Problem Set 2

Lorenzo Orsenigo, Luca Geminiani

November 6^{th}, 2023

Exercise 1

Consider a data-generating process:

$$y_t = \alpha + y_{t-1} + \varepsilon_t$$

were we assume $\alpha = 4$ and ε_t is white noise with variance $\sigma_\varepsilon^2 = 0.2$.
We want to estimate the following regression

$$y_t = \alpha + \rho y_{t-1} + \varepsilon_t$$

and test the null hypothesis $H_0 : \rho = 1$.

```
t_test = zeros(100000,1);
alpha=4;
for i = 1:100000
    y = zeros(250,1);
    y(1)= alpha;
    epsilon = sqrt(0.2) * randn(250, 1);
    for t = 2:250
        y(t) = alpha+ y(t-1) + epsilon(t);
    end
    vx = [ones(249,1) y(1:249)];
    vy = y(2:250);
    [beta,res,cov_b] = OLS(vy,vx);
    t_test(i) = (beta(2)-1)/sqrt(cov_b(2,2));
end

critical = tinv(1-(0.05/2),250-2);
sum(abs(t_test)>critical)/100000
figure;histogram(t_test)
```

where the OLS function is defined as follows

```
function [b,res,cov_b] = OLS(y,x)

b = inv(x'*x)*x'*y ;
res = y-x*b;
cov_b = inv(x'*x)*cov(res);

end
```

We can immediately see that a t-test works well despite being unit root testing a non-stationary process. First of all, we can look at the empirical distribution of the t-statistic:

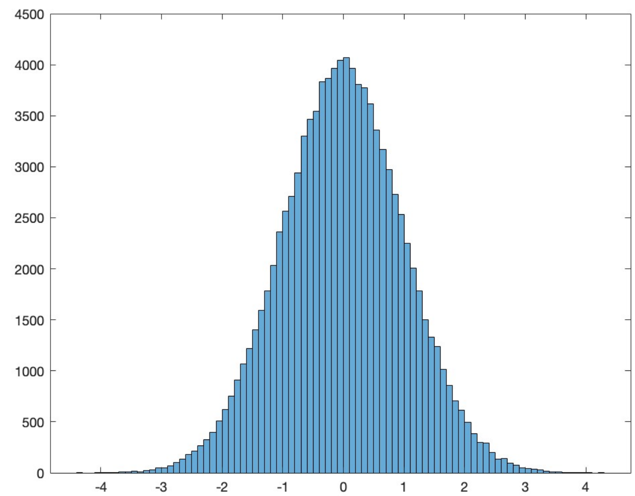

Figure 1: empirical distribution of the t-statistic

We see that the distribution is indeed a standard normal one and, more precisely, we can see that the t-test we performed is a correct 0.05-level test since we wrongly reject the null hypothesis only 0.0505% of the times. But why does the t-test work? why the coefficients are asymptotically gaussian?
Intuitively Enders provides the following explanation:

> *"If the data-generating process contains any deterministic regressors (i.e., an intercept or a time trend) and the estimating equation contains these deterministic regressors, inference on all coefficients can be conducted using a t-test or an F-test. This is because a test involving a single restriction across parameters with different rates of convergence is dominated asymptotically by the parameters with the slowest rates of convergence".*

In our case, it means that the regressor y_{t-1} is asymptotically dominated by the deterministic one, the intercept α. We can further confirm this fact by looking at the mean of the empirical distribution of the t-statistic

```
mean (t_test)
```

in our case -0.0095, almost 0, as it should be in a standard normal distribution
Now we can check if the intuition previously provided is still true, as it should, if we add a time-trend both to the generating model and the regression. Hence the DGP (where we assume $\delta = 1$) becomes

$$y_t = \alpha + y_{t-1} + \delta t + \varepsilon_t$$

and the regression

$$y_t = \alpha + \rho y_{t-1} + \delta t + \varepsilon_t$$

By testing the same null hypothesis $H_0 : \rho = 1$

```
t_test = zeros(100000,1);
alpha=4;
delta=1;
for i = 1:100000
    y = zeros(250,1);
    y(1)= alpha;
    epsilon = sqrt(0.2) * randn(250, 1);
    for t = 2:250
    y(t) = alpha+ y(t-1) + delta*(t-1) +epsilon(t);
    end
    vx = [ones(249,1) y(1:249) (1:249)'];
    vy = y(2:250);
    [beta,res,cov_b] = OLS(vy,vx);
    t_test(i) = (beta(2)-1)/sqrt(cov_b(2,2));
end

critical = tinv(1-(0.05/2),250-3);
sum(abs(t_test)>critical)/100000
figure;histogram(t_test)
```

we can see how the 0.05 t-test is still successfully providing a 0.05-level test as we commit a type I error 0.0496% of the times.

