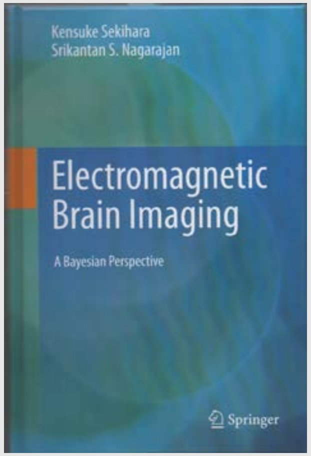
Causality Analysis
and
Cross-Frequency Analysis
in MEG Source Space

株式会社シグナルアナリシス
関原　謙介



Electromagnetic Brain Imagingの第8章・第9章
を解説.

脳の各部位間が関連して活動しているかどうかを調べる.

⇩

Connectivity Analysis

脳の各部位間の関連のみでなく，情報の流れ（どちらの活動が原因でもう片方の脳活動が引き起こされるのか）までも調べようとする方法.

⇩

Causality Analysis

Biomagなどの学会で研究発表はあるが，方法として確立しているものはない．（どころか，将来方法として確立するのかも不明）

とは言っても，研究発表などは結構あるので，MEG分野のengineerとして，どんな名前でよばれているどんな方法が有るのかについて概略は知っておく必要があるか？

異なる脳の部位が関連して活動し，どちら向きにどの程度の情報が流れているかを評価できる量を計算し，その値によって脳活動間の情報の流れを求める．

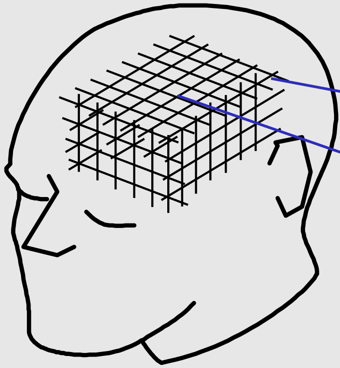
⇧

Causality measure

いくつかのcausality measureが提案されている．

Phase-slope index

直感的に解りやすいものとして: coherenceの位相の符号を調べる.



voxel time courses

voxel spectra

$s_k(t) \longrightarrow \sigma_k(f) = \int s_k(t) e^{-2\pi i f t} dt$

$s_j(t) \longrightarrow \sigma_j(f) = \int s_j(t) e^{-2\pi i f t} dt$

Cross spectrum: $\varphi_{j,k} = \left\langle \sigma_j(f)\sigma_k(f)^* \right\rangle = \left| \left\langle \sigma_j(f)\sigma_k(f)^* \right\rangle \right| \exp(2\pi i f \tau)$

位相項は周波数 f と時間差 τ に比例．ここで，τ は情報が第 j ボクセルから第 k ボクセルに伝わる時間である．

$$\text{Phase slope: } \frac{\partial}{\partial f} \text{phase}(\varphi_{j,k}) \sim \tau.$$

$$\text{Phase slope: } \frac{\partial}{\partial f}\text{phase}(\varphi_{j,k}) \sim \ \tau.$$

phase slopeの符号は時間差 τ の符号に等しく，情報が k から j へ流れるのか，あるいは， j から k へ流れるのかを示す

phase slopeは以下のphase slope index(PSI)から推定できる：

$$\Psi_{j,k} = \Im\left[\sum_{f\in F_w} \varphi_{j,k}(f)^* \varphi_{j,k}(f+\delta f)\right]$$

なぜならば，これに $\varphi_{j,k} = \left|\left\langle |\sigma_j(f)|^2 \right\rangle\right| \exp(2\pi i f\tau)$ を代入すれば，

$$\Psi_{j,k} = \Im\left[\sum_{f\in F_w} \left|\left\langle |\sigma_j(f)|^2 \right\rangle\right|^2 \exp(2\pi i \delta f \tau)\right] = \left[\sum_{f\in F_w} \left|\left\langle |\sigma_j(f)|^2 \right\rangle\right|^2\right] \Im\left[\exp(2\pi i \delta f\tau)\right]$$
$$= \left[\sum_{f\in F_w} \left|\left\langle |\sigma_j(f)|^2 \right\rangle\right|^2\right] \sin(2\pi\delta f\tau) \approx \left[\sum_{f\in F_w} \left|\left\langle |\sigma_j(f)|^2 \right\rangle\right|^2\right] 2\pi\delta f\tau$$

となり， $\Psi_{j,k} \propto \tau$ が得られる．ただし，$\Psi_{j,k}$ はまた，信号強度 $\left[\sum_{f\in F_w} \left|\left\langle |\sigma_j(f)|^2 \right\rangle\right|^2\right]$ にも依存する．

そこで，Nolteの論文ではPSIは以下のように定義されている.

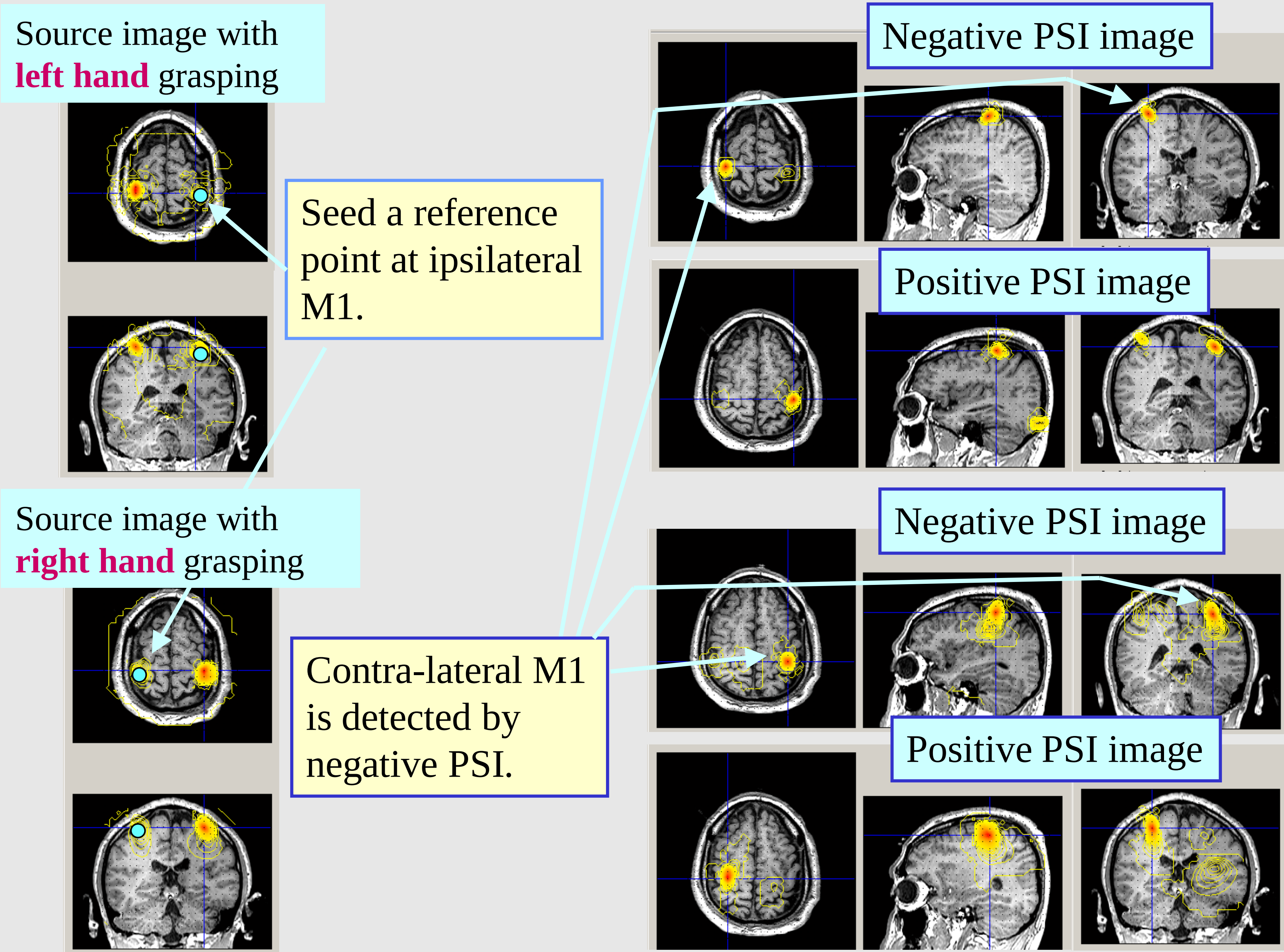
$$\Psi_{j,k} = \Im\left[\sum_{f\in F_w} \eta_{j,k}(f)^* \eta_{j,k}(f+\delta f)\right]$$

ここで，$\eta_{j,k}(f)$ はコヒーレンス： $\eta_{j,k}(f) = \left\langle \varphi_{j,k}(f) \right\rangle / \sqrt{\varphi_{j,j}(f)\varphi_{k,k}(f)}$
である.

この場合，$\Psi_{j,k}$ は $\left[\sum_{f\in F_w} \left| \left\langle | \sigma_j(f) |^2 \right\rangle \right|^2 \right]$ に依存しない量となる.

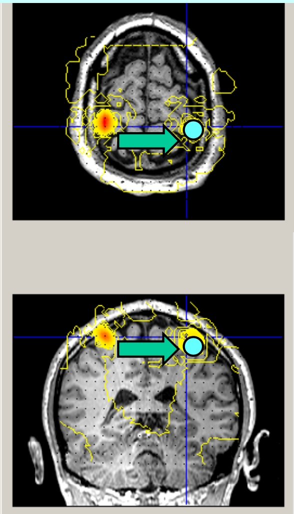
G. Nolte et al.: Phys. Rev. Lett. 100, pp.234101, 2008

Application of PSI to hand-motor MEG data



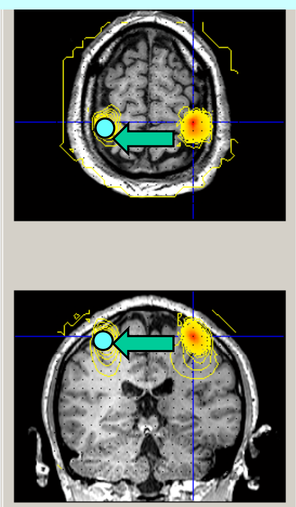
Application of PSI to hand-motor MEG data

Source image with **left hand** movement



- These results indicate that the contra-M1 is first activated and the ipsi-M1 is activated later.
- Information flows from the contra-M1 to ipsi-M1.
- Such results agree with our physiological knowledge, and demonstrate the effectiveness of the use of PSI.

Source image with **right hand** movement



PSIはこの例のように強い脳活動に対しては reasonableな結果を与える場合も有るが, 安静時脳活動のような場合で意味の有る結果が得られるかは未知(疑問)である.

MVAR-modeling-based methods

Granger-causality measureとその親類

Granger causality

- Causality measureとしては，Granger causalityが良く知られている.

- Granger causality はデータのmultivariate vector-autoregressive（MVAR）modelingをもとにして導かれる.

MVAR modeling

Q-channelのvoxel time seriesを考え，それらを以下で表す.

$$\boldsymbol{y}(t) = [y_1(t), \ldots, y_Q(t)]^T$$

MVAR modelingではこの $\boldsymbol{y}(t)$ が過去の値を用いて以下のように表されるとする.

$$\boldsymbol{y}(t) = \sum_{p=1}^{P} \boldsymbol{A}(p)\boldsymbol{y}(t-p) + \boldsymbol{e}(t)$$

ここで，$\boldsymbol{A}(p)$は$Q\times Q$の係数行列．$\boldsymbol{e}(t)$は残差を表すベクトル，Pはモデルオーダーと呼ばれる.

もっとも簡単なbivariate（チャンネルが2個）の場合で考える.

time series $y_1(t)$ と $y_2(t)$ について:

case#1:それぞれのtime seriesがおのおのの過去にしか依存しないとしてモデル化.

$$y_1(t) = \sum_{p=1}^{P} A_{1,1}(t-p)y_1(t-p) + e_1(t)$$
$$y_2(t) = \sum_{p=1}^{P} A_{2,2}(t-p)y_2(t-p) + e_2(t)$$

case#2:それぞれのtime seriesがお互いの過去にも依存するとしてモデル化.

$$y_1(t) = \sum_{p=1}^{P} A_{1,1}(t-p)y_1(t-p) + \sum_{p=1}^{P} A_{1,2}(t-p)y_2(t-p) + \varepsilon_1(t)$$
$$y_2(t) = \sum_{p=1}^{P} A_{2,1}(t-p)y_1(t-p) + \sum_{p=1}^{P} A_{2,2}(t-p)y_2(t-p) + \varepsilon_2(t)$$

残差の分散： $V(e_1(t)) = \Sigma_e^1,\quad V(e_2(t)) = \Sigma_e^2,\quad V(\varepsilon_1(t)) = \Sigma_\varepsilon^1,\quad V(\varepsilon_2(t)) = \Sigma_\varepsilon^2$

Granger causality

$G_{2\to1} = \log\left[\Sigma_e^1 / \Sigma_\varepsilon^1\right]$（$y_1(t)$の予測値が$y_2(t)$を考慮することでどの程度良くなったか）

$G_{1\to2} = \log\left[\Sigma_e^2 / \Sigma_\varepsilon^2\right]$（$y_2(t)$の予測値が$y_1(t)$を考慮することでどの程度良くなったか）

全く同じ考え方を用いてMultivariate Granger causalityを定義できる.

