Symbols

General Notation

| \( |M| \) | number of objects in a finite set \( M \)
| \( || \cdot || \) | Euclidean distance

Sets

\( \mathbb{N} \) | natural numbers
\( \mathbb{R}, \mathbb{R}_0^+ \) | real numbers, positive real numbers including zero
\( A_j, \mathcal{A} \) | attribute, finite set of attributes
\( \Omega \) | direct product of attributes
\( V \) | data set
\( C_i, \mathcal{C} \) | cluster, \( k \)-cluster set (finite set of disjoint clusters)
\( \wp(\Omega) \) | power set of \( \Omega \)
\( I, J \) | index subset
\( \mathcal{A}(J) \) | reduced set of attributes (only \( A_j \) with \( j \in J \))
\( \Omega(J) \) | direct product of attributes in \( \mathcal{A}(J) \)
\( V(J) \) | canonical projection of \( V \) on \( \Omega(J) \)
\( \Theta_s \) | partition
\( \Theta \) | decomposition (finite set of disjoint partitions)
\( B_j \) | subset of attribute \( A_j \)
\( B, \Delta_s \) | box
\( \Delta, \Delta_I \) | set of boxes, reduced set of boxes (only \( \Delta_s \) with \( s \in I \))
\( W \) | codebook
\( \mathcal{C}(W) \) | compressed clustering
\( \hat{\mathcal{C}} \) | extended clustering
\( \Theta_W \) | decomposition based on SOM codebook
\( \hat{W}_s \) | codebook box

Matrices

\( S \) | stochastic matrix
\( \hat{S} \) | coupling matrix
\( D \) | weighting matrix
### Variables

- $q$: dimension of $\Omega$
- $v, v_i$: data object in $V$
- $n$: number of data objects in $V$
- $k$: number of clusters
- $v(J)$: projection of $v$ on $\Omega(J)$
- $n_k$: number of decomposition partitions
- $w_s$: codebook vector
- $T, L$: time steps
- $k$: upper bound of $n_k$
- $z_s$: grid position of neuron $s$
- $l_i, r_i$: left and right boundaries of interval in $\mathbb{R}$
- $X$: random variable
- $u$: average number of codebook updates
- $\lambda_i, Y_i$: eigenvalue, eigenvector

### Functions

- $f$: frequency function
- $h$: homogeneity function
- $h_{\max}(V)$: maximal value of homogeneity function in $V$
- $\Gamma_{f,h}$: weighted intra-cluster homogeneity
- $d$: distance function
- $h_d$: homogeneity function based on distance function
- $S$: conditional transition probability function
- $\hat{S}$: set extension of $S$
- $h_S$: homogeneity function based on transition probability function
- $\chi_M$: characteristic function of set $M$
- $r$: membership rule (set)
- $\hat{\varrho}_{f,h}$: decomposition error
- $\hat{f}$: compressed frequency function
- $\hat{h}$: compressed homogeneity function
- $\hat{f}$: set extension of $f$
- $\hat{h}$: set extension of $h$
- $\rho$: probability density function
- $P_\rho$: probability function corresponding to $\rho$
- $\alpha$: learning rate
- $\gamma$: neighborhood radius function
- $\eta$: grid distance function
- $E(X)$: conditional expectation value of $X$
- $P$: weighted homogeneity function
- $\hat{P}$: set extension of $P$