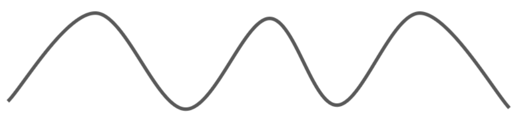
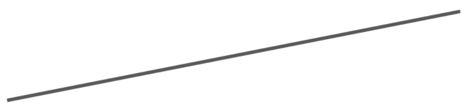
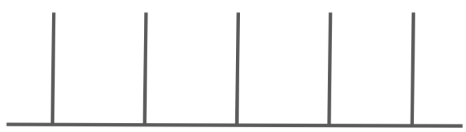
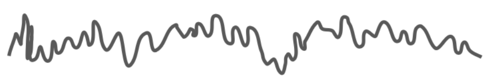
Alberto Andrés Valdés González.
Degree: Mathematical Engineer.
Work position: Data Scientist.
Mail: anvaldes@uc.cl/alberto.valdes.gonzalez.96@gmail.com
Location: Santiago, Chile.

Time Series Review

When we analyze a time serie, we can decompose mainly in 4 components:

	Stationary Pattern $S(t)$
	Trend $M(t)$
	Holiday Effect $H(t)$
	Noise $N(t)$

$$\boxed{X(t) = M(t) + S(t) + H(t) + N(t)}$$

Definition: We can see a time series for the observed data as a sequence of random variables $\{X_t\}$ (order matters).

White Noise: The more simple model of time series is which in what the variables are independent and identically distributed with mean 0.

Differentiation:

$$\boxed{Y_t = X_t - X_{(t-d)}}$$

Definition: Let be $\{X_t\}$ a time series with $\mathbb{E}[X_t^2] < \infty$. The mean of X_t is given by:

$$\boxed{\mu_X(t) = \mathbb{E}[X_t]}$$

And the covariance is defined by:

$$\boxed{Cov(X_r, X_s) = \gamma_X(r,s) = \mathbb{E}\left[(X_r - \mu_X(r)) \cdot (X_s - \mu_X(s))\right]}$$

Definition: A time series $\{X_t\}$ is stationary in the <u>second order</u> sense if:

(I) $\mu_X(t)$ is independent of t.

(II) $\gamma_X(t, t+k)$ is independent of t for all integer k.

Definition: A time series $\{X_t\}$ is <u>strictly stationary</u> if the random vectors $(X_1, ..., X_n)$ and $(X_k, ..., X_{(n+k)})$ have the same distribution for all integer k and $n > 0$.

$$\boxed{\text{Strictly stationary} \Rightarrow \text{2nd order stationary}}$$

Definition: Let be $\{X_t\}$ a time series. Its autocovariance function is:

$$\boxed{\gamma_X(k) = Cov(X_t, X_{(t+k)})}$$

And its autocorrelation function is:

$$\boxed{\rho_X(k) = \frac{\gamma_X(k)}{\gamma_X(0)} = corr(X_t, X_{(t+k)})}$$

Box-Cox Transformation: Box Cox transformation is a transformation of non-normal dependent variables into a normal shape.

$$\boxed{Y_t = \begin{cases} \dfrac{X_t^\lambda - 1}{\lambda} & \text{if } \lambda \neq 0 \\ ln(X_t) & \text{if } \lambda = 0 \end{cases}}$$

White noise:

$$\boxed{\gamma_X(k) = \gamma_X(t, t+k) = \begin{cases} \sigma^2 & \text{if } k = 0 \\ 0 & \text{if } k \neq 0 \end{cases}}$$

Example: Random Walk.

