

2023.3.7



The University of Tokyo

ベイズ推定のスキーム

～全てのユーザーのために～

東京大学
岡田研究室

片上 舜

－基調講演－

13:05-13:35 「ベイズ計測」 岡田 真人（東京大学）

一般講演

13:35-13:55 「メスバウアーハミルトニアン推定」
森口 椋太（東京大学）

13:55-14:15 「X線小角散乱プロファイルのベイズ推定」
林 悠偉（東京大学）

14:15-14:35 「時分割XRDのベイズ推定による化学反応モデルの理解」
横山 優一（JASRI）

14:35-14:55 「ベイズ統合 ～マルチモーダル測定へ～」
水牧 仁一朗（JASRI）

14:55-15:15 「ベイズ推定のスキーム～全てのユーザーのために～」
片上 舜（東京大学）

自己紹介

- 東京大学・大学院理学系研究科 岡田研 (2016 ~ 2022)
 - 学位論文「ベイズ推論による物理モデルに対するパラメータ分布推定」
- 東京大学・大学院新領域創成科学研究科 助教 (2022/04 ~)

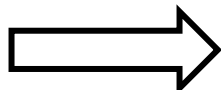
アンケート

- スペクトルや画像データからフィッティングを行なっている
 - そのフィッティングの際に、パラメータを手打ちで決めている。最急降下法などを使っているが、うまくいかない。
 - フィッティング用のモデルが複数あって、事前にどれを使うかを決めておかないといけない。
 - S/Nが悪いデータや欠損データをなんとかした。
 - 複数計測の統合を行いたい。
-
- そのような方は、一度ベイズ計測をお試してください。

