Auto-Associative Models a non-linear PCA

Valentin Pibernus[1,2] · Sylvain Girard[1] · Stéphane Girard[2] · Serge Iovleff[3]

[1]Phimeca Engineering, Paris, France

[2]INRIA, Grenoble - Rhone-Alpes, France [3]UTBM, Belfort-Montbeliard, France

Class of the problem

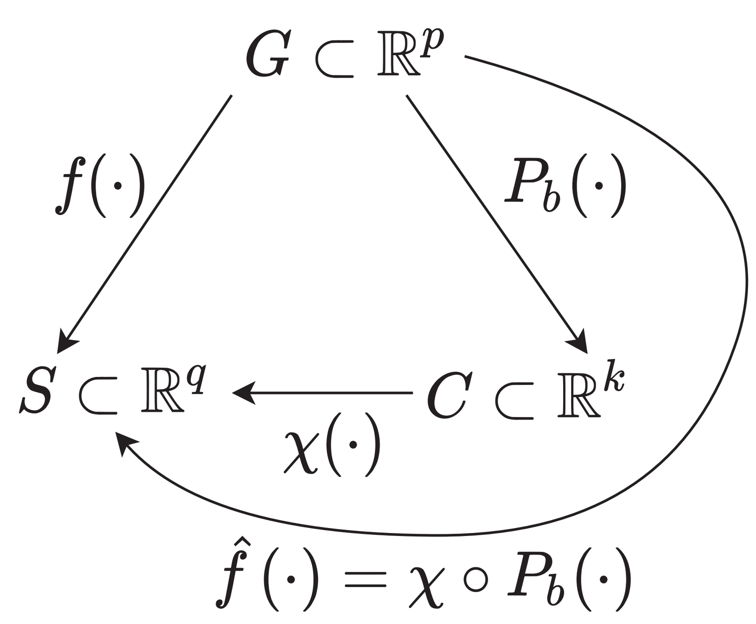

Figure 1. Class of problems

We study the relation $f(\cdot)$ between G and S. G is a k-manifold with $k \ll p$. The objective is to build a map between G and C in order to perform a dimension reduction.

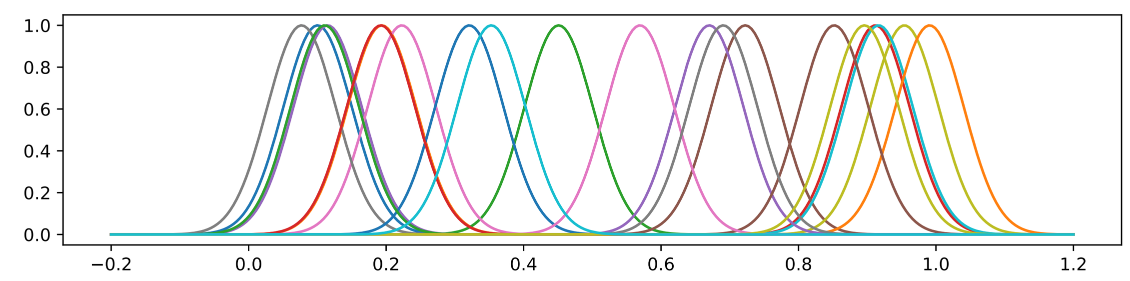

Figure 2. Shifting mean Gaussians example

Motivation

A machine learning

- frugal when possible
- robust
- rooted in physics
- interpretable and theoretically grounded

Acknowledgements

This work is partly funded by AMIES through the PEPS programme.

Auto-Associative Models

Auto-Associative Models are thought of as a non-linear PCA. These methods are dedicated to the approximation of the dataset by a manifold.

The algorithm is the following. For $i = 1, \dots, k$:

Step 1: Find direction

- Find a direction b_i optimizing a given criterion $I_{b_i}(X^{(i-1)})$.

Step 2: Project

- Compute $C^{(i)} = P_{b_i}(X^{(i-1)})$ the i^{th} coordinate in C.

Step 3: Recover

- Estimate the recovery function : $s_i(t) = \mathbb{E}(X^{(i-1)}|C^{(i)} = t)$.

Step 4: Iterate !

- Compute the residual and restart. $X^{(i)} = X^{(i-1)} - s_i \circ P_{b_i}(X^{(i-1)})$

Composing the regressions enables the reconstruction of complex shaped manifolds.

A powerfull projection criterion

The use of a criterion preserving the first-order neighbourhood structure yields interesting results.

$$I_b(X) = \sum_{i=1}^{n} P_b(X)^2 \Big/ \sum_{i=1}^{n}\sum_{j=1}^{n} m_{i,j} P_b(X_i - X_j)^2$$

The matrix $M = (m_{i,j})$ is a first order contiguity matrix whose value is 1 when X_j is the nearest neigbor of X_i, and 0 otherwise.

