



Contribution ID: 81

Type: **not specified**

## **A novel online PCA algorithm for large variable space dimensions**

*Wednesday, 15 September 2021 12:40 (20 minutes)*

Principal component analysis (PCA) is a basic tool for reducing the dimension of a space of variables. In modern industrial environments large variable space dimensions up to several thousands are common, where data are recorded live in high time resolution and have to be analysed without time delay. Classical batch PCA procedure start from the full covariance matrix and construct the exact eigenspace of the space defined by the covariance matrix. The latter approach is infeasible under large dimensions, and even if feasible live updating of the PCA is impossible. Several so-called online PCA algorithms are available in the literature who try to handle large dimensions and live updating with different approaches. The present study compares the performance of available online PCA algorithms and suggests a novel online PCA algorithm. The algorithm is derived by solving a simplified maximum trace problem where the optimisation is restricted on the curve on the unit sphere, which directly connects the respective old principal component estimation with a projection of the newly observed data point. The algorithm scales linearly in runtime and in memory with the data dimension. The advantage of the novel algorithm lies in providing exactly orthogonal vectors whereas other algorithms lead to approximately orthogonal vectors. Nevertheless, the runtime of the novel algorithm is not worse and sometimes even better than the one of existing online PCA algorithms.

### **Keywords**

batch PCA, online PCA

### **Special/invited session**

**Primary authors:** FROEHLICH, Philipp (University of Wuerzburg); GÖB, Rainer

**Presenters:** FROEHLICH, Philipp (University of Wuerzburg); GÖB, Rainer

**Session Classification:** Process 2

**Track Classification:** Process