ベイズ計測

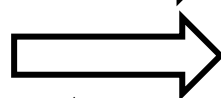
順アプローチ

計測データ



実験の結果

モデル



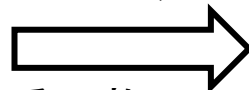
理論の結果



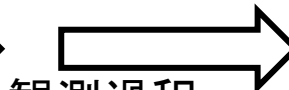
$p(Y|\theta, K)$ 解析計算, 数値計算

逆アプローチ

対象とする
物理系



系の物理
モデル



観測過程
計測機器の特性

計測データ

$p(\theta, K)$

$p(\theta, K|Y)$



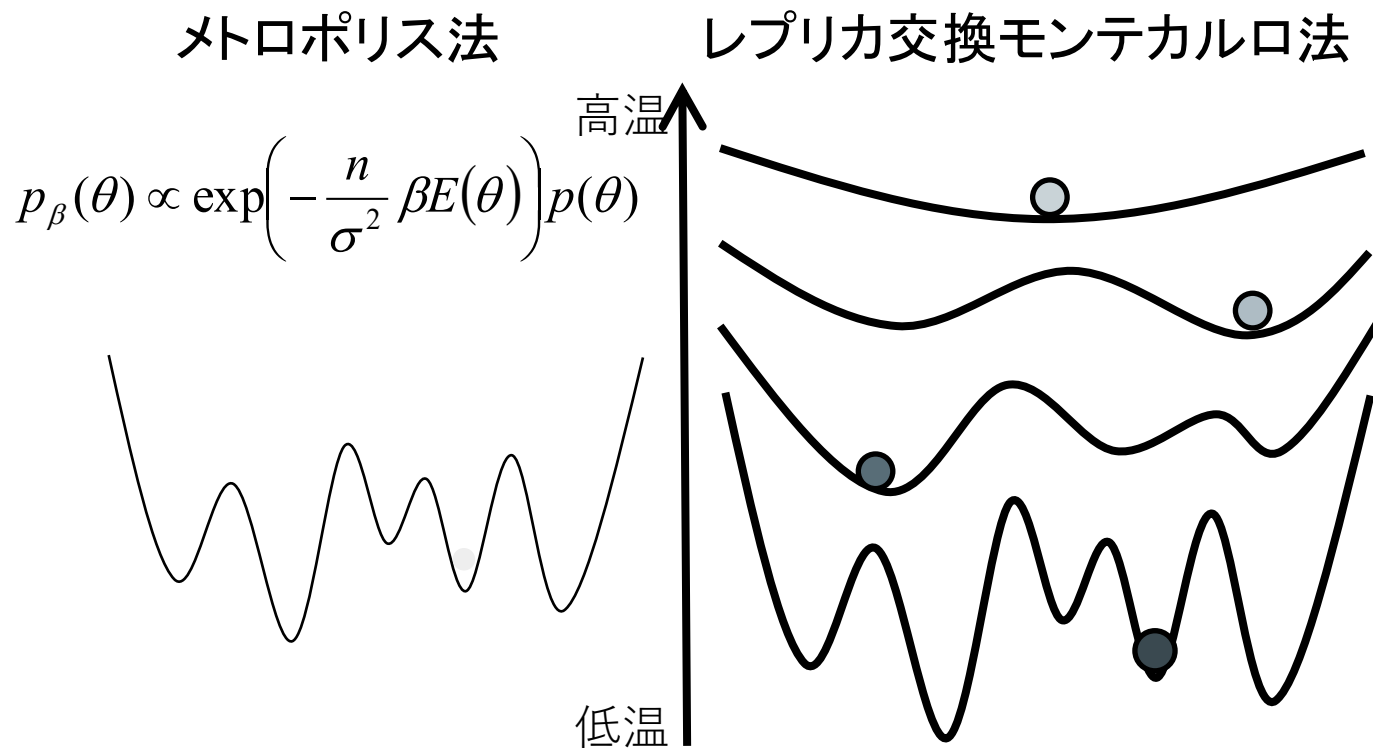
全てをモデル化し

ベイズの定理で因果をさかのぼる



ベイズ推定って、どうやって実装したらいいのか。

レプリカ交換モンテカルロ法 ランダムスピンの知見から



K. Hukushima, K. Nemoto, *J. Phys. Soc. Jpn.* **65** (1996).



ベイズ推定って、どうやって実装したらいいのか。



ベイズ計測オープンソースソフトウェアの構築



マルコフ連鎖モンテカルロ法 (MCMC) のライブラリ

- Stan
- PyMC3
- JAGS
- emcee
- BUGS

交換モンテカルロ法のライブラリ

- Tensorflow
- emcee
- ptmcmcsampler



求められる機能性

- 基本的なベイズ推論ワークフローを完備
- 迅速に実装可能かつ柔軟なモデル構築が可能なUI
- 実行解析結果の可視化
- ベイズ推論の高速な実行

ベイズ推論

$$P(\theta|D) \propto P(D|\theta) \times P(\theta)$$

事後分布 尤度(モデル) 事前分布

θ : 物理量 (モデルパラメータ)

