```
% Run the data generation script
datagen_ts

% Part a
N = length(yk);
[cdf, ~] = ksdensity(yk, [a, b], 'Function', 'cdf');
prob_est = cdf(2)-cdf(1);

% Part b

% High order determination
[pacf, ~, bounds] = parcorr(yk);
order = find(pacf<bounds(2) | pacf>bounds(1), 1, 'last');

% AR Model
Mdl = arima(order, 0, 0);
Mdl.Constant = 0;
EstMdl = estimate(Mdl, yk);
ek = infer(EstMdl, yk);
lbqtest(ek)

% LR Model
Y = yk(4:end) - ek(4:end);
X = [ek(3:end-1), ek(2:end-2), ek(1:end-3)];
LRmdl = fitlm(X, Y, 'y ~ x1 + x2 + x3 - 1');
Chat = LRmdl.Coefficients.Estimate;

% Part C
pVals = LRmdl.Coefficients.pValue;
Cflags = pVals<0.05;

prob_ans = 1;
```

ARIMA(18,0,0) Model (Gaussian Distribution):

	Value	*StandardError*	*TStatistic*	*PValue*
Constant	*0*	*0*	*NaN*	*NaN*
AR{1}	*1.2173*	*0.035714*	*34.083*	*1.317e-254*
AR{2}	*-0.68711*	*0.056548*	*-12.151*	*5.6724e-34*
AR{3}	*-0.2048*	*0.063306*	*-3.2351*	*0.0012158*
AR{4}	*0.83166*	*0.063516*	*13.094*	*3.5794e-39*
AR{5}	*-0.87025*	*0.06901*	*-12.611*	*1.848e-36*
AR{6}	*0.34626*	*0.074353*	*4.6569*	*3.2095e-06*
AR{7}	*0.28571*	*0.070168*	*4.0718*	*4.6647e-05*
AR{8}	*-0.73009*	*0.073631*	*-9.9156*	*3.5614e-23*
AR{9}	*0.57684*	*0.08015*	*7.197*	*6.1529e-13*
AR{10}	*-0.14878*	*0.078715*	*-1.8902*	*0.058737*
AR{11}	*-0.33298*	*0.074936*	*-4.4435*	*8.8495e-06*
AR{12}	*0.49773*	*0.078074*	*6.3751*	*1.8283e-10*

```
    AR{13}      -0.32676     0.079769     -4.0964     4.1963e-05
    AR{14}     -0.010669     0.073233    -0.14568        0.88417
    AR{15}       0.20961     0.062782      3.3386     0.00084192
    AR{16}      -0.17469     0.061753     -2.8288       0.004672
    AR{17}      0.097821      0.05907       1.656       0.097716
    AR{18}     -0.010873     0.038146    -0.28504        0.77561
    Variance      1.9465      0.10897      17.863     2.2949e-71


ans =

  logical

   0
```

Published with MATLAB® R2020a

Table of Contents


```
% SCRIPT TO GENERATE DATA FOR QUIZ 3 MATLAB GRADER problem
%
% For CH5115: Parameter and State Estimation
%
% Arun K. Tangirala
% November 27, 2020
```

Specify invertible model

```
c2 = unifrnd(0.16,0.9); c1 = unifrnd(1-c2,1+c2);
dgp_mod = arima('MA',{c1 c2},'Constant',0);
dgp_mod.Variance = 2;
```

Generate data

```
% Sample size
N = randsample(200:1000,1);
% Simulate
yk = simulate(dgp_mod,N);
```

Intervals for probability computation

```
a =
 min(sign(min(yk))*0.6*abs(min(yk)),sign(max(yk))*0.4*abs(max(yk)));
b = max(sign(max(yk))*0.6*abs(max(yk)),1.2*a);

clear c1 c2 dgp_mod N
```

Published with MATLAB® R2020a

INDIAN INSTITUTE OF TECHNOLOGY MADRAS
Department of Chemical Engineering

CH5115 Parameter and State Estimation
Quiz 3 Solutions

1

Given the following data generating process,

$$x[k] = x^*[k] + \varepsilon_x[k] \tag{1}$$
$$y^*[k] = \alpha_0 x^*[k] \tag{2}$$
$$y[k] = y^*[k] + \varepsilon_y[k] \tag{3}$$

1.a

Calculation of μ_x^*

$$x[k] = x^*[k] + \varepsilon_x[k] \tag{4}$$
$$E(x[k]) = E(x^*k) - \mu_x \tag{5}$$
$$\hat{\mu}_x = \frac{1}{N}\sum_{k=1}^{N} x[k] \tag{6}$$

