Kernel Methods in Orthogonalization of Near-Infrared Hyperspectral Images of Maize Kernels

Allan Aasbjerg Nielsen

Technical University of Denmark
DTU Space – National Space Institute
Presently located at
DTU Informatics – Informatics and Mathematical Modelling
Richard Petersens Plads, Building 321, DK-2800 Lyngby, Denmark
Tel +45 4525 3425, Fax +45 4588 1397
E-mail aa@space.dtu.dk, http://www.imm.dtu.dk/∼aa

Abstract

Principal component analysis (PCA) may be considered as the mother of all linear orthogonal transformations for data compression and dimensionality reduction of correlated multivariate data. This contribution describes a kernel version of PCA and it also sketches kernel versions of maximum autocorrelation factor (MAF) analysis and minimum noise fraction (MNF) analysis. In kernel methods the actual observations are replaced by high (in principal maybe even infinite) dimensional nonlinear mappings of the observations. In dual formulation or Q-mode PCA which is applied in the kernel versions, these mappings are implicitly defined by a kernel function. This type of analysis allows for nonlinearities in the original data, and especially the kernel MAF/MNF transformations tend to focus on extreme observations on a slower changing background. Application areas include hyperspectral NIR images for food quality control, analysis of simple or generalized differences of multi- and hyperspectral Earth observation image data, and irregularly spaced geochemical data for geological mapping. The example shown here illustrates the application of linear and kernel PCA and MAF analysis to hyperspectral NIR images of maize kernels.

The 149 rows by 370 columns image is recorded by a line-scan NIR camera. In the analysis 153 spectral bands in the 900-1700 nm region are used.

In all images shown the fronts and backs of eight maize kernels are shown. The original hyperspectral images are (obviously) recorded separately and stitched together later. In all

images all bands shown are stretched linearly between mean minus and plus three standard deviations.

Figure 1 shows linear principal components 1, 2 and 3 as RGB. Figure 2 shows linear maximum autocorrelation factors 1, 2 and 3 as RGB. Figure 3 shows kernel principal components 1, 2 and 3 as RGB. Figure 4 shows kernel maximum autocorrelation factors 1, 2 and 3 as RGB. In all images exept Figure 4 with the kernel MAFs the border between the front image and the back image is visible and some shadow effect is present. Also, the kernel MAFs provide a better discrimination between the structural components of the kernels which generally can be divided into three classes denoted Endosperm, Germ and Pedicel. The analysis demonstrates that the kernel MAF analysis outperforms the linear methods as well as kernel PCA in producing interesting projections of the data.

A few relevant references in this context are listed below [1, 2, 3, 4, 5, 6, 7].

References

[1] A. A. Green, M. Berman, P. Switzer, and M. D. Craig, "A transformation for ordering multispectral data in terms of image quality with implications for noise removal," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 26, no. 1, pp. 65–74, 1988.

[2] B. Schölkopf, A. Smola, and K.-R. Müller, "Nonlinear component analysis as a kernel eigenvalue problem," *Neural Computation*, vol. 10, no. 5, pp. 1299–1319, 1998.

[3] J. Shawe-Taylor and N. Cristianini, *Kernel Methods for Pattern Analysis*, Cambridge University Press, 2004.

[4] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.

[5] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, third edition, Cambridge University Press, 2007.

[6] R. Larsen, M. Arngren, P. W. Hansen, and A. A. Nielsen, "Kernel based subspace projection of near infrared hyperspectral images of maize kernels," Scandinavian Image Analysis Conference, SCIA, Oslo, Norway, 15-18 June 2009, Internet http://www.imm.dtu.dk/pubdb/p.php?5744.

[7] A. A. Nielsen, "Kernel maximum autocorrelation factor and minimum noise fraction transformations," submitted, 2010.

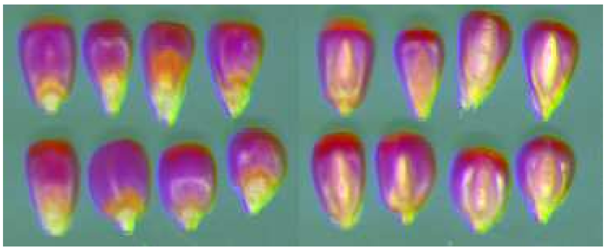

Figure 1: Linear principal components 1, 2 and 3 as RGB.

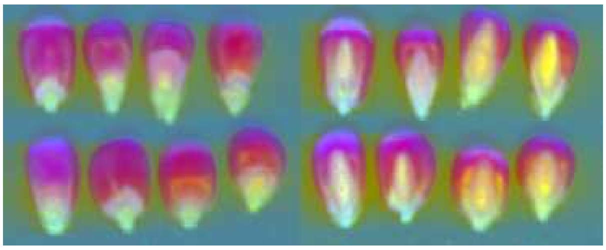

Figure 2: Linear maximum autocorrelation factors 1, 2 and 3 as RGB.

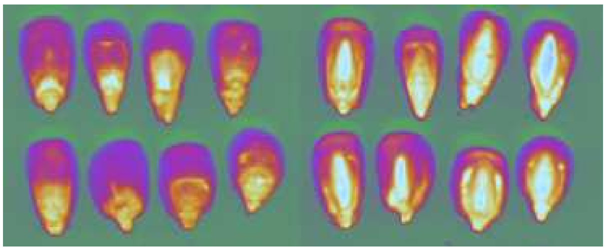

Figure 3: Kernel principal components 1, 2 and 3 as RGB.

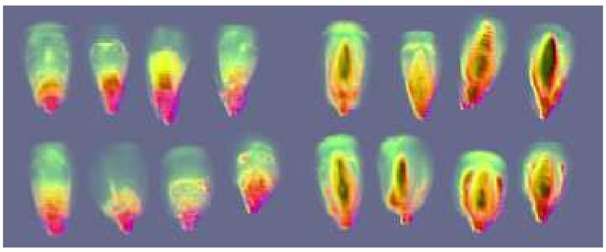

Figure 4: Kernel maximum autocorrelation factors 1, 2 and 3 as RGB.