Evolutionary Factor Analysis of Brain Signals

Maastricht University
The Department of Quantitative Economics

BROWN Public Health Program

Giovanni Motta Hernando Ombao

g.motta@maastrichtuniversity.nl

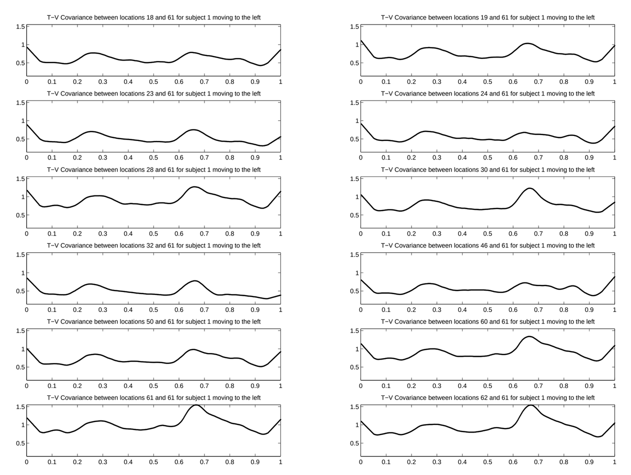
Abstract

In this paper (see [3]) we characterize and estimate the dynamic structure of multi-channel electroencephalograms (EEGs) in a motor-visual task experiment. Preliminary analyses of our data suggest that both the variance of each channel and cross-covariance between a pair of channels evolve over time. Moreover, the cross-covariance profiles display a common structure across all pairs. Based on these observations, we apply the methods of Evolutionary Factor Analysis (EFA) to the multi-channel EEG data. EFA (see [4]) is a statistical tool recently developed to study multivariate non-stationary stochastic processes that are driven by common factors. EFA provides a new class of factor models with time-varying factor loadings. The basic idea is to consider these loadings as smooth functions of time rendering the process nonstationary while keeping the factors stationary. The assumption that loadings are smooth enables to estimate the model using nonparametric methods. The estimation of these nonstationary factor models makes use of the generalization of the properties of the principal components techniques to the time-varying framework. One key feature of the dataset is that multi-channel EEGs were recorded over many trials. In this paper we develop a new model derived from EFA so that the factors share common features across trials thus allowing us to pool information across trials. We establish conditions for identification and estimation of loadings, factors and common components. In our analysis, we explain the common co-movements of EEG signals through the existence of a small number of common factors. These latent factors are primarily responsible for processing the visual-motor task which, through the loadings, drive the behavior of the signals observed at different channels.

Keywords: *Electroencephalography, Local Stationarity, Principal Components, Source Localization, Spatio-temporal modeling.*

1. Motivation: Dimension-Reduction for Multivariate Non-Stationary Time Series

- **Non-stationarity**. Time-homogeneity models describing mean and (co-)variance structure of EEG multivariate time series data are too restrictive.
- **Factor Structure**. The series show similar behavior over time: they are driven by common factors.



Time-varying covariances of EEG recorded during a motor-visual task experiment.

Goals:
- explain the outcome of P variables using fewer ($q < P$) *latent factors*;
- split the influences of the factors into common and specific ones;
- exploit the additional information at hand, given by multiple trials.

2. The Model

Observed time series at time t, $t = 1, \dots, T$, and r-th trial, $r = 1, \dots, R$:

$$\boldsymbol{Y}_{PT}(t,r) = \boldsymbol{X}_{PT}(t,r) + \boldsymbol{Z}_P(t,r) \qquad (1)$$
$$= \boldsymbol{\Lambda}_P\big(\tfrac{t}{T}\big)\boldsymbol{F}(t,r) + \boldsymbol{Z}_P(t,r); \qquad (2)$$

$\boldsymbol{Y}_{PT}(t)$ $P \times 1$ **evolutionary stochastic process**,
$\boldsymbol{X}_{PT}(t)$ $P \times 1$ **evolutionary common components**,
$\boldsymbol{\Lambda}_P\big(\tfrac{t}{T}\big)$ $P \times q$ **factor loadings**, $\lim_{P\to\infty} \sup_{u\in[0,1]} \Big\| \frac{\boldsymbol{\Lambda}'_P(u)\boldsymbol{\Lambda}_P(u)}{P} - \boldsymbol{\Sigma}^{\Lambda}(u) \Big\|$=0,
$\boldsymbol{F}(t)$ $q \times 1$ **orthogonal factors,**
$\boldsymbol{Z}_P(t)$ $P \times 1$ **idiosyncratic errors**.

