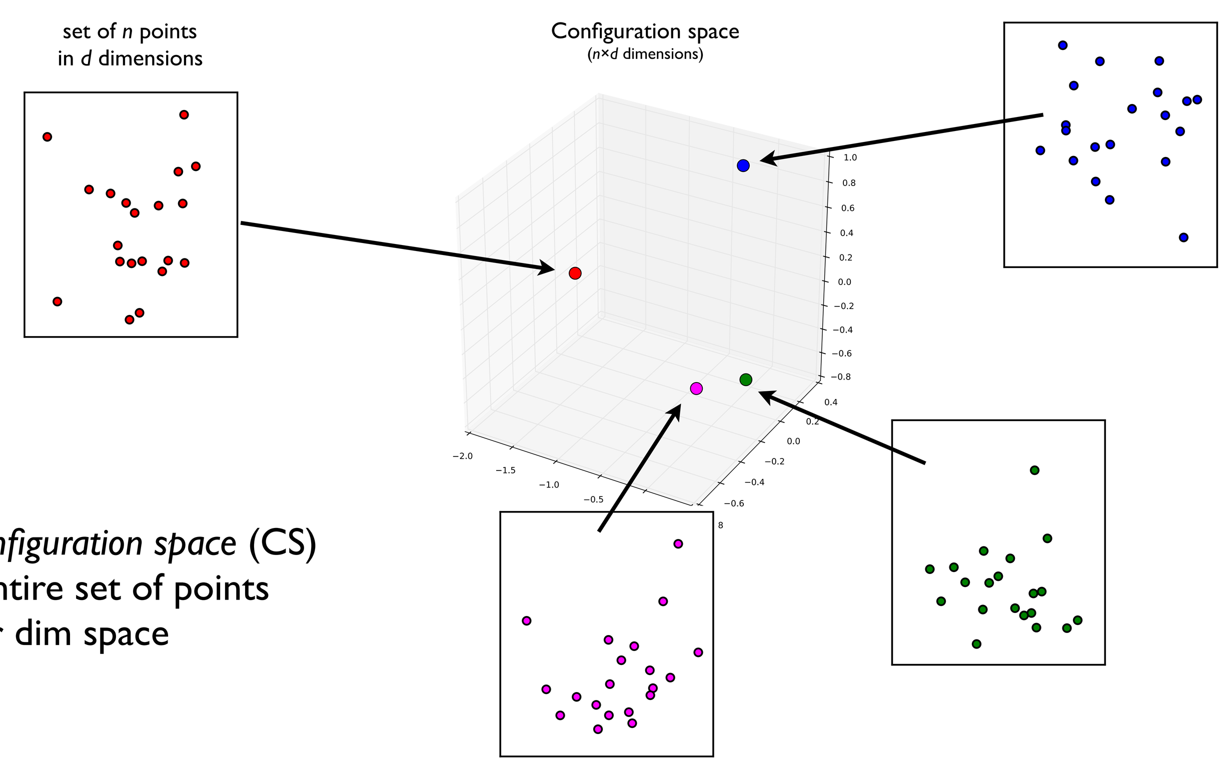
A State-Space Model on Interactive Dimensionality Reduction

Ignacio Díaz[1], Abel A. Cuadrado[1] and Michel Verleysen[2] *

1- Electrical Engineering Dept. University of Oviedo
Edif. Dept. 2, campus de Viesques s/n 33204, Gijón, SPAIN
2- Univ. Catholique de Louvain - Machine Learning Group
ICTEAM/ELEN - Place du Levant, 3 1348 Louvain-la-Neuve, Belgium

Abstract. In this work, we present a conceptual approach to the convergence dynamics of interactive dimensionality reduction (iDR) algorithms from the perspective of a well stablished theoretical model, namely *state-space* theory. The expected benefits are twofold: 1) suggesting new ways to import well known ideas from the state-space theory that help in the characterization and development of iDR algorithms and 2) providing a conceptual model for user interaction in iDR algorithms, that can be easily adopted for future interactive machine learning (iML) tools.

