

脳波位相と正答確率の関数推定に向けたロジット変換を用いたガウス過程回帰

小野島 隆之¹・神保 雅一²

(受付 2025 年 6 月 28 日；改訂 9 月 19 日；採択 9 月 25 日)

要 旨

人間の脳活動は脳波として計測でき、その状態が課題成績に影響を与えることから、特定の周波数帯域の活動の脳波位相が知覚に関係すると考えられている。この関係を推定するには、被験者ごとに対象とする周波数帯域の脳波の位相と課題成績の関数関係を推定する必要がある。近年では、脳波位相依存刺激法などの脳波位相の状態に基づく刺激法も提案されており、脳波位相と脳機能の関係を被験者ごとに柔軟に推定できる手法が求められている。本研究では、刺激提示に対する各試行の二値応答(正答/誤答あるいは検出/未検出)を観測データとし、脳波位相を説明変数とした正答確率を柔軟に推定する枠組みを構築した。具体的には、ロジット変換をリンク関数とし、周期カーネルを用いたガウス過程回帰によって脳波位相に依存して変化する正答確率を推定するベルヌーイ・ロジットモデルを構築した。このモデルにおけるサンプルサイズの増加に伴うグラム行列の退化と計算時間の増加に対処するためにモデルを改良し、新たに二項・ロジットモデルを提案した。本研究は、脳波位相と正答確率の関数関係を推定するための2つの統計モデルを構築し、数値実験により検証したものである。

キーワード：脳波、ガウス過程回帰、ロジット変換、周期カーネル。

1. はじめに

1.1 研究背景

人間の脳活動はしばしば周期的な性質を持ち、頭皮脳波(electroencephalogram: EEG)として非侵襲的に計測することが可能である。この周期的な脳活動は外部刺激や異なる脳領域の活動と同期するとされ、脳機能において重要な神経基盤として議論されている。ここでいう脳波位相とは、特定の周波数帯域の波形に対する瞬間的な位相を指し、刺激提示のタイミングとの関係で知覚や反応に影響を及ぼすと考えられている。

近年では、特定の脳波の位相がある値のときにのみ刺激を与える脳波位相依存刺激法が提案されている(Zrenner et al., 2016)。特に経頭蓋磁気刺激法(Transcranial Magnetic Stimulation: TMS)と組み合わせた場合には、刺激に対する応答が位相依存的に変化することが報告されている(Zrenner et al., 2018)。しかし、脳波位相依存刺激法における応答の位相依存性や最適な位相には個人差があり(Torrecillos et al., 2020; Zrenner et al., 2023; Kirchoff et al., 2024)、対

¹ 滋賀大学 データサイエンス・AI イノベーション研究推進センター：〒522-8522 滋賀県彦根市馬場1丁目1-1; takayuki-onojima@biwako.shiga-u.ac.jp

² 統計数理研究所：〒190-8562 東京都立川市緑町10-3; jimbo@ism.ac.jp

象とする脳波やその位相の状態を事前を選択する方法が実験を遂行する上で課題となっている。このような課題に対し、TMS研究などではベイズ最適化を用いて少数試行から最適な刺激位相を推定する枠組みが提案されている (Tervo et al., 2020, 2022; Kirchoff et al., 2024)。ベイズ最適化は、目的関数の不確実性推定を活用して効率的に次の刺激条件を選択できるため、試行数の限られる行動実験や脳刺激実験で有効である。

一方で、視覚刺激を脳波位相に基づいて提示する実験課題 (Shakeel et al., 2021) や、入眠過程において残存する α 帯域 (8–12Hz) の脳波位相をリアルタイムで検出し、そのタイミングに合わせて音刺激を与える刺激法 (Hebron et al., 2024) も報告されている。これらの視覚・聴覚の感覚モダリティでは、刺激提示直前の脳波位相が課題成績に影響を与えることが報告されており (Mathewson et al., 2009; Busch and VanRullen, 2010; Ng et al., 2012; Zoefel and VanRullen, 2015; Onojima et al., 2017)、脳波と認知機能の関係解明において瞬時位相の状態が重要であることが示唆されている。

このように脳波位相と課題成績の関係や、脳波位相を基にした刺激法などが提案されてきたが、被験者ごとの脳波位相と課題成績の関数関係は不明瞭なままである。特に視覚や聴覚などの感覚刺激を用いた行動実験において、課題成績と脳波位相の関数関係を推定する手法は十分に確立されていない。したがって、この関数を推定するとともに、その信用区間を求めることが重要である。

1.2 本研究の目的

本研究の目的は、脳波位相と課題の応答の関係を柔軟に推定できる統計モデルの構築である。ここで対象とする課題は Yes/No 課題や二者択一課題 (two alternative forced choice task: 2AFC) などの 2 値応答を伴うものであり、各試行において脳波位相 $x \in (-\pi, \pi]$ と、正答 $z(x) = 1$ または誤答 $z(x) = 0$ が観測される。このような課題における正答確率 $p(x)$ の形状は被験者や実験条件によって異なる可能性があるため、特定の関数形を仮定せずに柔軟に表現できるノンパラメトリックな手法が望ましい。

そこで本稿では、ロジット変換をリンク関数とし、周期カーネルを用いたガウス過程回帰 (Logit-Linked Periodic Gaussian Process Regression) によって $p(x)$ を推定する枠組みを構築する (ベルヌーイ・ロジットモデル)。また、この推定手法を用いる際にサンプルサイズの増加に伴って起こるグラム行列の退化と計算時間の増加に対する対策として、新たに二項・ロジットモデルを提案する。本研究は、脳波位相と課題成績の関数関係を推定するための 2 つの統計モデルを構築し、数値実験により比較した。

2. 統計モデルと推定手法

初めに、本研究で推定する関数について説明する。ここでの脳波位相とは、計測された脳波から特定の周波数帯域を抽出し、ヒルベルト変換などを用いて求めた瞬時位相であり、 $(-\pi, \pi]$ の範囲の値を持つ (図 1a 下段)。被験者に課題を提示する場合、刺激直前の脳波位相の状態が課題成績に影響を与えることが報告されている (Mathewson et al., 2009; Busch and VanRullen, 2010; Ng et al., 2012; Zoefel and VanRullen, 2015; Onojima et al., 2017)。これらの研究では脳波計測を行いながら視覚刺激 (ドットや図形) や聴覚刺激 (ビーブ音や音声) を被験者に提示し、その刺激を検出できたか、または、正しく判別できたのかを二値応答として計測している。その上で実験後に刺激提示前の脳波の位相が課題成績に影響を与えるのかを検証している。本研究では、脳波位相と課題の二値応答の正答確率の関係を被験者ごとに関数として推定することを目的とする (図 1b)。

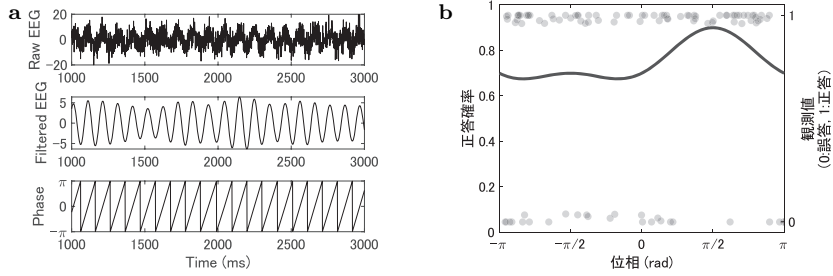


図 1. 瞬時位相と推定する関数の概念図：(a)生の EEG 信号(Raw EEG, 上段), 帯域通過フィルタ後の EEG 信号(Filtered EEG, 中段), およびヒルベルト変換によって推定された瞬時位相(Phase, 下段)の例. (b)脳波位相と正答確率の関数. 位相に依存して変動する正答確率(実線)を, 観測値(灰色の点)から推定することを想定している. この正答確率の関数形は, 3章で説明する数値実験で用いるものと同じであり, シミュレーションにより作成したものである. 観測値は描画時の重なりを避けるため, 縦方向に小さなジッターを加えて表示している. 左軸は正答確率(未知), 右軸は 0/1 の観測値を表す.

本研究が対象とする課題は, 二者択一課題(2AFC)や Yes/No 課題, 検出課題などの二値応答を伴うものであり, 各試行において脳波位相 $x \in (-\pi, \pi]$ と, 正答 $z(x) = 1$ または誤答 $z(x) = 0$ が観測される. Yes/No 課題や検出課題の場合は正答/誤答の代わりに検出/未検出の応答を伴う. この場合は正答確率ではなく検出率(hit rate)や誤報率(false alarm rate)などを推定することになるが, 本研究ではこれらの確率を包括的に正答確率として扱う. 被験者ごとに n 回の刺激を与え刺激に対する応答と刺激提示前の脳波位相を計測することを想定する. 実験試行ごとに観測される $z(x)$ は, 位相依存の正答確率 $p(x)$ に基づくと仮定する.

本稿では, ロジット変換をリンク関数とし, 周期カーネルを用いたガウス過程回帰(Logit-Linked Periodic Gaussian Process Regression)によって $p(x)$ を推定する手法を構築する. 本研究ではベイズ最適化や適応的実験設計などへの応用を目指すものであるが, 推定手法自体の構築と評価に焦点を当て, ベルヌーイ・ロジットモデルと二項・ロジットモデルを構築し, 数値実験により推定結果を比較した.

