THE BOX – JENKINS PROCEDURE

Following the idea of Box and Jenkins (1970, 1976), the practical construction of a time series model can be developed in three steps: 1. Specification*; 2. Estimation; 3. Diagnostic checking. Once a particular model has been specified, estimated, and checked, it can be used for forecasting purposes (which, as we have seen, is one of the main objectives of time series analysis).

1. SPECIFICATION

To specify a model in the general class of ARIMA(p, d, q) models is to determine the most appropriate values of the parameters (p, d, q) [redacted]. If a time series is stationary (d=0), we can use ARMA(p, q) models; if it is non-stationary, we must transform it into a new time-series W_t which is stationary. To do this, as we have seen, we must apply a sufficient number of times the operator Δ to the original series Y_t. If, for instance, we find that

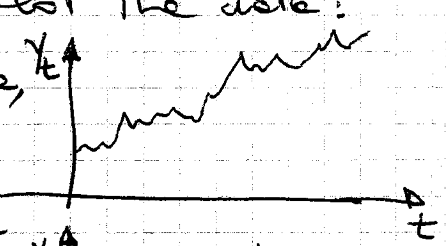
---

* Box and Jenkins use the term "identification" instead of "specification". As, in general, in econometrics analysis the term "identification" refers to ~~another~~ a different kind of problem, we will use the term "specification".

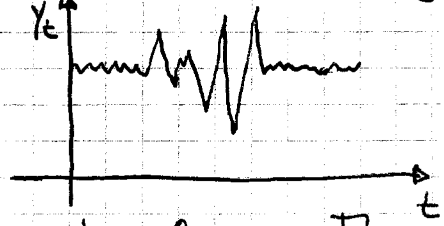
y_t is non-stationary, and that $w_t = \Delta y_t$ is still non-stationary, but $z_t = \Delta w_t = \Delta^2 y_t$ is stationary we will set $d=2$ and ~~go~~ go on to determine the values of the parameters (p,q) for an ARMA(p,q) on the new time-series z_t (which is equivalent to say that we will proceed to determine the value of (p,q) for an ARIMA$(p,2,q)$ on the time-series y_t).*

So, what we need is a number of tools to check the stationarity of a time-series. Only once we will have obtained a stationary time-series we can use ARMA(p,q) models.

•) A first useful thing to do is to plot the data:
- if ~~xxx~~ a trend is clearly detectable, the series is not stationary.



- If there is evidence that the variance is not constant, the series is also non stationary.



•) If there is no evidence of non-stationarity from the graph, we can consider the sample autocorrelation

---

*Notice that if x_t is a stationary time series, then also $y_t = \Delta x_t$ is stationary (and the same holds for any $w_t = \Delta^d x_t$ for $d = 1, 2, \ldots$). This means that the stationarity of $\Delta^d x_t$ is not sufficient for x_t to be ~~I(d)~~ integrated of order d.

function (also known as the correlogram).
We know that if a process is stationary the auto-correlation function should "decline relatively rapidly" to zero when K is increased, and thus we expect a similar behavior for the sample autocorrelation function. If the correlogram $\hat{\rho}_K$ does not fall off quickly as K increases, this is an indication of non stationarity.
After this preliminary qualitative analysis, a formal test migh be performed.

) UNIT ROOT TEST: A common test for non-stationarity is the Dickey-Fuller test. Let us examine its logic.
We know that a random walk is non-stationary integrated of order 1* ~~Assume that~~
Now, consider the simple AR(1) process on our time series Y_t:
$$Y_t = \delta + \phi Y_{t-1} + \varepsilon_t \qquad \varepsilon_t \sim WN$$
If $\phi = 1$, the above process will coincide with the random walk process: $Y_t = \delta + Y_{t-1} + \varepsilon_t$
which would imply that
our time-series Y_t is non stationary integrated of order 1.
In principle, thus, we might test whether ϕ is significantly different from 1 in order to ~~r~~ refuse the hypothesis that Y_t follows a random walk process.

---

* Indeed the process $W_t = \Delta Y_t$ is a white noise if Y_t is a random walk.