New Hypothesis (**Local Stationarity** by [1]): the **loadings** are now allowed to be **smooth functions of time**.

Assumption 1 (Evolutionary factor model)**.** *The sequence of vectors $\boldsymbol{Y}_{PT}(t)$, $1 \le t \le T$, with $T, P \in \mathbb{N}$ is a family of stochastic processes given by* (1) *and* (2)*, and satisfying the following conditions:*

(i) $\mathbb{E}\big[\boldsymbol{F}(t)\big] = \mathbf{0}$, $\mathbb{E}\big(\boldsymbol{F}(t)\boldsymbol{F}(t)'\big) = \boldsymbol{\Sigma}^F$, $\boldsymbol{\Sigma}^F$ *is diagonal;*

(ii) the entries of $\boldsymbol{\Lambda}_P(u)$ *are continuously differentiable and, for all* $u \in [0,1]$ $\mathrm{rk}\big[\boldsymbol{\Lambda}_P(u)\big] = q$*;*

(iii) $\mathbb{E}\big[\boldsymbol{Z}_P(t)\boldsymbol{F}(t-k)'\big] = \mathbf{0}$ *for all* $P \in \mathbb{N}$ *and* $t, k \in \mathbb{Z}$.

Assumption 2 (Evolutionary covariance matrix)**.**

(i) For all $u \in [0,1]$*, there exists a* $P \times P$ *matrix* $\boldsymbol{\Sigma}_P(u)$ *such that*

$$\boldsymbol{\Sigma}_P\big(\tfrac{t}{T}\big) := \mathbb{V}\mathrm{ar}\big[\boldsymbol{Y}_{PT}(t,r)\big] = \boldsymbol{\Lambda}_P\big(\tfrac{t}{T}\big)\boldsymbol{\Sigma}^F\boldsymbol{\Lambda}_P\big(\tfrac{t}{T}\big)' + \boldsymbol{\Sigma}^Z_P = \boldsymbol{\Sigma}^X_P\big(\tfrac{t}{T}\big) + \boldsymbol{\Sigma}^Z_P, \qquad (3)$$

for all $r = 1, \dots, R$*, where* $\boldsymbol{\Sigma}^X_P$ *and* $\boldsymbol{\Sigma}^Z_P$ *are the covariance matrices of the common and idiosyncratic components, respectively;*

(ii) $\boldsymbol{\Sigma}^Z_P$ *is a sequence of covariance matrices with uniformly bounded eigenvalues, that is,* $\sup_P \mathrm{v}^Z_{1P} < \infty$, $P \in \mathbb{N}$*, where* v^Z_{1P} *denotes the largest eigenvalue of* $\boldsymbol{\Sigma}^Z_P$.

3. Estimation Steps

1. The pre-etimator $\widetilde{\boldsymbol{\Sigma}}_P(u,r)$ is the average (over time) of the squared observations at trial r for those values of $\frac{s}{T}$ around u:

$$\widetilde{\boldsymbol{\Sigma}}_P(u,r) := \sum_{s=1}^{T} \boldsymbol{Y}_{PT}(s,r)\boldsymbol{Y}_{PT}(s,r)' K_h\big(u - \tfrac{s}{T}\big),$$

where $K_h(\cdot) := \frac{1}{h}K\big(\frac{\cdot}{h}\big)$ is the rescaled version of a second order kernel, and $h = h_T$ is the sequence of smoothing bandwidth.