The empirical distribution of the t-statistic is indeed Gaussian.

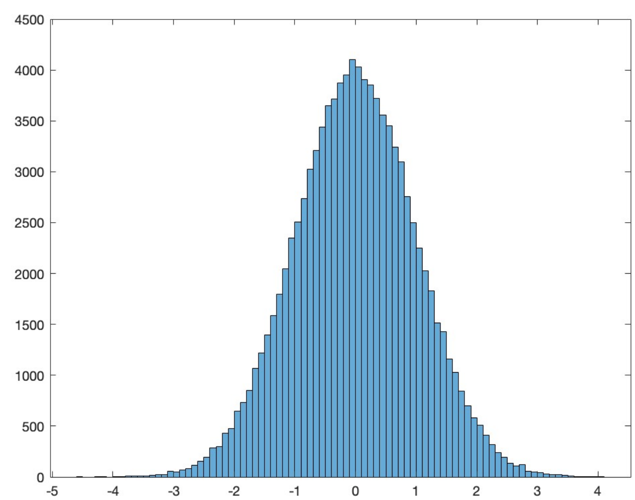

Figure 2: empirical distribution of the t-statistic

Exercise 2

In order to show that the regression coefficient, the t-test and the R^2 are meaningless in the case of a spurious regression we can look at what happens in four different cases.
To begin, we look at a standard regression model:

$$y_t = \alpha_0 + \alpha_1 z_t + e_t$$

where the two sequences $\{z\}$ and $\{y\}$ are two independent AR(1) process:

$$y_t = \rho_1 y_{t-1} + \varepsilon_{y,t}$$

and

$$z_t = \rho_2 z_{t-1} + \varepsilon_{z,t}$$

where we assume $\varepsilon_{z,t}$ and $\varepsilon_{y,t}$ are white noise with standard deviation equal to 1.
In the first case, we want to consider the case in which both $\{z\}$ and $\{y\}$ are stationary and we assume $\rho_1 = 0.7$ and $\rho_2 = 0.2$, hence $e_t = \rho_1 y_{t-1} + \varepsilon_{y,t}$.
In the second case assume that one of the sequences $\{z\}$ is non-stationary, so $\rho_2 = 1$, while the other sequence remains a stationary AR(1) with $\rho_1 = 0.7$.
In the third example, we consider a scenario where both processes are non-stationary and thus $\rho_1 = \rho_2 = 1$.
In the fourth and last case, we look at two random walks with a perfectly correlated white noise hence $\rho_1 = \rho_2 = 1$ as before and $\varepsilon_{z,t} = \varepsilon_{y,t}$.
Let's see, assuming y_0 and z_0 equal to 0 for simplicity, how these different scenarios affect the regression coefficients, the t-test and the R^2 with the standard regression models:

```
%%%% 1st case %%%%
T = 250;
rho1 = 0.7;
rho2 = 0.2;
MC = 5000;
rho = 1;
Y = zeros(T,1);
for t = 1:MC
    e = randn(T,1);
    e(1)=0;
    w = randn(T,1);
    w(1)=0;
    y = filter(1,[1 -rho1],e);
    Z = filter(1,[1 -rho2],w);
    X = [ones(T, 1), Z];
    beta_hat1(:,t) = X \ y;
```

```
    iXX = inv(X'*X);
    sig_hat(t) = cov(y - X*beta_hat1(:,t));
    s_beta = iXX(2,2)*sig_hat(t);
    t_ratio1(t) = (beta_hat1(2,t))/sqrt(s_beta);
    residuals = y - X * beta_hat1(:, t);
    y_bar = mean(y);
    tss = sum((y - y_bar).^2);
    rss = sum(residuals.^2);
    r_squared1(t) = 1 - rss / tss;
end

%%%% 2n case %%%%

T = 250;
rho1 = 0.7;
rho2 = 1;
MC = 5000;
alpha = 0.6;
rho = 1;
Y = zeros(T,1);
for t = 1:MC
    e = randn(T,1);
    e(1)=0;
    w = randn(T,1);
    w(1)=0;
    y = filter(1,[1 -rho1],e);
    Z = filter(1,[1 -rho2],w);
    X = [ones(T, 1), Z];
    beta_hat2(:,t) = X \ y;
    iXX = inv(X'*X);
    sig_hat(t) = cov(y - X*beta_hat2(:,t));
    s_beta = iXX(2,2)*sig_hat(t);
    t_ratio2(t) = (beta_hat2(2,t))/sqrt(s_beta);
    residuals = y - X * beta_hat2(:, t);
    y_bar = mean(y);
    tss = sum((y - y_bar).^2);
    rss = sum(residuals.^2);
    r_squared2(t) = 1 - rss / tss;
end

%%%% 3rd case %%%%

T = 250;
```

```
rho = 1;
MC = 5000;
alpha = 0.6;
Y = zeros(T,1);
for t = 1:MC
    e = randn(T,1);
    e(1)=0;
    w = randn(T,1);
    w(1)=0;
    y = filter(1,[1 -rho],e);
    Z = filter(1,[1 -rho],w);
    X = [ones(T, 1), Z];
    beta_hat3(:,t) = X \ y;
    iXX = inv(X'*X);
    sig_hat(t) = cov(y - X*beta_hat3(:,t));
    s_beta = iXX(2,2)*sig_hat(t);
    t_ratio3(t) = (beta_hat3(2,t))/sqrt(s_beta);
    residuals = y - X * beta_hat3(:, t);
    y_bar = mean(y);
    tss = sum((y - y_bar).^2);
    rss = sum(residuals.^2);
    r_squared3(t) = 1 - rss / tss;
end

%%%% 4th case %%%%

T = 250;
rho = 1;
MC = 5000;
Y = zeros(T,1);
for t = 1:MC
    w = randn(T,1);
    w(1)=0;
    y = filter(1,[1 -rho],w);
    Z = filter(1,[1 -rho],w);
    X = [ones(T, 1), Z];
    beta_hat4(:,t) = X \ y;
    iXX = inv(X'*X);
    sig_hat(t) = cov(y - X*beta_hat4(:,t));
    s_beta = iXX(2,2)*sig_hat(t);
    t_ratio4(t) = (beta_hat4(2,t))/sqrt(s_beta);
    residuals = y - X * beta_hat4(:, t);
    y_bar = mean(y);
```

```
    tss = sum((y - y_bar).^2);
    rss = sum(residuals.^2);
    r_squared4(t) = 1 - rss / tss;
end

%%%% results %%%%

figure
[f,xi] = ksdensity(beta_hat1(2,:));
plot(xi,f,'LineWidth', 2);
hold on
[f,xi] = ksdensity(beta_hat2(2,:));
plot(xi,f,'LineWidth', 2);
[f,xi] = ksdensity(beta_hat3(2,:));
plot(xi,f,'LineWidth', 2);
hold off

critical = tinv(1-(0.05/2),250-2);
sum(abs(t_ratio1)>critical)/5000
sum(abs(t_ratio2)>critical)/5000
sum(abs(t_ratio3)>critical)/5000
sum(abs(t_ratio4)>critical)/5000

figure
[f,xi] = ksdensity(r_squared1);
plot(xi,f,'LineWidth', 2);
hold on
[f,xi] = ksdensity(r_squared2);
plot(xi,f,'LineWidth', 2);
[f,xi] = ksdensity(r_squared3);
plot(xi,f,'LineWidth', 2);
hold off

mean(r_squared1)
mean(r_squared2)
mean(r_squared3)

mean(beta_hat4(2,:)
mean(r_squared4)
```

From these Monte-Carlo experiments, we can now analyse the results.

- In the first case, the regression model is appropriate. Indeed the estimated value is approximately 0, we wrongly reject null hypothesis $H_0: \alpha_1 = 0$ at a 5% significance level only 8% of times and the R^2 is very low, approximately 0. The e_t of the regression is equal to $\sum_{i=1}^{t} \rho_1^i \varepsilon_{y,t} - \alpha_1 \sum_{i=1}^{t} \rho_2^i \varepsilon_{z,t}$ and hence a stationary process.

- In the second case the residual is instead non-stationary due to the non-stationarity of $\{z\}$, it has a stochastic trend, the deviations from the model are permanent, indeed $e_t = \sum_{i=1}^{t} \varepsilon_{y,t} - \alpha_1 \sum_{i=1}^{t} \rho_2^i \varepsilon_{z,t}$. Variance explodes as t increases and hence standard test cannot be applied. The estimate is still on average 0 but the null hypothesis $H_0: \alpha_1 = 0$ at a 5% significance level approximately 40% of times and the on average R^2 is higher than before

- In the third case the residual e_t is equal to $\sum_{i=1}^{t} \varepsilon_{y,t} - \alpha_1 \sum_{i=1}^{t} \varepsilon_{z,t}$ which, again, is non-stationary, so we cannot apply standard regression model as the results will be unreliable, all errors are permanent and the variance explodes as t increases. Indeed we wrongly reject the same null hypothesis 86% of times, and on average the R^2 increases even more.
  We can see the distribution of the estimation (case 1 in red, case 2 in blue, case 3 in yellow):