2.1 ロジット変換によるガウス過程回帰

位相 x は $\Omega = (-\pi, \pi]$ を定義域とし, 観測値 $z(x)$ は $\{0, 1\}$ の 2 値をとるとする. この $z(x)$ は位相 x ごとに $p(x)$ が与えられたもとの, 独立なベルヌーイ分布 $\text{Bin}(1, p(x))$ に従って得られる観測値であるとする. ここで $p(x)$ は位相 x ごとの正答確率であり $[0, 1]$ に値をとる連続的な周期関数とする. $p(x)$ のロジット変換

$$g(x) = \log \frac{p(x)}{1 - p(x)}$$

は Ω 上で $(-\infty, \infty)$ に値をとる定常ガウス過程 $GP(\mu_0, k(\cdot, \cdot))$ に従うとする. ただし, k はカーネル共分散関数であり, $p(x)$ と $g(x)$ は 1 対 1 に対応しており, 逆変換は

$$p(x) = \frac{e^{g(x)}}{1 + e^{g(x)}}$$

と書ける. したがって, $p(x)$ は $[0, 1]$ に値を取る確率過程であり, その見本関数が正答確率関数である. ロジット変換をリンク関数とするガウス過程回帰の統計モデルの推定・予測については Rue et al. (2009); Wang and Shi (2014); Sofro and Oktaviarina (2018) で

議論されており，本稿では $\Omega = (-\pi, \pi]$ 上で周期カーネルを用いた同様のモデルを構築し，このベルヌーイ分布とロジット変換を用いたガウス過程回帰をベルヌーイ・ロジットモデルと呼ぶこととする．位相 x_i ($i = 1, \dots, n$) での観測値 $z_i = z(x_i)$ は正答確率 $p_i = p(x_i)$ のベルヌーイ分布より得られる．また， g_i を p_i のロジット変換とする．ここで $\mathbf{z}_n = (z_1, \dots, z_n)$, $\mathbf{g}_n = (g_1, \dots, g_n)$, $\mathbf{p}_n = (p_1, \dots, p_n)$ とし，標本を $D_n = ((z_1, x_1), \dots, (z_n, x_n))$ とする．我々は任意の x に対して $p(x)$ の事後確率密度関数 $f(p(x) | D_n)$ を得たい．そのためにロジット変換 $g(x)$ の事後確率密度関数 $f(g(x) | D_n) = f(g(x) | \mathbf{z}_n)$ を求める．まず， z_i は p_i が与えられたもとで，独立にベルヌーイ分布 $\text{Bin}(1, p_i)$ に従うため，

$$f(\mathbf{z}_n | \mathbf{g}_n) = f(\mathbf{z}_n | \mathbf{p}_n) = \prod_{i=1}^n p_i^{z_i} (1 - p_i)^{1 - z_i} = \prod_{i=1}^n \frac{e^{g_i z_i}}{1 + e^{g_i}}$$

である．また， \mathbf{g}_n の事前分布は n 変量正規分布に従うことを仮定しているため，

$$f(\mathbf{g}_n) = \frac{1}{(2\pi)^{n/2} |K_n|^{1/2}} \exp \left[-\frac{1}{2} (\mathbf{g}_n - \mu_0 \mathbf{1})^\top K_n^{-1} (\mathbf{g}_n - \mu_0 \mathbf{1}) \right]$$

であり， $K_n = (k(x_i, x_j))$ の各要素はカーネル関数である．また， μ_0 は事前分布として考えられる正答確率 p_0 をロジット変換した値 $\mu_0 = \log \frac{p_0}{1 - p_0}$ である．ここでは次の周期カーネルを用いる：

$$(2.1) \quad k(x_i, x_j) = \theta_1 \exp \left[-\frac{(1 - \cos(x_i - x_j))}{\theta_2} \right] = \theta_1 \exp \left[-\frac{2 \sin^2(\frac{x_i - x_j}{2})}{\theta_2} \right]$$

θ_1, θ_2 はハイパーパラメータである．このとき， \mathbf{g}_n の事後分布は

$$f(\mathbf{g}_n | \mathbf{z}_n) = \frac{f(\mathbf{z}_n | \mathbf{g}_n) f(\mathbf{g}_n)}{f(\mathbf{z}_n)} \\ \propto \left(\prod_{i=1}^n \frac{e^{g_i z_i}}{1 + e^{g_i}} \right) \frac{1}{(2\pi)^{n/2} |K_n|^{1/2}} \exp \left[-\frac{1}{2} (\mathbf{g}_n - \mu_0 \mathbf{1})^\top K_n^{-1} (\mathbf{g}_n - \mu_0 \mathbf{1}) \right]$$

である．ここで，

$$(2.2) \quad h(\mathbf{g}_n) = \left(\prod_{i=1}^n \frac{e^{g_i z_i}}{1 + e^{g_i}} \right) \exp \left[-\frac{1}{2} (\mathbf{g}_n - \mu_0 \mathbf{1})^\top K_n^{-1} (\mathbf{g}_n - \mu_0 \mathbf{1}) \right]$$

とおくと， $f(\mathbf{g}_n | \mathbf{z}_n) \propto h(\mathbf{g}_n)$ であり， \mathbf{g}_n の事後分布は正規分布とはならない．通常，事後分布を求めるにはマルコフ連鎖モンテカルロ法 (Markov chain Monte Carlo methods: MCMC) などの手法を用いるが⁸，MCMC では状態空間の次元 n はサンプルサイズとともに増加し，計算量もそれともなって増加する (Rue et al., 2009)．また，本研究では将来的にベイズ最適化と組み合わせる際にリアルタイムで高速で関数推定を行うことを想定しているため，ラプラス近似により \mathbf{g}_n の事後分布を正規分布で近似する (ラプラス近似の精度については Rue et al., 2009; Wang and Shi, 2014 及び付録 A を参照)． $h(\mathbf{g}_n)$ の対数をとると，

$$(2.3) \quad \log h(\mathbf{g}_n) = \sum_{i=1}^n g_i z_i - \sum_{i=1}^n \log(1 + e^{g_i}) - \frac{1}{2} (\mathbf{g}_n - \mu_0 \mathbf{1})^\top K_n^{-1} (\mathbf{g}_n - \mu_0 \mathbf{1})$$

であり，(2.3) の極大点，すなわち

$$\nabla \log h(\mathbf{g}_n) = \mathbf{z}_n - \mathbf{p}_n - K_n^{-1} (\mathbf{g}_n - \mu_0 \mathbf{1}) = \mathbf{0}$$

の解を \mathbf{g}_n^* とする． \mathbf{g}_n^* のまわりでラプラス近似を行うと \mathbf{g}_n の条件付き分布は

$$f(\mathbf{g}_n | \mathbf{z}_n) \propto \exp \left[-\frac{1}{2} (\mathbf{g}_n - \mathbf{g}_n^*)^\top A_n^* (\mathbf{g}_n - \mathbf{g}_n^*) \right]$$

となり，多変量正規分布 $\mathcal{N}(\mathbf{g}_n^*, (A_n^*)^{-1})$ で近似できる．ただし， $\mathbf{g}_n^* = (g_1^*, \dots, g_n^*)$ であり，

$$p_i^* = \frac{e^{g_i^*}}{1 + e^{g_i^*}}$$

$$A_n^* = -\text{diag}(p_1^*(1-p_1^*), \dots, p_n^*(1-p_n^*)) + K_n^{-1}$$

である．

2.2 新たな x に対する $p(x)$ の推定

本稿ではベイズ最適化を用いた実験を想定し， \mathbf{z}_n が与えられたもとの，新たな x_{n+1} に対する $p_{n+1} = p(x_{n+1})$ の分布を求めたい．そのために新たな x_{n+1} に対する $g_{n+1} = g(x_{n+1})$ の分布を求める．そこで，Wang and Shi (2014) と同様に g_{n+1} の事後分布 $f(g_{n+1} | \mathbf{z}_n)$ を求め，対応する p_{n+1} の信用区間を求める．まず， $f(g_{n+1} | \mathbf{g}_n)$ は多変量正規分布 $\mathcal{N}(\mu_{n+1}, K_{n+1})$ の条件付き分布であり， $\mathcal{N}(\mu_{n+1}(\mathbf{g}_n), \sigma_{n+1}^2)$ となる．ただし，

$$\mu_{n+1}(\mathbf{g}_n) = \mathbb{E}[g_{n+1} | \mathbf{g}_n] = \mu_0 + \mathbf{k}_{n+1}^\top K_n^{-1} (\mathbf{g}_n - \mu_0 \mathbf{1})$$

$$\sigma_{n+1}^2 = \text{Var}[g_{n+1} | \mathbf{g}_n] = k(x_{n+1}, x_{n+1}) - \mathbf{k}_{n+1}^\top K_n^{-1} \mathbf{k}_{n+1}$$

$$\mathbf{k}_{n+1} = (k(x_1, x_{n+1}), \dots, k(x_n, x_{n+1}))^\top$$

である．ここで，

$$\varepsilon_1 = g_{n+1} - \mu_{n+1}(\mathbf{g}_n) = g_{n+1} - \mu_0 - \mathbf{k}_{n+1}^\top K_n^{-1} (\mathbf{g}_n - \mu_0 \mathbf{1})$$

とおくと， $\varepsilon_1 \sim \mathcal{N}(0, \sigma_{n+1}^2)$ に従う．

次に， $f(\mathbf{g}_n | \mathbf{z}_n)$ が $\mathcal{N}(\mathbf{g}_n^*, (A_n^*)^{-1})$ で近似できるので， \mathbf{z}_n が得られたもとので，

$$\varepsilon_2 = \mathbf{g}_n - \mathbf{g}_n^* \sim \mathcal{N}(\mathbf{0}, (A_n^*)^{-1})$$

したがって g_{n+1} は

$$g_{n+1} = \mu_0 + \mathbf{k}_{n+1}^\top K_n^{-1} (\mathbf{g}_n^* - \mu_0 \mathbf{1} + \varepsilon_2) + \varepsilon_1$$

