MACHINE LEARNING FOR COMPUTATIONAL FINANCE
ASSIGNMENT 4

KARTHIK IYER

***Problem**: State Space Model for ARMA(1,1)*

Consider ARMA(1,1) model, which can be written as the recursion

$$\xi_t = \epsilon_t + \phi\xi_{t-1} + \theta\epsilon_{t-1},$$

where ϵ_t are independent Gaussian variables with mean 0 and variance σ^2.

There are multiple ways to express this problem as a state space model. We denote x_t as the state and y_t as the observation,

(1) Letting $x_t = \xi_t$, we write the process equation

$$\begin{aligned} x_1 &= \bar{x} + w_1 \\ x_t &= \phi x_{t-1} + w_t, \quad t = 2, \dots, T, \end{aligned}$$

where,

$$w_1 = \epsilon_1, \quad w_t = \epsilon_t + \theta\epsilon_{t-1}, \quad t = 2, \dots, T.$$

Is w_t still a Gaussian variable? Calculate $\mathbb{E}[w_t]$, $\text{Var}(w_t)$ and $\text{Cov}(w_t, w_{t-1})$ for $t = 2, \dots, T$.

Solution: Clearly w_1 is Gaussian. For $t \geq 2$, note that since $w_t = \epsilon_t + \theta\epsilon_{t-1}$, it is a sum of indepdent Gaussian random variables implying that w_t itself is Gaussian. (Sum of indepdent Gaussian random variables is always Gaussian. The cleanest way to prove/see this is via the use of moment generating function. For the sake of brevity, we omit the proof and simply use this result.)

Clearly $\mathbb{E}[w_1] = 0$. For $t \geq 2$, $\mathbb{E}[w_t] = \mathbb{E}[\epsilon_t] + \theta\mathbb{E}[\epsilon_t] = 0$.

$\text{Var}(w_1) = \text{Var}(\epsilon_1) = \sigma^2$, $\text{Var}(w_t) = \text{Var}(\epsilon_t) + \theta^2\text{Var}(\epsilon_{t-1})$; since ϵ_t and ϵ_{t-1} are independent. Thus for $t \geq 2$, $\text{Var}(w_t) = \sigma^2(1 + \theta^2)$.

$\mathbb{E}[w_t w_{t-1}] = \mathbb{E}[\epsilon_t\epsilon_{t-1} + \theta\epsilon_t\epsilon_{t-2} + \theta(\epsilon_{t-1}^2) + \theta^2\epsilon_t\epsilon_{t-1}]$. By independence of ϵ_t and ϵ_s for $t \neq s$ and the fact that $\mathbb{E}[\epsilon_t] = 0$, $\mathbb{E}[w_t w_{t-1}] = \theta\sigma^2$. Since $\mathbb{E}[w_t] = 0$, $\text{Cov}(w_t, w_{t-1}) = \theta\sigma^2$.

■

(2) To express the model compactly, take

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_T \end{bmatrix}, \quad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_T \end{bmatrix}, \quad \eta = \begin{bmatrix} \bar{x} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad w = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_T \end{bmatrix}.$$

Now the state space model can be written

$$\begin{aligned} Gx - \eta &= w, \\ Fx - y &= 0. \end{aligned}$$

What are G and F? Compute W, the full covariance matrix for w.

Solution: Clearly,

$$G = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ -\phi & 1 & 0 & \dots & 0 \\ 0 & -\phi & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots -\phi & 1 \end{bmatrix}$$

and $F = I_{T\times T}$.

Note that for $s = 2, 3, .., t-1$, w_s is only correlated with w_{s-1} and w_{s+1}; w_1 is only corrleated with w_2 and w_T is only correlated with w_{T-1}. Moreover, from Part (1), we know for $s = 2, 3, ..., T$, $\text{Cov}(w_s, w_{s-1}) = \theta\sigma^2$ and that for $s = 1, 2..., T$, $\text{Cov}(w_s, w_s) = \text{Var}(w_s) = \sigma^2(1+\theta^2)$. Hence, the covariance matrix

$$W = \begin{bmatrix} \sigma^2 & \theta\sigma^2 & 0 & \dots & 0 & 0 \\ \theta\sigma^2 & \sigma^2(1+\theta^2) & \theta\sigma^2 & \dots & 0 & 0 \\ 0 & \theta\sigma^2 & \sigma^2(1+\theta^2) & \theta\sigma^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \theta\sigma^2 \\ 0 & 0 & 0 & 0 & \dots \theta\sigma^2 & \sigma^2(1+\theta^2) \end{bmatrix}$$

∎

(3) Form the maximum likelihood formulation to estimate ϕ and θ given y. Simplify the form as much as you can and finish the implementation in **Jupyter Notebook**.