Configuration space



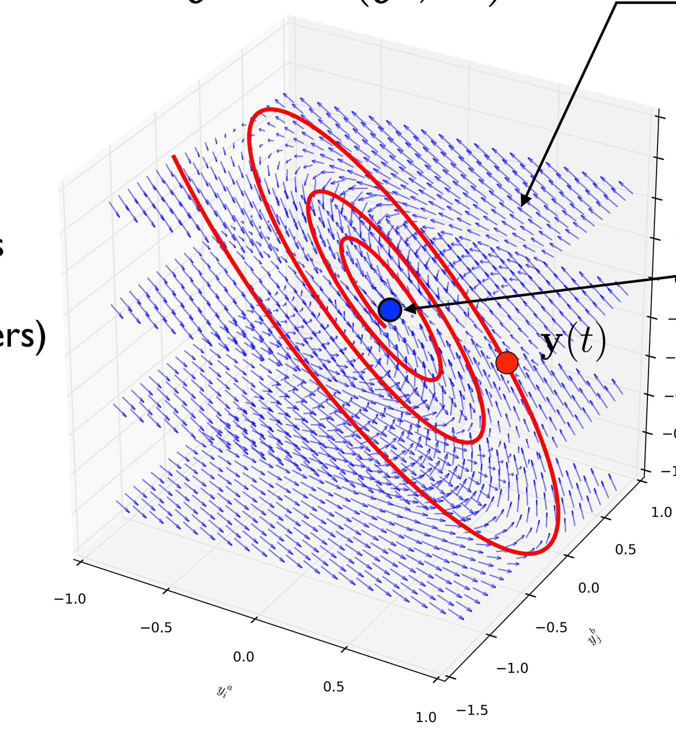
Every point in a *configuration space* (CS) represents an entire set of points in a lower dim space

State-space dynamics in the configuration space

$$\dot{\mathbf{y}} = \mathbf{f}(\mathbf{y},\mathbf{u})$$

During DR algorithm execution the state (i.e. the whole projection set) evolves until convergence

The motion law in the projection CS depends on the current state of the projection and exogenous factors (input data and parameters)



steady state (convergence of DR algorithm)

$\dot{\mathbf{y}} = \mathbf{0}$

$\mathbf{0} = \mathbf{f}(\mathbf{y}^0, \mathbf{u}^0)$

Interactive DR in the configuration space

DR Projections **y**(t) evolves according to **f** (natural dynamics) and u (forced dynamics)

context vector

exogenous (forced) inputs	
input data set in its config. space	algorithm parameters

$$\mathbf{u} = (\mathbf{x},\mathbf{w}) = (x_1^1, x_1^2, ..., x_1^D, x_2^1, x_2^2, ..., x_2^D, ..., x_Q^1, x_Q^2 ..., x_Q^D, w^1, w^2, ..., w^m)$$

problem dependent — user steerable

discrete approximation:

$$\mathbf{y}(t+1) = \mathbf{y}(t) + \Delta t \cdot \mathbf{f}(\mathbf{y}(t), \mathbf{u}(t))$$

f can be interpreted as the gradient of the cost function of the DR algorithm

$$\mathbf{f}(\mathbf{y},\mathbf{u}) \approx \frac{\partial J(\mathbf{y},\mathbf{u})}{\partial \mathbf{y}}$$

$$\dot{\mathbf{y}} = \mathbf{f}(\mathbf{y},\mathbf{u})$$

$\mathbf{x}(t)$

$\mathbf{x}_1(t), \ldots, \mathbf{x}_Q(t)$

if **x**(t) reaches a steady state

$$\mathbf{x}^0 \text{ for } t \to \infty \longrightarrow \mathbf{0} = \mathbf{f}(\mathbf{y}^0, (\mathbf{x}^0, \mathbf{w}))$$

assuming the system is locally stable, this implies that for a given $\mathbf{x}_0$ DR projection converges to $\mathbf{y}_0$

