Notation from Lutkepohl 2005:

(Section 2.1.1., p.13)

The object of interest in the following is the VAR(p) model (VAR model of order p),

$$y_t = \nu + A_1 y_{t-1} + \cdots + A_p y_{t-p} + u_t, \quad t = 0, \pm 1, \pm 2, \ldots,$$

where

- $y_t = (y_{1t}, \ldots, y_{Kt})'$ is a $(K \times 1)$ random vector,
- the A_i are fixed $(K \times K)$ coefficient matrices,
- $\nu = (\nu_1, \ldots, \nu_K)'$ is a fixed $(K \times 1)$ vector of intercept terms allowing for the possibility of a nonzero mean $E(y_t)$,
- $u_t = (u_{1t}, \ldots, u_{Kt})'$ is a K-dimensional white noise or innovation process:
  - $E(u_t) = 0$,
  - $E(u_t u_t') = \Sigma_u$,
  - $E(u_t u_s') = 0$ for $s \neq t$,
  - The covariance matrix Σ_u is assumed to be nonsingular if not otherwise stated.

It is assumed that a time series $y_1, \ldots, y_T$ of the y variables is available, that is, we have a sample of size T for each of the K variables for the same sample period. In addition, p presample values for each variable, $y_{p+1}, \ldots, y_0$, are assumed to be available. Partitioning a multiple time series into sample and presample values is convenient in order to simplify the notation.

$$\begin{aligned}
Y &:= (y_1, \ldots, y_T) && (K \times T) \\
B &:= (v, A_1, \ldots, A_p) && (K \times (Kp+1)) \\
Z_t &:= \begin{bmatrix} 1 \\ y_t \\ \vdots \\ y_{t-p+1} \end{bmatrix} && ((Kp+1) \times 1)
\end{aligned}$$