reasons of *invertibility* (more later).

Second-order Moving Average Process

- A moving average of order 2, denoted $MA(2)$, is defined as:

$$Y_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2}$$

- It can be shown that, for the $MA(2)$ process,

$$\gamma_0 = var(Y_t) = (1 + \theta_1^2 + \theta_2^2)\sigma_e^2$$

$$\gamma_1 = cov(Y_t, Y_{t-1}) = (-\theta_1 + \theta_1\theta_2)\sigma_e^2$$

$$\gamma_2 = cov(Y_t, Y_{t-2}) = -\theta_2\sigma_e^2$$

Autocorrelations for Second-order Moving Average Process

- The autocorrelation formulas can be found in the usual way from the autocovariance and variance formulas.
- For the specific case when $\theta_1 = 1$ and $\theta_2 = -0.6$, $\rho_1 = -0.678$ and $\rho_2 = 0.254$.
- And $\rho_k = 0$ for $k = 3, 4, \ldots$.
- The strong negative lag-1 autocorrelation, weakly positive lag-2 autocorrelation, and zero lag-3 autocorrelation can be seen in plots of Y_t versus Y_{t-1}, Y_t versus Y_{t-2}, etc., from a simulated $MA(2)$ process.

Extension to $MA(q)$ Process

- For the moving average of order q, denoted $MA(q)$:

$$Y_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}$$

- It can be shown that

$$\gamma_0 = var(Y_t) = (1 + \theta_1^2 + \theta_2^2 + \cdots + \theta_q^2)\sigma_e^2$$

- The autocorrelations ρ_k are zero for $k > q$ and are quite flexible, depending on $\theta_1, \theta_2, \ldots, \theta_q$, for earlier lags when $k \leq q$.

Autoregressive Processes

- Autoregressive (AR) processes take the form of a regression of Y_t on itself, or more accurately on past values of the process:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + e_t,$$

  where e_t is independent of $Y_{t-1}, Y_{t-2}, \ldots, Y_{t-p}$.
- So the value of the process at time t is a linear combination of past values of the process, plus some independent "disturbance" or "innovation" term.

The First-order Autoregressive Process

- The $AR(1)$ process is (note we do not need a subscript on the ϕ here) a stationary process with:

$$Y_t = \phi Y_{t-1} + e_t,$$

- Without loss of generality, we can assume $E(Y_t) = 0$ (if not, we could replace Y_t with $Y_t - \mu$ everywhere).
- Note $\gamma_0 = var(Y_t) = \phi^2 var(Y_{t-1}) + var(e_t)$ so that $\gamma_0 = \phi^2\gamma_0 + \sigma_e^2$ and we see:

$$\gamma_0 = \frac{\sigma_e^2}{1 - \phi^2},$$

where $\phi^2 < 1 \Rightarrow |\phi| < 1$.

Autocovariances of the $AR(1)$ Process

- Multiplying the $AR(1)$ model equation by Y_{t-k} and taking expected values, we have:

$$\begin{aligned} E(Y_{t-k}Y_t) &= \phi E(Y_{t-k}Y_{t-1}) + E(e_t Y_{t-k}) \\ \Rightarrow \gamma_k &= \phi\gamma_{k-1} + E(e_t Y_{t-k}) \\ &= \phi\gamma_{k-1} \end{aligned}$$

  since e_t and Y_{t-k} are independent and (each) have mean 0.
- Since $\gamma_k = \phi\gamma_{k-1}$, then for $k = 1$, $\gamma_1 = \phi\gamma_0 = \phi\sigma_e^2/(1-\phi^2)$.
- For $k = 2$, we get $\gamma_2 = \phi\gamma_1 = \phi^2\sigma_e^2/(1-\phi^2)$.
- In general, $\gamma_k = \phi^k\sigma_e^2/(1-\phi^2)$.