2. We define the estimator of the time-varying covariance in (3) as the average over trials of the estimates $\widetilde{\boldsymbol{\Sigma}}_P\big(u,r\big)$, i.e.:

$$\widehat{\boldsymbol{\Sigma}}_P(u) := \tfrac{1}{R}\sum_{r=1}^{R} \widetilde{\boldsymbol{\Sigma}}_P(u,r), \quad u \in (0,1). \qquad (4)$$

3. Before estimating the $P \times q$ matrix of loadings, we need to determine the number $\widehat{q}$ of factors. This is detailed in section 5.
4. Extract the $P \times q$ matrix of eigenvectors $\widehat{\boldsymbol{\Lambda}}_P(u)$ corresponding to the largest eigenvalues of the $P \times P$ matrix $\widehat{\boldsymbol{\Sigma}}_P(u)$ collected in the $q \times q$ diagonal matrix $\widehat{\mathbf{V}}(u)$:

$$\widehat{\boldsymbol{\Sigma}}_P(u)\widehat{\boldsymbol{\Lambda}}_P(u) = \widehat{\boldsymbol{\Lambda}}_P(u)\widehat{\mathbf{V}}(u), \quad u \in (0,1). \qquad (5)$$

5. Define the $q \times 1$ principal components at time t for the r-th trial, as the projection of the data $\boldsymbol{Y}_{PT}(t,r)$ at time t and trial r on the orthonormal eigenvectors $\widehat{\boldsymbol{\Lambda}}_P\big(\frac{t}{T}\big)$ at time t:

$$\widehat{\boldsymbol{F}}(t,r) = \tfrac{1}{P}\widehat{\boldsymbol{\Lambda}}_P(\tfrac{t}{T})'\boldsymbol{Y}_{PT}(t,r).$$

6. Define the estimated common components as

$$\widehat{\boldsymbol{X}}_{PT}(t,r) = \widehat{\boldsymbol{\Lambda}}_P(\tfrac{t}{T})\widehat{\boldsymbol{F}}(t,r). \qquad (6)$$

4. Asymptotic Theory

The asymptotic results hold for $T \to \infty$, $h_T \to 0$, $P \to \infty$ and $R \to \infty$ in such a way that $Th_T \to \infty$ and $Th_T^3 \to 0$. We benefit form the theory provided by [5]. Moreover, we increase the rate of convergence by $\sqrt{R}$.

Theorem 1 (Consistency of the Covariance-Matrix Estimator)**.**

$$P^{-1}\big\|\widehat{\boldsymbol{\Sigma}}_P(u) - \boldsymbol{\Sigma}_P(u)\big\| = O_p\big[\big(RTh_T\big)^{-\frac{1}{2}}\big] \qquad (7)$$

Theorem 2 (Consistency of the Time-Varying Eigenvalues)**.**

$$\min\big(P, \sqrt{RTh_T}\big)\big\|\widehat{\mathbf{V}}(u) - \mathbf{V}(u)\big\| = O_p(1), \qquad (8)$$

where $\mathbf{V}(u)$ *is a diagonal matrix containing the eigenvalues of* $\boldsymbol{\Sigma}^{\Lambda}(u)\boldsymbol{\Sigma}^F$.

Non-uniqueness of the factors and the factor loadings:

Let $\mathbf{G}$ be $q \times q$ orthogonal: the factor model (2) with factors $\mathbf{G}'\boldsymbol{F}(t,r)$ and loadings $\boldsymbol{\Lambda}\big(\frac{t}{T}\big)\mathbf{G}$ is also true. Thus, $\widehat{\boldsymbol{\Lambda}}(u)$ is an estimator of a *rotation* of the loadings $\boldsymbol{\Lambda}(u)\mathbf{H}(u)$, where $\mathbf{H}(u)$ is invertible for all u.

Theorem 3 (Consistency of the Time-Varying Eigenvectors)**.**

$$\min\big(\sqrt{P}, \sqrt{RTh_T}\big)\left\{\frac{1}{\sqrt{P}}\big\|\widehat{\boldsymbol{\Lambda}}_P(u) - \boldsymbol{\Lambda}_P(u)\mathbf{H}(u)\big\|\right\} = O_p(1), \qquad (9)$$

where $\mathbf{H}(u) = \big(\boldsymbol{\Sigma}^F\big)^{\frac{1}{2}}\boldsymbol{\Upsilon}(u)\mathbf{V}^{-\frac{1}{2}}(u)$*, and where* $\boldsymbol{\Upsilon}(u)$ *contains the eigenvectors of* $\big(\boldsymbol{\Sigma}^F\big)^{\frac{1}{2}}\boldsymbol{\Sigma}^{\Lambda}(u)\big(\boldsymbol{\Sigma}^F\big)^{\frac{1}{2}}$