$\mathbf{x}(\lambda)$

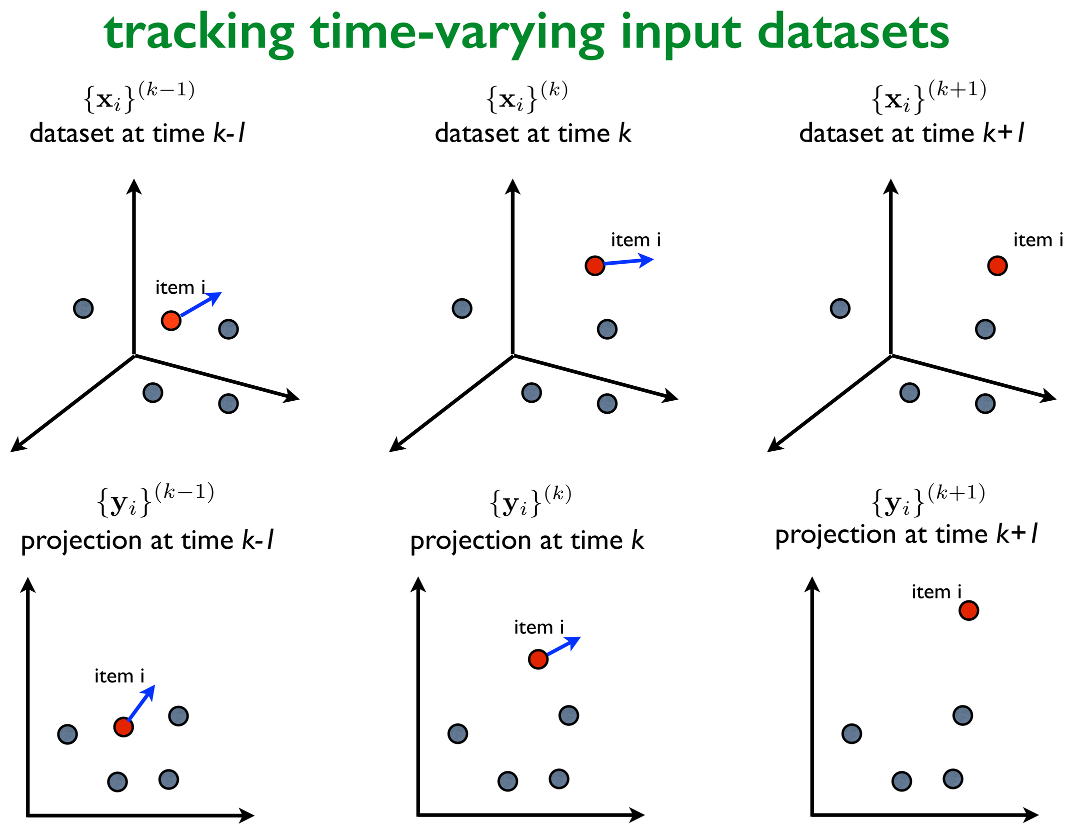
user-steerable λ

$$\mathbf{x}_e(\lambda) = [(1-\lambda)\mathbf{x},\ \lambda\bar{\mathbf{x}}_{c(\mathbf{x})}]$$

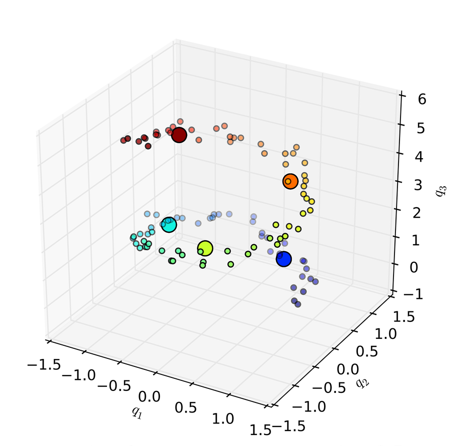
user-steerable DR parameters (e.g. $w^i = \omega_{ii}$)

$$\|\mathbf{x}_i - \mathbf{x}_j\|^2_{\mathbf{\Omega}} := \sum_r \sum_s (x_i^r - x_j^r)\omega_{rs}(x_i^s - x_j^s)$$

