Introduction to Time Series Analysis. Lecture 3.
Peter Bartlett

1. Review: Autocovariance, linear processes
2. Sample autocorrelation function
3. ACF and prediction
4. Properties of the ACF
A time series \( \{X_t\} \) has mean function \( \mu_t = \mathbb{E}[X_t] \) and autocovariance function

\[
\gamma_X(t + h, t) = \text{Cov}(X_{t+h}, X_t) \\
= \mathbb{E}[(X_{t+h} - \mu_{t+h})(X_t - \mu_t)].
\]

It is stationary if both are independent of \( t \).
Then we write \( \gamma_X(h) = \gamma_X(h, 0) \).
The autocorrelation function (ACF) is

\[
\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)} = \text{Corr}(X_{t+h}, X_t).
\]
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 $$\mu, \psi_j$$ are parameters satisfying

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

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

Examples:
- White noise: \( \psi_0 = 1 \).
- MA(1): \( \psi_0 = 1, \psi_1 = \theta \).
- AR(1): \( \psi_0 = 1, \psi_1 = \phi, \psi_2 = \phi^2, \ldots \)
Suppose that \( \{X_t\} \) is a stationary time series. Its mean is
\[
\mu = \mathbb{E}[X_t].
\]
Its autocovariance function is
\[
\gamma(h) = \text{Cov}(X_{t+h}, X_t) \\
= \mathbb{E}[(X_{t+h} - \mu)(X_t - \mu)].
\]
Its autocorrelation function is
\[
\rho(h) = \frac{\gamma(h)}{\gamma(0)}.
\]
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

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

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.
Sample ACF for white Gaussian (hence i.i.d.) noise
We can recognize the sample autocorrelation functions of many non-white (even non-stationary) time series.

<table>
<thead>
<tr>
<th>Time series</th>
<th>Sample ACF:</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>zero</td>
</tr>
<tr>
<td>Trend</td>
<td>Slow decay</td>
</tr>
<tr>
<td>Periodic</td>
<td>Periodic</td>
</tr>
<tr>
<td>MA(q)</td>
<td>Zero for $</td>
</tr>
<tr>
<td>AR(p)</td>
<td>Decays to zero exponentially</td>
</tr>
</tbody>
</table>
Sample ACF: Trend
Sample ACF: Trend

(why?)
Sample ACF

Time series: Sample ACF:

White zero
Trend Slow decay
Periodic Periodic
MA(q) Zero for $|h| > q$
AR(p) Decays to zero exponentially
Sample ACF: Periodic
Sample ACF: Periodic

![Sample ACF: Periodic Graph]

- The graph shows a periodic signal and signal plus noise.
- The x-axis represents the time index, ranging from 0 to 100.
- The y-axis represents the amplitude, ranging from -4 to 6.

Legend:
- Red: signal
- Blue: signal plus noise
Sample ACF

Time series: Sample ACF:

- White: zero
- Trend: Slow decay
- Periodic: Periodic
- MA(q): Zero for $|h| > q$
- AR(p): Decays to zero exponentially
ACF: MA(1)

MA(1): $X_t = Z_t + \theta Z_{t-1}$

$\theta/(1+\theta^2)$
Sample ACF: MA(1)
Sample ACF

Time series:

- White: zero
- Trend: Slow decay
- Periodic: Periodic
- MA(q): Zero for $|h| > q$
- AR(p): Decays to zero exponentially
ACF: AR(1)

AR(1): \( X_t = \phi X_{t-1} + Z_t \)
Sample ACF: AR(1)
Introduction to Time Series Analysis. Lecture 3.

1. Sample autocorrelation function
2. ACF and prediction
3. Properties of the ACF
ACF and prediction

- ACF and Sample ACF
- White noise and MA(1) comparison
ACF of a MA(1) process
ACF and least squares prediction

Best least squares estimate of $Y$ is $\mathbb{E}Y$:

$$\min_c \mathbb{E}(Y - c)^2 = \mathbb{E}(Y - \mathbb{E}Y)^2.$$  

Best least squares estimate of $Y$ given $X$ is $\mathbb{E}[Y|X]$:

$$\min_f \mathbb{E}(Y - f(X))^2 = \min_f \mathbb{E}[\mathbb{E}((Y - f(X))^2|X)]$$

$$= \mathbb{E}[\mathbb{E}[(Y - \mathbb{E}[Y|X])^2|X]]$$

$$= \text{var}[Y|X].$$

Similarly, the best least squares estimate of $X_{n+h}$ given $X_n$ is $f(X_n) = \mathbb{E}[X_{n+h}|X_n]$. 