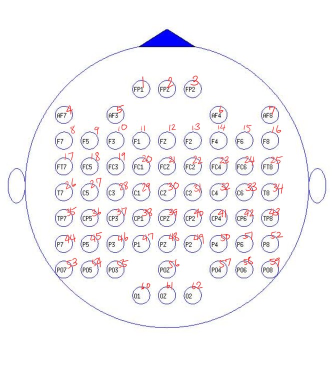
Theorem 4 (Consistency of the Factors)**.**

$$\min\big(\sqrt{P}, \sqrt{RTh_T}\big)\big\|\widehat{\boldsymbol{F}}(t,r) - \mathbf{H}^{-1}(u)\boldsymbol{F}(t,r)\big\| = O_p(1) \qquad (10)$$

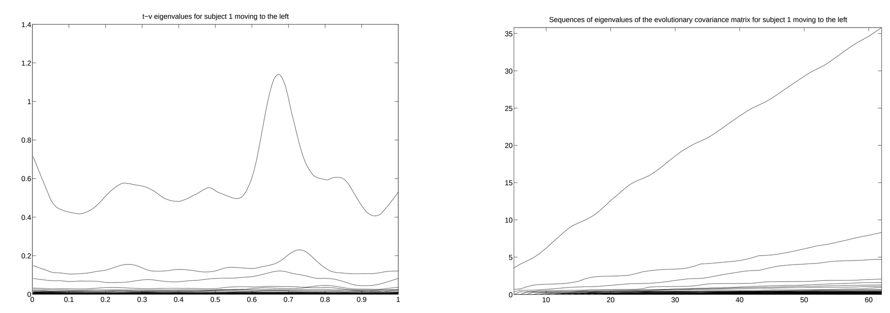
Corollary 1 (Consistency of the Common Components)**.**

$$\min\big(\sqrt{P}, \sqrt{RTh_T}\big)\big|\widehat{X}_{iT}(t,r) - X_{iT}(t,r)\big| = O_p(1) \qquad (11)$$