Multivariate time series $\boldsymbol{x}(t)$ と $\boldsymbol{y}(t)$ を定義:

$$\boldsymbol{x}(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_R(t) \end{bmatrix} \qquad \boldsymbol{y}(t) = \begin{bmatrix} y_1(t) \\ \vdots \\ y_Q(t) \end{bmatrix}$$

case#1:それぞれのグループのtime seriesがおのおののグループの過去にしか依存しない場合.

$$\boldsymbol{x}(t) = \sum_{p=1}^{P} \boldsymbol{A}_x(t-p)\boldsymbol{x}(t-p) + \boldsymbol{e}_x(t)$$
$$\boldsymbol{y}(t) = \sum_{p=1}^{P} \boldsymbol{A}_y(t-p)\boldsymbol{y}(t-p) + \boldsymbol{e}_y(t)$$

case#2:それぞれのグループのtime seriesがお互いの過去にも依存する場合.

$$\boldsymbol{x}(t) = \sum_{p=1}^{P} \boldsymbol{A}_x(t-p)\boldsymbol{x}(t-p) + \sum_{p=1}^{P} \boldsymbol{B}_y(t-p)\boldsymbol{y}(t-p) + \boldsymbol{\varepsilon}_x(t)$$
$$\boldsymbol{y}(t) = \sum_{p=1}^{P} \boldsymbol{A}_y(t-p)\boldsymbol{y}(t-p) + \sum_{p=1}^{P} \boldsymbol{B}_x(t-p)\boldsymbol{x}(t-p) + \boldsymbol{\varepsilon}_y(t)$$

Multivariate Granger causality--続き

残差ベクトルの共分散行列を定義：

$$\boldsymbol{\Sigma}_e^x = \left\langle \boldsymbol{e}_x(t)\boldsymbol{e}_x^T(t) \right\rangle,\ \boldsymbol{\Sigma}_e^y = \left\langle \boldsymbol{e}_y(t)\boldsymbol{e}_y^T(t) \right\rangle,\ \boldsymbol{\Sigma}_\varepsilon^x = \left\langle \boldsymbol{\varepsilon}_x(t)\boldsymbol{\varepsilon}_x^T(t) \right\rangle,\ \boldsymbol{\Sigma}_\varepsilon^y = \left\langle \boldsymbol{\varepsilon}_y(t)\boldsymbol{\varepsilon}_y^T(t) \right\rangle$$

Granger causality

$$G_{\boldsymbol{y}\to\boldsymbol{x}} = \log\left[\left|\boldsymbol{\Sigma}_e^x\right| / \left|\boldsymbol{\Sigma}_\varepsilon^x\right|\right]$$ (***x***の予測値が***y***を考慮することでどの程度良くなったか)

$$G_{\boldsymbol{x}\to\boldsymbol{y}} = \log\left[\left|\boldsymbol{\Sigma}_e^y\right| / \left|\boldsymbol{\Sigma}_\varepsilon^y\right|\right]$$ (***y***の予測値が***x***を考慮することでどの程度良くなったか)

Multivariateの場合には以下も定義できる.

Total-interdependence： $\boldsymbol{z}(t) = [\boldsymbol{x}^T(t), \boldsymbol{y}^T(t)]^T$ として,

$$\boldsymbol{z}(t) = \sum_{p=1}^{P} \boldsymbol{A}_z(t-p)\boldsymbol{z}(t-p) + \boldsymbol{\varepsilon}_z(t)$$

もし, ***x*** と ***y*** にinteractionが無ければ $\boldsymbol{\Sigma}_\varepsilon^z = \left\langle \boldsymbol{\varepsilon}_z(t)\boldsymbol{\varepsilon}_z^T(t) \right\rangle = \begin{bmatrix} \boldsymbol{\Sigma}_e^x & 0 \\ 0 & \boldsymbol{\Sigma}_e^y \end{bmatrix}$ である.

Total-interdependence --続き

したがって，$\boldsymbol{\Sigma}_{\varepsilon}^{z}$ の，行列 $\begin{bmatrix} \boldsymbol{\Sigma}_{e}^{x} & 0 \\ 0 & \boldsymbol{\Sigma}_{e}^{y} \end{bmatrix}$ からのずれは $\boldsymbol{x}$ と $\boldsymbol{y}$ のinteractionの程度を与える．

Total-interdependence: $$I_{\{\boldsymbol{x},\boldsymbol{y}\}} = \log \frac{\left| \begin{bmatrix} \boldsymbol{\Sigma}_{e}^{x} & 0 \\ 0 & \boldsymbol{\Sigma}_{e}^{y} \end{bmatrix} \right|}{\left| \boldsymbol{\Sigma}_{\varepsilon}^{z} \right|} = \log \frac{\left| \boldsymbol{\Sigma}_{e}^{x} \right| \left| \boldsymbol{\Sigma}_{e}^{y} \right|}{\left| \boldsymbol{\Sigma}_{\varepsilon}^{z} \right|}$$

Total-interdependenceは相互情報量（mutual information）に等しいことを示せる．

Total-interdependenceからGranger causality成分を引いてみる．

$$I_{\{\boldsymbol{x},\boldsymbol{y}\}} - G_{y\to x} - G_{x\to y} = \log \frac{\left| \boldsymbol{\Sigma}_{\varepsilon}^{x} \right| \left| \boldsymbol{\Sigma}_{\varepsilon}^{y} \right|}{\left| \boldsymbol{\Sigma}_{\varepsilon}^{z} \right|} = I_{\{x\bullet y\}}$$ と定義する．

これが何を意味するかは結構悩ましい．Gewekeはinstantaneous dependenceと呼んでいる．

$$I_{\{\boldsymbol{x},\boldsymbol{y}\}} = G_{y\to x} + G_{x\to y} + I_{\{x\bullet y\}}$$

すなわち，total interdependenceは2つのGranger causality成分と向きを持たない成分に分解できる!!

Instantaneous dependence: $I_{\{x \bullet y\}} = \log \frac{\left|\boldsymbol{\Sigma}_\varepsilon^x\right|\left|\boldsymbol{\Sigma}_\varepsilon^y\right|}{\left|\boldsymbol{\Sigma}_\varepsilon^z\right|}$ について

Time seriesに共通の現在の時間にのみ依存する項が含まれる場合

$$\boldsymbol{x}(t) = \sum_{p=1}^{P} \boldsymbol{A}_x(t-p)\boldsymbol{x}(t-p) + \sum_{p=1}^{P} \boldsymbol{B}_y(t-p)\boldsymbol{y}(t-p) + \boldsymbol{v}(t) + \boldsymbol{\varepsilon}_x(t)$$
$$\boldsymbol{y}(t) = \sum_{p=1}^{P} \boldsymbol{A}_y(t-p)\boldsymbol{y}(t-p) + \sum_{p=1}^{P} \boldsymbol{B}_x(t-p)\boldsymbol{x}(t-p) + \boldsymbol{v}(t) + \boldsymbol{\varepsilon}_y(t)$$

かなり長い導出の後，$I_{\{x \bullet y\}} = \log \frac{\left|\left\langle \boldsymbol{v}\boldsymbol{v}^T \right\rangle\right|}{\left|\boldsymbol{\Sigma}_\varepsilon^z\right|}$ を示すことができる(多分).

これは，コヒーレンス解析のseed blurのところで議論した2つのタイムコースに共通に含まれる成分による（見かけの）相互依存項を表す.

であれば，次のような解釈が成り立つ：

imaginary coherenceが検出する相互依存性

$$I_{\{\boldsymbol{x},\boldsymbol{y}\}} = \overbrace{G_{y\to x} + G_{x\to y}} + I_{\{x \bullet y\}}$$

magnitude coherenceが検出する相互依存性

coherenceのreal partが検出する相互依存性

MVAR model周波数領域表現

$$\boldsymbol{y}(t) = \sum_{p=1}^{P} \boldsymbol{A}(p)\boldsymbol{y}(t-p) + \boldsymbol{e}(t) \quad \underset{FT}{\rightleftharpoons} \quad \boldsymbol{y}(f) = \sum_{p=1}^{P} \boldsymbol{A}(p)e^{-2\pi ifp}\boldsymbol{y}(f) + \boldsymbol{e}(f)$$

ただし：

$\boldsymbol{y}(t) \underset{FT}{\rightleftharpoons} \boldsymbol{y}(f)$

$\boldsymbol{e}(t) \underset{FT}{\rightleftharpoons} \boldsymbol{e}(f)$

データベクトル $\boldsymbol{y}(f)$ でまとめると

$$\underbrace{\left[\boldsymbol{I} - \sum_{p=1}^{P} \boldsymbol{A}(p)e^{-2\pi ifp}\right]}_{\bar{\boldsymbol{A}}(f)} \boldsymbol{y}(f) = \boldsymbol{e}(f)$$

したがって，$\bar{\boldsymbol{A}}(f)\boldsymbol{y}(f) = \boldsymbol{e}(f)$ を得る.

さらに，$\boldsymbol{H}(f) = \bar{\boldsymbol{A}}(f)^{-1}$ として，$\boldsymbol{y}(f) = \boldsymbol{H}(f)\boldsymbol{e}(f)$ を得る.

「入力 $\boldsymbol{e}(f)$ に伝達関数 $\boldsymbol{H}(f)$を乗じて，出力 $\boldsymbol{y}(f)$ を得る．」という解釈が成り立つ.

周波数領域のcausality measure

MVAR model周波数領域表現

$$\boldsymbol{y}(f) = \boldsymbol{H}(f)\boldsymbol{e}(f) : \ \boldsymbol{H}(f) = \tilde{\boldsymbol{A}}(f)^{-1}$$

したがって

$$\boldsymbol{S}(f) = \boldsymbol{H}(f)\boldsymbol{\Sigma}\boldsymbol{H}(f)^H$$ ただし，$\boldsymbol{S}(f) = \left\langle \boldsymbol{y}(f)\boldsymbol{y}(f)^H \right\rangle, \quad \boldsymbol{\Sigma} = \left\langle \boldsymbol{e}(f)\boldsymbol{e}(f)^H \right\rangle$

であるので，j チャンネルと k チャンネルのコヒーレンス：

$$\eta_{j,k} = \frac{S_{j,k}(f)}{\sqrt{S_{j,j}(f)S_{k,k}(f)}} = \frac{\boldsymbol{h}_j^H \boldsymbol{\Sigma}\boldsymbol{h}_k}{\sqrt{[\boldsymbol{h}_j^H \boldsymbol{\Sigma}\boldsymbol{h}_j][\boldsymbol{h}_k^H \boldsymbol{\Sigma}\boldsymbol{h}_k]}}$$

となる.

ただし，$\boldsymbol{H}(f)^H = [\boldsymbol{h}_1, \ldots, \boldsymbol{h}_P]$ である.

Spectral Granger Causality (Geweke Measure)

Bivariateの場合を考える.

$$\begin{bmatrix} x(f) \\ y(f) \end{bmatrix} = \begin{bmatrix} H_{xx} & H_{xy} \\ H_{yx} & H_{yy} \end{bmatrix} \begin{bmatrix} e_x(f) \\ e_y(f) \end{bmatrix}$$

$$\boldsymbol{S}(f) = \begin{bmatrix} \left\langle \left| x(f) \right|^2 \right\rangle & \left\langle x(f) y(f)^* \right\rangle \\ \left\langle x(f)^* y(f) \right\rangle & \left\langle \left| y(f) \right|^2 \right\rangle \end{bmatrix} = \boldsymbol{H} \begin{bmatrix} \left\langle \left| e_x(f) \right|^2 \right\rangle & \left\langle e_x(f) e_y(f)^* \right\rangle \\ \left\langle e_x(f)^* e_y(f) \right\rangle & \left\langle \left| e_y(f) \right|^2 \right\rangle \end{bmatrix} \boldsymbol{H}^H = \boldsymbol{H} \begin{bmatrix} \Sigma_x & \gamma \\ \gamma^* & \Sigma_y \end{bmatrix} \boldsymbol{H}^H$$

Total interdependence:

$$I_{\{\boldsymbol{x},\boldsymbol{y}\}} = \log \frac{\left\| \begin{bmatrix} \Sigma_x & 0 \\ 0 & \Sigma_y \end{bmatrix} \right\|}{\left\| \begin{bmatrix} \Sigma_x & \gamma \\ \gamma^* & \Sigma_y \end{bmatrix} \right\|} = \log \frac{\left\langle \left| x(f) \right|^2 \right\rangle \left\langle \left| y(f) \right|^2 \right\rangle}{\left\langle \left| x(f) \right|^2 \right\rangle \left\langle \left| y(f) \right|^2 \right\rangle - \left| \left\langle x(f)^* y(f) \right\rangle \right|^2} = -\log(1 - \left| \eta \right|^2)$$

ここで, η はcoherenceであり, $\left| \eta \right|^2 = \left| \left\langle x(f)^* y(f) \right\rangle \right|^2 / \left[\left\langle \left| x(f) \right|^2 \right\rangle \left\langle \left| y(f) \right|^2 \right\rangle \right]$ である.

Spectral Granger Causalityを導く

$$\begin{bmatrix} x(f) \\ y(f) \end{bmatrix} = \begin{bmatrix} H_{xx} & H_{xy} \\ H_{yx} & H_{yy} \end{bmatrix} \begin{bmatrix} e_x(f) \\ e_y(f) \end{bmatrix}$$ から，若干意味不明の導出を経て，

パワースペクトルの分解：

$$\left\langle \left| x(f) \right|^2 \right\rangle = (H_{xx} + \frac{\gamma^*}{\Sigma_x} H_{xy}) \Sigma_x (H_{xx}^* + \frac{\gamma}{\Sigma_x} H_{xy}^*) + H_{xy} (\Sigma_y - \frac{\left|\gamma\right|^2}{\Sigma_x}) H_{xy}^*$$ を得る．

x(自分自身) からの寄与　　*y*(相手)のからの寄与

したがって，Spectral Granger Causalityは：

$$f_{y \to x}(f) = \log \frac{\text{トータルのスペクトルパワー}}{\text{トースペクトルパワーの内で}x\text{(自分自身)からの寄与分}}$$

$$= -\log \left[1 - \frac{H_{xy} (\Sigma_y - \frac{\left|\gamma\right|^2}{\Sigma_x}) H_{xy}^*}{\left\langle \left| x(f) \right|^2 \right\rangle} \right]$$ である．$f_{x \to y}(f)$ も同じ考え方で導ける．

ここまでで述べた周波数領域のGranger causality(Geweke measureと呼ばれる場合もある)は：

- 3チャンネル以上の場合について拡張ができない.
- 途中の導出もわかりにくい.

等の問題がある.

一方で，周波数領域のCausality measureについてはもっと解りやすいものが知られている.