D : 計測データ

モデルの実装

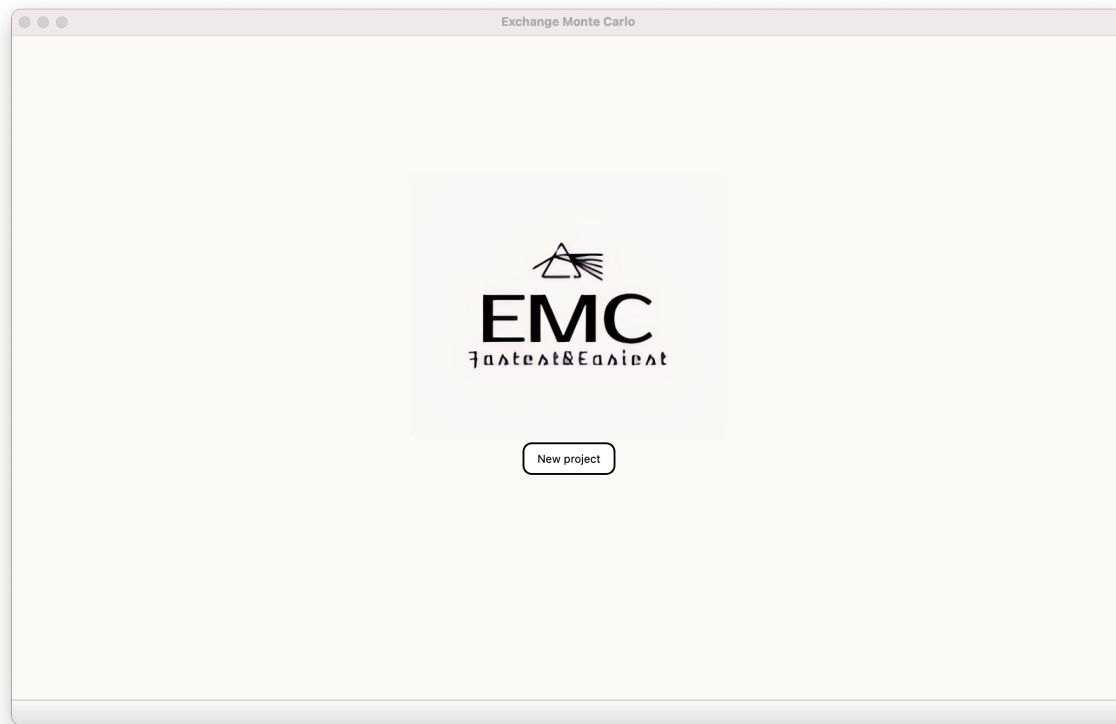


データの取り込み

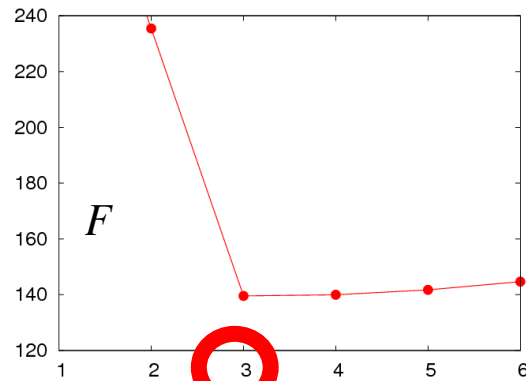
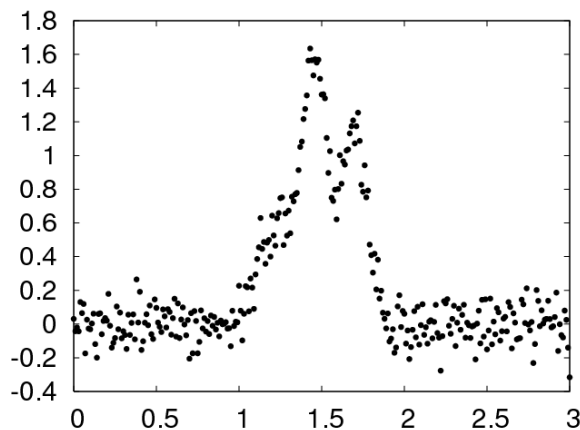


解析結果の確認

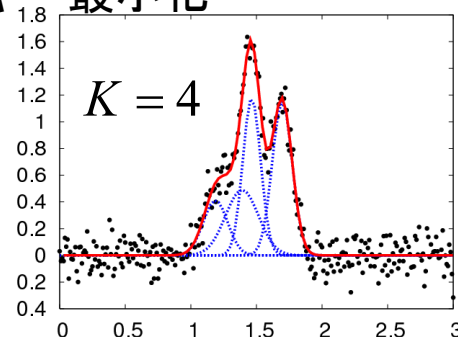
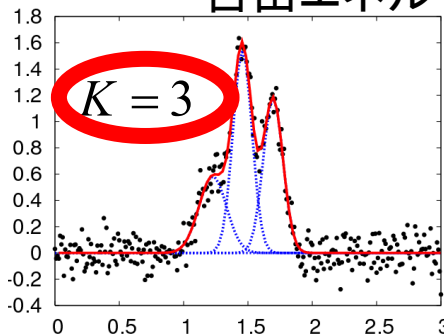
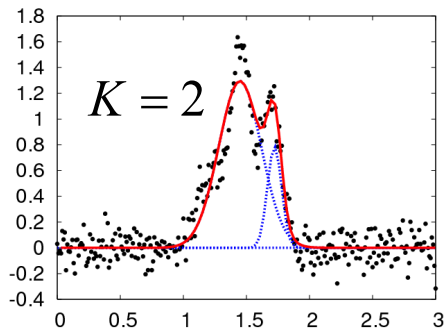
ベイズ推論ワークフロー



スペクトル分解



最適な K をデータだけが決める
自由エネルギー最小化



Nagata, Sugita and Okada, Bayesian spectral deconvolution with the exchange Monte Carlo method, *Neural Networks* 2012