Autocorrelations of the $AR(1)$ Process

- Since $\rho_k = \gamma_k/\gamma_0$, we see:

$$\rho_k = \phi^k, \text{ for } k = 1, 2, 3, \ldots$$

- Since $|\phi| < 1$, the autocorrelation gets closer to zero (weaker) as the number of lags increases.
- If $0 < \phi < 1$, all the autocorrelations are positive.
- Example: The correlation between Y_t and Y_{t-1} may be strong, but the correlation between Y_t and Y_{t-8} will be much weaker.
- So the value of the process is associated with very recent values much more than with values far in the past.

More on Autocorrelations of the $AR(1)$ Process

- If $-1 < \phi < 0$, the lag-1 autocorrelation is negative, and the signs of the autocorrelations alternate from positive to negative over the further lags.
- For ϕ near 1, the overall graph of the process will appear smooth, while for ϕ near -1, the overall graph of the process will appear jagged.
- See the R plots for examples.

The $AR(1)$ Model as a General Linear Process

- Recall that the $AR(1)$ model implies $Y_t = \phi Y_{t-1} + e_t$, and also that $Y_{t-1} = \phi Y_{t-2} + e_{t-1}$.
- Substituting, we have $Y_t = \phi(\phi Y_{t-2} + e_{t-1}) + e_t$, so that $Y_t = e_t + \phi e_{t-1} + \phi^2 Y_{t-2}$.
- Repeating this by substituting into the past "infinitely" often, we can represent this by:

$$Y_t = e_t + \phi e_{t-1} + \phi^2 e_{t-2} + \phi^3 e_{t-3} + \cdots$$

- This is in the form of the general linear process, with $\psi_j = \phi^j$ (we require that $|\phi| < 1$).

Stationarity

- For an $AR(1)$ process, it can be shown that the process is stationary if and only if $|\phi| < 1$.
- For an $AR(2)$ process, one following $Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + e_t$, we consider the AR *characteristic equation*:

$$1 - \phi_1 x - \phi_2 x^2 = 0.$$

- The $AR(2)$ process is stationary if and only if the solutions of the AR characteristic equation exceed 1 in absolute value, i.e., if and only if

$$\phi_1 + \phi_2 < 1, \phi_2 - \phi_1 < 1, \text{ and } |\phi_2| < 1.$$

Autocorrelations and Variance of the $AR(2)$ Process

- The formulas for the lag-k autocorrelation, ρ_k, and the variance $\gamma_0 = var(Y_t)$ of an $AR(2)$ process are complicated and depend on ϕ_1 and ϕ_2.
- The key things to note are:
  - the autocorrelation ρ_k dies out toward 0 as the lag k increases;
  - the autocorrelation function can have a wide variety of shapes, depending on the values of ϕ_1 and ϕ_2 (see R examples).

The General Autoregressive Process

- For an $AR(p)$ process:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + e_t$$

  the *AR characteristic equation* is:

$$1 - \phi_1 x - \phi_2 x^2 + \cdots + \phi_p x^p = 0.$$

- The $AR(p)$ process is stationary if and only if the solutions of the AR characteristic equation exceed 1 in absolute value.

Autocorrelations in the General AR Process

- If the process is stationary, we may form what are called the Yule-Walker equations:

$$\rho_1 = \phi_1 + \phi_2\rho_1 + \phi_3\rho_2 + \cdots + \phi_p\rho_{p-1}$$
$$\rho_2 = \phi_1\rho_1 + \phi_2 + \phi_3\rho_1 + \cdots + \phi_p\rho_{p-2}$$
$$\vdots$$
$$\rho_p = \phi_1\rho_{p-1} + \phi_2\rho_{p-2} + \phi_3\rho_{p-3} + \cdots + \phi_p$$

and solve numerically for the autocorrelations $\rho_1, \rho_2, \ldots, \rho_k$.

The Mixed Autoregressive Moving Average (ARMA) Model

- Consider a time series that has both autoregressive and moving average components:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}.$$

- This is called an *Autoregressive Moving Average* process of order p and q, or an $ARMA(p, q)$ process.