As, in general, it is easier to perform a test to check if a parameter is significantly different from zero, we can rearrange the AR(1) model as follows:

$$Y_t - Y_{t-1} = \delta + \phi Y_{t-1} - Y_{t-1} + \varepsilon_t$$

$$\Delta Y_t = \delta + (\phi - 1) Y_{t-1} + \varepsilon_t$$
$$\Delta Y_t = \delta + \alpha Y_{t-1} + \varepsilon_t$$

having defined $\alpha = \phi - 1$.
To say that ϕ is significantly different from one is to say that α is significantly different from zero.
So the null hypothesis will be $H_0: \alpha = 0$ *
to be tested against the alternative hypothesis $H_1: \alpha < 0$
In principle to perform such a test we might use an OLS regression and consider the t-test on the parameter α. However, Dickey and Fuller have shown that in this case, under the null hypothesis, the t-statistics has a different distribution (which implies that critical values are different).
So instead of referring to the t-test critical values, we will refer to the critical values calculated by Dickey and Fuller.

Notice that the test is unidirectional. Indeed in the alternative hypothesis we want the process to be stationary and we know that this will be true if $|\phi| < 1$.
Now, $|\phi| < 1 \Rightarrow \alpha < 0$.
To put it differently, imagine that $\alpha > 0$: this will imply that $\phi > 1$, but in this case the AR(1) process would still be non stationary, contradicting our purpose that in the alternative hypothesis the process is stationary.

Two observations on the Dickey-Fuller Test:

(1) A failure to reject the assumption of ~~same~~ non stationarity is only weak evidence in favor of the random walk hypothesis (This depends on the fact that the test is not very powerful).

(2) There are a number of variants of the test in which also a trend is included in the AR(1) model assumed as reference for ~~the~~ testing whether ϕ is significantly different from 1.
In addition, many econometric packages calculate more sophisticated versions of the Dickey-Fuller test, called Augmented Dickey Fuller tests. As the Dickey-Fuller test is not very powerful, it is, in general, a good idea to consider the results of all the available Dickey-Fuller and Augmented Dickey-Fuller tests.

Once the parameter d has been determined, and the series y_t has been opportunetely transformed to the new series $W_t = \Delta^d y_t$, ~~the~~ it is possible to [redacted] of the ARMA(p,q) for the time series W_t. To do this we can consider the sample autocorrelation function and the sample partial autocorrelation function of the time-series W_t. Having studied the behavior of the theoretical ACF and PACF of the different AR(p), MA(q), and ARMA(p,q) models, what we have to do now, is to understand to which of these models our time series W_t resembles more. ~~In~~ So if, for instance, we ~~can recognize~~ obtain a sample ACF and a sample PACF that resemble closely to the theoretical ACF and PACF of an ARMA(1,1), we will choose as a possible model to estimate the model ARMA(1,1).

In general, as the sample ACF and PACF are sample functions they will not exhibit the precise behavior of any specific model even in those cases in which, for some reason, we might assume to know exactly the particular parameters (p,q) of the generating process.

A good suggestion is then to proceed with tentative guesses and to leave open the possibility that a number of models might be useful to be estimated.*

Another thing to keep in mind is that although increasing the order of a model leads to a better fit of actual data, models with a lot of parameters might be more difficult to interpret economically.** In principle one should look for the model which has the smallest number of parameters that are sufficient to produce white noise residuals.

So if a model can be simplified and the results are almost as good, one might try to work with the simplified model.