と書ける．また，

$$\begin{aligned} \text{Cov}[\mathbf{g}_n, \varepsilon_1] &= \mathbb{E}[(\mathbf{g}_n - \mu_0 \mathbf{1}) \varepsilon_1] \\ &= \mathbb{E}[(\mathbf{g}_n - \mu_0 \mathbf{1})(g_{n+1} - \mu_0 - \mathbf{k}_{n+1}^\top K_n^{-1} (\mathbf{g}_n - \mu_0 \mathbf{1}))] \\ &= \mathbf{k}_{n+1} - K_n K_n^{-1} \mathbf{k}_{n+1} = \mathbf{0} \end{aligned}$$

であるから， ε_1 と \mathbf{g}_n は無相関(正規性より独立)である．したがって ε_1 と ε_2 も独立であり，また， \mathbf{z}_n と ε_1 も独立であり， $\mathbb{E}[\varepsilon_1 | \mathbf{z}_n] = \mathbb{E}[\varepsilon_1] = 0$ である．ゆえに， $\mathbb{E}[\varepsilon_2 | \mathbf{z}_n] = \mathbf{0}$ に注意すると，

$$\begin{aligned} \mathbb{E}[g_{n+1} | \mathbf{z}_n] &= \mu_0 + \mathbf{k}_{n+1}^\top K_n^{-1} (\mathbf{g}_n^* - \mu_0 \mathbf{1}) = \mu_{n+1}(\mathbf{g}_n^*) \\ \text{Var}[g_{n+1} | \mathbf{z}_n] &= \text{Var}[\mathbf{k}_{n+1}^\top K_n^{-1} \varepsilon_2 + \varepsilon_1] \\ &= \mathbf{k}_{n+1}^\top K_n^{-1} \text{Var}[\varepsilon_2] K_n^{-1} \mathbf{k}_{n+1} + \text{Var}[\varepsilon_1] \\ &= \mathbf{k}_{n+1}^\top K_n^{-1} (A_n^*)^{-1} K_n^{-1} \mathbf{k}_{n+1} + \sigma_{n+1}^2 \end{aligned}$$

となり, $f(g_{n+1} | z_n)$ は正規分布

$$\mathcal{N}(\mu_0 + \mathbf{k}_{n+1}^\top K_n^{-1} (\mathbf{g}_n^* - \mu_0 \mathbf{1}), \mathbf{k}_{n+1}^\top K_n^{-1} (A_n^*)^{-1} K_n^{-1} \mathbf{k}_{n+1} + \sigma_{n+1}^2)$$

で近似できる. これを用いて g_{n+1} の期待値と信用区間を求め, ロジット変換の逆変換によって p_{n+1} の期待値と信用区間を求める.

2.3 グラム行列の退化と正則化

本稿では, 定義域 $\Omega = (-\pi, \pi]$ 上の周期カーネルのグラム行列を考えている. 3章以降で行う数値実験では, サンプルサイズ n の位相 x を一様分布 $U(-\pi, \pi]$ に従って生成し, x_1, \dots, x_n を得ることで, グラム行列 K_n を構成する. しかし, x_i と x_j が近い値を取る場合, グラム行列が退化し, 逆行列が計算できなくなる可能性がある. 特にサンプルサイズが大きくなると, 定義域内に多くの点が集中するため, この問題が顕著になる. また, ベイズ最適化のように関数の最大値などを探索する手法では, 特定の位相 x の周辺にサンプルが集中することがあり, さらに退化のリスクが高まる. このため, 2.1 および 2.2 節で述べた手法を用いる際には, カーネルのグラム行列の退化を想定しておく必要がある.

まず, このような数値的な不安定性を評価するために, 本稿ではカーネルのグラム行列 K の 1-ノルムに基づく条件数の逆数を指標とした(付録B参照). この指標 $\kappa_1^{-1}(K)$ が小さい場合(例えば $\kappa_1^{-1}(K) < 10^{-10}$), 行列 K は数値的に特異または準特異と判断される. 図2に, 周期カーネルについて, 区間 $\Omega = (-\pi, \pi]$ におけるカーネル関数とサンプルサイズの増加に対する条件数の逆数の変化を示す. 周期カーネルの θ_2 が大きくなると x と x' が離れていても相関が大きくなる. さらに, サンプルサイズが数十から百以上になると, 数値的に不安定になる傾向が見られる(図2右参照). カーネル関数のグラム行列 K_n の退化の問題は周期カーネルのように定義域 Ω が有界区間の場合に顕著に現れると思われる. この問題に対して2つの方法でグラム行列の退化の対策を行なった.

- 正則化項を加えることによる退化の防止
- 位相 Ω のビン分けと二項分布を用いた二項・ロジットモデルの構築

まず1つ目の方法では, グラム行列の対角要素に微小な正則化項 εI を加えて退化を防ぐ:

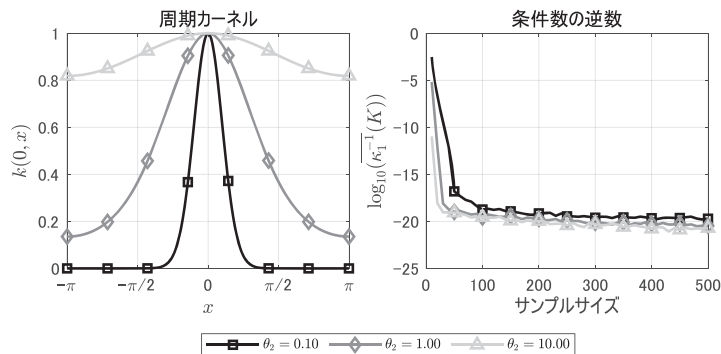


図2. ハイパーパラメータごとの周期カーネル(Circular)の相関関数とグラム行列の条件数の逆数の平均値: 左図は $\theta_1 = 1$ とし, $\theta_2 = 0.1, 1, 10$ のときの周期カーネル関数の形状を示す. 右図は各カーネル関数に対するグラム行列 K における条件数の逆数の10回の試行結果の平均値 $\overline{\kappa_1^{-1}(K)}$ の変化を示す. サンプルサイズが増加するに連れて, 条件数の逆数が小さくなる傾向が見られる.

$$K'_n = K_n + \varepsilon I \quad (\varepsilon = 10^{-6})$$

ここで I は n 次の単位行列である．この正則化により，グラム行列の逆行列が求まり， p_{n+1} の期待値および信用区間の計算が数値的に安定して行えるようになる．ベルヌーイ・ロジットモデルではこの方法でグラム行列の退化を防止している．

2つ目の方法は，次節で詳しく説明するが， Ω をいくつかのビンに分け，ビンごとの正答数が二項分布に従うと仮定することで，サンプルサイズの増加によるグラム行列の退化が発生しないモデルを構築する方法である．

2.4 z_i が二項分布に従う場合

ロジット変換をリンク関数とするガウス過程回帰を用いることで，位相と正答確率の関係を推定できるが，2.3節で述べたようにサンプルサイズが増加すると計算時間の増加とグラム行列の退化の問題がある．そこで， Ω を b 個のビン c_1, \dots, c_b に等間隔に分割する．各ビン $c_i = (c_{iL}, c_{iU}]$ に対して複数回の標本が得られる場合，(i) c_i に属する標本の平均値(あるいは中央値)をビン c_i の位相 x_i とする，あるいは，(ii) 区間中央 $\frac{c_{iL} + c_{iU}}{2}$ を位相 x_i とする．(i)の方が(ii)の代表値の決め方より推定精度が高くなることが期待される．一方，リアルタイムでの高速計算が必要な場合(ii)の方法を用いると事前に逆行列を計算しておく利点があり，本稿では代表値として(ii)の方法を採用する． m_i をその標本のサイズ，その中の正答数を z_i とし，標本を $D'_b = \{(z_1, m_1, x_1), (z_2, m_2, x_2), \dots, (z_b, m_b, x_b)\}$ とする．このとき， z_i は二項分布 $\text{Bin}(m_i, p(x_i))$ に従うため，

$$(2.4) \quad f(\mathbf{z}_b | \mathbf{g}_b) = f(\mathbf{z}_b | \mathbf{p}_b) = \prod_{i=1}^b \binom{m_i}{z_i} p_i^{z_i} (1-p_i)^{m_i-z_i} = \prod_{i=1}^b \binom{m_i}{z_i} \frac{e^{g_i z_i}}{(1+e^{g_i})^{m_i}}$$

である．ここで g_i はベルヌーイ・ロジットモデルのときと同様にガウス過程に従い，観測値 z_b が得られたもとで \mathbf{g}_b の事後分布は

$$\begin{aligned} f(\mathbf{g}_b | \mathbf{z}_b) &= \frac{f(\mathbf{z}_b | \mathbf{g}_b) f(\mathbf{g}_b)}{f(\mathbf{z}_b)} \\ &= \frac{1}{f(\mathbf{z}_b)} \left(\prod_{i=1}^b \binom{m_i}{z_i} \frac{e^{g_i z_i}}{1+e^{g_i}} \right) \frac{1}{(2\pi)^{b/2} |K_b|^{1/2}} \exp \left[-\frac{1}{2} (\mathbf{g}_b - \mu_0 \mathbf{1})^\top K_b^{-1} (\mathbf{g}_b - \mu_0 \mathbf{1}) \right] \end{aligned}$$

である． $K_b = (k(x_i, x_j))$ の各要素はカーネル関数であり，カーネル関数は周期カーネルを用いている．ここで，

$$(2.5) \quad h(\mathbf{g}_b) = \left(\prod_{i=1}^b \binom{m_i}{z_i} \frac{e^{g_i z_i}}{1+e^{g_i}} \right) \exp \left[-\frac{1}{2} (\mathbf{g}_b - \mu_0 \mathbf{1})^\top K_b^{-1} (\mathbf{g}_b - \mu_0 \mathbf{1}) \right]$$

