Time series in frequency domain
Singular spectrum analysis
Entropy methods

Time Series Analysis

4. Model free methods

Andrew Lesniewski

Baruch College
New York

Fall 2017

Time series in frequency domain
Singular spectrum analysis
Entropy methods

Outline

1 Time series in frequency domain

2 Singular spectrum analysis

3 Entropy methods

Time series in frequency domain
Singular spectrum analysis
Entropy methods

Time series in frequency domain

- So far, we have discussed various models within the parametric approach to time series analysis. The key element of this approach is to specify a time series model with a number of free parameters which are determined via estimation from a data set.

- While this approach will remain the focus of these lectures, we will now take a brief side trip into the non-parametric (or model free) approach to time series analysis. In particular, we will focus of analyzing time series by means of expansion in various basis functions.

- The first approach that we discuss, namely *time series analysis in frequency domain* (in contrast to the *time domain* approach taken so far), is reminiscent of Fourier transform approach in signal processing. The idea is to decompose the underlying time series into components, each of which corresponds to evolution *cycles* of different frequencies.

- The appropriate basis functions are the trigonometric functions $\cos(\omega t)$ and $\sin(\omega t)$ or, equivalently, the complex exponential function $e^{i\omega t}$.

Time series in frequency domain
Singular spectrum analysis
Entropy methods

Spectral density function

- Let X_t be a covariance stationary time series, such that

Xt: univariate

$$\sum_{t=-\infty}^{\infty} |\Gamma_t| < \infty. \tag{1}$$

- The *spectral density function* (or *population spectrum*) of X_t is defined as

$$s_X(\omega) = \frac{1}{2\pi} \sum_{t=-\infty}^{\infty} \Gamma_t e^{-i\omega t}. \tag{2}$$

It is essentially the Fourier transform of X_t.

- From the trigonometric representation of complex numbers, and the fact that $\Gamma_{-t} = \Gamma_t$, we can write this in terms of purely real valued quantities:

$$s_X(\omega) = \frac{1}{2\pi} \left(\Gamma(0) + 2 \sum_{t=1}^{\infty} \Gamma_t \cos(\omega t) \right). \tag{3}$$

Time series in frequency domain
Singular spectrum analysis
Entropy methods

Spectral density function for white noise

- The easiest example is that of a white noise, $X_t = \varepsilon_t$. In this case,

$$\Gamma_t = \begin{cases} \sigma^2, & \text{if } t = 0, \\ 0, & \text{otherwise.} \end{cases}$$

- As a consequence, the SDF is constant,

$$s_X(\omega) = \frac{\sigma^2}{2\pi}. \tag{4}$$

Time series in frequency domain
Singular spectrum analysis
Entropy methods

Spectral density function for $AR(1)$

● As a next example, let us determine the spectral density function of the $AR(1)$ process. From equation (14) in Lecture Notes #1,
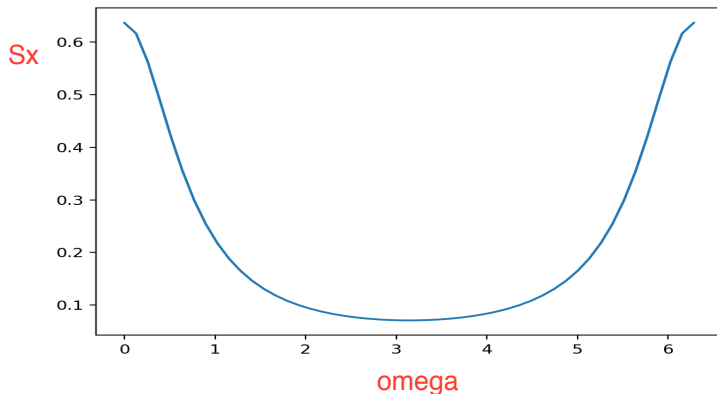
$$
\begin{aligned}
s_X(\omega) &= \frac{1}{2\pi} \sum_{t=-\infty}^{\infty} \Gamma_0 \beta^{|t|} e^{-i\omega t} \\
&= \frac{\Gamma_0}{2\pi} \left(1 + \sum_{t=1}^{\infty} \beta^t e^{i\omega t} + \sum_{t=1}^{\infty} \beta^t e^{-i\omega t} \right) \\
&= \frac{\Gamma_0}{2\pi} \left(1 + \frac{\beta e^{i\omega}}{1 - \beta e^{i\omega}} + \frac{\beta e^{-i\omega}}{1 - \beta e^{-i\omega}} \right).
\end{aligned}
$$