---

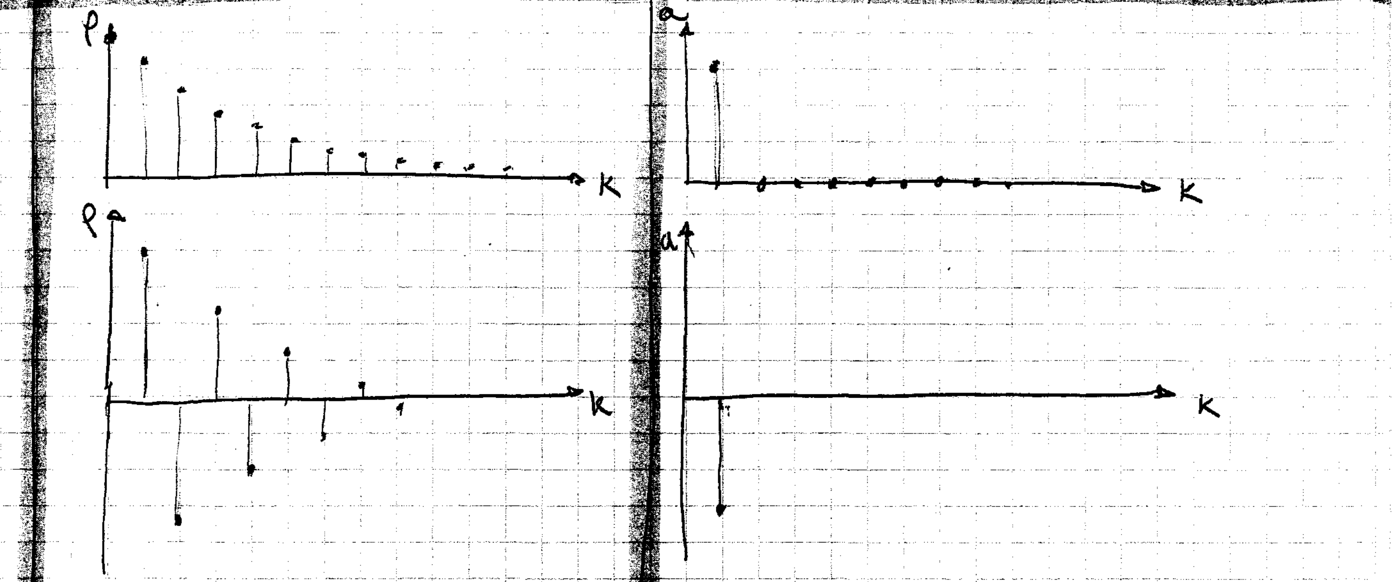
* Once we will have estimated these models, by analysing the diagnostic tests perhaps we will be able to restrict the number of possible models, by eliminating those that do not pass the tests.

