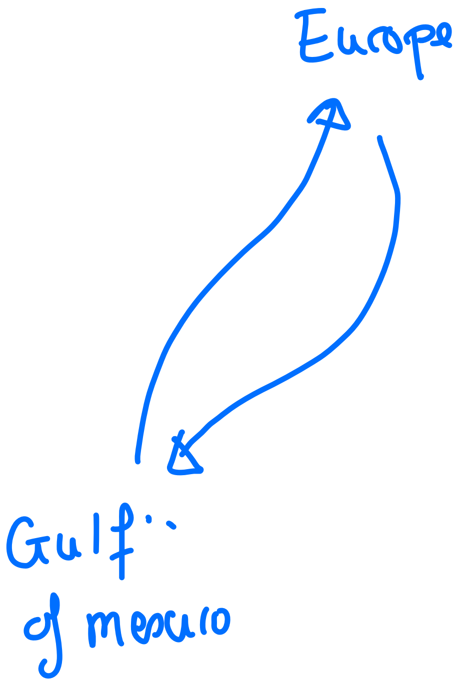
Graphical models for multivariate time series

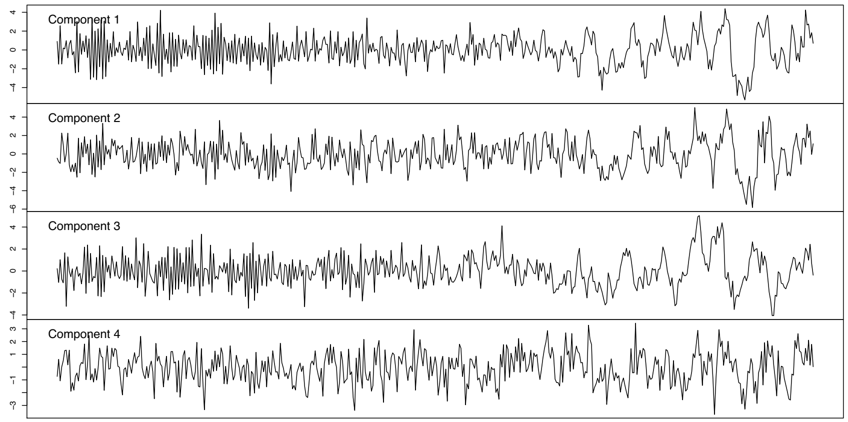
(or Interpreting the inverse of very big matrices)

Kemp's Ridley Turtle



Motivating example

Consider the following realisations from the multivariate time series $\{X_t^\top = (X_t^{(1)}, X_t^{(2)}, X_t^{(3)}, X_t^{(4)})\}_{t=1}^n$



Pairwise relationships:

- Visually, it appears that every component is correlated with the others.
- Visually, it is clear that every component is nonstationary.

Formally, we would test and (a) H_0 : no correlation vs H_1: correlation and (b) H_0 : second order stationary (i.e. for all $t.\tau \in \mathbb{Z}$ $\mathrm{cov}[X_t^{(a)}, X_\tau^{(b)}] = c_{a,b}(t-\tau)$) vs H_1 : second order nonstationary.

- All tests strongly suggests the time series is both pairwise correlated *and* nonstationary. But visually this is quite "obvious".

- Aim Move away from "pairwise" relationships. Conditional or "system-wide" relationships may give us greater insight into dependence structures in the time series.

  $\Rightarrow$ Take the GGM route.

Review: Gaussian Graphical Models (GGM)

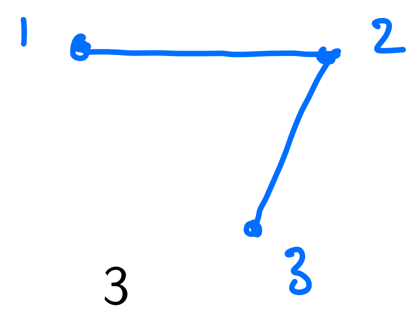
- $X = (X^{(1)}, \ldots, X^{(p)})'$ is a Gaussian random vector with $\mathrm{var}[X] = \Sigma$.

- A graph is defined by the node set $V = \{1, \ldots, p\}$ and edge set $E \subseteq V \times V$.

- Define the vector $X^{V \setminus \{a,b\}}$ which is X without $X^{(a)}$ and $X^{(b)}$.

- Conditional Graph definition An undirected conditional graph $G = (V, E)$ places an edge between two nodes a and b if $\mathrm{cov}[X^{(a)}, X^{(b)} | X^{V \setminus \{a,b\}}] \neq 0$, where $X^{V \setminus \{a,b\}}$ is the random vector with $(X^{(a)}, X^{(b)})$ removed.

$p = 3$



Workhorse of GGM: the precision matrix

(i) $\Sigma = \mathrm{var}[X]$,

$$\Sigma^{-1} = \begin{pmatrix} d^{(1,1)} & d^{(1,2)} & \dots & d^{(1,p)} \\ d^{(2,1)} & d^{(2,2)} & \dots & d^{(2,p)} \\ \vdots & \vdots & \ddots & \vdots \\ d^{(p,1)} & d^{(p,2)} & \dots & d^{(p,p)} \end{pmatrix}.$$

- Then $\mathrm{cov}[X^{(a)}, X^{(b)} | X^{V \setminus \{a,b\}}] = 0$ iff $[\Sigma^{-1}]_{a,b} (= d^{(a,b)}) = 0$:
- We below show why this is true. Classical results in multivariate analysis:

