Chapter 6

Maximum Likelihood Estimation

In the last chapter attention was given to the determination of the state vector \( \xi \) for given observations \( Y \) and known parameters \( \lambda \). In this chapter the maximum likelihood estimation of the parameters \( \lambda = (\theta', \rho', \xi_0')' \) of an MS–VAR model is considered. The aim of this chapter is (i.) to provide the reader with an introduction to the methodological issues of ML estimation of MS–VAR models in general, (ii.) to propose with the EM algorithm an estimation technique for all discussed types of the MS–VAR models, (iii.) to inform the reader about alternative techniques which can be used for special purposes or model extensions and (iv.) to give some basic asymptotic results.

Thus, this chapter is partly a survey, partly an interpretation, and partly a new contribution; preliminaries for the ML estimation are considered in the following two sections. Section 6.1 gives three alternative approaches to formulate the likelihood function of MS–VAR models which it will be seen, have turned out to be useful. Section 6.2 discusses the identifiability of MS(M)-VAR(\( p \)) models. An identifiability result for hidden Markov-chain models provided by Leroux [1992] is extended to our augmented setting. In Section 6.3 the normal equations of ML estimation of MS–VAR models are derived. At the center of interest is the EM algorithm which has been suggested by Hamilton [1990] for the statistical analysis of time series subject to changes in regimes. In the literature the regressions involved with the EM algorithm are developed only for vector systems without autoregressive dynamics. We analyze the critical points; in particular, we relax the limitation in the literature to MSI(M)-VAR(0) models thus allowing the estimation of genuine vector autoregressive models. It is shown that the implementation of the EM algorithm to MS(M)-VAR(\( p \)) models causes some problems. Therefore the discussion is restricted to an MS regression model, but one which captures all MSI specifications.
A concrete discussion of the ML estimation of the various model types via the EM algorithm is left for Chapter 9. Extensions and alternatives which have been proposed in the literature are considered in Section 6.5. In the closing Section 6.6, the asymptotic properties of the ML estimation of MS-VAR models are discussed; in particular, procedures for the estimation of the variance-covariance matrix of the ML estimates are suggested.

6.1 The Likelihood Function

In econometrics the so-called Markov model of switching regressions considered by GOLDFELD & QUANDT [1973]

\[ y_t = x_t' \beta_m + u_{mt}, u_{mt} \sim \text{NID}(0, \sigma_m^2) \text{ for } m = 1, 2 \]

has been one of the first attempts to analyze regressions with Markovian regime shifts. GOLDFELD & QUANDT claimed to derive maximum likelihood estimates by maximizing their “likelihood” function, which would be in terms of our model

\[ Q(\theta, \rho, \xi_0) = \prod_{t=1}^{T} \eta_t(\theta)' \xi_{t|0}(\rho, \xi_0), \]

where \( \eta_t \) is again an \((M \times 1)\) vector collecting the conditional densities \( p(y_t|Y_{t-1}, \theta_m), m = 1, \ldots, M \), and \( \xi_{t|0} = \mathbf{F}^t \xi_0 \) are the unconditional regime probabilities.

By using this function of prior regime probabilities \( \xi_{t|0}(\rho, \xi_0) \) which can be approximated by the ergodic probabilities \( \tilde{\xi}(\rho) \) for sufficiently large \( t \) instead of the “posterior” inference \( \hat{\xi}_{t|t-1} \), GOLDFELD & QUANDT are not required to provide filtering procedures to reconstruct the time-path of regimes. The model’s parameters are estimated by numerical methods.

Unfortunately, the function \( Q(\theta, \rho, \xi_0) \) is not the likelihood function as pointed out by COSSLETT & LEE [1985]. However, equipped with the results of Chapter 5 it is possible to derive the likelihood function as a by-product of the BLHK filter:

\[ L(\lambda|Y) := p(Y_T|Y_0; \lambda) \]