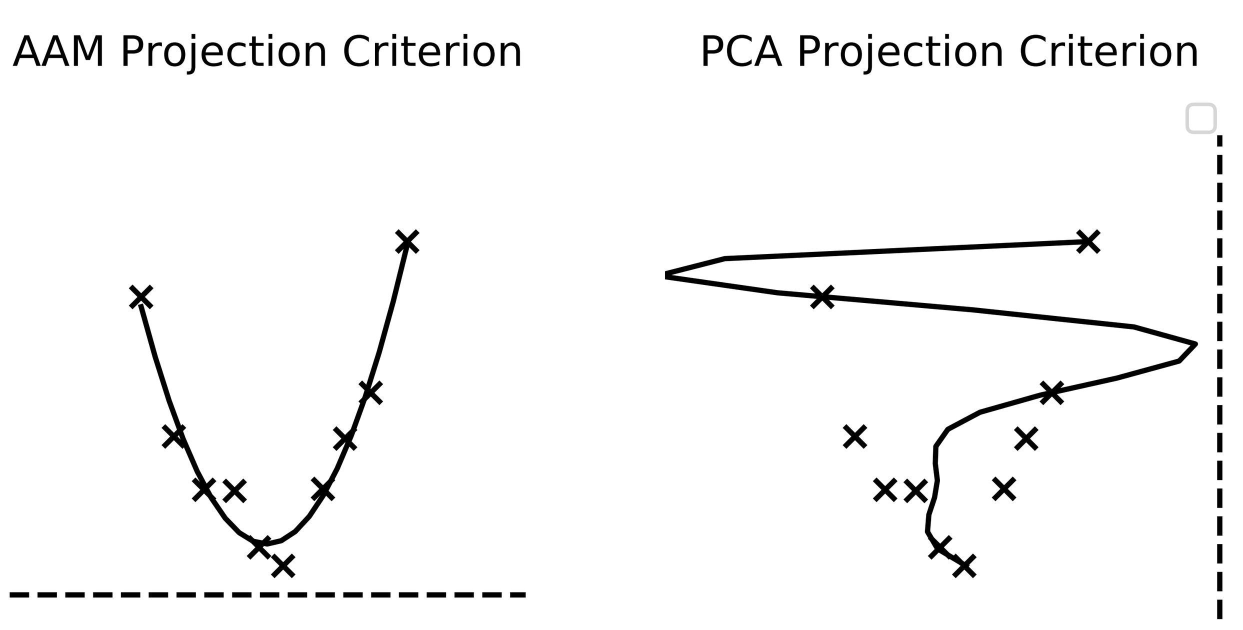

Figure 3. AAM vs PCA projection criterion

A non-linear recovery function

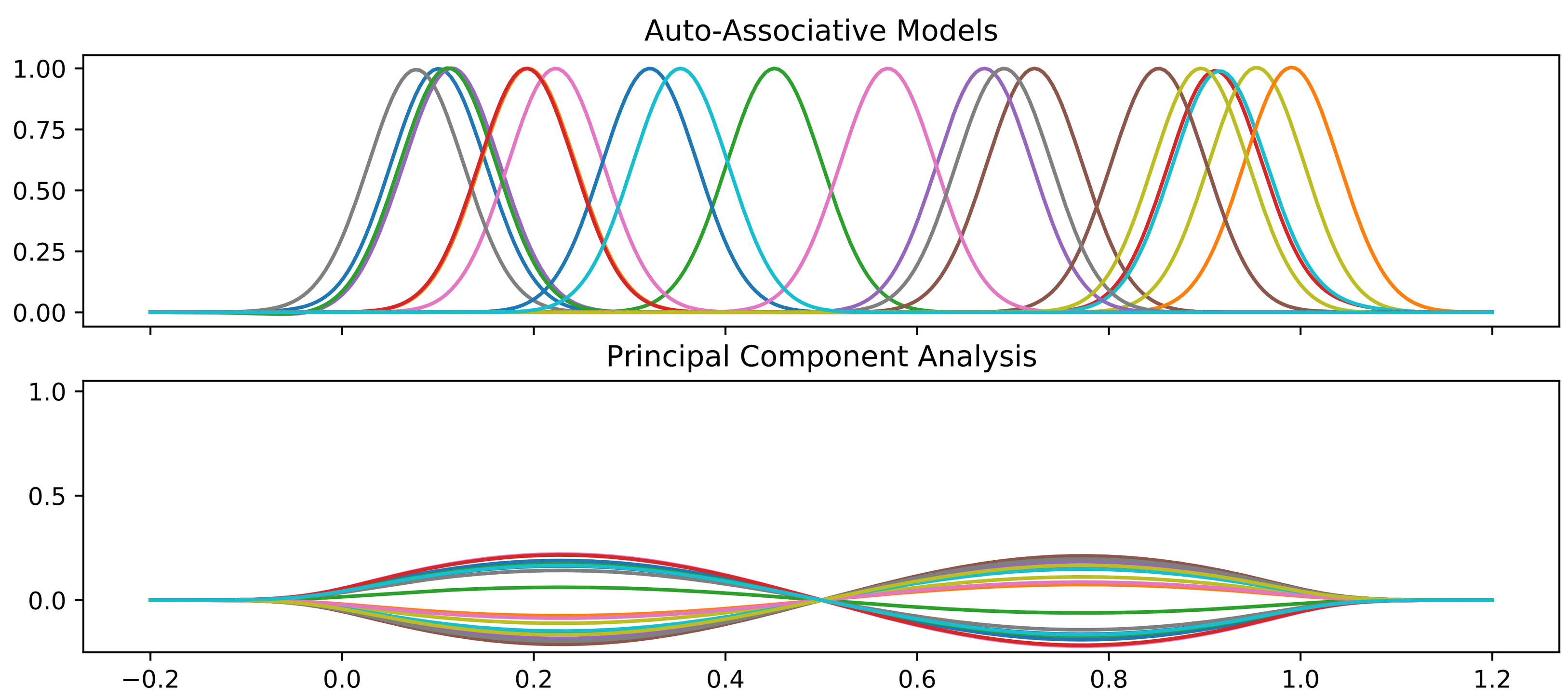

Figure 4. AAM vs PCA shifting Gaussian recovery

$s(t)$ is a univariate function and thus its estimation does not suffer from the curse of dimensionality.

Supervised Auto-Associative Models

Since the aim of dimension reduction is to simplify the regression from $X \in G$ to $Y \in S$, we can use the information in S to **supervise** the dimension reduction.