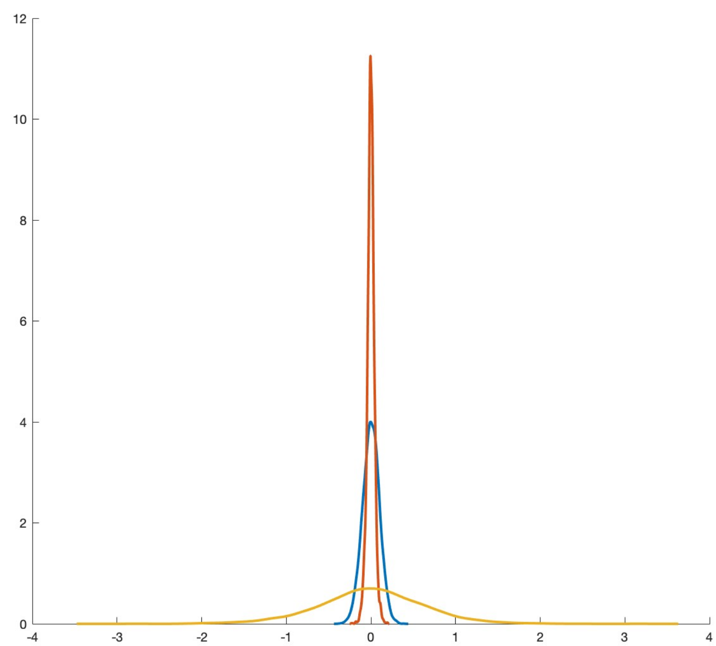

Figure 3: empirical distribution of $\hat{\alpha}_0$

and the distribution of the R^2 (case 1 in blue, case 2 in red, case 3 in yellow)

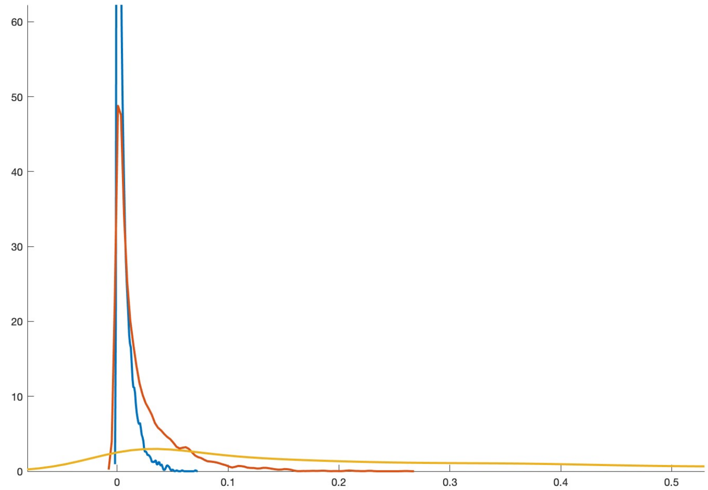

Figure 4: empirical distribution of the R^2

- In the last case we have perfectly correlated random white noises, which, with the assumptions of $z_0 = y_0 = 0$ leads to $\hat{\alpha}_0 = 1$ and $e_t = 0$, thus $R^2 = 1$ and an always rejected null hypothesis.

Exercise 3

First of all we create a matrix with the data in the file *Romer$_R$omer.xlsx* and eliminate the first column, the dates.

```
data1= xlsread("Romer_Romer.xlsx");
data=data1(:,2:end);
```

Hence the first column is the US inflation time series, the second Unemployment, the third US federal funds rate, and finally the Romer and Romer monetary policy shock time series. If we want to understand whether the shocks Granger-cause other variables we can do so by estimating a VAR(4). We want essentially to see whether past values (in this case 4 lags) of the Romer and Romer monetary policy shock help in predicting inflation, unemployment or federal funds rate given their past values.
On MATLAB we simply run 3 OLS regressions, where the dependent variable is respectively inflation, unemployment and FFR and the explanatory ones are the 4 lagged values for each variable.
First consider the regression, where inflation is $\{\pi\}$, unemployment $\{u\}$, FFR $\{i\}$ and the monetary policy shock $\{R\}$:

$$\pi_t = \alpha_1 \pi_{t-1} + \alpha_2 u_{t-1} + \alpha_3 i_{t-1} + \alpha_4 R_{t-1} + \alpha_5 \pi_{t-2} + \alpha_6 u_{t-2} + \alpha_7 i_{t-2} +$$

$$\alpha_8 R_{t-2} + \alpha_9 \pi_{t-3} + \alpha_{10} u_{t-3} + \alpha_{11} i_{t-3} + \alpha_{12} R_{t-3} + \alpha_{13} \pi_{t-4} + \alpha_{14} u_{t-4} + \alpha_{15} i_{t-4} + \alpha_{16} R_{t-4} + \varepsilon_{\pi,t}$$

hence we want to estimate $\{\alpha\}$:

```
depvar = data(5:end,1);
lag1 = data(4:end-1,1:end);
lag2 = data(3:end-2,1:end);
lag3 = data(2:end-3,1:end);
lag4 = data(1:end-4,1:end);
X = [lag1 , lag2 , lag3 , lag4];
b = inv(X'*X)*X'*depvar;
res = depvar-X*b;
cov_b = inv(X'*X)*cov(res);
```