Example

a = 1

b = 2

$$\text{var}\begin{pmatrix} X^{(1)} & |X^{V\backslash\{1,2\}} \\ X^{(2)} & \end{pmatrix}$$

$$= \begin{pmatrix} c^{(1,1)} & c^{(1,2)} \\ c^{(1,2)} & c^{(2,2)} \end{pmatrix} - \begin{pmatrix} c^{(1,3)} & \ldots & c^{(1,p)} \\ c^{(2,3)} & \ldots & c^{(2,p)} \end{pmatrix} \begin{pmatrix} c^{(3,3)} & \ldots & c^{(3,p)} \\ \vdots & \ddots & \vdots \\ c^{(p,3)} & \ldots & c^{(p,p)} \end{pmatrix}^{-1} \begin{pmatrix} c^{(1,3)} & c^{(2,3)} \\ \vdots & \vdots \\ c^{(1,p)} & c^{(2,p)} \end{pmatrix}$$

$$= \begin{pmatrix} d^{(1,1)} & d^{(1,2)} \\ d^{(1,2)} & d^{(2,2)} \end{pmatrix}^{-1}$$

- It is clear if $d^{(1,2)}$ is zero, then the above matrix results shows that the partial covariance $\text{cov}[X^{(1)}, X^{(2)}|X^{V\backslash\{1,2\}}] = 0$.

- Entries in the precision matrix are also connected to those in linear regression. Thus tools from linear regression are often used to estimate conditional graphs.

  Inverting matrices (the precision matrix) to learn conditional relationships is central to this talk.

Stationary Graphical Models (StGM)

- Suppose $\{X_t^\top = (X_t^{(1)}, \ldots, X_t^{(p)})\}_t$ is a p-dimensional (Gaussian) second order stationary time series where $\mathrm{cov}[X_t, X_\tau] = C_{t-\tau} = (c_{t-\tau}^{(a,b)}; 1 \leq a, b \leq p)$.

- Directly applying GGM to time series data would mean only considering contemporaneous interactions:

$$\mathrm{var}(X_t)^{-1} = C_0^{-1} = \begin{pmatrix} c_0^{(1,1)} & c_0^{(1,2)} & \ldots & c_0^{(1,p)} \\ c_0^{(2,1)} & c_0^{(2,2)} & \ldots & c_0^{(2,p)} \\ \vdots & \vdots & \ddots & \vdots \\ c_0^{(p,1)} & c_0^{(p,2)} & \ldots & c_0^{(p,p)} \end{pmatrix}^{-1}$$

- It would completely ignore lead lag correlations between different time points.

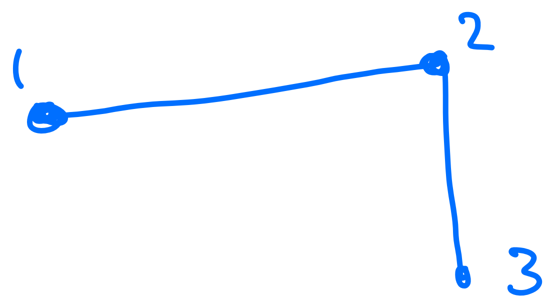
- Dahlhaus (2000) proposed StGM based on the partial correlation/covariance commonly used the multivariate time series.

- Define the linear space $X^{V\setminus\{a,b\}} = \overline{\text{sp}}(X_t^{(c)}, t \in \mathbb{Z}, 1 \leq c \leq p, c \neq a, b)$

- The time series partial covariance is defined as $\text{cov}[X_t^{(a)}, X_\tau^{(b)} | X^{V\setminus\{a,b\}}]$.

- Conditional Graph definition In StGM an edge exists between nodes a and b iff $\text{cov}[X_t^{(a)}, X_\tau^{(b)} | X^{V\setminus\{a,b\}}] \neq 0$ for some (t, τ).

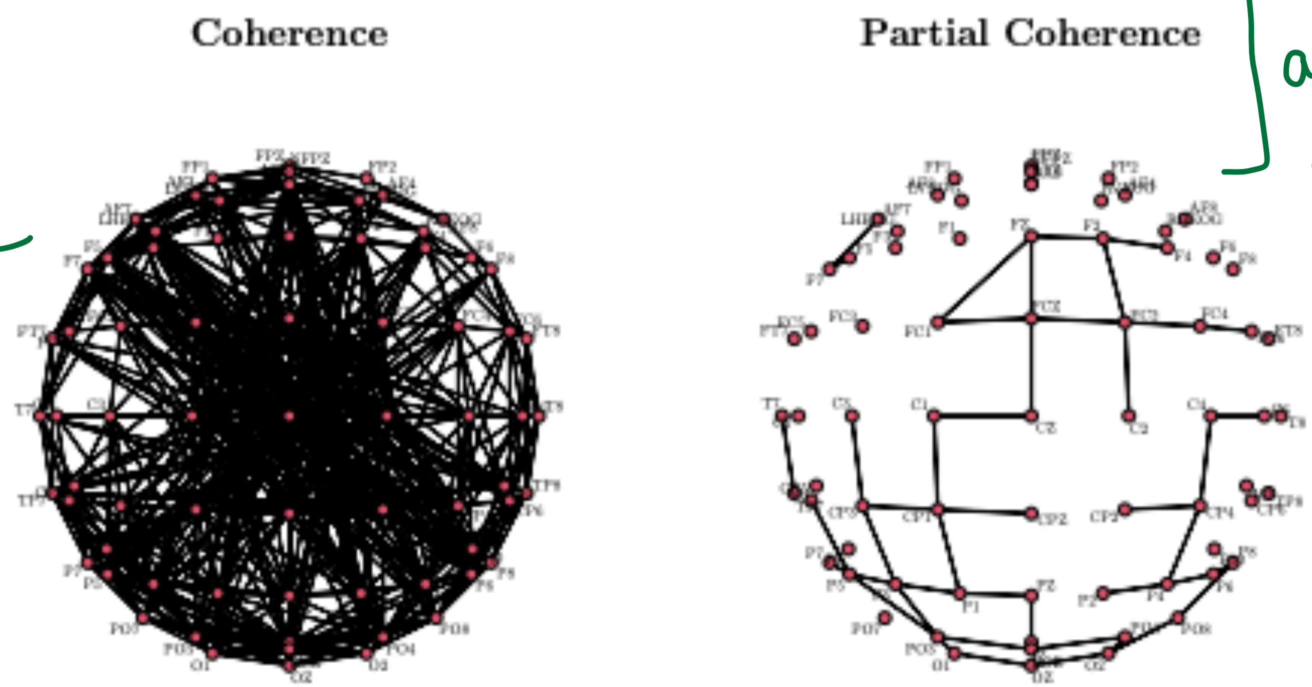
$p = 3$



Application to EEG data

Connection if there is correlation between different channels

Connection if there is correlation between channels after removing dependence on all other channels.