とおくと， $f(\mathbf{g}_b | \mathbf{z}_b) \propto h(\mathbf{g}_b)$ となる．この \mathbf{g}_b をベルヌーイ・ロジットモデルの時と同様に，ラプラス近似により正規分布で近似する． $h(\mathbf{g}_b)$ の対数

$$(2.6) \quad \log h(\mathbf{g}_b) = \sum_{i=1}^b g_i z_i - \sum_{i=1}^b m_i \log(1+e^{g_i}) - \frac{1}{2} (\mathbf{g}_b - \mu_0 \mathbf{1})^\top K_b^{-1} (\mathbf{g}_b - \mu_0 \mathbf{1})$$

の極大点，すなわち

$$\nabla \log h(\mathbf{g}_b) = \mathbf{z}_b - M_b \mathbf{p}_b - K_b^{-1} (\mathbf{g}_b - \mu_0 \mathbf{1}) = \mathbf{0}$$

の解を \mathbf{g}_b^{**} とする．なお， $M_b = \text{diag}(m_1, \dots, m_b)$ とする． \mathbf{g}_b^{**} のまわりでラプラス近似を行う

と g_b の事後分布は

$$f(g_b | z_b) \propto \exp \left[-\frac{1}{2} (g_b - g_b^{**})^\top A_b^{**} (g_b - g_b^{**}) \right]$$

となり、多変量正規分布 $\mathcal{N}(g_b^{**}, (A_b^{**})^{-1})$ で近似できる。ただし、

$$A_b^{**} = -\text{diag}(m_1 p_1^{**}(1 - p_1^{**}), \dots, m_b p_b^{**}(1 - p_b^{**})) + K_b^{-1}$$

である。この事後分布を用いることで、新たな x_{n+1} に対する $p_{n+1} = p(x_{n+1})$ を 2.2 節と同様の手順で推定することが可能であり、 $f(g_{n+1} | z_b)$ は

$$f(g_{n+1} | z_b) \sim \mathcal{N}(\mu_0 + \mathbf{k}_{n+1}^\top K_b^{-1} (g_b^{**} - \mu_0 \mathbf{1}), \mathbf{k}_{n+1}^\top K_b^{-1} (A_b^{**})^{-1} K_b^{-1} \mathbf{k}_{n+1} + \sigma_{n+1}^2)$$

である。ただし

$$\begin{aligned} \mathbf{k}_{n+1} &= (k(x_1, x_{n+1}), \dots, k(x_b, x_{n+1}))^\top \\ \sigma_{n+1}^2 &= \text{Var}[g_{n+1} | g_b] = k(x_{n+1}, x_{n+1}) - \mathbf{k}_{n+1}^\top K_b^{-1} \mathbf{k}_{n+1} \end{aligned}$$

である。この g_{n+1} の事後分布から、 p_{n+1} の期待値と信用区間を推定することが可能である。この手法を二項・ロジットモデルと呼ぶこととする。この場合はサンプルサイズによらずカーネルのグラム行列 K_b は一定であり、計算量の増加や退化の懸念は生じない。ただし、 b を大きく設定すると K_b が退化する可能性がある。これを避けるためには周期カーネルのパラメータ θ_2 を小さく取る必要がある。例えば、各ビンの代表値を区間の中央とする場合、 $b = 12$ のときは $\theta_2 = 1$ でも退化しないが、 $b = 30$ のときは $\theta_2 = 1$ だと退化し、 $\theta_2 = 0.25$ だと退化しない。また、各ビンの代表値を平均値とする場合も同様の結果が得られる。

3. 数値実験

本研究では、ベルヌーイ・ロジットモデルを構築し、2.3 節(図 2)においてグラム行列が退化しやすいことを確認した。その対処として二項・ロジットモデルを新たに構築した。3 章では、これら 2 つのモデルの推定手法を比較・検証するための数値実験について説明する。数値実験の目的は、(i) 提案手法により関数 $p(x)$ を推定することができるかを確認すること、(ii) サンプルサイズの増加に伴う推定結果の変化を明らかにすること、(iii) g_{n+1} の期待値および 95% 信用区間から導かれる $p_{n+1} = p(x_{n+1})$ の信用区間を両モデルで比較することである。検証の流れは以下の通りである。まず、4 章では両モデルにおけるハイパーパラメータの選択方法を説明し、予備実験を想定した数値実験を行い、ハイパーパラメータの選択を行う。次に、5 章で選択したハイパーパラメータを用いてサンプルサイズ $n = 5, 50, 500, 5000$ のデータに対して推定を行い、結果を例示する。ここではサンプルサイズ $n = 500$ が本実験で想定しているデータサイズであり、 $n = 5$ は極端に小さい場合、 $n = 5000$ は極端に大きい場合を想定している。最後に、サンプルサイズの増加に伴う両モデルの推定性能の違いを比較する。

次の条件で数値実験を行い、 g_{n+1} の期待値と 95% 信用区間を推定し対応する $p_{n+1} = p(x_{n+1})$ の信用区間を求めることで推定手法を検証した。まず、サンプルサイズを n とし、一様分布 $U(-\pi, \pi]$ に従って x_1, \dots, x_n を生成して、真の正答確率 $p(x)$ を次のように設定した(図 1b)：

$$p(x) = p_0 + a_1 \sin x + a_2 \cos 2x$$

ここで $a_1 = 0.1$, $a_2 = 0.05$, $p_0 = 0.75$ とした。この p_0 は課題に対する正答確率が 0.75 であることを仮定し、その周りで脳波位相により正答確率が関数として変化していることを想定している。この正答確率は実際の実験課題の難易度に相当しており、予備実験等で事前に正答

確率 p_0 が 0.75 もしくは 0.5 になるように刺激の強度や難易度を調整していることを想定する．本研究では $p_0 = 0.75$ とし事前分布のパラメータ $\mu_0 = \log \frac{p_0}{1-p_0} \simeq 1.099$ としている．この $p(x)$ に基づき，ベルヌーイ分布 $\text{Bin}(1, p(x_i))$ から 0, 1 の値をとる z_i を生成し，データセット $D_n = \{(z_1, x_1), (z_2, x_2), \dots, (z_n, x_n)\}$ を生成した．二項・ロジットモデルの場合はこれらの (z_i, x_i) を $b = 12$ 個のビンに分類して $D'_b = \{(z'_1, m_1, x'_1), (z'_2, m_2, x'_2), \dots, (z'_{12}, m_{12}, x'_{12})\}$ とした．なお各ビンの代表値は区間中央 $x'_i = \frac{c_{iL} + c_{iU}}{2}$ とした．

4. ハイパーパラメータの選択

ベルヌーイ・ロジットモデルと二項・ロジットモデルはハイパーパラメータとして θ_1 と θ_2 を設定する必要がある．なお，付録 A にあるように θ_1 が大きくなるとラプラス近似の精度が悪くなるため，本研究では $\theta_1 \leq 1$ とした．また， θ_2 が大きくなるとグラム行列が退化する確率が大きくなる (図 2)．二項・ロジットモデルでは退化を避けるために $b = 12$ を考慮して $\theta_2 \leq 2$ とした．ベルヌーイ・ロジットモデルも二項ロジットモデルとの比較のために $\theta_2 \leq 2$ とした．本稿では交差検証と経験ベイズ法により，それぞれのモデルのハイパーパラメータを選択した．

4.1 交差検証による選択

d -fold 交差検証によりハイパーパラメータを選択するために誤差関数として次のような平均二乗誤差を用いた．

$$\text{CV-MSE} = \frac{1}{d} \sum_{k=1}^d \left\{ \frac{1}{N_{vld}^{(k)}} \sum_{i=1}^{N_{vld}^{(k)}} (z_i^{(k)} - \hat{p}^{(k)}(x_i^{(k)}))^2 \right\}$$

d はデータを分割するグループ数である． $d = 5$ とし，データを 5 分割したうちの 1 グループを検証データ，残りの 4 グループを訓練データとする． k を検証データのグループ番号とし，そのサンプルサイズを $N_{vld}^{(k)}$ とする． $z_i^{(k)}$ が検証データの観測値， $x_i^{(k)}$ が観測点である．訓練データで求めたモデルから計算された検証データの観測点 $x_i^{(k)}$ の正答確率を $\hat{p}^{(k)}(x_i^{(k)})$ とする．この正答確率と $z_i^{(k)}$ の平均二乗誤差を用いて 5-fold 交差検証を行った．本シミュレーションでは，まず，交差検証に用いるデータとしてサンプルサイズ $n = 200$ のデータを生成し，交差検証によるハイパーパラメータの選択を行った．なお，本研究では本実験で 500 程度のサンプルサイズを得ることを想定し，予備実験として 200 程度のサンプルサイズのデータを用意できることを想定している．

図 3a 左はベルヌーイ・ロジットモデルにおける交差検証によるハイパーパラメータ選択の結果である．ハイパーパラメータは $\theta_1 = [0.05, 1]$ と $\theta_2 = [0.05, 2]$ の範囲で 0.05 刻みで CV-MSE を計算し，値が最小のパラメータ $\theta_1 = 0.3$ ， $\theta_2 = 0.2$ を選択した．次に新たに $n = 200$ のデータを生成し，推定されたハイパーパラメータを用いて正答率の推定を行った (図 3a 右)．ベルヌーイ・ロジットモデルと同じデータセットを用いて二項・ロジットモデルにおける交差検証によるパラメータ選択をした結果が図 3b 左である．二項・ロジットモデルにおいてもほぼ同じようなパラメータ $\theta_1 = 0.3$ ， $\theta_2 = 0.15$ が選択された．このように予備実験等でハイパーパラメータを選択のためのデータセットを用意することで交差検証によりハイパーパラメータを選択することができる．

4.2 経験ベイズ法による選択

d -fold 交差検証を用いればハイパーパラメータの選択は可能であるが， d 回の再学習が必要であり，ある程度の計算時間が必要である．また， d 分割するため，パラメータ選択のために