● As a result,

$$
s_X(\omega) = \frac{\sigma^2}{2\pi} \frac{1}{1 - 2\beta \cos \omega + \beta^2} \, . \tag{5}
$$

A. Lesniewski   **Time Series Analysis**

Time series in frequency domain
Singular spectrum analysis
Entropy methods

Spectral density function for $AR(1)$

- Below is the plot of (5) with $\beta = 0.5$ and $\sigma = 1$.

Time series in frequency domain
Singular spectrum analysis
Entropy methods

Spectral density function for *MA*(1)

- Let us now consider an *MA*(1) model. Using equation (62) in Lecture Notes #1, we see that

$$s_X(\omega) = \frac{1}{2\pi} \left((1 + \theta^2)\sigma^2 + \theta\sigma^2 e^{i\omega} + \theta\sigma^2 e^{-i\omega}\right).$$
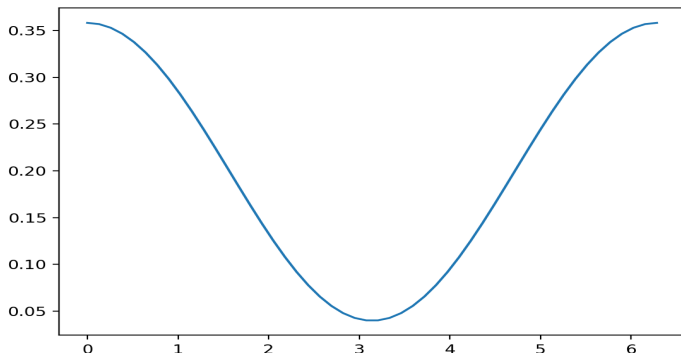
- This implies that the spectral density function of an *MA*(1) process is

$$s_X(\omega) = \frac{\sigma^2}{2\pi} \left(1 + 2\theta\cos\omega + \theta^2\right). \tag{6}$$

Time series in frequency domain
Singular spectrum analysis
Entropy methods

Spectral density function for *MA*(1)

● Below is the plot of (6) with $\theta = 0.5$ and $\sigma = 1$.

Time series in frequency domain
Singular spectrum analysis
Entropy methods

Spectral density for $ARMA(p, q)$

- The calculations above can be generalized to produce an expression for the $ARMA(p, q)$ model:

$$\psi(L)X_t = \alpha + \varphi(L)\varepsilon_t, \tag{7}$$

where our notation follows Lecture Notes #1.

- Namely, as you will establish in Homework Assignment #4, the spectral density is then given by

$$s_X(\omega) = \frac{\sigma^2}{2\pi} \left| \frac{\varphi(e^{i\omega})}{\psi(e^{i\omega})} \right|^2. \tag{8}$$

- If we factorize the polynomials $\psi(z)$ and $\varphi(z)$,

$$\psi(z) = (1 - \lambda_1 z) \ldots (1 - \lambda_p z),$$
$$\varphi(z) = (1 - \mu_1 z) \ldots (1 - \mu_q z),$$

then

$$s_X(\omega) = \frac{\sigma^2}{2\pi} \frac{(1 - 2\mu_1 \cos\omega + \mu_1^2) \ldots (1 - 2\mu_q \cos\omega + \mu_q^2)}{1 - 2\lambda_1 \cos\omega + \lambda_1^2) \ldots (1 - 2\lambda_p \cos\omega + \lambda_p^2)}. \tag{9}$$