To see whether R_t Granger-causes ϕ_t we need to look at $\hat{\alpha}_4$,$\hat{\alpha}_8$, $\hat{\alpha}_{12}$ and $\hat{\alpha}_{16}$ and we can test the null hypothesis $H_0 : \alpha = 0$ through a t-test with a Bonferroni correction given the numerous parameters(16):

```
t_test=zeros(16,1);
for t=1:16
t_test(t)=b(t)/sqrt(cov_b(t,t));
end
critical = tinv(1-(0.05/32),108-16);
abs(t_test)>critical
```

We can see one lag of R_t helps in predicting y_t since with $\hat{\alpha}_8$ we reject the null hypothesis with a 95% confidence. Thus R_t Granger-causes inflation.

Second, consider the following regression:

$$u_t = \alpha_1\pi_{t-1} + \alpha_2 u_{t-1} + \alpha_3 i_{t-1} + \alpha_4 R_{t-1} + \alpha_5\pi_{t-2} + \alpha_6 u_{t-2} + \alpha_7 i_{t-2} +$$

$$\alpha_8 R_{t-2} + \alpha_9\pi_{t-3} + \alpha_{10} u_{t-3} + \alpha_{11} i_{t-3} + \alpha_{12} R_{t-3} + \alpha_{13}\pi_{t-4} + \alpha_{14} u_{t-4} + \alpha_{15} i_{t-4} + \alpha_{16} R_{t-4} + \varepsilon_{\pi,t}$$

hence we want to estimate $\{\alpha\}$:

```
depvar = data(5:end,2);
lag1 = data(4:end-1,1:end);
lag2 = data(3:end-2,1:end);
lag3 = data(2:end-3,1:end);
lag4 = data(1:end-4,1:end);
X = [lag1 , lag2 , lag3 , lag4];
b = inv(X'*X)*X'*depvar;
res = depvar-X*b;
cov_b = inv(X'*X)*cov(res);
```

To see whether R_t Granger-causes u_t we need to look at $\hat{\alpha}_4$,$\hat{\alpha}_8$, $\hat{\alpha}_{12}$ and $\hat{\alpha}_{16}$ and we can test the null hypothesis $H_0 : \alpha = 0$ through a t-test with a Bonferroni correction given the numerous parameters(16):

```
t_test=zeros(16,1);
for t=1:16
t_test(t)=b(t)/sqrt(cov_b(t,t));
```

```
end
critical = tinv(1-(0.05/32),108-16);
abs(t_test)>critical
```

We can see 0 lags of R_t help in predicting y_t thus R_t does not Granger-cause unemployment.This indicates that monetary policy changes do not significantly affect unemployment in the immediate term. Unemployment might be affected, in the long-run, by other causal channels spurring from the shock.
Third, consider the following regression:

$$i_t = \alpha_1 \pi_{t-1} + \alpha_2 u_{t-1} + \alpha_3 i_{t-1} + \alpha_4 R_{t-1} + \alpha_5 \pi_{t-2} + \alpha_6 u_{t-2} + \alpha_7 i_{t-2} +$$

$$\alpha_8 R_{t-2} + \alpha_9 \pi_{t-3} + \alpha_{10} u_{t-3} + \alpha_{11} i_{t-3} + \alpha_{12} R_{t-3} + \alpha_{13} \pi_{t-4} + \alpha_{14} u_{t-4} + \alpha_{15} i_{t-4} + \alpha_{16} R_{t-4} + \varepsilon_{\pi,t}$$

hence we want to estimate $\{a\}$:

```
depvar = data(5:end,3);
lag1 = data(4:end-1,1:end);
lag2 = data(3:end-2,1:end);
lag3 = data(2:end-3,1:end);
lag4 = data(1:end-4,1:end);
X = [lag1 , lag2 , lag3 , lag4];
b = inv(X'*X)*X'*depvar;
res = depvar-X*b;
cov_b = inv(X'*X)*cov(res);
```

To see whether R_t Granger-causes i_t we need to look at $\hat{\alpha}_4$,$\hat{\alpha}_8$, $\hat{\alpha}_{12}$ and $\hat{\alpha}_{16}$ and we can test the null hypothesis $H_0 : \alpha = 0$ through a t-test with a Bonferroni correction given the numerous parameters(16):

```
t_test=zeros(16,1);
for t=1:16
t_test(t)=b(t)/sqrt(cov_b(t,t));
end
critical = tinv(1-(0.05/32),108-16);
abs(t_test)>critical
```

We can see one lag of R_t helps in predicting i_t since with $\hat{\alpha}_{16}$ we reject the null hypothesis with a 95% confidence. Thus R_t Granger-causes Federal Funds Rate.

