Dim. Reduction of High Freq. and High Dim. Data in Time and Space

HU-MATHS-IN
Hungarian Service Network for Mathematics in Industry and Innovations

The Industrial Problem

The objective of this research is to find the underlying main signals or driving latent sources detected by many motion sensors worn by different subjects, mainly athletes.

Health Care, Sport Analytics

Research group

BME Statistics and Mathematical Modeling Consulting Group

Statisztikai és Matematikai Modellezési Tanácsadó Csoport

MŰEGYETEM 1782

It offers statistical consulting and provides data science research and development service to corporate and academic partners.

Company

I-QRS International

i-QRS

It is a R&D company studying ECG and its relation to other physiological parameters. With their unique in-house created hardware, they monitor wirelessly and precisely vital and motion parameters of subjects.

SZÉCHENYI 2020
MAGYARORSZÁG KORMÁNYA
Európai Unió
Európai Strukturális és Beruházási Alapok
BEFEKTETÉS A JÖVŐBE

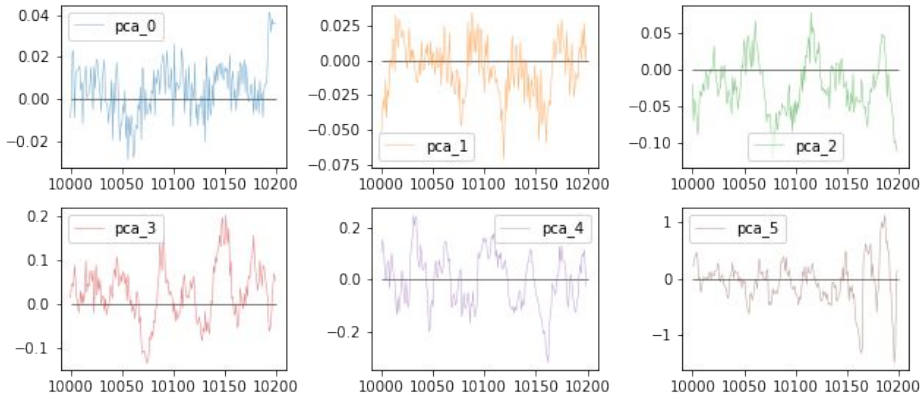
Dim. Reduction of High Freq. and High Dim. Data in Time and Space

HU-MATHS-IN
Hungarian Service Network for Mathematics in Industry and Innovations

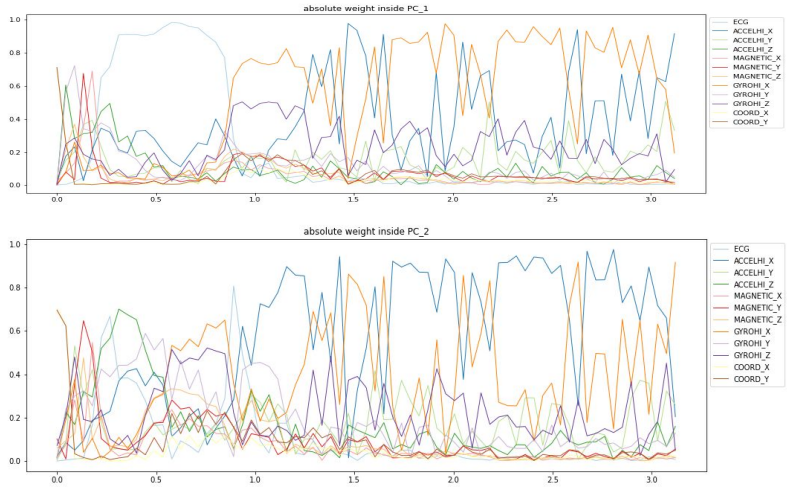
Health Care,
Sport Analytics

Challenges & Goals

- High and mixed frequency multidimensional time-series data from noisy sensors.
- The size and nature of the data are challenging for simple algorithms.
- The goal is to find dynamical principal components (PCs) of the underlying signals.
  - These have many nice properties, and advantage over static techniques.
  - They provide explainable dimension reduction by the loadings.