Calculation of $\hat{\alpha}_0$

$$x[k] = x^*[k] + \varepsilon_x[k] \tag{7}$$
$$y[k] = \alpha_0[x[k] - \varepsilon_x] + \varepsilon_y[k] \tag{8}$$
$$E(y[k]) = \alpha_0 E(x[k]) \ (E(\varepsilon_x) = E(\varepsilon_y) = 0) \tag{9}$$
$$\hat{\alpha}_0 = \frac{\sum_{k=1}^{N} y[k]}{\sum_{k=1}^{N} x[k]} = \frac{\hat{\mu}_y}{\hat{\mu}_x} \tag{10}$$

Calculation of $\hat{\sigma}_x^2$

$$x[k] = x^*[k] + \varepsilon_x[k] \tag{11}$$
$$y[k] = \alpha_0 x^*[k] + \varepsilon_y[k] \tag{12}$$
$$E(y[k]x[k]) = \alpha_0 E(x^*)^2 \tag{13}$$
$$\hat{\sigma}_x^2 = \frac{\sum_{k=1}^{N} y[k]x[k]}{\hat{\alpha}_0} - \hat{\mu}_x^2 \tag{14}$$

Calculation of $\hat{\lambda}_x$ and $\hat{\lambda}_y$

$$x[k] = x^*[k] + \varepsilon_x[k] \tag{15}$$

$$E(x[k] - \mu_x)^2 = \sigma_x^2 + \lambda_x \tag{16}$$

$$\hat{\lambda}_x = \frac{1}{N}\sum_{k=1}^{N} x^2[k] - \hat{\mu}_x^2 - \hat{\sigma}_x^2 \tag{17}$$

$$y[k] = \alpha_0 x^*[k] + \varepsilon_y[k] \tag{18}$$

$$E(y[k] - \mu_y)^2 = \alpha_0^2\sigma_x^2 + \lambda_y \tag{19}$$

$$\hat{\lambda}_y = \frac{1}{N}\sum_{k=1}^{N} y^2[k] - \hat{\mu}_y^2 - \hat{\alpha}_0^2\hat{\sigma}_x^2 \tag{20}$$

MATLAB codes to check the consistency of $\hat{\alpha}$

```
% Generation of x*[k]
mu_xstar = 1, sigma_xstar =  1;
xk_star = normrnd(mu_xstar,sigma_xstar.^2,1000,1);

% Generation of e_x[k]
lambda_x = 2;
ek_x = normrnd(0,sqrt(lambda_x),1000,1);

% Generation of e_y[k]
lambda_y = 3;
ek_y = normrnd(0,sqrt(lambda_y),1000,1);

% Generation of x[k]
xk = xk_star + ek_x;

% Generation of y*[k]
alpha_0 = 1.4;
yk_star = alpha_0*xk_star;

% Generation of y[k]
yk = yk_star + ek_y;

N = length(yk);

% MOM estimate of alpha
alpha_est_mom = sum(yk)/sum(xl)

% Consistency of MOM estimate of alpha
for i = 1:N
    alpha_est_mom(i) = sum(yk(1:i))/sum(xk(1:i));
end

% LS estimate of alpha

alpha_est_ls = sum(yk.*xk)/sum(xk.^2)

% Consistency of LS estimate of alpha
```

```
for i = 1:N
    alpha_est_ls(i) = sum(yk(1:i).*xk(1:i))./sum(xk(1:i).^2);
end

figure;plot(alpha_est_mom,'LineWidth',1.5)
hold on;plot(alpha_est_ls,'k','LineWidth',1.5)
hold on;plot(alpha_0*ones(length(alpha_est_ls),1),'-.','LineWidth'
    ,1.5);
xlabel('Sample size (N)')
ylabel('Estimated \alpha')
title('Consistency of estimated \alpha')
ax = gca;
ax.FontSize = 16;
legend('Estimated \alpha (MOM)','Estimated \alpha (LS)','\alpha_0')
```