** Notice also that by increasing the number of the parameters to estimate the degrees of freedom decrease.

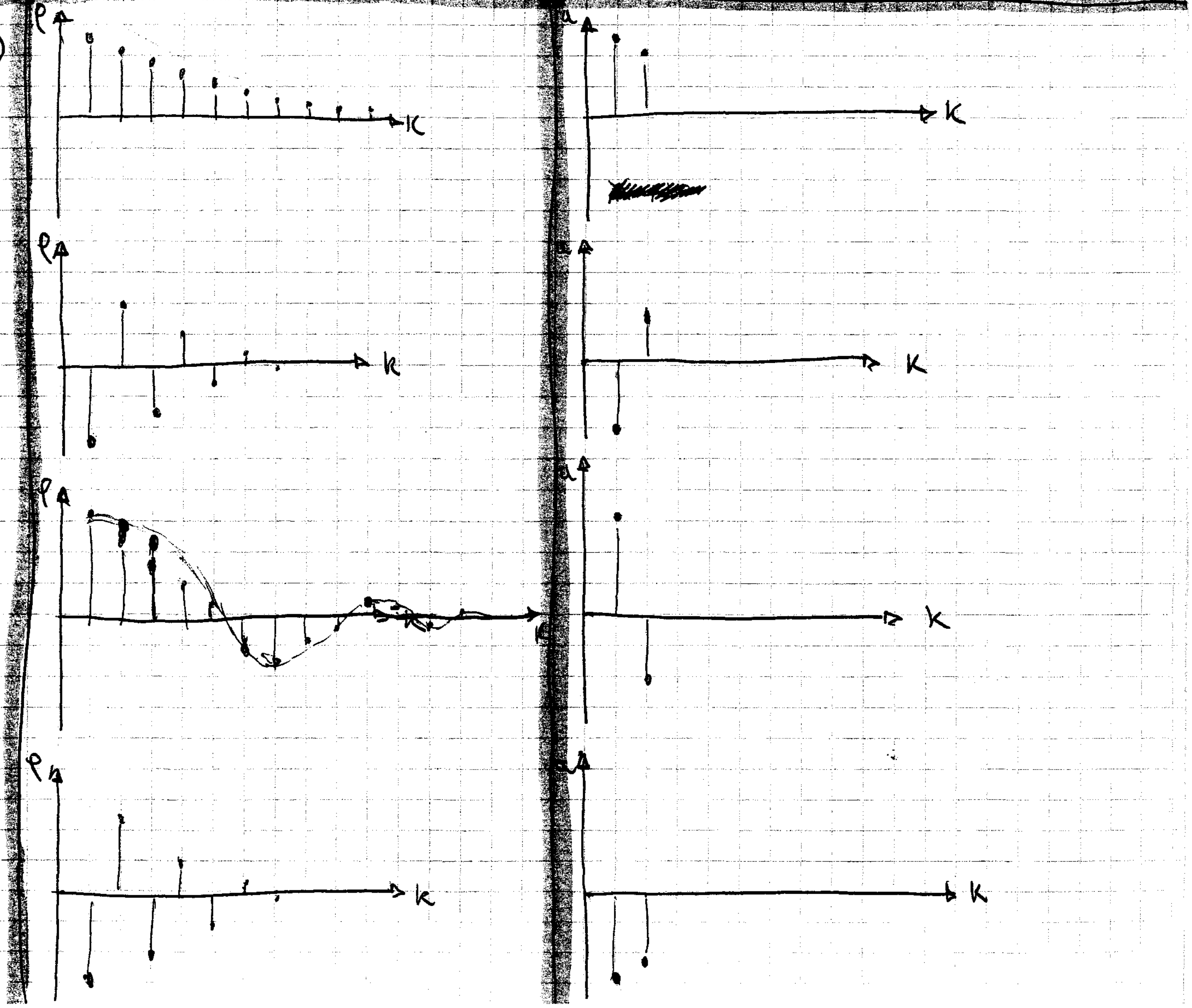
Having said this, let us now present compactly the behavior of the main ARMA(p,q) ACF and PACF.

ACF (ρ_k)	PACF (a_k)

(1)



(2)



To obtain the analogous configurations for the MA processes, simply invert, the ACF and the PACF of the ~~AR~~ ~~preceding~~

columns of

preceding figure -

As concerns the general case of AR(p) and MA(q) we can limit the characterisation of the ACF and the PACF to their qualitative features -

	ACF	PACF
AR(P) :	It decreases to zero in absolute value -	It has the first p terms different from zero and all the others equal to zero
MA(q) :	It has the first q terms different from zero and all the others equal to zero -	It decreases to zero in absolute value -

2. ESTIMATION

Once one or more models have been specified, we have to estimate them. The estimation method is different for models with no moving average component (pure AR models) and for models which contain a moving average component (pure MA models and ARMA models). Let us consider those two cases in sequence:

PURE AR MODELS: AR(p): $Y_t = \delta + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \varepsilon_t$

Remember that ~~whereas~~ in the specification procedure what we tried to do was to build a model such that the residuals ε_t ~~were~~ were as close as possible to the WN assumption.*

As we know, when we have a model with the WN residuals the best linear unbiased estimator is provided by the OLS method. In this case, then, we have just to estimate the AR (P) model by mean of OLS.

---

*To put it differently, what we did was to explain all the systematic behavior of the time series by the parameters $\delta, \phi_1, \dots \phi_p$ and to assume that the non-systematic behavior was a pure random component of WN type (i.e. completely unpredictable).

RE MA and ARMA MODELS.

