Latent Space Exploration Using Generative Kernel PCA

David Winant[1], Joachim Schreurs[1], and Johan Suykens[1]

KU Leuven, Department of Electrical Engineering (ESAT),
STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics,
Kasteelpark Arenberg 10, B-3001 Leuven, Belgium
{david.winant,joachim.schreurs,johan.suykens}@kuleuven.be

1 Introduction

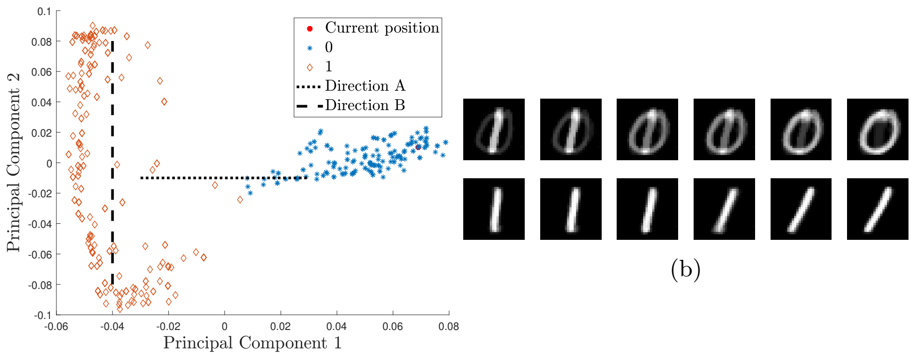
Latent spaces provide a representation of data by embedding the data into an underlying vector space. Exploring these spaces allows for deeper insights in the structure of the data distribution and the relationships between data points. The focus of this paper will be on how the synthesis of new data with generative kernel PCA [4] can help with understanding the latent features extracted from a dataset.

2 Generative Kernel PCA

Kernel PCA, as first described in [3], is a well-known feature extractor method often used for denoising and dimensionality reduction of datasets. In [5], kernel PCA was cast within the framework of Restricted Kernel Machines (RKMs) which allows for an interpretation in terms of hidden and visible units similar to Restricted Boltzmann Machines (RBMs) [1]. This connection between kernel PCA and RBMs was later used to explore a generative mechanism for the kernel PCA [4]. In practice generative kernel PCA works as follows: first kernel PCA is performed to find the hidden features of the dataset. After choosing an initial hidden unit as starting point, the values for each component of the hidden unit can be varied to explore the latent space. The corresponding newly generated data point in the input space is estimated using the kernel smoother approach.

3 Latent space exploration

The goal of this paper is to explore the latent feature space extracted by kernel PCA, in an effort to interpret the components. This has led to the development of a Matlab tool which can be used to visualise the latent space of the kernel PCA method along its principal components. Along with the newly generated point, a partial visualisation of the latent space projected onto two principal components is shown. Continuously varying the values of the components of the selected hidden unit allows for the exploration of the extracted latent space by visualising the resulting variation in the input space. In Fig. 1 an example use case for the MNIST handwritten digits dataset [2] is shown.



(a) (b)

Fig. 1: Latent space exploration for 1000 digits 0 and 1 from the MNIST dataset using generative kernel PCA. (a) Latent space projected on the first two principal components. (b) Generated digits along the directions A and B. The generated digits in the top row allow for the interpretation of the first component as transitioning from 1 to 0, while the bottom row indicates that the second component smoothly rotates the digit. This explains the limited variation of the hidden units for the digits zero along this direction as they are largely invariant under rotations.

4 Conclusion

Generative kernel PCA can be used for exploring the latent space. Gradually moving in the feature space allows for the interpretation of components and consequently additional insight into the underlying latent space. For this purpose a Matlab tool has been developed which can easily be used for additional datasets as well as aid interpretation in the context of novelty detection.

References

1. Hinton, G.E.: A practical guide to training restricted Boltzmann machines. In: Neural networks: Tricks of the trade, pp. 599–619. Springer (2012)
2. LeCun, Y., Bottou, L., Bengio, Y., Haffner, P., et al.: Gradient-based learning applied to document recognition. Proceedings of the IEEE **86**(11), 2278–2324 (1998)
3. Schölkopf, B., Smola, A., Müller, K.R.: Kernel principal component analysis. In: International conference on artificial neural networks. pp. 583–588. Springer (1997)
4. Schreurs, J., Suykens, J.A.K.: Generative kernel pca. ESANN 2018 pp. 129–134 (2018)
5. Suykens, J.A.K.: Deep restricted kernel machines using conjugate feature duality. Neural computation **29**(8), 2123–2163 (2017)