$$S_t = \sum_{i=1}^{t} X_i \text{ with } X_i \overset{\text{iid}}{\sim} D(0, \sigma^2)$$

$$\gamma_S(t, t+k) = Cov\left(S_t, S_{(t+k)}\right) = Cov\left(\sum_{i=1}^{t} X_i, \sum_{j=1}^{t+k} X_j\right) = \sum_{i=1}^{t}\sum_{j=1}^{t+k} Cov(X_i, X_j)$$

$$= \sum_{i=1}^{t}\left[\sum_{j=1}^{t} Cov(X_i, X_j) + \sum_{j=t+1}^{t+k} Cov(X_i, X_j)\right]$$

$$= \sum_{i=1}^{t}\sum_{j=1}^{t} Cov(X_i, X_j) + \sum_{i=1}^{t}\sum_{j=t+1}^{t+k} Cov(X_i, X_j)$$

$$= \sum_{i=1}^{t}\sum_{j=1}^{t} Cov(X_i, X_j) + \sum_{i=1}^{t}\sum_{j=t+1}^{t+k} 0 = \sum_{i=1}^{t}\sum_{j=1}^{t} Cov(X_i, X_j)$$

$$= \sum_{i=1:i\neq j}^{t}\sum_{j=1}^{t} Cov(X_i, X_j) + \sum_{i=1}^{t} Cov(X_i, X_i)$$

$$= \sum_{i=1:i\neq j}^{t}\sum_{j=1}^{t} 0 + \sum_{i=1}^{t} \sigma^2 = t \cdot \sigma^2$$

$\Rightarrow$

$$\boxed{\gamma_S(t, t+k) = t \cdot \sigma^2}$$

And for that the **Random Walk** isn't stationary.

Moving Average MA(1):

$$\boxed{X_t = Z_t + \theta \cdot Z_{(t-1)} \text{ with } t \in \mathbb{Z}}$$

Where $\{Z_t\} \overset{\text{iid}}{\sim} D(0, \sigma^2)$.

$$\boxed{E[X_t] = 0}$$

$$\boxed{\gamma_X(k) = \begin{cases} \sigma^2 \cdot (1+\theta^2) & \text{if } |k| = 0 \\ \sigma^2 \cdot \theta & \text{if } |k| = 1 \\ 0 & \text{if } |k| > 1 \end{cases}}$$

$$\boxed{\rho_X(k) = \begin{cases} 1 & \text{if } |k| = 0 \\ \dfrac{\theta}{1+\theta^2} & \text{if } |k| = 1 \\ 0 & \text{if } |k| > 1 \end{cases}}$$

Proof:

$$\mathbb{E}[X_t] = \mathbb{E}[Z_t + \theta \cdot Z_{(t-1)}] = \mathbb{E}[Z_t] + \theta \cdot \mathbb{E}[Z_{(t-1)}] = 0 + \theta \cdot 0 = 0$$

$\Rightarrow$

$$\boxed{\mathbb{E}[X_t] = 0}$$

Now we going to compute the autocorrelation:

$$\gamma_X(t, t+k) = Cov(X_t, X_{(t+k)}) = Cov(Z_t + \theta \cdot Z_{(t-1)}, Z_{(t+k)} + \theta \cdot Z_{(t+k-1)})$$

$$= Cov(Z_t, Z_{(t+k)}) + \theta \cdot Cov(Z_{(t-1)}, Z_{(t+k)}) + \theta \cdot Cov(Z_t, Z_{(t+k-1)}) + \theta^2 \cdot Cov(Z_{(t-1)}, Z_{(t+k-1)})$$

$$= \gamma_Z(k) + \theta \cdot \gamma_Z(k+1) + \theta \cdot \gamma_Z(k-1) + \theta^2 \cdot \gamma_Z(k)$$

$$= (1+\theta^2) \cdot \gamma_Z(k) + \theta \cdot [\gamma_Z(k+1) + \gamma_Z(k-1)]$$

$\Rightarrow$

$$\boxed{\gamma_X(k) = (1+\theta^2) \cdot \gamma_Z(k) + \theta \cdot [\gamma_Z(k+1) + \gamma_Z(k-1)]}$$

<u>If $k = 0$:</u>

$$\gamma_X(k) = (1+\theta^2)\cdot\gamma_Z(0) + \theta\cdot[\gamma_Z(1)+\gamma_Z(-1)] = (1+\theta^2)\cdot\sigma^2 + \theta\cdot[0+0] = (1+\theta^2)\cdot\sigma^2$$