Solution: Note that there is no noise in our observations and hence there is no measurement error that needs to be accounted for. We only need to factor in the innovation error. Since the covariance matrix W for w is symmetric, we can thus formulate the maximum likelihood as the following optimization problem:

$$\min_{\phi,\theta} \frac{1}{2}\|W^{-\frac{1}{2}}(Gy-\eta)\|_F^2$$

We obtain $\phi = 0.52$ *and* $\theta = 0.12326$. ∎

(4) Another way to view ARMA is as a lagged sum of AR time series. Define AR(1) process,

$$\begin{aligned} \zeta_1 &= \bar{\zeta} + \epsilon_1, \\ \zeta_t &= \phi\zeta_{t-1} + \epsilon_t, \quad t = 2, \dots, T. \end{aligned}$$

Then we build a time series based on $\{\zeta_t\}_{t=1}^T$,

$$\begin{aligned} \xi_1 &= \zeta_1, \\ \xi_t &= \zeta_t + \theta\zeta_{t-1}, \quad t = 2, \dots, T. \end{aligned}$$

Show that $\{\xi_t\}_{t=1}^T$ is a ARMA(1,1) process with parameters ϕ and θ.

Proof. For $t = 3, 4, ..., T$,

$$\begin{aligned}\xi_t &= \zeta_t + \theta\zeta_{t-1} \\ &= (\phi\zeta_{t-1} + \epsilon_t) + \theta\zeta_{t-1} \\ &= \epsilon_t + \phi\zeta_{t-1} + \theta(\phi\zeta_{t-2} + \epsilon_{t-1}) \\ &= \epsilon_t + \theta\epsilon_{t-1} + \phi(\zeta_{t-1} + \theta\zeta_{t-2}) \\ &= \epsilon_t + \theta\epsilon_{t-1} + \phi\xi_{t-1}\end{aligned}$$

For $t = 2$, $\xi_2 = \zeta_2 + \theta\zeta_1 = \epsilon_2 + \phi\zeta_1 + \theta\zeta_1 = \epsilon_2 + \phi\zeta_1 + \theta(\bar{\zeta} + \epsilon_1)$. Provided, $\bar{\zeta} = 0$, we obtain $\zeta_2 = \epsilon_2 + \theta\epsilon_1 + \phi\zeta_1$.

We have hence showed that $\{\xi_t\}_{t=1}^T$ is a ARMA(1,1) process with parameters ϕ and θ. ■

(5) Using the sum view, we will treat the underlying AR(1) process as the state, and the observe the lagged sum:

$$\begin{cases} x_1 = \bar{x} + \epsilon_1, \\ x_t = \phi x_{t-1} + \epsilon_t, \quad t = 2, \dots, T. \end{cases} \qquad \begin{cases} y_1 = x_1, \\ y_t = x_t + \theta x_{t-1}, \quad t = 2, \dots, T. \end{cases}$$

Following the same notation as in (2), we want to write the model as,

$$\begin{aligned} Gx - \eta &= \epsilon, \\ Fx - y &= 0. \end{aligned}$$

What are G and F? And what is the maximum likelihood problem?

Solution: It is easy to see that, as before, we obtain

$$G = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ -\phi & 1 & 0 & \dots & 0 \\ 0 & -\phi & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots -\phi & 1 \end{bmatrix}$$

It is equally easy to see that

$$F = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ \theta & 1 & 0 & \dots & 0 \\ 0 & \theta & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots \theta & 1 \end{bmatrix}$$

As in problem 3, we only need to consider the innovation error as there is no noise in the obsrvation. Moreover, the covariance matrix for ϵ is $\sigma^2 I_{T\times T}$ as the errors are iid $N(0, \sigma^2)$. We can thus formulate the maximum likelihood problem as:

$$\min_{\phi,\theta} \frac{1}{2}\|W^{-\frac{1}{2}}(Gx - \eta)\|_F^2 = \min_{\phi,\theta} \frac{1}{2}\|(GF^{-1}y - \eta)\|_F^2,$$

where G is defined above. ■