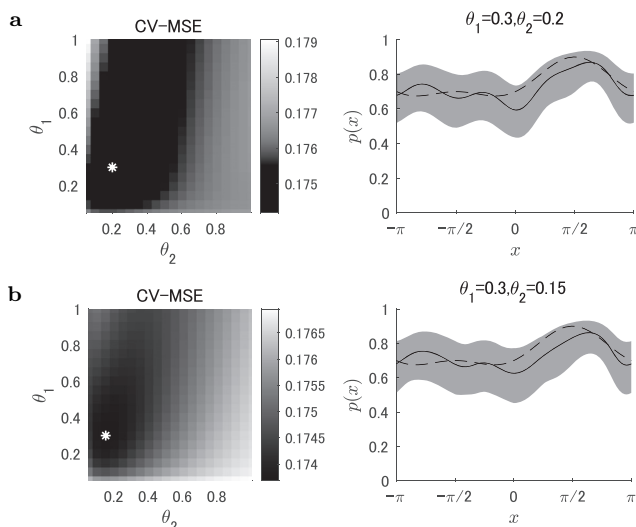


図3. 交差検証によるハイパーパラメータの選択：(a)ベルヌーイ・ロジットモデル．左図は θ_1 および θ_2 (≤ 1) ごとに交差検証で計算した平均二乗誤差を示す．白い星印は、このマップ上での最小値を表す ($\theta_1 = 0.3, \theta_2 = 0.2$)．右図は推定された正答率の平均 (実線) と 95% 信用区間 (灰色帯) を示し、破線は真の正答率を示す．(b)二項ロジットモデル．平均二乗誤差が最小のときのハイパーパラメータは $\theta_1 = 0.3, \theta_2 = 0.15$ ．左図と右図の構成は (a) と同様．

用意したすべてのデータを用いて学習と検証が行えない．本研究では将来的にベイズ最適化により少ないサンプルサイズで関数推定を行うこと目指しており、ハイパーパラメータの選択のために十分なサンプルサイズのデータを用意できないことも想定される．そこで経験ベイズ法を用いてハイパーパラメータ $\theta = (\theta_1, \theta_2)$ の選択を行う．この方法では次の周辺尤度

$$f(z|\theta) = \int f(z_n, g_n|\theta) dg_n = \int f(z_n|g_n, \theta) f(g_n|\theta) dg_n$$

を考え、この関数の値が最大となる θ を選択する．ベルヌーイ・ロジットモデルにおいてラプラス近似を用いると

$$\begin{aligned} f(z_n|\theta) &= \int f(z_n|g_n, \theta) f(g_n|\theta) dg_n \\ &= \int \prod_{i=1}^n \frac{e^{z_i g_i}}{1 + e^{g_i}} \frac{1}{\sqrt{2\pi^n |K_n|^{1/2}}} \exp\left[-\frac{1}{2} (g_n - \mu_0 \mathbf{1})^\top K_n^{-1} (g_n - \mu_0 \mathbf{1})\right] dg_n \\ &= \frac{1}{\sqrt{2\pi^n |K_n|^{1/2}}} \int h(g_n) g_n = \frac{1}{\sqrt{2\pi^n |K_n|^{1/2}}} h(g_n^*) \frac{(\sqrt{2\pi})^n}{|A_n^*|^{1/2}} \\ &= \frac{h(g_n^*)}{|K_n A_n^*|^{1/2}} \end{aligned}$$

となる．なお、二項・ロジットの場合も対応する $h(g_b^{**})$ に置き換わるだけである．ここでベルヌーイ・ロジットモデルの式(2.2)と二項・ロジットモデルの式(2.5)に g_n^* と g_b^{**} を代入した $h(g_n^*)$ と $h(g_b^{**})$ を用いると、ベルヌーイ・ロジットモデルの対数周辺尤度は

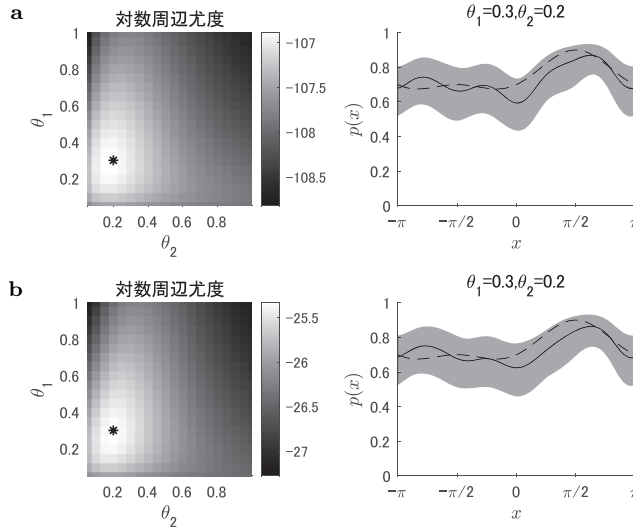


図 4. 経験ベイズ法によるハイパーパラメータの選択：(a)ベルヌーイ・ロジットモデル．左図は θ_1 および θ_2 (≤ 1) ごとの対数周辺尤度を示す．黒の星印はこのマップ上での最大値を表す ($\theta_1 = 0.3$, $\theta_2 = 0.2$)．右図は推定された正答率の平均(実線)と 95% 信用区間(灰色帯)を示し、破線は真の正答率を示す．(b)二項ロジットモデル．対数周辺尤度が最大のときのハイパーパラメータは $\theta_1 = 0.3$, $\theta_2 = 0.2$ ．左図と右図の構成は (a)と同様．

$$\begin{aligned}
 \log f(z_n|\boldsymbol{\theta}) &= \sum_{i=1}^n [g_i^* z_i - \log(1 + e^{g_i^*})] - \frac{1}{2} (\mathbf{g}_n^* - \boldsymbol{\mu}_0)^\top K_n^{-1} (\mathbf{g}_n^* - \boldsymbol{\mu}_0) \\
 &\quad - \frac{1}{2} \log |K_n A_n^*|
 \end{aligned}$$

であり、二項・ロジットモデルの対数周辺尤度は

$$\begin{aligned}
 \log f(z_b|\boldsymbol{\theta}) &= \sum_{i=1}^b \left[\log \binom{m_i}{z_i} + g_i^{**} z_i - m_i \log(1 + e^{g_i^{**}}) \right] \\
 &\quad - \frac{1}{2} (\mathbf{g}_b^{**} - \boldsymbol{\mu}_0)^\top K_b^{-1} (\mathbf{g}_b^{**} - \boldsymbol{\mu}_0) \\
 &\quad - \frac{1}{2} \log |K_b A_b^{**}|
 \end{aligned}$$

である．これらの対数周辺尤度が最大となるハイパーパラメータを選択し同様に推定を行った結果が図 4 である．経験ベイズ法でも交差検証のときに用いた $n = 200$ のデータに対して $\theta_1 = [0.05, 1]$ と $\theta_2 = [0.05, 2]$ の範囲で 0.05 刻みで対数周辺尤度の計算を行い、値が最大となるハイパーパラメータを選択した．ベルヌーイ・ロジットモデルでは、 $\theta_1 = 0.3$, $\theta_2 = 0.2$ が選択された(図 4a)．同様に二項・ロジットモデルでも偶然ではあるが同一の値 $\theta_1 = 0.3$, $\theta_2 = 0.2$ が選択された(図 4b)．5 章では、両モデルにおいて比較のためハイパーパラメータをこれらの値に固定してサンプルサイズの増加に伴う変化の比較を行った．本稿では 1 つのデータセットを用いてハイパーパラメータを選択したが、データセットに依存して選択されるハイパーパラメータは異なる可能性がある．実際に、 $n = 200$ のデータを 100 回生成し、経験ベイズ法によるハイパーパラメータの選択を行った結果、ベルヌーイ・ロジットモデルでは θ_1 の平均 0.31

(標準偏差 0.26), θ_2 の平均 0.55 (標準偏差 0.47), 二項・ロジットモデルでは θ_1 の平均 0.28 (標準偏差 0.23), θ_2 の平均 0.56 (標準偏差 0.48) となり, データごとに一定のばらつきがみられた. 実際の脳波データを用いた推定を行う場合, 検証用データや予備実験から得られたデータに基づき, ハイパーパラメータを事前に選択する必要がある.

5. 推定結果

サンプルサイズ n が, 5, 50, 500, 5000 のデータセットに対してベルヌーイ・ロジットモデル (図 5) と二項・ロジットモデル (図 6) を用いた推定を実施した. 実際の実験ではサンプルサイズ n は数十から数百程度であり $n = 5000$ とすることは困難であると思われるが, 推定精度の確認のために $n = 5000$ も実施した. ベルヌーイ・ロジットモデルの場合は, $\kappa_1^{-1}(K) < 10^{-10}$ を満たすときに特異行列とみなし, サンプルサイズが $n = 5000$ の場合においても行列が退化しないように $\varepsilon = 10^{-6}$ と設定している. $\varepsilon = 0$ の場合のグラム行列の条件数の逆数 $\kappa_1^{-1}(K)$ は $n = 50$ で $\kappa_1^{-1}(K_{50}) = 4.5891 \times 10^{-19} < 10^{-10}$ となり特異行列になる. 一方, $\varepsilon = 10^{-6}$ のグラム行列の条件数の逆数は $n = 5000$ でも $(K'_{5000}) = 1.1827 \times 10^{-9} > 10^{-10}$ となりグラム行列が退化しない. ただし, サンプルサイズが大きくなるにつれて, K_n に εI を加えても κ_1^{-1} は徐々に小さくなる.