Time series in frequency domain
Singular spectrum analysis
Entropy methods

Spectral density function

- In general, the spectral density function $s_X(\omega)$ has the following properties:
    - (i) It is non-negative.
    - (ii) It is a periodic function of ω with period 2π (assuming $h = 1$).
    - (iii) It is continuous in ω.
- The autocovariance can be calculated from the population spectrum by means of

$$\Gamma_t = \int_{-\pi}^{\pi} s_X(\omega) e^{i\omega t} d\omega. \tag{10}$$

- This is an immediate consequence of the fact that

$$\int_{-\pi}^{\pi} e^{i\omega(t-s)} d\omega = \begin{cases} 2\pi, & \text{if } t = s' \\ 0, & \text{otherwise.} \end{cases} \tag{11}$$

- Alternatively,

$$\Gamma_t = \int_{-\pi}^{\pi} s_X(\omega) \cos(\omega t) d\omega. \tag{12}$$

Time series in frequency domain
Singular spectrum analysis
Entropy methods

Spectral density function

- In particular,

$$\Gamma_0 = \int_{-\pi}^{\pi} s_X(\omega) d\omega, \tag{13}$$

i.e. the variance of X_t is equal to the area under the population spectrum between $-\pi$ and π.
- This also leads to the interpretation of $s_X(\omega)$ as the fraction of the variance that is attributable to cycles of frequency ω.

Time series in frequency domain
Singular spectrum analysis
Entropy methods

Spectral representation theorem

- There is a general result that states that any covariance-stationary time series process can be expressed in terms of its spectral data.
- Namely, there exists a unique complex valued stochastic function $z_X(\omega)$, such that

$$X_t = \mu + \int_{-\pi}^{\pi} e^{i\omega t} z_X(\omega) d\omega, \tag{14}$$

  where $\mu = E(X_t)$.
- Since X_t is real valued, the random function $z_X(\omega)$ must have the following symmetry property: bar: complex contribute

$$\overline{z_X(\omega)} = z_X(-\omega). \tag{15}$$

- Furthermore, $z_X(\omega)$ has the following properties:
  - (i) For all ω,

$$E(z_X(\omega)) = 0. \tag{16}$$

  - (ii) For all ω, ω',

$$E(z_X(\omega)\overline{z_X(\omega')}) = s_X(\omega)\delta(\omega - \omega'), \tag{17}$$

    where $\delta(\omega - \omega')$ denotes Dirac's delta function.
- This result is known as the *spectral representation theorem* or *Cramer's theorem*.

Time series in frequency domain
Singular spectrum analysis
Entropy methods

Spectral representation theorem

- The spectral representation theorem can also be written in terms of real quantities only. Namely, we define

$$a_X(\omega) = \operatorname{Re} z_X(\omega), \qquad \text{Re: real part}$$
$$b_X(\omega) = -\operatorname{Im} z_X(\omega) \qquad \text{IM: imagine part} \tag{18}$$

(the negative sign is just for convenience).

- Note that the random functions $a_X(\omega)$ and $b_X(\omega)$ have the following properties:
  (i)

$$a_X(-\omega) = a_X(\omega),$$
$$b_X(-\omega) = -b_X(\omega). \tag{19}$$

This is simply a conseqence of (15).

  (ii)

$$a_X(\omega)^2 + b_X(\omega)^2 = |z_X(\omega)|^2. \tag{20}$$

- As a result, we can write

$$X_t = \mu + \int_{-\pi}^{\pi} \left(cos(\omega t) a_X(\omega) + sin(\omega t) a_X(\omega) \right) d\omega. \tag{21}$$

Time series in frequency domain
Singular spectrum analysis
Entropy methods

Sample periodogram

- A complete proof of the spectral representation theorem is a bit technical, and can be found in specialized mathematical literature. Instead, we will interpret it in terms sample data.

