

- └ Learning, Forecasting and Retrospection: Examples
- └ Time-varying autoregressions and decompositions

A TVAR(p) process is given by

$$y_t = \phi_{1,t}y_{t-1} + \dots + \phi_{p,t}y_{t-p} + \epsilon_t, \quad \epsilon_t \sim N(0, v_t).$$

- *Evolution of AR coefficients.*

$$\phi_t = \phi_{t-1} + \mathbf{w}_t, \quad \mathbf{w}_t \sim N(\mathbf{0}, \mathbf{W}_t),$$

with $\phi_t = (\phi_{1,t}, \dots, \phi_{p,t})'$.

- *DLM representation:* $\{\mathbf{F}_t, \mathbf{I}_p, v_t, \mathbf{W}_t\}$ with $\mathbf{F}_t' = (y_{t-1}, \dots, y_{t-p})$.

- └ Learning, Forecasting and Retrospection: Examples
 - └ Time-varying autoregressions and decompositions

$$y_t = x_t + \nu_t, \quad x_t = \mathbf{F}'\boldsymbol{\theta}_t, \quad \boldsymbol{\theta}_t = \mathbf{G}_t\boldsymbol{\theta}_{t-1} + \mathbf{w}_t,$$

- Let \mathbf{G}_t be a $p \times p$ matrix with p different eigenvalues $\alpha_{t,1}, \dots, \alpha_{t,p}$. Assume C pairs of complex eigenvalues $r_{t,j}\exp(\pm i\omega_{t,j})$, $j = 1 : C$, and $R = p - 2C$ real eigenvalues $r_{t,j}$, $j = (C + 1) : (C + R)$.
- Then,

$$\mathbf{G}_t = \mathbf{B}_t \mathbf{A}_t \mathbf{B}_t^{-1}$$

where \mathbf{A}_t is the diagonal matrix of eigenvalues (in arbitrary but fixed order) and \mathbf{B}_t is a corresponding matrix of eigenvectors.

- Learning, Forecasting and Retrospection: Examples
 - Time-varying autoregressions and decompositions

Define $\mathbf{H}_t = \text{diag}(\mathbf{B}_t' \mathbf{F}) \mathbf{B}_t^{-1}$. Transform θ_t and \mathbf{w}_t via $\gamma_t = \mathbf{H}_t \theta_t$ and $\delta_t = \mathbf{H}_t \mathbf{w}_t$. Then,

$$y_t = x_t + \nu_t, \quad x_t = \mathbf{1}' \gamma_t, \quad \gamma_t = \mathbf{A}_t \mathbf{K}_t \gamma_{t-1} + \delta_t,$$

where $\mathbf{1}' = (1, \dots, 1)$ and $\mathbf{K}_t = \mathbf{H}_t \mathbf{H}_{t-1}^{-1}$.

General decomposition: $y_t = x_t + \nu_t$, with

$$x_t = \sum_{j=1}^C x_{t,j}^{(1)} + \sum_{j=1}^R x_{t,j}^{(2)}.$$

- ▶ $x_{t,j}^{(1)}$ real process related to $r_{t,j} \exp(\pm i \omega_{t,j})$, $j = 1 : C$.
- ▶ $x_{t,j}^{(2)}$ real process related to $r_{t,C+j}$, $j = 1 : R$.

AR(p) decomposition

Using the DLM representation $\{\mathbf{E}_p, \mathbf{G}(\phi), 0, \mathbf{W}\}$ we have:

- ▶ The eigenvalues of $\mathbf{G}(\phi)$ are the reciprocal roots of the characteristic polynomial, $\alpha_1, \dots, \alpha_p$, with C pairs of complex reciprocal roots $\alpha_{2j-1} = r_j \exp(-i\omega_j)$, $\alpha_{2j} = r_j \exp(i\omega_j)$, for $j = 1 : C$, and $\alpha_j = r_j$ for $j = (2C + 1) : p$.
- ▶ $\mathbf{A}_t = \mathbf{A} = \text{diag}(\alpha_1, \dots, \alpha_p)$, $\mathbf{B}_t = \mathbf{B}$, and $\mathbf{K}_t = \mathbf{I}$.
- ▶ **Decomposition:** $y_t = \sum_{j=1}^C x_{t,j}^{(1)} + \sum_{j=1}^R x_{t,j}^{(2)}$, with
 - ▶ $x_{t,j}^{(1)} \sim \text{ARMA}(2, 1)$ with **constant** modulus r_j and frequency ω_j , $j = 1 : C$;
 - ▶ $x_{t,j}^{(2)} \sim \text{AR}(1)$ with **constant** modulus r_{C+j} , $j = 1 : R$.

TVAR(p) decomposition

Using the DLM representation $\{\mathbf{E}_p, \mathbf{G}_t(\phi), 0, \mathbf{W}\}$ with

$$\mathbf{G}_t(\phi) = \begin{pmatrix} \phi_{1,t} & \phi_{2,t} & \phi_{3,t} & \dots & \phi_{p-1,t} & \phi_{p,t} \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \vdots & & & & & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \end{pmatrix}.$$

we have $\mathbf{A}_t = \text{diag}(\alpha_{1,t}, \dots, \alpha_{p,t})'$, \mathbf{B}_t the corresponding matrix of eigenvectors, and $\mathbf{K}_t = \mathbf{H}_t \mathbf{H}_{t-1}^{-1}$.

If ϕ_t does not vary a lot over time $\mathbf{K}_t \approx \mathbf{I}_p$. Assuming that C and R are constant over time we have the following decomposition...

TVAR(p) decomposition

$$y_t = \sum_{j=1}^C x_{t,j}^{(1)} + \sum_{j=1}^R x_{t,j}^{(2)},$$

- ▶ $x_{t,j}^{(1)} \approx \text{TVARMA}(2, 1)$ with **time-varying** modulus $r_{t,j}$ and frequency $\omega_{t,j}$, $j = 1 : C$;
 - ▶ $x_{t,j}^{(2)} \approx \text{TVAR}(1)$ with **time-varying** modulus $r_{t,C+j}$, $j = 1 : R$.
1. The structure of the underlying processes is *approximate*, but calculation of component processes is exact. The approximation depends on how close \mathbf{K}_t is to \mathbf{I}_p .
 2. R and C do not need to be constant over time.

Interpreting latent TVAR structure

How close are the latent processes in the decomposition to TVAR(1) and TVARMA(2, 1) obtained when $\mathbf{K}_t = \mathbf{I}_p$?

$$\begin{aligned} \mathcal{M}_1 : \quad y_t &= x_t + \nu_t & \mathcal{M}_2 : \quad y_t &= x_t + \nu_t \\ x_t &= \mathbf{1}'\gamma_t & x_t &= \mathbf{1}'\gamma_t \\ \gamma_t &= \mathbf{A}_t\mathbf{K}_t\gamma_{t-1} + \delta_t & \gamma_t &= \mathbf{A}_t\gamma_{t-1} + \delta_t. \end{aligned}$$

$$\begin{aligned} f_t^{(1)}(h) &= E(x_{t+h} \mid \gamma_t, \mathcal{M}_1) = \mathbf{1}'\mathbf{A}_{t+h}\mathbf{K}_{t+h}\dots\mathbf{A}_{t+1}\mathbf{K}_{t+1}\gamma_t, \\ f_t^{(2)}(h) &= E(x_{t+h} \mid \gamma_t, \mathcal{M}_2) = \mathbf{1}'\mathbf{A}_{t+h}\mathbf{A}_{t+h-1}\dots\mathbf{A}_{t+1}\gamma_t. \end{aligned}$$