The $ARMA(1,1)$ Model

- The simplest type of $ARMA(p,q)$ model is the $ARMA(1,1)$ model:
$$Y_t = \phi Y_{t-1} + e_t - \theta e_{t-1}.$$
- The variance of a Y_t that follows the $ARMA(1,1)$ process is:
$$\gamma_0 = \frac{1 - 2\phi\theta + \theta^2}{1 - \phi^2}\sigma_e^2$$
- The autocorrelation function of the $ARMA(1,1)$ process is, for $k \geq 1$:
$$\rho_k = \frac{(1 - \theta\phi)(\phi - \theta)}{1 - 2\theta\phi + \theta^2}\phi^{k-1}$$

Autocorrelations of the $ARMA(1,1)$ Process

- The autocorrelation function ρ_k of an $ARMA(1,1)$ process decays toward 0 as k increases, with *damping factor* ϕ.
- Under the $AR(1)$ process, the decay started from $\rho_0 = 1$, but for the $ARMA(1,1)$ process, the decay starts from ρ_1, which depends on θ and ϕ.
- The shape of the autocorrelation function can vary, depending on the signs of ϕ and θ.

Other Properties of the $ARMA(1,1)$ and $ARMA(p,q)$ Processes

- The $ARMA(1,1)$ process (and the general $ARMA(p,q)$ process) can also be written as a general linear process.
- The $ARMA(1,1)$ process is stationary if and only if the solution to the AR characteristic equation $1 - \phi x = 0$ is greater than 1, i.e., if and only if $|\phi| < 1$.
- The $ARMA(p,q)$ process is stationary if and only if the solutions to the AR characteristic equation all exceed 1.
- The values of the autocorrelation function ρ_k for an $ARMA(p,q)$ process can be found by numerically solving a series of equations that depend on either $\phi_1, \ldots, \phi_p$ or $\theta_1, \ldots, \theta_q$.

Invertibility

- Recall that the $MA(1)$ process is *nonunique*: We get the same autocorrelation function if we replace θ by $1/\theta$.
- A similar nonuniqueness property holds for higher-order moving average models.
- We have seen that an AR process can be represented as an infinite-order MA process.
- Can an MA process be represented as an AR process?
- Note that in the $MA(1)$ process, $Y_t = e_t - \theta e_{t-1}$. So $e_t = Y_t + \theta e_{t-1}$, and similarly, $e_{t-1} = Y_{t-1} + \theta e_{t-2}$.
- So $e_t = Y_t + \theta(Y_{t-1} + \theta e_{t-2}) = Y_t + \theta Y_{t-1} + \theta^2 e_{t-2}$.
- We can continue this substitution "infinitely often" to obtain:

$$e_t = Y_t + \theta Y_{t-1} + \theta^2 Y_{t-2} + \cdots$$

More on Invertibility

- Rewriting, we get

$$Y_t = -\theta Y_{t-1} - \theta^2 Y_{t-2} - \cdots + e_t$$

- If $|\theta| < 1$, this $MA(1)$ model has been *inverted* into an infinite-order AR model.
- So the $MA(1)$ model is *invertible* if and only if $|\theta| < 1$.
- In general, the $MA(q)$ model is invertible if and only if the solutions of the *MA characteristic equation*

$$1 - \theta_1 x - \theta_2 x^2 - \cdots - \theta_q x^q = 0$$

all exceed 1 in absolute value.
- We see invertibility of MA models is similar to stationarity of AR models.

Invertibility and the Nonuniqueness Problem

- We can solve the nonuniqueness problem of MA processes by restricting attention only to invertible MA models.
- There is only one set of coefficient parameters that yield an *invertible* MA process with a particular autocorrelation function.
- Example: Both $Y_t = e_t + 2e_{t-1}$ and $Y_t = e_t + 0.5e_{t-1}$ have the same autocorrelation function.
- But of these two, only the second model is invertible (its solution to the MA characteristic equation is -2).
- For $ARMA(p, q)$ models, we restrict attention to those models which are both *stationary and invertible*.