Directed transfer function (DTF)とPartial directed coherence (PDC)

bivariateな場合で考えてみる.

$$y_1(t) = \sum_{p=1}^{P} A_{1,1}(t-p)y_1(t-p) + \sum_{p=1}^{P} A_{1,2}(t-p)y_2(t-p) + \varepsilon_1(t)$$
$$y_2(t) = \sum_{p=1}^{P} A_{2,1}(t-p)y_1(t-p) + \sum_{p=1}^{P} A_{2,2}(t-p)y_2(t-p) + \varepsilon_2(t)$$

$y_2(t)$から$y_1(t)$への影響は$A_{1,2}(p)$, $p=1,\ldots P$で評価できる．すなわち，2→1のcausal influenceは（例えば）$\sqrt{\sum_{p=1}^{P} A_{1,2}(p)^2}$ により評価できる.

周波数領域の量で$A_{1,2}(p)$, $p=1,\ldots P$を情報として含んでいるのは

および

$H_{1,2}(f)$ ➡ Directed transfer function (DTF)

$\bar{A}_{1,2}(p)$ ➡ Partial directed coherence (PDC)

が導かれる.

MVAR model周波数領域表現ー復習

$$\boldsymbol{y}(t) = \sum_{p=1}^{P} \boldsymbol{A}(p)\boldsymbol{y}(t-p) + \boldsymbol{e}(t) \quad \underset{FT}{\rightleftharpoons} \quad \boldsymbol{y}(f) = \sum_{p=1}^{P} \boldsymbol{A}(p)e^{-2\pi i f p}\boldsymbol{y}(f) + \boldsymbol{e}(f)$$

データベクトル $\boldsymbol{y}(f)$ でまとめると

$$\underbrace{\left[\boldsymbol{I} - \sum_{p=1}^{P} \boldsymbol{A}(p)e^{-2\pi i f p}\right]}_{\bar{\boldsymbol{A}}(f)}\boldsymbol{y}(f) = \boldsymbol{e}(f)$$

$\boxed{\text{ただし}}$:

$$\boldsymbol{y}(t) \underset{FT}{\rightleftharpoons} \boldsymbol{y}(f)$$
$$\boldsymbol{e}(t) \underset{FT}{\rightleftharpoons} \boldsymbol{e}(f)$$

したがって，$\bar{\boldsymbol{A}}(f)\boldsymbol{y}(f) = \boldsymbol{e}(f)$ を得る.

さらに，$\boldsymbol{H}(f) = \bar{\boldsymbol{A}}(f)^{-1}$ として，$\boldsymbol{y}(f) = \boldsymbol{H}(f)\boldsymbol{e}(f)$ を得る.

「入力 $\boldsymbol{e}(f)$ に伝達関数 $\boldsymbol{H}(f)$を乗じて，出力 $\boldsymbol{y}(f)$ を得る.」という解釈が成り立つ.

Directed transfer function (DTF)

$$\boldsymbol{H}(f) = \bar{\boldsymbol{A}}(f)^{-1} \quad : \quad \bar{\boldsymbol{A}}(f) = \boldsymbol{I} - \sum_{p=1}^{P} \boldsymbol{A}(p) e^{-2\pi i f p}$$ と定義された$\boldsymbol{H}(f)$を用いる.

bivariateの場合：

$$H_{1,2}(f) = \frac{\bar{A}_{1,2}(f)}{\left|\tilde{\boldsymbol{A}}(f)\right|}$$

であるので，$\bar{\boldsymbol{A}}(f)$を用いるのと同じ結果になる.

multivariateの場合：

$$H_{1,2}(f) = \frac{\left|\boldsymbol{M}_{2,1}(f)\right|}{\left|\bar{\boldsymbol{A}}(f)\right|}$$

$\longleftarrow \bar{\boldsymbol{A}}(f)$の(2,1)要素に対するマイナー

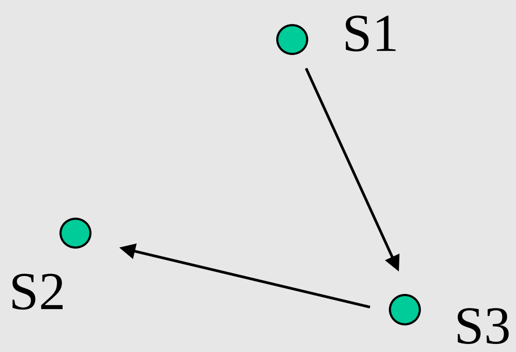
trivariateの場合：

$$H_{1,2}(f) = \frac{\left|\bar{A}_{1,2}\bar{A}_{3,3} - \bar{A}_{3,1}\bar{A}_{2,3}\right|}{\left|\bar{\boldsymbol{A}}(f)\right|}$$

$H_{1,2}(f)$は2→1の情報の流れが無くても，つまり，$\bar{A}_{1,2} = 0$ でもゼロにはならない.

Directed transfer function (DTF): a trivariate case

$$H_{1,2}(f)=\frac{\left|\bar{A}_{1,2}\bar{A}_{3,3}-\bar{A}_{3,1}\bar{A}_{2,3}\right|}{\left|\bar{\boldsymbol{A}}(f)\right|}$$ の意味



「S1からS3に情報の流れがあり，さらに，S3からS2に情報の流れがあるが，S1からS2へは直接情報の流れが無い」場合を考える.

S1からS2へは直接情報の流れが無い ⇒ $\bar{A}_{1,2}=0$

S1からS3に情報の流れがあり，さらにS3からS2に情報の流れがある ⇒ $\bar{A}_{3,1}\neq 0$, $\bar{A}_{2,3}\neq 0$

すなわち，DTFはS1からS3を経由したS2への情報の流れを(indirect influence)検出してしまう．DTFはdirect influenceとindirect influenceを区別できない.

Directed transfer function (DTF)

次のnormalized formが用いられる：

$$\mu_{k\to j} = \frac{H_{j,k}(f)}{\sqrt{\boldsymbol{h}_j^H \boldsymbol{h}_j}} = \frac{H_{j,k}(f)}{\sqrt{\sum_{m=1}^{Q} \left| H_{j,m}(f) \right|^2}}$$

DTFを用いてcoherenceをDTFの積と和で表すことができる

$$\eta_{j,k}(f) = \frac{S_{j,k}}{\sqrt{S_{j,j} S_{k,k}}} = \frac{\boldsymbol{h}_j^H \boldsymbol{\Sigma} \boldsymbol{h}_k}{\sqrt{\left[\boldsymbol{h}_j^H \boldsymbol{\Sigma} \boldsymbol{h}_j\right]\left[\boldsymbol{h}_k^H \boldsymbol{\Sigma} \boldsymbol{h}_k\right]}} \qquad (\ \boldsymbol{\Sigma} = \boldsymbol{I}\ \text{と仮定して})$$

$$= \frac{\boldsymbol{h}_j^H \boldsymbol{h}_k}{\sqrt{\left[\boldsymbol{h}_j^H \boldsymbol{h}_j\right]\left[\boldsymbol{h}_k^H \boldsymbol{h}_k\right]}} = \frac{\sum_{m=1}^{Q} H_{j,m}(f) H_{k,m}^*(f)}{\sqrt{\left[\boldsymbol{h}_j^H \boldsymbol{h}_j\right]\left[\boldsymbol{h}_k^H \boldsymbol{h}_k\right]}} = \sum_{m=1}^{Q} \mu_{m\to j} \mu_{m\to k}^*$$

このfactorizationの意味は，今一つはっきりしない．

Partial Directed Coherence (PDC)

Transfer functionではなく，MVAR coefficient matrixを直接使う．

以下のように定義する．

$$\pi_{k\to j} = \frac{\bar{A}_{j,k}(f)}{\sqrt{\bar{\boldsymbol{a}}_j^H \bar{\boldsymbol{a}}_j}} : \quad \text{ここで，} \bar{\boldsymbol{A}}(f) = [\bar{\boldsymbol{a}}_1, \ldots, \bar{\boldsymbol{a}}_Q] \text{ である．}$$

MVAR coefficient matrixを直接使うため，indirect influenceを含まない．すなわち，$\bar{A}_{j,k}(f) = 0$ なら，$\pi_{k\to j} = 0$ である．

実は，coherenceとDTFの関係と同様な関係がpartial coherenceとPDCの間に成り立つ．

名前の由来

MVAR model周波数領域表現

$$\boldsymbol{S}(f) = \boldsymbol{H}(f)\boldsymbol{\Sigma}\boldsymbol{H}(f)^H \quad \text{ただし，} \boldsymbol{S}(f) = \left\langle \boldsymbol{y}(f)\boldsymbol{y}(f)^H \right\rangle, \quad \boldsymbol{\Sigma} = \left\langle \boldsymbol{e}(f)\boldsymbol{e}(f)^H \right\rangle$$

Coherence

$$\eta_{j,k} = \frac{S_{j,k}(f)}{\sqrt{S_{j,j}(f)S_{k,k}(f)}} = \frac{\boldsymbol{h}_j^H \boldsymbol{\Sigma}\boldsymbol{h}_k}{\sqrt{[\boldsymbol{h}_j^H \boldsymbol{\Sigma}\boldsymbol{h}_j][\boldsymbol{h}_k^H \boldsymbol{\Sigma}\boldsymbol{h}_k]}} = \sum_{m=1}^{Q} \boldsymbol{\mu}_{j\to m}\boldsymbol{\mu}^*_{m\to k}$$

CoherenceはDTFの積の全てのチャンネルによる和として表される.

Partial coherence

$$\kappa_{j,k} = \frac{\boldsymbol{M}_{j,k}(f)}{\sqrt{\boldsymbol{M}_{j,j}(f)\boldsymbol{M}_{k,k}(f)}} = \frac{\overline{\boldsymbol{a}}_j^H \boldsymbol{\Sigma}^{-1}\overline{\boldsymbol{a}}_k}{\sqrt{[\overline{\boldsymbol{a}}_j^H \boldsymbol{\Sigma}^{-1}\overline{\boldsymbol{a}}_j][\overline{\boldsymbol{a}}_k^H \boldsymbol{\Sigma}^{-1}\overline{\boldsymbol{a}}_k]}}$$

j番目の変数とk番目の変数に直接の相互依存関係が無い場合にはゼロとなる.

補足説明-Partial correlation coefficient（偏相関係数）

Partial coherenceとCoherenceの関係は，相関係数(Correlation)と偏相関係数(Partial correlation) の関係と同じである．偏相関係数については以下の例に説明される．

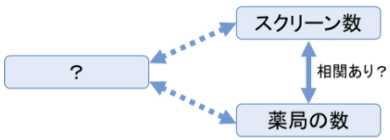
26-4. 偏相関係数

次のデータは2015年12月末時点の各都道府県内にある映画館のスクリーンの合計数と可住地面積100km^2当たりの薬局数を表したものです。このデータを用いて相関係数を算出すると、「0.82」でした。つまり、映画館のスクリーン数と薬局の数には強い相関があるという結果でした。

	A	B	C
1		映画館全スクリーン数	薬局数
2	北海道	113	10.3
3	青森県	44	18.1
4	岩手県	25	15.5
5	宮城県	64	35.2
43	長崎県	26	44.9
44	熊本県	49	29.7
45	大分県	35	31.4
46	宮崎県	18	31.9
47	鹿児島県	31	26.5
48	沖縄県	40	47.5
49			

出典：総務省統計局 社会生活統計指標－都道府県の指標－2015

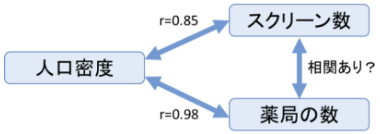
しかし、一般的に考えて都道府県ごとの映画館のスクリーン数と可住地面積100km^2当たりの薬局の数は直接的に関係がないような気がします。映画館のスクリーン数が多いから薬局の出店数が増えるわけでも、薬局の数が多いから映画館のスクリーン数が増えるわけでもないためです。このような場合には、「第3の因子」の存在を考慮する必要があります。



上のデータに各都道府県の人口密度のデータを加えてみます。

	A	B	C	D
1		映画館全スクリーン数	薬局数	人口密度
2	北海道	113	10.3	69.3
3	青森県	44	18.1	138.4
4	岩手県	25	15.5	84.8
5	宮城県	64	35.2	319.5
43	長崎県	26	44.9	340.2
44	熊本県	49	29.7	243.2
45	大分県	35	31.4	185.8
46	宮崎県	18	31.9	144.8
47	鹿児島県	31	26.5	182.8
48	沖縄県	40	47.5	621.5
49				

人口密度と映画館のスクリーン数、及び人口密度と薬局の数の相関係数はそれぞれ「0.85」と「0.98」でした。つまり、人口密度がスクリーン数と薬局の数それぞれと強い相関を持っているため、これらの影響を除いた上で映画館のスクリーン数と薬局の数との相関関係を調べる必要があります。



映画館のスクリーン数と薬局の数のような相関関係のことを「見かけ上の相関」や「疑似相関」といいます。見かけ上の相関がある場合は、相関係数ではなく第3の因子の影響を除いた相関係数である「偏相関係数」を用いて相関関係を評価します。1つ目の因子をx、2つ目の因子をy、3つ目の因子をzとおき、xとyの相関係数をr_{xy}、yとzの相関係数をr_{yz}、zとxの相関係数をr_{zx}とします。これらを用いると、zの影響を除いたxとyの偏相関係数$r_{xy\cdot z}$を次の式から求められます。

$$r_{xy\cdot z} = \frac{r_{xy} - r_{xz}r_{yz}}{\sqrt{1-r_{xz}^2}\sqrt{1-r_{yz}^2}}$$

上のデータの映画館のスクリーン数、薬局の数、人口密度をそれぞれx、y、zとおくと、相関係数はそれぞれ$r_{xy} = 0.82$、$r_{yz} = 0.98$、$r_{zx} = 0.85$となるので、偏相関係数$r_{xy\cdot z}$は「-0.13」となります。

$$r_{xy\cdot z} = \frac{r_{xy} - r_{xz}r_{yz}}{\sqrt{1-r_{xz}^2}\sqrt{1-r_{yz}^2}} = \frac{0.82 - 0.98 \times 0.85}{\sqrt{1-0.98^2}\sqrt{1-0.85^2}} = -0.13$$

この結果から、映画館のスクリーン数と薬局の数との相関は、実はあまり強くないことが分かります。

変数がQ個　$x_1, \ldots, x_Q$ の場合の偏相関係数(Partial correlation)は以下のように計算する.