スペクトル分解の定式化

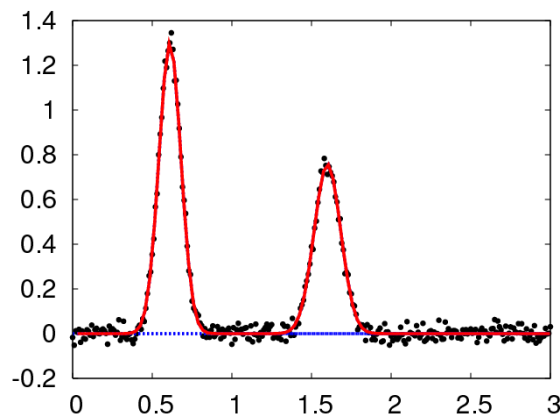
ガウス関数(基底関数)の足し合わせにより, スペクトルデータを近似

観測データ: $D = \{x_i, y_i\}_{i=1}^n$

x_i : 入力 y_i : 出力

$$f(x; \theta) = \sum_{k=1}^K a_k \exp\left(-\frac{b_k (x - \mu_k)^2}{2}\right)$$

$$\theta = \{a_k, b_k, \mu_k\} \quad k = 1, \dots, K$$



二乗誤差を最小にするようにパラメータをフィット(最小二乗法)

$$E(\theta) = \frac{1}{n} \sum_{i=1}^n (y_i - f(x_i; \theta))^2$$

01 EMCによるスペクトル分解実装



Exchange Monte Carlo

モデル設定

データ入力 x: 次元

データ出力 y: 次元

パラメータの数

フォワードモデル

	Name	Prior distribution
1	<input type="text"/>	<input type="text"/>

基底数 (BaseNUM)

次へ

フォワードモデル

$$f(x; \theta) = \sum_{k=1}^K a_k \exp\left(-\frac{b_k (x - \mu_k)^2}{2}\right)$$

02 EMCによるスペクトル分解実装



Exchange Monte Carlo

モデル設定

データ入力 x: 次元

データ出力 y: 次元

パラメータの数

	Name	Prior distribution
1	a	gamma(2,2)
2	mu	normal(160,2)
3	b	normal(10,2.5)

基底数 (BaseNUM)

フォワードモデル

```
for(int i=0; i<BaseNum;i++){  
  y+=a[i]*exp(-b[i])*pow(x[i]-mu[i],2)/2);  
}
```

次へ

フォワードモデル

$$f(x;\theta) = \sum_{k=1}^K a_k \exp\left(-\frac{b_k(x - \mu_k)^2}{2}\right)$$



パラメータの名前と事前分布

パラメータの数 3

	Name	Prior distribution
1	a	gamma(2,2)
2	mu	normal(160,2)
3	b	normal(10,2.5)

基底数 (BaseNUM) 3

フォワードモデル

$$f(x;\theta) = \sum_{k=1}^K a_k \exp\left(-\frac{b_k(x - \mu_k)^2}{2}\right)$$

```
for(int i=0; i<BaseNum;i++){  
    y += a(i)*exp(-b(i)*pow(x[0]-mu(i),2)/2)  
}
```




サンプリング設定とデータの設定

Exchange Monte Carlo

サンプリング設定 & データ設定

バーンイン数 1000

サンプリング数 1000

レプリカ数 8 gamma 1.00

解析データを指定

出力先パスを指定

Name	C	d
a		
b		
mu		

戻る

保存

実行

03 EMCによるスペクトル分解実装



サンプリング設定とデータの設定

Exchange Monte Carlo

サンプリング設定 & データ設定

バーンイン数

サンプリング数

レプリカ数 gamma

Name	C	d
a	<input type="text" value="5"/>	<input type="text" value="0.7"/>
mu	<input type="text" value="10"/>	<input type="text" value="0.7"/>
b	<input type="text" value="0.1"/>	<input type="text" value="0.7"/>

Inverse Temp. vs Replica Index

04 EMCによるスペクトル分解実装



Exchange Monte Carlo

サンプリング設定 & データ設定

バーンイン数

サンプリング数

レプリカ数 gamma

Name	C	d
a	<input type="text" value="5"/>	<input type="text" value="0.7"/>
mu	<input type="text" value="10"/>	<input type="text" value="0.7"/>
b	<input type="text" value="0.1"/>	<input type="text" value="0.7"/>

F: 6662.48

01 おわりに：ベイズ計測オープンソースソフトウェアの構築



- 基本的なベイズ推論ワークフローを完備
- 迅速に実装可能かつ柔軟なモデル構築が可能なUI
- 実行解析結果の可視化
- ベイズ推論の高速な実行

