Chapter 7. Estimation of VARMA Models

In this chapter maximum likelihood estimation of the coefficients of a VARMA model is considered. Before we can proceed to the actual estimation a unique set of parameters must be specified. In this context the problem of nonuniqueness of a VARMA representation becomes of importance. This problem is often referred to as the identification problem, that is, the problem of identifying a unique structure among many equivalent ones. It is treated in Section 7.1. In Section 7.2 the Gaussian likelihood function of a VARMA model is considered. A numerical algorithm for maximizing it and thus for computing the actual estimates is discussed in Section 7.3. The asymptotic properties of the ML estimators are the subject of Section 7.4. Forecasting with estimated processes and impulse response analysis are dealt with in Sections 7.5 and 7.6, respectively.

7.1 The Identification Problem

7.1.1 Nonuniqueness of VARMA Representations

In the previous chapter we have considered K-dimensional, stationary processes \( y_t \) with VARMA \((p, q)\) representation

\[
y_t = A_1 y_{t-1} + \cdots + A_p y_{t-p} + u_t + M_1 u_{t-1} + \cdots + M_q u_{t-q}. \tag{7.1.1}
\]

Since the mean term is of no importance for the presently considered problem we have set it to zero. Therefore no intercept term appears in (7.1.1). This model can be written in lag operator notation as

\[
A(L)y_t = M(L)u_t, \tag{7.1.2}
\]

where

\[
A(L) := I_K - A_1 L - \cdots - A_p L^p \quad \text{and} \quad M(L) := I_K + M_1 L + \cdots + M_q L^q.
\]

Assuming that the VARMA representation is stable and invertible the well-defined process described by the model (7.1.1) or (7.1.2) is given by

\[
y_t = \sum_{i=0}^{\infty} \Phi_i u_{t-i} = \Phi(L)u_t = A(L)^{-1}M(L)u_t.
\]

In practice it is sometimes useful to consider a slightly more general type of
VARMA models by attaching nonidentity coefficient matrices to \( y_t \) and \( u_t \), that is, one may want to consider representations of the type

\[
A_0 y_t = A_1 y_{t-1} + \cdots + A_p y_{t-p} + M_0 \bar{u}_t + M_1 \bar{u}_{t-1} + \cdots + M_q \bar{u}_{t-q}.
\] (7.1.3)

Such a form may be suggested by subject matter theory which may imply instantaneous effects of some variables on other variables. It will also turn out to be useful in finding unique structures for VARMA models. By the specification (7.1.3) we mean the well-defined process

\[
y_t = (A_0 - A_1 L - \cdots - A_p L^p)^{-1}(M_0 + M_1 L + \cdots + M_q L^q)\bar{u}_t.
\]

Such a process has a standard VARMA\((p, q)\) representation with identity coefficient matrices attached to the instantaneous \( y_t \) and \( u_t \) if \( A_0 \) and \( M_0 \) are nonsingular. To see this we left-multiply (7.1.3) by \( A_0^{-1} \) and define \( u_t = A_0^{-1} M_0 \bar{u}_t \) which gives

\[
y_t = A_0^{-1} A_1 y_{t-1} + \cdots + A_0^{-1} A_p y_{t-p} + u_t + A_0^{-1} M_1 M_0^{-1} A_0 u_{t-1} + \cdots
\]

\[
+ A_0^{-1} M_q M_0^{-1} A_0 u_{t-q}.
\]

Redefining the matrices appropriately this, of course, is a representation of the type (7.1.1) with identity coefficient matrices at lag zero which describes the same process as (7.1.3). The assumption that both \( A_0 \) and \( M_0 \) are nonsingular does not entail any loss of generality as long as none of the components of \( y_t \) can be written as a linear combination of the other components. We call a stable and invertible representation as in (7.1.1) a \textit{VARMA representation in standard form} or a \textit{standard VARMA representation} to distinguish it from representations with nonidentity matrices at lag zero as in (7.1.3). This discussion shows that VARMA representations are not unique, that is, a given process \( y_t \) can be written in standard form or in nonstandard form by just left-multiplying by any nonsingular \((K \times K)\) matrix. We have encountered a similar problem in dealing with finite order VAR processes which can, for instance, be written in standard form and in recursive form. However, once we consider standard VAR models only we have unique representations. This is in sharp contrast to the presently considered VARMA case where in general a standard form is not a unique representation as we will see shortly.

It may be useful at this stage to emphasize what we mean by equivalent representations of a process. Generally, two representations of a process \( y_t \) are equivalent if they give rise to the same realizations (except on a set of measure zero) and thus to the same multivariate distributions of any finite subcollection of variables \( y_t, y_{t+1}, \ldots, y_{t+h} \) for arbitrary integers \( t \) and \( h \). Of course, this just says that equivalent representations really represent the same process. If \( y_t \) is a zero mean process with canonical MA representation

\[
y_t = \sum_{i=0}^{\infty} \Phi_i u_{t-i}, \quad \Phi_0 = I_K,
\]

\[
= \Phi(L)u_t,
\] (7.1.4)