通常の相関係数： $r_{j,k} = \dfrac{\langle x_j x_k \rangle}{\sqrt{\langle x_j^2 \rangle \langle x_k^2 \rangle}}$　　相関行列： $\boldsymbol{R} = \begin{bmatrix} r_{1,1} & \cdots & r_{1,Q} \\ \vdots & \ddots & \vdots \\ r_{Q,1} & \cdots & r_{Q,Q} \end{bmatrix}$

x_j と x_kの偏相関係数 $\tilde{r}_{j,k}$ は：

$$\tilde{r}_{j,k} = \frac{-[\boldsymbol{R}^{-1}]_{j,k}}{\sqrt{[\boldsymbol{R}^{-1}]_{j,j}[\boldsymbol{R}^{-1}]_{k,k}}} = \frac{\left|\boldsymbol{M}_{j,k}\right|}{\sqrt{\left|\boldsymbol{M}_{j,j}\right|\left|\boldsymbol{M}_{k,k}\right|}}$$

と与えられる．ここで，$[\boldsymbol{R}^{-1}]_{j,k}$ は $\boldsymbol{R}$ の逆行列の(j,k)要素，$\left|\boldsymbol{M}_{j,k}\right|$ は $\boldsymbol{R}$ の(j,k)マイナーである.

Partial coherence

$$\kappa_{j,k} = \frac{\boldsymbol{M}_{j,k}(f)}{\sqrt{\boldsymbol{M}_{j,j}(f)\boldsymbol{M}_{k,k}(f)}} = \frac{\overline{\boldsymbol{a}}_j^H \boldsymbol{\Sigma}^{-1} \overline{\boldsymbol{a}}_k}{\sqrt{[\overline{\boldsymbol{a}}_j^H \boldsymbol{\Sigma}^{-1} \overline{\boldsymbol{a}}_j][\overline{\boldsymbol{a}}_k^H \boldsymbol{\Sigma}^{-1} \overline{\boldsymbol{a}}_k]}}$$

$\boldsymbol{\Sigma} = \boldsymbol{I}$ の仮定のもとで，Partial coherenceはPDCの積と和で表される.

$$\kappa_{j,k} = \frac{\overline{\boldsymbol{a}}_j^H \overline{\boldsymbol{a}}_k}{\sqrt{[\overline{\boldsymbol{a}}_j^H \overline{\boldsymbol{a}}_j][\overline{\boldsymbol{a}}_k^H \overline{\boldsymbol{a}}_k]}} = \frac{\sum_{m=1}^{Q} \overline{A}^*_{m,j} \overline{A}_{m,k}}{\sqrt{[\overline{\boldsymbol{a}}_j^H \overline{\boldsymbol{a}}_j][\overline{\boldsymbol{a}}_k^H \overline{\boldsymbol{a}}_k]}} = \sum_{m=1}^{Q} \pi^*_{j\to m} \pi_{k\to m}$$

DTFがcoherenceをfactorizeするのに対して，PDCはpartial coherenceをfactorizeする.
このことがPDCの名前の由来である.

Coherenceの場合

$$\eta_{j,k} = \frac{S_{j,k}(f)}{\sqrt{S_{j,j}(f) S_{k,k}(f)}} = \frac{\boldsymbol{h}_j^H \boldsymbol{h}_k}{\sqrt{[\boldsymbol{h}_j^H \boldsymbol{h}_j][\boldsymbol{h}_k^H \boldsymbol{h}_k]}} = \sum_{m=1}^{Q} \mu_{j\to m} \mu^*_{m\to k}$$

係数行列 $\boldsymbol{A}(p)$ の推定

第k要素についてMVAR processを書き下すと：

$$y_k(t) = \sum_{j=1}^{Q} A_{k,j}(1) y_j(t-1) + \sum_{j=1}^{Q} A_{k,j}(2) y_j(t-2) + \cdots + \sum_{j=1}^{Q} A_{k,j}(P) y_j(t-P) + e_k(t)$$

時間点 $t=1,\ldots K$ で計測が行なわれたとして， K が十分大きいとすれば，上式が$t=P+1$から$t=K$ まで成り立つので，以下の線形方程式(Yule-Walker equation)を得る.

$$\boldsymbol{y}_k = \boldsymbol{G}\boldsymbol{x}_k + \boldsymbol{e}_k \quad (k=1,\ldots,Q)$$

$$\boldsymbol{y}_k = [y_k(P+1), y_k(P+2), \ldots, y_k(K)]$$

$$\boldsymbol{x}_k = \left[A_{k,1}(1),\ldots,A_{k,Q}(1),\ldots,A_{k,1}(P),\ldots,A_{k,Q}(P)\right]$$

$$\boldsymbol{G} = \begin{bmatrix} y_1(P) & \cdots & y_Q(P) & \cdots & y_1(1) & \cdots & y_Q(1) \\ y_1(P+1) & \cdots & y_Q(P+1) & \cdots & y_1(2) & \cdots & y_Q(2) \\ & \vdots & & \vdots & & \vdots & \\ y_1(K) & \cdots & y_Q(K) & \cdots & y_1(K-P) & \cdots & y_Q(K-P) \end{bmatrix}$$

したがって，最小二乗解：$\hat{\boldsymbol{x}}_k = \left(\boldsymbol{G}^T\boldsymbol{G}\right)^{-1}\boldsymbol{G}^T\boldsymbol{y}_k$ を得る.

最小二乗解は観測データに(妨害信号やセンサーノイズなどの)ARプロセスに従わない成分を含む場合に大きな誤差を含む可能性が有る.

観測データに妨害信号等が含まれる場合によりrobustな解を得るには

Yule-Walker方程式のSparse Bayes解

「係数行列は，ほとんどの要素がゼロに近いsparse matrixである」はリーズナブルな制約であろう.

確率モデル：

事前分布： $f(\boldsymbol{x}_k) = N(\boldsymbol{x}_k \mid 0, \boldsymbol{\nu})$

データ尤度： $f(\boldsymbol{y}_k \mid \boldsymbol{x}_k) = N(\boldsymbol{y}_k \mid \boldsymbol{\Phi}\boldsymbol{x}_k, \boldsymbol{\Lambda})$

事後分布： $f(\boldsymbol{x}_k \mid \boldsymbol{y}_k) = N(\boldsymbol{y}_k \mid \overline{\boldsymbol{x}}_k, \boldsymbol{\Gamma})$

優決定の推定問題なので，EMアルゴリズムが使える

E-step:

$$\overline{\boldsymbol{x}}_k = \boldsymbol{\Gamma}^{-1}\boldsymbol{\Phi}^T\boldsymbol{\Lambda}\boldsymbol{y}_k$$

$$\boldsymbol{\Gamma} = \boldsymbol{\Phi}^T\boldsymbol{\Lambda}\boldsymbol{\Phi} + \boldsymbol{\nu}$$

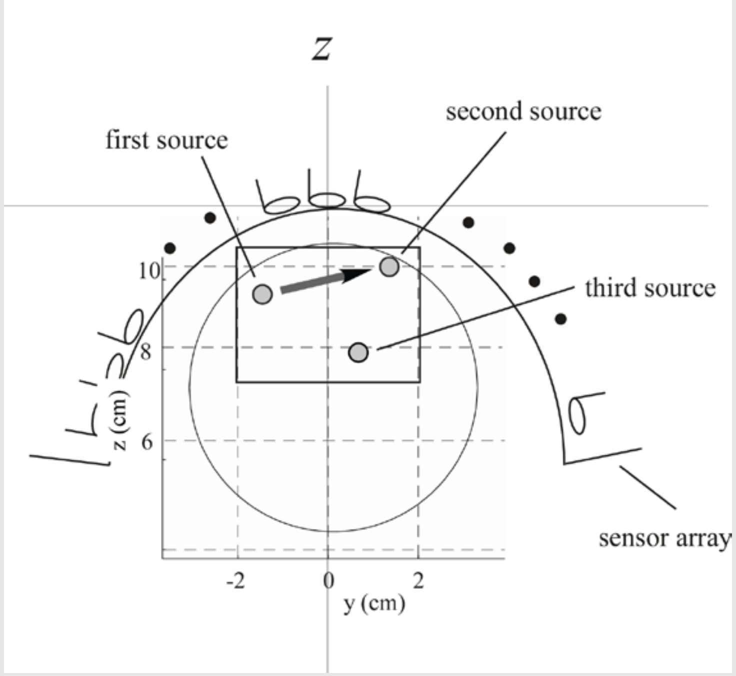
M-step:

$$\nu_{j,j} = \frac{1}{\left[\boldsymbol{R}_{xx}\right]_{j,j}}, \quad \boldsymbol{R}_{xx} = \overline{\boldsymbol{x}}_k\overline{\boldsymbol{x}}_k^T + \boldsymbol{\Gamma}^{-1}$$

$$\boldsymbol{\Lambda}^{-1} = \mathrm{diag}[\boldsymbol{R}_{yy} - \boldsymbol{R}_{xy}^T\boldsymbol{\Phi}^T - \boldsymbol{\Phi}\boldsymbol{R}_{xy} + \boldsymbol{\Phi}\boldsymbol{R}_{xx}\boldsymbol{\Phi}^T]$$

$$\boldsymbol{R}_{yy} = \boldsymbol{y}_k\boldsymbol{y}_k^T, \quad \boldsymbol{R}_{xy} = \overline{\boldsymbol{x}}_k\boldsymbol{y}_k^T$$

Computer simulation:
Source space causality analysis--bivariate case



ソースタイムコース

$$\begin{bmatrix} s_1(t) \\ s_2(t) \end{bmatrix} = \begin{bmatrix} 0.9 & 0 \\ 0.16 & 0.8 \end{bmatrix} \begin{bmatrix} s_1(t-1) \\ s_2(t-1) \end{bmatrix} + \begin{bmatrix} -0.5 & 0 \\ -0.2 & -0.5 \end{bmatrix} \begin{bmatrix} s_1(t-1) \\ s_2(t-1) \end{bmatrix} + \boldsymbol{e}(t)$$

$$s_3(t) \sim N(s_3(t) \mid 0, \sigma^2)$$

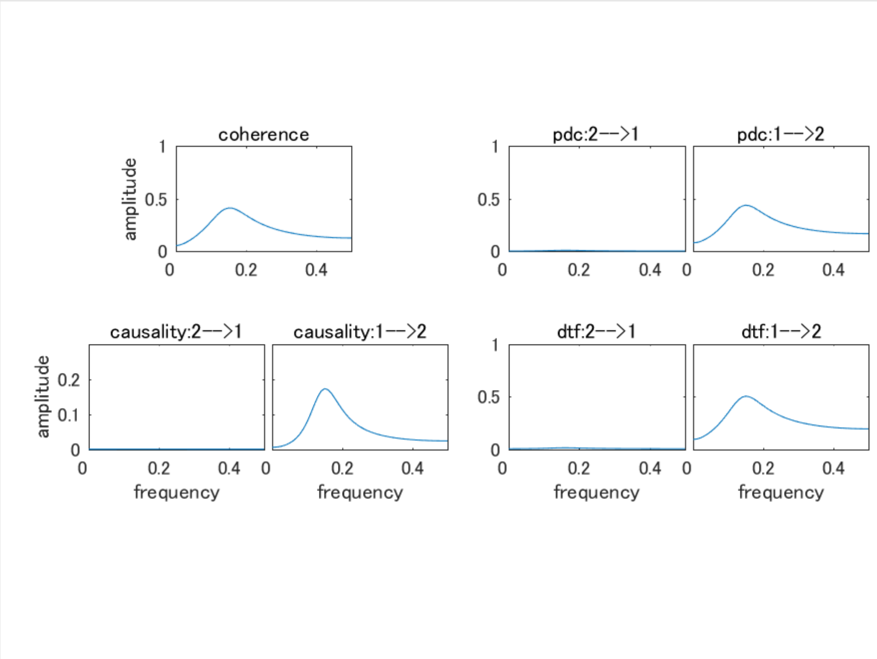
- S1→S2にdirect causal couplingを仮定.
- S3はS1，S2とcouplingを持たない.

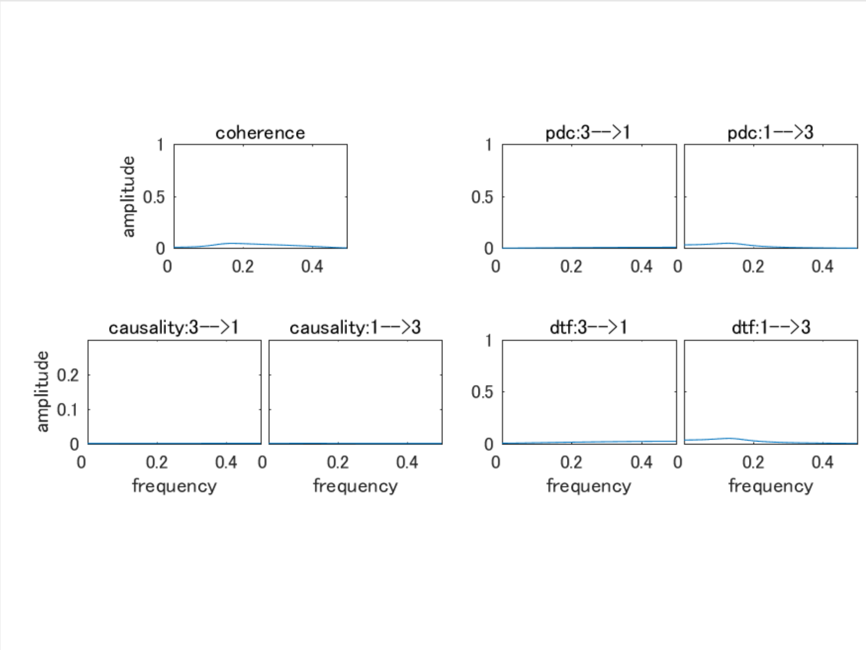
センサーデータ

$$\boldsymbol{b}(t) = \sum_{j=1}^{3} \boldsymbol{l}(\boldsymbol{r}_j) s_j(t) + \boldsymbol{\varepsilon}$$

- CTFアレイを仮定して磁場を計算しわずかなセンサーノイズを与え，センサーデータとする.
- Champageによる信号源再構成.
- 仮定したソース位置のもっとも近いボクセルのタイムコースを用い，$\hat{s}_1(t), \hat{s}_2(t), \hat{s}_3(t)$とする.