Taken from Krampe and Paparoditis https://arxiv.org/pdf/2206.02250.pdf

- Later in the talk I will explain how the multivariate partial correlation is related to an inverse covariance matrix (an infinite dimensional one).

- However, estimation of "big" matrices can be (very) difficult. Instead what is done is to transform to the so called frequency domain.

- Define the so called $p \times p$ spectral density matrix as

$$f(\omega) = \sum_{r=-\infty}^{\infty} C_r \exp(ir\omega) \qquad \omega \in [0, 2\pi].$$

- The spectral precision matrix is $f(\omega)^{-1}$ (it is a $p \times p$ matrix).

- It can be shown that $\mathrm{cov}[X_t^{(a)}, X_\tau^{(b)} | X^{V \backslash \{a,b\}}] = 0$ iff $[f(\omega)^{-1}]_{a,b} = 0$ for all $\omega \in [0, 2\pi]$ (we explain why later).

- Estimation of graph: Low dimension (fixed p) Eichler (2008), Böhm and van Sachs (2009). High dimensions (large p): Fiecas et. al. (2018), Basu et. al. (2022) and Krampe and Paparoditis (2022).

- Our objective: Generalize these notions to (second order) nonstationary multivariate time series.

Nonstationary StGM (NonStGM)

- Suppose $\{X_t\}_t$ is a zero mean p-dimensional multivariate nonstationary time series in the sense that $C_{t,\tau} = \mathrm{cov}[X_t, X_\tau]$. We do not place any modelling assumptions on $\{X_t\}_t$ and the approach is nonparametric.

- Our aim is to build a parsimonious network for $\{X_t\}_{t=1}^{n}$ that describes conditional relationships between the components of the time series.
  - Conditional (non)correlation (analogous to Gaussian Graphical Models).
  - Conditional (non)stationarity (a notion we introduce).

- We show that this information can be found in an infinite dimensional precision matrix. Like StGM, we show that frequency domain can better encode conditional relationships.

Motivation: Example for stationary case

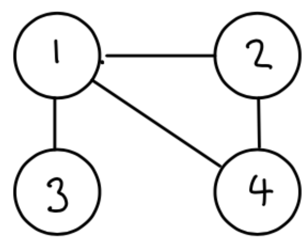
Graphical models for the stationary VAR(1)

$$\begin{pmatrix} X_t^{(1)} \\ X_t^{(2)} \\ X_t^{(3)} \\ X_t^{(4)} \end{pmatrix} = \begin{pmatrix} \alpha_1 & 0 & \alpha_3 & 0 \\ \beta_1 & \beta_2 & 0 & \beta_4 \\ 0 & 0 & \gamma_3 & 0 \\ \nu_1 & 0 & 0 & \nu_4 \end{pmatrix} \begin{pmatrix} X_{t-1}^{(1)} \\ X_{t-2}^{(2)} \\ X_{t-3}^{(3)} \\ X_{t-4}^{(4)} \end{pmatrix} + \varepsilon_t = A X_{t-1} + \varepsilon_t$$

$\{\varepsilon_t\}_t$ are iid N$(0, I_4)$.

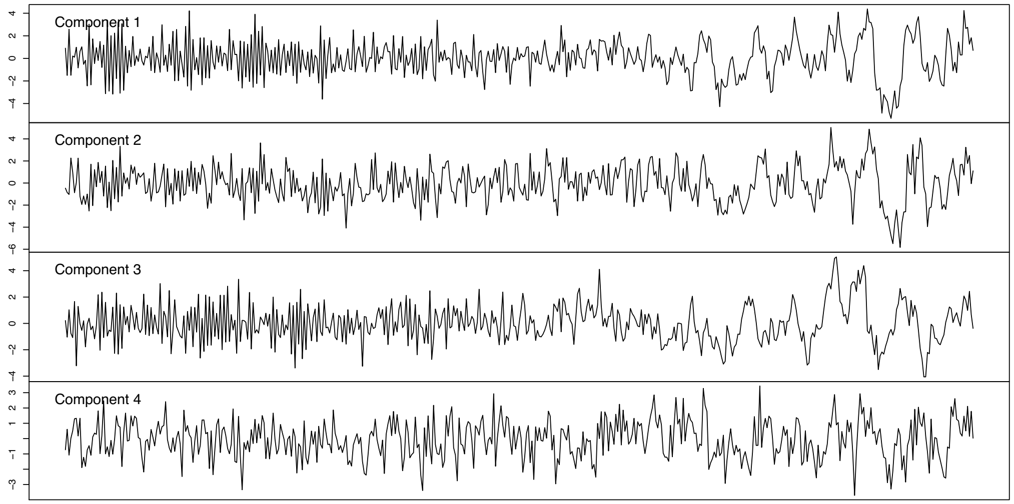
Brillinger (1996) and Dahlhaus (2000) connect the entries of the transition matrix to those of conditional uncorrelatedness.