- Let $x_1, \ldots, x_T$ be observations of X_t, and let $\widehat{\Gamma}_t$ denote the estimated autocovariance as defined by equation (5) in Lecture Notes #1. For any ω, the estimated sample spectral density function,

$$\widehat{s}_X(\omega) = \frac{1}{2\pi} \sum_{t=-(T-1)}^{T-1} \widehat{\Gamma}_t e^{-i\omega t}. \tag{22}$$

  is called the *sample periodogram*.

- We can then verify that

$$\widehat{\Gamma}_0 = \int_{-\pi}^{\pi} \widehat{s}_X(\omega) d\omega, \tag{23}$$

  i.e. the area under the periodogram is equal to the sample variance.

Time series in frequency domain
Singular spectrum analysis
Entropy methods

Sample periodogram

- In order to formulate the sample version of the spectral representation theorem, we assume that T is odd, and denote $\omega_j = 2\pi j / T$, for $j = -M, -M+1, \ldots, M$, where $M = (T-1)/2$.

- For each j, we define   If T is even, change T to T+1

$$\widehat{z}_X(\omega_j) = \frac{1}{T} \sum_{t=1}^{T} e^{-i\omega_j t} x_t - \widehat{\mu}. \tag{24}$$

Notice that

$$\widehat{z}_X(\omega_0) = 0. \tag{25}$$

- Then

$$x_t = \widehat{\mu} + \sum_{j=-M}^{M} e^{i\omega_j t} \widehat{z}_X(\omega_j). \tag{26}$$

Time series in frequency domain
Singular spectrum analysis
Entropy methods

Sample periodogram

- To see this, we multiply both sides of (24) by $e^{i\omega_j s}$ and sum over $j = 1, \dots, M$, and notice that

$$\sum_{j=-M}^{M} e^{i\omega_j(s-t)} = \begin{cases} T, & \text{if } s = t, \\ 0, & \text{otherwise.} \end{cases}$$

- Finally, notice that

$$\sum_{j=1}^{T}(x_t - \widehat{\mu})^2 = \sum_{j=-M}^{M} |\widehat{z}_X(\omega_j)|^2. \tag{27}$$

Time series in frequency domain
Singular spectrum analysis
Entropy methods

Singular spectrum analysis

- *Singular spectrum analysis* (SSA) is a model free feature extraction methodology, which may be thought of as a variant of the principal component analysis (PCA).
- Its extension to multivariate time series (not discussed here) is referred to as *multi channel singular spectrum analysis* (M-SSA).
- We consider a sample from a time series $X_1, \ldots, X_T$, and let $1 < l < T$ be the length of the rolling window. Then $k = T - l + 1$ is the number of lagged vectors.
- The basic algorithm of SSA consists of two stages:
    - (i) embedding,
    - (ii) reconstruction.

Time series in frequency domain
Singular spectrum analysis
Entropy methods

Singular spectrum analysis

- Embedding is carried out in two steps. First, we form the *trajectory matrix*:

$$\mathcal{X} = \begin{pmatrix} X_1 & X_2 & \dots & X_k \\ X_2 & X_3 & \dots & X_{k+1} \\ \vdots & \vdots & \dots & \vdots \\ X_l & X_{l+1} & \dots & X_T \end{pmatrix}. \tag{28}$$

  Note that $\mathcal{X}_{ij} = X_{i+j-1}$; matrices of this form are called *Hankel matrices*.
- The columns in the trajectory matrix correspond to the observations of the time series as the length l observation window slides forward.

Time series in frequency domain
Singular spectrum analysis
Entropy methods

Singular spectrum analysis

- Then, we perform the singular value decomposition (SVD) of the trajectory matrix $\mathcal{X}$:

  (i) Let $\mathcal{S} = \mathcal{X}\mathcal{X}^{\mathrm{T}}$. Then $\mathcal{S}$ is positive definite; we denote its eigenvalues by $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_l \geq 0$, and the corresponding orthonormal system of eigenvectors by $U_1, U_2, \ldots, U_l$. The numbers $\sqrt{\lambda_i}$ are called the *singular values* of $\mathcal{X}$.