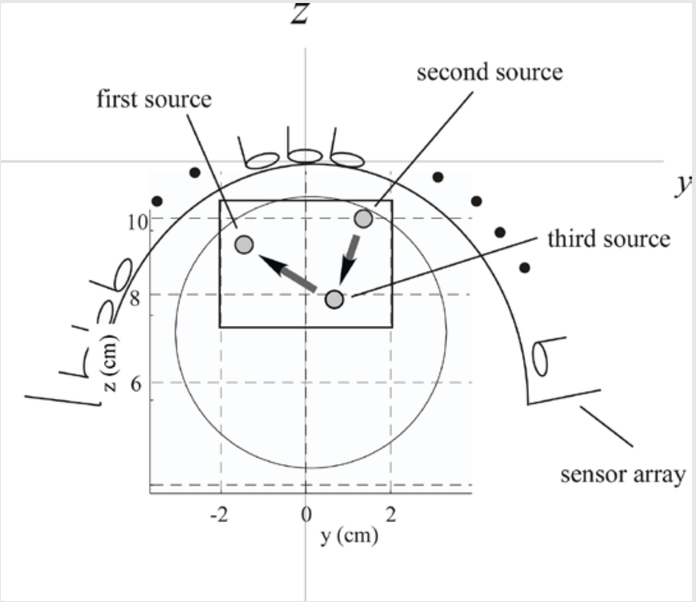
S1,S2間のinterdependence　　S1,S3間のinterdependence





- S1からS2への情報の流れがGeweke measure, pdc, dtfともに捕らえられている．
- S1とS3の間にはcausal interdependenceが小さい（無い）．
- Geweke causalityは（考え方として，よりシンプルな）pdcやdtfに対し，特にアドバンテージはない．

Computer simulation:
Source space causality analysis--trivariate case



ソースタイムコース

$$\begin{bmatrix} s_1(t) \\ s_2(t) \\ s_3(t) \end{bmatrix} = \begin{bmatrix} 0.8 & 0 & 0.4 \\ 0 & 0.9 & 0 \\ 0 & 0.5 & 0.5 \end{bmatrix} \begin{bmatrix} s_1(t-1) \\ s_2(t-1) \\ s_3(t-1) \end{bmatrix} + \begin{bmatrix} -0.5 & 0 & 0 \\ 0 & -0.8 & 0 \\ 0 & 0 & -0.2 \end{bmatrix} \begin{bmatrix} s_1(t-2) \\ s_2(t-2) \\ s_3(t-2) \end{bmatrix} + \boldsymbol{e}(t)$$

- S2→S3，S3→S1はdirect couplingを持つ.
- S2→S1はdirect couplingは無い.

センサーデータ

$$\boldsymbol{b}(t) = \sum_{j=1}^{3} \boldsymbol{l}(\boldsymbol{r}_j) s_j(t) + \alpha \boldsymbol{b}_I(t) + \boldsymbol{\varepsilon}$$

自発脳磁解を加える．α はSIRを調節する定数

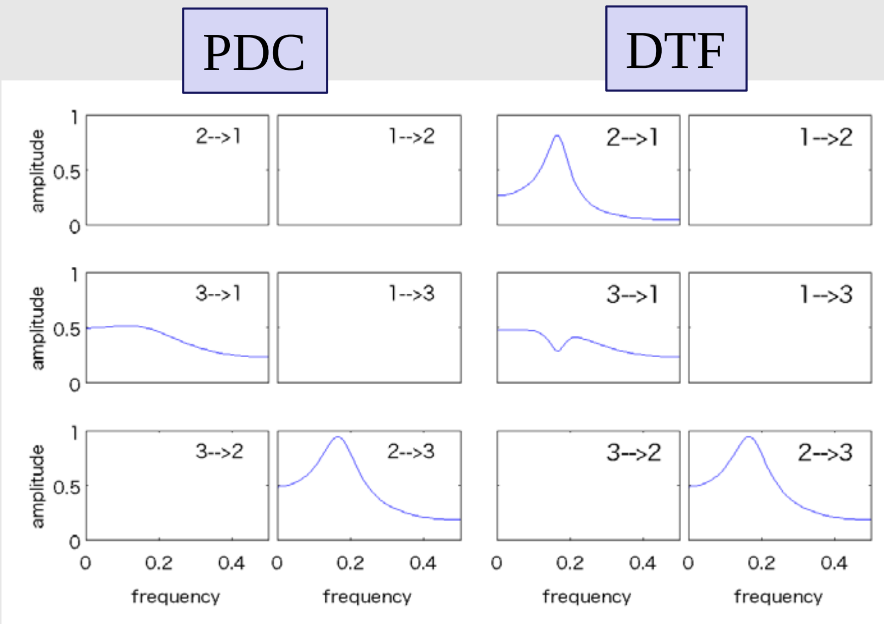
- CTFアレイを仮定し，センサーのイズを与えてMEGデータを発生.
- Champagneによる信号源再構成.
- 仮定したソース位置のもっとも近いボクセルのタイムコースを用い,

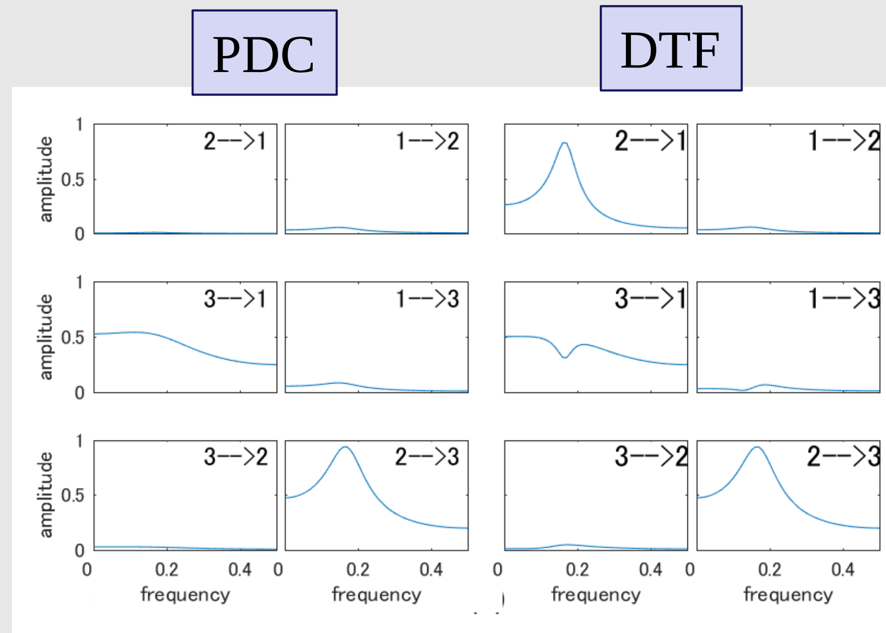
  $\hat{s}_1(t), \hat{s}_2(t), \hat{s}_3(t)$ とする.

ground truth:
データ発生に用いたARcoefficient
を用いてDTFとPDCを計算

ボクセルタイムコースから最小
自乗法によりAR coefficientを
推定し，DTFとPDCを計算.

SIR=8:（高SIRを仮定）



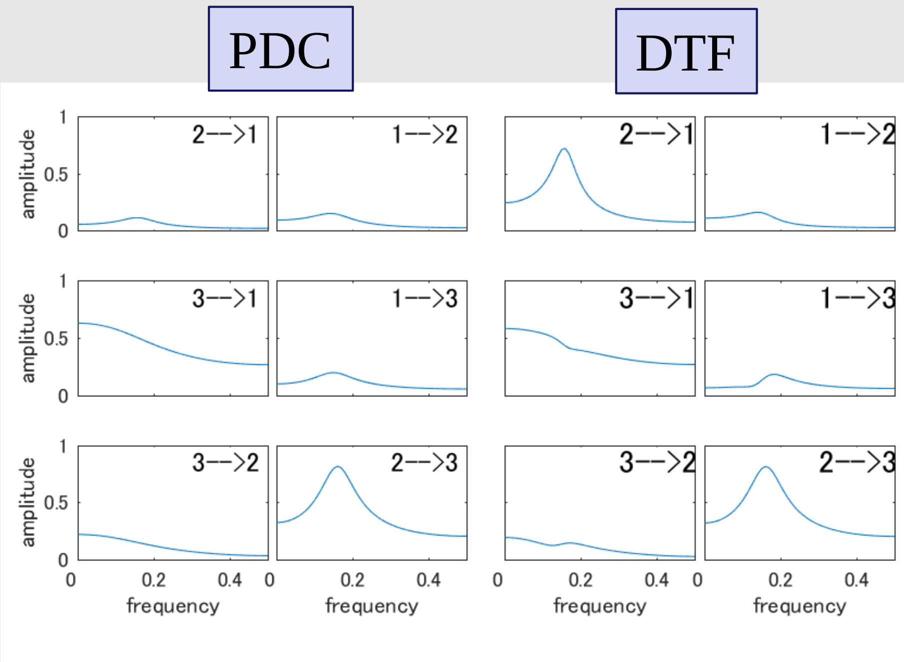


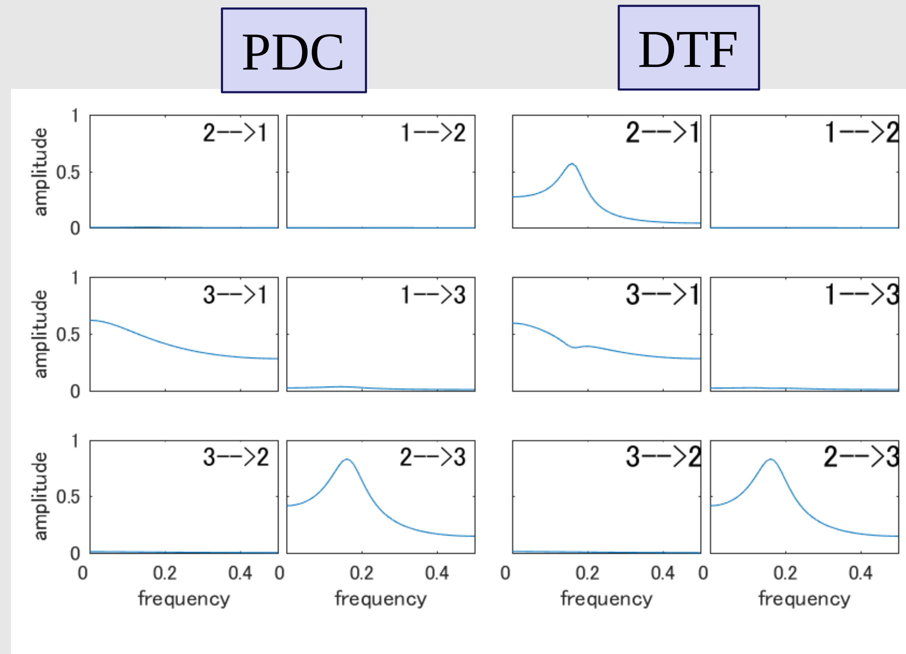
- DTFはS2→S1へのindirect couplingを検出してしまうが，PDCはdirect couplingのみを検出している.
- 高SIRの場合には最小二乗法でAR coefficientを推定しても大きな問題は無い.

低SIRでの実験（SIR=2）

ボクセルタイムコースから最小自乗法によりAR coefficientを推定し，DTFとPDCを計算.

ボクセルタイムコースからSparse Bayes法によりAR coefficientを推定し，DTFとPDCを計算.





- 低SIRの場合には最小二乗法によるAR coefficient推定は誤差を含み，すべてのconnectionで有る程度のcouplingを検出してしまっているが，Sparse Bayesを用いれば，誤差は大幅に低減される.
- しかし，いずれにしろ，得られた結果の有意性判定が必要か.

有意性判定（Statistical thresholding）

Surrogate-data bootstrapping

$$s(t) \underset{FT}{\rightarrow} \sigma(f)\exp[-2\pi i\varepsilon] \underset{IFT}{\rightarrow} s^*(t)$$

の手順でsurrogate dataを発生し，このデータを用いてAR coefficientを推定しcausality measureを計算．surrogate dataによって得られたnull distributionを用いて有意性を検定.