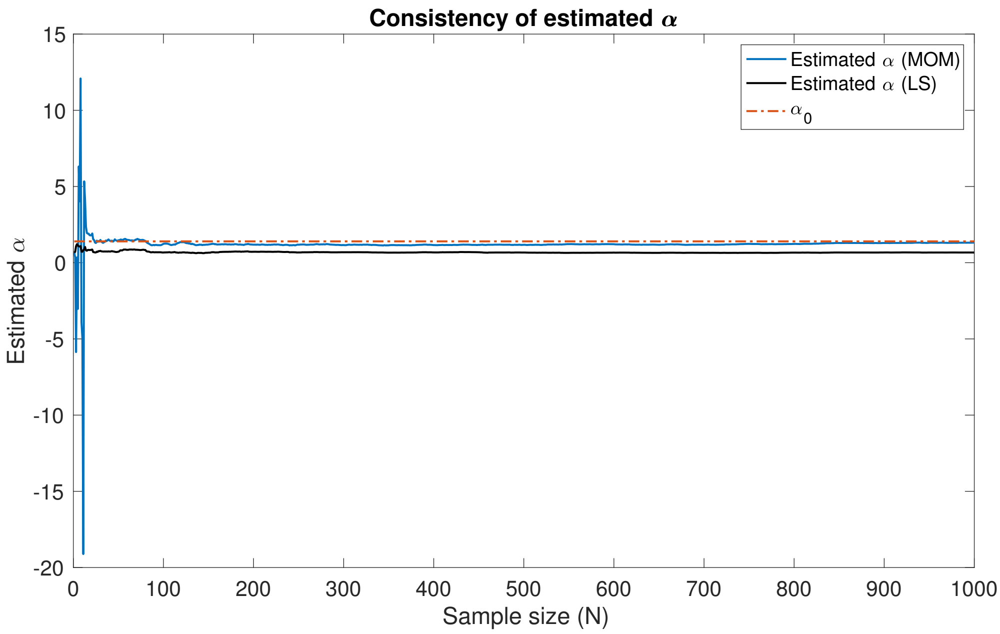

Figure 1: Consistency of $\hat{\alpha}_{\text{MOM}}$ and $\hat{\alpha}_{\text{LS}}$

If μ_x^* is 0, then $\hat{\alpha}_{\text{MOM}}$ can not be estimated using the first moment. It has to be estimated using second-order moment. However, given that we can only use the first and second-order moments, these parameters cannot be estimated uniquely becasue it results in under-determined problem with the need to estimate more number of parameters (4 in this case) using lesser number of equations (3 equations).

1. b

As observed from the Figure 1, $\hat{\alpha}_{\text{MOM}}$ is a consistent estimate while LS results in inconsistent estimate of α. [1]

[1] Since, most of the students attempted to arrive at a conclusion on the consistency of $\hat{\alpha}$ through theoretical formulation. An additional note has also been provided in section 3 for the same.Please note, since the question did not ask for a formal proof any logically consistent argument has been duly credited in this regard.

Calculation of $\hat{\alpha}_{\mathsf{LS}}$

$$O = \sum_{k=1}^{N} (y[k] - \alpha x[k])^2 \tag{21}$$

$$\left.\frac{dO}{d\alpha}\right|_{\alpha=\hat{\alpha}_{\mathsf{LS}}} = \sum_{k=1}^{N} \alpha x^2[k] - \sum_{k=1}^{N} y[k]x[k] = 0 \tag{22}$$

$$\hat{\alpha}_{\mathsf{LS}} = \frac{\sum\limits_{k=1}^{N} y[k]x[k]}{\sum\limits_{k=1}^{N} x^2[k]} \tag{23}$$

Consistency of α_{LS}: The regression model involving $y[k]$ and $x[k]$ can be written as

$$y[k] = \alpha_0(x[k]) + \varepsilon[k] \tag{24}$$

$$\text{where, } \varepsilon = \alpha\varepsilon_x + \varepsilon_y \tag{25}$$

The main condition for consistency of LS estimate is $x[k]$ should be uncorrelated with $\varepsilon[k]$ but in this scenario, $\varepsilon[k]$ contains $\varepsilon_x[k]$ which is correlated with $x[k]$. This violation leads inconsistent estimate of α.

2

2.1

Given $Y \sim Poison(\lambda)$

$$L(y;\lambda) = \frac{e^{-N\lambda}\lambda^{\sum\limits_{k=1}^{N} y[k]}}{\prod\limits_{k=1}^{N} y[k]!} \tag{26}$$

$$\frac{\partial \log L(y;\lambda)}{\partial\lambda} = -N\lambda + \sum_{k=1}^{N} y[k]\log\lambda \tag{27}$$

$$\frac{\partial \log L(y;\lambda)}{\partial\lambda}\Big|_{\lambda=\lambda^*} = -N\lambda + \sum_{k=1}^{N} y[k]\log\lambda = 0 \tag{28}$$

$$\lambda^* = \frac{\sum\limits_{k=1} y[k]}{N} \tag{29}$$

$$\text{Also, } \frac{d^2 \log L(y;\lambda)}{d\lambda^2} = -\frac{\sum\limits_{k=1}^{N} y[k]^2}{\lambda^2} < 0 \; \forall\lambda > 0 \tag{30}$$