  (ii) Let $r = \mathrm{rank}(\mathcal{X})$ (typically, $r = L$), and set $V_i = \frac{1}{\sqrt{\lambda_i}} \mathcal{X}^{\mathrm{T}} U_i$, for $i = 1, \ldots, l$.

  (iii) Then

  $$\mathcal{X} = \mathcal{X}_1 + \mathcal{X}_2 + \ldots + \mathcal{X}_r, \tag{29}$$

  where $\mathcal{X}_i = \sqrt{\lambda_i} U_i V_i^{\mathrm{T}}$ are rank 1 matrices, called *elementary matrices*. The triple $(\sqrt{\lambda_i}, U_i, V_i)$ is called an *eigentriple* (ET) of the SVD and the vectors $\sqrt{\lambda_i}\mathcal{X}_i = U_i V_i^{\mathrm{T}}$ are the *principal components*.

  (iv) The numpy implementation of SVD is called `numpy.linalg.svd`.

Time series in frequency domain
Singular spectrum analysis
Entropy methods

Singular spectrum analysis

- The reconstruction stage is performed in two steps. First, we partition the set of indices $I = \{1, \ldots, r\}$ into m disjoint subsets $I = I_1 \cup \ldots \cup I_m$. For each subset I_k, form the sum

$$\mathcal{X}_{I_k} = \sum_{i \in I_k} \mathcal{X}_i. \tag{30}$$

Clearly, this defines a decomposition of the trajectory matrix into components:

$$\mathcal{X} = \mathcal{X}_{I_1} + \ldots + \mathcal{X}_{I_m}. \tag{31}$$

- The final step is *diagonal averaging*. Each matrix $\mathcal{X}_{I_k}$ in the decomposition (31) is transformed into a new *reconstructed time series* $(\widetilde{X}_1^{(k)}, \widetilde{X}_2^{(k)}, \ldots, \widetilde{X}_T^{(k)})$ by means of the following procedure.

Time series in frequency domain
Singular spectrum analysis
Entropy methods

Singular spectrum analysis

● Let A be an $l \times k$-matrix, and let $T = l + k - 1$. We denote

$$A_{ij}^* = \begin{cases} A_{ij}, \text{ if } l < k, \\ A_{ji}, \text{ otherwise.} \end{cases} \tag{32}$$

Diagonal averaging transforms the matrix A into a time series $\widetilde{A}_1, \ldots, \widetilde{A}_T$ as follows:

$$A_j = \begin{cases} \frac{1}{j} \sum_{m=1}^{k} A_{m,j-m+1}^*, & \text{for } 1 \leq j < l \wedge k, \\ \frac{1}{l \wedge k} \sum_{m=1}^{l \wedge k} A_{m,j-m+1}^*, & \text{for } l \wedge k \leq j \leq l \vee k, \\ \frac{1}{N-j+1} \sum_{m=k-l \vee k+1}^{T-l \vee k+1} A_{m,j-m+1}^*, & \text{for } l \vee k \leq j \leq T. \end{cases} \tag{33}$$

● As a result, the original time series is represented as a sun of m reconstructed series;

$$X_t = \sum_{i=1}^{m} \widetilde{X}_t^{(i)}. \tag{34}$$

Time series in frequency domain
Singular spectrum analysis
Entropy methods

Singular spectrum analysis

- The choice of the rolling window length l is an important matter. It should be suffiently large so that each lagged time series incorporates the essential features of the original series $X_1, \ldots, X_N$.
- It is a good idea to perform SSA with different choices of l.

Time series in frequency domain
Singular spectrum analysis
Entropy methods

SSA of a simulated $I(1)$ process

- The figure below shows the results of SSA of the simulated $I(1)$ process given by the following specification:

$$X_t = 1.1 + X_{t-1} + 5.0\varepsilon_t, \tag{35}$$

where $\varepsilon_t \sim N(0, 1)$.