Exercise 4

We consider the following DGP:

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} 1 & L^2 \\ \frac{\beta}{1-\beta} & \frac{\beta^2}{1-\beta} + \beta L \end{bmatrix} \begin{bmatrix} \eta_t \\ \varepsilon_t \end{bmatrix}$$

with $u_t = [\eta \ \varepsilon]'$, $Cov(u_t) = \begin{bmatrix} 1 & 0 \\ 0 & 0.8 \end{bmatrix}$ and $\beta = 0.6$.

Now we want to estimate from 500 observations of $[x_t \ y_t]'$ a VAR(4) through the following regressions:

$$x_t = b_1 x_{t-1} + c_1 y_{t-1} + d_1 x_{t-2} + e_1 y_{t-2} + ... + \eta_t$$

$$y_t = a_2 x_t + b_2 x_{t-1} + c_2 y_{t-1} + d_2 x_{t-2} + e_2 y_{t-2} + ... + \varepsilon_t$$

then we want to compute the impulse response function.

Given that

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ a_2 & 1 \end{bmatrix} \begin{bmatrix} b_1 & c_1 \\ b_2 & c_2 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ a_2 & 1 \end{bmatrix} \begin{bmatrix} d_1 & e_1 \\ d_2 & e_2 \end{bmatrix} \begin{bmatrix} x_{t-2} \\ y_{t-2} \end{bmatrix} + ... + \begin{bmatrix} 1 & 0 \\ a_2 & 1 \end{bmatrix} \begin{bmatrix} \eta_t \\ \varepsilon_t \end{bmatrix}$$

and the companion form is

$$\begin{bmatrix} w_t \\ w_{t-1} \\ w_{t-2} \\ w_{t-3} \end{bmatrix} = \begin{bmatrix} A_1 & A_2 & A_3 & A_4 \\ I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & I & 0 \end{bmatrix} \begin{bmatrix} w_{t-1} \\ w_{t-2} \\ w_{t-3} \\ w_{t-4} \end{bmatrix} + \begin{bmatrix} C \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} \eta_t \\ \varepsilon_t \end{bmatrix}$$

where I = $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $0 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, $w_t = \begin{bmatrix} x_t \\ y_t \end{bmatrix}$, $C = \begin{bmatrix} 1 & 0 \\ a_2 & 1 \end{bmatrix}$, and $A_1 = C \begin{bmatrix} b_1 & c_1 \\ b_2 & c_2 \end{bmatrix}$.

The IRF of the shocks η_t and ε_t are equal to the upper-left $2 * 2$ block of these matrices:

- impact D
- one period ahead AD
- two periods ahead $A^2 D$
- ...
- k periods ahead $A^k D$