Let us consider directly the general case of an ARMA(p, q) model (taking in mind that if we have a pure MA(q) model we will have all the ϕ_K, $K = 1, 2, \ldots p$ parameters equal to zero $[\phi_1 = \phi_2 = \ldots = \phi_p = 0]$.

$$\text{ARMA}(p, q): \quad Y_t = \delta + \phi_1 Y_{t-1} + \ldots + \phi_p Y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} + \ldots - \theta_q \varepsilon_{t-q}$$

The parameters $(\delta, \phi_1, \ldots, \phi_p, \theta_1, \ldots \theta_q)$ can be estimated by minimising

$$S(\delta, \phi_1 \ldots \phi_p, \theta_1, \ldots \theta_q) = \sum_{t=1}^{T} \varepsilon_t^2 = \sum_{t=1}^{T} \left(Y_t - \delta - \phi_1 Y_{t-1} - \ldots - \phi_p Y_{t-p} + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q} \right)^2$$

At first glance, minimisation of the above equation appears to be a linear least squares problem. However, nonlinearities arise because there is a feed-back from $\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots, \varepsilon_{t-q}$ to ε_t when the equation is being computed.

In this case, thus, we must use non linear least squares estimates. *

Every time an econometrics package estimates a model including a MA component it computes non-linear least squares estimates. These estimates are the result of iterative procedures that explores the set of the possible values of the parameters to be estimated $(\delta, \phi_1, \ldots \phi_p, \theta_1, \ldots \theta_q)$ with the objective of minimising the function $S(\delta, \phi, \ldots \phi_p \theta_1 \ldots \theta_q)$

3. DIAGNOSTIC CHECKING

After a time-series model has been specified and its parameters estimated, one must test whether the original specification was correct. If this is the case, the residuals of our model should resemble a WN process. Remember, indeed, that the estimating procedure was based precisely on this assumption. In order to build the test, as usual, we have now to assume also a particular distribution for the residuals. So in addition to the hypothesis of WN we assume that the residuals are normally distributed.

Under this assumption we can examine the sample ACF of the residuals and test if it is close to the ACF of a WN. Let us denote the sample ACF of the residuals as $\hat{r}_k$.

$$\hat{r}_k = \frac{\sum \hat{\varepsilon}_t \hat{\varepsilon}_{t-k}}{\sum \hat{\varepsilon}_t^2} \qquad k = 1, 2, \ldots$$

---

n general these procedures are based on algorithm that start from articular values of the parameters and try to modify them in the ttempt to minimise the function S. This allows to consider nly a subset of all the possible values of the parameters (the whole et is, in general, impossible to be explored). On the other hand, hese algoritms have to converge to particular values of the parameters $\delta, \phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q)$ so that the function $S(\delta, \phi, \ldots, \phi_p, \theta_1, \ldots, \theta_q)$ is minimized and this does not necessarily occur. Non-linear least squares estimates are obtained only when convergence occur.

We have Two Tests:

(1) r_k can be shown To have an approximate normal distribution with mean zero and variance $\frac{1}{T}$. Consequently one can test The significance for individual r_k, by seeing if The calculated value $\hat{r}_k$ is significantly different from zero.

$r_k = 0$

$r_k \neq 0$

$N\left(0, \frac{1}{T}\right)$

Remember That for a WN process we expect the ACF To be close To zero for any $K > 0$.
As the convergence To the normal distribution is only for great values of K, This test should be conducted, let us say, for any K greater Than 5.

(2) A second Type of Test considers the joint null hypothesis that all $\hat{r}_k$ are equal to zero. Under this circumstance, The statistic

$r_k = 0 \ \forall k = 1, 2, .. K$

$r_k \neq 0$ for some k

$$Q = T \sum_{k=1}^{K} \hat{r}_k^2$$

χ^2_{K-p-q}

is approximately distributed as a χ^2 with $K-p-q$ degrees of freedom. *

Notice That although it has nothing To do with the checking of The appropriateness of our model, sometimes we might be interested in testing The significance of The parameters of The ACF or of The PACF of our <u>time series</u> (on The contrary The above tests are conducted on The sample ACF of The <u>residuals</u>).
In This case we can use basically The same kinds of →

test, by observing ~~xxxx~~ what follows:

- As concerns the ACF: (1) $H_0: \rho_K = 0$ $\qquad \rho_K \sim N\left(0, \frac{1}{T}\right)$
  $H_1: \rho_K \neq 0$

  (2) $H_0: \rho_K = 0 \quad \forall K$ $\qquad Q = T \sum_{K=1}^{K} \rho_K^2 \sim \chi^2_K$
  $H_1: \rho_K \neq 0$ for some K

- As concerns the PACF: (1) $H_0: a_K = 0$ $\qquad a_K \sim N\left(0, \frac{1}{T}\right)$
  $H_1: a_K \neq 0$

  (2) There is no joint test for $a_K = 0$ for any K because $a_1, a_2, \ldots$ are estimates from different models.