- The upper left plot shows the actual time series, while the remaining ones show the first five SSA components.
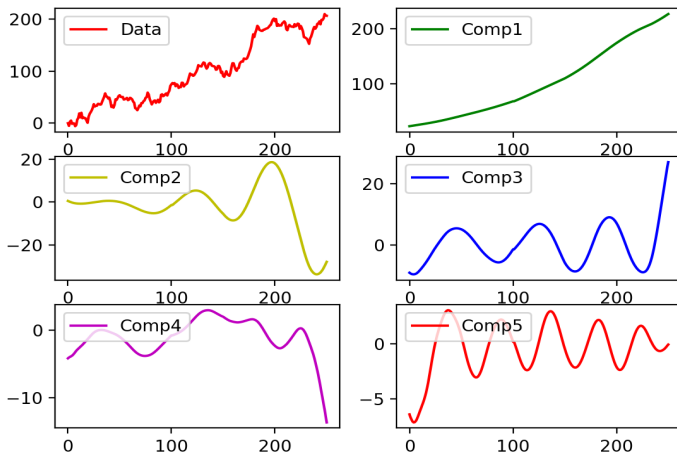- The cumulative weights, defined as

$$CW_j = \frac{\lambda_1 + \ldots + \lambda_j}{\lambda_1 + \ldots + \lambda_l}, \tag{36}$$

of the plotted components are:

$$
\begin{aligned}
CW_1 &= 0.595, \\
CW_2 &= 0.653, \\
CW_3 &= 0.698, \\
CW_4 &= 0.720, \\
CW_5 &= 0.737.
\end{aligned} \tag{37}
$$

Time series in frequency domain
Singular spectrum analysis
Entropy methods

SSA of a simulated $AR(1)$ process

Time series in frequency domain
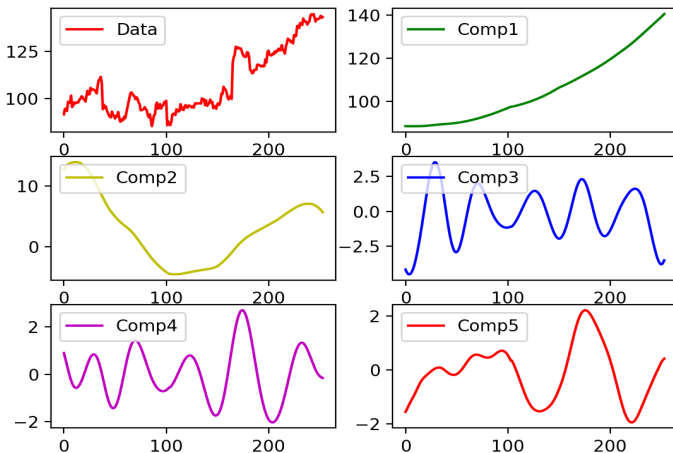**Singular spectrum analysis**
Entropy methods

SSA of the Netflix share price

- The next figure below shows the results of SSA of the share price of Netflix (NFLX) during the one-year period 02-24-2016 through 02-24-2017.
- As before, the upper left plot shows the actual time series, while the remaining ones show the first five SSA components.
- The cumulative weights of the plotted components are:

$$
\begin{aligned}
CW_1 &= 0.725, \\
CW_2 &= 0.763, \\
CW_3 &= 0.778, \\
CW_4 &= 0.793, \\
CW_5 &= 0.805.
\end{aligned}
\tag{38}
$$

- Notice that the first component (trend) is responsible for 72.5% of the dynamics.