$$\begin{pmatrix} \alpha_1 & 0 & \alpha_3 & 0 \\ \beta_1 & \beta_2 & 0 & \beta_4 \\ 0 & 0 & \gamma_3 & 0 \\ \nu_1 & 0 & 0 & \nu_4 \end{pmatrix}$$



Nonstationary graphical models: Running example

$$\begin{pmatrix} X_t^{(1)} \\ X_t^{(2)} \\ X_t^{(3)} \\ X_t^{(4)} \end{pmatrix} = \begin{pmatrix} \alpha(t) & 0 & \alpha_3 & 0 \\ \beta_1 & \beta_2 & 0 & \beta_4 \\ 0 & 0 & \gamma(t) & 0 \\ 0 & \nu_2 & 0 & \nu_4 \end{pmatrix} \begin{pmatrix} X_{t-1}^{(1)} \\ X_{t-1}^{(2)} \\ X_{t-1}^{(3)} \\ X_{t-1}^{(4)} \end{pmatrix} + \varepsilon_t = A(t)X_{t-1} + \varepsilon_t,$$



- Pairwise each component in the time series is *nonstationary*.

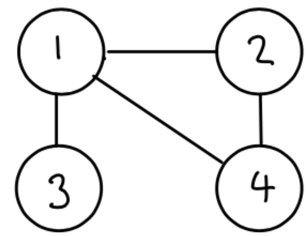
- However, components (1) and (3) are driving the nonstationarity which propogates through to component (2) and (4).

A potential network for nonstationary data?

As the zeros in the stationary AR transition matrix match those in the time-varying AR it seems plausible that the same network holds:

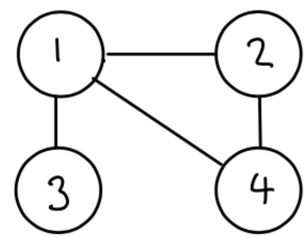
(i) Stationary $X_t = AX_{t-1} + \varepsilon_t$

$$\begin{pmatrix} \alpha_1 & 0 & \alpha_3 & 0 \\ \beta_1 & \beta_2 & 0 & \beta_4 \\ 0 & 0 & \gamma_3 & 0 \\ \nu_1 & 0 & 0 & \nu_4 \end{pmatrix}$$

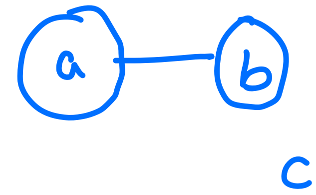


(ii) Nonstationary $X_t = A(t)X_{t-1} + \varepsilon_t$?

$$\begin{pmatrix} \alpha(t) & 0 & \alpha_3 & 0 \\ \beta_1 & \beta_2 & 0 & \beta_4 \\ 0 & 0 & \gamma(t) & 0 \\ 0 & \nu_2 & 0 & \nu_4 \end{pmatrix}$$



- Question What network conveys meaningful information about nonstationary time series?



Def: Conditional graph for nonstationary time series

- Two nodes a and b are said to be **conditionally uncorrelated** if $\mathrm{cov}[X_t^{(a)}, X_\tau^{(b)} | X^{V \setminus \{a,b\}}] = 0$ for all $t, \tau \in \mathbb{Z}$.

- A **node** a is **conditionally stationary** if for all t, τ $\rho_{t-\tau}^{(a,a)|-\{a\}} = \mathrm{cov}[X_t^{(a)}, X_\tau^{(a)} | X^{V \setminus \{a\}}]$

- An **edge** (a, b) is **conditionally time invariant** if for all t, τ $\rho_{t-\tau}^{(a,b)|-\{a,b\}} = \mathrm{cov}[X_t^{(a)}, X_\tau^{(b)} | X^{V \setminus \{a,b\}}]$.

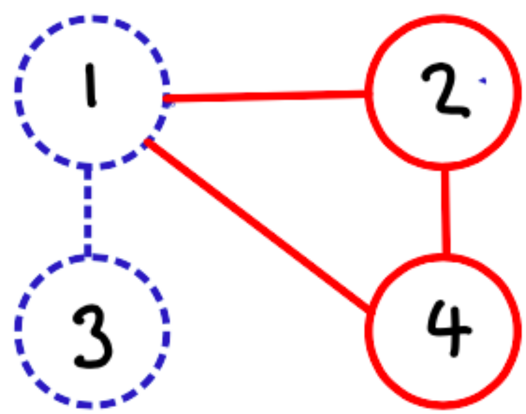
- A **node or edge** (a, b) is **conditionally nonstationary/time-varying** if

  $\rho_{t,\tau}^{(a,b)|-\{a,b\}} = \mathrm{cov}[X_t^{(a)}, X_\tau^{(b)} | X^{V \setminus \{a,b\}}]$ (is not shift invariant).

Running example: NonStGM Network

$$\begin{pmatrix} X_t^{(1)} \\ X_t^{(2)} \\ X_t^{(3)} \\ X_t^{(4)} \end{pmatrix} = \begin{pmatrix} \alpha(t) & 0 & \alpha_3 & 0 \\ \beta_1 & \beta_2 & 0 & \beta_4 \\ 0 & 0 & \gamma(t) & 0 \\ 0 & \nu_2 & 0 & \nu_4 \end{pmatrix} \begin{pmatrix} X_{t-1}^{(1)} \\ X_{t-1}^{(2)} \\ X_{t-1}^{(3)} \\ X_{t-1}^{(4)} \end{pmatrix} + \underline{\varepsilon}_t = A(t)\underline{X}_{t-1} + \underline{\varepsilon}_t,$$

Using the Cholesky decomposition, the infinite dimensional inverse covariance can be deduced from the model.
It can be shown that the conditional relationships are described by this network:



- Like GGM, we show these partial covariances are encoded within a precision matrix.