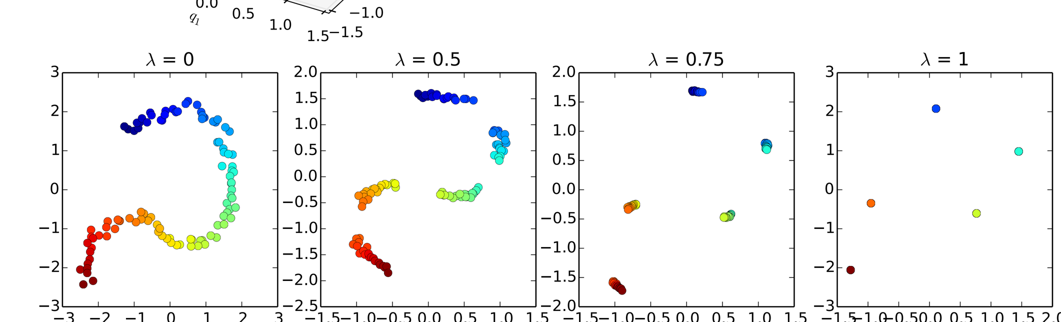
tracking time-varying input datasets



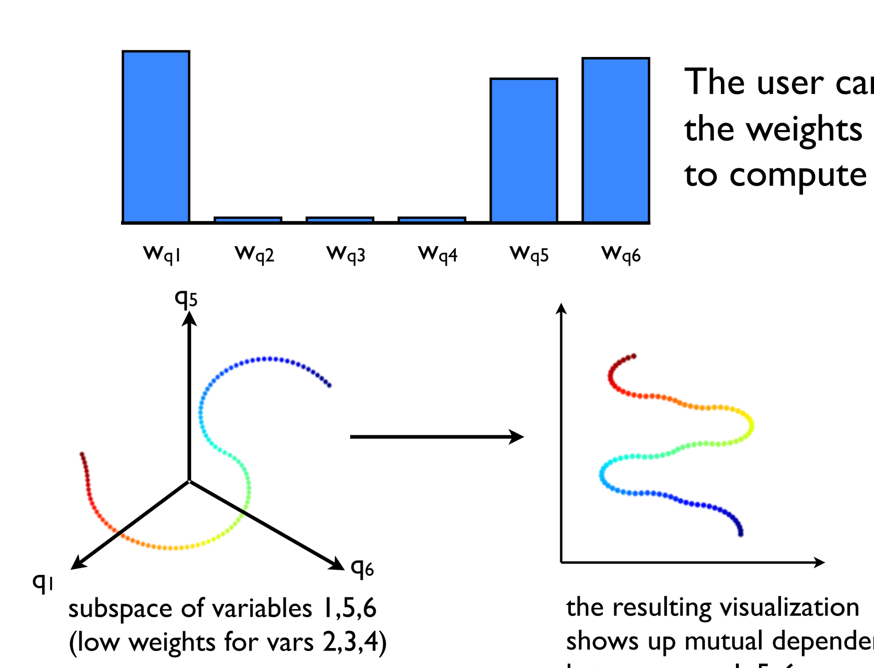
introducing class knowledge



The original dataset and class centroids in the input space can be combined into an extended input space of 2n dimensions. A user steered parameter λ can be used to interactively weight the subspace of the original dataset against the subspace of the centroids.