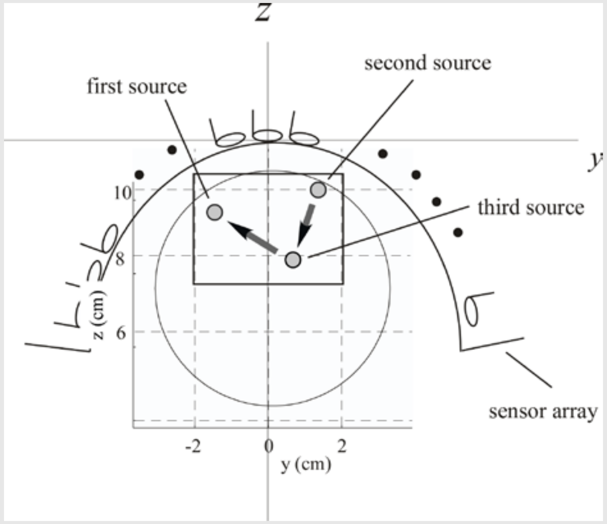
Permutation test（ control dataの存在を前提とする ）

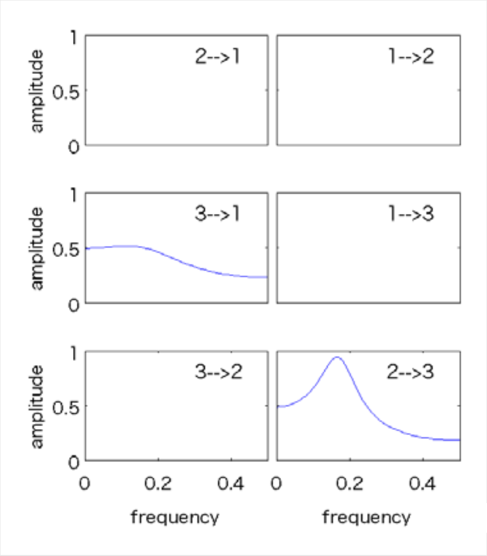
task dataとcontrol dataがそれぞれN_{e} 個のtrialとして計測されているとする.
それぞれから計算したAR coefficientを

$$A_{j,k}^{(m)},\ \tilde{A}_{j,k}^{(m)} \quad m = 1, \ldots, N_e$$

とする．ランダムに選んだNe/2個の $A_{j,k}^{(m)}$ と $\tilde{A}_{j,k}^{(m)}$ から $\left\langle A_{j,k}^{(m)} \right\rangle$ を計算し，これを用いてcausality measureを計算する．選び方を変えてcausality measureのnull distributionを求め，これを用いて有意性を検定.
（多くの場合control dataを得るのがむずかしい）

Computer simulation--trivariate case
Statistical thresholding





仮定したAR coefficient から求めたPDC

Interacting source scenario

ソースタイムコース(ここまでのスライドと同じ)

$$\begin{bmatrix} s_1(t) \\ s_2(t) \\ s_3(t) \end{bmatrix} = \begin{bmatrix} 0.8 & 0 & 0.4 \\ 0 & 0.9 & 0 \\ 0 & 0.5 & 0.5 \end{bmatrix} \begin{bmatrix} s_1(t-1) \\ s_2(t-1) \\ s_3(t-1) \end{bmatrix} + \begin{bmatrix} -0.5 & 0 & 0 \\ 0 & -0.8 & 0 \\ 0 & 0 & -0.2 \end{bmatrix} \begin{bmatrix} s_1(t-2) \\ s_2(t-2) \\ s_3(t-2) \end{bmatrix} + \mathbf{e}(t)$$

Non-interacting source scenario

$s_1(t)$: 正規ノイズを[0,0.3]の帯域のLow-pass filterを適用

$s_2(t)$: 正規ノイズを[0,0.2]の帯域のLow-pass filterを適用

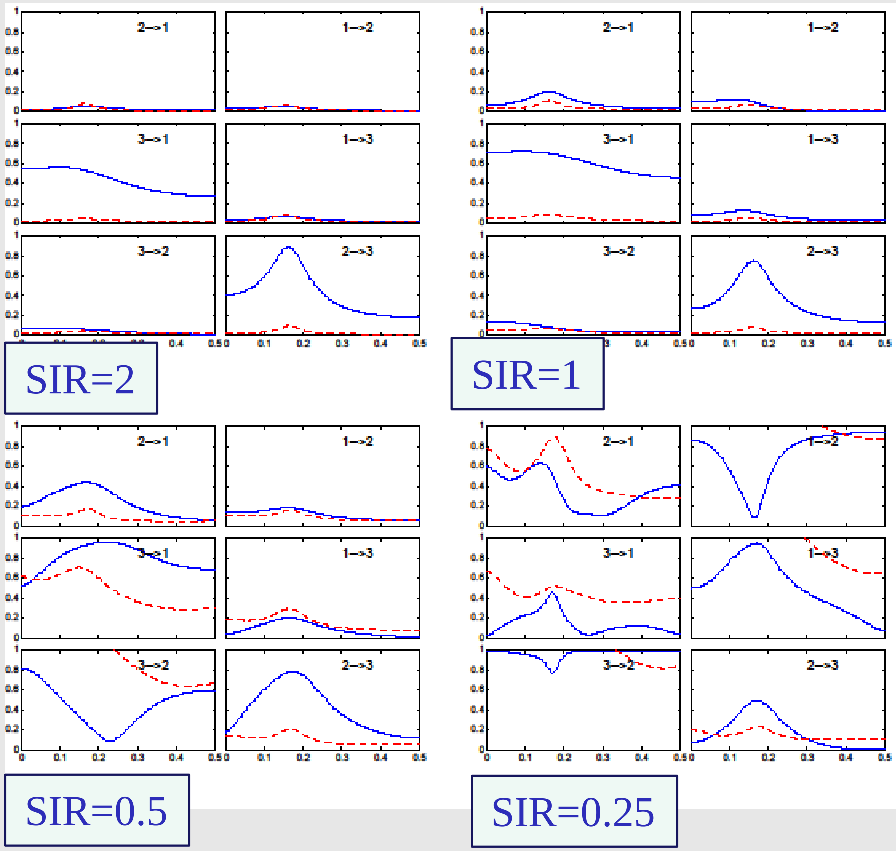
$s_3(t)$: 正規ノイズを[0,0.15]の帯域のLow-pass filterを適用

実は，こうすると大きな偽のcausalityを生じることを知っていた.

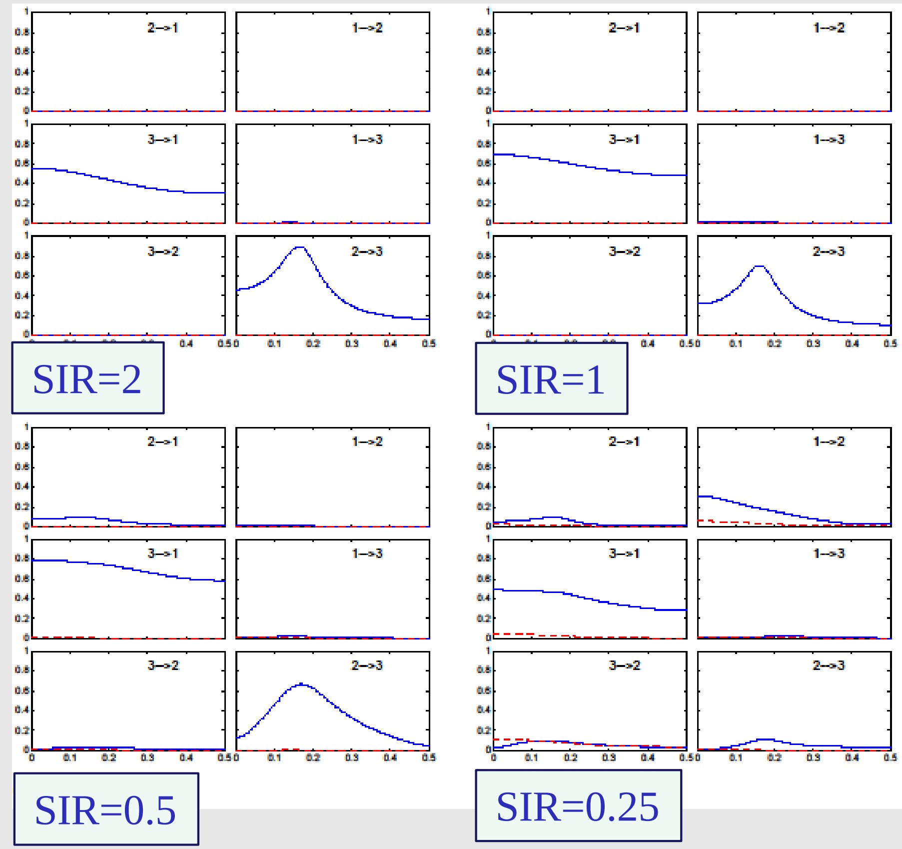
Statistical thresholding using surrogate-data bootstrapping

95%-significance level

最小自乗法によるAR coefficient推定



Sparse BayesによるAR coefficient推定



- Sparse BayesによるAR coefficient推定は最小二乗法に比べrobustである.
- surrogate-data bootstrappingによるthresholdはSparse Bayes法に対しては低すぎる.

赤い点線：statistical threshold, 青い実線：causality measureの値

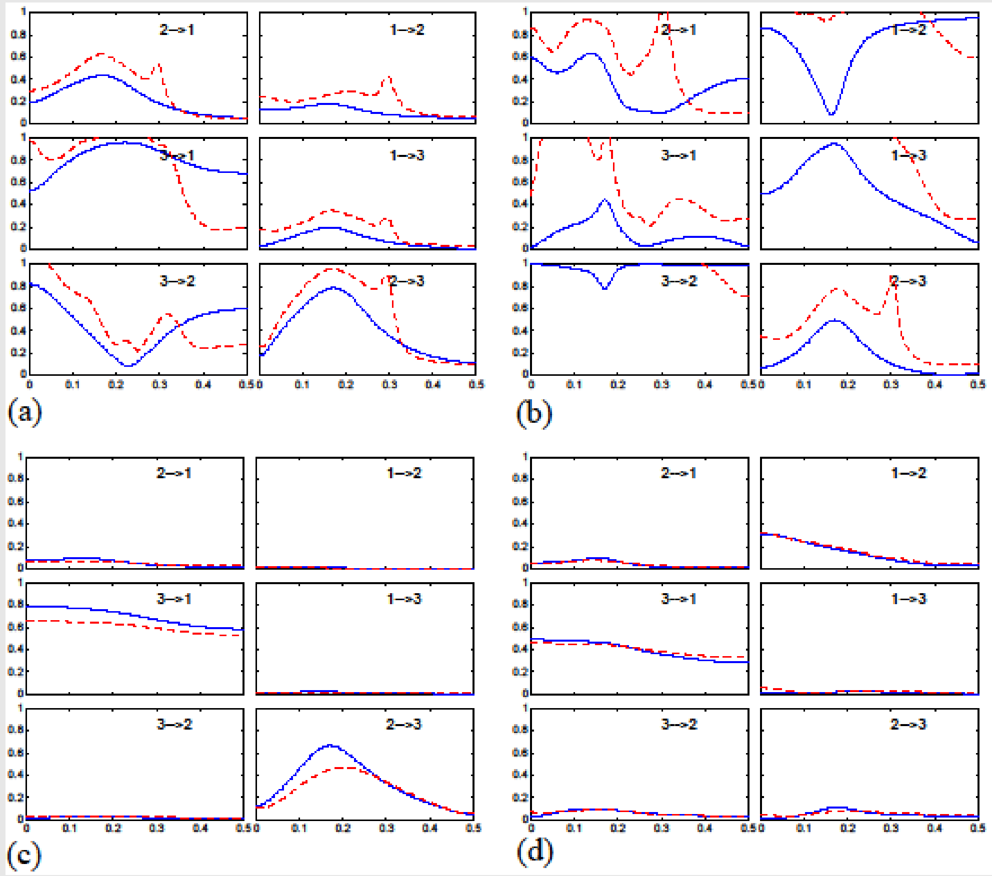
Statistical thresholding using permutation test

SIR=0.5

SIR=0.25

最小自乗法による
AR coefficient推定

Sparse Bayesによる
AR coefficient推定



- Permutation testによるstatistical thresholdはsurrogate法より厳しく，擬陽性を避ける観点ではsurrogate法の与えるthresholdよりも望ましい.

赤い点線：statistical threshold, 青い実線：causality measureの値

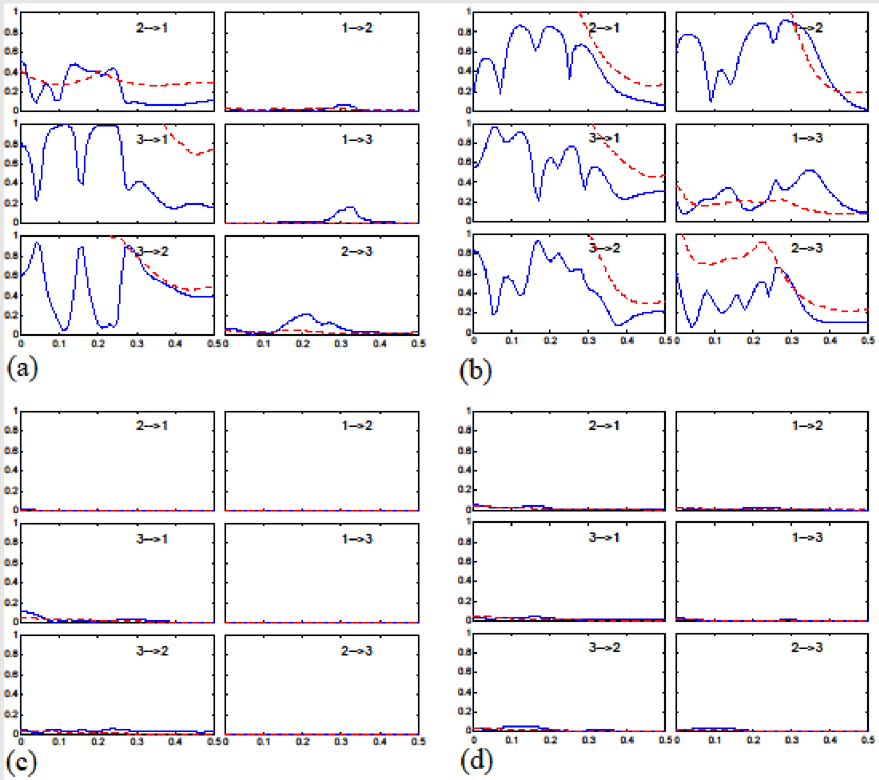
Non-interacting source scenarioでの擬陽性

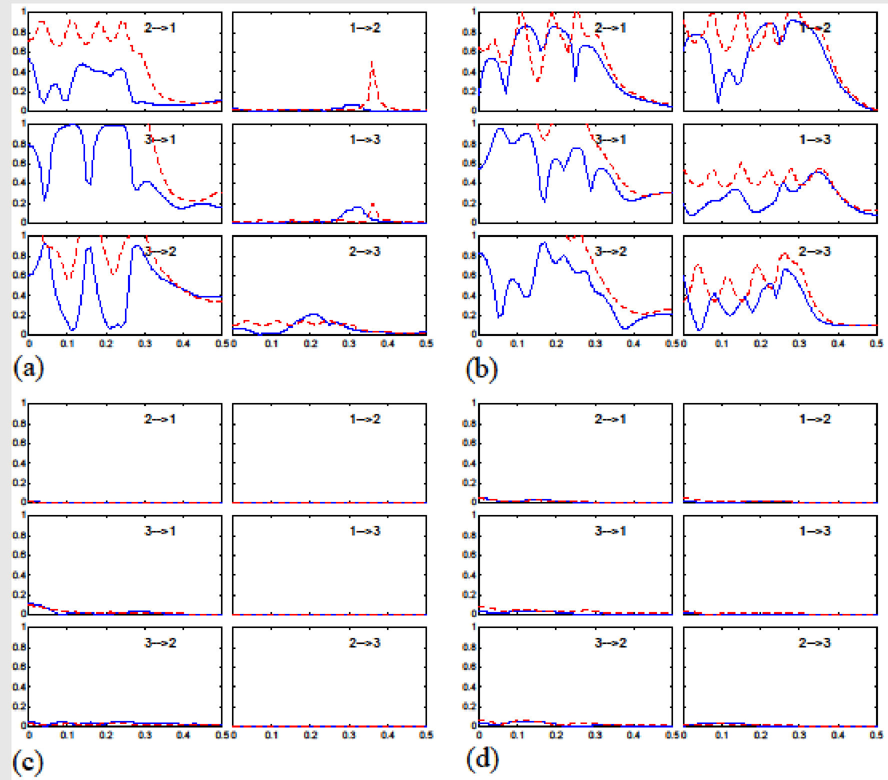
Surrogate-data bootstrapping | **permutation test**

SIR=1 | SIR=0.25 | SIR=1 | SIR=0.25

LS

SB





LS: 最小自乗法によるAR coefficient推定
SB:Sparse BayesによるAR coefficient推定

最小自乗法では大きな擬陽性を生じるが，なんらかの統計検定を行なう限り，擬陽性は「有る程度」除くことができる．

Transfer Entropy－AR modeling不要のcausality measure

２つの観測ベクトル $\boldsymbol{x}(t)$ と $\boldsymbol{y}(t)$ に対して，過去の値をつなげた以下のベクトル

$$\tilde{\boldsymbol{x}}(t) = \begin{bmatrix} \boldsymbol{x}(t-1) \\ \boldsymbol{x}(t-2) \\ \vdots \\ \boldsymbol{x}(t-P) \end{bmatrix} \qquad \tilde{\boldsymbol{y}}(t) = \begin{bmatrix} \boldsymbol{y}(t-1) \\ \boldsymbol{y}(t-2) \\ \vdots \\ \boldsymbol{y}(t-P) \end{bmatrix}$$

を用いて，2種のconditional entropyを定義する.