Encoding in terms of matrices

- Define the infinite dimensional matrix operator C containing all the covariances of $\{X_t\}_{t\in\mathbb{Z}}$

$$C = \begin{pmatrix} C_{1,1} & C_{1,2} & \dots & C_{1,p} \\ C_{2,1} & C_{2,2} & \dots & C_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ C_{p,1} & C_{p,2} & \dots & C_{p,p} \end{pmatrix},$$

$$C_{1,2} = \left\{ \mathrm{cov}\left(X_t^{(1)}, X_\tau^{(2)}\right); t, \tau \in \mathbb{Z} \right\}$$

Infinite dimensional matrix.

where $C_{a,b}$ contains all the cross covariances between $X^{(a)}$ and $X^{(b)}$.

Difference to most covariance matrices each "entry" in this matrix is infinite dimensional.

- The central assumption is the eigenvalues of C are bounded away from zero and are finite.

Encoding these relationships in the precision matrix

- The conditional relationships (described above) are encoded in an *infinite* dimensional precision matrix

$$
\begin{aligned}
D = C^{-1} \quad &= \quad \begin{pmatrix} C_{1,1} & C_{1,2} & \dots & C_{1,p} \\ C_{2,1} & C_{2,2} & \dots & C_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ C_{p,1} & C_{p,2} & \dots & C_{p,p} \end{pmatrix}^{-1} \\
&= \quad \begin{pmatrix} D_{1,1} & D_{1,2} & \dots & D_{1,p} \\ D_{2,1} & D_{2,2} & \dots & D_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ D_{p,1} & D_{p,2} & \dots & D_{p,p} \end{pmatrix}.
\end{aligned}
$$

Comparison between regular precision and infinite precision

- Recall the regular partial covariance result:

$$\mathrm{var}\left(\begin{array}{c} X^{(a)} \\ X^{(b)} \end{array} | X^{V\backslash\{a,b\}} \right) = \left(\begin{array}{cc} d^{(a,a)} & d^{(a,b)} \\ d^{(b,a)} & d^{(b,b)} \end{array} \right)^{-1}$$

- We can show a similar result for infinite dimensional matrices:

(i) $\mathrm{var}\left[X_t^{(a)|V\backslash\{a\}}; t \in \mathbb{Z}\right] = D_{a,a}^{-1}$.

(ii) If $a \neq b$, then

$$\mathrm{var}\left[X_t^{(c)|V\backslash\{a,b\}}; t \in \mathbb{Z}, c \in \{a,b\}\right] = \left(\begin{array}{cc} D_{a,a} & D_{a,b} \\ D_{b,a} & D_{b,b} \end{array} \right)^{-1}$$

- These are the partial covariances of a multivariate time series.

The conditional graphs in terms of the precision infinite dimensonal matrix

- Two nodes a and b are **conditionally uncorrelated** iff $D_{a,b} = 0$ (the zero matrix).

- A **node** a is **conditionally stationary** iff $D_{a,a}$ is Toeplitz. [1]

- An **edge** (a, b) is **conditionally invariant** iff $D_{a,b}$ is a Toeplitz matrix.

- A node a or edge (a, b) is **conditionally nonstationary/time-varying** if $D_{a,a}$ or $D_{a,b}$ is *not* Toeplitz.

[1]Definition: a Toeplitz matrix is a matrix whose rows are successive shifts of one sequence.

Toeplitz Matrix

$$
= \begin{pmatrix}
 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\
\cdots & c_{-2} & c_{-1} & c_0 & c_1 & c_2 & c_3 & \cdots \\
\cdots & c_{-3} & c_{-2} & c_{-1} & c_0 & c_1 & c_2 & \cdots \\
\cdots & c_{-4} & c_{-3} & c_{-2} & c_{-1} & c_0 & c_1 & \cdots \\
\cdots & c_{-5} & c_{-4} & c_{-3} & c_{-2} & c_{-1} & c_0 & \cdots \\
 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots &
\end{pmatrix}
$$

Sparsity relations in the Fourier domain

- Directly estimating D in order to estimate the conditional graph is difficult.

- Instead, we transform the matrix, which can "sparsify" it (estimating "large" sparse matrices is easier).

- Recall, if $\mathrm{var}(X) = C$ and A is a deterministic, then the linear transformation AX is such that $\mathrm{var}(AX) = ACA^\top$.

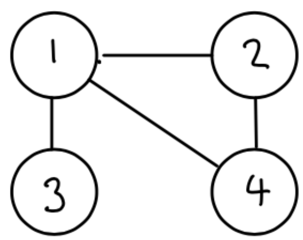
- If one takes the Fourier transform matrix $\mathcal{F}$ (a special transform consisting of signs and cosines) and $\mathrm{var}(X) = C$, where C is a Toeplitz matrix, then $\mathrm{var}(AX) = \mathcal{F}C\mathcal{F}^* = \mathrm{diagonal\ ''matrix''}$