$\Rightarrow$

$$\boxed{\gamma_X(k) = (1+\theta^2)\cdot\sigma^2}$$

<u>If $k = 1$:</u>

$$\gamma_X(k) = (1+\theta^2)\cdot\gamma_Z(1) + \theta\cdot[\gamma_Z(2)+\gamma_Z(0)] = (1+\theta^2)\cdot 0 + \theta\cdot[0+\sigma^2] = \theta\cdot\sigma^2$$

$\Rightarrow$

$$\boxed{\gamma_X(k) = \theta\cdot\sigma^2}$$

<u>If $k = -1$:</u>

$$\gamma_X(k) = (1+\theta^2)\cdot\gamma_Z(-1) + \theta\cdot[\gamma_Z(0)+\gamma_Z(-2)] = (1+\theta^2)\cdot 0 + \theta\cdot[\sigma^2+0] = \theta\cdot\sigma^2$$

$\Rightarrow$

$$\boxed{\gamma_X(k) = \theta\cdot\sigma^2}$$

<u>If $|k| > 1$:</u>

$$\gamma_X(k) = (1+\theta^2)\cdot\gamma_Z(k) + \theta\cdot[\gamma_Z(k+1)+\gamma_Z(k-1)] = (1+\theta^2)\cdot 0 + \theta\cdot[0+0] = 0$$

$\Rightarrow$

$$\boxed{\gamma_X(k) = 0}$$

And with that we get the autocorrelation function.

Auto Regressive AR(1):

$$\boxed{X_t = \phi \cdot X_{(t-1)} + Z_t \text{ with } t \in \mathbb{Z}}$$

Where $\{Z_t\} \overset{\text{iid}}{\sim} D(0, \sigma^2)$ and $|\phi| < 1$.

$$\boxed{E[X_t] = 0}$$

$$\boxed{\gamma_X(k) = \frac{\phi^{|k|} \cdot \sigma^2}{1 - \phi^2}}$$

$$\boxed{\rho_X(k) = \phi^{|k|}}$$

Proof:

Note that:

$$X_t = \phi \cdot X_{(t-1)} + Z_t = \phi \cdot \left[\phi \cdot X_{(t-2)} + Z_{(t-1)}\right] + Z_t$$

$$= \phi^2 \cdot X_{(t-2)} + \phi \cdot Z_{(t-1)} + Z_t$$

$\Rightarrow$

$$\boxed{X_t = \sum_{i=0}^{\infty} \phi^i \cdot Z_{(t-i)}}$$

$$\mathbb{E}\left[X_t\right] = \mathbb{E}\left[\sum_{i=0}^{\infty} \phi^i \cdot Z_{(t-i)}\right] = \sum_{i=0}^{\infty} \mathbb{E}\left[\phi^i \cdot Z_{(t-i)}\right] = \sum_{i=0}^{\infty} \theta^i \cdot \mathbb{E}\left[Z_{(t-i)}\right]$$

$$= \sum_{i=0}^{\infty} \phi^i \cdot \mathbb{E}\left[Z_{(t-i)}\right] = \sum_{i=0}^{\infty} \phi^i \cdot 0 = 0$$

$\Rightarrow$

$$\boxed{\mathbb{E}\left[X_t\right] = 0}$$

$$\gamma_X(t, t+k) = Cov\left(\sum_{i=0}^{\infty} \phi^i \cdot Z_{(t-i)}, \sum_{j=0}^{\infty} \phi^j \cdot Z_{(t+k-j)}\right)$$

$$= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} Cov\left(\phi^i \cdot Z_{(t-i)}, \phi^j \cdot Z_{(t+k-j)}\right)$$

$$= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \phi^{(i+j)} \cdot Cov\left(Z_{(t-i)}, Z_{(t+k-j)}\right)$$

$$= \sum_{i=0:i\neq(j-k)}^{\infty} \sum_{j=0}^{\infty} \phi^{(i+j)} \cdot Cov\left(Z_{(t-i)}, Z_{(t+k-j)}\right) + \sum_{i=0}^{\infty} \phi^{(i+[k+i])} \cdot Cov\left(Z_{(t-i)}, Z_{(t+k-[k+i])}\right)$$

$$= \sum_{i=0:i\neq(j-k)}^{\infty} \sum_{j=0}^{\infty} \phi^{(i+j)} \cdot 0 + \sum_{i=0}^{\infty} \phi^{2i} \cdot \phi^{k} \cdot Cov\left(Z_{(t-i)}, Z_{(t-i)}\right)$$

$$= \phi^k \cdot \sum_{i=0}^{\infty} \phi^{2i} \cdot \sigma^2 = \phi^k \cdot \sigma^2 \cdot \sum_{i=0}^{\infty} (\phi^2)^i = \phi^k \cdot \sigma^2 \cdot \frac{1}{1-\phi^2}$$