現在の値 $\boldsymbol{y}$ の過去の値 $\tilde{\boldsymbol{y}}$ を知った上での不確定さ.

$$H(\boldsymbol{y} \mid \tilde{\boldsymbol{y}}) = -\iint p(\boldsymbol{y}, \tilde{\boldsymbol{y}}) \log p(\boldsymbol{y} \mid \tilde{\boldsymbol{y}}) d\boldsymbol{y} d\tilde{\boldsymbol{y}}$$

現在の値 $\boldsymbol{y}$ の過去の値 $\tilde{\boldsymbol{x}}$ と $\tilde{\boldsymbol{y}}$ を知った上での不確定さ.

$$H(\boldsymbol{y} \mid \tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}}) = -\iint p(\boldsymbol{y}, \tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}}) \log p(\boldsymbol{y} \mid \tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}}) d\boldsymbol{y} d\tilde{\boldsymbol{x}} d\tilde{\boldsymbol{y}}$$

すると，以下をTransfer Entropyと呼ぶ.

$$H_{\boldsymbol{x}\to\boldsymbol{y}} = H(\boldsymbol{y} \mid \tilde{\boldsymbol{y}}) - H(\boldsymbol{y} \mid \tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}}) = -\iint p(\boldsymbol{y}, \tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}}) \log \frac{p(\boldsymbol{y} \mid \tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}})}{p(\boldsymbol{y} \mid \tilde{\boldsymbol{y}})} d\boldsymbol{y} d\tilde{\boldsymbol{x}} d\tilde{\boldsymbol{y}}$$

$\boldsymbol{x}$ の過去の値を知ることによる，$\boldsymbol{y}$ の不確定さの減少

Transfer EntropyはGranger Causalityと等価な量であることを示せる

正規性の仮定のもとで，Transfer Entropyは以下のように表せる.

$$H_{\boldsymbol{x}\to\boldsymbol{y}} = H(\boldsymbol{y} \mid \tilde{\boldsymbol{y}}) - H(\boldsymbol{y} \mid \tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}}) = \frac{1}{2}\log\frac{\left|\boldsymbol{\Sigma}_{yy} - \boldsymbol{\Sigma}_{y\tilde{y}}\boldsymbol{\Sigma}_{\tilde{y}\tilde{y}}^{-1}\boldsymbol{\Sigma}_{y\tilde{y}}^{T}\right|}{\left|\boldsymbol{\Sigma}_{yy} - \boldsymbol{\Sigma}_{y\tilde{z}}\boldsymbol{\Sigma}_{\tilde{z}\tilde{z}}^{-1}\boldsymbol{\Sigma}_{y\tilde{z}}^{T}\right|}$$

ただし，$\tilde{\boldsymbol{z}} = [\tilde{\boldsymbol{x}}^T, \tilde{\boldsymbol{y}}^T]^T$ である.

一方，Granger Causalityは，

$$\boldsymbol{y}(t) = \sum_{p=1}^{P} \boldsymbol{A}(p)\boldsymbol{y}(t-p) + \boldsymbol{e}(t) = \boldsymbol{A}\tilde{\boldsymbol{y}}(t) + \boldsymbol{e}$$

$$\boldsymbol{y}(t) = \sum_{p=1}^{P} \boldsymbol{A}(p)\boldsymbol{y}(t-p) + \sum_{p=1}^{P} \boldsymbol{B}(p)\boldsymbol{x}(t-p) + \boldsymbol{\varepsilon} = \boldsymbol{A}\tilde{\boldsymbol{y}}(t) + \boldsymbol{B}\tilde{\boldsymbol{x}}(t) + \boldsymbol{e} = \boldsymbol{C}\tilde{\boldsymbol{z}}(t) + \boldsymbol{\varepsilon}$$

とすれば，(ここで，$\boldsymbol{C} = [\boldsymbol{A}, \boldsymbol{B}]; \boldsymbol{A} = [\boldsymbol{A}(1), \ldots, \boldsymbol{A}(P)], \boldsymbol{B} = [\boldsymbol{B}(1), \ldots, \boldsymbol{B}(P)]$ とした)

C.3.3節議論より，

$$G_{\boldsymbol{x}\to\boldsymbol{y}} = \log\left[\left|\boldsymbol{\Sigma}_e^y\right| / \left|\boldsymbol{\Sigma}_\varepsilon^y\right|\right] = \log\frac{\left|\boldsymbol{\Sigma}_{yy} - \boldsymbol{\Sigma}_{y\tilde{y}}\boldsymbol{\Sigma}_{\tilde{y}\tilde{y}}^{-1}\boldsymbol{\Sigma}_{y\tilde{y}}^{T}\right|}{\left|\boldsymbol{\Sigma}_{yy} - \boldsymbol{\Sigma}_{y\tilde{z}}\boldsymbol{\Sigma}_{\tilde{z}\tilde{z}}^{-1}\boldsymbol{\Sigma}_{y\tilde{z}}^{T}\right|}$$

を示せる．つまり，Transfer EntropyとGranger Causalityは定数倍を除いて等しい.

まとめ

- causality analysisについて：
（１）脳活動のcausalな面についての知見が非常に少ない(はっきり言って無い).
（２）方法論的にも未成熟である.

というのが現状である（臨床応用に用いられるのはかなり先の話か）．

- したがって，当面はconnectivity analysisの臨床応用に傾注し， 経験の蓄積を待つという方針でいいと思う（私の個人的見解）．

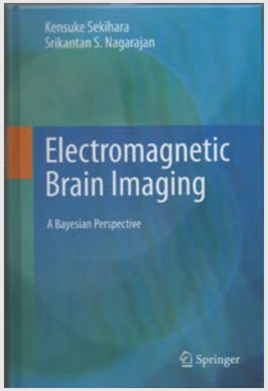
方法論的な問題点：

- Causality measureの計算に際して，関連した全てのノードが知られているという仮定が必要かどうかはっきりしない．AR coefficient行列を使うmeasure(DTFやPDC)は当然この仮定が必要と考えられ，その場合，missing nodeの問題が存在.

- この， missing nodeの問題に対するrobustnessの観点で考えると， （おそらく）Granger causalityをvoxel pairwiseで計算する方法がbestか.

Introduction to MEG Cross-Frequency Analysis

このスライド以降の部分はElectromagnetic Brain Imagingの第9章を解説したものである.



ただし,technicalなポイントのみの概略
詳細は，リコー社内の専門家（奥村・淺川）にプレゼンしてもらうべき.

脳活動における（アルファリズムなどの）周期的な活動の役割を明らかにする.

⇩

作業仮説：

脳活動には（あたかもコンピュータのベースクロックのような）基本となるリズム（多分アルファ波に近い帯域，その昔は40Hz活動と言われていた）があり，そのリズムで脳全体の活動をコントロールしている．個別の脳活動はより高い周波数（ガンマ帯域かそれ以上か）のMEGを示す．そうだとすれば，低い周波数のMEGに対応した脳活動と高い周波数のMEGに対応した脳活動には何らかの関連があるはず.

したがって，低い周波数のMEGとより高い周波数のMEGの関連を調べれば上に述べた脳活動のモデルについて知見が得られるはず.

用いられる用語

cross-frequency coupling (CFC):異なる周波数帯域の脳活動のcoupling

same-frequency coupling (SFC):同じ周波数帯域の脳活動のcoupling

phase-amplitude coupling (PAC):

低い周波数の脳信号(LF signal)の位相と高い周波数の脳信号(HF signal)の振幅にある相関関係．前述の作業仮説を調べるのに中心的な役割をはたす.

amplitude-amplitude coupling (AAC):

2つの周期的脳信号の振幅間に見られる相関関係.

phase-phase coupling (AAC):

2つの異なる周波数帯にある周期的脳信号の位相間に見られる相関関係.

当然ながら，PACが主な解析対象である．PACはさらに以下に分類される.

Local PAC (lPAC):LF signalとHF signalが同じ部位から計測される場合.

Cross-location PAC (xPAC):LF signalとHF signalが異なる部位から計測される場合.

LF信号とHF信号との間のPAC検出：PAC measure

$$\text{LF signal: } x_L(t), \quad \text{HF signal: } x_H(t)$$

Analytic signalへ変換して：

$$A[x_L(t)] = A_L(t)e^{i\theta_L(t)}, \quad A[x_H(t)] = A_H(t)e^{i\theta_H(t)}$$

Amlitude-phase diagram

LF信号の位相に対して，HF信号の振幅がどのように関連しているかを表す量．まず，位相角[0,2π]をq 分割する．

位相角を領域 $\Delta_1, \Delta_2, \ldots, \Delta_q$ に分割する．ここで，$\Delta_j = [2\pi \frac{j-1}{q}, 2\pi \frac{j}{q}]$である．

$\theta_L(t)$ が Δ_j に存在するときの $A_H(t)$ の平均を計算し，Ψ_j とする．

つまり， $\Psi_j = \left\langle A_H(t) \right\rangle_{\theta_L(t)\in\Delta_j}$ である．

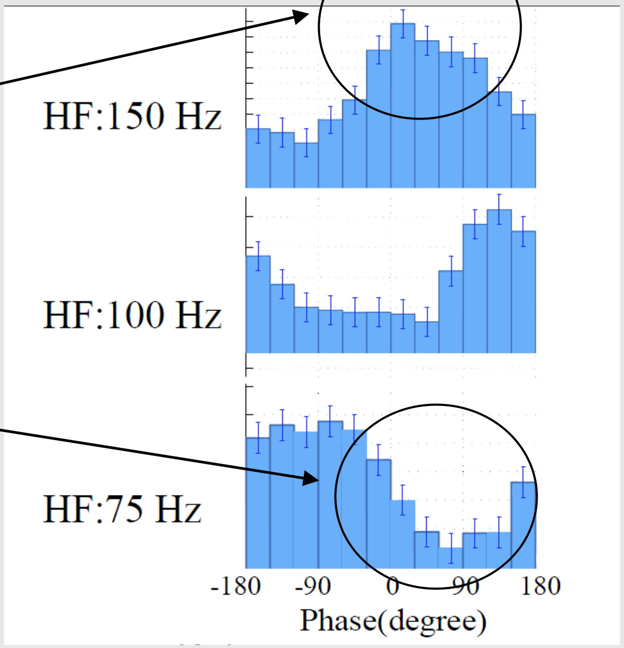
Ψ_j を全ての phase bin $\Delta_1, \ldots, \Delta_q$ に対して求めた $\Psi_1, \ldots, \Psi_q$ を amplitude-phase diagramと呼ぶ．

Amplitude-phase diagram

phase bin $\varDelta_1,\ldots,\varDelta_q$ に対して求めた $\varPsi_1,\ldots,\varPsi_q$ のプロット

実際の計測例

12.5HzのLF信号に対して，150Hz近辺のHF信号の振幅が，位相30-120度付近で有意に大きいことを示す．



LF:12.5±0.5

12.5HzのLF信号に対して，75Hz近辺のHF信号の振幅が，位相90度付近で有意に小さいことを示す．

これらの結果は，ガンマ帯の周波数を示す脳活動はアルファ帯のLF信号によってコントロールされているという解釈を示唆する．

Modulation Index (MI)

Amplitude-phase diagramを定量化したもの.

もし，LF信号とHF信号に関連が無ければ，Amplitude-phase diagramはフラットになるはず.

⇩

したがって，Amplitude-phase diagramの「フラットさ」からのずれがLFとHF信号の関連の強さを表すはず!!