どちらの結果もサンプルサイズが $n = 5$ の場合は真の関数の値 (破線) と期待値の値 (実線) のずれが大きいが, 95% 信用区間は広がっており, その中におおむね真の関数を含む結果となった. サンプルサイズが大きくなるにつれて, 期待値は真の値に近づき $n = 5000$ のときにはおおむね一致しておりまた信用区間も狭くなっているのがわかる.

二項・ロジットモデルの場合はピンのサイズを 12 に固定しているため, サンプルサイズが大きくなってもグラム行列のサイズも値も変化しない. そのため, サンプルサイズの増加に伴う行列の退化が起きなかった. 実際, グラム行列の条件数の逆数 $\kappa_1^{-1}(K)$ は $\kappa_1^{-1}(K_{12}) = 0.0582$ である. 図 5 と図 6 は同じデータセットをもとに計算しているため, ベルヌーイ・ロジットモデルと二項・ロジットモデルのいずれの場合も, 期待値と信用区間の推定の結果がおおむね一

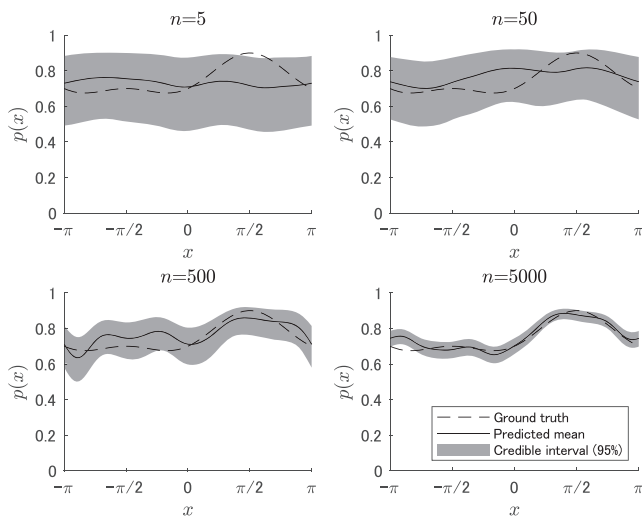


図 5. サンプルサイズごとのベルヌーイ・ロジットモデルによる推定例: 実線は推定された正答率の平均と灰色帯は 95% 信用区間を示し, 破線は真の正答率を示す.

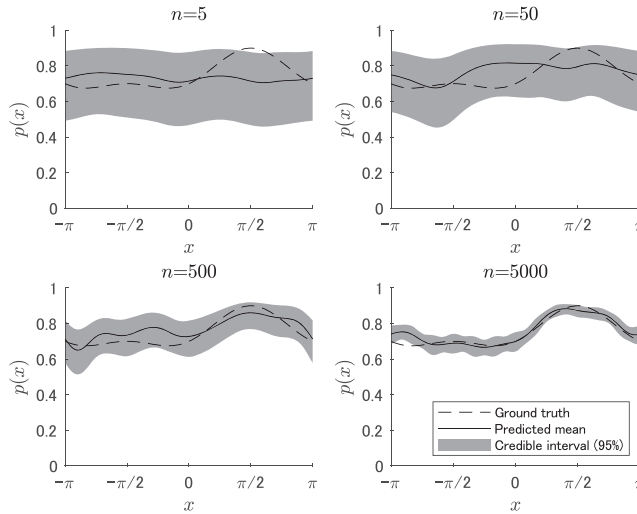


図 6. サンプルサイズごとの二項・ロジットモデルによる推定例：実線は推定された正答率の平均と灰色帯は 95% 信用区間を示し、破線は真の正答率を示す。

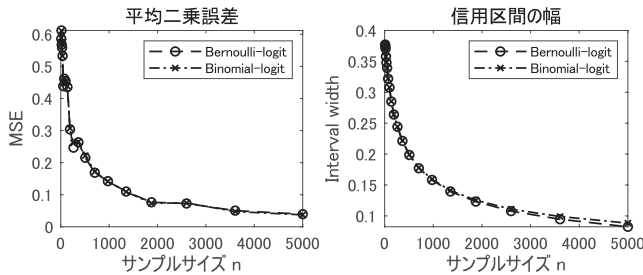


図 7. サンプルサイズの増加に伴う平均二乗誤差と信用区間の変化：左図は推定された正答率の平均と真の正答率との二乗誤差の平均を示し、右図は信用区間の幅の平均を示す (50 回)。サンプルサイズ n は 10 から 5000 の範囲で対数スケールに基づき等間隔に 20 点選び、整数に丸めた値を用いた。

致している。ただし、サンプルサイズ $n = 5000$ の結果を見るとベルヌーイ・ロジットモデルの結果と比較して二項・ロジットモデルの信用区間が揺らいでいるのがわかる。これは 12 個のビンに分けたためにデータが存在する x の値とそうでない x の値で推定値の分散が異なることが原因である。ただし、 $n \leq 500$ ではそのゆらぎが顕著ではない。

最後に、 n を増やしながらか D_n を同様の生成方法で 50 パターンずつ生成し、正答率の真の関数と推定された正答率の平均の二乗誤差の平均と信用区間の大きさの平均の変化を検証した (図 7)。ベルヌーイ・ロジットモデルの場合と二項・ロジットモデルの場合それぞれの結果を比較したがほぼ変わらなかった。ここでベルヌーイ・ロジットモデルでは特異行列になることを避けるために正則化を行っており、これが推定精度に影響を与える可能性がある。また、二項・ロジットモデルでは $b = 12$ としてビン分けを行っており、これが推定精度に影響を与える可能性がある。ビン分けの影響はサンプルサイズが大きいときに信用区間に現れており、ベルヌーイ・ロジットモデルのほうが信用区間の幅が少し小さくなっている。平均二乗誤差に関してはベルヌーイ・ロジットモデルと二項・ロジットモデルはほぼ同じ結果を示した。

ただし、ベルヌーイ・ロジットモデルはサンプルサイズが大きくなるにつれてグラム行列のサイズが大きくなり計算に必要な時間も長くなるのに対して、二項・ロジットモデルはグラム行列のサイズが変化しないので行列計算の演算に必要な時間がほぼ一定である。

6. 考察

6.1 ベルヌーイ・ロジットモデルと二項・ロジットモデルの比較

本稿では、脳波位相と正答確率の関数を推定するためにベルヌーイ・ロジットモデルと二項・ロジットモデルの両者を構築した。ベルヌーイ・ロジットモデルは各試行ごとの 0/1 応答をそのままベルヌーイ分布に従うものとしてモデル化している。このモデルはサンプルサイズの増加に伴い、グラム行列が退化する可能性が高くなり、数値安定性の確保には正則化を必要としている。また、サンプルサイズが増加するにつれてグラム行列のサイズも大きくなるため、計算量も増加する。一方、二項・ロジットモデルでは位相区間を複数のビンに分割し、各ビンで得られた正答数を二項分布に従うと仮定している。これにより、サンプルサイズが増えてもグラム行列のサイズはビン数に依存して一定のままであり、計算量や退化の問題が回避できる。特にリアルタイムでの高速計算に適していると考えている。

また、本研究は将来的に脳波位相依存刺激法におけるベイズ最適化での利用を想定している。ベイズ最適化を用いた先行研究として、脳波位相と TMS による運動誘発電位 (Motor Evoked Potential: MEP) の関数推定に関するベイズ最適化の試みが報告されている (Kirchhoff et al., 2024)。彼らも周期カーネルを用いたガウス過程回帰により脳波位相と MEP の関数を推定しているが、MEP は連続値として扱えるため、通常のガウス過程回帰で実装可能であった。一方、本研究では二者択一課題などの 2 値応答を対象とするため、ロジット変換を導入した統計モデルを構築した。さらに、ベイズ最適化への応用を見据えて、MCMC では計算負荷が大きいという問題に対してラプラス近似を導入し、また、二項・ロジットモデルによるビン分けを用いることでガウス過程回帰を行う際のグラム行列の退化に対する対処と、リアルタイムでの高速計算が行えるように改良した。

6.2 推定方法と数値的安定性の課題

本研究で用いたロジット変換をリンク関数としたガウス過程回帰ではラプラス近似により正規分布として事後分布を求めている。推定したいモデルが解析的に求めることが難しい場合、MCMC などの手法により推定する方法もあるが、ベルヌーイ・ロジットモデルでは状態空間の次元 n はサンプルサイズとともに増加し、計算量もそれに伴って増加する (Rue et al., 2009)。また、二項・ロジットモデルでも将来的にベイズ最適化と組み合わせる際にリアルタイムで高速で関数推定を行うことを想定しており、計算時間が大きくなる MCMC によらない手法として構築した。ただし、ハイパーパラメータである θ_1 が大きいとラプラス近似による正規近似の精度に注意が必要である。本研究の数値実験の設定では、交差検証と経験ベイズ法のどちらの結果も θ_1 が 1 以下の値を選択しており、この範囲においては大きな問題は生じないが、 θ_1 に 1 以上の値を設定したい場合は正規近似の近似精度が低下する可能性がある (付録 A 参照)。

さらに、両モデルは θ_2 が大きいとグラム行列が退化しやすくなる問題がある。このパラメータを小さく設定したとしてもベルヌーイ・ロジットモデルの場合、グラム行列はサンプルサイズの増加に伴い退化する確率が高くなるため ε を設定する必要がある。またベイズ最適化に用いる場合は獲得関数により、意図的に近い観測点をサンプリングする可能性があり、より退化する確率が高くなる。リアルタイムでの高速計算が必要な場合にも、サンプルサイズに依存して行列のサイズが大きくなることは好ましくなく実装上の障害となる。本稿では、この問題に

