Measuring Interactions from Time Series

Information-theory-based techniques
measuring interactions

**basic idea:** interaction $\iff$ information flow

- the "stronger" the information flow, the stronger the interaction
- information flow from one system to another indexes directionality

characterize information with Shannon entropy

$$H = - \sum_i p_i \log p_i$$

$p$ is the (normalized) probability for an event / state / amplitude /… to occur

estimate probability with

$$p_i = \lim_{N \to \infty} \frac{N_i}{N}$$

where $N$ is the total number of events / states / amplitudes /…
measuring interactions

strength of interaction:

given systems $X$ and $Y$, the mutual information is defined as:

$$ I(X, Y) = H(X) + H(Y) - H(X, Y) $$

information generated by system $X$ is characterized by the Shannon entropy:

$$ H(X) = -\sum_i^N p_X(i) \log p_X(i) $$

joint information is characterized by the Shannon entropy:

$$ H(X, Y) = -\sum_{i,j}^N p_{X,Y}(i, j) \log p_{X,Y}(i, j) $$
measuring interactions

strength of interaction:

relative entropy (also known as Kullback-Leibler divergence):

$$H_{\text{rel}}(X|Y) = - \sum_i^N p_X(i) \log \frac{p_X(i)}{p_Y(i)}$$

- characterizes the similarity between the probability distributions.
- relative entropy is asymmetric: $H_{\text{rel}}(X|Y) \neq H_{\text{rel}}(Y|X)$
- in general $H_{\text{rel}}$ is positive, and zero for identical systems

→ alternative definition of mutual information:

$$I_{\text{rel}}(X|Y) = - \sum_{i,j}^N p_{X,Y}(i,j) \log \frac{p_{X,Y}(i,j)}{p_X(i)p_Y(j)}$$

characterizes relative difference between respective probability density distributions and the joint distribution density

measuring interactions

**strength of interaction:**

properties of mutual information:
- symmetric: $I(X, Y) = I(Y, X)$
- $I(X, Y) = 0$ for independent (non-interacting) systems
- $I(X, Y) = \max$ for identical (fully synchronized) systems
- $I(X, Y)$ increases monotonically with increasing coupling strength $\rightarrow$ data-driven estimator for strength of interaction

disadvantages:
- only considers (single/joint) probability density distributions
- no information about dynamics
- can not explicitly distinguish between information exchange and joint information (e.g. due to common input or joint past)
measuring interactions

**strength of interaction:**

extensions:

- time-delayed mutual information
  (Kaneko, Physica D 23, 436, 1986)

- partial (or conditional) mutual information
  (S. Frenzel & B. Pompe, PRL 99, 204101, 2007)
  - part of mutual information of two random quantities that is not contained in a third one
  - similar to partial correlation
  - can also detect directionality*

$$I(X, Y|Z) = H(X, Z) + H(Y, Z) - H(Z) - H(X, Y, Z)$$

Fundamentals of Analyzing Biomedical Signals

**Interactions**

measuring interactions

**direction of interaction:**

aim: characterize flow of information between systems $X$ and $Y$

idea: replace (static) probability density distributions by transition probability densities (cf. Entropies)

given: time series $\mathbf{v}$: $v_1, v_2, \ldots, v_N$ of some observable $x$ and time series $\mathbf{w}$: $w_1, w_2, \ldots, w_N$ of some observable $y$

1) incorporate time-dependence by relating previous samples $v_i$ and $w_i$ to predict the next value $v_{i+1}$ (cf. N. Wiener),

2) consider generalized Markov condition ($p =$ transition probability density):

$$p(v_{i+1} \mid \mathbf{v}_i, \mathbf{w}_i) = p(v_{i+1} \mid \mathbf{v}_i)$$

measuring interactions

direction of interaction:

3) if systems $X$ and $Y$ independent $\rightarrow$ Markov condition fulfilled

4) use relative entropy concept to quantify incorrectness of Markov condition; with this, transfer entropy is defined as:

$$ T_{Y \rightarrow X} = \sum_{i}^{N} p\left( v_{i+1}, v_{i}^{(k)}, w_{i}^{(l)} \right) \log \frac{ p\left( v_{i+1} | v_{i}^{(k)}, w_{i}^{(l)} \right) }{ p\left( v_{i+1} | v_{i}^{(k)} \right) } $$

$(l,k)$ denote orders of Markov processes
$T_{X \rightarrow Y}$ defined in complete analogy
measuring interactions

**direction of interaction:**

**information-theory**

**transfer entropy**

**properties of transfer entropy**

- can detect direction of information flow since $T_{Y \rightarrow X} \neq T_{X \rightarrow Y}$
- unbounded, needs suitable definition of directionality, e.g.

$$T := T_{Y \rightarrow X} - T_{X \rightarrow Y}$$

\[
\begin{cases}
  > 0 : \text{Y drives X} \\
  = 0 : \text{no or symmetric bidir. coupling} \\
  < 0 : \text{X drives Y}
\end{cases}
\]

- depends on coupling strength $\rightarrow$ data-driven estimator for direction of interaction
- for Gaussian distributed data, transfer entropy equals Granger causality
- similar to conditional mutual information
  (replace system $Z$ by e.g., past of system $Y$)
measuring interactions

direction of interaction:

information-theory

transfer entropy

extensions:

- multivariate (partial) transfer entropy

- various estimation techniques

- estimators for transient signals* and delay-systems**


measuring interactions

strength and direction of interaction

how to estimate probability density distributions and the joint distribution densities from time series?