フラットさからのずれは以下のようにKL-distanceを用いて定量化できる

$$\bar{\Psi}_j = \Psi_j \Big/ \sum_{i=1}^{q} \Psi_i$$ (normalized version) $$u_j = 1/q$$ (uniform distribution)

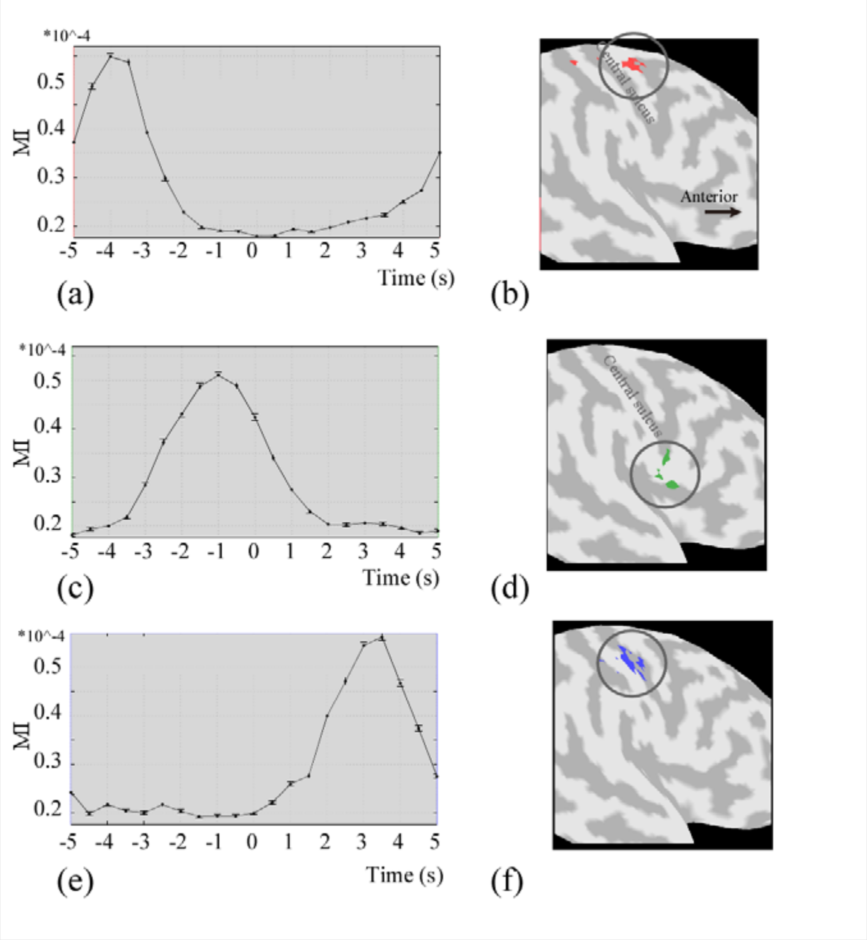
Modulation Index(MI): $$M_I = K_L[\bar{\psi}_j \,||\, u_j] = \sum_{j=1}^{q} \bar{\psi}_j \log \frac{\bar{\psi}_j}{u_j}$$

適当な時間窓でMIを計測し，時間窓をずらしてMIの時間変化 $M_I(t_1), \ldots, M_I(t_K)$ を求めればこれらはPACの時間変化を記述するものになる.

Modulation Index (MI):測定結果

「10秒ごとに左手を握る(hand-grasping data)を64回繰り返す．その間自発脳磁解を計測.

IPACの検出結果

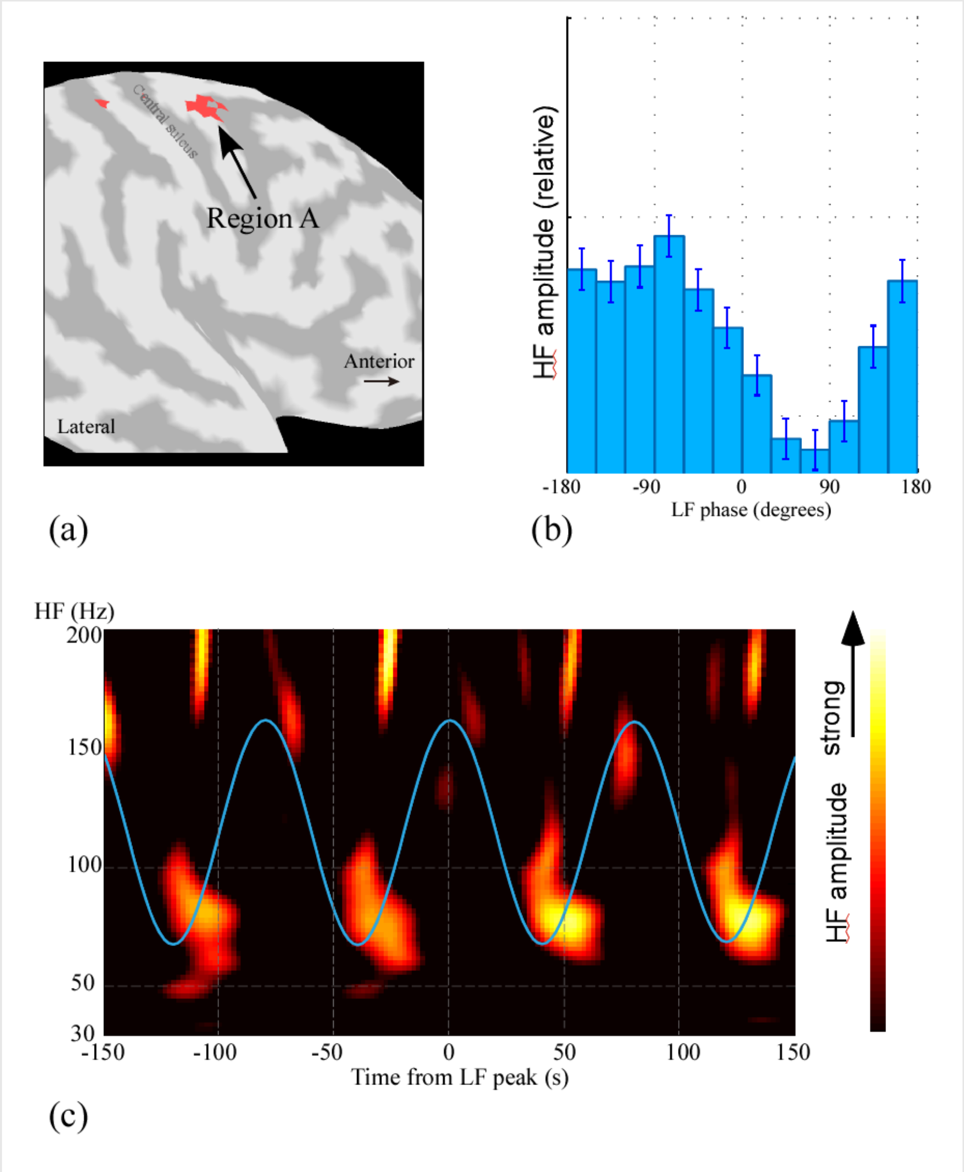


右側に統計的に有意な（surrogate法）IPACを検出したボクセルを，左側にその部位におけるMIのタイムコースを示す.

統計的に有意なIPACが見出されたのは:
(b),(f)ではLF:12.5±0.5Hz, HF:75±25Hz.
(d)ではLF:10.5±0.5, HF:75±25Hz.

(b)のボクセルに対するさらに詳細な解析結果を次に示す

hand-grasping dataにおけるIPAC



RegionAと示されたボクセルにおけるIPACのamplitude-phase diagram

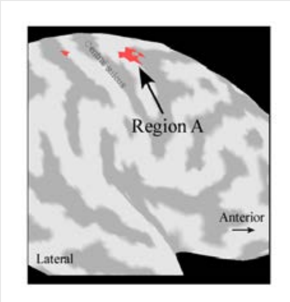
LF:12.5±0.5Hz,
HF:75±25Hz.

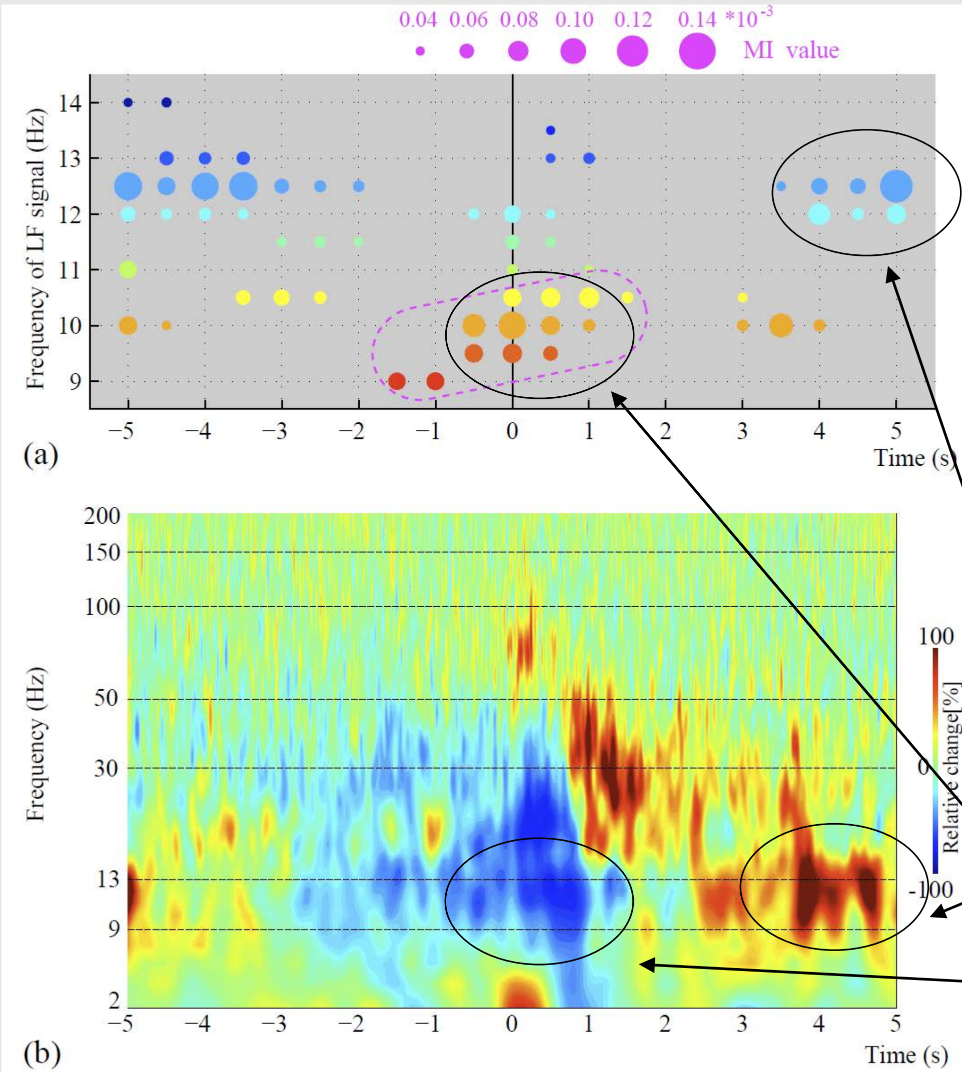
Phase-informed time-frequency map:

HF信号のTF mapにLF信号から求めた $\cos(\theta_L(t))$ をオーバーレイ

LF信号の位相とHF信号の活動強度の相関が良くわかる.

RegionAからのIPACについての追加の解析





HF信号：75±25Hzに固定

LF信号を9±0.5Hzから14±0.5Hzまで11通り変えてMIを計算しプロット.

LF信号(中心周波数10.5と12.5）において大きなMI値を持つ.

脳は２つのベースリズムを持つのか?

RegionAからのMEG信号のtime frequency map

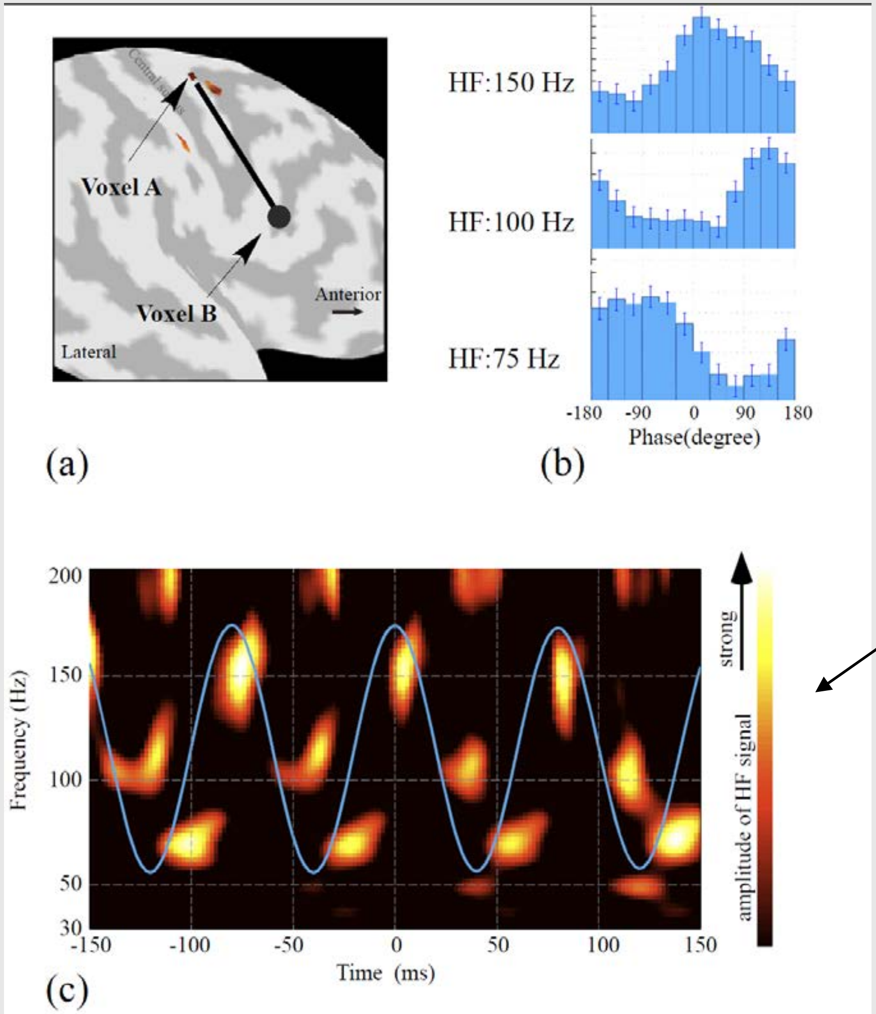
同じ時間周波数領域に存在

同じ時間周波数領域に存在

hand-grasping dataにおけるxPAC

ボクセルBにおけるLF信号とボクセルAにおけるHF信号を用いて解析

LF：12.5±0.5Hz　HF:75±25, 100±25, 125±25Hz



３種類のHF subbandにおける
amplitude-phase diagram

Phase-informed time-frequency map:

LF信号のある特別な位相でHFの活動が特に強くなることが明瞭に描出されている.