steering DR optimization parameters

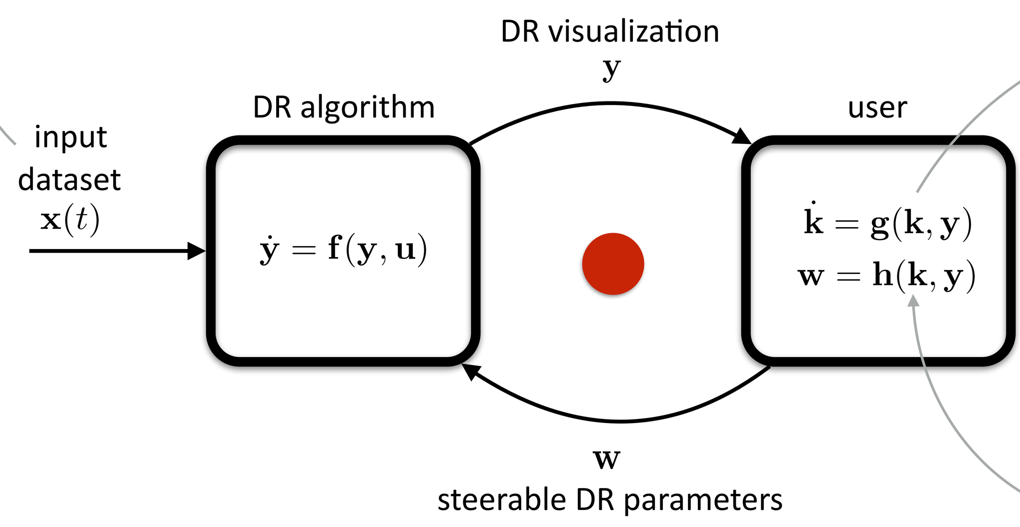


The user can interactively tune the weights ω_{ii} of the metric used to compute the point distances.

e.g. with nonzero weights for pairwise (possibly nonlinear) correlated variables, a 1D manifold projection is obtained

Conceptual model for user interaction

The final result $(\mathbf{k}^0, \mathbf{y}^0)$ mainly depends on the input data **x** to be analyzed, assuming that the user is not influenced by other factors.



The knowledge of the user evolves depending on the visualization and the current knowledge state as $\dot{\mathbf{k}} = g(\mathbf{k},\mathbf{y})$

Based on the current state of the user's knowledge **k** and the current projection state **y**, the user steers the DR configuration parameters as $\mathbf{w} = \mathbf{h}(\mathbf{k},\mathbf{y})$ to recompute the projection, according to new criteria that better meet the user's interest.

This results in a coupled system containing the current user's knowledge and the current projection, that models the interaction between the user and the DR algorithm.

The quality of the final knowledge $\mathbf{k}^0$ depends on a good design exploiting the synergy between the algorithm $\mathbf{f}()$ the user's mind capabilities $\mathbf{g}()$ and the interaction $\mathbf{h}()$.

In control theory, this equation can be seen as a *control law*, since it defines an input value **u** (depending on **w**) that manipulates the dynamical system, according to some target (here, maximizing user's knowledge), based on the information of the current system state.

Stability analysis of DR algorithms

Analysis of local dynamics of DR

The state space approach allows to analyze the DR from a dynamic point of view. Small variations around the equilibrium point 0 allow us to consider the linear model

$$\dot{\mathbf{y}} = \mathbf{A}\mathbf{y} + \mathbf{B}\mathbf{u}$$

This paves the way for rigorous local analysis of stability and dynamical behavior of iDR convergence, based on eigenmode analysis of the state matrix

$$\mathbf{A} = \left.\frac{\partial f(\mathbf{y},\mathbf{u})}{\partial \mathbf{y}}\right|_o$$

Potential applications

Existence of convergence (stability)

The existence of convergence checking for the stability condition $\mathrm{Re}(\lambda_i) < 0$ for all eigenvalues.
Eigenvalue tracking could be used to visualize the influence of context (e.g. algorithm parameters or even input data configurations) in stability

Comparison of DR algorithms

Numerical characterization of DR algorithms dynamics by K dominant eigenmodes could potentially be used for comparing algorithms according to their dynamical behavior, allowing also to elaborate taxonomies.

Improvement of DR algorithms

A better understanding of the dynamics of DR algorithms and the possibility to explicitly determine the local dynamical behavior, paves the way for strategies to improve the global stability using, for instance, adaptive tuning of DR algorithms

Acknowledgements

The authors would like to thank financial support from the Spanish Ministry of Economy (MINECO) and FEDER funds from the EU under grant DPI2015-69891-C2-2-R.

MINISTERIO DE CIENCIA E INNOVACIÓN