Example: Multivariate stationary time series

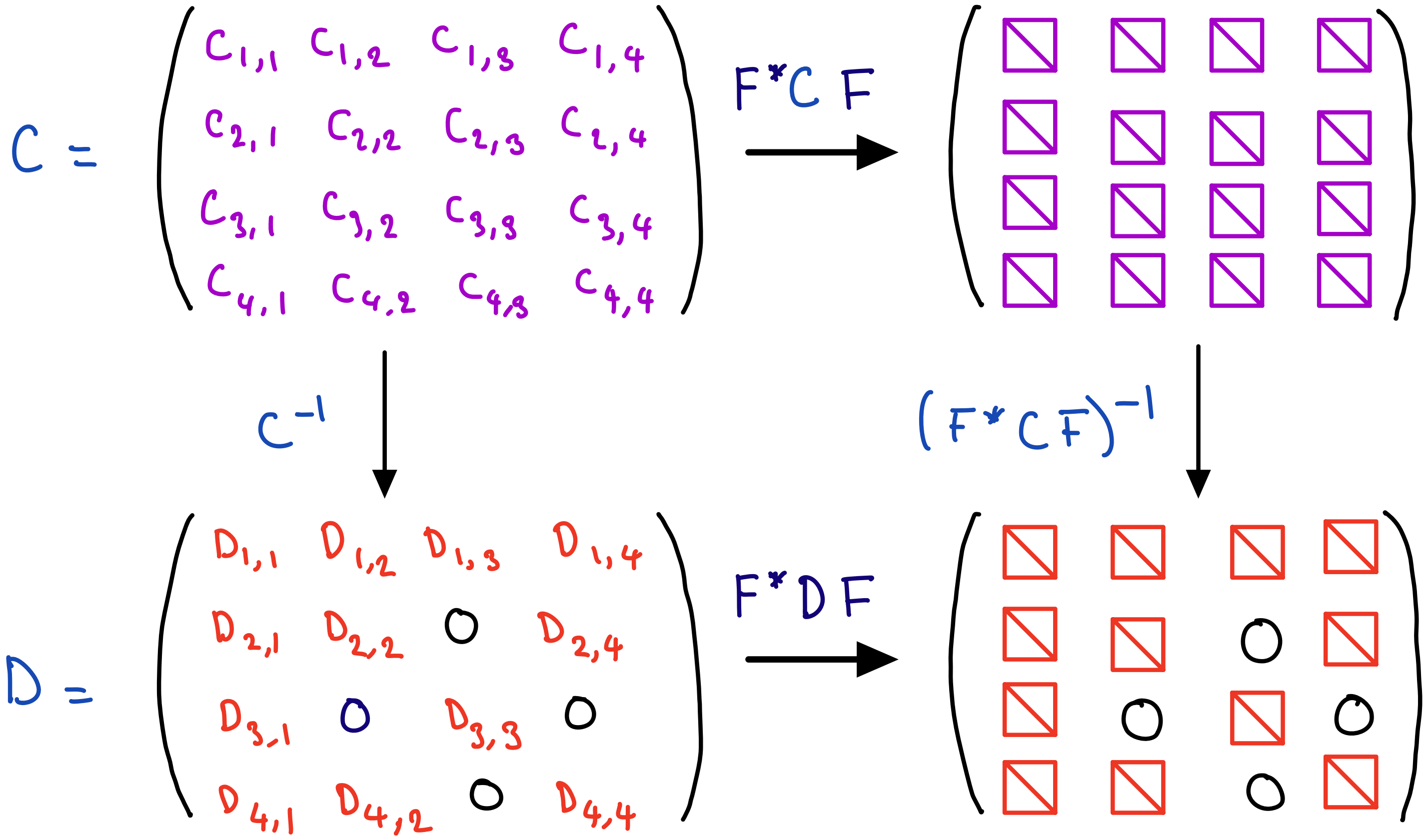
- Consider, as an example, the multivariate stationary VAR model

$$\begin{pmatrix} X_t^{(1)} \\ X_t^{(2)} \\ X_t^{(3)} \\ X_t^{(4)} \end{pmatrix} = \begin{pmatrix} \alpha_1 & 0 & \alpha_3 & 0 \\ \beta_1 & \beta_2 & 0 & \beta_4 \\ 0 & 0 & \gamma_3 & 0 \\ \nu_1 & 0 & 0 & \nu_4 \end{pmatrix} \begin{pmatrix} X_{t-1}^{(1)} \\ X_{t-2}^{(2)} \\ X_{t-3}^{(3)} \\ X_{t-4}^{(4)} \end{pmatrix} + \varepsilon_t = AX_{t-1} + \varepsilon_t$$

$\{\varepsilon_t\}_t$ are iid $var(\varepsilon_t) = I_4$.



Stationarity means that the covariance matrix and its inverse are both Toeplitz matrices



Each diagonal matrix ⧅ corresponds the spectral density

$$⧅_{a,b} = \text{diag}\left(f_{a,b}(\omega) = \sum_{r=-\infty}^{\infty} c_{a,b}(r)\, e^{ir\omega} ;\ \omega \in [0, 2\pi] \right)$$

Whereas each diagonal matrix ⧅ corresponds the inverse spectral density

$$⧅_{a,b} = \text{diag}\left(g^{(a,b)}(\omega) = \sum_{r=-\infty}^{\infty} d_{a,b}(r)\, e^{ir\omega} ;\ \omega \in [0, 2\pi] \right)$$

↑ sometimes called partial spectral coherence.

(*) Zero matrices get mapped to zero diagonal function

$= \Sigma_1(\omega)^{-1}$

$$\mathcal{F}D\mathcal{F}^* = \mathrm{diag}\left(\begin{pmatrix} g^{(1,1)}(\omega) & g^{(1,2)}(\omega) & g^{(1,3)}(\omega) & g^{(1,4)}(\omega) \\ g^{(2,1)}(\omega) & g^{(2,2)}(\omega) & 0 & g^{(2,4)}(\omega) \\ g^{(3,1)}(\omega) & 0 & g^{(3,3)}(\omega) & 0 \\ g^{(4,1)}(\omega) & g^{(4,2)}(\omega) & 0 & g^{(4,4)}(\omega) \end{pmatrix}; \omega \in [0, 2\pi]\right)$$

- We use the Discrete Fourier Transform (DFT) to estimate the conditional graph.