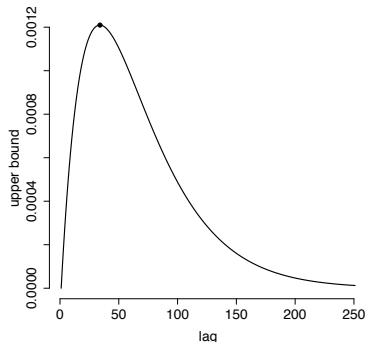
$$\frac{|f_t^{(1)}(h) - f_t^{(2)}(h)|}{\|\gamma_t\|_\infty} \leq (\lambda^*)^h \times [(1 + \epsilon^*)^h - 1],$$

with $\lambda^* = \max_{0 \leq j \leq h-1} (\max_{1 \leq i \leq p} |\lambda_{((t+h-j), i)})|$;

$\epsilon^* = \max_{1 \leq j \leq h} \|\mathcal{E}_{t+j}\|_\infty$, $\mathbf{K}_{t+j} = \mathbf{I}_p + \mathcal{E}_{t+j}$.

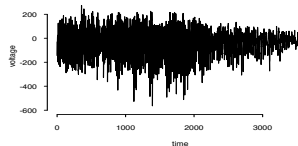
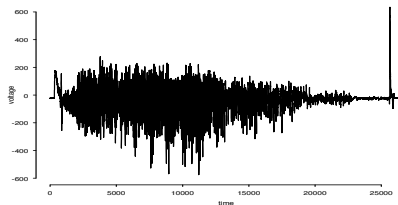
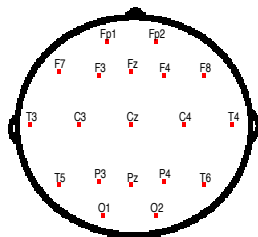
TVAR(p) such that, for all $j > t$, $\|\mathcal{E}_{t+j}\|_{\infty} \leq 10^{-4}$ and that all the characteristic reciprocal roots have moduli less than 0.97.

$$\frac{|f_t^{(1)}(h) - f_t^{(2)}(h)|}{\|\gamma_t\|_{\infty}} \leq 0.97^h \times [(1.0001)^h - 1].$$



Example: Inferring latent structure in EEGs

- **Data:** 3,600 EEG observations recorded on a patient who received electroconvulsive therapy (ECT). Subsampled central portion of channel Cz. 18 additional channels.



Example: Inferring latent structure in EEGs► **Model:**

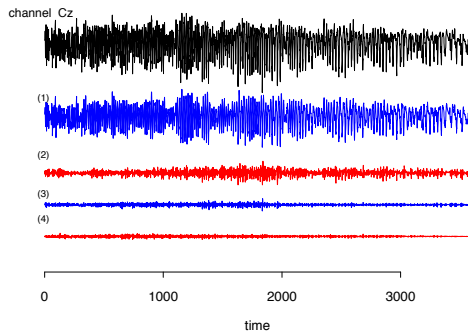
$$\begin{aligned}
 y_t &= \sum_{j=1}^p \phi_{j,t} y_{t-j} + \nu_t, \quad \nu_t \sim N(0, \mathbf{v}_t), \\
 \phi_t &= \phi_{t-1} + \xi_t, \quad \xi_t \sim N(\xi_t \mid \mathbf{0}, \mathbf{U}_t(\delta_\phi)), \\
 \mathbf{v}_t &= \delta_v \mathbf{v}_{t-1} / \eta_t, \quad \eta_t \sim Be(\eta_t \mid a_t, b_t),
 \end{aligned}$$

with δ_ϕ and δ_v discount factors in $(0, 1]$. Optimal values of δ_ϕ , δ_v , and p obtained by maximizing

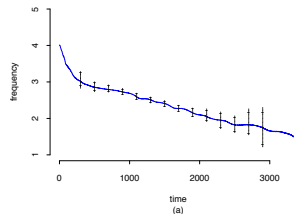
$$l(\delta_\phi, \delta_v, p) \equiv \log[y_{(p^*+1):T} \mid \mathcal{D}_{p^*}] = \sum_{t=p^*+1}^T \log[p(y_t \mid \mathcal{D}_{t-1})],$$

over $[0.9, 1] \times [0.9, 1] \times [4, 20]$, with $p^* = 20$. $\hat{\delta}_\phi = 0.994$, $\hat{\delta}_v = 0.95$, and $\hat{p} = 12$.