対してビン分けと二項分布を仮定した二項・ロジットモデルを用いることにより、サンプルサイズが増加しても、グラム行列のサイズが一定であるモデルを構築することで対処している。二項・ロジットモデルはグラム行列のサイズがサンプルサイズに対して一定であるため、計算時間も変化しないことからリアルタイムで実装する適応的な手法に適していると考えている。また、ビン分けを等間隔とし、区間中央を代表値とする場合、グラム行列が対称な循環行列となり、ビンの数が多い場合もその逆行列の計算を高速フーリエ変換(FFT)で計算することが可能となる。

本研究の数値実験においては、位相を12のビンに分割しているにもかかわらず、二項・ロジットモデルはベルヌーイ・ロジットモデルとほぼ同様の性能を示している。平均二乗誤差に関してはサンプルサイズが5000でもほぼ一致しており、信頼区間は僅かにベルヌーイ・ロジットモデルの方が狭いが、二項・ロジットモデルも大きくは異なる(図7)。実データにおいて同様の結果が得られるかは正答確率の関数の形にもよるが、実際の被験者ごとの実験において、1つの条件の実験試行が5000を超えることは稀であると考えられる。また、本研究では脳波位相が観測されていることを前提としているが、実際には脳波位相の推定にも不確実性がありビン分けをしていない状態の位相の精度が十分に高いとは限らない。特に脳波位相依存刺激法による実験では、位相が特定の値の近傍で刺激を与える。この場合、ある範囲の位相は同じ状態とみなし刺激を与えているため、位相をビン分けした統計モデルを構築することは実験に用いる上で妥当であると考えられる。

7. おわりに

本稿では、ロジット変換をリンク関数とした周期カーネルを用いたガウス過程回帰により、位相ごとの正答確率関数を推定する手法を構築した。

本研究の統計モデルでは脳波の位相を観測点として扱ったが、この位相は必ずしも正確に求められるものではない。特に脳波位相依存刺激法を用いる場合には、リアルタイムで過去の信号から未来を予測し、現在の瞬時位相を決定する必要がある(Chen et al., 2013; Onojima and Kitajo, 2021)。より実践的なモデルを考える場合、脳波位相を確率変数として扱う必要があるが、本研究では観測点として与えられると仮定した。今後は、位相推定の不確実性を含めた統計モデルの構築や実データでの検証を進める必要がある。

さらに、本稿で構築した手法はノンパラメトリックなガウス過程回帰に基づくものであるが、今後は実データにおいてスプライン回帰などの他のノンパラメトリックな手法や、周期関数を用いたパラメトリックな手法との比較を、獲得関数を導入したベイズ最適化の枠組みにおいて検討する必要がある。

謝 辞

本研究は、統計数理研究所が主催する統計エキスパート人材育成コンソーシアムの支援を受けて実施されました。ここに深く感謝申し上げます。また、本研究はJSPS 科研費 JP24K18601, JP22K11943 の助成を受けたものです。有益なご指摘と建設的なコメントをくださった査読者の皆様に感謝申し上げます。特に、ビンの代表値の取り方について、平均値や中央値を用いることにより推定精度が向上する可能性をご指摘くださったコメントに感謝いたします。

付 録

A. ラプラス近似の精度

Rue et al. (2009)は一般の潜在ガウスモデルにおけるラプラス近似の精度の改良について議論している. Wang and Shi (2014)は彼らの結果を用いてロジスティック回帰の場合のラプラス近似に言及しており, 一般には $\log h(\mathbf{g}_n)$ 全体をラプラス近似するより, $\log f(z_i|g_i)$ を一つずつテラー展開を用いて2次近似するほうが近似精度が良くなることを述べている. しかし, 我々のベルヌーイ・ロジットモデルの場合には $\log f(z_i|g_i)$ はすべて同じ形のため全体をラプラス近似することと彼らの逐次近似法は同値である.

なお, $h(\mathbf{g}_n)$ のラプラス近似の精度はハイパーパラメータ θ_1 に依存している. まず $n = 1$ のときにそのことを確認しておこう. 本稿ではラプラス近似を用いて(2.2)式

$$(A.1) \quad h(\mathbf{g}_n) = \prod_{i=1}^n \frac{e^{z_i g_i}}{1 + e^{g_i}} \exp \left[-\frac{1}{2} (\mathbf{g}_n - \mu_0 \mathbf{1})^\top K_n^{-1} (\mathbf{g}_n - \mu_0 \mathbf{1}) \right]$$

の正規近似を考えている. この近似は, もとの非正規分布関数を次のように正規分布で近似している.

$$h(\mathbf{g}_n) \simeq h(\mathbf{g}_n^*) \exp \left[-\frac{1}{2} (\mathbf{g}_n - \mathbf{g}_n^*)^\top A_n^* (\mathbf{g}_n - \mathbf{g}_n^*) \right],$$

$$A_n^* = -H_n^* = \text{diag}(p_1^*(1 - p_1^*), \dots, p_n^*(1 - p_n^*)) + K_n^{-1}$$

\mathbf{g}_n^* は $h(\mathbf{g}_n)$ の極大点である. (A.1)式の正規近似の精度を確認するために $n = 1$ のときの近似を考える. $k(x, x) = \theta_1$ とおくと, (A.1)式は

$$h(g_1) = \frac{e^{g_1 z_1}}{1 + e^{g_1}} \exp \left[-\frac{1}{2} \frac{(g_1 - \mu_0)^2}{\theta_1} \right]$$

と書ける.

$$\nabla \log h(g_1) = z_1 - p_1 - \frac{g_1 - \mu_0}{\theta_1} = z_1 - \frac{e^{g_1}}{1 + e^{g_1}} - \frac{g_1 - \mu_0}{\theta_1} = 0$$

の解を g_1^* とすると,

$$A_1^* = p_1^*(1 - p_1^*) + \frac{1}{\theta_1} = \frac{e^{g_1^*}}{(1 + e^{g_1^*})^2} + \frac{1}{\theta_1}$$

であるから, $h(g_1)$ は $\mathcal{N}(g_1^*, (A_1^*)^{-1})$ の定数倍で近似される. この関数と $h(g_1)$ を重ねて表示すると, $\mu_0 = \log 3 (p_0 = 0.75)$, $\theta_1 = 1$, $z = 0, 1$ のとき, 図 8a, b のように, $z = 0, 1$ のいずれの場合にもほぼ一致しており, よく近似できている. しかし, $\theta_1 = k(x, x) = 4$ の時は, 図 8c, d のように, すこし近似が悪くなる. $\theta_1 = 16$ の時は, 図 8e, f のように $h(g_1)$ の歪が大きくなり, より近似が悪くなる. θ_1 が大きくなるに従い, 近似精度は悪くなる. 次に $n = 20$ とし, 観測点 x_i を $(-\pi, \pi]$ で等間隔にとり, $\mathbf{x} = (x_1, \dots, x_{20})$ での観測値を $\mathbf{z} = (1, 1, 0, 0, 0, 0, 1, 1, 1, 1, 0, 1, 1, 1, 1, 1, 1, 1, 0, 1)$ として, 周期カーネルのハイパーパラメータが $\theta_1 = 1$, $\theta_1 = 16$ のそれぞれの場合に MCMC を 100 万回ずつおこない, g_{15} の周辺事後分布を求めた. その結果が図 8g, h である. $\theta_1 = 1$ のときはほぼ正規分布であるが, $\theta_1 = 16$ のときは歪みが生じラプラス近似の精度が悪くなる.

B. 条件数の逆数

特異行列の判定のために MATLAB (The MathWorks Inc., 2022) の組み込み関数 `rcond` を用

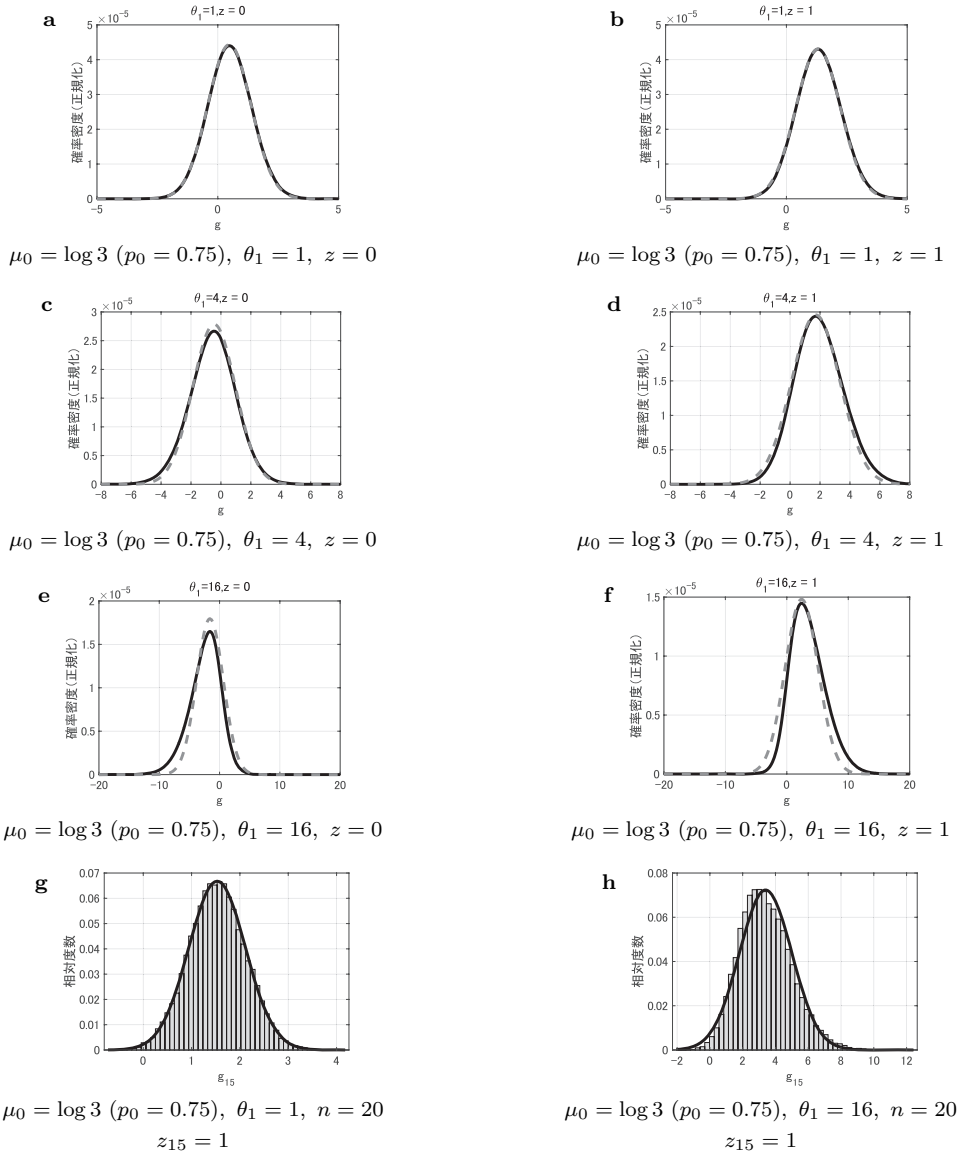


図 8. 異なる θ_1 と z の条件における分布の比較.