We use a contiguity matrix $M_S = (m_{i,j})$ whose value is 1 when Y_j is the nearest neighbour of Y_i and 0 otherwise. This projection criterion tries to project the isolign of $f(\cdot)$ on a point. This makes the regression from C to S straightforward.

SIR vs SAAM

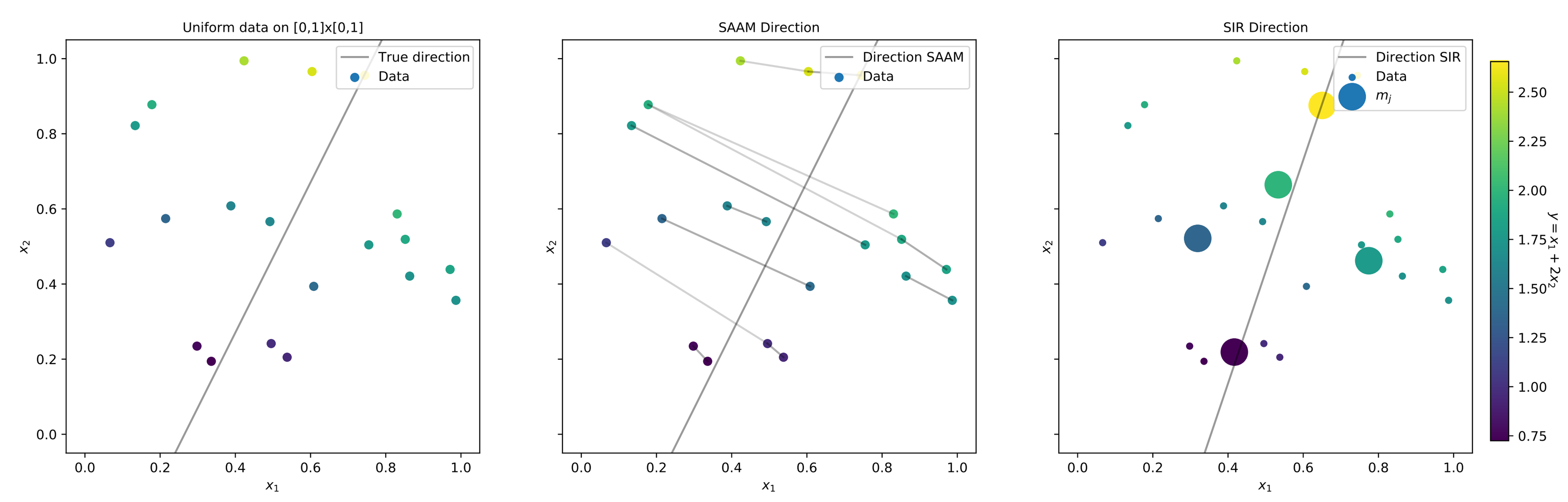

Figure 5. SIR against SAAM

SIR : Apply PCA on the barycenter of grouped data.

SAAM : Minimize the projected distance in G from the nearest neighbour in S.

Take home message

AAM and SAAM exploit different ideas :

- an "inversible" projection
- a non linear recovery function
- a topological thinking of dimension reduction