MULTIVARIATE MODELS (5)

In The preceding chapters we have already discussed multivariate models. We now have to discuss this subject by Taking account of the fact That our variables are Time-series.

A first problem that can arise is That if we regress one nonstationary Time-series against another non-stationary Time-series we might have spurious results. In This case, OLS might produce an apparently highly significant relationship between the variables even when The Two variables are completely independent. For instance we might obtain a good fit of actual data (a high R^2), significant t-tests and so on.* But This might be due simply to The fact That, for instance, The variables might both have a trend, or might both be RWs. Under these conditions OLS do not provide consistent estimates and The distribution of The t-tests might be different from the one assumed under the hypothesis of WN ~~errors~~ errors.

One possibility might Then be To Transform our series to stationary series and To work with The latter. However, differencing may result in a loss of information about The long-run relationship between The variables.

---

* One test That should Indicate The possibility of spurious regression is The Durbin-Watson, since The residuals ~~show~~ should hardly fit The WN assumption.

If, for instance we want to analyse the relationship between income and consumption and we transform these series in order to obtain stationarity, the regression of the two series will not give us any idea about the propensity to consume*, but will give us an idea only about how much changes consumption when income changes.

*(i.e. what share of income is consumed).

Fortunately, sometimes we can manage such a problem without any loss of information. Indeed, it can happen that although two variables are not stationary, a linear combination of them might be stationary. If this is the case we say that these two variables are [redacted] and we call the parameters of the linear combination cointegrating parameters.*

To test whether two (or more) time series are cointegrated is simple: we run the OLS regression $x_t = \alpha + \beta y_t + \varepsilon_t$ and we test whether the residuals $\hat{\varepsilon}_t$ are stationary. We already know a test for non-stationarity; the Dickey-Fuller test. So what we do is to test the hypothesis that $\hat{\varepsilon}_t$ is not stationary, i.e. the hypothesis of

---

*In practice we normalise the first parameter to one, so that with a two-variables model we have just one cointegrating parameter: for instance if x_t and y_t are cointegrated, it means that there exist α and β such that: $\alpha x_t + \beta y_t$ is stationary. But then we can consider the equivalent linear combination $x_t + \gamma y_t$ (where $\gamma = \frac{\beta}{\alpha}$), which

no cointegration.
The conclusion of this discussion is that if we have nonstationary cointegrated variables we can use the simple OLS to obtain consistent estimates.
Our purpose, now is to generalise the analysis of ARIMA(p, d, q) models to the multivariate case. Such a generalisation leads to very complicated models. Thus most econometric research has developed by generalising only a limited representation of this class of models, namely the AR(p) models. The generalisation of AR(p) model to the multivariate case is done with Vector Autoregressive models (VAR).
The main differences between the AR and the VAR representations are the following:
(1) Y_t is now a vector $(m \times 1)$;
(2) the AR parameters $\phi_1, \dots, \phi_p$ are now matrices $(m \times m)$;
(3) ε_t is now a vector $(m \times 1)$.
With these conventions a VAR(p) model can be written exactly as an AR(p) model:

$$VAR(p): \quad Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t \quad \varepsilon_t \sim IID(0, \Sigma) \quad \forall t.$$

the matrix Σ ~~shows~~ is $(m \times m)$ and shows variances and ~~covariances~~ contemporaneous covariances for the individual elements of ε_t.
This model generalises the univariate AR(p) to the case of

Let us consider the following example with just to variables ($m=2$).

