Time Series and Sequence Learning

A closer look at AR(1)

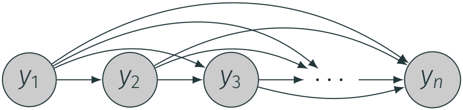
Fredrik Lindsten, Linköping University

Stochastic process

A fundamental approach to time series analysis is to model the data as a stochastic process,

$$\{y_t : t = 1, 2, \dots\}$$

Probabilistic graphical model:



Mean and autocovariance functions

A complete probabilistic description of a stochastic process is given by the *joint probability density function*

$$p(y_{1:n}) = p(y_1, y_2, \ldots, y_n).$$

Derived quantities of interest:

1. Marginal distributions, $p(y_t)$, $t = 1, \ldots, n$
2. Mean function, $\mu(t) := \mathbb{E}[y_t]$
3. Autocovariance function,
   $\gamma(s, t) := \mathrm{Cov}(y_s, y_t) = \mathbb{E}[(y_s - \mu(s))(y_t - \mu(t))]$

And, **of particular interest**, the autocorrelation function (ACF).

$$\rho(s, t) := \frac{\gamma(s, t)}{\sqrt{\gamma(s, s)\gamma(t, t)}}$$

Properties of WN

Let $y_t = \varepsilon_t$, $\quad \varepsilon_t \overset{iid}{\sim} N(0, \sigma_\varepsilon^2)$

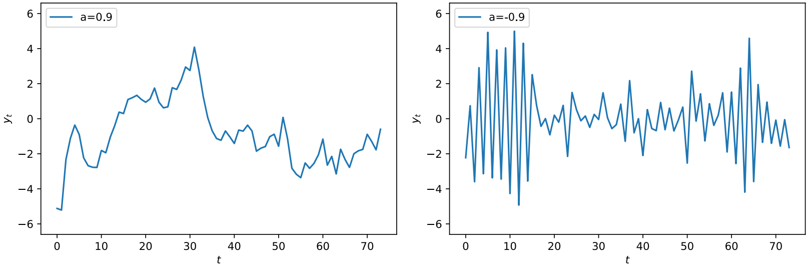
$$\mu(t) = \mathbb{E}[y_t] = 0 \qquad \forall \; t$$

$$\gamma(s,t) = \mathbb{E}[y_s y_t] = \begin{cases} \mathbb{E}[y_t^2] = \sigma_\varepsilon^2 & t = s \\ \underbrace{\mathbb{E}[y_s]}_{=0} \cdot \underbrace{\mathbb{E}[y_t]}_{=0} = 0 & t \neq s \end{cases}$$

$$\rho(s,t) = \begin{cases} 1 & \text{if} \quad s = t \\ 0 & \text{if} \quad s \neq t \end{cases}$$

ex) AR(1): $y_t = a y_{t-1} + \varepsilon_t$

Intuitively: $\mathrm{Corr}(y_{t-1}, y_t) = \rho(t-1, t) = a$



Properties of AR(1)

$Y_t = a Y_{t-1} + \varepsilon_t \,, \quad \varepsilon_t \overset{iid}{\sim} N(0, \sigma_\varepsilon^2)$

$$\mu(t) = \mathbb{E}[Y_t] = \mathbb{E}[a Y_{t-1} + \varepsilon_t] = a \underbrace{\mathbb{E}[Y_{t-1}]}_{\mu(t-1)} + \underbrace{\mathbb{E}[\varepsilon_t]}_{=0}$$

$$\boxed{\therefore \ \mu(t) = a \mu(t-1)}$$

- If $\mu(1) = 0$, then $\mu(t) = 0 \quad \forall t$
- If $|a| < 1$, then $\mu(t) \to 0$ (exponentially fast)

Properties of AR(1)

$$\gamma(s,t) = \mathbb{E}[(y_s - \underbrace{\mu(s)}_{=0})(y_t - \underbrace{\mu(t)}_{=0})] = \mathbb{E}[y_s y_t]$$

$$\mathrm{Var}(y_t) = \gamma(t,t) = \mathbb{E}[y_t^2] = \mathbb{E}[a^2 y_{t-1}^2 + 2a\, y_{t-1}\varepsilon_t + \varepsilon_t^2]$$

$$= a^2 \underbrace{\gamma(t-1,t-1)}_{\mathrm{Var}(y_{t-1})} + 2a\,\mathbb{E}[y_{t-1}]\underbrace{\mathbb{E}[\varepsilon_t]}_{=0} + \underbrace{\mathbb{E}[\varepsilon_t^2]}_{\sigma_\varepsilon^2}$$

$$= a^2 \gamma(t-1,t-1) + \sigma_\varepsilon^2 \qquad \circledast$$

$\therefore$
- If $|a| \geq 1$, then the variance increases to infinity as $t \to \infty$
- If $|a| < 1$, then $\circledast$ has a fixed-point solution

$$\gamma(t-1,t-1) = \gamma(t,t) = \frac{\sigma_\varepsilon^2}{1-a^2}$$

Properties of AR(1)

Assume
- $\mathbb{E}[y_1] = \mu(1) = 0$
- $Var(y_1) = \gamma(1,1) = \dfrac{\sigma_\varepsilon^2}{1-a^2}$

$$\Rightarrow \boxed{\begin{aligned} \mathbb{E}[y_t] &= \mu(t) = 0 \\ Var(y_t) &= \gamma(t,t) = \frac{\sigma_\varepsilon^2}{1-a^2} \end{aligned}}$$

Properties of AR(1)

"Lag 1":

$$\gamma(t,t+1) = \mathbb{E}[y_t y_{t+1}] = \mathbb{E}[y_t(a y_t + \varepsilon_{t+1})]$$

indep.

$$= a \operatorname{Var}(y_t) = a \frac{\sigma_\varepsilon^2}{1-a^2}$$

independent of t!

"Lag 2":

$$\gamma(t,t+2) = \mathbb{E}[y_t y_{t+2}] = \mathbb{E}[y_t(a y_{t+1} + \varepsilon_{t+2})]$$

$$= a\,\mathbb{E}[y_t y_{t+1}] = a\,\gamma(t,t+1) = a^2 \frac{\sigma_\varepsilon^2}{1-a^2}$$

"Lag h":

$$\gamma(t,t+h) = \mathbb{E}[y_t y_{t+h}] = \dots = a^h \frac{\sigma_\varepsilon^2}{1-a^2}$$

Properties of AR(1) - summary

For a first-order AR model $y_t = a y_{t-1} + \varepsilon_t$ **with $|a| < 1$**:

Assume that,

- $\mathbb{E}[y_1] = \mu(1) = 0$, and
- $\mathrm{Var}(y_1) = \gamma(1,1) = \frac{\sigma_\varepsilon^2}{1-a^2}$

We then have, **for all $t \geq 1$.**

- $\mathbb{E}[y_t] = \mu(t) = 0$, and
- $\mathrm{Var}(y_t) = \gamma(t,t) = \frac{\sigma_\varepsilon^2}{1-a^2}$
- $\mathrm{Cov}(y_t, y_{t+h}) = \gamma(t, t+h) = a^h \frac{\sigma_\varepsilon^2}{1-a^2}$