Geophysical Research Abstracts
Vol. 13, EGU2011-9072-1, 2011
EGU General Assembly 2011

Non-negative Matrix Factorisation in Context

Albrecht Schmidt (1), Erwan Tréguier (1), Frédéric Schmidt (2), Saïd Moussaoui (3), and Nicolas Dobigeon (4)

(1) European Space Agency, SRE-OS, Villanueva de la Cañada, Spain (albrecht.schmidt@esa.int), (2) IDES, UMR 8148, Université Paris Sud, CNRS, Orsay, France (frederic.schmidt@u-psud.fr), (3) IRCCYN, UMR CNRS 6597, Ecole Centrale Nantes, France (said.moussaoui@irccyn.ec-nantes.fr), (4) IRIT/INPT-ENSEEIHT, Toulouse, France (nicolas.dobigeon@enseeiht.fr)

Various methods which implement non-negative matrix factorisation have been used successfully in recent years for analysing hyperspectral images. Assuming linear mixings, the methods estimate the source spectrograms which are used to approximate the individual pixels of the hyperspectral image, given the number of sources one expects to find in the data. This presentation will try to put matrix factorisation techniques into context and compare them to other unsupervised data analysis methods and concepts, such as nearest neighbours, k-means, principal components, and convex hulls.

The general setting of the work is framed as follows: we consider P pixels of a hyperspectral image which are acquired at L frequency bands and which are represented as a PxL data matrix X. Each row of this matrix contains a measured spectrum at a pixel with spatial index p=1..P, i.e. the original topology is disregarded. Since we assume linear mixing, the p-th spectrum, 1<=p<=P, can be expressed as a linear combination of r, 1<=r<=R, pure spectra of the surface components. Thus, X=AxS+E, E being an error matrix, which should be minimised, and X, A, and S only have non-negative entries. The rows of matrix S are the estimated surface spectra of the R components, and each entry of A expresses the strength of the r-th component in the pixel with spatial index p.

A distinguishing feature of matrix factorisation techniques is that they allow a constructive interpretation of matrices A and S: the rows of S can be interpreted as the physical source spectrograms that would appear in some pixels if instrumental and geometrical limitations allowed this. The rows of A can be seen as quantitative descriptions of how the various rows of S are mixed to obtain the original image. Thus, the underlying processes are abstracted and fully described by the above mathematical representation. Note that all operations take place in the original space which remains untransformed throughout the computation.

We use both theorical examples and case-studies from the planetary spectrometers to illustrate the potential of the techniques and the physical implications they imply. We will also try to illustrate how to optimise computations by taking advantage of well-known concepts such as nearest neighbours and convex hulls to accelerate the computations.