

Let us consider a matrix  $\mathbf{X}$  of  $p$  numerical variables observed on  $n$  observations. Two objectives for the clustering of the variables may be distinguished. Firstly, if the aim is to separate variables that are highly, but negatively, correlated each cluster must be defined locally around a latent variable that has the same orientation as the variables in the cluster. On the contrary, if the aim is to group together correlated variables in the same cluster, whatever the sign of the correlation coefficient, each cluster is to be defined directionally around a new axis.

Both cases are associated with a maximization problem [17]. The criterion to be maximized (1) involves the covariance between the variables and the latent variables for local groups:

$$S = \sum_{k=1}^K \sum_{j=1}^p \delta_{kj} \text{cov}(x_j, t_k) \quad \text{with } t_k^t t_k = 1 \quad (1)$$

For directional groups, the criterion in (2) involves the squared covariance:

$$T = \sum_{k=1}^K \sum_{j=1}^p \delta_{kj} \text{cov}^2(x_j, t_k) \quad \text{with } t_k^t t_k = 1 \quad (2)$$

In (1) and (2),  $K$  stands for the number of groups in the partition  $\{G_1, \dots, G_K\}$ .  $x_j$  represents the  $j^{\text{th}}$  variable and  $t_k$  the latent variable associated with the  $k^{\text{th}}$  group.  $\delta_{kj}$  is an indicator equal to 1 if the variable  $j$  belongs to the group  $k$ , and equal to 0 otherwise.

The case of directional groups was considered in this case study. It is noteworthy that, in this (2), the optimum value of  $T$  is obtained when the latent variable  $t_k$  within each cluster  $G_k$  is the first normalized principal component of the variables belonging to this cluster.

The maximization of the CLV criteria for  $K$  fixed is achieved by an iterative alternating procedure whose monotonicity can be proven. Moreover, partitions of growing size, from one (all the variables into only one group) to  $p$  groups (each variable in its own group) may be extracted in the course of an hierarchical algorithm. The CLV approach has been implemented in an R-package, ClustVarLV, freely available on the CRAN website [18].