We run 500 simulations and compute the average IRF for 4 periods to then compare it with the true one.

```
T = 600;
vary = 0.8;
beta = 0.6;
N = 500;
Imp0 = cell([1 N]);
Imp1 = cell([1 N]);
Imp2 = cell([1 N]);
```

```
Imp3 = cell([1 N]);
Imp4 = cell([1 N]);

for i = 1:N
   eta = randn(T,1);
   epsilon = randn(T,1)*sqrt(vary) ;
   X = zeros(T,1);
   Y = zeros(T,1);
   for t = 3:T
       x(t) = eta(t) + epsilon(t-2);
       y(t) = (beta/(1-beta))*eta(t) + (beta^2/(1-beta))*epsilon(t)+beta*epsilon(t-1);
   end
   G = x(101:600);
   Y = y(101:600);
   data = [G ; Y]';

   depvar = data(5:end,1);  %Regression for X
   lag1 = data(4:end-1,:);
   lag2 = data(3:end-2,:);
   lag3 = data(2:end-3,:);
   lag4 = data(1:end-4,:);
   X = [lag1 , lag2 , lag3 , lag4];
   b = inv(X'*X)*X'*depvar;
   res1 = depvar-X*b;
   cov_b = inv(X'*X)*cov(res1);

   depvar = data(5:end,2);  %Regression for y
   lag1 = data(4:end-1,:);
   lag2 = data(3:end-2,:);
   lag3 = data(2:end-3,:);
   lag4 = data(1:end-4,:);
   X = [data(5:end,1), lag1 , lag2 , lag3 , lag4];
   b2 = inv(X'*X)*X'*depvar;
   res2 = depvar-X*b2;
   cov_b = inv(X'*X)*cov(res2);

   V = [res1 res2];
   I = eye(2);
   C = [1 0; b2(1) 1];
   A1 =C*[b(1) b(2); b2(2) b2(3)];
   A2 = C*[b(3) b(4);b2(4) b2(5)];
   A3 = C*[b(5) b(6);b2(6) b2(7)];
   A4 = C*[b(7) b(8);b2(8) b2(9)];
```

```
    O = zeros(2,2);
    A = [A1 A2 A3 A4; I O O O; O I O O; O O I O];
    D = [C; O; O; O];

    a_1 = A*D;
    a_2 = (A^2)*D;
    a_3 = (A^3)*D;
    a_4 = (A^4)*D;

    Imp0{i} = [D(1,:); D(2,:); D(3,:); D(4,:)];
    Imp1{i} = [a_1(1,:); a_1(2,:); a_1(3,:); a_1(4,:)];
    Imp2{i}= [a_2(1,:); a_2(2,:); a_2(3,:); a_2(4,:)];
    Imp3{i} = [a_3(1,:); a_3(2,:); a_3(3,:); a_3(4,:)];
    Imp4{i} = [a_4(1,:); a_4(2,:); a_4(3,:); a_4(4,:)];

end

%Compute means of the impulse responses.

ImpAvg0 = zeros(4,2);
ImpAvg1 = zeros(4,2);
ImpAvg2 = zeros(4,2);
ImpAvg3 = zeros(4,2);
ImpAvg4 = zeros(4,2);

for j = 1:N
     ImpAvg0 = ImpAvg0 + Imp0{j};
     ImpAvg1 = ImpAvg1 + Imp1{j};
     ImpAvg2 = ImpAvg2 + Imp2{j};
     ImpAvg3 = ImpAvg3 + Imp3{j};
     ImpAvg4 = ImpAvg4 + Imp4{j};
end

MeanImp0 = ImpAvg0/N ;
MeanImp1 = ImpAvg1/N ;
MeanImp2 = ImpAvg2/N ;
MeanImp3 = ImpAvg3/N ;
MeanImp4 = ImpAvg4/N ;
```

and hence we can see the difference between the estimated (grey line) and the true impact (red line):

```
%Impulse Response of X to a unitary change in eta
subplot(2,3,1);
time = [0:4];
```

```
X1ETA = [MeanImp0(1,1) MeanImp1(1,1) MeanImp2(1,1) MeanImp3(1,1) MeanImp4(1,1)];
plot(time, X1ETA, 'black');
hold on
TXETA = [1 0 0 0 0]; % True impulse response
plot(time, TXETA, 'red');
xlabel('time')
ylabel('x')
title('Impulse␣Response␣of␣X␣to␣a␣unitary␣change␣in␣eta')
hold off

%IRF of x to a 1 unit change in epsilon
subplot(2,3,2);
X1EPS = [MeanImp0(1,2) MeanImp1(1,2) MeanImp2(1,2) MeanImp3(1,2) MeanImp4(1,2)];
plot(time, X1EPS, 'black');
hold on
TXEPS = [0 0 1 0 0]; % True impulse response
plot(time, TXEPS, 'red');
xlabel('time')
ylabel('x')
title('IRF␣of␣x␣to␣a␣1␣unit␣change␣in␣epsilon')
hold off

%IRF of x to a 1 unit change in both eta and epsilon
subplot(2,3,3);
X1EPSETA = [MeanImp0(1,1)+MeanImp0(1,2), MeanImp1(1,1)+ MeanImp1(1,2),
MeanImp2(1,1)+ MeanImp2(1,2), MeanImp3(1,1)+MeanImp3(1,2), MeanImp4(1,1)+MeanImp4(1,2)];
plot(time, X1EPSETA, 'black')
hold on
TXEPSETA = [1 0 1 0 0]; % True impulse response
plot(time, TXEPSETA, 'red');
xlabel('time')
ylabel('x')
title('IRF␣of␣x␣to␣a␣1␣unit␣change␣in␣both␣eta␣and␣epsilon')
hold off

%IRF Of y to a unit change in eta
subplot(2,3,4);
Y1ETA = [MeanImp0(2,1) MeanImp1(2,1) MeanImp2(2,1) MeanImp3(2,1) MeanImp4(2,1)];
plot(time, Y1ETA, 'black');
hold on
TYETA = [beta/(1-beta) 0 0 0 0]; % True impulse response
plot(time, TYETA, 'red');
xlabel('time')
```

```
ylabel('y')
title('IRF␣of␣y␣to␣a␣unit␣change␣in␣eta')
hold off

%IRF of y to a 1 unit change in epsilon
subplot(2,3,5);
Y1EPS = [MeanImp0(2,2) MeanImp1(2,2) MeanImp2(2,2) MeanImp3(2,2) MeanImp4(2,2)];
plot(time, Y1EPS, 'black');
hold on
TYEPS = [beta^2/(1-beta) beta 0 0 0]; % True impulse response
plot(time, TYEPS, 'red');
xlabel('time')
ylabel('y')
title('IRF␣of␣y␣to␣a␣1␣unit␣change␣in␣epsilon')
hold off

%IRF of y to a 1 unit change in both eta and epsilon
subplot(2,3,6);
Y1EPSETA = [MeanImp0(2,1)+MeanImp0(2,2), MeanImp1(2,1)+ MeanImp1(2,2),
MeanImp2(2,1)+ MeanImp2(2,2), MeanImp3(2,1)+MeanImp3(2,2), MeanImp4(2,1)+MeanImp4(2,2)];
plot(time, Y1EPSETA, 'black')
hold on
TYEPSETA = [beta^2/(1-beta)+beta/(1-beta), beta, 0  0 0]; % True impulse response
plot(time, TYEPSETA, 'red');
xlabel('time')
ylabel('y')
title('IRF␣of␣y␣to␣a␣1␣unit␣change␣in␣both␣eta␣and␣epsilon')
hold off
```