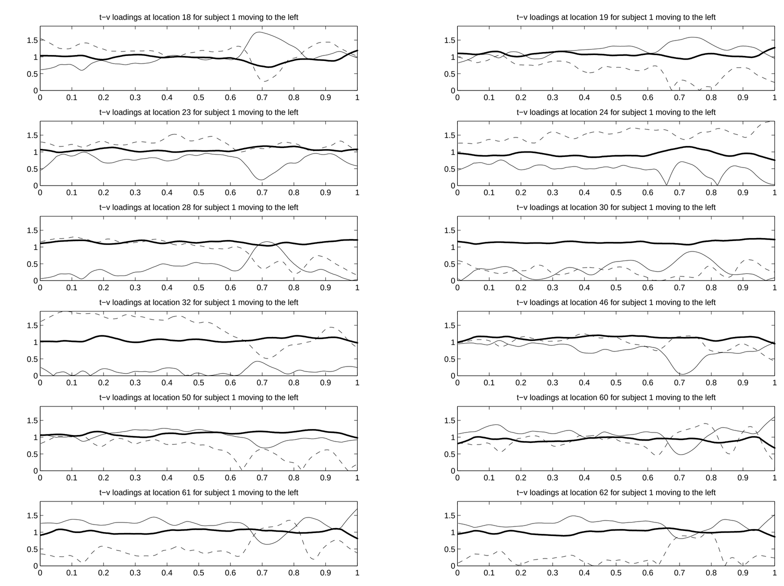
5. Empirical Results



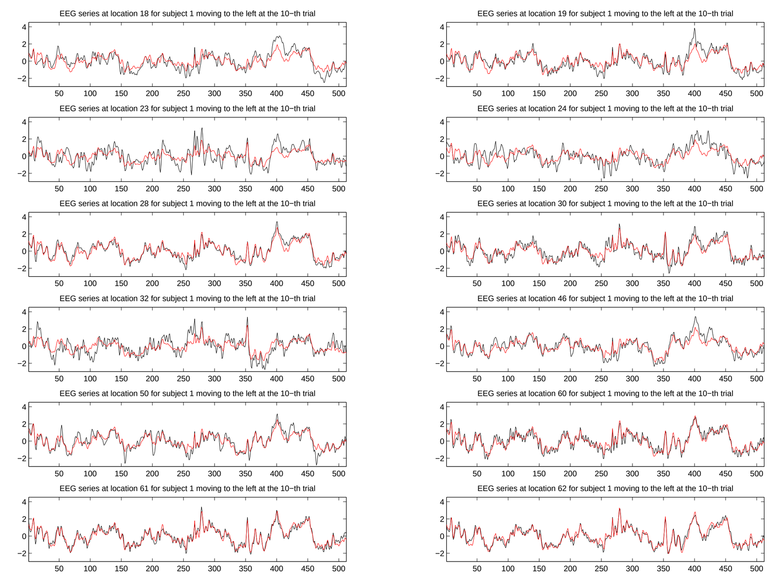
We observe EEG data recorded for an interval of one second, over $P = 62$ channel locations on the scalp. The subjects were commanded to move the joystick many times, with a total of $R = 118$ trials. For each trial, a time series is recorded. The start of the time series is 500 milliseconds before the command. The end of the time series is 500 milliseconds after the command. Here, one trial will have a total of $T = 512$ time points.



Left: time-varying eigenvalues $\widehat{\mathbf{V}}(u)$. Right: sequences of eigenvalues $\overline{\mathbf{V}}_p$, average over time of $\widehat{\mathbf{V}}_p(u)$ in (5), obtained from $\widehat{\boldsymbol{\Sigma}}_p(u)$; $p = 5, \dots, P = 62$. We select $\widehat{q} = 3$ factors.



Estimated matrix of the $(1 \times \widehat{q})$ time-varying loadings vectors $\widehat{\boldsymbol{\lambda}}_i(u)$ over $\widehat{q} = 3$ factors. Among $P = 62$ channels, we select 12 representative locations: $i = 18, 19, 23, 24, 28, 30, 32, 46, 50, 60, 61, 62$.



Black: observed data; red: common components defined in (6).

6. Conclusions

1. The series are non-stationary and they move together: this motivates the use of Evolutionary Factor Analysis
2. Our methodology gives mean-squared consistent results. Moreover, it combines information from several trials in a statistically meaningful and computationally efficient manner. As a result, we benefit from a faster rate of convergence of the estimators. The increase in the speed of the rate is given by the square root of the number of trials.
3. It seems there there are three main factors (located in the occipital lobe) driving the whole process: the first is a *background factor*, the others are *location factors*.
4. These three factors are weighted by some coefficients that are time-varying. In particular, it seems that the higher the weight on one factor the smaller the weight on the other two factors. This behavior of the loadings reflects two important properties of the dynamics of brain signals: smooth (time-changing) evolutions and orthogonality of the latent sources.
5. Time-varying source localization: the higher the loading weight $\lambda_{ij}(u)$ of channel i on factor j, the stronger the correlation between the signal observed at channel i and factor j.
6. Future research: generalization to the dynamic case of [2], where the loadings become time-varying filters and the common components are dynamic processes: $\boldsymbol{X}_{PT}(t) = \sum_{k=0}^{\infty}\boldsymbol{\Lambda}(\frac{t}{T}, k)\boldsymbol{F}(t-k)$.

Acknowledgement to Jerome Sanes (Neuroscience, Brown University) for sharing the EEG data set for illustration.

References

[1] DAHLHAUS, R. (1997). Fitting Time Series Models to Nonstationary Processes. *The Annals of Statistics*, Vol. 25, No. 1, 1-37.
[2] EICHLER, M., MOTTA, G. AND VON SACHS, R. (2008). Fitting Dynamic Factor Models to Non-stationary time series. Forthcoming in the *Journal of Econometrics*.
[3] MOTTA, G. AND OMBAO, H. (2010). Evolutionary Factor Analysis of Replicated Time Series. *Preprint*.
[4] MOTTA, G. (2009). Evolutionary Factor Analysis. *PhD Thesis*, Institute de Statistique, UCL, Belgium.
[5] MOTTA, G., HAFNER, C. AND VON SACHS, R. (2006). Locally Stationary Factor Models: Identification and Nonparametric Estimation. Forthcoming in *Econometric Theory*.