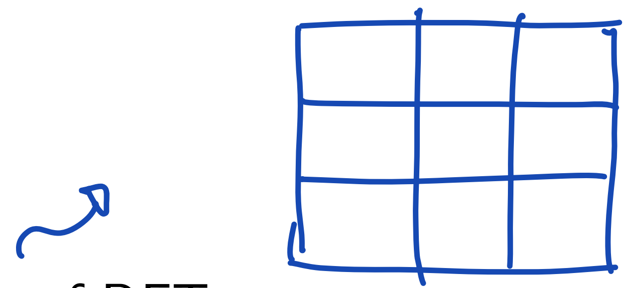
  The DFT of the univariate time series $\{X_t^{(a)}\}_{t=1}^n$ is defined as $J_k^{(a)} = \frac{1}{n}\sum_{t=1}^n X_t^{(a)} \exp(it\omega_k)$ where $\omega_k = 2\pi k/n$ and $\underline{J}_k = (J_k^{(a)})$.

- Since $\Sigma(\omega_k)^{-1} \approx \mathrm{var}[\underline{J}_k]^{-1}$, we estimate the entries in $\Sigma(\omega_k)^{-1}$ using regression of the DFTs.

Returning to nonstationary multivariate time series

- For $1 \leq a \leq p$ string all the frequencies in a vector $\mathbf{J}_n^{(a)} = (J_1^{(a)}, J_2^{(a)}, \ldots, J_n^{(a)})'$ and define the concatonated np-dimensional vector

$$\mathbf{J}_n = \text{vec}\left(\mathbf{J}_n^{(1)}, \mathbf{J}_n^{(2)}, \ldots, \mathbf{J}_n^{(p)} \right).$$



- $\text{var}[\mathbf{J}_n]$ is a $p \times p$ block variance matrix of DFTs.

- Idea In GGM conditional relationships are encoded in the inverse (precision) covariance matrix.

  In NonStGM, $D = C^{-1}$ encodes nonstationary relationships. Could the inverse of the variance of the Fourier transform, $(\text{var}[\mathbf{J}_n])^{-1}$ contain “useable” information too?

The DFT precision matrix and conditional relationships

- $(\mathrm{var}[\mathbf{J}_n])^{-1}$ is $p \times p$-dimensional block matrix, each block has dimension $n \times n$:

$$K_n = (\mathrm{var}[\mathbf{J}_n])^{-1} = \begin{pmatrix} (K_n)_{1,1} & (K_n)_{1,2} & \dots & (K_n)_{1,p} \\ (K_n)_{2,1} & (K_n)_{2,2} & \dots & (K_n)_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ (K_n)_{p,1} & (K_n)_{p,2} & \dots & (K_n)_{p,p} \end{pmatrix}.$$

  where $(K_n)_{a,b} = (K_{a,b}(\omega_{k_1}, \omega_{k_2}); 1 \leq k_1, k_2 \leq n)$.

- Theorem Conditional relationships between (a, b) are encoded in $(K_n)_{a,b}$.

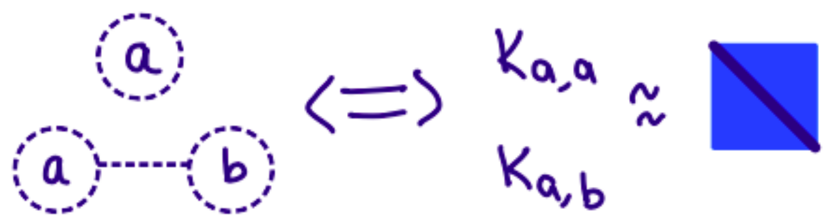
Theorem: Sparsity in the Fourier domain

- Two nodes a and b are said to be conditionally uncorrelated if $(K_n)_{a,b} \approx 0$.

- A node a is conditionally stationary if $(K_n)_{a,b}$ is approx diagonal.

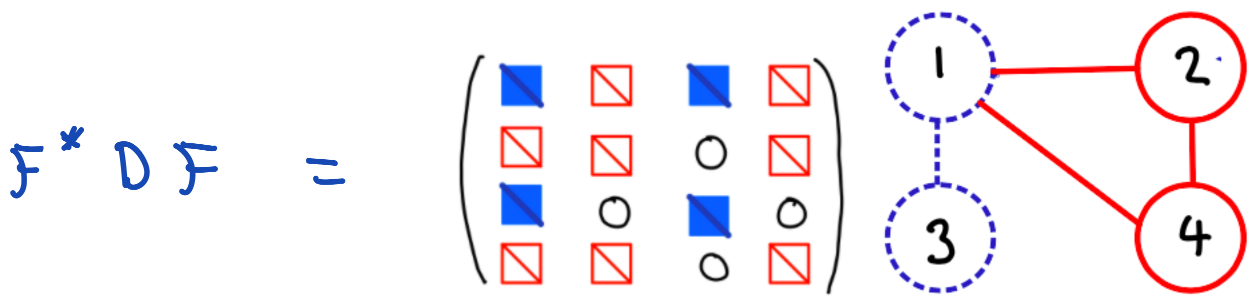
- An edge (a, b) is conditionally time-invariant if $(K_n)_{a,b}$ is approx diagonal.

- A node or edge (a, b) is conditionally nonstationary/time varying if $(K_n)_{a,b}$ is not diagonal.