Since it is a convex programing and the score is single-root the best estimate of λ is given as

$$\lambda^* = \begin{cases} \frac{\sum_{k=1}^{N} y[k]}{N} & \text{if } \frac{\sum_{k=1}^{N} y[k]}{N} \geq b \\ b & \text{Otherwise} \end{cases}$$

Another way: classic convex programming approach using KKT condition. Since it is a constrained programming, we invoke Lagrange coefficients (λ_1, g_1) to solve it

$$O_{Cons} = -N\lambda + \sum_{k=1}^{N} y[k]\log(\lambda) - \log(\prod_{k=1}^{N} y[k]!) + \lambda_1(\lambda - b) \tag{31}$$

$$\tag{32}$$

Using KKT conditions

$$-N + \frac{\sum\limits_{k=1}^{N} y[k]}{\lambda} + \lambda_1 = 0 \tag{33}$$

$$\lambda - b = 0 \tag{34}$$

$$\tag{35}$$

Possible cases (interior set)

- If $\lambda = b$ then for λ_1 to be positive $N > \frac{\sum\limits_{k=1}^{N} y[k]}{b}$
- Otherwise, set λ_1 ineffective and zero. In that case, the optimal solution for λ is $\frac{\sum\limits_{k=1}^{N} y[k]}{N}$

The solution for λ^*

$$\lambda^* = \begin{cases} \frac{\sum\limits_{k=1}^{N} y[k]}{N} & \text{if } \frac{\sum\limits_{k=1}^{N} y[k]}{N} \geq b \\ b & \text{Otherwise} \end{cases}$$

2.2

Given, $\mathbf{y}_N \sim \Gamma(1, \theta)$

$$L(\mathbf{y}_N; \theta) = \frac{e^{-\frac{\sum\limits_{k=1}^{N} y[k]}{\theta}}}{\theta^N} \tag{36}$$

$$I(\theta) = -E\left(\frac{d^2 \log(L(\mathbf{y}_N; \theta))}{d\theta^2}\right) \tag{37}$$

$$I(\theta) = \frac{N}{\theta^2} \tag{38}$$

So, for $\pi(\theta)$ to be a Jeffrey's prior, it has to be proportional to the square root of $I(\theta)$.

$$\pi(\theta) = K\frac{1}{\theta}$$

From this argument, it is evident that Parignya's choice of $\pi'(\theta) \propto \frac{1}{\theta^2}$ is not a Jeffrey's prior.

$$f(\theta|\mathbf{y}_N)f(\mathbf{y}_N) = f(\mathbf{y}_N|\theta)\pi(\theta) \tag{39}$$

$$f(\theta|\mathbf{y}_N) = \frac{K_\theta e^{-\frac{\sum\limits_{k=1}^{N} y[k]}{\theta}}}{\theta^{N+1}} \tag{40}$$

For $\lambda = \frac{1}{\theta}$

$$I(\theta) = \left(\frac{\partial\lambda}{\partial\theta}\right)^2 I(\lambda) \tag{41}$$

$$\frac{N}{\theta^2} = \frac{I(\lambda)}{\theta^4} \tag{42}$$

$$I(\lambda) = \frac{N}{\lambda^2} \tag{43}$$

Calculation of MMSE for λ

$$f(\lambda|\mathbf{y}_N)f(\mathbf{y}_N) = f(\mathbf{y}_N|\lambda)\pi(\lambda) \tag{44}$$

$$f(\lambda|\mathbf{y}_N) = K_\lambda \lambda^{N-1} e^{-\lambda \sum\limits_{N}^{k=1} y[k]} \tag{45}$$

$$\hat{\lambda}_{\mathsf{MMSE}} = E(f(\lambda|\mathbf{y}_N)) = K_\lambda \frac{\Gamma_{N+1}}{(\sum\limits_{k=1}^{N} y[k])^{(N+1)}} = K_\lambda \frac{N!}{(\sum\limits_{k=1}^{N} y[k])^{(N+1)}} \tag{46}$$

3 Checking the consistency of $\hat{\alpha}$

Given $\hat{\alpha} = \frac{\bar{y}}{\bar{x}}$ whereas $\alpha_0 = \frac{\mu_y}{\mu_x}$. We know for large sample size, $\bar{y} \sim \mathcal{N}(\mu_y, \sigma^2_{\bar{y}})$ and $\bar{x} \sim \mathcal{N}(\mu_x, \sigma^2_{\bar{x}})$. Further,

$$\lim_{N\to\infty} E(\bar{y}_N - \mu_y)^2 = l.i.m_{N\to\infty}\bar{y}_N - \mu_y = 0$$

where, $l.i.m.$ refers to limit in mean square sense.