$\Rightarrow$

$$\boxed{\gamma_X(k) = \frac{\phi^k \cdot \sigma^2}{1-\phi^2}}$$

And with that we get the autocorrelation function.

Auto Regressive $AR(q)$:

$$\boxed{X_t = \phi_1 \cdot X_{(t-1)} + \phi_2 \cdot X_{(t-2)} + ... + \phi_q \cdot X_{(t-q)} + Z_t}$$

Moving Average $MA(p)$:

$$\boxed{X_t = \theta_1 \cdot Z_{(t-1)} + \theta_2 \cdot Z_{(t-2)} + ... + \theta_p \cdot Z_{(t-p)} + Z_t}$$

Combination $ARMA(p, q)$:

$$\boxed{X_t = [\theta_1 \cdot Z_{(t-1)} + \theta_2 \cdot Z_{(t-2)} + ... + \theta_p \cdot Z_{(t-p)}] + [\phi_1 \cdot X_{(t-1)} + \phi_2 \cdot X_{(t-2)} + ... + \phi_q \cdot X_{(t-q)}] + Z_t}$$

Lineal process: A time series $\{X_t\}$ is a lineal process if have a representation:

$$\boxed{X_t = \sum_{i=-\infty}^{\infty} \psi_i \cdot Z_{(t-i)}}$$

Where $\{Z_t\} \overset{\text{iid}}{\sim} D(0, \sigma^2)$ and $\{\psi_i\}_{i=-\infty}^{\infty}$ such that:

$$\boxed{\sum_{i=-\infty}^{\infty} |\psi_i| < \infty}$$

Proposition: If $\{Y_t\}$ is a stationary time series of mean 0 and covariance function γ_Y.
If $\sum_{i=-\infty}^{\infty} |\psi_i| < \infty$ then the next time series:

$$\boxed{X_t = \sum_{i=-\infty}^{\infty} \psi_i \cdot Y_{(t-i)}}$$

Is stationary of mean 0 and covariance function:

$$\boxed{\gamma_X(k) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \psi_i \cdot \psi_j \cdot \gamma_Y(k + [i - j])}$$

Estimation of $\gamma(\cdot)$ and $\rho(\cdot)$:

On the particular case of the $ARMA$ models, we have the next:

$$\boxed{\hat{\rho} \sim N\left(\rho, \frac{W}{n}\right)}$$

$$\boxed{W_{i,j} = \sum_{k=1}^{\infty}[\rho(k+i) + \rho(k-i) - 2 \cdot \rho(i) \cdot \rho(k)] \cdot [\rho(k+j) + \rho(k-j) - 2 \cdot \rho(j) \cdot \rho(k)]}$$

$$\boxed{\hat{\gamma}(k) = \frac{1}{n} \cdot \sum_{t=1}^{n-|k|} (X_{t+|k|} - \bar{X}) \cdot (X_t - \bar{X})}$$

$$\boxed{\bar{X} = \frac{1}{n} \cdot \sum_{t=1}^{n} X_t}$$

Definition of partial autocorrelation (PACF):

$$\boxed{PACF(Y_t, Y_{(t-k)}) = \frac{Cov(Y_t|Y_{(t-1)}, Y_{(t-2)}, ..., Y_{(t-k+1)}, Y_{(t-k)}|Y_{(t-1)}, Y_{(t-2)}, ..., Y_{(t-k+1)})}{\sigma_{Y_t|Y_{(t-1)}, Y_{(t-2)}, ..., Y_{(t-k+1)}} \cdot \sigma_{Y_{(t-k)}|Y_{(t-1)}, Y_{(t-2)}, ..., Y_{(t-k+1)}}}}$$