- counting (cumbersome)
- various binning techniques
- nearest neighbor estimators (e.g. Kozachenko-Leonenko)
- correlation sum (via phase-space embeddings)
- symbolization (e.g. based or permutation entropy*)

measuring interactions

**strength and direction of interaction**

estimators based on the concept of symbolic dynamics and on symbolization

**symbolic dynamics:** modeling a smooth dynamical system by a *discrete space* consisting of infinite *sequences of symbols*, each of which corresponds to a state of the system, with the dynamics (evolution) given by the shift operator

**symbolization:** generate symbols via delay embedding

\[ S_i := (v_i+(j_1-1)\tau, v_i+(j_2-1)\tau, \ldots, v_i+(j_m-1)\tau) \]

where

\[ v_i+(j_1-1)\tau \leq v_i+(j_2-1)\tau \leq \cdots \leq v_i+(j_m-1)\tau \]

→ symbol

\[ \hat{s}_i := (\hat{j}_1, \hat{j}_2, \ldots, \hat{j}_m) \]

measuring interactions

strength and direction of interaction

embedded data: (3, 5, 9) (10, 1, 6)

symbols ⇒ (1, 2, 3) (1, 3, 2)

permutation entropy: \[ H(m) = - \sum_{i=1}^{m!} \hat{s}_i \log \hat{s}_i \]

normalization: \[ 0 \leq H = \frac{H(m)}{\log(m!)} \leq 1 \]

\( H \to 0 \) for deterministic systems, \( H \to 1 \) for stochastic systems
measuring interactions

strength and direction of interaction

given time series of systems X and Y:
- estimate permutation entropy from windowed data
- investigate changing tendency of permutation entropies

\[
S(w_i) = \begin{cases} 
+1 : & \text{if } H(w_i) < H(w_{i+1}) \\
-1 : & \text{else}
\end{cases}
\]

- characterize in-step behavior of pairs of permutation entropies

\[
\gamma := \sum_{i=1}^{N_w} S_X(w_i) S_Y(w_i)
\]

- \( \gamma = 0 \) for independent systems; \( \gamma \rightarrow 1 \) for synchronized systems; \( \gamma \) increase monotonically with increasing coupling strength \( \rightarrow \) data-driven estimator for strength of interaction

measuring interactions

strength and direction of interaction

an example

given time series of systems X and Y:
- for estimating probability density distributions and the joint distribution densities:
  replace probabilities of data with probabilities of symbols count symbols (→ very fast)
- symbolic transfer entropy:
  \[ T_{Y \rightarrow X}^S = \sum_{\hat{v}_{i+1}, \hat{v}_i^{(k)}, \hat{w}_i^{(l)}} p\left(\hat{v}_{i+1}, \hat{v}_i^{(k)}, \hat{w}_i^{(l)}\right) \log \frac{p\left(\hat{v}_{i+1} | \hat{v}_i^{(k)}, \hat{w}_i^{(l)}\right)}{p\left(\hat{v}_{i+1} | \hat{v}_i^{(k)}\right)} \]
  
- see properties of transfer entropy
- easy-to-use data driven estimator for direction of interaction

measuring interactions

strength and direction of interaction

information-theory

an example

measuring interactions

strength and direction of interaction

an example

diffusively coupled Rössler oscillators (100 realizations)

\[
\begin{align*}
\dot{x}_{1,2} &= -\Omega_{1,2}y_{1,2} - z_{1,2} + c_{1,2}(x_{2,1} - x_{1,2}) \\
\dot{y}_{1,2} &= \Omega_{1,2}x_{1,2} + 0.165y_{1,2} \\
\dot{z}_{1,2} &= 0.2 + z_{1,2}(x_{1,2} - 10) \\
\Omega_{1,2} &\in \mathcal{N}(0.89; 0.1) \\
c_1 &= 0 \\
c_2 &= c
\end{align*}
\]

measuring interactions

**strength and direction of interaction**

**information-theory**

**an example**

diffusively coupled Rössler-Lorenz oscillators

(100 realizations of driver-responder system)

\[
\begin{align*}
\dot{x}^R &= -\Omega^R (y^R - z^R) \\
\dot{y}^R &= \Omega^R (x^R + 0.2y^R) \\
\dot{z}^R &= \Omega^R (0.2 + z^R (x^R - 5.7)) \\
\dot{x}^L &= 10(y^L - x^L) \\
\dot{y}^L &= \Omega^L x^L - y^L - x^L z^L + c(y^R)^2 \\
\dot{z}^L &= x^L y^L - \frac{8}{3} z^L \\
\Omega^R &\in \mathcal{N}(6, 0.1) \\
\Omega^L &\in \mathcal{N}(28, 1)
\end{align*}
\]

measuring interactions

**strength and direction of interaction**

an example

driver-responder relationships in EEG dynamics from epilepsy patients

Fundamentals of Analyzing Biomedical Signals

Interactions

measuring interactions

strength and direction of interaction

permutation-entropy-based estimators

advantages

- easy-to-use, fast-to-calculate
- high robustness against noise (symbolization)

disadvantages

- symbolization may lead to loss of information
- require appropriate choice of embedding parameter
- choice of window-size, finiteness of available symbols
- “faster” system (eigen-frequency, noise) → driver
  (need reliable surrogate test for directionality)
- may be fooled by (unobserved) third system