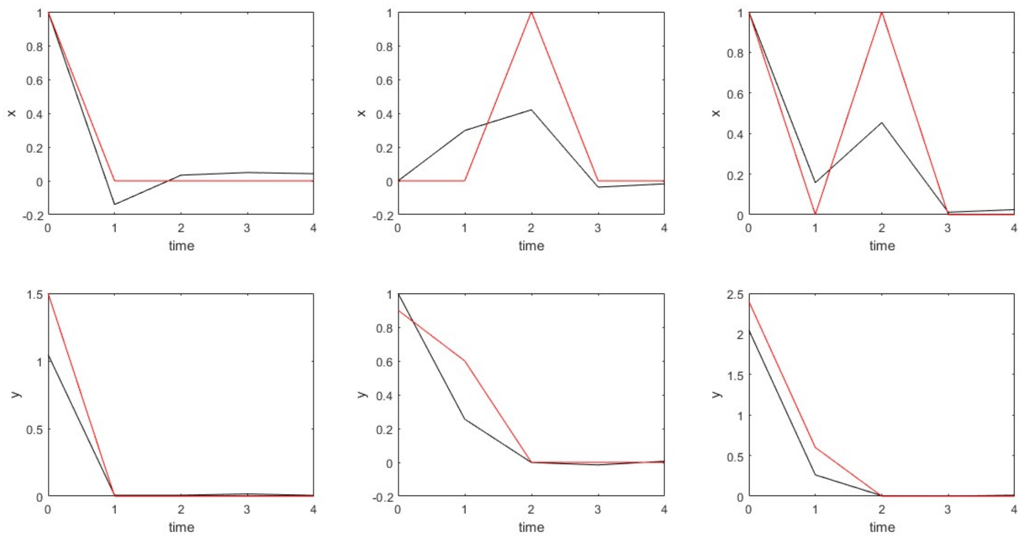

Figure 5: comparison of true and estimated IRFs

In the graph above the first row shows, from left to right, the true (red) and the average of estimated impulse responses (grey) of X due to a unitary change in eta, epsilon, and epsilon + eta. The second provides, analogously, plots for the estimated and true impulse responses of Y. The average of estimated impulse responses matches the true impulse responses for both X and Y. However, some slight discrepancies appear. By looking closely at the DGP, we notice a problem that might bias the estimation of the regressions. The issue resides in $\frac{\beta}{1-\beta}$ of the matrix;

$$\begin{bmatrix} 1 & L^2 \\ \frac{\beta}{1-\beta} & \frac{\beta^2}{1-\beta} + \beta L \end{bmatrix}$$

it means that y_t is correlated with the error term η_t, therefore in the regression of x_t, since we didn't control for a simultaneous effect of y_t, the estimates are slightly biased. Looking at the distributions of our estimates, and computing percentiles, we can see how they are not exactly centered around the true value of the impulse response. As an example, the following code provides the 5th and 95th percentile of the first parameter in the 2x2 upper left corner of the AD matrix, corresponding to the unitary effect of a change in η on x one step ahead. Estimates appear not to be centered around 0 (the true value),but slightly tilted to negative values.

```
c1 = zeros(1,N);a
for j = 1:N
    p = Imp1{j};
    c1(j) = p(1,1);
end
P1 = prctile(c1,5);
P2 = prctile(c1,95);
```