$$\begin{pmatrix} Y_t^1 \\ Y_t^2 \end{pmatrix} = \begin{pmatrix} \phi_{11}^1 & \phi_{12}^1 \\ \phi_{21}^1 & \phi_{22}^1 \end{pmatrix} \begin{pmatrix} Y_{t-1}^1 \\ Y_{t-1}^2 \end{pmatrix} + \dots + \begin{pmatrix} \phi_{11}^p & \phi_{12}^p \\ \phi_{21}^p & \phi_{22}^p \end{pmatrix} \begin{pmatrix} Y_{t-p}^1 \\ Y_{t-p}^2 \end{pmatrix} + \begin{pmatrix} \varepsilon_t^1 \\ \varepsilon_t^2 \end{pmatrix}$$

$$Y_t = \phi_1 \cdot Y_{t-1} + \dots + \phi_p \cdot Y_{t-p} + \varepsilon_t$$

We can also write the system in terms of two equations:

$$\begin{cases} Y_t^1 = \phi_{11}^1 Y_{t-1}^1 + \phi_{12}^1 Y_{t-1}^2 + \dots + \phi_{11}^p Y_{t-p}^1 + \phi_{12}^p Y_{t-p}^2 + \varepsilon_t^1 \\ Y_t^2 = \phi_{21}^1 Y_{t-1}^1 + \phi_{22}^1 Y_{t-1}^2 + \dots + \phi_{21}^p Y_{t-p}^1 + \phi_{22}^p Y_{t-p}^2 + \varepsilon_t^2 \end{cases}$$

Perhaps rearranging the terms, the representation is more intuitive:

$$\begin{cases} Y_t^1 = \underbrace{\phi_{11}^1 Y_{t-1}^1 + \phi_{11}^2 Y_{t-2}^1 + \dots + \phi_{11}^p Y_{t-p}^1}_{\text{lags of } Y^1} + \underbrace{\phi_{12}^1 Y_{t-1}^2 + \phi_{12}^2 Y_{t-2}^2 + \dots + \phi_{12}^p Y_{t-p}^2}_{\text{lags of } Y^2} + \varepsilon_t^1 \\ Y_t^2 = \overbrace{\phi_{21}^1 Y_{t-1}^1 + \phi_{21}^2 Y_{t-2}^1 + \dots + \phi_{21}^p Y_{t-p}^1} + \overbrace{\phi_{22}^1 Y_{t-1}^2 + \phi_{22}^2 Y_{t-2}^2 + \dots + \phi_{22}^p Y_{t-p}^2} + \varepsilon_t^2 \end{cases}$$

To say it with words: we have two variables Y^1 and Y^2 and we regress each of them on itself (lagged $1,2,\dots p$) and on the other one (lagged $1,2,\dots p$).
In practice, what we will do is to estimate each single equation individually, by observing that if $\varepsilon_t \sim \text{iid}(0,\Sigma)$, then each $\varepsilon_t^i \sim WN$, $i=1,2,\dots m$, so that we can use OLS to obtain consistent estimates of the ϕ parameters.

APPENDIX - INVERTIBILITY

Sometimes it is useful to consider a ~~pa~~ MA(q) model as a particular form of an AR model. Let us analyze when such an interpretation of the MA models is possible and what it implies.

Consider the simplest case of a MA(1).

$$MA(1): \quad Y_t = \delta + \varepsilon_t - \theta\varepsilon_{t-1} \qquad \varepsilon_t \sim WN$$

We already know that this process, like all MA processes, is stationary for any value of the parameters θ, δ.

Now consider a particular AR(∞) process with geometrically declining parameters:

$$AR(\infty): \quad Y_t = \delta + \theta\delta + \theta^2\delta + \theta^3\delta + \dots + \theta Y_{t-1} + \theta^2 Y_{t-2} + \theta^3 Y_{t-3} + \dots + \varepsilon_t$$

$$\varepsilon_t \sim WN.$$

$$\Rightarrow \quad Y_{t-1} = \delta + \theta\delta + \theta^2\delta + \theta^3\delta + \dots + \theta Y_{t-2} + \theta^2 Y_{t-3} + \dots \varepsilon_{t-1}$$

$$\Rightarrow \quad \theta Y_{t-1} = \theta\delta + \theta^2\delta + \theta^3\delta + \dots + \theta^2 Y_{t-2} + \theta^3 Y_{t-3} + \dots + \theta\varepsilon_{t-1}$$