Time series in frequency domain
Singular spectrum analysis
Entropy methods

SSA of the Netflix share price

Time series in frequency domain
Singular spectrum analysis
Entropy methods

Transfer entropy

- The concept of Granger causality defined earlier in these lectures can be reformulated in a model free manner, using the concept of *transfer entropy*.
- The price for the model freeness is a bit of formalism required. In order to lighten up on the math, we will assume that X_t can take on only one of finitely many state values in $A = \{x^1, \ldots, x^K\}$.
- The *Shannon entropy* of the probability distribution $p_j = P(X_t = x^j)$ is given by:

  Shannon entropy from information theory

  Physics: Gibbs entropy

  $$H(X_t) = -\sum_{j=1}^{K} p_j \log p_j. \tag{39}$$

- The Shannon entropy is always nonnegative. Its value is 0, if one of the p_j's is 1. It reaches is maximum value $\log K$, if the distribution is uniform, $p_j = 1/K$, for all $j = 1, \ldots, K$.
- It is interpreted as a measure of information in the probability distribution: the lower the entropy, the higher its information content.

Time series in frequency domain
Singular spectrum analysis
Entropy methods

Transfer entropy

- Suppose $q_j, j = 1, \ldots, K$ is another probability distribution (possibly an *a priori* guess of p). A useful measure of distance between p and q is given by the *Kullback-Leibler divergence*, a.k.a. *relative entropy*:

  Practice: p observed from real data

  q: model parameters

$$D_{p \| q}(X_t) = \sum_{j=1}^{K} p_j \log \frac{p_j}{q_j} . \tag{40}$$

This D is not symmetry by construction

  - One can show that $D_{p \| q}(X_t) \geq 0$, and $D_{p \| q}(X_t) = 0$, iff $p = q$.

  - Consider now a second process Y_t with state values in $B = \{y^1, \ldots, y^K\}$. We define the *joint entropy*:

$$H(X_t, Y_s) = - \sum_{i,j=1}^{K} p_{i,j} \log p_{i,j}, \tag{41}$$

  where $p_{i,j} = P(X_t = x^i, Y_s = y^j)$ is the joint probability distribution.

Time series in frequency domain
Singular spectrum analysis
Entropy methods

Transfer entropy

- The mutual entropy of X_t and Y_t is defined by

$$M(X_t, Y_s) = H(X_t) + H(Y_s) - H(X_t, Y_s)$$
$$= \sum_{i,j=1}^{K} P(X_t = x^i, Y_s = y^j) \log \frac{P(X_t = x^i, Y_s = y^j)}{P(X_t = x^i) P(Y_s = y^j)} . \quad (42)$$

- The mutual entropy is thus identital with the Kullback-Leibler distance between the joint distribution of X_t, Y_s and the product of the marginal distributions, $q_{i,j} = P(X_t = x^i) P(Y_s = y^j)$.

Time series in frequency domain
Singular spectrum analysis
Entropy methods

Entropy rate of a process

- The dynamic character of a time series is captured by the transisition probabilities $P(X_{t+1} = x^j | x_{t-l+1:t})$. The associated Shannon entropy is given by

  t-l+1:t   l: lag

$$H(X_{t+1} | x_{t-l+1:t}) = -\sum_{j=1}^{K} P(X_{t+1} = x^j | x_{t-l+1:t}) \log P(X_{t+1} = x^j | x_{t-l+1:t}). \quad (43)$$

- The entropy concept below (which is related to Shannon's *entropy rate*) measures the information content in these transition probabilities.

- Namely, the entropy rate of a time series is defined as the expected value of $H(X_{t+1} | x_{t-l+1:t})$ with respect to all histories $x_{t-l+1:t}$:

$$\begin{aligned}
H(X_{t+1} | X_{t-l+1:t}) &= \sum_{x_{t-l+1:t}} p(x_{t-l+1:t}) H(X_{t+1} | x_{t-l+1:t}) \\
&= -\sum_{x_{t-l+1:t+1}} p(x_{t-l+1:t+1}) \log p(x_{t+1} | x_{t-l+1:t}).
\end{aligned} \quad (44)$$

Time series in frequency domain
Singular spectrum analysis
Entropy methods

Entropy rate of a process

- If q is another probability distribution, we can define the following Kullback-Leibler divergence:

$$D_{p\|q}(X_{t+1}|X_{t-l+1:t}) = \sum_{x_{t-l+1:t+1}} p(x_{t-l+1:t+1}) \log \frac{p(x_{t+1}|x_{t-l+1:t})}{q(x_{t+1}|x_{t-l+1:t})} . \qquad (45)$$

- The same concepts can be extended to multivariate time series, with the corresponding increase in the notational complexity.