A small segment (~0.4s) of the dynamic PCs estimated from the motion sensor data, reduction from 12 to 6 dimensions



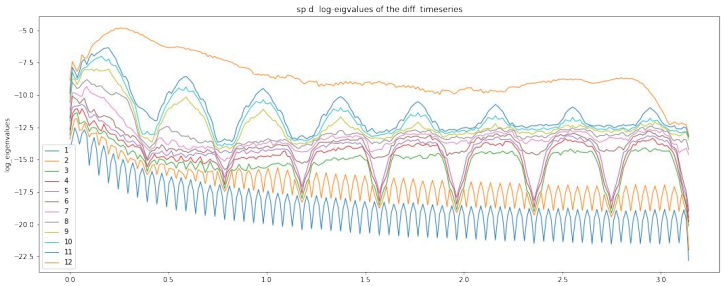
Absolute values of the complex weights (loadings of each dimension) inside the first two PCs on small segment of the data (~40s)

Dim. Reduction of High Freq. and High Dim. Data in Time and Space

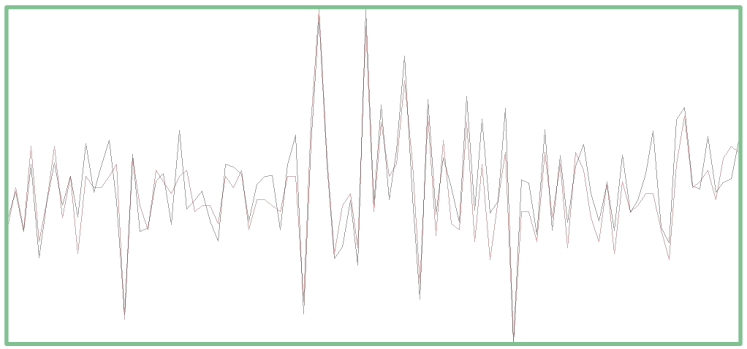
HU-MATHS-IN
Hungarian Service Network for Mathematics in Industry and Innovations

Mathematical and computational methods and techniques applied

- The **spectral theory** of multivariate stationary time series.
  - It is developed mainly for one-dimensional processes a century ago, can be extended in many ways. Importantly for this project, it can be extended to dimension reduction of multivariate, discrete-time processes, to dynamic principal component analysis (d-PCA).
- **Linear algebra**: eigenvalue decomposition, projections, complex linear interpolation.
- **Fourier analysis**: fast Fourier transformation (FFT).
- **Statistics**: estimation of autocovariances, standardization.



Log-eigenvalues of the spectral density estimates of the mixed freq. data



A small segment (~0.2s) of one sensor (red) with its low-rank approximation (black)

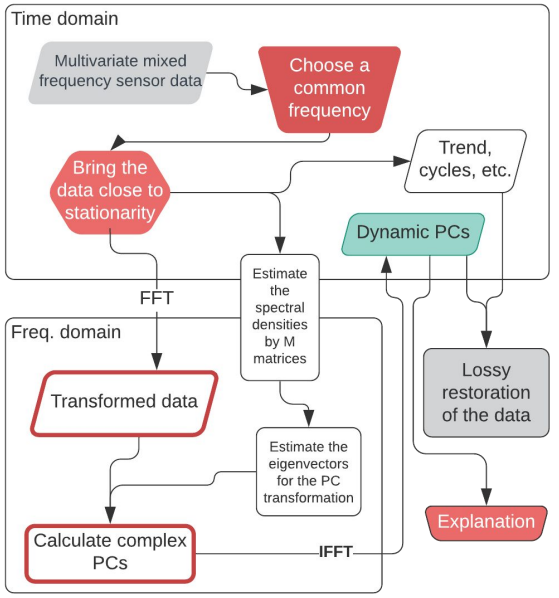
Dim. Reduction of High Freq. and High Dim. Data in Time and Space

HU-MATHS-IN
Hungarian Service Network for Mathematics in Industry and Innovations

Results & Benefits to the company

We developed and implemented an algorithm in Python to calculate/estimate the dynamic principal components of the multidimensional motion sensor data of the athletes.

- It may help the company understand their data even better in the future.
- The algorithm can be utilized as a lossy data compression technique.



Schematic workflow of the algorithm.

The company is planning to further investigate the delivered algorithm/implementation in the future.