~~Now consider $Y_t - \theta Y_{t-1}$ by taking them in the first and the third expressions of our AR(∞) model:~~

~~$AR(\infty): Y_t - \theta Y_{t-1} = \delta + \dots$~~

Now, subtract from equation (2) the value θY_{t-1} on both the left and the right hand sides:

$$) \quad Y_t - \theta Y_{t-1} = \delta + \theta\delta + \theta^2\delta + \dots + \theta^2 Y_{t-2} + \theta^3 Y_{t-3} + \dots + \varepsilon_t$$

Now substitute the value of θY_{t-1}, which we have calculated in equation (4) and simplify the terms which appear on both the ~~l~~ left and the ~~right~~ right hand sides:

AR(∞): $\quad Y_t = \delta + \varepsilon_t - \theta \varepsilon_{t-1}$

which coincides with the model MA(1), which we considered at the beginning. This proves that any MA(1) can be seen as an AR(∞) with a particular structure of the parameters: in particular we have assumed that such parameters decline geometrically and that

$$\lim_{s \to \infty} \theta^s = 0.$$

These conditions mean that the parameter θ of our MA(1) model must satisfy the following condition:

$$|\theta| < 1.$$

Indeed if this condition is violated: $\lim_{s \to \infty} \theta^s \neq 0$ and the MA(1) process cannot be written as an AR(∞).

On the other hand if we have an AR(1) process, we can see it as a particular MA(∞) process. Let us see how.

$$\text{AR}(1): \quad Y_t = \delta + \phi Y_{t-1} + \varepsilon_t \qquad \varepsilon_t \sim WN$$

$$Y_t = \delta + \phi(\delta + \phi Y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t$$

$$= \delta + \phi\delta + \phi^2 Y_{t-2} + \phi\varepsilon_{t-1} + \varepsilon_t$$

$$Y_t = \delta + \phi\delta + \phi^2(\delta + \phi Y_{t-3} + \varepsilon_{t-2}) + \phi\varepsilon_{t-1} + \varepsilon_t$$

$$= \delta + \phi\delta + \phi^2\delta + \phi^3 Y_{t-3} + \phi^2\varepsilon_{t-2} + \phi\varepsilon_{t-1} + \varepsilon_t$$

y iterating these substitutions at the infinite we will we that the lagged variables of Y_t will disappear on e right hand side and we will have only a number : constants (whose summation can be indicated by δ^*) and terms $\varepsilon_t + \phi\varepsilon_{t-1} + \phi^2\varepsilon_{t-2} + \dots$, which is precisely a

MA(∞) process of This form:

$$MA(\infty): Y_t = \delta^* + \varepsilon_t + \phi\varepsilon_{t-1} + \phi^2\varepsilon_{t-2} + \phi^3\varepsilon_{t-3} + \dots$$

$$\delta^* = \delta + \phi\delta + \phi^2\delta + \phi^3\delta + \dots$$

Notice That we have not asked any conditions on The parameters in order To invert The AR(1) To This particular MA(∞). This means That an AR(1) process in invertible To a MA(∞) independently from The values of δ and ϕ.

These results can be generalised:

- Any MA(q) process can be inverted To an AR(∞) process (with a particular structure on The parameters $\phi_1, \phi_2, \dots$) if some conditions on The parameters $\theta_1, \dots \theta_q$ hold.
- Any AR(p) process can be inverted To a MA(∞) process (with a particular structure on The parameters $\theta_1, \theta_2, \dots$) without any conditions on The parameters $\phi_1, \dots \phi_p$ being required.

So, ~~in general~~ to summarise invertibility and stationarity conditions we have The following results:

1) Invertibility of AR(p) needs no conditions on $\phi_1, \dots \phi_p$
2) Stationarity of AR(p) needs particular conditions on $\phi_1, \dots \phi_p$
3) Invertibility of MA(q) needs particular conditions on $\theta_1, \dots \theta_q$
4) Stationarity of MA(q) needs no conditions on $\theta_1, \dots \theta_q$.