- In particular, if X_t, Y_t is a bivariate time series, the Kullback-Leibler divergence with respect to

$$q(x_{t+1}, y_{t+1}|x_{t-l+1:t}, y_{t-l+1:t}) = p(x_{t+1}|x_{t-l+1:t})p(y_{t+1}|y_{t-l+1:t})$$

is the corresponding mutual entropy.

Time series in frequency domain
Singular spectrum analysis
Entropy methods

Entropy rate of a process

- Explicitly,

$$M(X_{t+1}, Y_{t+1} | X_{t-l+1:t}, Y_{t-l+1:t}|)$$
$$= \sum_{x_{t-l+1:t+1}} \sum_{y_{t-l+1:t+1}} p(x_{t-l+1:t+1}, y_{t-l+1:t+1}) \log \frac{p(x_{t+1}, y_{t+1} | x_{t-l+1:t}, y_{t-l+1:t})}{p(x_{t+1} | x_{t-l+1:t}) p(x_{t+1} | x_{t-l+1:t})} .$$

Symmetric

Symmetric

Switch X and Y does not change result

- While this concept is an approprite information measure for the bivariate time series X_t, Y_t, it has the property that it is symmetric in X_t and Y_t. As a consequence, it cannot be used for quantifing causal dependence of one series on the other.

- The causal dependence measure related the concepts introduced above is *transfer entropy*.

Time series in frequency domain
Singular spectrum analysis
Entropy methods

Transfer entropy

- If two processes X_t and Y_t are independent, then, for any number of lags j,

$$p(x_{t+1}|x_{t-j+1:t}) = p(x_{t+1}|x_{t-j+1:t}, y_{t-j+1:t}) \tag{46}$$

- This leads to the following concept of *transfer entropy*:

$$
\begin{aligned}
&T(X_{t+1}|X_{t-l+1:t}, Y_{t-l+1:t}|) \\
&= \sum_{x_{t-l+1:t+1}} \sum_{y_{t-l+1:t}} p(x_{t-j+1:t+1}, y_{t-j+1:t}) \log \frac{p(x_{t+1}|x_{t-j+1:t}, y_{t-j+1:t})}{p(x_{t+1}|x_{t-j+1:t})} .
\end{aligned} \tag{47}
$$

No symmetric any more

- This also can be rewritten as

$$
\begin{aligned}
&T(X_{t+1}|X_{t-l+1:t}, Y_{t-l+1:t}|) \\
&= H(X_{t+1}|X_{t-j+1:t}) - H(X_{t+1}|X_{t-j+1:t}, Y_{t-j+1:t}).
\end{aligned} \tag{48}
$$

Time series in frequency domain
Singular spectrum analysis
Entropy methods

Transfer entropy

- Transfer entropy is a very elegant and economic concept of causal dependence among time series.
- It applies to time series models that are not necessarily linear, or whose disturnaces are necessarily normally distributed.
- In case of autoregressive models with normally distributed disturbances, transfer entropy is essentially identical with the statistics used to test Granger causality.
- Estimation of transfer entropy from observed data is a bit of a challenge, as reliable estimates require large sample sets.

Time series in frequency domain
Singular spectrum analysis
Entropy methods

References

📑 Golyandina, N., and Zhiglyavsky, A.: *Singular Spectrum Analysis for Time Series*, Springer (2013).

📑 Hamilton, J. D.: *Time Series Analysis*, Princeton University Press (1994).

📑 Schreiber, T.: Measuring Information transfer, *Phys. Rev. Lett.*, **85** ,461 - 464 (2000).