3.1 Checking the convergence of $\frac{1}{\bar{x}_N}$

Since, $\bar{x}_N$ is Gaussian the distribution of $Z_N = \frac{1}{\bar{x}_N}$ $(g(Z_N))$ can be written as

$$\begin{aligned}
g(Z_N = z_N) =& \frac{d}{dz}[g(Z_N \le z_N)] = \frac{d}{dz_N}\int_{-\infty}^{\frac{1}{z_N}} f(\bar{x}_N)d\bar{x}_N \\
g(Z_N = z_N) =& \frac{\partial \frac{1}{z_N}}{\partial z_N} f(1/z_N) \\
g(Z_N = z_N) =& -\frac{1}{z_N^2\sqrt{2\pi}\sigma_{\bar{x}_N}} e^{-\frac{(\frac{1}{z_N}-\mu_x)^2}{2\sigma^2_{\bar{x}_N}}} \\
\lim_{N\to\infty,\sigma_{\bar{x}_N}\to 0} E(z_N) =& \frac{1}{\mu_x} \\
\lim_{N\to\infty,\sigma_{\bar{x}_N}\to 0} E(z_N)^2 =& \frac{1}{\mu_x^2} \\
\lim_{N\to\infty,\sigma_{\bar{x}_N}\to 0} Var(z_N) =& 0 = \lim_{N\to\infty,\sigma_{\bar{x}_N}\to 0} E(z_N) - \frac{1}{\mu_x} \\
& \implies l.i.m_{N\to\infty} z_N - \frac{1}{\mu_x} = 0
\end{aligned}$$

So, it can be seen that z_N converges to $\frac{1}{\mu_x}$ in a mean square sense.

Now, using the fundamental theorems of mean square calculus, we can write if $\bar{y_N} \to \mu_y$ and $z_N \to \frac{1}{\mu_x}$ then $\lim\limits_{N\to\infty} E(\bar{y_N}\bar{z_N}) = E(\frac{\mu_y}{\mu_x}) = \frac{\mu_y}{\mu_x}$

Since, $\bar{y_N}$ is Gaussian and converges to μ_y it can be shown that $\bar{y_N}^2$ also converges to μ_y^2 (Use similar approach mentioned above). Same holds for z_N as well. Following this we can write

$$\lim_{N\to\infty} E(\bar{y_N}^2 z_N^2) = E(\frac{\mu_y^2}{\mu_x^2}) = \frac{\mu_y^2}{\mu_x^2} \tag{47}$$

To check the consistency of $\hat{\alpha}$ in mean square sense amounts to evaluate the following

$$\implies \lim_{N\to\infty} E(\bar{y}_N z_N - \frac{\mu_y}{\mu_x})^2$$
$$\implies \lim_{N\to\infty} E(\bar{y}_N z_N - E(\bar{y}_N z_N) + E(\bar{y}_N z_N) - \frac{\mu_y}{\mu_x})^2$$
$$\implies \lim_{N\to\infty} E(\bar{y}_N z_N - E(\bar{y}_N z_N) + E(\bar{y}_N z_N) - \frac{\mu_y}{\mu_x})^2$$
$$\implies \lim_{N\to\infty} E(\bar{y}_N z_N - E(\bar{y}_N z_N))^2 + E((\bar{y}_N z_N) - \frac{\mu_y}{\mu_x})^2 + 2\underbrace{E((\bar{y}_N z_N) - \frac{\mu_y}{\mu_x})(\bar{y}_N z_N - E(\bar{y}_N z_N))}_{=0}$$
$$\implies \lim_{N\to\infty} E(\bar{y}_N z_N - E(\bar{y}_N z_N))^2 + \underbrace{E((\bar{y}_N z_N) - \frac{\mu_y}{\mu_x})^2}_{=0,\ \bar{y_N} z_N \to \frac{\mu_y}{\mu_x}}$$

$$\lim_{N\to\infty} E(\bar{y}_N z_N - E(\bar{y}_N z_N))^2 = \lim_{N\to\infty} E(\bar{y}_N z_N)^2 - (E(\bar{y}_N z_N))^2$$

Using equation 47 we can write

$$\lim_{N\to\infty} E(\bar{y}_N z_N)^2 - (E(\bar{y}_N z_N))^2 = 0$$

So, from here it can be shown that $\hat{\alpha}$ is consistent in MSE sense.

Interestingly, if $\mu_x = 0$ then the distribution for $\hat{\alpha}$ becomes bimodal in which case, the first and second order moment are undefined.