25
Suppose that \( X = (X_1, \ldots, X_{n+h}) \) is jointly Gaussian:

\[
f_X(x) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp \left( -\frac{1}{2} (x - \mu)' \Sigma^{-1} (x - \mu) \right).
\]

Then the joint distribution of \((X_n, X_{n+h})\) is

\[
N \left( \begin{pmatrix} \mu_n \\ \mu_{n+h} \end{pmatrix}, \begin{pmatrix} \sigma_n^2 & \rho \sigma_n \sigma_{n+h} \\ \rho \sigma_n \sigma_{n+h} & \sigma_{n+h}^2 \end{pmatrix} \right),
\]

and the conditional distribution of \(X_{n+h}\) given \(X_n\) is

\[
N \left( \mu_{n+h} + \rho \frac{\sigma_{n+h}}{\sigma_n} (x_n - \mu_n), \sigma_{n+h}^2 (1 - \rho^2) \right).
\]
So for Gaussian and stationary $\{X_t\}$, the best estimate of $X_{n+h}$ given $X_n = x_n$ is

$$f(x_n) = \mu + \rho(h) (x_n - \mu),$$

and the mean squared error is

$$E(X_{n+h} - f(X_n))^2 = \sigma^2 (1 - \rho(h)^2).$$

Notice:

- Prediction accuracy improves as $|\rho(h)| \to 1$.
- Predictor is linear: $f(x) = \mu(1 - \rho(h)) + \rho(h)x$. 
Consider a linear predictor of $X_{n+h}$ given $X_n = x_n$. Assume first that \( \{X_t\} \) is stationary with $\text{E}X_n = 0$, and predict $X_{n+h}$ with $f(x_n) = ax_n$. The best linear predictor minimizes

$$
\text{E}(X_{n+h} - ax_n)^2 = \text{E}(X_{n+h}^2) - 2\text{E}(aX_{n+h}X_n) + \text{E}(a^2X_n^2)
$$

$$
= \sigma^2 - 2a\gamma(h) + a^2\sigma^2,
$$

and this is minimized when $a = \rho(h)$, that is,

$$
f(x_n) = \rho(h)X_n.
$$

For this optimal linear predictor, the mean squared error is

$$
\text{E}(X_{n+h} - f(X_n))^2 = \sigma^2 - 2\rho(h)\gamma(h) + \rho(h)^2\sigma^2
$$

$$
= \sigma^2(1 - \rho(h)^2).
$$
ACF and least squares linear prediction

Consider the following **linear predictor** of $X_{n+h}$ given $X_n = x_n$, when \{X_n\} is stationary and $EX_n = \mu$:

$$f(x_n) = a(x_n - \mu) + b.$$  

The linear predictor that minimizes

$$E(X_{n+h} - (a(X_n - \mu) + b))^2$$

has $a = \rho(h)$, $b = \mu$, that is,

$$f(x_n) = \rho(h)(X_n - \mu) + \mu.$$  

For this optimal linear predictor, the mean squared error is again

$$E(X_{n+h} - f(X_n))^2 = \sigma^2(1 - \rho(h)^2).$$
Least squares prediction of $X_{n+h}$ given $X_n$

\[ f(X_n) = \mu + \rho(h)(X_n - \mu). \]
\[ \text{E}(f(X_n) - X_{n+h})^2 = \sigma^2(1 - \rho(h)^2). \]

- If $\{X_t\}$ is stationary, $f$ is the \textbf{optimal linear predictor}.
- If $\{X_t\}$ is also Gaussian, $f$ is the \textbf{optimal predictor}.
- Linear prediction is optimal for Gaussian time series.
- Over all stationary processes with that value of $\rho(h)$ and $\sigma^2$, the optimal mean squared error is maximized by the Gaussian process.
- Linear prediction needs only second order statistics.
- Extends to longer histories, $(X_n, X_n - 1, \ldots)$. 
Introduction to Time Series Analysis. Lecture 3.

1. Sample autocorrelation function
2. ACF and prediction
3. Properties of the ACF
Properties of the autocovariance function

For the autocovariance function $\gamma$ of a stationary time series $\{X_t\}$,

1. $\gamma(0) \geq 0$, (variance is non-negative)
2. $|\gamma(h)| \leq \gamma(0)$, (from Cauchy-Schwarz)
3. $\gamma(h) = \gamma(-h)$, (from stationarity)
4. $\gamma$ is positive semidefinite.

Furthermore, any function $\gamma : \mathbb{Z} \to \mathbb{R}$ that satisfies (3) and (4) is the autocovariance of some stationary time series.
A function $f : \mathbb{Z} \rightarrow \mathbb{R}$ is *positive semidefinite* if for all $n$, the matrix $F_n$, with entries $(F_n)_{i,j} = f(i - j)$, is positive semidefinite.

A matrix $F_n \in \mathbb{R}^{n \times n}$ is positive semidefinite if, for all vectors $a \in \mathbb{R}^n$,

$$a' F a \geq 0.$$ 

To see that $\gamma$ is psd, consider the variance of $(X_1, \ldots, X_n)a$. 

---

**Properties of the autocovariance function**
Properties of the autocovariance function

For the autocovariance function $\gamma$ of a stationary time series $\{X_t\}$,

1. $\gamma(0) \geq 0$,
2. $|\gamma(h)| \leq \gamma(0)$,
3. $\gamma(h) = \gamma(-h)$,
4. $\gamma$ is positive semidefinite.

Furthermore, any function $\gamma : \mathbb{Z} \to \mathbb{R}$ that satisfies (3) and (4) is the autocovariance of some stationary time series (in particular, a Gaussian process).

e.g.: (1) and (2) follow from (4).
Introduction to Time Series Analysis. Lecture 3.

1. Sample autocorrelation function
2. ACF and prediction
3. Properties of the ACF