The running example in the frequency domain

$$\begin{pmatrix} X_t^{(1)} \\ X_t^{(2)} \\ X_t^{(3)} \\ X_t^{(4)} \end{pmatrix} = \begin{pmatrix} \alpha(t) & 0 & \alpha_3 & 0 \\ \beta_1 & \beta_2 & 0 & \beta_4 \\ 0 & 0 & \gamma(t) & 0 \\ 0 & \nu_2 & 0 & \nu_4 \end{pmatrix} \begin{pmatrix} X_{t-1}^{(1)} \\ X_{t-1}^{(2)} \\ X_{t-1}^{(3)} \\ X_{t-1}^{(4)} \end{pmatrix} + \underline{\varepsilon}_t = A(t)\underline{X}_{t-1} + \underline{\varepsilon}_t,$$



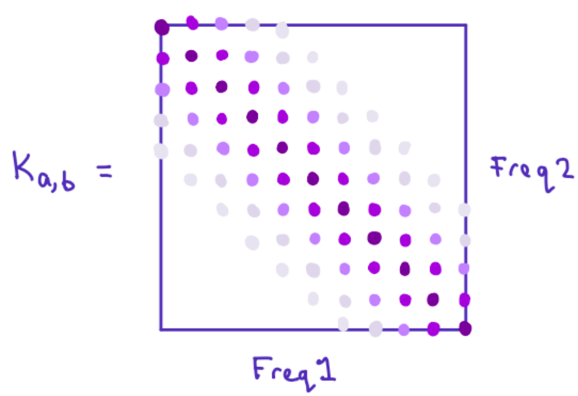
- Aim Understand the structure of $K_{a,b}$ when conditionally nonstationary.

Structure of $K_{a,b}$ under local stationarity

- We focus on time series whose covariance structure changes smoothly (or piecewise smoothly) over time.

- Dahlhaus (1997) developed the asymptotic machinery for studying such processes, which he showed were "locally stationary".

- If $\{X_t\}$ is locally stationary and (a, b) is conditionally nonstationary/time-varying, then:



  - $K_{a,b}$ is non-zero off the diagonal.

  - Sub-diagonals *close* to the diagonal are large and taper off further from the diagonal: $K_{a,b}(\omega_k, \omega_{k+r}) = O(|r|^{-1})$.

Relationship between linear regression and K_n

- In classical GGM (or StGM) we often estimate the entries of the precision matrix using linear regression.

- Similarly we can estimate the entries of $K_n = (\mathrm{var}[\mathbf{J}_n])^{-1} = F_n^* D_n F_n$ by projecting/regressing one DFT on all the other DFTs:

$$P_{\mathcal{G}_n - J_k^{(a)}}(J_k^{(a)}) = \sum_{b=1}^{p} \sum_{r \neq 0} \beta_{(b,k+r) \succ (a,k)} J_{k+r}^{(b)},$$

where $\mathcal{G}_n - J_k^{(a)} = \overline{\mathrm{sp}}(J_s^{(c)}; 1 \leq c \leq p, 1 \leq s \leq n, (c,s) \neq (a,k))$.

- The coefficients related to $K_{a,b}$ through the relation: $\beta_{(b,k+r) \succ (a,k)} = -K_{a,b}(\omega_k, \omega_{k+r}) / K_{a,a}(\omega_k, \omega_k)$.

Motivation: Spectral precision matrix estimation

- Suppose that $\{X_t\}$ is a multivariate stationary time series with spectral density $\Sigma(\omega)$ and spectral precision matrix $\Gamma(\omega) = \Sigma(\omega)^{-1}$.

- Since $\Gamma(\omega)$ is smooth across frequencies, like spectral density estimation, localised least squares is often used to estimate the entries of the spectral precision matrix $\Gamma(\omega)$:

$$\widehat{\beta}_{a,k} = \arg\min_{\gamma} \sum_{\ell=-M}^{M} W\left(\frac{\ell}{M}\right) \left| J_{k+\ell}^{(a)} - \sum_{b \neq a}^{p} \gamma_{(b,k)\nsucc(a,k)} J_{k+\ell}^{(b)} \right|^2$$

- $\widehat{\beta}_{(b,k)\nsucc(a,k)}$ is an estimator of $-\Gamma_{a,b}(\omega_k)/\Gamma_{a,a}(\omega_k)$

Estimation of K_n

- Likewise, to estimate the entries of K_n we use localised least squares.

- However, each row of K_n has dimension np, which is infeasible to estimate without regularisation.

- Under local stationarity the coefficients $\beta_{(b,k+r)\rightsquigarrow(a,k)}$ decay as $|r|$ grows. Thus we can truncate the number of regressors to estimate in the localised regression

$$\widehat{\beta}_{a,k} = \arg\min_{\gamma} \sum_{\ell=-M}^{M} W\left(\frac{\ell}{M}\right) \left| J_{k+\ell}^{(a)} - \sum_{b=1}^{p} \sum_{|r|\leq\nu} \gamma_{(b,k+r)\rightsquigarrow(a,k)} J_{k+\ell+r}^{(b)} \right|^2$$

- $\widehat{\beta}_{(b,k+r)\rightarrow(a,k)}$ is an estimator of $-K_{a,b}(\omega_k, \omega_{k+r})/K_{a,a}(\omega_k, \omega_k)$.

- We show that under suitable conditions, a Gaussian approximation of

$$\left(\widehat{\beta}_{(b,k+r)\rightarrow(a,k)}; |r| \leq d\right)$$

  is possible where we allow d to grow with n.

- This result can be used to test the values in the row of $K_{a,b}$ and thus to test for particular features in the graph.

Challenges in inference and testing

- Truncation to ν induces a bias in the parameter estimators. Asymptotically we need to allow for the truncation ν to grow with the sample size.

- The asymptotic analysis requires a delicate and intricate study of the DFTs in the case of nonstationarity.

- The variance of the regression estimators also needs to be estimated in order to implement the test (to build the graph).

Conclusion

- The main aim in this work was to build a graph for nonstationary time series based on conditional relationships.

- The paper for the main part of this talk is based on (Basu and SSR, 2023, Annals of Statistics).

- The relationship between locally stationarity (in terms of the covariances) and its inverse covariance operator can be found in Krampe and SSR, 2023, Bernoulli.

- Estimation of the network: Krampe and SSR, 2023, Preprint.

- Future work: Estimation of network for high dimensional time series.