いた. この関数は行列 K の 1-ノルムに基づく条件数の逆数を指標とした. 条件数は次式で定義される:

$$\kappa(K) = \|K\| \cdot \|K^{-1}\|$$

ここで $\|\cdot\|$ は行列ノルムであり, 本稿では 1-ノルムを用いた. 1-ノルムは, 行列の各列の絶対値の合計のうち最大のものとして定義される:

$$\|K\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^m |k_{ij}|$$

この定義に基づき、1-ノルム条件数の逆数は次のように近似される：

$$\kappa_1^{-1}(K) \approx \frac{1}{\|K\|_1 \cdot \|K^{-1}\|_1}$$

`rcond` は、LAPACK (Anderson et al., 1999) の条件数推定アルゴリズムに基づき、LU 分解を用いて $\|K^{-1}\|_1$ を間接的に評価することで、条件数の逆数 $\kappa_1^{-1}(K)$ を高速かつ数値的に安定に算出している。

参 考 文 献

- Anderson, E., Bai, Z., Bischof, C., Blackford, S., Demmel, J., Dongarra, J., Du Croz, J., Greenbaum, A., Hammarling, S., McKenney, A. and Sorensen, D. (1999). *LAPACK Users' Guide*, 3rd ed., Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania.
- Busch, N. A. and VanRullen, R. (2010). Spontaneous EEG oscillations reveal periodic sampling of visual attention, *Proceedings of the National Academy of Sciences*, **107**(37), 16048–16053, <https://doi.org/10.1073/pnas.1004801107>.
- Chen, L. L., Madhavan, R., Rapoport, B. I. and Anderson, W. S. (2013). Real-time brain oscillation detection and phase-locked stimulation using autoregressive spectral estimation and time-series forward prediction, *IEEE Transactions on Biomedical Engineering*, **60**(3), 753–762.
- Hebron, H., Lugli, B., Dimitrova, R., Jaramillo, V., Yeh, L. R., Rhodes, E., Grossman, N., Dijk, D.-J. and Violante, I. R. (2024). A closed-loop auditory stimulation approach selectively modulates alpha oscillations and sleep onset dynamics in humans, *PLOS Biology*, **22**(6), <https://doi.org/10.1371/journal.pbio.3002651>.
- Kirchhoff, M., Humaidan, D. and Ziemann, U. (2024). Closed-loop phase selection in EEG-TMS using Bayesian optimization, *2024 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 1730–1735, <https://doi.org/10.1109/SMC54092.2024.10831723>.
- Mathewson, K. E., Gratton, G., Fabiani, M., Beck, D. M. and Ro, T. (2009). To see or not to see: Prestimulus α phase predicts visual awareness, *Journal of Neuroscience*, **29**(9), 2725–2732, <https://doi.org/10.1523/JNEUROSCI.3963-08.2009>.
- Ng, B. S. W., Schroeder, T. and Kayser, C. (2012). A precluding but not ensuring role of entrained low-frequency oscillations for auditory perception, *Journal of Neuroscience*, **32**(35), 12268–12276, <https://doi.org/10.1523/JNEUROSCI.1877-12.2012>.
- Onojima, T. and Kitajo, K. (2021). A state-informed stimulation approach with real-time estimation of the instantaneous phase of neural oscillations by a Kalman filter, *Journal of Neural Engineering*, **18**(6), <https://doi.org/10.1088/1741-2552/ac2f7b>.
- Onojima, T., Kitajo, K. and Mizuhara, H. (2017). Ongoing slow oscillatory phase modulates speech intelligibility in cooperation with motor cortical activity, *PLOS ONE*, **12**(8), <https://doi.org/10.1371/journal.pone.0183146>.
- Rue, H., Martino, S. and Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations, *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **71**(2), 319–392, <https://doi.org/10.1111/j.1467-9868.2008.00700.x>.
- Shakeel, A., Onojima, T., Tanaka, T. and Kitajo, K. (2021). Real-time implementation of EEG oscillatory phase-informed visual stimulation using a least mean square-based AR model, *Journal of Personalized Medicine*, **11**(1), <https://doi.org/10.3390/jpm11010038>.
- Sofro, A. and Oktaviarina, A. (2018). Gaussian process regression model in spatial logistic regression, *Journal of Physics: Conference Series*, **947**(1), <https://doi.org/10.1088/1742-6596/947/1/012005>.
- Tervo, A. E., Metsomaa, J., Nieminen, J. O., Sarvas, J. and Ilmoniemi, R. J. (2020). Automated

- search of stimulation targets with closed-loop transcranial magnetic stimulation, *Neuroimage*, **220**, <https://doi.org/10.1016/j.neuroimage.2020.117082>.
- Tervo, A. E., Nieminen, J. O., Lioumis, P., Metsomaa, J., Souza, V. H., Sinisalo, H., Stenroos, M., Sarvas, J. and Ilmoniemi, R. J. (2022). Closed-loop optimization of transcranial magnetic stimulation with electroencephalography feedback, *Brain Stimulation*, **15**(2), 523–531, <https://doi.org/10.1016/j.brs.2022.01.016>.
- The MathWorks Inc. (2022). MATLAB version: 9.13.0.2320565 (R2022b), <https://www.mathworks.com/help/releases/R2022b/> (最終アクセス日 2025 年 5 月 1 日).
- Torrecillos, F., Falato, E., Pogosyan, A., West, T., Di Lazzaro, V. and Brown, P. (2020). Motor cortex inputs at the optimum phase of beta cortical oscillations undergo more rapid and less variable corticospinal propagation, *Journal of Neuroscience*, **40**(2), 369–381, <https://doi.org/10.1523/JNEUROSCI.1953-19.2019>.
- Wang, B. and Shi, J. Q. (2014). Generalized Gaussian process regression model for non-Gaussian functional data, *Journal of the American Statistical Association*, **109**(507), 1123–1133, <https://doi.org/10.1080/01621459.2014.889021>.
- Zoefel, B. and VanRullen, R. (2015). The role of high-level processes for oscillatory phase entrainment to speech sound, *Frontiers in Human Neuroscience*, **9**, <https://doi.org/10.3389/fnhum.2015.00651>.
- Zrenner, C., Belardinelli, P., Mueller-Dahlhaus, F. and Ziemann, U. (2016). Closed-loop neuroscience and non-invasive brain stimulation: A tale of two loops, *Frontiers in Cellular Neuroscience*, **10**, <https://doi.org/10.3389/fncel.2016.00092>.
- Zrenner, C., Desideri, D., Belardinelli, P. and Ziemann, U. (2018). Real-time EEG-defined excitability states determine efficacy of TMS-induced plasticity in human motor cortex, *Brain Stimulation*, **11**(2), 374–389, <https://doi.org/10.1016/j.brs.2017.11.016>.
- Zrenner, C., Kozák, G., Schaworonkow, N., Metsomaa, J., Baur, D., Vetter, D., Blumberger, D. M., Ziemann, U. and Belardinelli, P. (2023). Corticospinal excitability is highest at the early rising phase of sensorimotor μ -rhythm, *NeuroImage*, **266**, <https://doi.org/10.1016/j.neuroimage.2022.119805>.

Gaussian Process Regression with a Logit Link to Estimate the Relationship between Neural Oscillatory Phase and the Probability of a Correct Response

Takayuki Onojima¹ and Masakazu Jimbo²

¹Data Science and AI Innovation Research Promotion Center, Shiga University

²The Institute of Statistical Mathematics

Human brain activity can be measured using electroencephalography (EEG), and the ongoing brain state is known to affect task performance. It has been suggested that the phase of EEG oscillations within specific frequency bands is related to perceptual processes. To estimate this relationship, it is necessary to model the functional dependence between EEG phase in a target frequency band and task performance on a subject-by-subject basis. In recent years, stimulation techniques that rely on the ongoing EEG phase, such as phase-dependent stimulation, have been proposed, highlighting the need for methods that can flexibly estimate the relationship between EEG phase and brain function for each individual. In this study, we constructed a statistical framework for flexibly estimating the probability of a correct response, using trial-by-trial binary outcomes (correct/incorrect or detected/undetected) as observed data and EEG phase as an explanatory variable. Specifically, we developed a bernoulli–logit model that employs the logit link function and Gaussian process regression with a periodic kernel to estimate the probability of correct responses as a function of EEG phase. To address matrix degeneration and increased computational cost with larger sample sizes, we further refined the model and proposed a binomial–logit model. This study presents two statistical models for estimating the functional relationship between EEG phase and the probability of a correct response, and verifies their effectiveness through numerical simulations.