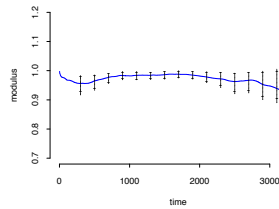
EEG decomposition



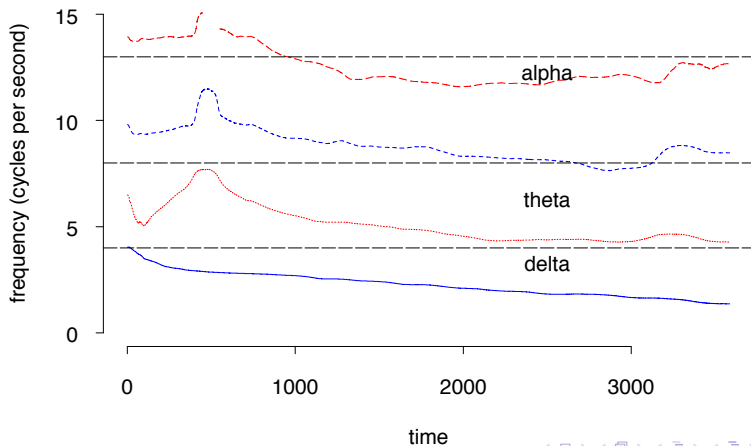
Frequency



Modulus

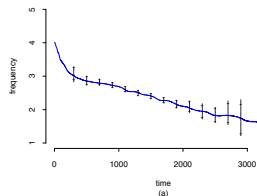
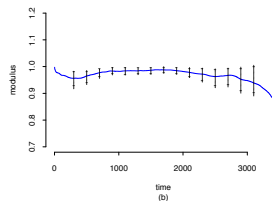
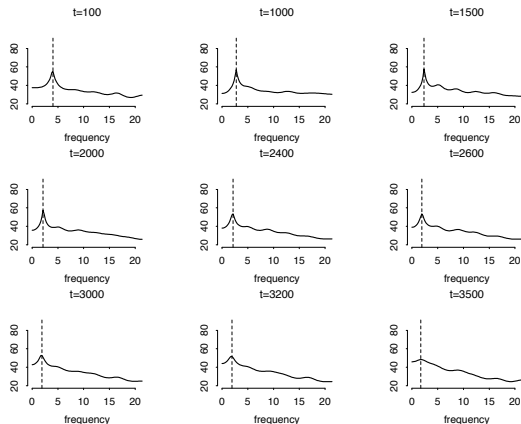


Ordering of the latent components



EEG data: Estimated time-varying spectra

$$r_{t,j}, \omega_{t,j}$$

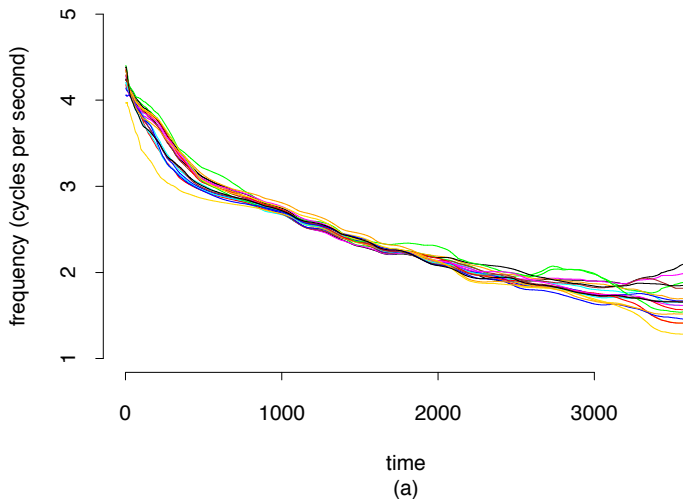


TVAR models for multiple time series

Case study: ECT data

- ▶ Fit TVAR(12) models to each of the 19 EEG signals.
- ▶ Extract dominant frequencies and moduli.

EEG data: Trajectories of dominant frequencies



EEG data: Moduli of dominant frequencies

